



**HAL**  
open science

# Revealing Hidden Climate Indices from the Occurrence of Hydrologic Extremes

Benjamin Renard, M. Thyer

► **To cite this version:**

Benjamin Renard, M. Thyer. Revealing Hidden Climate Indices from the Occurrence of Hydrologic Extremes. *Water Resources Research*, 2019, 55 (9), pp.7662-7681. 10.1029/2019WR024951 . hal-02610059

**HAL Id: hal-02610059**

**<https://hal.inrae.fr/hal-02610059v1>**

Submitted on 16 Sep 2020

**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.

# Water Resources Research

## RESEARCH ARTICLE

10.1029/2019WR024951

### Key Points:

- A Bayesian hierarchical model is proposed to identify hidden climate indices from the occurrence of hydrologic extremes
- Synthetic and real-life case studies illustrate the potential of the model to identify the climate drivers of hydrologic extremes
- In regions where standard climate indices are poor predictors, hidden climate indices can reliably describe space-time extreme occurrences

### Supporting Information:

- Supporting Information S1
- Data Set S1
- Data Set S2
- Data Set S3
- Figure S1
- Figure S2
- Figure S3
- Figure S4
- Figure S5
- Figure S6

### Correspondence to:

B. Renard,  
benjamin.renard@irstea.fr

### Citation:

Renard, B., & Thyer, M. (2019).  
Revealing hidden climate indices from  
the occurrence of hydrologic extremes.  
*Water Resources Research*, 55, 7662–7681.  
<https://doi.org/10.1029/2019WR024951>

Received 7 FEB 2019

Accepted 19 AUG 2019

Accepted article online 21 AUG 2019

Published online 6 SEP 2019

## Revealing Hidden Climate Indices from the Occurrence of Hydrologic Extremes

Benjamin Renard<sup>1,2</sup>  and Mark Thyer<sup>2</sup> 

<sup>1</sup>Irstea, UR Riverly, Lyon, France, <sup>2</sup>School of Civil, Environmental and Mining Engineering, University of Adelaide, Adelaide, South Australia, Australia

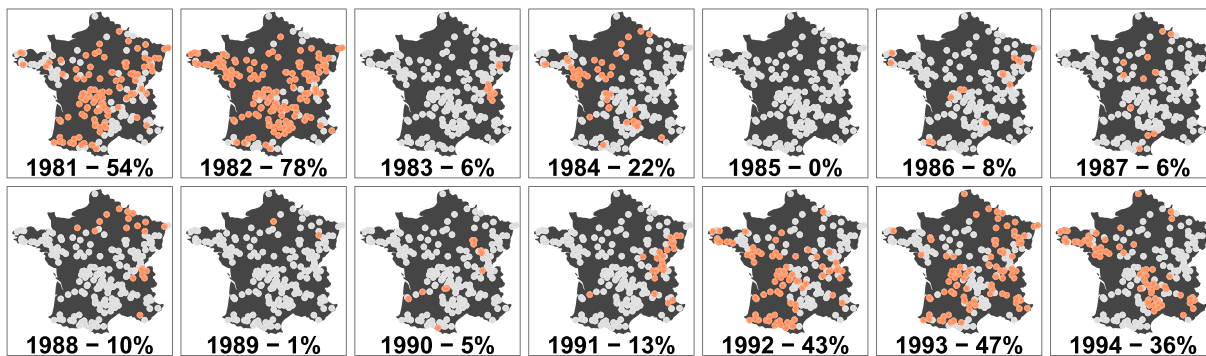
**Abstract** Describing the space-time variability of hydrologic extremes in relation to climate is important for scientific and operational purposes. Many studies demonstrated the role of large-scale modes of climate variability such as the El Niño–Southern Oscillation (ENSO) or the North Atlantic Oscillation (NAO), among many others. Climate indices have hence frequently been used as predictors in probabilistic models describing hydrologic extremes. However, standard climate indices such as ENSO/NAO are poor predictors in some regions. Consequently, this paper describes an innovative method to avoid relying on standard climate indices, based on the following idea: the relevant climate indices are effectively unknown (they are *hidden*), and they should therefore be estimated directly from hydrologic data. In statistical terms, this corresponds to a Bayesian hierarchical model describing extreme occurrences, with hidden climate indices treated as latent variables. This approach is illustrated using three case studies. A synthetic case study first shows that identifying hidden climate indices from occurrence data alone is feasible. A second case study using flood occurrences at 42 east Australian sites confirms that the model correctly identifies their ENSO-related climate driver. The third case study is based on 207 sites in France, where standard climate indices poorly predict flood occurrence. The hidden climate indices model yields a reliable description of flood occurrences, in particular their clustering in space and their large interannual variability. Moreover, some hidden climate indices are linked with specific patterns in atmospheric variables, making them interpretable in terms of climate variability and opening the way for predictive applications.

## 1. Introduction

### 1.1. Space-Time Variability of Hydrologic Extremes

Managing hydrologic hazards such as floods and intense precipitation requires a good knowledge of their space-time variability. As an illustration, consider the occurrence of flood events at 207 stations in France shown in Figure 1. This figure illustrates several key properties of how hydrologic extremes occur in space and time. First, they tend to cluster in space, occurrences forming coherent spatial patterns (Merz et al., 2018). Figure 1 also suggests a possible tendency to cluster in time, with flood-rich periods (1981–1982 and 1992–1994) alternating with flood-poor periods (1983–1991; Kiem et al., 2003; Merz et al., 2016; G. M. McMahon & Kiem, 2018). As a consequence of these properties, the spatial extent of extremes varies very irregularly from year to year: on any given year, the percentage of stations affected by an extreme tends to be either very low or very high, but is rarely close to its interannual average of 20% (corresponding to the 5-year flood threshold). In layman's (and exaggerated) terms, managing flood hazard at this spatial scale looks like an alternation of long runs of problem-free years and shorter runs of catastrophic years, rather than a smooth sail through average years.

The irregularity described above poses great challenges to socio-economic sectors that need to manage hydrological risk at a national or even continental scale. For instance, international (re)insurance companies need to secure enough capital to withstand occasional huge losses (European Environment Agency, 2010). Alternatively, disaster response institutions (civil protection and humanitarian aid) need to deploy their actions over a large spatial extent and are hence calling for early warning systems that may facilitate logistics and trigger early actions (Coughlan de Perez et al., 2016). Beyond these two specific examples, a recent World Bank report (World Bank, 2014) emphasizes the importance of understanding the space-time variability of extremes and its relation with climate.



**Figure 1.** Autumn flood occurrence in France, 1981–1994. An occurrence (orange dot) corresponds to a flood exceeding the local 5-year autumn flood (empirical estimate based on at least 40 years of data). In each panel, the percentage of stations with a flood occurrence is shown.

### 1.2. The Usual Suspects: ENSO, NAO, and Other Modes of Climate Variability

In many areas of the world, hydrologic extremes are influenced by modes of climate variability such as the El Niño–Southern Oscillation (ENSO) or the North Atlantic Oscillation (NAO), among many others. Since the latter tend to have consistent effects at the regional scale, they are good candidates to explain the variability of hydrologic extremes in both time and space. Several authors provide global-scale overviews of the effect of ENSO/NAO on hydrologic extremes: see, for example, Kenyon and Hegerl (2010) and Sun, Renard, et al. (2015) for extreme precipitation or Ward et al. (2014) and Emerton et al. (2017) for floods. In addition, many studies demonstrate the effect of ENSO/NAO or other modes of climate variability at a smaller regional scale. While a comprehensive review lies beyond the scope of this paper, examples are provided in the following sections.

### 1.3. Probabilistic Modeling of Extremes Using Standard Climate Indices

In practice, modes of climate variability are quantified by means of climate indices, which are one-dimensional time series typically resulting from a dimensionality reduction exercise (e.g., Barnston & Livezey, 1987). A large number of climate indices have been proposed in the literature (see, e.g., National Center for Atmospheric Research, 2019). Throughout this paper, they are collectively referred to as “standard climate indices” (SCIs).

Since a SCI is a simple time series, it can conveniently be used as a covariate in a climate-informed frequency analysis model, where the distribution of hydrologic extremes is assumed to change according to the value taken by the SCI. This is typically achieved by assuming that some parameters of this distribution are a function of the SCI (see, e.g., the generalized additive models for location, scale, and shape of Stasinopoulos & Rigby, 2007). Examples of climate-informed flood frequency analysis models include (among many others) Sankarasubramanian and Lall (2003), Kwon et al. (2008), López and Francés (2013), or Lima et al. (2015). These at-site models have been extended to the regional scale, typically by making additional assumptions on the spatial variability of the parameters (including those controlling SCI effects; e.g., Aryal et al., 2009; Gregersen et al., 2013; Lima & Lall, 2010; Steinschneider & Lall, 2015; Sun et al., 2014; X. Sun & Lall, 2015; Sun, Lall, et al., 2015; Sun, Renard, et al., 2015).

### 1.4. Hidden Climate Indices

The examples provided in the previous section illustrate that SCIs can often be used as skillful predictors of hydrologic extremes. However, SCIs also have limitations. For instance, they may focus on capturing a mode of climate variability that is not relevant for hydrologic predictands in the target study region (Lavers et al., 2013). The combined use of several SCIs should account for the fact that distinct SCIs may be strongly related to each other (e.g., Westra et al., 2008; Westra & Sharma, 2009; Westra et al., 2015). Most notably, several authors reported that all SCIs are poor predictors in some regions and seasons (e.g., Giuntoli et al., 2013; Grantz et al., 2005; Renard & Lall, 2014). This does not imply that hydrologic extremes are unrelated to climate but rather that this relation cannot be expressed through SCIs for such region/season combinations. By restricting to SCIs, one may hence miss a

genuine relation with climate; in a sense, this is akin to looking for lost keys under a lamppost because the light is better here.

Consequently, alternative methods have been proposed to avoid relying on a predefined list of off-the-shelf SCIs, based on the idea that the relevant climate indices may effectively be unknown (i.e., they may be hidden). A hidden climate index (HCI) affecting the target hydrologic predictand data should leave a trace in it, and should therefore be identifiable directly from the hydrologic predictand. For instance, Grantz et al. (2005), Zeng et al. (2017), or Schlef et al. (2018) defined “customized” predictors (that can be viewed as HCIs) by means of some correlation analysis between the hydrologic predictand and large-scale climate fields (e.g., sea surface temperatures or atmospheric pressures).

The notion of hidden climate indices can also be embedded into a formal statistical model: for instance, hidden Markov models postulate the existence of a hidden climate state (e.g., a binary wet/dry indicator) that drives the distribution of the hydrologic predictand (Bracken et al., 2016; Thyer & Kuczera, 2000, 2003a, 2003b): this climate state can be interpreted as a discrete HCI. Renard and Lall (2014) proposed a Bayesian hierarchical model to extend this approach to a continuous HCI, and highlighted much more skillful predictions than traditional approaches using SCIs.

The model proposed by Renard and Lall (2014) was based on the assumption that the HCI to be identified has a common effect for all target sites: this restricts its application to small regions where the effect of climate variability can be considered as homogenous. Ahn et al. (2017) relaxed this assumption by means of a parsimonious description of the variation of the HCI effect in space. They assumed that the HCI effect is maximal at some unknown location of the study area, and then decreases as a deterministic function of the distance to this location. While this parsimonious model allows considering larger, moderately heterogeneous regions, it is still not applicable at a large spatial scale (national, continental or even global). Indeed, climate effects are unlikely to follow the deterministic pattern used by Ahn et al. (2017) at such a large spatial scale. Moreover, a single HCI may not be sufficient to explain the variability of hydrologic extremes. There is therefore a need to develop a probabilistic HCI model that could be applied at a large spatial scale. This entails allowing the identification of several HCIs whose effects may realistically vary in space: the objective of this paper is precisely to fill this research gap.

### 1.5. Detailed Objectives and Outline of the Paper

This paper describes and evaluates a Bayesian hierarchical model describing the space-time variability of extreme occurrences at a large spatial scale. This development pursues the following specific objectives:

1. Identifying (several) HCIs that drive the temporal variability of extreme occurrences;
2. Proposing a realistic description of the spatial variation of the HCI effects at a large spatial scale;
3. Deriving an efficient inference strategy to estimate parameters and their uncertainties;
4. Testing the internal consistency of the model using a synthetic case study;
5. Applying the model to real-life data sets.

The remainder of this paper is organized as follows. Section 2 describes the probabilistic model and the inference strategy used to estimate its parameters. Section 3 describes a synthetic case study mostly aimed at verifying that the estimation of hidden climate indices from occurrence data alone is feasible. In sections 4 and 5, the model is applied to two contrasting case studies based on flood occurrences: eastern Australia (where the influence of ENSO is well established) and France (where standard climate indices are poor predictors). Section 6 discusses limitations of the proposed model and suggests avenues for improvement. Finally, the main outcomes of this paper are summarized in section 7.

