Marcos-García, P.; Brown, C.; Pulido-Velazquez, M. (2020). Development of Climate Impact Response Functions for highly regulated water resource systems. Journal of Hydrology. 590:1-14. https://doi.org/10.1016/j.jhydrol.2020.125251

DEVELOPMENT OF CLIMATE IMPACT RESPONSE FUNCTIONS FOR HIGHLY REGULATED WATER RESOURCE SYSTEMS

Patricia Marcos-Garcia[a,b*], Casey Brown[c], Manuel Pulido-Velazquez[a]

[a] *Research Institute of Water and Environmental Engineering (IIAMA), Universitat Politècnica de València, Camí de Vera s/n, 46022 Valencia, Spain*

[c] *Department of Civil and Environmental Engineering, University of Massachusetts, 12B Marston Hall, 130 Natural Resources Road, Amherst, Massachusetts 01003, USA*

Corresponding author: Patricia Marcos-Garcia (Patricia.MARCOS-GARCIA@ec.europa.eu). Present address: [b] *European Commission–Joint Research Centre, via E. Fermi 2749, 21027 Ispra, VA, Italy*

ABSTRACT

Climate Impact Response Functions (CIRFs) could be useful for exploring potential risks of system failure under climate change. The performance of a water resource system could be synthesized through a CIRF that relates climate conditions to system behavior regarding a certain threshold of deliveries to demands or environmental flow requirements. However, in highly regulated water resource systems this relationship could be quite complex, depending on storage capacity and system operation. In this paper we define a CIRF for these types of systems through a multivariable logistic regression (LR) model where a binary variable (system response) is explained by two continuous variables or predictors (precipitation and temperature). The approach involves generating multivariate synthetic inflow time series and relate them to specific climate conditions. Next, these inflows are used as inputs in a water management model, and the outcome is coded as a dichotomous variable (failure or its absence) depending on selected vulnerability criteria. To identify the time span before the failure event in which climate variables are relevant, we characterized drought development stages through relative standardized indices. Mean values of precipitation and temperature for

the selected time span are computed and used as explanatory variables through a LR model, which is validated using data from several climate models and scenarios. Results show that the predictive capacity of LR models is acceptable, so that they could be used as screening tools to detect challenging climate conditions for the system which would require adaption actions.

Keywords: water management; climate change; climate impact response functions; synthetic streamflow generation; multivariable logistic regression.

1. INTRODUCTION

A Climate Impact Response Function (CIRF) could be defined as a function explaining the relationship between changes in the climate variables and the environmental and socioeconomic impacts resulting of those changes (Toth et al., 2000). According to Füssel et al. (2003), there are three main application modes for CIRFs: 1) the "forward mode", which determines the likely impacts of a specific climate state or scenario; 2) the "overview mode", which allows the detection of possible non-linear responses of a system to changes in the forcing variables and; 3) the "inverse mode", which identifies the subset of climate states where a previously defined impact threshold is not violated.

The sources of climate information involved in CIRF definition are different depending on the mode. Traditionally, the computation of a CIRF in the "forward mode" involves the application of a geographically explicit impact model to a representative subset of plausible future climate states (Füssel et al., 2003). This impact-oriented or top-down approach moves from the global to the local scale: climate variables for different scenarios are derived from climate models (usually Global Climate Models (GCM), downscaled for a particular region through Regional Climate Models (RCM)), and used as inputs for a particular impact model (e.g. hydrological model). According to Wilby and Dessai (2010), the whole process could be defined as a "cascade of uncertainty", because uncertainty is propagated and enlarged from one step to the following one. In recent years, many authors have cast serious doubts on the adequacy of

climate-model driven approaches to tackle climate change adaptation (Stainforth et al., 2007; Koutsoyiannis et al., 2008; Blöschl and Montanari, 2010) and top-down combined with participatory-based “bottom-up” approaches have been proposed as an alternative (Bhave et al., 2013; Girard et al., 2015).

Nevertheless, climate models cannot be regarded as the only available source of climate information. Ray and Brown (2015) defined two types of climate scenarios: ex-ante scenarios (climate-model driven) and ex-post scenarios (independent of climate model outcomes). The later involve two steps: identification of historical climate conditions related to system problems and scenario generation through parametric or stochastic variation of the climate. As ex-post scenarios cover a wider range of climate conditions than ex-ante scenarios, they are more able to identify system susceptibility to failure. Therefore, they could be suitable to define CIRFs in the "inverse mode". This vulnerability-oriented (also called “bottom-up” or scenario-free) approach is being increasingly applied to address water management issues under climate change conditions (e.g. Cunderlik and Simonovic, 2004; Brown et al., 2012; Steinschneider et al., 2015; Poff et al, 2015; Soundharajan et al., 2016). On the one hand, this approach has the advantage of avoiding the credibility problems of GCM-based climate projections at the local scale (Brown and Wilby, 2012), expanding the analysis of climate threats to the system to a broader range of possible climate outcomes; on the other hand, when it comes to define the most likely future climate conditions, it still depends on the information generated by climate models.

The definition of inverse CIRFs in water resource systems could be challenging, since their ability to supply water demands is not only related to climate factors, but also to non-climate ones (such as the available infrastructure or management rules and constraints; Martin-Carrasco and Garrote, 2006). Highly regulated systems with large storage capacity could show greater inertia against climate variability: temporary meteorological droughts (low

precipitation periods) not always lead to water scarcity and, even if they do, there could be a significant time lag between the climate trigger and the impacts on water supply.

To simplify, the problem could be ultimately defined by two continuous, explanatory variables (precipitation and temperature) and a dependent dichotomous variable (system failure, defined as the inability to meet a certain performance criteria). Herein, we propose the use of a multivariable logistic regression (LR) model to define a CIRF for the inverse analysis of highly regulated water resource systems, using precipitation and temperature in the previous months as predictors of system failure. Once implemented, the LR model could be used as a screening model to identify climate change conditions that pose potential risks to the system, instead of the traditional model chain. Another advantage of this method is that it could be applied in any water resource system.

The selected case study and the climate series are described in Section 2. Section 3 explains in detail the selected methodology. To build the LR model, we generated synthetic streamflow time series through a Simulated Annealing algorithm (Section 3.3) and used them as inputs for a previously calibrated water management model (Section 3.1). Under the selected vulnerability criteria (Section 3.2), we were able to obtain the system response (coded as a binary variable) for each synthetic time series. The next step was relating synthetic streamflows to precipitation and temperature (Section 3.5). In Section 3.6, we characterized drought development stages through relative standardized indices, in order to identify the time span before the failure event in which precipitation and temperature conditions are relevant. Section 3.7 is devoted to LR model development and calibration, while Section 3.8 describes the validation process using climate change time series. Section 4 shows the main results for the selected approach. Finally, Section 5 discuss further research lines while Section 6 is devoted to the main conclusions.

2. MATERIALS

2.1 Case study description

The case study is the Jucar River Basin, a Mediterranean basin of 22,261 km^2 in Eastern Spain (Fig. 1). The basin shows two main climate zones, continental in the upper part and Mediterranean in the coastal area. During the period from 1980/81 to 2011/12, the mean annual precipitation was 475.2 mm/year (495.5 mm in 1940/41 to 2011/12), while mean annual temperature was 14.2^oC (13.8^oC in 1940/41 to 2011/12) (Jucar River Basin Authority (CHJ), 2015). The system is highly regulated through 3 main reservoirs: Alarcon and Contreras (located in parallel in the upper basin) and Tous, downstream. Regarding to groundwater, the main water body is La Mancha Oriental, one of the most extensive carbonate aquifers in Southern Europe (7,260 million m^3) and hydraulically connected to the Jucar river. Agriculture is the most prominent water user in the basin (with a share of about 80% of the total demand), and its principal withdrawals (Agricultural Demand Units, ADUs) are located in the lower part basin (except for groundwater irrigation in Mancha Oriental). Main urban demands (Urban Demand Units, UDUs) correspond to the cities of Valencia, Albacete and Sagunto.

Water availability has historically been a main issue in the region, where a frail equilibrium between resources and demands already exists: 1713.4 million m^3 of average inflow in the period 1980/81 to 2011/12 against a total demand of 1648.4 million m^3 (CHJ, 2015). Drought events are relatively frequent and, when they develop into water scarcity, significant economic, social and environmental impacts arise (CHJ, 2018). Besides, in a climate change context water resources in the basin could experiment an important decrease (Chirivella Osma et al., 2015; Marcos-Garcia and Pulido-Velazquez, 2017), along with an increase in drought frequency, magnitude and intensity (Marcos-Garcia et al., 2017; Escriva-Bou et al., 2017) due to the combined effects of rainfall reduction and evapotranspiration increase. Using a hydro-economic approach, Escriva-Bou et al. (2017) showed that the system is very vulnerable to

climate and land use change, especially over the mid and long term, and that innovative adaptation actions can significantly reduce the potential economic losses.

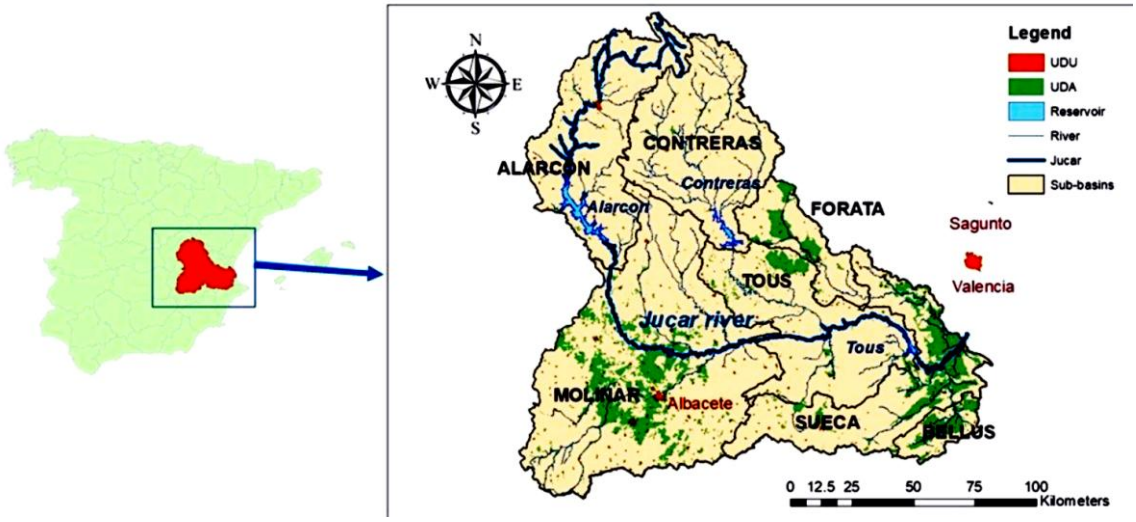
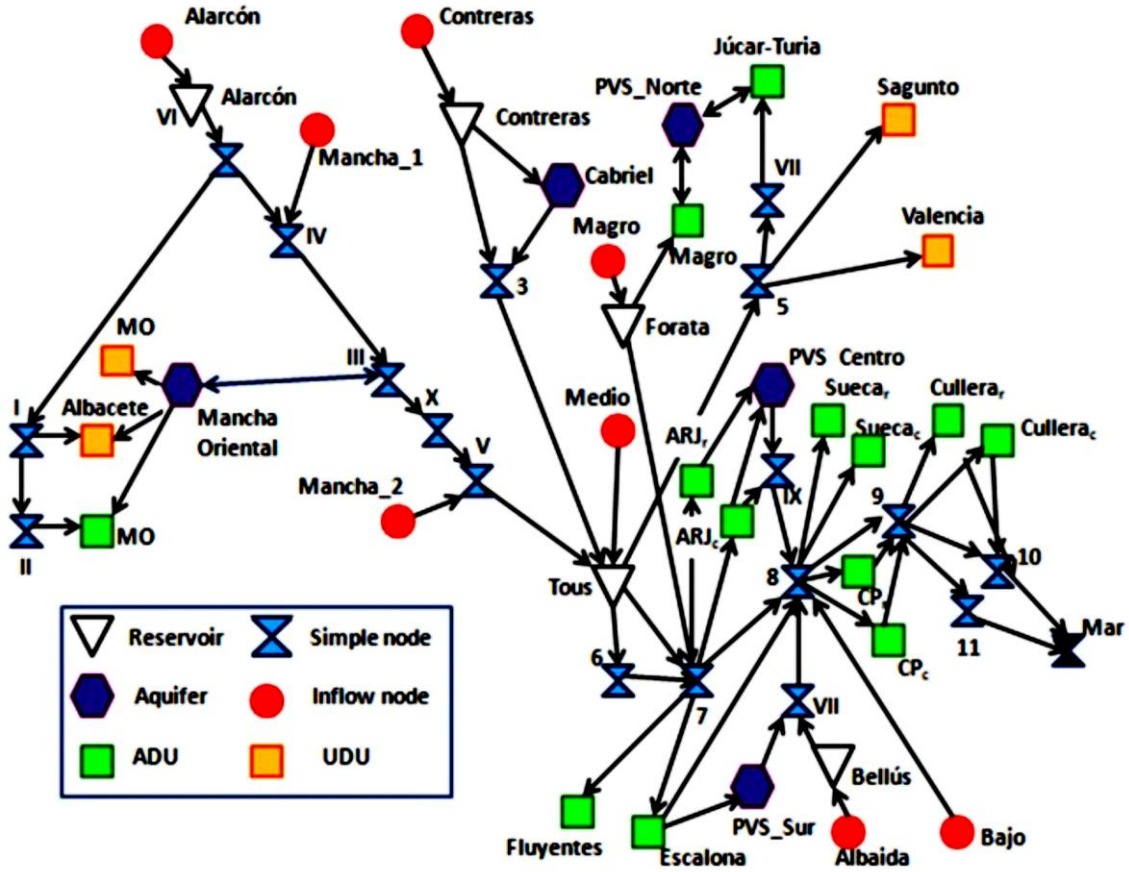


