

Accepted Manuscript

Non-stationary rainfall and natural flows modeling at the watershed scale

M. Egüen, C. Aguilar, S. Solari, M.A. Losada

PII: S0022-1694(16)30254-2

DOI: <http://dx.doi.org/10.1016/j.jhydrol.2016.04.061>

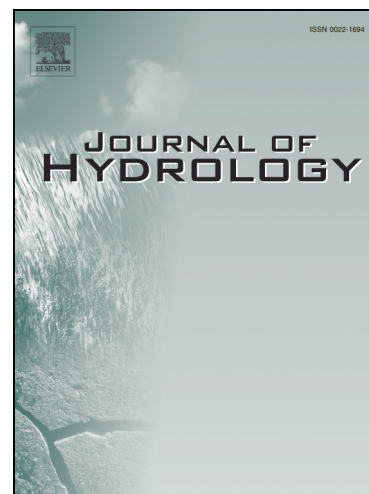
Reference: HYDROL 21231

To appear in: *Journal of Hydrology*

Received Date: 6 July 2015

Revised Date: 27 April 2016

Accepted Date: 28 April 2016



Please cite this article as: Egüen, M., Aguilar, C., Solari, S., Losada, M.A., Non-stationary rainfall and natural flows modeling at the watershed scale, *Journal of Hydrology* (2016), doi: <http://dx.doi.org/10.1016/j.jhydrol.2016.04.061>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Non-stationary rainfall and natural flows modeling at the watershed scale

M. Egüen¹, C. Aguilar¹, S. Solari² and M.A. Losada³

[1]{University of Granada, Fluvial Dynamics and Hydrology - Andalusian Institute of Earth System Research, Granada, Spain}

[2]{Universidad de la República, Instituto de Mecánica de los Fluidos e Ingeniería Ambiental, J. Herrera y Reissig 565, 11300 Montevideo, Uruguay}

[3]{University of Granada, Environmental Fluid Dynamics - Andalusian Institute of Earth System Research, Granada, Spain}

Correspondence to: M. Egüen (meguen@ugr.es)

Abstract

In areas in which natural water resources are variable over time, tools that determine the probability distribution of hydrological variables are required to evaluate various management alternatives. In this article, a stochastic simulation framework of hydrological variables through atmospheric pressure modeling is proposed. This methodology employs the mean value of the atmospheric pressure in the winter to differentiate the wet, medium and dry years in terms of rainfall and flow at different temporal scales. Monthly mean and daily maximum rainfall and flow data series are stochastically replicated. To achieve this replication, a non-stationary parametric mixture distribution model that combines a Weibull and a Normal

distribution is fitted to the univariate distribution of the atmospheric pressure. This model includes interannual variability through two covariables: extraterrestrial solar radiation and the NAO index. This model is applied to the Guadalete River Basin in southern Spain, in which the river flow regime is influenced by the highly seasonal precipitation regime typically found in the Mediterranean area. The non-stationary parametric mixture distribution model with the two covariables showed a good fit to the observed sea level pressure, displaying an important reduction on the BIC. A good correlation was obtained between the average sea level pressure in winter and the accumulated precipitation and flow ($r = -0.8$ for monthly values and -0.6 for maximum daily values). The statistical similarity indicated that the synthetic series of precipitation and flow preserved the distribution trends in the observed data. The identical methodology can be applied in other watersheds once the direct relationship between the mean atmospheric pressure and the hydrology of the area is known.

Keywords: non-stationary model; atmospheric pressure; precipitation; river flow; Monte Carlo simulation; Iberian Peninsula

1 Introduction

Ascertaining the available water resources in a certain study site is required by different stakeholders to accomplish their goals. For example, the proper allocation of water resources to different uses by water planners requires tools for the computation of monthly flows. Reliable estimations of extreme rainfall and flows are needed to define design criteria of infrastructures for flood management purposes (Vasiliades et al., 2015). Environmental flows must considerer monthly flows as well as maximum and minimum extreme flows that could create a potential danger for the ecosystem downstream. Therefore, a variety of hydrological

models have been developed in the last century with different levels of input data requirements.

Climate is the main force acting on runoff flow at the watershed scale. Thus, the characterization of the dynamics of meteorological variables is essential to properly forecast the hydrological response of the system in the short and medium term. The temporal variability of meteorological variables depends on the latitude of the study site.

In semiarid areas, the well-known substantial variations in precipitation at the monthly, seasonal and annual scales determine the temporal variability of water flows and the typical occurrence of dry and wet years. Monthly and seasonal changes in meteorological variables are associated with the high frequency variations in the solar radiation received at the surface of the Earth because of the translation movement around the Sun throughout the year. By contrast, the inter-annual variability is mostly related to climatic factors with a low frequency such as the periodic fluctuations in the incoming extraterrestrial radiation through the top of the atmosphere because of the variations in the number of sunspots (Roy and Haigh, 2011) and other no-periodic factors that change the solar radiation budget at the land surface (e.g. volcanic eruptions, fires, land-uses changes...). Over the last few decades, multiple studies have demonstrated teleconnection patterns that affect the climate in different regions of the planet, such as the El Niño Southern Oscillation (ENSO), the North Pacific Oscillation (NPO), the North Atlantic Oscillation (NAO) and the Western Mediterranean Oscillation (WeMo). The majority of these climatic indexes are based on spatial gradients of atmospheric pressure. The NAO is a large scale seesaw in atmospheric masses between the subtropical high and the polar low. The NAO is also the dominant mode of winter climate variability in the North Atlantic region, although this process also affects all seasons (Hurrell, 1995; Hurrell et al., 2003).

Numerous authors have analyzed the relationships between the climatic indexes and hydrological variables in different areas throughout the world. For example, Kiem and Franks (2001) studied the relationship between the ENSO and rainfall and runoff in a catchment in Australia; Wang et al., (2006) elucidated connections between ENSO events and decreasing water discharges in the Yellow River in China; Ward et al., (2014) evaluated the statistical relationships between the ENSO and the peak daily discharge in a given year globally; Cullen et al. (2002) explored the relationship between the NAO and streamflow of major rivers in the Middle East; and López-Moreno et al. (2007) analyzed the effects of the extreme phases of the winter NAO on the water resources in the Tagus River basin in Spain. In Southern Europe, certain relationship exists between the NAO and WeMO and the rainfall and flow regimes, notably in the winter season (López-Moreno et al., 2007; Lorenzo-Lacruz et al., 2011; Martin-Vide and Lopez-Bustins, 2006; Trigo et al., 2004). The winter NAO index is widely recognized as the most important pattern for explaining the climatic and hydrological variability in Western Europe. In Southwestern Europe, drier years and lower water flows are reported when the winter NAO is in a positive phase, and wetter years and higher flows are reported during negative phases. In contrast, the opposite response has been found in the Northwestern Europe area (López-Moreno and Vicente-Serrano, 2008; López-Moreno et al., 2007; Pires and Perdigão, 2007; Salgueiro et al., 2013; Trigo et al., 2004).

Although the connections of precipitation with large-scale atmospheric pressure is well known, most studies, in the case of Europe, establish relationships between the NAO index and the precipitation and not between the precipitation and atmospheric pressure directly (Lavers et al., 2013). The non-stationary characteristics of atmospheric pressure at different temporal scales complicates the study, notably when seasonal assessments are performed and may explain the scarcity of available studies in this respect. In previous studies of geophysical

variables (Solari and Losada, 2011), mixed statistical distribution functions can be applied to reproduce the seasonal variability when the statistical distribution function of a non-stationary variable does not follow a simple theoretical distribution. In a previous study, Egüen (2014) noted a seasonal pattern in the sea level pressure at different locations in the Iberian Peninsula; the 28-day moving average values were higher than the annual mean value in the wet season and lower than the annual mean value in the dry season. The interannual variability pattern that determines the occurrence of wet/dry periods should then be added to the seasonal pattern. In the development of non-stationary statistical models of geophysical variables, covariables are commonly used with time-dependent parameters to reflect the interannual variability of the variable (e.g., Cannon, 2010; Izaguirre et al., 2010; Maraun et al., 2010; Solari and Losada, 2011; Trambly et al., 2012; Vasiliades et al., 2015).

In line with the studies mentioned above that link the winter NAO with the hydrological behavior in Southern Europe (López-Moreno et al., 2007; Trigo et al., 2004), the direct effect of the winter mean atmospheric pressure on the wet/dry character of the year can constitute a tool for water resources management and planning. Once reliable relationships between the mean atmospheric pressure and hydrological variables are found, the former can be used to forecast rainfall and flow-related variables.

Stochastic models are often applied in hydrology to generate different samples of meteorological and hydrological data equally likely with respect to the observed series. The simulation is particularly useful in the context of water resources management because the process yields multiple equally likely series that can be used to evaluate different alternatives in water resources management (Kottegoda, 1980; Kottegoda et al., 2014; Montanari et al., 1997, 1999). Once a statistical model for the generation of the atmospheric pressure is available in a certain study site, a stochastic simulation can be used to generate an equally

likely data series of atmospheric pressure that constitutes the forcing to the stochastic generation of hydrological variables.

This article proposes a methodological framework for the stochastic generation of monthly and maximum daily winter rainfall and flow related variables through atmospheric pressure modeling of commonly available data. This aim involves fitting a non-stationary parametric mixture distribution to model the univariate distribution of the atmospheric mean sea level pressure (including climatic covariables) to reflect the interannual variability of the atmospheric pressure and fitting an autoregressive model to reproduce the short-term autocorrelation of the pressure series. Therefore, the NAO index and extraterrestrial radiation are considered as the main covariables capturing the interannual differences in the historical records of the atmospheric pressure. Then, the mean winter atmospheric pressure is used to classify the simulated years into wet, mean and dry categories, and a simulation of the rainfall and the flow-related variables is performed conditional to this classification. This process is applied to the Guadalete river basin in southern Spain, in which the river flow is influenced by the highly seasonal precipitation regime typically found in the Mediterranean area.

The paper is organized as follows. A description of the study area and the dataset used is shown in Section 2. The methodology applied is described in Section 3. The main results are presented and discussed in Section 4. Finally, the main conclusions are drawn in Section 5. Additionally, supplementary materials is presented in Appendix A. An overview of the monthly adjustment of sea level pressure to theoretical distributions is shown in Appendix B along with a list of abbreviations and symbols used throughout the paper (Appendix C).

2 Study Area and Available Data Description

The Guadalete River Basin, with an extent of 3677 km², is located in the southwest of the Iberian Peninsula (Fig. 1) and is surrounded by the Grazalema Mountains in the east and the Gulf of Cadiz in the west. The main river channel originates in the Grazalema Mountains at 1100 m.a.s.l and runs into the Atlantic Ocean after 157 km. The mean flow in the main channel is approximately 10 m³s⁻¹ and reaches 100 m³s⁻¹ in extreme events.

The basin displays a typical Mediterranean climate with warm-temperate conditions and fluctuations in the rainfall regime influenced by changes in the NAO index (Wolf et al., 2014). The mean temperature of the basin is 18 °C and 9.5 °C in mountainous areas. The mean annual precipitation is 645 mm, and varies between 345 and 1150 mm as dry and wet years occurs. The majority of the precipitation occurs mostly between autumn and spring when moist air is transported by the westerly winds driven by the NAO (Wolf et al., 2014). The spatial distribution of the annual precipitation in the basin is highly heterogeneous because of its topographic gradients and geographic location upwind of the Betic Range. Therefore, over 2000 mm of precipitation occurs in the Grazalema Mountains, which constitutes the location with the highest amount of rainfall recorded in Spain. As a result of torrential events, discharge peaks often reach high values with serious consequences such as flooding in the estuary area and surrounding areas upstream, and compromising the biological status of the water if the maximum threshold in the environmental flow regime is exceeded, which poses a potential danger to the ecosystem. In addition, important episodes of turbidity may occur (Contreras and Polo, 2012).

Six-hour averaged sea-level pressure (*6hr-SLP*) data are used. These data are provided by the ERA-40 (September 1957-August 2002) and ERA-Interim (January 1979-current) reanalysis datasets produced by the European Centre for Medium-Range Weather Forecast (ECMWF)

(Dee et al., 2011). These are gridded datasets (0.75° spatial resolution) for the entire globe. Because ERA-Interim is an improvement of ERA-40, *6hr-SLP* data for the period of September 1957-December 1978 from ERA-40 and for the period of January 1979-August 2014 from ERA-Interim were used in this study. A grid point located in the ocean in the Gulf of Cadiz (36.75°N , 6.75°W) was selected (Fig. 1a).

Monthly North Atlantic Oscillation climatic index (NAO) and monthly measurements of the integrated emission from the solar disc at 2800 MHz (10.7 cm wavelength), known as the daily solar flux (SF), from the Climate Prediction Center (CPC) and from the National Geophysical Data Center (NGDC) of the National Oceanic and Atmospheric Administration (NOAA) were used in this application over the study period (September 1957-August 2014).

Daily rainfall datasets recorded at 58 weather stations distributed throughout the basin were used to generate the spatially averaged daily rainfall over the study period (September 1957-August 2014). Data are available from the Spanish Meteorological Agency (AEMET) and the Agroclimatic Information Network (RIA) from the CLIMA subsystem of the Andalusian Government. A careful review of the original data series was performed to check and validate the quality to generate a robust dataset covering the longest time period possible. Distributed daily rainfall at the basin (with a grid 90×90 m) was computed with WiMMed model (Egüen et al., 2009; Polo et al., 2009) through specific algorithms for the spatial interpolation of meteorological variables (e.g., corrections with height) (Aguilar et al., 2014; Egüen et al., 2012; Herrero, 2007; Herrero et al., 2014). Finally, the basin-averaged daily rainfall series was computed for the reliable period.

