



A spatio-temporal weather generator for the temperature over France

Caroline Cognot^{1,2}, Liliane Bel², David Métivier³, and Sylvie Parey¹

¹EDF R&D, 7 Bd Gaspard Monge, 91120 Palaiseau, France

²MIA Paris-Saclay, AgroParisTech, INRAE, Université Paris-Saclay,
22 place de l'Agronomie, 91120 Palaiseau, France

³MISTEA, Université de Montpellier, INRAE, Institut Agro, Montpellier, France

Correspondence: Caroline Cognot (caroline.cognot@edf.fr)

Received: 9 December 2024 – Revised: 8 August 2025 – Accepted: 8 August 2025 – Published: 9 October 2025

Abstract. Stochastic weather generators are efficient statistical models producing synthetic weather series by replicating key statistical properties without the computational cost of physical models. However, for applications requiring temperature simulation over a large area, challenges arise due to non-stationarity over time and spatial-temporal dependencies. This paper introduces a daily stochastic weather generator for temperature with arbitrary spatial resolution. The non-stationarity issue is addressed using a decomposition method to separate deterministic terms (trends and seasonality), from the stochastic part representing the underlying climate variability. We extend the existing local decomposition method to extrapolate to any point in space. The spatial-temporal dependence is modeled through a Gaussian field with a non-separable covariance function, accommodating complex interactions between time and space. Our generator, calibrated on a few French weather stations, is validated using several spatio-temporal indicators. First, we evaluate the generator's performance at the fitting stations, comparing simulated and observed indicators. Subsequently, we compare our spatial simulations to a high resolution gridded observation dataset. Results demonstrate that the proposed generator accurately captures the observed spatio-temporal statistics, even for extreme events such as large scale persistent heat waves.

1 Introduction

Many fields are impacted by climate change, from agriculture and health to energy management. These impacts can be studied through risk evaluation, which often relies on probabilistic models that use meteorological variables as inputs. While historical records can provide such data, they are often limited in both space and time and may not represent future conditions accurately. Another limitation is that they typically offer only a single realization, which can lead to an underestimation of variability.

Physical climate models can generate additional data, but they are computationally expensive, making it difficult to perform many simulations. Moreover, their coarse spatial resolution makes them unsuitable for local studies. In this context, stochastic weather generators have emerged as powerful tools for simulating weather data that captures both spatial and temporal variability (see IPCC, 2023, Sect. 10.3.3.7).

These generators consider observations of meteorological variables as one of the infinite possible trajectories of a stochastic model. The primary objective is to simulate numerous plausible sequences of the desired variables that share common statistical properties with the observations. These simulations are fast and allow for a more comprehensive sampling of climate variables than physical models.

In the context of climate adaptation in the energy sector, it is crucial for decision-makers to understand the potential risks associated with events that affect large areas simultaneously or involve a set of specific sites distributed across the area, such as production sites. A desirable model must effectively capture the spatial dependence in temperature and be capable of generating simulations for sites where no data is available. Another key objective is to separate the contribution of the changing climate (long-term trend) from stationary effects. This distinction allows the model, once calibrated

to the present climate, to be adapted for future climate conditions.

Since the development of the first single-site weather generator by Richardson (1981), many others have been proposed using various approaches to produce spatially correlated series. Resampling methods do not require a model for the spatial relationship, instead, the values of all variables at all sites at time $t + 1$ are generated based on the nearest neighbors at time t (Rajagopalan and Lall, 1999; Buishand and Brandsma, 2001; Prairie et al., 2006). The resampling procedure can be conditioned on weather states (Apipattanasit et al., 2007) or circulation variables (Beersma and Buishand, 2003). More recent techniques use analogues (Yiou, 2014). Although these non-parametric methods have the advantage of being distribution-free, their limitation lies in their inability to generate new, unobserved values, which often leads to simulations that are overly similar to historical events and underestimate variability. Semi-parametric procedures mitigate this problem by fitting single-site margins, generate simulations, and reconstruct the relationships between all variables and sites using a shuffling method, related to empirical copulas (Clark et al., 2004; Li, 2014; Li and Babovic, 2019; Li et al., 2019; Wang et al., 2021). Recent works have also explored the use of neural networks (Ji et al., 2024; Sha et al., 2024).

When constructing a model for the spatial relationship, two different approaches can be separated. Multisite models focus on a small number of distributed sites and treat each location as a separate variable and rely on multivariate statistics to represent inter-site correlations, employing methods such as multivariate autoregressive models (Wilks, 2009; Dubrovsky et al., 2020) or copulas (Kroiz et al., 2021; Dawkins et al., 2022). These methods have two main limitations: they cannot generate data for new locations, and they often require estimating a large number of parameters, making them difficult to scale. While feasible for small numbers of sites, they do not efficiently utilise spatial information and become impractical for larger grids. The use of Empirical Orthogonal Functions (EOFs) as in Sparks et al. (2018) offers dimensionality reduction but does not resolve location dependence issues. Contrary to that, approaches making use of spatial statistics focus on dense, local grids. They use Gaussian process models, which treat dependence as a spatial or spatio-temporal field, allowing for data generation at unobserved locations once the model is fitted (Wilks, 2009; Kleiber et al., 2012; Baxevani and Lennartsson, 2015; Verdin et al., 2015; Bourotte et al., 2016; Sparks et al., 2018; Verdin et al., 2019). However, their application as of now has mostly been limited to specific regions, while broader applications require simulations across larger areas.

The temporal dependence for each location is usually modeled using the same autoregressive (AR) structure introduced in Richardson (1981). It captures both the temporal dependence and the distribution of the variables (Wilks, 2009; Dubrovsky et al., 2020). The AR model can be modified to

accommodate various distributions such as the skew exponential power distribution used by Evin et al. (2019), to better represent the tails of the distribution. The parameters can depend on time to represent the spatial-temporal interactions. In a spatial model, the AR approach can be replaced by a covariance model that quantifies variability between two points separated in space, for every time difference, and the space-time interactions.

Weather variables have strong non-stationarity. When available, covariates can be introduced in the model to tackle both spatial and temporal non-stationarity using Generalised Linear Models (GLMs), Generalised Additive Models (GAMs), or related methods (Furrer and Katz, 2007; Holsclaw et al., 2016; Chandler, 2020; Dawkins et al., 2022). Common covariates for this purpose include large-scale information such as sea surface temperature (SST) or climate indices. One drawback of these approaches is that they may be unable to generate events that differ significantly from observations, as they constrain generation to observed patterns. A proxy for this large-scale information is the use of weather types which are fitted directly on the weather data; see Allard et al. (2015) for a review, Gobet et al. (2025) for a recent application. The temporal non-stationarity include seasonal effects. They can be addressed by fitting a separate set of parameters for each month, as in Verdin et al. (2015), or by using smooth periodic functions (Evin et al., 2019). Global warming can be taken into account by using a linear model to represent the increase in mean temperature over the past century. This increase can include a breakpoint to capture recent acceleration (Touron, 2019; Evin et al., 2019).

In this work, we build from the preprocessing approach of Hoang (2010) that decomposes the temperature series for one location into deterministic and stochastic terms. We propose a spatial-temporal model for temperature by extending this concept to a spatial grid.

The chosen approach uses non-parametric LOESS (Locally Estimated Scatterplot Smoothing; Cleveland, 1979) regression to extract long-term trends from station data spread across France. Compared to simpler approaches based on linear trends, it better represents interannual variability caused by large-scale atmospheric patterns. It also captures the varying pace of warming over time more effectively. The seasonality of the temperature is obtained through parametric regression with trigonometric functions, in the mean and the variance. We propose an interpolation method to obtain spatial trends and seasonality from components computed separately for each location.

Residuals from the decomposition are assumed to be a centered Gaussian process for each location. Here, they are modeled as a spatio-temporal stationary Gaussian process. The interactions between space and time are modeled using a non-separable covariance function as in Allard et al. (2022). While previous applications of such a model have focused on smaller regions, we show that this approach allows to ac-

count for the strong spatial dependence in temperature across a large area.

We also propose indicators to validate the model, focusing on important properties of the temperature. We compare these indicators computed on observations to simulations generated with our model. These indicators include correlations in space and time. We show that our model accurately reproduces these correlations, showing a good capacity to reproduce the spatial-temporal relationships. We develop a particular focus on extremes, using first pairwise threshold exceedance. We also develop an indicator for spatial events called the quantile exceedance ratio, that is used to represent the spatial size of heat events, and show that the simulations from our model slightly underestimate the real size of spatial events, producing heat episodes that are quite as large as observed events.

The paper is organised as follows. In Sect. 2, we describe the data used to train our model. The model is described in detail in Sect. 3. In Sect. 4, we first simulate and test the model on weather stations used during the training. Then, we validate it on a high resolution gridded dataset at the scale of France. Finally, we conclude with a discussion of future work in Sect. 5.

2 Data

The generator developed in this work is applied to French daily average temperature data. We use weather station data from the ECA&D dataset to fit the model parameters and validate it by comparing gridded spatial simulations to the E-OBS gridded dataset.

2.1 The ECA&D dataset

The ECA&D dataset (<https://www.ecad.eu/>, last access: March 2023; Klein Tank, 2002) is a high-quality European dataset of observed daily station data, covering periods that range from centuries for some stations to decades for others. We used the February 2023 version of the dataset, which includes 2068 stations across Europe with less than 8 % missing data for the period 1985–2015. From these, we selected 41 stations located in France, as shown in Fig. 1, forming the set of locations in space $\mathcal{S}_{\text{stations}} \subset \mathcal{D}$, where $\mathcal{D} = [6^\circ \text{W}; 10^\circ \text{E}] \times [42^\circ \text{N}; 52^\circ \text{N}]$.

These stations are distributed regularly in space, with distances between pairs ranging from 70.95 to 1263.78 km, as illustrated in Fig. 2 (left). They are mostly situated in low-elevation areas (less than 500 m), as depicted in the histogram in Fig. 2 (right).

For each location, missing values in the temperature series were replaced with the average value for the corresponding calendar day. As there are only few stations with missing data, and at most 6.99 % of missing values in that case, using this method does not have a significant impact on the variance of the data. Days corresponding to the 29 February were

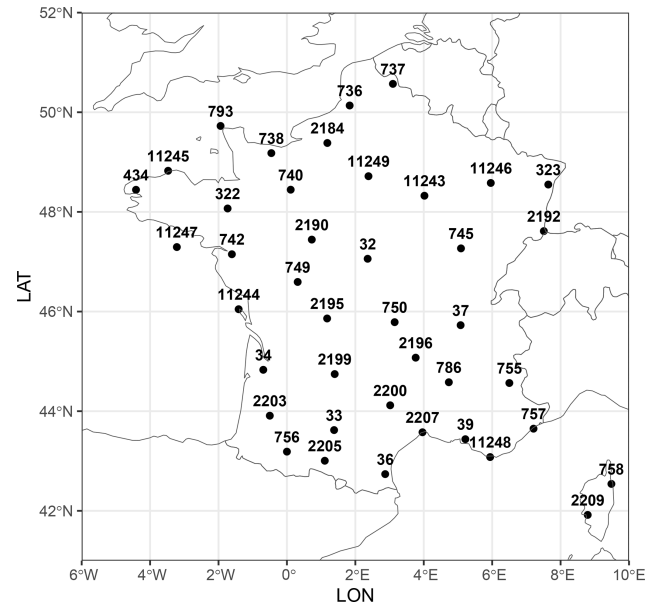


Figure 1. Selected 41 stations from the ECA&D dataset.

removed for simplification. The remaining data is indexed by day $t \in \mathcal{T} = [[1, n_T]]$, with $n_T = 11\,315$ the number of days in the series.

2.2 The E-OBS dataset

The E-OBS dataset (Cornes et al., 2018) is a gridded version of the ECA&D dataset. It is obtained from the station data by statistical learning methods. Each day is interpolated independently first by fitting a spatial trend with a Generalised Additive Model (GAM) on the coordinates and background field, and then a kriging of the residuals with an exponential variogram and a fixed nugget parameter representing measurement uncertainty. The background field represents the monthly values of temperature. The GAM uses smoothing splines in order to model the large-scale spatial trends. The E-OBS dataset exists in 2 versions with different resolutions, 0.1 or 0.25°, we use the 0.25° version of April 2023 (version 27.0e). In the following, the set of grid points will be named $\mathcal{S}_{\text{grid}} \subset \mathcal{D}$, with $|\mathcal{S}_{\text{grid}}| = 1071$.

3 Methodology

In this section, we present the methodology for constructing a spatio-temporal model for daily average temperature. This model extends the single-site temporal model introduced by Hoang (2010) to a spatial context. We adapt the procedures for estimation, simulation, and validation to accommodate the spatio-temporal framework.

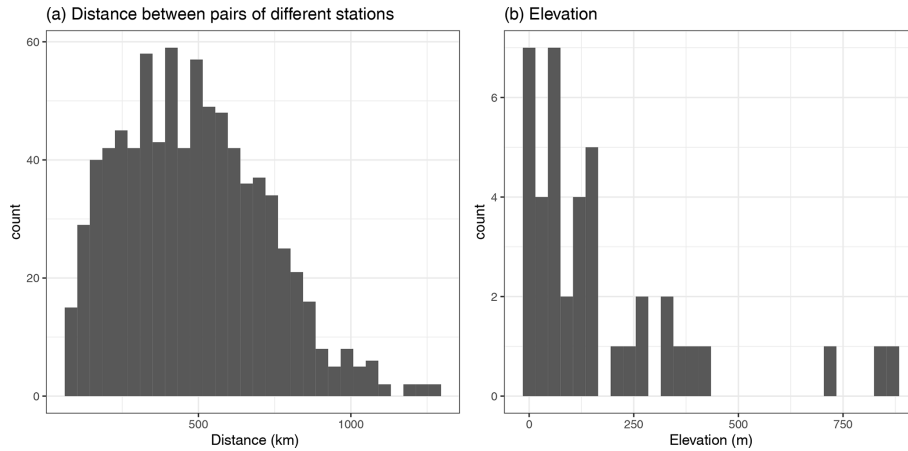


Figure 2. Histograms of the distances between stations (a) and of the elevation within the dataset (b).

3.1 Model

Let $X(s, t)$ be the temperature field at location $s \in \mathcal{D}$ and day $t \in \mathcal{T}$. $X(s, t)$ is written as a combination of deterministic and stochastic terms. As in the very first weather generator of Richardson (1981), the temperature series is standardised by its variance. Following Hoang (2010), our decomposition also accounts for the annual cycle of variance and its temporal trend.

$$X(s, t) = T_m(s, t) + S_m(s, t) + \sqrt{T_{\sigma^2}(s, t)}\sqrt{S_{\sigma^2}(s, t)}Z(s, t), \quad s \in \mathcal{D}, t \in \mathcal{T}, \quad (1)$$

where $T_m(s, t)$ and $S_m(s, t)$ are the trend and the seasonality in the mean, $T_{\sigma^2}(s, t)$ and $S_{\sigma^2}(s, t)$ are the trend and seasonality in the variance. Z is the stochastic part called residuals. The trend terms are interpreted as representing long-term climate variations, and the seasonality terms represent the cyclic influence of the sun's radiation. Both of them represent the deterministic components of the temperature, while the stochastic part represents the intra-annual climate variability. The main assumption made in this work is that spatial and temporal non-stationarities are in the margins, while the dependence structure can be assumed stationary. Consequently, all the non stationarities are taken into account in the trend which is a function of time and space, and the dependence structure is handled in the residuals covariance which is stationary.

3.1.1 Deterministic terms

The trends in temperature over time for a given point in space s , $T_m(s, \cdot)$, $T_{\sigma^2}(s, \cdot)$ are commonly modeled using linear regression or constants. However, to better capture long-term dynamics, we prefer as in Hoang (2010) to use LOESS (Locally Estimated Scatterplot Smoothing; Cleveland, 1979). LOESS regression fits local linear regression lines at time t to the data at time $t_1, \dots, t_n$ based on a chosen span param-

eter sp_s , with tri-cubic weighting proportional to $(1 - |\frac{t-t_i}{sp_s}|^3)^3$. This non-parametric approach provides a more accurate representation of long-term variations compared to a simple linear trend.