## 2. Theory

### 2.1. Probabilistic Model

#### 2.1.1. Notation

Let  $Y(x,t)$  denote the random variable representing the occurrence of an event at location  $x$  and time  $t$ . Corresponding observations are available at  $N_x$  locations during  $N_t$  time steps, yielding the data set  $\mathbf{y} = (y(x_i, t_j))_{i=1:N_x, j=1:N_t}$ , which can be stored as a 0/1 (nonoccurrence/occurrence) matrix. The event could for instance be defined as the seasonal maximum streamflow exceeding some threshold, but

“nonextreme” events may also be considered, such as the seasonal average precipitation being lower than normal.

### 2.1.2. Data Layer

Since  $Y(x,t)$  represents occurrence data, it is assumed to follow a Bernoulli distribution whose parameter  $\theta(x,t)$  (equal to the probability of occurrence) varies in space and time as follows:

$$\text{Logit}(\theta(x,t)) = \lambda_0(x) + \lambda_1(x)\tau_1(t) + \dots + \lambda_K(x)\tau_K(t) \quad (1)$$

The basic idea behind equation (1) is to describe the temporal variation of  $\theta(x,t)$  using a set of  $K$  hidden climate indices ( $\tau_1(t), \dots, \tau_K(t)$ ). The climate indices are hidden in the sense that they are assumed unknown and hence have to be inferred from the occurrence data. When a large spatial domain is analyzed, the effect of each HCI is not expected to be the same for all sites: the spatial parameters ( $\lambda_1(x), \dots, \lambda_K(x)$ ) therefore control the site-specific effect of each HCI. The spatial parameter  $\lambda_0(x)$  can be interpreted as the “normal” probability of occurrence (i.e., when all HCIs are equal to zero). Finally, the *Logit* transformation is used to ensure that parameter  $\theta(x,t)$  always remains between 0 and 1. The *Logit* transformation and its inverse are defined as

$$\text{Logit}(p) = \ln\left(\frac{p}{1-p}\right), p \in (0; 1); \quad \text{Logit}^{-1}(u) = \frac{1}{1 + e^{-u}}, u \in (-\infty; +\infty) \quad (2)$$

Note that the number of HCIs  $K$  is assumed to be known; strategies to select this number will be discussed in the case studies.

### 2.1.3. Spatial Process Layer

Consider the spatial parameter  $\lambda_k(x)$  controlling the effect of the  $k$ th HCI. This parameter is likely to be spatially structured. For instance, it may depend on the elevation of the site  $x$ , or on some other covariate (e.g., windward/leeward orientation and distance to sea). In addition, it is expected to show some degree of smoothness in space, since close sites tend to have similar climate effects. A similar reasoning may hold for the spatial parameter  $\lambda_0(x)$ . Consequently, all spatial parameters ( $\lambda_0(x), \lambda_1(x), \dots, \lambda_K(x)$ ) are assumed to be realizations from Gaussian spatial processes, which can be written as follows:

$$\begin{aligned} &\text{for any } k = 0 : K, \text{ and for any set of sites } (x_1, \dots, x_n) : \\ &(\lambda_k(x_1), \dots, \lambda_k(x_n)) \sim MG(\mathbf{z}_k \boldsymbol{\beta}_k, \boldsymbol{\Sigma}_k) \end{aligned} \quad (3)$$

where  $MG(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  denotes the multivariate Gaussian distribution with mean  $\boldsymbol{\mu}$  and covariance matrix  $\boldsymbol{\Sigma}$ .

In equation (3),  $\mathbf{z}_k$  is a  $n \times N_Z$  matrix of  $N_Z$  site descriptors (e.g., elevation and distance to sea) and  $\boldsymbol{\beta}_k$  is a  $N_Z \times 1$  vector of parameters. The product  $\mathbf{z}_k \boldsymbol{\beta}_k$  therefore accounts for the linear effect of covariates on parameter  $\lambda_k(x)$ . Note that a spatially constant mean can be obtained by setting  $\mathbf{z}_k$  to a  $n \times 1$  matrix of ones.

The covariance matrix  $\boldsymbol{\Sigma}_k$  controls the spatial variability of parameter  $\lambda_k(x)$ , including its spatial smoothness. Following a common approach used in geostatistics, it is parameterized as a function of the pairwise intersite distance:

$$(\boldsymbol{\Sigma}_k)_{ij} = f_k(\text{dist}(x_i, x_j); \boldsymbol{\gamma}_k) \quad (4)$$

The function  $f_k(d; \boldsymbol{\gamma}_k)$  is the covariance function or covariogram, and corresponds to an assumption of isotropic spatial process. It can be chosen among the many valid covariogram models existing in the geostatistical literature (e.g., exponential, Gaussian, and Matérn). As an example, the exponential covariogram that will be used in the case studies is defined as

$$f(d; \gamma_1, \gamma_2) = \gamma_1^2 \exp(-d/\gamma_2) \quad (5)$$

In the context of a hierarchical model such as the one presented here, the Gaussian spatial process of equation (3) is referred to as the hyperdistribution, and the parameters  $\boldsymbol{\beta}_k$  and  $\boldsymbol{\gamma}_k$  controlling its properties are the hyperparameters.

#### 2.1.4. Independence Assumptions

To complete the specification of the probabilistic model, several additional independence assumptions are made. The first assumption is that the effects of two distinct HCIs are independent, or in other words, that there is no cross correlation between the spatial processes  $\lambda_{k_1}(x)$  and  $\lambda_{k_2}(x)$  (for any  $k_1 \neq k_2$ ). The validity of this assumption can be empirically checked after the inference by analyzing the estimated values  $(\hat{\lambda}_{k_1}(x_i))_{i=1:N_x}$  and  $(\hat{\lambda}_{k_2}(x_i))_{i=1:N_x}$ .

The second assumption is that the occurrences  $Y(x,t)$  are independent in both space and time *conditionally* on parameters  $\theta(x,t)$ . It is stressed that the word “conditionally” is extremely important here: in particular, this conditional independence assumption does not imply that the occurrences  $Y(x,t)$  are independent in space and time. To see this, consider the model with a single HCI:

$$Y(x, t) \sim \text{Bernoulli}(\theta(x, t)); \text{Logit}(\theta(x, t)) = \lambda_0(x) + \lambda_1(x)\tau_1(t)$$

Consider a particular site  $x_0$  where the influence of the HCI is highly positive ( $\lambda_1(x_0) \gg 0$ ). Moreover, assume that the HCI  $\tau_1(t)$  has some strong autocorrelation, with a tendency for high values to cluster in time: in this case, high probabilities  $\theta(x_0, t)$  will also have the same tendency to cluster in time, and so will the observed occurrences. In other words, the possible temporal dependence structure existing in the HCI  $\tau_1(t)$  will propagate to the occurrences  $Y(x_0, t)$ .

Similarly, consider a particular time step  $t_0$  where the value of the HCI is highly positive ( $\tau_1(t_0) \gg 0$ ). Since the process  $\lambda_1(x)$  shows some spatial dependence (see section 2.1.3), high values of  $\lambda_1(x)$  will be clustered in space and so will high probabilities  $\theta(x, t_0)$ . As previously, the spatial dependence structure existing in the HCI effect  $\lambda_1(x)$  will propagate to the occurrences  $Y(x, t_0)$ : years with high values of the HCI will tend to see more occurrences in those (clustered) locations where  $\lambda_1(x)$  is high.

These considerations illustrate that the conditional independence assumption does not correspond to assuming independent occurrences in space and time, quite the contrary: It should rather be viewed as a way to indirectly model space/time dependences, using the hidden temporal variables  $\tau_k(t)$  and their corresponding spatial effects  $\lambda_k(x)$ .

#### 2.2. Identifiability Constraints

In order to make the model in equation (1) identifiable, it is necessary to use two constraints on the mean and the variance of the HCIs. To see this, consider again a model with a single HCI, and observe that it can trivially be rewritten as

$$\text{Logit}(\theta(x, t)) = \lambda_0(x) + \lambda_1(x)\tau_1(t) = \lambda_0(x) + \frac{\lambda_1(x)}{c} (c\tau_1(t))$$

where  $c$  is a nonzero constant. This equation implies that the scale of the HCI is irrelevant: multiplying the HCI  $\tau_1(t)$  and dividing its effect  $\lambda_1(x)$  by the same constant leads to exactly the same probability of occurrence  $\theta(x,t)$ . Consequently, all parameter pairs of the form  $(\frac{\lambda_1(x)}{c}; c\tau_1(t))$  are indistinguishable from the data, leading to nonidentifiability. For similar reasons, the mean of the HCI is also irrelevant. Consequently, two identifiability constraints are used to force the HCI  $\tau_k(t)$  to have mean zero and variance one:

$$\forall k = 1 : K, \sum_{j=1}^{N_t} \tau_k(t_j) = 0 \text{ and } \frac{1}{N_t} \sum_{j=1}^{N_t} \tau_k^2(t_j) = 1 \quad (6)$$

These two constraints imply that only the first  $N_t - 2$  values of  $\tau_k(t)$  need to be inferred, with values  $\tau_k(t_{N_t-1})$  and  $\tau_k(t_{N_t})$  being deduced from the others so as to meet the constraints. Some elementary algebra leads to the following formula:

$$\tau_k(t_{N_i-1}) = \frac{-u - \sqrt{2N_i - 2v - u^2}}{2}; \tau_k(t_{N_i}) = \frac{-u + \sqrt{2N_i - 2v - u^2}}{2} \quad (7)$$

$$\text{where } u = \sum_{j=1}^{N_i-2} \tau_k(t_j); v = \sum_{j=1}^{N_i-2} \tau_k^2(t_j)$$

Note that equation (7) is undefined if the quantity below the square root is strictly negative: in this case, the vector  $(\tau_k(t_j))_{j=1:N_i-2}$  is considered as infeasible.

### 2.3. Inference

#### 2.3.1. Stepwise Strategy

In the model described in the previous section, the following quantities need to be inferred:

1. The spatial parameter  $\lambda_0 = (\lambda_0(x_i))_{i=1:N_x}$  controlling the “normal” probability of occurrence, and the hyperparameters controlling its spatial properties  $\beta_0$  and  $\gamma_0$ ;
2. For  $k = 1 : K$ , the  $k$ th HCI  $\tau_k = (\tau_k(t_j))_{j=1:N_i-2}$ , its spatial effect  $\lambda_k = (\lambda_k(x_i))_{i=1:N_x}$ , and the hyperparameters controlling the spatial properties of the latter  $\beta_k$  and  $\gamma_k$ .

A stepwise Bayesian inference strategy is adopted to estimate these parameters. The time-invariant model with no HCI ( $\text{Logit}(\theta(x,t)) = \lambda_0(x)$ ) is estimated first. The model with a single HCI is then estimated conditionally on the previously estimated values  $\hat{\lambda}_0(x)$  (i.e.,  $\text{Logit}(\theta(x,t)) = \hat{\lambda}_0(x) + \lambda_1(x)\tau_1(t)$ ). The model with two HCIs is then estimated conditionally on  $\hat{\lambda}_0(x)$ ,  $\hat{\lambda}_1(x)$ , and  $\hat{\tau}_1(t)$  (i.e.,  $\text{Logit}(\theta(x,t)) = \hat{\lambda}_0(x) + \hat{\lambda}_1(x)\hat{\tau}_1(t) + \lambda_2(x)\tau_2(t)$ ) and so forth until the desired number of HCIs  $K$  is reached. This stepwise procedure is used because it is much better behaved than a single global procedure estimating all parameters at once. The drawbacks of the stepwise approach are discussed in section 6.3, along with potential avenues to improve the behavior of the global estimation procedure.