The river flow regime is influenced by the highly seasonal precipitation regime typically found in the Mediterranean area. Winter and spring flows account for the majority of the fluvial regime (85%), followed by a relatively long and dry period. Natural daily flows were

restituted at the outlet of the main river channel to the estuary (control point in Fig. 1b) during September 1960-August 2014 following a similar methodology to that applied by Jiongxin, (2005) or Wang et al., (2006). Daily inflows into and outflows from the dams along the basin from the Automatic Water Information Network (SAIH) were used to obtain the natural flow ($Q_{natural}$) at the target point:

$$Q_{natural} = Q_{measured} + Q_{stored} \quad (1)$$

where $Q_{measured}$ is the measured daily flow at the study point, and Q_{stored} is the amount of water stored (+)/released (-) daily from the reservoirs upstream of the control point. In this case, the sum of the natural flow calculated following this procedure for the two dams located more downstream in the two principal streams, the Guadalete and Majaceite Rivers (Fig. 1), is considered as the natural flow to the estuary (Egüen, 2014).

Table A.1 summarizes the data sets used and their main characteristics.

3 Methods

First, the non-stationary model for the atmospheric pressure is described. The details of how both seasonal and interannual variability are incorporated into the model are provided. Then, the relationship between the mean atmospheric pressure and the rainfall and flows that constitutes the basis for the stochastic generation is provided. Finally, the different steps to generate new data series are described.

3.1 6hr-SLP time series simulation

3.1.1 Non-stationary distribution

A first approximation of the statistical behavior of the 6hr-SLP revealed that this variable mostly followed either a Normal or a Weibull distribution function depending on the month of the year (Appendix B) (Egüen, 2014). Thus, the following parametric mixture model with a Weibull-Normal probability density function (f_{W-N}) is proposed (Eq. 2.):

$$f_{W-N}(x, t) = \alpha_w(t) f_w(x, t) + (1 - \alpha_w(t)) f_N(x, t) \quad (2)$$

where x equals the 6hr-SLP atmospheric pressure, f_w is the Weibull probability density function, f_N is the Normal probability density function and $\alpha_w(t)$ and $(1 - \alpha_w(t))$ provides the weight of Weibull and Normal distribution with time, respectively. The model is thus composed by five time variable parameters: the weight of Weibull distribution, $\alpha_w(t)$; the two parameters of the Weibull distribution, $a(t)$ and $b(t)$; and the two parameters of the Normal distribution, $\mu(t)$ and $\sigma(t)$. The parameters are estimated by fitting the f_{W-N} distribution to the historical records as follows. First, the seasonal variation in the parameters is considered. Then, the inclusion of the interannual variability completes the non-stationary model of the atmospheric pressure in the study site.

3.1.2 Seasonal variability

Fourier series are used to represent the annual pattern of each of the five parameters of the mixed distribution function: $\alpha_w(t)$, $a(t)$, $b(t)$, $\mu(t)$ and $\sigma(t)$. For a generic parameter, $\theta(t)$, the following expression is adopted:

$$\theta(t) = \theta_{a0} + \sum_{k=1}^{N_k} \theta_{ak} \cos(2\pi kt) + \theta_{bk} \sin(2\pi kt) \quad (3)$$

where $\theta(t)$ is the parameter (α_w , a , b , μ or σ), N_k is the order of the series, k varies between 1 and N_k , θ_{a0} is the stationary term, θ_{ak} and θ_{bk} are the coefficients of the non-stationary terms and t is time measured on a yearly scale, varying between 0 (January 1st 1957 00:00 hrs) and 58 (December 31st 2014 24:00 hrs).

The coefficients θ_{a0} , θ_{ak} and θ_{bk} are estimated for every parameter $\alpha_w(t)$, $a(t)$, $b(t)$, $\mu(t)$ and $\sigma(t)$ using maximum likelihood. The parameters were estimated by progressively increasing the order of approximation of the Fourier series. To evaluate the significance of the improvement in the fit when the order is increased, the Bayesian Information Criterion (BIC) (Eq. 4) was computed (Schwarz, 1978):

$$BIC = -2LLF + \log N \ D \quad (4)$$

where LLF is the log likelihood, N is the number of available observations and D is the number of coefficients in the model, i.e., the total number of coefficients θ_{jk} , with $j = \{a, b\}$ and k varying from 0 to N_k when $j = a$ and from 1 to N_k when $j = b$, i.e., $D = 5(2N_k + 1)$. The maximum order of the approximation analyzed in this study is $N_k = 4$, corresponding to a minimum cycle of three months. The model that minimizes the BIC is judged to be the simplest among the candidates for the studied samples.

Once the order of the series is determined and the parameters are estimated with *6hr-SLP* reanalysis data for the entire study period (September 1957-August 2014, a total of 57 hydrological years), the goodness of fit of the seasonal model is evaluated by applying the Kolmogorov-Smirnov test at a 0.01 significance level.

3.1.3 Interannual variability

Following previous studies, interannual variations are incorporated in the parameters of the distribution function using a series of covariables $C_i(t)$ to reflect the non-cyclic behavior of the *6hr-SLP* series (Coles, 2001; Izaguirre et al., 2010; Solari and Losada, 2011):

$$\theta(t) = \theta_{a0} + \sum_{k=1}^{N_k} \theta_{ak} \cos(2\pi kt) + \theta_{bk} \sin(2\pi kt) + \sum_{i=1}^{N_i} f_i(C_i(t), t) \quad (5)$$

where N_i is the number of covariables to be included and f_i is the function that defines how each covariable affects the parameter $\theta(t)$.

Because the climate depends on a large-scale of low-frequency patterns explained by teleconnection indices, the interannual variability can be included with covariables when the teleconnection indices are incorporated into the theoretical probability distributions with time-dependent parameters (Vasiliades et al., 2015). In this study, the monthly NAO index and the monthly SF are included as covariables. Three different interannual models are tested by considering both covariables individually ($i=1$ in Eq. 5), and finally by considering both covariables together ($i=2$ in Eq. 5). In all cases, the effect of the covariables on the parameters is assumed to be linear, i.e., $f_i = a_i C_i(t)$, where a_i is estimated by a maximum likelihood with θ_{jk} .

Among the three different adjustments applied, the best non-stationary *Weibull-Normal (W-N)* model is determined using log-likelihood and the Bayesian Information Criterion.

3.1.4 Short-term autocorrelation

To model short-term autocorrelation of the *6hr-SLP* time series an autorregressive (AR) model of order p is used. An AR(p) model is given by

$$Z_t = \phi_1 Z_{t-1} + \dots + \phi_p Z_{t-p} + \varepsilon_t \quad (6)$$

where ϕ_i are the autoregressive coefficients, ε_t are independent, identically distributed realizations of a normal distribution with zero mean and σ_ε^2 variance, and Z_t are the normalized *6hr-SLP*, obtained as

$$Z_t = \Phi^{-1} F_{W-N}^{-1}(x|\theta, t) \quad (7)$$

where Φ is the standard normal distribution.

Autoregressive coefficients ϕ_i and variance σ_ε^2 are obtained through least squares (see e.g. Lütkepohl, 2005).

3.2 Relationship between the atmospheric pressure and selected hydrologic variables

Following the approach by Trigo et al., (2004) in their relationship between the winter NAO index and monthly river flows in the Iberian Peninsula, the correlation between the mean winter sea level atmospheric pressure and the selected hydrological variables was analyzed: monthly rainfall (P_m , $m=1:12$, calendar months), monthly flow (Q_m), maximum daily winter rainfall (P_{maxd}) and maximum daily winter flow (Q_{maxd}). The analyses are computed for the hydrological year (September-August for the latitude of the study site) to avoid splitting winter rainfall in two different years, as suggested by Trigo et al., (2004) in other Iberian rivers.

First, a global correlation analysis was performed between the observed mean atmospheric pressure in the winter months (December-January-February, DJF), SLP_w , and both the accumulated rainfall in the winter (DJF), P_w , and the water flow with one-month delay (January-February-March, JFM), Q_w . The different time-windows (DJF or JFM) applied in each variable is justified by the results obtained by López-Moreno et al., (2007). This previous study found that the influence of extreme NAO winters acts with different time lags

on the different variables: immediately and significantly for December to March in precipitation, but with a lag of at least one month and with a longer duration in water flows. The analysis was performed on the normalized variables (P_{wn} , Q_{wn} , P_{maxdn} , Q_{maxdn}).

Second, three different types of years, wet, medium and dry, are identified in terms of the SLP_w value relative to the mean value for the entire period (September 1957-August 2014), $\overline{SLP_w}$. For that, two thresholds in the SLP_w values should be fixed to distinguish the wet, medium and dry years in the study area. The 40 and 60 percentiles of observed precipitation and flow are chosen to determine the type of year in terms of the hydrological variables. These quantiles have been used in other studies in Mediterranean areas (López-Moreno et al., 2011). This process conditions the type of year and constitutes the forcing agent in the stochastic simulation of the hydrological variables. Thus, the different behaviors of the observed hydro-meteorological variables (P_m , Q_m , P_{maxd} and Q_{maxd}) per type of year (e.g., P_{m_Dry} , P_{m_Med} and P_{m_Wet}) were previously analyzed.

Finally, the statistical distribution of the observed hydro-meteorological variables was evaluated. Because P_m does not follow any theoretical distribution function (results not shown here), empirical cumulative distribution functions (F_e) are computed per month and type of year (wet-medium-dry) (i.e., $F_e(P_{m_Dry})$, $F_e(P_{m_Med})$, $F_e(P_{m_Wet})$). The identical consideration is imposed for Q_m (i.e., $F_e(Q_{m_Dry})$, $F_e(Q_{m_Med})$, $F_e(Q_{m_Wet})$). By contrast, P_{maxd} and Q_{maxd} follow a Gamma distribution. Thus, a Gamma distribution function (F_g) is computed per variable and type of year (wet-medium-dry) (e.g., $F_g(Q_{maxd_Dry})$). The F_e and F_g of each variable and type of year are used in the stochastic simulation process detailed in the next section.

3.3 Stochastic simulation of variables

Once a relationship is found between the SLP_w and the hydrological variables analyzed, the approach followed in this paper is to simulate the $6hr$ - SLP time series and to use the series as a basis for the simulation of the hydro-meteorological variables. All simulations are performed with the identical series of covariables, NAO and SF.

NR synthetic samples of N_y -years were generated ($N_y = 57$, equal to the observed series during September 1957-August 2014). In this case $N_y = 57$ years is used in order to avoid the simulation of new time series of the covariables (NAO and SF). On the other hand, the choice of NR will depend on the probability of occurrence of the particular event to be tested with the generated synthetic series (see e.g. Kottegoda and Rosso 2008; chap. 8). In this work the objective is to compare the simulated and the original series in order to show the performance of the simulation methodology, so $NR = 100$ is chosen. The steps of the stochastic simulation are listed below.

1. A N_y -year long time series of $6hr$ - SLP is simulated using the fitted f_{W-N} distribution and AR model.
2. The mean winter SLP_w is estimated for each of the N_y years previously simulated. Then, the SLP_w is used to classify each of the N_y years as either wet, medium or dry.
3. Depending on the type of year obtained in step 2 (wet, medium or dry), the appropriate F_e and F_g distributions are used to simulate the mean monthly (P_{sm} and Q_{sm}) and maximum daily (P_{s_maxd} and Q_{s_maxd}) winter values of precipitation and flow for each of the N_y years.

Steps 1 to 3 are repeated NR times. The result is a Monte Carlo simulation of NR N_y -year long time series of $6hr$ - SLP series, NR N_y -year long time series of monthly values P_{sm} and Q_{sm} , and NR N_y -year long series of winter maximum daily values P_{s_maxd} and Q_{s_maxd} . The simplicity of the stochastic approach can be appreciated from the step-by-step procedure given in Fig. A.1.

Statistics for the synthetic series were computed to explore the ability of the proposed methodology to reproduce the characteristics of historic data adequately. The resulting ensembles of the estimates for each variable were then analyzed to determine the mean, standard deviation and quantiles. These parameters were then compared with the corresponding observed values. This represents the first validation of the model.

Additionally, an extreme value analysis is applied to the observed data, P_{maxd} and Q_{maxd} , and compared to the identical analysis applied to simulated values obtained from the *NR* simulated series, P_{s_maxd} and Q_{s_maxd} . This analysis provides a second validation of the modeling approach.

4 Results

4.1 Atmospheric pressure model

4.1.1 Seasonal variability

The five parameters of the non-stationary distribution function f_{W-N} were estimated with the *6hr-SLP* reanalysis series for the selected point located in the in the Gulf of Cadiz (36.75°N, 6.75°W). The Fourier series were evaluated (Eq. 3) with different orders of approximation (from 1 to 4). In this case, the minimum BIC value was obtained when using a third order Fourier series to model the seasonal cycle of the parameters of the f_{W-N} distribution, i.e., a Fourier series of order three is used to estimate $\alpha_w(t)$, $a(t)$, $b(t)$, $\mu(t)$ and $\sigma(t)$.

Figure 2 shows the annual temporal evolution (0 for January 1st 00:00 hrs and 1 for December 31st 24:00 hrs) of the five parameters adjusted in the seasonal *W-N* model. The principal component is the annual period. Therefore, the parameters $a(t)$ and $\mu(t)$ show a typical

seasonal behavior, reaching maximum values in the winter months (1025 and 1028 hPa, respectively). This behavior occurs when the Weibull distribution becomes more important than the Normal distribution (according to the high $\alpha_w(t)$ values, exceeding 0.8 in Eq. 2). By contrast, minimum $a(t)$ and $\mu(t)$ values (1015 hPa) are found in the central period of the year when the Normal distribution is the dominant distribution (denoted by the low $\alpha_w(t)$ values of approximately 0.1 in Eq. 2) (see Appendix B). This seasonal behavior is in line with the pattern in the 28-day moving average values of the sea level pressure found in Egüen (2014).

