Check for updates

Geophysical Research Letters

RESEARCH LETTER
10.1029/2024GL113540

**Key Points:**

- A convolutional neural network (UNET) is set and trained to predict temperature anomalies
- This pilot study is carried out in an idealized framework using only data from the large MIROC6 ensemble
- UNET exhibits excellent skills, with correlations higher than 0.9 over Europe, higher than analogs

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

E. Cariou,
enora.cariou@meteo.fr

**Citation:**

Cariou, E., Cattiaux, J., Qasmi, S., Ribes, A., Cassou, C., & Doury, A. (2025). Linking European temperature variations to atmospheric circulation with a neural network: A pilot study in a climate model. *Geophysical Research Letters*, *52*, e2024GL113540. https://doi.org/10.1029/2024GL113540

Received 14 NOV 2024
Accepted 19 FEB 2025

© 2025. The Author(s).
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

# Linking European Temperature Variations to Atmospheric Circulation With a Neural Network: A Pilot Study in a Climate Model

**Enora Cariou[1], Julien Cattiaux[1], Saïd Qasmi[1], Aurélien Ribes[1], Christophe Cassou[2,3], and Antoine Doury[1]**

[1]Centre National de Recherches Météorologiques, Université de Toulouse, CNRS, Météo-France, Toulouse, France, [2]LMD/IPSL, CNRS, Ecole Normale Supérieure Université PSL, Paris, France, [3]CECI, CNRS-Cerfacs, Toulouse, France

**Abstract** In Europe, temperature variations are mainly driven by the North Atlantic atmospheric circulation. Here, with data from the MIROC6 large ensemble, we investigate a convolutional neural network (a UNET) for reconstructing daily temperature anomalies in Europe from Sea Level Pressure (SLP) as a proxy of the atmospheric circulation, and we compare the results with a traditional analogs approach. We show an excellent ability of the UNET to estimate temperature variations given information from SLP only. This novel method outperforms the analogs method, at both daily and inter-annual time scales. Our study also shows that during the training, the UNET learns information such as the seasonal cycle of the relationship between sea-level pressure and temperature anomalies, which could explain part of its excellent scores. This exploratory work opens up promising prospects for estimating the contribution of atmospheric variability to observed temperature variations.

**Plain Language Summary** Day-to-day variations in European temperatures are strongly linked to fluctuations in the large-scale atmospheric circulation over the North Atlantic basin. Europe is one of the fastest warming regions in the world, and it is essential to understand the contribution of atmospheric circulation to these recent temperature trends. This requires a proper quantification of the relationship between a given atmospheric circulation (e.g., a SLP map) and the associated temperature anomaly over Europe. Here we present a novel approach to address this issue, based on artificial intelligence techniques. In an idealized framework (i.e., numerical simulations of a climate model), we show that this approach has excellent results and outperforms the traditional analogs circulation method. We show that the neural network used in this study is able to learn a lot of information from a single pressure map, such as the season, and even the day of the year with only a small error. Our results are very promising for further research on the contribution of atmospheric variability to temperature variations.

## 1. Introduction

Inter-annual to multidecadal variations in European temperatures are mainly linked to the variability of the North Atlantic dynamics on top of anthropogenically forced trends. The North Atlantic Oscillation (NAO) is the most prominent atmospheric pattern explaining a significant fraction of climate variability, especially during boreal winter (Deser et al., 2016). Featuring changes in position and intensity of the Azores High and the Iceland Low, the NAO influences mean wind speed and direction over the entire North Atlantic basin due to the related latitudinal shift of the stormtrack and modulation of the trade winds. NAO-altered heat and moisture transports have a direct impact on temperatures over Europe (Hurrell et al., 2003). Europe is currently warming particularly rapidly with respect to other regions (https://climate.copernicus.eu/climate-indicators/temperature), especially in summer (Schumacher et al., 2024; Vautard et al., 2023). Estimating the contribution of the North Atlantic atmospheric dynamics (e.g., the NAO) to these recent trends is essential, as it can provide insights to disentangle the forced and internal components of the European warming. It requires a precise quantification of the relationship between a given atmospheric circulation and the associated temperature on a daily basis across all seasons.

