

Document downloaded from:

<http://hdl.handle.net/10251/166256>

This paper must be cited as:

Macian-Sorribes, H.; Molina González, J.L.; Zazo-Del Dedo, S.; Pulido-Velazquez, M. (2021). Analysis of spatio-temporal dependence of inflow time series through Bayesian causal modelling. *Journal of Hydrology*. 597:1-14. <https://doi.org/10.1016/j.jhydrol.2020.125722>



The final publication is available at

<https://doi.org/10.1016/j.jhydrol.2020.125722>

Copyright Elsevier

Additional Information

# Analysis of spatio-temporal dependence of inflow time series through Bayesian Causal Modelling

Hector Macian-Sorribes <sup>(1)</sup>, Jose-Luis Molina <sup>(2)</sup>, Santiago Zazo <sup>(3)</sup> and Manuel Pulido-Velázquez <sup>(1)</sup>

(1) Universitat Politècnica de València, Research Institute of Water and Environmental Engineering (IIAMA), Ciudad Politécnica de la Innovación, Camino de Vera, 46022 Valencia, Spain.

(2) IGA Research Group. Salamanca University, High Polytechnic School of Engineering Avila, Department of Hydraulic Engineering, Av. de los Hornos Caleros, 50, 05003 Ávila, Spain. Corresponding author, e-mail: [jlmolina@usal.es](mailto:jlmolina@usal.es)

(3) Universitat Politècnica de València, Department of Hydraulic Engineering and Environment, C/ Camino de Vera s/n, 46022 Valencia, Spain

## Abstract

This paper aims to assess fully the spatio-temporal dependence dimensions of inflow across two adjacent and parallel basins and among different time steps through Causality. This is addressed from the perspective of Causal Reasoning, supported by Bayesian modelling, under a novel framework named Bayesian Causal Modelling (BCM). This is applied, through a “concept-proof”, to the Jucar River Basin (the second largest basin of Eastern Spain, characterized by long and severe drought conditions). In this “concept-proof” a double goal is evaluated; first dedicated to a lumped analysis of dependence and second a specific one over dry periods focused on time-horizon of the Jucar basin typical drought (3 years). These challenges comprise the development of two fully connected Bayesian Networks (BNs), one for each challenge populated/trained from historical-inflow records. BNs were designed at a season-scale and consequently, time was upscaled and grouped into Irrigation and Non-Irrigation periods, according to Jucar River Basin Authority operational practices. Results achieved showed that BCM framework satisfactorily captured the spatio-temporal dependencies of systems. Furthermore, BCM is able to answer to some key questions over interdependencies between adjacent and parallel subbasins. Those questions may comprise, the amount of spatial dependences among time series, the temporarily conditionality among subbasins and the spatio-temporal dependence among basins. This provides a relevant insight on the intrinsic spatio-temporal dependence structure of inflow time series in complex basins systems. This approach could be very valuable for water resources planning and management, due to its application power for predicting extreme events (e.g. droughts) as well as improving and optimizing the reservoirs operation rules.

**Keywords:** Causality, Causal reasoning, Bayesian spatio-temporal dependence, Stochastic hydrology, Jucar river basin, Historical inflow time series.

## 1. Introduction

The inherent nonlinear nature of hydrologic series and their associated processes is widely known (Wei et al., 2012; Molina et al., 2016; Peng et al., 2017). Furthermore, the nonstationarity of records (annual or even larger time scale) is rapidly growing worldwide over the last years due to modifications/alterations in the climatic patterns (Allan and Soden 2008; Trenberth 2011; Kalra et al., 2013; Chang et al., 2015; Wasko and Sharma, 2015; Donat et al., 2016;) as result of the influence of climate change into small-scale meteorological processes (Marotzke et al., 2017; Pfahl et al., 2017). This global situation is materialized in the increasing occurrence, intensity and persistence of extreme events (Jyrkama and Sykes, 2007; Marcos-Garcia et al., 2017; Molina and Zazo, 2018) including droughts, which can produce significant economic losses (Gil et al., 2011; Kahil et al., 2015; Freire-González et al. 2017; Lopez-Nicolas et al., 2017). In order to reduce these negative impacts, it is necessary to better understand the internal structure of dependencies that inherently underlies in the runoff time series (Molina et al., 2019). The meteorological-hydrological basin response is captured by means of this internal structure (Wang et al., 2009; Hao and Singh 2016; Molina and Zazo, 2018; Zazo et al., 2019). Furthermore, the spatio-temporal changes in weather patterns are likely to further aggravate the appearance and persistence of drought events (Mishra and Singh, 2010).

Likewise, the dependence of hydrological processes is a traditional studied matter, given its importance to assess, model and forecast the behavior and availability of water resources, especially in a context of droughts (Mishra and Singh, 2011). Classic ways to evaluate these dependences are mainly related to statistical procedures, including regression (e.g. Hao and Singh, 2016), time series modeling (e.g. Salas et al., 1980; Hipel and McLeod, 1994; Mishra and Desai, 2005), recurrence analysis (e.g. Estrela et al., 2000), and principal component analysis (e.g. Vicente-Serrano et al., 2004; Vicente-Serrano and Cuadrat-Prats, 2007). For a deeper explication of these techniques, the reader is referred to the work Mishra and Singh (2011).

Regarding temporal dependence of hydrological time series, this has traditionally received a more attention by the scientific community, for example by means of Persistence (inherent statistical property of time series), which is strongly related to the measure of the time series long-term memory through Hurst coefficient (Hurst, 1951), or storage and drought statistics (Salas et al., 1980), which are currently benchmarks; or

other ones such as Copula applications (De Michele and Salvadori, 2003), the usage of multivariate distributions in modeling different dependence structures (Sarabia-Alzaga and Gómez-Déniz, 2008), or novel analysis strategies based on Artificial Intelligence (AI), such as Artificial neural networks (ANNs; Ochoa-Rivera, 2008), Bayesian Networks (BN; Molina and Zazo, 2017; Zazo et al., 2019) or Dynamic Bayesian networks (DBNs; Molina et al., 2013). However, spatial and spatio-temporal dependence for hydrologic science and engineering is much poorer studied (Holmström et al., 2015) and even more through a Bayesian perspective (Wikle et al., 1998; Jona Lasinio et al., 2005). This is because of: (a) complexity of characterizing and differentiate water sub-systems, (b) scarcity of spatial data availability and (c) difficulties on the application of spatial statistical methods, among others.

On the other side, in the last years, the emergence of AI and Information Theory techniques have allowed tackling new ways of approaching, characterizing and quantifying dependences (Wang et al., 2009; Hejazi and Cai, 2011; Singh, 2011; Wu et al., 2015; Lima et al., 2016; Molina et al., 2016; Kousari et al., 2017; Yang et al., 2016, 2017a,b). These techniques are useful for exploring the evolution of meteorological and hydrological dynamics in order to find suitable drought predictors at different spatio-temporal scales (Molina et al., 2020).

On the other hand, Causality in hydrological records has not been deeply studied and it could be done by means of the joint use of different forms of reasoning patterns. These forms are Causal Reasoning (CR), Evidential Reasoning (ER) and Intercausal Reasoning (IR) (Koller and Friedman, 2009; Pearl, 2009). CR is used when the approach is done from top to bottom. In this sense, the analysis is focused on the cause and the objective comprises the prediction of the effect or consequence. Consequently, the queries in form of conditional probability, where the “downstream” effects of various factors are predicted, are instances of causal reasoning or prediction. Classical Stochastic Hydrology has been very focused along several decades on the characterization of temporal behavior for runoff or streamflow in different basins worldwide. In this sense, CR approach becomes especially useful when hydrological dependence analysis takes place because it is about an evaluation of the influence of past events on posterior hydrological processes. ER comprises bottom-up reasoning, so the analysis is focused on the consequence (effect) and the cause is inferred (Bayesian Inference). This approach would be appropriate to the analysis of catastrophic events such as flooding where the most important thing to be

studied is the consequence. IR is probably the hardest concept to understand. It comprises the interaction of different causes for the same effect. This type of reasoning is very useful in hydrology, where a consequence can be generated or explained from several causes. This approach would be very appropriate for developing Decision Support Systems based on Causality and Causal Modelling (Molina et al., 2010) where there is a high heterogeneity of variables' nature and high complexity of causal relationships. Furthermore, one of the most exciting prospects in recent years has been the possibility of using the theory of BNs to discover causal structures in raw data (historical runoff record) (Pearl, 2014). This is done in this research through the usage of historical inflow data to train and populate the BN model. Consequently, AI techniques such as CR and ER and/or IR provide new horizons for this type of studies.

Recently, the potential of BNs to find temporal causal structures into raw statistical data (Spirtes, 2010) has begun to be applied to discover and characterize the logical and non-trivial structure of dependencies that inherently underlies historical records (Zazo, 2017; Molina and Zazo, 2018; Molina et al., 2019, Zazo et al., 2019). Consequently, there is a general clear necessity of strengthen the spatio-temporal dependence studies on water systems (Holmström et al., 2015).

This study has been applied in Jucar River Basin (the second largest basin on Mediterranean side of Spain). This river basin is characterized by an irregular hydrology, with long drought episodes, an intense irrigation activity and the existence of large reservoirs able to provide carryover storage (Marcos et al., 2017). In order to achieve an efficient system operation, within a context of sustainable and safe management of water resources of a basin, it is necessary to adequately characterize, both in space and time, the hydrological discharge to its main reservoirs (Alarcon and Contreras) and their evolution, as well as deepening in the knowledge of their spatial-temporal behavior through the discovery and analysis of the non-trivial and logical structure of interdependence, which intrinsically underlies the historical records and that define of its behavior.

This paper is mainly aimed to provide a new methodological approach, named Bayesian Causal Modelling (BCM). This is done by means of an innovative “concept-proof”, focused on Causality, here addressed by Causal Reasoning and supported by Bayesian modelling. This provides not only temporal dependence behavior but also the spatial and spatio-temporal dependencies behavior of inflow across river basins. This

knowledge is practically unexplored in the field of water engineering and science and that is what the novelty is about. In this sense, the application of BCM both to the prediction of droughts and to operation rules improvement of reservoirs, by means of this approach seems quite straightforward in the case of in parallel interconnected basins.

On the other side, this work, although addressed through a “concept-proof”, is a further step in the research initiated with Molina et al. (2016) and Zazo (2017). This research activity is largely characterized by the application of BNs, in a dynamic and stochastically way, to increase knowledge over water resources behavior of basins.

This paper is organized as follows: after this Introduction, next Section 2 covers a description of the applied materials and methods for this research. Then, Section 3 is devoted to case studies description and available hydrological historical records. General methodology is shown and explained in depth in Section 4. After that, Section 5 presents the main experimental results drawn for the research. In Section 5, the results are discussed in detail. Finally, Section 6 addresses a Discussion and Conclusions from the study, where the most important aspects learnt from this research are shown.

## **2. Materials and methods**

### **2.1 Bayesian networks**

Bayesian Networks (BN), also known as Bayesian Belief Networks or Belief Networks, belong to the AI methods, such as artificial neural networks, fuzzy logic systems and decision trees (Molina et al., 2016). They are Probabilistic Graphical Models (PGM) in which a visual representation of a reasoning problem is performed to infer a new knowledge into an uncertainty context (Cabañas de Paz, 2017).