#### 2.3.2. Estimation of the Time-Invariant Model With No HCI

The stepwise procedure starts by estimating parameters  $\lambda_0$ ,  $\beta_0$ , and  $\gamma_0$ . Their posterior distribution, given occurrence data  $\mathbf{y}$  and spatial covariates  $\mathbf{z}$ , is given by

$$p(\lambda_0, \beta_0, \gamma_0 | \mathbf{y}, \mathbf{z}) \propto p(\mathbf{y} | \lambda_0, \beta_0, \gamma_0, \mathbf{z}) p(\lambda_0, \beta_0, \gamma_0 | \mathbf{z})$$

$$= \underbrace{p(\mathbf{y} | \lambda_0)}_{\text{likelihood}} \underbrace{p(\lambda_0 | \beta_0, \gamma_0, \mathbf{z})}_{\text{hierarchical term}} \underbrace{p(\beta_0, \gamma_0 | \mathbf{z})}_{\text{prior}} \quad (8)$$

Equation (8) corresponds to a standard expression of the posterior distribution in a Bayesian hierarchical framework. The first row corresponds to the application of the Bayes theorem. The second term of this first row is then decomposed into hierarchical and prior terms, using basic conditional probability algebra. Moreover, the likelihood term is simplified by noting that it only depends on parameters  $\lambda_0$ , but not on hyperparameters  $\beta_0$  and  $\gamma_0$ . Indeed, the likelihood is equal to

$$p(\mathbf{y} | \lambda_0) = \prod_{i=1}^{N_x} \prod_{j=1}^{N_i} (\theta(x_i, t_j))^{y(x_i, t_j)} (1 - \theta(x_i, t_j))^{1 - y(x_i, t_j)} \quad (9)$$

$$\text{where } \theta(x_i, t_j) = \text{Logit}^{-1}(\lambda_0(x_i))$$

In equation (9), each term within the double product is equal to the probability of occurrence  $\theta(x_i, t_j)$  if the event occurs ( $y(x_i, t_j) = 1$ ) and to  $1 - \theta(x_i, t_j)$  if the event does not occur ( $y(x_i, t_j) = 0$ ). The double product is a consequence of the conditional independence assumption in space and time. It has the positive side effect of making the handling of missing data straightforward: if a value  $y(x_i, t_j)$  is missing, the corresponding term can simply be omitted from the double product.

The hierarchical term in equation (8) is computed as a multivariate Gaussian pdf  $f_{MG}$  with mean vector  $\mathbf{z}_0\beta_0$  and covariance matrix  $\Sigma_0(\gamma_0)$ , as explained in section 2.1.3:

$$p(\lambda_0 | \beta_0, \gamma_0, \mathbf{z}) = f_{MG}(\lambda_0; \mathbf{z}_0\beta_0, \Sigma_0(\gamma_0)) \quad (10)$$

Finally, the prior term in equation (8) has to be specified by the user and is case specific. Note that in most cases, this prior will not depend on the spatial covariates  $\mathbf{z}$ , and independent priors will be used, leading to  $p(\beta_0, \gamma_0 | \mathbf{z}) = p(\beta_0)p(\gamma_0)$ .

### 2.3.3. Estimation of the $k$ th HCI Model

Now assume that components 0 to  $k - 1$  of the model have been estimated, with corresponding estimated values  $\hat{\lambda}_0$ ,  $\hat{\lambda}_{1:k-1}$ , and  $\hat{\tau}_{1:k-1}$ . These values correspond to the *maxpost* estimates, that is, the parameter vector that maximizes the posterior pdf. Analogous to the preceding section, the posterior pdf of the parameters of the  $k$ th component is then given by

$$p(\tau_k, \lambda_k, \beta_k, \gamma_k | \hat{\lambda}_0, \hat{\lambda}_{1:k-1}, \hat{\tau}_{1:k-1}, \mathbf{y}, \mathbf{z}) \propto \underbrace{p(\mathbf{y} | \tau_k, \lambda_k, \hat{\lambda}_0, \hat{\lambda}_{1:k-1}, \hat{\tau}_{1:k-1})}_{\text{likelihood}} \underbrace{p(\lambda_k | \beta_k, \gamma_k, \mathbf{z})}_{\text{hierarchical term}} \underbrace{p(\beta_k) p(\gamma_k) p(\tau_k)}_{\text{prior}} \quad (11)$$

The likelihood is computed as follows:

$$p(\mathbf{y} | \tau_k, \lambda_k, \hat{\lambda}_0, \hat{\lambda}_{1:k-1}, \hat{\tau}_{1:k-1}) = \prod_{i=1}^{N_x} \prod_{j=1}^{N_t} (\theta(x_i, t_j))^{y(x_i, t_j)} (1 - \theta(x_i, t_j))^{1 - y(x_i, t_j)} \quad (12)$$

where  $\theta(x_i, t_j) = \text{Logit}^{-1}(\hat{\lambda}_0(x_i) + \hat{\lambda}_1(x_i)\hat{\tau}_1(t_j) + \dots + \hat{\lambda}_{k-1}(x_i)\hat{\tau}_{k-1}(t_j) + \lambda_k(x_i)\tau_k(t_j))$

The hierarchical term can be computed in a similar way as in the previous section.

### 2.3.4. MCMC Sampling

The posterior distribution in equation (11) is explored by Markov Chain Monte Carlo (MCMC) sampling. An adaptive block Metropolis sampler is used, with block of length one: the vector of inferred parameters is updated one component at a time (see Marshall et al., 2004, for a description of such block MCMC samplers). The adaption strategy is based on monitoring the acceptance rate during iterations and increasing/decreasing the variance of the univariate Gaussian jump distributions until acceptance rates lie within user-specified bounds (typically between 10% and 50%, see Renard et al., 2006, for a more detailed description). In addition, the sampler is customized to the particular structure of the posterior distribution (11), and in particular its double-product likelihood (equation (12)). Indeed, many simplifications occur in the computation of the Metropolis ratio, and taking advantage of these dramatically improves the computational efficiency of the sampler: this is further described in Text S1 in the supporting information.

## 3. A synthetic Case Study

The objective of this synthetic case study is to illustrate the model described in section 2, and to verify that it is able to reliably recover the HCIs used to generate occurrence data, solely based on the latter.

### 3.1. Setup

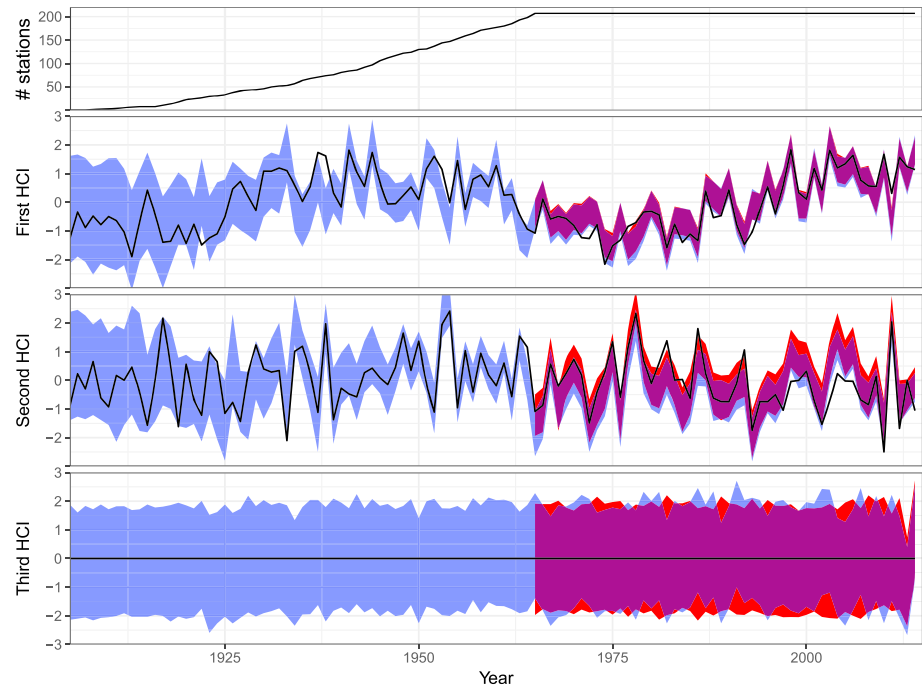
#### 3.1.1. Data

Synthetic data are generated from the Bernoulli model of equation (1). To achieve this, the following elements need to be specified:

1. Hidden climate indices  $\tau_1(t), \dots, \tau_k(t)$ : We use a model with  $K = 2$  HCIs, shown as black lines in the second and third panels of Figure 2. The first HCI shows some clear low-frequency variability, while the second is mostly interannual.
2. Site locations:  $N_x = 207$  sites in France are used (Figure 3), corresponding to the hydrometric stations that will be studied in the real data case study of section 5.
3. A “normal” probability of occurrence set to 0.2 for all sites (leading to  $\lambda_0(x) = \text{Logit}(0.2) \approx -1.39$ ).
4. HCI spatial effects  $\lambda_1(x), \lambda_2(x)$ : The effect of each HCI at the 207 sites is shown in the top panels of Figure 3. The effect follows a zonal gradient for the first HCI, and a meridional gradient for the second one. In practice the values were generated from a Gaussian spatial field.

With these elements, it is possible to compute the probability of occurrence for any location and time step using equation (1) and to generate a 0/1 value from the corresponding Bernoulli distribution. Note that the conditional independence assumption made in section 2.1.4 implies that each value can be generated independently (in both space and time). Finally, a missing data pattern resembling that of real-life data sets is implemented (Figure S1 and top panel in Figure 2). The simulated data are released and illustrated in the supporting information (Table S1 and Figure S2).





**Figure 2.** Synthetic case study: hidden climate indices (HCIs)  $\tau_1$ ,  $\tau_2$ , and  $\tau_3$ . The black line denotes the real values; colored bands denote 95% posterior intervals using data from the short (red) or the full period (blue). The top panel shows the number of available stations.

### 3.1.2. Model

The HCI model developed in section 2 is applied to the synthetic occurrence dataset with  $K = 3$  components. The objective behind using a third component (whereas only two were used to generate the data) is to evaluate how such an unnecessary component is estimated: this may suggest heuristic ways to select the number of components. For all HCI effects, the underlying Gaussian spatial process (equation (3)) is assumed to have a constant mean ( $\mu = \mathbf{1} \cdot \beta$ ) and an exponential covariogram (equation (5)). Very vague priors are used for the associated hyperparameters. A Gaussian distribution with mean zero and standard deviation 2.5 is used for the hypermean parameter  $\beta$ : as an order of magnitude, a mean HCI effect  $\beta = 2.5$  would lead to probabilities of occurrence varying from less than 0.01 to more than 0.99 for “extreme” values of the HCI  $-2$  and  $2$  (remind that each HCI has standard deviation 1). An inverse chi-squared distribution with a single degree of freedom is used for covariogram parameters  $\gamma_1$  and  $\gamma_2$ . This prior distribution is weakly informative and is aimed at avoiding values of  $(\gamma_1, \gamma_2)$  that would tend toward zero, which may create numerical problems (see section 9.2 in Spiegelhalter et al., 1996, for a discussion on this topic). Flat priors are used for all HCIs  $\tau_k$ .

### 3.1.3. Estimation

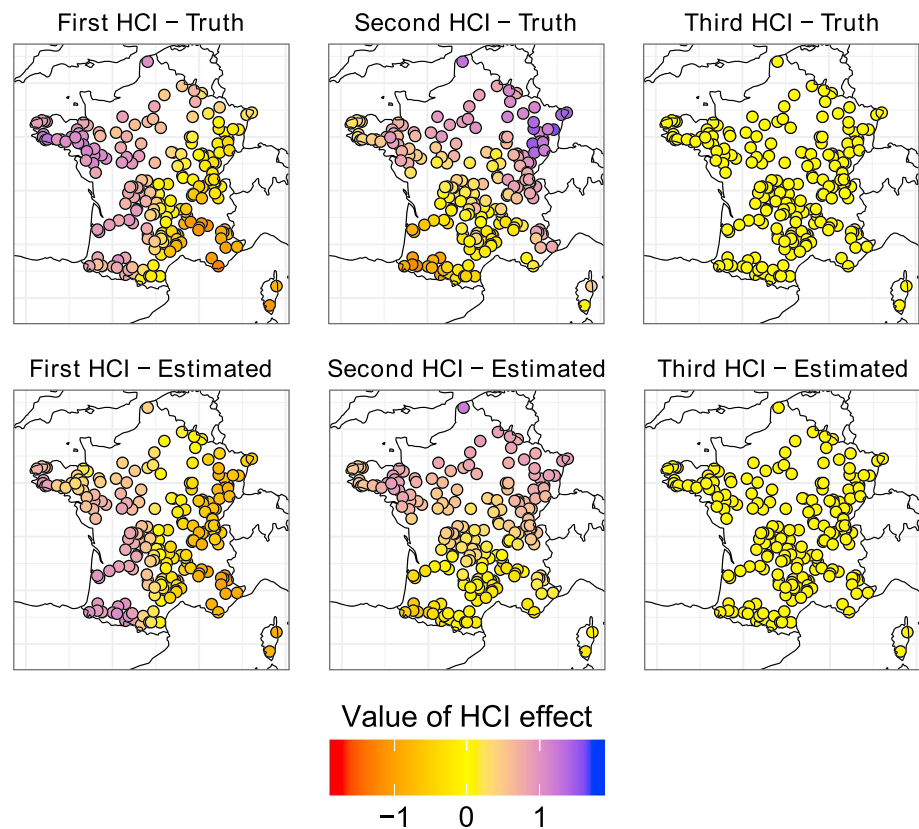
Two estimation periods are used for inference: a 50-year short period (1965–2014) with no missing data across the 207 sites, and the full 110-year (1905–2014) period, with an increasing frequency of missing data as one goes back in time. The objective is to evaluate whether the model is able to identify early values of the HCIs despite decreasing data availability, without deteriorating the estimation over the recent period.

