

European Congress of Actuaries
Paris, June 2026

**A WASSERSTEIN GAN-BASED CLIMATE
SCENARIO GENERATOR FOR RISK
MANAGEMENT AND INSURANCE: THE CASE OF
SOIL SUBSIDENCE**

Olivier LOPEZ

Scientific Director Detralytics/
Researcher CREST – ENSAE,
Paris

Antoine HERANVAL

Biostatistiques et Processus
Spatiaux (BioSP), INRAE,
Avignon

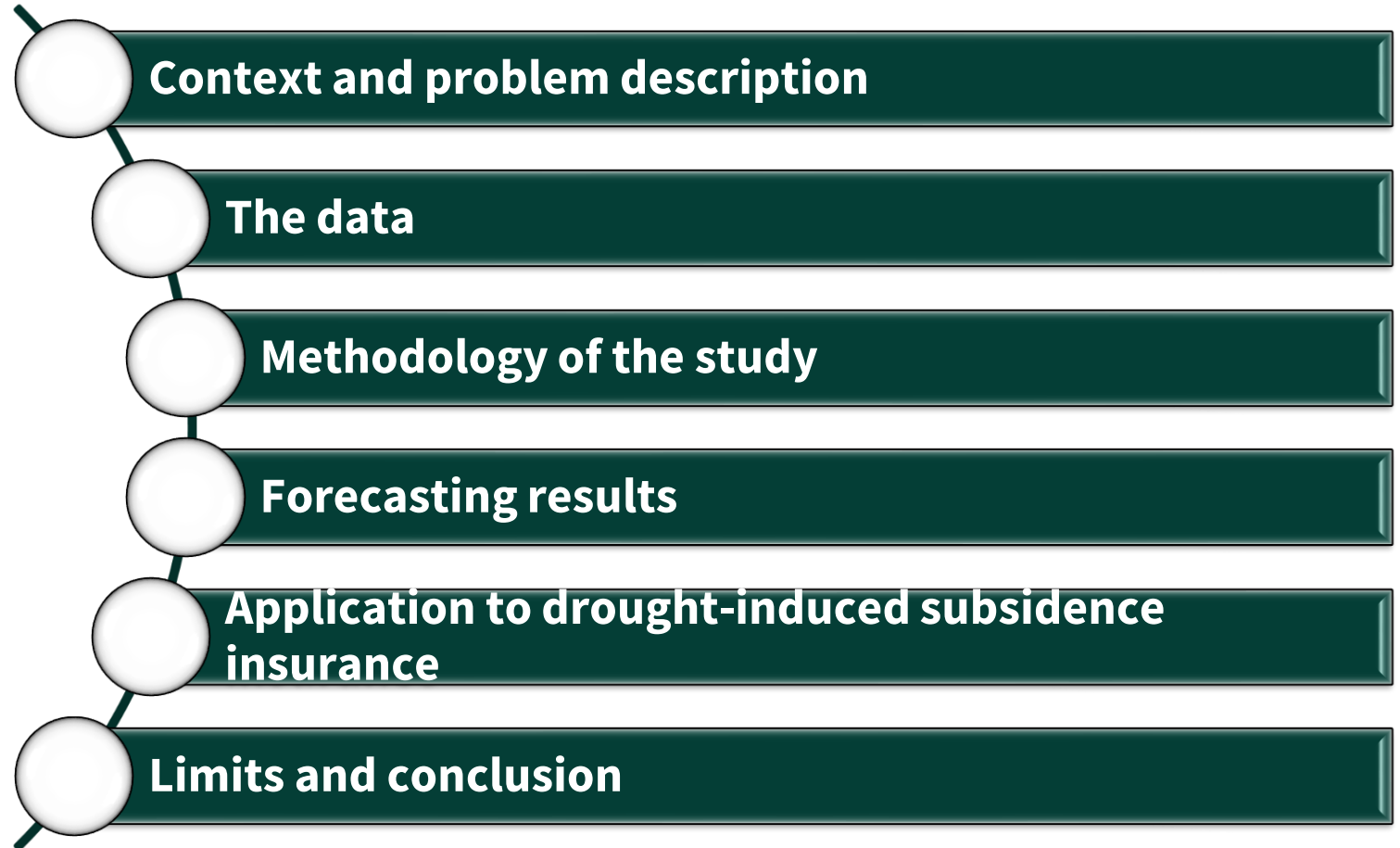
Daniel NKAMENI

TCP Consultant Detralytics/
PhD Candidate CREST –
ENSAE, Paris

Didier NGATCHA

Institut Louis Bachelier,
Paris

PLAN OF THE PRESENTATION

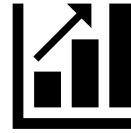


CONTEXT AND PROBLEM DESCRIPTION (1/7)

Why build a climate scenario generator?

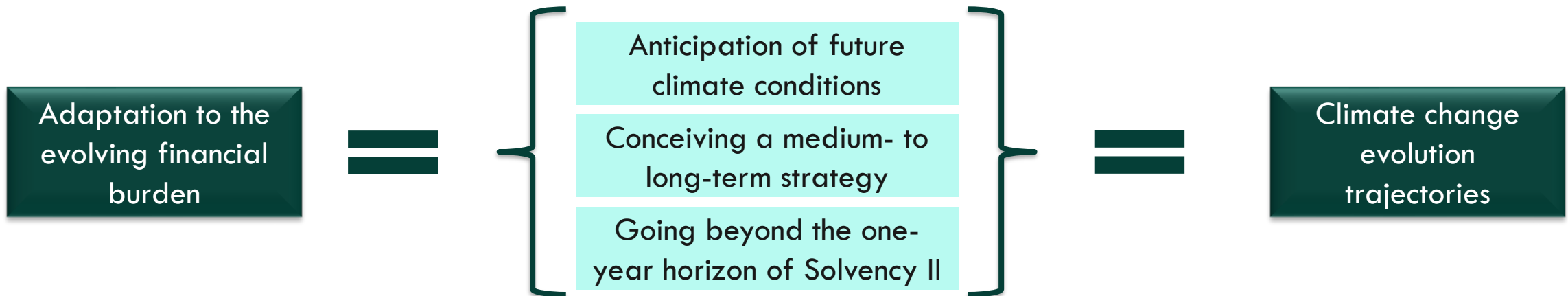
Average yearly cost of natural catastrophes according to the United Nations Office for Disaster Risk Reduction (2025):

70-80 billions USD between 1970 and 2000



180-200 billions between 2001 and 2020

Question: How can insurance against natural catastrophes cover a risk that constitutes an increasing financial burden?



The past is not always a reliable indicator of the future when it comes to natural catastrophes

CONTEXT AND PROBLEM DESCRIPTION (2/7)

The context of drought-induced subsidence

Projecting the medium- to long-term evolution of a climate insurance portfolio often relies on the ability to forecast climate indices or variables.

Drought-induced subsidence causes the shrinking and swelling of soils in areas where clay is highly present in the ground. Building foundations and structures are weakened by soil swelling due to rainfall, followed by shrinkage during drought episodes.



Average annual losses in the French insurance market increased from €400 million over the period 1989–2015 to €1 billion between 2016 and 2020, making subsidence one of the sector's major concerns.

The United Kingdom, Spain, and certain U.S. states such as Texas, California, and Colorado are also affected by this problem.



CONTEXT AND PROBLEM DESCRIPTION (3/7)

The context of drought-induced subsidence

In France, compensation for subsidence claims is governed by a weather index called: 📌

The Uniform Soil Wetness Index (SWI)



Developed by Météo-France, this index quantifies the intensity of a drought episode.



The evolution of this index is key to the future of protection against damage caused by geotechnical drought.

Projecting the distribution of the SWI can help answer questions such as:

- Should we modify the threshold used to characterize events to ensure that subsidence insurance remains sustainable?
- What level of preventive effort is required to contain the evolution of the risk?
- What coping strategies should be implemented to anticipate future drought episodes?
- Are current insurance products adapted to the evolution of this risk?

Note that our methodology could be extended to any situation in which an index describes the severity of a natural disaster.

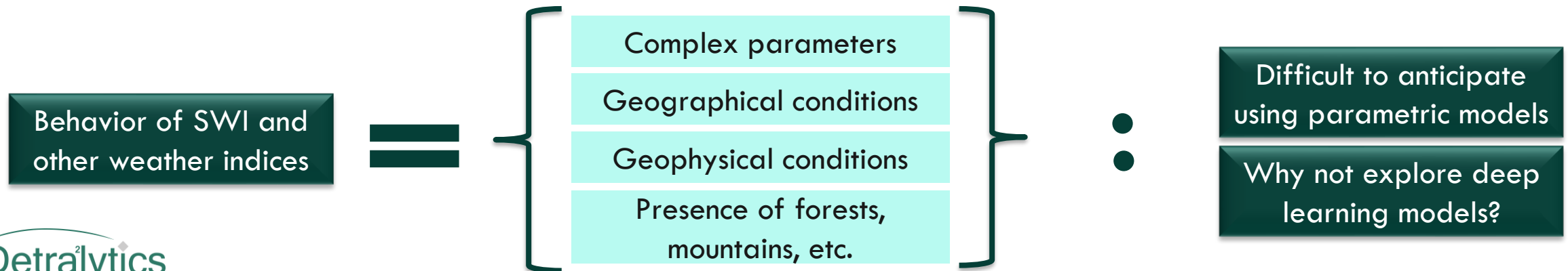
CONTEXT AND PROBLEM DESCRIPTION (4/7)

Our research idea

In our research project, we develop a generic methodology to simulate the evolution of the SWI using Generative Adversarial Networks (GANs).

Our approach is motivated by the desire:

- to obtain future distributions of the index of interest (the SWI)
- to derive a distribution of future financial damages caused by drought-induced subsidence
- to build a projection model that accounts for both the spatial and temporal dimensions of the index of interest
- to leverage existing climate projection models for variables that influence drought
- to perform fast simulations without incurring excessive complexity in representing the phenomena

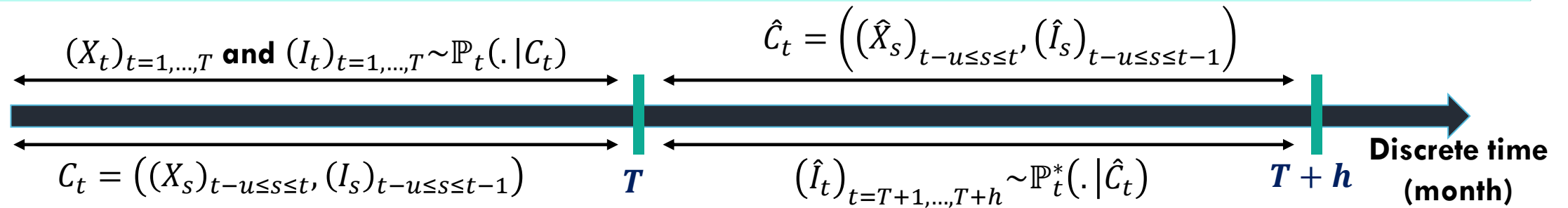


CONTEXT AND PROBLEM DESCRIPTION (5/7)

Problem description

Let $I_t \in \mathbb{R}^d$ be the index we aim to project, and $X_t \in \mathbb{R}^d$ the covariates that inform us about the state of the climate at time t .