BNs combine probability theory with graph theory (Vogel et al., 2014). In this sense, they are based on: 1) a graphical representation comprising nodes and links, of a given set of random variables with their conditional interdependencies; and 2) a probabilistic model consisting of probabilistic expressions that describe the probability distributions/functions of the variables involved and the relationships between them depicted in the links (Cain, 2001; Castelletti and Soncini-Sessa, 2007; Sperotto et al., 2017). Furthermore, BNs aim to model the joint probability distribution of all considered variables. This is performed by means of propagation of (in)dependencies between the variables throughout the whole graphical structure, which can be seen as a general

description of the behavior of a system (Vogel et al., 2018). In other words, PGMs offer a compact representation of the joint probability distribution over sets of random variables (Said, 2006; Castelletti and Soncini-Sessa, 2007).

On the other side, each node in the network represents a variable. The input variables to a BN are graphically represented through the parent nodes (nodes without arriving links) and the output variables are computed following the links between the root nodes and the leaf nodes (nodes without departing links, which represent the output variables in the network). For each link, the probabilistic expressions representing the relationship between the departing and arriving nodes are applied. The input variables can be provided as single values or as probability functions, while the output variables are obtained as probability functions. This probabilistic assessment is the main distinctive feature of a BN and enables the possibility of characterizing uncertainty, perform risk analyses and support decision-making processes (Castelletti and Soncini-Sessa, 2007; Vogel et al., 2013, 2012; Sperotto et al., 2017).

Formally, BN  $N = (G, P)$  consists of a Direct Acyclic Graph (DAG; (Lappenschaar et al., 2012)) denoted by  $G = (V, E)$  and a set of probability distributions  $P$  (Molina et al., 2013), therefore, a BN  $N = (G, P)$  defines a joint probability distribution over all the analyzed variables. The DAG is defined by a set of nodes (or variables)  $V$  and a set of links (or edges)  $E$ . The edges join the variables and are oriented, indicating the causal dependence between the connected variables ( $A/B$  indicates  $A$  causes  $B$  or  $B$  is the effect of  $A$ ) (Madsen et al., 2003). The set of probability distributions  $P$  includes “a priori” probabilities and a set of conditional probabilities (expressed in form of Conditional Probabilities Tables; CPTs). Theorem of Bayes is performed as updater of “a priori” probabilities adding the observed evidence and providing “a posteriori” probabilities, mathematically expressed as (Bayes, 1763):

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)}$$

where  $P(B|A)$  is the conditional probability of B for a given state of variable A;  $P(A|B)$  the other-way-round conditional probability;  $P(B)$  the probability of B; and  $P(A)$  the probability of A.

Specifically applied to “concept-proof” over BCM of this research, the nodes in this case represent the hydrological space and time variables, which summarize the response of a basin to the hydrological cycle (Zazo, 2017) due to the fact that the inflow values intrinsically encompasses the influence of key factors such as rainfall-temperature variability, orography, fluvial morphodynamics, and anthropic factor (e.g. land-use change) as well as river-aquifer interaction that condition of the water flow in a point (Molina and Zazo, 2018), in this case inflow at a reservoir.

On the other hand, BNs present relevant advantages such as: 1) the explicit graphical representation, that lets any user see the variables and the relationships analyzed by them; and 2) the propagation of uncertainty, something very valuable in risk assessment and decision-making. They are also able to quantify the interdependencies existing between variables and they allow defining, in an easy and automatically way, relationships into complex systems (Molina et al., 2010) and proving more accurate results in the modeling of natural processes (See and Openshaw, 2000; Jain and Kumar, 2007). In contrast, their data requirements are big, their complexity grows as the systems and processes being modeled increase, and the fact that it is difficult to perform a quantitative validation on the results (Molina et al., 2010; Sperotto et al., 2017).

## **2.2 Spatio-temporal dependence analysis through Causality**

Spatio-temporal dependence analyses in hydrology are mainly related to the use of statistics and time series modeling procedures, which require them as a preliminary step (Salas et al., 1980; Hipel and McLeod, 1994). A crucial part of these approaches is the postulation of the stochastic model to be used (e.g. AR 1, ARMA 1,1, AR 2), and the choice of the statistical distribution followed by the variables considered by the model. Both decisions constrain the modeling of spatio-temporal dependences, since they assume that the analyzed variables are structured in a specific way and that this structure does not change over time, or that it changes following a specific pathway. In fact, the goal of these choices is to look for a spatio-temporal pattern close enough to the one shown by the available data records of the modeled variables. AI models, on the other hand, do not need to make assumptions on the data, as they are able to find out the properties and spatio-temporal patterns shown by it, even if they could not be identified before the analysis (Molina et al., 2016; Molina and Zazo, 2017). Although, AI approaches require large amount of data that not always is easy to obtain (Zazo, 2019; Molina et al., 2020).



At this point, it is worth highlighting that BNs are able to discover, extract and quantify the logical and non-trivial time-dependency relationships existing among variables by CR (Molina and Zazo, 2018; Molina et al., 2019). This research comprises an advance over the previous ones as it is able to discover and extract both spatial and spatio-temporal dependencies. Furthermore, once the BN is built, it is possible to determine how strong are the strength of relationships between variables by modifying the probability distribution of one of them and seeing how it affects the probability distributions of the remaining ones (Molina et al., 2016). This modification may be successfully addressed by means of maximizing its highest interval, and its impact is measured quantifying the change in the expected value of the other variables (the average) found when the first one is maximized (Molina et al., 2016). Under this approach, small changes mean that the relationship is weak, while large modifications refer to strong bonds between the variables represented in those nodes. If the BN represents the values of a given variable for different locations and time stages, the relationships obtained from the analysis are spatial, temporal and spatio-temporal.

### **2.3 Potential applications of BN for spatio-temporal analysis**

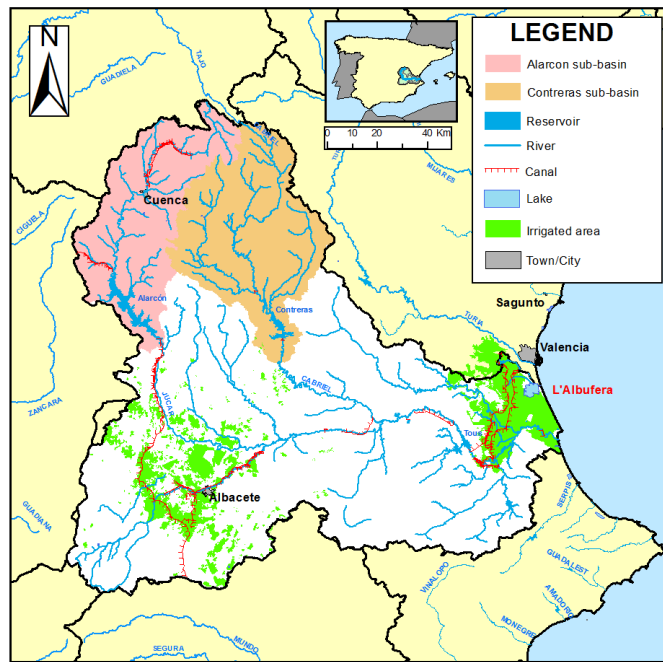
The use of BNs to perform spatio-temporal analyses shows a main advantage in comparison with classic statistics as the spatio-temporal dynamics of these relationships can be easily assessed. Their potential applications include:

1. Situations in which it is desired to explore if two or more variables show a causal relationship (e.g. between dry or wet cycles and climatic indicators such as the ENSO).
2. Situations in which the spatio-temporal pattern of the relationship cannot be known in advance or cannot be adequately modelled with traditional statistics (because stochastic models able to fit the given pattern would be very complex and thus, less parsimonious).
3. Situations in which fitting a probability distribution to the variables cannot be reliably done (because no probability function is found adequate enough or the chosen one would be very complex).
4. Situations in which the spatio-temporal dependencies vary over time (e.g. river basins subject to changing climatic patterns).

### 3. Case study: the Jucar River Basin in Spain

#### 3.1 Description

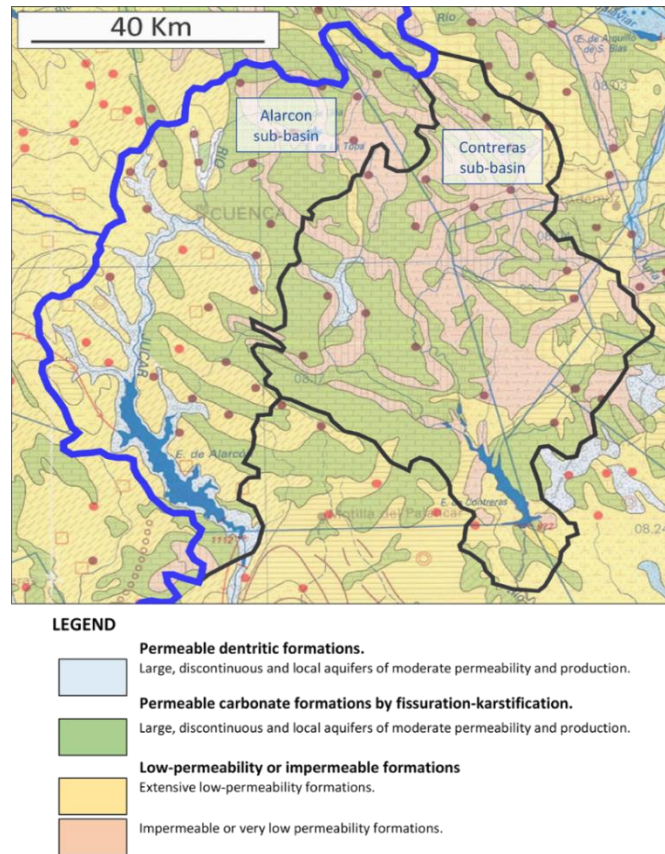
The case study is located at Jucar River Basin, a Mediterranean basin of 22,261 km<sup>2</sup> in Eastern Spain (Fig 1). It flows through 497 Km between the Iberian Mountains and the Mediterranean Sea. Following a typical Mediterranean pattern (peaks in autumn and very little values during summer), the annual rainfall spans from 309 to 717 mm, with an average value of 473 mm. Its mean annual resource is 1,605 Mm<sup>3</sup>/year, in which 70% comes from groundwater outflow through springs and stream-aquifer interaction (CHJ, 2015). Its main reservoirs are Alarcon (1,088 Mm<sup>3</sup> useful capacity), Contreras (429 Mm<sup>3</sup>), they are the most one's capacity of Jucar River Basin. Alarcón and Contreras are placed at the outlets of the upper basins of the Jucar and Cabriel rivers, while Tous is located at the limit of the Jucar floodplain, in which most urban and agricultural demands concentrate.



