A Hidden Climate Indices Modeling Framework for Multivariable Space-Time Data

B. Renard, M. Thyer, D. Mcinerney, D. Kavetski, M. Leonard, S. Westra

Water Resources Research

RESEARCH ARTICLE
10.1029/2021WR030007

A Hidden Climate Indices Modeling Framework for Multivariable Space-Time Data

B. Renard[1,2], M. Thyer[2], D. McInerney[2], D. Kavetski[2], M. Leonard[2], and S. Westra[2]

[1]INRAE, RiverLy, Lyon, France, [2]School of Civil, Environmental and Mining Engineering, University of Adelaide, Adelaide, SA, Australia

Key Points:

- We propose a general probabilistic model for describing the space-time variability of multivariable data
- The model is based on hidden climate indices extracted from the target data, as opposed to predefined standard climate indices
- The model is general and flexible, and can handle a wide range of hydrometeorological data sets

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

B. Renard,
benjamin.renard@inrae.fr

Citation:

Renard, B., Thyer, M., McInerney, D., Kavetski, D., Leonard, M., & Westra, S. (2021). A hidden climate indices modeling framework for multivariable space-time data. *Water Resources Research*, *57*, e2021WR030007. https://doi.org/10.1029/2021WR030007

Received 17 MAR 2021
Accepted 9 DEC 2021

Abstract Risk assessment for climate-sensitive systems often relies on the analysis of several variables measured at many sites. In probabilistic terms, the task is to model the joint distribution of several spatially distributed variables, and how it varies in time. This paper describes a Bayesian hierarchical framework for this purpose. Each variable follows a distribution with parameters varying in both space and time. Temporal variability is modeled by means of hidden climate indices (HCIs) that are extracted from observed variables. This is to be contrasted with the usual approach using predefined standard climate indices (SCIs) for this purpose. In the second level of the model, the HCIs and their effects are assumed to follow temporal and spatial Gaussian processes, respectively. Both intervariable and intersite dependencies are induced by the strong effect of common HCIs. The flexibility of the framework is illustrated with a case study in Southeast Australia aimed at modeling "hot-and-dry" summer conditions. It involves three physical variables (streamflow, precipitation, and temperature) measured on three distinct station networks, with varying data availability and representing hundreds of sites in total. The HCI model delivers reliable and sharp time-varying distributions for individual variables and sites. In addition, it adequately reproduces intervariable and intersite dependencies, whereas a corresponding SCI model (where hidden climate indices are replaced with standard ones) strongly underestimates them. It is finally suggested that HCI models may be used as downscaling tools to estimate the joint distribution of several variables at many stations from climate models or reanalyzes.

Plain Language Summary The management of hydroclimatic hazards such as droughts and heatwaves relies on the estimation of probabilities of occurrence for extreme events. Standard approaches are available for this task when one given hazard is studied at one particular location. However, this is often not sufficient. For instance, the impact of droughts or heatwaves strongly depends on their spatial extent, which requires analyzing them over many sites. It is also useful to analyze droughts and heatwaves together rather than separately, because their joint occurrence creates favorable conditions for other hazards such as bushfires to occur. In this paper, we propose a methodological framework to analyze in a probabilistic way data sets describing several hazards at many sites over many years. The principle of this approach is to identify unobserved processes called "hidden climate indices" that are pulling the strings that make data vary. This is illustrated with a case study analyzing "hot-and-dry" summers in Southeast Australia (https://vimeo.com/600898709).

1. Introduction

Resources and risk management for environmental and technological systems often requires information on multiple dependent and spatially distributed variables whose properties vary in time under the influence of climate variability and change. For instance, wildfire depends on the intensity, spatial extent, and possible co-occurrence of droughts, heat waves, and high winds (Barbero et al., 2014; Sharples et al., 2016). Each of these components may be affected by some climate-related trend (e.g., Perkins-Kirkpatrick et al., 2016, for heat waves) or multidecadal variability (e.g., Kiem et al., 2016, for droughts). As another example, renewable energy production depends on the space-time variability of multiple variables such as wind, solar radiation, and river streamflow (Engeland et al., 2017; François et al., 2014).

Quantitative risk assessment for such systems therefore relies on a probabilistic model for several spatially distributed variables. In particular, this model should adequately describe spatial and intervariable dependencies in order to obtain reliable risk estimates. It should also account for possible climate-related temporal variability or trend. Finally, this model should be flexible enough to accommodate the complicating factors frequently

accompanying station-based data sets. Factors include varying data availability that induces a large number of missing values, noncolocated variables (i.e., distinct variables being measured on distinct station networks), the use of both discrete and continuous variables (e.g., event count and intensity) and possibly the presence of censored data. A probabilistic framework meeting all these requirements is currently lacking, and the main objective of this article is to propose one.

1.1. Probabilistic Models for Resources and Risk Management

Probabilistic models are widely used for risk assessment and resources management in climate-sensitive systems. The most basic model assumes that for a single variable at a single site, data are independent and identically distributed (*iid*) realizations from a distribution with unknown parameters. Parameter estimation and uncertainty quantification can be achieved in many well-documented ways (e.g., Commonwealth of Australia, 2019; Interagency Advisory Committee on Water Data, 1982; Ramachandra Rao & Hamed, 2019). Such a model focuses on the marginal distribution of data and is typically used to design some infrastructure (e.g., for flood protection, Botto et al., 2017). However, the assumption of identical distribution does not allow using external information from climate covariates, for instance. Alternatively, it may be unrealistic because of some trend affecting the data (e.g., Nogaj et al., 2006).

To address above issues, a time-varying conditional distribution can be obtained by assuming that the parameters vary as a function of some temporal covariates. This approach is now well established, with generic tools available (e.g., Carpenter et al., 2017; Stasinopoulos & Rigby, 2007). Time itself can be used as a covariate, resulting in a nonstationary distribution (Perreault et al., 2000a, 2000b). Other typical covariates include large-scale climate information such as global temperature (Westra et al., 2012) or climate indices (Steirou et al., 2019), synoptic-scale information such as weather type (Garavaglia et al., 2010) or airflow descriptors (Maraun et al., 2010), paleoclimate information (Devineni et al., 2013; Ho et al., 2015) or even nonclimatic information (Prosdocimi et al., 2015). Note that such models may or may not be stationary, depending on the stationarity of the covariates themselves.

It is sometimes necessary to analyze the variable of interest at multiple sites. For instance, a hydroelectricity producer operating several dams needs to model streamflow at several sites. In this case, estimating a marginal or conditional distribution at a single site is not sufficient, and a multisite model is required. Two specific issues need to be addressed in this case. The first issue is to model intersite data dependence, which can be achieved using, e.g., a spatial copula (Bracken, Rajagopalan, Cheng et al., 2016; Ghosh & Mallick, 2011; Renard, 2011; Sang & Gelfand, 2010; Sun, Renard, et al., 2015; Sun et al., 2014) or a max-stable process for extremes (Le et al., 2018; Padoan et al., 2010; Ribatet et al., 2012; Westra & Sisson, 2011). The second issue is to describe the variability of parameters in space, which is typically achieved by means of a Gaussian spatial process which may include a regression with spatial covariates (Cooley et al., 2007; Dyrrdal et al., 2015). This approach leads to a hierarchical model where the first level describes the distribution of data and the second level describes the spatial hyperdistribution of parameters. When some parameters in this second level control the effect of temporal covariates, the resulting model produces distributions varying in both space and time (Aryal et al., 2009; Gregersen et al., 2013; Lima & Lall, 2010b; Ossandon et al., 2021; Steinschneider & Lall, 2015; Sun, Lall, et al., 2015).

Multivariate models can similarly result from multiple types of variables (Zscheischler et al., 2018), rather than multiple sites. Such models have also become a well-established approach for the single-site case, with intervariable dependence being typically described using copulas (Favre et al., 2004; Salvadori & De Michele, 2004) or extreme-specific models (De Haan & De Ronde, 1998; Heffernan & Tawn, 2004). Time-varying multivariable models have also been proposed (Bracken et al., 2018; Sarhadi et al., 2016).

The literature review shows that "full" space-time multivariable models are very few. We refer here to models combining all three properties discussed in the previous paragraphs, namely: (a) time-varying, (b) spatially varying, and (c) multivariable. Such models have been mostly derived within a geostatistical framework (De Iaco, 2011; Sideris et al., 2014). Unfortunately, the underlying Gaussian assumption is restrictive as it is not adapted to extreme or discrete variables, for instance. A key objective of this paper is to address this limitation by proposing a framework that is flexible enough to accommodate such variables, while concurring with the three properties above.

1.2. Hidden Climate Indices Models

The approaches described in the previous section model temporal variability by means of known time-varying covariates. Standard climate indices (SCIs) such as the Southern Oscillation Index (SOI; Ropelewski & Jones, 1987), the North Atlantic Oscillation (NAO) index (Hurrell & Van Loon, 1997), and many others (NCAR, 2019) are typically chosen for this purpose. However several authors reported that identifying relevant SCIs is sometimes difficult (Giuntoli et al., 2013; Grantz et al., 2005; Renard & Lall, 2014; Westra & Sharma, 2009). In such cases, an alternative is to treat the time-varying covariates as unknown temporal latent variables that need to be inferred from the target data. Hidden Markov models provide an example of such an approach, in which the temporal latent variable takes the form of a categorical variable describing the unknown climate state, with a hyperdistribution controlling the probability of transition between states (Bracken, Rajagopalan, & Woodhouse, 2016; Thyer & Kuczera, 2003a, 2003b). Replacing the categorical climate state with a continuous variable leads to a hidden climate index (HCI) model, as proposed by Renard and Lall (2014) and Ahn et al. (2017). These earlier models made strong assumptions (a single temporal HCI with constrained spatial effects), effectively limiting their application to fairly small regions. Renard and Thyer (2019) addressed these limitations by enabling the use of several HCIs and by using a more realistic description of the spatial variability of HCI effects. They also demonstrated, using synthetic and real-life case studies, that estimating both hidden climate indices and their effects from a multisite data set is feasible and not prone to overfitting when the number of sites is large.

While HCI models have been originally motivated by the need to overcome the poor predictive ability of SCIs for some data sets, Renard and Thyer (2019) suggested that HCI models may also be of interest to indirectly model spatial dependence. Spatial dependence between data is not modeled explicitly in HCI models, but instead is induced indirectly by nearby sites being affected by the same temporal HCIs. In other words, spatial dependence arises from temporal covariability. This indirect treatment is of great practical interest for the following reasons:

1. the treatment of missing values is straightforward, which is a major advantage for highly irregular station-based data sets
2. discrete variables can easily be handled, which offers advantages over copula-based approaches for which specific difficulties arise with discrete variables (e.g., nonuniqueness, nonidentifiability, see Genest & Nešlehová, 2007; Faugeras, 2017)

The potential of HCI models to represent dependence is further highlighted by some similarities with several approaches from the statistical literature. In particular, standard methods such as principal component analysis (PCA) or canonical correlation analysis (CCA) have been reinterpreted as Gaussian models with unknown latent variables (Bach & Jordan, 2005; Klami et al., 2013; Tipping & Bishop, 1999). In the case of a space-time data set, these latent variables play a similar role to the HCIs (but within a restrictive Gaussian framework). Alternatively, the max-stable spatial model of Reich and Shaby (2012) also uses time-varying latent variables to describe spatial dependence between extreme data.

This literature review highlights the potential of HCI approaches to model space-time variability for a wide range of data sets. However, the models proposed to date are restricted to specific distributions that may limit their practical use: Gaussian distribution for PCA/CCA-like models, max-stable distribution in the model of Reich and Shaby (2012), Bernoulli distribution in the model of Renard and Thyer (2019). There is scope to develop a more general framework that would enable the use of any user-specified distribution. In addition, this generalization should also account for the multivariable case, where several distinct distributions might be necessary to model several variables.

1.3. Objectives

This paper describes a general HCI framework for modeling space-time multivariable data. It generalizes existing HCI models by considering several variables following any user-specified distribution. A case study illustrates its application for jointly modeling hydrological droughts, meteorological droughts, and heat waves at many stations in Southeast Australia. The key objectives are as follows:

1. Develop a HCI framework that can handle a wide range of data types and data sets, including
   (a) continuous and discrete variables, or potentially a combination of both in the same data set
   (b) irregular station-based data sets, including missing values, noncolocated variables and censored data

2. Demonstrate that this HCI framework is able to
   (a) reliably describe temporal variability
   (b) reliably describe both spatial and intervariable dependence
   (c) improve upon SCI-based models on both points

Compared to our previous work on this topic (Renard & Thyer, 2019), the HCI framework developed in this paper is much more general. It moves from a single-variable to a multivariable setup, and it enables the use of: (a) any distribution (our previous work was restricted to a Bernoulli distribution); (b) temporal processes for HCIs; and (c) censoring.

The remainder of the paper is organized as follows. Section 2 describes the HCI modeling framework, including the specification of the model, its inference, and its use for making predictions. Section 3 describes the data and models used in a case study aimed at modeling hot-and-dry summer conditions in Southeast Australia, with Section 4 describing the results. Section 5 discusses current limitations and avenues for future work, and the conclusion Section 6 summarizes the main outcomes of this work.

2. Theory

This section begins with a short example to motivate HCI modeling (Section 2.1) and introduce its main principles. The full HCI modeling framework is then described in Sections 2.2–2.4.

2.1. Motivating Example

Consider the two time series $(Y_{1,t})_{t=1,\ldots,n}$ and $(Y_{2,t})_{t=1,\ldots,n}$ representing, for instance, annual streamflow anomalies observed at two locations. Assume that both series are linked to the same temporal covariate $\tau_t \overset{iid}{\sim} \mathcal{N}(0,1)$ (e.g., a climate index) by the following linear relations:

$$\begin{cases} Y_{1,t} = \lambda_1 \tau_t + \varepsilon_{1,t} \;\text{ with }\; \varepsilon_{1,t} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right) \\ Y_{2,t} = \lambda_2 \tau_t + \varepsilon_{2,t} \;\text{ with }\; \varepsilon_{2,t} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right). \end{cases} \tag{1}$$

Parameters λ_1 and λ_2 are called the "effects of the covariate" and control the strength of its influence at each site. Random variables $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$, with common standard deviation σ, describe the variability unexplained by the covariate τ_t. The common temporal covariate induces dependence between the two time series. Indeed, assuming mutual independence between τ_t, $\varepsilon_{1,t}$, and $\varepsilon_{2,t}$, the correlation between $Y_{1,t}$ and $Y_{2,t}$ can be derived as

$$\rho = \frac{\lambda_1 \lambda_2}{\sqrt{\lambda_1^2 + \sigma^2}\sqrt{\lambda_2^2 + \sigma^2}}. \tag{2}$$

The assumption of independence between $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ may not always be realistic in practice, but it allows illustrating the fact that spatial dependence can arise purely from temporal covariability. Figure 1 illustrates this simple example for a few values of spatial effects λ_1, λ_2. It shows that spatial correlation is strong when the effects (λ_1, λ_2) are large compared to the standard deviation σ (top panels). In other words, spatial dependence occurs as a consequence of sites following the same temporal pattern. Conversely, dependence vanishes when at least one of the effects is negligible with respect to σ (bottom panels).

