



HAL
open science

Bayesian modelling of water resource stress and drought monitoring in southeastern France

H. Di Costanzo, Philippe Martin

► **To cite this version:**

H. Di Costanzo, Philippe Martin. Bayesian modelling of water resource stress and drought monitoring in southeastern France. First Ozcar - Tereno International Conference, Oct 2021, Strasbourg, France. hal-03426959

HAL Id: hal-03426959

<https://hal.science/hal-03426959>

Submitted on 12 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Expert Decision Support System for drought monitoring in Cevennes (French South-East): a use case of Bayesian Networks

Scientific and operational problems

In the Mediterranean region, a strong increase in temperatures and a decrease in precipitation in certain areas was observed during the 20th century (Norrant and Douguedroit, 2006; Stagge and al., 2017). These trends should be confirmed at the end of the 21st century (Milano and al., 2013). **In theory, a decrease in the quantity of water available in the Mediterranean watersheds is expected. In addition, the population is growing. A gradual increase in the pressure on water resources during the 21st century is to be feared.**

The **Cevennes** (figure 1) is a pilot area for studying drought in small and medium-sized Mediterranean watersheds (between 50 and 300 km²). It is a territory subject to hydrological extremes (Herring and al., 2015). The literature on Cevennes floods is abundant, but not for droughts. However, **low water** in this area can be particularly **severe** (Martin and al., 2015, 2016, 2017, 2018, 2019) due to the small size of the aquifer reserves (mineral geological formation with low permeability). A prolonged period of low water can affect agricultural yields

and the supply of drinking water. **Tourism** implies an increase in the local population that can **double** during the **summer**. In fact, The Cevennes are exposed to water stress problems that lead to conflicts between the various water users in the basin.

In this PHD project, we propose an **operational methodological approach for monitoring drought in a Mediterranean watershed**. It implies to consider both the available **water resource** and the **anthropogenic demand** (Charre, 1977; Martin and al., 2020) *i.e.* to reason on **water stress** (environmental supply versus social demand). In a context of **uncertainty** on the measurement (quality, gaps, spatial representativity...) and **lack of knowledge** of the studied system (structuring element and causal relationship), Bayesian networks are an adapted tool for the management of water resources (Phan and al., 2016) and environmental risk in a more general way (Sperotto and al., 2017). They allow the integration of different types of knowledge (variables and causal relations) in the form of **probability distributions**. This **probabilistic approach** is proposed as an alternative to the calculation of the water balance ($Q + ETR = P \pm R$) which is difficult to estimate for intermediate size basins.

The use of Bayesian networks is recent in environmental studies (Kaikkonen et al., 2021). They have mainly been used to compare the **effectiveness** of environmental management policies, **evaluate** the effects of planning on a territory, explore **alternative policies**, or determine the weight of **uncertainties on estimates**. The technical approaches differ according to the objectives (inference, classification, regression). In the case of drought, we use an influence network for the decision making approach. A **Bayesian Decision Network (BDN)** is an efficient tool for incorporating different aspects related to **decision making**, *i.e.* analyzing the expected effects of a management policy while taking into account the uncertainties of each variable.