Figure 1. Location of the Jucar basin (left) and sub-basins and main demands (right)

For the representation of the main elements the water resource management model of the system (WRSM model), we used a classic flow network approach with nodes and links. Several types of nodes were considered: 8 inflow nodes (which match the sub-basin division of Figure 1, although Molinar inflow is split in two nodes), 5 surface reservoir nodes (Alarcon, Contreras, Tous, Forata and Bellus), 5 aquifer nodes (La Mancha Oriental aquifer, Cabriel aquifer and 3 nodes corresponding to La Plana de Valencia aquifer), 18 junction nodes, 4 urban demand nodes or UDUs (Albacete, La Mancha Oriental, Valencia and Sagunto) and 13 agricultural demand nodes or ADUs (La Mancha Oriental, Magro, Jucar-Turia, Flowing, Acequia Real, Cuatro Pueblos, Cullera, Escalona and Sueca). The agricultural demands of Acequia Real, Cuatro Pueblos, Cullera and Sueca have 2 corresponding nodes each, considering the two main crop types: citrus fruits and rice (Figure 4):



140

141 *Figure 2. Water management system scheme*

142 According to the Spanish Water Law, urban supply has the highest priority in water system
143 management, followed by irrigation demand (MMA, 2001). Environment is not considered as a
144 user, but as a restriction to water allocation. Therefore, in case of water scarcity, agricultural
145 demands are going to suffer a water shortage before urban water supply is concerned.
146 Besides, in the Jucar basin there are also priorities among agricultural users due to the Alarcon
147 Reservoir Agreement, which establishes a reserve curve in favor of Acequia Real, Cuatro
148 Pueblos, Cullera, Escalona, Sueca and Flowing irrigation demands. This reserve curve and other
149 system operating rules (integrated in the management model presented before) are further
150 described in Macian-Sorribes et al. (2017).

151 ***2.2 Observed data***

To characterize the climate variables (precipitation and temperature) for the selected historical period (1980-2012), we used three sources of information: 1) SPAIN02 v4 dataset (Herrera et al., 2010), which provides daily time series with high spatial resolution (0.11^{o}) for peninsular Spain from 1971 to 2007; 2) ECA&D dataset (Haylock et al., 2008), which supplies daily time series at the European level from 1950 to the current date, with a spatial resolution of 0.25^{o} and; 3) meteorological records from the State Meteorological Agency (AEMET).

Regarding the inflows for the period 1980-2012, monthly naturalized time series were provided by the Jucar river basin authority (CHJ) for each of the 7 sub-basins in which the case study is divided (Figure 1). The CHJ also provided reservoir storage time series, which were used to calibrate the water management model for the period 2003-2012. Finally, water demands were extracted from the 2009-2015 basin plan (CHJ, 2014).

2.3 Climate change data

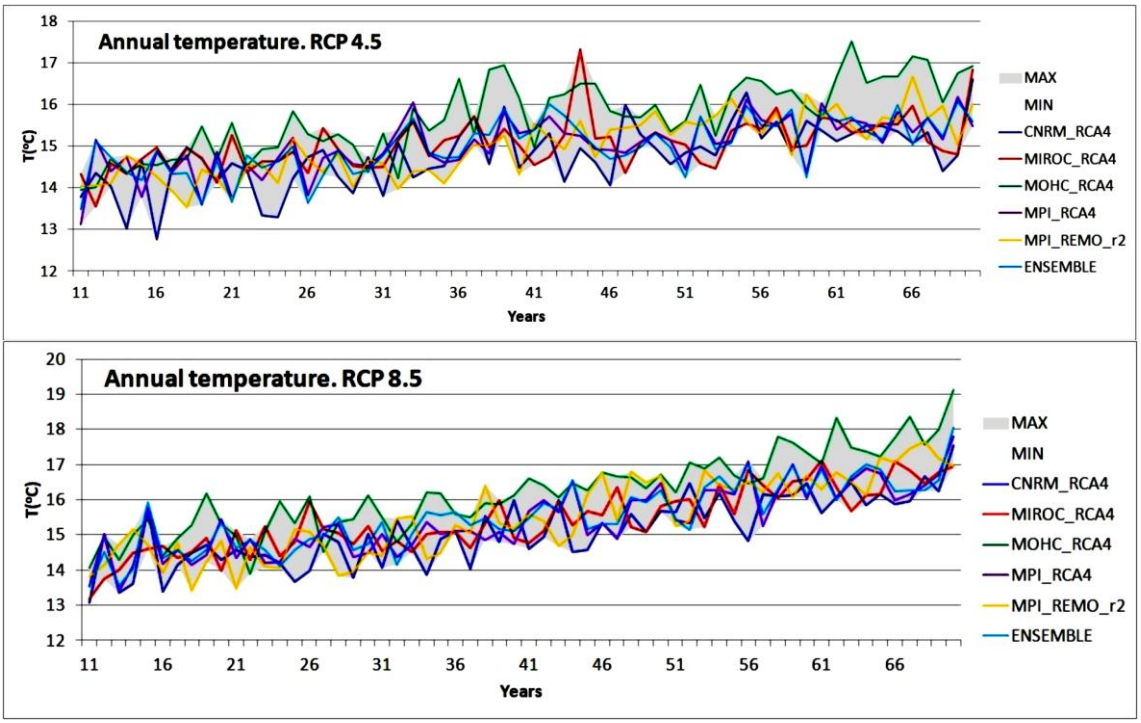
Figure 2 represents the evolution of the precipitation and temperature from 2011 to 2070, predicted by 6 climate model combinations and two emissions scenarios (RCPs 4.5 and 8.5). The selected combinations of GCMs and RCMs are: CNRM-CERFACS-CNRM-CM5/SMHI-RCA4 (CNRM_RCA4), MIROC-MIROC5/SMHI-RCA4 (MIROC_RCA4), MOHC-HadGEM2-ES/SMHI-RCA4 (MOHC_RCA4), MPI-M-MPI-ESM-LR/SMHI-RCA4 (MPI_RCA4), MPI-M-MPI-ESM-LR/MPI-CSC-REMO2009 (MPI_REMO_r2) and an ensemble (ENSEMBLE) of several climate models. These data at the monthly scale were obtained from the research presented in Marcos-Garcia and Pulido-Velazquez (2017).

Table 1 shows the minimum, mean and maximum annual temperature and precipitation in the Jucar basin for the periods 1980-2012, 2011-2040 and 2041-2070. Regarding temperature, all statistics increase in the RCP scenarios in comparison to the observed period. This increase is larger in the midterm than in the short term, and for the RCP 8.5 scenario.

Table 1. Annual temperature (T) and precipitation (P) in the Jucar basin

Variable	Statistic	1980-2012 Observed	2011-2040		2041-2070	
			RCP 4.5	RCP 8.5	RCP 4.5	RCP 8.5
T (°C)	Minimum	13.00	13.78	13.53	14.95	15.45
	Mean	14.33	14.67	14.78	15.50	16.27
	Maximum	15.27	15.88	15.59	16.24	17.74
P (mm)	Minimum	350	308	242	267	246
	Mean	521	541	536	507	478
	Maximum	798	792	843	819	758

Regarding precipitation, there is a decrease of the minimum annual precipitation in all the future scenarios. This decrease is larger in the midterm than in the short term, and for the RCP 8.5 scenario. The mean annual precipitation could increase slightly in the short term, although it decreases in the midterm. Finally, some of the scenarios show an increase of the maximum annual precipitation.



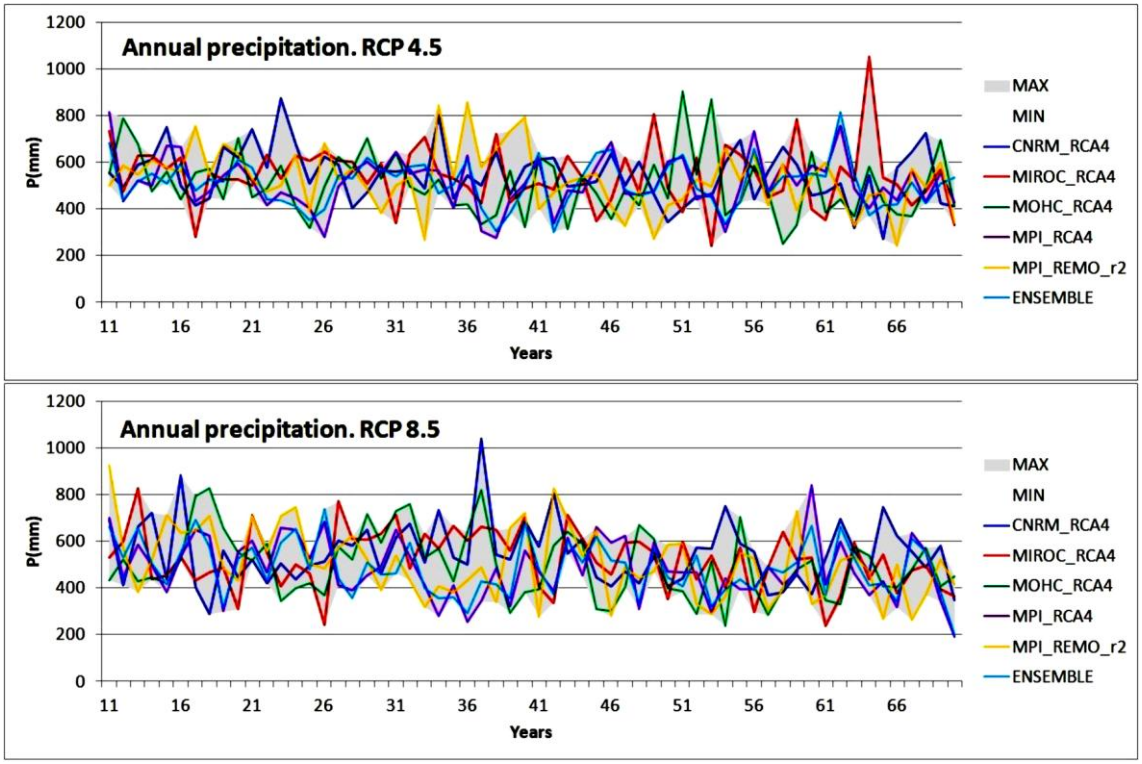


Figure 3. Annual precipitation and temperature from 2011 to 2070 for RCP 4.5 and 8.5

3. METHODS

The selected approach (Figure 3) involves 8 main steps, further described in the following sections:

