

MSTWeatherGen: a multivariate Stochastic Weather Generator

Denis Allard, with Ahmed Boualam, Lionel Benoit Jean-François Rey, Saïd Obakrim

Biostatistique et processus Spatiaux (BioSP), MathNum, INRAE
Avignon, France

Geolearning Scientific day – April 8, 2026



Non extreme climate events, but with huge impact



- ▶ **2016, 2024:** around 25% yield loss on wheat in the "Breadbasket" (France)
- ▶ Succession of adverse conditions at each stage of the wheat season: sowing, growing and harvesting
- ▶ **Stochastic simulation as a tool to address this complexity**

SWGs

Adapted from Yiu (2024)

SWG are tools that generate random series of meteorological variables such as precipitation, temperature, wind speed, etc., with statistics similar to those of recorded data:

- ▶ Mean, variance, quantiles, skewness, extremes
 - ▶ Covariance (dependence) between variables
 - ▶ Temporal dependence / coherence (persistence)
 - ▶ Spatial dependence / coherence
-
- ▶ Calibrated on recorded series
 - ▶ Computational efficiency \Rightarrow long series and/or large number of realizations

For what purpose?

Used in impact studies

Outputs of SWGs are used as alternative, plausible weather time series in process-based models

- ▶ Explore unmeasured sequences of a given climate
- ▶ Assessing complex, non linear, responses to climate in agro-ecological systems s
- ▶ Explore plant / ecosystem models as functions of climate variability
- ▶ Optimal decision under uncertainty: simulate up to year $t + k$, optimize decision
- ▶ Also used with energy demand models, hydrological models, insurance models, ...

Some challenges that SWGs pose for spatial statistics

- ▶ Building models quantifying spatial, temporal and spatio-temporal variations ← **this talk**
- ▶ Doing stochastic simulations, both for the bulk and of the tail
- ▶ Building models and methods for multivariate, spatio-temporal extreme events
- ▶ Devising new approaches for assessing return levels of impactful compound events

General architecture

Two main assumptions

- ▶ Finite number of weather types over the region, $\mathbf{k} = 1, \dots, K$
- ▶ In each weather type k , the weather variables $Y_i(\mathbf{s}, t)$ are modeled as transformed latent Gaussian random fields $Z_i(\mathbf{s}, t)$

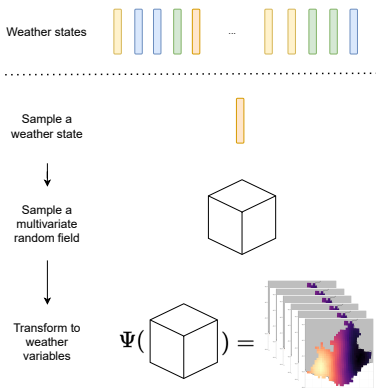
$$Y_i(\mathbf{s}, t) = \psi_{\mathbf{k}, i, \mathbf{s}}(Z_{\mathbf{k}, i}(\mathbf{s}, t)), \quad \mathbf{k} = \mathbf{X}(t)$$

- ▶ Weather states, $\mathbf{X}(t)$ are modeled as 1st order Markov chain with transition matrices $\pi(t)$
- ▶ We use empirical transformation for $\psi_{\mathbf{k}, i}$ with tail adjustments (Peterson and Cavanaugh, 2019)
- ▶ Multivariate spatio-temporal GPs for $\mathbf{Z} = (Z_i)_{i=1, p}$

Estimation step

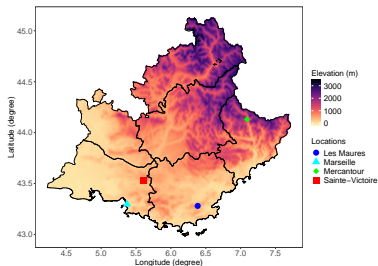
1. Find the weather types using a clustering algorithm
2. Estimate all 365 transition matrices $\pi(t)$
3. Estimate all transform functions $\psi_{k,i,s}$
4. In each weather type, estimate the parameters of the multivariate spatio-temporal covariance function of \mathbf{Z}

Simulation



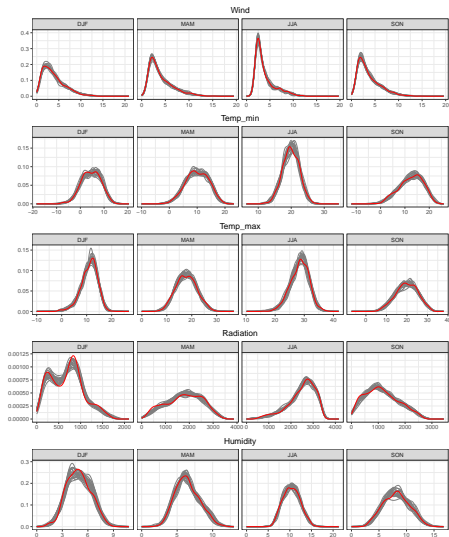
Our domain of interest

- ▶ Region of interest: PACA, highly non-stationary
- ▶ 6 daily variables: precipitation, humidity, radiation, wind, min and max temperature
- ▶ SAFRAN reanalysis data from 2012 to 2021
- ▶ $498 \text{ pixels} \times 6 \text{ variables} \times 3652 \text{ days}$
 $\approx 10.9 \text{ M data}$