The empirical density functions of both the observed and modeled *SLP* throughout the year are shown in Fig. 3a. In general, the calculated probability density function for the seasonal *W-N* model, f_{W-N} , agrees with the empirical density function, f_e , and also reproduces the seasonal variation stated in the parameters of the mixed model. Therefore, the dispersion of values is higher (between 990 and 1040 hPa) in the winter when the weight of the Weibull distribution is dominant. Additionally, the highest frequencies are found to be skewed in the density function towards higher-pressure values of approximately 1028 hPa (colored light green in Fig. 3a). This season is where the major fluctuations in *6hr-SLP* appear and where minimum pressure values are achieved. The opposite relationship is noted in the central part of the year when the influence of the Normal distribution is stronger, as denoted by the smaller range of values (1006-1023 hPa) and the centered location of the highest relative frequency (0.18) of approximately 1016 hPa (colored yellow in Fig. 3a).

Figure 3b shows the 28-day moving average of observed values, SLP_{28} , versus the expected values obtained with the seasonal *W-N* model and the 90% confidence bounds. The close adjustment of the model to observed data is displayed. The results of the Kolmogorov-Smirnov goodness of fit test based on the maximum difference between the *W-N* model distribution function (F_{W-N}) and the sample distribution function (F_e) (Fig. 4a) was below the

0.01 significance level, validating the model. Additionally, the negligible deviation from the 1:1 line in the P-P plot (a slope of 0.98 (Fig. 4b)) confirmed the validity of the seasonal model.

4.1.2 Interannual variability

The five parameters of the non-stationary distribution function f_{W-N} were anew estimated by incorporating the covariables. In Fig. 5, the monthly mean of the observed data and the expected value of the modeled series are represented for the three interannual models: when considering both covariables individually (Fig. 5a and 5b) and considering both covariables jointly (Fig. 5c). The seasonal model previously generated is also shown for comparison (Fig. 5d).

Certain differences in the estimations of each of the three different non-stationary interannual $W-N$ models can be observed. In years with higher SLP values than the average (e.g., 1983, 1984, 1989, 1995), the correction applied by the incorporation of the SF as a covariable is not sufficient (highlighted by the red boxes in Fig. 5) because the model does not properly capture those maximum values. Otherwise, when the SLP values are near the average (e.g., 1960, 1966, 1977, 1988, 2006), the SF better reflects the interannual variability (highlighted by the blue boxes in Fig. 5) than the simulations obtained by the model with the NAO index as the single covariable (Fig. 5a). These problems are globally solved when both the SF and NAO index are included as covariables (Fig. 5c).

Figure 6 shows log-likelihood (LLF) and Bayesian Information Criterion (BIC) estimated for the four $W-N$ models mentioned above. As expected, the inclusion of the covariables improves the fit of the model (LLF increases); however, an important reduction on the BIC is only achieved when both covariables (NAO and SF) are included in the model. Therefore, this

multiple covariable model is selected and applied for the stochastic replications of the SLP in the study site.

4.2 Statistical dependence of the hydrological variables and atmospheric pressure

4.2.1 Annual dependence

Figure 7 shows the interannual variability of the SLP_{wn} , from $6hr$ - SLP reanalysis datasets at the studied point and each of the four hydrological variables analyzed, P_{wn} (Fig. 7a), Q_{wn} (Fig. 7b), P_{maxdn} (Fig. 7c) and Q_{maxdn} (Fig. 7d), with their correlation coefficients. The subscript n indicates that each variable is normalized with its respective mean and standard deviation. Moreover, the SLP_{wn} is multiplied by -1 to facilitate the visual comparison in Fig. 7.

In the four cases, a relationship was noted, especially in the cases of P_{wn} and Q_{wn} . Therefore, years with low SLP_w (high SLP_{wn} values in Fig. 7), such as 1964, 1998 or 2010, were characterized by high rainfall (SLP_{wn} and P_{wn} values > 1.5 in Fig. 7a) and high flow levels (SLP_{wn} and Q_{wn} values > 1 in Fig. 7b) in the winter. These relationships confirm the expected occurrence of high rainfall-runoff events after low pressure events. Similarly, years with the highest P_{maxd} and Q_{maxd} values (e.g., 1963, 1970 and 2010 in Fig. 7c and 7d) were inversely related to SLP_w . As obtained by Trigo et al. (2004) in other Iberian rivers during the 1980s and 1990s, neither the rainfall nor flows decreased to the low levels expected from the highest SLP_w values (lowest SLP_{wn} values in Fig. 7) during the entire study period. Although positive extreme normalized hydrological values (P_{wn} and $Q_{wn} > 0.5$) seem to present wide oscillations of the identical order of magnitude of extreme positive SLP_{wn} values (up to 3 in Fig. 7a and over 3 in Fig. 7b), in years with extremely high SLP_w values (extremely low SLP_{wn} values in

Fig. 7), the hydrological variables are not reduced to the identical extent and seem to be limited by a constant threshold of approximately -0.9 for Q_{wn} and Q_{maxdn} (in Fig. 7b and 7d).

As expected, a lower correlation was obtained between the winter mean SLP (SLP_{wn}) and the daily maximum values (P_{maxdn} and Q_{maxdn} , with $r = -0.5$ and -0.6 , respectively) than between the SLP_{wn} and the winter accumulated values (P_{wn} and Q_{wn} , with $r = -0.82$ and -0.76 , respectively) because the daily maximum (an instantaneous variable) has higher variability than the monthly or seasonal means (medium term average of an instantaneous variable). Nevertheless, the absolute correlation values above 0.6 confirm the relationship of the hydrological response of the watershed with the seasonal SLP_w .

Two thresholds in the SLP_w values should be fixed to distinguish the wet, medium and dry years in the study area. To do this, different thresholds were tested. Fig. A.2 displays the number of inappropriate years classified according to the different thresholds of SLP_w . In the study site, the thresholds that provide the best classification correspond to 1 hPa above the mean SLP_w for the entire period ($\overline{SLP_w}$) for dry years and 1.5 hPa below the $\overline{SLP_w}$ for wet years. Years with SLP_{wn} within $\overline{SLP_w} + 1$ hPa and $\overline{SLP_w} - 1.5$ hPa are classified as medium years. For the study period and the thresholds applied, the numbers of dry, medium and wet years were 21, 19 and 17, respectively. The results of the classification with the normalized thresholds are also shown in Fig. 7 in which the inappropriate classification of a year in terms of the normalized hydrological values is represented with an open circle (e.g., 1966, 1986, 2001, 2005, 2006 and 2013).

The total number of inappropriate years was below 20% of the number of years in the period, with higher errors values in wet years (e.g., 2001 or 2013).

This relationship supports the use of the SLP_w to forecast the type of hydrological year in the study site and allows the methodology to be applicable in other semi-arid coastal basins.

4.2.2 Monthly dependence

Figure 8 represents the mean annual cycle of the monthly rainfall and flows calculated for the dry, medium and wet years classified using the SLP_w . Significant differences can be appreciated in the monthly rainfall from December to February (Fig. 8a) and between December and March (Fig. 8b) in the monthly flows. A unimodal monthly distribution of the monthly values is observed during the wet years with a peak in December for the rainfall (200 mm/month in Fig. 8a) and in January for the flow (180 hm^3/month in Fig. 8b). This distribution transforms into a bi-modal distribution with maximum values in the autumn (November) and spring (April for the rainfall and March for the flow) for both variables during dry years (Fig. 8a and 8b). This is given by the fact that the condition of wet or dry year in this study is determined by the precipitation that occurs during the winter months, as explained in section 3.2. Therefore the order of magnitude of monthly rainfall in the winter in wet years which are significantly higher than the values along rest of the months shifts the bimodal distribution into an unimodal distribution. Additionally, the pattern in autumn is less regular than the one in spring, as noted in other studies in the Mediterranean area (e.g., Montanari et al., 1997). This pattern can be appreciated through the higher standard deviations in the autumn months than in the spring months (e.g., in dry years, $\sigma = 67.8$ mm and 75.4 mm for P in October and November, respectively, versus $\sigma = 44.5$ mm and 33.9 mm for P in April and May, respectively) (Table 1).

Regarding the maximum winter rainfall and flows (Table 2), increasing statistical values are generally found from the driest to the wettest conditions (e.g., 119.8 to 465.5 m^3/s in the mean values in Q_{maxd}). This relationship is also reflected by the Gamma distribution fitted to the

observed values per type of year (Fig. 9) in which increasing values for higher return periods are found from the dry to wet years.

4.3 Stochastic simulation

One-hundred 57-year long time series are simulated following the methodology described in section 3.3. Covariables (NAO and SF time series) are identical in all simulations over the measured series for the period of September 1957-August 2014. Figure 10 shows the temporal evolution of the observed and simulated series of the variables mean monthly SLP , mean monthly precipitation (P_m) and mean monthly flow (Q_m). Important characteristics, such as runs of periods with low/high SLP values and thus high/low rainfall values and extreme events/low flows, are reproduced. Therefore, years 1969, 1996 and 2010, displaying low SLP values, were preserved in the simulated series (up to 1020 hPa in Fig. 10a). Similarly, in these years, both the observed and simulated P_m and Q_m variables displayed high values (up to 400 mm/month and 380 hm^3 /month, respectively, in Fig. 10b and 10c). By contrast, years 1989, 1992-1994, 2002, displaying high SLP values (up to 1028 hPa in Fig. 10a) in both the observed and simulated series, presented low values for the observed and simulated P_m and Q_m (up to 150 mm/month and 100 hm^3 /month, respectively, in Fig. 10b and 10c).

From Fig. 11, the cumulative distribution functions of the mean monthly variables are well reproduced using this methodology. The basic statistics of both the observed and simulated series appear in Table 3, in which the statistics of the simulated series represent the average of the 100 replicates generated for each variable. The stochastic model adequately preserves the mean (e.g., $P_m = 90.9$ mm/month, $P_{sm} = 84.2$ mm/month in February), standard deviation (e.g., $Q_m = 25.4$ hm^3 /month, $Q_{sm} = 21.3$ hm^3 /month in April) and quantiles (e.g., $P_m = 179.8$ mm/month, $P_{sm} = 179.2$ mm/month for the 90th percentile in February) of the historic monthly rainfall and flow series. The Kolmogorov-Smirnov goodness of fit test between the observed

and simulated monthly variables passed the 0.01 significance cutoff in 87% and 94.5% of the cases for the P_m and Q_m , respectively.

As in the case of monthly variables, the cumulative distribution functions of the simulated maximum daily winter variables agree with the distribution of the observed data (Fig. 12). A comparison of the basic statistics of both the observed and simulated series is presented in Table 4. The quantiles, mean and standard deviation of the simulated series are of the identical order as the observed series, notably for precipitation (e.g., 34.3 mm/day for the mean P_{s_maxd} and P_{maxd}). The P_{maxd} values are within the 95% confidence limits of P_{s_maxd} . Nevertheless, the Q_{maxd} values are below the 5% quantile of Q_{s_maxd} .

Figure 13 shows a Gamma distribution fitted to the observed extreme values of rainfall (P_{maxd}) (Fig. 13a) and flow (Q_{maxd}) (Fig. 13b) plotted against the return periods, Tr . The simulated series are also shown. The simulated precipitation values, P_{s_maxd} , are equidistantly distributed above and below the Gamma fit of the observed values. The simulated Q_{maxd} values are more skewed toward high values when compared to the Gamma fit of the observed values. Table 5 shows the parameters for the Gamma distribution function fitted to the observed and simulated series of the maximum daily winter precipitation per type of year (dry, medium and wet).

5 Conclusions

The proposed methodology enables the simulation of the stochastic series of monthly and maximum daily rainfall and flow at a basin scale using a regional geophysical variable: the sea level atmospheric pressure. The availability of the pressure from a complete and globally accessible database could therefore be used to predict the hydrological information in basins

in which the hydrological response is clearly related to the dynamics of the sea level atmospheric pressure.

A non-stationary mixed model consisting of a Weibull and a Normal distribution was described to estimate the temporal distribution of the atmospheric sea level pressure (SLP). The mean annual cycle of the distribution is modeled by a Fourier series approximation of the parameters for the distributions. Interannual variations are introduced using covariables: NAO and the less usually used, SF. The increase on the log-likelihood function (LLF) and the important reduction on Bayesian Information Criterion (BIC) confirmed the validity of the proposed multiple covariable model when both covariables are introduced. Therefore, this model is selected for the stochastic simulation of the SLP in the study area.

A high linear correlation was found between the mean winter sea level pressures (SLP_w) and the monthly means of several hydrological variables. Thus, years with low SLP_w , correspond to years with high records of accumulated rainfall and flow in the winter (P_w and Q_w , with $r = -0.8$ for both hydrological variables). Similar trends were found for the correlation between the SLP_w and the winter maximum of daily variables (P_{maxd} and Q_{maxd}).

The identified SLP_w thresholds classify the dataset into dry, medium or wet years with an error of less than 20%. This classification established the base of the stochastic simulation of the precipitation and flow. Thus, from the generation of the synthetic series of SLP by the non-stationary model, the stochastic series of precipitation and flow can be derived, given the type of year.

Simpler models than the non-stationary mixed model used here for the SLP could had been used to simulate the series of dry/medium/wet years directly. However, this pressure model generates a wealth of additional information (namely: six-hourly time series of SLP) which, although not used here, may be used in other studies; e.g. as an input for a synthetic weather

generator model that requires the full time series of SLP. Therefore, other hydrological variables commonly used in water resources planning could be derived (e.g. number of rainfall-runoff events, duration, maximum values, total volume, runoff coefficients, etc.). The authors are currently working in this line of research. Empirical distribution functions of the simulated P_{sm} and Q_{sm} series preserved the trends of the distribution of the observed series as shown by the Kolmogorov-Smirnov test and the proximity of statistical variables.