Next, consider the practical situation where the modeler wishes to derive a probabilistic model for the data shown in Figure 1. At least three approaches are available:

1. The scatterplot representation (panels A and D) suggests modeling the joint distribution of data using a bivariate Gaussian distribution, as given in Equation 3

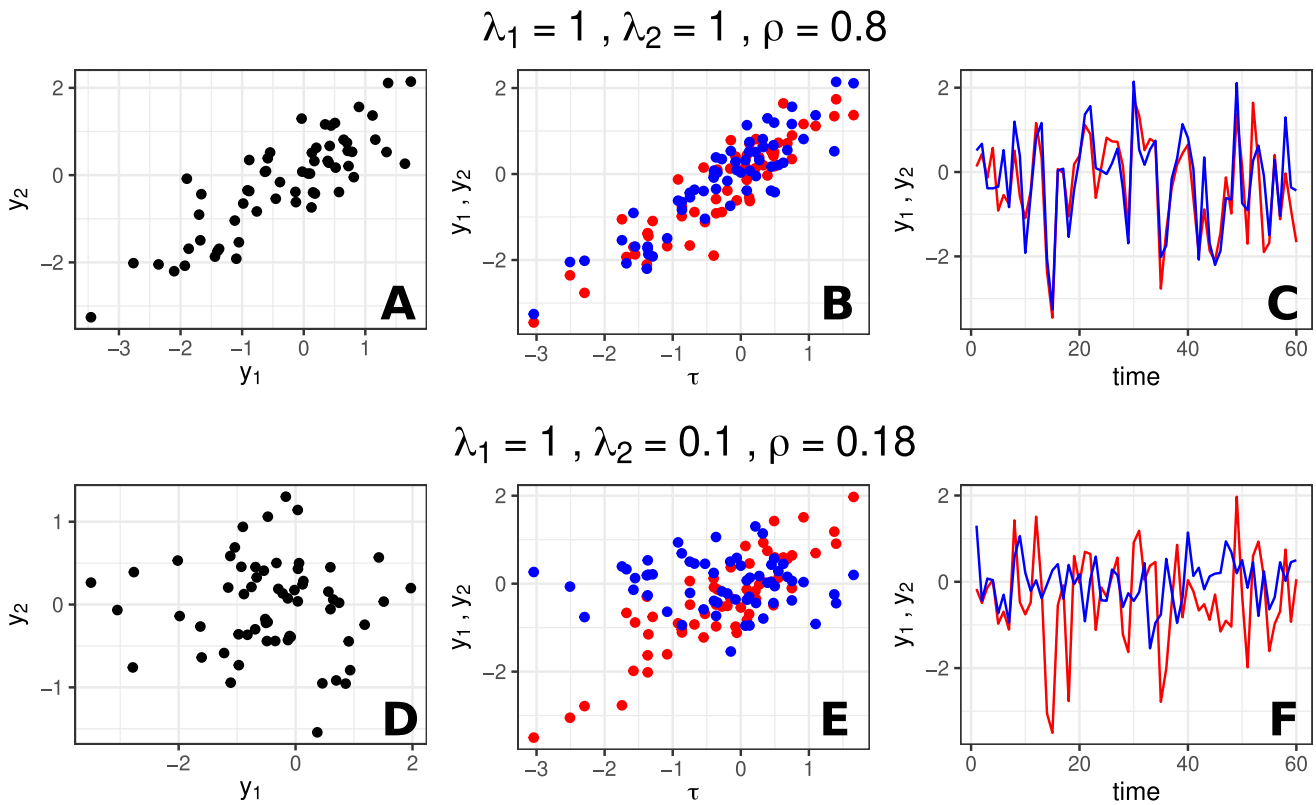

Figure 1. Data $\boldsymbol{y}_1$ (red) and $\boldsymbol{y}_2$ (blue) simulated from the model in Equation 1 with residual standard deviation $\sigma = 0.5$. The covariate effects (λ_1, λ_2) and the correlation ρ between $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ are shown in the titles. (a and d) Scatterplot representation; (b and e) covariate representation; and (c and f) time series representation.

$$(Y_{1,t}, Y_{2,t}) \overset{iid}{\sim} \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}\right). \tag{3}$$

2. The covariate representation (panels B and E) suggests modeling the distribution of data conditionally on the temporal covariate τ_t (Equation 4). This approach is more general in the sense that it leads to exactly the same joint distribution as in Equation 3 (with $\sigma_i^2 = \lambda_i^2 + \sigma^2$) while enabling more precise time-varying predictions according to the covariate values. However, this approach does assume that the data-generating covariate τ_t is known perfectly, which is not the case in general

$$\begin{cases} (Y_{1,t}|\tau_t = u) \overset{i}{\sim} \mathcal{N}\left(\lambda_1 u, \sigma^2\right) \\ (Y_{2,t}|\tau_t = u) \overset{i}{\sim} \mathcal{N}\left(\lambda_2 u, \sigma^2\right) \end{cases}. \tag{4}$$

3. Acknowledge that the data-generating temporal covariate is hidden, which is the approach proposed in this paper. This approach still uses Equation 4 but considers the values taken by τ_t as unknown quantities that need to be estimated from the data. For an intuitive illustration, consider panel C in Figure 1. While the time series τ_t is not shown, it is suggested by the covariability of the time series $y_{1,t}$ and $y_{2,t}$. Compared with the first two approaches, treating the temporal covariate as hidden introduces many unknown quantities: with $p = 2$ series of size n each, there are $q = n$ (τ's) + 2 (λ's) + 1 (σ) quantities to be estimated (versus only 3 for the first two approaches). While such a large number of unknown parameters might be prohibitive and prone to overfitting given data from only two sites, it becomes more tractable as the number of sites increases because $q = n + p + 1$ becomes small relative to the number of data points (np)

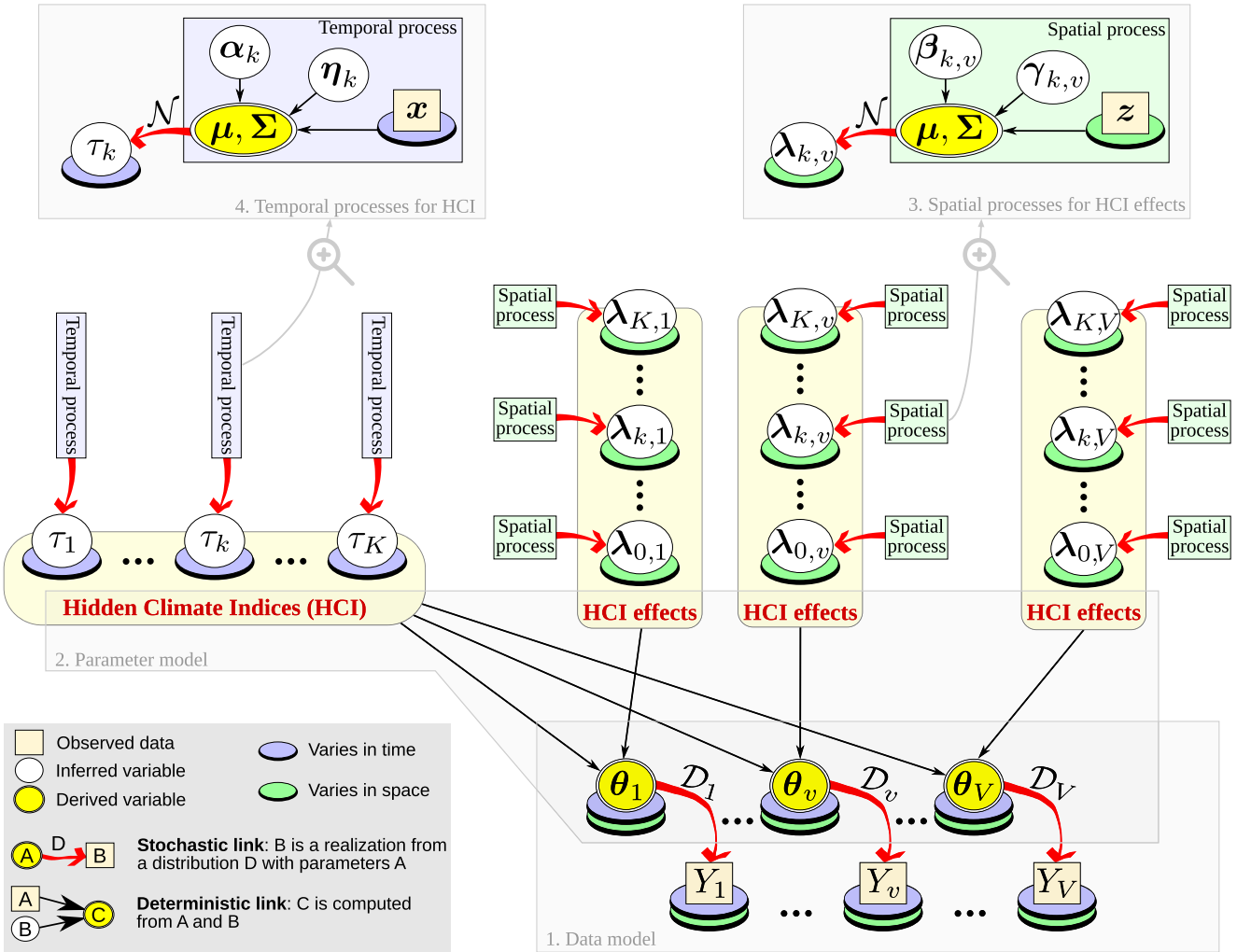

Figure 2. Schematic of the hidden climate index (HCI) modeling approach.

The HCI modeling framework derived in the next sections represents a generalization of Approach 3 above. In particular, it allows modeling multiple variables with non-Gaussian distributions at many sites using multiple HCIs.

2.2. Model Formulation

Let $Y_v(\boldsymbol{s}, t)$ denote the random variable representing variable v (=1, …, V) at time t and site $\boldsymbol{s}$ (typically, $\boldsymbol{s} = (lon, lat)$). The assumptions made when building an HCI model are illustrated in Figure 2, and can be summarized as follows:

1. Data model: For each variable v, $Y_v(\boldsymbol{s}, t)$ follows a distribution with parameters $\boldsymbol{\theta}_v(\boldsymbol{s}, t)$ varying in space and time
2. Parameter model: Parameters $\boldsymbol{\theta}_v(\boldsymbol{s}, t)$ can be retrieved from a set of hidden climate indices $\boldsymbol{\tau}(t)$ which vary in time and their effects $\boldsymbol{\lambda}(\boldsymbol{s})$ which vary in space
3. Space variability: The effects of HCIs $\boldsymbol{\lambda}(\boldsymbol{s})$ are realizations from spatial random processes
4. Time variability: The HCIs $\boldsymbol{\tau}(t)$ themselves are realizations from temporal random processes

These assumptions can be formalized as follows. First, the data model is (box 1 in Figure 2)

$$Y_v(\boldsymbol{s}, t) \sim \mathcal{D}_v\left(\boldsymbol{\theta}_v(\boldsymbol{s}, t)\right). \tag{5}$$

The distribution $\mathcal{D}_v$ is variable-specific. For instance if Y_1 denotes an annual number of flood events and Y_2 denotes an annual average flow, then it would be sensible to use the Poisson distribution for $\mathcal{D}_1$ and the lognormal distribution for $\mathcal{D}_2$.

The parameter vector $\boldsymbol{\theta}_v(\boldsymbol{s}, t)$, with size P_v, is also variable-specific. Each component $\theta_{v,c}$ ($c = 1, \ldots, P_v$) is allowed to vary in space and time as follows (box 2 in Figure 2)

$$g_{v,c}\left(\theta_{v,c}(\boldsymbol{s}, t)\right) = \lambda_{0,v,c}(\boldsymbol{s}) + \lambda_{1,v,c}(\boldsymbol{s})\tau_1(t) + \cdots + \lambda_{K,v,c}(\boldsymbol{s})\tau_K(t). \tag{6}$$

Similar to generalized linear models (McCullagh & Nelder, 1989), the link function $g_{v,c}$ is used to map the range of parameter $\theta_{v,c}$ to $(-\infty, +\infty)$. For instance the logarithm function can be used if $\theta_{v,c} > 0$, or the logit function if $\theta_{v,c} \in (0, 1)$. The transformed parameter is then derived from a set of K time-varying HCIs $\boldsymbol{\tau}$ and their space-varying effects $\boldsymbol{\lambda}$. The following comments can be made:

1. Unlike the link function $g_{v,c}$ and the spatial effect $\lambda_{k,v,c}$, the temporal HCI τ_k is assumed to be the same for all variables. The motivation is that using a common set of HCIs for all variables can induce intervariable dependence, as illustrated in the motivating example of Section 2.1
2. A nonlinear formulation could be used in lieu of Equation 6 (see discussion Section 5.3)
3. If a parameter does not vary in time, then one can set $\lambda_{k,v,c} = 0$ for all $k \geq 1$

Each spatial HCI effect $\lambda_{k,v,c}$ is assumed to vary according to a spatial Gaussian process (box 3 in Figure 2). By definition, this continuous process is characterized by its application to any finite set of sites resulting in a multivariate Gaussian distribution. Hence, for any set of m sites

$$\left(\lambda_{k,v,c}(\boldsymbol{s}_1), \ldots, \lambda_{k,v,c}(\boldsymbol{s}_m)\right) \sim \mathcal{N}\left(\boldsymbol{\mu}\left[\boldsymbol{\beta}_{k,v,c}\right], \boldsymbol{\Sigma}\left[\boldsymbol{\gamma}_{k,v,c}\right]\right). \tag{7}$$

The mean of this multivariate Gaussian distribution is defined through a mean function parameterized by $\boldsymbol{\beta}$ (k, v, and c subscripts have been removed to reduce cluttering). The simplest case is to use a constant mean $\boldsymbol{\mu}[\beta] \equiv \beta$, but covariates such as elevation $\boldsymbol{z} = (z(\boldsymbol{s}_1), \ldots, z(\boldsymbol{s}_m))$, for instance, can be accounted for through a regression such as $\boldsymbol{\mu}[\boldsymbol{\beta}] = \beta_0 + \beta_1 \boldsymbol{z}^T$. A covariance function parameterized by $\boldsymbol{\gamma}$ is also used, with $\boldsymbol{\gamma}$ typically representing sill and range parameters. For instance, an exponential covariance function with sill γ_1 and range γ_2 can be written as a function of the intersite distance $d_{i,j}$ as follows:

$$\left(\boldsymbol{\Sigma}[\boldsymbol{\gamma}]\right)_{i,j} = \gamma_1^2 \exp(-d_{i,j}/\gamma_2). \tag{8}$$

Similar to spatial HCI effects being modeled with a spatial Gaussian process, each temporal HCI τ_k is assumed to vary according to a temporal Gaussian process (box 4 in Figure 2)

$$\left(\tau_k(t_1), \ldots, \tau_k(t_m)\right) \sim \mathcal{N}\left(\boldsymbol{\mu}\left[\boldsymbol{\alpha}_k\right], \boldsymbol{\Sigma}\left[\boldsymbol{\eta}_k\right]\right). \tag{9}$$

The mean of this process may be constant or it may depend on some time-varying covariate $\boldsymbol{x} = (x(t_1), \ldots, x(t_m))$, for instance through a linear regression $\boldsymbol{\mu}[\boldsymbol{\alpha}] = \alpha_0 + \alpha_1 \boldsymbol{x}^T$. The covariance function may be specified using a model similar to Equation 8 (replacing intersite distance by time lag) or a standard time series model (e.g., autoregressive). Note that in the special case where τ_k's are known covariates (e.g., some predefined climate indices), the model defined by Equations 5–7 reduces to a standard hierarchical regression model (Congdon, 2010), also known as a random effect model (Longford, 1993). The originality of the HCI approach is to model time variability using (unknown) latent variables rather than (known) covariates.

The formulation of the model is completed by assuming that all $Y_v(\boldsymbol{s}, t)$ are independent, conditional on the values taken by the temporal HCIs $\boldsymbol{\tau}$ and their spatial effects $\boldsymbol{\lambda}$. In mathematical terms, this means that for any pair $\left(Y_{v_1}(\boldsymbol{s}_1, t_1), Y_{v_2}(\boldsymbol{s}_2, t_2)\right)$ (with $v_1 \neq v_2$ or $\boldsymbol{s}_1 \neq \boldsymbol{s}_2$ or $t_1 \neq t_2$ so that variables are distinct), the conditional joint pdf evaluated at values (y_1, y_2) is equal to

$$p(y_1, y_2 | \boldsymbol{\tau}, \boldsymbol{\lambda}) = p(y_1 | \boldsymbol{\tau}, \boldsymbol{\lambda})\, p(y_2 | \boldsymbol{\tau}, \boldsymbol{\lambda}). \tag{10}$$

It is stressed that such a conditional independence assumption does not imply that $Y_v(\boldsymbol{s}, t)$ are unconditionally independent, as illustrated in the motivating example of Section 2.1. In mathematical terms

$$p(y_1, y_2) \neq p(y_1)\, p(y_2). \tag{11}$$

2.3. Inference

2.3.1. Posterior Distribution of Inferred Parameters

Let $\boldsymbol{y} = (y_i)_{i=1,\ldots,N}$ denote the calibration data set stored in a "long vector" format, leading to a total of N data points. Each value y_i is associated with the following three elements: (a) a variable index v_i in 1, …, V; (b) a site $\boldsymbol{s}_i$ amongst the S calibration sites; and (c) a time step t_i amongst the T calibration time steps. In the particular case where all variables are measured at all sites and all time steps, $N = V \times S \times T$. In general, N is smaller than this product because distinct variables may be measured on distinct measurement networks and time series may comprise missing values.

