This is a post-peer-review, pre-copyedit version of an article published in *Stochastic Environmental Research and Risk Assessment*. The final authenticated version is available online at: https://doi.org/10.1007/s00477-019-01733-8

Quantifying the role of individual flood drivers and their correlations in flooding of coastal river reaches

María Bermúdez [1,2,*], Luis Cea [2] and Javier Sopelana [3]

[1] University of Granada, Environmental Fluid Dynamics Group, Andalusian Institute for Earth System Research, Av. del Mediterráneo s/n, 18006, Granada, Spain.

[2] University of A Coruña, Environmental and Water Engineering Group, Department of Civil Engineering, Elviña, 15071, A Coruña, Spain

[3] Aquática Ingeniería Civil, Vigo, Spain

* Correspondence: mariabermudez@ugr.es / mbermudez@udc.es

ORCID iDs: 0000-0003-3189-4791 (M Bermúdez) / 0000-0002-3920-0478 (L Cea)

Abstract

Flooding in coastal river reaches is the result of complex interactions between coastal and inland drivers. Flood hazard assessments need to consider how these drivers interact in space and time, for which a standard method is currently lacking. A complex hydrodynamic model is required to reproduce the physics of the combined forcings and, at the same time, to fully explore the combinations of drivers that can occur in order to determine extreme flood frequencies. In this work, we explore the individual role of astronomical tide, storm surge and river discharge and their correlations in the extreme flood levels of a coastal river reach. We apply a computationally efficient surrogate model of a 2D shallow water model based on least squares support vector machines (LS-SVM) regression to reconstruct 10000 years-long time series of water levels in the reach. As input to the model, we consider an ensemble of synthetic time series of the flood drivers, which differ in the number of variables considered and in their correlations. Probabilities of exceedance of water levels are then computed and compared. The proposed methodology can give a better understanding of the flooding processes in a multivariable environment, as low-lying coastal urban areas typically are, and can provide guidance on where to focus modelling efforts when developing flood hazard assessments in such areas.

Keywords: coastal river reach, compound flooding, extreme floods, flood inundation modelling, LS-SVM

Acknowledgments: María Bermúdez gratefully acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement Nº 754446 and UGR Research and Knowledge Transfer Found – Athenea3i

1. Introduction

Standard methodology for flood hazard assessment relies on the estimation of floods of different return period, and the subsequent simulation of such floods with a hydrodynamic model to produce inundation maps (de Moel et al. 2015). When the inundation depth depends only on the river discharge, a one-to-one relation can be assumed between both variables for any given return period, i.e. the 100-year discharge will generate the 100-year depth. However, this is not true when several drivers are involved in the flooding process, as it is the case in coastal urban areas, which are typically exposed to multiple flood drivers such as sea water level, waves, river discharge, and local precipitation. Flooding may thus arise from a combination of sea-induced and precipitation-induced inland flooding, in what is sometimes referred to as coincident or compound flooding. In such a compound flood, the individual contributing variables may not be extreme themselves, but it is their combination that renders an event exceptional and leads to an extreme water depth. The correlations and dependencies between flood drivers (e.g., coastal surge and river discharge or rainfall) can be significant along riverine and estuary areas, as evidenced by Svensson and Jones (2002) in eastern Britain, Petroliagkis et al. (2016) in Europe, or Wahl et al. (2015) in the USA. Ignoring the dependence among them would result in an underestimation of risk (Bevacqua et al. 2017), whereas assuming total dependence would be too conservative and therefore inadequate for flood risk planning.

Compound hazards have gained attention in climate science, and current research is advancing on frameworks and tools for their characterization (Hawkes 2008; Seneviratne et al. 2012; Leonard et al. 2014; Zscheischler et al. 2018; Sadegh et al. 2018). One possibility is to use continuous long term simulation methods, which consist of running a hydrodynamic flood model during a long period of time (several years) driven by simultaneous time series of the flood drivers. In this way, the statistical dependence between the drivers does not need to be explicitly analyzed, which represents a clear advantage over multivariate analysis methods such as copulas or Bayesian methods (Van Den Hurk et al. 2015; Xu et al. 2017, 2019; Sadegh et al. 2018). Methods that rely on this approach have recently been developed by Falter et al. (2015) for large-scale basins or by Schumman et al. (2016) for the whole Australian continent. This method has also been successfully applied by the authors to estimate extreme inundation caused by both high sea levels and river discharges in a coastal town (Sopelana et al. 2018).

A typical constraint of continuous simulation methods is the lack of availability of simultaneous time series of the drivers, long enough to be statistically representative. To this regard, in the last years efforts have been done to improve the availability of long-term instrumental records (Peterson and Manton 2008; Brunet and Jones 2011; Brunet et al. 2014), as well as new climate data sources such as model simulations, model-based reanalyses and remote sensors (Overpeck et al. 2011). It is also possible to generate synthetic time series for providing sufficiently large samples or ensembles of different time series of the same process (Keylock 2012; Efstratiadis et al. 2014; Sopelana et al. 2018).

Another drawback of continuous simulation approaches is the high computational time required to perform long term simulations. In spite of the various High Performance Computing techniques available nowadays, which have been applied to 2D shallow water codes (Vacondio et al. 2014; Liu et al. 2018; García-Feal et al. 2018), it is still unfeasible to perform detailed 2D inundation simulations spanning hundreds of years with a time resolution of a few minutes. This is why most studies that apply this method rely on simple approaches to compute the flood characteristics (e.g., rating curves or 1D hydrodynamic models), rather than performing 2D hydrodynamic simulations (Falter et al. 2016). Physical processes like complex channel-floodplain interactions, flow paths through urban areas or superposition of flood waves at river confluences can, however, only be properly reproduced through spatially-detailed 2D shallow water model simulations (de Almeida et al. 2018; Bermúdez and Zischg 2018). In cases where such processes are relevant, the development of fast surrogate models that emulate the original 2D model responses can be a suitable strategy to reduce in a significant manner the computational time. In this line, artificial intelligence (AI) based methods such as support vector machines (SVM) or artificial neural networks (ANN) are becoming increasingly popular as function approximation techniques in water resources modelling (Razavi et al. 2012; Kasiviswanathan and Sudheer 2013; Yaseen et al. 2015). Although applications to flood inundation modelling are still scarce, previous studies applying ANNs

(Chang et al. 2010; Bermúdez et al. 2018), Radial Basis Function ANNs (Sopelana et al. 2018) and SVMs (Lin et al. 2013; Liu and Pender 2015; Cea et al. 2016; Jhong et al. 2017; Bermúdez et al. 2019) for this purpose have obtained comparable predictions to physically-based models, while significantly increasing the simulation speed.

In this work we apply a continuous long term simulation approach to quantify the individual role of astronomical tide, storm surge and river discharge and their correlations in the flooding of a coastal river reach. The continuous simulation approach uses a computationally efficient surrogate model of a 2D shallow water model based on LS-SVM regression. In this way we reconstruct 10000 years-long time series of water levels from synthetic time series of the flood drivers, using a limited number of 2D shallow water simulations (250 days) to calibrate the LS-SVM model. To quantify the role of the flood drivers, we consider an ensemble of time series which differ in the number of flood drivers considered and the correlations between them. Probabilities of exceedance of water levels are then computed and compared from the resulting long-term water level series. The ultimate aim of the study is to provide guidance on where to focus modelling efforts when developing flood hazard assessment approaches in low-lying coastal areas.

