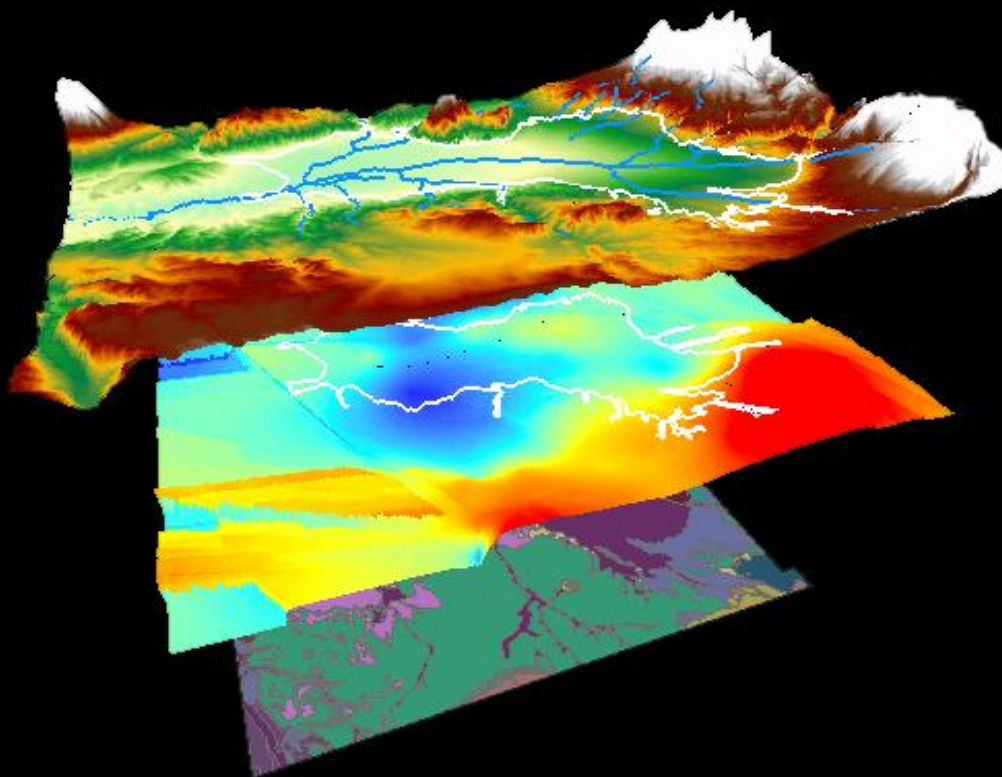




*Universidad de Granada
Dpto. Geodinámica*

*Metodología basada en SIG para el Desarrollo
de un Sistema Soporte de Decisión en la gestión
de la calidad de los recursos hídricos
subterráneos de la “Vega de Granada”*



*Tesis Doctoral
María José García Soldado
Granada, 2009*

METODOLOGÍA BASADA EN SIG PARA EL
DESARROLLO DE UN SISTEMA SOPORTE DE
DECISIÓN EN LA GESTIÓN DE LA CALIDAD
DE LOS RECURSOS HÍDRICOS SUBTERRÁNEOS
DE LA "VEGA DE GRANADA"

Tesis doctoral presentada en la Universidad de Granada
para optar al grado de doctor
Facultad de Ciencias

María José García Soldado
Departamento de Geodinámica

16 de junio de 2009

Editor: Editorial de la Universidad de Granada
Autor: María José García Soldado
D.L.: GR 512-2013
ISBN: 978-84-9028-081-2

El trabajo de investigación que se expone en la presente memoria, titulada: METODOLOGÍA BASADA EN SIG PARA EL DESARROLLO DE UN SISTEMA SOPORTE DE DECISIÓN EN LA GESTIÓN DE LA CALIDAD DE LOS RECURSOS HÍDRICOS SUBTERRÁNEOS DE LA "VEGA DE GRANADA", que para aspirar al grado de Doctor en Geología presenta por María José García Soldado, ha sido realizado en la Universidad de Granada gracias a la beca concedida por el Ministerio de Educación y Ciencia "Formación de Personal Investigador (convocatoria 2003)", bajo la dirección de:

Vº Bº:

El Director:

Vº Bº:

El Director:

Mario Chica Olmo

Juan Antonio Luque Espinar

Aspirante al Grado de Doctor:

Lcda. María José García Soldado

Agradecimientos

La investigación y escritura de la tesis doctoral es un trabajo duro y en muchos casos imposible sin la ayuda y el apoyo de muchas otras personas. En estas líneas me gustaría agradecer a todas esas personas que han hecho posible la realización de esta tesis doctoral, ya sea mediante sus directrices científicas o mediante su apoyo moral, ya que ambas cosas han sido en cada momento igual de necesarias.

Agradecer a mi director de tesis, Mario Chica, que me acogió e incluyó en su grupo de trabajo RSGIS. Todos estos años han dado para mucho y afortunadamente me llevo muchas cosas buenas. Desde aquella primera reunión en su despacho, en la que Mario esbozaba la idea de la tesis hasta hoy, me parece que ha pasado casi una vida. Y es que he han sido muchas las dificultades que hemos compartido en el camino. Debo decir que me siento privilegiada de haber tenido un gran maestro de director y una bellísima persona como amigo.

A Juan Antonio Luque, mi codirector y probablemente en parte responsable de mi inquietud por la investigación. Juan Antonio ha sido fundamentalmente un amigo y un apoyo. Siempre recordaré aquellos días de campo, junto con María Dolores, Luis, Antonio Castillo, Irene y Juan Pedro. Era el comienzo de la investigación, y a pesar del frío, la lluvia y ciertos problemillas con el coche en el barro, no los más duros.

A Enrique Barahona, a quien más echaré de menos y a quien no volveré a ver para decirle cuanto lo admiraba. Enrique nos acogió en el Zaidín, y nos presentó a la doctora María Dolores Mingorance que nos ayudó en los primeros momentos. Enrique nos guió, a Irene y a mí en la realización de los análisis de los suelos y además no dudó en asesorarnos en todas nuestras dudas edafológicas. Él era una de esas personas que disfrutaba compartiendo su conocimiento, explicando y sobre todo aprendiendo. Una de las más admirables de Enrique es su dinamismo, siendo capaz de aprender todos los días.

A Gabriel Delgado Calvo-Flores, por sus aclaraciones en la edafología de la Vega de Granada, su disponibilidad y su amabilidad.

A Manuel Lozano, por su guía en la aplicación de métodos complementarios a la hidrogeología como son los métodos de inteligencia artificial, por su amistad, por su ánimo, por sus *meng* y por acogerme en esa encantadora familia.

A Javier Cruz San Julián, por sus consejos en todos los sentidos, por sus correcciones, por sus aclaraciones tanto hidrológicas como gramaticales, por su tiempo y su carácter afable y cercano ante cualquier duda que le he planteado.

A Víctor, mi compañero, por su actitud colaboradora siempre, por sus explicaciones de teledetección, por contar conmigo, por entenderme, porque me siento afortunada de tenerte como compañero y amigo. Gracias.

En general a todos los miembros de mi grupo de investigación, por su disponibilidad, por su tiempo, por su ánimo, apoyo y por sus conocimientos que han servido como base de este trabajo. Mario, Eulogio, Juan Pedro, Juan Antonio, Manolo, Irene, Víctor. Gracias a todos.

Este trabajo se ha visto enriquecido también por la estancia realizada en la Universidad de Jaén, con el doctor Francisco Javier Ariza. A quien agradezco la acogida y los consejos prestados en relación fundamentalmente a la parte de análisis multicriterio y toma de decisiones.

También quiero agradecer a mis compañeros del CEAMA su amistad y su tiempo y a mis amigos. Marian gracias por solventarme todo el tema de papeleo en tantas ocasiones, Luis gracias por tu apoyo y tu ánimo, Juanlu, Jaime, Fran, Juan Alfredo, Susana, Penélope, Hassan, Curro, Olga, Eli, Damián, Josemi, gracias por vuestra amistad, por vuestro tiempo, por el ánimo que me habéis dado y por vuestros consejos en los momentos más duros, sin duda habría sido más difícil sin vuestra compañía y sin nuestros cafés.

Mi último agradecimiento va dirigido a mi familia, a mis padres Jesús y María José, a mi hermano y a mi tío Pepe. Ellos se han llevado la peor parte, ya que sé que he sido difícil y "he puesto de los nervios" a todos en algunos momentos. Muchas gracias por entenderme y saber animarme.

Finalmente quiero dedicar este trabajo a mi marido y a mi hijo, de quienes he robado más tiempo. Gracias David por ayudarme en tantos momentos en los que perdí las fuerzas, por devolverme la luz en los momentos más duros y por estar conmigo. Sin ti no habría sido capaz de terminar esta tesis.

A David, mi marido y

a mi hijo

Índice general

1. Introducción	1
1.1. Aspectos generales	1
1.2. Antecedentes	4
1.3. Objetivos	8
1.4. Organización de la memoria	10
2. El área de estudio: La Vega de Granada	13
2.1. Localización geográfica	13
2.2. Marco hidrogeológico	13
2.2.1. Geología y geometría del acuífero	13
2.2.2. Funcionamiento hidrogeológico y piezometría	23
2.2.3. Aspectos hidrogeoquímicos	32
2.2.3.1. Variables físicas	36
2.2.3.2. Variables químicas	46
2.2.3.3. Microbiología	64
2.3. Climatología y datos foronómicos	72
2.4. Aspectos edafológicos	78
2.4.1. Textura	86
2.4.2. Materia orgánica	91
2.4.3. pH del suelo	95
2.4.4. Humedad del suelo	97
2.4.5. Especies químicas de la fase líquida del suelo	99
2.5. Imágenes de teledetección espacial y aerotransportada	109
2.5.1. Imágenes multiespectrales: Landsat Thematic Mapper	110
2.5.2. Imágenes hiperespectrales: Airbone Hyperspectral Scanner	111
2.6. Marco socioeconómico	113
2.7. Marco ambiental	120
2.8. Otros datos complementarios	121

3. Bases metodológicas para el desarrollo del S.S.D.	125
3.1. Introducción	125
3.2. Aspectos introductorios de la base de datos geoespacial	127
3.2.1. Tipos de bases de datos	129
3.2.2. Diseño conceptual de la base de datos	132
3.2.3. Diseño lógico de la base de datos	140
3.2.4. Diseño físico de la base de datos	142
3.2.5. Implementación de la base de datos	142
3.3. Métodos Geoestadísticos de análisis espacial	147
3.3.1. Estimación de la estructura de variabilidad espacial: el variograma	153
3.3.2. El krigeaje Ordinario (KO) y sus variantes	156
3.4. Tratamiento de imágenes de Teledetección	163
3.4.1. Procesamiento digital de imágenes	166
3.4.1.1. Preprocesamiento: corrección de imágenes	167
3.4.1.2. Realces geométricos	167
3.4.1.3. Realces radiométricos	168
3.4.1.4. Composiciones de color	168
3.4.1.5. Transformaciones de la imagen	169
3.4.2. Clasificación digital de imágenes	170
3.5. Métodos basados en Inteligencia Artificial	175
3.5.1. Aspectos generales	175
3.5.2. Modelado de sistemas con Sistemas Basados en Reglas Difusas .	178
3.5.2.1. Componentes de un SBRD	182
3.5.2.2. Aprendizaje de SBRDs	183
3.5.3. Redes Neuronales Artificiales	184
3.5.3.1. La neurona artificial	186
3.5.3.2. Redes neuronales artificiales	188
3.5.3.3. Modelos de redes neuronales artificiales	190
3.6. El Sistemas de Información Geográfica como herramienta de integración de datos espaciales	191
3.6.1. Generalidades	191
3.6.2. Operaciones y análisis de datos en los SIG	198
3.7. La teoría de la decisión en la gestión de los recursos hídricos, y su apli- cación en el S.S.D.	201
 4. Análisis temático y espacial de los datos	 207
4.1. Introducción	207
4.2. Análisis Exploratorio de Datos	210

4.3.	Análisis estadístico multivariante	227
4.3.1.	Estudio estadístico multivariante de datos hidroquímicos	227
4.3.2.	Estudio multivariante de datos edafológicos	239
4.4.	Análisis espacial de los datos	251
4.4.1.	Regionalización de los parámetros químicos del agua y suelo	252
4.4.2.	Regionalización de parámetros hidrodinámicos del acuífero	258
4.4.3.	Espacialización geostatística de índices de calidad del agua	262
5. Integración de la información en SIG y aplicaciones al estudio de la calidad del agua subterránea		269
5.1.	Introducción	269
5.2.	Integración SIG mediante modelos basados en el conocimiento experto	271
5.2.1.	Creación de mapas de calidad del agua mediante lógica booleana	272
5.2.2.	Creación de mapas de calidad del agua mediante suma ponderada	277
5.2.3.	Creación de mapas de calidad del agua mediante lógica difusa	284
5.2.4.	Creación de mapas de calidad del agua mediante agregación exponencial	285
5.2.5.	Aplicación de la Evaluación Multicriterio al estudio de la vulnerabilidad	294
5.2.5.1.	Evaluación de la vulnerabilidad en la Vega de Granada	296
5.2.5.2.	Creación del mapa de vulnerabilidad	303
5.3.	Integración SIG mediante modelos basados en los datos	308
5.3.1.	Estimación de la calidad del agua subterránea mediante redes neuronales artificiales	308
5.3.2.	Estimación de la vulnerabilidad mediante Sistemas de Reglas Basados en Lógica Difusa	315
6. Conclusiones		331
Bibliografía		337

Índice de figuras

1.1. Esquema de la metodología del trabajo.	11
2.1. Situación geográfica de la Vega de Granada.	14
2.2. Relieve y ortofoto de la Vega de Granada.	15
2.3. Situación de la región estudiada en un mapa geológico en el que se representan las principales unidades estudiadas de la cordillera bética.	16
2.4. Mapa geológico del acuífero de la Vega de Granada y su entorno.	17
2.5. Mapa de coropletas de la litología de la zona no saturada del acuífero de la Vega de Granada (estimación mediante krigeaje)	22
2.6. Mapa de cota del sustrato del acuífero de la Vega de Granada.	23
2.7. Cortes transversales y longitudinales del acuífero de la Vega de Granada.	24
2.8. Distribución de la red de acequias en el área de estudio.	26
2.9. Distribución de la precipitación en el área de estudio.	27
2.10. Manantiales Fuente Fontana y Martinete.	27
2.11. Mapa de coropletas de la piezometría e isopiezas de Abril de 1980.	28
2.12. Mapa de espesor no saturado y espesor saturado del acuífero.	29
2.13. Estimación geoestadística del gradiente para abril de 1980. (a) Estimación del módulo del gradiente. (b) Estimación de la dirección del gradiente máximo.	30
2.14. Detalle de la dirección del gradiente en un sector localizado al noroeste del acuífero. (Ver recuadro figura 2.13).	31
2.15. Estimación de la conductividad hidráulica y de la transmisividad.	32
2.16. Fotografía del muestreo de aguas superficiales realizado en el río Genil con GPS Trimble Pathfinder.	33
2.17. Distribución espacial de datos experimentales de aguas subterráneas.	37
2.18. Distribución espacial de la turbidez en el área de estudio.	38
2.19. Distribución espacial de sólidos en suspensión en el área de estudio	39
2.20. Distribución espacial de valores de conductividad del agua subterránea.	41

2.21. Red de muestreo de la temperatura de las aguas subterráneas, distribución de la T ^a y situación de los perfiles de T ^a y conductividad.	41
2.22. Perfiles de temperatura y conductividad del acuífero.	43
2.23. Mapa de la distribución espacial del pH en las aguas subterráneas.	45
2.24. Distribución espacial de calcio y magnesio del agua subterránea.	46
2.25. Distribución espacial de sodio y potasio del agua subterránea.	49
2.26. Distribución espacial de bicarbonatos y sulfatos en el agua subterránea.	50
2.27. Distribución espacial de cloruros y nitratos en las aguas subterráneas.	53
2.28. Distribución espacial de fluoruros y boro en las aguas subterráneas.	55
2.29. Distribución espacial de plomo y cromo en las aguas subterráneas.	58
2.30. Distribución espacial del contenido en hierro y cobre en las aguas subterráneas.	60
2.31. Distribución espacial del contenido en manganeso y antimonio en las aguas subterráneas.	62
2.32. Distribución espacial del contenido en arsénico en las aguas subterráneas.	65
2.33. Distribución espacial del contenido en bacterias aerobias de las aguas subterráneas.	69
2.34. Distribución espacial del contenido en coliformes totales y fecales de las aguas subterráneas.	70
2.35. Distribución espacial del contenido en estreptococos fecales y D.Q.O. de las aguas subterráneas.	71
2.36. Climograma de la Vega de Granada.	73
2.37. Distribución espacial de las estaciones termopluviométricas en la Vega de Granada.	74
2.38. Precipitaciones medias mensuales para la Vega de Granada.	75
2.39. Precipitaciones máximas medias en la Vega de Granada.	76
2.40. Hidrogramas de los principales ríos de la Vega de Granada.	77
2.41. Evolución piezométrica (cota sobre el nivel del mar), foronómica y precipitación mensual en la zona SE de la Vega de Granada.	79
2.42. Análisis espectrales de algunas estaciones de la Vega de Granada.	80
2.43. Análisis espectrales de algunas estaciones de la Vega de Granada.	81
2.44. Ejemplos de los tipos de suelos de la Vega de Granada: (a) poco evolucionados, (b) medianamente evolucionados (c) muy evolucionados, y (d) y con horizonte Ck	83
2.45. Mapa de suelos de la Vega de Granada (LUCDEME). <i>Nota: La división horizontal del mapa corresponde a la división de las hojas 1009 y 1026 que no han sido unificadas.</i>	84

2.46. Fotografía de la preparación de las muestras de suelos de la Vega de Granada (estufa de secado y tamiz de suelos).	85
2.47. Fotografía de los análisis de las muestras en la Estación Experimental del Zaidín. Determinación de las arcillas y batidora de las muestras.	86
2.48. Distribución espacial del muestreo de suelos en la Vega de Granada.	87
2.49. Distribución espacial del % arenas y % arcillas en los suelos de la Vega de Granada.	88
2.50. Mapa del % de limos de los suelos de la Vega de Granada.	89
2.51. Mapa de la textura de los suelos de la Vega de Granada.	89
2.52. Triángulo textural USDA y representación de los suelos analizados en la Vega de Granada.	90
2.53. Mapas del % de arcillas en el horizonte superior e inferior de los suelos de la Vega de Granada.	91
2.54. Mapa del % de arenas en los horizontes superior e inferior de los suelos de la Vega de Granada.	92
2.55. Mapa del % de limos en los horizontes superior e inferior de los suelos de la Vega de Granada.	92
2.56. Distribución espacial del porcentaje de carbono orgánico en los suelos de la Vega de Granada en un horizonte de 1 m y perfiles en profundidad del % C.O.	94
2.57. Distribución espacial del % carbono orgánico en suelo de los horizontes superior e inferior de la Vega de Granada	95
2.58. Distribución espacial del pH promedio en el suelo de la Vega de Granada.	96
2.59. Distribución espacial del pH en suelo del horizonte superior e inferior de la Vega de Granada.	97
2.60. Distribución espacial del % humedad promedio (0-1 m) en el suelo.	98
2.61. Distribución espacial del % humedad del suelo en el horizonte superior e inferior de la Vega de Granada.	98
2.62. Distribución espacial de la conductividad en la solución del suelo de la Vega de Granada.	100
2.63. Mapas del contenido en potasio y sodio en la solución del suelo de la Vega de Granada.	101
2.64. Mapas del contenido en magnesio y calcio en la solución del suelo de la Vega de Granada.	102
2.65. Mapas del contenido en cloro y sulfatos en la solución del suelo de la Vega de Granada.	103
2.66. Mapas del contenido en nitratos y nitritos en la solución del suelo de la Vega de Granada.	103

2.67. Mapas del contenido en manganeso y cinc en la solución del suelo de la Vega de Granada.	106
2.68. Mapas del contenido en cobre y cromo en la solución del suelo de la Vega de Granada.	107
2.69. Mapas del contenido en hierro y plomo en la solución del suelo de la Vega de Granada.	108
2.70. Composición TM 3,2,1 (RGB) "color verdadero" año 1985.	111
2.71. Composición TM 3,2,1 (RGB) "color verdadero" año 1993.	112
2.72. Composición TM 3,21 "color verdadero" año 2001.	112
2.73. Imagen AHS en "falso color" (Ch:13,Ch:8,Ch:4) con resolución espacial 6x6 m.	114
2.74. Imagen AHS en "falso color" (Ch:13,Ch:8,Ch:4) con resolución espacial 2x2 m.	115
2.75. Evolución de la población 1900-2006 en los municipios de la Vega de Granada.	117
2.76. Densidad de población por municipios en la Vega de Granada (enero 2005)	117
2.77. Unidades ganaderas y porcentaje de cabezas de ganado por tipo en los municipios de la Vega de Granada (1999).	118
2.78. Explotaciones agrarias en hectáreas y aprovechamiento por cultivo en los municipios de la Vega de Granada (1999).	119
2.79. Distribución del consumo de agua (m^3/dia).	119
2.80. Distribución de los focos de contaminación en la Vega de Granada. . . .	123
3.1. Un ejemplo del diagrama entidad-relación (E-R).	130
3.2. Ejemplo de elemento estructural de tipo virtual y clave externa de la base de datos.	132
3.3. Ejemplo de elemento primario (capa vectorial) y derivado (capa ráster). . . .	138
3.4. Organización de la base de datos HIDROGIS junto con algunos de los metadatos de la misma.	143
3.5. Ejemplo de metadatos para una de las capas de información de la base de datos.	144
3.6. Ejemplo de una estimación geoestadística para una variable hidrogeológica experimental	148
3.7. Mapas de coropletas correspondiente a la estimación por krigeaje puntual de la variable piezometría y de bloque.	151

3.8.	Variograma experimental (puntos) y variograma ajustado (líneas) para un ejemplo de anisotropía zonal calculada para la dirección de máxima variabilidad (130°) y mínima variabilidad (40°) de la permeabilidad del acuífero de la Vega de Granada.	152
3.9.	Mapa de errores de la estimación por krigeaje de la variable transmisividad.	152
3.10.	Ejemplo de variograma de la variable % de arcillas del suelo, donde se observa el efecto de la deriva para valores altos de distancia.	154
3.11.	Variograma experimental y modelo ajustado correspondiente a la variable humedad de saturación del suelo.	156
3.12.	Modelos teóricos de variograma utilizados en este trabajo.	157
3.13.	Elementos de un sistema de teledetección.	164
3.14.	Espectro electromagnético.	165
3.15.	Contraste espectral de la vegetación sana en las bandas del R e IRC del espectro.	170
3.16.	Índice de Vegetación Normalizado para la Vega de Granada y su contorno (NDVI) del año 2004 (Sensor ETM+).	171
3.17.	Localización de los polígonos de entrenamiento escogidos para la clasificación digital de la imagen Landsat ETM+.	173
3.18.	Espectros medios de reflectividad de los polígonos de entrenamiento.	174
3.19.	Mapa clasificado de Imagen Landsat ETM+ (año 2004) por el método de máxima verosimilitud.	174
3.20.	Mapa clasificado de Imagen Landsat ETM+ (año 2004) por el metodo de maxima verosimilitud y con filtro de mediana.	175
3.21.	Funciones de pertenencia de conjuntos.	179
3.22.	Estructura básica de un SBRD.	182
3.23.	Ejemplo de aprendizaje por el método de Wang y Mendel.	184
3.24.	Método de Wang y Mendel.	185
3.25.	Ejemplo de red neuronal de propagación hacia delante.	189
3.26.	El ciclo de tratamiento de la información georreferenciada.	192
3.27.	Ejemplo de creación de un mapa de zonas encharcables.	193
3.28.	Ejemplo de distintas clases de capas en un SIG.	196
3.29.	Rasterización del mapa vectorial de los usos del suelo de la Vega de Granada.	197
3.30.	Operaciones de los SIG.	199
3.31.	Operaciones básicas del SIG: ejemplo de creación de mapas mediante consultas a la base de datos.	200
3.32.	El SIG como herramienta metodológica en el Sistema Soporte de Decisión de la Vega de Granada.	201

4.1.	Histogramas y situación espacial de las distintas poblaciones de algunos de los parámetros fisicoquímicos del agua subterránea.	215
4.2.	Histogramas de algunos metales pesados del agua subterránea.	216
4.3.	Histogramas de las variables microbiológicas de las aguas subterráneas.	217
4.4.	Histogramas y regionalización de algunos de las variables del suelo.	218
4.5.	Histogramas de algunos parámetros relacionados con el acuífero.	219
4.6.	Diagramas de caja (box-plot) para algunas de las variables de las aguas subterráneas.	221
4.7.	Diagramas de caja (box-plot) para las variables del suelo.	222
4.8.	Diagramas de caja (box-plot) para algunos de los parámetros del acuífero.	223
4.9.	Gráficos Q-Q normales de algunos elementos químicos mayoritarios de las aguas subterráneas.	224
4.10.	Gráficos Q-Q normales para algunos metales pesados y microbiológicos.	225
4.11.	Gráficos Q-Q normales para algunos elementos químicos y físicos de los suelos.	226
4.12.	Gráficos Q-Q normales para los parámetros relacionados con el acuífero.	228
4.13.	Dendograma resultado del HCA para las muestras de aguas subterráneas y diagramas de Stiff para la concentración media de cada clúster.	231
4.14.	Gráfico de Collins para los tres clústeres.	234
4.15.	Diagrama de Piper para los tres clústeres.	234
4.16.	Regionalización en el acuífero de los clústeres hidrogeoquímicos obtenidos mediante HCA.	235
4.17.	Diagrama Piper para las aguas superficiales.	236
4.18.	Variogramas y modelos teóricos de los factores del ACP.	240
4.19.	Mapas estimados de los componentes del ACP.	241
4.20.	Dendograma de los análisis del suelo.	242
4.21.	Situación espacial de los clústeres de los análisis del suelo.	243
4.22.	Mapas de estimación geoestadística correspondientes a los componentes principales y variables más significativas para cada uno.	248
4.23.	Variogramas y modelos ajustados de algunas de las variables de la calidad del agua subterránea.	253
4.24.	Ejemplos de comportamientos aleatorios.	255
4.25.	Mapas de errores en la estimación geoestadística del contenido en calcio de las aguas subterráneas.	256
4.26.	Variogramas y modelos ajustados de algunas variables químicas del suelo.	257
4.27.	Histograma de la transmisividad.	260
4.28.	Variograma experimental y ajuste del modelo teórico log-T de la transmisividad.	261

4.29. Mapas de transmisividad realizados mediante distintas estimaciones geo-estadísticas.	263
4.30. Variograma de la indicatriz mediana para el índice de calidad.	264
4.31. Mapas de probabilidad espacial de la calidad del agua según el índice basado en Poch.	266
4.32. Cálculo de funciones de distribución local del índice de calidad para tres sectores del acuífero.	267
5.1. Clasificación de técnicas de EMC.	273
5.2. Mapas de probabilidad creados mediante krigeaje de indicatrices de algunas de las variables fisicoquímicas del agua subterránea.	274
5.3. Mapas binarios de aptitud según <i>RD 140/2003</i> construido mediante criterios de probabilidad de algunas de las variables fisicoquímicas del agua subterránea.	276
5.4. Modelo de integración de capas de tipo booleano para la creación de un mapa de calidad.	278
5.5. Mapa binario resultado de la superposición de zonas aptas (1) y no aptas (0) de calidad del agua subterránea según <i>RD 140/2003</i> mediante criterios de probabilidad.	278
5.6. Mapa de calidad del agua resultado de la suma de los mapas binarios de probabilidad de aptitud para consumo público del agua. (Leyenda: 0-6 Variables que no cumplen los criterios de potabilidad).	279
5.7. Variogramas experimentales de las variables indicatriz mediana correspondientes a los índices de calidad del agua de Poch (izquierda) y Mingo Magro (derecha) Mediana(Poch)=90.12, Mediana(CE)=80.27	281
5.8. Mapa de isoprobabilidad (krigeaje de indicatrices) de dos índices de calidad del agua con un valor indicatriz igual a la mediana de la distribución.	281
5.9. Funciones de pertenencia sigmoidales.	285
5.10. Variograma experimental de la variable indicatriz mediana correspondiente al índice ID de calidad del agua creado mediante funciones difusas. Mediana=89.99	286
5.11. Mapa de isoprobabilidad (krigeaje de indicatrices) del índice de calidad del agua ID construido con funciones de pertenencia difusas para un valor indicatriz igual a la mediana de la distribución.	287
5.12. Funciones de transformación de los subíndices de los nitratos y el plomo.	289
5.13. Funciones de transformación de los subíndices de la conductividad y el pH.	289
5.14. Variograma experimental de la variable indicatriz mediana correspondiente al índice de calidad del agua creado mediante agregación exponencial. Mediana(WQI)=29.51.	291

5.15. Mapa de isoprobabilidad (krigeaje de indicatrices) de un índice de calidad del agua construido mediante la agregación exponencial para un valor indicatriz igual a la mediana de la distribución.	292
5.16. Funciones de distribución para seis de los municipios estudiados según cuatro índices de calidad.	293
5.17. Triángulo de decisión.	296
5.18. Variograma experimental y teórico de la piezometría de abril de 1980. . .	300
5.19. Mapa del espesor no saturado normalizado mediante una función de pertenencia.	301
5.20. Mapa de la recarga normalizado mediante una función de pertenencia. . .	301
5.21. Mapa de la topografía reclasificado mediante una función de pertenencia monótonamente decreciente.	303
5.22. Mapa de la litología de la zona no saturada reclasificado mediante una función de pertenencia monótonamente creciente.	304
5.23. Mapa de la conductividad hidráulica reclasificado mediante una función de pertenencia monótonamente decreciente.	304
5.24. Vulnerabilidad del acuífero evaluada mediante OWA.	307
5.25. Vulnerabilidad del acuífero evaluada mediante una combinación lineal de pesos (WLC).	307
5.26. Distribución espacial del contenido en nitratos del agua subterránea. . .	308
5.27. Esquema de la red neuronal utilizada para estimar la asociación entre variables de entrada y aptitud.	314
5.28. Estimación de la aptitud del agua para consumo público mediante redes neuronales.	315
5.29. Mapa de aptitud del agua subterránea según las estimaciones geoestadísticas.	316
5.30. Mapa de aptitud del agua en los puntos de muestreo (<i>según RD 140/2003</i>).317	317
5.31. Función de pertenencia triangular.	319
5.32. Mapas del contenido en nitratos del agua subterránea obtenidos mediante SBRD con dos variables de entrada.	322
5.33. Estimación del contenido de nitratos del agua subterránea mediante SBRD (experimento 2).	323
5.34. Mapas del contenido en nitratos del agua subterránea obtenidos mediante SBRD con cinco variables de entrada.	327
5.35. Estimación del contenido en nitratos del agua del acuífero mediante SBRD.328	328
5.36. Mapas del contenido en nitratos del agua subterránea obtenidos mediante SBRD con ocho variables de entrada.	329

5.37. Mapas del contenido en nitratos del agua subterránea obtenidos mediante
SBRD con cinco variables de entrada. 330

Índice de tablas

2.1. Reclasificación de la litología de las columnas de sondeos.	21
2.2. Métodos analíticos utilizados para el análisis de las muestras de aguas subterráneas.	34
2.3. Relación de caudales y aportaciones medias de cauces y canales. Fuente: Antonio Castillo Martín. 1986	78
3.1. Ejemplo de algunos de los datos primarios y secundarios disponibles en la base de datos del acuífero de la Vega de Granada.	134
3.1. Modelos teóricos de variograma utilizados en este trabajo.	157
3.2. Elementos de un sistema de teledetección (figura 3.13).	165
3.3. Funciones de activación utilizadas en esta investigación.	187
4.1. Estadísticos básicos de los elementos fisicoquímicos primarios del agua.	211
4.2. Estadísticos básicos de los metales pesados del suelo.	211
4.3. Estadísticos básicos de la microbiología del agua.	211
4.4. Estadísticos básicos de los elementos químicos de la solución del suelo.	212
4.5. Estadísticos básicos de los parámetros relacionados con el acuífero.	212
4.6. Resumen estadístico de las variables del agua analizadas no aptas para consumo público.	214
4.7. Valores medios de variables hidroquímicas y del acuífero correspondientes a los tres clústeres.	232
4.8. Cargas de los componentes principales y varianza para los 5 componentes con la rotación Quartimax.	237
4.9. Tabla resumen de los modelos teóricos ajustados a los factores del ACP.	239
4.10. Variables usadas en el análisis clúster de los parámetros físicos del suelo.	242
4.11. Valores promedios de las variables del suelo para cada uno de los clústeres.	244
4.12. Resultados del ACP: varianza total y porcentaje que explica cada componente.	245
4.13. Resultados del ACP: matriz de componentes rotados. (<i>En negrita los vectores propios mayores para cada componente</i>).	247

4.14. Ajustes de los variogramas de los factores.	249
4.15. Parámetros de los modelos esféricos ajustados a las variables de las aguas subterráneas.	254
4.16. Ajustes de modelos esféricos de variograma de las variables físicas y químicas del suelo. <i>1: horizonte 80-100 cm; 2: Horizonte 0-20 cm; 3: Horizonte medio. *: Transformación logarítmica.</i>	259
4.17. Parámetros estadísticos de las distintas transformaciones de la transmisividad.	263
5.1. Criterios de aptitud y porcentaje de aptitud para las variables de calidad del agua analizadas.	277
5.2. Intervalos de transformación y pesos del índice basado en Poch por intervalos.	282
5.3. Intervalos de transformación para el índice ICG y pesos.	283
5.4. Ajustes a modelos teóricos de las variables indicatriz de los índices de calidad del agua.	284
5.5. Intervalos de transformación y pesos del índice creado a través de funciones sigma.	286
5.6. Ajuste del modelo teórico de las variable indicatriz del índice de calidad del agua creado mediante funciones difusas.	287
5.7. Ajustes del modelo teórico de la variable indicatriz del índice de calidad del agua.	291
5.8. Resumen de los datos de origen utilizados para calcular la vulnerabilidad.	298
5.9. Ejemplo de asignación de pesos de orden (OWA).	299
5.10. Parámetros del modelo teórico del ajuste del variograma de la piezometría de abril de 1980.	300
5.11. Matriz de comparación pareada para la asignación de pesos de vulnerabilidad.	306
5.12. Pesos ordenados para la evaluación de la vulnerabilidad.	307
5.13. Variables utilizadas en el modelo de estimación de la calidad del agua subterránea.	311
5.14. Variables incluidas en el SBRD para la inferencia de la vulnerabilidad.	320
5.15. Tabla de errores de lógica difusa.	321
5.16. Base de reglas del experimento 2. (<i>2 variables de entrada y 5 etiquetas</i>).	321
5.17. Base de reglas con 5 variables de entrada y 5 etiquetas.	323

Capítulo 1

Introducción

"Lo importante no es dominar la Naturaleza, para al final acabar por destruirla, sino que el hombre debe asociarse a ella para conservarla en todo su potencial, y ponerla a disposición de las generaciones futuras para su ordenado disfrute."

René Dumont, La utopía o la muerte.

1.1. Aspectos generales

El agua es un recurso indispensable para el desarrollo socio-económico tanto de los países industrializados como en vías de desarrollo. El estudio de su calidad y cantidad es un tema que preocupa cada vez más a la sociedad actual, y es por ello que los organismos tanto nacionales como internacionales tratan de establecer políticas de desarrollo sostenible dirigidas hacia una gestión integral de los recursos hídricos, donde la prevención y el aprovechamiento de los mismos son el eje fundamental (directiva 2000/60/CE).

En este sentido, el estudio de la variabilidad espacio-temporal de la calidad es un aspecto de especial relevancia en la gestión global de los recursos hídricos. En este contexto, este trabajo pretende aportar nuevas herramientas al estudio clásico de la calidad del agua, integrando métodos avanzados de extracción de conocimiento en un entorno espacial (Sistemas de Información Geográfica o SIG), que permitan una mejor gestión, actualización y uso de esos datos en el territorio. Para este fin, se requiere un compendio muy amplio de diferentes tipos de datos.

En el conjunto de información necesaria para abordar un trabajo de esta naturaleza se pueden establecer diferentes categorías de información, así no cabe duda que el pilar principal está referido directamente a los datos químicos experimentales del agua del acuífero y de los ríos que lo alimentan, seguido del estudio de la granulometría y composición de los suelos a través de los cuales se infiltra una parte de estos recursos que

alimentan el acuífero. A continuación, en otro plano también muy importante, puesto que dentro de la preservación de la calidad de las aguas subterráneas uno de los aspectos más importantes, es el de vulnerabilidad, se encuentran todas las variables que están en relación con las propiedades hidrogeológicas e hidrodinámicas del acuífero, como son la piezometría, la conductividad hidráulica, el espesor no saturado, la geología, la litología de la zona no saturada, etc. Por último, en el plano de la gestión se encuentra la información socio-económica vinculada al acuífero estudiado, donde entran tanto la demanda de agua de los municipios, la distribución de los polígonos industriales, usos del suelo, las actividades ganaderas, las agrarias, los focos de contaminación, etc. como los valores medioambientales de interés social.

La zona de estudio elegida corresponde al acuífero de la Vega de Granada y su entorno próximo, que desde el punto de vista hidrográfico forma parte de la subcuenca del Alto Genil, dentro de la cuenca hidrográfica del Guadalquivir. El área corresponde a una extensa planicie que se extiende a ambos márgenes del río Genil, entre las poblaciones de Cenes de la Vega, al este y Láchar, al oeste. Este acuífero está compuesto por materiales detríticos cuaternarios que rellenan parte de la cuenca intramontañosa de la depresión de Granada y provienen de la erosión de los relieves circundantes. Todo este depósito ha generado una potente masa acuífera (acuífero multicapa) con una extensión próxima a los 200 km^2 , unos recursos de más de $150 \text{ hm}^3/\text{a}$ y unas reservas explotables máximas del orden de 1000 hm^3 . Este acuífero es uno de los mejor estudiados y conocidos en Andalucía por ser también uno de los más importantes, si bien aún persisten algunas dudas sobre sus características y potencialidad hidráulica y sobre la cuantificación de ciertas partidas de balance.

En el caso de la Vega de Granada, la integración de los datos socio-económicos es de gran interés debido a las importantes actividades agrícolas, industriales y urbanísticas vinculadas al acuífero entre las cuales se establece una fuerte competencia por el uso del territorio. En conjunto, una parte considerable de estas actividades usan directamente los recursos de este acuífero, aunque en este momento se está cambiando el modelo de gestión del mismo, como se mencionará más adelante. Es de destacar que las cualidades y situación estratégica del área propiciaron, desde época secular, el asentamiento de numerosos núcleos urbanos. Hoy día son cerca de la treintena los existentes, incluida Granada capital; en ellos se concentra una población estable de 400.000 habitantes, dando lugar a la mayor densidad demográfica de la provincia; 40.000 personas dependen hoy día, en exclusividad, del abastecimiento de aguas subterráneas, estando previsto el aumento de esta dependencia con la realización de nuevas captaciones. La abundancia de aguas y la fertilidad del suelo ha propiciado además, un intenso aprovechamiento agrícola de las aguas, con una superficie de regadío de más de 15.000 ha. Hoy día, los cultivos de regadío se reparten entre cereal, choperas, hortalizas, maíz y tabaco. La industria

está escasamente implantada en el área, no tanto por el número de instalaciones, como, sobre todo, por el tamaño, grado de transformación y naturaleza de las mismas. Desde la década de los ochenta este área está sumida por una paulatina crisis agrícola, originada por la baja rentabilidad de los productos y métodos de cultivo tradicionales lo que está produciendo un imparable empuje urbanizador, ocupando excelentes tierras de vega, que al margen de los impactos medioambientales que está provocando, es de muy discutible idoneidad por los riesgos sísmicos, de subsidencia y avenidas e inundaciones a los que se puede ver sometido. Por otra parte, y desde la óptica de los recursos del acuífero, supone un aumento considerable de riesgo químico y pérdida de capacidad de infiltración.

Pese a que el área posee unas magníficas aptitudes para la agricultura, uno de los problemas actuales, como en otros muchos lugares, está en la deficiente calidad que empiezan a tener las aguas de riego por aumento progresivo de la salinidad y aportaciones de aguas superficiales de peor calidad. Las de superficie, procedentes de la escorrentía de Sierra Nevada son de excelente calidad, pero su garantía de suministro es relativamente baja. Las residuales depuradas no tienen la calidad suficiente para permitir riegos de productos de primor (huerta), ni el empleo de técnicas modernas de fertirrigación localizadas ni automatizadas. Es por ello que una de las grandes bazas de este vasto espacio agrícola de regadío está en el empleo de aguas subterráneas del acuífero. Si bien la cantidad del recurso es más que suficiente (el acuífero es excedentario) y las elevaciones no son altas, son cada vez más frecuentes los problemas de calidad, que conviene afrontar, ya que son limitantes para algunos cultivos y usos. Así por ejemplo, la contaminación de nitratos, que no es nueva, empieza a alcanzar valores y extensiones considerables que impiden su uso para los abastecimientos a poblaciones, localizados en la parte central del acuífero y extremo oriental, esta zona como suministro emergencia fundamentalmente. Por otra parte los cloruros, cuyos niveles son aún moderados, está sufriendo un incremento en los últimos años, aunque mal caracterizado todavía, debido a cambios antrópicos en la hidrodinámica del acuífero; ello ocasiona serios problemas al cultivo del tabaco, uno de los más extendidos de la Vega de Granada, aunque en rápida disminución. En consecuencia, la calidad de las aguas subterráneas está sufriendo una degradación progresiva como resultado de la actividad agrícola y urbana.

La evaluación del grado de afección medioambiental que estos factores tienen sobre la calidad de los recursos hídricos del acuífero es un aspecto de especial importancia puesto de relieve por los organismos gestores, no existiendo en la actualidad una herramienta que facilite las tareas de gestión y planificación integral de la calidad de dichos recursos.

A esta problemática se pretende responder en este trabajo de tesis, donde se desarrolla una metodología basada en las nuevas técnicas de tratamiento de información para implementar un sistema de soporte de decisiones para la gestión integral de la calidad de los recursos hídricos estudiados. En este sentido, los Sistemas de Información

Geográfica, la Teledetección Espacial, la Geoestadística y la Inteligencia Artificial constituyen la base metodológica escogida para abordar la problemática planteada, validada en muchos campos de los recursos naturales y en la que se fundamenta este trabajo.

1.2. Antecedentes

La investigación del acuífero de la Vega de Granada se inicia en los años 60, cuando la FAO (1972, [1]) lleva a cabo el primer estudio hidrogeológico de importancia, que sentó las bases para la observación piezométrica y del control hidrogeoquímico e hidrométrico, que aún perduran en la actualidad. Desde este inicio ha sido el IGME, principalmente, la institución que se ha encargado del seguimiento de las investigaciones hidrogeológicas. Desde la Universidad de Granada también se han impulsado investigaciones en tal sentido a través del Instituto del Agua y de otros grupos de investigación. En concreto, las investigaciones más relevantes corresponden a la tesis realizada por Castillo Martín (1986, [30]), que aportan una de las visiones más completas del funcionamiento del acuífero. Posteriormente, otros estudios más específicos han sido realizados en la zona, en relación con la piezometría del acuífero (Luque Espinar, 2001 [110]), el impacto del riego con aguas residuales en la zona no saturada (El Mabrouki, 2003 [68]), el fraccionamiento de los metales del suelo (Campos Fernández-Fígarez, 1997[27]) o las transformaciones recientes del espacio agrario (Menor Toribio, 1998 [122]). En todos estos casos, de forma indirecta la investigación trata de responder a un mismo problema social, como es el riesgo ambiental que se está provocando debido al rápido crecimiento socio-económico e industrial de la zona. Riesgos, que pueden afectar directamente a la salud pública, cuando hablamos del agua subterránea, ya que este recurso es la única fuente de agua potable para muchos municipios de la comarca en caso de emergencia.

En este contexto, radica el interés de este trabajo, que trata de integrar la abundante información existente y la recogida durante estos años de investigación, para una correcta gestión de la calidad de los recursos hídricos. La correcta gestión requiere establecer relaciones con los factores medioambientales y antrópicos que pueden estar alterando dicha calidad, lo cual supone una visión novedosa hasta la fecha.

Con el avance la informática, las tecnologías de análisis de datos espaciales como los SIG y la Teledetección han experimentado un gran desarrollo, particularmente en el ámbito del estudio de los recursos naturales. Tanto es así, que actualmente muchos organismos están introduciendo estas herramientas para la gestión, el estudio, el análisis y la evaluación tanto de la información disponible como de sus resultados. Algunos ejemplos de estas aplicaciones son: el SIAS¹, Sistema de Información del Agua Subterránea, que

¹SIAS <http://www.igme.es/internet/ServiciosMapas/siasweb/sias-an.html>

pone a disposición de los usuarios las bases de datos hidrogeológicas; el SIGMA², Sistema de Información Geológico-Minero de Andalucía, del servicio de Industria, Energía y Minas; el (SIGA)³ Sistema de Información Geográfica de datos Agrarios ; el REDIAM⁴, Red de Información Ambiental de Andalucía, o el SIG oleícola⁵, entre otros.

A la vista de las voluminosas bases de datos espacio-temporales disponibles, se hace necesaria una herramienta más efectiva para la comunicación de la información ambiental, que extraiga información que sea interpretable por los gestores y ayude a la toma de decisiones. Estas herramientas son los Sistemas de Soporte de Decisión (SSD), cada vez más utilizadas en este área e integrados en los SIG. Los sistemas de ayuda a la decisión surgen de la necesidad de afrontar los problemas relacionados con la gestión de modo racional y coherente, valorando los múltiples aspectos a tener en consideración, utilizando algún método que le permita tener en cuenta las distintas ventajas e inconvenientes.

Los Sistemas de Soporte de Decisión (SSD) o de ayuda a la gestión tienen una característica principal, y es su actual auge. La generalización de las herramientas informáticas, junto con la mundialización de la información y los métodos científicos, ha favorecido enormemente la extensión de este tipo de sistemas por la mayoría de los países ricos. Por otra parte, esta extensión está directamente ligada a su variedad y heterogeneidad tanto en los objetivos, como en la estructura y los métodos utilizados. Analizando el conjunto de sistemas de ayuda a la decisión existentes, se puede observar que cada uno hace hincapié en un aspecto determinado de su funcionamiento, descuidando quizás otros. Hay sistemas cerrados que resuelven problemas concretos y que no pueden ser usados para nada más, aunque también hay otros más versátiles y que pueden ser aplicados a multitud de situaciones diferentes. En ocasiones nos encontramos sistemas con una gran potencia a la hora de generar modelos, pero débilmente conectados con un SIG, o a la inversa. Es posible que esto se deba a que se trata de una disciplina relativamente nueva en la que todavía no hay normas fijas; las bases teóricas todavía no están del todo afianzadas, ni los métodos, ni tan siquiera los lenguajes de programación. En definitiva abundan las distintas soluciones para problemas similares. La "selección natural" (o el azar y la publicidad probablemente) actuará poco a poco dejando aquellos sistemas que sean realmente útiles para los gestores.

Entre algunos de los sistemas cerrados, podemos encontrar los siguientes:

- Sistema H2O que es un programa en entorno Windows desarrollado por Guillermo Collazos, que sirve para configurar (crear o modificar) modelos de Sistemas de

²SIGMA <http://www.juntadeandalucia.es/empleoydesarrollotecnologico/w08/sigma/>

³SIGA <http://www.mapa.es/es/sig/pags/siga/intro.htm>

⁴REDIAM <http://www.juntadeandalucia.es/medioambiente/site/web/menuitem.504860762adc832f99bea88060425ea0/?vgnnextoid=178411a12a7ae010VgnVCM100000624e50aRCRD>

⁵SIGOleico http://w3.mapya.es/dinatierra_v3/

Recursos Hídricos (S.R.H.), y simular u optimizar la asignación de agua cruda. Es un programa genérico, que permite representar cualquier cuenca o SRH que el usuario desee estudiar. En la planificación y gestión sostenible del agua, los SSD permiten obtener indicadores de funcionamiento del sistema analizado, para comparar de forma homogénea las alternativas existentes; dando precisión a la experiencia e intuición del analista. Un SSD ayuda a tomar decisiones, permitiendo responder preguntas del tipo: ¿Qué pasaría si....? en un SRH modelado.

- El sistema MOIRA, que comienza a desarrollarse en el año 1996 como parte de un proyecto europeo de colaboración entre las universidades de Uppsala (Suecia) y Oslo (Noruega), el ENEA (Italia), las empresas KEMA (Holanda) y Studsvik (Suecia) y el Departamento de Tecnología Nuclear y el Grupo de Análisis de Decisiones de la Universidad Politécnica de Madrid, con el objetivo de su aplicación a ecosistemas altamente contaminados en los citados países. Este sistema de ayuda (SAD) se desarrolla para un problema de contaminación en caso de introducción accidental de sustancias radiactivas en ecosistemas hídricos, siendo el objetivo básico la construcción de un SAD basado en modelos predictivos validados entre los que se incluyó un módulo de evaluación para la ordenación de las posibles alternativas de restauración de acuerdo con sus potenciales beneficios.
- El IctNeo que surgió como fruto de la colaboración entre el Hospital General Universitario Gregorio Marañón de Madrid y el Grupo de Análisis de Decisiones de la Facultad de Informática de la Universidad Politécnica de Madrid y está enfocado a un problema médico, como es la ictericia, tratando de conseguir un equilibrio entre el infra y el sobretreatmento, además de reducir los costes de diagnóstico, riesgos de pruebas, preferencias de médicos, etc.

Todos estos sistemas como ya se ha comentado antes, requieren la especificación de un problema concreto, aunque son muchos otros los sistemas comerciales o libres abiertos que existen para la toma de decisiones. Sin embargo, en el proceso de detallada búsqueda encontramos algunas herramientas bastante interesantes, pero en muchos casos, su coste, su escasa adaptabilidad al entorno SIG o la limitada capacidad de cálculo matemático limitaron su uso, optando así al desarrollo propio de las metodologías que se consideraron más interesantes, no sin previamente destacar y conocer los aspectos más interesantes de cada uno de ellos. Entre los sistemas de soporte de decisión abiertos más interesantes encontramos los siguientes:

- El sistema EMDS (Ecosystem Management Decision Support) (Reynolds, 1998 [150]), financiado por el departamento de agricultura de Estados Unidos (USDA) bajo la dirección del Dr. Keith Reynolds, responsable de la oficina noreste del departamento forestal de su país. El EMDS permite la integración de un generador

de modelos basado en objetos con un SIG comercial, pero el principal interés de este sistema es que su motor matemático se fundamenta en las bases de conocimiento y el tipo de razonamiento que derivan de ellas, basándose en la lógica difusa, permitiendo la evaluación de aspectos abstractos y complejos introduciendo la información, en muchos casos cualitativa de la que disponen los gestores.

- El sistema DPL (Decision Programming Language) de Call y Miller (ADA Decision Systems, Menlo Park, CA) es un sistema diseñado para problemas de Análisis de Decisiones. Contiene un editor integrado para especificar los problemas en un lenguaje propio, un compilador, una unidad de ejecución, una gran variedad de formas de mostrar los resultados y el análisis, y varias posibilidades de análisis de sensibilidad. El principal problema de este sistema es que su integración con un SIG es más complicado y a pesar de su potencialidad, solamente es útil para su trabajo con árboles y tablas de decisión.
- El módulo ModelBuilder de ESRI (2000, [69]) para el software comercial ArcView, que facilita la construcción de modelos sencillos espaciales. Este programa también dispone de una interfaz gráfica amigable desde la cual se van añadiendo objetos que corresponden a las distintas entidades que se van a analizar, las cuales son capas propias de un SIG. En este sentido, la aplicación es muy interesante, pero por otra parte flaquea a la hora de realizar cálculos más complejos que simples operaciones aritméticas de suma, resta, producto, etc.

A la vista de estos ejemplos de Sistemas de Soporte de Decisión, en este trabajo se decide no optar por ningún sistema concreto, aunque es necesario tenerlos presentes para justificar los métodos y análisis que se aplicarán en este trabajo.

El motivo por el cual se ha decidido crear un SSD no limitado a un software concreto proviene de la complejidad del caso de estudio. La evaluación del acuífero de la Vega de Granada, como se ha comentado en los aspectos generales, requiere de datos de distinta naturaleza, y diverso origen, que requieren ser tratados específicamente para obtener información útil y versátil que mejore el conocimiento sobre el acuífero.

Hoy día las grandes bases de datos sobre el territorio se realizan en formato espacial (REDIAM, SIGA, SIGMA, etc.) y es el SIG el núcleo de nuestro SSD. Como se ha visto en los distintos ejemplos de SSD, éstos pretenden ayudar al gestor a tomar una decisión, ya sea incorporando el conocimiento del gestor al sistema (modelos basados en conocimiento experto) o ya sea extrayendo información de los datos (modelos basados en los datos). Estos dos puntos de vista se abordan en este trabajo de tesis y se incluyen en el capítulo 5. Los modelos basados en conocimiento experto han sido ampliamente aplicados a los recursos hídricos desde diversas perspectivas, con diversos ejemplos para el estudio de vulnerabilidad (Gemitzi et al, 2005; Doerfliger, et al, 1999; Jiang y East-

man, 2000; Van der Werf y Zimmer, 1998; y Roussel, 2000 [60, 77, 102, 183, 158]). En estos ejemplos tienen cabida tanto las técnicas multicriterio como la lógica difusa, y en todos se trata de evaluar la vulnerabilidad de un acuífero supeditado al conocimiento experto.

Sin embargo, en ejemplos con la ingente cantidad de información como en el caso de la Vega de Granada, el experto generalmente tiene dificultades para ponderar y discriminar la información. Es por esta razón, por la que los sistemas basados en Inteligencia Artificial (IA) capaces de extraer conocimiento sin necesidad de supuestos sobre los datos están siendo cada vez más utilizados en todas las disciplinas que incluyen las ciencias de la Tierra. Los principales métodos basados en IA son la lógica difusa, las redes neuronales artificiales y los algoritmos genéticos. En estas líneas también se han desarrollado actualmente diversos trabajos relacionados con la hidrogeología, fundamentalmente en la evaluación de la vulnerabilidad de los acuíferos (Dixon, 2002 y 2005a y b; Andujar et al, 2006; Li Huang, et al, 2007; Wang et al, 2006; Muhammetoglu et al, 2006 [59, 58, 57, 11, 106, 190, 125]), aunque también en otros campos afines como la predicción de la contaminación en las aguas subterráneas o la predicción de los niveles piezométricos de los acuíferos (Gümrah et al, 2004; Nourani et al, 2008 [80, 129]).

1.3. Objetivos

El objetivo final que se plantea en la elaboración de este trabajo, es el de avanzar en el desarrollo de una metodología multidisciplinar para el estudio de la calidad de los recursos hídricos subterráneos de la "Vega de Granada", mediante la aplicación de Sistemas de Información Geográfica, Teledetección, Geoestadística e Inteligencia Artificial, entre los más destacados. En este estudio se tendrán en cuenta los factores medioambientales que están alterando la calidad de las aguas de este acuífero, uno de los más importantes de Andalucía.

El enfoque multitemático de la investigación va dirigido a la elaboración de métodos de integración de los parámetros de calidad, hidrogeoquímicos y medioambientales, principalmente, así como otras variables complementarias referentes al medio físico y al contexto socio-económico.

Todo ello constituye una aportación novedosa e integradora de la información geoespacial, tanto existente como experimental, fundamentada en la interpretación de imágenes digitales de satélite y Sistemas de Información Geográfica como herramienta de integración. El resultado final constituye una herramienta de apoyo a la toma de decisiones, de interés para futuros programas de control hidrogeoquímico y medioambiental del acuífero.

Esta ambiciosa propuesta requiere establecer una serie de objetivos concretos de acuerdo con la información disponible y experimental a estudiar y de acuerdo con los estudios ya realizados en áreas similares, los cuales se detallan a continuación:

- Recopilación de la información disponible del área de estudio. Así mismo, asimilación de las bases científicas necesarias: hidrogeología, hidroquímica, edafología, etc. para el estudio y el interés de los nuevos datos a adquirir.
- Análisis de los fundamentos teóricos de las herramientas informáticas de asistencia en la toma de decisiones. Este trabajo comienza haciendo una revisión exhaustiva a conceptos como la teoría de la decisión, los Sistemas de Soporte de Decisión, los Sistemas de Información Geográfica, la Teledetección, la Inteligencia Artificial, etc., además de estudiar las posibles ventajas o potenciales limitaciones de los distintos métodos, y sobre todo, de los resultados aparentes.
- Análisis de los distintos sistemas de apoyo a la decisión aplicados a la gestión del medio natural que se usan en la actualidad. Revisión de las distintas aproximaciones que se han puesto en marcha en la actualidad y que pudieran tener aspectos comunes a este trabajo, tanto en su aproximación global como en aspectos más específicos de aplicación.
- Desarrollo y construcción de una base de datos geoespacial que integre la información disponible y analizada del área de estudio (espacial o no) en un formato que permita su actualización y extracción, teniendo en cuenta futuros posibles usos de la misma.
- Análisis de datos estadística y espacialmente, caracterización de tendencias, relaciones entre variables y determinación de datos erróneos o aberrantes (outliers).
- Integración de datos mediante la elaboración de diferentes índices, tanto de calidad como de vulnerabilidad del acuífero, que permitan una gestión de la información más sencilla y más rápida por parte de los gestores
- Creación de diferentes modelos de predicción de la calidad de los recursos hídricos subterráneos, comparación de los métodos y de sus resultados.
- Aplicación de las nuevas metodologías basadas en Inteligencia Artificial al proceso de predicción de la calidad del agua.
- Simulación de diferentes escenarios en la evaluación de la calidad del agua basados en SIG.

1.4. Organización de la memoria

En este apartado se describe brevemente la organización de la memoria, que podemos considerar dividida en cuatro partes. Una primera parte de introducción y datos (Capítulo 1 y 2), una segunda parte de metodologías (Capítulo 3), una tercera parte de análisis (Capítulo 4) y una cuarta parte de ensayo (Capítulo 5). Además las conclusiones más relevantes se incluyen en el apartado final.

La primera parte, correspondiente al capítulo 2, es la parte más importante en la que se describen los datos utilizados para la elaboración del trabajo. La importancia de este capítulo estriba en la gran variedad y heterogeneidad de los datos, los cuales han debido de estandarizarse y actualizarse tanto en formato como en análisis, ya que en este apartado se incluyen tanto los datos experimentales tomados durante las distintas campañas, así como otros datos experimentales o digitalizados más antiguos, cedidos por diferentes organismos o personas. El trabajo con este tipo de datos más antiguos ha requerido en muchos casos, tratar de discernir la información útil y fiable entre informes antiguos y no organizados. En este apartado se muestran los datos una vez estudiados y analizados, y se comparan con los resultados obtenidos en los trabajos anteriores, además de incluir los nuevos resultados obtenidos a la luz de los datos actuales.

En el capítulo 3, se detallan las bases metodológicas fundamentales de las cuales se parte, comenzando con la construcción de la base de datos geoespacial, fundamental para una correcta manipulación de la información. Las técnicas y métodos con principal relevancia son la Geoestadística, la Teledetección, la Inteligencia Artificial y los Sistemas de Información Geográfica, todos ellos integrados en la base de datos geoespacial. La integración de estos métodos en un contexto espacial (SIG) está enfocada a la extracción de conocimiento o "información" necesaria para la toma de decisiones, que es el objetivo final del análisis.

El capítulo 4, explica con más detalle algunos de los ejemplos más relevantes obtenidos en la aplicación de los distintos métodos, como siguiente apartado y fase primera en el tratamiento de los datos, se muestran algunos ejemplos de los análisis, estadísticos, geoestadísticos uni y multivariantes, con sus correspondientes ejemplos.

En el capítulo 5 se aplican las técnicas más novedosas de integración partiendo de la modelización booleana, pasando por los métodos de evaluación multicriterio y terminando con la implementación de una red neuronal para la evaluación de la vulnerabilidad del acuífero y un sistema basado en reglas difusas con el mismo objetivo.

Finalmente en el capítulo 6 se exponen detalladamente las principales conclusiones del trabajo, donde se realiza una discusión general y se dan algunas recomendaciones y detalles sobre posibles nuevas vías de trabajo.

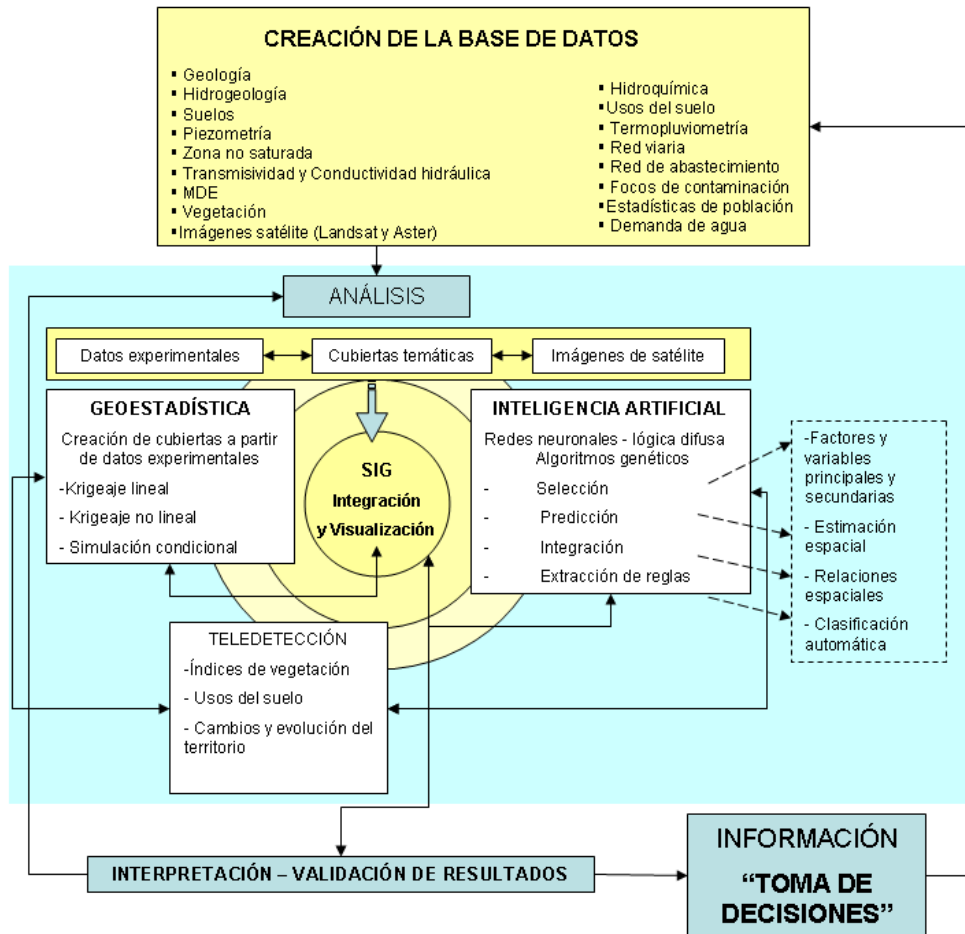


Figura 1.1: Esquema de la metodología del trabajo.

Una visión más esquemática de la metodología seguida para la adquisición de los objetivos planteados, se detalla en la figura 1.1, donde cabe destacar la retroalimentación para los distintos tipos de análisis.

Capítulo 2

El área de estudio: La Vega de Granada

2.1. Localización geográfica

El acuífero de la Vega de Granada está situado al sur de la Península Ibérica, dentro de la provincia de Granada, al oeste y sur de Granada ciudad (figura 2.1). Tomadas como referencia las hojas topográficas de escala 1:50.000 del Servicio Cartográfico del Ejército se localiza en las hojas 19-41 (1009) de Granada, y 19-42 (1026) de Padul.

Desde el punto de vista hidrográfico se sitúa en la Cuenca Alta del río Genil, donde confluyen además del río Genil, los ríos Aguas Blancas, Darro, Beiro, Cubillas, Velillos, Monachil y Dílar, todos ellos afluentes del primero

Morfológicamente, el acuífero coincide con la denominada Vega de Granada y ocupa una superficie esencialmente plana de unos 200 km^2 , que se inclina hacia el oeste en el sentido de drenaje del acuífero, coincidiendo este sector con la población de Láchar. Superados estos límites morfológicos, el relieve se hace más o menos abrupto dependiendo de la dirección. Una visualización del área de estudio a partir de la ortofoto y del modelo digital del terreno (figura 2.2), nos da una idea aproximada de la situación del acuífero de la Vega de Granada así como del relieve y los usos del suelo.

2.2. Marco hidrogeológico

2.2.1. Geología y geometría del acuífero

La depresión de la Vega de Granada, donde se localiza el área de estudio, presenta un relleno de naturaleza postorogénica de edad Neógeno a Cuaternario. Estos depósitos ocultan en este sector el contacto entre dos grandes dominios estructurales, las zonas Externas e Internas de las Cordilleras Béticas (figura 2.3).

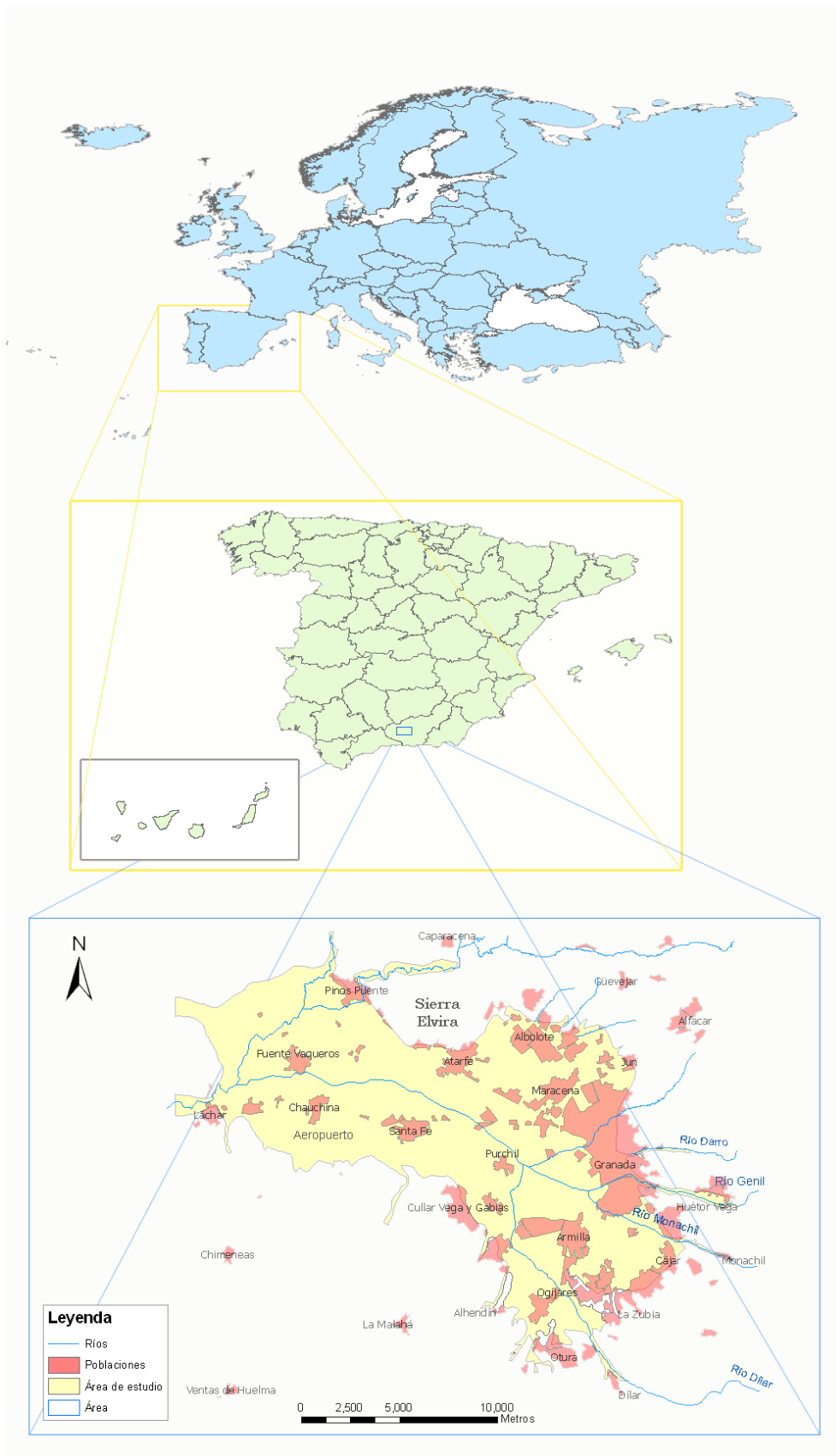


Figura 2.1: Situación geográfica de la Vega de Granada.



Figura 2.2: Relieve y ortofoto de la Vega de Granada.

La geología de esta depresión es relativamente compleja, dados los continuos cambios de facies y de potencia de sus materiales, unido a la dificultad de datación de los mismos. En este apartado se describen resumidamente la serie tipo de los materiales que constituyen el entorno de la Vega de Granada, como apoyo al capítulo de hidrogeología y al conocimiento hidrogeológico de la región.

Como referencia fundamental se han tomado los mapas de la figuras 2.4, aunque para una descripción más detallada de los materiales presentes se han utilizado las Memorias de las hojas geológicas de Granada (19-41) y Padul (19-42) de la serie Magna, editadas por el Servicio de Publicaciones del Ministerio de Industria y Energía en los años 1985 y 1980, respectivamente. También se han tenido en cuenta los trabajos de otros autores como FAO e IGME (1972 y 1981, [1, 3]), Castillo Martín (1986 [30]), Luque Espinar (2001[110]) y Roldán García (1981 [156]).

Los materiales colindantes más representativos del entorno se citan a continuación en orden cronológico:

1. Zonas internas:

- a) Trías medio superior: Dolomías del manto del Trevenque y calizas masivas y tableadas del Manto de Víboras.

Estos materiales pertenecen al complejo Alpujárride, las calizas del manto de Víboras alcanzan un espesor de 180 m y las dolomías (de tonos blancos y grises azulados) generalmente se encuentran intensamente brechificadas.

- b) Mioceno medio: Margas versicolores, margocalizas y calizas bioclásticas.

Es un conjunto muy heterogéneo de materiales, con una edad que comprende desde el Maastrichtiense inferior al Mioceno medio, presenta una homogenei-

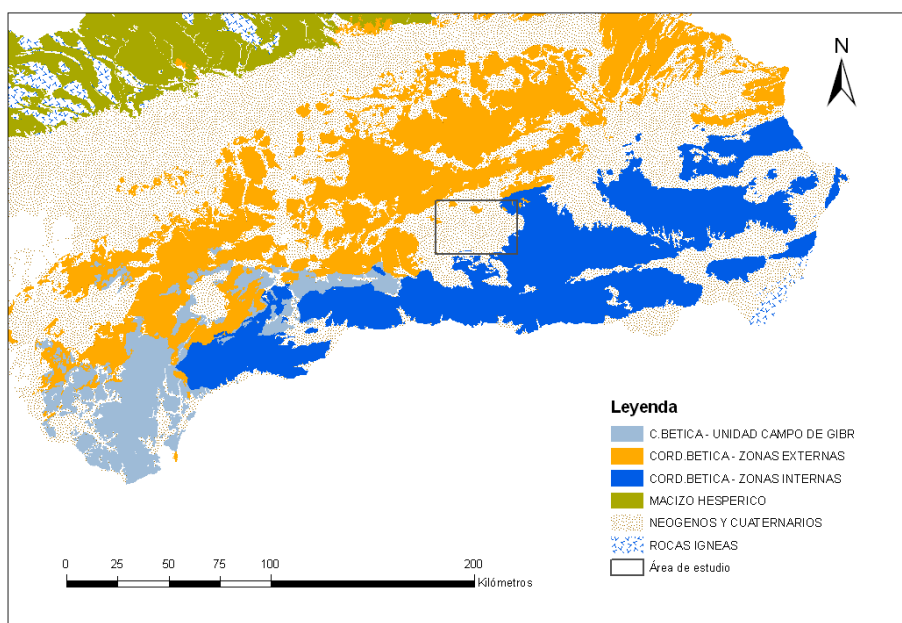


Figura 2.3: Situación de la región estudiada en un mapa geológico en el que se representan las principales unidades estudiadas de la cordillera bética.

Modificado: Mapa Geológico de Andalucía (Junta de Andalucía, 1992)

dad litológica y una complejidad en los afloramientos considerable, por lo que en su conjunto se integran en el Flysh Circumbético.

2. Zonas externas:

a) Trías: Arcillas, margas y yesos (facies Keuper del Trías germano-andaluz) con intrusiones de andesitas. Carniolas al techo.

Afloran en forma de extrusión diapírica y se componen de margas y arcillas versicolores con intercalaciones de niveles de areniscas y masas de yeso. Entre estas arcillas triásicas se pueden observar intrusiones de rocas ígneas de naturaleza dolerítica y dolomías y carniolas.

b) Lías: Dolomías, calizas con sílex y detríticas. Pillow lavas intercaladas.

Encima de las margas y arcillas triásicas se localizan unas dolomías grisáceas mal estratificadas, incluso normalmente brechificadas, con una potencia superior a los 200 m. A continuación, se dispone un paquete de 80 m de potencia de calizas bien estratificadas con laminaciones debidas a la acumulación de algas. Culminando la serie, se colocan unas calizas detríticas de crinoides que a techo presentan un hardground cuya potencia no llega a los 25 m.

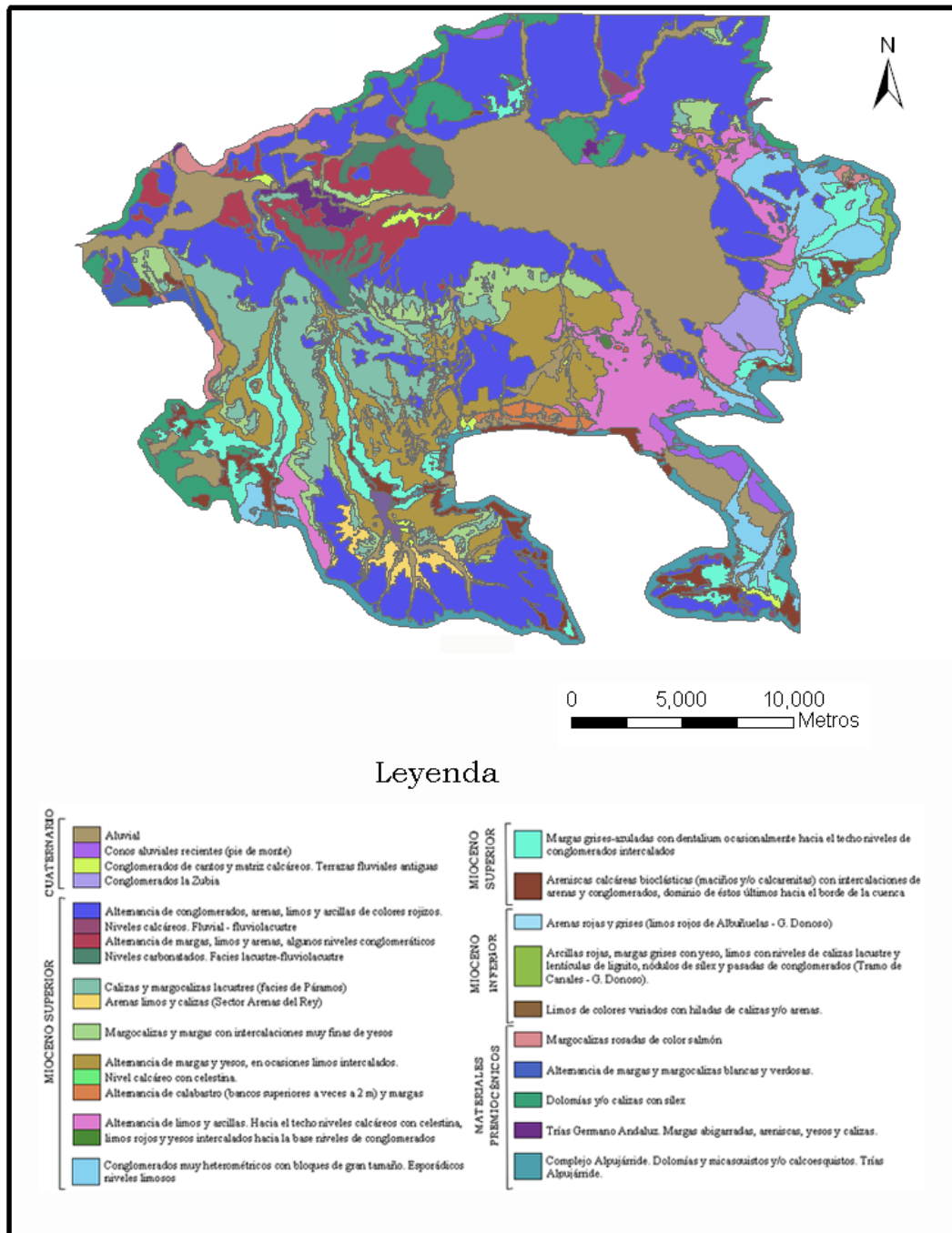


Figura 2.4: Mapa geológico del acuífero de la Vega de Granada y su entorno.

- c)* Jurásico superior: Margocalizas, radiolaritas y calizas margosas.
Se trata de una serie compuesta por margas y margocalizas con una potencia total de 90 m. Le siguen unas margocalizas silíceas y radiolaritas con unos 30 m de espesor de edad Dogger que culmina con unas calizas margosas con sílex que llegan a los 40 m de potencia.
- d)* Cretácico inferior: Margas y margocalizas.
Aparecen en términos bien estratificados en bancos de 10-15 cm con fósiles característicos del Neocomiense. En algunos sectores se aprecia mayor contenido en margas y ausencia de macrofauna. En series levantadas hacia el oeste de la Vega, la fauna encontrada llega hasta el Cretácico superior.

3. Depresión de Granada:

- a)* Mioceno inferior (Tortonense): Limos, areniscas con yesos y conglomerados.
Predominan los limos que presentan algunas intercalaciones de areniscas y arcillas e incluso yesos y otros materiales.
- b)* Mioceno superior (Turolense inferior): Conglomerados y arenas. Formación Pinos Genil.
Unidad constituida por cantos metamórficos y carbonatados con matriz arenolimsa muy rica en micas. Estos cantos son normalmente redondeados y con un diámetro que oscila entre 5 y 15 cm. Frecuentemente se intercalan entre los mismos niveles lenticulares de arena.
- c)* Mioceno superior (Turolense medio): Limos micáceos arenas y gravas.
Estos materiales son de origen lacustre y entre ellos también se puede encontrar yeso fibroso de origen secundario entre las fracturas existentes. Las gravas y las arenas abundan en la base de la secuencia en relación con los canales.
- d)* Mioceno superior (Turolense medio-superior): Arcillas grises, limos y arenas.
Serie arcillosa con niveles locales de arenas y limos micáceos, así como lignitos y, además, pueden aparecer niveles de yeso que abundan hacia techo.
- e)* Mioceno superior (Turolense superior): Yesos con lutitas y carbonatos.
Este conjunto está formado por niveles que se localizan en el tramo anteriormente descrito. Se pueden encontrar dos tipos de facies. En primer lugar niveles de yesos masivos intercalados en la parte baja del tramo arcilloso, y en segundo lugar, alternancias rítmicas de yeso con lutitas, y ocasionalmente, carbonatos. Estas intercalaciones afloran en sectores más occidentales de la depresión de Granada, con espesores que varían entre algo menos de 1 m hasta casi 40 m.

- f)* Plioceno basal: Calizas, margocalizas y limos lacustres.
Estos materiales de origen lacustre presentan calizas fundamentalmente oque-
rosas y de tonos claros (oeste de Santafé), mientras que hacia el este aparece
una alternancia de limos con yesos y margocalizas.
- g)* Plioceno: Margas y conglomerados.
Los materiales predominantes son lutitas y margas blancas y grises azula-
das con frecuentes episodios conglomeráticos y arenosos, cuyo ambiente de-
posicional corresponde a un lago somero alimentado por sedimentos de los
sistemas fluviales adyacentes.
- h)* Plioceno-Pleistoceno: Arcillas, limos rojos y conglomerados.
Conjunto compuesto, esencialmente, por conglomerados con cantos de gran
tamaño que pueden superar el metro y medio de diámetro en las proximidades
de los relieves circundantes. Al alejarnos de estos relieves el tamaño se reduce
hasta los 20 cm, y presentando un tamaño limo-arena en las facies distales,
con intercalaciones esporádicas de paquetes conglomeráticos.
- i)* Plioceno superior-Pleistoceno: Conglomerados y arenas. Formación Alham-
bra.
Estos materiales reposan en una discordancia angular sobre parte del tér-
mino anterior. En áreas proximales a los relieves circundantes las facies más
típicas son conglomerados desorganizados con cantos que llegan a tener 1,5
m de diámetro. En sectores más alejados de estos relieves, las facies más
frecuentes consisten en canalizaciones con tendencia granodecreciente hacia
techo, que se integran en megasecuencias estrato y granodecrecientes hacia
techo que terminan con el desarrollo de suelos. Dentro del esquema de este
conjunto, se diferencia unas facies proximales y otras intermedias, con uno
o varios abanicos que migrarían lateralmente originando las megasecuencias
descritas con anterioridad. Estos abanicos estarían nutridos de Sierra Nevada
y otros relieves.
- j)* Pleistoceno: Conglomerados. Formación Zubia.
Se trata de una potente masa de conglomerados de cemento calcáreo y can-
tos de igual naturaleza, de carácter heterométrico, redondeados y tamaño
superior a los 10 cm con intercalaciones de arenisca y cuya potencia puede
alcanzar los 160 m.
- k)* Pleistoceno-Holoceno: Arcillas y costras calcáreas. Glacis.
Es un nivel de arcillas rojas residual resultante de la inmensa meteorización
de la formación detrítica preexistente que corresponde a la fase de erosión
del dispositivo aluvial y lacustre que se inició con las margas claras y arcillas

del turolense terminal-Pleistoceno. Los niveles de encostramiento son calizos, discontinuos, muy porosos y se piensa que pertenecen a un episodio de aumento de la aridez, que culmina en el nivel superior de glacia.

- l) Pleistoceno superior - Holoceno inferior: Arcillas rojas, gravas y arenas. Paleosuelo.

Este término coincide con la denominada Vega Alta y está constituido por sedimentos aluviales con gran desarrollo de la llanura de inundación que se ordenan en secuencias positivas que presentan en su base un episodio canalizado de gravas y arenas, con un espesor en torno al metro. Por encima, se observan 2 m de limos y arcillas con frecuentes procesos de edafización que constituyen paleosuelos.

- m) Cuaternario reciente: Aluvial.

El aluvial está compuesto por un conjunto de materiales que conforman la Vega Baja, denominación que se realiza desde un punto de vista morfológico. Estos depósitos están ligados al funcionamiento actual o subactual de la red fluvial, con espesores observados en sondeos de hasta 300 m. Esta zona de depósito es actualmente subsidente según un eje NO-SE, formando una depresión que sirve de colector hacia el que discurren los afluentes del río Genil, los cuales presentan escaso desarrollo en sus depósitos aluviales. En el conjunto de estas formaciones aluviales se han podido diferenciar un número máximo de tres niveles de terrazas a una altura respecto a los cauces actuales entre 4 y 20 m.

Esta última secuencia litológica está recubierta de un potente suelo húmico arcilloso que oculta la geología de superficie y que no permite la cartografía de los cambios de facies bajo el mismo. Esta información ha sido inferida a través del estudio de los perfiles de geofísica eléctrica realizados dentro del proyecto FAO y a partir de diferentes sondeos realizados por el IGME, que sólo en algunos casos llegaban al sustrato. El estudio de estos sondeos, en concreto en un total de 222 puntos, ha permitido realizar para este trabajo, un análisis de la litología de la zona no saturada de este aluvial como se muestra a continuación.

La interpretación realizada por los sondistas de la litología de los testigos se ha dividido en seis clases (tabla 2.1), y posteriormente se ha ponderado teniendo en cuenta el espesor de cada material y la profundidad del sondeo, obteniendo un único valor correspondiente a la media de los materiales. Esta información ha sido krigeada (interpolación espacial) para elaborar una capa de información con la litología de la zona no saturada tomando como valores un mapa de clases. En este mapa los valores más cercanos a cero corresponden a materiales más finos y más cercanos a seis, corresponden a mate-

Reclasificación categórica	Litología
6	Grava limpia
5	Arena gruesa limpia
4	Mezcla de arena
3	Arena fina
2	Arena limosa
1	Limo
0	Arcilla

Tabla 2.1: Reclasificación de la litología de las columnas de sondeos.

riales más gruesos (figura 2.5). La distribución espacial de este promedio de materiales es coherente con el modelo deposicional interpretado por Castillo (1986, [30]), que en líneas generales describe como una franja central delimitada por el curso del río Genil con preponderancia en la vertical de gravas gruesas y una gradación hasta fracciones tipo arcilla al alejarnos de la misma. No obstante, la estimación mostrada en la figura 2.1 aporta algunas diferencias. La distribución espacial está menos suavizada, más acorde con la información aportada por los sondeos. Por otra parte, matiza el modelo de aportes de acarreo procedentes de los relieves circundantes puesto que muestra mejor la influencia de algunos cauces y torrentes en la granulometría de los depósitos de la Vega.

El substrato de este paquete de materiales aluviales no es conocido salvo en los bordes, donde lo constituye el resto de los materiales neógenos e incluso triásicos, jurásicos y cretácicos. Siendo escasos los sondeos que lo han cortado, por lo que se supone que debe corresponder a materiales limo-arcillosos del Messiniense y del Plioceno-Pleistoceno, pues las formaciones de Pinos Genil y Alhambra parecen más restringidas geográficamente. Sin embargo, la geometría de este substrato se conoce relativamente bien gracias a los estudios de prospección geofísica eléctrica realizados por la FAO en los años 60. Una vez interpretados los datos geofísicos, FAO e IGME (1972, [1]) confeccionan un mapa de isorresistividades verdaderas y otro de la topografía del sustrato. Este último mapa, se ha elaborado de nuevo bajo un enfoque geoestadístico a partir de los datos de los sondeos verticales de ese proyecto cuyo resultado se muestra en la figura adjunta (figura 2.6). La diferencia de este mapa frente a los mapas obtenidos en el proyecto de la FAO no son significativas, aunque cabe destacar ciertas diferencias locales en la zona de Maracena, donde se localiza un máximo y en la zona al sur de Granada, donde se localiza un mínimo relativo, no siendo el resultado en este último caso estudiado lineal a lo largo del río, sino con un mínimo en la zona sur de Chauchina y un aumento prácticamente lineal hacia el este y el oeste.

Esta geometría puede verse claramente en la figura 2.7, en la que se muestran varios perfiles longitudinales y transversales del acuífero. En la misma figura se representa

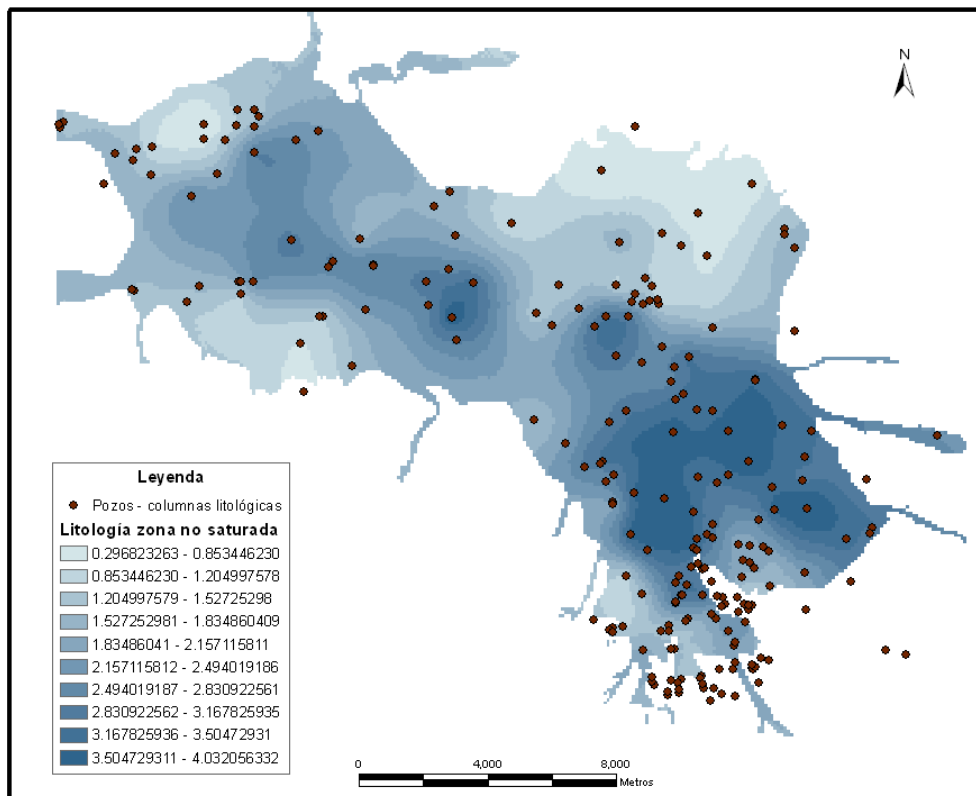


Figura 2.5: Mapa de coropletas de la litología de la zona no saturada del acuífero de la Vega de Granada (estimación mediante krigeaje)

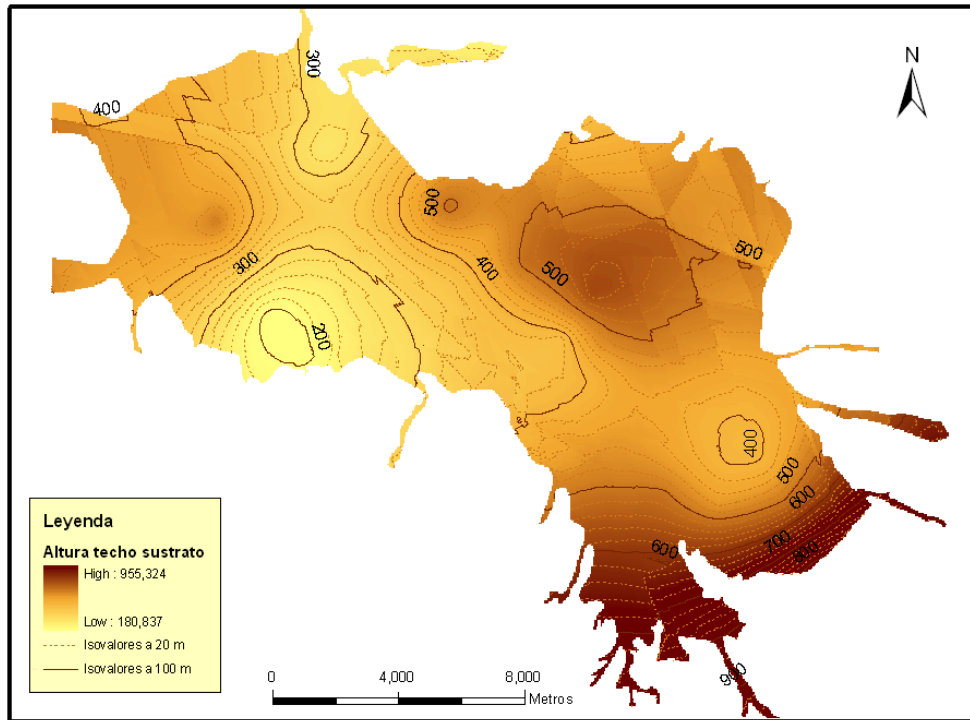


Figura 2.6: Mapa de cota del sustrato del acuífero de la Vega de Granada.

también la piezometría, que será descrita en detalle en la siguiente sección. En esta figura se puede observar como a lo largo del perfil 6 el nivel piezométrico disminuye hacia el oeste, lo cual concuerda con la variación del gradiente, que se verá más adelante (2.2.2). El espesor no saturado también disminuye en el sentido indicado anteriormente. Además, se puede observar que existe una estructura de cuenca con un levantamiento a la altura de Chauchina donde sube el sustrato y disminuye el espesor saturado. A pesar de mantenerse los rasgos principales mencionados en la tesis de Castillo Martín (1986, [30]), que explica que existe un mayor espesor de relleno aluvial en la parte central del perfil 3 y disminuye longitudinalmente a lo largo del perfil 6, como se observa en los perfiles 1 y 2, estas estructuras se muestran más complejas a la luz de los nuevos tratamientos geoestadísticos realizados.

2.2.2. Funcionamiento hidrogeológico y piezometría

En este apartado se resume el funcionamiento hidrogeológico del acuífero de la Vega de Granada, con objeto de que aquellas interpretaciones realizadas a la luz de este trabajo de investigación, tengan el contraste suficiente con el modelo de comportamiento hidrogeológico de los materiales representados en el acuífero.

De especial utilidad han resultado los trabajos realizados por Castillo Martín (1989, [30]) y Luque Espinar (2001, [110]), los cuales recogen la mayor parte de la información

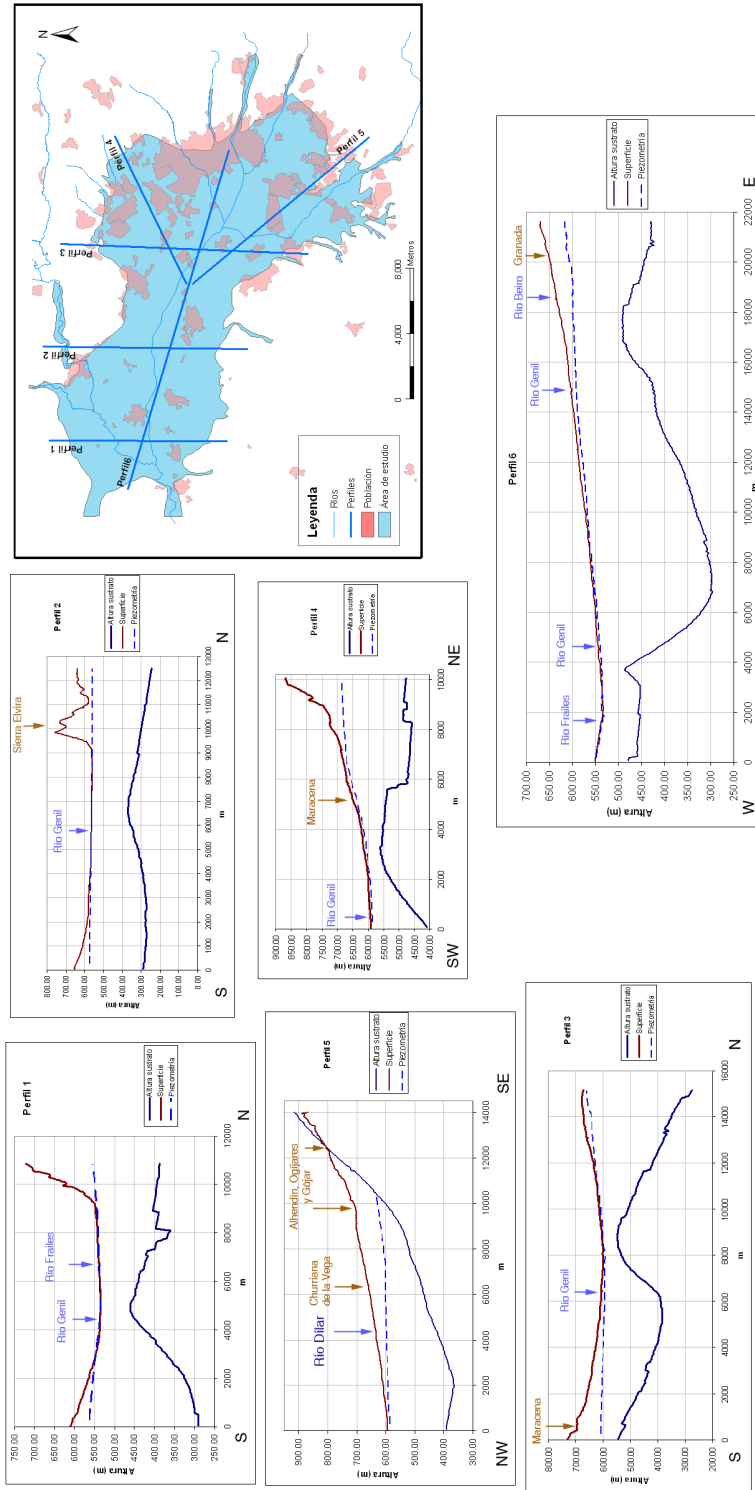


Figura 2.7: Cortes transversales y longitudinales del acuífero de la Vega de Granada.

e investigación hidrogeológica de la Vega de Granada realizada en los últimos años y a los que se puede acudir para un mayor detalle. Asimismo, se presentan algunos datos y mapas elaborados, bajo la nueva perspectiva de los métodos de extracción de conocimiento presentados en este trabajo.

Como ya se ha comentado antes, el acuífero de la Vega de Granada se localiza en la parte alta de la Cuenca del río Genil, el cual constituye el eje de drenaje de esta región y recorre la superficie del acuífero de este a oeste. La superficie permeable del acuífero, que presenta forma alargada según la dirección del río Genil, supone unos 200 km^2 . Aproximadamente el espesor de este relleno permeable puede superar los 300 m en su parte central y es progresivamente menor hacia los bordes. Los materiales de relleno son fundamentalmente gravas y arenas en la parte central, aumentando las intercalaciones de arcillas y limos hacia los bordes y disminuyendo el contenido de gravas. Los límites laterales del acuífero están constituidos por materiales más antiguos, pertenecientes a la depresión de Granada, e incluso a su substrato subbético. De todos ellos dominan los limos y arcillas del Messiniense y Plioceno-Pleistoceno, respectivamente, que constituyen además el muro generalizado del mismo.

La alimentación y descarga de este acuífero se puede considerar a grandes rasgos dividida en dos sectores. Un sector oriental que corresponde al área de alimentación y un sector occidental correspondiente al área de descarga.

La alimentación del acuífero procede mayoritariamente de las aguas de escorrentía y regadío que acceden a esta vega desde su amplia cuenca vertiente, a la entrada de los principales ríos como a la densa red de acequias (figura 2.8) sin revestir y al uso de riego por encharcamiento, que hoy día sigue utilizándose en muchas parcelas y hazas. Así, la mayor parte de la alimentación del acuífero se produce en el sector de cabecera con la entrada de los principales ríos que vienen de Sierra Nevada y de sus áreas de influencia de regadío. El mayor aporte de los cursos superficiales coincide por lo tanto con los meses de deshielo (abril a junio). En la mitad occidental, la alimentación por este concepto está restringida a los ríos Cubillas y Velillos y al Canal de Cacán y a las escorrentías del río Genil y sus derivaciones, única época en la cual algunos años el río Genil lleva agua más allá del puente de los Vados.

En orden de importancia, le sigue la infiltración de la lluvia útil sobre los 200 km^2 de superficie permeable que presenta el acuífero. La media de precipitación calculada entre los años 1975 y 2003 es de 426,5 mm y su distribución temporal se puede observar en la figura 2.9, aunque estos datos serán tratados más profundamente en la sección de climatología (Sección 2.3).

Para otros autores, (ITGE, 1989 [100]), las aportaciones ocultas (escorrentías subterráneas) revisten mayor importancia que la lluvia útil. Estas entradas ocultas provienen

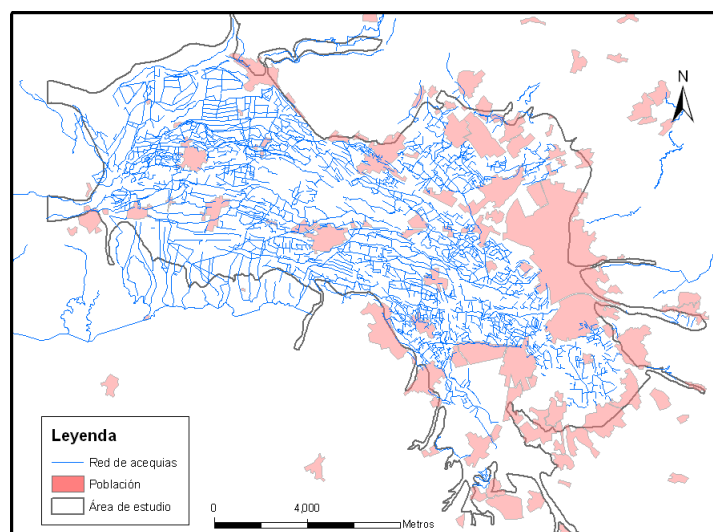


Figura 2.8: Distribución de la red de acequias en el área de estudio.

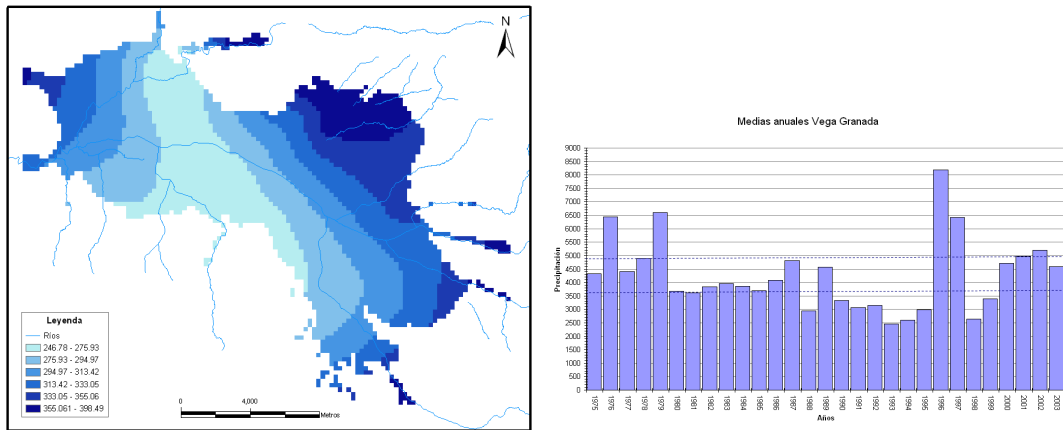
de acuíferos limítrofes, es decir, Sierra Elvira, Formación Alhambra y Formación Zubia, fundamentalmente.

La descarga del sistema se produce fundamentalmente a través del río Genil a partir del puente de los Vados, en la que el río deja de ser perdedor y comienza a ser ganador, como principal eje de descarga del acuífero, especialmente entre Chauchina y Valderrubio. En esta zona también cobra especial importancia el papel del río Cubillas, como colector de los recursos que se drenan. Así, el triángulo formado por los pueblos Chauchina, Valderrubio y Láchar es la principal área de drenaje situada entre las cotas de 540 y 525 m.

La descarga a los ríos Genil y Cubillas, que se intuye en la época de aguas altas, es visible en la época de estiaje, donde los aportes de las aguas subterráneas a los mismos, es mayoritario.

Además de la descarga producida a partir de los ríos, existen numerosas salidas a través de manantiales de carácter pseudoantrópico, que desde tiempo inmemorial fueron captados mediante zanjias para su utilización en regadío, formando áreas casi permanentemente inundadas. Una imagen de estas captaciones para el riego se puede ver en la figura adjunta 2.10. Estos manantiales drenaban un caudal medio de 1500 l/s (ITGE, 1989 [100]), sin embargo, en la actualidad este caudal se ha reducido hasta el punto que en algunos momentos se encuentran prácticamente secos (CHG e ITGE, 1993 [5]).

Otro aspecto importante a tener en cuenta en la descarga del acuífero es el más de un millar de captaciones que lo explotan, cuya finalidad mayoritaria es el regadío, con un aumento importante de su uso (por el aumento de la población) para el consumo



(a) Distribución espacial de la precipitación para un año medio (1985). (b) Evolución de la precipitación (decimas de mm) en los años 1975-2003 en la Vega de Granada.

Figura 2.9: Distribución de la precipitación en el área de estudio.



Figura 2.10: Manantiales Fuente Fontana y Martinete.

humano y para usos recreativos con el llenado de piscinas. Estas salidas se producen principalmente entre los meses de mayo a septiembre.

Otra consideración que se debe de tener en cuenta en la descarga, señalada por Castillo Martín (1986, [30]) es la posible evapotranspiración a partir de aquellas áreas con nivel piezométrico muy próximo a la superficie, las cuales suponen una elevada extensión en el caso de la Vega de Granada como se podrá ver más adelante en el mapa de espesor no saturado (figura 2.12), además de la evapotranspiración por cultivos como las choperas, con gran poder transpirante y localizadas generalmente en zonas donde el nivel freático es más alto.

Uno de los parámetros hidrodinámicos más importantes en el estudio del funcionamiento hidrogeológico del acuífero es la piezometría. El análisis de esta variable se realiza con todo detalle en el trabajo de Luque-Espinar (2001, [110]), en él se puede encontrar el estudio tanto de su variabilidad espacial como temporal. Como ejemplo, se

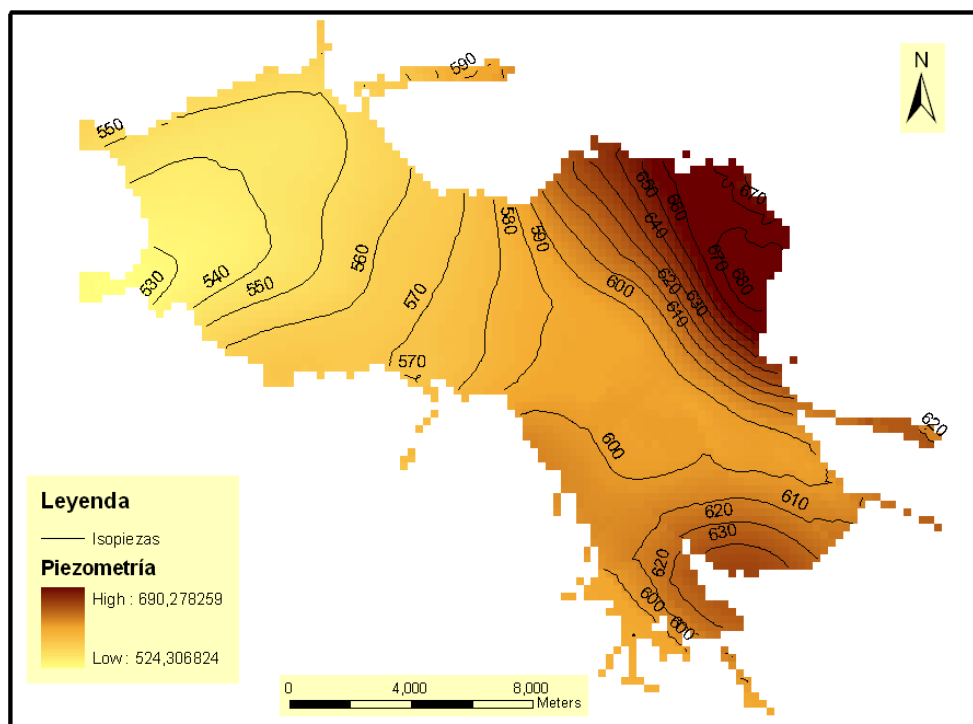


Figura 2.11: Mapa de coropletas de la piezometría e isopiezas de Abril de 1980.

muestra el mapa de isopiezas que he elaborado para un año medio (figura 2.11), a partir de un total de 180 medidas distribuidas homogéneamente en el acuífero y posteriormente espacializadas mediante geoestadística.

La interpretación del mapa de isopiezas indica un gradiente de flujo en sentido este a oeste, desde la zona de alimentación a la zona de descarga. La disposición perpendicular de las isopiezas en buena parte de los bordes del acuífero es indicativo de ausencia de alimentación en estas zonas y existencia de alimentación en los casos en los que se muestran más o menos paralelas. La disposición convexa de las isolíneas hacia el sentido del flujo bajo los cauces actuales muestra la existencia de un flujo hacia el río y el mayor espaciado una mayor transmisividad en esos ejes del río.

Desde el punto de vista temporal, el análisis espectral de las series piezométricas ha mostrado ciclos anuales con niveles piezométricos bajos entre agosto y noviembre y altos entre marzo y mayo. Además de estos ciclos anuales, se constata la existencia de otros ciclos de tipo interanual los cuales se relacionan con el clima de la región. En concreto, fenómenos como la NAO (North Atlantic Oscillation) se reflejan en el funcionamiento del acuífero (Luque Espinar, J. A. et al, 2007 [111]).

El estudio conjunto de la piezometría, del modelo digital del terreno y de la altura del sustrato ha permitido realizar una estimación del espesor saturado y no saturado del acuífero. El espesor saturado es un dato de gran interés hidrogeológico puesto que

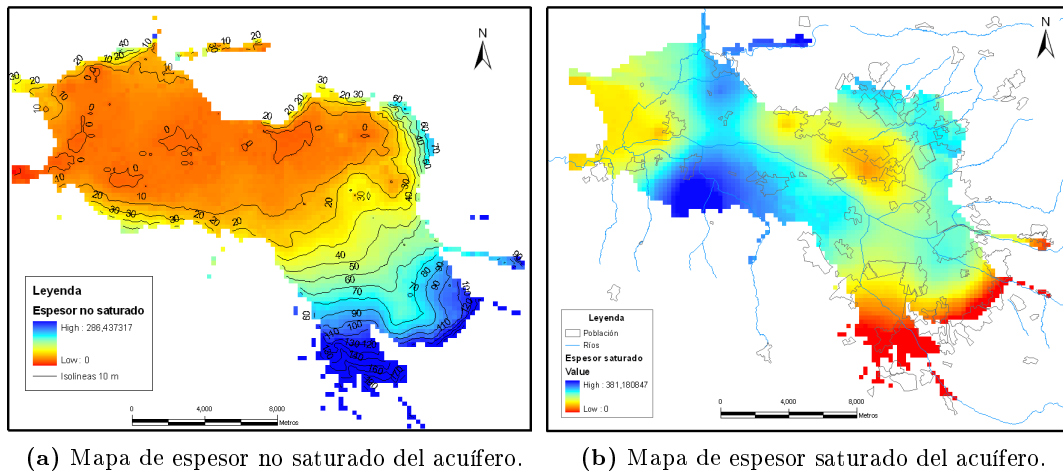


Figura 2.12: Mapa de espesor no saturado y espesor saturado del acuífero.

permite cuantificar el volumen de reservas del acuífero, conocido igualmente el coeficiente de almacenamiento. La dificultad de cálculo del coeficiente de almacenamiento y los escasos datos que existen para el área de estudio no han permitido realizar su estimación, aunque la aproximación realizada por Castillo Martín (1986, [30]) puede servir como referencia para su cálculo. Según esta aproximación, el volumen de reservas del acuífero sería de unos 140 hm^3 según los valores de infiltración.

El espesor no saturado es una variable clave para evaluar la vulnerabilidad de un acuífero, el cual es un aspecto tratado en capítulos posteriores. Esta variable se deduce empíricamente de la sustracción de la piezometría al modelo digital del terreno (MDT) (figura 2.12). En el acuífero, los valores más bajos de espesor no saturado están en la parte occidental y central del acuífero, zona que también coincide con la zona de drenaje. Los mínimos de espesor no saturado se localizan al sur de Atarfe, también al este y al sur de Fuente Vaqueros, y en un sector cercano a Láchar. Estas zonas con nivel freático tan cercano a la superficie son de gran interés en el estudio de la descarga, ya que tal y como planteó Castillo Martín (1986, [30]) es posible que exista una evapotranspiración directa del acuífero, dada la relativa proximidad de las aguas subterráneas a la superficie.

A partir del estudio de la piezometría es posible obtener una representación espacial del gradiente hidráulico por el método geoestadístico propuesto por Pardo-Iguzquiza y Chica-Olmo (2004 [140]). En este trabajo se ha realizado una aplicación de este método para calcular el módulo y la dirección del gradiente a partir de los datos experimentales disponibles. El mapa mostrado en este capítulo corresponde a la estimación geoestadística del gradiente correspondiente a un año medio, a partir de 182 medidas de piezometría, lo que nos da una visión novedosa de la distribución espacial del gradiente en el área de estudio (figura 2.13).

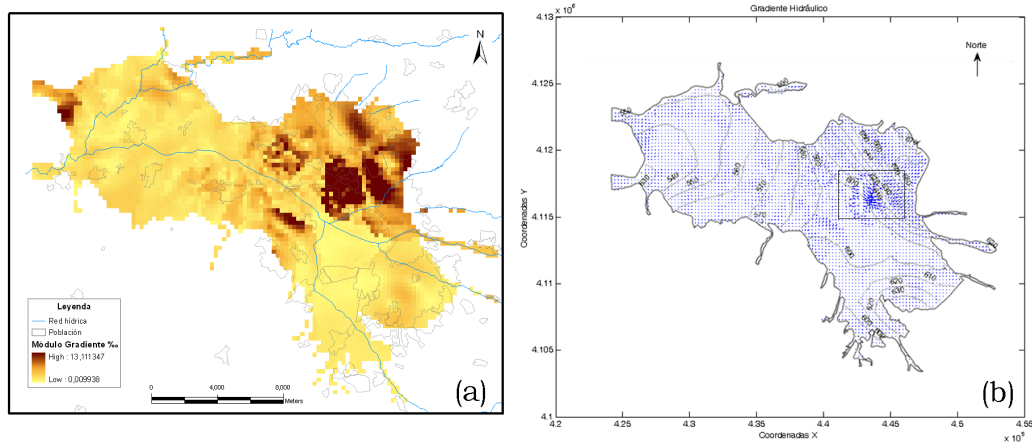


Figura 2.13: Estimación geoestadística del gradiente para abril de 1980. (a) Estimación del módulo del gradiente. (b) Estimación de la dirección del gradiente máximo.

Según Luque-Espinar (2001, [110]), los mayores valores del gradiente hidráulico se observan en el sector noreste del acuífero, con un valor medio del 15 ‰, con valores del 5 ‰ en la zona central. Los datos obtenidos con los cálculos realizados han resultado en valores ligeramente distintos a los mencionados, así el máximo se localiza en la zona de Maracena con un valor de 13‰, algo más bajo que el estimado por Luque-Espinar y el valor mínimo obtenido para la estimación es de 0,01‰. Estos valores se localizan a la entrada del río Monachil en la zona de cabecera del acuífero y en gran parte del acuífero. Otra diferencia notable es que los valores superiores a 2‰ sólo se encuentran en la franja norte-sur comprendida entre los pueblos de Gabia y de Albolote.

De esta manera, puede decirse que la mayor parte del acuífero presenta unos valores promedio de gradiente entre 0,01‰ y 2‰. Dentro de estos rangos de valores, se distinguen dos máximos bien diferenciados, el primero se sitúa en la zona noreste del acuífero al norte de la confluencia entre el río Beiro y Genil y el segundo máximo en el área superior oeste. El primer máximo se relaciona con la zona de baja permeabilidad de los materiales que existen en esa área (figura 2.15 a), que se observa igualmente en el mapa de isopiezas (figura 2.13 b) donde las líneas aparecen mucho más apretadas. Un poco más al sur se encuentra un máximo local correspondiente, probablemente, a una explotación importante que produce un cono de depresión, un detalle de la dirección del gradiente en esta zona se representa en la figura 2.14. El segundo máximo corresponde a la entrada de aguas del noroeste del área de estudio, ligado a zonas de baja permeabilidad de los materiales del acuífero.

La zona de gradientes mínimos aparecen al sureste, desde la entrada de los ríos de cabecera hasta aproximadamente el cruce de los ríos Genil, Beiro y Dílar. Continúa en

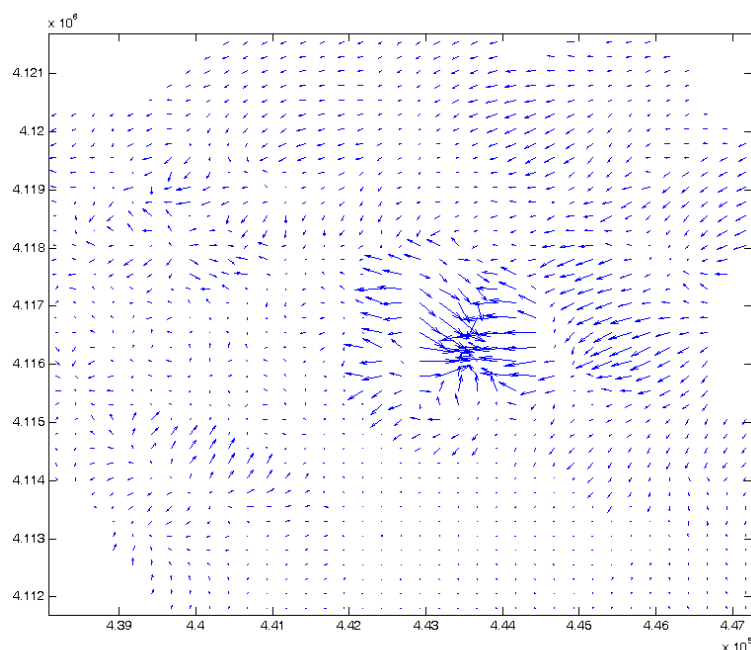
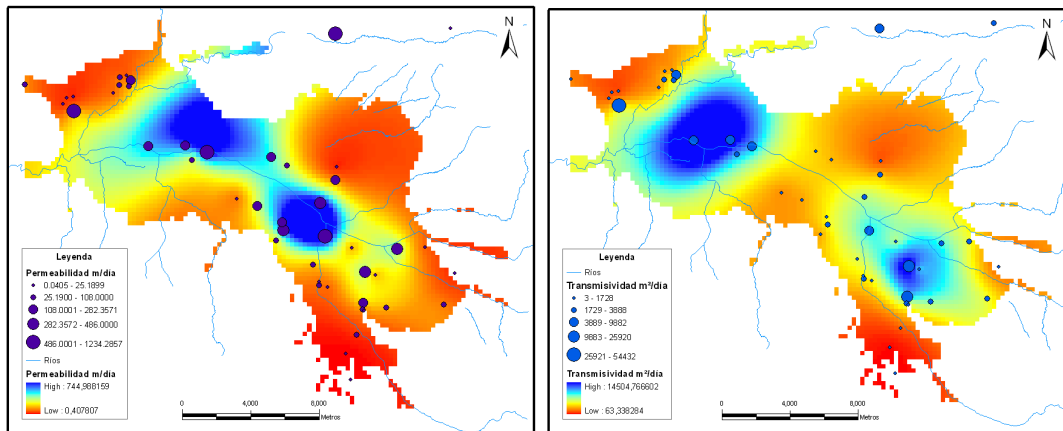


Figura 2.14: Detalle de la dirección del gradiente en un sector localizado al noroeste del acuífero. (Ver recuadro figura 2.13).

una franja estrecha a lo largo del río Genil hasta Santa Fe, donde de nuevo forma un área más amplia coincidiendo con la zona de máxima permeabilidad.

El estudio de la transmisividad hidráulica en el acuífero de la Vega de Granada se remonta al proyecto FAO (1972, [1]), según el cual los valores de transmisividad a lo largo del río Genil varían entre 900 y 9000 $m^2/día$, con máximos de 40.000 $m^2/día$. Para el estudio de la transmisividad, en este trabajo, se ha contado con un total de 46 datos cedidos por el IGME, obtenidos a partir de diferentes ensayos de bombeo. Algunos de estos datos requirieron una corrección espacial según Martín Asín (1990,[117]) debido a que su sistema de coordenadas estaba en el antiguo datum Madrid. La estimación geostatística la transmisividad para el sector estudiado se puede ver en la figura adjunta 2.15, junto con la estimación de la permeabilidad. Para estas estimaciones, se ha usado la corrección propuesta por Journel y Huijbregts (1978, [103]), que establece una corrección en el caso de distribuciones no normales, mediante una transformación logarítmica. Esta corrección también fue usada por Chica-Olmo (1981, [2]) para el estudio de la transmisividad del acuífero en esta área.

Los valores de transmisividad según los datos estimados varían entre 60 y 15.000 $m^2/día$, los valores máximos se localizan en dos áreas distintas, una al sur de Granada capital y otra a la altura de Fuente Vaqueros. Los valores máximos de la permeabilidad



(a) Mapa de permeabilidad del acuífero.

(b) Mapa de transmisividad del acuífero.

Figura 2.15: Estimación de la conductividad hidráulica y de la transmisividad.

sin embargo están localizados a lo largo y norte del río Genil y alcanzan valores de cerca de los 800 m/día (figura 2.15 b). La diferencia en la situación de los máximos de permeabilidad con los de transmisividad es consecuencia de la variación en el espesor saturado fundamentalmente en la zona central del acuífero (figura 2.7). Estas variaciones también podrían deberse a niveles colgados en la zona de Armilla que enmascaren las medidas de transmisividad y permeabilidad.

2.2.3. Aspectos hidrogeoquímicos

El estudio de la calidad de las aguas subterráneas del acuífero es uno de los objetivos principales de este trabajo. Para abordar este objetivo se requiere un conocimiento experimental detallado de las variables hidroquímicas, así como de otras variables complementarias relacionadas con la contaminación. De acuerdo con esta premisa, se han diseñado distintas campañas de muestreo orientadas al estudio de una parte importante de las características relacionadas con la calidad del agua subterránea, particularmente con sus propiedades físicas, químicas y microbiológicas.

Las diversas campañas se han planificado de forma exhaustiva, contemplando la cantidad de muestras que se pretendía recoger, teniendo en cuenta los aspectos básicos de organización, así como la elección de las técnicas y métodos para la optimización del muestreo. Uno de los principales problemas encontrados ha sido disponer de una base de datos de pozos georreferenciada y actualizada, que permitiera seleccionar el conjunto de puntos a muestrear. Esta labor se realizó con la ayuda del GPS-SIG (Trimble Pathfinder, figura 2.16), lo que supuso una ayuda inestimable a la hora de realizar la rejilla de muestreo. Un elevado detalle de las características de cada punto acuífero (pozo), era necesario puesto que se preveía una segunda y tercera campaña de toma de muestras,



Figura 2.16: Fotografía del muestreo de aguas superficiales realizado en el río Genil con GPS Trimble Pathfinder.

lo que fue conseguido exitosamente gracias a la meticulosidad del tratamiento de la información y al equipamiento informático.

La labor de campo comenzó en septiembre del 2003. Para el muestreo se disponía de un GPS Trimble Pathfinder con un colector TSCE, un conductímetro, pH-metro, sonda piezométrica y un termómetro de mercurio. La creación de la base de datos de campo, se implementó adecuadamente para almacenar la información de medidas y adicional en cada uno de los pozos de muestreo. Esta información adicional a los datos analíticos corresponde tanto a la posición espacial exacta, fotografías, horarios, persona de contacto, etc, así como a datos tomados in situ durante la campaña de fecha, hora, pH, conductividad, temperatura y piezometría (en el caso de ser posible su toma). Este muestreo es el mas detallado y en él se recogieron 209 muestras de agua subterráneas y 43 muestras de aguas superficiales. Las muestras fueron analizadas en el Instituto del Agua de la Universidad de Granada por los métodos de determinación mostrados en la tabla 2.2.

La densidad del muestreo realizada en esta campaña corresponde a un punto por km^2 . La distribución espacial de las muestras según la rejilla de muestreo se observa en la figura 2.17.

El segundo muestreo se realizó en octubre del 2003, en el cual, se eligieron un total de 51 puntos para realizar los análisis microbiológicos, una vez elaborada la base de

datos, de los cuales se tomaron 41 puntos de aguas subterráneas, 6 puntos de aguas superficiales y 4 de aguas residuales urbanas. Las variables analizadas son: bacterias aerobias, coliformes totales, coliformes fecales, estreptococos fecales, demanda bioquímica de oxígeno (DQO) y sólidos en suspensión (SS). Estas muestras fueron igualmente analizadas por el Instituto del Agua de la Universidad de Granada cuyos métodos se resumen en la tabla 2.2.