The seasonality terms are modeled as trigonometric polynomials of the following form:

$$S_{\theta}^{(\beta_{\theta})}(s, t) = \beta_{\theta,1}(s) + \sum_{i=1}^{d_{\theta}} \left[\beta_{\theta,2i}(s) \cos\left(\frac{2i\pi t}{365}\right) + \beta_{\theta,2i+1}(s) \sin\left(\frac{2i\pi t}{365}\right) \right], \quad s \in \mathcal{D}, t \in \mathcal{T}, \quad (2)$$

with $\beta_{\theta} \in \mathbb{R}^{2d_{\theta}+1}$ the seasonality parameters, with $\theta = m$ or σ^2 either the mean or the variance, and d_{θ} the degree of the trigonometric polynomial. The value of d_{θ} has to be chosen large enough to capture seasonal details (summer, winter, transitions) but not too large to avoid overfitting.

3.1.2 A spatio-temporal model for the residuals

Due to the decomposition of Eq. (1), we assume that the spatio-temporal residuals $Z(s, t)$, $s \in \mathcal{D}$, $t \in \mathcal{T}$ is a Gaussian random field (Cressie and Wikle, 2015) with zero mean, stationary, and isotropic properties. Under these assumptions, the residuals are fully characterised by their covariance function

$$C(h, u) = \text{Cov}(Z(s+h, t+u), Z(s, t)), \quad h \in \mathbb{R}^+, u \in \mathbb{R}^+, \quad (3)$$

where h is the spatial distance and u is the temporal lag.

Choosing a covariance model that reflects the properties of the data is not an easy task. There are numerous covariance models that can describe spatio-temporal dependence, see Porcu et al. (2021).

A straightforward approach to constructing a spatio-temporal covariance function is to use the product of a spatial

covariance function C_S and a temporal covariance function C_T :

$$C_{\text{sep}}(h, u) = C_S(h)C_T(u). \quad (4)$$

This model is known as separable and is computationally convenient because the full covariance matrix can be expressed as the Kronecker product of two smaller matrices. However, as noted in the literature (De Iaco et al., 2019; Porcu et al., 2021), separable models may not adequately capture complex interactions between space and time, particularly for environmental or meteorological variables. In such cases, more sophisticated models that account for space-time interactions are required.

Non-separable spatio-temporal covariance functions can be constructed easily from spatial and temporal covariance functions. For instance the product-sum model is obtained using the two covariance functions C_S and C_T and positive weights a_1, a_2, a_3 by constructing the spatio-temporal function $C_{ST}(h, u) = a_1 C_S(h) + a_2 C_T(u) + a_3 C_S(h)C_T(u)$. Another construction can be made using a single covariance function C and two positive scaling factors a_1, a_2 by $C_{ST}(h, u) = C(\sqrt{a_1^2 h^2 + a_2^2 u^2})$. $\sqrt{a_1^2 h^2 + a_2^2 u^2}$ can be seen as a space-time distance. This construction considers time as a 3rd dimension of space, which oversimplifies the fundamentally different behaviors and units of space and time.

To tackle this problem, Gneiting (2002) proposes a construction method for classes of non-separable covariance function, that is, covariance functions that do not satisfy Eq. (4). It was extended by Allard et al. (2020) for the class of Gneiting-type covariance functions. Given ϕ a completely monotone function, γ a continuous variogram, i.e., a conditionally negative semidefinite function (Chilès and Delfiner, 2012),

$$C_{\gamma, \phi}(h, u) = \frac{\sigma^2}{(\gamma(u) + 1)} \phi\left(\frac{\|h\|^2}{\gamma(u) + 1}\right), \quad h \in \mathbb{R}^2, u \in \mathbb{R}, \quad (5)$$

is a spatio-temporal non-separable covariance function that propose more flexibility in the interaction between the decay of spatial correlation over time than separable models. Usually, the spatial function $\phi(h)$ is of the form $\phi_0(h/r)$ with r a spatial scale parameter. In the case of Eq. (5), the space-time relationship can be considered as using a spatial scale parameter that varies with the temporal lag, with the variogram γ representing how the temporal distance interacts with the spatial distance. Another advantage of this construction is the existence of ad hoc simulation methods as described in Sect. 3.3.2.

We use the following functions for ψ, γ as in Allard et al. (2022):

$$\phi(x) = \mathcal{M}(\sqrt{x}; r, \nu),$$

where $\mathcal{M}(\cdot, r, \nu)$ is the Matérn function with spatial range $r > 0$, regularity $\nu > 0$.

$$\gamma(t) = \left(\left(\frac{t}{a} \right)^{2\alpha} + 1 \right)^b - 1, \quad a > 0, \quad b \in [0, 1], \\ \alpha \in [0, 1]$$

This leads to the model defined as:

$$C_{\gamma, \phi}(h, u) = \frac{\sigma^2}{\left(\left(\frac{u}{a} \right)^{2\alpha} + 1 \right)^b} \mathcal{M}\left(\frac{h}{\sqrt{\left(\left(\frac{u}{a} \right)^{2\alpha} + 1 \right)^b}}; r; \nu \right) \quad (6)$$

The parameter b is called the separability parameter. When it equals to 0, the model reduces to a separable function. When $b > 0$, the model is non-separable.

One drawback of this model is that the temporal covariance $C(0, u)$, which describes the covariance at a fixed spatial location, can be overly constrained by the chosen parameters. A solution to this problem is to multiply Eq. (6) by a purely temporal covariance function of the form $\left(\left(\frac{u}{a} \right)^{2\alpha} + 1 \right)^{-\delta}$. By introducing a slight reparametrization, $\tau = \delta + b$, this yields the Gneiting-Matern covariance model Gneiting (2002, Eq. 16)

$$C_{\text{GM}}(h, u) = \frac{\sigma^2(1 - \eta^2)}{\left(\left(\frac{u}{a} \right)^{2\alpha} + 1 \right)^\tau} \mathcal{M}\left(\frac{h}{\sqrt{\left(\left(\frac{u}{a} \right)^{2\alpha} + 1 \right)^b}}; r; \nu \right) + \sigma^2 \eta^2 1_{h=0, u=0} \quad (7)$$

The nugget effect η^2 is added to account for small-scale variability and measurement error, as a proportion of the total variance.

This model has recently found applications and extensions, particularly for meteorological variables. Gneiting et al. (2010) introduced a multivariate spatial version of the model. Subsequently, Bourotte et al. (2016) and Allard et al. (2022) refined it by proposing multivariate spatio-temporal classes.

3.2 Estimation

Estimating the parameters of the trend and seasonality terms β_θ and the parameters $\sigma^2, \eta^2, \alpha, a, b, r, \delta, \nu$ of the covariance function simultaneously is theoretically possible maximising the whole likelihood function. However, estimating that many parameters at the same time may cause computational challenges due to identifiability issues. We prefer proceeding by a multi-stage approach and estimate “deterministic terms” first (site-by-site), introducing robustness in the estimation. This approach comes down to estimate marginal parameters first, then the covariance parameters, ensuring a stable and interpretable decomposition of the spatiotemporal structure.

3.2.1 Deterministic terms

The different components describing the temperature $X(s, t)$ in Eq. (1) are first obtained for each station location $s \in \mathcal{S}_{\text{stations}}$, for $t \in \mathcal{T} = [[1 : n_T]]$ in 4 steps:

1. First, a LOESS regression on $X(s, \cdot)$ is used to get $T_m(s, \cdot)$. The span parameter sp_s is chosen using the Modified Partitioned Cross-Validation (MPCV) of Hoang (2010).
2. The trigonometric polynomial coefficients $\beta_m(s)$ (Eq. 2) involved in the seasonality of the mean $S_m^{(\beta_m)}(s, \cdot)$, are obtained by linear regression on $X(s, \cdot) - T_m(s, \cdot)$. The order of the polynomial $d_m(s)$ is chosen using the Akaike criterion.
3. The variance term defined as $V(s, \cdot) = (X(s, \cdot) - T_m(s, \cdot) - S_m(s, \cdot))^2$. $T_{\sigma^2}(s, t)$, is obtained by LOESS regression on the variance $V(s, \cdot)$ with the same span parameter sp_s as in the trend in mean.
4. The trigonometric polynomial coefficients $\beta_{\sigma^2}(s)$ in $S_{\sigma^2}^{(\beta_{\sigma^2})}(s, \cdot)$, are obtained by linear regression on $\frac{V(s, \cdot)}{T_{\sigma^2}(s, \cdot)}$. The order of the polynomial $d_{\sigma^2}(s)$ is also chosen with the Akaike criterion and can be different from $d_m(s)$.

This decomposition approach is described in detail in Hoang (2010). It relies on a modified partitioned validation criterion suited to heteroscedastic and dependent data to choose the optimal smoothing parameter. The obtained residuals Z are zero-mean, stationary in time for each site s and with variance that averages to 1 over the year. The order in which the components are estimated bears some importance: if the smoothing parameter is fixed, estimating seasonality first then trends leads to the same results as estimating trends first then seasonality. However, if seasonality is identified first, it includes some part of the trend, and the optimal parameter may not be the same as the one obtained when identifying the trend first.

3.2.2 Residuals

Once the trends and seasonality components are obtained, the values of the residual field $Z(s, t)_{s \in \mathcal{S}_{\text{stations}}, t \in [[1 : n_T]]}$ can be calculated and used to estimate the parameters of the Gneiting-Matern covariance model in Eq. (7).

The covariance matrix of the vector residuals Z is of size $(|\mathcal{S}_{\text{stations}}| \times n_T)^2$, and computing its inverse becomes infeasible when $|\mathcal{S}_{\text{stations}}|$ and n_T exceed a few tens, which limits the use of full likelihood methods. Composite Likelihood methods (Lindsay, 1988; Varin et al., 2011), which involve only pairwise likelihoods, are much easier to compute and exhibit good convergence properties. For a dataset $(y_1, \dots, y_n)$ and a vector of parameters θ , this approach means replacing the optimisation of the likelihood $L(y_1, \dots, y_n; \theta)$ by the pairwise likelihood

$PL(y_1, \dots, y_n; \theta) = \prod_{i < j} L(y_i, y_j; \theta)$. To further reduce the number of terms in the pairwise likelihood, we employ the Weighted Composite Likelihood (WCL), assigning a weight of 0 to pairs of sites that are far apart based on a threshold distance, chosen as half the maximum observed distance (around 650 km in this study) and a weight of 1 otherwise. This choice of half the observed distance is a usual practice in the field of geostatistics. Bourotte et al. (2016) for example have studied the relative efficiency of the pairwise likelihood estimator, showing that it first increases then decreases as a function of this cutoff parameter.

In the pairwise estimation procedure, the score function $\nabla L(y_1, \dots, y_n; \theta)$ is replaced by the composite score $\nabla PL(y_1, \dots, y_n; \theta)$. In that case, the classical Fisher information matrix is no longer appropriate for variance estimation. Instead, we rely on the Godambe information matrix (Lindsay, 1988) defined as

$$\mathbf{G}(\theta) = \mathbf{H}(\theta)^{-1} \mathbf{J}(\theta) \mathbf{H}(\theta)^{-1}, \quad (8)$$

where $\mathbf{H}(\theta)$ is the sensitivity matrix (the negative expected derivative of the pairwise score function) and $\mathbf{J}(\theta)$ is the variability matrix (the variance of the pairwise score). Because these components are not directly available from the pairwise likelihood, they are estimated via empirical averages of the observed scores and their derivatives (Varin et al., 2011).

3.3 Simulation

Simulations are to be performed over an arbitrary set of points in space $\mathcal{S}$, such as a grid of points $\mathcal{S}_{\text{grid}}$ covering the country, according to Eq. (1), where the deterministic components are extrapolated on each point, and the residuals are simulated as a spatio-temporal Gaussian field with covariance function determined by Eq. (7).

3.3.1 Spatial extension of the deterministic terms

Spatial extrapolation of parameters computed at station locations is achieved using ordinary kriging (Cressie, 1991) a statistical interpolation method that leverages spatial correlation between data points. The key to kriging is the variogram model $\gamma(h)$, which quantifies the spatial dependence between points at distance h . For a spatial variable W , the variogram is defined as $\gamma(h) = \frac{1}{2} E[(W(x) - W(x+h))^2]$. Prediction at a location x_0 is a weighted sum of the observed values $W^*(x_0) = \sum_{i=1}^n \lambda_i W(x_i)$, where the weights depend on the chosen variogram model.

Seasonality

The seasonality coefficients $\beta_m(s), \beta_{\sigma^2}(s)$ for $s \in \mathcal{S}$ are obtained by ordinary kriging of the coefficients calculated at locations $s \in \mathcal{S}_{\text{stations}}$ using a Gaussian variogram fitted to the

observed data. The Gaussian variogram is defined as follows:

$$\gamma_g(h) = g_1 + g_2 \left(1 - \exp \left(-\frac{h^2}{g_3^2} \right) \right), \quad g_1, g_2, g_3 \geq 0. \quad (9)$$

The annual cycles are then obtained by plugging these estimated coefficients in the seasonality functions defined in Eq. (2).

Trends

To provide the trends at each grid point, we first correct the trend in mean at sites $s \in \mathcal{S}_{\text{stations}}$ using the International Standard Atmosphere (ISA) model (ISO, 1975) which describes the evolution of the standard temperature as a function of altitude:

$$T(z) = T_0 - cz, \quad (10)$$

where $c = 6.5 \times 10^{-3} \text{ } ^\circ\text{C m}^{-1}$ and T_0 is the temperature at sea level (0 elevation). An equivalent temperature trend at $z = 0$ is given by

$$T_{\text{eq}}(s, t) = T_{\text{m}}(s, t) + cz(s), \quad s \in \mathcal{S}_{\text{stations}}. \quad (11)$$

T_{eq} is interpolated from $\mathcal{S}_{\text{stations}}$ to $\mathcal{S}$ by kriging with a linear variogram $\gamma_1(h) = g + h$, $g \geq 0$. The trend at any grid point s is then recovered by:

$$T_{\text{m}}(s, t) = T_{\text{eq}}(s, t) - cz(s), \quad s \in \mathcal{S}. \quad (12)$$

3.3.2 Spatio-temporal Gaussian process simulation

Simulating from the covariance function in Eq. (7) at sites $\mathcal{S}$ and times $t \in \mathcal{T}$ means simulating from a Multivariate Gaussian process with $|\mathcal{S}| \times |\mathcal{T}|$ components. Working with large spatiotemporal grids makes it unmanageable to simulate a Gaussian process using standard procedures like the Cholesky decomposition, which has in general a complexity in $O(|\mathcal{S}|^3 \times |\mathcal{T}|^3)$, and requires the manipulation of matrices of size $|\mathcal{S}|^2 \times |\mathcal{T}|^2$. To address this issue, we use two different methods: an iterative method and a spectral method.

Iterative simulation

We take advantage of the temporal structure of the data, noting that time steps far apart are unlikely to be dependent. Using this observation, simulations at time t are obtained iteratively, conditioned only on the ℓ previous time steps, according to the conditional distribution of Gaussian vectors. The choice of ℓ is consistent to the temporal range parameter estimated in the covariance model. The Cholesky decomposition is then computed on matrices of size $(|\mathcal{S}| \times \ell)^2$ at most, which is reasonable if ℓ is not too large.

Spectral simulation

Allard et al. (2020) propose various algorithms based on spectral decomposition for simulating spatial and spatio-temporal processes with covariance models of Gneiting type, such as in Eq. (5). We use the algorithm known as the substitution approach.