This index is multivariate, as we observe its values at d spots, each spot of I_t corresponding to a fixed geographical position.



$\hat{X}_t$ represents anticipated values of covariates produced by high-level climate scenarios, such as **RCP scenarios**, or by **stochastic weather generators** like that of Obakrim et al. [2025].

The generator G of our GAN can be viewed as a function that transforms a random noise vector $Z \in \mathbb{R}^Z$ (drawn from a known distribution) and covariates $\hat{C}_t$ into a quantity in $\mathbb{R}^d$ such that:

$$\hat{I}_t = G(Z; \hat{C}_t) \in \mathbb{R}^d$$

CONTEXT AND PROBLEM DESCRIPTION (6/7)

Problem description

The generator G is constructed such that $\mathbb{P}_t^*(\cdot | \mathcal{C}_t = c)$ is **close to** $\mathbb{P}_t(\cdot | \mathcal{C}_t = c)$ **for all values of c** . We then use $\mathbb{P}_t^*(\cdot | \hat{\mathcal{C}}_t)$ to generate future values of the index.

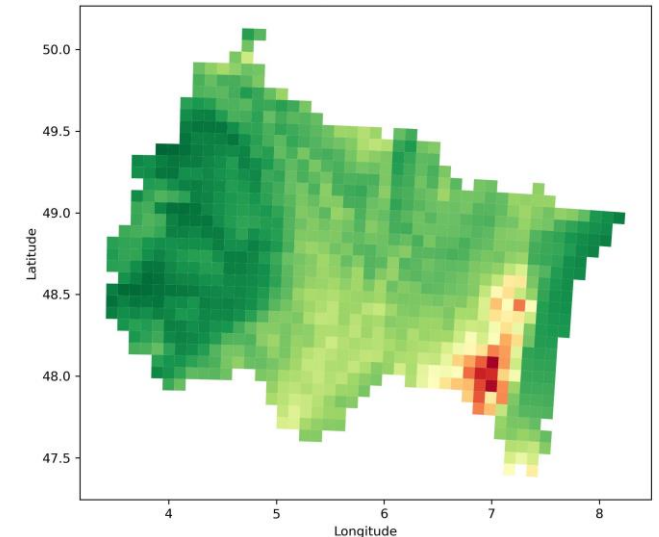
d coordinates of I_t
corresponding to d
geographical locations



A map in which each pixel
corresponds to the value of the index
at its location

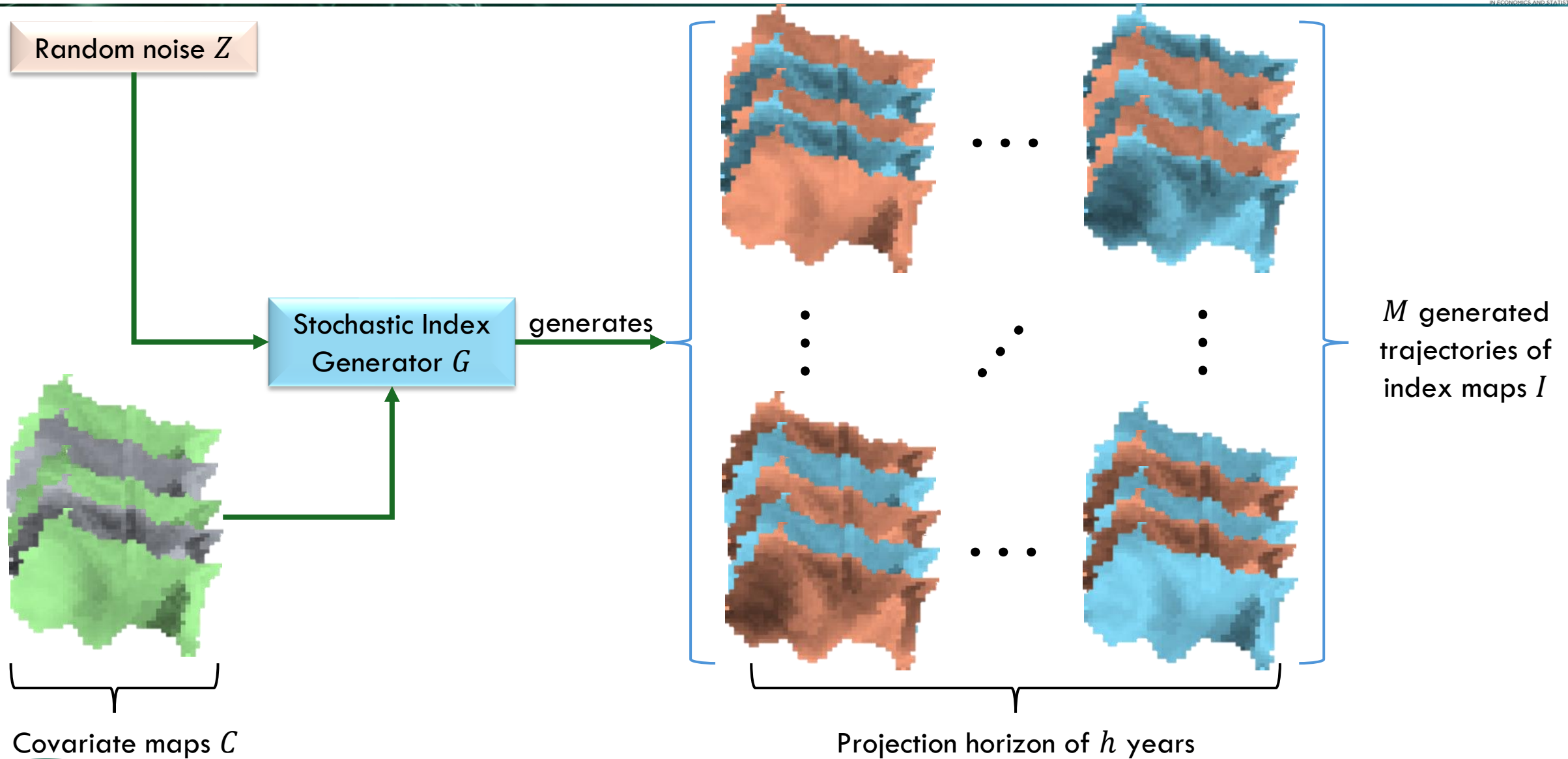


We then leverage advances in the field of image generation using GANs by applying them to our weather index projection problem.



Note: Our aim is not to forecast a single value or a single map, but rather to estimate the **distribution of** $(I_t)_{t=T+1, \dots, T+h}$ by drawing **M different values of Z** .

CONTEXT AND PROBLEM DESCRIPTION (7/7)

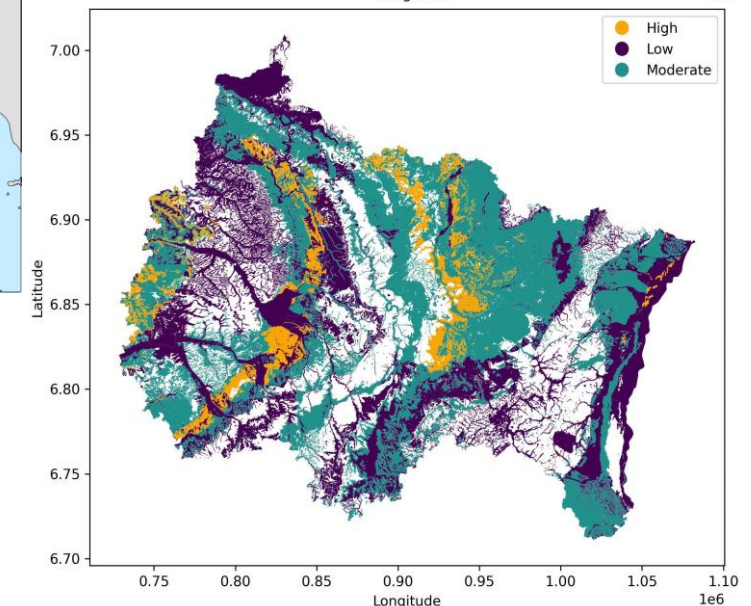
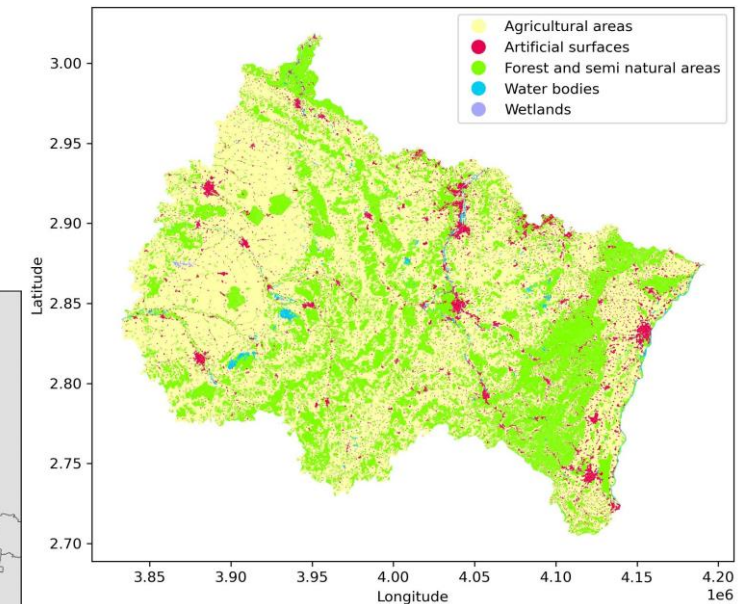


THE DATA (1/4)