El tercer muestreo se organizó en marzo del 2004 para estudiar los metales pesados, en el cual se seleccionaron un total de 76 puntos, de los cuales 70 puntos son de agua subterránea y 7 de aguas superficiales (figura 2.17). Se analizaron ocho metales, correspondientes a plomo, cromo, cobre, hierro, manganeso, cadmio, arsénico y antimonio. Los seis primeros fueron analizados en el Centro de Instrumentación Científica de la Universidad de Granada por espectroscopia de radiación atómica, y los dos segundos fueron analizados en los laboratorios de la empresa Cavendish por el mismo método.

Además de estos datos, hay que contar con la información suministrada por el Instituto Geológico y Minero de España, correspondiente a varias campañas realizadas en el mismo año, con un total de 50 puntos donde se analizaron los parámetros fisicoquímicos mayoritarios en intervalos de cuatro meses.

Por último, cabe señalar que también se ha realizado un muestreo orientado al estudio de plaguicidas en las aguas subterráneas de la Vega en un conjunto de 70 puntos. Estas muestras se están analizando en el Centro de Instrumentación Científica de la Universidad de Granada, pero lamentablemente debido a ciertos problemas ocurridos ajenos a nuestra labor, no se han podido incluir en esta investigación.

Los mapas que a continuación se muestran, que representan la distribución espacial de las variables hidrogeoquímicas, han sido obtenidos mediante estimación geoestadística. Estos mapas, resultado del análisis geoestadístico de las variables experimentales, se presentan aquí para ilustrar mejor la explicación de los aspectos hidroquímicos y completar la visión general relativa a la calidad hidroquímica de las aguas del acuífero. El detalle de la metodología desarrollada para la elaboración de los mismos, sin embargo, se comenta en los capítulos 3 y 4.

Tabla 2.2: Métodos analíticos utilizados para el análisis de las muestras de aguas subterráneas.

Variable	Unidad	Método de análisis
<i>Conductividad del agua</i>	microS/cm	Conductímetro Hanna (HI-7071)
<i>pH</i>		Peachímetro Hanna (HI-9025)

(continúa en la siguiente página)

(continuación de la página anterior)

Variable	Unidad	Método de análisis
<i>Temperatura del agua “in situ”</i>	°C	Termómetro de mercurio de 0 a 50°C, con precisión de 0,1 °C
<i>Turbidez</i>	UNF	Turbidímetro Hach (2100 P)
<i>Sólidos en suspensión</i>	mg/l	Filtración pesando partículas > 0,45 µm
<i>Cloruros</i> <i>Nitratos</i> <i>Sulfatos</i> <i>Fosfatos</i>	mg/l	Cromatógrafo iónico Dionex DX 300
<i>Calcio</i> <i>Sodio</i> <i>Magnesio</i> <i>Potasio</i>	mg/l	Cromatógrafo iónico Dionex DX 300 con columna de separación de cationes Ionpac CS 12A
<i>Bicarbonato</i> <i>Carbonato</i>	mg/l	Valoración con ácido clorhídrico
<i>Nitritos</i>	mg/l	Espectrofotometría
<i>Boro</i>	mg/l	Determinación con azometina H, por espectrofotometría a $\lambda = 410$ nm
<i>Amonio</i>	mg/l	Espectrofotometría a ($\lambda = 640$ nm), con kit spectroquant Merk
<i>Fluoruros</i>	mg/l	Fluorímetro HANNA HI 93729
<i>Cromo</i> <i>Plomo</i> <i>Cobre</i> <i>Hierro</i> <i>Manganeso</i> <i>Antimonio</i> <i>Arsénico</i>	ppb	Espectroscopía de radiación atómica (con un PERKIN ELMER mod.5100, con Cámara de grafito PERKIN ELMER mod. 5100 ZL ZEEMAN.
<i>Bacterias aerobias</i>	Ud. En 1 ml	Filtración por membrana (0.45 µ) y cultivo en agar nutritivo TSA (Soja triptona Agar). Incubación 24 horas. Temperatura: 28°C
<i>Coliformes totales</i>	UFC en 100 ml	Filtración por membrana (0.45 µ) y cultivo en medio específico (m-ENDO Agar). Incubación 24 horas. Temperatura: 37 °C

(continúa en la siguiente página)

(continuación de la página anterior)

Variable	Unidad	Método de análisis
<i>Coliformes fecales</i>	UFC en 100 ml	Filtración por membrana (0.45 μ) y cultivo en medio específico (m-ENDO Agar). Incubación 24 horas. Temperatura: 44.5 °C
<i>Streptococos fecales</i>	UFC en 100 ml	Filtración por membrana (0.45 μ) y cultivo en medio específico (Agar-KF para Streptococos fecales). Incubación 48 horas. Temperatura: 37 °C
<i>DQO</i>	ppm	Oxidación con dicromato potásico en medio ácido
<i>Clostridios sulfitorreductores</i>	UFC en 50 ml	Medio Agar sulfito de hierro. Calentando hasta 80°C durante 5 min, para eliminar formas vegetativas.

2.2.3.1. Variables físicas

El primer grupo de variables, son las variables físicas que incluyen: la conductividad, la temperatura, la turbidez y los sólidos en suspensión. Este grupo de variables originalmente respondía a una serie de características que podían ser determinadas o apreciadas fácilmente in situ, sin necesidad de instrumentación específica de laboratorio. Hoy día, la mayoría de estas variables, excepto la temperatura, se determina en laboratorio de forma sistemática, ofreciendo cuantificaciones estandarizables. Sin embargo, esta clasificación sigue estando vigente y la tomaremos aquí haciendo referencia simplemente a su facilidad de medida en campo y añadiendo además el pH.

La materia presente en el agua se puede encontrar en tres estadios: disuelta, en suspensión (o particulada) y/o en forma coloidal. El primer estadio corresponde a las sustancias dispersas de forma homogénea en el líquido, de manera que hay una sola fase. El segundo implica la existencia de dos fases: la correspondiente al líquido (agua) y la correspondiente a las partículas sólidas. Finalmente, la tercera se encuentra entre las dos, de forma que las partículas son demasiado pequeñas para ser eliminadas por procedimientos normales de sedimentación o filtración. Este tercer estadio se caracteriza por causar el efecto tyndall, que se basa en la reflexión de la luz por parte de las partículas presentes en la disolución (Poch, 2003 [145]).

La **turbidez** es un parámetro utilizado habitualmente en aguas naturales como indicador de la presencia de sólidos, especialmente coloidales. En este caso, lo que se mide es la extensión con la que un rayo de luz es reflejado a su paso por el agua, con un ángulo de 90°. La extensión se encuentra relacionada con la cantidad de materia en suspensión. Esta relación no es estrictamente lineal, ya que el proceso de dispersión

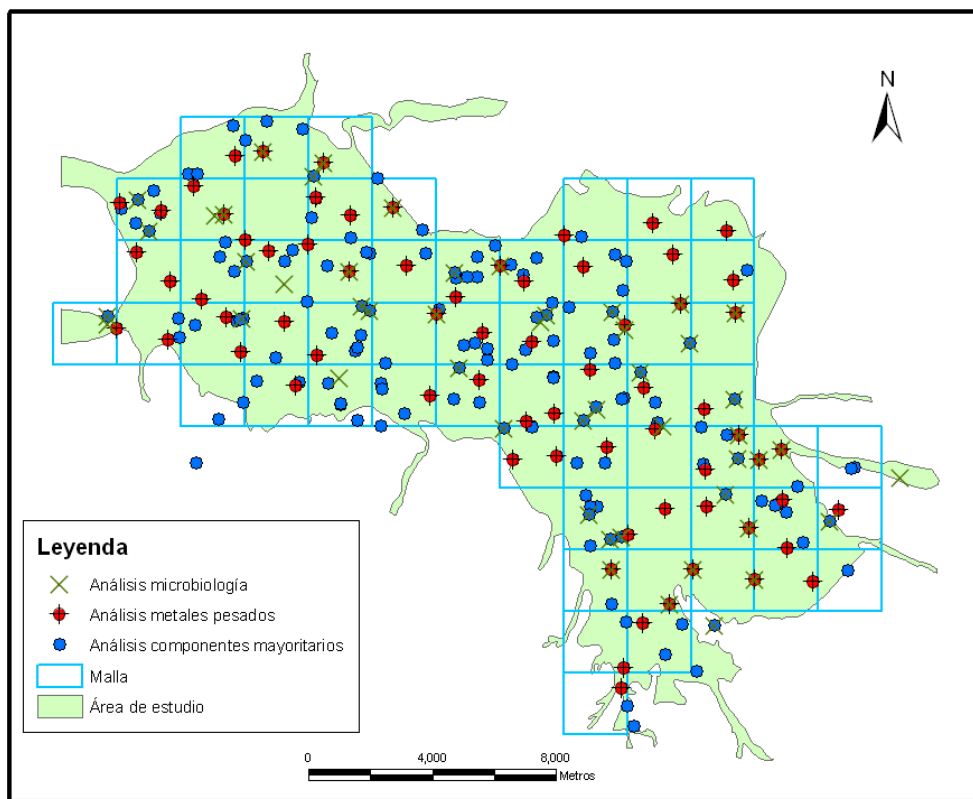


Figura 2.17: Distribución espacial de datos experimentales de aguas subterráneas.

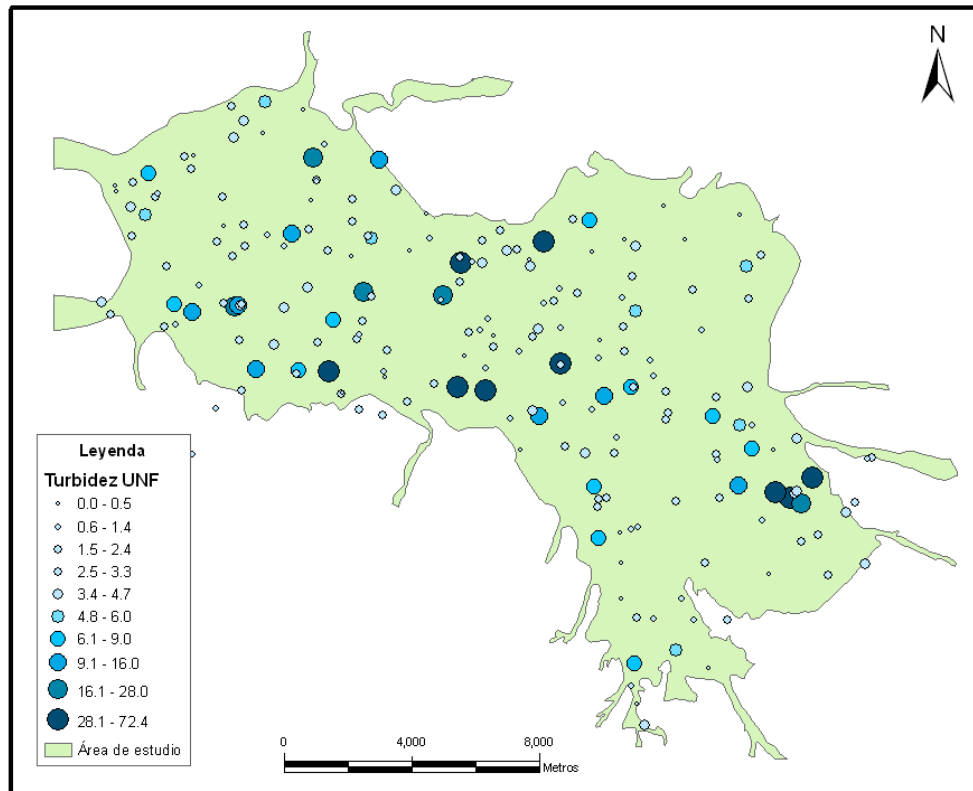


Figura 2.18: Distribución espacial de la turbidez en el área de estudio.

de la luz se encuentra influenciado tanto por el tamaño como por las características superficiales de la materia presente.

La fuente principal de turbidez en aguas naturales procede de la erosión y transporte de materia coloidal (arcilla, fragmentos de roca, sustancias del lecho, etc.) por parte de los ríos en su recorrido. Y otra parte de fibras vegetales, microorganismos o aguas residuales (domésticas e industriales). En las aguas subterráneas sin embargo, la turbidez suele ser baja, ya que el propio suelo y la zona vadosa actúan como filtro natural, siendo la fuente principal de turbidez los parámetros relacionados con el propio pozo, tales como el material acuífero o el correcto uso del mismo, en muchos casos permitiendo la entrada directa de materia superficial.

Debido a la influencia local de estos factores del pozo, parece que esta variable no muestra una estructura espacial de los datos, por lo que su representación en el mapa se ha realizado para cada punto (figura 2.18).

Por otra parte, la medida de la turbidez puede no ser muy fiable cuando aumenta la cantidad de sólidos en suspensión. En este caso es mejor recurrir a una medida directa como es la determinación de los sólidos en suspensión de la muestra. Esta determinación

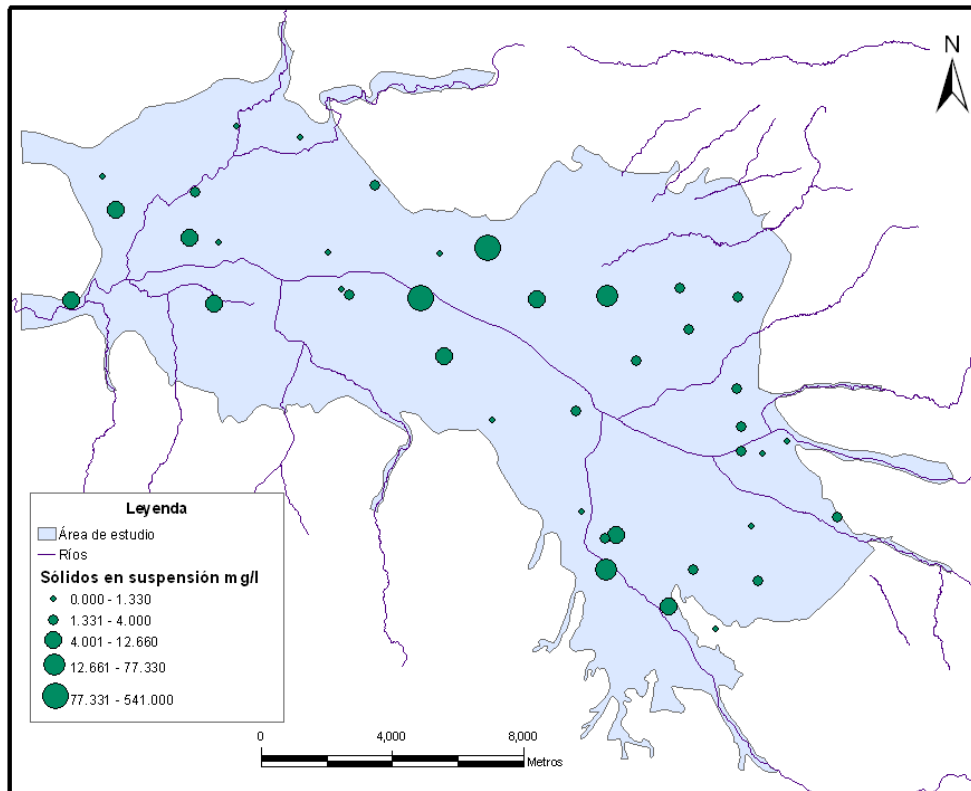


Figura 2.19: Distribución espacial de sólidos en suspensión en el área de estudio

se realiza filtrando primero el agua y pesando después los sólidos que han quedado retenidos en el filtro, eliminando previamente el agua que pudieran retener las partículas.

La cantidad de **sólidos en suspensión** en el agua del acuífero, igual que ocurre con la turbidez, suele ser baja y en muchos casos también influenciada por factores locales como el tipo de pozo, la limpieza del mismo, etc. El bajo número de muestras para este caso (50 puntos), así como las anomalías locales comentadas anteriormente no han permitido identificar de forma clara una estructura de variabilidad espacial, por lo que se ha preferido representar los puntos experimentales de manera individual (figura 2.19). En esta figura se aprecia que el 90 % de las medidas de sólidos en suspensión son inferiores a 10 mg/l. El resto de medias (5 puntos) corresponden a pozos localizados en el suelo sin ningún tipo de barrera física que impida la entrada de suciedad a la boca del mismo, por lo que cabe pensar que se trata de datos erróneos.

La **conductividad** se define como la capacidad del agua para conducir la electricidad, debido a las sales que lleva disueltas. Al ser un parámetro relacionado con el conjunto de iones, no específico, sirve como indicador del conjunto de iones totales que lleva el agua en disolución. No obstante, al no tener todos los cationes y aniones la

misma movilidad, el valor de la conductividad dependerá también de la proporción de los diferentes tipos presentes.

Esta variable está condicionada en primer lugar por las características de los suelos y los materiales por donde circula el agua. En el caso de la Vega de Granada, se trata de suelos calcáreos, por lo que su valor será en general alto, con valores que varían entre 485 y 4200 $\mu S/cm$ medida a 25°C. Otra de las causas que puede influir en el valor de la conductividad es la presencia de vertidos residuales, que se caracterizan por su elevada salinidad o filtraciones a través del suelo de gran cantidad de determinados elementos de origen antrópico (iones) relacionados con la agricultura, como los nitratos, fosfatos, etc. De este modo, esta variable es un indicador de la calidad del agua.

El aumento la conductividad puede limitar su uso tanto para consumo humano como para riego, por lo que se deberá estudiar su idoneidad según el tipo de uso. El límite para establecer la potabilidad, según la normativa española (Real Decreto 140/2003) del agua en función de la conductividad es, 2500 $\mu S/cm$, valor por encima del cual su uso para abastecimiento no es recomendable.

Los valores máximos de esta variable en el área de estudio (figura 2.20) se localizan fundamentalmente en el borde norte, al sur de Sierra Elvira y en el borde sur, en la zona del aeropuerto. Los valores más bajos, y más adecuados para consumo humano se sitúan en la zona de recarga al sureste del área de estudio.

En el mapa indicado se puede observar una clara regionalización de la variable que muestra una identificación de la zona de recarga en el sureste del acuífero con aguas superficiales menos salinas que mantienen esta característica a lo largo del curso del río Genil. El aumento de la conductividad en el acuífero está influida probablemente por el aporte o mezcla de aguas más salinas provenientes de los bordes evaporíticos del acuífero así como a la filtración de las aguas de riego (en muchos casos residuales).

La **temperatura** del agua subterránea se determinó en la campaña de muestreo de 2003 en un total de 53 puntos (figura 2.21a). La distribución espacial de la temperatura, realizada mediante krigeaje, muestra una franja central coincidente con el flujo del río Genil de temperaturas más frías y que corresponde a las aguas de infiltración de los ríos procedentes de Sierra Nevada. En esta franja se distinguen tres áreas de menor temperatura, que según Castillo Martín (1986, [30]) se relacionan con un sistema de flujo local de las aguas de cabecera. Otra zona de mínimos se sitúa en el borde noroeste que probablemente esté ligado a flujos subterráneos más profundos. Esta hipótesis se explicará con los perfiles de conductividad y temperatura más adelante.

Las áreas con valores altos de temperatura al norte del acuífero están influidas por las aguas termales de Sierra Elvira, donde se llegan a alcanzar temperaturas de hasta 30°. Estas aguas termales tiene su origen en una circulación más profunda, que algunas

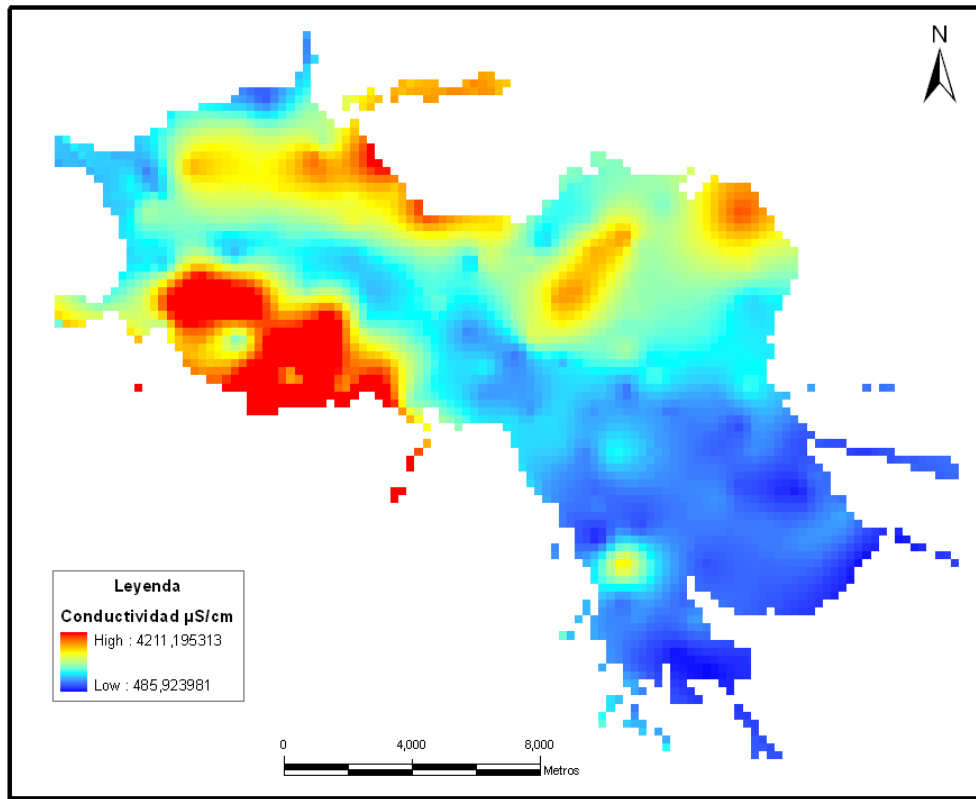
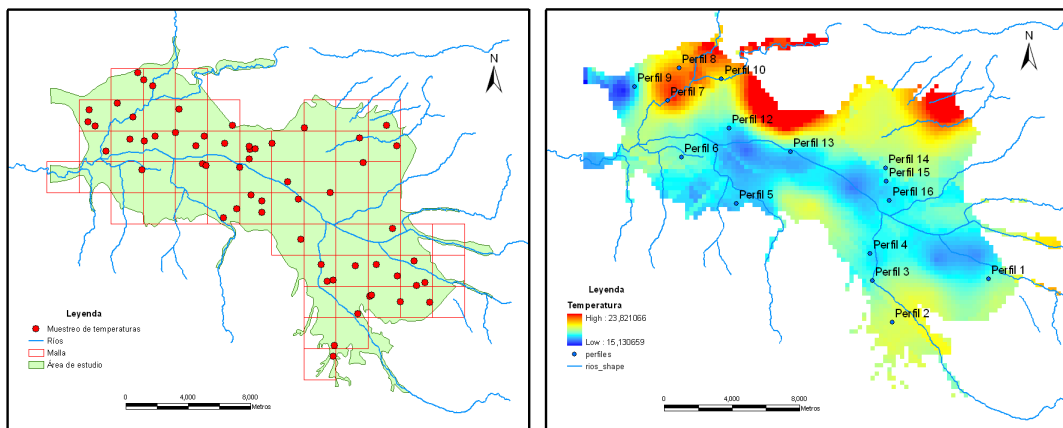


Figura 2.20: Distribución espacial de valores de conductividad del agua subterránea.



(a) Red de muestreo de la temperatura de las aguas subterráneas. (b) Mapa de la temperatura de las aguas subterráneas del acuífero y situación de los perfiles de T^a y conductividad.

Figura 2.21: Red de muestreo de la temperatura de las aguas subterráneas, distribución de la T^a y situación de los perfiles de T^a y conductividad.

hipótesis asocian a una alimentación procedente de alguna de las unidades subbéticas situadas más al norte.

Otras áreas con temperaturas altas pero más moderadas (entre 18 y 20^o C) localizada al norte de Maracena y a lo largo del río Cubillas pueden tener su origen en una recarga lateral.

En el mapa de distribución de temperaturas realizado mediante krigeaje se observa un mayor gradiente de temperaturas en la dirección norte-sur que se identifica en el variograma como una deriva marcada en esta dirección. Por otra parte, en la dirección este-oeste, la variabilidad es mínima y el gradiente de temperaturas es más bajo. Otro dato interesante procedente de la regionalización de la variable es el alcance, que no es excesivamente alto (4000 m), lo que indica que existen bastantes cambios a escala local, como se puede observar igualmente en la figura 2.21.

Otros datos recopilados y estudiados en este trabajo son los perfiles de conductividad eléctrica y temperatura del agua realizados en algunos de los piezómetros del acuífero. El estudio de los perfiles en profundidad de estas variables ayudan al conocimiento de los flujos subterráneos o las recargas producidas en el acuífero. No obstante, estos resultados se han tomado con cierta reserva, pues no se ha verificado en dichos puntos el grado real de renovación del agua en el interior de los piezómetros. Se realizaron 16 perfiles en distintas áreas del acuífero con una sonda de conductividad y temperatura en noviembre del 2003 (situación en la figura 2.21 b).

Las gráficas de la figura 2.22 representan los valores de temperatura y conductividad observados en función de la cota piezométrica. Se ha optado por una representación según una cota absoluta para mostrar con mayor claridad los posibles gradientes de unos puntos a otros.

Un aspecto interesante relativo a la temperatura del agua en los perfiles es que la mayoría de los pozos tienen una franja superior de agua más cálida. Esta franja de temperaturas más cálida es desde medio grado hasta casi dos grados superior al resto del perfil en profundidad y puede comprender entre 5 y 20 metros de espesor (figura 2.22a). Los pozos que presentan esta franja cálida son: 1, 3, 4, 6, 7, 8, 9, 11, 12, 13 y 16, es decir, en la mayoría de los pozos, excepto los de menor profundidad y los perfiles 2 y 5. Las aguas superficiales en noviembre están más frías que las subterráneas, por lo que creemos que el aumento de temperatura es debido probablemente al calentamiento de la tubería metálica del piezómetro por la acción del sol que transmite el calor a la parte superior del agua del piezómetro y al agua. Así, en los pozos menos profundos, no se observa este incremento porque la renovación de las aguas más profundas no se llega a constatar. Por este motivo, consideramos que los perfiles 14 y 15 no son representativos y su análisis se ha dejado hasta contrastar la información en otros perfiles cercanos en próximos estudios.

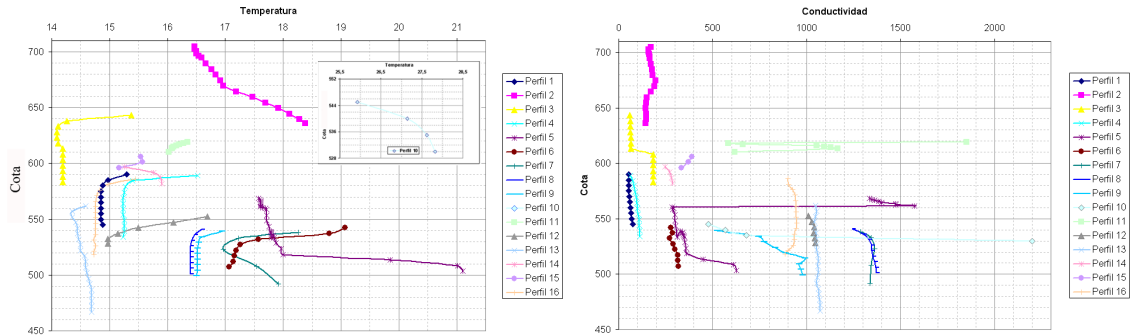


Figura 2.22: Perfiles de temperatura y conductividad del acuífero.

Los perfiles cercanos al cauce del Genil y Monachil (puntos 1, 3, 4, 13 y 16) presentan una distribución en profundidad similar con temperaturas más bajas en función de la proximidad al río, y menores aún en perfiles más lejanos a los cauces del río (16, 1 y 4). Los perfiles 6, 7, 8, 9, 10 y 12, situados en la zona de descarga del acuífero y cercanos a Sierra Elvira, presentan mayores temperaturas en general, lo que es coherente con lo mencionado para la distribución espacial de la temperatura. Los puntos 10, 8, 9 y 7 muestran una mayor temperatura influida por las aguas termales, sin embargo, el punto 6, que está mucho más alejado también tiene una elevada temperatura fundamentalmente en superficie, pudiendo estar influida, además de por lo comentado anteriormente, por una mayor cercanía de las aguas subterráneas a la superficie (aquí la piezometría es muy somera).

Un comportamiento contrario se da en los perfiles 2 y 5. El perfil 5 muestra un aumento de la temperatura de 18^o a 21^oC entre las cotas 520-500 m. Este aumento indica un aporte de aguas más cálidas en profundidad. Hay varias hipótesis para explicar este incremento. En primer lugar cabría atribuirlo a aportes procedentes de un sondeo surgente termal cercano y que no se cementó correctamente. Una segunda posibilidad es que la influencia proceda de retornos de riegos realizados con aguas del canal de Cacán, que circulan en superficie y son calentadas por el sol. Por último, también existe la posibilidad de que se produzcan estos aportes a través de niveles permeable, por lo que el origen de los mismos estaría alejado de esta zona.

Además de este aumento de temperaturas, también existe un aumento de conductividad en el perfil 5, lo que corrobora un origen distinto de las aguas. La conductividad aumenta en la misma franja que la temperatura (520-500m) de valores de 300 $\mu S/cm$ a 620 $\mu S/cm$. Este área de hipertermia no se observa en mapa de la distribución de temperatura, debido probablemente a que no existen datos cercanos a esa zona en esta campaña. Sin embargo, esta anomalía térmica sí es citada por Castillo Martín (1986, [30]), aludiendo a infiltraciones del arrollo Salado (que en principio cabría descartar

porque el aumento de T^a y conductividad es en profundidad) o alimentación lateral oculta de algún nivel de sales evaporíticas del Mioceno.

Por último, el perfil nº 2 que tiene una distribución similar al 5, presenta un aumento desde la superficie más moderado y un aumento más brusco a los 40 m de profundidad. Desde la superficie hasta los 40 m el aumento de la temperatura es de $1^{\circ}C$ (cota de 710 a 660 m), es decir, aumenta $0.025^{\circ}C/m$ y a partir de los 40 m (cota 660-630 m) la temperatura aumenta $0.05^{\circ}C/m$. La diferencia de temperaturas no se manifiesta en una diferencia en la conductividad por lo que en principio no se interpreta un distinto origen de las aguas. Así, el aumento en profundidad puede ser debido a una circulación más lenta del agua subterránea en la parte más baja del acuífero y un flujo más rápido en la parte más superficial con aguas más frías e influidas por la recarga del río Dílar.

El **pH** de las aguas naturales viene determinado por las interacciones entre las sustancias ácidas y básicas presentes en la misma. Su valor se sitúa en un rango comprendido entre 6 y 9, con tendencia a mantener un valor casi constante. Esta regulación del pH se debe al efecto tampón y ocurre por la existencia de ácidos o bases muy débilmente ionizados que tienden a mantener casi constante el pH del agua cuando se intenta cambiarlo añadiendo pequeñas cantidades de ácidos o bases fuertes. Estas reacciones son relativamente frecuentes en la naturaleza debido a la presencia en el agua de CO_2 disuelto y HCO_3^- . De esta manera se explica que la mayoría de las aguas subterráneas tengan un pH entre 6,5 y 8.

La concentración del ion hidrógeno en las aguas naturales tiene un papel relevante en las reacciones químicas asociadas con la formación, alteración y disolución de minerales. El pH de la disolución determinará, por tanto, el sentido de estos procesos. Existen diversos procesos que influyen en la variación del pH, pero uno de los más importantes en nuestro estudio son los procesos de nitrificación, que conllevan una disminución del valor del pH, mientras que la desnitrificación tiende a incrementarlo.

La ubicuidad de las rocas carbonatadas y las reacciones de equilibrio del dióxido de carbono hacen que los bicarbonatos estén presentes en la mayoría de las aguas naturales, aún más como es el caso del área de estudio, al tratarse de un acuífero detrítico carbonatado. También hay que tener en cuenta los fenómenos como la lluvia ácida que pueden modificar el pH del acuífero o la filtración de aguas residuales de procesos industriales, las cuales contienen normalmente diferentes ácidos y bases, fuertes y débiles, así como sus sales, que pueden afectar al equilibrio del pH del agua. Así por ejemplo, la disolución de ácido clorhídrico o de hidróxido de sodio, modifica fuertemente el aporte de hidrógenos e hidroxilos produciéndose un agua ácida o básica respectivamente y la disolución de muchas sustancias neutras, como la calcita que da una solución algo básica o alcalina, pues se establece la reacción $CO_3^{2-} + H^+ \rightleftharpoons HCO_3^-$, quedando un exceso de OH^- que al aumentar el grado de disociación del agua resulta en un pH mayor.

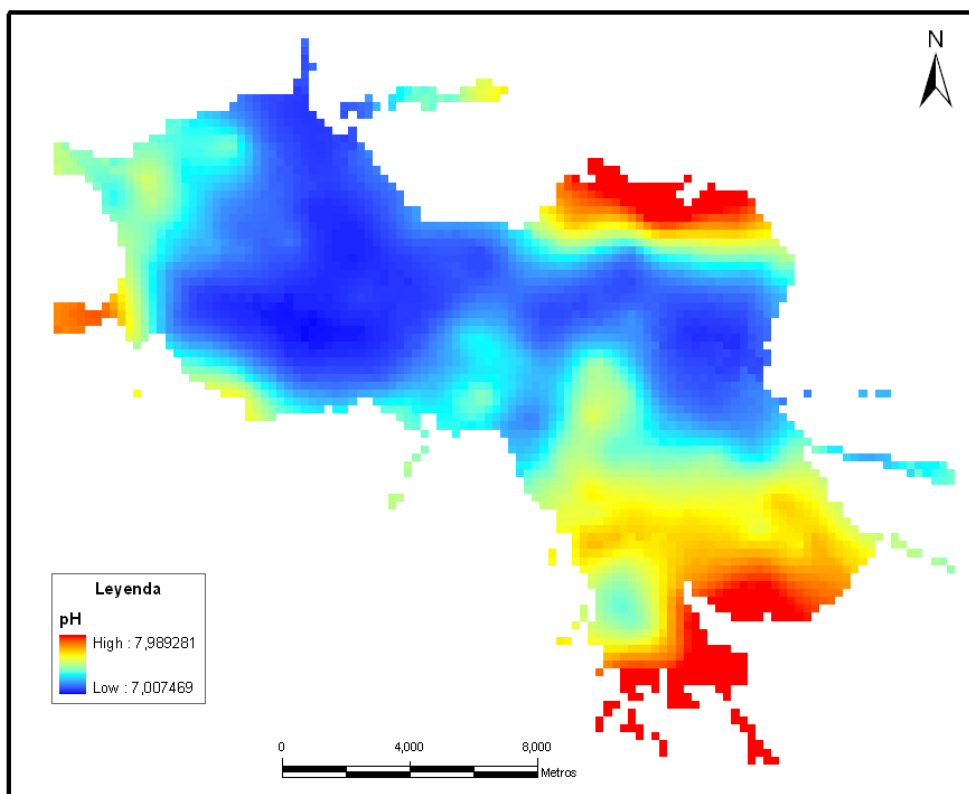


Figura 2.23: Mapa de la distribución espacial del pH en las aguas subterráneas.

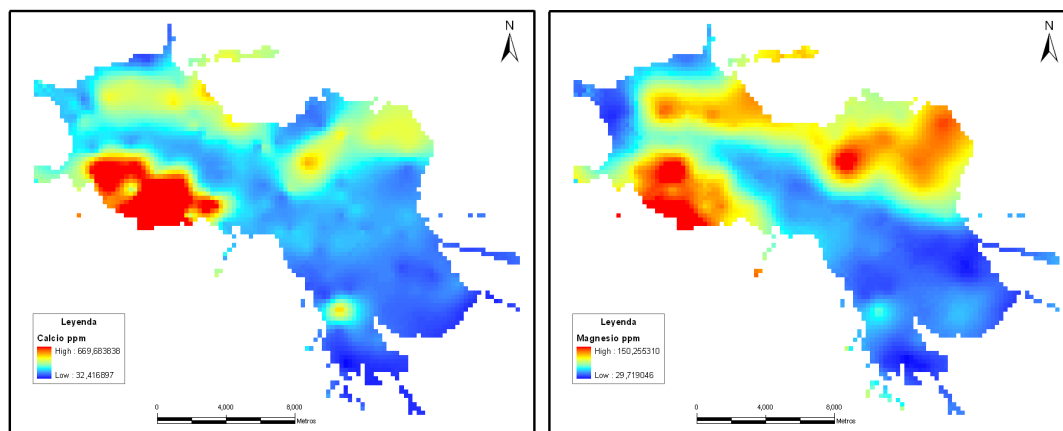
La medida de pH para el área de estudio está referida a 25°C y su rango oscila entre 7 y 7,9 (figura 2.23), ambos límites dentro del intervalo que estipula la reglamentación española para el consumo humano (de 6,5 a 9,5).

En la Vega de Granada, los mayores valores se sitúan al sureste, en la zona de recarga del acuífero, donde las aguas provienen principalmente de Sierra Nevada y al nornordeste donde la mayoría de las aguas provienen de los relieves cercanos y donde se concentra mayormente industria de la zona. También conviene anotar el aumento del pH en el borde oeste, fundamentalmente en la zona de descarga. Estos valores pueden estar influidos por la posible recarga lateral del borde oeste del acuífero

o ya comentado anteriormente para el mapa de temperaturas (figura 2.21b) y en el perfil 9 de la figura 2.22.

El análisis geoestadístico de esta variable indica una elevada aleatoriedad (50 %) aunque el alcance también elevado (11000 m). Igual que para la temperatura, el análisis direccional para el pH muestra una mayor deriva en la dirección norte-sur y una menor variabilidad este oeste, influido por un aumento general de los valores hacia los bordes.

Para la agricultura los valores de esta variable son también importantes, ya que a pH inferiores a 6 y superiores a 9 se dan problemas en la mayoría de los cultivos. Y para



(a) Mapa del contenido en calcio de las aguas subterráneas. (b) Mapa del contenido en magnesio de las aguas subterráneas.

Figura 2.24: Distribución espacial de calcio y magnesio del agua subterránea.

la industria, es importante igualmente ya que valores bajos pueden producir corrosión y valores altos pueden dar problemas de incrustación.

2.2.3.2. Variables químicas

Dentro de este grupo, se distinguen entre componentes mayoritarios y secundarios, según su concentración natural en las aguas. Dentro de los elementos mayoritarios se han incluido en el estudio: calcio, magnesio, sodio, potasio, bicarbonatos, carbonatos, cloruros, sulfatos, nitratos, potasio, amonio y nitritos; (concentración en ppm). De las concentraciones en cada componente se ha elaborado un mapa de su distribución espacial en el acuífero, cuya interpretación se ha fundamentado en los trabajos de Custodio (1996, [50]), Catalán Lafuente (1981, [33]) y Earnshaw (1997, [65]).

El **calcio** es generalmente el catión más abundante en las aguas naturales de salinidad media y procede fundamentalmente de la disolución de los principales minerales que lo contienen, que son la calcita, la dolomita, el yeso y la anhidrita, así como de cementos calcáreos de muchas rocas. En terrenos metamórficos puede proceder del ataque de feldespatos y otros silicatos cálcicos. El agua de lluvia lo aporta en cantidades mínimas.

Está frecuentemente en estado de saturación y su estabilidad en solución depende del equilibrio $CO_2 \rightleftharpoons CO_3H \rightleftharpoons CO_3^{=}$. Puede precipitar con facilidad y le afecta mucho el cambio iónico. Debido a los efectos del ion común y del cambio iónico la concentración o disolución de las distintas sustancias pueden verse muy afectadas.

Los valores fijados por la reglamentación española y europea para las aguas de consumo público no estipulan ningún límite, pero su elevada concentración puede causar

problemas de incrustación y dureza que pueden ser inconvenientes para su uso tanto doméstico como para la industria.

La concentración de este catión en las aguas del acuífero de la Vega está comprendida entre 32 y 669 ppm, las concentraciones máximas se localizan al suroeste del área de estudio (figura 2.24 a), en la zona del aeropuerto y el río Salado, probablemente influenciadas por los materiales evaporíticos de este borde. En el borde norte, se encuentran también elevadas concentraciones, influenciadas igualmente por materiales evaporíticos, a la entrada de los ríos Cubillas y Beiro, aunque la concentración no llega a ser tan alta debido probablemente al tipo y cantidad de rocas evaporíticas en esta zona. La concentración de calcio en las aguas subterráneas es mínima en la zona de cabecera con la entrada de aguas procedentes de Sierra Nevada y va aumentando a lo largo del río Genil hasta la zona de descarga al mezclarse con las aguas de borde más cargadas de calcio en solución.

El **magnesio**, que tiene un comportamiento geoquímico muy similar al del ion calcio, es más soluble que éste (aunque se disuelve lentamente), y tiende a permanecer en solución cuando el calcio precipita. Igualmente es muy afectado por el cambio iónico.

No existe libre en la naturaleza. Se encuentra combinado formando carbonatos como la magnesita o giobertita ($MgCO_3$), o sulfatos y cloruros formando parte de sales dobles en diversos yacimientos. Los silicatos son corrientes, entre los principales: el talco (piedra de jabón) de fórmula $H_2Mg_3(SiO_3)_4$, el asbesto $Mg_3Ca(SiO_3)_4$, la sepiolita o espuma de mar ($Mg_2Si_3O_8 \cdot 2H_2O$), la serpentina ($Mg_3Si_2O_7 \cdot 2H_2O$), y el olivino o peridoto (Mg_2SiO_4).

Su concentración en la naturaleza varía entre 1 y 100 ppm en aguas dulces, aunque en aguas salinas o salmueras puede alcanzar algunos miles de ppm.

En las aguas se encuentra generalmente en cantidades mucho menores que el calcio, pero su importancia biológica es grande ya que es indispensable en el desarrollo de ciertos sistemas enzimáticos y en la constitución de los huesos. Una persona adulta debe tomar por término medio de 300 a 500 mg por día. Si la cantidad de magnesio en el agua es muy grande, por encima de 125 ppm, este agua puede tener efecto laxante e incluso adquirir un sabor amargo, sobre todo si el contenido en ion sulfato es también elevado.

Una particularidad del magnesio es que el aporte del mismo mediante los alimentos, no es suficiente para cubrir las necesidades diarias, por lo que se ha propuesto añadir magnesio a la sal de mesa, igual que se hace con el yodo y el flúor. Se ha encontrado una incidencia alta, en enfermedades de coronaria de corazón y síndromes de muertes de niños por shock de histamina, en zonas donde la concentración de magnesio en agua era menor de 15 ppm, y poca incidencia en áreas que contenían 35 ppm o más (Catalán Lafuente, 1981 [33]).

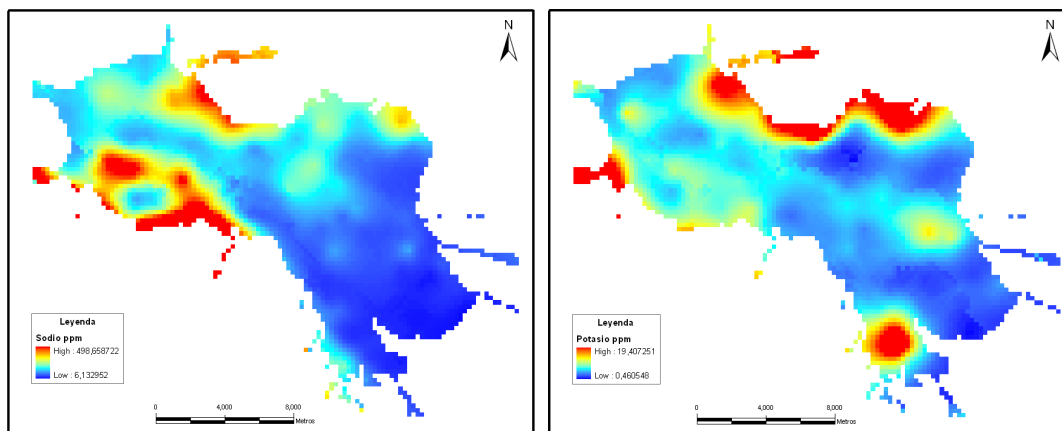
A pesar de ser más soluble que el calcio, la concentración del magnesio en el área de estudio es menor, y varía entre 29 y 150 ppm. La distribución espacial es muy similar a la del calcio (figura 2.24b), con máximos en la zona sur de Sierra Elvira y en la zona del río Salado. La reglamentación española y europea vigente no estipulan ningún límite para su uso en consumo público, ya que su toxicidad es muy baja.

El **sodio** tiene una solubilidad muy elevada y es muy difícil de precipitar. Es un elemento afectado muy fácilmente por el cambio de bases y suele ir asociado al anión cloro, aunque no siempre sucede así. Las aguas naturales con altos contenidos en sodio suelen tener también altos contenidos en flúor y su concentración en aguas dulces generalmente varía entre 1 y 150 ppm, no siendo raro encontrar contenidos mucho mayores de hasta varios miles de ppm. Su procedencia en aguas subterráneas puede estar ligado al ataque de feldespatos, feldespatoides y otros silicatos, aunque en aguas ligadas a rocas de origen sedimentario es más común que proceda del lavado de sedimentos de origen marino y cambio de bases con arcillas del mismo origen. Localmente también puede proceder del lavado de la sal gema (cloruro sódico), carbonato de sodio (sosa trona), borato de sodio (bórax), nitrato de sodio (nitrato de Chile) o de sulfato sódico. Su origen puede estar ligado además a la contaminación urbana (detergentes y otros productos) e industrial (lámparas de vapor de sodio, aleaciones que lo contengan, etc.).

Su elevada concentración en las aguas naturales puede tener efectos perjudiciales en las plantas al reducir la permeabilidad del suelo y es especialmente nocivo si además las concentraciones de calcio y magnesio son bajas. Para su uso industrial también puede resultar inconveniente, sobre todo si se trata de aguas carbonatadas, ya que el bicarbonato sódico puede causar problemas de espuma en calderas.

El límite para el consumo humano se estipula en 200 ppm según el Real Decreto 140/2003, aunque en el área de estudio se alcanzan concentraciones de hasta cerca de 500 ppm. Los mayores valores de las concentraciones en este catión (figura 2.25a) están influenciadas por la existencia de las formaciones sedimentarias que constituyen el borde del acuífero, tanto al sur en la zona del arroyo Salado como al norte en la zona de Sierra Elvira, áreas donde los valores de las concentraciones son máximos. En la zona del arroyo Salado las concentraciones máximas se alinean a lo largo del río. En general en el acuífero las concentraciones aumentan hacia la zona de descarga al mezclarse con las aguas de borde, igual que en los anteriores cationes.

El **potasio** tiene un comportamiento es muy parecido al del sodio. Su solubilidad es muy elevada y es muy difícil de precipitar. Es afectado por el cambio de bases y es adsorbido de forma muy poco reversible por las arcillas en formación (paso de montmorillonita a illita), para formar parte de su estructura, circunstancia que lo diferencia notablemente del sodio. Por ello en aguas naturales, a pesar de la mayor abundancia de



(a) Mapa del contenido en sodio de las aguas subterráneas. (b) Mapa del contenido en potasio de las aguas subterráneas.

Figura 2.25: Distribución espacial de sodio y potasio del agua subterránea.

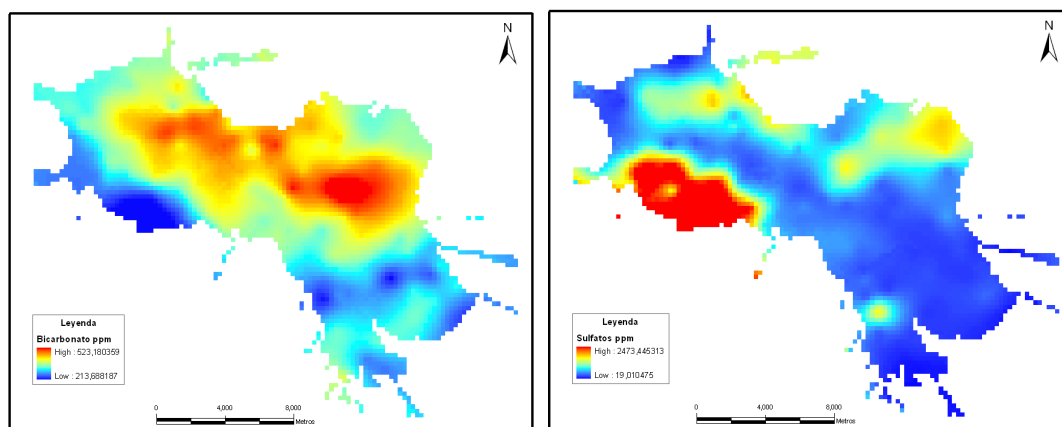
potasio en muchas rocas, la concentración de potasio en las aguas es mucho menor que la de sodio, excepto las muy diluidas.

Su concentración en las aguas proviene del ataque de la ortosa y otros silicatos (micas, arcillas, etc.) y localmente de la disolución de sales potásicas naturales (silvinita, carnalita, etc). También puede proceder de contaminación agrícola, industrial, urbana y minera. La procedencia de potasio por contaminación agrícola se debe a su uso como fertilizante en forma de sulfato de potasio o de cloruro de potasio. La contaminación urbana o industrial de potasio se debe su uso en jabones líquidos (hidróxido de potasio) o blandos (carbonato de potasio). También se usa en la industria del vidrio, en los fósforos y en los fuegos pirotécnicos.

La concentración de potasio en aguas dulces naturales suele variar entre 0,1 y 10 ppm, aunque excepcionalmente puede alcanzar algunos cientos de ppm. El potasio no presenta problemas especiales a las concentraciones habituales, ya que es uno de los nutrientes esenciales para la vida de las plantas y de las personas. Así, la legislación española y europea no establece ningún límite tóxico o peligroso para la salud.

En el área de estudio los valores oscilan entre 0,4 y 19,4 ppm, las máximas concentraciones en el agua del acuífero se localizan a lo largo de todo el borde norte, probablemente influidos por las rocas sedimentarias que forman ese borde. Otros valores máximos se sitúan en el borde sur y en la zona cercana a Granada capital, cuyos orígenes podrían estar ligados a una contaminación industrial y agrícola (jardines) respectivamente.

El **bicarbonato**, así como los carbonatos, proceden de la disolución del CO_2 atmosférico o del suelo, y de la disolución de las calizas y dolomías ayudada por el anhídrido carbónico y/o los ácidos naturales (ácidos húmicos, oxidación de sulfuros, etc.).



(a) Mapa del contenido en bicarbonatos de las aguas subterráneas. (b) Mapa del contenido en sulfatos de las aguas subterráneas.

Figura 2.26: Distribución espacial de bicarbonatos y sulfatos en el agua subterránea.

Existe una estrecha relación entre el CO_2 gas, el CO_2 disuelto, HCO_3^- , $CO_3^{=}$ y el pH. Estos iones comunican alcalinidad al agua en el sentido de que proporcionan capacidad de consumo de ácido al crear una solución tampón e influyen en la solubilidad de la calcita. Estos iones no son oxidables ni reducibles y pueden precipitar muy fácilmente como carbonato cálcico. Si en el acuífero no existen aportes de CO_2 el contenido en carbono disuelto permanece constante y muchas veces puede suponerse que el contenido en bicarbonato es casi constante ya que es el ion dominante. El carbono disuelto es fácilmente precipitable por concentración, por cambio en la presión parcial del CO_2 o por efecto de ion común.

El ion bicarbonato suele tener una concentración en agua dulce entre 50 y 350 ppm con máximos en algunos casos de hasta 800 ppm. El ion carbonato, sin embargo, presenta concentraciones mucho menores y se consideran cero si el pH es inferior a 8,3.

Su elevada concentración no presenta problemas de toxicidad, por lo que no existe regulación de su contenido en la normativa española o europea para agua de consumo público. Sin embargo, para el riego, las aguas bicarbonatadas sódicas son malas debido a la fijación del sodio en el terreno y a la creación de un medio alcalino.

La distribución espacial de la concentración de bicarbonato en la Vega se observa en la figura 2.26a, en la que los mayores contenidos se localizan en la parte central de la vega y con un desplazamiento de los máximos al norte, la concentración varía entre 212 y 523 ppm. El contenido en carbonatos es cero limitado por el pH de las aguas subterráneas de la Vega.

Los **sulfatos** en las aguas subterráneas proceden del lavado de terrenos formados en ambiente marino o de gran aridez, de la oxidación de los sulfuros de rocas ígneas, me-

tamórficas o sedimentarias o de la disolución de minerales evaporíticos (yeso, anhidrita, thenardita, glauberita, etc.). Son sales moderadamente solubles a muy solubles, excepto las de estroncio y bario. En agua pura el límite de solubilidad del sulfato cálcico se alcanza a 1500 ppm, sin embargo en aguas salinas este límite se alcanza a concentraciones mucho más altas y puede llegar hasta 7.200 ppm. En aguas dulces su concentración varía entre 2 y 150 ppm, en aguas salinas su concentración alcanza los 5.000 ppm si existe calcio y hasta 200.000 si está asociado a magnesio y sodio en ciertas salmueras.

Las aguas selenitosas (con elevado contenido en sulfato) no quitan la sed y tienen sabor poco agradable y amargo. Por sí mismo o si va asociado a magnesio o sodio en cantidad importante puede comunicar propiedades laxantes. La normativa española establece su límite de potabilidad en 250 ppm.

Para la agricultura y las plantas puede ser perjudicial, y en más de algunos centenares de ppm puede incluso perjudicar a la resistencia del hormigón y del cemento, por lo que el uso de aguas con elevados contenidos de sulfato en cementeras debe ser estudiado adecuadamente.

La distribución de la concentración de este ion en la Vega de Granada (figura 2.26 b) muestra que los mayores contenidos se encuentran al sur en la zona del aeropuerto con valores de hasta 2.473 ppm, por lo que se supera ampliamente el límite fijado por la normativa para consumo humano. Estas elevadas concentraciones están ligadas a las formaciones de yeso del Mioceno superior localizadas en el borde sur de la Vega. En la zona norte también existe un aumento de las concentraciones de sulfato aunque inferior, con valores alrededor de 800 ppm, ligado a las formaciones de yeso del Trias germano-andaluz y otras formaciones con contenidos evaporíticos, menos abundantes y menos extensas. Los valores mínimos (desde 19 ppm) se sitúan en las zonas de entrada de las aguas de recarga, al sureste del acuífero.

Los **cloruros** son sales en general muy solubles, altamente estables en disolución y difícilmente precipitables. El ion cloruro no se oxida ni se reduce en aguas naturales, y en general va asociado al ion sodio, especialmente en aguas salinas.

En aguas dulces naturales su concentración suele variar entre 10 y 250 ppm, no siendo raro encontrar contenidos mucho mayores hasta varios miles de ppm. Los cloruros proceden del lavado de terrenos de origen marino; las aguas congénitas y fósiles pueden aportar cantidades importantes del mismo, también pueden proceder del agua de lluvia o del lavado de otro tipo de terrenos. Su concentración en la corteza terrestre es escasa por lo que las aportaciones procedentes del ataque de rocas, son pequeñas. Localmente pueden aparecer en las aguas subterráneas por el lavado de rocas evaporíticas (sal gema, silvina, carnalita, etc.) y por el ataque de ciertos minerales asociados a rocas ígneas y metamórficas (sodalita, apatito, etc.). También pueden estar relacionados con problemas de salinidad en acuíferos costeros (Padilla et al., 1997 y López Geta et al., 2003

[134, 108]), ya que en el agua del mar sus concentraciones son muy altas. En las aguas continentales, por lo tanto, es menos importante por lo que no se suele alcanzar el límite de solubilidad casi nunca. Es además, muy difícilmente alterable por cambio iónico u otro tipo de acción. Cabe mencionar aquí que la concentración de cloruros tiende a ir aumentando en la evolución de la mineralización regional de las aguas subterráneas, así en primer lugar el ion carbonato alcanza el límite de solubilidad saturando la concentración en este ion, más tarde el ion sulfato y posteriormente el ion cloruro que no suele llegar a alcanzar dicho límite, adquiriendo el agua de esta manera una composición muy distinta de la roca por la que circula.

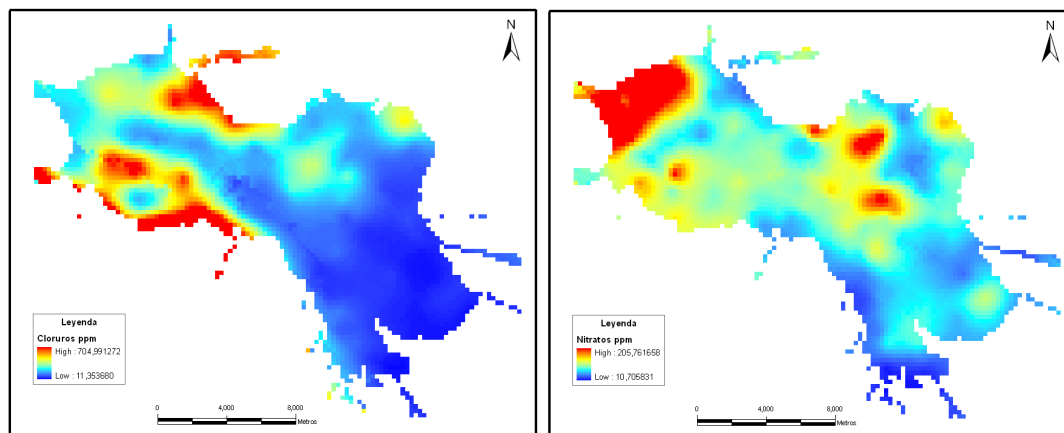
El origen antrópico de este ion puede estar ligado a la contaminación por vertidos urbanos e industriales, que pueden aportar cantidades importantes, ya que el uso del cloro en distintos compuestos está muy extendido por su acción blanqueante y desinfectante en la industria del papel y los pesticidas.

Con una concentración superior a 300 ppm el agua toma un sabor salado, aunque no es perjudicial por lo menos hasta unos miles de ppm. Es muy importante para la vida ya que resulta esencial en el proceso digestivo. Para las plantas, los contenidos elevados son perjudiciales y además comunican corrosividad al agua.

El mapa de la distribución de las concentraciones de cloruros en las aguas subterráneas de la Vega, se ha representado en la figura 2.27a. Al igual que para el sodio, los valores máximos se sitúan en los bordes norte y sur del área de estudio, por lo que probablemente su origen esté ligado a la disolución de materiales evaporíticos (halita) del Mioceno y del Triás. Los valores máximos en estas zonas alcanzan los 704 ppm, por lo que el límite de potabilidad establecido por la normativa española en 250 ppm se supera en un 200 %. Los valores mínimos presentan una zonación muy clara a lo largo del río Genil y con un aumento hacia la zona de descarga. Esta lineación tan clara también es debida al importante papel que juega el cloro como trazador químico. Según Sánchez-Díaz (2004 [170]), a lo largo de los últimos años se ha producido en la Vega de Granada un incremento a nivel general de la concentración de cloruros debida principalmente a una disminución en el aporte de aguas menos salinas. Es decir, por la regulación de las aguas de cabecera provenientes de Sierra Nevada y por la recirculación de aguas más salinas.

Los **nitratos** pueden aparecer de manera natural en los suelos y posteriormente pasar al agua debido a los procesos de nitrificación de las plantas leguminosas y por la descomposición de la materia orgánica. Sin embargo, el aporte más común de los nitratos en las aguas subterráneas suele estar ligado a la contaminación agrícola por abonos, a la contaminación ganadera a partir de los purines y a la industrial.

También, más raramente, pueden aparecer por la descomposición de ciertos minerales nitrogenados, emanaciones volcánicas o lavado de suelos antiguos.



(a) Mapa del contenido en cloruros de las aguas subterráneas. (b) Mapa del contenido en nitratos de las aguas subterráneas.

Figura 2.27: Distribución espacial de cloruros y nitratos en las aguas subterráneas.

Los nitratos son sales muy solubles y por lo tanto son muy difícilmente precipitables. Aunque tienen tendencia a ser estables aún en medios reductores, puede pasar a N_2 o NH_4^+ , y excepcionalmente a NO_2^- . La mayoría de los compuestos pasan a NO_3^- en medio oxidante, aunque el NH_4^+ puede perderse en parte como N_2 . El terreno puede fijar el NO_3^- , especialmente en suelo vegetal. Además comunican corrosividad al agua y producen interferencias en fermentaciones.

La concentración de nitratos normalmente está comprendida entre 0,1 y 10 ppm, pero en aguas polucionadas puede llegar a los 200 ppm y en algún caso hasta 1000 ppm. Muchas veces es un indicador de la contaminación, en cuyo caso suele estar estratificado (NWRC, 1984 [4]), en estos casos las concentraciones mayores se disponen en la parte alta del acuífero libre. Si las concentraciones son menores a 10 ppm se considera que son bajas, aunque en aguas subterráneas no contaminadas es raro superar algunos ppm.

Las concentraciones elevadas en agua de consumo público pueden ser peligrosas, pues pueden producir cianosis en los niños y metaheglominemia. El límite superior de potabilidad según la normativa española y europea, se establece en 50 ppm. Además, comunican corrosividad al agua y comunican interferencias en fermentaciones.

En la Vega el mayor problema de contaminación detectado hasta hoy son los nitratos, hasta el extremo de que las concentraciones llegan en algunos casos a valores de 205 ppm en la zona noroeste y superiores a 100 ppm en varios puntos entre Maracena y Atarfe (figura 2.27 b). Los valores mínimos y por tanto las zonas con menor riesgo para la salud se localizan en la zona sureste a la entrada de los ríos Dílar, Genil y Monachil.

Los **fluoruros** se encuentran en la naturaleza formando parte de algunas rocas, en las arcillas y en el carbón. Forman parte de algunos minerales, fundamentalmente el

fluoruro de calcio (fluorita), el fluorapatito ($Ca_5(PO_4)_3F$) y la criolita (Na_3AlF_6) y pueden aparecer en el agua subterránea por el lavado de estos minerales. Los fluoruros de los metales alcalinos son solubles en agua e higroscópicos; sin embargo, los de los metales alcalinotérreos, especialmente el calcio y el bario, son poco solubles. El flúor se hidroliza ligeramente por lo que su solubilidad tiene gran importancia en la alcalinidad del agua.

La concentración de flúor en la corteza terrestre se estima que es del 0,065 %, por lo que es casi tan abundante como el carbono o el nitrógeno. Los fluoruros son liberados al aire cuando el viento arrastra el suelo. También pueden pasar al aire como fluoruro de hidrógeno en los procesos de combustión de las industrias que acabará depositándose en el suelo y/o pasará al agua.

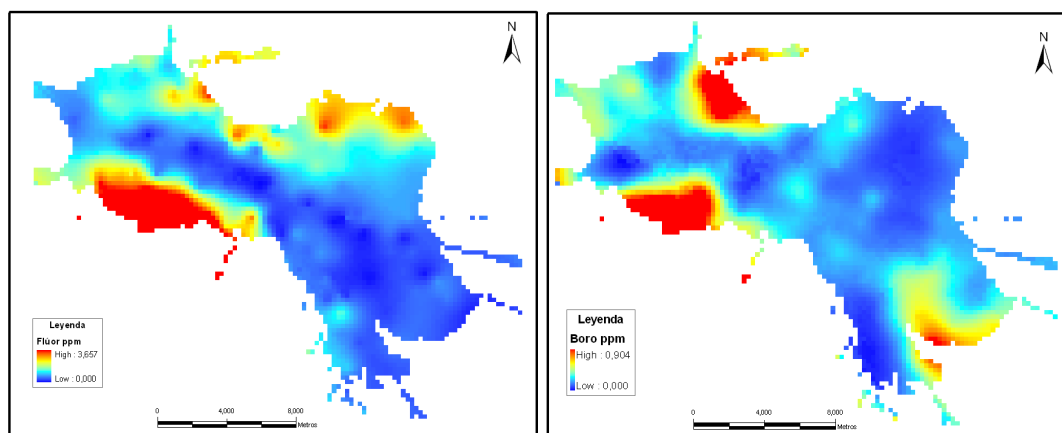
Su origen también puede estar ligado a la contaminación antrópica a partir de vertidos industriales, abonos, etc.

En aguas dulces su concentración frecuentemente oscila entre 0,1 y 1 ppm, en algunos casos su concentración puede alcanzar hasta 10 ppm y raramente 50 ppm en aguas muy sódicas y con muy poco calcio, ya que el calcio limita la concentración de flúor.

En el medio ambiente el flúor no puede ser eliminado, por lo que se acumula en el suelo o las plantas, y puede provocar daños en las hojas e influir en el crecimiento si la planta en cuestión es sensible al flúor, incluso para bajas concentraciones.

El flúor es esencial para la vida y está muy relacionado con la solidez de los huesos y el decaimiento de los dientes, por este motivo suele añadirse artificialmente al agua hasta concentraciones de 1 ppm. Sin embargo la fluoración sistemática del agua es aún cuestionable por el aumento total de flúor procedente de varias fuentes como la pasta fluorada o los alimentos (Leveret, 1986 [105]). Una absorción elevada o demasiado frecuente del flúor puede tener graves contraindicaciones, como la fluorosis, provocar caries, osteoporosis y daños en los riñones, huesos, nervios y músculos. También puede provocar la disminución de la cantidad de alimento tomada por el estómago, provocar bajo peso al nacer y en los animales, alterar el desarrollo de las garras (Thomas G. Reeves, 2002 [178]). La OMS establece una serie de requisitos básicos para que la fluoración pueda ser considerada una medida eficaz: grado suficiente de desarrollo económico, constancia de que la población bebe agua de la red pública, moderada frecuencia de caries en la comunidad, etc. Siguiendo estos criterios, en nuestro país se han desarrollado algunas experiencias piloto (Marín Camaches, M.D., 1998 [115]), pero la reglamentación española no establece un límite mínimo, siendo estas experiencias más bien anecdóticas. Sin embargo, debido a la toxicidad del flúor sí se establece un límite máximo en 1,5 ppm.

El mapa de distribución de este ion en la Vega (figura 2.28a) muestra una variación de los contenidos entre cero y 3,6 ppm, localizándose el máximo en la zona del aeropuerto (borde sur) donde las aguas no son aptas para el consumo público ya que se supera el



(a) Mapa del contenido en fluoruros del agua subterránea. (b) Mapa del contenido en boro del agua subterránea.

Figura 2.28: Distribución espacial de fluoruros y boro en las aguas subterráneas.

límite de potabilidad en más de un 100 %. Castillo Martín (2004 [31]) en un estudio de los fluoruros de la Vega, encuentra un aumento de su contenido en el año 2003 respecto al año 1984, lo que explica por una disminución de las aguas más pobres en fluoruros (de recarga de los ríos Genil, Monachil y Dílar) y un posible aumento del aporte del río Salado, así como de las aguas de retorno más salinas.

El **boro** procede de la precipitación atmosférica, que contiene pequeñas cantidades de boro en disolución y del vulcanismo y/o la erosión de rocas con concentraciones variables de boro. El boro es un elemento sublimable, con baja preferencia por las fases fundidas, es un elemento incompatible con la fase sólida (pegmatitas) muy móvil en fase acuosa y con alta afinidad por los minerales arcillosos (elemento litófilo). Por este motivo es mucho más abundante en rocas sedimentarias (300 ppm) que en rocas ígneas (3 ppm). Los minerales principales en los que se encuentra el boro son en su mayoría constituyentes de las rocas evaporitas, como el bórax ($Na_2B_4O_7 \cdot 10H_2O$), altamente soluble en agua, la colemanita, la kernita y la ulexita.

El boro en las aguas subterráneas se encuentra disuelto como BO_3H_3 , parcialmente disociado a $BO_3H_2^-$ y su presencia contribuye algo a la alcalinidad. Su origen, en general, está en relación con el lavado de minerales evaporíticos, aunque también puede pasar al acuífero como producto de contaminación antrópica industrial en actividades relacionadas con la fabricación de vidrios, de plantas generadoras de energía por combustión de carbón, la fundición del cobre y por contaminación agrícola debido a su uso como pesticidas y fertilizantes agrícolas.

El boro puede pasar a los animales a través de las plantas que lo absorben del suelo, aunque no se acumula en los tejidos animales. También puede afectar a los seres

humanos por la exposición al polvo de borato en el lugar de trabajo en lugares cercanos a las zonas de extracción de bórax, plantas de refinación y zonas donde se produce ácido bórico.

La ingesta elevada de boro por los animales puede causar problemas en los órganos reproductivos masculinos, y la exposición al boro durante el embarazo puede producir defectos de nacimiento y problemas en el desarrollo. Para los seres humanos también es tóxico en elevadas cantidades, pudiendo infectar el estómago, el hígado, los riñones y el cerebro, e incluso eventualmente puede llevar a la muerte. La exposición en pequeñas cantidades tiene lugar irritación de la nariz, garganta y ojos. La normativa española fija su contenido en 1 ppm para las aguas de consumo público.

La distribución de este metaloide en la Vega se muestra en el mapa de la figura 2.28b. Los máximos se sitúan cercanos a los bordes de Sierra Elvira y al sur del aeropuerto, lo que indicaría un origen ligado a las rocas evaporíticas de estos bordes. Cabe destacar un pequeño máximo situado en la zona de recarga entre los ríos Dílar y Genil, cuyo origen puede estar relacionado con la disolución de minerales de boro de rocas metamórficas de Sierra Nevada o con alguna contaminación antrópica, extremo este último para el que no hay indicios suficientes.

El **plomo** se encuentra en la naturaleza en forma de carbonato de plomo $PbCO_3$ (cerusita), sulfuro PbS (galena), sulfato $PbSO_4$ (anglesita), etc. Es un metal blando y pesado que por la acción del anhídrido carbónico de la atmósfera y en presencia de humedad se oxida y se transforma en carbonato de plomo, de manera que impide ataques ulteriores. Si se pone en contacto con aguas muy salinas, se forman las sales insolubles (sulfatos, carbonatos de plomo, calcio, etc.) que también impiden un ataque posterior. Sin embargo, el agua destilada o la procedente de terrenos graníticos o de lluvia (en general la poco mineralizada), con oxígeno disuelto, atacan al metal formando hidróxido de plomo $Pb(OH)_2$ que es ligeramente soluble.

También son agresivas al plomo las aguas ácidas, favoreciendo igualmente el ataque la presencia de nitratos, así como cuando se encuentran juntos el oxígeno y el anhídrido carbónico. Igualmente, la agresividad al plomo aumenta con la temperatura.

En aguas superficiales naturales se encuentra en cantidades del orden de 0,01 ppm que pueden ser superadas debido a la polución, p. e. la acción de disolución del agua en las tuberías de plomo posibilita una concentración de hasta 0,1 ppm, el doble de lo permitido por la OMS y cuatro veces lo que marca la legislación española hasta el año 2013 (25 ppb), ya que posteriormente el límite será de 10 ppb.

En el agua natural, el plomo se encuentra en la forma Pb^{2+} , $Pb(OH)^+$ y $Pb(CH_3)_4$. En este medio los complejos más significativos, si el agua contiene carbonatos, son: $PbCO_3$, $Pb(OH)_2$ y $Pb_3(OH)_2(CO_3)_2$, con un intervalo de solubilidad entre los pH de

5 a 8 en el caso del carbonato, superior a 8,5 en el caso del hidróxido y en un rango de 7,5 a 8,5 en el caso del hidrocarbonato de plomo.

El plomo se encuentra de forma natural en el ambiente, pero las mayores concentraciones encontradas en el ambiente son el resultado de las actividades humanas.

Uno de los usos más amplios del plomo se encuentra en la fabricación de acumuladores, junto con la fabricación de tetraetilplomo, forros para cables, elementos de construcción, pigmentos, soldadura suave y municiones.

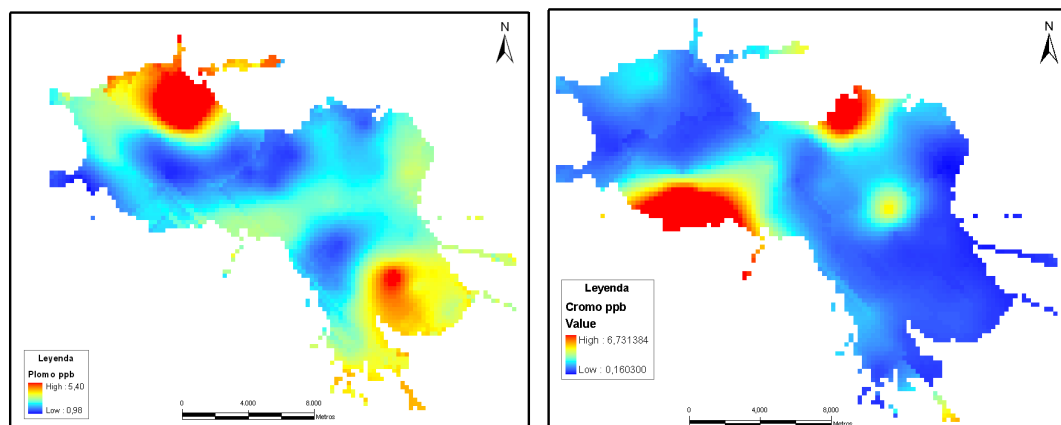
El uso del plomo en la gasolina favorece la puesta en marcha de un ciclo no natural de este metal, motivo por el cual ha sido retirado de la misma. Otras actividades humanas, como la combustión del petróleo, procesos industriales, combustión de residuos sólidos, también contribuyen a su liberación al medio ambiente. El plomo puede terminar en el agua y en los suelos a través de la corrosión de las tuberías de plomo en los sistemas de transportes y a través de la corrosión de pinturas que contienen plomo.

El aporte de plomo a los organismos puede ser por medio de los alimentos o del aire o del agua. Aparentemente no tiene efectos beneficiosos para la salud humana y todas sus sales son muy tóxicas, ya que producen saturnismo. Su ingestión diaria mediante agua de bebida, con concentraciones de 1 mg/l, puede producir una intoxicación crónica, debido a que se acumula en los tejidos, especialmente en el óseo. También puede entrar en el feto a través de la placenta de la madre y puede causar daños al sistema nervioso y al cerebro de los niños en la gestación.

En el medio ambiente también puede ser muy tóxico por la acumulación del mismo en los organismos acuáticos y los organismos del suelo. Estos organismos pueden sufrir envenenamiento por plomo y pasar a la cadena alimenticia. A causa de su fuerte toxicidad, los niveles de plomo en las aguas de bebida deben mantenerse tan bajos como sea posible, motivo por el cual está en estudio bajar la concentración máxima permitida hasta 10 ppb en el año 2014 en España.

La distribución de este metal en las aguas subterráneas de la Vega (figura 2.29a) varía entre valores inferiores a 1 ppb hasta 5,4 ppb, por debajo del límite mínimo exigido en todos los casos. Los valores más altos se localizan en la zona noroeste y al sur de Granada capital. La situación de estos máximos no parece mostrar una influencia relacionada con los núcleos de población (gasolineras con plomo), ya que las mayores densidades de población se sitúan al norte de la mencionada zona (Granada capital). Debido a su situación y a su concentración, estos máximos habría que asociarlos a una contaminación industrial puntual o a un origen natural por disolución de las rocas limítrofes o del sustrato del acuífero.

El **chromo** no existe libre en la naturaleza, el mineral más abundante que lo contiene es la cromita ($Fe(CrO_2)_2$) y se encuentra en los alimentos y el humo de los cigarrillos.



(a) Mapa del contenido en plomo del agua subterránea. (b) Mapa del contenido en cromo de las aguas subterráneas.

Figura 2.29: Distribución espacial de plomo y cromo en las aguas subterráneas.

En las aguas naturales no existe normalmente, por lo que su presencia suele deberse a una polución de origen industrial, principalmente de talleres de cromado, tenerías, peleterías, industrias textiles, pinturas, y otras aplicaciones industriales. A través de la combustión de carbón también se libera cromo que eventualmente se disolverá.

Este metal se puede presentar en varios estados de oxidación, pero en los sistemas acuosos sólo se encuentran el trivalente y el hexavalente. El cromo trivalente se encuentra en el agua principalmente como catión ya que el hidróxido es muy insoluble. El cromo hexavalente se encuentra como anión en las formas $HCrO_4^{2-}$, CrO_4^- o $Cr_2O_7^{2-}$ dependiendo una u otra del pH. Todas estas formas son muy solubles y las más habituales.

Hay evidencias científicas de que, dependiendo de la cantidad de cromo en el agua, su consumo puede ser saludable o no para la salud humana. El cromo es esencial para el metabolismo y actúa como cofactor con la insulina para mantener la tolerancia normal de la glucosa. Existe una clara asociación entre la concentración de cromo en el agua de bebida y el aumento de la diabetes juvenil. Además, el cromo puede actuar también como protector frente a la enfermedad coronaria.

El cromo se encuentra en el organismo en forma trivalente aunque sea el cromo hexavalente el que mejor se asimila. Por otra parte el cromo hexavalente es un tóxico muy fuerte si se inhala ya que puede producir cáncer de pulmón en las personas expuestas a él, así como sensibilizaciones en la piel.

El nivel de cromo en el aire y en el agua es generalmente bajo. En el agua de bebida el nivel de cromo es usualmente bajo, así como en las aguas subterráneas, pero el agua de pozo contaminada puede contener el peligroso cromo hexavalente. El principal meca-

nismo natural de consumo de cromo es a través del alimento, principalmente hortalizas, frutas, carnes, levaduras y semillas, en las que se encuentra en la forma trivalente. La manera de preparar la comida y su almacenaje pueden alterar el contenido de cromo en la comida. Así, cuando la comida es almacenada en tanques de acero o latas, las concentraciones de cromo pueden aumentar.

Los cultivos contienen sistemas para gestionar la toma de cromo para que ésta sea lo suficientemente baja como para no causar cáncer. Pero cuando la cantidad de cromo en el suelo aumenta, o existe acidificación, la concentración puede aumentar. Las plantas usualmente absorben sólo cromo trivalente, que es esencial, pero al exceder cierto valor, sus efectos pueden ser negativos.

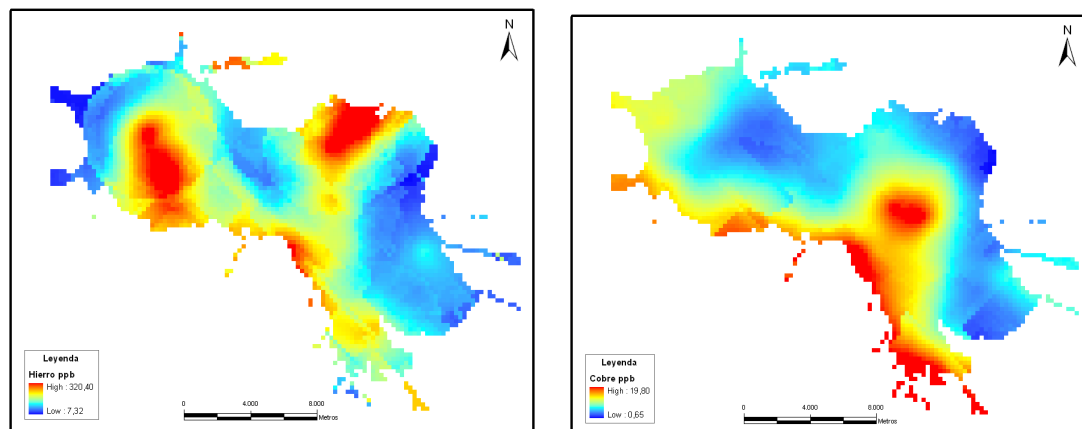
El límite máximo permitido por la legislación española para aguas de consumo público se sitúa en 50 ppb. La diferencia entre las distintas legislaciones es que en unas se considera el cromo total, como en la española, y en otras sólo el hexavalente.

La distribución del cromo total en la Vega de Granada (2.29b) presenta los mayores contenidos al sureste de Sierra Elvira y en la zona del río Salado con valores cercanos a los 10 ppb y muy por debajo del límite de potabilidad. El origen más probable del máximo de concentración situado en la zona norte es posiblemente la contaminación antrópica originada por los polígonos industriales que se localizan cercanos a esa área. Por otra parte, el máximo reflejado en la zona del río Salado puede tener un origen natural relacionado con las mineralizaciones de estroncio de Montevives (al sur del área de estudio).

El **hierro** es un elemento muy extendido en la naturaleza y se encuentra presente en todos los terrenos. No existe libre; sus compuestos más importantes son: hematites roja Fe_2O_3 ; hematites parda o limonita $Fe_2O_3 \cdot nH_2O$; magnetita Fe_3O_4 ; Siderita $FeCO_3$; Pirita FeS_2 , etc.

El hierro se encuentra disuelto en muchas aguas naturales, fundamentalmente en las aguas subterráneas, ya que las sales solubles son en general las ferrosas. Principalmente se encuentra disuelto en forma de bicarbonato ferroso $Fe(HCO_3)_2$. En contacto con el oxígeno disuelto en el agua, estas sales pasan a férricas y precipitan en forma de hidróxido férrico, cuando el pH es superior a 7.5 la precipitación es instantánea. En las aguas superficiales se encuentra a veces en el estado de complejos organoférricos y en casos raros como sulfuros. También es frecuente que se presente en forma coloidal en cantidades apreciables.

En las aguas con pH mayor de 2,2 el hidróxido férrico es insoluble. El ion ferroso lo es en las aguas con pH mayor de 6. Según esto, las aguas subterráneas, que por estar fuera del contacto con el aire se encuentran en un medio natural fuertemente reductor, podrán tener en solución cantidades notables de hierro ferroso.



(a) Mapa del contenido en hierro de las aguas subterráneas. (b) Mapa del contenido en cobre de las aguas subterráneas.

Figura 2.30: Distribución espacial del contenido en hierro y cobre en las aguas subterráneas.

Por otra parte, a veces su origen no es natural, sino que proviene de la disolución de las tuberías fabricadas con este metal. En este sentido, en la Vega muchos pozos cuentan con instalaciones fabricadas con este metal por lo que los contenidos detectados pueden provenir de la disolución del hierro del mismo pozo.

Es un constituyente normal del organismo (forma parte de la hemoglobina) y sus sales no son tóxicas en pequeñas cantidades. Comunica un sabor al agua astringente, aunque no quiere decir que el agua no sea potable. Juega un papel importante en el ciclo de los fosfatos, lo que hace que su importancia sea muy grande desde el punto de vista biológico. En la naturaleza se presenta bajo las formas de asimilable y no asimilable.

Las aguas que contienen hierro suelen ser inadecuadas para determinados procesos industriales a causa de las propiedades catalizantes que el mismo tiene, por lo que debe ser eliminado totalmente.

La legislación española permite un máximo de 200 ppb para las aguas de consumo público, el cual es superado en 5 puntos de la Vega, 2 situados al norte de Chauchina, 2 al norte de Maracena y uno en Cúllar Vega. Estos puntos a priori se asocian a la disolución de la propia canalización, ya que la mayoría de ellos corresponden a pozos de bombeo manual o cortijos con antiguas instalaciones. La distribución del hierro en la Vega se puede ver en la figura 2.30a, donde los máximos corresponden a las zonas nombradas anteriormente.

El **cobre** se encuentra en la naturaleza libre y combinado, formando numerosos compuestos: cuprita (Cu_2O), calcopirita ($CuFeS_2$), malaquita ($CuCO_3$), etc. El cobre es esencial para la formación de hemoglobina. Es un elemento normal en la materia viviente y presenta un gran interés biológico. Entra en la composición de las enzimas

y se encuentra en la hemocianina, pigmento respiratorio de los moluscos y de algunos crustáceos.

Se presenta muy raramente en aguas dulces (de 5 a 1000 ppb) y es interesante su estudio dado que en el estado de trazas es indispensable para el organismo. Las canalizaciones de las conducciones de agua son además fabricadas con este metal, el cual es atacado por la mayor parte de las mismas. Este ataque es favorecido por el oxígeno, la acidez, la temperatura, etc.

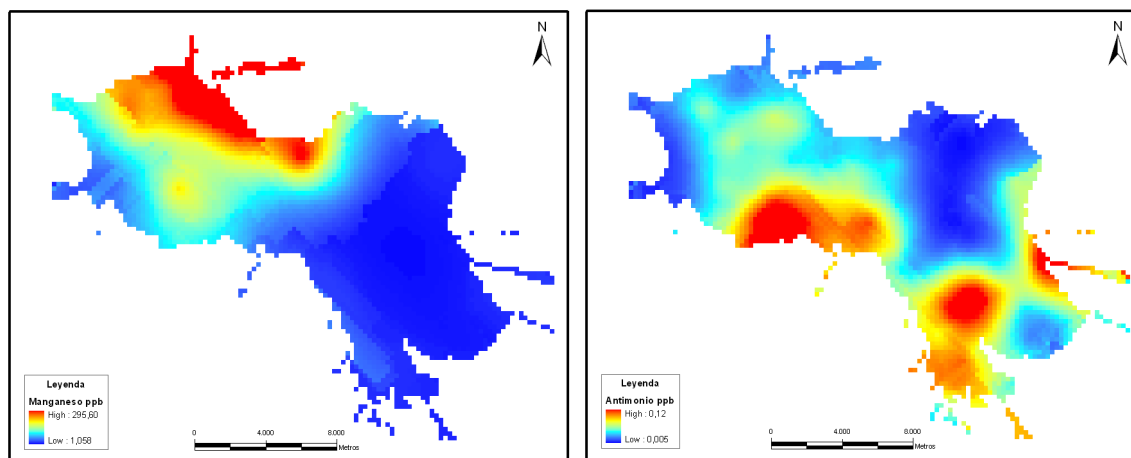
El cobre se encuentra en el agua asociado principalmente con materia orgánica coloidal. En el agua se encuentran principalmente los iones Cu^+ , Cu^{2+} , $Cu(OH)^+$, $Cu(NH_3)_x^{2+}$. En dosis relativamente importantes es tóxico, aunque su ingestión improbable debido a que por encima de 1 ppm presenta un sabor desagradable.

Uno de los compuestos más importantes en las aguas es el sulfato de cobre hidratado ($Cu(SO_4) \cdot 5H_2O$), de color azul, que se emplea, por su acción bactericida y algicida, para tratar el agua fundamentalmente de los depósitos, piscinas e incluso embalses que se encuentran sometidos a la acción del sol, con dosis de 0,1 a 2 ppm de ion cobre según los casos. Estas dosis son inofensivas y no presentan riesgo alguno, ya que, por otra parte, están en el orden de magnitud de las necesidades de este ion para las personas adultas.

El cobre también puede pasar a las aguas subterráneas a partir de fuentes antrópicas de contaminación. Los usos más frecuentes se dan en la agricultura, donde es usado como fungicida e insecticida; en la industria para prevenir la putrefacción de la madera, las telas, cuerdas y redes de pesca; en pigmentos; en soluciones galvanoplásticas; como mordentes en teñido y como catalizadores.

Las concentraciones de cobre permitidas por los distintos países en el agua de consumo público varían desde cero a 3 ppm. En España el límite superior se establece en 2 ppm y en la Vega su contenido oscila entre 0,7 y 19,8 ppb, en todos los casos por debajo del límite que marca la Ley. La distribución del cobre en las aguas subterráneas de la Vega (figura 2.30b), señala los máximos contenidos a lo largo de todo el borde sur de la Vega, con un máximo más localizado al oeste de Granada. Esta distribución puede indicar una diferencia en el contenido en cobre de las rocas de borde o de los sedimentos transportados desde ese borde con respecto a los de la zona norte.

El **manganeso** no se encuentra libre en la naturaleza, sino que aparece formando minerales como pirolusita (MnO_2), braunita (Mn_2O_3), blenda de manganeso (MnS), etc. Este elemento se encuentra normalmente en el organismo como un activador de ciertas enzimas, aunque si se ingiere en grandes dosis es un veneno que afecta al sistema nervioso central.



(a) Mapa del contenido en manganeso de las aguas subterráneas. (b) Mapa del contenido en antimonio de las aguas subterráneas.

Figura 2.31: Distribución espacial del contenido en manganeso y antimonio en las aguas subterráneas.

No se encuentra normalmente en las aguas naturales. Solamente en casos excepcionales aparece en las aguas ácidas y generalmente se encuentra asociado al hierro y en el caso de contaminación industrial.

Su presencia en las aguas provoca el desarrollo de ciertas bacterias que producen la formación de depósitos insolubles de estas sales, en las que pasa del estado manganesoso en solución al estado mangánico en el precipitado. Esta acción es realizada igualmente por el hierro.

Cuando existe en cantidades apreciables, produce en el agua un sabor desagradable, lo que hace que su presencia sea notada al beber y su acción tóxica más fácilmente evitada. La legislación española establece su límite máximo en 50 ppb, valor que es superado en 7 de los puntos estudiados y alcanza valores de hasta casi 300 ppb (figura 2.31a). Como se verá más adelante el manganeso también aparece en la solución del suelo, con valores alrededor de los 30 ppb y su origen se asocia a la formación de margas por hidromorfía.

El **antimonio** se puede encontrar en los suelos, en el aire y en el agua en pequeñas cantidades y contamina principalmente los suelos. En la naturaleza se encuentra como estibinita, antimonita (Sb_2S_3) y valentinita (Sb_2O_3), como producto de la descomposición de la estibinita. Es un mineral calcófilo y forma parte de los minerales de cobre, plata y plomo y también se encuentra en los antimonuros metálicos $NiSb$ (breithauptita), $NiSbS$ (ulmanita) y Ag_2Sb (dicrasita); también existen numerosos tioantimoniatos como el Ag_3SbS_3 (pirargirita). Su concentración media en la corteza terrestre se estima entre 0,2 y 0,5 ppm por lo que es un mineral muy minoritario.

Es liberado al medio ambiente desde fuentes naturales e industriales y puede permanecer en el aire adherido a partículas muy pequeñas por muchos días. La mayoría del antimonio del aire acaba depositándose en el suelo, donde se adhiere firmemente a partículas que contienen hierro, manganeso y aluminio.

La exposición a elevados contenidos de antimonio es altamente tóxica para los animales y los seres humanos y puede causar problemas de irritación ocular, caída del pelo, problemas cardíacos, pulmonares e incluso la muerte. Los mayores problemas de intoxicación se producen principalmente al respirar antimonio enlazado con hidrógeno en la fase gaseosa, pero su exposición puede ser también a través del agua, la comida o por contacto cutáneo con tierra o agua que lo contenga.

Su existencia en la naturaleza puede ser debida a fenómenos naturales o a procesos de contaminación, los cuales están generalmente ligados a su uso en las aleaciones (principalmente plomo), los acumuladores, el revestimiento de cables, cojinetes antifricción y diversas clases de metales de consumo.

La legislación española establece su límite superior en 5 ppb. En todos los casos estudiados en la Vega la concentración está por debajo de este límite y el valor máximo alcanzado es de 0,12 ppb. Las mayores concentraciones se localizan (figura 2.31a) al sur del área de estudio, aunque las concentraciones son tan bajas que apenas es apreciable este dato.

El **arsénico** se encuentra libre en la naturaleza. Combinado se encuentra en forma de diversos minerales, entre los cuales están el sulfuro arsenioso rojo, As_2S_2 (rejalgar), sulfuro arsénico amarillo As_2S_3 (oropimente), pirita arsenical $FeAsS$, etc.

Todos los compuestos solubles son venenosos. No obstante, los organismos que los consumen desarrollan cierta tolerancia a estos compuestos, por lo que algunos tienen aplicación en medicina.

La necesidad de limitar la concentración del arsénico en las aguas destinadas al consumo y usos domésticos viene impuesto por de su alta toxicidad. Su origen natural queda minimizado debido al aporte que se hace artificialmente como consecuencia de su utilización en pesticidas (los arsenitos como herbicidas y los arseniatos como insecticidas).

En las aguas naturales el arsénico se encuentra normalmente como anión ácido trivalente o pentavalente. La valencia y la especie dependen de las características físico-químicas del medio, del pH, potencial redox, etc. Las especies solubles de arsenito pueden ser H_3AsO_3 , $H_2AsO_3^-$ y $HAsO_3^{2-}$ y las de arseniato H_3AsO_4 , $H_2AsO_4^-$, $HAsO_4^{2-}$ y AsO_4^{3-} donde las formas predominantes, en el intervalo de pH de 4 a 10, es H_3AsO_4 para los arsenitos y son $H_2AsO_4^-$ y $HAsO_4^{2-}$ para los arseniatos. En general, las formas reducidas se encuentran en las aguas subterráneas, donde imperan las condiciones anaerobias, y las oxidadas en las aguas superficiales, donde predominan las condiciones

aerobias. En las aguas superficiales se encuentran cantidades de arsénico entre trazas y 1,0 ppm y mayores. En las aguas subterráneas se han encontrado concentraciones mayores de 1,50 ppm

No se conoce que el arsénico sea esencial para la salud de los hombres o de los animales. Sin embargo, sí es muy conocido el peligro que encierra el arsénico para la salud humana. Los compuestos inorgánicos son más tóxicos que los orgánicos y de entre aquellos la forma trivalente más que la pentavalente.

El arsénico inorgánico es fácilmente adsorbido por el tracto gastrointestinal, pulmón y piel y se distribuye a lo largo de los tejidos y fluidos del organismo. Su eliminación se realiza por vía urinaria, heces, sudor y epitelio de la piel.

Además de la forma química en que se encuentra, la toxicidad del arsénico depende del tiempo de exposición, de la vía de penetración, etc. y se han detectado concentraciones de hasta 100 ppb de arsénico sin haberse manifestado efectos adversos para la salud. De aquí que, si consideramos que el hombre ingiere 2 litros de agua por día, si la misma contiene 20 ppb, lo ingerido es aproximadamente el 10 % del arsénico total que puede ingerir sumando el aire más los alimentos, por lo que establecer el límite de 10 ppb, que marca la ley española proporciona un buen margen de seguridad.

Las concentraciones de arsénico en la Vega son en todos los casos inferiores a 10 ppb, con un valor máximo medido de 1,6 ppb. La distribución de las concentraciones sitúa las mayores concentraciones en los bordes sureste, noroeste y en los bordes norte y sur de la zona central (figura 2.32). Igual que para el antimonio, estas concentraciones no son significativas debido a su bajísima concentración pues la mayoría de los puntos analizados están por debajo del límite de detección de medida.

2.2.3.3. Microbiología

Una de las propiedades más restrictivas del agua para su uso en consumo público es la relativa a los microorganismos presentes en ella, los cuales pueden transmitir y producir un amplio número de enfermedades. Los acuíferos, a pesar de ser medios menos vulnerables que las aguas superficiales, también están expuestos a la contaminación microbiológica; de hecho, los acuíferos libres, y en especial los kársticos y fisurados, presentan una elevada vulnerabilidad. Incluso acuíferos aislados por formaciones impermeables, que estarían protegidos de forma natural, dejan de estarlo cuando la incorrecta ejecución de sondeos los pone en comunicación con niveles de peor calidad o con el exterior.

Los microorganismos presentes en el agua constituyen un grupo extraordinariamente heterogéneo y se agrupan bajo la determinación de protistas debido a su sencilla organización. Se pueden diferenciar dos tipos: los superiores, que son eucariotas y poseen mitocondrias y un núcleo diferenciado, donde se incluyen las arqueobacterias, las

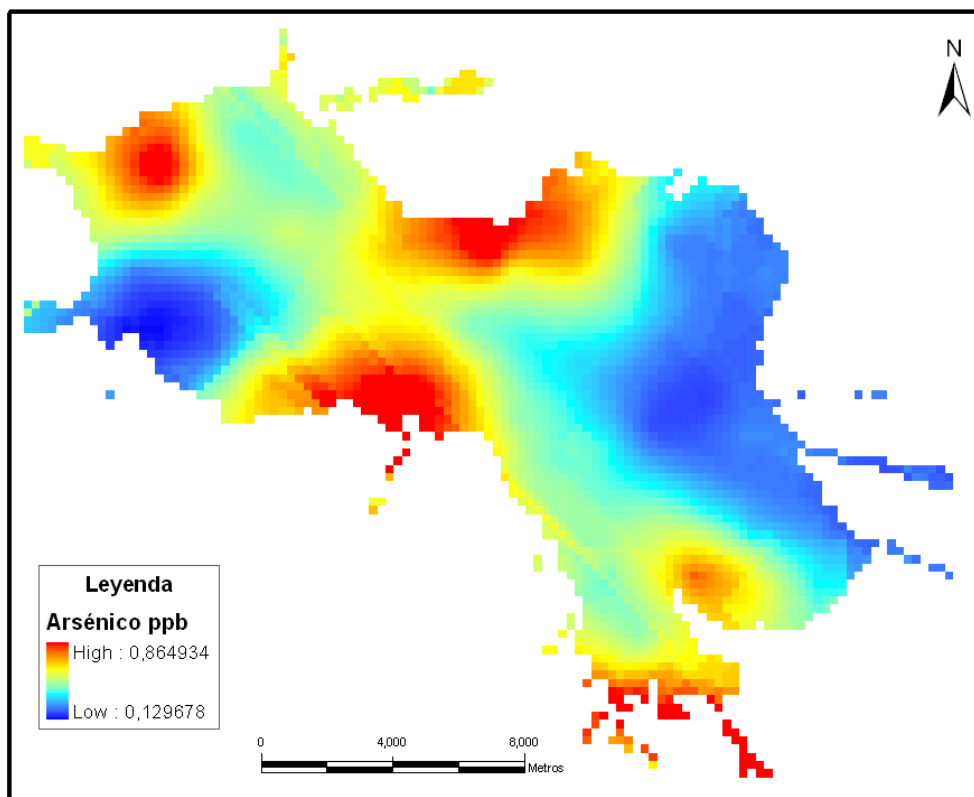


Figura 2.32: Distribución espacial del contenido en arsénico en las aguas subterráneas.

bacterias y las cianobacterias; y los inferiores, que son procariotas sin estructuras celulares aisladas por membranas y donde se incluyen los protozoos, los hongos y las algas. Dentro del grupo de microorganismos también estarían los virus, que son acelulares y se consideran la expresión más simple de vida.

Las bacterias son los microorganismos más pequeños que contienen la maquinaria necesaria para crecer y multiplicarse. La estructura interna de las bacterias es una membrana celular que regula el intercambio de nutrientes y desechos con el exterior y una pared celular que rodea a la membrana, útil en medios de baja presión osmótica que evita que estalle la célula. Sin embargo, carecen de núcleo y de otros órganos rodeados por membrana (como las mitocondrias). Las dimensiones más frecuentes de estos organismos oscilan entre 0,5 a 1 por 2 a 5 μm . Debido a este pequeño tamaño en relación con la superficie/volumen se produce un rápido intercambio de sustancias con el exterior que resulta en una gran velocidad de desarrollo.

Sus necesidades nutricionales y energéticas son muy variables y, dependiendo de éstas, se pueden agrupar en distintos grupos: fotoautótrofos (obtienen energía de la luz solar, como las clorobacterias y bacterias púrpura), quimioautótrofos (la obtienen de reacciones químicas, como las azobacterias, sulfobacterias y siderobacterias), y el grupo mayoritario que son las heterótrofas (que obtienen la energía de la materia orgánica). Los elementos necesarios para el desarrollo de los microorganismos se agrupan en dos categorías, los elementos mayoritarios (C, O, N, P, S, K, Na, Ca y Fe) y los elementos minoritarios o traza, necesarios sólo en cantidades muy pequeñas (Mn, Mg, Cu, Sn y Mo). El carbono necesario para la fabricación de metabolitos intermedios y en general para los esqueletos carbonatados se obtiene del CO_2 o de los compuestos orgánicos. El azufre, necesario para la síntesis de aminoácidos y proteínas, se obtiene de compuestos orgánicos azufrados o de sulfatos y sulfuros. El fósforo, empleado en la síntesis de fosfolípidos se obtiene de compuestos inorgánicos (fosfatos).

Dentro del grupo de las bacterias heterótrofas se encuentran los agentes causantes de enfermedad en el hombre, animales y plantas, que constituyen además la mayor parte de la población microbiana que nos rodea. En el agua, se encuentran principalmente bacterias Gramnegativas, que se caracterizan por ser capaces de vivir en medios oligotróficos.

Para esta sección, nos centraremos en el estudio de las bacterias, ya que son los microorganismos más importantes en el acuífero detrítico de la Vega, pues, debido a su tamaño, son los únicos que pueden pasar a la capa freática a través del filtro de la zona vadosa. Otros organismos superiores también pueden ser encontrados, pero en la mayoría de los casos su presencia en estas aguas no será consecuencia de una causa de origen natural, sino más bien ligada a aspectos antrópicos como el deficiente mantenimiento de los pozos.

Como la separación e identificación de todos los microorganismos de una muestra sería un proceso muy largo y costoso, se han elegido unos "indicadores" que permiten la determinación rápida y fiable de la existencia de contaminación, para luego realizar un estudio de detalle, si se precisa. Las características que debe cumplir un microorganismo para ser empleado como indicador de contaminación son las siguientes (Moreno Merino, L. 1998 [124]):

- No debe ser un poblador natural de las aguas de estudio.
- Debe ser específico del proceso contaminante.
- La técnica de análisis ha de ser sencilla, poco costosa y lo más específica posible, mediante métodos oficiales o estándares a fin de garantizar la reproductibilidad del ensayo.

Según estos criterios, un análisis bacteriológico rutinario incluye las siguientes determinaciones:

- Recuento total de bacterias aerobias.
- Coliformes totales.
- Coliformes fecales (*E. coli*).
- *Streptococcus* fecales.
- Esporas.

Según la reglamentación española, los indicadores microbiológicos de riesgo para la salud, los cuales no deben existir en ningún momento en las aguas de consumo público son: el *Escherichia coli*, el enterococo, el *Clostridium perfringens* (incluidas las esporas) y las bacterias coliformes.

Para este trabajo se han escogido tres de los cuatro indicadores microbiológicos estipulados por la reglamentación española para el estudio de las aguas de consumo público junto con la mayoría de los indicadores bacteriológicos rutinarios. El conjunto de indicadores son: las bacterias aerobias, los coliformes totales y los fecales, los *Streptococcus* fecales, los *Clostridium* sulfito reductores, el *Escherichia coli* y la demanda química de oxígeno (D.Q.O.), que se incluye en esta sección debido a la frecuente relación que guarda con la demanda bioquímica de oxígeno (D.B.O.). La demanda química de oxígeno es la cantidad química de oxígeno que se necesita para oxidar los materiales orgánicos presentes en el agua, aunque no diferencia entre la materia biodegradable y el resto y no suministra información sobre la velocidad de degradación en condiciones naturales.

Las bacterias aerobias son las que necesitan el oxígeno para su metabolismo. El número de bacterias en un medio natural concreto está condicionado por la limitación

de nutrientes; así, para aguas con mayor cantidad de nutrientes aumenta el número de bacterias. Las aguas subterráneas suelen tener un potencial redox oxidante (medio aerobio), aunque en determinadas condiciones se pueden dar procesos típicos de medios anaerobios derivados de la adherencia de las bacterias a las partículas de materia orgánica en suspensión, ya que es posible que dentro del flóculo en suspensión haya condiciones anaerobias. De este modo es posible el transporte y supervivencia de bacterias anaerobias en medios aerobios.

La distribución espacial de este tipo de bacterias en las aguas subterráneas de la Vega muestra en la figura 2.33, en la que se aprecia una regionalización bastante clara influenciada en gran parte por los distintos ríos que acceden a la Vega. Los valores máximos se sitúan cercanos a los bordes norte y sur, con especial influencia del río Salado, donde está el máximo local.

Probablemente, la regionalización de estas bacterias está influida porque tienen una gran adaptación al medio con escasos o nulos nutrientes y un alto tiempo de pervivencia (Castillo Martín, 2002 [32]), por lo que su concentración a la hora de relacionar su distribución con un posible indicador de contaminación es poco orientativa sin un adecuado análisis temporal, aunque sí nos muestra las condiciones más favorables para su reproducción.

En el resto de variables microbiológicas, su espacialización es mucho más complicada, y no existe en la mayoría de los casos una regionalización de la variable, debido al escaso tiempo de supervivencia del resto de las bacterias y a que la mayoría, afortunadamente, no suelen poder multiplicarse en las condiciones del acuífero.

Las enfermedades entéricas causadas por bacterias *coliformes* se transmiten casi exclusivamente por contaminación de origen fecal en aguas y alimentos, ya que este tipo de microorganismos se encuentra radicado en el tracto intestinal de mamíferos y animales de sangre caliente. La transmisión vía contaminación de los suministros de agua es la forma de infección porcentualmente más importante, y ha sido la causa de epidemias masivas, particularmente de fiebre tifoidea y cólera. De aquí la importancia de los análisis rutinarios para comprobación de la inexistencia de este tipo de bacterias en aguas de bebida y de consumo público.

No todas las bacterias englobadas bajo la denominación genérica de "coliformes" son patógenos para el hombre, pero si no se detectan en un análisis de agua, se garantizará que no hay ningún tipo de organismo entérico, patógeno o no, ya que al ser mucho menos resistentes que los coliformes en un medio acuático, su ausencia es un indicador suficiente que avalaría la óptima pureza bacteriológica del agua.

Los coliformes totales, que según la legislación vigente española para las aguas de consumo deberían ser inexistentes, tienen un número bastante elevado en algunos de los puntos muestreados de estas aguas subterráneas, con valores de hasta 340 UFC para

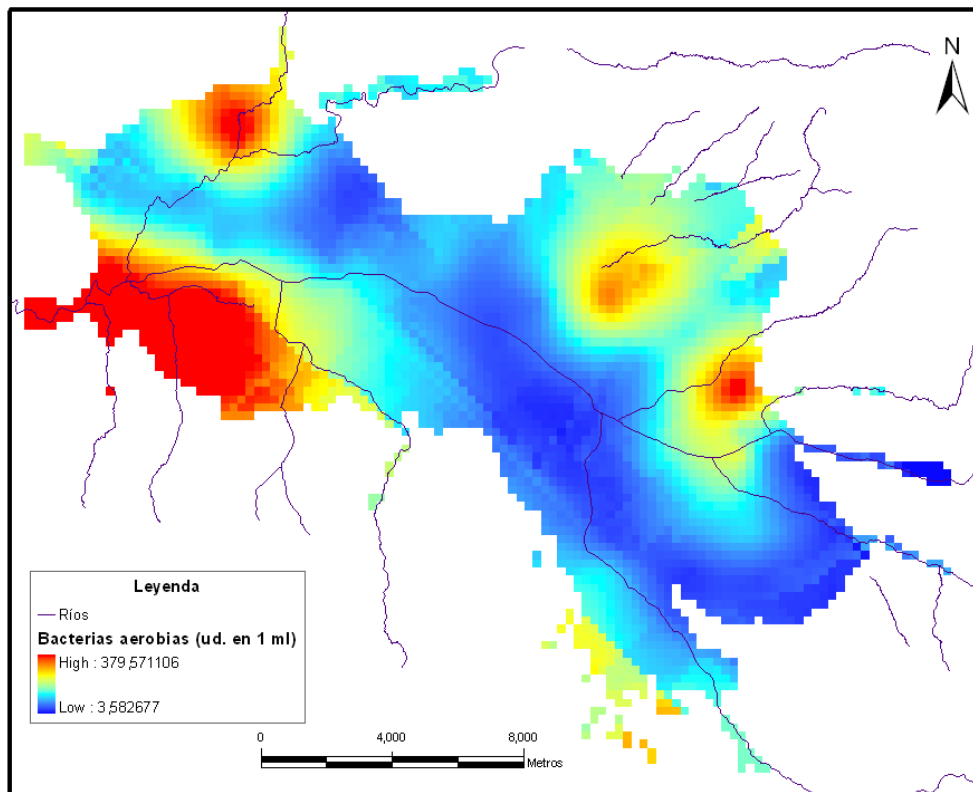
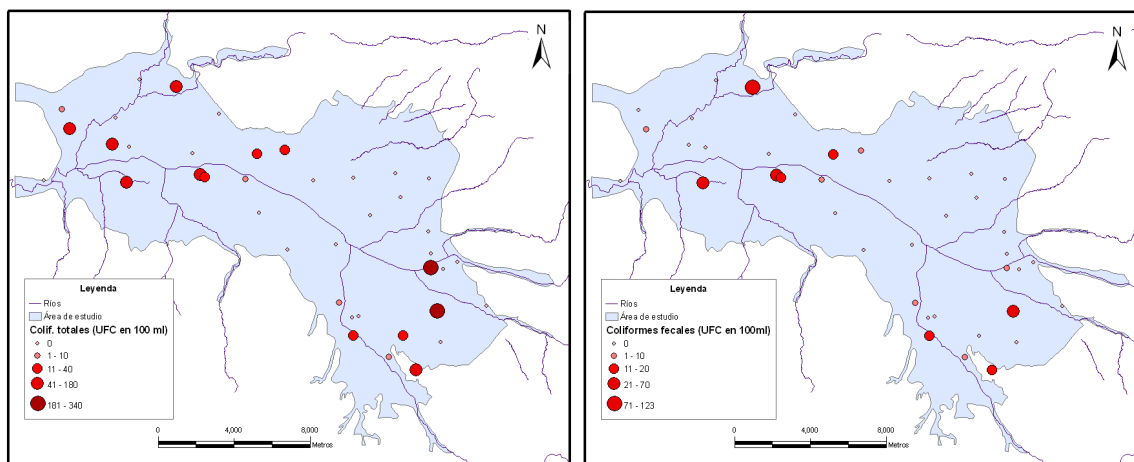


Figura 2.33: Distribución espacial del contenido en bacterias aerobias de las aguas subterráneas.



(a) Distribución espacial del contenido en coliformes totales de las aguas subterráneas. (b) Distribución espacial del contenido en coliformes fecales de las aguas subterráneas.

Figura 2.34: Distribución espacial del contenido en coliformes totales y fecales de las aguas subterráneas.

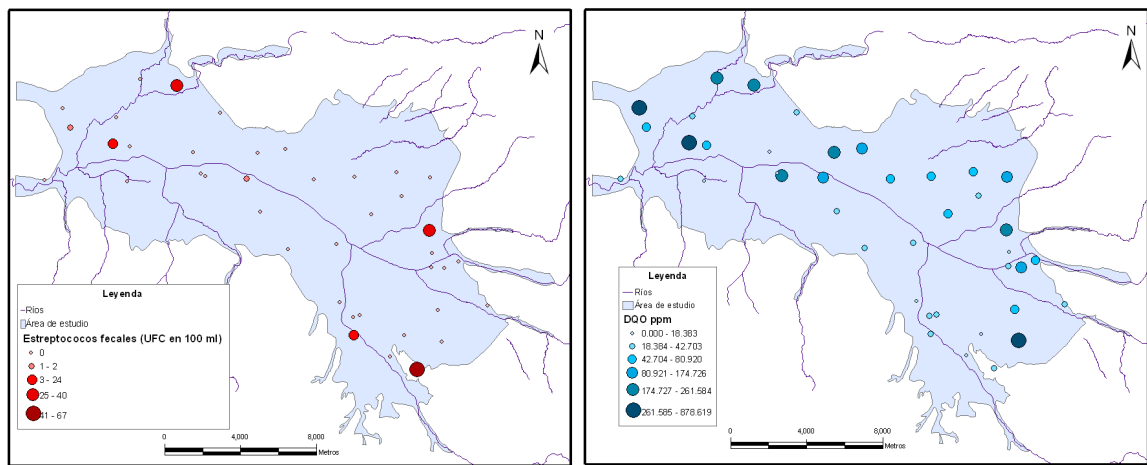
100 ml, lo cual es indicativo de contaminación (figura 2.34a). Los coliformes totales son un grupo muy heterogéneo de bacterias, que agrupa a bacilos Gram negativos, fermentativos, etc. y cuya procedencia puede ser fecal (se encuentran en el intestino del hombre y de los animales) o también proceder del suelo, polvo y agua. Por ello, si se quiere precisar si la contaminación es fecal, se recurre a los coliformes fecales y más comúnmente al *E. coli*, pues su presencia en el agua indica contaminación reciente.

La presencia de coliformes fecales también es bastante elevada en algunos de los puntos, llegando a valores de hasta 123 UFC para 100 ml, lo que además indica una gran vulnerabilidad del acuífero, ya que estos organismos, debido a su considerable tamaño, pueden ser filtrados por niveles de arcillas. (figura 2.34b)

Como era de esperar, la mayoría de los puntos estudiados que presentan contaminación por coliformes se relaciona con un origen fecal (coliformes fecales), ligado en la mayoría de los casos a pozos negros, fugas de saneamiento y en otros casos granjas de vacuno y porcino en intensivo.

El género *streptotococcus* es uno de los géneros representativos de las bacterias del ácido láctico. En concreto, se trata de anaerobios facultativos incapaces de realizar un metabolismo respiratorio, que crecen fácilmente sobre la superficie de medios sólidos expuestos al aire, pero requieren para su ciclo vital de complejos factores de crecimiento integrados por varias vitaminas y aminoácidos.

Las bacterias lácticas suelen ser muy tolerantes a los ácidos, a consecuencia de su modo de producción de energía. Esta circunstancia permite su enriquecimiento en medios azucarados que progresivamente se acidifican por incremento de la concentración



(a) Distribución del contenido en estreptococos fecales de las aguas subterráneas. (b) Distribución de la Demanda Química de Oxígeno de las aguas subterráneas.

Figura 2.35: Distribución espacial del contenido en estreptococos fecales y D.Q.O. de las aguas subterráneas.

de ácido láctico, así como por el subsiguiente uso industrial de las bacterias del ácido láctico en la producción de alimentos como yogures, queso, vinos, cervezas, etc.

Además, las bacterias de este género pueden vivir en el tracto intestinal del hombre y otros animales, en la nasofaringe y la vagina: de ahí su potencial uso como indicadores fecales bajo la denominación no taxonómica de estreptococos fecales.

En el orden práctico, y referidos a aguas, el orden de mortalidad de organismos indicadores fecales típicos es de Coliformes totales, Coliformes fecales, *Estreptococos* y *Clostridium*. De este modo, si un agua no tiene Coliformes y sí *Estreptococos* suele inferirse que se trata de una contaminación antigua o lejana. Además, los *Estreptococos* son menos resistentes que los Coliformes fecales a un ambiente hostil, ya que mueren rápidamente en aguas moderadamente salinas. Finalmente, los *Clostridium* son los más resistentes del grupo, por su capacidad de enquistamiento.

Para algunos autores existe una relación entre el valor del cociente "*Coliformes fecales*"/"*Estreptococos fecales*" para determinar aproximadamente la procedencia de la contaminación fecal en un agua: cocientes superiores a 4 indicarían contaminación de carácter humano, mientras cocientes inferiores a 1 parecerían sugerir contaminaciones de fuentes no humanas (Marín Galvín, 2003 [116]). No obstante, recientemente se ha sugerido que la anterior suposición puede no ser cierta siempre, puesto que los índices de supervivencia de las distintas especies de *estreptococcus* difieren drásticamente.

Según esta relación, para los datos analizados, el cociente es siempre inferior a cuatro para los puntos donde existen *Estreptococos*, aunque ya se ha mencionado que hay que tener en cuenta la especie de *estreptococo*. Para las aguas del área de estudio,

en la mayoría de los casos donde hay presencia de *Estreptococos* no hay presencia de *Coliformes fecales*, por lo que se intuye que se trata de un ambiente hostil para los *Estreptococos* o aguas moderadamente salinas, como es el caso. La evidencia de una contaminación antigua o lejana es poco habitual, como se puede observar en la figura 2.35a, ya que son sólo dos del total de los puntos los que presentan *estreptococos* y no presentan *coliformes fecales*: uno situado en Granada capital y otro al oeste de Fuente Vaqueros.

La demanda de oxígeno de un agua es la cantidad de oxígeno en ppm que es consumido por las sustancias contaminantes (inorgánicas u orgánicas) presentes durante un cierto tiempo. Concretamente, la *demanda química de oxígeno* (D.Q.O.) es la cantidad de oxígeno consumido por los elementos reductores en agua, mediante oxidantes químicos, sin intervención de organismos vivos. Se ideó a fin de soslayar la dificultad que presenta la determinación de la DBO.

La medida de esta magnitud se ha utilizado como un procedimiento para determinar la cantidad total de materia orgánica. Sin embargo, a partir de este dato es muy difícil obtener un valor cuantitativo del contenido de materia orgánica, debido a que las sustancias presentes tienen diverso grado de oxidabilidad. Por otra parte, esta técnica se estudió para aguas negras, que no tienen reductores inorgánicos y no es de aplicación en el caso de pequeños valores de materia orgánica o en presencia de reductores inorgánicos, entre ellos la de cantidades apreciables de cloruros, ya que éstos son reductores.

En cualquier caso, salvando esta problemática inicial, se puede observar que la D.Q.O. toma valores más bajos en la franja central del área de estudio y valores más altos en los bordes este y oeste (figura 2.35b), curiosamente coincidiendo en el borde oeste con la presencia de bacterias *coliformes* y no así en el oeste, por motivos que habría que estudiar más en detalle.

Los *clostridios*, muestreados también en estas aguas, son inexistentes en las aguas subterráneas, ya que aunque están presentes en las heces fecales, su número es apreciablemente menor que los correspondientes a *coliformes* o *estreptococos fecales*. Las esporas de estas bacterias pueden resistir en el agua mayor tiempo que los *coliformes*, pues son además más resistentes a la desinfección que aquéllos; sin embargo, la filtración es un método eficaz para su eliminación, causa también por la cual no se encuentran en el acuífero.

2.3. Climatología y datos foronómicos

El clima en la Vega de Granada es mediterráneo continentalizado debido a la baja influencia del mar, fundamentalmente por el aislamiento producido por la orografía y también por su altitud (valor medio de 670 m). La combinación de estos elementos

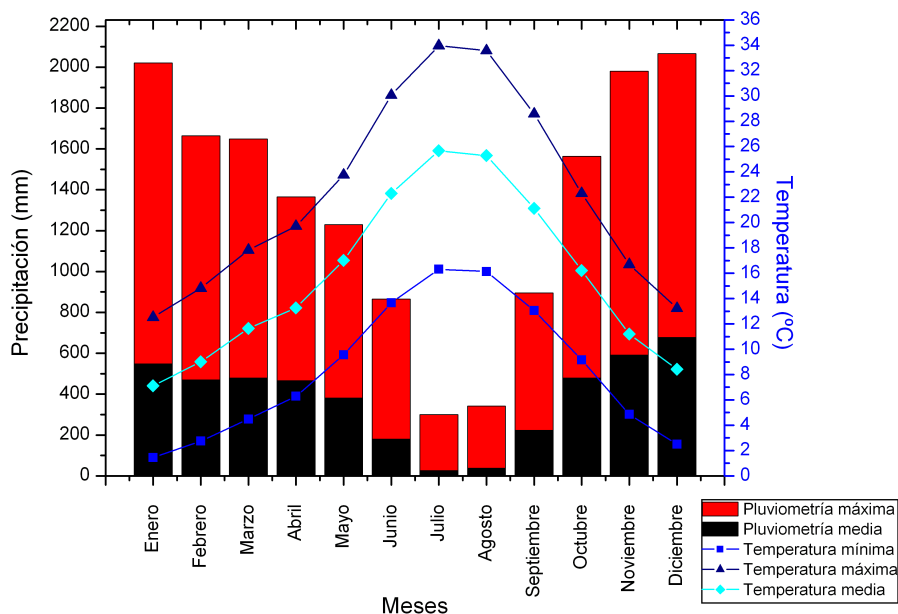


Figura 2.36: Climograma de la Vega de Granada.

favorece la aparición de un clima de características propias o microclima, cuyos aspectos más relevantes son los suelos con altos grado de humedad, una influencia suave del clima mediterráneo, gradiente térmico elevado y gran duración de las estaciones extremas (invierno y verano). Otra característica propia del área es la escasez de precipitaciones y la concentración de las mismas en los meses más fríos con una fuerte sequía estival. Estos aspectos se evidencian en el climograma medio del total de las estaciones de la Vega de Granada estudiadas (figura 2.36), donde se representan las precipitaciones medias y máximas y las temperaturas medias, máximas y mínimas.

Se ha dispuesto de 34 estaciones, 26 pluviométricas y 8 termopluviométricas existentes en la Vega (figura 2.37), pertenecientes al Instituto Nacional de Meteorología. El registro total de años depende de la estación; en el caso de la precipitación el registro más antiguo es de 1920 en la estación de La Cartuja y los datos en la mayoría de las estaciones terminan en diciembre del 2001. En lo que se refiere a las temperaturas, el dato más antiguo pertenece también a esta estación, con registros desde el año 1902.

La distribución espacial de las precipitaciones medias mensuales (figura 2.38) muestra que en los meses más secos (junio), la precipitación se concentra fundamentalmente en el área norte, y en los meses más húmedos (diciembre y octubre) la precipitación se concentra al noreste.

El conocimiento de la distribución de las pluviometrías máximas es también una información interesante, ya que posibles problemas ligados a la erosión pueden ser la consecuencia de las lluvias torrenciales (precipitaciones máximas) del área de estudio.

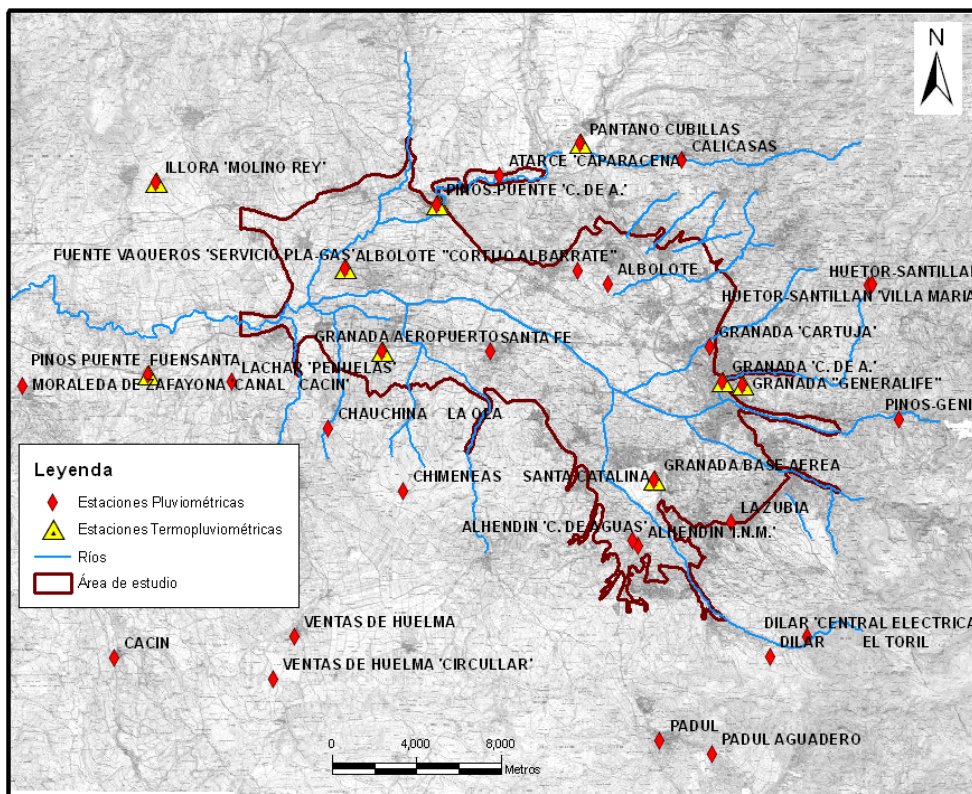


Figura 2.37: Distribución espacial de las estaciones termopluviométricas en la Vega de Granada.