**Fig 1. Jucar River Basin location**

In lithological terms, Alarcon and Contreras sub-basins are characterized by existing of important limestone formations, dolomites, conglomerates, sandstones and Tertiary detritus (IGME, 2020a). Hydrogeology speaking, Alarcon sub-basin presents impermeable or very low permeability formations, together with extensive, discontinuous and local aquifers of moderate permeability and production. In contrast, Contreras sub-

basin is characterized by very permeable aquifers, generally extensive and productive with other ones impermeable or very low permeability zones (IGME, 2020b). Moreover, both sub-basins show a piezometric gradient towards the coastal area, from West to East (CHJ, 2006,2020) (Fig. 2).



**Fig 2. Hydrogeological maps of Alarcon and Contreras sub-basin. Source: IGME (2020b)**

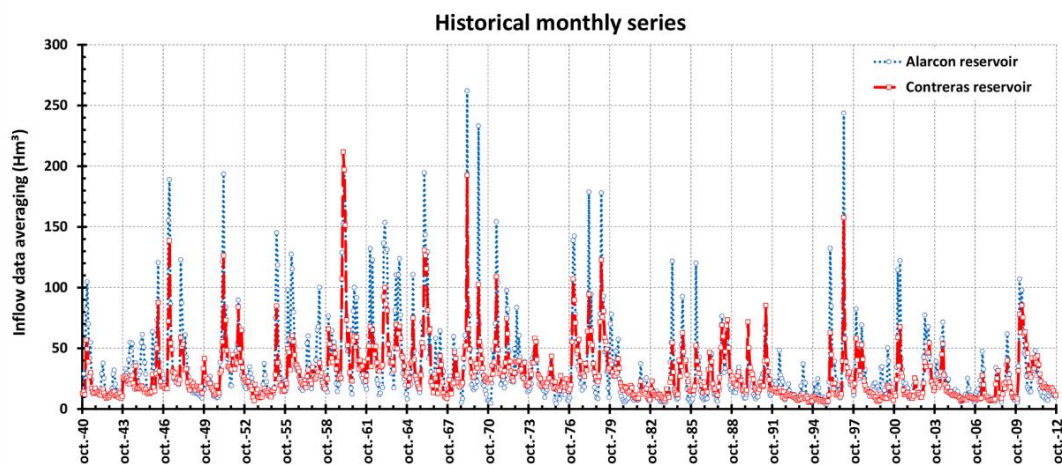
The annual consumptive demand of the Júcar river system is 1,557 Mm<sup>3</sup> for the 2015-2021 period (CHJ, 2015). The major consumptive use is agriculture (90%), followed by urban (8%) and industrial (2%) uses. The most important urban districts correspond to Valencia, Albacete and Sagunto. Surface irrigation demands concentrate on the lower basin (Riberas del Júcar and Canal Júcar-Turia), while the main groundwater-irrigated area is the Mancha Oriental, in the middle basin. The aquifer overdraft from this demand has caused an inversion of the stream-aquifer interaction found between the Mancha Oriental aquifer and the Júcar river, moving from gaining to losing river. Furthermore, minimum flows are set on 18 streams by environmental reasons (CHJ, 2015).

The planning and operation of the Júcar River Basin is the main responsibility of the Júcar River Basin Authority (in Spanish *Confederacion Hidrográfica del Júcar*, CHJ).

Once environmental requirements are met, the current operating rules give priority to urban uses over agricultural areas. Among the latter, the users with elder rights from the lower basin, gathered together in the users' union USUJ (*Unidad Sindical de Usuarios del Júcar*), are given the highest surface priority. The remaining demands, with access to both surface and ground waters, can only use surface water if the Alarcon storage is higher than the limits established by the Alarcon Agreement (CHJ, 2015).

### 3.2 Available and processed data

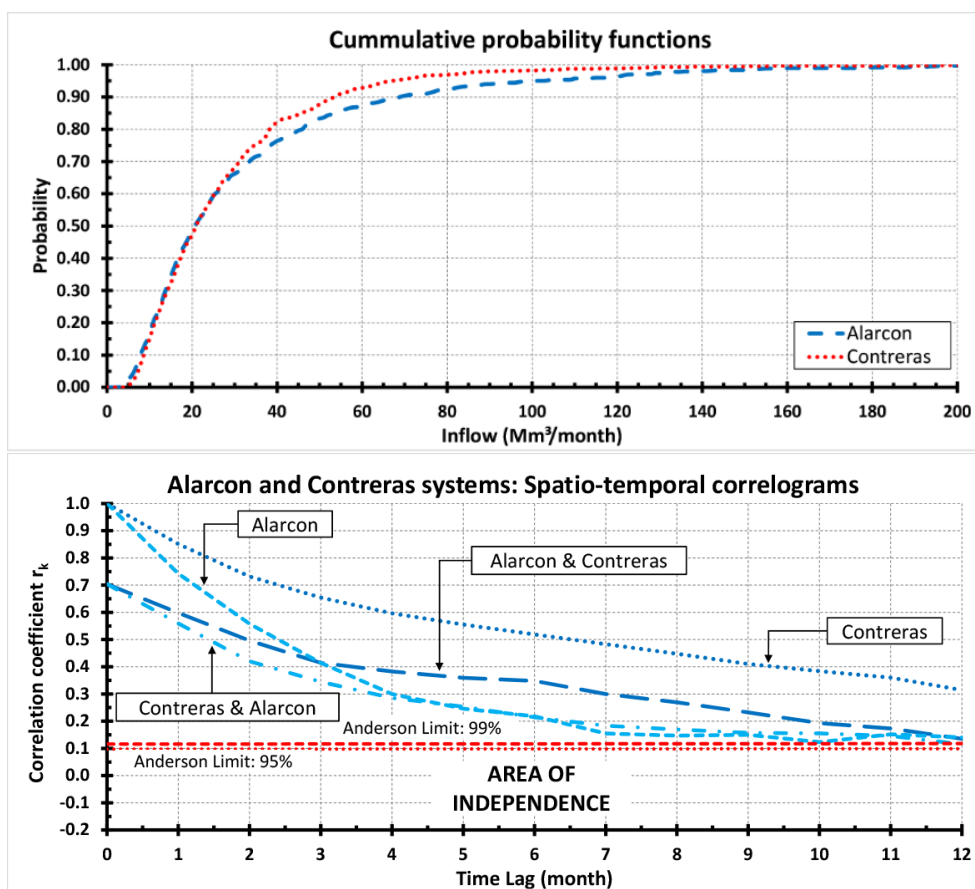
The BCM analysis of spatio-temporal inflow dependence has been applied to the Alarcon and Contreras sub-basins, which correspond to the upper part of the Jucar River Basin and are the main contributors to the water discharge that can be regulated through the Alarcon and Contreras reservoirs (see Fig 1). The available data consists of monthly inflow time series to the reservoirs for the 1940-2012 period (72 hydrological years; please note: in Spanish hydrological years the first month is October and the last is September of following year), obtained from the Spanish Ministry of Agriculture, Fishery, Food and Environment (<http://ceh-flumen64.cedex.es/anuarioaforos/default.asp>). They have been naturalized by removing the effect of evaporation and seepage losses in Alarcon and Contreras, given that anthropic modifications in the upper basins of the Jucar and the Cabriel rivers are negligible. Fig. 3 and Table 1 display the monthly time series considered and their main statistic parameters respectively, as well as the probability distribution functions of both time series and their correlograms are presented in Fig. 4.



**Fig 3. Alarcon and Contreras historical monthly series after removing effect of evaporation and seepage losses, according to Spanish hydrological years (first month October and the last one is September of following year).**

Statistic parameters	Sub-basins	
	Alarcon	Contreras
Monthly mean (Hm <sup>3</sup> ):	33.01	28.51
Annual mean (Hm <sup>3</sup> ):	396.09	342.12
Minimum (Hm <sup>3</sup> ):	3.86	5.52
Maximum (Hm <sup>3</sup> ):	262.14	211.70
Range (Hm <sup>3</sup> ):	258.28	206.18
Variance:	1099.39	532.31
Standard deviation (Hm <sup>3</sup> ):	33.16	23.07
Variation coefficient (%):	100.5	80.9
Skewness coefficient (*):	2.78	2.98
	Positive skewness	Positive skewness
Kurtosis:	10.12	14.30

**Table 1.** Alarcon and Contreras monthly time series. Main statistic parameters.



**Fig 4.** Preliminary assessment of the spatial and temporal dependencies of the Alarcon and Contreras sub-basins. Upper: Probability distribution functions (own elaboration). Bottom: Temporal and Spatio-temporal correlograms with Anderson probability limits for an independent series (95 and 99 percent probability level). Please note that the Anderson limits define independence area of a correlogram.

The similarity exposed by their probability distributions is the result of its spatial proximity and their resemblance in physical, meteorological and hydrological properties. Their correlograms show a strong spatial and spatio-temporal dependence. The temporal dependence of Contreras runoff is higher than in Alarcon. On the contrary, the spatio-

temporal correlograms show a higher influence of Alarcon on Contreras inflow values than in the opposite. The main reason behind these differences seems to be the role of groundwater discharge (CHJ, 2006, 2020). These relationships will be explored more deeply with the BCM analysis.

#### 4. Methodology

This research was articulated in three consecutive and interrelated sequential steps (Fig. 5). Step-1 comprises the treatment and analysis of the raw inflow data, in order to adapt it to the subsequent steps and to the goals pursued by the study. After that, Step-2 corresponds to the building of the BN\_PGM models applied to perform the spatio-temporal BCM analysis; and they models are populated and trained by the treated data previously. Finally, Step-3 is fully entirely devoted to calculating and analyzing the causal relationships through “concept-proof” BCM. This last step is crucial because of it will allow discovering, extracting and quantifying of the logical and non-trivial time-dependency structure that inherently underlies into hydrological series and that define of the behavior of the adjacent sub-basins.

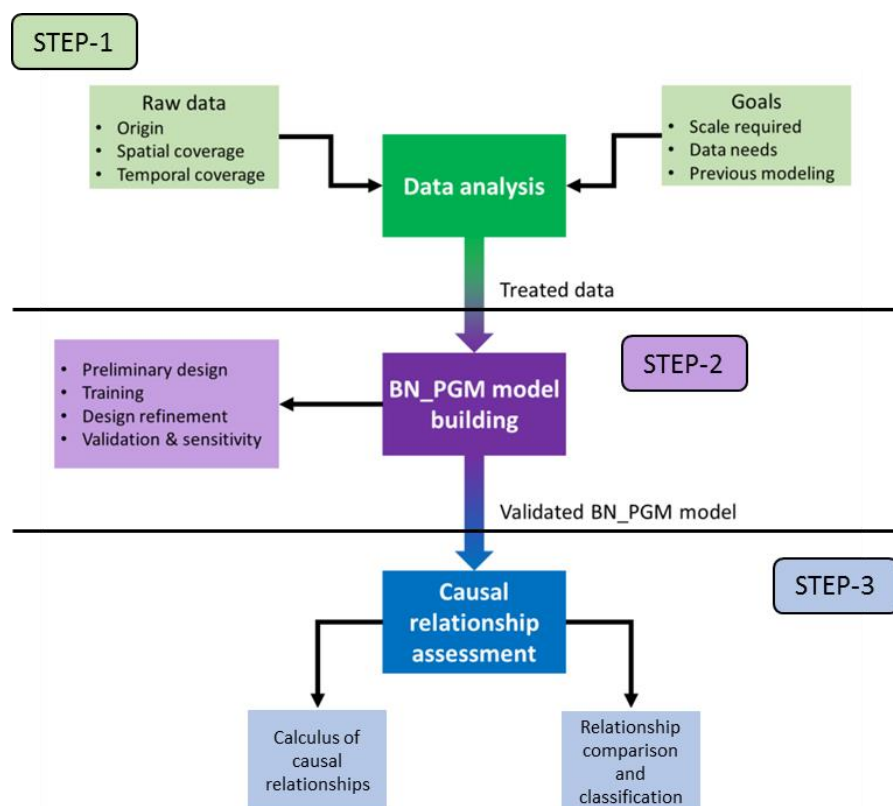


Fig 5. General methodology.