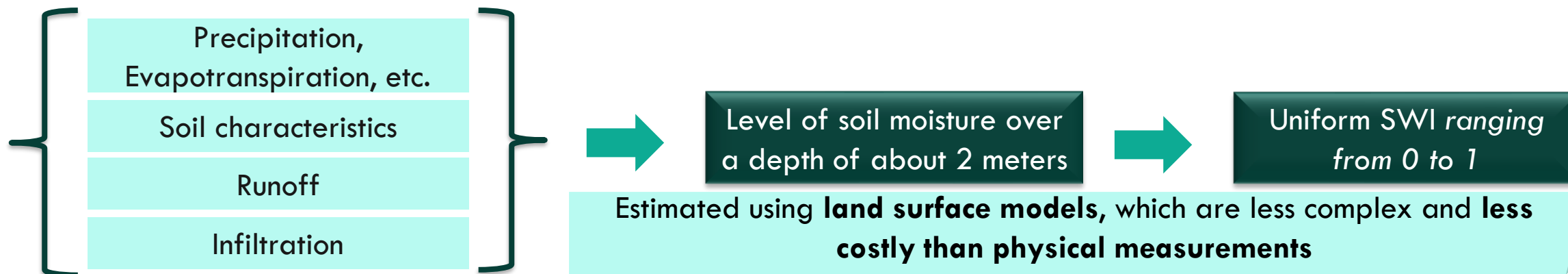
The study area

The study area to which SwiGAN is applied is the Grand Est region of France. This region has the following characteristics:

- A surface area of approximately 57,430 km² (about 10% of metropolitan France)
- Population of about 5.55 million
- 59.3% of agricultural land
- 33.8% is covered by forests and semi-natural environments
- The remaining urban area has a significant proportion of **clay soils**
- An expected increase in the probability of heatwaves and drier summers in the future
- SWI historical data are available from 1960 to 2024 at a monthly frequency



Description of the Soil Wetness Index



The SWI (which we will use to refer to the uniform SWI henceforth) is calculated by Météo-France using a specific configuration of the SIM2 land surface model. It has the following characteristics:

- It is the main indicator used by the French government to declare a state of natural disaster due to drought
- It is made available by Météo-France from 1960 to 2024 at a monthly frequency
- The spatial resolution of the data is $8 \text{ km} \times 8 \text{ km}$
- The data for the Grand Est region has dimensions of $37 \times 44 = 1,628$ pixels

We selected the period **1960 to 2020 for training**, **2021 to 2022 for validation**, and **2023 to 2024 for testing**.

THE DATA (3/4)

Description of the covariates

Variable	Description	Unit
huss	Specific humidity near the surface	kg/kg
prtot	Total precipitation	kg/m ² /s
rlds	Incoming longwave radiation	W/m ²
rsds	Incoming shortwave (visible) radiation	W/m ²
sfcWind	Horizontal surface wind speed	m/s
tas	Daily mean near-surface air temperature	K
tasmax	Daily maximum near-surface air temperature	K
tasmin	Daily minimum near-surface air temperature	K
evspsblpot	Potential evapotranspiration (Penman–Monteith, SICLIMA)	—

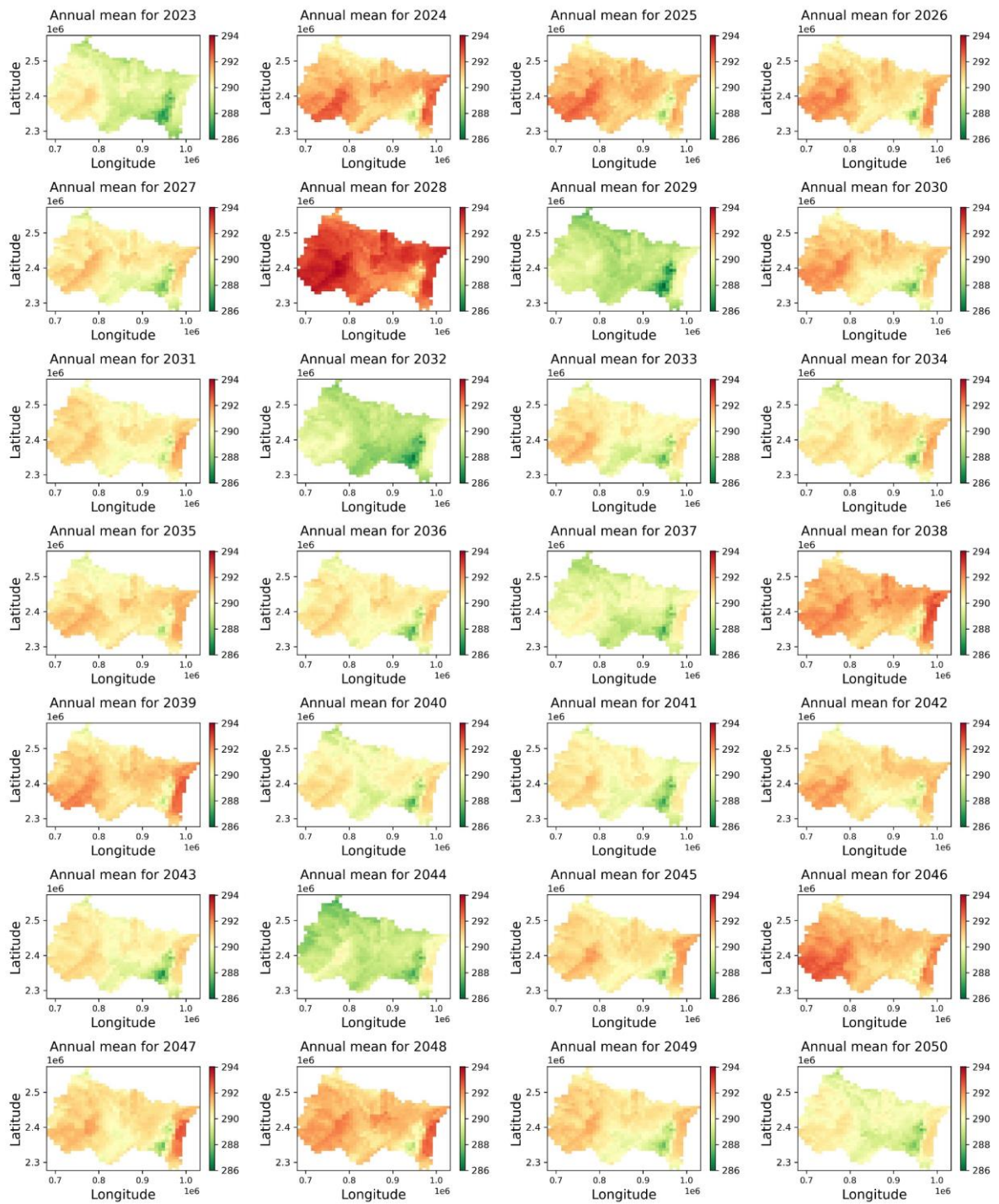
All covariates originate from global climate model outputs and are available for both historical and future periods under the **RCP 4.5** and **RCP 8.5** scenarios..

The covariates are available at the **same spatial resolution as the SWI (8 km × 8 km)**, but at a **daily frequency**.