Similarly, the P_{s_maxd} and Q_{s_maxd} series obtained from the Gamma distribution functions properly reproduced the statistics of the observed series (e.g., 34.3 mm/day for both the observed and simulated mean daily maximum winter rainfall).

In this study the explanatory variable is the winter sea level pressure, SLP_w . However, in another study site, the mean SLP at the proper time scale that affects the hydro-meteorology must be computed. By generating a series of SLP from the adjusted parameters at each step of the simulation sequence, an unlimited set of trend curves as possible climatic scenarios can be obtained. It must be emphasized the easy availability of input data requirements because freely available data from the ERA reanalysis datasets produced by the ECMWF that can be used even for rainfall and flow data in the most limiting situations (Balsamo et al., 2015). This methodology provides an efficient synthetic weather generator focused on monthly mean and annual extreme variables, that allows for the generation of several realizations of the stochastic processes, once the relationship between the mean atmospheric pressure and the hydrology of the area has been properly modeled. These new realizations of the process can be used to analyze how a given system would behave under different feasible conditions (e.g. Gómez-Beas et al., 2012; Polo et al., 2014a, 2014b; Wilks, 2009).

This methodology is limited when subsurface flow plays a major role in a watershed as the hydrological response is not clearly driven by changes in SLP. Also in very regulated

watersheds where anthropogenic actions (dams, irrigation, etc.) greatly modify the natural flow.

Moreover, the inclusion of the SF as a covariable indirectly involved in the variability of hydrological variables is an uncommon practice that in this work has proven to be beneficial.

Acknowledgments

We would like to acknowledge the ECMWF for providing their reanalysis and to the NOAA for providing the NAO and SF datasets. Tremendous appreciation is extended to the Spanish and Andalusian Institutions (AEMET, SAIH and Andalusian Environmental Department) that maintain and make accessible environmental datasets to scientific research.

References

- Aguilar, C., Herrero, J., Millares, A., Losada, M.A., Polo, M.J., 2014. Meteomap: generation of meteorological variables for distributed physically-based hydrological modeling, in: 11th International Conference on Hydroinformatics. New York. http://academicworks.cuny.edu/cc_conf_hic/278
- Balsamo, G., Albergel, C., Beljaars, a, Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., Vitart, F., 2015. ERA-Interim/Land: a global land surface reanalysis data set. *Hydrol. Earth Syst. Sci.* 389–407. doi:10.5194/hess-19-389-2015
- Cannon, A.J., 2010. A flexible nonlinear modelling framework for nonstationary generalized extreme value analysis in hydroclimatology. *Hydrol. Process.* 24, 673–685. doi:10.1002/hyp.7506

Contreras, E., Polo, M. J. 2012. Measurement frequency and sampling spatial domains required to characterize turbidity and salinity events in the Guadalquivir estuary (Spain), *Nat. Hazards Earth Syst. Sci.*, 12, 2581-2589, doi:10.5194/nhess-12-2581-2012, 2012. Cullen, H.M., Kaplan, A., Arkin, P. a., DeMenocal, P.B., 2002. Impact of the North Atlantic Oscillation on Middle Eastern climate and streamflow. *Clim. Change* 55, 315–338. doi:10.1023/A:1020518305517

Dee, D.P., Uppala, S.M., Simmons, a. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. a., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, a. C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, a. J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, a. P., Monge-Sanz, B.M., Morcrette, J.J., Park, B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N., Vitart, F., 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* 137, 553–597. doi:10.1002/qj.828

Egüen, M., 2014. Sources of variability in the forcing agents of the dynamics in Atlantic estuaries in Andalusia (in Spanish). University of Granada.

Egüen, M., Aguilar, C., Herrero, J., Millares, a., Polo, M.J., 2012. On the influence of cell size in physically-based distributed hydrological modelling to assess extreme values in water resource planning. *Nat. Hazards Earth Syst. Sci.* 12, 1573–1582. doi:10.5194/nhess-12-1573-2012

Egüen, M., Aguilar, C., Moreno, I., Herrero, J., Millares, A., Polo, M.J., Losada, M.A., 2009. WiMMed, a distributed physically-based watershed model (II): Application examples, in: *Environmental Hydraulics - Theoretical, Experimental and Computational Solutions:*

Proceedings of the International Workshop on Environmental Hydraulics, IWEH09, 29 & 30 October 2009, Valencia, Spain. Valencia, pp. 229–231.

Gómez-Beas, R., Moñino, A., Polo, M. J. 2012. Development of a management tool for reservoirs in Mediterranean environments based on uncertainty analysis, *Nat. Hazards Earth Syst. Sci.*, 12, 1789-1797, doi:10.5194/nhess-12-1789-2012, 2012. Herrero, J., 2007. Physical model of accumulation and snowmelt. Application in Sierra Nevada (Spain). University of Granada.

Herrero, J., Millares, A., Aguilar, C., Egüen, M., Losada, M.A., Polo, M.J., 2014. Coupling spatial and time scales in the hydrologic modeling of mediterranean regions: WiMMed, in: 11th International Conference on Hydroinformatics. New York, pp. 11–16.

Hurrell, J.W., 1995. Decadal trends in the north atlantic oscillation: regional temperatures and precipitation. *Science* 269, 676–679. doi:10.1126/science.269.5224.676

Hurrell, J.W., Lyons, W.B., Minschwaner, K.R., Strobel, D., Young, W.R., 2003. Climatic Significance and Environmental Impact The North Atlantic Oscillation. *Geophys. Monogr.* 134.

Izaguirre, C., Mendez, F.J., Menendez, M., Luceño, A., Losada, I.J., 2010. Extreme wave climate variability in southern Europe using satellite data. *J. Geophys. Res.* 115, 1–13. doi:10.1029/2009JC005802

Jiongxin, X., 2005. The water fluxes of the Yellow River to the sea in the past 50 years, in response to climate change and human activities. *Environ. Manage.* 35, 620–631. doi:10.1007/s00267-004-3094-y

Kiem, A.S., Franks, S.W., 2001. On the identification of ENSO-induced rainfall and runoff variability: a comparison of methods and indices. *Hydrol. Sci. J.* 46, 715–727. doi:10.1080/02626660109492866

Kottegoda, N., 1980. Stochastic water resources technology. London.

Kottegoda, N.T., Natale, L., Raiteri, E., 2014. Monte Carlo Simulation of rainfall hyetographs for analysis and design. *J. Hydrol.* 519, 1–11. doi:10.1016/j.jhydrol.2014.06.041
Kottegoda, N.T., Rosso, R. 2008. Applied statistics for civil and environmental engineers. 2nd edition. Blackwell Publishing Ltd. ISBN: 978-1-4051-7917-1.

Lavers, D., Prudhomme, C., Hannah, D.M. 2013. European precipitation connections with large-scale mean sea-level pressure (MSLP) fields. *Hydrolog. Sci. J.*, 58 (2), 310–327. doi:10.1080/02626667.2012.754545.

López-Moreno, J.I., Beguería, S., Vicente-Serrano, S.M., García-Ruiz, J.M., 2007. Influence of the North Atlantic Oscillation on water resources in central Iberia: Precipitation, streamflow anomalies, and reservoir management strategies. *Water Resour. Res.* 43, 1–17. doi:10.1029/2007WR005864

López-Moreno, J.I., Vicente-Serrano, S.M., 2008. Positive and negative phases of the wintertime North Atlantic Oscillation and drought occurrence over Europe: A multitemporal-scale approach. *J. Climate* 2, 1220-1243.

doi: <http://dx.doi.org/10.1175/2007JCLI1739.1> Lorenzo-Lacruz, J., Vicente-Serrano, S.M.,

López-Moreno, J.I., González-Hidalgo, J.C., Morán-Tejeda, E., 2011. The response of Iberian rivers to the North Atlantic Oscillation. *Hydrol. Earth Syst. Sci.* 15, 2581–2597. doi:10.5194/hess-15-2581-2011

Lütkepohl, H., 2005. New introduction to multiple time series analysis. Springer-Verlag.

- Maraun, D., Rust, H.W., Osborn, T.J., 2010. Synoptic airflow and UK daily precipitation extremes. *Extremes* 13, 133–153. doi:10.1007/s10687-010-0102-x
- Martin-Vide, J., Lopez-Bustins, J.A., 2006. The western mediterranean oscillation and rainfall in the Iberian Peninsula. *Int. J. Climatol.* 26, 1455–1475. doi:10.1002/joc.1388
- Montanari, A., Longoni, M., Rosso, R., 1999. A seasonal long-memory stochastic model for the simulation of daily river flows. *Phys. Chem. Earth, Part B Hydrol. Ocean. Atmos.* 24, 319–324. doi:10.1016/S1464-1909(99)00007-6
- Montanari, A., Rosso, R., Taqqu, M.S. 1997. Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation, and simulation. *Water Resour. Res.* 33, 1035. doi:10.1029/97WR00043
- Pires, C.A., Perdigão, R.A.P. 2007. Non-Gaussianity and asymmetry of the winter monthly precipitation estimation from the NAO. *Mon. Weather Rev.* 135.2: 430-448.
- Polo, M.J., Aguilar, C., Millares, A., Herrero, J., Gómez-Beas, R., Contreras, E., Losada, M.A. 2014a. Assessing risks for integrated water resource management: coping with uncertainty and the human factor. *Evolving Water Resources Systems: Understanding, Predicting and Managing Water-Society Interactions.* 364: 285-291. IAHS Press.DOI: 10.5194/piahs-364-285-2014
- Polo, M.J., Egüen, M., Nieto, S., Herrero, J., Baquerizo, A., Losada, M. A. 2014b. Weather Generator Tool For The Assessment Of Water Resource Planning In Mediterranean Watersheds. in: 11th International Conference on Hydroinformatics. New York. http://academicworks.cuny.edu/cc_conf_hic/416
- Polo, M.J., Herrero, J., Aguilar, C., 2009. WiMMed, a distributed physically-based watershed model (I): Description and validation, in: *Environmental Hydraulics: Theoretical, Experimental and Computational Solutions* -

Proceedings of the International Workshop on Environmental Hydraulics, IWEH09, 29 & 30 October 2009, Valencia, Spain. Valencia, pp. 225–228.

Roy, I., Haigh, J.D., 2011. The influence of solar variability and the quasi-biennial oscillation on lower atmospheric temperatures and sea level pressure. *Atmos. Chem. Phys.* 11, 11679–11687. doi:10.5194/acp-11-11679-2011

Salgueiro, A.R., Machado, M.J., Barriendos, M., Pereira, H.G., Benito, G., 2013. Flood magnitudes in the Tagus River (Iberian Peninsula) and its stochastic relationship with daily North Atlantic Oscillation since mid-19th Century. *J. Hydrol.* 502, 191–201. doi:10.1016/j.jhydrol.2013.08.008

Schwarz, G., 1978. Estimating the Dimension of a Model. *The Annals of Statistics* 6 (2), 461–464

Smith, E.P., 2002. An Introduction to Statistical Modeling of Extreme Values. *Technometrics*. doi:10.1198/tech.2002.s73

Solari, S., Losada, M. a., 2011. Non-stationary wave height climate modeling and simulation. *J. Geophys. Res. Ocean.* 116, 1–18. doi:10.1029/2011JC007101

Tramblay, Y., Neppel, L., Carreau, J., Sanchez-Gomez, E., 2012. Extreme value modelling of daily areal rainfall over Mediterranean catchments in a changing climate. *Hydrol. Process.* 26, 3934–3944. doi:10.1002/hyp.8417

Trigo, R.M., Pozo-Vázquez, D., Osborn, T.J., Castro-Díez, Y., Gámiz-Fortis, S., Esteban-Parra, M.J., 2004. North Atlantic oscillation influence on precipitation, river flow and water resources in the Iberian Peninsula. *Int. J. Climatol.* 24, 925–944. doi:10.1002/joc.1048

- Vasiliades, L., Galiatsatou, P., Loukas, A., 2015. Nonstationary Frequency Analysis of Annual Maximum Rainfall Using Climate Covariates. *Water Resour. Manag.* 29, 339–358. doi:10.1007/s11269-014-0761-5
- Wang, H., Yang, Z., Saito, Y., Liu, J.P., Sun, X., 2006. Interannual and seasonal variation of the Huanghe (Yellow River) water discharge over the past 50 years: Connections to impacts from ENSO events and dams. *Glob. Planet. Change* 50, 212–225. doi:10.1016/j.gloplacha.2006.01.005
- Ward, P.J., Eisner, S., Flo Rke, M., Dettinger, M.D., Kummerow, M., 2014. Annual flood sensitivities to el nintild;O-Southern Oscillation at the global scale. *Hydrol. Earth Syst. Sci.* 18, 47–66. doi:10.5194/hess-18-47-2014
- Wilks, D. S. 2009. A gridded multisite weather generator and synchro- nization to observed weather data, *Water Resour. Res.*, 45, 1–11, doi:10.1029/2009WR007902.
- Wolf, D., Seim, A., Faust, D., 2014. Fluvial system response to external forcing and human impact - Late Pleistocene and Holocene fluvial dynamics of the lower Guadalete River in western Andalucía (Spain). *Boreas* 43, 422–449. doi:10.1111/bor.12044

Appendix A: Supplementary materials

Appendix B: Analysis of the distribution function of the SLP

The distribution function of the *6hr-SLP* over the entire study period (September 1957–August 2014) was evaluated. The empirical distribution function does not fit simple theoretical distribution functions. Fig. B.1 shows the adjustment to a Normal and a Weibull distribution.