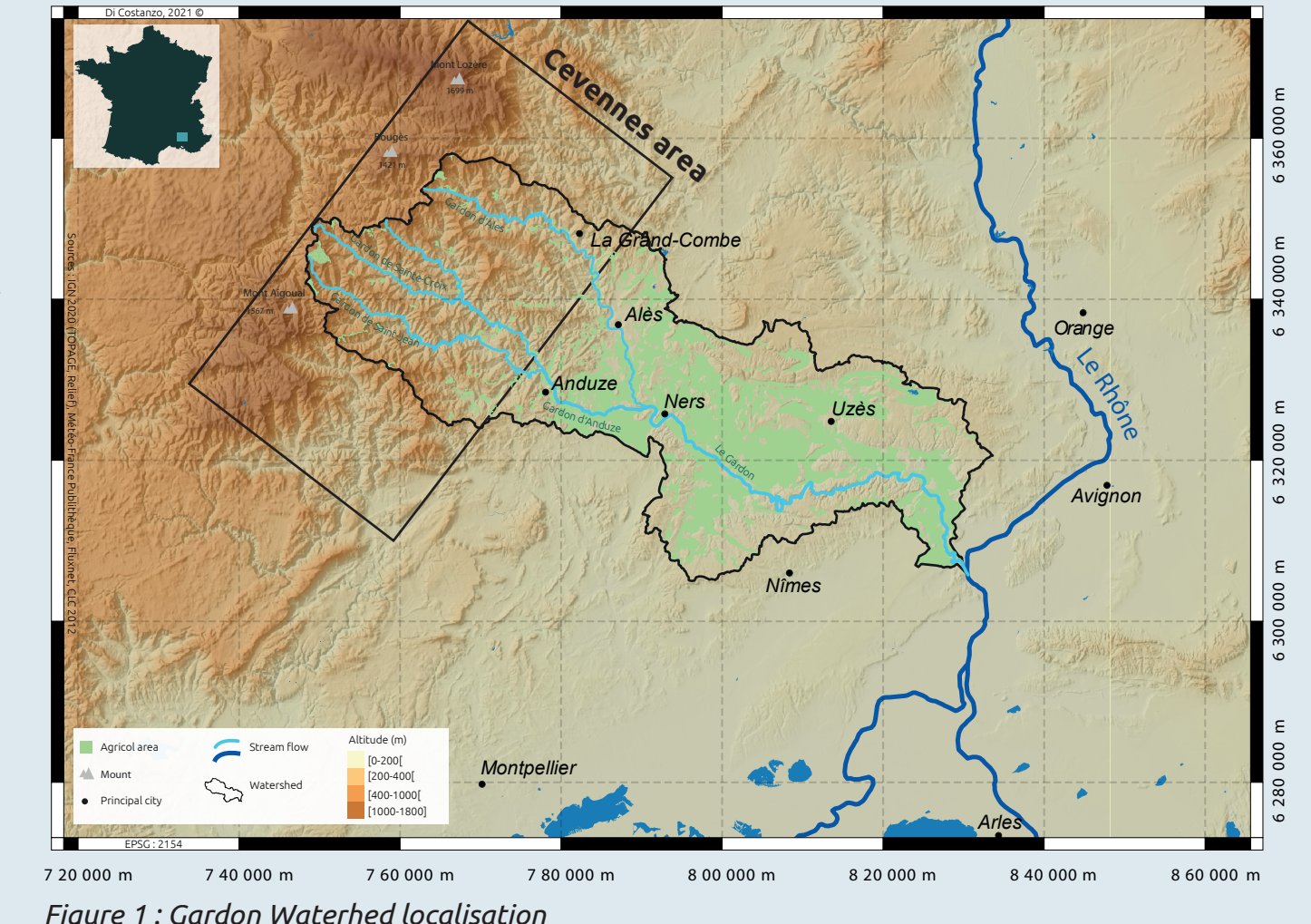


Figure 1 : Gardon Watershed localisation

Data and methods

Different datasets (figure 2), centralized in a Spatial Relational Database (SRDB), provide representative information on the different states of **hydraulic deficit** (atmospheric, edaphic, hydrological drought) and on the **social demand** in the basin. Data is aggregated at basin scale.

A **Bayesian network** (Pearl, 2008) is defined as a **multivariate statistical model** composed of a set of variables $X = \{X_1, \dots, X_n\}$. It consists of two components :

- A qualitative which correspond to a directed acyclic graph where each vertex corresponds to a variable of the model and each link materializes a **dependence** between the vertices.
- A quantitative one which defines the **conditional distributions** $p(x_i|pa(x_i))$ for each variable X_i , given its parents in the graph, noted $pa(x_i)$. Taking into account the independence encoded by the acyclic graph, the joint distribution on all variables is equal to the product of the conditional distributions attached to each node. This calculation is based on **Bayes' theorem**:

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

Where $P(A)$ and $P(B)$ are the observed probabilities of the independent nodes A and B; $P(B|A)$ is the conditional probability of B knowing A; and $P(B|A)/P(B)$ the Bayes or likelihood factor.