2. Materials and Methods

2.1. Study site

The coastal town of Betanzos (Figure 1), located in the NW of Spain, was used as the study site for this work. Betanzos is located at the confluence of two rivers (Mandeo and Mendo) in the inner part of a macrotidal estuary with a spring tidal range of roughly 4.5 m. The Mandeo River has a length of about 50 km. Its drainage basin occupies a total area of 450 km^2, of which 100 km^2 correspond to its tributary, the Mendo River. The Mendo river flows across a wide floodplain (approximately 400 meters) dominated by marsh.

Several interventions carried out during the last century have completely changed the natural state of the area. The river has been artificially channelized downstream the confluence, the floodplains have been developed for urban use, and a railway embankment has been built just downstream the confluence. These changes have increased the vulnerability of the area and consequently the flood risk.

A set of 8 representative control points distributed along the Mandeo River were defined for the analysis of results (Figure 1). The location of the control points was chosen in order to sample regions with a different exposure to the sea level influence. Point 1 is located in the estuary, point 2 in the middle of the artificial channel, point 3 at the confluence with the river Mendo, and points 4 to 8 along the Mandeo river.

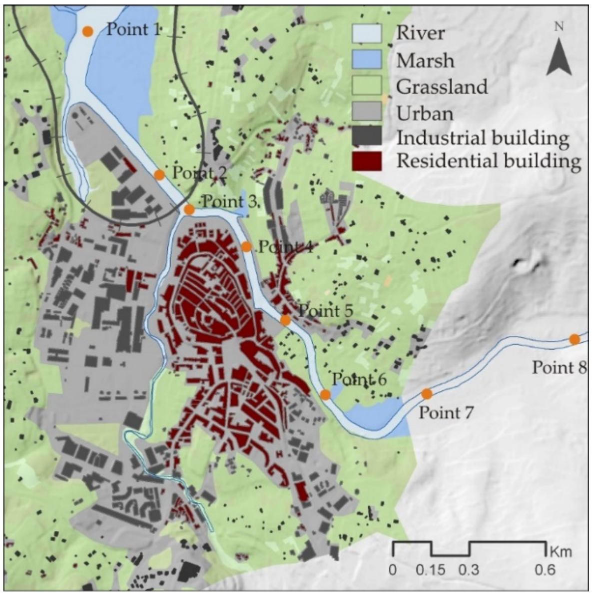

Figure 1. Schematic map of the town of Betanzos, showing land uses and building locations. Location of control points along the river Mandeo are also shown.

2.2. Long-term time series of the flood drivers

The lack of observed data during a time period long enough to be statistically representative is one of the main problems in extreme flood analysis. For this reason, data selection and preparation are probably the most important elements in extreme value analyses (Hawkes 2008) and thus, the way in which these data are generated makes an important difference between methodologies. Another important issue related to input data is its time resolution. In order to define representative time series of the drivers, the time scale of each physical phenomenon that has an effect on the inundation must be taken into account.

The continuous simulation technique used in this work to evaluate flood hazard is based on the generation of synthetic long-term time series of the flood drivers (predictors) with a daily resolution. Those series must be simultaneous, and they must reflect the observed seasonality and mutual correlations between predictors, in order to correctly account for the probability of simultaneous occurrence of extreme values of the flood drivers.

In the coastal river reach analyzed in this work, inundation depths are affected by the tidal level but not by ocean waves. In this case, the most relevant flood predictors at a daily scale are: the daily astronomical tidal range (TR), the daily maximum storm surge (Sd), the daily peak discharge (Qd) and the time lag between peak discharge and high tide (Tlag). In order to have a statistically representative sample of these four predictors for the estimation of extreme events, we generated daily time series that span over 10000 years. Such a length is necessary because in the proposed method, return values are estimated from the long-term reconstructed time series of water depth using a simple plotting position formula, without fitting any statistical distribution. This kind of direct estimation is robust for return periods much lower than the length of the reconstructed time series. The methodology followed to generate the synthetic time series is described in detail in (Sopelana et al. 2018) and therefore, only a brief overview is given in the following.

The astronomical tide is a deterministic variable and therefore, the TR time series were generated from the tidal harmonic constituents at the study site, obtained from historical records of sea level measured at a tidal gauge located in the outer estuary.

The synthetic time series of Qd were obtained from a regression regional hydrological model based on the following descriptors: mean annual precipitation, catchment area, mean catchment slope and mean SCS curve number. The model was calibrated using observed discharge data at 18 gauge stations located in the hydrological region where the study site is located.

Regarding the storm surge, it has a strong seasonal variability (higher in winter and lower in summer), and at the same time it is somewhat correlated with the river discharge, with a correlation coefficient that varies from one month to another. Both effects have been considered in the generation of the synthetic time series of Sd, that take into account the seasonality and the monthly correlation between Qd and Sd. The procedure to generate Sd is based on the observed mean and standard deviation of the daily surge and its observed correlation with Qd, computed on a monthly basis from historical time series from 1992 to 2014. The tidal harmonic constituents and surge were extracted by Pérez-Gómez (2014). For more details, the reader is referred to (Sopelana et al. 2018).

The largest tidal constituent in this area is the principal lunar semidiurnal (or M2), with a period of approximately 12.42 hours. The synthetic series of Tlag were generated as a random value between 0 and 12.42 h following a uniform distribution, since the daily tidal range and the river discharge are completely uncorrelated and have the same probability of occurrence at any time within the day.

The previous procedure produces synthetic time series that reproduce the observed seasonality and correlations between the Qd, Sd and TR, as shown in (Sopelana et al. 2018).

2.3. Two-dimensional shallow water equation modelling

The 2D inundation model Iber (Bladé et al. 2014; García-Feal et al. 2018) was used to transform the flood drivers described above in water depths in the study region. The model solves the 2D depth-averaged shallow water equations using a high resolution unstructured finite volume solver. The numerical mesh used in the computations has 126,266 elements with an average size of 11 m. The mesh resolution in the river and urban area is higher, with a mesh size of the order of 5 m. The topography is defined from a Digital Surface Model

(DSM), obtained by combining the river and estuary bathymetries with Light Detection and Ranging (LiDAR) terrain data. Bed roughness is defined with a variable Manning coefficient obtained from a land use chart. Six different land uses were defined, with Manning values ranging from 0.02 s/m$^{1/3}$ in the main river channels to 0.15 s/m$^{1/3}$ in the residential areas.

Considering that river discharge and sea level vary significantly within a day, the daily values of the flood drivers Qd, Sd and TR must be downscaled to a higher time resolution in order to predict flood hazard from the maximum instantaneous water depths and velocities. The tidal range was downscaled using the tidal harmonic constituents at the study site. The surge was assumed to be constant over 24 hours and is implemented as an increase in the mean sea level. The river discharge was downscaled using the Soil Conservation Service (SCS) unit hydrograph. (Fill and Steiner 2003; Taguas et al. 2008).

The computation of the maximum water depths with the inundation model was done in a daily basis, using the corresponding downscaled flood drivers as boundary conditions. The downscaled tidal level and surge were used as the downstream boundary condition, while the river hydrograph was imposed at the upstream inlet boundary (Figure 2). An appropriate offset was introduced in the time series of the boundary conditions in order to respect the time lag between peak discharge and high tide. For the purposes of this work, the output of the model is the daily maximum water depth at the 8 control points defined in Figure 1.

