Stochastic weather generator and deep learning approach for predicting and sampling extreme European heatwaves

George Miloshevich, Dario Lucente, Pascal Yiou and Freddy Bouchet

LSCE, UMR 8212 CEA-CNRS-UVSQ, IPSL, Orme des Merisiers, Gif-sur-Yvette, 91191, France

Context and Goals

- Extreme events are rare but impactful
- Sampling heatwaves with GCMs costly
- Training on imbalanced datasets is hard
- We benchmark probabilistic prediction
- We consider dimensionality reduction for efficient analogs computation
- We generate surrogate extremes and validate them with a long control run

Heat Wave (HW) definition

- HW: $1-\overline{p} :=$ 95th percentile of time averaged running mean 2 meter temperature anomalies:

$$A_T(t) = \frac{1}{T}\int_t^{t+T} \frac{1}{|\mathcal{D}|}\int_{\mathcal{D}} \left(T_{2m} - \mathbb{E}\left(T_{2m}\right)\right) \,\mathrm{d}\vec{r}\mathrm{d}u \qquad (1)$$

Data: 8000 years of Plasim

- Intermediate complexity climate model
- configuration: fixed ocean, repeating SST

Return times formula

- Sort summer extremes $\{a_m\}_{1\leqslant m \leqslant M}$

$$r = \frac{1}{\log\left(1 - \mathrm{rank}(a)/M\right)}, \qquad (2)$$

SWG return times vs control

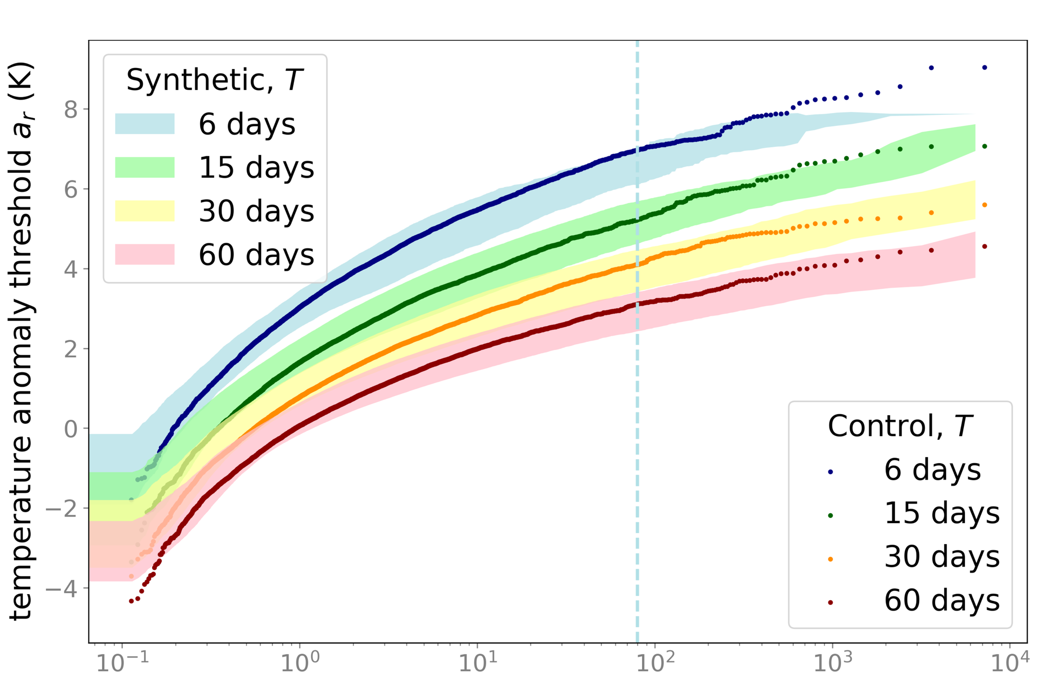


Figure: Return time plot: synthetic vs control

- Synthetic trajectories *(shaded lines)* are obtained from 80 years of a reference run
- Validation achieved comparing synthetic to 7200 years-long control run *(dots)*

Contact Information

- Website: georgemilosh.github.io
- Email: gmiloshe@lsce.ipsl.fr
- Scan for GitHub link (code repo):

Figure: Climate-Learning

Stochastic Weather Generator

- Markov chain → surrogate trajectory [3]