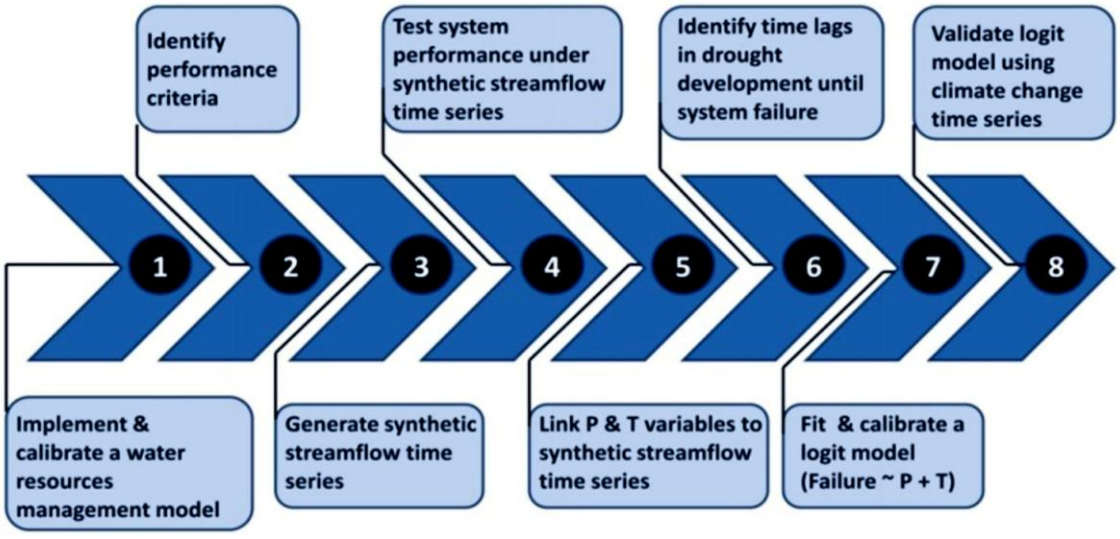


Figure 4. Overall approach scheme

3.1 Implementation and calibration of the water resource system management model

The system network was implemented using the Hydra Modeller software (build on the foundation of Hydra Platform; Meier et al., 2014), and then exported to General Algebraic Modeling System (GAMS; GAMS Development Corporation, 2013) format. Using GAMS software, a monthly simulation model of the system was built, considering environmental restrictions, water allocation rules under Spanish legal framework, and current agreements. The model was calibrated for the period 2003-2012, based on the observed storage time series of the three main reservoirs (Alarcon, Contreras and Tous) and the historical releases from Tous reservoir to supply the downstream demands. Due to the selected calibration period, the demands are those considered in CHJ (2014).

3.2 Performance assessment and failure criteria

In Spain, the Water Planning Act (MARM, 2008) introduces the legal definition of system failure in relation to agricultural water uses through 3 criteria: the 1-year criterion, when the annual deficit is greater than 50% irrigation demand; the 2-year criterion, when the deficit in 2 consecutive years is greater than 75% irrigation demand) and; and the 10-year criterion, when the deficit in 10 consecutive years is greater than 100% irrigation demand). For the purpose of the present work, only the 1-year criterion has been considered.

3.3 Generation of synthetic streamflow time series

The approach selected for this step was the one proposed by Borgomeo et al. (2015). This method involves obtaining a random sampling from observed streamflow records and then swapping the values of this sampling through the Simulated Annealing algorithm (SA, Figure 5). The shuffling stops when the generated series matches the hydrological properties imposed by the objective function. Besides, the objective function can be easily altered to generate

streamflow time series with specific characteristics (e.g. monthly mean decrease, standard deviation increase) to test system vulnerability.

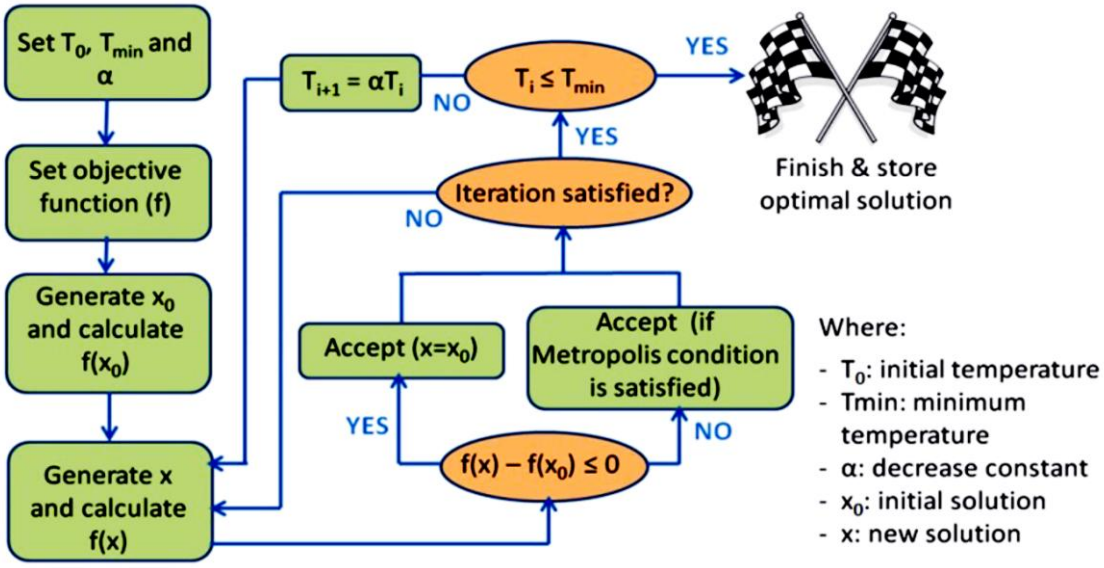


Figure 5. Simulated Annealing algorithm

The selected hydrological properties of the observed time series (1980-2012) whose statistics we aim to reproduce were: 1) monthly mean (M); 2) monthly standard deviation (SD); 3) quantile 90% (Q); 4) monthly temporal correlation (auto-correlation, AC) and; 5) Pearson correlation coefficient between sub-basins (spatial cross-correlation of sub-basin "i" with sub-basin "j", PC_{ij}). Here, it should be noted that spatial correlation was not considered in the precedent work by Borgomeo et al., (2015). Eq. 1 shows the objective function (F) selected to generate the synthetic streamflow time series, where a, b, c are coefficients (e.g. a=0.7 would mean a decrease of 30% in the monthly mean). Sub-index "0" is related to the observed time series and sub-index "s" to the synthetic one:

Eq. 1 $$\boldsymbol{F = a(M_0 - M_s) + b(SD_0 - SD_s) + c(Q_0 - Q_s) + (AC_0 - AC_s) + \sum_{j=1}^{n}(PC_{ij_0} - PC_{ij_s})}$$

Here it should be noted that statistics values are previously normalized to have the same order of magnitude.

3.4 Testing of the system performance under synthetic streamflow time series

Using the SA, 210 synthetic streamflow time series of 30 years time span were generated (30 for each of the 7 sub-basins). We performed 30 simulations using the new time series as inputs for the water management model described in Section 3.1. The existence of failure was coded as 1 when the 1-year criterion is met for at least one of the agricultural demands described in Section 3.2, and 0 in any other case.

3.5 Linking precipitation and temperature variables to synthetic streamflow time series

To link climate conditions to the synthetic streamflow time series, in first place we explored the temporal correlation between precipitation and streamflow for the observed time series (1980-2012) at the monthly scale. Table 2 shows that the highest correlations are found for time lag 0, except for the headwaters basins (Alarcon and Contreras), where streamflow in a particular month is slightly more correlated with the precipitation in the previous month (time lag 1).

Table 2. Correlation coefficients between precipitation and inflows for several time lags

Lag (month)	Alarcon	Contreras	Molinar	Tous	Forata	Sueca	Bellus
0	0.50	0.45	**0.26**	**0.38**	**0.66**	**0.65**	**0.74**
1	**0.58**	**0.50**	0.10	0.17	0.30	0.31	0.24
2	0.40	0.33	0.07	0.13	0.13	0.17	0.17
3	0.21	0.22	0.03	0.10	0.08	0.08	0.08

For each of 210 synthetic streamflow time series, we obtained two vectors of precipitation and temperature replicating the transformation that SA did (Figure 6). We considered a time lag 0 between precipitation and temperature conditions and streamflow production for each month.

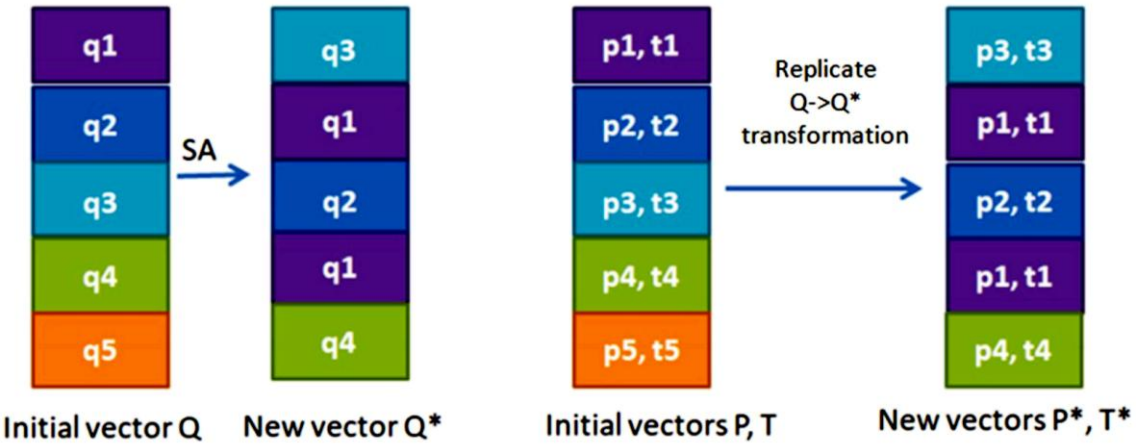


Figure 6. Procedure to link precipitation (P) and temperature (T) to inflows (Q)

3.6 Characterization of the time lag between drought development and system failure

In highly regulated water resource system such as the Jucar basin, meteorological and hydrological droughts can begin months and even years before they cause a system failure. Therefore, characterizing the time lag between a meteorological/hydrological drought onset and the system failure could be important to identify previous climate sequences that head the system to fail. Villalobos (2007) studied the three main stages of drought development (meteorological, hydrological and operational drought) in the Jucar basin, applying standardized indices to precipitation, inflows and reservoir storage. He concluded that the Standardized Precipitation Index (SPI; McKee et al., 1993) has certain predictive capacity regarding operational droughts, when aggregation periods of 12 and 24 months were considered. According to his results, if the 24 months aggregated SPI identifies a meteorological drought, an operational drought is likely to appear 18 months later.

Here, we use two standardized drought indices: the Standardized Precipitation & Evapotranspiration Index (SPEI, Vicente-Serrano et al., 2010) to identify meteorological droughts and the Standardized Streamflow Index (SSI) for the hydrological ones. We fitted a Log-Logistic distribution to effective precipitation (precipitation minus potential evapotranspiration, PET) and a Log-Normal distribution streamflow time series, considering

the observed values during the period 1980-2012 and an aggregation time window of 12 months (Villalobos, 2007; Marcos-Garcia et al., 2017). PET was computed using the Thornthwaite method (Thornthwaite, 1948). Finally, we applied the same distributions to the synthetic time series, in order to obtain relative standardized indices (Dubrovsky et al., 2009; Marcos-Garcia et al., 2017).

