Noise Enhanced Students

Alessandro Catalano

PhD in complex systems for physical, socio-economic and life sciences, University of Catania

Gabriele La Milia

Bachelor degree in Management Engineering, Università degli Studi di Palermo

Gabriele Lo Cascio

Bachelor degree in Computer Engineering, Università degli Studi di Palermo

TUNNEL EURALPIN LYON TURIN

2015 2025

Task

The challenge has a very specific goal: predict the climate impact on the territory of the Susa and Maurienne Valley. Specifically, the development of a custom Climate Model for predictions up to 2050.

The hybrid nature of this challenge allowed a multidisciplinary team, such as ours, to approach this task from different perspectives, from Physics (climate as a complex system) to Statistics and Data Science.

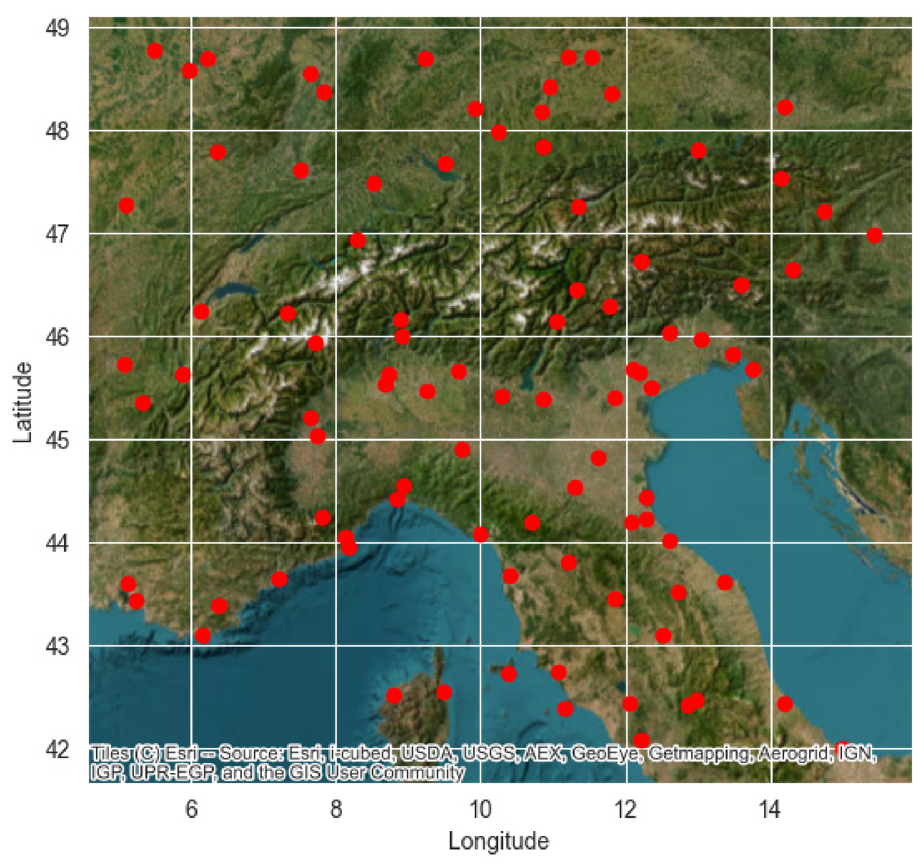
Dataset Journey

- Climate models need lots of giga of data and lots of numerical simulations.
- ERA5 dataset and others are too big to be downloaded and have so many features to analyze.
- Our choice is to work with data from IOWA STATE UNIVERSITY.
- The IOWA dataset maintains an ever-growing archive of automated airport weather observations from around the world.

Link: https://mesonet.agron.iastate.edu/request/download.phtml

Data Processing

- We focus our analysis on a specific geographic area [1] to reduce the amount of data and, in this way, we also create a custom model for the Valleys.
- From the IOWA dataset we consider the stations with a number of missing consecutive days which is less than 3% over the period (1980-2024).
- We consider only these features: temperature, dew point, relative humidity and wind direction. Other features were not available.
- We note also a presence of outliers. So, we apply winsorization, a robust technique to handle outliers. Imputation of missing values has been done with KNNImputer.