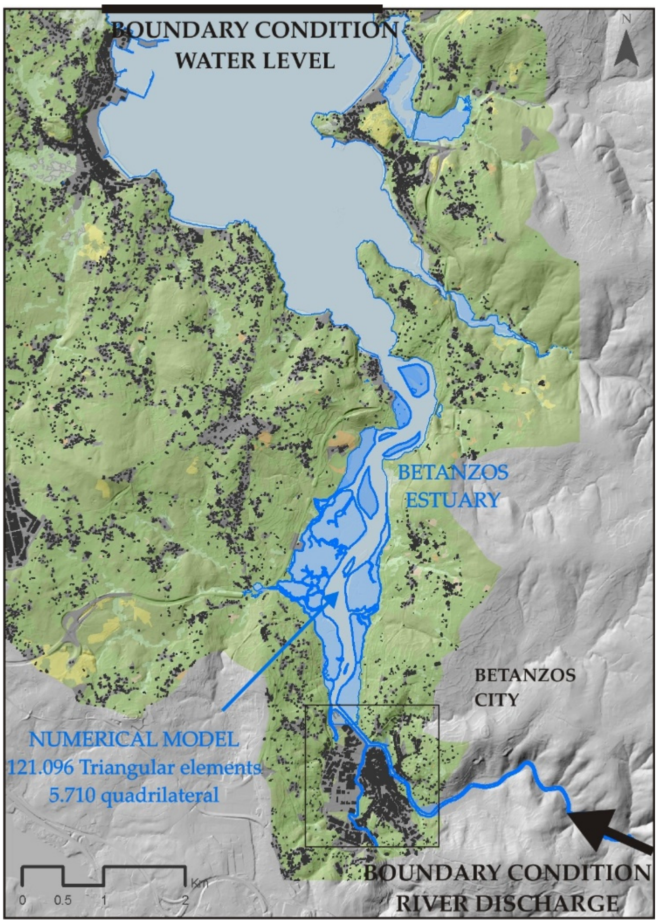

Figure 2. Spatial domain of the numerical simulations and location of the open boundaries.

2.4. Least-squares support vector machine regression modelling

In the present study case, the 2D high-resolution inundation model takes about 1 hour of CPU time to simulate one day of real time. This prevents the use of this model to simulate the water depths during 10000 years, since such a computation would take around 400 years of CPU time. Instead, a limited number of representative days were simulated with the 2D inundation model, and the water depth results obtained in those simulations were used to calibrate and validate an efficient surrogate model based on least-squares support vector machine (LS-SVM) regression (Vapnik 1998; Suykens et al. 2002). LS-SVMs are a non-parametric regression technique derived from the original SVM model (Vapnik 1998), which share the small sample learning and generalization abilities of SVM, but have the additional advantage of transforming a quadratic programming problem into a linear one. In this work, models were developed with StatLSSVM toolbox in Matlab software (Brabanter et al. 2013). The LS-SVM technique was used to transform the long-term daily time series of the flood drivers into long-term daily series of maximum water depths at the control points.

The calibration of the LS-SVM model was done using the numerical results obtained in 250 simulations performed with the 2D inundation model, with values of the flood drivers selected randomly over their range of variation. The tidal range is between 4.0 and 0.8 m and the storm surge between -0.53 m and 0.82 m. Daily peak discharge ranges from 0 to 446 m^3/s, and its time lag with high tide takes values from 0 to 12.42 h. To validate the predictions of the LS-SVM model, a representative sample of 100 characteristic days was selected from the long-term time series of the drivers, using for that purpose the Maximum Dissimilarity selection algorithm (Kennard and Stone 1969). This algorithm identifies a subset of cases that represents the diversity of the data, based on the Euclidean distance from each other in the multi-dimensional space of input data. As shown by Camus et al. (2011) for wave climate analysis, the subset selected is distributed fairly evenly across the space with some points selected in the outline of the data space. In this way, not only mean conditions but also extremes that can result in flood events are represented in the subset. The 100 characteristic days selected in such a way were simulated with the 2D inundation model, and the maximum water depths computed were compared with those predicted by the calibrated LS-SVM model.

Once the LS-SVM model was calibrated and validated at each control point, it was applied to reconstruct the 10000-year time series of daily maximum water depths at reduced computational cost. The calibration process takes about a second for each point. The calculation time required to reconstruct a time series of 10000 data is in the order of 20 seconds.

3. Results and discussion

3.1. Sensitivity analysis

Before applying the LS-SVM regression model, a sensitivity analysis of the maximum water depth at the control points to the flood drivers was conducted using the results of a series of characteristic cases modelled with the 2D inundation model. Sensitivity indices, based on both linear and non-linear non-parametric regression, were calculated to quantify the effects of each individual flood driver on the maximum water depth at each control point.

The Standardized Regression Coefficients (SRC) obtained from a multivariable linear regression of maximum water depths at each control point are shown in Figure 3. The coefficient of determination R^2 of the multiple linear regression for control points P1 to P8 is respectively 0.93, 0.91, 0.94, 0.95, 0.95, 0.96, 0.97, 0.97. These relatively high values of R2 confirm the validity of the SRC sensitivity measures (Storlie et al. 2009). According to the SRC values, the most relevant parameter at all the control points is the river discharge (Qd), followed by the tidal range (TR) and the surge (Sd). As expected, they all take positive values, meaning that an increase in the parameter implies an increase in the water depth. The sensitivity to the river discharge increases significantly as we move upstream the river reach (from control point 1 to 8), while the sensitivity to the sea level related parameters (TR and Sd) decreases. The maximum water depths show a very low sensitivity to the time lag (Tlag), regardless of the control point considered. Similar results can be inferred from the first order and total effects variance-based sensitivity indices obtained from the ANOVA decomposition (S_i and $S_{T,i}$ values in Figure 3 and Table 1). The total effect of a given parameter on model output ($S_{T,i}$ in Table 1) is given by all the first and second order terms of the ANOVA decomposition in which the parameter appears. The differences between the first order and total effects reflect non-linear interactions among model parameters. At all points, the total effect of the time lag is very close to zero, which corroborates the limited influence of this parameter on the water depth in our study case.

Table 1. First order effects and total effects for the water depth at control points.