The posterior distribution of unknown temporal HCIs $\boldsymbol{\tau}$, their spatial effects $\boldsymbol{\lambda}$, and hyperparameters $(\boldsymbol{\alpha}, \boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma})$ can be written as

$$p(\boldsymbol{\tau}, \boldsymbol{\lambda}, \boldsymbol{\alpha}, \boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma} | \boldsymbol{y}) \propto \underbrace{p(\boldsymbol{y}|\boldsymbol{\tau}, \boldsymbol{\lambda})}_{\text{likelihood}} \underbrace{p(\boldsymbol{\tau}|\boldsymbol{\alpha}, \boldsymbol{\eta})\, p(\boldsymbol{\lambda}|\boldsymbol{\beta}, \boldsymbol{\gamma})}_{\text{time/space hyperdist.}} \underbrace{p(\boldsymbol{\alpha}, \boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma})}_{\text{priors}}. \tag{12}$$

Due to the conditional independence assumption, the likelihood is a simple product of the contributions l_i of each data point

$$p(\boldsymbol{y}|\boldsymbol{\tau}, \boldsymbol{\lambda}) = \prod_{i=1}^{N} \underbrace{f_{v_i}\left(y_i; \boldsymbol{\theta}_{v_i}(\boldsymbol{s}_i, t_i)\right)}_{l_i}, \tag{13a}$$

$$\boldsymbol{\theta}_{v_i}(\boldsymbol{s}_i, t_i) = \left(\theta_{v_i,1}, \ldots, \theta_{v_i,P_{v_i}}\right), \tag{13b}$$

$$\theta_{v_i,c} = g_{v_i,c}^{-1}\left(\lambda_{0,v_i,c}(\boldsymbol{s}_i) + \lambda_{1,v_i,c}(\boldsymbol{s}_i)\tau_1(t_i) + \cdots + \lambda_{K,v_i,c}(\boldsymbol{s}_i)\tau_K(t_i)\right), \tag{13c}$$

where f_{v_i} in Equation 13a is the pdf of distribution $\mathcal{D}_{v_i}$, and Equations 13b and 13c derive the associated parameter vector $\boldsymbol{\theta}_{v_i}(\boldsymbol{s}_i, t_i)$ using Equation 6.

Note that this likelihood can be easily modified to use censored data. If all that is known about the ith data is that it lies in the interval $[a, b]$ (with possibly $a = -\infty$ or $b = +\infty$), then its contribution l_i to the likelihood uses the cdf, rather than the pdf, of distribution $\mathcal{D}_{v_i}$, as shown in Equation 14 (e.g., Payrastre et al., 2011). This approach can also be used to model continuous variables with reachable bounds. Typically, for precipitation or streamflow, a zero value can be considered as belonging to the interval $(-\infty, 0]$. This is similar to a "Tobit regression" approach (Chib, 1992)

$$l_i = F_{v_i}\left(b; \boldsymbol{\theta}_{v_i}(\boldsymbol{s}_i, t_i)\right) - F_{v_i}\left(a; \boldsymbol{\theta}_{v_i}(\boldsymbol{s}_i, t_i)\right). \tag{14}$$

Following Equations 7 and 9, the terms related to the space and time hyperdistributions are computed using the pdf $f_{\mathcal{N}}(\boldsymbol{z}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ of a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$

$$p(\boldsymbol{\tau}|\boldsymbol{\alpha}, \boldsymbol{\eta}) = \prod_{k=1}^{K} f_{\mathcal{N}}\left(\boldsymbol{\tau}_k; \boldsymbol{\mu}[\boldsymbol{\alpha}_k], \boldsymbol{\Sigma}[\boldsymbol{\eta}_k]\right), \tag{15}$$

$$p(\boldsymbol{\lambda}|\boldsymbol{\beta}, \boldsymbol{\gamma}) = \prod_{k=0}^{K} \prod_{v=1}^{V} \prod_{c=1}^{P_v} f_{\mathcal{N}}\left(\boldsymbol{\lambda}_{k,v,c}; \boldsymbol{\mu}[\boldsymbol{\beta}_{k,v,c}], \boldsymbol{\Sigma}[\boldsymbol{\gamma}_{k,v,c}]\right). \tag{16}$$

2.3.2. Overcoming Nonidentifiability

In Equation 6, there are infinitely many values of HCI τ_k and effect λ_k $(k > 0)$ that yield exactly the same value of parameter θ (e.g., dividing τ_k and multiplying λ_k by the same value does not change the value of θ). As a consequence, the parameters in the likelihood function (Equation 13) are not identifiable, leading to an ill-posed posterior distribution. Renard and Thyer (2019) discuss the origin of this nonidentifiability in more depth, and propose a two-part practical solution to overcome it:

1. Apply identifiability constraints: For each k, the temporal HCI estimated at T calibration time steps $\left(\tilde{t}_1, \ldots, \tilde{t}_T\right)$ is constrained to have zero mean and unit standard deviation. These two constraints imply that only the first $T-2$ values are estimated, with $\tau_k(\tilde{t}_{T-1})$ and $\tau_k(\tilde{t}_T)$ being computed from $\left(\tau_k(\tilde{t}_1), \ldots, \tau_k(\tilde{t}_{T-2})\right)$ so as to meet the constraints (see Renard & Thyer, 2019, for precise formulae)
2. Use a stepwise inference: A model with a single HCI is estimated first, leading to estimates $\hat{\lambda}_0$, $\hat{\lambda}_1$, and $\hat{\boldsymbol{\tau}}_1$ (maximum-posterior estimates are used here). These estimates are then used as known, fixed values in a two-HCI model; this leads to estimates $\hat{\lambda}_2$ and $\hat{\boldsymbol{\tau}}_2$. This process is repeated until the chosen number of components is reached (see Section 2.3.4 for strategies to select this number)

2.3.3. Markov Chain Monte Carlo Sampling

The posterior distribution of Equation 12 is explored by means of a block Metropolis sampler. The parameter vector is updated one component at a time (this is sometimes referred to as a Metropolis-within-Gibbs sampler). The sampler periodically adapts the size of the univariate Gaussian jump distributions to achieve user-specified acceptance rates. In addition, the sampler takes advantage of the many simplifications occurring in the Metropolis ratio to reduce the computational cost. This sampler is very similar to the one described in Renard and Thyer (2019), generalized to account for multiple variables. A detailed description is therefore not repeated here.

2.3.4. Selecting the Number of Components

While there is no unique optimal way to select the number of components in the HCI model, several tools exist to help making this decision. A first strategy is to stop adding components when the effects $\hat{\boldsymbol{\lambda}}$ become small for all variables. This can be quantified by computing, for each variable v and component k, the standard effect $\zeta_{k,v}$ across the S_v available sites (Renard & Thyer, 2019)

$$\zeta_{k,v} = \sqrt{\frac{1}{S_v}\sum_{i=1}^{S_v}\left(\hat{\lambda}_{k,v}(s_i)\right)^2}. \tag{17}$$

In addition, Renard and Thyer (2019) showed, based on a synthetic case study, that MCMC convergence tends to deteriorate when attempting to infer "useless" HCIs (i.e., more HCIs than used to generate the synthetic data). Finally, it is also possible to rely on standard model-checking procedures (e.g., comparing observed frequencies and estimated probabilities, setting up cross-validation experiments) to select a fit-for-purpose number of components.

An automatic model selection strategy, such as selecting the number of components that minimizes the Deviance Information Criterion (DIC, Spiegelhalter et al., 2002), could also be used in principle. However, the stepwise estimation procedure proposed in this paper makes this strategy questionable, because fixing the parameters estimated at the previous step may unduly favor models with many components. We therefore favor the nonautomatic but more flexible approaches based on monitoring standard effects, MCMC convergence, and applying model-checking procedures.

2.4. Predictions

2.4.1. Predictions at Calibration Sites and Time Steps

Once temporal HCIs $\hat{\boldsymbol{\tau}}$ and their spatial effects $\hat{\boldsymbol{\lambda}}$ have been estimated as described in the previous section, the HCI model can be used to make probabilistic predictions.

For a given variable v, a probabilistic prediction at site $\boldsymbol{s}^*$ and time t^* can be made by computing the parameters of distribution $\mathcal{D}_v$ using Equation 6. More precisely, the pdf associated with this probabilistic prediction is given by the equation below. Monte Carlo samples can be generated from this distribution to generate ensemble predictions.

$$f_v\left(u; \hat{\boldsymbol{\theta}}_v(\boldsymbol{s}^*, t^*)\right), \tag{18a}$$

$$\hat{\boldsymbol{\theta}}_v(\boldsymbol{s}^*, t^*) = \left(\hat{\theta}_{v,1}, \ldots, \hat{\theta}_{v,P_v}\right), \tag{18b}$$

$$\hat{\theta}_{v,c} = g_{v,c}^{-1}\left(\hat{\lambda}_{0,v,c}(\boldsymbol{s}^*) + \hat{\lambda}_{1,v,c}(\boldsymbol{s}^*)\hat{\tau}_1(t^*) + \cdots + \hat{\lambda}_{K,v,c}(\boldsymbol{s}^*)\hat{\tau}_K(t^*)\right). \tag{18c}$$

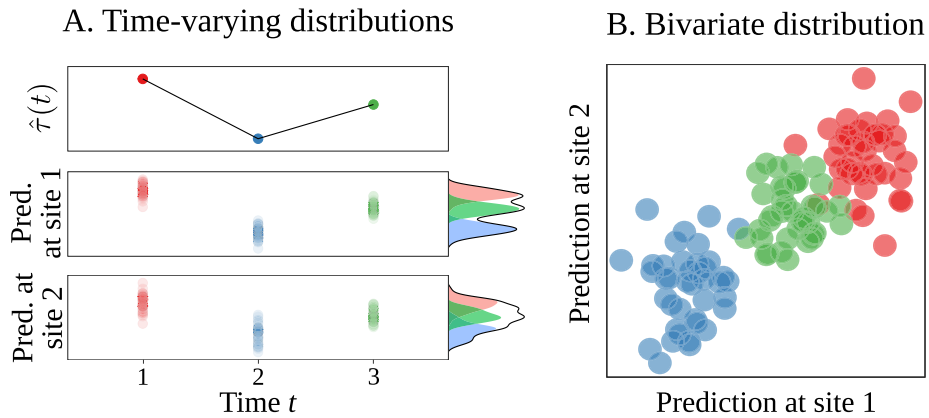

Figure 3. Schematic illustration of probabilistic predictions. (a) Simulation conditional on the values taken by the hidden climate index (HCI) at three time steps. The resulting mixture marginal pdfs are shown on the right axis. (b) Bivariate distribution resulting from the simulations at two sites, illustrating the existence of spatial dependence.

A probabilistic prediction at several time steps (but still a single variable and site) can be obtained by sampling from the distribution in Equation 18a independently for each time step. The time-varying nature of this prediction is induced by the time-varying HCIs $\hat{\boldsymbol{\tau}}$, as schematized in Figure 3a. It is also possible to obtain the mixture marginal distribution over a period of interest by treating the Monte Carlo samples at multiple time steps as a single set of samples.

Thanks to the conditional independence assumption, generating multivariable and/or multisite predictions is equally straightforward. It simply consists of generating independent replicates for several variables and/or sites, conditional on the same HCI values. As illustrated in Figure 3b, spatial dependence will be present for sites where the HCI effects $\hat{\boldsymbol{\tau}}$ are similar. The exact same mechanism applies to intervariable dependence.

Note that all predictions discussed above only require estimates of temporal HCIs ($\hat{\boldsymbol{\tau}}$) and their spatial effects ($\hat{\boldsymbol{\lambda}}$) and can therefore be made for any combination of site and time step that are part of the calibration data set. This includes combinations that were potentially unobserved. For instance, consider a hypothetical data set with variable v_1 being observed during the period 1951–2020, but variable v_2 during the period 1961–2020 only. Predictions for variable v_2 during the missing decade 1951–1960 can still be made since estimates are available for both the HCIs $\hat{\boldsymbol{\tau}}$ during 1951–1960 (based on v_1 data) and the HCI effects $\hat{\boldsymbol{\lambda}}$ at v_2 sites. This provides a mechanism to transfer information from an observed variable to an unobserved one.

2.4.2. Predictions at Ungauged Sites or Unobserved Time Steps

Prediction at an ungauged site $\boldsymbol{s}^*$ requires interpolating the HCI effects estimated at calibration sites. This can be achieved by using the spatial process in Equation 7. More precisely, since this process is assumed to be Gaussian, the interpolated value for an HCI effect λ (subscripts have been dropped to simplify notation) is also Gaussian with mean and variance given below (Renard, 2011) and corresponding to standard Kriging formulas

$$\mu_* = \boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}\hat{\boldsymbol{\lambda}}^\top, \tag{19a}$$

$$\sigma_*^2 = \sigma^2 - \boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Omega}^\top, \tag{19b}$$

where $\boldsymbol{\Omega}$ is the $1 \times S$ covariance vector between the ungauged site and the calibration sites, $\boldsymbol{\Sigma}$ is the $S \times S$ covariance matrix between calibration sites, $\hat{\boldsymbol{\lambda}}$ is the $1 \times S$ estimated values of λ at calibration sites, σ^2 is the marginal variance of the process and all variances/covariances are computed using the covariance function assumed for the process (e.g., Equation 8).

Prediction for an unobserved time step can be made in a similar way by using the temporal process in Equation 9 to extrapolate the HCIs a few time steps ahead. However, a potentially more skillful alternative would be to predict the HCIs from large-scale climate information, as discussed in Section 5.4.

Table 1
Description of the Data Sets Used in the Case Study

Variables	Data set name	Time	Space	References
Temperature	ACORN-SAT[a]	1918–2018	49 stations	Trewin (2013, 2018)
Precipitation	High-quality daily rainfall[b]	1918–2018	76 stations	Bureau of Meteorology (2020a)
Streamflow	Hydrologic Reference Stations[c]	1951–2019	271 stations	Zhang et al. (2014) and Bureau of Meteorology (2020b)
ENSO	NINO4 index[d]	1870–2018	—	NOAA (2020d)
IOD	DMI index[e]	1870–2019	—	NOAA (2020c)
SAM	SAM index[f]	1851–2011	—	NOAA (2020a)
SAM	AAO index[g]	1979–2020	—	NOAA (2020b)

[a]http://www.bom.gov.au/climate/data/acorn-sat. [b]ftp://ftp.bom.gov.au/anon/home/ncc/www/change/HQdailyR/. [c]http://www.bom.gov.au/water/hrs/. [d]https://www.esrl.noaa.gov/psd/gcos_wgsp/Timeseries/Data/nino4.long.anom.data. [e]https://psl.noaa.gov/gcos_wgsp/Timeseries/Data/dmi.had.long.data. [f]https://www.esrl.noaa.gov/psd/data/20thC_Rean/timeseries/monthly/SAM/sam.20crv2c.long.data. [g]https://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/aao/monthly.aao.index.b79.current.ascii.table.

