# Assessing the impact of climate variability on seasonal streamflow forecasting in the Iberian Peninsula

Memoria presentada para optar al grado de

Doctor en Física

por:

Jose Manuel Hidalgo Muñoz

Directoras:

Dra. Yolanda Castro Díez

Dra. María Jesús Esteban Parra

Dra. Sonia Raquel Gámiz Fortis

Programa de Doctorado en Física
Departamento de Física Aplicada
Facultad de Ciencias
Universidad de Granada

Editor: Universidad de Granada. Tesis Doctorales
Autor: José Manuel Hidalgo Muñoz
ISBN: 978-84-9125-049-4
URI: http://hdl.handle.net/10481/39895

# Assessing the impact of climate variability on seasonal streamflow forecasting in the Iberian Peninsula

Directoras:

**Dra. Yolanda Castro Díez**
Catedrática
Universidad de Granada

**Dra. María Jesús Esteban Parra**
Profesora Titular
Universidad de Granada

**Dra. Sonia Raquel Gámiz Fortis**
Profesora Titular
Universidad de Granada

Memoria presentada para optar al grado de:

**Doctor en Física**

Por

**Jose Manuel Hidalgo Muñoz**
Licenciado en Física

Granada, enero de 2015

# Agradecimientos

Una tesis doctoral no solo engloba fases de documentación, desarrollos metodológicos y obtención de resultados, sino que también hace acopio de una parte importante de tu vida, gente que te rodea y lugares que visitas en el trascurso de la misma. Se trata de una etapa de crecimiento profesional y personal, que no hubiese sido posible sin la oportunidad que me brindaron mis directoras: Yolanda, María Jesús y Sonia, a las que agradezco enormemente su dedicación, apoyo y guía, su confianza en mi, y su abrazo de ánimo en los momentos más complicados.

Ahora que esta etapa llega a su fin, echo la vista atrás y recuerdo mis primeros momentos en el mundo de la investigación, mis primeros compañeros de despacho: Andy, Borja, Penélope y Dani. Con ellos he compartido muy buenos ratos, cafés y partidos de fútbol. Dani, mi hermano mayor en este mucho de la investigación (con todo lo que conlleva ser hermano mayor, ¡regañinas incluidas!), que tanto me ha ayudado, y que me presentó a quien a la postre se convertiría en uno de mis grandes pilares durante esta etapa, el señor *Google*. Cómo te eché de menos cuando te fuiste. Y aún sigo haciéndolo.

Muchos otros compañeros han ido pasando por el despacho, como Enrique, gran investigador, mejor persona, con el que hemos pasado grandes ratos en el despacho y en el 'joyo'. Otros que han pasado por allí y han redirigido sus caminos profesionales, como Dani Calandria (que nos echó un gran cable con MATLAB!) y cómo no, Maria. Qué de ratos hemos pasado juntos, lo que hemos maldecido, y esas charlas sobre temas de rabiosa actualidad. ¡Moltes gràcies! (tan pronto como dejes de moverte voy a visitarte!). Un huequecito aquí se merecen Samir y Reiner, con los que he pasado muy buenos momentos, que me han ayudado

mucho a la hora de darle un buen empujón a esta tesis y que espero que ellos pronto se vean en estas lides. También un recuerdo a otros estudiantes que vinieron a compartir con nosotros unos meses como Clemente y Farid, que espero tengan mucha suerte. Agradecer también a Ana su paciencia conmigo (¡o falta de ella!), con quien he tenido la suerte no solo de compartir ratos en el despacho sino también fuera de él (gracias por darme cobijo), ánimo y tranquila que eres buena y te va a ir bien, y también a Mati, que está empezando su andadura en el mundillo y que espero que tenga mucha suerte y que modele bien. Seguro que sí. Para cualquier cosa, por aquí me tenéis.

Quisiera agradecer a todos aquellos colegas de profesión que han aportado su granito de arena en mi investigación, todos aquellos que he conocido en congresos y cursos. Also, I want to give special thanks at CRU staff and students (Tim, Jonathan, Giorgio, …) with whom I shared really good times in UK and learnt a lot. I would also like to thank the MetOffice team for giving me the opportunity to continue researching.

También quisiera hacer mención a todos aquellos que me han ayudado a recargar pilas después de cada día de trabajo, a Juanma, con quien he compartido tantos años y tantas cosas, a Jose Carlos, Álvaro, Luna, a las peñas de futbol sala con las que he compartido muchos partidos (Oscar, Dani, la gente del IAA), y al grupo de 'Granada x el running', que fueron un gran apoyo para mi durante el último año de tesis, y que me transmitieron el amor por ese deporte. Espero veros en muchas carreras!

Dejo para el final lo más importante, agradecerle a mi familia todo su apoyo y sacrificio para que llegue este momento. A mi madre, que tanto ha luchado por mi. Todo lo que he logrado es gracias a su esfuerzo. A Miguel, que desde donde me esté viendo le mando un fuerte abrazo. A mis abuelos, ¡cuánto os echo de menos! A Bea, que ha estado a mi lado todos estos años, sufriéndome, en especial durante los últimos tiempos de final eterno. ¡Pronto emprenderemos nuevas aventuras! Ídem.

*"Brindemos que hoy es siempre todavía, que nunca me gustaron las despedidas"*

(I. Serrano).

# CONTENTS

# ABSTRACT

The water shortage represents for some regions, in particular for the Iberian Peninsula (IP), one of the most severe limiting factors to maintain a sustainable development, since water resources play a crucial role in various socio-economic and environmental needs, such as agriculture, industry, hydropower industry, and tourism sector. The problem of water scarcity is likely to become more severe according to the projected decrease/increase in water availability/demand. Then, the motivation of this dissertation relies on the necessity to improve the understanding of the large-scale climate variability that drives the streamflow variability on the IP, which becomes the basis for developing streamflow forecast at scales that are of paramount importance for reservoir operations and irrigation management decisions, protection of the environment or in the reduction of expenses in flood and drought mitigation.

During the first part of this work, an analysis of the streamflow database was performed. A process to identified stations with inhomogeneities in data series, mainly derived from the regulation processes, was carried out using a combined methodology based on Pettitt test and Common Area Index over original database of 1380 gauging stations. As result, a total 382 stations were selected for this study, covering the period from October 1975 to September 2008. Also, the main spatial and temporal characteristics of streamflow variability in the IP were described.

The second part of this Thesis consisted on identifying the main climate factors that have a noteworthy influence on near future (lagging from one to four seasons) seasonal streamflow variability of the IP Rivers and provides an insight into the possible mechanisms and physical processes behind these relationships. Firstly, teleconnection indices, which represent most of the

dominant sources of climate variability, were evaluated as potential predictors. Secondly, Singular Value Decomposition (SVD) technique was employed to identify and isolate the main modes of covariability between seasonal streamflow and the climate variables (sea surface temperature, geopotential height at 500 hPa in Northern Hemisphere and global temperature and precipitation) that precede it from one to four seasons. Once the main climatic predictors were identified, predictions based on them were conducted. A leave-one-out cross-validation approach based on a multiple linear regression approach that combining Variance Inflation Factor and Stepwise Backward selection was used to avoid multicollinearity and select the best subset of predictors. The forecasting methodology was developed for four forecasting scenarios, related to the number of seasons prior to which the forecasting is made, from one year (4S scenario) until one season (1S scenario) in advance, updating and improving the predictions seasonally. The correlation coefficient (RHO), Root Mean Square Error Skill Score (RMSESS) and the Gerrity Skill Score (GSS) were used to evaluate the forecasting skill.

For the predictions made based on teleconnection indices, in case of autumn streamflow, good forecasting skill (RHO > 0.5, RMSESS > 20%, GSS > 0.4) was found for a third of the stations located in the Mediterranean Andalusian Basin, the North Atlantic Oscillation of the previous winter being the main predictor. Also, fair forecasting skill (RHO > 0.44, RMSESS > 10%, GSS > 0.2) was found in stations in the northwestern IP (16 of these located in the Douro and Tagus Basins) with two seasons in advance. For winter streamflow, fair forecasting skill was found for one season in advance in 168 stations, with the Snow Advance Index as the main predictor. Finally, forecasting was poorer for spring streamflow than for autumn and winter, since only 16 stations showed fair forecasting skill in with one season in advance, particularly in the northwestern of IP.

The use of SVD improved the forecasting skills of the autumn streamflow, in particular relevant for 3S scenario. In this case, up to 42 stations present fair forecasting skills (RHO > 0.44, RMSESS > 10%, and GSS > 0.2), particularly in the Mediterranean Andalusian Basin, with Zdjf2 and Adjf2 (related to winter NAO) as predictors, but also in the north-northwestern IP (Douro, Miño-Sil, Cantabrian and upper Ebro Basins), being Rson1, Adjf1 and Rdjf1 (linked to ENSO) the main predictors. Also, a refinement in spring streamflow forecasting is observed, specially in 3S scenario, when values of RHO > 0.44, RMSESS > 10%, and GSS > 0.2 are obtained in 21 stations, mainly locate in the northeastern quadrant of IP. In those cases Pjja1 (linked to summer ENSO phenomenon), Zjja1 and Tjja1 (which could be associated with the summer Northern Annular Mode and with criosphere variability) were used as predictors.

Conversely, winter streamflow was not forecasted in almost any stations. In this case, SAI appeared to be the only reliable predictor.

In conclusion, this Thesis presents a valuable contribution to the studies regarding seasonal streamflow forecasting of the IP Rivers. Some distinguishing features are that it relies on a long, complete, reliable and spatially well-distributed streamflow database, which enables to describe with a high spatial resolution the potential use of different climate signals as predictors of seasonal streamflow in different areas of the IP, which become very useful for making local decision in water resources management. Also, it explores the links between climate signal and streamflow variability of the IP Rivers, not only evaluating the most commonly used climate indices but also exploring further relationships between climate variability and streamflow in the following seasons. Finally, this study can provide a more comprehensive view of relationship of climate variability and streamflow on seasonal timescales in a way that can significantly contribute to streamflow forecasting purposes (with various forecasting schemes, according to the time in advance the predictions are made), providing the option of developing water-management policies some seasons in advance and with the possibility of modifying or adjusting these strategies as the predictions are updated.

# RESUMEN

La escasez de agua representa para algunas regiones, en especial para la Península Ibérica (PI), uno de los factores limitantes más severos para poder mantener un desarrollo sostenible, debido a que los recursos hídricos juegan un papel fundamental en la satisfacción de las necesidades socioeconómicas y ambientales, tales como la agricultura, industria, la industria hidroeléctrica y el sector turístico. Asimismo, el problema de escasez del recurso hídrico es probable que se agrave en el futuro, de acuerdo a las proyecciones de disminución y de la disponibilidad de agua y el aumento en la demanda. En consecuencia, la motivación de este trabajo de Tesis radica en la necesidad de avanzar en el entendimiento de la variabilidad climática relacionada con la variabilidad del caudal de los ríos ibéricos, lo cual se convierte en la base para la predicción del caudal a escalas temporales de gran importancia para la gestión de embalses, manejo de regadíos, protección medioambiental o mitigación de los daños provocados por sequías e inundaciones.

Durante la primera parte de este trabajo se llevo a cabo un profundo análisis de la base de datos de caudales disponible. En particular, se realizó un estudio para identificar posibles inhomogenidades en las series de datos, las cuales derivan principalmente de la actuación de procesos de regulación de caudales (coincidiendo muchas de ellas con puestas en marcha de embalses en las cabeceras de determinados ríos). La metodología llevada a cabo para identificar estas inhomogeneidades consistió en una combinación del test de Pettitt y de un test que analiza el porcentaje de área común que encierra las curvas del caudal intraanual antes y después del posible punto de ruptura de la serie. Como resultado, de la base de datos original formada por 1380 estaciones de medida de caudal, 382 fueron seleccionadas como adecuadas para este

estudio, cumpliendo, además del criterio de homogeneidad utilizado, que cubran el periodo temporal 1975-2008 con menos de un 10% de datos faltantes.

La segunda parte de este trabajo consistió en identificar los factores climáticos más importantes que influyen en la variabilidad del caudal estacional de los ríos ibéricos en un futuro cercano (entre 1 y 4 estaciones), así como en discutir sobre los posibles mecanismos físicos que hay detrás de estas relaciones. En primer lugar se evaluó la potencial predictibilidad de los índices de teleconexión, los cuales son una representación de los principales modos de variabilidad climática. En segundo lugar, es utilizó la técnica de descomposición del valor singular (SVD, por sus siglas en inglés) para determinar los principales modos de variabilidad acoplada entre el caudal estacional y una serie de variables climáticas (temperatura de la superficie del mar, altura geopotencial a 500 hPa en el Hemisferio Norte, y valores de temperatura y precipitación globales) precediendo al caudal entre 1 y 4 estaciones. Una vez que se identificaron los principales predictores climáticos, se llevó a cabo un ejercicio de predicción del caudal estacional. Para ello se optó por un enfoque de validación cruzada y se utilizaron modelos de regresión lineal múltiple, combinando el test de factor de inflación de la varianza y la selección por pasos para evitar problemas derivados de la correlación excesiva entre predictores y elegir el mejor conjunto de predictores. La metodología de predicción se basa en la creación de 4 escenarios de predicción, de acuerdo a diferentes (hasta 4) numero de estaciones de adelanto con las cuales las predicciones son efectuadas, desde cuatro estaciones (escenario 4S) hasta una estación previa (escenario 1S) al caudal estacional a predecir. El coeficiente de correlación (RHO), el coeficiente de mejora respecto a la climatología de la raíz cuadrada del error cuadrático medio (RMSESS) y el parámetro de Gerrity (GSS) se utilizaron para evaluar la calidad de las predicciones.

Para el caso de predicciones basadas en los índices de teleconexión, y en el caso del caudal de otoño, predicciones de buena calidad (RHO > 0.5, RMSESS > 20%, GSS > 0.4) se encontraron para un tercio de las estaciones de caudal pertenecientes a la Cuenca Mediterránea Andaluza, siendo la Oscilación del Atlántico Norte del invierno previo el principal predictor. Predicciones moderadamente buenas (RHO > 0.44, RMSESS > 10%, GSS > 0.2) se obtuvieron para el escenario 3S en estaciones de caudal localizadas en el cuadrante noroeste de la PI (16 de las cuales estaban localizadas en las cuencas del Duero y Tajo). En el caso del caudal de invierno, moderadas predicciones se encontraron en hasta 168 estaciones de caudal (de hecho en bastantes de ellas se podrían clasificar de buenas) para el escenario 1S, siendo el Índice de Avance de la Nieve del Octubre previo el principal predictor. Finalmente, peores predicciones se

obtuvieron para el caudal de primavera en comparación con el de otoño e invierno, pues solo 16 estaciones mostraron una calidad moderada en las predicciones en el escenario 1S, mayormente ubicadas en el cuadrante noroeste de la PI.

El uso de la metodología basada en SVD mejoró la calidad de las predicciones del caudal de otoño, en particular en el caso del escenario 3S. En este caso se incrementaron hasta 42 las estaciones con moderada predictibilidad (RHO > 0.44, RMSESS > 10%, and GSS > 0.2), sobre todo en la Cuenca Mediterránea Andaluza, siendo Zdjf2 y Adjf2 (relacionados con la Oscilación del Atlántico Norte) los principales predictores, así como en las cuencas del noroeste de la PI (Duero, Miño-Sil, Cantábrica y alto Ebro), siendo Rson1, Adjf1 y Rdjf1 (relacionados con el fenómeno del Niño). Así mismo, las predicciones del caudal de primavera mejoraron, especialmente en el escenario 3S, donde valores de RHO > 0.44, RMSESS > 10%, and GSS > 0.2 se obtuvieron en 21 estaciones de medida, mayormente localizadas en el cuadrante noroeste de la PI. En esos casos, Pjja1 (relacionado al desarrollo de fenómeno del Niño en verano), Zjja1 y Tjja (los cuales podrían estar ligados a la variabilidad del Modo Anular del Norte y la criosfera durante el verano, respectivamente). Por el contrario, en el caso del invierno apenas de encontró capacidad predictiva en casi ninguna estación de caudal, por lo cual el Índice de Avance de la nieve parece ser el único predictor válido en este caso de entre todos los estudiados.

En conclusión, esta Tesis presenta una valiosa contribución a los estudios relacionados con la predicción estacional del caudal de los ríos ibéricos. Algunas de sus características diferenciadoras son el hecho de usar una base de datos de caudales extensa, completa y cuya calidad ha sido comprobada, permitiendo describir con un alto grado de resolución espacial el potencial uso de diferentes señales climáticas como predictores del caudal estacional en diferentes regiones de la PI, lo cual es de gran utilidad para efectuar decisiones a nivel local acerca de gestión de los recursos hídricos. De igual modo, se exploran las relaciones entre señales climáticas y la variabilidad del caudal de los ríos ibéricos, no sólo evaluando los índices climáticos más comunes, sino analizando otras relaciones más allá de las cubiertas por los mismos. Finalmente, este estudio puede proporcionar una visión más completa de la relación entre la variabilidad climática y el caudal estacional de los ríos ibéricos de una forma que pueda ser utilizada con fines predictivos (con varios esquemas de predicción de acuerdo al tiempo de antelación con el cual las predicciones son efectuadas), dando lugar a tener la opción de desarrollar políticas de gestión de recursos hídricos con algunas estaciones de antelación, así como de modificar o ajustarlas a medida que se acerca la estación en cuestión.

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF ACRONYMS

| | |
|---|---|
| AMJ | April-May-June |
| AMO | Atlantic Multidecadal Oscillation |
| AO | Arctic Oscillation |
| ATL | Atlantic Ocean |
| CAI | Common Area Index |
| CCA | Canonical Correlation Analysis |
| CEDEX | Centro de Estudios Hidrográficos |
| COR | Correlation between 'predicted' left and right expansion coefficients |
| CPC | Climate Prediction Center |
| EA | East Atlantic |
| EA-WR | East Atlantic - Western Russia |
| EMI | El Niño Modoki Index |
| ENSO | El Niño-Southern Oscillation |
| EOF | Empirical Orthogonal Function |
| EP-NP | East Pacific - North Pacific |
| GEV | Generalized Extreme Value |
| GLO | Generalized Logistic |

| | |
|---|---|
| GP | Generalized Pareto |
| GSS | Gerrity Skill Score |
| IOD | Indian Ocean Dipole |
| IP | Iberian Peninsula |
| JFM | January-February-March |
| KS | Kolmogorov-Smirnov statistic |
| LN3 | Lognormal 3 parameters |
| MC | Monte Carlo |
| MD | Minimum orthogonal Distance |
| MCA | Maximum Covariance Analysis |
| NAH | North Atlantic Horseshoe |
| NAM | Northern Annual Mode |
| NAO | North Atlantic Oscillation |
| NAT | North Atlantic Tripole |
| NCEP | National Center for Environmental Predictions |
| NPGO | North Pacific Gyre Oscillation |
| OLS | Ordinary Least Squares |
| OND | October-November-December |
| PAC | Pacific Ocean |
| PCA | Principal Components Analysis |
| PCR | Principal Components Regression |
| PDO | Pacific Decadal Oscillation |
| PE3 | Pearson type 3 |
| PLS | Partial Least Squares regression |
| PNA | Pacific North Atlantic |
| RHO | Correlation coefficient |

| | |
|---|---|
| RMSESS | Root Mean Square Error Skill Score |
| RR | Precipitation |
| SAI | Snow Advance Index |
| SCAND | Scandinavian |
| SCF | Square Covariance Fraction |
| SLP | Sea Level Pressure |
| SNHT | Standard Normal Homogeneity Test |
| SNIRH | Sistema Nacional de Informaçao de Recursos Hídricos |
| SST | Sea Surface Temperature |
| STR | Strengh of coupling between fields in SVD |
| TMP | Temperature |
| VIF | Variance Inflation Factor |
| WeMO | Western Mediterranean Oscillation |
| WP | Western Pacific |
| Z500 | Geopotential Height at 500 hPa |

# CHAPTER 1

# INTRODUCTION

In this chapter, the motivation that supports the research carried out in this work is presented, especially in the context of the importance of long-term prediction of the availability of water resources in a region subjected to recurrent water stress. Also, the objectives and outline of the dissertation are described.

## 1.1 Water resources management. A challenging issue

The water shortage represents for some countries one of the most severe limiting factors to maintain a sustainable development, since water resources play a crucial role in various socio-economic and environmental needs, such as agriculture, industry, hydropower industry, and tourism sector. The problem of water scarcity becomes more complex in the Mediterranean areas, where water demand from agriculture and tourism is mostly concentrated during summer, the drier season.

A recent study about changes in climate extremes and their impacts (Seneviratne et al., (2012) found that since the 1950s some regions of the world have experienced trends toward more intense and longer droughts, in particular in southern Europe and West Africa. Also, they found that anthropogenic influence has contributed to some changes in the drought patterns observed in the second half of the 20$^{th}$ century, based on its attributed impact on precipitation

and temperature changes. As European regions regard, Dai et al. (2004) found an increase in dryness for most of the European continent based on Palmer Drought Severity index (PDSI). On the contrary, Lloyd-Hughes and Saunders (2002) and van der Schrier et al. (2006) analysed the variability of Standard Precipitation Index (SPI) and self-calibrating PDSI for the 20th century (for 1901-1999 and 1901-2002, respectively), concluding that no statistically significant changes were observed in extreme and moderate drought conditions in Europe, with the exception of the Mediterranean region in van der Schrier et al. (2006). Sheffield and Wood (2008) also found contrasting dryness trends in Europe, with increases in the southern and eastern part of the continent, but decreases elsewhere. Other authors have also found increases in warm-dry conditions in the central and southern part of Europe (Beniston, 2009; Alexander et al., 2006). Focusing in river discharge, Stahl et al. (2010) investigated streamflow data across Europe and found negative trends (lower streamflow) in southern and eastern regions, and generally positive trends (higher streamflow) elsewhere (especially in northern latitudes).

Consequently, the observed increases of drought events in the Mediterranean region further increase the complexity of water scarcity management over this area. A clear example is Spain, one of the European areas that suffer from hydric stress, where major droughts have severe effects on the agricultural economy and have evidenced the lack of reliability of urban water supply to the majority of the Spanish territory, either due to the lack of regulation (north) or restrictions (south and central). Particularly, Lorenzo Lacruz et al., (2012) found a marked decrease in annual, winter, and spring streamflows in most of the Iberian sub-basins, especially those in the south. In terms of streamflow droughts, Lorenzo-Lacruz et al., (2010) found a marked reduction in water availability in Tagus Basin, related to the more frequent droughts. López-Moreno et al. (2009) examined the effects of a large dam on hydrological droughts in the transboundary Tagus River, central Spain and Portugal, concluding that because of the reservoir construction, the Portuguese part of the basin has experienced more severe droughts than the upstream part, in terms of both magnitude and duration. Furthermore, Lorenzo-Lacruz (2013a) studied the spatial and temporal patterns of streamflow droughts in the IP finding a trend toward increased drought severity in the majority of regions. Additionally, Vicente-Serrano et al., (2014) found that hydrological drought frequency and severity have also increased in the past five decades in natural, regulated and highly regulated basins.

The situation of water stress to which it is subjected the IP, it is not only because the decrease in water availability but also de increase in water demand. Usually, water demands consist on urban supply, industrial demand and irrigation and also environmental requirements.

Each demanding sector presents different characteristic regarding quantitative and qualitative requirements. From the National Hydrological Plan of 2001 (PHN, 2001), some statistics can be extracted in relation to water demands in the IP:

- Urban supply includes the demand for domestic consumptions and other activities related to industrial services based on urban areas. In Spain, it is about 4.700 $hm^3$/yr, which means around 15% of the total demand. Its spatial distribution is closely related with the density of population. This is particularly notorious in case of the southern Mediterranean countries where tourist population is very significant. Tourist water consumption in the southern Mediterranean countries is about three times higher than local demands (EEA, 2000). Although it is highly seasonal, the related activity increases permanent water demand for facilities and leisure structures. Particularly, Spain is experiencing a tourist and second-home boom all across the Mediterranean coast and together with the 275 golf courses (75 more are on project only along the Mediterranean coast), represent a demand increase of about 30 million $m^3$. In conclusion, some of the areas with higher natural hydrological constraints present a remarkable process of urban concentration and a strong development of the tourism sector. Despite its proportion in relation with the total of water demand is small, the risk of shortages is notable.
- The water requirements for industries represent around 1700 $hm^3$ in Spain (which is about 5% of total demands). These consumptions do not include those related to urban distribution networks or for energy production (cooling and hydropower), whose net consumptions are scarce.
- Irrigation is, from a quantitative point of view, the main water use in Spain, with a water demand above 24000 $hm^3$ per year (~ 80% of the total consumptions). The million hectares for the traditional irrigable area was tripled, especially in the mildest parts of the country, and the water resources demand increased in all regions especially in the driest ones. During the last years, there has also been an important reduction in the trend of irrigated land in Valencia and Barcelona but an increase has been observed in certain areas (Albacete, Almería, Murcia, among others) (Programa AGUA, 2004). Among the new irrigated areas in some central parts of Spain, common dry-crops (cereals) irrigation through pumping wells is currently taking place.

In order to mitigate the effects of the unreliable water available resulting from the high spatial and seasonal hydrological variability across the Iberian Peninsula (IP), it has been developed a complex network of dams and channels, particularly during the second half of the

twentieth century (Berga-Casafont, 2003). These storage structures (e.g., reservoirs, canals) are developed to store water to ride out the dry spells and to provide socio-economic benefits through power generation, ensure irrigation (which is the major aim for most of reservoirs) and urban water supplies during periods of water scarcity or reducing the impact of floods and droughts. While reservoirs have limited capacity, the demands are diverse and often in competition (e.g., the conflict between the amount of water stored and the timing of releases to meet agriculture, hydropower generation and environmental needs). Additionally, dam operations have an impact on the natural functioning of rivers and related subsystems. On this regard, Morán-Tejeda et al. (2012) investigated the effect of reservoirs on river regimes in the Duero Basin. They found that the degree of regulation was highly dependent on annual inflows into the reservoir, and consequently alterations to river regimes were more intense during dry years. Also, they observed an absence of a common approach to reservoir management, and the dominance of other interests over environmental concerns, particularly in the context of hydrological change in the basin. Then, skilful long-lead streamflow forecasting (e.g., seasonal/monthly volume, timing of peak flow) are key to efficient and sustainable water resources development, providing information for appropriate reservoir management.

The increasing demands on limited water resources during the last decades are contemplated to continue in future (WWAP, 2014). In fact, the last Intergovernmental Panel on Climate Change Assessment Report (IPCC, 2013: Summary for Policymakers), it is emphasized that warming of the climate system is unequivocal, and since the 1950s, many of the observed changes are unprecedented over decades to millennia. For example, 1983–2012 was likely the warmest 30-year period of the last 1400 years in the Northern Hemisphere. Also, the upper ocean has warmed, the amounts of snow and ice have diminished, sea level has risen, and the concentrations of greenhouse gases have increased. Furthermore, under continued emissions of greenhouse gases, global surface temperature change for the end of the 21st century is likely to exceed 1.5°C relative to 1850 to 1900 and it will continue beyond 2100 for most Representative Concentration Pathways (RPC) scenarios. Warming will continue to exhibit interannual-to-decadal variability and will not be regionally uniform. Concerning the effects on water cycle, it is virtually certain that, in the long term (Collins et al., 2013), global precipitation will increase with increased global mean surface temperature. These changes will exhibit substantial spatial variation. General pattern of change indicates that high-latitude land masses are likely to experience greater amounts of precipitation due to the increased specific humidity of the warmer troposphere as well as increased transport of water vapour from the tropics by the end of this

century under the RCP8.5 scenario (the pathway with the highest greenhouse gas emissions). Many mid-latitude and subtropical arid and semi-arid regions will likely experience less precipitation and many moist mid-latitude regions will likely experience more precipitation by the end of this century under the RCP8.5 scenario. Decreases in annual runoff are likely in parts of southern Europe, the Middle East, and southern Africa by the end of the 21st century under the RCP8.5 scenario. An increase of agricultural drought are likely in presently dry regions and are projected with medium confidence by the end of the 21st century under the RCP8.5 scenario. Soil moisture drying in the Mediterranean, southwest USA and southern African regions is consistent with projected changes in the Hadley Circulation and increased surface temperatures, so surface drying in these regions as global temperatures increase is likely with high confidence by the end of this century under the RCP8.5 scenario.

Concerning the impacts of future climate change on water resources in the Mediterranean Basin, an ensemble of regional climate models driven by several GCMs using the A1B scenario have a robust decrease in runoff emerging only after 2050 (Sanchez-Gomez et al., 2009). Feyen and Dankers (2009) analysed the impact of global warming on streamflow drought in Europe, finding that frost-free season streamflow droughts will become more severe and persistent in most parts of Europe by the end of this century, except in the most northern and northeastern regions. Estrela et al., (2012) studied the impacts of climate change on water resources in Spain, concluding that runoff is expected to be reduced between 10 and 30% for the whole country through the 21st century. Also, Arguëso et al. (2012) indicated that a substantial decrease of precipitation in Spain is likely expected by the end of the 21$^{st}$ century. Forzieri et al. (2014) addressed the issue of future developments in streamflow drought characteristics across Europe under a multimodel ensemble projections approach. This analysis shows that streamflow droughts will become more severe and persistent in many parts of Europe due to climate change, except for the northern and northeastern parts of Europe. In particular, southern regions will face strong reductions in low flows. Koriala et al. (2014) evaluated the changes in streamflow at the end of 21st century by using runoff outputs from 11 atmosphere–ocean general circulation models (AOGCMs) participating in the fifth phase of Coupled Model Intercomparison Project (CMIP5). They concluded that mean and low flows are both projected to decrease in Europe, particularly in the south, under both RCP8.5 and RCP4.5 scenarios.

Taking into consideration the observed and postulated changes in water resources, particularly in the Mediterranean region, maximizing water management efficiency based on streamflow forecasting plays a key role in water resources management and planning. While

short-term streamflow forecasting such as hourly or daily forecasting is crucial for flood warning and defence, long-term forecasting based on monthly, seasonal or annual time scales is very useful in reservoir operations and irrigation management decisions, protection of the environment or in the reduction of expenses in flood and drought mitigation. Furthermore, every year, management decisions for operating river diversions and dams are made early in the year in anticipation of the forthcoming seasonal streamflow. Because of this, in recent years, interest in long-range predictability of river discharge variability has increased markedly in most of the world regions.

### 1.2 Overview of current streamflow prediction framework

The hydrological system acts as spatial and temporal integrator of precipitation (rain and snow), temperature, and related evapotranspiration over a specific region. Hence, variations in these fields are amplified in streamflow, and in general, it is easier to detect a change in discharge that directly in the basic climatic variables (Dettinger and Diaz, 2000; Trigo et al., 2004).

The skill of the long-range forecasts largely rests on the influence of slow variations in the Earth's surface conditions of soil moisture, snow cover, sea-ice and ocean surface temperature (Shukla and Kinter, 2005), and somehow also the stratosphere (Marshall and Scaife, 2009; 2010). In addition, on seasonal timescales, anomalous atmospheric conditions are often linked with seasonal variations in the river streamflows and reservoir storages, via variations in precipitation and temperature (Dettinger and Diaz, 2000; Cullen et al., 2002; Trigo et al., 2004).

Most of the dominant sources of climate variability are naturally recurring and persistent phenomena referred to as "teleconnection patterns" (Barnston and Livezey, 1987). These patterns, associated with some large-scale oscillations in atmospheric dynamics, may influence temperature, rainfall, storm tracks, and other meteorological phenomena over vast areas. A comprehensive review can be found in Hurrell et al. (2003) and some updates have been carried out in from Quadrelli and Wallace (2004) and Trenberth et al. (2005).

Two of the most important phenomena that influence streamflow variability are El Niño/ Southern Oscillation [ENSO; Philander (1990); Halpert and Ropelewski (1992); Neelin et al. (1998)] and the North Atlantic Oscillation [NAO; Barnston and Livezey (1987); Hurrell (1995); Visbeck et al. (2002)]. The ENSO pattern encompasses two linked phenomena, a quasi periodic warming behaviour in the tropical Pacific near South America around and an adhering effect which is the zonal anomaly of SLP in the tropical Pacific (the Southern Oscillation), that is the

atmospheric component of El Niño (Bjerknes, 1969). The North Atlantic Oscillation (NAO) is the most prominent and recurring extratropical teleconnection pattern in the Northern Hemisphere (NH). Generally, the NAO refers to the meridional seesaw, or the dipole structure, typically associated with the north center near Iceland and the south center near the Azores of the SLP field. These main teleconnection patterns have been commonly used as potential predictors of hydrological variables. For example, several studies have shown significant relationships between ENSO events and streamflows at global (Dettinger et al., 2000; Chiew and McMahon, 2002), and, particularly, for the European area (van Oldenborgh et al., 2000; Mariotti et al., 2002; Park, 2004; Zanchettin et al., 2008; Shaman and Tziperman, 2010; García-Serrano et al., 2011) and the IP has been also addressed (Rodó et al., 1997; Pozo-Vázquez et al., 2005; Vicente-Serrano 2005; Lorenzo et al., 2010; Vicente-Serrano et al., 2011). Significant lag-correlations were identified between NAO index and several river streamflow anomalies from the IP (Trigo et al., 2004) and Tigris-Euphrates streamflow anomalies (Cullen et al., 2002). Also, the effects of NAO in Iberian rivers have been addressed (Lopez-Moreno et al., 2007; Lorenzo-Lacruz et al., 2011; Moran-Tejeda et al., 2011). Rimbu et al. (2004) found significant lag-correlation between NAO and ENSO indices and Danube streamflow. However, the association between NAO and ENSO and streamflow for the IP (Trigo et al., 2004) and for south-east Europe (Cullen et al., 2002; Rimbu et al., 2004) is non-stationary, i.e. the strength of the correlation between these two phenomena and streamflow anomalies has changed over time. On this regard, López-Moreno and Vicente-Serrano (2008) evaluates the non-stationary influence of the North Atlantic Oscillation on European precipitation, finding a general trend toward a strengthening of the NAO-precipitation relationship over most of Europe has been detected for the twentieth century. Because of this, the identification of stable relationships between atmospheric and oceanic modes of variability and the hydrological variables to be predicted becomes a major step in forecasting experiments.

Additionally, atmospheric and oceanic variability in regions different from the used to define the main teleconnection patterns could have a non-negligible influence on the climate in a particular area. Therefore, researchers have emphasised in the analysis of the relationship between seasonal or interannual streamflow variability in many large river basins with changes in the large-scale oceanic and atmospheric phenomena to make potential previsions. Some examples of studies using climatic information, from both teleconnection indices and other climatic fields, from previous seasons or months to make long-range streamflow forecasting in many regions of the world are: in North America (Hanson et al., 2004; Tootle and Piechota, 2004; Grantz et al.,

2005; Opitz-Stapleton et al., 2007; Tootle et al., 2007; Soukup et al., 2009; Timilsena et al., 2009; Kalra et al., 2009, 2012; Bracken et al., 2010; Lamb et al., 2010; Oubeidillah et al., 2011; Tang et al., 2011; Anderson et al., 2012), South America (Gutierrez and Dracup, 2001; Tootle et al., 2008; Córdoba-Machado et al., 2014), Europe (Rimbu et al., 2005; Ionita et al., 2008, 2011; Bierknes and van Beek, 2009; Gámiz-Fortis et al., 2010, 2011; Oubeidillah et al., 2012; Hernández-Martínez et al., 2014) , Asia (Shrestha and Kostaschuk,2005; Chandimala and Zubair, 2007; Maity and Kumar, 2008), Africa (Eldaw et al., 2003; Sittichok et al., 2014) or Australia (Piechota et al., 2001; Chiew et al., 2003; Ruiz et al., 2007).

Due to the importance of hydrologic forecasting, a considerable number of forecasting models and methodologies have been developed and applied in streamflow forecasting. The streamflow forecasting models may fall into two general classes, *process-driven methods* and *data-driven methods* (Wang, 2006). *Process-driven models* treat a streamflow process as the output of a watershed system. Some examples are the rainfall-runoff models such as lumped, semi-distributed and distributed, and snowmelt-runoff and low flow models. *Data-driven models*, on the other hand, identify the relationship between the inputs and outputs on a mathematical ground, without taking into account the physical mechanism. This Thesis focuses on the data-driven models to identify the relationships between the atmospheric-oceanic variability and the seasonal streamflow variability and develop forecasting models able to make accurate seasonal streamflow predictions.

Among the statistical forecasting models, regression models are probably the most common in climate applications (Goddard et al., 2001; Zwiers and von Storch, 2004). Also, other models commonly used are Canonical Correlation Analysis (CCA) or Singular Value Decomposition (SVD), which expands simple regression, one variable upon another, to multidimensional vectors. These models are more useful because they can be quite simple and applicable, so they have become more popular in streamflow forecasting, due to the increase in data availability from stations, real-time data retrieval, and increasing computational capability with the development of more robust methods and computer techniques (Wang, 2006). Instead of using linear modeling, some recent studies advocate non-linear modeling since it offers more flexibility and possibly better performance; more importantly, non-linearity may be closer to the nature of climate mechanisms. Among these non-linear models, some of the most used are the artificial neural network (ANN) (Hsieh and Tang, 1998; Tangang et al., 1998; Tang et al., 2000), the non-linear generalizations of CCA (Hsieh, 2001) or the discriminant analysis (Mason, 1998) and specific techniques from machine learning (Lima et al., 2009; Lima and Lall, 2010). However, the

utilization of non-linear models is subjected to the availability of streamflow time series long enough.

Regarding the IP, some authors have recently centered their efforts on produce advances on the study of the seasonal or interannual streamflow forecasting. In particular, Gámiz-Fortis et al. (2008a) examined the interannual variability and predictability of the winter streamflow of the main IP international rivers (Douro, Tagus, and Guadiana) for the period 1923–2004. These authors used a singular spectral analysis to isolate the main oscillatory components of the streamflow series. Then, they fitted autoregressive-moving-average (ARMA) models to the Singular Spectrum Analysis (SSA) filtered streamflow in order to conduct an intereannual forecast experiment. Additionally, in a companion paper, Gámiz-Fortis et al. (2008b) analysed the role of the Atlantic summer and autumn sea surface temperatures (SSTs) on the predictability of these winter IP River flows. A similar approach, combining the predictability skill of SST in seasonal and interannual streamflow variability with the knowledge of main oscillatory modes of variability presented in streamflow time series, was used to investigate the predictability of the Douro (Gámiz-Fortis et al., 2010), the Ebro River flow anomalies (Gámiz-Fortis et al., 2011), and streamflow in the Internal Catalonian Basins (Hernández-Martínez et al., 2014).

### 1.3 Proposed research

The main goal of the research presented here is to contribute to the evaluation of medium and long-range scales seasonal streamflow forecasting of the IP Rivers. The key contribution of the work focuses on providing a global perspective to the entire area, using the greater density coverage of stations as possible and exploring the atmospheric-oceanic forcings driving seasonal streamflow. Specifically, the proposed research framework has four major components.

1) To develop a streamflow database covering the entire IP (or as many areas as possible), able to be used in climate studies. To do that, a balance between density of available gauging stations records and their length is needed. A careful check for inhomogeneities in streamflow time series has to be done to ensure that gauging stations belonging this database do not present important alterations in the natural regime. Additionally, the main spatial and temporal characteristic of the IP Rivers are studied.
2) To understand and identify the main global land-ocean-atmospheric drivers of seasonal variability of IP Rivers. This involves analyzing the multivariate space-time data of large-scale climate variables to find any statistically significant relationship between

different climatic variables (e.g. sea surface temperature, precipitation) and the seasonal variability of discharge anomalies. The large-scale climate variables explored are not only restricted to the main climate patterns (known as 'teleconnection patterns') but also the identification of the main modes of covariability between climate variables and seasonal streamflow and investigate the potential use of these new regions in long lead-time hydrologic forecasting. In addition, the problem of non-stationarity of climate variability and streamflow variability relationship is addressed by identifying stable predictors.

3) To develop statistical forecasting models to make long-range predictions with different updatable forecasting scenarios. So, as the lag between predictor and predictand is reduced, the new climatic information is incorporated to the statistical models and a new set of predictions are made. This work focuses on the developing of linear models to establish the relationships between predictors and predictands. Forecasting based on the identification of non-linear signals between the climate system and the river discharge variability requires long records, which are not available in most of the Iberian River Basins. Since, to provide a forecasting perspective covering the IP is one of the main goal, a linear modelling approach is preferred until longer time series are available for most of the streamflow records.
4) To evaluate the skill of streamflow forecasts using a set of different verification measures in order to identify those basins and seasons where more accurate predictions can be obtained. The quality of the forecast is not only about how well is the fitting of the predicted time series but also with how many months or seasons is the prediction given in advance.

### 1.4 Outline of this thesis

The work is divided into eight chapters, including the introductory and conclusion chapters. An outline of the chapters that follow is provided here.

- Chapter 2 gives the details of the study area, its geographical and climatic features and a brief description of the different basins.
- Chapter 3 provides a description of the original streamflow database, together with the procedure followed to discard gauging stations with non-homogeneous behaviour.
- Chapter 4 presents an overview of the spatial and temporal variability of the seasonal

streamflow in the IP. Multivariate methods are used to determine the main spatial patterns, whereas a non-parametric approach is employed to detect temporal trends.

- Chapter 5 attempts to find stable predictors among most common teleconnection indices representing main large-scale atmospheric and oceanic patterns.
- Chapter 6 describes the connections between seasonal streamflow and Atlantic and Pacific sea surface temperature, geopotential height at 500 hPa and global temperature and precipitation. These connections are explored using Singular Value Decomposition analysis.
- Chapter 7 presents a forecasting experiment of seasonal streamflow using as predictors the teleconnection indices and significant coupled modes identified as stable predictors in Chapters 5 and 6, respectively. Predictions are made with various lead times, using linear regression models.
- Chapter 8 summarizes the main conclusions and future work.

# CHAPTER 2

# STUDY AREA

A general overview of the geographic, climatic and hydrological features of the Iberian Peninsula is carried out in this chapter. A special emphasis is made in relation with the effect that the development of wide dams network during the second half of twentieth century have had on the Iberian Peninsula hydrological system.

## 2.1 Geographical and climatic features of the Iberian Peninsula

The Iberian Peninsula is located between 36ºN - 44ºN and 3ºE - 10ºW. The region is surrounded by the Atlantic Ocean to the west, the European continent to the northeast, the Mediterranean Sea to the east and the Sahara desert to the south. Its complex orography, with crucial geographical features such as the Strait of Gibraltar or localized mountainous systems, become an important key for the climate of the region. Figure 2.1 displays the location and main topographic features of IP.

The climate of the IP is characterized by high intra-annual variability. The origin of the air masses makes its temperature and water content vary dramatically, which represent an important component of local variability for both precipitation and temperature. According to Köppen climate classification, the climatic regimes in the IP are: temperate climate, such as Csa (temperate with dry or hot summer, over the majority of the southern-central plateau region, and

the Mediterranean coastal regions, with the exception of the arid zones in the southeast), Cfa (temperate with a dry season and hot summer, mainly in the northeast, within an area of medium altitude which surrounds the Pyrenees and the Iberian mountains) or Cfb (temperate with a dry season and temperate summer, in the Cantabrian mountains, in the Iberian mountain ranges, as well as part of the northern central plateau region and a large part of the Pyrenees, with the exception of areas of high altitude), dry climates, such as Bsk (in the southeast of the IP) and also cold climates in the mountains. A more detailed definition of the IP climates can be found in AEMET (2011). The main factors affecting the climate of the IP are its location, in mid-latitudes of the NH, the influence of two important bodies of water (the Atlantic Ocean and the Mediterranean Sea) and its rugged terrain, with the mountain ranges acting as barriers and passageways that determine local climates. These factors generate a northwest-southeast gradient in annual precipitation distribution (De Castro et al., 2005; González-Hidalgo et al., 2011), varying from more than 2000 mm/year in north-western to less than 300 mm/year in the south-eastern. The large-scale precipitation is mainly modulated by the position of the Azores anticyclone that acts as a blocking structure in the summer, preventing the low-pressure systems from reaching the IP, and bringing hot and dry weather. The scarce summer precipitation is mostly due to local factors and convective storms (Serrano et al., 1999). On the contrary, during winter months, which account about 40% of annual precipitation, the Azores high moves toward the tropics accompanied by a southward displacement of the jet-stream letting the low-pressure systems get into the IP, resulting in wetter conditions. Note that in the eastern façade, mesoscale convective systems are also responsible for high rainfall rates, especially in transient seasons (Paredes et al., 2006; García-Herrera et al., 2005). Additionally, spatial variability is also remarkable, with a strong contrast between the northern (wet) and southern (arid) areas. Regarding its variability in temperature, high temperature values are reached in river valleys during summer (mainly in southern basins), whereas very low records are registered in the elevated mountainous areas in winter. Also, temperature range varies from inland areas (higher) to coastal areas (lower).

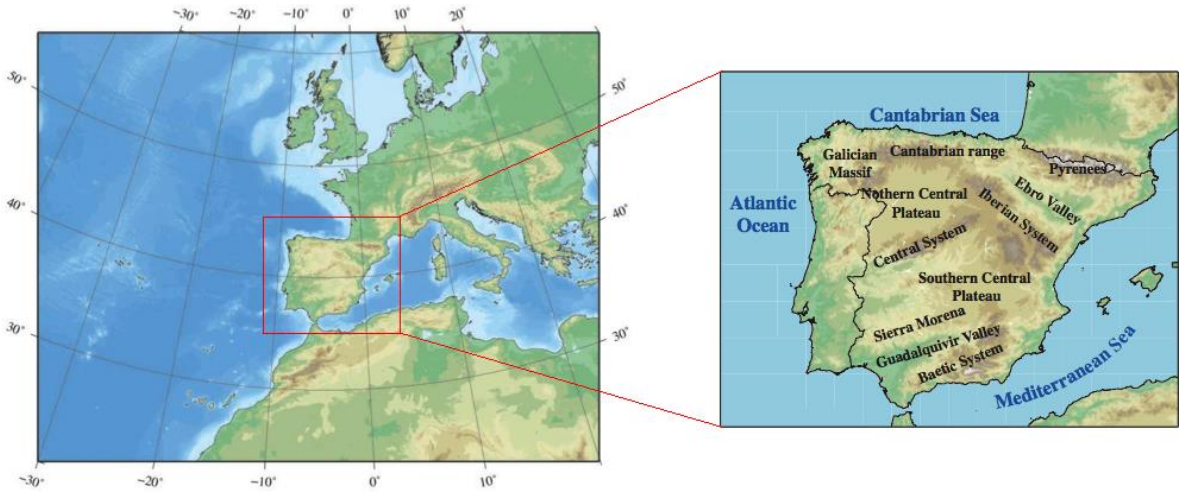


Figure 2.1. Location of the IP (left) and main geographical features (right).

## 2.2 Hydrology of the Iberian Peninsula

The river system in the IP has many peculiarities and strong contrasts, due to a rich and diverse environment and landscape, being the irregularity and the asymmetry of the slopes the main features defining the IP rivers. The major IP rivers run in the same direction of the parallels, flowing towards the Atlantic Ocean (the Miño, Douro, Tagus, Guadiana and Guadalquivir Basins) and the Mediterranean Sea (the Segura, Júcar and Ebro Basins), except for the Cantabrian and the southern peninsular rivers, which born in mountains near the sea and follow the direction of the meridians. In natural regime, the Spanish river system accounts for 111000 $hm^3$/yr, which represents about a third of mean annual precipitation and almost total cumulative flow form rivers (Water in Spain, 2003), although this value is unevenly distributed along time and space. River basins in northern sector of Atlantic watershed present higher mean annual flow (of 10570 $hm^3$/yr for the Miño River, 13788 $hm^3$/yr for the Douro River, and 12350 $hm^3$/yr for the Tagus River) than rivers in the southern sector of Atlantic watershed (4039 $hm^3$/yr for the Guadiana River and 3780 $hm^3$/yr for the Guadalquivir River). On the contrary, most of streamflows belonging to basins in the Mediterranean watershed (the Mediterranean Andalusian, Segura, Júcar, and Inland Catalonian Basins) are generally lower. The exception is the Ebro Basin, which has an abundant flow (mean 12279 $hm^3$/yr) that is generated in the Cantabrian Range and in the Pyrenees.

Water resources in the IP, which play a crucial role in the economy of the area, are under recurrent water stress. In addition, the climate models project a decrease in precipitation and an increase in evapotranspiration in the IP during this century (IPCC 2013, summary for

policymakers). Together with these expectations, water demands still grow insatiably, especially where resources are scarce (the tourist sector in the Mediterranean areas or a competitive export-oriented agricultural sector). The unbalanced distribution of water resources within the IP together with the rising demand for water have culminated in conflict among users and regions (Quiroga et al., 2011), especially in the Mediterranean area, and in the planning of a variety of inter-basin water transfer formulas (such as the water law reform in 1999, enabling water market transaction; the EU Water Framework Directive in 2000; the 2001 National Hydrological Plan and its subsequent reform in 2004, with the controversial repealing of Ebro water transfer (Arroyo-Ilera, 2007).

Also, in addition to the situation of recurrent water stress, the increasing frequency of drought events in the IP since the 1970s (Vicente-Serrano et al., 2006a; Iglesias et al., 2007) has resulted in the construction of a complex network of dams and channels to optimize the use of available water resources. In the early twentieth century, there is reference of about 100 dams (with a capacity of 100 $hm^3$), being 60 of them are currently in operation (Berga-Casafont, 2003). During the first half of the twentieth century 218 large dams were built, increasing the reservoir capacity in 6034 $hm^3$. However, it was in the second half of the century when the construction of dams experienced a substantial increase, especially between 1955 and 1975, multiplying by six the reservoir capacity (mainly due to hydroelectric purposes). Currently there are about 1.300 reservoirs in Spain, with a capacity of about 60.000 $hm^3$ (which is approximately equal to the mean annual streamflow of the 8 major rivers of the IP), and a regulated volume of about 46000 $hm^3$/yr (41% of natural resources). The 98.2% of the reservoir capacity is concentrated in 347 reservoirs with a capacity greater than 10 $hm^3$. Figure 2.2 shows the location and capacity of major reservoirs in Spain. Table 2.1, from Berga-Casafornt (2003), summarizes the evolution of the major dams in the second half of the twentieth century.

Advantages of reservoirs construction are multiple, such as providing socio-economic benefits through power generation, ensuring irrigation (which is the major aim for most of reservoirs) and urban water supplies during periods of water scarcity or reducing the impact of floods and droughts. However, there are some drawbacks related to these infrastructures. In particular, they involve substantial modifications to the river system from a hydrological, ecological or geomorphological perspective. Hence, the comparison of natural and altered regimes in the IP is very useful to evaluate the effects of reservoirs on the entire river system (such as in Batalla et al., 2004, López-Moreno, 2009a).

| INCREASE IN MAJOR DAMS IN 1950-2000 | | | | |
|---|---|---|---|---|
| **Period** | **Nº major dams** | **Annual mean increase** | **Storage capacity (hm3)** | **Annual mean storage capacity (hm3)** |
| **< 1950** | 276 | 4 | 6142 | 120 |
| **50-60** | 464 | 19 | 18167 | 1200 |
| **60-70** | 666 | 20 | 36919 | 1875 |
| **70-80** | 859 | 19 | 41717 | 480 |
| **80-90** | 1016 | 16 | 49313 | 760 |
| **90-00** | 1195 | 18 | 56500 | 720 |

Table 2.1. Evolution in the number of major dams during the second half of twentieth century (from Berga-Casafornt, 2003)

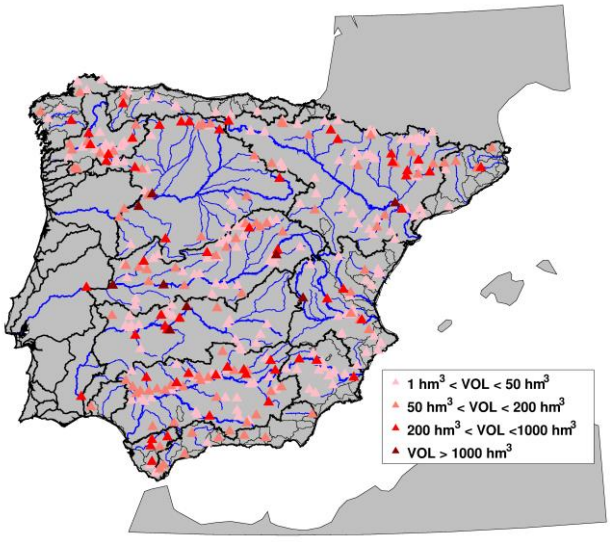


Figure 2.2. Location and capacity of major reservoirs (greater than 1 $hm^3$) in Spain.

In the Atlantic sector, the trans-boundary character of the Duero River (which flows from Spain into Portugal) creates difficulties for the management of water and reservoirs in the basin, and often causes conflicts between users and policymakers in both countries. The case of Tagus

River, probably the most affected by the construction of large reservoirs, is also controversial. There are a total of 40 reservoirs with more than 15 $hm^3$ of capacity in the Tagus Basin, which affect most of the main tributaries (Flores, 2004). In the headwaters, the Entrepeñas (802 $hm^3$) and Buendía (1638 $hm^3$) reservoirs provide water for the Tagus-Segura transfer (Gómez-Mendoza and Mata, 1999; Morales *et al.*, 2005). This inter-basin water transfer between the Tagus and Segura Rivers in the upper Tagus is undertaken to solve the problem of the near-chronic water deficit experienced in coastal areas of the Spanish Mediterranean. Since the construction of the inter-basin channel in 1979, the annual discharge of the Tagus has been generally below the historical average, and water transfer has been only able to supply about 40% of the planned volume of 600 $hm^3$/yr (Morales *et al.*, 2005). This water transfer has aggravated the observed decreasing in Tagus streamflow (Gallart and Llorens, 2002), which has changed from 7515 $hm^3$/yr during the pre-dam period (1943-1969) to 6208 $hm^3$/yr for the post-dam period (1970-2003). The determination of the amount of available water for transferring during dry years has caused severe social and political conflicts in Spain. The trans-boundary nature of the Tagus Basin also leads to difficulties in flow management, mainly during large floods and drought periods (Azevedo *et al.*, 2004). Another important reservoir, in terms of its potential to modify the river regime downstream, is the Alcántara reservoir (with a gross storage capacity of 3162 $hm^3$). It is located close to the border between Spain and Portugal, and it was built in 1969 for hydropower generation. In the Guadiana catchment (whose headwater is one of the driest areas of Europe), the increase of water storage (from almost zero in 1954 to 4000 $hm^3$ in 1963 (Brandão and Rodrigues, 2000) and to 12000 $hm^3$ in 1991) to satisfy the large water consumption from agricultural projects and public water supply, has significantly modified the natural regime of the river.

As regards of the river basins in the Mediterranean sector, in the Segura Basin is one of the most highly regulated, especially since 1957, because of the major reservoirs constructed in the headwaters. It is worth to mention the negative impact on the natural flow that the inter-basin transfer Tagus-Segura has had. Also, major rivers in the Jucar Basin are strongly regulated, while short watercourses are less affected, although the quantitative assessment of this type of pressure is vague in absolute terms as well as associated impact (Gil-Olcina, 2006). Concerning the situations of water transfer and regulation of the Ebro River, an annual amount of 150 $hm^3$ is transferred for Bilbao metropolitan area and an extra of 55 $hm^3$/yr for Tarragona's lands, in Catalonia. In the last years, a dispute regarding the projected, and thereafter revoked, water transfer from the Ebro to negative balanced Mediterranean Basins has involved several politic

and social disputes. How water demands from these basins are going to be cover is still under study.

## 2.3 Description of the IP main basins

The Spanish hydrographical map is integrated by different river basins and districts, as a result of the application of Royal Decree 125/2007 and subsequent amendments and other constraints established from the regional arrangements. Figure 2.3 shows the spatial delimitation of the river basins considered for this study. The different river basins considered, together with a brief description (main geographical features, climatology and water resources and demands) are summarized in Table 2.2 (this information has been compiled from the website of each river basin authority and the Spanish National Hydrological Plan).


Figure 2.3. Identification of the river basins used in this work. Line separating Douro, Tagus and Guadiana basins represent the border between Spanish and Portuguese parts.

In addition to the Spanish River basins described in Table 2.2, gauging stations belonging the main basins located in Portugal were also considered. Particularly the gauging stations located in the basins shared between Portugal and Spain (Douro, Tagus and Guadiana), but also stations belonging the main Portuguese basins, such as the Mondego (between the Portuguese part of Douro and Tagus Basins) and Sado (between the Portuguese part of Tagus and Guadiana basins). The Mondego River is the greatest Portuguese river (237 km), with a drainage area of 6670 $km^2$ and a discharge of 411 $hm^3$/yr in average. The Sado River basin has a catchment area of 7640 $km^2$, and in general, its streamflow discharge is smaller (253 $hm^3$/yr).

| BASIN | Hydrological features | Climatology | Regulation and Water demands |
|---|---|---|---|
| **Miño-Sil** | CA = 17757 $km^2$<br>MC = 8393 $hm^3$/yr | Temperate and humid oceanic climate, with soft winters and cold summers.<br>AMP = 1175 mm/yr<br>AMT = 11.3ºC | Total storage capacity is above 3000 $hm^3$ (mainly for hydroelectrical purposes).<br>Total demand 436.4 $hm^3$/yr, especially for agriculture (70%) and urban supplies (26%). |
| **Douro** | CA = 97290 $km^2$ (Spanish part, 81%)<br>MC = 13500 $hm^3$/yr<br>Length of Douro river: 897 km (572 km Spanish part) | Mediterranean climate (highly continental), with cold winter (freeze) and mild summers.<br>AMP = 618 mm/yr irregularly distributed, mainly in autumn and spring, almost non-existent in summer.<br>AMT ranges between 10-12 ºC, with peaks of almost -25ºC and 45ºC. Great variability in the diurnal temperature range, from 3.5ºC (cold months) to 20.5ºC (warm months). | Storage dams capacity of 8000 $hm^3$.<br>The total water demand is about 4884 $hm^3$/yr, being 92% for agriculture (particularly irrigation) and the rest for human demand and industries. |
| **Tagus** | CA = 88700 $km^2$ (55800 $km^2$ in Spain),<br>MC = 12350 $hm^3$/yr (in Spain).<br>The length of Tagus River is 1100 km (827 km in Spain). | Climate varies from Mediterranean, in the east, to Atlantic conditions, in the west. Marked seasonal variability (dry and warm summers, and cold winters).<br>AMP is 650 mm, ranging between 1100 mm/yr in mountainous areas and 450 mm/yr in arid zones.<br>AMT is about 11ºC, with clear | The total capacity of the reservoirs is 14500 hm3 (12500 hm3 belong to Spain).<br>Supplies 11 millions of people (6 million in Spain), and 120000 (110000) ha of public (private) irrigation land. Industries demand 250 $hm^3$/yr |

| | | | |
|---|---|---|---|
| | | spatial variations (5-7ºC in the mountain areas, and 14–15ºC in the central part) | |
| **Guadiana** | CA = 60256 $km^2$ (48656 $km^2$ Spain).<br>MC = 5860 $hm^3$/yr<br>Length Guadiana River is 744 km. | Mediterranean-continental climate.<br>AMP: 550 mm (mainly in period October-April)<br>AMT: 16-17º C (with spatial variations between, 14ºC in header, and 18-19ºC in the river mouth. | Total storage capacity of 8189.4 $hm^3$.<br>Water demand reaches 2238.57 $hm^3$/yr. Irrigation demands are notable (90% of total water demand). Problems derived of overexploitation of water resources. |
| **Guadalquivir** | CA = 57527 $km^2$<br>MC = 7230 $hm^3$/yr.<br>Length of Guadalquivir River is 722 km. | Mediterranean climate.<br>AMP is 550 mm, being irregular, torrential in some cases, and with recurrent summer drought.<br>Mean annual temperature is about 16.8ºC, with high records during summer. | Highly regulated, with a total storage capacity of 8782 $hm^3$.<br>High evapotranspiration levels. Water demands, for agriculture and animal husbandry (87%) and for human supplies (11 %) are above 3600 $hm^3$, generating a 'structural deficit'. |
| **Cantabrian** | CA = Western sub-basin 19000 $km^2$, Eastern sub-basin 5788 $km^2$<br>MC = 4659 $hm^3$/yr (eastern subbasin) and 11849 $hm^3$/yr (western subbasin) | Soft winters and cold summers, with a cold Pyrenean regime in mountainous areas.<br>AMP is 1250 mm (ranging from 800 to 1700 mm in the basin), with a homogeneous annual distribution, being scarcer in summer. | Major dams are located in western sub-basin (Salime, with 266 $hm^3$ and Doiras with 114.6 $hm^3$).<br>Water demands are 353 $hm^3$/yr in the eastern subbasin (68 % for urban and 30 % for industrial supplies) and 474.1 $hm^3$/yr in western subbasin (50 % for urban supply, 35 % for industry and 15 % for agriculture) |

| | | | |
|---|---|---|---|
| **Jucar** | CA = 42.851 $km^2$.<br>MC: 4142 $hm^3$/yr | Mediterranean climate, with warm summers and soft winters.<br>AMP is 500 mm, ranging spatially form 300 to 750 mm.<br>AMT between 14-16.6 ºC | Dam storage 3000 $hm^3$, allowing annual regulation of 1200 $hm^3$.<br>Water demands come from agriculture (80 %) and urban uses (17 %) |
| **Segura** | CA: 18870 $km^2$.<br>MC = 640 $hm^3$ /yr (surface waters) and 220 $hm^3$/yr of groundwater.<br>Tagus-Segura transfer: 545 $hm^3$/yr. | Mediterranean climate.<br>AMP about 500 mm, with spatial ranges from 300-750 mm.<br>AMT ranging between 14-16.5 ºC. Maximum annual temperature during summer (and dry) months. | Dam storage capacity 1.359 $hm^3$.<br>Negative water balance, (overexploitation of aquifers). 89 % of water demands comes from agriculture, 10 % for urban supplies and 1 % for industries.<br>Uncontrolled expansion of irrigation derived from inter-basin water transfer. |
| **Ebro** | CA = 85997 $km^2$<br>MC = range 7305-12279 $hm^3$/yr, for dry-rainy years. | Heterogeneous climatology, although predominately Mediterranean with different subtypes. AMP is about 600 mm, with spatial variability of 261-2.187 mm, and AMT is about 12.6ºC, with spatial variability of 5.6-15.8ºC. | Storage capacity 8360 $hm^3$.<br>Water demands reach 50000 $hm^3$/yr, for urban and industrial supply, hydroelectric and nuclear power generation, and, especially, for irrigation of 800000 ha (with an estimated water demand of 6310 $hm^3$/yr). |
| **Internal Catalonian** | CA = 16600 $km^2$.<br>MC = 2315 $hm^3$/yr | Mediterranean climate, with soft winters and dry and warm summers. Scarce precipitation, predominantly in mid seasons, with spatial variability from 400 mm/yr in central depression to | There are 10 reservoirs in total in the internal basins giving the ability to regulate 695 $hm^3$.<br>Provides 1186 $hm^3$/yr (38% of total consumptions in |

| | | | |
|---|---|---|---|
| | | 1200 mm/yr in the Pyrenees. High spatial temperature variability due to its contrasting topography. | Catalonia, mainly for urban and industrial uses) |
| **Mediterranean Andalusian** | CA = 17952 $km^2$<br>MC = 1187 $hm^3$/yr<br>2102 km of water masses (with a maximum river length of 70 km). | Because of its contrasting topography, it presents a highly variable climate (Mediterranean subtropical, sub-desert, semi-continental, mountainous) and precipitation regime (from 2.000 mm/yr in western slopes, to 200 mm/yr in eastern façade). | The storage dams capacity of 1368 $hm^3$.<br>In general, the water balance is negative. Water demands are 1410 $hm^3$/yr, being 75% for agriculture (irrigation) and 18% for urban supply. |

Table 2.2. Summary of the main features of the analysed basins, with CA = Catchment Area, MC = Mean Contribution, AMP = mean total annual precipitation and AMT = annual mean temperature.

# CHAPTER 3

# STREAMFLOW DATABASE

This chapter presents a description of the original database and the quality control processes followed to select, among all gauging stations available, those which overcome the homogeneity tests, the percentage of missing values allowed and the minimum length required.

## 3.1 The original database

Streamflow time series (on a monthly scale) of gauging stations in each river basin were established from the following water agencies:

- Centro de Estudios Hidrográficos (CEDEX). *http://hercules.cedex.es/anuarioaforos/default.asp*
- Agència Catalana de l'Aigua. *http://aca-web.gencat.cat/aca/appmanager/aca/aca/*
- Agencia Andaluza del Agua. *http://www.agenciamedioambienteyagua.es/*
- Sistema Nacional de Informaçao de Recursos Hídricos (SNIRH) de Portugal. *http://snirh.pt/*

Data files with the information of the river basins boundaries, rivers location and the capacity and location of the reservoirs were obtained from the website of the Spanish Ministry of Agriculture, Food and Environment, (*http://servicios2.marm.es/sia/visualizacion/descargas/mapas.jsp*)

The streamflow records from CEDEX and SNIRH are measured in $hm^3$ (contributions) at monthly scale. In hydrology, contributions refer to the volume of water supplied by the river at the measuring point for a month (in case of monthly data). By contrast, data of rivers in the Internal Catalonian Basins were expressed in $m^3/s$ (which refers to the average daily value), so they were translated into monthly contributions.

In order to increase the spatial coverage of gauging stations (scarce in some places, particularly in the south half of the IP), those stations located at the entrances of the main reservoirs were added (obtained from CEDEX website). It worth to be note how these entrances are calculated. Generally, two values are measured in reservoirs, the outputs and the water mass stored. The entrances are not directly measured, but they are calculated balancing the outputs and the water mass stored. These data should be treated cautiously, since they are not a direct measure, and effects as evaporation in reservoir are not taking into account. Despite these limitations, records in entrances of reservoirs were considered so as to complete de database in areas where quality and quantity of gauging stations were insufficient.

The compiled dataset comprises 808 gauge stations in Spain, 246 in Portugal, and 326 reservoir entrances in Spain, totalling 1380 data series. Figure 3.1 shows their spatial distribution. From this Figure, it could be seen how the incorporation of data from reservoir entrances complete the spatial coverage, particularly in Guadalquivir and Guadiana Basins, where the density of gauging stations available is inferior.

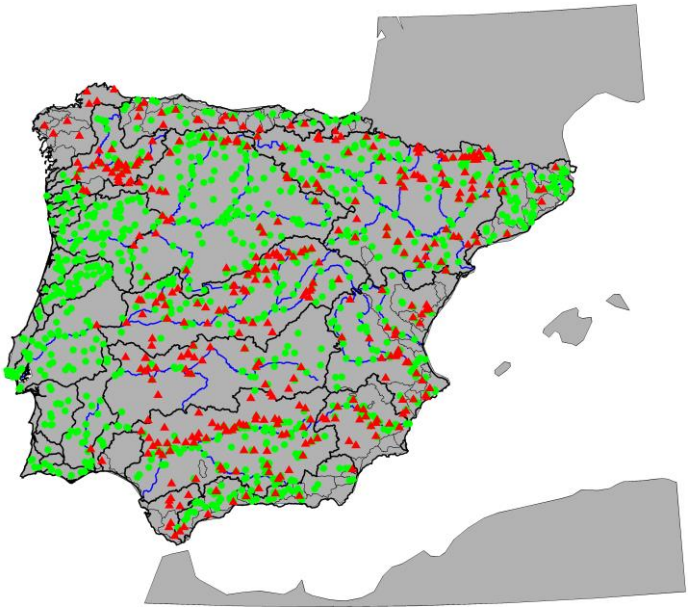

Figure 3.1. Spatial distribution of gauging stations (green circles) and reservoir entrances (red triangles) of the original database.

## 3.2 Data quality control. Methodology

### 3.2.1 Balance between temporal coverage and missing values

Establishing the most suitable time period for this study was the first question addressed. This period had to be the one that maximize the number of years (at least 30 years for climatic studies), the number of stations, its spatial density, minimize the number of missing values in time series and to be common for all stations and the as recent as possible.

### 3.2.2 Homogeneity test

The second stage of the quality control consisted in assessing the homogeneity of seasonal time series. Homogeneity testing is necessary in climate studies, since inaccurate conclusions can be derived from the analysis of time series that present unnatural behaviour. Some common causes of a lack of homogeneity in the climate data series are changes in the location of the stations, or environmental changes around the station. Particularly for the streamflow time series in the IP, the high degree of dam regulation in Spain may be the cause of alterations on the natural regime of streamflow, introducing a non-homogeneous behaviour in time series.

Detection of inhomogeneities or abrupt changes in hydrological series is a difficult issue to address, because the diversity of tests that can be found in the literature. The importance of specific information on the timing of changes in the station observing system or microclimate cannot be overemphasized. Most of this information, usually called *metadata*, becomes a very useful tool for determining inhomogeneities. However, when this information is not available, users can employ some of the several methods available for detecting inhomogeneities in climate records in the absence of information about the history of the observations. Although extremely useful, these methods should be taken with precaution, because without information of metadata the inhomogeneities detected could not be properly identified or corrected. An interesting review of the methodology more commonly employed in the study of changes in hydrologic series can be found in Kundzewicz sand Robson (2004). These authors recommended the use of various change-detection tests, and also presenting the results accompanied with the careful graphical inspection of the time series behaviour itself, as well as background knowledge (possible impacts from reservoir changes in land use or type of instrumentation) or metadata to help complete the study. Other authors such as Wijngaard et al. (2003) use the results of several tests, such as the Standard Normal Homogeneity Test, SNHT (Alexandersson, 1996), Buishand (Buishand, 1982), Pettitt (Pettitt, 1979) and von Neumann ratio test (Neumann, 1941), to classify the series based on the accordance between the different tests.

Some of the statistical tests developed to assess the homogeneity of time series assume normality in these time series. In this regard, it is known that meteorological time series, such as precipitation or streamflow, usually suffer from a lack of normality. Then, addressing the normality of the seasonal streamflow time series is a necessary previous step. Razali and Wah (2011) compared four statistical tests to evaluate the normality assumption of the data: Shapiro-Wilk (Shapiro and Wilk, 1965), Kolmogorov-Smirnov (Massey, 1951), Lilliefors (Lilliefors, 1967) and Anderson-Darlin (Stephens, 1974), concluding that Shapiro-Wilk was the most powerful. In addition, Steinskog et al. (2007), in a cautionary note on the use of the Kolmogorov-Smirnov test for normality, put on show that, a power comparison of eight tests for normality favoured Jarque-Bera (Jarque and Bera, 1987) and Shapiro-Wilk tests. In line with these findings, three tests were used to check the normality of the streamflow time series: the Lilliefors test, the Jarque-Bera and the Shapiro-Wilk. It was concluded that a time series follows a normal distribution if the three tests agree.

Bearing in mind the likely non-normal behaviour of hydrological time series, the Pettitt test was selected among the proposed test by Kundzewicz and Robson (2004) to analyse the

possible existence of abrupt changes in the seasonal streamflow time series. The main reason of using this test is its non-parametric nature, so it does not assume that the data fit any probability distribution, and that it is more sensible to detect break points in the middle of the series. The Pettitt test has been used in multiple studies to determine the existence of change points in hydrological series (Aka et al., 1996; Tu et al., 2004; Salarijazi et al., 2012, among others) and it is considered robust in terms of changes in the shape of the distribution of data.

The Pettitt test is essentially a sequential version of the non-parametric Mann-Whitney test, designed to search for points of abrupt changes in a time series. If we consider a sequence of variables $X_1$, $X_2$, ..., $X_T$, it is said to have a change point at time $\tau$ if $X_t$, for t = 1, ..., $\tau$ has a distribution function $F_1(x)$ and, $X_t$ for $t = \tau + 1, \ldots T$ has a distribution function $F_2(x)$ such as $F_1(x) \neq F_2(x)$. No further assumptions about the shape of the distributions are made except that they are continuous. It tests the null hypothesis of no change against the alternative of a change at time $\tau$ through the statistic $K_t$. Let $r_1$, ..., $r_T$ the ordered sequence of variables $X_1$, ... $X_T$, the statistic is defined as

$$K_t = 2\sum_{i=1}^{t} r_i - t(T+1) \tag{3.1}$$

*where t=1,…T,* being

$$K = \max_{1 \le t \le T} |K_t| \tag{3.2}$$

and the probability associated with the statistic $K$ is approximated by:

$$P_{0A} = 2e^{\left(\frac{-6K^2}{T^3+T^2}\right)}, \quad for\ T \to \infty \tag{3.3}$$

If $P_{0A}$ is less than 0.05 (95% confidence level), then the null hypothesis of 'no change' is rejected and the alternative hypothesis of an abrupt change in the series at the point $t$, where $K_t$ is maximized, is accepted.

When nonhomogeneous behaviour of seasonal time series is detected, the results from the Pettitt test indicate the year of the break point (at time $t$, where $K$ takes its maximum).

The Pettitt test was applied to the ‘extended’ seasonal streamflow time series. The term ‘extended’ means that the all records available where considered, not only the period from 10/1975 to 9/2008. The reason behind this is that the effects of river regulation could appear before 1975, and the time series for the period 1975-2008 could not reflect the natural behaviour of the river flow.

The main gaol of this quality control is identifying those stations that are strongly affected by dam regulation, so the original signal can be masked. For instance, an additional indicator was proposed (following the recommendations of Kundzewics and Robson (2004) of using multiple tests) to determine if the abrupt change found for Pettitt test could be associated with a change in natural regime as a consequence of dam regulation. The proposed indicator, called the 'Common Area Index' (CAI), is defined as the percentage of common area between the curves of the annual cycles calculated using data from before and after the change point. Hence, the CAI was calculated for each year where a seasonal time series presented an abrupt change (i.e., there were calculated as much CAI as seasons showing an abrupt change). Those stations that present any of these CAI inferior to 50 % were considered as inhomogenous, since the natural streamflow regime was strongly affected after the break.

### 3.2.3 Gap filling

A simple method based on linear regression has been developed to fill each gap. As predictor, it was chosen the station with the higher correlation with the station to be filled (predictand), with the restriction that this correlation had to be greater than 0.8. In those cases where there was not available any predictor well correlated, the gap was filled using the mean value of the correspondent season.

## 3.3 Results

The identification of missing values in time series was the first step (Figure 3.2). It could be seen from Figure 3.2 that the amount of stations available increases significantly after 1970s for gauging stations within Spanish territory (Figure 3.2a). This is the case of stations in Cantabrian and Tagus Basins, where a remarkable number of gauging stations are available from the seventies. In addition, stations with larger periods present some lacks during decades of 1920s and 1930s, especially during years of Spanish Civil War. In case of reservoirs entrances, as it is shown in Figure 3.2b, the first records started in mid 1940s and early 1950s, with the notable development in these infrastructures during 1970s. Regarding the degree of missing values in Portuguese gauging stations (Figure 3.2c), the number of stations available with a remarkable record length and quality was notably reduced. According to this evaluation, a period between 1970s until present could be considered appropriate for this study, because is long

enough for climate studies (more than 30 years) and display an appropriated spatial density of stations (most of basins were represented).