	P1		**P2**		**P3**		**P4**		**P5**		**P6**		**P7**		**P8**	
	S_i	$S_{T,i}$	S_i	$S_{T,i}$	S_i	$S_{T,i}$	S_i	$S_{T,i}$	S_i	$S_{T,i}$	S_i	$S_{T,i}$	S_i	$S_{T,i}$	S_i	$S_{T,i}$
Qd	0.301	0.477	0.343	0.534	0.469	0.641	0.602	0.745	0.724	0.819	0.774	0.865	0.793	0.878	0.809	0.887
TR	0.281	0.310	0.247	0.283	0.171	0.202	0.107	0.129	0.059	0.072	0.041	0.054	0.034	0.046	0.029	0.041
Sd	0.131	0.129	0.114	0.114	0.079	0.080	0.049	0.052	0.027	0.027	0.019	0.020	0.016	0.016	0.014	0.013
Tlag	0.000	0.001	0.000	0.001	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

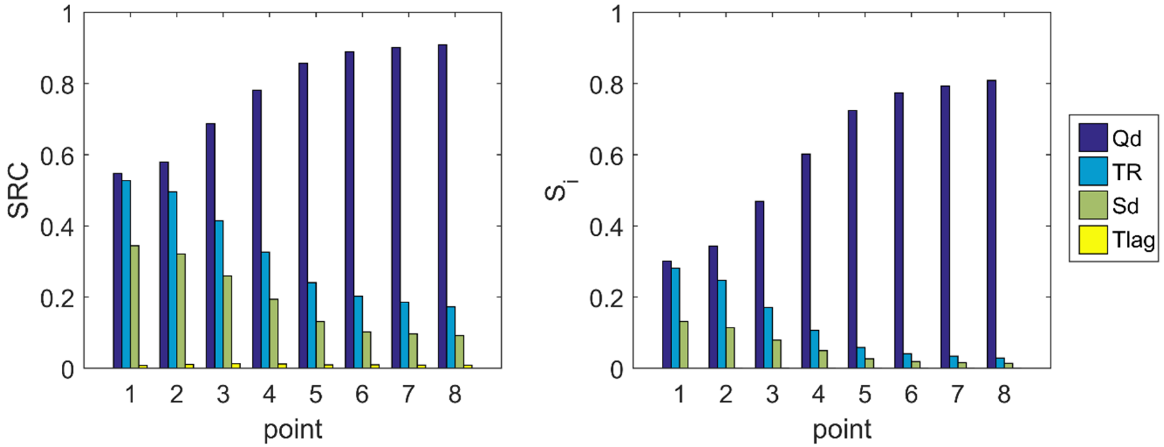

Figure 3. Standardized regression coefficient and first order sensitivity index for the water depth at each control point.

3.2. LS-SVM model performance

Based on the results of the sensitivity analysis, a LS-SVM model was calibrated using only Qd, TR and Sd as predictor variables. Given the little influence of Tlag on the water depth predictions, it was not considered as an input parameter in the LS-SVM model.

The performance of the LS-SVM model was quantified comparing the daily maximum water surface elevation (wse) predictions with those obtained with the 2D inundation model in 100 characteristic days (Figure 4). Considering the ensemble of validation runs, the mean absolute error (MAE) is below 10 cm at all points in validation. The global MAE on the 8 points is 7.2 cm. Thus, the LS-SVM model can be considered as an efficient surrogate of the 2D inundation model in order to reconstruct time series of water levels in the river reach from time series of the flood drivers.

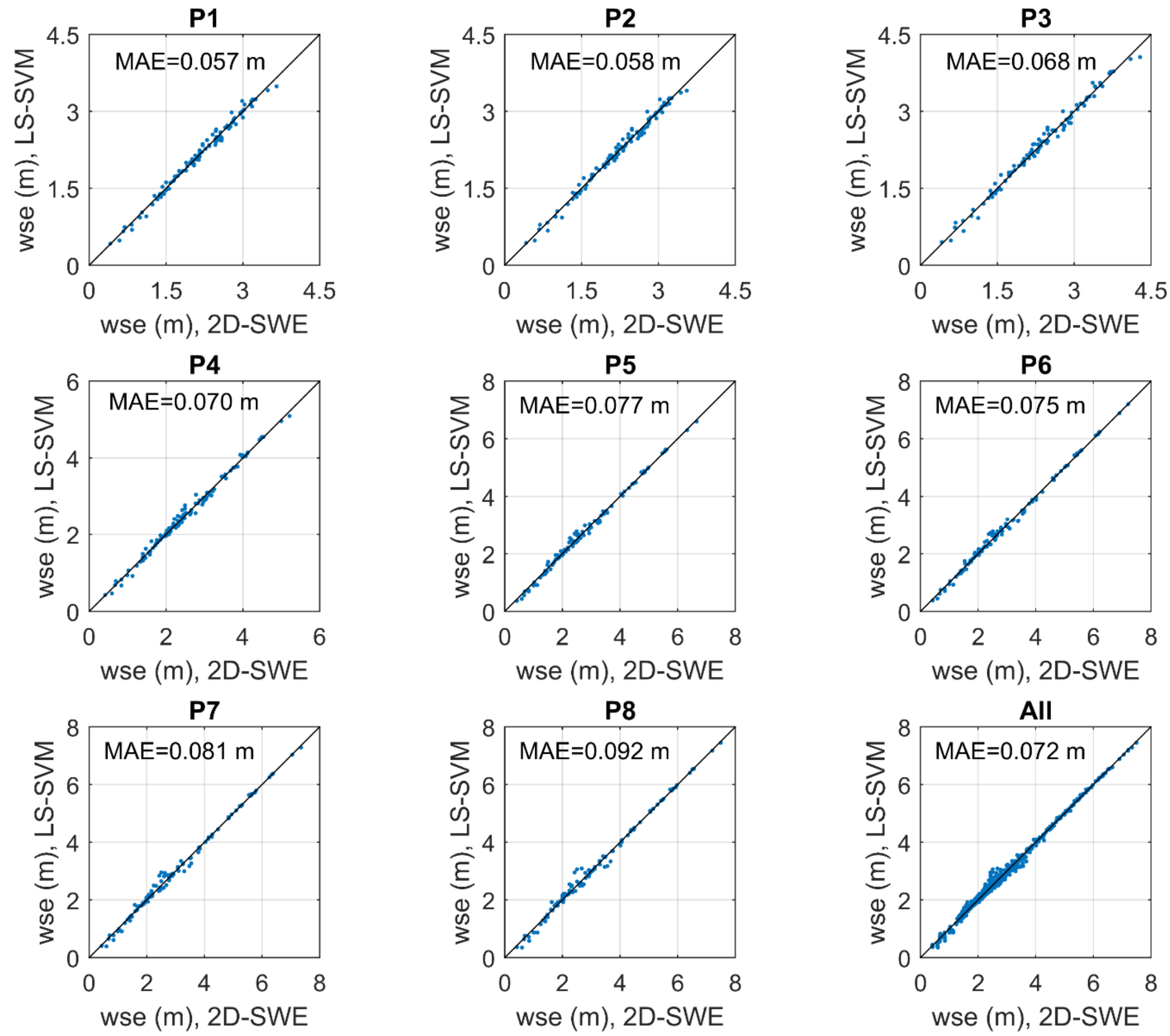

Figure 4. Scatter plot of water levels computed with the 2D-SWE model and the LS-SVM model in validation, with the 1:1 line plotted for reference in black. The mean absolute error in each point and the global mean absolute error, considering the 8 control points, for the 100 validation runs is indicated in the upper left-hand corner of the corresponding subfigure.

3.3. Annual maximum water depths

In order to analyse the combinations of the flood drivers that are responsible of the maximum water surface elevations, the values of Qd, TR and Sd associated with the annual maxima at each control point were extracted from the 10000 year-long time series. The results of this analysis at control point 1 are shown in Figure 5. At points of the estuary such as this one, annual maximum water levels are in general associated with high tidal ranges and positive surges. However, annual maxima also occur during low tidal range conditions, associated with high river flows. A certain discharge threshold can be established to distinguish between these two types of maxima. As illustrated in Figure 5 for point 1, the value of the threshold can be set at approximately 210 m^3/s at this point. If this threshold is exceeded, inundation levels are driven mainly by the river discharge, regardless of the sea level conditions.