3. Case Study: Data and Models

The proposed framework is demonstrated using a case study describing "hot-and-dry" summer conditions in Southeast Australia. An HCI model is used to describe streamflow, precipitation, and temperature variables measured at hundreds of sites during 100 years. The proposed HCI model is compared with a more standard model that uses three SCIs that have a well-documented effect on Australian summer weather. This case study has been selected because it allows illustrating the versatility of the HCI framework, while being connected to very impactful hazards in Australia: heat waves constitute the deadliest natural hazard in the country and their severity shows an increasing trend that is expected to continue in the future (Perkins-Kirkpatrick et al., 2016); drought is also a recurrent hazard, with large impacts on agricultural and other economic sectors (Kiem et al., 2016); both heat waves and droughts are important components of the bushfire hazard (Sharples et al., 2016), which is widespread in Australia as illustrated by the 2019–2020 bushfire season (Abram et al., 2021).

3.1. Data

3.1.1. Surface Variables

High-quality temperature (T), precipitation (P), and streamflow (Q) daily time series are provided by the Australian Bureau of Meteorology (Table 1). Each variable is measured on a different station network, as shown in the maps in Figure 4. Data availability varies strongly in time and between variables. In particular, Q stations are more numerous but are available for a much shorter time period than T/P stations. Moreover, the period pre-1951 is only "partially observed" in the sense that only T and P data are available. Despite this, the analysis period is 1918–2018, and Q predictions are possible before 1951, as will be shown subsequently.

Four variables are extracted from the daily time series to quantify summer (DJF) heat-and-drought conditions:

1. Low flow duration Qd (–) is computed as the number of days with streamflow below the 10th percentile, divided by the number of days during the DJF season (90 or 91 for leap years). This duration is hence expressed as a fraction of the season length between 0 and 1. If any daily data is missing during the season, the variable is considered as missing (the same applies to all four variables)
2. Dry-day duration Pd (–) is computed as the number of dry days ($P < 1$ mm), divided by the number of days during the DJF season
3. Heat wave intensity Tx (°C) is computed as the maximum threshold excess during a heat wave event. A heat wave corresponds to successive days above a high-temperature threshold taken as the 99th percentile, which is slightly more extreme than the standard definition proposed by Nairn et al. (2013). Successive events should be separated by at least 3 days and the temperature should drop 5° below the threshold, otherwise they are considered a single event. Note that Tx is not an absolute temperature but rather a threshold excess (i.e., the temperature minus the 99th percentile threshold)

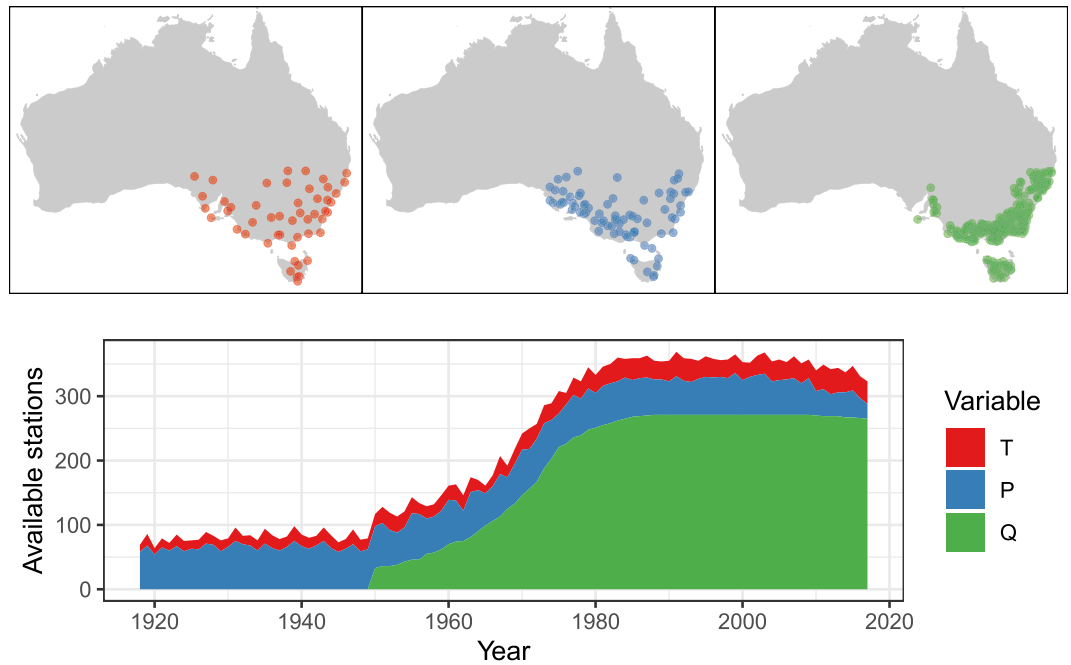

Figure 4. Location of stations used in the case study and data availability.

4. *Tn* (–) is computed as the number of heat wave events occurring during the season

Note that for all variables, large values are associated with dry/hot summers.

3.1.2. Standard Climate Indices

Three DJF-averaged indices are used to characterize climate drivers that are known to influence heat waves and droughts in Australia (Westra et al., 2016), as described in Table 1: the El-Nino Southern Oscillation (ENSO), the Indian Ocean Dipole (IOD), and the Southern Annular Mode (SAM). Note that the Southern Annular Mode could be characterized using two alternative indices (Ho et al., 2012): the SAM index (availability: 1851–2011) and the AAO index (1979–2020). We use the former due to its earlier starting date and reconstructed recent years by means of a linear regression ($r^2 = 0.84$) with the latter. The final SCIs are centered and scaled to unit standard deviation. Since HCIs also have zero mean and unit standard deviation due to the identifiability constraints (Section 2.3.2), this preprocessing enables the comparison of SCI and HCI effects for each studied variable.

3.2. Models and Inference

3.2.1. HCI Model

The HCI model used to describe the four target variables is shown in Equation 20. The number of HCIs was fixed to three in order to enable comparisons with a model based on the three SCIs described in Section 3.1.2. The adequacy of this number will be discussed in the results.

Variables *Qd* and *Pd* are both derived from the number of days below a threshold. At first sight, integer-valued distributions such as the binomial or the negative binomial distributions may be considered as candidates, but they are in fact not well suited for two main reasons: (a) values from successive days can hardly be considered as results from independent and identically distributed trials; (b) the discrete nature of the variables is mostly an artifact induced by the daily time step of the original series; the actual quantity of interest is the continuous duration below the threshold, which would be accessible with a varying-time-step series. We therefore decided to use continuous Gaussian distributions for both *Qd* (Equation 20a) and *Pd* (Equation 20b), with the mean varying in both space and time and the standard deviation varying in space only. The Gaussian distribution is censored

below zero and above one as described in Section 2.3 to account for the nonzero probabilities of seeing durations equal to zero and one. Note that a logit transformation cannot be applied to *Qd* or *Pd* because zero/one values are frequently observed and would be sent to $\pm\infty$.

Heat wave intensities *Tx* are defined as threshold excesses and are therefore modeled with a 2-parameter Generalized Pareto Distribution (Equation 20c). The scale parameter varies in both space and time and the shape parameter varies in space only. Note that an alternative version of this model, with a third time-varying threshold parameter, was tried but abandoned. Indeed, small exceedances are generally selected by the procedure, even during particularly hot years also containing very large exceedances. The threshold is therefore not free to vary much because it is constrained every year by the smallest exceedance.

Finally, the number of heat waves *Tn* (Equation 20d) is modeled with a Poisson distribution with the rate parameter varying in both space and time

$$\begin{cases} Qd(\boldsymbol{s},t) \sim \mathcal{N}\left(\mu_1(\boldsymbol{s},t), \sigma_1(\boldsymbol{s})\right) \\ \mu_1(\boldsymbol{s},t) = \lambda_{0,1,1}(\boldsymbol{s}) + \sum_{k=1}^{3} \lambda_{k,1,1}(\boldsymbol{s})\tau_k(t) \\ \log\left(\sigma_1(\boldsymbol{s})\right) = \lambda_{0,1,2}(\boldsymbol{s}) \end{cases} \tag{20a}$$

$$\begin{cases} Pd(\boldsymbol{s},t) \sim \mathcal{N}\left(\mu_2(\boldsymbol{s},t), \sigma_2(\boldsymbol{s})\right) \\ \mu_2(\boldsymbol{s},t) = \lambda_{0,2,1}(\boldsymbol{s}) + \sum_{k=1}^{3} \lambda_{k,2,1}(\boldsymbol{s})\tau_k(t) \\ \log\left(\sigma_2(\boldsymbol{s})\right) = \lambda_{0,2,2}(\boldsymbol{s}) \end{cases} \tag{20b}$$

$$\begin{cases} Tx(\boldsymbol{s},t) \sim \mathcal{GPD}\left(\sigma_3(\boldsymbol{s},t), \xi_3(\boldsymbol{s})\right) \\ \log\left(\sigma_3(\boldsymbol{s},t)\right) = \lambda_{0,3,1}(\boldsymbol{s}) + \sum_{k=1}^{3} \lambda_{k,3,1}(\boldsymbol{s})\tau_k(t) \\ \xi_3(\boldsymbol{s}) = \lambda_{0,3,2}(\boldsymbol{s}) \end{cases} \tag{20c}$$

$$\begin{cases} Tn(\boldsymbol{s},t) \sim \mathcal{P}\left(\mu_4(\boldsymbol{s},t)\right) \\ \log\left(\mu_4(\boldsymbol{s},t)\right) = \lambda_{0,4,1}(\boldsymbol{s}) + \sum_{k=1}^{3} \lambda_{k,4,1}(\boldsymbol{s})\tau_k(t) \end{cases} \tag{20d}$$

Each spatial term $\lambda_{k,v,c}$ is modeled using its own Gaussian hyperdistribution as described in Equation 7, with a constant mean function and an exponential covariance function (Equation 8). This leads to three unknown hyperparameters (mean, sill, and range) for each spatial term. Likewise, each HCI τ_k is modeled using its own Gaussian hyperdistribution as described in Equation 9. Following preliminary analyses that indicated no clear temporal structure in the estimated HCIs, we used *iid* Gaussian hyperdistributions with constant mean vectors and covariance matrices proportional to the identity matrix. Moreover, the identifiability constraints allowed fixing all means and marginal standard deviations to zero and one, respectively.

Priors for the spatial hyperparameters need to be specified to complete the model. Flat priors are used for mean and sill hyperparameters. For range hyperparameters, we use an exponential prior with scale parameter equal to 1,000 km.

3.2.2. SCI Model

An SCI model is also used for comparison purposes with the HCI model. The equations for the SCI model are identical to those of the HCI model in Equation 20, with the key exception that the unknowns $(\tau_1(t), \tau_2(t), \tau_3(t))$

are replaced with the known values taken by the indices (NINO4(t), DMI(t), SAM(t)). All spatial terms $\lambda_{k,v,c}$ are estimated and the same Gaussian hyperdistributions as described in the previous section are used. Priors for the spatial hyperparameters are also identical.

3.2.3. MCMC Sampling

Four chains of size 120,000 each are run in parallel. The first half of each chain is discarded as burn-in, and the remaining iterations are further thinned by a factor of 60, leaving 1,000 iterations per chain to generate the results presented hereafter. MCMC convergence is assessed by monitoring MCMC traces and computing the Gelman-Rubin criterion (Gelman & Rubin, 1992).

3.3. Results Analysis Strategy

The results of the case study are analyzed in four main steps. The first step focuses on the estimated HCIs $\hat{\boldsymbol{\tau}}$ and their spatial effects $\hat{\boldsymbol{\lambda}}$. The HCIs are described in terms of basic statistical properties such as trend or autocorrelation, and cross-correlations with the three predefined SCIs are computed. The effects of HCIs and SCIs are also compared in terms of strength and spatial distribution.

In the second step, time-varying predictions resulting from both HCI and SCI models are analyzed and compared in terms of reliability and sharpness. Reliability is assessed using the widely used PIT diagram (Laio & Tamea, 2007). Sharpness is characterized by a ratio representing the predictive variance explained by interannual variations, and computed as follows. For a single site and a single variable, let $(w_{t,i})_{t=1\ldots T}^{i=1\ldots N_{rep}}$ denote $N_{rep} = 1{,}000$ Monte Carlo replicates from the distribution estimated at time steps $t = 1\ldots T$, as described in Section 2.4. The sharpness ratio ρ is the ratio between the variance of the predictive means $w_{t,\cdot}$ and the total predictive variance

$$\rho = \frac{\mathrm{Var}[(w_{t,\cdot})_{t=1\ldots T}]}{\mathrm{Var}[(w_{t,i})_{t=1\ldots T}^{i=1\ldots N_{rep}}]}. \tag{21}$$

The sharpness ratio varies between 0 and 1. For a time-invariant climatology prediction, the temporal variance in the numerator of Equation 21 is null, leading to $\rho = 0$. A deterministic prediction would correspond to all replicates being identical, so that the variance in the denominator is purely temporal and is equal to the variance in the numerator, leading to $\rho = 1$.

The third step of the results analysis strategy evaluates the ability of both HCI and SCI models to represent spatial and intervariable dependencies. This is done by computing the probabilities of joint or conditional events according to the models and comparing them to the corresponding observed frequencies.

Finally, the fourth step is a split-sample experiment based on a reduced calibration data set built as follows: (1) for each variable, 1/3 of the sites are removed; (2) for variable *Pd* only, all data from the period 1971–1990 are removed. The HCI model is then used to make predictions at these left-out sites or years, as described in Section 2.4. Comparing these predictions with observed values assesses the ability of the HCI model to transfer information between sites (1) or variables (2).

4. Case Study: Results

4.1. Overview of MCMC Sampling

Computing time for running the MCMC simulations described in Section 3.2.3 on a standard laptop is about 12 hr, which roughly corresponds to 1 hr per 10,000 MCMC iterations. While this is a nonnegligible computational effort, this is not a major impediment in the context of this paper given the absence of any real-time constraint. This computing time also needs to be interpreted in relation to the scale of the analysis: hundreds of sites are considered over a 100-year period, leading to $\sim 26,600$ data points to infer $\sim 1,000$ unknown quantities.

MCMC convergence is overall excellent. Values of the Gelman-Rubin criterion are well below 1.1 for nearly all inferred quantities. A few higher values around 1.2 are noted for values of τ_1 at the beginning of the period or for a couple of spatial hyperparameters. The corresponding MCMC traces are shown in Figure S1 in Supporting Information S1 along with those of other randomly selected parameters.

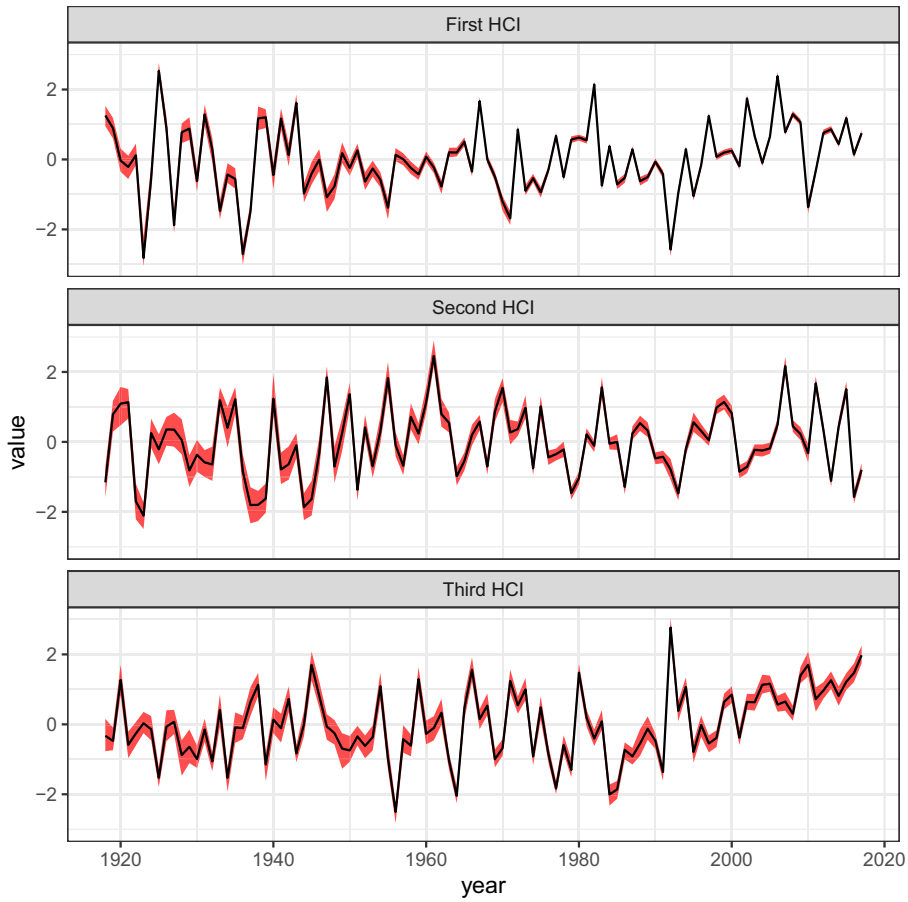

Figure 5. Estimated hidden climate indices (HCIs; posterior median) with 90% credibility intervals.