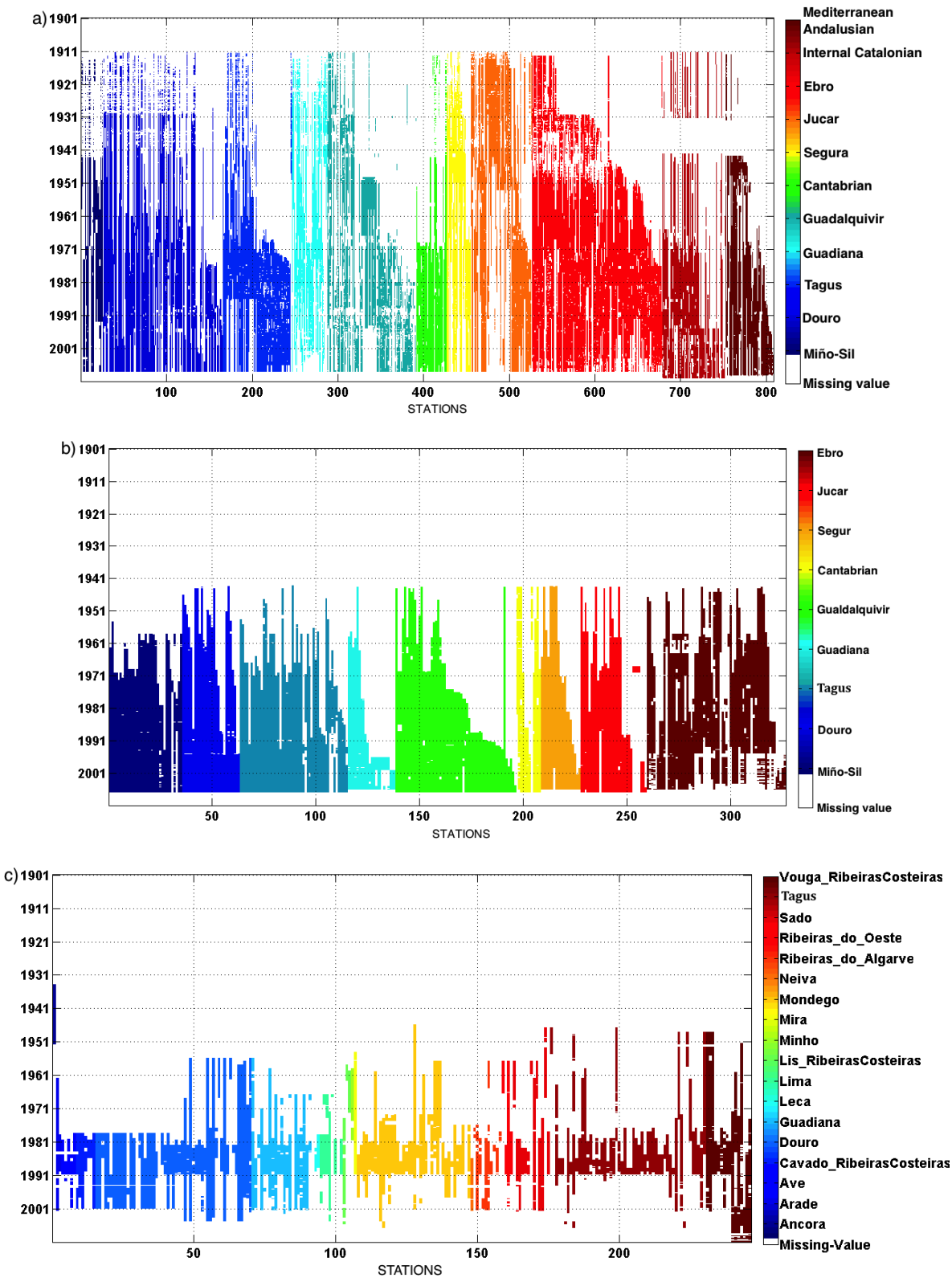


Figure 3.2. Location of non-missing values (different colours mean different basins, white means missing value), for gauging stations and stations in reservoirs in Spanish territory, a) and b), respectively, and for gauging stations in Portugal, c).

In terms of completeness of time series, only those stations with less than 10% of missing data in the selected period were considered. Then, different time periods (starting from late sixties until late seventies) were considered, and balanced between number of stations, percentage of missing values and its spatial distribution. As a result, the period starting in October 1975 and ending in September 2008 (years are considered as hydrological years, it means from October to September) was considered. The number of stations that present less than 10% of missing values on this period was 325 of the 808 stations belonging the gauging stations in Spanish territory, 9 of the 246 gauging stations in Portugal and 170 of 326 reservoir entrances in Spain, totalling 504 stations. Figure 3.3 shows the spatial location of the 504 stations selected. As it can be seen in Figure 3.3, the southern Portugal and the Guadiana Basin are the areas with poorer spatial coverage of stations.

Subsequently, monthly streamflow time series were seasonally averaged. Often, in climate studies, seasons are defined as: winter (December to February, DJF), spring (March to May, MAM), summer (June to August, JJA), and autumn (September to November, SON). However, the components of the hydrological cycle interact in multiple ways in a catchment, which introduces time lags between a decline in precipitation and when this becomes evident in other components of the hydrological cycle (Wilhite and Glantz, 1985; Vicente-Serrano and López-Moreno, 2005; Lorenzo-Lacruz et al., 2010). Thus, it is reasonable to consider a possible lag response of streamflow with respect to precipitation. Accordingly, Trigo et al. (2004) studied the response of precipitation and streamflow in the IP to the NAO and found that, whereas DJF precipitation is well related to DJF NAO, the JFM mean river flow proved to be more closely associated with the 1-month leading (DJF) NAO index than was the simultaneous (DJF) river flow. Additionally to these results, the month-to-month streamflow correlation was evaluated (see Figure 3.4). From this figure, it could be seen that March-April correlation presents smaller values than most of the month-to-month correlation, so it does not seem appropriate to include both months together in the seasonal average. Taking into account these results together with the findings from Trigo et al. (2004), March was decide to be grouped together with February and January for winter streamflow definition. Consequently, spring season was defined as April to June, summer as July to September and autumn as October to December.

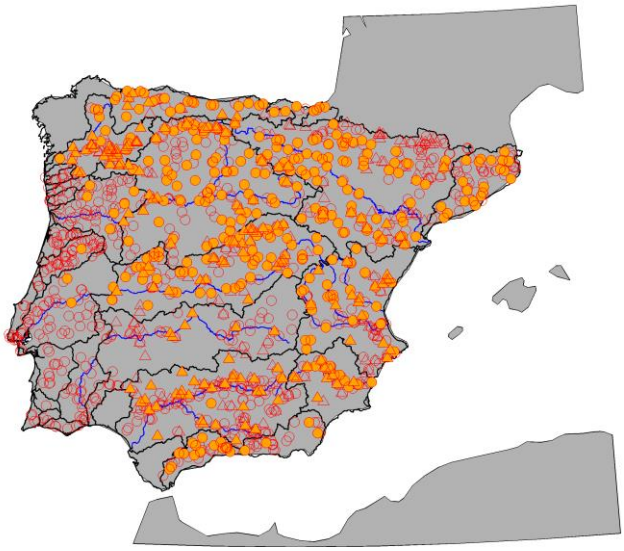

Figure 3.3. Over the raw database of 1380 stations (in red), and filled in orange, the 504 stations selected for this study (circles for gauging station and triangles for entrance reservoir stations).

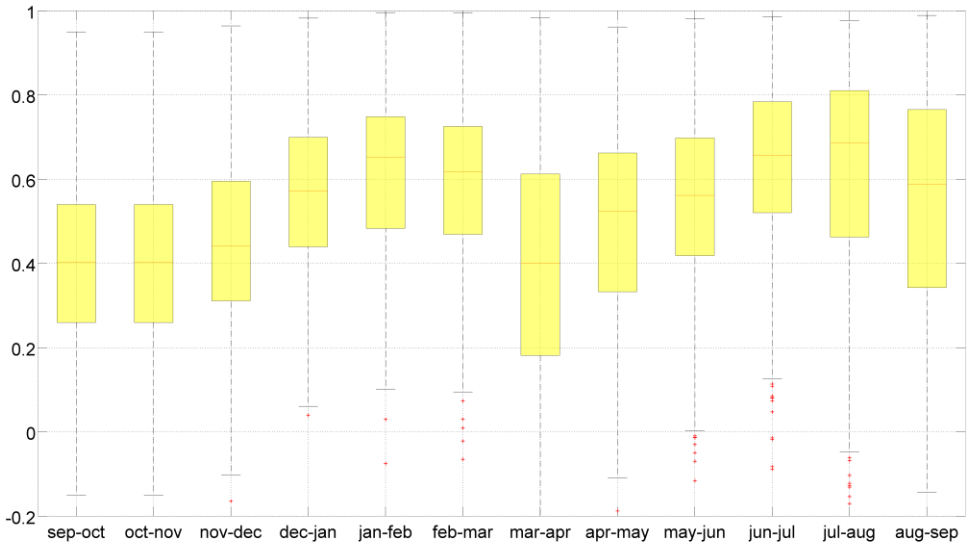


Figure 3.4. Box plot of the streamflow serial correlation (1 month lag).

Prior to apply the Pettitt test to the seasonal streamflow time series, the normality of the time series was checked by applying the Lilliefors, Jarque-Bera and Shapiro-Wilk tests. The results from the normality tests are summarized in Table 3.1 and Figure 3.5. The percentage of time series that fit to a normal distribution is quite low (from 6.9 % in case of autumn to 14.1% for spring), and mainly located in the northern sector. When a previous logarithmic transformation is done, the number of time series fitted to a normal distribution increase notably

(being between 43.6 % of stations in case of summer and 61.9 % in case of spring). However, there is still a notably percentage of stations (between 32.9 % for spring and 46.9 % for summer) that do not fit to a normal or log-normal distribution (especially located in the eastern of the IP). Hence, since around more than a third of data do not follow a normal or log-normal distribution, it seems reasonable use test that do not assume normality for homogeneity analysis, such as the Pettitt test.

| | JFM | AMJ | JAS | OND |
|---|---|---|---|---|
| **NORMAL** | 9.3 | 14.1 | 12.5 | 6.9 |
| **LOG-NORMAL** | 53.6 | 61.9 | 43.6 | 57.9 |
| **NORMAL or LOG-NORMAL** | 61.1 | 67.1 | 53.1 | 61.9 |

Table 3.1. Percentage of stations (relative to the 504 total stations analysed) where the hypothesis of normality or log-normality (or one of them) is accepted by the three tests used (Jarque-Bera, Lilliefors and Shapiro-Wilk).

Results from the Pettitt test indicated that 375 of the 504 stations selected shown a change point in at least one seasonal time series. The Figure 3.6a displays the results of Pettitt test indicating the year of the change point. In autumn, change points were found in stations located in the headwaters of Tagus, Douro and Guadalquivir and also in the Mediterranean sector (except in Catalonian Basin), especially in Segura and Jucar Basins. These changes mainly happened in the sixties (in Douro Basin) and late seventies (in the rest of the areas). In winter, the number of change points lightly increases with respect to autumn and they occurred in late seventies and early eighties in most stations. The number of change points detected presents a considerable increase in spring, particularly in the Douro, Cantabrian and Catalonian Basins. In this season, the year of the breaks varies from sixties (especially in some stations in the Douro and Cantabrian basins), late seventies and even nineties (particularly in the Cantabrian Basin and upper Ebro). Finally, summer is the season where most breakpoints were detected, in almost all of the territory, with a happening in a wide range of years (since sixties until late nineties). This result was, in somehow, expected, because the low flow levels during this season

makes that the effect of regulation (to cover the necessities for agricultural and tourism sector) be more important in relative terms.

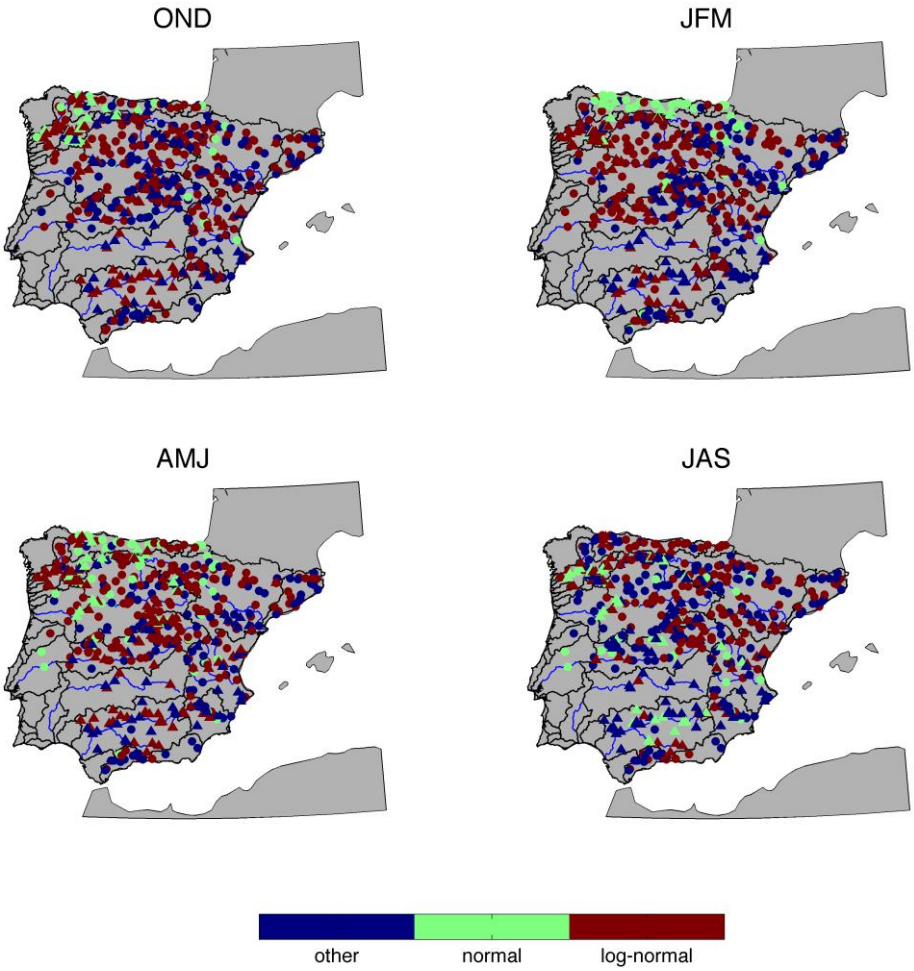


Figure 3.5. Results from the normality tests. In green, seasonal time series following a normal distribution. In red, seasonal time series following a normal distribution after a logarithmic transformation. In green, those station that neither the seasonal time series or its logarithmic transformation follow a normal distribution.

From the CAI analysis (Figure 3.6b), it could be seen as the CAI values related to break points in stations located in the north-west of the IP are above 50 %, indicating that there is not a clear sign of serious alteration in the natural regime of the river flow. However, stations located in the Mediterranean slope and also in the headwaters of Douro, Tagus, Guadiana and Guadalquivir Rivers present CAI below 50 % between the flow before and after the year of the break detected.

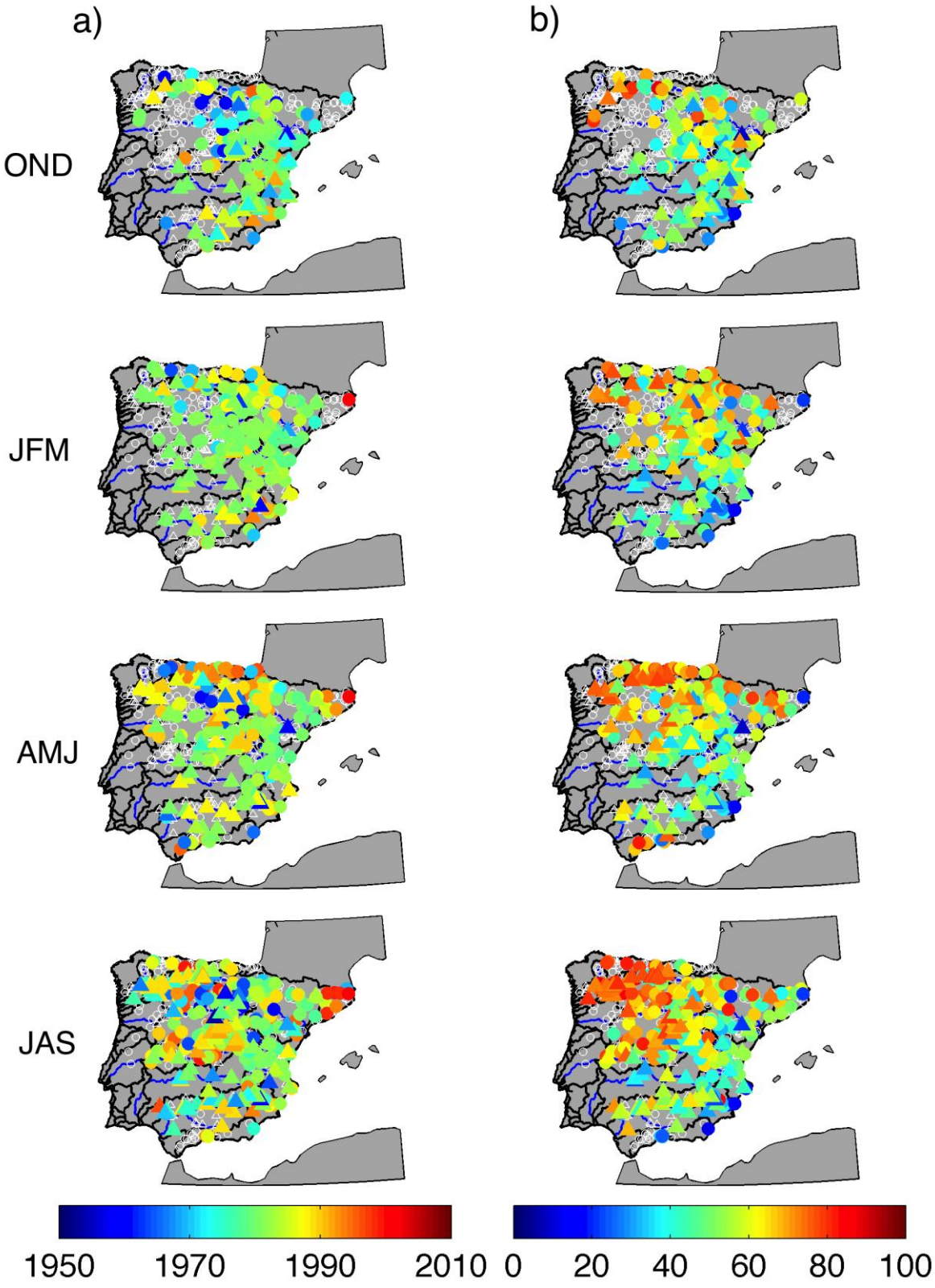


Figure 3.6. a) Year of the break point detected by Pettitt test applied to the seasonal time series. b) CAI value calculated using the data before and after the year of the break identified in the seasonal time series. Unfilled white circles indicate no break.

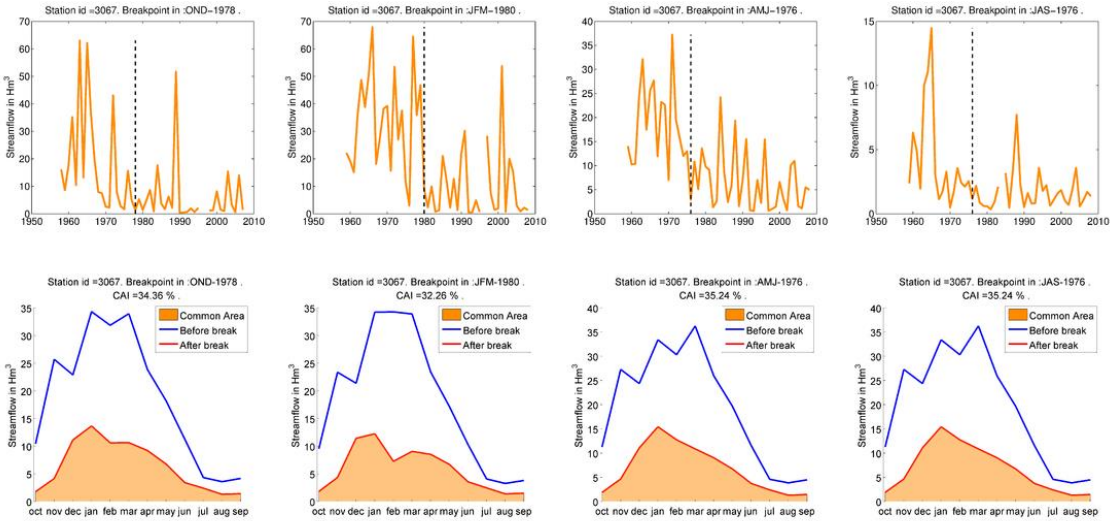


Figure 3.7. a) Seasonal time series for station id=3067. The year of the break detected by Pettitt's test is indicated with a dashed vertical line. b) Intraannual streamflow regime before (blue) and after (red) the break. The common area between both curves is colored.

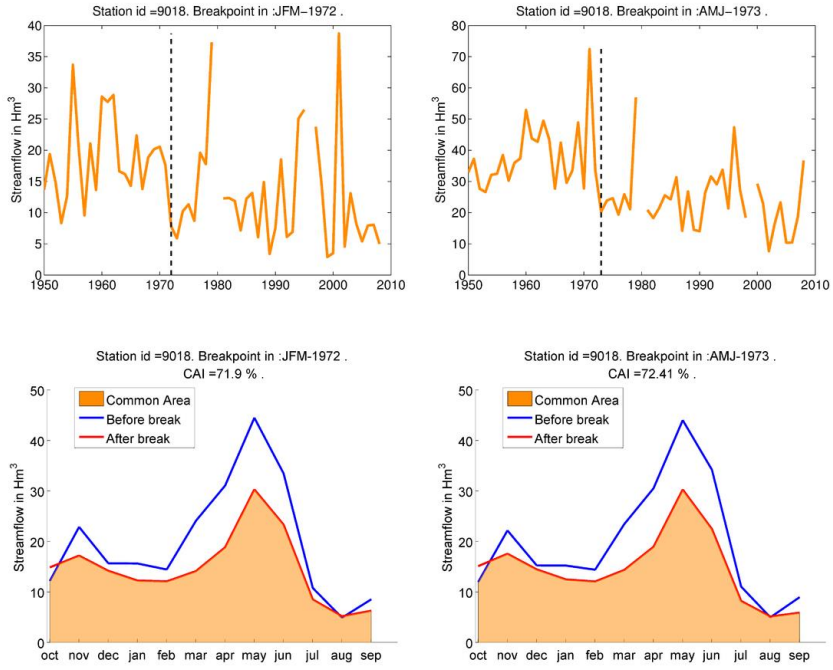


Figure 3.8. As Figure 3.7 but for station id=9018.

To sum up the process described to detect inhomogenous time series, two examples are shown. The first example (Figure 3.7) represents the analysis for a station (id=3067) located in the Tagus Basin. From the Pettitt test, a break point was found in all four seasonal time series (Figure 3.7a), around late seventies and early eighties. In all cases, the seasonal streamflow after

the break was notably inferior to the observed before that. The CAI calculated for each year identified as change point (1976 in case of autumn, spring and summer and 1980 for winter) was always inferior to 50 % (Figure 3.7b), indicating a serious alteration of the natural regime after the break. On the other hand, Figure 3.8a shows that, for the station id=9018 (in the Ebro Basin), two change points were found, in winter and spring, located in 1972 and 1973 respectively. However, the CAI computed according to these break points (Figure 3.8b) was greater than 50 %. Hence, there was not a clear evidence of alteration of the natural regime after the break point.

Finally, after the quality control carried out, 122 of the 504 stations selected were considered that present a non homogenous behaviour. Hence, a total of 382 stations overcome the quality control and were considered as adequate to this study. The location of these stations is shown in Figure 3.9.

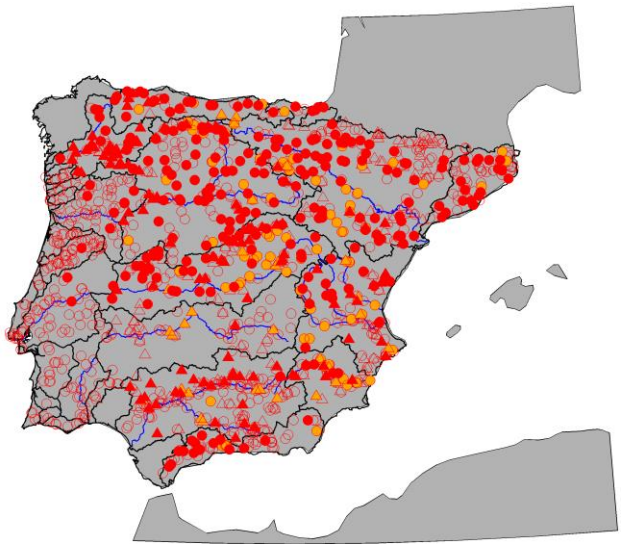

Figure 3.9. Filled in red, the 382 stations that overcome the quality control. Filled in orange, the 122 stations considered as no homogeneous. Unfilled the remaining stations up to the 1380 total original database.

## 3.4. Summary

From the original database comprises 1380 gauging stations, a trade off between number of available stations, the percentage of missing values and the length of the period was carried out, resulting in 504 stations covering the period 1975-2008, with less than 10% of missing values. Note that most of stations located in Portugal did not overcome these criteria. Moreover, the quality and homogeneity of streamflow time series was evaluated as a prerequisite for carried out reliable studies of streamflow variability and predictability. To investigate the presence of

inhomogeneities in hydrological time series (particularly because of regulation effects), a combined methodology based on Pettitt test (which not assume normality in data series) and Common Area Index was applied to the 504 seasonal time series selected. Consequently, 122 of them were discarded because of the existence of a non-natural behaviour. These stations were mainly located in mainstream of the Tagus River, Guadiana, Guadalquivir, Segura and Jucar Basins. Changes in time series were mainly identified in summer, with increase in discharge in years after the construction of a reservoir upstream (probably because of reservoir management to ensure water availability during the drier months). Sometimes this change is related to a decrease in river discharge in wet season (for water storage in reservoirs), which can be significantly enough to result in a great alteration of annual cycle. Then, for the remaining 382 stations, a reconstruction scheme was performed to fill in missing values by linear regression.

# CHAPTER 4

# SPATIAL AND TEMPORAL SEASONAL STREAMFLOW VARIABILITY

In this chapter, the temporal and spatial variability of seasonal streamflow was studied in order to identify the main spatio-temporal patterns of variability, providing a more complete description of the streamflow over the IP.

## 4.1 Introduction

Trenberth (1999) postulated an intensification of the hydrological cycle due to the increase in greenhouse gases which intensifies downwelling infrared radiation, leading higher surface temperatures that signifies greater water-holding capacity of the atmosphere as well as increased evaporation. The combination of both effects should increase atmospheric moisture. In fact, some observations and modelled simulations agree with this hypothesis (Groisman et al., 2005; Kharin and Zwiers, 2005; Tebaldi et al., 2006). In addition, some significant impacts on water resource management can be derived from increased temperature values. For example, increasing winter temperatures could reduce the amount of snow in a basin (e.g., more precipitation falling as rain than snow), as it has been observed in several parts of western United States (Aguado et al., 1992; Dettinger and Cayan, 1995). Also, higher spring temperatures could initiate earlier runoff and peak streamflow in snowmelt-dominated basins (Aguado et al., 1992;

Cayan et al., 2001). Other impacts related to the alteration of the hydrological cycle include the alteration in the magnitude and timing of runoff, frequency and intensity of floods and droughts, and also in the regional water availability.

The study of the changes in the streamflow and its temporal variability within a global change context plays an important role for water managers develop successful water policies. On this regard, a number of studies worldwide have been undertaken to characterize the variability and trends in observed records of streamflow, such as Fu et al. (2009) in United States, Burn et al. (2010) in Canada, Genta et al. (1998) in South America, Petrone et al. (2010) in Australia and Hu et al. (2011) in China. In Europe, Stahl et al. (2010) analyzed the evolution of 441 river basins in the last 40 years and showed a generalized pattern of decreasing streamflow trends in the southern part of Europe. These results are in agreement with those found for the Iberian Peninsula by Lorenzo-Lacruz et al. (2012). Also other studies regarding streamflow trends in different European regions can be found in Birsan et al. (2005), for the rivers in Switzerland, Hannaford and Marsh (2006) in the United Kingdom, Hyvarinen (2003) in Finland, Khaya and Kalaycy (2004) in Turkey, Lindstrom and Bergstrom (2004) in Sweden, Wilson et al. (2010) in the Nordic countries, and Massei et al. (2010) in France. Hence, the identification of hydrologic impacts of climate change for catchment in the IP was addressed in this chapter. To do this, the possible existence of significant trends in seasonal streamflow, their quantification and the analysis of the spatial distribution of catchments exhibiting or not trends was evaluated.

Another important aspect in the analysis of streamflow variability on the IP is the identification spatio-temporal patterns of variability that allow establishing regions with common climatic evolution. This classification is necessary because streamflow conditions over a region can vary from one area to other, even inside the same basin.

## 4.2 Methodology

### 4.2.1 Intra-annual variability

Brief descriptions of the intra-annual variability in each basin (at seasonal and monthly time scales) have been calculated to have a clearer understanding about the streamflow regimes in each basin, which become a background for the subsequent analysis carried out on this section.

### 4.2.2 Interannual variability. Trend estimation

Trends of seasonal streamflows have been estimated using the Sen's estimator of slope, also called the "median of pair-wise slopes" or Theil-Sen estimator (Sen, 1968). The statistical significance of the trends was tested using the Mann-Kendall test (Kendall, 1975; Mann, 1945) at 0.05 significance level, against the null hypothesis of no trend. This method is a rank-based test, robust to outliers and does not depend on the assumption of a Gaussian distribution of residuals. This methodology is commonly used in trend detection of hydroclimatic variables (Khaya and Kalaycy 2004; Hidalgo-Muñoz et al., 2011; Hu et al., 2011; Lorenzo-Lacruz et al., 2012; among others).

Prior to studying temporal trends, the effects of autocorrelations in the series were removed, because of this increase the probability that the test detect a significant trend (von Storch and Navarra, 1995), with an approach similar to the used for Lorenzo-Lacruz et al (2012). These authors used the pre-whitening procedure to correct the effects of autocorrelation in trends. This method is based on the 1-lag autocorrelation coefficient of the series (Burn and Hag Elnur, 2002), as follows:

$$X_t^p = X_{t+1} - r_1 X_t \qquad (4.1)$$

where $X_t^p$ is the value of the pre-whitened series for the *t* interval, *X* is the value of the original series for the *t* interval, and $r_1$ is the estimated autocorrelation coefficient.

Prior to the correction of the effects of auto-correlation, we followed the trend-free pre-whitening method proposed by Yue et al. (2002) to remove the trend (if present). The procedure involved the following steps.

1. Calculation of the slope (*b*) of the original series using a linear regression. If $b \neq 0$, this is removed from the time series as follows: $X_t^p = X_{t+1} - r_1 X_t$

$$X_t^b = X_t - bt \qquad (4.2)$$

where $X_t^b$ is the value of the detrended series for the *t* interval, $X_t$ is the original value of the series for the *t* interval, *b* is the slope value for the *t* interval.

2. Calculation of the autocorrelation coefficient *r*, between the detrended series $X_t$ and the lag-1 detrended series $X_{t-1}^b$. Removal of the auto- correlation coefficient of the detrended series was performed as follows:

$$Y_t^b = X_t^b - r_1 X_{t-1}^b \tag{4.3}$$

where the resulting series $Y_t^b$ is independent.

3. Adding the linear trend to the $Y_t^b$:

$$Y_t = Y_t^b + bt \tag{4.4}$$

This procedure ensures that the resulting series preserves the original trend and was not affected by autocorrelation.

Finally, the Mann-Kendall test was applied together with the Sen's estimator for the magnitude of trends. Since the Sen's estimator if given in $hm^3/yr$, and in order to ensure comparability among gauging stations (as magnitude of seasonal streamflow presents a notable spatial variability), it was necessary to use a relative measure of trends in the magnitude of streamflow. To do this, the strength of the change was determined by dividing the slope by the average of each series, and multiplying by 100 (to express it as percentage).

### 4.2.3 Spatial variability. Principal Components Analysis

Among the methodologies developed to classify regions with similar temporal behavior of a particular climatic variable, Principal Components Analysis (PCA; Preisendorfer, 1988) is probably the most widespread used. Examples of its application in the streamflow regionalization context can be found in Khaya et al. (2008) and Lorenzo-Lacruz et al. (2013b), among others. Here, the PCA has been used to identify the main modes of variability of seasonal streamflow. The PCA is a useful tool to reduce the dimensionality of the data, identifying dominant modes of variability. It consists in an orthogonal linear transformation of the original set into a new coordinate system. So the new variables are linear combinations of the original dataset, subject to a maximization of the covariance (or correlation).

When the measurement units of the variables differ in size and type, as it is the case of streamflow data (where there are large spatial differences between the variances), the correlation

matrix should be used instead of the covariance, because of those stations whose variances are largest will tend to dominate the first few PCs (Jollife, 2002). Therefore, taking into account the variance of stations located in mainstream of major rivers in the IP, and, especially, in the stations near to the mouth of the Douro, Tagus and Ebro Rivers, the PCA based on the correlation matrix has been carried out. To this end, streamflow were standardized to z-scores before conducting the PCA. Being mindful of the fact that the hydrological time series, in particular, seasonal streamflow time series, aim of this study, do not follow a normal distribution in most cases (as it was shown in Chapter 3), the procedure needed to standardize them requires more steps than subtract the mean and divide by the standard deviation (which assumes that data series follow a Gaussian distribution). Some studies concerning streamflow variability of the IP rivers have used different distributions to fit the streamflow time series, such as Pearson type 3 (Vicente-Serrano, 2006b; López-Moreno et al., 2009a; Lorenzo-Lacruz et al., 2010; Gámiz-Fortis et al., 2010) or lognormal (Gámiz-Fortis et al., 2011). In this study, the approach proposed in Vicente-Serrano et al. (2011) was followed. These authors selected the most appropriated distribution for each time series between a subset of three parametric distributions commonly used in hydrology. See the Appendix B for more details about this step.

Once the seasonal time series were standardized, the PCA was carried out. Although this technique is commonly used in the field of climatology, a brief explanation of PCA procedure is given in order to familiarize the lector with this.

Let us assume that the matrix $F$ is formed by the standardized time series of each station. So the dimension of $F$ is [$N$ (times) × $K$ (stations)].

We form the correlation matrix $F$ by calculating:

$$R = F^T F \tag{4.5}$$

where $F^T$ is the transpose matrix of $F$.

Note that this definition of correlation matrix is the one followed by Preisendorfer (1988). Other authors, as Peixoto and Oort (1992) use instead

$$R = \frac{1}{N} F^T F \tag{4.6}$$

However there is no difference in eigenvectors calculation but a constant factor.

The problem of finding a new coordinates system which maximizes the correlations is reduced to solve an eigenvalues problem:

$$RE = EL \tag{4.7}$$

where $L$ is a diagonal matrix containing the eigenvalues $l_i$ of $R$ (non negatives). The $E_i$ columns vectors of $E$ are the eigenvectors of $R$ corresponding to the eigenvalue $l_i$. The size of both $E$ and $L$ is [$K \times K$].

The property $EE^T = E^T E = I$ (where $I$ is the identity matrix) means that the eigenvectors are uncorrelated over space (they are orthogonal to each other). Because of this, they are commonly named Empirical Orthogonal Functions (EOFs).

The eigenvalues in the diagonal of $L$ are in descending order, so the first eigenvector is aligned in the direction in which the data vectors jointly exhibit the most variability (understood as spatial correlation in this case), the second one is calculated so that it is orthogonal to the previous principal component and represents the second spatial structure more coherent, and so on.

The trace of $L$ is equal to the trace of $R$, and represents the maximum variance of the dataset. Hence, the percentage of variance explained in the direction of the eigenvector $E_i$ is:

$$Variance = \frac{l_i}{trace(L)} 100 \tag{4.8}$$

Each eigenvector represents a standing pattern when is plotted as a map (for each station, a weight is given, and a map is displayed by plotting these weights in the location of each station). The time evolution of these eigenvectors is calculated by projecting the original dataset of standardized time series onto the new coordinate system:

$$Pc = FE \tag{4.9}$$

where columns of the matrix $PC$, dimensions [$N \times K$] represent the time evolution of the pattern related to each eigenvector. The columns of the $Pc$ matrix are usually called Principal Components (PCs) or Expansion Coefficients, and, as the EOFs are uncorrelated in space, they are uncorrelated in time. So, a PC groups the temporal behaviour of the set of stations identified by the related EOF, and its variance is associated with its eigenvalue.

A common way to represent the EOFs is through the correlation map between the PCs associated with the eigenvector and the original dataset $F$. These correlations are called loading factors. When the matrix of correlations is used, instead covariances, these loading factors are a direct interpretation of spatial variability patterns (they are proportional, by the square root of the eigenvalue, to the eigenvectors).

The identification of the number of principal components that can be considered as significant is still a matter of controversy. There are different methods to address this issue, such as a method called *The North Rule of Thumb* (North et al., 1982), based on the degeneration of eigenvalues. That is, if the spacing between two consecutive eigenvalues is smaller than the error of the first one, then they cannot be considered as different true eigenvalues. Other selection rules are the *scree plot* (a graphical tool), where the eigenvalues are plotted and the number of component to be retained are the number of components before a break in the curve; *The Kaiser Rule*, which considers significant those eigenvalues greater than the average value of eigenvalues (usually 70% of this value), or also the number of components that account a minimum (fixed) threshold of the total variance.

Bearing in mind that the atmospheric processes are not independent, the orthogonal constraint might sometimes hamper the physical interpretation because the examination of independent modes of variability might not be justified (North, 1984; Wilks, 2006). In order to circumvent this problem, the significant eigenvectors are often rotated with the aim to produce a simple structure in the results. Among the different techniques that exist to rotate the eigenvectors (Richman, 1986), the Varimax rotation was employed here.

## 4.3 Results

### 4.3.1 Seasonal streamflow regime. Intra-annual Variability

Before to present the spatial and temporal variability, the seasonal averaged streamflow values at the gauging stations are shown (Figure 4.1). It could be appreciated as gauging stations located in the mainstream of the major IP Rivers present larger values of these parameters (whose differences goes to various degrees of magnitude). These high differences are not only found for the average values but also for the variances (not shown). As it was mentioned in the description of the basins in Chapter 2, it could be seen that stations belonging rivers in Atlantic slope present higher seasonal discharges than those located in the Mediterranean and Cantabrian slopes. On this regard, the northwestern quadrant of the IP presents the higher seasonal streamflow regimes.

Additionally, the intra-annual variability is also presented (Figure 4.2), showing the monthly and seasonal streamflow values (averaged between all stations) in each basin. From this Figure, it is discernible a common feature in the Miño-Sil, Douro, Tagus, Guadiana and Gualdaquivir Basins, where winter months (from December to March) present the higher

streamflow averages, diminishing through spring months till reaching their minimum values during summer. In the Cantabrian and Ebro Basins, these high values showed in winter are extended till May, mainly due to the snowmelt contribution (from Pyrenees and Iberian System). However, it could be appreciated as the Júcar and Segura basins present fairly regular interannual contributions, being particularly noteworthy the case of Segura basin, where summer months show high average values. This is likely due to these basins are subjected to high regulation regimes to ensure water availability in summer months (especially because demands for tourism and irrigation purposes are high in this area). Also, it is worth to mention that in the Internal Catalonian Basins, the autumn, winter and spring averages are fairly similar. This could be explained from the fact that some rivers have their headwaters in the Pyrenees, which ensure high levels during spring because of snowmelt contributions, whereas other gauging stations near to the coast present high values in autumn (since the Mediterranean sector presents high precipitation values in this season as a consequence of convective storms in the late summer and early autumn), such as in the case of the Mediterranean Andalusian Basin.

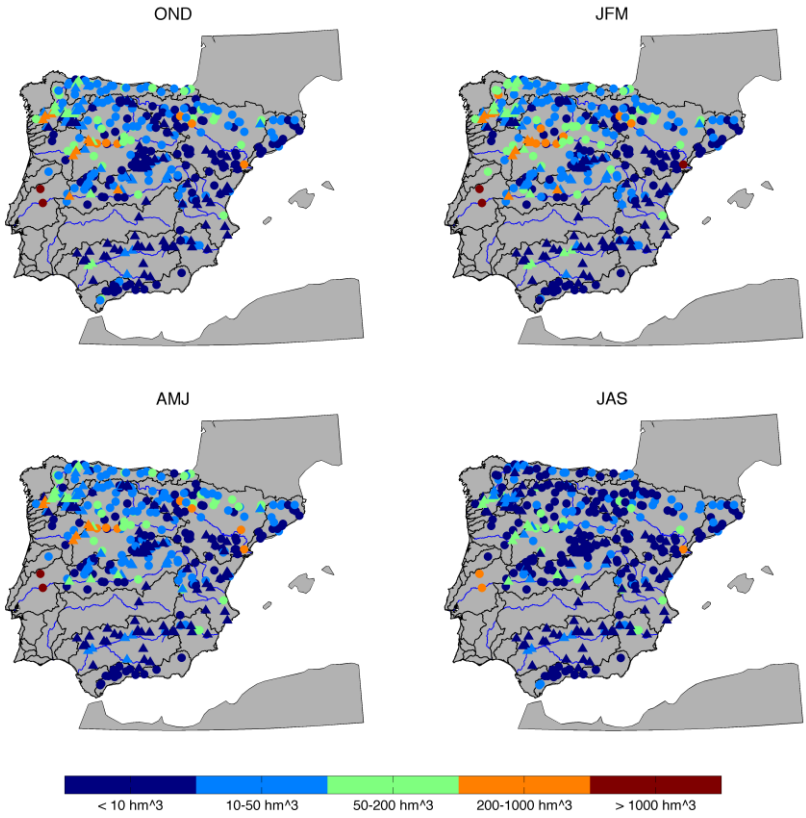


Figure 4.1. Seasonal streamflow averages for the different streamflow gauging stations analysed for the period 1975-2008.

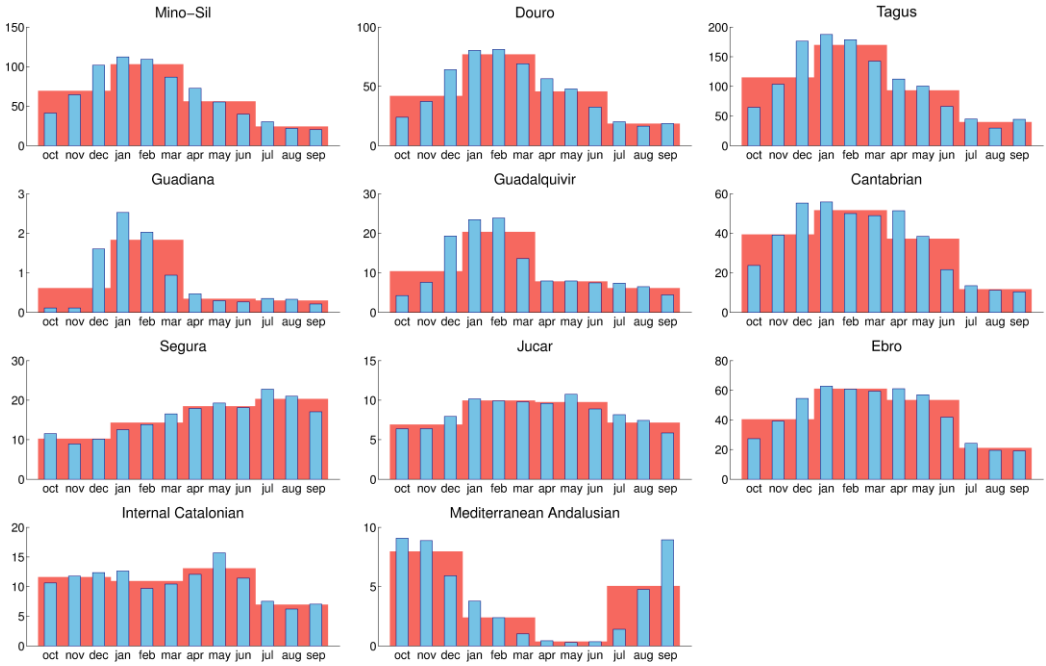


Figure 4.2. Seasonal (red bars at the background) and monthly (blue bars at the front) streamflow average for all stations in each basin.

Seasons with the maximum streamflow contributions are presented in Figure 4.3 for each basin. In particular, gauging stations of most part of the IP exhibit the maximum contribution during winter. The stations that present their maximum during spring are located in the Ebro, (because of the high contributions of snowmelt from surrender ranges, Pyrenees and Iberian system). Autumn is the season when stations located in the Mediterranean slope exhibits their maximum contributions, whereas only few stations present their maximum in summer, which could be due to effects of regulation.

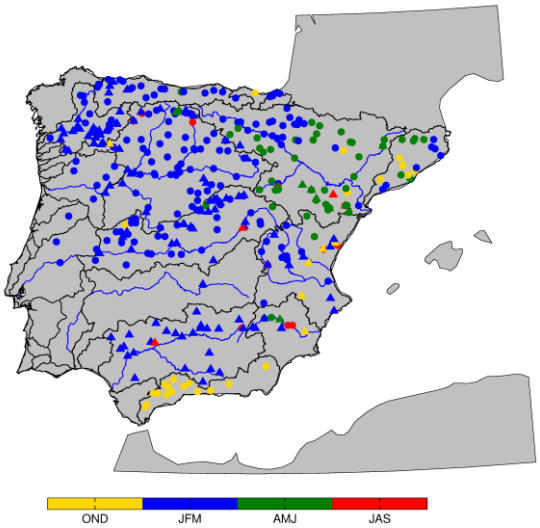


Figure 4.3. Season with the highest streamflow averages at the different gauging stations.

### 4.3.2 Temporal variability of seasonal streamflow

The interannual variability of the averaged (between all gauging stations) seasonal streamflow for each basin has been displayed in Figure 4.4. From this Figure, it could be noted a dry period in early nineties is observed in all basins. Also, wet periods are observed, as in winter of 1996 (in all basins) and 2001 (mainly in the Miño-Sil, Douro and Tagus Basins). Last seventies present wet years (in almost all basins), along with autumn and winter of 1996/1997 hydrological year, with high streamflow values in south-western basins.

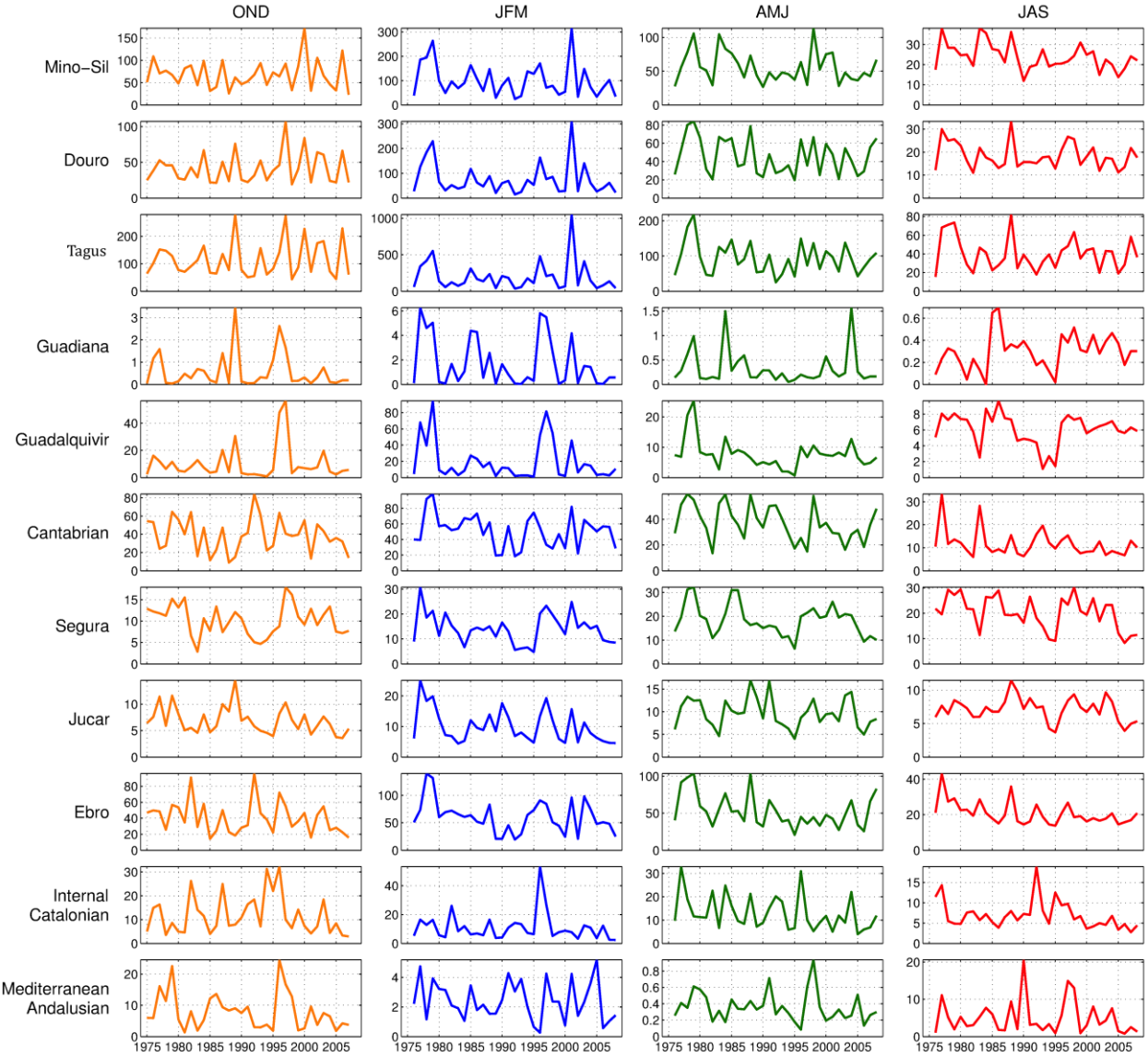


Figure 4.4. Interannual variability of the seasonal streamflow averages ($hm^3$) for each basin.

Figure 4.5 displays the results obtained from the trends analysis. From this Figure, it can be observed that significant trends located in the eastern half of the IP show most of them a

negative sign, with magnitudes ranging from -4 to 4 per cent by year. In particular, in autumn, significant negative trends are located in eastern part of IP, in headwaters of the Segura and Jucar Basins, along the Ebro River, in upper Tagus, and also in some stations in the Guadalquivir Basin and Mediterranean Andalusian Basin. On the contrary, some stations near the Mediterranean coast present high significant positive trends, which could be related to an increase in convective precipitation on Mediterranean façade during autumn, which also is the season with maximum streamflow in this area. Also, significant upward trends appear in stations placed in the Douro basin, especially in headwaters of the northern tributaries of Douro River. In winter, significant trends were found in the Miño Basin, upper part of the Douro Basin, headwater of the Tagus River, upper and medium part of the Jucar, Segura and Guadalquivir Basins, and except for some gauging stations, most of these trends present negative sign. In the case of spring streamflow, the significant trends exhibit a similar spatial pattern to that found in winter, with predominantly negative trends, and with the difference that a minor amount of stations presents significant trend in the Miño Basin and, on the contrary, most of stations in the Cantabrian Basin presents significant downward trends. During summer, significant trends are found in most of the basins. Whereas the northern and eastern half of the IP presents, generally, decreasing trends, as in previous seasons, it is remarkable the stations with significant upward trends observed in in the central and western of the IP, in particular in the Douro, Tagus and Guadalquivir Basins.

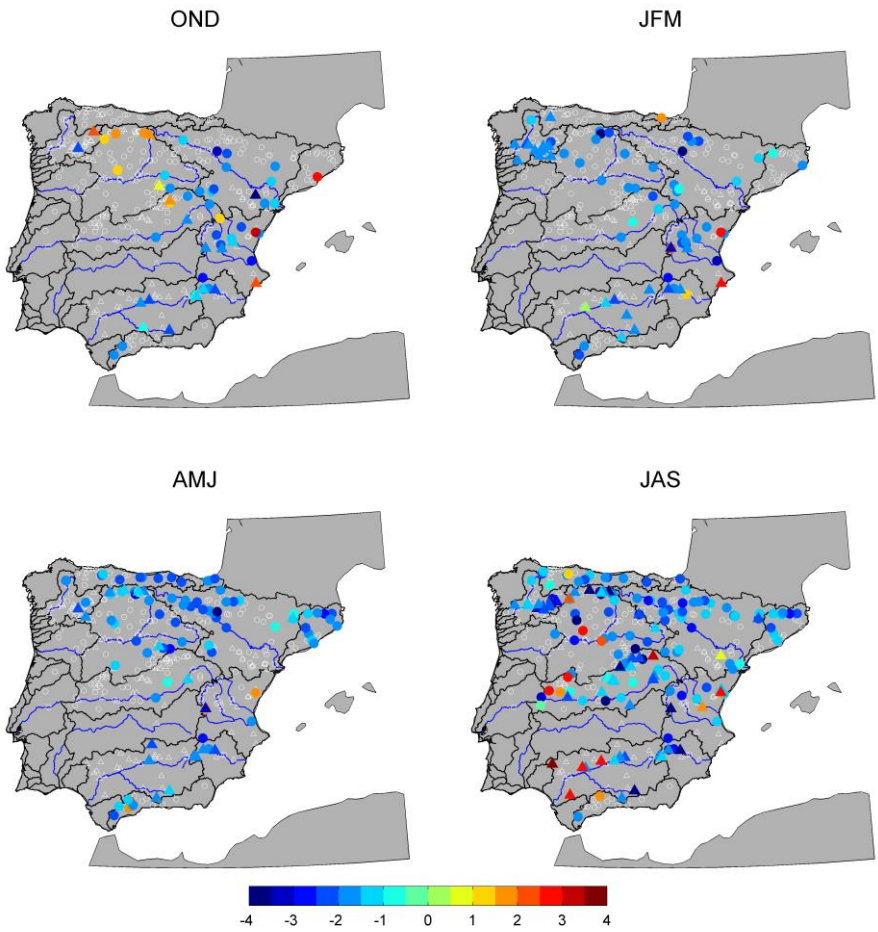


Figure 4.5. Spatial distribution of significant seasonal trends. The marks are filled when the trends are significant at 95% confidence level. Values are in percentage of change per year.

### 4.3.3 Spatial variability of seasonal streamflow

After an evaluation of the information provides by the tests for determining the number of significant components in PCA, five components were considered as significant in the four principal components analyses (one for each season). The percentage of total variance explained by the five first components are: 71.38% for autumn, 78.41% for winter, 66.11% for spring and 51.43% for summer streamflow.

The loading factors (correlation between the time series of each principal component and the seasonal streamflow) related to the rotated Principal Components are shown in Figure 4.6 as an indicator of the stations represented by each principal component. Taking into consideration that the sign of loading factors only refer to the eigenvector sense, spatial patterns for autumn and winter are rather similar, where EOF1 represents stations in the central and northern sectors of the IP rivers with Atlantic slope, accounting for 33.90% (45.57%) of the variance in autumn

(winter). EOF2 is associated with the northern IP rivers in the Cantabrian slope, explaining 13.94% (11.90%) of the variance for autumn (winter). Southern half of the IP is represented by EOF3, with differences between winter and autumn. Despite both EOF3, for autumn (10.53% of variance) and winter (9.05% of variance), group stations in southern and central part of the IP, in case of autumn this pattern include most of stations of the Mediterranean Andalusian Basin, whereas EOF3 of winter does not (these stations are grouped into the winter EOF5). Internal Catalonian Basins and some stations in lower Ebro are classified into EOF4 for both seasons, presenting a slightly difference in the explained variances (7.19% in autumn and 6.92% in winter). Finally, EOF5 (which explains 5.80% in autumn and 5.39% in winter) is related to stations located near the Ebro Delta (and some stations of the Mediterranean Andalusian Basin in winter, as it was noted before). Conversely to what happens for autumn and winter, spatial patterns identified for spring and summer present more dissimilarity, and also there is less difference between the percentages of variance associated with each pattern, mainly in summer. In the case of spring, the first EOF explains 20.64% of the variance and groups stations in the north-western of the IP. The second EOF (14.18% of the variance) represents stations in the central and western of the IP. The third EOF is mainly associated with stations in the Guadalquivir Basin and explains 13.09% of the variance. The fourth EOF (13.08%) groups stations in the Ebro Valley and Internal Catalonian Basins. The fifth EOF seems to group stations in the Mediterranean Andalusian Basin, although the spatial pattern is not as clear as the previous ones. Finally, the spatial patterns obtained for summer group the north-western of the IP (EOF1, 17.14% of variance), the central and southern half of the IP (EOF2, 13.00% of variance), the Ebro Delta and Catalonian Basins (EOF3, 9.10% of variance) and the Mediterranean Andalusian Basin (EOF4, 6.82% of variance). The EOF5 does not make a clear grouping of stations.

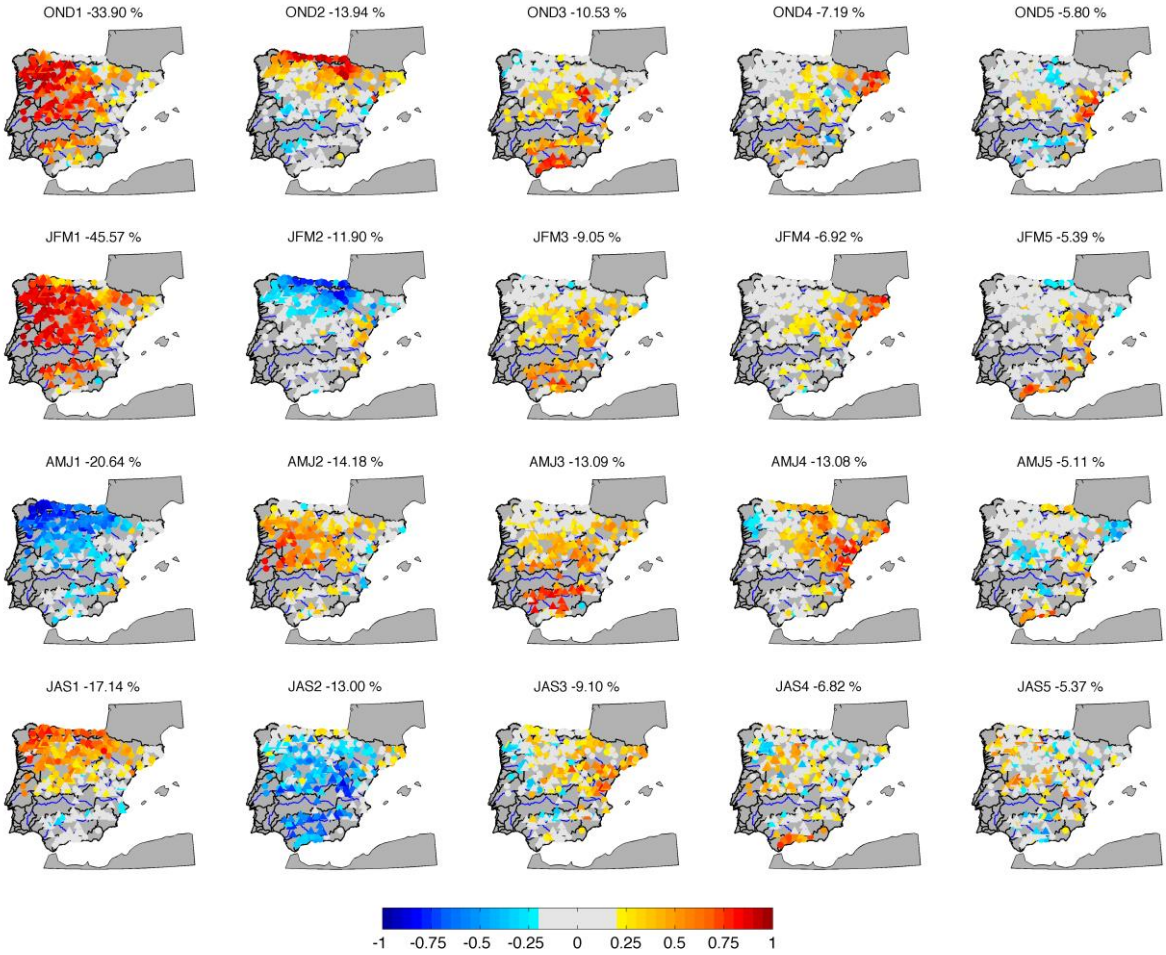


Figure 4.6. Loading factors of rotated components 1-5 (in columns) from the PCA analysis of the seasonal streamflow series (in rows).

## 4.4 Summary

To sum up, the analysis of temporal trends in seasonal streamflow reveals a common pattern of decreasing of streamflow, especially in the eastern part of the IP, in all seasons. These negative trends agree with the result observed by other authors. In particular, Lorenzo-Lacruz et al. (2012), using a database of 187 stations covering 1945-2005, found a marked decrease in winter and spring streamflows in most of the IP Basins during the period 1945-2005. Also, Morán-Tejeda et al. (2011b) and López-Moreno et al. (2011) highlighted dominant negative trends in various sub-catchments of the Duero and Ebro Basins, respectively, during the second half of the Twentieth Century. In addition, these decreasing trends agree with the found by other authors in catchments in the south of Europe, such as in Stahl et al. (2010), Mavromatis and Stathis (2010) and Lespinas et al. (2010). Conversely with this general finding, some stations present positive significant trends, such as some stations in northern part of the Douro Basin and

the Miño-Sil Basin during autumn. Also, positive trends are observed in some stations during summer, which are likely influenced by reservoir operations; the filling of reservoirs occurs during the maximum annual flows generated by winter and spring precipitation (López-Moreno et al., 2009a), and water is progressively released during the dry season to guarantee the summer water supply for human consumption and irrigation. One of the factor responsible of these negative trends could be the climate variability in the Mediterranean region, which has been associated with generally negative rainfall trends during the second half of the 20th century (Xoplaki et al., 2004; Dünkeloh and Jacobeit, 2003), which is more pronounced in winter (López-Moreno et al., 2009b). Specifically for the IP, it has been observed a decrease in precipitation (Rodrigo and Trigo, 2007; Mourato et al., 2010, González-Hidalgo et al., 2011) and in winter snow accumulation (Sanz-Elorza et al., 2003; López-Moreno and García-Ruiz, 2004; López-Moreno, 2005), which can directly influence the decline in spring flows. In addition to the observed decrease in precipitation and snow accumulation, there are other factors than can contribute to the observed decrease in streamflow, such as water management strategy (filling the reservoirs during the wet season to guarantee water supply in summer), changes in land cover/land use (increasing the forested land within the IP, and expanding of the irrigated surface).

Regarding the spatial patters of seasonal streamflow, it is found a similar grouping in autumn and winter (the wettest seasons) according to PCA results, when stations in the Miño-Sil, Douro, Tagus, Guadiana and Guadalquivir Basins are grouped into the first mode, stations in the Cantabrian and Upper Ebro Basins in the second mode, and stations in the Mediterranean Andalusian and Internal Catalonian Basins in the third and fourth modes. On the other hand, spatial patterns differ in spring and summer (drier half of the year), displaying a north-south gradient in the first two or three modes. In addition, it is reduced the difference between the variance explained by the first modes with respect to winter and autumn.

# CHAPTER 5

# ASSESSING TELECONNECTION INDICES AS POTENTIAL PREDICTORS OF SEASONAL STREAMFLOW OF THE IP RIVERS

In this chapter, the main teleconnection patterns controlling the large-scale atmospheric and oceanic variability are evaluated as potential predictors of autumn, winter and spring streamflow of the IP Rivers. Several lag times (from one up to four seasons) between teleconnection indices and streamflow are considered. The main plausible physical causes explaining the relationships found are also discussed.

## 5.1 Introduction

The global atmospheric circulation presents a number of preferred spatial patterns of variability. This variability occurs in well-defined spatial patterns (Wallace and Gutzler, 1981; Barnston and Livezey, 1987), especially during the boreal winter over the Northern Hemisphere (NH). Such variations can be associated with some large-scale oscillations in atmospheric dynamics (e.g., planetary waves, jet streams, and monsoons) resulting in simultaneous or lagging variations in weather and climate over widely separated points over the earth. These variations

are commonly referred as "teleconnections", a term that is defined in The Glossary of Meteorology (Glickman, 2000) as:

1. *A linkage between weather changes occurring in widely separated regions of the globe.*
2. *A significant positive or negative correlation in the fluctuations of a field at widely separated points.*

Climate analysis is facilitated by the construction of a teleconnection map, which describes the linkage between a region of interest and all other points in the domain that are farther than the de-correlation length scale of the variable. Teleconnection maps thus provide information about the structure of recurrent climate variability that is characterized by the correlation-at-a-distance feature.

Teleconnections, because of its definition as climate links between geographically separated regions, have been commonly used as predictor of variables such as precipitation, temperature or streamflow, specially through the variability of other variables, such as Sea Level Pressure (SLP) or Sea Surface Temperature (SST) in remote areas. The main goal of this chapter is to identify which of the large-scale variability modes (teleconnection patterns) have an influence on seasonal streamflow in the IP some seasons later.

## 5.2 Teleconnection indices

### 5.2.1 Definition

In this chapter, the main teleconnection indices that summary the main large-scale atmospheric and oceanic variability indices used are: Artic Oscillation (AO), North Atlantic Oscillation (NAO), East Atlantic pattern (EA), Pacific/North American pattern (PNA), Western Pacific pattern (WP), East Pacific North Pacific pattern (EP-NP), Scandinavian pattern (SCAND), East Atlantic Western Russian pattern (EA-WR), Atlantic Multidecadal Oscillation (AMO), Pacific Decadal Oscillation (PDO), North Pacific Gyre Oscillation (NPGO), Indian Ocean Dipole (IOD), Southern Oscillation Index (SOI), ENSO indices (specifically: El Niño-1+2, Niño-3, Niño-4, and Niño-3.4 and SOI), El Niño Modoki (EMI), Western Mediterranean Oscillation (WeMO) and the Snow Advance Index (SAI). The monthly time series of AO, NAO, EA, EA-WR, SCAND, AMO, WP, EP-NP, PNA, Niño1+2, Niño3, Niño4, Niño3.4, SOI and PDO were obtained from the Climate Prediction Center (CPC) at the National Center of

Environmental Predictions (NCEP; www.cpc.noaa.gov). The procedure used for CPC to identify the Northern Hemisphere teleconnection patterns and indices is the Rotated Principal Component Analysis -RPCA (Barnston and Livezey 1987) of Geopotential Height at 500 hPa. The monthly time series of EMI and IOD were obtained from the Japan Agency for Marine-Earth Science and Technology website (www.jamstec.go.jp/). The monthly time series of WeMO index were downloaded from the website http://www.ub.edu/gc/English/wemo.htm. The time series of NPGO were obtained from the website http://www.o3d.org/npgo/. The time series of SAI were kindly provided by the Justin Jones in personal communication. A brief definition of each index together with a schematic map indicating its location (Figure 5.1) is provided. Some of them, such as NAO, AO, SCAND, EA, WeMO, ENSO indices or SAI, have been selected because they are specifically related to the climate in the IP (it will be widely discussed in the next section). The rest of the teleconnection indices were selected for exploratory purposes. That is, although the literature does not contain specific evidence that connects them to the IP climate, it would be worthwhile to determine whether they are useful for improving long-range forecast over the study area.

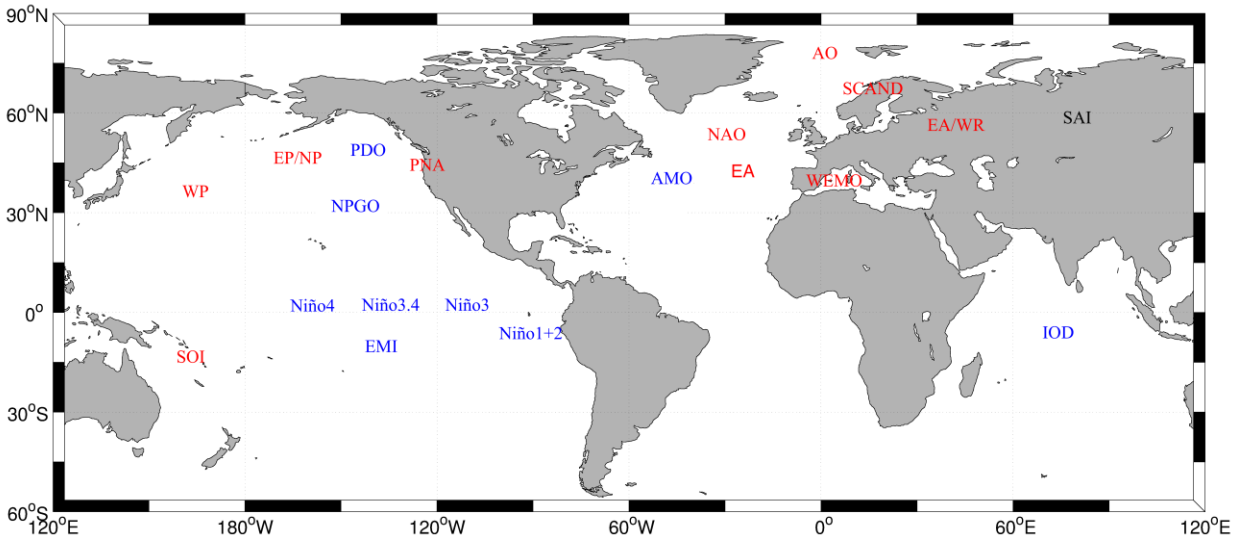


Figure 5.1. Approximate location of the main areas associated to the teleconnection indices used in this study. In blue, indices based on data from sea surface temperature. In red, indices calculated using data from atmospheric pressure at different levels. In black, indices computed using other variables.

Artic Oscillation (AO)

The AO index is usually defined as the first EOF of the mean sea level pressure field in the Northern Hemisphere, and it is a robust result from EOF analysis of this field on timescales from weeks to decades in any season (Thompson and Wallace, 1998). It describes the relative intensity of a semi-permanent low-pressure centre over the North Pole. A band of upper-level winds circulates around this centre, forming a vortex. When the AO index is positive and the vortex intense, the winds tighten like a noose around the North Pole, locking cold air in place. However, with a negative AO and a weak vortex, it is allowed intrusions of cold air to plunge southward into North America, Europe, and Asia.

North Atlantic Oscillation (NAO)

The North Atlantic Oscillation is one of the most prominent teleconnection patterns in all seasons (Barnston and Livezey, 1987), which especially influences on the northern Atlantic Ocean (Visbeck et al., 2001). The NAO consists of a north-south dipole of pressure anomalies, with one centre located over Greenland and the other centre of opposite sign spanning the central latitudes of the North Atlantic, between 35°N and 40°N. It has also been identified as part of the Northern Annular Mode, the Arctic Oscillation (Marshall et al., 2001). The positive phase of the NAO reflects below-normal heights and pressure across the high latitudes of the North Atlantic and above-normal heights and pressure over the central North Atlantic, the eastern United States and Western Europe. The negative phase reflects an opposite pattern of height and pressure anomalies over these regions. Both phases of the NAO are associated with basin-wide changes in the intensity and location of the North Atlantic jet stream and storm track, and in large-scale modulations of the normal patterns of zonal and meridional heat and moisture transport (Hurrell, 1995), which in turn results in changes in temperature and precipitation patterns often extending from eastern North America to western and central Europe (Walker and Bliss, 1932; van Loon and Rogers, 1978; Rogers and van Loon, 1979).

East Atlantic pattern (EA)

The East Atlantic pattern is the second main pattern of low-frequency variability over the North Atlantic, and appears as a leading mode in all months, although is more prominent in winter. Structurally is similar to the NAO, with a north-south dipole of pressure anomaly centres spanning the North Atlantic from east to west. The anomaly centres of the EA pattern are displaced southeastward to the approximate nodal lines of the NAO pattern (often interpreted as a “southward shifted” NAO pattern). However, the lower-latitude centre contains a strong

subtropical link in association with modulations in the subtropical ridge intensity and location. This subtropical link makes the EA pattern distinct from its NAO counterpart. The EA pattern exhibits very strong multi-decadal variability in the 1950-2004 record, with the negative phase prevailing during much of 1950-1976, and the positive phase occurring during much of 1977-2004.

East Atlantic-Western Russia pattern (EA-WR)

The East Atlantic/West Russia pattern is one of three prominent teleconnection patterns that affect Eurasia throughout year. The EA-WR pattern consists of four main geopotential height anomaly centres. The positive (negative) phase is associated with positive (negative) height anomalies located over Europe and northern China, and negative (positive) height anomalies located over the central North Atlantic and north of the Caspian Sea.

Atlantic Multi-decadal Oscillation (AMO)

The Atlantic Multi-decadal Oscillation represents a persistent SST pattern observed in the northern Atlantic Ocean between 0º – 70ºN bounded by the continents (Gray et al., 2004). In contrast to the other indices, the AMO does not vary drastically, and it is by far the index with the longest cycle period. It is estimated from modelling studies to be between 50-70 years, and from tree-ring and Arctic ice regression as closer to 80 years (Schlesinger and Ramankutty, 1994; Kerr, 2000). Analyses of global climate from measurements dating back to the nineteenth century show the AMO as a leading large-scale pattern of multidecadal variability in surface temperature. This AMO cycles are known to alternately mask and amplify the anthropogenic impact of global warming (Knight et al., 2006). Over the instrumental period (1856-Present) the AMO exhibited warm phases at roughly 1860 – 1880 and 1930– 1960 and cool phases during 1905 – 1925 and 1970 – 1990. The AMO appears to have returned to a warm phase beginning in the mid 1990s.

Scandinavian pattern (SCAND)

The Scandinavian pattern consists of a primary circulation centre over Scandinavia, with weaker centres of opposite sign over Western Europe and eastern Russia/western Mongolia. The positive phase of this pattern is associated with positive height anomalies, sometimes reflecting major blocking anticyclones over Scandinavia and western Russia, while the negative phase of the pattern is associated with negative height anomalies in these regions.

Western Pacific pattern (WP)

The Western Pacific pattern is a primary mode of low-frequency atmospheric variability over the North Pacific in all months. During winter and spring, the pattern consists of a north-south dipole of anomalies, with one centre located over the Kamchatka Peninsula and another broad centre of opposite sign covering portions of southeastern Asia and the western subtropical North Pacific. Therefore, strong positive or negative phases of this pattern reflect pronounced zonal and meridional variations in the location and intensity of the entrance region of the Pacific (or East Asian) jet stream. These anomalies exhibit a strong northward shift from winter to summer, which is consistent with the observed northward shift of the East Asian jet stream. A third anomaly centre is located over the eastern North Pacific and southwestern United States in all seasons.

East Pacific-North Pacific pattern (EP-NP)

The East Pacific/North Pacific pattern is a spring-summer-fall pattern with three main anomaly centres. The positive phase of this pattern features positive height anomalies located over Alaska/Western Canada and negative anomalies over the central North Pacific and eastern North America. Strong positive phases of the EP-NP pattern are associated with a southward shift and intensification of the Pacific jet stream from eastern Asia to the eastern North Pacific, followed downstream by an enhanced anticyclonic circulation over western North America, and by an enhanced cyclonic circulation over the eastern United States. Strong negative phase of the pattern is associated with circulation anomalies of opposite sign in these regions.

Pacific/North American pattern (PNA)

The Pacific/North American teleconnection pattern is one of the most prominent modes of low-frequency variability in the Northern Hemisphere extratropics. The PNA is comprised of four teleconnection centres at 500mb (Wallace and Gutzler, 1981) or 700 hPa geopotential height fields (Barnston and Livezey, 1987). The positive (negative) phase of the PNA pattern is defined as two above-average (below-average) height centres located in the vicinity of Hawaii (20ºN, 160ºW) and Alberta (55ºN, 115ºW), along with two below-average (above-average) height centres located in the vicinity of the northern Pacific Ocean (45ºN, 165ºW) and the Gulf Coast region (30ºN, 85ºW). The linear combination of the normalized height anomalies at these centres specifies the PNA index. The PNA pattern is associated with strong fluctuations in the strength and location of the East Asian jet stream. The positive phase is associated with an enhanced East Asian jet stream and with an eastward shift in the jet exit region toward the

western United States. The negative phase is associated with a westward retraction of that jet stream toward eastern Asia, blocking activity over the high latitudes of the North Pacific, and a strong split-flow configuration over the central North Pacific.