3.7 Fitting and calibrating a logistic regression model

A LR model is able to describe the relationship between a binary variable (which is the response or dependent variable) and a set of continuous, independent variables (predictors or explanatory variables). In our case, the binary variable is the existence or the absence of system failure (coded as 1 and 0, respectively) and the explanatory variables are precipitation and temperature. Eq. 2 describes the general equation of the LR model and Eq. 3 the probability of system failure for a particular year:

Eq. 2 $$Logit(p) = b_0 + b_1 P + b_2 T$$

Eq. 3 $$p = \frac{\exp\,(b_0 + b_1 P + b_2 T)}{1 + \exp\,(b_0 + b_1 P + b_2 T)}$$

Where:

p: probability of system failure for a particular year

b_0, b_1, b_2: regression coefficients

P: mean annual precipitation during a selected time period before the year under analysis

T: mean annual temperature during a selected time period before the year under analysis

The selection of a time period to compute P and T must be relevant for the system and it is based on the assessment described in Section 3.6. According to the results we obtained (Section 4.3), four LR models were fitted computing P and T values for 2, 3, 4 and 5 years. The yearly time series of the response and the explanatory variables have been previously

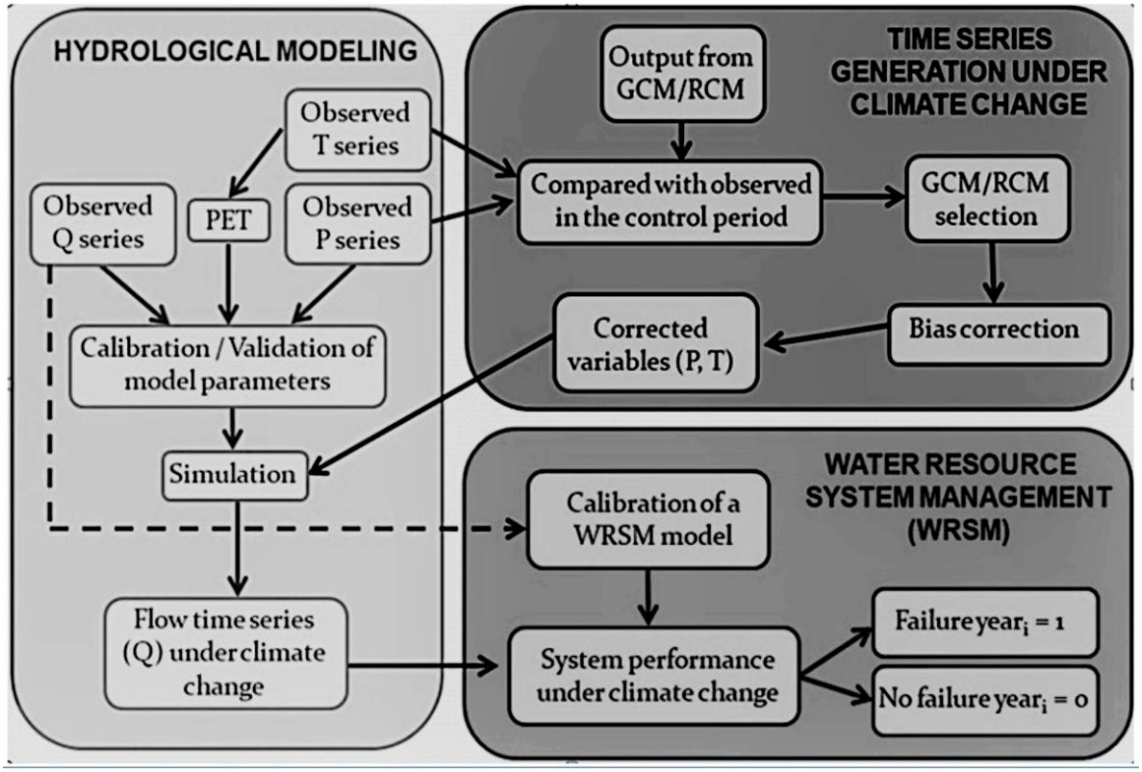
obtained as described in Sections 3.4 and 3.5. To limit the influence of the system's initial conditions (e.g. reservoir and aquifer volume), we removed the four first years of the computed variables before using them to build the LR model. The explanation is the following: if the reservoirs are full at the beginning of the period, the system will not fail to fully meet the demands even if the precipitation in the first four years is really scarce, so the absence of failure in those years could not be related to the climate conditions but to the influence of the system's initial conditions.

A LR model could be considered as well calibrated if the predicted probabilities match the observed proportions of the response (Nattino et al., 2017). To assess the goodness of fit of the LR models, three approaches were selected: 1) Hosmer-Lemeshow test (Hosmer and Lemeshow, 1980); 2) pseudo R-squared measures (Cox-Snell (Cox and Snell, 1989), McFadden (McFadden, 1974), Nagelkerke (Nagelkerke, 1991) and Tjur (Tjur, 2009); 3) GiViTI calibration belt (Nattino et al., 2016; Nattino et al., 2017).

Finally, we compare the goodness of fit results with the outcome of a stepwise logistic regression (performed through the stepAIC function of the MASS package in R (Venables and Ripley, 2002), where the possible variables to select are the mean annual precipitation in the previous 2, 3, 4 and 5 years.

3.8 Validation of the LR model using climate change time series

The validation of the LR model was done through climate change time series using the traditional top-down approach (Fig. 7). This approach involves the use of a chain of models: a GCM is downscaled to obtain a RCM. The bias of the climate variables from the RCM are corrected to match the observed values during the control period, and next they are used as inputs for a hydrological model. Finally, system behavior under climate change conditions is simulated using a water management model.



316

317 *Figure 7. Top-down approach to characterize system failure under climate change*

318 Precipitation, temperature and inflow time series under climate change were generated by
319 Marcos-Garcia and Pulido-Velazquez (2017) for the Jucar basin, using several combinations
320 and an ensemble of GCM-RCM models, two climate change scenarios (RCP 4.5 and 8.5) in the
321 short term (2011-2040) and in the mid-term (2041-2070) and the Temez hydrological model
322 (Temez, 1977). Here, we used 6 of these inflow time series as inputs to the WRSM model
323 described in Section 3.1. To characterize the existence/absence of failure, the vulnerability
324 criteria and codification were the same that the ones explained in Section 3.2 and 3.4,
325 respectively. Besides, we obtained the mean annual values of P and T for 3 and 4 years periods
326 (see Section 4.4).

327 We used the calibrated "3 years" and "4 years" LR models to predict the failure probability
328 regarding the P and T time series obtained from a particular climate model, and compared this
329 probability with the real incidence (coded as 0 and 1) from the WRSM model simulation for the

same climate model. As explained in Section 2.3, part of the climate conditions of the validation time series were outside the range of values of the calibration period..

4. RESULTS

4.1 Water resource system management model

Figure 8 plots the evolution of water storage in the main reservoirs (Alarcon, Contreras and Tous) and the releases from Tous reservoir during the calibration period (2003-2012), for the observed (Obs., blue line) and the simulated values (Sim., red dotted line). R^2 values range between 0.97 and 0.74, so it is possible to conclude that the model is able to accurately reproduce the system operating rules during the calibration period.

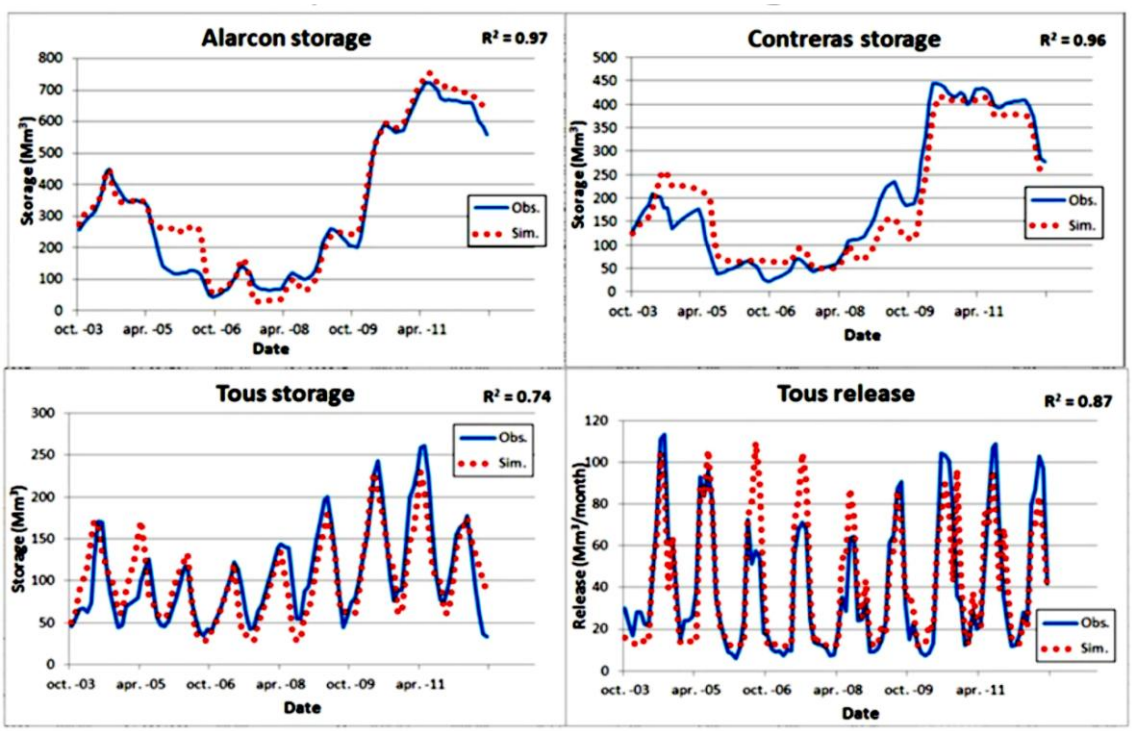


Figure 8. Goodness of fit of the WRSM model (calibration period)

4.2 Synthetic streamflow time series

Figure 9 compares the monthly mean and standard deviation of the observed streamflow time series with the same statistics of one of the synthetic time series for Alarcon sub-basin.

According to the graphs, the SA algorithm is able to reproduce the monthly structure of both statistics, at the same time that it decreases the mean and increases the standard deviation (see Section 3.3). Nevertheless, monthly mean decrease is less than the 30% reduction specified in the objective function, as the SA is limited by the available values in the sampling.

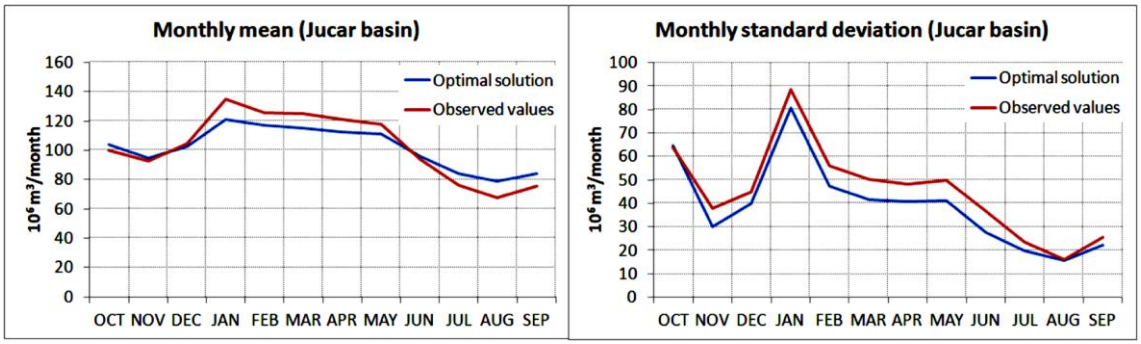


Figure 9. Monthly mean and standard deviation for observed and synthetic time series

Figure 10 plots the autocorrelation function for the observed and the synthetic time series, after deseasonalizing them. The SA algorithm is not only able to reproduce the temporal autocorrelation, but also the spatial correlation between sub-basins. Although Pearson correlation coefficient between Alarcon and Contreras sub-basins (0.60) is lesser than the one obtained from the observed time series (0.86), it is necessary to take into account that the SA looks for a trade-off between the hydrological properties specified in the objective function. If we were more interested in the spatial correlation than in other characteristics, we could always increase the weight of this term to the detriment of the others.