Various methods have been proposed for this purpose, such as the circulation analogs method introduced by Lorenz (1969). This method consists in estimating the temperature of a given day on the basis of temperatures observed during days in the past with the most similar atmospheric circulations (analog days). This method named "nearest neighbors" has already been used by Cattiaux et al. (2010, 2012), Vautard and Yiou (2009) to study the

10.1029/2024GL113540

dynamical contribution to long-term trends and extreme events in European temperatures. A known limitation is that the nearest neighbors or closest analogs always differ, even slightly, from the situation studied, so that temperature reconstructions are imperfect. Variations of the analogs method have been developed to address this issue, such as the reconstructed analogs approach which consists of forming a pool of idealized analogs constructed from the nearest neighbors themselves. Deser et al. (2016) perform a so-called dynamical adjustment on winter temperature trends over North America by linearly combining these analogs, in order to quantify the atmospheric dynamics contribution treated as internal-driven variability and better assess anthropogenically forced long-term signals. This study, carried out on a monthly basis, was then adapted by Terray (2021) at daily timescales to study extreme events in Europe, such as the 2009–2010 cold European winter and the 2010 Russian heat wave. Sippel et al. (2019) propose an improvement of this dynamical adjustment by linear regression between Sea Level Pressure (SLP) and temperature or precipitation, using the regularization linear model which makes the statistical model more robust to noise.

A novel alternative, motivated by the rapid advances in the field of artificial intelligence (AI), is to quantify the circulation-temperature relationship using a neural network. The large amount of data available in climate and weather numerical models makes deep learning a promising candidate for this purpose. Choices between different types of neural networks are specific to the targeted problems (Reichstein et al., 2019). Convolutional neural networks (CNNs) are proven to be effective to recognize objects and patterns in images (maps in this study). They are used for example, to predict ENSO events in the field of seasonal forecast (Ham et al., 2019), to differentiate internal and forced climate variability from air surface temperature observations (Bône et al., 2024) or to perform climate downscaling (Doury et al., 2023). They have also been used to predict North America regional average daily maximum temperature, using several atmospheric variables including SLP, soil moisture and geopotential height from general circulation models (Trok et al., 2024). This last study shows very effective skill with an average $R^2$ of 0.96 with CNNs trained on CanESM5 and tested on ERA5. However, CNNs have never been used to determine the contribution of atmospheric circulation, from SLP only to European temperature variations.

In this paper, our objective is to explore an AI approach based on neural networks to reconstruct daily temperature anomalies in Europe based on the knowledge of the North-Atlantic atmospheric circulation only, and to compare the performance of this reconstruction with the analogs approach traditionally used in the literature. Our analysis is performed in a so-called perfect model framework, using a large ensemble of simulations to have a large set of circulation and temperature data for training algorithms. We characterize the circulation with daily SLP maps, and we chose the UNET, a fully CNN introduced by Ronneberger et al. (2015) for biomedical image segmentation. Methods and data are detailed in Section 2. The results of the comparison between UNET and analogs are presented in Section 3, followed by sensitivity tests to further explore the robustness of results obtained with UNET. Finally, Section 4 provides conclusions and discusses the perspectives of our study.

## 2. Data and Methods

### 2.1. Data

We use daily SLP and surface air temperature maps from the large ensemble MIROC6 (Shiogama et al., 2023; Tatebe et al., 2019), containing 50 ensemble members between 1880 and 2100, with historical forcings until 2014, followed by SSP5-8.5 scenario ones (O'Neill et al., 2016). We divide the large ensemble into 40 members to form the learning sample on which we train the UNET or in which we search for the circulation analogs and the rest, that is, 10 members used as test sample to be reconstructed over 1950–2022, that is, typically the period where observations are available. Data are given on a 1.4° × 1.4° grid and temperature anomalies are reconstructed at each grid point over a domain whose geographical boundaries are 28°N–72 N and 11°W–32°E to include all of Europe. The SLP domain covers a large portion of the North Atlantic, from the North America eastern coast to Europe included and from the subtropics to polar regions (28°N–72°N and 87°W−32°E) (see Figure S1 in Supporting Information S1). Using such a large domain is common practice in analogs-based methods as it captures the main centers of action of the large-scale atmospheric dynamics. In Section 3.2 we also consider smaller SLP domains for our AI-based approach.