El Niño-Southern Oscillation (ENSO)

El Niño-Southern Oscillation (ENSO), see overviews by Philander (1990) and Neelin et al. (1998), is an anomalous large-scale ocean-atmosphere phenomenon associated with strong fluctuations in ocean currents and surface temperatures. ENSO is a major example of the interrelationship between ocean currents and atmospheric conditions and it consists of two components: El Niño and the Southern Oscillation. El Niño (Halpert and Ropelewski, 1992) involves warming of the tropical Pacific surface waters, weakening the usually strong SST gradient across the Pacific, with associated changes in ocean circulation. It's linked with an atmospheric counterpart, the Southern Oscillation (SO), which involves changes in trade winds, tropical circulation and precipitation. During El Niño events (warmer than normal central and eastern equatorial Pacific SSTs), the low-level atmospheric winds are weaker along the equator and there is an enhanced convection across the entire equatorial Pacific. During La Niña events (cooler than normal central and eastern equatorial Pacific SSTs), the low-level atmospheric winds are stronger along the equator and convection decreases across the entire equatorial Pacific.

Currently, there are several indices used to define the ENSO phenomenon. The most commonly used indices include several sea surface temperature (SST) areas such as the Niño1+2 (80º-90ºW,0º-10ºS), Niño3 (90ºW-150ºW, 5ºN-5ºS), Niño3.4 (120ºW-170ºW, 5ºN-5ºS) and Niño4 (150ºW-160ºE, 5ºN-5ºS), and an index created from sea level pressure (SLP) differences between Darwin (Australia) and Tahiti, called the Southern Oscillation Index (SOI).

El Niño Modoki (EMI)

El Niño Modoki is a coupled ocean-atmosphere phenomenon in the tropical Pacific (Ashok et al., 2007). This phenomenon appears as the second dominant mode of interannual variability in the tropical Pacific (being El Niño the first one). Conventional El Niño is characterized by a strong anomalous warming in the eastern equatorial Pacific, whereas El Niño Modoki is associated with a strong anomalous warming in the central tropical Pacific and a cooling in the eastern and western tropical Pacific. Also, teleconnections are very different from those related to the conventional El Niño.

Pacific Decadal Oscillation (PDO)

The Pacific Decadal Oscillation is the second most common climatic index in the Pacific. The PDO is defined as the first mode of the principal component analysis of the Pacific Ocean SSTs poleward of 20ºN. During a warm or positive phase, the west Pacific becomes cool and a part of the eastern ocean warms; during a cool (or negative) phase, the opposite pattern occurs. PDO fluctuations have a period between 50-70 years with a given warm or cool phase persisting for about 15–25 years (Chao et al., 2000; Minobe, 1997). Several independent studies find evidences for just two full PDO cycles in the past century: cool PDO regimes prevailed from 1890-1924 and again from 1947-1976, while warm PDO regimes dominated from 1925-1946 and from 1977 through (at least) the mid-1990's.

North Pacific Gyre Oscillation (NPGO)

The North Pacific Gyre Oscillation is a climate pattern that emerges as the second dominant mode of sea surface height (SSH) variability in the Northeast Pacific (Di Lorenzo et al., 2008; 2010). The NPGO mode closely tracks the second EOF of North Pacific SST anomalies, also referred to as the "Victoria Mode" (Bond et al., 2003). The NPGO is distinct from the Pacific Decadal Oscillation (PDO) mode (Mantua et al., 1997), which emerges as the first mode of SST anomalies and SSH anomalies in the Northeast Pacific.

Indian Ocean Dipole (IOD)

Large interannual variability of SST in the Indian Ocean has been associated with the Indian Ocean Dipole (IOD), also referred to as the Indian Ocean Zonal Mode (IOZM; Saji et al., 1999; Webster et al., 1999). This pattern manifests through the difference in sea surface temperature between two areas, the Arabian Sea (western Indian Ocean) and the eastern Indian Ocean (south of Indonesia). This dipole, characterized by cooling (warming) in the eastern ocean and warming (cooling) in the western ocean, is referred to as the positive (negative) phase of the IOD. The phase changes in the IOD imply the anomaly of moisture transport, which impacts many regions significantly.

Western Mediterranean Oscillation (WeMO)

The Western Mediterranean Oscillation was defined by Martin-Vide and Lopez-Bustins (2006) within the synoptic framework of the western Mediterranean basin and its vicinities. It is a pressure dipole between the Po Plain, in the north of the Italian Peninsula, an area with a relatively high barometric variability, and the Gulf of Cádiz, in the southwest of the IP, often

subject to the influence of the Azores anticyclone and, episodically, to the cut off of circumpolar lows or to its own cyclogenesis In positive phase of WeMO, below/above normal pressure anomalies appear in the Po Plain/Gulf of Cádiz, with a prevalence of northwesterly winds from the Atlantic Ocean. During negative phase, above/below normal pressure anomalies appear in the Po Plain/Gulf of Cádiz, favouring the airflow entrance from the Mediterranean.

Snow Advance Index (SAI)

The Snow Advance Index (SAI), accounts how rapidly the Eurasian snow cover increases in October (Cohen and Jones, 2011). Specifically, it is a standardized index that measures the rate of increase (by means of the regression coefficient of the least square fit) of Eurasian snow cover in October, as described by the regression coefficient of the least squares fit of the daily (available since 1997) or weekly (available from 1973, from NOAA's satellite-sensed observations) Eurasian snow cover extension equatorward of 60ºN.

### 5.2.2 Influence of teleconnection indices on European region

Two of the most important teleconnection patterns are El Niño/Southern Oscillation (ENSO) and the North Atlantic Oscillation (NAO). The ENSO encompasses two linked phenomena, a quasi-periodic warming in the tropical Pacific near South America and an adhering effect, the zonal anomaly of Sea Level Pressure (SLP) in the tropical Pacific (the Southern Oscillation). The NAO is the most prominent and recurring extratropical teleconnection pattern in the Northern Hemisphere (NH). Generally, the NAO refers to the meridional seesaw, or the dipole structure, typically associated with the north center near Iceland and the south center near the Azores of the SLP field. These main teleconnection patterns have been commonly used as potential predictors of hydrological variables. For example, several studies have shown significant relationships between ENSO events and precipitation or streamflow variability at the global scale (Dettinger et al., 2000; Chiew and McMahon, 2002), and, particularly, for the European area (van Oldenborgh et al., 2000; Mariotti et al., 2002; Park, 2004; Zanchettin et al., 2008; García-Serrano et al., 2011; Shaman and Tziperman, 2010) and the IP has been also addressed (Rodó et al., 1997; Pozo-Vázquez et al., 2005; Vicente-Serrano, 2005; Lorenzo et al., 2010; Frías et al., 2010; Vicente-Serrano et al., 2011). The importance of winter NAO as the major atmospheric mechanism controlling the precipitation during wintertime in the central and western IP has been described by several authors (Zorita et al., 1992; Corte-Real et al., 1995; Rodó et al., 1997; Rodríguez-Puebla et al., 2001; González-Rouco et al., 2000;

Trigo and Palutikof, 2001). Also, effects of NAO on winter streamflow in Iberian Rivers have been identified (Trigo et al., 2004; López-Moreno et al., 2007; Lorenzo-Lacruz et al., 2011; Morán-Tejeda et al., 2011a).

There are other teleconnection patterns that account for an important percentage of the climatic variability and have an influence on IP climate. For example, Rodríguez-Puebla et al. (2001) studied the influence of North Atlantic Oscillation (NAO), East-Atlantic (EA), the East-Atlantic/Western-Russia (EA-WR) and the Polar/Eurasian (POL) patterns on winter precipitation in the IP. In addition, it has been found that the EA and EA-WR patterns represent a significant contribution for the precipitation over the northern IP (Sáenz et al., 2001) as well as parts of the eastern Mediterranean areas (Quadrelli et al., 2001; Xoplaki et al., 2004; Krichak et al., 2002). De Castro et al. (2006) found a significant contribution of the NAO and also Scandinavian (SCAND), EA-WR and EA patterns in explaining the main variability of precipitation and river-flow regimes in Galicia (north-west of IP). Casanueva et al. (2014) examined the relationship between the variability of extreme precipitations events in Europe and different teleconnection patterns, such as the Arctic Oscillation (AO), NAO, SCAND, EA, EA-WR, the Southern Oscillation Index (SOI), the Atlantic Multidecadal Oscillation (AMO) and the Madden–Julian Oscillation (MJO). The Western Mediterranean Oscillation (WeMO; Martin-Vide and Lopez-Bustins, 2006) has been related to precipitation episodes in the Mediterranean façade of IP (Martin-Vide and Lopez-Bustins, 2006; Lopez-Bustins et al., 2008; Martin-Vide et al., 2008; Hidalgo-Muñoz et al., 2011) and with IP river discharge (Martín i Díaz, 2010; Lorenzo-Lacruz et al., 2013b).

During the last years, there has been an increase in studies about the influence of the variability in Siberian snow cover during autumn on the climate of the following winter on the northern hemisphere (NH) and its connection to the winter Artic Oscillation (Cohen et al., 2001; Bojariu and Gimeno, 2003; Cohen and Saito, 2003; Saito and Cohen, 2003; Cohen and Fletcher, 2007). For instance, Cohen and Entekhabi (1999) showed that the Eurasian snow cover in early fall is significantly correlated with the following winter Arctic Oscillation (AO). These authors hypothesize that a possible dynamical mechanism linking Eurasian snow cover anomalies and North Atlantic climate variability is through the strength and position of the Siberian high. The conceptual model is explained by Cohen (2011) and briefly summarized here. The presence of snow cover can increase the amount of sunlight reflected back into space from 20% to 80% (Cohen and Rind, 1991). Also snow cover is a good insulator or a thermal blanket, preventing heat from the ground escaping into the atmosphere. These radiative properties of snow cover

cold the atmosphere above the earth's surface. The presence of snow cover can lead to much colder temperatures than the absence of snow cover. This cold and dense air promotes a stronger and more expansive Siberian high. As a consequence, the amount of energy transfer from troposphere (or lower stratosphere) to the stratosphere (or upper troposphere) can be increased. Often, when such an excess energy exists, this is absorbed in the polar stratosphere, leading a warming of the polar stratosphere and a weakening of the polar vortex (a fast stream of air that flows west to east around the pole and it derives its energy from the temperature gradient pole-equator in the upper troposphere). In case of weakening of polar vortex, the flow of air is weaker turning north-south (rather than west-east), favouring that cold air from the Arctic penetrates into the mid-latitudes and warm air from the subtropics is carried into the Arctic. This mixing of air masses favours more storms and snow in the mid-latitudes. During a weak polar vortex, high pressure occurs in the Arctic region (negative AO phase). In contrast, below normal snow cover anomalies during fall are related to a positive phase of AO in the following winter. This index has proved being a reliable predictor of winter climate conditions over the IP (Brands et al., 2012; 2013).

In addition to these commonly used teleconnection indices, others related to the North Atlantic climate have been examined in this chapter. The North Pacific has also been hypothesized to influence decadal variations in different geographical regions, through the Pacific Decadal Oscillation (PDO; Baik and Paek, 1998; Tomita et al., 2001). Particularly, the teleconnection between the PDO and rivers in Europe has been investigated by various authors (Dettinger and Diaz, 2000; Rimbu et al., 2002; Bardin and Voskresenskaya, 2007; Labat, 2010; Gámiz-Fortis et al., 2011). Also, The Pacific North Atlantic (PNA) has been proposed as one of the links between the tropical forcing and the extratropical circulation response in the North Atlantic area (Pozo-Vázquez et al., 2005). In addition to these indices, the ability of other patterns related to the North Pacific Ocean, such as the Western Pacific pattern (WP), the East Pacific-North Pacific pattern (EP-NP) and the North Pacific Gyre Oscillation (NPGO, Di Lorenzo et al., 2010). Also, the Indian Ocean has been identified as possibly modulating ENSO variability through an extension of the Walker Circulation to the west and associated flow of warm tropical ocean water from the Pacific into the Indian Ocean (e.g. Jansen et al., 2009; Izumo et al., 2010; Frauen and Dommenget, 2012). The potential influence of the Indian Ocean on IP climate, especially over river-flow variability, has been pointed out (Hernández-Martínez et al., 2014). Here, we used the Indian Ocean Dipole (IOD), a western-eastern dipole of sea surface temperature differences in the Indian Ocean in order to evaluate the teleconnection with Indian

Ocean. Finally, the Atlantic Multidecadal Oscillation (AMO) was included, which has also been related to streamflow variability in France (Oubeidillah et al., 2012).

## 5.3 Stable predictors selection

The seasonal teleconnection indices were defined as averages of 3-month periods, autumn (September to November, 'son'), winter (December to February, 'djf'), spring (March to May, 'mam') and summer (June to August, 'jja'). Teleconnection indices from up to four seasons in advance were evaluated as potential predictors. For example, in case of spring streamflow forecast, the teleconnection indices of the previous spring, summer, autumn and winter were evaluated. When a '1' is added to these acronyms means the season belongs to the previous year. The identification of the teleconnection indices that can be used as predictors for seasonal streamflow at each gauging station was carried out by evaluating the significance and stability of point linear correlation between the seasonal streamflow and the teleconnection indices.

Before determining the critical correlation coefficient, it has to be addressed if serial correlation could affect its calculation. Note that if serial correlation is not accounted for, positive serial correlation could lead to too many type-1 errors because of an artificial lowering of the correlation coefficient's p-value, arising from the fact that the number of temporally independent data pairs is lower than the sample size (Trenberth, 1984; Kristjánsson et al., 2002). The effective sample size ($N_{eff}$) is given, according Bretherton et al. (1999), by:

$$N_{eff} = N\left(\frac{1-r_1 r_2}{1+r_1 r_2}\right) \qquad (5.1)$$

where $r_1$ and $r_2$ are the serial correlation (lag 1) of the time series 1 and 2 (streamflow and teleconnection index), respectively, which assumes that the time series follow a first-order autoregressive process.

Finally, the significance of the correlation between the streamflow data and the potential predictor considered is determined using

$$r_{cri} = \frac{t_{95}}{\sqrt{r_{95}^2 + N_{eff} - 2}} \qquad (5.2)$$

where $r_{cri}$ is the critical correlation coefficient; $t_{95}$ is the 95$^{th}$ quantile of the t-distribution with $N_{eff}$ -2 degrees of freedom; and $N_{eff}$ is the effective sample size. A correlation value is considered as significant if the calculated correlation coefficient $r$ (in absolute value) is bigger than $r_{cri}$.

A seasonal time series corresponding to a teleconnection index was considered a stable predictor if it showed stable teleconnection with at least 20 stations (around 5% of the total) or with a set grouped stations. To evaluate the stability of the correlations, a similar approach to the one used by Ionita et al. (2008) and Gámiz-Fortis et al. (2010) was followed. The correlation between seasonal streamflow and teleconnection indices was calculated for 19 moving windows of 15 years (running 1 year). The correlation between predictor and predictand is considered to be strongly stable (stable) when more than 80% of the total 19 moving windows present significant correlations at 90% (80%) level, are two sided, and show no change in correlation signs. Note that the serial correlation was taken into account in determining the critical value for correlation significance

## 5.4 Results

Prior the evaluation of the stability of the significant correlations between seasonal teleconnection indices leading seasonal streamflow time series, a special reference to the SAI is needed. The SAI measures the ratio of the increase in the extent of the snow cover during October in Eurasia (through the regression coefficient between snow cover and time). In this regard, there are two versions of the SAI (both were kindly provided by Dr Justin Jones). The first one, which we called ‘SAIw’, covered the period from October 1972 to October 2010 and was calculated using weekly data of snow cover (i.e. four points, as weeks in October, are used in the regression fit). In the second, which we called ‘SAId’, each value was calculated as the regression coefficient using daily data of the extent of snow cover (i.e. 31 points are used in the regression fit), available since 1997. It was correlated the seasonal streamflow with both indices to evaluate the possible variations related to the differences in how they were calculated (basically in the number of points used in the regression fit). The most noteworthy results were found in the case of winter streamflow, with 248 stations showing significant correlations (with values above 0.344, significant at 95% confidence level with 33 years) with ‘weekly SAI’, 0.478 being the average of all these values. On the other hand, 255 stations showed significant correlation for the ‘daily SAI version’ (with values above 0.602, significant at 95% confidence level with 11 years) their average being 0.761. At this point, it seems that the SAId shows higher

correlation with winter streamflow. Additionally, the correlations using SAIw in the same period as SAId (i.e., 1997-2007) were computed and compared with SAId. In this case, only a few stations significantly correlated with '1997-2007 SAIw', reinforcing the idea that the daily version of SAI provides more trustworthy measures of the advance of snow cover in October in Eurasia and presents higher correlations with winter streamflow in the IP. Therefore, to make the potential forecasting ability provided by SAId useful (since the length of the SAId time series is still short), we created a combined SAI index, composed of the data from the 'weekly' version of SAI from 1974 to 1996 and the 'daily' version of SAI from 1997 to 2007. Hereafter, the results refer to this 'combined' SAI.

An example of the stability analysis is shown in Figure 5.2. It could be seen as SAI is a stable predictor of winter (JFM) streamflow of gauging station id=3153 (in Tagus Basin) since all 15-year running window correlations are above the critical threshold. This critical threshold at a given significance level (80% level, two-sided, in this study) is calculated accounting the effect of possible serial correlation in both time series (streamflow and predictor). As it can be seen in Figure 5.2, this threshold is the same in case of winter and spring but different (slightly higher) in case of autumn (the existence of serial correlation diminishes the degree of freedom). Similarly, NAO of previous winter appears as a stable predictor for autumn (OND) streamflow in the gauging station id=10020 (sited in Mediterranean Andalusian Basin). On the contrary, SOI of previous summer was not considered as a stable predictor of spring (AMJ) streamflow in gauging station id=1607 (in Cantabrian basin) since there was less than 80% of running windows with a significant correlation at the established significance level and, also, the sign of correlation in running windows changed after 1991-2006 (from negative to positive).

A summary of the stability of correlations between seasonal streamflow and teleconnection indices is shown in Table 5.1 and Figure 5.3. This figure indicates the number of gauging stations that present significant and stable (and also strongly stable) correlations with teleconnection indices corresponding to different seasons. An asterisk indicates the seasonal teleconnection indices that present stable correlations with at least 20 stations (~5% of the total) or, in case of less than 20, if they are grouped together belonging to a basin represented by less than 20 stations in the database used here (such as the Mediterranean Andalusia or the Internal Catalonian Basin). From this figure, it is worth highlighting the large number of stations where the OND streamflow is significantly correlated with ENSO indices of the previous autumn. However, the number of stations where these correlations are stable was lower. This feature (significant but not stable correlations) was also found in correlations between OND streamflow

and WeMO in the previous spring and SCAND and SOI in the previous summer. A total of 18 teleconnection indices (those with an asterisk in Figure 5.3) of the 77 evaluated are considered stable predictors for autumn streamflow. For JFM streamflow, it bears noting that few teleconnection indices (only 8 of the 77 analyzed) were identified as stable predictors (with asterisk in Figure 5.3). For example, there is a notable number of stations with significant but unstable correlations with WP and SCAND in previous winter and summer, respectively. A remarkable case is that of SAI, which was found to be a stable predictor of winter streamflow in 238 stations. In the case of AMJ streamflow, 16 of the 77 seasonal teleconnection indices evaluated were considered stable predictors (with asterisk). Notably, none of the teleconnection indices in the previous spring was considered to be a stable predictor, most of them being in previous summer. Additionally, significant but unstable correlations appear with SOI in the previous summe

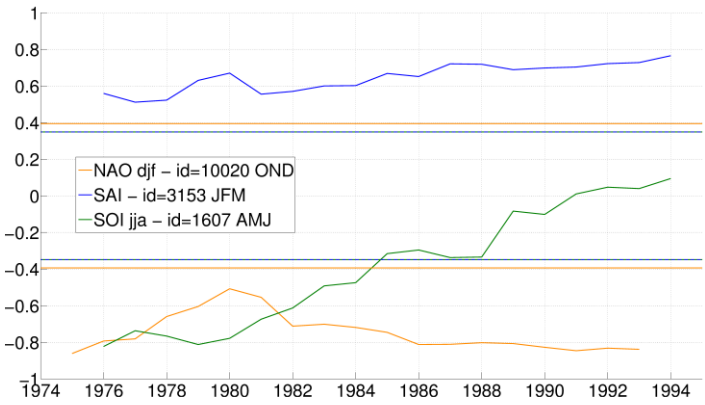


Figure 5.2. Running correlation (15-year windows) between autumn streamflow in station id=10020 and previous winter NAO (in dark orange), winter streamflow in station id=3153 and previous October SAI (in blue) and spring streamflow in station id=1607 and previous summer SOI (in green). Straight dark orange, blue and green lines represent the threshold of 80% confidence level for the previous cases. The value of correlation for each moving window is indicated in the first year (i.e., correlation in 1994 is referred to 1994-2008 period).

The correlation maps between stable predictors (with asterisk in Figure 5.3) and seasonal streamflow are displayed in Figures 5.4 to 5.6 for autumn, winter and spring, respectively. Regarding the predictors corresponding to the previous autumn (Figure 5.4), the most relevant is the presence of various indices related to the ENSO phenomenon (Niño3, Niño4, El Niño3.4, and SOI). Although with some minor differences, they present stable correlations with stations located in the Douro and Tagus Basins, but also some stations located in the upper Jucar and Guadalquivir Basin). El Niño4 is the index that presents stable correlation with the largest

number of stations (45, Table 5.1). These correlations are not high in magnitude (around 0.4 in absolute value). The SAI of the previous October displays a stable correlation with stations in the northern mountains of the Douro Basin and the Mediterranean Andalusian Basin (reaching values of around 0.56 in this basin). The NAO and AO of the previous winter were found to be stable predictors, particularly for stations in Mediterranean Andalusian Basin. In this case, the correlation values reach up to 0.7 and 0.6 for NAO and AO, respectively. Although the winter, spring, and summer WeMO appear as a potential predictor, they do not display a similar correlation map. Winter WeMO correlates stably with stations in the Mediterranean Andalusian Basin, whereas spring and summer WeMO present stable correlations with stations in the northern basins (mostly in Miño-Sil, Douro, and upper Ebro). Spring and summer East Atlantic were found to be a stable predictor for stations in Miño-Sil Basin and also some located in the headwaters of the Douro and Tagus. The EP-NP of the previous spring also shows stable correlations with a remarkable number of stations (74, Table 5.1) situated in the northwestern quadrant of the IP. The Niño3.4 and SOI in summer present stable correlations with some stations in northern half of IP. However, it bears noting that the sign of these correlations is the opposite of that found for the previous autumn ENSO indices.

| | **OND** | | | | **JFM** | | | | **AMJ** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | son1 | djf | mam | jja | djf1 | mam1 | jja1 | son1 | mam1 | jja1 | son1 | djf |
| **NAO** | 1 | 15 | 7 | 6 | 1 | 2 | 1 | 0 | 4 | 2 | 3 | 17 |
| **EA** | 0 | 1 | 21 | 26 | 0 | 2 | 7 | 1 | 7 | 8 | 0 | 2 |
| **WP** | 14 | 6 | 15 | 0 | 3 | 67 | 0 | 0 | 1 | 0 | 3 | 1 |
| **EP-NP** | 2 | 0 | 74 | 6 | 0 | 18 | 2 | 4 | 5 | 20 | 2 | 2 |
| **PNA** | 10 | 0 | 2 | 1 | 2 | 1 | 13 | 0 | 0 | 33 | 6 | 3 |
| **EA-WR** | 0 | 0 | 3 | 3 | 1 | 0 | 1 | 0 | 1 | 2 | 2 | 7 |
| **SCAND** | 1 | 5 | 1 | 6 | 2 | 3 | 6 | 45 | 1 | 9 | 1 | 18 |
| **PDO** | 1 | 0 | 0 | 9 | 4 | 1 | 1 | 5 | 0 | 4 | 10 | 1 |
| **AMO** | 2 | 1 | 1 | 1 | 2 | 1 | 1 | 1 | 16 | 13 | 3 | 2 |
| **AO** | 0 | 20 | 1 | 3 | 1 | 2 | 0 | 0 | 0 | 53 | 4 | 10 |
| **Niño1+2** | 9 | 3 | 1 | 1 | 1 | 1 | 0 | 0 | 4 | 37 | 54 | 28 |
| **Niño3** | 20 | 4 | 9 | 8 | 6 | 0 | 0 | 3 | 4 | 33 | 5 | 4 |
| **Niño4** | 45 | 22 | 2 | 5 | 4 | 2 | 4 | 2 | 5 | 7 | 4 | 2 |
| **Niño3.4** | 29 | 7 | 8 | 16 | 4 | 2 | 1 | 2 | 5 | 11 | 4 | 2 |
| **SOI** | 41 | 5 | 3 | 16 | 4 | 5 | 2 | 1 | 1 | 7 | 6 | 3 |
| **EMI** | 2 | 3 | 0 | 0 | 5 | 1 | 3 | 2 | 3 | 6 | 10 | 10 |
| **NPGO** | 3 | 0 | 1 | 9 | 0 | 0 | 0 | 0 | 4 | 4 | 12 | 3 |
| **IOD** | 9 | 1 | 0 | 3 | 0 | 2 | 2 | 4 | 8 | 10 | 27 | 4 |
| **WeMO** | 2 | 8 | 47 | 17 | 6 | 5 | 1 | 9 | 2 | 2 | 0 | 2 |
| **SAI** | 21 | - | - | - | - | - | - | 238 | - | - | 7 | - |

Table 5.1 Number of stations with stable correlations between seasonal streamflow (OND, JFM and AMJ) and teleconnection indices in the preceding seasons (son1, djf, mam, jja in case of OND, djf1, mam1, jja1 and son1 for JFM and mam1, jja1, son1 and djf for AMJ).

For winter streamflow, Figure 5.5 displays the correlation maps for the stable predictors. Although previous winter EMI and WeMO shows stable correlation with few stations, they are grouped. A notable number of stations (67, Table 5.1) present stable correlations with previous spring WP, particularly in the Douro, Tagus, and Guadalquivir Basins. However, the magnitude of these correlations is weak (0.40 in average). Previous spring EP-NP, summer PNA, and autumn SCAND also show stable but weak correlations (around 0.40 and 0.45 on average), mainly in the Douro, Tagus, and Ebro for spring EP-NP, the Douro and Tagus for summer PNA and, in case of autumn SCAND, the Miño-Sil, Douro, Tagus, and Ebro Basins. Previous autumn WeMO also present stable correlations in stations on the Mediterranean slope. Finally, the most remarkable results are found for the SAI of the previous October, which has stable correlations with 238 stations (Table 5.1), except with those located near the Cantabrian and Mediterranean slopes. In addition, the average of these correlations is higher (around 0.6) than observed with all the previous indices.

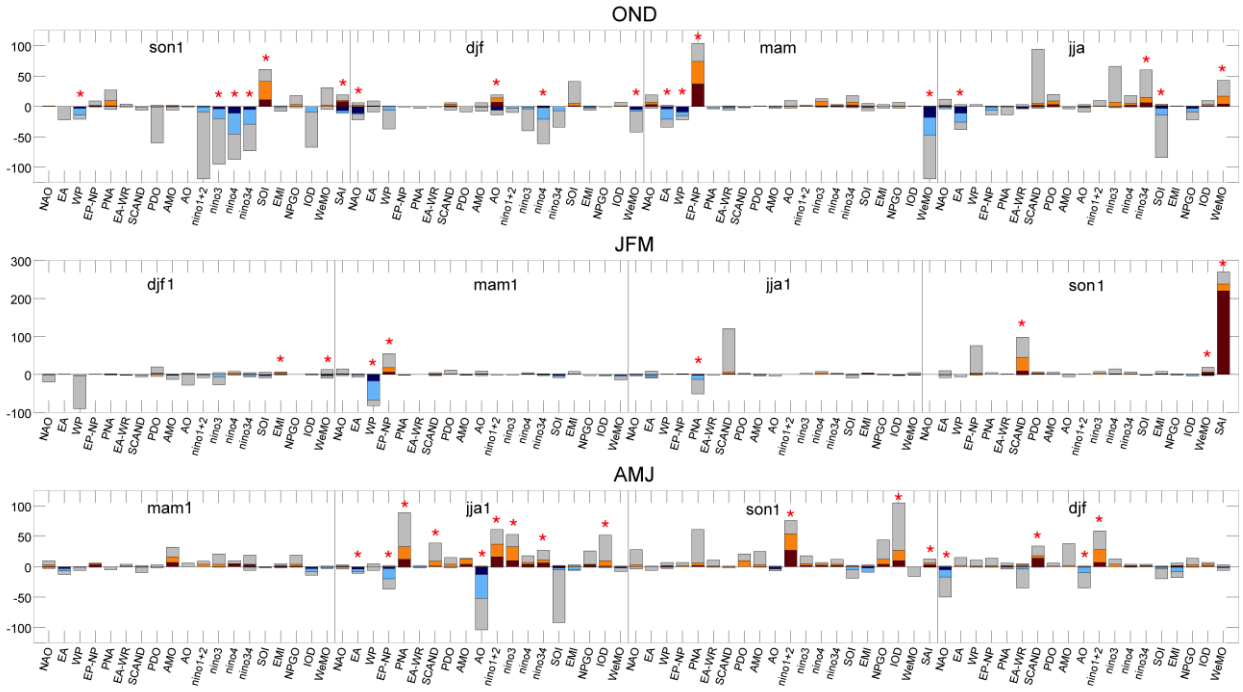


Figure 5.3 Number of stations that present a significant (at 95% confidence level) (in grey bars), stable (light blue and orange bars) and strongly stable (dark blue and dark red bars) correlation between teleconnection indices and lagging seasonal streamflow for autumn (OND) streamflow (upper panel); winter (JFM) streamflow (middle panel) and spring (AMJ) streamflow (lower panel) are presented; son1, djf1, mam1 and jja1 correspond to autumn, winter, spring and summer of the previous year, respectively. Note that negative numbers mean sum of stations with negative correlations.

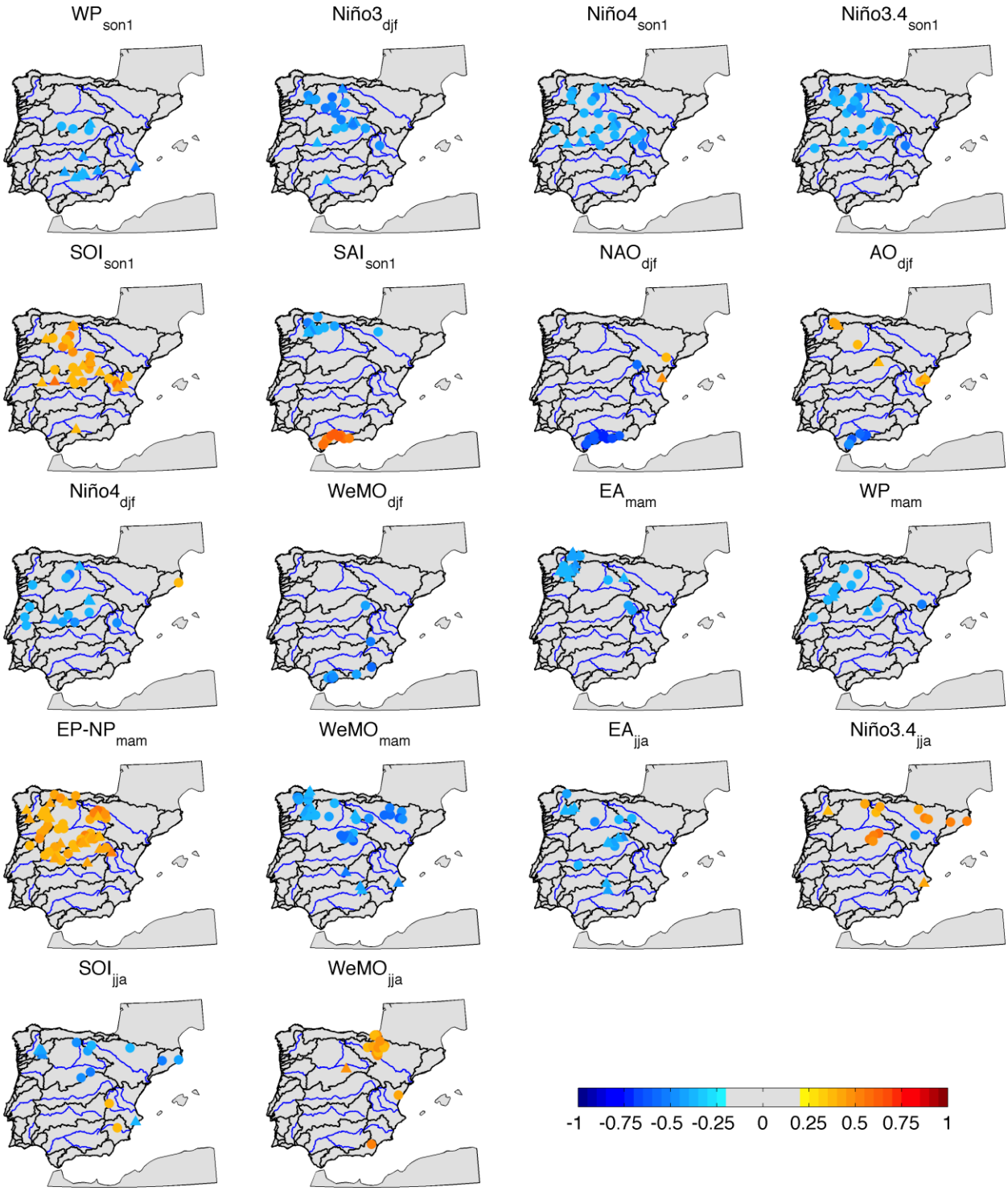


Figure 5.4. Maps of the correlation between the teleconnection indices selected a predictor and autumn (OND) streamflow. In the title of each map, the acronym of the corresponding teleconnection index and its season (subindex) is shown. Only significant and stable correlations are displayed. Circles and triangles mean gauging stations located mainstreams of rivers and reservoirs respectively.

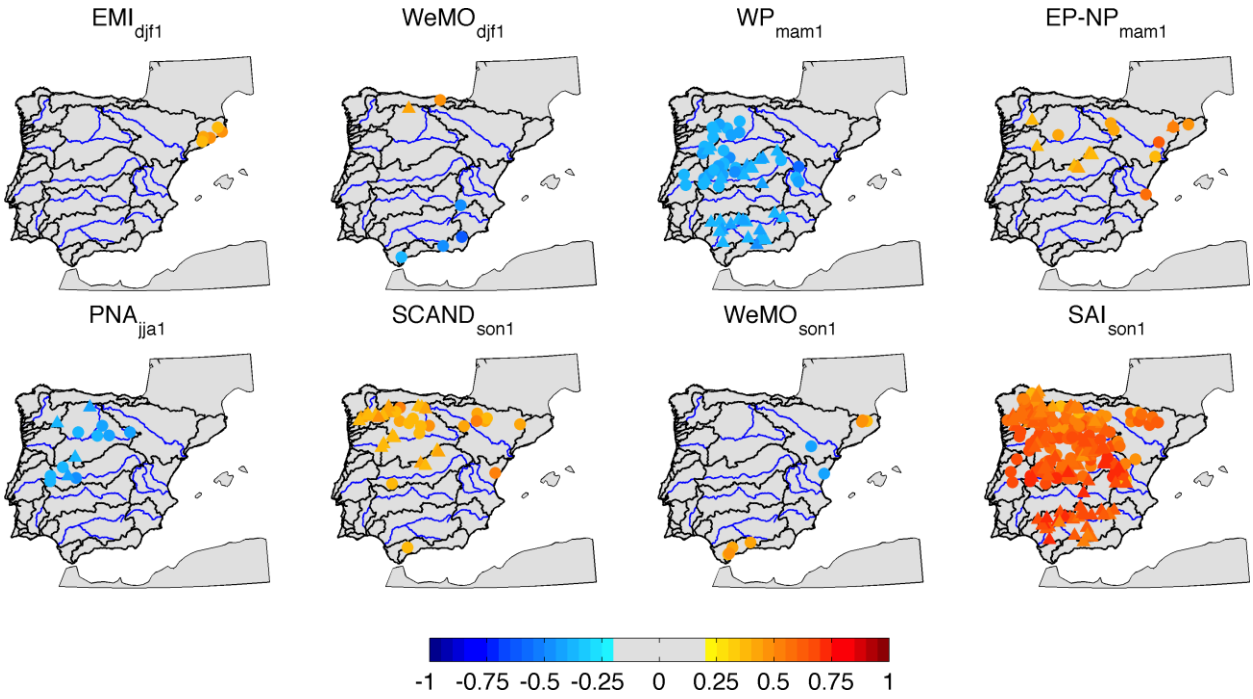


Figure 5.5. As Figure 5.4 but for winter (JFM) streamflow.

In case of spring streamflow (Figure 5.6), while no index in the previous spring is identified as a predictor, up to 9 teleconnection indices of the previous summer do. In particular, summer EA correlates stably with stations in the Douro Basin. The summer EP-NP correlates with the stations in the eastern IP, the summer SCAND with stations in the Guadalquivir Basin, and summer PNA appears as a potential predictor of spring streamflow for stations located in the Miño-Sil, Douro, and Tagus Basins. ENSO-related indices in summer (Niño1+2, Niño3, Niño3.4 and IOD) present stable correlations with spring streamflow in stations located in the Miño-Sil, Cantabrian, and Douro Basins as well as some in the upper Ebro Basin. The presence of ENSO-related indices as a stable predictor extend through the autumn (Niño1+2 and IOD) and winter (Niño1+2). Also previous October SAI and winter NAO/AO correlate stably with stations in the Douro, Tagus, and Guadalquivir Basins, especially in case of NAO and AO, and winter SCAND correlate stably with stations in the Miño-Sil, Douro, and Tagus Basins. The magnitude of these correlations is, in general, relatively weak (around 0.4 and 0.45), the previous winter NAO being the index presenting the highest correlations (-0.48 in average).

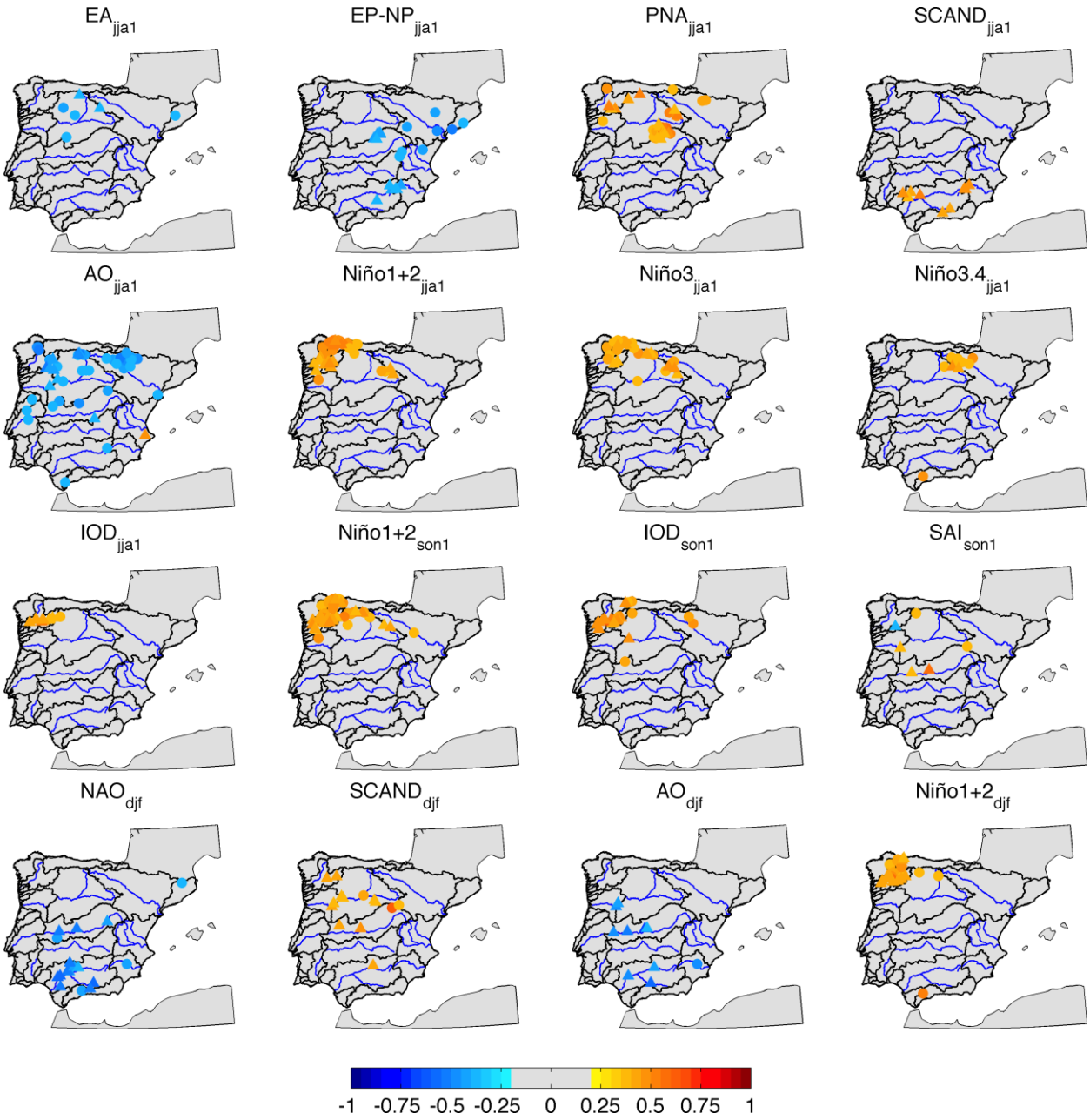


Figure 5.6. As Figure 5.4 but for spring (AMJ) streamflow.

## 5.5 Discussion

In this chapter, it has been identified stable correlations between seasonal teleconnection indices and lagging seasonal streamflow in the IP. Based on these stable correlations, teleconnection indices leading seasonal streamflow have been evaluated as potential predictors. However, understanding the physical mechanisms responsible for these relationships requires additional analysis. The connection between teleconnection indices leading seasonal streamflow in some months or seasons in the IP might be explained through changes in slowly varying

boundary conditions, such as SST or ice-sea extension in the Arctic or the snow cover in high-latitude land masses. Moreover, the individual features of each basin play an important role in the lag response to the teleconnection indices, in particular on shorter time scales (one season) responses. To evaluate this, a hydrological model would be needed for each individual basin. With the large number of teleconnection indices and seasonal streamflow analyzed, a detailed description of the physical basis of these relationships becomes a substantial amount of work and is beyond the scope of this study. Nonetheless, based on the literature, we postulate on which of the physical links might explain the most noteworthy relationships.

One of the most remarkable results is the acceptable winter streamflow prediction in a high number of gauging stations found when the SAI of the previous October is used as the predictor. These results agree with those of Brands et al. (2012, 2013), who found reasonable forecasting skill for winter precipitation over the central and western IP when using the SAI as the predictor. The role that Eurasian snow cover plays in the winter Northern Hemisphere climate has attracted greater attention in the last two decades (Cohen et al., 2001; Cohen and Saito., 2003; Gong et al., 2003; Cohen and Flecher, 2007; Cohen et al., 2007, among others). The links between observed snow cover and Northern Hemisphere wintertime circulation were also reproduced by model simulations (Orsolini and Kvamstø, 2009). Cohen and Entekhabi (1999) showed that the Eurasian snow cover in early fall is significantly correlated with the following AO. These authors hypothesize that a possible dynamical mechanism linking Eurasian snow cover anomalies and North Atlantic climate variability involves the strength and position of the Siberian high. The conceptual model, explained in Cohen (2011), proposes that the presence of snow cover can increase the amount of sunlight reflected back into space from 20 to 80% and also avoid the heat escaping from the earth's surface, leading to colder air temperatures. This cold and dense air promotes a stronger and more expansive Siberian high. As a consequence, the amount of energy transfer from the troposphere (or lower stratosphere) to the stratosphere (or upper troposphere) can be increased. Often, when such an excess energy exists, it is absorbed in the polar stratosphere; warming the polar stratosphere and weakening of the polar vortex. In the case of weakening of polar vortex, the flow of air is weaker and turns north-south (rather than west-east), encouraging the cold Arctic air to penetrate the mid-latitudes while the warm air from the subtropics is carried into the Arctic. The mixture of these air masses foment more storms and snow in the mid-latitudes. During a weak polar vortex, high pressure occurs in the Arctic region (negative phase of AO). By contrast, below-normal snow-cover anomalies during fall are related to a positive phase of AO in the following winter. As winter

NAO (regional manifestation of AO) is a major driver of winter precipitation in the western IP, the ability of SAI to predict the following winter AO is reflected in high correlation values with winter streamflow in the IP. However, the stationary relationship of Siberian snow-cover extent in the fall and following winter Arctic Oscillation has recently been questioned (Peings et al., 2013), indicating that it could be modulated by the Quasi-Biennal Oscillation in the equatorial stratosphere. Thus, the SAI-AO relationship should be taken in careful consideration and examined when more data are available, especially if it is to be used for forecasting purposes.

In addition, it worth to be noted the strong correlation (above 0.6 in absolute values) found between October SAI and winter NAO and AO with following autumn streamflow in the Mediterranean Andalusian Basin (Figure 5.4). As it was previously remarked, the SAI is considered a useful index for predicting winter conditions in the North Atlantic, in particular winter AO and NAO, which are major drivers of winter precipitation in the western IP. However, the relationship between NAO and precipitation in the Mediterranean slope of the IP is weak. We examined the correlation between winter (December to February) NAO and winter (January to March) streamflow in the Mediterranean Andalusian Basin, which is not significant in most of stations (not shown). A possible link between winter NAO/AO and October SAI with following OND streamflow in the Mediterranean Andalusian Basin could be through later spring and summer snowmelt contribution. This basin is located between the Mediterranean Sea and the Betics System, a high-mountain ridge (with the highest peaks of the IP, i.e. over 3000 m high), which accumulate large amounts of snow during winter (it acts as a natural barrier for fronts coming from Atlantic Ocean in winter). Thus, the snowmelt during late spring and summer months could contribute to an increase in runoff during the summer months, and this effect could remain until the next autumn. To corroborate this hypothesis, it was evaluated the correlation between summer (June to August) and autumn (October to December) streamflow in this basin. As a result, only 3 of the 15 stations presented a significant correlation, so this hypothesis can be discarded. Consequently, the explanation for the connection between these indices and autumn streamflow in the Mediterranean Andalusian Basin needs additional analyses. A possible physical explanation of this relationship might be in the line of the results found by Báez et al. (2013). These authors postulated a relationship between winter NAO and AO conditions and following summer SST in the Alboran Sea (a part of the Mediterranean Sea close to the Mediterranean Andalusian Basin), through the accumulated snow in Sierra Nevada, which increases the input of continental freshwater in the Alboran Sea during summer, maintaining the mixed layer at a greater depth and increasing the SST. With this hypothesis borne in mind,

warmer than average SSTs in the Mediterranean could lead to enhanced local evaporation, and hence to increased lower-tropospheric humidity. This additional moisture is likely to be advected from the western Mediterranean to the southern IP during late summer and autumn and, together with the rugged terrain, promote greater rainfall over this area. Note that an atmospheric configuration at sea level providing advection from the Mediterranean together with a cut-off low in the upper levels, have been associated with intense rainfall events over southeastern Spain during autumn (Hidalgo-Muñoz et al., 2011).

Another notable result concerns the stable correlations (negative) identified between summer Arctic Oscillation and seasonal streamflow in following spring (Figure 6), particularly in western or northwestern IP. One possible physical mechanism responsible for this lag relationship could be the associated with summer Arctic Sea ice extension. Matsumura et al. (2014) evaluated the influence of spring Eurasian snow cover on summer Arctic atmospheric circulation and the possible linkage with the increased Arctic sea-ice loss. Specifically, they indicate that the reduced albedo due to earlier spring snowmelt over Eurasia leads to a warmer land surface, which amplifies stationary Rossby waves, decelerating the subpolar jet. As a consequence of the rising motion enhanced over land, compensating subsidence and adiabatic heating occur in the Arctic troposphere, forming a negative Northern Annular Mode (NAM). Note that AO and NAM resemble similar height anomalies pattern. The recent trends towards a negative phase of NAM during summer have been linked to the accelerating trend of sea-ice extension retreat in the Arctic during summer (Ogi and Yamazaki, 2010). In addition, the effects of Arctic sea-ice reduction in summer in the mid-latitudes weather in the following winter have received special attention in recent years (two noteworthy reviews have been recently published by Cohen et al. (2014) and Vihma (2014). In Cohen et al. (2014), the authors indicate three possible dynamical pathways through which Arctic amplification may influence mid-latitude weather: changes in storm tracks in the North Atlantic sector, in the characteristic jet stream and in the planetary wave configurations. Some of the studies reviewed in Vihma (2014), such as the one made by Francis et al. (2009) and Overland and Wang (2010), suggest that summer sea ice declining in the Arctic causes a deepening of the atmospheric boundary layer, a warming and destabilization of the lower troposphere, and a decreasing of the poleward gradient in the thickness of the atmospheric layer between 500 and 1000 hPa pressure levels, which weakens the polar jet stream. This configuration is similar to a negative phase of Arctic Oscillation. In addition, Outten and Esau (2012), on the basis of ERA-Interim reanalysis, identified cooling trends of wintertime near-surface air temperature along a band that extends across mid-latitude

Eurasia. By their interpretation, the weakened north–south temperature gradient has decreased wintertime westerly winds over mid-latitude Eurasia, resulting in lower temperatures. The cooling in northern Eurasia resulting from less sea ice extent in the Arctic in September is also supported by Rinke et al. (2013) and Peings and Magnusdottir (2013) using model experiments. With these findings in mind, it could be postulated that this cooling in winter combined with more precipitation could result in an increased snowfall during winter. Consequently, an increase in spring streamflow could be expected by snowmelt contribution. A correlation analysis was conducted between summer AO and winter and spring temperature (using CRUTS3.0 temperature dataset, with 0.5º degree of spatial resolution) and precipitation (using GPCC precipitation dataset, with 1º degree of spatial resolution) over Europe. It was found significant positive correlation (not shown) with winter temperature over Finland and western Russia. Also some areas in the IP present positive but no significant correlations. In case of spring temperatures (Figure 5.7a), the above positively correlated area in western Russia increases in extension, and also appears a positive significant correlation in the western half of the IP but a negative correlation with precipitation (Figure 5.7b). This positive/negative correlation with spring temperature/precipitation over the IP means that negative AO anomalies (or NAM) in summer is linked with below-normal temperature in the western IP in spring (late winter and early spring) and with positive precipitation anomalies and perhaps increased snowfall in mountain ranges because of the lower temperatures. This could explain an above-normal streamflow during spring as a consequence of both higher precipitation and snowmelt contribution. However, uncertainties remain concerning these dynamical pathways (Cohen et al., 2014). For example, although observational analysis has shown links between summer sea-ice loss in the Arctic and winter NAO/AO, modeling simulations do not. Also, there are other factors, besides the near-surface meridional temperature gradient, influencing the zonal jet that should be considered. In addition, more statistically robust results regarding the influence of alteration in planetary waves over Eurasia are needed.

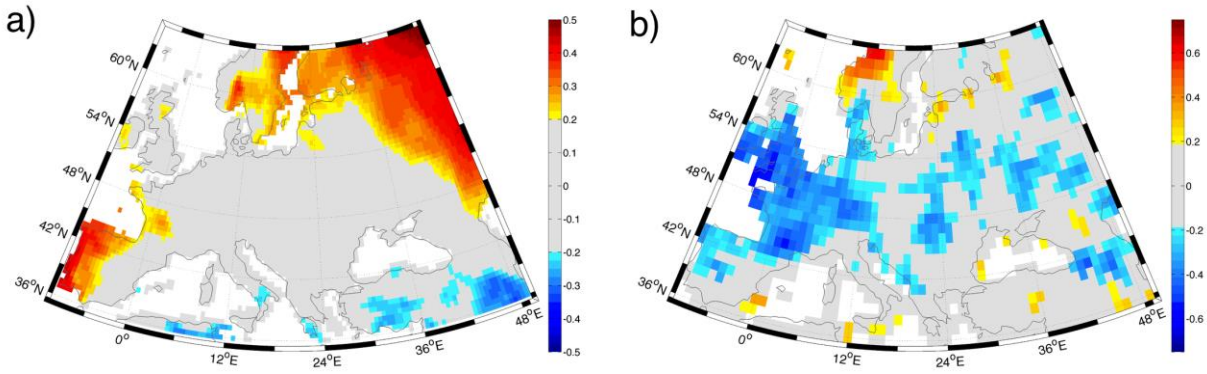


Figure 5.7. Correlation of summer AO with spring (March to May) a) temperature and b) precipitation. Temperature data at 0.5º of spatial resolution were obtained from CRU TS v.3.1 dataset (Harris et al., 2014). Precipitation data at 1º of spatial resolution where obtained from: GPCC Precipitation (Schneider et al., 2014), NAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd/).

The correlation found between winter NAO and SCAND with following spring streamflow might be associated with the lagging response in each basin and the contribution of snowmelt that accumulated during winter in high-mountain systems, as NAO and SCAND are major teleconnection indices driving winter precipitation in the western IP.

The WeMO index has also been identified as a stable predictor for seasonal streamflow (with different lags). The WeMO index has been associated with extreme precipitation events over the Mediterranean façade of the IP (Martin-Vide et al., 2008; Hidalgo-Muñoz et al., 2011), particularly during autumn and early winter. The correlation between WeMO and streamflow in autumn was calculated (not shown), finding stable positive correlations in the Cantabrian range and negative in the Mediterranean slope. The positive phase of WeMO resembles a low/high pressure centre in Padua/Gibraltar. This configuration favors the entrance of humid winds from the Atlantic Ocean/Cantabrian Sea that bring about precipitation on the northern Slope of the Cantabrian ranges. Correspondingly, positive correlations with the Cantabrian Basin were found when September to November WeMO was correlated with October to December streamflow (not shown). Similarly, Figure 5.4 displays a positive correlation in the Upper Ebro (near the Cantabrian coast) between summer WeMO and following autumn streamflow. Thus, it seems that the origin of this teleconnection is likely due to a lagging response of streamflow. However, the teleconnection between previous winter and spring WeMO with autumn streamflow is not well understood as a lagging response to the coetaneous relationship between WeMO and

precipitation in the IP. This sea-level pressure configuration related to WeMO phases could be a regional reflection of a larger-scale phenomenon. Nevertheless, further analyses are needed.

In relation with the lagging influence of teleconnection indices located in the Pacific, ENSO indices have been also found stably correlated with autumn (Figure 5.4) and spring (Figure 5.6) streamflow, especially with stations located in the Northwestern quadrant of the IP. Numerous studies have demonstrated statistical associations between ENSO and precipitation in the Mediterranean Basin. In case of the IP, precipitation has been shown to increase during late summer, autumn, and early winter when El Niño conditions are present in the Pacific (Rodó et al., 1997; Mariotti et al., 2002; Park, 2004; Pozo-Vázquez et al. 2005; Vicente-Serrano, 2005; Sordo et al., 2008). In agreement with this, positive correlations were found between summer ENSO indices (which represent the beginning of the phenomenon) and following OND streamflow (Figure 5.4). Also, it was conducted a correlation analysis between September to November ENSO indices and October to December streamflow, finding an increase in the number of stations with positive correlations. Hence, the link between autumn and early winter ENSO with autumn and early winter streamflow in the IP begins in summer, with the beginning of the phenomenon. Also notable was the contrary sign observed between previous autumn and winter ENSO indices with autumn streamflow of the following year with respect to the correlation between summer and autumn ENSO indices and autumn streamflow of the same year. This change in the sign of the correlation, also observed by Córdoba-Machado et al. (2014) in a study of the relationship between Pacific SST and precipitation in Colombia, could be related to a 2-year periodicity observed in El Niño (Yu and Kim, 2010). Rodó et al. (1997) and Knippertz et al. (2003) found an association between boreal winter ENSO conditions and spring precipitation over Spain and southwestern Europe, respectively. Furthermore, Mariotti et al. (2002) found that precipitation in the Mediterranean area in late winter and spring is suppressed during an El Niño event. However, Lorenzo et al. (2010) concluded that the negative phase of ENSO, "La Niña", almost always announces dry springs in northwestern IP, whereas the positive phase of ENSO, "El Niño", does not anticipate the appearance of wet springs. In the line of these results (negative ENSO associated with negative precipitation), we found stable positive correlations in spring streamflow the northwestern basins (Figure 5.6) with ENSO related indices (such as Niño1+2, Niño3, Niño3.4, IOD, PNA) in summer and autumn. The influence of tropical pacific SST variability on North Atlantic sector has been widely studied (Giannini et al., 2001; Shaman et al., 2009; Shaman and Tziperman, 2010; Shaman, 2014; among others). For instance, in Giannini et al. (2001), a physical explanation of the teleconnection between the SST

anomalies in tropical Pacific related to an ENSO events and spring precipitation in north Atlantic sector is provided. Accordingly, in a warm ENSO event SLP is lower than average over the warm waters of the central and eastern equatorial Pacific, and higher than average over the equatorial Atlantic. Related to this anomalous seesaw in SLP is a weakening of the meridional SLP gradient in both hemispheres of the Atlantic basin, consistent with reduced trade winds at subtropical latitudes. During boreal winter, the weakening of the meridional SLP gradient is reinforced at sub-tropical latitudes in the Northern Hemisphere by a low-pressure center over the southeastern United States and southwestern North Atlantic, related to the Pacific–North American (PNA) wave train (Horel and Wallace 1981; Wallace and Gutzler 1981). The end result is the oceanic component of the teleconnection: a warming of the tropical North Atlantic Ocean surface that peaks in boreal spring, that is, with a delay of about one season with respect to the attainment in winter of maximum SST anomalies in the tropical Pacific. More recently, Shaman and Tziperman (2010) observed that ENSO modulates southwestern European precipitation during the September–December wet season. They indicated that the precipitation anomalies are associated with changes in large-scale atmospheric fields to the west of Iberia that alter low-level westerly winds and onshore moisture advection from the Atlantic. The interannual variability of fall and early winter precipitation over southwestern Europe is linked to ENSO variability in the eastern Pacific via an eastward-propagating atmospheric stationary barotropic Rossby wave train.

Other teleconnection patterns related to atmospheric variability in the Pacific have also been identified as stable predictors, such as summer PNA (for following winter and spring streamflow), spring EP-NP (for following autumn) and spring WP (for following autumn and winter). Studies linking these teleconnection patterns with precipitation or streamflow in southwestern Europe are still scarce. One explanation relating these atmospheric patterns with following autumn to spring is that they reflect changes in tropical Pacific SST or the extent of sea ice in the Arctic. For example, in the discussion in Pozo-Vázquez et al. (2005), the PNA is proposed as one of the links between tropical forcing and extratropical circulation response in the North Atlantic area, triggering a standing wave train that propagates downstream to the North Atlantic area. In fact, Lau and Nath (2001) showed that the influence of ENSO is not limited to the tropical Atlantic but is also significant in the extratropical North Atlantic, which supports the idea that the PNA mechanism plays an important role in the Atlantic response to ENSO (Alexander et al., 2002). In the case of WP, recent works, such as in Linkin and Nigam (2008), have related the WP with the sea-ice extent in marginal zones in the Arctic, establishing

relationships with PNA and ENSO in the Pacific. However, how this pattern influences streamflow in late autumn and winter is still a matter of study.

Concluding, this chapter provides a useful evaluation of the ability of main teleconnection indices as potential predictors of seasonal streamflow on the IP, which becomes a helpful tool for forecasting seasonal streamflow. This will be the main goal of Chapter 7. However, there exist some limitations regarding the use of teleconnection indices as potential predictors. For example, the predefined character of teleconnection indices (calculated as the average of a climatic variable over an area or the difference of a climatic variable between two locations), which could be overcome by exploring the relationship between climate variables (such as sea surface temperature or atmospheric pressure) with seasonal streamflow. Chapter 6 focus on this, on studying the possible influence of climate variability on remote areas (beyond the defined teleconnection indices) on seasonal streamflow of IP Rivers.

# CHAPTER 6

# COVARIABILITY ANALYSIS

In this chapter, the singular value decomposition technique is used to identify climate-hydrology response relationships using gridded global climatic data such as sea surface temperature, geopotential height, temperature and precipitation over land, and lagging seasonal streamflow. Additionally, the significance of these relationships is evaluated and the main physical bases supporting them are discussed. Finally, the ability of these main coupled modes as stable predictors of seasonal streamflow in the IP is addressed.

## 6.1 Introduction

In the previous chapter, the ability of main teleconnection indices (such as ENSO, SAI or NAO) as potential predictors of seasonal streamflow of IP Rivers was addressed. Although each of these oceanic/atmospheric phenomena represent part of the ocean or atmospheric variability, they are referred to a specific and spatially predetermined region (e.g., tropical Pacific Ocean, northern Pacific Ocean, northern Atlantic Ocean, a particular mode of atmospheric variability or the pressure differences between two particular areas), which may not take into account additional information of climatic system that could have an influence in seasonal streamflow on the IP. In addition, the non-stationarity of the teleconnection patterns (Trigo et al., 2004; De Castro et al., 2006; Vicente-Serrano and López-Moreno, 2008) compels us to seek new variables

that can provide information regarding the variability of climate system. These variables could have an influence on streamflow fluctuations and, therefore, be used as climatic predictors. Variables that could be used as predictors for streamflow forecasting are air temperature, humidity, wind or precipitation. For example, precipitation is related to the impact of a large-scale climatic phenomenon, and could provide a clearer signal of this phenomenon for forecasting purposes, temperature and humidity are very much related to the amount of moisture in the air, wind, which is a determinant factor for moisture transport in the atmosphere, or geopotential height at 500 hPa, which is a direct indicator of the atmospheric conditions that originates precipitation or its absence. Bearing in mind that the oceans are the largest resources of water moisture of the earth, the ocean dynamics should be considered since it plays a significant role in streamflow variability. For example, the passage of a cold front over a cold ocean area induces the stabilization of the air column, diminishing associated rainfall, while a positive SST anomaly increases rainfall, via enhanced evaporation and the decrease of the vertical stability. Hence, SST can provide important predictive information about hydrologic variability in regions around the world. Therefore, the identification of coupling modes between these climatic variables and the seasonal streamflow variability could be a useful tool to improve long lead-time streamflow forecasting. On this regard, Gámiz-Fortis et al. (2010, 2011) and Hernández-Martínez et al. (2014) evaluated the influence of SST regions on forecasting IP Rivers discharge, Rimbu et al. (2005) used SST as predictor of seasonal discharge in Danube River or Ionita et al. (2008), which predicted spring discharge in Elbe River using Global SST, temperature and precipitation as predictors.

In this Chapter, the possible links between the seasonal streamflow in the IP Rivers and climatic variables such as SST, geopotential height or land surface temperature and precipitation data for entire regions (Pacific and Atlantic Oceans in case of SST, the Northern Hemisphere in case of geopotential height or global in case of temperature and precipitation) are going to be evaluated. This could result in new oceanic, atmospheric or surface regions being identified as having coupled impacts.

Various methods are available to determine coupled relationships between two spatial-temporal fields, such as SST and climatic variables. For example, Ionita et al. (2008) used linear correlation between the gauging station representing Elbe River and global SST, temperature and precipitation. However, when multiple stations are simultaneously evaluated, this strategy is not practical, and other methods are required to find possible links between climate fields and streamflow variability over the IP. Bretherton et al. (1992) compared four methods for isolating

the important coupled modes of variability between two fields. The methods compared were principal component analysis with combined fields (CPCA), canonical correlation analysis (CCA), principal component analysis of one field followed by correlation of its component amplitude with the second field (SFPCA) and singular value decomposition of the covariance matrix between two fields (SVD). In general, CCA and SVD present advantages such as their lack of systematic bias and good general performance. Also they produce explicit measures of relatedness between the derived coupled patters (unlike CPCA). In addition, their results indicated that SVD was simpler to apply than CCA and preferable for general use since it does not require user-chosen variables. Wallace et al. (1992) also used SVD to study the inter-annual relationship between the Pacific Ocean SST anomalies and 500mb atmospheric pressure in wintertime and found that SVD isolates the most important modes of covariability. Hence, the SVD is a powerful method used to identify pairs of coupled spatial patterns and their temporal variations. Each pair explains a fraction of the covariance between the two jointly analysed fields. This decomposition allows the extraction of dominant modes of coupled variability between the two analysed fields. The SVD (also referred in literature as Maximum Covariance Analysis, MCA) have been widely used to identify air-sea interactions. For example, to establish the relation between Atlantic SST and the atmospheric circulation in the north Atlantic European sector (Czaja and Frankignoul, 1999, 2002; Drévillon et al., 2001; Frankignoul et al., 2003; Frankignoul and Kastenare, 2005; García-Serrano et al., 2008; Gastineau et al., 2013; Robertson et al., 2000; Rodwell and Folland, 2002; Rodríguez-Fonseca et al., 2006) and also in other regions, such as the Indo-Pacific SST relationships (Sewell and Landman, 2001), or the impact of North Pacific SST on the atmosphere (Liu et al., 2006). Additionally, and in the line with the aim of this chapter, the coupling between atmospheric or oceanic fields and hydrologic variability has also been analysed by using SVD technique for several authors. For example, considering the following hydrological variables:

**Drought:** Rajagopalan et al. (2000) utilized SVD and applied a lagging approach to evaluate global SST impacts on continental U.S. drought (e.g., Palmer Drought Severity Index, PDSI). Shabbar and Skinner (2004) utilized a lagging approach in which the link between winter global SST and summer Canadian drought (PDSI) was evaluated. Feng et al. (2011) addressed the influence of Atlantic SST on the persistent droughts in North America.

**Precipitation:** Ting and Wang (1997) established a relationship between summertime US precipitation and the Pacific SST. Uvo et al. (1998) applied SVD to evaluate Pacific and Atlantic Ocean SSTs and northeast Brazilian precipitation. The Pacific and Atlantic Oceans were

evaluated independently using both a simultaneous and lagging approach. Wang and Ting (1999) analysed the covariability of winter US precipitation and Pacific SST. Wang and Ting (2000) evaluated Pacific Ocean SST and winter U.S. precipitation for concurrent (overlapping) time periods. They found that the precipitation patterns associated with North Pacific SST anomalies account for more U.S. precipitation variability than those associated with the tropical SST variations, except when only El Niño and La Niña years are considered. Rodríguez-Fonseca and de Castro (2002) used SVD to study the relationship between monthly winter precipitation anomalies in Europe and North Africa with North Atlantic Ocean SST and sea level pressure (SLP). The authors argued that the lagging connection between the different periods could make possible to use the SST and SLP for long lead-times forecasting of the precipitation. Liu (2003) used SVD to improve predictions of monthly seasonal precipitation over East Asia by coupling with previous soil moisture. Joly et al. (2007) explored the teleconnections between tropical SSTs and African monsoon at interannual to multidecadal time scales via a SVD analysis applied to low and high frequency signals of the dataset. Polo et al. (2008) presented a description of the coupling between tropical Atlantic SST variability modes and the West African rainfall during the monsoon season. Peings and Douville (2010) reviewed the possible influence of winter/spring Eurasian snow cover on Indian summer monsoon. Taschetto and England (2009) used SVD to assess the impacts of El Niño Modoki pattern on Australian rainfall. Omondi et al. (2013) investigated the linkages between global SST and decadal rainfall variability over Eastern Africa region. Wei et al. (2012) explored the relationships between SLP over east Asia and summer precipitation in eastern China during the period 1850-2008 using SVD. Finally, Córdoba-Machado et al. (2014) studied the forecast skill provided by the tropical Pacific SST associated with El Niño and El Niño Modoki over seasonal precipitation in Colombia through a lagging SVD.

**Streamflow:** Grantz et al. (2005) and Soukup et al. (2009) identified 500 hPa geopotential height (Z500) values as a useful long lead-time predictor of streamflow in Truckee-Carson River System and North Platte River, in United States, respectively. Tootle and Piechota (2006) evaluated the relationship between Atlantic and Pacific Ocean SST with U.S. streamflow variability, considering cold and warm phases in both oceans (through positive or negative phases of PDO and AMO). Tootle et al. (2008) evaluated the relationships between the Pacific and Atlantic Ocean SSTs and Colombian streamflow variability applying SVD in a lead-time approach. They found that the use of SSTs for entire regions eliminates any spatial bias as to which oceanic SST region impact hydrology. Anderson et al. (2012) added the contribution of

Pacific Ocean climate variability (via SVD) to tree rings chronologies in order to improve hydrologic reconstructions in the Upper Green River Basin (located in North America). Oubeidillah et al. (2011) studied the use of oceanic and atmospheric variables as predictors for long lead-time forecasting of the streamflow and snowpack in the Upper Colorado River and Great Basin (U.S.). Lamb et al. (2010) identified a Pacific Ocean region when evaluating the streamflow in Colorado River Basin (U.S.). Oubeidillah et al. (2012) studied the relationship between the Atlantic SST and the streamflow in Adour-Garonne Basin (France).

The goal of this work is to improve the understanding of the climate drivers for the seasonal streamflow variability in the IP in order to develop an improved long-lead forecasting technique. For this purpose the lagging SVD approach was used for the combined fields of streamflow and the potential predictor variables SST, Z500, temperature and precipitation. The significance and the predictability skill of the covariability coupled modes obtained by SVD were also evaluated.

## 6.2 Datasets

The large-scale atmospheric and oceanic variables, covering from September of 1974 to February of 2008, were obtained from the following datasets:

- The monthly data for the sea surface temperature (SST) were obtained from The Hadley Centre Global Sea Ice and Sea Surface Temperature (HadISST, Rayner et al., 2003), with a resolution of 1º x 1º. The region of Pacific Ocean SST data used for the analysis is [120°E-70°W, 40°S-70°N], while the region of Atlantic Ocean SST data used is [80°W-30°E, 40°S-70°N].
- The monthly mean geopotential height at 500 hPa (Z500) was downloaded from the NCEP Reanalysis (Kalnay et al., 1996), with a resolution of 2.5º x 2.5º. Only Z500 data northern than 20ºN are used here.
- The monthly values for global precipitation (RR) were downloaded from the Global Precipitation Climatology Centre (GPCC) version 6- Total Full (Scheneider et al., 2014). The dataset has a 1º x 1º resolution.
- The monthly global surface temperature data over land (TMP) were taken from the high-resolution database of the Climatic Research Unit (CTU-TS.3.10, Harris et al., 2014), with a 0.5º x 0.5º of spatial resolution.

For computational efficiency, all dataset were regridded into a coarse grid of 2.5º x 2.5º. Seasonal time series were created by averaging the months: September-October-November (son) for autumn, December-January-February (djf) for winter, March-April-May (mam) for spring and June-July-August (jja) for summer.

## 6.3 Methodology

### 6.3.1 Coupling patterns. Singular Value Decomposition (SVD)

In order to perform the SVD technique, the temporal cross-covariance matrix between the two space-time distributed data fields needs to be computed. Let's $X$ and $Y$ be the left and right data fields representing the climatic variable (SST, Z500, TMP or RR) and seasonal streamflow respectively. The dimension of $X$ is $N$ x $M_x$ and the dimension of $Y$ is $N$ x $M_y$, where $N$ represents times (years), $M_x$ the number of grid points and $M_y$ the number of gauging stations. The temporal cross-covariance matrix can be constructed as:

$$C = X^T Y = \begin{pmatrix} X_1 Y_1 & \cdots & X_1 Y_{M_y} \\ \vdots & \ddots & \vdots \\ X_{M_x} Y_1 & \cdots & X_{M_x} Y_{M_y} \end{pmatrix} \tag{6.1}$$

which has $M_x$ x $M_y$ dimension, with each element of the matrix, $\langle X_i Y_j \rangle$, corresponding to the spatial cross-covariance between the time series $X_i$ and $Y_j$ at grid $i$ and station $j$, respectively. Based on the cross-covariance matrix, $C$, matrices $U$, $V$ and $L$ can be computed such that:

$$C = ULV^T \tag{6.2}$$

The singular vectors for $X$ are the columns of $U$ (often-called left patterns), and the singular vectors of $Y$ are the columns of $V$ (right patterns). Each pair of singular vectors is a mode of covariability between the fields $X$ and $Y$.

These vectors, a set of $N_x$ dimensional orthogonal vectors $U_k$ ($k$=1,.., $N_x$) for $X$, and a set of $N_y$ dimensional orthogonal vectors $V_q$ ($q$=1,…, $N_y$) for $Y$, are determined so the covariance between the projections of the fields on them is maximized, subjected to orthogonalty:

$$UU^T = I \tag{6.3}$$

$$VV^T = I \tag{6.4}$$

The matrix $L$ is a diagonal matrix whose dimension is of $M_x$ x $M_y$, with its elements being zero except the first $R$ (with $R \leq \min(M_x, N_y)$ ) diagonal elements, which are referred as singular values. The expansion coefficients, which describe the time variability of each mode, can be obtained by projecting each field onto their respective singular vectors:

$$A = XU \tag{6.5}$$

$$B = YV \tag{6.6}$$

The $k^{th}$ columns of the $A$ and $B$ matrices contain the expansion coefficients for the $k^{th}$ SVD mode for fields $X$ and $Y$, respectively. Since both $U$ and $V$ are orthogonal, the original fields can be easily reconstructed as:

$$X = AU^T \tag{6.7}$$

$$Y = BV^T \tag{6.8}$$

Hence, using equations (6.1), (6.2), (6.7) and (6.8), the temporal covariance between the two fields can be expressed in terms of the expansion coefficients as follows:

$$A^T B = L \tag{6.9}$$

That is, an expansion coefficient of the left field is correlated only with the time series of the same mode in the right field.

The relative importance of each SVD mode is indicated by the percentage of square covariance (*SC*) for the associated mode. If $l_k = L(k,k)$ is the $i^{th}$ singular value, the fraction of the squared covariance (*SCF*) explained by this mode is:

$$SCF_k = \frac{l_k^2}{\sum_{i=1}^{R} l_i^2} \tag{6.10}$$

The strength of the coupling is represented by the correlation between the expansion coefficients corresponding to the left and right fields ($A$ and $B$, respectively). For example, in case of the $k^{th}$ mode:

$$STR_k = r_{A_k B_k} \tag{6.11}$$

In order to present the spatial patterns corresponding to the SVD modes, the homogeneous and heterogeneous maps are used. The $k^{th}$ homogeneous correlation map is

defined as the vector of correlation values between the expansion coefficient of the $k^{th}$ mode of a field and the values of the same field at each point. It is a useful indicator of the spatial localization of the covarying part between the field and its $k^{th}$ mode. The $k^{th}$ heterogeneous map is defined as the vector of correlation values between the expansion coefficient of the $k^{th}$ mode of a field and the values of the other field. It indicates how well the grid point values of the second field can be predicted from the knowledge of the expansion coefficient of the first field.

Particularly, in this study, the homogenous maps refer to the correlations between the left expansion coefficients and the values of the same field (in this case, SST, Z500, TMP or RR), while the heterogeneous maps refer to the correlations between the left expansion coefficients (associated to SST, Z500, TMP or RR) and the right field (streamflow).