Chen and Pollino (2012) details the different steps required in the model building process. The objective of the proposed model is to allow an **estimation of the water stress** from disponibilities data on the watershed. Among the great variety of drought indicators already existing (Zargar and al., 2011), we have selected one or more indicators of atmospheric, edaphic, hydrological drought adapted to our sector to estimate its water status (figure 2). The data that can approximate the social demand for water is scarce, it is for the most part determined by **expert opinion**. We have chosen to use a **hybrid BDN**. This technical choice allows to take into account the diversity of the distribution families of the variables used and to model the statistical relations between these different distribution families. The structure of the model was built **in an expert way**. The relationships between measured dependent variables are statistically determined by generalized linear models. 2015-2020 is the period when all the measured data on the basin are available. The cross-validation method is used on this period to calibrate and validate the robustness of the model over time.

The computation of conditional distributions is done with the Rjags package (<https://CRAN.R-project.org/package=rjags>). Jags is a program allowing to approximate conditional distributions (Markov Chain Monte Carlo algorithm) in the case of presence of discrete and continuous variables in the network. The Hydenet package

(<https://CRAN.R-project.org/package=HydeNet>) is used to parameterize the decision network.

Drought types

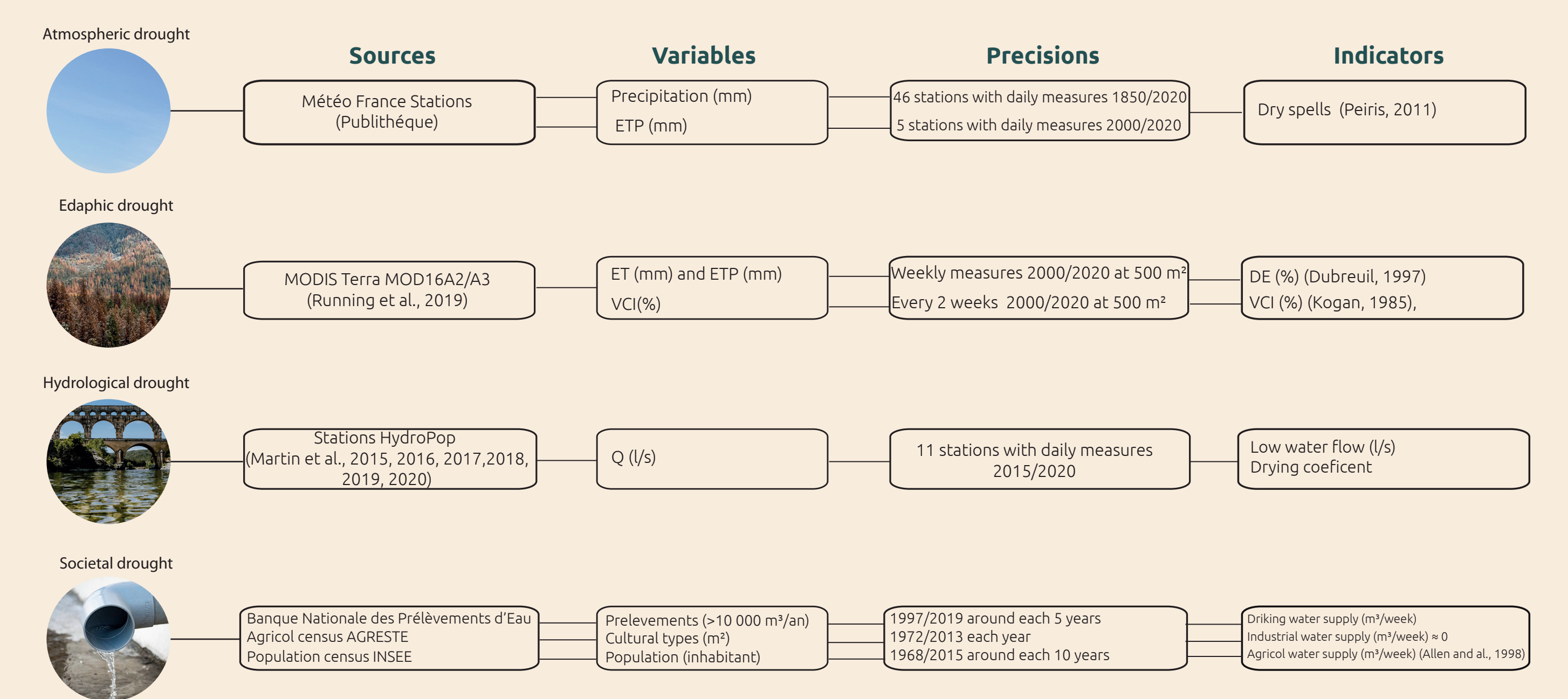


Figure 2 : data and indicators used. With Q water flow; ET Evapotranspiration; ETP Potential Evapotranspiration; VCI Vegetation Condition Index; DE Deficit Evaporation.

First résultats

Two branches of the model (figure 4) are to be distinguished : **natural resources** (utility node WR) and **anthropogenic needs** (utility node AC) for the calculation of water stress. **Utility node** correspond to a variable who want to minimize or maximise the value in the network.

Caliano et al (2017) distinguish **three types of water use**: industrial, agricultural and water supply. As the area is no industrialized, we make the assumption that water consumption for industrial activities are equal to zero. The FAO crop coefficients method (Allen and al., 1998) is used to estimate the amount of water needed (the agricultural demand) in the case of an optimal crop yield (AGRESTE 2010 census). We simplify the drinking water demand by the population variable. Anthropogenic consumption is approximated by the joint distribution of the three types of estimated consumption.

The water resource is approximated by the joint distribution of indicators describing the water state of each component of the water cycle. A water deficit state is identified when all the indicators estimate a water deficit situation (extreme values of the distributions).

The decision node "Arrete" allows to simulate the effect of the implementation of the abstraction restrictions in the basin (French arrêté sécheresse). The network allows to propose which level of restriction (0 to 3) should be implemented to minimize the tension on the water resource according to the situation. The weekly decision node allows to integrate the temporal dimension in the monitoring of the stress on the water resource.

The Fresh Water System Performance and Agricol Water System Performance nodes were determined by expert opinion (Plan Gestion de la ressource en eau de EPTB Gardon). The credibility of these values is modeled by a probability distribution.

The drought matrix (Figure 3) is a way of representing the trajectories of water stress in the Cevennes socio-environmental system over the year. It allows a quick reading of the evolution of the water stress situation for basin managers.

Figure 3 : Drought matrix in hypothetical case of water resource joint distribution and anthropogenic consumption joint distribution are normal.