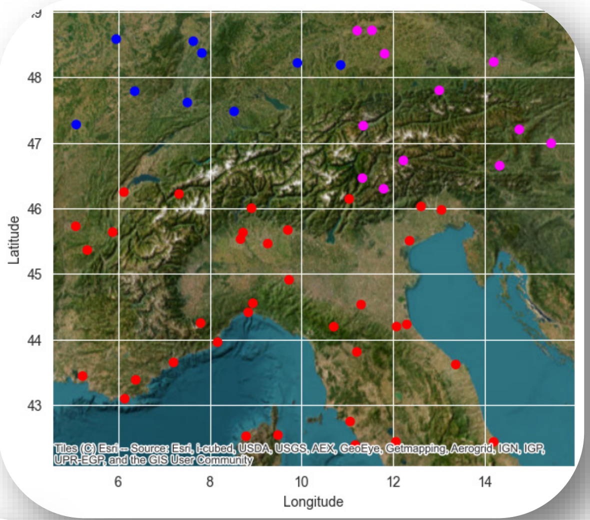


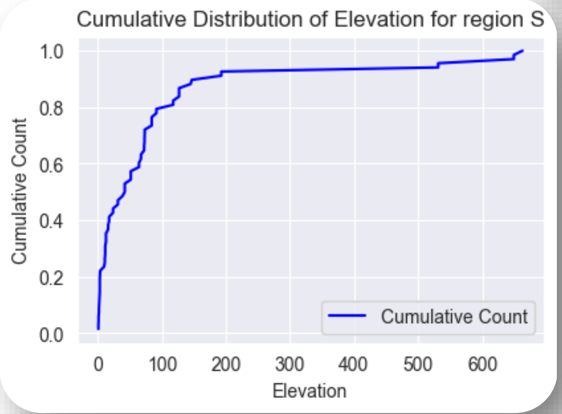
[1] Kotlarski, S., Gobiet, A., Morin, S. et al. 21st Century alpine climate change. Clim Dyn 60, 65–86 (2023). https://doi.org/10.1007/s00382-022-06303-3

Spatial and Temporal Heterogeneity

Choosing the appropriate spatial and temporal scale is crucial for obtaining accurate and meaningful predictions:

- About latitude and longitude, we divided the total area in three parts: the south in red (strong Mediterranean influence); the northwestern in blue (strong Atlantic influence); the northeastern in magenta (continental character).
- About elevation, we created three groups only for the south (considering the cumulative distribution).
- At the same time, since we must predict the next 25 years, we lay our analysis on the past 45.
- Every season is different, so we distinguish them all.

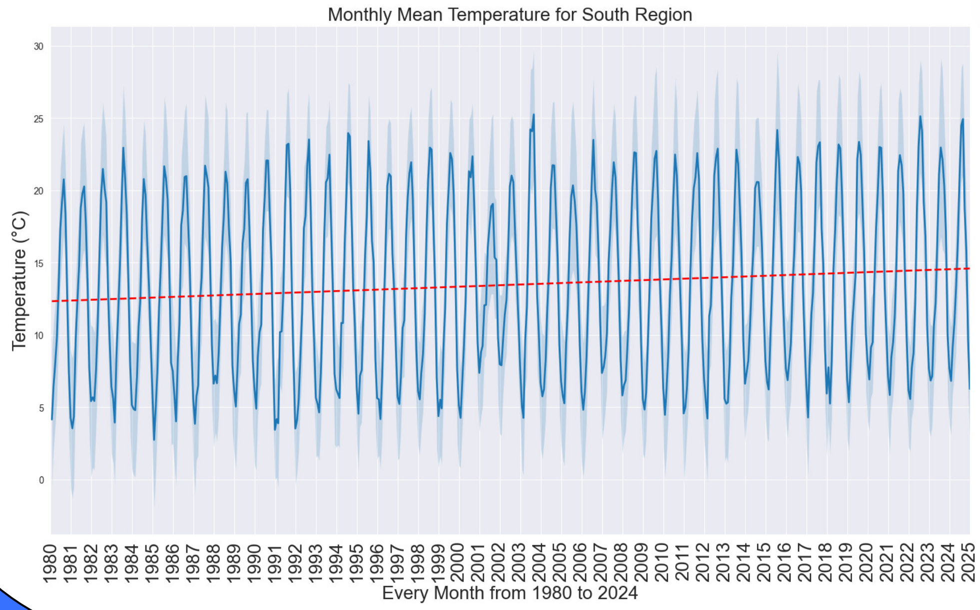




Data Analysis

In short:

- Temperature is rising.
- Dew Point is almost costant.
- Wind Direction is shifting from the south.
- Relative Humidity decreases.