Significance of SVD coupled patterns

Although SVD is a useful technique in the statistical analysis of two spatio-temporal fields, there are some conditions that must be satisfied for SVD to be used correctly (Newman and Sardeshmukh, 1995). These authors indicate that the relationship between the two fields (*X* and *Y*) in the underlying dynamical system is unclear. The assumption always made is that the relationship between the two data fields being analysed is fundamentally a linear one. The success of SVD in recovering the relationship between two variables depends not only on the relationship but also on the covariance structure of the data (e.g., two analysis separated with a given set of variables, such as precipitation and SST, could be relatively successful one time but unsuccessful the next). SVD is able to isolate the most frequently occurring pairs of patterns of *X* and *Y*. It could be thought that a SVD pair, say $X_1$ and $Y_1$, occur 'simultaneously' or 'go together' in some sense. Even this interpretation needs to be treated with caution. Three possibilities can be considered in this context:

1. The coefficient time series of $X_1$ and $Y_1$ are not highly correlated. Then, the notion of the simultaneous occurrence of $X_1$ and $Y_1$ is clearly dubious.
2. The coefficient time series are highly correlated but $X_1$ and $Y_1$ do not project strongly on the dominant EOFs of *X* and *Y*. In this case $X_1$ and $Y_1$ might occur simultaneously but they are not important because the total covariability of *X* and *Y* is small.
3. The coefficient time series are highly correlated and $X_1$ and $Y_1$ project strongly on dominant EOFs of *X* and *Y* respectively. In this case, SVD might appear useful.

Additionally, Cherry (1996) recommend be cautious when interpreting results from SVD and CCA methods, because both have a high potential to produce spurious spatial patterns, especially with small size samples.

According to that, the question of whether there is a physical connection between the two fields together with the evaluation of the predictability skill that left field could have on right field have been carefully addressed. This was handled in similar way to Czaja and Frankignoul (1999), Martín et al. (2004) or Rodwell and Folland (2002) by adopting the null hypothesis that there is no physical connection. If true, then temporally shuffling the sequences of the fields and applying the SVD analysis should not lead to a statistically different outcome. In particular, the approach used in Martín et al. (2004) was followed here. So, Monte Carlo realizations were carried out shuffling the left field (climatic variables, such as SST or Z500), keeping the same order in the right field (streamflow). From the combined action of shuffling and applying the SVD 100 times, a Monte Carlo-style test is conducted based on the value of *Square Covariance* (*SC*), calculating the percentage of *C*s from the test that exceed the true *SC*. This percentage is an estimate of the significance level (*SL*) of the strongest test that is satisfied. The smaller the *SL*, the stronger the test passed and the more confident one can be of the existence of a physical connection between the covarying patterns. Although, other authors have applied the test to other statistics such as Covariance Fraction (*CF*; Rodwell and Folland, 2002), or the Squared Covariance Fraction (*SCF*; Iwasaka and Wallace, 1995), Martín et al. (2011) suggested that whereas the *SCF* and *STR* are indirect measures of the relationship between the SVD modes, the square covariance is a direct measure of the relationship between the climatic variables and streamflow and between the coupled SVD patterns. One may find a large *SCF* and *STR* in the leading SVD modes of the two weakly related fields, even though these modes actually account for little squared covariance (Wallace et al., 1992). Therefore, the *SCF* and *STR* are only meaningful when they are associated with a significant square covariance. Additionally, in order to improve the understanding of the physical meaning of the identified SVD patterns, the left expansion coefficients (associated with the climatic variables) were correlated with the EOFs of their corresponding fields (see Appendix C), and also with the main teleconnection indices (those used in Chapter 5). For example, the left expansion coefficients of the SVD modes coupling winter Z500 and following autumn streamflow were correlated with winter EOFs of Z500 and winter teleconnection indices.

Further, the usefulness of the forecast system based on the lagging SVD technique was also estimated (following Rodwell and Folland, 2002). The procedure is similar to standard

'cross-validation' tests for other statistically based forecasting systems (Livezey, 1995). Considering the original input data, *X* and *Y*, one common year, *t*, is removed to give $X^*_t$ and $Y^*_t$. After that, the SVD analysis to these reduced time series is applied to obtain the pair of patterns $U^*_t$, $V^*_t$.

The 'cross-validated time series' values are then defined by:

$$A^*(t) = X^*_t U^*_t \tag{6.12}$$

$$B^*(t) = Y^*_t V^*_t \tag{6.13}$$

Repeating this procedure 33 times (number of years) the cross-validated time series are built up. The cross-validated time series based on the leading (i.e. earlier) field is the 'left time series'. The cross-validated time series based on the lagging (i.e. later) field is the 'right time series'. The correlation between the two cross-validated time series for the $k^{th}$ mode is:

$$COR_k = r_{A^*_k, B^*_k} \tag{6.14}$$

The patterns that emerge from a SVD analysis have an arbitrary joint sign, so that the lagging SVD analysis could produce $-(U^*_t, V^*_t)$ rather than $(U^*_t, V^*_t)$ for a particular mode. Here, the joint sign is determined by ensuring that the lagging pattern 'projects' positively onto the lagging pattern from the full 33-year analysis. 'Projection' is defined here as the scalar product of two fields.

Note that the cross-validated time series ($A^*_t$ and $B^*_t$) could be somewhat different from the non-cross validated time series *A* and *B*, and the correlation $r_{A_k, B_k}$ is found to be quite a bit larger than *COR*. This is no doubt an artefact of the covariance maximisation of the SVD technique and it is the reason for cross-validating when assessing true predictive skill. Also, negative values of *COR* imply no link detected, so no predictability skill is expected.

### 6.3.2 Selection of stable predictors

After identifying the significant modes of covariability between the climatic variables and seasonal streamflow, the next step consisted on studying their potential as stable predictors. To do that, a similar method to that used in Chapter 5 was followed, i.e. evaluating the stability of the correlation between predictors and predictands at each gauge station through a running window correlation approach (see Chapter 5, Section 5.3 for further details).

## 6.4 Results

### 6.4.1 Significance of covariability modes

The seasonal values of the atmospheric and oceanic fields were passed to anomalies by subtracting the mean, whereas the seasonal streamflow were standardized. The use of standardized anomalies in seasonal streamflow time series is because it exists a great disparity in variance between streamflow in gauging stations located in lower part of main rivers and those located in the headwaters or in small rivers. Both left and right fields were detrended to remove any trends in the data sets that may bias the analysis and mask the underlying variability. In addition, prior to construct the cross-covariance matrix, the seasonal anomalies of the left field were weighted by the square root of the cosine of latitude, which ensures that equal areas are given equal weight in the analysis.

For each seasonal streamflow (autumn, winter and spring), the SVD was performed considering the atmospheric and oceanic fields leading the seasonal streamflow for the previous four seasons (i.e., in case of autumn streamflow, the SVD was carried out with seasonal Atlantic and Pacific SST, Z500, TMP and RR corresponding to the previous autumn, winter, spring and summer). This accounts for a total 20 SVD computations (5 variables and 4 seasons) for each seasonal streamflow.

In all SVD analyses only the three first modes were evaluated, because of they account for 70% or more of the *SCF* in most cases, and the *SCF* for the fourth and subsequent modes usually drops below 10%, which means that the amount of information accounted for these modes respecting the two fields is not very relevant. Among all coupling modes determined in SVD analysis between the oceanic and atmospheric fields and the seasonal streamflow, only those modes in which *COR* is significant at 90% of confidence level (0.29) have been considered and described. The reasons to do that are, on one hand, a reduction in the number of modes (bearing in mind the five leading fields, with four lags and the three first modes in each SVD, totals 60 coupled modes for each seasonal streamflow), and, on the other hand, to focus on those modes that have real predictability skills (given by a significant *COR* value).

Modes are named with the first capital letter indicating the climatic field (where ‘P’ correspond to Pacific Ocean SST, ‘A’ to Atlantic Ocean SST, ‘Z’ to geopotential height at 500 hPa, ‘T’ to global temperature and ‘R’ to global precipitation), with the next three letters meaning the season of the climatic field (son, djf, mam, and jja) and with the last number denoting the mode order (ordered descending in *SCF* values). The modes are presented from

those found in SVD analysis considering the maximum lag (one year) until the minimum lag (previous season) between climatic variables and seasonal streamflow.

Covariability with autumn (OND) streamflow

The covariability SVD modes selected (those with *COR* value significant at 90%) for autumn (OND) streamflow are summarized in Table 6.1. This table presents the information of the *SCF*, the significance level provided by Monte Carlo test (*MC*) and also the strength of the coupling (*STR*) and the potential predictability skill (*COR*).

| **Mode** | ***SCF*** | ***MC*** | ***STR*** | ***COR*** |
|---|---|---|---|---|
| Pson1 | 75.2 | 98 | 0.51 | 0.32 |
| Ason2 | 18.0 | 87 | 0.79 | 0.49 |
| Zson2 | 28.2 | 35 | 0.60 | 0.51 |
| Tson1 | 57.1 | 93 | 0.58 | 0.30 |
| Tson3 | 8.9 | 85 | 0.72 | 0.34 |
| Rson1 | 51.8 | 98 | 0.76 | 0.35 |
| Adjf1 | 54.5 | 99 | 0.72 | 0.45 |
| Adjf2 | 14.9 | 87 | 0.80 | 0.50 |
| Zdjf2 | 18.3 | 88 | 0.72 | 0.48 |
| Rdjf1 | 48.2 | 94 | 0.73 | 0.31 |
| Rdjf3 | 9.1 | 43 | 0.92 | 0.29 |
| Pmam2 | 27.9 | 95 | 0.54 | 0.50 |
| Amam2 | 20.9 | 79 | 0.73 | 0.29 |
| Rmam2 | 17.1 | 95 | 0.81 | 0.29 |
| Ajja2 | 23.1 | 89 | 0.66 | 0.43 |
| Tjja2 | 20.7 | 94 | 0.75 | 0.34 |

Table 6.1. Summary of the square covariance fraction (*SCF*), confidence level given by Monte Carlo test (*MC*), both in percentage, the strength of the coupling (*STR*), and the predictability skill (*COR*) corresponding to each selected mode from SVD between climatic variables and autumn (OND) streamflow.

These selected coupled modes of Table 6.1 are represented in Figure 6.1 by plotting the homogeneous maps, their left and right normalized expansion coefficient time series, and the heterogeneous maps. Additionally, Figure 6.2 shows the correlations between left expansion coefficients of the selected SVD modes shown in Table 6.1 and the teleconnection indices for the same season, while Figure 6.3 shows the correlations between these expansion coefficients and the principal components of the respective atmospheric and oceanic fields.

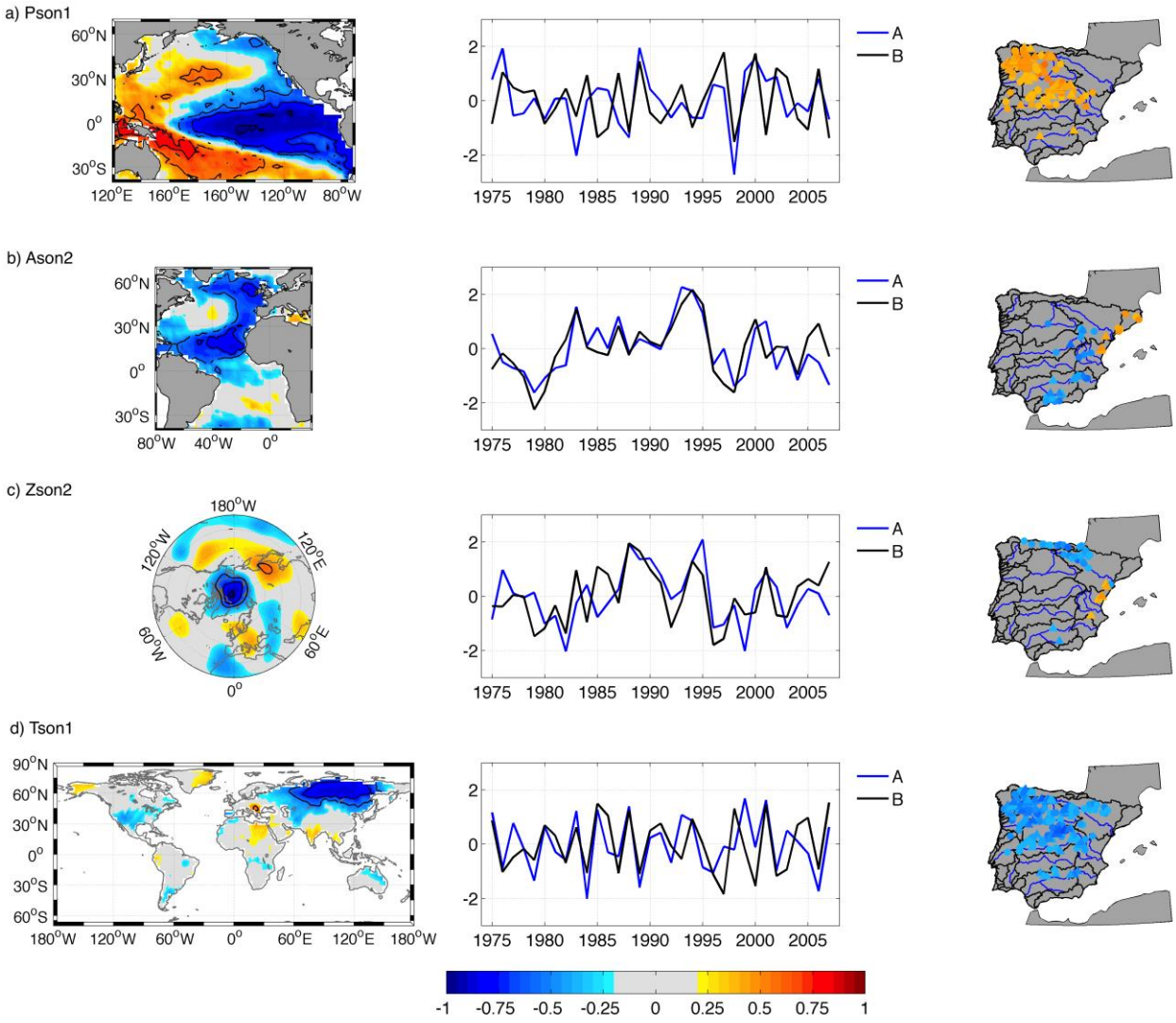


Figure 6.1. a)-p) Selected modes from lagging SVD analysis between seasonal atmospheric and oceanic variables leading autumn streamflow. Left panels display the homogeneous maps (contours indicate areas with significant correlation), central panels show the left (A, blue line) and right (B, black line) normalized expansion coefficient time series, and right panels display the heterogeneous maps.

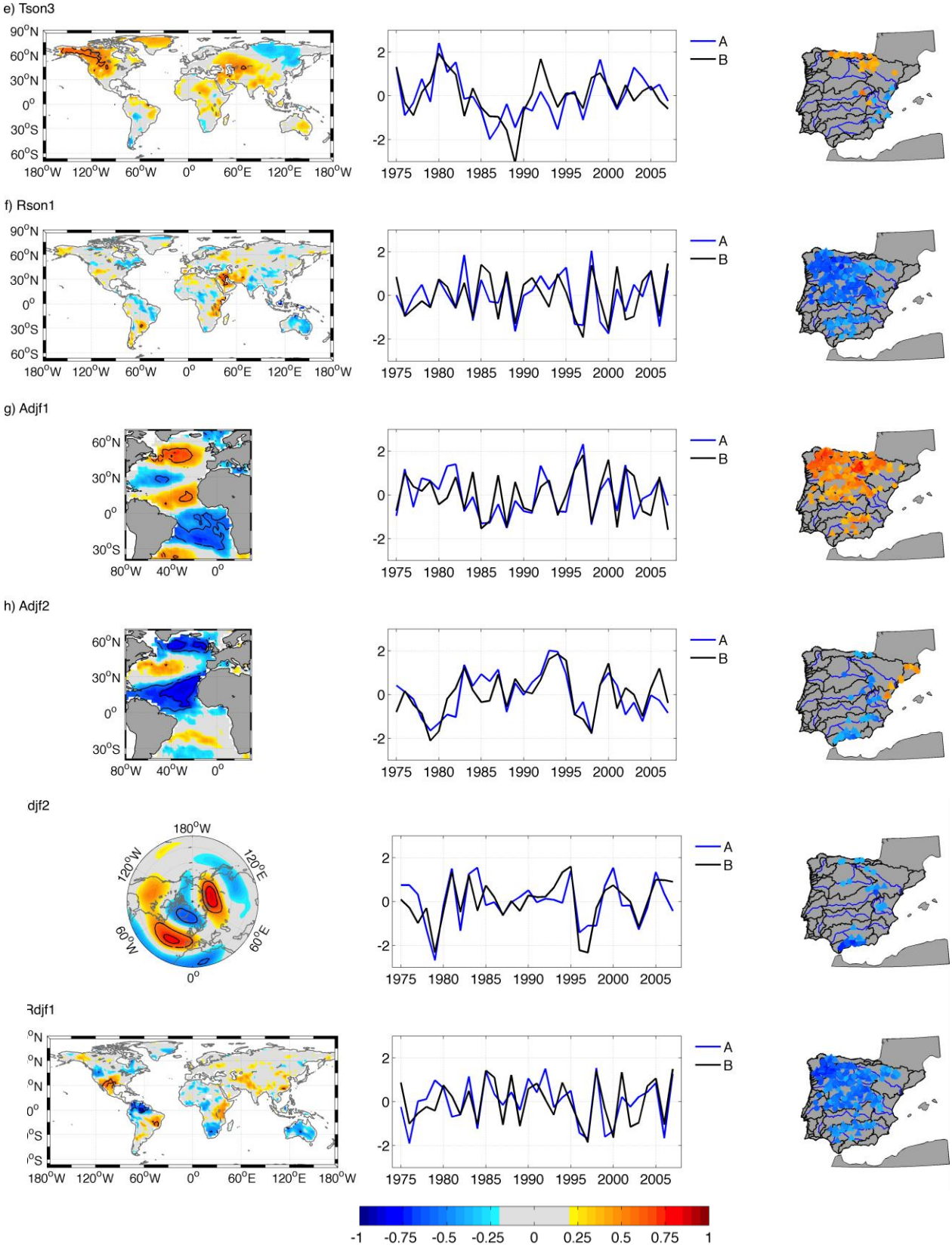


Figure 6.1. Cont.

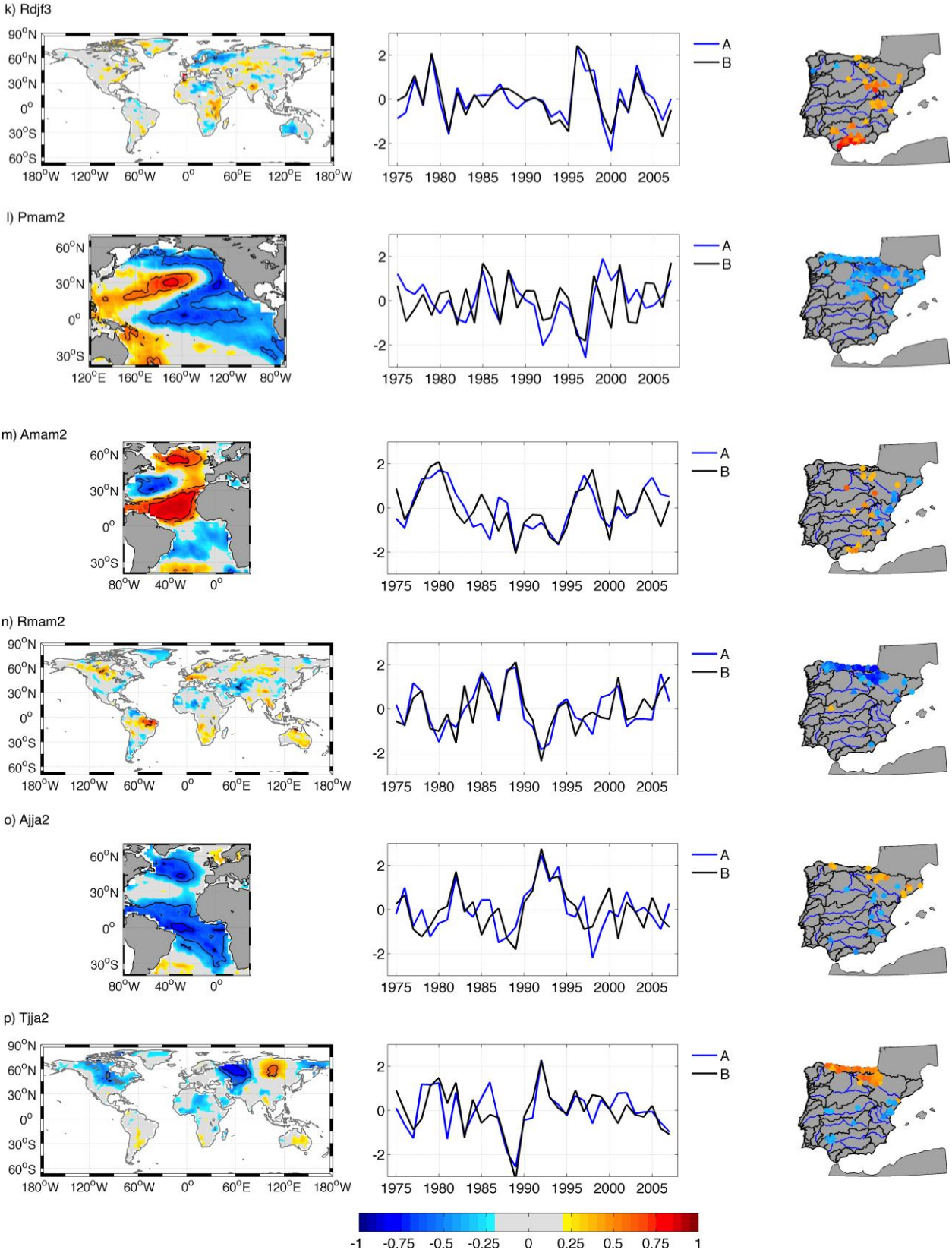


Figure 6.1 Cont

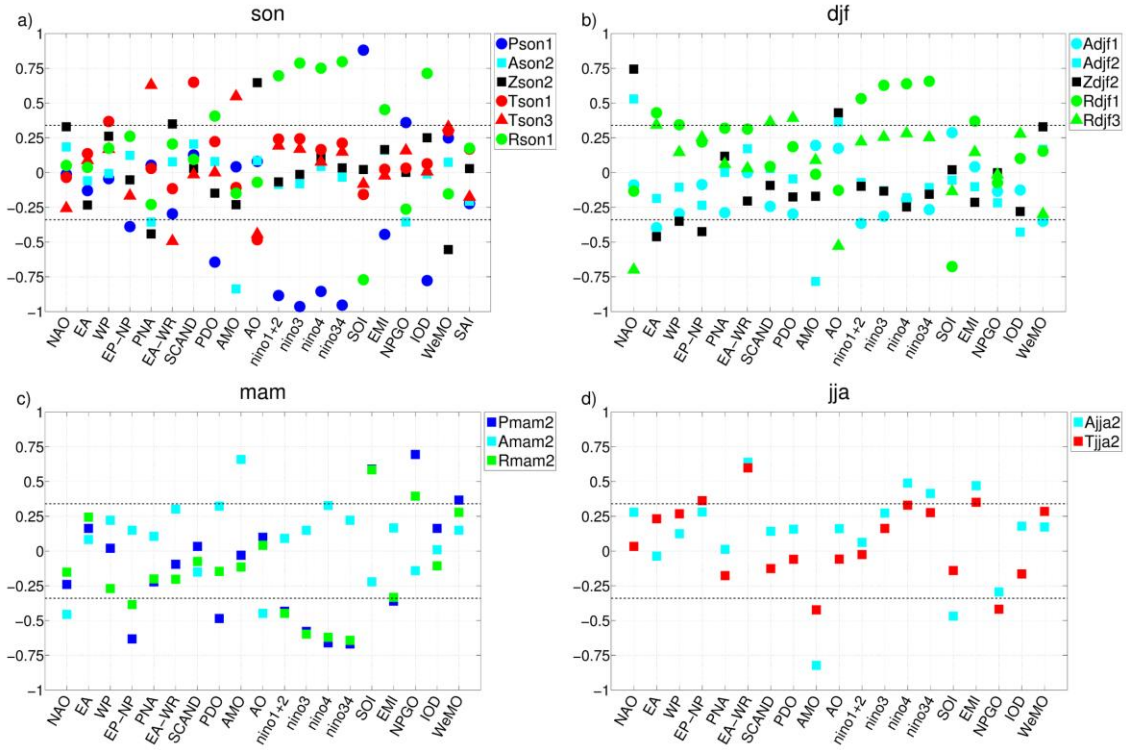


Figure 6.2. Correlations between left expansion coefficients of the selected SVD modes shown in Table 6.1 and teleconnection indices for the same season. a) autumn (son), b) winter (djf), c) spring (mam) and d) summer (jja).

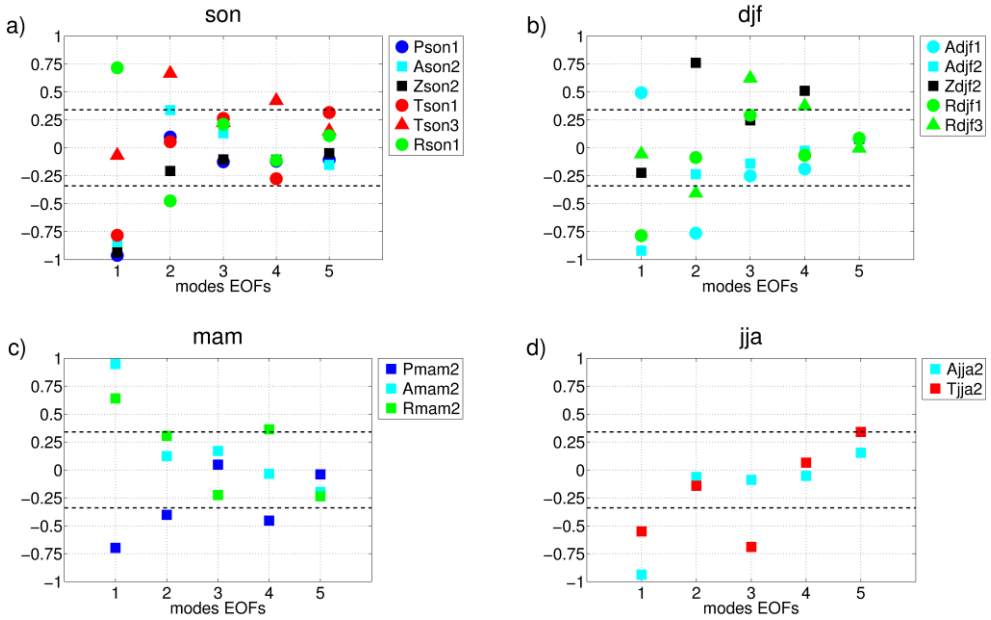


Figure 6.3. Correlations between left expansion coefficients of the selected SVD modes shown in Table 6.1 and the principal components of the respective atmospheric and oceanic fields for the same season. a) autumn (son), b) winter (djf), c) spring (mam) and d) summer (jja).

Subsequently, a description of the modes summarized in Table 6.1 is presented.

a) **Pson1:** The first SVD mode corresponding to the coupling between OND streamflow and Pacific SST of previous autumn (son) is significant at 98% of confidence level and comprises 75.2% of the square covariance fraction. According to the homogenous map (Figure 6.1a, left panel), this mode represents the Equatorial Pacific SST area, which is related to the El Niño phenomenon. In fact, the left expansion coefficient correlates significantly (above 0.8 in absolute values) with indices El Niño1+2, El Niño3, El Niño3.4 and SOI in autumn (Figure 6.2a). In addition, it presents a high correlation with the first EOF of Pacific SST variability in autumn (Figure 6.3a). The associated heterogeneous map links this mode to the following autumn streamflow variability in the north-western quadrant of the IP (Figure 6.1a, right panel). This result also agrees with that found in Chapter 5, where teleconnection indices related to autumn ENSO correlate with following autumn streamflow. The strength of the coupling is not very high (0.51), and the *COR* value is 0.32 (significant at 90%).

b) **Ason2:** It is the second mode of the SVD between previous autumn Atlantic SST and following autumn streamflow in the IP accounts 18.0% of *SCF*. However, the Monte Carlo test indicates that is only significant at 87% of confident level. Figure 6.1b shows the homogeneous and heterogeneous maps corresponding to this mode. The homogeneous map displays a like 'horse-shoe pattern' in the North Atlantic Ocean. The expansion coefficient of the left pattern presents a high correlation value with AMO index (Figure 6.2a). In addition, this mode correlates notably (higher than 0.80) with the first EOF of autumn Atlantic SST (Figure 6.3a). The heterogeneous map groups stations in the Mediterranean slope. The correlation between left and right expansion coefficient (the strength of the coupling) is 0.79 and the *COR* value is 0.49.

c) **Zson2:** This mode accounts 28.2% of the square covariance, however, the level of significant is only 35 %, according to Monte Carlo test. Although the strength of the coupling is not very high (0.60), the *COR* shows the greatest value between all selected modes (0.51). When analysing the homogeneous map of this mode (Figure 6.1c), significant correlations appear with opposite signs in the Arctic and the east of China. From Figure 6.2, this mode also presents significant correlation with AO and with WeMO indices. This last finding could be related to the negative and positive centres (though no significant) identified near Canary Islands and central Europe, similar to the pressure centres that define the WeMO index. The expansion coefficient of the left field

(autumn Z500) correlates significantly with the first EOF from the PCA of autumn Z500 (Figure 6.3a), showing some similitudes with its spatial pattern (although the intensity of action centres is smaller than in case of the EOF mode). The heterogeneous map identifies significant negative correlation in the Cantabrian Basin and upper Ebro Basin and positive in lower Ebro area.

**d) Tson1:** The first mode from the SVD between autumn temperature over land and following autumn streamflow is selected for the purpose of this study since *COR* value is 0.30. This mode, significant at 95% level according to Monte Carlo test, comprises 57.1% of the square covariance, and the strength of the coupling is 0.58. It presents a significant correlation with autumn SCAND index (Figure 6.2a), which can be explained since the region highly correlated with the expansion coefficient of the left field (autumn temperature) including most of Siberia (Figure 6.1d, left), which is also identified when correlating SCAND index with land temperature (figure available in Climate Prediction Centre, CPC, website). In addition, this pattern is highly correlated with the first EOF of autumn land temperature (Figure 6.3a). The heterogeneous map shows significant negative correlations with OND streamflow of most of the IP (except in the Mediterranean and Cantabrian slopes). This correlation pattern differs from that obtained when correlating with autumn SCAND index.

**e) Tson3:** Despite neither the *SCF* explained by this mode and the significant level are very high (8.9% and 85%, respectively), the strength of the coupling is (0.72). Also, *COR* value is significant (0.34). The homogeneous map for this mode (Figure 6.1e) identifies positive correlations in a region located in western Canada and Alaska and also in a small region in central Asia. This mode presents a correlation value around 0.7 with the second EOF of autumn land temperature (Figure 6.3a). In addition its correlation with PNA, and AMO indices is reasonably high (0.65 and 0.50, Figure 6.2a) and with AO (-0.4, Figure 6.2a). In fact, positive correlation in Alaska and west of Canada appears in the temperature pattern associated with October PNA index (from CPC website). Positive correlation with the expansion coefficient of the left field appears in stations located in the Cantabrian and upper Ebro Basins and negative in some stations of the Segura and Jucar Basins (Figure 6.1e). This correlation pattern presents some similarities with the correlation map obtained between autumn PNA index and autumn streamflow.

**f) Rson1:** This mode explains 51.8% of the square covariance, with a significance level of 98%. The strength of the coupling is 0.76 and *COR* value is 0.35. According the correlation between the left expansion coefficient and teleconnection indices of previous

autumn, high values (above 0.70 in absolute value) are found with ENSO, SOI and IOD indices (Figure 6.2a). The positive phase of IOD is related to above normal precipitation over western Indian Ocean and below normal precipitation over Indonesia and Australia. This spatial pattern is similar to the homogenous map corresponding to this mode (Figure 6.1f). Additionally, the impact of El Niño events on precipitation (see CPC website) displays a correlation map similar to the homogenous map of this mode. Also the Indian Ocean has been identified as a candidate for the modulation of ENSO variability (e.g. Jansen et al., 2009; Izumo et al., 2010; Frauen and Dommenget, 2012). Hence, it seems that this mode represents in some way the influence of atmospheric mechanism related to ENSO and IOD in autumn streamflow of IP Rivers (in most of them, according to the heterogeneous correlation map shown in Figure 6.1f). Additionally, it presents a high correlation (again greater than 0.70) with the first EOF of autumn global precipitation (Figure 6.3a).

**g) Adjf1:** A total of 54.5% of the *SCF* is explained by this mode and the Monte Carlo test indicates that is significant at 99%. The strength of the coupling is 0.72 and *COR* value is 0.45. The homogeneous map (Figure 6.1g) shows the existence of four areas with significant correlations. In the north Atlantic, there are three centres that display a pattern similar to the 'Atlantic tripole'. A fourth centre (with stronger correlation than the previous ones) appears in the tropical Atlantic (between 0-30ºS) as significant. When correlating the left expansion coefficient with the EOFs of winter Atlantic SST, the highest correlation is found with the second EOF (whose pattern mainly represents the 0-30ºS region). The correlation of this mode with the teleconnection indices in winter (Figure 6.2b) does not show any relevant result (considering as relevant correlations higher than 0.5). The heterogeneous correlation map exhibits significant correlations is most of the IP (Figure 6.1g).

**h) Adjf2:** This mode accounts for 14.9% of the square covariance and the strength of coupling is 0.80, the *COR* value is found significant (0.50), while the *SCF* is only significant at 87% of confidence level. However, the correlation of the expansion coefficient of left field with the first EOF of winter Atlantic SST is remarkable (above 0.9). This dominant mode of North Atlantic SST variability is associated with the NAO variability (Visbeck et al., 2001). Also, correlation with winter AMO index (Figure 6.2b) is high (above 0.75 in absolute value). The homogenous and heterogeneous maps of this mode (Figure 6.1h) are reasonably similar to those found for Ason2 mode (Figure 6.1b),

so it seems logical to consider it as a temporal evolution of the coupling between Atlantic SST and autumn IP streamflow.

i) **Zdjf2:** The *SCF* explained by this mode reaches up 18.3%, with a significance of 88%, according to Monte Carlo. The strength of the coupling is 0.72 and *COR* value is 0.48. The homogeneous map (Figure 6.1i) describes three centres of action. One centre (negative) is located over Greenland and another one (positive) is located over Azores. This configuration is similar to the north-south dipole of anomalies corresponding to winter NAO. The difference lies on a third centre (positive) located over Siberia. The heterogeneous map reveals significant negative correlation of this mode with stations located in some areas of the Iberian System and, especially, in the Mediterranean Andalusian Basin. This finding is rather similar to the obtained in Chapter 5, where NAO index of previous winter was found a good predictor of following autumn streamflow for stations in the Mediterranean Andalusia Basin. Additionally, the correlation with teleconnection indices (Figure 6.2b) shows a high value with NAO index (around 0.75) and it presents a remarkable correlation with the second EOF obtained from the PCA analysis of the winter geopotential height at 500 hPa (Figure 6.3b).

j) **Rdjf1:** This mode, significant at 94% of confidence level, comprises 48.2% of the square covariance fraction. The strength of the coupling is 0.73 and the *COR* value is 0.31. The left expansion coefficient presents a high correlation with winter ENSO signal (Figure 6.2b). In addition to this, the homogeneous map (Figure 6.1j) is rather similar to the impacts of El Niño events on winter precipitation (see CPC website). A positive phase (El Niño) is related to an increase of precipitation in California and Mexico, due to a more southerly, zonal, storm track. Also, drier conditions are found in South Africa and the northern coast of South America. Conversely, in Uruguay and southern Brasil rainfall increases on average. Hence, this pattern identified through SVD analysis, seems to be closely related with El Niño phenomenon during winter. The heterogeneous map indicates high negative correlations of the expansion coefficient with most stations (except Mediterranean and Cantabrian slopes) of the IP (Figure 6.1j).

k) **Rdjf3:** The third mode of the SVD between winter precipitation and following autumn streamflow in the IP only account for 9.1% of the *SCF*, with Monte Carlo test throwing a significance level of 43%, the strength of coupling is 0.92 and *COR* value is 0.29 (significant at 90%) (0.29). The homogeneous map (Figure 6.1k) identifies a high significantly correlated region over the IP, which means that this mode is linked to the winter precipitation over the IP. In fact, when correlating this mode with winter

teleconnection indices (Figure 6.2b), it can be seen a high correlation with NAO index (which is known as the main driven of winter precipitation over the IP). The heterogeneous map (Figure 6.1k) shows high correlation mainly with stations in Mediterranean Andalusian Basin and with stations located along the mountain ranges of the Iberian System. As it was remarked when mode Zdjf2 was described (Figure 6.1i), this heterogeneous map looks like similar to that found in Chapter 5 when we analysed the correlation of winter NAO index with following autumn streamflow. Furthermore, it presents a notable correlation (about 0.7) with the third EOF of winter global precipitation (Figure 6.3b).

l) **Pmam2:** This mode represents 27.9% of the square covariance, with a significance of 95%. The strength of the coupling is 0.54, and *COR* value is 0.50. The homogeneous map (Figure 6.1l) shows high correlations in North Pacific and Central Equatorial Pacific. The correlation with teleconnection indices (Figure 6.2c) reveals the highest values for El Niño-3.4, NPGO and EP-NP indices. The homogenous map (Figure 6.1l) displays a spatial structure similar to the second EOF of north Pacific SST (also known as Victoria mode), which is related to NPGO index (Di Lorenzo, 2010). Also the significantly correlated area in central Pacific is close to the region defined by El Niño-3.4, and, in relation with EP-NP, the homogeneous map resemblance the map of the correlation between EP-NP and surface temperature departures in April (CPC website), with opposite sign. The heterogeneous map (Figure 6.1l) presents high correlations in Cantabrian Basin, upper Tagus, Douro and Ebro Basins and also in some stations in the Internal Catalonian Basin.

m) **Amam2:** The second mode from SVD between spring Atlantic SST and following autumn streamflow explains 20.9 % of the square covariance. The Montecarlo test indicates a level of significance of the square covariance of 79%. The strength of the coupling is 0.73, whereas *COR* value is 0.29. From Figure 6.3c, the expansion coefficient for left field correlates highly with the first EOF of spring Atlantic SST (above 0.9), and also presents a high correlation with AMO index (Figure 6.2c). In addition, the homogeneous and heterogeneous maps (Figure 6.1m) looks like those found for the Adjf2 (Figure 6.1h) and Ason2 (Figure 6.1b) modes, which shows signs to be all related to the same oceanic variability mechanism.

n) **Rmam2:** A total of 17.1 % of the square covariance is explained by this mode, with a significance level of 95%. Although the strength of coupling is 0.81, the *COR* value is notably lower (0.29). The homogeneous map (Figure 6.1n) shows positive significant

correlation with precipitation over Brazil, Centre of Canada and Indonesia, and negative in some area of central Asia. This correlation map presents some similarities with the map of ENSO impact on precipitation in spring. Also, the correlation with ENSO indices (Figure 6.2c) is remarkable (around 0.6), which indicates that the physical mechanism behind it, is related to the ENSO phenomenon. In fact, from Chapter 5, correlation of some indices, such as El Niño-3 and El Niño-3.4, with autumn streamflow were found significant in stations located in the eastern Cantabrian Basin and upper Ebro Basin. However, these correlations are less extended than the ones described by the heterogeneous map (Figure 6.1n).

**o) Ajja2:** This mode account for 23.1 % of the square covariance, with a significance level of 89% according to Monte Carlo test and a strength of coupling of 0.66. Additionally, the *COR* values is also relatively high (0.43). The homogenous map (Figure 6.1o) reflects a highly correlated area below Greenland and also a wide area in the tropical Atlantic. This mode presents a high correlation with the first EOF of summer Atlantic SST (Figure 6.3d) and also with AMO index and with EA-WR pattern (Figure 6.2d). Note that one of the centres of summer EA-WR is placed near to the area in north Atlantic that presents high correlation with this index. The heterogeneous map shows significant correlations with stations located in the Mediterranean sector.

**p) Tjja2:** The *SCF* explained by this mode is 20.7%, being significant at 94% significance level. The strength of the coupling is 0.75 whereas the *COR* value is 0.34. In addition, from Figure 6.2d, the teleconnection index with the highest correlation value with this mode is the EA-WR index (around 0.6). Also, the homogeneous map (Figure 6.1p) displays some common areas with the map of EA-WR impacts on temperature in July (CPC website), which could indicate some kind of relationship with this pattern. When focusing on the heterogeneous map (Figure 6.1p), stations in Cantabrian Basin and upper Ebro present positive correlations whereas some stations in lower Ebro and the southern half of IP present negative correlations. This pattern differs from that obtained by correlation analysis between previous summer EA-WR index and autumn streamflow. Hence, despite some similarities found with EA-WR pattern, there should be an additional physical mechanism behind.

Covariability with winter (JFM) streamflow

Similarly to the previous autumn streamflow analysis, Table 6.2 gives a summary for the SVD modes selected (those with COR above 0.29) from the analysis of the covariability patterns

between winter streamflow and the climate variables used in this study. In this Table, the information of the SCF, together with the significance level provided by Monte Carlo test (MC), the strength of the coupling (STR) and also the potential predictability skill (COR) are presented.

| **Mode** | ***SCF*** | ***MC*** | ***STR*** | ***COR*** |
|---|---|---|---|---|
| Pdjf3 | 6.5 | 82 | 0.71 | 0.45 |
| Adjf2 | 15.4 | 94 | 0.81 | 0.35 |
| Adjf3 | 11.3 | 99 | 0.83 | 0.51 |
| Zdjf1 | 54.1 | 22 | 0.49 | 0.29 |
| Tdjf2 | 27.8 | 87 | 0.44 | 0.36 |
| Pmam3 | 10.8 | 99 | 0.68 | 0.44 |
| Amam2 | 23.2 | 85 | 0.65 | 0.34 |
| Tmam3 | 8.3 | 66 | 0.79 | 0.60 |

Table 6.2. Summary of the square covariance fraction (*SCF*), the significance given by Monte Carlo test (*MC*), both in percentage, the strength of the coupling (*STR*), and the predictability skill (*COR*) corresponding to each selected mode from SVD between climatic variables and winter (JFM) streamflow.

The selected coupled modes shown in Table 6.2 are represented in Figure 6.4 by plotting the homogeneous maps, their left and right normalized expansion coefficient time series, and the heterogeneous maps. Figures 6.5 and 6.6 display the correlations between left expansion coefficients of the selected SVD modes and the teleconnection indices for the same season, and the correlations between these expansion coefficients and the principal components of the respective atmospheric and oceanic fields, respectively.

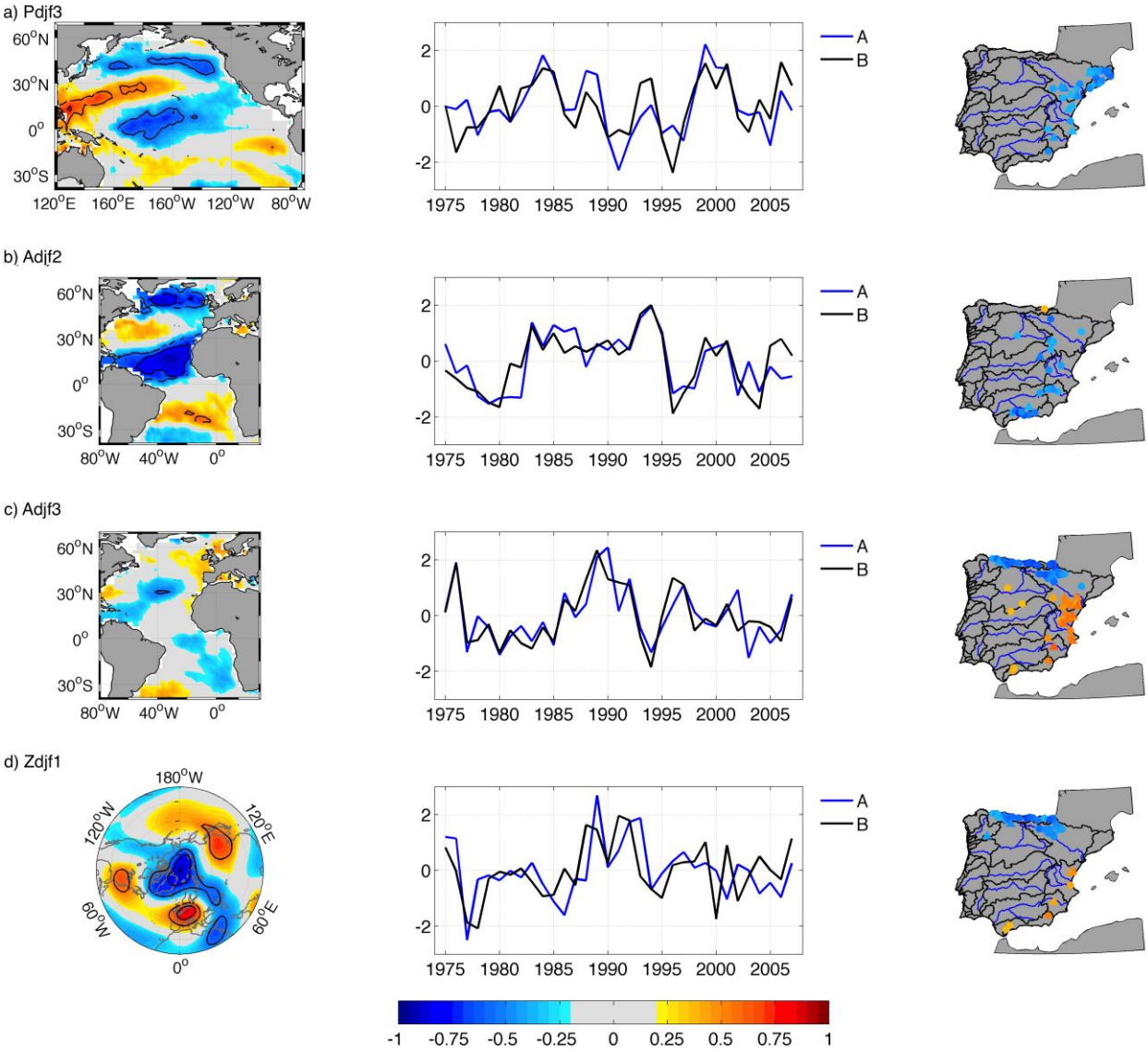


Figure 6.4. a)-h) Selected modes from lagging SVD analysis between seasonal atmospheric and oceanic variables leading winter streamflow. Left panels display the homogeneous maps (contours indicate areas with significant correlation), central panels show the left (A, blue line) and right (B, black line) normalized expansion coefficient time series, and right panels display the heterogeneous map.

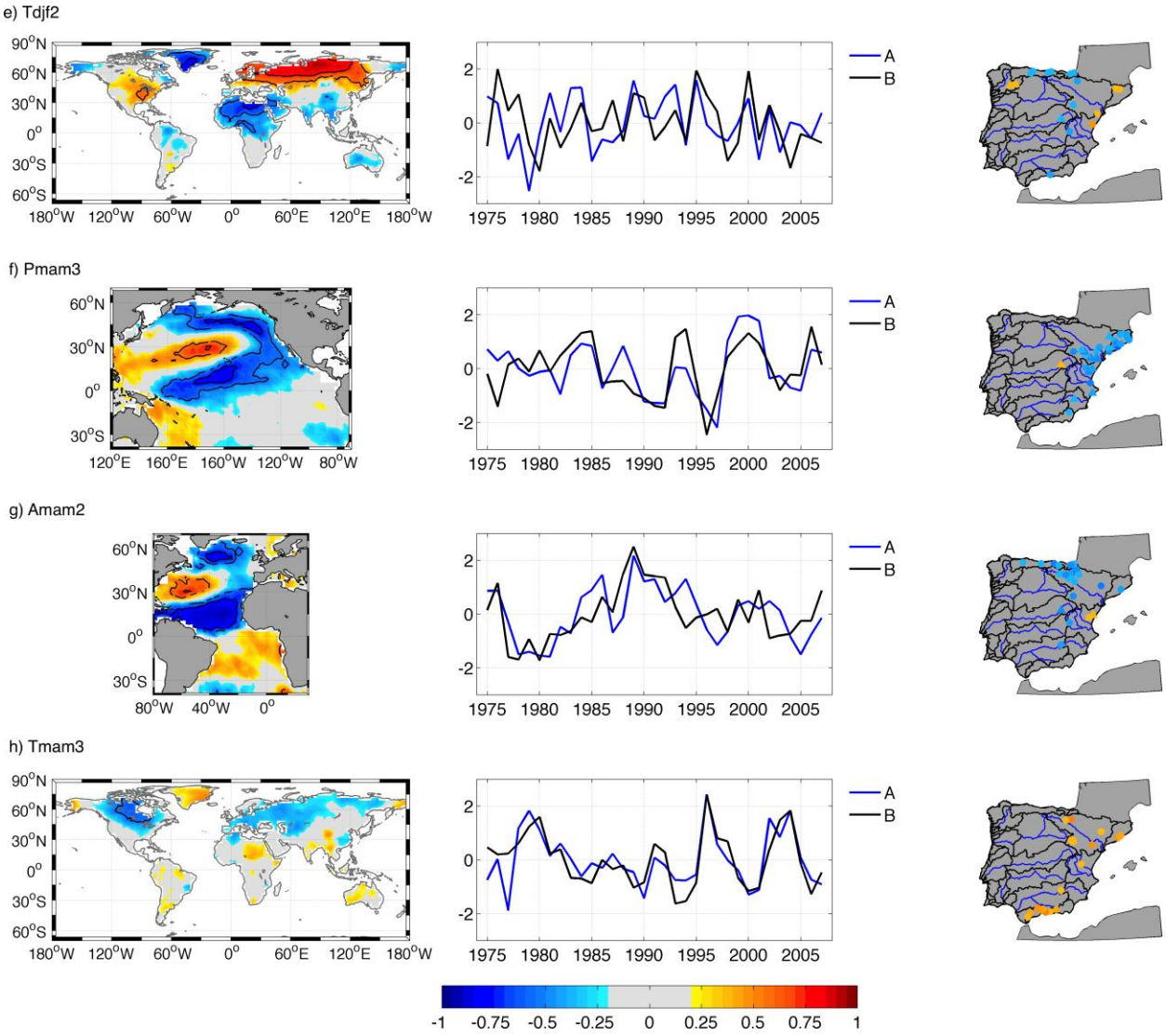


Figure 6.4. Cont.

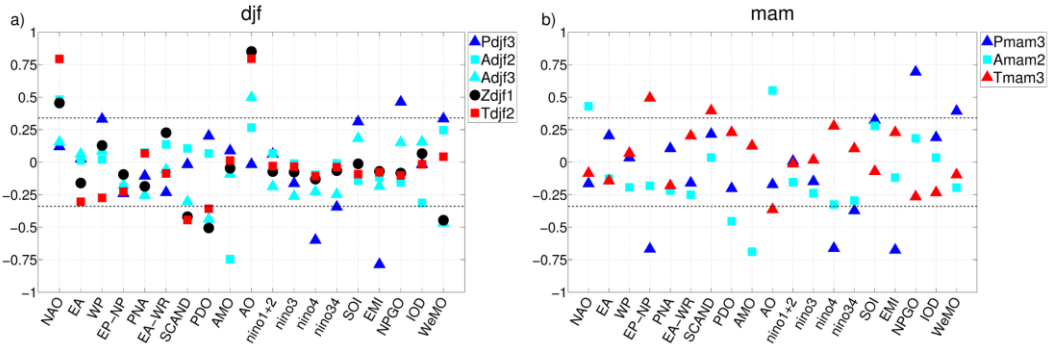


Figure 6.5. Correlations between left expansion coefficients of the selected SVD modes shown in Table 6.2 and teleconnection indices for the same season. a) winter (djf), b) spring (mam).

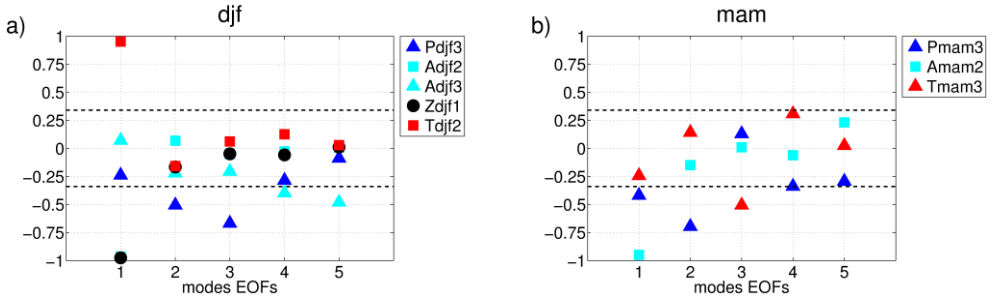


Figure 6.6. Correlations between left expansion coefficients of the selected SVD modes shown in Table 6.2 and the principal components of the respective atmospheric and oceanic field for the same season. a) winter (djf), b) spring (mam).

Subsequently, a description of the modes summarized in Table 6.2 is provided:

**a)** **Pdjf3:** This mode presents a strong coupling (0.71) and a relative high *COR* value (0.45), but only accounts 6.5% of the square covariance, with a low significance level (82%). However, there are some reasons that suggest this mode should be considered in this study. In particular, this mode presents a high correlation with EMI, El Niño-4 and NPGO indices (Figure 6.5a). In the definition of EMI index given in Chapter 5, this index is related to a strong anomalous warming in central tropical Pacific (also the SST region where El Niño-4 index is defined). This region shows high correlation values in the homogeneous map (Figure 6.4a, left panel). In addition, the two highly correlated areas in North Pacific (with opposite signs) are similar to the NPGO pattern from the North Pacific SST. Di Lorenzo (2010) postulated that Central Pacific warming (related to EMI) drives changes in the large-scale atmospheric circulation in the central North Pacific, which is integrated by the ocean to yield NPGO. The heterogeneous map (Figure 6.4a, right panel) indicates that this mode is linked with the winter streamflow variability mainly in lower Ebro and Catalonian Basin. In addition, significant correlations with winter EMI were found in Chapter 5 in Catalonian Basin when correlating winter EMI with following winter streamflow.

**b)** A**djf2:** The *SCF* represented by this mode is 15.4%, with a significance level of 94% according to Monte Carlo test. The coupling is strong (0.81), but the *COR* value is moderately inferior (0.35). The homogenous map (Figure 6.4b) looks like the Atlantic tripole SST pattern, similar to the Adjf2 mode coupled with autumn streamflow (Figure 6.1h), and the SST pattern related to winter NAO. Moreover, the heterogeneous map (Figure 6.4b) resembles the obtained for Adjf2 coupled with autumn streamflow

(although in this case the Catalonian Basin are not linked with this mode). Also, it strongly correlates with the first EOF of winter Atlantic SST (Figure 6.6a) and with AMO and NAO indices (Figure 6.5a).

c) **Adjf3:** Despite the *SCF* corresponding to this mode (11.3%) is significant (99%) and the strength of the coupling and *COR* value are high (0.83 and 0.51, respectively), some doubts arise about the physical meaning of this mode when observing the homogenous map (Figure 6.4c). On this map, only a small area in the central North Atlantic presents significant correlation with the left expansion coefficient. Also, the left expansion coefficient does not present a high correlation with the EOFs of winter Atlantic SST (Figure 6.6a). Additionally, only AO index presents a reasonable high (above 0.5) correlation with this mode (Figure 6.5a). On the other hand, the heterogeneous map (Figure 6.6c) reflects negative correlations with most stations in the Cantabrian slope and positive with stations in the Mediterranean slope, particularly in the lower Ebro area.

d) **Zdjf1:** This mode accounts for a large fraction of the square covariance (54.1%), which is not found significant according to Monte Carlo test (22% of confidence level). Despite the strength of the coupling is not very high (0.49) the *COR* value remains significant (0.29). The homogeneous map (Figure 6.4d) shows a pattern similar to the winter AO, with presence of a negative correlated area in the North Pole and a positive correlated area at mid-latitudes (from IP to the east coast of North America), and a centre of positive correlation in east Asia-west Pacific. In addition, it presents a high correlation with the first EOF found from a PCA of winter Z500 (Figure 6.6a), which correspond to the AO, and with winter AO index (Figure 6.5a). The heterogeneous map (Figure 6.4d) reveals negative correlations with stations in the Cantabrian slope and positive in Mediterranean area (except in Ebro and Catalonian Basins). This map is rather similar to that obtained from the correlation analysis between winter AO index with following winter streamflow in Chapter 5. Hence, despite the *SCF* is not significant, this SVD mode shows sign of being related to the AO index.

e) **Tdjf2:** The fraction of the square covariance explained by this mode is 27.8%, with a significance level of 87%. The coupling is not very strong (0.44), although *COR* value is significant (0.36). The homogeneous map (Figure 6.4e) displays positive correlations over northern Europe-Asia, and the east coast of United States, and negative correlations in northern Africa and Greenland. This pattern has the appearance of the impact of January NAO on temperature (see CPC website). Furthermore, the left expansion coefficient correlates highly with winter NAO index (Figure 6.5a) and with the first EOF

of winter global temperature (Figure 6.6a). For this reason, this mode seems to be associated with winter NAO. The heterogeneous map indicates that it correlates negatively with some stations in Cantabrian range and positively with stations in Miño-Sil and Catalonian Basins.

**f) Pmam3:** This mode explains 10.8% of the square covariance (with a significance level of 99%). Also, the *STR* and *COR* values are significant (0.68 and 0.44, respectively). The homogeneous and heterogeneous maps (Figure 6.4f) are similar to the ones associated with Pdjf3 mode. Moreover, this mode presents as high correlations with NPGO, EMI and El Niño-4 as Pdjf3 did. In addition, from Figure 6.5b it could be also seen high correlations with EP-NP index (note that this pattern is a spring-summer-autumn pattern, which explains why it was not found correlation with it in winter).

**g) Amam2:** A total of 23.2% of the square covariance is explained by this mode, with a significance level of 85%. The strength of the coupling is 0.65 and *COR* reaches up 0.34. Similarly to what was commented for Adjf2, the homogenous map (Figure 6.4g) looks like the 'tripole pattern' of North Atlantic SST. However, the heterogeneous map (Figure 6.4g) presents some differences with respect to that found for Adjf2 mode. In this case, stations significantly correlated with this mode are located mostly in Cantabrian Basin and upper Ebro. This mode exhibits a high correlation value with AMO (Figure 6.5b) and with the first EOF of North Atlantic spring SST (Figure 6.6b).

**h) Tmam3:** This mode explains only 8.3% of the square covariance, with a significance level of 66%. The strength of the coupling is 0.79 and the *COR* is the highest of all SVD modes coupled with winter streamflow selected (0.60). The homogeneous map (Figure 6.4h) shows significant correlations in Canada. When correlating the left expansion coefficient of this mode with teleconnection indices (Figure 6.5b) and spring land temperature EOFs (Figure 6.6b), only relevant (above 0.5) correlations are found with EP-NP and the third EOF, respectively. Regarding the possible connection with EP-NP, the correlation map of this index with spring land temperature (see CPC website) also displays an area with high correlation in western North-America. Stations significantly correlated with this mode are mainly located in the Mediterranean Andalusian area.

Covariability with spring (AMJ) streamflow

Table 6.3 summarizes the information of the *SCF*, together with the significance level provided by Monte Carlo test (*MC*) and also the strength of the coupling (*STR*) and the potential predictability skill (*COR*) for the SVD modes selected. Modes are named following the previously indicated nomenclature.

| **Mode** | ***SCF*** | ***MC*** | ***STR*** | ***COR*** |
|---|---|---|---|---|
| Rmam3 | 12.2 | 95 | 0.85 | 0.52 |
| Pjja1 | 56.2 | 84 | 0.55 | 0.29 |
| Ajja2 | 21.1 | 78 | 0.74 | 0.43 |
| Zjja1 | 74.4 | 100 | 0.58 | 0.42 |
| Tjja1 | 43.9 | 89 | 0.82 | 0.30 |
| Tjja2 | 16.9 | 95 | 0.72 | 0.29 |
| Ason1 | 43.3 | 69 | 0.68 | 0.32 |
| Zson3 | 12.7 | 97 | 0.69 | 0.38 |
| Adjf2 | 21.7 | 100 | 0.72 | 0.58 |
| Zdjf1 | 57.5 | 77 | 0.55 | 0.29 |
| Tdjf1 | 56.3 | 89 | 0.58 | 0.30 |
| Rdjf3 | 11.7 | 100 | 0.84 | 0.55 |

Table 6.3. Summary of the square covariance fraction (*SCF*) and its significance given by Monte Carlo test (*MC*), both in percentage, the strength of the coupling (*STR*), and the predictability skill (*COR*) correspondent to each selected mode from SVD between climate variables and spring (AMJ) streamflow.

Figure 6.7 represents the homogeneous and heterogeneous maps (left and right panels, respectively) and the temporal series of the expansion coefficients (central panel) for the selected modes. And Figures 6.8 and 6.9 show the correlations between left expansion coefficients of the selected SVD modes and the teleconnection indices and the principal components of the respective atmospheric and oceanic fields, respectively, for the same season.

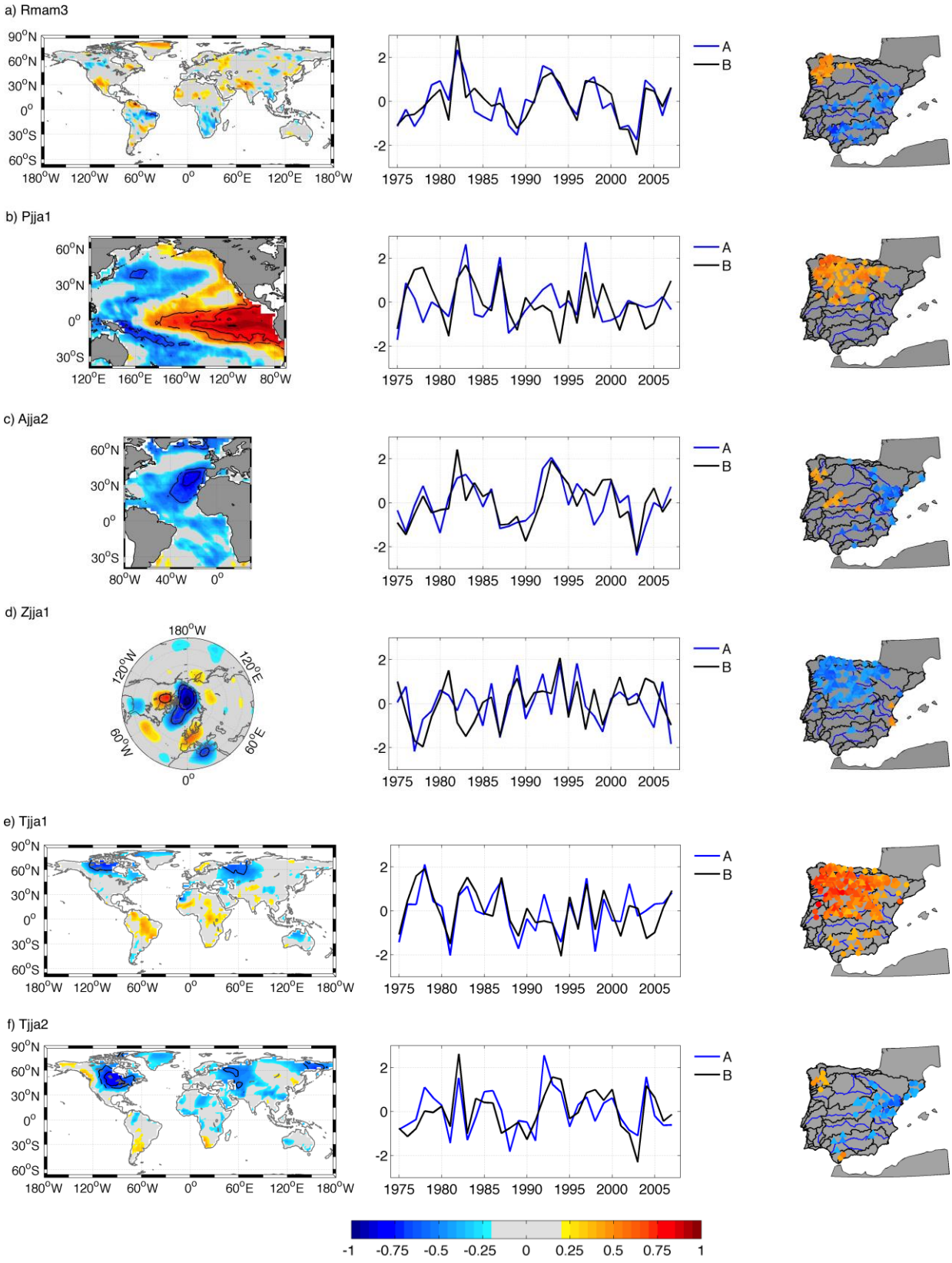


Figure 6.7. a)-l) Selected modes from lagging SVD analysis between seasonal atmospheric and oceanic variables leading spring streamflow. Left panels display the homogeneous maps (contours indicate areas with significant correlation), central panels show the left (A, blue line) and right (B, black line) normalized expansion coefficient time series, and right panels display the heterogeneous maps.

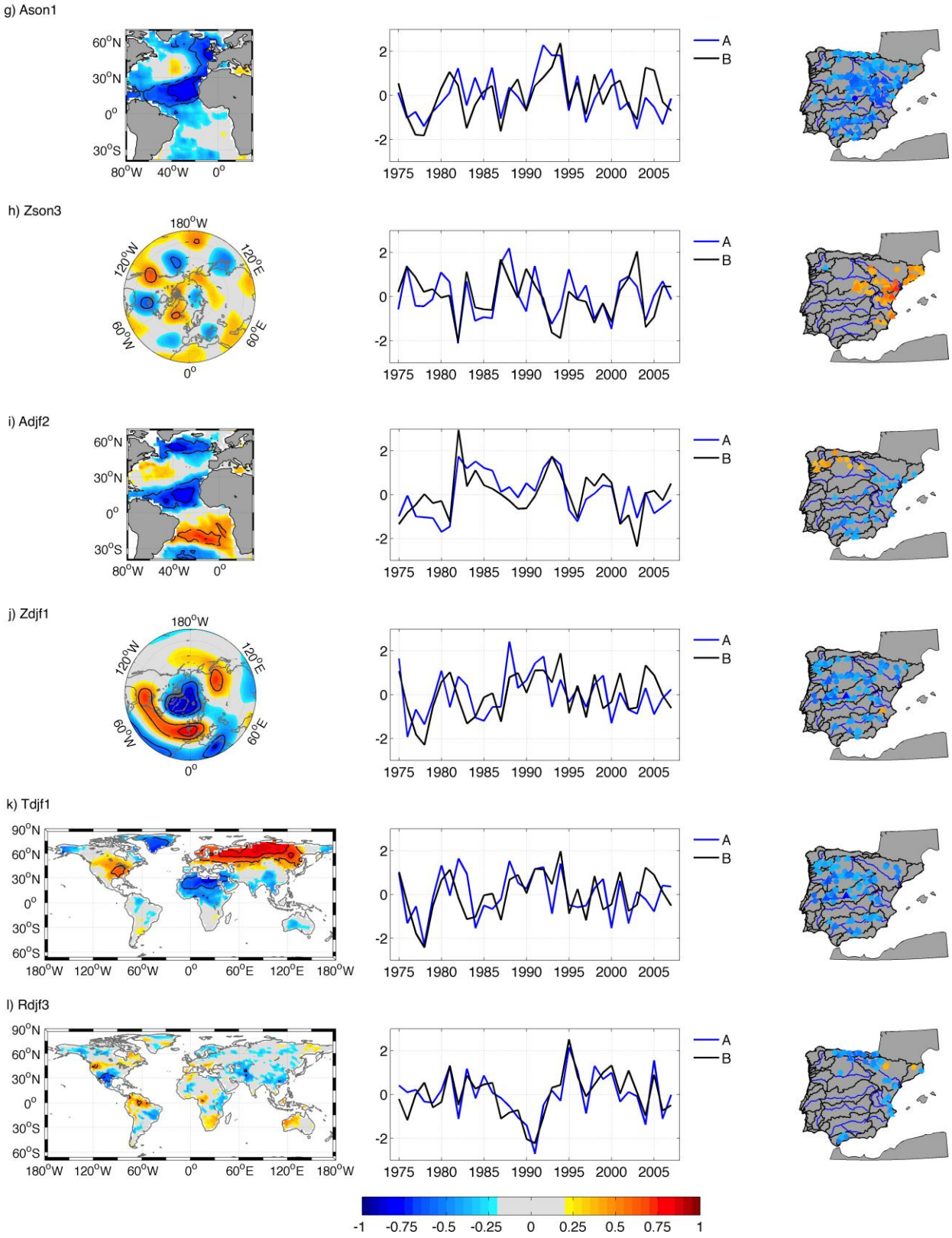


Figure 6.7 Cont

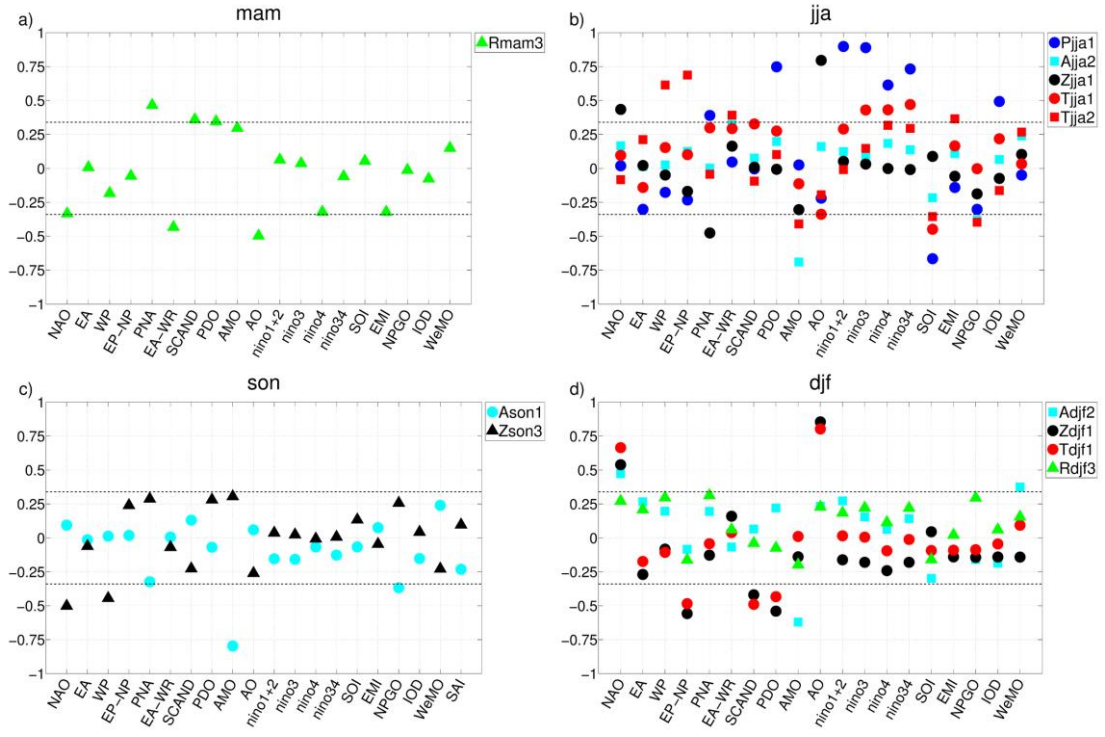


Figure 6.8. Correlations between left expansion coefficients of the selected SVD modes shown in Table 6.3 and teleconnection indices for the same season. a) spring (mam) and b) summer (jja), c) autumn (son), d) winter (djf).

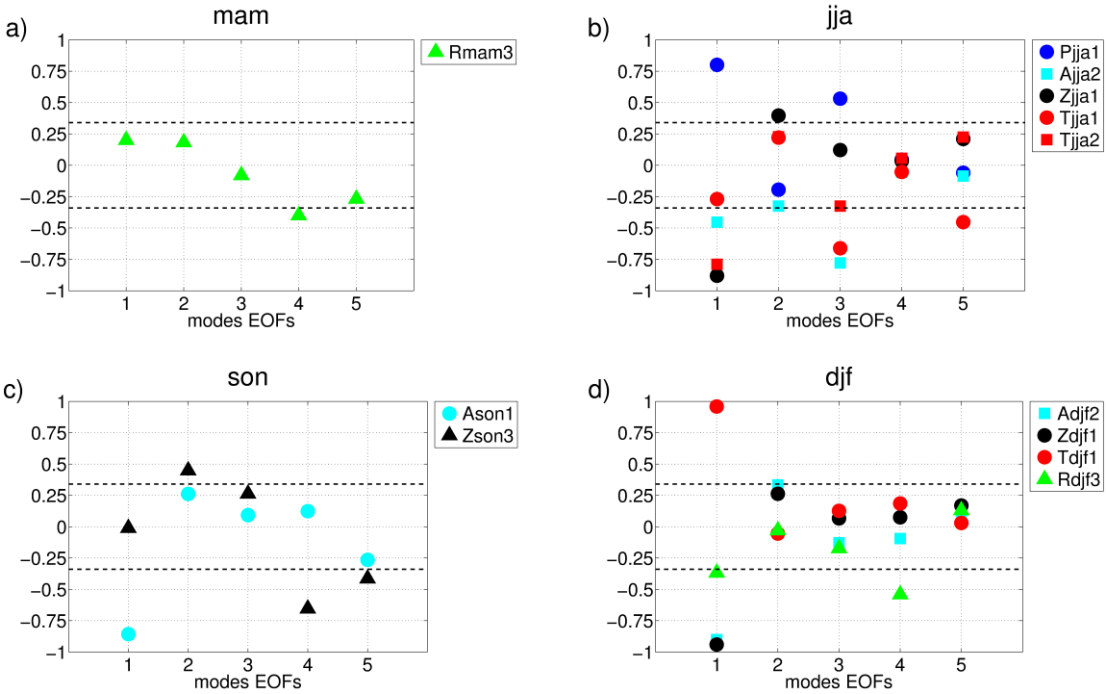


Figure 6.9. Correlations between left expansion coefficients of the selected SVD modes shown in Table 6.3 and the principal components of the respective atmospheric and oceanic field for the same season. a) spring (mam) and b) summer (jja), c) autumn (son), d) winter (djf).