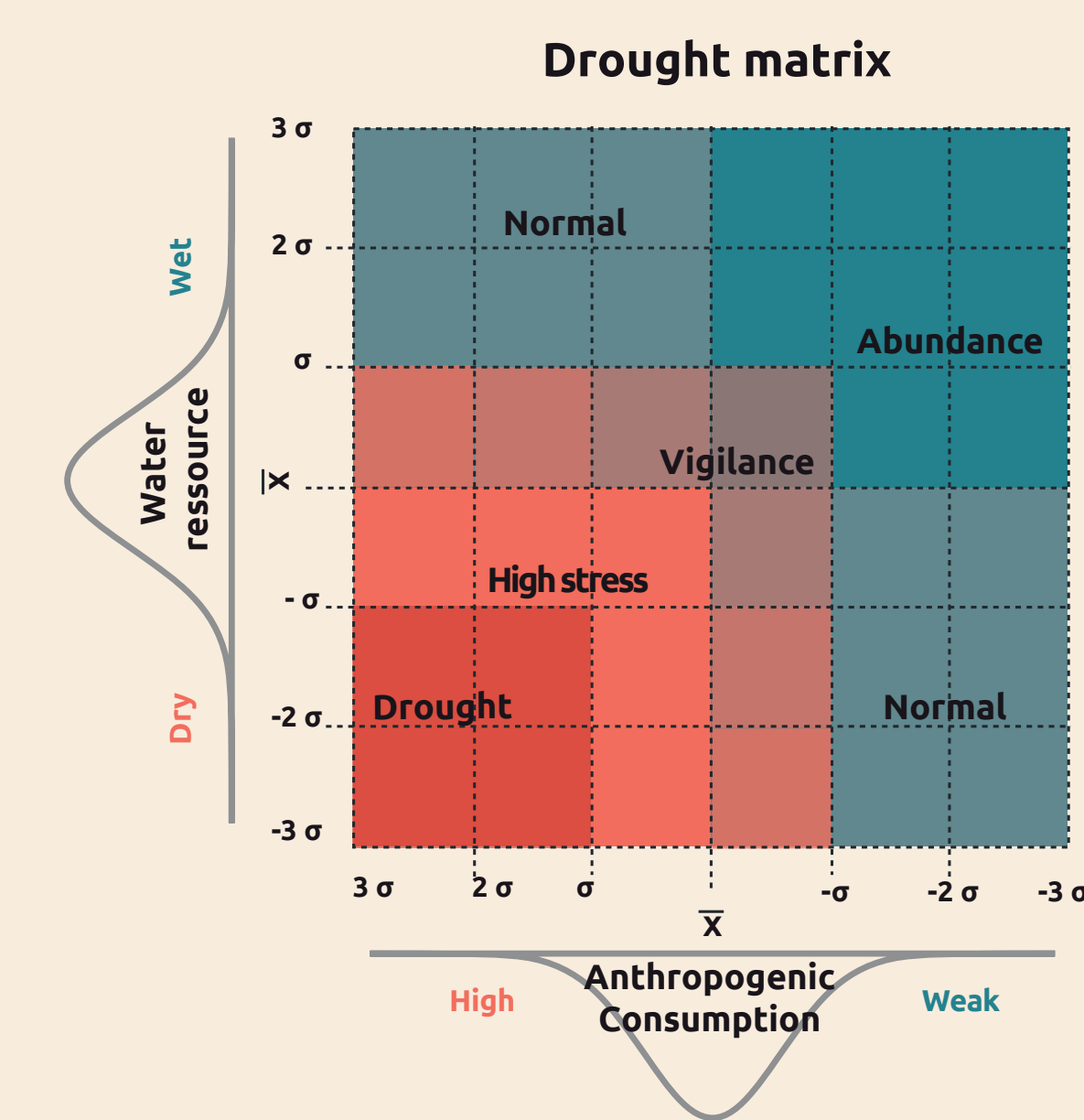