Let Y be the spatio-temporal process defined by

$$Y(s, t) = \sqrt{-2 \log(U)} \cos(\sqrt{2R} \langle \Omega, x(s) \rangle + |\Omega|/\sqrt{2}W(t) + \Phi), \quad s \in \mathcal{S}, t \in \mathcal{T}, \quad (13)$$

where, with the notations of Eq. (5), $x(s)$ are the cartesian coordinates of s , $U \sim U([0, 1])$, $R \sim \mu$, $\Omega \sim \mathcal{N}(0, I_2)$, $W \sim \gamma$, $\Phi \sim U([0, 2\pi])$ and the measure μ is the measure related to the Bernstein representation of ϕ . This process has the covariance function in Eq. (5). For p copies of such a process Y_i , $i \in [1, \dots, p]$, the Central Limit Theorem states that

$$\tilde{Z}(s, t) = \frac{1}{\sqrt{p}} \sum_{i=1}^p Y_i(s, t), \quad s \in \mathcal{S}, t \in \mathcal{T}, \quad (14)$$

converges to a Gaussian process with the covariance function of Eq. (5) when $p \rightarrow \infty$.

This approach separates the simulation of the spatial component (simulated from μ) and the temporal component (simulated from γ), thereby reducing the dimensionality of the problem in both space and time. It is more scalable than the iterative method, which is faster for smaller grids but becomes very slow as the grid size increases. Conversely, the spectral simulation remains tractable even for larger grids, with a complexity in $O(p|\mathcal{S}| \times |\mathcal{T}|)$.

3.4 Validation

To assess the validity of the weather generator, several indicators are computed for the simulations and compared to those obtained from the observations. We will compare the envelope of the simulated values with the observed values to evaluate the generator.

For spatio-temporal models, informative indicators are correlation between pairs of stations $(s_i, s_j) \in \mathcal{S}_{\text{stations}}^2$ both in the temperature $\text{Cor}(X(s_i, \cdot), X(s_j, \cdot))$ and in the residuals $\text{Cor}(Z(s_i, \cdot), Z(s_j, \cdot))$ (Sparks et al., 2018; Wilks, 1998).

Because of the non-separable assumption, it is also important to evaluate the spatio-temporal correlations, between different stations and at different time lags: $\text{Cor}(Z(s_i, \cdot + \Delta t), Z(s_j, \cdot))$.

Finally, to evaluate the ability of the model to generate joint spatial extremes, we consider the pairwise conditional threshold exceedances for high (or low) quantiles, $\bar{p}_{i,j}^\alpha$ (and $p_{i,j}^\alpha$): for $0 < \alpha < 1$ and $q_\alpha(i)$ such that $\alpha = \mathbb{P}(X(s_i, \cdot) < q_\alpha(i))$ and a pair of stations $(s_i, s_j) \in \mathcal{S}_{\text{stations}}^2$

$$p_{i,j}^\alpha = \mathbb{P}(X(s_i, \cdot) < q_\alpha(i) | X(s_j, \cdot) < q_\alpha(j)) \quad \text{and} \quad \bar{p}_{i,j}^\alpha = \mathbb{P}(X(s_i, \cdot) > q_{1-\alpha}(i) | X(s_j, \cdot) > q_{1-\alpha}(j)), \quad (15)$$

They provide insight into how pairs of stations behave during extreme temperature events and are empirically estimated by

$$\hat{p}_{i,j}^{\alpha} = \frac{\sum_{t \in \mathcal{T}} 1_{X(s_i,t) < q_{\alpha}(i) \cap X(s_j,t) < q_{\alpha}(j)}}{\sum_{t \in \mathcal{T}} 1_{X(s_j,t) < q_{\alpha}(j)}} \quad \text{and} \quad \bar{\hat{p}}_{i,j}^{\alpha} = \frac{\sum_{t \in \mathcal{T}} 1_{X(s_i,t) > q_{1-\alpha}(i) \cap X(s_j,t) > q_{1-\alpha}(j)}}{\sum_{t \in \mathcal{T}} 1_{X(s_j,t) > q_{1-\alpha}(j)}}. \quad (16)$$

The values of $\hat{p}_{i,j}^{\alpha}$ for the winter season and $\bar{\hat{p}}_{i,j}^{\alpha}$ for the summer season, for all pairs $\{i, j\}$, are then compared with the values from the simulations. Note that $q_{\alpha}(i)$ is not necessarily equal to $q_{\alpha}(j)$.

In order to better understand the capacity of the model to show spatial dynamics, let us define the quantile exceedance ratio $R_{\alpha}(t)$. It represents the proportion of locations in a set of locations $\mathcal{S}$ exceeding a given quantile at any time. For $t \in \mathcal{T}$, $0 < \alpha < 1$, this proportion is written as

$$R_{\alpha}(t) = \frac{\sum_{s \in \mathcal{S}} 1_{X(s,t) > q_{1-\alpha}(i)}}{|\mathcal{S}|}. \quad (17)$$

This indicator serves two main purposes. First, its survival function $\mathbb{P}(R_{\alpha}(t) \geq \cdot)$, computed over all time samples, describes the model's ability to generate "spatial hot days". These are days when at least a certain (high) percentage of the domain exceeds their (high) quantile of a specified order. Second, the indicator can be used to examine the temporal evolution of this proportion, which relates to persistent heat waves. Heat waves with the most significant impact (large spatial extent over several days) are expected to exhibit a plateau for most quantiles, including even higher ones.

It is similar to the heatwave indicator of Miloshevich et al. (2023), where a measure of persistence and amplitude of anomalies is given by

$$A(t) = \frac{1}{T} \int_t^{t+T} \frac{1}{\mathcal{D}} \int (X(t', s) - \mathbb{E}(X(t', s))) ds dt', \quad (18)$$

where T is the time scale of interest, $\mathcal{D}$ the spatial domain, X the temperature. $\mathbb{E}(X(t', s))$ is the expected value of temperature at day t' . This indicator averages the temperature anomaly over both space and a time period of length T . Compared to our quantile exceedance ratio of Eq. (17), it can yield non-zero values even for mild events, as long as the temperature remains above its expected value over the given period and space. However, the resulting value may be low if extreme temperatures occur only in part of the domain, while the rest of the area experiences mild or even cool temperatures, making it difficult to distinguish between different types of events. In contrast, our indicator reflects the presence of extreme values within part of the area. Additionally, this indicator is limited to detecting events that occur on a time scale close to T . Isolated hot days may go undetected.

4 Application: temperature over France

4.1 Estimated parameters

4.1.1 Deterministic components

The decomposition for each station in the ECA&D dataset is computed on the 1960–2015 data when available, or otherwise on the longest available period. Because of the initial choice of stations, it is ensured that at least the 1985–2015 period is covered. The decomposition produces the non-parametric trends and the seasonality parameters for each station.

An example of single-site decomposition is shown in Fig. 3 for the Toulouse-Blagnac station. The seasonality reveals the typical winter/summer oscillation. The trend in the mean indicates a general increase with occasional low peaks, showing an average increase of 2° over the years of observations. This highlights the necessity of including a trend and using a non-linear model. For this station, as for the others, the seasonal variance is higher in winter compared to summer. The trend in variance reflects contrasting patterns in winter and summer, along with interannual variability. The seasonality term in the variance acts as a modulating factor on the trend component, together forming the overall variance structure. The trend captures long-term changes and can be considered approximately constant throughout a single year. The seasonal component adjusts this baseline, meaning that the variance is more pronounced in winter than in summer. The slight increase in variance observed during summer in Toulouse in Fig. 3 likely reflects an increasing trend in summer compared to the annual trend, possibly because extreme events with higher amplitude tend to occur more frequently in summer in the context of French climatology. Other sites may not have this increase, instead have something like a low plateau in the summer.

4.1.2 Spatial model

The decomposition into deterministic and stochastic terms has removed most of the seasonality from the signal. We assume hence that the model is stationary in time. However, initial investigations revealed differences between seasons in the empirical spatial covariance structure of the residuals Z . To address this, the model for the residuals is built separately for two seasons: the extended winter (October, November, December, January, February, March – ONDJFM) and the extended summer (April, May, June, July, August, September – AMJJAS). The residuals Z are assumed to be stationary in those seasons.

The estimation of covariance parameters was conducted for all years at once, for each season, by implementing the pairwise likelihood in Julia. The optimization was done using the LBFGS algorithm and automatic differentiation (AD) using an appropriate package compatible with AD for the Matérn function (Geoga et al., 2022). A maximum distance

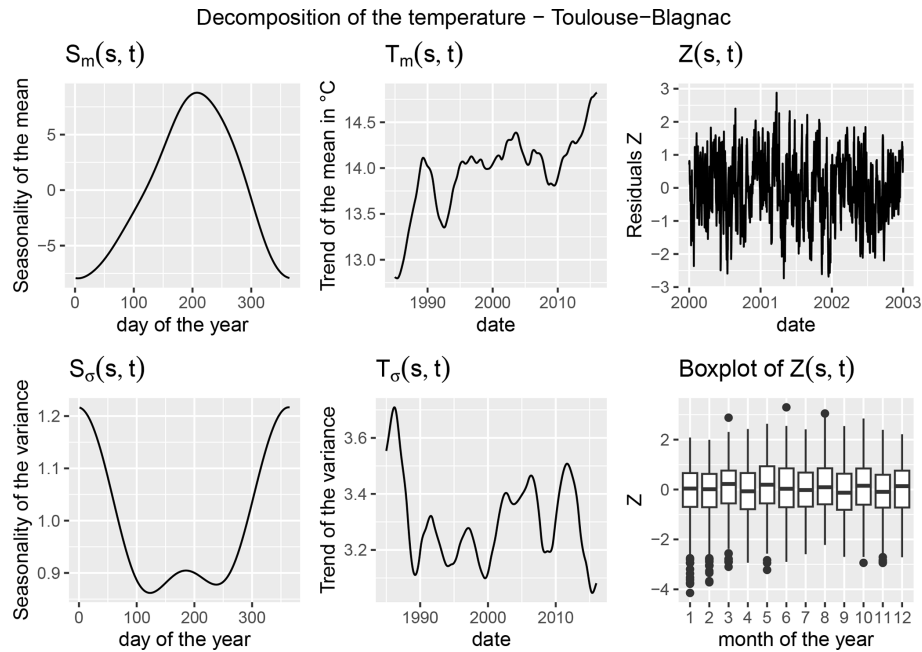


Figure 3. Decomposition at the Toulouse station (south-west of France). From left to right and top to bottom: seasonality in mean, trend in mean, 3 years of residuals Z , seasonality in variance, trend in variance and monthly boxplots of the residuals Z .

cutoff of 650 km in space and 5 d in time was applied to reduce the number of pairs and avoid rare spatial differences.

We also estimated the variance of our pairwise likelihood estimator, using the Godambe information matrix (see Varin et al., 2011). The $\mathbf{H}$ matrix in Eq. (8) is estimated using the Hessian of the objective function for the observations. The $\mathbf{J}$ matrix is estimated by computing the variance-covariance matrix of the gradient on 100 simulations.

To check the robustness of the parameter estimation, the training data is split into 4 groups of equal size, with each group removed once from the estimation stage to construct 4 models in addition to the complete model estimated using all available station data. This provides insight on the impact of different input data on the estimation and simulations. The obtained parameters are shown in Table 1. Except for the spatial range parameter r , there are not many differences between estimations using the full data or estimations removing a fourth of the data.

Several strategies for estimation were conducted by fixing different sets of parameters to better understand their relationships, the likelihood, and the corresponding covariance function.

The first thing to notice is that adding the flexibility in the temporal covariance with the parameter δ was in the end not necessary. In practice, the best value for this parameter is 0. As this is the minimum allowed value, estimating the hessian and gradient is not relevant and the estimated standard deviation for this parameter is not meaningful.

The spatial shape parameter ν is estimated close to 1 in both seasons, meaning the spatial regularity is not very differ-

ent in the two seasons. It is interesting to see that the Matérn model for $\nu = 1$ corresponds to a linear $\times$ exponential covariance model, which is coherent with the observed empirical curves. The fact that ν is close between the seasons allows the comparison of the spatial scale parameter r , related to the typical correlation distance. It is estimated at around 800 km for both seasons, though higher in the winter than in the summer. In a Matérn model with $\nu \simeq 1$, the correlation drops below 0.05 at around $h = 5.6r$. In our case, this means that for $u = 0$ the stations become uncorrelated for a distance of more than 4000 km, which is quite large, but expected as the temperature variable exhibits strong large-scale correlation. The estimated spatial scale is inherently dependent on the data used to estimate it. When removing a fourth of the stations from the dataset, the estimated r can be quite increased in some cases, however when that happens, the summer and winter values stay coherent.

The temporal scale parameter a is estimated as higher in the winter than in the summer, but the associated standard deviation of the estimate is also quite larger. This highlights the difficulty to fit a model in the winter, where the data has both large-scale episodes and sudden variations.

The separability parameter $b \in [0, 1]$ is estimated as 0.663 in the summer and 0.883 in the winter. Both of these values are far from 0 and indicate a clear non-separable behaviour of the spatio-temporal dependence.

The obtained variogram functions $\Gamma(h, u) = \frac{1}{2} \text{Var}[Z(s, t) - Z(s + h, t + u)]$, along with the empirical function computed on the observed data $\hat{\Gamma}(h, u)$, are shown in Fig. 4. The Gneiting-Matérn model fits the data

better than the separable model, which struggles to represent space-time interactions even for small time differences. The nugget effect accounts for approximately 10 % of the total variability and reflects short-range variation caused by several factors, including measurement error, micro-topographic features, and spatial variability at scales smaller than the sampling resolution.

4.2 Simulation at the training sites

Temperature series were simulated 100 times at each of the 41 fitting data points by first simulating residuals over 31 summers and 31 winters from the fitted spatio-temporal covariance model (Eq. 7) and then adding the deterministic parts in Eq. (1).

For the stochastic component, different simulation methods were employed. The Gneiting-Matérn model was simulated using both the iterative and spectral methods. Additionally, the iterative method was applied to simulate from two alternative models: first, a purely temporal version of the covariance model (no spatial component), aimed at assessing how much spatial information is conveyed by the deterministic trend and seasonality versus the spatial model in the stochastic component. A separable exponential model was also used to demonstrate the limitations of an overly simplified model.

4.2.1 An example of simulation

An example of 4 consecutive days in one simulation is presented in Fig. 5. The first column displays the observations, the second column shows a simulation based on a separable spatio-temporal covariance model, and the next two columns present simulations generated using the non-separable covariance function (via the iterative and spectral methods, respectively). The last column illustrates a simulation according to a spatially independent model. While a direct comparison of the values is not meaningful, we can evaluate whether the range of values is reasonable and whether the spatial pattern is adequately reproduced. Obviously, the last column shows a more discontinuous pattern compared to the others, highlighting the importance of incorporating a spatial component in the model.

4.2.2 Marginal properties

Figure 6 presents the violin plots of the observed and simulated temperatures at several stations for both extended seasons (summer on the top row, winter in the bottom row). Across all models and simulation methods, the distribution of simulated temperatures closely matches the distribution of observed temperatures, demonstrating the model’s ability to accurately reproduce the temperature distribution.

Table 1. Estimated covariance parameters and their estimated standard deviation using a bootstrapped Godambe information matrix.