#### **4.1 Step-1. Data analysis**

The key factors of this step are the origin of the data (historical records, outputs from hydrological modeling, results obtained by stochastic models and so on), its spatial coverage (e.g. the entire hydrological basin as a lumped unit, the entire basin divided into sub-basins, certain sub-basins, ...) and its temporal coverage (period of analysis, time scale of the data, etc.). They should be shaped according to the objective of the analysis. For example, the evaluation of climate change impacts would require the obtaining of future inflow scenarios through hydrological modeling, while analyses based on the current hydrology may be done employing historical records. One aspect that should also be considered is the data needs of a BN, which may require to extend the available records using modelling techniques (e.g. Molina et al., 2016).

In the case of the Jucar river basin, the analysis has been developed at the seasonal scale, dividing the monthly inflows into two different periods per year. This division has been made in accordance with the operational practices of the Jucar River Basin Authority (CHJ). The Irrigation Period (IP; to May from September) corresponds to the irrigation season, in which agricultural demands concentrate and rainfall is reduced, so reservoirs are lowered to satisfy them. The Non-Irrigation Period (NIP; to October from April of next year) is characterized by low agricultural demands and higher rainfalls than the summer one. The Jucar river system operation by the CHJ Operation Office is supported by projections on inflows during the irrigation period based on the hydrological discharge on the past months (Macian-Sorribes and Pulido-Velazquez, 2017). Based on them, deliveries from the reservoirs to the farmers are scheduled for the upcoming irrigation season. Consequently, this division aligns the analysis with the practices of the system operators.

#### **4.2 Step-2. Causal model building**

This step involves developing BN-PGM model by means of other four sub-steps: 1) Preliminary design; 2) training; 3) design refining; and 4) validation and sensitivity analysis. At the beginning of the process, a first design of the PGM is proposed. A training process on this draft is performed to refine the design of the BN-PGM model and make sure it captures the spatio-temporal relationships shown by the data. Once the model is set, it is necessary to validate and to perform a sensitivity analysis, based on the perspective of the Theory of Information.

Given the novelty of application of this BCM methodology, there were not previous references on how to proceed regarding the application of the PGMs for the triple dimension of analysis pursued by this research, so it was built in an iterative way. After a period of analysis within the preliminary design phase, an initial draft was developed. Once the network design was set, supported by a powerful training through Learning Wizard (HUGIN® version 7.3), a fully connected PGM was built for wholly capturing and modelling the double dimension (spatio-temporal) of both sub-basins. This is consistent with the strong relationships found in the previous data analysis (please see Fig 4 bottom, correlograms). Finally, model validation was developed through the calculation of some important parameters belonging to Information Theory discipline. In particular, it has been used Total Entropy, Conditional Entropy and Mutual information. In the Jucar river basin two different BN-PGM models have been built: one describing the average and lumped behavior of the upper sub-basins and another specific one used to obtain the spatio-temporal relationships during a dry period.

### **4.3 Step-3. Bayesian Causal Modelling**

This is a key step in this research and consists in calculating and analyzing the causal relationships using the BCM developed in the last step, following the process explained in section 2.2. The estimation of the relationships embedded in each BN-PGM consisted in alternatively modifying (maximizing) one node of the model and evaluate how the remaining nodes change in response (Molina et al., 2016). A remarkable change in one particular node associated with a change in another would imply a strong causal relationship between them, whereas a little one would mean the absence of a bond. The analysis of these relationships consists in comparing and sorting the changes previously calculated, considering the type of relationship that each pair of nodes has (e.g. spatial, temporal, spatio-temporal).

This step was also tackled from a triple dimension that comprises spatial, temporal and spatio-temporal analysis. Drawn results from these dimensions show important differences among them. Furthermore, due to the flexibility of design and the back propagation of the probability, the BN-PGM was able to evaluate the influence of both sub-basins each other in both ways. This is a relevant characteristic, which is essential for developing an accurate and rigorous water rivers basins management. This is highlighted through two different analysis. First one dedicated to a lumped analysis of dependence,



and the second one by means of a specific analysis over dry periods focused on time-horizon of the Jucar basin typical drought (3 years; (CHJ, 2007)). Both analyses comprise the development of two fully connected Bayesian Networks (BNs), designed at a seasonal scale and grouped in two different periods, Temporal Irrigation and Non-Irrigation ones, according to Jucar River Basin Authority operational practices. Furthermore, it was considered a year as dry year if either Alarcon or Contreras runoff during the winter season falls below the percentile 33 of the whole historical period.

A novel variation ratio (*VR*) has been proposed that allows a dependence analysis between the 2 subbasins at the three dimensions (spatial, temporal and spatio-temporal) at seasonal scale. This index is obtained as the division between the expected value of the probability function after maximization of the probability intervals from designed BN-PGM models and the one from before. In order to facilitate its compression, the outcomes were normalized, in this way that variation ratio had the same physical sense as a classical correlation coefficient. That is, the higher the value departs from zero (0), the stronger the dependence is. In contrast, values lower than 0 indicate an inverse dependence, and 0 value displays non-dependence.

#### **4.4 BCM Concept-proof. Applied conceptual frameworks**

Due to the absence of previous references on this type of application, it is essential to describe the applied conceptual frameworks (Fig 6). Each node of the BN represents the runoff of an upper Jucar sub-basin for a given season. Alarcon NIP is the runoff of the Alarcon sub-basin for the non-irrigation period; Contreras NIP is the runoff of Contreras for the non-irrigation period; Alarcon IP corresponds to the runoff from the Alarcon sub-basin during the irrigation season; and Contreras IP represents the runoff from Contreras for the irrigation period.

On the other side, the data used to populate and train this BN corresponds to the available monthly historical records presented in Section 3.2., obtained from the Jucar River Basin Authority (CHJ), upscaled to seasonal periods (months from October to April added to form the non-irrigation runoff and months from May to September to form the irrigation runoff).

Consequently, in order to propagate the Bayes' Theorem in both space and time the Fig. 6 summarizes the applied conceptual frameworks both for lumped analysis and specific ones focused on dry cycle analysis.

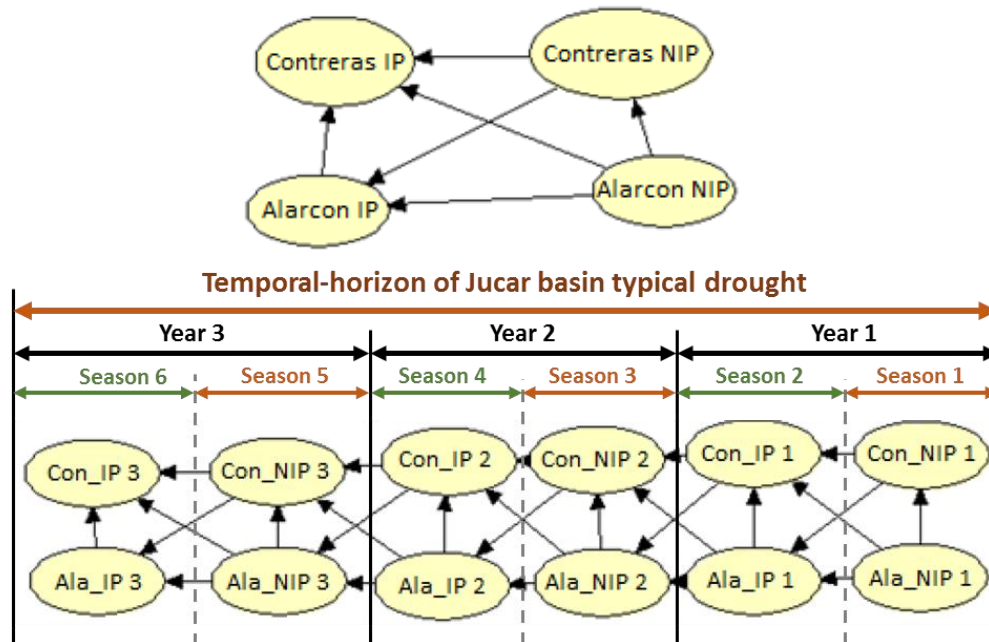


Fig 6. BCMs conceptual frameworks to upper Jucar river sub-basins. Direct Acyclic Graph considered between one season and the subsequent one. Upper lumped analysis. Bottom specific analysis over dry cycle. Abbreviations used: Non-irrigation period (NIP). Irrigation period (IP). Contreras reservoir (Con). Alarcon reservoir (Ala). Please note that the numbers indicate the year of the dry cycle analyzed.

The BCM was built using the Learning Wizard of HUGIN Expert ® (V 7.3), which generated probabilistic distributions of the variables associated with each node, as well as a logic structure according to the internal dependences and relationships detected. The Learning Wizard result was a fully connected network (each node interacts with the remaining three), which is in accordance to the strong dependences shown by the statistical analysis performed before (please see Fig 4 bottom, correlograms). Therefore, each node of the network shows a causal relationship with the remaining three. Another reason for this fully connected design is that it is the best way for assuring the complete capture of the runoff dependence for the double dimension: space and time.

Finally, although this conceptual frameworks, by itself could be understood as results of the research, it has been considered appropriate to show them as part of the methodological section, given the novelty of these approaches, in this way the reader is facilitated to understand of the developed process.

#### 4.5 Concept-proof. Sensitivity Analysis and Validation of results

Sensitivity analysis and validation of results is addressed through the development of an Entropy approach, based on the perspective of the Theory of Information. In this sense, Conditional Entropy, Mutual information is calculated for the Bayesian causal models.

Sensitivity analysis can be performed using two types of measures; entropy and Shannon's measure of mutual information (Pearl, 1988). The entropy measure assumes that the uncertainty or randomness of a variable  $X$ , characterized by probability distribution  $P(x)$ , can be represented by the entropy function  $H(X)$  (Molina et al., 2016):

$$H(X) = - \sum_{x \in X} P(x) \log P(x)$$

Entropy of a probability distribution can be defined as a measure of the associated uncertainty to that random process that this distribution describes. Consequently, a score of uncertainty/certainty level of events can be made attending to this entropy,  $H(X)$ .