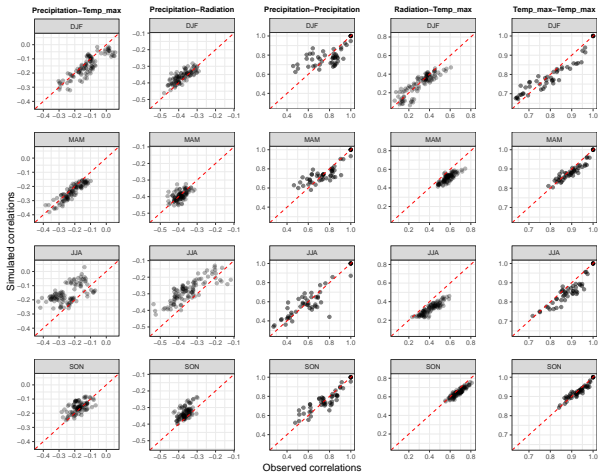
Marginals

All seasons, all continuous variables



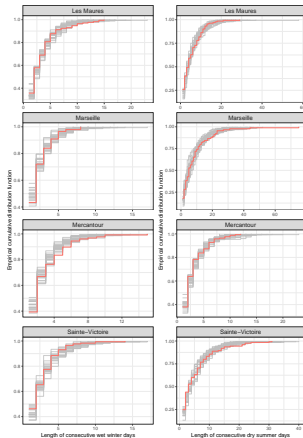
Correlations

Winter, 10 random locations



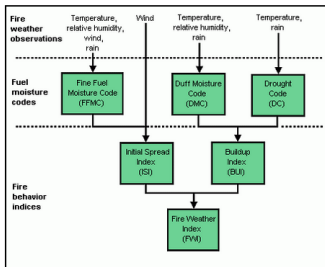
Empirical cumulative distribution function of spells

Wet winter spells, Dry summer spells



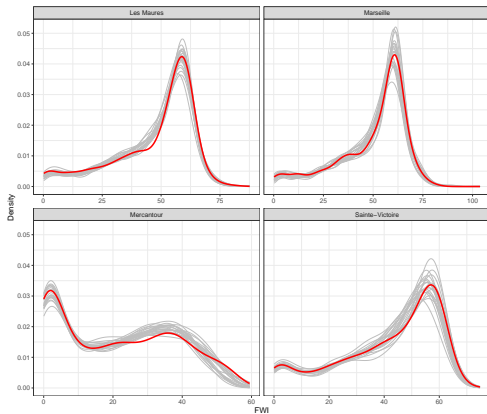
Fire Weather Index

- ▶ Developed and maintained by the Canadian Forest Service
- ▶ consists of six components that account for the effects of fuel moisture and weather conditions on fire behavior.
- ▶ Updated daily using FWI(day before) and Temp, Rain, Wind, RH
- ▶ Used in France for the alert system



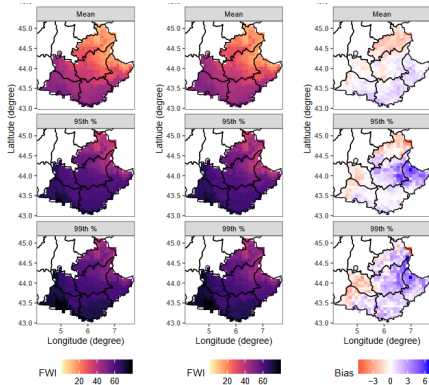
Fire Weather index

Summer density



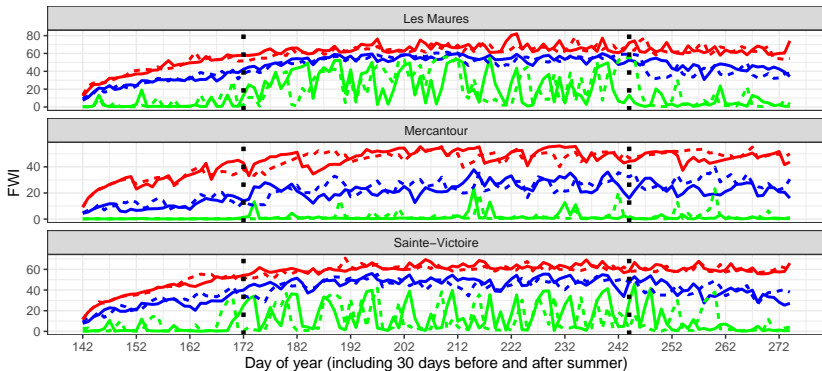
Fire Weather Index

Summer average



Fire Weather Index

Summer



Statistic — Mean — Min — Max Type — Observed — Simulated

Beyond MSTWeatherGen

MSTWeatherGen

✓ Obakrim S., Benoit L., Allard D. (2025) A multivariate and space-time stochastic weather generator using a latent Gaussian framework. *Stochastic Environmental Research and Risk Assessment* (In press).

✓ Obakrim S, Benoit L, Allard D, Rey J (2024). MSTWeatherGen: Multivariate Space-Time Weather Generator. R package

<https://sobakrim.github.io/MSTWeatherGen/index.html>

- ▶ CDD Ahmed Boualam (2 ans) à BioSP financé par la chaire Geolearning
- ▶ Première mission: consolider MSTWeatherGen et l'interface utilisateur
- ▶ Au-delà:
 - ▶ Intégrer les résultats obtenus dans la chaire
 - ▶ Passer à l'échelle sur des territoires plus grands