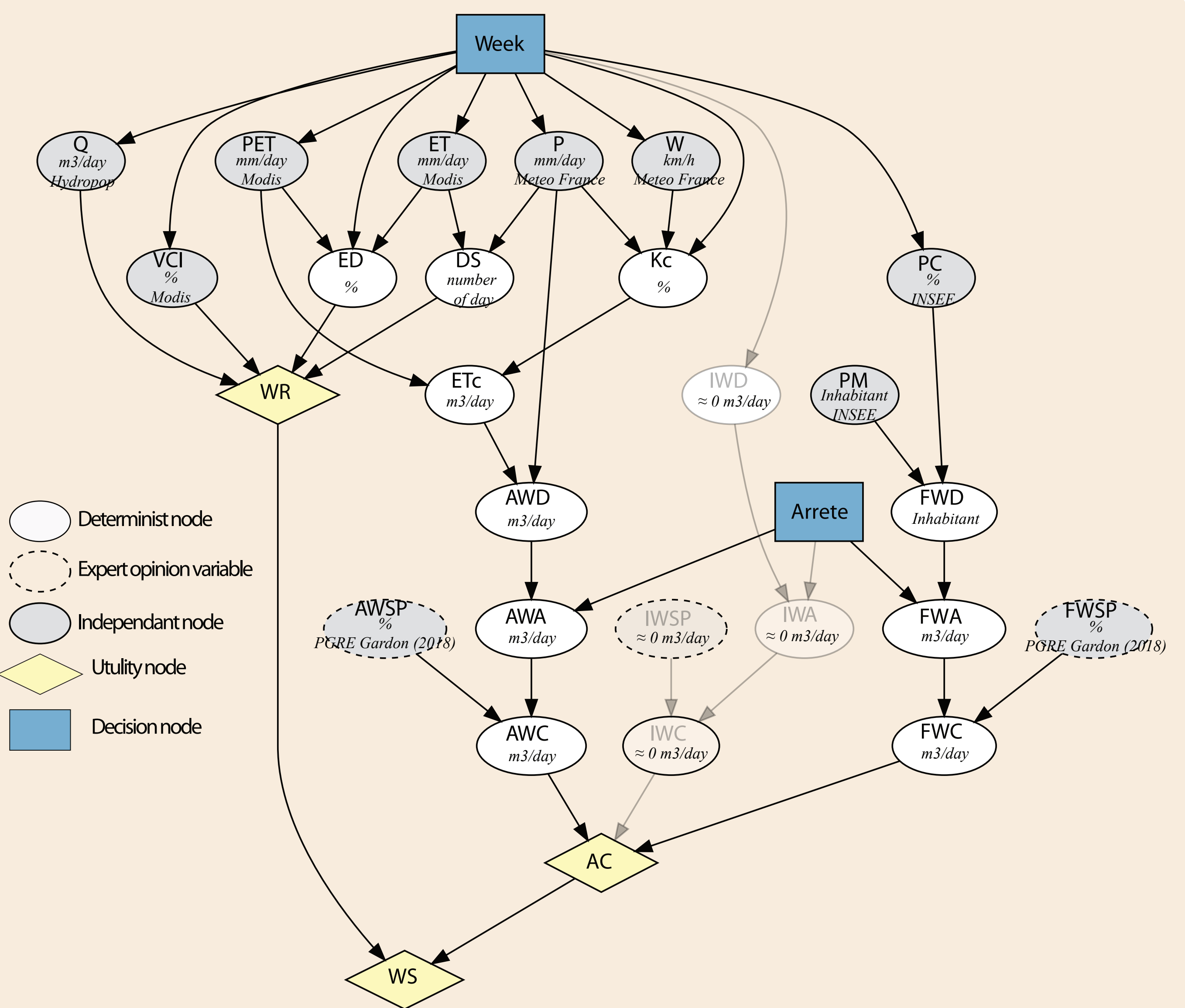