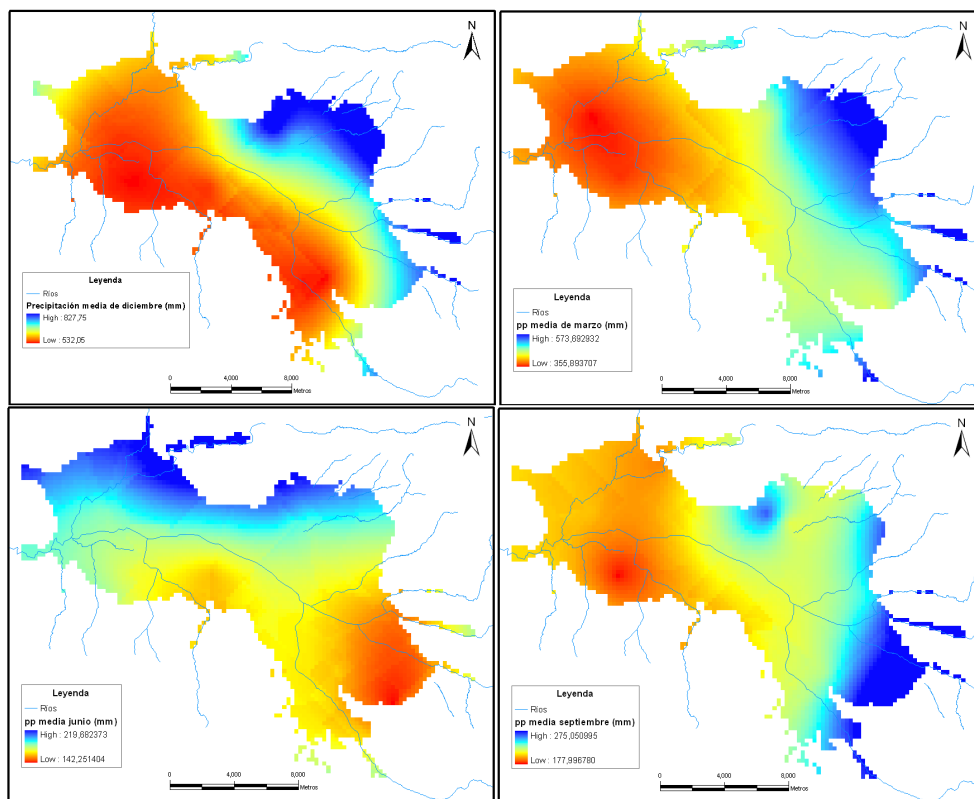


Figura 2.38: Precipitaciones medias mensuales para la Vega de Granada.

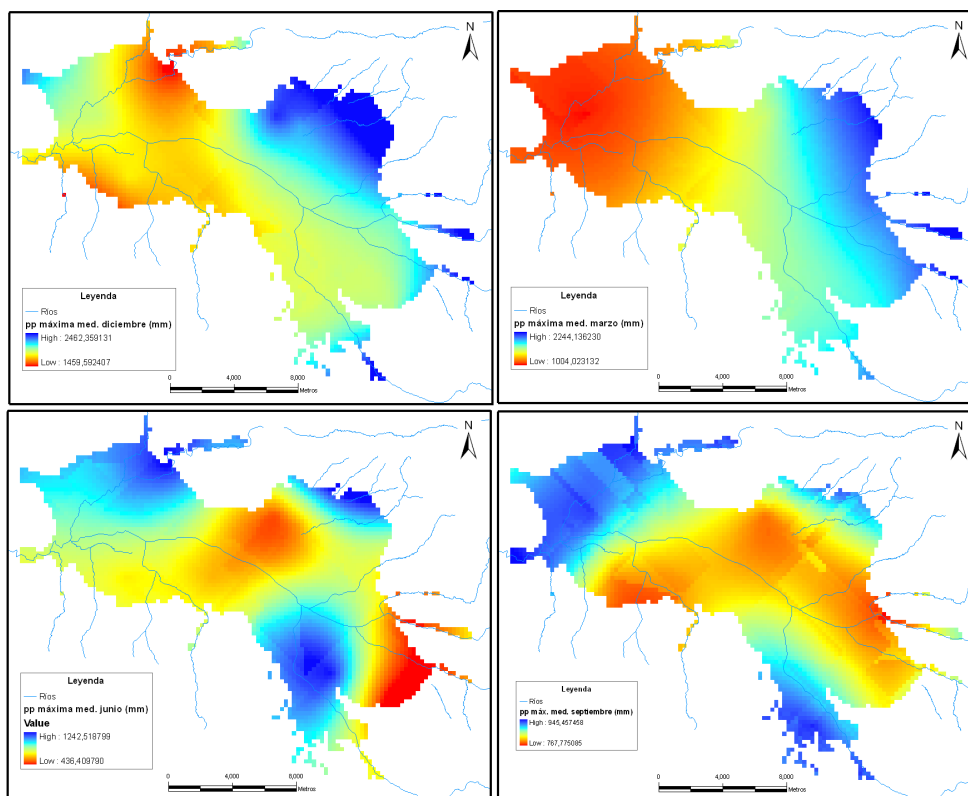


Figura 2.39: Precipitaciones máximas medias en la Vega de Granada.

En el mapa de la figura 2.39 se representa la media mensual de las precipitaciones máximas en el área de estudio, con mayor torrencialidad general en los bordes, donde se sitúan los principales relieves, y una menor torrencialidad en la zona central.

Ya se ha comentado anteriormente la importancia de la precipitación en el balance hídrico, ya que es después de la escorrentía superficial la principal aportación de agua al acuífero. Una representación de los caudales aportados al acuífero a través de los cauces superficiales frente a los caudales drenados hacia la red superficial, se puede observar a partir de los aforos realizados por la Confederación Hidrográfica del Guadalquivir, representados en la figura 2.40. Según Castillo (1986, [30]), las aportaciones medias para los distintos cauces de superficie y canales de riego son los representados en la tabla 2.3, donde la última columna representa el caudal total de $\text{hm}^3/\text{año}$ que se infiltra al acuífero.

La importancia de la recarga del acuífero proveniente de Sierra Nevada se muestra con la representación de la precipitación y el caudal (gráfico 2.41) a partir de un piezómetro situado al norte de la base aérea de Armilla. La precipitación tomada en esa misma estación y el aforo del canal de la Espartera, refleja en la evolución piezométrica una coincidencia de picos entre la precipitación, el caudal y la recarga subterránea. Un

Ríos y Canales	Caudal medio l/s	Aporte anual $h^3/año$	% Infiltrado	Aporte al acuífero $h^3/año$
Genil	2584	81	55	44.55
Aguas blancas	1427	45	35	15.75
Monachil	946	30	35	10.5
Dílar	885	28	35	9.8
Darro	500	16	35	5.6
Velillos	1328	42	15	6.3
Colomera	1015	32	15	4.8
Cubillas	1903	60	35	21
C. de Cacín	-	29	25	7.25
C. de Albolote	-	20	25	5
Resto	-	40	25	10
Total		423		140.55

Tabla 2.3: Relación de caudales y aportaciones medias de cauces y canales.
Fuente: Antonio Castillo Martín. 1986

análisis más detallado, en el que he participado para estudiar la respuesta de los piezómetros a los fenómenos meteorológicos para el área de estudio ha sido realizado según los métodos de Luque-Espinar y otros (2005, 2007 y 2007, [112, 111, 113]), a partir de varios análisis espectrales se detecta la existencia de hasta 5 ciclos distintos de 22, 11, 3, 2, 1 y 0,5 años. El detalle del análisis de los picos, junto con la representación de algunas series se muestra en las figuras 2.42 y 2.43. Por otra parte, los ciclos de 22 y 11 años cabe relacionarlos con ciclos solares; sin embargo, la mayoría de los autores consideran que el ciclo de 11 años, junto con el de 3 años, se deben al fenómeno conocido como NAO (North Atlantic Oscillation), que marca el comportamiento climático de las latitudes en las que nos encontramos. No obstante, algunos autores creen que el fenómeno conocido como ENSO (El Niño South Oscillation) también puede influir en esta parte del planeta (Rodo, 1997 y Pozo-Vázquez 2000, [153, 146]). El ciclo de un año de duración cabe relacionarlo con el ciclo hidrológico. Por último, el ciclo de medio año también tiene un claro reflejo en alguna series, aunque no es razonable que también esté relacionado con NAO, como ya se ha apuntado antes.

2.4. Aspectos edafológicos

El estudio de las propiedades edafológicas de los suelos de la Vega de Granada constituye un aspecto importante de este trabajo, debido al papel que juega la zona no saturada en el acuífero. Las características de la zona no saturada, junto con el tipo e intensidad de las actividades socioeconómicas desarrolladas en superficie, son factores que afectan a la calidad de las aguas subterráneas. La excelente calidad agrícola

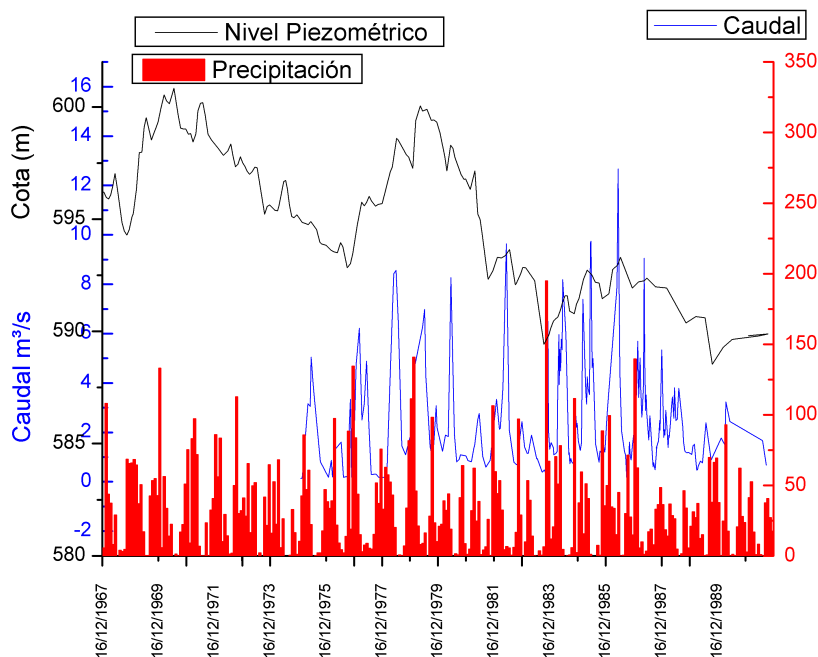


Figura 2.41: Evolución piezométrica (cota sobre el nivel del mar), foronómica y precipitación mensual en la zona SE de la Vega de Granada.

de los suelos de la Vega de Granada ha propiciado el cultivo intensivo de diferentes labores, donde las mayoritarias son: maíz, tabaco, trigo y otros cereales, alamedas y, en menor proporción, frutales. Con el fin de conseguir altos rendimientos, se adicionan con frecuencia cantidades de abonos y plaguicidas, de forma poco apropiada, y se riega con aguas residuales, procedentes tanto de las fábricas asentadas en la Vega, como de los pueblos y del núcleo urbano de Granada capital. Por todo ello, no es difícil pensar que el suelo, soporte de la vegetación, podría alcanzar un alto grado de contaminación debido a la presencia excesiva de abonos (nitratos, sulfatos, etc.), metales pesados o plaguicidas, los cuales se filtran, con distintas velocidades, a través de la zona no saturada hasta el acuífero.

Los suelos de la Vega de Granada son jóvenes, o poco desarrollados, debido a que están en continua renovación por el aporte de material procedente de los relieves limítrofes, transportado, fundamentalmente, por el río Genil y sus afluentes. En general son suelos donde existe un horizonte de acumulación de carbonatos y con distintas secuencias más o menos generales según el grado de evolución. Según esta evolución temporal podríamos encontrar tres tipos de suelos atendiendo a su subdivisión en horizontes.

El menos evolucionado es el equivalente a un fluvisol calcáreo, donde se distingue: un horizonte superficial Ap, correspondiente al horizonte de labrado, de textura poco uniforme, producido por la adición continuada de estiércoles muy terrosos o compost; un

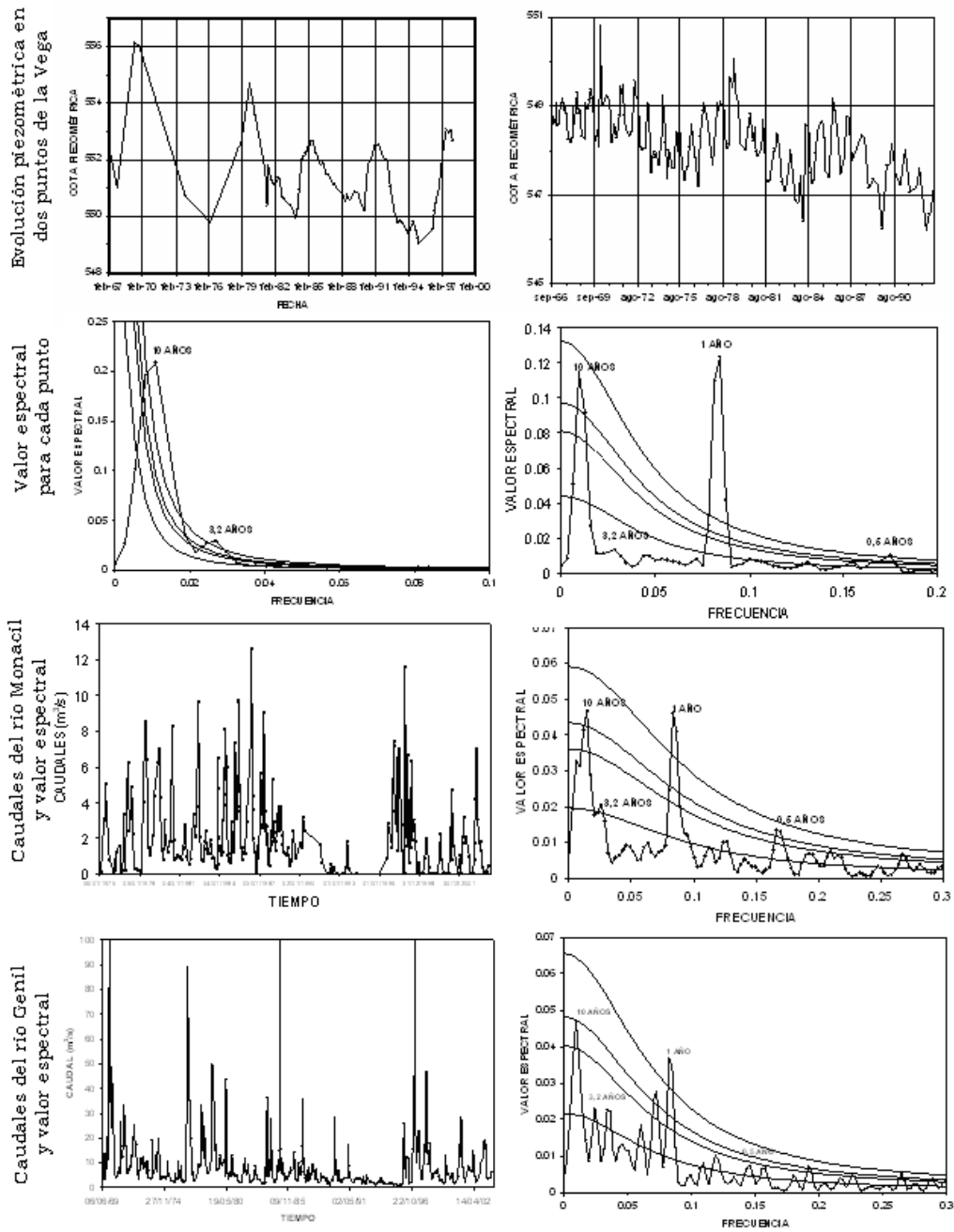


Figura 2.42: Análisis espectrales de algunas estaciones de la Vega de Granada.

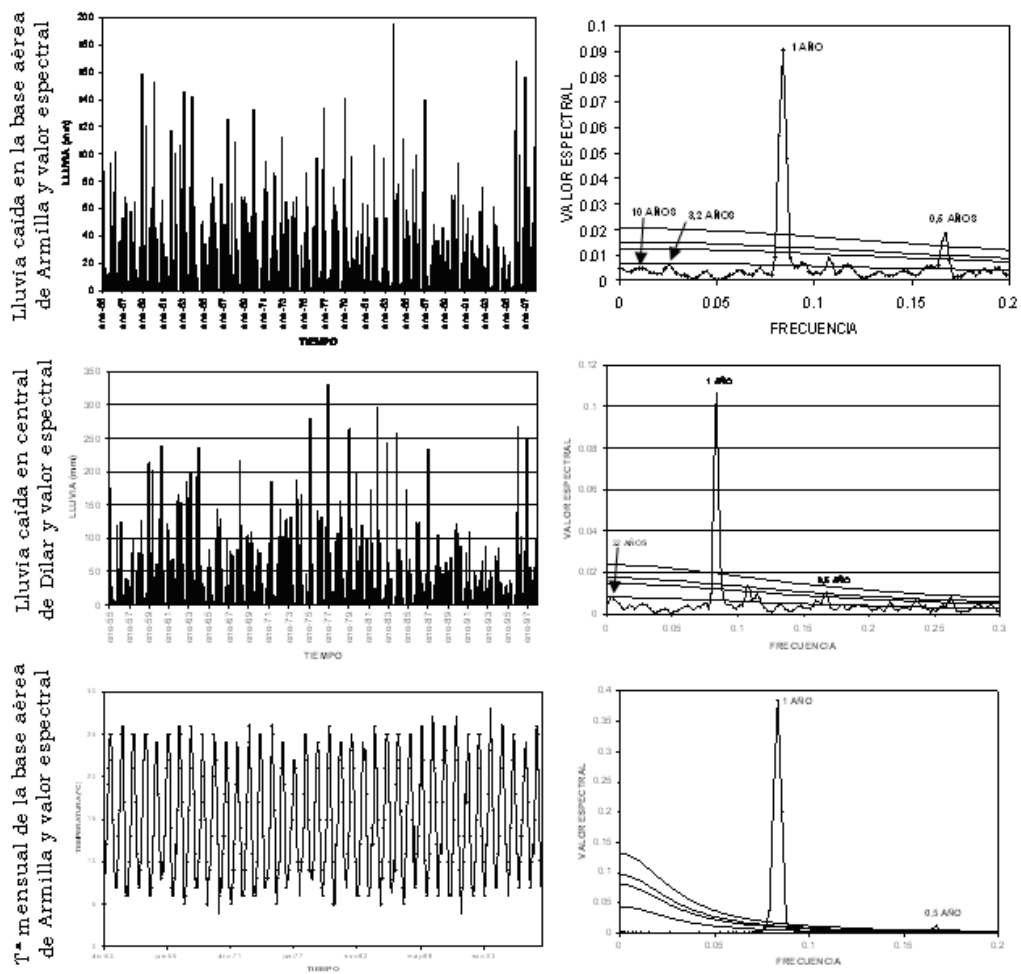


Figura 2.43: Análisis espectrales de algunas estaciones de la Vega de Granada.

horizonte C1, de arena suelta y estructura masiva; y un horizonte C2, menos evolucionado. Con un grado de evolución intermedia se encuentra un suelo con el mismo horizonte Ap anterior; un horizonte Bw más estructurado y con más arcilla; y un horizonte C masivo. Este tipo de suelo se denomina Vega Parda y corresponde a un cambisol. Y, por último, el suelo más evolucionado, con 5 horizontes, que correspondería a la Vega Roja, que es equivalente a los cambisoles y los luvisoles. En estos suelos se pueden distinguir: el horizonte de cultivo superficial Ap; un horizonte Bw o Bt de alteración o iluviación, de colores rojizos debido a la alteración de hematites (sin H_2O) o de goethitas (hidratado); un horizonte Ck, de acumulación de carbonatos (nódulos calcáreos); un horizonte Bt, de iluviación que recoge el material procedente del lavado anterior; y, finalmente, un horizonte Cmk, de consistencia pétreo (horizonte petrocálcico), de recristalización de un horizonte cálcico. Algunos ejemplos de los suelos recogidos durante la campaña de campo se pueden ver en la figura 2.44, correspondiente a: un suelo poco evolucionado (a), un suelo representativo de Vega Parda, con alto contenido en arcilla (b), un suelo más evolucionado (c) de colores rojizos (Vega Roja) y un suelo con el característico horizonte Ck de acumulación de carbonatos (d).

Una representación general de la distribución de los suelos en la Vega es la realizada por el proyecto Lucdeme (figura 2.45), que diferencia fundamentalmente dos tipos de suelos: los regosoles y los calcisoles. Este mapa presenta una línea de división de este a oeste que corresponde al límite entre las hojas 1009 y 1026 de los mapas originales que no han sido unificados. Estos suelos (regosoles calcáricos y cambisoles cálcicos) corresponden a suelos poco y medianamente evolucionados, que no muestran una manifiesta horizonación. Conforme aumenta la distancia al aluvial de la Vega aparecen los calcisoles áplicos, suelos desarrollados sobre materiales calizos en zonas de mayor pendiente y con buen drenaje; también pueden aparecer en aquellas zonas donde se ha destruido el horizonte petrocálcico como consecuencia del arado intensivo.

En las zonas de borde, donde ya no existe la influencia de material aluvial, los suelos predominantes son mucho más someros, limitados en profundidad por la roca madre, por este motivo abundan los leptosoles líticos, calcisoles pétricos y calcisoles háplicos, protegidos de la erosión por costras calcáreas endurecidas o con restos de esta costra destruida. Por último, en el sureste, coincidiendo con la entrada de los ríos, los suelos son más arenosos: predominan los litosoles y los arenosoles.

En este contexto, cabe destacar que el estudio integral de la calidad de las aguas requiere datos adicionales que correlacionen los datos hidroquímicos analizados con los procesos que tienen lugar en la superficie. El estudio de los procesos o cambios que se producen en la calidad del agua desde que se infiltra hasta que llega a la zona saturada se puede abordar de diferentes maneras, pero, sin duda, uno de los factores más



Figura 2.44: Ejemplos de los tipos de suelos de la Vega de Granada: (a) poco evolucionados, (b) medianamente evolucionados (c) muy evolucionados, y (d) y con horizonte Ck

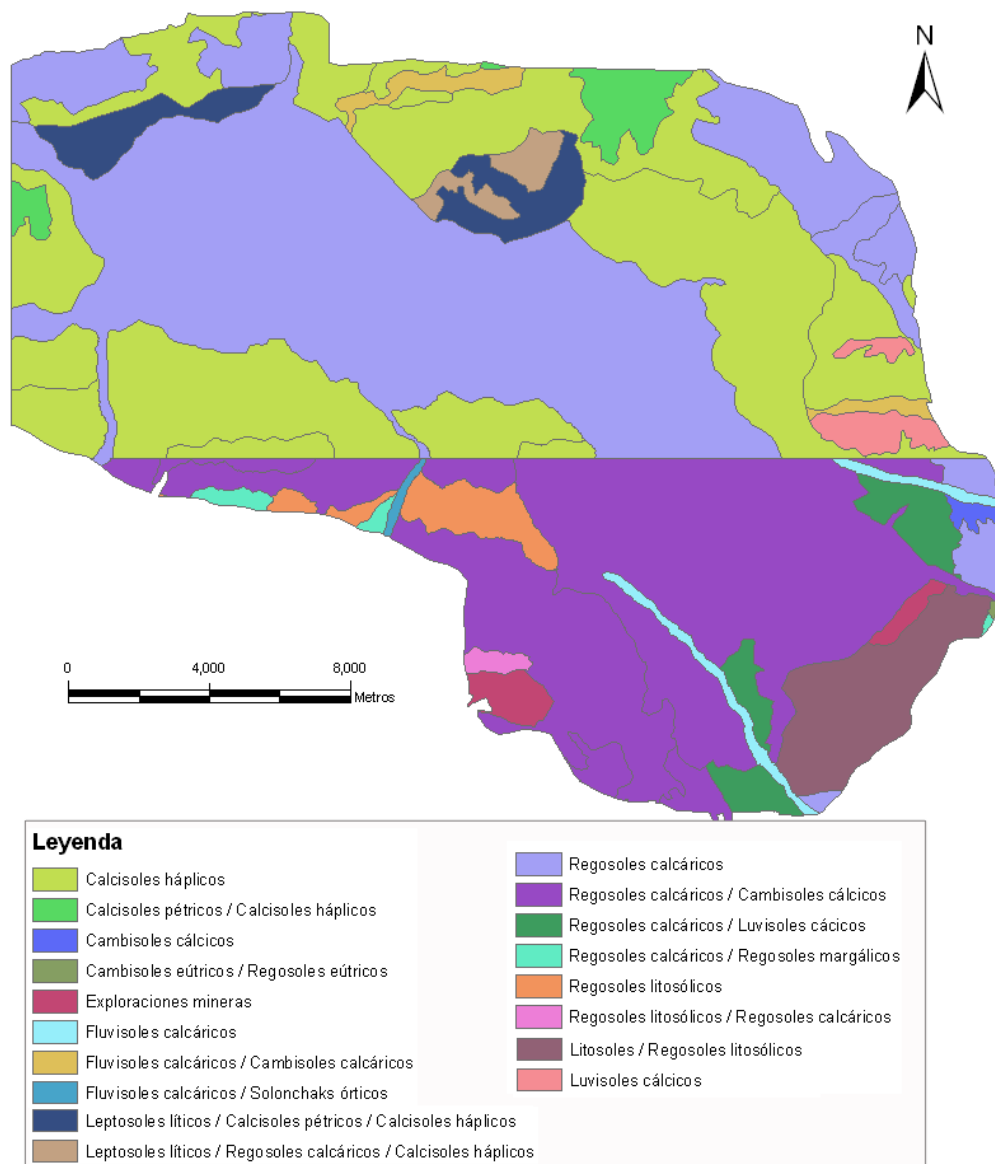


Figura 2.45: Mapa de suelos de la Vega de Granada (LUCDEME).

Nota: La división horizontal del mapa corresponde a la división de las hojas 1009 y 1026 que no han sido unificadas.



Figura 2.46: Fotografía de la preparación de las muestras de suelos de la Vega de Granada (estufa de secado y tamiz de suelos).

importantes a la hora de correlacionar los datos hidroquímicos del agua, es el relativo a los datos del suelo.

Conviene aclarar que la campaña de recogida de muestras de suelo se preparó simultáneamente a la campaña de metales, en marzo del 2004. El total de muestras es igualmente 76, coincidiendo en la mayoría de los casos con la localización de los puntos donde se analizan los metales pesados. La muestra de suelo se recogió lo más cercana posible a cada uno de los pozos estudiados. El procedimiento de muestreo consistió en la toma de una columna de un metro de profundidad, separando cada uno de los horizontes existentes. Para este muestreo, también se tuvo en cuenta la cobertura del suelo, el color de los horizontes o la cantidad de gravas. Todos estos datos están incluidos igualmente en la base de datos de campo, junto con una fotografía de los horizontes.

El análisis de los suelos se realizó en la Estación Experimental del Zaidín, bajo la supervisión del profesor D. Enrique Barahona. Una vez separados los horizontes, tamizados y secados en estufa (figura 2.46), el total de muestras asciende a 300. De cada una de ellas se ha medido la humedad, el pH, el carbono orgánico, la textura (% arcillas, % arenas y % limos) y la humedad de saturación. Además de estas variables, del último horizonte 80-100 cm, se realizó un extracto de saturación. En este último horizonte se analizaron: conductividad, flúor, cloruros, dióxido de cloro, nitratos, nitritos, fosfatos, sulfatos, magnesio, calcio, sodio, potasio, plomo, cobre, hierro, manganeso, zinc, cadmio y cromo. Estas determinaciones fueron realizadas en el Centro de Instrumentación Científica de la Universidad de Granada, donde los aniones de componentes mayoritarios se analizaron por cromatografía de gases, y los cationes y metales por espectroscopía de radiación atómica (figura 2.47).

La distribución de estas muestras, como ya se ha comentado anteriormente, estuvo condicionada a los puntos de muestreo de aguas subterráneas, tratando de disponer de una muestra de suelo en cada celda de 2 km de lado (figura 2.48).



Figura 2.47: Fotografía de los análisis de las muestras en la Estación Experimental del Zaidín. Determinación de las arcillas y batidora de las muestras.

Los resultados de estos análisis junto con su estimación espacial por medio de geostatística para elaborar los mapas temáticos de distribución espacial, completan la información recopilada sobre las propiedades y características de los suelos. La descripción de estas variables comienza con la textura, parámetro básico para entender el funcionamiento del suelo como sistema dinámico y abierto. La textura comprende la fase sólida del suelo y determina su porosidad. En los huecos que deja la estructura formada por la parte sólida, se encuentra la fase líquida, agua principalmente, que contiene iones y sustancias en disolución, y la fase gaseosa o atmósfera del suelo.

2.4.1. Textura

En el suelo, las interacciones entre la fase sólida y líquida adquieren especial relevancia, al existir en la fase sólida partículas cargadas eléctricamente y de elevada superficie específica. Los procesos de adsorción e intercambio iónico son fundamentales para la actuación del suelo como depurador natural, para la vida en el suelo y en concreto para la nutrición de las plantas. Las fases líquida y gaseosa se hallan en cambio constante, tanto en cantidad como en composición. La lluvia, el riego y las filtraciones incorporan agua al sistema y la evapotranspiración constituye la principal salida al margen que puede representar la percolación hacia la zona saturada.

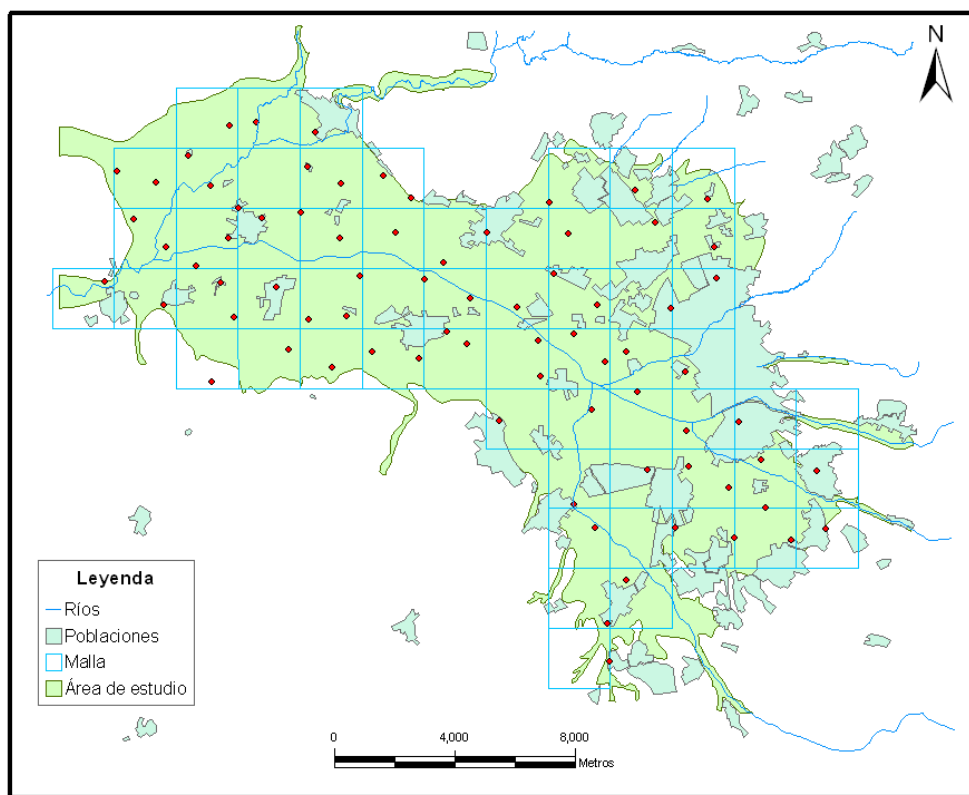


Figura 2.48: Distribución espacial del muestreo de suelos en la Vega de Granada.

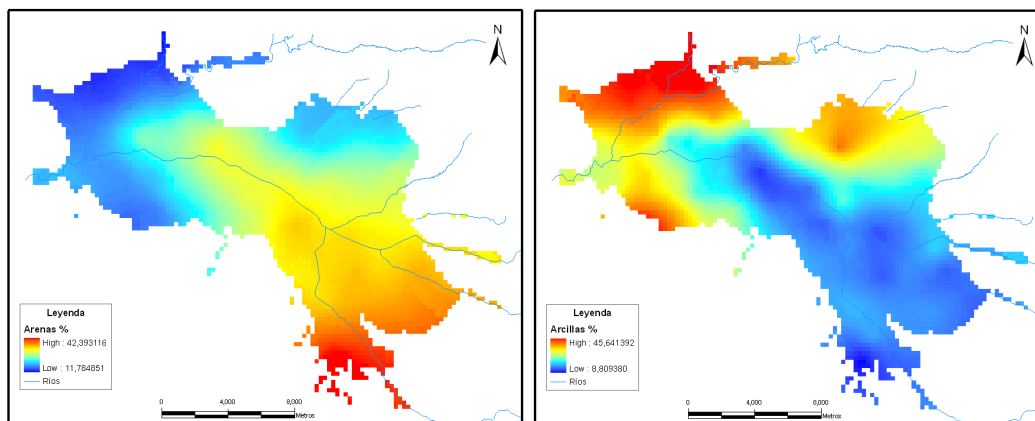


Figura 2.49: Distribución espacial del % arenas y % arcillas en los suelos de la Vega de Granada.

La **textura** o granulometría expresa las proporciones relativas de las distintas partículas minerales inferiores a 2 mm, agrupadas por clases de tamaños, por lo que supone una útil expresión sintética de las características de cada horizonte. El interés de conocer la granulometría reside en que permite inferir otras propiedades y características directamente relacionadas con el uso y comportamiento del suelo, además de ser una de las características más estables a lo largo del tiempo.

El porcentaje de cada una de estas partículas (arenas, arcillas y limos, según la clasificación USDA simple) corresponde a la media de los horizontes del primer metro de suelo (figuras 2.49 y 2.50). En estos mapas se comprueba que existe una relación inversa en la distribución de arenas y arcillas: así el mayor contenido en arenas y mínimo de arcillas se sitúa a la entrada de la zona de los ríos Dilar y Monachil; la proporción de arenas disminuye, en consecuencia, y la de arcillas aumenta, con la proximidad a la zona de descarga y a los bordes. Esta zonación es coherente con la evolución de los suelos, de esta manera, los suelos más evolucionados son los que tienen mayor contenido en arcillas (borde norte y oeste del acuífero) y suelos menos evolucionados, los de mayor contenido en arenas, y por tanto mayor capacidad de drenaje.

Por otra parte, el porcentaje de limos es mayor en algunos cauces de ríos, como por ejemplo entre el río Genil y el río Salado, y en el entorno de la confluencia de los ríos Genil, Monachil y Beiro. Esta localización de los máximos seguramente está asociada al depósito de los limos en llanuras de inundación.

El análisis granulométrico puede expresarse de una manera más sintética según el triángulo textural propuesto por el USDA (figura 2.52). La clasificación de los porcentajes de arena, arcilla y limo en clases texturales permite la identificación de áreas con análogo comportamiento frente al agua y frente a la respuesta de las plantas. En esta clasificación (figura 2.51) se observa que la mayor parte de la superficie de la Vega

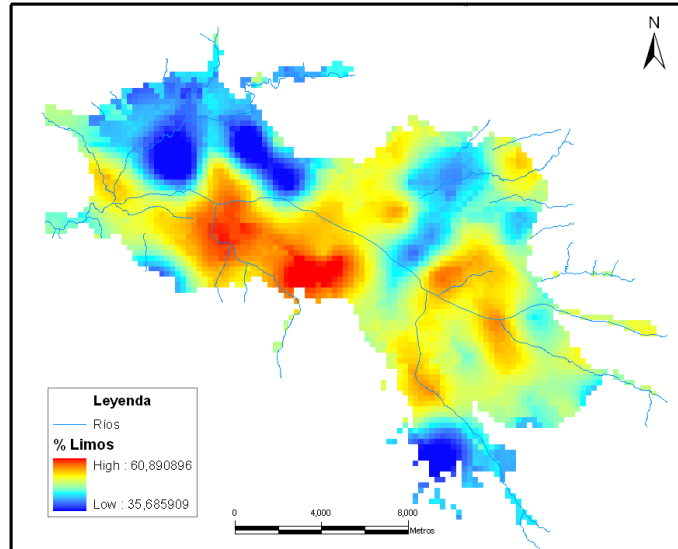


Figura 2.50: Mapa del % de limos de los suelos de la Vega de Granada.

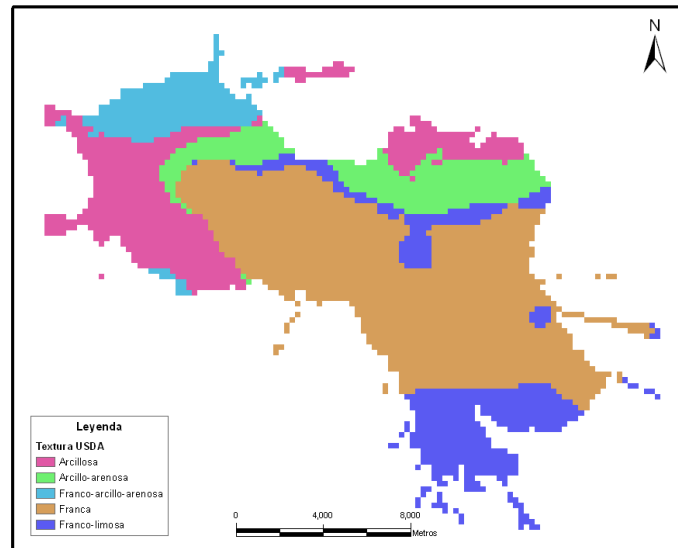


Figura 2.51: Mapa de la textura de los suelos de la Vega de Granada.

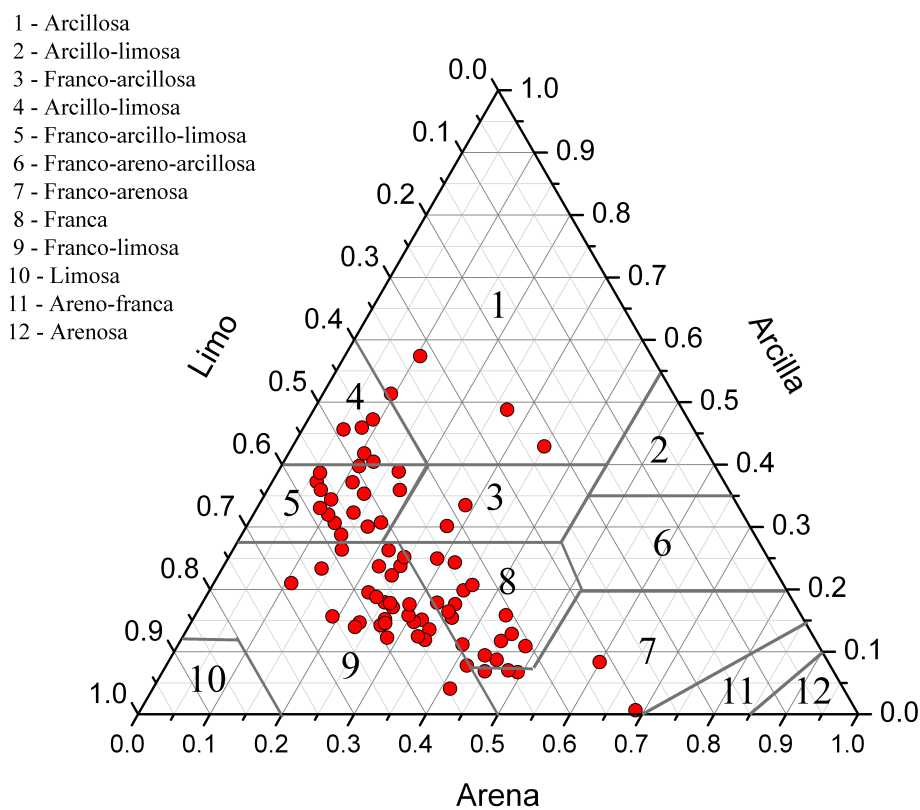
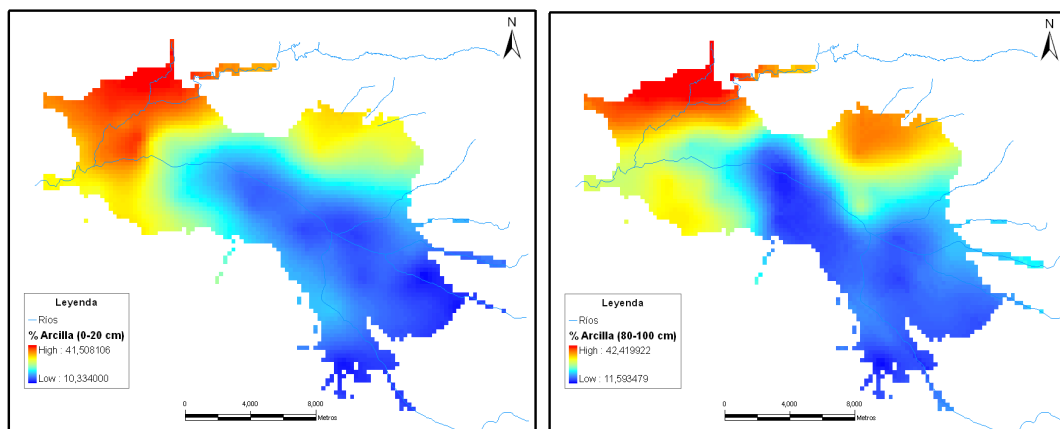


Figura 2.52: Triángulo textural USDA y representación de los suelos analizados en la Vega de Granada.

corresponde a una textura franco limosa a lo largo del río Genil, que cambia con la proximidad al borde oeste hasta una textura franco-arcillo-limosa y arcillo-limosa debido al aumento de arcillas y a franca o equilibrada en la zona de recarga con el aumento de arenas.

La distribución espacial de la textura del suelo en profundidad (horizontes superior e inferior) (figuras 2.53, 2.54 y 2.55) es muy similar en las arcillas y con algunas diferencias en las arenas. El porcentaje de limos, sin embargo, presenta mayores diferencias en profundidad. El máximo % limos en el horizonte superficial se sitúa al noroeste (en la zona de Fuente Vaqueros) y en el horizonte inferior a la altura de Santa Fé. Estas diferencias pueden ser debidas que los limos se encuentran en depósitos de llanuras de inundación, las cuales han cambiado de situación a lo largo de los años. Concretamente, la zona de máximo para el primer horizonte, situada al noroeste de la Vega, podría relacionarse con una zona inundable que desecaron los musulmanes entre los años 711 y 713 y que se desecó en 1940 cuando pasó a ser propiedad del duque de Wellington. Por otra parte, el máximo situado en la zona del río Salado correspondiente al horizonte



(a) Distribución espacial del% de arcillas en el horizonte superior (0-20 cm) de los suelos de la Vega de Granada. (b) Distribución espacial del% de arcillas en el horizonte inferior (80-100 cm) de los suelos de la Vega de Granada.

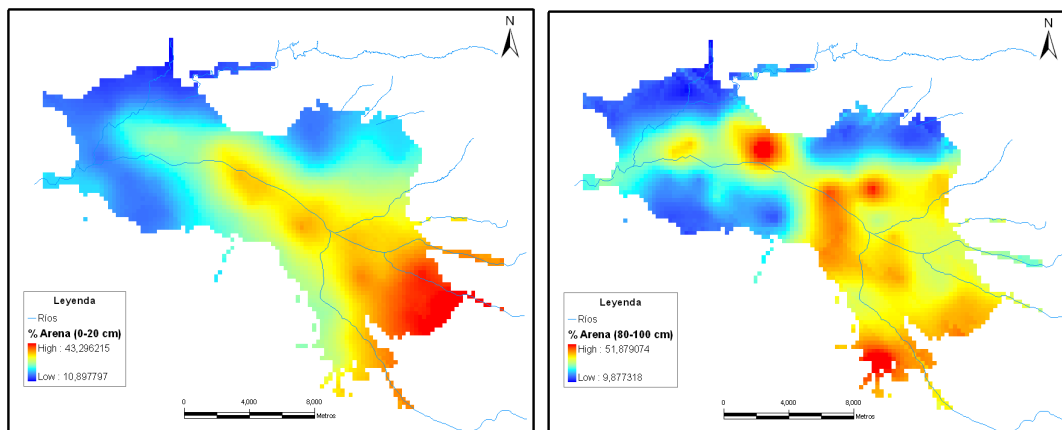
Figura 2.53: Mapas del % de arcillas en el horizonte superior e inferior de los suelos de la Vega de Granada.

más profundo podría estar influido por una mayor cantidad de aportes de la red fluvial perteneciente al río Salado en otra época.

2.4.2. Materia orgánica

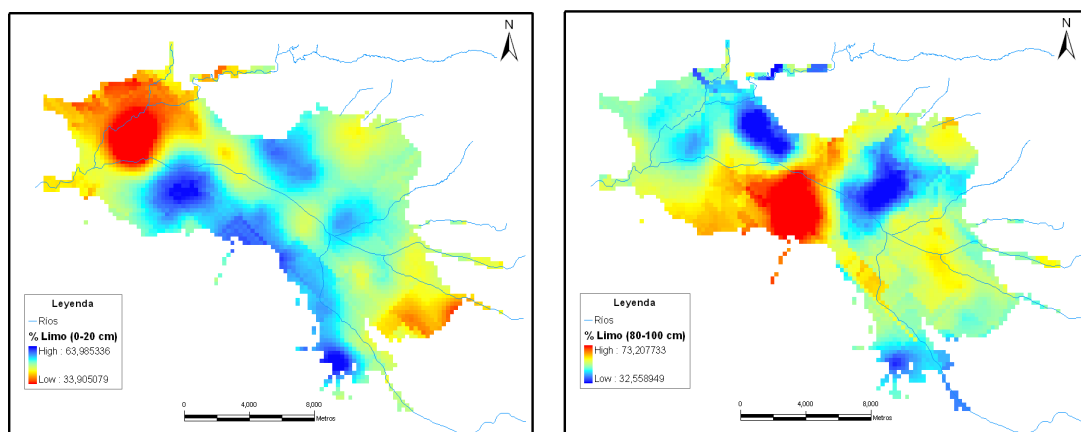
Las características físicas del suelo (textura, porosidad, etc.) pueden verse influenciadas por el contenido en materia orgánica. La materia orgánica, por su capacidad aglutinante, proporciona una cierta cohesión a los suelos arenosos y, en cambio, comunica cierta ligereza a los arcillosos, por lo que se considera un amortiguador de la textura del suelo. Además, la materia orgánica es un constituyente y un indicador de la calidad del suelo, pues las funciones del suelo se ven directamente afectadas por la cantidad y calidad de la materia orgánica que contiene, con incidencia en propiedades físicas como la estructura, la porosidad o la erosionabilidad. También afecta a las propiedades químicas como el intercambio iónico, la estabilización de nutrientes o la formación de complejos; y en propiedades biológicas como la actividad enzimática y diferentes interacciones como el secuestro de C de muchos suelos.

Los componentes orgánicos de los suelos proceden de aportes y de síntesis biótica y abiótica, y se distinguen: la biomasa que vive en el suelo y su necromasa; la acumulación de restos de plantas y animales (biomasa senescente o materiales de origen biológico aportados por el hombre); los productos xenobióticos resultantes de la síntesis industrial como los fitosanitarios, o los plásticos, el papel y el carbón; las macromoléculas resultantes de procesos de alteración y de humificación; y la síntesis abiótica.



(a) Distribución espacial del% de arenas del horizonte superior (0-20 cm) de los suelos de la Vega de Granada. (b) Distribución espacial del% de arenas del horizonte inferior (80-100 cm) de los suelos de la Vega de Granada.

Figura 2.54: Mapa del% de arenas en los horizontes superior e inferior de los suelos de la Vega de Granada.



(a) Distribución espacial del% de limos del horizonte superior (0-20 cm) de los suelos de la Vega de Granada. (b) Distribución espacial del% de limos del horizonte inferior (80-100 cm) de los suelos de la Vega de Granada.

Figura 2.55: Mapa del% de limos en los horizontes superior e inferior de los suelos de la Vega de Granada.

El contenido de materia orgánica de estos suelos, se expresa como carbono orgánico y ha sido medido mediante el análisis clásico de oxidación con dicromato potásico en medio sulfúrico y corregida posteriormente en función de la humedad. Los valores de carbono orgánico son bajos (entre 0.12 y 2.8 %), debido probablemente al hecho de que la Vega ha estado sujeta a un cultivo secular. A modo de ejemplo, en la figura 2.56 se muestra el mapa de la distribución del contenido orgánico medio en el horizonte de un metro y algunos perfiles de su variación con la profundidad, que expresan la disminución del carbono orgánico según distintos patrones, aunque, en la mayoría de ellos, la materia orgánica no se incorpora a los horizontes más profundos, en algunos de ellos llega a valores cercanos a 1,5 % a una profundidad de 40 cm. En los suelos de cultivo, la incorporación de la materia orgánica a horizontes más profundos es generalmente fruto del laboreo, lo que favorece la mezcla con la materia mineral y la masa de microorganismos del suelo.

La distribución espacial de la materia orgánica, por otra parte, revela una cierta alineación en sentido sureste a noroeste, con valores medios a lo largo del río, con valores más altos al norte y menores al sur. Esta diferencia, aunque es bastante pequeña, es significativa y podría ser debida a una diferencia en la composición mineralógica de las arcillas transportadas por los ríos de la zona norte y de la zona sur, ya que una afinidad distinta de adsorción o complejación de las moléculas orgánicas puede suponer un efecto protector frente a la acción de microorganismos. La distribución espacial de la materia orgánica en el horizonte superior (0-20cm) y en el inferior (80-100 cm) presenta un rango de variación entre 0,5 % y 1,37 % y entre 0,2 % y 0,7 %, respectivamente. Es decir, existe una disminución en el contenido en materia orgánica tal y como muestran algunos de los perfiles en profundidad (figura 2.57). Sin embargo, estas distribuciones según horizontes muestran débilmente la zonación noreste-sureste, comentada anteriormente, lo que es debido, probablemente, a un factor intrínseco, como podría ser la composición mineralógica de las arcillas comentada antes. El contenido de materia orgánica en un suelo puede verse influenciado por condiciones muy variadas y mucho más locales, como la textura, la vegetación, la estructura y el laboreo, que hacen complicada una interpretación más regional.

En el horizonte superior, de este modo, se observan los máximos localizados alrededor de Sierra Elvira y a la entrada de los ríos Genil y Monachil, lo que puede ser coherente con una variación en la composición de las arcillas. Sin embargo, para el horizonte inferior, la distribución está mucho más limitada a ciertas áreas: este efecto puede deberse al lavado, la estructura del suelo, o la vegetación.

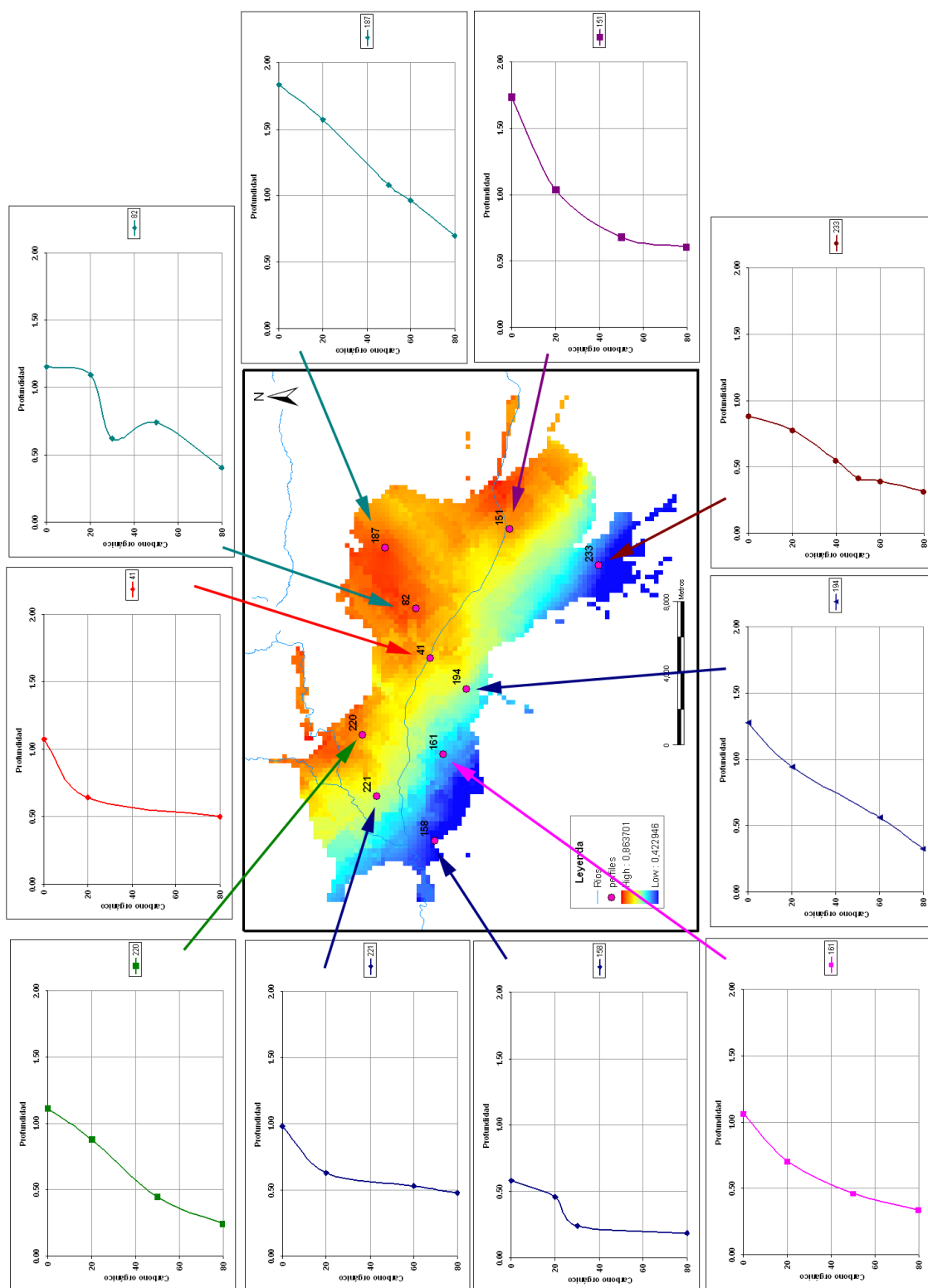


Figura 2.56: Distribución espacial del porcentaje de carbono orgánico en los suelos de la Vega de Granada en un horizonte de 1 m y perfiles en profundidad del % C.O.

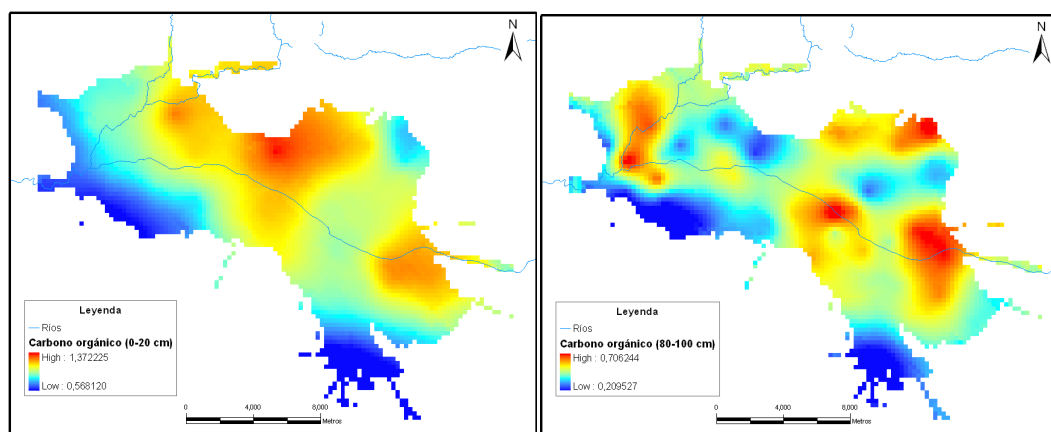


Figura 2.57: Distribución espacial del % carbono orgánico en suelo de los horizontes superior e inferior de la Vega de Granada .

2.4.3. pH del suelo

Conocer el valor del **pH** en los suelos es un aspecto fundamental en el estudio de zonas agrícolas, ya que la elección del cultivo a implantar, así como el manejo, están condicionados por el grado de acidez o basicidad de cada horizonte expresado por medio del pH. Por ello, conocer el valor del pH puede evitar fracasos al llevar a cabo revegetaciones en áreas forestales o a la hora de elegir los abonos en agricultura o jardinería.

Uno de los motivos de esta adaptación de las plantas a un pH determinado, es el hecho de que el pH del suelo afecta a la disponibilidad de los nutrientes. En el caso de suelos alcalinos-calcáreos, que es el caso de la Vega, abundan los nutrientes Na, K, Ca y Mg, pero es fácil que sean deficientes en elementos traza disponibles (excepto Mo). Por otro lado, estos suelos tendrán menos problemas de toxicidad, ya que los metales pesados tienden a estar precipitados en forma insoluble (Campos Fernández, 1997 [27]).

El estudio de este parámetro en este trabajo se hizo en agua, obteniendo un rango de variación bastante escaso debido a que se trata de suelos muy ricos en carbonatos, los cuales actúan como tampón impidiendo la bajada del pH. El poder tampón o poder amortiguador del suelo es la resistencia que presenta éste a modificar su pH cuando se le añaden ácidos o bases. El poder tampón por lo tanto permite mantener el pH dentro de unos rangos muy estrechos, por lo que se evitan modificaciones radicales en el ambiente del suelo que podrían afectar desfavorablemente a las plantas o a la liberación de nutrientes y metales. La capacidad de amortiguación está directamente relacionada con su contenido coloidal (arcillas de diámetro inferior a una micra) y con la capacidad de intercambio que posea: cuanto mayor sea ésta, mayor será su poder de amortiguación. De esta manera, la textura del suelo influye directamente en el poder tampón, y se puede

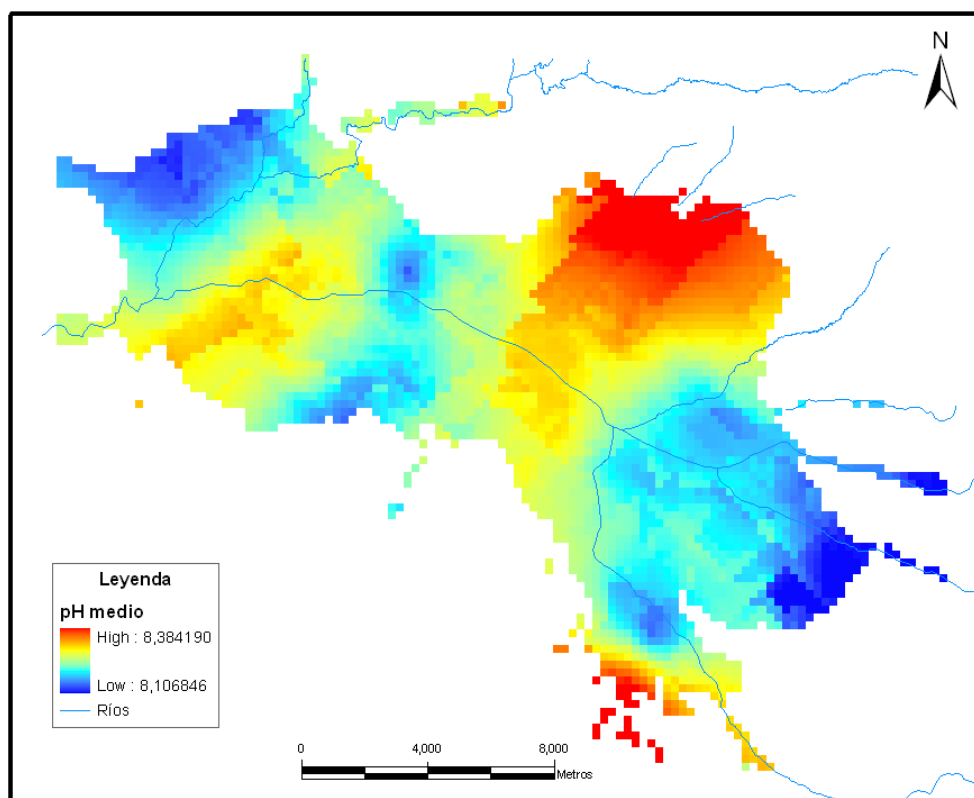
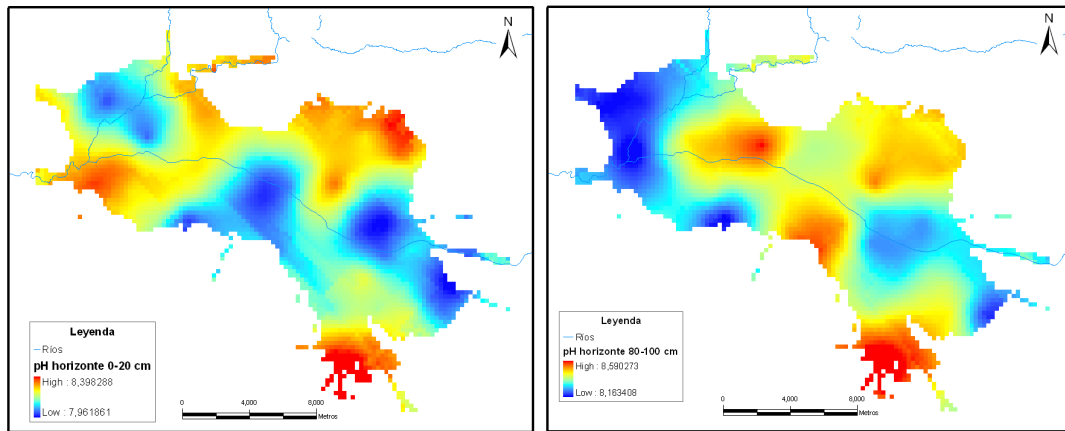


Figura 2.58: Distribución espacial del pH promedio en el suelo de la Vega de Granada.

enumerar de mayor a menor poder tampón desde húmico, arcilloso, franco a arenoso (Navarro Blaya, 2000 [127]).

A partir de los datos de cada horizonte, se ha determinado que la variación de pH se sitúa entre 7,7 y 9,2, es decir, que el suelo es mayoritariamente básico y ligeramente alcalino. En una distribución promedio en profundidad de un metro (figura 2.58), los máximos se sitúan al noreste y al sur del área de estudio, de modo que los perfiles superiores e inferiores son bastante similares a grandes rasgos (figura 2.59). Los mínimos del perfil promedio se sitúan en los bordes oeste y este, lo que coincide también a grandes rasgos con los horizontes superior e inferior. Sin embargo, al comparar el primer y último horizonte se aprecian casos opuestos. Así, en la zona central sur, se observa un mínimo de pH para el horizonte superior y un máximo en el horizonte inferior. Estas ligeras diferencias de pH vienen determinadas por los cationes adsorbidos sobre los minerales de la arcilla y la materia orgánica (Porta Casanellas, 2003 [29]), que podría ser una consecuencia más local debida a diferentes contenidos o composiciones de la arcilla o de la materia orgánica.



(a) Distribución espacial del pH en suelo del horizonte superior (0-20 cm). (b) Distribución espacial del pH en suelo del horizonte inferior (80-100 cm).

Figura 2.59: Distribución espacial del pH en suelo del horizonte superior e inferior de la Vega de Granada.

2.4.4. Humedad del suelo

Otro factor que condiciona la vida de las plantas en el suelo es la **humedad**. Para medir la humedad del suelo, en este trabajo, se ha utilizado un método gravimétrico previa eliminación del agua en una estufa a 105 °C. Sin embargo, esa cantidad de agua que tiene el suelo, debe expresarse en función de la fuerza con la que es retenida, ya que su comportamiento va a ser muy distinto en función de las fuerzas de retención a que se encuentre sometida. De este modo, si la mayor parte del agua está débilmente retenida ésta se puede mover y será asimilable para las plantas, mientras que si toda el agua está fuertemente retenida, carecerá de movilidad y será un agua inútil para las plantas.

El porcentaje del agua medido de esta manera, corresponde por lo tanto a una humedad natural sin distinguir la adsorbible de la no adsorbible por las plantas, ya que a 105°C no se consigue eliminar las moléculas de agua atrapadas entre las láminas de arcilla y se pierde parte del agua de cristalización de componentes importantes como el yeso ($CaSO_4 \cdot H_2O$), el sulfato sódico hidratado ($Na_2SO_4 \cdot H_2O$) y otras sales (suelos alcalinos) (Porta Casanellas, 2003 [29]).

La medida de esta variable, por lo tanto, se realiza en el sentido del agua como principal transporte de elementos solubles y contaminantes hacia la zona no saturada y saturada del acuífero. Su distribución espacial (figura 2.60) es bastante similar a la de las arcillas, pero con un máximo localizado en el noroeste y borde norte del área de estudio. La variación de este parámetro según los horizontes no es significativamente diferente como se muestra en la figura 2.61.

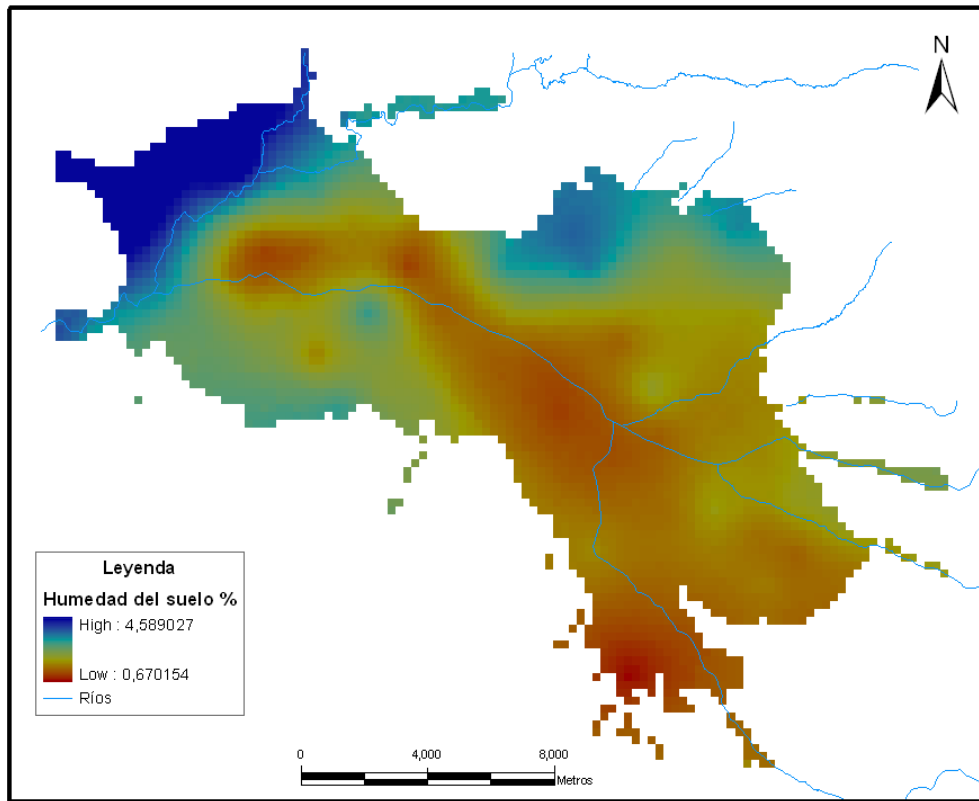


Figura 2.60: Distribución espacial del % humedad promedio (0-1 m) en el suelo.

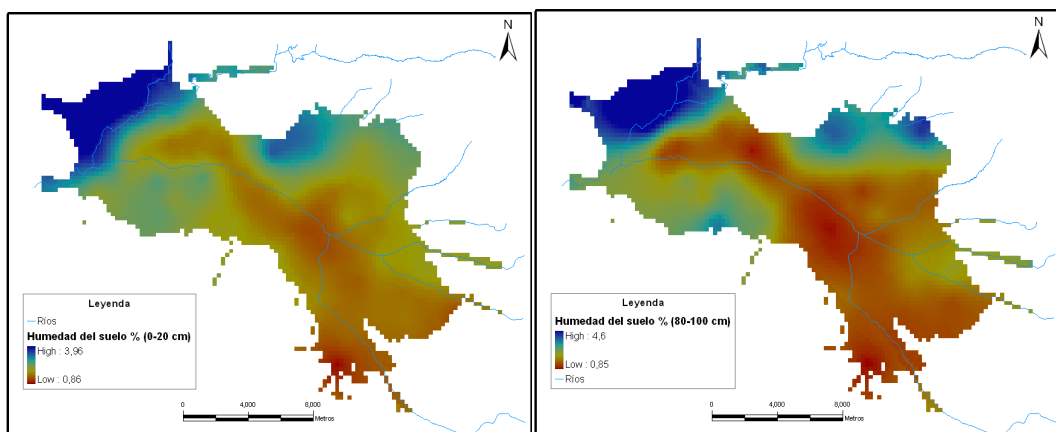


Figura 2.61: Distribución espacial del % humedad del suelo en el horizonte superior e inferior de la Vega de Granada.

2.4.5. Especies químicas de la fase líquida del suelo

Como se ha comentado anteriormente, el análisis de los parámetros del suelo se han realizado con el objetivo de estudiar su interrelación con las características de calidad del agua subterránea del acuífero. También se ha comentado antes la capacidad de la parte sólida del suelo como filtro natural. Sin embargo, también es interesante el estudio de la parte líquida o solución del suelo, ya que juega un papel fundamental en el acuífero al ser el vehículo de transporte de la mayor parte de sustancias hacia horizontes más profundos.

La parte líquida del suelo está formada fundamentalmente por agua, que circula por el espacio de macroporos y pasa a ocupar total o parcialmente los poros capilares donde puede ser retenida. Este agua disuelve y transporta elementos nutritivos, sales solubles y contaminantes, y hace posible su absorción por las raíces. Los iones disueltos precipitan al secarse el suelo y al añadir agua vuelven a disolverse. La reacción de un ión con agua se conoce como hidratación, proceso por el cual una serie de moléculas de agua quedan rodeando a cada ión. Estos elementos de la fase líquida presentan, por lo tanto, una elevada movilidad, y pueden ser transportados dentro del suelo (traslocados) o bien fuera de él (perdidos por lavado). Estos elementos móviles lavados de la solución del suelo, al ser arrastrados hacia capas más profundas pasarán a la zona no saturada del acuífero y posteriormente a la capa freática. El estudio de estos elementos es de especial interés ya que su concentración y abundancia influye directamente en la composición y calidad de las aguas subterráneas del acuífero.

Con el fin de correlacionar estos datos de suelos con los datos analizados del acuífero, se ha realizado un extracto de saturación del horizonte inferior del suelo (80-100 cm), en el cual la absorción de iones por las plantas es mínima, ya que a esta profundidad sólo llegaran las raíces de algunos árboles o algunas plantas típicas de zonas más áridas, por lo que la mayor parte de este contenido se integra en la zona vadosa. El extracto de saturación se ha preparado para muestras tamizadas a 2 mm y secadas en estufa a 30°C. La extracción se realiza en el punto de capacidad de campo (saturación) por medio de un rizom, que extrae la solución a una presión equivalente a la que una raíz la absorbería. Se obtiene de este modo una solución filtrada y microbioestéril (sin bacterias ni materia orgánica). Este método de extracción es la forma menos agresiva de extraer la solución del suelo y correspondería a la situación más cercana a lo que ocurre en la realidad, cuando el agua de riego o de lluvia arrastra los elementos hacia la capa freática.

El resultado de las especies químicas que encontramos en este extracto corresponde, por lo tanto, a aquellas formas que pueden pasar al ciclo biológico y que se encuentran en la solución del suelo. Por este motivo, los resultados de estos análisis se expresan en *mg/l* o *µg/l* en agua en lugar de en *mg/kg* de suelo, por lo que se obtiene así una idea

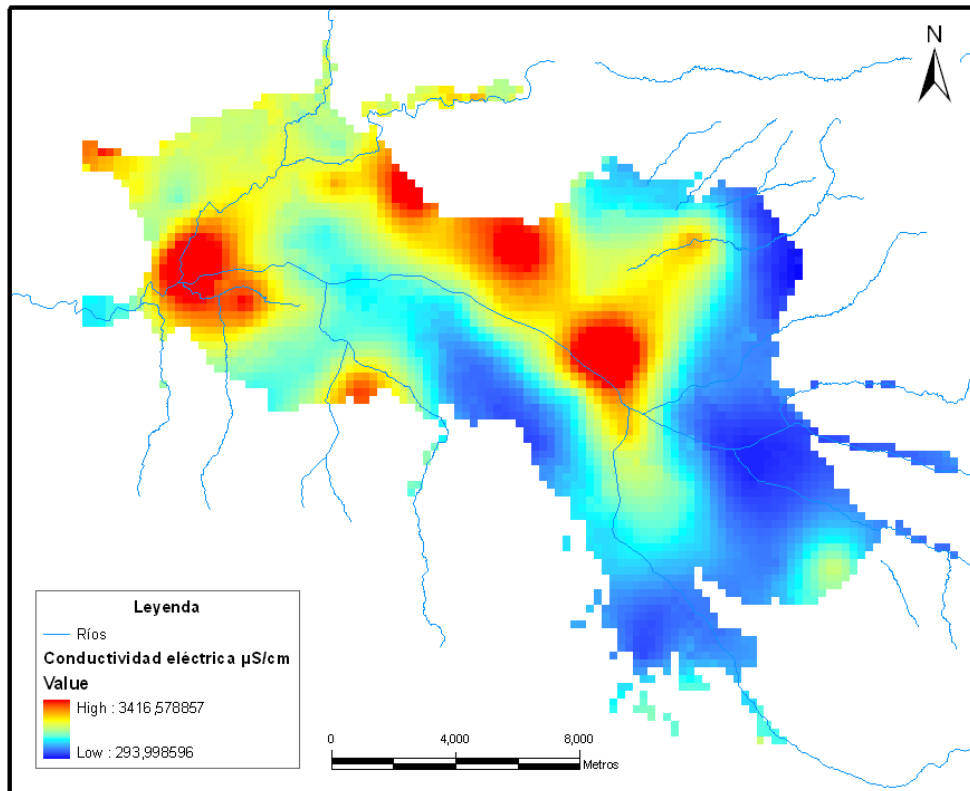


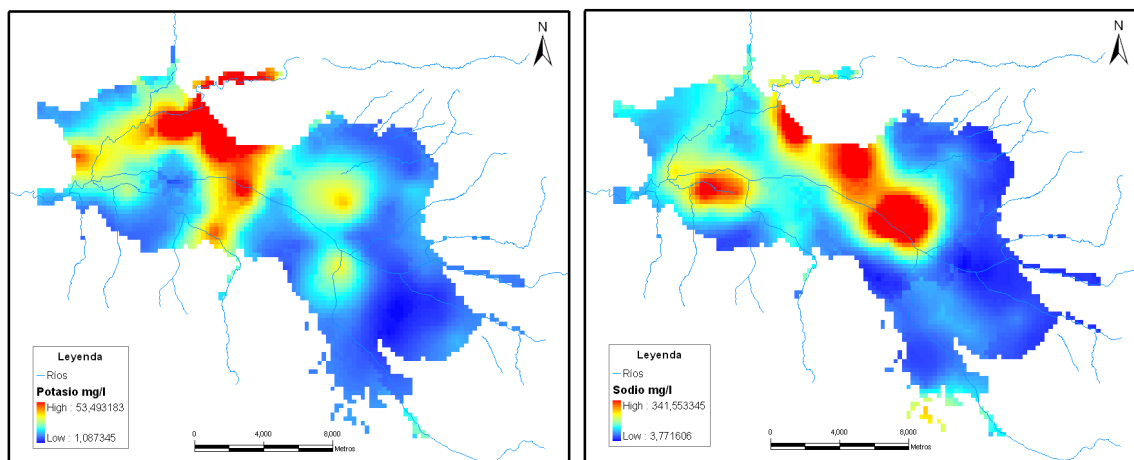
Figura 2.62: Distribución espacial de la conductividad en la solución del suelo de la Vega de Granada.

más intuitiva de la concentración en líquido para su correlación con las variables del agua subterránea.

La **conductividad** expresa una idea del total de aniones existentes en el suelo. La distribución espacial de la misma muestra valores mínimos en la zona de recarga o borde este, y valores máximos al norte del río Genil y en la zona de Láchar y río Salado (borde oeste y borde sur) (figura 2.62).

Las variables en el suelo las podemos dividir en dos grupos: los iones y cationes mayoritarios y los metales. La existencia de iones a esta profundidad puede deberse tanto a la roca madre como a procesos de lavado, queluvación o lixiviación. Estos procesos de lixiviación o emigración de sales solubles, afectan fundamentalmente a los cationes más móviles como son los alcalinos y los alcalinotérreos (sodio, potasio, magnesio y calcio) que existen en las soluciones del suelo en estado de equilibrio con las formas "de cambio" retenidas por el complejo adsorbente.

El anión "arrastrante" puede ser inorgánico (como el ácido nítrico o carbónico) u orgánico (como el ácido láctico). Los suelos que contienen además carbonatos (calizas



(a) Distribución espacial del contenido en potasio de la solución del suelo. (b) Distribución espacial del contenido en sodio en la solución del suelo.

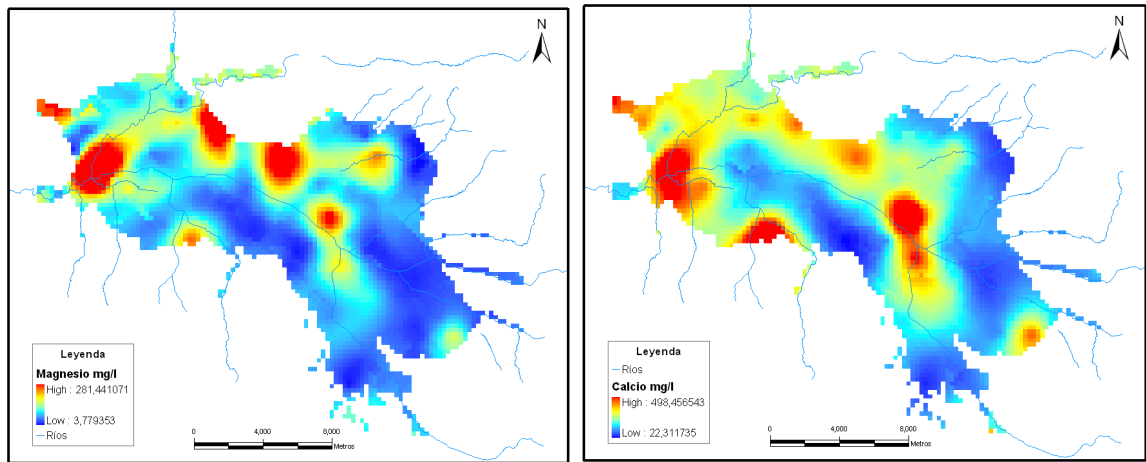
Figura 2.63: Mapas del contenido en potasio y sodio en la solución del suelo de la Vega de Granada.

y dolomías) sufren un proceso particular de lixiviación, la descarbonatación, que tiene lugar generalmente (pero no exclusivamente) por la acción del CO_2 disuelto.

También en este caso, la norma general, en clima húmedo, son las pérdidas por drenaje profundo; sin embargo, en clima más seco, con una fuerte evapotranspiración, la precipitación del bicarbonato cálcico se produce en profundidad, y da lugar a la formación de un horizonte iluvial cálcico o incluso gípsico, cuando la precipitación se produce en forma de yeso (Duchaufour, 1984 [64]). El origen del calcio iónico o de cambio (el que se encuentra en la solución del suelo) puede ser doble: la alteración de los minerales cálcicos y la biodegradación de los complejos órgano-minerales de la hojarasca (proceso limitado en la Vega a las choperas).

La distribución de los cationes solubles se muestra en las figuras 2.63 y 2.64.

La distribución de estos cationes en la solución tiene probablemente un origen relacionado con la desintegración, descomposición o disolución de las rocas que los contienen (en este caso los carbonatos, sulfatos y sales), aunque la mayoría de ellos pueden ser incluidos en el suelo por medio de diversos fertilizantes. La distribución de estos cationes es bastante similar, y se observan varios puntos en común, como un máximo situado al sur de Sierra Elvira, a la desembocadura del río Genil a la altura de Láchar, a la entrada del río Salado y en la confluencia del río Genil con el Beiro y el Dílar. La situación al sur de Sierra Elvira y a la entrada del río Salado, nos da una idea de este origen asociado al material originario, pues como se comentaba en la descripción geológica, en esos sectores existen materiales evaporíticos como la halita, silvinita, carnalita y yeso.



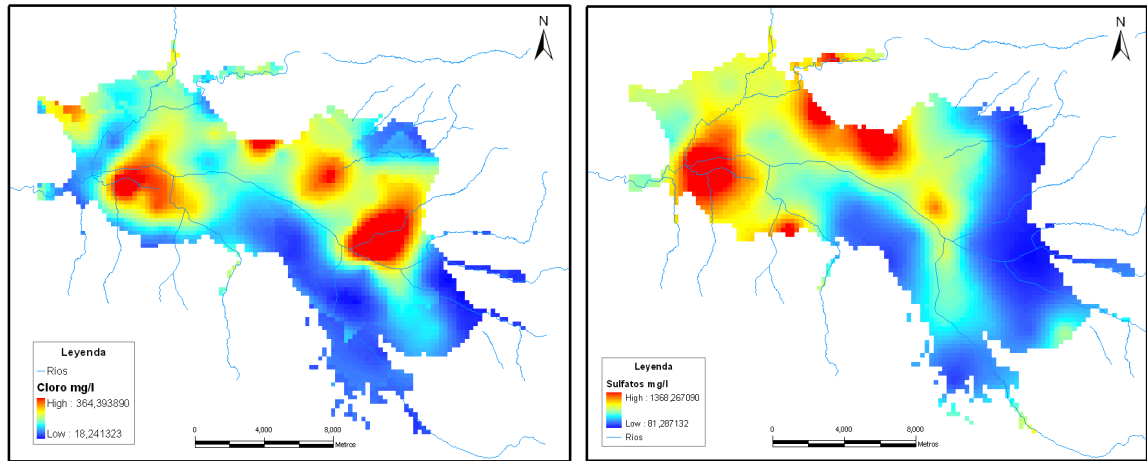
(a) Distribución espacial del contenido en magnesio de la solución del suelo. (b) Distribución espacial del contenido en calcio de la solución del suelo.

Figura 2.64: Mapas del contenido en magnesio y calcio en la solución del suelo de la Vega de Granada.

Si se compara el comportamiento de los iones entre sí, el de los monovalentes difiere profundamente de los divalentes. Los primeros son más móviles en clima húmedo, y salvo por una minoría de iones K^+ que pueden ser fijados por las arcillas, son arrastrados prioritariamente y eliminados casi siempre fuera del perfil. Los iones divalentes por el contrario, son menos móviles y su arrastre es más lento; además, parte de ellos son retenidos en profundidad, en los horizontes B_t , por el complejo adsorbente. Para las regiones más áridas, donde el drenaje es prácticamente nulo, las soluciones se enriquecen mucho en cationes divalentes, pero no puede producirse el arrastre debido a la adsorción de los monovalentes por el complejo adsorbente y el paso de los divalentes a la solución (Duchaufour, 1984[64]).

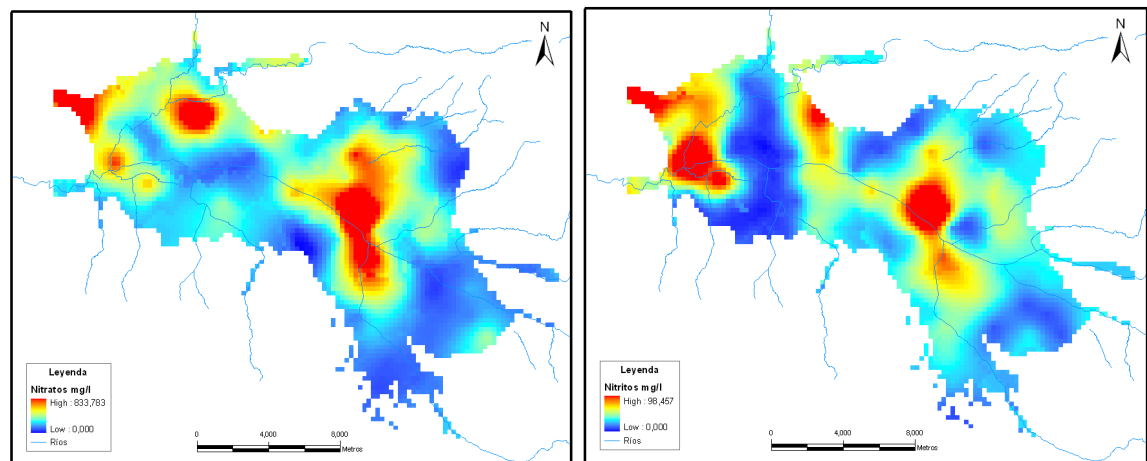
El ión **sodio** sólo puede subsistir en el perfil, debido a su elevada solubilidad, en clima seco cuando la fuerte evapotranspiración potencial impide su drenaje. La abundancia de este ión forma suelos salsódicos, típicos de regiones esteparias o semidesérticas, que no es el caso de la Vega de Granada, pero su origen sí puede estar ligado a la presencia de halita. La presencia de halita es probablemente la causa de su distribución en la Vega tal y como se infiere al comparar las figuras 2.63 y 2.65a, ya que ambas especies, el cloro y el sodio, muestran una clara correlación espacial a simple vista.

En condiciones naturales el **nitrógeno** del suelo no proviene de la degradación de la roca madre, si no que deriva del que existe en la atmósfera terrestre a través de los distintos procesos de fijación, fundamentalmente de tipo biológico. Pero en suelos cultivados, como es el caso de la Vega, la mayor parte de este nitrógeno proviene del empleo de fertilizantes nitrogenados. Además de la variación que puede existir debido a



(a) Distribución espacial del contenido en cloro de la solución del suelo. (b) Distribución espacial del contenido en sulfatos de la solución del suelo.

Figura 2.65: Mapas del contenido en cloro y sulfatos en la solución del suelo de la Vega de Granada.



(a) Distribución espacial del contenido en nitratos de la solución del suelo. (b) Distribución espacial del contenido en nitritos de la solución del suelo.

Figura 2.66: Mapas del contenido en nitratos y nitritos en la solución del suelo de la Vega de Granada.

la diferente adición de fertilizantes, su concentración también varía dependiendo de las condiciones de drenaje, topografía, textura del suelo y prácticas de cultivo. Según este esquema, observamos que la mayor lixiviación de nitratos y nitritos en los horizontes inferiores sucede en tres sectores: la confluencia del Genil con el río Dílar, en el borde oeste del área de estudio y al este del río Cubillas (figura 2.66). La elevada concentración de los nitratos a esta profundidad del suelo se debe al bajo poder de adsorción que presentan la mayoría de los suelos para las partículas cargadas negativamente, motivo también por el cual, las formas en las que se puede encontrar el nitrato en el suelo son casi exclusivamente como iones libres en la disolución del suelo (Navarro Blaya, 2000[127]).

El origen del **sulfato** en la solución del suelo, se explica por dos factores principales, la disolución de yesos originarios de la roca madre, y la adición de sulfatos como abono (sulfato de magnesio o de calcio con diferentes grados de hidratación). Se observa en la distribución espacial del ión sulfato que existe cierta correlación espacial con ambos iones (calcio y magnesio), por lo que su origen puede asociarse a la litología conocida, con un origen más ligado así a la roca madre en la zona sur de Sierra Elvira y la zona del aeropuerto, y un origen relacionado con las enmiendas del suelo en la zona de Láchar y confluencia del Genil y Dílar.

Para finalizar este apartado se comentarán algunos aspectos sobre la presencia, origen, funcionamiento y distribución de los metales analizados. La distribución de los metales en los suelos se muestra en las figuras 2.67, 2.68 y 2.69.

El **Manganeso** es esencial en la vida de la plantas ya que interviene en numerosos procesos metabólicos, asemejándose, según su comportamiento químico, en algunos aspectos al calcio y al magnesio y en otros a los oligoelementos hierro y cinc. Algunos de estos procesos son la fotosíntesis, el metabolismo del nitrógeno y la síntesis del ácido adsorbico (Navarro Blaya, 2000 [127]). El manganeso de los suelos es originado principalmente por la descomposición de las rocas ferromagnéticas, aunque forma parte de numerosos minerales (sulfuros, óxidos, carbonatos, silicatos, fosfatos, arseniatos y otros). Por otra parte, su origen antrópico puede estar ligado a los diferentes usos del Mn dentro de la industria metalúrgica, pilas alcalinas, cerámica, tintes, etc.

Las formas del manganeso en el suelo son complejas, ya que existe en distintos estados de oxidación, y puede formar óxidos según estos estados, que no siempre son estequiométricos. Además, los óxidos superiores pueden estar en forma cristalina y pseudocristalina. Los hidróxidos de Fe y Mn tienen un comportamiento anfótero. Los procesos de oxidación y reducción que sufren las formas del manganeso están causados por factores tanto químicos como microbiológicos. En los suelos la distribución varía mucho según el pH. En suelos de pH bajo, predominan las formas solubles en agua, cambiables y ligadas a la materia orgánica, mientras que en suelos con pH alto existe

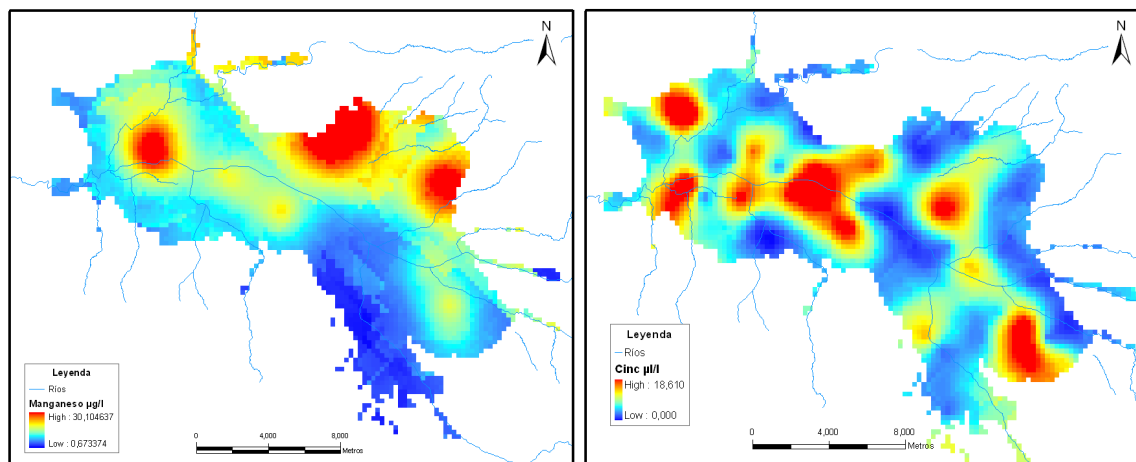
gran cantidad de manganeso reducible, menos soluble que el anterior. En suelos calcáreos el Mn puede ser fijado por el $CaCO_3$ y más tarde precipitar como carbonato.

En el área de estudio (figura 2.67), los valores más altos se sitúan en el borde noreste y al oeste. Esta última zona también aparece en el estudio del primer horizonte de suelo realizado por Campos Fernández-Fígares (1997 [27]) y la autora lo relaciona con el área de influencia de las margas que contienen depósitos de manganeso causados por fenómenos de hidromorfía no actuales. En la zona de Sierra Elvira, además, el manganeso podría estar asociado también a los carbonatos de Sierra Elvira, aunque también podría estar relacionado con la presencia de los polígonos industriales que existen en el área.

El **Cinc** es esencial para la vida de las plantas y es usado como fertilizante. Se acumula en las plantas pero no a niveles tóxicos para el consumo humano o animal. La toxicidad es más frecuente en suelos ácidos. Es fácilmente complejado en el suelo y su principal origen natural son los sulfuros de Zn (esfalerita y blenda), aunque también se pueden citar como compuestos principales los silicatos (willemita y hemimorfita), los óxidos (óxido de zincita) y los carbonatos (smithsonita). Como posibles fuentes de contaminación se pueden citar la industria del vidrio y del galvanizado, la combustión de petróleo, incineración de basura, minería, neumáticos y pantallas de televisión, cementeras, fertilizantes e insecticidas. En las rocas sedimentarias el cinc es un constituyente menor junto con el hierro. El nivel de Zn en los suelos varía mucho según su origen, entre 10 y 300 mg/kg , aunque hay que resaltar que su contenido útil suele situarse en la capa arable, ya que suele quedar adsorbido por las arcillas y la materia orgánica, lo que impide de su emigración descendente (Navarro Blaya, 2000 [127]).

Su solubilidad está en relación inversa a la presencia de Ca y de P; los complejos solubles con la materia orgánica pueden, sin embargo, permitirle mantenerse en solución a pH alcalino. La distribución espacial de la concentración de Zn en el área de estudio (figura 2.67) muestra los máximos valores en las zonas más cercanas a los cursos del río y los mínimos en las zonas más cercanas a los bordes. La mayor parte de cinc en el suelo permanece adherido a las partículas de suelo y no se disuelve en agua, por lo que el mayor contenido en las zonas de los ríos pudiera reflejar un origen relacionado con un mayor lavado de esas zonas.

El **Cobre** es esencial para la vida de las plantas, pero puede ser tóxico, sobre todo en suelos ácidos, como la mayoría de los metales. Se acumula en las plantas, pero no a niveles tóxicos para el consumo. Se encuentra fundamentalmente como sulfuro (calcosita, calcopirita y covelita), óxido (cuprita y tenorita) y carbonato (malaquita y azurita). El cobre es isomórfico con Fe^{2+} , Zn^{2+} y Mg^{2+} . Es fácilmente complejable en el suelo. Tiene una amplia utilización en aleaciones y se usa como fertilizante, fungicida y bactericida. El nivel normal de la concentración de cobre en los suelos de cultivo es de 20 a 55 mg/kg . El pH influye fuertemente en la solubilidad del cobre, y queda reducida



(a) Distribución espacial del contenido en manganeso de la solución del suelo. (b) Distribución espacial del contenido en cinc de la solución del suelo.

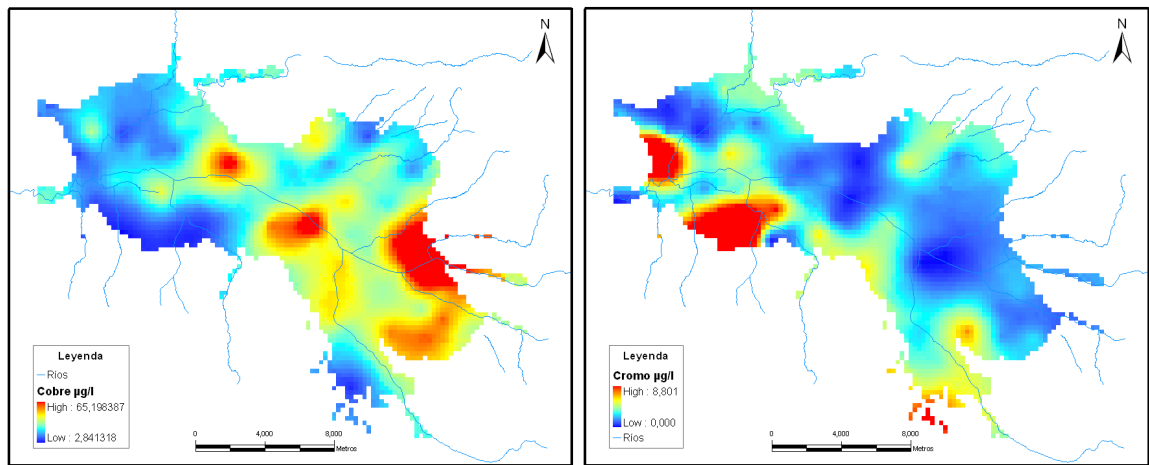
Figura 2.67: Mapas del contenido en manganeso y cinc en la solución del suelo de la Vega de Granada.

fuertemente por encima de 7. Sin embargo puede estar en forma soluble en suelos de pH alto, formando complejos orgánicos, de modo que la presencia de materia orgánica tiene un papel fundamental. De todos los metales divalentes, el cobre es el que forma con la materia orgánica los complejos más estables que, junto con los óxidos de Fe y Mn, son las formas que controlan la fijación del cobre en el suelo (Campos Fernández-Fígarez, 1997[27]).

La distribución del cobre en el suelo se localiza fundamentalmente en la zona este de la Vega, centrándose el máximo principal en la zona de Granada. Este máximo se corresponde con la distribución que Campos Fernández-Fígarez (1997, [27]) encuentra en el primer horizonte de suelo, y al que alude como anomalía geoquímica o contaminación. Los valores mínimos de cobre se sitúan en la zona oeste de la Vega.

El cobre no se degrada en el medio ambiente, pero tiende a fijarse a las partículas de materia orgánica, arcilla, tierra o arena. La mayor concentración de cobre, sobre todo, en la zona de recarga, podría indicar un origen relacionado con las rocas metamórficas de Sierra Nevada y con un porcentaje más bajo de arcillas, que no favorece la complejación de este metal y permite su lavado hasta el horizonte más profundo.

El **Cromo** aparece naturalmente en el medio ambiente con estado de valencia 3, y aparece fundamentalmente en forma de cromita. La movilidad del cromo en la pedosfera solamente puede evaluarse si se consideran la capacidad de adsorción y reducción de los suelos y de los sedimentos. Los hidróxidos de cromo (III), una vez sedimentados y fijados en el sedimento acuático, difícilmente vuelven a moverse, dado que la oxidación de los compuestos de cromo (III) para formar compuestos de cromo (VI) prácticamente no



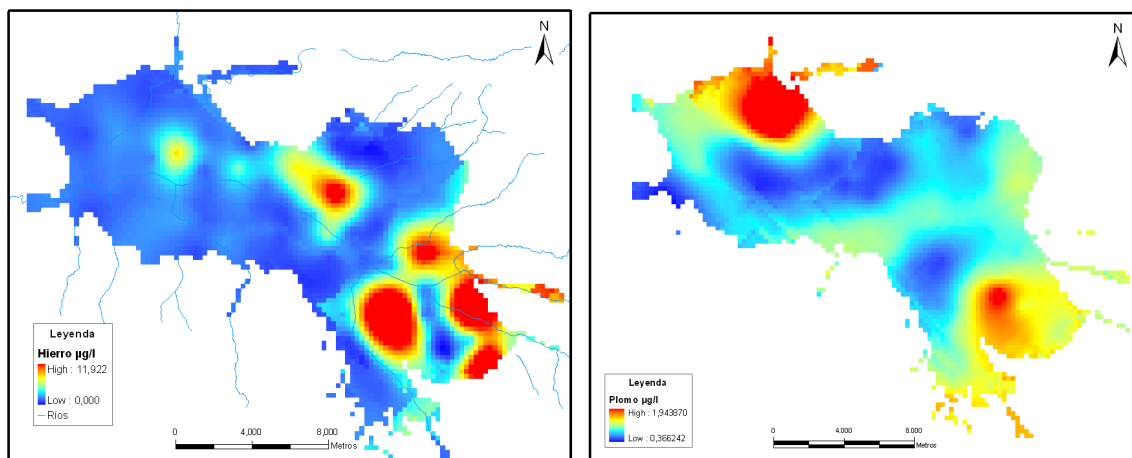
(a) Distribución espacial del contenido en cobre de la solución del suelo. (b) Distribución espacial del contenido en cromo de la solución del suelo.

Figura 2.68: Mapas del contenido en cobre y cromo en la solución del suelo de la Vega de Granada.

ocurre en forma natural. El cromo (VI), aun en concentraciones relativamente bajas, ya resulta tóxico, por lo que el pH del suelo es un factor fundamental. El uso de abonos fosfatados incrementa el ingreso de cromo en el suelo.

Para los suelos de la Vega se distinguen dos máximos muy localizados a la entrada del río Salado y a la salida del río Genil. Estos máximos corresponden probablemente a un factor de naturaleza antrópica.

El **Hierro** es un elemento esencial para la vida humana y las plantas. La concentración de hierro en el suelo varía mucho más de la que el agua de riego pueda aportar. Es difícil que sea tóxico aunque, cuando hay un alto nivel de sus formas amorfas, puede inducir una deficiencia de P y Mo. Sus propiedades geoquímicas son similares a las del Co y el Ni. Su presencia es abundante en la naturaleza, no se considera un elemento traza. Con respecto a la función en las plantas ocupa una posición intermedia entre macro y micronutriente. Tiene un papel muy importante en el comportamiento de los demás metales pesados del suelo. El hierro del suelo procede en general del proceso de edafización, y determina la composición de los minerales constituyentes de diferentes tipos de rocas. Su presencia en el suelo es, fundamentalmente, en forma de óxido e hidróxido, como pequeñas partículas o en la superficie de otros minerales, lo cual le da una coloración rojiza en climas más áridos (contienen óxidos no hidratados) y amarillo-pardo para las zonas templadas o húmedas (contienen óxidos más hidratados). Su movilidad o su fijación dependen, esencialmente, de las características redox del suelo y del pH. Las condiciones oxidantes y el pH alcalino provocan su precipitación. En condiciones reductoras y de acidez se facilita su disolución. Los compuestos de hierro, tanto orgáni-



(a) Distribución espacial del contenido en hierro de la solución del suelo. (b) Distribución espacial del contenido en plomo de la solución del suelo.

Figura 2.69: Mapas del contenido en hierro y plomo en la solución del suelo de la Vega de Granada.

cos como inorgánicos, se transforman fácilmente. El contenido en suelos normalmente oscila entre 0,5 y 5 %. En climas áridos es fácil que aparezca una deficiencia de Fe en los suelos, especialmente en suelos calcáreos y alcalinos.

El contenido de este metal en forma soluble en la Vega es bastante escaso, y presenta la mayor proporción al oeste del área de estudio, con una distribución muy similar al cobre. Igual que ocurre con el cobre, una mayor proporción de arenas puede producir una menor adsorción del hierro por los suelos, lo que permite que el hierro pase a horizontes más profundos.

El origen natural del **plomo** en el suelo suele venir asociado a la meteorización de minerales como la galena, que por oxidación da lugar a minerales secundarios de plomo como la piromorfita, la cerusita, la plumbogumita, la anglesita y la mimetita. Es un metal pesado carente de acciones benéficas para los organismos y se presenta como Pb^{2+} y Pb^{4+} . Su biodisponibilidad está fuertemente condicionada por el pH del suelo, de modo que aumenta al disminuir el pH y es especialmente móvil a pHs ácidos (Prueb, 1997 [147]). Los compuestos de Fe, Mn y Al presentan una fuerte capacidad de absorber Pb. Con la materia orgánica forma complejos estables a pH altos, como los quelatos de Pb, que representan un alto porcentaje entre las especies solubles del Pb. Tiene una fuerte tendencia a unirse a los fosfatos para formar compuestos insolubles (Alloway, 1995 [9]).

La concentración de este metal en la Vega a la profundidad de estudio es muy baja, y está en la mayoría de los casos por debajo del límite de detección. No obstante otros estudios (Campos, 1997 [27]) muestran elevadas concentraciones (hasta 170 mg/kg) de

plomo en el primer horizonte de suelo (0-30 cm) en la zona de Granada capital y que se relacionan probablemente con las emisiones de gas de los coches de gasolina con plomo. Sin embargo, estas concentraciones no se observan a la profundidad estudiada, lo que plantea la situación de que exista una posible precipitación del mismo en niveles intermedios.

2.5. Imágenes de teledetección espacial y aerotransportada

La teledetección se define, en sentido amplio, como la adquisición de información de un fenómeno, objeto o material, mediante un dispositivo de medida que no está en contacto físico directo con el mismo (Chuvieco Salinero, 1996 [41]), y engloba no sólo los procesos que permiten obtener una imagen, sino también su posterior tratamiento, en el contexto de una determinada aplicación (Chuvieco Salinero, 2002 [42]).

La aplicación de las imágenes de satélite han sido utilizadas con éxito en muchos ámbitos de las geociencias, aunque en el campo de los recursos hídricos subterráneos estas aplicaciones han sido más tardías, por lo que la aceptación de esta tecnología como herramienta de trabajo para el hidrogeólogo ha tenido cierto retraso. La aplicación de estas técnicas, aunque siempre como apoyo y complemento fundamentalmente en el área de la gestión y la planificación, supone una herramienta muy novedosa como entrada en modelos complejos de evaluación o gestión de acuíferos subterráneos. Algunos de los datos más interesantes que se pueden obtener a partir del estudio de las imágenes de satélite son, por ejemplo, la demanda de agua, la evapotranspiración o la cartografía de los tipos de cultivo, aunque estas aplicaciones se comentarán más adelante.

Uno de los aspectos más interesantes en el uso de imágenes de teledetección en el estudio de los procesos medioambientales es la globalidad de la información en cuanto a la extensión en el territorio y su homogeneidad, en referencia a situaciones comparables (mismo sensor, misma altura, etc). La comparación entre imágenes de satélite, no sólo tiene un contexto espacial (diferentes sitios) sino que también permite realizar estudios multitemporales que ofrezcan información sobre la evolución a lo largo del tiempo de distintos procesos, tales como desertificación, análisis de inundaciones, escorrentía en coberturas de nieve, urbanización, control de las extracciones de agua subterránea, estimación de la cantidad de agua necesaria según tipo de cultivo, etc. y por supuesto, la dinámica de los fenómenos meteorológicos.

El principal problema en relación con la teledetección que se aborda en esta tesis, es la construcción de cartografía temática enfocada fundamentalmente a la identificación de los usos del suelo. La definición de las actividades desarrolladas sobre la superficie del acuífero permite identificar posibles focos de contaminación, así como la posibilidad

de relacionar el tipo de contaminación difusa (nitratos, sulfatos, etc.) con ciertos tipos de cultivos.

La selección del conjunto de imágenes, así como su resolución espacial, espectral, radiométrica o temporal son puntos clave ya que en función de los objetivos que se pretendan conseguir, se deberá escoger uno u otro tipo de imagen. La resolución espacial se refiere a la distancia mínima entre dos objetos que permite su diferenciación; también definida como el tamaño del objeto más pequeño discernible. Numéricamente, viene expresado por el tamaño de píxel. La resolución espectral se refiere a la capacidad de un sensor de definir y detectar pequeños intervalos de longitudes de onda. Por tanto, una resolución espectral más fina implica un rango de longitudes de onda más estrecho para una banda o canal. Muchos sistemas de teledetección registran la energía en distintos rangos de longitudes de onda separados y con distintas resoluciones espectrales, y constituyen los llamados sensores multiespectrales. La resolución radiométrica de un sensor describe su capacidad para discriminar ligeras diferencias en energía. Cuanto más fina es la resolución radiométrica de un sensor, más sensible es para detectar pequeñas diferencias en la energía recibida en 256 niveles (2^8), lo que supone una resolución radiométrica relativamente buena. La resolución temporal se refiere al intervalo de tiempo mínimo, normalmente varios días, que ha de transcurrir para obtener dos imágenes del mismo sector (diferente a la resolución espacial absoluta, igual al periodo orbital del satélite).

Las imágenes que se han incluido en la base de datos de esa investigación corresponden a distintos tipos de sensores correspondientes a distintas resoluciones. De manera resumida son los siguientes: una serie de imágenes Landsat Thematic Mapper de los meses de verano desde 1985 al 2002, y una imagen AHS (Airborne Hyperspectral Scanner) del año 2004. A continuación se hace una pequeña descripción más detallada de estas imágenes.

2.5.1. Imágenes multiespectrales: Landsat Thematic Mapper

El programa Landsat comienza en los años 70 dentro de un proyecto de la agencia espacial norteamericana con el objeto de estudiar los recursos terrestres. El gran éxito del programa, radica en la alta calidad de sus sensores, el carácter global y periódico que realizan y su buena comercialización.

Las imágenes Landsat utilizadas en este trabajo corresponden a los satélites Landsat 5 y 7 y tienen cierta heterogeneidad. Desde el año 1985 hasta el año 1999, las imágenes eran tomadas por el sensor Thematic Mapper (TM), incorporado en el Landsat 5, y diseñado, como su nombre indica, para la cartografía temática. Es un sensor de barrido multiespectral incluido en la nueva generación de sensores. Se caracteriza por su relativamente buena resolución espectral, con 7 bandas y por ofrecer información en zonas

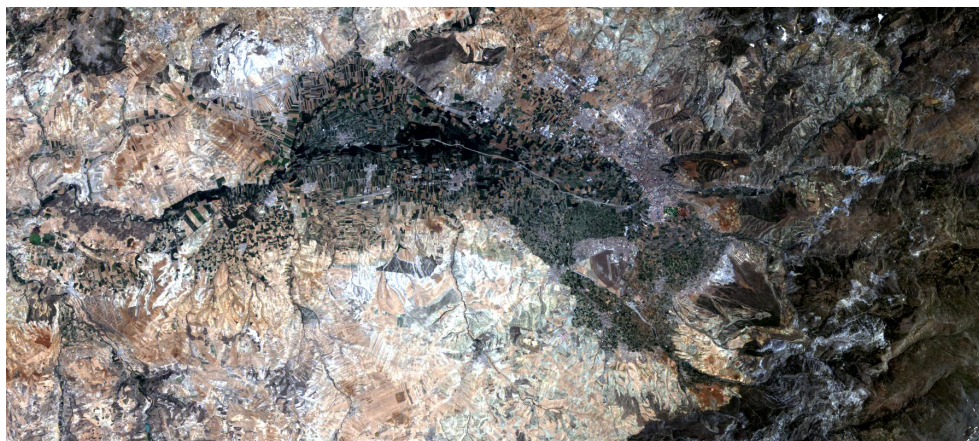


Figura 2.70: Composición TM 3,2,1 (RGB) "color verdadero" año 1985.

del espectro de gran interés para la investigación (visible e infrarrojo). En este sensor, las bandas visibles (1,2 y 3) tienen mejores disposiciones para la detección de aspectos antrópicos, como áreas urbanas o vías de comunicación, a la vez que proporcionan una interesante información sobre la calidad del agua. Las dos bandas en el infrarrojo (4 y 5), por otra parte, pretenden registrar parámetros vitales en la actividad de las plantas, además de ser especialmente sensibles a la humedad. Su resolución espacial es de 30 m y su resolución radiométrica de 8 bits.

En el año 1999 se lanza el satélite Landsat 7, que incluye el sensor ETM+ (Enhanced Thematic Mapper) que mejora las características del TM al añadir una banda pancromática de 15 m de resolución y aumentar la resolución de la banda térmica a 60 m.

El incremento de la resolución espacial y espectral que ha ofrecido el programa Landsat en sus dos últimos sensores, tanto el TM como el ETM+, ha permitido mejorar notablemente la precisión obtenida en múltiples ámbitos de la cartografía temática, a la vez que ha permitido ampliar el rango de aplicaciones operativas a partir de sensores espaciales. Estas nuevas aplicaciones son, por ejemplo, la detección de turbidez y contaminantes en el agua, contenidos de humedad en las plantas y el suelo, detección de focos de calor, etc., muy importantes por su relación con la temática de este trabajo.

En algunos de los ejemplos más representativos de estas imágenes (figuras 2.70, 2.71 y 2.72), se distingue a simple vista tanto el crecimiento de algunos núcleos de población, como la disminución de la vegetación a lo largo de los años.

2.5.2. Imágenes hiperespectrales: Airbone Hyperspectral Scanner

Las imágenes aerotransportadas (AHS) o recogidas por el radiómetro hiperespectral aerotransportado de barrido mecánico, fueron tomadas por el Instituto Nacional de



Figura 2.71: Composición TM 3,2,1 (RGB) "color verdadero" año 1993.

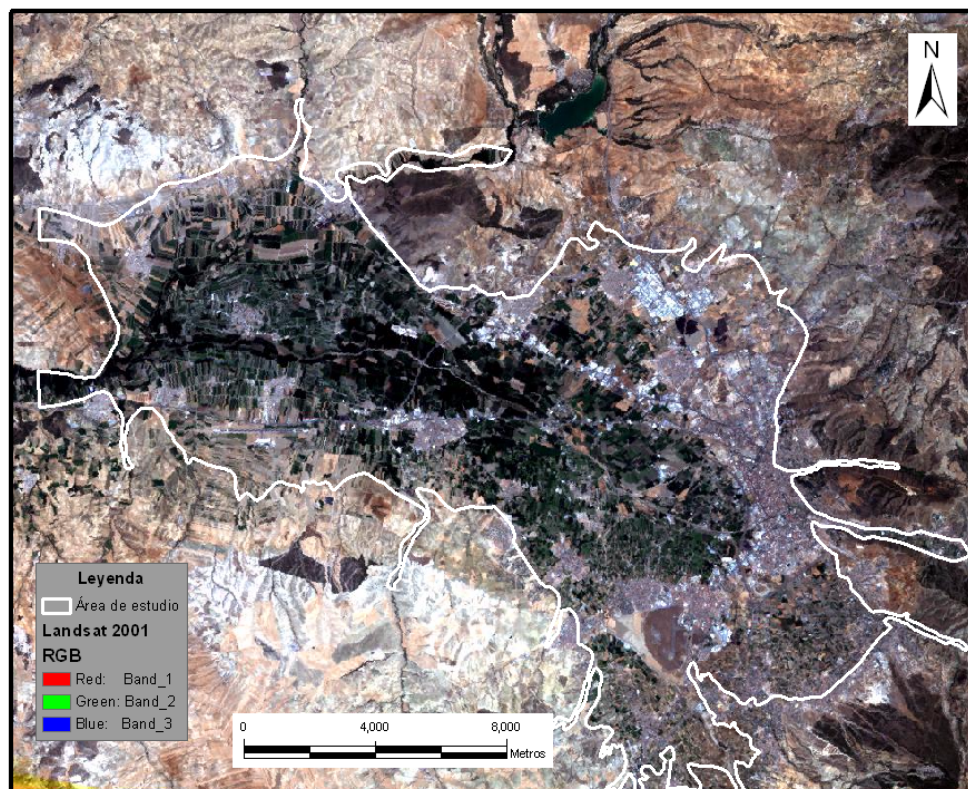


Figura 2.72: Composición TM 3,2,1 "color verdadero" año 2001.

Técnica Aeroespacial (INTA) en Julio del 2004 a bordo de la aeronave CASA C-212 e integradas con un GPS. Este radiómetro tiene cobertura desde el espectro visible al térmico, y pretende ser una herramienta de alta resolución espectral, radiométrica y espacial muy interesante para la comunidad científica.

Las imágenes se adquirieron en el nivel 1c, lo que corresponde a un primer procesamiento de los datos brutos (nivel 0) más una corrección a radiancia en el sensor (1b) y geolocalizados (1c).

El número de bandas de la imagen AHS es de 80, que cubren el rango espectral de longitudes de onda de 0.43 a 12.7 micrómetros, y el total de imágenes adquiridas han sido tres. La primera de ellas está tomada a una altura media de 793.9 metros y tiene un tamaño de píxel de 6x6 metros. La siguiente está tomada a 775.2 metros de altura media, con un tamaño de píxel de 4x4 metros y, finalmente, la última está tomada a 757.5 metros de altura media y el tamaño de píxel es de 2x2 metros. El área comprendida por estas imágenes es un segmento bastante estrecho perteneciente a Granada capital y parte de la Vega, como se ve en las figuras 2.73 y 2.74, y comprende la altura máxima y mínima de vuelo y, por tanto, la máxima y mínima superficie registrada en las imágenes.

Algunas anotaciones interesantes que pueden hacerse a simple vista son la diferenciación de cultivos (de color rojo) en el margen oeste de la autovía A-44 con el margen este, así como con la vegetación natural perteneciente a la zona de la Alhambra. Estas imágenes permitirían, por su elevada resolución espectral, diferenciar los diferentes tipos de cultivo muy exactamente y elaborar una cartografía de detalle. Sin embargo, el elevado coste sólo nos permite realizar un estudio meramente experimental de zonas piloto como es el caso de estas imágenes.

2.6. Marco socioeconómico

Históricamente la comarca de la Vega de Granada ha sido un área rural, pero recientemente, debido a la crisis generalizada del sector agrario y a la pérdida de importancia del mismo dentro de la economía general del país, está sufriendo importantes transformaciones. La comarca de la Vega de Granada, al tratarse de un núcleo periurbano, se ve fuertemente afectada por los procesos de crecimiento demográfico y expansión urbanística del núcleo principal de población, que es la ciudad de Granada, por lo que se establece una seria competencia entre los usos tradicionalmente agrarios y los usos residenciales, recreativos e industriales. El incesante desarrollo urbano de la ciudad de Granada, al que se une, desde la década de los años 80, el de gran parte de los municipios de la Vega, que constituyen su área de influencia más inmediata, explica la magnitud de tales transformaciones y la celeridad de la evolución [121, 1997].

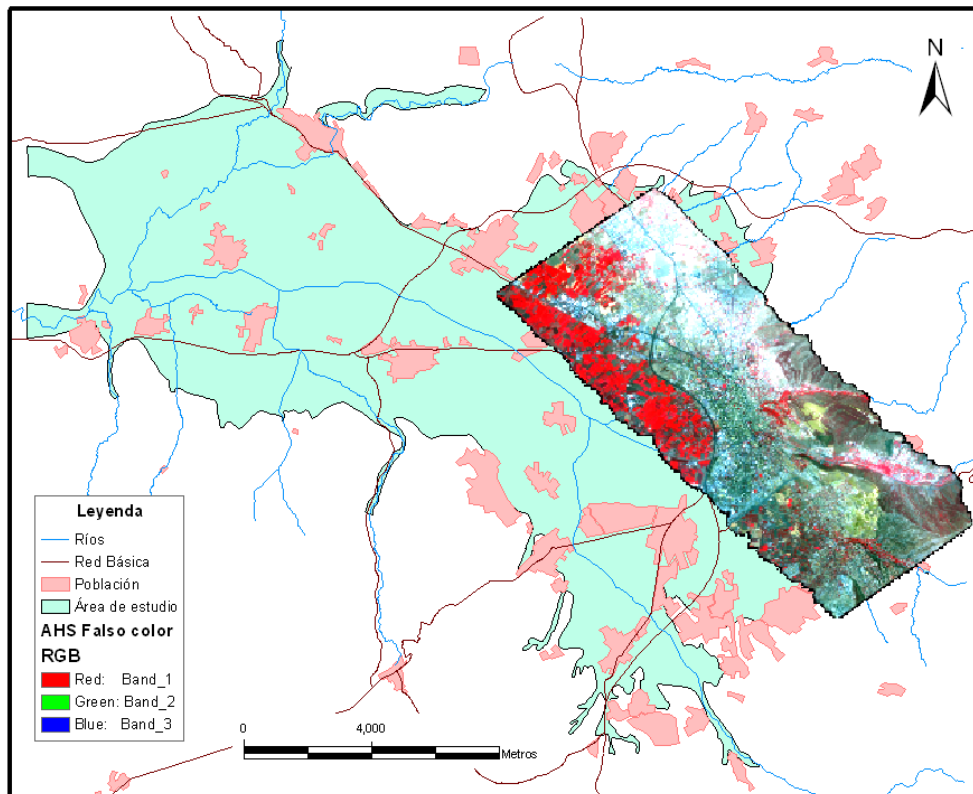


Figura 2.73: Imagen AHS en "falso color" (Ch:13,Ch:8,Ch:4) con resolución espacial 6x6 m.

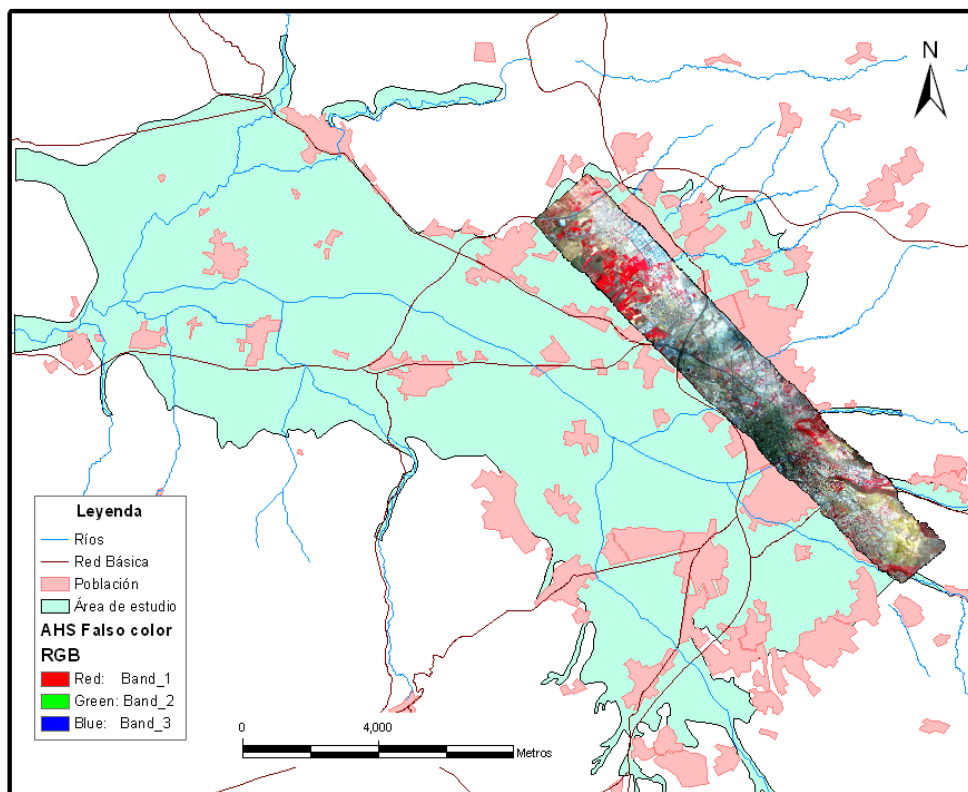


Figura 2.74: Imagen AHS en "falso color" (Ch:13,Ch:8,Ch:4) con resolución espacial 2x2 m.

El área metropolitana de Granada, también conocida como cinturón de Granada, es una aglomeración urbana formada por 55 municipios de la provincia, alrededor de su núcleo central que es ciudad de Granada. La población total asciende a 569.423 (INE, año 2006), a lo que hay que añadir aproximadamente 40.000 estudiantes foráneos residentes la mayor parte del año. De estos 55 municipios, el área de estudio comprende 41 de ellos, para los cuales la distribución de la población por densidad queda reflejada en la figura adjunta (2.76). Se aprecia una mayor densidad en Granada capital y los pueblos adyacentes, que disminuye conforme aumenta la distancia a la capital.

Debido al auge de la construcción en los últimos años y al aumento del precio de las viviendas en Granada capital, la población de los núcleos del cinturón se ha visto incrementada, y desde el año 2001 supera a la población de la capital (figura 2.75).

Otro dato interesante respecto a la economía de la zona es la distribución tanto del sector agrario como del sector ganadero. Los datos mostrados en el presente documento, tratan de dar una idea general de la distribución de la agricultura, así como de la distribución ganadera. La figura adjunta (2.78), muestra la distribución espacial de las hectáreas de cultivo de cada municipio, así como una distribución de los cultivos prioritarios. De este modo, podemos distinguir que en la zona sur el cultivo mayoritario es el de frutales, al igual que sucede aisladamente en Fuente Vaqueros, mientras que en la zona norte el porcentaje mayoritario de la agricultura corresponde al olivar.

La distribución de la ganadería se indica en la figura 2.77, que nos permite concluir que la mayor cantidad de unidades ganaderas corresponde a la zona central del área de estudio, cantidad que va disminuyendo hacia los bordes noreste y sureste. Según la distribución por tipo de ganado se puede decir, a grandes rasgos, que existe una concentración de explotaciones avícolas en la franja central del área de estudio, y que las mayores concentraciones de ganado se dan en Granada capital y Fuente Vaqueros, en concreto de ganado bovino casi en exclusiva.

Otro dato significativo es el consumo de agua en invierno y en verano, ya que éste es un elemento objetivo de análisis y valoración de las dotaciones locales. Para tener una idea general del consumo de agua, se presentan dos mapas, cuyas unidades están en $m^3/día$, y cuya espacialización se muestra en las figuras 2.79 a y b. En estas figuras no se aprecia una diferencia notable en relación con el consumo de verano, de modo que los que consumen más en invierno, también lo hacen en verano. Aún así, destacan los municipios al norte de la Vega de Pinos Puente, Caparacena y Albolote, con un consumo bastante elevado junto con Granada, Vegas del Genil (Ambroz, Belicena y Vegas del Genil), Armilla y Monachil.