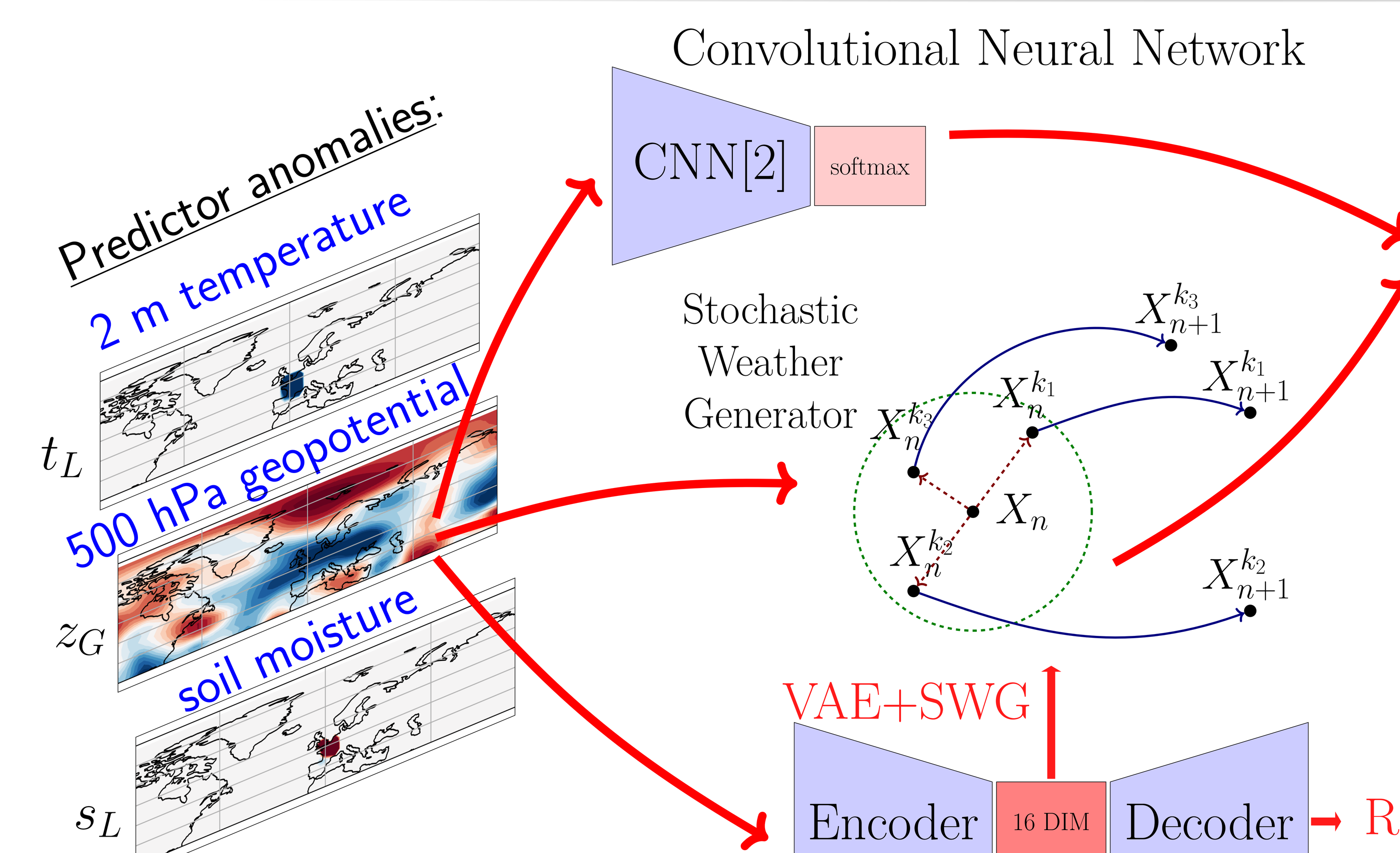
$$P(X(t+\delta t)|X(s)) \propto \begin{cases} 1, & \text{if } X(t) \sim X(s) \\ 0, & \text{otherwise} \end{cases} \qquad (3)$$

Euclidean distance analogs

Analogs sought using $X_{n_\star} = \underset{\{X_n\}}{\mathrm{argmin}}\left\{d\left(x, X_n\right)\right\}$ (4)

$$d^2 = \frac{\left(\Delta T_L^I\right)^2}{\sigma_T^2} + \frac{\left(\Delta S_L^I\right)^2}{\sigma_S^2} + \alpha \sum_{i=1}^{\dim(Z)_G} \frac{\left(\Delta Z_G^I\right)^2}{\sigma_Z^2 \dim(Z_G)} \qquad (5)$$

Data pipeline for learning probabilistic forecasting: CNN vs SWG



committor function: probability of an event

$$\hat{P}(A_T(t) > a | X = x, \tau) \qquad (6)$$

Normalized Logarithmic Score (NLS)

- Suitable for rare events:

$$\mathrm{NLS} := 1 - \frac{\mathbb{E}\left\{\sum_{k=0}^{1} y_k \log\left[\hat{P}_k(x,\tau)\right]\right\}}{\sum_i \overline{p}_i \log \overline{p}_i} \qquad (7)$$

Figure: Possible field inputs $X = \{t_M, z_G, s_M\}$ in stacked architecture which works better for CNN