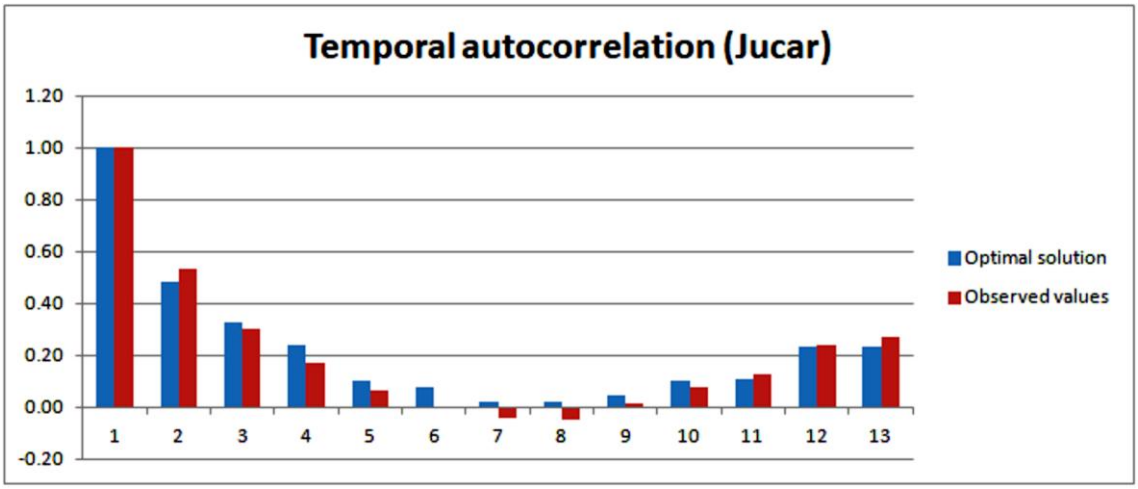


Figure 10. Temporal correlation in the synthetic time series

4.3 Time lag between drought development and system failure

Figure 11 a) shows the drought indices evolution for the period 2000-2011 of the historical time series. Meteorological and hydrological drought started 30 and 29 months before the first system failure (in 2005), respectively. Therefore, while a meteorological drought developed quickly into a hydrological drought, it took 2.5 years to result in a system failure. The same analysis has been done for each of the synthetic time series (Figure 11 b and c), considering only the time lag between the meteorological/hydrological drought and the first failure if there were several consecutive ones. For a total amount of 44 failure events, the time lag oscillated between 24 month (2 years) and 60 months (5 years), being the mean value 40 months. Thus, to cover the entire range, we built four LR models considering the precipitation and temperature conditions in the 2, 3, 4 and 5 previous years (Section 3.7).

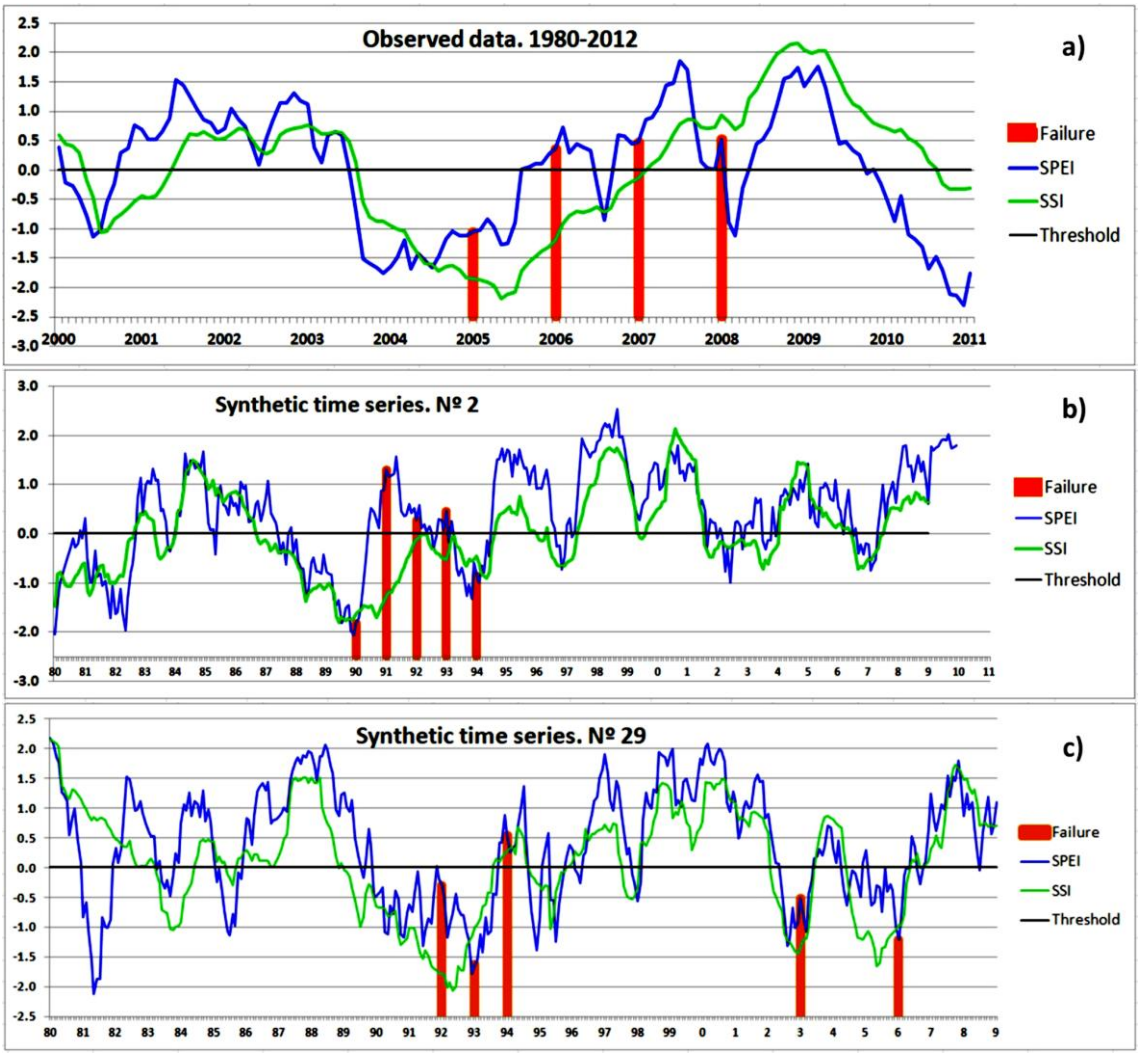
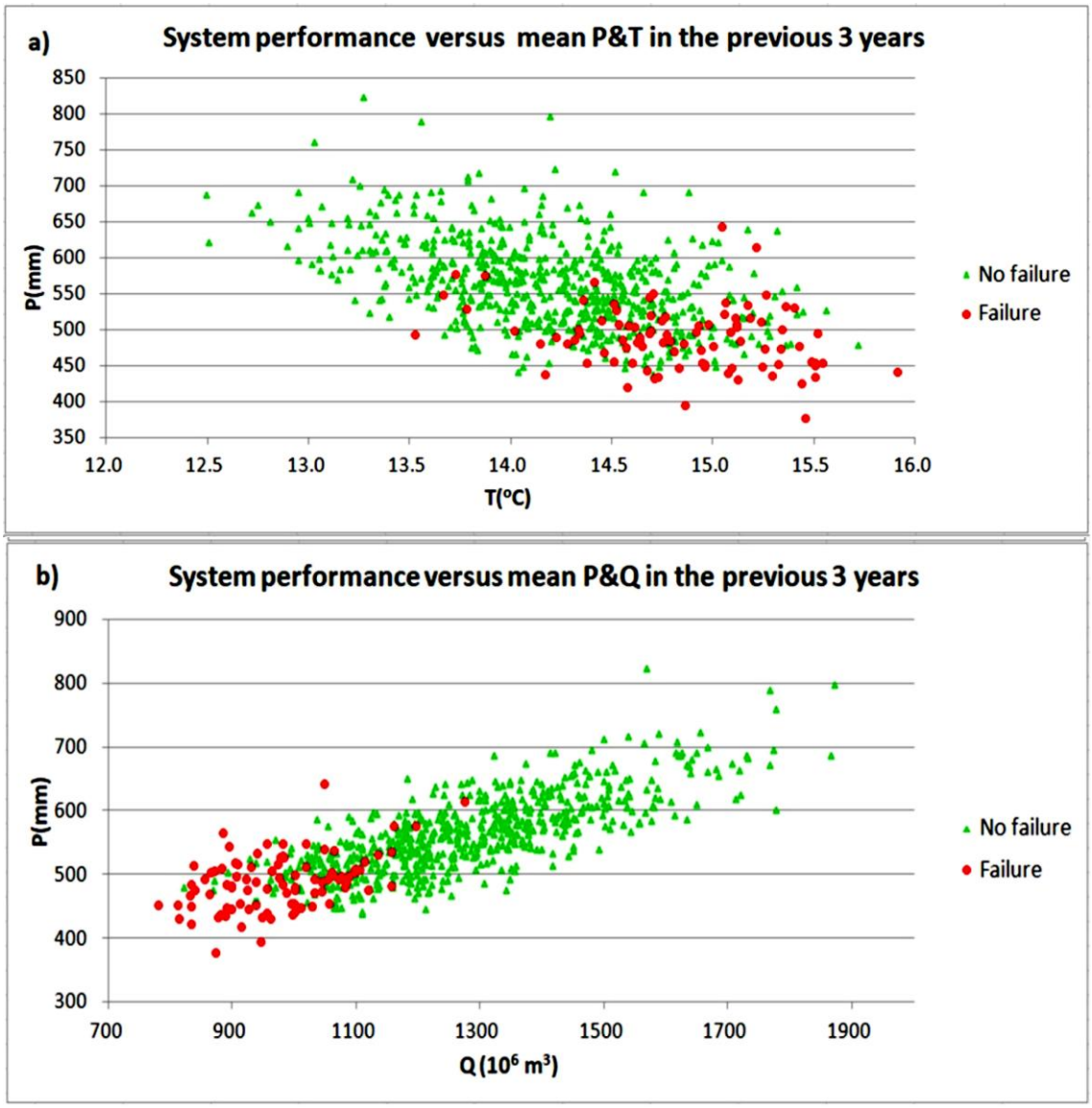


Figure 11. SPEI and SSI evolution and system failures

4.4 System performance under synthetic streamflow time series

Figure 12 shows the system performance regarding the mean annual precipitation (P), temperature (T) and inflow (Q) of the previous 3 years. The average annual values of P, T and Q for the failure events are 14.82°C, 490 mm and 974 million m^3, while for the absence of failure is 14.18°C, 566 mm and 1280 million m^3.



380

381

382 *Figure 12. System performance versus mean annual P, T and Q in the previous 3 years*

383 ***4.5 Goodness of fit of the LR model (calibration)***

384 Table 3 shows the outcome of the Hosmer-Lemeshow test (H-M test), the Cox-Snell (C-S),
385 McFadden (M-F), Nagelkerke (N) and Tjur (T) pseudo R^2 values and the p-value for the GiViTI
386 calibration belt (GCB).

387 *Table 3. Parameters and goodness of fit results of the LR models*

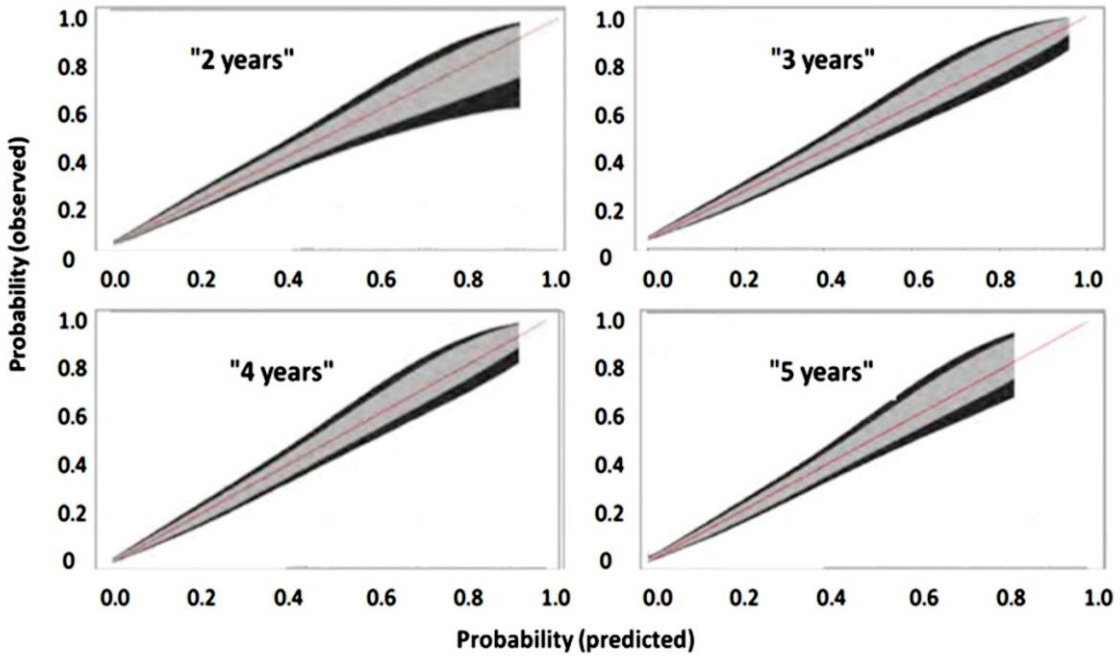
Parameters	H-M test*	Pseudo R^2	GCB

Model	b_0	b_1	b_2	χ^2	p-value	C-S	MF	N	T	p-value
2 yr	-6.743	-0.018	0.986	58.896	0.000	0.18	0.25	0.33	0.22	0.935
3 yr	-12.238	-0.025	1.613	17.467	0.026	0.22	0.33	0.41	0.30	0.476
4 yr	-19.297	-0.025	2.114	5.623	0.689	0.23	0.33	0.42	0.31	0.501
5 yr	-20.727	-0.020	2.034	13.231	0.104	0.19	0.26	0.34	0.25	0.521

*number of quartiles of risk = 10

According to the Hosmer-Lemeshow test and considering a confidence level of 99% (p=0.01), we only can reject the null hypothesis for model "2 years". Besides, the highest values of the four pseudo R2 are shown by models "3 years" and "4 years".