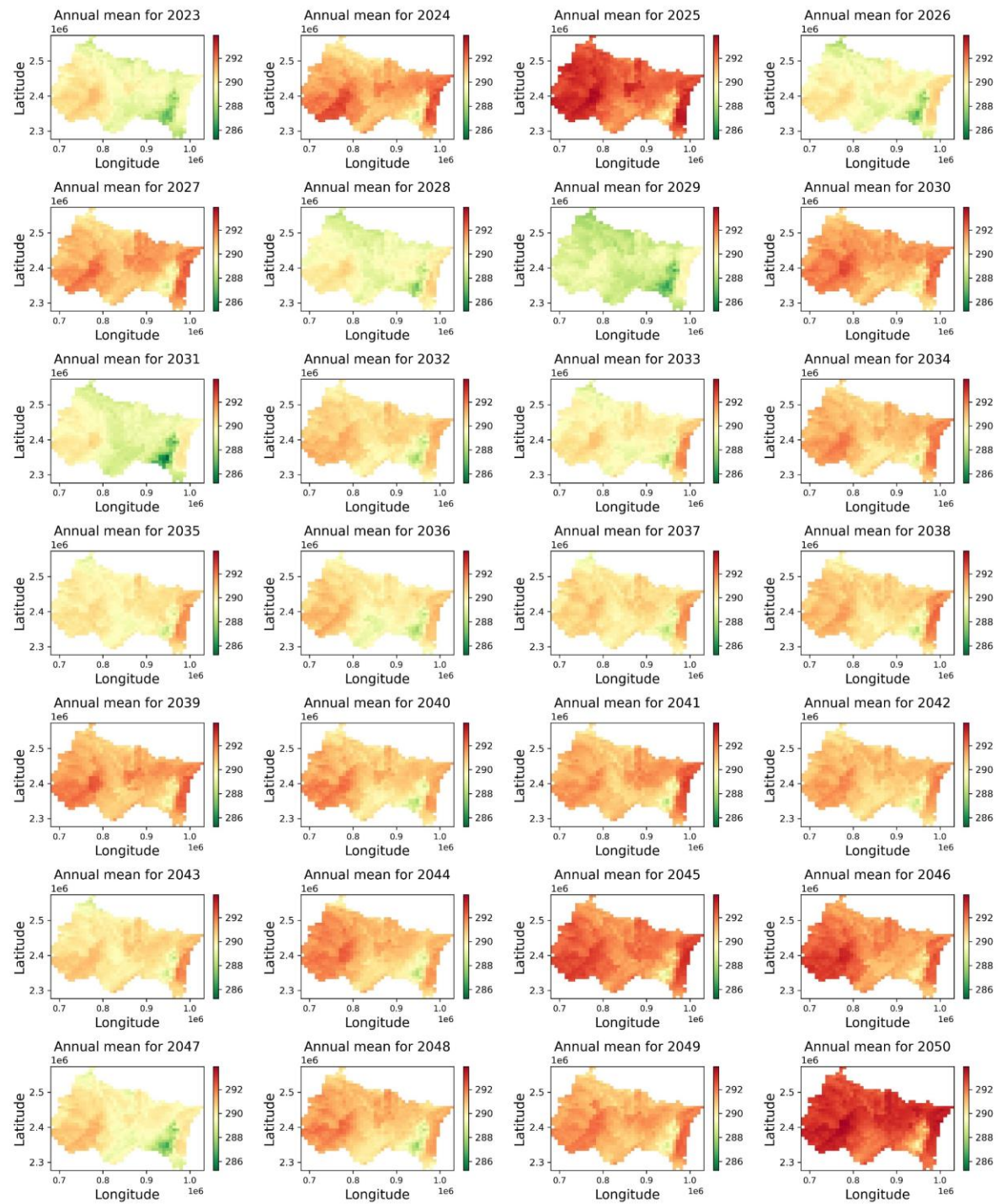




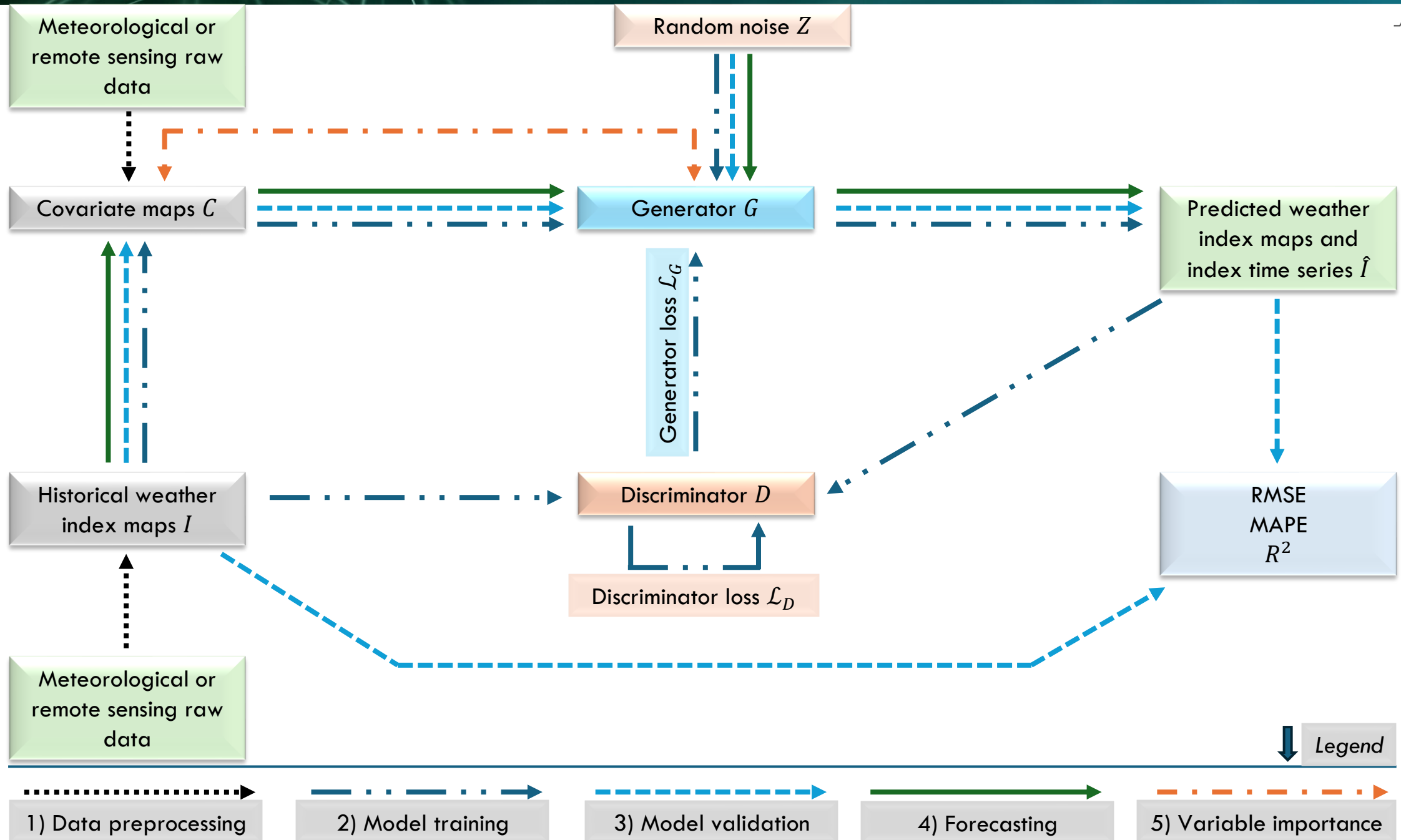
Mean annual temperature under RCP 4.5



Mean annual temperature under RCP 8.5

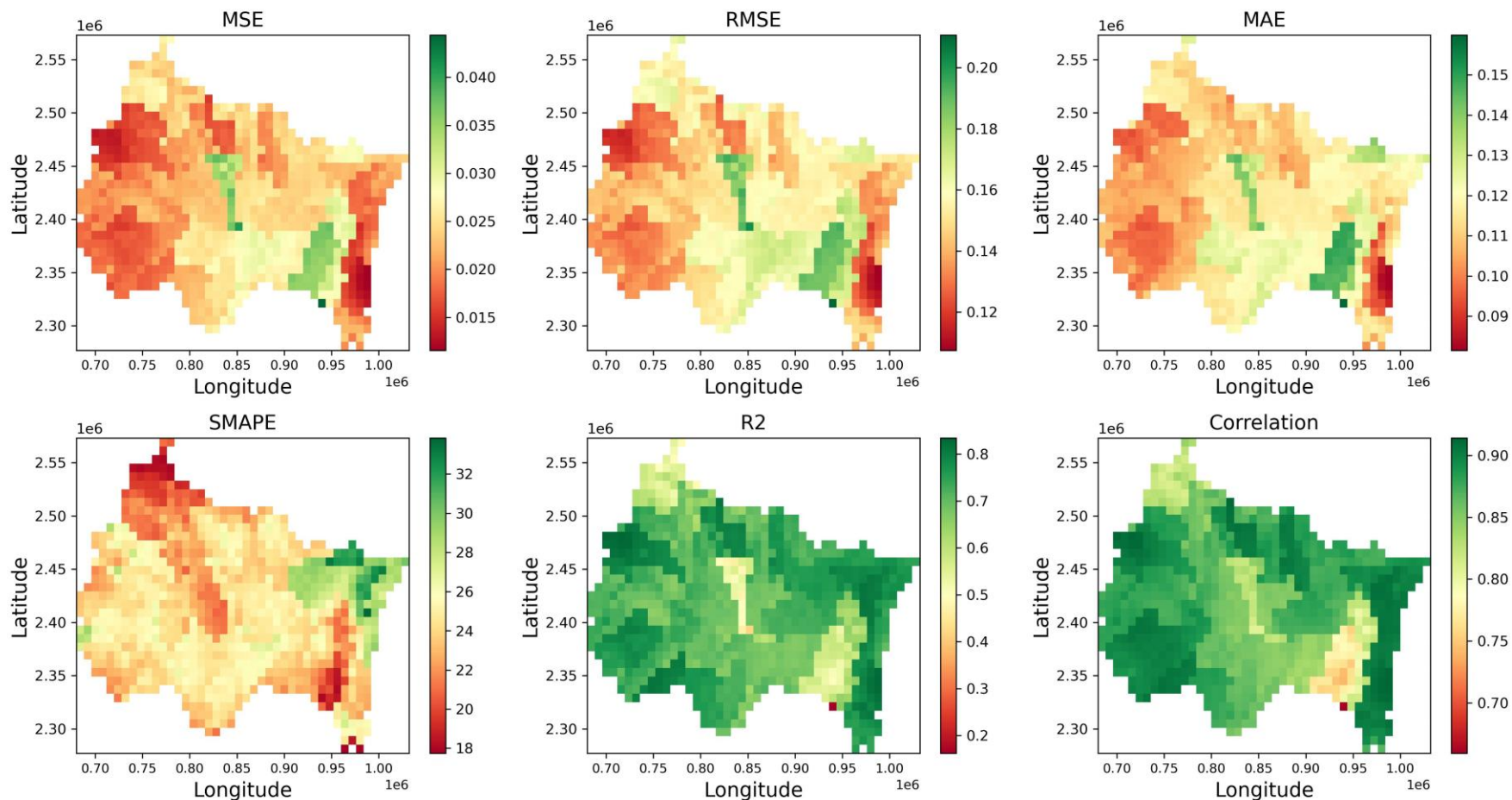


METHODOLOGY OF THE STUDY



FORECASTING RESULTS (1/5)

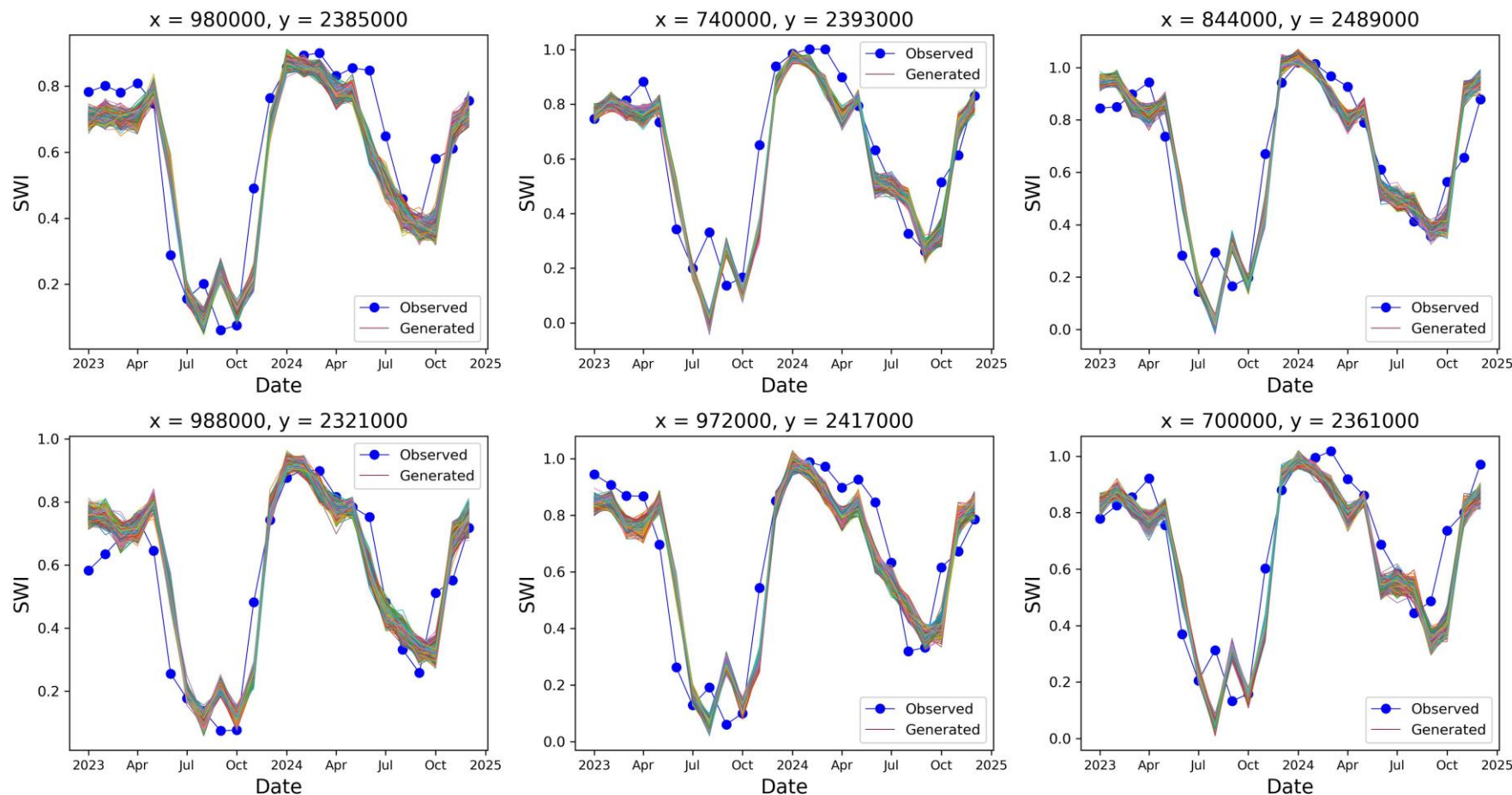
Validation of final model



Overall, the performance of the SwiGAN model is strong: more than 80% of pixels have SMAPE values below 25.59%, R^2 values above 0.67, and correlations above 0.85.