The MCMC sampler is run for 300,000 iterations. The first 100,000 iterations are discarded as a burn-in period, and the remaining 200,000 iterations are further thinned by a factor of 200 to avoid storage issues and long postprocessing computing times. Convergence is assessed visually by examining MCMC traces for all inferred quantities (Figure S3).

## 3.2. Results

### 3.2.1. Ability to Identify the True HCIs

Figure 2 compares the true HCIs with the estimated ones (represented as 95% posterior intervals). Overall, the agreement is good, suggesting that inferring HCIs solely from occurrence data is feasible. Focusing on



**Figure 3.** Synthetic case study: effect of the hidden climate indices (HCIs)  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$ : true values (top) versus estimated values (posterior median, bottom).

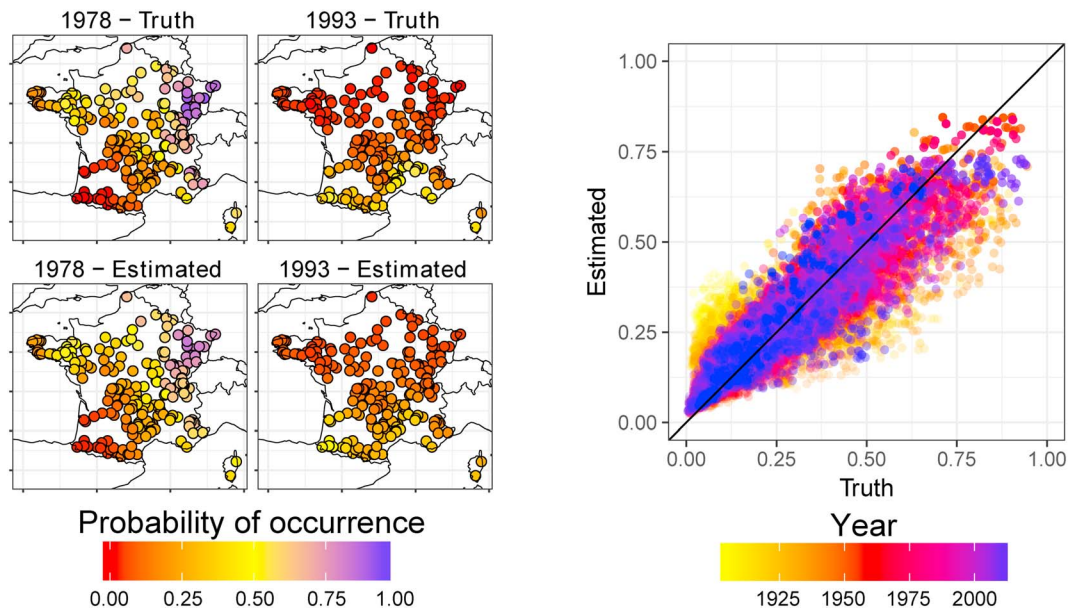
the first HCI, the full-period estimation is able to identify the underlying low-frequency signal. The uncertainty around HCI values grows as one goes back in time, reflecting the increasing frequency of missing data. Over the recent years, the full-period estimation is almost identical to the short-period estimation, which illustrates that including old albeit incomplete data does not deteriorate the inference. Note that restricting on the 1965–2014 period may erroneously suggest an upward trend; the full-period estimation more correctly reveals the multidecadal nature of this HCI. The estimated effect of this first HCI is mapped in Figure 3. The full-period estimate is reported, the short-period estimate being almost identical (not shown). The estimated pattern adequately reproduces the true zonal gradient but seems slightly smoother (this might be due to the use of posterior medians).

Similar comments can be made for the second HCI (third panel of Figure 2): with the full-period estimation, the uncertainty in HCI values grows as one goes back in time; the full-period and short-period estimations are consistent over recent years, although the difference is larger than for the first HCI; the estimated effect (middle panels of Figure 3) reproduces the meridional gradient characterizing the true effect, although in a smoother way.

A more formal evaluation of the reliability of estimated HCIs using PIT diagrams (probability integral transform—see, e.g., Laio & Tamea, 2007) indicates an acceptable reliability, despite a slight underestimation of uncertainty (Figure S4). This is most probably due to the stepwise estimation strategy: estimating the  $k$ th component conditionally on a point estimate of the  $k$ th-1 component leads to ignoring the uncertainty in the latter. While this is not entirely satisfying, Figure S4 suggests that this underestimation is limited and may be acceptable for practical purposes. This issue is further discussed in section 6.3.

### 3.2.2. Selecting the Number of Components

The last row of Figure 2 shows the third estimated HCI (remind that only two HCIs were used to simulate the data). Both short-period and full-period estimates show no discernable temporal variability. Note that the uncertainty interval does not collapse to zero because of the identifiability constraint forcing all HCIs to



**Figure 4.** Synthetic case study: true versus estimated probabilities of occurrence. Left: true and estimated (posterior median) probability maps for years 1978 and 1993. Right: true versus estimated probabilities for all years and all stations; the transparency of each point is proportional to the posterior standard deviation.

have a standard deviation equal to one (see section 2.2). However, the effect of the third HCI does collapse to zero (right panels of Figure 3). This suggests that computing the standard effect of each HCI might be useful to choose the number of components:

$$SE_k = \sqrt{\frac{1}{N_x} \sum_{i=1}^{N_x} (\lambda_k(x_i))^2}$$

Here, this standard effect is equal to 0.61 for the first HCI and 0.51 for the second one but drops by a factor of 10 for the third one (0.06), confirming that two components are sufficient. It is also worth noting that MCMC convergence is rather poor for the effects of the third HCI  $\lambda_3(x)$  (see last page in Figure S3), which may also be indicative of the uselessness of this third component.

### 3.2.3. Estimation of Occurrence Probabilities

Once HCIs and their effects are estimated, it is possible to estimate probabilities of occurrence using the formula in equation (1). The left part of Figure 4 shows two examples for the years 1978 and 1993. In both cases, the estimated probabilities of occurrence closely reproduce the true spatial pattern: In 1978, the model correctly identifies the northeast area with very high probabilities of occurrence, while in 1993 it recognizes lower-than-usual probabilities in the Northern half of the country. The right part of Figure 4 summarizes these findings for all sites and all years. It shows that the estimation of probabilities of occurrence is accurate. Moreover, the model is able to predict a wide range of probabilities. In particular, probabilities close to zero and as high as 0.8 are predicted, which is to be compared with the “normal” probability of occurrence set to 0.2 in this synthetic case study.

## 4. Spring floods in Eastern Australia

ENSO is an important driver of hydroclimatic variability in Australia (Chiew et al., 1998; Verdon et al., 2004). In particular, it impacts floods on the East Coast (Kiem et al., 2003), especially during summer and spring (Liu et al., 2018). The objective of this case study is therefore to analyze spring flood occurrences in Eastern Australia and to verify that the HCI identified by the proposed model is indeed related to ENSO.

### 4.1. Setup

#### 4.1.1. Data

Daily discharge series from 42 stations are used, over the period 1951–2014. These stations are part of the Hydrologic Reference Stations data set (Zhang et al., 2014) and were selected to cover coastal New South

**Table 1**  
*Climate Data Used to Interpret the Hidden Climate Indices*

Acronym	Full name	Data source	Availability	Reference
		ENSO indices		
SOI	Southern Oscillation Index	NOAA CPC <sup>a</sup>	1951–pst	(Ropelewski & Jones, 1987)
NINO1+2	Extreme eastern tropical pacific SST	NOAA ESRL <sup>b</sup>	1950–pst	(Trenberth & Stepaniak, 2001)
NINO3	Eastern tropical pacific SST	NOAA ESRL <sup>b</sup>	1950–pst	(Trenberth & Stepaniak, 2001)
NINO3.4	East central tropical pacific SST	NOAA ESRL <sup>b</sup>	1950–pst	(Trenberth & Stepaniak, 2001)
NINO4	Central tropical pacific SST	NOAA ESRL <sup>b</sup>	1950–pst	(Trenberth & Stepaniak, 2001)
	Standard climate indices relevant for western Europe			
NAO	North Atlantic Oscillation	NOAA CPC <sup>c</sup>	1950–pst	(Barnston & Livezey, 1987)
SCAND	Scandinavia pattern	NOAA CPC <sup>c</sup>	1950–pst	(Wibig, 1999)
EA	East Atlantic pattern	NOAA CPC <sup>c</sup>	1950–pst	(Barnston & Livezey, 1987)
EAWR	East Atlantic-Western Russia pattern	NOAA CPC <sup>c</sup>	1950–pst	(Ionita, 2014)
AMOr	Atlantic Multicadal Oscillation, raw	NOAA ESRL <sup>d</sup>	1856–pst	(Enfield et al., 2001)
AMOs	Atlantic Multicadal Oscillation, smoothed	NOAA ESRL <sup>d</sup>	1856–pst	(Enfield et al., 2001)
	Atmospheric fields over the North Atlantic			
Z850	Geopotential height at 850 hPa	NCEP/NCAR <sup>e</sup>	1948–pst	(Kalnay et al., 1996)
U850	Zonal wind component at 850 hPa	NCEP/NCAR <sup>e</sup>	1948–pst	(Kalnay et al., 1996)
V850	Meridional wind component at 850 hPa	NCEP/NCAR <sup>e</sup>	1948–pst	(Kalnay et al., 1996)
CAPE	Convective available potential energy	20CR <sup>f</sup>	1871–2012	(Compo et al., 2011)
	Sea surface temperature (SST)			
SST	Kaplan low-resolution global SST	NOAA ESRL <sup>g</sup>	1856–pst	(Kaplan et al., 1998)
SSTmed	High-resolution Mediterranean SST	CMEMS <sup>h</sup>	1955–2015	(Fратиanni et al., 2015)
	Extreme events			
AR	Duration (day) with an atmospheric river detected over Western France	Universidad de Cantabria <sup>i</sup>	1900–2012	(Brands et al., 2017)
MEDI	Number of medicanes formed in the Western Mediterranean region	Data sent by authors	1950–2011	(Cavicchia et al., 2014)

Note. NOAA = National Oceanic and Atmospheric Administration; NCEP = National Centers for Environmental Prediction; NCAR = National Center for Atmospheric Research; ESRL = Earth System Research Laboratory; CPC = Climate Prediction Center.

<sup>a</sup><https://www.cpc.ncep.noaa.gov/data/indices/soi>. <sup>b</sup><https://www.esrl.noaa.gov/psd/data/climateindices/list/>. <sup>c</sup><http://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml>.

<sup>d</sup><https://www.esrl.noaa.gov/psd/data/timeseries/AMO/>. <sup>e</sup><https://www.esrl.noaa.gov/psd/data/reanalysis/reanalysis.shtml>. <sup>f</sup>[https://www.esrl.noaa.gov/psd/data/20thC\\_Rean/](https://www.esrl.noaa.gov/psd/data/20thC_Rean/). <sup>g</sup>[https://www.esrl.noaa.gov/psd/data/gridded/data.kaplan\\_sst.html](https://www.esrl.noaa.gov/psd/data/gridded/data.kaplan_sst.html). <sup>h</sup>[https://doi.org/10.25423/MEDSEA\\_REANALYSIS\\_PHY\\_006\\_009](https://doi.org/10.25423/MEDSEA_REANALYSIS_PHY_006_009). <sup>i</sup><http://www.meteo.unican.es/atmospheric-rivers>.

Wales and southeast Queensland. At each station, the series of austral spring maximum discharge is extracted (September-October-November, SON). A flood is considered to occur if the spring maximum exceeds the empirical 0.8-quantile computed over all maxima available at this station (thus approximately corresponding to a 5-year spring flood). All stations therefore have approximately the same frequency of occurrences (20%) and nonoccurrences (80%). The data set is released in Table S2 and illustrated in Figures S1 and S5. Additional climate data are used to interpret the results of this case study: global sea surface temperatures (SST) and standard ENSO indices (see Table 1 for a more detailed description).