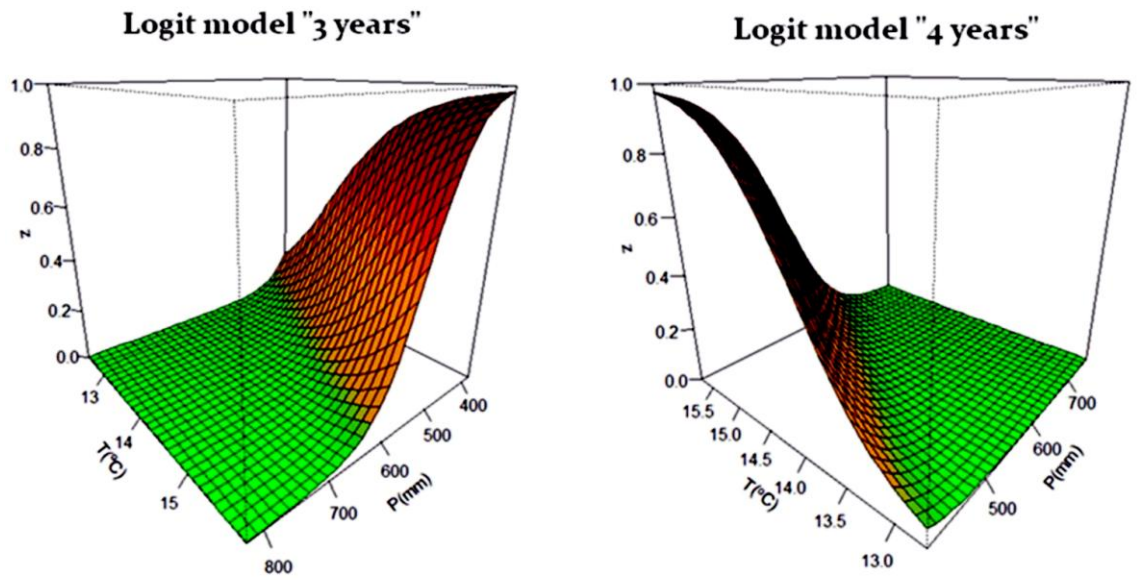
Regarding the GiviTI calibration belt (Figure 13), no evidence of lack of calibration emerges from any of the four models, because the belt encompasses the bisector in the whole 0-1 range (although is not defined for some probability values between 0.8 and 1, mainly for the models "2 years" and "5 years"). The p-values for the four models suggests that the calibration of the models is acceptable. Finally, it should be noted that the bands convey the uncertainty in the estimated relationship between predictions and the probabilities of the true response (Nattino et al., 2017). Therefore, model "2 years" shows the greatest uncertainty, while "3 years" and "4 years" models' could be considered less uncertain.



400

401 *Figure 13. GiViTI calibration belts*

402 The discussed goodness of fit measures for models "3 years" and "4 years" were coherent with
403 the average time lag identified in Section 4.3 (36 months < 40 months < 48 months). Therefore,
404 we selected the models "3 years" and "4 years for the validation stage. Moreover, the model
405 selection procedure exposed in Sections 3.6 and 3.7 could be substituted by a stepwise logistic
406 regression, which was able to identify the mean annual precipitation during the previous 3 and
407 4 years as the most suitable predictors. Figure 14 plot the relationship between P, T and failure
408 probability (z) of both models.



409

410 *Figure 14. 3D plot relating P, T and failure probability for the calibrated LR models*

411 ***4.6 Prediction performance of the LR model (validation)***

412 Table 4 shows the results of the GiViTI calibration belt test for the LR models "3 years" and "4
413 years", regarding the validation time series obtained from the climate change models.
414 Considering a confidence level of 99% (p=0.01), the obtained p-values suggest that both "3
415 years" and "4 years" models are able to predict the system failures in the case of CNRM_RCA4,
416 MIROC_RCA4 and MPI_REMO_r2 data. Besides, the "3 years" model shows also predictive
417 ability for MPI_RCA4 data (p-value larger than 0.01). Nevertheless, both "3 years" and "4
418 years" models show p-values lesser than 0.01 for ENSEMBLE and MOHC_RCA4 data, so it is not
419 possible to affirm that they show a good predictive ability in those cases.

420 *Table 4. Goodness of fit for the validation time series*

Model	CNRM_RCA4	ENSEMBLE	MIROC_RCA4	MOHC_RCA4	MPI_RCA4	MPI_REMO_r2
"3 years"	0.911	0.005	0.049	0.005	0.142	0.134
"4 years"	0.11	<0.001	0.457	<0.001	0.006	0.226

421

Figure 15 plots the GiViTI calibration belts for both LR models and the ENSEMBLE climate change data. These plots have an easy interpretation, when there is part of the confidence band under the bisector, the model is overestimating failure probability, and when the band is over the bisector, it is underestimating it. For example, regarding the ENSEMBLE data and the 99% confidence level, failure probability is overestimated for high probabilities (approximately greater than 0.97 for the "3 years" model and 0.75 for the "4 years" one). If we take into account the 95% confidence level for the same data, the models are also underestimating failure probability for low probabilities (less than 0.10 for the "3 years" model and 0.09 for the "4 years" one. Validation results regarding the rest of climate change data is included as Supplementary Material.

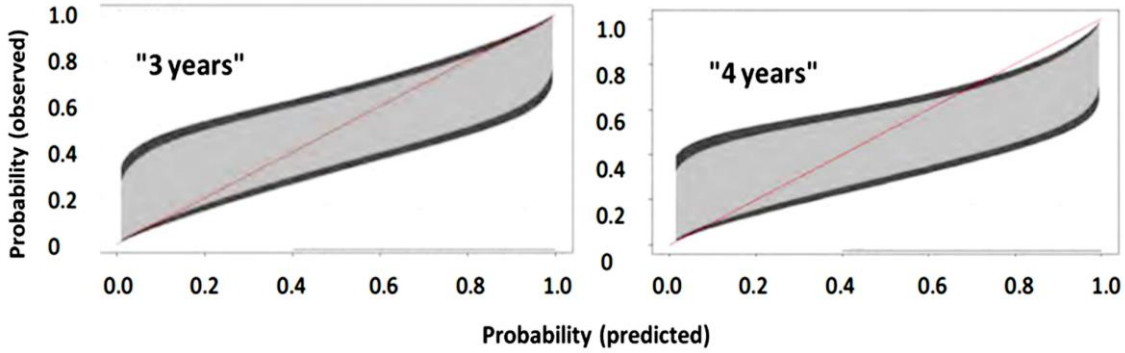
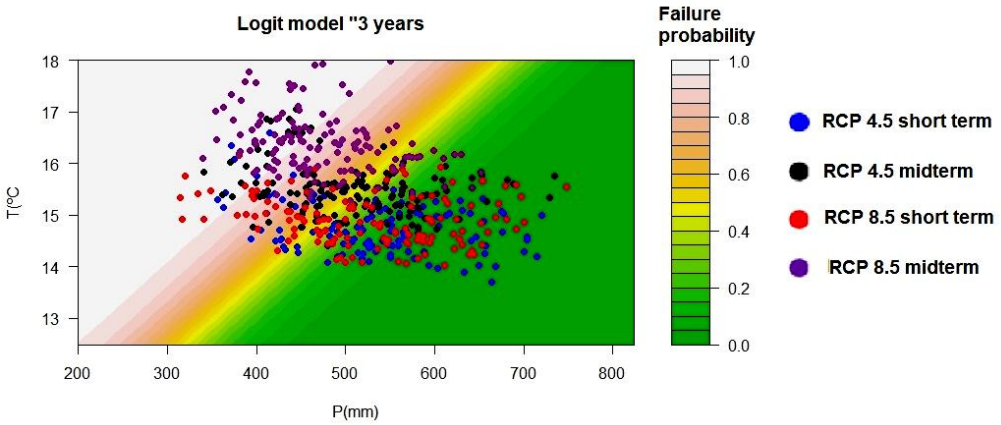


Figure 15. GiViTI calibration belt for ENSEMBLE

Figure 16 plots the climate response map derived from the LR model "3 years" and the mean annual precipitation and temperature for the historical data (from 1980 to 2012) and the climate change scenarios (RCP 4.5 and 8.5 in the short and midterm). According to our results, average failure probabilities are 0.13 for the historical period; 0.27 for RCP 4.5 in the short term; 0.32 for RCP 8.5 in the short term; 0.50 for RCP 4.5 in the midterm and; 0.80 for RCP 8.5 in the midterm. Therefore, failure probability is higher in the midterm than in the short term, especially for RCP 8.5, with the largest changes in mean annual temperature (increase) and mean annual precipitation (decrease) (see Section 2.3).



442

443 *Figure 16. Climate response map and climate models' output*

444 **5. DISCUSSION**

445 CIRFs could be used to assess the risks of a system under climate change, as well as the
446 magnitude and possibility of potential adaptation efforts (Füssel et al., 2003). Therefore, CIRFs
447 are valuable tools to help the decision making process, since they allow to make informed
448 judgments about potential impacts of climate change and the value of different adaptation
449 strategies.

450 Here, we have defined a CIRF through a LR model, using precipitation and temperature in the
451 previous years as predictors of system failure. The proposed method involves several steps:
452 implementation of a water resource management (WRSM) model for assessing the
453 performance of the system for different scenarios; generation and validation of synthetic time
454 series to be used as inputs for the WRSM model; definition of vulnerability criteria and;
455 calibration and validation of a LR model to link the system response to the climatic explanatory
456 variables (precipitation and temperature). In the case of a highly regulated water resource
457 system, this paper has shown CIRF to be a suitable tool to identify climate conditions that lead
458 the system to fail in meeting certain pre-established performance conditions. Results show
459 that the probability of system failure is higher in the midterm than in the short term, especially

460 for RCP 8.5 (which presents the largest increase in mean annual temperature and the largest
461 decrease in mean annual precipitation).

462 Characterizing a system response to climate involves finding a function $f(y|x)$ which
463 approximates the value of the dependent variable y (without knowing the probability density
464 $p(y|x)$), for any value of the predictive variable x. Models such as logistic regression (LR) and
465 Classification and Regression Trees (CART) are able to approximate such functions.

466 However, sometimes it could be more interesting to identify regions in the input space
467 associated with a very high (or low) value of y. For example, the "Patient" Rule Induction
468 Method (PRIM) proposed by Friedman and Fisher (1999) directly seeks these regions and skips
469 the finding of $f(y|x)$. Regarding climate scenarios, PRIM has been widely used in scenario
470 discovery, to find clusters of future states to which the systems are vulnerable (Lempert et al.,
471 2006; Groves and Lempert, 2007; Kwakkel and Cunningham, 2016), although it could present
472 some shortcomings, i.e.: 1) specific defects of high climbing optimization algorithms (local
473 optima, plateaus, ridges and valleys), (Kwakkel and Cunningham, 2016); 2) it could not be
474 suitable for all types of cluster shape and configuration within the multi-dimensional space of
475 futures (Lempert et al., 2006); 3) it strives when the uncertain factors are a mix of data types
476 (Kwakkel and Jaxa-Rozen, 2016). Here, we propose LR as a simple and fast approach to identify
477 climate scenarios that lead to system failure. This approach was also used by Kim et al. (2019)
478 for decision-centric assessment of climate change impacts on a complex river system.
479 However, further research on comparing LR, PRIM and CART performance is suggested.