Reducing  $H(X)$  by collecting information in addition to the current knowledge about variable  $X$  is interpreted as reducing the uncertainty about the true state of  $X$  (Molina et al., 2016; Zazo, 2017). The entropy measure therefore enables an assessment of the additional information required to specify a particular alternative. Shannon's measure of mutual information is used to assess the effect of collecting information about one variable ( $Y$ ) in reducing the total uncertainty about variable  $X$  using:

$$I(Y.X) = H(Y) - H(Y | X)$$

where  $I(Y.X)$  is the mutual information between variables. This measure reports the expected degree to which the joint probability of  $X$  and  $Y$  diverges from what it would be if  $X$  were independent of  $Y$ . If  $I(Y.X) = 0$ ,  $X$  and  $Y$  are mutually independent.  $H(Y|X)$  is conditional entropy which means the uncertainty that remains about  $Y$  when  $X$  is known to be  $x$ .

According to the entropy approach presented, mutual information values higher than 0 would imply some dependence between the variables analyzed, while the opposite would mean that they are independent from each other (Pearl, 1988). Therefore, this analysis represents an alternative way to characterize the dependences between variables.

In this research, this validation is useful because its results can be compared with the temporal dependence analysis carried out.

Assessments of the entropy associated to each variable, the conditional entropy and mutual information for each variable and connection were performed supported by HUGIN Expert ® software.

## 5. Results

### 5.1 Average spatio-temporal dependences

The average spatio-temporal dependence analysis has been done by maximizing the probability distribution of one variable of the BN each time and measuring how this perturbation affects the probability distributions of the rest of the network. This dependence has been evaluated using a novel and normalized developed variation ratio, for each remaining node of the network, as the division between the expected value of the probability function after the perturbation and the one from before (Fig 7). Moreover, the dependences obtained have been divided into spatial (same season, different sub-basin), temporal (different season, same sub-basin) and spatio-temporal (different season, different sub-basin). The considered seasons are non-irrigation (NIP) and irrigation (IP); while the sub-basins correspond to Alarcon (A) and Contreras (C).

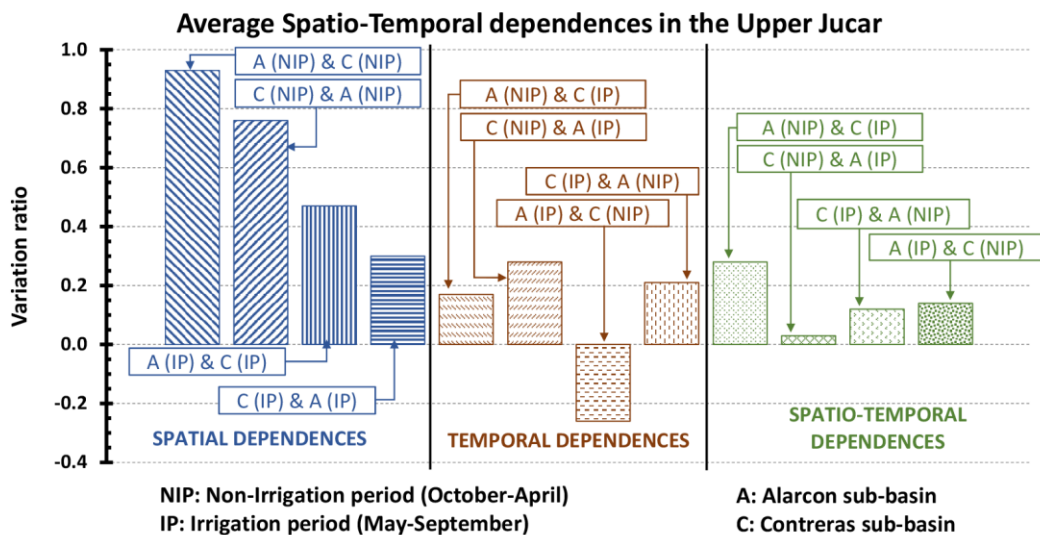


Fig 7. Average spatio-temporal dependences for the upper Jucar river basin

It can be observed that spatial dependences (blue bars) are the strongest ones, with values distinctly higher than 0 and higher than the temporal and the spatio-temporal dependences. This bond is in accordance with the geographical proximity of both sub-basins. Furthermore, dependences are higher for the non-irrigation season than during irrigation. This may be caused by the higher rainfalls in the non-irrigation season, since its geographic proximity drives similar rainfall events, strengthening the bond between them. Another interesting aspect (also noticed in the correlograms, see Fig 4 bottom) is the fact that the influence of Alarcon subbasin on Contreras' inflows is higher than in the opposite way, for both seasons. This asymmetry in their dependences may be caused by groundwater flows (please see Fig 2), which show an underground discharge from Alarcon sub-basin to Contreras sub-basin, according to piezometric levels as it is shown CHJ (2006,2020).

Considering the temporal dependences, Contreras shows a higher bond than Alarcon, given the higher variation ratio noticed in Contreras. This agrees with the correlograms, in which this influence is clearly shown by the significant statistical evidence (see Fig. 4 bottom; Alarcon & Contreras correlogram displays higher correlation coefficients,  $r_k$ , than Contreras & Alarcon correlogram in all analyzed monthly time lags). The reason behind this may be the geological differences found between both sub-basins, which can be explored using the information provided by the Geological Survey of Spain (Instituto Geológico y Minero de España, IGME, <http://www.igme.es/default.asp>). In this sense, the groundwater bodies under the Alarcon sub-basin present a significant portion of detrital sedimentary rocks, while the aquifers related with the Contreras sub-basin are mainly chemical sedimentary rocks. Given that detrital rocks are associated with quicker hydrological responses than chemical ones, the geological differences noticed may cause a slower hydrological response (and thus with longer time spans) in Contreras.

Regarding the spatio-temporal dependences, the prevalence pattern observed for the spatial ones (more influence on Alarcon in Contreras than the opposite) is maintained from the non-irrigation to the irrigation season. In fact, Contreras runoff on non-irrigation does not seem to play any influence on the Alarcon runoff during irrigation, as the variation ratio is approximately 0. The reason behind this behavior may be the maintenance across time of the bond observed in the spatial analysis. The lack of influence of Contreras on Alarcon seems to be caused by the limited groundwater flow from the Contreras sub-basin to the Alarcon one. It is also interesting notice that the

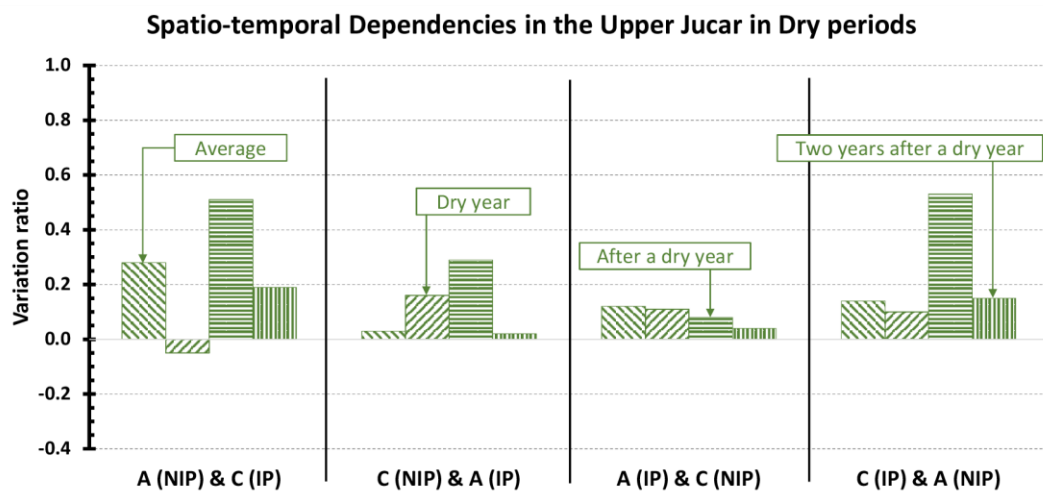
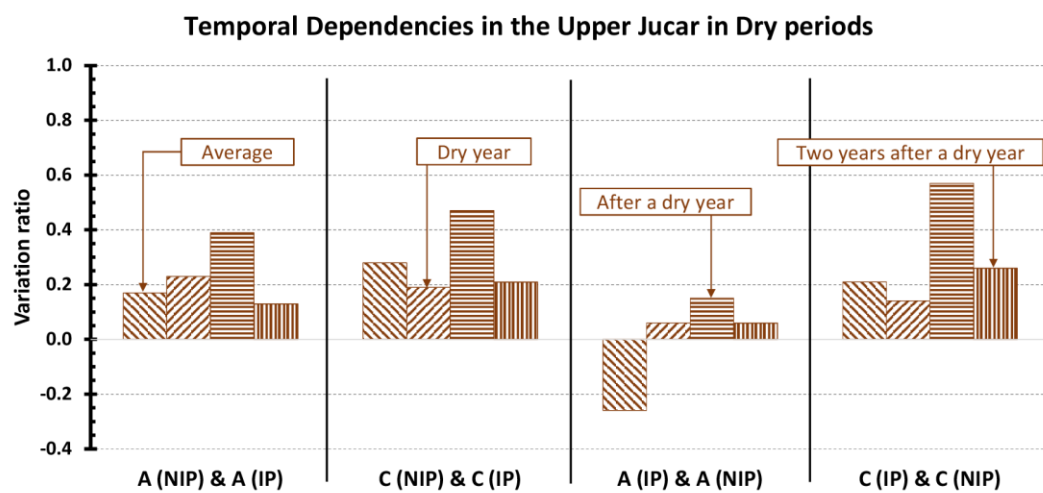
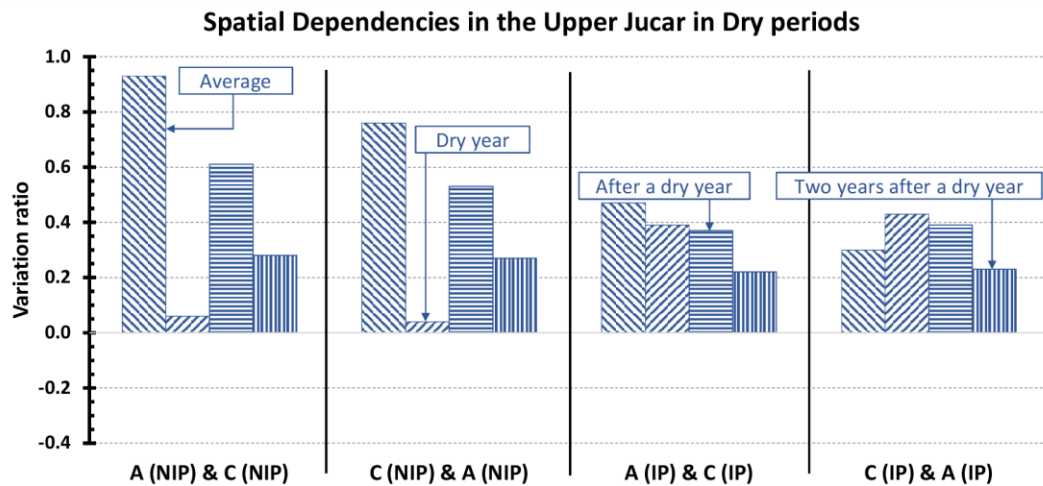
spatio-temporal dependence from Alarcon to Contreras (A (NIP) & C (IP)) is similar to the temporal dependence of Contreras (C (NIP) & C (IP)).