Posterior correlations are generally very small, with 99% of the off-diagonal terms of the posterior correlation matrix being below 0.1. The highest correlations are observed for spatial hyperparameters, which may explain their slower MCMC convergence. Posterior correlations between parameters related to a same site are all close to zero.

While the number of HCIs was primarily chosen to enable comparisons with a 3-SCI model, it also appears to be an adequate choice. A model including a fourth HCI was estimated and it shows that by the fourth HCI the standard effects have dropped for all variables (Figure S2 in Supporting Information S1). MCMC convergence also deteriorates for this fourth HCI, suggesting that it may not be very informative.

4.2. Estimated HCIs and Their Effects

Figure 5 shows that the three HCIs are precisely estimated in the sense that the credibility intervals are small with respect to the interannual variability. Intervals also tend to be wider at the beginning of the analysis period (Figure S3 in Supporting Information S1), reflecting the smaller number of available stations (Figure 4). The three HCIs represent modes of climate variability that are quite distinct from those represented by the SCIs: cross-correlations are generally small, with the largest one (in absolute value) being equal to 0.34 between the second HCI and the SAM index (not shown).

The first two HCIs show no significant trends or autocorrelation. The third HCI, however, shows a significant upward trend (p-value<0.001, Mann-Kendall test), suggesting that it may be associated with some long-term change

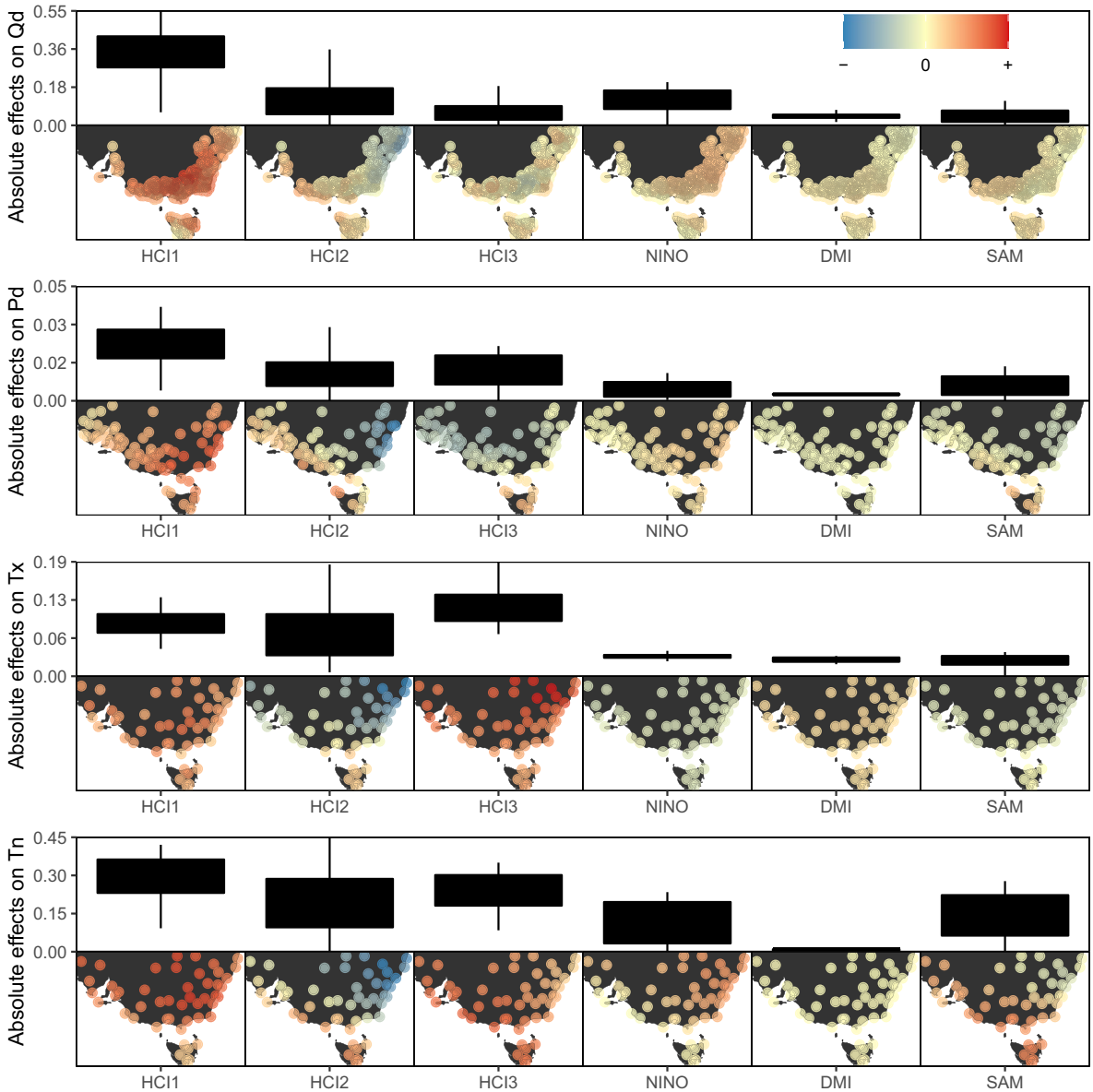

Figure 6. Estimated effects (posterior median) of hidden climate indices (HCIs) and standard climate indices NINO, DMI, and SAM for the four variables of the case study. Boxplots represent the absolute values of the effects and maps allow visualizing spatial patterns. For a given variable, boxplots and maps can be compared across indices.

affecting at least part of the data set. Some autocorrelation coefficients are also significant, but they become non-significant after detrending (not shown). This suggests that low-frequency variability, if it exists, is rather weak.

Figure 6 maps the effects of HCIs and SCIs for the four studied variables. Recall that all HCIs and SCIs are centered and scaled, so that their effects can be compared for a given variable. Overall, HCI effects λ are much larger (in absolute value) than SCI effects, suggesting that the former account for more of the observed variability than the latter. The first HCI has a global effect on all variables and in the whole studied region. The second HCI has both a smaller effect and a more spatially contrasted one, with generally negative effects in the East and positive ones in the West. The effect of the third HCI is quite small for the streamflow variable *Qd* (more than four times smaller than the effect of the first HCI on average). Conversely, it is quite high for temperature variables *Tx* and *Tn*: it is the largest effect amongst the three HCIs for 78% and 37% of the sites, respectively. The fact that this third HCI mostly affects temperature variables suggests that its upward trend may be interpreted as a warming signal.

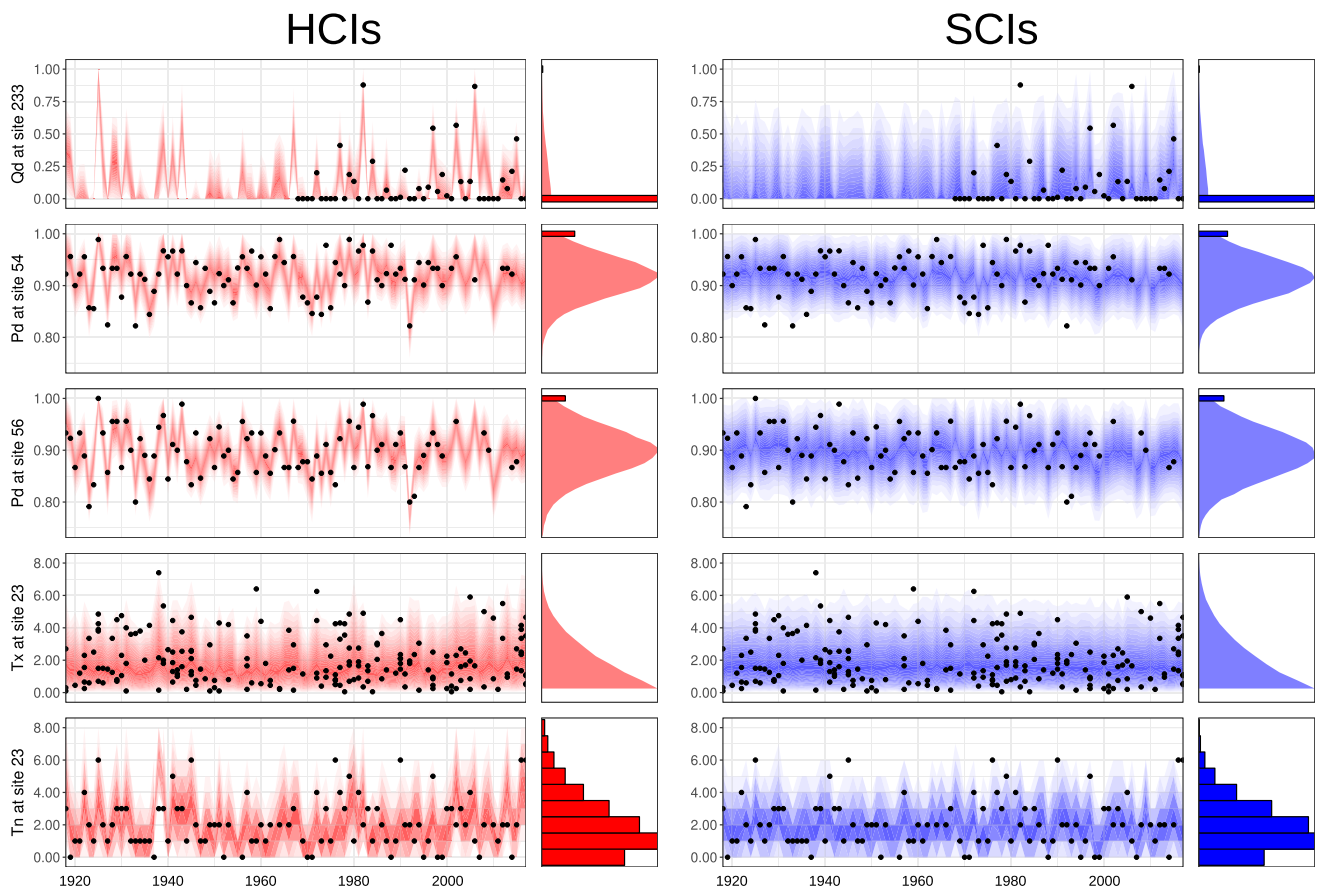

Figure 7. Examples of time-varying predictions obtained using the hidden climate indices (HCI) and the standard climate indices (SCI) models. Colored bands represent nested prediction intervals up to the 95% level, black dots are observed values. Rotated panels on the right represent the mixture marginal distribution (shaded area = continuous pdf, bars = discrete probabilities).

In addition to HCI effects, the model also estimates spatial processes controlling intercepts or time-invariant parameters (λ_0's in Equation 20). Corresponding maps are provided in Figure S4 in Supporting Information S1 and illustrate well-known climate gradients in Australia. For example, the dry-day duration *Pd* tends to get smaller and more variable as one moves eastward in the studied region.

4.3. Time-Varying Predictions

Figure 7 shows the time-varying predictions obtained using the HCI and the SCI models for the four variables at a few sites, as described in Section 2.4. These sites have been selected because they allow illustrating some key features of the models. The three sites for variables *Qd* and *Pd* are located in northern Victoria and are quite close to each other (≈200 km). The temperature site is located ≈600 km further away to the Northeast, to the North of Sydney.

Overall, HCI predictions are sharper than SCI ones: they are more precise and their temporal variability is larger. The latter is a direct consequence of HCI effects being much larger than SCI effects, hence inducing larger temporal variations. Figure 8a shows the distribution of the sharpness ratio at all sites. It indicates that HCI predictions are overall sharper than the SCI predictions across all sites, not just at those shown earlier in Figure 7. The difference in sharpness is particularly salient for variables *Qd* and *Pd*, for which the ratio is five times higher for HCI than for SCI predictions. On the other hand, the sharpness ratio is low for both the HCI and SCI models for variable *Tx*. This is probably due to the fact that the time-varying parameter is the scale parameter of the GPD distribution for this variable (see Equation 20c), which induces temporal variation in the variance of the distribution. This is to be compared with the other three variables for which the time-varying parameter is the location

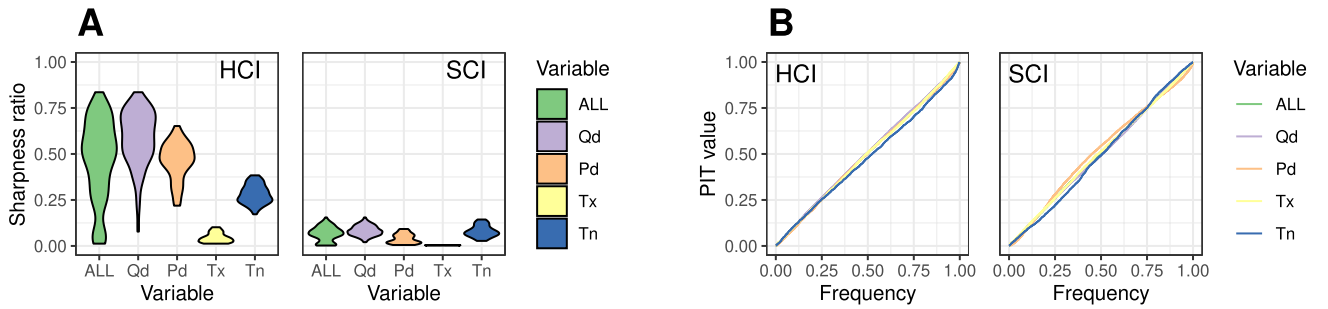

Figure 8. (a) Distributions (across stations) of the sharpness ratio. (b) PIT diagram to evaluate the reliability of predictions. For one given variable, the PIT curve is derived from all stations.

parameter (see Equation 20), inducing temporal variations in the mean. Finally, the PIT diagrams in Figure 8b indicate that both the HCI and SCI predictions are reliable across all sites. This implies that the larger sharpness of HCI predictions compared with SCI ones comes at no cost in terms of reliability.

Some covariability is evident in the HCI predictions of Figure 7. Predictions for *Pd* at sites 54 and 56 have very similar temporal patterns, which is indicative of spatial dependence. There is also some similarity between the predictions of *Pd* and *Qd*. Near-zero *Qd* predictions are generally associated with low predictions for *Pd*, which indicates some intervariable dependence.

Note that for variable *Qd*, predictions can be made at the beginning of the period, despite the fact that no streamflow data was available before 1951 (at any site). These predictions correspond to the "missing period" situation discussed in Section 2.4 and are feasible because the HCIs have been estimated based on the other variables available at this time. This suggests a potentially interesting application of HCI modeling to perform probabilistic reanalyzes of multivariable data sets.

The rotated panels in Figure 7 represent the marginal mixture distribution (see Section 2.4). Both the HCI and SCI models are able to handle continuous and discrete distributions. The censoring mechanism also allows computing the probability for a continuous variable to reach a physical bound. As an illustration, the variable *Qd* is continuous with a discrete mass at zero, and the HCI and SCI models are able to estimate the probability of a zero duration (HCI: 0.55, SCI: 0.59, observed frequency: 0.56). Overall the HCI and SCI marginal distributions are remarkably similar, despite their time-varying distributions being very different.