TUNNEL EURALPIN LYON TURIN

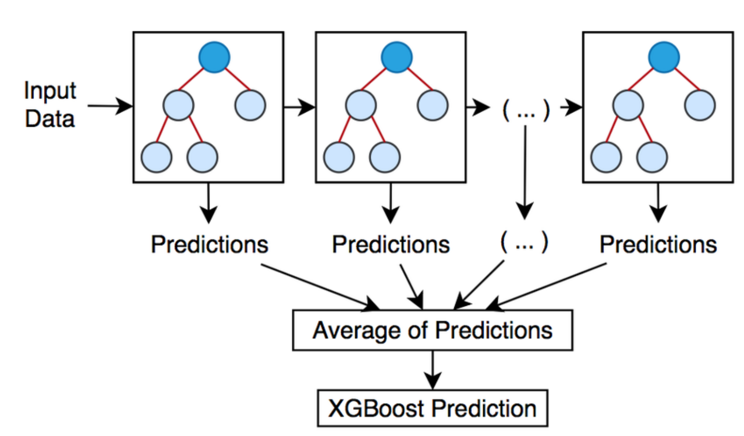
2015 2025

Model Development

We focus on the south region, where Susa and Maurienne Valleys are. We consider these features: station, date, **longitude**, **latitude**, elevation, elevation group, region, season.

So, we train 4 models with these features as input and a different target for each model: temperature, dew point, relative humidity and wind direction.

Since we found some seasonality in the data, we decide to predict mean values of temperature, dew point, relative humidity and wind direction for each month, using the XGBRegressor [2].



Explainable AI for Interpretable Credit Scoring - Scientific Figure on ResearchGate.

We have 2 group of models:

- The first four models are trained on all the south region dataset.
- The second four models are trained on all the south region dataset filtered with Elevation Group H3 (the highest elevation).

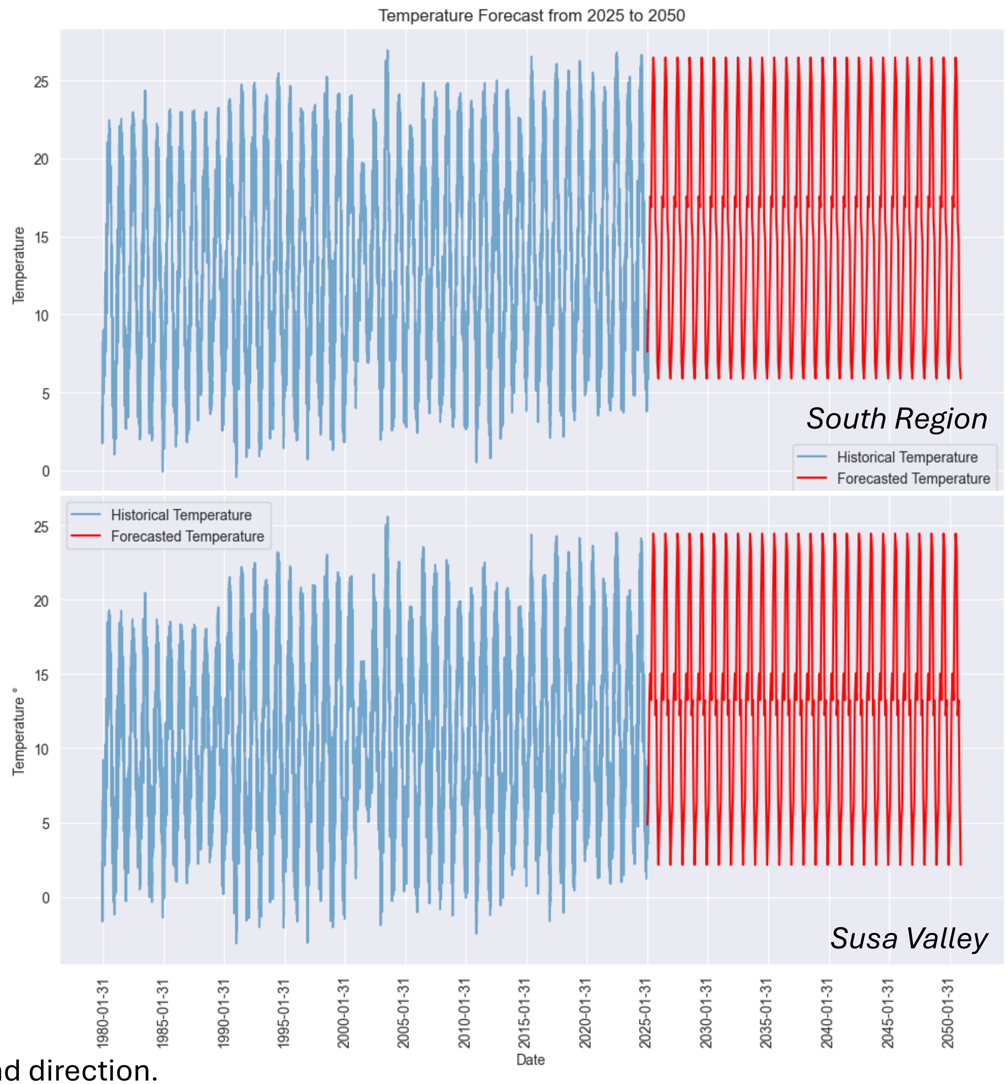
[2] hyperparameters: n_estimators=1000, learning_rate=0.05, random_state=177, max_depth=16