Figure 4 : First BDN conceptual proposition of water stress in up-stream Gardon Watershed. In order for each node : variable code, unit and source. With Q is water flow, PET is potential evapotranspiration, ET is evapotranspiration, P is cummul precipitation, W is the median wind speed, VCI is Vegetation Condition Index, ED is evaporation deficit, DS is dry spell, Kc is cultural coefficient, WR is water resource, ETC agricultural plant transpiration, PC week variation coefficient of population, PM population of municipalities, AWD is agricultural water demand, IWD is industrial water demand, FWD is fresh water demand, AWSP is agricultural water system performance, IWSP is industrial water system performance, FWSP is fresh water system performance, AWA is agricultural water abstraction, IWA industrial water abstraction, FWSP is fresh water abstraction, AWC is agricultural water consumption, FWC is fresh water consumption, IWC is industrial water consumption, AC is Anthropogenic consumption and WS is water stress.



First conclusions, some limitation and perspectives

The proposed BDN allows a conceptual formalization of the complex calculation of water stress. This estimation method is **exploratory and network form could be upgrade**. It presents the variables necessary for its calculation, including those for which no measurements are available (expert opinion). Each variable is associated with a **credence interval**. This leads to reasoning with **probabilistic reasoning**. The **quality and precision** of BDN estimates should tend to **improve** with the increase in the temporal length of the data and the addition of new data sets that will eventually replace the distributions determined by expert opinion. When the calibration is good, this type of model has the advantage of estimating the **missing elements** in the calculation knowing the elements available in the network (inference). It allows, according to the values taken by the decisional nodes, to **give suggestions** on the water resource management policies of the watershed. In the case of the drought in France, the "arrêté sécheresse" is a policy of control of the abstraction water. Some statistical relationships between indicators of

water deficit monitoring need to be further explored to improve the inference of the environmental branch of the BDN. The temporal hindsight on hydrometric data limits the quality of water deficit estimation for surface rivers at the moment. An estimate of groundwater hydrologic drought is absent. No piezometers representative of the area and measuring over the long term are available. In Cevenne context ("natural rivers"), we have exclu in the model big hydraulic aménagement impact, industrial impact and water quality. Agricol water need estimated by FAO method overestimate probably the demand and calcul repose on larges simplificatrice hypotheses of plant developpement and on a approximate surface of cultures types. In case the robustness of the model is validated, **environmental scenarios** (effect of climate change on the mode and dispersion of the distribution laws of the indicators) and **social evolution** (increase in population, tourist attractiveness...) will be engaged.