Monthly data analysis reveals that the distribution function has a non-stationary behavior. Therefore, during November to March, the best fit obtained is to a Weibull distribution function, whereas for April to October, the data better fit a Normal distribution function (Fig. B.2). A similar behavior has been noted in other coastal areas of the Iberian Peninsula (results not shown in this paper).

The results of this preliminary analysis are the basis of the non-stationary mixed model proposed in Section 3.1.

Appendix C: Lists of abbreviations and symbols

Table 1. Mean ($\bar{x}$) and standard deviation (σ_x) of the monthly rainfall, P_m , (mm/month), and flows, Q_m , (hm^3/month), per type of year.

			Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
P (mm/month)	Dry	$\bar{x}$	33.7	94.5	101.6	59.6	62.3	51.8	58.8	69.2	44.1	9.7	0.9	6.0
		σ_x	23.5	67.8	75.4	58.8	44.4	43.4	68.5	44.5	33.9	17.6	2.0	15.5
	Medium	$\bar{x}$	24.8	74.2	106.2	136.1	83.3	98.3	86.9	61.2	44.9	14.2	1.1	3.2
		σ_x	29.1	62.6	90.3	87.0	61.8	69.2	65.3	41.6	39.6	13.8	2.8	5.0
	Wet	$\bar{x}$	23.8	59.1	92.8	204.8	168.0	130.9	71.6	50.4	35.4	22.4	1.9	3.7
		σ_x	22.8	49.5	62.7	140.0	100.3	82.4	43.1	29.0	31.6	19.4	4.1	5.3
Q (hm^3/month)	Dry	$\bar{x}$	2.3	11.6	31.0	28.7	26.4	19.5	25.3	19.0	7.4	3.3	1.3	1.4
		σ_x	2.2	20.1	48.0	54.6	37.4	26.2	54.8	25.7	7.5	6.9	2.7	2.3
	Medium	$\bar{x}$	2.0	8.3	32.1	67.1	68.0	74.8	61.6	29.7	18.0	4.3	2.0	3.5
		σ_x	2.3	10.1	46.3	80.7	81.2	69.5	56.3	25.8	25.5	7.3	4.3	10.6
	Wet	$\bar{x}$	3.1	7.6	20.8	124.3	181.3	146.0	77.7	27.1	15.7	6.5	2.5	6.5
		σ_x	3.1	9.4	21.4	150.6	156.4	128.0	63.2	24.8	15.6	7.2	3.0	14.7

Table 2. Maximum, minimum, mean ($\bar{x}$) and standard deviation (σ_x) of the maximum daily winter rainfall, P_{maxd} , and flow, Q_{maxd} ; a_G and b_G are parameters of the Gamma distribution per type of year.

Statistics	P_{maxd} (mm/day)				Q_{maxd} (m ³ /s)			
	min	$\bar{x}$	max	σ_x	min	$\bar{x}$	max	σ_x
Dry	3.5	25.9	42.6	10.6	2.9	119.8	585.2	141.7
Medium	16.7	37.9	88.7	15.7	29.8	331.1	971.3	227.9
Wet	16.0	40.7	67.5	13.0	69.2	465.5	1348.3	317.0
Gamma Parameters	a_G		b_G		a_G		b_G	
Dry	4.4		5.8		0.9		131.3	
Medium	8.0		4.7		1.9		171.9	
Wet	9.6		4.2		2.6		181.4	

Table 3. The 5th, 50th, 90th percentiles (F_5 , F_{50} , F_{90}), mean ($\bar{x}$) and standard deviation (σ_x) of the observed and simulated monthly rainfall (P_m and P_{sm} , mm/month) and flows (Q_m and Q_{sm} , hm³/month). The statistics of the simulated series represent the average over the 100 replicates generated for each variable.

		Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
P_m (mm/month)	F_5	0.7	2.0	8.5	7.5	0.1	0.6	6.0	7.4	2.1	0.0	0.0	0.0
	F_{50}	20.7	59.6	88.8	96.8	79.0	83.3	56.2	51.9	28.6	7.3	0.1	0.7
	F_{90}	60.5	160.8	220.9	284.7	190.4	179.8	144.5	101.7	94.2	38.5	2.8	10.1
	$\bar{x}$	27.8	77.2	100.5	128.4	100.8	90.9	72.0	60.9	41.8	15.0	1.3	4.4
	σ_x	25.3	61.8	76.2	113.3	82.7	72.3	61.0	39.5	34.9	17.5	3.0	10.2
P_{sm} (mm/month)	F_5	0.3	2.0	6.3	7.4	1.0	1.2	4.5	5.8	2.3	0.1	0.0	0.0
	F_{50}	18.5	56.3	83.3	92.6	73.8	72.7	55.5	50.4	27.4	7.0	0.1	0.6
	F_{90}	58.5	149.7	190.9	260.9	181.1	179.2	133.1	100.3	91.4	34.1	2.8	9.5
	$\bar{x}$	25.2	71.7	93.5	117.3	90.6	84.2	66.5	56.9	38.9	13.0	0.9	3.4
	σ_x	23.0	58.2	71.6	101.0	74.3	68.3	53.1	36.5	32.6	14.7	2.2	6.2
Q_m (hm ³ /month)	F_5	0.0	0.3	0.4	1.8	2.2	2.1	0.7	1.6	0.1	0.0	0.0	0.0
	F_{50}	1.7	3.7	9.5	21.1	44.3	39.2	28.0	13.3	8.0	1.2	0.2	0.2
	F_{90}	6.8	24.9	66.0	183.5	255.6	196.6	143.1	57.0	26.7	13.1	5.8	5.7
	$\bar{x}$	2.5	9.3	28.1	70.2	87.5	75.7	52.6	24.7	13.1	4.6	1.9	3.6
	σ_x	2.5	14.5	40.5	106.9	119.1	97.0	61.3	25.4	17.4	7.1	3.3	10.2
Q_{sm} (hm ³ /month)	F_5	0.0	0.2	0.3	1.9	1.9	1.6	1.0	1.3	0.1	0.0	0.0	0.0
	F_{50}	1.5	3.6	8.8	18.9	29.7	34.9	25.7	14.2	7.9	1.1	0.2	0.2
	F_{90}	5.4	20.7	65.8	157.2	201.1	177.9	130.5	49.7	24.6	11.7	4.9	5.4
	$\bar{x}$	2.1	7.8	24.4	59.0	73.5	66.6	46.6	22.4	11.8	3.8	1.5	2.6
	σ_x	2.2	11.2	33.1	87.7	99.6	82.2	51.3	21.3	13.9	5.8	2.7	7.2

Table 4. The 5th, 50th, 90th percentiles (F_5 , F_{50} , F_{90}), mean ($\bar{x}$) and standard deviation (σ_x) of the observed and simulated maximum daily winter rainfall (P_{maxd} and P_{s_maxd}) and flows (Q_{maxd} and Q_{s_maxd}). The statistics of the simulated series represent the average over the 100 replicates generated for each variable. The 95% confidence limits are shown in brackets.

	P_{maxd} (mm/day)	P_{s_maxd} (mm/day)	Q_{maxd} (m ³ /s)	Q_{s_maxd} (m ³ /s)
F_5	12.6	12.2 (12.0-15.8)	7.9	13.8 (11.8-32.2)
F_{50}	32.2	32.6 (32.3-36.8)	174.8	214.3 (213.9-268.3)
F_{90}	46.5	53.1 (52.5-60.5)	450.6	636.4 (622.0-771.5)
$\bar{x}$	34.3	34.3 (34.1-36.9)	246.2	290.5 (289.4-333.0)
σ_x	14.5	14.6 (14.5-17.9)	242.7	263.0 (258.6-331.5)

Table 5. Parameters for the Gamma distribution function fitted to the observed and simulated series of maximum daily winter precipitation, P_{maxd} , and flow, Q_{maxd} , per type of year (dry, medium and wet). The parameters of the simulated series represent the average over the 100 replicates generated for each variable.

	Observed		Simulated	
	a	b	a	b
<u>P_{maxd}</u>				
Dry	4.4	5.8	4.5	5.9
Medium	8.0	4.7	7.6	4.9
Wet	9.6	4.2	9.5	4.3
Total	5.2	6.6	5.1	6.6
<u>Q_{maxd}</u>				
Dry	0.9	131.4	0.9	128.3
Medium	1.9	171.9	2.0	163.8
Wet	2.6	181.4	2.6	175.2
Total	0.9	270.1	1.1	265.3

Table A.1. Summary of the datasets used, including the source, time step, units and magnitude (see Table C.1 for abbreviations).

Variable	Source	Time step	Date range	Units	max	med	min
SLP	ECMWF	6 hourly	Sep. 1957- Aug. 2002	hPa	1039.6	1017.5	982.9
NAO	CPC-NOAA	Monthly	Sep. 1957- Aug. 2014	-	3.06	0	-3.14
SF	NGDC- NOAA	Monthly	Sep. 1957- Aug. 2014	sfu	2812	1259	665
P	AEMET/RIA	Daily	Sep. 1957- Aug. 2014	mm	88.6	2	0
Q	SAIH	Daily	Sep. 1960- Aug. 2014	m ³ /s	1348	12	0

Table C.1. List of abbreviations

Abbreviation	Description
AR	Autoregressive model
BIC	Bayesian information criterion
NAO	North Atlantic Oscillation
RMSE	Root mean square error
SF	Solar Flux
SLP	Sea level pressure
W-N	Weibull-Normal

TableC.2. List of symbols

Symbol	Description
$a(t), b(t)$	Parameters of the Weibull distribution
a_G, b_G	Parameters of the Gamma distribution
$C_i(t)$	Covariables
D	Total number of coefficients $\theta(t)$ in Fourier series
F	Normally distributed random value between 0 and 1
f_e	Empirical density function
F_e	Empirical cumulative distribution function
F_g	Gamma distribution function
F_{Pm}	Cumulative distribution functions of mean monthly rainfall
F_{Qm}	Cumulative distribution functions of mean monthly flow
f_N	Normal probability density function
f_W	Weibull probability density function
f_{W-N}	Weibull-Normal probability density function
F_{W-N}	Weibull-Normal cumulative distribution function
i	Number of covariables
k	Order of the series of Fourier
LLF	Log likelihood function
m	Calendar months $m = 1:12$
m_{P-P}	Slope of P-P plot
n	Normalized
N	Number of available observations of SLP
N_k	Maximum Fourier orden
N_y	Number of years equal to that of the original series (September 1957-August 2014).
NR	Number of repetitions
P_m	Monthly rainfall
$P_{m_Dry}, P_{m_Med}, P_{m_Wet}$	Monthly rainfall in dry, médium and wet period
P_{maxd}	Maximum daily winter rainfall
$P_{maxd,n}$	Maximum daily winter rainfall normalized
P_{sm}	Simulated monthly rainfall
P_{s_maxd}	Simulated maximum daily winter rainfall
P_w	Accumulated rainfall in the winter (DJF)
P_{wn}	Accumulated rainfall in the winter (DJF) normalized
Q_m	Monthly flow
$Q_{m_Dry}, Q_{m_Med}, Q_{m_Wet}$	Monthly flow in dry, médium and wet period
Q_{maxd}	Maximum daily winter flows
$Q_{maxd,n}$	Maximum daily winter flows normalized

$Q_{measured}$	Measured daily water at the control point
$Q_{natural}$	Natural flow estimated at the control point
Q_{sm}	Simulated monthly flows
Q_{s_maxd}	Simulated maximum daily winter flows
Q_{stored}	Amount of water daily stored in the reservoirs
Q_w	Winter water flows (JFM)
Q_{wn}	Winter water flows (JFM) normalized
r	Correlation coefficient
SLP_w	Mean winter sea level atmospheric pressure (DJF)
$\overline{SLP_w}$	Mean value of SLP_w for the whole study period (September 1957-August2014)
SLP_{28}	28-days moving average of observed $6hr$ - SLP values
t	Time period within the year
t_{sp}	Time step along the study period (September 1957-August 2014)
$\alpha_w(t)$	Weight of each distribution along the time
$\mu(t), \sigma(t)$	Parameters of the Normal distribution
θ_{a0}	Coefficient of the stationary term in Fourier series
θ_{ak}, θ_{bk}	Coefficients of the non-stationary terms in Fourier series
$6hr$ - SLP	6 hourly average sea-level pressure

Figure 1. a) Location of the Guadalete River Basin in the Iberian Peninsula and the pressure measurement point; b) details of the Guadalete Basin with the locations of the two main rivers, the dams and the control point of flow. Elevation is given in meters.

Figure 2. Time evolution of a) $a(t)$, b) $b(t)$, c) $\mu(t)$, d) $\sigma(t)$, and e) $\alpha_w(t)$ parameters for the seasonal $W-N$ model.

Figure 3. a) Empirical SLP probability density function, f_e , and probability density function calculated with the seasonal $W-N$ model, f_{W-N} . b) 28-day moving average, SLP_{28} , of the observed values and mean modeled SLP . The 90% confidence bounds are displayed in broken lines.

Figure 4. a) $W-N$ model distribution function, F_{W-N} , and sample distribution function, F_e . b) P-P plot and 1:1 line.

Figure 5. Mean monthly observed SLP and expected value of the non-stationary mixture distribution f_{W-N} fitted using as co-variable: a) NAO index (NAO), b) Solar Flux (SF), c) NAO index and solar flux (NAO & SF), and d) none (seasonal model).

Figure 6. Comparison of log-likelihood (LLF) and Bayesian Information Criterion (BIC) obtained with the non-stationary $W-N$ model with and without the inclusion of the covariables.