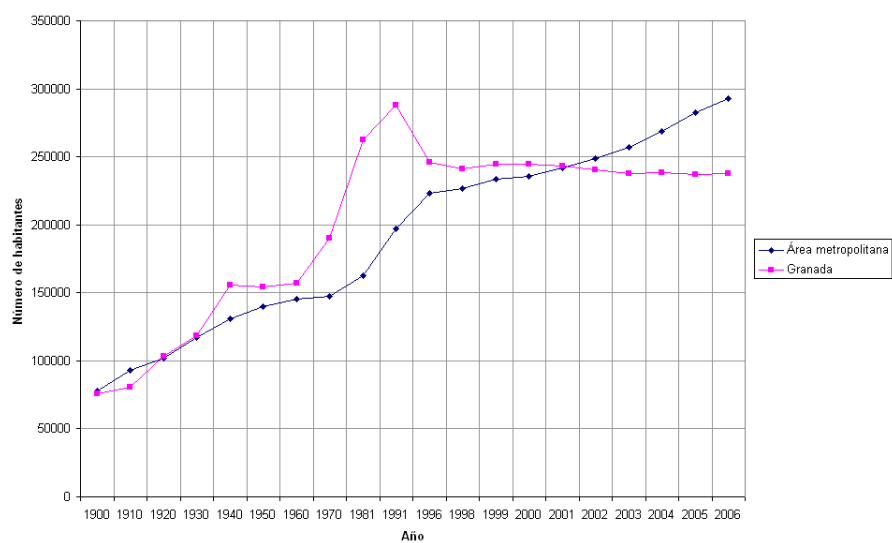


Figura 2.75: Evolución de la población 1900-2006 en los municipios de la Vega de Granada.

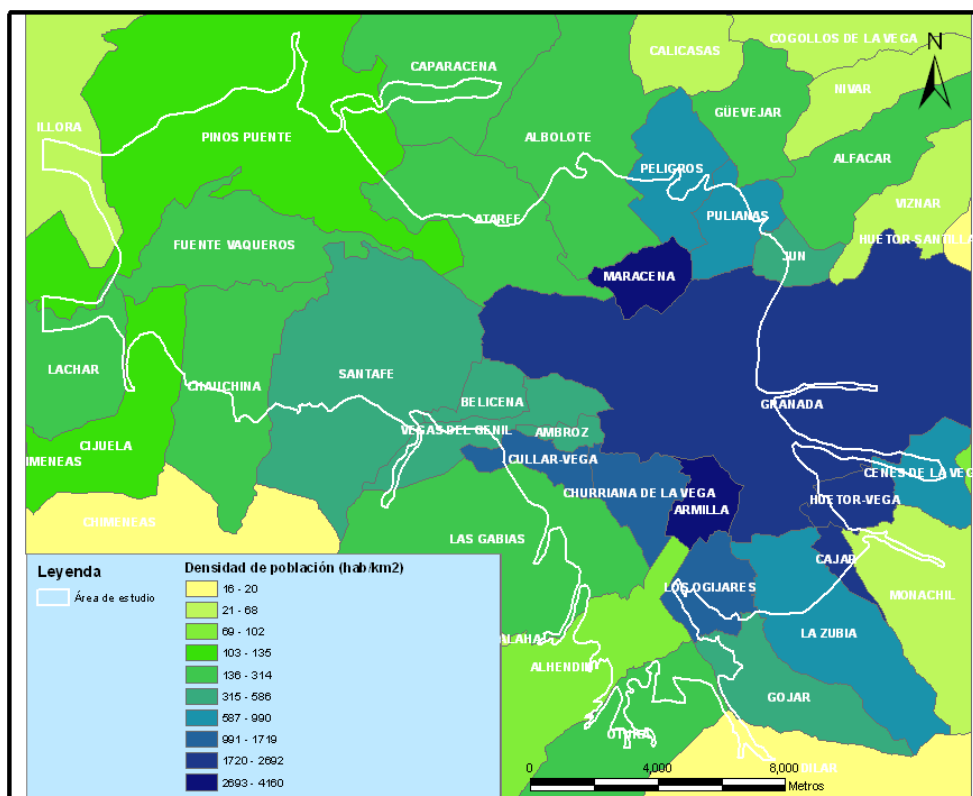


Figura 2.76: Densidad de población por municipios en la Vega de Granada (enero 2005)

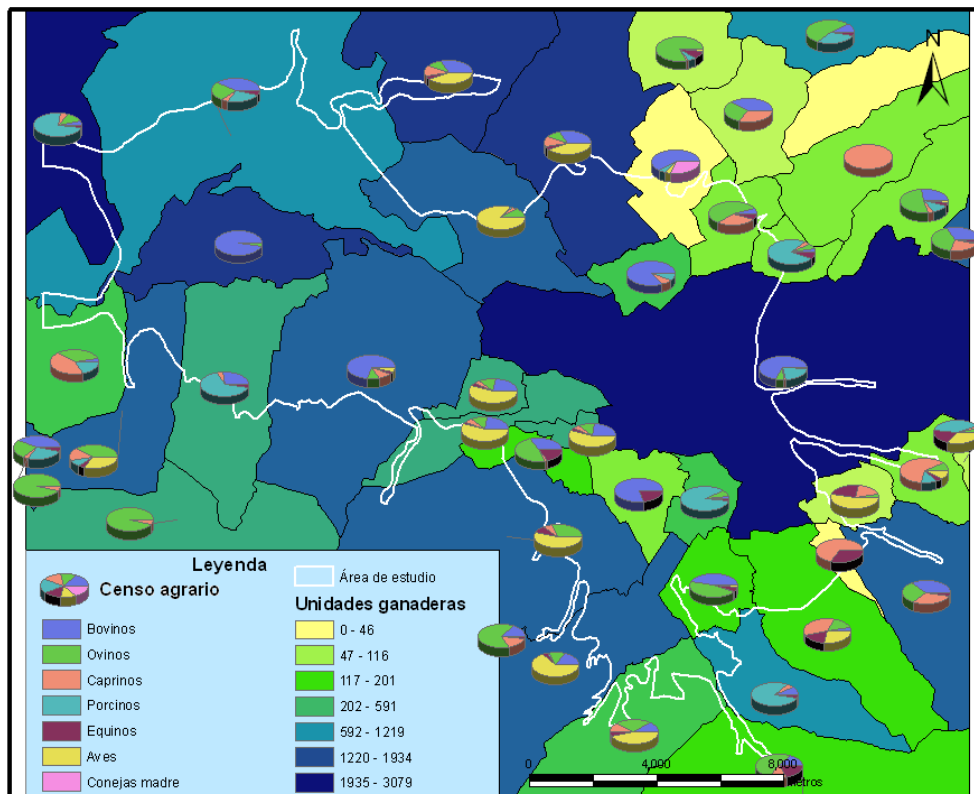


Figura 2.77: Unidades ganaderas y porcentaje de cabezas de ganado por tipo en los municipios de la Vega de Granada (1999).

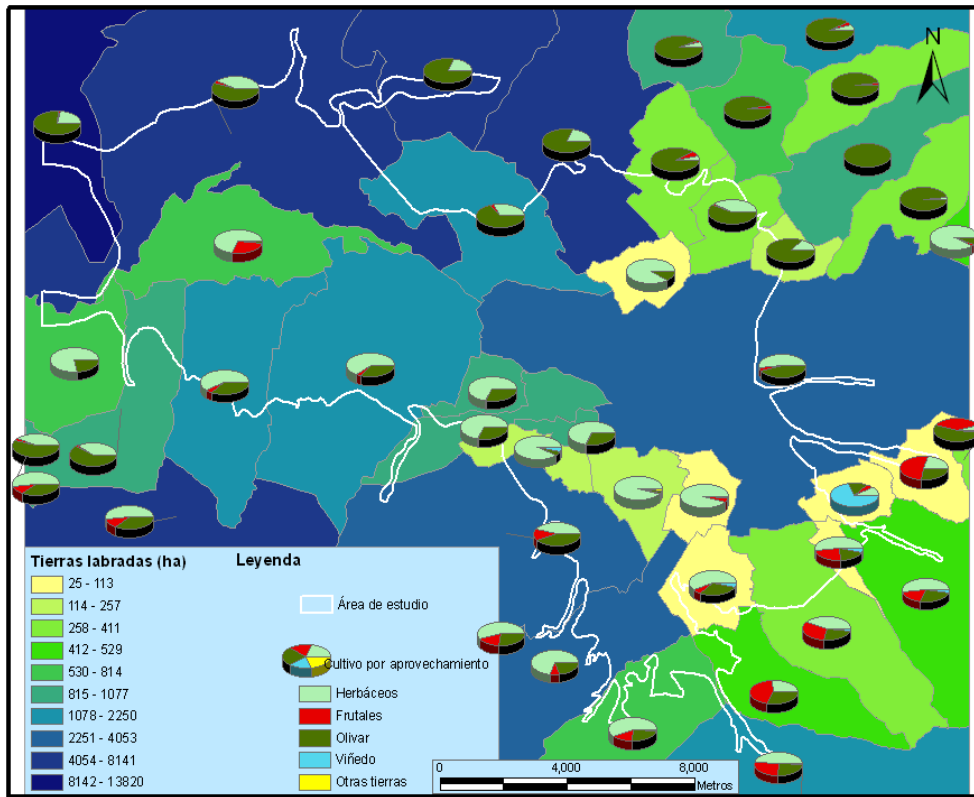
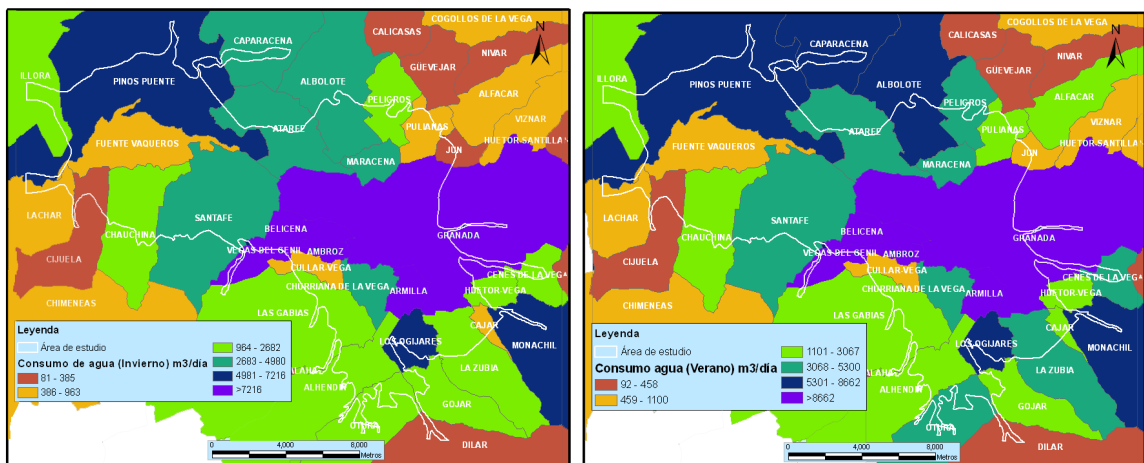


Figura 2.78: Explotaciones agrarias en hectáreas y aprovechamiento por cultivo en los municipios de la Vega de Granada (1999).



(a) Consumo de agua en invierno.

(b) Consumo de agua en verano.

Figura 2.79: Distribución del consumo de agua ($m^3/día$).

2.7. Marco ambiental

El marco ambiental del área de estudio está recogido dentro del Plan de Ordenación del Territorio de la Aglomeración Urbana de Granada (Publicado en Boja de 9 de Enero de 1999) como espacio sometido a protección y mejora, enfocando principalmente su actuación en los siguientes términos:

- Protección de su paisaje: Búsqueda de estrategias que racionalicen el consumo de los recursos ambientales origen de la pérdida de los valores paisajísticos.
- Protección de sus recursos hídricos: Racionalización del agua como recurso escaso e imprescindible.
- Potenciación de su uso público: Adecuación la estructura territorial a la ciudad funcional y construcción del sistema de espacios libres de Granada y su entorno.
- Protección de sus valores productivos: Potenciación de los valores productivos territoriales, es decir los suelos de la vega, que presentan un alto valor productivo agrícola, con actividades propias de su naturaleza y evitar aquellos que puedan ocasionar la pérdida de los mismos.
- Potenciación de su capacidad productiva: Puesta al alcance de objetivos y medios que racionalicen el sistema de regadío de la vega.

Dentro de este marco de protección, cabe destacar el interés ambiental, geológico y paisajístico del parque de Sierra Elvira, donde se sitúan los baños de Sierra Elvira, conocidos desde época romana, así como el resto de manantiales de la Vega de Granada los cuales están sufriendo su desecación a lo largo de los años.

Además de ésta iniciativa, existen diferentes grupos que actualmente han presentado la propuesta de declaración de la Vega de Granada como sitio histórico, y es que el área de estudio dispone de una indiscutible unidad patrimonial propiciada tanto por sus características morfológicas y productivas, como por el incuestionable reconocimiento social y ciudadano de un territorio diferenciado y unitario. El río Genil en este área se ha convertido en el elemento aglutinador de todo el ámbito, entre otras razones, por estructurar las unidades del relieve y por incidir en la disposición de los asentamientos urbanos tradicionales. Sin embargo, en la actualidad, la transformación de éste río, debido a la contaminación, fundamentalmente urbana y también agrícola e industrial, ha producido efectos contrarios, convirtiéndolo en un lugar no tan deseado tanto para asentamientos como para ocio o actividades agrícolas o ganaderas.

Algunas de las áreas de interés, para su reconocimiento como sitio histórico, pasan desde la naturaleza de sus suelos o de su vegetación, especialmente las choperas, hasta los culturales, donde se encuentran vestigios arquitectónicos, técnicos y espaciales de

la malograda industrialización emprendida en torno a la remolacha a primera mitad del siglo XX, que se constatan por las azucareras como la del Genil o San Isidro en Granada, y la de San Pascual y Nuestra Señora del Carmen en Pinos Puente.

Debido a la evolución de las diferentes fases históricas, sobre la Vega de Granada se ha dibujado un densa red hidráulica con numerosos restos constructivos de importante valor, Partidor (Alitaje), Azud de la Media Luna y Molino de Santa Matilde (Pinos Puente), Acequia del Molino Nuevo y Molino de la Aurora (Otura), Torre de Roma (Chauchina), etc. Además de relevante y diverso patrimonio arquitectónico y urbano (iglesia de Santa Ana en Ogíjares, Baños árabes en Churriana, Palacio del Cuzco en Víznar, etc), la enorme variedad y riqueza de la arquitectura agrícola o rural (cortijos, huertas, explotaciones agrícolas, etc) y finalmente el interés y diversidad de las actividades, usos, técnicas y conocimientos asociados a las diferentes formas de explotación agrícola de la Vega.

Todo este conjunto de valores, que poco a poco se están empezando a reconocer, requieren por lo tanto un cuidadoso estudio para su mantenimiento y restauración. Este fin es principalmente uno de los objetivos de la ordenación del territorio, pero para una correcta ordenación del territorio es necesario un conocimiento intenso y profundo de todas las variables que históricamente han propiciado la evolución de la Vega, así como de las que actualmente están impidiendo una correcta explotación. Estas variables son los suelos, las aguas (subterráneas y superficiales) y los focos de contaminación actual. Así, este trabajo contribuye no sólo bibliográficamente al catálogo de estudios sobre la Vega, sino también activamente a una correcta ordenación.

2.8. Otros datos complementarios

Además de la información experimental explicada en los epígrafes anteriores, se han incluido otros datos de naturaleza diversa en la base de datos del Sistema de Soporte de Decisión, algunos de ellos ya mostrados en los apartados de socioeconomía y marco hidrogeológico.

Estos datos han sido recopilados de muy distintas fuentes. La naturaleza de toda esta información puede ser consultada en los metadatos de la base de datos como se verá en el capítulo 4. Entre los más relevantes cabe destacar:

- Modelo digital del terreno (MDT): realizado por la Junta de Andalucía en el año 2002 con una malla regular de 10 x 10 metros de espaciado (resolución planimétrica).
- Focos de contaminación del área de estudio: digitalizados a partir de la cartografía 1:10.000 y GPS de campo. El total de puntos controlados es de 456, diferenciando

los dentro de siete grupos: actividades ganaderas, industriales, mineras, cementerios, aguas residuales urbanas, residuos sólidos urbanos y estaciones depuradoras de aguas residuales (figura 2.80).

- Hidrogeología, geología, usos del suelo y cultivos y aprovechamientos del área de estudio y el mapa de cultivos y regadíos: incluidos igualmente en la base de datos del sistema de soporte de decisión, y correspondientes respectivamente a la digitalización de los mapas hidrogeológico de Antonio Castillo [30], los mapas geológicos 1009 y 1026 editados por el IGME, de la serie Magna 1:50.0000, el mapa de usos y aprovechamientos (Ministerio de Agricultura y Pesca), el mapa de usos del suelo (Proyecto LUCDEME) y el mapa de cultivo y regadíos (Ministerio de Agricultura y Pesca).
- Datos climatológicos: series pluviométricas desde los años 1956 al 2003, de precipitación anual y mensual, y máxima mensual; temperaturas máxima y mínima absoluta total y mensual, y mínima y máxima media; evaporación total mensual. Todos estos datos recopilados por el Instituto Nacional de Meteorología.
- Datos foronómicos: datos de aforo pertenecientes a la Confederación Hidrográfica del Guadalquivir.
- Datos de estructuras humanas: autovías, carreteras, poblaciones, términos municipales, red de ferrocarriles, etc, que fueron extraídos de la Ortofoto digital de Granada y del Modelo Digital del Terreno de Andalucía.

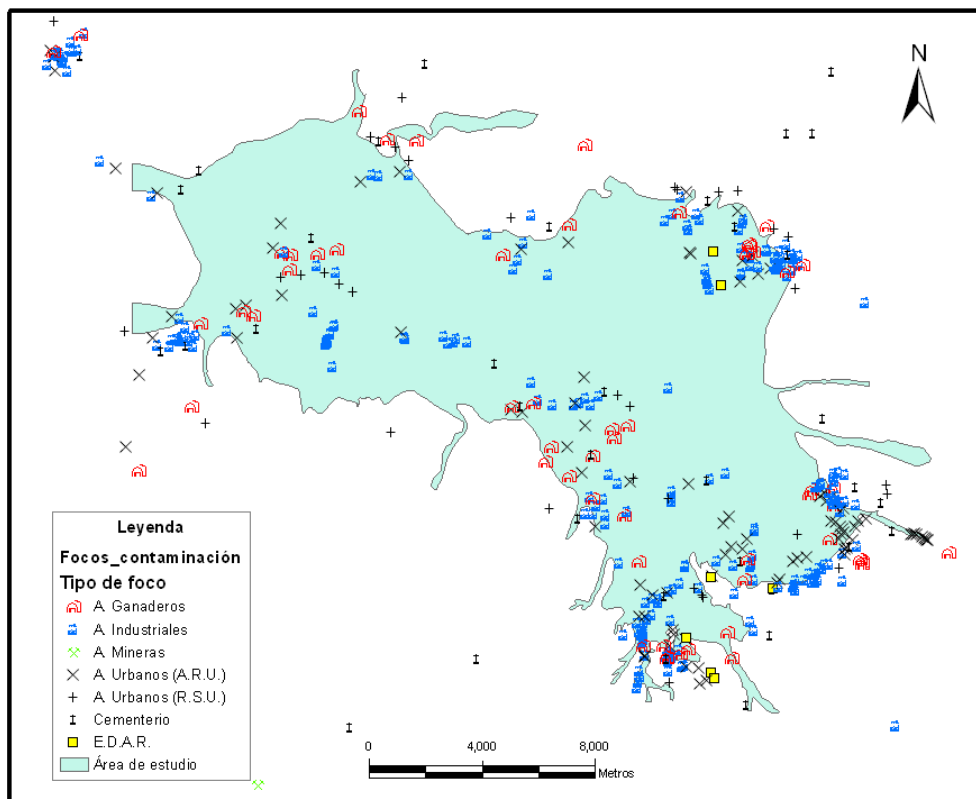


Figura 2.80: Distribución de los focos de contaminación en la Vega de Granada.

Capítulo 3

Bases metodológicas para el desarrollo del S.S.D.

3.1. Introducción

El desarrollo de un proyecto de investigación de la calidad del agua subterránea tanto para consumo público como para otro tipo de usos (industrial, agrícola, etc.) debe tener en cuenta diversos factores propios del acuífero: hidroquímicos, edafológicos, geológicos, etc. Además, deben considerarse los factores antrópicos relacionados con la explotación de este recurso natural, y de las actividades agrícolas, industriales, etc. que se desarrollan en superficie y que puedan conducir a problemas de contaminación del agua subterránea.

Por ello, y desde la perspectiva de este trabajo de investigación, considero como aspecto inicial necesario y fundamental la creación de una base de datos espacial que recoja la diversidad de información ("knowledge base") relativa tanto a la naturaleza del acuífero como a la explotación socioeconómica del mismo. No cabe duda que disponer de una visión conjunta de toda esta información constituye una ayuda inestimable para que los gestores tomen las mejores decisiones.

El término de "basado en el conocimiento" o "knowledge base" es comúnmente entendido como un cuerpo de conocimiento sobre algún tipo de problema. En este trabajo, usaremos el término en su más estricto sentido para significar un cuerpo organizado de conocimiento que provea una especificación formal lógica para la interpretación de la información (Walters y Nielsen, 1988 [187]). En este sentido estricto, las bases de conocimiento pueden ser entendidas como meta-databases y constituye el pilar principal de nuestro soporte de decisión, ya que la base de datos con núcleo SIG almacena y organiza la información disponible y necesaria para una gestión óptima de los recursos hídricos subterráneos.

De este modo, la la recopilación de la información relevante referida al área de estudio, expuesta en el capítulo anterior, es una de las tareas más importantes. Esta tarea aunque pudiera parecer "a priori" sencilla y sin ningún tipo de dificultad, realmente conlleva la mayor parte del esfuerzo. En muchos casos el recopilar la información almacenada por distintas administraciones supone trabajar con distintos formatos y con escasa información sobre el origen de la información, por lo que la estandarización e integración de la información supone en sí misma una tarea de investigación. A lo largo de este trabajo, nos hemos visto con numerosas dificultades tales como: la diversidad de los distintos sistemas de georreferenciación usados; mapas temáticos sin coordenadas o sin leyenda; numerosos y extensos listados de características almacenados a lo largo de los años con formatos heterogéneos o con información incompleta; distintos criterios en la elaboración de un mismo tipo de informe o ficha técnica; y otros muchos problemas. Por este motivo, la homogeneización de toda esta información para su posterior uso ha sido una de las tareas más importantes en la generación del Sistema de Soporte de Decisión de la Vega de Granada. En los estudios e informes consultados hemos encontrado diversa cartografía generada por distintos métodos, sin embargo, por cuestión de homogeneización, actualización y mejora, en la mayoría de los casos se ha optado por descartar la cartografía antigua y, en los casos posibles, volver a generarla mediante nuevos y sofisticados métodos. No obstante, en todos los casos ha existido una comparación y evaluación de los resultados que ganaron detalle mediante los nuevos tratamientos.

Esta primera parte de recopilación, también lleva implícita la generación de nueva información, fundamentalmente en lo que se refiere al paso de información puntual (experimental) a continua (mapa), que se trata con más detalle en el apartado de métodos geoestadísticos y en el de análisis espacial. En resumen, podemos decir que la creación de la base de datos espacial por lo tanto ha consistido en la recopilación, homogeneización y elaboración de toda la cartografía disponible, bajo un mismo "encuadre", haciendo posible así la superposición y la comparación de los distintos atributos estudiados. Esta tarea se ha hecho minuciosamente detallando el origen o "camino" de cada uno de los datos (metadatos), ya que se considera fundamental conocer este tipo de detalles para la utilización correcta de la información recopilada ya sea, en el marco de este proyecto o trabajo o para futuros usos de la misma.

La segunda parte de este capítulo está dedicada a las principales disciplinas de análisis espacial y temático, en las que se fundamenta este Sistema Soporte de Decisión. El estudio de los factores que intervienen y afectan a la calidad del agua requiere herramientas y métodos analíticos específicos de diversa naturaleza orientados al área de conocimiento concreta.

Se ha comenzado con los principales fundamentos de la Geoestadística, cuya principal aplicación es resolver el problema del paso de información puntual a continua, sin

dejar de mencionar su utilidad en aspectos como la creación de mapas de probabilidad o su versatilidad para estimaciones tales como el gradiente hidráulico, que permite combinar estimaciones a partir de datos de piezometría con datos experimentales del gradiente hidráulico. Otro de los aspectos interesantes integrados en el Sistema de Soporte de Decisión es la inclusión de imágenes de satélite. El tratamiento de estas imágenes se detalla en el apartado de Teledetección y su uso está orientado a la cartografía de las áreas de regadío de la Vega. La identificación de las áreas de regadío permite realizar un control indirecto de los volúmenes de agua necesarios para el desarrollo de los cultivos cuantificando los volúmenes de agua empleados para cada uno de ellos. No obstante, una cuantificación de detalle sobre los volúmenes procedentes de las aguas superficiales y de las aguas subterráneas para su uso en cultivos requiere de otros datos adicionales con los que no se cuenta en la actualidad, pero que son posibles de integrar en la base de datos SIG. Estos datos adicionales estarían referidos a datos de aforos en la red de acequias, cuantificación de extracciones de pozos realizadas por las distintas comunidades de regantes, etc.

Los métodos basados en Inteligencia Artificial aplicados en este trabajo han sido la lógica difusa y las redes neuronales. La lógica difusa como sistema basado en reglas difusas permite la modelización de sistemas mediante el uso de un lenguaje de descripción basado en predicados difusos del tipo SI-ENTONCES. Este tipo de modelos basados en el razonamiento humano permite interpretar fácilmente los modelos capturando la vaguedad del problema así como su incertidumbre, por lo que resultan modelos idóneos para su uso en recursos naturales y específicamente en los recursos hídricos subterráneos. Las redes neuronales artificiales, por otra parte, tienen la capacidad de aprender y son, por tanto, herramientas de generalización utilizadas en el contexto de las aguas subterráneas.

En el último apartado de este capítulo se hace un breve resumen sobre la teoría de la Decisión, donde se esbozan los principales conceptos en los que se basan los métodos multicriterio.

Todas estas metodologías se han integrado en un entorno tecnológico basado en SIG, lo que permite la interrelación de la información almacenada y una sencilla recuperación. Finalmente, este conjunto formado por la "Información" y los "Métodos de Análisis" es lo que constituye el Sistema Soporte de Decisión.

3.2. Aspectos introductorios de la base de datos geoespacial

La base de datos geográfica (SIG) es el núcleo de este proyecto, por lo que es evidente que las posibilidades de utilización de esta herramienta dependen ciertamente la forma

de construcción de la base de datos. Esto es así, hasta el punto que una incorrecta organización de la misma puede dar lugar, incluso, a limitaciones severas en la adecuada utilización del sistema. Es por ello bien conocido que el éxito o el fracaso de un proyecto SIG está determinado por la robustez del diseño de la base de datos y, por tanto, debe dedicarse suficiente tiempo a esta actividad (Worboys, 1995 [194]). La importancia en la construcción de la base de datos SIG se debe a que los elementos de la misma están íntimamente interrelacionados y por consiguiente necesitan estar estructurados para una fácil integración y recuperación de la información. El tipo de aplicación SIG realizada en esta investigación responde a lo que se denominan SIG orientados a proyectos. En este caso concreto se trata de una base de datos relativamente voluminosa orientada a los recursos hídricos subterráneos. Los requerimientos de una base de datos de este tipo están lejos de las grandes bases de datos corporativas por lo que el proceso de diseño es, en consecuencia, relativamente sencillo. No obstante, las consideraciones para una base de datos genérica son válidas en el dominio de los SIG, y son especialmente relevantes en cuanto a los diferentes tipos de naturaleza de los datos que han de ser organizados y almacenados.

En general, el desarrollo de una base de datos SIG para la gestión de un acuífero depende en gran medida de la disponibilidad de los datos y de la facilidad de incorporación de los mismos en el SIG. No se debe olvidar que la adquisición y preparación de los datos para un proyecto SIG supone frecuentemente el mayor esfuerzo tanto en recursos como en tiempo de todas las etapas necesarias para su implementación (Burrough, 1986 [25]; Aronoff, 1989 [13]; Bosque Sendra, 1997 [21]). Este paso inicial en el desarrollo de un SIG implica una serie de procesos, como son, por ejemplo, el establecimiento de la extensión espacial del área de estudio, la decisión de una proyección de trabajo adecuada o el ensamblaje de los distintos datos espaciales y no espaciales para que éstos puedan ser usados finalmente en el estudio en forma digital.

La base de datos SIG, como cualquier base de datos alfanumérica, se ha diseñado adecuadamente para poder responder a las necesidades de las aplicaciones para las que se pretende utilizar (Aronoff, 1989 [13]). Su diseño debe proporcionar un marco general a la base de datos, así como identificar los datos esenciales y apropiados y filtrar los datos irrelevantes (Yeung, 1998 [195]). Como se ha señalado antes, es muy importante ser consciente de que el proceso cobra mayor importancia de forma paralela a la envergadura de la base de datos. El objetivo es, en definitiva, conseguir que la base de datos tenga la información necesaria, pero que no sea redundante y que esté adecuadamente representada. Las principales etapas que se han seguido para la construcción de la base de datos consisten en la definición de los objetos de interés, geográficos en este caso, en términos de sus características y sus relaciones con otros objetos. Estas fases son: (1)Diseño conceptual, (2) Diseño lógico y (3) Diseño físico.

La primera etapa de diseño conceptual de la base de datos consiste en la definición en términos amplios y genéricos del objetivo y los requerimientos de la base de datos. La etapa de diseño lógico implica la especificación de la visión que el usuario va a tener de la base de datos con una clara definición de los atributos y relaciones. La etapa de diseño físico conlleva la especificación de la estructura interna de almacenamiento y la organización de los ficheros de datos. A partir de las propuestas de Yeung (1998 [195]) las tres fases se pueden esquematizar de forma genérica como se comenta a continuación.

El diseño conceptual de la base de datos es básicamente el establecimiento o formulación de los requerimientos de la aplicación y la especificación de la utilización final de la base de datos. El diseño conceptual es independiente del hardware y del software y puede ser una lista con los objetivos de la misma. Se realiza habitualmente utilizando el modelo entidad-relación (figura 3.1) e implica cuatro niveles de trabajo: (a) identificación de entidades, (b) identificación de atributos, (c) determinación de relaciones, y (d) creación de un diagrama entidad-relación (E-R). El diseño lógico es el proceso en virtud del cual el modelo conceptual es refinado y consolidado, de forma que la base de datos propuesta es examinada y revisada de forma completa para identificar problemas como: (a) datos irrelevantes que no serán usados, (b) datos omitidos o no disponibles, (c) representaciones inadecuadas de entidades, (d) falta de integración entre distintas partes de la base de datos, (e) aplicaciones no soportadas, o (f) tiempo y gastos potenciales adicionales para revisar la base de datos. El producto final de este proceso es un esquema lógico, que es el desarrollo trasladando el modelo conceptual (diagrama) a un esquema adaptado al software que se va a utilizar. Finalmente, el diseño físico consiste en la determinación de cómo las tablas que van a almacenar los datos son definidas en términos de formato, requerimientos de almacenamiento (memoria, espacio en disco, etc.), localización física, para generar un esquema físico dependiente del software y del hardware a utilizar.

3.2.1. Tipos de bases de datos

Antes de proceder al diseño conceptual de la base de datos se ha escogido el modelo de datos más adecuado a nuestra aplicación. Para ello en este apartado se hace un breve resumen de los tipos de bases de datos, que se pueden resumir en: jerárquico, en red, relacional e híbrido.

a) Modelo jerárquico

El modelo jerárquico es una estructura tipo árbol, donde cada entidad "padre" puede contener distintas entidades "hijo", pero que sólo permite cada "hijo" tener un único "padre". De este modo, por ejemplo, cada municipio podría tener varios núcleos de

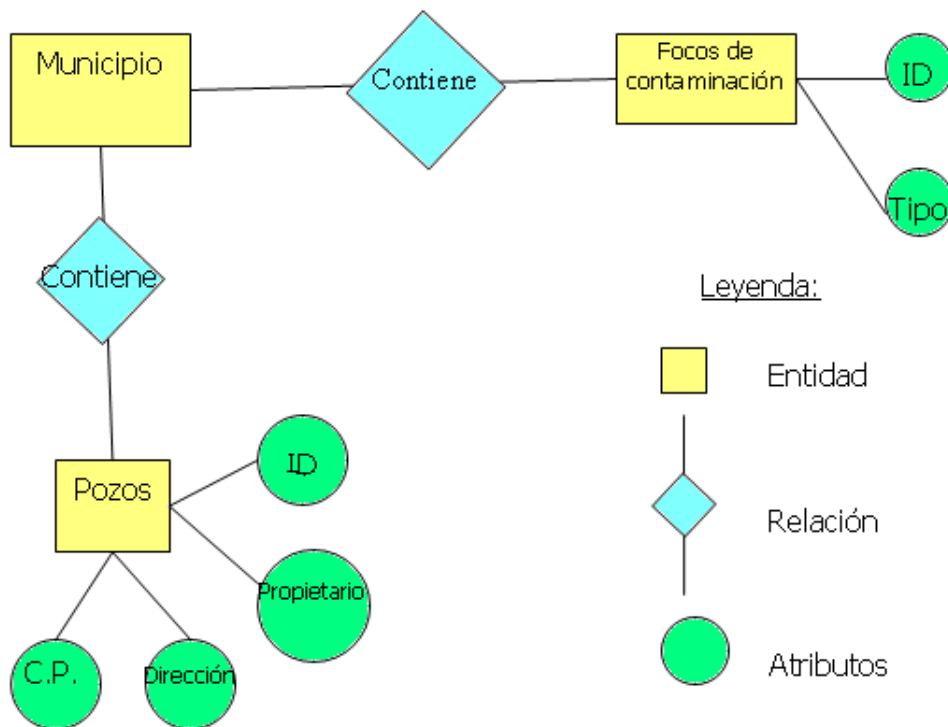


Figura 3.1: Un ejemplo del diagrama entidad-relación (E-R).

población que a su vez tuvieran diversos pozos de abastecimiento, pero cada uno de estos pozos sólo podría pertenecer a un núcleo de población, y cada núcleo a un municipio. Este modelo tiene el problema, por lo tanto, de que para poder representar múltiples pertenencias (por ejemplo, dos núcleos de población son abastecidos por el mismo pozo) los datos deberían de estar duplicados.

b) Modelo en red

Este modelo es similar al modelo jerárquico, excepto por el hecho de que los registros "hijo" pueden pertenecer a varios registros "padre", por lo que se resuelve el problema de las múltiples pertenencias resolviendo gran parte de los problemas del antiguo sistema. Sin embargo, al tratarse de un sistema basado en marcadores, el usuario y el programador, deben conocer profundamente el modelo de datos para ser capaz de "navegar" por los marcadores para obtener información. Por este motivo, a pesar de la gran difusión de este tipo de bases de datos en la industria, los fabricantes de SIG no han adoptado ninguno de estos dos modelos para su software, escogiendo en este caso el modelo de datos relacional.

c) Modelo relacional

El modelo relacional es un modelo basado en la teoría de Codd, que al contrario que los otros dos modelos anteriores cuyas teorías surgieron para explicar la realidad empírica. El modelo de Codd tiene tres elementos fundamentales un elemento estructural que describe la forma en que pueden guardarse los datos, un elemento de manipulación que describe el número de operadores que permiten a los usuarios procesar los datos y un elemento de integridad que propone reglas para asegurarnos que los datos se gestionan con validez y consistencia. El elemento estructural es el más sencillo, ya que todos los datos se almacenan en tablas, aunque con ciertos requisitos establecidos por las leyes de Codd, de este modo cada tabla consta de una cabecera y un cuerpo, así como de una clave externa para enlazar unas tablas con otras (ejemplo figura 3.2). El elemento de manipulación son los distintos comandos para operar sobre los datos almacenados, estos comandos operan sobre toda la tabla en lugar de hacerlo registro a registro, por lo tanto, la entrada de cada comando es una o más tablas y la salida es siempre una nueva tabla. La declaración matemática mediante el álgebra relacional se realiza mediante el lenguaje SQL (Structured Query Language). Un ejemplo de una sentencia formulada mediante este lenguaje en nuestra base de datos podría ser el siguiente:

```
SELECT FROM an_fq03 WHERE Nitratos>=50 AND Conductividad>=2500 ->
      pozos_contaminados
```

Esta consulta seleccionaría los pozos donde los nitratos y la conductividad superan los valores establecidos por la ley (50 ppm y 2500 $\mu S/cm$) y construiría una nueva tabla llamada "pozos_contaminados".

La integridad de una base de datos se refiere a la exactitud y grado de corrección de los datos almacenados en ella. Se suele hablar de dos tipos de integridad, la integridad de las entidades que hace referencia a la clave primaria y consiste en que cada una de estas tablas tenga una entrada válida y única y la integridad referencial que requiere que los valores de una clave externa existan como clave primaria en otra tabla.

Este modelo ha sido el elegido para nuestra investigación debido a que es el modelo mayormente integrado en los software comerciales y uno de los más sencillos con los que trabajar.

d) Modelo híbrido (orientado a objetos)

Aunque el modelo relacional ha tenido una gran aceptación por su sencillez, se considera que tiene algunas dificultades para capturar algunos aspectos de la realidad. Actualmente, las nuevas tendencias en el desarrollo de bases de datos persiguen nuevos modelos como las BBDD orientadas a objetos. Este tipo de modelo consigue un nivel

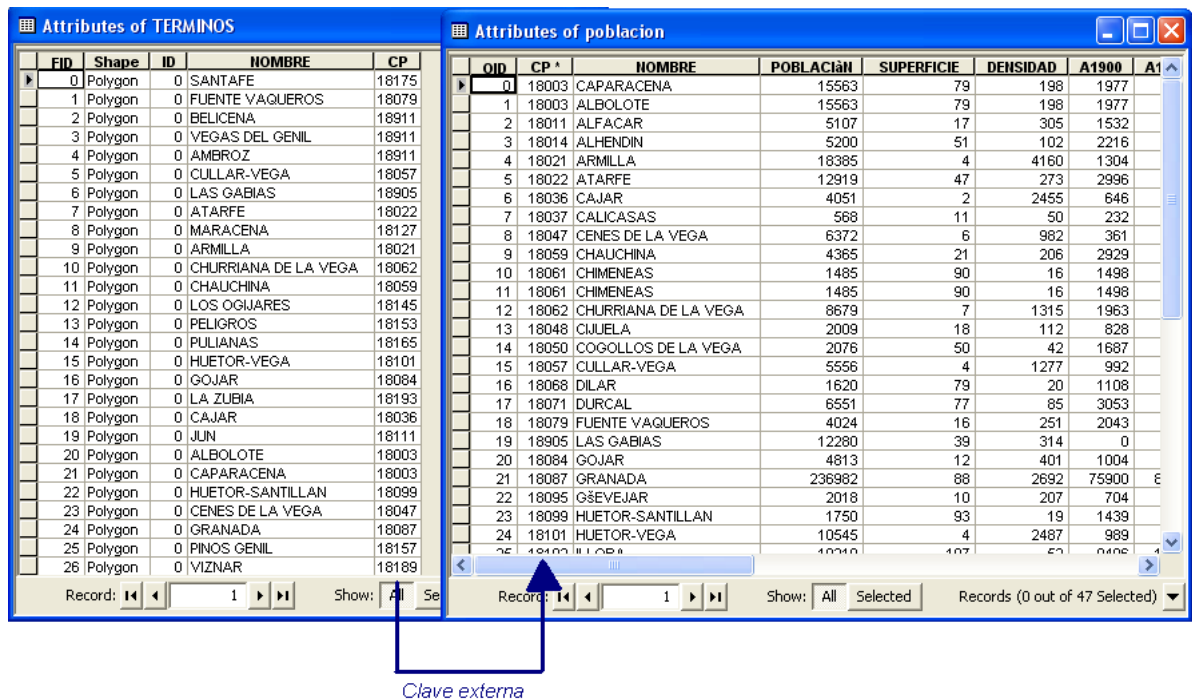


Figura 3.2: Ejemplo de elemento estructural de tipo virtual y clave externa de la base de datos.

de abstracción superior al relacional y lo acerca mucho más al mundo real, es decir, consigue representar las entidades y los eventos de una forma mucho más parecida a como los humanos los concebimos mentalmente.

3.2.2. Diseño conceptual de la base de datos

La creación de la base de datos de este trabajo comienza con el diseño conceptual. Este diseño considera las necesidades de la base de datos y las características y particularidades de los datos para así poder albergarlos con las máximas garantías de acceso, almacenamiento e integridad.

Para ello es necesario tener en cuenta que el proceso que se va a modelizar es un fenómeno natural complejo (calidad de las aguas subterráneas de un acuífero) en el que intervienen numerosas variables con diferentes características (zona no saturada, zona saturada, piezometría, puntos de agua, puntos de suelo, ríos, focos de contaminación, geología, etc.) Así, las entidades de la base de datos han de representar, en buena parte, campos continuos, como son: la piezometría, la concentración de determinado elemento químico en agua o suelo, el espesor saturado del acuífero o el porcentaje de arcilla de los suelos, aunque también objetos discretos como ríos, fallas o pozos. Al mismo tiempo, se ha considerado que muchos de los atributos de las entidades de la base de datos se

han obtenido mediante muestreo. Por tanto el tipo de datos utilizado para almacenar los atributos de las entidades espaciales será predominantemente de tipo intervalo o de razón, aunque también nominal u ordinal, además de topológico. Por otra parte, la base de datos no ha requerido grandes capacidades específicas para la gestión de datos multitemporales, ya que en la mayoría de los casos se trata de información estática para los objetivos de este proyecto. Aunque sí se ha tenido en cuenta la integración y actualización de ciertos datos como la concentración de elementos en el agua, en el suelo, la piezometría y los focos de contaminación. Por este motivo resulta conveniente convertir todos los datos a un mismo formato, en este caso a formato ráster (capa ráster), para facilitar así el análisis y modelización espacial (Aronoff, 1989; Berry, 1993; Berry, 1995 [13, 17, 18]).

Básicamente el diseño conceptual de la base de datos SIG requiere definir tanto las necesidades de la aplicación como el objetivo final de la base de datos. Generalmente, esto es definido de forma difusa, aunque suele cristalizar y evolucionar a la vez que la base de datos SIG progresa. Sin embargo, cuanto más claro y mejor definido esté el diseño conceptual más fácil es el diseño lógico de la base de datos SIG (Worboys, 1995; Date, 1995 [194, 51]). A lo largo de la memoria se han ido introduciendo ciertas necesidades, así como el objetivo final de la base de datos de los recursos hídricos. No obstante, algunos de los aspectos fundamentales que merecen ser considerados para el diseño conceptual son comentados a continuación.

Un primer aspecto a considerar es la especificación del uso último de la base de datos SIG como una sola declaración. En nuestro caso sería "base de datos SIG para la gestión de la calidad del agua de la Vega de Granada". Otro aspecto es el nivel de detalle de la base de datos SIG, indicando la escala o nivel de los contenidos de los datos de la base de datos. En general, una base de datos diseñada para un estudio a media o gran escala (por ejemplo, 1:50.000-1:10.000), como era nuestro caso, requerirá más detalle que una diseñada para aplicaciones a pequeña escala (por ejemplo 1:500.000).

También se deben determinar los elementos espaciales de la base de datos SIG. Éstos dependen obviamente del uso final de la misma y definen los conjuntos de mapas o datos espaciales que compondrán dicha base de datos. Los elementos espaciales son específicos de la aplicación y su elección en aplicaciones como ésta, tal como se ha argumentado a lo largo de la memoria, debería estar guiada por el modelo del fenómeno que se está estudiando. Los elementos espaciales están constituidos principalmente por mapas obtenidos de diferentes fuentes, que pueden ser categorizados en elementos primarios: que han sido digitalizados o introducidos directamente en la base de datos; y en elementos derivados: que son deducidos de los elementos primarios sobre la base de una operación espacial. Por ejemplo, los datos de muestreo hidroquímicos son elementos primarios, pero la superficie generada por interpolación mediante krigeaje a partir de esos puntos

es un elemento derivado. Esta distinción entre los elementos primarios y secundarios es útil en la estimación de la carga de trabajo en la creación de la base de datos y también en la planificación de las operaciones SIG. Cabe mencionar, que en la mayoría de los casos el elemento primario corresponde a una estructura vectorial (capa vectorial) y el derivado a una ráster (capa ráster).

Tabla 3.1: Ejemplo de algunos de los datos primarios y secundarios disponibles en la base de datos del acuífero de la Vega de Granada.

Nombre	Variable	Unidad	Origen
Variables del pozo			
Nº	Número de pozo	-	Datos del GPS
NOMBRE	Nombre del pozo	-	
NORTE	Coord Y	-	
ESTE	Coord X	-	
FECHA	Fecha muestreo agua	-	
Variables fisicoquímicas del agua subterránea			
CONDUCTIVI	Conductividad agua	microS/cm	Medidas tomadas in situ
PH	pH		
Temperatura	Temperatura del agua in situ	°C	
TURBIDEZ	Turbidez	UNF	Análisis de las muestras tomadas en puntos experimentales
CLORUROS	Cloruros	mg/l	
NITRATOS	Nitratos	mg/l	
NITRITOS	Nitritos	mg/l	
SULFATOS	Sulfatos	mg/l	
FOSFATOS	Fosfatos	mg/l	
BORO	Boro	mg/l	
CO3H_	Bicarbonato	mg/l	
CO3_	Carbonato	mg/l	
CALCIO	Calcio	mg/l	
SODIO	Sodio	mg/l	
MAGNESIO_	Magnesio	mg/l	
POTASIO	Potasio	mg/l	
AMONIO	Amonio	mg/l	
FLUORUROS	Fluoruros	mg/l	
Cromo_agua	Cromo	ppb	
Pb_agua	Plomo	ppb	
Cu_agua	Cobre	ppb	

(continúa en página siguiente)

(viene de página anterior)

Nombre	Variable	Unidad	Origen
Fe_agua	Hierro	ppb	
Mn_agua	Manganeso	ppb	
Sb_agua	Antimonio	ppb	
As_agua	Arsénico	ppb	
Variables químicas del suelo			
Ca_suelo	Calcio	mg/l	Análisis de las muestras tomadas en puntos experimentales
Cinc_suelo	Cinc	mg/l	
cloro_suelo	Cloro	mg/l	
Cobre_suelo	Cobre	mg/l	
Conductividad_suelo	Conductividad	mg/l	
Cromo_suelo	Cromo	mg/l	
Potasio_suelo	Potasio	mg/l	
magnesio_suelo	Magnesio	mg/l	
Manganeso_suelo	Manganeso	mg/l	
Nitratos_suelo	Nitratos	mg/l	
Nitritos_suelo	Nitritos	mg/l	
sodio_suelo	Sodio	mg/l	
sulfatos_suelo	Sulfatos	mg/l	
Variables del acuífero			
Litologia_zona _no_sat	Litología zona no saturada		Datos cedidos por el IGME
Transmisividad	Transmisividad del acuífero		
Permeabilidad	Permeabilidad del acuífero		
Espesor_no _saturado	Espesor no saturado	m	
Pendiente %	Pendiente	%	
alt_techo	Altura techo del sustrato	m	
Grad90	Gradiente en la dirección 90°		
Grad0	Gradiente en la dirección 0°		
Variables físicas del suelo			

(continúa en página siguiente)

(viene de página anterior)

Nombre	Variable	Unidad	Origen
Arcilla_h1	Arcillas horizonte primero (0-20 cm)	%	Análisis de muestras tomadas en puntos experimentales
Arcillas	Arcillas horizonte promedio (0-1m)	%	
Arcilla_hult	Arcillas último horizonte (80 - 100 cm)	%	
arena_h1	Arenas horizonte primero (0-20 cm)		
Arenas	Arenas horizonte promedio (0-1m)		
arena_hult	Arenas horizonte último (80 - 100 cm)		
limo_h1	Limos horizonte primero (0-20 cm)		
Limos	Limos horizonte promedio (0-1m)		
limo_hult	Limos horizonte último (80-100 cm)		
pH_h1	pH horizonte primero (0 - 20 cm)		
pH_suelo	pH horizonte promedio (0 -1 m)		
pH_hult	pH horizonte último (80 -100 cm)		
Hum_h1	Carbono orgánico horizonte primero (0 - 20 cm)		
% Humedad_suelo	Carbono orgánico horizonte promedio (0 -1 m)		
Hum_hult	Carbono orgánico horizonte último (80 -100 cm)		

(continúa en página siguiente)

(viene de página anterior)

Nombre	Variable	Unidad	Origen
CO_h1	% Humedad horizonte primero (0 - 20 cm)		
C_org_suelo	% Humedad horizonte promedio (0 -1 m)		
CO_hult	% Humedad horizonte último (80 -100 cm)		
H. Saturac_suelo	% Humedad de saturación		
Focos de contaminación			
ARU	Puntos de Aguas Residuales Urbanas	m	Digitalizados en pantalla y georreferenciados con GPS
EDAR	Estaciones Depuradoras de Aguas residuales	m	
Ganaderias	Granjas de ganado ovino, bovino y vacuno	m	
Intustria	Áreas industriales	m	
Minas	Puntos de zonas mineras activas	m	
RSU	Puntos de Residuos Sólidos Urbanos	m	
cementerio	Puntos donde se localizan cementerios	m	
aceq	Red de acequias	m	
Otras capas			
Uso_suelo	Usos de suelo	1:50.000	Digitalización - Proyecto LUCDEME
Geologia	Geología	1:50.000	Digitalización - IGME

(continúa en página siguiente)

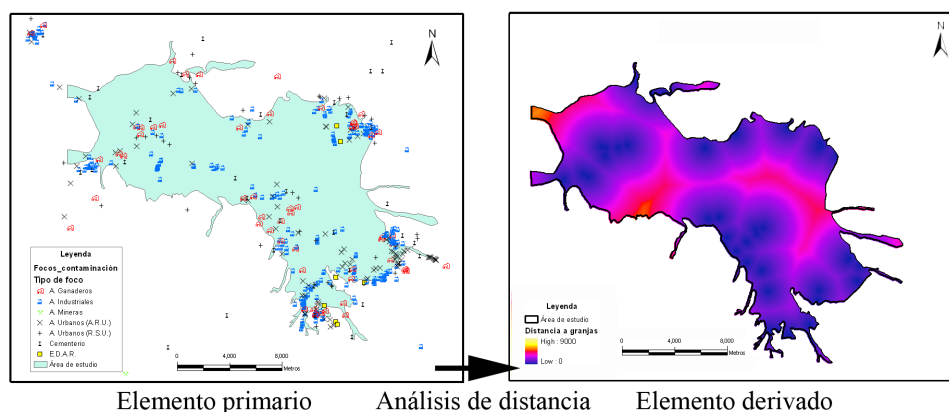


Figura 3.3: Ejemplo de elemento primario (capa vectorial) y derivado (capa ráster).

(viene de página anterior)

Nombre	Variable	Unidad	Origen
MDT	Modelo digital del terreno	1:20.000	Digitalización - Junta de Andalucía
Usos y aprovechamientos	Mapa de usos y aprovechamientos del suelo	1:50.000	Digitalización - Ministerio de Agricultura y Pesca
Cultivos_regadíos	Mapa de cultivos y regadíos	1:50.000	Digitalización - Ministerio de Agricultura y Pesca
Hidrogeología	Mapa hidrogeológico de A. Castillo		Digitalización

La tabla 3.1 muestra algunos de los elementos primarios más importantes contenidos en la base de datos de gestión de la calidad de las aguas subterráneas. Los elementos derivados de la misma corresponden en muchos casos a la interpolación geostadística de cada uno de éstos elementos primarios a un mapa ráster, así como a la generación de capas ráster mediante otras aplicaciones espaciales que resultaron interesantes como por ejemplo la creación de mapas de distancias a los focos de contaminación (ejemplo figura 3.3), mapas de aptitud del agua, mapas del módulo del gradiente hidráulico, mapas de vulnerabilidad, etc. Muchos de estos elementos derivados han sido ya mostrados en el apartado de datos y algunos otros se mostrarán en los sucesivos capítulos. Todos los datos incluidos en la base de datos y, fundamentalmente, los derivados, mantienen la homogeneidad en cuanto a área, tamaño de píxel y tipo de coordenadas, ya que esto facilitará la integración de las distintas capas mediante los diversos métodos estudiados (capítulo 5).

Como se ha comentado, la necesidad de interpolación en la mayoría de las capas primarias, que corresponde a elementos hidrológicos, edáficos o parámetros del acuífero, responde a la necesidad de una correcta representación de la realidad, ya que estos datos representan variables continuas en el espacio. La elaboración de cada una de las capas continuas a través de los diversos tipos de interpolación es una elección personal y orientada a los objetivos de este estudio. No obstante, es posible que para otro tipo de estudio se requiera de un tratamiento distinto de los datos, por lo que en la base de datos se ha incluido un fichero anexo donde se explica con detalle los pasos para la creación del mapa.

Además de los elementos espaciales definidos anteriormente, es necesario definir un apartado para otro tipo de datos complementarios no espaciales como las fotografías, perfiles de la zona no saturada, estadísticas relativas a la región, etc., que también se incluyen en la base de datos y que están asociados a elementos individuales de una capa. Por este motivo, ha sido necesario establecer una estructura que asocie cada una de estas tablas, datos, fotografías, etc. a cada uno de los elementos individuales al que se asocia, ya sea punto, línea o polígono (elementos topológicos vectoriales).

En este proyecto podemos distinguir varios grupos de datos según su origen. El primero está compuesto por los datos experimentales (muestreo de campo). Otro grupo de datos son los procedentes de diferentes administraciones y, por tanto, heterogéneos y con un tratamiento intermedio para su integración en la base de datos. Las principales administraciones son: el Instituto Geográfico y Minero de España, el Instituto Nacional de Meteorología, la Consejería de Medio Ambiente de la Junta de Andalucía y el Instituto Geográfico Nacional. El tercer grupo lo integran las imágenes de satélite y aerotransportadas que se adquirieron a través del Instituto Nacional de Técnicas Aeroespaciales (INTA). El último grupo corresponde a la digitalización de distinta información referente al acuífero, elaborada en tesis, tesinas o informes técnicos utilizados.

En algunos aspectos, la antigüedad de los datos puede ser un aspecto bastante importante para el diseño de nuestra base de datos, ya que define la edad de la misma y la hace útil o inútil para la aplicación final particular. En este punto podemos distinguir dos tipos de datos referentes a nuestro proyecto particular. Los datos relativos a los distintos parámetros del acuífero, los cuales no exigen un intervalo de tiempo dado su carácter estático, y los relativos a la calidad de las aguas y el suelo (parámetros químicos y microbiológicos fundamentalmente), así como los posibles usos del suelo. Estos datos, como ya se ha comentado con anterioridad, proceden directamente de las campañas de campo realizadas en el marco del proyecto, y los diferentes datos de calidad digitalizados corresponden en general a las fechas de estudio. Existe también otro tipo de dato digitalizado para fechas anteriores y que también se ha incluido en la base de datos, aunque su mayor importancia sin duda será el estudio temporal de los mismos,

el cual no ha sido elaborado debido a la elevada cantidad de datos ya disponible y el exceso de tiempo para su correcto tratamiento. Algunos estudios temporales sí incluidos han sido los mostrados en el apartado de datos, donde se han incluido solamente los parámetros relativos a la piezometría y precipitación.

Otro aspecto a tener en cuenta es el dominio de los datos espaciales en relación al marco básico de referencia de los distintos mapas o conjuntos de datos espaciales. En este trabajo todos los datos han sido creados o, posteriormente, transformados al sistema de coordenadas (Elipsoide Internacional, Datum Europeo) que era el estándar cuando comenzó a elaborarse este trabajo. La definición del área de estudio ya se definió en los primeros capítulos, y es otro de los elementos importantes en el diseño lógico de los elementos espaciales. Uno de los requisitos para la construcción de la base de datos fue el tener acceso a todos los datos originales en su mayor nivel de detalle e integridad, ya no sólo enfocando los usos en este proyecto, si no a posibles futuros proyectos. Este requisito se estableció debido a que los mapas derivados creados en este trabajo corresponden a unos objetivos concretos, por lo que el uso de los datos en otro tipo de proyecto necesitará la construcción a partir de los datos primarios de datos secundarios específicos y enfocados a los futuros usos.

3.2.3. Diseño lógico de la base de datos

En cualquier proceso de diseño, incluido el de un SIG, el desarrollo comienza con el análisis de los objetivos, pasando por niveles de detalle creciente, hasta que la información es recogida y se realiza la implementación. Al progresar en el diseño se incrementa el detalle, se añaden definiciones de datos y se asignan las estructuras de datos espaciales adecuadas. En este caso las estructuras proporcionadas por ArcGIS, que ha sido el software utilizado fundamentalmente para la realización de este proyecto.

Un punto importante a tener en cuenta es el modo en que las coordenadas van a ser almacenadas en el paquete SIG. La mayoría de programas SIG ofrecen un abanico de sistemas de coordenadas en función del sistema de proyección empleado. Es evidente que el sistema de coordenadas de la base de datos SIG debe estar en unidades apropiadas que representen las características geográficas en su verdadera forma y tamaño, por lo que el sistema de coordenadas quedará definido por el dominio espacial de la base de datos SIG. A la luz de lo anterior, se adoptó un área de estudio rectangular, correspondiente a un área ligeramente mayor a las dos hojas del IGN 1009 y 1026, con el sistema de proyección UTM (Huso 30). El resultado final es que todos los elementos espaciales de la base de datos están referenciados a un sistema de coordenadas uniforme, que permite la fácil integración de conjuntos de datos en los procesos de análisis espacial. Sin embargo, todos los mapas mostrados en este trabajo corresponden exclusivamente al "detrítico de la Vega", ya que a pesar de haberse realizado el estudio para el área mencionada, los

errores asociados a las diferentes interpolaciones y transformaciones son excesivamente elevados fuera del denominado "detrítico de la Vega" por la ausencia o poca densidad de datos experimentales.

En el primer paso del diseño de la base de datos lógica, se han identificado las entidades, esto es, los "objetos" o "cosas" sobre las que se necesita recopilar información y cuya existencia es "independiente". Para nuestro caso de estudio, estos objetos son los distintos muestreos y datos digitalizados en relación al acuífero, de modo que los pozos, sondeos y puntos de muestreo serían nuestras entidades así como los sondeos y localización de la información puntual referida a muestras. Otro tipo de datos importantes se refiere a distintas cartografías temáticas sobre vegetación, suelos, litología, geología, flora, etc. Por lo tanto, estas serán nuestras otras entidades identificadas como polígonos, siendo cada una de estas características una entidad así como los ríos, fallas, etc. que serán entidades lineales, entre el resto de información incluida en la base de datos.

En el siguiente paso se han identificado las relaciones entre las distintas entidades. Así, por ejemplo, el área de estudio se ha dividido en distintos municipios, y cada municipio contiene distintos pozos en los que se han muestreado cada una de las variables experimentales, distintos núcleos de población con su densidad de población, así como las características específicas de cada uno. Además, podemos identificar distinta información relativa a la geología, suelo, uso del suelo, ríos, acequias, etc. del área de estudio.

En último lugar se han identificado los atributos asociados a cada una de las entidades. Estos atributos, en el caso de los datos experimentales, se refieren al conjunto de características asociadas al punto experimental, y en el caso de las cartografías serán los distintos tipos de variables temáticas como: vegetación, litología, geología, etc., asociada a cada objeto poligonal. Estos atributos forman una colección organizada de códigos con su descripción correspondiente.

La relación entre los datos espaciales y no espaciales debe ser igualmente específica y definida. Estos enlaces e interrelaciones son un elemento clave para la organización de la base de datos SIG ya que definen las relaciones que pueden crearse entre los distintos elementos y tablas de datos (Gooldich y Kemp, 1990; Worboys, 1995; Yeung, 1998). En el modelo relacional de ArcGIS las relaciones se establecen definiendo un campo común (campo clave) en la tabla de la entidad espacial y la tabla de atributos (temáticos) que permita relacionar ambas. En la tabla 3.2 se muestra un ejemplo típico de las relaciones necesarias para enlazar una entidad espacial de la base de datos del acuífero de la Vega de Granada, como es la densidad del núcleo de población con el municipio o con los análisis químicos realizados para cada uno de los puntos establecidos.

Para concluir este apartado, debo hacer la siguiente consideración, referente al "uso" de la base de datos: ¿Esta va a ser usada por varios o por uno sólo usuario? Esta cuestión

puede complicar o facilitar mucho la tarea de diseño y acceso a la base de datos. En el caso de este proyecto de tesis, la mayor parte del tratamiento y almacenamiento ha estado orientado a un solo usuario que se encargaría de crear y actualizar la base de datos, por lo que las distintas modificaciones y actualizaciones no repercuten a terceros; de esta forma, el establecimiento del flujo de trabajo es mucho más sencillo.

3.2.4. Diseño físico de la base de datos

El diseño físico está relacionado con aspectos esencialmente informáticos como la estimación de la carga, requerimiento de espacio en disco, requerimiento de memoria, acceso, velocidad etc. Todo esto depende, como ya se ha comentado anteriormente, de la plataforma SIG en la que se opera, por lo que no hay estándares universales sobre los aspectos del diseño físico por lo que en gran parte ha de basarse en la experiencia.

Hoy día los requerimientos de espacio en disco no son un problema, ya que cada vez las capacidades son mayores, permitiendo almacenamientos más que suficientes. En el trabajo que aquí se presenta, un sólo disco de 350 GB permitió almacenar la totalidad de los datos utilizados en el proyecto (tanto primarios como derivados, aunque no algunos intermedios que se excluyeron de la base de datos por simplicidad). La carga de la base de datos es también difícil de estimar pero en este caso podemos hablar de aproximadamente un total de 444 ficheros incluyendo tablas, capas vectoriales y ráster.

En este trabajo se utilizó el sistema de base de datos implementado en ArcGIS 9.2, el *file Geodatabase*, que permitía un almacenamiento de hasta 1 TB, lo que se consideró más que suficiente para los requerimientos del proyecto. Para proyectos mayores se suelen usar sistemas gestores de bases de datos relacionales externos como Oracle, SQL, etc. pero hoy día con el avance de la informática, las implementaciones realizadas en este software han sido más que suficientes. En el sistema *file Geodatabase* se almacenaron las tablas con toda la información referente a las entidades espaciales, los datos no espaciales en forma de tablas o atributos, así como información sobre símbolos cartográficos, descripciones, etc.

La organización de la base de datos se muestra en la figura 3.4, donde se observa que se ha dividido el conjunto de la información en grupos de variables ráster y grupos vectoriales, fundamentalmente. El detalle de cada uno de estos conjuntos se explica con más profundidad en el apartado siguiente.

3.2.5. Implementación de la base de datos

Considerados los aspectos anteriores del diseño conceptual, lógico y físico, se procedió a la creación de la base de datos HIDROGIS. El modelo elegido para la implementación de la base de datos fue el *file Geodatabase* incluido en el software ArcInfo 9.2. De este

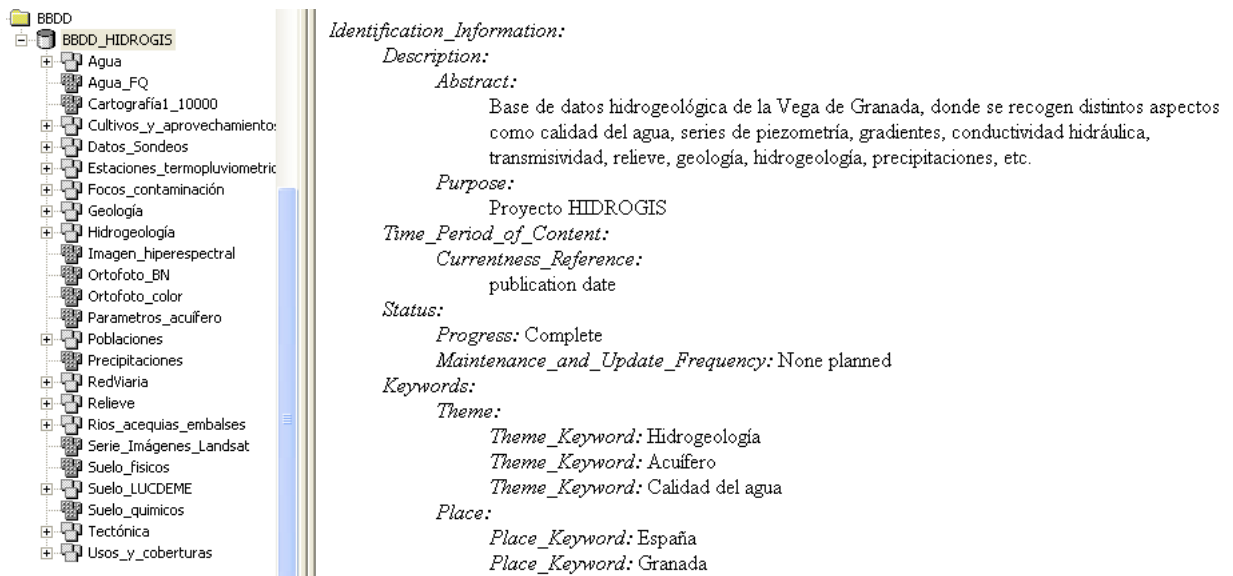


Figura 3.4: Organización de la base de datos HIDROGIS junto con algunos de los metadatos de la misma.

modo, una vez construidas la mayoría de las capas de información se agruparon las distintas variables en función de su origen y su tipo. Además, se crearon los metadatos correspondientes a estas cubiertas. Para cada capa realizada se intentó cumplir la norma ISO 19115:2003, la cual divide la información de los metadatos en 10 apartados: información acerca del propio metadato (lenguaje, punto de contacto, fecha, estándar, etc.), la identificación del dato (descripción, propósito, estado, resolución espacial, clasificación temática, etc.), la extensión (espacial, temporal y vertical), el sistema de referencia espacial (proyección, elipsoide y datum), la representación espacial (vectorial o ráster), la calidad (contexto de aplicación, origen y procesamiento del dato, precisión, etc.), las restricciones (uso, legales y de seguridad), el contenido (descripción del catálogo o cobertura), las citas y finalmente la distribución (formato, distribuidor, etc.). Esta implementación de los metadatos se realizó a través del propio software ArcGIS para el que ya existía una herramienta especializada a través del NEM (Núcleo Español de Metadatos) y recoge las directrices establecidas por las dos normas más importantes de metadatos, la ya nombrada ISO internacional 19115:2003 y la Dublin Core Metadata Element Set 15836:2003, teniendo en cuenta los proyectos más importantes en curso de generación cartográfica, la Directiva Marco del Agua y las iniciativas IDE regionales.

Un ejemplo de parte de los metadatos de una de las capas puede observarse en la figura 3.5, los cuales concretamente se refieren a una parte de la descripción de los datos, como ya se comentó anteriormente. Como puede verse, existen diversos apartados que describen a los datos que no se listan aquí por motivos de espacio, aunque otro ejemplo relativo a los datos de la base de datos se ve en la figura 3.4.

Description:
Abstract:
Datos fisicoquímicos recogidos en la campaña de Octubre del 2003 de la Vega de Granada
Purpose:
Proyecto HIDROGIS
Time_Period_of_Content:
Time_Period_Information:
Single_Date/Time:
Calendar_Date: 2004-2008
Currentness_Reference:
publication date
Status:
Progress: Complete
Maintenance_and_Update_Frequency: None planned
Spatial_Domain:
Bounding_Coordinates:
West_Bounding_Coordinate: -3.836788
East_Bounding_Coordinate: -3.562612
North_Bounding_Coordinate: 37.333949
South_Bounding_Coordinate: 37.084535
Keywords:
Theme:
Theme_Keyword_Thesaurus: HIDROGIS
Theme_Keyword: analisis fisicoquímicos
Theme_Keyword: Vega de Granada
Theme_Keyword: aguas subterráneas
Theme_Keyword: acuífero
Theme:
Theme_Keyword_Thesaurus: Análisis fisicoquímicos

Figura 3.5: Ejemplo de metadatos para una de las capas de información de la base de datos.

Se establecieron 24 grupos de variables (figura 3.4), entre los que se distinguen variables ráster y variables vectoriales. La mayoría de estos datos han sido comentados en el capítulo 2; no obstante, se ha considerado adecuado hacer a continuación un breve repaso a cada uno de los grupos de variables vectoriales creados, así como a su organización en la base de datos.

Aguas: este grupo de datos corresponde a los análisis hidroquímicos realizados para las aguas subterráneas del acuífero y adquiridos en las campañas de muestreo. Cada una de estas campañas se almacena en tablas individuales, diferenciando los datos de aguas superficiales y subterráneas. Todos los datos fueron georreferenciados con un GPS de la marca Trimble y posteriormente corregidos del error atmosférico con una estación de referencia alcanzando una exactitud de 0,5 m en las coordenadas de localización.

Suelos: igual que el caso anterior, este grupo de datos corresponde a los análisis edafológicos realizados para los suelos; poseen una exactitud en localización idéntica al anterior conjunto de datos. Uno de los principales problemas es que hubo que asignar a cada punto la información contenida en cada uno de los horizontes, de modo que se reflejara lo mas fielmente posible la información para cada punto, no siendo en todos los casos el número de horizontes el mismo. Una de las soluciones para su integración directa en el SIG (además de añadir tablas individuales para cada punto en la base de datos) fue la creación de capas continuas homogeneizando el total de los horizontes en uno solo de 0 a 1 m en función de su espesor, así como la creación de una capa correspondiente al horizonte superior (0-20 cm) y al inferior (80-100 cm) los cuales eran comunes en todos los suelos.

Cultivos y aprovechamientos: pertenecientes al mapa de cultivos y aprovechamientos del suelo de España a escala 1:50.000 inicialmente elaborado por la Secretaría General de Producciones y Mercados Agrarios y durante los años 1998 y 1999 actualizada para el proyecto SIG Oleícola.

Datos sondeos: Estos datos corresponden mayormente a la digitalización de los datos en papel correspondientes al proyecto de la FAO (1972 [1]) para el desarrollo agrícola de la Vega de Granada y a otra serie de datos de sondeos pertenecientes al IGME. En este apartado se distinguen cuatro tipos de datos: datos de altura del sustrato, datos granulométricos de la zona no saturada, datos de transmisividad y datos de permeabilidad. Los datos pertenecientes a la altura del sustrato proceden de diversas campañas geofísicas realizadas en el marco del proyecto anteriormente mencionado, así como puntualmente de algunos de los sondeos más profundos. Los datos granulométricos de la zona no saturada corresponden a los distintos informes realizados para cada uno de los sondeos tal y como se comentó en el capítulo de datos. Los datos de transmisividad y permeabilidad corresponden generalmente a datos de ensayos de bombeo y, en algu-

nos casos, a estimaciones a partir del espesor saturado correspondiente a ese ensayo en particular.

Piezometría: Los datos piezométricos incluidos en la base de datos corresponden, tanto a los muestreados en campo de las distintas campañas realizadas en el marco de esta tesis, como a datos de campañas más antiguas cedidos por el IGME.

Datos de estaciones termopluviométricas: estos datos corresponden a las precipitaciones máximas, medias, mínimas mensuales, anuales y diarias de toda la serie de datos disponible del Instituto Nacional de Meteorología, junto con los datos de temperatura media, máxima y mínima y de evapotranspiración para el mismo período. Este conjunto de datos, a pesar de ser un archivo digital en forma de texto y debido a su formato específico para la base de datos del INM, no permitía una importación o exportación directa, por lo que fue necesario crear un programa específico que leyera esa base de datos y separara cada una de las estaciones en ficheros individuales que se asignarían posteriormente a cada estación.

Focos de contaminación: los focos de contaminación provienen de la digitalización en pantalla a partir de la cartografía 1:10.000, así como de la observación en campo de diversos puntos (de vertido, fundamentalmente) georreferenciados con la ayuda de GPS. En este grupo de datos se han distinguido distintos tipos de focos, como son los correspondientes a actividades ganaderas, actividades industriales, actividades mineras, puntos de vertido de aguas residuales urbanas, puntos de vertido de residuos sólidos urbanos, cementerios y estaciones depuradoras de aguas residuales.

Geología: el grupo de datos de geología corresponde tanto a la cartografía Magna del Instituto Geológico y Minero de España, como a la geología realizada dentro del proyecto de exploración geológica minera de la cuenca lignitífera de Granada (Roldán García, 1981 [156]) y a la geología regional de Andalucía.

Hidrogeología: en este apartado fundamentalmente se encuentra la cartografía realizada por Castillo Martín (1986 [30]), así como la información relativa al tipo de material acuífero. Esta cartografía se encontraba en formato de papel, por lo que hubo que proceder a su digitalización con la ayuda de una tableta digitalizadora.

Poblaciones: la digitalización de los límites municipales pertenece a la cartografía incluida en la base de datos de la Junta de Andalucía incluida en la Ortofoto color. Los núcleos de población también pertenecen a esta cartografía, aunque hubieron de ser modificados para ampliar y detallar algunos de los núcleos más pequeños que no aparecían en la misma, mediante una digitalización en pantalla de la cartografía 1:10.000 y con la ayuda de la foto aérea, así como de las imágenes de satélite.

Red Viaria: este conjunto de capas vectoriales responde a las vías principales (autovías), secundarias (carreteras), otras vías, los nudos y el ferrocarril e igualmente procede

de los datos de la ortofoto color de la Junta de Andalucía con ligeras modificaciones y añadidos.

Relieve: en este apartado se incluye el modelo digital de elevaciones de la Junta de Andalucía para el área de estudio, así como las curvas de nivel generadas a partir del mismo y un mapa reclasificado de pendientes.

Ríos: este grupo de mapas vectoriales corresponden fundamentalmente a los ríos principales, secundarios, acequias y a los polígonos relativos a los embalses del área de estudio y su entorno. Estos datos han sido digitalizados a partir de la cartografía 1:10.000 del área de estudio, algunos de los cuales han sido digitalizados en pantalla y otros se han digitalizado en tableta digitalizadora.

Suelo LUCDEME: esta capa de información se ha obtenido mediante la digitalización con ayuda de tableta digitalizadora de la cartografía de suelos realizada en el proyecto LUCDEME para el área de la Vega de Granada.

Tectónica: esta capa también fue digitalizada mediante tableta digitalizadora a partir de los mapas geológicos de la serie Magna 1:50.000 del área de la Vega de Granada.

Usos y coberturas: Esta cartografía vectorial corresponde a la cartografía realizada por la Junta de Andalucía para el proyecto CORINE Landcover a partir de distintas imágenes Landsat de los años 1995 y 1999.

Imágenes de satélite: Los datos de las imágenes de satélite corresponden a las capas raster multiespectrales Landsat TM4, ETM+ y hiperespectrales AHS y a los resultados obtenidos con estas imágenes tras su tratamiento, como son los mapas de vegetación, la clasificación de los usos del suelo, etc.

En resumen, podemos decir que todos estos *Datos* son la base de nuestro Sistema Soporte de Decisión del acuífero de la Vega de Granada pues recoge prácticamente toda la información disponible relativa al acuífero. Corresponde a continuación exponer brevemente las bases metodológicas para el "Análisis" de los datos, con la finalidad de convertirlos en "Información" y con ello facilitar la toma de decisiones sobre la gestión de los recursos hídricos subterráneos de la Vega de Granada. A este menester se dedican los siguientes apartados de este capítulo.

3.3. Métodos Geoestadísticos de análisis espacial

El análisis espacial de datos constituye un pilar fundamental en el desarrollo de una aplicación SIG. Los métodos geoestadísticos son la herramienta más potente para este tipo de análisis pues se fundamenta en el estudio de la variabilidad espacial de las variables estudiadas.

En nuestro caso, la mayoría de los datos hidrogeológicos estudiados como el nivel piezométrico, la conductividad hidráulica, la transmisividad, la porosidad, el coeficiente

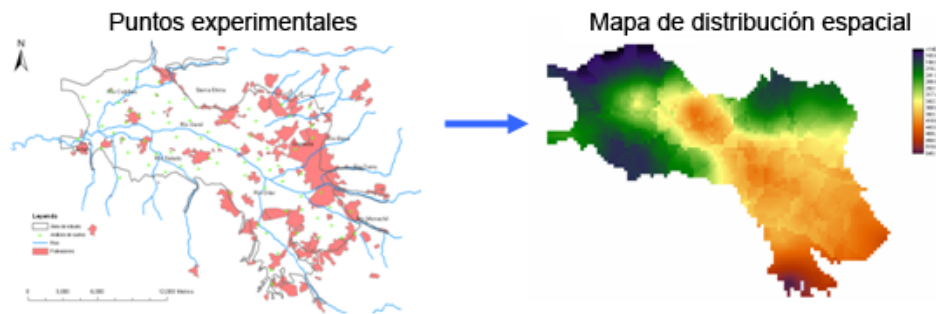


Figura 3.6: Ejemplo de una estimación geoestadística para una variable hidrogeológica experimental .

de almacenamiento, el espesor de una formación geológica, la infiltración, la precipitación, etc., son funciones que varían a lo largo del espacio, esto es, del acuífero. Esta variabilidad espacial puede ser mas menos importante, pero en general no es totalmente aleatoria, sino que tiene carácter estructurado o de correlación espacial (función variograma). Esta correlación espacial es la que nos muestra el carácter estructurado de estas variables denominadas "variables regionalizada" que Matheron (1970) desarrolla con su Teoría de las Variables Regionalizadas.

Los orígenes de la Geoestadística se encuentran en los estudios mineros, para resolver problemas de estimación de recursos y reservas mineras a partir de datos de leyes mineras procedentes de sondeos. Sin embargo, rápidamente se puso de manifiesto la aplicabilidad general de las técnicas geoestadísticas a otras áreas de las ciencias de la Tierra, y en particular al estudio de la variabilidad espacial de variables hidroclimáticas e hidrogeológicas, como se muestra en la tesis doctoral de Jean-Pierre Delhomme (1976, [53]), trabajo pionero en este campo. Desde entonces y hasta la fecha, muchos han sido los avances y las aplicaciones de la Geoestadística en Hidrogeología, como se comentará en este trabajo.

En el contexto de este trabajo de investigación, el uso de esta metodología se ha dirigido a resolver un problema muy común con el que se enfrenta el hidrogeólogo: la obtención de un mapa de una variable hidrogeológica regionalizada a partir de datos experimentales de campo (figura 3.6). Estos datos experimentales pueden ser, por ejemplo, medidas del nivel piezométrico en puntos de agua, transmisividades medidas a partir de ensayos de bombeo, lluvia medida en pluviómetros, datos de calidad hidroquímica del agua etc., Se quiere resolver el problema de forma óptima de forma que el mapa de la variable regionalizada muestre, de la forma más fidedigna posible, su distribución espacial en el acuífero.

El kriging es la técnica que nos proporciona el método geoestadístico para resolver de forma eficaz el problema planteado, es decir, pasar de información puntual discreta a

información continua en el plano (mapa), a partir de valores experimentales. Su forma más general presenta las siguientes características:

1. Es un estimador lineal, esto es, una combinación lineal de los valores experimentales.

A partir de n valores experimentales $\{z(u_1), z(u_2), \dots, z(u_n)\}$, el valor de la variable en una localización arbitraria u_0 vale:

$$Z^*(u_0) = \sum_{i=1}^n \lambda_i^0 z(u_i) \quad (3.1)$$

donde:

$z(\cdot)$: valor experimental de la variable regionalizada de interés, p. e. nivel piezométrico.

$u_i = \{x_i, y_i\}$: localización en el plano del i -ésimo valor experimental.

λ_i^0 : peso o ponderador asignado al i -ésimo valor experimental en la estimación de la variable en la localización u_0 .

Conviene remarcar en este punto como otros métodos de estimación espacial bien conocidos como los polígonos de Thiessen o el inverso de la distancia, también se pueden expresar en forma de ecuación lineal. Lo que va a caracterizar el krigearje es que los pesos se van a obtener de modo "óptimo", al considerar la estructura de variabilidad espacial de la variable regionalizada, así como al imponer condiciones de no sesgo y de varianza del error de estimación mínimo.

2. Es un estimador insesgado.

El error de estimación $R(u_0)$ se puede expresar por la relación trivial:

$$R(u_0) = Z(u_0) - Z^*(u_0) \quad (3.2)$$

donde las letras mayúsculas denotan variables aleatorias en orden a que tengan sentido las operaciones estadísticas de esperanza matemática, esto es cálculo de la media y de la varianza.

El no sesgo del estimador implica:

$$E\{R(u_0)\} = 0 \quad (3.3)$$

de donde al sustituir en la ecuación 3.3 las ecuaciones 3.2 y 3.1 se llega a la condición de que la suma de los pesos ha de valer uno:

$$\sum_{i=1}^n \lambda_i^0 = 1 \quad (3.4)$$

donde $E \{.\}$ es el operador de la esperanza matemática.

3. Es un estimador que minimiza la varianza de estimación. La varianza del error de estimación

$$Var \{R(u_0)\} \quad (3.5)$$

se puede minimizar sujeto a la condición 3.4 para obtener los valores buscados de los pesos de ponderación.

4. Es un estimador que no depende de la distribución de los datos experimentales.
El estimador de krigeaje se obtiene a partir de las condiciones de no sesgo y minimización de la varianza de estimación, sin ninguna asunción respecto a la función de distribución de la variable regionalizada de interés.
5. Tiene en cuenta el soporte de la información, tanto el soporte de la información experimental como el soporte de estimación.

El soporte de información se define como el segmento, área o volumen (según se trate de un problema uni-, bi- o tridimensional) sobre el cual se mide la variable regionalizada. Por ejemplo, la medida de lluvia en un pluviómetro es un valor medio para un área circular de unos 15 cm de diámetro, mientras el valor de lluvia a partir de una medida de aforo, informa de lluvia media para la cuenca vertiente a dicha estación de aforo. Del mismo modo, si el interés es estimar el valor medio de la variable hidrogeológica para una cubierta ráster con una celda cuadrada de 30 m de lado, la variable a estimar tiene un soporte de 900 m². Si el soporte de la información experimental es muy pequeño en comparación con el área de estudio, se puede considerar, a efectos prácticos, como puntual (figura 3.7).

6. Tiene en cuenta el agrupamiento y correlación de la información experimental.
Datos agrupados y muy correlacionados proporcionan una información redundante y el krigeaje tiende a conceder pesos bajos a dichas disposiciones efectuando en la práctica un efecto de desagrupamiento.
7. Tiene en cuenta la correlación entre la localización a estimar y la información experimental.

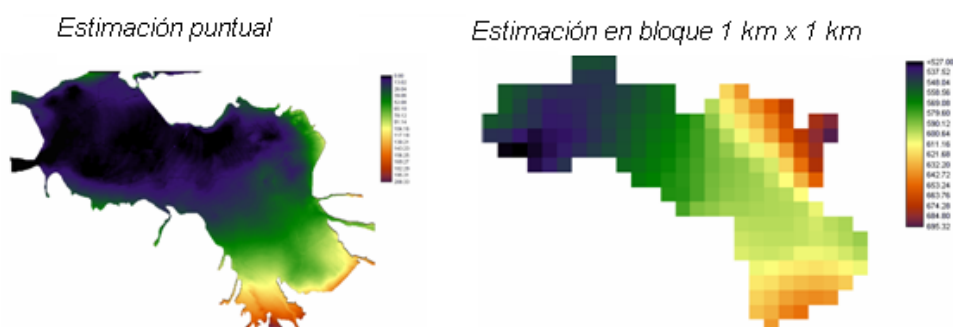


Figura 3.7: Mapas de coropletas correspondiente a la estimación por krigeaje puntual de la variable piezometría y de bloque.

En general, muestras más cercanas tienden a recibir mayores pesos de ponderación que valores experimentales más alejados debido a la correlación espacial y en respuesta a la intuición desarrollada en métodos ad hoc como los polígonos de Thiessen o el inverso de la distancia, utilizados fundamentalmente en el estudio hidrogeológico.

8. Tiene en cuenta el modelo de variabilidad espacial de la variable a estimar.

La variabilidad espacial de la variable regionalizada viene caracterizada por una función estadística, a introducir posteriormente, y conocida como variograma. Dicho variograma permite introducir en el proceso de estimación características de variabilidad de la variable regionalizada de interés, como pueden ser la anisotropía de la variabilidad espacial (figura 3.8) o su grado de continuidad espacial, entre otros.

9. Proporciona un mapa de incertidumbre (varianza del error de estimación) del mapa estimado.

El krigeaje proporciona como un subproducto del proceso de estimación, un mapa de la incertidumbre de estimación evaluada como la varianza del error de estimación (figura 3.9). Este mapa tiene diferentes utilidades, como el poder construir estimadores de intervalo y la identificación de zonas donde el muestreo puede ser deficiente que, en caso de ser posible, requiere incrementar el número de valores experimentales.

Algunos ejemplos de la utilidad del krigeaje, en su mayoría ya presentados en el capítulo de datos, se listan a continuación.

- Cálculo del mapa de isopiezas a partir de medidas de nivel piezométrico en puntos de agua. Lo mismo es aplicable para otras variables hidrológicas como lluvia, infiltración, etc.

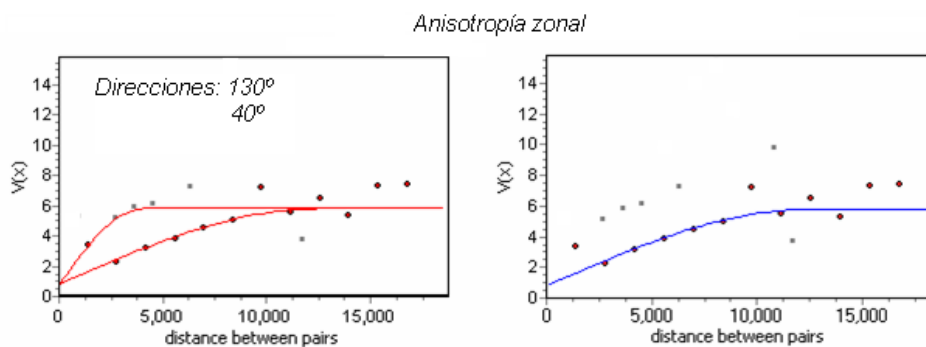


Figura 3.8: Variograma experimental (puntos) y variograma ajustado (líneas) para un ejemplo de anisotropía zonal calculada para la dirección de máxima variabilidad (130°) y mínima variabilidad (40°) de la permeabilidad del acuífero de la Vega de Granada.



Figura 3.9: Mapa de errores de la estimación por krigeaje de la variable transmisividad.

- Estimación de mapas de calidad de agua subterránea a partir de muestras de agua tomadas en piezómetros.
- Obtención del mapa de espesor de una formación geológica a partir de medidas de sondeos.
- Estimación de parámetros hidrogeológicos (permeabilidad, coeficiente de almacenamiento, porosidad, etc) a partir de medidas en puntos acuíferos.

Además, el krigeaje permite otra serie de cálculos muy interesantes para el hidrogeólogo, como por ejemplo:

- Obtener intervalos de confianza de la variable estimada a partir de la varianza de estimación.
- Estimación de valores medios como los valores de un mapa ráster (valores medios de celdas regulares) o el valor medio para áreas irregulares o el valor medio para todo el acuífero.
- Localización de los lugares más favorables para situar nuevas medidas experimentales.
- Calcular la probabilidad de superar un valor de corte, como la probabilidad de superar un límite de calidad de agua a partir de los valores experimentales medios.
- Mejorar el proceso de estimación con la inclusión de variables secundarias, especialmente si existe información abundante de dichas variables (cokrigeaje).
- Estimación de funciones de la variable estudiada, como por ejemplo el gradiente hidráulico a partir de la piezometría.

En cualquier caso, el krigeaje está basado en la estructura de variabilidad espacial de la variable regionalizada que, en general, es desconocida a priori, por lo que se ha de estimar a partir de los valores experimentales. El proceso de estimación consta de dos etapas:

1. Estimación de la estructura de variabilidad espacial "variograma".
2. Estimación del mapa de la variable por krigeaje haciendo uso del variograma.

3.3.1. Estimación de la estructura de variabilidad espacial: el variograma

La estimación de la estructura de variabilidad espacial es un proceso de inferencia estadística para el cual se ha de adoptar un modelo probabilístico para la variable

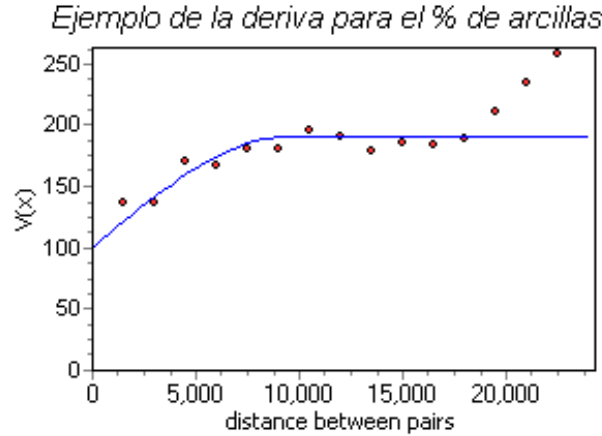


Figura 3.10: Ejemplo de variograma de la variable % de arcillas del suelo, donde se observa el efecto de la deriva para valores altos de distancia.

regionalizada. De este modo, la variable regionalizada se considera como una función aleatoria $Z(u)$.

$Z(\cdot)$: Representa la variable de interés: nivel piezométrico, contenido en calcio del agua subterránea, espesor saturado, transmisividad, precipitación, infiltración, etc.

$u \in \chi \subset R^2$: Es cualquier punto de una región finita del plano, por ejemplo cualquier punto dentro de los límites del acuífero. De este modo, como ya se ha señalado anteriormente $u = \{x, y\}$ representa la coordenada bidimensional arbitraria.

La función aleatoria $Z(u)$, $u \in \chi \subset R^2$, se considera que consta de dos componentes:

$$Z(u) = m(u) + R(u) \quad (3.6)$$

donde $m(u)$ es la derivada o tendencia que representa la variabilidad regional o de baja frecuencia y $R(u)$ es el residuo que representa la variabilidad local o de alta frecuencia. La deriva o residuo suele tener una interpretación física muy clara, como puede ser la precipitación orográfica que se incrementa con la altitud o el gradiente potencial que mantiene el flujo regional en un acuífero. El residuo representa la variabilidad local en torno a la deriva.

Para modelar la deriva es suficiente considerar un modelo polinomial de bajo orden (constante, lineal o cuadrático):

$$m(u) = \sum_{k=0}^K a_k f_k(u) = a_0 + a_1x + a_2y + a_3x^2 + a_4y^2 + a_5xy \quad (3.7)$$

donde se pueden identificar las funciones de base $f_k(u)$ para k de 0 a 5 con los monomios $(1, x, y, x^2, y^2, xy)$ respectivamente.

Para una deriva constante, se tiene que $a_1 = a_2 = a_3 = a_4 = a_5 = 0$, mientras que para una deriva lineal $a_3 = a_4 = a_5 = 0$.

Estos coeficientes se pueden estimar por mínimos cuadrados generalizados y la estimación de la deriva permite obtener una estimación del residuo.

El residuo se considera como un proceso aleatorio estacionario de media nula:

$$E \{R(u)\} = 0 \quad (3.8)$$

y covarianza (variograma) dependientes sólo del vector que separa dos localizaciones pero no de la posición absoluta de dichas localizaciones.

$$C(h) = E \{R(u) \cdot R(u+h)\} \quad (3.9)$$

$$C(h) = C(0) - \gamma(h) \quad (3.10)$$

$$\gamma(h) = \frac{1}{2} E \left\{ [R(u) - R(u+h)]^2 \right\} \quad (3.11)$$

El variograma dado en la ecuación 3.11, puede ser estimado por el método de los momentos utilizando la expresión:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [R(u_i) - R(u_i+h)]^2 \quad (3.12)$$

donde $N(h)$ es el número de parejas de datos que están separados por una distancia h .

El variograma estimado no sólo es útil para el krigeaje, sino que proporciona una valiosa información acerca de la variabilidad espacial de la variable regionalizada (figura 3.11):

1. En primer lugar hay que decir que el variograma experimental no se utiliza directamente, sino que es preciso ajustarle un modelo teórico.
2. El variograma puede presentar una discontinuidad en el origen debido a errores de medida o a una estructura de variabilidad que no se puede evaluar para distancias pequeñas, esto es, la distancia entre los datos más cercanos. Este valor de discontinuidad en el origen se llama varianza o efecto de pepita.
3. El variograma puede presentar un crecimiento sin límite o el crecimiento puede ser limitado hasta alcanzar un valor de "meseta".
4. La distancia para la cual el variograma alcanza la meseta se denomina "alcance".

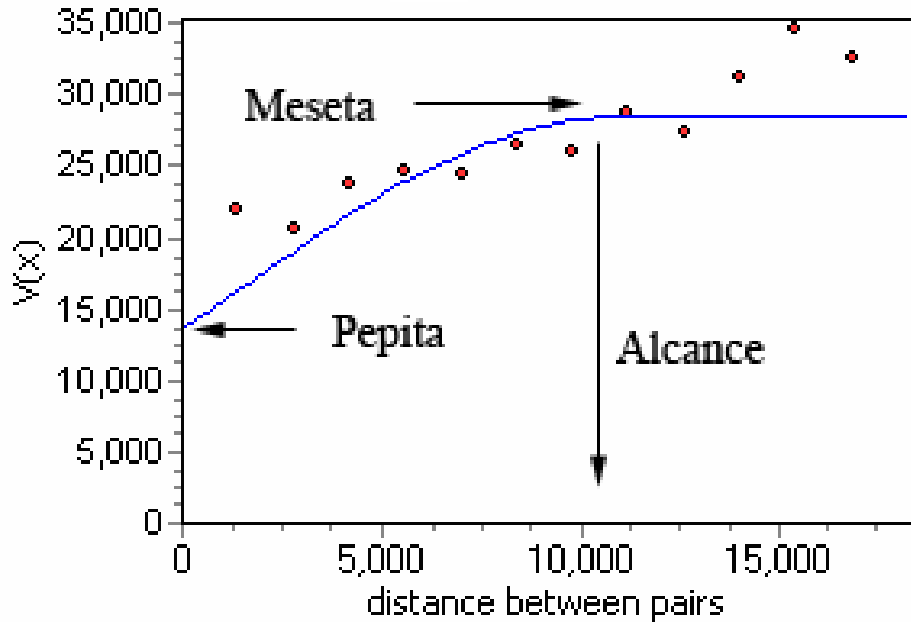


Figura 3.11: Variograma experimental y modelo ajustado correspondiente a la variable humedad de saturación del suelo.

5. El alcance puede variar según la dirección del espacio por lo que ayuda a interpretar los comportamientos anisótropos de la variable con la distancia.
6. La continuidad del campo aleatorio está relacionado con el comportamiento del variograma cerca del origen.

La relación 3.10 implica que se puede trabajar indistintamente con el variograma o con la covarianza (correlación) del residuo según conveniencia.

3.3.2. El krigeaje Ordinario (KO) y sus variantes

Una vez estimado el variograma y ajustado un modelo teórico (tabla 3.1 y figura 3.12), es posible minimizar la varianza de estimación (ecuación 3.5) sujeto a la condición de no sesgo (ecuación 3.3) para obtener el sistema de krigeaje (ordinario):

$$\sum_{i=1}^n \lambda_i^0 \gamma(h_{ij}) + \mu = \gamma(h_{i0}) \quad j = 1, \dots, n \quad (3.13)$$

siendo la suma de los pesos igual a uno (3.4) y donde:

h_{ij} : vector distancia entre los valores experimentales $z(u_i)$ y $z(u_j)$.

μ : parámetro de Lagrange.

Modelos teóricos de variograma	Nombre del modelo
$\gamma(h) = \begin{cases} 0 & \text{si } h = 0 \\ M & \text{si } h > 0 \end{cases}$	Efecto de pepita puro
$\gamma(h) = \begin{cases} M \left(\frac{3h}{2a} - \frac{1h^3}{2a^3} \right) & \text{si } h < a \\ M & \text{si } h \geq a \end{cases}$	Esférico
$\gamma(h) = M \left\{ 1 - e^{-\frac{h}{a}} \right\}$	Exponencial
$\gamma(h) = M \left\{ 1 - e^{-\left(\frac{3h}{a}\right)^2} \right\}$	Gaussiano
$\gamma(h) = Ch^b \text{ con } 0 < b \leq 2$	Potencial
$\gamma(h) = Ch$	Lineal

Tabla 3.1: Modelos teóricos de variograma utilizados en este trabajo.

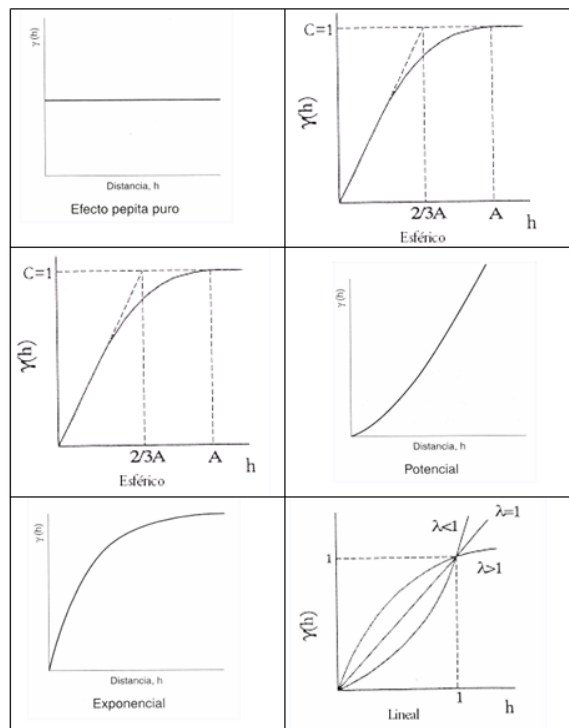


Figura 3.12: Modelos teóricos de variograma utilizados en este trabajo.

De la resolución del sistema lineal 3.13 con $n+1$ ecuaciones y el mismo número de incógnitas, se obtiene el valor de las ponderaciones o pesos λ_i^0 que se utilizarán en la estimación de acuerdo a la ecuación 3.1.

La varianza del error de estimación vale:

$$Var \{R(u_0)\} = \sum_{i=1}^n \lambda_i^0 \gamma(h_{i0}) + \mu \quad (3.14)$$

El sistema básico del krigeaje ordinario mostrado en la ecuación 3.13 se puede entender para considerar varios casos especiales. En primer lugar, puede ser de interés el estimar el valor medio de una variable regionalizada sobre una superficie de geometría arbitraria, como puede ser la lluvia media en una cuenca vertiente o el valor medio del nivel piezométrico en un acuífero. El sistema de krigeaje ordinario toma la misma forma que en 3.13 pero sustituyendo el variograma $\gamma(h_{i0})$ por $\gamma_{\cdot V}(h_{i0})$ que representa el variograma medio entre el valor experimental i -ésimo y la geometría a estimar V .

En el caso de que la deriva no sea constante, sino que sea lineal o cuadrática, el sistema de krigeaje se denomina sistema de krigeaje universal y toma la forma

$$\sum_{i=1}^n \lambda_i^0 \gamma(h_{ij}) + \sum_{k=0}^K \mu_k f_k(u_i) = \gamma(h_{i0}) \quad j = 1, \dots, n \quad (3.15)$$

$$\sum_{i=1}^n \lambda_i^0 f_k(u_i) = f_k(u_0) \quad (3.16)$$

Otro aspecto interesante es la estimación de probabilidades espaciales como por ejemplo ¿cuál es la probabilidad de que el contenido en nitratos en una localización supere un determinado valor umbral a partir de los valores medidos en los piezómetros vecinos?. Esto es cuánto vale la probabilidad:

$$Q(u; z) = 1 - F(u; z) = 1 - P(Z(u) \leq z) \quad (3.17)$$

con la estimación de $P(Z(u) \leq z)$ por krigeaje como:

$$P^*(u_0) = \sum_{i=1}^n \lambda_i^0 I(u_i) \quad (3.18)$$

donde

$$I(u_i) = \begin{cases} 1 & z(u_i) \leq z \\ 0 & z(u_i) > z \end{cases} \quad (3.19)$$

es la función indicatriz para la variable regionalizada $Z(u)$ y z el valor de corte. Los pesos necesarios de la ecuación 3.18 se obtienen de un sistema de krigeaje idéntico al dado en la ecuación 3.13 pero sustituyendo el variograma de la variable regionalizada por el variograma de la variable indicatriz para dicha variable regionalizada y un determinado umbral. El variograma de la función indicatriz se estima como:

$$\gamma_I(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [I(u_i) - I(u_i + h)]^2 \quad (3.20)$$

Varios ejemplos de este tipo de krigeaje se verán en capítulos posteriores, donde se calculará la probabilidad de que la calidad del agua supere ciertos valores establecidos como aptos para el consumo humano.

Otra mejora de la estimación se puede plantear en términos de introducción en el proceso de estimación de una o varias variables secundarias, como puede ser la estimación de la precipitación utilizando medidas de lluvia en pluviómetros y datos de altitud. Los datos de altitud son más fáciles de medir y pueden tener una cobertura completa del área de estudio, por ejemplo utilizando un modelo digital de elevaciones (MDE).

En este caso, además del variograma de la precipitación y el variograma de la altitud, es necesario considerar el variograma cruzado entre ambas variables que describe su variación conjunta:

$$\gamma_{ZY}(h) = \frac{1}{2N(h)} \sum_{i=1}^{n(h)} [Z(u_i) - Z(u_i + h)][Y(u_i) - Y(u_i + h)] \quad (3.21)$$

El estimador por krigeaje utilizando la variable principal $Z(u)$ y secundaria $Y(u)$, denominado cokrigeaje, se puede escribir como:

$$Z^*(u_0) = \sum_{i=1}^n \lambda_i^0 Z(u_i) + \sum_{j=1}^m \beta_j^0 Y(u_j) \quad (3.22)$$

donde β_j^0 es el peso asignado a la j -ésima variable experimental secundaria en la estación de la variable principal en la localización u_0 . De la imposición de la condición de estimador insesgado se obtienen las condiciones de los pesos:

$$\sum_{i=1}^n \lambda_i^0 = 1 \quad (3.23)$$

$$\sum_{j=1}^m \beta_j^0 = 0 \quad (3.24)$$

De la minimización de la varianza de estimación sujeta a las condiciones 3.23 y 3.24 se obtiene el sistema de cokrigeaje:

$$\sum_{i=1}^n \lambda_i^0 \gamma_Z(h_{ik}) + \sum_{j=1}^m \beta_j^0 \gamma_{ZY}(h_{jk}) + \mu_1 = \gamma_Z(h_{k0}) \quad k = 1, \dots, n \quad (3.25)$$

$$\sum_{i=1}^n \lambda_i^0 \gamma_{YZ}(h_{is}) + \sum_{j=1}^m \beta_j^0 \gamma_Y(h_{js}) + \mu_2 = \gamma_{YZ}(h_{s0}) \quad s = 1, \dots, m \quad (3.26)$$

donde:

$\gamma_Z(\cdot)$: variograma de la variable primaria.

$\gamma_Y(\cdot)$: variograma de la variable secundaria.

$\gamma_{YZ}(\cdot) = \gamma_{ZY}(\cdot)$: variograma cruzado entre variable primaria y secundaria.

μ_1, μ_2 : multiplicador de Lagrange.

Hasta este punto se ha pasado revisión del método de interpolación de krigeaje ordinario y de algunas de sus variantes:

- krigeaje de bloques para estimación de valores medios en polígonos de forma arbitraria.
- krigeaje de indicatrices para estimación de probabilidades de superar un valor umbral.
- krigeaje universal para considerar la presencia de una deriva polinomial.
- Cokrigeaje para la estimación multivariable.

En todos los casos, un mapa estimado, incluso la estimación por krigeaje, proporciona una versión suavizada de la realidad desconocida, esto es, una versión con menor varianza y con una estructura de correlación más continua que la realidad. Con objeto de generar simulaciones que reproduzcan la variabilidad espacial experimental, se elaboró la técnica de simulación condicional. Una simulación condicional es una realización de la función aleatoria $Z(u)$ que tiene las siguientes características:

- En los puntos experimentales realidad y simulación son coincidentes.
- La esperanza matemática de la simulación es idéntica a la media experimental.
- La varianza de la simulación es coincidente con el histograma experimental.
- El variograma de la simulación es coincidente con el variograma experimental.

Existen dos tipos de procedimientos: los algoritmos que permiten generar realizaciones no condicionales y el condicionamiento se efectúa por krigeaje y los algoritmos que permiten generar directamente simulaciones condicionales. En el primer grupo están el

método de las bandas rotantes y los métodos espectrales, entre otros, y en el segundo grupo está el método de la simulación secuencial.

Existen numerosas técnicas que permiten la generación de realizaciones. El método de simulación secuencial es el más sencillo y permite generar directamente simulaciones condicionales. Consiste fundamentalmente en cinco pasos:

1. Dada una malla de puntos a simular, se genera una secuencia aleatoria de visita de los puntos de la malla.
2. En el primer punto de la secuencia, se estima por krigeaje el valor más probable y la varianza de estimación en función de los valores experimentales.
3. Se genera un valor aleatorio con la media y varianza anteriores, en general a partir de una distribución gaussiana.
4. El valor simulado se añade a la información experimental para formar la información experimental extendida.
5. Se repiten los tres pasos anteriores para los demás puntos de la secuencia.

La simulación se utiliza en estudios de incertidumbre con la generación de un número alto de realizaciones (escenarios) que se pueden utilizar como entrada a simulaciones deterministas como la simulación de flujo y transporte de solutos. El resultado de la simulación de flujo, por ejemplo nivel piezométrico en un punto, ya no es un único valor sin un histograma que contiene la incertidumbre asociada al desconocimiento de la transmisividad.

Otros avances en la aplicación de la geoestadística en hidrogeología son:

- La inclusión del tiempo como una variable adicional en el estudio. Esto plantea diferentes problemas, entre los cuales está la definición de modelos de variograma espacio-temporales.

La variable regionalizada espacio-temporal se puede modelar con la función aleatoria $Z(u, t)$. El variograma espacio-temporal es:

$$\gamma(h, \tau) = \frac{1}{2} E \left\{ [Z(u, t) - Z(u + h, t + \tau)]^2 \right\} \quad (3.27)$$

La conjunción de la geoestadística espacial y los procedimientos utilizados en series temporales es todavía un campo de investigación activa.

- La utilización de métodos de inferencia del variograma, como la estimación por máxima verosimilitud, especialmente indicado cuando el número de datos experimentales es pequeño.

Este procedimiento es paramétrico, y asume que los datos siguen una distribución normal multivariable:

$$p(z; \beta, \theta) = (2\pi)^{-\frac{n}{2}} |V|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (z - X\beta)' V^{-1} (z - X\beta) \right] \quad (3.28)$$

$z' = \{z(u_1), z(u_2), \dots, z(u_n)\}$: n valores experimentales.

X : $n \times p$ vector de funciones de base de la deriva (monomios previamente introducidos)

β : $p \times 1$ vector de coeficiente de la deriva.

θ : Parámetros de la covarianza que se quieren estimar.

Los parámetros estimados por máxima verosimilitud son los que maximizan la expresión 3.28. Como el logaritmo es una función monótona, el máximo de la expresión 3.28 es también el máximo del logaritmo:

$$L(\beta, \theta; z) = -\frac{n}{2} \ln(2\pi) - \frac{1}{2} |V| - \frac{1}{2} (z - X\beta)' V^{-1} (z - X\beta) \quad (3.29)$$

Los detalles se pueden consultar en Pardo-Igúzquiza (1997 y 1998, [138, 137, 139]).

- La utilización de cokrigeaje para la resolución del problema inverso en hidrogeología.

Consiste en la estimación de la transmisividad (o la permeabilidad) mediante la utilización de datos de transmisividad (permeabilidad) y datos de nivel piezométrico que son, por lo general, más abundantes. El principal problema reside en obtener una covarianza cruzada que relacione ambas variables, transmisividad y nivel piezométrico, que están físicamente relacionadas. Una vez obtenida dicha covarianza, se puede obtener el mapa de transmisividad por cokrigeaje.

- La utilización del cokrigeaje para la estimación del gradiente hidráulico.

El gradiente hidráulico se puede definir como:

$$\nabla Z(u) = (D_i \{Z(u)\}, D_j \{Z(u)\}) \quad (3.30)$$

esto es, una magnitud vectorial, cuyas dos componentes son las derivadas direccionales a lo largo de los ejes principales X e Y respectivamente. Una derivada direccional a lo largo de una dirección arbitraria se estima como una combinación lineal de los datos experimentales:

$$D_e \{Z(u_0)\} = \sum_{i=1}^n \lambda_i^0 Z(u_i) \quad (3.31)$$

los detalles sobre el modo de obtener los pesos de modo óptimo por krigeaje se pueden consultar en Pardo Igúzquiza y Chica-Olmo (2004 y 2007 [137, 141]) (ver capítulo 2).

- Aumento óptimo de redes de muestreo y optimización de redes de muestreo.

Existen diferentes opciones para el problema de optimizar una red de muestreo. Lo habitual es tratar con una red de muestreo ya existente y los problemas más interesantes son el aumento de la red de muestreo de modo óptimo o la optimización de la disposición existente. En Pardo-Igúzquiza (1998 [139]) se muestra la solución de dichos problemas mediante geoestadística y simulated annealing.

3.4. Tratamiento de imágenes de Teledetección

Desde hace ya algunos años, la teledetección se ha convertido en una de las fuentes de información habituales en muchos de los estudios realizados sobre los distintos aspectos de la superficie terrestre, y en especial, en los trabajos de cartografía de recursos y procesos geológicos. Entre las excelencias de esta técnica hay que señalar su capacidad para proporcionar información a escalas relativamente grandes (de detalle) sobre amplias regiones de la superficie terrestre (Chuvieco, 1996), así como su notable relación calidad/coste. El procesamiento digital de imágenes puede además convertir imágenes brutas en mapas temáticos que destaquen ciertas características de los materiales superficiales, así como estructuras asociadas con elementos de interés (por ejemplo, acuíferos o depósitos minerales). Existen numerosos estudios en que imágenes Landsat, por ejemplo, se han utilizado para elaborar mapas geológicos básicos, producir mapas de estructuras (fracturas, etc.) regionales y locales que pueden haber controlado o controlan el movimiento de fluidos (por ejemplo, agua subterránea), detectar zonas alteradas asociadas con depósitos minerales, etc.

Desde el punto de vista físico, la teledetección parte del principio de la existencia de una perturbación, en este caso la energía electromagnética, que el sistema observado produce en el medio. Esta energía se transmite al sistema receptor que capta una señal que será registrada, almacenada y posteriormente interpretada. Desde el punto de vista práctico, la teledetección tiene por objeto el reconocimiento de las características de la superficie terrestre y los fenómenos que en ella tienen lugar a partir de los datos registrados por el sensor. El principal reto con que se enfrentan quienes trabajan en el

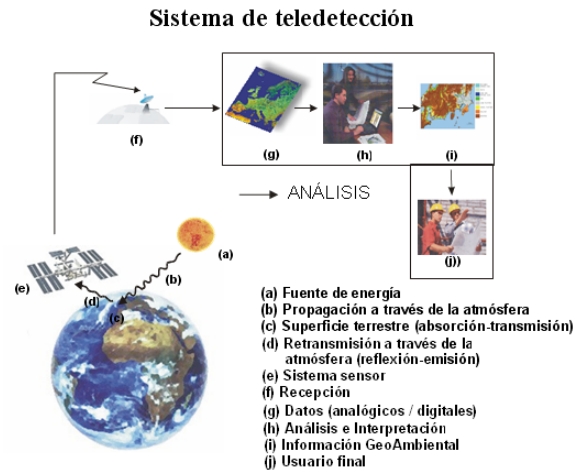


Figura 3.13: Elementos de un sistema de teledetección.

campo de la teledetección consiste en la transformación de los datos registrados por el sensor en magnitudes físicas y, en consecuencia, en información útil al usuario.

Los principales componentes de un sistema de teledetección son los siguientes (figura 3.13 y tabla 3.2):

- Fuente de energía: Supone el origen del flujo energético detectado por el sensor. Si el foco es externo al sensor, como el Sol (principal fuente de energía), se habla de teledetección pasiva, y si el sensor emite energía que se refleja en la tierra y es recogida de nuevo, se habla de teledetección activa (radar, lidar, etc).
- Cubierta terrestre: Formada por distintas masas de vegetación, rocas, suelos, agua, etc, que reciben la señal de la fuente de energía y la reflejan o emiten de acuerdo a sus características físicas.
- Sistema sensor: Incluye el sensor y la plataforma (satélite, avión, etc.) que lo sustenta. Los datos son luego enviados a los sistemas de recepción terrestres.
- Sistema de tratamiento y análisis de datos (visual y digital): para producir la información que llega al usuario.

Existe además una interacción significativa entre la energía electromagnética y el medio por el que viaja, es decir, la atmósfera. El efecto neto de la atmósfera depende de diversos factores (longitud de onda, condiciones atmosféricas, magnitud de la señal energética, trayecto de las ondas) y está controlado principalmente por los mecanismos de absorción y dispersión atmosférica.

En teledetección aplicada la radiación electromagnética más utilizada corresponde a las longitudes de onda de la radiación visible, infrarrojo próximo, medio y térmico,

Nº	Elemento del sistema
1	Fuente de energía
2	Propagación a través de la atmósfera
3	Superficie terrestre (absorción-emisión)
4	Retransmisión a través de la atmósfera (reflexión-emisión)
5	Sistema sensor
6	Recepción
7	Datos analógicos/digitales
8	Análisis e interpretación
9	Información
10	Usuario final

Tabla 3.2: Elementos de un sistema de teledetección (figura 3.13).

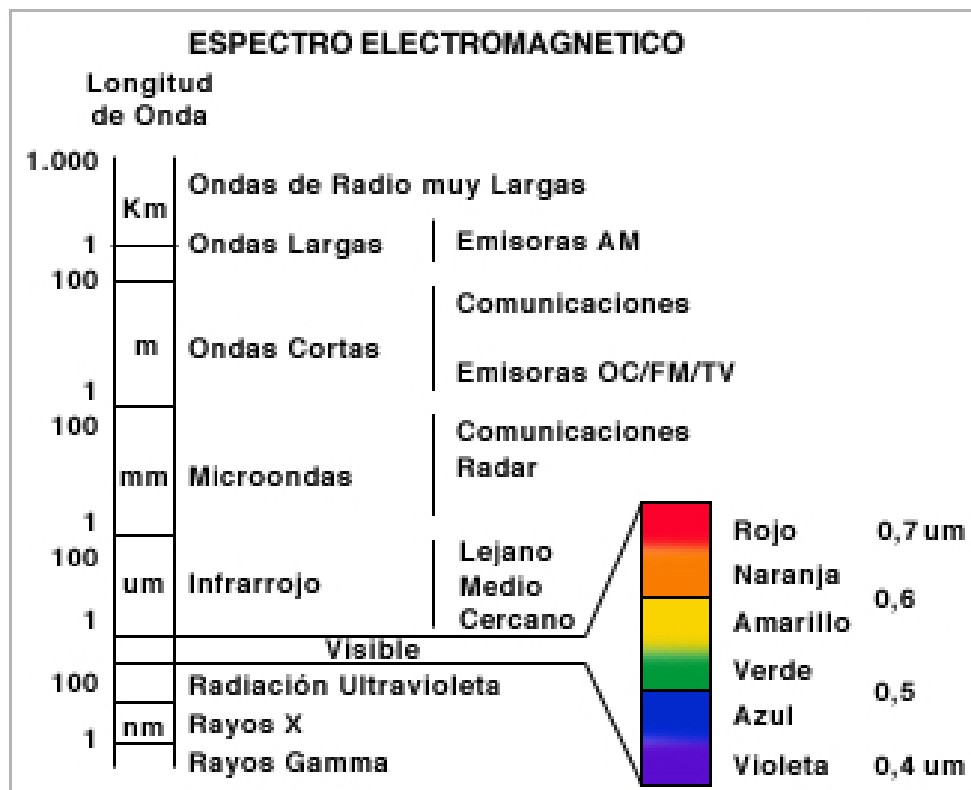


Figura 3.14: Espectro electromagnético.

así como al radar y a las microondas (figura 3.14). Las curvas que muestran la intensidad de energía emitida o reflejada por un objeto en diferentes longitudes de onda son denominadas curvas espectrales (y firmas o signatures a sus aspectos más distintivos) y son uno de los elementos clave en los estudios de teledetección.

Las aplicaciones de teledetección pasan por diferentes etapas, que incluyen básicamente (Drury, Chuvieco, Sabins, 1993, 1996 y 1996 [63, 41, 161]):

1. La definición del problema, donde se trata de identificar las diversas características físicas o biológicas de los procesos y fenómenos involucrados, para entender así las posibilidades de manifestación del fenómeno resultante en las imágenes de teledetección. Por ejemplo: el estudio de la geología estructural donde los principales fenómenos a estudiar serían las formas tectónicas del terreno o los lineamientos, o el estudio de la vegetación (tipo y fenológico, densidad de biomasa).
2. El análisis de los requerimientos de resolución (de lo que ya se habló en la sección 2.5)
3. La selección de los conjuntos de datos: tanto los datos de teledetección (imágenes) como la información del terreno necesaria (datos estructurales, topográficos, suelos, vegetación, etc.)
4. El procesamiento de datos, su interpretación y su aplicación.

3.4.1. Procesamiento digital de imágenes

En general, para llevar a cabo un estudio basado en teledetección es necesario realizar una serie de transformaciones en los datos originales. Estas transformaciones incluyen tratamientos más o menos estándares que sirven para corregir posibles distorsiones en los datos (correcciones), o bien, para poner de manifiesto ciertos objetos de interés, como pueda ser la vegetación, roca desnuda, determinadas litologías o estructuras, etc. Pero en muchos casos es necesario también realizar tratamientos ad hoc más avanzados de cara a destacar características concretas de un territorio.

El realce de imágenes consiste en una serie de técnicas para mejorar la interpretación visual de las imágenes, es decir, la modificación de una imagen para alterar su impacto sobre el intérprete. Se trata de disponer los datos mejor para su análisis visual, de tal forma que sean más evidentes las características de interés en nuestra aplicación. Normalmente, el realce se hace a costa de las características relativamente sin interés en el contexto de la aplicación. Los métodos de realce se pueden aplicar separadamente a cada banda de una imagen multispectral, por lo que sólo las bandas que se usan para generar una salida gráfica para asistir en las tareas de interpretación serían realizadas, son los realces geométricos y radiométricos, o bien, se pueden aplicar a varias bandas

a la vez, dando lugar a los llamados realces multibanda, espectrales o transformaciones de la imagen (cocientes, ACP, etc). Antes de aplicar los realces es siempre conveniente conocer la distribución estadística de los datos en las imágenes como paso preparatorio (histogramas, diagramas bivariantes, etc.). Esto puede ser realizado muy fácilmente con cualquier programa de tratamiento de imágenes.

En este trabajo se han sometido las imágenes disponibles a los procedimientos más o menos clásicos de un estudio aplicado, como son correcciones y realces radiométricos y geométricos, composiciones de color, cocientes de bandas, clasificación, generación de índices de vegetación, etc.

3.4.1.1. Preprocesamiento: corrección de imágenes

La calidad de una imagen de teledetección tiene dos aspectos: radiométricos y geométricos, y ambos influyen conjuntamente en la cantidad de información que puede ser extraída de la misma. Normalmente, el primer paso de un estudio de teledetección suele ser la corrección de las distorsiones radiométricas y geométricas que presentan las imágenes digitales. Una excepción a esta regla es cuando se pretende realizar una clasificación digital de una imagen; en este caso el proceso de corrección (geométrica) se suele realizar sobre la imagen ya clasificada.

En general, la calidad radiométrica de las imágenes puede verse afectada por diversos factores, pero uno de los más importantes, puesto que siempre está presente, es el efecto producido por la atmósfera. Para reducir este efecto en las distintas bandas de las imágenes, se procedió a realizar una corrección atmosférica basada en la aplicación de la técnica de la sustracción del objeto oscuro (Chavez, 1988 y 1989, [34, 35]). Se trata de un método muy sencillo y fácil de aplicar siempre que se pueda localizar en la imagen una masa de agua profunda, ya que se comporta de forma similar a un cuerpo negro si está relativamente limpia (sin carga de sedimentos). Este método suele dar un resultado aceptable en las aplicaciones más comunes. En este caso se utilizaron píxeles situados sobre el mar para efectuar la corrección.

Por otra parte, en el propio proceso de adquisición de las imágenes aparecen una serie de distorsiones geométricas que es necesario corregir, sobre todo si se pretende integrar las imágenes con otros datos en forma de mapas o bien derivar mapas de las imágenes. Así, una vez eliminado el efecto atmosférico se procedió a corregir geoméricamente las imágenes, a la vez que se referían todas a un sistema de georreferenciación común, concretamente UTM.

3.4.1.2. Realces geométricos

El proceso de realce de imágenes es un procedimiento monobanda, y consiste en aplicar sobre los números digitales originales una ventana o matriz móvil de coeficientes

de filtraje (kernel). El tamaño de la matriz es variable, de modo que a matrices de mayor tamaño el efecto deseado, generalmente realce o suavizado, es mayor. La teoría de filtros y su aplicación es un campo muy amplio y se puede consultar para mayor detalle en Schowengerdt (1997 [164]), por otra parte también es un campo extensamente utilizado y está disponible en la mayoría de los paquetes de tratamiento digital de imágenes.

En este estudio se examinaron diversos tipos de filtros orientados a la clasificación, para finalmente escoger un filtro de mediana que suavizara algunos contrastes y permitiera una mejor clasificación, como se verá más adelante en la clasificación.

3.4.1.3. Realces radiométricos

El realce radiométrico es una técnica de mejora de la calidad de la imagen visual, que trata de disponer mejor los datos para su análisis visual de modo que sean más evidentes los rasgos de interés. El principal realce radiométrico es el realce de contraste, referido al rango de valores de una imagen (o banda, si es que es multiespectral). Frecuentemente el rango de valores de reflectancia medidos por los sensores más comunes no ocupa todo el rango que los dispositivos de visualización permiten mostrar (0 a 255). El realce de contraste, por lo tanto permite ampliar el rango de valores que presenta una banda hasta ocupar el rango de 256 colores de los dispositivos, destacando de esta forma las características de la imagen, al hacer un uso óptimo de los colores disponibles. Estas técnicas están incorporadas en la mayor parte de los software de tratamiento de imágenes y son utilizadas rutinariamente por cualquier intérprete, siendo las más comunes el realce lineal y la ecualización del histograma.

3.4.1.4. Composiciones de color

La naturaleza multiespectral de las imágenes de satélite Landsat y AHS permite el uso simultáneo de varias bandas para su estudio mediante la superposición de varias bandas en una sola imagen y discernir así información difícilmente visible en una sola imagen. Esta superposición de bandas se realiza normalmente mediante el llamado método aditivo, en el que es posible generar cualquier color sumando los tres colores básicos: rojo, verde y azul (RGB), de manera que asignando los colores a tres bandas se genera una nueva imagen o composición en color. Las composiciones que hacen uso de bandas diferentes al visible son llamadas en falso color y son las más interesantes en geología. Este es un método rutinario, pero no por ello deja de ser un método efectivo y cómodo de presentación de imágenes multiespectrales, que produce un realce de las características de la imagen y facilita la interpretación por tres razones básicas (Lillesand, 1999 [107]): 1) La sensibilidad del ojo humano es mucho mayor al color que a los tonos de gris; 2) el número de variables disponibles pasa de una (brillo) a 3 (intensidad, tono y saturación); 3) la posibilidad de mostrar multibandas conjuntamente.

El sensor ETM+ con 7 bandas y la imagen AHS fundamentalmente, presentan grandes posibilidades para la combinación de información multibanda y el diseño de composiciones de color. Lo más común es intentar varias composiciones de bandas sobre la base de los objetivos del estudio y el conocimiento de las características espectrales de los objetos de la escena, para finalmente seleccionar la composición que incluya la información más relevante y presente el mejor aspecto visual (por ejemplo, se sabe que en geología son comunes las composiciones RGB 7/4/1 ó 5/4/2, y para la clasificación de usos del suelo la 4/3/2, mostrando las longitudes de onda mayores en el rojo y las menores en el azul según su disposición en el espectro).