Overall, results show a relevant asymmetry on the spatial and spatio-temporal influence of both basins each other. This may be due to the fact that Alarcon sub-basin plays more influence on Contreras than the opposite, probably by the existence of groundwater flows following mainly the Alarcon-Contreras direction (piezometric gradient towards the coastal area, from West to East (CHJ, 2006,2020)).

## **5.2 Spatio-temporal dependence on dry years**

Given that dry spells in the Jucar river basin last more than one year, a new BCM has been developed extending its temporal-horizon up to 3 years (6 seasons, see Fig. 6 Bottom). According to dry year definition a set of 27 years were considered. In the most of them, both Alarcon and Contreras inflow during winter fall below the threshold, as well as during the summer season.

Fig 8 shows the results of the causal relationship analysis done for dry periods, starting at a dry year and spanning to 2 years after a dry one. Most of the years after a dry year were dry too, but a 3-year dry spell was only found for the case of an extreme drought.



NIP: Non-Irrigation period (October-April)  
 IP: Irrigation period (May-September)

A: Alarcon sub-basin  
 C: Contreras sub-basin

**Fig 8. Spatio-temporal dependence in the upper Jucar river basin during dry periods**

The main difference between the average dependences and the ones in a dry year is the absence of significant spatial relationships during the non-irrigation season, as the variation ratios are similar to 0 (please see Fig. 8 upper; A (NIP) & C (NIP) and C (NIP) & A (NIP)). This may be caused by the absence of rainfall and the lowering of groundwater tables. The decrease in heads is steeper in the Alarcon sub-basin due to having more detrital rock areas than Contreras, causing the interruption of the groundwater flow from Alarcon to Contreras. During irrigation, on the other hand, spatial dependences are similar to the average ones. This seems to be caused by the fact that rainfall during the irrigation season is low even during wet years, so the differences among years are less relevant than the ones found for the non-irrigation season. Another change between normal and dry years is that Contreras plays more influence on Alarcon than Alarcon on Contreras during dry years (C (IP) & A (IP) is greater than A (IP) & C (IP)). This inversion is also shown in the spatio-temporal relationships (C (NIP) & A (IP) is greater than A (NIP) & C (IP)). The temporal relationships do not suffer significant modifications, although Contreras does not show higher temporal dependences than Alarcon.

The causal relationships observed after a dry year (which in most cases corresponds to a dry year) present a remarkable increase in temporal and spatio-temporal relationships with respect to the average dependences. This increase in the relationships with temporal component may be caused by groundwater discharge, since its importance grows during long periods without rainfall. Spatial relationships show higher levels than the ones found for the first dry year, but do not reach the levels shown on average. This seems to point that groundwater exchanges between sub-basins increase after a dry year, may be because the geological differences between both sub-basins cause an unbalance between piezometric heads again (due to different hydrological response times).

Two years after the start of the dry spell (which except for the most severe droughts found, is not dry) the temporal and spatio-temporal dependences are, on a broader view, similar to the average. This is consistent with the ending of a dry period. However, spatial dependences are lower than the average while showing similar levels regardless of the sub-basin and season. This may be caused by a rise in precipitations, since its influence in the bond between sub-basins is the same regardless of the direction of the relationship (from Alarcon to Contreras or from Contreras to Alarcon).

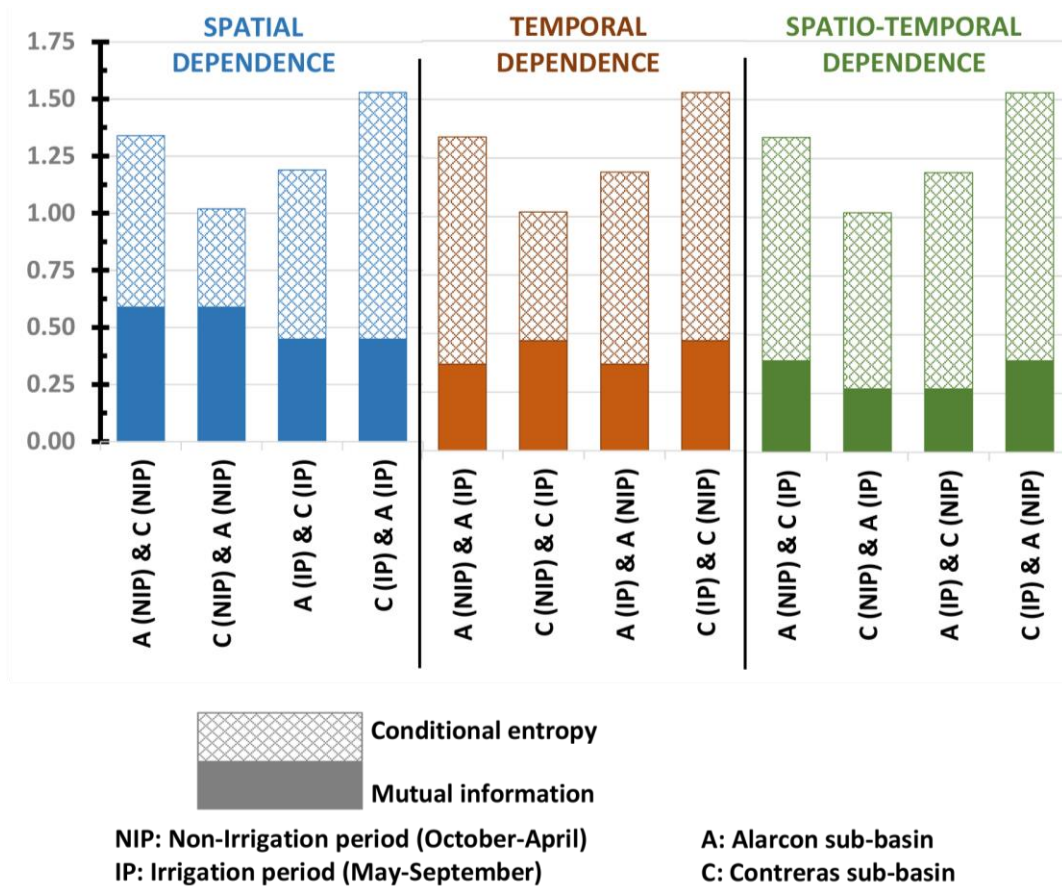


### 5.3 Sensitivity and Validation Analysis

Calibration of causal reasoning process is done through a comparative analysis with traditional techniques such as correlation. In this sense, the results obtained from the causal reasoning analysis for the average BN-PGM model resemble the ones provided by correlation (please see Fig 4 Bottom), as explained previously.

On the other side, assessments of the entropy associated to each variable, the conditional entropy and mutual information for each variable and connection were done for the average BN-PGM model (please see Fig 6 Upper, BCM conceptual framework for lumped analysis). In this sense, Fig 9 summarizes the main achieved results. These outcomes of the entropy analysis agree with the causal relationship analysis (see Fig 7). In particular, the same prevalence pattern is observed: spatial dependences show the highest values of mutual information, then temporal and finally spatio-temporal dependences, whose values are the lowest.

Regarding the mutual information values of the spatial dependences are higher in the non-irrigation period, as the causal relationships. Similarly, mutual information values on the temporal dependences of Contreras are higher than the ones in Alarcon. Regarding the spatio-temporal dependences, the one between A (NIP) and C (IP) has the highest mutual information value, in accordance with the strongest causal relationship found between both nodes in comparison to the rest of spatio-temporal dependences. In spite of the similarity between them, there exist some differences. The most important is that the results of the mutual information between two nodes are the same regardless of the direction considered (e.g. A (NIP) & C (NIP) has the same value than C (NIP) & A (NIP)), while the causal relationship analysis is sensitive to it. Moreover, the remarkable differences found between the spatial and the remaining causal relationships do not appear in the entropy analysis, in which only slight modifications are found.



**Fig 9. Sensitivity analysis results for the average behavior of Alarcon (Ala) and Contreras (Con) sub-basins**

On the other hand, Fig. 10 shows an entropy analysis of the BN-PGM model for a dry period, which is also supported through HUGIN Expert<sup>®</sup> software. The entropy results clearly reflect the temporal mitigation of dependences, since the mutual information values decrease when moving forward or backward in time. The mutual information values distinctly decrease when moving more than one year apart from the hypothesis node. This is in line with the correlograms (see Fig 4), which show dependence for 12 months. The fact that Contreras has more temporal dependence than Alarcon is also shown in the entropy analysis, since mutual information values associated with temporal dependences are higher in Contreras than in Alarcon (although with slight increases).

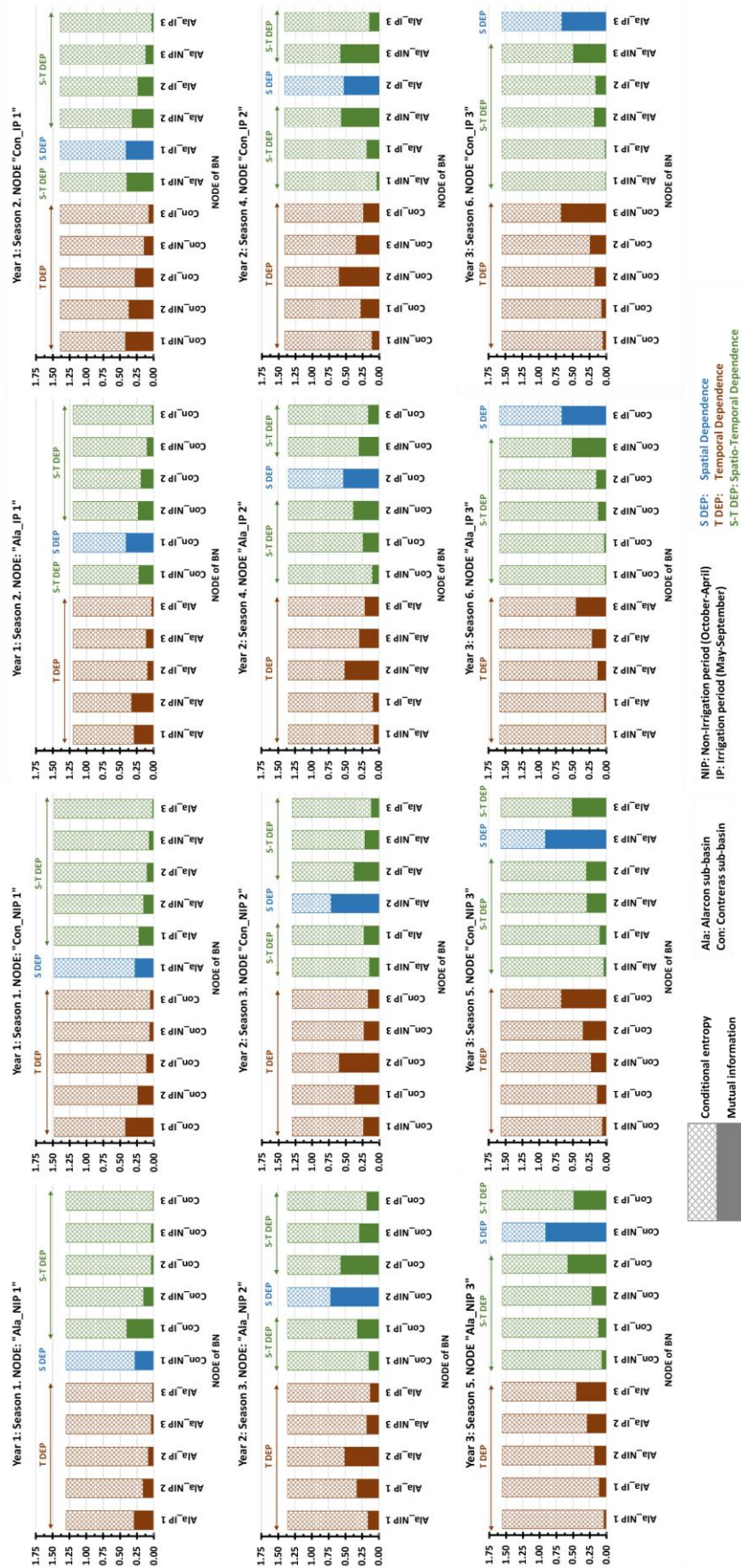


Fig 10. Entropy analysis over dry cycle analysis performed over time-horizon of 3 year (6 seasons).