4.4. Multivariate Predictions

Figure 9 shows bivariate distributions for a few combinations of the variables shown in Figure 7. Both the HCI and SCI models are able to handle complex dependence structures, including dependence between discrete and continuous distributions. This is made possible by the conditional independence assumption, with dependence being induced indirectly by the effect of common covariates (hidden or not).

The first two rows of Figure 9 indicate that the HCI model leads to a stronger dependence than the SCI model, which is a consequence of the HCI effects being larger, leading to more covariability. This is particularly evident for variable *Pd* at sites 54 and 56, where the SCI model leads to a nearly independent bivariate distribution which is not consistent with the observed scatterplot.

The ability to describe dependence can be evaluated in more depth by computing joint or conditional probabilities from the bivariate distributions. The bottom left panel of Figure 9 shows the probability for the dry duration *Pd* to be larger than some value at both sites 54 and 56. The HCI model achieves a good fit to the observed frequencies, while the SCI model systematically underestimates this probability, due to the underestimation of spatial dependence. The second panel shows the probability of seeing a hydrological drought ($Qd > 0$) given that the dry duration *Pd* is larger than some value. The HCI model correctly recognizes that longer dry periods lead to a larger probability of hydrological drought. By contrast, the SCI model suggests a nearly constant conditional probability. A similar finding is obtained for the number of heat waves *Tn*: when the dry duration increases, the

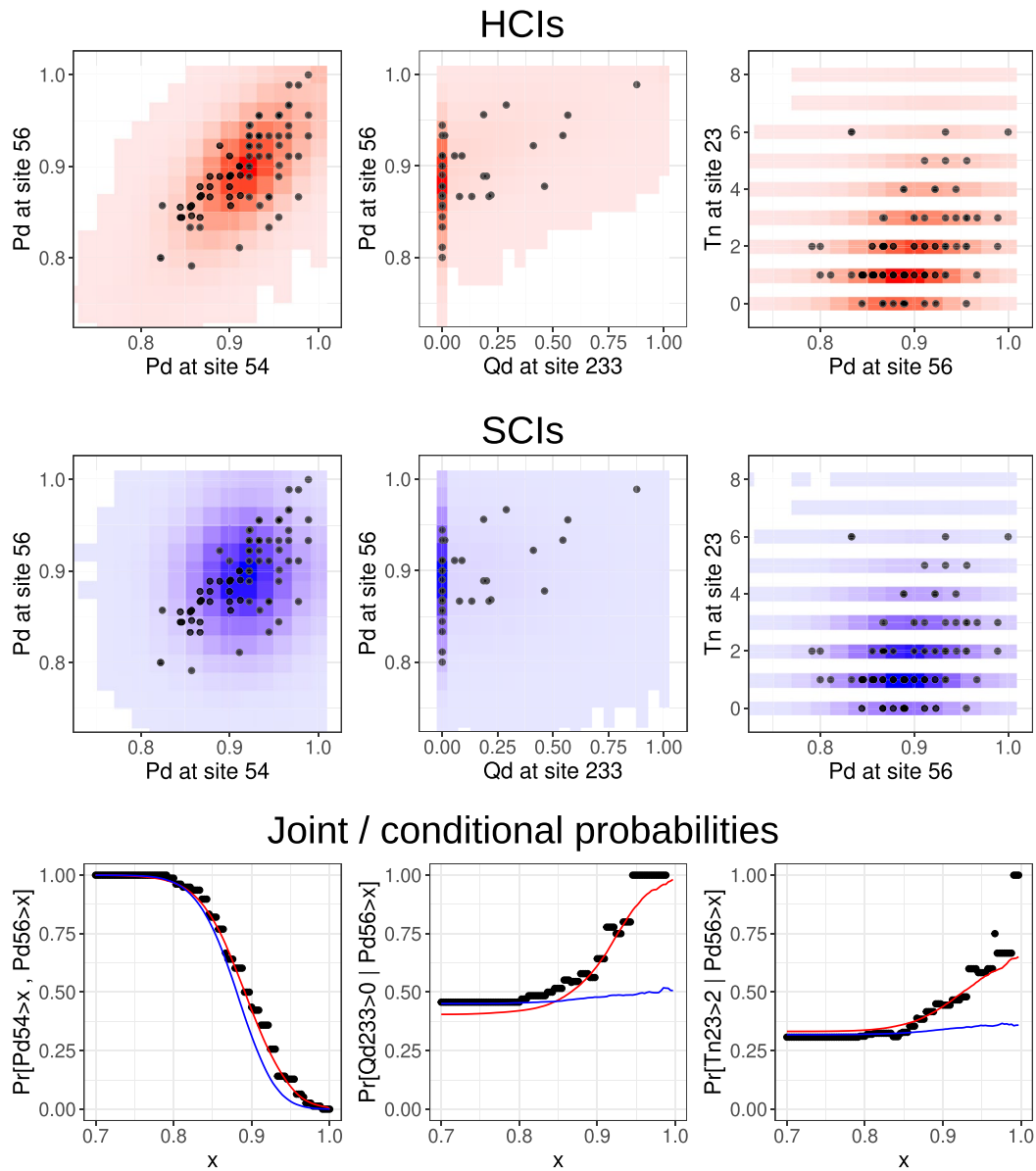

Figure 9. Top two rows: bivariate distributions derived from the hidden climate indices (HCI) and the standard climate indices (SCI) models. The colored area represents the bivariate pdf (estimated by counting Monte Carlo simulations over a two-dimensional grid), the points are the observed values. Bottom row: examples of joint and conditional probabilities extracted from the joint distributions (red = HCI, blue = SCI), and comparison with observed frequencies (points).

probability of seeing more than two heat waves increases according to the data and the HCI model, while the SCI model suggests a near-constant probability, thus underestimating the dependence between these two variables.

The ability to describe dependence can be evaluated beyond the bivariate case by considering events involving all variables at all sites. For instance, consider the question of estimating the fraction of sites where conditions are drier (variables *Qd* and *Pd*) or hotter (*Tx* and *Tn*) than some predefined threshold (e.g., a local p-quantile). Figure 10 shows that this fraction varies around $1 - p$ for both the HCI and SCI models. However, the HCI model leads to a much larger temporal variability of this fraction, in agreement with the observed frequencies. As an illustration, the probability of seeing >75% of the sites with above-median conditions is equal to 0.07 according to the HCI model, but is equal to 0.0 according to the SCI model (observed frequency: 0.05). The better reliability of the HCI model also holds for more extreme thresholds. For instance, only the HCI model is able to recognize that half of the stations may be simultaneously affected by a 10-year event (bottom panel). Reliably estimating the

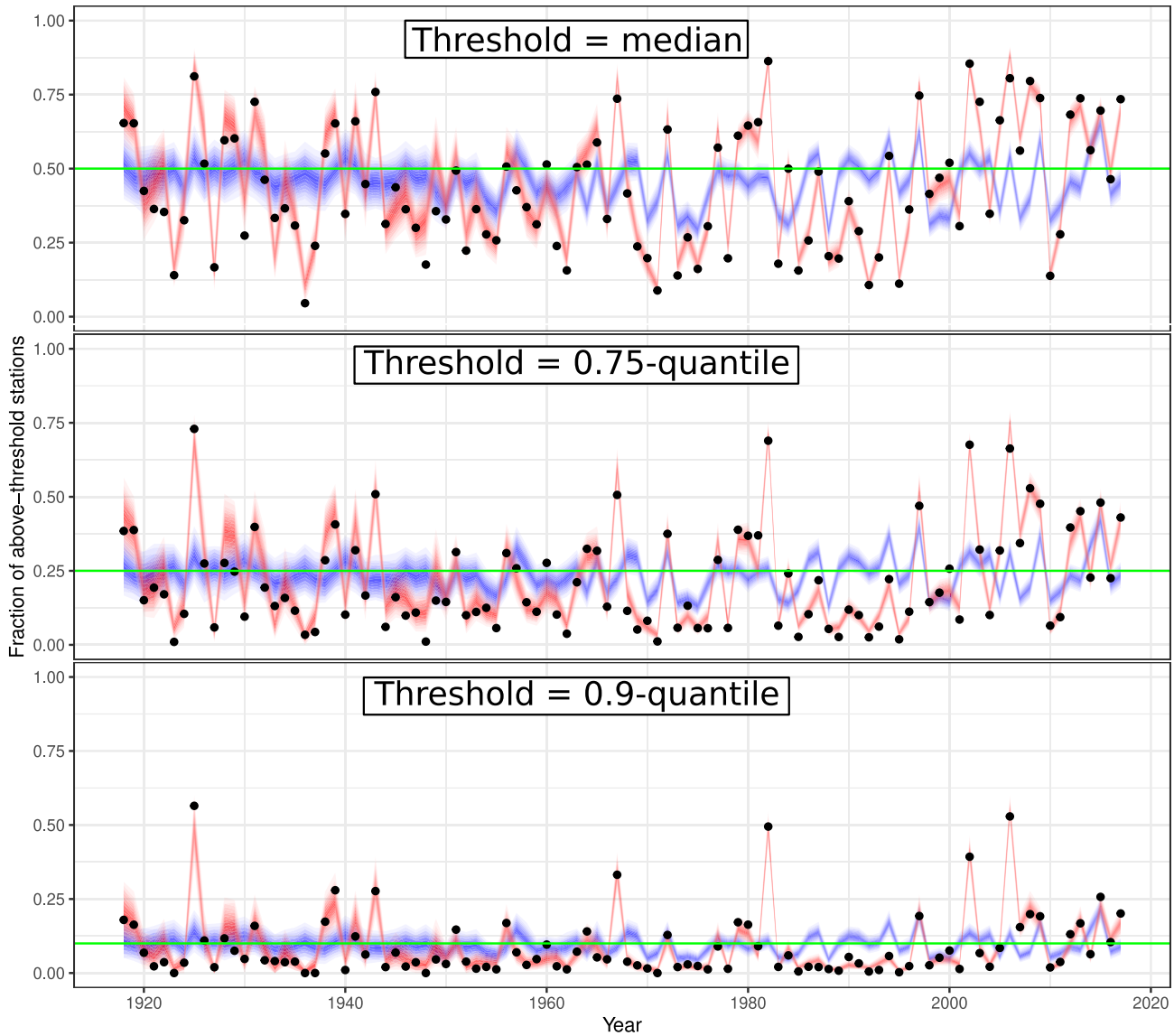

Figure 10. Modeled versus observed fraction of above-threshold stations. Colored bands represent nested prediction intervals up to the 95% level (red = hidden climate indices (HCI), blue = standard climate indices (SCI)), black dots are observed values.

large temporal variability of such hazard measures is important for practical purposes. Indeed, hazard response should be designed to cope with an alternation of severe and relatively problem-free years, rather than with the repetition of a hypothetical "average" year.

4.5. Split-Sample Experiment

Results of the split-sample experiment are presented by comparing the predictions obtained with: (a) the reduced calibration data set described in Section 3.3, leaving out 1/3 of the sites for all variables and all values from the period 1971–1990 for variable *Pd*; (b) the full calibration data set.

Figure 11 focuses on predictions for 1971–1990 *Pd* values, which are included in the full calibration data set but not in the reduced one. It suggests that the ability of the model to predict unobserved decades for one variable is satisfactory. Indeed, the reduced calibration data set does not lead to any marked loss of performance compared with the full calibration data set. In particular, Figure 11a shows a slight increase of the scatter, mostly visible for

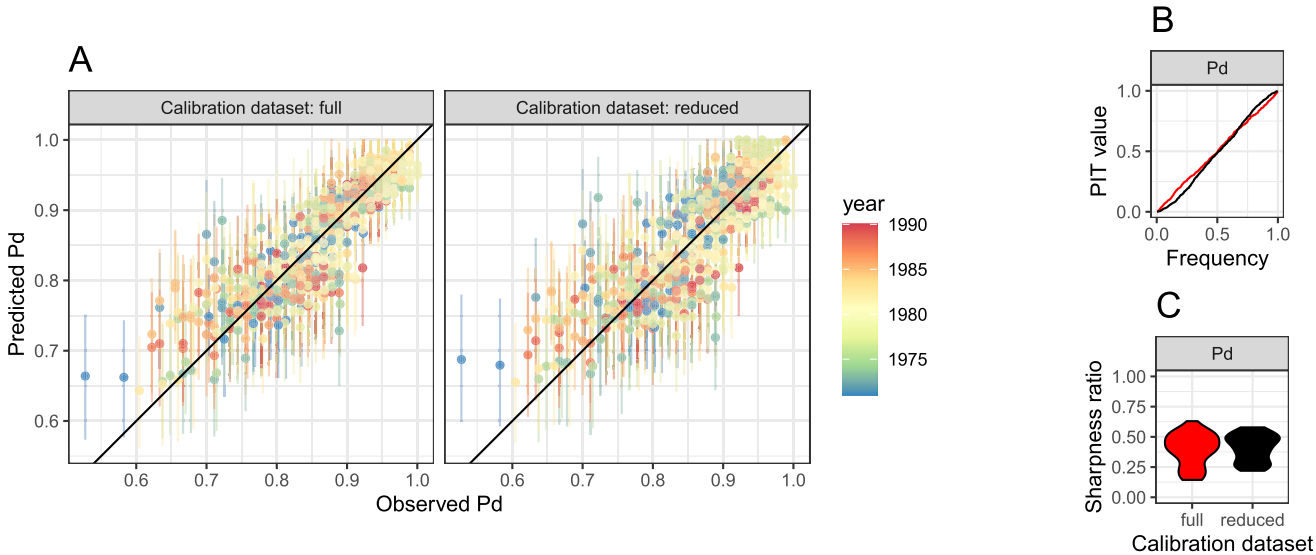

Figure 11. Split-sample experiment: ability to predict variable *Pd* during the unobserved decades 1971–1990. (a) Observed versus predicted values; (b) PIT diagrams (red = full calibration data set, black = reduced calibration data set); (c) distribution of sharpness ratios.

high *Pd* values, and no systematic underestimation or overestimation for specific years. Figures 11b and 11c also indicate that the reliability and sharpness of predictions remain similar.

Figure 12 focuses on predictions at left-out sites. Figure 12a shows results for variable *Pd* only. As previously, the scatter is slightly larger when prediction sites are ungauged (reduced calibration data set), but the performance remains satisfactory. Results for other variables are summarized in Figures 12b and 12c. While reliability is very similar for both calibration data sets, sharpness is slightly decreased for all variables when prediction sites are ungauged. This is due to the propagation of the Kriging variance (Equation 19), which leads to more uncertain (but still reliable) predictions. A similar figure was made for the stations located in Tasmania to verify that the stationary covariance function used in this case study was reasonable in an island configuration. Figure S5 in Supporting Information S1 suggests that reliability indeed remains acceptable.

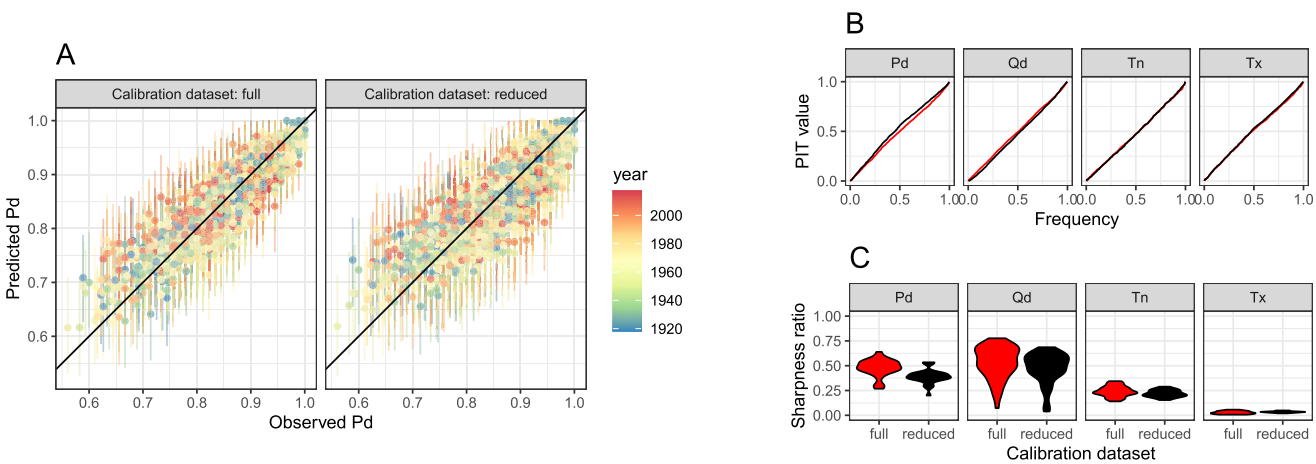

Figure 12. Split-sample experiment: ability to predict variables at ungauged sites. (a) Observed versus predicted values for variable *Pd*; (b) PIT diagrams for all variables (red = full calibration data set, black = reduced calibration data set); (c) distribution of sharpness ratios for all variables.