3.4.1.5. Transformaciones de la imagen

El hecho de trabajar en una zona fundamentalmente de cultivo repercute en que las bandas de la imagen landsat correspondientes al rojo y al infrarrojo próximo (ETM3 y ETM4) sean las más utilizadas para el estudio de la superficie. Esto es debido a que los componentes a analizar muestran características específicas que permiten su discriminación en esta región espectral y, a su vez, se puede decir que aproximadamente el 90 % de la información relativa a la vegetación está contenida en estas dos bandas (Bannari et al. 1995 [16]).

La discriminación de las masas vegetales por tanto deriva del peculiar comportamiento radiométrico de la vegetación. Como se observa en la figura 3.15 la signatura espectral característica de la vegetación sana muestra un claro contraste entre las bandas del visible (especialmente en la roja) y el infrarrojo cercano (IRC). Mientras en el visible los pigmentos de la hoja absorben la mayor parte de la energía que reciben, estas sustancias apenas afectan al IRC, lo que permite separar con relativa claridad, la vegetación sana de otras cubiertas (Chuvieco Salinero, 2002 [42]). Por lo tanto, al observar la figura 3.15, podemos concluir que a mayor vigor vegetal, mayor será la diferencia entre las reflectividades de las bandas IRC y R, por lo que bajos contrastes indican una vegetación enferma o senescente o con poca densidad hasta llegar a los suelos descubiertos o el agua, que presentan una reflectividad muy similar en ambas longitudes de onda o incluso inferior en esta última.

Los índices de vegetación por lo tanto se construyen de acuerdo a estas características presentadas. Los más empleados son los denominados cociente simple entre bandas y el índice de vegetación normalizado (NDVI, Normalized Difference Vegetation Index) :

$$NDVI_i = \frac{\rho_{i,IRC} - \rho_{i,R}}{\rho_{i,IRC} + \rho_{i,R}}$$

donde $\rho_{i,IRC}$ y $\rho_{i,R}$ indican las reflectividades del píxel i en la banda del infrarrojo cercano y del rojo, respectivamente (TM4 y TM3). El mayor interés del índice nor-

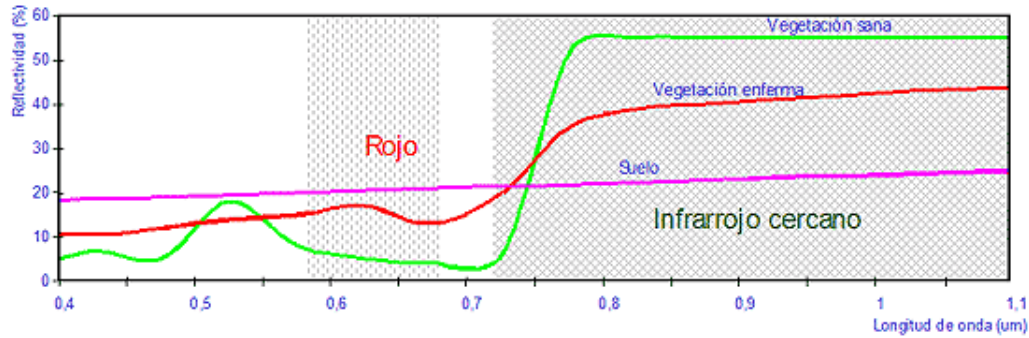


Figura 3.15: Contraste espectral de la vegetación sana en las bandas del R e IRC del espectro.

malizado de vegetación es que el rango de variación del mismo varía entre -1 y +1, lo que facilita su interpretación. De este modo se puede señalar como umbral crítico para cubiertas vegetales un valor de NDVI en torno a 0,1 y para vegetación densa entre 0,5 y 0,7 (Holben, 1986 [97]).

El NDVI se ha relacionado exitosamente con multitud de parámetros agronómicos, como por ejemplo el contenido de clorofila en las hojas (Curran, 1989; Westman, 19994; Zarco-Tejada et al., 2001 [49, 192, 198]), evapotranspiración potencial (Cihlar et al., 1991 ; Sánchez y Chuvieco, 2000 [43, 169]) entre algunos de los más interesantes, y en relación con este estudio e incluso en algunos casos con parámetros hidrogeológicos como puede ser la piezometría de un acuífero (Pei-Yu et al., 2006 [144]).

El NDVI se ha calculado para la zona de estudio (figura 3.16), para su inclusión en el área de estudio y para posibles estudios futuros en relación con la vegetación existente en la superficie del acuífero.

3.4.2. Clasificación digital de imágenes

La clasificación de la imagen es el proceso de asignación de píxeles individuales de una imagen multiespectral a clases discretas de tipo temático, a las que los valores radiométricos correspondientes a los distintos canales del sensor mejor se adecuan. En términos estadísticos esto supone reducir la escala de medida, de una variable continua (los ND originalmente detectados por el sensor) a una escala categórica (tipos de vegetación) u ordinal (densidades de vegetación) (Chuvieco Salinero, 2002 [42]). Para la mayor parte de los usuarios el proceso de clasificación supone la fase culminante del análisis de imágenes, hacia la que se orientan buena parte de los realces y correcciones previamente realizados y algunas transformaciones. De este modo, fruto de la clasificación es posible obtener una cartografía e inventario de las categorías objeto de estudio (Drury, 1993; Sabins 1996; Lillesand 1999 [63, 161, 107]).

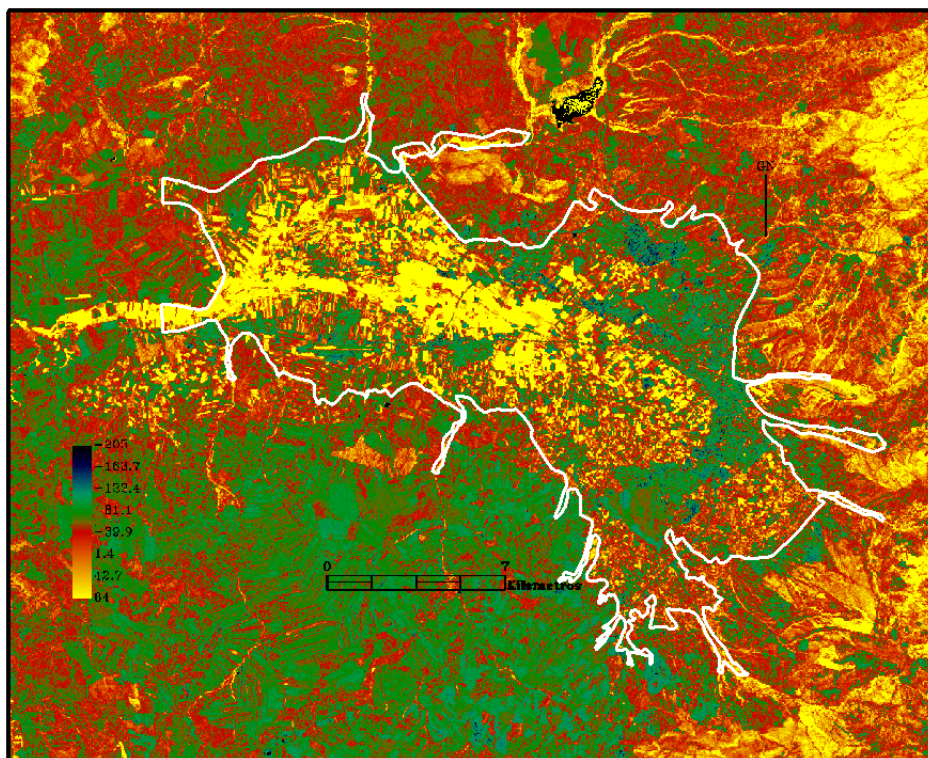


Figura 3.16: Índice de Vegetación Normalizado para la Vega de Granada y su contorno (NDVI) del año 2004 (Sensor ETM+).

La clasificación digital, de tipo supervisado, se inicia definiendo cada una de las áreas que se pretenden discriminar (Schowengerdt, 1983 [165]).

La clasificación supervisada, aplicada en este estudio, parte de un conocimiento previo del terreno, a partir del cual se seleccionan las muestras para cada una de las categorías. Los valores espectrales de los píxeles en las zonas de entrenamiento se utilizan para definir el espacio de decisión de cada clase. Para la clasificación supervisada, el éxito de la clasificación y por lo tanto, la validez de la información generada a partir de ella, depende directamente de la calidad del proceso de selección de áreas de entrenamiento (Lillesand, 1999 [107]). Éstas deben ser ejemplos representativos y homogéneos de cada categoría que se ha de utilizar en la clasificación. Estos ejemplos serían los patrones que definen en la imagen las distintas categorías objetivo, por lo que se trataría de obtener o identificar el rango de ND que identifica cada categoría, para todas las bandas que intervienen en la clasificación.

Estos rangos corresponden a las firmas espectrales de cada categoría (figura 3.18) las cuales se han elaborado en base a la media de los valores de los píxeles de la zona de entrenamiento. La elección de las zonas de entrenamiento se ha realizado según varios criterios, que son: una cubierta homogénea y una inclinación mínima y orientados de modo que se evite la presencia de sombras según la hora de paso del satélite. Además de estos criterios, se han escogido varias parcelas de tamaño pequeño, ya que si se escogen únicamente parcelas grandes la heterogeneidad de las distintas clases no se verá correctamente representada, lo que irá en detrimento de la clasificación (Campbell, 1987 [26]).

Varios autores han comprobado que los resultados de la clasificación están mucho más influidos por la definición previa de las categorías que por el criterio con el que éstas son posteriormente discriminadas (Hixson et al., 1980; Story y Campbell, 1986 [96, 173]). En otras palabras, la fase de entrenamiento constituye la columna vertebral de la clasificación numérica. En virtud de las anteriores consideraciones podemos observar que no existe superposición entre las clases escogidas al observar los espectros de cada una de las clases para las bandas de la imagen (figura 3.18). Las clases escogidas para este estudio han sido arbolado, bosque, embalses, matorral, olivar, regadíos, secano y urbano. Cada una de estas clases ha sido definida a partir de fotointerpretación y trabajo de campo. Los polígonos de entrenamiento usados para este estudio se muestran en la figura 3.17.

Para este trabajo se ha escogido una clasificación supervisada por el método de máxima verosimilitud (Richards, 1999 [151]). El método de clasificación de máxima verosimilitud asume que los valores de cada clase en cada banda están normalmente distribuidos, y calcula la probabilidad de que un píxel pertenezca a una determinada clase según la función:

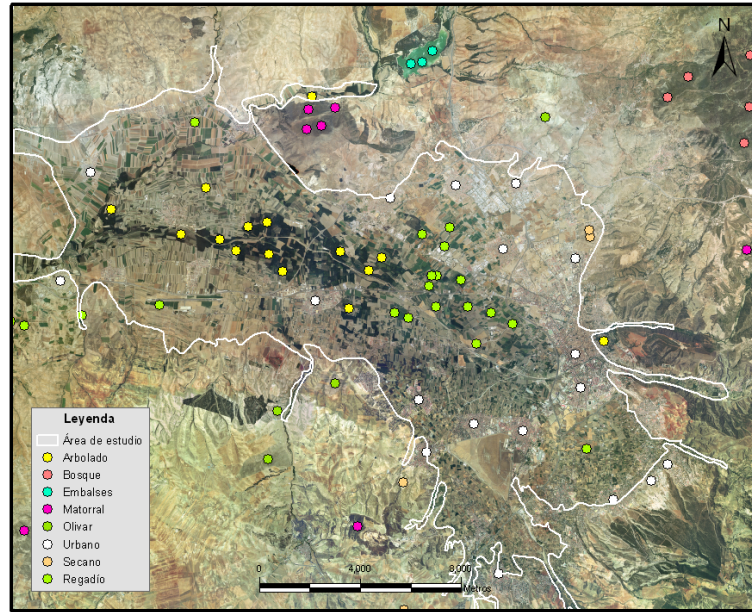


Figura 3.17: Localización de los polígonos de entrenamiento escogidos para la clasificación digital de la imagen Landsat ETM+.

$$g_i(x) = \ln p(w_i) - \frac{1}{2} \ln \left| \sum_i \right| - \frac{1}{2} (x - m_i)^t \sum_i^{-1} (x - m_i) \quad (3.32)$$

donde:

i = clase

x = dato n-dimensional (donde n es el número de bandas)

$p(w_i)$ = probabilidad de que la clase w_i ocurra en la imagen y se asume la misma para todas las clases

$\left| \sum_i \right|$ = determinante de la matriz de covarianza de los datos en la clase w_i .

\sum_i^{-1} = la matriz inversa

m_i = vector medio

Este método ha permitido clasificar la imagen del área de estudio en 8 clases, con un porcentaje de acierto según los polígonos de test de 95.5% al testear con las áreas de entrenamiento usadas para clasificar, y con un acierto del 90.7% al usar áreas de test independientes a las de la clasificación. Este resultado además se ve notablemente mejorado si añadimos un filtro de mediana (figura 3.20) a un 92.96% de exactitud.

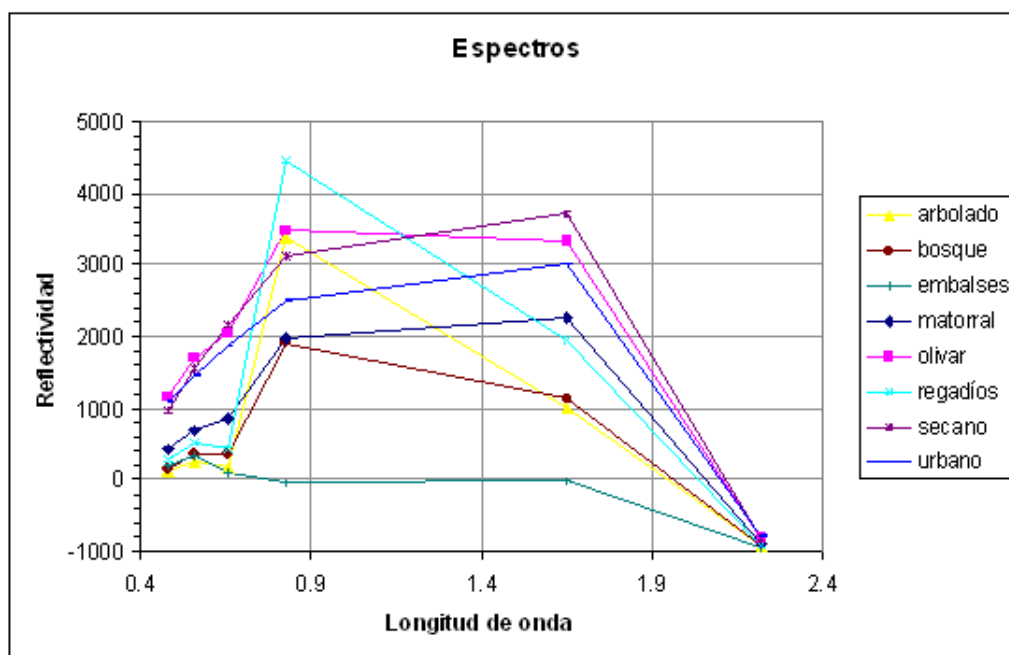


Figura 3.18: Espectros medios de reflectividad de los polígonos de entrenamiento.

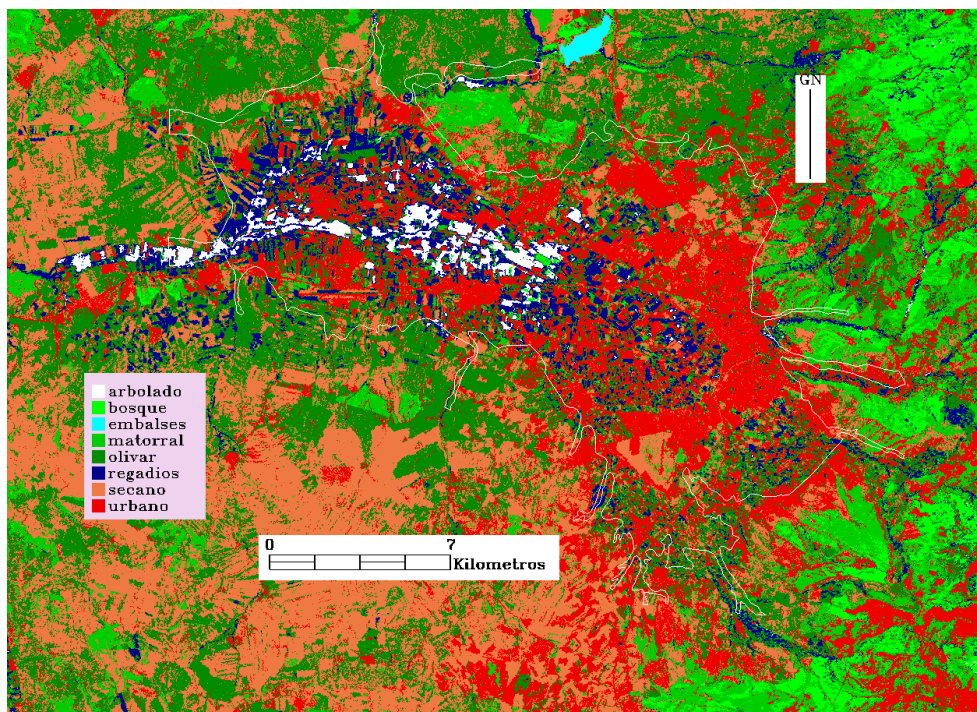


Figura 3.19: Mapa clasificado de Imagen Landsat ETM+ (año 2004) por el método de máxima verosimilitud.

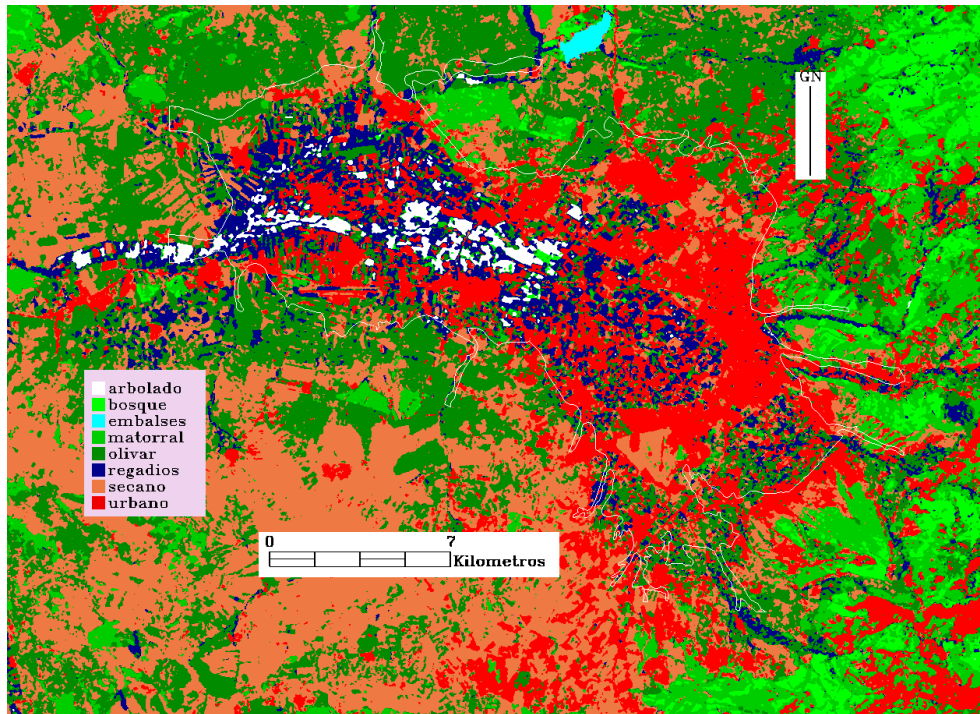


Figura 3.20: Mapa clasificado de Imagen Landsat ETM+ (año 2004) por el método de máxima verosimilitud y con filtro de mediana.

3.5. Métodos basados en Inteligencia Artificial

3.5.1. Aspectos generales

El Soft Computing (SC) es un área de la Inteligencia Artificial que se está convirtiendo en una alternativa a los métodos clásicos usados en el análisis espacial de la información (SIG-Ciencia). La SC es una mezcla de distintos métodos, que de una forma u otra, pueden cooperar desde sus fundamentos. Su principal objetivo es aprovechar la tolerancia que conllevan la imprecisión y la incertidumbre, para conseguir manejabilidad, robustez y soluciones de bajo costo. Los principales ingredientes de la SC son la Lógica Difusa, las Redes Neuronales Artificiales y los Algoritmos Genéticos. La SC puede considerarse como antítesis de lo que podríamos denominar Hard Computing (métodos clásicos de la Inteligencia Artificial, menos flexibles y con un costo computacional muy alto), de manera que podría verse como un conjunto de técnicas y métodos que permitan tratar las situaciones prácticas reales de la misma forma que suelen hacerlo los seres humanos, es decir, en base a inteligencia, sentido común, consideración de analogías, aproximaciones, etc.

La Soft Computing está llamada a jugar un papel cada día más importante en las Ciencias de la Tierra. Esto se debe, en parte, a la propia naturaleza de las reglas que

gobiernan muchos fenómenos físicos dentro de estas ciencias. La incertidumbre asociada a los datos, la inmensa cantidad de datos a tratar y la diversidad de tipos de datos son factores importantes que indican la idoneidad para utilizar la Soft Computing. Estas características se dan en los datos de este trabajo de investigación.

En el caso de la SIG-Ciencia, las características específicas de los datos espaciales que maneja (las observaciones no son independientes y la incertidumbre de los datos y los errores están normalmente estructurados espacialmente) propician que la Soft Computing sea una metodología potencialmente útil para aportar soluciones dentro de ésta, gracias a su tolerancia a la imprecisión e incertidumbre (Painho et al., 2005 [135]). Por otro lado, la Soft Computing dispone de instrumentos capaces de transformar un número muy elevado de datos (almacenados en grandes bases de datos) en información, y esta información en conocimiento (Knowledge Discovery). El espectacular incremento de los datos geoespaciales ocasionado por el desarrollo de la Teledetección y los Sistemas de Información Geográficos enfatiza la importancia de la Soft Computing como herramienta para el análisis y el modelado espacial.

Se describen a continuación de forma breve los principales métodos basado en Inteligencia Artificial.

a) Lógica Difusa

La base principal sobre la que se apoya la Lógica Difusa (LD) (Zadeh, 1965 [197]) es el concepto de conjunto difuso. A diferencia del conjunto clásico, un conjunto difuso es un conjunto sin frontera concisa. La transición de la pertenencia de los elementos al conjunto a la no pertenencia es gradual, y se especifica mediante funciones de pertenencia que dan una flexibilidad a los conjuntos difusos para el modelado de expresiones lingüísticas. Los conjuntos difusos permiten el modelado de conocimiento impreciso y cualitativo y la transmisión y manejo de la incertidumbre.

En la actualidad, una de las áreas de aplicación más importantes de la LD la componen los Sistemas Basados en Reglas Difusas (SBRDs). Este tipo de sistemas constituyen una extensión de los Sistemas Basados en Reglas de la lógica clásica, ya que emplean reglas de tipo SI-ENTONCES en las que los antecedentes y consecuentes están compuestos por proposiciones difusas (basadas en conjuntos difusos), en lugar de proposiciones lógicas clásicas. En un sentido muy general, un SBRD es un sistema basado en reglas en el que la LD se puede emplear tanto como herramienta para representar distintas formas de conocimiento sobre el problema a resolver, como para modelar las interacciones y relaciones existentes entre las variables del mismo. La estructura lógica de las reglas difusas SI-ENTONCES facilita la comprensión y el análisis de los modelos en forma muy parecida a la forma de razonar de los seres humanos sobre el mundo real. En resumen,

los SBRD permiten identificar sistemas y obtener modelos claramente interpretables por el ser humano.

b) Redes Neuronales Artificiales

Las Redes Neuronales Artificiales (RNA) (Bishop, 1995 [19]) modelan funciones no lineales muy complejas basándose en el funcionamiento del cerebro humano. Externamente, una RNA se comporta como una función a la que se aplican una serie de valores de entrada y proporciona un conjunto de valores de salida. Para actuar así, la RNA representa conocimiento sobre la función al utilizar una estructura formada por varias unidades de procesamiento simples (neuronas) interconectadas entre sí. La dinámica particular con que actúa cada neurona viene dada por una función (función de activación) que describe el cambio de la salida de la neurona. La dinámica global de la RNA aparecerá como resultado de la acumulación de las dinámicas particulares de cada neurona ponderadas mediante pesos asociados a las conexiones entre las neuronas.

La parte más importante de una RNA es el aprendizaje. En el contexto de las RNAs, el aprendizaje es el proceso de actualizar (aprender) la estructura y los pesos de las conexiones de la red, con objeto de que realice eficientemente una tarea específica. El objetivo fundamental del algoritmo de aprendizaje es producir redes que generalicen correctamente casos de un determinado dominio, y que adquieran esa facultad a partir de un aprendizaje previo con un conjunto de casos típicos de entrenamiento dentro de ese dominio (conjunto de ejemplos). En definitiva, los pesos de las conexiones se modifican de forma que los elementos de cómputo asocien patrones o representen características emergentes del dominio en cuestión.

c) Algoritmos Genéticos

Los Algoritmos Genéticos (AGs) (Goldberg, 1989 [83]) son algoritmos de búsqueda de propósito general que se basan en principios inspirados en la genética de las poblaciones naturales para llevar a cabo un proceso evolutivo sobre soluciones de problemas. Los AGs mantienen una población de cromosomas, los cuales representan soluciones candidatas a un problema, que evolucionan con el tiempo a través de tres procesos: el mecanismo de selección, los operadores de cruce y mutación. Estos algoritmos han demostrado ser, tanto desde un punto de vista teórico como práctico, una herramienta que puede proporcionar una búsqueda robusta en espacios complejos, y ofrecer un enfoque válido para solucionar problemas que requieran una búsqueda eficiente y eficaz. Para su aplicación, es necesario diseñar una función de adaptación para cada problema que se desee resolver. Dado un cromosoma, esta función devuelve un valor numérico proporcional al grado de bondad de la solución que dicho cromosoma codifica.

Los AGs se han aplicado con mucho éxito en problemas de búsqueda y optimización. La razón de gran parte de este éxito se debe a su habilidad para explotar la información que van acumulando sobre el espacio de búsqueda que manejan, lo que les permite redirigir posteriormente la búsqueda hacia zonas del espacio más prometedoras. Esta capacidad de adaptación es su característica principal, especialmente en espacios de búsqueda grandes, complejos y con poca información disponible, en los que las técnicas clásicas de búsqueda (enumerativas, heurísticas, etc.) no presentan buenos resultados.

d) Hibridación entre los Componentes de la Soft Computing

Cada una de las tres componentes de la SC descritas anteriormente ofrece sus propias ventajas y conlleva ciertas debilidades. Aunque comparten algunas características comunes, pueden considerarse como complementarias, ya que las características deseables que le faltan a una están presentes en otra. Esta propiedad justifica el creciente interés en el diseño de sistemas híbridos obtenidos mediante una combinación simbiótica (cooperación) entre componentes de la SC. Algunos ejemplos son: las RNAs difusas, la aplicación de los AGs para determinar la topología y pesos de una RNA, el uso de los SBRDs para adaptar parámetros de control de los AGs, la utilización de los AGs para diseñar SBRDs, etc.

3.5.2. Modelado de sistemas con Sistemas Basados en Reglas Difusas

En la actualidad, una de las áreas de aplicación más importantes de la Teoría de Conjuntos Difusos y de la Lógica Difusa, enunciadas por Zadeh en 1965 [197], la componen los SBRDs. Este tipo de sistemas constituye una extensión de los Sistemas Basados en Reglas de la lógica clásica, puesto que emplean reglas del tipo comentado SI-ENTONCES en las que los antecedentes y consecuentes están compuestos por proposiciones difusas en lugar de proposiciones booleanas clásicas.

En los conjuntos clásicos algo está incluido completamente en él o no lo está en absoluto. Por ejemplo, asignar a una calidad del agua buena o mala podría asimilarse a asignar 1 a todos los elementos incluidos en el conjunto o 0 a los no incluidos. La función que asigna estos valores se denomina función de inclusión o pertenencia. La función de inclusión de conjuntos clásicos, sería equivalente a la que se muestra en la figura 3.21a. En el caso de los conjuntos borrosos es posible describir el grado de pertenencia o inclusión de un objeto al concepto dado por la etiqueta (buena-mala calidad) mediante la asignación de un número real entre 0 y 1, así por ejemplo un agua podrá ser medio buena o medio mala. La función de pertenencia en el caso de los conjuntos difusos sería equivalente a la que se muestra en la figura 3.21b. Así, a un agua con un contenido en

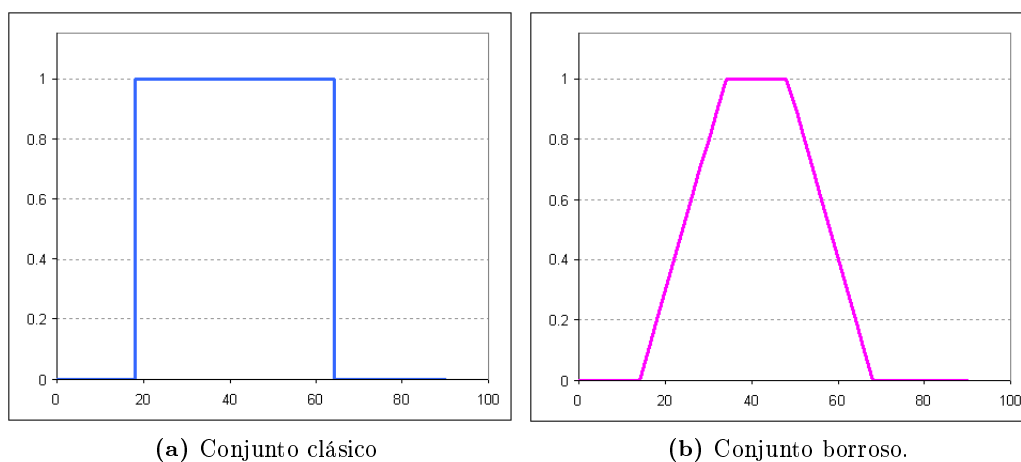


Figura 3.21: Funciones de pertenencia de conjuntos.

nitratos de 80 ppm habría que definirla como mala, pero en términos borrosos, podría decirse que se incluye en aproximadamente un 0.6 (60 %) al conjunto del agua mala.

De este modo, en términos matemáticos, sea U un conjunto de objetos, por ejemplo $U = R^n$, que se denomina Universo de Discurso, un conjunto borroso F en U queda caracterizado por la función de inclusión μ_F que toma valores en el rango $[0,1]$, es decir, $\mu_F : U \rightarrow [0, 1]$; donde $\mu_F(u)$ representa el grado en el que $u \in U$ pertenece al conjunto borroso F . Esto representa una generalización del concepto clásico de conjunto, donde la función de pertenencia toma sólo valores 0 ó 1; por el contrario, para un conjunto borroso, la función puede tomar valores intermedios.

En un sentido muy general, un SBRD es un sistema basado en reglas en el que la lógica difusa se puede emplear tanto como herramienta para representar distintas formas de conocimiento sobre el problema a resolver, como para modelar las interacciones y relaciones existentes entre las variables del mismo. La principal aplicación de estos sistemas inteligentes es el modelado de sistemas (Bárdossy y Duckstein 1995; Pedrycz 1996; Sugeno y Yasukawa 1993; [23, 143, 174]). Esto se debe a que existe una amplia gama de sistemas reales para los cuales, debido a su nivel de complejidad o imprecisión, no resulta posible obtener buenos resultados con las herramientas clásicas de modelado. Los SBRDs se han aplicado con éxito a una gran cantidad de problemas reales a lo largo de los últimos años (Bárdossy y Duckstein 1995; Doonatilake y Treleaven 1995; Hirota 1993; Pedrycz 1996 y Wang 1994; [23, 88, 95, 143, 188]).

En el área de las aguas subterráneas han sido muchos los trabajos en los que la lógica difusa ha sido aplicada exitosamente para la predicción de la vulnerabilidad junto con las técnicas SIG (Dixon 2005; Jiang y Eastman 2000; Gemitz y et al. 2005 y Andújar et al. 2006 [58, 57, 102, 77, 11]). Para la mayoría de estos trabajos, se establecen una

serie de variables a criterio experto, que generalmente están basadas en el índice de vulnerabilidad DRASTIC, mediante la asignación de una función de pertenencia a cada una y unos pesos también estimados a criterio experto.

Las técnicas neurodifusas pueden aprender el comportamiento de un sistema suficientemente grande de datos, y automáticamente generar reglas difusas y conjuntos difusos de un nivel de precisión especificado.

El interés de integrar el SIG con las técnicas neurodifusas radica en la creación de una aproximación inteligente, simple e interactiva, para la modelización de la vulnerabilidad de las aguas subterráneas, con un amplio campo de aplicación donde la fácil visualización de la misma podría hacer más sencilla la compleja tarea de la cartografía y mapeo de la vulnerabilidad para una gestión y planeamiento urbanístico mejor.

Con la introducción de los SIG en la hidrogeología, la adopción de métodos de índices y superposición como herramientas comúnmente usadas para evaluar la vulnerabilidad de las aguas subterráneas a escala de territorio ha crecido enormemente. Sin embargo, la aproximación por índices y superposición ha mostrado algunas limitaciones. Estas limitaciones están debidas probablemente a causas como la base física de estos métodos de superposición, así como a la inhabilidad de los SIG simples de tratar con la incertidumbre. Por ello, el uso conjunto de estas técnicas es particularmente útil cuando se modelan entradas difusas como son los parámetros hidrogeológicos, ya que estos toleran la imprecisión y la incertidumbre y muestran una disminución de la pérdida de información cuando son usadas con técnicas de SIG sencillas.

En la mayoría de los casos, el aprendizaje y el testeado se hace con datos de campo, por ejemplo con el empleo del contenido en nitratos. El uso de esta variable como medida de contaminación y vulnerabilidad tiene también ciertas limitaciones. El principal problema que supone es que su adición en el suelo del acuífero no es homogénea y tampoco se correlaciona con el tipo de cultivo. Es decir, es "al gusto del agricultor"; otro problema, es la consideración del sistema como estable (análisis píxel a píxel) cuando el sistema de flujos puede modificar y modifica la concentración de los nitratos en el píxel considerado. Sin embargo es de las pocas variables fácilmente medible en número suficiente para realizar este tipo de análisis, cuyos problemas se pueden obviar según ciertos supuestos. Por ejemplo, mediante la reclasificación de los nitratos se pueden establecer niveles de muy contaminado, a poco o nada contaminado. La consideración del sistema como estable, se resuelve suponiendo que el movimiento de la contaminación de un punto a otro supone igualmente una manera de autodepuración.

El modelado de sistemas mediante SBRDs es una vía para modelar un sistema haciendo uso de un lenguaje de descripción basado en la lógica difusa con predicados difusos (Sugeno y Yasukawa, 1993 [174]). El modelado con el empleo de conceptos propios de la lógica difusa es el típico ejemplo de una técnica que hace uso del conocimiento

y de los procesos de deducción propios de los seres humanos. En consecuencia, mediante el uso de la lógica difusa es posible capturar la vaguedad del modelo y la incertidumbre del problema, de forma tal que dicha representación aproximada se encuentre más ajustada a la realidad que otras formas que no puedan capturar estas cuestiones. Es posible llevar a cabo distintos tipos de modelado empleando distintos tipos de SBRDs, en función del grado de descripción y precisión que deseemos que tenga el futuro modelo. Estos dos requisitos, el grado de interpretabilidad del modelo para el ser humano y el grado en que ese modelo aproxima al sistema real, suelen estar enfrentados (Zadeh, 1973 [196]).

Quisiera dirigir la atención a la relación entre variables por medio de reglas difusas de tipo SI-ENTONCES por su interés en este trabajo. El ejemplo planteado corresponderá a la expresión:

SI *el espesor no saturado es bajo* **ENTONCES** *la contaminación será alta*

Estas reglas permiten establecer una relación lógica entre variables cualitativas de un sistema. Dichos valores cualitativos poseen una interpretación lingüística clara, que se denomina usualmente etiqueta o término lingüístico. El significado de los términos lingüísticos de las variables de entrada y salida está definido por conjuntos difusos, más precisamente por sus funciones de pertenencia. En este sentido, los conjuntos difusos o sus funciones de pertenencia proveen una interfaz entre las variables numéricas de entrada/salida y las correspondientes variables difusas de las reglas. Este concepto de modelado y análisis de sistemas por medio de variables lingüísticas fue introducido por Zadeh (1973 [196]) y ha sido desarrollado considerablemente en estos últimos años.

La estructura lógica de las reglas difusas SI-ENTONCES facilita la comprensión y el análisis de los modelos en forma muy parecida a como razonan los seres humanos sobre el mundo real. En este sentido, el solapamiento de las funciones de pertenencia asegura la generalización de situaciones no capturadas completamente por las reglas, esto es, la inferencia puede considerarse como un proceso de interpolación entre los resultados de las reglas individuales. Estos modelos con SBRDs tienen varias ventajas (Babuska, 1998[15]) sobre otros enfoques inteligentes como las Redes Neuronales:

- El modelado mediante SBRDs integra el procesamiento numérico y el simbólico en un esquema común y en consecuencia no se restringe solamente a áreas que requieren especialización o conocimiento humano. En este sentido, los modelos con SBRDs pueden hacer un uso efectivo de algoritmos de aprendizaje basados en ejemplos (datos de entrada-salida) y ser combinados con técnicas convencionales de regresión (Takagi y Sugeno, 1985; Wang, 1994 [177, 188]).
- La estructura basada en reglas de los Sistemas Difusos resulta de gran utilidad en el análisis de los modelos con SBRDs desarrollados a partir de datos numéricos,

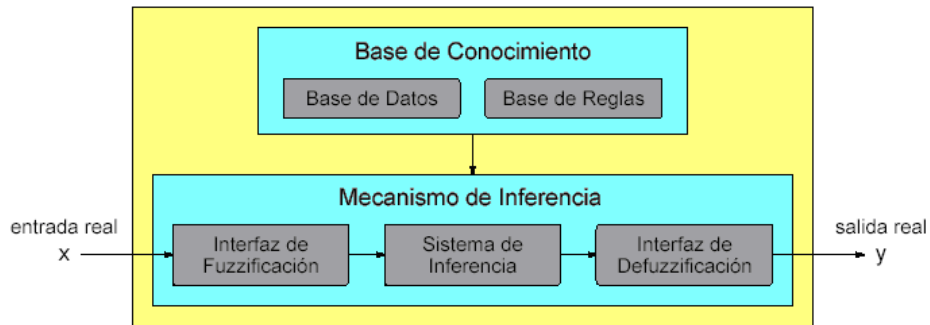


Figura 3.22: Estructura básica de un SBRD.

debido a que estas reglas pueden revelar una descripción adecuada del sistema que generan los datos. Este tipo de descripción puede ser confrontada y posiblemente combinada con conocimiento experto. Todo esto colabora en la comprensión del sistema y en la validación del modelo al mismo tiempo.

- El uso de términos difusos en las reglas puede verse como una clase de cuantificación, esto es, resulta posible desarrollar modelos de diferentes niveles de abstracción y precisión para un sistema dado, dependiendo del número de valores cualitativos considerado (granularidad).

3.5.2.1. Componentes de un SBRD

Un SBRD lingüístico está formado por los siguientes componentes (figura 3.22):

- Base de Conocimiento. Consta de: una Base de Reglas, que contiene las reglas en la estructura simbólica mencionada; y una Base de Datos, que almacena los conjuntos de términos lingüísticos y las funciones de pertenencia que definen sus significados.
- Interfaz de fuzzificación. Se encarga de transformar los datos de entrada precisos en valores utilizables en el proceso de razonamiento difuso, es decir, en algún tipo de conjunto difuso.
- Sistema de inferencia. Emplea estos valores y la información contenida en la base para llevar a cabo dicho proceso
- Interfaz de defuzzificación. Transforma la acción difusa resultante del proceso de inferencia en una acción precisa que constituye la salida global del SBRD.

Para obtener una descripción detallada sobre cada uno de estos componentes se puede consultar el apéndice.

3.5.2.2. Aprendizaje de SBRDs

Una vez seleccionadas las entradas y salidas y definida la base de datos, la relación entre las mismas debe establecerse mediante las reglas difusas. Existen dos maneras de construir la base de reglas difusas:

1. Mediante el uso del conocimiento experto. Con este método, la composición de la base de conocimiento se hace a partir de la información disponible de expertos. La persona experta especifica las etiquetas lingüísticas asociadas a cada variable lingüística, la estructura de las reglas y el significado de cada etiqueta. Este método es el más sencillo de aplicar cuando el experto es capaz de expresar su conocimiento en forma de reglas lingüísticas.
2. Mediante el uso de un método de aprendizaje automático basado en la existencia de información numérica. Debido a las dificultades que encuentran los expertos en la creación de la base de datos de reglas cuando el número de variables es elevado, se han desarrollado diversos métodos de aprendizaje inductivo en los últimos años para los distintos tipos de SBRDs. Con el uso del aprendizaje automático las relaciones relevantes serán automáticamente determinadas, y pueden suponer un nuevo punto de vista para la comprensión de la compleja interacción entre las variables del sistema.

En la actualidad, los SBRDs se obtienen mediante un proceso de aprendizaje automático a partir de un conjunto de pares de datos entrada-salida que reflejan el comportamiento del sistema real que se desea modelar. Existe una gran cantidad de métodos de este tipo que combinan distintos paradigmas de aprendizaje, tales como las redes neuronales (Fuller, 2000; Nauck et al. 1997 [76, 126]) o los AGs (Cordón et al., 2001; Herrera y Verdegay, 1996; Pedrycz, 1997)[46, 93, 143]).

En particular, en este trabajo usamos el método de Wang y Mendel (1992 [189]) para obtener la Base de Reglas de un SBRD. Es un método de aprendizaje eficiente y simple que se guía por criterios de envoltura de los datos del conjunto de ejemplos para realizar la extracción de reglas. Además, disfruta de dos ventajas adicionales: 1) es un algoritmo fácil de entender e implementar, y 2) lleva a cabo un proceso de aprendizaje muy rápido.

En la mayoría de los trabajos aplicados a la hidrogeología, y específicamente a la vulnerabilidad, el aprendizaje de las reglas se hace a criterio experto (Gemitzi et al., 2005; Jiang y Eastman, 2000; Van de Werf y Zimmer, 2000; Roussel et al., 2000 [77, 102, 183, 158]), exceptuando algunos trabajos de Dixon (2002 y 2005 [59, 58, 57]) y algunos otros con relación a ríos y contaminación agrícola (Andujar et al., 2006; Li et al., 2007 [11, 106]). Otro grupo de trabajos se centran la comparación de los métodos de

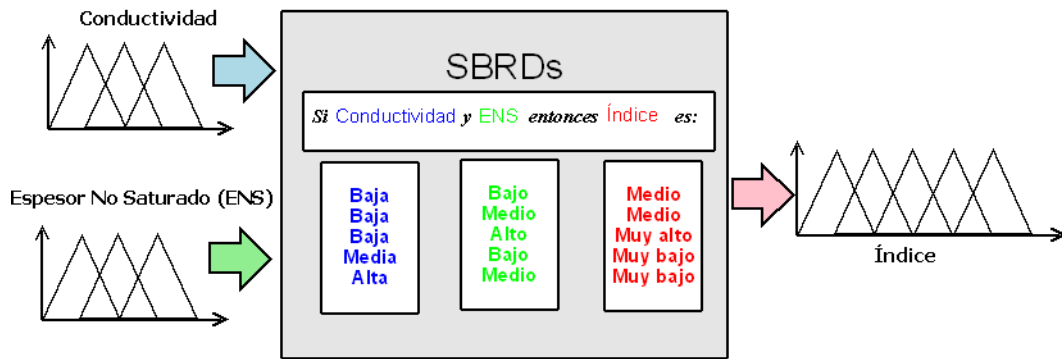


Figura 3.23: Ejemplo de aprendizaje por el método de Wang y Mendel.

defuzificación para el cálculo de índices, donde la generación de las reglas se hace también a criterio experto (Wang et al., 1994; Muhammetoglu y Yardimci, 2006 [188, 125]).

En la figura 3.23 se presenta un ejemplo de aprendizaje de una serie de reglas difusas por el método del algoritmo de Wang y Mendel. En este ejemplo, la conductividad y el espesor no saturado son las variables de entrada y se clasifican mediante tres etiquetas lingüísticas: bajo, medio y alto. La variable de salida es un índice de calidad del agua para el acuífero construido con todas las variables analizadas y donde los valores altos significan la mejor calidad. Este índice se clasifica mediante 5 etiquetas: muy bajo, bajo, medio, alto, muy alto.

El algoritmo de Wang y Mendel (figura 3.24) consiste en dividir los espacios de entrada, y salida en regiones difusas, generar la regla lingüística que mejor envuelve a cada ejemplo y, finalmente, para cada región de entrada difusa, seleccionar entre aquellas reglas con antecedente común (reglas conflictivas) la que mayor grado de importancia tenga (calculado según el grado de envoltura).

3.5.3. Redes Neuronales Artificiales

El aspecto central de una red neuronal es la capacidad de aprender de la experiencia acumulada de un conjunto de ejemplos (valores de entrada), de modo que se emula en cierta forma la capacidad del cerebro para aprender. De esta manera es posible resolver aquellos problemas que no pueden expresarse a través de fórmulas lineales o algoritmos, bien porque no se conocen, bien porque son demasiado complejos. Así, una red neuronal se puede considerar como un sistema para el tratamiento de la información, cuya unidad básica de procesamiento está inspirada en la célula fundamental del sistema nervioso humano, esto es la neurona.

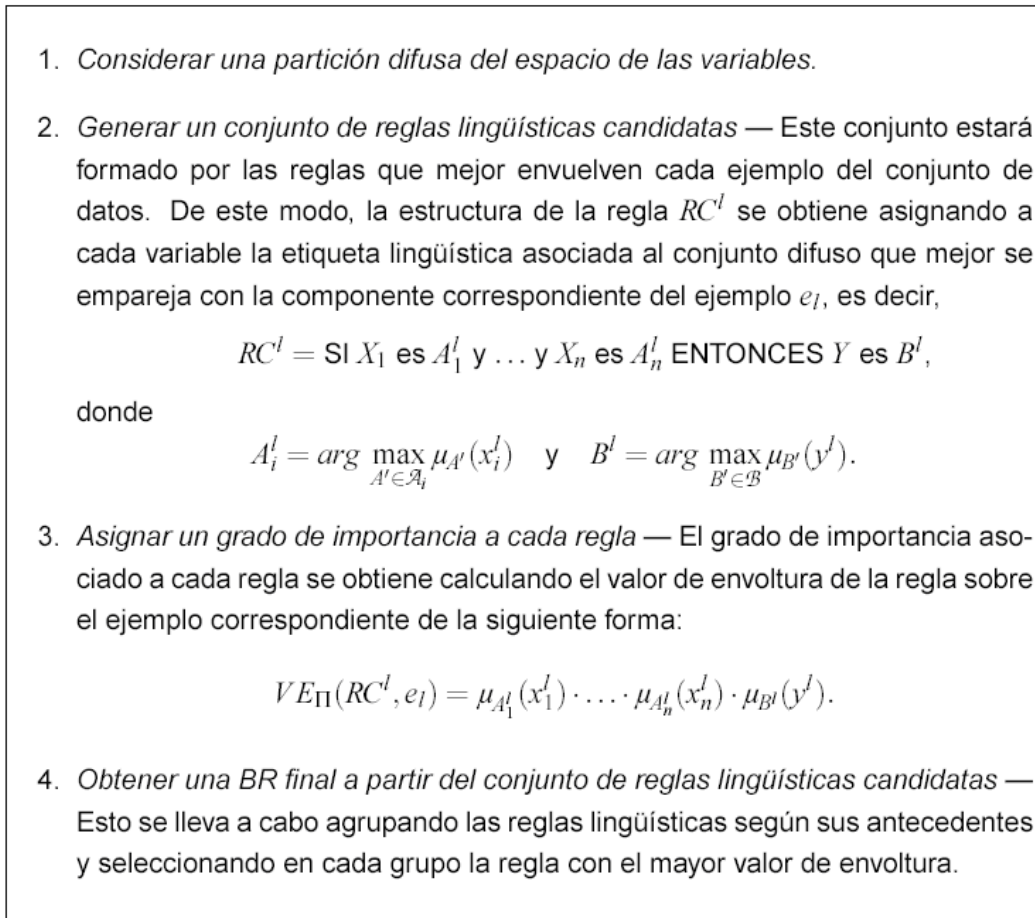


Figura 3.24: Método de Wang y Mendel.

3.5.3.1. La neurona artificial

El proceso de una neurona biológica se puede sintetizar en una serie de estímulos o entradas que se agregan de forma ponderada (fuerza sináptica), excitando a la neurona. Cuando la excitación alcanza un cierto umbral, permite la activación de la neurona que emite una señal, o salida, que a su vez puede estimular a otras neuronas. De forma análoga, una neurona artificial se puede considerar un dispositivo de estímulo-respuesta que actúa como un procesador elemental, ya que las operaciones que realiza son muy simples, y que al interrelacionarse con otras neuronas, vía una arquitectura definida previamente, da lugar a una red neuronal artificial (RNA) (Nilsson, 2001 [128]).

Se puede considerar entonces que una neurona artificial es un procesador elemental que recibe una serie de entradas con diferentes ponderaciones que, al agregarse, producen un nivel de excitación S . Esta activación se produce mediante la aplicación de una función de activación F , generalmente no lineal, a cada nivel de excitación, proporcionando una salida.

Las entradas suelen ser binarias o continuas y se suelen representar mediante vectores: $X = (x_1, x_2, \dots, x_n)$. La excitación S se calcula a partir de este vector de entrada, y las ponderaciones se definen con un vector de pesos $W = (w_1, w_2, \dots, w_n)$. La forma habitual es la suma de de entradas ponderadas:

$$S = XW = \sum_{i=1}^n x_i \cdot w_i$$

La función de activación determina el estado de actividad de una neurona al transformar la excitación en un valor de activación, que es la salida de la neurona, y cuyo rango se encuentra generalmente en el intervalo $[0,1]$ ó $[-1,1]$, ya que se considera que una neurona puede estar totalmente activa (1) o totalmente inactiva (0 ó -1).

De este modo podemos resumir que una neurona estándar consiste en:

- Un conjunto de entradas x_j y pesos sinápticos w_i .
- Una regla de propagación donde la suma ponderada es la más común $\sum_{j=1}^n x_j \cdot w_j$.
- Una función de activación $y_i(t) = f_i(h_i(t))$ que representa simultáneamente la salida de la neurona y su estado de activación.

Existen diferentes tipos de funciones de activación, las más usadas se representa en la tabla 3.3. Las funciones de activación proporcionan el estado de activación (y) actual a partir del potencial postsináptico (x) y del propio estado de activación anterior.



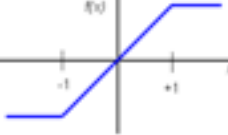
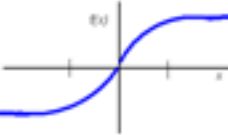


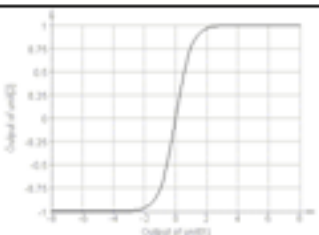
	Función	Rango	Gráfica
Identidad	$y = x$	$[-\infty, +\infty]$	
Escalón	$y = \text{sign}(x)$ $y = H(x)$	$\{-1, +1\}$ $\{0, +1\}$	
Lineal a tramos	$y = \begin{cases} -1, & \text{si } x < -1 \\ x, & \text{si } -1 \leq x \leq 1 \\ +1, & \text{si } x > 1 \end{cases}$	$[-1, +1]$	
Sigmoidea	$y = \frac{1}{1+e^{-x}}$ $y = \text{tg}^2(x)$		
Gaussiana	$y = Ae^{-\pi x^2}$	$[0, +1]$	
Sinusoidal	$y = A \text{sen}(\omega x + \varphi)$	$[-1, +1]$	
Tangente hiperbólica	$f(x) = \text{tg}^2(x)$	$[-1, +1]$	

Tabla 3.3: Funciones de activación utilizadas en esta investigación.

3.5.3.2. Redes neuronales artificiales

La interconexión del conjunto de neuronas mediante una estructura y arquitectura predefinida es lo que da origen a una red neuronal. La red neuronal se puede definir de distintos modos, pero la definición más precisa sería "un conjunto de procesadores elementales interconectados, ni lineal ni estacionariamente, que realiza al menos alguna de las siguientes funciones: Aprendizaje, Memorización, Generalización o Abstracción de características esenciales, a partir de un conjunto de ejemplos (Requena [149]).

En una red neuronal se pueden considerar tres aspectos esenciales:

- **Diseño estructural:** Contempla el tipo de neurona, sus entradas y funciones de activación, el tipo de conexiones entre las mismas y la selección de un conjunto adecuado de datos para el entrenamiento y para el testeo.
- **Aprendizaje:** Consiste en el proceso de adaptación de los pesos para que la respuesta de la RNA sea la adecuada tanto para los datos de los conjuntos de entrenamiento como de testeo, así como para otros datos de entrada no contemplados en estos conjuntos. Los métodos de aprendizaje se pueden dividir en dos grupos: supervisado, donde la red se entrena en base a un conjunto de datos de entrada y salida; y no supervisado, donde la red se entrena con ejemplos que incluyen sólo la entrada.
- **Recuperación de la información:** Consiste en recuperar la información almacenada en los pesos durante el proceso de entrenamiento para que sea utilizada en la aplicación, donde los datos de entrada pueden o no coincidir con los incluidos en el proceso de aprendizaje. Aún si los datos de entrada no coinciden con los de entrenamiento se espera que la respuesta de la RNA (previamente entrenada) sea lo más aproximada posible a la posible respuesta aprendida.

Uno de los modelos de red neuronal más utilizado es el de propagación hacia adelante. En este tipo de red, las neuronas se agrupan en capas paralelas, con distinto número de neuronas por capa, tal que los enlaces y conexiones se dan entre las neuronas de una capa y la siguiente solamente. Se pueden diferenciar tres tipos de capa: capa de entrada (donde no hay proceso de los datos), capas ocultas y capa de salida. El número de capas ocultas no conectadas directamente a las entradas o salidas puede ser variable.

La figura 3.25 representa la estructura de una red multicapa totalmente conectada y de propagación hacia adelante, con una capa de entrada $X = (x_1, x_2, \dots, x_n)$, una capa oculta $Z = (z_1, z_2, \dots, z_k)$ y una capa de salida $Y = (y_1, y_2, \dots, y_m)$.

Uno de los aspectos esenciales en una red neuronal lo constituye el proceso de aprendizaje y verificación, donde la red hace generalizaciones a partir de un conjunto de ejemplos y aprende a calcular la salida deseada. La esencia de este proceso es la adaptación

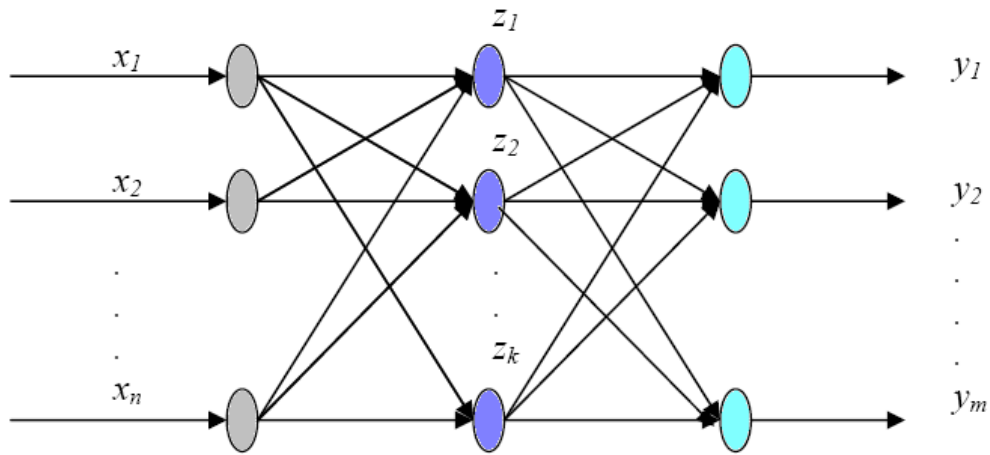


Figura 3.25: Ejemplo de red neuronal de propagación hacia delante.

de los pesos o ponderaciones en las conexiones de la RNA. Existen muchos procesos de aprendizaje, aunque para este trabajo nos vamos a centrar en el aprendizaje supervisado.

En el aprendizaje supervisado los datos de entrenamiento están constituidos por conjuntos de parejas (entrada, salida). El aprendizaje, llamado también proceso de entrenamiento de la red, consiste en la optimización de la función que mide las discrepancias entre las salidas calculadas por la red y las salidas esperadas (maestro). Para conseguir esta optimización se ajustan los valores de los pesos de las conexiones de forma iterativa, de acuerdo con algún criterio basado generalmente en técnicas matemáticas de optimización no lineal. Normalmente la función de discrepancia suele ser una medida del error (error cuadrático medio, generalmente). Así, el objetivo del entrenamiento es encontrar los pesos óptimos para minimizar el error entre los valores del maestro y las salidas de respuesta actual (Andina de la Fuente, 2005 [10]).

En el proceso de aprendizaje debe producirse el ajuste de los pesos de las conexiones entre las neuronas de acuerdo a una determinada regla. Las reglas más comunes de aprendizaje se mencionan a continuación, aunque una explicación mas detallada se puede encontrar en cualquier libro dedicado al estudio de las redes neuronales artificiales (Martín del Brío y Sanz Molina, 1997; Isasi Viñuela y Pedro Galván, 2004 [118, 99]). El aprendizaje hebbiano, según la regla de Hebb (1949 [91]) cuya idea básica sería que cuando dos neuronas conectadas están activadas simultáneamente, su interconexión se refuerza. Otro de los modelos de aprendizaje muy común es el perceptrón simple (Rosenblatt, 1962; Hertz, 1991 [157, 94]), que se inspira en las primeras etapas de procesamiento de los sistemas sensoriales de los animales (por ejemplo el de visión) en los cuales la información va atravesando sucesivas capas de neuronas que realizan

un procesamiento progresivamente de más alto nivel. Este algoritmo de aprendizaje se denomina de corrección por errores, y los algoritmos de este tipo ajustan los pesos en proporción a la diferencia existente entre la salida actual de la red y la salida deseada con el objetivo de minimizar el error de la red. Otro de los modelos de aprendizaje sería el de la regla delta o LMS y constituye el algoritmo de aprendizaje asociado al modelo de red neuronal denominado adalina (Widrow, 1960 [193]). La regla delta o de Widrow-Hoff es válida sólo para redes sin capas ocultas, y busca minimizar el error cuadrático medio siguiendo la idea del gradiente descendente.

3.5.3.3. Modelos de redes neuronales artificiales

En este apartado se presentan unos breves resúmenes de algunos de los tipos y arquitecturas de redes neuronales más frecuentemente utilizados en los campos de la hidrología y la hidrogeología.

- **Redes de propagación hacia adelante o feedforward:** Es el modelo de red neuronal artificial más simple, introducido por Rosenblatt (1962 [157]). Tiene propagación hacia adelante y como función de activación utiliza la función umbral (tabla 3.3). Tiene dos capas, la de entrada y la de salida, y no posee capas ocultas. Su entrenamiento es supervisado en base a un conjunto de ejemplos binarios con valores de entrada y salida conocidas, lo que permite ajustar los pesos de conexión entre la capa de entrada y salida para obtener el valor y la respuesta esperada. Un perceptrón permite realizar tareas de clasificación, y sólo permite discriminar entre dos clases linealmente separables, es decir, cuyas regiones de decisión pueden ser separadas mediante una única condición lineal. La importancia histórica del perceptrón radica en su carácter de dispositivo entrenable.
- **La adalina:** Este es otro de los modelos clásicos de redes neuronales (Widrow, 1960 [193]). Este modelo utiliza una neurona similar a la del perceptrón, pero de respuesta lineal, cuyas entradas pueden ser continuas. No obstante, la diferencia más importante con el perceptrón reside en la regla de aprendizaje que implementa, conocida como LMS (Least Mean Square) o regla delta.
- **El perceptrón multicapa:** Es una red de propagación hacia adelante, con capa de entrada, capa de salida y una o más capas ocultas y se entrena con el algoritmo de retropropagación. Aunque la primera mención de este tipo de red fue de Werbos (1974 [191]) fue el grupo del PDP (*Parallel Distributed Processing Group*) (Rumelhart et al., 1986 [159]) el primero en desarrollarla. Esta red precisa de funciones de activación continuas y diferenciables y utiliza la regla delta generalizada para adaptar los pesos. Existen distintas versiones de este algoritmo, y son muy

diversos los tipos de arquitecturas y conexiones que pueden ser implementadas. La más común es una arquitectura con una capa oculta y una función de activación de tipo sigmoidea.

Algunas aplicaciones de estas metodologías de Inteligencia Artificial se aplican en el capítulo 5.

3.6. El Sistemas de Información Geográfica como herramienta de integración de datos espaciales

3.6.1. Generalidades

Los Sistemas de Información Geográfica (SIG) son, básicamente, herramientas informáticas que procesan y analizan datos con alguna componente espacial. En nuestro Sistema de Soporte de Decisión de las aguas subterráneas de la Vega de Granada los SIGs han valido como tecnología de integración del conjunto de datos, métodos y técnicas comentados en los apartados anteriores de este capítulo (Geoestadística, Teledetección, Lógica Difusa, Redes Neuronales, etc.). En resumen, nuestro Sistema de Soporte de Decisión comprende el conjunto de técnicas y metodologías específicas que mejoran el tratamiento de la información para la obtención de resultados y su integración en un sistema (SIG) que permite el fácil acceso y recuperación de la información para la toma de decisiones (figura 3.26).

La definición quizás más aceptada y más extendida viene dada por el National Center for Geographic Information and Analysis de USA, que los define como: un sistema de hardware, software y procedimientos elaborados para facilitar la obtención, gestión, manipulación, análisis modelado, representación y salida de datos espacialmente referenciados, para resolver problemas complejos de planificación y gestión. Una variante de esta definición sería la dada por Gooldich (2000 [86]), que los define como una tecnología integradora que une varias disciplinas con el objetivo común del análisis, creación adquisición, almacenamiento, edición, transformación, visualización, distribución, etc. de la información geográfica.

Cabe decir que su característica principal es que trabajan con mapas y, a diferencia de otros programas que también lo hacen, como son los de topografía, los de cartografía y los sistemas CAD/CAM, los SIG pueden realizar operaciones de análisis espacial bastante sofisticadas en algunos casos, mediante el uso de los datos espaciales y sus atributos almacenados en el propio sistema, que permiten obtener nuevos mapas a partir de una única fuente de datos (ejemplo en figura 3.27).

El principal problema que pueden llegar a plantear la utilización de los SIG es precisamente consecuencia de la facilidad en su manejo, lo cual posibilita que usuarios

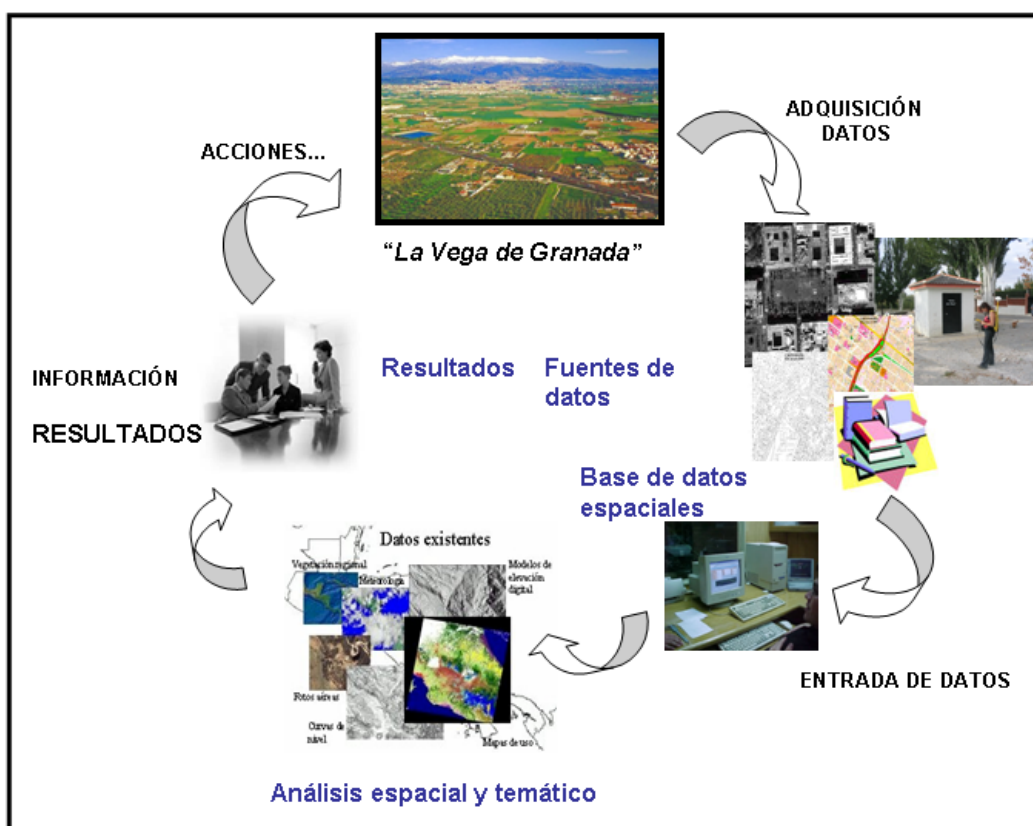


Figura 3.26: El ciclo de tratamiento de la información georreferenciada.

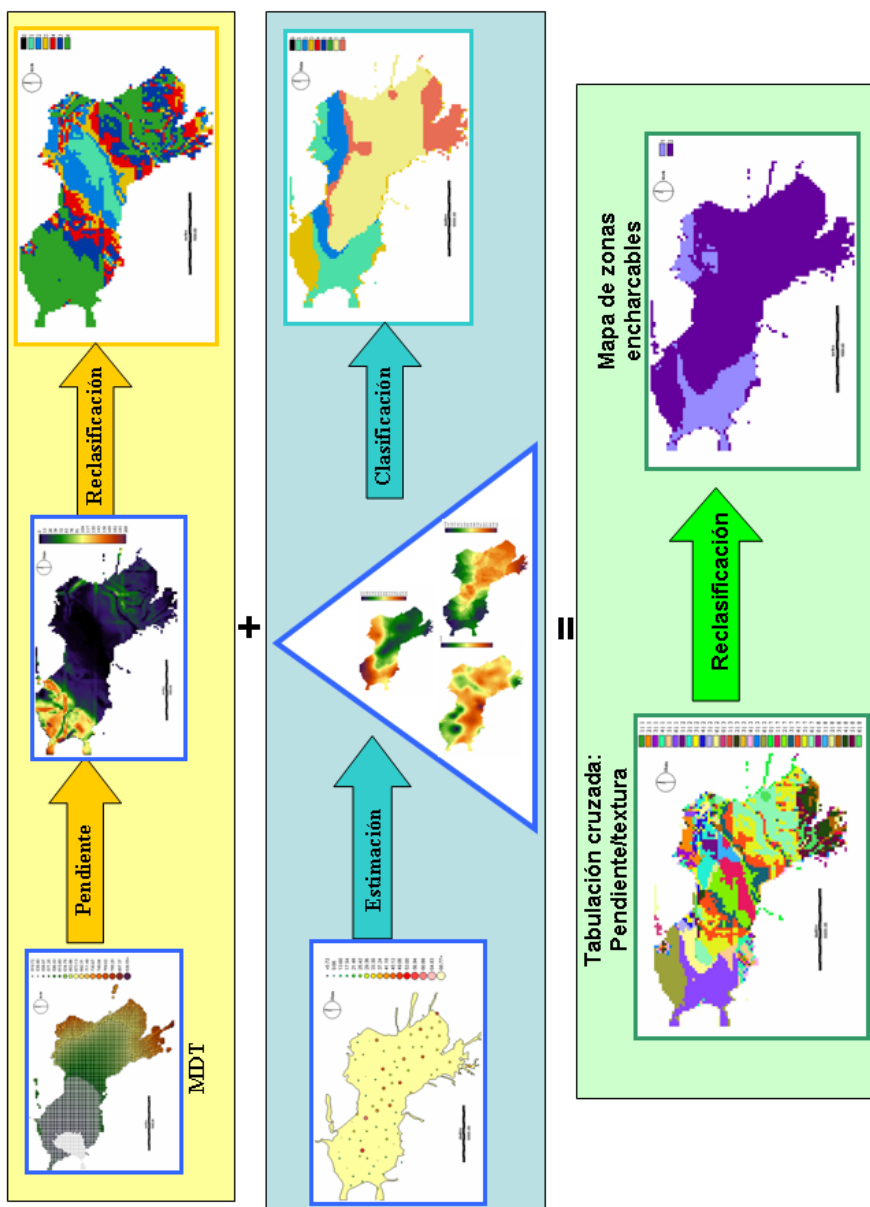


Figura 3.27: Ejemplo de creación de un mapa de zonas encharcables.

sin conocimientos suficientes sobre la naturaleza de los datos geográficos y sobre las funciones de análisis que incorporan puedan utilizar estas herramientas y obtener resultados aparentemente correctos, presentados como prolijos listados numéricos o mapas incuestionables a tenor de su apariencia, pero con escaso o nulo valor científico-técnico.

Actualmente existe mucha documentación específica sobre el tema en revistas científicas dedicadas parcial o totalmente a los SIG (International Journal of Geographic Information Science, Photogrammetric Engineering & Remote Sensing, Computer and Geosciences, International Journal of Remote Sensing), así como libros y monografías. Pero en muchos casos, la abundante información pierde al usuario medio entre complicadas fórmulas y teorías. Por este motivo es necesaria información complementaria que ayude a discernir y a comprobar los resultados de un mapa antes de tomarlo como verdadero para su uso en nuestros objetivos, ya que puede dar lugar a resultados y ejecuciones incorrectas en la ejecución de nuestro trabajo. Una de las herramientas que han aparecido para solucionar este tipo de problemas por usar cartografía no creada por uno mismo o no creada por un experto, son los metadatos. Los metadatos son archivos en los cuales se almacena toda la información referente a un mapa en cuestión, tal como la escala, la fuente, el error, los pasos para su elaboración, etc., información necesaria e imprescindible y que debe de estar estandarizada y actualizada si se pretende un uso de la información por parte del público.

De esta manera actúan muchos de los organismos responsables de la infraestructura cartográfica de los países más desarrollados que disponen de información digital y bases de datos accesibles al gran público, que son, al fin y al cabo, el alimento de los SIG. Sin embargo, actualmente las bases de datos son tan voluminosas que, en ocasiones, hace inviable su utilización en organizaciones pequeñas y proyectos con bajos presupuestos, puesto que cada vez son necesarias máquinas más potentes y con mayor espacio solamente para manejar la base de datos.

En los últimos años se está produciendo una explosión en la utilización profesional de estas herramientas, englobadas dentro de las aplicaciones multimedia, con acceso remoto desde redes globales (Internet) a bases de datos extensivas, lo que supone un importante cambio cualitativo en los métodos de trabajo clásicos, al tiempo que evita el "tedioso" trabajo de guardar, actualizar y mover la información.

Aunque a veces se piensa que un SIG es un software integral, realmente suele estar compuesto por una serie de ficheros ejecutables de una forma sencilla, de tal modo que cada módulo realiza un tipo de operación. Según el tipo de software que se elija las funcionalidades serán unas u otras, de modo que en muchos casos es necesario migrar los datos de un tipo de extensión a otra distinta para poder realizar según qué tipo de operaciones. Estas diferencias entre los distintos softwares se deben en muchos casos al tipo de trabajo al que está orientado, incluyendo en algunos casos herramientas más

potentes de visualización o de cálculo e investigación. Sin embargo todos ellos deben tener unos componentes y unas capacidades mínimas para ser considerados un SIG.

Los componentes esenciales de un SIG se podrían reducir en seis. El principal y más importante es la base de datos para almacenar y gestionar los datos y sus atributos, seguido en importancia de un sistema para la representación cartográfica y un sistema de análisis espacial. Ésta última parte es en la que se encuentran mayores diferencias entre los distintos software comerciales, ya que se incluyen en este apartado en la mayoría mientras que aparecen como módulo aparte en otros. El sistema de tratamiento de imágenes, el análisis estadístico y el análisis geoestadístico completan los componentes esenciales.

A pesar de las diferencias ya comentadas, en la parte esencial (base de datos, representación y análisis espacial) no todos los software son iguales, y varían tanto en su grado de desarrollo y capacidad como en la facilidad de manejo.

El núcleo central del sistema lo constituyen las bases de datos espacial y temática, en las cuales se almacenan de forma estructurada los objetos cartográficos (posición, tamaño y forma) y sus características no geométricas (atributos). En algunos sistemas, ambos datos están claramente separados y son de distinto tipo. En otros, se integran en una sola entidad y permiten además la utilización de bases de datos externas como Dbase, Access, MySQL y Oracle entre otros, para almacenar los atributos temáticos. El hecho de permitir este flujo de información entre distintas fuentes, es una de las mayores potencialidades de un SIG, ya que, en muchos casos, los cambios de formato y extensiones son un problema añadido para la integración de datos con distinto origen, época, escala y proyección.

En cualquier caso, una vez creada la base de datos, el siguiente punto importante es la gestión de los mismos, ya que un SIG debe permitir que con los datos almacenados en tablas se establezcan relaciones entre ellos y se creen nuevas tablas con los resultados obtenidos. Al relacionar siempre estas tablas con el atributo espacial se obtienen mapas temáticos. El tipo de operaciones que se pueda realizar con estas tablas también limita o extiende la capacidad de trabajo con estas herramientas.

Para este trabajo el software escogido ha sido ArcGis, que es uno de los software comerciales más extendidos y populares, y es en el que se ha instalado la base de datos. Sin embargo, debido en muchos casos a sus limitaciones como software comercial y menos enfocado a la investigación, el tratamiento de diferentes capas ha sido hecho en Idrisi, para algunos casos de lógica difusa y teoría de decisión; en ENVI o Erdas, para el tratamiento de las imágenes de satélite e incluso Matlab para la generación de la cartografía de textura del suelo. Sin embargo, en la mayoría de los casos, una vez realizado el tratamiento, los nuevos datos (mapas) obtenidos, han vuelto a incluirse en la base de datos espacial de ArcGIS (geodatabase), y se ha añadido como información

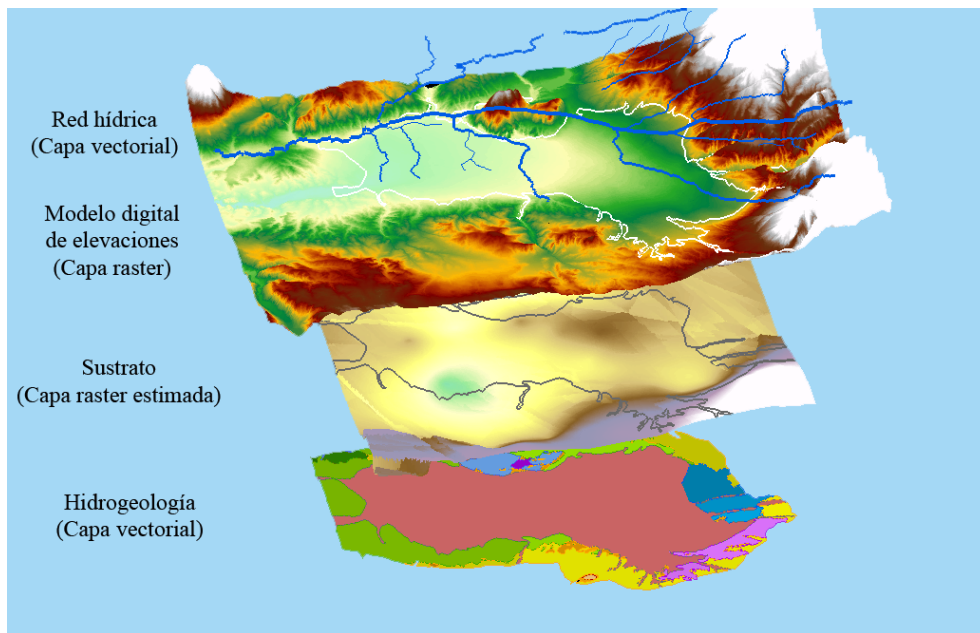


Figura 3.28: Ejemplo de distintas clases de capas en un SIG.

adicional los pasos seguidos para la consecución de los nuevos datos (metadatos) en caso de necesidad de repetir algún proceso.

Además de la variedad de tratamientos espaciales realizados, para este trabajo se ha requerido la aplicación de herramientas más complejas y nuevas, que en la mayor parte de los casos no se incluyen en los SIG, ya que estos sistemas no ofrecen en ocasiones soluciones adecuadas para determinados problemas de análisis y modelado espacial (O'Sullivan y Unwin, 2002; Fischer y Nijkamp, 1992; Gooldich et al., 1992 [132, 72, 87]). Por lo tanto, para el tratamiento de la información espacial con métodos fundamentalmente de inteligencia artificial ha sido necesario exportar los datos a distintos formatos, para poder incluirlos de este modo en programas creados por nosotros mismos o de código abierto (Weka y Keel) y que posteriormente, gracias a las capacidades de integración de los SIG, los resultados se han podido incorporar en la base de datos espacial como mapas.

Los mapas son modelos de la realidad que registran de forma simplificada aquellos aspectos que más nos interesan en función del objetivo del mapa y de la escala. Los SIGs organizan el conjunto de la información georreferenciada en capas (mapas) que se superponen unas a otras (figura 3.28) y permiten la interpretación y recopilación de la información espacial en un área mediante las operaciones de análisis y modelado. Normalmente se utilizan dos métodos para representar esta realidad, que dan lugar a otros tantos de modelos de datos geográficos: el modelo vectorial y el raster.

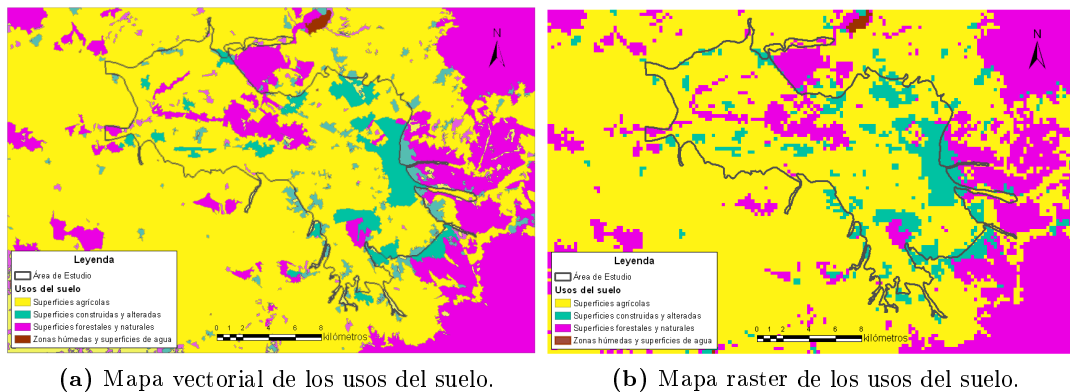


Figura 3.29: Rasterización del mapa vectorial de los usos del suelo de la Vega de Granada.

En el modelo vectorial se registran únicamente las fronteras de los objetos espaciales, aproximándolas por medio de líneas delimitadas por puntos que se localizan por sus coordenadas en un sistema de referencia. El resultado son mapas en los que aparecen tres objetos básicos: puntos, líneas y polígonos. Algunos ejemplos de este tipo de mapas son los puntos de muestreo, los ríos o la cartografía de los polígonos industriales del área de estudio.

En el modelo raster lo que se registra no son las fronteras de los objetos espaciales, sino su contenido, quedando sus límites implícitamente representados. Para ello se divide el dominio geográfico en una malla regular de celdas (píxeles), normalmente cuadradas, y se asigna a cada celda un valor numérico que representa el atributo que se está registrando, mientras que la situación geográfica de cualquier punto viene definida por la posición de la celda correspondiente en un sistema de coordenadas cartesiano. En este tipo de mapas, los cálculos aritméticos u operaciones matemáticas se suelen realizar píxel a píxel, de modo, que una correcta estandarización de los datos para un mismo trabajo requiere tamaños de celda similares (preferentemente iguales) que permitan correlacionar y operar con datos de la misma escala y precisión.

En la elaboración de la base de datos de este trabajo, la mayoría de los datos vectoriales, además de estar incluidos en la base de datos como elementos de puntos, líneas o polígonos, se ha transformado a formato raster por medio de geostatística (para los puntos) o por simple asignación de valores enteros (rasterización) para mapas de polígonos como es por ejemplo el mapa geológico o de usos del suelo (figura 3.29). Estos mapas raster tienen todos una resolución similar con un píxel de 250 m x 250 m, en coordenadas UTM y proyección European Datum 1950, que era el estándar cuando se comenzó a realizar este trabajo.

3.6.2. Operaciones y análisis de datos en los SIG

Como complemento a los aspectos generales del SIG comentados, se indican a continuación, de forma resumida, las funciones básicas del SIG como sistema de integración de la información espacial recopilada en este trabajo.

Las operaciones de análisis se han esquematizado seis grupos genéricos de operaciones (figura 3.30): entrada de datos, operaciones sobre el aspecto temático, operaciones sobre el aspecto espacial, operaciones sobre ambos componentes, modelado cartográfico, modelado espacial y salida de datos.

El grupo de operaciones concerniente al análisis de datos espaciales y temáticos es posiblemente el conjunto de operaciones que diferencia a los SIG de los sistemas CAM/CAD y es a partir de este conjunto de operaciones de los que se puede generar nuevos datos, introduciéndonos así en el campo del modelado cartográfico y espacial. La potencialidad de este sistema ha permitido identificar áreas de interés especial con mayor facilidad y menor coste. Por ejemplo, pongamos que interesa evaluar los municipios con más riesgo a la contaminación por nitratos en base a su densidad de población. Los sistemas SIG permiten realizar consultas temáticas a su base de datos con refejo en su componente espacial. Por lo que tras diversas consultas podemos obtener un mapa de pozos utilizados como abastecimiento contaminados por nitratos. La superposición del mapa de densidad de población municipal permite identificar las zonas con mayor riesgo de contaminación a la población (figura 3.31).

El modelado espacial pueden definirse de distintas maneras. Una definición bastante amplia es la propuesta por Unwin (1981 [181]) que lo define como un amplio conjunto de procedimientos de estudio de los datos geográficos, en los que se considera de alguna manera sus características espaciales. Sin embargo debemos tener otras definiciones en cuenta como la de Goodchild (1987 [85]) que lo define como un conjunto de técnicas cuyos resultados son dependientes de la localización y los atributos de los objetos, esto es, a ambos componentes del dato espacial. Esta cuestión se verá más extensamente en el apartado del análisis espacial de los datos.

Por otra parte, el modelado espacial comprende la construcción de modelos para producir resultados espaciales (Fotheringham y Wegener, 2000; O'Sullivan y Unwin, 2002 [75, 132]).

Finalmente, la integración y distintas operaciones espaciales permiten que los modelos espaciales antes mencionados se puedan concebir también como la unión entre los métodos basados en modelos y las operaciones de los SIG para proveer herramientas adecuadas para asistir a la toma de decisión en contextos espaciales (figura 3.32). Es decir, los modelos espaciales serían el punto de unión entre las operaciones aritméticas, espaciales o algorítmicas normalmente no implementadas en los SIG (aunque relativas

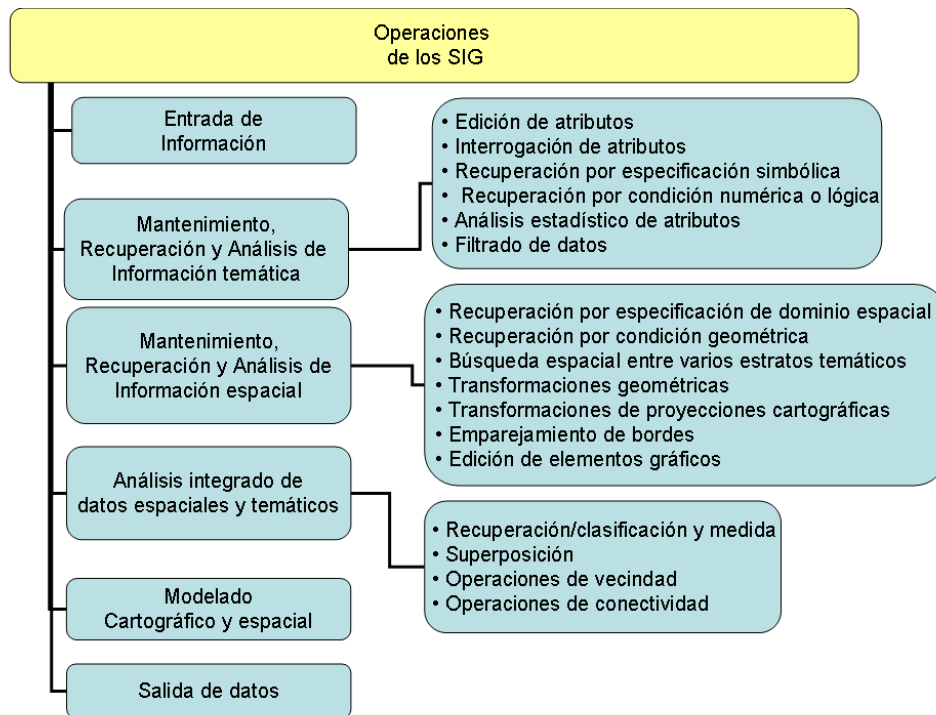


Figura 3.30: Operaciones de los SIG.

a los datos geográficos) y las operaciones de los SIG, las cuales son fundamentalmente la superposición y creación de capas de información espacial.

Este enfoque por lo tanto es el que dirige este trabajo, y son los SIG el punto de encuentro entre las distintas disciplinas. Los SIG nos permiten tanto la visualización, como la búsqueda, edición, gestión, creación y obtención de información relativa al acuífero de la Vega de Granada. De este modo, la integración de las distintas disciplinas a través del Sistema de Información Geográfica facilita la gestión y la integración de la información orientada a un usuario final (posible gestor) y es el núcleo principal del sistema de ayuda a la decisión.

Las técnicas de modelado cartográfico se definen como el uso de las operaciones básicas de un SIG en una secuencia lógica para resolver problemas espaciales complejos y donde se definen dos tipos básicos de técnicas, las descriptivas y las prescriptivas (Tomlin, 1990 [179]). El modelo cartográfico descriptivo se refiere a aquellas técnicas cuyo propósito es lograr descripciones acerca de la información espacial, distinguiendo entre técnicas que analizan y técnicas que sintetizan datos espaciales. Por otra parte el modelado cartográfico prescriptivo consta de tres fases:

- Planteamiento del problema
- Generación de soluciones al problema

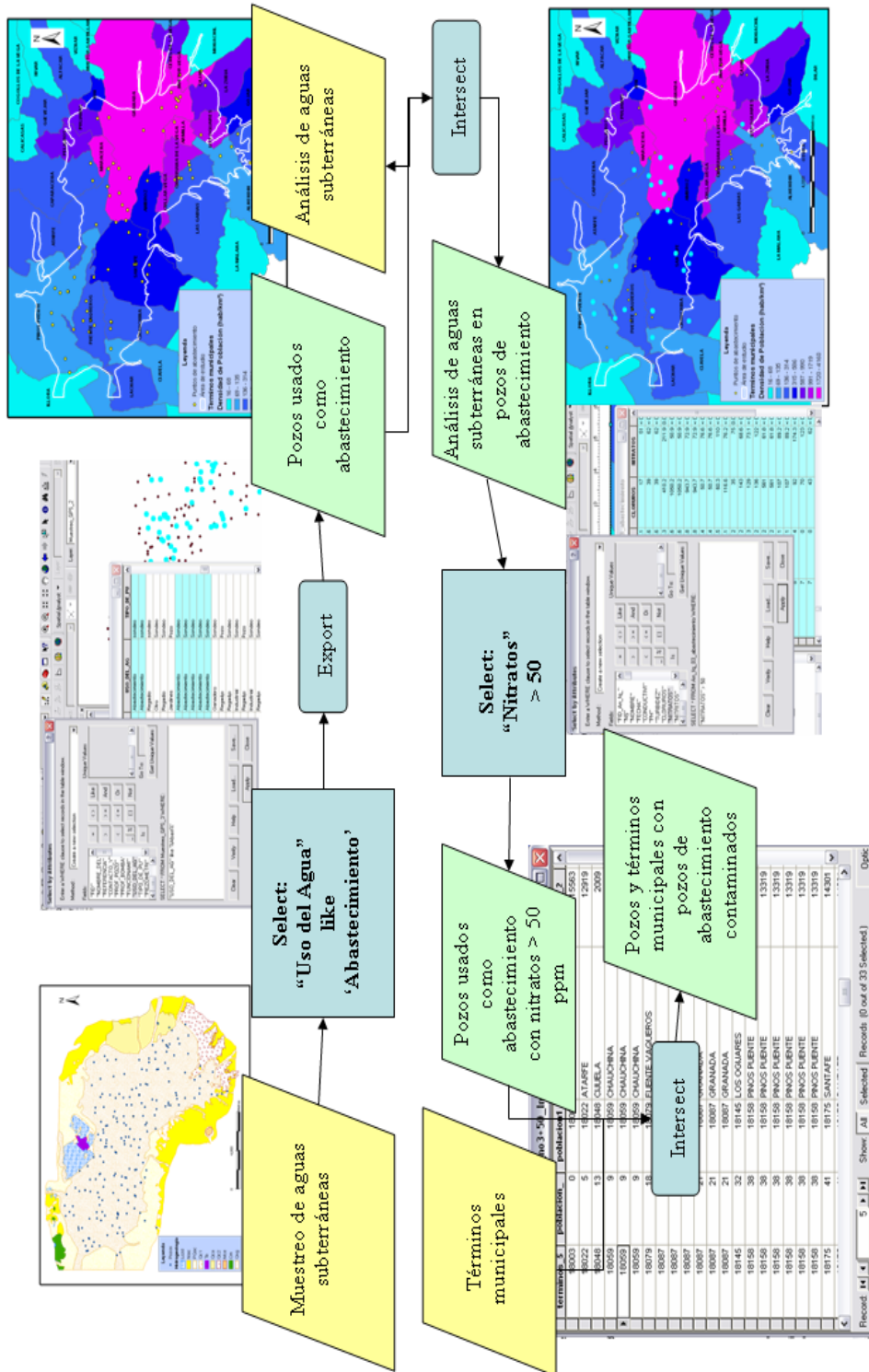


Figura 3.31: Operaciones básicas del SIG: ejemplo de creación de mapas mediante consultas a la base de datos.

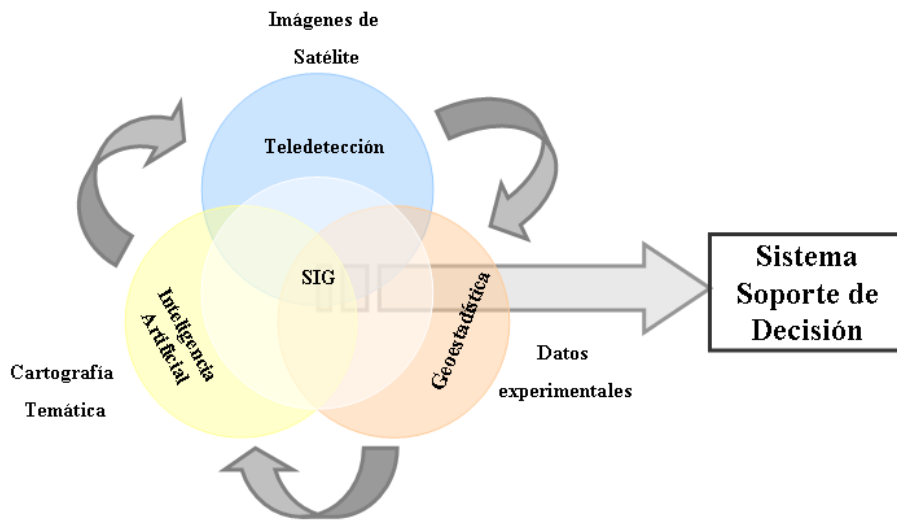


Figura 3.32: El SIG como herramienta metodológica en el Sistema Soporte de Decisión de la Vega de Granada.

- Evaluación de la soluciones

De las anteriores, la evaluación de soluciones es la fase fundamental del modelado prescriptivo y asociado a ciertas formas de localización geográfica. Por ejemplo la situación de un pozo de abastecimiento a un municipio. Esta localización generalmente se basa en la reunión de ciertas condiciones temáticas de asentamiento de actividades con el objetivo de determinar los lugares más convenientes para la actividad planteada. Así pues esta localización deberá reunir los requisitos de una contaminación mínima, una vulnerabilidad baja, y una distancia máxima a las zonas de riesgo (focos de contaminación, carreteras, etc.). Es en relación a este último aspecto es donde la mayoría de los SIG presentan ciertas limitaciones y donde destacan las técnicas de evaluación multicriterio como herramienta para asistir a la toma de decisiones espaciales, y concretamente en el ámbito que nos interesa que es la toma de decisiones en los recursos hídricos subterráneos.

3.7. La teoría de la decisión en la gestión de los recursos hídricos, y su aplicación en el S.S.D.

Para finalizar este capítulo dedicado a las bases metodológicas se ha querido incluir un apartado referente a la teoría de la decisión. La complejidad de los sistemas naturales, particularmente en el caso de la calidad de los recursos hídricos subterráneos, es bastante elevada debido a los numerosos procesos y factores que alteran o modifican dicha calidad, así como los factores de riesgo social y económico que la degradación de este recurso

puede suponer a la sociedad. Por lo tanto cualquier tipo de herramienta metodológica que ayude a la toma de decisiones es de especial interés cuando evaluamos soluciones en el contexto de las aguas subterráneas de la Vega de Granada.

En este apartado se esbozan algunas ideas generales sobre la teoría de la decisión, ya que a pesar de no tratarse de un concepto nuevo, su introducción como método automatizado, cuantificable y matemático, sí es una novedad. De este modo, un problema de decisión se entenderá como una abstracción cuantitativa o lógica de una parte de la realidad que se crea y utiliza para ayudarnos a tomar una buena decisión.

En la fase de análisis de Decisiones conviene distinguir entre lo que se denominan fases cualitativas y cuantitativas. El análisis cualitativo se basa fundamentalmente en el aspecto intuitivo que el decisor tiene del problema, a partir de su experiencia en la resolución de problemas similares o de la simplicidad del problema al que se enfrenta. Por el contrario, si el problema es complejo o el decisor no tiene suficiente experiencia con problemas similares, entonces resulta útil la realización de un análisis cuantitativo para tomar la decisión final.

En general los gerentes tienen aptitudes, que aumentan con la experiencia, para el análisis cualitativo pero, sin lugar a dudas, si comprenden la importante contribución del análisis cuantitativo y pueden incrementar su efectividad en la toma de decisiones. Esto es, si se conocen los procedimientos de la toma de decisiones cuantitativa se está en mejor posición para comparar y evaluar las fuentes de recomendaciones, tanto cualitativas como cuantitativas, y combinar las dos para tomar la mejor decisión posible y maximizar así la efectividad (Rodríguez Cotilla, 2000 [154]).

Es necesario tener en cuenta, además, que generalmente en la estructura de una decisión están presentes cursos de acción alternativos entre los que se debe hacer la elección, determinar las consecuencias de cada una de estas elecciones y establecer los objetivos. Por ejemplo, supongamos que queremos disminuir la contaminación de un río mediante la instalación de centrales de depuración de aguas residuales. La elección de un elevado número de depuradoras tendrá como consecuencia una mejor calidad de las aguas que llegan a los ríos. Establecer una calidad mínima de las aguas que llegan a los ríos es un objetivo y el coste ocasionado por este número de depuradoras es una consecuencia. El proceso de la toma de decisiones comienza por la observación cuidadosa del problema y su formulación para construir un modelo científico que intenta abstraer la esencia de la situación real. Su contribución proviene precisamente de lograr la estructuración de una situación de la vida real como un modelo matemático, con abstracción de los elementos esenciales para que se alcance una solución que concuerde con los objetivos del decisor. La solución que concuerde con los objetivos del decisor será la instalación del número óptimo de depuradoras, es decir mínimo en costes, alcanzando una calidad buena de las aguas de los ríos. Estos modelos permiten combinar muchos

factores a la vez para experimentar, buscando soluciones a problemas complejos que pueden no ser resueltos por otros enfoques quizás más convencionales. Los factores que influirían en ajustar el número de depuradoras serían por ejemplo elegir el tipo de tratamiento que se va a hacer a las aguas residuales (físicos, químicos, biológicos, etc.), los costes de mantenimiento de cada una de las depuradoras, la calidad obtenida de cada tratamiento, etc. Los análisis de sensibilidad permiten además realizar variaciones de los datos de entrada y así observar lo que sucede al resultado sin entorpecer la operación que se modela. La variación de los datos de entrada podría deberse a un aumento de población y aumento de las aguas residuales o una disminución de la misma. Esta herramienta sin duda constituye un método efectivo para modelar situaciones dinámicas.

En los años setenta se produce en gran parte de los países desarrollados, un indiscutible progreso en el campo de la teoría de la decisión, con el surgimiento del llamado paradigma decisional multicriterio, que tiene en cuenta el conjunto de objetivos usualmente en conflicto y pretende buscar un equilibrio o compromiso entre ellos, o la satisfacción en la medida de lo posible de un conjunto de metas asociadas a dichos objetivos.

La teoría de la decisión y los métodos de Evaluación Multicriterio (EMC) comenzaron a desarrollarse en el campo de los SIG en 1993 (Eastman [67]) donde se establecen las bases conceptuales de esta herramienta en el ámbito SIG.

Bajo este enfoque de la teoría de la decisión multicriterio planteamos nuestro problema concreto de mejorar la calidad de las aguas de un río mediante el emplazamiento de una o varias centrales de depuración de aguas residuales. El concepto de decisión suele implicar una cierta distribución de los recursos en un sentido u otro, es decir es necesario dotar de recursos dicha decisión. La persona o personas capaces de tomar decisiones suele denominarse centro decisor. Al aumentar el nivel jerárquico del centro decisor lo hacen también los riesgos asumidos al tomar la decisión y la cantidad de recursos puestos en juego en la misma. Este centro decisor suele tomar decisiones en función de una serie de objetivos, que no son más que lo que se pretende alcanzar mediante la distribución de recursos que implica la toma de decisiones. De este modo podemos comparar dos conceptos que suelen confundirse, una decisión y un objetivo. Así, mejorar la calidad de las aguas de un río concreto, no es una decisión, si no un objetivo. Para lograr dicho objetivo habrá que tomar muchas decisiones que serán correctas o incorrectas. Para intentar tomar la mejor decisión, es decir, la decisión correcta, los centros decisores se valen de una serie de criterios que se definen como aspectos considerados importantes para tomar la decisión. Estos criterios son los que incorporan un mayor grado de discrecionalidad al proceso de toma de decisiones y los que más dificultan su modelización, ya que estos criterios varían según formación, intereses personales, intereses políticos, sociales, etc.

El proceso mediante el cual el centro decisor combina sus objetivos, con los criterios que tiene para lograrlos, y que conduce a una decisión final, se denomina análisis de la decisión. De forma más o menos consciente y estructurada todos llevamos a cabo un análisis de la situación antes de tomar una decisión determinada. Esto es, evaluamos las ventajas e inconvenientes de una decisión, analizamos las posibles consecuencias, los costes que ocasionará, etc. En definitiva, simulamos mentalmente el proceso de toma de decisiones, generando un modelo que nos permite evaluar a priori el resultado de la decisión.

Hay otros conceptos muy importantes a la hora de describir el proceso de toma de decisiones. En primer lugar necesitamos contar con varias alternativas para compararlas y decidir cual es la que mejor se ajusta a los objetivos. Estas alternativas en el caso de la aplicación en SIG están representadas por objetos o unidades espaciales (localización de las centrales depuradoras). El proceso de análisis consiste en comparar varias alternativas que se diferencian en la forma en la que se combinan los criterios, la importancia relativa de cada uno, etc. Los criterios medidos en forma continua son factores y en el caso de la instalación de una central de depuración de aguas residuales son por ejemplo la distancia a los distintos núcleos de población de los que recogerá las aguas. Así una menor menor distancia supondrá un menor coste económico. Los criterios de tipo limitante restringen la disponibilidad de las alternativas en función de la actividad evaluada. Por ejemplo, las áreas de vegetación protegida serían descartadas para la ubicación de la estación depuradora y el resto de usos del suelo sería válido.

Una vez que nos decidimos por una alternativa, surge uno de los grandes problemas del proceso de toma de decisiones: la incertidumbre. A veces se llama suerte, pero en realidad la incertidumbre es un concepto inherente a todo sistema complejo. La incertidumbre surge de la incapacidad que tenemos de controlar todas las variables que están implicadas en un proceso concreto. Aunque es imposible evitarla, podemos evaluar la incertidumbre mediante métodos estadísticos de complejidad variable. Esta incertidumbre está definida por Zadeh (1973 [196]) mediante el criterio de incompatibilidad: conforme aumenta la complejidad de un sistema, se reduce nuestra capacidad de hacer afirmaciones precisas y suficientemente significativas sobre su comportamiento, hasta llegar a un umbral sobre el cual la precisión y la relevancia son conceptos mutuamente excluyentes. La combinación del proceso de análisis de la decisión, con la incertidumbre inherente al sistema, dan lugar a un resultado. El hecho de que la incertidumbre forme parte del resultado hace que éste también sea incierto. En otras palabras, no podemos asegurar que el resultado se ajuste al objetivo por muy elaborado que sea el proceso de análisis de la situación. Por tanto, cuando el decisor opta por una alternativa, asume cierto riesgo de que el resultado no se ajuste a sus objetivos. El riesgo asumido, como

veremos más adelante, puede controlarse mediante el tipo de método escogido para la evaluación de la solución.

Este ejemplo recoge los principales conceptos utilizados en la teoría de la decisión multicriterio, los cuales serán aplicados en capítulos posteriores. La complejidad de la toma de decisiones en el medio ambiente, y concretamente en el caso de la calidad de los recursos hídricos, justifican por otra parte la aplicación de estos métodos a diversas situaciones en el entorno de la gestión de la calidad del agua debido fundamentalmente a las numerosas circunstancias, procesos y fenómenos que participan en la modificación de las características hidrogeoquímicas del acuífero y que no son perfectamente conocidas, así como a los posibles efectos económicos o socioeconómicos que puede tener la modificación de la calidad de estos recursos.

Capítulo 4

Análisis temático y espacial de los datos

4.1. Introducción

En el contexto general de esta investigación, se pretende establecer la relación entre *datos*, *análisis* y *(geo)información*, en forma de una ecuación aparentemente sencilla pero, ciertamente, difícil de resolver. Los datos mostrados en los capítulos anteriores suponen un volumen importantísimo de información, que han costado un esfuerzo económico y humano notables. Llegar a obtener la *(geo)información*, y consecuentemente poder tomar decisiones sobre la misma, pasa obligadamente por seleccionar las mejores técnicas, métodos o procedimientos de análisis. Este asunto se abordaba metodológicamente en el capítulo anterior y, de modo aplicado, es el que se aborda en este y el siguiente capítulo.

El análisis en su doble vertiente, espacial y temática de datos implica básicamente dos tipos de actividades: la descripción de patrones espaciales y el análisis de relaciones entre patrones espaciales (Unwin, 1981 [181]). Normalmente, el análisis descriptivo supone un primer paso que suele conducir a la formulación de cuestiones más complejas acerca de las relaciones entre los patrones espaciales (por ejemplo, ¿cuales son las principales variables espaciales que explican la distribución de la contaminación de los suelos?) (Openshaw, 1991 [131]). Estas actividades pueden comportar tanto análisis univariantes como multivariantes, que a menudo, aunque no exclusivamente, son de naturaleza estadística, y en muchos casos incluyen la modelización matemática y otras formas ad hoc de análisis geográfico. En este sentido, el análisis espacial de datos es un amplio campo que incluye no sólo la estadística, espacial (geoestadística) y no espacial (clásica), sino que comprende también a la modelización espacial, implicando esta a su vez un extenso conjunto de modelos (determinísticos, estocásticos, empíricos, etc.) (An-

selin y Getis, 1993; Fisher et al., 1996 [12, 73]). Estos modelos, por ejemplo, simulación numérica del flujo subterráneo están fuera del objetivo de este trabajo.

El análisis estadístico de datos está constituido por dos grandes etapas: el análisis exploratorio y el análisis confirmatorio (Haining, 1990 [89]). El Análisis Exploratorio de Datos (AED), es una primera fase en la que el objetivo principal es la indentificación y descripción de las propiedades estadísticas de los datos, la identificación de potenciales estructuras presentes en los datos y la formulación de hipótesis a partir de los mismos. El AED se caracteriza por la utilización de métodos robustos, es decir, poco sensibles a valores extremos (muy interesantes en la evaluación temporal de la calidad del agua subterránea) y en la utilización de gráficos para presentar informes de los datos de forma que sea fácil su interpretación. En el caso de los datos univariantes, las medidas de estructura pueden incluir propiedades de localización, escala, sesgo y distribución, mientras que en un contexto multivariante esto puede extenderse a la identificación de relaciones entre variables (diagramas de dispersión, grado de correlación, etc.). El Análisis Confirmatorio de Datos (ACD) está, en cambio, más próximo a la inferencia estadística tradicional (proporcionando intervalos de confianza y comprobando hipótesis), pero también incluye análisis de sensibilidad o influencia (evaluando la robustez del ajuste de un modelo y las conclusiones estadísticas de los datos), así como el análisis de los residuales (Haining, 1990 [89]). En la práctica estas dos fases no son utilizadas de modo secuencial, sino alternante.

En nuestro caso, debemos tener en cuenta que los sistemas naturales, en concreto los sistemas acuíferos, son objetos complejos en donde las hipótesis utilizadas por los métodos tradicionales de tratamiento de la información son a menudo demasiado simples. Así, los datos espaciales presentan en particular ciertas características espaciales que dificultan sensiblemente su análisis mediante las técnicas estadísticas clásicas, las cuales están relacionadas esencialmente con la dependencia espacial y la heterogeneidad espacial (Fisher et al., 1996 [73]). La primera se refiere a la relación entre datos georreferenciados, ya que los datos que están próximos en el espacio tienden a ser más parecidos entre sí que los que están lejos (Cressie, 1993 [48]), y por tanto tienden a no ser independientes, supuesto básico de la estadística clásica. La heterogeneidad espacial surge cuando no existe uniformidad espacial en los efectos de la dependencia espacial y/o de las relaciones entre variables concretas (Haining, 1990; Fisher et al., 1996 [89, 73]). Estos efectos espaciales pueden invalidar muchos de los procedimientos estadísticos estándares y dar lugar a errores en la inferencia estadística, falseando las indicaciones de la validez de los modelos, etc. Es bien conocido, por ejemplo, que la autocorrelación de los datos espaciales provoca modificaciones en los intervalos de confianza de los test estadísticos (Haining, 1990; Bosque Sendra y Moreno Jiménez, 1994 [89, 22]). Consecuentemente,

el análisis espacial de datos debe ir más allá del análisis estadístico estándar (Anselin y Getis, 1993 [12]).

Por otra parte, la aplicación de técnicas de análisis temático y espacial de datos se ha visto notablemente facilitada desde la incorporación de este tipo de herramientas en los SIG. Numerosos autores afirman, sin embargo, que una de las mayores deficiencias en los SIG actuales es la falta de funciones analíticas y de modelización, y afirma incluso que el futuro éxito de la tecnología SIG dependerá en gran medida de la incorporación de capacidades de este tipo más potentes (Openshaw, 1991; Fisher et al., 1996 [130, 73]). Desde entonces numerosos softwares comerciales han incluido en su aplicación diversas herramientas estadísticas, geoestadísticas y de modelización, aunque cada uno de los software se ha especializado en un tipo de datos. Así ArcGIS se ha especializado en las herramientas de análisis vectorial, ENVI y ERDAS en el análisis raster fundamentalmente de imágenes de satélite, Idrisi en el análisis raster de datos geográficos, etc. Sin embargo, una característica común a todos ellos es la posibilidad de incluir scripts desarrollados por el usuario, los cuales permiten incorporar al análisis de datos la técnica deseada.

En esta línea de visión futura queda la inclusión de nuevas herramientas innovadoras de análisis espacial que sean especialmente adecuadas al ambiente SIG. Entre estas nuevas herramientas están las redes neuronales, los árboles de decisión, y distintas técnicas provenientes del campo de la inteligencia artificial, que se han comentado en el capítulo anterior. El principal interés de estas técnicas es la ausencia de supuestos sobre los datos espaciales y su capacidad de procesar datos ruidosos y modelizar relaciones no lineales (Rigol Sánchez, 2000 [152]).

El presente capítulo está dedicado, fundamentalmente, al análisis descriptivo de los datos hidrogeoquímicos, aunque también se incluyen análisis de relaciones entre variables (análisis multivariante). Se aplicaron para ello técnicas tanto univariantes (histogramas, variografía, etc.) como multivariantes (clustering, análisis de componentes principales, etc.) Se tratan aquí, por tanto, aspectos relacionados con el análisis (exploratorio) estadístico espacial, mientras que la elaboración de modelos (predictivos) sobre los datos, o modelización espacial es abordada en el capítulo 5. Uno de los objetivos de la caracterización de los patrones de distribución de las variables hidrogeoquímicas será la detección de zonas anómalas de posible contaminación, así como la clasificación e identificación de los principales mecanismos que influyen la química de las aguas subterráneas.

Seguidamente se expondrán de forma sintética los resultados del análisis temático de los datos experimentales, principalmente, de aguas y suelos. Para ello, se ha aplicado un procedimiento de análisis exploratorio uni y multivariante. Se completa este análisis

abordando el aspecto espacial o de regionalización a través de los métodos geoestadísticos.

4.2. Análisis Exploratorio de Datos

El análisis temático de los datos comienza por una exploración preliminar que permita una identificación inicial de las propiedades estadísticas de los mismos, antes de proceder al análisis de datos multivariante. En esta fase se hace énfasis en la descripción estadística de las variables (valores máximo y mínimo, media, varianza, función de distribución, etc.) de los datos, cuyo conocimiento resultará útil tanto en el desarrollo de hipótesis como a la hora de dar forma a los sucesivos análisis estadísticos (Haining, 1990; Cressie, 1993; Bosque Sendra y Moreno Jiménez, 1994 [89, 48, 22]). Además del interés que tienen los datos en sí mismos (sumarios estadísticos), este análisis es necesario para indicar si serán necesarias transformaciones especiales de los datos para subsecuentes análisis estadísticos (por ejemplo, transformación logarítmica).

Con este enfoque se ha realizado un Análisis Exploratorio de los Datos (AED) experimentales de la Vega de Granada recogidos en este trabajo, especialmente de los datos hidrogeoquímicos y edafológicos. Debido al volumen tan importante de variables y de resultados, éstos se expondrán de forma muy resumida para aquellas variables que he considerado más representativas.

La interpretación de los datos puntuales se fundamenta habitualmente en la elaboración de los histogramas de frecuencias, la observación de los estadísticos básicos, los diagramas de caja y los gráficos normales Q-Q. Estos métodos son sencillos y fácilmente aplicables y permiten obtener un notable grado de conocimiento acerca del comportamiento de las variables. Además, es posible en numerosas ocasiones alcanzar conclusiones similares a las que se derivan de la aplicación de complejos métodos multivariantes (Rigol-Sánchez, 2000 [152]).

El análisis univariante del conjunto de variables estudiadas se presenta dividido en varios grupos, atendiendo a la naturaleza de las variables: hidroquímicas, hidrogeológicas y edafológicas. Algunas de las variables hidrogeológicas mostradas en el capítulo 1 (gradiente, elevación del terreno, espesor no saturado, etc.) son resultado de operaciones diversas entre variables algunas de las cuales se muestran en este capítulo.

Las tablas 4.1, 4.2, 4.3, 4.4 y 4.5 sintetizan los estadísticos básicos de todas las variables, incluyendo medidas de dispersión, de tendencia central y de forma. También se realizaron y examinaron histogramas, diagramas de caja (box-plot) y gráficos Q-Q (cuantiles). Para el análisis de las muestras estudiadas en laboratorio, así como para las obtenidas a partir de otras fuentes se utilizaron distintos paquetes estadísticos comerciales (SPSS, Excel, Matlab y los paquetes incluidos en los SIG ArcGIS e Idrisi)

	Rango	Minimo	Máximo	Media		Mediana	Desv. tip.	Varianza	Asimetría	Curtosis
	Estadístico	Estadístico	Estadístico	Estadístico	Error típico	Estadístico	Estadístico	Estadístico	Estadístico	Estadístico
Conductividad	8396	434	8830.00	1446.12	62.56	1222.00	900.13	810234.88	3.40	21.67
pH	2.5	6.5	9.00	7.25	0.03	7.00	0.38	0.14	1.39	1.57
Turbidez	72.4	0	72.40	5.59	0.84	2.00	12.04	144.86	4.22	18.52
Cloruros	2245	4	2249.00	95.87	13.43	52.00	193.29	37362.43	7.93	79.06
Nitratos	341.5	0.5	342.00	61.93	3.42	50.20	49.20	2420.85	2.19	6.70
Sulfatos	5343.6	6.5	5350.10	458.86	41.69	237.40	599.83	359795.32	3.67	21.60
Boro	2	0	2.00	0.12	0.02	0.00	0.25	0.06	3.79	18.69
Bicarbonato	390.6	185.6	576.20	365.68	5.48	362.80	78.78	6206.86	0.10	-0.51
Carbonato	2.3	0	2.30	0.01	0.01	0.00	0.16	0.03	14.39	207.00
Calcio	664.7	27.6	692.30	169.71	9.03	123.00	129.97	16892.15	2.15	4.44
Sodio	1495.4	2.2	1497.60	69.00	9.06	41.30	130.31	16981.04	7.66	74.40
Magnesio	273.9	19.7	293.60	64.13	2.57	53.80	36.94	1364.54	2.64	11.62
Potasio	40.3	0.3	40.60	3.66	0.34	2.50	4.88	23.85	4.78	26.82
Fluoruros	3.86	0	3.86	0.65	0.05	0.34	0.75	0.57	2.27	5.51

Tabla 4.1: Estadísticos básicos de los elementos fisicoquímicos primarios del agua.

	Rango	Minimo	Máximo	Media	Mediana	Desv. tip.	Varianza	Asimetría	Curtosis
Plomo	5.379	0.021	5.400	0.704	0.705	0.709	0.503	4.503	28.305
Cromo	9.930	0.000	9.930	0.781	0.470	1.371	1.880	5.281	31.479
Cobre	19.800	0.000	19.800	2.728	1.225	4.148	17.206	2.593	7.248
Hierro	320.372	0.028	320.400	49.095	17.895	73.362	5381.956	2.152	4.117
Manganeso	295.600	0.000	295.600	18.774	1.685	55.320	3060.302	3.741	14.164
Antimonio	0.119	0.000	0.119	0.018	0.000	0.028	0.001	1.539	1.793
Arsénico	1.653	0.000	1.653	0.409	0.270	0.368	0.136	1.749	3.063

Tabla 4.2: Estadísticos básicos de los metales pesados del suelo.

que permiten realizar fácilmente los cálculos de estas funciones. En general, los resultados muestran que la mayoría de los elementos mayoritarios presentan distribuciones normales y los minoritarios asimétricas de tipo positivo (en la mayoría de los casos log-normales), con valores máximos anómalos. También se observan en algunos casos mezclas de distintas poblaciones, identificables en distintas zonas del área de estudio. Debido a la fuerte asimetría en la distribución de algunos elementos, es necesario recurrir al cálculo de estadísticos más robustos como son la media recortada, el estadístico-M y estudio de percentiles.

Los gráficos de caja o box-plot facilitan la visualización de un conjunto de estadísticos como: valores extremos, percentiles y mediana, que proporcionan una información bastante completa, entre otras cosas, del grado de dispersión de los datos y el grado de asimetría de la distribución.

	Rango	Minimo	Máximo	Media	Mediana	Desv. tip.	Varianza	Asimetría	Curtosis
Bacterias aerobias	450	0	450	86.85	29.5	127.119	16159.362	1.747	1.998
Coliformes totales	340	0	340	37.85	0.0	79.243	6279.464	2.667	7.139
Coliformes fecales	123	0	123	9.50	0.0	23.759	564.513	3.538	13.945
Estreptococos fecales	67	0	67	4.60	0.0	13.504	182.349	3.422	12.411
Clostridios sulfitorreductores	0	0	0	.00	0.0	.000	.000	.	.
E. Coli	40	0	40	4.93	0.0	9.542	91.046	2.176	4.458
DQO	878.619	.000	878.619	131.60225	47.9	217.553771	47329.643	2.814	7.308
Sólidos en suspensión	541.000	.000	541.000	26.87258	2.1	96.317745	9277.108	4.686	22.879

Tabla 4.3: Estadísticos básicos de la microbiología del agua.

	Rango	Mínimo	Máximo	Media	Mediana	Desv. tip.	Varianza	Asimetría	Curtosis
Conductividad	4811.900	188.100	5000.000	1289.133	1062.000	957.502	916809.633	1.809	4.001
Magnesio	306.500	4.500	311.000	40.408	27.600	45.542	2074.119	3.676	17.673
Calcio	604.100	21.000	625.100	141.239	108.700	114.043	13005.724	2.363	6.767
Sodio	427.900	3.600	431.500	72.776	45.300	90.274	8149.329	2.601	6.997
Potasio	368.400	0.800	369.200	16.718	4.860	54.206	2938.263	5.851	34.448
Flúor	0.000	0.000	0.000	0.000	0.000	0.000	0.000		
Cloruros	394.910	21.059	415.969	93.726	64.625	84.526	7144.684	1.767	3.177
Nitritos	118.837	0.000	118.837	17.973	11.751	22.821	520.796	2.417	7.709
Nitratos	1302.871	0.000	1302.871	148.468	69.371	206.171	42506.643	3.368	14.257
Sulfatos	2226.573	21.833	2248.406	412.193	269.330	435.574	189724.644	2.158	5.333
Plomo	1.094	0.000	1.094	0.526	0.508	0.301	0.091	0.104	-0.790
Cobre	71.500	1.700	73.200	19.103	16.350	13.173	173.528	1.569	3.627
Hierro	22.147	0.003	22.150	1.673	0.689	3.451	11.908	4.023	18.595
Manganeso	77.186	0.044	77.230	8.449	2.750	15.546	241.676	3.070	9.659
Zinc	21.531	0.009	21.540	5.229	2.830	5.404	29.209	1.234	0.572
Cadmio	0.199	0.000	0.200	0.098	0.097	0.059	0.004	0.093	-1.187
Cromo	8.890	0.050	8.940	1.552	0.965	1.713	2.935	2.757	8.641

Tabla 4.4: Estadísticos básicos de los elementos químicos de la solución del suelo.

	Rango	Mínimo	Máximo	Media	Desv. tip.	Varianza	Asimetría	Curtosis
Cota piezométrica	157.9	525.1	683.0	585.063	34.1542	1166.512	839	.792
Transmisividad	544.29	3	544.32	4355.44	8979.253	80626984.3	4.488	22.978
Permeabilidad	1234.245	.041	1234.286	169.542	264.173	69787.587	2.361	5.887
Litología zona no saturada	4.72	.00	4.72	2.1536	1.08096	1.168	.175	-7.18

Tabla 4.5: Estadísticos básicos de los parámetros relacionados con el acuífero.

En las **variables hidrogeoquímicas** estudiadas, son frecuentes los histogramas que representan distribuciones bimodales para los elementos mayoritarios y algunas propiedades físicas del agua subterránea. La representación de los histogramas se ha realizado identificando espacialmente el conjunto de datos de los casos más interesantes (generalmente las colas del histograma o las distintas poblaciones). De esta forma se consiguen identificar los valores anómalos o la concentración de valores debida a factores externos (por ejemplo, posible contaminación antrópica) (figura 4.1). Como dato de interés, ligado a la contaminación de las aguas para consumo público, cabe decir que un 70 % de las muestras de agua analizadas no cumplen los criterios establecidos por la ley para consumo público. Destacan los nitratos y los sulfatos (tabla 4.6) con un 50 % de los casos, aproximadamente, con valores superiores a los establecidos por la ley. Es decir, el 50 % de las muestras de agua analizadas no son potables debido a los nitratos y a los sulfatos. Otro variable importante que resta potabilidad al agua es la turbidez, para la que un 20 %, aproximadamente, tampoco cumple las normas de potabilidad. Esta variable puede estar influenciada por factores relacionados con la suciedad del pozo y no del acuífero, por lo que no es realmente indicativa. Con un porcentaje cercano al 10 % de muestras de agua no apta está la conductividad, los fluoruros y el manganeso. Y en casos mucho más aislados: los cloruros, el boro y el sodio.

Desde esta perspectiva podemos asegurar que ciertas zonas del acuífero tienen problemas graves debido a la contaminación del agua del acuífero, fundamentalmente por la actividad agraria y ganadera (nitratos y sulfatos).

Otra perspectiva interesante de los datos es identificar el número de poblaciones en las variables (figura 4.1 y figura 4.2). En general podemos establecer al menos tres poblaciones de datos que se identifican en la mayoría de las variables mayoritarias (aquí mostramos el magnesio, el calcio, la conductividad y el flúor). En todas ellas se localiza una primera inflexión (valores más bajos) que recoge a las muestras localizadas en la zona de recarga del acuífero por los ríos Genil, Darro, Dílar, Monachil y Beiro y corresponde a las aguas con menor concentración de cationes y aniones y por lo tanto de mejor calidad, procedentes de los relieves circundantes. También se observa una segunda población de valores intermedios y que localizada en la mitad oeste del acuífero y que corresponden a la mezcla de aguas de recarga de los ríos con la recarga de riego y con las aguas de borde. Las aguas de borde, fundamentalmente en la zona del río Salado y al sur de Sierra Elvira, se sitúan en las colas de la distribución del histograma, con valores altos o muy altos debido a la disolución de los materiales evaporíticos que atraviesan estas aguas subterráneas hasta el acuífero. En el caso de los metales pesados se observa una única población y no muestra una zonación tan clara en el espacio, como en el caso de los elementos mayoritarios.

En el caso de las **variables microbiológicas** tenemos un 40 % de las muestras de agua subterránea que no son potables según los límites de potabilidad debido a los coliformes totales y un 35 % debido a los coliformes fecales y al *escherichia coli*. La distribución espacial de estos datos ya se ha mostrado en el apartado correspondiente a los datos, donde se indicó que no era posible realizar una estimación espacial de estas variables (excepto para las bacterias aerobias) debido al poco tiempo de supervivencia que estos organismos tienen en las aguas subterráneas. Al observar el histograma (figura 4.3) comprobamos que es asimétrico positivo con abundancia de valores nulos, motivo por el que no se consigue modelar correctamente el variograma. Cabe comentar que afortunadamente la ausencia de microorganismos indica una buena calidad microbiológica.

Las **variables edafológicas** tienen histogramas más heterogéneos que los de las aguas subterráneas (figura 4.4) e influidos por ejemplo en el caso del sodio, los nitritos, la conductividad y los cloruros por la agricultura y los sistemas de regadío, o el mecanismo de deposición de los suelos en el caso de las arenas, la humedad de saturación y la humedad natural.

Finalmente en los histogramas de los **parámetros hidrogeológicos** (4.5) se aprecia la fuerte asimetría de las variables de transmisividad y permeabilidad, motivo por el que su espacialización requiere de una transformación previa (apartado 4.4.2). Los

Variable analizada	Nº de muestras no potables	% de muestras no potables
Conductividad	22	10.68
pH	1	0.49
Turbidez	40	19.42
Cloruros	12	5.83
Nitratos	104	50.49
Nitritos	0	0.00
Sulfatos	98	47.57
Fosfatos	0	0.00
Boro	9	4.37
Bicarbonato	0	0.00
Carbonato	0	0.00
Calcio	0	0.00
Sodio	10	4.85
Magnesio	0	0.00
Potasio	0	0.00
Amonio	2	0.97
Fluoruros	22	10.68
Cromo	0	0.00
Plomo	0	0.00
Cobre	0	0.00
Hierro	0	0.00
Manganeso	22	10.68
Antimonio	0	0.00
Arsénico	0	0.00
<i>Total de variables</i>	161	78.16
<i>Total variables (excluyendo la turbidez)</i>	143	69.42

Tabla 4.6: Resumen estadístico de las variables del agua analizadas no aptas para consumo público.