Season	Stations removed	σ^2	(SD)	η^2	(SD)	r	(SD)	a	(SD)	α	(SD)	b	(SD)	δ	(SD)	ν	(SD)
AMJIAS	None	0.999	(0.026)	0.121	(0.005)	746.22	(32.623)	1.703	(0.427)	0.948	(0.093)	0.663	(0.056)	0	(0.196)	0.978	(0.036)
AMJIAS	Group 1	0.999	(0.031)	0.116	(0.005)	653.68	(27.707)	1.704	(0.424)	0.952	(0.092)	0.658	(0.068)	0	(0.171)	1.111	(0.038)
AMJIAS	Group 2	1.000	(0.027)	0.100	(0.004)	695.78	(35.801)	1.642	(0.383)	0.951	(0.086)	0.656	(0.069)	0	(0.164)	1.044	(0.040)
AMJIAS	Group 3	0.998	(0.029)	0.107	(0.007)	922.33	(43.050)	1.740	(0.541)	0.944	(0.113)	0.677	(0.054)	0	(0.277)	0.769	(0.025)
AMJIAS	Group 4	0.998	(0.029)	0.133	(0.006)	972.65	(44.960)	1.716	(0.430)	0.944	(0.092)	0.662	(0.066)	0	(0.208)	0.750	(0.029)
ONDJFM	None	1.015	(0.031)	0.100	(0.007)	870.91	(60.763)	2.599	(2.037)	0.814	(0.170)	0.883	(0.066)	0	(0.747)	1.003	(0.035)
ONDJFM	Group 1	1.015	(0.034)	0.098	(0.005)	779.88	(45.925)	2.601	(1.313)	0.818	(0.107)	0.877	(0.105)	0	(0.426)	1.102	(0.035)
ONDJFM	Group 2	1.014	(0.031)	0.090	(0.005)	882.48	(56.705)	2.517	(1.610)	0.814	(0.137)	0.871	(0.077)	0	(0.566)	0.967	(0.036)
ONDJFM	Group 3	1.016	(0.031)	0.100	(0.011)	973.54	(73.861)	2.565	(2.497)	0.822	(0.218)	0.858	(0.069)	0	(0.962)	0.904	(0.026)
ONDJFM	Group 4	1.015	(0.032)	0.098	(0.036)	1029.93	(355.96)	2.670	(10.576)	0.805	(0.868)	0.915	(0.155)	0	(3.971)	0.848	(0.039)

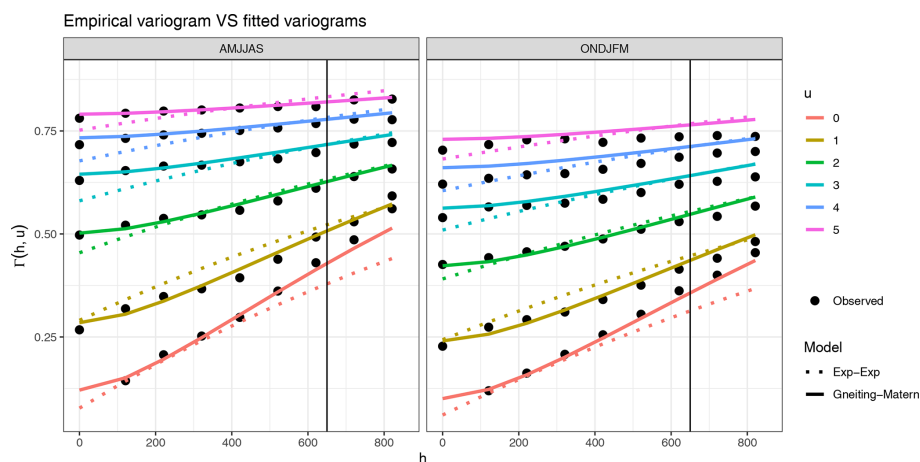


Figure 4. Variogram functions corresponding to the fitted parameter in winter (ONDJFM) and in summer (AMJJAS) for the separable model (Exp-Exp) and the Gneiting-Matérn model (Gneiting-Matérn). The x -axis represents the distance in space, the y -axis holds the values of the variogram for several values of temporal distances u denoted each by a colored curve. The empirical variogram in the data is represented by the circle points. The vertical line indicates the spatial cutoff, beyond which the pairs were not taken into account in estimation. The value for $h = u = 0$ is missing from the empirical values because it should be the value of the nugget effect and cannot be computed in the empirical variogram.

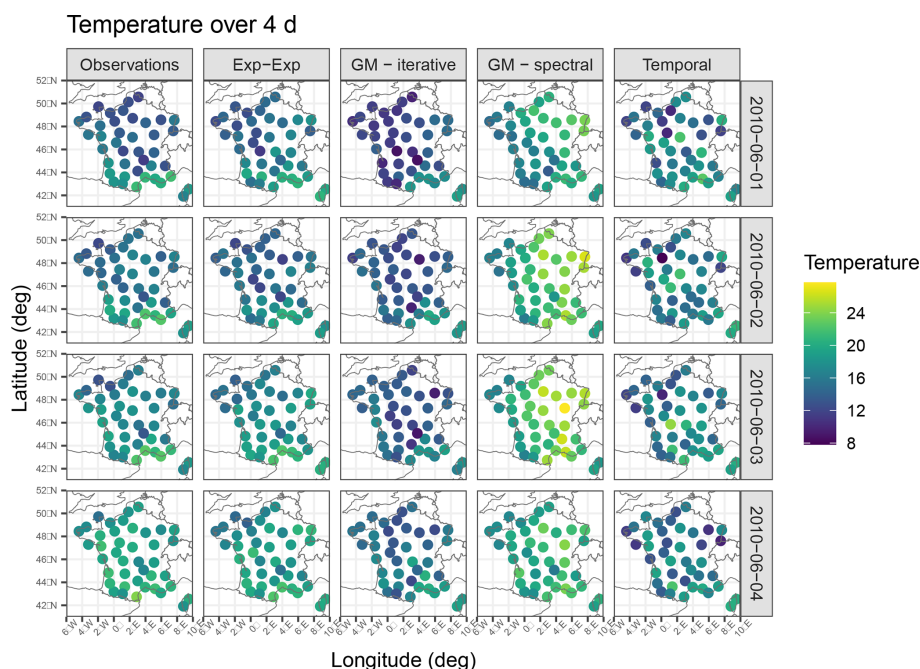


Figure 5. Observations and simulations using the separable model (Exp-Exp), the 10 d-memory simulation algorithm (GM-iterative), the spectral method (GM-spectral) and the purely temporal model (Temporal).

4.2.3 Spatial correlation

Figure 7 (left) illustrates the pairwise correlation coefficients between stations, comparing simulated correlation in the residuals $\text{Cor}(Z(s_i, \cdot), Z(s_j, \cdot))$ (right) and temperatures $\text{Cor}(X(s_i, \cdot), X(s_j, \cdot))$ (left) with their observed counterpart. For both simulation methods based on the spatio-temporal Gneiting-Matérn model – conditional simulation over the

previous 10 d and spectral simulation – the observed and simulated correlations align closely with the 1 : 1 line. This indicates that this choice of correlation function effectively captures the spatial dependence structure in the data. In contrast, using a purely temporal model results in a correlation of zero for the residuals (Z). Although the corresponding simulated temperature values are non-zero due to the de-

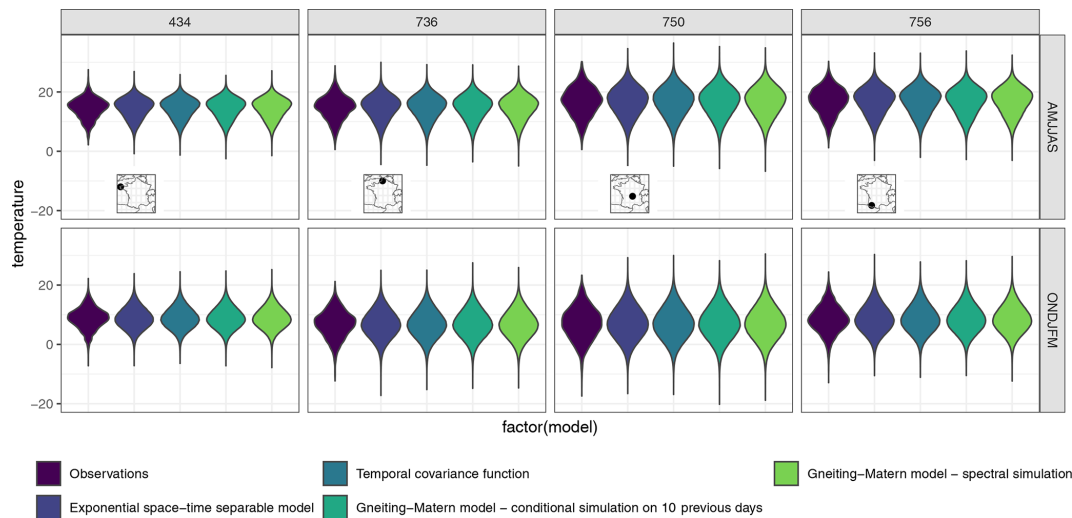


Figure 6. Violin-plot of the temperature in the two extended seasons (summer on the top row, winter in the bottom row) for 4 stations, with their corresponding locations. For simulations, 100 simulations were made using a temporal only model, a separable model and the non-separable model using the methods described in Sect. 3.3.2.

terministic components, they deviate significantly from the observed values, emphasizing the benefit of incorporating a spatial model. The separable model performs similarly to the non-separable model but overestimates lower correlations. Although the significance of the non-separability property is not immediately evident in this specific graph, its advantages may become clearer when analyzing more complex spatial patterns or longer time scales.

To illustrate where the non-separability property enhances the model's performance, Fig. 8 presents the covariance function $C(h, u)$ for the residuals Z at various time lags u . This function is computed from both the observations and one simulation for each of the three models: temporal only, separable, and non-separable. The purely temporal model fails to reproduce the observed correlations, as expected. Because the observed spatial covariance is scattered, making direct comparison difficult, a LOESS smoothing line is added to help visualise the difference between the separable and non-separable models. For $u = 0$, both spatio-temporal models (separable and non-separable) overestimate the covariance for stations that are far apart, reflecting the previous observations from pairwise correlations. However, the non-separable model demonstrates a modest improvement from the separable model in these cases. This overestimation likely results from the fact that far-apart station pairs are under-represented in the dataset and were not used during model training. For non-zero time lags, the separable model tends to overestimate high covariances and underestimate low covariances compared to the non-separable model. This discrepancy emphasizes the necessity of using a non-separable model, as further detailed in Appendix B, where we investigate empirically the separability of the data.

Figure 7 (right) shows the pairwise conditional probabilities of threshold exceedance in the observations compared to the simulations. For all models except the spatially uncorrelated model, the values in simulation are in the range of the observations. For the highest quantile (0.99), the small probabilities are overestimated, and the large probabilities are underestimated. This means that the dependence is overestimated when it is high, and underestimated when low. Such behaviour was to be expected as the Gaussian distribution is known to be asymptotically independent (Coles, 2001). Nonetheless, once the effects of trends and seasonalities are accounted for, the resulting extremal dependence in the simulated temperatures aligns reasonably well with the observed data, except at the very most extreme quantiles. Enhancing the model accordingly would require at least two improvements. First, adopting a distribution that better captures the behavior in the extremes, such as (extended) Pareto distributions. However, this shift would introduce new complexities during parameter estimation. An alternative would be to explicitly model the extremes in a separate statistical framework. Second, choosing a dependence structure for one that exhibits stronger dependence in the extremes such as max-stable models which are asymptotically dependent or one intermediate.

4.2.4 Quantile Exceedance Ratio

Focusing on the summer months, with quantiles $q_i(\alpha)$ and Quantile Exceedance Ratios $R_\alpha(t)$ (as defined in Eq. 17) computed only on June, July and August (JJA) data and simulations, the survival function of $R_\alpha(t)$ for several values of α is represented in Fig. 9. There is about 10 % chance that at least 15 out of the 41 stations exceed their 90 % quantile

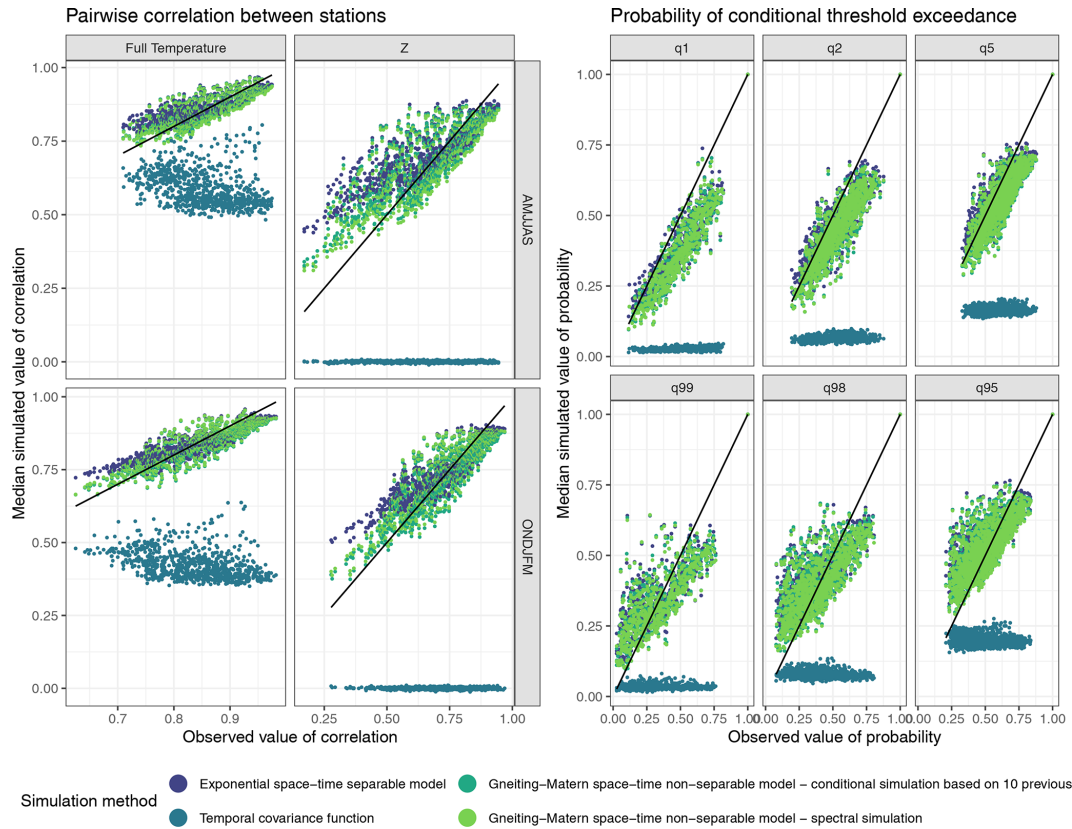


Figure 7. Correlations and conditional probabilities of threshold exceedance. x -axis: observed values, y -axis: median simulated values for 100 simulations. Left: Pairwise correlation coefficient for all pairs of stations, left in the full temperature $\text{Cor}(X(s_i, \cdot), X(s_j, \cdot))$ and right in the residuals $\text{Cor}(Z(s_i, \cdot), Z(s_j, \cdot))$. Top row represents the result in the extended summer, second row in the winter. Right: Empirical pairwise conditional probabilities from Eq. (15). Top: $\mathbb{P}(X(s_i, \cdot) < q_\alpha(i) | X(s_j, \cdot) < q_\alpha(j))$ for $\alpha = 0.01, 0.02, 0.05$. Bottom: $\mathbb{P}(X(s_i, \cdot) > q_{1-\alpha}(i) | X(s_j, \cdot) > q_{1-\alpha}(j))$, for $\alpha = 0.01, 0.02, 0.05$.