FORECASTING RESULTS (2/5)

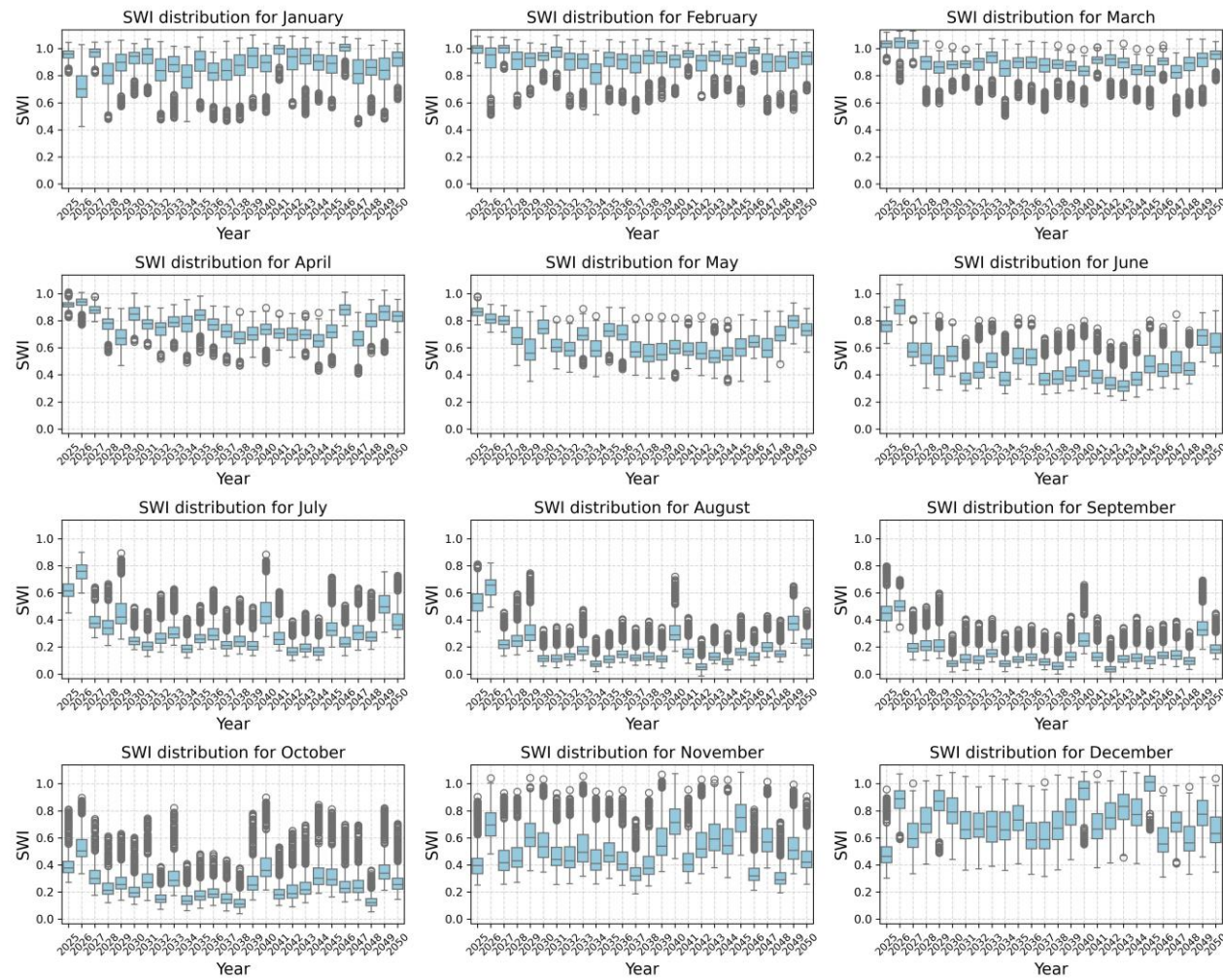
Validation of final model



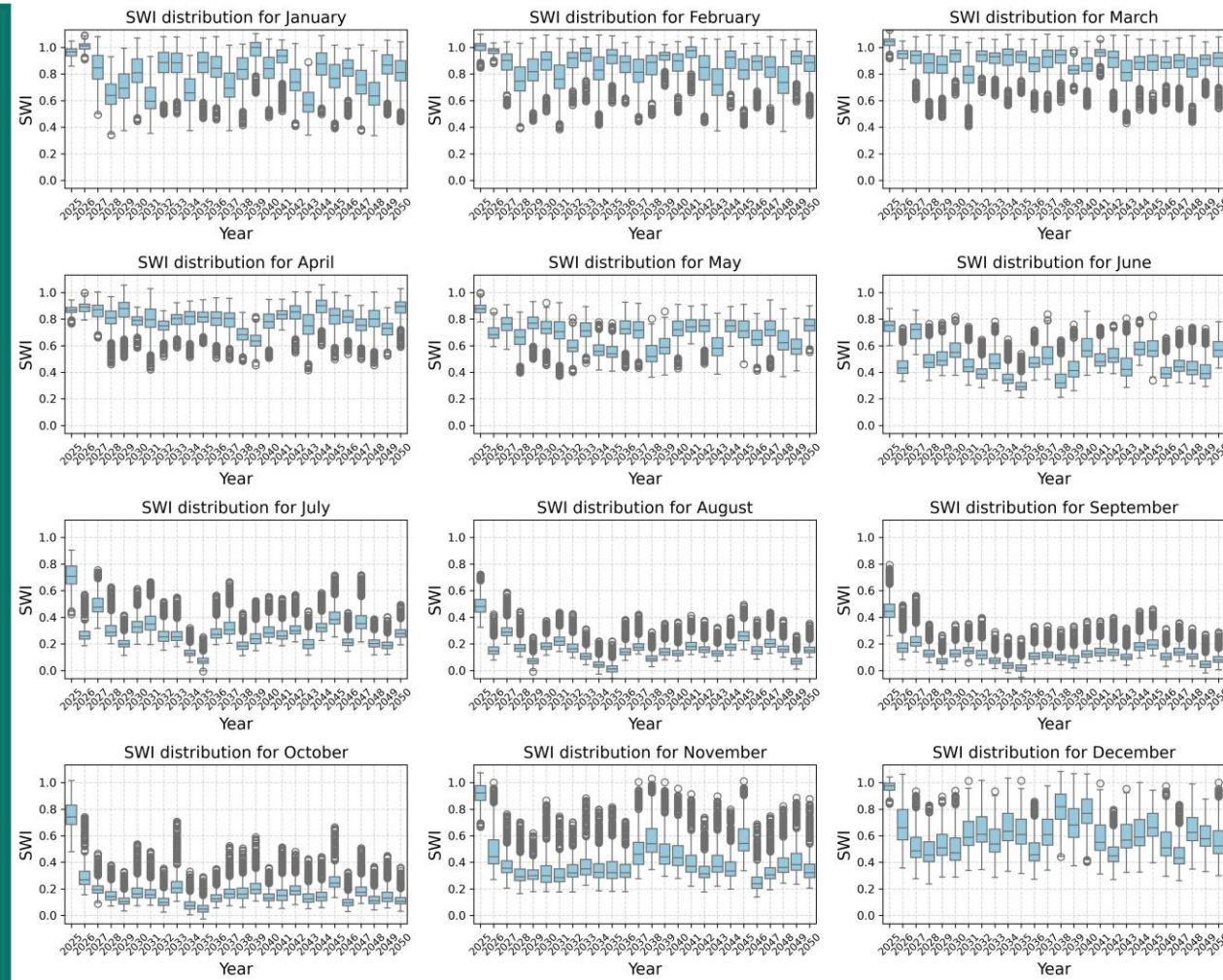
The generated trajectories follow the same trend as the observed trajectory, as indicated by the values of R^2 and the correlation coefficient.