480 Besides, the selected approach based on the generation of synthetic time series has also the
481 advantage of not requiring a calibrated hydrological model to "translate" climate variables into
482 inflows, because synthetic streamflow time series are first generated and then linked to
483 precipitation and temperature conditions. However, it is necessary to note that the intra-
484 annual behavior of precipitation and temperature is not necessary preserved using the same

485 resampling applied to the streamflow time series. This fact is not relevant for the purpose of
486 this study, because the synthetic precipitation and temperature time series are only used in
487 the LR model (which takes into account the mean annual values instead of monthly values),
488 but it could be significant for other applications. In addition, the procedure to generate the
489 climate variables may add uncertainty to the response function. In this regard, it could be
490 interesting to explore the use of a weather generator combined with a hydrological model for
491 building a new LR model, and assess if its predictive capacity improves in relation to the same
492 validation time series.

493 Once we identify the climate conditions in which the system is prone to fail, a possible step
494 forward could be to explore when these conditions are likely to take place, for example
495 through the link between them and teleconnections. In the case of the North Atlantic
496 Oscillation (NAO), it is well known that positive NAO phases are closely related to precipitation
497 amounts lower than normal in the Iberian peninsula, whereas negative phases are linked to
498 wetter conditions (Muñoz-Díaz and Rodrigo (2003); Trigo et al. (2004); Vicente-Serrano and
499 Cuadrat (2007); Queralt et al. (2009); Vicente-Serrano et al. (2009)). However, with respect to
500 temperature, Lopez-Moreno et al. (2011) observed a positive correlation in the European
501 Mediterranean area.

502 In addition, and regarding adaptation to climate change, this approach could be used to
503 explore how the system response changes after implementing adaptation measures (i.e.
504 decrease of failure probability in relation to the same climate conditions) and to relate
505 effectiveness and cost for the selection of a portfolio of adaptation options. Nevertheless,
506 climate change will not only have impacts on the supply but also on the demands, so it would
507 be interesting to develop further research lines which include both sides of the problem.
508 Finally, different CIRFs for each user could be used to assess the current allocation rules in a
509 context of great reduction of the available resources and formulate new ones.

6. CONCLUSIONS

A new method has been proposed for obtaining CIRFs in highly regulated systems. The method uses a LR model to relate climate variables to system failure. The main strength of finding a CIRF for a particular system is that it allows to identify climate change conditions which pose potential risks of system failure using a single model, instead of the traditional model chain in which uncertainty is propagated from one step to the following one. This approach could be applied to any water resource system, and it could be useful to explore how the system response changes after the implementation of adaptation measures.

ACKNOWLEDGEMENTS

This study has been supported by the IMPADAPT project (CGL2013-48424-C2-1-R), funded with Spanish MINECO (Ministerio de Economía y Competitividad) and European FEDER funds, and for the earlier ADAPTAMED project (RTI2018-101483-B-I00), funded by the Ministerio de Ciencia, Innovación y Universidades (MICINN) of Spain. Patricia Marcos-Garcia has been also supported by a FPI grant from the PhD Training Program (BES-2014-070490) of the former MINECO. The authors thank AEMET (Spanish Meteorological Office) and University of Cantabria for the data provided for this work (dataset Spain02).

REFERENCES

Bhave, A.G., Mishra, A., Raghuwanshi, N.S.. 2013. A combined bottom–up and top– down approach for assessment of climate change adaptation options. J. Hydrol.. Vol. 518, Part A, pp. 150-161, ISSN 0022-1694, https://doi.org/10.1016/j.jhydrol.2013.08.039.

Blöschl, G., Montanari, A. 2010. Climate change impacts - throwing the dice? Hydrological Processes. 24(3), 374-381. doi:10.1002/hyp.7574

Borgomeo, E., Farmer, C.L., Hall, J.W. 2015. Numerical rivers: a synthetic streamflow generator for water resources vulnerability assessments. Water Resources Research, Vol. 51, Issue 7, 5382-5405

534 Brown, C., Ghile, Y., Laverty, M., Li, K. 2012. Decision scaling: Linking bottom-up vulnerability analysis
535 with climate projections in the water sector. Water Resources Research, Vol. 48, W09537.
536 doi:10.1029/2011WR011212

537 Brown, C., Wilby, R. L. 2012. An alternate approach to assessing climate risks. Eos, Transactions
538 American Geophysical Union, 93(41), 401-402.

539 Chirivella Osma, V., Capilla Romá, J.E., Pérez Martín, M.A., 2015. Modelling regional impacts of climate
540 change on water resources: the Júcar basin (Spain). Hydrol. Sci. J. 60 (1), 30–
541 49.http://dx.doi.org/10.1080/02626667.2013.866711

542 CHJ. 2014. Plan Hidrológico de la Demarcación Hidrográfica del Júcar. Ciclo de planificación hidrológica
543 2009-2015 (in Spanish)

544 CHJ. 2015. Plan Hidrológico de la Demarcación Hidrográfica del Júcar. Ciclo de planificación hidrológica
545 2015-2021 (in Spanish)

546 CHJ. 2018. Plan Especial de Sequía. Demarcación Hidrográfica del Júcar. Memoria. (In Spanish)

547 Cox, D. R., and Snell, E. J. 1989. The analysis of binary data (2nd ed.). London: Chapman and Hall.

548 Cunderlik, J.M., Simonovic, S.P., 2004. Inverse Modeling of Water Resources Risk and Vulnerability to
549 Changing Climatic Conditions. Proceedings, 57th, Canadian Water Resources Association Annual
550 Congress: Water and Climate Change: Knowledge for Better Adaptation, CWRA, Montreal.

551 Dubrovsky, M., Svoboda, M.D., Trnka, M., Hayes, M.J., Wilhite, D.A., Zalud, Z., Hlavinka, P., 2009.
552 Application of relative drought indices in assessing climate change impacts on drought conditions in
553 Czechia. *Theor. Appl. Climatol*. (96), 155–171. http://dx.doi.org/10.1007/s00704-008-0020-x.

554 Escriva-Bou, A., Pulido-Velazquez, M., Pulido-Velazquez, D. 2017. The economic value of adaptive
555 strategies to global change for water management in Spain's Jucar Basin. J. Water Resources Planning
556 and Management 143(5), 1-1. doi: 10.1061/(ASCE)WR.1943-5452.0000735.

557 Friedman, J.H., Fisher, N.I. 1999. Bump hunting in high-dimensional data. Statistics and Computing 9,
558 123–143.

559 Füssel, H.M., Toth, F.L., Van Minnen, J.G., Kaspar, F. 2003. Climate Impact Response Functions as impact
560 tools in the tolerable windows approach. Climatic Change, 56: 91-117.

561 GAMS Development Corporation. 2013. General Algebraic Modeling System (GAMS). Release 24.2.1.
562 Washington, DC, USA.

563 Girard, C., Pulido-Velazquez, M., Rinaudo JD., Page, C., Caballero, Y., 2015. Integrating top-down and
564 bottom-up approaches to design global change adaptation at the river basin scale. Global Environmental
565 Change, 34 (132-146). doi:10.1016/j.gloenvcha.2015.07.002

566 Groves, D.G., Lempert, R.J. 2007. A new analytic method for finding policy-relevant scenarios. Global
567 Environmental Change 17, 73-85.

568 Haylock, M.R., Hofstra, N., Klein Tank A.M.G., Klok E.J., Jones P.D., New, M. 2008. A European
569 daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006.
570 Journal of Geophysical Research: Atmospheres, Vol. 113, Issue D20.
571 https://doi.org/10.1029/2008JD010201

572 Herrera, S., Gutiérrez, J.M., Ancell, R., Pons, M.R., Frías, M.D., Fernández, J. 2010. Development
573 and analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain (Spain 02).
574 Int. J. Climatol. http://dx.doi.org/10.1002/joc.2256/

575 Hosmer, D.W. and Lemeshow, S. 1980. A goodness-of-fit test for the multiple logistic regression
576 model. Communications in Statistics-Theory and Methods, 9, 1043–1069.

577 Kim, D., Chun, J.A., Choi, S.J. 2019. Incorporating the logistic regression into a decision-centric
578 assessment of climate change impacts on a complex river system. Hydrol. Earth Syst. Sci., 23, 1145–
579 1162, https://doi.org/10.5194/hess-23-1145-2019

580 Koutsoyiannis, D., Efstratiadis, A., Mamassis, N., Christofides, A. 2008. On the credibility of climate
581 predictions. Hydrological Sciences Journal, Vol. 53, Issue 4, pp. 671-684.
582 https://doi.org/10.1623/hysj.53.4.671

583 Kwakkel, J.H., Cunningham, S.C. 2016. Improving scenario discovery by bagging random boxes.
584 Technological Forecasting and Social Change, Elsevier, vol. 111(C), pages 124-134.

585 Kwakkel, J.H., Jaxa-Rozen, M. 2016. Improving scenario discovery for handling heterogeneous
586 uncertainties and multinomial classified outcomes. Environmental Modelling and Software, Volume 79,
587 311-321. https://doi.org/10.1016/j.envsoft.2015.11.020

588 Lempert, R.J., Groves, D.G., Popper, S.W., Bankes, S.C. 2006. A general, analytic method for generating
589 robust strategies and narrative scenarios. Management Sciences 52 (4), 514–528.

590 López-Moreno, J.I., Vicente-Serrano, S.M., Morán-Tejeda, E., Lorenzo-Lacruz, J., Kenawy, A.,
591 Beniston, M., 2011. Effects of the North Atlantic Oscillation (NAO) on combined temperature and
592 precipitation winter modes in the Mediterranean mountains: Observed relationships and
593 projections for the 21st century. Global and Planetary Change 77, 62–76

594 Macian-Sorribes, H., Tilmant, A., Pulido-Velazquez, M. 2017. Improving operating policies of large-scale
595 surface-groundwater systems through stochastic programming. Water Resources Research, 53(2), 1407-
596 1423, doi: 10.1002/2016WR019573.

597 Marcos-Garcia P., Pulido-Velazquez, M., 2017. Cambio climático y planificación hidrológica: ¿es
598 adecuado asumir un porcentaje único de reducción de aportaciones para toda la demarcación?
599 Ingeniería del agua, [S.l.], v. 21, n. 1, p. 35-52. ISSN 1886-4996 (in Spanish with English summary)

600 Marcos-Garcia, P., Lopez-Nicolas, A., Pulido-Velazquez, M. 2017. Combined use of relative drought
601 indices to analyze climate change impact on meteorological and hydrological droughts in a
602 Mediterranean basin. Journal of Hydrology 554, 292–305

603 MARM. 2008. Orden ARM/2656/2008, de 10 de septiembre, por la que se aprueba la Instrucción de
604 Planificación Hidrológica. BOE nº. 229, 22 September 2008, 38472-38582 (in Spanish)

605 Martin-Carrasco, F.J., Garrote, L. 2006. Drought-induced water scarcity in water resources systems. In:
606 Vasiliev O., van Gelder P., Plate E., Bolgov M. (eds) Extreme Hydrological Events: New Concepts for
607 Security. NATO Science Series, vol 78. Springer, Dordrecht

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), Frontiers in econometrics (pp. 104-142). New York: Academic Press

McKee, T.B., Doesken, N.J., Kleist, J. 1993. The relationship of drought frequency and duration of time scales. In: Eighth Conference on Applied Climatology, American Meteorological Society, Jan17-23, 1993, Anaheim CA, pp. 179–186.

Meier, P., Knox, S., Harou, J. 2014. Linking water resource network models to an open data management platform. D.P. Ames, N.W.T. Quinn, A.E. Rizzoli (Eds.), 7th Intl. Congress on Env. Modelling and Software. International Environmental Modelling and Software Society (IEMSs), San Diego, California, USA , pp. 463-469. http://www.iemss.org/society/index.php/iemss-2014-proceedings

MMA. 2001. Real Decreto Legislativo 1/2001, de 20 de julio, por el que se aprueba el Texto Refundido de la Ley de Aguas. BOE nº. 176, 24 July 2001 (in Spanish)

Muñoz-Díaz, D., Rodrigo, F.S, 2003. Effects of the North Atlantic Oscillation on the probability for climatic categories of local monthly rainfall in Southern Spain. Int. J. Climatol. 23: 381–397

Nagelkerke, N. 1991. A note on a general definition of the coefficient of determination. Biometrika, 78, 691–692.