### 2.2. Computation of Temperature Anomalies

To obtain daily temperature anomalies, we remove an estimate of the daily climatological normal from the raw temperature value. Climate normals are non-stationary to account for anthropogenically forced warming over the



19448007, 2025, 9, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL113540 by Cochrane France, Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

period of interest (1880–2100) (Rigal et al., 2019). More specifically, the anomalous temperature, Ta, of a given day ($d$) and a given year ($y$) can be written as

$$T_a(d,y) = T(d,y) - (f(d) + g(y) \times h(d)) \tag{1}$$

with the second term representing the so-called non-stationary normal. f(d) stands for the daily seasonal cycle estimated without climate change and $g(y)$ $h(d)$ describes the effects of climate change ($g$ represents the yearly forced response and $h$ the distortion of the annual cycle). First, we estimate $g$ from the 50-member ensemble mean of annual mean modeled temperature. Then, $f$ and $h$ are estimated by linear regression and periodic smoothing (details in Rigal et al. (2019)). We have tested different degrees of freedom for smoothing and arbitrarily chosen 12 for $g$ (), 20 for $f$ () and 24 for $h$ (), to capture the amplitude of seasonal variations and reduce the high frequency noise. We verify that our results do not strongly depend on these choices.

### 2.3. Analogs Method

The analogs method consists of reconstructing a temperature anomaly from a raw SLP field. For a given day ($d$) of a given testing member, the temperature anomaly Ta($d$) is reconstructed by averaging the temperature anomalies of the 20 days whose SLP pattern is the closest to SLP($d$) based on Euclidean distance as similarity criterion. This methodology thus accounts for the amplitude of the spatial patterns and gradients of the SLP field, which are important from a synoptic perspective. Analogs are searched among the 40 training members in a calendar neighborhood to account for the seasonality of the relationship between circulation and temperature anomalies (Cassou & Cattiaux, 2016) and we here use a 31 day window centered on $d$. As a result, for a given day, analog days are searched in a 274,040 day collection (31 days × 221 years × 40 members).

### 2.4. UNET

We use a UNET, a CNN (Ronneberger et al., 2015), to estimate the transfer function F between daily temperature anomaly and SLP map

$$T_a = F(\text{SLP}) \tag{2}$$

The name UNET comes from its $U$ shape (See Figure S2 in Supporting Information S1). The descending part of the $U$ corresponds to the encoder. It consists of convolution, activation and max-pooling operations which reduce the spatial dimension of the input image (here a daily map of SLP) and increase the pixel depth. The ascending part of the $U$ corresponds to the decoder. It uses transposed convolutions, concatenations and convolutions to decode the encoded information and reconstruct the image (here a daily map of temperature anomaly). Connections are established between the encoder and the decoder to preserve the spatial information that can be lost during the encoding process. UNET has been documented to be well suited to deal with spatial and instantaneous fields as detailed in Murphy (2022). In this study we adapt the code developed by (Doury et al., 2023) for downscaling purposes for the construction of our UNET to link temperature and pressure fields.

To train the UNET, we randomly select 80% of the data (hereafter training sample) from the 40 training members. The other 20% (validation sample) is used to monitor the performance of the model on unseen data during training, allowing parameter (learning rate, number of epochs) adjustments to prevent overfitting. The network optimizes its parameters by learning the relationship between a normalized SLP map and the associated temperature anomaly image. Similarly to the analogs, the network learns the relationship between a raw SLP map and the associated temperature anomaly. But for computational reasons, the first step of the UNET is to normalize the input data, as it facilitates convergence in the optimization process. Here, to normalize the pressure maps, we calculate at each grid point the mean and standard deviation of the daily time series concatenated across all time steps and all 40 training members. For the training process we choose a batch size of 1,000 and the optimization process uses the Adam optimizer (Murphy, 2022) with a learning rate of 0.001. The number of epochs is set to 100, but using the Early Stopping criterion (stop training after 10 epochs with no improvement in validation loss), training stops once the model's performance on the validation dataset stops improving. The loss used for the training is the Mean Squared Error (MSE).



19448007, 2025, 9, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL113540 by Cochrane France, Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