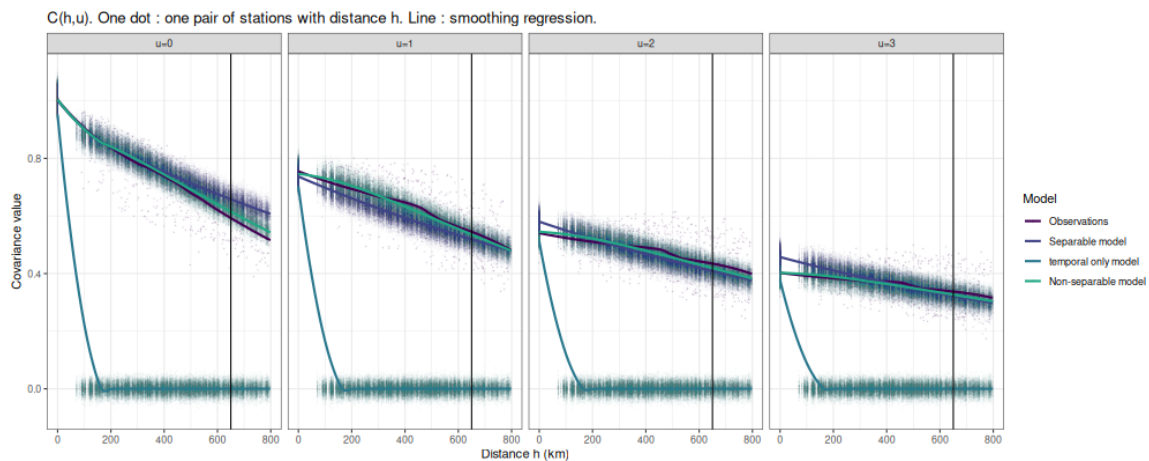


Figure 8. Empirical covariance in the reduced variable Z , $C(h, u)$, for different time lags $u = 0, 1, 2, 3$. Each dot denotes the covariance between a pair of stations separated by the distance h . The small dots are the covariance for each pair, the solid line the loess values. The colored dots correspond to 100 simulations for the observations and all 3 possible models: temporal-only, separable, non-separable.

at the same time. In the simulations, it has about 9 % chance. Overall, the simulations tend to underestimate the probability to exceed a threshold for several locations at the same time. This is coherent with the previous observations on pairwise threshold exceedance probabilities, where we found that the simulations underestimated the high values.

4.3 Simulations on a grid

This section shows the use of the generator trained on the station data for simulation over the E-OBS grid. The simulations are then compared to the E-OBS data that serves as a reference.

4.3.1 1985–2015 simulations

We generated 100 simulations over the same 1985–2015 period. Examples of these simulations are shown in Fig. 10, alongside the corresponding values from the E-OBS database. Among the 100 simulations, we selected examples closest to the median temperature of E-OBS, as well as the coldest and warmest simulations. Although direct comparison of day-to-day values is not meaningful, the overall patterns are similar. The range of simulated temperatures is broad, accommodating both cool and hot days with strong temporal correlation.

The simulated temperatures appear less smooth than those from the E-OBS dataset, which is expected for several reasons. First, E-OBS is an interpolated product based on station measurements, resulting in inherently smoother spatial fields. In contrast, our simulation model includes a nugget effect, which explicitly captures short-range variability – such as measurement error or micro-scale environmental effects – as a proportion of the total variance.

The spatial resolution of E-OBS is approximately 25 km, while the minimum distance between stations in the training data was closer to 100 km. This relatively coarse station density limited the model's ability to capture very short-range spatial structure, likely leading to an overestimation of the nugget effect and, consequently, rougher simulations.

However, the presence of a nugget effect in our model is not necessarily a drawback. It reflects a realistic component of spatial variability at small scales and ensures that simulations account for local heterogeneity. With denser station coverage, this short-range behavior could be estimated more precisely, potentially yielding smoother simulated fields. Alternatively, removing the nugget effect would produce smoother outputs, but at the risk of under-representing local variability and generating unrealistically homogeneous fields.

The quantile exceedance ratio $R_\alpha(t)$ can be analyzed in the same manner. Figure 11 shows the survival function for various quantiles. These quantiles are extracted from the E-OBS database for each grid cell. The simulated survival function is represented by a ribbon spanning the 0.1–0.9 quantile range

of the simulations, with a dotted line indicating the median value, and the survival function in E-OBS is represented by a full line. The simulations tend to overestimate the number of heat days over smaller areas while underestimating the frequency of events covering larger areas, compared to the observed values.

It is also important to reproduce not only isolated hot days but also sequences of hot days with length comparable to the observed events. Figure 12 illustrates the evolution of $R_\alpha(t)$ for several quantiles in some observed summers and some simulated summers. It illustrates the model's ability to reproduce summer heatwave characteristics in terms of duration, intensity, and spatial extent. Each panel shows the proportion of the domain exceeding a temperature quantile (color-coded) over the days of summer. The two left panels show observed summers of 2003 and 2009, with 2003 notable for its intense and long-lasting heatwave affecting much of France, used here as a reference. Its exceptional spatial extent and duration are highlighted with arrows. The other panels show examples of simulated summers. These demonstrate the model's capacity to generate heat events of varying types: short and localised events (all graphs), short events that are large in space (bottom right), events resembling the 2003 heatwave in both intensity and duration (top middle and top right), and even a long-duration heatwave with slightly less intensity than 2003 (bottom middle). These examples support the model's ability to capture realistic variability in heatwave behavior.

We consider spatial heat episodes as sequences of days when the temperature exceeds a specified threshold in at least a given proportion of the domain. Figure 13 shows the distribution of heat episode lengths for both E-OBS (red dots) and the 100 simulations (black bars).

For heat episodes that cover at least 40 % or 60 % of the domain, and for moderate temperature quantiles such as the 0.85 or 0.9 quantiles, the observed values fall within the range of simulated values. The model even succeeds in generating events longer than those observed. However, for higher spatial proportions, the simulations struggle more to generate extended episodes, indicating some limitations in replicating the most extreme spatial heat events. This is coherent with the conclusion given by the previous Fig. 11.

Cold events can be similarly defined as sequences of days when the temperature in winter (December–January–February) is under a given threshold for a given proportion of the domain. Figure 14 shows the distribution of the length of cold events in the observations and the simulations. For events that are below the 0.05 quantile, the model generates an appropriate number of single-day events for all considered spatial extents, but does not reproduce events with high temporal length. However, for such a low quantile and for a large spatial extent, the observations contain only a few events. For milder quantiles, the model is able to generate longer events but overestimates the number of single day events.

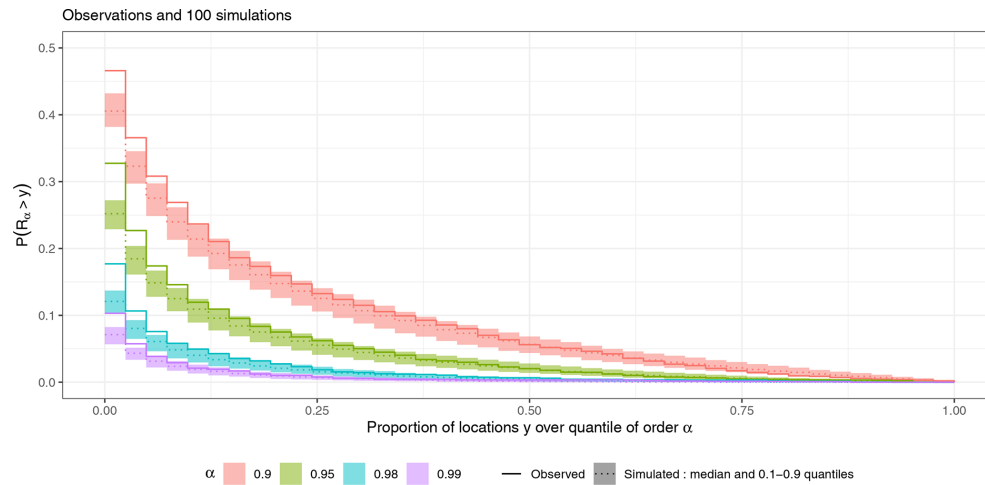


Figure 9. Survival function of the Quantile Exceedance Ratio, $\mathbb{P}(R_\alpha(t) \geq y)$ for several quantiles α , computed on the observations (ECA&D) and simulations from the non-separable model.

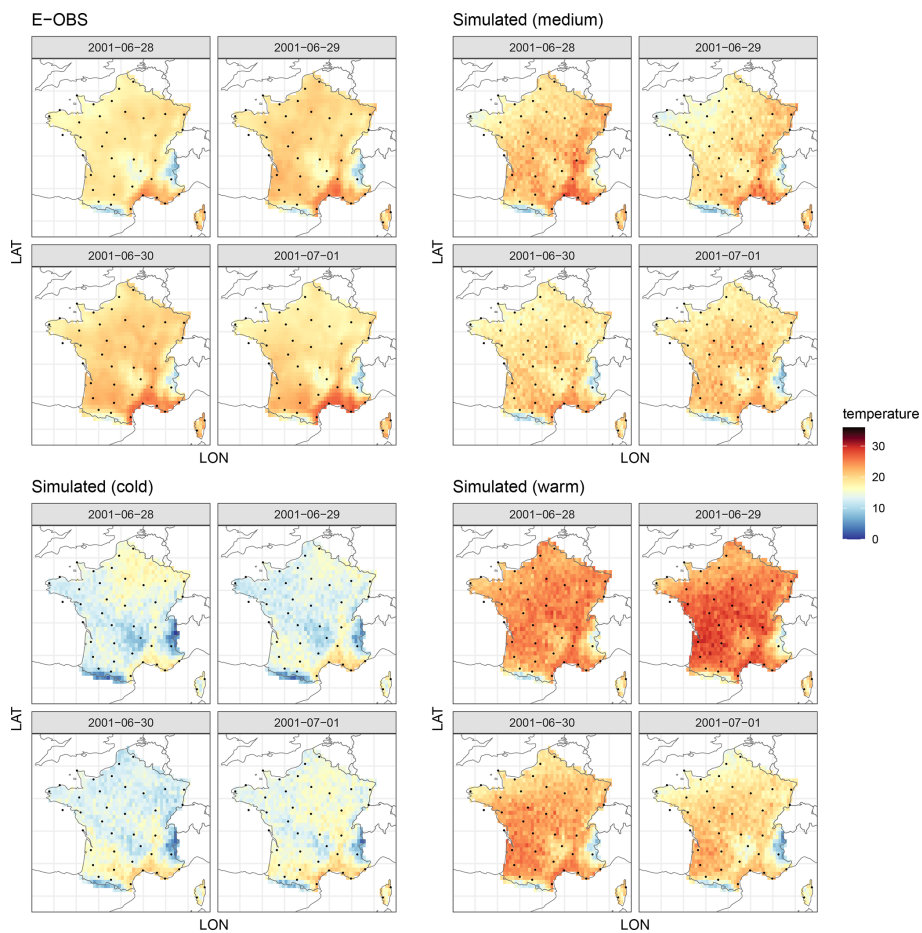


Figure 10. From left to right: 4 d in E-OBS, an example of the same simulated 4 d, coldest and warmest simulation for the same days.

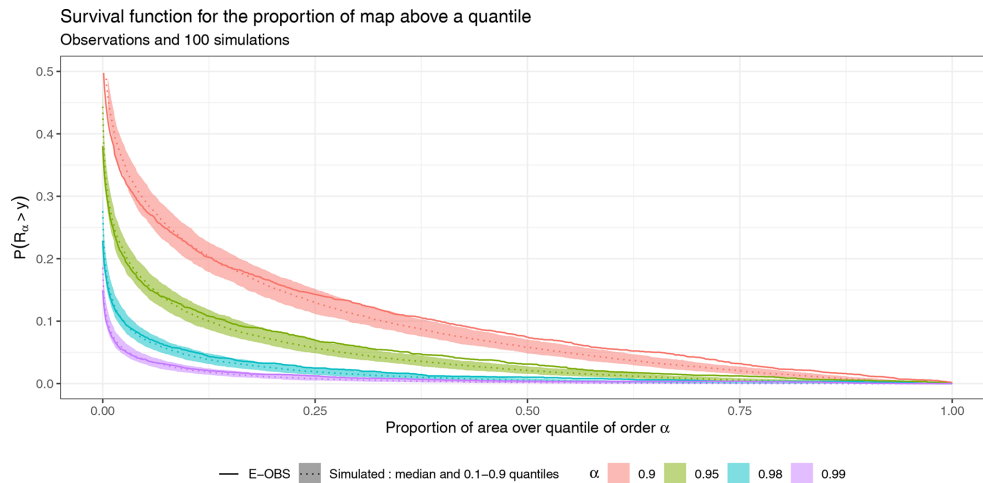


Figure 11. Survival function $\mathbb{P}(R_\alpha(t) \geq y)$ for several quantiles α , computed on E-OBS and gridded simulations.

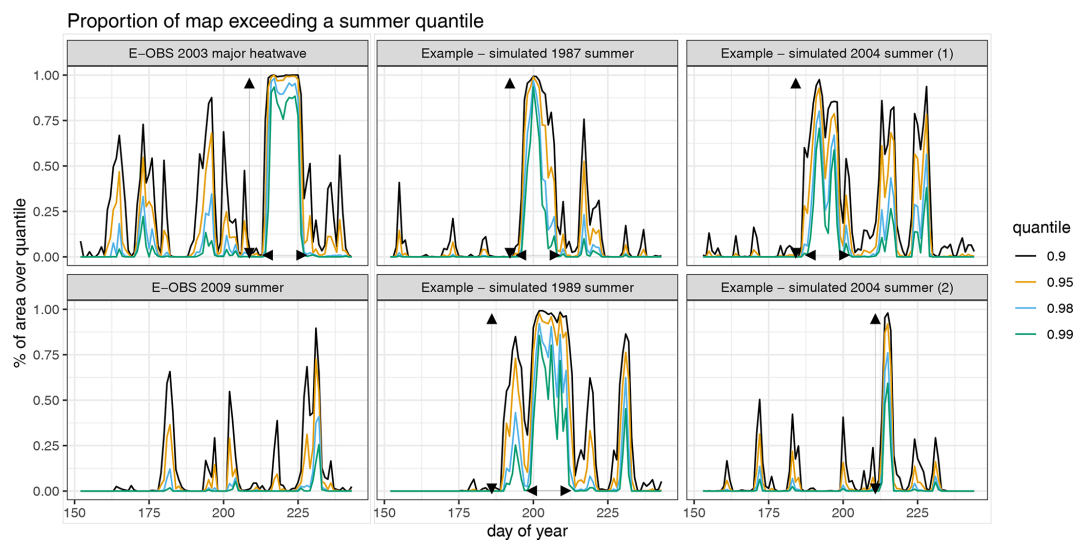


Figure 12. Spatial coverage of extreme heat over time, $R_\alpha(t)$, for several quantiles α , across two observed summers and four simulated ones.

Using the E-OBS observational database as a reference, we analyze the model's ability to replicate heat and cold events, defined by the exceedance of specific temperature quantiles over large spatial areas. The model captures the overall behavior of temperature extremes, including heat-wave sequences similar to the notable 2003 European heat-wave. However, while it successfully reproduces the frequency and duration of moderate heat episodes, the model tends to overestimate the occurrence of heat events in smaller regions and underestimates those covering larger areas. For cold episodes, the model generates realistic short-term events but struggles to reproduce longer events, particularly those covering large spatial extents.

4.3.2 Projections for 2016–2023

In this section, we demonstrate how the complete model can be utilised for near future projections. The following reasonable assumptions are made:

- The seasonality in mean and variance will remain the same in the near future. This is because they are estimated mainly as the result of the incoming solar radiation linked to earth rotation. The temperature change over time is included in the trend part which is removed first.
- The spatio-temporal dependence in the residuals Z will remain unchanged in the future, as far as only the near future is considered.