FORECASTING RESULTS (3/5)

Spatial distribution of mean generated SWI up to 2050



Under RCP 4.5

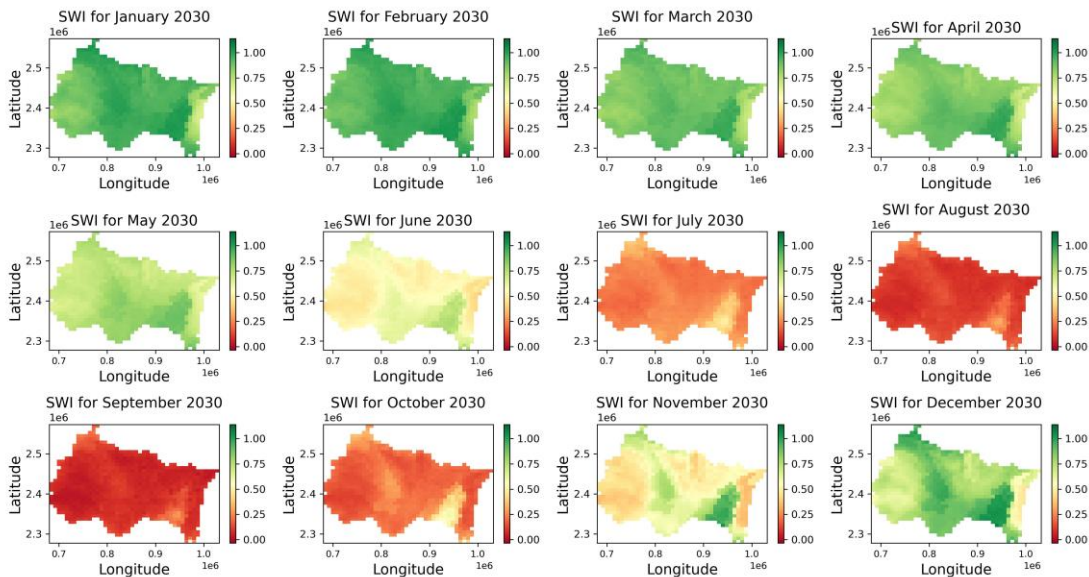


Under RCP 8.5

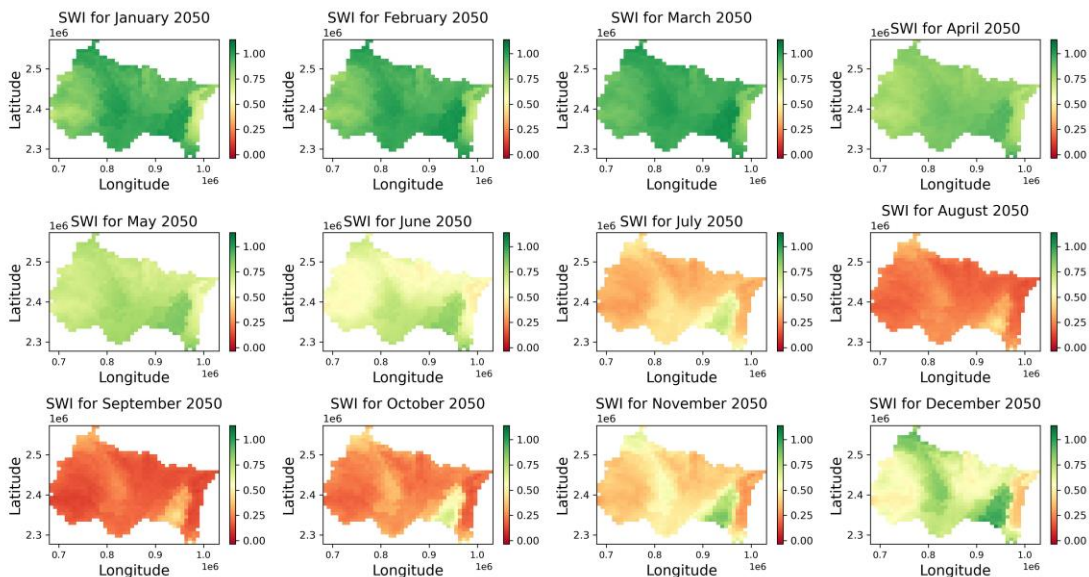
FORECASTING RESULTS (4/5)