A brief description of the modes summarized in Table 6.3 is carried out:

a) **Rmam3:** This mode explains 12.2% of the total square covariance with a significance level of 95%. It also presents a strong coupling (0.85) and a high *COR* value (0.52). However, the expansion coefficient corresponding to the left field (precipitation) does not present a notable correlation value with teleconnection indices (Figure 6.8a) or with the main modes of variability of this field in spring (Figure 6.9a). The homogeneous map associated with this mode (Figure 6.7a) shows a couple of significantly correlated small areas (positive around Venezuela and negative in northeast of Brazil), which are associated with spring streamflow in Miño-Sil Basin and with stations in the southern half of IP (according the heterogeneous map shown in Figure 6.7a).
b) **Pjja1:** The *SCF* associated with this mode is 56.2% (with a significance level of 84%), the strength of the coupling is 0.55 and *COR* value is 0.29. The homogenous map (Figure 6.7b) describes a highly correlated area in equatorial Pacific, similar to the related to El Niño phenomenon. In fact, the correlation of the left expansion coefficient with summer teleconnection indices (Figure 6.8b) is high with the PDO, El Niño-1+2, El Niño-3, El Niño-4, El Niño-3.4 and SOI indices. In addition, it presents high correlation with the first mode of Pacific summer SST (Figure 6.9b). The heterogeneous map (Figure 6.7b) shows significant correlations in the northwest quadrant of the IP. These findings are in agreement with the results showed in Chapter 5 and also with that found by other authors (Lorenzo et al., 2010), about the influence of equatorial Pacific summer SST on streamflow or rainfall on the northwestern of IP in the following spring.
c) **Ajja2:** The fraction of the square covariance explained by this mode reaches up 21.1%, at 78% of significance level. The strength of the coupling is 0.74 and *COR* value is 0.43. The homogenous map (Figure 6.7c) shows an Atlantic SST region in front of the coast of north-western Africa and the IP and also near Greenland and Iceland. This pattern presents some similitudes with the third EOF of summer Atlantic SST (which has a third centre located between the two commented, with opposite sign), correlating above 0.75 in absolute value (Figure 6.9b). Also, it shows high correlation with AMO (Figure 6.8b). The heterogeneous map (Figure 6.7c) displays positive correlations in stations placed in Miño-Sil and lower Tagus Basins, and negative in lower Ebro, Jucar, Segura and Guadalquivir Basins.
d) **Zjja1:** This mode explains 74.4% of the square covariance fraction, being significant at 100% level. The strength of the coupling is 0.58 and *COR* is 0.42. The homogeneous

map (Figure 6.7d) displays negative correlations in North Pole and Mediterranean Sea, and positive ones in Canada. This configuration looks like summer AO pattern (or Northern Annular Mode). Actually, this mode correlates highly with summer AO index (above 0.75, according to Figure 6.8b) and with the first EOF of summer geopotential height in Northern Hemisphere (Figure 6.9b). The heterogeneous map (Figure 6.7d) indicates that this mode significantly correlates with stations located in north-western quadrant of IP (Tagus, Douro, Miño-Sil, Cantabrian and upper Ebro Basins). This correlation pattern looks like that found when we computed the correlation between the summer AO index and the spring streamflow in Chapter 5.

**e) Tjja1:** A total of 43.9% of the square covariance is explained by this mode, with a significance level of 89%. The coupling strength is 0.82 and *COR* value is 0.30. The homogeneous map (Figure 6.7e) shows two zones with significant correlations, one located over north Canada and other over north-western Russia. There is not a high correlation between this mode and summer teleconnection indices (in Figure 6.8b it can be seen as only indices related to ENSO present significant correlations, but always below 0.5). It correlates (Figure 6.9b) with the third EOF of summer land temperature variability (which seems to be related with summer temperature in western Russia). The heterogeneous map indicates that this mode is highly associated with streamflow variability in most of the IP (except for the Mediterranean slope).

**f) Tjja2:** This mode accounts for 16.9% of the *SCF* (with a significance level of 95%). *STR* and *COR* values are 0.72 and 0.29, respectively. The homogenous map (Figure 6.7f) presents significant negative correlations in western Russia, middle-east and eastern Canada, and north-eastern United States (surrounding Hudson Bay). This mode is highly correlated with the WP and EP-NP indices (Figure 6.8b). Certainly, the impacts of both summer WP and EP-NP patterns in surface temperature departures (CPC website) present negative values over western North America). In addition, this pattern presents high correlation with first EOF of summer surface temperature (Figure 6.9b). The heterogeneous map (Figure 6.7f) presents significant positive correlation with stations in Miño-Sil and Mediterranean Andalusian Basins, and negative in lower Ebro, Catalonian and Guadalquivir Basins. A similar correlation pattern with spring streamflow was found in Chapter 5 when we calculated the correlation with summer WP and EP-NP indices (with the exception of the Miño-Sil Basin, which did not present significant correlations).

**g) Ason1:** The *SCF* related to this mode is 43.3% (with a significance level of 69%), coupling strength and *COR* values are 0.68 and 0.32, respectively. According to the

homogeneous map (Figure 6.7g), this pattern looks like the Atlantic SST tripole pattern, which is similar to the Ason2 spatial pattern found when coupling autumn Atlantic SST with following autumn streamflow (Figure 6.1b). On the contrary, the heterogeneous map presents some differences. In this case, it shows negative correlations with most of stations (except in Miño-Sil and some parts of Douro, Segura and Jucar Basins). The left expansion coefficient significantly correlates with AMO index and with the first EOF of autumn Atlantic SST variability (Figures 6.8c and 6.9c, respectively).

**h) Zson3:** This mode explains 12.7% of the *SCF* (at 97% of significance level) and is coupled with strength of 0.69. The *COR* value is 0.38. The homogenous map (Figure 6.7h) represents a configuration similar to the Pacific Transition, which is a leading mode during late summer/early autumn (August and September) that captures anomalous wave-train at 500 hPa heights extending from central Pacific to eastern United States. In particular, this pattern presents five significant correlation centres (three positives placed in central Pacific, western United States and Greenland, and two negatives located in front of the coast of Alaska and eastern United States). In addition, the correlation between the left expansion coefficient and September Polar Transition is 0.53. Also, from Figure 6.8c, it shows significant correlation with NAO index (mainly related to the centre of anomalies located in Greenland and eastern United States, which is similar to the configuration of NAO in autumn) and WP index (related to the anomalous centres located in the Pacific Ocean). The mode also correlates significantly with the fourth EOF of autumn Z500 variability (Figure 6.9c). The heterogeneous map couples this mode with streamflow variability in upper Tagus, Jucar, Ebro and Catalonian Basins.

**i) Adjf2:** The *SCF* corresponding to this mode is 21.7% (with a significance level of 100%). It presents a coupling strength of 0.72 and *COR* value is the highest of the selected SVD modes (0.58). The homogeneous map (Figure 6.7i) shows a pattern rather similar to the found for the Ason1 mode (the Atlantic SST tripole pattern), but also appears some significant correlations in South Atlantic Ocean. This mode is highly correlated with the first EOF of winter Atlantic SST variability and correlates moderately with winter AMO as well (Figure 6.9d and 6.8d, respectively). The heterogeneous map (Figure 6.7i) displays positive correlations with stations placed in Miño-Sil and Douro Basins and negatives with stations located in Tagus, Guadalquivir, Segura, Jucar and lower Ebro Basins.

**j) Zdjf1:** This mode accounts for 57.5% of the square covariance (with a significance level of 77%). The coupling strength is 0.55 and *COR* value is 0.29. Although the *SCF* is not

significance at 90%, the homogenous map (Figure 6.7j) is rather similar to the winter AO/NAO pattern. In fact, the correlations with these teleconnection indices are around 0.8 and 0.55, respectively (Figure 6.8d). In addition, it correlates significantly with the first EOF of winter geopotential height at 500 hPa variability (Figure 6.9d). The heterogeneous map (Figure 6.7j) looks like that found when we correlated winter AO index with spring streamflow in Chapter 5, showing significant correlations in stations throughout the IP (except for the Cantabrian and Mediterranean slopes).

**k) Tdjf1:** A total of 56.3% of the square covariance is explained by this mode (with a significance level of 89%). The strength of the coupling is 0.58 and *COR* value is 0.30. These values are similar to that found for Zdjf1 mode. This is not the only point they share, since both present similar high correlation values with winter NAO and AO indices (Figure 6.8d). If focusing on the homogenous map (Figure 6.7k), it is rather similar to the impacts of winter NAO on surface temperature (see CPC website), showing positive correlations in north Europe, Siberia and east coast of United States, and negative ones in Greenland and northern Africa. Also, this pattern is highly correlated with first EOF of winter land surface temperature (Figure 6.9d). The heterogeneous map (Figure 6.7k) is almost the same that the observed for Zdjf1 mode, with negative correlations in most of IP, except for the Cantabrian and Mediterranean slopes.

**l) Rdjf3:** This coupled pattern is responsible for the 11.7% of the *SCF*, with a significance level of 100%. The strength of the coupling and the *COR* value are remarkable (0.84 and 0.55, respectively). The homogenous map (Figure 6.7l) presents two zones with significant positive correlations (one in the northwestern United States and the other in north of Brazil), and a more extended region with significant negative correlations in southwestern United States. On the other hand, correlations with teleconnecton indices do not clearly associate this precipitation pattern to any of them (Figure 6.8d). This precipitation pattern, despite does not present high correlation with the EOFs of winter global precipitation over land (Figure 6.9d), looks like similar to the one associated with the fourth EOF. The heterogeneous map shows significant correlations with stations located in the Cantabrian Basin and with some of them in the Mediterranean slope.

### 6.4.2 Stability analysis

Figure 6.10, 6.11 and 6.12 show the correlation maps between SVD modes and seasonal streamflow in those stations where correlations are found stable, for autumn, winter and spring streamflow, respectively. The most interesting results can be observed when comparing these maps with the corresponding heterogeneous maps displayed before (Figures 6.1, 6.4 and 6.7, right panels).

In case of autumn streamflow (Figure 6.10), the correlation with the Amam2 and Ajja2 modes was not found stable in most stations. Regarding Pson1 mode, stations in Miño-Sil Basin did not correlate stably with this mode. For Ason2 and Adjf2 modes, stations in Internal Catalonian Basins showed significantly but no stable correlations. In case of Tson1 and Adjf1, they did not correlate stably with stations in Tagus Basin. Only stations in Cantabrian Basin correlate stably with Tson3. In the remaining modes, most of the stations that correlate significantly with them also do it stably.

In case of winter streamflow (Figure 6.11), the most remarkable finding is that Tdjf2, Amam2 and Tmam3 modes do not correlate stably with almost any station.

Results from correlation stability analysis between SVD modes and spring streamflow (Figure 6.12) indicate that most of the stations that correlate significantly with Rmam3 do not do it stably. Also, it is worth to mention that some stations in Douro Basin correlate significant but not stable with Pjja1. Tjja2 do not correlates stably with stations in Internal Catalonian and Miño-Sil Basins. Similarly, stations in Miño-Sil, the headwater of Douro River and Guadalquivir Basin correlate significantly but not stably with Tdjf1, and Adjf2 only presents stable correlation with stations in Jucar and Segura Basin.

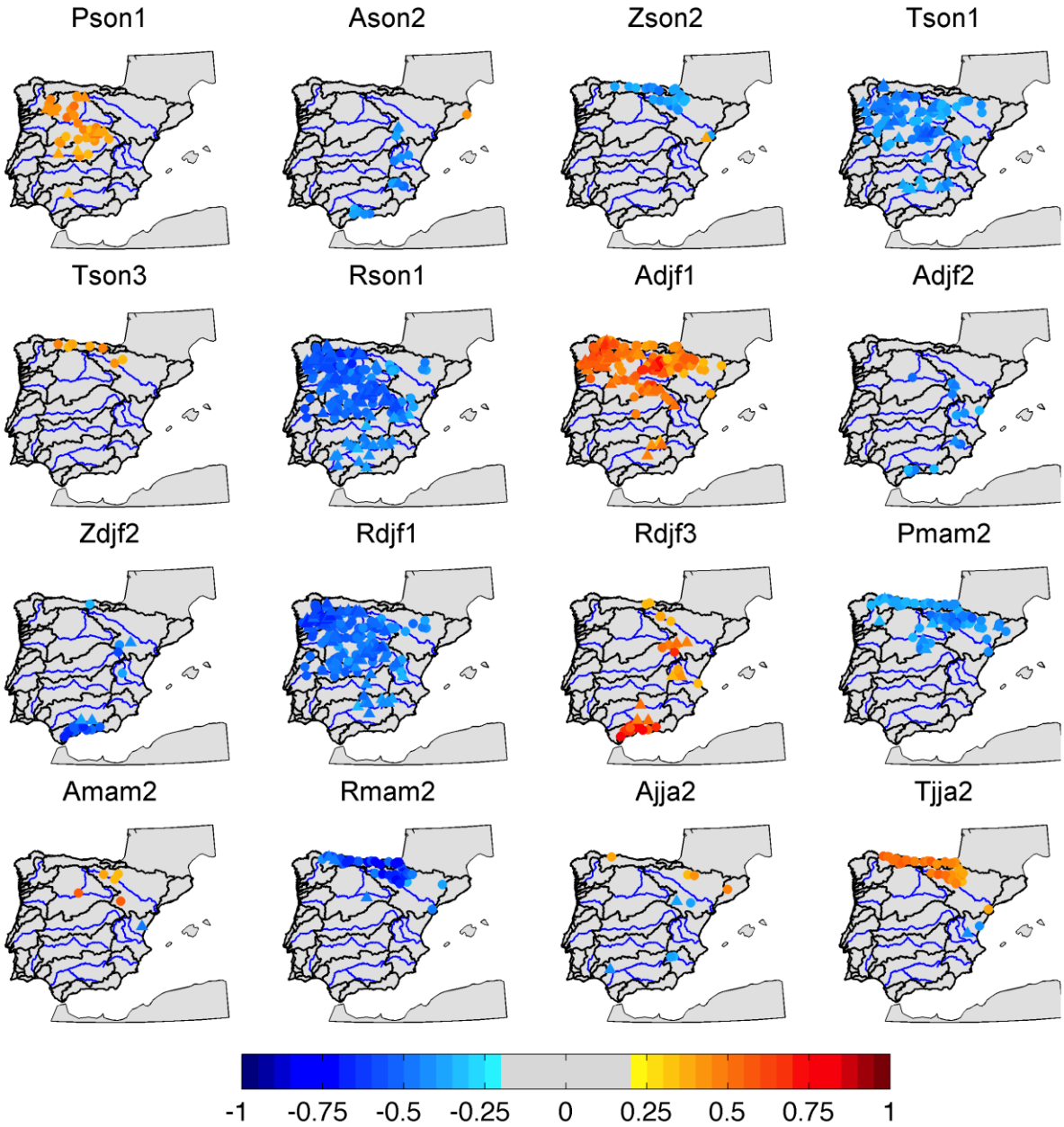


Figure 6.10. Maps of the correlation between the SVD modes selected and autumn (OND) streamflow. Only significant and stable correlations are displayed.

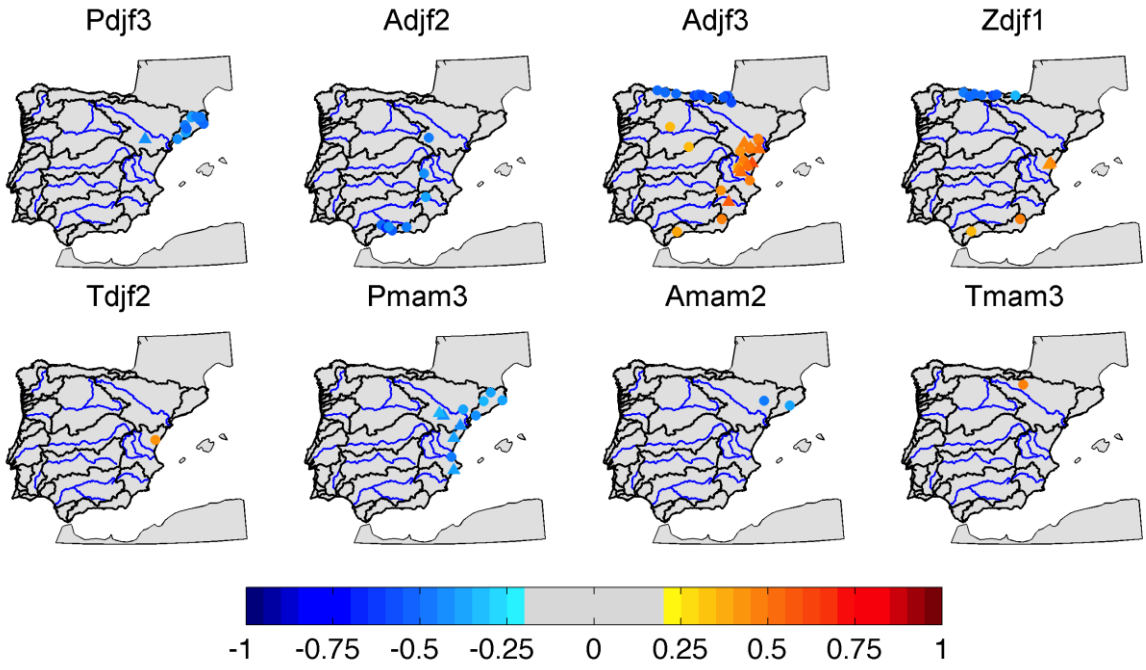


Figure 6.11. As Figure 6.10 but for winter (JFM) streamflow.

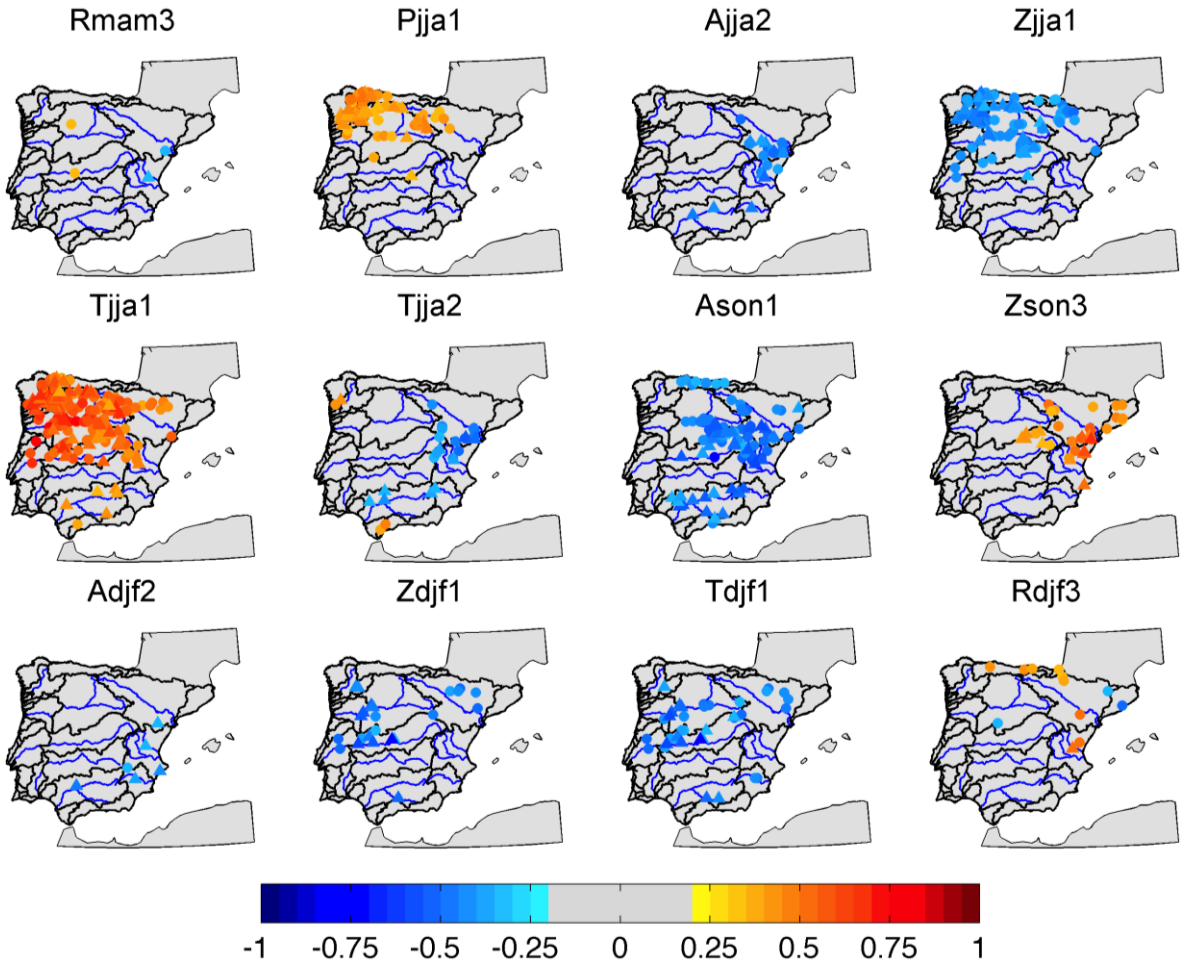


Figure 6.12. As Figure 6.10 but for spring (AMJ) streamflow.

## 6.5 Discussion

A thorough investigation regarding the influence of climatic variables (Atlantic and Pacific sea surface temperature, geopotential height at 500 hPa in the Northern Hemisphere and global temperature and precipitation over land surface) of previous seasons (until four preceding seasons) on seasonal streamflow (autumn, winter and spring) of IP Rivers has been carried out to identify the primary drivers of seasonal streamflow in the IP. In this section, a summary of the main significant coupled modes, their interrelation and some considerations regarding the physical basis behind them are presented.

As it can be noticed from the description of the selected SVD modes, some of them are related to the same atmospheric or oceanic mechanism. In order to quantify this, a cross-correlation analysis between them was conducted. Figures 6.13 to 6.15 display the correlation values between SVD modes corresponding to the coupling with autumn, winter and spring streamflow, respectively. The summary results of this analysis are started herein.

In particular for autumn streamflow, three modes were discarded as its significant levels according to Montecarlo tests dropped below 80% (Zson2, Rdjf3 and Amam2) and another one because of the few number of stations that correlated stably (Ajja2). On the contrary, some modes with significance levels between 80-90% were selected because of they are related to a well-known climatic phenomena (Ason2, Tson3, Adjf2 and Zdjf2). According previous studies, some comments regarding the physical basis behind the statistical links are proposed: Pson1, Rson1 and Rdjf1 modes (left expansion coefficients) are highly correlated to each other, and with El Niño indices. Moreover, they display similar heterogeneous maps (particularly the northwest and centre of the IP). The relationship between ENSO events in autumn and IP Rivers discharge in following autumn (one year lagging response) was also found in Chapter 5 with teleconnection indices. Relationship between the SST anomalies in tropical Pacific related to ENSO and precipitation in the **N**orth Atlantic sector has been observed by several authors (Rodó et al., 1997; van Oldenborgh et al., 2000; Rimbu et al., 2004; Mariotti et al., 2002; Park, 2004; Pozo-Vázquez et al., 2005; Vicente-Serrano, 2005; Zanchettin et al., 2008; García-Serrano et al., 2010; Shaman and Tziperman, 2010). Most of them are related to the autumn and spring precipitation response to an El Niño event. They found positive precipitation anomalies in the Mediterranean area when El Niño phenomenon is presented and negative when La Niña conditions prevail. In this case, the lag in the response is larger (about one year), and positive SST anomalies in autumn/winter in equatorial Pacific are associated with negative streamflow anomalies in north-west quadrant of the IP during the following autumn. This is contrary to the

sign observed between previous autumn and winter ENSO indices with autumn streamflow of the following year with respect to the correlation between summer and autumn ENSO indices and autumn streamflow of the same year. This change in the sign of the correlation was also observed in Chapter 5, using ENSO indices and by Cordoba-Machado et al. (2014) in a study of the relationship between Pacific SST and precipitation in Colombia, and could be related to a 2 years periodicity observed in El Niño (Yu and Kim, 2010).

The Ason2 and Adfj2 modes are also highly correlated between them, with AMO and with winter NAO. The homogenous maps of these modes display the Atlantic SST horseshoe and tripole pattern, respectively. The transition between North Atlantic Horseshose (NAH) in summer/autumn and the North Atlantic tripole (NAT) in winter and its relation to precipitation in North-Atlantic European sector in winter have been considered in Czaja and Frankignoul (1999; 2002), Garcia-Serrano et al. (2008) or Gastineau et al. (2013), among others. These authors indicated the possibility of a positive feedback between the NAT pattern and the NAO, which begins before the established NAO phase. The gap in the covariance between autumn NAH and winter NAO have been also found in previous studies (Czaja and Frankignoul, 1999, 2002; Rodwell and Folland, 2002; Frankignoul and Kestenare, 2005). García-Serrano et al. (2008) also suggested that the damping shown by the NAH through autumn could be understood as a decaying forced by alterations in the Atlantic Hadley cell, and associated with the subtropical warming, via Rossby wave extension from the Amazon convergence zone. In late autumn, the SST anomaly weakly resembles the NAT, and the subtropical forcing leads to a displacement of the Intertropical Convergence Zone (ITCZ) and changes in the upper divergence over northwestern South America. Such anomalous outflow forces another Rossby wave train of opposite polarity from the Caribbean, affecting the trade winds (cyclonic barotropic circulation), which also acts during early winter. Also during early winter the transient eddy activity adds its effects by affecting the airmass exchange between the Azores high and Iceland low through changes in the zonal extension of the subtropical Atlantic jet. Finally, the anomalous circulation seems to be the response of the atmosphere to the subtropical SST anomaly, adding the direct Gill and forced Rossby responses, not via perturbations in the ITCZ. The transient activity acts as a positive feedback inducing a large cyclonic anomaly over Western Europe. They also found that the larger amplitude of NAT through early spring (which could be related to the mode Amam2) is due to atmospheric forcing. The correlation of both modes, Ason2 and Adjf2 with winter NAO (coetaneous with Adjf2) is 0.50 and 0.53, respectively. Once it has been manifested the link between NAH-NAT-NAO, the next step is to understand the basis of the relationship

between them and streamflow in following autumn. On this regard, in case of the stations in the Mediterranean Andalusian Basin, the possible link between winter NAO and following autumn streamflow was pointed out in Chapter 5. Regarding other stations in the Mediterranean slope displayed in the heterogeneous maps of Ason2 and Adf2, in addition to the before mentioned physical link, the possibility of a contribution via regulation should be considered, because the level of regulation in Segura and Jucar Basins is non negligent. Not surprisingly, Adj2 correlates with Zdjf2 and Rdjf3, which are both related to the winter NAO, and also present similar heterogeneous maps, with high correlation for the streamflow of the Mediterranean Andalusian Basin. However, only Zdjf2 has been selected since its level of significance is 88%, while in case of Rdjf3 is only 43%. The heterogeneous maps displayed in both cases look like that found in Chapter 5 for winter NAO and the same possible physical explanation is extended here.

The Adjf1 mode exhibits high negative correlation with Rdjf1 mode and both display similar heterogeneous maps. As it was remarked before, the Rdjf1 presents a homogeneous map similar to El Niño impacts on winter precipitation (see CPC website) and also positive correlation with El Niño indices (around 0.6),. However, Adjf1 does not correlate significantly with ENSO indices (except with El Niño1+2, with a correlation value of -0.37). However, it exists significant correlation between Adjf1 and El Niño indices of following autumn (in particular, 0.46 with El Niño1+2, 0.52 with El Niño3 and 0.47 with El Niño3.4). In addition, the correlation map of El Niño indices in autumn with OND streamflow (not shown) looks like similar to the heterogeneous map of Adjf1. Then, the relationship between Adjf1 and OND streamflow is likely through ENSO, with Atlantic SST acting as precursor of ENSO events. On this regard, Kayano et al. (2013) found, during 1975-200 period, a positive correlation between the first mode of an EOF analysis of tropical Atlantic SST (which features a negative dipole with the negative centre at 15° S, 0ºW and the positive at 37.5°S, 25ºW) and the first mode of an EOF analysis in the tropical Pacific (which reflects mainly the ENSO-related SST variations in the tropical Pacific), when Atlantic SST dipole leads from 1 to 4 seasons. So, a negative (positive) South Atlantic SST dipole precedes by one season to 1 year the establishment of an El Niño (a La Niña) event. The physical explanation given by these authors indicate that once the of negative (positive) South Atlantic dipole persists, the persistent descending (ascending) motion in the equatorial Atlantic related to the cooling (warming) there gradually drives an anomalous east–west circulation in the equatorial–vertical plane with an ascending (descending) branch in the eastern equatorial Pacific. This east–west circulation contributes to weaken (enhance) the surface divergence in the eastern equatorial Pacific, leading to the sinking (shallowing) the

equatorial termocline. Under this persistent condition, an El Niño (a La Niña) is established one season to 1 year later. This agrees with the relation between Adjf1 and OND streamflow via following autumn El Niño indices (i.e. positive values of Adjf1, which maps the negative dipole indicated by Kayano et al., 2013, correlate positive and significantly with El Niño indices in following autumn). Moreover, Kucharski et al. (2014) found significant time-delayed impact of the tropical Atlantic on tropical Pacific sea surface temperatures, leading to an anticorrelation between the tropical Atlantic and the eastern Pacific if the Atlantic is leading by about 10 months. They postulated that this teleconnection is likely to be trigger by an atmospheric bridge and a modification of the Walker circulation, which causes low-level wind anomalies in the central-western Pacific that in turn trigger oceanic Kelvin waves that move to the east and initiate the Bjerknes feedback in the eastern equatorial Pacific (Rodríguez- Fonseca et al., 2009; Ding et al., 2012; Frauen and Dommenget, 2012).

The physical basis supporting the links between the remaining SVD modes (Tson1, Tson3, Pmam2, Rmam2, and Tjja2) and autumn streamflow are not totally understood and require more complex analyses. For example, Pmam2 and Rmam2 also establish a relationship between Pacific SST and autumn streamflow (in northern IP). However, both modes are not clearly associated with the 'canonical El Niño', particularly in case of Pmam2, but with the central equatorial Pacific and the north Pacific mode related to NPGO (which has been linked with the second mode of SST variability in North Pacific, the Victoria Mode, by Di Lorenzo et al., 2010) and reveal a variability with a stronger decadal component. Additionally, it is remarkable a feature that involves two modes (Rmam2 and Tjja2), where both left and right expansion coefficients exhibit a strong shift around 1990 and the heterogeneous maps are also similar (particularly including stations in Cantabrian Basin). At this point, it worth to note that the coetaneous WeMO is the major atmospheric driver of autumn streamflow over Cantabrian slope (not shown here), and this shift around 1990 also appear in the seasonal time series of autumn WeMO. How they could be related to a local atmospheric pattern as WeMO is not clear and need further investigations.

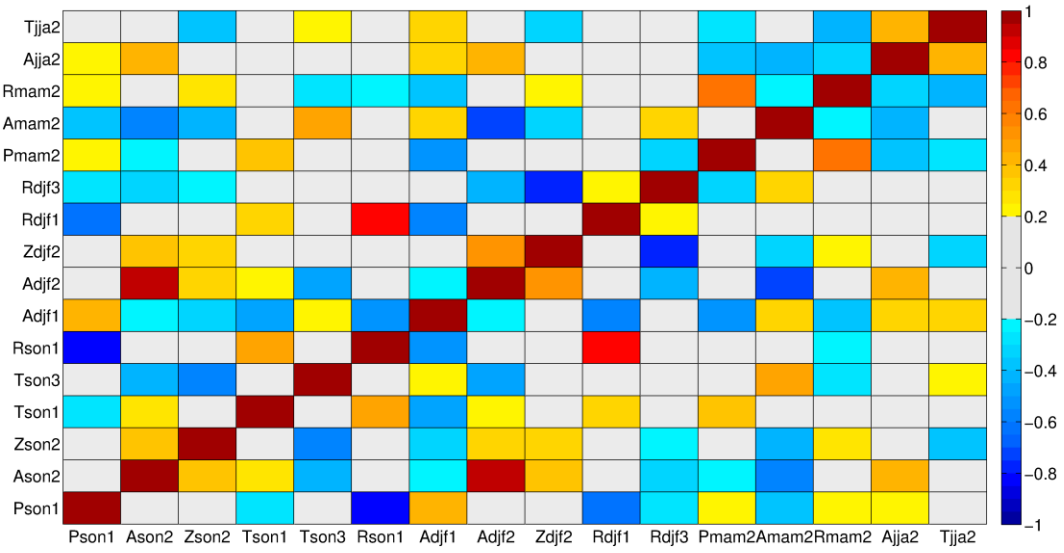


Figure 6.13. Correlations between the expansion coefficients of the SVD modes corresponding to autumn streamflow.

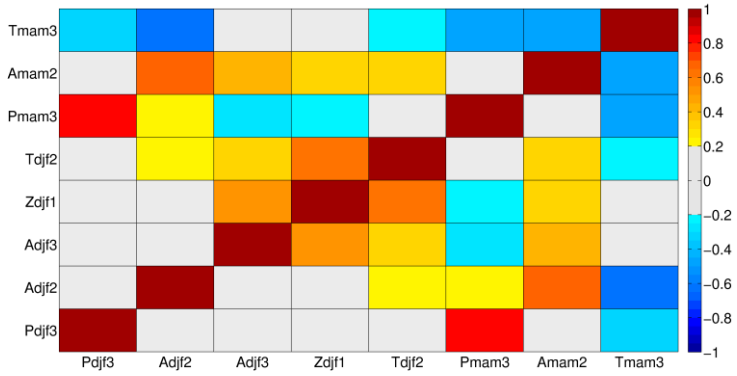


Figure 6.14. As figure 6.13 but for SVD modes corresponding to winter streamflow.

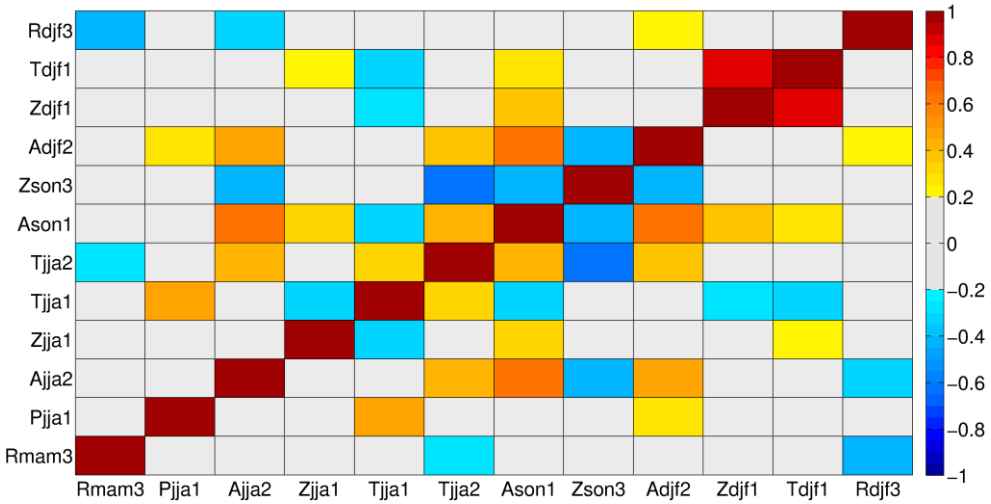


Figure 6.15. As Figure 6.13 but for SVD corresponding to spring streamflow.

Also, two SVD modes have significant *SC* but no significant *COR* (Pjja1 and Tdjf2) so they were not shown studied here, as they lacked of predictability skill. The Pjja1 mode (*MC* = 95% and *COR* = 0.22) linked the variability in Pacific SST in El Niño region with autumn streamflow variability in north-western quadrant of the IP, similarly as Pson1, but with contrary sign. The link between ENSO in summer (initialization of the phenomenon) with following winter streamflow was also found in Chapter 5. In addition, Tdjf2 (*MC* = 94% and *COR* = 0.17) displays a homogeneous pattern similar to the impact of NAO on winter temperature and a heterogeneous map sharing some characteristic with that associated with Zdjf2.

In case of winter streamflow, Figure 6.11 displays the correlation between the SVD modes. Two modes, Zdjf1 and Tmam3 were discarded due to the lack of significance (22% and 66% respectively, according to Montecarlo test). A third mode (Adjf3) was not selected because the homogeneous map does not represent a wide and well-known mode in the Atlantic Ocean. In fact, only a small portion of grid points in the Atlantic SST correlates significantly with this mode. Finally, two modes (Tdjf2 and Amam2) were not considered as they correlated stably with few stations.

Pdjf3 and Pmam3 modes, which are highly correlated to each other (about 0.8), are linked with the same physical mechanism, the anomalies in central Pacific SST (El Niño Modoki) and their influence on north Pacific SST (in particular in the region associated with the NPGO pattern). These modes also have similar heterogeneous maps (showing stations in Mediterranean slope, mainly in lower Ebro and Catalonian Basin). These results agree with that obtained in Chapter 5 where correlations between EMI conditions in winter and following winter streamflow in Internal Catalonian Basin were found.

Adjf2 (which also correlates with Tdjf2 and Amam2) presents significant correlation with winter NAO. Also, Adjf2 presents high correlation values with the Ason2 and Adjf2 modes linked with autumn streamflow (0.86 and 0.94, respectively) and the heterogeneous map of Adjf2 is similar to that found for autumn streamflow, which indicates that the link could be due to an extension of this teleconnection to wintertime.

In case of spring streamflow, from the 12 SVD modes described, four were discarded. Ajja2, Ason1 and Adjf1 were not considered since the significance of these modes was inferior to 80%, and Rmam3 was not accounted for subsequent analysis because it was not well related to any known pattern of variability (teleconnection indices or EOFs) and the lack of stability in correlation with spring streamflow. Regarding the significant SVD modes considered, it can be discussed some characteristic that could help to understand the physical mechanism behind.

Tjja2 and Zson3 modes present strong negative correlation and, as it was previously commented, also correlate significantly with WP pattern. Then, both modes seem to be associated with a similar atmospheric mechanism and display similar heterogeneous maps (especially regarding correlations with stations in lower Ebro, Catalonia, Jucar and upper Tagus Basins). Additionally, Zson3 presents significant correlation with Pacific Transition pattern in September (0.53). This pattern captures anomalous wave-train of 500-hPa heights extending from the central subtropical North Pacific to the eastern United States. Both patterns, Tjja2 and Zson3 exhibit heterogeneous maps similar to Ajj2 and Ason1 (which were not found significant). These two non-significant Atlantic modes are related to the NAH pattern, and correlates significantly with Zson3 and Adjf1 (see Figure 6.12). Furthermore, Adjf2 (significant according to Monte Carlo test) also shows a similar heterogeneous map. This pattern presents a structure similar to NAT in the North Atlantic sector, and correlates significantly with Ason1 and Ajja2 (see all comments related to the NAH-NAT transition given before). Moreover, Adjf2 presents a high correlation with the Adjf2 mode found in the SVD analysis coupling Atlantic SST in winter with autumn and winter streamflow (0.78 and 0.82, respectively). This pattern also presents an SST dipole in south tropical Atlantic. However, how the wave train-like pattern associated with Zson3 and also with Tjja2 is related to spring streamflow, particularly in the Mediterranean sector, and if it has also a link with NAH-NAT pattern, require further investigation.

A high correlation is also found between Zdjf1 and Tdjf1. Both modes are related to a configuration similar to the winter AO/NAO, which is the main driver of winter precipitation over most of the IP (as observed in heterogeneous maps of both modes). The lagging response could be due to streamflow persistence, snowmelt contribution and a persistence of atmospheric winter conditions driving precipitation over IP in early spring. This finding agrees with the correlation analysis using teleconnection indices summarized in Chapter 5.

Pjja1 mode relates tropical Pacific SST (El Niño region) in summer with following spring streamflow variability in the northwestern quadrant of the IP. According homogeneous and heterogeneous maps, positive (negative) anomalies in equatorial Pacific SST are related to positive (negative) spring streamflow anomalies in northwestern IP. This relationship was also identified in Chapter 5 through the correlation with teleconnection indices. Significant precipitation anomalies related to ENSO has been pointed out mostly for winter and spring seasons. Particularly for the IP, Rodó et al. (1997) found a significant connection between winter ENSO events and spring precipitation over the Spain, and Lorenzo et al. (2010) demonstrated that spring precipitation in the northwestern IP is related to ENSO. Although in these studies

refers to winter ENSO, the Pjja1 mode found here refers to summer Pacific SST, being related to the initialization of the phenomenon, which acquires its mature state in winter. In fact, SVD coupled modes between autumn and winter Pacific SST with spring streamflow in IP revealed a similar pattern, but they were not shown since they were not found significant. Bulic and Kucharski (2012) indicated that Rossby wave propagation mechanism (see Trenberth et al., 1998, for a review of mid–latitude teleconnections) is too fast to be responsible for such a delayed impact of winter sea-surface temperature (SST) anomalies in tropical Pacific on spring precipitation over the North Atlantic/European (NAE). Instead they suggest that a chain of air-sea interaction may explain the delayed response. The link between the tropical Pacific and extratropical SST anomalies may be established via the so-called 'atmospheric bridge' (see e.g. Lau and Nath, 1996). In detail, they posit that a positive (negative) ENSO event leads to a quasi-barotropic trough (ridge) in the North Atlantic region. The resulting wind and cloud changes cause anomalies in the surface heat fluxes that result in negative (positive) SST anomalies in the central North Atlantic and anomalies of the opposite sign further to the south. The SST anomalies persist into spring and the atmospheric response to these anomalies is an extension of the ENSO-induced trough (ridge) into the European region, leading to enhanced (reduced) moisture flux and low-level convergence (divergence) and thus positive (negative) precipitation anomalies.

Zjja1 mode is related with the northern annular mode in summer. In fact, it presents high correlation with AO. Furthermore, the heterogeneous map is rather similar to that found when we correlated summer AO with following spring streamflow (Chapter 5). The plausible physical basis behind this relationship is the same that were debated in Chapter 5 for summer AO (see Chapter 5 for further explanation).

Tjja1 and Rdjf3 were considered since they were significant in both terms, Montecarlo and *COR* results, although there are not clearly established relationships with well-know climatic phenomena. For example, in spite of Tjja1 presents significant correlation with El Niño indices and with Pjja1 and also similar heterogeneous pattern, the magnitude of this correlation is not high enough to establish a clear relationship between Tjja1 and ENSO.

To sum up, the SVD technique applied in this chapter has been able to identify climate-hydrology response relationships between global data (sea surface temperature, geopotential height, temperature and precipitation over land) and lagging seasonal streamflow. Among the SVD modes obtained, those considered significant and stably correlated with a sufficient number of stations were the following:

- Linked with OND streamflow: Pson1, Ason2, Tson1, Tson3, Rson1, Adjf1, Adjf2, Zdjf2, Rdjf1, Pmam2, Rmam2 and Tjja2.
- Linked with JFM streamflow: Pdjf3, Adjf2 and Pmam3.
- Linked with AMJ streamflow: Pjja1, Zjja1, Tjja1, Tjja2, Zson3, Adjf2, Tdjf1 and Rdjf3.

The relevance of these results lies not only in helping to clarify the underlying physical mechanisms to explain the variability of streamflow, but also in becoming the basis for streamflow forecasting models. This task is addressed in the next chapter.

# CHAPTER 7

# SEASONAL STREAMFLOW FORECASTING BASED ON LARGE-SCALE CLIMATE VARIABILITY

In this chapter, the ability of teleconnection indices and the significant covariability modes between seasonal streamflow and atmospheric and oceanic variables, identified in Chapters 5 and 6, respectively, as potential predictors of seasonal streamflow of IP Rivers are investigated. A forecasting scheme based on linear regression models is used to elaborate seasonal streamflow predictions with different seasons in advance. The skill of predictions is evaluated through a set of verification scores.

## 7.1 Introduction

In Chapter 5, the ability of main teleconnection indices as potential predictor of seasonal streamflow in the IP was addressed. As a result, those teleconnection indices that showed stable correlation with a remarkable number of the stations were considered as potential predictors. In Chapter 6, the covariability modes between atmospheric and oceanic fields (sea surface temperature –SST–, geopotential height at 500 hPa –Z500–, global temperature –TMP– and precipitation –RR–) and seasonal streamflow of the IP Rivers were analyzed through singular

value decomposition (SVD) method. As in Chapter 5, from all significant modes of covariability found, only those that correlate stably with an important number of stations were considered. In this chapter, it is pretended to evaluate the skill of these indices and modes of covariability in forecasting the seasonal streamflow. With the aim of addressing the individual performance of each set of predictors in forecasting the seasonal streamflow, the evaluation will be carried out for each set of predictors independently and for a combination of both.

Although atmospheric processes are basically nonlinear, linear statistical forecast models can contribute toward understanding the mechanisms associated with large-scale teleconnection patterns of the general atmospheric and ocean circulation (Kung and Sharif, 1980). They are particularly interesting when time series are relatively short (in terms of climatological studies) and finding non-linear relations is challenging. Given the relatively short length of the streamflow records used in this study (33 years), the utilization of linear models was considered the most appropriate for forecasting the seasonal streamflow (autumn –OND–, winter –JFM–, and spring –AMJ–) at each gauging station.

Teleconnection indices have been used as predictors in forecasting streamflow studies in many different areas across the world such as in North America (Tootle and Piechota 2004; Grantz et al., 2005; Timilsena et al., 2009; Soukoup et al., 2009; Oubeidillah et al., 2011; Tang et al., 2011; Kalra et al., 2013), South America (Gutiérrez and Dracup, 2001; Tootle et al., 2008), Europe (Ionita et al., 2008, 2011; Gámiz-Fortis et al., 2011; Oubeidillah et al., 2012), Asia (Chandimala and Zubair, 2007), Africa (Sittichok et al., 2014) or Australia (Piechota et al., 2001; Chiew et al., 2003), using different forecasting methodologies (for example, Artificial Neural Networks, Partial Least Squares Regression, Principal Component Regression or Multiple Linear Regression, among others).

Regarding the SVD technique used in Chapter 6, different ways to predict one field using information of the other coupled field can be found in a literature review. In particular, Wei et al. (2012) applied SVD analysis to reconstruct summer SLP over East Asia from 1470 to 2008 using the summer precipitation in eastern China as predictor. They obtained the SLP field by projecting the first *n* expansion coefficients on their corresponding eigenvectors of the SLP fields. The first *n* expansion coefficients of precipitation were computed using their obtained first *n* eigenvectors and the observed precipitation record Oubeidillah et al. (2011) used the Pacific Ocean SST and Z500 regions identified from SVD analysis with the Upper Colorado River and the Great Basin streamflow and snowpack to create indices (averaging the values in the area represented by these indices) that were then used as predictors in a non-parametric forecasting

model. Anderson et al. (2012) also followed this approach, to determine the region in the Pacific SST linked with the streamflow in the northern portion of the Upper Colorado River Basin. Once a region was identified they created an index by averaging the SST values of this region and used it as predictor.

Other authors, such as Lamb et al. (2010) and Soukup et al. (2009) used the left expansion coefficients obtained from the SVD between Pacific SST and streamflow in the Colorado River Basin and the North Platte River, respectively, as predictor in a non-parametric model to develop continuous exceedance probability forecasts. Also, Tootle et al. (2008) postulated the ability of left expansion coefficients (corresponding to Pacific and Atlantic SST) as predictors of Colombian streamflow. However, Lamb et al. (2010) manifested the question if making forecast using a model that requires the identification of a forecast point is or not adequate. This is the case when left expansion coefficients are calculated using data from the whole period, and they are used as predictors in a cross-validation exercise. This fact was also been argued by García-Serrano and Frankignoul (2014) in a study about the predictability of winter Euro-Atlantic climate form cryospheric variability. They used the expansion coefficients of the left field (determined via MCA using the whole subset) in the one-year-out cross-validated hindcast, so the regression coefficients and predictor value of the statistical model are estimated assuming the knowledge of the MCA fields in the year out. Then, they repeated all calculations performing MCA on the subset of data without the target year, repeating all the procedure with each withdrawn year. The predictability skill obtained under this approach was lower than the previous one, so they concluded that the first methodology is likely to produce an overestimation of the cross-validation skill, because of the a priori knowledge of the forecast point.

Following this line, Wang and Ting (1999) predicted US precipitation for summer of 1997 and winter 1997/98 using tropical Pacific SST through SVD. The prediction algorithm consisted on perform a SVD analysis between the observed tropical Pacific SST and US precipitation with the target year removed. Then, the SST for the target year is projected on the SVD vectors to obtain a value for the expansion coefficient of this field. Using the relation between left and right expansion coefficient given by SVD, the expansion coefficient for precipitation is obtained and, subsequently, the precipitation values. A similar approach was followed by Liu (2003), removing from the original subset each year to be predicted and conducting a SVD between soil moisture leading precipitation in United States. Finally the precipitation values are calculated using soil moisture values of the withdrawn year. Also,

Córdoba-Machado et al. (2014) used a similar technique for forecasting precipitation in Colombia.

Taking into consideration the high number of predictors and predictands available and the limited length of the time series, a multiple linear regression using leave-one-out was adopted for developing forecasting models. When SVD modes were used as predictors, an algorithm was built based on SVD carried out in each iteration of the cross-validation process, predicting the removed year and repeating the process for all years (similar to Wang and Ting, 1999; Liu, 2003 and Córdoba-Machado et al., 2014). In addition, in order to evaluated the overestimation in forecasting skills derived from using as predictors the left expansion coefficient calculated using the entire period, all the calculation were repeated based using this approach.

## 7.2 Methodology

Three separated forecasting exercises have been carried out in this chapter, considering up to three different subsets of predictors (only the significant and stably correlated teleconnection indices identified in Chapter 5, only the significant and stable correlated SVD modes identified in Chapter 6 and a combination of both). The procedure followed for making seasonal predictions consists of the establishment of four forecasting scenarios. Each scenario refers to the seasons in advance for which the predictions are made. For example, in case of spring streamflow forecast based on teleconnection indices, in the first scenario (referred as '4S' hereafter) the predictions are made four seasons in advance, using the teleconnection indices of the previous spring as predictors. The following scenario ('3S') uses as predictors the teleconnection indices from four to three seasons beforehand, i.e. the teleconnection indices of the previous spring and summer. The next scenario ('2S') uses as predictors the teleconnection indices from four to two seasons beforehand, i.e. the seasonal teleconnection indices of the previous spring, summer and autumn. The last scenario ('1S') uses as predictors the teleconnection indices from four to one season beforehand, i.e. the teleconnection indices of the previous spring, summer, autumn, and winter. Then, for each forecasting scenario, seasonal streamflow to be predicted at each gauging station, the predictors and predictands will be separated into calibration and validations subsets, and the calibration subsets will be used in building a forecasting model based on linear regression. Finally, validation subsets will be used to assess the forecasting skills. Before proceeding with a more detailed explanation of the

forecasting methodology, a brief explanation about how the calibration and validation subsets are built is provided.

### 7.2.1 Leave-one-out cross-validation

In order to avoid artificial skill effects in forecasting, data are usually divided into two subperiods: 'calibration' and 'validation'. Then, the model is fitted to the calibration data and tested on the independent information contained in the validation subset. However, in small samples, this procedure could not perform effectively. In these cases, cross-validation is arguably the best approach for estimating true skill, because it makes all data available for validation. In this approach, the first step consists on withholding a small part of the sample. The model is fitted to the data that are retained and is used to make forecasts or specifications of the data that are withheld. These steps are performed separately, either until no new verification data sets can be selected or until there are enough forecast/verification or specification/verification pairs to estimate skill accurately. Therefore, many different (but highly related) models are used to make independent forecasts. According to the relatively small sample size in streamflow records, a leave-one-out cross-validation approach was performed here.

In the procedure followed to generate the 'forecasted' time series, for the seasonal streamflow at each station, each model is formulated based on *N-1* years and tested on the remaining year not used in the construction of the model. Assuming the dataset $y_t$, with $t = 1,...,N$, a particular value $y_i$ is removed from the set and a model is fitted based on the remaining $N - 1$ values. Then, the model is utilized to forecast the $i^{th}$ value; that is, $f_i$ is obtained. This procedure is repeated $N$ times and the forecast dataset $f_1, \ldots, f_N$ is determined. To exclude the possibility that the results of the leave-one-out cross-validation could be biased by linear trends, the predictor/predictand samples used are linearly de-trended and centred to have zero mean. To eliminate a further potential source of dependency/artificial skill, the removal of the trend and the mean is repeated in each step of the cross-validation (von Storch and Zwiers, 1999). That is, the $i^{th}$ forecast is obtained from predictor/predictand samples having zero linear trend and zero mean in any case. The trend and mean values obtained in the $i^{th}$ step of the cross-validation are also removed from the $i^{th}$ withheld predictors and predictand values, respectively. Note that, although the word 'hindcast' could be more appropriated when calibration and validation dataset (via cross-validation) belong to the past, the term 'forecast' is used in this study in relation with the validation subset.

Calibration and validation for predictors from SVD analysis

When the forecasting procedure uses as predictors the left expansion coefficients of the significant SVD modes identified in Chapter 6, the methodology presents some differences with respect to the use of teleconnection indices as predictors. The reason is that SVD analysis was repeated in each one-year-out iteration of the cross-validation exercise. This is mandatory, because when using the entire dataset to conduct the SVD analysis, if you use the left expansion coefficients obtained in a leave-one-out cross-validation exercise (dropping one year, which it will be used in the validation process, and using the remaining years to calibrate the model) the years used for calculating the parameters of the regression models 'knows' about the dropped year (because all years were used in the SVD), so an overestimation of the cross-validation skill is likely to be produced, because a priori knowledge of the forecast point exists. To overcome this problem, for each seasonal streamflow to be predicted, each forecasting scenario and in each leave-one-out cross-validation step, the prediction algorithm developed consists on the following steps:

1. To remove the target year from the streamflow time series and from the atmospheric/oceanic datasets (Pacific and Atlantic SST, Z500, TMP and RR).
2. To identify the significant SVD modes (see Chapter 6) used as predictors according the forecasting scenario, performing a new SVD between the corresponding climatic field and the seasonal streamflow. For each one of the significant SVD modes, to correlate the corresponding eigenvector of this new SVD analysis with that calculated using the whole period to ensure the new and the original mode are the same (without alteration in modes order). To calculate the new left expansion coefficients (predictors) projecting the field values onto the eigenvectors.
3. To calculate the associated value of the left expansion coefficient used for validation by projecting the value of climate field of the dropped year onto the associated eigenvector.

Steps 1-3 are repeated 33 times (for each year being out) until left expansion coefficients values for calibration (33 time series of 32 years) and validation (33 years time series) processes are obtained, for each station, scenario and seasonal streamflow to be predicted.

### 7.2.2 Multiple Linear Regression

Taking into account the large number of predictands and predictors, as well as the limited length of the time series, an approach based on linear relationship (via multiple linear regression)

between predictors and predictands was followed. This linear statistical technique allows to find the best relationship between a variable (dependent, predictant) and several other variables (independent, predictor) through the least square method. Multiple regression models can be presented by the following equation:

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_n X_n \qquad (7.1)$$

Where $Y$ is the predictand, $\beta_i$ are the regression coefficients, estimated as $\beta = (X^t X)^{-1} X^t Y$, and $X_i$ are the predictors.

When a large set of predictors is used to build a regression model, an excessive correlation among explanatory variables (collinearity) can complicate or prevent the identification of an optimal set of explanatory variables for a statistical model. This situation can be found in this work, for example, when various of the teleconnection indices related to ENSO are used as predictors. Hence, it is mandatory to address the possible problems related to collinearity between predictors.

Von Storch and Zwiers (1999) indicated that two or more factors are multicollinear when the corresponding columns in the design matrix point in similar directions in $R^n$, i.e. when they are strongly correlated. The effect of multicollinearity is to make the matrix $X^tX$ nearly uninvertible (the denominator in an ordinary least-squares approach to estimate the regression coefficients), resulting in highly variable parameter estimators and making it difficult to diagnose the factors that are most important in specifying $Y$. Furthermore, multicollinearity increases the standard errors of the coefficients, so that coefficients for some independent variables may be found not to be significantly different from 0, whereas without multicollinearity and with lower standard errors, these same coefficients might have been found to be significant and the researcher may not have come to null findings in the first place.

To answer the question about how to address multicollinearity in hydrologic regression models, Kroll and Song (2013) compared four techniques: ordinary least squares (OLS), OLS with variance inflation factor screening (VIF), principal component regression (PCR), and partial least-squares regression (PLS). They observed the performance of these four techniques for varying sample sizes, correlation coefficients between the explanatory variables, and model error variances consistent with regional hydrologic regression models, concluding that negative effects of multicollinearity are magnified at smaller sample sizes, higher correlations between the variables, and larger model error variances (smaller $R^2$). Also they found that if the true model is known, multicollinearity is present, and the estimation and statistical testing of regression

parameters are of interest, then PCR or PLS should be employed. However, when the model is unknown, or when the interest is solely on model predictions, then it would be recommended that OLS be employed since using more complicated techniques offers no improvement in model performance.

Bearing in mind the recommendations given in Kroll and Song (2013), it was opted by a simple approach to identify collinearity, the variance inflation factor (VIF), and combining its results in a backward stepwise multiple-linear regression to select the most appropriate set of predictors. Stepwise regression has also been used in streamflow reconstruction (Woodhouse et al., 2006; Barnett et al., 2010).

Formally, VIF measures how much the variance of the estimated coefficients is increased over the case of no correlation among the *X* variables. VIF calculations are straightforward and easily comprehensible; the higher the value, the higher the collinearity. A VIF for a single explanatory variable is calculated using the r-squared value of the regression of that variable against all other explanatory variables.

$$\mathrm{VIF}_j = \frac{1}{1 - R_j^2} \tag{7.2}$$

where the VIF for variable *j* is the reciprocal of the inverse of $R_j^2$, coefficient of determination from a regression of predictor *j* on the remaining predictors.

When the variation of predictor *j* is largely explained by a linear combination of the other predictors, $R_j^2$ is close to 1, and the VIF for that predictor is correspondingly large. The inflation is measured relative to a $R_j^2$ of 0 (no collinearity), and a VIF of 1. A VIF is calculated for each explanatory variable and those with high values are removed. The definition of 'high' is somewhat arbitrary, and depends merely on the requirements of the tolerance level. Most practical suggestions put an acceptable tolerance at lower than 0.2 or 0.1 ($R^2$ higher than 0.9 or 0.8) to detect the problem. These levels corresponding to a VIF should be larger than 10 and 5 values (Menard, 1995; Hair et al., 1995; Mason, 1998; Kennedy, 1992). In small samples, less than 50 points, a threshold of 5 is preferable. The solution proposed here is similar to the approach suggested by Zuur et al. (2010). For the original subset of predictors, calculate the VIF. If there are two or more variables with a VIF greater than 5, proceed by removing the variable with the highest VIF, recalculate all VIF values with the new set of variables (VIF values changes after a variable is removed), and again remove the variable with the next highest value if it exceeds the threshold, repeating the process until all VIF values are below the threshold.

After removing the predictors with VIF higher than 5, the next step is to identify which ones of the final set should be used to build the regression model. To do this, we used a stepwise regression approach. Stepwise regression is an iterative semi-automated process of building a model by successively adding or removing variables based on the *F*-statistics of their estimated coefficients. The *F*-statistic tests whether the change in the sum of squared errors by the regression model for the predictor under consideration is significantly large compared to the mean of the squared errors after including the predictor. The greater the value of the *F*-statistic (exceeding a predefined threshold according to a given level of significance), the more significant is the predictor.

The stepwise option lets us either starts with *no* variables in the model and proceed forwards (adding one variable at a time), or start with all potential variables in the model and proceed backwards (removing one variable at a time). In the forward selection, the common criterion is the ratio of the reduction in residual sum of squares caused by the next candidate variable to be considered to the residual mean square from the model including that variable. This criterion can be expressed in terms of a critical "*F*-to-enter", where *F* is the "*F*-test" of the partial sum of squares of the variable being considered. The forward selection terminates when no variable outside the model meets the criterion to enter. In this chapter, it was used the backward-elimination procedure, which operates similarly to the forward-selection procedure. In this case, the stopping rule is the "*F*-test" of the smallest partial sum of squares of the variables remaining in the model. Again, this criterion can be stated in terms of an "*F*-to-stay". Backward elimination terminates when all variables remaining in the model meet the criterion to stay.

### 7.2.3. Forecasting skill evaluation. Verification Scores

Verification measures enable us to evaluate the quality of the forecast, its weakness and strengths. A large number of verification measures have been suggested to this end (Jolliffe and Stephenson, 2003). In this study, a set of verification measures has been used: Pearson's correlation coefficient (called 'RHO' hereafter), the Root Mean Square Error Skill Score (RMSESS) and the Gerrity Skill Score (GSS). The RHO is a measure of the linear relation. The RMSESS provides a measure of error. It is referenced to the climatology mean, so that positive (negative) values of RMSESS indicate better (worse) forecasting skill than climatological mean. The GSS identifies the accuracy in forecasting streamflow that are in the same category as the observations. To the calculation of GSS, three categories were considered: below normal,

normal, and above normal, by the 33rd and 66th terciles. A more detailed description of GSS is given in Appendix A.

## 7.3 Results

The teleconnection indices and SVD modes used as predictors in each forecasting scenario are displayed in Table 7.1 and Table 7.2, respectively. These predictors are called according the abbreviation used as in Chapters 5 and 6. In case of teleconnection indices, the letter in the subindex refer to the season ('d' for winter, 's' for spring, 'j' for summer and 'a' for autumn) and when a '1' is added to these acronyms means the season belongs to the previous year. Note that in this table, the use of 'd' instead 'djf' to refer winter seasons was just considered for reducing the length of the predictor name. The same happens to the other seasons.

| SCENARIOS | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| S4 | S3 | | S2 | | | S1 | | | | |
| $WP_{s1}$ | $WP_{s1}$ | $NAO_d$ | $WP_{s1}$ | $NAO_d$ | $EA_m$ | $WP_{s1}$ | $NAO_d$ | $EA_m$ | $EA_j$ | |
| $Niño3_{s1}$ | $Niño3_{s1}$ | $AO_d$ | $Niño3_{s1}$ | $AO_d$ | $WP_m$ | $Niño3_{s1}$ | $AO_d$ | $WP_m$ | $Niño3.4_j$ | |
| $Niño4_{s1}$ | $Niño4_{s1}$ | $Niño4_d$ | $Niño4_{s1}$ | $Niño4_d$ | $EP\text{-}NP_m$ | $Niño4_{s1}$ | $Niño4_d$ | $EP\text{-}NP_m$ | $SOI_j$ | |
| $Niño3.4_{s1}$ | $Niño3.4_{s1}$ | $WeMO_d$ | $Niño3.4_{s1}$ | $WeMO_d$ | $WeMO_m$ | $Niño3.4_{s1}$ | $WeMO_d$ | $WeMO_m$ | $WeMO_{ja}$ | **OND** |
| $SOI_{s1}$ | $SOI_{s1}$ | | $SOI_{s1}$ | | | $SOI_{s1}$ | | | | |
| $SAI_{s1}$ | $SAI_{s1}$ | | $SAI_{s1}$ | | | $SAI_{s1}$ | | | | |
| $EMI_{d1}$ | $EMI_{d1}$ | $WP_{m1}$ | $EMI_{d1}$ | $WP_{m1}$ | $PNA_{j1}$ | $EMI_{d1}$ | $WP_{m1}$ | $PNA_{j1}$ | $SCAND_{s1}$ | |
| $WeMO_{d1}$ | $WeMO_{d1}$ | $EP\text{-}NP_{m1}$ | $WeMO_{d1}$ | $EP\text{-}NP_{m1}$ | | $WeMO_{d1}$ | $EP\text{-}NP_{m1}$ | | $WeMO_{s1}$ | **JFM** |
| | | | | | | | | | $SAI_{s1}$ | |
| | $EA_{j1}$ | | $EA_{j1}$ | $Niño1+2_{s1}$ | | $EA_{j1}$ | $Niño1+2_{s1}$ | $NAO_d$ | | |
| | $EP\text{-}NP_{j1}$ | | $EP\text{-}NP_{j1}$ | $IOD_{s1}$ | | $EP\text{-}NP_{j1}$ | $IOD_{s1}$ | $SCAND_d$ | | |
| | $PNA_{j1}$ | | $PNA_{j1}$ | $SAI_{s1}$ | | $PNA_{j1}$ | $SAI_{s1}$ | $AO_d$ | | |
| | $SCAND_{j1}$ | | $SCAND_{j1}$ | | | $SCAND_{j1}$ | | $Niño1+2_d$ | | |
| | $AO_{j1}$ | | $AO_{j1}$ | | | $AO_{j1}$ | | | | **AMJ** |
| | $Niño1+2_{j1}$ | | $Niño1+2_{j1}$ | | | $Niño1+2_{j1}$ | | | | |
| | $Niño3_{j1}$ | | $Niño3_{j1}$ | | | $Niño3_{j1}$ | | | | |
| | $Niño3.4_{j1}$ | | $Niño3.4_{j1}$ | | | $Niño3.4_{j1}$ | | | | |
| | $IOD_{j1}$ | | $IOD_{j1}$ | | | $IOD_{j1}$ | | | | |

Table 7.1. Teleconnection indices selected used as predictors according each forecasting scenario each seasonal streamflow predicted.

For SVD modes, the first capital letter indicates the climatic field ('P' correspond to Pacific Ocean SST, 'A' to Atlantic Ocean SST, 'Z' to geopotential height at 500 hPa, 'T' to global temperature, and 'R' to global precipitation), the next three letters mean the season of the

climatic field (son, djf, mam and jja) and with the last number being the mode (ordered descending in SCF values).

| SCENARIOS | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **S4** | **S3** | | **S2** | | | **S1** | | | | |
| Pson1 | Pson1 | Adjf1 | Pson1 | Adjf1 | Pmam2 | Pson1 | Adjf1 | Pmam2 | Tjja2 | **OND** |
| Ason2 | Ason2 | Adjf2 | Ason2 | Adjf2 | Rmam2 | Ason2 | Adjf2 | Rmam2 | | |
| Tson1 | Tson1 | Zdjf2 | Tson1 | Zdjf2 | | Zson2 | Zdjf2 | | | |
| Tson3 | Tson2 | Rdjf1 | Tson2 | Rdjf1 | | Tson1 | Rdjf1 | | | |
| Rson1 | Rson1 | | Rson1 | | | Tson2 | Rdjf3 | | | |
| | | | | | | Rson1 | | | | |
| Pdjf3 | Pdjf3 | Pmam3 | Pdjf3 | Pmam3 | | Pdjf3 | Pmam3 | | | **JFM** |
| Adjf2 | Adjf2 | | Adjf2 | | | Adjf2 | | | | |
| | Pjja1 | | Pjja1 | Zson3 | | Pjja1 | Zson3 | Adjf2 | | **AMJ** |
| | Zjja1 | | Zjja1 | | | Zjja1 | | Tdjf1 | | |
| | Tjja1 | | Tjja1 | | | Tjja1 | | Rdjf3 | | |
| | Tjja2 | | Tjja2 | | | Tjja2 | | | | |

Table 7.2. SVD modes selected in the analysis carried out in Chapter 6 used as predictors according each forecasting scenario and each seasonal streamflow predicted.

In order to summarize the results from the forecasting skill evaluation, a classification was made according the values of the verification scores to evaluate the accuracy of the forecasting. The classes are defined as follow:

- Bad: RHO <= 0.34 or RMSESS <= 0 or GSS <= 0
- Poor: RHO > 0.34 and RMSESS > 0 and GSS > 0
- Fair: RHO > 0.44 and RMSESS > 10 % and GSS > 0.2
- Good: RHO > 0.5 and RMSESS > 20 % and GSS > 0.4

Note that the values 0.34 and 0.39 used as thresholds in the case of RHO score, correspond to the 90% and 95% significance levels in Pearson's correlation coefficient.

### 7.3.1 Prediction based on teleconnection indices

Prior to evaluate the quality of the forecasting results, it is briefly mentioned how the possible problems derived from the multicollinearity were addressed when calibrating the

regression model used to predict the seasonal streamflow at each gauging station. The results from the combined VIF and backward selection iterative procedures were carefully checked, particularly when various of the indices representing the ENSO phenomenon were considered as predictors.

An illustrative example of the iterative process that combines VIF and backward stepwise selection criteria is presented. Let's assume a scheme where it is pretended to forecast the autumn (OND) streamflow in the station id=1734 (Miño-Sil Basin) for the 1S scenario. The stable predictors (from the previous autumn) are:

- $Niño3_{son1}$
- $Niño4_{son1}$
- $Niño3.4_{son1}$
- $SOI_{son1}$

Because of the high cross-correlation values between predictors (see Table 7.3), a problem of multicollinearity could arise.:

| | **$Niño3_{son1}$** | **$Niño4_{son1}$** | **$Niño3.4_{son1}$** | **$SOI_{son1}$** |
|---|---|---|---|---|
| **$Niño3_{son1}$** | **1.000** | 0.767 | 0.964 | -0.796 |
| **$Niño4_{son1}$** | 0.767 | **1.000** | 0.890 | -0.812 |
| **$Niño3.4_{son1}$** | 0.964 | 0.890 | **1.000** | -0.848 |
| **$SOI_{son1}$** | -0.796 | -0.812 | -0.848 | **1.000** |

Table 7.3. Cross-correlation between predictors for the autumn streamflow in the station id=1734 for the 1S forecating scenario.

To evaluate this, the VIF was calculated:

- VIF ($Niño3_{son1}$) = 32.724
- VIF ($Niño4_{son1}$) = 11.538
- VIF ($Niño3.4_{son1}$) = 65.570
- VIF ($SOI_{son1}$) = 3.773

As a threshold of 5 has been chosen for VIF, there are 3 predictors with VIF>5. Then, the predictor with the largest VIF value (Niño3.4) was removed and the calculation of cross correlation matrix was repeated. The new VIF values were:

- VIF ($Niño3_{son1}$) = 3.092
- VIF ($Niño4_{son1}$) = 3.326
- VIF ($SOI_{son1}$) = 3.731

In this case all the predictors have a VIF<5, so, according with the established threshold for VIF, it is concluded that the level of collinearity between predictors is not strong enough to notably affect the regression parameters calculation.

Subsequently the backward stepwise procedure was carried out to select the best subsets of predictors according to the 'F-criterion' introduced in the methodology section. In this case, for all iteration (for all the years been withdrawn) in the leave-one-out cross-validation exercise, the backward selection only the Niño3 predictor was retained.

A summary of this procedure is shown in Tables 7.4 to 7.6 (for autumn, winter and spring streamflow respectively), for all forecasting scenarios, which show the number of stations where predictors (teleconnection indices) present stable correlations, stable correlation and overcome the VIF criterion and present stable correlation, overcome the VIF criterion and also the backward stepwise selection. For example, for autumn streamflow forecasting in 4S scenario (Table 7.4), after applying the VIF criterion, $Niño3.4_{s1}$, in 4S scenario, was considered as stable predictor in 11 stations (originally it showed stable correlation with 29 stations), and in one of them, it did not passed the stepwise backward selection. Also, it could be seen that something similar happens with other ENSO related indices (such as $Niño3_{s1}$ and $Niño4_{s1}$). Note that in these tables the subindices indicating the seasons present a reduced format ('s' instead 'son', 'd' instead 'djf', 'm' instead 'mam' and 'j' instead 'jja'). In case of the winter streamflow forecasting (Table 7.5), the most notable feature regarding the predictors selection is that $PNA_{j1}$ was not selected as predictor in any station after combining the VIF and backward stepwise criteria (originally it correlates stably with 13 stations). As it happened with the predictor selection for the autumn streamflow forecasting, the ENSO related indices are the most affected by these criteria in case of the spring streamflow (Table 7.6).

| | 4S | 3S | 2S | 1S |
|---|---|---|---|---|

| | stable | VIF | stepwise | stable | VIF | stepwise | stable | VIF | stepwise | stable | VIF | stepwise |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $WP_{s1}$ | 14 | 14 | 14 | 14 | 14 | 14 | 14 | 13 | 11 | 14 | 13 | 11 |
| $Niño3_{s1}$ | 20 | 18 | 14 | 20 | 12 | 11 | 20 | 14 | 9 | 20 | 14 | 9 |
| $Niño4_{s1}$ | 45 | 45 | 19 | 45 | 27 | 12 | 45 | 31 | 16 | 45 | 31 | 15 |
| $Niño3.4_{s1}$ | 29 | 11 | 10 | 29 | 13 | 9 | 29 | 9 | 6 | 29 | 9 | 6 |
| $SOI_{s1}$ | 41 | 41 | 30 | 41 | 40 | 29 | 41 | 35 | 28 | 41 | 35 | 29 |
| $SAI_{s1}$ | 21 | 21 | 21 | 21 | 21 | 14 | 21 | 21 | 13 | 21 | 21 | 13 |
| $NAO_{d}$ | - | - | - | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 | 15 |
| $AO_{d}$ | - | - | - | 20 | 20 | 7 | 20 | 20 | 4 | 20 | 20 | 5 |
| $Niño4_{d}$ | - | - | - | 22 | 22 | 15 | 22 | 18 | 15 | 22 | 18 | 14 |
| $WeMO_{d}$ | - | - | - | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| $EA_{m}$ | - | - | - | - | - | - | 21 | 21 | 18 | 21 | 21 | 18 |
| $WP_{m}$ | - | - | - | - | - | - | 15 | 13 | 10 | 15 | 13 | 10 |
| $EP\text{-}NP_{m}$ | - | - | - | - | - | - | 74 | 71 | 69 | 74 | 71 | 70 |
| $WeMO_{m}$ | - | - | - | - | - | - | 47 | 45 | 38 | 47 | 45 | 34 |
| $EA_{j}$ | - | - | - | - | - | - | - | - | - | 26 | 26 | 18 |
| $Niño3.4_{j}$ | - | - | - | - | - | - | - | - | - | 16 | 16 | 10 |
| $SOI_{j}$ | - | - | - | - | - | - | - | - | - | 16 | 16 | 12 |
| $WeMO_{ja}$ | - | - | - | - | - | - | - | - | - | 17 | 17 | 17 |

Table 7.4. Number of stations where the predictors (teleconnection indices) of the autumn streamflow present stable correlations and overcome the VIF criterion and the backward stepwise selection for the four forecasting scenarios.

| | 4S | | | 3S | | | 2S | | | 1S | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | stable | VIF | stepwise | stable | VIF | stepwise | stable | VIF | stepwise | stable | VIF | stepwise |
| $EMI_{d1}$ | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| $WeMO_{d1}$ | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| $WP_{m1}$ | - | - | - | 67 | 67 | 66 | 67 | 67 | 65 | 67 | 67 | 62 |
| $EP\text{-}NP_{m1}$ | - | - | - | 18 | 18 | 16 | 18 | 18 | 16 | 18 | 18 | 18 |
| $PNA_{j1}$ | - | - | - | - | - | - | 13 | 13 | 11 | 13 | 13 | 0 |
| $SCAND_{s1}$ | - | - | - | - | - | - | - | - | - | 45 | 45 | 12 |
| $WeMO_{s1}$ | - | - | - | - | - | - | - | - | - | 9 | 9 | 9 |
| $SAI_{s1}$ | - | - | - | - | - | - | - | - | - | 238 | 238 | 238 |

7.5. As Table 7.4 but for winter streamflow.

| | 4S | | | 3S | | | 2S | | | 1S | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **stable** | **VIF** | **stepwise** | **stable** | **VIF** | **stepwise** | **stable** | **VIF** | **stepwise** | **Stable** | **VIF** | **stepwise** |
| $EA_{j1}$ | - | - | - | 20 | 20 | 20 | 20 | 20 | 20 | 8 | 8 | 7 |
| $EP\text{-}NP_{j1}$ | - | - | - | 33 | 33 | 32 | 33 | 33 | 31 | 20 | 20 | 20 |
| $PNA_{j1}$ | - | - | - | 9 | 9 | 9 | 9 | 9 | 9 | 33 | 33 | 31 |
| $SCAND_{j1}$ | - | - | - | 53 | 53 | 53 | 53 | 53 | 53 | 9 | 9 | 8 |
| $AO_{j1}$ | - | - | - | 37 | 37 | 31 | 37 | 37 | 25 | 53 | 53 | 53 |
| $Niño1+2_{j1}$ | - | - | - | 33 | 31 | 11 | 33 | 31 | 7 | 37 | 37 | 27 |
| $Niño3_{j1}$ | - | - | - | 11 | 11 | 9 | 11 | 11 | 7 | 33 | 31 | 7 |
| $Niño3.4_{j1}$ | - | - | - | 10 | 10 | 9 | 10 | 10 | 0 | 11 | 11 | 6 |
| $IOD_{j1}$ | - | - | - | - | - | - | 54 | 45 | 22 | 10 | 10 | 0 |
| $Niño1+2_{s1}$ | - | - | - | - | - | - | 27 | 27 | 25 | 54 | 33 | 16 |
| $IOD_{s1}$ | - | - | - | - | - | - | 7 | 7 | 7 | 27 | 27 | 25 |
| $SAI_{s1}$ | - | - | - | - | - | - | - | - | - | 7 | 7 | 4 |
| $NAO_{d}$ | - | - | - | - | - | - | - | - | - | 17 | 17 | 16 |
| $SCAND_{d}$ | - | - | - | - | - | - | - | - | - | 18 | 18 | 14 |
| $AO_{d}$ | - | - | - | - | - | - | - | - | - | 10 | 10 | 7 |
| $Niño1+2_{d}$ | - | - | - | - | - | - | - | - | - | 28 | 23 | 5 |

Table 7.6. As Table 7.4 but for spring streamflow.