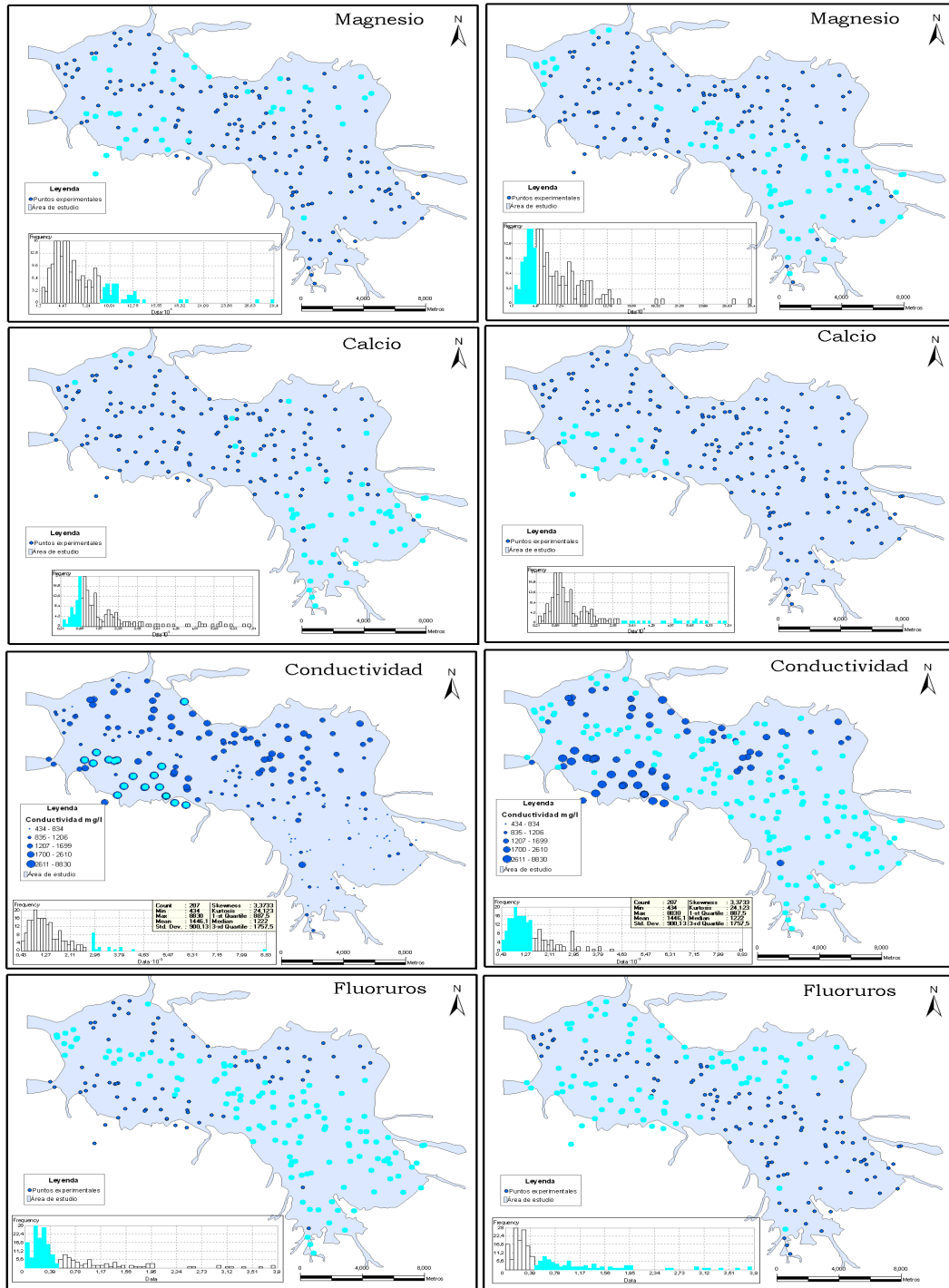


Figura 4.1: Histogramas y situación espacial de las distintas poblaciones de algunos de los parámetros fisicoquímicos del agua subterránea.

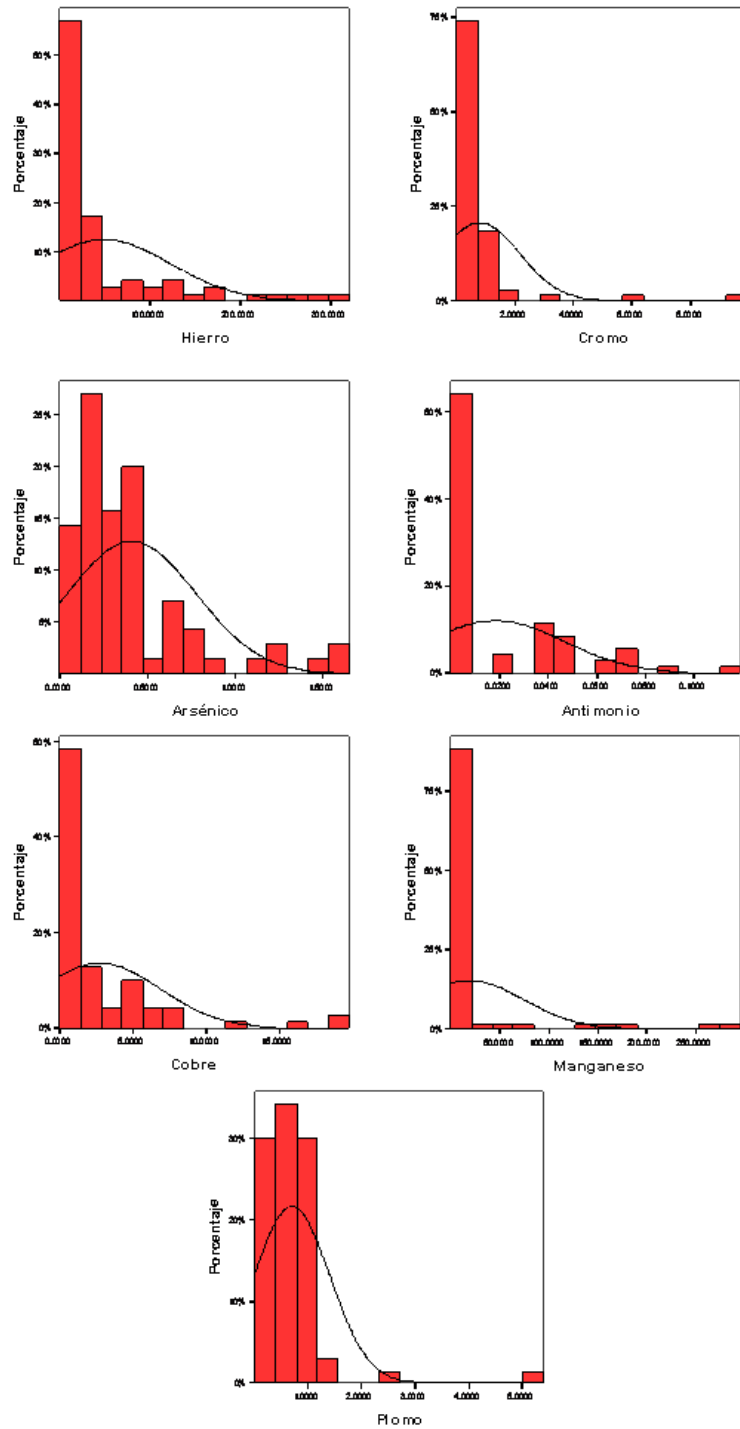


Figura 4.2: Histogramas de algunos metales pesados del agua subterránea.

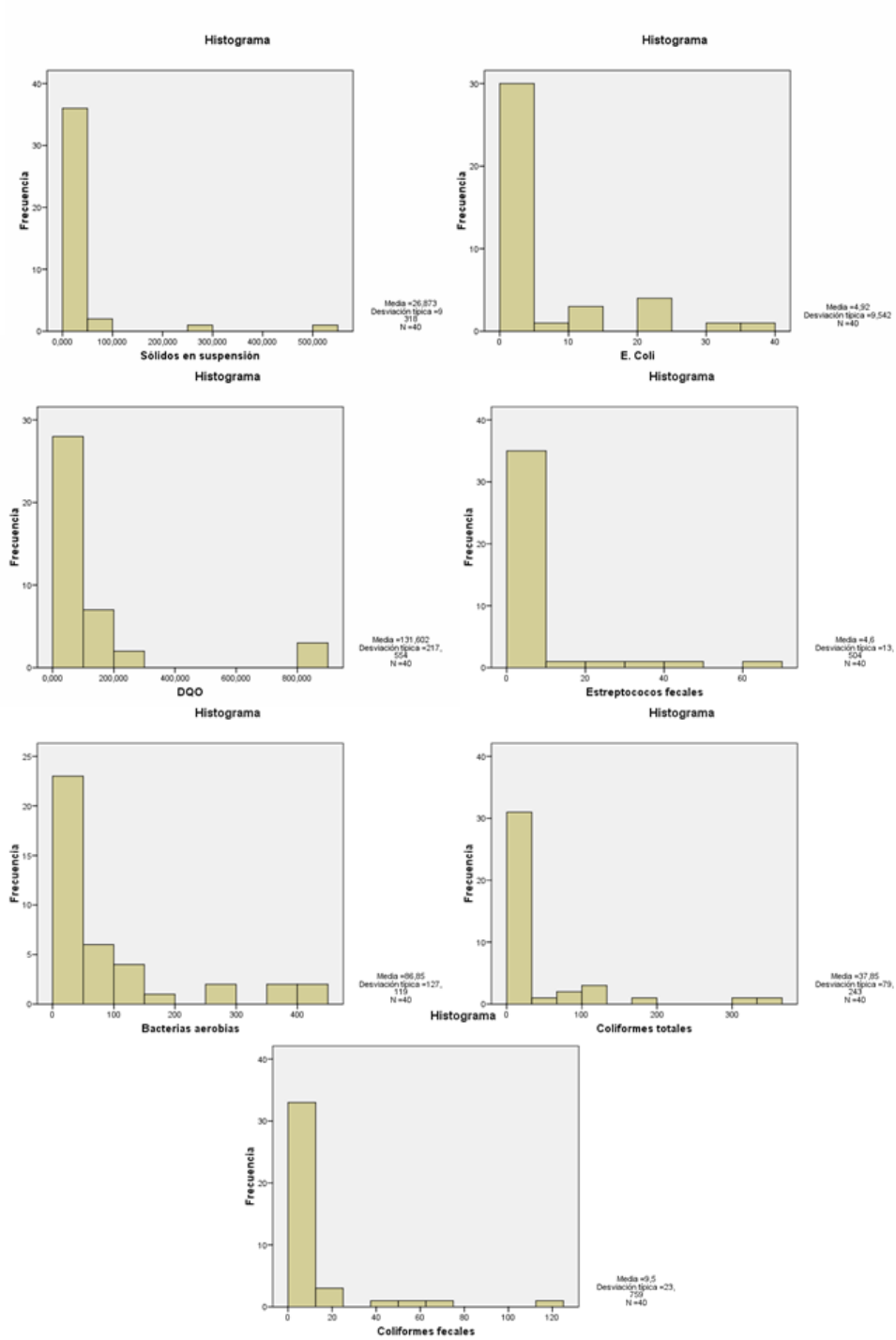


Figura 4.3: Histogramas de las variables microbiológicas de las aguas subterráneas.

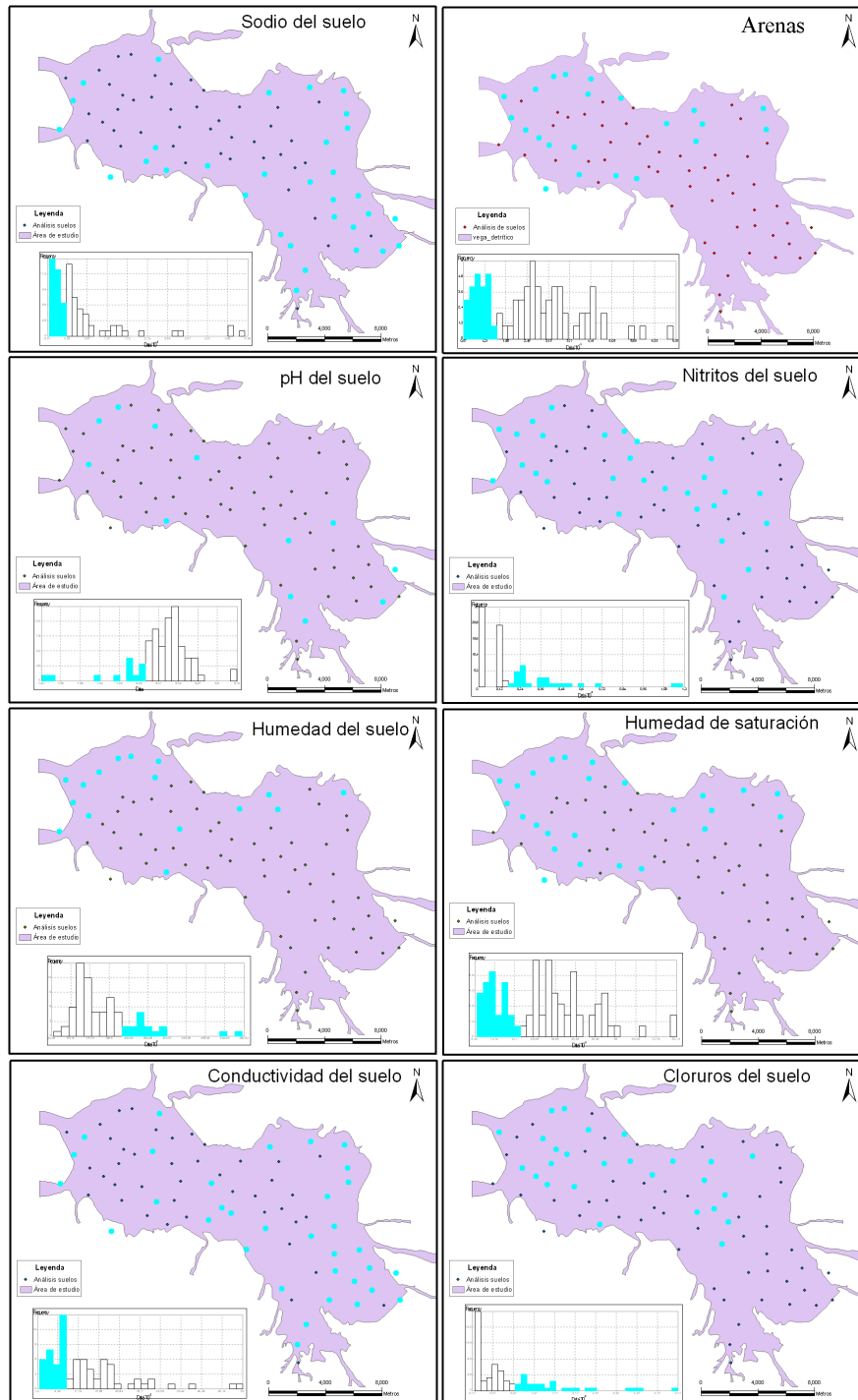


Figura 4.4: Histogramas y regionalización de algunos de las variables del suelo.

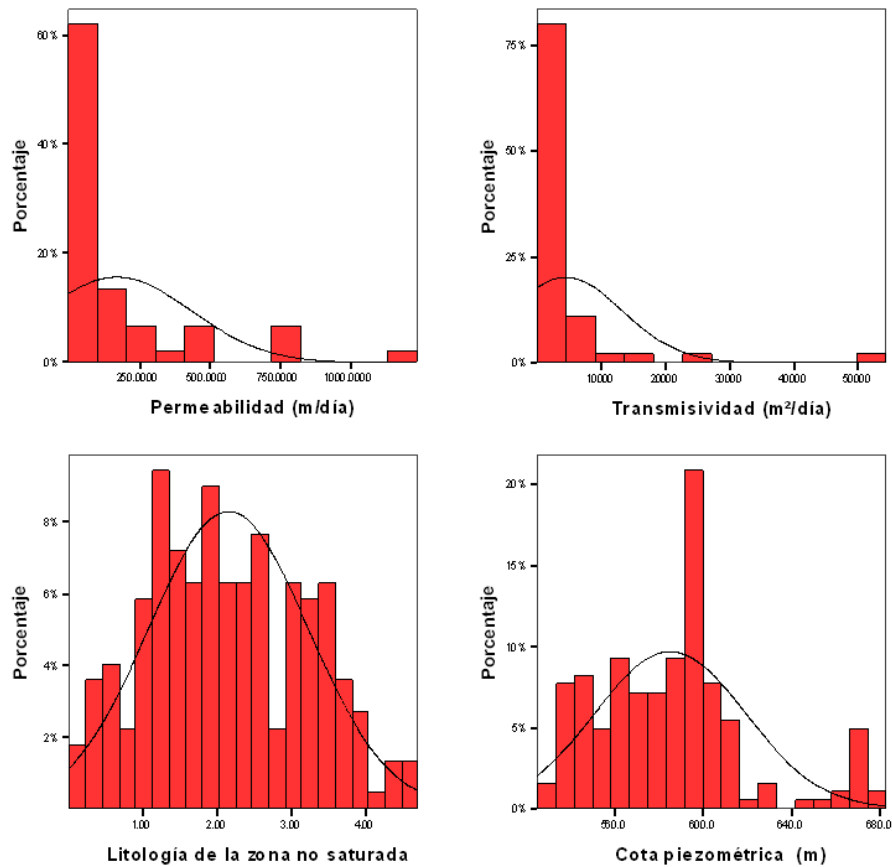


Figura 4.5: Histogramas de algunos parámetros relacionados con el acuífero.

histogramas de la piezometría y de la litología de la zona no saturada son normales y por tanto no requieren de transformaciones para su espacialización.

Otra representación que ayuda a la interpretación de los datos experimentales son los gráficos box-plot o diagramas de caja, que fueron desarrollados por Turkey (1977 [180]) y consisten en la representación de la distribución de modo que el 50% de los casos se representa en el interior de la caja, en las patillas o bigotes el 25% superior e inferior y los casos atípicos y los extremos con un círculo y un asterisco, respectivamente. La línea que cruza la caja es la mediana y la línea que señala el segmento de la patilla o el bigote es el valor más grande que no llega a ser atípico. Los casos atípicos son los valores alejados más de 1,5 longitudes de caja del percentil 75, y los extremos alejados más de 3 longitudes de caja del percentil 75. Es una forma alternativa al histograma para representar una distribución, algunos ejemplos para propiedades físicas, elementos mayoritarios y metales de suelos y aguas tanto como microbiología de aguas se representan en la figura .

En la figura 4.1 ya se observaron algunos de los histogramas de las **variables hidroquímicas del agua**, sin embargo, la identificación de umbrales que permitan diferenciar las poblaciones atípicas y extremas es un aspecto de gran importancia para posteriores tratamientos multivariantes y geoestadísticos de la calidad del agua (figura 4.6).

Las **variables físicas del suelo** no presentan distribuciones asimétricas y excepcionalmente algunos casos extremos (ejemplos: *limos, pH y carbono orgánico*; figura 4.7).

Sin embargo, las **variables químicas del suelo** (incluyendo los metales pesados) sí que presentan este tipo de distribuciones asimétricas positivas (figura 4.7) como por ejemplo: la *conductividad, los nitratos y el cobre*. Esto es debido a la manipulación antrópica de los suelos de cultivo mediante la adición de fertilizantes y por lo tanto, adición de elementos químicos.

La fuerte asimetría de algunos **parámetros relativos al acuífero** (transmisividad y permeabilidad) queda presente en los gráficos box-plot con la presencia de 3 datos extremos para cada una de las variables, además del amplio rango de valores que van desde cero a 1200 *m/día* y de cero a 50.000 *m²/día*. La cota piezométrica presenta algunos valores atípicos y la litología de la zona no saturada no presenta casos atípicos ni extremos, ya que al tratarse de una reclasificación de tamaños de grano ponderada por el espesor, los valores quedan muy normalizados.

Otro aspecto que debe considerarse en el análisis exploratorio de datos es el ajuste de la distribución empírica de los datos a alguna de las distribuciones clásicas en estadística, como por ejemplo Normal y Log-normal

La normalidad de las distribuciones se estudia a través de estadísticos como el de Kolmogorof-Smirnov y el de Shapiro-Wilk, aunque para muestreos grandes, como es en este caso, son más recomendables los gráficos Q-Q normales. En estos gráficos se compara el valor observado frente a la puntuación típica que teóricamente le correspondería a ese valor en una distribución normal estandarizada. Así, cuando los puntos se agrupan en torno a la diagonal, es indicativo de que la muestra procede de una población normal. De este modo, podemos resumir que la mayor parte de las variables químicas y físicas del agua corresponden a poblaciones log-normales, siendo la transformación de la variable al logaritmo natural, distribuciones Q-Q perfectas (figuras 4.9 y 4.10) y exceptuando dos casos puntuales como el de los bicarbonatos o el magnesio, elementos que suelen presentar distribuciones normales (Clotier et al., 2008 [45]).

Las variables químicas de los suelos también se distribuyen como log-normales (figura 4.11), sin embargo, las variables físicas como la materia orgánica, el % de arenas, limos y arcillas o la humedad, tienen distribuciones normales (figura 4.11).

La evidencia de dos poblaciones en la representación Q-Q de los **parámetros hidrogeológicos** en gráficos Q-Q (4.12) de la transmisividad y la permeabilidad es clara.

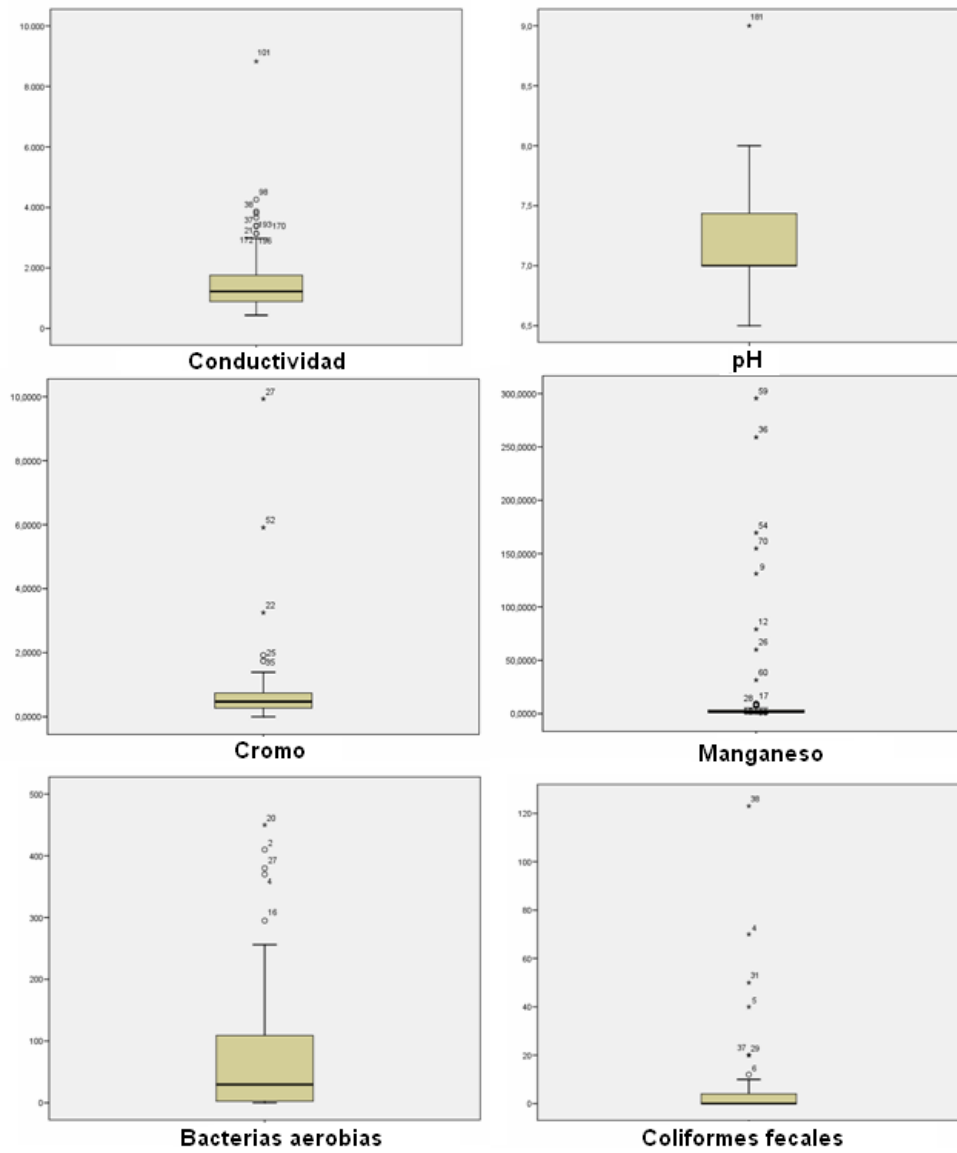


Figura 4.6: Diagramas de caja (box-plot) para algunas de las variables de las aguas subterráneas.

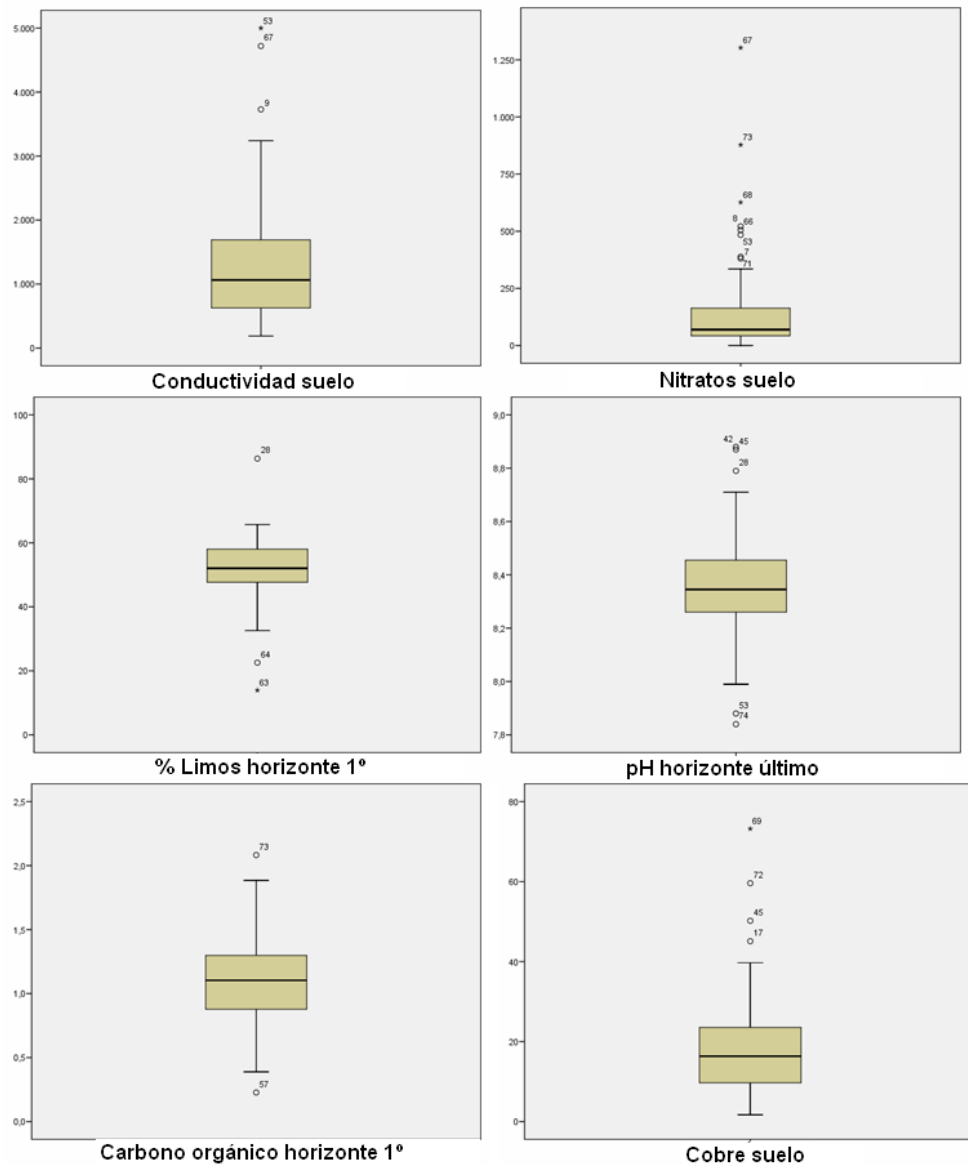


Figura 4.7: Diagramas de caja (box-plot) para las variables del suelo.

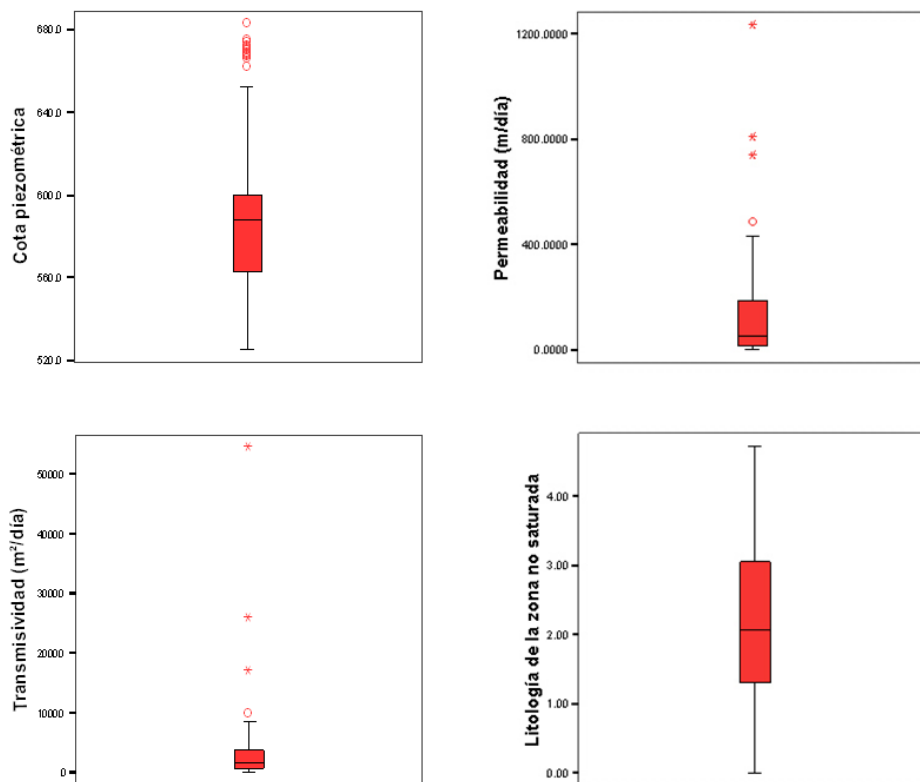


Figura 4.8: Diagramas de caja (box-plot) para algunos de los parámetros del acuífero.

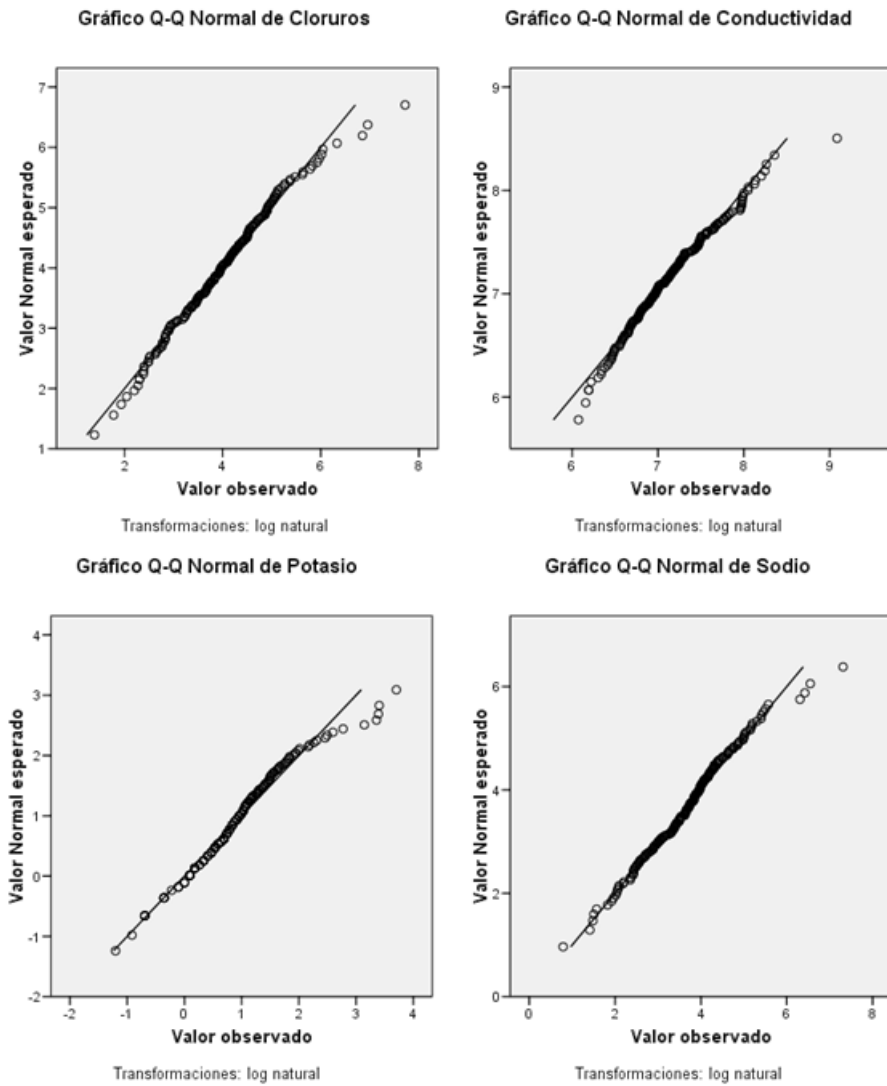


Figura 4.9: Gráficos Q-Q normales de algunos elementos químicos mayoritarios de las aguas subterráneas.

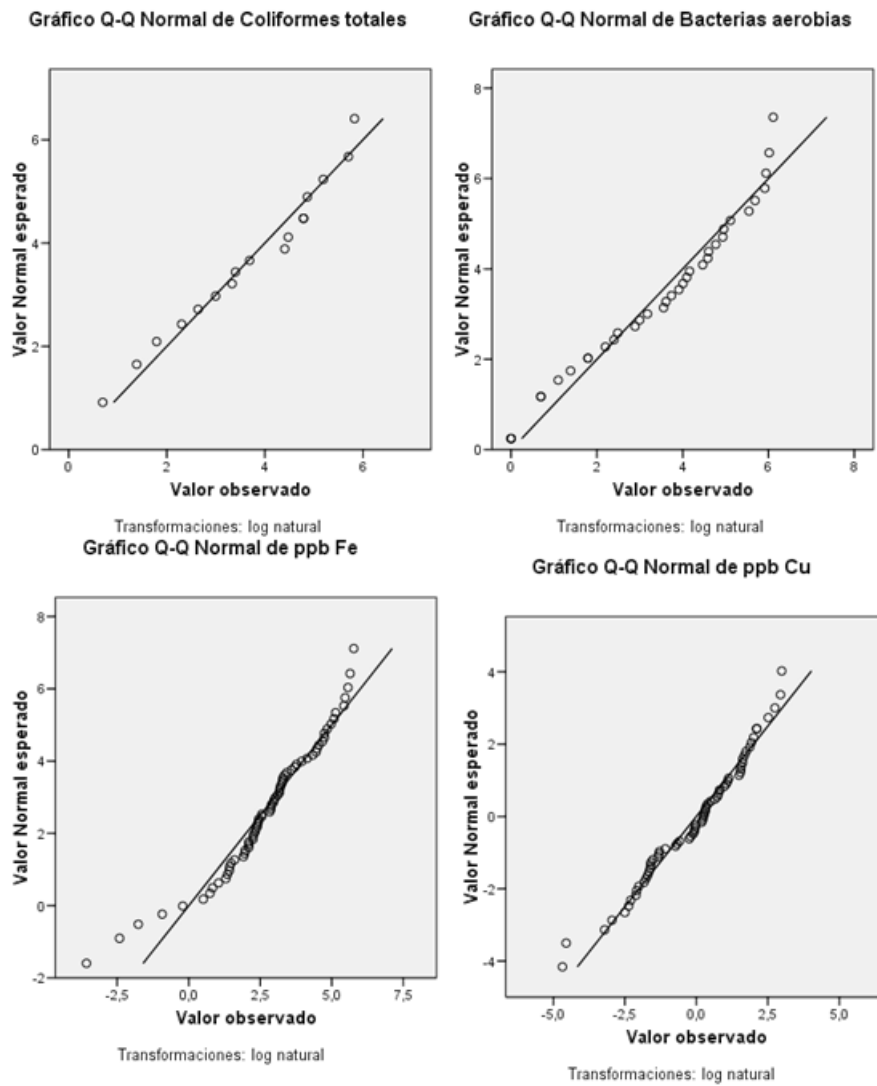


Figura 4.10: Gráficos Q-Q normales para algunos metales pesados y microbiológicos.

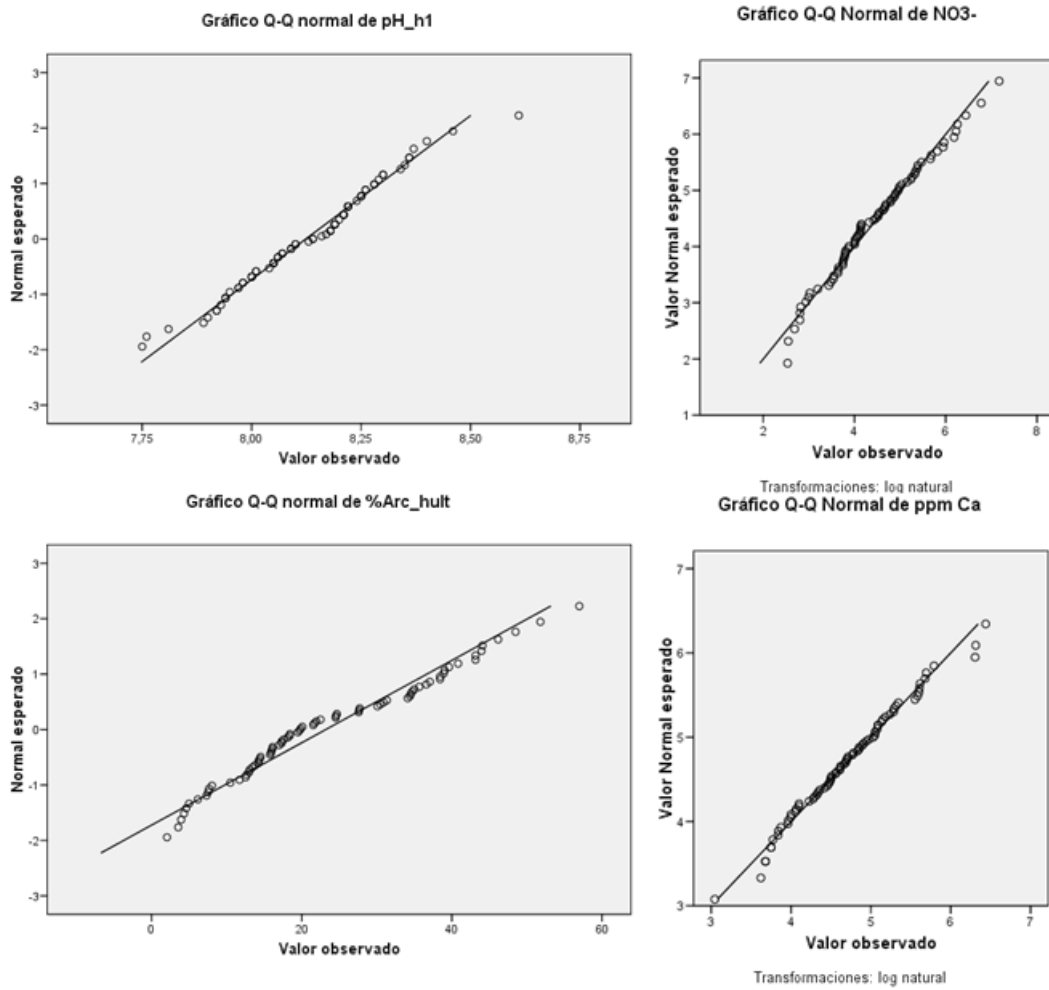


Figura 4.11: Gráficos Q-Q normales para algunos elementos químicos y físicos de los suelos.

Sin embargo su transformación logarítmica ajusta a una distribución normal, que es la transformación elegida para su estimación geoestadística. Los parámetros cota piezométrica y espesor no saturado no requerirán de ningún tratamiento para su espacialización, aunque también se evidencian varias poblaciones en el parámetro de cota piezométrica.

4.3. Análisis estadístico multivariante

El análisis de datos multivariante es una metodología de gran utilidad en estudios ambientales, en los que cada muestra o unidad observacional está caracterizada por un conjunto de variables con valores cuantitativos o cualitativos. Un ejemplo típico son las medidas hidrogeoquímicas o bien edafológicas, donde se dispone de muestras multielemento. Las técnicas estadísticas multivariantes permiten analizar cambios y relaciones entre variables (dependencia). Tienen limitaciones, puesto la relación entre las variables suele ser compleja, por lo que es difícil proporcionar una interpretación física a los resultados. Esto es especialmente cierto los datos edafológicos que son el resultado de diversos fenómenos que interactúan entre sí.

Seguidamente se exponen los resultados correspondientes al estudio estadístico multivariante de los datos hidroquímicos y edafológicos.

4.3.1. Estudio estadístico multivariante de datos hidroquímicos

El análisis de datos multivariantes es una metodología frecuentemente utilizada en estudios hidrogeoquímicos. Steinhort y Williams (1985 [171]) usan el análisis estadístico multivariante de los datos químicos del agua en estudios para identificar los orígenes del agua subterránea. En su aplicación del análisis multivariante y datos químicos, Usunoff y Guzmán-Guzmán (1989 [182]) demuestran la utilidad de aproximaciones multivariantes para el estudio preliminar de estudios hidroquímicos. Melloul y Collin (1992 [120]) usan Análisis de Componentes Principales (ACP) como suplemento a los métodos geoquímicos clásicos como los diagramas de Scholler y Piper, para identificar grupos principales de aguas y factores que afectan a la calidad del agua en el acuífero. Schot y van der Wall (1992 [163]) aplican componentes principales y análisis clúster a datos hidroquímicos para mostrar el impacto regional de las actividades humanas en la composición de las aguas subterráneas. En el estudio de Farnham et al. (2003 [70]), la aplicación del análisis multivariante a los elementos químicos traza de las aguas subterráneas ayuda a identificar los procesos de interacción agua-roca y las condiciones redox. Estos estudios son algunos de los muchos existentes que muestran que el análisis estadístico multivariante ayuda significativamente a la interpretación hidrogeoquímica de las aguas subterráneas. Además, cuando dicha interpretación hidroquímica se combina con el conocimiento geológico e hidrogeológico del área de estudio, los métodos estadísticos

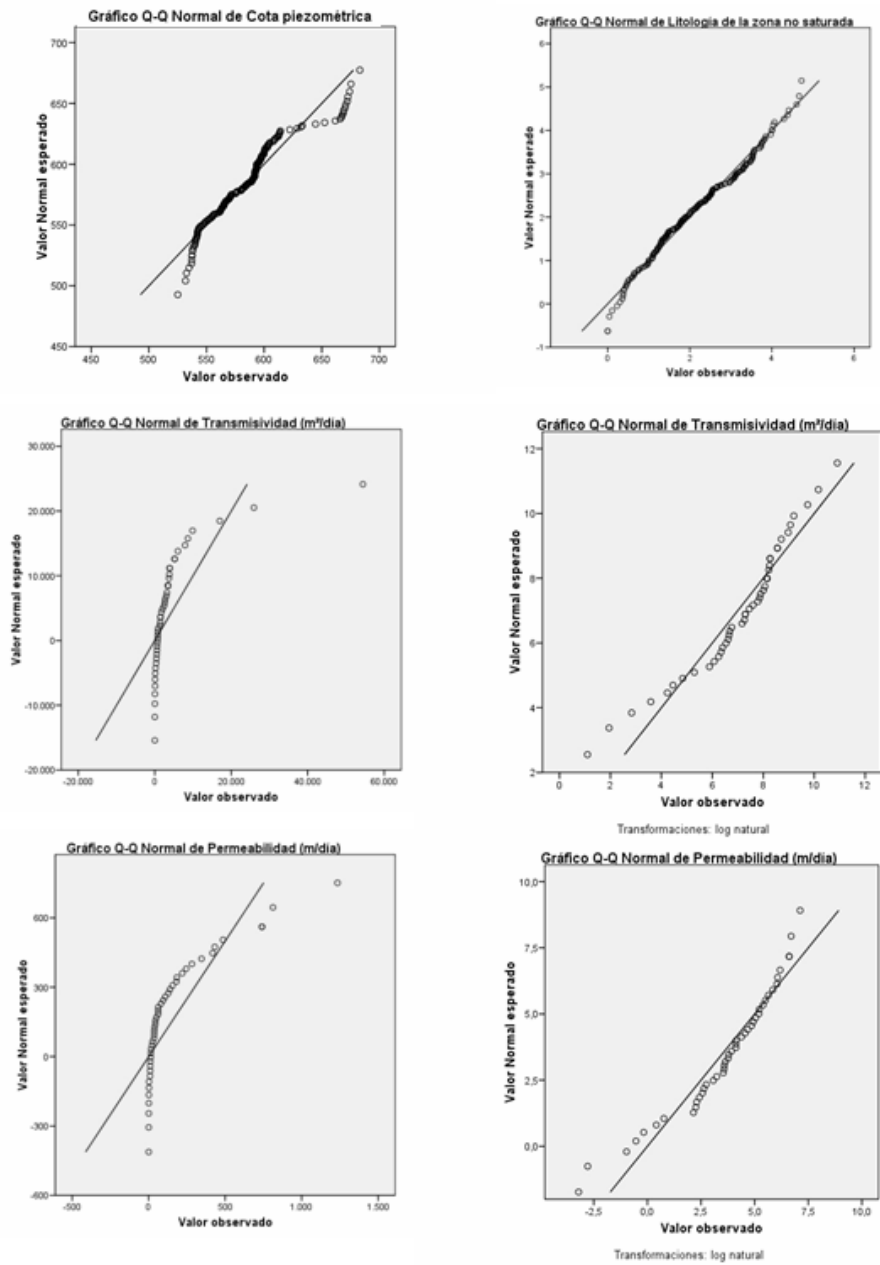


Figura 4.12: Gráficos Q-Q normales para los parámetros relacionados con el acuífero.

multivariantes pueden también ayudar a entender el flujo de las aguas subterráneas en sistemas acuíferos complejos (Farnham et al. 2000; Stetzenbach et al. 2001 [71, 172]).

Un aspecto previo al comienzo del análisis multivariante es la adecuación de las tablas de dato. Concretamente, se ha excluido un pequeño grupo de variables que presentaban más del 80 % de los valores por debajo del límite de detección (LD). En el resto de casos el valor inferior al LD se ha reemplazado por otro. Existen una gran cantidad de métodos para reemplazar los valores por debajo del límite de detección (Alley, 1993; Güler et al. 2002 [8, 78]). En este estudio se ha decidido usar valores aleatorios inferiores al límite. En las variables: cromo (LD: 0.1 ppb), cobre (LD: 0.7 ppb), hierro (LD: 0.4 ppb) y manganeso (LD: 0.2 ppb); 33 datos de un total de 304 (76 por muestra) han sido reemplazados mediante este método. Las variables con más del 80 % de valores por debajo del límite de detección son: nitritos, fosfatos, carbonatos, amonio, plomo (LD: 1 ppb), cadmio (LD: 0.1 ppb), arsénico (LD: 0.1 ppb) y antimonio (LD: 2 ppb).

Se han aplicado dos métodos estadísticos multivariantes para extracción de información: Análisis Clúster Jerárquico (Hierarchical Cluster Analysis, HCA) y Análisis de Componentes Principales (ACP). El HCA es uno de los más aplicados en las Ciencias de la Tierra (Davis 1986 [52]); y a menudo usado en la clasificación de datos hidroquímicos (Steinhorst y Williams, 1985; Schot y van der Wal, 1992; Güler et. al, 2002; Cloutier et. al, 2008 [171, 163, 78, 45]).

El principal resultado del análisis HCA puede expresarse en un dendograma, como el representado en la figura 4.13 correspondiente a las muestras hidroquímicas del agua. Para este estudio se ha escogido la distancia euclídea como medida de distancia, o medida de similaridad entre las muestras de las aguas. Las medidas con mayor similaridad han sido las primeras en agruparse. Después, los grupos de muestras han sido unidos mediante una regla de unión, y los pasos se han repetido hasta que todas las observaciones han sido clasificadas. Para este grupo de muestras, el método de Ward ha sido el más exitoso por los clústeres obtenidos, que son relativamente homogéneos y geoquímicamente interpretables. Los resultados difieren un tanto en los clústeres obtenidos por otros métodos, como la agrupación de medianas, aunque también cabe decir que los resultados obtenidos son muy similares al método de agregación intragrupos. El método de Ward es distinto a otras reglas de unión porque usa un análisis de aproximación de la varianza para evaluar las distancias entre los clústers, y es usado habitualmente en otros estudios por ser el método más discriminante (Schot y van der Wall, 1992; Güler et. al, 2002 [163, 78]), además de ser el que está sujeto a una pérdida de información mínima (Pardo y Ruíz, 2002 [136]).

La clasificación de las muestras dentro de los clústers está basada en una observación visual del dendograma. En este estudio, la línea que divide los distintos grupos está dibujada a una distancia de, aproximadamente, 9. De este modo, las muestras con una

distancia de unión inferior a 9 quedan agrupadas en el mismo clúster. Esta posición de la línea permite una división del dendograma en 3 clústers (C1, C2 y C3). Como ya mencionó Güler et. al (2002 [78]), un mayor o menor número de clústers podría ser definido moviendo la línea arriba o abajo del dendograma. Esta evaluación subjetiva hace que el HCA sea un método semi-objetivo. Para este trabajo, la agrupación en tres clústers distintos arroja los resultados más satisfactorios, como son la formación más clara de distintos clústers geoquímicos y espaciales.

El estudio del dendograma revela, además, algunos otros datos como es el nivel de similaridad entre clústeres y posibles sub-clústeres, así los clústers C2 y C3 presentan una mayor similaridad entre ellos que con el C1. Esto es debido a que estas aguas están influenciadas por los mismos ríos (figura 4.16). El clúster C3 sin embargo, está más influenciado por la composición de las aguas de los ríos Velillos, Cubillas y Salado. Para describir las características de cada clúster de muestras, se han representado los valores medios de los datos geoquímicos y geofísicos en la tabla 4.7, así como el resto de parámetros descartados y algunos otros del acuífero medidos para esos puntos, como son: la litología de la zona no saturada¹, el espesor no saturado, la pendiente, la geología² y la profundidad del techo del sustrato. Un dato interesante es el valor del espesor no saturado que se hace más pequeño desde el C1 al C3, así como el tamaño de grano en la zona no saturada (litología de la zona no saturada) que también disminuye (véase capítulo 2).

El HCA permite asociar un tipo de agua distinta a cada clúster, caracterizada por una composición química diferente, como muestran los diagramas de Stiff. El **C1** tiene una composición de las aguas fundamentalmente bicarbonatada cálcica con valores bajos en la mayoría de las variables, excepto en el potasio y el plomo que tienen valores ligeramente más elevados. La composición del **C2** es cálcica-sulfatada-bicarbonatada, la mayoría de los parámetros empiezan a tomar valores más altos (figuras 4.14 y 4.15) con valores medios superiores al límite establecido por ley de potabilidad en el cromo, nitratos y sulfatos (tabla 4.7). La composición del **C3** es sulfatada-cálcica con valores medios de los parámetros superiores a las medias del clúster 2, también superiores a la concentración establecida por la ley en los mismos parámetros del clúster 2 añadiendo el manganeso.

Como cabría esperar, las muestras de cada clúster presentan una regionalización espacial en el acuífero. Especialmente las muestras de **C1** (figura 4.16) se localizan en la zona de cabecera del acuífero donde se produce la recarga por los principales ríos (Genil y Dílar). Las aguas subterráneas del clúster 1, en comparación con los otros dos, tienen una concentración más baja en aniones y cationes con predominio del bicarbonato y el

¹*Estimación según se detalla en el capítulo 2.

²**Valores de clase

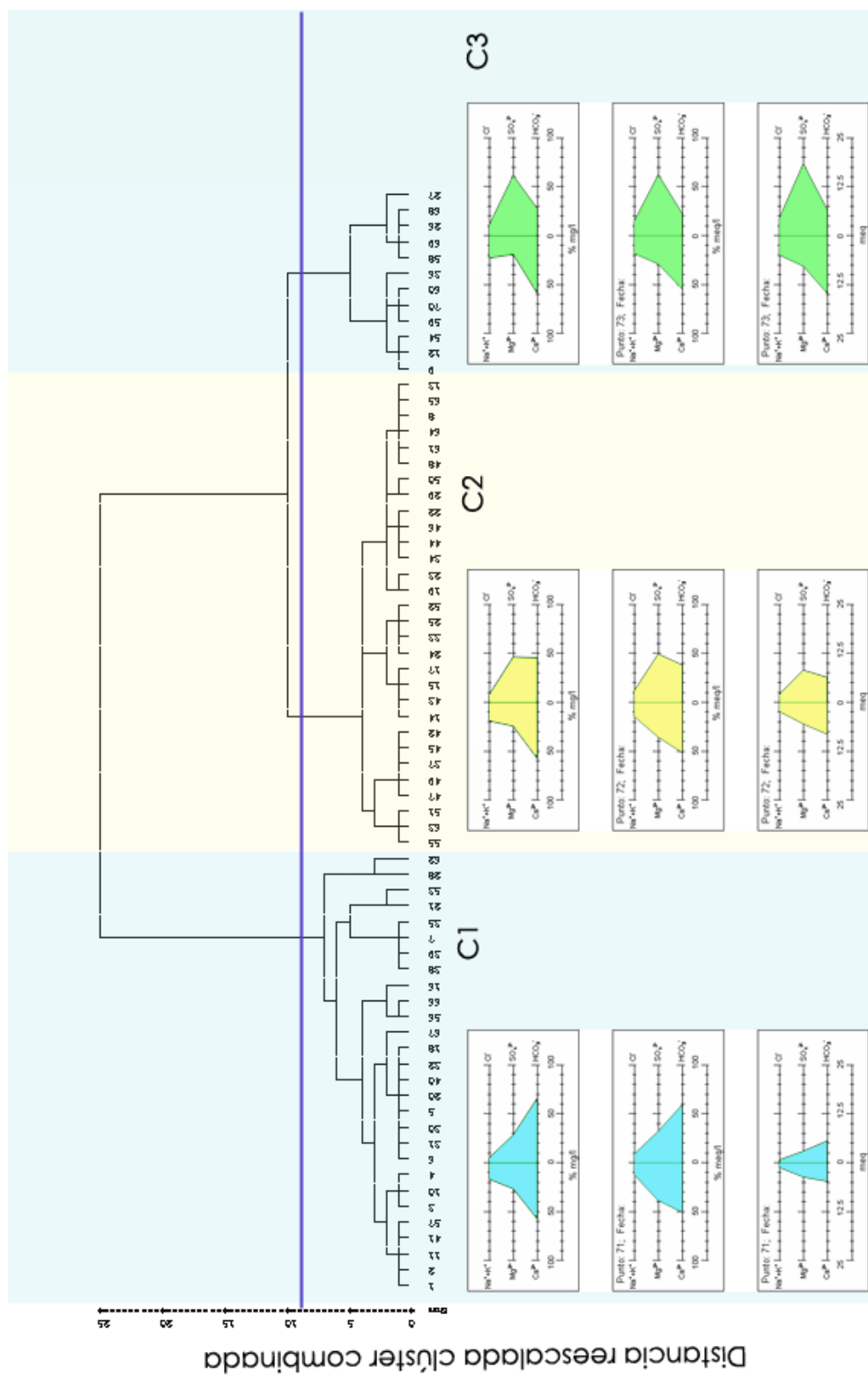


Figura 4.13: Dendrograma resultado del HCA para las muestras de aguas subterráneas y diagramas de Stiff para la concentración media de cada clúster.

Tabla 4.7: Valores medios de variables hidroquímicas y del acuífero correspondientes a los tres clústeres.

Variable	Unidad	Clúster 1	Clúster 2	Clúster 3
Conductividad	µS/cm	865.43	1424.20	2108.67
pH	-	7.47	7.37	7.17
Turbidez	UNF	2.01	1.86	2.39
Cloruros	ppm	29.10	70.50	158.50
Nitratos	ppm	42.05	80.10	58.35
Nitritos	ppm	0.01	0.01	0.01
Sulfatos	ppm	146.48	400.45	880.73
Fosfatos	ppm	0.00	0.02	0.01
Boro	ppm	0.18	0.17	0.36
Bicarbonatos	ppm	341.94	394.07	400.27
Carbonatos	ppm	0.00	0.00	0.00
Calcio	ppm	92.75	159.70	293.68
Sodio	ppm	23.40	48.25	107.27
Magnesio	ppm	42.88	66.96	94.01
Potasio	ppm	4.09	3.47	5.18
Amonio	ppm	0.16	0.28	0.15
Fluoruros	ppm	0.21	0.59	1.11
Plomo	ppb	0.67	0.54	0.91
Cromo	ppb	0.39	0.94	1.31
Cobre	ppb	1.28	4.27	2.41
Hierro	ppb	35.11	50.32	78.65
Manganeso	ppb	1.87	2.23	99.57
Cadmio	ppb	0.04	0.05	0.06
Antimonio	ppb	0.02	0.01	0.03
Arsénico	ppb	0.29	0.46	0.55
<i>Litología zona no saturada</i>	^a	2.772	1.778	1.688
<i>Espesor no saturado</i>	m	57.359	18.589	4.057
<i>Pendiente</i>	%	2.643	1.685	1.276
<i>Geología</i>	^b	23	54	54
<i>Altura techo sustrato</i>	m	477.958	443.212	331.566

^aEstimación según se detalla en el capítulo 2.

^bValores de clase

calcio. La representación de la composición química de las aguas de los ríos de cabecera en un diagrama Piper (figura 4.17) muestra también una composición bicarbonatada cálcica, lo que es coherente con la recarga.

Las muestras de **C2** se localizan en la zona central del acuífero. Estas muestras tienen una concentración más elevada en magnesio, sulfatos y calcio. Estas aguas proceden de la infiltración del río Genil y de los arroyos Juncaril, la Teja, etc. provenientes, estos últimos, de materiales con alternancias de yeso del Mioceno superior. En esta zona también existe infiltración por riego de cultivos. Los análisis de las aguas superficiales del río Genil (figura 4.17) en la zona del C2 tienen una composición sulfatada magnésica, cálcica y sódica, por lo que la infiltración procedente del Genil, en esta zona con mezcla de aguas, tiende a incrementar el contenido en los iones mencionados de las aguas subterráneas.

Las muestras de **C3** se localizan en la zona más occidental y en relación con la recarga de los ríos Cubillas y Velillos al norte y Salado al sur. Los cursos de estos ríos, así como las aguas de recarga laterales, atraviesan a su paso materiales evaporíticos, fundamentalmente yesíferos, lo que le da a estas aguas subterráneas un elevado contenido en sulfatos. Además del contenido evaporítico, el empeoramiento de la calidad general (aumento generalizado en la mayoría de cationes y aniones), se puede ver influenciado por otros factores (tabla 4.7) como el espesor no saturado o la pendiente. La composición de las aguas superficiales del río Salado es sulfatada sódica y clorurada, debido a su paso por materiales evaporíticos. La composición de los ríos Cubillas y Velillos es sulfatada magnésica, cálcica y sódica, con mayor influencia de materiales yesíferos. La infiltración de estos ríos sobre los bordes más occidentales producen por tanto empeoramiento de la calidad inicial correspondiente a las aguas de recarga de Sierra Nevada.

El Análisis de Componentes Principales (ACP) es otra técnica importante de extracción de información de datos multivariantes, pues ayuda a comprender la estructura de interrelación subyacente que se asume existe en el conjunto de datos (Davis, 1986 [52]). Para este análisis se han usado las mismas variables que para el HCA. Del conjunto estandarizado de datos, se han extraído los Componentes Principales de la matriz de correlación simétrica de las 17 variables analizadas.

El número de componentes extraídos se basa en el criterio de Kaiser, según el cual sólo los componentes con autovalores mayores a la unidad se mantienen. Como resultado, todos los componentes que contienen una varianza mayor que las variables originales estandarizadas se guardan. Para este estudio los 6 primeros componentes extraídos tienen una varianza mayor que uno, y suponen un 77,61 % del total de la varianza de la carga de información. Para maximizar la varianza de los principales ejes, se ha aplicado una rotación Varimax. La tabla 4.8 presenta las cargas principales para estos 6 componentes, así como la varianza explicada para cada uno. Las cargas o pesos representan

Gráfico de Collins

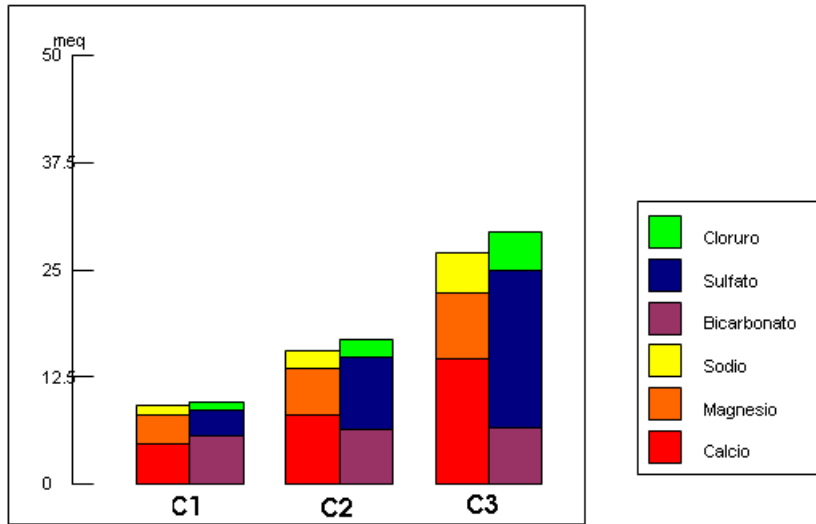


Figura 4.14: Gráfico de Collins para los tres clústeres.

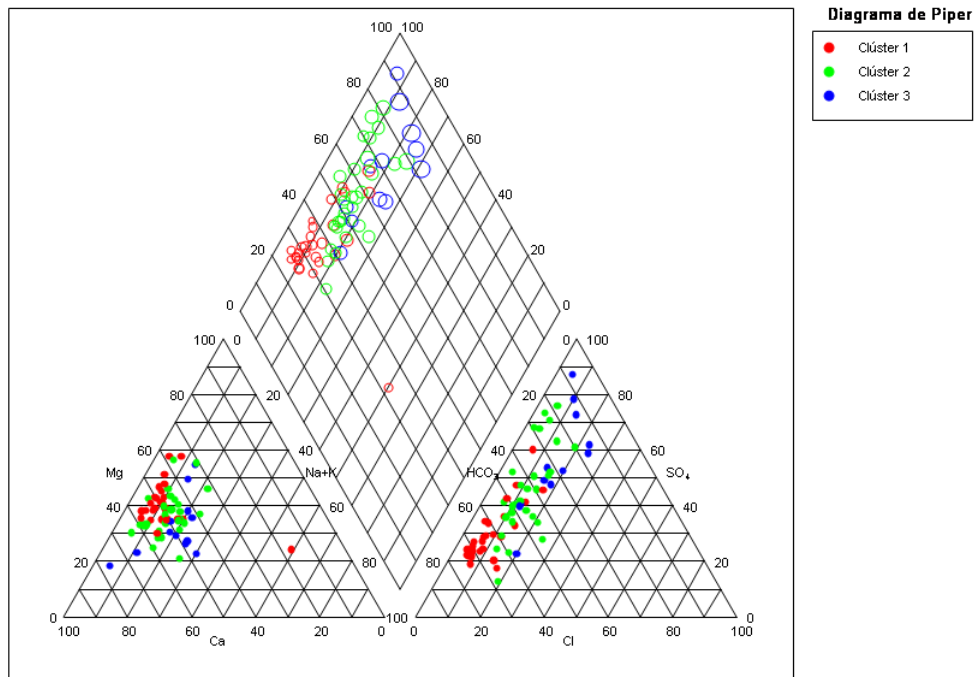


Figura 4.15: Diagrama de Piper para los tres clústeres.

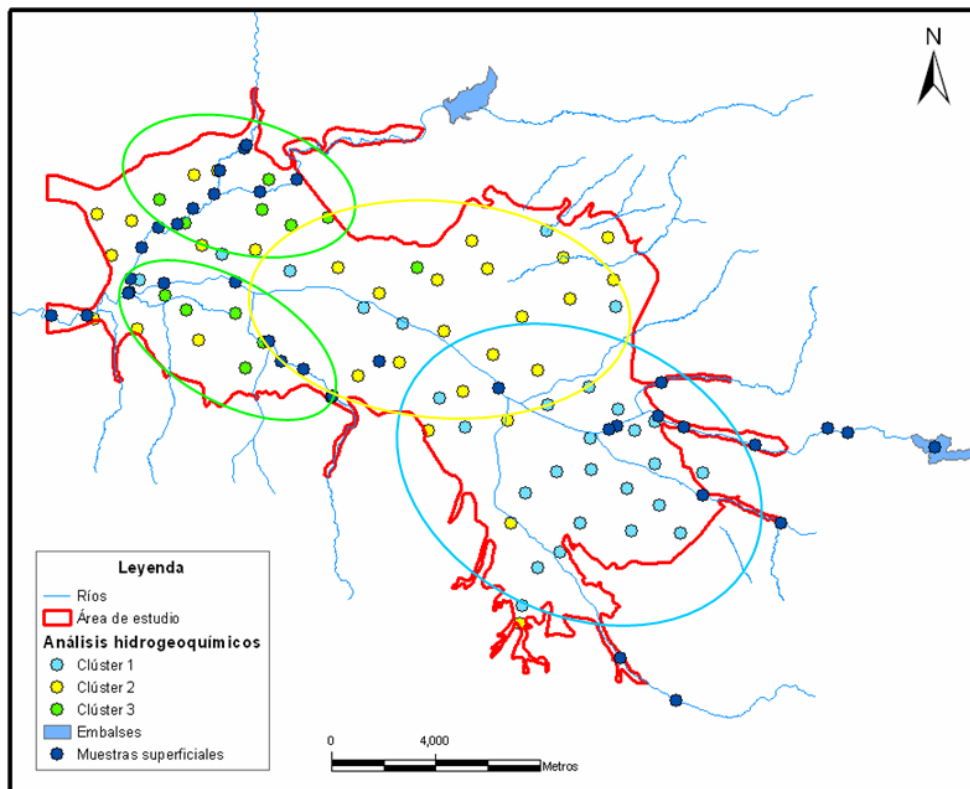


Figura 4.16: Regionalización en el acuífero de los clústeres hidrogeoquímicos obtenidos mediante HCA.

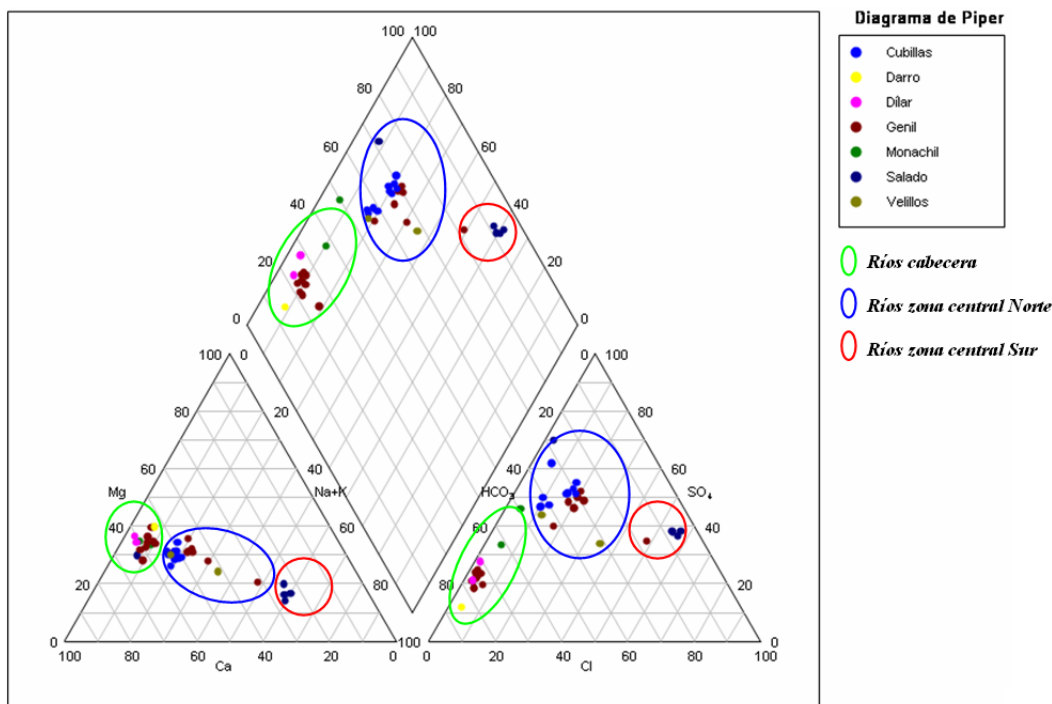


Figura 4.17: Diagrama Piper para las aguas superficiales.

la importancia de las variables para cada componente y están en negrita aquellas cuyos valores son superiores a 0,5. Así, cada componente se caracteriza por unas cargas altas y muchas cargas cercanas a cero, lo que supone el objetivo de la rotación. El primer componente explica el 30 % de la varianza total de los datos, el 2 y el 3 explican cada uno un 10 % de la varianza y los otros 3 explican cada uno porcentajes inferiores al 10 % de la varianza. En resumen, el primer componente es el que agrupa la mayoría de la varianza explicada, seguido de los componentes 2 y 3.

El CP1 está caracterizado por cargas altas y positivas en la conductividad, cloruros, sulfatos, calcio, sodio, flúor y magnesio. Estos iones están ligados a los materiales geológicos del acuífero (carbonatos, yesos, evaporitas, etc.) con la conductividad como variable aditiva. Por lo tanto, el CP1 se puede definir como los iones naturales. El CP2 está caracterizado por el pH, el bicarbonato y los nitratos. Esta asociación es debida probablemente a la relación, ya comentada en el capítulo 2, entre la concentración de nitratos y el pH, con una carga del pH negativa y la de los nitratos positiva, lo que justifica la disminución del pH con la nitrificación de las aguas. La CP3 está caracterizada por el potasio, el boro y el sodio, cationes procedentes de la disolución de rocas evaporitas o de compuestos agrícolas, para el caso del potasio y el boro. La CP4 agrupa los metales hierro y manganeso. El hierro y el manganeso pueden tener orígenes similares ya que aparecen asociados cuando se trata de contaminación industrial, o pueden ser

Matriz de componentes rotados(a)

	Componente					
	1	2	3	4	5	6
Conductividad	.921	.190	.251	.112	.039	.042
pH	-.283	-.819	.045	.010	.142	.044
Turbidez	.061	-.043	-.143	-.033	-.848	.024
Cloruros	.776	.179	.463	.049	.009	.085
Nitratos	.127	.676	-.174	-.193	.360	.204
Sulfatos	.939	.032	.013	.080	.037	.102
Boro	.083	.033	.548	-.020	-.356	.517
Calcio	.899	.128	.048	.001	.135	.158
Sodio	.724	.116	.547	.056	-.060	.082
Potasio	.223	-.171	.743	.037	.296	-.063
Flúor	.668	-.342	-.190	.259	-.137	-.078
Cromo	.361	-.018	-.184	.101	.418	.543
Cobre	-.002	.000	.020	.134	-.017	.787
Hierro	.110	-.186	-.052	.776	.365	.187
Manganeso	.315	.114	.153	.803	-.212	.097
Magnesio	.795	.241	.006	.247	-.101	-.069
Bicarbonatos	.011	.578	.474	.295	-.036	-.227
Varianza explicada	5.14	1.82	1.80	1.55	1.49	1.39
% de la varianza	30.23	10.68	10.61	9.14	8.77	8.16
% acumulado	30.23	40.92	51.53	60.67	69.45	77.61

Método de extracción: Análisis de componentes principales.

Método de rotación: Normalización Varimax con Kaiser.

a. La rotación ha convergido en 29 iteraciones.

Tabla 4.8: Cargas de los componentes principales y varianza para los 5 componentes con la rotación Quartimax.

consecuencia de una disolución natural del terreno cuando se produce un vertido orgánico debido al carácter reductor de la disolución (Luque Montilla, 1998 [114]). Ambos elementos provocan el desarrollo de bacterias que producen la formación de depósitos insolubles de sus sales, ya que ambos metales tienen comportamientos geoquímicos similares origen de su agrupación. La componente 5 recoge únicamente la turbidez, que al ser un parámetro muy influido por condiciones locales del pozo, no está relacionado con el resto de variables geoquímicas. La CP6 está caracterizada por cargas altas en el cobre, el cromo y el boro. La asociación de estos tres elementos puede estar influida por la contaminación industrial de actividades relacionadas con el acero o el cobre, pero al ser este el último factor la agrupación de variables es una componente más errática pues representa una varianza menor.

Con el fin de favorecer la interpretación espacial de la información anterior, se han elaborado mediante krigeaje los mapas estimados de la distribución espacial de cada componente principal en el acuífero. Los modelos teóricos de variograma ajustados (tabla 4.9 y figura 4.18) son esféricos, con alcances que oscilan entre 3500 m para el CP6, que es el que parece estar influenciado por la contaminación industrial y 11000 m para el CP3 que está influenciado por geología regional. El efecto de pepita o grado de aleatoriedad oscila entre el valor mínimo de 0.05 (5 %) para el CP1, lo que es coherente puesto que este componente representa los iones naturales y por tanto su espacialización es muy buena, y un máximo de 0.65 (65 %) para el CP4, componente que agrupa los metales con una distribución más local y aleatoria.

Los mapas de distribución espacial de los factores (figura 4.19) resumen la información aportada por las variables. El CP1 hemos comentado es una variable resumen de la química del acuífero, con las cargas altas en positivo. Por lo tanto los valores altos del CP1 significan alta concentración de iones y se localizan en la zona de descarga y en el borde noreste, coincidiendo con las zonas de rocas evaporíticas. El CP2 tiene las cargas de nitratos y bicarbonatos en positivo y el pH en negativo, por lo que valores altos en esta componente significan alta concentración de nitratos y bicarbonato y un pH bajo. Estos valores altos se localizan fundamentalmente al suroeste de Sierra Elvira, cercanos a la zona de descarga. En esa zona el espesor no saturado es bajo y al estar cerca de la zona de descarga los nitratos acumulados a lo largo del acuífero son también elevados. El CP3 tiene cargas positivas en el boro, el sodio y el potasio y los valores máximos se localizan en la zona central del río Genil. Como su origen está ligado tanto a abonos como a la disolución de las rocas evaporíticas de la zona norte su distribución toma este aspecto. El CP4 agrupa los metales hierro y manganeso que suelen ir asociados en el caso de la polución industrial y los máximos se localizan en la zona central del borde norte. El CP5 tiene carga elevada negativa en la turbidez, por lo que los máximos de turbidez corresponden a los mínimos del componente, en el borde norte y en en la zona

	Paso	Nº Pasos	Modelo	Pepita	Meseta	Alcance
Factor 1	1100	10	Esférico	0.05	0.93	9700
Factor 2	1000	10	Esférico	0.4	0.6	8000
Factor 3	1600	10	Esférico	0.3	0.7	11000
Factor 4	1900	10	Esférico	0.65	0.35	4000
Factor 5	1500	10	Esférico	0.6	0.4	6000
Factor 6	1000	10	Esférico	0.1	0.9	3500

Tabla 4.9: Tabla resumen de los modelos teóricos ajustados a los factores del ACP.

del Salado. El CP6 tiene cargas positivas en el cobre, el cromo y el boro y los máximos se localizan en el borde oeste del acuífero y en algunos puntos cercanos a Granada capital y Armilla. Estos máximos además de por contaminación industrial pueden estar relacionados con su uso como bactericida en piscinas y posterior riego de jardines con estas aguas.

Finalmente, el estudio de los factores de las variables y su espacialización aporta un nuevo punto de vista que permite identificar el posible origen de la contaminación, así como su relación con otras variables. Además, permite sintetizar la información aportada por el conjunto de 17 variables en seis factores que representan casi el 80 % de la variabilidad total inicial. El estudio de la matriz de componentes y el análisis geoestadístico de cada uno de los factores, permite además profundizar en el conocimiento de cada una de las variables y de sus relaciones.

4.3.2. Estudio multivariante de datos edafológicos

En relación con los datos experimentales edafológicos también se ha aplicado un HCA y un Análisis Factorial/ACP. Las variables seleccionadas han sido las medidas en todos los horizontes del suelo (tabla 4.10). El clúster se ha realizado por el método de Ward que usa un análisis de aproximación la varianza para evaluar las distancias y es el que mejores resultados ofrece. La medida de distancia escogida es la distancia euclídea al cuadrado y la línea de división de los distintos grupos se ha situado aproximadamente en 7 (figura 4.20). Para este valor de distancia se han obtenido tres grupos clústers que representan espacialmente la evolución de los suelos (figura 4.21), comentada en el primer capítulo, diferenciando entre suelos más evolucionados (C2) y menos evolucionados (C1). El C3 representa suelos menos comunes, que aparecen aislados en el espacio y se corresponderían a zonas con textura más arenosa en promedio (tabla 4.11).

El Análisis Factorial (AF) se ha realizado con el total de variables químicas y físicas analizadas en los suelos. El ACPI, comúnmente usado en esta disciplina (Golobocanin et al., 2004; Lucho-Constantino, et al., 2005; Sena et al., 2002; y Carroll y Oliver, 2005 [84, 109, 166, 28]), permite reducir la dimensionalidad del problema y crear un modelo

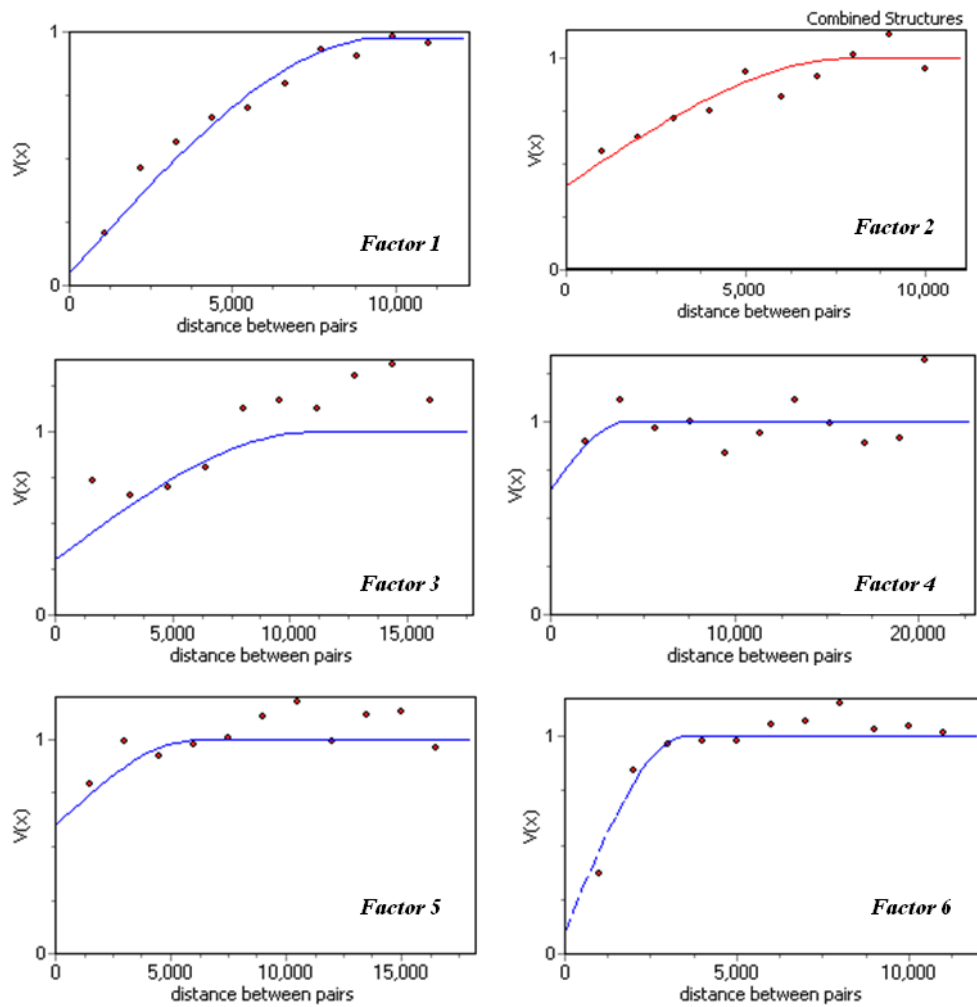


Figura 4.18: Variogramas y modelos teóricos de los factores del ACP.

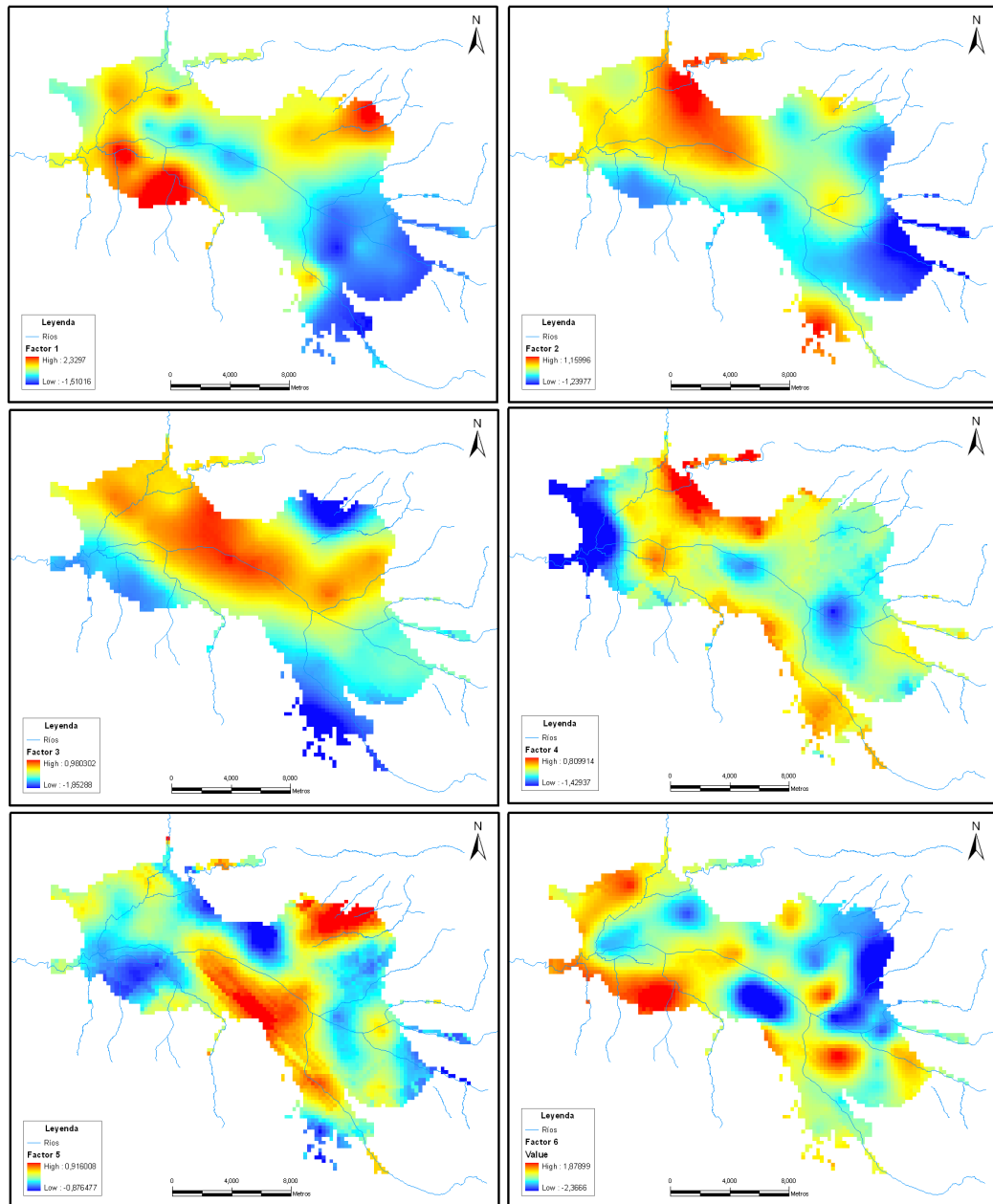


Figura 4.19: Mapas estimados de los componentes del ACP.

Nombre variable	Variable	horizonte analizado
COrg_h1	% Carbono Orgánico	horizonte 0-20 cm
pH_h1	pH	horizonte 0-20 cm
%H_h1	% Humedad	horizonte 0-20 cm
%Aren_h1	% Arenas	horizonte 0-20 cm
%Lim_h1	% Limos	horizonte 0-20 cm
%Arc_h1	% Arcillas	horizonte 0-20 cm
COrg_hult	% Carbono Orgánico	horizonte 80-100 cm
pH_hult	pH	horizonte 80-100 cm
%H_hult	% Humedad	horizonte 80-100 cm
%Aren_hult	% Arenas	horizonte 80-100 cm
%Lim_hult	% Limos	horizonte 80-100 cm
%Arc_hult	% Arcillas	horizonte 80-100 cm
COrg_medio	% Carbono Orgánico	horizonte 0-1m
pH_medio	pH	horizonte 0-1m
%H_medio	% Humedad	horizonte 0-1m
%Aren_medio	% Arenas	horizonte 0-1m
%Lim_medio	% Limos	horizonte 0-1m
%Arc_medio	% Arcillas	horizonte 0-1m

Tabla 4.10: Variables usadas en el análisis clúster de los parámetros físicos del suelo.

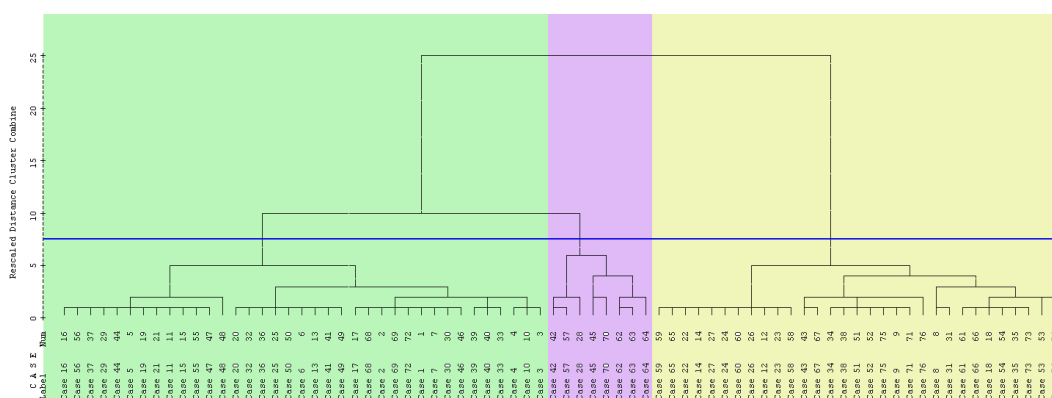


Figura 4.20: Dendrograma de los análisis del suelo.

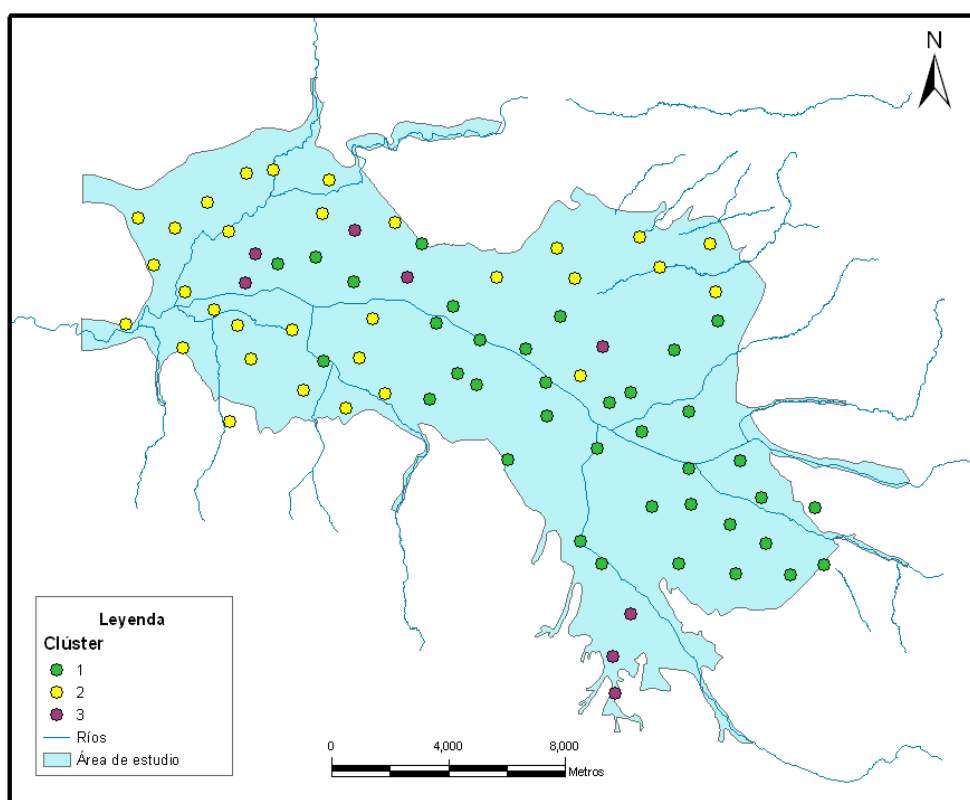


Figura 4.21: Situación espacial de los clústeres de los análisis del suelo.

<i>Variable</i>	<i>Horizonte</i>	<i>C1</i>	<i>C2</i>	<i>C3</i>
Carbono orgánico	0-20 cm	1.188	1.057	0.824
pH	0-20 cm	8.048	8.168	8.296
% Humedad	0-20 cm	1.387	2.457	1.139
% Arenas	0-20 cm	31.671	15.147	30.113
% Limos	0-20 cm	53.751	53.015	42.413
% Arcillas	0-20 cm	14.579	31.838	27.476
Carbono orgánico	80-100 cm	0.509	0.507	0.270
pH	80-100 cm	8.366	8.271	8.629
% Humedad	80-100 cm	1.208	2.571	0.639
% Arenas	80-100 cm	33.295	13.275	54.660
% Limos	80-100 cm	52.517	51.000	28.361
% Arcillas	80-100 cm	14.187	35.724	16.979
Carbono Orgánico	0-1 m	0.763	0.752	0.510
pH	0-1 m	8.248	8.227	8.222
% Humedad	0-1 m	1.275	2.561	0.940
% Arenas	0-1 m	32.498	14.415	41.994
% Limos	0-1 m	52.891	50.323	36.685
% Arcillas	0-1 m	14.611	35.262	21.321
Humedad de saturación	0-1 m	395.716	175.198	512.724

Tabla 4.11: Valores promedios de las variables del suelo para cada uno de los clústeres.

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	7.199	35.995	35.995	7.199	35.995	35.995	6.711	33.554	33.554
2	3.689	18.444	54.439	3.689	18.444	54.439	3.985	19.923	53.477
3	1.884	9.422	63.862	1.884	9.422	63.862	1.811	9.054	62.531
4	1.431	7.155	71.017	1.431	7.155	71.017	1.560	7.799	70.330
5	1.140	5.698	76.715	1.140	5.698	76.715	1.277	6.385	76.715
6	.970	4.850	81.565						
7	.841	4.204	85.769						
8	.587	2.935	88.704						
9	.541	2.704	91.408						
10	.440	2.202	93.610						
11	.373	1.867	95.477						
12	.299	1.497	96.974						
13	.206	1.030	98.003						
14	.148	.742	98.745						
15	.123	.613	99.359						
16	.081	.406	99.765						
17	.031	.156	99.921						
18	.012	.058	99.979						
19	.004	.021	100.000						
20	- 7.929E- 17	-3.965E- 16	100.000						

Tabla 4.12: Resultados del ACP: varianza total y porcentaje que explica cada componente.

de relación de las variables a través de la interpretación de factores. Sus aplicaciones son diversas, ya sea la interpretación del origen de la contaminación en los suelos, evaluación los distintos métodos de gestión y su influencia en los suelos o determinar relaciones entre distintas propiedades físicas o químicas.

Como concepto previo, se recordará que la técnica multivariante de ACP parte del cálculo de la matriz de covarianzas o de correlaciones (elegida en este estudio para estandarizar los datos) que expresa la variabilidad conjunta de las variables (parámetros medidos), y de la extracción posterior de los valores (Tabla 4.12) y vectores propios. Un valor propio (autovalor) representa un conjunto de coeficientes (cargas o pesos) que multiplica a las variables originales correlacionadas para obtener una nueva variable no correlacionada (ortogonal), llamada Componente Principal (CP), la cual es una combinación lineal ponderada de las variables originales.

Los Componentes Principales (CP) son variables sintéticas que se obtienen de proyectar los datos en unos nuevos ejes, es decir, es el producto del dato original por el vector propio. En esta transformación de coordenadas, se escoge como primera componente la que tiene mayor varianza, la segunda varianza mayor es la segunda componente, y así sucesivamente (Adams, 1998; Otto, 1998; y Vega, 1998 [6, 133, 184]).

En ACP existen tantas componentes como variables originales, sin embargo los CP suministran información de los parámetros más representativos que describen el total de nuestros datos, de modo que ofrece una reducción de datos con la mínima pérdida de información. Para este estudio se han elegido las 5 primeras componentes que representan un 76.7% de la varianza total explicada, que corresponden a las componentes con autovalores superiores a 1 (Tabla 4.12).

En la práctica el Análisis Factorial suele aplicarse con posterioridad al ACP. El principal objetivo del AF es reducir el aporte de variables menos significativas con el fin de simplificar aún más la estructura proveniente del ACP. Este propósito se logra rotando los ejes definidos por el ACP, de acuerdo con las reglas establecidas y construyendo nuevos grupos de variables, llamadas varifactores (VF). Se debe decir que un CP es una combinación lineal de las variables experimentales del suelo, mientras que un varifactor puede incluir variables no observables, hipotéticas y latentes (Otto, 1998; Vega, 1998; y Helena, 2000 [133, 184, 92]).

Para el cálculo del AF/ACP, se ha realizado una rotación varimax de los CP (tabla 4.8) que vienen de las variables estandarizadas, con el objetivo de reducir la contribución de las variables con menor significado, simplificando así la estructura de los datos. Según esta tabla, el primer factor corresponde esencialmente a las variables químicas del suelo, el segundo representa la variabilidad textural, el tercero la influencia de la materia orgánica y formación de quelatos, en el cuarto se agrupan los abonos (excepto los nitratos y nitritos, que se agrupan en la primera componente), y en el último, la relación de los metales pesados con el pH.

Para facilitar la interpretación espacial de las características edafológicas del acuífero se han elaborado mediante estimación geoestadística los mapas estimados de cada factor (figura 4.22). También se han representado los mapas estimados de las principales variables relacionadas con los factores a efectos de comparación. La estimación espacial de cada uno de los factores, así como de las variables individuales se representa en las tablas 4.14 y 4.16 ya que la interpretación de los mismos está también muy relacionada espacialmente.

El *factor 1*, explica un 33.6 % de la varianza total y recoge principalmente las variables: conductividad, magnesio, sodio, potasio, cloro, nitratos, nitritos y sulfatos (tabla 4.13, valores en negrita), todas con signo positivo en el factor. Estas variables representan la composición de elementos químicos mayoritarios en el suelo.

Los modelos de variograma correspondientes a las variables que participan en este factor (tablas 4.14 y 4.16; y figura 4.22), tienen alcances un tanto parecidos, comprendidos entre 3.0 y 8 km. En algunos casos, aparece una ligera deriva en las variables que se refleja en el modelo variográfico del factor 1, cuyo modelo consta de dos modelos

Variable	Componente				
	1	2	3	4	5
Arenas (%)	.012	-957	.131	-.073	-.005
Limos (%)	-.405	.616	-.027	.466	.101
Arcillas (%)	.422	.635	-.149	-.404	-.102
pH	-.358	-.650	-.082	-.210	.249
Humedad (%)	.058	.757	-.020	-.212	-.152
C. Orgánico_suelo (%)	.280	.501	.580	.294	-.041
H. Saturac_suelo (%)	.057	.900	-.088	.087	.146
Conductividad (μS/cm a 25°)	.966	.106	-.011	.161	.007
Magnesio (ppm)	.864	.017	-.062	.430	-.069
Calcio (ppm)	.933	.151	-.024	.060	-.092
Sodio (ppm)	.767	.023	.080	-.157	.188
Potasio (ppm)	.532	.042	-.028	.692	-.086
Cloro (ppm)	.883	-.005	-.015	.013	.186
Nitritos (ppm)	.771	.025	.190	-.036	-.119
Nitratos (ppm)	.775	.259	.027	-.246	-.170
Sulfatos (ppm)	.866	.039	-.079	.381	.054
Cobre (ppb)	-.076	-.302	.810	.072	-.009
Manganeso (ppb)	-.012	-.149	-.032	-.120	.695
Cinc (ppb)	.052	.091	.249	.128	.686
Cromo (ppb)	-.036	.035	-.804	.173	-.241

Tabla 4.13: Resultados del ACP: matriz de componentes rotados. (*En negrita los vectores propios mayores para cada componente*).

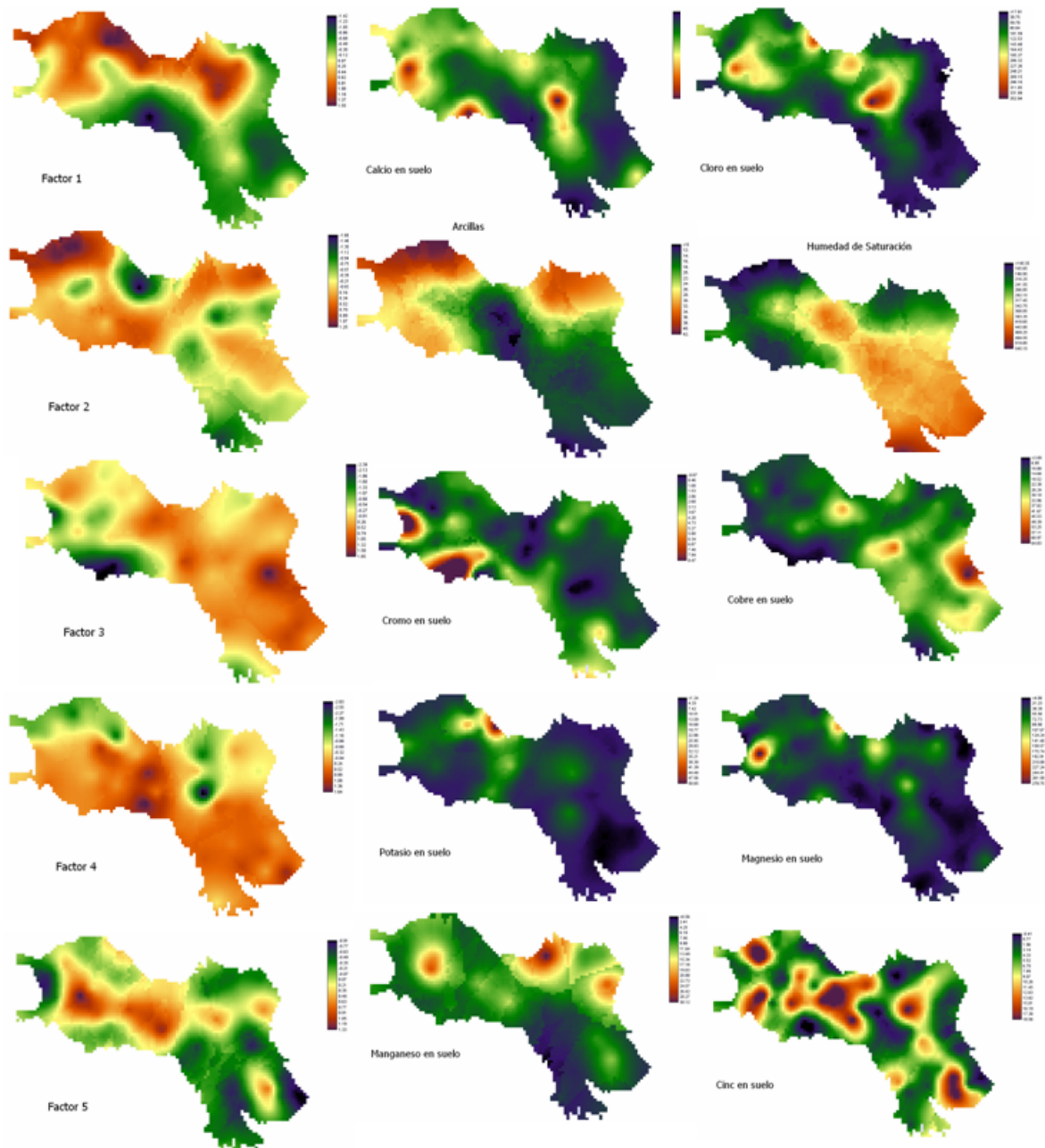


Figura 4.22: Mapas de estimación geostatística correspondientes a los componentes principales y variables más significativas para cada uno.

Variable	Modelo	Pepita	Alcance	Meseta	% Aleatoriedad
Factor 1	Esférico	0.07	3500	0.28	9.33
			11000	0.4	
Factor 2	Esférico	0.45	4000	0.651	40.87
Factor 3	Esférico	0.1	3000	0.2	10.00
			9000	0.7	
Factor 4	Esférico	0.05	2000	0.125	4.44
			8000	0.95	
Factor 5	Esférico	0.5	3500	0.6	45.45

Tabla 4.14: Ajustes de los variogramas de los factores.

esféricos con alcances de 3.5 km (comportamiento local) y de 11 km (comportamiento regional) (tabla 4.16).

En la distribución espacial (estimación) del factor la zona de máximos se localiza al norte y oeste del área de estudio, y la zona de mínimos a la entrada de los ríos principales en la Vega (sector SE). Esta distribución es coherente con la distribución de las variables que participan en este factor, donde los valores máximos se localizan fundamentalmente en la zona de descarga y al sur de Sierra Elvira. El arrastre de material procedente de Sierra Elvira (detritus) con elevados contenidos en rocas evaporitas pone en la disolución del suelo cationes y aniones como el calcio, magnesio, potasio, cloro y sulfatos que participan en este factor y son elevados en esta zona. Además del factor intrínseco debido a la composición del suelo, la mayor capacidad de intercambio iónico de las arcillas (variable alta en este factor) favorece una mayor adsorción de cationes y aniones en esta zona más arcillosa.

Las variables más representativas del *factor 2* (tabla 4.13) corresponden, principalmente, a las variables físicas del suelo: arenas, arcillas, limos, pH, humedad, humedad de saturación y carbono orgánico. Las arenas y pH tienen signo negativo y el resto positivo. Este comportamiento se puede explicar ya que la acidez o la basicidad de un suelo vienen determinadas por los cationes absorbidos por los minerales de la arcilla y la materia orgánica (Porta 2003 [29]). Por lo tanto, en suelos más arenosos, el pH predominante está controlado por el material detrítico de relleno de la cuenca y, al tratarse de un relleno detrítico carbonatado, el pH es más básico a mayores cantidades de arenas y más ácido en las zonas más arcillosas.

Los modelos de variogramas correspondientes a las variables del *factor 2* (tabla 4.14), muestran que hay tres tipos de variables según el grado de correlación espacial.

Las variables de corto alcance que corresponden a Carbono Orgánico, arenas y humedad del suelo, con un alcance de 5.5 km; el pH, que se sitúa con un alcance intermedio de 6.4 km; y los limos, arcillas y humedad de saturación que tienen un alcance superior a los 7.5 km. Por otra parte, el variograma del *factor 2* tiene un corto alcance 4 km y un alto porcentaje de aleatoriedad en el origen observado en la mayoría de las variables que lo componen (tabla 4.16 y figura 4.22).

El mapa estimado del *factor 2* presenta los máximos al oeste y noreste de la zona de estudio y los mínimos a la entrada de los ríos Dilar, Genil, Monachil y Beiro y al sur de Sierra Elvira. Los máximos de este factor corresponden a máximos en el porcentaje de arcillas, donde la energía del agua del río es menor y deposita las arcillas que lleva en suspensión. Los mínimos corresponden a los depósitos con mayores proporciones de arenas en la zona de entrada de los ríos, donde el flujo del río es más enérgico y deja en suspensión los materiales más finos y deposita las arenas.

El *factor 3* corresponde al grupo de metales cobre, cromo y la materia orgánica (tabla 4.13); el cromo tiene signo negativo. La asociación de estas tres variables es debida a la relación del cobre con los quelatos, que son capaces de inmovilizar cationes metálicos como el cobre, y la correlación negativa a la mayor movilidad del cromo en presencia de materia orgánica, que lava este metal.

La correlación espacial de las variables que componen este factor (tabla 4.16) muestran un alcance similar, cercano a los 5 km (cobre, cromo y materia orgánica). El efecto de pepita para estas variables es bajo, lo que indica que existe continuidad espacial para distancias pequeñas. El variograma del factor 3 presenta dos estructuras, la primera con un alcance de 3 km y la otra con alcance de 9 km (tabla 4.14).

El mapa estimado del factor 3 presenta zonas de máximos (máximos de cobre y carbono orgánico y mínimos de cromo) a la entrada del acuífero y de mínimos al suroeste (máximos de cromo). Los máximos de cobre están influidos, probablemente, por el uso de fertilizantes o fungicidas en zonas ajardinadas. En el borde suroeste del área de estudio se localizan los máximos de cromo, pueden relacionarse con las mineralizaciones de estroncio de Montevives.

El *factor 4* recoge el grupo de elementos formado por magnesio, potasio, sulfatos, limos y arcillas (con signo negativo) (tabla 4.13). Este factor representa la carga de abono de los suelos, y su correlación con los limos es evidente, ya que los suelos limosos son suelos con pocos nutrientes (Porta 2003 [29]), al contrario que los suelos arcillosos.

Los variogramas ajustados a las variables (tabla 4.16) muestran que: el potasio, el magnesio y los sulfatos tienen un alcance corto (continuidad espacial moderada), algo mayor en el potasio y los limos y las arcillas tienen un alcance largo (gran continuidad espacial). Esta estructura está presente también en el factor 4 (tabla 4.14), donde hay una estructura de menor alcance (2 km) que corresponde a los elementos usados como

abonos (potasio, sulfatos y magnesio) y una estructura de largo alcance (8 km) en relación con la textura (limos y arcillas).

En los mapas de estimación del factor 4, los máximos se sitúan fundamentalmente al sur del área de estudio y los mínimos al noroeste y noreste, coincidiendo con las zonas más limosas (y menos fértiles).

El *factor 5* (tabla 4.13) representa el comportamiento de los metales frente al pH, y corresponde a las variables manganeso y cinc. El cinc, aumenta su solubilidad con la acidez, al igual que le ocurre al manganeso.

En el análisis de los variogramas y los ajustes (tabla 4.16) encontramos que tanto las variables como el factor tienen un alcance similar (3-4 km) (tabla 4.14). En la estimación observamos que los máximos están en la zona central del oeste del acuífero y los mínimos en los bordes este y oeste.

Según estos resultados, podemos caracterizar los suelos del acuífero de la Vega de Granada en cinco factores: la capacidad de intercambio catiónico, la textura, la materia orgánica, el abonado y el pH. Cada uno de estos factores, así como las variables edáficas asociadas, presentan características de variabilidad propias. También es importante resaltar que la combinación de técnicas multivariantes y geoestadísticas constituye una herramienta de especial interés para el estudio y la caracterización de los suelos, que facilita la comprensión de las complejas relaciones edafológicas y su distribución espacial.

4.4. Análisis espacial de los datos

El análisis temático por medio de métodos estadísticos uni y multivariantes se ha mostrado en los apartados anteriores. En los mismos también se ha presentado la espacialización de algunos de los resultados estadísticos (p.e. componentes principales y factores). En este apartado se presenta una síntesis de los resultados del análisis espacial de las variables experimentales de calidad del agua, suelo y acuífero. Con el objeto de no hacer de este apartado excesivamente extenso debido al elevado número de variables estudiadas, pero manteniendo el rigor científico, se mostrarán algunos ejemplos de la estimación espacial utilizada en este trabajo. Se ha tratado de abarcar la mayoría de las técnicas utilizadas y algunos de los resultados más importantes en lo referente a la construcción de mapas temáticos.

Uno de los principales objetivos del análisis espacial de datos es la creación de mapas temáticos, a partir de muestras experimentales; es decir, se pretende pasar la información puntual experimental a continua. Esta espacialización se puede realizar por distintos métodos, que están mayormente disponibles en los softwares comerciales, aunque principalmente nos centraremos en la estimación espacial por los métodos geoestadísticos. Ya se comentó en las bases metodológicas que el tipo de variable analizada

puede modificar el tipo de estudio debido a que es necesario un conocimiento profundo de la variable para realizar una correcta estimación geoestadística.

En síntesis se ha trabajado con tres tipos de variables:

- *Variables tipo*: Parámetros químicos, físicos y microbiológicos de las muestras de agua y suelo, así como de parámetros climáticos o hidrogeológicos.
- *Variables síntesis*: Engloban tanto las construidas a partir de los métodos estadísticos descritos en el apartado de metodología, como los índices creados a partir de diferentes directrices, normativas o criterio experto.
- *Variables indicatrices*: Es un tipo particular de variable, que consiste en una transformación de la variable inicial a una nueva variable binaria "cero" y "uno", previa aplicación de una serie de valores de corte "z".

Como ya se comentó anteriormente, el proceso de estimación geoestadística por krigeaje requiere de varias fases: en primer lugar el cálculo de la función variograma o de covarianza espacial; en segundo lugar el ajuste del variograma experimental a determinados modelos teóricos y finalmente el proceso de interpolación mediante krigeaje (ordinario, simple, logarítmico, de indicatriz, etc.).

El cálculo de variogramas experimentales se ha realizado para todas las variables experimentales analizadas. Los variogramas han permitido determinar la estructura de variabilidad espacial de las muestras, determinar el tamaño de la ventana de búsqueda utilizada para interpolar, para detectar efectos de anisotropía direccional en los datos y finalmente, para interpolar los datos. Puede decirse que tanto los variogramas experimentales como los modelos teóricos ajustados son la "radiografía topoprobabilística" de las variables estudiadas.

Finalmente, cabe mencionar que los resultados de la estimación de las variables experimentales han sido expresados como mapas temáticos (formato ráster). Con ello se facilita la interpretación de la distribución de la variable en el área de estudio. Seguidamente se explicarán con más detalle algunos aspectos de este proceso de construcción de mapas temáticos mediante métodos geoestadísticos.

4.4.1. Regionalización de los parámetros químicos del agua y suelo

Las estimaciones para las **muestras de las aguas** se pueden dividir en varios grupos en función de su variabilidad espacial. Para hacer esta clasificación es necesaria una primera interpretación de cada uno de los parámetros de ajuste. Así, el componente de aleatoriedad (o componente pepítica) está asociado con un comportamiento aleatorio. De este modo, dos puntos muy próximos pueden presentar una diferencia de varianza al menos igual al efecto de pepita (C_0). Un porcentaje de aleatoriedad del 100 %

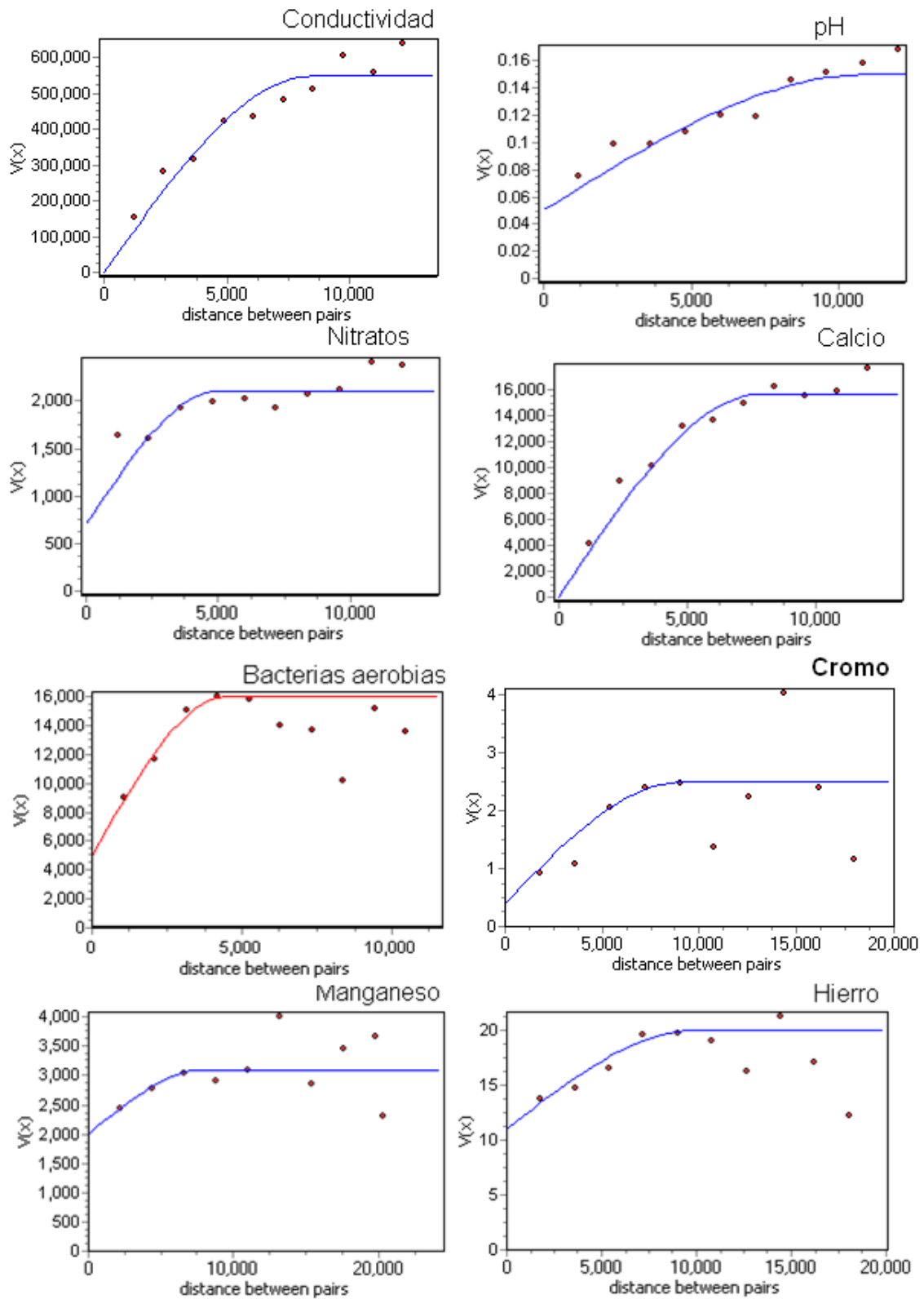


Figura 4.23: Variogramas y modelos ajustados de algunas de las variables de la calidad del agua subterránea.

<i>Variable (Z)</i>	<i>Varianza (σ^2)</i>	<i>Paso</i>	<i>Pepita (C_0)</i>	<i>Meseta (C)</i>	<i>Alcance (A)</i>	<i>% Aleatoriedad</i>
pH	0,15	1200	0,05	0,1	11000	33,3
Conductividad	548409	1200	0	548000	8800	0,0
Sulfatos	245156	1200	0	245000	8000	0,0
Calcio	15694,8	1200	0	15600	7800	0,0
Bicarbonato	6172,4	1200	1000	5200	7000	16,1
Fluoruro	0,53	1200	0	0,53	7000	0,0
Cloruros	14885,8	1200	0	15000	6000	0,0
Sodio	7090,4	1200	0	7100	6000	0,0
Potasio	21	1200	5	16	6000	23,8
Magnesio	1155,3	1200	300	800	5500	27,3
Nitratos	2427	1200	700	1400	5000	33,3
Boro	0,05	1200	0	0,047	4500	0,0
Plomo	0,5	2150	0,35	0,25	5000	58,33
Arsénico	0,135	1650	0,07	0,06	4700	53,85
Antimonio	0,00079	2000	0,00042	0,0004	5000	51,22
Manganeso	3060,37	2200	2000	1100	8000	64,52
Hierro	5381,47	1700	3900	1800	4000	68,42
Cobre	17,26	1800	11	9	10000	55,00
Cromo	1,87	1800	0,4	2,1	9100	16,00
Bacterias aerobias	16159	1050	5000	11000	4500	31,25

Tabla 4.15: Parámetros de los modelos esféricos ajustados a las variables de las aguas subterráneas.

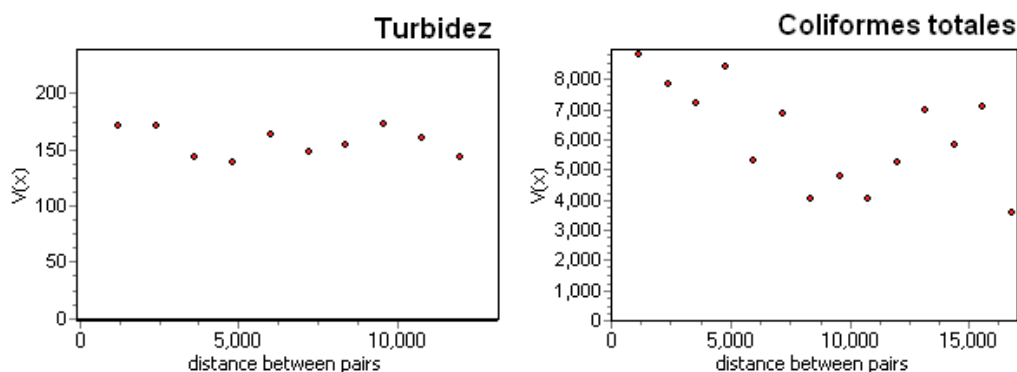
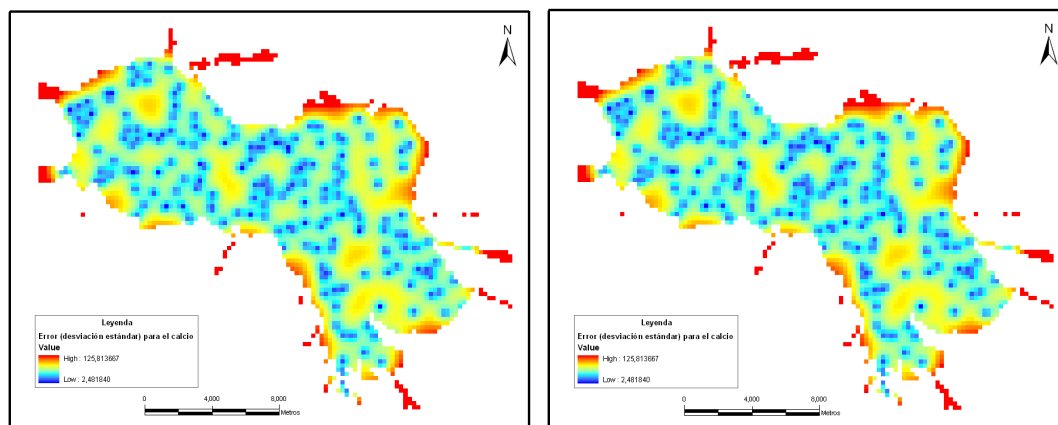


Figura 4.24: Ejemplos de comportamientos aleatorios.

significaría que la variable es completamente errática y no existe correlación espacial alguna entre los datos. Esta situación ha sido comentada en los primeros capítulos, cuando se vió la inexistencia de una estructuración de la variable observada, por ejemplo, en las variables microbiológicas y en la turbidez (figura 4.24).

El alcance o zona de influencia tiene un sentido hidrogeológico muy preciso e importante, referido al concepto de "zona de influencia" o "continuidad geológica". De este modo, el crecimiento más o menos rápido del variograma refleja el grado de deterioro de la influencia de un dato respecto a otro a medida que aumenta la distancia que los separa. Cuando el fenómeno es estacionario, el variograma crece hasta un límite llamado "meseta", a partir del cual el variograma se estacionariza. La distancia para la cual se alcanza el valor de meseta se denomina "alcance" ("memoria" si se trabaja en el tiempo), y corresponde a la distancia para la cual desaparece la correlación entre dos datos $Z(x)$ y $Z(x+h)$. Cuando el fenómeno no es estacionario, el variograma no presenta este valor límite, y crece por encima del valor de dispersión a priori. El valor de meseta coincide teóricamente con el valor de la varianza experimental de los datos (σ^2), esto es, con el valor de la covarianza $C(h)$ en el origen C_0 , o con el variograma en el infinito.

Una vez hechas estas aclaraciones teóricas cabe distinguir distintos tipos de comportamiento espacial de variables de calidad del agua subterránea (tabla 4.15 y figura 4.23). En primer lugar, podemos distinguir las variables con gran continuidad espacial, es decir, con mínima aleatoriedad y máximo alcance, que corresponderían a las variables más continuas presentes en el agua subterránea. Ejemplo de tales son: la conductividad, los sulfatos, el calcio, los fluoruros, los cloruros, el sodio y el boro; esto es, se trata principalmente de variables químicas mayoritarias y cuya presencia, generalmente, está exclusivamente ligada a las propiedades litológicas del acuífero. Aquellas variables con una aleatoriedad media (15- 50 %) y un alcance medio serían el cromo, el bicarbonato, el potasio, el magnesio, el pH y los nitratos, es decir, variables con moderada o elevada influencia antrópica o afectadas por condiciones muy particulares como puede ser



(a) Mapa del error típico de estimación del contenido en calcio de las aguas subterráneas. (b) Mapa del error relativo de estimación del contenido en calcio de las aguas subterráneas.

Figura 4.25: Mapas de errores en la estimación geostatística del contenido en calcio de las aguas subterráneas.

el bicarbonato, el pH o el magnesio. Y, finalmente, las variables con elevado grado de aleatoriedad ($>50\%$), y, por tanto, corto alcance, como el arsénico, el cobre, el plomo, el manganeso y el hierro, que corresponden a la mayoría de los metales, los cuales no presentan un origen regionalizado, sino ligado generalmente a focos puntuales de contaminación antrópica.

Uno de los resultados más interesantes del krigeaje es la el mapa de incertidumbre (varianza del error de estimación) del mapa estimado 4.25. En esta figura se representa tanto el error típico como el error relativo (%) de estimación.

El conjunto de parámetros de ajuste de los variogramas experimentales de las **variables edáficas** se ha representado en la tabla 4.16. En general los variogramas experimentales de las variables de suelo están peor estructurados que los de las variables del agua subterránea (figura 4.26), debido en parte a que el número de valores experimentales es menor y en parte a la facilidad de modificación de las variables edáficas por aspectos antrópicos, como puede ser la adición de ciertos componentes, ya sean químicos o físicos, o a la alteración del suelo natural mediante movimientos de tierra.