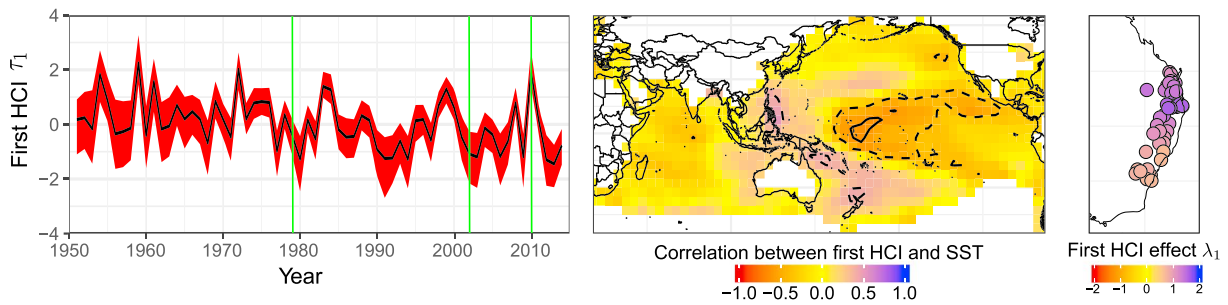
#### 4.1.2. Model and Estimation

We use the same HCI model as described in section 3.1.2 and apply it to the spring floods dataset using all available data. The model was estimated with up to  $K = 3$  components, but computing the standard effects as proposed in section 3.2.2 suggests that a single component is sufficient (1.23, dropping to 0.54 and 0.37). MCMC traces are shown in Figure S6.

### 4.2. Results

#### 4.2.1. Identification of the HCI

The left panel of Figure 5 shows the estimated HCI. Three particular years are also reported in order to interpret this temporal signal. Spring 2010 corresponds to the start of a strong La Niña event (Boening et al., 2012), leading to many flood occurrences in the region (Figure 6); the corresponding HCI value is high ( $\tau_1 \approx 2$ ). Conversely, an El Niño event occurred in 2002 (McPhaden, 2004), with eastern Australia experiencing no spring flooding; the corresponding HCI value is low ( $\tau_1 \approx -1$ ). Year 1979 was ENSO neutral, and the corresponding HCI value is close to zero. This suggests that the estimated HCI can be interpreted as a La Niña indicator. This is further confirmed by high absolute correlations between the HCI and standard



**Figure 5.** East Australia case study. Left: hidden climate index (HCI)  $\tau_1$  (black line = posterior median, red band = 95% credibility interval, and vertical green lines: particular years illustrated in next Figure 6). Center: correlation map between hidden climate index  $\tau_1$  and sea surface temperatures (SST) computed over the period 1971–2014 (dashed and plain contours denote correlations larger than 0.4 and 0.6, respectively). Right: effect  $\lambda_1$  of the hidden climate index (posterior median).

ENSO climate indices, reaching  $-0.63$  with the NINO4 index (not shown). In addition, the correlation map between the HCI and spring SST (central panel of Figure 5) reveals an ENSO-like spatial pattern, with high HCI values being associated with negative SST anomalies in the equatorial Pacific (especially its western part), that is, to La Niña conditions. Overall, these results indicate that the HCI approach is able to identify the well-known influence of ENSO on floods in this region and during this season.

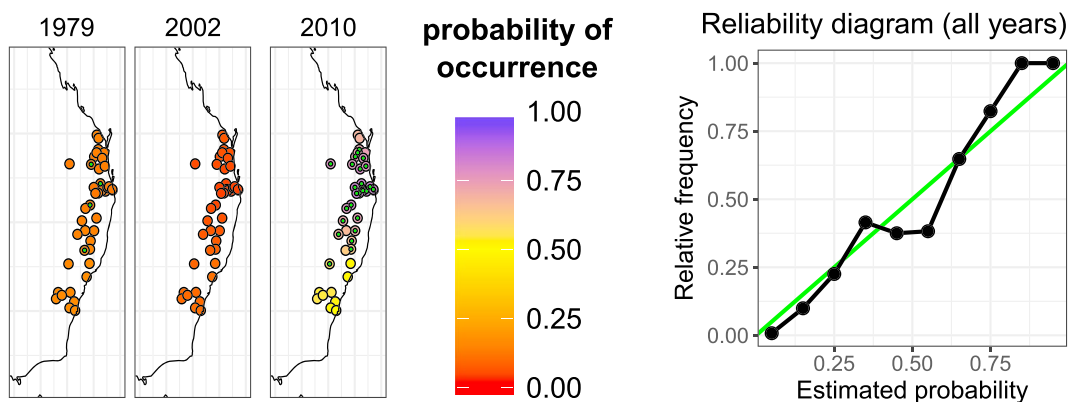
#### 4.2.2. Estimation of Occurrence Probabilities

The right panel of Figure 5 indicates that the HCI effect is positive at all stations (i.e., high values of the HCI lead to higher-than-usual flood probabilities). It also shows a meridional gradient suggesting that stations in Queensland (north) are more strongly affected than stations in New South Wales (south).

The HCI and its effect can be combined to derive probabilities of occurrence using the formula in equation (1). This is illustrated in the left panels of Figure 6. Year 1979 correspond to a neutral state  $\tau_1 \approx 0$  and to “normal” probabilities of occurrence close to 0.2. For El Niño year 2002 ( $\tau_1 \approx -1$ ), flood probabilities range between 0.05 and 0.1, which is 2 to 4 times smaller than normal; no flood occurred in the region. Conversely, La Niña year 2010 ( $\tau_1 \approx 2$ ) is characterized by flood probabilities ranging from 0.5 to 0.8, which is up to 4 times higher than normal; a flood occurred at 67% of the sites. The agreement between estimated probabilities and actual flood occurrences can be evaluated more formally by means of a reliability diagram (Bröcker, 2007): the right panel of Figure 6 confirms the good reliability of estimated probabilities across all sites and all years.

### 5. Autumn Floods in France

The NAO is an important driver of hydrologic variability in many European regions (Pociask-Karteczka, 2006), especially in autumn and winter. It has strong and opposite effects in Northern and Southern Europe, but tends to be less impactful in central regions, including France (Lavers et al., 2013). Beyond



**Figure 6.** East Australia case study: estimated probabilities of occurrence for years 1979, 2002, and 2010 (posterior medians). The green dots denote stations where a flood did occur. Right panel: reliability diagram for all years.

NAO, Renard and Lall (2014) documented that no standard climate index delivered skillful predictions of autumn floods in Mediterranean France. This case study therefore aims at illustrating the benefits of a hidden climate indices approach in such a region.

## 5.1. Setup

### 5.1.1. Flood Occurrence Data

Daily discharge series from 207 French stations are used, from the earliest available data up to 2016. These stations form the reference hydrologic network used by, for example, Giuntoli et al. (2012) and Hodgkins et al. (2017). Autumn (October–November–December, OND) flood occurrences are defined exactly as in previous section 4.1.1. The data set is released in Table S3 and illustrated in Figure S7. Data availability varies greatly between stations, as shown in Figure S1: the oldest stations started in the early twentieth century, but most stations started in the mid-1960s.

### 5.1.2. Climate Data

In order to interpret the HCIs identified from flood occurrences, several types of climate data are used. They can be classified into four categories (see Table 1 for a more detailed description):

1. Standard climate indices whose influence in Western Europe is well documented;
2. Atmospheric variables over a North Atlantic domain (80°W–20°E, 5°N–70°N);
3. Sea surface temperatures at the global scale and over the Mediterranean basin;
4. Extreme event indices denoting the occurrence of atmospheric rivers over western France and medicanes in the western Mediterranean.

All climate data are averaged over the same OND season as that used for defining flood occurrences.

### 5.1.3. Model and Estimation

The same HCI model as described in section 3.1.2 was estimated with up to  $K = 6$  components. The first component is by far the most important: its standard effect is equal to 1.74, then drops to 0.6 for the second component before slowly decreasing for subsequent components. However, for illustrative purposes, results are shown for the first three components. MCMC traces are shown in Figure S8.

## 5.2. Results

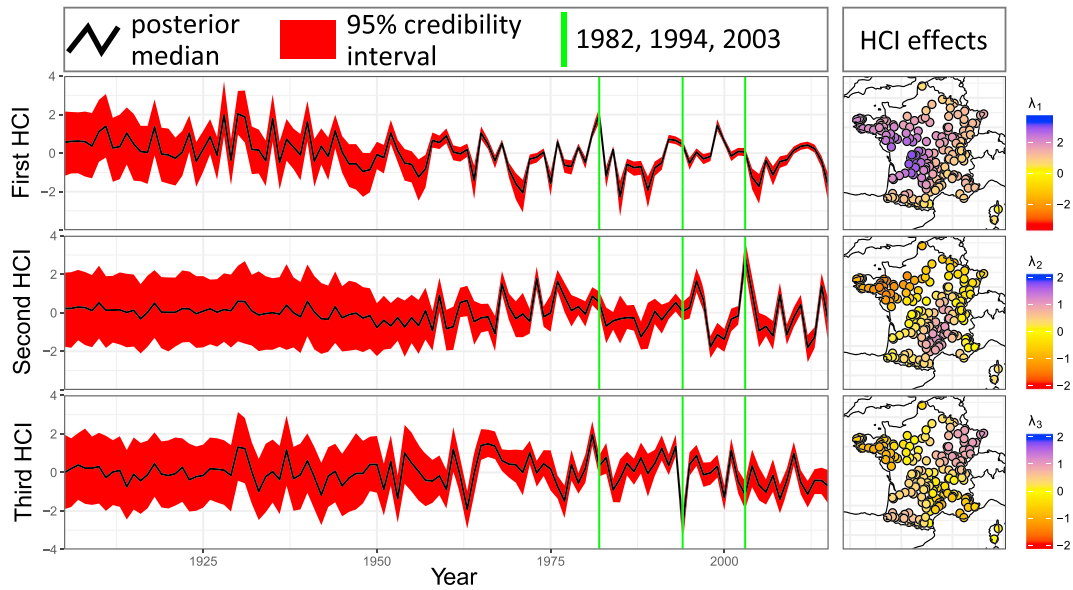
### 5.2.1. Identification of HCIs

Figure 7 shows the estimated HCIs and their effects at all stations. For the three components shown in this figure, HCI values are estimated quite precisely after 1960, but uncertainties are much larger before this date, due to lower data availability. The first HCI controls flood occurrence over a large part of France: its effect is positive at all stations, and is particularly large over the Oceanic part of the country (centerwest). The second HCI effect shows an opposition between Brittany (northwest) and the Mediterranean region (southeast): higher-than-usual probabilities in one region are associated with lower-than-usual probabilities in the other. Also note that many stations are not affected by this second HCI (yellow dots in the southwest and northeast, corresponding to near-zero effects). The third HCI effect highlights the “four corners” of France, with positive effects in the northeast and southwest, and negative effects in the northwest and southeast. However these effects start becoming relatively small, suggesting a moderate influence on probabilities of occurrence.

### 5.2.2. Ability of the HCI Model to Reproduce the Observed Space-Time Variability of Flood Occurrences

The left panels of Figure 8 show the estimated probabilities of occurrence for three particular years (denoted by vertical green lines in Figure 7). Year 1982 is characterized by a large value of the first HCI; consequently, probabilities of occurrence are higher-than-usual in most of the country, and in particular in the Oceanic part, consistently with the HCI effect shown in Figure 7. Similar comments can be made for years 2003 (large value for the second HCI) and 1994 (large negative value for the third HCI).

Overall, the agreement between the estimated probabilities of occurrence and the actual flood occurrences (green dots) is adequate in the sense that floods tend to occur where probabilities are high (blue dots) and rarely occur where probabilities are close to zero (red dots). The reliability diagram in the right panel of Figure 8 confirms the excellent reliability of estimated probabilities. More qualitatively, Figure 8 illustrates that the HCI model is able to explain the tendency of flood occurrences to cluster in space: the spatial dependence of the HCI effects (Gaussian spatial model of equation (3)) propagates to probabilities of occurrence, and to occurrences themselves.

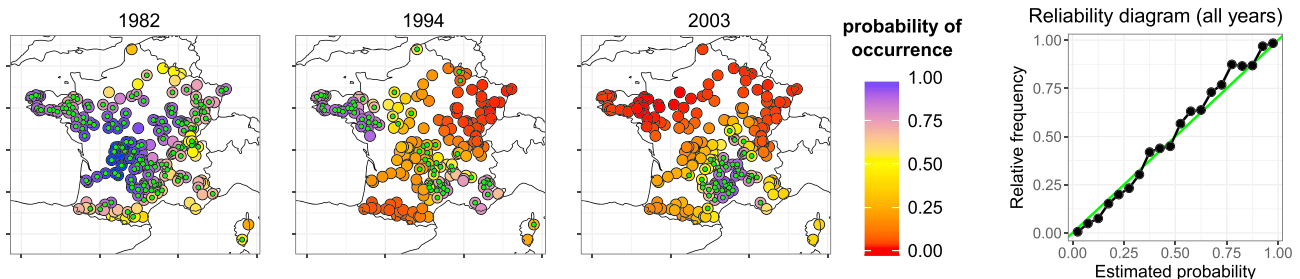


**Figure 7.** France case study: hidden climate indices (HCIs;  $\tau_1$ ,  $\tau_2$ , and  $\tau_3$ ) and their effect ( $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$ ). The vertical green lines denote the three particular years illustrated in next Figure 8.