The number of annual maxima corresponding primarily to high discharges increases as we move upstream the river Mandeo (e.g., from control point 2 to point 4 and then 6, as shown in Figure 6). In parallel, the value of the discharge threshold decreases (it is approximately 200 m^3/s at control point 2, 180 m^3/s at point 4 and 140 m^3/s at point 6), since the river discharge tends to dominate the inundation levels at the upstream locations.

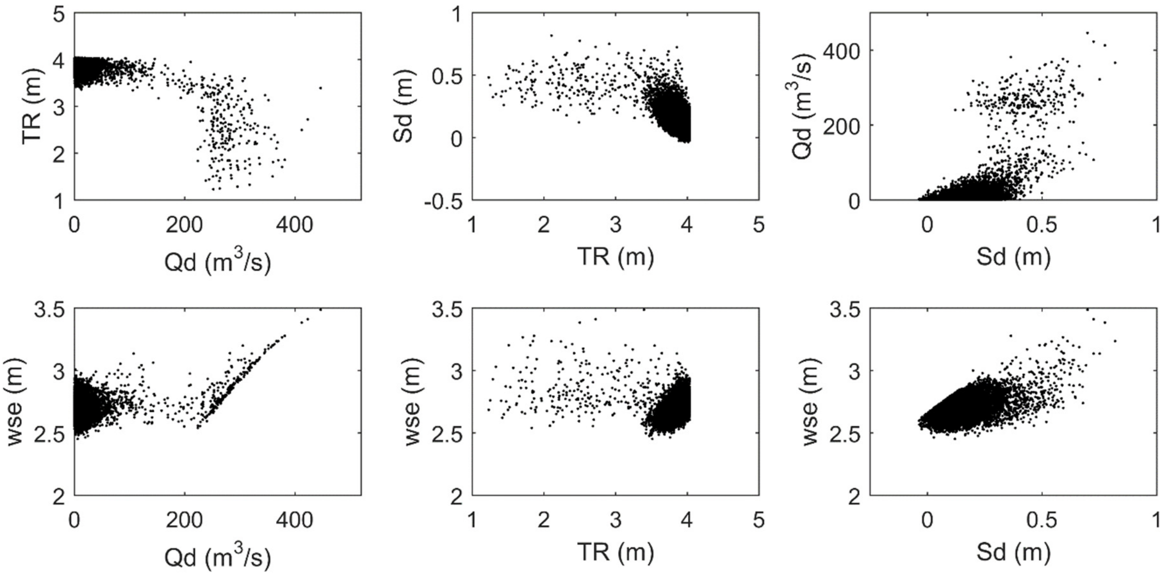

Figure 5. Annual maximum water surface elevation at control point 1 and associated daily peak discharge (Qd), astronomical tidal range (TR) and storm surge (Sd) conditions.

3.4. Water depth exceedance probabilities

The probability of exceedance of a given water depth was computed from the annual maxima of the 10000-year time series. Given the length of the series, return period values up to 500-years were obtained without the need of fitting a statistical distribution. Water level frequencies at four control points along the reach, expressed as return period, are plotted in Figure 7. It should be noted that the methodology used allows the estimation of the return period, jointly considering the relevant flood drivers and the combinations of them that can coexist. It differs from standard univariable approaches for defining return level events, based on a single flood driver, which have proven insufficient in other coastal river reaches (Serafin et al. 2019). It also moves away from simplified multivariable approaches in which the return period of the water level is assumed to be the same as that of its drivers (MARM 2011), or methods that use a relatively arbitrary combination of return periods of the flood drivers (Hawkes 2006). The reader is referred to (Serinaldi 2015) for discussion about multivariate return periods.

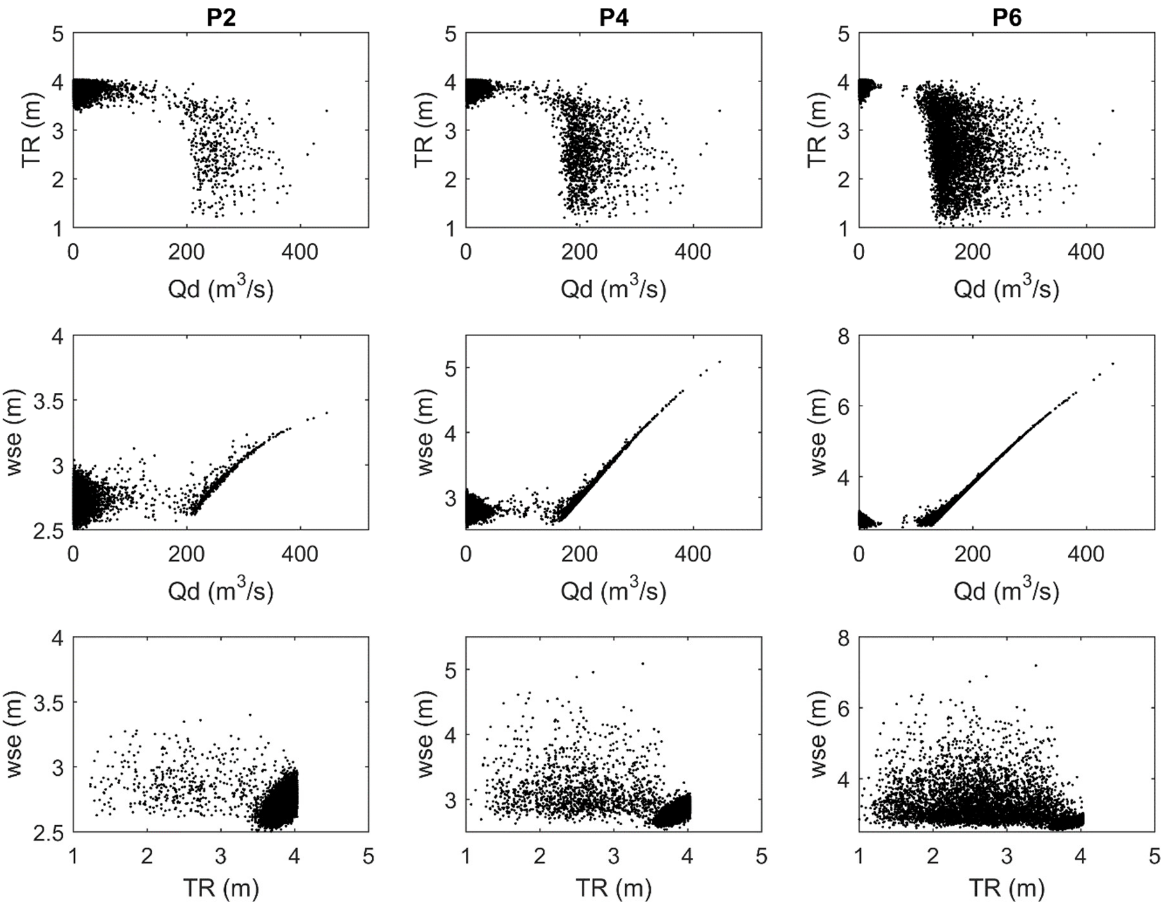

Figure 6. Annual maximum water surface elevation and associated daily peak discharge (Qd) and astronomical tidal range (TR) conditions at control point 2 (first column), control point 4 (middle column) and control point 6 (right column).