5. Discussion

5.1. Comparison of HCI and SCI Models

The basic idea behind HCI models is to extract the relevant temporal covariates directly from the target variables, rather than to rely on predefined SCIs. The case study confirms that HCIs can be precisely estimated from the target variables, thus identifying the strong temporal structures in the data that cannot be easily described by SCIs. Moreover, the case study shows that HCIs have much stronger effects on the target variables than SCIs. This finding is not surprising, since HCIs are estimated and are hence "tuned" to the target variables. Stronger HCI effects induce larger temporal variability, leading to sharper time-varying predictions. In addition, they induce larger covariability between sites and/or variables, leading to a more reliable description of spatial and/or intervariable dependencies.

However, there is a price to pay for this additional explanatory power: HCIs are highly specific to the target predictand data set. In a sense, they can be considered as "disposable" temporary variables that are used to reproduce the variabilities and dependencies existing in the target data set. However, they should not be reused for other predictand data sets—they would not be hidden anymore! By contrast, SCIs are extracted from atmospheric and oceanic circulation variables, which are the main drivers of surface variables. They have hence more potential to be useful for many variables in many places of the world. In addition, the mechanisms behind at least some SCIs have been thoroughly studied and are well understood (e.g., the role of atmospheric/oceanic dynamics and thermocline variability for ENSO). This gives SCIs some legitimacy as distinct climatic phenomena, compared with HCIs which are purely data-driven devices. Another practical limitation of HCIs is that their identification relies on the availability of many time series (i.e., many sites and/or variables). HCI models are therefore not adapted to studies involving a small number of time series, whereas SCI models can be applied to even a single time series.

5.2. On the Indirect Treatment of Dependence in HCI Modeling

The HCI model treats both spatial and intervariable dependence in an indirect way: dependence is induced by the strong and spatially structured effects of common HCIs, but all data are assumed conditionally independent. This assumption has strong practical advantages. First, it easily accommodates "nonrectangular" data sets, where data availability strongly varies in time. This is an important strength for station-based hydroclimatic data sets, since it simplifies the use of all available data, including old data which are often discarded because spatial coverage is deemed insufficient. By contrast, approaches modeling data dependence more directly (e.g., geostatistical models) require implementing a specific treatment of missing data (e.g., imputation, case-deletion, data augmentation, etc.). In addition, HCI modeling also accommodates data sets containing both continuous and discrete variables with no extra effort, whereas such data sets require a specialized and nontrivial treatment with geostatistical approaches (Emery & Silva, 2009; Faugeras, 2017; Genest & Nešlehová, 2007; Leblois & Creutin, 2013).

The ability of HCI models to accommodate data sets with strongly varying data availability allows using them as reanalysis tools. Indeed, when the HCI model has been calibrated with data from S sites over T time steps, probabilistic predictions can be made for all $S \times T$ combinations, including combinations for which data were missing. For instance, streamflow data Qd are not available during the "partially observed" period pre-1951 in the case study. However, Qd predictions can still be made for this period because the HCIs have been estimated from other available variables.

As noted in Section 1, the HCI framework shares similarities with principal component analysis (PCA). In both cases, dependence in the data set is summarized by means of unobserved factors (here, HCIs). When only a small number of factors is used, the representation of dependence is necessarily approximate. In the case study of Section 3, three HCIs were sufficient to provide a reasonable description of the overall dependence structure; however, small-scale dependence might be underestimated unless many HCIs are used. A geostatistical approach might be more adapted if small-scale dependence is deemed important for the analysis at hand, since it naturally induces strong dependencies at short distances. Note that, in principle, both approaches could be combined; however, giving up the conditional independence assumption would cancel the practical advantages described in the previous paragraphs.

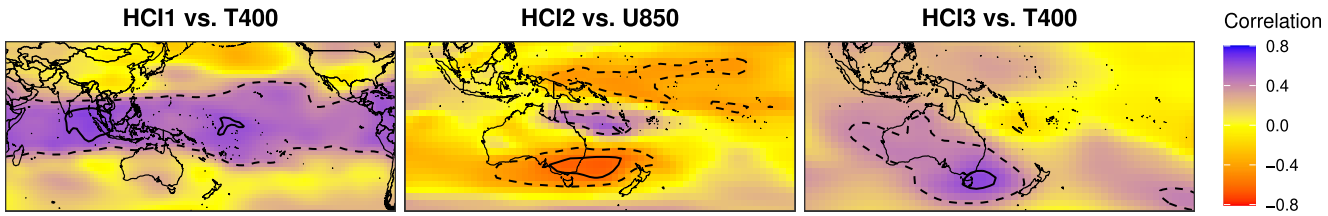

Figure 13. Examples of correlation maps between hidden climate indices (HCIs) and DJF-averaged atmospheric variables from the NCEP-NCAR reanalysis (Kalnay et al., 1996). Contour plots denote correlation values of 0.4 (dashed) and 0.6 (solid). T400 = temperature at 400 hPa, U850 = zonal wind component at 850 hPa.

5.3. Model Improvements

Some assumptions made in this paper may be relaxed without fundamentally changing the general HCI framework. Consider for instance the linear space-time decomposition used in Equation 6 and the Gaussian hyperdistributions for the HCIs (Equation 9) and their effects (Equation 7). These default assumptions are reasonable but could be modified if necessary. For instance, implementing the max-stable model of Reich and Shaby (2012) is feasible but requires both a nonlinear space-time decomposition and a non-Gaussian "positive-stable" hyperdistribution for the HCI. Nonstationary spatial covariance functions could also be needed in configurations including islands, large elevation range or orographic barriers, etc. Likewise, *iid* temporal processes were deemed sufficient for the purpose of this paper but other applications may require introducing trends, seasonality, and autocorrelation structures.

The covariance matrix in the spatial hyperdistribution (Equation 7) requires $\mathcal{O}(n^3)$ operations, which becomes intractable beyond a few hundreds of sites. A similar bottleneck may exist for the temporal covariance matrix when analyzing either very long or fine-time-step data sets. This computational bottleneck could be overcome by using adapted models such as the nearest neighbor Gaussian process (Banerjee, 2017; Datta et al., 2016a, 2016b). Note, however, that using a high number of variables does not induce a similar issue thanks to the conditional independence assumption prevailing between variables. This is to be contrasted with the treatment of multivariable data sets in geostatistics, which require specifying cross-covariance functions. Further work is needed to assess whether this theoretical advantage really holds in practice.

Finally, estimation of the HCI model could benefit from improvements in two directions. First, the K HCIs could be estimated jointly rather than through a stepwise procedure. To achieve this, another identifiability constraint is needed, namely that HCIs are orthogonal (in addition to having mean zero and unit standard deviation, see e.g., Murphy, 2012, chapter 12). Unfortunately, this constraint is not straightforward to implement, but recent work by Pourzanjani et al. (2020) may offer a solution. Second, HCI models could be implemented in general-purpose Bayesian frameworks such as STAN (Carpenter et al., 2017). While the handling of identifiability constraints makes this task not straightforward, such implementations would allow benefiting from built-in STAN functionalities such as alternative MCMC samplers, convergence diagnostics, approximate variational inference tools, etc.

5.4. Predicting HCIs From Large-Scale Climate Information and Potential Applications

Previous studies suggested that while being extracted from the target station data, HCIs may also be associated with specific patterns in large-scale atmospheric or oceanic variables (Ahn et al., 2017; Renard & Lall, 2014; Renard & Thyer, 2019). In order to investigate such association for the Australian case study of this paper, Figure 13 shows correlation maps between HCIs and selected atmospheric variables. At each pixel, the color denotes the correlation between two time series: (a) the atmospheric variable at this pixel and (b) the HCI (Figure 5). The first HCI is correlated with temperatures in the tropics, while the other two HCIs are more locally correlated with wind and temperature.

These correlation maps are indicative of a relation between HCIs and large-scale climate that could be exploited to predict the former from the latter. If sufficient predictability exists, the model of Equation 20 could then be used in a second stage to make predictions for all variables at individual stations. This would result in a downscaling tool between large-scale climate information and station data which would open the way for several interesting applications. For instance, a seasonal forecasting tool (e.g., Lima & Lall, 2010a; Lima et al., 2015) could be derived by downscaling the outputs of an atmospheric-oceanic forecasting system (e.g., Johnson et al., 2019).

Alternatively, past reconstructions (e.g., Caillouet et al., 2016, 2017) or future projections (e.g., Schlef et al., 2018; Tramblay et al., 2012) could be obtained from long reanalyzes such as 20CR (Compo et al., 2011) or from GCM outputs (e.g., Eyring et al., 2016).

Note that predicting HCIs from large-scale climate variables is not straightforward because the number of potential predictors is very large (whole climate fields for many atmospheric or oceanic variables) and is much larger than the number of predictands (a few HCIs). Candidate methods include regularized regression approaches such as ridge, lasso, or elastic net (Hastie et al., 2015), inverse regression (Devijver & Perthame, 2020), and possibly many other machine learning algorithms (Bishop, 2006; Giannakis & Majda, 2012). Developing such an additional predictive layer constitutes a key future objective.

6. Conclusion

A general probabilistic framework was introduced to model the time-varying joint distribution of multivariable space-time data. Instead of using predefined standard climate indices to model time variability as usually done in similar frameworks, the proposed approach extracts a set of hidden climate indices from the observed variables. These hidden climate indices correspond to temporal latent variables in a Bayesian hierarchical model. The framework can be applied to both discrete and continuous variables, and any distribution can be used. Data are assumed to be conditionally independent, which has strong practical advantages in terms of accommodating irregular station-based data sets. Intersite and intervariable dependencies are not modeled directly, but are both induced by the strong effect of common HCIs.

The proposed framework is applied to a case study in Southeast Australia aimed at jointly modeling hydrological droughts, meteorological droughts, and heat waves at many stations. The case study first illustrates that the HCI framework can easily be applied to a typical station-based data set that includes several noncolocated variables with very different spatial and temporal coverage. Moreover, the HCI model delivers reliable time-varying distributions that are much sharper than the ones resulting from an equivalent SCI model. In addition, the HCI model is able to reliably describe both spatial and intervariable dependencies, while an equivalent SCI model clearly underestimates them. Such dependencies include complex dependence structures between continuous and discrete variables.

This paper illustrated that the HCI framework is flexible and generic, and can be applied to a wide range of different hydrometeorological variables. Future work will investigate methods to predict HCIs from large-scale climate data. This would allow using the HCI framework as a downscaling tool to estimate the joint distribution of several variables at many stations from climate models or reanalyzes.

Data Availability Statement

The data and the scripts used for setting up models, analyzing results, and preparing figures are available in a Zenodo repository (https://doi.org/10.5281/zenodo.5721783). Original and updated versions of the data can also be downloaded using the url's given in Table 1. MCMC simulations have been performed with the computing code STooDs v0.1.0 (Renard, 2021b, https://github.com/STooDs-tools/STooDs) and its R interface RSTooDs v0.1.1 (Renard, 2021a, https://github.com/STooDs-tools/RSTooDs). Data and hidden climate indices are illustrated in a sonified animation at https://vimeo.com/600898709

Acknowledgments
We would like to thank the Editor, the Associate Editor, and the Reviewers for their insightful and constructive comments. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie Grant Agreement No. 835496.

References

Abram, N. J., Henley, B. J., Sen Gupta, A., Lippmann, T. J. R., Clarke, H., Dowdy, A. J., et al. (2021). Connections of climate change and variability to large and extreme forest fires in Southeast Australia. *Communications Earth & Environment*, *2*(1), 8. https://doi.org/10.1038/s43247-020-00065-8

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*, 241–253. https://doi.org/10.1175/2008JHM1007.1

Bach, F. R., & Jordan, M. I. (2005). *A probabilistic interpretation of canonical correlation analysis* (No. Technical Report 688). Retrieved from https://www.stat.berkeley.edu/~jordan/688.pdf

Banerjee, S. (2017). High-dimensional Bayesian geostatistics. *Bayesian Analysis*, *12*(2), 583–614. https://doi.org/10.1214/17-BA1056R

Barbero, R., Abatzoglou, J. T., Steel, E. A., & Larkin, N. K. (2014). Modeling very large-fire occurrences over the continental United States from weather and climate forcing. *Environmental Research Letters*, *9*, 124009. https://doi.org/10.1088/1748-9326/9/12/124009

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer. Retrieved from https://www.microsoft.com/en-us/research/uploads/prod/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf

Botto, A., Ganora, D., Claps, P., & Laio, F. (2017). Technical note: Design flood under hydrological uncertainty. *Hydrology and Earth System Sciences*, *21*, 3353–3358. https://doi.org/10.5194/hess-21-3353-2017

Bracken, C., Holman, K. D., Rajagopalan, B., & Moradkhani, H. (2018). A Bayesian hierarchical approach to multivariate nonstationary hydrologic frequency analysis. *Water Resources Research*, *54*, 243–255. https://doi.org/10.1002/2017WR020403

Bracken, C., Rajagopalan, B., Cheng, L., Kleiber, W., & Gangopadhyay, S. (2016). Spatial Bayesian hierarchical modeling of precipitation extremes over a large domain. *Water Resources Research*, *52*, 6643–6655. https://doi.org/10.1002/2016WR018768

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

Bureau of Meteorology. (2020a). *About Australian climate change site networks*. Retrieved from http://www.bom.gov.au/climate/change/hqsites/about-hq-site-data.shtml

Bureau of Meteorology. (2020b). *Hydrologic reference stations—Update 2020*. Retrieved from http://www.bom.gov.au/water/hrs/update_2020.shtml

Caillouet, L., Vidal, J. P., Sauquet, E., Devers, A., & Graff, B. (2017). Ensemble reconstruction of spatio-temporal extreme low-flow events in France since 1871. *Hydrology and Earth System Sciences*, *21*, 2923–2951. https://doi.org/10.5194/hess-21-2923-2017

Caillouet, L., Vidal, J. P., Sauquet, E., & Graff, B. (2016). Probabilistic precipitation and temperature downscaling of the twentieth century reanalysis over France. *Climate of the Past*, *12*, 635–662. https://doi.org/10.5194/cp-12-635-2016

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., et al. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*, 1–32. https://doi.org/10.18637/jss.v076.i01

Chib, S. (1992). Bayes inference in the Tobit censored regression model. *Journal of Econometrics*, *51*, 79–99. https://doi.org/10.1016/0304-4076(92)90030-U

Commonwealth of Australia. (2019). In J. Ball, M. Babister, R. Nathan, W. Weeks, E. Weinmann, M. Retallick, & I. Testoni (Eds.), *Australian rainfall and runoff: A guide to flood estimation*. Retrieved from http://arr.ga.gov.au/

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*, 1–28. https://doi.org/10.1002/qj.776

Congdon, P. D. (2010). *Applied Bayesian hierarchical methods*. Chapman and Hall/CRC.