Observing the years 2030 and 2050

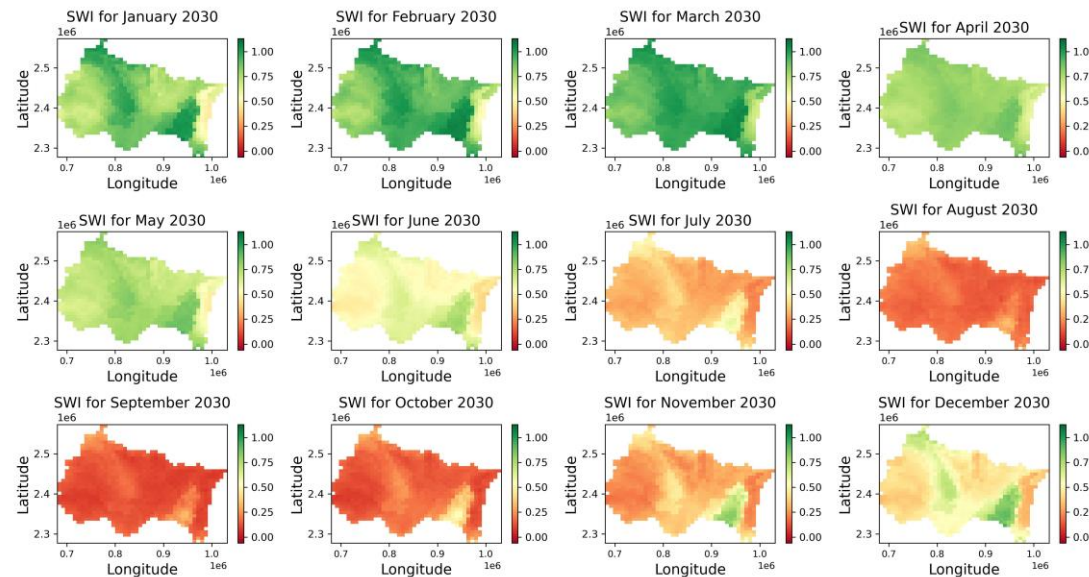
2030 under RCP 4.5



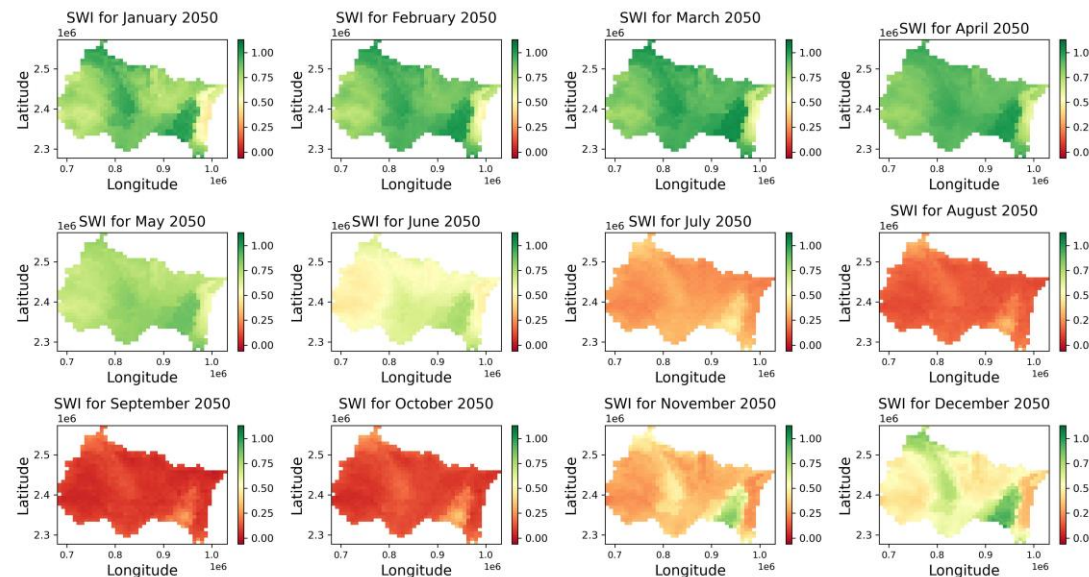
2050 under RCP 4.5



2030 under RCP 8.5

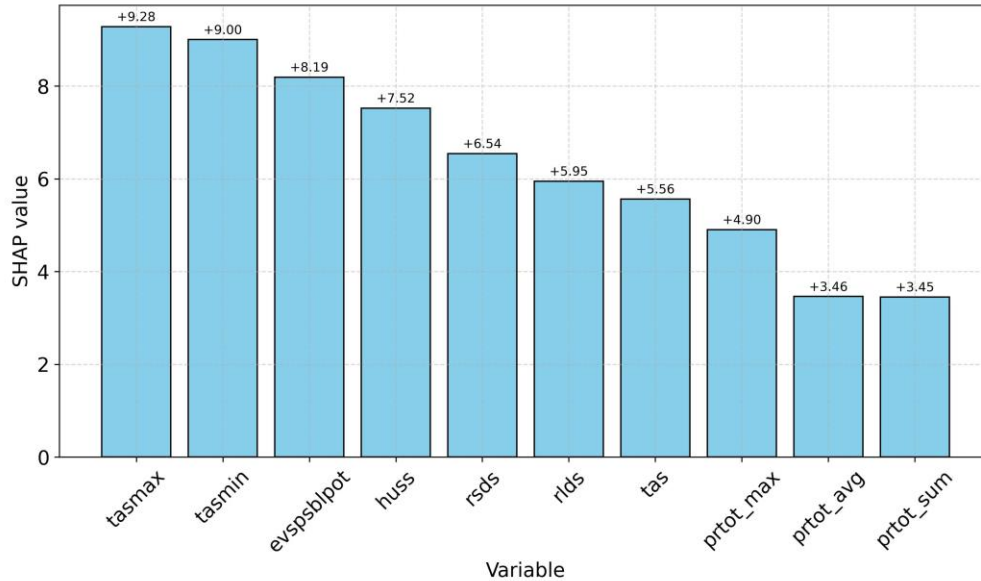


2050 under RCP 8.5

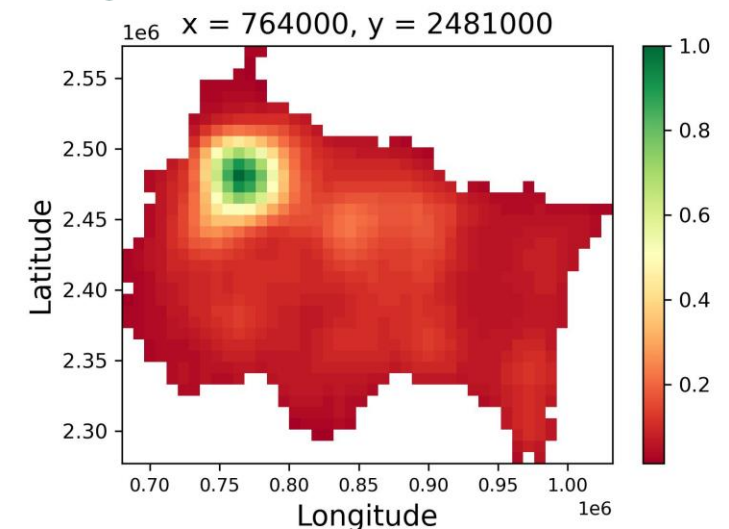
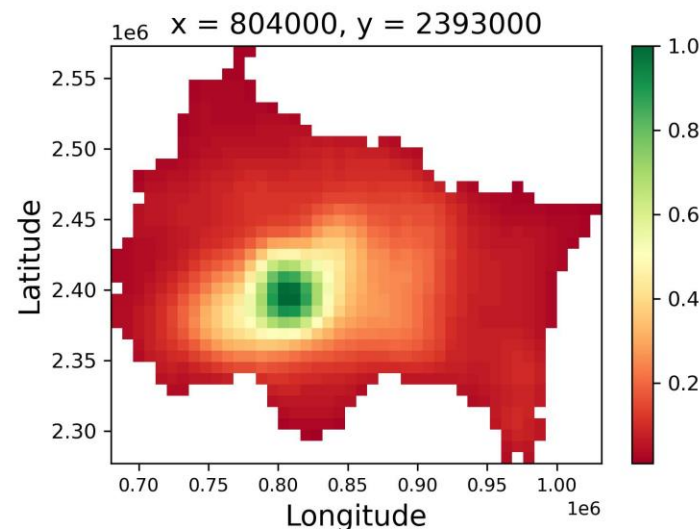
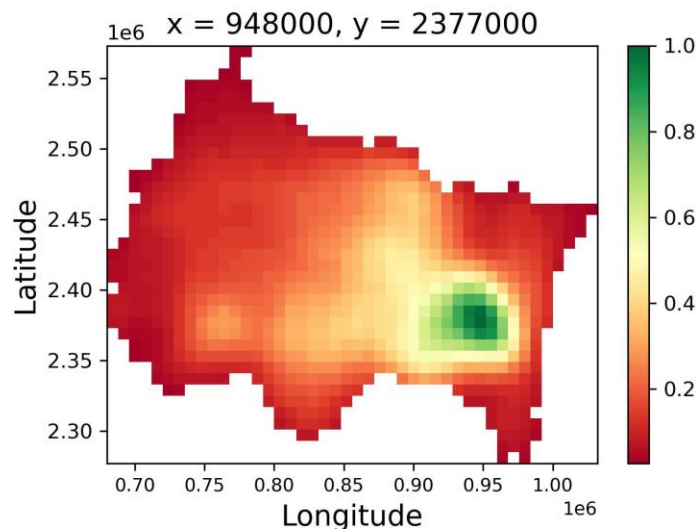


FORECASTING RESULTS (5/5)

Variable and spatial importance

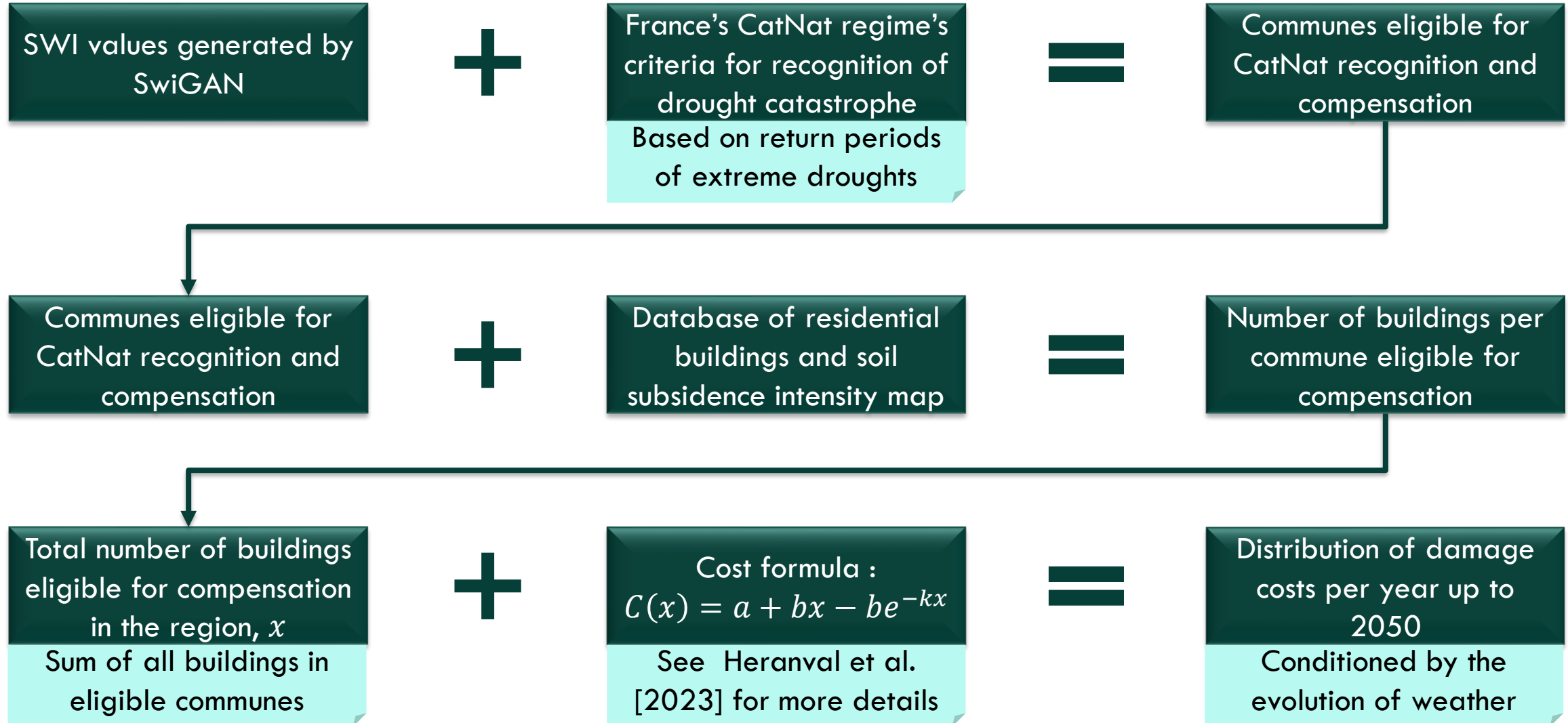


- As expected, temperature and evapotranspiration have the greatest impact on the variability of the SWI
- An unexpected low impact of precipitation on the variability of soil moisture (see Xue and Wu, 2025)
- All variables exhibit similar explanatory power
- SwiGAN is able to capture spatial information from relatively large regions surrounding the predicted pixels.
- SwiGAN adapts to the geographical location of each pixel

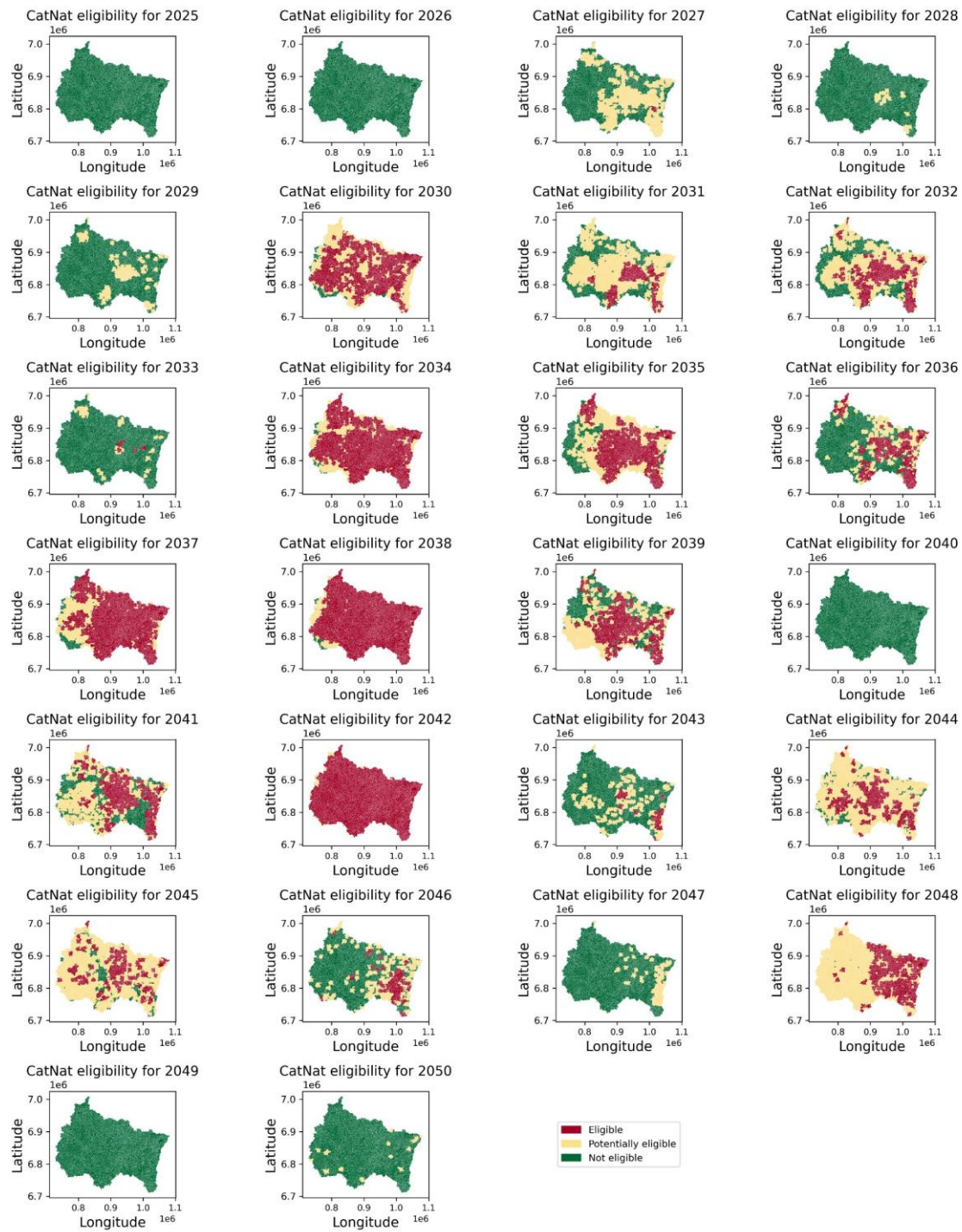


APPLICATION TO DROUGHT-INDUCED SUBSIDENCE INSURANCE (1/3)