In order to further analyse the influence of the flood drivers on the inundation depths, two additional synthetic 10000-years long time series of the flood drivers were generated and converted in water depths using the LS-SVM model. In the first synthetic time series the storm surge was set to zero and therefore, the sea level only depends on the tidal range. In the second synthetic time series the correlation between the river discharge and the storm surge was neglected. The existing correlation between these two parameters in the original time series implies that high storm surges coincide with high river discharges, whereas low storm surges occur during low flows in the river. In the second synthetic time series, storm surge values are generated independently of discharge values and, compared to the original time series, are lower during high flows (and higher during low flows). The water level exceedance curves obtained with these two additional simulations are plotted in Figure 7.

At points close to the sea, neglecting the storm surge contribution results in a significant underestimation of the inundation level. The underestimation is observed for all return periods, although it is lower for high return periods. For example, at control point 1 the differences are reduced for return periods above 100 years (Figure 7). Such high return periods are associated with very high discharges, that not always coincide with very high sea levels, as mentioned in the sensitivity analysis. Also, as shown in Figure 5, the highest water levels are always associated to high river discharges, but not necessarily to high sea levels.

The influence of the sea level condition decreases as we move away from the river mouth. At control point 3, located at the confluence with the river Mendo, neglecting the storm surge altogether in the simulation makes practically no difference in the water levels associated to a return period equal or higher than 50 years. Further upstream, at control point 5, the underestimation related to the exclusion of the storm surge occurs only for return periods below 5 years, and its magnitude is negligible compared to points closer to the sea. Upstream of point 5, the effect of the storm surge on the water levels is negligible for all return periods. Therefore, this procedure allows locating the boundary between sea-influenced and river-influenced water levels in a coastal river reach.

Regarding the correlation between storm surge and river discharge, its influence on the inundation water levels is much lower. At control point 1 it leads to a very slight underestimation of the water depth (of the order of 5 cm) or return periods above 100 years. At other control points its effect is negligible (Figure 7).

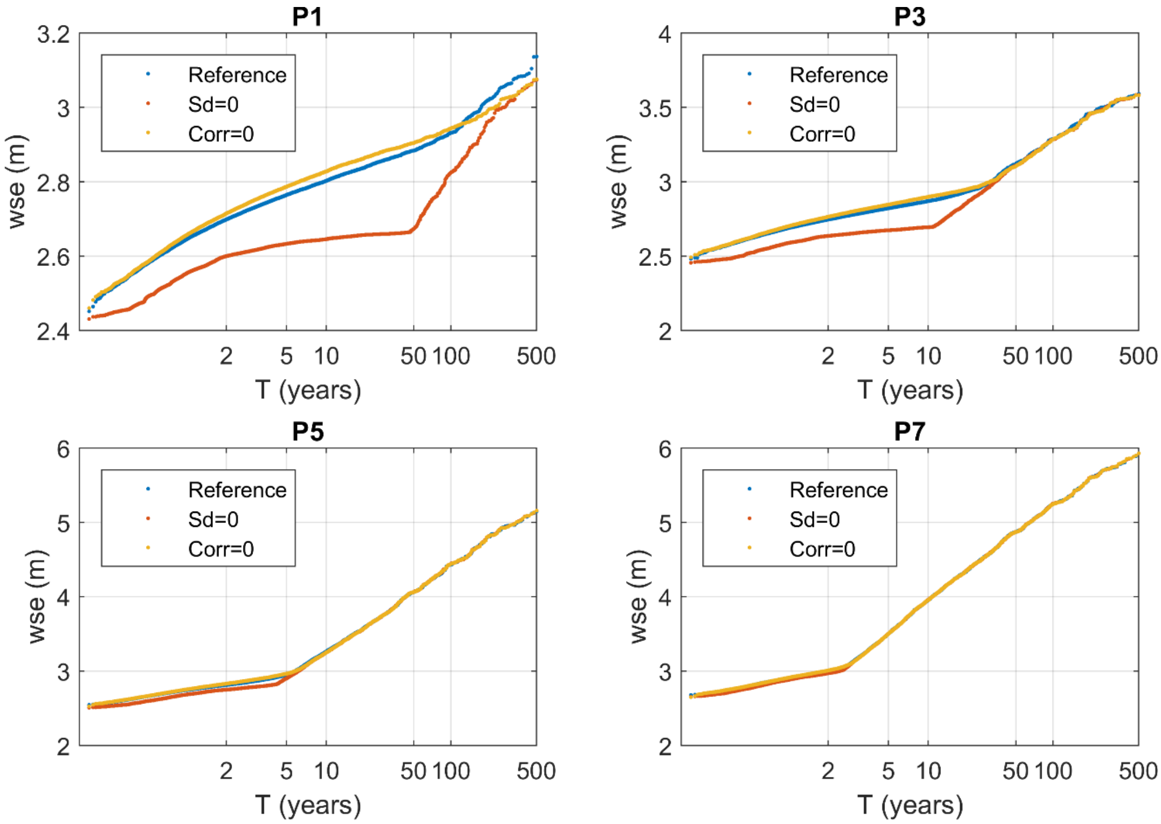

Figure 7. Water level frequency distribution obtained at points P1, P3, P5 and P7 with the original simulation (Reference), the simulation that neglects the storm surge (Sd=0) and the simulation that neglects the correlation between the storm surge and the river discharge (Corr=0).

The inundation levels obtained with the original and synthetic time series are compared in terms of associated return periods in Figure 8. This allows us to evaluate the changes in return period (instead of water levels) obtained with the different input time series. The results shown in Figure 8 confirm the necessity of taking into account the contribution of storm surge on extreme water levels in the river mouth. For example, the water level that has a 10-year return period considering the original time series at control point 1 would have a return period of nearly 100 years if the storm surge was not considered. This gives a better idea of how return water level estimates can vary if the contribution of certain drivers or correlations between drivers are neglected, depending on the location within the river reach.