Forecasting heatwaves, benchmarking: CNN, SWG, VAE+SWG

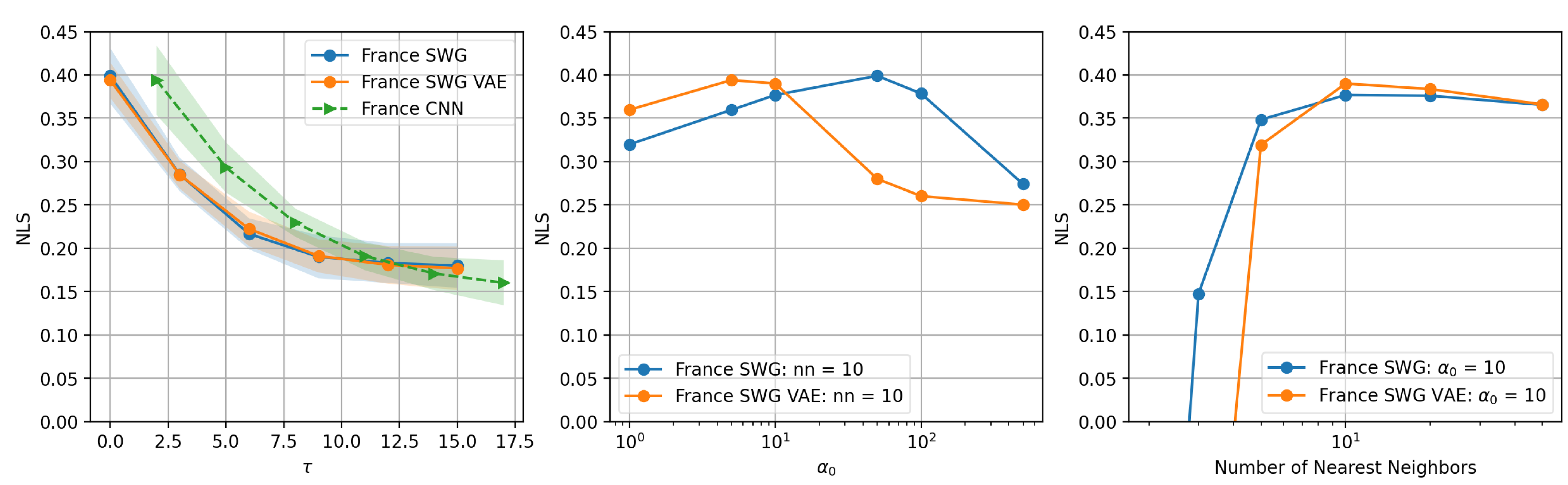
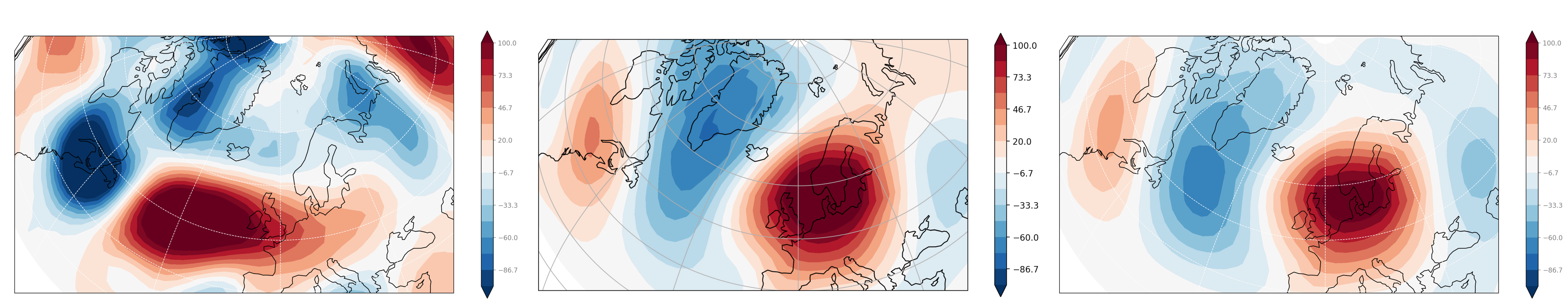


Figure: NLS benchmarks: (left panel) NLS vs τ (center panel) NLS vs α (right panel) NLS vs number of nearest neighbors

Key Takeaways

For *prediction* CNN performs better than SWG. SWG estimates *extreme return times and teleconnections*

Validating extreme surrogate 500 hPa geopotential teleconnection patterns



(a) Single event in the reference run (b) 7200 years control run (c) Composite from SWG run

Figure: Composite maps for $A(t) > 4.5K$ threshold France heatwaves at $\tau = 0$. The reference run has 80 years.

Conclusions and future work

- CNN [1] produces better probabilistic forecasts of heatwaves than SWG
- The prediction skill improves on order of hundreds of years [2]
- This lack of data [2] indicates the rare event algorithm could be useful
- Analogs are computed faster with dimensional reduction
- s_L provides long-term prediction skill, z_G - shorter term
- SWG samples well extreme return times and teleconnection patterns
- The need to re-train on ECMWF reanalysis (transfer learning)

References

[1] https://www.cnrs.fr/fr/changements-climatiques-une-meilleure-prediction-des-canicules-grace-lia.

[2] G. Miloshevich, B. Cozian, P. Abry, P. Borgnat, and F. Bouchet. *Phys. Rev. Fluids*, 8:040501, Apr. 2023.

[3] P. Yiou. *Geoscientific Model Development*, 7(2):531–543, 2014.