Figure 7. Interannual variability of the normalized mean winter observed (DJF), SLP_{wn} , and: a) mean winter (DJF) rainfall, P_{wn} ; b) mean winter river flow with one month of delay (JFM), Q_{wn} ; c) maximum daily winter rainfall, P_{maxdn} ; and d) maximum daily winter flow, Q_{maxdn} . All variables have been normalized; r is the correlation coefficient.

Figure 8. a) Mean monthly rainfall (P_m) and b) mean monthly river flow (Q_m) during the dry years ($SLP_w > (\overline{SLP_w} + 1 \text{ hPa})$) (dotted curve), wet years ($SLP_w < (\overline{SLP_w} - 1.5 \text{ hPa})$) (dashed curve) and medium years ($\overline{SLP_w} + 1 \text{ hPa} < SLP_w < \overline{SLP_w} - 1.5 \text{ hPa}$) (solid line).

Figure 9. Gamma distribution (f_g) per type of year for a) P_{maxd} and b) Q_{maxd} .

Figure 10. Mean monthly SLP , rainfall and flow for the study period (September 1957-August 2014): observed (in black) and confidence intervals of 95% of 100 simulated series (in grey).

Figure 11. Empirical cumulative distribution functions of the mean monthly rainfall (F_{Pm}) and flow (F_{Qm}) over the study period (September 1957-August 2014): observed (in black) and 100 simulated series (in grey).

Figure 12. Cumulative distribution functions of the maximum daily winter rainfall (F_{Pmaxd}) and flow (F_{Qmaxd}) over the study period (September 1957-August 2014): observed (in black) and 100 simulated series (in grey).

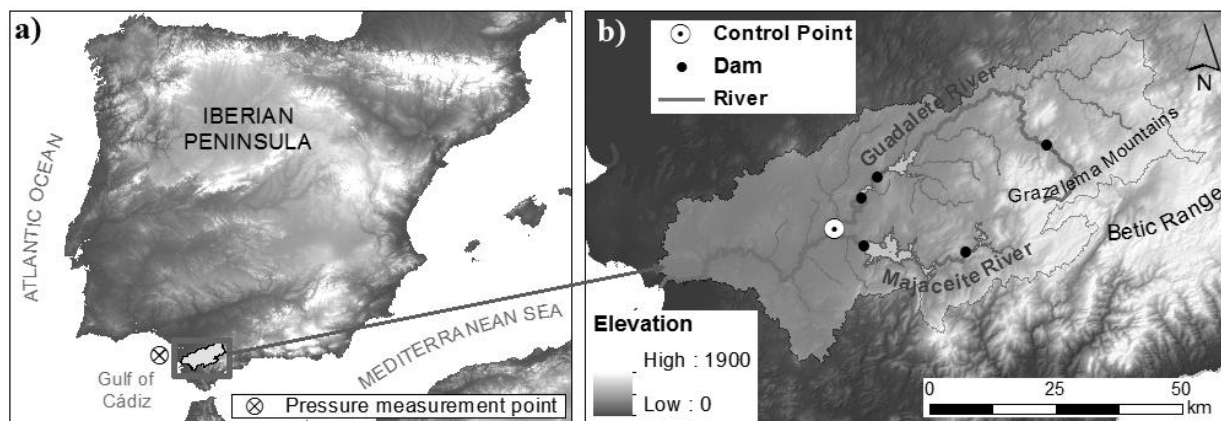
Figure 13. Maximum daily winter observed (circles) and simulated (grey shaded) values of a) precipitation, P_{maxd} , and b) flow, Q_{maxd} . The red line is the Gamma distribution fitted to the observed data.

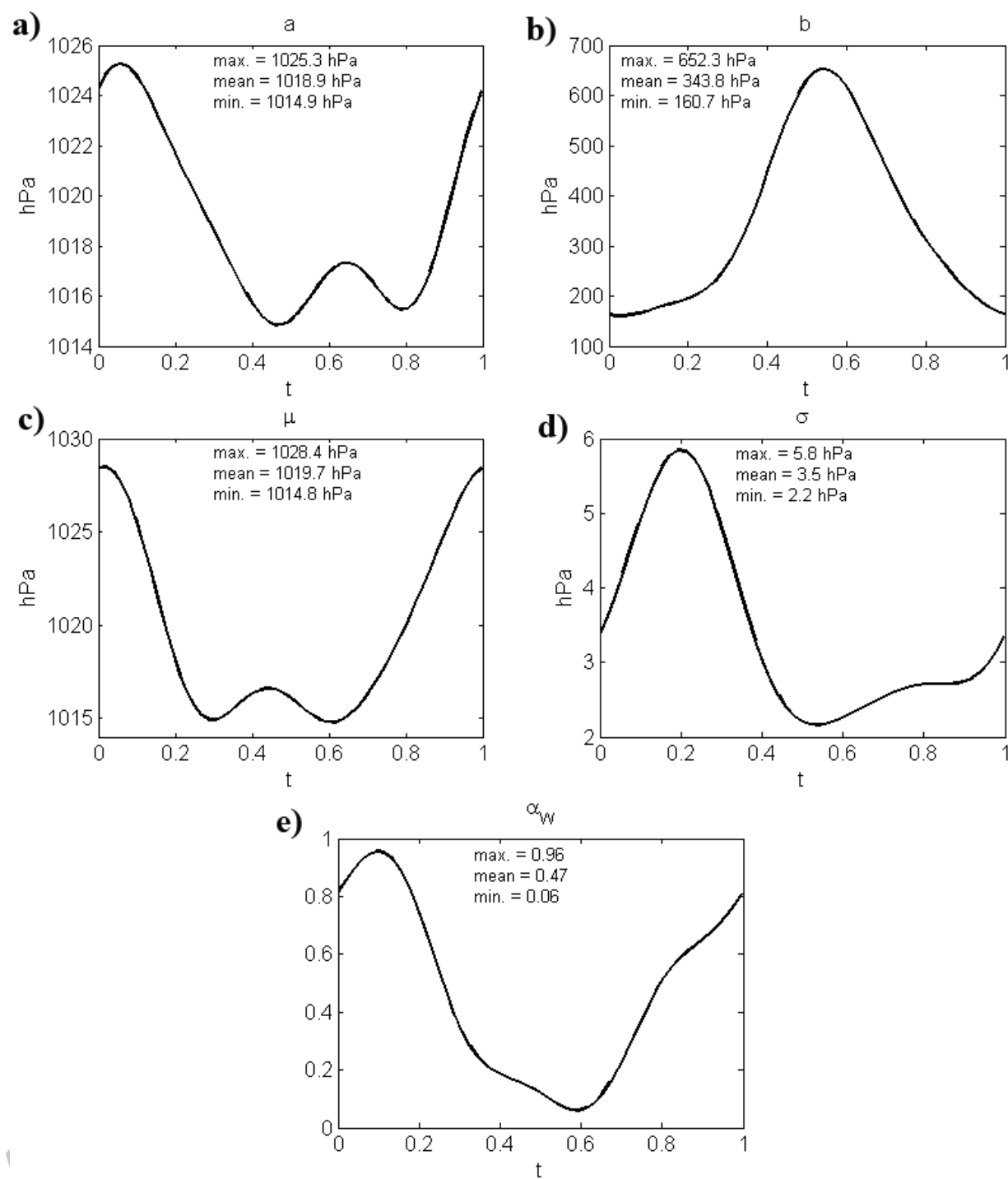
Figure A.1. Flow chart for the stochastic simulation of monthly and maximum daily winter P or $Q(X)$ from SLP .

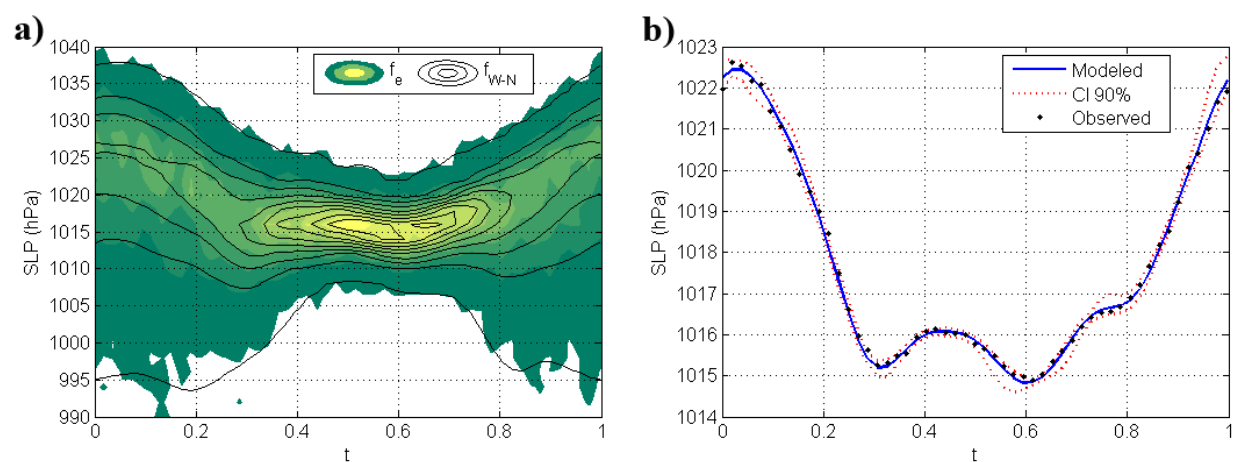
Figure A.2. Minimum error (minimum number of wrong years) obtained for different thresholds of SLP_w for the classification of wet, medium and dry years. Errors for a) monthly precipitation, P_m ; b) monthly flow, Q_m ; and c) total errors (sum of both types of errors).

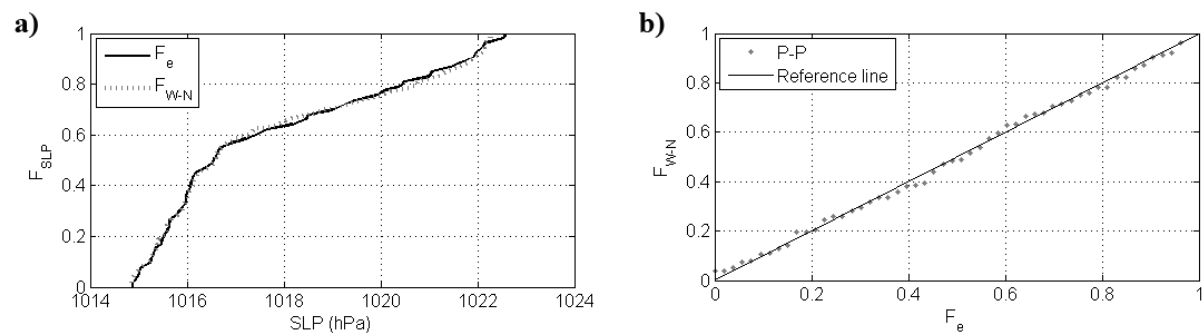
Figure B.1. Empirical probability density function, f_e , of the 6-hourly mean observed data of the SLP and the fit to a Normal, f_N , and a Weibull, f_W , distribution function.

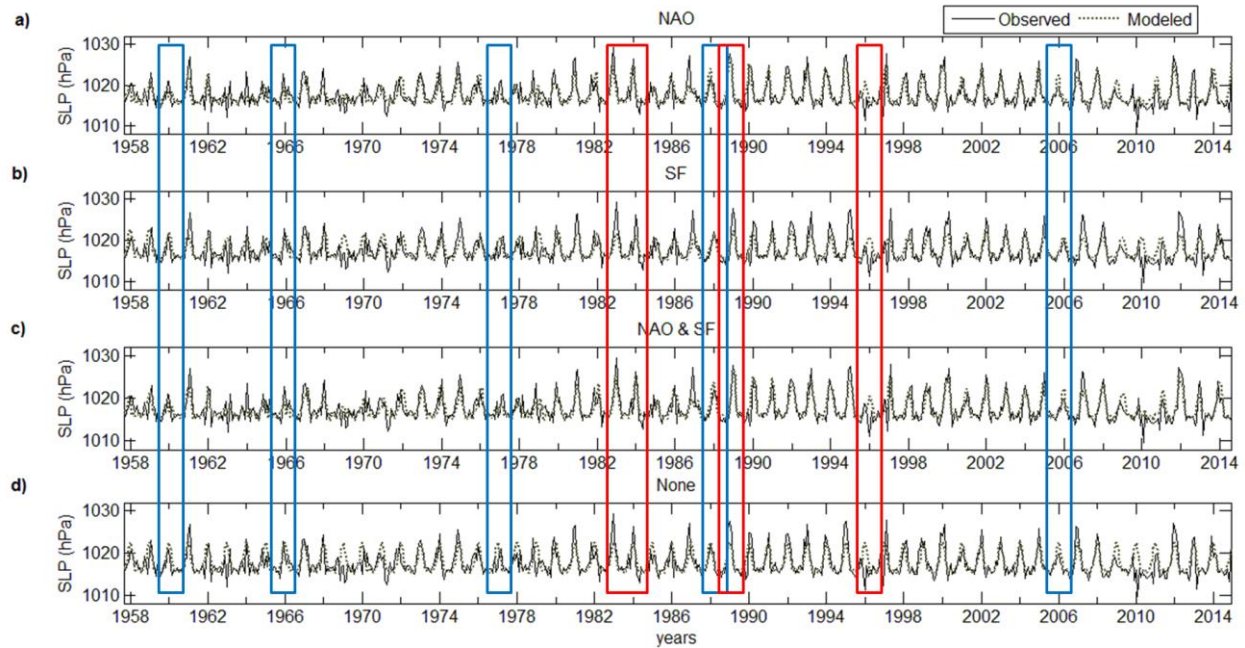
Figure B.2. Monthly empirical probability density function, f_e , of the $6hr-SLP$ and the fit to a Normal, f_N , and a Weibull, f_W , distribution function.