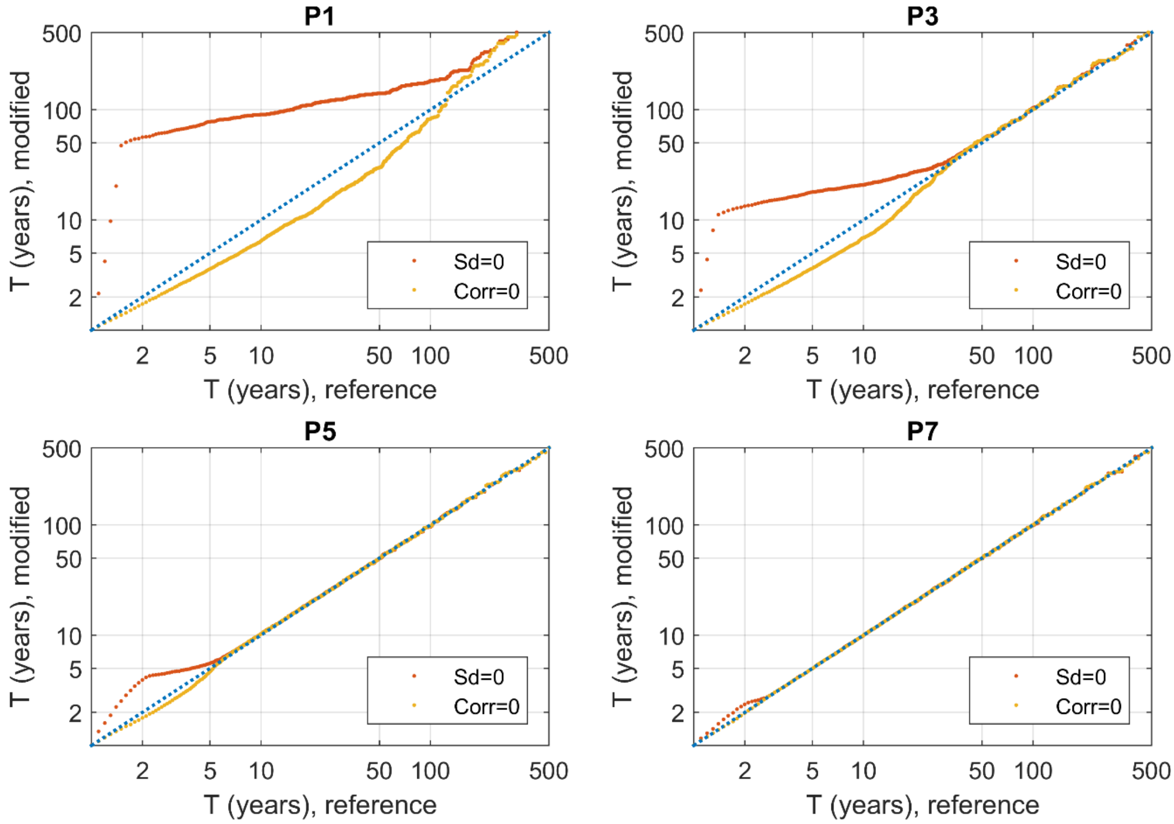

Figure 8. Changes in return period between the original simulation (x-axis) and the simulation that neglects the storm surge (Sd=0), and the simulation that neglects the correlation between the storm surge and the river discharge (Corr=0) (y-axis). The 1:1 line representing no changes in return period is plotted in blue.

4. Conclusions

We have developed a method to quantify the role and interactions between coastal and inland flood drivers in a coastal river reach. The methodology relies on the reconstruction of water levels from the flood drivers, by means of the combination of a 2D inundation model and a LS-SVM regression model. With the proposed approach it is possible to:

(1) Reconstruct long-term time series of water depth from synthetic time series of the flood drivers using a calibrated LS-SVM regression model.
(2) Determine return water levels jointly considering the relevant flood drivers and the combinations of them that can coexist. Since the methodology is computationally very efficient, it could be extended to generate flood maps for different return periods by simply increasing the number of control points considered (Bermúdez et al. 2019).
(3) Define the extension of the sea influence in a coastal river reach and, more specifically, the spatial domain where the estimation of extreme inundation levels requires considering the interaction between sea-level and river discharge.
(4) Quantify the role of individual flood drivers and their correlations for generating extreme flood events along the reach.

In the present study case, the most relevant flood predictors are the river discharge, the tidal range and the storm surge. Although the sensitivity to the river discharge decreases significantly as we move downstream the river reach, it is the most relevant parameter at all the control points. Discharge thresholds were identified to distinguish between sea-influenced and river-influenced annual maxima at each point. The analysis confirms the necessity of taking into account the contribution of storm surge to extreme water levels in the river mouth, and the possibility of neglecting its correlation with the river discharge.

In other river reaches the relative influence of the flood drivers might differ from our study case, and other processes as ocean waves or the time lag between high tide and peak discharge might have a significant influence on the inundation levels. Nevertheless, the methodology of analysis proposed here would still be valid in those cases.

References

Bermúdez M, Cea L, Puertas J (2019) A rapid flood inundation model for hazard mapping based on least squares support vector machine regression. J Flood Risk Manag e12522. doi: 10.1111/jfr3.12522

Bermúdez M, Ntegeka V, Wolfs V, Willems P (2018) Development and Comparison of Two Fast Surrogate Models for Urban Pluvial Flood Simulations. Water Resour Manag 32:2801–2815. doi: 10.1007/s11269-018-1959-8

Bermúdez M, Zischg AP (2018) Sensitivity of flood loss estimates to building representation and flow depth attribution methods in micro-scale flood modelling. Nat Hazards 92:1633–1648. doi: 10.1007/s11069-018-3270-7

Bevacqua E, Maraun D, Hobæk Haff I, et al (2017) Multivariate statistical modelling of compound events via pair-copula constructions: analysis of floods in Ravenna (Italy). Hydrol Earth Syst Sci 21:2701–2723. doi: 10.5194/hess-21-2701-2017

Bladé E, Cea L, Corestein G, et al (2014) Iber: herramienta de simulación numérica del flujo en ríos. Rev Int Métodos Numéricos para Cálculo y Diseño en Ing 30:1–10. doi: 10.1016/j.rimni.2012.07.004

Brabanter K De, Suykens JAK, Moor B De (2013) Nonparametric Regression via **StatLSSVM**. J Stat Softw 55:1–21. doi: 10.18637/jss.v055.i02

Brunet M, Jones P (2011) Data rescue initiatives: bringing historical climate data into the 21st century. Clim Res 47:29–40. doi: 10.3354/cr00960

Brunet M, Jones PD, Jourdain S, et al (2014) Data sources for rescuing the rich heritage of Mediterranean historical surface climate data. Geosci Data J 1:61–73. doi: 10.1002/gdj3.4

Camus P, Mendez FJ, Medina R, Cofiño AS (2011) Analysis of clustering and selection algorithms for the study of multivariate wave climate. Coast Eng 58:453–462. doi: 10.1016/j.coastaleng.2011.02.003

Cea L, Bermúdez M, Puertas J, et al (2016) Rapid flood inundation modelling in a coastal urban area using a surrogate model of the 2D shallow water equations. In: Proceedings of the 4th European Congress of the International Association of Hydroenvironment engineering and Research, IAHR 2016. pp 850–855

Chang L-C, Shen H-Y, Wang Y-F, et al (2010) Clustering-based hybrid inundation model for forecasting flood inundation depths. J Hydrol 385:257–268. doi: 10.1016/j.jhydrol.2010.02.028

de Almeida GAM, Bates P, Ozdemir H (2018) Modelling urban floods at submetre resolution: challenges or opportunities for flood risk management? J Flood Risk Manag 11:S855–S865. doi: 10.1111/jfr3.12276

de Moel H, Jongman B, Kreibich H, et al (2015) Flood risk assessments at different spatial scales. Mitig Adapt Strateg Glob Chang 20:865–890. doi: 10.1007/s11027-015-9654-z

Efstratiadis A, Dialynas YG, Kozanis S, Koutsoyiannis D (2014) A multivariate stochastic model for the generation of synthetic time series at multiple time scales reproducing long-term persistence. Environ Model Softw 62:139–152. doi: 10.1016/J.ENVSOFT.2014.08.017

Falter D, Dung NV, Vorogushyn S, et al (2016) Continuous, large-scale simulation model for flood risk assessments: proof-of-concept. J Flood Risk Manag 9:3–21. doi: 10.1111/jfr3.12105