On a broader view, the results of the entropy analysis agree with the causal relationship analysis. The mutual information values for the first year of a dry cycle are lower than on average, in accordance with the causal relationship analysis, as well as the increase noticed in the mutual information values after the first year. Similarly, temporal mutual information values after a dry year increase, as causal relationships do. However, there are also points in which both analyses depart. For example, the mutual information values for the third year are in general similar to the second, while the causal relationships showed a decrease between those years.

With respect to the spatio-temporal relationships, the mutual information values for the non-irrigation season are higher in Alarcon than in Contreras, while during the irrigation season they are similar except for the second year, in which Contreras shows higher values. This is consistent with the causal relationship analysis (see Fig 8) except for the first dry year, in which they disagree (entropy points at a higher influence from Alarcon on Contreras while the causal relationship shows the opposite). In spite of the differences found, most of the pattern changes noticed between the average behavior and a dry year found by the causal relationship analysis can be drawn from the entropy analysis. Given this, the BN-PGM for dry years can be considered valid.

## **6. Discussion and conclusions**

Bayesian Causal Modelling, evaluated from Bayesian Networks perspective, has demonstrated in this “concept-proof” to be able to identify which relationships between variables are important, as well as how they change between an average behavior and a dry year. Compared with classic statistical analysis, neither prior knowledge nor assumptions on the probability distribution of the variables involved is required. On the contrary, Bayesian Networks learn it by themselves, being also able to recursively add more information as soon as it is available (Molina et al. 2016).

The spatio-temporal dependences in the upper Jucar river sub-basins analyzed through causal reasoning have been compared with an entropy analysis. Although some differences were found, both of them agreed in most of the main patterns found in the data, as well as the modifications suffered by these for a dry year. However, CR considers the direction of each relationship. This advantage is a distinct feature of causal reasoning, making it able to determine which variables are playing the highest influence on the spatio-temporal relationships found in the system. If the influence of each variable can be

assessed and ranked with respect to the others, one can determine where investments in reducing uncertainty (e.g. increasing the monitoring or the quality of the measurements) would be efficient.

BN models developed in this research have been populated, trained and validated using historical data. This could be done due to the existence of enough records (72 years for the whole historical period and 27 dry years). However, they would exist situations in which historical records would not reach the minimum size to ensure a safe training. In this case, suitable alternatives to extend the available time series would consist in using stochastic processes like Monte Carlo simulation or stochastic models, as done by Molina et al. (2016). However, these techniques assume a specific spatio-temporal dependence structure, so they would influence the analysis of the causal relationship performed, given that the generated data would follow a pattern that may not be in line with the one intrinsic to the historical records.

Another alternative to extend the available runoff data would be the use of hydrological models, whose assumptions are less tight in terms of spatio-temporal dependences. However, these models would require enough historical records of meteorological variables (e.g. temperature and rainfall) and information about the hydrological properties of the sub-basins to be analyzed. Furthermore, hydrological models usually need more time to be set up than stochastic models.

The changes found out by BCM at the beginning of the dry period open the possibility of predicting droughts from examining the evolution of these dependences. In fact, BNs' assessment of dependence is able to determine how each variable depends on previous measurements, enabling its use as a predictor (Molina and Zazo 2017b). For the Jucar river basin in particular, the absence of spatial dependence between Alarcon and Contreras sub-basins can be used to foresee water scarcity periods and start activating drought management measures in advance.

Using the BCM defined in combination with water resource management models, one can automatically trigger these measures according to the evolution of the causal relationship. These possibilities of predictive potential will be explored in upcoming research.

From the methodology developed in this research and its application to the case study, the following conclusions can be drawn:

- Causal reasoning has been successfully applied to the upper Jucar river basin. The BN\_PGM model developed has been able to identify the dependences existing between runoff variables in both time (in-year seasons) and space (sub-basins). These dependences are similar to the ones provided by an entropy analysis and agree with the ones found by a previous statistical analysis.
- Causal reasoning identified how dependences change between average and dry years for the upper Jucar river basin. They were also able to characterize how these dependences evolve during a dry period. This ability to determine the entrance and exit of a drought enables the use of BNs as drought forecasting tool or as drought scenarios generation technique.

Additionally, this research demonstrates the suitability of BCM methodology for spatial-temporal analysis of water resource behavior of one or several basins, in a dynamic and stochastic way, because of BCM is able to reveal a logical, hidden and non-trivial structure of interdependence underlying in the hydrological historical records.

On the other hand, the spatial approach proposed in this “concept-proof” based on BCM, and focused on dry periods (droughts), may be a suitable starting point for its application to other types of extreme events such as floods, within a context of flood damage prediction.

Finally, this “concept-proof” reinforces that past information will provide prior knowledge of the future, particularly useful in water resources planning and management of highly dependent basins.

## **7. Acknowledgements**

This work was supported by the by the IMPADAPT project (CGL2013-48424-C2-1-R) with Spanish MINECO (Ministerio de Economía y Competitividad) and European FEDER funds; and the European Union’s Horizon 2020 research and innovation programme under the IMPREX project (GA 641.811).

## 8. References

- Allan, R. P., and Soden, B. J. (2008). "Atmospheric Warming and the Amplification of Precipitation Extremes." *Science*, 321(5895), 1481–1484.
- Bayes, T. (1763). "An Essay towards solving a Problem in the Doctrine of Chances." *Philosophical Transactions of the Royal Society of London*, 53, 370–418.
- Cabañas de Paz, R. (2017). New Methods and Data Structures for Evaluating Influence Diagrams. Ph.D. Thesis, University of Granada, Granada, Spain, 360 pp.
- Cain, J. (2001). Guidelines for using Bayesian Networks to support the planning and management of development programmes in the water sector and beyond. Centre for Ecology & Hydrology Wallingford, Wallingford, UK.
- Castelletti, A., Soncini-Sessa, R. (2007). "Bayesian Networks and participatory modelling in water resource management". *Environmental Modelling & Software*. 22, 1075-1088.
- Chang, B., Guan, J., and Aral, M. M. (2015). "Scientific Discourse: Climate Change and Sea-Level Rise." *Journal of Hydrologic Engineering*, 20(1), A4014003.
- CHJ. (2006). Mapa piezométrico general de la confederación hidrográfica del Júcar. In "Comprobación y evaluación en la Cuenca Piloto del río Júcar de las Guías desarrolladas en el marco de la Estrategia común para la implementación de la Directiva Marco del Agua (in S. Valencia).
- CHJ. (2007). Plan Especial de Alerta y Eventual Sequia en la Confederación Hidrográfica del Júcar (in Spanish). Valencia.
- CHJ. (2015). Plan Hidrológico de la Demarcación Hidrográfica del Júcar. Valencia.
- CHJ. (2020). Sistema de Información del Agua (SIA) de la Confederación Hidrográfica del Júcar.Red Piezométrica Operativa.
- De Michele, C., Salvadori, G. (2003). A Generalized Pareto intensity-duration model of storm rainfall exploiting 2-Copulas. *Journal of Geophysical Research-Atmospheres*. 108, 4067.

- Donat, M. G., Lowry, A. L., Alexander, L. V., O’Gorman, P. A., and Maher, N. (2016). “More Extreme Precipitation in the World’s Dry and Wet Regions.” *Nature Climate Change*, 6(5), 508–513.
- Estrela, M. J., Peñarrocha, D., and Millán, M. (2000). “Multi-annual drought episodes in the Mediterranean (Valencia region) from 1950-1996. A spatio-temporal analysis.” *International Journal of Climatology*, John Wiley & Sons, Ltd., 20(13), 1599–1618.
- Freire-González, J., Decker, C., and Hall, J. W. (2017). “The Economic Impacts of Droughts: A Framework for Analysis.” *Ecological Economics*, 132, 196–204.
- Gil, M., Garrido, A., and Gómez-Ramos, A. (2011). “Economic analysis of drought risk: An application for irrigated agriculture in Spain.” *Agricultural Water Management*, 98(5), 823–833.
- Hao, Z., and Singh, V. P. (2016). “Review of Dependence Modeling in Hydrology and Water Resources.” *Progress in Physical Geography*, 40(4), 549–578.
- Hejazi, M. I., and Cai, X. (2011). “Building more realistic reservoir optimization models using data mining - A case study of Shelbyville Reservoir.” *Advances in Water Resources*, 34(6), 701–717.
- Hipel, K. W., and McLeod, A. I. (1994). *Time Series Modelling of Water Resources and Environmental Systems*. Elsevier, Amsterdam, The Netherlands.
- Holmström, L., Ilvonen, L., Seppä, H., and Veski, S. (2015). “A Bayesian spatiotemporal model for reconstructing climate from multiple pollen records.” *Institute of Mathematical Statistics in The Annals of Applied Statistics*, 9(3), 1194–1225.
- Hurst, H.E. (1951). Long-term storage capacity of reservoirs. *Transactions of Amer.Soc.Civil Eng. (ASCE)*. 116, 770-808.
- IGME. (2020a). Map of Permeabilities of Spain scale 1/1000.000. <<https://igme.maps.arcgis.com/home/webmap/viewer.html?webmap=da8eb570845b41bbb5548c8266eaed0d>> (Accessed 01.03.20).