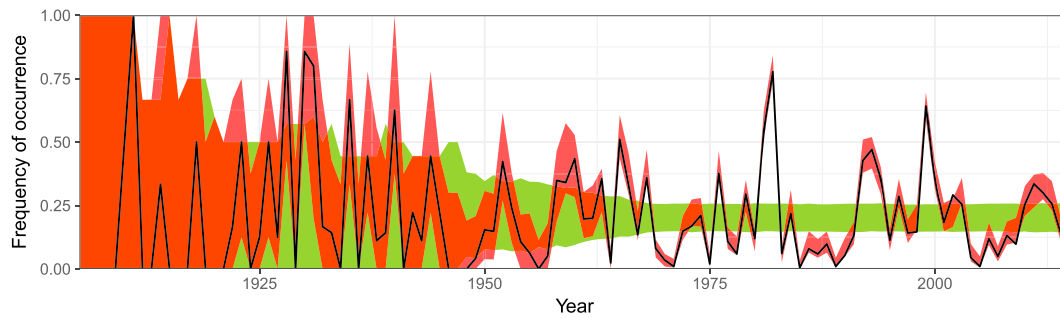
It is also of interest to assess the ability of the HCI model to reproduce the temporal variability of flood occurrences. To this aim, the percentage of available stations showing a flood occurrence can be computed each year, yielding the time series shown as a solid line in Figure 9. If occurrences were iid, the number of stations with a flood occurrence at year  $t$  would be a realization from a binomial distribution with success probability 0.2 and number of trials equal to the number of available stations  $n_t$ . The green area in Figure 9 shows the corresponding 95% probability interval expressed in percentage of stations: it fails to capture the large inter-annual variability of flood occurrences, due to the fact that the latter are not iid. Under the assumptions of the HCI model, flood occurrences are not iid since the probability of occurrence varies in both space and time. On any given year  $t$ , the number of stations with a flood occurrence corresponds to the number of successes among  $n_t$  trials, with varying success probabilities ( $\pi_1, \dots, \pi_{n_t}$ ). The corresponding distribution is known as the Poisson-binomial distribution (Hong, 2013; Serinaldi & Kilsby, 2018). The red area in Figure 9 shows the corresponding 95% probability interval, which closely follows the observed frequency of occurrence. Overall, the results shown in this section emphasize the ability of the HCI model to capture the observed variability of flood occurrences is both space and time.

### 5.2.3. Interpretation of HCIs

This section evaluates whether the three HCIs shown in Figure 7 can be interpreted in relation to climate. All analyses described thereafter restrict to the period 1960–2014, for which the uncertainty in estimated HCIs is small and relatively constant.



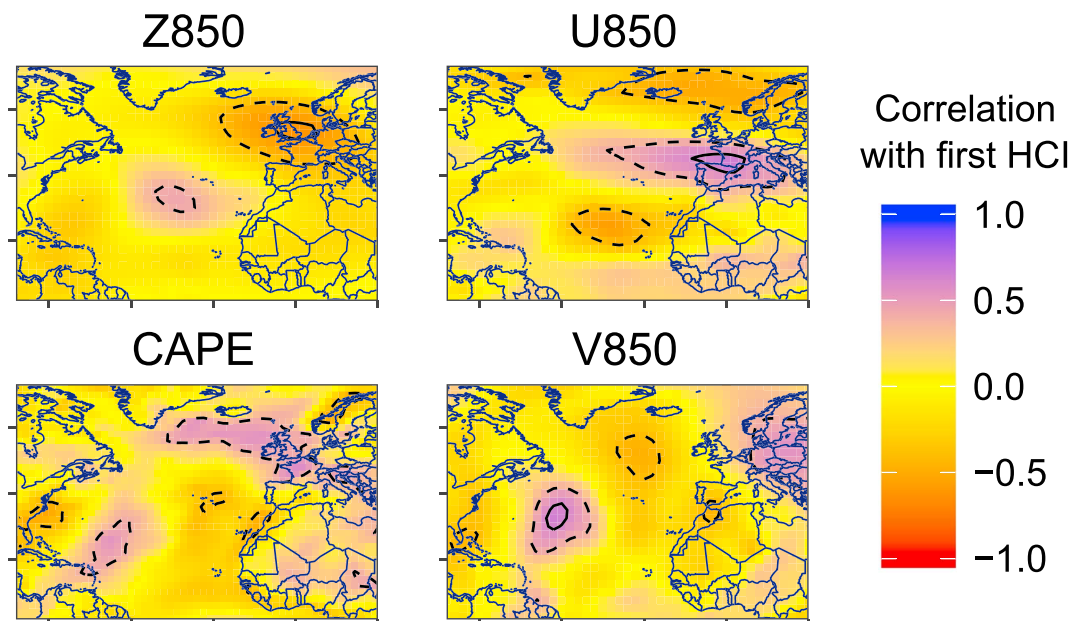
**Figure 8.** France case study: estimated probabilities of occurrence for years 1982, 1994, and 2003 (posterior medians). The green dots denote stations where a flood did occur. Right panel: reliability diagram for all years.



**Figure 9.** France case study: annual frequencies of occurrence. The solid line denotes observed frequencies, the green band denotes a 95% interval under a Binomial assumption (corresponding to iid occurrences), and the red band denotes a 95% interval derived from the hidden climate index model.

The first interesting observation is that all three HCIs of Figure 7 are essentially iid: A Mann-Kendall test (Kendall, 1975; Mann, 1945) reveals no significant trend (5% level), and no autocorrelation coefficient differs significantly from zero. The three HCIs therefore mostly describe interannual variability, with no evidence of long-term trend, low-frequency variability, or temporal clustering behavior. It is also worth noting that the three HCIs remain quite stable when the threshold defining flood occurrence varies between the 2- and the 20-year autumn flood (see sensitivity analysis in Figure S10).

Correlations between HCIs and the standard climate indices listed in Table 1 are in most cases negligible (not shown): as expected, they are not related to these particular modes of climate variability. In order to explore the relation between HCIs and climate in more depth, Figure 10 maps the correlation between the first HCI and the atmospheric climate variables listed in Table 1. The top panels show that large values of the first HCI (and hence large flood probabilities over most of France) are associated with low geopotential heights over the British Islands and strong westerly winds above the Gulf of Biscay: these can be interpreted as the traces of large oceanic storms directly hitting France. The bottom panels suggest a more distant teleconnection: large values of the first HCI are associated with large convection potential North of the Caribbean Islands and with strong southerly winds in the West Atlantic. This might correspond to conditions favoring the export of tropical moisture toward Western France. This hypothesis is also supported by



**Figure 10.** France case study: correlation maps between the first hidden climate index  $\tau_1$  and the four atmospheric variables described in Table 1, computed over the period 1960–2014. Dashed and plain contours denote correlations larger than 0.4 and 0.6, respectively. HCI = hidden climate index.



a moderate correlation (0.44) between the first HCI and the presence of atmospheric rivers over western France during the OND season (variable AR in Table 1).

These relations between the first HCI and large-scale atmospheric variables are of interest because they suggest that the former could be predicted to some extent from the latter (see section 6.4 for further discussions). However, other results corresponding to a lack of correlation between HCIs and large-scale climate are also worth reporting:

1. No noticeable correlation was found between the first HCI and global SST; this casts doubt on the feasibility of seasonal forecasting for flood occurrences, since seasonal predictability is often linked with the inertia of oceans (although antecedent moisture conditions may also play a role in some catchments, e.g., Singla et al., 2012).
2. No noticeable correlation was found between the second HCI and any of the climate variables listed in Table 1, which includes variables specific to the Mediterranean area (SST and number of Medicanic events). This suggests that the predictability of this second HCI from large-scale climate information aggregated at a seasonal time step is limited, possibly because it corresponds to flood events that are highly localized in both space and time.
3. Likewise, no noticeable correlation was found between the third HCI and any of the climate variables of Table 1.
4. Lagging SCIs or climate variables does not improve correlations with HCIs (see Figure S9 for an illustration).

## 6. Discussion

### 6.1. Relation With Probabilistic Principal Component Analysis

The reader familiar with principal component analysis (PCA, also known as empirical orthogonal functions analysis in climate sciences) may have noted conceptual similarities with the HCI model. Indeed, both approaches yield similar outputs when applied to space-time data: a set of  $K$  time series and associated spatial patterns. Interestingly, Tipping and Bishop (1999) demonstrated that PCA can be interpreted as the output from a specific hierarchical Gaussian model which shares similarities with the HCI model presented here. In a way, the HCI model can hence be interpreted as an adaptation of this probabilistic PCA model for Bernoulli occurrence data (as opposed to Gaussian data). Another difference is that the HCI model takes explicit advantage of the spatial nature of analyzed data by using a spatial model to describe the HCI effects, which can be used to make predictions at ungauged sites. More generally, the advantage of building an explicit probabilistic model over PCA-like descriptive methods is that it allows making probabilistic statements while subjecting the explicit probabilistic assumptions to empirical scrutiny.

### 6.2. Improving the Estimation Procedure

In this paper, parameters of the HCI model are estimated using a stepwise procedure (one HCI at a time). This approach is not fully satisfying because at one given step, estimated parameters are conditional on the particular parameter values estimated at the previous step. Potential interactions between parameters of different HCIs are therefore ignored, which may lead to an underestimation of uncertainty (although results from the case studies suggest it is limited). A more satisfying solution would hence be to estimate the whole set of  $K$  HCIs at once. While deriving the full posterior distribution for a set of  $K$  HCIs poses no difficulty, preliminary attempts at exploring it through MCMC sampling indicate that it is poorly behaved, probably due to a nonidentifiability problem. The two identifiability constraints of equation (6) (HCIs have mean zero and variance one) are hence sufficient to identify a single HCI at a time, but not several HCIs at once. It is conjectured that the following third constraint may solve this identifiability problem: all HCIs are uncorrelated. Unfortunately, implementing this third constraint is not straightforward because it does not lead to an explicit solution such as equation (7).

Improvements in terms of computational efficiency may also be required to apply HCI models to thousands of sites (rather than the hundreds of sites considered in this paper). Indeed, the computational bottleneck is related to inversion/multiplications of the  $N_x \times N_x$  covariance matrix of HCI effects (see MCMC algorithm in Text S1 in the supporting information). The complexity of such operations is  $O(N_x^3)$ , which becomes intractable when  $N_x$  is large. This is a common problem in Geostatistics, and various solutions have been proposed

in the literature, as reviewed by Sun et al. (2012): heuristic speed-up approximations, use of a special class of covariance models with good computational properties, etc.

### 6.3. Generalizing the HCI model

Several assumptions were made for the sake of simplicity in this paper but could in principle be generalized. For instance, the linear space-time decomposition of equation (1) could be replaced by a nonlinear one. We stress however that this linearity assumption acts on a probability of occurrence and does not imply any linearity in the flood generating process. A similar argument can be made for the effect of covariates on the hyperdistribution of HCI effects (equation (3)). The isotropic model used for the covariance of this hyperdistribution (equation (4)) could also be extended by including directional effects or by using a distance accounting for covariates such as elevation (e.g., Blanchet & Davison, 2011). The normality of this hyperdistribution could also be evaluated using a normality test (e.g., Mardia, 1980), and a non-Gaussian copula-based spatial process could be used if need be (e.g., Bárdossy & Li, 2008). Alternatively, it could be of interest to use a time series model as a hyperdistribution for the HCI (similar to the spatial model used as a hyperdistribution for the HCI effects). Such a time series model could then be used to forecast the value of the HCI a few time steps ahead.

Finally, a more ambitious endeavor would be to generalize the (occurrence data/Bernoulli distribution) setup described here, by considering for instance the case of (seasonal maxima/Generalized Extreme Value distribution). This would be useful to consider climate effects on higher extremes (e.g., the 100-year flood), for which occurrence data are not well adapted. An even more general framework would enable the consideration of any pair (variable/distribution). Such general frameworks already exist to describe the effect of *known* covariates on the parameters of some distribution (e.g., Stasinopoulos & Rigby, 2007) but would need to be extended to handle hidden covariates. It is also noted that HCI-like models can be of interest for modeling spatial extremes: indeed, the idea of using latent variables (here, HCIs and their effects) to indirectly model spatial dependence has been used by, for example, Reich and Shaby (2012) to derive a new class of max-stable models for precipitation extremes. More generally, this indirect treatment of spatial dependence should be compared with approaches modeling it more directly (e.g., the copula-based approach used by Renard, 2011).