Forecasting

Our setup let us to choose latitude and longitude and make predictions of any given point in the map up to 2050.

In the south region our mean predicted temperature is higher than Susa or Maurienne Valleys. As expected, different points in the map follow different trend, and we are able to capture them.

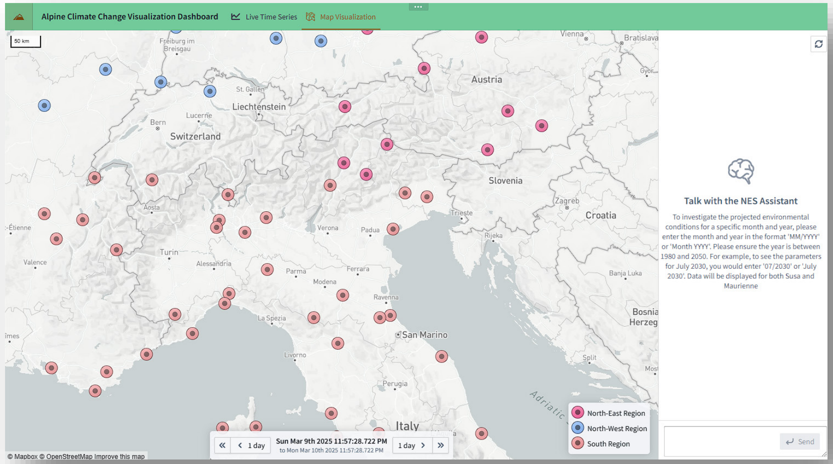
We understand how each point in the map can be grouped based on its spatial characteristics (region and elevation). In this way we reduce heterogeneity (we have only homogeneous records), and this allow us to make high resolution predictions (local scale).



Note that we have done forecasting also for dew point, relative humidity and wind direction.

Visualization and Awareness

We realized an interactive dashboard using Palantir Foundry: an interactive map with stations is shown with a chatbot that uses an LLM (giving input date, it returns the predictions about temperature, dew point, wind direction and relative humidity for Susa and Maurienne Valley). Another Tab is used to plot climate data from 1980 to 2050 with filters.



The challenge also required a prototype app for students and citizens. KlimaKindle:

- Simplifies explanations with interactive quizzes
- Displays projections and consequences
- Easy to use and accessible
- Leaderboard and short online battles with friends

Our Contributions

- A multidisciplinary team of physicist and engineers allowed for an exchange of expertise.
- We built a dataset considering the European area and historical period (1980-2024) consistent with the scientific literature about the Alpine climate change.
- Alpine climate has been considered as a complex system. In this sense, much attention was paid to the spatial and temporal heterogeneity of the available data, in order to obtain accurate and representative **estimates** and **predictions** of the area.
- Data analysis shows an increase in average monthly temperature from 1980 to 2024 in all regions (S, NW, NE).
- An increasing maximum temperature trend in summer and minimum temperature trend in winter (in addition different altitudes have a significant impact on them).
- A reduction in average monthly relative humidity and a shift in wind direction in the southern region from 1980 to 2024.
- Our setup let us to choose latitude and longitude of Susa and Maurienne Valleys to make predictions, but also of any other given point in the map up to 2050, solving the task of the hackathon.
- We built an interactive app to display the predictions made and to raise awareness of climate change among young people for educational and playful purposes