Cooley, D., Nychka, D., & Naveau, P. (2007). Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, *102*(479), 824–840. https://doi.org/10.1198/016214506000000780

Datta, A., Banerjee, S., Finley, A. O., & Gelfand, A. E. (2016a). Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets. *Journal of the American Statistical Association*, *111*, 800–812. https://doi.org/10.1080/01621459.2015.1044091

Datta, A., Banerjee, S., Finley, A. O., & Gelfand, A. E. (2016b). On nearest-neighbor Gaussian process models for massive spatial data. *Wiley Interdisciplinary Reviews: Computational Statistics*, *8*, 162–171. https://doi.org/10.1002/wics.1383

De Haan, L., & De Ronde, J. (1998). Sea and wind: Multivariate extremes at work. *Extremes*, *1*, 7–45. https://doi.org/10.1023/a:1009909800311

De Iaco, S. (2011). A new space-time multivariate approach for environmental data analysis. *Journal of Applied Statistics*, *38*, 2471–2483. https://doi.org/10.1080/02664763.2011.559206

Devijver, E., & Perthame, E. (2020). Prediction regions through inverse regression. *Journal of Machine Learning Research*, *21*(113), 1–24. Retrieved from http://jmlr.org/papers/v21/19-535.html

Devineni, N., Lall, U., Pederson, N., & Cook, E. (2013). A tree-ring-based reconstruction of Delaware river basin streamflow using hierarchical Bayesian regression. *Journal of Climate*, *26*, 4357–4374. https://doi.org/10.1175/JCLI-D-11-00675.1

Dyrrdal, A. V., Lenkoski, A., Thorarinsdottir, T. L., & Stordal, F. (2015). Bayesian hierarchical modeling of extreme hourly precipitation in Norway. *Environmetrics*, *26*, 89–106. https://doi.org/10.1002/env.2301

Emery, X., & Silva, D. A. (2009). Conditional co-simulation of continuous and categorical variables for geostatistical applications. *Computers & Geosciences*, *35*, 1234–1246. https://doi.org/10.1016/j.cageo.2008.07.005

Engeland, K., Borga, M., Creutin, J.-D., François, B., Ramos, M.-H., & Vidal, J.-P. (2017). Space-time variability of climate variables and intermittent renewable electricity production—A review. *Renewable and Sustainable Energy Reviews*, *79*, 600–617. https://doi.org/10.1016/j.rser.2017.05.046

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*, 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Faugeras, O. P. (2017). Inference for copula modeling of discrete data: A cautionary tale and some facts. *Dependence Modeling*, *5*, 121–132. https://doi.org/10.1515/demo-2017-0008

Favre, A. C., El Adlouni, S., Perreault, L., Thiemonge, N., & Bobee, B. (2004). Multivariate hydrological frequency analysis using copulas. *Water Resources Research*, *40*, W01101. https://doi.org/10.1029/2003WR002456

François, B., Borga, M., Anquetin, S., Creutin, J. D., Engeland, K., Favre, A. C., et al. (2014). Integrating hydropower and intermittent climate-related renewable energies: A call for hydrology. *Hydrological Processes*, *28*, 5465–5468. https://doi.org/10.1002/hyp.10274

Garavaglia, F., Gailhard, J., Paquet, E., Lang, M., Garcon, R., & Bernardara, P. (2010). Introducing a rainfall compound distribution model based on weather patterns sub-sampling. *Hydrology and Earth System Sciences*, *14*, 951–964. https://doi.org/10.5194/hess-14-951-2010

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472. https://doi.org/10.1214/ss/1177011136

Genest, C., & Nešlehová, J. (2007). A primer on copulas for count data. *ASTIN Bulletin*, *37*, 475–515. https://doi.org/10.1017/S0515036100014963

Ghosh, S., & Mallick, B. K. (2011). A hierarchical Bayesian spatio-temporal model for extreme precipitation events. *Environmetrics*, *22*, 192–204. https://doi.org/10.1002/env.1043

Giannakis, D., & Majda, A. J. (2012). Nonlinear Laplacian spectral analysis for time series with intermittency and low-frequency variability. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 2222–2227. https://doi.org/10.1073/pnas.1118984109

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

Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical learning with sparsity: The lasso and generalizations*. CRC Press, Taylor & Francis Group.

Heffernan, J. E., & Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society*, *66*, 497–546. https://doi.org/10.1111/j.1467-9868.2004.02050.x

Ho, M., Kiem, A. S., & Verdon-Kidd, D. C. (2012). The southern annular mode: A comparison of indices. *Hydrology and Earth System Sciences*, *16*, 967–982. https://doi.org/10.5194/hess-16-967-2012

Ho, M., Kiem, A. S., & Verdon-Kidd, D. C. (2015). A paleoclimate rainfall reconstruction in the Murray-Darling basin (MDB), Australia: 1. Evaluation of different paleoclimate archives, rainfall networks, and reconstruction techniques. *Water Resources Research*, *51*, 8362–8379. https://doi.org/10.1002/2015WR017058

Hurrell, J. W., & Van Loon, H. (1997). Decadal variations in climate associated with the North Atlantic oscillation. *Climatic Change*, *36*, 301–326. https://doi.org/10.1007/978-94-015-8905-5_4

Interagency Advisory Committee on Water Data. (1982). *Guidelines for determining flood-flow frequency: Bulletin 17b of the hydrology subcommittee* (Office of Water data Coordination, Ed.). U.S. Geological Survey.

Johnson, S. J., Stockdale, T. N., Ferranti, L., Balmaseda, M. A., Molteni, F., Magnusson, L., et al. (2019). SEAS5: The new ECMWF seasonal forecast system. *Geoscientific Model Development*, *12*, 1087–1117. https://doi.org/10.5194/gmd-12-1087-2019

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), 437–472. https://doi.org/10.1175/1520-0477(1996)077<0437:tnyrp>2.0.co;2

Kiem, A. S., Johnson, F., Westra, S., van Dijk, A., Evans, J. P., O'Donnell, A., et al. (2016). Natural hazards in Australia: Droughts. *Climatic Change*, *139*, 37–54. https://doi.org/10.1007/s10584-016-1798-7

Klami, A., Virtanen, S., & Kaski, S. (2013). Bayesian canonical correlation analysis. *Journal of Machine Learning Research*, *14*, 965–1003.

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

Le, P. D., Leonard, M., & Westra, S. (2018). Modeling spatial dependence of rainfall extremes across multiple durations. *Water Resources Research*, *54*, 2233–2248. https://doi.org/10.1002/2017WR022231

Leblois, E., & Creutin, J.-D. (2013). Space-time simulation of intermittent rainfall with prescribed advection field: Adaptation of the turning band method. *Water Resources Research*, *49*, 3375–3387. https://doi.org/10.1002/wrcr.20190

Lima, C. H. R., & Lall, U. (2010a). Climate informed long term seasonal forecasts of hydroenergy inflow for the Brazilian hydropower system. *Journal of Hydrology*, *381*, 65–75. https://doi.org/10.1016/j.jhydrol.2009.11.026

Lima, C. H. R., & Lall, U. (2010b). Spatial scaling in a changing climate: A hierarchical Bayesian model for non-stationary multi-site annual maximum and monthly streamflow. *Journal of Hydrology*, *383*, 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

Longford, N. T. (1993). *Random coefficient models*. Oxford University Press.

Maraun, D., Rust, H. W., & Osborn, T. J. (2010). Synoptic airflow and UK daily precipitation extremes development and validation of a vector generalised linear model. *Extremes*, *13*, 133–153. https://doi.org/10.1007/s10687-010-0102-x

McCullagh, P., & Nelder, J. A. (1989). *Generalized linear models* (2nd ed.). Chapman and Hall/CRC.

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. MIT Press.

Nairn, J. R., Fawcett, R. G., & Day, K. A. (2013). *Defining heatwaves: Heatwave defined as a heat-impact event servicing all community and business sectors in Australia*. Centre for Australian Weather and Climate Research.

NCAR. (2019). *The climate data guide: Overview: Climate indices*. Retrieved from https://climatedataguide.ucar.edu/climate-data/overview-climate-indices

NOAA. (2020a). *20CR climate indices: Southern annular mode (SAM)*. Retrieved from https://psl.noaa.gov/data/20thC_Rean/timeseries/monthly/SAM/

NOAA. (2020b). *Daily Antarctic oscillation index*. Retrieved from https://www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/aao/aao_index.html

NOAA. (2020c). *Dipole mode index (DMI)*. Retrieved from https://psl.noaa.gov/gcos_wgsp/Timeseries/DMI/

NOAA. (2020d). *Niño 4 SST index*. Retrieved from https://psl.noaa.gov/gcos_wgsp/Timeseries/Nino4/

Nogaj, M., Yiou, S., Parey, S., Malek, F., & Naveau, P. (2006). Amplitude and frequency of temperature extremes over the North Atlantic region. *Geophysical Research Letters*, *33*, L10801. https://doi.org/10.1029/2005GL024251

Ossandon, A., Rajagopalan, B., & Kleiber, W. (2021). Spatial-temporal multivariate semi-Bayesian hierarchical framework for extreme precipitation frequency analysis. *Journal of Hydrology*, *600*, 126499. https://doi.org/10.1016/j.jhydrol.2021.126499

Padoan, S. A., Ribatet, M., & Sisson, S. A. (2010). Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association*, *105*, 263–277. https://doi.org/10.1198/jasa.2009.tm08577

Payrastre, O., Gaume, E., & Andrieu, H. (2011). Usefulness of historical information for flood frequency analyses: Developments based on a case study. *Water Resources Research*, *47*, W08511. https://doi.org/10.1029/2010WR009812

Perkins-Kirkpatrick, S. E., White, C. J., Alexander, L. V., Argüeso, D., Boschat, G., Cowan, T., et al. (2016). Natural hazards in Australia: Heatwaves. *Climatic Change*, *139*, 101–114. https://doi.org/10.1007/s10584-016-1650-0

Perreault, L., Bernier, J., Bobee, B., & Parent, E. (2000a). Bayesian change-point analysis in hydrometeorological time series. Part 1. The normal model revisited. *Journal of Hydrology*, *235*, 221–241. https://doi.org/10.1016/S0022-1694(00)00270-5

Perreault, L., Bernier, J., Bobee, B., & Parent, E. (2000b). Bayesian change-point analysis in hydrometeorological time series. Part 2. Comparison of change-point models and forecasting. *Journal of Hydrology*, *235*, 242–263. https://doi.org/10.1016/S0022-1694(00)00271-7

Pourzanjani, A. A., Jiang, R. M., Mitchell, B., Atzberger, P. J., & Petzold, L. R. (2020). Bayesian inference over the Stiefel Manifold via the givens representation. *Bayesian Analysis*, *16*, 639–666. https://doi.org/10.1214/20-BA1202

Prosdocimi, I., Kjeldsen, T. R., & Miller, J. D. (2015). Detection and attribution of urbanization effect on flood extremes using nonstationary flood-frequency models. *Water Resources Research*, *51*, 4244–4262. https://doi.org/10.1002/2015WR017065

Ramachandra Rao, A., & Hamed, K. H. (2019). In K. Hamed, & A. R. Rao (Eds.), *Flood frequency analysis* (1st ed.). CRC Press. https://doi.org/10.1201/9780429128813

Reich, B. J., & Shaby, B. A. (2012). A hierarchical max-stable spatial model for extreme precipitation. *Annals of Applied Statistics*, *6*, 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. (2021a). *STooDs-tools/RSTooDs: RSTooDs package v0.1.1*. https://doi.org/10.5281/zenodo.5075760
Renard, B. (2021b). *STooDs-tools/STooDs: STooDs engine v0.1.0*. https://doi.org/10.5281/ZENODO.5075586
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
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
Ribatet, M., Cooley, D., & Davison, A. C. (2012). Bayesian inference from composite likelihoods, with an application to spatial extremes. *Statistica Sinica*, *22*, 813–845. https://doi.org/10.5705/ss.2009.248
Ropelewski, C. F., & Jones, P. D. (1987). An extension of the Tahiti-Darwin southern oscillation index. *Monthly Weather Review*, *115*, 2161–2165. https://doi.org/10.1175/1520-0493(1987)115<2161:aeotts>2.0.co;2
Salvadori, G., & De Michele, C. (2004). Frequency analysis via copulas: Theoretical aspects and applications to hydrological events. *Water Resources Research*, *40*, W12511. https://doi.org/10.1029/2004WR003133
Sang, H., & Gelfand, A. E. (2010). Continuous spatial process models for spatial extreme values. *Journal of Agricultural, Biological, and Environmental Statistics*, *15*, 49–65. https://doi.org/10.1007/s13253-009-0010-1
Sarhadi, A., Burn, D. H., Concepción Ausín, M., & Wiper, M. P. (2016). Time-varying nonstationary multivariate risk analysis using a dynamic Bayesian copula. *Water Resources Research*, *52*, 2327–2349. https://doi.org/10.1002/2015WR018525
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
Sharples, J. J., Cary, G. J., Fox-Hughes, P., Mooney, S., Evans, J. P., Fletcher, M.-S., et al. (2016). Natural hazards in Australia: Extreme bushfire. *Climatic Change*, *139*, 85–99. https://doi.org/10.1007/s10584-016-1811-1
Sideris, I. V., Gabella, M., Erdin, R., & Germann, U. (2014). Real-time radar-rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland: Real-time radar-rain-gauge merging. *Quarterly Journal of the Royal Meteorological Society*, *140*, 1097–1111. https://doi.org/10.1002/qj.2188
Spiegelhalter, D. J., Best, N. G., Carlin, B. R., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society-Series B: Statistical Methodology*, *64*, 583–639. https://doi.org/10.1111/1467-9868.00353
Stasinopoulos, D. M., & Rigby, R. A. (2007). Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software*, *23*(7), 1–46. https://doi.org/10.18637/jss.v023.i07
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
Steirou, E., Gerlitz, L., Apel, H., Sun, X., & Merz, B. (2019). Climate influences on flood probabilities across Europe. *Hydrology and Earth System Sciences*, *23*, 1305–1322. https://doi.org/10.5194/hess-23-1305-2019
Sun, X., Lall, U., Merz, B., & Dung, N. V. (2015a). 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. (2015b). 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
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*, 12–26. 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*, 27–48. https://doi.org/10.1016/s0022-1694(02)00411-0
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
Tramblay, Y., Badi, W., Driouech, F., El Adlouni, S., Neppel, L., & Servat, E. (2012). Climate change impacts on extreme precipitation in Morocco. *Global and Planetary Change*, *82–83*, 82–114. https://doi.org/10.1016/j.gloplacha.2011.12.002
Trewin, B. (2013). A daily homogenized temperature data set for Australia. *International Journal of Climatology*, *33*, 1510–1529. https://doi.org/10.1002/joc.3530
Trewin, B. (2018). *The Australian climate observations reference network-surface air temperature (ACORN-SAT) version 2*. Retrieved from http://www.bom.gov.au/research/publications/researchreports/BRR-032.pdf
Westra, S., Alexander, L. V., & Zwiers, F. W. (2012). Global increasing trends in annual maximum daily precipitation. *Journal of Climate*, *26*(11), 3904–3918. https://doi.org/10.1175/jcli-d-12-00502.1
Westra, S., & Sharma, A. (2009). Probabilistic estimation of multivariate streamflow using independent component analysis and climate information. *Journal of Hydrometeorology*, *10*, 1479–1492. https://doi.org/10.1175/2009jhm1121.1
Westra, S., & Sisson, S. A. (2011). Detection of non-stationarity in precipitation extremes using a max-stable process model. *Journal of Hydrology*, *406*, 119–128. https://doi.org/10.1016/j.jhydrol.2011.06.014
Westra, S., White, C. J., & Kiem, A. S. (2016). Introduction to the special issue: Historical and projected climatic changes to Australian natural hazards. *Climatic Change*, *139*, 1–19. https://doi.org/10.1007/s10584-016-1826-7
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 *Hydrology and Water Resources Symposium 2014 (HWRS 2014)—Conference Proceedings* (pp. 1–8).
Zscheischler, J., Westra, S., van den Hurk, B. J. J. M., Seneviratne, S. I., Ward, P. J., Pitman, A., et al. (2018). Future climate risk from compound events. *Nature Climate Change*, *8*(6), 469–477. https://doi.org/10.1038/s41558-018-0156-3