### 6.4. Predicting HCIs and Potential Applications

In this paper, the HCIs are estimated from occurrence data, which allows evaluating the ability of the HCI model to reproduce the observed space-time variability of flood occurrences. While this corresponds to an interesting descriptive exercise, an important development would be to develop an additional predictive layer, where the value taken by the HCIs (and hence the probabilities of occurrence of extremes) could be predicted from, for example, the large-scale climate predictors used in section 5.2.3. Depending on the origin of the climate predictors, this could lead to several practical applications. A seasonal forecasting application could be explored if climate predictors are lagged with respect to the hydrologic predictand (e.g., predicting autumn floods using summer SST) or arise from a climate seasonal forecasting system. Future projections of flood hazard could be obtained by using climate predictors arising from global circulation models projections. Likewise, past reconstructions of flood hazard could be derived from climate predictors arising from long-term reanalyses such as ERA20C or 20CR: this could reveal flood “hot spots” and “hot moments” during the late nineteenth to early twentieth century (at a time when hydrometric networks were very sparse), which could be compared with historical and paleoflood evidence.

The development of this additional predictive layer is considered beyond the scope of this paper but will certainly be explored in future work given its potential for practical applications. Data mining methods may be useful for this purpose, for example, self-organizing maps (Schlef et al., 2019; Verdon-Kidd & Kiem, 2009), wavelets (Dieppois et al., 2013), empirical mode decomposition (T. A. McMahon et al., 2008), and information theory approaches (Tozer & Kiem, 2017).

## 7. Conclusion

This paper proposes a Bayesian hierarchical model based on the notion of hidden climate indices to describe the space-time variability of hydrologic extreme occurrences. The model postulates the existence of several hidden climate indices that drive the temporal variability of extreme occurrences within a set of sites. The

effect of each hidden climate index is allowed to vary in space, making the model applicable at a fairly large space scale. An efficient stepwise inference procedure is also proposed to estimate the model parameters.

A synthetic case study is used to demonstrate the feasibility of the proposed approach. In particular, hidden climate indices can reliably be recovered using occurrence data alone. Moreover, the space-time-varying probabilities of occurrence estimated from the model are in close agreement with the true probabilities of occurrence used to generate the synthetic data.

The proposed approach is then applied to real-life case studies based on flood occurrences in eastern Australia and France. The first case study confirms that the HCI approach is able to recover the well-known ENSO driver behind flood occurrences. The second case study illustrates the benefit of the approach in a region where standard climate indices are poor flood predictors. In particular, HCIs are able to explain the observed space-time variability of flood occurrences, including their tendency to cluster in space and their large interannual variability. Moreover, a detailed analysis of the hidden climate indices yields valuable insights on the climate drivers behind flood occurrences. It is first possible to assess the existence of trends or low-frequency variabilities affecting flood occurrences at a large spatial scale (both were found to be non-significant in this particular case study). In addition, correlation analyses with large-scale atmospheric fields reveal specific climate patterns associated with the hidden climate indices. This suggests that the latter may be predicted to some extent from large-scale climate information. In turn, this opens the way for additional applications such as seasonal forecasting, past reconstruction or future projection of flood hazard.

#### Acknowledgments

We would like to thank the Editor, the Associate Editor and the Reviewers for their insightful and constructive comments. Streamflow data are provided by the HYDRO database (France) and the Bureau of Meteorology (Australia). Occurrence data are provided as supporting information. Daily series can be accessed (i) through the HYDRO (<http://www.hydro.eaufrance.fr/>) and the BOM (<http://www.bom.gov.au/water/hrs/>) websites, (ii) by request to the Global Runoff Data Center ([http://www.bafg.de/GRDC/EN/Home/homepage\\_node.html](http://www.bafg.de/GRDC/EN/Home/homepage_node.html)), and (iii) by email request to the author. Atmospheric climate indices are provided by the NOAA Climate Prediction Center from their Web site at <http://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml>. AMO indices, NCEP Reanalysis Derived data, and Kaplan SST V2 data are provided by the NOAA NOAA/OAR/ESRL PSD from their Web site at <https://www.esrl.noaa.gov/psd/>. Support for the Twentieth Century Reanalysis Project data set is provided by the U.S. Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and Experiment (DOE INCITE) program, and Office of Biological and Environmental Research (BER) and by the National Oceanic and Atmospheric Administration Climate Program Office. This study has been conducted using E.U. Copernicus Marine Service Information (Mediterranean SST reanalysis). The Santander Meteorology group is acknowledged for releasing the atmospheric river archive. Leone Cavicchia is gratefully acknowledged for sending medicane data. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 835496. The analyses carried out in this paper use Dmitri Kavetski's numerical library (DMSL). Thanks D.

#### References

- Ahn, K.-H., Palmer, R., & Steinschneider, S. (2017). A hierarchical Bayesian model for regionalized seasonal forecasts: Application to low flows in the northeastern United States. *Water Resources Research*, *53*, 503–521. <https://doi.org/10.1002/2016WR019605>
- Aryal, S. K., Bates, B. C., Campbell, E. P., Li, Y., Palmer, M. J., & Viney, N. R. (2009). Characterizing and modeling temporal and spatial trends in rainfall extremes. *Journal of Hydrometeorology*, *10*(1), 241–253. <https://doi.org/10.1175/2008JHM1007.1>
- Bárdossy, A., & Li, J. (2008). Geostatistical interpolation using copulas. *Water Resources Research*, *44*, W07412. <https://doi.org/10.1029/2007WR006115>
- Barnston, A. G., & Livezey, R. E. (1987). Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Monthly Weather Review*, *115*(6), 1083–1126. [https://doi.org/10.1175/1520-0493\(1987\)115<1083:CSAPOL>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<1083:CSAPOL>2.0.CO;2)
- Blanchet, J., & Davison, A. C. (2011). Spatial modeling of extreme snow depth. *The Annals of Applied Statistics*, *5*(3), 1699–1725. <https://doi.org/10.1214/11-AOAS464>
- Boening, C., Willis, J. K., Landerer, F. W., Nerem, R. S., & Fasullo, J. (2012). The 2011 La Niña: So strong, the oceans fell. *Geophysical Research Letters*, *39*, L19602. <https://doi.org/10.1029/2012GL053055>
- Bracken, C., Rajagopalan, B., & Woodhouse, C. (2016). A Bayesian hierarchical nonhomogeneous hidden Markov model for multisite streamflow reconstructions. *Water Resources Research*, *52*, 7837–7850. <https://doi.org/10.1002/2016WR018887>
- Brands, S., Gutiérrez, J. M., & San-Martín, D. (2017). Twentieth-century atmospheric river activity along the west coasts of Europe and North America: algorithm formulation, reanalysis uncertainty and links to atmospheric circulation patterns. *Climate Dynamics*, *48*(9–10), 2771–2795. <https://doi.org/10.1007/s00382-016-3095-6>
- Bröcker, J. (2007). Increasing the reliability of reliability diagrams. *Weather and Forecasting*, *22*(3), 651–661. <https://doi.org/10.1175/WAF993.1>
- Cavicchia, L., von Storch, H., & Gualdi, S. (2014). A long-term climatology of medicanes. *Climate Dynamics*, *43*(5–6), 1183–1195. <https://doi.org/10.1007/s00382-013-1893-7>
- Chiew, F. H. S., Piechota, T. C., Dracup, J. A., & McMahon, T. A. (1998). El Niño/Southern Oscillation and Australian rainfall, streamflow and drought: Links and potential for forecasting. *Journal of Hydrology*, *204*(1–4), 138–149. [https://doi.org/10.1016/S0022-1694\(97\)00121-2](https://doi.org/10.1016/S0022-1694(97)00121-2)
- Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., et al. (2011). The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, *137*(654), 1–28. <https://doi.org/10.1002/qj.776>
- Coughlan de Perez, E., van den Hurk, B., van Aalst, M. K., Amuron, I., Bamanya, D., Hauser, T., et al. (2016). Action-based flood forecasting for triggering humanitarian action. *Hydrology and Earth System Sciences*, *20*(9), 3549–3560. <https://doi.org/10.5194/hess-20-3549-2016>
- Dieppois, B., Durand, A., Fournier, M., & Massei, N. (2013). Links between multidecadal and interdecadal climatic oscillations in the North Atlantic and regional climate variability of northern France and England since the 17th century. *Journal of Geophysical Research: Atmospheres*, *118*, 4359–4372. <https://doi.org/10.1002/jgrd.50392>
- Emerton, R., Cloke, H. L., Stephens, E. M., Zsoter, E., Woolnough, S. J., & Pappenberger, F. (2017). Complex picture for likelihood of ENSO-driven flood hazard. *Nature Communications*, *8*(1), 14796. <https://doi.org/10.1038/ncomms14796>
- Enfield, D. B., Mestas-Nunez, A. M., & Trimble, P. J. (2001). The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US. *Geophysical Research Letters*, *28*(10), 2077–2080. <https://doi.org/10.1029/2000GL012745>
- European Environment Agency (2010). *Mapping the impacts of natural hazards and technological accidents in Europe: An overview of the last decade*. Copenhagen: EEA.
- Fратиanni, C., Simoncelli, S., Pinaridi, N., Cherchi, A., Grandi, A., & Dobricic, N. S. (2015). Mediterranean RR 1955–2015 (Version 1), [dataset] Copernicus Monitoring Environment Marine Service (CMEMS).
- Giuntoli, I., Renard, B., & Lang, M. (2012). Floods in France. In Z. W. Kundzewicz (Ed.), *Changes in flood risk in Europe* (Chap. 10, pp. 199–211). Boca Raton, Florida: IAHS Press.
- Giuntoli, I., Renard, B., Vidal, J. P., & Bard, A. (2013). Low flows in France and their relationship to large scale climate indices. *Journal of Hydrology*, *482*, 105–118. <https://doi.org/10.1016/j.jhydrol.2012.12.038>