En términos generales, podemos distinguir entre las variables químicas y las físicas ya que presentan marcadas diferencias en sus variogramas:

Variables químicas: presentan en general un corto alcance y una peor una regionalización, ya que al tratarse del extracto de saturación del horizonte inferior, la textura, el grado de agregación o la cantidad de agua de infiltración influyen en la permanencia o lavado de éstos elementos químicos. El alcance de la mayoría de los elementos químicos es cercano a los 4 km exceptuando los sulfatos (8 km), cuyo origen está directamente relacionado con los materiales sedimentarios ricos en yesos de la roca madre. El resto de

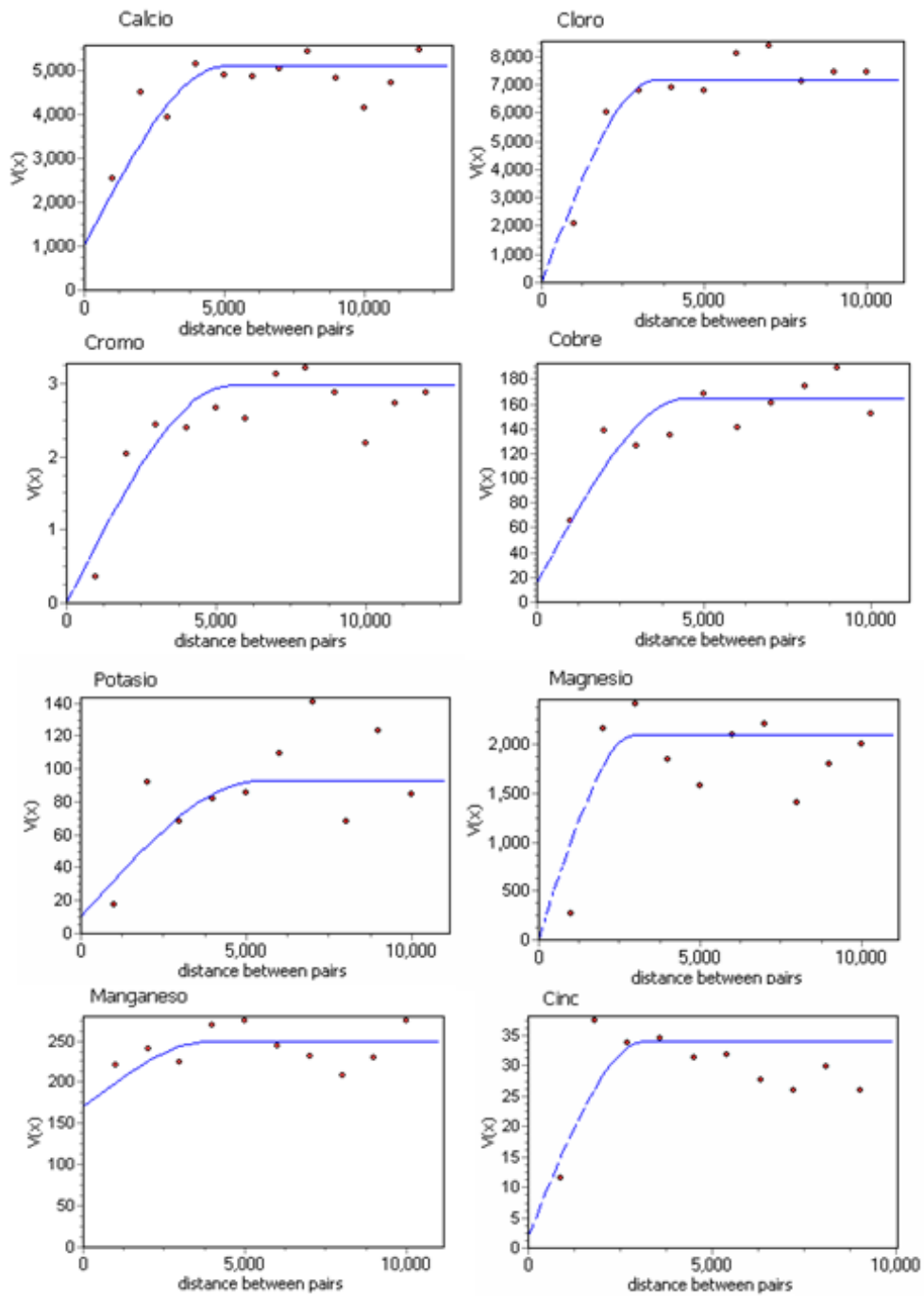


Figura 4.26: Variogramas y modelos ajustados de algunas variables químicas del suelo.

variables químicas están influidas por otros factores como las variables físicas que influyen en los procesos de depuración y por procesos antrópicos derivados de la agricultura que modifican la química de los horizontes superiores.

Variabes físicas: la regionalización de los variogramas experimentales de las variables físicas del suelo es significativamente mejor y se distinguen dos grupos en relación con su alcance. Por una parte están las variables con un alcance alto (>8 km), como la humedad promedio del suelo, las arcillas del horizonte promedio y del horizonte superior, las arenas del horizonte promedio y del horizonte superior, la humedad de saturación del horizonte inferior, el pH promedio del suelo y el carbono orgánico promedio del suelo. Estas variables con elevado alcance corresponden a las variables con buena regionalización en el acuífero y con menor influencia puntual. Estas variables son fundamentalmente las relacionadas directamente con la textura del suelo (arenas, arcillas humedad media y humedad de saturación) y las variables promedio como el pH y el carbono orgánico. La posible modificación antrópica, sin embargo, se observa en la aleatoriedad, siendo más alta ($>40\%$) en las arenas y pH del suelo. Las variables de menor alcance son aquellas estimadas para uno de los horizontes del suelo, y por tanto con una peor regionalización, y son el pH, carbono orgánico, humedad del suelo de los horizontes primero y último, los limos, las arcillas y arenas del último horizonte (80-100 cm) cuyo depósito, aunque a igual profundidad, no tiene por qué ser contemporáneo.

4.4.2. Regionalización de parámetros hidrodinámicos del acuífero

La modelización de un acuífero requiere disponer de datos relativos a los parámetros hidrodinámicos del mismo, entre los cuales destacan la conductividad hidráulica o permeabilidad, el espesor no saturado, la transmisividad, la piezometría, la litología de la zona no saturada, el gradiente hidráulico e incluso la topografía. En las secciones 2.2.1 y 2.2.2 se mostraron los datos experimentales recopilados referentes a cada una de estas variables, así como el proceso de su estimación espacial. En este apartado se quieren detallar algunos aspectos de la estimación espacial que resultan interesantes, así como la justificación del método elegido. Para ello se muestra la estimación de la transmisividad. Este parámetro es uno de los más importantes y complejos debido a la fuerte asimetría en su distribución, lo que ha requerido la transformación logarítmica de los datos, y su posterior corrección una vez estimados los datos. Esta transformación se ha aplicado también a otros datos, aunque por cuestión de espacio e importancia se ha decidido no extender este apartado y limitar el texto a la explicación del krigeaje logarítmico de la transmisividad.

La mayor parte de los datos de transmisividad proceden de sondeos realizados por la FAO (1972 [1]) y por el IGME a lo largo de todos estos años. Sin embargo, la información relativa a los sondeos no estaba procesada ni actualizada y no existía un

<i>Variable (Z)</i>	<i>Varianza (σ^2)</i>	<i>Paso</i>	<i>Pepita (C_0)</i>	<i>Meseta (C)</i>	<i>Alcance (A)</i>	<i>% Aleatoriedad</i>
Calcio	13005.7	1000	1000	4100	5000	19.61
Cloro	7144.68	1000	500	6700	4000	6.94
Conductividad	916810	1000	200000	730000	4600	21.51
Cromo	2.917	1000	0	2.97	5500	0.00
Cobre	173.528	1000	15	150	4500	9.09
Potasio	93.3888	1000	10	80	5500	11.11
Magnesio	2074.12	1000	0	2100	3000	0.00
Manganeso	241.962	1000	170	80	4000	68.00
Sodio	8149.33	1100	1500	7000	4500	17.65
Nitritos	520.7	1000	40	500	3500	7.41
Nitratos	42506.6	1050	3000	14500	4500	17.14
Sulfatos *	1.13	1000	0.31	0.75	8000	29.25
Cinc	29.4541	900	2	32	3200	5.88
Carbono orgánico ¹	0.03114	1050	0.009	0.022	5000	29.03
Carbono orgánico ²	0.1218	1100	0.003	0.026	5500	10.34
Carbono orgánico ³	0.0446	1400	0.018	0.032	16000	36.00
Humedad de saturación ¹	28390.3	1400	13500	15000	11000	47.37
Humedad del suelo ¹	1.17856	1500	0.5	0.51	5500	49.50
Humedad del suelo ²	0.8682	1300	0.18	0.4	6400	31.03
Humedad del suelo ³	0.834729	1500	0.1	0.65	9000	13.33
pH del suelo ¹	0.0366173	1100	0.009	0.029	6400	23.68
pH del suelo ²	487.667	1200	0.014	0.013	4800	51.85
pH del suelo ³	0.0391164	1400	0.018	0.027	13000	40.00
Arenas ¹	342.828	1200	145	200	4900	42.03
Arenas ²	164.123	1500	70	90	10000	43.75
Arenas ³	192.844	1500	110	120	20000	47.83
Arcillas ¹	181.216	1200	90	95	8000	48.65
Arcillas ²	160.43	1700	50	110	13500	31.25
Arcillas ³	165.942	1500	40	120	9500	25.00
Limos ¹	186.954	1400	120	75	3000	61.54
Limos ²	96.79	1250	52	45	3500	53.61
Limos ³	77.2878	1200	30	47	3000	38.96

Tabla 4.16: Ajustes de modelos esféricos de variograma de las variables físicas y químicas del suelo. 1: horizonte 80-100 cm; 2: Horizonte 0-20 cm; 3: Horizonte medio. *: Transformación logarítmica.

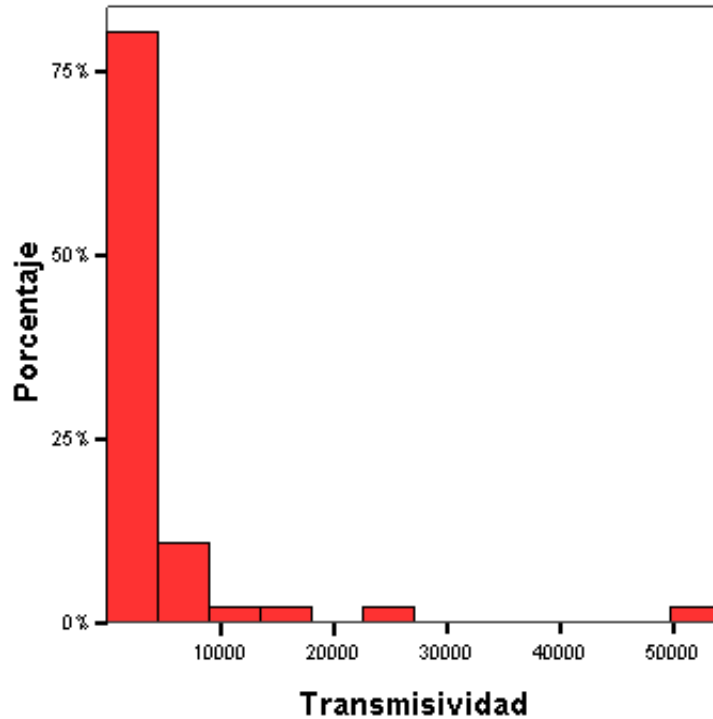


Figura 4.27: Histograma de la transmisividad.

registro digital de la misma, por lo que a partir de la documentación suministrada por del IGME (un total de 222 sondeos) se logró finalmente identificar 46 medidas de conductividad hidráulica o de transmisividad útiles. Estas 46 medidas fueron realizadas en un intervalo amplio de 50 años, por lo que a pesar de que los puntos estaban localizados espacialmente existía una gran diversidad de proyecciones, datums, etc. lo que requirió diversas correcciones hasta representarlos todos en el mismo sistema de coordenadas. En la mayoría de los casos existían datos de transmisividad y permeabilidad para cada sondeo, en los casos en que faltaba alguno de los dos datos, se calculó el parámetro que faltaba a partir de la aproximación $T \simeq K \cdot b$, siendo T la transmisividad, K la conductividad hidráulica y b el espesor saturado.

El histograma de los datos de transmisividad (figura 4.27) muestra la fuerte asimetría positiva de tipo logarítmico.

En la sección 3.3, se comentaba que el método de krigeaje es un estimador óptimo e insesgado, esto significa que la estimación no tiene sesgo y que la varianza del error es mínima. Sin embargo, al tratar con variables log-normales su optimalidad se puede degradar considerablemente (Samper-Calvete y Carrera-Ramírez, 1990 [162]). Para

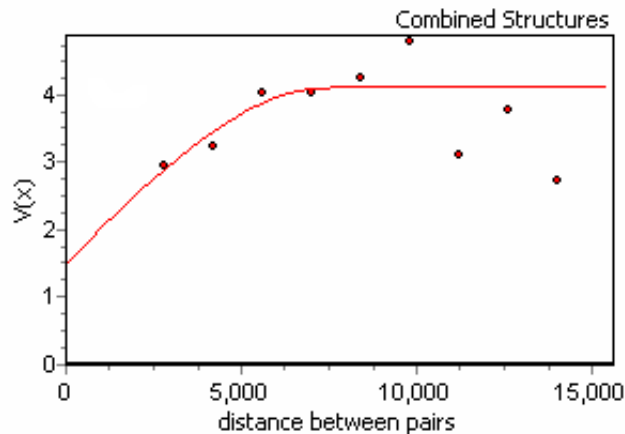


Figura 4.28: Variograma experimental y ajuste del modelo teórico log-T de la transmisividad.

superar esta limitación es conveniente realizar una transformación de T de forma que siga una distribución normal. De este modo la metodología a seguir puede resumirse en varios pasos:

1. Transformar logaritmo \ln todos los datos: $Y_i = f(T_i)$, $i = 1, \dots, n$.
2. Transformación logarítmica de T , $Y = \ln(T)$.
3. Análisis variográfico de Y .
4. Estimar el semivariograma de Y .
5. Estimación por krigeaje de la variable Y para obtener $Y^* = E(T)$ y $Var(T - T^*)$.

El cálculo de $E(T)$ y su varianza no es trivial, por lo que lo más común es hacer todos los cálculos con la variable Y , aunque debe tenerse presente que $E(T)$ no es igual a $f^{-1}(E(T))$.

Para el caso de las distribuciones log-normales es posible relacionar la media y función de covarianza de T con las de Y mediante las expresiones de Sichel (1947 [167]), pero el problema surge al tratar de dar un estimador para T . Si seguimos la teoría en sentido estricto, por lo tanto, obtenemos (Journel and Huibregts, 1978 [103]):

$$\hat{T} = e^{mY + \frac{1}{2}\sigma_Y^2}$$

$$Var(\hat{T}) = m_Z^2 e^{\sigma_Y^2(1 - e^{-\sigma^2 k Y})}$$

donde \hat{T} se sustituye por T^* para indicar que ese valor no representa el estimador de krigeaje ordinario (no es un estimador lineal). Este estimador representaría la esperanza

de T condicionada a las medidas T_1, \dots, T_n . Sin embargo, esta expresión presenta algunas dificultades como el hecho de que no siempre satisface el sesgo nulo, en el sentido de que puede diferir significativamente de la media derivada de los valores de T . Para ello Journel y Huijbregts (1978 [103]) proponen emplear como estimador $T^{**} = K \cdot \hat{T}$, donde K se determina imponiendo que la media aritmética de los valores estimados de \hat{T} coincida con m_T .

De acuerdo con este procedimiento se ha calculado el variograma experimental de los datos transformados mediante logaritmo neperiano y ajustado un modelo esférico omnidireccional con alcance de 7500 m, una meseta de 2.6 y un efecto de pepita de 2.5 (figura 4.28). Este modelo es usado para el krigeaje logarítmico.

A efectos de comparar los distintos tipos de estimación se calculan los estadísticos básicos de las transformaciones de los datos y se representan los mapas de las distintas estimaciones (4.29).

Estos mapas muestran algunos de los inconvenientes planteados y justifican la elección del modelo. Estas transformaciones se representan en la tabla 4.17 y representan los parámetros de: los datos experimentales (T), el logaritmo neperiano de los datos experimentales (Y), la estimación del logaritmo neperiano de los datos experimentales (Y^*), la inversa del logaritmo neperiano de los datos experimentales (T^*), la inversa de la estimación del logaritmo neperiano de los datos experimentales con la corrección propuesta por Journel y Huijbregts (\hat{T}), la transformación anterior con la corrección de sesgo (T^{**}).

Finalmente al comparar los resultados en todos los casos de estimación, observamos que la corrección propuesta (T^{**}) es el mejor ajuste ya que tanto la media, como la varianza y el valor mínimo y máximo de los datos son los que más se asemejan a la realidad. No obstante, cabe decir que a pesar de tratarse de la mejor estimación, los valores máximos y mínimos quedan aún lejos de la realidad, por lo que existe una gran diferencia para los datos extremos.

4.4.3. Espacialización geoestadística de índices de calidad del agua

En las secciones anteriores se ha visto la utilidad de los métodos de espacialización para la obtención de mapas de distribución de algunas de las variables más importantes del acuífero, a las que se ha denominado *variables tipo*. En esta sección quiero, además, mostrar el interés de la espacialización de índices de calidad del agua en términos de probabilidad mediante otro procedimiento geoestadístico "krigeaje de indicatrices". La principal utilidad de este tipo de aproximación es el valor sintético que ofrece, ya que el mapa elaborado no representa valores de la variable, si no la probabilidad de que las

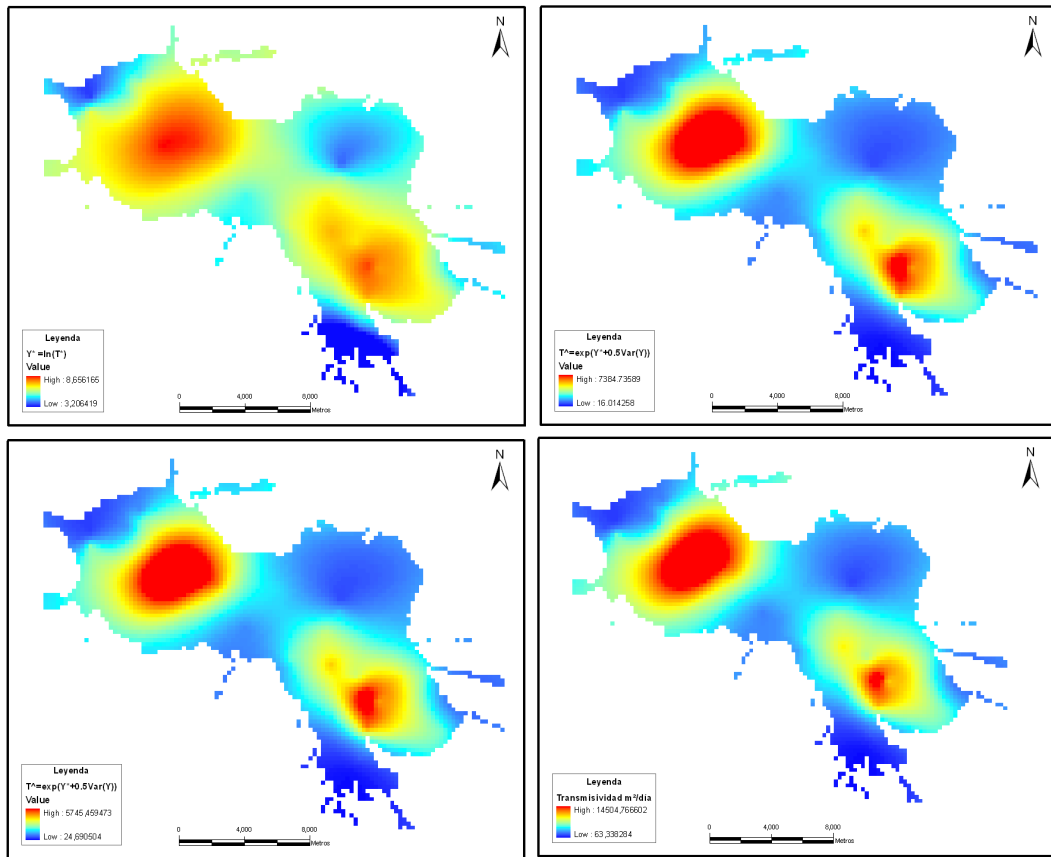


Figura 4.29: Mapas de transmisividad realizados mediante distintas estimaciones geoestadísticas.

	T	$\ln(T) = Y$	$\ln(T^*) = Y^*$	$T^* = \exp(Y^*)$	$T^* = e^{Y^* + \frac{1}{2}\sigma_Y^2}$	$T^{**} = T \cdot T^*$
Media	4,278.20	7.05	7.1	1582.1	1,893.87	4,278,20
Varianza	79,110,351.36	4.09	0.7	1,222,572.5	1,637,516.9	8,450,998.04
Mínimo	3.000	1.22	3.2	24.7	16.01	63,34
Máximo	54,432.00	10.91	8.6	5,745.5	7,348.74	14,504,77

Tabla 4.17: Parámetros estadísticos de las distintas transformaciones de la transmisividad.

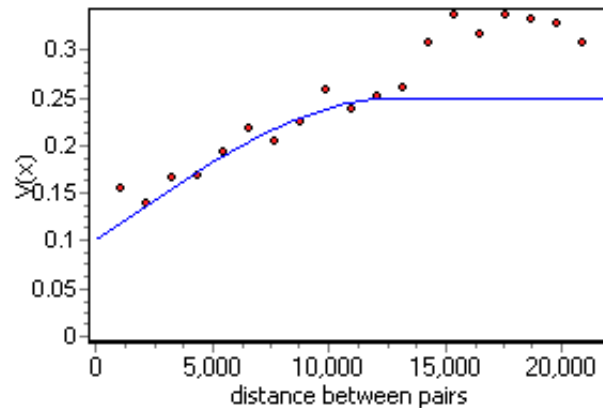


Figura 4.30: Variograma de la indicatriz mediana para el índice de calidad.

aguas del acuífero superen un valor establecido como valor alerta, a partir del cual las aguas ya no son aptas para el consumo público, además del cálculo de la función de distribución que explicamos a continuación.

En base a las expresiones mostradas en el apartado 3.3.2, se calcula la función de distribución aplicando un conjunto de valores de corte $z(u_i)$ a la variable experimental (índice de calidad del agua).

El índice calculado corresponde al elaborado según Chica-Olmo et al. (2004 [37]) y basado en el RD 140/2003. Este índice varía en un rango de 1 a 4 y establece un valor de alerta en 2.5 a partir del cual el agua no es apta para consumo público.

Para cada valor de corte establecido (10 valores tomados cada 0.2 unidades en el intervalo 1.6 a 3.4) obtenemos una indicatriz experimental que debe ser objeto de un análisis variográfico y de una estimación local, y se resuelve el correspondiente sistema de ecuaciones de krigeaje. Este procedimiento es ciertamente largo y tedioso, por lo que Deutsch y Journel (1993 [54]) sugieren usar un único variograma indicatriz, correspondiente a la mediana, para estimar el conjunto de variables indicatrices. De este modo, se asume que todos los variogramas son proporcionales por lo que, consecuentemente, los pesos obtenidos de los sistemas de krigeaje son iguales.

Según lo anterior, calculamos el variograma indicatriz de la mediana (valor = 2.69) y ajustamos el modelo: pepita=0.1, meseta=0.15 y alcance=13 km (figura 4.30). Posteriormente realizamos el krigeaje de cada una de las nuestras variables indicatriz y obtenemos el conjunto de mapas mostrados en la figura 5.2.

El conjunto de mapas krigeados (figura 5.2) permite el cálculo de la función de distribución local con una simple composición de las probabilidades estimadas. A modo de ejemplo, se han escogido tres zonas distribuidas a lo largo del cauce del río Genil, que pueden servir para estudiar las variaciones locales en la calidad del agua (Granada Sur, Purchil-Ambroz y Fuente Vaqueros) (figura 4.32).

Esta figura (figura 4.32) muestra que la calidad de las aguas del acuífero disminuye en el sentido del flujo del río principal SE-O (Granada Sur a Fuente Vaqueros) y podemos establecer que, por ejemplo, en el conjunto del municipio de Granada la probabilidad de que el agua del acuífero tenga una calidad aceptable para el consumo público es del 90 %, mientras en Purchil-Ambroz y FuenteVaqueros esta probabilidad disminuye hasta un 43 % y 30 % respectivamente. Otro dato interesante es la variación en el índice de la calidad del agua. En el municipio de Granada, esta calidad varía fundamentalmente entre los valores del índice 2 y 2.4, mientras que en Purchil-Ambroz varía entre 2.2 y 3.2 y en Fuente Vaqueros entre 2.4 y 3.2, por lo que se deduce que la calidad de las aguas en Granada capital es buena con un rango muy estrecho de variación, mientras que en Fuente Vaqueros y Purchil-Ambroz tienen en general una peor calidad del agua aunque con una amplia variación de unas zonas a otras.

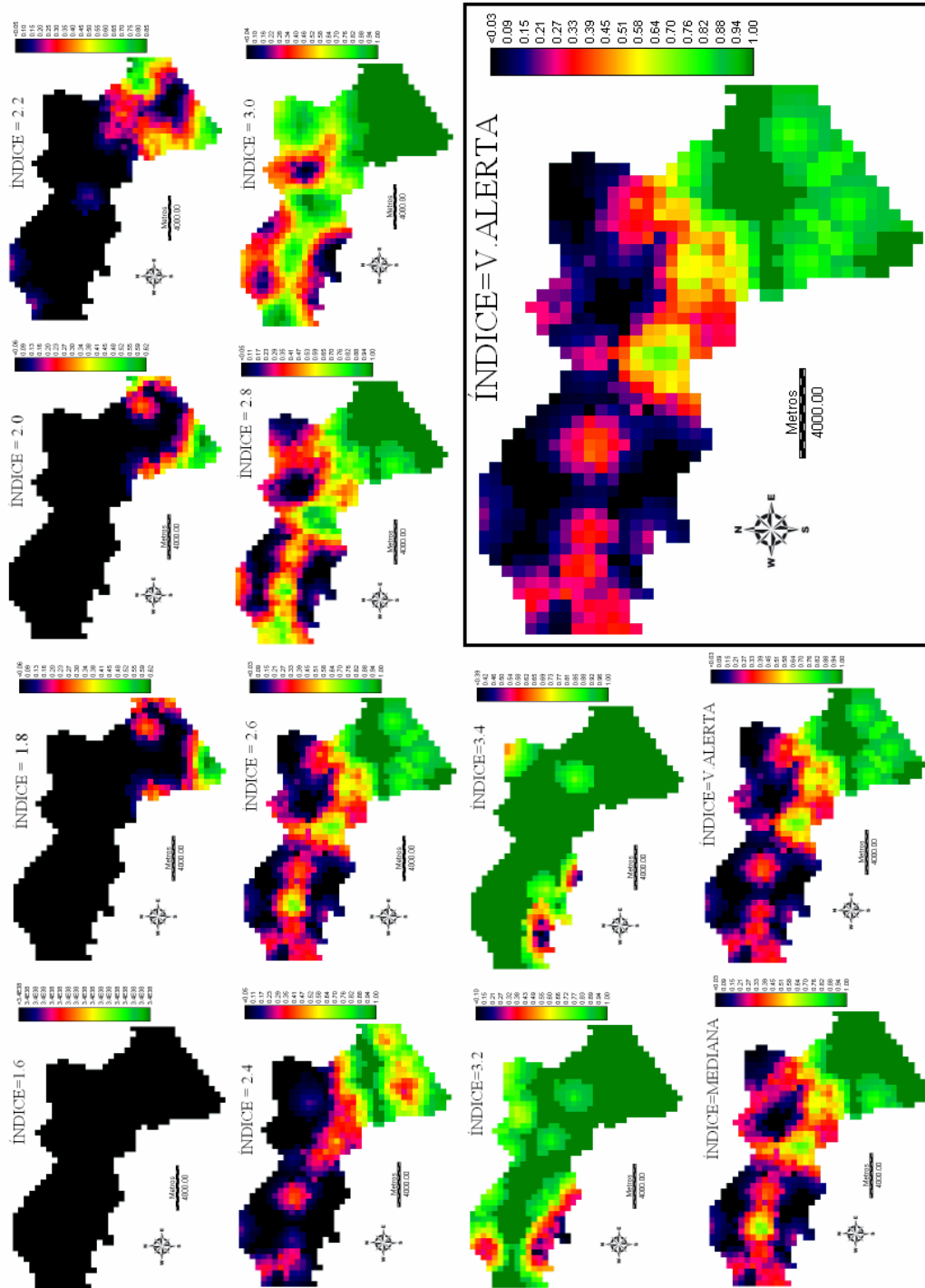


Figura 4.31: Mapas de probabilidad espacial de la calidad del agua según el índice basado en Poch.

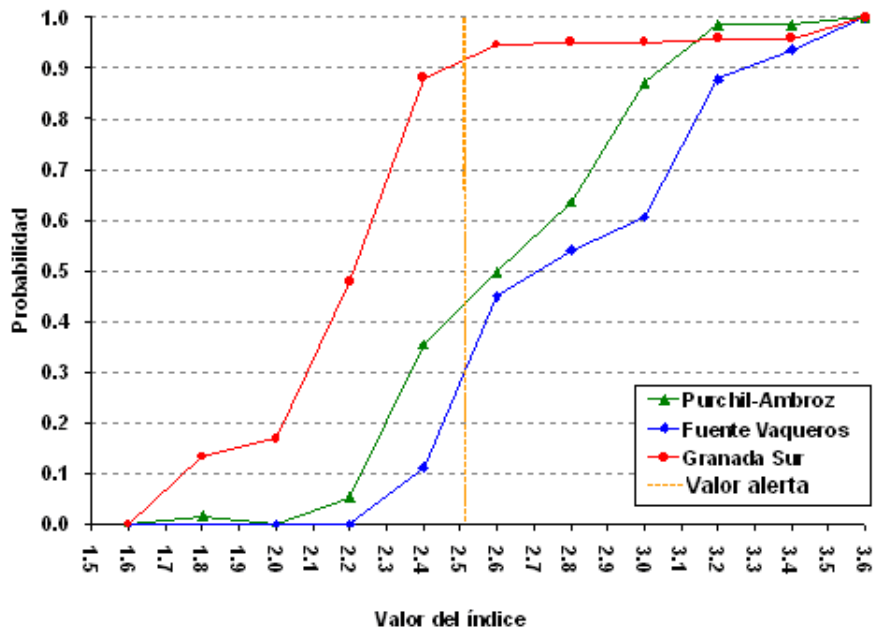


Figura 4.32: Cálculo de funciones de distribución local del índice de calidad para tres sectores del acuífero.

Capítulo 5

Integración de la información en SIG y aplicaciones al estudio de la calidad del agua subterránea

5.1. Introducción

Los métodos de análisis de datos espaciales, en general, y particularmente las técnicas de modelización, han sido objeto de estudio desde los años sesenta dentro del campo de las ciencias de la tierra (Agterber y Cabilio, 1969; Agterberg et al. 1972; Botbol et al., 1978; Chung y Agterberg, 1980; Bonham-Carter y Chung, 1983; Bonham-Carter et al., 1988; Moon et al., 1991; Pan y Harris, 1992; Goodacre et al., 1993; Fabbri y Chung, 1996; Harris et al., 1998). La principal razón de ese interés radica, posiblemente, en la necesidad que existe en estas disciplinas de extraer información que representada espacialmente permita descubrir y caracterizar, adecuadamente, las relaciones espaciales existentes entre los datos (geológicos, mineros, ambientales, etc.) y así poder utilizarlas para lo que son los objetivos últimos, la predicción y la subsiguiente toma de decisiones. En el capítulo anterior se mostraron diversos aspectos relacionados con las variables geoambientales estudiadas, así como las relaciones existentes más importantes entre ellas. En este capítulo, se pretende dar un paso adelante entrando en el ámbito de la creación de modelos espaciales de predicción, en los que a través de distintos casos de estudio se van a combinar e integrar todas las variables geoambientales en relación al acuífero de la Vega de Granada. Con estos ejemplos, he querido abordar aquí diversos procedimientos de modelización espacial sobre la base de herramienta SIG, que van desde la aplicación de sencillas reglas lógicas hasta técnicas no lineales, pasando por métodos tan conocidos como la regresión lineal múltiple o el método denominado "peso de la evidencia" basado en una sencilla suma ponderada.

Los modelos de información basados en SIG han sido clasificados en función de la metodología empleada o el objetivo perseguido. Así, Tomlin (1991) diferencia entre modelos prescriptivos y modelos descriptivos; los primeros se caracterizan por la aplicación de reglas bien establecidas en el cuerpo de doctrina en el que se enmarca el estudio (por ejemplo, los perímetros de protección); y los segundos por estar regidos por reglas surgidas de los datos. Bonham-Carter (1994 [20]) habla también de modelos prescriptivos, cuyo objetivo es el descubrimiento o predicción de nuevos "objetos". En términos generales, un "modelo SIG" puede ser considerado como el proceso de combinación de un conjunto de mapas o capas de entrada para producir un mapa de salida (Burrough, 1986; Aronoff, 1989; Berry, 1993 [25, 13, 17]).

$$\text{Mapa}_{\text{salida}} = f(\text{mapa}_1, \text{mapa}_2, \text{mapa}_3, \dots, \text{mapa}_n)$$

La función f , que puede presentar formas distintas, es categorizada por Bonham-Carter (1994 [20]) en tres tipos dependiendo de la naturaleza de la relación expresada: (a) basada en teorías y principios de la física y la química, (b) empírica, basada en observaciones de los datos (estadística o heurística), o (c) algún tipo de mezcla entre teoría y empirismo, surgiendo así otra clasificación de los modelos teóricos, empíricos e híbridos, respectivamente.

Los modelos utilizados para la evaluación de la calidad de las aguas subterráneas del acuífero de la Vega de Granada, objetivo de este trabajo, son típicamente modelos predictivos de tipo empírico (estadísticos y heurísticos) y modelos de tipo híbridos, basados tanto en teorías como en el empirismo ya que el número de variables que interviene en la calidad de las aguas subterráneas es muy elevado y los procesos que permiten la autodepuración de los contaminantes en el acuífero son muy complejos.

La evaluación de este tipo de recursos generalmente se basa en el análisis de determinadas variables, cuya importancia se asume en función de diversas hipótesis difícilmente corroborables. Como consecuencia, la selección de la forma y parámetros (pesos o coeficientes) de la función f de combinación de dichas capas mediante estos modelos empíricos, ha de basarse, bien en conocimiento experto, bien en criterios estadísticos sobre los datos. Esto da lugar a una división adicional de los modelos espaciales empíricos en dos tipos: "basados en el conocimiento" (knowledge-driven) y "basados en los datos" (data-driven) (Bonham-Carter, 1994; Pendock y Nedeljkovic, 1997 [20]). En el primer caso, los parámetros son estimados sobre la base de la opinión de un experto en el tema, y en el segundo se derivan del análisis de las relaciones espaciales entre las capas independientes (factores que afectan a la calidad) y la capa dependiente o capa de interés (calidad del agua). Los modelos basados en el conocimiento hacen uso de funciones de integración tales como la lógica booleana, la suma ponderada o la lógica difusa, mientras que los modelos basados en los datos utilizan funciones como la

regresión múltiple, el análisis discriminante, métodos probabilísticos bayesianos, redes neuronales o los SBRD.

En este contexto podemos decir que, dentro de los métodos basados en el conocimiento, se encuentran los índices, y en particular los índices de calidad del agua. En la evaluación de la calidad de las aguas, los índices son las herramientas más importantes y frecuentemente utilizadas de manera que un agua es potable o no si supera o no un valor establecido por el índice. Desde nuestro punto de vista, los índices son considerados *variables síntesis* de fácil manejo en los que se integra un conjunto más o menos amplio de variables de interés. De aquí que en la literatura existan multitud de métodos para su evaluación. Estos procedimientos se fundamentan en la elección de un grupo de variables para evaluar la calidad del agua que se transforman a subíndices (s_i) por diversas funciones y, posteriormente, se integran con distintos métodos de agregación. La mayor parte de los índices están basados en el Water Quality Index (WQI) desarrollado por Brown et al. (1970), y mejorado por Deininger para la National Sanitation Foundation (1975), donde se escogen nueve variables de calidad del agua: DO, coliformes fecales, pH, demanda bioquímica de oxígeno, cambio de temperatura, fosfatos totales, nitratos, turbidez y sólidos totales. Estos parámetros se transforman en valores de cero a cien y se integran por medio de una suma ponderada. La creación de los subíndices depende del uso del agua, en agricultura, pesca, industria, abastecimiento, etc.

Diferenciando los tipos de índices según el método de agregación, podemos distinguir entre los índices de media aritmética ponderada (Horton 1965; Brown et al. 1970; Dinius 1972, Mingo Magro 1972 [98, 24, 55, 123]), los índices de media geométrica ponderada (Landwehr et al. 1974; Walski and Parker 1974; Dinius 1987 [104, 186, 56]), la agregación mediante la raíz cuadrada de la media armónica de los cuadrados de los subíndices (Dojlido et al. 1994 [62]), el operador mínimo como función de agregación (Smith 1990 [168]) y otras variantes como la propuesta por Swamee y Tyagi (2007 [176]).

En la mayoría de los casos estos índices están basados en criterios expertos establecidos en base al objetivo perseguido. Por ejemplo, en el caso de índices de potabilidad de las aguas el criterio será el riesgo para la salud de determinadas sustancias, o riesgo de toxicidad para el cultivo en el caso de índices de calidad de aguas de riego.

5.2. Integración SIG mediante modelos basados en el conocimiento experto

Los modelos basados en conocimiento experto se caracterizan por la aplicación de métodos de combinación o integración de capas de naturaleza supervisada, en los que la aportación e intervención del analista o experto en el dominio de la aplicación resulta

esencial e indispensable. Los modelos analizados a continuación incluyen, entre otros, los basados en lógica booleana, en métodos de suma ponderada y en la lógica difusa.

El conjunto de métodos que involucran el conocimiento experto se agrupan dentro de las técnicas de Evaluación Multicriterio (EMC), donde las distintas técnicas se diferencian fundamentalmente en los procedimientos aritmético-estadísticos que se realizan sobre las matrices de evaluación y las prioridades a partir de las cuales se obtiene la evaluación final de las alternativas.

Las distintas operaciones pueden ser simples, como las del método de Sumatoria Lineal Ponderada, aplicado frecuentemente en modelos desarrollados con SIG, o pueden requerir operaciones de mayor complejidad como el Análisis del Punto Ideal (API), el Análisis Concordancia-Discordancia (ACD), el método de Optimización Jerárquica (OJ), la Programación Lineal y otros.

Existen diversas clasificaciones de los distintos métodos de evaluación multicriterio, Jankowski (1995 [101]) presenta una clasificación desde el punto de vista del tipo de procedimientos que desarrollan, siendo clasificados de acuerdo al nivel de proceso de demanda cognitiva. Se diferencian dos grupos de técnicas, las compensatorias y las no compensatorias. Así, las compensatorias son las que requieren de un mayor proceso cognitivo ya que el centro decisor debe especificar los pesos de los criterios como valores cardinales o funciones de prioridad, mientras que las no compensatorias demandan un menor proceso cognitivo del centro decisor, ya que éstas por lo general requieren una jerarquización ordinal de los criterios basada en las prioridades del centro decisor.

Otra de las clasificaciones de los diversos métodos de EMC es la realizada por Gómez Delgado y Barredo Cano (2005 [79]) (figura 5.1), que incluye las técnicas borrosas en un grupo a parte de las técnicas compensatorias, sin embargo en los últimos años se ha incrementado el uso de algunas de las técnicas compensatorias y aditivas adaptadas a la filosofía de la lógica borrosa. Dentro de este grupo de métodos se encontraría la sumatoria lineal ponderada borrosa o el método OWA, el cual se aplica en este capítulo.

5.2.1. Creación de mapas de calidad del agua mediante lógica booleana

La modelización espacial mediante lógica booleana es posiblemente el método más sencillo y consiste en la simple combinación de mapas binarios mediante los operadores estándar AND, OR, NOT y XOR del álgebra de Boole. En el análisis de la calidad del agua subterránea, los mapas binarios son cada una de las variables implicadas en la evaluación de la aptitud del agua. Los mapas de estas variables (conductividad, cloruros, metales, etc.) se construyen aplicando el operador condicional del tipo *if...then* (si...entonces) (por ejemplo, si Nitratos > 50ppm entonces, 0, si no 1) y asignando así el

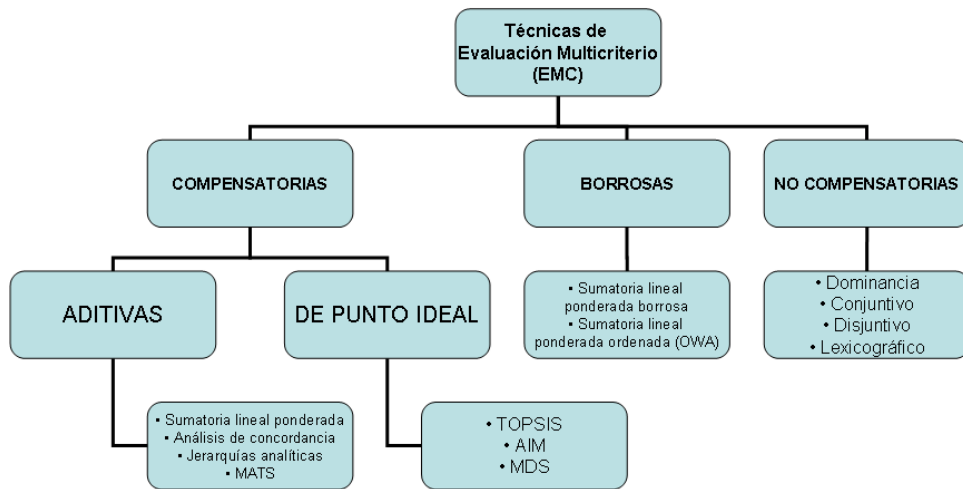


Figura 5.1: Clasificación de técnicas de EMC.

valor 0 a los píxeles que no cumplen la directiva de consumo público y a valores 1 para los que sí la cumplen.

En este apartado se muestra una pequeña variación del ejemplo anterior realizando una evaluación potencial de la contaminación del agua subterránea mediante un criterio de probabilidad. En la construcción de mapas de calidad del agua, la elección de un criterio de probabilidad facilita la identificación de zonas conflictivas a un posible gestor y amplía la información de la aptitud del agua en términos de riesgo o probabilidad. Esto es, en los mapas de las distintas variable la escala es común y su significado es la probabilidad, en %, de que se supere el valor establecido por la ley o valor límite (VL según *RD 140/2003*).

Por lo tanto, los mapas binarios se construyen a partir de mapas de probabilidad cuyo rango oscila entre 0 y 1. El valor 0 indica que la probabilidad de que la variable Z supere el valor de corte o valor límite (VL) es nula (calidad del agua apta) y el valor 1 que la probabilidad es máxima (calidad del agua no apta). El cálculo de la aptitud se realiza a cada una de las variables en cada punto experimental asignando un valor 1 a los puntos que cumplen el criterio y un valor 0 a los puntos que no cumplen este criterio.

La estimación espacial de la probabilidad de superar un valor límite (VL) en cada una de las celdillas del mapa, se realiza mediante krigeaje indicatriz (KI) (Chica-Olmo y Luque Espinar, 2002 y 2003 [38, 39, 40]). Este método se fundamenta en la transformación previa de la variable experimental en una variable binomial, variable indicatriz, $Z(x_0, vl) = \{1 \text{ si } Z(x_0) \leq vl ; 0 \text{ si } Z(x_0) > vl\}$. De este modo, un krigeaje simple de $Z(x_0, vl)$ da la estimación de la función de probabilidad $Pr[Z(x_0) \leq vl]$, que en este caso la probabilidad hace referencia que cada una de las variables no supere el valor límite

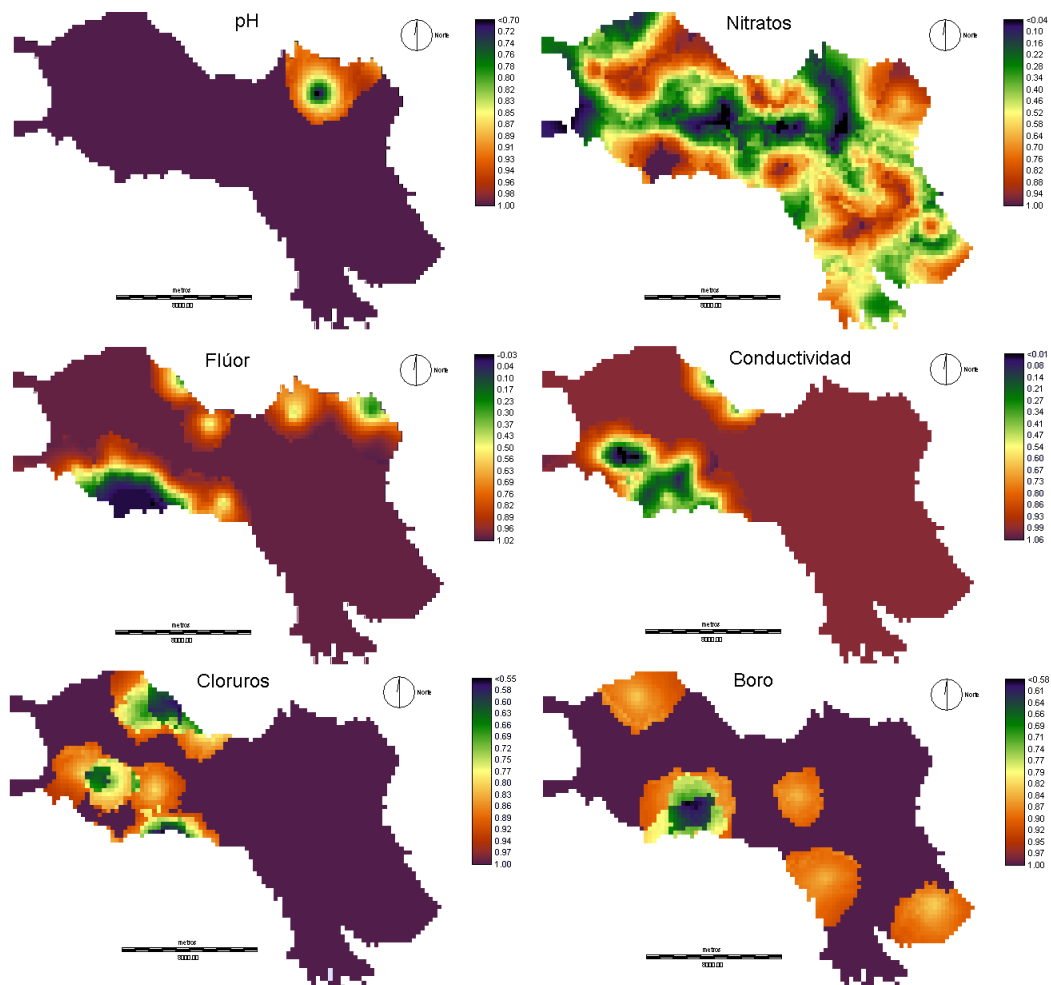


Figura 5.2: Mapas de probabilidad creados mediante krigeaje de indicatrices de algunas de las variables fisicoquímicas del agua subterránea.

establecido por el RD, es decir la probabilidad de que para cada una de las variables de calidad del agua sea apta para el consumo humano.

Cada indicatriz experimental es objeto de un análisis variográfico y de una estimación local, resolviendo el correspondiente el sistema de ecuaciones de krigeaje. Los valores de corte VL son los reglados según el RD. Sin embargo, puede ocurrir que para valores altos de VL se produzca una desestructuración de la indicatriz, lo que puede dar lugar a problemas en la creación del modelo de regionalización de la indicatriz (variograma). En estas situaciones se ha recurrido a la simplificación propuesta por Deutsch y Journel (1993) que sugieren utilizar el valor de corte correspondiente a la mediana para la modelización del variograma de las indicatrices. Las variables para las que se ha usado esta simplificación en la modelización han sido los nitratos, el pH y el amonio.

Con este procedimiento se han construido los respectivos mapas de probabilidad de las variables (figura 5.2) a partir de los cuales se han creado los mapas binarios que finalmente serán agregados mediante un criterio booleano. La creación de los mapas binarios de calidad se realiza de nuevo mediante reclasificación del tipo *si...entonces*, estableciendo dos valores de corte en función del tipo de variable. Según el RD 140/2003 los parámetros microbiológicos, químicos y radiactivos (correspondientes al anexo I, A,B y D) deben ser cumplidos estrictamente para considerar un agua apta, y son considerados como criterios de contaminación. Sin embargo, los parámetros indicadores (correspondientes al anexo I, C) deben estudiarse en cada caso concreto, en función del riesgo para la salud. En base a estas directrices se han elaborado los mapas de aptitud binarios de cada variable (figura 5.3). En el caso de los criterios de contaminación se ha establecido un 95 % de probabilidad de no superar el VL y un 75 % para los criterios de corte indicadores (tabla 5.1). Los valores 95 % y 75 % se han marcado como valores arbitrarios probabilidad mínima de contaminación del agua según los criterios de contaminación (criterios con mayor riesgo para la salud) y probabilidad baja de contaminación del agua según los criterios indicadores (criterios con riesgo moderado para la salud).

El mapa resultante de este modelo (figura 5.5) corresponde a una regionalización de la calidad del agua similar a un índice de calidad que refleja, según los criterios de probabilidad establecidos, si existe o no existe contaminación de la misma. Este mapa hace uso de un total de 23 variables o capas de entrada (excluyendo la turbidez debido a la ausencia de regionalización de esta variable). Las áreas (píxeles) consideradas como aptas son reducidas debido a la gran severidad de los criterios de probabilidad establecidos.

Desde el punto de vista metodológico la creación del mapa resultado consta de varias etapas cuyo esquema se resume en la figura 5.4. La ventaja de esta agregación de tipo booleano con criterios probabilísticos es la construcción de un mapa binario de calidad del agua, el cual es sencillo de entender y puede ser modificado fácilmente según criterios menos estrictos de probabilidad. Por ejemplo, si establecemos un valor de corte de probabilidad para determinados criterios indicadores en un 60 % en lugar de un 75 %, aumentaremos las áreas consideradas aptas.

Otra utilidad de los mapas construidos mediante lógica booleana es su integración mediante álgebra clásica. La reclasificación de los mapas intercambiando la asignación de 0 y 1, permite construir un mapa cuya suma de mapas booleanos permita la identificación de las zonas donde existe más de una variable que no cumple los criterios del RD en las probabilidades establecidas (figura 5.6). Este mapa resultado tiene como leyenda las categorías 0 a 6, lo que significa que puede haber desde 0 a 6 variables incumplan los criterios de potabilidad del agua. Un mayor número de variables conflictivas supone

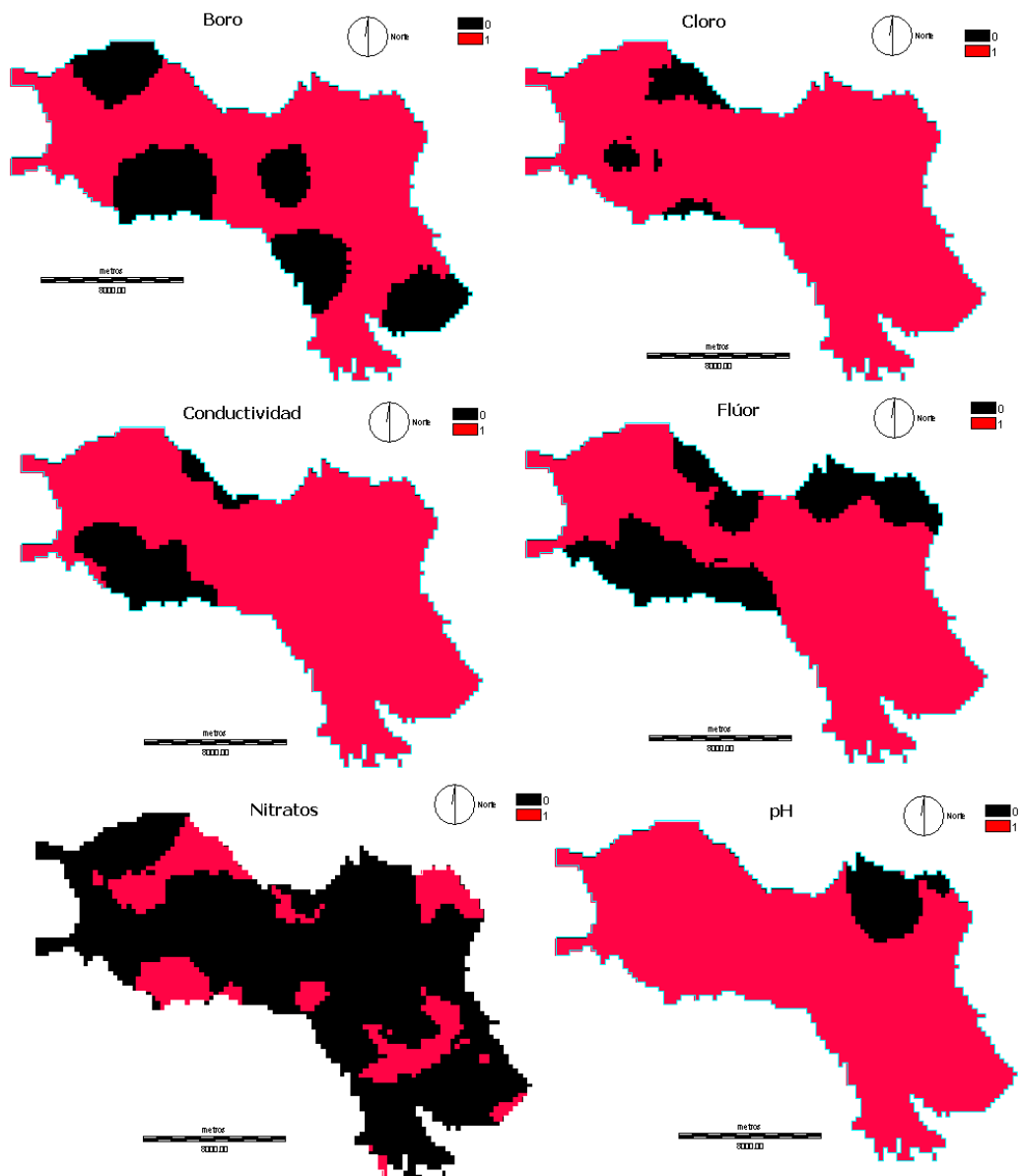


Figura 5.3: Mapas binarios de aptitud según *RD 140/2003* construido mediante criterios de probabilidad de algunas de las variables fisicoquímicas del agua subterránea.

Variable	Unidad	% puntos no aptos	Valor de corte	Tipo de variable
Conductividad	μS/cm	10.63	2500	Indicador
pH	-	0.48	<6.5 y >9.5	Indicador
Turbidez	UNF	19.32	5	Indicador
Cloruros	ppm	5.80	250	Indicador
Nitratos	ppm	50.24	50	Contaminante
Nitritos	ppm	0.00	0.1	Contaminante
Sulfatos	ppm	47.34	250	Indicador
Fosfatos	ppm	0.00	-	-
Boro	ppm	4.35	1	Contaminante
Bicarbonatos	ppm	0.00	-	-
Carbonatos	ppm	0.00	-	-
Calcio	ppm	0.00	-	-
Sodio	ppm	4.83	200	Indicador
Magnesio	ppm	0.00	-	-
Potasio	ppm	0.00	-	-
Amonio	ppm	0.97	0.5	Indicador
Fluoruros	ppm	10.63	1.5	Contaminante
Cromo	ppb	0.00	50	Contaminante
Plomo	ppb	0.00	25	Contaminante
Cobre	ppb	0.00	2000	Contaminante
Hierro	ppb	0.00	200	Indicador
Manganeso	ppb	10.63	50	Indicador
Antimonio	ppb	0.00	5	Contaminante
Arsénico	ppb	0.00	10	Contaminante

Tabla 5.1: Criterios de aptitud y porcentaje de aptitud para las variables de calidad del agua analizadas.

mayor dificultad en la depuración de esas aguas subterráneas, por lo que es interesante conocer e identificar estas áreas.

5.2.2. Creación de mapas de calidad del agua mediante suma ponderada

Este procedimiento está basado en el método "weight of evidence" (Bonham-Carter, 1994 [20]). Su fundamento es sencillo, parte de la selección de las variables (capas de información) y de la asignación de un peso o factor de ponderación por parte del experto, para posteriormente realizar la suma ponderada de las capas. El resultado es una nueva capa de integración de variables, que representan en escala relativa la favorabilidad de cada zona respecto a las hipótesis que trata de evaluar el modelo. A diferencia del método booleano, esta aproximación posibilita una selección más flexible de las zonas de interés, permitiendo además tener en cuenta la distribución espacial de las mismas en el proceso de selección (Bonham-Carter, 1994 [20] y Chica-Olmo, et al, 2002 [36]).

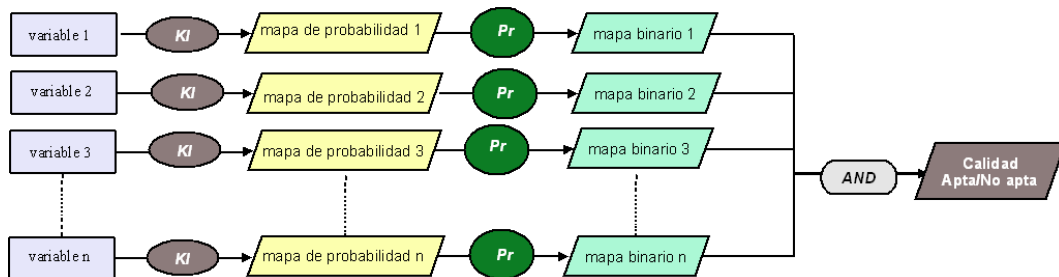


Figura 5.4: Modelo de integración de capas de tipo booleano para la creación de un mapa de calidad.



Figura 5.5: Mapa binario resultado de la superposición de zonas aptas (1) y no aptas (0) de calidad del agua subterránea según *RD 140/2003* mediante criterios de probabilidad.

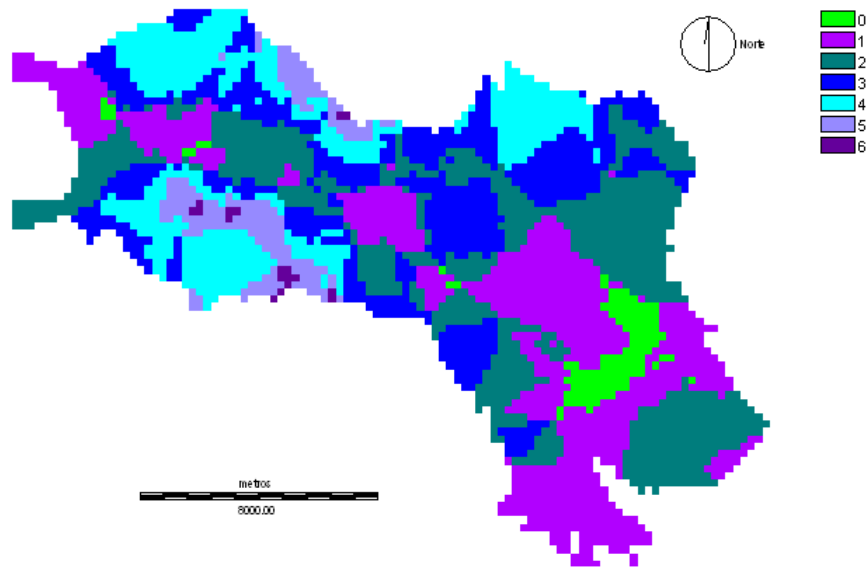


Figura 5.6: Mapa de calidad del agua resultado de la suma de los mapas binarios de probabilidad de aptitud para consumo público del agua. (Leyenda: 0-6 Variables que no cumplen los criterios de potabilidad).

El método de la sumatoria lineal ponderada es una técnica aditiva, donde el valor obtenido para cada alternativa es el producto del peso del criterio y las puntuaciones de los criterios, siendo la alternativa que obtenga el valor más alto la más adecuada o la mejor para la actividad evaluada. Es uno de los métodos más empleados en el ámbito de los SIG, siendo también el más común para la realización de índices debido a que es sencillo, intuitivo y fácil de implementar. La obtención del nivel de adecuación de cada alternativa se deduce de la siguiente fórmula:

$$r_i = \sum_{j=1}^n w_j v_{ij}$$

donde:

r_i es el nivel de adecuación de la alternativa i .

w_j es el peso del criterio j .

v_{ij} es el valor ponderado de la alternativa i en el criterio j .

Dentro de las aplicaciones de evaluación de la suma ponderada como método multicriterio, podemos encontrar la creación del índices de calidad del agua, en los cuales las distintas alternativas se refieren al valor de cada variable estudiada en el territorio y los pesos de los criterios serían la peligrosidad de cada tipo de variable al uso indicado. La transformación de cada una de las variables analizadas a un rango determinado (generalmente de 0 a 100) se suele hacer por distintos métodos y teniendo en cuenta el uso del agua, como agricultura, pesca, industria, abastecimiento, etc.

Para este apartado se han escogido como ejemplo 2 índices de calidad propuestos en la literatura, cuya diferencia radica en el modo de creación de los subíndices (transformación de la variable al rango 0-1). Las transformaciones escogidas han sido las siguientes: reclasificación a intervalos, una transformación sigmoideal y la transformación mediante funciones lineales en intervalos (Mingo Magro 1981 [123]), que es el estándar en España.

El primer índice, basado en Poch (1999 [145]), se ha creado a partir de los intervalos y pesos que se indican en la tabla 5.2, los pesos se han establecido según la importancia que el autor asigna a cada variable. La transformación por intervalos de esta variable sería el equivalente a una transformación por medio de funciones lineales en intervalos, cuyo valor de pendiente es 0 en todos los casos.

El segundo índice utilizado es el Índice de Calidad General (ICG) propuesto por Mingo Magro (1981 [123]), donde la creación de los subíndices y su ponderación se hace según las funciones de la tabla 5.3 como establece el autor.

Puesto que la aplicación del índice se realiza para cada punto experimental y la finalidad es obtener información cartográfica, es preciso aplicar un procedimiento geoestadístico de estimación espacial que consta de tres etapas. El cálculo del variograma, que ha mostrado que no existe una anisotropía espacial marcada en la variación de la calidad del agua. Por lo que a escala local se ha considerado que las variables son isotropas (figura 5.7). El ajuste de los variogramas se ha hecho considerando un modelo esférico, cuyos parámetros, efecto de pepita, meseta y alcance, quedan indicados en la tabla 5.4. El efecto de pepita representa la componente aleatoria (aproximadamente un 30 %), mientras que el modelo esférico representa la parte estructurada (70 %). La estimación de los mapas de los dos índices por krigeaje (figura 5.8) ha permitido identificar ciertas diferencias en la espacialización de la calidad del agua a pesar de la similitud en la creación de los índices.

Según Mingo Magro (1986 [123]), un valor del índice ICG superior a 0.6 supone un valor crítico bajo el cual las aguas no pueden ser utilizadas sin un tratamiento corrector. Esto es equivalente a delimitar las zonas de mejor aptitud como las distribuidas en el intervalo [0-0.6].

Un problema general en la construcción de índices de calidad es el eclipsamiento (Swamee y Tyagi, 2000 y Prabhata et al, 2007 [175, 176]). Esto significa que si un subíndice muestra una mala calidad, la agregación de la suma ponderada puede no reflejar estos valores bajos. Por lo tanto es posible identificar zonas de mala calidad del agua subterránea que según el índice sean buenas.

Para tratar de solventar este problema se han propuesto en la literatura diversos métodos de agregación libres de eclipsamiento. Entre ellas encontramos la media geométrica ponderada (Landwehr et al., 1974; Walski y Parker, 1974 y Dinius, 1987 [104, 186, 56]), la raíz cuadrada de la media armónica de los cuadrados de los subíndices (Dojlido et

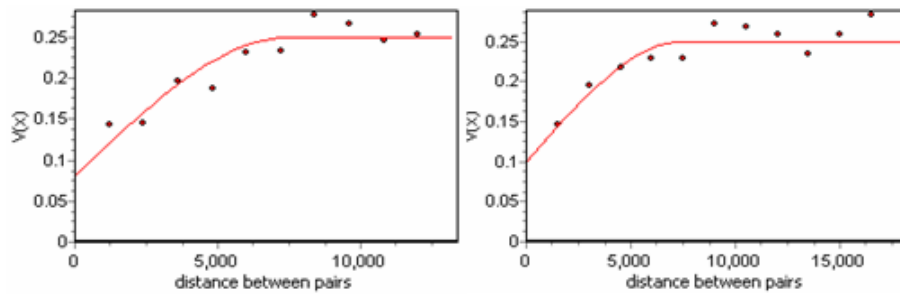
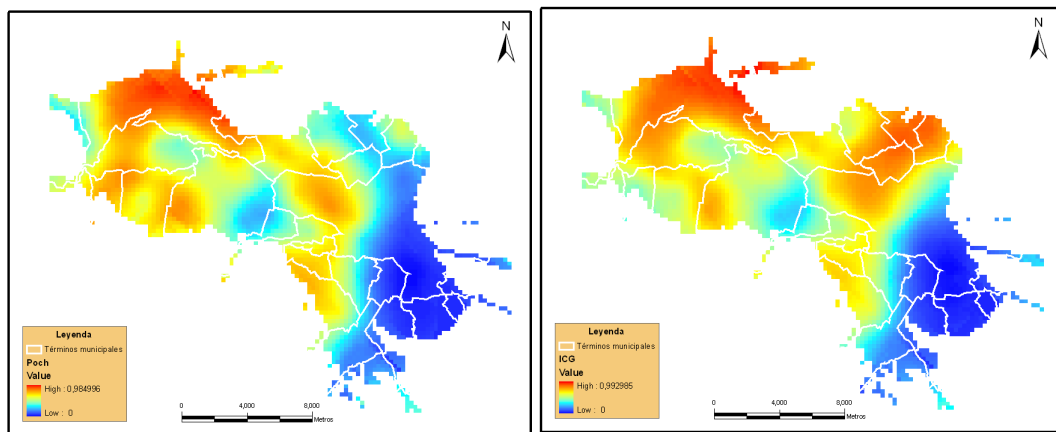


Figura 5.7: Variogramas experimentales de las variables indicatriz mediana correspondientes a los índices de calidad del agua de Poch (izquierda) y Mingo Magro (derecha) Mediana(Poch)=90.12, Mediana(CE)=80.27



(a) Mapa de calidad del agua basado en Poch (b) Mapa de calidad del agua ICG (transformación de subíndices por intervalos) (transformación de subíndices mediante funciones a trozos)

Figura 5.8: Mapa de isoprobabilidad (krigeaje de indicatrices) de dos índices de calidad del agua con un valor indicatriz igual a la mediana de la distribución.

al. 1994 [62]), el operador mínimo (Smith, 1990 [168]), etc. Sin embargo, hoy día, la suma ponderada sigue siendo uno de los métodos más usados debido a su sencilla interpretación y a su utilidad en la monitorización y seguimiento de tendencias, ya que métodos como el operador mínimo no tiene el problema del eclipsamiento, pero falla a la hora de componer una escena de la calidad del agua. El operador mínimo no refleja si hay cambios en más de una variable que no cumple la condición dada y tampoco sirve para comparar tendencias. Otros como la media geométrica ponderada pueden tener eclipsamiento en pesos cercanos a cero, etc.

Variable	Unidad	Pesos	Intervalo 1	Valor	Intervalo 2	Valor	Intervalo 3	Valor	Intervalo 4	Valor	Intervalo 5	Valor
Conductividad	(μ S/cm) a 25°C	0.05	<625	100	<1250	75	<1875	50	<2500	25	>2500	0
pH		0.07	0.07	0	6.5 - 7.5	65	7.5 - 8.5	100	8.5 - 9.5	65	9.5 - 14	0
Turbidez	(U.N.F.)	0.00	<1.25	100	<2.5	75	<3.75	50	<5	25	>5	0
Cloruros	(mg/l)	0.00	<62.5	100	<125	75	<187.5	50	<250	25	>250	0
Nitratos	(mg/l)	0.00	<12.5	100	<25	75	<37.5	50	<50	25	>50	0
Nitritos	(mg/l)	0.07	<0.125	100	<0.25	75	<0.375	50	<0.5	25	>0.5	0
Sulfatos	(mg/l)	0.05	<62.5	100	<125	75	<187.5	50	<250	25	>250	0
Fosfatos	(mg/l)	0.00	<0.25	100	<0.5	75	<0.75	50	<1	25	>1	0
Boro	(mg/l)	0.07	<0.25	100	<0.5	75	<0.75	50	<1	25	>1	0
CO3H-	(mg/l)	0.00	-	-	-	-	-	-	-	-	-	-
CO3-	(mg/l)	0.00	-	-	-	-	-	-	-	-	-	-
Calcio	(mg/l)	0.02	<62.5	100	<125	75	<187.5	50	<250	25	>250	0
Sodio	(mg/l)	0.05	<50	100	<100	75	<150	50	<200	25	>200	0
Magnesio	(mg/l)	0.00	<62.5	100	<125	75	<187.5	50	<250	25	>250	0
Potasio	(mg/l)	0.07	-	-	-	-	-	-	-	-	-	-
Amonio	(mg/l)	0.00	<0.125	100	<0.25	75	<0.375	50	<0.5	25	>0.5	0
Fluoruros	(mg/l)	0.07	<0.375	100	<0.75	75	<1.125	50	<1.5	25	>1.5	0
Plomo	ppb	0.07	<6.25	100	<12.5	75	<18.75	50	<25	25	>25	0
Cromo	ppb	0.07	<12.5	100	<25	75	<37.5	50	<50	25	>50	0
Cobre	ppb	0.00	<0.5	100	<1	75	<1.5	50	<2	25	>2	0
Hierro	ppb	0.00	<50	100	<100	75	<150	50	<200	25	>200	0
Manganeso	ppb	0.07	<12.5	100	<25	75	<37.5	50	<50	25	>50	0
Cadmio	ppb	0.07	<1.25	100	<2.5	75	<3.75	50	<5	25	>5	0
Arsénico	ppb	0.05	<2.5	100	<5	75	<7.5	50	<10	25	>10	0
Antimonio	ppb	0.07	<1.25	100	<2.5	75	<3.75	50	<5	25	>5	0

Tabla 5.2: Intervalos de transformación y pesos del índice basado en Poch por intervalos.

Variable	Unidad	Pesos	Funciones		
Conductividad	(mS/cm) a 25°C	0.03	$Si\ x < 1000$; $s_i = 120 - 0.08\ x$	$Si\ x < 2000$; $s_i = 80 - 0.04\ x$	$Si\ x > 2000$; $s_i = 0$
pH	-	0.03	$Si\ x < 4.5$; $s_i = 0$	$Si\ x < 6.5$; $s_i = 30\ x - 135$	$Si\ x < 8$; $s_i = 40\ x - 200$
Turbidez	(U.N.F.)	0.00	$Si\ x < 5$; $s_i = 100 - 20\ x$	$Si\ x > 5$; $s_i = 0$	
Cloruros	(mg/l)	0.05	$Si\ x < 400$; $s_i = 100 - 0.1\ x$	$Si\ x < 2000$; $s_i = 75 - 0.0375\ x$	$Si\ x > 2000$; $s_i = 0$
Nitratos	(mg/l)	0.08	$Si\ x < 250$; $s_i = 100 - 0.4\ x$	$Si\ x > 250$; $s_i = 0$	
Nitritos	(mg/l)	0.00	$Si\ x < 0.5$; $s_i = 100 - 200\ x$	$Si\ x > 0.5$; $s_i = 0$	
Sulfatos	(mg/l)	0.05	$Si\ x < 1250$; $s_i = 100 - 0.08\ x$	$Si\ x > 1250$; $s_i = 0$	
Fosfatos	(mg/l)	0.08	$Si\ x < 3.75$; $s_i = (1500 - 400\ x)/15$	$Si\ x > 3.75$; $s_i = 0$	
Boro	(mg/l)	0.05	$Si\ x < 1$; $s_i = 100 - 100\ x$	$Si\ x > 1$; $s_i = 0$	
Bicarbonato	(mg/l)	0.00	-	-	-
Carbonato	(mg/l)	0.00	-	-	-
Calcio	(mg/l)	0.08	$Si\ x \leq 5$; $s_i = 50$	$Si\ x \leq 200$; $s_i = 100$	$Si\ x < 2000$; $s_i = (2000 - x)/30$
Sodio	(mg/l)	0.10	$Si\ x < 200$; $s_i = 100 - 0.25\ x$	$Si\ x < 1000$; $s_i = (1000 - x)/16$	$Si\ x > 1000$; $s_i = 0$
Magnesio	(mg/l)	0.10	$Si\ x < 150$; $s_i = 100 - 4x/15$	$Si\ x < 1000$; $s_i = 6/85 (1000 - x)$	$Si\ x > 1000$; $s_i = 0$
Potasio	(mg/l)	0.00	-	-	-
Amonio	(mg/l)	0.00	$Si\ x < 0.5$; $s_i = 100 - 0.005\ x$	$Si\ x > 0.5$; $s_i = 0$	
Fluoruros	(mg/l)	0.00	$Si\ x < 1.5$; $s_i = 100 - 66.7\ x$	$Si\ x > 1.5$; $s_i = 0$	
Plomo	(mg/l)	0.05	$Si\ x < 0.2$; $s_i = 100 - 500\ x$	$Si\ x > 0.2$; $s_i = 0$	
Cromo	(mg/l)	0.03	$Si\ x < 0.125$; $s_i = 100 - 800\ x$	$Si\ x > 0.125$; $s_i = 0$	
Cobre	(mg/l)	0.05	$Si\ x \leq 0.01$; $s_i = 100$	$Si\ x < 1$; $s_i = 101 - 101\ x$	$Si\ x > 1$; $s_i = 0$
Hierro	(mg/l)	0.05	$Si\ x < 200$; $s_i = 100 - 0.5\ x$	$Si\ x > 200$; $s_i = 0$	
Manganeso	(mg/l)	0.05	$Si\ x < 50$; $s_i = 100 - 0.2\ x$	$Si\ x > 50$; $s_i = 0$	
Cadmio	(mg/l)	0.03	$Si\ x < 0.01$; $s_i = 100 - 10000\ x$	$Si\ x > 0.01$; $s_i = 0$	
Arsénico	(mg/l)	0.05	$Si\ x < 10$; $s_i = 100 - 10\ x$	$Si\ x > 10$; $s_i = 0$	
Antimonio	(mg/l)	0.05	$Si\ x < 5$; $s_i = 100 - 20\ x$	$Si\ x > 5$; $s_i = 0$	

Tabla 5.3: Intervalos de transformación para el índice ICG y pesos.

Índice	Pepita	Meseta	Alcance (m)
Poch (intervalos)	0.08	0.17	7500
ICG (Mingo Magro)	0.1	0.15	7500

Tabla 5.4: Ajustes a modelos teóricos de las variables indicatriz de los índices de calidad del agua.

5.2.3. Creación de mapas de calidad del agua mediante lógica difusa

Como ya se comentó en el apartado de metodología, la lógica difusa extiende el razonamiento clásico de la teoría de conjuntos, basado en la pertenencia excluyente a un conjunto (pertenece o no pertenece, 0 ó 1), para considerar estados intermedios en los que un elemento puede presentar un grado de pertenencia intermedio a un conjunto (entre 0 y 1) (Zadeh, 1965 [197]). La lógica booleana es, por tanto, un caso especial de la lógica difusa, concretamente cuando hay certidumbre sobre los datos y los valores de pertenencia se reducen a 0 y 1 (certidumbre), de forma que el conjunto se convierte en discreto y los resultados coincidirán con los obtenidos con la lógica booleana (Cox, 1994 [47]).

Los métodos de lógica difusa han sido aplicados en distintos aspectos de la modelización espacial con SIG, en este sentido la lógica difusa supone un punto intermedio entre los rígidos modelos booleanos y los modelos continuos de la suma ponderada. En el contexto de este trabajo relacionado con la calidad del agua subterránea la aptitud de una variable es considerada como un concepto difuso. Con ello se indica el grado de pertenencia de una variable en los conjuntos "bueno" o "malo", de forma que una clase "muy buena", es seguramente una variable del conjunto "bueno", mientras que una clase "mala", no sería considerada del conjunto "bueno".

Desde esta perspectiva, se ha establecido un índice difuso (ID) de calidad basado en los criterios de la Directiva Marco (DMA) y el RD 140/2003, en el que los subíndices se han creado por medio de funciones de pertenencia difusas sigmoidales (s-shaped), asignando un valor cero en el momento en el que se superan los límites de la normativa. La función s-shaped se define por cuatro parámetros: a , b , c y d (figura 5.9) representados en la tabla 5.5. La función que define la transformación sigmoidal es:

$$\mu(x) = \cos^2 \alpha$$

en el caso de que se aplique una función sigmoidal decreciente el valor α es igual a:

$$\alpha = \frac{x - c \cdot \pi}{d - c \cdot 2}$$

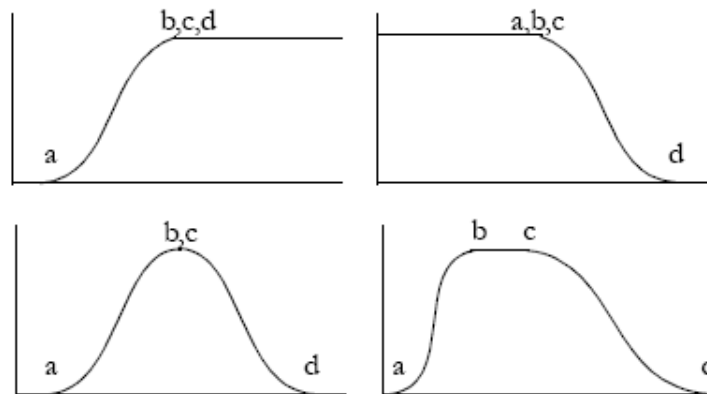


Figura 5.9: Funciones de pertenencia sigmoideas.

cuando $x < c$; $\mu(x) = 1$. En el caso de una función monótonamente creciente, α vale:

$$\alpha = \frac{1 - (x - a)}{b - a} \cdot \frac{3}{2}$$

cuando $x > b$; $\mu(x) = 1$.

El objetivo de transformar las variables mediante funciones de pertenencia difusas es obtener un conjunto de variables en un mismo rango sin perder la información de los valores intermedios. De este modo, una suma ponderada nos puede aportar mayor información acerca de la calidad del agua. Por esta razón, la aplicación se ha realizado considerando la suma ponderada de los subíndices de las variables transformadas mediante funciones difusas. La cartografía del índice, como en los casos anteriores, se ha realizado mediante estimación por krigeaje. (figuras 5.10 y 5.11; y tabla 5.6).

Debido a la transformación de las variables mediante funciones difusas la continuidad espacial de este índice de la calidad del agua, respecto a los dos anteriores, es mayor (alcance 11000 m). La comparación del mapa resultado (figura 5.11) con los mapas anteriores (figura 5.8) muestran que la distribución espacial de los tres índices de calidad son similares en el acuífero. No obstante, se aprecian algunas diferencias locales particularmente en la zona noreste del acuífero. En esta zona, el índice Poch asigna una mejor calidad, el índice ICG una calidad peor y el índice basado en funciones difusas asigna una calidad intermedia.

5.2.4. Creación de mapas de calidad del agua mediante agregación exponencial

Otro método de creación de índices de calidad del agua es la agregación exponencial. El método elegido es el propuesto por Swamee y Tyagi (2007 [176]), que presentan

Variable	Unidad	Pesos	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>
Conductividad	(mS/cm) a 25°C	0.04			500	2500
pH		0.02	1	7.5	8.5	13
Turbidez	(U.N.F.)	0.00			0.5	5
Cloruros	(mg/l)	0.04			25	250
Nitratos	(mg/l)	0.07			5	50
Nitritos	(mg/l)	0.00			25	250
Sulfatos	(mg/l)	0.07			25	250
Fosfatos	(mg/l)	0.04			0.1	1
Boro	(mg/l)	0.07			0.1	1
Bicarbonato	(mg/l)	0.00			-	-
Carbonato	(mg/l)	0.00			-	-
Calcio	(mg/l)	0.04			25	250
Sodio	(mg/l)	0.04			20	200
Magnesio	(mg/l)	0.04			25	250
Potasio	(mg/l)	0.00			-	-
Amonio	(mg/l)	0.00			0.05	0.5
Fluoruros	(mg/l)	0.00			0.15	1.5
Plomo	ppb	0.07			2.5	25
Cromo	ppb	0.07			5	50
Cobre	ppb	0.07			0.2	2
Hierro	ppb	0.07			20	200
Manganeso	ppb	0.07			5	50
Cadmio	ppb	0.07			0.5	5
Arsénico	ppb	0.07			1	10
Antimonio	ppb	0.07			0.5	5

Tabla 5.5: Intervalos de transformación y pesos del índice creado a través de funciones sigma.

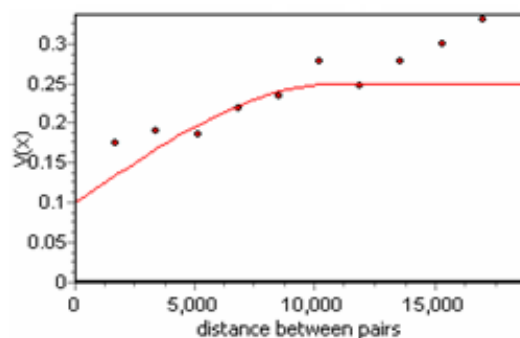


Figura 5.10: Variograma experimental de la variable indicatriz mediana correspondiente al índice ID de calidad del agua creado mediante funciones difusas. Mediana=89.99

Índice	Pepita	Meseta	Alcance (m)
CE (fuzzy)	0.1	0.15	11000

Tabla 5.6: Ajuste del modelo teórico de las variable indicatriz del índice de calidad del agua creado mediante funciones difusas.

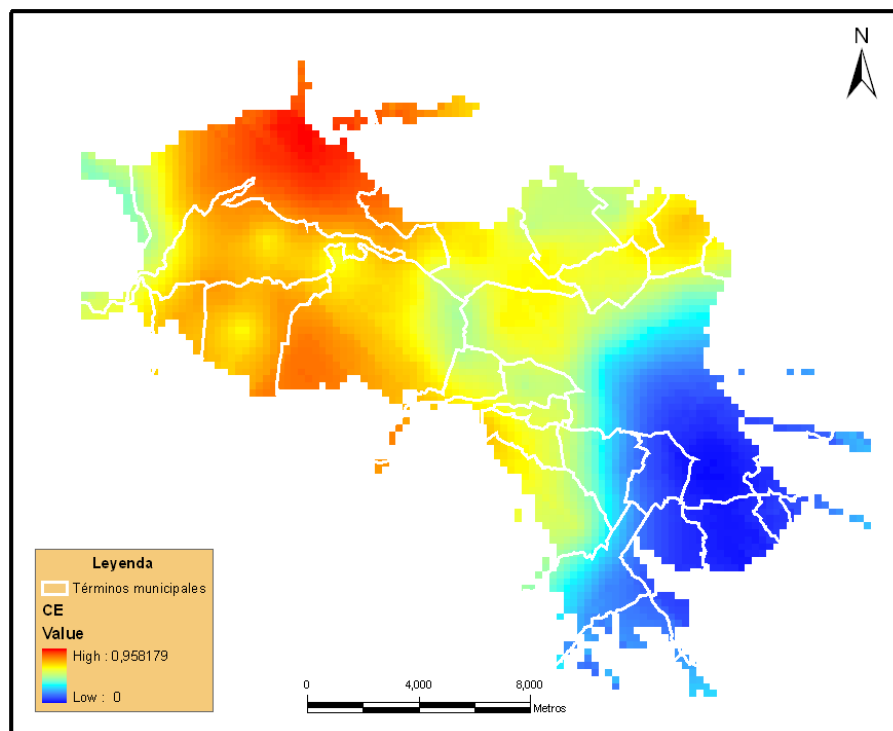


Figura 5.11: Mapa de isoprobabilidad (krigeaje de indicatrices) del índice de calidad del agua ID construido con funciones de pertenencia difusas para un valor indicatriz igual a la mediana de la distribución.

un método libre de ambigüedad, eclipsamiento y rigidez, en el que la creación de los subíndices se ha hecho de acuerdo con los límites impuestos por la ley (RD 140/2003).

De forma análoga a lo visto en los epígrafes anteriores, el procedimiento de integración comienza con la creación de los subíndices mediante diferentes funciones. La creación de subíndices monótonamente decrecientes, es decir, variables cuyo valor de calidad es mejor cuanto menor sea la concentración, como es, por ejemplo, el caso de los nitratos (figura 5.12a), la ecuación del subíndice es:

$$s = \left(1 + \frac{q}{q_c}\right)^{-m}$$

donde q = concentración de una variable del agua, q_c = valor característico, y m es un exponente positivo, que es el resultado de ajustar el valor límite de la concentración a un subíndice de 0.25 a partir del cual la calidad no sería apta. Por ejemplo, el límite de los nitratos está en 50 ppm, así que aplicando la transformación mediante la función monótonamente decreciente el valor del subíndice para esa concentración sería 0.25. Los valores inferiores estarán en el rango 0.25-1 y los superiores estarán en el rango 0-0.25. Por lo que la mejor calidad tendrá valores próximos a 1.

La ecuación genérica para índices no uniformemente decrecientes es:

$$s = \frac{1 + \left(\frac{q}{q_t}\right)^4}{1 + 3\left(\frac{q}{q_t}\right)^4 + 3\left(\frac{q}{q_t}\right)^8}$$

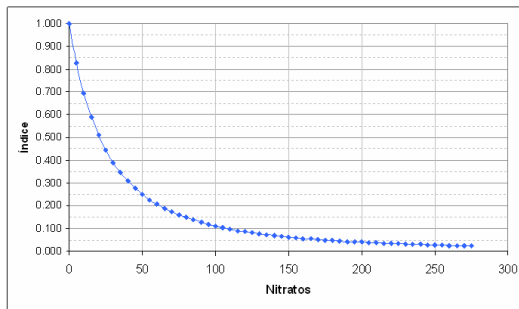
donde q_t es la concentración límite a partir de la cual el valor del subíndice disminuye rápidamente, como por ejemplo el plomo (figura 5.12b).

Los subíndices unimodales con un máximo en $s = 1$ para $q = q^*$ se calculan mediante la ecuación:

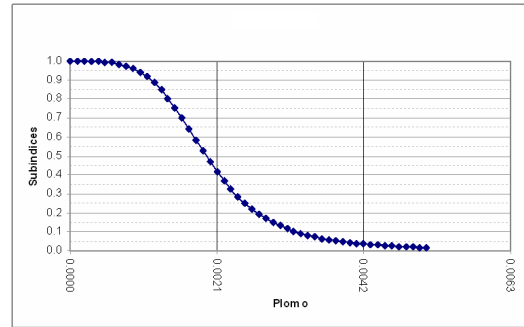
$$s = \frac{pr + (n + p)(1 - r)\left(\frac{q}{q^*}\right)^n}{p + n(1 - r)\left(\frac{q}{q^*}\right)^{n+p}}$$

donde r = subíndice para $q = 0$; q^* = valor característico de q , y n y p son exponentes. Estos valores se han calculado del mismo modo que para la ecuación monótonamente decreciente, donde los valores no aptos corresponden a subíndices entre el rango 0-0.25 y los aptos están en el rango 0.25-1. Como ejemplo, se representa la función unimodal de la conductividad y del pH (figura 5.13a y 5.13b).

El siguiente paso corresponde al proceso de agregación para la construcción del índice. Así, un índice I libre de ambigüedad debe cumplir que si todas excepto una de las variables de calidad del agua son de buena calidad, las variables con calidad aceptable no deben enmascarar a la variable buena y la agregación debe reflejar el subíndice de la

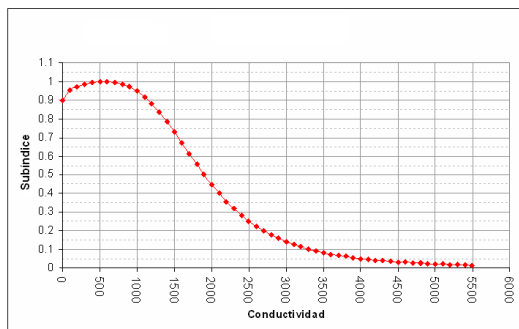


(a) Función monótonamente decreciente para los nitratos.

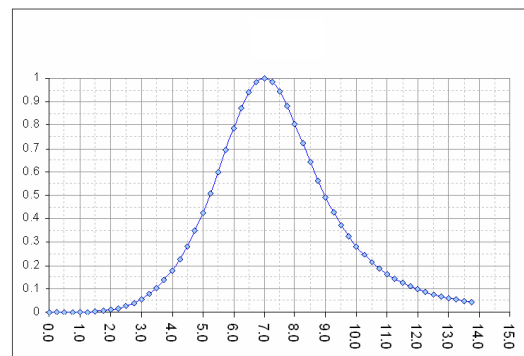


(b) Función no uniformemente decreciente para el plomo.

Figura 5.12: Funciones de transformación de los subíndices de los nitratos y el plomo.



(a) Función unimodal de la conductividad.



(b) Función unimodal del pH.

Figura 5.13: Funciones de transformación de los subíndices de la conductividad y el pH.

variable de mala calidad. Esta propiedad hace que el proceso de agregación sea insensible al número de variables involucradas. Esto es:

$$I(1, 1, \dots, 1, s_i, 1, \dots, 1) = s_i$$

Además, para que el índice I esté libre del eclipsamiento, la agregación debe estar sesgada a los subíndices de mala calidad. Es decir, debe cumplir:

$$I(s_1, s_2, s_3, \dots, s_{i-1}, 0, s_{i+1}, \dots, s_n) = 0$$

Ya que I es una función monótonamente creciente de subíndices, la reducción de los subíndices de 1 a s debe reducir la agregación, esto es:

$$I(s, s, s, \dots, s) \geq s$$

De este modo Swamee y Tyagi (2000 y 2007 [175, 176]) proponen el siguiente método de agregación exponencial:

$$I = \left(1 - N + \sum_{i=1}^N s_i^{\frac{-1}{k}} \right)^{-k}$$

donde N es el número de subíndices que participan el índice I , y k se determina:

$$k = \frac{1}{\log_2(N - 1)}$$

de modo que si el número de subíndices aumenta, el índice de agregación I no se hace más pequeño y se solventa el problema de rigidez y ambigüedad, que se produciría usando un k positivo y constante para cualquier N (Swamee y Tyagi, 2007 [176]).

Vista la etapa de agregación, el método finaliza aplicando la metodología geoestadística para elaborar un mapa de la distribución espacial del índice de la calidad del agua del acuífero de la Vega de Granada. Las ventajas de este procedimiento es que el método de krigeaje permite obtener tanto el valor más probable del índice en cualquier punto del acuífero (Krigaje Ordinario), como el mapa de probabilidad de que el índice supere un valor dado (Krigaje de Indicatrices) (Chica-Olmo y Luque Espinar, 2002 y 2003 [38, 40]) (figura 5.15) comentado en el apartado 5.2.1.

En el caso particular de este estudio, quiero mostrar otra de las ventajas que proporciona la aplicación del método geoestadístico. Concretamente, se trata de realizar un análisis local de los índices de calidad del agua en términos de probabilidad espacial (cálculo de la función de probabilidad local). De modo resumido, se ha calculado la función de probabilidad local de los índices previamente establecidos en cada uno de los municipios de la Vega (suma ponderada: CE y Poch; lógica difusa: ID; y agregación

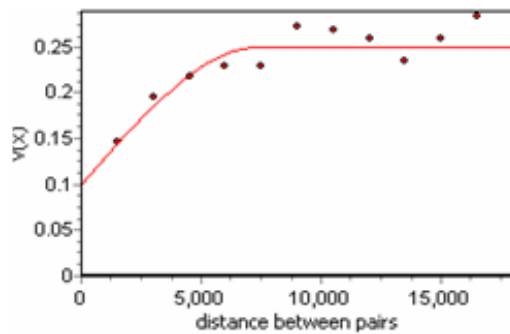


Figura 5.14: Variograma experimental de la variable indicatriz mediana correspondiente al índice de calidad del agua creado mediante agregación exponencial. Media-na(WQI)=29.51.

Índice	Pepita	Meseta	Alcance (m)
WQI (Samwee y Tyagi, [176])	0.11	0.14	8000

Tabla 5.7: Ajustes del modelo teórico de la variable indicatriz del índice de calidad del agua.

exponencial: WQI) según el procedimiento descrito por Chica-Olmo y Luque-Espinar (2002 y 2003 [38, 40]).

El modelo ajustado a los datos experimentales del índice WQI (figura 5.14 y tabla 5.7) tiene un alcance intermedio en comparación con los variogramas de los índices anteriores. El mapa de distribución espacial del índice WQI, sin embargo, presenta menos valores intermedios con zonas extremas de buena o mala calidad más extensas (figura 5.15).

Las funciones de probabilidad local se han calculado para algunos de los municipios más significativos de la Vega de Granada (5.16). Los índices de calidad se han reescalado en el intervalo [0-100], donde 0 es la peor calidad y 100 la mejor. Debido a los diferentes métodos de creación de los subíndices y diferentes métodos de agregación, observamos distintos comportamientos. Aún así, los índices Poch e ICG se comportan de una manera muy parecida, e incluso con los índices ID y WQI observamos bastantes similitudes. Por ejemplo, para el municipio de Santa Fé, la función de probabilidad local presenta distintos escalones que se repiten con un desfase en x para cada índice. Estos escalones muestran diferentes "saltos" en la calidad del agua y el desfase en x sólo representa que un índice es más estricto (próximo a cero) o menos estricto (próximo a 100).

Por lo tanto, podemos decir que, en general, a pesar de las variaciones en el valor del índice en los cuatro índices, el aumento de probabilidad con la calidad, se comporta de la misma manera, es decir, a grandes rasgos, la forma de las funciones de pertenencia son muy similares aunque desplazadas en x . Esta similitud es debida a que todos los índices están construidos con los mismos datos experimentales (variables) y los valores

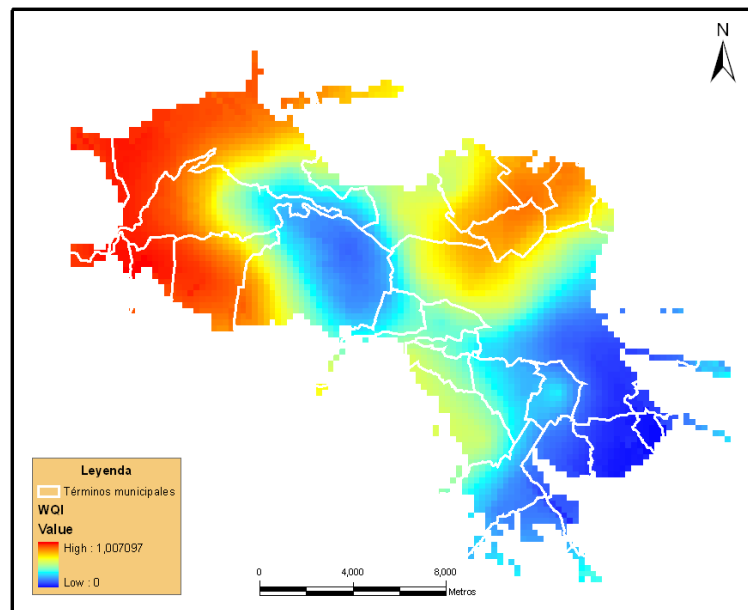


Figura 5.15: Mapa de isoprobabilidad (krigeaje de indicatrices) de un índice de calidad del agua construido mediante la agregación exponencial para un valor indicatriz igual a la mediana de la distribución.

límite son los mismos en todos los índices. Por tanto, a pesar de que el valor del índice absoluto pueda variar en su significado, la distribución espacial de la calidad del agua viene a ser la misma en términos relativos.

Obviamente, las pequeñas diferencias observadas a escala local son debidas a diferencias en los pesos asignados así como a la diferencia en la creación de los subíndices o al tipo de agregación de las variables en la construcción de los índices.

Si comparamos los mapas de probabilidad de calidad del agua en la mediana (figuras 5.11, 5.8 y 5.15) en todos los casos destacan las zonas de buena calidad en la parte oriental del acuífero, donde se está produciendo la recarga más importante a través de la red fluvial, principalmente el río Genil, y a lo largo del cual la calidad del agua va disminuyendo al mezclarse con aguas de borde de peor calidad. Estas zonas de peor calidad están ligadas, básicamente, a los materiales evaporíticos que afloran en esas áreas.

En resumen podemos concluir que los índices son una buena herramienta para la evaluación de la calidad del agua subterránea. No obstante, aunque en términos absolutos es importante la elección de un índice concreto para establecer una calidad y así evitar el eclipsamiento de valores no aptos en algunas variables del agua y la ambigüedad en los valores de calidad, en todos los casos es posible discernir las áreas de mejor calidad, así como un posible empeoramiento en el tiempo (tendencias, etc.). Por lo tanto, concluimos que un estudio de la calidad del agua en profundidad requiere de

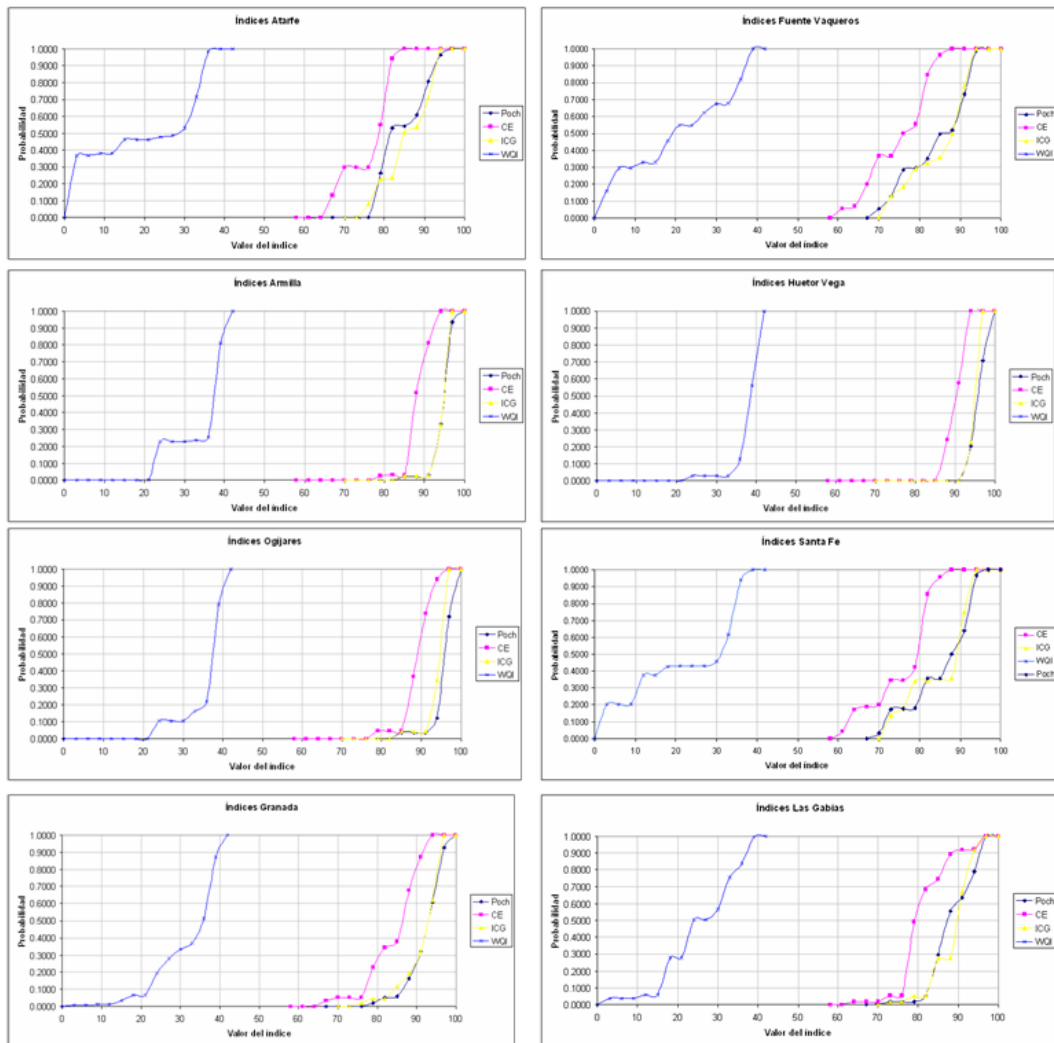


Figura 5.16: Funciones de distribución para seis de los municipios estudiados según cuatro índices de calidad.

herramientas como la función de probabilidad local que permite reconocer en detalle la variación de la calidad del agua en un área concreta (municipio en nuestro caso) y de este modo analizar los posibles problemas y posibles soluciones.

5.2.5. Aplicación de la Evaluación Multicriterio al estudio de la vulnerabilidad

La Evaluación Multicriterio (EMC) es una herramienta de soporte a la decisión muy utilizada en el estudio de problemas de evaluación en los que tienen cabida múltiples criterios. Una decisión es una elección entre alternativas (tales como alternativas de localización de un pozo, la localización de zonas vulnerables, etc.). La base de la decisión se conoce como criterio. En una evaluación multicriterio, se intenta combinar un grupo de criterios para conseguir una combinación sencilla conforme a un objetivo específico, por tanto sus posibilidades de aplicación en el estudio de los recursos hídricos subterráneos son amplias.

Los criterios pueden ser de dos tipos: factores y restricciones. Los factores son generalmente de naturaleza continua (tales como el espesor no saturado del acuífero, o la pendiente del terreno) e indican la aptitud relativa de ciertas áreas. Las restricciones, por otra parte, son siempre de carácter booleano y afectan a "objetos" (tales como espacios naturales protegidos, zonas industriales, etc.), estos sirven para excluir ciertas áreas de la consideración. Los factores y las restricciones son combinados en el análisis multicriterio mediante distintos métodos. En el caso de los criterios booleanos (restricciones), la solución del índice de evaluación suele ser la unión (lógica OR) o la intersección (lógica AND) de las condiciones. Sin embargo, para factores continuos, una combinación lineal ponderada (Voogd, 1983 [185]) es la más comúnmente usada. En una combinación lineal ponderada, los factores se combinan aplicando un peso a cada uno y, seguidamente, una sumatoria a los resultados para obtener el mapa objetivo.

La asignación de estos pesos se puede hacer por diversos métodos. El método escogido en este estudio consiste en la asignación de pesos a los factores de cada grupo, según el Proceso Analítico Jerárquico (Analytical Hierarchy Process o AHP) (Saaty, 1977 [160]). Este método de asignación de pesos da un enfoque más objetivo que la simple asignación de pesos del ejemplo anterior, ya que se aplica una combinación pareada en la cual sólo dos criterios se consideran a la vez. De este modo es más probable producir un grupo de pesos robusto (Eastman, 2003 [66]). La implementación de la técnica de la comparación pareada de factores fue desarrollada por Saaty (1977 [160]) en el contexto de un proceso de toma de decisiones conocido como AHP (Pavlikakis y Tsihrintzis, 2003 [142]). En el procedimiento de evaluación multicriterio que se muestra como ejemplo en este apartado, se ha usado una combinación lineal de pesos cuya suma es 1. En la técnica de Saaty (1977 [160]), los pesos de los factores pueden ser derivados a partir de los

autovalores principales de una matriz cuadrada recíproca de comparaciones pareadas entre criterios. Las comparaciones atañen a la importancia relativa de los dos criterios involucrados en determinar la aptitud para el objetivo indicado (Eastman, 2003 [66]). Un índice de consistencia (CR) (Saaty, 1977 [160]), se deduce de la matriz de comparación pareada que contiene múltiples caminos por los cuales la importancia relativa de los criterios puede ser evaluada. De este modo, se determina el grado de consistencia que ha sido usado en el desarrollo de los pesos. El CR representa la probabilidad de que los pesos de la matriz hayan sido aleatoriamente generados e indica que matrices con CR superiores a 0,1 deben ser reevaluadas. Además del ratio de consistencia total es posible analizar la matriz para determinar dónde están las mayores inconsistencias (Eastman, 2003 [66]).

Una aplicación más allá del típico AHP es la asignación de un segundo grupo de pesos, conocida como sumatoria lineal ponderada ordenada (Order Weights Average, OWA). En este caso, los pesos no se aplican a ningún factor específico, sino que son aplicados en una base píxel a píxel a las puntuaciones como se determina por su orden de rango a partir de cada localización (píxel). El método consiste en utilizar, además de los pesos de los factores, otra serie de pesos "ordenados" que controlan como se agregan esos factores ponderados. De esta manera es posible determinar el nivel total de compensación permitido.

Este método sería una generalización de tres tipos básicos de funciones de agregación: la intersección de conjuntos borrosos, la unión de conjuntos borrosos y los operadores de medias. Por lo tanto en los dos extremos podríamos elegir entre una técnica no compensatoria (utilización del operador mínimo o Y lógico) o una compensatoria (cualquiera de las otras posibilidades en las que se permite compensación entre los distintos factores).

Para aclarar este concepto de pesos de orden (Ordered Weighted Average o OWA), deben aclararse un par de conceptos: *compensación (tradeoff)* y *riesgo*.

- *Compensación (tradeoff)*: Los pesos de los factores son pesos que se aplican a factores específicos, es decir, todos los píxeles de un factor concreto reciben el mismo peso. Estos indican el grado relativo de importancia que cada factor juega en el resultado final.
- *Riesgo*: Las aproximaciones booleanas son funciones extremas cuyo resultado son soluciones de poco riesgo cuando se usa el operador AND o de elevado riesgo cuando se usa el operador OR. Con el operador AND una aptitud con elevada puntuación de agregación, para un píxel determinado, es sólo posible si todos los factores tienen altas puntuaciones. Para el operador OR, una alta puntuación en cualquier factor dará en una elevada puntuación de agregación, incluso si los otros factores tienen puntuaciones muy bajas. Así, el operador AND puede ser definido

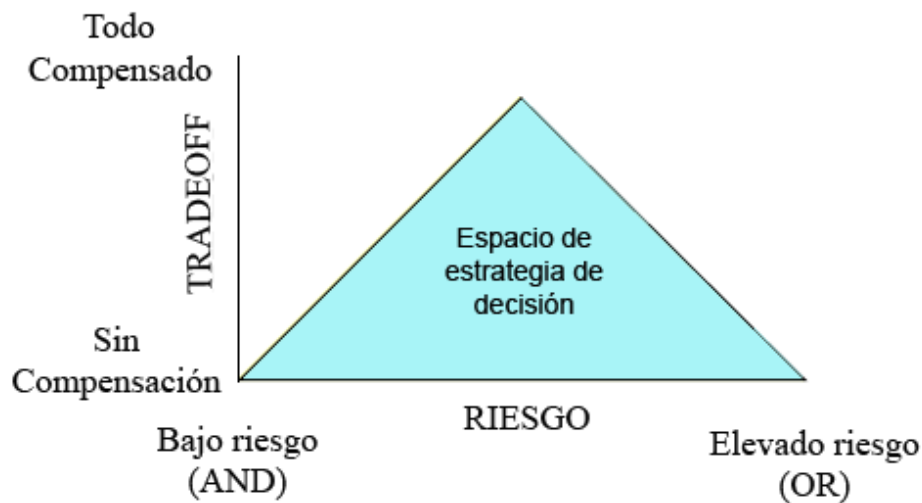


Figura 5.17: Triángulo de decisión.

como el mínimo (muy seguro: risk-averse) y el operador OR se puede definir como el máximo (alto riesgo: risk-taking).

Así, una aproximación de Combinación Lineal de Pesos (CLP) es una técnica de promedio que suaviza las duras decisiones de la aproximación booleana y evita los extremos. De hecho, dado un riesgo continuo de un mínimo a un máximo, la CLP cae justo en el medio, no siendo ni muy segura ni muy arriesgada.

El uso de los pesos ordenados permite soluciones de agregación que caen en el continuo de riesgo entre AND y OR (figura 5.17). Los pesos de orden se aplican en una base píxel a píxel a puntuaciones de factores, como determina su orden de rango para cada localización (píxel). Así, el peso de orden 1 se asigna al factor con el peso más bajo para ese píxel (i.e., al factor con la mínima puntuación), el peso de orden 2 al siguiente factor de mayor peso y así sucesivamente.

5.2.5.1. Evaluación de la vulnerabilidad en la Vega de Granada

La evaluación de la vulnerabilidad del acuífero de la Vega de Granada es un aspecto de gran interés en el análisis y gestión de la calidad de las aguas subterráneas del acuífero. Las áreas con mayor vulnerabilidad precisan mayor atención en las políticas de protección. El término de vulnerabilidad de un acuífero incluye dos aspectos básicos: la vulnerabilidad intrínseca y la específica (Gogu y Dassargues, 2000 [81]). El primero define la vulnerabilidad del acuífero generada por actividades humanas, teniendo en cuenta únicamente los parámetros hidrogeológicos del mismo, siendo independiente de

la naturaleza de los contaminantes. El segundo es específico para un grupo particular de contaminantes.

Hasta ahora han sido diversos los métodos desarrollados para evaluar la vulnerabilidad intrínseca de un acuífero. Algunos de los más comunes son el DRASTIC (Aller et al. 1987 [7]), el GOD (Foster, 1987 [74]), el AVI (Van Stempvoort et al., 1993[?]), el SINTACS ([44]) y el EPIK ([61]), entre otros. Gogu et al. (2003 [82]) realizan una comparación de algunos de los métodos mencionados arriba mostrando la amplia variedad en los resultados obtenidos por cada método, y en muchos casos el desacuerdo entre los resultados. La causa de estas diferencias es que la vulnerabilidad no es una medida experimental cuantificable, por lo que la selección entre los métodos es una tarea bastante ambigua.

Por este motivo, en este trabajo planteamos el uso conjunto de los Sistemas de Información Geográfica con métodos de lógica difusa y de evaluación multicriterio como herramienta útil para estudiar el problema de la diversidad de metodologías en la evaluación de la vulnerabilidad. Este interés está en la normalización de las variables que se incluyen en la mayoría de estos métodos y en la realización de una asignación de pesos consistente.

El acuífero de la Vega de Granada se ha discretizado usando un tamaño de celdilla de 250 m x 250 m. Inicialmente, todos los factores se han estandarizado a un rango de 0-255 (nivel byte) el cual provee la máxima diferenciación posible cuando se analizan los datos en tipo byte. De este modo el 0 se le asigna a las áreas menos vulnerables y el 255 a las áreas más vulnerables, transformando las diferentes unidades de medida de las capas SIG, en valores comparables utilizando las funciones de pertenencia difusas (apartado 5.2.3).

Para calcular la vulnerabilidad del acuífero se han escogido 6, de las 7 variables, que incluye el método DRASTIC (Aller et al., 1987 [7]) (tabla 5.8). La variable que no se ha utilizado es la naturaleza del acuífero (A), debido a que a la escala de trabajo esta variable es homogénea y no aporta información. El objetivo final de la aplicación de esta metodología no es la comparación en términos absolutos con otros acuíferos, sino la identificación de las diferencias de vulnerabilidad en el mismo acuífero

El cálculo de los pesos se ha asignado en función de la importancia de cada factor, y para que esta comparación sea más objetiva se ha utilizado el método de las comparaciones pareadas, en el cual sólo dos factores se consideran a la vez. De este modo es más sencillo producir un conjunto de pesos robustos (Eastman, 2003 [66]).

La implementación de la técnica de la comparación por parejas fue desarrollada por Saaty (1977 [160]) en el contexto del proceso de la decisión conocido como Proceso Analítico Jerárquico (Analytical Hierarchy Process o AHP). El proceso de la EMC de este estudio, que usa una combinación lineal de pesos, requiere que los pesos asignados a

Variable	Origen	Estimación
<i>D</i> – Espesor no saturado	MDT – Resolución a 10 m. Campaña piezométrica IGME en 1980 en 183 puntos	Geoestadística
<i>R</i> – Recarga neta	Datos precipitación INM 1980 en 26 estaciones. Datos Temperatura INM en 13 estaciones	Geoestadística
<i>S</i> – Tipo de suelo	Campaña de muestreo de suelos en 76 puntos	Geoestadística
<i>T</i> – Topografía	MDT – Resolución a 10 m (mapa de pendientes)	MDT
<i>I</i> – Impacto de la zona no saturada	Columnas litológicas de sondeos en 220 puntos	Geoestadística
<i>C</i> – Conductividad Hidráulica o permeabilidad	Ensayos de bombeo en 45 puntos	Geoestadística

Tabla 5.8: Resumen de los datos de origen utilizados para calcular la vulnerabilidad.

los factores sumen uno. En la técnica desarrollada por Saaty, los pesos de esta naturaleza derivan del autovector dominante de una matriz cuadrada recíproca de comparaciones binarias de los criterios. Estas comparaciones entre los pares de criterios establecerán la importancia relativa que un criterio tiene sobre su par, en una escala continua del 1 al 9, donde 1 representa igualdad de importancia y 9 representa la máxima importancia relativa de una propiedad sobre otra.

Igualmente se ha determinado el índice de consistencia (Consistency Ratio o CR) usado al desarrollar los pesos. El CR indica la probabilidad de que las puntuaciones de la matriz hayan sido generadas aleatoriamente, e indica que matrices con CR mayores de 0.10 deben ser reevaluadas; además, permite analizar la matriz para determinar dónde aumentan las inconsistencias (Eastman 2003 [66]).

En este estudio, también se ha aplicado el procedimiento conocido como "pesos ordenados" (OWA) que ha sido aplicado en una base de píxel a píxel. Este método da la posibilidad de ajustar el cálculo de acuerdo al nivel de riesgo de la decisión, es decir, en la cantidad en la que los valores de los píxeles se dirigen hacia los valores mínimos o máximos de vulnerabilidad. Tomando como ejemplo algunas de las variables escogidas, la asignación de pesos ordenados resultaría como se muestra en la tabla 5.9 en la que se identifican las diferentes situaciones de riesgo.

En el primer caso, los pesos se distribuyen del mismo modo para todos los factores, independientemente de su posición de orden. El resultado se sitúa exactamente en medio en términos de riesgo. En el segundo caso, los factores con la mínima puntuación, es

Factores	Topografía	Impacto de la zona vadosa	Espesor no saturado
<i>Nivel moderado de riesgo - Decisión de riesgo moderada</i>			
Pesos Ordenados	0.33	0.33	0.33
Orden	1 ^o	2 ^o	3 ^o
<i>Nivel moderado de riesgo - Decisión no estricta</i>			
Pesos Ordenados	1	0	0
Orden	1 ^o	2 ^o	3 ^o
<i>Nivel moderado de riesgo - Decisión estricta</i>			
Pesos Ordenados	0	0	1
Orden	1 ^o	2 ^o	3 ^o

Tabla 5.9: Ejemplo de asignación de pesos de orden (OWA).

decir, los menos vulnerables, reciben todo el peso del factor de pesos asignado. Este resultado incorpora un alto nivel de riesgo, ya que a todos los píxeles del área de estudio se les asigna el valor del factor menos vulnerable. Por el contrario, en el tercer caso se le asigna todo el peso del factor con la máxima puntuación para cada píxel, resultando un valor de píxel igual al factor con la máxima vulnerabilidad. Este es el caso más conservativo ya que todos los valores asignados en el mapa de vulnerabilidad se toman del factor de máxima vulnerabilidad. Así, el sesgo de los pesos ordenados hacia factores con máxima o mínima puntuación resultan en todo el abanico de posibilidades entre riesgo y severidad.

La aplicación de las funciones difusas a cada una de las variables se comenta a continuación:

Factor D: Espesor no saturado

El cálculo del espesor no saturado del acuífero se ha realizado a partir de las medidas de piezometría de la campaña más extensa disponible, correspondiente a abril de 1980, que coincide con un año de precipitación media para el área de estudio (figura 2.9). A partir de estos datos se ha creado una cubierta continua de la piezometría mediante geoestadística (figura 5.18 y tabla 5.10). El espesor no saturado se ha obtenido como diferencia entre el modelo digital del terreno de Andalucía de las Consejerías de Obras Públicas y Transportes, Agricultura y Pesca, y Medio Ambiente de 2005 (generado a partir de fotografías aéreas a escala 1:20:000) y la cubierta estimada mediante krigeaje de la piezometría. Así, una función sigmoideal monótonamente decreciente se ha aplicado para transformar los valores del espesor no saturado a un rango de valores de 0 a 255. Los puntos de control de la función para esta variable son: 10, dónde la función toma valor 1 (es decir, a 10 m de distancia del agua la función toma el valor máximo) y 150 donde la función toma el valor de 0 (a más de 150 m de distancia al agua, la función

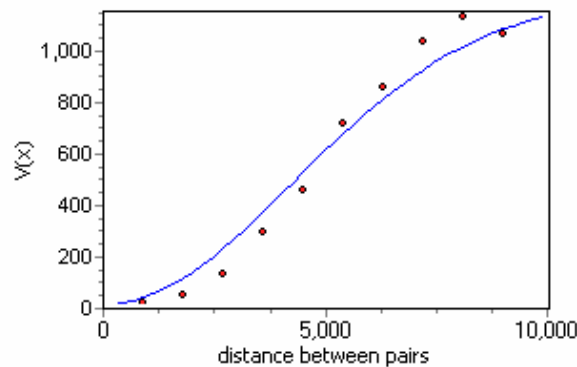


Figura 5.18: Variograma experimental y teórico de la piezometría de abril de 1980.

Variable	Modelo	Paso	Meseta (C)	Alcance (A) Km	Efecto de pepita (Co)
Piezometría	Gausiano	1200	1200	6.0	15

Tabla 5.10: Parámetros del modelo teórico del ajuste del variograma de la piezometría de abril de 1980.

toma el valor mínimo). Así, valores inferiores a 10 reciben el mismo valor de pertenencia de la función que 1 y valores superiores a 150 el mismo que 0 (figura 5.19).

Factor R: Recarga

Como se ha comentado en el primer capítulo, la recarga del acuífero se produce fundamentalmente por los ríos de cabecera que proceden desde su amplia cuenca vertiente, así como de las infiltraciones de lluvia y riego a través de la densa red de acequias. Esta recarga no es en absoluto despreciable pero sí limitada a los cursos de los ríos y a las acequias y a las zonas de riego por regadío. Debo aclarar que el análisis de la recarga en este caso de estudio no pretende ser un análisis cuantitativo, ya que para esta aproximación sólo interesa la cantidad de agua que puede infiltrarse al acuífero por unidad de píxel en términos relativos, ya que finalmente el mapa final es una normalización. Este tipo de simplificaciones son comúnmente utilizadas en la evaluación SIG de la vulnerabilidad (Uddameri, V. and Honnunar, V., 2007 ; Al-Zabet, T., 2002 y Gemitzy, A. et al., 2006 [77])

De esta manera, las zonas de regadío se han identificado en la imagen de satélite clasificada (3.20). Como en el caso de las áreas de regadío tenemos una capa de regadío o no regadío, asignamos un valor arbitrario de 125 (vulnerabilidad intermedia).

Como siguiente factor en importancia está la recarga de la precipitación. Así, cada píxel del área de estudio recibe una cantidad de lluvia (figura 2.9). La precipitación

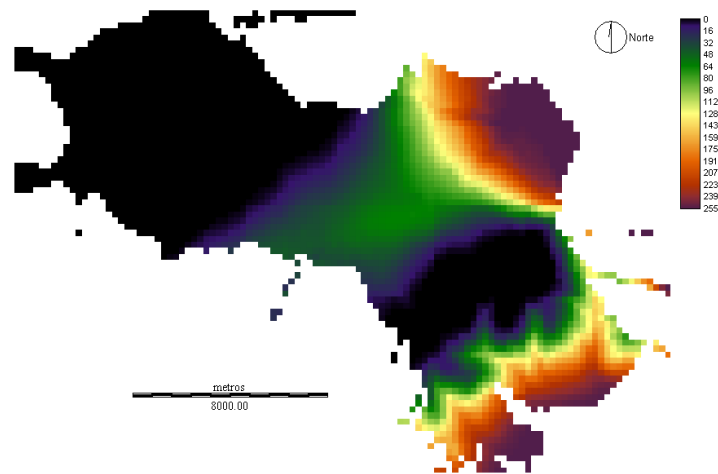


Figura 5.19: Mapa del espesor no saturado normalizado mediante una función de pertenencia.

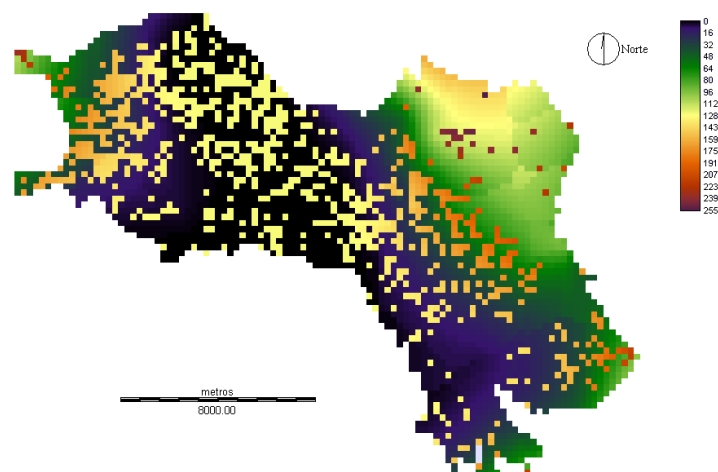


Figura 5.20: Mapa de la recarga normalizado mediante una función de pertenencia.

media del área de estudio para el año 1980 (año medio) es 367 mm y la desviación estándar es de 92 mm. Estableciendo la condición de que la máxima vulnerabilidad se da para valores superiores a la media más una desviación estándar y la mínima en la media menos una desviación estándar, establecemos los puntos de control para una función monótonamente creciente en 274 y 459.

Para incluir una única imagen, simplificamos el factor de recarga con la suma de los dos mapas obtenidas, ya que la recarga total será la suma de la recarga de precipitación más la recarga por infiltración de regadío. Al realizar la suma los valores máximos se salen del rango 0-255, por lo que asignamos un valor de 255 a los valores superiores al rango, es decir, máxima vulnerabilidad (figura 5.20).

Factor S: Tipo de suelo

Esta variable es importante en la cantidad de agua infiltrada que llega al acuífero y en el movimiento del contaminante en la vertical a través de la zona no saturada. Cada suelo presenta una vulnerabilidad diferente a la contaminación, dependiente del tipo de arcilla que lo conforma, su capacidad de hinchamiento, el tamaño de grano que presenta y la materia orgánica del mismo. De manera aproximada se puede aceptar que a mayor capacidad de hinchamiento y menor tamaño de grano, menor será la vulnerabilidad.

Los suelos de la Vega de Granada están asociados principalmente a depósitos aluviales del río Genil, por lo que su composición es relativamente homogénea, así que la mayor o menor vulnerabilidad en relación a este parámetro estará principalmente ligada a la cantidad de arcillas que contenga el suelo y que actúan como barrera impermeable. De este modo, asignamos una vulnerabilidad baja con un valor arbitrario de 25 a las zonas de textura arcillosa, una vulnerabilidad también baja a los suelos franco limosos con un valor de 75, una vulnerabilidad alta (200) a las zonas de textura franco-arcillo-arenosa, y una vulnerabilidad intermedia (125) al resto de texturas (figura 2.50).

Factor T: Topografía

La topografía de la zona de estudio tiene incidencia en el arrastre del contaminante por la escorrentía superficial o infiltración y en el desarrollo del suelo. Como regla general, cuanto mayor sea la pendiente de la zona, la vulnerabilidad será menor. En la Vega de Granada este parámetro realmente no es demasiado importante, ya que excepto los bordes, se trata de una depresión y por tanto con un relieve muy suave. La pendiente máxima del área de estudio es de 13 %, por lo que cabe comentar que la mayor parte del área tiene una baja pendiente. La función de pertenencia monótonamente decreciente asignada tiene como puntos de control $c=0$ y $d=18$ de acuerdo con las directrices indicadas por Martínez Navarrete y García García (2003 [119]) (figura 5.21).