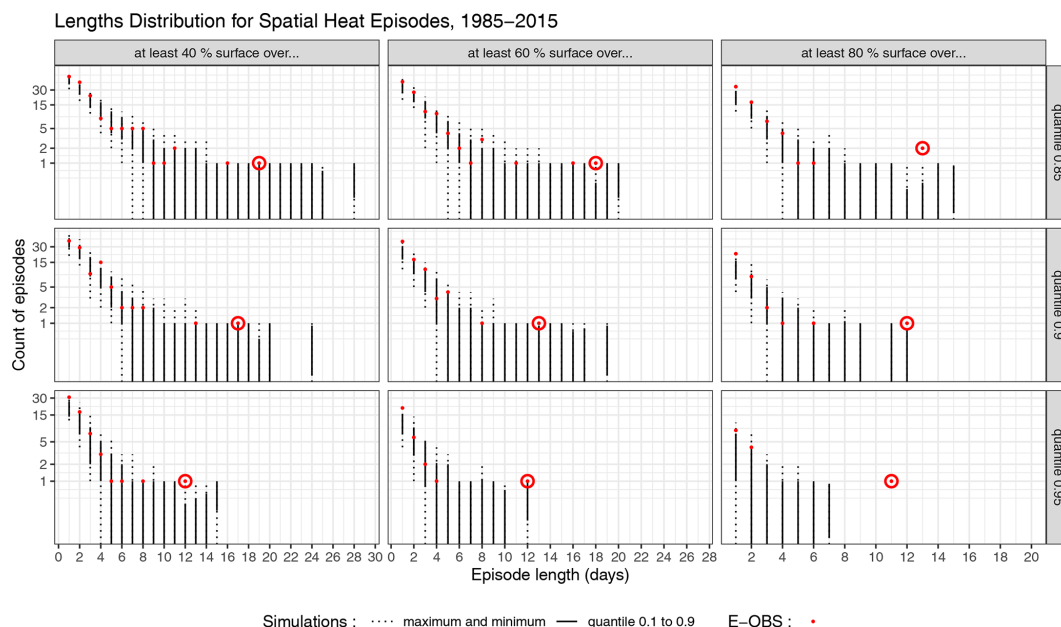


Figure 13. Length distribution of heat events for the observations and for 100 simulations. From left to right : spatial size of event (cumulative % of surface over the quantile threshold), from top to bottom: 0.85, 0.9, and 0.95 quantile. For each plot, the x -axis represents the episode length and the y -axis holds the number of events of this length in one 1985–2015 series. Log-scale on the y -axis. The dotted bars indicate the whole range of values found in the simulations, while the full bars represent the 0.1–0.9 quantile range of simulated values. The red dots indicate the E-OBS values, the circled dots represent the 2003 heatwave.

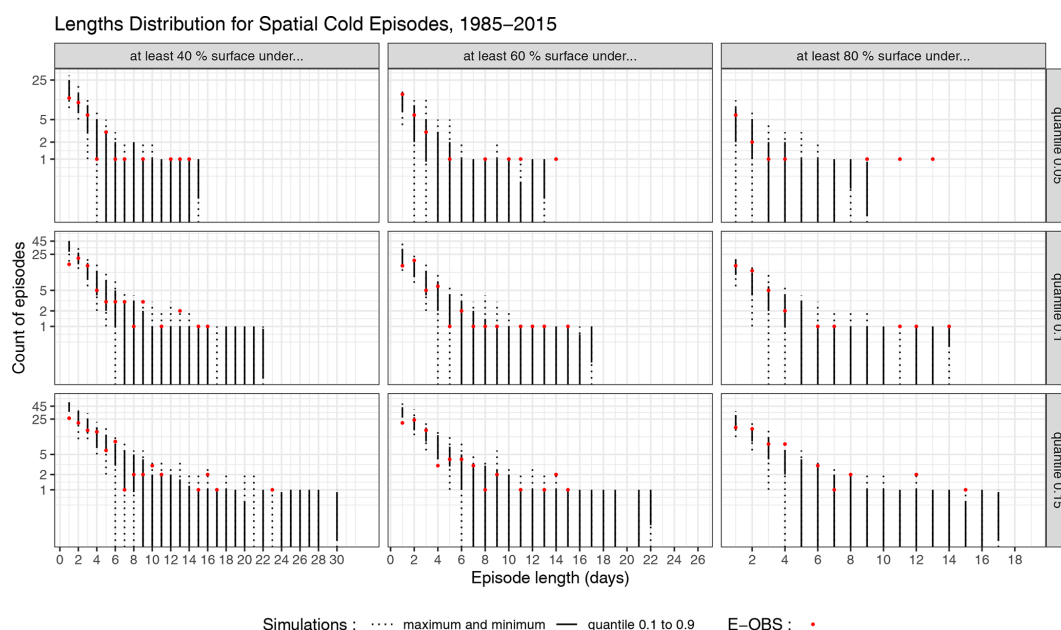


Figure 14. Distribution of the length of cold events in the observations and in 100 simulations. From left to right: spatial size of event (cumulative % of surface over the quantile threshold), from top to bottom: considered quantile. For each plot, the x -axis represent the episode length and the y -axis holds the number of events of this length in one 1985–2015 series. Log-scale on the y -axis. The dotted bars indicate the whole range of values found in the simulations, while the full bars represent the 0.1–0.9 quantile range of simulated values. The red dots indicate the E-OBS values.

- Future trends in mean and variance have been derived here from the observations, but can be obtained from climate projections using various scenarios if farther future periods are targeted.
- Should the projections concern farther future periods, the seasonality in mean and variance and the spatio-temporal structure of the residuals may have to be fitted again on climate projections.

The stochastic model is calibrated using data from 1985 to 2015. To validate the model's projection capabilities, we apply it to the period from 2016 to 2022. Trends in mean and variance for this future period can be derived using the same decomposition method, but with updated data from 1985 to 2022, while retaining the seasonal patterns from the earlier period.

We generated 100 simulations based on the updated trends and previous seasonality and stochastic models. The resulting mean temperatures averaged across the France grid are illustrated in Fig. 15 by their medians and 95 % intervals, and compared with the E-OBS values. The coverage rate of the interval is 92.21 %, which is close to the expected 95 %, with a small underestimation of the real variability. This demonstrates the model's capability to accurately project future temperatures, provided that the trends are known. This shows that future climate projected trends can be plugged into the model if farther future is considered.

The year 2022 experienced many important heatwaves. In Fig. 16, the distribution of the length of spatial heat episodes are represented for different spatial thresholds. The quantiles used to define the “hot days” are the 1985–2015 quantiles.

Most observed episodes fall within the range of the simulations. For all considered quantiles and spatial extents, the model successfully generates heat episodes that are longer than those observed. This suggests that the model can account for the possibility of longer heat waves than those recorded.

5 Conclusions

The spatio-temporal stochastic weather generator presented in this study and applied to French station data is a tool to generate realistic temperature series across a large area. By leveraging spatial statistics, the model effectively captures spatial-temporal correlations, enabling the simulation of plausible large-scale temperature events across the study area. This capability is particularly valuable for applications requiring temperature modeling across various locations, whether data is available for specific points, and even when those points are spread out spatially. To our knowledge, previous studies using spatial statistics have focused on more localised regions and have not addressed temperature trends with the same level of precision.

To study a changing climate, the trend components of the model can be adjusted based on climate simulation outputs

and various scenarios. By either keeping the seasonal cycles and spatial covariance structure stationary, or retraining them based on corrected model output, this approach allows for generating numerous simulations corresponding to different scenarios or specific areas of interest. This method significantly reduces computation time compared to physical models, thereby complementing physical simulations and providing valuable insights into future climate conditions.

Further investigation into the marginal distributions may be necessary to improve the representation of extreme conditions, such as extreme heat events. For many applications, accurately sampling these unfavorable conditions is crucial. Some further investigation on the margins may also be needed. For instance, the skewed-exponential-power distribution used in Evin et al. (2019) could provide a better fit for the temperature distribution, especially in the tails.

In this paper, while the Gaussian model was used, and although it is known to lack extremal dependence, the spatial extremes were found to be reasonably well represented in the simulations. However, there was a slight underestimation of the high-temperature extremal pairwise correlations. Future research should explore different non-Gaussian dependence structures, as the Gaussian approach, while adequate for temperature alone, might not be suitable for other variables or for modeling cross-correlations.

Additionally, there are plans to extend this model to cover the entire European continent. This expansion may require addressing issues of non-stationarity due to varying climates across different regions.

Many applications require modeling additional variables such as precipitation, wind speed, or solar radiation, which necessitates accounting for their spatial and temporal cross-correlations. While the Gneiting-Matérn model used in this study is well-suited for temperature, extending it to multiple variables presents challenges. Authors such as Bourotte et al. (2016) and Allard et al. (2022) have proposed flexible multivariate covariance functions that extend the univariate Gneiting-Matérn class. However, these models have not been tested over large areas like ours and may not be fully suitable. As for the marginals, they cannot be considered Gaussian. For precipitation amounts, while mixtures of Gamma or Exponential distributions are commonly used, they may not always adequately capture the right tail of the distribution. The Extended Pareto distribution, as proposed by Naveau et al. (2016), could be a valuable alternative, as it is designed to model both the bulk and the tail of the distribution effectively.

Another approach involves incorporating weather types into the model. Weather type modeling breaks down the dependence into different types of days, allowing for simpler models to be applied to each type (Allard et al., 2015; Stoner and Economou, 2020; Dawkins et al., 2022; Gobet et al., 2025). The sequence of weather types can be modeled using a Markov chain. Weather types can be learned from the data, with the simplest classification being dry/wet states, while more complex models might include hidden or unobserved

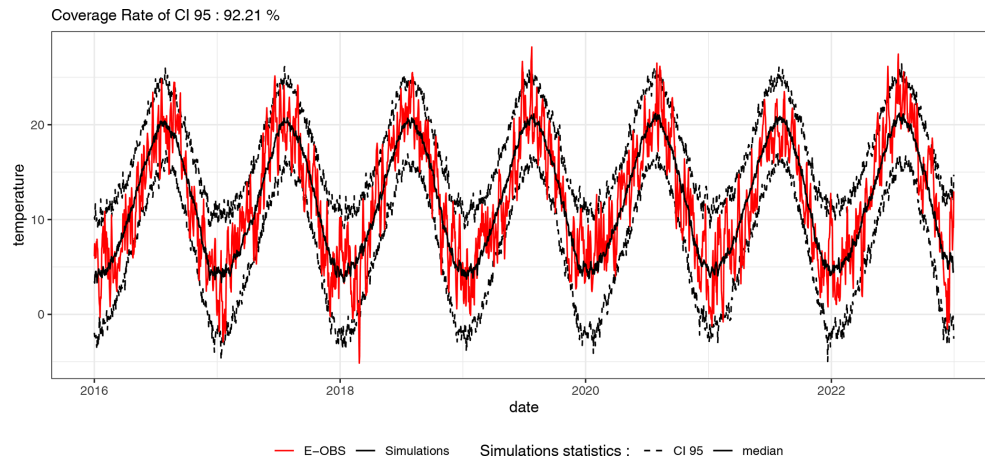


Figure 15. Mean temperature over the France grid for 2016–2022. The 100 simulations are obtained using a model fitted on the 1985–2015 period. Observed temperatures are shown in red. Simulations are represented with dotted black lines for the quantiles of order 0.025 and 0.975 of the simulated temperatures, and a solid black line for the median simulated temperature.

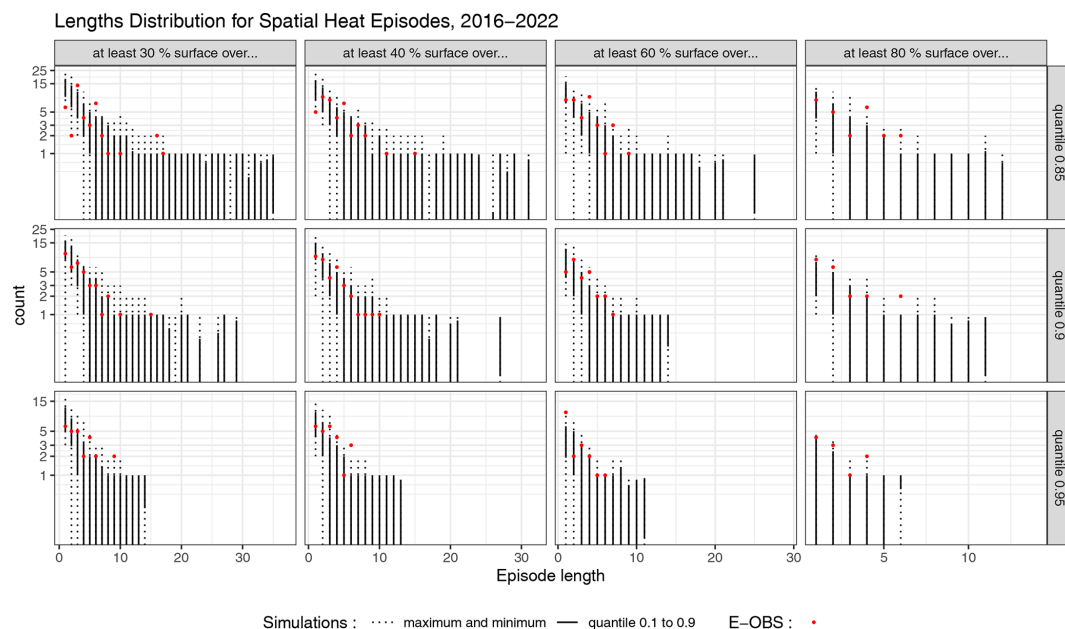


Figure 16. Distribution of the length of heat events in the observations (red dots) and in 100 simulations (black bars), for the period 2016–2022.

states. Future work will build on the extensive literature on multi-site and spatial weather type models to integrate the temperature generator from this study into a comprehensive multivariate generator.

Appendix A: Data locations

The Table A1 presents the stations used in this study: station identifier STAID, name STANAME, coordinates in the LON/LAT format and elevation. Additional columns hold the start and end date of the temperature file, and the column “missing” holds the proportion of missing data over the period 1985–2015.

Table A1. List of the French ECA&D stations.

STAID	STANAME	LAT	LON	ELEVATION	Start date (yyyy-mm-dd)	End date (yyyy-mm-dd)	Missing
32	BOURGES	47.05917	2.3594444	161	1945-03-01	2023-01-31	0.0000000
33	TOULOUSE-BLAGNAC	43.62083	1.3788889	151	1947-01-01	2023-01-31	0.0000000
34	BORDEAUX-MERIGNAC	44.83056	−0.6913889	47	1924-01-01	2023-01-31	0.0000000
36	PERPIGNAN	42.73694	2.8727778	42	1924-11-17	2023-01-31	0.0000000
37	LYON – ST EXUPERY	45.72639	5.0777778	235	1975-04-20	2023-01-31	0.0000000
39	MARIGNANE	43.43778	5.2158333	9	1921-01-01	2023-01-31	0.0000000
322	RENNES-ST JACQUES	48.06889	−1.7338889	36	1945-01-01	2023-01-31	0.0000000
323	STRASBOURG-ENTZHEIM	48.54917	7.6402778	150	1924-01-01	2023-01-31	0.0000000
434	BREST-GUIPAVAS	48.44417	−4.4116667	94	1945-01-01	2023-01-31	0.0001825
736	ABBEVILLE	50.13583	1.8338889	69	1922-05-01	2023-01-31	0.0000000
737	LILLE-LESQUIN	50.57000	3.0975000	47	1944-12-01	2023-01-31	0.0000000
738	CAEN-CARPIQUET	49.18000	−0.4558333	67	1945-10-01	2023-01-31	0.0000913
740	ALENCON	48.44556	0.1102778	143	1946-01-01	2023-01-31	0.0000000
742	NANTES-BOUGUENAI	47.15000	−1.6088889	26	1945-05-01	2023-01-31	0.0000000
745	DIJON-LONGVIC	47.26778	5.0880556	219	1921-05-29	2023-01-31	0.0000000
749	POITIERS – BIARD	46.59389	0.3141667	123	1921-05-01	2023-01-31	0.0000000
750	CLERMONT-FERRAND	45.78667	3.1491667	331	1923-01-01	2023-01-31	0.0000000
755	EMBRUN	44.56556	6.5019444	871	1947-06-24	2021-02-17	0.0000000
756	TARBES – OSSUN	43.18806	0.0000000	360	1946-01-01	2023-01-31	0.0000000
757	NICE	43.64889	7.2088889	2	1942-07-01	2020-12-31	0.0000000
758	BASTIA	42.54056	9.4852778	10	1947-01-01	2023-01-31	0.0000000
786	MONTÉLIMAR	44.58083	4.7330556	73	1920-08-17	2023-01-31	0.0000000
793	PTE DE LA HAGUE	49.72528	−1.9394444	6	1921-01-01	2023-01-31	0.0003651
2184	ROUEN – BOOS	49.38306	1.1816667	151	1968-03-01	2023-01-31	0.0000000
2190	TOURS	47.44417	0.7269444	108	1959-11-01	2023-01-31	0.0000000
2192	BASEL-MULHOUSE	47.61417	7.5100000	263	1947-01-01	2023-01-31	0.0000000
2195	LIMOGES – BELLEGARDE	45.86083	1.1750000	402	1973-01-01	2023-01-31	0.0000000
2196	LE PUY – LOUDES	45.07417	3.7638889	833	1984-01-01	2023-01-31	0.0000000
2199	GOURDON	44.74500	1.3966667	260	1961-02-01	2023-01-31	0.0000000
2200	MILLAU	44.11833	3.0183333	712	1964-04-01	2023-01-31	0.0000000
2203	MONT-DE-MARSAN	43.90944	−0.5002778	59	1945-01-01	2023-01-31	0.0000000
2205	ST-GIRONS	43.00528	1.1066667	414	1949-01-01	2023-01-31	0.0000000
2207	MONTPELLIER-AEROPORT	43.57667	3.9630556	2	1946-01-01	2023-01-31	0.0000000
2209	AJACCIO	41.91806	8.7927778	5	1949-10-01	2023-01-31	0.0000913
11243	TROYES-BARBEREY	48.32444	4.0197222	112	1975-05-01	2023-01-31	0.0000000
11244	PTE DE CHASSIRON	46.04667	−1.4116667	11	1917-01-01	2023-01-31	0.0000000
11245	PLOUMANACH	48.82556	−3.4730556	55	1947-08-04	2023-01-31	0.0004563
11246	NANCY-OCHEY	48.58083	5.9594444	336	1966-11-01	2023-01-31	0.0000000
11247	BELLE ILE – LE TALUT	47.29417	−3.2180556	34	1930-04-01	2023-01-31	0.0000000
11248	CAP CEPET	43.07917	5.9405556	115	1968-10-01	2020-02-07	0.0000913
11249	ORLY	48.71667	2.3841667	89	1921-03-01	2023-01-31	0.0000000

Appendix B: Separability

The choice of covariance structure reflects the non-separability in the data.