Falter D, Schröter K, Dung NV, et al (2015) Spatially coherent flood risk assessment based on long-term continuous simulation with a coupled model chain. J Hydrol 524:182–193. doi: 10.1016/J.JHYDROL.2015.02.021

Fill HD, Steiner AA (2003) Estimating Instantaneous Peak Flow from Mean Daily Flow Data. J Hydrol Eng 8:365–369. doi: 10.1061/(ASCE)1084-0699(2003)8:6(365)

García-Feal O, González-Cao J, Gómez-Gesteira M, et al (2018) An Accelerated Tool for Flood Modelling Based on Iber. Water 10:1459. doi: 10.3390/w10101459

Hawkes PJ. SC (2006) Use of Joint Probability Methods in Flood Management: A guide to best practice. T02-06-17

Hawkes PJ (2008) Joint probability analysis for estimation of extremes. J Hydraul Res 46:246–256. doi: 10.1080/00221686.2008.9521958

Jhong B-C, Wang J-H, Lin G-F (2017) An integrated two-stage support vector machine approach to forecast inundation maps during typhoons. J Hydrol 547:236–252. doi: 10.1016/j.jhydrol.2017.01.057

Kasiviswanathan KS, Sudheer KP (2013) Quantification of the predictive uncertainty of artificial neural network based river flow forecast models. Stoch Environ Res Risk Assess 27:137–146. doi: 10.1007/s00477-012-0600-2

Kennard RW, Stone LA (1969) Computer Aided Design of Experiments. Technometrics 11:137. doi: 10.2307/1266770

Keylock CJ (2012) A resampling method for generating synthetic hydrological time series with preservation of cross-correlative structure and higher-order properties. Water Resour Res 48:. doi: 10.1029/2012WR011923

Leonard M, Westra S, Phatak A, et al (2014) A compound event framework for understanding extreme impacts. Wiley Interdiscip Rev Clim Chang 5:113–128. doi: 10.1002/wcc.252

Lin G-F, Lin H-Y, Chou Y-C (2013) Development of a real-time regional-inundation forecasting model for the inundation warning system. J Hydroinformatics 15:1391–1407. doi: 10.2166/hydro.2013.202

Liu Q, Qin Y, Li G, et al (2018) Fast Simulation of Large-Scale Floods Based on GPU Parallel Computing. Water 10:589. doi: 10.3390/w10050589

Liu Y, Pender G (2015) A flood inundation modelling using v-support vector machine regression model. Eng Appl Artif Intell 46:223–231. doi: 10.1016/j.engappai.2015.09.014

MARM (2011) Guía Metodológica para el Desarrollo del Sistema Nacional de Cartografía de Zonas Inundables. Ministerio de Medio Ambiente y Medio Rural y Marino, Centro de Publicaciones

Overpeck JT, Meehl GA, Bony S, Easterling DR (2011) Climate data challenges in the 21st century. Science 331:700–2. doi: 10.1126/science.1197869

Pérez Gómez B, Begoña (2014) Design and implementation of an operational sea level monitoring and forecasting system for the Spanish coast. University of Cantabria

Peterson TC, Manton MJ (2008) MONITORING CHANGES IN CLIMATE EXTREMES: A Tale of International Collaboration. Bull. Am. Meteorol. Soc. 89:1266–1271

Petroliagkis TI, Voukouvalas E, Disperati J, Bidlot J (2016) Joint probabilities of storm surge, significant wave height and river discharge components of coastal flooding events utilising statistical dependence methodologies & techniques. European Commission. Joint Research Centre. Publications Office of the European Union

Razavi S, Tolson BA, Burn DH (2012) Review of surrogate modeling in water resources. Water Resour Res 48:W07401. doi: 10.1029/2011WR011527

Sadegh M, Moftakhari H, Gupta H V., et al (2018) Multihazard Scenarios for Analysis of Compound Extreme Events. Geophys Res Lett 45:5470–5480. doi: 10.1029/2018GL077317

Schumann GJ-P, Stampoulis D, Smith AM, et al (2016) Rethinking flood hazard at the global scale. Geophys Res Lett 43:10,249-10,256. doi: 10.1002/2016GL070260

Seneviratne SI, Nicholls N, Easterling D, et al (2012) Changes in climate extremes and their impacts on the natural physical environment. In: Intergovernmental Panel on Climate Change Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. Cambridge University Press, Cambridge, UK, New York, NY, USA

Serafin KA, Ruggiero P, Parker KA, Hill DF (2019) What’s streamflow got to do with it? A probabilistic simulation of the competing oceanographic and fluvial processes driving extreme along-river water levels. Nat Hazards Earth Syst Sci Discuss 1–30. doi: 10.5194/nhess-2018-347

Serinaldi F (2015) Dismissing return periods! Stoch Environ Res Risk Assess 29:1179–1189. doi: 10.1007/s00477-014-

0916-1
Sopelana J, Cea L, Ruano S (2018) A continuous simulation approach for the estimation of extreme flood inundation in coastal river reaches affected by meso- and macrotides. Nat Hazards 93:1337–1358. doi: 10.1007/s11069-018-3360-6
Storlie CB, Swiler LP, Helton JC, Sallaberry CJ (2009) Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. Reliab Eng Syst Saf 94:1735–1763. doi: DOI: 10.1016/j.ress.2009.05.007
Suykens JAK, Van Gestel T, De Brabanter J, et al (2002) Least Squares Support Vector Machines. World Scientific, Singapore
Svensson C, Jones DA (2002) Dependence between extreme sea surge, river flow and precipitation in eastern Britain. Int J Climatol 22:1149–1168. doi: 10.1002/joc.794
Taguas EV, Ayuso JL, Pena A, et al (2008) Testing the relationship between instantaneous peak flow and mean daily flow in a Mediterranean Area Southeast Spain. CATENA 75:129–137. doi: 10.1016/J.CATENA.2008.04.015
Vacondio R, Dal Palù A, Mignosa P (2014) GPU-enhanced Finite Volume Shallow Water solver for fast flood simulations. Environ Model Softw 57:60–75. doi: 10.1016/J.ENVSOFT.2014.02.003
Van Den Hurk B, Van Meijgaard E, De Valk P, et al (2015) Analysis of a compounding surge and precipitation event in the Netherlands. Environ Res Lett 10:. doi: 10.1088/1748-9326/10/3/035001
Vapnik VN (1998) Statistical learning theory. Wiley
Wahl T, Jain S, Bender J, et al (2015) Increasing risk of compound flooding from storm surge and rainfall for major US cities. Nat Clim Chang 5:1093–1097. doi: 10.1038/nclimate2736
Xu H, Xu K, Lian J, Ma C (2019) Compound effects of rainfall and storm tides on coastal flooding risk. Stoch Environ Res Risk Assess 1–13. doi: 10.1007/s00477-019-01695-x
Xu Y, Huang G, Fan Y (2017) Multivariate flood risk analysis for Wei River. Stoch Environ Res Risk Assess 31:225–242. doi: 10.1007/s00477-015-1196-0
Yaseen ZM, El-shafie A, Jaafar O, et al (2015) Artificial intelligence based models for stream-flow forecasting: 2000–2015. J Hydrol 530:829–844
Zscheischler J, Westra S, van den Hurk BJJM, et al (2018) Future climate risk from compound events. Nat Clim Chang 8:469–477. doi: 10.1038/s41558-018-0156-3