- IGME. (2020b). Hydrogeological map of Spain scale 1/1000.000. <  
<https://igme.maps.arcgis.com/home/item.html?id=036292dc5b8946bd979a7dc47d2f8561>> (Accessed 01.03.20).
- Jain, A., Kumar, A.M. (2007). Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*. 7, 585-592.
- Jona Lasinio, G., Sahu, S. K., and Mardia, K. V. (2005). *Modeling rainfall data using a Bayesian Kriged-Kalman model*. Southhampton, UK.
- Jyrkama, M. I., and Sykes, J. F. (2007). “The Impact of Climate Change on Spatially Varying Groundwater Recharge in the Grand River Watershed (Ontario).” *Journal of Hydrology*, 338(3–4), 237–250.
- Kahil, M. T., Dinar, A., and Albiac, J. (2015). “Modeling water scarcity and droughts for policy adaptation to climate change in arid and semiarid regions.” *Journal of Hydrology*, 522, 95–109.
- Kalra, A., Li, L., and Ahmad, S. (2013). “Improving Streamflow Forecast Lead Time using Oceanic-Atmospheric Oscillations for Kaidu River Basin, Xinjiang, China.” *Journal of Hydrologic Engineering*, 18(8), 1031–1040.
- Koller, D., and Friedman, N. (2009). *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press, Cambridge, Massachusetts, USA.
- Kousari, M. R., Hosseini, M. E., Ahani, H., and Hakimelahi, H. (2017). “Introducing an Operational Method to Forecast Long-Term Regional Drought Based on the Application of Artificial Intelligence Capabilities.” *Theoretical and Applied Climatology*, 127(1–2), 361–380.
- Lappenschaar, M., Hommersom, A., Lucas, P.J.F. (2012). Qualitative chain graphs and their use in medicine. Proceedings of the Sixth European Workshop on Probabilistic Graphical Models. Proceedings of the Sixth European Workshop on Probabilistic Graphical Models, Granada, Spain, 2012, 179-185.

- Lima, A. R., Cannon, A., and Hsieh, W. W. (2016). "Forecasting Daily Streamflow using Online Sequential Extreme Learning Machines." *Journal of Hydrology*, 537, 431–443.
- Lopez-Nicolas, A., Pulido-Velazquez, M., and Macian-Sorribes, H. (2017). "Economic risk assessment of drought impacts on irrigated agriculture." *Journal of Hydrology*, 550, 580–589.
- Macian-Sorribes, H., and Pulido-Velazquez, M. (2017). "Integrating Historical Operating Decisions and Expert Criteria into a DSS for the Management of a Multireservoir System." *Journal of Water Resources Planning and Management*, American Society of Civil Engineers, 143(1).
- Madsen, A.L., Lang, M., Kjaerulff, U.B., Jensen, F. (2003). The Hugin Tool for learning Bayesian networks. *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, Proceeding. 2711, 594-605.
- Marcos, P., Lopez-Nicolas, A., Pulido-Velazquez, M. (2017). Analysis of climate change impact on meteorological and hydrological droughts through relative standardized indices. *EGUGA.*, 1391.
- Marcos-Garcia, P., Lopez-Nicolas, A., Pulido-Velazquez, M. (2017). Combined use of relative drought indices to analyze climate change impact on meteorological and hydrological droughts in a Mediterranean basin. *Journal of Hydrology*. 554, 292-305.
- Mishra, A. K., and Desai, V. R. (2005). "Drought forecasting using stochastic models." *Stochastic Environmental Research and Risk Assessment*, 19(5), 326–339.
- Mishra, A. K., and Singh, V. P. (2010). "A review of drought concepts." *Journal of Hydrology*, 391(1–2), 202–216.
- Mishra, A. K., and Singh, V. P. (2011). "Drought modeling - A review." *Journal of Hydrology*, 403(1–2), 157–175.
- Molina, J.L., Bromley, J., García-Aróstegui, J. L., Sullivan, C., and Benavente, J. (2010). "Integrated water resources management of overexploited hydrogeological systems using Object-Oriented Bayesian Networks." *Environmental Modelling & Software*, 25(4), 383–397.

- Molina, J.L., Pulido-Velazquez, D., Luis Garcia-Arostegui, J., Pulido-Velazquez, M. (2013). Dynamic Bayesian Networks as a Decision Support tool for assessing Climate Change impacts on highly stressed groundwater systems. *Journal of Hydrology*. 479, 113-129.
- Molina, J.L., Zazo, S., Rodríguez-Gonzálvez, P., and González-Aguilera, D. (2016). "Innovative Analysis of Runoff Temporal Behavior through Bayesian Networks." *Water*.
- Molina, J.L., Zazo, S. (2017). "Causal Reasoning for the Analysis of Rivers Runoff Temporal Behavior." *Water Resour.Manage.*, 1-13.
- Molina, J.L., Zazo, S. (2018). "Assessment of Temporally Conditioned Runoff Fractions in Unregulated Rivers." *J.Hydrol.Eng.* 23, 04018015.
- Molina, J.L., Zazo, S., Martín-Casado, A.M. (2019). "Causal Reasoning: Towards Dynamic Predictive Models for Runoff Temporal Behavior of High Dependence Rivers." *Water*. 11, 877.
- Molina, J.L., Zazo, S., Martín-Casado, A., Patino-Alonso, M. (2020). Rivers' Temporal Sustainability through the Evaluation of Predictive Runoff Methods. *Sustainability*. 12, 1720.
- Ochoa-Rivera, J.C. (2008). Prospecting droughts with stochastic artificial neural networks. *Journal of Hydrology*. 352, 174-180.
- Pearl, J. (1988). *Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Francisco, USA.
- Pearl, J. (2009). *Causality: models, reasoning and inference*. Cambridge University Press, New York City, NY, USA.
- Pearl, J. (2014). "Graphical Models for Probabilistic and Causal Reasoning." *Computing Handbook, Third Edition: Computer Science and Software Engineering*, T. Gonzalez, J. Diaz-Herrera, and A. Tucker, eds., Chapman & Hall / CRC, Boca Raton, Florida, USA.

- Peng, T., Zhou, J., Zhang, C., Fu, W., (2017). Streamflow Forecasting Using Empirical Wavelet Transform and Artificial Neural Networks. *Water*. 9, 406.
- Said, A. (2006). The implementation of a Bayesian network for watershed management decisions. *Water Resour.Manage.* 20, 591-605.
- Salas, J. D., Delleur, J. W., Yevjevich, V. M., and Lane, W. L. (1980). *Applied Modeling of Hydrologic Time Series*. Water Resources Publications, Littleton, Colorado, USA.
- Sarabia-Alzaga, J., Gómez-Déniz, E. (2008). Construction of multivariate distributions: a review of some recent results. *Statistics and Operations Research Transactions*. 32, 3-36.
- See, L., Openshaw, S. (2000). A hybrid multi-model approach to river level forecasting. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*. 45, 523-536.
- Singh, V. P. (2011). “Hydrologic Synthesis using Entropy Theory: Review.” *Journal of Hydrologic Engineering*, 16(5), 421–433.
- Sperotto, A., Molina, J.L., Torresan, S., Critto, A., and Marcomini, A. (2017). “Reviewing Bayesian Networks potentials for climate change impacts assessment and management: A multi-risk perspective.” *Journal of Environmental Management*, Academic Press, 202, 320–331.
- Spirtes, P. (2010). Introduction to Causal Inference. *Journal of Machine Learning Research*. 11, 1643-1662.
- Trenberth, K. (2011). “Changes in Precipitation with Climate Change.” *Climate Research*, 47(1–2), 123–138.
- Vicente-Serrano, S. M., and Cuadrat-Prats, J. M. (2007). “Trends in drought intensity and variability in the middle Ebro valley (NE of the Iberian peninsula) during the second half of the twentieth century.” *Theoretical and Applied Climatology*, 88(3), 247–258.
- Vicente-Serrano, S. M., González-Hidalgo, J. C., de Luis, M., and Ravents, J. (2004). “Drought patterns in the Mediterranean area:” *Climate Research*, Inter-Research Science Center, 26(1), 5–15.

- Vogel, K., Riggelsen, C., Merz, B., Kreibich, H., and Scherbaum, F. (2012). "Flood damage and influencing factors: A Bayesian network perspective." In Proceedings of the 6th European workshop on probabilistic graphical models (PGM 2012), University of Granada, 625, Granada, Spain, pp. 347-354.
- Vogel, K., Riggelsen, C., Scherbaum, F., Schröter, K., Kreibich, H., and Merz, B. (2013). "Challenges for Bayesian network learning in a flood damage assessment application." In G. Deodatis, B. R. Ellingwood, and D. M. Frangopol (Eds.), Safety, reliability, risk and life-cycle performance of structures and infrastructures. London: CRC Press, pp. 3123-3130.
- Vogel, K., Riggelsen, C., Korup, O., and Scherbaum, F. (2014). "Bayesian network learning for natural hazard analyses." *Natural Hazards and Earth System Science*, 14(9), 2605-2626. <https://doi.org/10.5194/nhess-14-2605-2014>.
- Vogel, K., Weise, L., Schröter, K., and Thielen, A.H. (2018). "Identifying Driving Factors in Flood-Damaging Processes Using Graphical Models." *Water Resour.Res.* 54, 8864-8889.
- Wang, W., Chau, K., Cheng, C., and Qiu, L. (2009). "A Comparison of Performance of several Artificial Intelligence Methods for Forecasting Monthly Discharge Time Series." *Journal of Hydrology*, 374(3-4), 294-306.
- Wasko, C., and Sharma, A. (2015). "Steeper Temporal Distribution of Rain Intensity at Higher Temperatures within Australian Storms." *Nature Geoscience*, 8(7), 527-U166.
- Wei, S., Zuo, D., Song, J. (2012). Improving prediction accuracy of river discharge time series using a Wavelet-NAR artificial neural network. *J.Hydroinf.* 14, 974-991.
- Wikle, C. K., Berliner, L. M., and Cressie, N. (1998). "Hierarchical Bayesian space-time models." *Environmental and Ecological Statistics*, 5, 117-154.
- Wu, F., Zhan, J., Su, H., Yan, H., and Ma, E. (2015). "Scenario-Based Impact Assessment of Land use/Cover and Climate Changes on Watershed Hydrology in Heihe River Basin of Northwest China." *Advances in Meteorology*, 410198.

- Yang, T., Asanjan, A. A., Faridzad, M., Hayatbini, N., Gao, X., and Sorooshian, S. (2017a). "An Enhanced Artificial Neural Network with a Shuffled Complex Evolutionary Global Optimization with Principal Component Analysis." *Information Sciences*, 418, 302–316.
- Yang, T., Asanjan, A. A., Welles, E., Gao, X., Sorooshian, S., and Liu, X. (2017b). "Developing Reservoir Monthly Inflow Forecasts using Artificial Intelligence and Climate Phenomenon Information." *Water Resources Research*, 53(4), 2786–2812.
- Yang, T., Gao, X., Sorooshian, S., and Li, X. (2016). "Simulating California reservoir operation using the classification and regression-tree algorithm combined with a shuffled cross-validation scheme." *Water Resources Research*, 52(3), 1626–1651.
- Zazo S., Macian-Sorribes H., Sena-Fael C.M., Martín-Casado, A.M., Molina J.L. and Pulido-Velazquez, M. (2019). "Qualitative approach for assessing runoff temporal dependence through geometrical symmetry". (Contributed Paper). Proceedings International Congress on Engineering. Engineering for Evolution (ICEUBI2019). 27 - 29 November 2019, Covilhã, Portugal.
- Zazo, S. (2017). Analysis of the Hydrodynamic Fluvial Behaviour through Causal Reasoning and Artificial Vision. Ph.D. Thesis, University of Salamanca, Ávila, Spain, 159 pp.