Factor I: Impacto de la zona no saturada

Del mismo modo que el suelo, la zona no saturada interviene en los procesos de atenuación de los contaminantes, ya sea producida por biodegradación, por volatilización, o por reacciones químicas, en el transporte del contaminante y su tiempo de permanencia en la zona y en el desarrollo del suelo. El factor I viene a reflejar el tamaño de grano, el grado de fracturación, y la clasificación granulométrica entre otras características. En la Vega de Granada la litología de la zona no saturada, igual que ocurre con los suelos, tiene una composición relativamente homogénea, por lo que el factor más influyente es el tamaño de grano. El tamaño de grano, normalizado a un valor categórico de 6 y ponderado por el espesor de cada una de las columnas, se ha representado en la

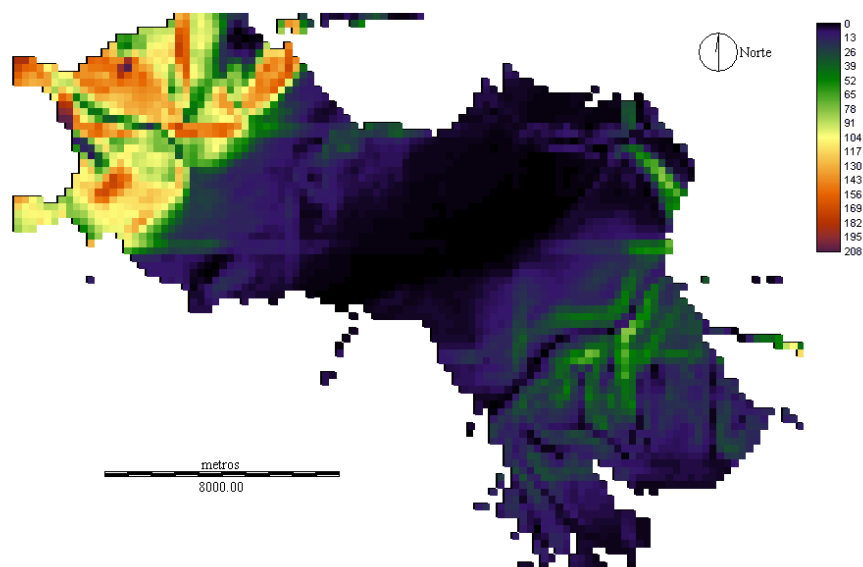


Figura 5.21: Mapa de la topografía reclasificado mediante una función de pertenencia monótonamente decreciente.

figura 2.5. A priori, un menor tamaño de grano significará una menor vulnerabilidad y viceversa, por lo que se asignan los valores de $a=0$ y $b=6$ a una función de pertenencia monótonamente creciente (figura 5.22).

Factor C: Conductividad hidráulica

La conductividad hidráulica condiciona la velocidad de movimiento del agua y por tanto la del contaminante, así como la renovación del agua en el interior del acuífero. De este modo, la vulnerabilidad será mayor cuanto mayor sea la conductividad hidráulica. Esta variable se ha estimado geoestadísticamente a partir de 46 valores de ensayos de bombeo facilitados por el IGME (figura 2.15). Los parámetros de control escogidos para esta variable son $c=10$ (máxima vulnerabilidad) y $d=100$ (mínima vulnerabilidad) para una función sigmoideal decreciente (figura 5.23).

5.2.5.2. Creación del mapa de vulnerabilidad

Una vez estandarizados todos los factores a un intervalo común de 0-255 por medio de funciones de pertenencia difusas, se han asignado los pesos a cada factor. Los pesos indican la importancia relativa del factor con respecto a los demás y controlan como los factores se compensan entre sí. En otras palabras, el grado en el cual un factor es compensado por otro se determina a partir del peso de esos factores. En el caso de la determinación de la vulnerabilidad de las aguas subterráneas a la contaminación, los

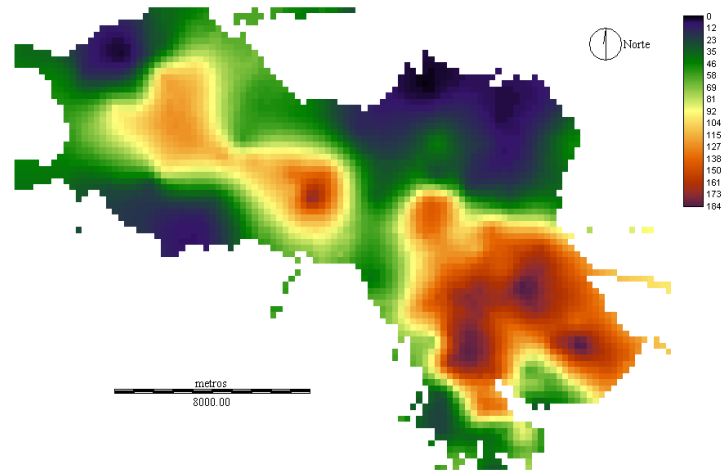


Figura 5.22: Mapa de la litología de la zona no saturada reclasificado mediante una función de pertenencia monótonamente creciente.

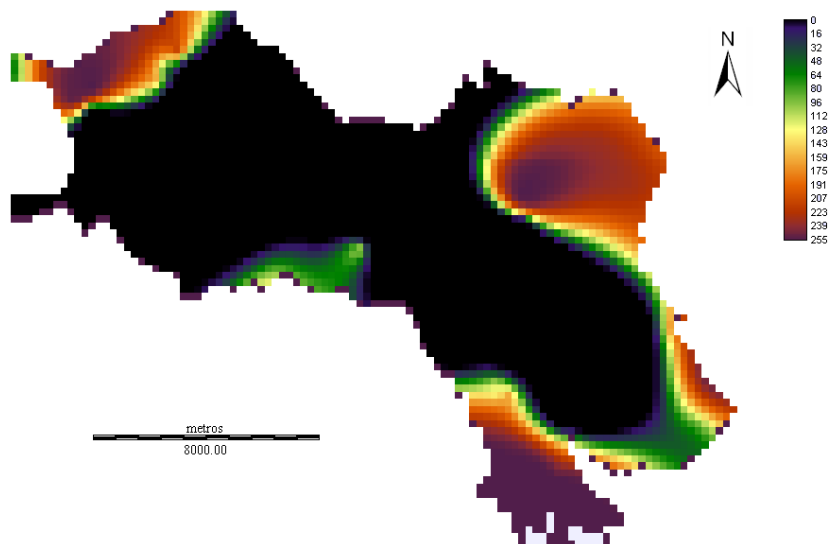


Figura 5.23: Mapa de la conductividad hidráulica reclasificado mediante una función de pertenencia monótonamente decreciente.

factores con baja vulnerabilidad a la contaminación en una localización dada, pueden ser compensados por otros factores con una alta vulnerabilidad en la misma localización.

En todos los casos los pesos de los factores deben sumar 1. Existen distintas técnicas para asignar los pesos de los factores. Una de las técnicas más eficientes para producir los pesos de los factores es el AHP (Saaty, 1977 y Pavlikakis y Tsihrintzis, 2000 [160, 142]). El método AHP ha sido implementado en el SIG de modo que cada par de factores es examinado al mismo tiempo, en términos de importancia relativa (tabla 5.11). De todas las posibles combinaciones de los dos factores, el módulo de cálculo proporciona un grupo de pesos que suma 1 y el coeficiente CR (consistency ratio). Este ratio indica algunas inconsistencias que pueden haber sido realizadas durante el proceso de comparación pareada, es decir, la probabilidad de que el peso del factor haya sido asignado aleatoriamente (Eastman 2003 [66]). Un CR mayor que 0.10 indica que los pesos de los factores deben ser reevaluados (Saaty 1977 [160]).

Una vez se han aplicado los pesos a los factores originales, los resultados son ordenados de baja a alta vulnerabilidad del acuífero a la contaminación para cada localización. Al factor con la puntuación mínima de vulnerabilidad se le asigna el primer peso de orden, al factor con la siguiente puntuación más alta se le asigna el siguiente peso de orden y así con todos los demás. Esto tiene el efecto de basar los pesos de los factores para cada localización en su rango de valores mínimo (baja vulnerabilidad) a máximo (alta vulnerabilidad). El sesgo relativo de mínimo o máximo de los pesos controla el nivel de riesgo de la evaluación (Eastman 2003 [66]). Además, el grado en el cual los pesos de orden son equitativamente distribuidos en todas las posiciones controla el grado en el cual los pesos de los factores influyen.

Los pesos ordenados (tabla 5.12) se asignan a cada uno de los factores. Los pesos asignados muestran un nivel moderado de riesgo y una decisión estricta (los tres factores con la máxima puntuación comparten el 70 % del peso, mientras los demás están pesados equitativamente). La vulnerabilidad evaluada mediante esta asignación ordenada de pesos (figura 5.24) presenta tres zonas de máxima vulnerabilidad situadas en los bordes (borde noreste, noroeste y sureste) y las zona de mínima vulnerabilidad se localiza en el área del río Salado. Utilizando los mismos pesos, pero agregando los factores mediante una combinación lineal de pesos (Weighted Linear Combination) el mapa resultado (figura 5.25) es menos estricto y el área de vulnerabilidad alta localizada en el borde noroeste desaparece. Comparando la distribución de los contenidos en nitratos del agua subterránea (figura 5.26) con los mapas de vulnerabilidad (figuras 5.24 y 5.25) observamos un mejor ajuste mediante el método de los pesos ordenados.

A pesar de este primer ajuste de tipo visual, la realidad es que no encontramos una alta correlación entre las zonas de alta vulnerabilidad con las zonas de elevado contenido en nitratos, debido a que aunque existan áreas con elevada vulnerabilidad, no existe

	<i>Espesor no saturado</i>	<i>Topografía</i>	<i>Permeabilidad</i>	<i>Recarga</i>	<i>Tipo de suelo</i>	<i>Impacto ZNS</i>	<i>Pesos</i>
<i>Espesor no saturado</i>	1						0.415
<i>Topografía</i>	1/9	1					0.0366
<i>Permeabilidad</i>	1/3	7	1				0.1928
<i>Recarga</i>	1/9	1/3	1/7	1			0.0255
<i>Tipo de suelo</i>	1/5	5	1/3	3	1		0.0802
<i>Impacto ZNS</i>	1/3	7	1	9	7	1	0.2499

Tabla 5.11: Matriz de comparación pareada para la asignación de pesos de vulnerabilidad.

una contaminación por nitratos. Debemos tener en cuenta que la vulnerabilidad a la contaminación es una propiedad relativa y por lo tanto no medible (Gogu y Dassargues, 2000 [81]), por lo que estos mapas sólo muestran la probabilidad o el riesgo de que la contaminación llegue a las aguas subterráneas tras la introducción en la superficie del contaminante, por lo que hay que tener precaución al comparar la contaminación observada con los mapas de vulnerabilidad.

A pesar de las diferencias de los mapas obtenidos, los métodos presentados suponen una mejora a los métodos clásicos de evaluación de la vulnerabilidad (DRASTIC, EPIK, GOD, etc.). El método aquí presentado permite una mayor diferenciación para cada una de las variables estableciendo una función lineal (no intervalos discretos) como método de transformación de los factores. Además, el método AHP permite establecer los pesos de un modo más objetivo mediante una comparación pareada y la agregación mediante los pesos ordenados (OWA) define el nivel de compensación preferido para la agregación de las variables.

No obstante, no hay que olvidar que muchas de las inexactitudes y problemas de estos métodos pueden ser solventados con modelos de flujo de aguas subterráneas y de transporte, aunque el conjunto de técnicas es probablemente la mejor manera de obtener un conocimiento completo del acuífero, ya que este tipo de métodos es muy útil para establecer los posibles asentamientos de actividades humanas, desarrollo urbano y agrario y gestión del agua.

<i>Pesos ordenados</i>	0.1	0.1	0.1	0.15	0.2	0.35
<i>Orden</i>	1 ^o	2 ^o	3 ^o	4 ^o	5 ^o	6 ^o

Tabla 5.12: Pesos ordenados para la evaluación de la vulnerabilidad.

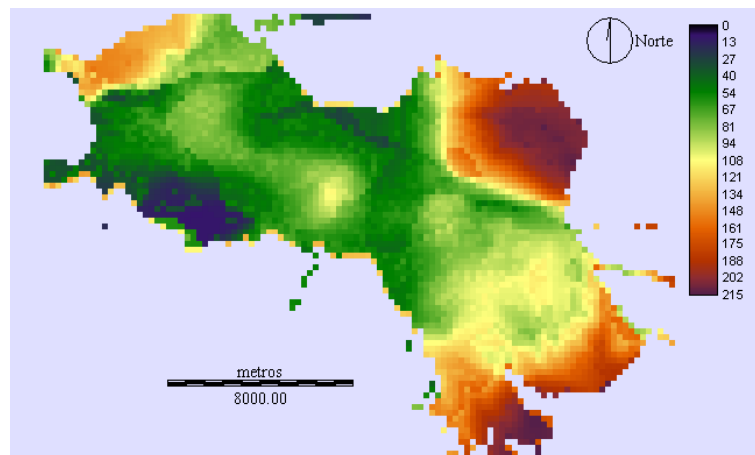


Figura 5.24: Vulnerabilidad del acuífero evaluada mediante OWA.

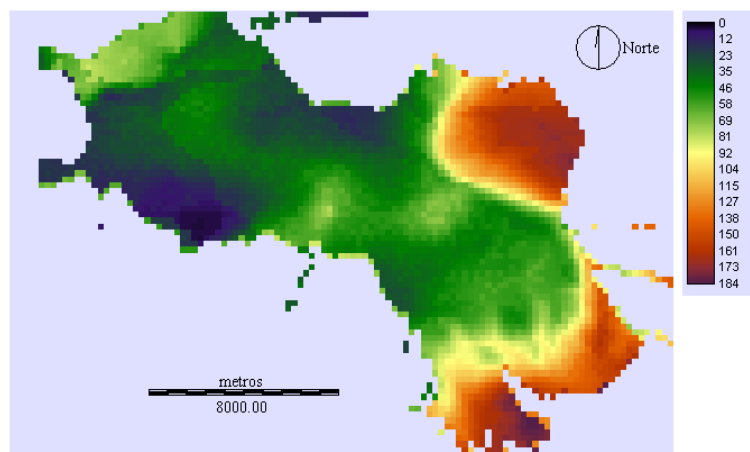


Figura 5.25: Vulnerabilidad del acuífero evaluada mediante una combinación lineal de pesos (WLC).

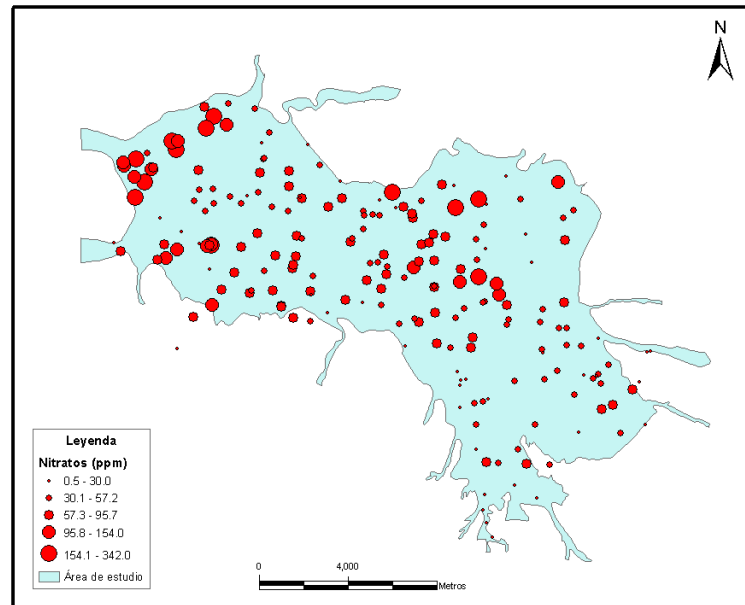


Figura 5.26: Distribución espacial del contenido en nitratos del agua subterránea.

5.3. Integración SIG mediante modelos basados en los datos

Los acuíferos son sistemas naturales complejos en los que las hipótesis utilizadas por los métodos de tratamiento de la información hidrogeológica y ambiental clásicos, como son los métodos estadísticos, son a menudo demasiado simples. Resulta por tanto necesario explorar nuevas posibilidades de análisis y modelización espacial que tengan en cuenta estas peculiaridades. En este sentido, ciertas metodologías provenientes de la Inteligencia Artificial, tales como las Redes Neuronales Artificiales y la Lógica Difusa, están recibiendo una gran atención por parte de la comunidad geocientífica.

5.3.1. Estimación de la calidad del agua subterránea mediante redes neuronales artificiales

Las Redes Neuronales Artificiales (RNA) presentan, en concreto, una serie de ventajas importantes sobre las técnicas de análisis espacial de datos hidrogeológicos tradicionales, tales como la ausencia de supuestos sobre los datos espaciales, su capacidad de procesar datos ruidosos e interacciones entre factores, así como de modelizar relaciones no lineales. Las RNA son sistemas implementados normalmente como programas de software diseñados para modelizar o simular la actividad de sistemas neuronales biológicos (Haykin, 1994 [90]). Aunque las RNA están inspiradas en organismos vivos, la mayor parte de las arquitecturas de redes neuronales no tratan de imitar fidedignamente

su modelo biológico, sino que más bien pueden ser consideradas simplemente como una clase de algoritmos paralelos (Bishop, 1995 [19]). Las RNA son muy utilizadas en la actualidad en diversos campos de la ciencia debido a las propiedades indicadas arriba, que las hacen muy atractivas, y a su simplicidad conceptual a la hora de aplicarlas. Las tareas que realizan las RNA se pueden englobar bajo el nombre genérico de “reconocimiento de patrones”, término que incluye a técnicas matemático-estadísticas más clásicas y conocidas de regresión y clasificación como el análisis discriminante (similar a la regresión logística).

En esta investigación se utiliza el tipo de red conocido como de retropropagación, que es el más extendido y utilizado en aplicaciones científico-técnicas.

La construcción de la red neuronal artificial (RNA) ha requerido tres pasos principales (Haykin, 1994):

1. Selección de la arquitectura.
2. Aprendizaje (entrenamiento).
3. Evaluación de la habilidad de generalización.

El límite entre ellos es, no obstante, difuso puesto que, por ejemplo, para seleccionar la arquitectura adecuada hace falta entrenar la red en cuestión. En la selección de la arquitectura el objetivo es encontrar una arquitectura de red (número de unidades, de capas ocultas, etc.) que produzca buenas predicciones tanto para datos observados como nuevos, por lo que la aproximación más simple para la comparación de distintas redes es evaluar el error, utilizando datos que son independientes de los usados para el entrenamiento. Una herramienta estándar en estadística, conocida como validación cruzada es frecuentemente usada como guía para este propósito (Bishop, 1995 [19]). Otro método muy utilizado es el denominado hold-out (Bishop, 1995 [19]), en el que el conjunto de datos disponible se particiona aleatoriamente en dos subconjuntos, uno para la selección y entrenamiento del modelo y otro para la evaluación de su funcionamiento (test). En la fase de entrenamiento los pesos de la red son ajustados mediante el proceso ya descrito de retropropagación, con el objetivo de minimizar una medida de la diferencia entre la salida real de la red y la salida deseada (Rumelhart et al., 1986 [159]). Finalmente, el modelo seleccionado es entrenado utilizando el conjunto completo de datos, y se mide la capacidad de generalización de la red resultante usando el conjunto de datos test.

Para la creación del modelo de predicción de la calidad del agua en el acuífero de la Vega de Granada se utilizó un conjunto de 206 patrones (datos) correspondientes a observaciones en pozos. En este caso el objetivo planteado es aprovechar la capacidad de las redes neuronales para el ajuste de funciones con el fin de estimar la función posiblemente no lineal que relaciona las variables fisicoquímicas del agua, las características

del suelo y del acuífero y la geología con la calidad del agua subterránea. Esta aproximación puede ser considerada como un modelo de regresión pero sin supuestos sobre los datos (normalidad, independencia) ni de linealidad de la relación. El valor a estimar consistió en la aptitud del agua calculada a partir de todas las variables disponibles, es decir, el conjunto más extenso de variables experimentales del área de estudio. Este conjunto de variables comprende el total de análisis fisicoquímicos realizados para el agua subterránea, cuyo uso es limitado por la normativa española RD 140/2003. El valor 0 se ha asignado en el caso de aguas no potables y 1 en aguas potables. Los datos de entrada del modelo son un subconjunto de datos reducido de fácil adquisición.

En un primer experimento se utilizaron como variables de entrada a la red un subconjunto de 42 variables del total de las más de 100 disponibles (Tabla 3.1) para estimar la aptitud o calidad del agua subterránea. Las variables usadas para la estimación de la aptitud del agua del acuífero (tabla 5.13) son: las coordenadas del punto x e y ; la conductividad, el pH y la temperatura del agua, fácilmente medibles en campo y de bajo coste; los parámetros físicos y químicos de los suelos; los parámetros hidrodinámicos del acuífero, excluyendo la piezometría porque se considera redundante ya que el espesor no saturado y el saturado nos aporta esa información; y las distancias a distintos focos de contaminación, así como a los afloramientos de materiales yesíferos. Como se aprecia se incluyen variables que están correlacionadas entre sí y variables que presentan distribuciones que no son gaussianas.

El modelo de RNA construido tiene una capa de entrada, una capa intermedia y una de salida (aptitud), por lo que dispone de 42 unidades en la capa de entrada y una única neurona en la capa de salida (el número de unidades en la capa intermedia es un parámetro a determinar en la selección del modelo).

En un segundo experimento se utilizaron solamente 5 variables de entrada caracterizadas por ser fácilmente obtenibles en la boca de pozo (conductividad, pH, temperatura, coordenada X , coordenada Y). La idea subyacente era tratar de estimar la calidad del agua (una calidad calculada utilizando un elevado número de variables, muchas de ellas de obtención compleja o costosa) utilizando sólo unas pocas variables que pueden ser medidas fácilmente, de forma económica y rápida, in situ. En este caso el modelo a construir tiene 5 unidades en la capa de entrada y una en la de salida (aptitud). Para ajustar el modelo se utilizó la variante con momento del algoritmo estándar de retropropagación disponible en el paquete de simulación SNNS. En todos los ensayos la función de activación utilizada fue la función tangente hiperbólica (tabla 3.3), y el parámetro de aprendizaje establecido a 0.01 (véase Bishop, 1995 [19], para una descripción detallada). Los procesos de entrenamiento se repitieron 100 veces y se hizo una media sobre los resultados para eliminar oscilaciones debidas a diferentes inicializaciones aleatorias de los pesos. Puesto que el número de unidades de entrada y salida del modelo estaban

<i>Variable</i>	<i>Descripción</i>
NORTE	Coordenada Y del punto
ESTE	Coordenda X del punto
CONDUCTIVI	Conductividad del agua subterránea
pH	pH del agua subterránea
Temperatura	Temperatura del agua subterránea
Ca_suelo	Calcio en el último horizonte de suelo
Cinc_suelo	Cinc en el último horizonte de suelo
cloro_suelo	Cloro en el último horizonte de suelo
Cobre_suelo	Cobre en el último horizonte de suelo
Conductividad_suelo	Conductividad media de los horizontes del suelo
Cromo_suelo	Cromo en el último horizonte de suelo
Potasio_suelo	Potasio en el último horizonte de suelo
magnesio_suelo	Magnesio en el último horizonte de suelo
Manganeso_suelo	Manganeso en el último horizonte de suelo
Nitratos_suelo	Nitratos en el último horizonte de suelo
Nitritos_suelo	Nitritos en el último horizonte de suelo
sodio_suelo	Sodio en el último horizonte de suelo
sulfatos_suelo	Sulfatos en el último horizonte de suelo
Transmisividad	Transmisividad del acuífero en ese punto
Permeabilidad	Permeabilidad del acuífero en ese punto
Espesor_no_saturado	Espesor no saturado del acuífero
Pendiente%	Pendiente en ese píxel
Arcillas%	Porcentaje de arcillas del suelo
Arenas%	Porcentaje de arenas del suelo
pH_suelo	pH de los horizontes medios del suelo
% Humedad_suelo	Porcentaje de humedad del suelo
C_org_suelo	Carbono orgánico de los horizontes medios del suelo
H. Saturac_suelo	Humedad de saturación del último horizonte del suelo
alt_techo	Cota del sustrato del acuífero
Grad90	Gradiente hidráulico en la dirección 90º
Grad0	Gradiente hidráulico en la dirección 0º
D_yesos	Distancia a yesos
D_rios	Distancia a ríos
D_pob	Distancia a poblaciones
D_aru	Distancia a focos de Aguas Residuales Urbanas
D_edar	Distancia a Estación Depuradora de Aguas Residuales
D_ganad	Distancia a granjas de ganado
D_indust	Distancia a industrias
D_minas	Distancia a minas
D_rsu	Distancia a focos de Residuos Sólidos Urbanos
D_cementerio	Distancia a cementerios
D_aceq	Distancia a acequias
Variable de salida	
Aptitud_todas_variables	Aptitud del agua subterránea para consumo público

Tabla 5.13: Variables utilizadas en el modelo de estimación de la calidad del agua subterránea.

definidos, la selección de la arquitectura consistió en la selección del número adecuado de unidades en la capa oculta.

En esta investigación se utilizó una variante del método de validación cruzada denominada k-fold (Bishop, 1995 [19]) especialmente diseñada para situaciones en las que el número de patrones (muestras) es escaso. En la validación cruzada tipo k-fold el conjunto de datos disponible se divide en dos conjuntos: entrenamiento y test. Se utilizó una partición del 90 % para el conjunto de entrenamiento y 10 % para el conjunto test.

El conjunto de entrenamiento se subdivide a su vez en dos subconjuntos: (a) un subconjunto para la estimación del modelo (es decir, entrenamiento de la red mediante el ajuste de los pesos), y (b) un subconjunto para la validación del rendimiento del modelo (típicamente el 10 %). Se tienen por tanto tres conjuntos denominados típicamente de entrenamiento, de validación y test.

El proceso de selección de la arquitectura más adecuada consiste en entrenar distintas redes con diversas configuraciones (números de unidades) minimizando el error del subconjunto de selección mediante retropropagación. En el proceso de entrenamiento se va presentando también a la red el conjunto de validación, sin que se modifiquen los pesos (sin entrenamiento), y se observa el error que produce la red para este conjunto. Este error se denomina error de validación puesto que es el error obtenido para datos que no están siendo directamente utilizados para ajustar los pesos de la red. La configuración que produzca el error de validación más bajo será la más adecuada, ya que producirá los errores más bajos para datos con los que no ha tenido contacto directo. A esto se le denomina capacidad de generalización, y corresponde al término interpolación en estadística clásica. El aprendizaje se para cuando se alcanza el error mínimo del conjunto de validación. En ese punto la red generaliza mejor. Cuando el aprendizaje no es interrumpido a tiempo, ocurre lo que se llama sobreentrenamiento (overtraining) y el rendimiento de la red sobre todos los datos disminuye, a pesar de que el error en el conjunto de selección de arquitectura todavía sigue disminuyendo. Después de finalizar la fase de entrenamiento, la red debería ser finalmente comprobada con un tercer conjunto de datos, el conjunto test. El procedimiento utilizado es el siguiente:

1. Los pesos de la red se inicializan a valores aleatorios en el rango $[-1,+1]$.
2. Se presenta el conjunto de datos de entrenamiento a la red, patrón a patrón y en orden aleatorio. Se calcula la salida de la red para el patrón.
3. Se calcula el error entre la salida de la red y el valor real deseado para ese patrón. El error se propaga hacia atrás a través de la red y los pesos son ajustados.
4. Después de un paso a través del conjunto de entrenamiento, se presentan los datos del conjunto de validación manteniendo los pesos fijos y se calcula la salida, para

finalmente derivar el error cuadrático medio (ECM o MSE) entre la salida y el valor deseado para los patrones del conjunto de validación.

5. Si el MSE del conjunto de validación es el menor obtenido hasta el momento, los valores actuales de los pesos de la red son guardados.
6. Los pasos 2 a 5 se repiten hasta que no se observe mejora en el error de validación MSE durante 200 ciclos de entrenamiento sobre los datos del conjunto de entrenamiento.
7. La red entrenada obtenida al finalizar el proceso (configuración que minimiza el error de validación) es evaluada mediante la presentación del conjunto test. El error obtenido para este conjunto sirve como criterio objetivo del poder de generalización de la red y puede ser comparado con otros valores así obtenidos por otras configuraciones.

Como se ha indicado, el rendimiento se mide normalmente utilizando estadísticos tales como el coeficiente de determinación, R^2 , o el error cuadrático medio (mean square error, MSE). En el método k-fold todo el proceso es repetido k veces realizando cada vez una partición aleatoria del conjunto de entrenamiento. En este caso se utilizó $k=10$, de forma que la red era entrenada con 168 patrones y validada sobre 18 patrones, que iban cambiando aleatoriamente (sin repetición) hasta completar los $k=10$ ciclos (el conjunto test contenía 20 patrones, es decir, 20 pozos). Así, la arquitectura con menor error medio de validación sobre las distintas particiones fue la seleccionada. El paso final del proceso consistió en el entrenamiento de la red seleccionada (por ejemplo, 5 unidades de entrada, 2 en una capa intermedia y 1 en la capa de salida, o 5-2-1) utilizando el conjunto completo de datos de entrenamiento para ajustar los pesos, y el conjunto test para decidir cuando parar el proceso (en este caso se usa como validación). La red así entrenada está lista para poder estimar nuevos patrones de forma adecuada.

Los resultados obtenidos de estos análisis indicaron para el primer experimento, que la red más simple que proporcionaba un buen rendimiento (MSE=0.42) en la generalización fue una red 4-2-1, es decir, con 2 unidades en la capa oculta (figura 5.27). La adición de más unidades no producía mejora apreciable en el error de validación. Esta arquitectura fue entonces entrenada utilizando el conjunto de entrenamiento con 186 patrones, mientras que el conjunto test con 20 patrones se utilizaba para parar en el momento adecuado el proceso, utilizando como guía el número de iteraciones y el error mínimo de entrenamiento obtenido en el proceso de selección de la arquitectura. Para el segundo experimento, la red más simple con un buen rendimiento en la generalización era una red 5-1-1, es decir, con una unidad en la capa oculta intermedia. Una vez entrenada la red para cada uno de los experimentos se generó la superficie de probabilidad

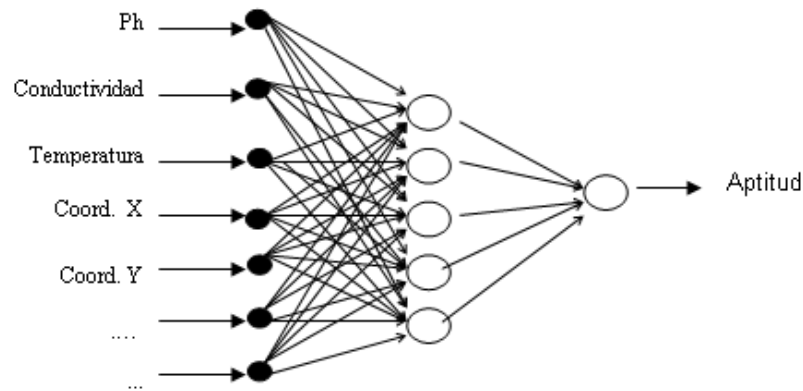


Figura 5.27: Esquema de la red neuronal utilizada para estimar la asociación entre variables de entrada y aptitud.

indicando el valor estimado de la calidad del agua en la Vega simplemente presentando a la red los valores de todas las celdas del área de estudio (cada celda corresponde a un patrón) y recogiendo el valor estimado por esta. La capa resultante presenta valores en el intervalo $[-1,1]$, aunque transformándolos linealmente al intervalo $[0,1]$ pueden interpretarse como probabilidades.

El rendimiento de la red neuronal como estimador fue analizado utilizando el coeficiente de determinación R^2 obtenido en la fase de entrenamiento sobre el conjunto test totalmente independiente. Tanto en el primer como en el segundo experimento el coeficiente presentó valores en torno a 0.6, sugiriendo que la cantidad de información de entrada en ambos experimentos era similar a los efectos de la estimación de la aptitud calculada a partir de todas las variables disponibles. No obstante el mapa de calidad generado con las redes de los dos experimentos presentan distribuciones con ciertas diferencias (figuras 5.28 a y b).

El mapa generado con las 42 variables recoge una zona más amplia de buena calidad a lo largo del río Genil y de peor calidad a ambos márgenes del río y en la zona de descarga del acuífero. Por otra parte, el mapa generado con 5 variables recoge una zona más estrecha a lo largo del río y destacan tres sectores de mala calidad, dos en la zona de recarga y uno en el borde oeste.

El mapa de aptitud del agua subterránea, calculado a partir de los mapas estimados mediante geoestadística de cada una de las variables, (figura 5.29) es más similar a la predicción realizada con 5 variables que a la predicción realizada con las 42 variables. Sin embargo, si comparamos la aptitud real en cada uno de los puntos de muestreo (figura 5.30) con los mapas estimados a partir de redes neuronales, se aprecia una mayor similitud con el mapa calculado a partir de las 37 variables.

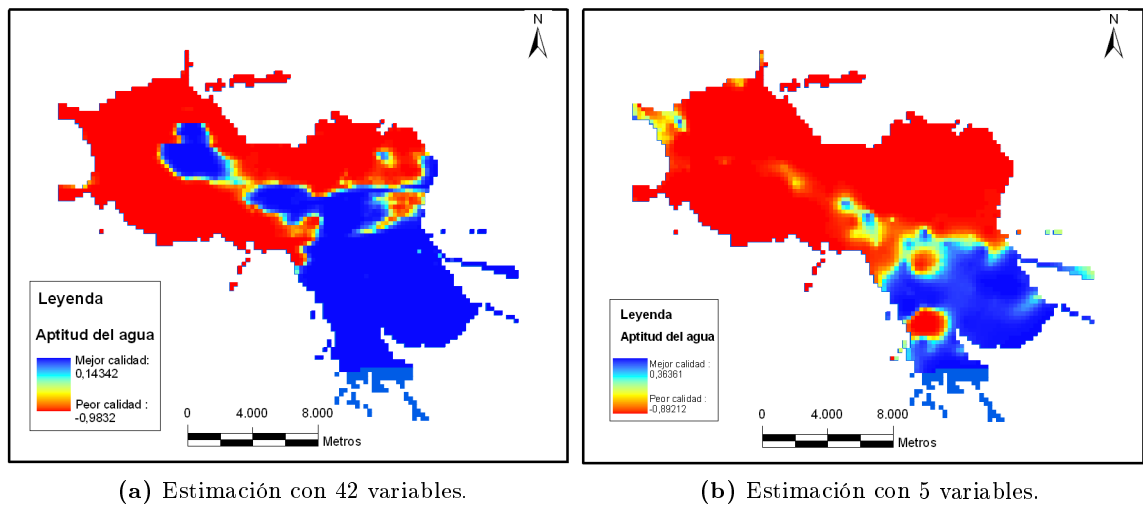


Figura 5.28: Estimación de la aptitud del agua para consumo público mediante redes neuronales.

De estas observaciones, puede concluirse, que la mayor similitud de la segunda estimación (con 5 variables) con el mapa de aptitud de superposición de capas estimadas por geoestadística es debida a la técnica de espacialización (krigeaje) de las variables, ya que ésta provoca un suavizado de los valores. Por otra parte, la primera estimación (con 37 variables) es capaz de recoger pequeñas anomalías, relacionadas con el resto de variables que obviamente no capta el segundo modelo de estimación.

Finalmente, cabe comentar que los experimentos realizados en la investigación sobre la Vega de Granada, indican que la red neuronal es una herramienta útil en la estimación de la aptitud de las aguas subterráneas. Las redes neuronales, además, no realizan supuestos sobre los datos ni sobre el tipo de relación entre las variables, como otros métodos estadísticos, lo que la convierte en una herramienta muy versátil. Otra de las ventajas de esta técnica es su habilidad para generalizar, pues una red convenientemente entrenada puede clasificar o estimar datos de la misma clase que los datos de entrenamiento aunque no los haya visto anteriormente. Por todo esto, la naturaleza computacional de esta metodología se considera ideal para la producción de mapas finales integrados de calidad del agua subterránea.

5.3.2. Estimación de la vulnerabilidad mediante Sistemas de Reglas Basados en Lógica Difusa

Como ya se ha comentado en el capítulo de las bases metodológicas, en los últimos años distintos estudios han usado las redes neuronales, la lógica difusa y técnicas neuro-difusas para hacer predicciones. Estas técnicas son capaces de aceptar incerti-

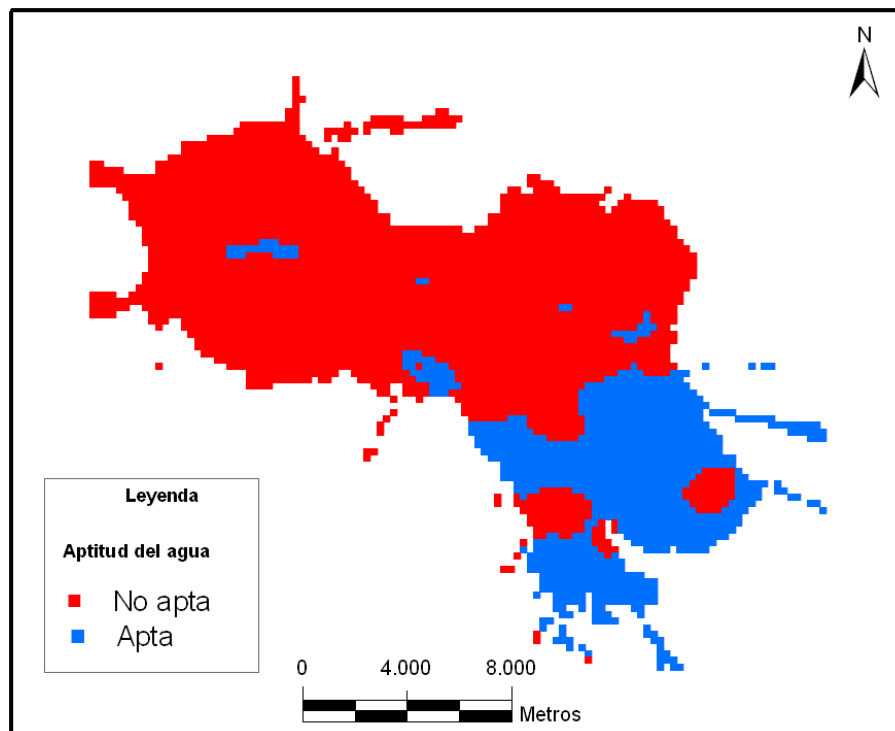


Figura 5.29: Mapa de aptitud del agua subterránea según las estimaciones geoestadísticas.

dumbres en los datos de entrada y permiten extraer información de datos incompletos o contradictorios (Rashid et al., 1992; Rogers et al., 1995; Dixon et al., 2002; Dixon, 2005; Wang et al., 2006 [148, 155, 59, 58, 57, 190]). Por este motivo se considera que el empleo de estas técnicas, en este estudio particular, es de especial interés, ya que la cantidad de información disponible en la Vega de Granada es ingente y en algunos casos contradictoria.

La principal ventaja en el empleo de los sistemas de reglas basados en lógica difusa es la interpretabilidad del modelo, frente a otras técnicas como las redes neuronales que actúan como una "caja negra". El conjunto de técnicas de SIG y reglas basadas en lógica difusa es especialmente útil cuando se modelan entradas, como los parámetros hidrodinámicos del acuífero y parámetros edafológicos, ya que además de obtener mapas de calidad del agua, estos sistemas permiten extraer un conjunto de reglas. Las reglas permiten identificar el conocimiento extraído por estos sistemas y aplicarlo al estudio de la calidad del agua, que es el caso que nos ocupa.

Supongamos que desconocemos parte del funcionamiento de un acuífero, aunque intuimos algunas de las variables que pueden influir en la contaminación del mismo. Por ejemplo, parece razonable que el espesor no saturado y el contenido en nitratos del suelo influyan en el contenido de nitratos del agua del acuífero. Etiquetamos cada una

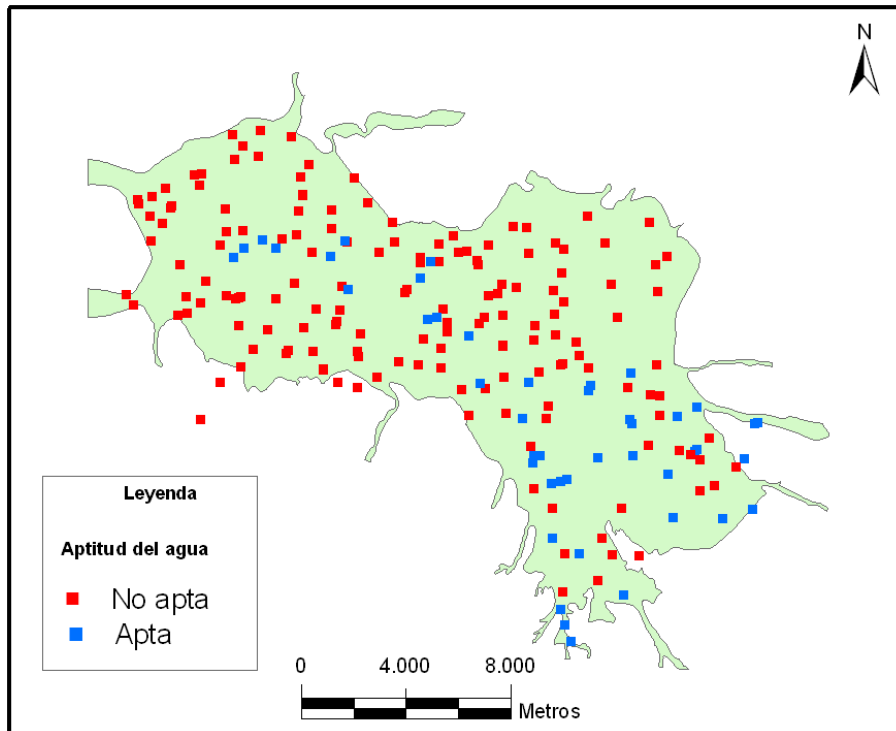


Figura 5.30: Mapa de aptitud del agua en los puntos de muestreo (según RD 140/2003).

de las variables de entrada como bajo, medio y alto, y la variable de salida como no contaminado, poco contaminado y contaminado.

En la evaluación de la vulnerabilidad de un acuífero es frecuente que los expertos usen expresiones como: "Si el espesor no saturado del acuífero es bajo y la concentración en nitratos es alta, entonces el agua del acuífero estará contaminada". En el lenguaje difuso esto se enunciaría:

Regla 1: Si Esp_no_sat es bajo & nitratos_suelo es alto entonces nitratos_agua es Contaminado.

Del mismo modo, se pueden enunciar otras reglas. Por ejemplo:

Regla 2: Si Esp_no_sat es alto & nitratos_suelo es alto entonces nitratos_agua es poco contaminado.

Regla 3: Si Esp_no_sat es alto & nitratos_suelo es bajo entonces nitratos_agua es no contaminado.

La robustez del sistema depende del número y calidad de las reglas. Como ya se comentó en el apartado de fundamentos, un SBRD se puede emplear tanto como herramienta para representar distintas formas de conocimiento sobre el problema a resolver, como para modelar las interacciones y relaciones existentes entre las variables del mismo. El principal interés de la modelización se debe a la dificultad de modelar las interaccio-

nes de sistemas naturales por su elevado grado de complejidad e imprecisión. De este modo, se han elaborado diferentes algoritmos para la extracción de conocimiento (extracción de reglas) y la eficacia o robustez del SBRD dependerá también del algoritmo elegido. En nuestro caso el algoritmo es el de Wang y Mendel (1992 [189]).

El principal objetivo de este apartado es, por lo tanto, la creación de mapas de vulnerabilidad a partir de las variables disponibles para el estudio. Como se comentó en el capítulo de fundamentos, la mayor parte de los estudios de vulnerabilidad con lógica difusa se basan en el índice DRASTIC (Aller et al. 1987 [7]) y como variable a estimar utilizan los nitratos, ya que es una de las variables más fácilmente medible en número suficiente.

En este trabajo se han hecho diversos experimentos, aunque los resultados más interesantes corresponden a modelos con pocas variables de entrada ya que, de este modo, el número de reglas obtenido es pequeño y, por lo tanto, son más sencillos de interpretar.

Uno de los primeros interrogantes a la hora de modelar un sistema mediante reglas, además del número de variables de entrada que, a priori, está condicionado por la disponibilidad de las mismas, es el número de etiquetas y la función de pertenencia. Algunos estudios muestran que la función de pertenencia trapezoidal es más estable que la triangular o la gaussiana (Dixon, 2005 [57]). Sin embargo, en la mayoría de los estudios hidrogeológicos de lógica difusa se usa la función de pertenencia triangular (Muhammetoglu, A. & Yardimci, A. 2006; Awasthi, A. et al. 2005; Li, et al. 2007 [125, 14, 106]). La función de pertenencia triangular (figura 5.31) se especifica por los parámetros, a , b y c , y es la utilizada en este estudio:

$$(x : a, b, c) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & \text{Si } x \in [a, b], \\ \frac{c-x}{c-b} & \text{Si } x \in [b, c], \\ 0 & x > c \end{cases}$$

El número de etiquetas es otro parámetro que va a modificar considerablemente el resultado de nuestro sistema y, por lo tanto, un aspecto importante a la hora de escoger un número de etiquetas adecuado. Una condición práctica, es escoger un número de etiquetas impar (por ejemplo: bajo, medio, alto), que uno par (por ejemplo: muy bajo, bajo, alto, muy alto) ya que es mas asimilable al razonamiento humano.

El primer conjunto de experimentos por lo tanto trata de establecer la idoneidad del sistema para modelar la variable vulnerabilidad mediante el contenido en nitratos. Para ello se han escogido 18 variables relacionadas con la vulnerabilidad del sistema según criterio experto (tabla 5.14) y se ha tomado como variable de salida el contenido

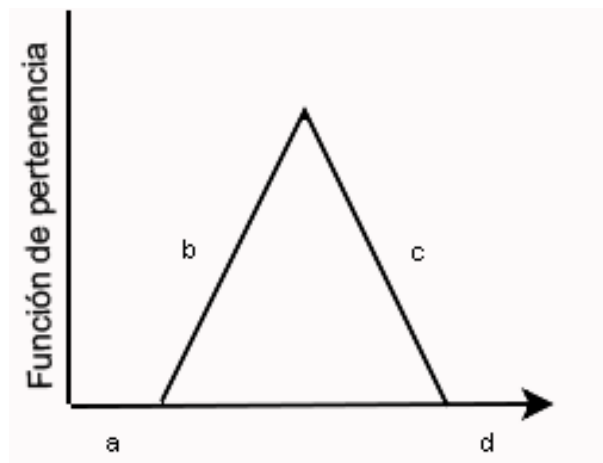


Figura 5.31: Función de pertenencia triangular.

en nitratos del agua del acuífero (Dixon, 2002 y 2005 [59, 58, 57]). El conjunto de las variables se han tipificado en el intervalo de -1 a +1, por lo que al normalizar la función de distribución de los valores de la variable, la asignación de las etiquetas es mas homogénea. Se han realizado diversos experimentos variando el número de reglas desde 3 a 11 (un número más elevado es poco entendible), y hemos variado también el número y las variables utilizadas para cada experimento (tabla 5.15). Los resultados más interesantes corresponden a los experimentos con pocas variables de entrada ya que estos modelos permiten aproximar la contaminación del acuífero a partir de un conjunto de datos fácil y económico de obtener.

El primer conjunto de experimentos se ha realizado con las variables de entrada: conductividad y pH del agua del acuífero. En este caso el menor error se consigue para los mapas creados con 7 y 9 etiquetas (tabla 5.15). Por otra parte, la distribución de los contenidos máximos y mínimos de los nitratos coinciden mayoritariamente para el conjunto de estimaciones (figura 5.32).

El mapa elaborado con 5 etiquetas y con los valores experimentales superpuestos (5.33) muestra una elevada coincidencia de los máximos y los mínimos, a pesar de tener un error más elevado. El aspecto más interesante de este ejemplo es el conjunto de reglas obtenidas (tabla 5.16). Se han deducido nueve reglas fácilmente interpretables, aunque observamos que no hay reglas para casos en los que la variable de salida nitratos sea alta o muy alta. Esto puede ser debido a que este tipo de reglas no se dan, es decir, el conjunto de ejemplos con valores altos o muy altos no responde a ninguna regla.

El segundo conjunto de experimentos se ha realizado con las variables de entrada: conductividad y pH del agua del acuífero, nitratos del suelo, espesor no saturado del acuífero y % de arcillas del primer metro de suelo. En este caso el menor error se consigue para el mapa creados con 11 etiquetas (tabla 5.15), también se obtiene un error bajo para

Nombre de la variable	<i>Experimento 1</i>	<i>Experimento 2</i>	<i>Experimento 3</i>	<i>Experimento 4</i>	<i>Experimento 5</i>
<i>Conductividad del agua</i>	X	X	X	X	X
<i>pH del agua</i>	X	X	X	X	X
<i>Temperatura del agua</i>					X
<i>Conductividad del suelo</i>				X	X
<i>Nitratos del suelo</i>		X	X	X	X
<i>Sulfatos del suelo</i>			X	X	X
<i>Litología de la zona no saturada</i>				X	X
<i>Conductividad hidráulica</i>					X
<i>Espesor no saturado</i>		X	X	X	X
<i>% Arcillas del suelo</i>		X	X	X	X
<i>% Arenas del suelo</i>				X	X
<i>pH del suelo</i>				X	X
<i>% Humedad del suelo</i>					X
<i>Carbono Orgánico del suelo</i>				X	X
<i>% Humedad de saturación del suelo</i>			X	X	X
<i>Altura de la base del acuífero</i>					X
<i>Módulo del gradiente hidráulico</i>			X	X	X

Tabla 5.14: Variables incluidas en el SBRD para la inferencia de la vulnerabilidad.

<i>Experimento 5</i>				
Número de variables = 17				
<i>Número de etiquetas</i>	3	5	7	9
<i>Número de reglas</i>	144	156	156	156
<i>Error</i>	0.695	0.638	0.344	0.398
<i>Experimento 4</i>				
Número de variables = 13				
<i>Número de etiquetas</i>	3	5	7	9
<i>Número de reglas</i>	113	146	153	155
<i>Error</i>	0.723	0.781	0.677	0.440
<i>Experimento 3</i>				
Número de variables = 8				
<i>Número de etiquetas</i>	3	5	7	9
<i>Número de reglas</i>	59	116	141	149
<i>Error</i>	0.741	0.630	0.631	0.539
<i>Experimento 2</i>				
Número de variables = 5				
<i>Número de etiquetas</i>	3	5	7	9
<i>Número de reglas</i>	27	75	120	139
<i>Error</i>	0.877	0.572	0.218	0.213
<i>Experimento 1</i>				
Número de variables = 2				
<i>Número de etiquetas</i>	3	5	7	9
<i>Número de reglas</i>	6	9	16	22
<i>Error</i>	0.685	1.003	0.402	0.428

Tabla 5.15: Tabla de errores de lógica difusa.

Conductividad del agua	pH del agua	Nitratos del agua
muy bajo	medio	muy bajo
muy bajo	bajo	muy bajo
bajo	bajo	muy bajo
muy bajo	muy bajo	muy bajo
medio	medio	bajo
medio	bajo	bajo
muy alto	bajo	muy bajo
bajo	medio	bajo
muy bajo	muy alto	muy bajo

Tabla 5.16: Base de reglas del experimento 2. (2 variables de entrada y 5 etiquetas).

Dos variables de entrada

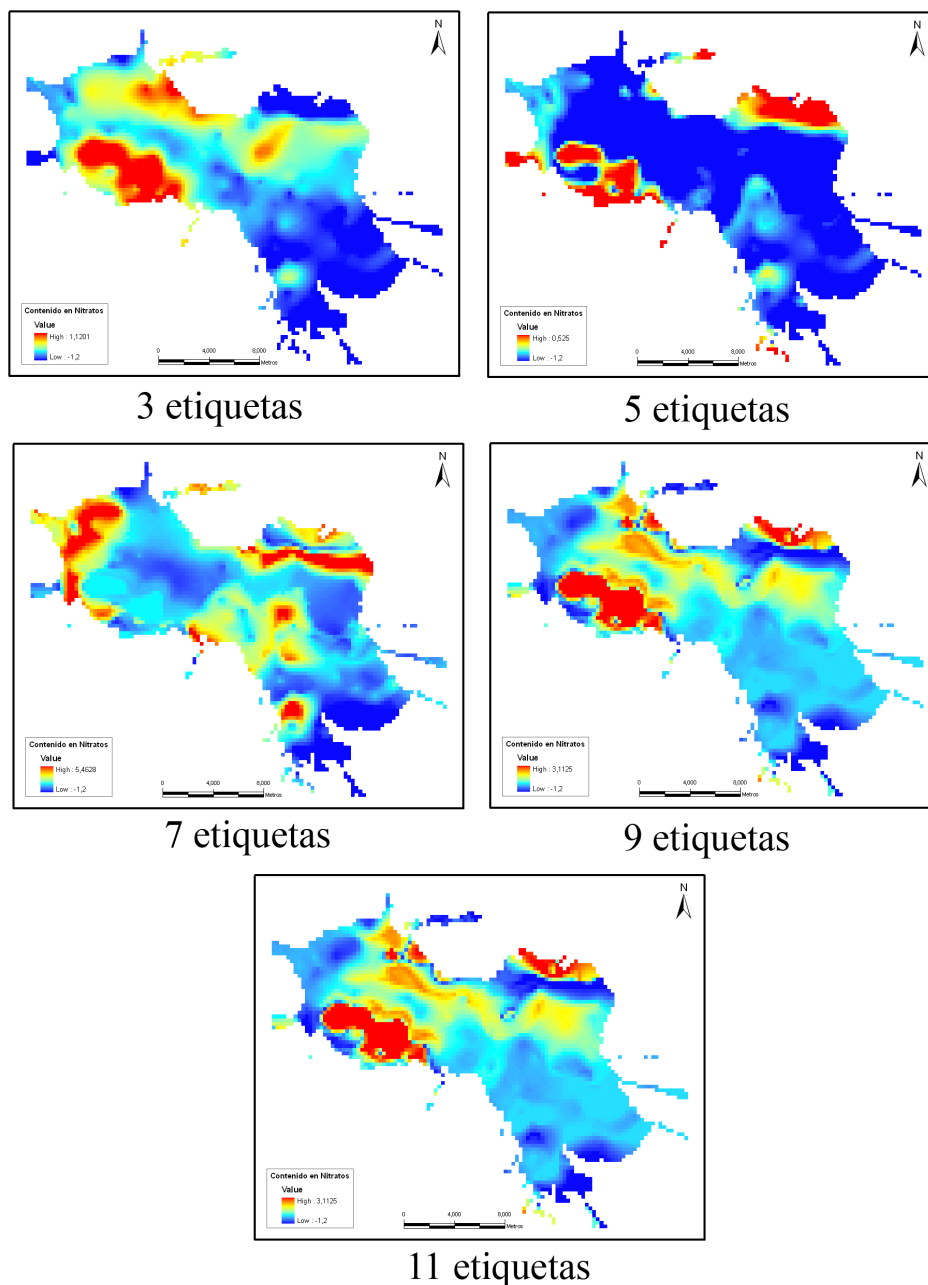


Figura 5.32: Mapas del contenido en nitratos del agua subterránea obtenidos mediante SBRD con dos variables de entrada.

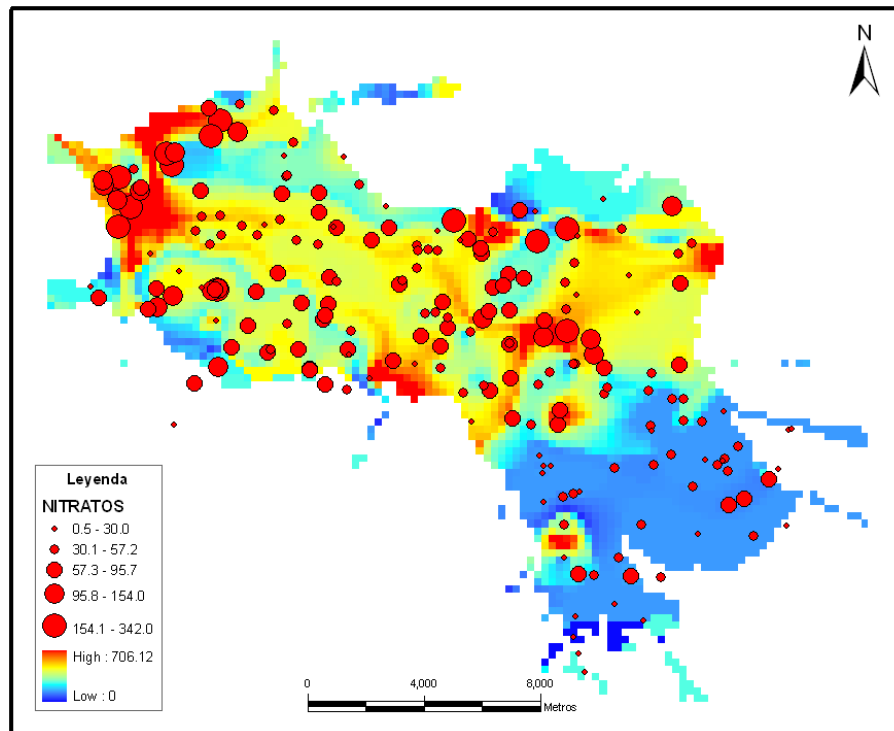


Figura 5.33: Estimación del contenido de nitratos del agua subterránea mediante SBRD (experimento 2).

los mapas con 7 y 9 etiquetas. En este caso, igual que para el anterior, la distribución de los contenidos máximos y mínimos de los nitratos coinciden mayoritariamente para el conjunto de estimaciones (figura 5.34), sin embargo existe una diferencia clara con el conjunto de estimaciones para los que se ha usado tan sólo 2 variables de entrada.

El conjunto de reglas con 5 variables aumenta considerablemente (tabla 5.15), para el ejemplo con 5 etiquetas obtenemos un total de 75 reglas (tabla 5.17). Este experimento es uno de los que tienen menor error. El mapa obtenido presenta buena correspondencia con los valores experimentales de las muestras (figura 5.35). Las mayores concentraciones de nitratos en el agua del acuífero se sitúan fundamentalmente en el borde oeste y localmente en la zona de Fuente Vaqueros.

Tabla 5.17: Base de reglas con 5 variables de entrada y 5 etiquetas.

Cond. Agua	pH agua	NO3 suelo	ENS	% A suelo	NO3 agua
muy bajo	medio	bajo	medio	muy bajo	muy bajo
muy bajo	bajo	muy bajo	medio	muy bajo	bajo
muy bajo	medio	bajo	bajo	muy bajo	muy bajo

(continúa en página siguiente)

(viene de página anterior)

Cond. Agua	pH agua	NO3 suelo	ENS	% A suelo	NO3 agua
muy bajo	bajo	bajo	bajo	muy bajo	muy bajo
muy bajo	bajo	bajo	bajo	bajo	bajo
muy bajo	bajo	bajo	muy bajo	bajo	bajo
muy bajo	medio	bajo	muy bajo	muy bajo	muy bajo
muy bajo	bajo	bajo	muy bajo	muy bajo	bajo
bajo	bajo	bajo	muy bajo	medio	bajo
bajo	bajo	alto	muy bajo	alto	muy bajo
bajo	bajo	muy bajo	muy bajo	medio	muy bajo
muy bajo	bajo	alto	muy bajo	medio	bajo
bajo	bajo	alto	muy bajo	medio	bajo
muy bajo	bajo	muy bajo	muy bajo	bajo	bajo
muy bajo	bajo	bajo	muy bajo	medio	muy bajo
muy bajo	muy bajo	bajo	muy bajo	alto	muy bajo
muy bajo	bajo	muy alto	muy bajo	alto	muy alto
bajo	bajo	bajo	bajo	medio	muy bajo
muy bajo	bajo	bajo	muy bajo	alto	bajo
muy bajo	medio	muy bajo	medio	muy bajo	muy bajo
medio	medio	bajo	bajo	medio	bajo
muy bajo	bajo	medio	muy bajo	muy bajo	bajo
muy bajo	bajo	bajo	medio	bajo	muy bajo
bajo	bajo	medio	muy bajo	medio	bajo
bajo	bajo	muy alto	muy bajo	muy alto	medio
bajo	bajo	muy bajo	muy bajo	alto	bajo
muy bajo	bajo	muy bajo	alto	muy bajo	muy bajo
muy bajo	medio	medio	alto	muy bajo	muy bajo
muy bajo	bajo	bajo	alto	muy bajo	muy bajo
muy bajo	bajo	bajo	medio	muy bajo	bajo
muy bajo	medio	medio	bajo	bajo	muy bajo
bajo	bajo	muy alto	muy bajo	alto	bajo
muy bajo	medio	bajo	muy bajo	alto	muy bajo
muy bajo	bajo	muy bajo	muy bajo	medio	muy bajo
bajo	bajo	bajo	muy bajo	alto	muy bajo
muy bajo	medio	bajo	muy bajo	medio	muy bajo
bajo	bajo	muy bajo	muy bajo	bajo	bajo

(continúa en página siguiente)

(viene de página anterior)

Cond. Agua	pH agua	NO3 suelo	ENS	% A suelo	NO3 agua
muy bajo	bajo	bajo	bajo	medio	bajo
muy bajo	bajo	muy alto	muy bajo	muy alto	medio
muy bajo	bajo	muy alto	bajo	muy alto	bajo
bajo	bajo	bajo	muy bajo	bajo	bajo
bajo	bajo	bajo	muy bajo	muy bajo	bajo
muy bajo	bajo	medio	muy bajo	bajo	bajo
muy bajo	bajo	muy alto	muy bajo	bajo	bajo
medio	bajo	medio	muy bajo	medio	bajo
muy alto	bajo	medio	medio	muy alto	muy bajo
muy bajo	bajo	bajo	alto	bajo	muy bajo
muy bajo	medio	muy alto	muy bajo	alto	bajo
muy bajo	medio	muy alto	muy bajo	muy alto	muy bajo
bajo	medio	bajo	muy bajo	medio	bajo
muy bajo	bajo	medio	muy bajo	alto	muy bajo
bajo	bajo	muy alto	muy bajo	medio	bajo
bajo	bajo	alto	muy bajo	bajo	bajo
bajo	bajo	medio	muy bajo	alto	bajo
muy bajo	medio	bajo	alto	muy bajo	muy bajo
muy bajo	medio	muy bajo	alto	muy bajo	muy bajo
muy bajo	medio	muy bajo	medio	bajo	bajo
muy bajo	medio	bajo	muy bajo	bajo	muy bajo
muy bajo	medio	bajo	bajo	bajo	muy bajo
bajo	medio	muy alto	muy bajo	muy alto	medio
bajo	bajo	medio	muy bajo	muy bajo	bajo
muy bajo	bajo	alto	muy bajo	bajo	bajo
muy bajo	medio	medio	medio	muy bajo	bajo
muy bajo	bajo	medio	muy bajo	medio	muy bajo
bajo	medio	bajo	bajo	medio	bajo
muy bajo	medio	muy bajo	muy bajo	alto	bajo
bajo	bajo	bajo	medio	bajo	muy bajo
muy bajo	medio	muy bajo	muy alto	muy bajo	muy bajo
bajo	medio	bajo	bajo	alto	bajo
bajo	medio	muy alto	bajo	muy alto	bajo
muy bajo	bajo	alto	muy bajo	muy alto	bajo

(continúa en página siguiente)

(viene de página anterior)

Cond. Agua	pH agua	NO3 suelo	ENS	% A suelo	NO3 agua
muy bajo	muy alto	bajo	muy bajo	alto	muy bajo
bajo	bajo	medio	muy bajo	bajo	bajo
muy bajo	bajo	muy bajo	medio	bajo	muy bajo
medio	bajo	alto	muy bajo	medio	medio

El tercer conjunto de experimentos se ha realizado con las variables de entrada: conductividad y pH del agua del acuífero, nitratos y sulfatos del suelo, espesor no saturado del acuífero, % de arcillas del primer metro de suelo, humedad de saturación del último horizonte del suelo y módulo del gradiente hidráulico. Como en los casos anteriores el menor error se ha conseguido con el mayor número de etiquetas (11 etiquetas) (tabla 5.15). Los mapas inferidos mediante SBRD con ocho variables (figura 5.34) presentan una clara diferenciación para estimaciones con más de 5 etiquetas. Estos mapas empiezan a ser menos realistas y con mayor dispersión espacial (heterogeneidad), a pesar de que las distribuciones generales se mantienen.

Finalmente, la representación de las estimaciones del contenido en nitratos mediante SBRD con trece variables de entrada (conductividad y pH del agua del acuífero; conductividad, pH, nitratos, sulfatos y carbono orgánico del primer metro de suelo; litología de la zona no saturada; espesor no saturado del acuífero; % de arcillas y de arenas del primer metro de suelo; humedad de saturación del último horizonte del suelo; y módulo del gradiente hidráulico) reflejan más claramente el efecto que se intuía en los mapas anteriores. El elevado número de etiquetas y de variables de entrada generan unos mapas nada realistas y poco útiles.

La integración SIG y las reglas difusas permiten la construcción de mapas de vulnerabilidad y facilitan mostrar los resultados en un contexto espacial. El conjunto de experimentos realizados permite detectar cual es el número de variables y etiquetas más idóneo para la generación de mapas de vulnerabilidad del agua subterránea. Los resultados muestran que un elevado número de variables no es adecuado para la estimación de los nitratos del agua subterránea. El conjunto óptimo de variables de entrada para los experimentos realizados es de 5 variables y 11 etiquetas, aunque con 9 y 5 se consigue también un buen error, y el mapa obtenido es muy similar. También observamos que un número superior a ocho variables de entrada no es adecuado para la generación de mapas de vulnerabilidad mediante SBRD, ya que el sistema genera mapas poco realistas y nada suavizados.

Las técnicas difusas integradas con SIG son en sí mismas sofisticadas herramientas de construcción de índices y superposición de capas. Estas herramientas tienen la capacidad de generar mapas de vulnerabilidad del agua subterránea y funcionar con

Cinco variables de entrada

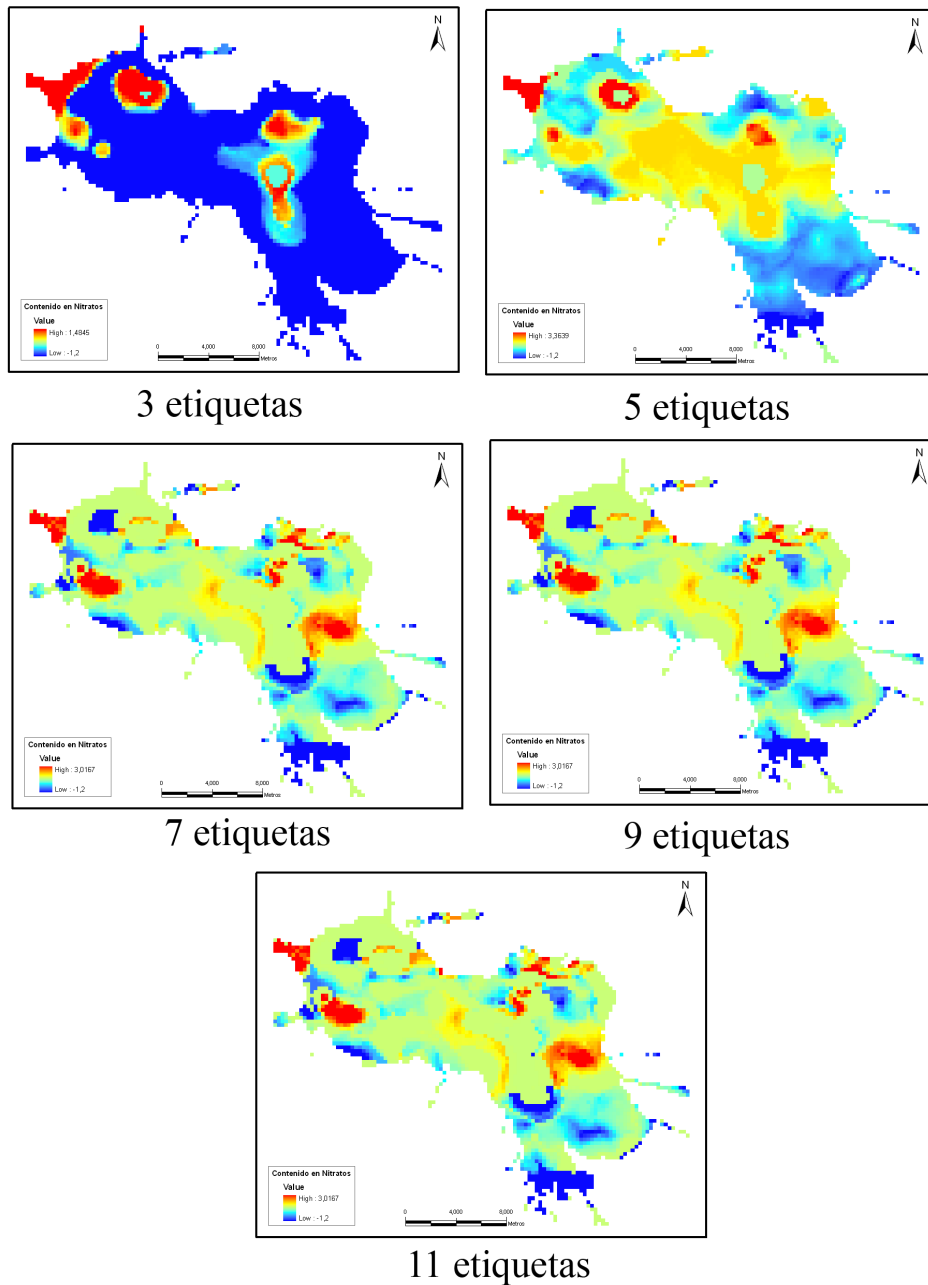


Figura 5.34: Mapas del contenido en nitratos del agua subterránea obtenidos mediante SBRD con cinco variables de entrada.



Figura 5.35: Estimación del contenido en nitratos del agua del acuífero mediante SBRD.

datos imprecisos. Por otra parte, estos modelos también tienen limitaciones que son comunes a las aproximaciones de superposición e índices, por lo que debemos tomar los resultados obtenidos mediante estas técnicas (técnicas de evaluación multicriterio y de construcción de índices) con precaución debido a la naturaleza subjetiva de estos métodos.

Por otra parte, los modelos basados en inteligencia artificial (SBRD y RNA) carecen de carga subjetiva y su limitación es debida a que es necesario un volumen suficiente de datos para generalizar y a que existe cierta complejidad en su aplicación. El uso de los modelos planteados en este capítulo (AHP, OWA, creación de índices, SBRD y RNA) puede verse favorecido con la ayuda de modelos numéricos complementarios de flujo de las aguas subterráneas y de transporte de solutos. Estas aproximaciones deben ser utilizadas como una herramienta más dentro del conjunto de un SIG integrado por métodos como la teledetección, modelos de transporte de solutos, geoestadística, etc. Sin embargo, el gran interés de estos métodos es que ofrecen un medio de tratar con información imprecisa, y que permiten realizar aproximaciones viables para estudios regionales donde prevalecen los datos imprecisos.

Ocho variables de entrada

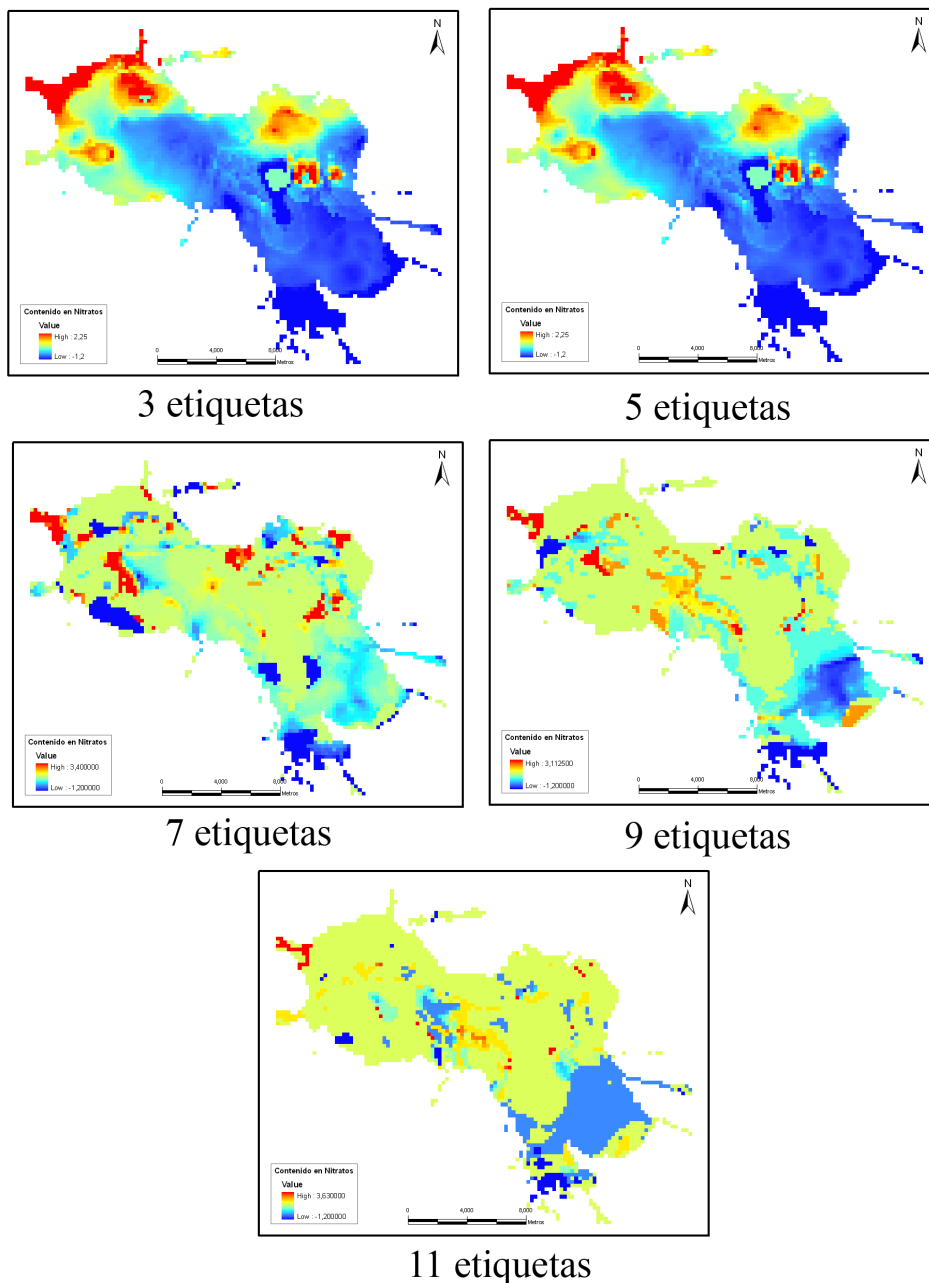


Figura 5.36: Mapas del contenido en nitratos del agua subterránea obtenidos mediante SBRD con ocho variables de entrada.

Trece variables de entrada

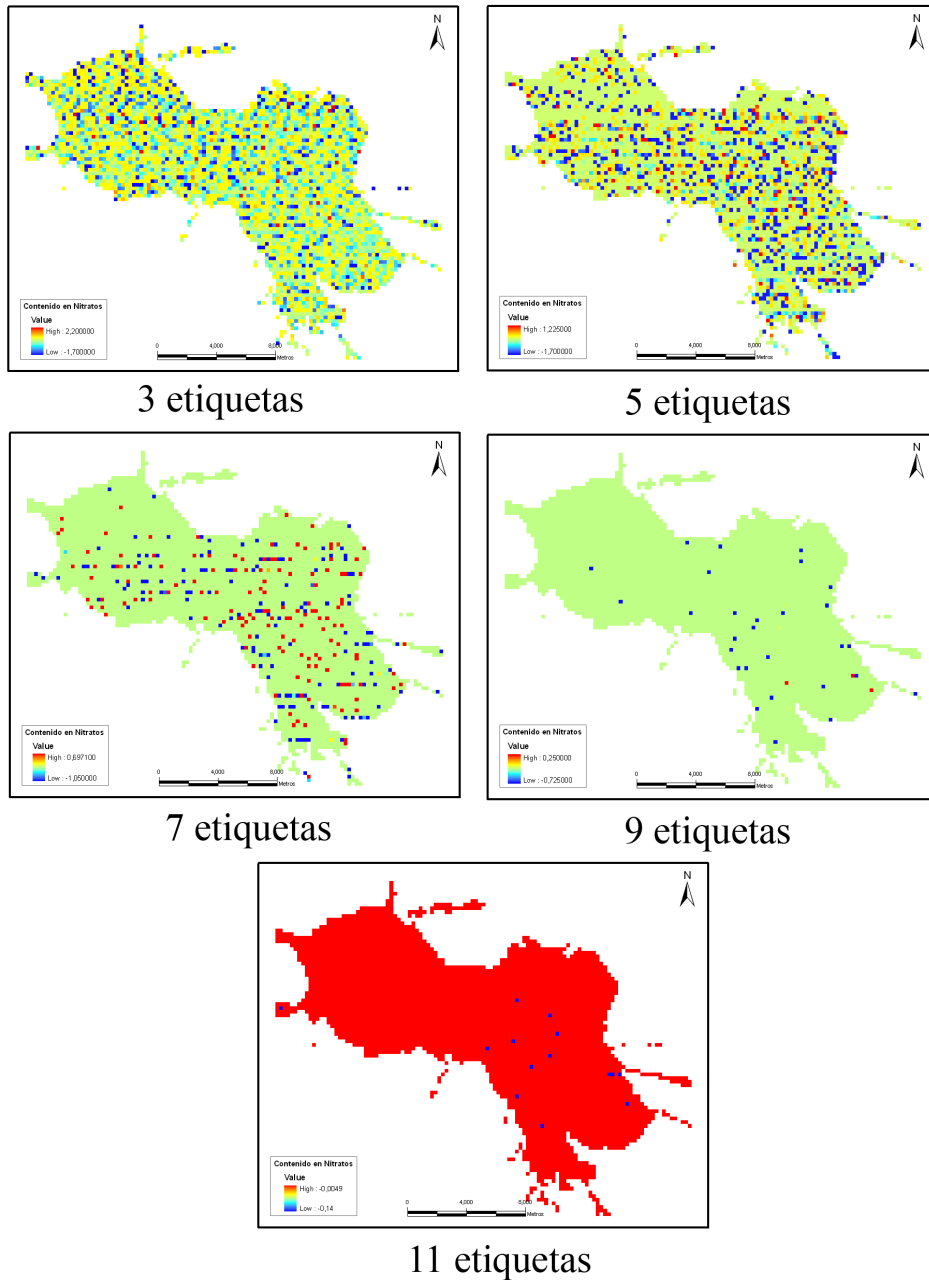


Figura 5.37: Mapas del contenido en nitratos del agua subterránea obtenidos mediante SBRD con cinco variables de entrada.

Capítulo 6

Conclusiones

En esta tesis doctoral se ha llevado a cabo el desarrollo de una metodología basada en SIG para el estudio de la calidad de los recursos hídricos en el acuífero de la Vega de Granada. El interés de esta metodología radica en su capacidad predictiva en estudios dirigidos a la evaluación espacial de dichos recursos como para el análisis de la vulnerabilidad de acuíferos a diferentes procesos de contaminación.

La metodología integra diferentes métodos y técnicas de análisis espacial de datos, entre los que cabe destacar los procedimientos estadísticos uni y multivariantes para la descripción de los datos, los métodos geoestadísticos para la estimación espacial de los datos, los métodos de inteligencia artificial para la integración de variables, etc. Todos ellos se han incorporado al SIG lo que ha facilitado la modelización espacial y temática de la información.

Con esta metodología se pretende aportar conocimiento científico que facilite y complemente los estudios hidrogeológicos del acuífero y , en consecuencia, ayude al gestor en la tarea compleja de toma de decisión referente a la gestión de las aguas subterráneas. Esta aportación se debe contemplar desde la óptica de la amplia variedad de métodos numéricos aplicados orientados a la generación e integración de información (por ejemplo, índices de calidad).

Quiero resaltar que la metodología propuesta tiene carácter general, por lo que puede ser aplicada al estudio de cualquier acuífero de interés; además, puede ser modificada incorporando nuevas herramientas de análisis de acuerdo con las necesidades de cada caso de estudio.

Vistas las conclusiones anteriores que tienen carácter general, seguidamente se expondrán una serie de conclusiones específicas surgidas en las diferentes etapas del desarrollo del trabajo.

En primer lugar, cabe destacar la enorme Base de Datos Geoespacial generada en este estudio relativa al conocimiento hidrogeológico de la Vega de Granada. Esta base de

datos es en sí misma la herramienta más importante para el estudio y la evaluación de los recursos del acuífero de la Vega de Granada. La integración en SIG del total de análisis de aguas, de suelos, de parámetros hidrodinámicos y de información complementaria socioeconómica, hidrogeológica y de imágenes de satélite ha sido crucial para el posterior desarrollo de las metodologías. La homogeneidad de la información ha sido el punto más importante en la creación de la base de datos, pues facilita la integración y creación de información.

Como metodología de análisis, la estadística y la geoestadística han aportado un conocimiento en profundidad de la distribución espacial de las variables estudiadas, así como la generación de mapas continuos en el espacio de cada una de ellas. En general, estadísticamente hablando hemos comprobado que la mayor parte de las variables hidrogeoquímicas del agua y las variables químicas de los suelos presentan poblaciones lognormales, lo que ha permitido una correcta espacialización de los datos. Se distinguen distintos comportamientos espaciales gracias al análisis del variograma que identifica las variables de gran continuidad espacial, las de alcance medio y las de elevado grado de aleatoriedad como los metales. La identificación espacial del tipo de variable permite establecer redes de muestreo optimizadas. Además, cuando la interpretación hidroquímica se combina con los métodos estadísticos multivariantes también facilita una mejor comprensión del flujo de las aguas subterráneas en sistemas acuíferos complejos. Los métodos estadísticos multivariantes también permiten definir la zonación geoquímica del sistema acuífero teniendo en cuenta los efectos de la litología, el espesor no saturado, la textura de la zona no saturada y la pendiente. La integración de todos los métodos de clasificación dentro del contexto geológico permite refinar la interpretación de la evolución geoquímica de las aguas subterráneas. El análisis multicriterio de los suelos ha identificado distintos factores donde destacamos los relacionados con la contaminación de los suelos por la carga de abonos y su relación con los suelos más limosos y con menos nutrientes.

Otro resultado también interesante es la combinación de la geoestadística y el análisis espectral en la identificación de los ciclos anuales e interanuales relacionados con ciclos climáticos y la influencia de las propiedades hidráulicas de los materiales en la conservación de los mismos.

La teledetección ha permitido la clasificación digital de imágenes para la elaboración de mapas de usos del suelo diferenciando entre: arbolado, bosques, embalses, matorral, olivar, regadíos, secano y urbano. Estas clases son de especial importancia en la gestión de los recursos hídricos subterráneos de la Vega de Granada, ya que el abuso de los abonos nitrogenados en la agricultura es una de las principales causas de contaminación del acuífero. Además, de la contaminación por nitratos de la agricultura, la recarga del acuífero por aguas de retorno de riego es una variable muy importante en la mo-

delización del acuífero. Además de conocer las áreas de máxima infiltración de agua al acuífero (zonas de regadío) sería necesario disponer de información cuantitativa respecto al origen, cantidad y frecuencia de riego de los cultivos para una correcta modelización, así como del abonado de los tipos de cultivo. Esta tarea queda pendiente para próximos estudios y que se podría completar con un modelo de flujo del agua.

Una aproximación muy útil en el estudio de la calidad del agua es el análisis de la probabilidad de que ésta sea apta para el consumo público. Esta aplicación probabilística combinada con lógica booleana muestra una calidad del agua muy degradada localmente en el acuífero debido fundamentalmente a los nitratos, aunque también debido a otras variables conflictivas como el boro, los cloruros, el flúor y la conductividad. Los mapas generados permiten conocer la probabilidad de encontrar un agua no apta para el consumo público así como la probabilidad de que este agua no se apta en más de una variable.

Otra metodología ampliamente utilizada para la evaluación de los recursos hídricos, así como para la evaluación de tendencias, son los índices de calidad del agua. Se han realizado diversas aproximaciones mediante los métodos más habituales que han presentado diferencias en los valores de los índices a pesar de estar en una misma escala. Sin embargo, a escala local (función de probabilidad local) la mayor parte de los métodos presentan las mismas funciones y por lo tanto la zonación en las áreas de calidad es similar. A pesar de ello, el método que se ha considerado más interesante es el método de agregación exponencial, ya que tanto en términos absolutos como relativos permite una mejor identificación de la calidad del agua.

La evaluación de la vulnerabilidad recoge distintos criterios y distintas aproximaciones que deben ser tenidas en cuenta. Las técnicas combinadas de SIG, lógica difusa y evaluación multicriterio han resultado adecuadas en la obtención de mapas de vulnerabilidad. Se han obtenido dos mapas de vulnerabilidad agregados mediante dos técnicas distintas, WLC y OWA. El resultado de la agregación de pesos ordenados es más satisfactorio, a pesar de no existir grandes diferencias, lo que indica que la metodología es moderadamente sensible al método de agregación utilizado. Las ventajas de esta técnica multicriterio es la estandarización de los factores a una escala común, la aplicación de pesos mediante un criterio de ponderación pareada (AHP), que es más objetivo, y la agregación de los pesos ordenados que sirve para resaltar las particularidades locales, haciendo el método más eficiente a una escala mayor.

Para evitar la carga de subjetividad del experto en la integración de datos, se han propuesto los métodos de Inteligencia Artificial. Estos métodos han permitido extraer conocimiento y encontrar relaciones no lineales para simular la actividad de razonamiento humano. Las redes neuronales han facilitado el reconocimiento de patrones, así como la regresión y la clasificación, siendo muy útiles para amplios conjuntos de datos

donde el razonamiento humano está más sesgado. Otra ventaja de las redes neuronales es su habilidad para generalizar, una vez entrenada la red neuronal aplicada en este trabajo ésta puede clasificar datos de la misma clase que los de entrenamiento aunque no los haya visto anteriormente. De este modo la red neuronal de retropropagación elaborada ha permitido estimar la función no lineal que relaciona las variables fisicoquímicas del agua y del suelo, los parámetros hidrodinámicos del acuífero y la geología con la calidad del agua subterránea. Se han realizado dos experimentos incluyendo 37 y 5 variables de entrada. A priori, los resultados son similares ya que el error de estimación en los dos casos es en torno a 0.6, sin embargo la observación de los mapas obtenidos permite detectar diferencias que son consecuencia de pequeñas anomalías relacionadas con el resto de variables que obviamente no capta el segundo modelo de estimación.

Los sistemas de reglas basados en lógica difusa, además de aceptar incertidumbres, como la red neuronal, permiten extraer conocimiento en forma de reglas lingüísticas. Los experimentos realizados en este trabajo han mostrado que la mejor aproximación corresponde a un conjunto de 5 variables de entrada y 7 etiquetas lingüísticas. Otro resultado interesante es que un número excesivamente elevado de variables de entrada generan mapas poco realistas y nada útiles, ya que el sistema tiende a generalizar demasiado mostrando mapas de valores intermedios y homogéneos. Otro dato que es interesante notar es que en el conjunto de reglas no obtenemos ninguna para la salida "nitratos muy alto", lo que corrobora lo mencionado anteriormente sobre utilizar los nitratos como variable de salida o como variable de test. El problema de utilizar los nitratos como indicador de contaminación se debe a que, a pesar de ser el mejor indicador de contaminación antrópica, la no existencia de contaminación por nitratos no significa la vulnerabilidad sea muy baja, sino que lo más probable es que no haya habido adición de nitratos en superficie.

Estas técnicas neurodifusas integradas con SIG han permitido la construcción de mapas de vulnerabilidad y han permitido mostrar los resultados en un contexto espacial. Es necesario alertar que estas sofisticadas herramientas funcionan con datos imprecisos y permiten ayudar en la identificación de áreas conflictivas cuando no existe otra información del área estudiada, por lo tanto son un complemento ideal a las técnicas clásicas de modelos de transporte de solutos.

Finalmente, se concluye que el conjunto de técnicas presentadas en esta tesis son herramientas adecuadas que permiten la generación de información espacial específica relativa a las diversas fuentes de datos (usos del suelo, calidad del agua, tipos de suelos, hidrodinámica del acuífero, clima, socioeconomía, etc.) así como el tratamiento individual o multivariante de los diferentes tipos de datos. Las técnicas presentadas se pueden clasificar en dos conjuntos: el análisis de los datos y la integración de los mismos. El análisis estadístico, espacial y multivariante de las variables ofrece resultados actuali-

zados de los datos y de sus relaciones que en la integración de los mismos son de gran utilidad ya que validan los resultados obtenidos de la integración de la información, ya sea mediante criterio experto o mediante modelos basados en los datos. Los novedosos métodos aplicados en este trabajo han permitido profundizar en el estudio de la hidrogeología y de la calidad de los recursos hídricos de la Vega de Granada aportando información novedosa y nuevos tratamientos de la información.

Las técnicas presentadas además han permitido mejorar y completar la información del acuífero de la Vega de Granada integrándola en una base de datos espacial actualizada, de fácil recuperación y homogénea. El conjunto de la información ha sido evaluada y detallada en sus tratamientos lo cual facilita al gestor la tarea de identificar y resolver problemas. Este ciclo de la información constituye nuestro Sistema de Soporte de Decisión aplicado a los Recursos Hídricos de la Vega de Granada.

Bibliografía

- [1] Proyecto piloto de utilización de aguas subterráneas para el desarrollo agrícola de la cuenca del Guadalquivir. utilización de las aguas subterráneas para la mejora del regadío de la vega de Granada. Inédito, FAO & IGME, 1972.
- [2] Etude géostatistique des données de transmissivité de la nappe. Technical report, Centre de Géostatistique et de morphologie mathématique., 1981.
- [3] Exploración geológica minera de la cuenca lignífera de Granada. Technical report, Instituto Geológico y Minero de España, 1981.
- [4] Implication of temporal variations and vertical stratification of groundwater nitrate-nitrogen in the Hall County special use area. Technical report, Nebraska Water Resources Center, 1984.
- [5] Propuesta de normas de explotación de las unidades hidrogeológicas con afección a embalses de regulación y fuentes de abastecimiento a poblaciones de la cuenca del Guadalquivir. Technical report, CHG e ITGE, 1993.
- [6] M.J. Adams. *Analytical Methods of Food Authentication*, chapter The principles of multivariate data analysis, page 350. P.R. Ashurst and M.J. Dennis, 1998.
- [7] L. Aller, T. Bennet, J.H. Lehr, R.J. Petty, and G. Hackett. *DRASTIC: a standardized system for evaluating ground water pollution potential using hydrogeologic settings*. U.S. Environmental Protection Agency, 1987.
- [8] William M. Alley. *Regional Groundwater Quality*, chapter Establishing a Conceptual Framework, pages 23–63. John Wiley and Sons, 1993.
- [9] B.J. Alloway. *Heavy metals in soils*. Blackie Academic and Professional, 1995.
- [10] D. Andina de la Fuente. *Redes neuronales artificiales*. Universidad Politécnica de Madrid, 2005.

-
- [11] J.M Andujar, J. Aroba, M.L. de la Torre, and J.A. Grande. Contrast of evolution models for agricultural contaminants in ground waters by means of fuzzy logic and data mining. *Environmental Geology*, 49:458–466, 2006.
- [12] L. Anselin and A. Getis. *Geographic Information Systems, Spatial Modelling and Policy Evaluation*, chapter Spatial statistical analysis and geographic information systems, pages 35–50. Fischer, M.M. and Nijkamp, P., 1993.
- [13] S. Aronoff. *Geographic Information System: a management perspective*. WDL, 1989.
- [14] A.K. Awasthi, O.P. Dubey, A. Awasthi, and S. Sharma. A fuzzy logic model for estimation of groundwater recharge. In *Annual Meeting of the North American Fuzzy Information Processing Society*, 2005.
- [15] R. Babuška. *Fuzzy modeling for control*. Kluwer Academic, Norwell, MA, EE. UU, 1998.
- [16] A. Bannari, D. Morin, F. Bonn, and A.R. Huete. A review of vegetation indexes. *Remote Sensing Reviews*, 13:95–120, 1995.
- [17] J.K. Berry. *Beyond Mapping: Concepts, Algorithms, and Issues in GIS*. GIS World Inc., 1993.
- [18] J.K. Berry. *Spatial reasoning for effective GIS*. GIS World Inc., 1995.
- [19] C.M. Bishop. *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.
- [20] G.F. Bonham-Carter. *Geographic Information System for geoscientist: modelling with GIS*. Pergamon, Ottawa, Canada, 1994.
- [21] J. Bosque Sendra. *Sistemas de Información Geográfica*. Rialp, 1997.
- [22] J. Bosque-Sendra and A. Moreno-Jiménez. *Prácticas de análisis exploratorio y multivariante de datos*. Oikos-Tau, Barcelona, 1994.
- [23] A. Bárdossy and L. Duckstein. *Fuzzy rule-based modeling with application to geophysical, biological and engineering systems*. CRC Press Boca Raton, 1995.
- [24] McClelland N. I. Deininger R. A. Brown, R. M. and R. G. Tozer. A water quality index-do we dare? *Water and Sewage Works*, 117:339–343, 1970.
- [25] P. A. Burrough. *Principles of Geographical Information Systems for Land resources assessment*. Claredon Press, Oxford University Press, 1986.

- [26] J.B. Campbell. *Introduction to remote sensing*. The Gulf Press, 1987.
- [27] E. Campos Fernández Figarez. *Estudio de la contaminación y fraccionamiento químico de metales pesados en suelos de la Vega de Granada*. PhD thesis, Universidad de Granada, 1997.
- [28] Z.L. Carroll and M.A. Oliver. Exploring the spatial relations between soil physical properties and apparent electrical conductivity. *Geoderma*, 128:354–374, 2005.
- [29] J. Porta Casanellas, M. López-Acevedo Reguerín, and C. Roquero de Laburu. *Edafología. Para la agricultura y el medio ambiente*. Madrid-Barcelona-México, 2003.
- [30] A. Castillo Martín. *Estudio hidroquímico del acuífero de la Vega de Granada*. PhD thesis, Universidad de Granada, 1986.
- [31] A. Castillo Martín. Cambios en la distribución de fluoruro en el acuífero de la vega de granada; consideraciones hidrodinámicas. *Geogaceta*, 37:103–106, 2004.
- [32] A. Castillo Martín and A. Ramos Cormenzana. Sobre la contaminación microbiana del acuífero de la vega de granada. *Geogaceta*, 32:191–194, 2002.
- [33] J. Catalán Lafuente. *Química del agua*. España, 1981.
- [34] P. S. Jr. Chavez. An improved dark-object substration technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24(3):459–479, 1988.
- [35] P. S. Jr. Chavez. Radiometric calibration of landsat thematic mapper multispectral images. *Photogrammetric Engineering & Remote Sensing*, 55(9):1285–1294, 1989.
- [36] M. Chica-Olmo, F. Abarca-Fernández, and J.P. Rigol-Sánchez. Development of a decision support system based on remote sensing and gis techniques for gold-rich area identification in se spain. *International Journal of Remote Sensing*, 2002.
- [37] M. Chica Olmo, I. Carpintero Salvo, M. J. García Soldado, J. A. Luque Espinar, E. Pardo Igúzquiza, and J.P. Rigol Sánchez. Una aproximación geoestadística al análisis espacial de la calidad del agua. *Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, 5:1578–5157, 2005.
- [38] M. Chica-Olmo and J. A. Luque-Espinar. Applications of the local estimation of the probability distribution function in environmental sciences by kriging methods. *Inverse Problems*, 18:25–36, 2002.

-
- [39] M. Chica-Olmo and J. A. Luque-Espinar. *Los Sistemas de información geográfica en la gestión de los riesgos geológicos y el medio ambiente*, chapter Interpolación espacial en la creación de cubiertas temáticas en S.I.G, pages 181–198. IGME, Madrid, 2002.
- [40] M. Chica-Olmo and J. A. Luque-Espinar. Creación de mapas de calidad de aguas subterráneas mediante métodos de krigeaje. *Boletín Geológico y Minero*, 114:299–310, 2003.
- [41] E. Chuvieco Salinero. *Fundamentos de Teledetección Espacial*. Rialp, 1996.
- [42] E. Chuvieco Salinero. *Teledetección ambiental*. Ariel Ciencia., 2002.
- [43] J. Cihlar, L. Laurent, and J.A. Dyer. Relation between the normalized difference vegetation index and ecological variables. *Remote Sensing of Environment*, 35:279–298, 1991.
- [44] M. Civita and C. De Regibus. Sperimentazione di alcune metodologie per la valutazione della vulnerabilità degli acquiferi. *Q Geol Appl Pitagora Bologna*, 3:63–71, 1995.
- [45] V. Clotier, R. Lefebvre, R. Therrien, and M. Savard. Multivariate statistical analysis of geochemical data as indicative of the hydro geochemical evolution of groundwater in a sedimentary rock aquifer system. *Journal of Hydrology*, 2008.
- [46] O. Cerdón, F. Herrera, F. Hoffmann, and L. Magdalena. *Genetic fuzzy systems: evolutionary tuning and learning of fuzzy knowledge bases*. World Scientific, Singapur, Singapur., 2001.
- [47] E. Cox. *The fuzzy systems handbook: A practitioner's Guide to Building, Using and Maintaining Fuzzy Systems*. Academic Press, 1994.
- [48] N.A.C. Cressie. *Statistics for Spatial Data*. John Wiley & Sons, New York, USA, 1993.
- [49] P. Curran. Remote sensing of foliar chemistry. *Remote Sensing of Environment*, 30:271–278, 1989.
- [50] E. Custodio and M.R. Llamas. *Hidrología subterránea*. Ediciones Omega, S.A., 1996.
- [51] C.J. Date. *An Introduction to Database Systems*. Addison-Wesley, 1995.
- [52] J.C. Davis. *Statistics and Data Analysis in Geology*. John Wiley & Sons Inc., 1986.

-
- [53] Jean Pierre Delhomme. *Applications de la théorie des variables régionalisées dans les sciences de l'eau*. PhD thesis, Université Pierre et Marie Curie, 1976.
- [54] C.V. Deutsh and A.G. Journel. *WGLIB. Geoestatistical Software Library and User's Guide*. Oxford University Press., 1992.
- [55] S. H. Dinius. Social accounting system for evaluating water resources. *Water Resources Research*, 8:1159–1177, 1972.
- [56] S. H. Dinius. Desing of an index of water quality. *Water Resources Bulletin*, 23:833–843, 1987.
- [57] B. Dixon. Applicability of neuro-fuzzy techniques in predicting ground-water vulnerability: A GIS-based sensitivity analysis. *Journal of Hydrology*, 309:17–38, 2005.
- [58] B. Dixon. Groundwater vulnerability mapping: A GIS and fuzzy rule based integrated tool. *Applied Geography*, 25:327–347, 2005.
- [59] B. Dixon, H.D. Scott, J.C. Dixon, and K.F. Steele. Prediction of aquifer vulnerability to pesticides using fuzzy rule-based models at the regional scale. *Physical Geography*, 23:130–153, 2002.
- [60] N. Doerfliger, P.Y. Jeannin, and F. Zwahlen. Water vulnerability assessment in karst environments a new method of defining protection areas using a multiattribute approach and gis tools (EPIK method). *Environmental Geology*, 39:165–176, 1999.
- [61] N. Doerfliger, P.Y. Jeannin, and F. Zwahlen. Water vulnerability assessment in karst environments a new method of defining protection areas using a multiattribute approach and GIS tools (EPIK method). *Environmental Geology*, 39:165–176, 1999.
- [62] J. Dojlido, J. Raniszewski, and J. Woyciechowska. Water quality index-Application for rivers in the Vistula River basin in Poland. *Water Sci. Technol.*, 30:57–64, 1994.
- [63] S. A. Drury. *Image interpretation in Geology*. Chapman & Hall, 1993.
- [64] P. Duchaufour and B. Souchier. *Edafología. Edafogénesis y clasificación*. Masson S.A., 1984.
- [65] A. Earnshaw and N. Greenwood. *Chemistry of the Elements*. Butterworth-Heinemann, 1997.

-
- [66] J. Eastman. *IDRISI Kilimanjaro. Guide to GIS and Image Processing*. Clark Labs, Clark University, Worcester Massachusetts, 2003.
- [67] J.R. Eastman, W. Jin, P.A. Kyem, and J. Toledano. *GIS and Decision Making*. United Nations Institute for Training and Research (UNITAR), 1993.
- [68] K. El Mabrouki. *Impacto del riego con aguas residuales sobre el suelo de la zona no saturada de la vega de Granada*. PhD thesis, Universidad de Granada, 2003.
- [69] ESRI. *ModelBuilder for ArcView Spatial Analyst 2.*, 2000.
- [70] I.M. Farnham, K.H. Johannesson, A.K. Singh, V.F. Hodge, and K.J. Stetzenbach. Factor analytical approaches for evaluating groundwater trace element chemistry data. *Analytical Chimica Acta*, 490:123–138, 2003.
- [71] I.M. Farnham, K.J. Stetzenbach, A.K. Singh, and K.H. Johannesson. Deciphering groundwater flow system in Oasis Valley, Nevada, using trace element chemistry, multivariate statistics, and geographical information system. *Mathematical Geology*, 32:943–968, 2000.
- [72] M. Fischer and P. Nijkamp. Geographic Information Systems and Spatial Modelling: potentials and bottlenecks. In *Proceedings EGIS 92 third European conference on Geographical Information Systems*, 1992.
- [73] M.M. Fisher, H. Scholten, and D. Unwin. *Spatial analytical perspectives on GIS in environmental and socioeconomic sciences*. Taylor & Francis, 1996.
- [74] S. Foster. *Vulnerability of soil and groundwater to pollutants*, chapter Fundamental concepts in aquifer vulnerability, pollution risk and protection strategy, pages 69–86. The Hague, 1987.
- [75] A.S. Fotheringham and M. Wegener. *Spatial models and GIS. New potential and new models*. Taylor & Francis, 2000.
- [76] R. Fullér. *Introduction to neuro-fuzzy systems*. Springer-Verlag, Heidelberg, Alemania, 2000.
- [77] A. Gemitzi, C. Petalas, V.A. Tsihrintzis, and V. Pisinaras. Assessment of groundwater vulnerability to pollution: a combination of GIS, fuzzy logic and decision making techniques. *Environmental Geology*, 49:653–673, 2005.
- [78] C. Güler, G.D. Thyne, J.E. McCray, and A.K. Turner. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeology Journal*, 10:455–474, 2002.

-
- [79] M. Gómez-Delgado and J.I. Barredo-Cano. *Sistemas de Información Geográfica y evaluación multicriterio en la ordenación del territorio*. RA-MA, 2005.
- [80] F. Gümrah, B. Öz, B. Güler, and S. Evin. The application of artificial neural networks for the prediction of water quality of polluted aquifer. *Water, Air, & Soil Pollution*, 119:275–294, 2004.
- [81] R.C. Gogu and A. Dassargues. Current trends and future challenges in groundwater vulnerability assessment using overlay and index methods. *Environmental Geology*, 39:549–559, 2000.
- [82] R.C. Gogu, V. Hallet, and A. Dassargues. Comparison of aquifer vulnerability assessment techniques. application to the ne´blon river basin (belgium). *Environmental Geology*, 44:881–892, 2003.
- [83] D.E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 1989.
- [84] D.D. Golobocanin, B.D. Skrbic, and N.R. Miljevic. Principal component analysis for soil contamination with PAHs. *Chemometrics and Intelligent Laboratory Systems*, 72:219–223, 2004.
- [85] M. Goodchild. A spatial analytical perspective on Geographical Information Systems. *International Journal of Geographical Information Systems*, 1:327–334, 1987.
- [86] M. Goodchild. *Social Sciences for a Digital World: Building Infrastructure and Databases for the Future*, chapter New horizons for the social sciences: Geographic Information Systems, pages 163–172. Organisation fo Economic Cooperation and Development, 2000.
- [87] M.F. Goodchild, R. Haining, and R. Wise. Integrating GIS and spatial data analysis: problems and possibilities. *International Journal of Geographical Information Systems*, 6:405–423, 1992.
- [88] S. Goonatilake and P. Treleaven, editors. *Intelligent Systems for Finance and Business*. John Wiley & Sons, Inc., 1995.
- [89] R. Haining. *Spatial Data Analysis in the Social Environmental Sciences*. Cambridge University Press, Cambridge, UK, 1990.
- [90] S. Haykin. *Neural Networks, a Comprehensive Foundation*. Macmillan College Publishing Company, 1st edition, 1994.

- [91] D. Hebb. *The Organization of Behaviour*. Wiley, 1949.
- [92] B. Helena, R. Pardo, M. Vega, E. Barrado, J.M. Fernández, and L. Fernandez. Temporal evolution of groundwater composition in an alluvial aquifer (pisuerga river, spain) by principal component analysis. *Water Resources*, 34:807–816, 2000.
- [93] F. Herrera and J. L. (Eds.) Verdegay. *Genetic algorithms and soft computing*. Physica-Verlag, Heidelberg, Alemania., 1996.
- [94] J. Hertz, A. Krogh, and R. Palmer. *Introduction to the theory of Neural Computations*. Addison-Wesley, 1991.
- [95] K. Hirota, editor. *Industrial applications of fuzzy technology*. Springer-Verlag, Heidelberg, Alemania, 1993.
- [96] M. Hixson, D. Scholz, N. Fuhs, and T.A. Kiyama. Evaluation of several schemes for classification of remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, 46:1547–1553, 1980.
- [97] Brent N. Holben. Characteristics of maximum-value composite images from temporal avhrr data. *International Journal of Remote Sensing*, 7:0143–1161, 1986.
- [98] R. K. Horton. An index number for rating water quality. *J. Water Pollut. Control Fed.*, 37:300–306, 1965.
- [99] P. Isasi Viñuela and I. Galván León. *Redes neuronales artificiales : un enfoque práctico*. Pearson Educación, 2004.
- [100] ITGE. *Serie manuales utilización de acuíferos*. Vega de Granada., itge edition, 1989.
- [101] P. Jankowski. Integrating geographical information systems and multiple criteria decision-making methods. *International Journal of Geographical Information Science*, 9:251–273, 1995.
- [102] H. Jiang and J.R. Eastman. Application of fuzzy measures in multi-criteria evaluation in gis. *International Journal of Geographical Information Science*, 14:173–184, 2000.
- [103] A.G. Journel and CH.J. Huijbregts. *Mining Geostatistics*. London: Academic Press, 1978.

-
- [104] J.M. Landwehr, R.A. Deininger, N.L. McClelland, and R.M. Brown. Discusión of an objective water quality index, by ralph d. hakins. *J. Water Pollut. Control Fed.*, 46:1804–1809, 1974.
- [105] D.H. Levereti. Prevalence of dental fluorosis in fluoridated and nonfluoridated communities-a preliminary investigation. *J Public Health Dent*, 46:184–187, 1986.
- [106] J. Li, G.H. Huang, G. Zeng, I. Maqsood, and Y. Huang. An integrated fuzzy-stochastic modeling approach for risk assessment of groundwater contamination. *Journal of Environmental Management*, 82:173–188, 2007.
- [107] T.M. Lillesand. *Remote Sensing and image interpretation, 4th ed.* Jonh Willey and Sons, New York, USA, 1999.
- [108] J.A. López Geta. *Tecnología de la intrusión de agua de mar en acuíferos costeros : países mediterráneos : estado de la intrusión de agua de mar en los acuíferos costeros del Mediterráneo y técnicas de evaluación.* Instituto Geológico y Minero de España, 2003.
- [109] C.A. Lucho-Constantino, M. Álvarez Suárez, M.I. Beltrán-Hernández, F. Prieto-García, and H.M. Poggi-Varaldo. A multivariate analysis of the accumulation and fractionation of major and trace elements in agricultural soils in hidalgo state, mexico irrigated with raw wastewater. *Environment International*, 31:313–323, 2005.
- [110] J. A. Luque Espinar. *Análisis Geoestadístico Espacio-Temporal de la variabilidad piezométrica. Aplicación a la Vega de Granada.* PhD thesis, Universidad de Granada., 2001.
- [111] J. A. Luque Espinar, M. Chica Olmo, E. Pardo Igúzquiza, and M. J. García Soldado. Climatological cycles in ground-water levels in detritic aquifer. *Journal of Geological Society of London*, (EN PRENSA), 2007.
- [112] J. A. Luque-Espinar, E. Pardo Igúzquiza, and M. Chica-Olmo. Análisis de ciclicidades de series hidrogeológicas en el estudio de los procesos de recarga. *VI Simposio del Agua en Andalucía*, pages 451–458, 2005.
- [113] J.A. Luque-Espinar, M. Chica-Olmo, E. Pardo-Iguzquiza, and M.J. García-Soldado. Influence of clymatological cycles in groundwater levels. *Journal of Hydrology*, 2007.
- [114] F. Luque Montilla. La contaminación de las aguas subterráneas por actividades urbanas e industriales: La visión de los usuarios. In *La contaminación de las aguas subterráneas: Un problema pendiente.*, 1998.

-
- [115] M.D. Marín Camaches, F. Pachecho Martínez, R. Martínez Gambín, E. Gómez Campoy, and C. Gutierrez Molina. La fluoración en las aguas de la región de murcia. *Revista Española de Salud Pública.*, 72:91–101, 1998.
- [116] R. Marín Galvín. *Fisicoquímica y microbiología de los medios acuáticos. Tratamiento y control de calidad de las aguas*. Ediciones Díaz de Santos, S.A., 2003.
- [117] F. Martín Asín. *Geodesia y Cartografía Matemática*. Madrid, 1990.
- [118] B. Martín del Brío and A. Sanz Molina. *Redes neuronales y sistemas borrosos*. Ra-Ma, 1997.
- [119] C. Martínez Navarrete and A. García García. *Perímetros de protección para captaciones de agua subterránea destinada al consumo humano. Metodología y aplicación al territorio*. Instituto Geológico y Minero de España, 2003.
- [120] A. Melloul and M. Collin. The 'principal components' statistical method as a complementary approach to geochemical methods in water quality factor identification; application to the coastal plain aquifer to israel. *Journal of Hydrology*, 140:49–73, 1992.
- [121] J. Menor Toribio. Transformaciones recientes en la organización territorial de la vega de granada: del espacio agrario territorial a la aglomeración urbana actual. *Estudios regionales*, 48:189–214, 1997.
- [122] J. Menor Toribio. *La vega de Granada: transformaciones recientes de un espacio agrario tradicional*. PhD thesis, Universidad de Granada., 1998.
- [123] J. Mingo Magro. *La vigilancia de la contaminación fluvial*. Serv. Publi MOPU. Madrid, 1981.
- [124] L. Moreno Merino, P. Navarrete Martínez, and L. Virgós Soriano. *Conceptos básicos de la microbiología de las aguas subterráneas*. IGME., 1998.
- [125] A. Muhammetoglu and A. Yardimci. A fuzzy logic approach to assess groundwater pollution levels below agricultural fields. *Environmental Monitoring and Assessment*, 118:337–354, 2006.
- [126] D. Nauck, F. Klawonn, and R. Kruse. *Foundations of neuro-fuzzy systems*. John Wiley & Sons, Nueva York, NY, EE. UU., 1997.
- [127] S. Navarro Blaya and G. Navarro García. *Química Agrícola. El suelo y los elementos químicos esenciales para la vida vegetal*. Ediciones Mundi-Prensa, 2000.

- [128] Nils J. Nilsson. *Inteligencia artificial : una nueva síntesis*. McGraw-Hill Interamericana de España, 2001.
- [129] V. Nourani, A.A. Mogaddam, and A.O. Nadiri. An ann-based model for spatiotemporal groundwater level forecasting. *Hydrological Processes*, 22:5054–5066, 2008.
- [130] S. Openshaw. A view on the GIS crisis in geography, or using GIS to put humpty dumpty back together again. *Environment and Planning A*, 23:621–628, 1991.
- [131] S. Openshaw and C. Openshaw. *Artificial Intelligence in Geography*. John Wiley and Sons Ltd, 1997.
- [132] D. O’Sullivan and D. Unwin. *Geographic Information Analysis*. London Wiley, 2002.
- [133] M. Otto. *Analytical Chemistry*, chapter Multivariate methods, page 916. Kellner R., Mermet J.M., Otto M., Widmer H.M., 1998.
- [134] F. Padilla, J. Benavente, and J. Cruz-Sanjulián. Simulación de diferentes alternativas de gestión de los recursos hídricos del acuífero costero del río verde (almuñecar, españa). *Estudios Geológicos*, 53:173–182, 1997.
- [135] M. Painho, A. Vasilakos, F. Bacao, and W. Pedrycz. Exploring spatial data through computational intelligence: A joint perspective. *Soft Computing*, 9:326–331, 2005.
- [136] M.A. Pardo, A. y Ruíz. *SPSS 11. Guía para el análisis de datos*. Mc Graw Hill, 2002.
- [137] E. Pardo-Igúzquiza. Mlreml: a computer program for the inference of spatial covariance parameters by maximum likelihood and restricted maximum likelihood. *Computers and Geosciences*, 23:153–162, 1997.
- [138] E. Pardo-Igúzquiza. Maximum likelihood estimation of spatial covariance parameters. *Mathematical Geology*, 30:95–108, 1998.
- [139] E. Pardo-Igúzquiza. Optimal selection of number and location of rainfall gauges for areal rainfall estimation using geostatistics and simulated annealing. *Journal of Hydrology.*, 210:206–220, 1998.
- [140] E. Pardo-Igúzquiza and M. Chica-Olmo. Estimation of gradients from sparse data by universal kriging. *Water Resources Research*, 40:1–17, 2004.

-
- [141] E. Pardo-Igúzquiza and M. Chica-Olmo. Krigradi: a cokriging program for estimating the gradient of spatial variables from sparse data. *Computers and Geosciences*, 33:497–512, 2007.
- [142] G.E Pavlikakis and V.A. Tsihrintzis. A quantitative method for accounting human opinion, preferences and perceptions in ecosystem management. *J. Environ. Manage.*, 68:193–205, 2003.
- [143] W. Pedrycz. *Fuzzy modelling: paradigms and practice*. Kluwer Academic, Norwell, 1996.
- [144] J.G.A. Pei-Yu Chen, M.V. Raghavan Srinivasan, and A. Peter M. Surveying ground water level using remote sensing: An example over the seco and hondo creek watershed in texas. *Ground Water Monitoring & Remediation*, 26:94–102, 2006.
- [145] M. Poch. *Las calidades del agua*. Rubes, 2003.
- [146] D. Pozo-Vázquez, M.J. Esteban-Parra, F.S. Rodrigo, and Y. Castro-Díez. An analysis of the variability of the north atlantic oscillation in the time and the frequency domains. *International Journal of Climatology*, 20:1675–1692, 2000.
- [147] A. Prueb. Action values mobile (nh4-no3-extractable) trace elements in soils based on german national standard din 19730. In *Contaminated Soils 3rd International Conference on Biochemistry of Trace elements*, pages 415–423, Paris, 1997.
- [148] A. Rashid, A. Aziz, and K-F. V. Wong. A neural network approach to the determination of aquifer parameters. *Ground Water*, 30:164–166, 1992.
- [149] I. Requena Sánchez. Neurocomputación (redes neuronales artificiales). Universidad de Granada.
- [150] K.M. Reynolds. *EMDS Users Guide (Version 2.0): Knowledge-based Decision Support for Ecological Assessment.*, 1998.
- [151] J.A. Richards. *Remote Sensing Digital Image Analysis*. Springer-Verlag, Berlin, 1999.
- [152] J.P. Rigol Sánchez. *Aplicación de Sistemas de Información Geográfica y Teledetección en Exploración Minera*. PhD thesis, Universidad de Granada, 2000.
- [153] X. Rodo, E. Baert, and F.A. Comin. Variations in seasonal rainfall in southern europe during the present century: relationships with the north atlantic oscillation and the niño-southern oscillation. *Climate Dynamics*, 13:275–284, 1997.

- [154] Z. Rodríguez Cotilla. Teoría de la decisión multicriterio: un enfoque para la toma de decisiones. *Economía y Desarrollo.*, 126:40–57, 2000.
- [155] S. J. Rogers, H. C. Chen, D.C. Kopaska-Merkel, and J. H. Fang. Predicting permeability from porosity using artificial neural network. *AAPG Bulletin*, 79:1786–1797, 1995.
- [156] F. J. Roldán García. *Exploración Geológico Minera de la Cuenca Lignitífera de Granada.*, chapter Mapa Geológico de la Cuenca de Granada., pages 14–36 pp. IGME, 1981.
- [157] F. Rosenblatt. *Principles of Neurodynamics*. Spartan Books, Nueva York, 1962.
- [158] O. Roussel, A. Cavelier, and H.M.G Van der Werf. Adaptation and use of a fuzzy expert system to assess the environmental effect of pesticides applied to field crops. *Agriculture Ecosystems & Environment*, 80:143–158, 2000.
- [159] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. *Parallel distributed processing*, chapter Learning internal representations by error propagation, pages 318–362. MIT Press Cambridge, 1986.
- [160] T.L. Saaty. A scaling method for priorities in hierarchical structures. *Journal of Math Psychol*, 15:234–281, 1977.
- [161] F. F. Sabins. *Remote Sensing: Principles and interpretation*. W. H. Freeman and Company, 1996.
- [162] F. J. Samper-Calvete and J. Carrera Ramírez. *Geoestadística. Aplicaciones a la hidrología subterránea*. CIMNE, 1990.
- [163] P.P. Schot and J. van der Wal. Human impact on regional groundwater composition through intervention in natural flow patterns and changes in land use. *Journal of Hydrology*, 134:297–313, 1992.
- [164] K.V. Schowengerdt. *Remote Sensing: Models and Methods for Image Processing*. Academic Press, Boston, USA, 1997.
- [165] R.A. Schowengerdt. *Techniques for Image Processing and Classification in Remote Sensing*. Academic Press, Boston, USA, 1983.
- [166] M.M. Sena, R.T.S. Frighetto, P.J. Valarini, H. Tokeshi, and R.J. Poppi. Discrimination of management effects on soil parameters by using principal component analysis: a multivariate analysis case study. *Soil and Tillage Research*, 67:171–181, 2002.

- [167] H.S. Sichel. An experimental and theoretical investigation of bias error in mine sampling with special reference to narrow gold reefs. *Trans. Inst. Min. Metall.*, 56:403–273, 1947.
- [168] D. G. Smith. A better water quality indexing system for rivers and streams. *Water Res.*, 24:1237–1244, 1990.
- [169] M.I. Sánchez and E. Chuvieco. Estimación de la evapotranspiración del cultivo de referencia, eto, a partir de imágenes noaa-avhrr. *Revista Española de Teledetección*, 14:11–21, 2000.
- [170] L. Sánchez-Díaz and A. Castillo-Martín. Distribución del ión cloruro en el acuífero de la vega de granada; consideraciones hidrodinámicas. *Hidrogeología y Recursos Hidráulicos*, 27:185–194, 2004.
- [171] R.K. Steinhort and R.E. Williams. Discrimination of groundwater sources using cluster analysis, manova, canonical analysis and discriminant analysis. *Water Resources Research*, 21:1149–1156, 1985.
- [172] K.J. Stetzenbach, V.F. Hodge, I.M. Farnham, and K.H. Johannesson. Geochemical and statistical evidence of deep carbonate groundwater within averlying volcanic rock aquifers/aquitards of southern nevada, usa. *Journal of Hydrology*, 243:254–271, 2001.
- [173] M. Story and J.B. Campbell. The effect of training data on classification accuracy. *ACSM-ASPRS convention*, 4:370–379, 1986.
- [174] M. Sugeno and T. Yasukawa. A fuzzy-logic-based approach to qualitative modeling. *IEEE Transactions on Fuzzy Systems*, 1:7–31, 1993.
- [175] Prabhata K. Swamee and Aditya Tyagi. Describing water quality with aggregate index. *Journal of Environmental Engineering*, 126:88–91, 2000.
- [176] Prabhata K. Swamee and Aditya Tyagi. Improved method for aggregation of water quality subindices. *Journal of Environmental Engineering*, pages 220–225, 2007.
- [177] T. Takagi and M. Sugeno. Fuzzy identification of systems and its application to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, 15:116–132, 1985.
- [178] P. E. Thomas G. Reeves. *Calidad y tratamiento del agua*, chapter Fluoración del agua, pages 977–997. España, 2002.

-
- [179] D. Tomlin. *Geographic Information Systems and Cartographic Modeling*. Prentice Hall, 1990.
- [180] J. Turkey. *Exploratory Data Analysis*. Addison-Wesley, Reading Massachussets, USA, 1977.
- [181] D.J. Unwin. *Introductory spatial analysis*. Methuen, London, UK., 1981.
- [182] E.J. Usunoff and A. Guzmán-Guzmán. Multivariate analysis in hydrochemistry: An example of the use of factor and correspondence analysis. *Ground Water*, 27:27–34, 1989.
- [183] H.M.G. Van der Werf and C. Zimmer. An indicator of pesticide environmental impact based on a fuzzy expert system. *Chemosphere*, 36:2225–2249, 1998.
- [184] M. Vega, R. Pardo, E. Barrado, and L. Debn. Assesment of seasonal and polluting effects on the quality of river water by exploratory data análisis. *Water Resources*, 32:3581–3592, 1998.
- [185] H. Voogd. *Multi-criteria evaluations for urban and regional planning*. London Princeton Univ., 1983.
- [186] T.M. Walski and F. L. Parker. Consumers water quality index. *Journal Environmental Engineering Division*, 100:593–611, 1974.
- [187] J. Walters and N. R. Nielsen. *Crafting Knowledge-Based Systems*. Wiley, 1988.
- [188] L. X. Wang. *Adaptive fuzzy systems and control, design and stability analysis*. Prentice Hall, Engleweed Cliffs, Engleweed Cliffs, NJ, EE. UU, 1994.
- [189] L. X. Wang and J. M. Mendel. Generating fuzzy rules by learning from examples. *IEEE Transactions on Systems, Man, and Cybernetics*, 22:1414–1427, 1992.
- [190] M.X. Wang, G.D. Liu, W.L. Wu, Y.H. Bao, and W.N. Liu. Prediction of agriculture derived groundwater nitrate distribution in north china plain with gis-based bpnn. *Environmental Geology*, 50:637–644, 2006.
- [191] P. Werbos. *Beyond regression: New tools for prediction and analysis in the behavioral sciences*. PhD thesis, Harvard University, 1974.
- [192] C. Westman. *Imaging Spectrometry, a tool for Environmental Observations*, chapter Estimating canopy biochemistry through imaging, pages 57–69. J. Hill and J. Megier, 1994.

- [193] Widrow and M. E. Hoff. Adaptive switching circuits. In *IRE WESCON Convention Record*, 1960.
- [194] M. F. Worboys. *GIS: A Computing Perspective*. Taylor & Francis, 1995.
- [195] Albert K. Yeung. *Data Organization and Structure, NCGIA Core Curriculum in GIScience*. <http://www.ncgia.ucsb.edu/giscc/units/u051/>, 1998.
- [196] L. A. Zadeh. Outline of a new approach to the analysis of complex systems and desision processes. *IEEE Transactions on Systems, Man, and Cybernetics*, 1:28–44, 1973.
- [197] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
- [198] P.J. Zarco-Tejada, J.R. Miller, T.L. Noland, G.H. Mohammed, and P.H. Sampson. Scaling-up and model inversion methods with narrowband optical indices for clorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 39:1491–1507, 2001.