Once selected the best subsets of predictor and avoided the inconvenient related to highly correlated predictors, the results obtained according with the established classification of forecasting skill based on the verification scores (bad, poor, fair and good forecasting skill) are summarized in Table 7.7 and Figure 7.1.

In case of the autumn streamflow forecasting, only 24 of the 96 predicted time series present some forecasting skill for the 4S scenario. Some stations located in the Miño-Sil, Douro, and Tagus Basins are poorly forecasted. In this case, the indices related to ENSO phenomenon in the previous autumn seem to be the main predictors used. Conversely, some stations located in the Mediterranean Andalusian Basin show fair or moderate forecasting skill, being the SAI index of previous October the predictor used for forecasting autumn streamflow at these stations. Regarding the 3S scenario, 30 of the 115 stations where the autumn streamflow was forecasted present accurate results. The most noteworthy feature is the improvement in forecasting at stations located in the Mediterranean Andalusian Basin, where most stations present fair and good forecasting results. For these stations, NAO and AO indices of previous winter are the predictors related to this improved forecasting skill. For 2S scenario, forecasting skill somewhat

improve (up to 63 stations present some forecasting skill, whereas 127 stations do not), particularly in the central and northwestern quadrant and the Ebro Basin. The increase in forecasting skill could be related to the inclusion of spring EP-NP and WeMO patterns. Forecasting skill does not vary notably in 1S scenario (79 stations present some forecasting skill, whereas 140 stations give bad forecasting results).

Winter streamflow is forecasted only in 11 stations in the 4S scenario, being skillful in 6 of them. In the 3S scenario, when the spring WP and EP-NP patterns were added to the winter EMI and WeMO as predictors, the winter streamflow is predicted in 90 stations. However, these predictors were considered skillful in only 17 of them (mostly poorly forecasted). In addition, there is no remarkable improvement in the forecasting results when summer PNA was added as the predictor for the 2S scenario. However, a substantial increase in forecasting skill is achieved in the 1S scenario. In this case, winter streamflow was forecasted at 269 stations, being skillful in 236 of them. In particular, at 166 stations the forecasting is considered fair or good, with RHO reaching values above 0.6, RMSESS up to 30% and GSS above 0.6 for some cases. The SAI is the predictor responsible for this noteworthy improvement in the forecasting results. The stations that present some forecasting skill are located in most of the IP, with the exception of the Cantabrian range and some areas in the Mediterranean slope.

In case of spring streamflow forecasting, there was no predictor for the 4S scenario. For 3S scenario, up to 9 predictors are used (from teleconnection indices in the previous summer). In this case, spring streamflow is predicted in 150 stations, but only in 52 of them with some skill (mostly poor). The stations where some forecasting skill is found are located mainly in the Miño-Sil, mountains surrounding the Douro Basin borders and Guadalquivir Basin. When predictors corresponding to the previous autumn are added to the forecasting model (2S scenario), there is no improvement in the forecasting results. In fact, the number of stations with skillful forecasted spring streamflow is slightly inferior (probably because the verification scores are low enough to fall below the classification thresholds). In case of the 1S scenario, the use of winter NAO, SCAND, and AO indices improve the forecasting results (up to 58 stations skillfully forecasted), particularly in the southern of IP.

| | *4S* | | | | *3S* | | | | *2S* | | | | *1S* | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | bad | Poor | Fair | good | bad | Poor | Fair | good | bad | Poor | Fair | good | bad | Poor | Fair | good |
| **OND** | 72 | 17 | 7 | 0 | 85 | 16 | 7 | 7 | 127 | 31 | 24 | 8 | 140 | 43 | 25 | 11 |
| **JFM** | 5 | 5 | 1 | 0 | 73 | 14 | 3 | 0 | 79 | 16 | 3 | 0 | 33 | 70 | 127 | 39 |
| **AMJ** | 0 | 0 | 0 | 0 | 98 | 46 | 6 | 0 | 125 | 37 | 7 | 0 | 133 | 43 | 15 | 1 |

Table 7.7. Number of stations belonging to each class defined for summarizing forecasting skill (bad, poor, fair and good). In rows the four scenarios considered and in columns the seasonal streamflow forecasted (autumn -OND-, winter -JFM-, and spring -AMJ-).

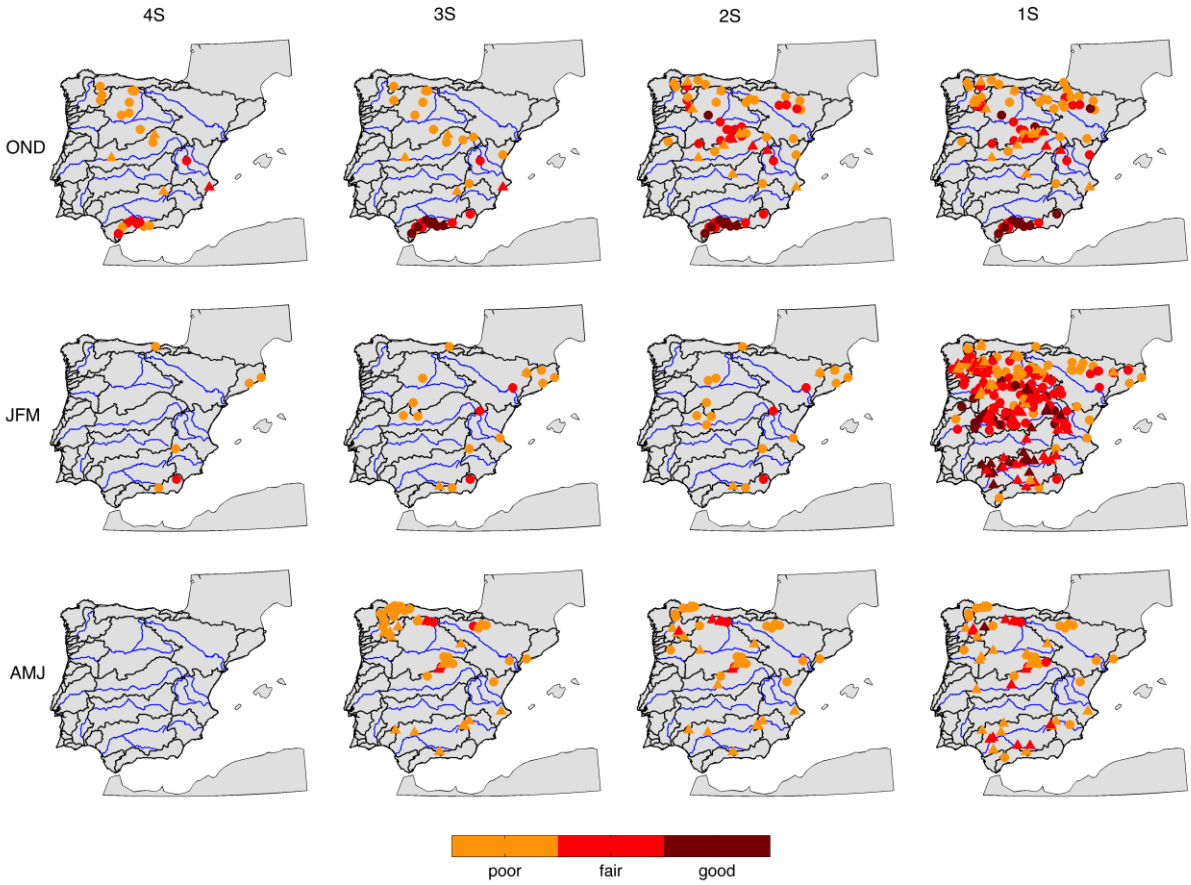


Figure 7.1. Maps of the forecasting skill classification (poor, fair or good) according with RHO, RMSESS and GSS values. They are displayed only in stations where some forecast skills were found. In rows the seasonal streamflow forecasted and in columns the forecasting scenario.

An example of the forecasting for seasonal time series (standardized) is shown in Figure 7.2, displaying the results for one station in each season (OND streamflow in station id=10020, JFM streamflow in station id=3144 and AMJ streamflow in station id=1710, where ‘id‘ indicates

the identifier of each station). The stations were selected among the more skilfully forecasted with the consideration to be located in different basins. The values of the verification scores are given in Table 7.8.

In Figure 7.2a, station id=10020 is used as an example of autumn streamflow forecasting in the 3S scenario. This station is located in the Mediterranean Andalusian Basin (Figure 7.2a, left panel). The predictor used (after backward stepwise selection) was the previous winter NAO. Figure 7.2a, central panel, shows the good agreement between the observed and forecasted autumn streamflow (RHO, RMSESS and GSS are 0.67, 21.1% and 0.45, respectively). That is also manifested in the right panel, where the time series of both observed and forecasted streamflow are plotted. In particular, the similarity in the 1970s and 1980s is noteworthy. On the other hand, the variability in the 1990s is not completely reproduced. Figure 7.2b presents the forecasting results for winter streamflow in station id=3144 in the Tagus Basin (left panel) for 1S forecasting scenario. In this case, October SAI and spring WP were the two predictors selected in the backward stepwise procedure (the SAI being the most correlated predictor). In this case, verification scores are even higher than in the previous example (RHO, RMSESS and GSS are 0.72, 29.8% and 0.66, respectively), as reflected in the central and right panels, where the good agreement between observed and forecasted time series is evident. Figure 7.2c shows an example of spring streamflow forecasting in scenario 1S. The station selected is located in the Miño-Sil Basin (left panel). The verification scores reach values up to 0.63, 20.6%, and 0.43 for RHO, RMSESS and GSS, respectively. The IOD of the previous autumn and the SCAND of previous winter constitute the predictors selected after the combination of VIF and backward stepwise selection. From the observed and forecasted time series displayed in the right panel, the forecasting skill improves after the 1990s, whereas the variability of the forecasted time series during the 1980s is significantly lower than the observed streamflow.

| Season | Station Id | Basin | Scenario | RHO | RMSESS | GSS | CLASS |
|---|---|---|---|---|---|---|---|
| **OND** | 10020 | Mediterranean Andalusian | 3S | 0.67 | 21.1 % | 0.45 | Good |
| **JFM** | 3144 | Tagus | 1S | 0.72 | 29.8 % | 0.66 | Good |
| **AMJ** | 1710 | Miño-Sil | 1S | 0.63 | 20.6 % | 0.43 | Good |

Table 7.8. Values of the verification scores (RHO, RMSESS and GSS) for the three stations selected as examples in Figure 7.2.

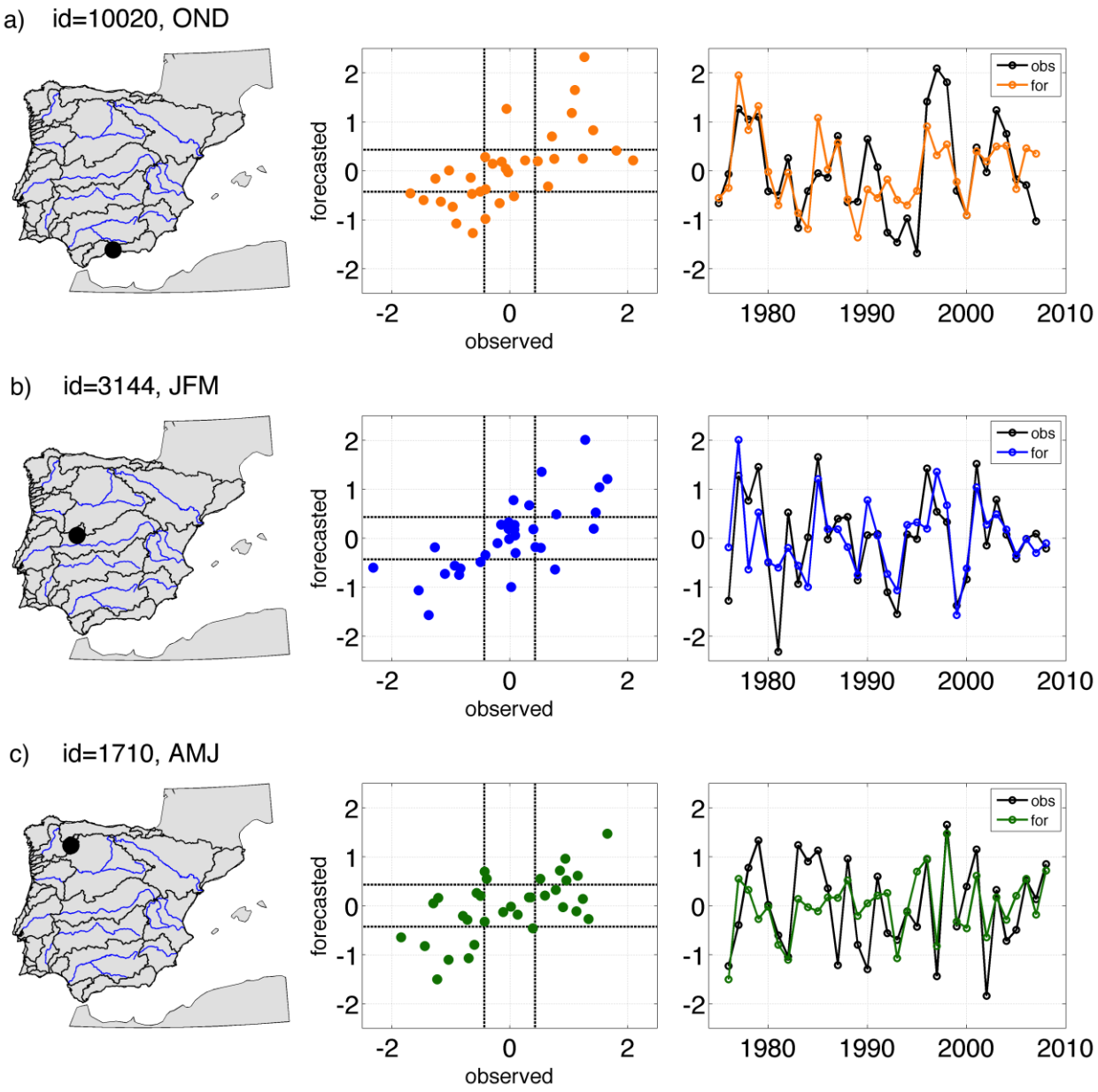


Figure 7.2. In rows, an example of forecasted seasonal streamflow time series for a gauging station in each season. a) For the autumn streamflow in station id=10020 (Mediterranean Andalusian Basin), b) for the winter streamflow in station id=3144 (Tagus Basin) and c) for the spring streamflow in station id=1710 (Miño-Sil Basin). In left panels, the location of these stations; in middle panels, scatter plots between the observed and forecasted seasonal streamflow in each station (black dashed line indicate the 33$^{rd}$ and 66$^{th}$ percentiles); and in right panels, the observed (black line) and forecasted (coloured lines) streamflow time series in each forecasting scenario.

### 7.3.2 Prediction based on SVD

Similarly to the previous section (prediction based on teleconnection indices) a summary of the main results of the VIF and backward stepwise selection steps is shown in Tables 7.9 to

7.11 (for autumn, winter and spring streamflow respectively), for all the forecasting scenarios, with the number of stations where the SVD modes selected as predictors overcome the VIF criterion and those where the predictor was selected after the backward stepwise selection.

For autumn streamflow forecasting (Table 7.9), the predictors most affected by the VIF criterion are Ason2 and Rson1 from 3S scenario. The reason is the relatively high correlation between Ason2 and Adjf2 (related to the horseshoe-tripole pattern) on one hand, Rson1, Pson1 and Rdjf1 (associated with the ENSO phenomenon). For most of SVD modes (excepts, perhaps, Zdjf2 and Rmam2), differences arise after applying the backward stepwise selection between the number of stations stably correlated and the number of stations that they are predictors for (for example, in the 3S scenario, Rson1 correlated stably with 200 stations, but it was finally used as predictor in 89 of them). In case of winter streamflow (Table 7.10), there are only three predictors, being Adjf2 the one that is not selected in most of stations as predictor after VIF and backward stepwise selection are applied. Also, for spring streamflow (Table 7.11), the most remarkable difference is observed for Pjja1, particularly due to backward stepwise selection.

| | **4S** | | | **3S** | | | **2S** | | | **1S** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **stable** | **VIF** | **stepwise** | **stable** | **VIF** | **stepwise** | **stable** | **VIF** | **stepwise** | **stable** | **VIF** | **stepwise** |
| Pson1 | 38 | 38 | 4 | 38 | 38 | 5 | 38 | 38 | 5 | 38 | 38 | 5 |
| Ason2 | 25 | 25 | 20 | 25 | 15 | 6 | 25 | 15 | 6 | 25 | 17 | 7 |
| Tson1 | 110 | 110 | 61 | 110 | 110 | 61 | 110 | 110 | 59 | 110 | 110 | 59 |
| Tson3 | 9 | 9 | 7 | 9 | 9 | 6 | 9 | 9 | 4 | 9 | 9 | 4 |
| Rson1 | 200 | 200 | 193 | 200 | 166 | 89 | 200 | 166 | 90 | 200 | 166 | 91 |
| Adjf1 | - | - | - | 150 | 150 | 110 | 150 | 150 | 89 | 150 | 150 | 87 |
| Adjf2 | - | - | - | 20 | 15 | 13 | 20 | 15 | 13 | 20 | 13 | 11 |
| Zdjf2 | - | - | - | 22 | 22 | 21 | 22 | 22 | 21 | 22 | 22 | 21 |
| Rdjf1 | - | - | - | 185 | 185 | 112 | 185 | 185 | 113 | 185 | 185 | 113 |
| Pmam2 | - | - | - | - | - | - | 81 | 81 | 28 | 81 | 81 | 28 |
| Rmam2 | - | - | - | - | - | - | 46 | 46 | 42 | 46 | 46 | 41 |
| Tjja2 | - | - | - | - | - | - | - | - | - | 42 | 42 | 24 |

Table 7.9 Number of stations where predictors (SVD modes) of autumn streamflow present stable correlations and overcome the VIF criterion and the backward stepwise selection for the four forecasting scenarios.

| | 4S | | | 3S | | | 2S | | | 1S | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | stable | VIF | stepwise | stable | VIF | stepwise | stable | VIF | stepwise | stable | VIF | stepwise |
| Pdjf3 | 12 | 12 | 11 | 12 | 12 | 11 | 12 | 12 | 11 | 12 | 12 | 11 |
| Adjf2 | 25 | 25 | 9 | 25 | 25 | 9 | 25 | 25 | 9 | 25 | 25 | 9 |
| Pmam3 | - | - | - | 11 | 11 | 6 | 11 | 11 | 6 | 11 | 11 | 6 |

Table 7.10 As Table 7.9 but for winter streamflow.

| | 4S | | | 3S | | | 2S | | | 1S | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | stable | VIF | stepwise | stable | VIF | stepwise | stable | VIF | stepwise | stable | VIF | stepwise |
| Pjja1 | - | - | - | 78 | 78 | 36 | 78 | 78 | 36 | 78 | 78 | 36 |
| Zjja1 | - | - | - | 104 | 104 | 85 | 104 | 104 | 86 | 104 | 104 | 86 |
| Tjja1 | - | - | - | 189 | 189 | 181 | 189 | 189 | 181 | 189 | 189 | 181 |
| Tjja2 | - | - | - | 27 | 27 | 26 | 27 | 27 | 21 | 27 | 27 | 21 |
| Zson3 | - | - | - | - | - | - | 46 | 46 | 34 | 46 | 46 | 34 |
| Adjf2 | - | - | - | - | - | - | - | - | - | 7 | 7 | 3 |
| Tdjf1 | - | - | - | - | - | - | - | - | - | 41 | 41 | 32 |
| Rdjf3 | - | - | - | - | - | - | - | - | - | 14 | 14 | 11 |

Table 7.11. As Table 7.9 but for spring streamflow.

Table 7.12 shows the number of stations belonging to each class for each seasonal streamflow and forecasting scenario. Figure 7.3 displays the stations where poor, fair or good forecasting is found, for each seasonal streamflow and forecasting scenario.

According to Table 7.12, autumn streamflow is forecasted for 234 stations in the 4S scenario, presenting some forecasting skill in 43 of them (classified as poor). From Figure 7.3, most of the stations are in Miño-Sil, Douro, and Tagus Basins, particularly in upper part of the basins, being Rson1 and Tson1 the main predictors accounting for these results. Interesting results are found in the scenario 3S, where it is found an interesting balance between quality of the forecast and seasons in advance that it is made. It is observed an increase in the number of forecasted stations (299), with forecasting skills found in 109 of them (in 67 stations is classified as poor skill, in 38 of them as fair and in 4 as good). The stations with reasonable skillful predictions are located in Miño-Sil, Cantabrian, upper Douro, Tagus and Ebro Basins and Mediterranean Andalusian Basin. In particular, 14 of the 15 stations belonging to the Mediterranean Andalusian Basin are forecasted in this scenario, with a fair or good forecasting

results in 13 of them. In this case, Zdjf2 is found the main predictor for autumn streamflow (which is related to the winter AO/NAO phenomenon). Also, the incorporation of Adjf1 and Rdjf1 is translated in an increase of the number of stations where forecasting results are slightly skillful in the northern corner of the IP. Although four predictors more are incorporated when forecasting for scenarios 2S and 1S, slightly modifications in forecasting results are found (only stations in the Cantabrian Basin present more skillful predictions).

As winter streamflow forecasting concerns, the main disadvantage is the reduced number of stable predictors found. Only three of the six SVD modes identified (Pdjf3, Adjf2 and Pmam3) can be used in forecasting winter streamflow. This circumstance limits notably the forecasting experiment. The only profit is that this forecasting can be made with three seasons in advance (these three predictors are linked to the climatic information of previous winter and spring). In fact, from Figure 7.3 and Table 7.1, it could be seen that winter streamflow is predicted in only 28 stations in 3S scenario (being these predictions skillful in 7 of them). These stations are mainly located in Catalonian and Mediterranean Andalusian Basins

Regarding spring streamflow forecasting, the most noteworthy features are, on one hand, that there is not any predictor for 4S scenario and, on the other hand, that the quality of number of stations forecasted increase as lag between predictors and predictands decrease. In the 3S scenario, a total of 230 stations are forecasted, with some skill identified in 77 of them (56 classified as poor and 21 as fair). These skillfully forecasted stations are mainly located in the northern half of IP, in particular in the Miño-Sil and Douro Basins. Pjja1, Tjja1 and Zjja1 are the predictors used in these stations. In the scenario 1S (Zson3, Adjf2, Tdjf1 and Rdjf3 are added as predictors), the number of stations forecasted increase up to 267 (being skillful in 100 of them, 68 poorly, 31 fairly and 1 good forecasted). This increase is more remarkable in the Mediterranean slope and in the Tagus Basin, and it is due to the incorporation of Tdjf1 (close related to each other and with NAO phenomenon) as predictor.

| | **4S** | | | | **3S** | | | | **2S** | | | | **1S** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Bad | Poor | Fair | Good | Bad | Poor | Fair | Good | Bad | Poor | Fair | Good | Bad | Poor | Fair | Good |
| **OND** | 191 | 43 | 0 | 0 | 190 | 67 | 38 | 4 | 189 | 77 | 35 | 4 | 189 | 76 | 40 | 4 |
| **JFM** | 14 | 4 | 3 | 0 | 21 | 4 | 3 | 0 | 21 | 4 | 3 | 0 | 21 | 4 | 3 | 0 |
| **AMJ** | 0 | 0 | 0 | 0 | 153 | 56 | 21 | 0 | 160 | 62 | 22 | 0 | 167 | 68 | 31 | 1 |

Table 7.12. Number of stations belonging to each class defined for summarizing forecasting skill (bad, poor, fair and good). In rows the four scenarios considered and in columns the seasonal streamflow forecasted (autumn -OND-, winter -JFM- and spring -AMJ-).

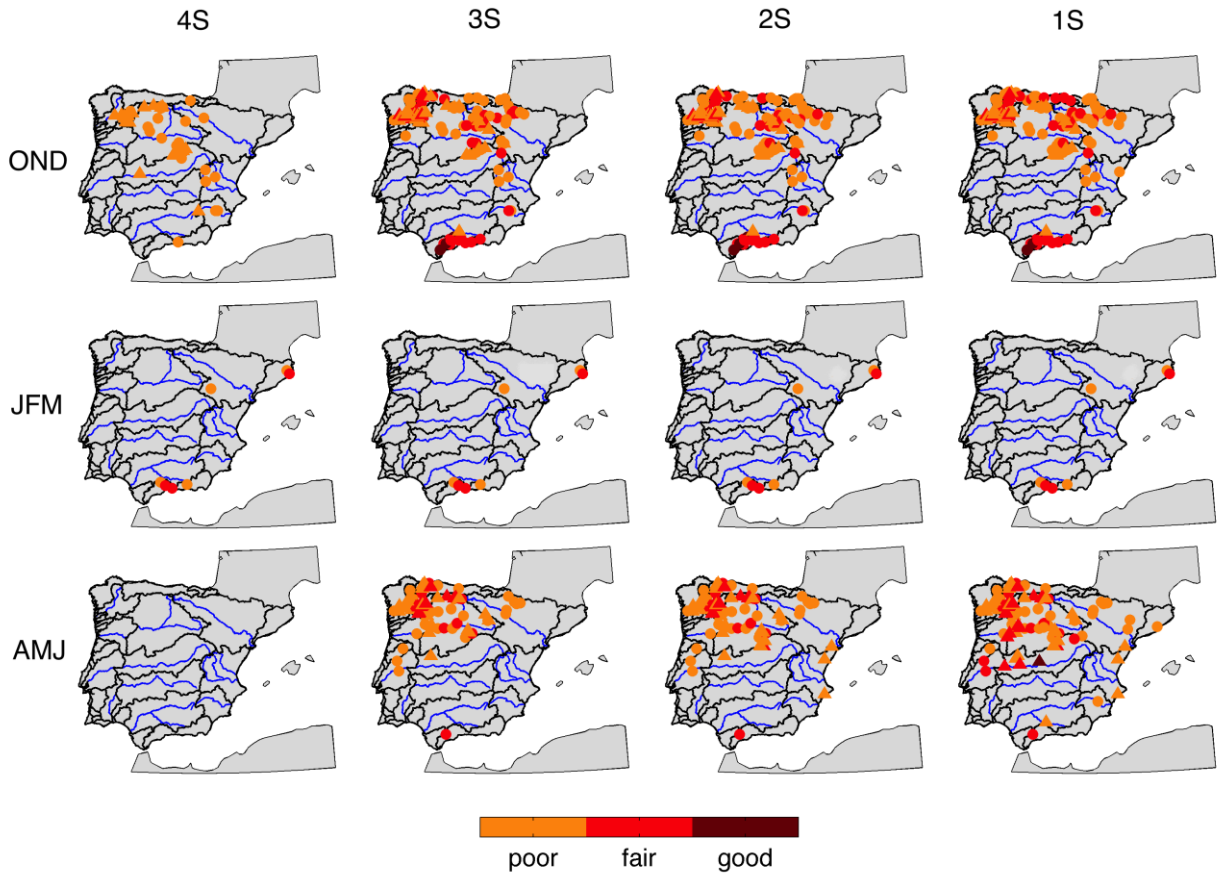


Figure 7.3. Maps of the forecasting skill classification (poor, fair or good) according with RHO, RMSESS and GSS values. They are displayed only stations where some forecast skills were found. In rows the seasonal streamflow forecasted and in columns the forecasting scenario.

As in case of the forecasting streamflow based on teleconnection indices, an example is shown in Figure 7.4, showing the forecasting results for one station in each season (OND streamflow in the station id=10028, JFM streamflow in the station id=46 and AMJ streamflow in the station id=3013, where 'id' indicates the identifier of each station). The stations were selected among the more skilfully forecasted with the consideration to be located in different basins. The values of the verification scores are given in Table 7.13.

In Figure 7.4a, station id=10028 is used as an example of autumn streamflow forecasting in the 3S scenario. This station is located in the Mediterranean Andalusian Basin (Figure 7.4a, left panel). The predictor used (after backward stepwise selection) was Zdjf2. Figure 7.4a, central panel, shows the good agreement between the observed and forecasted autumn streamflow (RHO, RMSESS and GSS are 0.69, 27.1% and 0.53, respectively). That is also manifested in the right panel, where the time series of both observed and forecasted streamflow

are plotted. Some notable details are the good agreement at the beginning and ending of the time series, whereas the variability of the forecasted time series during late eighties and early nineties is smaller than the observed and some peaks are not well reproduced. The more reduced variability observed in the middle of the time series is associated with the observed in winter NAO index during these years. Figure 7.4b presents the forecasting results for winter streamflow in station id=46 in the Internal Catalonian Basin (left panel) for 4S forecasting scenario. In this case, Pdjf3 was the predictor selected in the backward stepwise procedure and verification scores RHO, RMSESS and GSS are 0.53, 14.6% and 0.31, respectively. In general, low frequency variability is well reproduced although the low and high flows are usually underestimated, particularly in the middle of the time series. Figure 7.4c shows an example of spring streamflow forecasting in scenario 1S. The station selected (id=3013) is located in the Tagus Basin (left panel). The verification scores reach values up to 0.66, 24.7%, and 0.54 for RHO, RMSESS and GSS, respectively. The mode used as predictor was Tdjf1. From the observed and forecasted time series displayed in the right panel, it is seen a general good agreement in general, particularly in the late nineties and first years of twentieth century.

| Season | Station Id | Basin | Scenario | RHO | RMSESS | GSS | CLASS |
|---|---|---|---|---|---|---|---|
| **OND** | 10028 | Mediterranean Andalusian | 3S | 0.69 | 27.1 % | 0.53 | Good |
| **JFM** | 46 | Internal Catalonian | 4S | 0.53 | 14.6 % | 0.31 | Fair |
| **AMJ** | 3013 | Tagus | 1S | 0.66 | 24.7 % | 0.54 | Good |

Table 7.13. Values of the verification scores (RHO, RMSESS and GSS) for the three stations selected as examples in Figure 7.4.

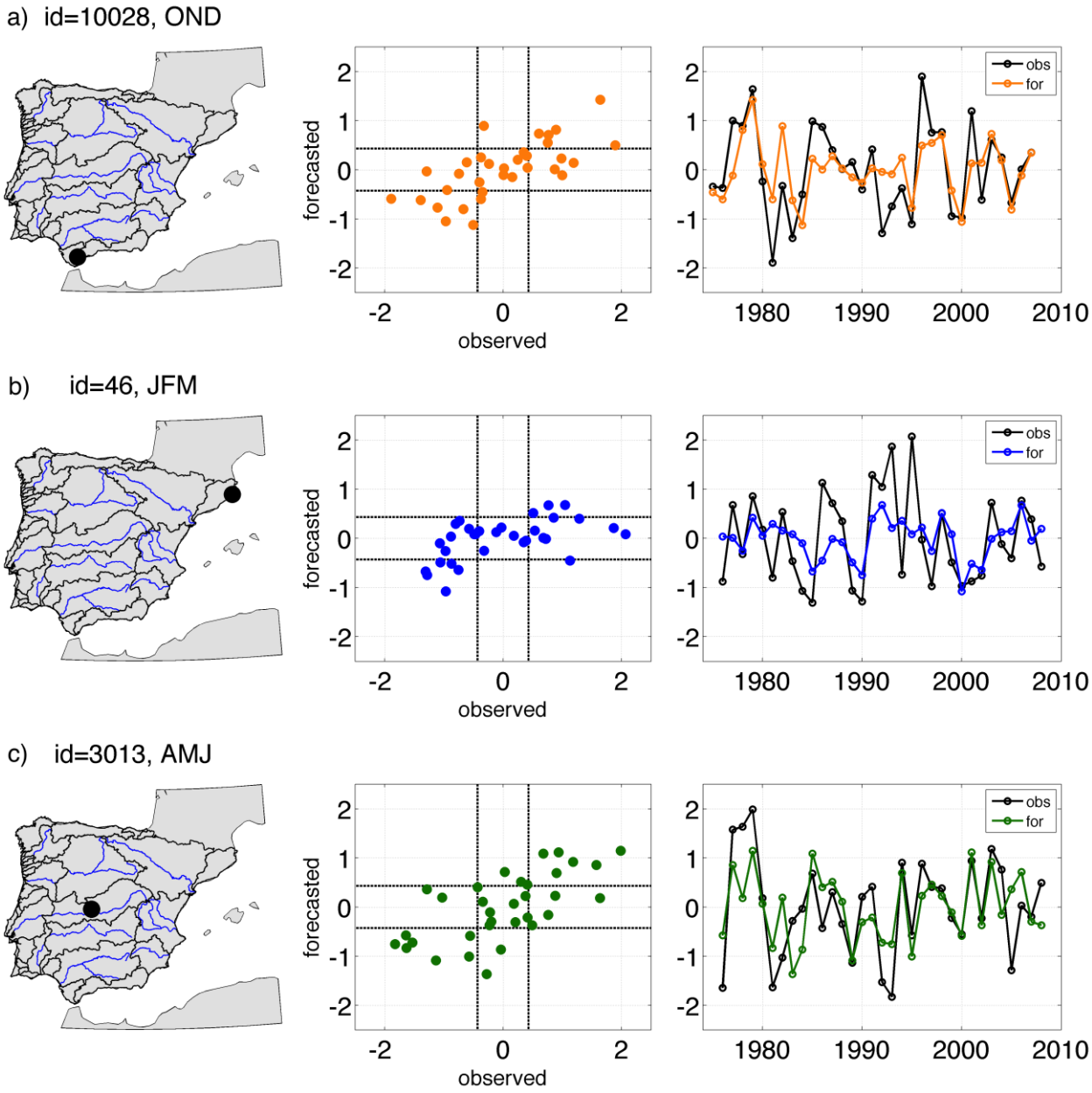


Figure 7.4 In rows, an example of forecasted seasonal streamflow time series for a gauging station in each season. a) for autumn streamflow in station id=10028 (Mediterranean Andalusian Basin), b) for winter streamflow in station id=46 (Internal Catalonian Basins) and c) for spring streamflow in station id=3013 (Tagus Basin). In left panels, the location of these stations; in middle panels, scatter plots between the observed and forecasted seasonal streamflow in each station (black dashed line indicate the 33$^{rd}$ and 66$^{th}$ percentiles); and in right panels, the observed (black line) and forecasted (coloured lines) streamflow time series in each forecasting scenario.

### 7.3.3 Prediction based on teleconnection indices and SVD modes

Tables 7.14 to 7.16 show the result from the application of the combining VIF and backward stepwise methods for selecting predictors (for autumn, winter and spring streamflow respectively), for all forecasting scenarios, when the pool of predictors is integrated by teleconnection indices and SVD modes. In these tables, it is displayed the number of stations where the teleconnection indices and the SVD modes are identified as stable predictor, those where the predictor overcome the VIF criterion and those where the predictor was selected after the backward stepwise selection.

| | 4S | | | 3S | | | 2S | | | 1S | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **stable** | **VIF** | **stepwise** | **stable** | **VIF** | **stepwise** | **stable** | **VIF** | **stepwise** | **stable** | **VIF** | **stepwise** |
| $WP_{s1}$ | 14 | 12 | 11 | 14 | 13 | 12 | 14 | 12 | 8 | 14 | 13 | 11 |
| $Niño3_{s1}$ | 20 | 13 | 10 | 20 | 12 | 5 | 20 | 4 | 1 | 20 | 1 | 0 |
| $Niño4_{s1}$ | 45 | 25 | 7 | 45 | 23 | 9 | 45 | 24 | 5 | 45 | 25 | 2 |
| $Niño3.4_{s1}$ | 29 | 10 | 8 | 29 | 1 | 0 | 29 | 2 | 0 | 29 | 1 | 0 |
| $SOI_{s1}$ | 41 | 34 | 27 | 41 | 23 | 14 | 41 | 31 | 11 | 41 | 26 | 5 |
| $SAI_{s1}$ | 21 | 21 | 18 | 21 | 21 | 13 | 21 | 21 | 11 | 21 | 20 | 13 |
| Pson1 | 38 | 29 | 1 | 15 | 15 | 0 | 38 | 24 | 9 | 38 | 24 | 8 |
| Ason2 | 25 | 25 | 7 | 25 | 25 | 3 | 25 | 25 | 2 | 25 | 19 | 9 |
| Tson1 | 110 | 108 | 29 | 110 | 105 | 69 | 110 | 106 | 24 | 110 | 97 | 39 |
| Tson3 | 9 | 9 | 6 | 9 | 9 | 1 | 9 | 9 | 3 | 9 | 9 | 4 |
| Rson1 | 200 | 196 | 24 | 200 | 194 | 24 | 200 | 166 | 55 | 200 | 166 | 87 |
| $NAO_d$ | - | - | - | 20 | 19 | 14 | 15 | 15 | 14 | 15 | 15 | 2 |
| $AO_d$ | - | - | - | 22 | 13 | 3 | 20 | 18 | 7 | 20 | 16 | 7 |
| $Niño4_d$ | - | - | - | 8 | 8 | 4 | 22 | 11 | 2 | 22 | 5 | 0 |
| $WeMO_d$ | - | - | - | 38 | 23 | 7 | 8 | 7 | 6 | 8 | 7 | 5 |
| Adjf1 | - | - | - | 150 | 148 | 59 | 150 | 140 | 22 | 150 | 142 | 77 |
| Adjf2 | - | - | - | 20 | 20 | 4 | 20 | 20 | 1 | 20 | 11 | 9 |
| Zdjf2 | - | - | - | 22 | 22 | 2 | 22 | 22 | 8 | 22 | 22 | 20 |
| Rdjf1 | - | - | - | 185 | 178 | 131 | 185 | 172 | 161 | 185 | 174 | 107 |
| $EA_m$ | - | - | - | - | - | - | 21 | 21 | 9 | 21 | 19 | 5 |
| $WP_m$ | - | - | - | - | - | - | 15 | 12 | 4 | 15 | 12 | 2 |
| $EP\text{-}NP_m$ | - | - | - | - | - | - | 74 | 67 | 55 | 74 | 61 | 46 |
| $WeMO_m$ | - | - | - | - | - | - | 47 | 42 | 15 | 47 | 42 | 11 |
| Pmam2 | - | - | - | - | - | - | 81 | 81 | 70 | 81 | 81 | 23 |
| Rmam2 | - | - | - | - | - | - | 46 | 12 | 8 | 46 | 46 | 39 |
| $EA_j$ | - | - | - | - | - | - | - | - | - | 26 | 25 | 6 |
| $Niño3.4_j$ | - | - | - | - | - | - | - | - | - | 16 | 13 | 7 |
| $SOI_j$ | - | - | - | - | - | - | - | - | - | 16 | 16 | 8 |
| $WeMO_{ja}$ | - | - | - | - | - | - | - | - | - | 17 | 17 | 10 |
| Tjja2 | - | - | - | - | - | - | - | - | - | 42 | 42 | 24 |

Table 7.14 Number of stations where predictors (teleconnection indices and SVD modes) of autumn streamflow present stable correlations and overcome the VIF criterion and the backward stepwise selection for the four forecasting scenarios.

In Table 7.14 (for autumn streamflow forecasting), although VIF exclude some predictors (such as those associated with ENSO), the most remarkable predictor exclusion is carried out by the backward stepwise selection. For example, in 4S scenario although Rson1 correlates with 200 stations, it was finally used as predictor in 24 of them. On the other hand, Rdjf1 remains as predictor in most of the stations. In general, it is observed a decreasing in the number of stations where the ENSO related indices are used as predictor (because they represent the same phenomenon, and the method choses the best one in each case). In the case of winter streamflow, predictors are not highly correlated between them, because of they do not usually represent the same event, so they are not usually dropped during the VIF and backward stepwise selection. For spring streamflow forecasting, something similar to autumn with ENSO related indices happen (it is observed pronounced decreasing in the number of stations where they are used as predictors). Also, It is noticeable a 'dance of numbers' between scenarios for a particular predictor (i.e., a particular teleconnection index or SVD mode can be selected as predictor in more stations in some scenarios than in other, which is due to the modification of predictor selection as a result of the incorporation of new predictors in the upcoming scenarios).

| | **4S** | | | **3S** | | | **2S** | | | **1S** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **stable** | **VIF** | **stepwise** | **stable** | **VIF** | **stepwise** | **stable** | **VIF** | **stepwise** | **stable** | **VIF** | **stepwise** |
| $EMI_{d1}$ | 5 | 5 | 4 | 5 | 5 | 4 | 5 | 5 | 4 | 5 | 5 | 1 |
| $WeMO_{d1}$ | 6 | 6 | 5 | 6 | 6 | 5 | 6 | 6 | 5 | 6 | 6 | 5 |
| Pdjf3 | 12 | 12 | 0 | 12 | 12 | 0 | 12 | 12 | 3 | 12 | 12 | 10 |
| Adjf2 | 25 | 25 | 1 | 25 | 25 | 9 | 25 | 25 | 3 | 25 | 25 | 10 |
| $WP_{m1}$ | - | - | - | 67 | 67 | 63 | 67 | 67 | 61 | 67 | 67 | 58 |
| $EP\text{-}NP_{m1}$ | - | - | - | 18 | 18 | 15 | 18 | 18 | 15 | 18 | 18 | 17 |
| Pmam3 | - | - | - | 11 | 11 | 4 | 11 | 11 | 9 | 11 | 11 | 4 |
| $PNA_{j1}$ | - | - | - | - | - | - | 13 | 13 | 10 | 13 | 13 | 0 |
| $SCAND_{s1}$ | - | - | - | - | - | - | - | - | - | 45 | 45 | 13 |
| $WeMO_{s1}$ | - | - | - | - | - | - | - | - | - | 9 | 9 | 8 |
| $SAI_{s1}$ | - | - | - | - | - | - | - | - | - | 238 | 238 | 237 |

Table 7.15 As Table 7.14 but for winter streamflow.

| | 4S | | | 3S | | | 2S | | | 1S | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | stable | VIF | stepwise | stable | VIF | stepwise | stable | VIF | stepwise | stable | VIF | stepwise |
| $EA_{j1}$ | - | - | - | 8 | 8 | 6 | 8 | 8 | 5 | 8 | 8 | 5 |
| $EP\text{-}NP_{j1}$ | - | - | - | 20 | 20 | 19 | 20 | 20 | 19 | 20 | 20 | 13 |
| $PNA_{j1}$ | - | - | - | 33 | 33 | 31 | 33 | 33 | 30 | 33 | 30 | 11 |
| $SCAND_{j1}$ | - | - | - | 9 | 9 | 8 | 9 | 9 | 8 | 9 | 9 | 7 |
| $AO_{j1}$ | - | - | - | 53 | 53 | 51 | 53 | 53 | 51 | 53 | 53 | 14 |
| $Niño1+2_{j1}$ | - | - | - | 37 | 37 | 23 | 37 | 37 | 19 | 37 | 21 | 14 |
| $Niño3_{j1}$ | - | - | - | 33 | 29 | 2 | 33 | 31 | 7 | 33 | 18 | 2 |
| $Niño3.4_{j1}$ | - | - | - | 11 | 10 | 8 | 11 | 11 | 6 | 11 | 10 | 1 |
| $IOD_{j1}$ | - | - | - | 10 | 10 | 1 | 10 | 10 | 0 | 10 | 6 | 1 |
| Pjja1 | - | - | - | 78 | 58 | 23 | 78 | 78 | 7 | 78 | 46 | 10 |
| Zjja1 | - | - | - | 104 | 103 | 40 | 104 | 104 | 34 | 104 | 99 | 75 |
| Tjja1 | - | - | - | 189 | 189 | 13 | 189 | 189 | 42 | 189 | 180 | 172 |
| Tjja2 | - | - | - | 27 | 27 | 6 | 27 | 27 | 4 | 27 | 27 | 19 |
| $Niño1+2_{s1}$ | - | - | - | - | - | - | 54 | 45 | 25 | 54 | 27 | 10 |
| $IOD_{s1}$ | - | - | - | - | - | - | 27 | 27 | 21 | 27 | 27 | 7 |
| $SAI_{s1}$ | - | - | - | - | - | - | 7 | 7 | 4 | 7 | 7 | 1 |
| Zson3 | - | - | - | - | - | - | 46 | 46 | 4 | 46 | 46 | 32 |
| $NAO_d$ | - | - | - | - | - | - | - | - | - | 17 | 17 | 12 |
| $SCAND_d$ | - | - | - | - | - | - | - | - | - | 18 | 18 | 8 |
| $AO_d$ | - | - | - | - | - | - | - | - | - | 10 | 9 | 1 |
| $Niño1+2_d$ | - | - | - | - | - | - | - | - | - | 28 | 20 | 9 |
| Adjf2 | - | - | - | - | - | - | - | - | - | 7 | 7 | 3 |
| Tdjf1 | - | - | - | - | - | - | - | - | - | 41 | 41 | 31 |
| Rdjf3 | - | - | - | - | - | - | - | - | - | 14 | 14 | 11 |

Table 7.16 As Table 7.14 but for spring streamflow.

Table 7.17 and Figure 7.5 display the result for the forecasting skill assessment. With respect to autumn streamflow forecasting, although it was forecasting in 243 stations in 4S scenario, only in 37 and 8 of them the quality of the forecast was classified as poor and fair, respectively. From Figure 7.5, most of the stations are in the Miño-Sil, Douro and, especially, Mediterranean Andalusian Basins. Rson1, Tson1 and SAI seem to be the main predictors accounting for these results. The forecasting skill increases when moving to 3S scenario (i.e., including predictor from previous winter). On one hand, the number of stations forecasted goes to 306, with a notable increase in the number of stations with satisfactory forecasting results (classified as poor, fair and good 70, 34 and 7 cases, respectively). In particular, it is noteworthy how the quality of the forecast improves in Mediterranean Andalusian Basin, where 12 stations are fair or good forecasted (as a consequence of the inclusion of winter NAO as predictor) and in the Miño-Sil, Cantabrian and upper Ebro Basins. In this case, Adjf1 and Rdjf1 give the

impression of being behind these results. Also, the number of stations skillful forecasted in the Mediterranean Basins increases. In case of 2S scenario, although the number of stations forecasted is similar (316), it is observed an improvement of the forecasting skill (from 111 to 137 stations), particularly in Tagus and Ebro Basins (as a result of adding EP-NP in spring and Rmam2 to the predictor pool). Finally, in 4S scenario, it is observed an increase in the quality of the forecast in Cantabrian Basin, which it seems to be related to the inclusion of Tjja2 as predictor.

Results of forecasting winter streamflow are not competent until 1S scenario, when SAI is included as predictor. Whereas in previous scenarios, only some stations in the Mediterranean slope exhibit some forecasting skills (for example, 22 in 3S scenario), in 1S scenario, a total of 285 stations (covering most of IP) were forecasted, with poor, fair and good forecasting skill in 100, 107 and 36, respectively.

In case of spring streamflow, predictions are made from three seasons in advance (3S scenario), because of there are not any predictor from previous spring. In 3S scenario, a total of 254 stations are forecasted, but only 95 of them present some noteworthy results (in 61 the forecast was classified as poor, in 32 as fair and in 2 as good). These skillfully forecasted stations are mainly located in the northwestern of IP, especially in Miño-Sil, Douro and Tagus Basins. The most relevant predictors in this case are likely summer AO and Zjja1 (closely related) and those associated with ENSO (such as Niño1+2, Pjja1 or PNA). The quality of the forecast increases slowly as moving to the target season. For instance, in 1S scenario 105 stations are skillfully forecasted (only 10 more than in 3S scenario). This improvement is particularly observed in Tagus Basin, after including winter NAO as predictor.

| | **4S** | | | | **3S** | | | | **2S** | | | | **1S** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Bad | Poor | Fair | Good | Bad | Poor | Fair | Good | Bad | Poor | Fair | Good | Bad | Poor | Fair | Good |
| **OND** | 198 | 37 | 8 | 0 | 195 | 70 | 34 | 7 | 179 | 91 | 38 | 8 | 181 | 81 | 49 | 9 |
| **JFM** | 16 | 6 | 5 | 0 | 88 | 15 | 7 | 0 | 94 | 17 | 7 | 0 | 42 | 100 | 107 | 36 |
| **AMJ** | 0 | 0 | 0 | 0 | 159 | 61 | 32 | 2 | 174 | 70 | 25 | 3 | 198 | 69 | 34 | 2 |

Table 7.17. Number of stations belonging to each class defined for summarizing forecasting skill (bad, poor, fair and good). In rows the four scenarios considered and in columns the seasonal streamflow forecasted (autumn -OND-, winter -JFM- and spring -AMJ-).

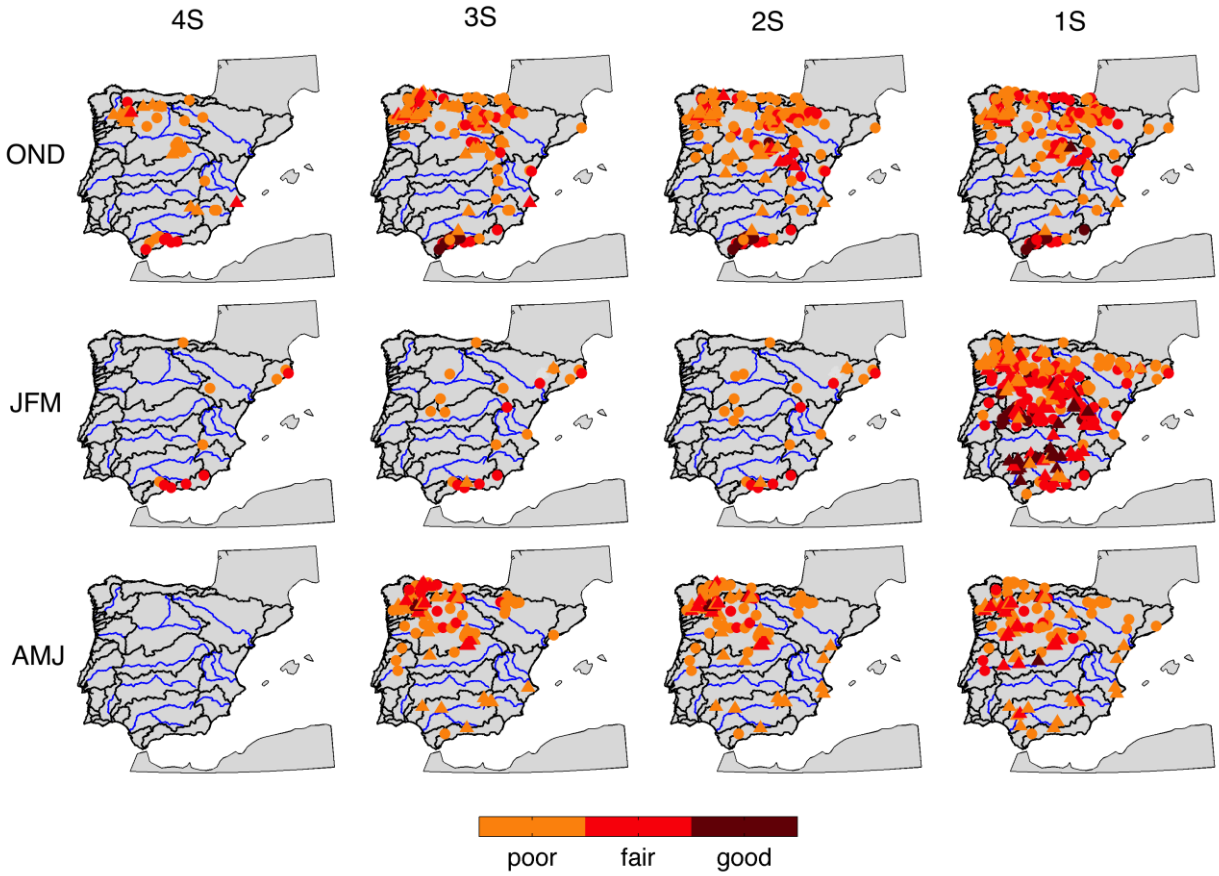


Figure 7.5. Maps of the forecasting skill classification (poor, fair or good) according with RHO, RMSESS and GSS values. They are displayed only stations where some forecast skills were found. In rows the seasonal streamflow forecasted and in columns the forecasting scenario.

Finally, as in the previous forecasting schemes, an example of forecasting results when combining both subsets of predictors, teleconnection indices and SVD modes is provided in Figure 7.6, showing forecasting results for one station in each season (OND streamflow in station id=10028, JFM streamflow in station id=3163 and AMJ streamflow in station id=1710, where 'id' indicates the identifier of each station). The stations were selected among the more skilfully forecasted with the consideration to be located in different basins. The values of the verification scores are given in Table 7.18.

In Figure 7.6a, station id=10028 is used as an example of autumn streamflow forecasting in the 3S scenario. This station is located in the Mediterranean Andalusian Basin (Figure 7.6a, left panel). The predictor used (after backward stepwise selection) was the previous winter NAO. Figure 7.6a, central panel, shows the good agreement between the observed and forecasted autumn streamflow (RHO, RMSESS and GSS are 0.73, 31.6% and 0.66, respectively). That is

also manifested in the right panel, which looks like similar to Figure7.6a, right panel, where a general good agreement in general is observed, particularly at the beginning and ending of the time series. Figure 7.6b presents the forecasting results for winter streamflow in station id=3163 in the Tagus Basin (left panel) for 1S forecasting scenario. In this case, WP in previous spring and SAI in previous autumn were the predictors selected in the 'VIF + backward stepwise' procedure and verification scores RHO, RMSESS and GSS are 0.73, 30.0% and 0.68, respectively. In general, low frequency variability is well reproduced although the low and high flows are usually underestimated, particularly in the middle of the time series. Figure 7.6c shows an example of spring streamflow forecasting in 3S scenario. The station selected (id=1710) is located in the Miño-Sil Basin (left panel). The verification scores reach values up to 0.62, 21.6%, and 0.47 for RHO, RMSESS and GSS, respectively, being Tjja1 the predictor used. From the observed and forecasted time series displayed in the right panel, although good overall agreement is found, variability in the forecasted time series is reduced when comparing with the observed, which means that most of higher peaks are not properly reproduced.

| Season | Station Id | Basin | Scenario | RHO | RMSESS | GSS | CLASS |
|---|---|---|---|---|---|---|---|
| **OND** | 10028 | Mediterranean Andalusian | 3S | 0.73 | 31.6 % | 0.66 | Good |
| **JFM** | 3163 | Tagus | 1S | 0.73 | 30.0 % | 0.68 | Good |
| **AMJ** | 1710 | Miño-Sil | 3S | 0.62 | 21.6 % | 0.47 | Good |

Table 7.18. Values of the verification scores (RHO, RMSESS and GSS) for the three stations selected as examples in Figure 7.6.

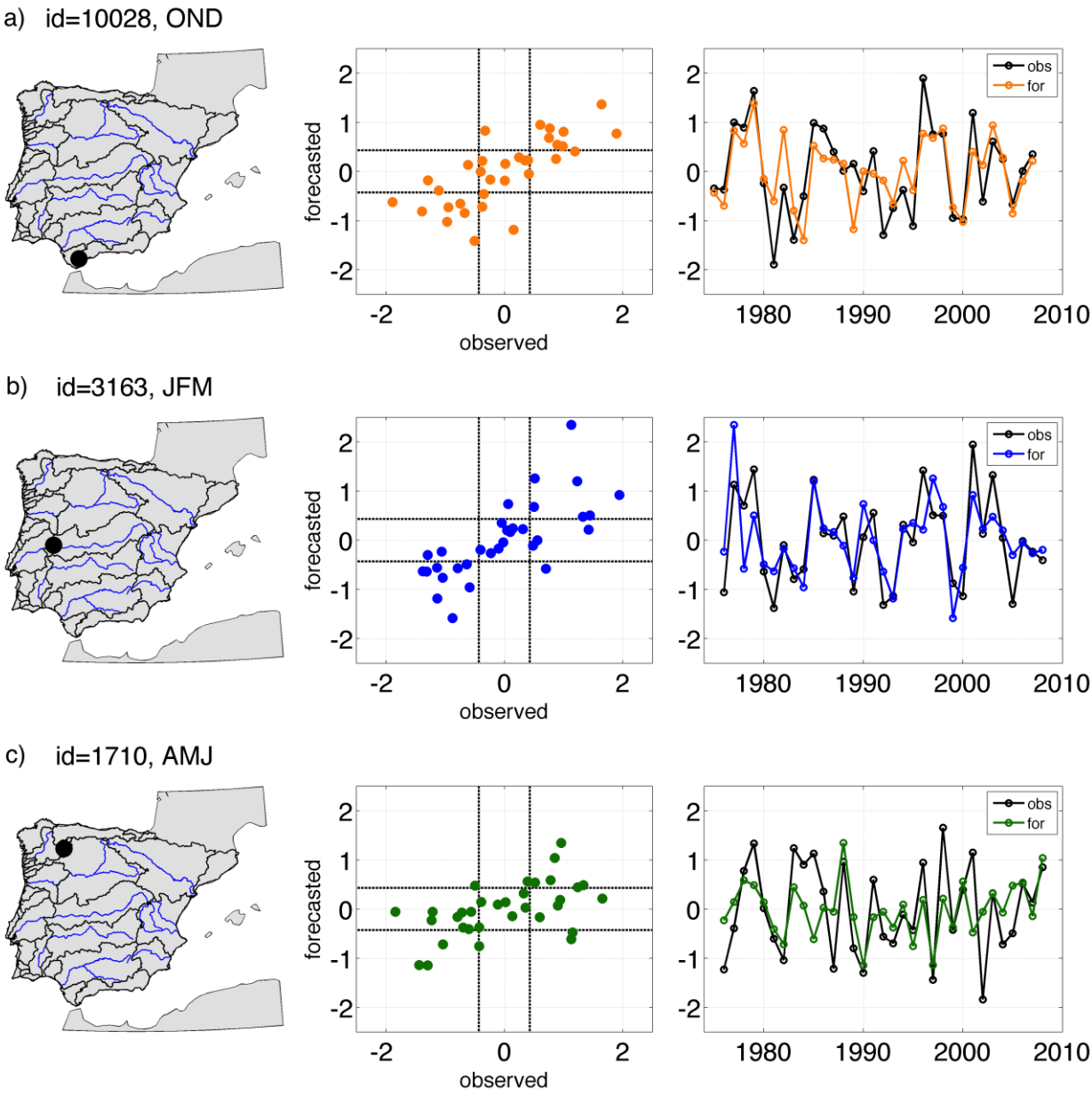


Figure 7.6. In rows, an example of forecasted seasonal streamflow time series for a gauging station in each season. a) For autumn streamflow in station id=10028 (Mediterranean Andalusian Basin), b) for winter streamflow in station id=3163 (Tagus Basin) and c) for spring streamflow in station id=1710 (Miño-Sil Basin). In left panels, the location of these stations; in middle panels, scatter plots between the observed and forecasted seasonal streamflow in each station (black dashed line indicate the 33$^{rd}$ and 66$^{th}$ percentiles); and in right panels, the observed (black line) and forecasted (coloured lines) streamflow time series in each forecasting scenario.

## 7.4 Summary

In this chapter it has been evaluated the ability of teleconnection indices and left expansion coefficients of the significant SVD modes as predictors of seasonal streamflow. Additionally it has been evaluated the advantage of using both datasets together. The forecasting methodology, based on multiple linear-regression models, was developed for four forecasting scenarios, related to the number of seasons prior to which the forecasting is made, from one year until one season in advance, updating and improving the predictions seasonally. This approach becomes a useful tool for decision making related to water management, since predictions are given with different seasons in advance. The question of multicollinearity among predictors was addressed by combining two iterative methods in the regression analysis: variance inflation fraction and backward stepwise selection.

For forecasting based on teleconnection indices:

- In case of autumn streamflow forecasting, reasonable good predictions (with RHO > 0.5, RMSESS > 20% and GSS > 0.4) are found for stations located in the Mediterranean Andalusian Basin, particularly from 3S scenario (i.e. considering the information provided by seasonal teleconnection indices from previous autumn and winter). In these cases, the SAI measured in October of the previous year, and the previous winter NAO and AO indices constitute the predictors used. Moderately accurate predictions (RHO > 0.44, RMSESS > 10% and GSS > 0.2) were also found in the 3S and 2S scenario for stations located in the northern basins of the IP (Miño-Sil, Douro, Ebro, and Tagus Basins), ENSO indices in the previous autumn, together with spring EP-NP and WeMO indices being the most widely used predictors.
- For winter streamflow forecasts, in the scenarios 4S, 3S, and 2S, forecasting skill was found only in stations located in the Mediterranean sector, although most of them presented poor forecasting skill. On the other hand, a significant number of stations (168) presented moderate or good predictions (with RHO > 0.44, RMSESS > 10% and GSS > 0.2) for the 1S scenario, especially in basins located in the central and western IP. The SAI was identified as the main predictor for winter streamflow, in both the number of stations significantly correlated and the highest correlation values. From this result, teleconnection indices are not able to predict winter streamflow more than two months in advance, and only when the October SAI is added, skillful predictions are made. These results agree with those of Brands et al. (2012, 2013), who forecasted winter precipitation with reasonable skill over the central and western IP by using the SAI as the predictor. In

addition to this, SAI is a very helpful predictor not only for following winter precipitation and streamflow but also for following autumn streamflow in Mediterranean Andalusian Basin. Bearing in mind that the temporal coverage of the 'daily' version of SAI (which presents higher correlations with JFM streamflow than the 'weekly' version) is still short, we conclude that the larger time series of SAI available, the more skillful winter streamflow forecasting.

- Spring presents the poorest forecasting results (only 17 stations showed moderate or good predictions in the most favorable scenario, 1S). Stable positive correlations were found with indices related to atmospheric or oceanic variability in the Pacific, such as Niño1+2, Niño3, Niño3.4, IOD, PNA or EP-NP in summer for stations in the northwest. Also, summer AO was found to be a stable predictor of stations in the northern half of the IP. In the case of stations located in central and southern basins of the IP, the previous winter SCAND and NAO are the teleconnection indices best correlated with them.

Results from forecasting experiment based on significant SVD modes indicate that:

- In case of autumn streamflow forecasting, moderately skillful predictions are obtained from 3S scenario, particularly in the northwestern of IP (Miño-Sil, Cantabrian, and upper Douro, Ebro and Tagus Basins) and also in the southern part (Mediterranean Andalusian Basin). In case of the Mediterranean Andalusian Basin, 13 of 15 stations in this Basin were fair or good forecasted (with RHO > 0.44, RMSESS > 10% and GSS > 0.2). The predictors associated with the acceptable results in northern basins are Pson1 and Rson1, related to El Niño phenomenon, Tson1, linked with the impacts of autumn SCA in global temperature, and Adjf1, which could be related to the influence of tropical Atlantic on tropical Pacific (as a precursor of El Niño phenomenon). On the other hand, the autumn streamflow in the central and eastern façade of the IP is not properly forecasted. The reason behind is basically that there was few stations located in the Mediterranean slope (except those in the Mediterranean Andalusian Basin and some in the Jucar and Segura Basins) with stable predictors. Then, despite the quality of predictions is not very high (except for stations in the Mediterranean Andalusian Basin), they are obtained with up to seven months in advance, which is an interesting gap for developing water management strategies.
- Regarding winter streamflow forecasting, the results in both number of stable predictors and number of stations with a reasonable good forecast are too low to make conclusions. Only some stations (mainly in the Catalonian and Mediterranean Andalusian Basins)

were forecasted. In case of stations in the Mediterranean Andalusian Basin, the predictor used in Adjf2, which resembles the 'tripole' pattern in North Atlantic, which is linked to NAO. This finding agrees with the found in the autumn streamflow forecasting in the Basin, indicating that the influence of winter NAO on streamflow on the Mediterranean Andalusian Basin could extent from following autumn to winter.

- As spring streamflow forecasting concerns, there is no predictor for the 4S scenario. The first predictions are found in the 3S scenario, particularly for stations in the Miño-Sil and Douro Basin (up to 57 stations in both basins). The Pjja1 (associated with ENSO), Zjja1 (linked to Northern Annular Mode) and Tjja1 are the main predictors for those stations. Moving to 2S and 1S scenario, there is a slight increase in the quality of predictions, which is more significant in case of 1S (particularly in the Douro and Tagus Basins), when Zdjf1 and Tdjf1 are added as predictors. Both modes are related to winter NAO, which is major driver of winter precipitation in western IP.

One last comment regarding forecasting streamflow based on SVD modes. It was noted in the 'Introduction' and 'Methodology' sections that the SVD must be conducted for each calibration subset to create the left expansion coefficients for each leave-one-out realization, because of the use of the coefficient calculated using the whole period could lead in a overestimation of forecasting skills produced when information about the year to be forecasted is used to identify and construct the predictors. To evaluate this 'overestimation', the entire forecasting methodology was repeated using as predictors the left expansion coefficients derived from the SVD carried out with the entire period, and removing the target year from these left expansion coefficients for calibrating the regression models. Figure 7.7 and table 7.19 summarize the classification of forecasting skills. It could be seen a remarkable increasing in both number of stations skillful predicted and quality of predictions, particularly in case of autumn and spring. Then, it is noteworthy how the incorrect use of SVD for forecasting purposes (i.e., using the information of the target years in the construction of SVD modes, which are used as predictors) produces an overestimation in forecasting skills.

| | 4S | | | | 3S | | | | 2S | | | | 1S | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Bad | Poor | Fair | Good | Bad | Poor | Fair | Good | Bad | Poor | Fair | Good | Bad | Poor | Fair | Good |
| **OND** | 46 | 53 | 99 | 40 | 58 | 80 | 106 | 58 | 46 | 68 | 123 | 74 | 49 | 69 | 119 | 79 |
| **JFM** | 4 | 12 | 6 | 0 | 11 | 13 | 7 | 0 | 11 | 13 | 7 | 0 | 11 | 13 | 7 | 0 |
| **AMJ** | 0 | 0 | 0 | 0 | 39 | 42 | 100 | 48 | 46 | 42 | 105 | 54 | 54 | 48 | 116 | 55 |

Table 7.19. Number of stations belonging to the each of the classes defined for summarizing forecasting skill (bad, poor, fair and good). In rows the four scenarios considered and in columns the seasonal streamflow forecasted (autumn-OND-, winter-JFM-and spring-AMJ-).

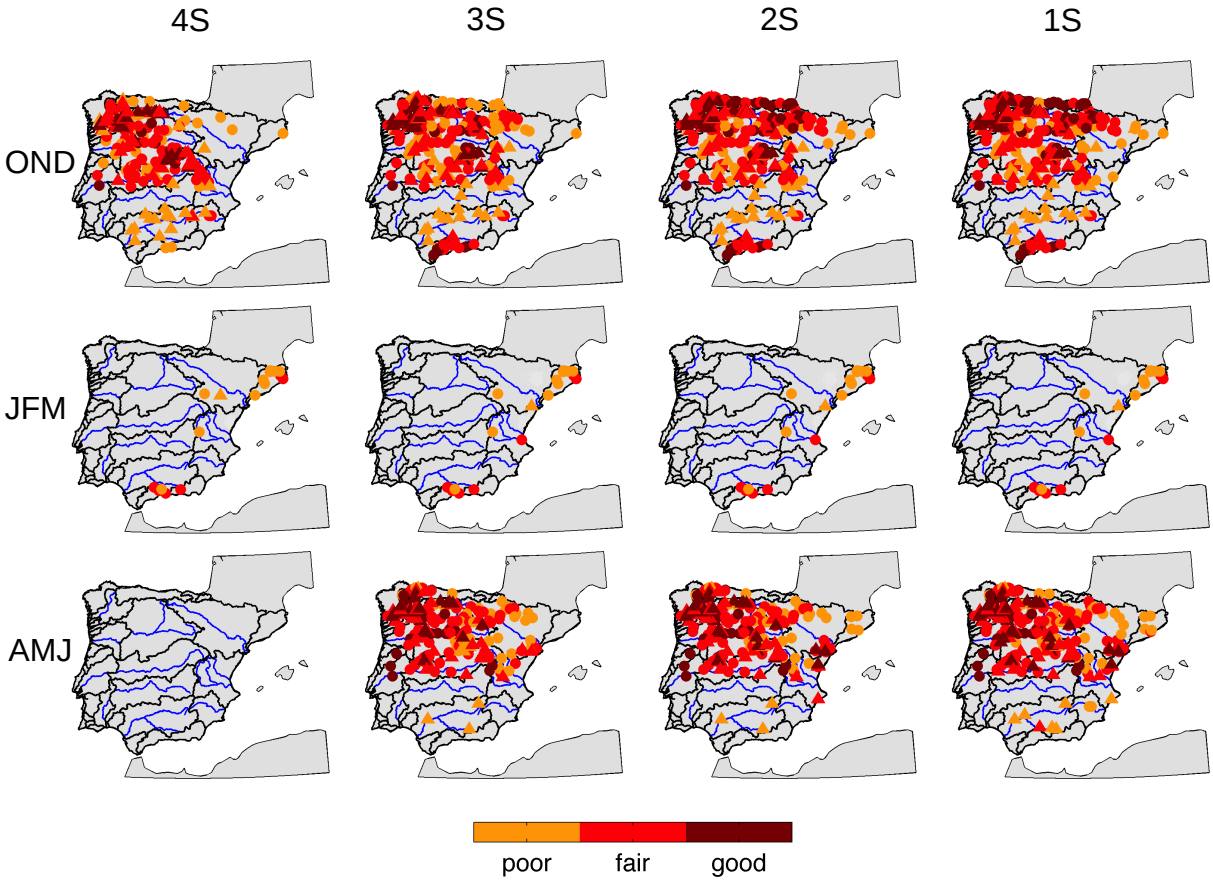


Figure 7.7. Maps of the forecasting skill classification (poor, fair or good) according with RHO, RMSESS and GSS values. Only stations with some forecast skills are displayed. In rows the seasonal streamflow forecasted and in columns the forecasting scenario.

Finally, the main conclusions from forecasting experiment combining teleconnection indices and significant SVD modes indicate that:

- In case of autumn streamflow, forecast based on SVD modes provide better results than the based on teleconnection indices, particularly in the northern IP. Nevertheless, adding teleconnection indices as predictors improve forecasting in some basins and scenarios, such as in the Mediterranean Andalusian Basin (for the 4S scenario) or the Tagus Basin (for the 2S and 1S scenarios).
- Predictability of winter streamflow is almost inexistent (excluding some stations in the Mediterranean slope) for all cases of predictors (SVD modes, teleconnection indices or a combination of both), except when SAI of previous October is considered as predictor. In that case, most station in the IP are reasonable well forecasted (except in the Cantabrian and Mediterranean slopes).
- Similarly to autumn streamflow, the use of SVD modes as predictors provides overall more skillful predictions of spring streamflow. However, when teleconnection indices are added as predictors, some improvements are observed in forecasting skills, particularly in the southern IP for the 3S, 2S and 1S scenarios.

# CHAPTER 8

# CONCLUSIONS

This chapter recapitulates the main conclusions obtained from this dissertation. It is structured in three parts: the first one summarizes the results related to the analysis of the raw database and the spatial and temporal description of the seasonal streamflow variability. The second part outlines the results derived from the analysis of potential climate predictors and their forecasting skills. Finally, the third section addresses the limitations of this study and the fundamental lines of future work.

Streamflow forecasting is challenging because of the complexity of the hydrologic system. Improving the quality of streamflow forecasting has always been an important task for researchers and hydrologic forecasters. Then, the motivation of this dissertation relies on the necessity to improve the understanding of the large-scale climate variability that drives the streamflow variability on the Iberian Peninsula, which becomes the basis for developing streamflow forecast at scales that are of paramount importance for population and natural environment.

The original database comprises 1380 gauging stations. From the trade off between number of available stations, covering period and percentage of gaps, 504 stations involving 1975-2008,

with less than 10% of missing values, were selected. In addition to data availability, the quality and homogeneity of streamflow time series are prerequisites for reliable and trustworthy assessment of seasonal streamflow forecasting. Hence, to investigate the presence of inhomogeneities in hydrological time series (particularly because of regulation effects), a combined methodology based on Pettitt test and Common Area Index was applied to the seasonal time series. Consequently, 122 gauging stations were discarded because of the existence of a non-natural behaviour. These stations were mainly located in the main stream of the Tagus River and the Guadiana, Guadalquivir, Segura and Jucar Basins. Changes in time series were mainly identified in summer, with increase in discharge in years after the construction of a reservoir upstream (probably because of reservoir management to ensure water availability during the drier months). Sometimes the change in summer is related to a decrease in river discharge in the wet season (for water storage in reservoirs), which could be significantly enough to result in a great alteration of annual cycle. Then, for the remaining 382 stations, a reconstruction scheme was performed to fill in missing values by linear regression.

The first part of this work attempts to explore spatial and temporal characteristics of streamflow variability in the Iberian Peninsula. Regarding the time-location of the highest intrannual discharge, stations located in the western IP (Cantabrian, Miño-Sil, Douro, Tagus, Guadiana, Guadalquivir, Jucar and upper Ebro Rivers) present their highest seasonal values in winter. These stations usually present a highly marked seasonality (with higher values in winter and smaller in summer). The Jucar and Segura Basins usually present a similar streamflow regime during all year, which is mainly due to regulation effects. Stations in the Ebro Basin and also some stations belonging the Internal Catalonian Basin (those nearest to the Pyrenees) show their maximum discharge values in spring (because of snowmelt contributions). Also, there are some stations located in the Mediterranean slope (particularly those in the Mediterranean Andalusian Basin) with their maximum discharges in autumn (as consequence of convective precipitation in that area).

Streamflow variability in autumn and winter (the wettest seasons) presents similar spatial patterns according to PCA results, grouping the stations in the Miño-Sil, Douro, Tagus, Guadiana and Guadalquivir Basins in the first mode, stations in the Cantabrian and Upper Ebro Basins in the second mode, and stations in the Mediterranean Andalusian and Internal Catalonian Basins in the third and fourth modes. On the other hand, spatial patterns differ in spring and summer (drier half of the year), when they reveal a clear north-south gradient.

From trend analysis, overall significant trends are negative in all seasons, particularly in the northern and eastern stations, with some exceptions. For example, positive trends are found in autumn in Douro Basin, which could be partially related to an observed increase in their variability during the second half of the study period. Negative trends found in winter and spring could be related to the persistence of positive NAO phase during the last two decades of twentieth century, translated in a minor precipitation and snow during winter in most of the IP. The positive trends found in summer are likely to be a consequence of hydrological regulation to ensure water requirements for urban supplies and irrigation.

The second part of this Thesis consisted on identifying the main climate factors that have a noteworthy influence on near future (lagging from one to four seasons) seasonal streamflow variability of IP Rivers. Summer was not considered in these analyses because most stations present the lowest streamflow discharge and variability in this season. Moreover, this part provides an insight into the possible mechanisms and physical processes behind these relationships. Once identified, they were used as predictors in a seasonal streamflow forecasting exercise. The forecasting methodology, based on multiple linear-regression models, was developed for four forecasting scenarios, related to the number of seasons prior to which the forecasting is made, from four seasons (4S scenario) until one season (1S scenario) in advance, updating and improving the predictions seasonally. This approach becomes a useful tool for decision making related to water management, since predictions are given with different seasons in advance.