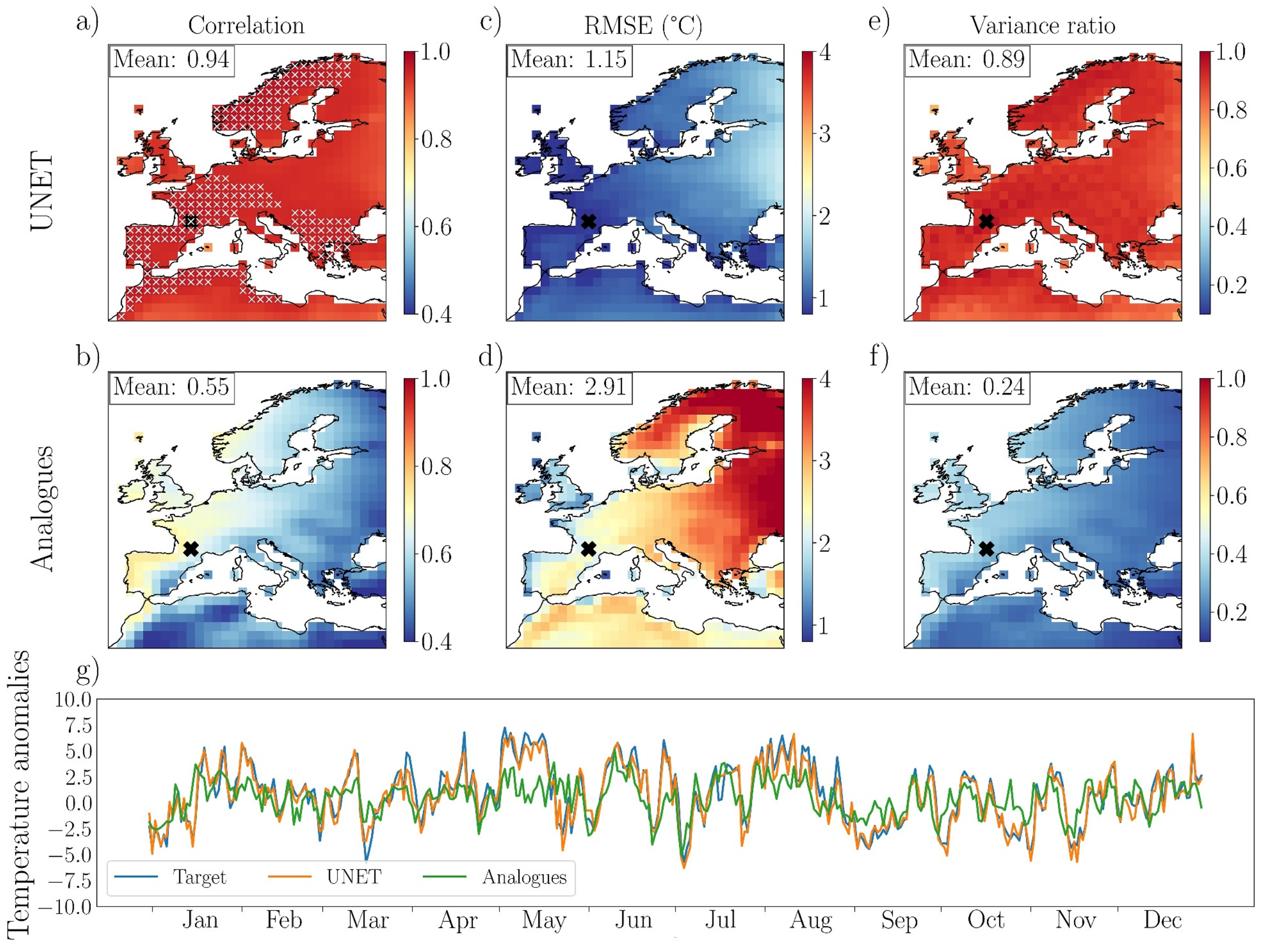


**Figure 1.** (a) and (b) Correlations (c) and (d) Root Mean Squared Error (RMSE), (e) and (f) variance ratio for reconstructions of daily temperature anomalies based on UNET (top) and analogs (bottom) over 1950–2022. White crosses correspond to correlations higher than 0.95. The average of scores of all grid points over the domain are shown in the top left corner of each figure, and (g) temperature anomalies reconstructed by the UNET (in orange), by the analogs (in green) and the true series (in blue) for 1 year of one member, on the black point represented on the maps. Scores calculated for these series are: $r = 0.96$, RMSE = 0.8, $v = 0.95$ for the UNET and $r = 0.63$, RMSE = 2.23, $v = 0.35$ for the Analogs.

Once the UNET has been trained, we test its skill to reconstruct temperature anomalies based on the daily SLP (also normalized with the mean and standard deviation maps previously calculated on the training members) from 1950 to 2022 of each of the 10 testing members.

### 2.5. Evaluation Metrics

To intercompare the reconstruction of the daily fields of temperature anomalies for the 10 ensemble members given by the analogs and UNET methods described above, we use the Root Mean Squared Error (RMSE), the Pearson correlation ($r$) and the variance ratio ($v$) as evaluation metrics. The variance ratio corresponds to the variance of the predicted series divided by the variance of the true series. In this paper, the scores are shown either at each grid point or at one specific grid point located in the South West of France (44.12°N–1.406°E) corresponding to the city of Toulouse in order to illustrate the quality of reconstructions on a local scale.

