Testing the use of deep learning techniques for emulating regional reanalysis

Antonio Pérez1, Mario Santa Cruz1, Javier Díez-Sierra2, Matthew Chantry3, András Horányi3, Mariana Clare3, and Cornel Soci3
1Predictia Intelligent Data Solutions S.L. (Predictia), Santander, Spain.
2Instituto de Física de Cantabria (IFCA), CSIC-Universidad de Cantabria, Santander, Spain.
3European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, United Kingdom.

Introduction
Reanalysis datasets merge historical weather data with predictive models for extensive climate monitoring. However, existing datasets like ERA5 lack the resolution for detailed local-scale analysis, while higher-resolution datasets like CERRA are computationally expensive. To address this, we developed a deep learning model to emulate CERRA’s 2m temperature field using ERA5 as input.

Data sources
ERA5 (0.25°) and CERRA (0.05°) reanalysis datasets are used, requiring a fivefold downsampling. A specific spatial domain has been selected to reduce computational costs (refer to Figure 1). ERA5 covers a geographic area ranging from -8.35° to 6.6° longitude and 35.50° to 46.45° latitude, whereas CERRA encompasses a slightly narrower region spanning from -6.85° to 5.1° longitude and 37° to 44.95° latitude.

Model architecture
Our model, based on the Swin v2 architecture, upscales the inputs by a factor of 8 using transformers.Swin2SRModel. In addition, the input is previously preprocessed by a Convolutional Neural Network (CNN) (see Figure 2). With over 12 million parameters, our model features 6 blocks of depth 6 with 6 heads each, optimized for validation loss. Patch size is 1 pixel, and window size is a common divisor of input dimensions (5).

Training procedure
A composite loss is used including terms for primary predictions, downsampled predictions, and blurred predictions to enhance accuracy. Trained over 100 epochs, our model employs the Adam optimizer with a learning rate of 0.0001 and a warm-up phase of 500 steps. Training spans 29 years of data, assessing performance at the end of each epoch against 3 years of independent data.

General results
The model shows a mean temperature prediction error of 0.08 Celsius, with the highest errors occurring in regions of greater variability (refer to Figure 3).

Figure 1: Spatial domain selected for the data sources.

Figure 2: Schematic representation of the model architecture.

Figure 3: The spatial representation of the mean error of our model during the testing period.

It outperforms the bicubic interpolation with a 85% reduction in MAE and around 30% lower RMSE (see Figure 4).

Figure 4: The absolute (top) and relative (bottom) improvement in RMSE (see Figure 4).

More specific results
We achieve significant improvements in complex landscapes. Validation at specific locations, such as Aneto mountain, shows a dramatic error reduction from -6.3°C to 0.06°C — a 99% improvement (refer to Figure 5). Similar enhancements are observed in Cantabrian Mountains, with Peña Vieja showing a 94% improvement (see Figure 6) and Peña Labra an 88% improvement.

Figure 5: The prediction error (in °C) distributions of ConvSwin2SR and baseline methods in the mountainous area of Pyrenees, Spain.

Figure 6: The prediction error (in °C) distributions of ConvSwin2SR and baseline methods in the mountainous area of Peaks of Europe, Spain.

Conclusions
In conclusion, our project shows promising results in emulating reanalysis data with AI, offering potential benefits in terms of computational efficiency and development of real-time applications. While initial results are encouraging, comprehensive validation against CERRA and broader domain expansion are crucial for confirming effectiveness and reliability.

Acknowledgements
The results presented were part of a project funded by the ECMWF Code4Earth 2023 initiative, an innovation program aimed at driving open source developments in the Earth sciences community. Scan the QR code to watch our ECMWF Code4Earth 2023 final presentation on YouTube.