Firstly, teleconnection indices that summarize the main modes of variability of climate system were evaluated as potential predictors of seasonal (autumn, winter and spring) streamflow through an analysis of the stability of the significant correlations found. Major findings are recapitulated as:

- Although El Niño indices in previous autumn appear as potential predictors of following autumn streamflow for stations in the northwestern corner of the IP, the quality of the predictions based on them is not very high (classified as poor forecast). Conversely, autumn streamflow in the Mediterranean Andalusian Basin is good predicted (RHO > 0.5, RMSESS > 20%, and GSS > 0.4) with three seasons in advance, when information about previous October SAI and winter NAO are available. Moderate predictions (RHO > 0.44, RMSESS > 10%, and GSS > 0.2) are also found in the 3S and 2S scenarios for stations located the northern basins of the IP (Miño-Sil, Douro, Ebro, and Tagus Basins),

when spring EP-NP and WeMO indices are added to ENSO indices of the previous autumn as predictors.

- For winter streamflow forecasts, in scenarios 4S, 3S, and 2S, forecast is skilful only in stations located in the Mediterranean sector, although most of them presented poor forecasting skill. On the other hand, a significant number of stations (168) presented moderate or good predictions (with RHO > 0.44, RMSESS > 10%, and GSS > 0.2) for the 1S scenario, especially in basins located in the central and western IP. The SAI is identified as the main predictor for winter streamflow, in both the number of stations significantly correlated and the highest correlation values. Consequently, teleconnection indices are not able to predict winter streamflow more than two months in advance (when the October SAI is used as predictor). Hence, SAI is found a very helpful predictor not only for following winter precipitation and streamflow but also for following autumn streamflow in the Mediterranean Andalusian Basin.
- Spring was only fair or good forecasted in 17 stations in the most favourable scenario (1S). Although several indices related to atmospheric or oceanic variability in the Pacific in summer, such as Niño1+2, Niño3, Niño3.4, IOD, PNA or EP-NP and also summer AO present stable correlations with stations in the northern IP, prediction based on them are mostly classified as poor (46 of the 52 that present forecasting skill). These forecasting skills slightly increase in 1S scenario (specially in the southern half of the IP), when previous winter SCAND and NAO indices are added as predictors.

Secondly, Singular Value Decomposition (SVD) technique was employed to identify and isolate the main modes of covariability between seasonal streamflow and the climate variables (sea surface temperature, geopotential height at 500 hPa in Northern Hemisphere and global temperature and precipitation) that precede it from one to four seasons. This procedure allows us to establish new modes of covariability apart from the represented by teleconnection indices, which can be used as predictors of seasonal streamflow. Next paragraphs summarize the significant modes of covariability found and their ability as potential predictors of seasonal streamflow:

- Up to 13 modes of covariability were used to predict autumn streamflow (most of them associated with climate variability in previous autumn and winter). In fact, forecasting results for 3S scenario (with information of previous autumn and winter) are quite acceptable (42 stations with RHO > 0.44, RMSESS > 10%, and GSS > 0.2), particularly in the Mediterranean Andalusian Basin, with Zdjf2 and Adjf2 (related to winter NAO) as

predictors, but also in the north-northwestern IP (Douro, Miño-Sil, Cantabrian and upper Ebro Basins), being Rson1, Adjf1 and Rdjf1 (linked to ENSO) the main predictors.

- Only 3 modes of covariability between climate variables and winter streamflow were selected as potential predictors (significant and stable). Forecasting results were poor and only few stations in the Mediterranean Andalusian and Internal Catalonian Basins showed some forecasting skills, being Adjf2 (related to winter NAO) and Pdjf3 (related to El Niño Modoki) the predictors used.
- In case of spring streamflow, 8 modes of covariability were used as predictors. Four of them are related to summer climate variability, Pjja1 (similar to El Niño SST pattern), Zjja1 (resembling the summer Northern Annular Mode), Tjja1 and Tjja2. Forecasting models based on these four predictors (3S scenario) provide some skilful predictions (up to 21 stations with RHO > 0.44, RMSESS > 10%, and GSS > 0.2) in the northeastern quadrant of IP (specially in the Miño-Sil and Douro Basins). Prediction skills increase in Tagus Basin when Tdjf1 (related to winter NAO) is added to the pool of predictors in 1S scenario.

To sum up, the use of SVD improves the forecasting skills of autumn and spring streamflow with respect to the use of teleconnection indices as predictors. This improvement, likely due to the inclusion of additional information of climate system not reflected in teleconnection indices, is particularly relevant for 3S scenario. Conversely, winter streamflow was not forecasted in almost any stations. In this case, SAI appears to be the only reliable predictor.

In conclusion, this study can be advantageous compared with earlier studies examining the potential predictability of large-scale climate variability on seasonal streamflow variability of the IP Rivers in several ways. First, it depends on a dataset of long, complete, reliable and spatially well-covered time series, which enable us to describe with a high spatial resolution the potential use of different climate signals as predictors of seasonal streamflow in different areas of the IP. This fact becomes interesting for making local decision in water resources management, highlighting the need to adopt their future policy and development plans on a more local scale to meet future demands. Previous studies were only restricted to very smaller number of gauging stations. Second, this study explores the links between climate signal and streamflow variability of the IP Rivers, not only evaluating the most commonly used climate indices but also exploring further relationships between climate variability and streamflow in the following seasons. Third, this study can provide a more comprehensive view of relationship of climate variability and

streamflow on seasonal timescales in a way that can significantly contribute to streamflow forecasting purposes (with various forecasting schemes, according to the time in advance the predictions are made). In particular, the use of an approach based on seasonally updated predictions allows the option of developing water-management policies some seasons in advance and with the possibility of modifying or adjusting these strategies as the predictions are updated. Having skilful forecasting with several months in advance provides helpful information for water-management strategies.

**Future work**

There are several areas of the proposed research that can be significantly improved in future work, which are listed below:

- Bearing in mind that the temporal coverage of the 'daily' version of SAI (which presents higher correlations with JFM streamflow than the 'weekly' version) is still short, it can be concluded that the larger time series of SAI available, the more skilful winter streamflow forecasting. However, as the stationary of the relationship between October SAI and following winter AO has been questioned, it must be checked in future studies.
- Improving the understanding of physical basis supporting the links between winter NAO and following autumn streamflow in the Mediterranean Andalusian Basin, such as the role of summer SST in the Mediterranean Sea close to this Basin or other related mechanisms, requires more thoughtful investigation.
- In addition to SAI, the role of the cryosphere, in particular the sea-ice Arctic variability in summer, in streamflow of the IP Rivers in subsequent seasons is a matter of further investigations. In particular, the relationships between sea-ice decline and Northern Annular Mode in summer and its effect on following winter climate.
- When larger time series of seasonal streamflow become available, it will be interesting to explore the non-linear relationships between ENSO and seasonal streamflow, with separate analysis of the response to El Niño, Neutral and La Niña years. Also, the modulation that low frequency variability modes, such as PDO or AMO, could result in ENSO or NAO conditions, and its repercussion on seasonal streamflow variability, could be evaluated.

- Also, larger time series are more appropriate for applying different forecasting methodologies, such as categorical forecast framework based on probabilities, which could explore further the relationships between predictors and predictands.
- With the basis established here, forecasting at smaller spatial scales (basin or subbasin) can be improved considering the effects of variables such as land-cover, the snowmelt contribution, soil moisture or topography factors, and also using different forecasting methodologies, such as those based on hydrological models.

# CHAPTER 8

# CONCLUSIONES

Este capítulo resume las principales conclusiones obtenidas de esta Tesis Doctoral. Se estructura en tres partes: en la primera se resumen los resultados relacionados con el análisis de la base de datos y la descripción espacial y temporal de la variabilidad del caudal estacional en la Península Ibérica En el segundo apartado se exponen los resultados derivados del análisis de los posibles predictores climáticos y las predicciones obtenidas a partir de ellos. Por último, la tercera parte se ocupa de las limitaciones de este estudio y las líneas fundamentales de futuros trabajos.

La predicción del caudal entraña una alta dificultad, como consecuencia de la complejidad del sistema hidrológico. Mejorar la calidad de las predicciones del caudal ha sido siempre una tarea importante para los investigadores y los hidrometeorólogos. Por consiguiente, la motivación de esta Tesis radica en la necesidad de mejorar la comprensión de la variabilidad climática de gran escala que influye en la variabilidad del caudal de los ríos ibéricos, que a su vez constituye la base para el desarrollo de predicciones de caudal a escalas temporales de suma importancia para la población y el medioambiente.

La base de datos original comprende 1.380 estaciones de aforo. Tras un pormenorizado análisis del número de estaciones disponibles, período que cubren y porcentaje de datos faltantes, se seleccionaron 504 estaciones para el periodo 1975-2008 con menos del 10% de los valores perdidos. Además de la disponibilidad de datos, la calidad y homogeneidad de las series temporales de caudales son requisitos previos para una evaluación fiable de la predictibilidad del caudal estacional. Por lo tanto, se ha investigado la presencia de inhomogeneidades (sobre todo debido a los efectos de regulación) utilizando una metodología combinada basada en el test de Pettitt y el Índice de Área Común. Como resultado, 122 de las anteriores estaciones fueron descartadas debido a la existencia de un comportamiento no natural en las mismas. Estas estaciones se localizan principalmente en las cuencas del Tajo, Guadiana, Guadalquivir, Segura y Júcar. Los puntos de ruptura identificados en dichas series se localizan principalmente en verano, lo cual está probablemente relacionado con aumentos del caudal a partir de la construcción de un embalse en la parte superior del río y la gestión del mismo con el fin de asegurar la disponibilidad de agua durante los meses más secos. En ocasiones, este cambio va acompañado de una disminución del caudal durante los meses de invierno (para el almacenamiento de agua en los embalses), que en algunos casos es lo suficientemente significativo como para dar lugar a una gran alteración de ciclo anual. Finalmente, los huecos encontrados en las 382 estaciones restantes se han rellenado usando métodos de regresión lineal.

Una vez preparada la base de datos, la primera parte de este trabajo tiene el objetivo de explorar las características espaciales y temporales de la variabilidad del caudal de los ríos ibéricos. En este sentido, las estaciones situadas en el oeste de la península (Cantábrico, Miño-Sil, Duero, Tajo, Guadiana, Guadalquivir, Júcar y alto Ebro) presentan sus valores más altos de caudal en invierno. Estas estaciones generalmente presentan una estacionalidad muy marcada (con valores más altos en invierno y más pequeños en verano). Por otro lado, las cuencas del Júcar y Segura presentan un régimen de caudal similar durante todo el año, lo que se debe principalmente a efectos de regulación. Las estaciones de la cuenca del Ebro y también algunas estaciones que pertenecen a las cuencas internas catalanas (las más cercanas a los Pirineos) muestran sus valores de descarga máxima en primavera (debido a las contribuciones del deshielo). Además, hay algunas estaciones ubicadas en la vertiente mediterránea (en particular las de la cuenca mediterránea andaluza) cuyos máximos de caudal se alcanzan en otoño (como consecuencia del mayor aporte de la precipitación de carácter convectivo en esa zona, que sucede a final del verano y principios de otoño).

En cuanto a la variabilidad espacial de los caudales en el otoño y el invierno (las estaciones más húmedas), ésta presenta patrones espaciales similares según los resultados del análisis PCA, agrupando las estaciones de las cuencas del Miño-Sil, Duero, Tajo, Guadiana y Guadalquivir en el primer modo, las estaciones del Cantábrico y Alto Ebro en el segundo modo, y estaciones de las cuencas mediterránea andaluza e internas catalanas en el tercer y cuarto modo, respectivamente. Por otra parte, estos patrones espaciales difieren más en la primavera y el verano (la mitad más seca del año), revelando un claro gradiente norte-sur.

Acerca del análisis de la variabilidad temporal del caudal estacional, en general, predominan las tendencias negativas, especialmente en las estaciones localizadas en cuencas del norte y del este, con algunas excepciones. Por ejemplo, las tendencias positivas encontradas en otoño en la cuenca del Duero, que podrían deberse a un aumento en su variabilidad observada en la segunda mitad del período de estudio. Las tendencias negativas encontradas en el invierno y la primavera podrían estar relacionadas con la persistencia de la fase positiva de la NAO durante las dos últimas décadas del siglo XX, lo que se traduce en una menor precipitación y nieve durante el invierno en la mayor parte de la península. Por el contrario, las tendencias positivas encontradas en verano es probable que sean una consecuencia de la regulación hidrológica para garantizar las necesidades de agua para abastecimiento urbano y regadíos.

La segunda parte de esta Tesis ha consistido en la identificación de los principales factores climáticos que influyen notablemente en la variabilidad del caudal en estaciones posteriores (con un retraso de entre una y cuatro estaciones). En este caso, el verano no ha sido considerado debido a que tanto el promedio del caudal como su variabilidad presentan sus menores valores en la mayoría de las estaciones de aforo. Por otra parte, otro objetivo abordado en este apartado ha sido dar una idea de los posibles mecanismos y procesos físicos que sirven como base de estas relaciones. Una vez identificados, han sido utilizados como predictores en un experimento de predicción estacional del caudal. La metodología de predicción se ha basado en la construcción de modelos de regresión lineal múltiple, y se ha desarrollado para cuatro escenarios de predicción, de acuerdo con el número de estaciones previas con las que se realiza la predicción, desde cuatro estaciones (escenario 4S) hasta una estación (escenario 1S) de adelanto, llevándose a cabo una actualización y mejora de las predicciones a medida que se avanza en el tiempo hacia el caudal a predecir. Este enfoque se convierte en una herramienta útil para la toma de decisiones relacionadas con la gestión del agua, ya que las predicciones se dan con la antelación correspondiente al escenario empleado.

En primer lugar se han evaluado como potenciales predictores los índices de teleconexión que resumen los principales modos de variabilidad del sistema climático, a través de un análisis de la estabilidad de las correlaciones significativas encontradas con el caudal estacional en las estaciones siguientes. Los principales hallazgos se recapitulan como sigue:

- Los índices relacionados con el fenómeno de El Niño en el otoño anterior aparecen como posibles predictores del caudal del otoño siguiente para estaciones de aforo ubicadas en el cuadrante noroeste de la península. Sin embargo, la calidad de las predicciones obtenidas no es muy alta. Por el contrario, el caudal de otoño en la cuenca mediterránea andaluza se predice mejor (RHO> 0.5, RMSESS> 20% y GSS> 0,4) con tres estaciones de antelación, cuando se utiliza la información sobre el SAI del octubre anterior y el índice NAO del invierno anterior. Se obtienen predicciones moderadas (RHO> 0,44, RMSESS> 10% y GSS> 0,2) en los escenarios 3S y 2S para estaciones ubicadas en las cuencas del norte de la península (Miño-Sil, Duero, Ebro y Tajo), cuando los índices EP-NP y WeMO de la primavera anterior se añaden a los índices del ENSO del otoño anterior como predictores.
- En relación con la predicción del caudal de invierno, en los escenarios 4S, 3S y 2S, sólo algunas estaciones ubicadas en el sector mediterráneo fueron susceptibles de predicción, aunque la mayoría de ellas presentan pobre habilidad predictiva. Sin embargo, en un número significativo de estaciones (168) se presentan predicciones moderadas o buenas (con RHO > 0,44, RMSESS > 10% y GSS > 0,2) para el escenario 1S, especialmente en las cuencas situadas en el centro y oeste peninsular. El SAI del octubre previo se identifica como el principal predictor para el caudal de invierno. Como consecuencia, los índices de teleconexión no son capaces de predecir los caudales de invierno con más de dos meses de antelación (cuando el SAI de octubre se utiliza como predictor). La capacidad predictiva del SAI sobre el caudal de invierno en la mayor parte de la península (excepto para estaciones en las vertientes cantábrica y mediterránea), se suma a la ya encontrada sobre el caudal del otoño siguiente en las estaciones de aforo de la cuenca mediterránea andaluza.
- Para el caudal de primavera, solamente en 17 estaciones, en el escenario más cercano (1S), se obtienen predicciones razonablemente buenas. A pesar de que varios índices relacionados con la variabilidad atmosférica u oceánica, como los índices Niño1+2, Niño3, Niño3.4, IOD, PNA, EP-NP y AO de verano presentan correlaciones estables con estaciones en el norte de la península, las predicciones obtenidas se clasifican como

pobres en la mayoría de los casos en los que se identifica cierta habilidad predictiva (46 de 52). La calidad de la predicción aumenta ligeramente en el escenario 1S (especialmente en la mitad sur), cuando se añaden los índices SCAND y NAO del invierno anterior como predictores.

En segundo lugar, se ha empleado la técnica de descomposición en valores singulares (SVD) para identificar y aislar los principales modos de variabilidad acoplada entre el caudal estacional y las variables climáticas (temperatura de la superficie del mar, altura geopotencial a 500 hPa en el Hemisferio Norte y la temperatura y precipitación globales) que preceden entre una y cuatro estaciones al caudal. Este procedimiento permite establecer nuevos modos de variabilidad acoplada adicionales a los representados por los índices de teleconexión, que pueden ser utilizados como predictores del caudal estacional. Los párrafos siguientes resumen los modos significativos de variabilidad acoplada encontrados, así como su capacidad como posibles predictores de caudal estacional.

- Se han utilizado hasta 13 modos de variabilidad acoplada para predecir el caudal de otoño (la mayoría de ellos asociados con la variabilidad climática en el otoño y el invierno previos). De hecho, los resultados de predicción para el escenario 3S (con información del otoño y el invierno previos) son bastante aceptables (42 estaciones con RHO > 0,44, RMSESS > 10% y GSS > 0,2), en particular en la cuenca mediterránea andaluza, con Zdjf2 y Adjf2 (relacionados con la NAO de invierno) como predictores, y en el norte-noroeste de la península (cuencas del Duero, Miño-Sil, Cantábrico y alto Ebro), con Rson1, Adjf1 y Rdjf1 (vinculados al fenómeno de El Niño) como principales predictores.
- Sólo 3 modos de variabilidad acoplada entre las variables climáticas y el caudal de invierno fueron seleccionados como posibles predictores. Debido a ello, sólo unas pocas estaciones en las cuencas mediterránea andaluza e internas catalanas presentan cierta habilidad predictiva, siendo Adjf2 (relacionado con la NAO de invierno) y Pdjf3 (relacionado con El Niño Modoki) los predictores utilizados.
- En el caso del caudal de primavera, 8 modos de variabilidad acoplada se han utilizado como predictores. Cuatro de ellos están relacionados con la variabilidad climática de verano, Pjja1 (relacionado con el fenómeno de El Niño), Zjja1 (parecido al Modo Anular del Norte en verano), Tjja1 y Tjja2. Los modelos de predicción basados en estos cuatro predictores (escenario 3S) proporcionan algunas predicciones hábiles (hasta 21 estaciones con RHO > 0.44, RMSESS > 10% y GSS > 0.2) en el cuadrante noreste de la

península (especialmente en las cuencas del Miño-Sil y del Duero). La calidad de la predicción aumenta en la cuenca del Tajo cuando Tdjf1 (relacionado con la NAO de invierno) se añade al conjunto de predictores en el escenario 1S.

En resumen, el uso de SVD mejora la capacidad predictiva del caudal de otoño y primavera con respecto al uso de los índices de teleconexión como predictores. Esta mejora, probablemente debida a la inclusión de información adicional del sistema climático que no se refleja en los índices de teleconexión, es particularmente relevante para el escenario 3S. Por el contrario, no se obtienen predicciones para el caudal de invierno en la mayoría de las estaciones como consecuencia del bajo número de predictores obtenido. En este caso, el SAI de otoño parece ser el único predictor fiable.

En conclusión, este estudio puede ser ventajoso en comparación con estudios anteriores que han examinado la potencial predictibilidad de la variabilidad del caudal estacional de los ríos de la Península Ibérica asociada a la variabilidad climática de gran escala en varios aspectos. En primer lugar, se basa de un conjunto de datos de series temporales suficientemente largas, con buena cobertura espacial y cuya calidad ha sido comprobada, lo que permite describir con una alta resolución espacial el uso potencial de las diferentes señales climáticas como predictores del caudal estacional en distintas áreas de la península. Este hecho se hace interesante para la toma de decisiones locales en la gestión de los recursos hídricos, destacando la necesidad de adoptar políticas y planes de desarrollo en una escala más local capaz de satisfacer las demandas futuras. A diferencia de este trabajo, otros estudios previos se han limitado al análisis de un pequeño número de estaciones de aforo. En segundo lugar, este estudio examina los vínculos entre la señal y la variabilidad climática del caudal de los ríos de la península, no sólo evaluando los índices climáticos más utilizados, sino también explorando nuevas relaciones entre la variabilidad climática y los caudales en estaciones posteriores. En tercer lugar, este estudio puede proporcionar una visión más completa de la relación entre la variabilidad del clima y el caudal en escalas de tiempo estacionales que puede contribuir de manera significativa al pronóstico de los caudales de los ríos (con diversos planes de previsión, de acuerdo con el tiempo de antelación con el que se hacen las predicciones). En particular, el uso de un enfoque basado en predicciones actualizadas estacionalmente permite la opción de desarrollar políticas de gestión del agua con algunas estaciones de antelación y con la posibilidad de modificar o ajustar estas estrategias conforme se actualizan las predicciones. La disponibilidad de predicciones adecuadas con varios meses de antelación proporciona una información muy útil para el establecimiento de estrategias de gestión del agua.

**Trabajo futuro**

Hay varias áreas de la investigación propuesta que se pueden mejorar de manera significativa en futuros trabajos, de acuerdo con las siguientes consideraciones:

- Teniendo en cuenta que la cobertura temporal de la versión 'diaria' del SAI (que presenta una mayor correlación con el caudal de invierno que la versión 'semanal') sigue siendo corta, se puede concluir que conforme aumente la longitud temporal disponible del mismo, aumentará la calidad de las predicciones. Sin embargo, el carácter estacionario de la relación entre SAI de octubre y AO/NAO de invierno (recientemente cuestionado) ha de ser revisado en el futuro.
- Se requiere una investigación más profunda para mejorar la comprensión de la relación entre la NAO de invierno y el caudal del otoño siguiente en la cuenca mediterránea andaluza, así como el papel que juega la temperatura de la superficie del mar Mediterráneo en verano y otros mecanismos relacionados.
- Además del SAI, el papel que juega la criosfera y en particular, la variabilidad el hielo marino del Ártico en verano, en el caudal de los ríos de la península en estaciones posteriores es una cuestión que requiere más investigación. Concretamente, las relaciones entre la disminución del hielo marino y el Modo Anular del Norte en verano y su efecto sobre el clima del invierno siguiente.
- Cuando la longitud de las series de caudal disponibles sea mayor, se podrán llevar a cabo investigaciones adicionales con el fin de explorar las relaciones no lineales entre el fenómeno de El Niño y el caudal, analizando por separado la respuesta del caudal a años de ocurrencia del fenómeno El Niño, años neutros y años de La Niña. Asimismo, la modulación que modos de variabilidad de baja frecuencia como la PDO o la AMO puedan originar en las condiciones del ENSO o la NAO, junto con su repercusión en la variabilidad del caudal estacional, podrían ser evaluados.
- Además, series de tiempo más largas son más apropiadas para la aplicación de diferentes metodologías de predicción, como las basadas en la probabilidad de ocurrencia de eventos, de manera que se puedan explorar más a fondo las relaciones entre los predictores y predictandos.
- Con la base establecida aquí, predicciones en escalas espaciales más pequeñas (de cuenca o subcuenca) pueden mejorarse teniendo en cuenta los efectos de variables tales como la cobertura del suelo, la contribución del deshielo, los factores de humedad del suelo o la topografía, así como con el uso de modelos hidrológicos.

# APPENDIX A

# STATISTICAL TESTS

## A.1 Lilliefors test

The Lilliefors test (LF), a normality test, is a modification of Kolmogorov-Smirnov used when the distribution parameters have been fitted to the same data used in the test. Given a sample of n observations, the Lilliefors statistic is defined as (Lilliefors, 1967):

$$D = max_x|S_n(x) - F^*(x)| \qquad \text{(a.1)}$$

where $S_n(x)$ is the empirical cumulative probability, estimated as $F_n(x_{(i)})=i/n$, for the $i^{th}$ smallest data value; and $F^*(x)$ is the theoretical cumulative distribution function evaluated at $x$, with $\mu = \bar{X}$ , the sample mean, and $s^2$, the sample variance.

Although the Lilliefors statistic is similar to the Kolmogorov-Smirnov statistic, the table for the critical values is different which leads to a different conclusion about the normality of the data (Razali and Wah, 2011). If D exceed the 1-$\alpha$ quantile as given by the table of quantiles, then the null hypothesis (sample data follow a normal distribution) can be rejected at the level of significance $\alpha$.

## A.2 Jarque-Bera test

The Jarque-Bera (JB) is a normality test where the departure from normality is detected by evaluating the sample moments of skewness and kurtosis. The JB statistic is given as

$$JB = n\left(\frac{\left(\sqrt{b_1}\right)^2}{6} + \frac{(b_2-3)^2}{24}\right) \quad \text{(a.2)}$$

where $\sqrt{b_1}$and $b_2$ are sample skewness and kurtosis, respectively.

## A.3 Shapiro-Wilk test

Shapiro and Wilk (1965) test (SW) for normality was originally restricted for sample size of less than 50. Given an ordered random sample, $y_1 < y_2 < \ldots < y_n$, the original Shapiro-Wilk test statistic is defined as,

$$W = \frac{\left(\sum_{i=1}^{n} a_i y_i\right)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2} \quad \text{(a.3)}$$

where $y_i$ is the $i^{th}$ order statistic

$\bar{y}$ is the sample mean,

$$a_i = (a_1, \ldots, a_n) = \frac{m^T V^{-1}}{\sqrt{m^T V^{-1} V^{-1} m}} \quad \text{(a.4)}$$

And $m_i = (m_1, \ldots m_n)^T$ are the expected values of the order statistics of independent and identically distributed random variables sampled from the standard normal distribution and $\boldsymbol{V}$ is the covariance matrix of those order statistics.

The value of $W$ lies between zero and one. Small values of $W$ lead to the rejection of normality whereas a value of one indicates normality of the data. This study use an approximation (algorithm AS R94) given by Royston (1992) improving the weights a, which can used for any n in the range $3 \le n \le 5000$.

## A.4 Root Mean Square Error Skill Score (RMSESS)

The Root Mean Square Error is a typical magnitude for forecast errors, sensible to large errors, and is given by:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(f_i - o_i\right)^2} \quad \text{(a.5)}$$

being $f_i$ and $o_i$ each pair of forecasted and observed values and n the sample size.

Range: 0 to ∞. Perfect score: 0.

The *RMSESS* is given by:

$$RMSESS = \frac{RMSE - RMSE_{cli}}{RMSE_{perf} - RMSE_{cli}} \quad \text{(a.6)}$$

In this case, the *RMSE* is interpreted as a percentage improvement over the reference forecasts (climatology).

Where $RMSE_{perf}$ is the value related to a perfect forecasts (zero). If $RMSE = RMSE_{perf}$, the skill score attains its maximum value of 100%. If $RMSE = RMSE_{cli}$ then $RMSESS = 0\%$, indicating no improvement over the climatology (reference forecasts). If the forecasts being evaluated are inferior to the climatology, $RMSESS < 0\%$, implies that model is worst than reference forecast.

## A.5 Gerrity Skill Score (GSS)

Gerrity (1992) has suggested a family of equitable, in the sense of Gandin and Murphy (1992), skill scores that are also sensitive to distance in this sense and appear to provide generally reasonable results for rewarding correct forecasts and penalizing incorrect ones (Livezey, 2003).

In the definition of Gerrity Skill Score (GSS) a scoring matrix $s_{ij}$ is used, which is a tabulation of the reward or penalty every forecast/observation outcome (represented by a contingency table) will be accorded:

$$GSS = \sum_{i=1}^{K}\sum_{j=1}^{K} p_{ij} s_{sj} \quad \text{(a.7)}$$

where the weights are calculated as follow:

$$s_{ij} = b\left(\sum_{r=1}^{i-1} a_r^{-1} + \sum_{r=i}^{K-1} a_r\right) \quad \text{(a.8)}$$

$$s_{ij} = b\left(\sum_{r=1}^{i-1} a_r^{-1} - (j-i) + \sum_{r=i}^{K} a_r\right); 1 \le i < j \le K \quad \text{(a.9)}$$

$$s_{ji} = s_{ij} \quad \text{(a.10)}$$

$$b = \frac{1}{K-1} \quad \text{(a.11)}$$

being

$$a_i = \frac{1 - \sum_{r=1}^{i} p_r}{\sum_{r=1}^{i} p_r} \qquad \text{(a.12)}$$

with the sample probabilities (observed frequencies) given by $p_i = N(o_i)/N$.

Equation $s_{ii}$ gives more credit for correct forecasts of rare events and less credit for correct forecasts of common events. Equation $s_{ij}$ also accounts for the intrinsic rarity of the $j$ events, and increasingly penalizes errors for greater differences between the forecast category $i$ and the observed category $j$, through the penalty term ($j$-$i$)

Range: -1 to 1, 0 indicates no skill. Perfect score: 1

The GSS does not depend on the forecast distribution, and is equitable (i.e., random and constant forecasts score a value of 0). Additionally, *GSS* does not reward forecast of common events, but rather rewards forecasts for correctly predicting of rare events. Smaller errors are penalized less than larger forecast errors. This is achieved through the use of the scoring matrix.

# APPENDIX B

# STANDARDIZATION

## B.1 Introduction

In the methodology used to standardize the seasonal streamflow time series, a major step is the determination of the probability distribution function that better fits to the time series. To this end, a set of the most commonly probability distribution function used in hydrology was evaluated: Generalized Extreme Value (GEV), Generalized Pareto (GP), Generalized Logistic (GLO), 3-parameters Log-normal (LN3) and Pearson type III (PE3). These distributions are three parametric (scale, location and shape), which let them be more flexible and fit better to the data series.

The first step consists on assigning a cumulative probability, $F(x)$, to each $x$ value. The expression of $F(x)$ for each distribution used is shown in Table B1. As it can be seen in Table B1, the expression of $F(x)$ depends on the distribution parameters (scale, shape and location). So, prior to calculate $F(x)$, the distribution parameters have to be calculated. To do that, the L-moments approach was used (a brief description of this method is given below). Once the L-moments are calculated, the parameters of the distribution and the cumulative probability for each value are determined according with the formulas displayed in Table B1. Finally, the standardized time series ($z$ scores) are obtained by applying the formula proposed in Abramowitz and Stegun (1964):

$$z = -\left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right) \quad for\ 0 < F(x) \leq 0.5 \quad \text{(b.1)}$$

$$z = +\left(t - \frac{c_0 + c_1 t + c_2 t^2}{1 + d_1 t + d_2 t^2 + d_3 t^3}\right) \quad for\ 0.5 < F(x) < 1 \quad \text{(b.2)}$$

where

$$t = \sqrt{ln\left[\frac{1}{(F(x))^2}\right]} \qquad for\ 0 < F(x) \leq 0. \qquad \text{(b.3)}$$

$$t = \sqrt{ln\left[\frac{1}{(1-F(x))^2}\right]} \quad for\ 0.5 < F(x) < 1 \qquad \text{(b.4)}$$

being $F(x)$ the cumulative distribution function, and

$c_0 = 2.515517$ $c_1 = 0.802853$ $c_2 = 0.010328$

$d_1 = 1.432788$ $d_2 = 0.189269$ $d_3 = 0.001308$

## B.2 Calculation of the parameters. L-moments approach

The parameters (scale, shape and location) of the five previously mentioned distributions have been estimated using the L-moments method. This approach, introduced by Hosking (1990) is increasingly being used by hydrologists. Although a brief summary is given here about the L-moments method, more details and properties can be found in Hosking and Wallis (1997).

L-moments, $\lambda_r$, are linear combinations of Probability Weighted Moments (PWMs) introduced by Greenwood et al. (1979). The reason behind use linear combinations of PWMs instead the original PWMs is because the PWMs are difficult to interpret directly as measures of the scale and shape of a probability distribution. The PWMs are defined as

$$\beta_r = E[x\{F(x)\}^r] \qquad \text{(b.5)}$$

which can be rewritten as

$$\beta_r = \int_0^1 x(F)F^r dF \quad \text{(b.6)}$$

where $F = F(x)$ is the cumulative distribution function (CDF) for $x$, $x(F)$ is the inverse CDF of $x$ evaluated at the probability $F$, and $r = 0, 1, 2,\ldots$ is a nonnegative integer.

For any distribution, the r$^{th}$ L moment, $\lambda_r$, is related to the r$^{th}$ PWM (Hosking, 1990) via

$$\lambda_{r+1} = \sum_{k=0}^{r} \beta_k (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \quad \text{(b.7)}$$

The first L-Moments are: $\lambda_1$, a measure of the location; $\lambda_2$, a measure of scale; $\lambda_3$, a measure of skewness; and $\lambda_4$, a measure of kurtosis (see Figure B1).

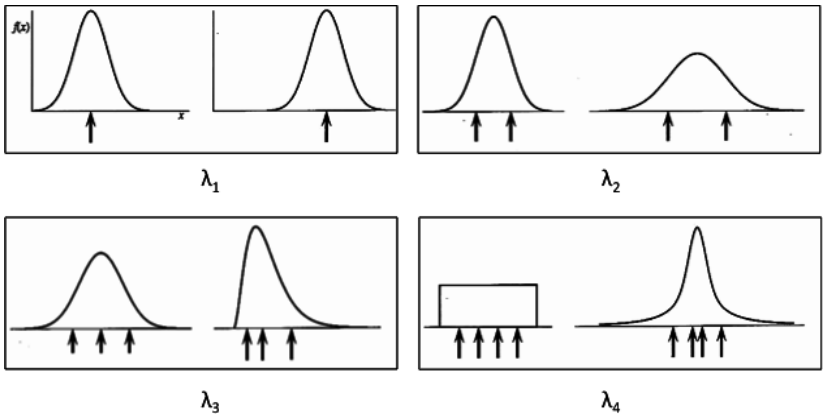


Figure B1. Definition sketch for the first four L-moments.

Although L-moments have been defined for a probability distribution, in practise, they are estimated from a finite sample. Let $x_{1:n} \leq x_{2:n} \leq \ldots \leq x_{n:n}$ be the ordered sample. According with Landwerh et al. (1979) an estimator (unbiased) of the PWMs is:

$$b_r = n^{-1} \binom{n-1}{r}^{-1} \sum_{j=r+1}^{n} \binom{j-1}{r} x_{j:n} \quad \text{(b.8)}$$

This may be alternatively written as

$$b_0 = n^{-1} \sum_{j=1}^{n} x_{j:n} \quad \text{(b.9)}$$

$$b_1 = n^{-1} \sum_{j=2}^{n} \frac{(j-1)}{(n-1)} x_{j:n} \quad \text{(b.10)}$$

$$b_2 = n^{-1} \sum_{j=3}^{n} \frac{(j-1)(j-2)}{(n-1)(n-2)} x_{j:n} \quad \text{(b.11)}$$

$$b_3 = n^{-1} \sum_{j=4}^{n} \frac{(j-1)(j-2)(j-3)}{(n-1)(n-2)(n-3)} x_{j:n} \quad \text{(b.12)}$$

and, in general

$$b_r = n^{-1} \sum_{j=r+1}^{n} \frac{(j-1)(j-2)\ldots(j-r)}{(n-1)(n-2)\ldots(n-r)} x_{j:n} \quad \text{(b.13)}$$

The sample L-moments are defined by:

$$\lambda_1 = b_0 \quad \text{(b.14)}$$

$$\lambda_2 = 2b_1 - b_0 \quad \text{(b.15)}$$

$$\lambda_3 = 6b_2 - 6b_1 + b_0 \quad \text{(b.16)}$$

$$\lambda_4 = 20b_3 - 30b_2 + 12b_1 - b_0 \quad \text{(b.17)}$$

and, in general

$$\lambda_{r+1} = \sum_{j=1}^{n} p^*_{r,k}\, b_k \quad \text{(b.18)}$$

where the $p^*_{r,k}$ coefficients are defined as:

$$p^*_{r,k} = \frac{(-1)^{(r-k)}(r+k)!}{(k!)^2(r-k)!} \quad \text{(b.19)}$$

An useful dimensionless version of L-moments is defined by dividing the higher order L-moments by the scale measure. So, L-moments ratios are defined (Hosking, 1990) as:

$$L - C_v = \tau_2 = \lambda_2/\lambda_1 \quad \text{(b.20)}$$

$$L - skew = \tau_3 = \lambda_3/\lambda_2 \quad \text{(b.21)}$$

$$L - kurtosis = \tau_4 = \lambda_4/\lambda_2 \quad \text{(b.22)}$$

L-moments ratios measure the shape of a distribution independently of its scale of measure. The L-moment ratio $\tau_2$ is the L-coefficient of variation (L-$C_v$), and the L-moment ratios $\tau_3$ and $\tau_4$ are the L-skewness and L-kurtosis, respectively.

In terms of numerical values, $\lambda_1$ can take any value, and $\lambda_2 \geq 0$. The L-$C_v$ is dimensionless measure of variability and, for a distribution that takes only positive values, is in the range $0 \leq \tau_2 < 1$. The L-skew and L-kurtosis are dimensionless measures of asymmetry and kurtosis. The absolute value of L-moment ratios $\tau_3$ and $\tau_4$ is lower than 1.

Despite the samples L-moments are unbiased estimators of L-moments, the L-moment ratios are not unbiased. Their bias are very small for moderate or large samples (Hosking and Wallis, 1997).

The main properties of L-moments (according Hosking 1989) are:

1) Existence: If the average of a distribution exists, then, all the L-moments of the distribution exist.

2) Uniqueness: If the average of a distribution exists, then, the L-moments just characterize that distribution. There are not two distributions with the same values of L-moments.

L-moments approach presents some advantages respect to other estimators. In particular, against the method of ordinary moments (MOM), these advantages are:

- The L moments exist whenever the mean of the distribution exists. This includes cases in which some of the higher moments fail to exist. (Hosking and Wallis, 1997).
- The L-moments are less sensitive to the presence of outliers (Hosking and Wallis, 1997).
- The L-moments provide better identification of the parent distribution (Hosking, 1990), especially for skewed distributions.
- Asymptotic approximations to sampling distributions are better for L-moments than for ordinary moments (Hosking, 1990).
- The L-moments allow characterizing a greater number of distributions. A total of 13 different distributions can be adjusted using the method of L-moments, of which 5 are generally used in regional frequency analysis (Sorman and Okur, 2000).
- The L-moments show little bias (nearly unbiased) for all distributions applied and their advantages over traditional moments are not limited for small samples (Vogel and Fennessey, 1993).

Additionally, L-moments approach also present some advantages respect to maximum likelihood estimator (MLE) method, which cannot be reduced to mere formulas, and is generally calculated using numerical methods. Also, MLE does not work well when the observations significantly deviate from the distribution to be adjusted.

| | RANGE | CUMULATIVE DISTRIBUTION FUNCTION | PARAMENTER ESTIMATION |
|---|---|---|---|
| **GEV** | if $\kappa > 0$: $-\infty<x\leq\xi+\alpha/\kappa$<br>if $\kappa = 0$: $-\infty < x < \infty$<br>if $\kappa < 0$: $\xi+\alpha/\kappa\leq x<\infty$ | $F(x) = e^{-e^{y}}$<br>if $\kappa \neq 0$;<br>$y = -\kappa^{-1}\ln(1-\kappa(x-\xi)/\alpha)$<br>if $\kappa = 0$;<br>$y = (x-\xi)/\alpha$ | $\xi$ (location), $\alpha$ (scale), $\kappa$ (shape)<br>$\kappa \approx 7.8590c + 2.9554c^2$<br>$c = \frac{2}{3+\tau_3} - 0.6309$<br>$\alpha = \frac{\lambda_2\kappa}{(1-2^{-\kappa})\Gamma(1+\kappa)}$<br>$\xi = \lambda_1 - \alpha\frac{1-\Gamma(1+\kappa)}{\kappa}$ |
| **GP** | if $\kappa > 0$: $\xi< x \leq\xi+\alpha/\kappa$<br>if $\kappa \leq 0$: $\xi \leq x < \infty$ | $F(x) = 1 - e^{-y}$<br>if $\kappa \neq 0$;<br>$y = -\kappa^{-1}\ln(1-\kappa(x-\xi)/\alpha)$<br>if $\kappa = 0$; $y = (x-\xi)/\alpha$ | $\xi$ (location), $\alpha$ (scale), $\kappa$ (shape)<br>$\kappa = \frac{1-3\tau_3}{1+\tau_3}$<br>$\alpha = (1+\kappa)(2+\kappa)\lambda_2$<br>$\xi = \lambda_1 - (2+\kappa)\lambda_2$ |
| **LN3** | if $\kappa > 0$: $-\infty<x\leq\xi+\alpha/\kappa$<br>if $\kappa = 0$: $-\infty < x < \infty$<br>if $\kappa > 0$: $\xi+\alpha/\kappa\leq x<\infty$ | $F(x) = f(y)$<br>(standard normal distribution)<br>if $\kappa \neq 0$;<br>$y=-\kappa^{-1}\ln(1-\kappa(x-\xi)/\alpha)$<br>if $\kappa = 0$;<br>$y = (x-\xi)/\alpha$ | $\xi$ (location), $\alpha$ (scale), $\kappa$ (shape),<br>$\kappa \approx -\tau_3\frac{E_0+E_1\tau_3^2+E_2\tau_3^4+E_3\tau_3^6}{1+F_1\tau_3^2+F_2\tau_3^4+F_3\tau_3^6}$<br>$E_0$=2.0466534 $F_1$=-2.0182171<br>$E_1$=-3.6544371 $F_2$=1.2420401<br>$E_2$=1.8396733 $F_3$=-0.21741801<br>$E_3$=-0.2360244<br>$\alpha = \frac{\lambda_2\kappa e^{-\kappa^2/2}}{1-2\Phi(-\kappa/\sqrt{2})}$<br>$\xi = \lambda_1 - \frac{\alpha}{\kappa}(1-e^{\kappa^2/2})$ |

| | | | |
|---|---|---|---|
| **GLO** | if $\kappa > 0$: $-\infty < x \leq \xi + \alpha/\kappa$<br>if $\kappa = 0$: $-\infty \leq x < \infty$<br>if $\kappa < 0$: $\xi + \alpha/\kappa \leq x < \infty$ | $F(x) = \frac{1}{1+e^{-y}}$<br>if $\kappa \neq 0$;<br>$y = -\kappa^{-1}\ln(1-\kappa(x-\xi)/\alpha)$<br>if $\kappa = 0$; $y = (x-\xi)/\alpha$ | $\xi$ (location), $\alpha$ (scale), $\kappa$ (shape)<br>$\kappa = -\tau_3$<br>$\alpha = \frac{\lambda_2 \sin(\kappa\pi)}{\kappa\pi}$<br>$\xi = \lambda_1 - \alpha\left(\frac{1}{\kappa} - \frac{\pi}{\sin(\kappa\pi)}\right)$ |
| **PE3** | If $\gamma \neq 0$,<br>let<br>$\alpha = 4/\gamma^2$, $\beta = (1/2)\sigma\lvert\gamma\rvert$,<br>$\xi = \mu - 2\sigma/\gamma$,<br>if $\gamma > 0$; $\xi \leq x < \infty$<br>if $\gamma < 0$; $-\infty \leq x < \xi$<br>if $\gamma = 0$; $-\infty < x < \infty$ | if $\gamma > 0$;<br>$F(x) = \frac{\int_\xi^x \frac{(x-\xi)^{\beta-1}}{\beta} e^{-\frac{(x-\xi)}{\beta}} dx}{\Gamma(\alpha)}$<br>if $\gamma < 0$;<br>$F(x) = 1 - \frac{\int_\xi^x \frac{(x-\xi)^{\beta-1}}{\beta} e^{-\frac{(x-\xi)}{\beta}} dx}{\Gamma(\alpha)}$<br>If $\gamma = 0$;<br>$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$<br>(the standard normal distribution) | $\mu$ (location), $\sigma$ (scale), $\gamma$ (shape)<br>if $1/3 \leq \lvert\tau_3\rvert < 1$; $z = 3\pi\tau_3^2$<br>$\alpha \approx \frac{1 + 0.2906z}{z + 0.1882z^2 + 0.0442z^3}$<br>if $1/3 \leq \lvert\tau_3\rvert < 1$<br>$\alpha \approx \frac{0.36067z - 0.59567z^2 + 0.25361z^3}{1 - 2.78861z + 2.56096z^2 - 0.77045z^3}$<br>Given $\alpha$, the parameters are:<br>$\gamma = 2\alpha^{-\frac{1}{2}} sign(\tau_3)$<br>$\sigma = \frac{\lambda_2 \pi^{\frac{1}{2}} \alpha^{\frac{1}{2}} \Gamma(\alpha)}{\Gamma(\alpha + 1/2)}$<br>$\mu = \lambda_1$ |

Table B1. Cumulative distribution functions of the Generalized Extreme Value (GEV), Generalized Pareto (GP), Lognormal (LN3), Generalized Logistic (GLO) and Pearson type 3 (PE3), along with the equations used to estimate the parameters according with the L-moments procedure (from Hosking and Wallis, 1997).

## B.3 Choosing an appropriate distribution

The evaluation of which probability distribution function better fits to each seasonal time series at each gauging station was realized through two indices (similar to those used in Vicente-Serrano et al., 2011):

*Minimum Orthogonal Distance*

The L-moments ratio diagram is a graphical measure about if the data samples from different sites are consistent with the fitted probability distribution functions (see Figure B2). The Minimum Distance (MD) is the orthogonal distance between the sample L-moments at a site and the L-moment curve for the distributions described in the Table B1. Following this criterion, the distribution with minimum MD is selected as the more appropriate.

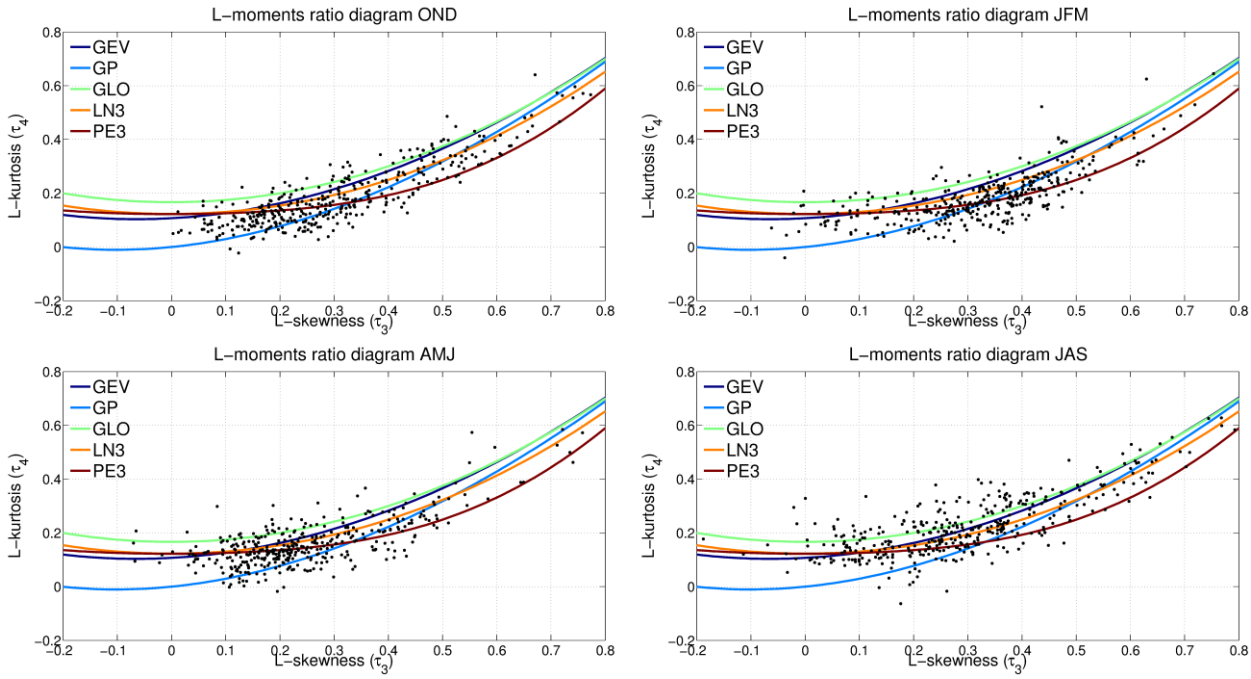


Figure B2. L-moments plot for samples and theoretical distributions used. The fitted distribution curves were drawn using the polynomial approximations given by Hosking (1990).

*The Kolmogorov-Smirnov statistic*

The Kolmogorov-Smirnov (KS) statistic is a measure of the difference between the empirical distribution and the cumulative distribution function $F(x)$, for a given probability distribution. The KS statistic is defined as the maximum vertical difference between the empirical and the cumulative distribution functions:

$$KS = \max(max_i|F(x_i) - \frac{r-1}{n}, \frac{r}{n} - F(x_i)|) \qquad \text{(b.23)}$$

where $r$ is the rank of the $i$ observation in ascending order.

According with this criterion, the distribution that presents the smaller *KS* statistic is selected.

It should be mentioned that, when a particular distribution could not be defined for some values (according the range of values specified in Table B1), this distribution was not selected. These *'out of range'* values use to appear when time series present large peaks, which do not follow the same theoretical distribution as the remaining data.

Hence, the distribution that presents the best fit, according to KS and MD stats, was chosen. To do that, the distributions were classified according values of KS and MD stats. The distribution with the best *'combined position'*, taking into account both statistics, was selected. In case the existence of *'out of range'* values for the selected distribution, the next distribution was chosen instead. Table B2 illustrates this method for a station selected as example (winter streamflow time series for station id=2101, in Douro Basin), which was standardized according the Pearson 3 distribution.

| | **KS stat** | **MD stat** | **KS position** | **MD position** | ***Combined position*** | **Out of range values** |
|---|---|---|---|---|---|---|
| **GEV** | 0.0838 | 0.0526 | 3 | 4 | 7 | 0 |
| **GP** | 0.0935 | 0.0154 | 4 | 2 | 6 | 1 |
| **GLO** | 0.0995 | 0.0833 | 5 | 4 | 9 | 0 |
| **LN3** | 0.0762 | 0.0398 | 2 | 3 | 5 | 0 |
| **PE3** | 0.0713 | 0.0134 | 1 | 1 | 2 | 0 |

Table B2. Example of the procedure followed to choose the distribution fitting better to the winter streamflow time series for station id=2101 (Douro Basin), using the KS and MD statistics.

Additionally, following with the previous example, Figure B3 illustrates the next steps carried out to standardize the time series, once the appropriated distribution (Pearson 3 in this example) was selected. To each data, a cumulative probability is assigned (according the CDF of Pearson type 3 fitted to the data series). The value of a normal distribution that corresponds to

this cumulative probability is given to the original data as its 'z-score'. In a red circle in the Figure B3 is shown a particular case.

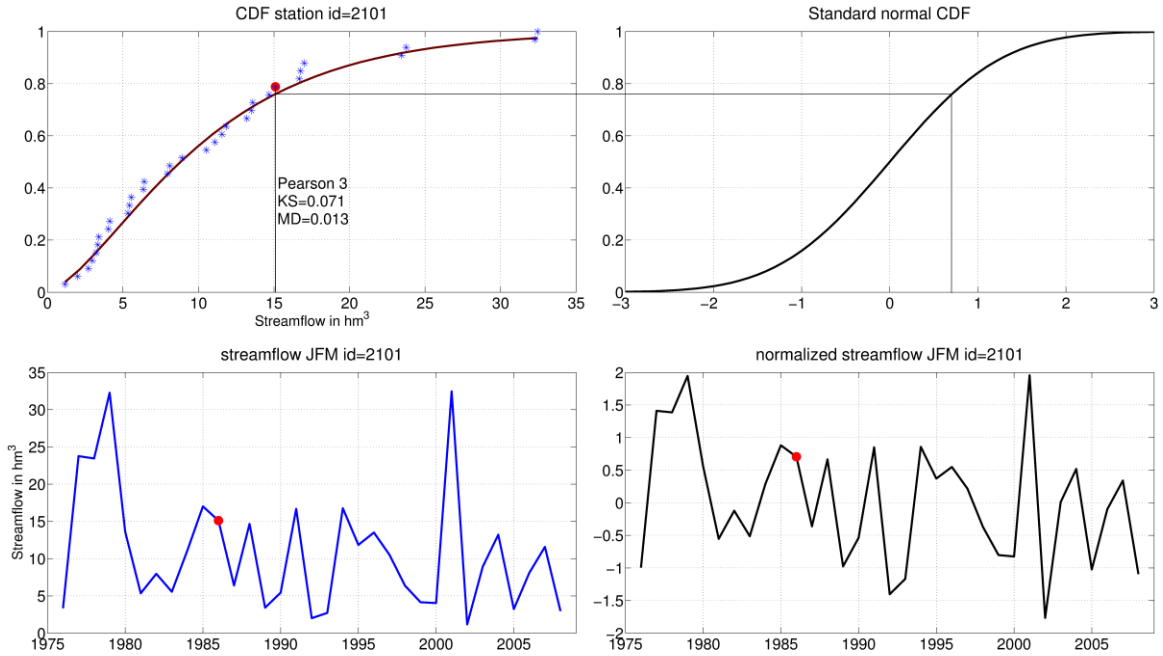


Figure B3. In top left panel, the empirical CDF of data series (in dots), and the CDF corresponding to the Pearson 3 distribution that fitted to the data series in a red line. Top right panel represent the CDF for a standard normal distribution. Bottom panels are the original winter streamflow time series (left) and the correspondent standardized time series.

Taking into account the above-described procedure, finally, Figure B4 displays, for each season, the probability distribution function selected for each station. Seasonal maps are accompanied with superposed bar graphs summarizing the number of stations that follows each distribution. This figure shows that there is not a distribution that presents the best fit to all stations. Also, despite there is no clear grouping of stations fitted by a particular distribution, it could be remarked as Log-normal 3 fits an important amount of stations in winter.

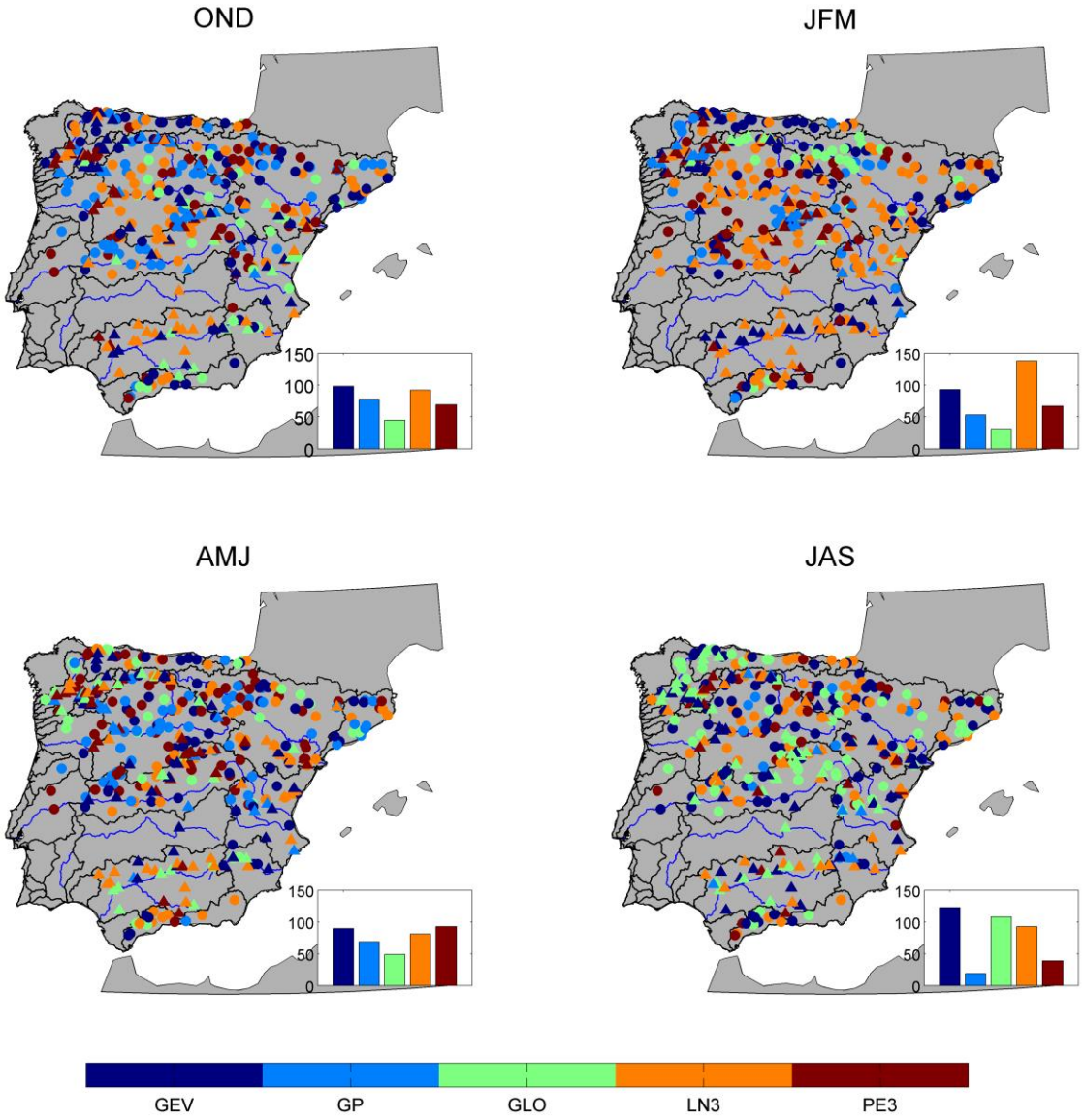


Figure B4. For all seasons, probability distribution functions selected to standardize seasonal streamflow time series and bar plots summing up the number of stations that follow each probability distribution function.

# APPENDIX C

## PRINCIPAL COMPONENT ANALYSIS OF CLIMATE FIELDS

In this Appendix, the main modes of seasonal variability of climate fields were isolated using Principal Components Analysis. A brief description of this methodology is provided in Chapter 4 and further details can be obtained in Preisendorfer (1988). This main modes of variaiblity are compared with the main SVD modes identified in Chapter 6 between seasonal streamflow and the different climate variables.

The large-scale atmospheric and oceanic variables, covering from September of 1974 to February of 2008, were obtained from the following datasets:

- The monthly data for the sea surface temperature (SST) were obtained from The Hadley Centre Global Sea Ice and Sea Surface Temperature (HadISST, Rayner et al., 2003), with a resolution of 1º x 1º. The region of Pacific Ocean SST data used for the analysis is [120°E-70°W, 40°S-70°N], while the region of Atlantic Ocean SST data used is [80°W-30°E, 40°S-70°N].
- The monthly mean geopotential height at 500 hPa (Z500) was downloaded from the NCEP Reanalysis (Kalnay et al., 1996), with a resolution of 2.5º x 2.5º. Only Z500 data northern than 20ºN are used here.

- The monthly values for global precipitation (RR) were downloaded from the Global Precipitation Climatology Centre (GPCC) version 6- Total Full (Schneider et al., 2014). The dataset has a 1º x 1º resolution.
- The monthly global surface temperature data over land (TMP) were taken from the high-resolution database of the Climatic Research Unit (CTU-TS.3.10, Harris et al., 2014), with a 0.5º x 0.5º of spatial resolution.

For computational efficiency, all dataset were regridded into a coarse grid of 2.5º x 2.5º. Seasonal time series were created by averaging the months: September-October-November (SON) for autumn, December-January-February (DJF) for winter, March-April-May (MAM) for spring and June-July-August (JA) for summer.

Figures C.1-C.5 show the 'loading factors' corresponding to each EOF. These loadings factors represent the correlation between the time series of each principal component and the seasonal time series of each field. In the title of each figure is displayed the percentage of explained variance of each mode.

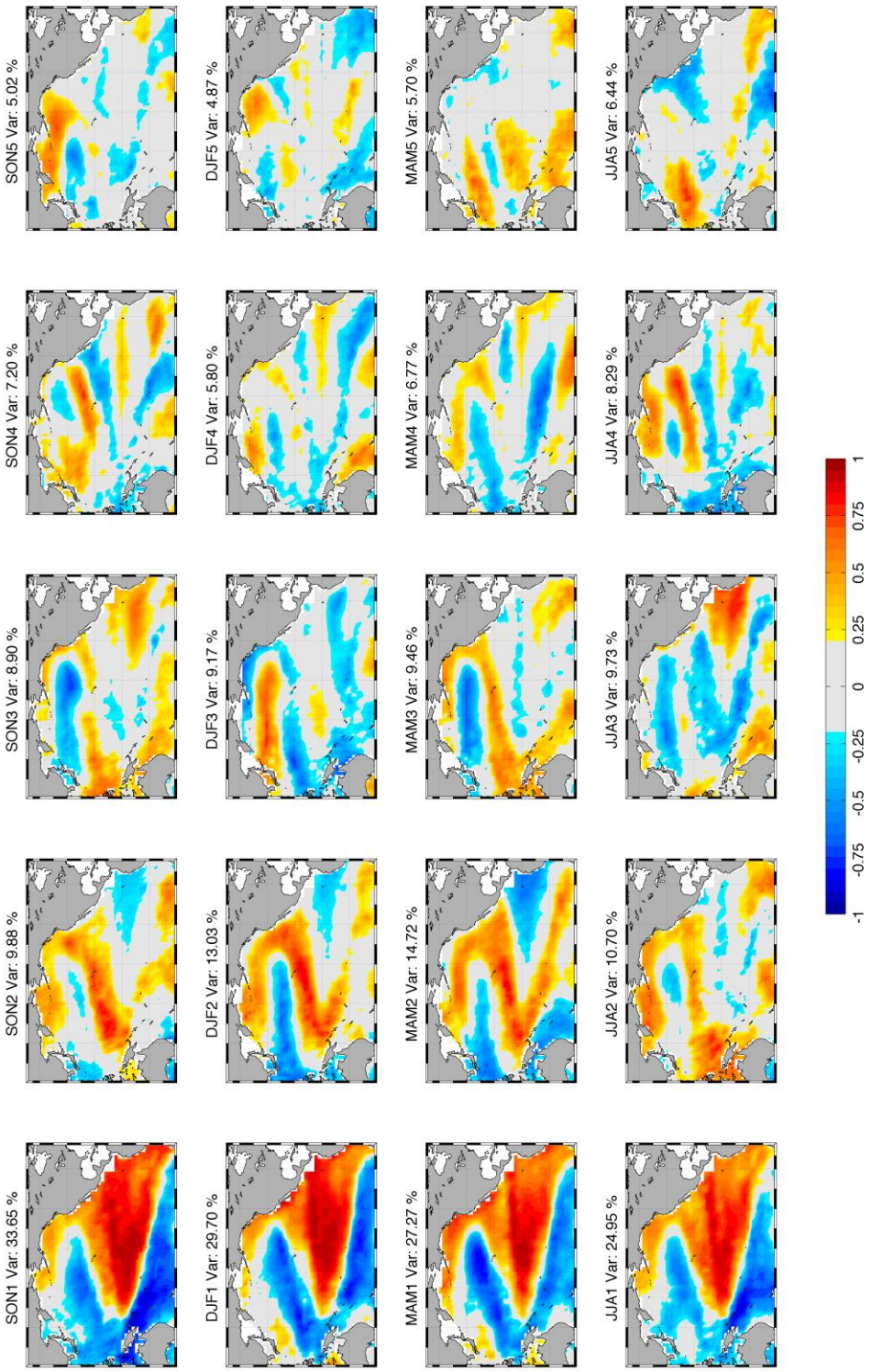


Figure C.1. Loading factors associated with principal components 1-5 (in columns) from the PCA analysis of the seasonal time series of Pacific Ocean surface temperature (in rows).

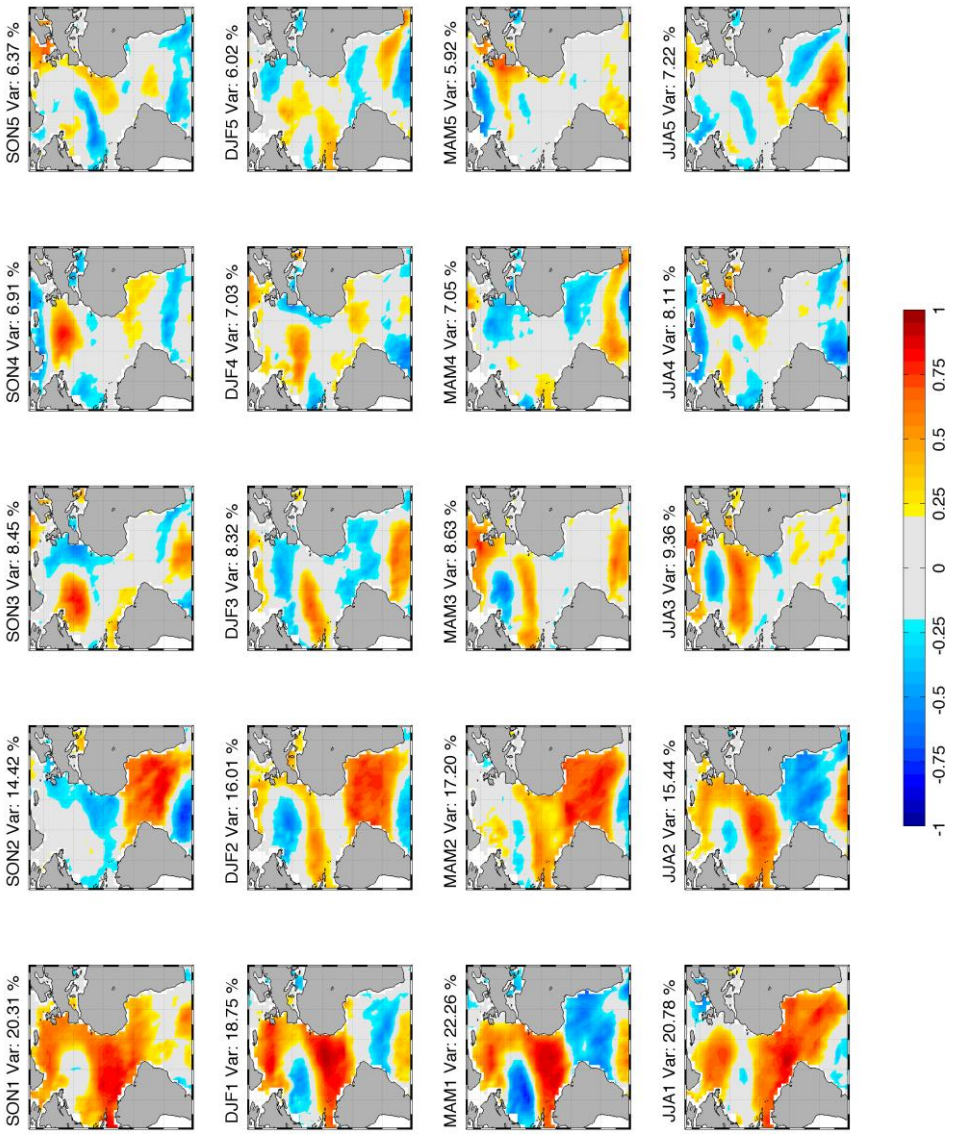


Figure C.2. Loading factors associated with principal components 1-5 (in columns) from the PCA analysis of the seasonal time series of Atlantic Ocean surface temperature (in rows).

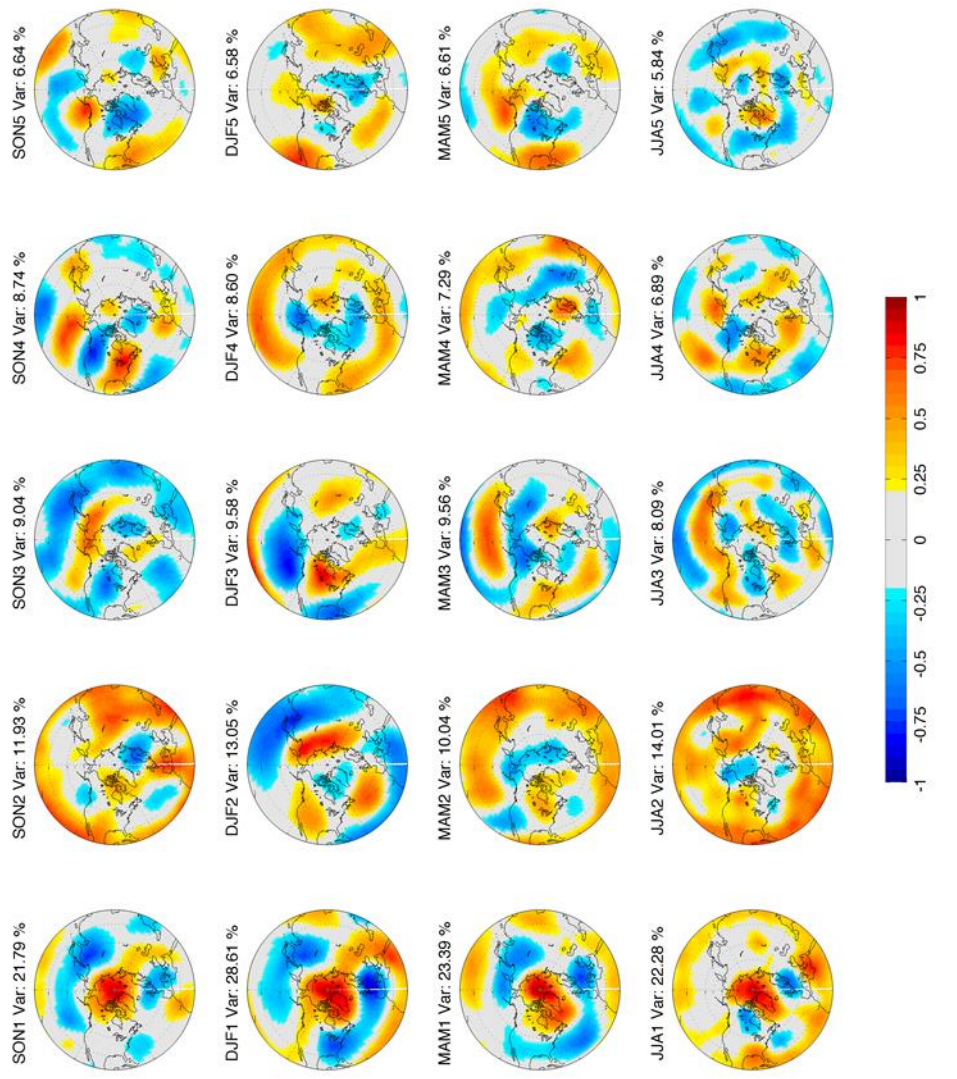


Figure C.3. Loading factors associated with principal components 1-5 (in columns) from the PCA analysis of the seasonal time series of geopotential height at 500 hPa (in rows).

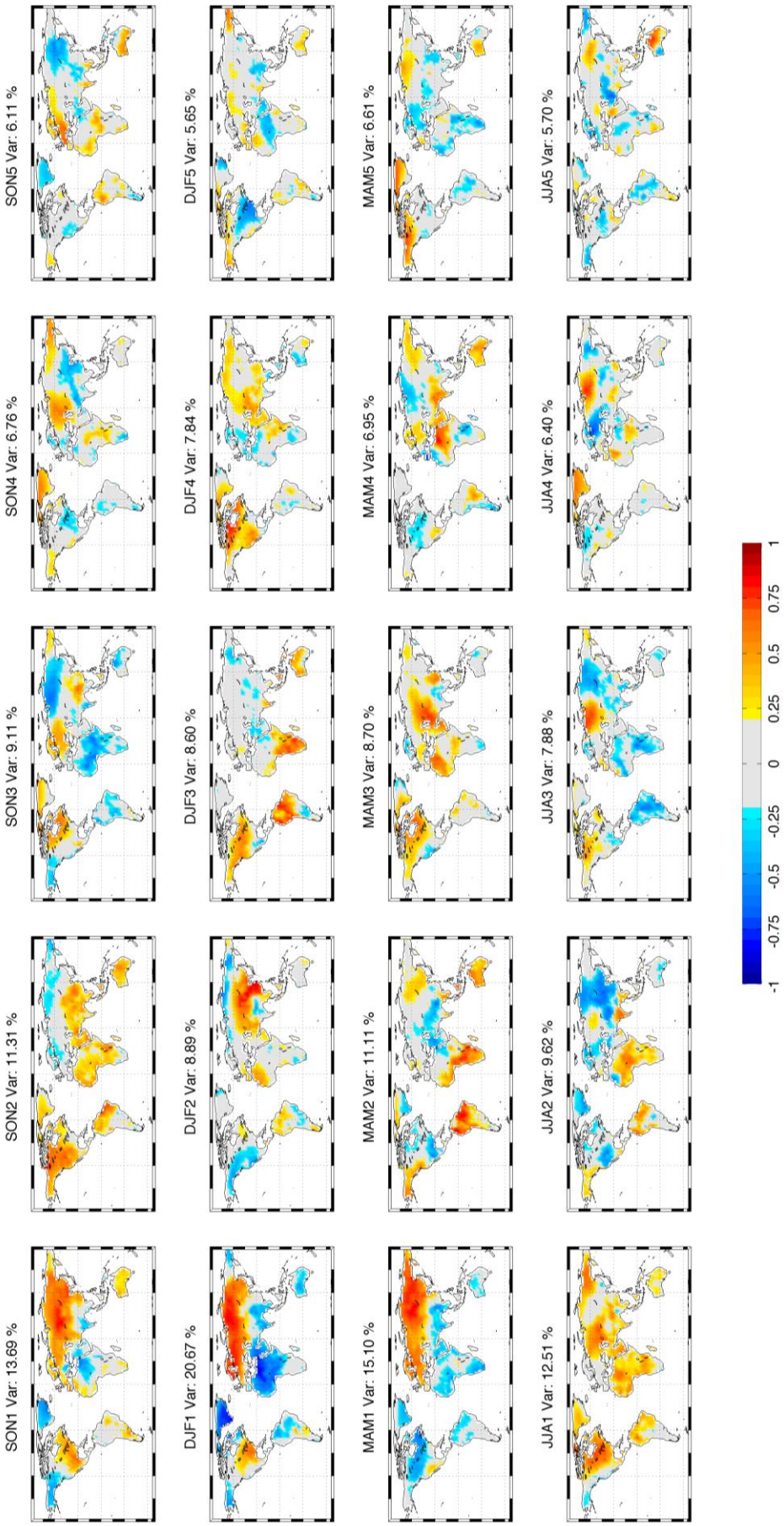


Figure C.4. Loading factors associated with principal components 1-5 (in columns) from the PCA analysis of the seasonal time series of global temperature (in rows).