- Grantz, K., Rajagopalan, B., Clark, M., & Zagona, E. (2005). A technique for incorporating large-scale climate information in basin-scale ensemble streamflow forecasts. *Water Resources Research*, 41, W10410. <https://doi.org/10.1029/2004WR003467>
- Gregersen, I. B., Madsen, H., Rosbjerg, D., & Arnbjerg-Nielsen, K. (2013). A spatial and nonstationary model for the frequency of extreme rainfall events. *Water Resources Research*, 49, 127–136. <https://doi.org/10.1029/2012WR012570>
- Hodgkins, G. A., Whitfield, P. H., Burn, D. H., Hannaford, J., Renard, B., Stahl, K., et al. (2017). Climate-driven variability in the occurrence of major floods across North America and Europe. *Journal of Hydrology*, 552, 704–717. <https://doi.org/10.1016/j.jhydrol.2017.07.027>
- Hong, Y. (2013). On computing the distribution function for the Poisson binomial distribution. *Computational Statistics & Data Analysis*, 59, 41–51. <https://doi.org/10.1016/j.csda.2012.10.006>
- Ionita, M. (2014). The impact of the East Atlantic/Western Russia pattern on the hydroclimatology of Europe from mid-winter to late spring. *Climate*, 2(4), 296–309. <https://doi.org/10.3390/cli2040296>
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-Year Reanalysis Project. *Bulletin of the American Meteorological Society*, 77(3), 437–471. [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2)
- Kaplan, A., Cane, M. A., Kushnir, Y., Clement, A. C., Blumenthal, M. B., & Rajagopalan, B. (1998). Analyses of global sea surface temperature 1856–1991. *Journal of Geophysical Research*, 103(C9), 18,567–18,589. <https://doi.org/10.1029/97JC01736>
- Kendall, M. G. (1975). *Rank correlation methods*, (p. 202). London: Griffin.
- Kenyon, J., & Hegerl, G. C. (2010). Influence of modes of climate variability on global precipitation extremes. *Journal of Climate*, 23(23), 6248–6262. <https://doi.org/10.1175/2010JCLI3617.1>
- Kiem, A. S., Franks, S. W., & Kuczera, G. (2003). Multi-decadal variability of flood risk. *Geophysical Research Letters*, 30(2), 1035. <https://doi.org/10.1029/2002GL015992>
- Kwon, H. H., Brown, C., & Lall, U. (2008). Climate informed flood frequency analysis and prediction in Montana using hierarchical Bayesian modeling. *Geophysical Research Letters*, 35, L05404. <https://doi.org/10.1029/2007GL032220>
- Laio, F., & Tamea, S. (2007). Verification tools for probabilistic forecasts of continuous hydrological variables. *Hydrology and Earth System Sciences*, 11(4), 1267–1277. <https://doi.org/10.5194/hess-11-1267-2007>
- Lavers, D., Prudhomme, C., & Hannah, D. M. (2013). European precipitation connections with large-scale mean sea-level pressure (MSLP) fields. *Hydrological Sciences Journal*, 58(2), 310–327. <https://doi.org/10.1080/02626667.2012.754545>
- Lima, C. H. R., & Lall, U. (2010). Spatial scaling in a changing climate: A hierarchical Bayesian model for non-stationary multi-site annual maximum and monthly streamflow. *Journal of Hydrology*, 383(3–4), 307–318. <https://doi.org/10.1016/j.jhydrol.2009.12.045>
- Lima, C. H. R., Lall, U., Troy, T. J., & Devineni, N. (2015). A climate informed model for nonstationary flood risk prediction: Application to Negro River at Manaus, Amazonia. *Journal of Hydrology*, 522, 594–602. <https://doi.org/10.1016/j.jhydrol.2015.01.009>
- Liu, J., Zhang, Y., Yang, Y., Gu, X., & Xiao, M. (2018). Investigating relationships between Australian flooding and large-scale climate indices and possible mechanism. *Journal of Geophysical Research: Atmospheres*, 123, 8708–8723. <https://doi.org/10.1029/2017JD028197>
- López, J., & Francés, F. (2013). Non-stationary flood frequency analysis in continental Spanish rivers, using climate and reservoir indices as external covariates. *Hydrology and Earth System Sciences*, 17(8), 3189–3203. <https://doi.org/10.5194/hess-17-3189-2013>
- Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica*, 13(3), 245–259. <https://doi.org/10.2307/1907187>
- Mardia, K. V. (1980). Tests of univariate and multivariate normality. In P. R. Krishnaiah (Ed.), *Handbook of Statistics I: Analysis of variance*, (pp. 279–320). Holland: Amsterdam. [https://doi.org/10.1016/S0169-7161\(80\)01011-5](https://doi.org/10.1016/S0169-7161(80)01011-5)
- Marshall, L., Nott, D., & Sharma, A. (2004). A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling. *Water Resources Research*, 40, W02501. <https://doi.org/10.1029/2003WR002378>
- McMahon, G. M., & Kiem, A. S. (2018). Large floods in South East Queensland, Australia: Is it valid to assume they occur randomly? *Australasian Journal of Water Resources*, 22(1), 4–14. <https://doi.org/10.1080/13241583.2018.1446677>
- McMahon, T. A., Kiem, A. S., Peel, M. C., Jordan, P. W., & Pegram, G. G. S. (2008). A new approach to stochastically generating six-monthly rainfall sequences based on empirical mode decomposition. *Journal of Hydrometeorology*, 9(6), 1377–1389. <https://doi.org/10.1175/2008JHM991.1>
- McPhaden, M. J. (2004). Evolution of the 2002/03 El Niño. *Bulletin of the American Meteorological Society*, 85(5), 677–696. <https://doi.org/10.1175/BAMS-85-5-677>
- Merz, B., Dung, N. V., Apel, H., Gerlitz, L., Schröter, K., Steirou, E., & Vorogushyn, S. (2018). Spatial coherence of flood-rich and flood-poor periods across Germany. *Journal of Hydrology*, 559, 813–826. <https://doi.org/10.1016/j.jhydrol.2018.02.082>
- Merz, B., Nguyen, V. D., & Vorogushyn, S. (2016). Temporal clustering of floods in Germany: Do flood-rich and flood-poor periods exist? *Journal of Hydrology*, 541, 824–838. <https://doi.org/10.1016/j.jhydrol.2016.07.041>
- NCAR (2019). In National Center for Atmospheric Research Staff (Ed.), *The climate data guide: Overview: Climate indices*. <https://climatedataguide.ucar.edu/climate-data/overview-climate-indices>
- Pociask-Karteczka, J. (2006). River hydrology and the North Atlantic Oscillation: A general review. *Ambio*, 35(6), 312–314. <https://doi.org/10.1579/05-S-114.1>
- Reich, B. J., & Shaby, B. A. (2012). A hierarchical max-stable spatial model for extreme precipitation. *The Annals of Applied Statistics*, 6(4), 1430–1451. <https://doi.org/10.1214/12-AOAS591>
- Renard, B. (2011). A Bayesian hierarchical approach to regional frequency analysis. *Water Resources Research*, 47, W11513. <https://doi.org/10.1029/2010WR010089>
- Renard, B., Garreta, V., & Lang, M. (2006). An application of Bayesian analysis and MCMC methods to the estimation of a regional trend in annual maxima. *Water Resources Research*, 42, W12422. <https://doi.org/10.1029/2005WR004591>
- Renard, B., & Lall, U. (2014). Regional frequency analysis conditioned on large-scale atmospheric or oceanic fields. *Water Resources Research*, 50, 9536–9554. <https://doi.org/10.1002/2014WR016277>
- Ropelewski, C. F., & Jones, P. D. (1987). An extension of the Tahiti–Darwin Southern Oscillation Index. *Monthly Weather Review*, 115(9), 2161–2165. [https://doi.org/10.1175/1520-0493\(1987\)115<2161:AEOTTS>2.0.CO;2](https://doi.org/10.1175/1520-0493(1987)115<2161:AEOTTS>2.0.CO;2)
- Sankarasubramanian, A., & Lall, U. (2003). Flood quantiles in a changing climate: Seasonal forecasts and causal relations. *Water Resources Research*, 39(5), 1134. <https://doi.org/10.1029/2002WR001593>
- Schlef, K. E., François, B., Robertson, A. W., & Brown, C. (2018). A general methodology for climate-informed approaches to long-term flood projection—Illustrated with the Ohio river basin. *Water Resources Research*, 54, 9321–9341. <https://doi.org/10.1029/2018WR023209>
- Schlef, K. E., Moradkhani, H., & Lall, U. (2019). atmospheric circulation patterns associated with extreme United States floods identified via machine learning. *Scientific Reports*, 9(1), 7171.
- Serinaldi, F., & Kilsby, C. G. (2018). Unsurprising surprises: The frequency of record-breaking and overthreshold hydrological extremes under spatial and temporal dependence. *Water Resources Research*, 54, 6460–6487. <https://doi.org/10.1029/2018WR023055>

- Singla, S., Céron, J. P., Martin, E., Regimbeau, F., Déqué, M., Habets, F., & Vidal, J. P. (2012). Predictability of soil moisture and river flows over France for the spring season. *Hydrology and Earth System Sciences*, *16*(1), 201–216. <https://doi.org/10.5194/hess-16-201-2012>
- Spiegelhalter, D., Thomas, A., Best, N., & Gilks, W. (1996). *BUGS 0.5: Bayesian inference using Gibbs sampling manual (version ii)*, (pp. 1–59). Cambridge, UK: MRC Biostatistics Unit, Institute of Public Health.
- Stasinopoulos, D. M., & Rigby, R. A. (2007). Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, *23*(7), 46.
- Steinschneider, S., & Lall, U. (2015). A hierarchical Bayesian regional model for nonstationary precipitation extremes in Northern California conditioned on tropical moisture exports. *Water Resources Research*, *51*, 1472–1492. <https://doi.org/10.1002/2014WR016664>
- Sun, X., & Lall, U. (2015). Spatially coherent trends of annual maximum daily precipitation in the United States. *Geophysical Research Letters*, *42*, 9781–9789. <https://doi.org/10.1002/2015GL066483>
- Sun, X., Lall, U., Merz, B., & Dung, N. V. (2015). Hierarchical Bayesian clustering for nonstationary flood frequency analysis: Application to trends of annual maximum flow in Germany. *Water Resources Research*, *51*, 6586–6601. <https://doi.org/10.1002/2015WR017117>
- Sun, X., Renard, B., Thyer, M., Westra, S., & Lang, M. (2015). A global analysis of the asymmetric effect of ENSO on extreme precipitation. *Journal of Hydrology*, *530*, 51–65. <https://doi.org/10.1016/j.jhydrol.2015.09.016>
- Sun, X., Thyer, M., Renard, B., & Lang, M. (2014). A general regional frequency analysis framework for quantifying local-scale climate effects: A case study of ENSO effects on Southeast Queensland rainfall. *Journal of Hydrology*, *512*, 53–68. <https://doi.org/10.1016/j.jhydrol.2014.02.025>
- Sun, Y., Li, B., & Genton, M. G. (2012). Geostatistics for large datasets. In E. Porcu, J.-M. Montero, & M. Schlather (Eds.), *Advances and challenges in space-time modelling of natural events*, (pp. 55–77). Berlin Heidelberg, Berlin, Heidelberg: Springer. [https://doi.org/10.1007/978-3-642-17086-7\\_3](https://doi.org/10.1007/978-3-642-17086-7_3)
- Thyer, M., & Kuczera, G. (2000). Modeling long-term persistence in hydroclimatic time series using a hidden state Markov model. *Water Resources Research*, *36*(11), 3301–3310. <https://doi.org/10.1029/2000WR00157>
- Thyer, M., & Kuczera, G. (2003a). A hidden Markov model for modelling long-term persistence in multi-site rainfall time series 1. Model calibration using a Bayesian approach. *Journal of Hydrology*, *275*(1-2), 12–26. [https://doi.org/10.1016/S0022-1694\(02\)00412-2](https://doi.org/10.1016/S0022-1694(02)00412-2)
- Thyer, M., & Kuczera, G. (2003b). A hidden Markov model for modelling long-term persistence in multi-site rainfall time series. 2. Real data analysis. *Journal of Hydrology*, *275*(1-2), 27–48.
- Tipping, M. E., & Bishop, C. M. (1999). Probabilistic principal component analysis. *Journal of the Royal Statistical Society, Series B: Statistical Methodology*, *61*(3), 611–622. <https://doi.org/10.1111/1467-9868.00196>
- Tozer, C. R., & Kiem, A. S. (2017). Large-scale ocean-atmospheric processes and seasonal rainfall variability in South Australia: accounting for non-linearity and establishing the hierarchy of influence. *International Journal of Climatology*, *37*(S1), 1180–1198. <https://doi.org/10.1002/joc.5074>
- Trenberth, K. E., & Stepaniak, D. P. (2001). Indices of El Niño evolution. *Journal of Climate*, *14*(8), 1697–1701. [https://doi.org/10.1175/1520-0442\(2001\)014<1697:LIOENO>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<1697:LIOENO>2.0.CO;2)
- Verdon, D. C., Wyatt, A. M., Kiem, A. S., & Franks, S. W. (2004). Multidecadal variability of rainfall and streamflow: Eastern Australia. *Water Resources Research*, *40*, W10201. <https://doi.org/10.1029/2004WR003234>
- Verdon-Kidd, D. C., & Kiem, A. S. (2009). On the relationship between large-scale climate modes and regional synoptic patterns that drive Victorian rainfall. *Hydrology and Earth System Sciences*, *13*(4), 467–479. <https://doi.org/10.5194/hess-13-467-2009>
- Ward, P. J., Eisner, S., Flörke, M., Dettinger, M. D., & Kummerow, M. (2014). Annual flood sensitivities to El Niño–Southern Oscillation at the global scale. *Hydrology and Earth System Sciences*, *18*(1), 47–66. <https://doi.org/10.5194/hess-18-47-2014>
- Westra, S., Renard, B., & Thyer, M. (2015). The ENSO–precipitation teleconnection and its modulation by the Interdecadal Pacific Oscillation. *Journal of Climate*, *28*(12), 4753–4773. <https://doi.org/10.1175/JCLI-D-14-00722.1>
- Westra, S., & Sharma, A. (2009). Probabilistic estimation of multivariate streamflow using independent component analysis and climate information. *Journal of Hydrometeorology*, *10*(6), 1479–1492. <https://doi.org/10.1175/2009JHM1121.1>
- Westra, S., Sharma, A., Brown, C., & Lall, U. (2008). Multivariate streamflow forecasting using independent component analysis. *Water Resources Research*, *44*, W02437. <https://doi.org/10.1029/2007WR006104>
- Wibig, J. (1999). Precipitation in Europe in relation to circulation patterns at the 500 hPa level. *International Journal of Climatology*, *19*(3), 253–269. [https://doi.org/10.1002/\(SICI\)1097-0088\(19990315\)19:3<253::AID-JOC366>3.0.CO;2-0](https://doi.org/10.1002/(SICI)1097-0088(19990315)19:3<253::AID-JOC366>3.0.CO;2-0)
- World Bank (2014). Understanding risk in an evolving world—Emerging best practices in Natural Disaster Risk Assessment, 224 pp, World Bank, Global Facility for Disaster Reduction and Recovery.
- Zeng, H., Sun, X., Lall, U., & Feng, P. (2017). Nonstationary extreme flood/rainfall frequency analysis informed by large-scale oceanic fields for Xidayang Reservoir in North China. *International Journal of Climatology*, *37*(10), 3810–3820. <https://doi.org/10.1002/joc.4955>
- Zhang, S. X., Bari, M., Amirthanathan, G., Kent, D., MacDonald, A., & Shin, D. (2014). Hydrologic Reference stations to monitor climate-driven streamflow variability and trends. In E. Australia, Barton (Ed.), *Hydrology and water resources symposium 2014*. Australia: ACT.