References

. Allen, R.G., Pereira, L.S., Raes, D., Smith, M. (1998) Crop evapotranspiration—guidelines for computing crop water requirements. FAO Irrigation and drainage paper 56. Food and Agriculture Organization, Rome.
 . Charre, J. A propos de sécheresse. geocar 52, 215–226 (1977).
 . Chen, S. H. & Pollino, C. A. Good practice in Bayesian network modelling. Environmental Modelling & Software 37, 134–145 (2006).
 . Dubreuil, V. La sécheresse dans la France de l'ouest : une contrainte climatique trop souvent oubliée. Sécheresse 8, 47–55 (1997).
 . Herring, S. C., Hoerling, M. P., Kossin, J. P., Peterson, T. C. & Stott, P. A. Extreme fall 2014 precipitation in cevennes mountains. 12.
 . Martin, P. et al. L'étage 2015 dans le Gardon amont (Cevennes). 55 (2015).
 . Martin, P., Didon-Lescot, J.-F., Ayrat, P.-A., Domergue, J.-M. & Grard, N. L'étage 2016 dans le Gardon amont (Cevennes). 63 (2016).
 . Martin, P., Didon-Lescot, J.-F., Ayrat, P.-A., Domergue, J.-M. & Grard, N. L'étage 2017 dans les Gardons amont (Cevennes). Projet pilote de suivi des débits d'étiages. 3eme rapport sur huit stations temporaires d'étiage et deux stations pérennes. 59 (2017).
 . Martin, P. et al. L'étage 2018 dans les Gardons amont (Cevennes). Projet pilote de suivi des débits d'étiage. Premier rapport sur huit stations temporaires d'étiage et deux stations pérennes. 71 (2018).
 . Martin, P. et al. L'étage 2019 dans les Gardons amont (Cevennes). Projet pilote de suivi des débits d'étiage. Rapport sur huit stations temporaires d'étiage et deux stations pérennes. 52 (2019).
 . Milano, M. et al. Current state of Mediterranean water resources and future trends under climatic and anthropogenic changes. Hydrological Sciences Journal 58, 498–518 (2013).
 . Norrant, C. & Douguedroit, A. Monthly and daily precipitation trends in the Mediterranean (1950–2000). Theor. Appl. Climatol. 83, 89–106 (2006).
 . Kaikkonen, L., Parviainen, T., Rahikainen, M., Uusitalo, L. & Lehtikoinen, A. Bayesian Networks in Environmental Risk Assessment: A Review. Integr. Environ. Assess. Manag. 17, 62–78 (2021).
 . Kogan, F. N. Droughts of the Late 1980s in the United States as Derived from NOAA Polar-Orbiting Satellite Data. Bulletin of the American Meteorological Society 76, 14 (1995).
 . Pearl, J. Probabilistic reasoning in intelligent systems: networks of plausible inference. (Kaufmann, 2008).
 . Peiris, T. Critical Evaluation of Dry Spell Research. 11, 8 (2011).
 . Phan, T. D., Smart, J. C. R., Capon, S. J., Hadwen, W. L. & Sahin, O. Applications of Bayesian belief networks in water resource management: A systematic review. Environmental Modelling & Software 85, 98–111 (2016).
 . Running, S. W., Mu, Q., Zhao, M. & Moreno, A. User's Guide MODIS Global Terrestrial Evapotranspiration (ET) Product (MOD16A2/A3 and Year-end Gap-filled MOD16A2GF/A3GF). (2019).
 . Sperotto, A., Molina, J.-L., Torresan, S., Critto, A. & Marcomini, A. Reviewing Bayesian Networks potentials for climate change impacts assessment and management: A multi-risk perspective. Journal of Environmental Management 202, 320–331 (2017).
 . Stagge, J. H., Kingston, D. G., Tallaksen, L. M. & Hannah, D. M. Observed drought indices show increasing divergence across Europe. Scientific Reports 7, (2017).