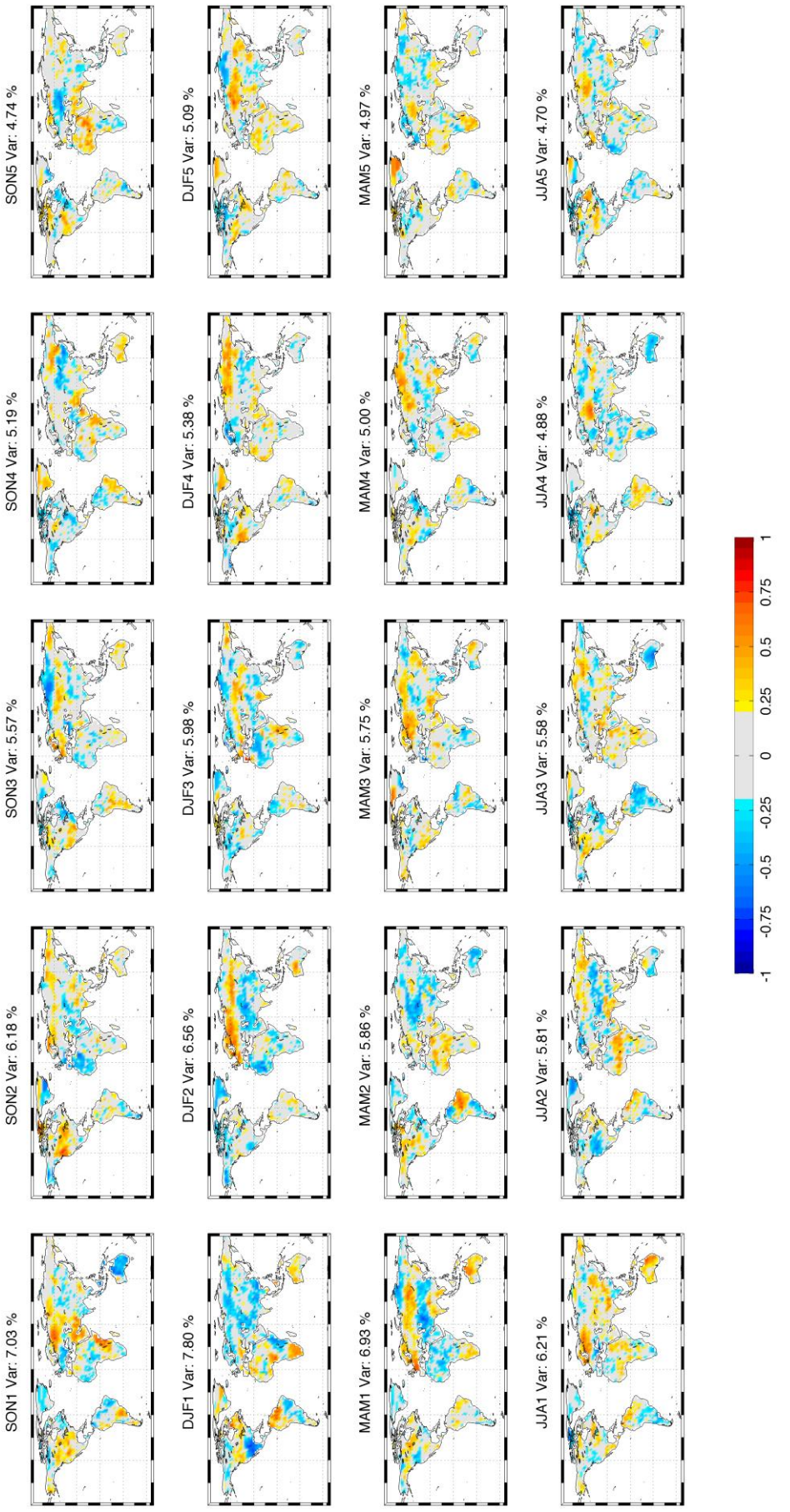


Figure C.5. Loading factors associated with principal components 1-5 (in columns) from the PCA analysis of the seasonal time series of global precipitation (in rows).

# REFERENCES

Abramowitz, M., and I. A. Stegun (1964), Handbook of mathematical functions with formulas, graphs, and mathematical tables. Dover Publications, New York.

AEMET (2011), Atlas Climático Ibérico, Ministerio de Medio Ambiente y Medio Rural y Marino. Agencia Estatal de Meteorología. 80 pp. ISBN: 978-84-7837-079-5

Aguado, E., Cayan, D., Riddle, L., and Roos, M. (1992), Climatic Fluctuations and the Timing of West Coast Streamflow, *Journal of Climate* 5, 1468–1481.

Aka, A. A., E. Servat, J. E. Paturel, B. Kouamé, H. Lubes, and J. M. Masson (1996), Analysis of the temporal variability of runoff in Ivory Coast: statistical approach and phenomena characterization, *Hydrological Sciences Journal*, 41(6), 959-970.

Alexander, L. V., X. Zhang, T. C. Peterson, J. Caesar, B. Gleason, A. M. G. Klein Tank, M. Hayloc, D. Collins, B. Trewin, F. Rahimzadeh, A. Tagipour, K. Rupa Kumar, J. Revadekar, G. Griffiths, L. Vincent, D. B. Stephenson, J. Burn, E. Aguilar, M. Brunet, M. Taylor, M. New, P. Zhai, M. Rusticucci, and J. L. Vazquez-Aguirre (2006), Global observed changes in daily climate extremes of temperature and precipitation, *Journal of Geophysical Research: Atmospheres*, 111(D5), D05109.

Alexander, M. A., I. Bladé, M. Newman, J. R. Lanzante, N.-C. Lau, and J. D. Scott (2002), The atmospheric bridge: The influence of ENSO teleconnections on air-sea interaction over the global oceans, *Journal of Climate*, 15(16), 2205-2231.

Alexandersson, H. (1986), A homogeneity test applied to precipitation data, *Journal of Climatology*, 6(6), 661-675.

Ambrizzi, T., and B. J. Hoskins (1997), Stationary Rossby-wave propagation in a baroclinic atmosphere, *Quarterly Journal of the Royal Meteorological Society*, 123(540), 919-928.

Anderson, S., O. Aziz, G. Tootle, H. Grissino-Mayer, and A. Barnett (2012), Using Pacific Ocean climatic variability to improve hydrologic reconstructions, *Journal of Hydrology*, 434-435(0), 69-77.

Argüeso, D., J. M. Hidalgo-Muñoz, S. R. Gámiz-Fortis, M. J. Esteban-Parra, and Y. Castro-Díez (2012), High-resolution projections of mean and extreme precipitation over Spain using the WRF model (2070–2099 versus 1970–1999), *Journal of Geophysical Research*, 117, D12108.

Arroyo-Ilera, F. (2007), Territorio, tecnología y capital, La regulación hidroeléctrica de los ríos españoles [1900–1970], *Treballs de la Societat Catalana de Geografía*, 63, 39–70.

Ashok, K., S. K. Behera, S. A. Rao, H. Weng, and T. Yamagata (2007), El Niño Modoki and its possible teleconnection, *Journal of Geophysical Research: Oceans*, 112(C11), C11007.

Azevedo, T. M., E. Nunes, and C. Ramos (2004), Some morphological aspects and hydrological characterization of the Tagus floods in the Santarem region, Portugal, *Natural Hazards*, 31(3), 587-601.

Báez, J. C., L. Gimeno, M. Gómez-Gesteira, F. Ferri-Yáñez, and R. Real (2013), Combined effects of the North Atlantic Oscillation and the Arctic Oscillation on sea surface temperature in the Alborán Sea, *PLoS ONE*, 8(4), e62201.

Baik, J. J. and J. S Paek (1998), A climatology of sea surface temperature and the maximum intensity of western North Pacific tropical cyclones, *Meteorological Society of Japan*, Tokyo, JAPON.

Barceló, D. (2008), Aguas continentales: Gestión de recursos hídricos, tratamiento y calidad del agua 280 pp., *Consejo Superior de Investigaciones Científicas*.

Bardin, M. Y., and E. N. Voskresenskaya (2007), Pacific Decadal Oscillation and European climatic anomalies, *Physical Oceanography*, 17(4), 200-208.

Barnett, F., S. Gray, and G. Tootle (2010), Upper Green River Basin (United States) streamflow reconstructions, *Journal of Hydrologic Engineering*, 15(7), 567-579.

Barnston, A. G., and R. E. Livezey (1987), Classification, seasonality and persistence of low-frequency atmospheric circulation patterns, *Monthly Weather Review*, 115(6), 1083-1126.

Batalla, R. J., C. M. Gómez, and G. M. Kondolf (2004), Reservoir-induced hydrological changes in the Ebro River basin (NE Spain), *Journal of Hydrology*, 290(1-2), 117-136.

Bates, B. C., Z. W. Kundzewicz, S. Wu, and J. P. Palutikof (2008), Climate change and water. *Technical Paper of the Intergovernmental Panel on Climate Change*. IPCC Secretariat, 210 pp., Geneva.

Beniston, M. (2009), Trends in joint quantiles of temperature and precipitation in Europe since 1901 and projected for 2100, *Geophysical Research Letters*, 36(7), L07707.

Berga-Casafont, L. (2003), Presas y embalses en la España del siglo XX, *Revista de Obras Públicas*, 3438, 37–40.

Bierkens, M. F. P., and L. P. H. van Beek (2009), Seasonal predictability of European discharge: NAO and hydrological response time, *Journal of Hydrometeorology*, 10(4), 953-968.

Birsan, M. V., P. Molnar, P. Burlando, and M. Pfaundler (2005), Streamflow trends in Switzerland, *Journal of Hydrology*, 314(1-4), 312-329.

Bjerknes, J. (1969), Atmospheric teleconnections from the equatorial Pacific, *Monthly Weather Review*, 97(3), 163-172.

Black, E., and R. Sutton (2007), The influence of oceanic conditions on the hot European summer of 2003, *Climate Dynamics*, 28(1), 53-66.

Bojariu, R., and L. Gimeno (2003), The role of snow cover fluctuations in multiannual NAO persistence, *Geophysical Research Letters*, 30(4), 1156-1156.

Bond, N.A., J.E. Overland, M. Spillane, and P. Stabeno, (2003), Recent shifts in the state of the North Pacific, *Geophysical Research Letters*, 30, doi: 10.1029/2003GL018597.

Bracken, C., B. Rajagopalan, and J. Prairie (2010), A multisite seasonal ensemble streamflow forecasting technique, *Water Resources Research*, 46(3), W03532.

Brandão, C., and R. Rodrigues (2000), Hydrological simulation of the International Catchment of Guadiana River, *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 25(3), 329-339.

Brands, S., S. Herrera, and J. M. Gutiérrez (2013), Is Eurasian snow cover in October a reliable statistical predictor for the wintertime climate on the Iberian Peninsula?, *International Journal of Climatology*, 34, 1615–1627, doi: 10.1002/joc.3788

Brands, S., R. Manzanas, J. M. Gutiérrez, and J. Cohen (2012), Seasonal predictability of wintertime precipitation in Europe using the Snow Advance Index, *Journal of Climate*, 25(12), 4023-4028.

Bretherton, C. S., C. Smith, and J. M. Wallace (1992), An intercomparison of methods for finding coupled patterns in climate data, *Journal of Climate*, 5, 541–560.

Bretherton, C. S., M. Widmann, V. P. Dymnikov, J. M. Wallace, and I. Bladé (1999), The effective number of spatial degrees of freedom of a time-varying field, *Journal of Climate*, 12(7), 1990-2009.

Bulic, I., and F. Kucharski (2012), Delayed ENSO impact on spring precipitation over North/Atlantic European region, *Climate Dynamics*, 38(11-12), 2593-2612.

Buishand, T. A. (1982), Some methods for testing the homogeneity of rainfall records, *Journal of Hydrology*, 58, 11-27.

Burn, D. H., and M. A. Hag Elnur (2002), Detection of hydrologic trends and variability, *Journal of Hydrology*, 255(1-4), 107-122.

Burn, D. H., M. Sharif, and K. Zhang (2010), Detection of trends in hydrological extremes for Canadian watersheds, *Hydrological Processes*, 24(13), 1781-1790.

Casanueva, A., C. Rodríguez-Puebla, M. D. Frías, and N. González-Reviriego (2014), Variability of extreme precipitation over Europe and its relationships with teleconnection patterns, *Hydrology and Earth System Sciences*, 18(2), 709-725.

Cayan, D. R., Kammerdiener, S. A., Dettinger, M. D., Caprio, J. M., and Peterson, D. H. (2001), Changes in the onset of spring in the western United States, *Bulletin of the American Meteorological Society*, 82, 399–415.

Chandimala, J., and L. Zubair (2007), Predictability of streamflow and rainfall based on ENSO for water resources management in Sri Lanka, *Journal of Hydrology*, 335(3-4), 303-312.

Chao, Y., M. Ghil, and J. C. McWilliams (2000), Pacific interdecadal variability in this century's sea surface temperatures, *Geophysical Research Letters*, 27(15), 2261-2264.

Cherry, S. (1996), Singular Value Decomposition Analysis and Canonical Correlation Analysis, *Journal of Climate*, 9(9), 2003-2009.

Chiew, F. H. S., and T. A. McMahon (2002), Global ENSO-streamflow teleconnection, streamflow forecasting and interannual variability, *Hydrological Sciences Journal*, 47(3), 505-522.

Chiew, F. H. S., S. L. Zhou, and T. A. McMahon (2003), Use of seasonal streamflow forecasts in water resources management, *Journal of Hydrology*, 270(1-2), 135-144.

Cohen, J. (2011), Eurasian snow cover variability and links with stratosphere-troposphere coupling and their potential use in seasonal to decadal climate predictions, *Climate Test Bed Joint Seminar Series*. US National Oceanic and Atmospheric Administration.

Cohen, J., and D. Rind (1991), The effect of snow cover on the climate, *Journal of Climate*, 4(7), 689-706.

Cohen, J., and D. Entekhabi (1999), Eurasian snow cover variability and northern hemisphere climate predictability, *Geophysical Research Letters*, 26(3), 345-348.

Cohen, J., K. Saito, and D. Entekhabi (2001), The role of the Siberian High in Northern Hemisphere climate variability, *Geophysical Research Letters*, 28(2), 299-302.

Cohen, J. L., and K. Saito (2003), Eurasian snow cover, more skillful in predicting U.S. winter climate than the NAO/AO?, *Geophysical Research Letters*, 30(23), 2190-2190.

Cohen, J., M. Barlow, P. J. Kushner, and K. Saito (2007), Stratosphere-troposphere coupling and links with Eurasian land surface variability, *Journal of Climate*, 20(21), 5335-5343.

Cohen, J., and C. Fletcher (2007), Improved skill of Northern Hemisphere winter surface temperature predictions based on land-atmosphere fall anomalies, *Journal of Climate*, 20(16), 4118-4132.

Cohen, J., and J. Jones (2011), A new index for more accurate winter predictions, *Geophysical Research Letters*, 38(21), L21701.

Cohen, J., J. A. Screen, J. C. Furtado, M. Barlow, D. Whittleston, D. Coumou, J. Francis, K. Dethloff, D. Entekhabi, J. Overland, and J. Jones (2014), Recent Arctic amplification and extreme mid-latitude weather, *Nature Geosciences*, 7(9), 627-637.

Collins, M., R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W.J. Gutowski, T. Johns, G. Krinner, M. Shongwe, C. Tebaldi, A.J. Weaver and M. Wehner (2013), Long-term Climate Change: Projections, Commitments and Irreversibility. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to

the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.), Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Córdoba-Machado, S., R. Palomino-Lemus, S. Gámiz-Fortis, Y. Castro-Díez, and M. J. Esteban-Parra (2014), Influence of tropical Pacific SST on seasonal precipitation in Colombia: prediction using El Niño and El Niño Modoki, *Climate Dynamics*, 1-18.

Corte-Real, J., X. Zhang, and X. Wang (1995), Dynamical Downscaling GCM information to regional scales: a non-parametric multivariate regression approach, *Climate Dynamics*, 11, 413-424.

Czaja, A., and C. Frankignoul (1999), Influence of the North Atlantic SST on the atmospheric circulation, *Geophysical Research Letters*, 26(19), 2969-2972.

Czaja, A., and C. Frankignoul (2002), Observed impact of Atlantic SST anomalies on the North Atlantic Oscillation, *Journal of Climate*, 15(6), 606-623.

De Castro, M., J. Martín-Vide, and S. Alonso (2005), El clima de España: pasado, presente y escenarios de clima para el siglo XXI. En Impactos del cambio climático en España. 64 pp. Ministerio Medio Ambiente, Madrid.

De Castro, M., N. Lorenzo, J. J. Taboada, M. Sarmiento, I. Álvarez, and M. Gomez-Gesteira (2006), Influence of teleconnection patterns on precipitation variability and on river flow regimes in the Miño River basin (NW Iberian Peninsula), *Climate Research*, 32(1), 63-73.

Dettinger, M. D., and D. R. Cayan (1995), Large-scale atmospheric forcing of recent trends toward early snowmelt runoff in California, *Journal of Climate*, 8, 606–623.

Dettinger, M. D., and H. F. Diaz (2000), Global characteristics of stream flow seasonality and variability, *Journal of Hydrometeorology*, (1978), 289-310.

Di Lorenzo, E., N. Schneider, K. M. Cobb, P. J. S. Franks, K. Chhak, A. J. Miller, J. C. McWilliams, S. J. Bograd, H. Arango, E. Curchitser, T. M. Powell and P. Rivière (2008), North Pacific Gyre Oscillation links ocean climate and ecosystem change, *Geophysical Research Letters*, 35(8), L08607.

Di Lorenzo, E., K. M. Cobb, J. C. Furtado, N. Schneider, B. T. Anderson, A. Bracco, M. A. Alexander, and D. J. Vimont (2010), Central Pacific El Nino and decadal climate change in the North Pacific Ocean, *Nature Geosciences*, 3(11), 762-765.

Diaz, H. F., and V. Markgraf (2000), El Niño and the Southern Oscillation: Multiscale variability and global and regional impacts, 496 pp., Cambridge University Press.

Ding, H., N. Keenlyside, and M. Latif (2012), Impact of the Equatorial Atlantic on the El Niño Southern Oscillation, *Climate Dynamics*, 38(9-10), 1965-1972.

Drévillon, M., L. Terray, P. Rogel, and C. Cassou (2001), Mid latitude Atlantic SST influence on European winter climate variability in the NCEP Reanalysis, *Climate Dynamics*, 18(3-4), 331-344.

Dünkeloh, A., and J. Jacobeit (2003), Circulation dynamics of Mediterranean precipitation variability 1948–98, *International Journal of Climatology*, 23(15), 1843-1866.

EEA, European Environment Agency (2000), Europe's environment: the third assessment. Copenhagen.

Eldaw, A. K., J. D. Salas, and L. A. Garcia (2003), Long-Range Forecasting of the Nile River Flows Using Climatic Forcing, *Journal of Applied Meteorology*, 42(7), 890-904.

Esteban-Parra, M. J., F. S. Rodrigo, and Y. Castro-Díez (1998), Spatial and temporal patterns of precipitation in Spain for the period 1880–1992, *International Journal of Climatology*, 18(14), 1557-1574.

Estrela, T., M. A. Pérez-Martin, and E. Vargas (2012), Impacts of climate change on water resources in Spain, *Hydrological Sciences Journal*, 57(6), 1154-1167.

Feng, S., Q. Hu, and R. Oglesby (2011), Influence of Atlantic sea surface temperatures on persistent drought in North America, *Climate Dynamics*, 37(3-4), 569-586.

Flores, J. (2004), 50 años de la Confederación Hidrográfica del Tajo. Confederación Hidrográfica del Tajo, Madrid.

Feyen, L., and R. Dankers (2009), Impact of global warming on streamflow drought in Europe, *Journal of Geophysical Research: Atmospheres*, 114(D17), D17116.

Forzieri, G., L. Feyen, R. Rojas, M. Flarke, F. Wimmer, and A. Bianchi (2014), Ensemble projections of future streamflow droughts in Europe, *Hydrology and Earth System Sciences*, 18(1), 85-108.

Francis, J. A., W. Chan, D. J. Leathers, J. R. Miller, and D. E. Veron (2009), Winter Northern Hemisphere weather patterns remember summer Arctic sea-ice extent, *Geophysical Research Letters*, 36(7), L07503.

Frankignoul, C., P. Friederichs, and E. Kestenare (2003), Influence of Atlantic SST anomalies on the atmospheric circulation in the Atlantic-European sector, *Annals of Geophysics*, 46(1), 71-85.

Frankignoul, C., and E. Kestenare (2005), Observed Atlantic SST anomaly impact on the NAO: An Update, *Journal of Climate*, 18(19), 4089-4094.

Frauen, C., and D. Dommenget (2012), Influences of the tropical Indian and Atlantic Oceans on the predictability of ENSO, *Geophysical Research Letters*, 39(2), L02706.

Frías, M. D., S. Herrera, A. S. Cofiño, and J. M. Gutiérrez (2010), Assessing the Skill of Precipitation and Temperature Seasonal Forecasts in Spain: Windows of Opportunity Related to ENSO Events, *Journal of Climate*, 23(2), 209-220.

Fu, G., M. E. Barber, and S. Chen (2010), Hydro-climatic variability and trends in Washington State for the last 50 years, *Hydrological Processes*, 24(7), 866-878.

Gallart, F., and P. Llorens (2002), Water resources and environmental change in Spain. A key issue for integrated catchment management, in Environmental Change and Water Sustainability, edited by J. M. García-Ruiz, J. J. A. Jones, and J. Arnáez, 340 pp., Inst. Pirenaico de Ecol., Zaragoza, Spain.

Gámiz-Fortis, S., D. Pozo-Vázquez, R. M. Trigo, and Y. Castro-Díez (2008a), Quantifying the predictability of winter river flow in Iberia. Part I: Interannual predictability, *Journal of Climate*, 21(11), 2484-2502.

Gámiz-Fortis, S., D. Pozo-Vázquez, R. M. Trigo, and Y. Castro-Díez (2008b), Quantifying the predictability of winter river flow in Iberia. Part II: Seasonal predictability, *Journal of Climate*, 21(11), 2503-2518.

Gámiz-Fortis, S. R., M. J. Esteban-Parra, R. M. Trigo, and Y. Castro-Díez (2010), Potential predictability of an Iberian river flow based on its relationship with previous winter global SST, *Journal of Hydrology*, 385(1-4), 143-149.

Gámiz-Fortis, S. R., J. M. Hidalgo-Muñoz, D. Argüeso, M. J. Esteban-Parra, and Y. Castro-Díez (2011), Spatio-temporal variability in Ebro river basin (NE Spain): Global SST as

potential source of predictability on decadal time scales, *Journal of Hydrology*, 409(3-4), 759-775.

Gandin, L. S., and A. H. Murphy (1992), Equitable skill scores for categorical forecasts, *Monthly Weather Review*, 120, 361-370.

García-Herrera, R., J. Díaz, R. M. Trigo, and E. Hernández (2005), Extreme summer temperatures in Iberia: health impacts and associated synoptic conditions, *Annales Geophysicae*, 23(2), 239-251.

García-Serrano, J., T. Losada, B. Rodríguez-Fonseca, and I. Polo (2008), Tropical Atlantic variability modes (1979-2002). Part II: Time-evolving atmospheric circulation related to SST-forced tropical convection, *Journal of Climate*, 21(24), 6476-6497.

García-Serrano, J., B. Rodríguez-Fonseca, I. Bladé, P. Zurita-Gotor, and A. de la Cámara (2010), Rotational atmospheric circulation during North Atlantic-European winter: the influence of ENSO, *Climate Dynamics*, 37(9-10), 1727-1743.

García-Serrano, J., and C. Frankignoul (2014), High predictability of the winter Euro-Atlantic climate from cryospheric variability, *Nature Geosciences*, 7(6), E1-E1.

Gastineau, G., F. D'Andrea, and C. Frankignoul (2013), Atmospheric response to the North Atlantic Ocean variability on seasonal to decadal time scales, *Climate Dynamics*, 40(9-10), 2311-2330.

Genta, J., G. Perez-Iribarren, and C. R. Mechoso (1998), A Recent increasing trend in the streamflow of rivers in southeastern South America, *Journal of Climate*, 11(11), 2858-2862.

Gerrity, J. P. (1992), A note on Gandin and Murphy's equitable skill score, *Monthly Weather Review*, 120(11), 2709-2712.

Giannini, A., M. A. Cane, and Y. Kushnir (2001), Interdecadal changes in the ENSO teleconnection to the caribbean region and the North Atlantic Oscillation, *Journal of Climate*, 14(13), 2867-2879.

Gil Olcina, A. (2006), Regímenes natural y artificial del Río Júcar. *Investigaciones Geográficas*, N. 40, pp. 5-34.

Glickman T.S., (2000), Glossary of Meteorology, 2nd Edition, American Meteororogical Society, Boston, 855 pp.

Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane (2001), Current approaches to seasonal to interannual climate predictions, *International Journal of Climatology*, 21(9), 1111-1152.

Gómez-Mendoza, J., and R. Mata (1999), Abastecimiento urbano, regadíos, trasvases y planificación en la cuenca del Tajo, in Los Usos del Agua en España, pp. 301 – 336, *Caja de Ahorros del Mediterráneo*, Alicante, Spain.

Gong, G., D. Entekhabi, and J. Cohen (2003), Relative impacts of Siberian and North American snow anomalies on the winter Arctic Oscillation, *Geophysical Research Letters*, 30(16), 1848.

González-Hidalgo, J. C., M. Brunetti, and M. de Luis (2011), A new tool for monthly precipitation analysis in Spain: MOPREDAS database (monthly precipitation trends December 1945–November 2005), *International Journal of Climatology*, 31(5), 715-731.

González-Rouco, J. F., H. Heyen, E. Zorita, and F. Valero (2000), Agreement between observed rainfall trends and climate shange Simulations in the southwest of Europe, *Journal of Climate*, 13(17), 3057-3065.

Grantz, K., B. Rajagopalan, M. Clark, and E. Zagona (2005), A technique for incorporating large-scale climate information in basin-scale ensemble streamflow forecasts, *Water Resources Research*, 41(10), W10410.

Gray, S. T., L. J. Graumlich, J. L. Betancourt, and G. T. Pederson (2004), A tree-ring based reconstruction of the Atlantic Multidecadal Oscillation since 1567 A.D, *Geophysical Research Letters*, 31(12), L12205.

Greenwood, J., J. Landwehr, N. Matalas, and J. Wallis (1979), Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research*, 15, 1049-1054.

Groisman, P. Y., R. W. Knight, D. R. Easterling, T. R. Karl, G. C. Hegerl, and V. A. N. Razuvaev (2005), Trends in intense precipitation in the climate record, *Journal of Climate*, 18 (9), 1326–1350.

Gutiérrez, F., and J. A. Dracup (2001), An analysis of the feasibility of long-range streamflow forecasting for Colombia using El Niño-Southern Oscillation indicators, *Journal of Hydrology*, 246(1-4), 181-196.

Hair, J. F. Jr., R. E. Anderson, R. L. Tatham, and W. C. Black (1995), Multivariate Data Analysis, 3rd ed. New York: Macmillan.

Halpert, M. S., and C. F. Ropelewski (1992), Surface temperature patterns associated with the Southern Oscillation, *Journal of Climate*, 5(6), 577-593.

Hannaford, J., and T. Marsh (2006), An assessment of trends in UK runoff and low flows using a network of undisturbed catchments, *International Journal of Climatology*, 26(9), 1237-1253.

Hanson, R. T., Newhouse, M. W., and Dettinger, M.W. (2004), A methodology to assess relations between climatic variability and variations in hydrologic time series in the southwestern United States, *Journal of Hydrology*, 287(1–4), 252–269.

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister (2014), Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset, *International Journal of Climatology*, 34(3), 623-642.

Hernández-Martínez, M., J. M. Hidalgo-Muñoz, S. R. Gámiz-Fortis, Y. Castro-Díez, and M. J. Esteban-Parra (2014), Temporal variability and potential predictability of the streamflow regimes in the northeastern Iberian Peninsula, *River Research and Applications*, n/a n/a, doi: 10.1002/rra.2825.

Hidalgo-Muñoz, J. M., D. Argüeso, S. R. Gámiz-Fortis, M. J. Esteban-Parra, and Y. Castro-Díez (2011), Trends of extreme precipitation and associated synoptic patterns over the southern Iberian Peninsula, *Journal of Hydrology*, 409(1-2), 497-511.

Hoerling, M. P., J. W. Hurrell, T. Xu, G. T. Bates, and a. S. Phillips (2004), Twentieth century North Atlantic climate change. Part II: Understanding the effect of Indian Ocean warming, *Climate Dynamics*, 23(3-4), 391-405.

Horel, J. D., and J. M. Wallace (1981), Planetary-scale atmospheric phenomena associated with the Southern Oscillation, *Monthly Weather Review*, 109, 813–829.

Hosking, J. R. M. (1989), Some theoretical results concerning L-moments. Research Report RC14492, *IBM Research Division*, Yorktown Heights, N.Y

Hosking, J. R. M. (1990), L-moments: analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52, 105-124.

Hosking, J. R. M., and J. R. Wallis (1997), Regional frequency analysis: an approach based on L-moments, Cambridge University Press.

Hsieh, W.W. (2001) Nonlinear canonical correlation analysis of the tropical Pacific climate variability using a neural network approach, *Journal of Climate*, 14, 2528-2539.

Hsieh, W. W., and B. Tang (1998), Applying neural network models to prediction and data analysis in Meteorology and Oceanography, *Bulletin of the American Meteorological Society*, 79(9), 1855-1870.

Hu, Y., S. Maskey, S. Uhlenbrook, and H. Zhao (2011), Streamflow trends and climate linkages in the source region of the Yellow River, China, *Hydrological Processes*, 25(22), 3399-3411.

Hurrell, J. W. (1995), Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation, *Science*, 269(5224), 676-679.

Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck (2003), The North Atlantic Oscillation: Climatic significance and environmental impact. *Geophyical Monographies. Series*, 134, 279 pp.

Hyvarinen, V. (2003), Trends and characteristics of hydrological time series in Finland. *Nordic Hydrology* 34(1-2): 71-90

Iglesias, A., L. Garrote, F. Flores, and M. Moneo (2007), Challenges to manage the risk of water scarcity and climate change in the Mediterranean, *Water Resources Management*, 21(5), 775-788.

Ionita, M., G. Lohmann, and N. Rimbu (2008), prediction of spring Elbe discharge based on stable teleconnections with winter global temperature and precipitation, *Journal of Climate*, 21, 6215-6226.

Ionita, M., G. Lohmann, N. Rimbu, and S. Chelcea (2011), Interannual variability of Rhine River streamflow and its relationship with large-scale anomaly patterns in spring and autumn, *Journal of Hydrometeorology*, 13(1), 172-188.

IPCC, 2013: Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Iwasaka, N. and J. M. Wallac (1995), Large-scale air sea interaction in the northern hemisphere from a view point of variations of surface heat flux by SVD analysis. *Journal Meteorological Society of Japan*, 73, 781–794

Izumo, T., J. Vialard, M. Lengaigne, C. de Boyer Montegut, S. K. Behera, J. J. Luo, S. Cravatte, S. Masson, and T. Yamagata (2010), Influence of the state of the Indian Ocean Dipole on the following year/'s El Nino, *Nature Geosciences*, 3(3), 168-172.

Jansen, M. F., D. Dommenget, and N. Keenlyside (2009), Tropical atmosphere-ocean interactions in a conceptual Fframework, *Journal of Climate*, 22(3), 550-567.

Jarque, C. M. and A. K. Bera (1987), A test for normality of observations and regression residuals, *Internat. Statst. Rev*. 55(2), pp. 163–172.

Jolliffe, I. T. (2002), Principal Component Analysis. Second ed. Springer Series in Statistics. New York: Springer-Verlag New York

Jolliffe, I. T., and D. B. Stephenson (2003), Forecast verification. A practitioner's guide in atmospheric sciences, John Wiley & Sons, Ltd.

Joly, M., A. Voldoire, H. Douville, P. Terray, and J.-F. Royer (2007), African monsoon teleconnections with tropical SSTs: validation and evolution in a set of IPCC4 simulations, *Climate Dynamics*, 29(1), 1-20.

Kahya, E., and S. Kalaycy (2004), Trend analysis of streamflow in Turkey, *Journal of Hydrology*, 289(1-4), 128-144.

Kahya, E. M. C. Demirel, and O. A Bég, (2008), Hydrologic homogeneous regions using monthly streamflow in Turkey, *Earth Sciences Research Journal*, 12(2), 181-193.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph (1996), The NCEP/NCAR 40-Year Reanalysis Project, *Bulletin of the American Meteorological Society*, 77(3), 437-471.

Kalra, A., W. P. Miller, K. W. Lamb, S. Ahmad, and T. Piechota (2013), Using large-scale climatic patterns for improving long lead time streamflow forecasts for Gunnison and San Juan River Basins, *Hydrological Processes*, 27(11), 1543-1559.

Kharin, V. V., and F. W. Zwiers (2005), Estimating extremes in transient climate change simulations, *Journal of Climate*, 18(8), 1156–1173.

Kayano, M. T., R. V. Andreoli, and R. A. Ferreira de Souza (2013), Relations between ENSO and the South Atlantic SST modes and their effects on the South American rainfall, *International Journal of Climatology*, 33(8), 2008-2023.

Kendall, M.G. (1975), Rank Correlation Methods, 4th edition, Charles Griffin, London.

Kennedy, P. (1992), A Guide to Econometrics. Oxford: Blackwell.

Kerr, R. A. (2000), A North Atlantic Climate Pacemaker for the Centuries, *Science*, 288(5473), 1984-1985.

Knight, J. R., C. K. Folland, and A. A. Scaife (2006), Climate impacts of the Atlantic Multidecadal Oscillation, *Geophysical Research Letters*, 33(17), L17706.

Knippertz, P., U. Ulbrich, F. Marques, and J. Corte-Real (2003), Decadal changes in the link between El Niño and springtime North Atlantic oscillation and European–North African rainfall, *International Journal of Climatology*, 23(11), 1293-1311.

Koriala, S., Y. Hirabayashi, R. Mahendran, and K. Shinjiro (2014), Global assessment of agreement among streamflow projections using CMIP5 model outputs, *Environmental Research Letters*, 9(6), 064017.

Krichak, S. O., P. Kishcha, and P. Alpert (2002), Decadal trends of main Eurasian oscillations and the Eastern Mediterranean precipitation, *Theoretical and Applied Climatology*, 72(3-4), 209-220.

Kristjánsson, J. E., A. Staple, J. Kristiansen, and E. Kaas (2002), A new look at possible connections between solar activity, clouds and climate, *Geophysical Research Letters*, 29(23), 2107.

Kroll, C. N., and P. Song (2013), Impact of multicollinearity on small sample hydrologic regression models, *Water Resources Research*, 49(6), 3756-3769.

Kucharski, F., F. S. Syed, A. Burhan, I. Farah, and A. Gohar (2014), Tropical Atlantic influence on Pacific variability and mean state in the twentieth century in observations and CMIP5, *Climate Dynamics*, 1-16.

Kundzewicz, Z. W., and A. J. Robson (2004), Change detection in hydrological records, *Hydrological Sciences Journal*, 49(1), 7-19.

Kung, E. C., and T. A. Sharif (1980), Regression forecasting of the onset of the Indian summer monsoon with antecedent upper air conditions, *Journal of Applied Meteorology*, 19(4), 370-380.

Labat, D. (2010), Cross wavelet analyses of annual continental freshwater discharge and selected climate indices, *Journal of Hydrology*, 385(1-4), 269-278.

Lamb, K., T. Piechota, O. Aziz, and G. Tootle (2010), Basis for extending long-term streamflow forecasts in the Colorado River Basin, *Journal of Hydrologic Engineering*, 16(12), 1000-1008.

Landwehr, J. M., N. C. Matalas, and J. R. Wallis (1979), Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles, *Water Resources Research*, 15(5), 1055-1064.

Lau, N. C., and M. J. Nath (1994), A modeling study of the relative roles of tropical and extratropical SST anomalies in the variability of the global atmosphere-ocean system. *Journal of Climate*, 7, 1184- 1207.

Lau, N. C., and M. J. Nath (1996), The role of the 'Atmospheric Bridge' in linking tropical Pacific ENSO events to extratropical SST anomalies, *Journal of Climate*, 9(9), 2036-2057.

Lau, N. C., and M. J. Nath (2001), Impact of ENSO on SST variability in the North Pacific and North Atlantic: Seasonal dependence and role of extratropical air-sea coupling, *Journal of Climate*, 14, 2846–2866.

Lespinas, F., W. Ludwig, and S. Heussner (2010), Impact of recent climate change on the hydrology of coastal Mediterranean rivers in Southern France, *Climatic Change*, 99(3-4), 425-456.

Lilliefors, H. W. (1967), On the Kolmogorov-Smirnov test for normality with mean and variance unknown, *Journal of the American Statistical Association*, 62(318), 399-402.

Lima, C. H. R., U. Lall, T. Jebara, and A. G. Barnston (2009), Statistical prediction of ENSO from subsurface sea temperature using a nonlinear dimensionality reduction, *Journal of Climate*, 22(17), 4501-4519.

Lima, C. H. R., and U. Lall (2010), Spatial scaling in a changing climate: A hierarchical Bayesian model for non-stationary multi-site annual maximum and monthly streamflow, *Journal of Hydrology*, 383(3-4), 307-318.

Lindstrom, G. and S. Bergstrom (2004), Runoff trends in Sweden 1807-2002. *Hydrological Sciences Journal*, 49, 69-83.

Linkin, M. E., and S. Nigam (2008), The North Pacific Oscillation-West Pacific teleconnection pattern: Mature-phase structure and winter impacts, *Journal of Climate*, 21(9), 1979-1997.

Livezey, R.E. (1995), The evaluation of forecasts, in Analysis of Climate Variability, edited by H. von Storch and A. Navarra, Springer, Berlin, p. 334.

Livezey, R.E. (2003), Categorical events, In *Forecast Verification*, Jolliffe, I.T., and D.B. Stephenson, eds. Wiley, New York, 77-96.

Liu, Y. (2003), Prediction of monthly-seasonal precipitation using coupled SVD patterns between soil moisture and subsequent precipitation, *Geophysical Research Letters*, 30(15), 1827.

Liu, Q., N. Wen, and Z. Liu (2006), An observational study of the impact of the North Pacific SST on the atmosphere, *Geophysical Research Letters*, 33(18), L18611.

Lloyd-Hughes, B., and M. A. Saunders (2002), Seasonal prediction of European spring precipitation from El Niño-Southern Oscillation and local sea-surface temperatures, *International Journal of Climatology*, 22(1), 1-14.

Lopez-Bustins, J. A., J. Martin-Vide, and A. Sánchez-Lorenzo (2008), Iberia winter rainfall trends based upon changes in teleconnection and circulation patterns, *Global and Planetary Change*, 63(2-3), 171-176.

López-Moreno, J. I. (2005), Recent variations of snowpack depth in the central Spanish Pyrenees, *Arctic, Antarctic, and Alpine Research*, 37(2), 253-260.

López-Moreno, J. I., and J. M. García-Ruiz (2004), Influence of snow accumulation and snowmelt on streamflow in the central Spanish Pyrenees, *Journal of Hydrological Sciences*, 49, 787– 802.

López-Moreno, J. I., S. Beguería, S. M. Vicente-Serrano, and J. M. García-Ruiz (2007), Influence of the North Atlantic Oscillation on water resources in central Iberia: Precipitation, streamflow anomalies, and reservoir management strategies, *Water Resources Research*, 43(9), W09411.

López-Moreno, J. I., and S. M. Vicente-Serrano (2008), Positive and negative phases of the wintertime North Atlantic Oscillation and drought occurrence over Europe: A multitemporal-scale Approach, *Journal of Climate*, 21(6), 1220-1243.

López-Moreno, J. I., S. M. Vicente-Serrano, S. Beguería, J. M. García-Ruiz, M. M. Portela, and A. B. Almeida (2009a), Dam effects on droughts magnitude and duration in a transboundary basin: The lower River Tagus, Spain and Portugal, *Water Resources Research*, 45(2), W02405.

López-Moreno, J. I., S. M. Vicente-Serrano, L. Gimeno, and R. Nieto (2009b), Stability of the seasonal distribution of precipitation in the Mediterranean region: Observations since 1950 and projections for the 21st century, *Geophysical Research Letters*, 36(10), L10703.

López-Moreno, J.I, S. M. Vicente-Serrano, E. Moran-Tejeda, J. Zabalza1, J. Lorenzo-Lacruz, and J. M. García-Ruiz (2011), Impact of climate evolution and land use changes on water yield in the ebro basin, *Hydrology and Earth System Sciences*, 15, 311-322.

Lorenzo, E. D., N. Scheneider, K. M. Cobb, J. Furtado, and M. Alexander (2012), ENSO and the North Pacific Gyre Oscillation: an integrated view of Pacific decadal dynamics, *Geophyscal Reearchs Letters*, submitted.

Lorenzo, M. N., J. J. Taboada, I. Iglesias, and M. Gómez-Gesteira (2010), Predictability of the spring rainfall in Northwestern Iberian Peninsula from sea surfaces temperature of ENSO areas, *Climatic Change*, 107(3-4), 329-341.

Lorenzo-Lacruz, J., S. M. Vicente-Serrano, J. I. López-Moreno, S. Beguería, J. M. García-Ruiz, and J. M. Cuadrat (2010), The impact of droughts and water management on various hydrological systems in the headwaters of the Tagus River (central Spain), *Journal of Hydrology*, 386(1-4), 13-26.

Lorenzo-Lacruz, J., S. M. Vicente-Serrano, J. I. López-Moreno, J. C. González-Hidalgo, and E. Morán-Tejeda (2011), The response of Iberian rivers to the North Atlantic Oscillation, *Hydrology and Earth System Sciences*, 15(8), 2581-2597.

Lorenzo-Lacruz, J., S. M. Vicente-Serrano, J. I. López-Moreno, E. Morán-Tejeda, and J. Zabalza (2012), Recent trends in Iberian streamflows (1945-2005), *Journal of Hydrology*, 414-415(0), 463-475.

Lorenzo-Lacruz, J., S. M. Vicente-Serrano, J. C. González-Hidalgo, J. I. López-Moreno, and N. Cortesi (2013a), Hydrological drought response to meteorological drought in the Iberian Peninsula, *Climate Research*, 58, 117-131.

Lorenzo-Lacruz, J., E. Morán-Tejeda, S. M. Vicente-Serrano, and J. I. López-Moreno (2013b), Streamflow droughts in the Iberian Peninsula between 1945 and 2005: spatial and temporal patterns, *Hydrology and Earth System Sciences*, 17(1), 119-134.

Mailier, P. J., D. B. Stephenson, C. A. T. Ferro, and K. I. Hodges (2006), Serial clustering of extratropical cyclones, *Monthly Weather Review*, 134(8), 2224-2240.

Maity, R., and D. Nagesh Kumar (2008), Probabilistic prediction of hydroclimatic variables with nonparametric quantification of uncertainty, *Journal of Geophysical Research: Atmospheres*, 113(D14), D14105.

Mann, H.B. (1945), Non-parametric tests against trend, *Econometrica*, 13, 163-171.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis (1997), A Pacific interdecadal climate oscillation with impacts on salmon production, *Bulletin of the American Meteorological Society*, 78, 1069-1079.

Mariotti, A., N. Zeng, and K. M. Lau (2002), Euro-Mediterranean rainfall and ENSO-a seasonally varying relationship, *Geophysical Research Letters*, 29(12), 59-1-59-4.

MARM (Ministry of Environment and Rural and Maritime Affairs) (2008), Informe calidad del agua de consumo humano en España, Ministerio de Medio Ambiente y Medio Rural y Marino, España.

Marshall, A. G., and A. A. Scaife (2009), Impact of the QBO on surface winter climate, *Journal of Geophysical Research: Atmospheres*, 114(D18), D18110.

Marshall, A. G., and A. A. Scaife (2010), Improved predictability of stratospheric sudden warming events in an atmospheric general circulation model with enhanced stratospheric resolution, *Journal of Geophysical Research: Atmospheres*, 115(D16), D16114.

Marshall, J., Y. Kushnir, D. Battisti, P. Chang, A. Czaja, R. Dickson, J. Hurrell, M. McCartney, R. Saravanan, and M. Visbeck (2001), North Atlantic climate variability: phenomena, impacts and mechanisms, *International Journal of Climatology*, 21(15), 1863-1898.

Martín i Díaz, J. (2010), Influence of low-frequency variability patterns WeMO and NAO on streamflow of the rivers in Catalonia (1951-2000), *Boletín de la Asociación de Geógrafos Español*es, 54, 237-257.

Martín, M. L., M. Y. Luna, A. Morata, and F. Valero (2004), North Atlantic teleconnection patterns of low-frequency variability and their links with springtime precipitation in the western Mediterranean, *International Journal of Climatology*, 24(2), 213-230.

Martín, M. L., F. Valero, A. Morata, M. Y. Luna, A. Pascual, and D. Santos-Muñoz (2011), Springtime coupled modes of regional wind in the Iberian Peninsula and large-scale variability patterns, *International Journal of Climatology*, 31, 880–895.

Martin-Vide, J., and J. A. Lopez-Bustins (2006), The Western Mediterranean Oscillation and rainfall in the Iberian Peninsula, *International Journal of Climatology*, 1475(11), 1455-1475.

Martin-Vide, J., A. Sanchez-Lorenzo, J. A. Lopez-Bustins, M. J. Cordobilla, A. Garcia-Manuel, and J. M. Raso (2008), Torrential rainfall in northeast of the Iberian Peninsula: synoptic patterns and WeMO influence, *Advances in Science and Research*, 2, 99-105.

Mason, S. J. (1998), Seasonal forecasting of South African rainfall using a non–linear discriminant analysis model, *International Journal of Climatology*, 18(2), 147-164.

Mason, R. L., R. F. Gunst, and, J. L. Hess (1989), Statistical design and analysis of Experiments: Applications to Engineering and Science. New York: Wiley

Massei, N., B. Laignel, J. Deloffre, J. Mesquita, A. Motelay, R. Lafite, and A. Durand (2010), Long-term hydrological changes of the Seine River flow (France) and their relation to the North Atlantic Oscillation over the period 1950–2008, *International Journal of Climatology*, 30(14), 2146-2154.

Massey, F. (1951), The Kolmogorov-Smirnov Test for Goodness of Fit, *Journal of the American Statistical Association*, 46(253), 68-78.

Matsumura, S., X. Zhang, and K. Yamazaki (2014), Summer Arctic atmospheric circulation response to spring Eurasian snow cover and its possible linkage to accelerated sea ice decrease, *Journal of Climate*, 27(17), 6551-6558.

Mavromatis, T., and D. Stathis (2011), Response of the water balance in Greece to temperature and precipitation trends, *Theoretical and Applied Climatology*, 104(1-2), 13-24.

Menard, S. (1995), Applied Logistic Regression Analysis: Sage University Series on Quantitative Applications in the Social Sciences, Thousand Oaks, CA: Sage.

MIMAM, (2003), Water in Spain (White Paper on Water in Spain). Ministerio de Medio Ambiente. I.S.B.N.: 84-8320-219-0, 607 pp

Minobe, S. (1997), A 50–70 year climatic oscillation over the North Pacific and North America, *Geophysical Research Letters*, 24(6), 683-686.

Morales, A., A. M. Rico, and M. Hernández (2005), El trasvase Tajo Segura, *Observatorio Medioambiental*, 8, 73-110.

Morán-Tejeda, E., A. Ceballos-Barbancho, and J. M. Llorente-Pinto (2010), Hydrological response of Mediterranean headwaters to climate oscillations and land-cover changes: The mountains of Duero River basin (Central Spain), *Global and Planetary Change*, 72(1-2), 39-49.

Morán-Tejeda, E., L.-M. I. López-Moreno, A. Ceballos.-Barbancho, and S. M. Vicente-Serrano (2011a), Evaluating Duero's basin (Spain) response to the NAO phases: spatial and seasonal variability, *Hydrological Processes*, 25(8), 1313-1326.

Morán-Tejeda, E., J. I. López-Moreno, A. Ceballos-Barbancho, and S. M. Vicente-Serrano (2011b), River regimes and recent hydrological changes in the Duero basin (Spain), *Journal of Hydrology*, 404(3-4), 241-258.

Morán-Tejeda, E., J. Lorenzo-Lacruz, J.I. López-Moreno, A. Ceballos-Barbancho, J. Zabalza, and S.M. Vicente-Serrano, (2012), Reservoir management in the Duero Basin (Spain): impact on river regimes and the response to environmental change, *Water Resources Management*, 26 (8), 2125-2146.

Mourato, S., M. Moreira, and J. Corte-Real (2010), Interannual variability of precipitation distribution patterns in Southern Portugal, *International Journal of Climatology*, 30(12), 1784-1794.

Neumann, J. V. (1941), Distribution of the ratio of the Mean Square Successive Difference to the variance, *The Annals of Mathematical Statistics*.

Newman, M., and P. D. Sardeshmukh (1995), A Caveat Concerning Singular Value Decomposition, *Journal of Climate*, 8(2), 352-360.

Neelin, J. D., Battisti, D. S., Hirst, A. C., Jin, F.-F., Wakata, Y., Yamagata, T., and Zebiak, S. E. (1998). ENSO Theory, *Journal of Geophysical Research*, 103, 14 261–14 290.

North, G. R. (1984), Empirical Orthogonal Functions and Normal Modes, *Journal of the Atmospheric Sciences*, 41(5), 879-887.

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng (1982), Sampling errors in the estimation of Empirical Orthogonal Functions, *Monthly Weather Review*, 110(7), 699-706.

Ogi, M., and K. Yamazaki (2010), Trends in the Summer Northern Annular Mode and Arctic Sea Ice, *Sola*, 6, 41-44.

Oldenborgh, G. J. V., G. Burgers, and A. K. Tank (2000), On the El Niño teleconnection to spring precipitation in Europe, *International Journal of Climatology*, 20(5), 565-574.

Omondi, P., L. A. Ogallo, R. Anyah, J. M. Muthama, and J. Ininda (2013), Linkages between global sea surface temperatures and decadal rainfall variability over Eastern Africa region, *International Journal of Climatology*, 33(8), 2082-2104.

Opitz-Stapleton, S., S. Gangopadhyay, and B. Rajagopalan (2007), Generating streamflow forecasts for the Yakima River Basin using large-scale climate predictors, *Journal of Hydrology*, 341(3-4), 131-143.

Orsolini, Y. J., and N. G. Kvamstø (2009), Role of Eurasian snow cover in wintertime circulation: Decadal simulations forced with satellite observations, *Journal of Geophysical Research: Atmospheres*, 114(D19), D19108.

Oubeidillah, A. A., G. A. Tootle, C. Moser, T. Piechota, and K. Lamb (2011), Upper Colorado River and Great Basin streamflow and snowpack forecasting using Pacific oceanic-atmospheric variability, *Journal of Hydrology*, 410(3-4), 169-177.

Oubeidillah, A. A., G. Tootle, and S. R. Anderson (2012), Atlantic Ocean sea-surface temperatures and regional streamflow variability in the Adour-Garonne basin, France, *Hydrological Sciences Journal*, 57(3), 496-506.

Outten, S. D., and I. Esau (2012), A link between Arctic sea ice and recent cooling trends over Eurasia, *Climatic Change*, 110(3-4), 1069-1075.

Overland, J. E., and M. Wang (2010), Large-scale atmospheric circulation changes are associated with the recent loss of Arctic sea ice, *Tellus A*, 62(1), 1-9.

Paredes, D., R. M. Trigo, R. Garcia-Herrera, and I. Franco Trigo (2006), Understanding precipitation changes in Iberia in early spring: Weather typing and storm-tracking approaches, *Journal of Hydrometeorology*, 7, 101-113.

Park, S. (2004), Remote ENSO influence on Mediterranean sky conditions during late summer and autumn: Evidence for a slowly evolving atmospheric bridge, *Quarterly Journal of the Royal Meteorological Society*, 130(602), 2409-2422.

Peings, Y., and H. Douville (2010), Influence of the Eurasian snow cover on the Indian summer monsoon variability in observed climatologies and CMIP3 simulations, *Climate Dynamics*, 34(5), 643-660.

Peings, Y., and G. Magnusdottir (2013), Response of the wintertime Northern Hemisphere atmospheric circulation to current and projected Arctic sea ice decline: A numerical study with CAM5, *Journal of Climate*, 27(1), 244-264.

Peings, Y., E. Brun, V. Mauvais, and H. Douville (2013), How stationary is the relationship between Siberian snow and Arctic Oscillation over the 20th century?, *Geophysical Research Letters*, 40(1), 183-188.

Peixoto, J. P., and A. H. Oort (1992), Physics of Climate, 520 pp., American Institute of Physics, New York.

Petrone, K. C., J. D. Hughes, T. G. Van Niel, and R. P. Silberstein (2010), Streamflow decline in southwestern Australia, 1950–2008, *Geophysical Research Letters*, 37(11), L11401.

Pettitt, A. N. (1979), A non-parametric approach to the change-point problem, *Applied statistics*, 28(2), 126-135.

Philander, S. G. H. (1990), El Niño, La Niña and the Southern Oscillation, pp 293, Academic Press, San Diego,.

Piechota, T., F. Chiew, J. Dracup, and T. McMahon (2001), Development of exceedance probability streamflow forecast, *Journal of Hydrologic Engineering*, 6(1), 20-28.

PHN, (2001). Plan Hidrológico Nacional. Ministerio de Medio Ambiente, España.

Polo, I., B. Rodríguez-Fonseca, T. Losada, and J. García-Serrano (2008), Tropical Atlantic Variability Modes (1979–2002). Part I: Time-Evolving SST Modes Related to West African Rainfall, *Journal of Climate*, 21, 6457–6475.

Pozo-Vázquez, D., S. R. Gámiz-Fortis, J. Tovar-Pescador, M. J. Esteban-Parra, and Y. Castro-Díez (2005), El Niño-southern oscillation events and associated European winter precipitation anomalies, *International Journal of Climatology*, 25(1), 17-31.

Preisendorfer, R. W. (1988): Principal Component Analysis in Meteorology and Oceanography, Elsevier, New York, 425 pp.

Programa AGUA (2004), Informe de sostenibilidad ambiental de las actuaciones urgentes del programa A.G.U.A. en las cuencas mediterráneas. Ministerio de Medio Ambiente, Vol. 2.

Quadrelli, R., V. Pavan, and F. Molteni (2001), Wintertime variability of Mediterranean precipitation and its links with large-scale circulation anomalies, *Climate Dynamics*, 17(5-6), 457-466.

Quadrelli, R., and J. M. Wallace (2004), A simplified linear framework for interpreting patterns of Northern Hemisphere wintertime climate variability, *Journal of Climate*, 17(19), 3728-3744.

Quiroga, S., L. Garrote, A. Iglesias, Z. Fernandez-Haddad, J. Schlickenrieder, B. de Lama, C. Mosso, and A. Sanchez-Arcilla (2011), The economic value of drought information for water management under climate change: a case study in the Ebro basin, *Natural Hazards and Earth System Sciences*, 11, 643–657, doi:10.5194/nhess-11-643-2011.

Rajagopalan, B., E. Cook, U. Lall, and B. K. Ray (2000), Spatiotemporal variability of ENSO and SST teleconnections to summer drought over the United States during the twentieth century, *Journal of Climate*, 13(24), 4244-4255.

Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent, and A. Kaplan (2003), Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *Journal of Geophysical Research: Atmospheres*, 108(D14), 4407.

Razali, N. M., and Y. B. Wah (2011), Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests, *Journal of Statistical Modeling and Analytics*, 2(1)21-33.

Richman, M. B. (1986), Rotation of principal components, *Journal of Climatology*, 6(3), 293-335.

Rimbu, N., C. Boroneant, C. Buta, and M. Dima (2002), Decadal variability of the Danube river flow in the lower basin and its relation with the North Atlantic Oscillation, *International Journal of Climatology*, 22(10), 1169-1179.

Rimbu, N., M. Dima, G. Lohmann, and S. Stefan (2004), Impacts of the North Atlantic Oscillation and the El Niño–Southern Oscillation on Danube river flow variability, *Geophysical Research Letters*, 31(23), L23203.

Rimbu, N., M. Dima, G. Lohmann, and I. Musat (2005), Seasonal prediction of Danube flow variability based on stable teleconnection with sea surface temperature, *Geophysical Research Letters*, 32(21), L21704.

Rinke, A., K. Dethloff, W. Dorn, D. Handorf, and J. C. Moore (2013), Simulated Arctic atmospheric feedbacks associated with late summer sea ice anomalies, *Journal of Geophysical Research: Atmospheres*, 118(14), 7698-7714.

Robertson, A. W., C. R. Mechoso, and Y. J. Kim (2000), The influence of Atlantic sea surface temperature anomalies on the North Atlantic Oscillation, *Journal of Climate*, 13(1), 122-138.

Rocha, A. (1999), Low-frequency variability of seasonal rainfall over the Iberian peninsula and ENSO, *International Journal of Climatology*, 19(8), 889-901.

Rodó, X., E. Baert, and F. A. Comín (1997), Variations in seasonal rainfall in Southern Europe during the present century: relationships with the North Atlantic Oscillation and the El Niño-Southern Oscillation, *Climate Dynamics*, 13(4), 275-284.

Rodrigo, F. S., and R. M. Trigo (2007), Trends in daily rainfall in the Iberian Peninsula from 1951 to 2002, *International Journal of Climatology*, 27(4), 513-529.

Rodríguez-Fonseca, B., and M. de Castro (2002), On the connection between winter anomalous precipitation in the Iberian Peninsula and North West Africa and the summer subtropical Atlantic sea surface temperature, *Geophysical Research Letters*, 29(18), 1863.

Rodríguez-Fonseca, B., I. Polo, E. Serrano, and M. de Castro (2006), Evaluation of the North Atlantic SST forcing on the European and Northern African winter climate, *International Journal of Climatology*, 26(2), 179-191.

Rodríguez-Fonseca, B., I. Polo, J. García-Serrano, T. Losada, E. Mohino, C. R. Mechoso, and F. Kucharski (2009), Are Atlantic Niños enhancing Pacific ENSO events in recent decades?, *Geophysical Research Letters*, 36(20), L20705.

Rodríguez-Puebla, C., A. H. Encinas, S. Nieto, and J. Garmendia (1998), Spatial and temporal patterns of annual precipitation variability over the Iberian Peninsula, *International Journal of Climatology*, 18(3), 299-316.

Rodríguez-Puebla, C., A. H. Encinas, and J. Sáenz (2001), Winter precipitation over the Iberian peninsula and its relationship to circulation indices, *Hydrology and Earth System Sciences*, 5(2), 233-244.

Rodwell, M. J., and C. K. Folland (2002), Atlantic air–sea interaction and seasonal predictability, *Quarterly Journal of the Royal Meteorological Society*, 128(583), 1413-1443.

Rogers, J. C., and H. Van Loon (1979), The seesaw in winter temperatures between Greenland and Northern Europe. Part II: Some oceanic and atmospheric effects in middle and high latitudes, *Monthly Weather Review*, 107(5), 509-519.

Royston, J. P. (1992), Approximating the Shapiro–Wilk's test for non-normality, *Statistics & Computing*, 2, 117–119.

Ruiz, J. E., I. Cordery, and A. Sharma (2007), Forecasting streamflows in Australia using the tropical Indo-Pacific thermocline as predictor, *Journal of Hydrology*, 341(3-4), 156-164.

Sáenz, J., C. Rodríguez-Puebla, J. Fernández, and J. Zubillaga, (2001) ,Interpretation of interannual winter temperature variations over southwestern Europe, *Journal of Geophysical Research*, 106, 20641–20652.

Salarijazi, M., A. Akhond-Ali, A. Adib, and A. Daneshkhah (2012), Trend and change-point detection for the annual stream-flow series of the Karun River at the Ahvaz hydrometric station, *African Journal of Agricultural Research*, 7(32), 4540-4552

Saito, K., and J. Cohen (2003), The potential role of snow cover in forcing interannual variability of the major Northern Hemisphere mode, *Geophysical Research Letters*, 30(6), 1302.

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata (1999), A dipole mode in the tropical Indian Ocean, *Nature*, 401(6751), 360-363.

Sanchez-Gomez, E., S. Somot, and A. Mariotti (2009), Future changes in the Mediterranean water budget projected by an ensemble of regional climate models, *Geophysical Research Letters*, 36(21), L21401.

Sanz-Elorza, M., E. D. Dana, A. González, and E. Sobrino (2003), Changes in the high-mountain vegetation of the central Iberian Peninsula as a probable sign of global warming, *Annals of Botany*, 92(2), 273-280.

Schlesinger, M. E., and N. Ramankutty (1994), An oscillation in the global climate system of period 65-70 years, *Nature*, 367(6465), 723-726.

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, and B. Rudolf (2014), GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, *Theorical Applied Climatology*, 115(1-2), 15-40.

Seierstad, I. A., D. B. Stephenson, and N. G. Kvamsto (2007), How useful are teleconnection patterns for explaining variability in extratropical storminess?, *Tellus A*, 59(2), 170-181.

Sen, P. K. (1968), Estimates of the regression coefficient based on Kendall's Tau, *Journal of the American Statistical Association*, 63(324), 1379-1389.

Seneviratne, S.I., N. Nicholls, D. Easterling, C.M. Goodess, S. Kanae, J. Kossin, Y. Luo, J. Marengo, K. McInnes, M. Rahimi, M. Reichstein, A. Sorteberg, C. Vera, and X. Zhang (2012), Changes in climate extremes and their impacts on the natural physical environment. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 109-230.

Serrano, A., J. A. García, V. L. Mateos, M. L. Cancillo, and J. Garrido (1999), Monthly modes of variation of precipitation over the Iberian Peninsula, *Journal of Climate*, 12(9), 2894-2919.

Sewell, R. D., and W. A. Landman (2001), Indo-Pacific relationships in terms of sea-surface temperature variations, *International Journal of Climatology*, 21(12), 1515-1528.

Shabbar, A., and W. Skinner (2004), Summer drought patterns in Canada and the relationship to global sea turface Temperatures, *Journal of Climate*, 17(14), 2866-2880.

Shaman, J. (2014), The seasonal effects of ENSO on European precipitation: Observational analysis, *Journal of Climate*, 27(17), 6423-6438.

Shaman, J., and E. Tziperman (2010), An atmospheric teleconnection linking ENSO and southwestern European precipitation, *Journal of Climate*, 24(1), 124-139.

Shaman, J., S. K. Esbensen, and E. D. Maloney (2009), The dynamics of the ENSO-Atlantic hurricane teleconnection: ENSO-related changes to the North African-Asian Jet affect Atlantic Basin Tropical Cyclogenesis, *Journal of Climate*, 22(9), 2458-2482.

Shapiro, S.S. and M.B Wilk (1965), An analysis of variance test for normality (complete samples), *Biometrika*, Vol. 52, No. 3/4, pp. 591-611.

Sheffield, J., and E. F. Wood (2008), Global trends and variability in soil moisture and drought characteristics, 1950–2000, from observation-driven simulations of the terrestrial hydrologic cycle, *Journal of Climate*, 21, 432–458.

Shrestha A, and R. Kostaschuk (2005), El Nino/Southern Oscillation (ENSO)-related variability in mean-monthly stream flow in Nepal, *Journal of Hydrology*, 308, 33–49.

Shukla, J., and J. L. Kinter (2005), Predictability of seasonal climate variations: A redagogical Review, In Palmer, T.N. and R. Hagedorn, eds, *Predictability of Weather and Climate*, Cambridge University Press.

Sittichok, K., A. Gado Djibo, O. Seidou, H. Moussa Saley, H. Karambiri, and J. Paturel (2014), Statistical seasonal rainfall and streamflow forecasting for the Sirba watershed, West Africa, using sea surface temperatures, *Hydrological Sciences Journa*l, DOI: 10.1080/02626667.2014.944526.

Sordo, C., M. D. Frías, S. Herrera, A. S. Cofiño, and J. M. Gutiérrez (2008), Interval-based statistical validation of operational seasonal forecasts in Spain conditioned to El Niño–Southern Oscillation events, *Journal of Geophysical Research: Atmospheres*, 113(D17), D17121.

Sorman, U. And Okur, A. (2000). Application of at site regional frequency analyses by using the L-moments technique, *Teknik Dergi*, 11(3): 2199-2218.

Soukup, T. L., O. A. Aziz, G. A. Tootle, T. C. Piechota, and S. S. Wulff (2009), Long lead-time streamflow forecasting of the North Platte River incorporating oceanic-atmospheric climate variability, *Journal of Hydrology*, 368(1), 131-142.

Stahl, K., H. Hisdal, J. Hannaford, L. M. Tallaksen, H. A. J. van Lanen, E. Sauquet, S. Demuth, M. Fendekova, and J. Jódar (2010), Streamflow trends in Europe: evidence from a dataset of near-natural catchments, *Hydrology and Earth System Sciences*, 14(12), 2367-2382.

Steinskog, D. J., D. B. Tostheim, and N. G. Kvamsto (2007), A cautionary note on the use of the Kolmogorov-Smirnov test for normality, *Monthly Weather Review*, 135(3), 1151-1157.

Stephens, M. A. (1974), EDF statistics for Goodness of fit and some comparisons, *Journal of the American Statistical Association*, 69(347), 730-737.

Tang, B., W. W. Hsieh, A. H. Monahan, and F. T. Tangang (2000), Skill comparisons between Neural Networks and Canonical Correlation Analysis in predicting the equatorial Pacific sea surface temperatures, *Journal of Climate*, 13(1), 287-293.

Tang, C., T. C. Piechota, and D. Chen (2011), Relationships between oceanic-atmospheric patterns and soil moisture in the Upper Colorado River Basin, *Journal of Hydrology*, 411(1-2), 77-90.

Tangang, F. T., B. Tang, A. H. Monahan, and W. W. Hsieh (1998), Forecasting ENSO events: A Neural Network-Extended EOF Approach, *Journal of Climate*, 11(1), 29-41.

Taschetto, A. S., and M. H. England (2009), El Niño Modoki impacts on Australian rainfall, *Journal of Climate*, 22(11), 3167-3174.

Tebaldi, C., K. Hayhoe, J. M. Arblaster, G. A. Meehl (2006), Going to the extremes, *Climate Change*, 79 (3–4), 185–211.

Thompson, D. W. J., and J. M. Wallace (1998), The Arctic oscillation signature in the wintertime geopotential height and temperature fields, *Geophysical Research Letters*, 25(9), 1297-1300.

Timilsena, J., T. Piechota, G. Tootle, and A. Singh (2009), Associations of interdecadal/interannual climate variability and long-term Colorado river basin streamflow, *Journal of Hydrology*, 365(3-4), 289-301.

Ting, M., and H. Wang (1997), Summertime U.S. precipitation variability and its relation to Pacific sea surface temperature, *Journal of Climate*, 10(8), 1853-1873.

Tomita, T., B. Wang, T. Yasunari, and H. Nakamura (2001), Global patterns of decadal-scale variability observed in sea surface temperature and lower-tropospheric circulation fields, *Journal of Geophysical Research: Oceans*, 106(C11), 26805-26815.

Tootle, G. A., and T. C. Piechota (2004), Suwannee river long range streamflow forecasts based on seasonal climate predictors, *Journal of the American Water Resources Association*, 40(2), 523-532.

Tootle, G. A., and T. C. Piechota (2006), Relationships between Pacific and Atlantic Ocean sea surface temperatures and U.S. streamflow variability, *Water Resources Research*, 42(7), W07411.

Tootle, G., A. Singh, T. Piechota, and I. Farnham (2007), Long lead-time forecasting of U.S. streamflow using Partial Least Squares Regression, *Journal of Hydrologic Engineering*, 12(5), 442-451.

Tootle, G. A., T. C. Piechota, and F. Gutiérrez (2008), The relationships between Pacific and Atlantic Ocean sea surface temperatures and Colombian streamflow variability, *Journal of Hydrology*, 349(3‚Äì4), 268-276.

Trenberth, K. E. (1984), Some effects of finite sample size and persistence on meteorological statistics. Part I: Autocorrelations, *Monthly Weather Review*, 112(12), 2359-2368.

Trenberth, K.E. (1999), Conceptual framework for changes of extremes of the hydrological cycle with climate change. *Climatic Change* 42 (1), 327–339.

Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N. C. Lau, and C. Ropelewski (1998), Progress during TOGA in understanding and modeling global teleconnections associated with tropical sea surface temperatures, *Journal of Geophysical Research: Oceans*, 103(C7), 14291-14324.

Trenberth, K. E., D. P. Stepaniak, and L. Smith (2005), Interannual variability of patterns of atmospheric mass distribution, *Journal of Climate*, 18(15), 2812-2825.

Trigo, R. M., and J. P. Palutikof (2001), Precipitation scenarios over Iberia: A comparison between direct GCM output and different downscaling techniques, *Journal of Climate*, 4422-4446.

Trigo, R. M., D. Pozo-Vázquez, T. J. Osborn, Y. Castro-Díez, S. Gámiz-Fortis, and M. J. Esteban-Parra (2004), North Atlantic Oscillation influence on precipitation, river flow and water resources in the Iberian Peninsula, *International Journal of Climatology*, 24(8), 925-944.

Tu, M., M. J. Hall, P. J. M. de Laat, and M. J.M. de Wit (2004), Detection of long-term changes in precipitation and discharge in the Meuse basin, GIS and Remote Sensing in Hydrology, Water Resources and Environment (Proceedings of ICGRHWE held at the Three Gorges Dam, China, September 2003), IAHS Publ., 289, 169-177.

Uvo, C. B., C. A. Repelli, S. E. Zebiak, and Y. Kushnir (1998), The Relationships between Tropical Pacific and Atlantic SST and Northeast Brazil Monthly Precipitation, *Journal of Climate*, 11(4), 551-562.

van der Schrier, G., K. R. Briffa, P. D. Jones, and T. J. Osborn (2006), Summer moisture variability across Europe, *Journal of Climate*, 19(12), 2818-2834.

van Loon, H., and J. C. Rogers (1978), The seesaw in winter temperatures between Greenland and Northern Europe. Part I: General Description, *Monthly Weather Review*, 106(3), 296-310.

van Oldenborgh, G. J., G. Burgers, and A. K. Tank (2000), On the El Niño teleconnection to spring precipitation in Europe, *International Journal of Climatology*, 20(5), 565-574.

Vicente Serrano, S.M. (2005), El Niño and La Niña influence on drought conditions at different time scales in the Iberian Peninsula, *Water Resources Research*, 41, W12415.

Vicente-Serrano, S. M. (2006a), Differences in spatial patterns of drought on different time scales: An analysis of the Iberian Peninsula, *Water Resources Management*, 20(1), 37-60.

Vicente-Serrano, S. M. (2006b), Spatial and temporal analysis of droughts in the Iberian Peninsula (1910-2000), *Hydrological Sciences Journal*, 51(1), 83-97.

Vicente-Serrano, S. M., and J. I. López-Moreno (2005), Hydrological response to different time scales of climatological drought: an evaluation of the Standardized Precipitation Index in a mountainous Mediterranean basin, *Hydrology and Earth System Sciences*, 9(5), 523-533.

Vicente-Serrano, S. M., and J. I. López-Moreno (2008), Nonstationary influence of the North Atlantic Oscillation on European precipitation, *Journal of Geophysical Research: Atmospheres*, 113(D20), D20120.

Vicente-Serrano, S. M., J. López-Moreno, S. Beguería, J. Lorenzo-Lacruz, C. Azorín-Molina, and E. Morán-Tejeda (2011), Accurate computation of a streamflow drought index, *Journal of Hydrologic Engineering*, 17(2), 318-332.

Vicente-Serrano, S. M., J. I. Lopez-Moreno, S. Beguería, J. Lorenzo-Lacruz, A. Sanchez-Lorenzo, J. M. García-Ruiz, C. Azorin-Molina, E. Morán-Tejeda, J. Revuelto, R. Trigo, F. Coelho, and F. Espejo (2014), Evidence of increasing drought severity caused by temperature rise in southern Europe, *Environmental Research Letters*, 9(4), 044001.

Vihma, T. (2014), Effects of Arctic sea ice decline on weather and climate: A review, *Surveys in Geophysics*, 35(5), 1175-1214.

Visbeck, M. H., J. W. Hurrell, L. Polvani, and H. M. Cullen (2001), The North Atlantic Oscillation: Past, present, and future, *Proceedings of the National Academy of Sciences*, 98(23), 12876-12877.

Vogel, R. M., and N. M. Fennessey (1993), L moment diagrams should replace product moment diagrams, *Water Resources Research,* 29(6), 1745-1752.

von Storch, H., and A. Navarra (1995), Analysis of climate variability: Applications of statistical techniques, Springer Verlag.

von Storch, and F. W. Zwiers (1999), Statistical analysis in climate research, 484 pp., Cambridge University Press, Cambridge.

Walker, G. T., and E. W. Bliss (1932), World Weather V. *Memoirs of the Royal Meteorological Society*, 4(36), 53-84.

Wallace, J. M., and D. S. Gutzler (1981), Teleconnections in the geopotential height field during the Northern Hemisphere winter, *Monthly Weather Review*, 109(4), 784-812.

Wallace, J. M., D. S. Gutzler, and C. S. Bretherton (1992), Singular value decomposition of wintertime sea surface temperature and 500-mb height anomalies, *Journal of Climate*, 5, 561–576.

Wang, H. L., and M. F. Ting (1999), Seasonal cycle of the climatological stationary waves in the NCEP-NCAR reanalysis, *Journal of the Atmospheric Sciences*, 56(22), 3892-3919.

Wang, H., and M. Ting (2000), Covariabilities of Winter U.S. Precipitation and Pacific sea surface temperatures, *Journal of Climate*, 13(20), 3711-3719.

Wang, W. (2006), Stochasticity, nonlinearity and forecasting of streamflow processes. Amsterdam: IOS Press.

Webster, P. J., A. M. Moore, J. P. Loschnigg, and R. R. Leben (1999), Coupled ocean-atmosphere dynamics in the Indian Ocean during 1997-98, *Nature*, 401(6751), 356-360.

Wei, F., L. Hu, G. Chen, Q. Li, and Y. Xie (2012), Reconstruction of summer sea level pressure over east Asia since 1470, *Journal of Climate*, 25(16), 5600-5611.

Wijngaard, J. B., A. M. G. Klein Tank, and G. P. Konnen (2003), Homogeneity of 20th century European daily temperature and precipitation series, *International Journal of Climatology*, 23(6), 679-692.

Wilhite, D. A., and M. H. Glantz (1985), Understanding the drought phenomenon: The role of definitions, *Water International*, 10(3), 111-120.

Wilks, D. S., 2006: Statistical methods in the Atmospheric Sciences. 2d ed. International Geophysics Series, Vol. 91, Academic Press, 627 pp.

Wilson, D., H. Hisdal, and D. Lawrence (2010), Has streamflow changed in the Nordic countries? Recent trends and comparisons to hydrological projections, *Journal of Hydrology*, 394(3-4), 334-346.

Woodhouse, C. A., S. T. Gray, and D. M. Meko (2006), Updated streamflow reconstructions for the Upper Colorado River Basin, *Water Resources Research*, 42(5), W05415.

WWAP (United Nations World Water Assessment Programme) (2014), The United Nations World Water Development Report 2014: Water and Energy, Paris, UNESCO.

Xoplaki, E., J. F. González-Rouco, J. Luterbacher, and H. Wanner (2004), Wet season Mediterranean precipitation variability: influence of large-scale dynamics and trends, *Climate Dynamics*, 23(1), 63-78.

Yu, J. Y., and S. T. Kim (2010), Relationships between extratropical sea level pressure variations and the central Pacific and eastern Pacific types of ENSO, *Journal of Climate*, 24(3), 708-720.

Yue, S., P. Pilon, B. Phinney, and G. Cavadias (2002), The influence of autocorrelation on the ability to detect trend in hydrological series, *Hydrological Processes*, 16(9), 1807-1829.

Zanchettin, D., S. W. Franks, P. Traverso, and M. Tomasino (2008), On ENSO impacts on European wintertime rainfalls and their modulation by the NAO and the Pacific multi-decadal variability described through the PDO index, *International Journal of Climatology*, 28(8), 995-1006.

Zorita, E., V. Kharin, and H. von Storch (1992), The atmospheric circulation and sea surface temperature in the North-Atlantic area in winter: their interaction and relevance for Iberian precipitation, *Journal of Climate*, 5, 1097-1108.

Zuur, A. F., E. N. Ieno, and C. S. Elphick (2010), A protocol for data exploration to avoid common statistical problems, *Methods in Ecology and Evolution,* 1(1), 3-14.

Zwiers, F. W., and H. von Storch (2004), On the role of statistics in climate research, *International Journal of Climatology*, 24(6), 665.-680.