The separability property holds if the covariance function of the process under study satisfies

$$C(h, u) = C(h, 0)C(0, u) \quad (\text{B1})$$

where $C(h, 0) = C_S(h)$ is the covariance in space and $C(0, u) = C_T(u)$ is the covariance in time.

Tests have been proposed to study this assumption such as in Li et al. (2007), but they fail in practice with our temperature data because of the strong correlation in time. However, it can be empirically seen in Fig. B1 that for most considered $h, u \in [1, \dots, T] \times \mathcal{D}$, $C(h, u) \geq C(h, 0)C(0, u)$. This means that not only is the covariance of the data non-separable, but also that according to the definition of De Iaco and Posa (2013), the non-separability is said to be “positive”.

Conveniently, any model of the Gneiting class satisfy this property, which is why this class was chosen.

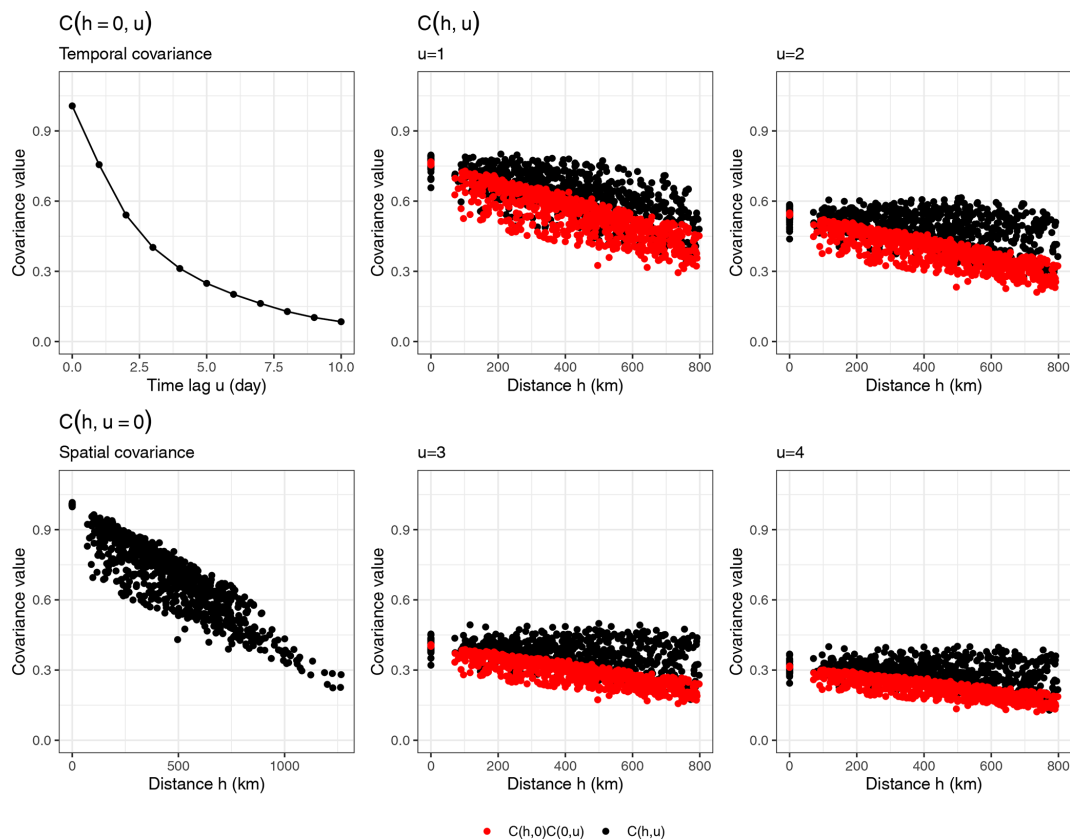


Figure B1. Comparison between the empirical spatial-temporal covariance function (black) and the product between the spatial and temporal covariance function (red). Each dot denotes the covariance between a pair of stations separated by a distance h .

Appendix C: Parameter estimates using station data

C1 Gridded trends

The Fig. C1 shows the interpolated temperature trend using the equivalent temperature that takes into account the elevation, compared to trends obtained from the much more complete E-OBS database. The addition of geographical information seems to correct quite well the values of the trend in mean.

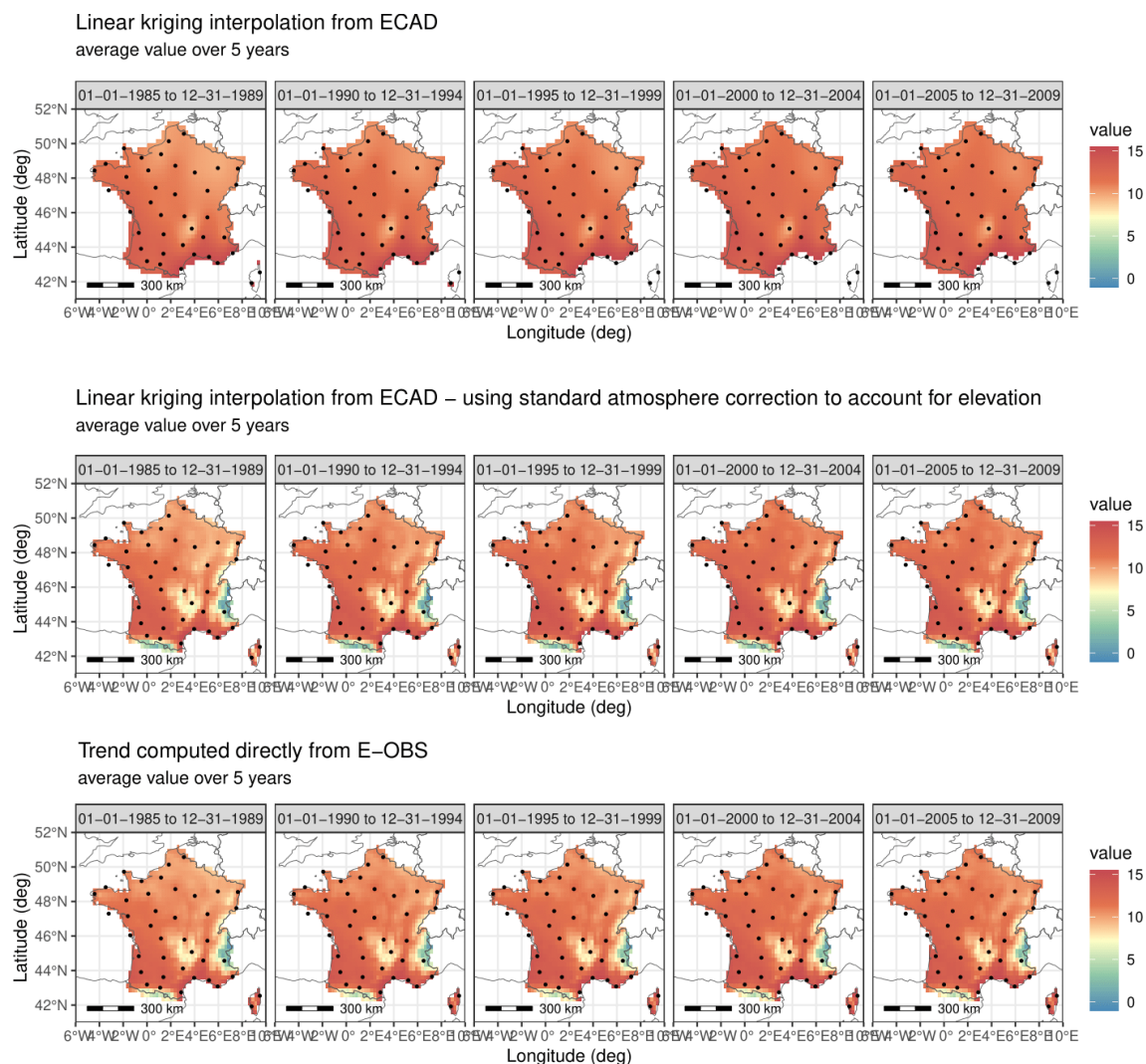


Figure C1. Trend in mean interpolated from site-by-site decomposition using ordinary kriging and altitude equivalent temperature, compared with trends obtained directly from the E-OBS database.

Figure C2 shows the average maps over 5 years of the standard deviation trend, using the kriging method. The stations are indicated by dots. The variance is stable in mean over 5 years.

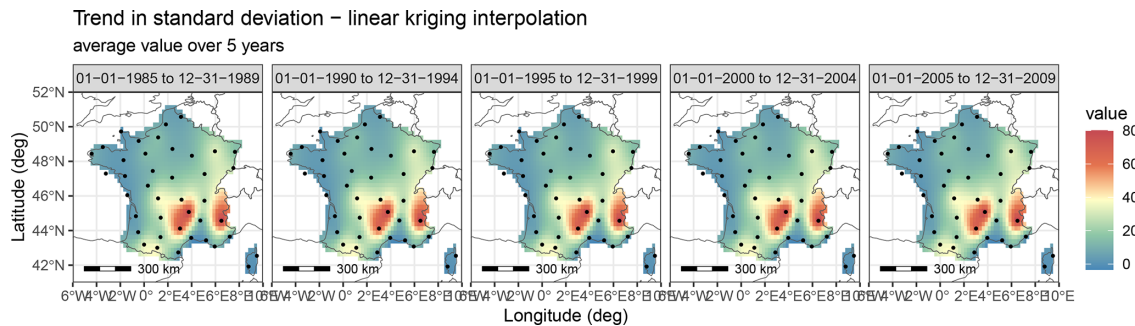


Figure C2. Trend in standard deviation, interpolated from site-by-site decomposition using ordinary kriging.

C2 Gridded seasonality coefficients

Kriging with a Gaussian variogram model was used to interpolate each seasonality coefficient. Resulting coefficients maps are given in Figs. C3, C4. Similar to what is observed when working with the trends, removing stations from the training set sometimes introduced significant error. It is mostly the case when several stations from the borders of the domain are missing, or in the case of the mountainous stations pointed out previously.

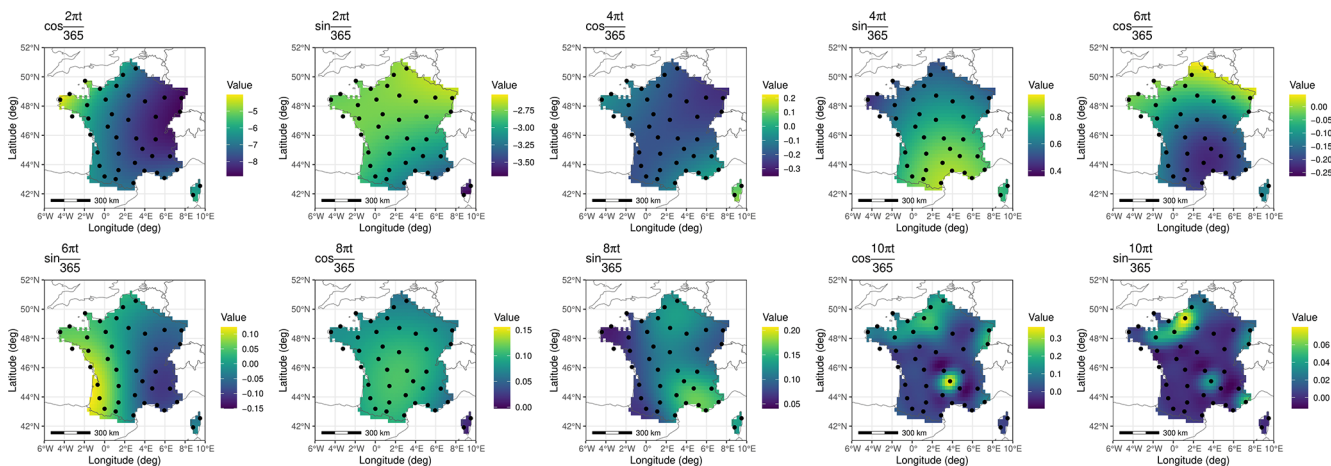


Figure C3. Seasonality coefficient for the mean, interpolated from site-by-site decomposition using ordinary kriging.

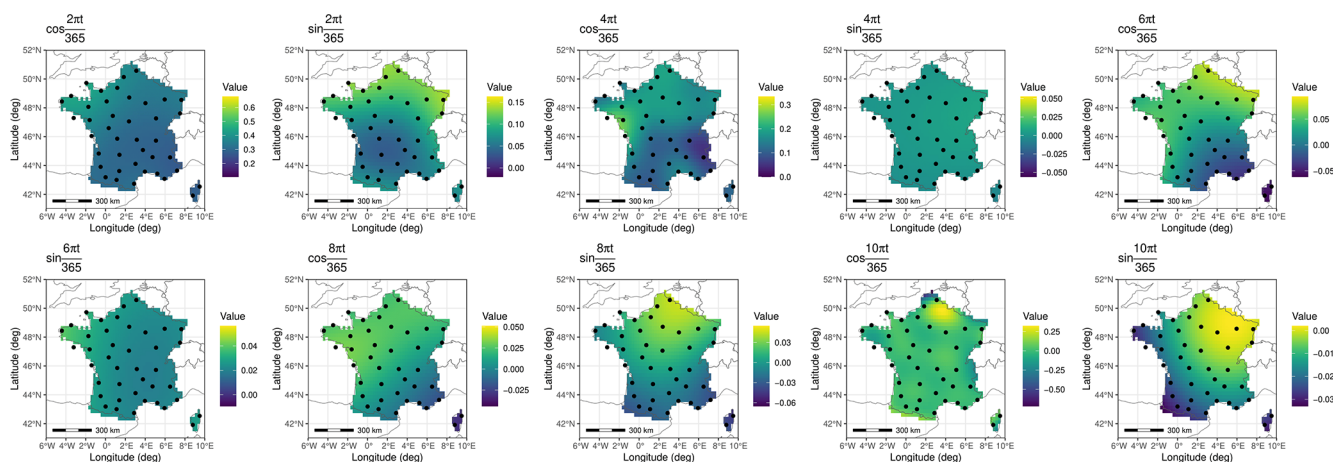


Figure C4. Seasonality coefficient for the standard deviation interpolated from site-by-site decomposition using ordinary kriging.