Nattino, G., Finazzi, S., Bertolini, G. 2016. A new test and graphical tool to assess the goodness of fit of logistic regression models. Statistics in Medicine. Vol. 35, Issue 5, pp. 709-720. https://doi.org/10.1002/sim.6744

Nattino, G., Finazzi, S., Rossi, C., Carrara, G., Bertolini, G. 2017. givitiR package: assessing the calibration of binary outcome models with the GiViTI calibration belt. Available online: https://cran.r-project.org/web/packages/givitiR/vignettes/givitiR.html. Last access: July 2018.

Poff, N.L., Brown, C.M., Grantham, T.E., Matthews, J.H., Palmer, M.A., Spence, C.M., Wilby, R.L., Haasnoot, M., Mendoza, G.F., Dominique, K.C., Baeza, A. 2015. Sustainable water management under future uncertainty with eco-engineering decision scaling. Nat. Clim. Change., 10.1038/nclimate2765

632 Queralt, S., Hernández, E., Barriopedro, D., Gallego, D., Ribera, P., Casanova, C., 2009. North
633 Atlantic Oscillation influence and weather types associated with winter total and extreme precipitation
634 events in Spain. Atmospheric Research 94, 675–683

635 Ray, P., Brown, C. 2015. Confronting Climate Uncertainty in Water Resources Planning and Project
636 Design. The Decision Tree Framework. Washington D.C., The World Bank.

637 Soundharajan, BS., Adeloye, A.J., Remesan, R. 2016. Evaluating the variability in surface water reservoir
638 planning characteristics during climate change impacts assessment. Journal of Hydrology, 538:625–639.
639 doi: 10.1016/j.jhydrol.2016.04.051

640 Stainforth, D.A., Downing, T.E., Washington, R., Lopez, A., New, M. 2007. Issues in the interpretation of
641 climate model ensembles to inform decisions. Philos. Trans. R. Soc. A, 365,2163–2177.

642 Steinschneider, S., McCrary, R., Wi, S., Mulligan, K., Mearns, L., Brown, C. 2015. Expanded
643 decision-scaling framework to select robust long-term water-system plans under hydroclimatic
644 uncertainties, J. Water Resour. Plann. Manage., doi:10.1061/(ASCE)WR.1943-5452.0000536

645 Témez Peláez, J.R., 1977. Modelo matemático de transformación precipitación-aportación. ASINEL,
646 1977. (in Spanish)

647 Thornthwaite, C.W. 1948. An approach toward a rational classification of climate. Geogr. Rev. 38, 55–94.

648 Tjur, T. 2009. Coefficients of Determination in Logistic Regression Models—A New Proposal: The
649 Coefficient of Discrimination. The American Statistician. Volume 63, Issue 4.
650 https://doi.org/10.1198/tast.2009.08210

651 Toth, F.L., Cramer, W., Hizsnyik, E. 2000. Climate Impact Response Functions: An Introduction. Climatic
652 Change, 46: 225. https://doi.org/10.1023/A:1005668420713

653 Trigo, R.M., Pozo-Vázquez, D., Osborn, T.J., Castro-Díez, Y., Gámiz-Fortis, S., Esteban-Parra, M.J. 2004.
654 North Atlantic Oscillation influence on precipitation, river flow and water resources in the Iberian
655 peninsula. Int. J. Climatol. 24: 925–944

656 Venables, W. N. and Ripley, B. D. 2002. Modern Applied Statistics with S. Fourth Edition. Springer, New
657 York. ISBN 0-387-95457-0

658 Vicente-Serrano, S.M., Cuadrat, J.M., 2007. North Atlantic oscillation control of droughts in north-east
659 Spain: evaluation since 1600 A.D. Climatic Change 85:357–379. DOI 10.1007/s10584-007-9285-9

660 Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., El Kenawy, A.M., Angulo-Martínez, M.,
661 2009. Daily atmospheric circulation events and extreme precipitation risk in northeast Spain: Role of the
662 North Atlantic Oscillation, the Western Mediterranean Oscillation, and the Mediterranean Oscillation.
663 Journal of Geophysical Research, Vol. 114, D08106, doi:10.1029/2008JD011492

664 Vicente-Serrano, S.M., Beguería, S., López-Moreno, J.I., 2010. A multiscalar drought index sensitive to
665 global warming: the Standarized Precipitation Evapotranspiration Index. J. Clim. 23, 1696–1718

666 Villalobos, A. 2007. Análisis y seguimiento de diferentes tipos de sequía en la cuenca del río Júcar. PhD
667 dissertation. Universitat Politècnica de València. Valencia, Spain. (In Spanish)

668 Wilby, R.L., Dessai, S. 2010. Robust adaptation to climate change. Weather; 65:180-185.

SUPPLEMENTARY MATERIAL

For MOHC_RCA4 data (Figure 3), the models again overestimate failure probability in the case of probabilities larger than 0.68 (model "3 years") and 0.41 (model "4 years"), while for MPI_RCA4 data only "4 years" model shows values under the bisector for probabilities larger than 0.74.

Regarding the rest of the climate data and the 99% confidence level, as the bands do not intersect the bisector in the whole 0-1 range, it could be concluded that the logit models are neither underestimating, nor overestimating failure probability. Nevertheless, for example model "4 years" tends to overestimate the failure probability more than model "3 years" for the CNRM_RCA4 data (Figure 1), while model "3 years" tends to underestimate it in the MIROC_RCA4 and MPI_REMO_r2 cases (Figures 2 and 5) in comparison to the "4 years" model. For MPI_RCA4 the "3 years" model tends to underestimate low probabilities and overestimate the high ones (Figure 4). In general, it could be concluded that the "3 years" model shows a better predictive capacity than the "4 years" one, except for MIROC_RCA4 and MPI_REMO_r2 data, where the performance of the "4 years" model is higher.

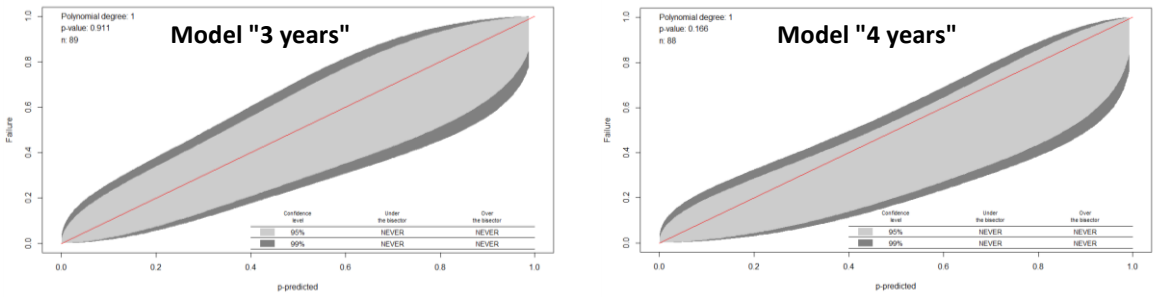
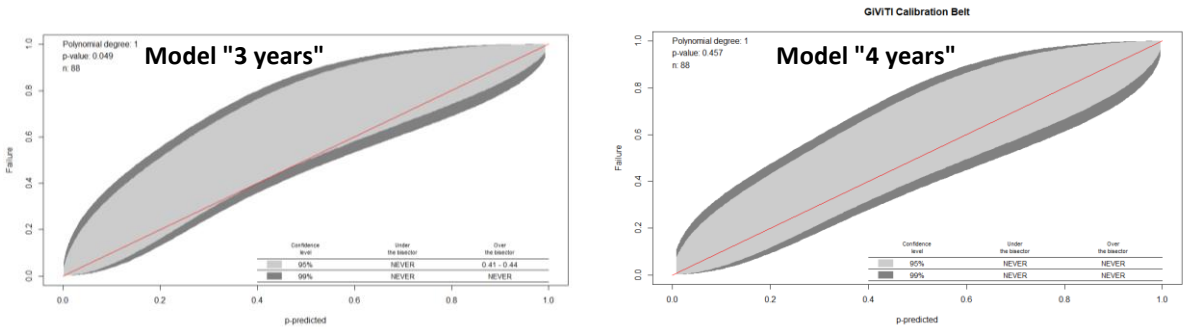
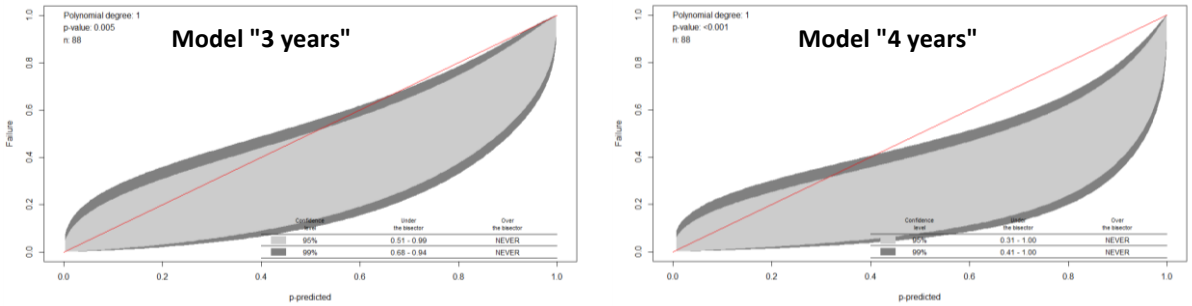


Figure 1. GiViTI calibration belt for CNRM_RCA4



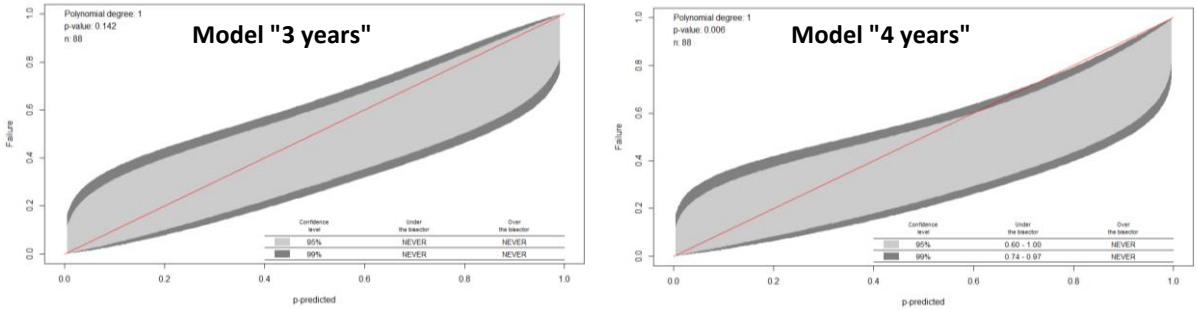
19

20 *Figure 2. GiViTI calibration belt for MIROC_RCA4*



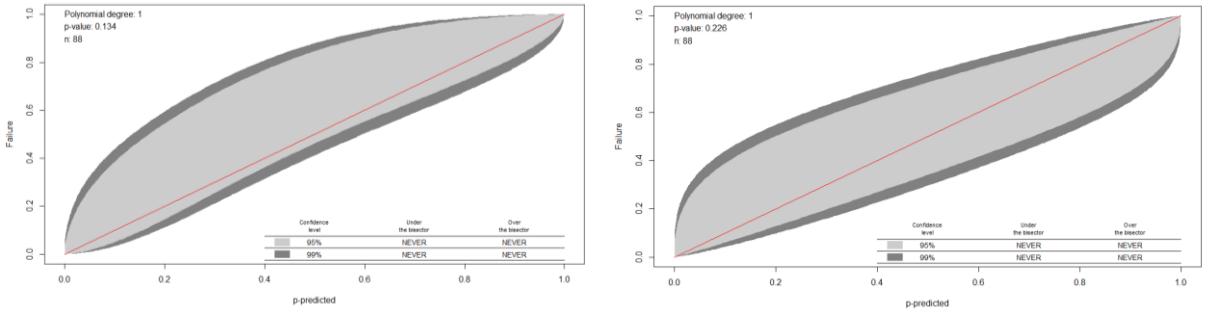
21

22 *Figure 3. GiViTI calibration belt for MOHC_RCA4*



23

24 *Figure 4. GiViTI calibration belt for MPI_RCA4*



25

26 *Figure 5. GiViTI calibration belt for MPI_REMO_r2*