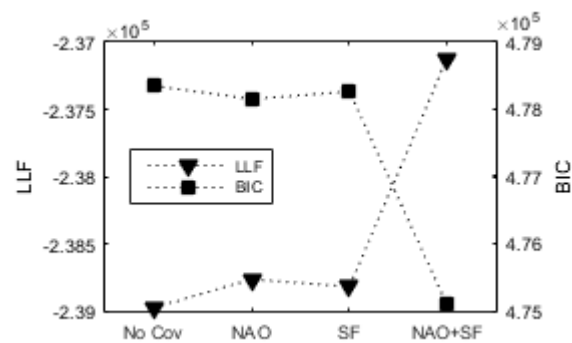


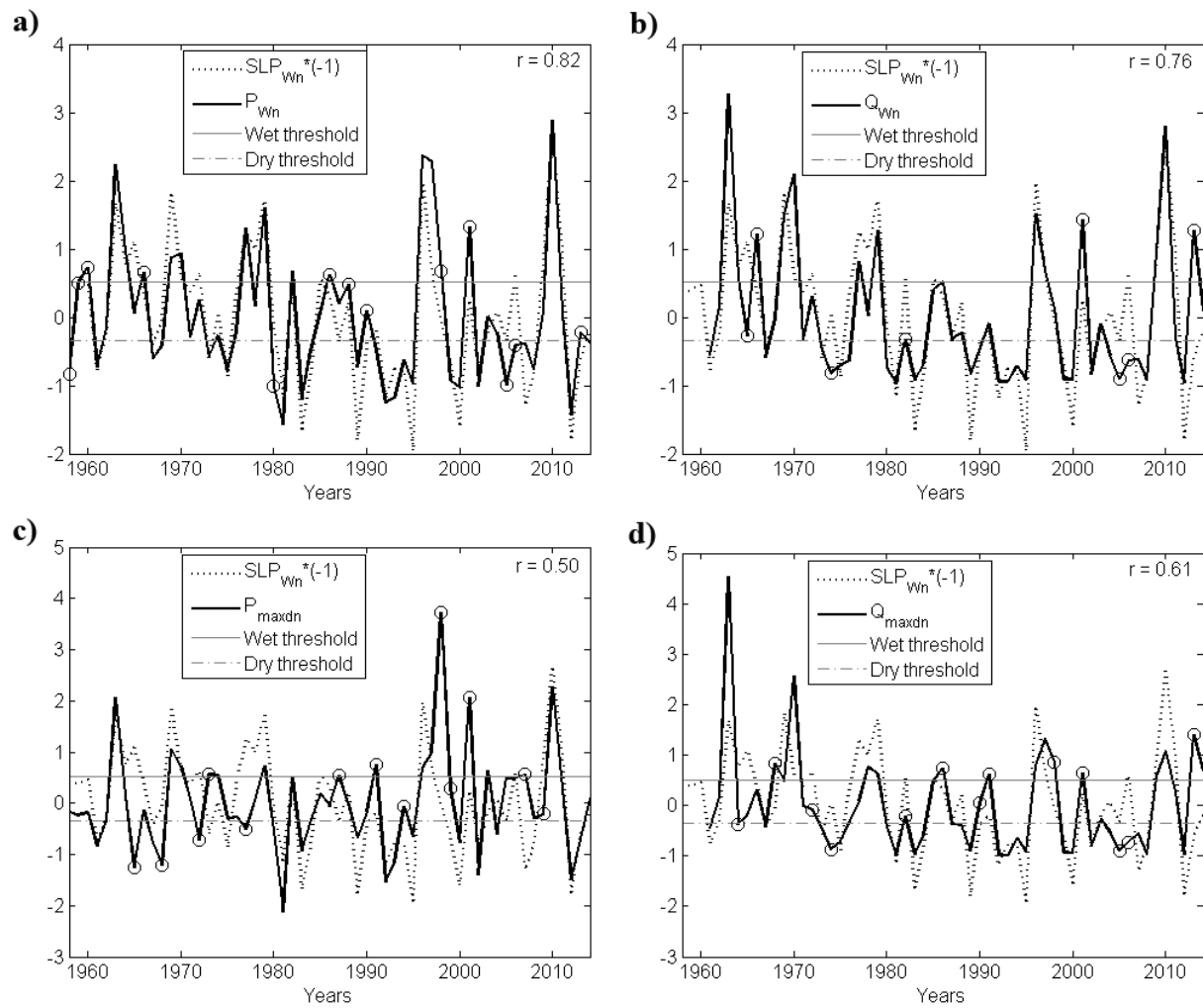


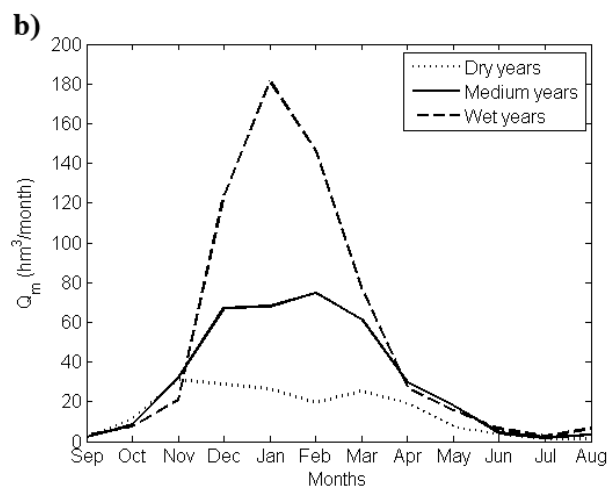
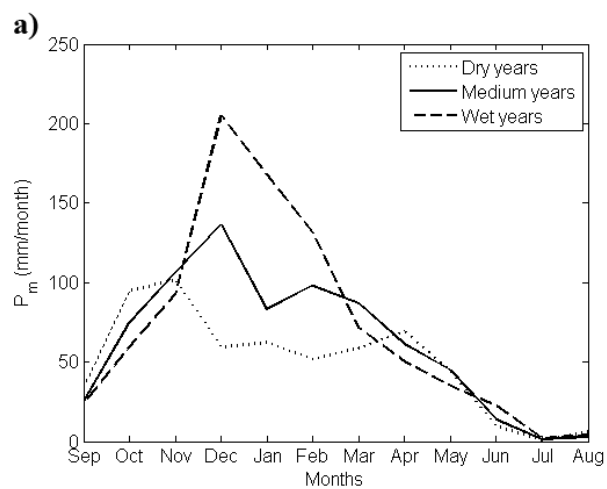