Code and data availability. Most of the code will be made available soon on GitHub. Both weather station data (ECA&D) and gridded data were downloaded from the European Climate Assessment & Dataset website (<https://www.ecad.eu/>, last access: March 2023; Klein Tank, 2002). In this paper, we used the March 2023 version of both datasets.

Author contributions. CC, LB and SP conceptualised the original idea. CC, LB, developed the mathematical methodology with domain knowledge from SP. CC performed the validation of the model and wrote the paper with feedback from all the authors. LB, SP and DM supervised the project. The original code was partly written in R by CC, with the estimation procedure implemented in Julia by CC and DM. Part of the gridded simulation algorithm was also implemented in Julia by DM.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors. Also, please note that this paper has not received English language copy-editing. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements. The authors thank Denis Allard for the useful feedback at various points in the work, as well as the implementation of the spectral simulation method.

Financial support. This research has been supported by the Association Nationale de la Recherche et de la Technologie (grant no. 2023/0061).

Review statement. This paper was edited by Soutir Bandyopadhyay and reviewed by three anonymous referees.

References

- Allard, D., Ailliot, P., Monbet, V., and Naveau, P.: Stochastic weather generators: an overview of weather type models, *Journal de la Société Française de Statistique*, 156, 101–113, 2015.
- Allard, D., Emery, X., Lacaux, C., and Lantuéjoul, C.: Simulating space-time random fields with nonseparable Gneiting-type covariance functions, *Statistics and Computing*, 30, 1479–1495, <https://doi.org/10.1007/s11222-020-09956-4>, 2020.
- Allard, D., Clarotto, L., and Emery, X.: Fully nonseparable Gneiting covariance functions for multivariate space-time data, *Spatial Statistics*, 52, 100706, <https://doi.org/10.1016/j.spasta.2022.100706>, 2022.
- Apipattanas, S., Podestá, G., Rajagopalan, B., and Katz, R. W.: A semiparametric multivariate and multisite weather generator, *Water Resources Research*, 43, <https://doi.org/10.1029/2006WR005714>, 2007.
- Baxevani, A. and Lennartsson, J.: A spatiotemporal precipitation generator based on a censored latent Gaussian field, *Water Resources Research*, 51, 4338–4358, <https://doi.org/10.1002/2014WR016455>, 2015.
- Beersma, J. and Buishand, T.: Multi-site simulation of daily precipitation and temperature conditional on the atmospheric circulation, *Climate Research*, 25, 121–133, <https://doi.org/10.3354/cr025121>, 2003.
- Bourotte, M., Allard, D., and Porcu, E.: A flexible class of non-separable cross-covariance functions for multivariate space-time data, *Spatial Statistics*, 18, 125–146, <https://doi.org/10.1016/j.spasta.2016.02.004>, 2016.

- Buishand, T. A. and Brandsma, T.: Multisite simulation of daily precipitation and temperature in the Rhine Basin by nearest-neighbor resampling, *Water Resources Research*, 37, 2761–2776, <https://doi.org/10.1029/2001WR000291>, 2001.
- Chandler, R. E.: Multisite, multivariate weather generation based on generalised linear models, *Environmental Modelling & Software*, 134, 104867, <https://doi.org/10.1016/j.envsoft.2020.104867>, 2020.
- Chilès, J. and Delfiner, P.: *Geostatistics: Modeling Spatial Uncertainty*, Wiley Series in Probability and Statistics, Wiley, 1st edn., ISBN 978-0-470-18315-1, <https://doi.org/10.1002/9781118136188>, 2012.
- Clark, M., Gangopadhyay, S., Hay, L., Rajagopalan, B., and Wilby, R.: The Schaake Shuffle: A Method for Reconstructing Space–Time Variability in Forecasted Precipitation and Temperature Fields, *Journal of Hydrometeorology*, 5, [https://doi.org/10.1175/1525-7541\(2004\)005<0243:TSSAMF>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0243:TSSAMF>2.0.CO;2), 2004.
- Cleveland, W. S.: Robust Locally Weighted Regression and Smoothing Scatterplots, *Journal of the American Statistical Association*, 74, 829–836, <https://doi.org/10.1080/01621459.1979.10481038>, 1979.
- Coles, S.: *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics, Springer London, London, ISBN 978-1-84996-874-4, <https://doi.org/10.1007/978-1-4471-3675-0>, 2001.
- Cornes, R. C., Van Der Schrier, G., Van Den Besselaar, E. J. M., and Jones, P. D.: An Ensemble Version of the E-OBS Temperature and Precipitation Data Sets, *Journal of Geophysical Research: Atmospheres*, 123, 9391–9409, <https://doi.org/10.1029/2017JD028200>, 2018.
- Cressie, N. and Wikle, C. K.: *Spatio-Temporal Statistics*, John Wiley & Sons, ISBN 978-0-471-69274-4, 2015.
- Cressie, N. A. C.: *Statistics for spatial data*, Wiley series in probability and mathematical statistics, Wiley, New York, ISBN 978-0-471-84336-8, 1991.
- Dawkins, L. C., Osborne, J. M., Economou, T., Darch, G. J. C., and Stoner, O. R.: The Advanced Meteorology Explorer: a novel stochastic, gridded daily rainfall generator, *Journal of Hydrology*, 607, 127478, <https://doi.org/10.1016/j.jhydrol.2022.127478>, 2022.
- De Iaco, S. and Posa, D.: Positive and negative non-separability for space–time covariance models, *Journal of Statistical Planning and Inference*, 143, 378–391, <https://doi.org/10.1016/j.jspi.2012.07.006>, 2013.
- De Iaco, S., Posa, D., Cappello, C., and Maggio, S.: Isotropy, symmetry, separability and strict positive definiteness for covariance functions: A critical review, *Spatial Statistics*, 29, 89–108, <https://doi.org/10.1016/j.spasta.2018.09.003>, 2019.
- Dubrovsky, M., Huth, R., Dabhi, H., and Rotach, M. W.: Parametric gridded weather generator for use in present and future climates: focus on spatial temperature characteristics, *Theoretical and Applied Climatology*, 139, 1031–1044, <https://doi.org/10.1007/s00704-019-03027-z>, 2020.
- Evin, G., Favre, A.-C., and Hingray, B.: Stochastic generators of multi-site daily temperature: comparison of performances in various applications, *Theoretical and Applied Climatology*, 135, 811–824, <https://doi.org/10.1007/s00704-018-2404-x>, 2019.
- Furrer, E. and Katz, R.: Generalized linear modeling approach to stochastic weather generators, *Climate Research*, 34, 129–144, <https://doi.org/10.3354/cr034129>, 2007.
- Geoga, C. J., Marin, O., Schanen, M., and Stein, M. L.: Fitting Matérn Smoothness Parameters Using Automatic Differentiation, 33, 48, <https://doi.org/10.1007/s11222-022-10127-w>, 2022.
- Gneiting, T.: Nonseparable, Stationary Covariance Functions for Space–Time Data, *Journal of the American Statistical Association*, 97, 590–600, <https://doi.org/10.1198/016214502760047113>, 2002.
- Gneiting, T., Kleiber, W., and Schlather, M.: Matérn Cross-Covariance Functions for Multivariate Random Fields, *Journal of the American Statistical Association*, 105, 1167–1177, <https://doi.org/10.1198/jasa.2010.tm09420>, 2010.
- Gobet, E., Métivier, D., and Parey, S.: Interpretable seasonal multisite hidden Markov model for stochastic rain generation in France, *Adv. Stat. Clim. Meteorol. Oceanogr.*, 11, 159–201, <https://doi.org/10.5194/ascmo-11-159-2025>, 2025.
- Hoang, T. T. H.: Modeling of non-stationary, non-linear time series. Application to the definition of trends on mean, variability and extremes of air temperatures in Europe, PhD thesis, Université Paris Sud – Paris XI, <https://theses.hal.science/tel-00531549> (last access: 30 May 2025), 2010.
- Holsclaw, T., Greene, A. M., Robertson, A. W., and Smyth, P.: A Bayesian Hidden Markov Model of Daily Precipitation over South and East Asia, *Journal of Hydrometeorology*, <https://doi.org/10.1175/JHM-D-14-0142.1>, 2016.
- Intergovernmental Panel On Climate Change (IPCC): *Climate Change 2021 – The Physical Science Basis: Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, 1st edn., ISBN 978-1-009-15789-6, <https://doi.org/10.1017/9781009157896>, 2023.
- International Organization for Standardization (ISO): *Standard Atmosphere*, ISO 2533:1975, <https://www.iso.org/standard/7472.html> (last access: 4 April 2024), 1975.
- Ji, H. K., Mirzaei, M., Lai, S. H., Dehghani, A., and Dehghani, A.: Implementing generative adversarial network (GAN) as a data-driven multi-site stochastic weather generator for flood frequency estimation, *Environmental Modelling & Software*, 172, 105896, <https://doi.org/10.1016/j.envsoft.2023.105896>, 2024.
- Kleiber, W., Katz, R. W., and Rajagopalan, B.: Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes: PRECIPITATION SIMULATION USING GAUSSIAN PROCESSES, *Water Resources Research*, 48, <https://doi.org/10.1029/2011WR011105>, 2012.
- Klein Tank, A. A. C.: Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment, *Int. J. Climatol.*, 22, 1441–1453, 2002.
- Kroiz, G. C., Majumder, R., Gobbert, M. K., Neerchal, N. K., Markert, K., and Mehta, A.: Daily Precipitation Generation using a Hidden Markov Model with Correlated Emissions for the Potomac River Basin, *PAMM*, 20, e202000117, <https://doi.org/10.1002/pamm.202000117>, 2021.
- Li, B., Genton, M. G., and Sherman, M.: A Nonparametric Assessment of Properties of Space–Time Covariance Functions, *Journal of the American Statistical Association*, 102, 736–744, <https://doi.org/10.1198/016214507000000202>, 2007.

- Li, X. and Babovic, V.: Multi-site multivariate downscaling of global climate model outputs: an integrated framework combining quantile mapping, stochastic weather generator and Empirical Copula approaches, *Climate Dynamics*, 52, 5775–5799, <https://doi.org/10.1007/s00382-018-4480-0>, 2019.
- Li, Z.: A new framework for multi-site weather generator: a two-stage model combining a parametric method with a distribution-free shuffle procedure, *Climate Dynamics*, 43, 657–669, <https://doi.org/10.1007/s00382-013-1979-2>, 2014.
- Li, Z., Li, J. J., and Shi, X. P.: A Two-Stage Multisite and Multivariate Weather Generator, *Journal of Environmental Informatics*, <https://doi.org/10.3808/jei.201900424>, 2019.
- Lindsay, B. G.: Composite likelihood methods, in: *Contemporary Mathematics*, edited by: Prabhu, N. U., Vol. 80, 221–239, American Mathematical Society, Providence, Rhode Island, ISBN 978-0-8218-5087-9, <https://doi.org/10.1090/conm/080/999014>, 1988.
- Miloshevich, G., Cozian, B., Abry, P., Borgnat, P., and Bouchet, F.: Probabilistic forecasts of extreme heatwaves using convolutional neural networks in a regime of lack of data, *Physical Review Fluids*, 8, 040501, <https://doi.org/10.1103/PhysRevFluids.8.040501>, 2023.
- Naveau, P., Huser, R., Ribereau, P., and Hannart, A.: Modeling jointly low, moderate, and heavy rainfall intensities without a threshold selection, *Water Resources Research*, 52, 2753–2769, <https://doi.org/10.1002/2015WR018552>, 2016.
- Porcu, E., Furrer, R., and Nychka, D.: 30 Years of space–time covariance functions, *WIREs Computational Statistics*, 13, e1512, <https://doi.org/10.1002/wics.1512>, 2021.
- Prairie, J. R., Rajagopalan, B., Fulp, T. J., and Zagana, E. A.: Modified K-NN Model for Stochastic Streamflow Simulation, *Journal of Hydrologic Engineering*, 11, 371–378, [https://doi.org/10.1061/\(ASCE\)1084-0699\(2006\)11:4\(371\)](https://doi.org/10.1061/(ASCE)1084-0699(2006)11:4(371)), 2006.
- Rajagopalan, B. and Lall, U.: A k-nearest-neighbor simulator for daily precipitation and other weather variables, *Water Resources Research*, 35, 3089–3101, <https://doi.org/10.1029/1999WR900028>, 1999.
- Richardson, C. W.: Stochastic simulation of daily precipitation, temperature, and solar radiation, *Water Resources Research*, 17, 182–190, <https://doi.org/10.1029/WR017i001p00182>, 1981.
- Sha, J., Chen, X., Chang, Y., Zhang, M., and Li, X.: A spatial weather generator based on conditional deep convolution generative adversarial nets (cDCGAN), *Climate Dynamics*, 62, 1275–1290, <https://doi.org/10.1007/s00382-023-06971-9>, 2024.
- Sparks, N. J., Hardwick, S. R., Schmid, M., and Toumi, R.: IMAGE: a multivariate multi-site stochastic weather generator for European weather and climate, *Stochastic Environmental Research and Risk Assessment*, 32, 771–784, <https://doi.org/10.1007/s00477-017-1433-9>, 2018.
- Stoner, O. and Economou, T.: A coupled hidden markov model for daily rainfall at multiple sites, in: *Proceedings of the 35th International Workshop on Statistical Modelling*, 20–24 July 2020, Bilbao, Basque Country, Spain, ISBN 978-84-1319-267-3, 228–232, <https://addi.ehu.eus/handle/10810/45863> (last access: 9 September 2024), 2020.
- Touron, A.: Modélisation multivariée de variables météorologiques, PhD thesis, Université Paris-Saclay, <https://theses.hal.science/tel-02319170> (last access: 8 June 2023), 2019.
- Varin, C., Reid, N., and Firth, D.: An Overview of Composite Likelihood Methods, *Statistica Sinica*, 5–42, <http://www.jstor.org/stable/24309261> (last access: 29 September 2025), 2011.
- Verdin, A., Rajagopalan, B., Kleiber, W., and Katz, R. W.: Coupled stochastic weather generation using spatial and generalized linear models, *Stochastic Environmental Research and Risk Assessment*, 29, 347–356, <https://doi.org/10.1007/s00477-014-0911-6>, 2015.
- Verdin, A., Rajagopalan, B., Kleiber, W., Podestá, G., and Bert, F.: BayGEN: A Bayesian Space-Time Stochastic Weather Generator, *Water Resources Research*, 55, 2900–2915, <https://doi.org/10.1029/2017WR022473>, 2019.
- Wang, Y., Xie, J., You, Y., Wang, Y., Xu, Y.-P., and Guo, Y.-X.: A new multi-site multi-variable stochastic model with inter-site and inter-variable correlations, low frequency attributes and stochasticity: A case study in the lower Yellow River basin, *Journal of Hydrology*, 599, 126365, <https://doi.org/10.1016/j.jhydrol.2021.126365>, 2021.
- Wilks, D. S.: Multisite generalization of a daily stochastic precipitation generation model, *Journal of Hydrology*, 210, 178–191, [https://doi.org/10.1016/S0022-1694\(98\)00186-3](https://doi.org/10.1016/S0022-1694(98)00186-3), 1998.
- Wilks, D. S.: A gridded multisite weather generator and synchronization to observed weather data, *Water Resources Research*, 45, <https://doi.org/10.1029/2009WR007902>, 2009.
- Yiou, P.: AnaWEGE: a weather generator based on analogues of atmospheric circulation, *Geosci. Model Dev.*, 7, 531–543, <https://doi.org/10.5194/gmd-7-531-2014>, 2014.