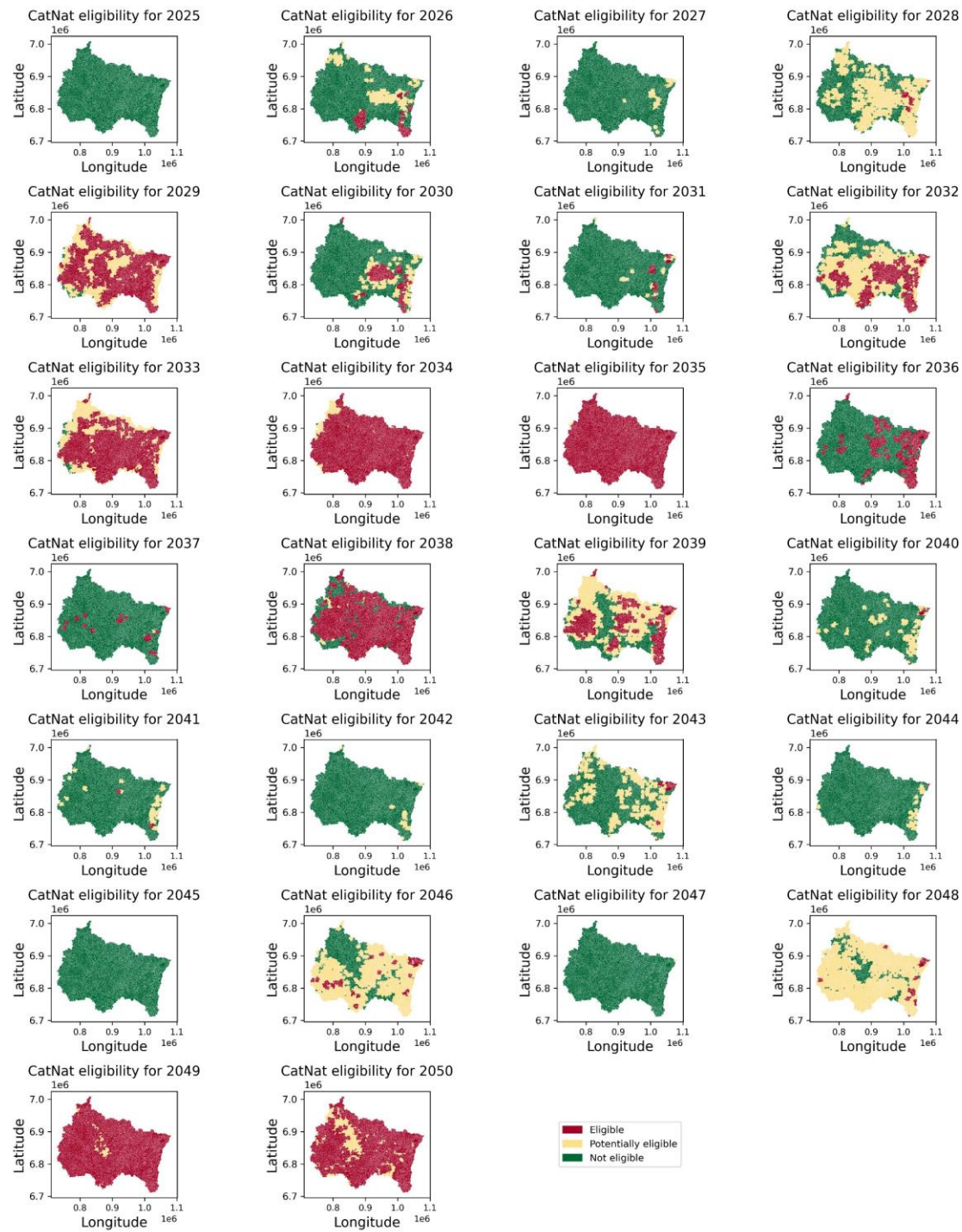
The methodology



Communes eligible for compensation under RCP 4.5

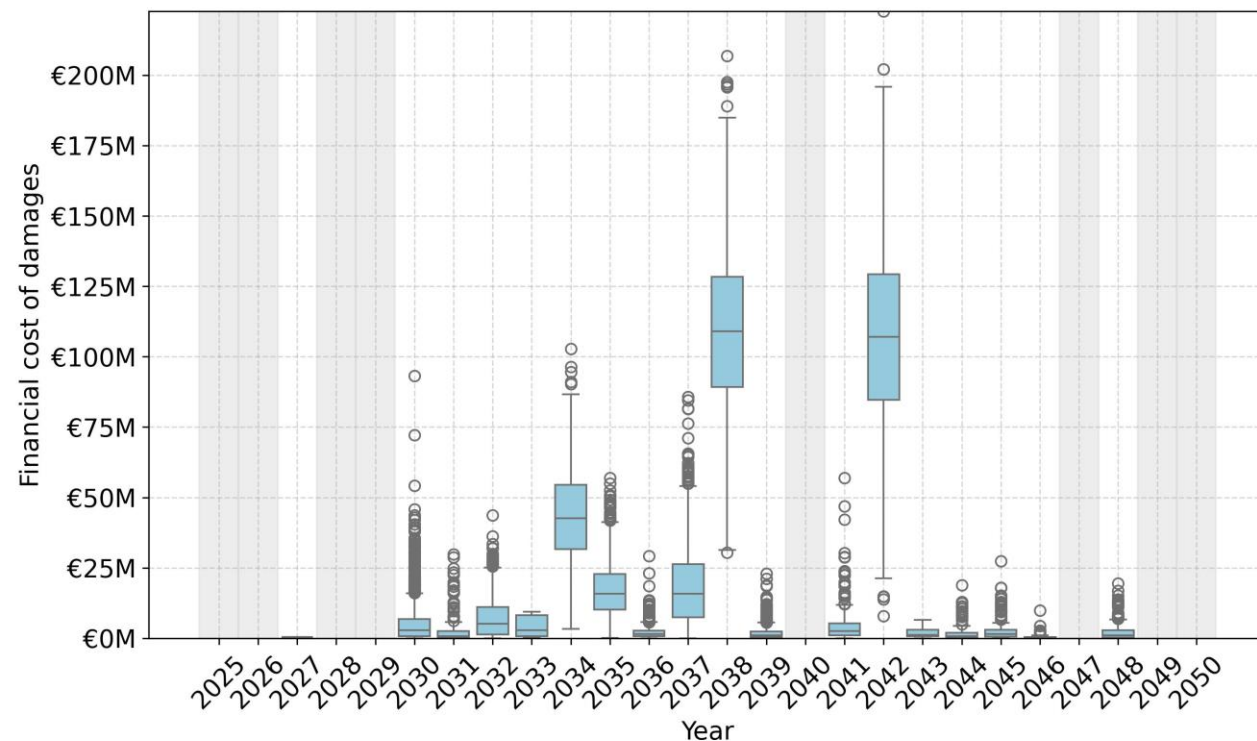


Communes eligible for compensation under RCP 8.5

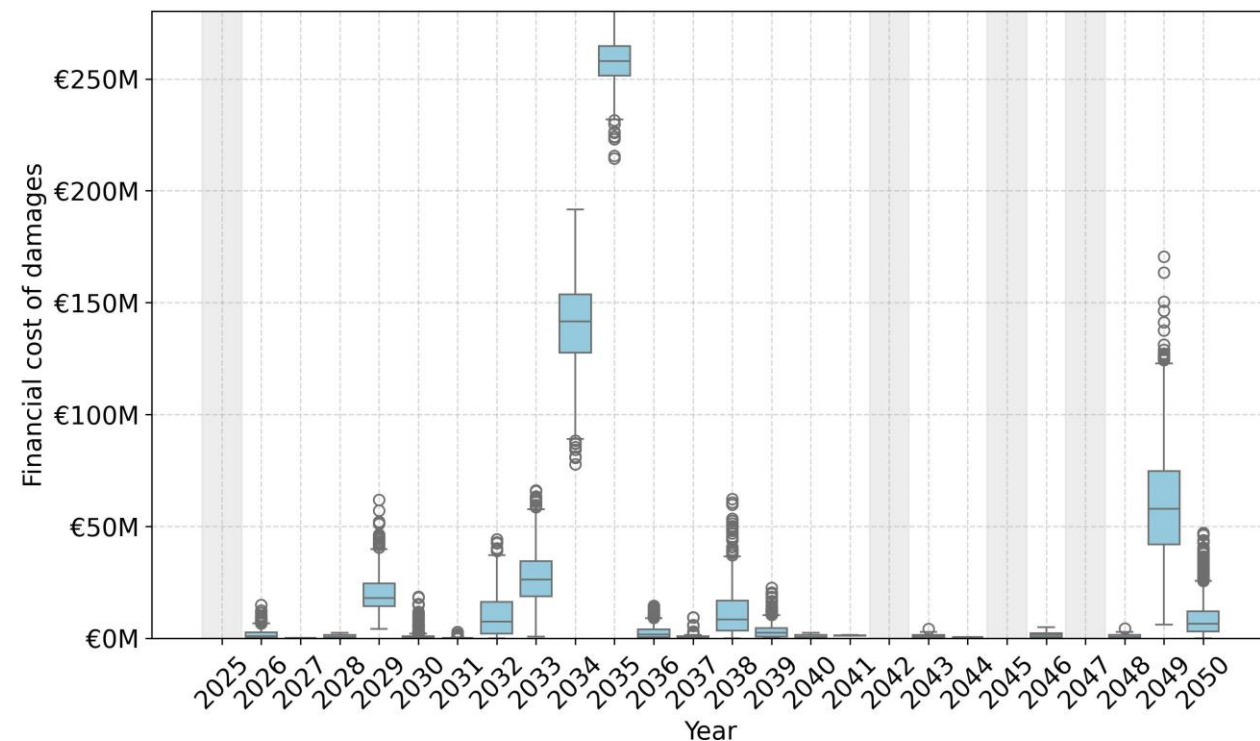


APPLICATION TO DROUGHT-INDUCED SUBSIDENCE INSURANCE (3/3)

Distribution of compensation costs



Future damage costs under RCP 4.5



Future damage costs under RCP 8.5

The evolution of compensation costs does not systematically increase over time. This is due to the number and complexity of the factors that influence the evolution of the SWI. The complexity of the eligibility criteria under France's CatNat regime can also explain the non-monotonic evolution of these costs.

Limits on the application to drought-induced subsidence insurance

- Assumption that exposure (population and number of buildings) remains fixed
- Assumption that there is no evolution in building standards
- Assumption that damaged houses are not removed from the exposure

Conclusion

- The Wasserstein GAN methodology introduced in this paper demonstrates the benefits of generative techniques for simulating weather indices related to insurance claims.
- The example developed for the Soil Wetness Index in the context of subsidence can be extended to other quantities (precipitation levels, hail severity indicators, or even complex indices used in parametric insurance).
- Moreover, it can adapt to the assumptions underlying different climate scenarios, allowing for straightforward comparison of their impact on insurance.
- Compared to fully physical models, generative models can produce simulations rapidly, enabling the efficient computation of relevant insurance metrics, such as Value-at-Risk.
- We believe that the generative methodology we develop can serve as a valuable tool for assessing the impact of various adaptation strategies and enhancing prevention in the field of natural disasters.

THANK YOU FOR YOUR ATTENTION!



QUESTIONS ?