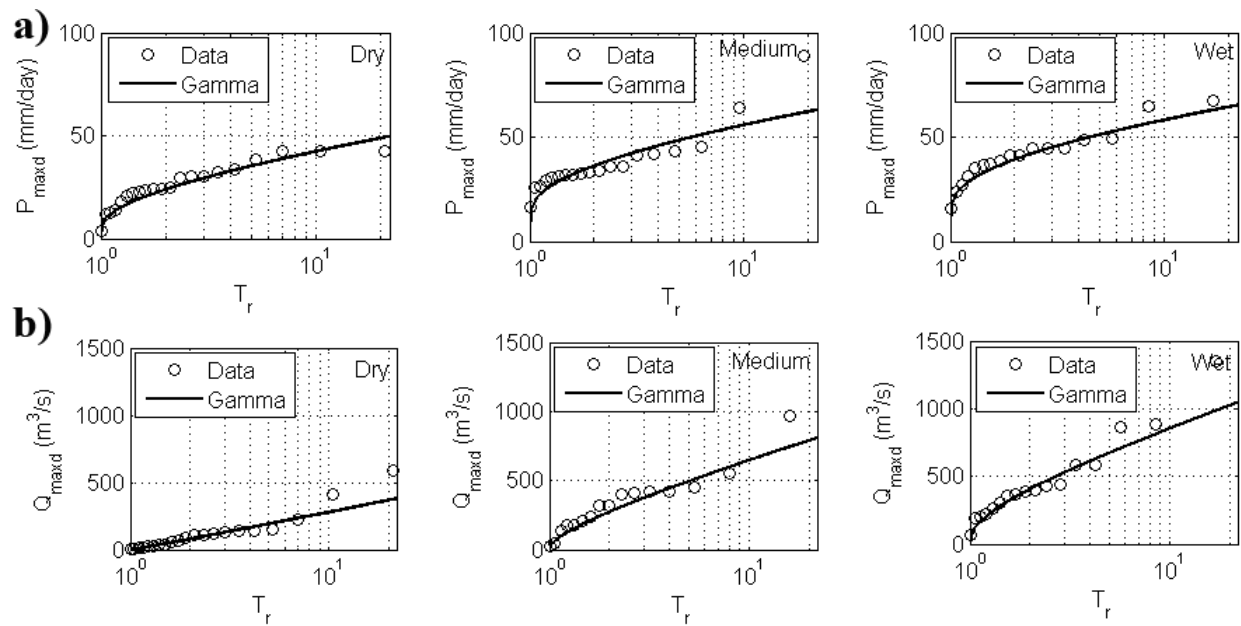


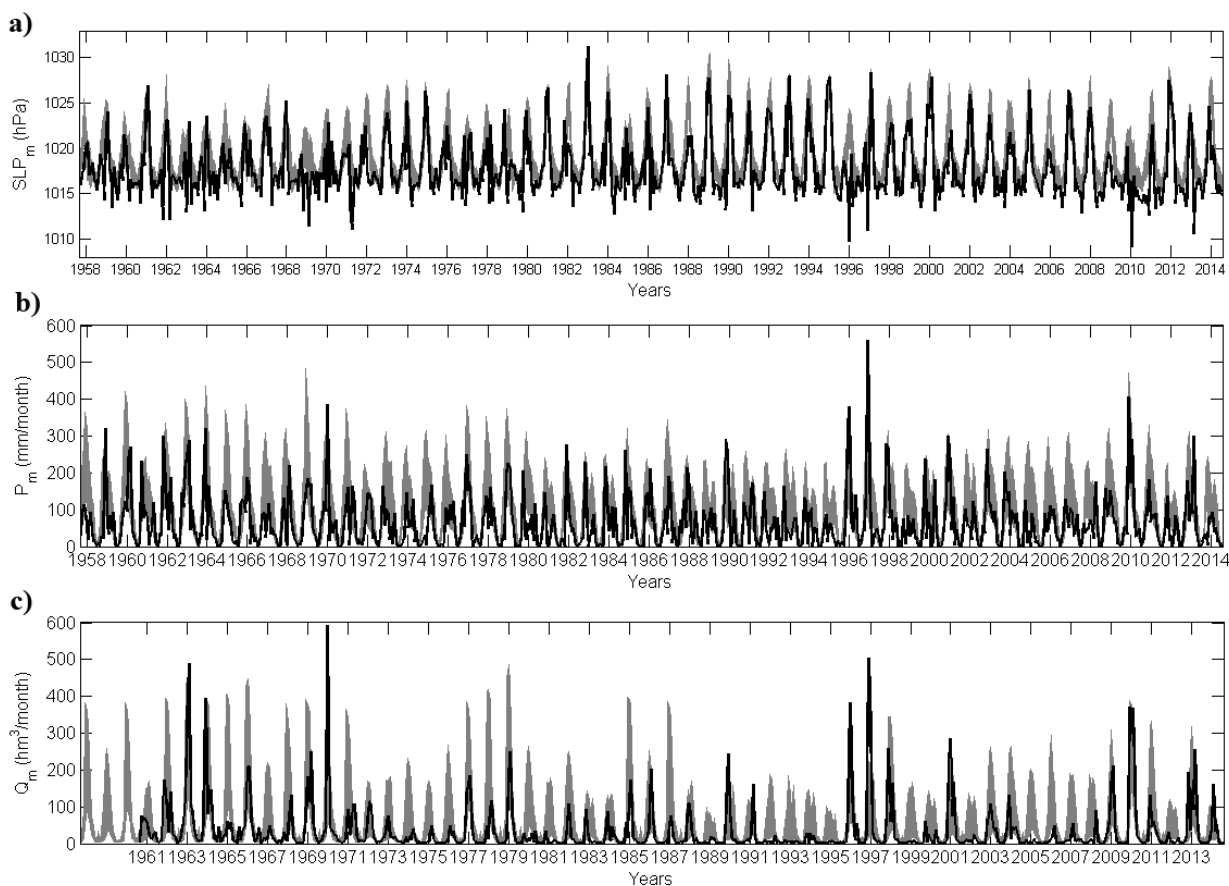




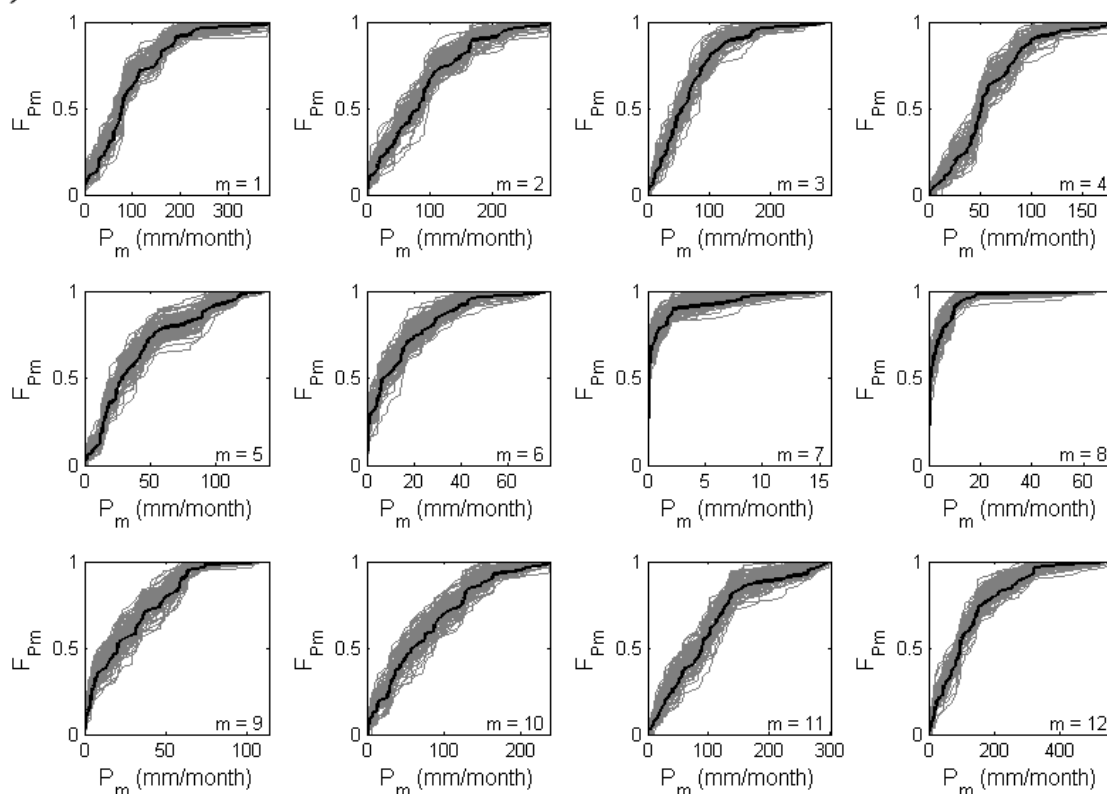




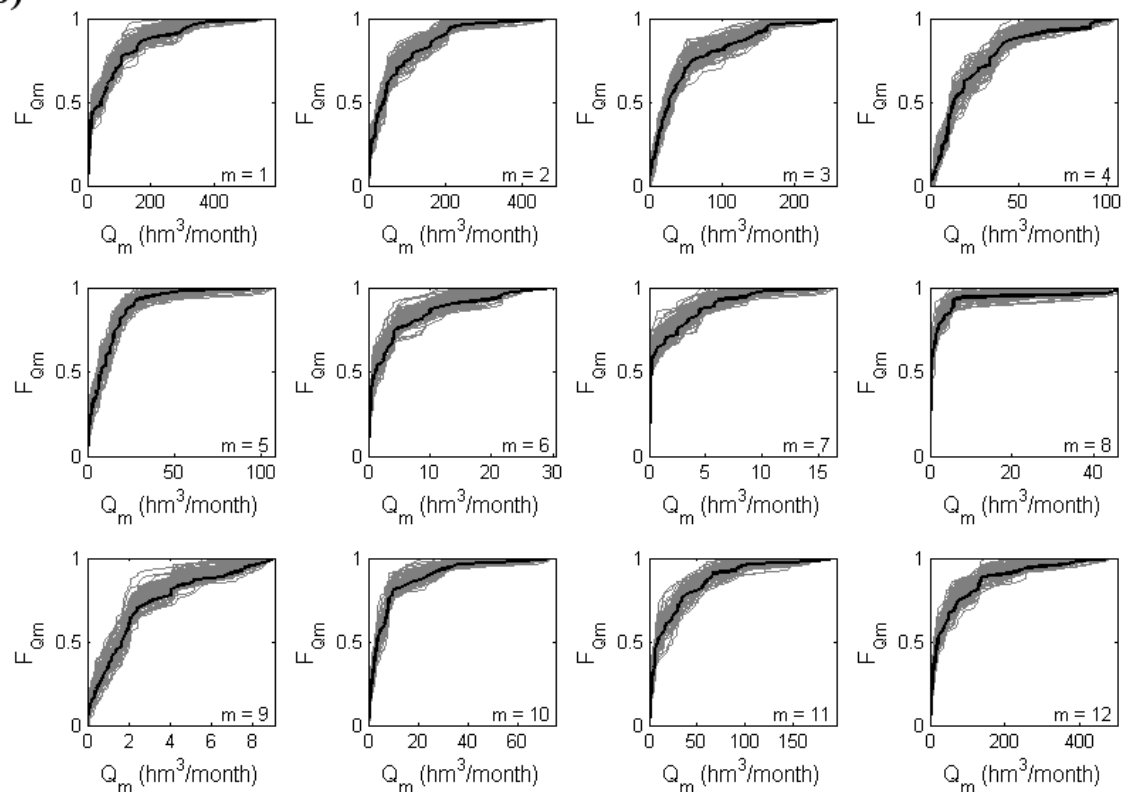


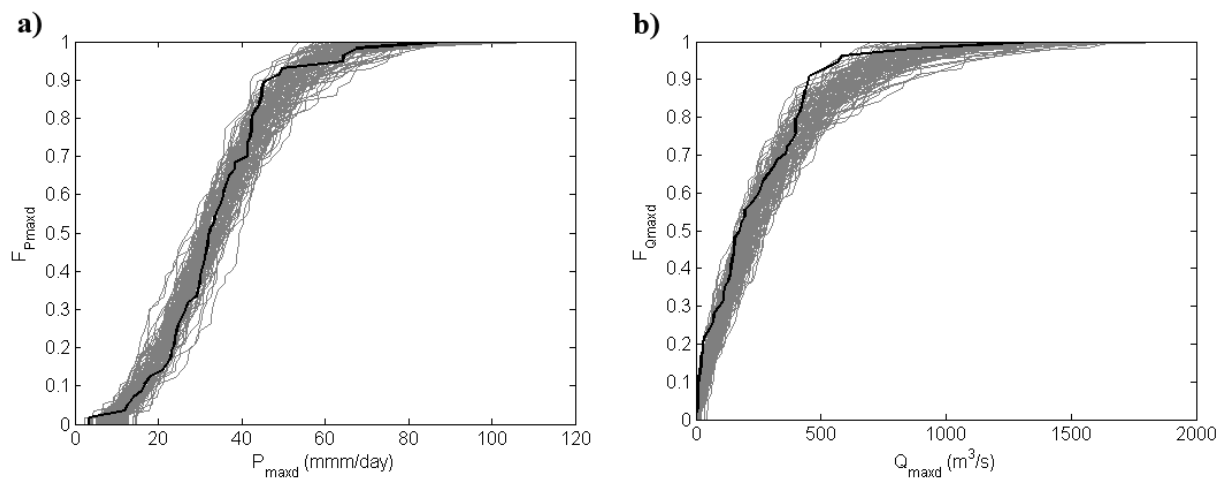


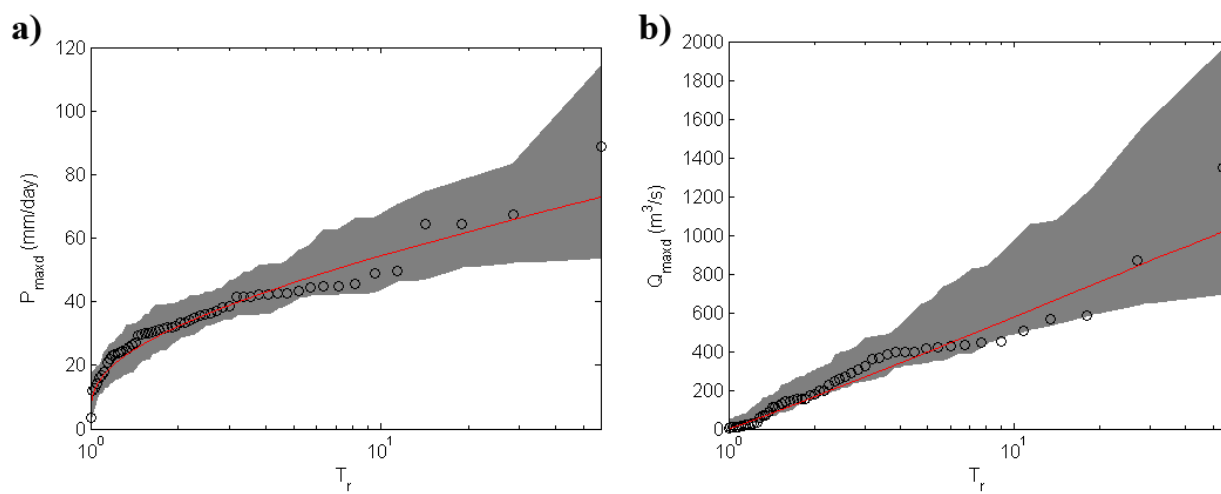
a)

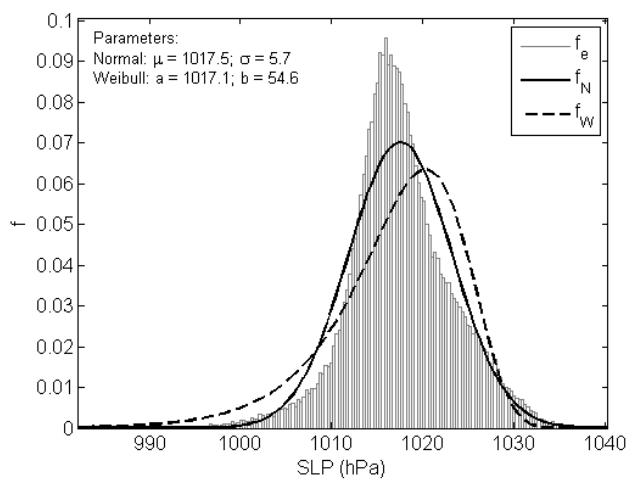


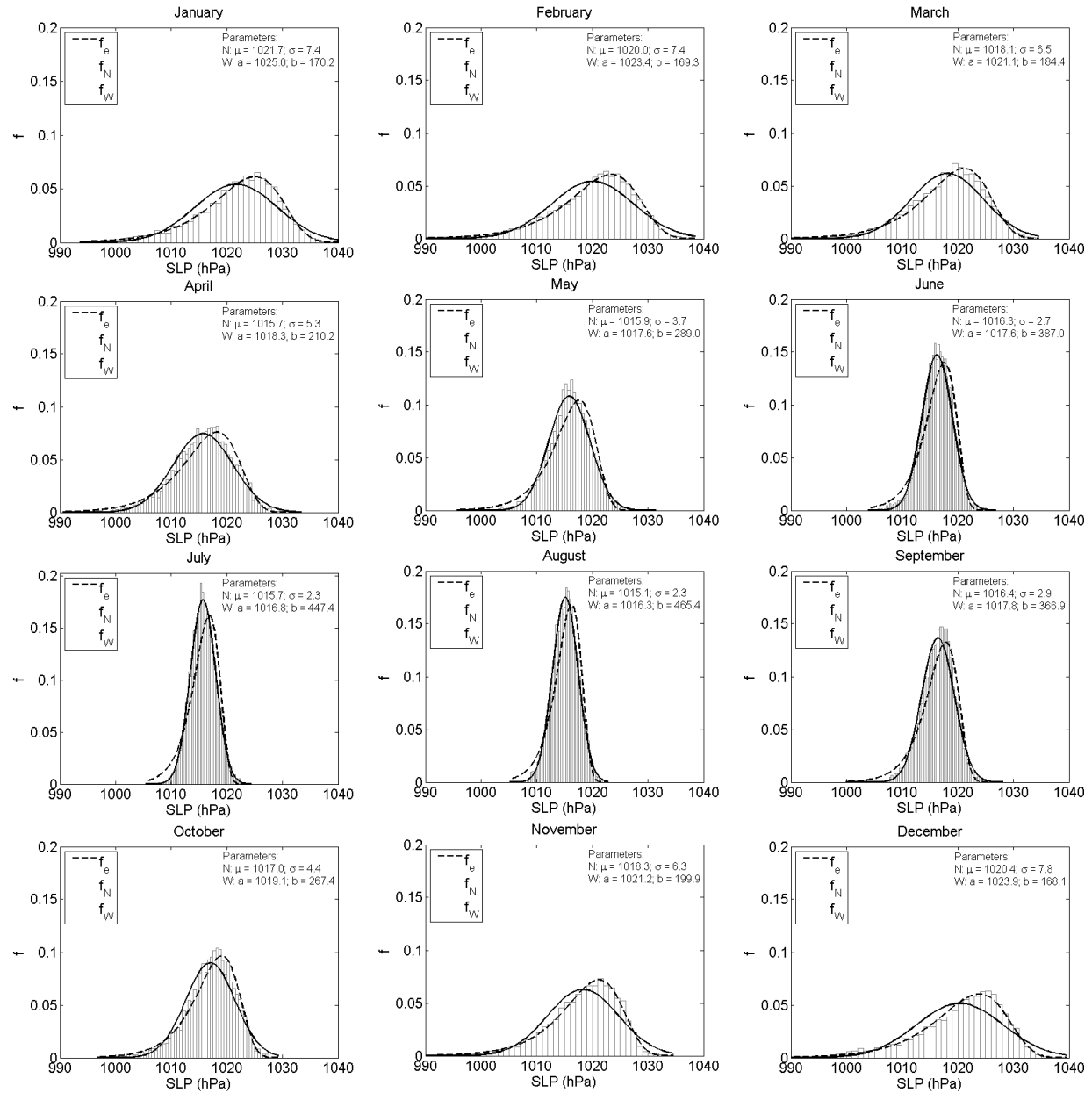
b)











- Stochastic simulation of hydrological variables from atmospheric pressure is proposed
- A non-stationary mixed distribution model is fitted to the atmospheric pressure
- Linear correlation between mean winter pressure and hydrological variables is found
- Monthly and maximum daily rainfall and flow data series are stochastically replicated