## 3. Results

### 3.1. Skill of UNET Versus Analogs

UNET provides excellent performances in reconstructing daily temperature anomalies from SLP maps only (Figure 1). Over the whole Europe, there are no points with correlation between original and reconstructed fields lower than 0.9, reaching over 0.95 in many areas such as France, Iberia, Scandinavia and the Balkans. Over a large



19448007, 2025, 9, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL113540 by Cochrane France, Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

19448007, 2025, 9, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL113540 by Cochrane France, Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

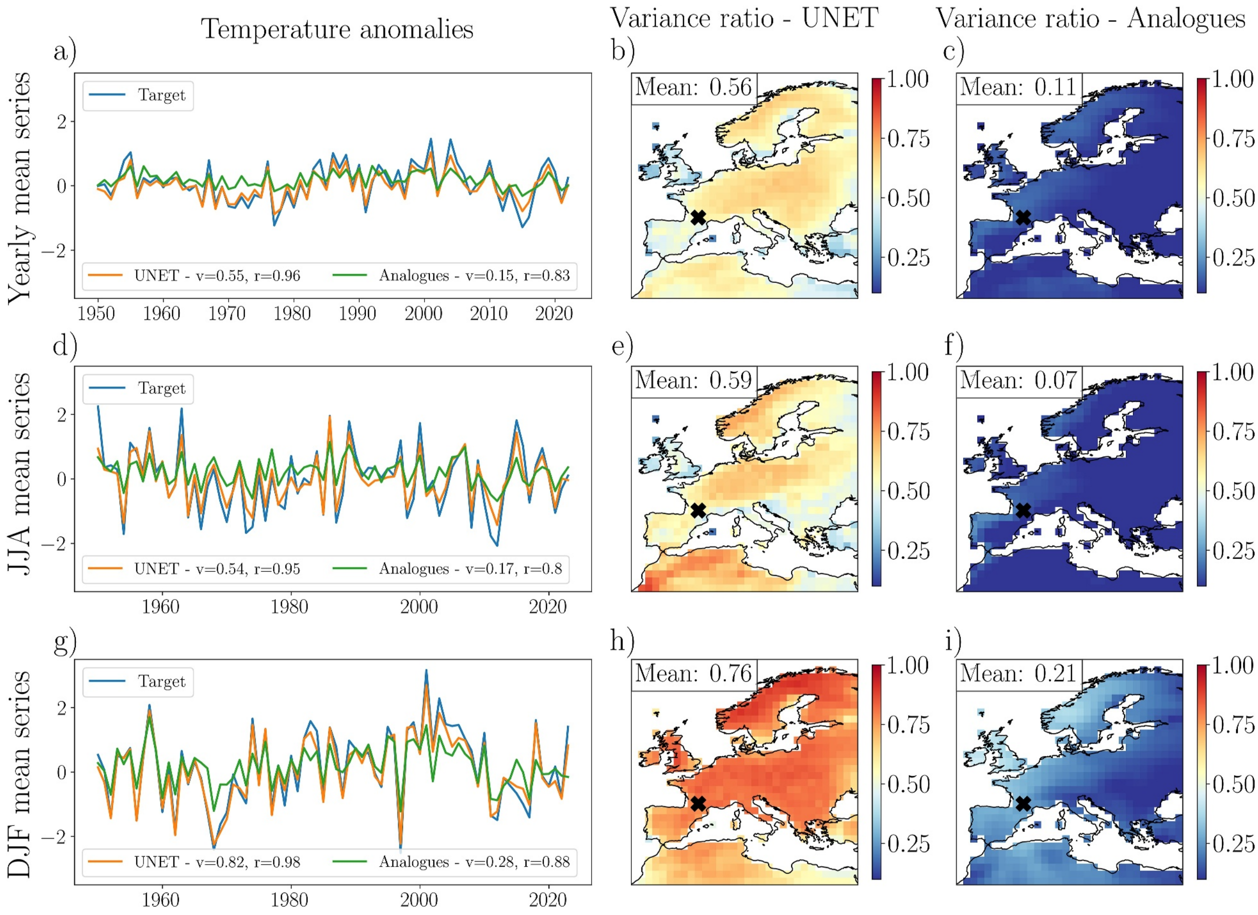


**Figure 2.** (a) Yearly mean of daily reconstructions of one member on the black point (city of Toulouse), with the variance ratio ($v =$ ) and the correlation ($r =$ ) with the target time series indicated, (b) and (c) variance ratio for yearly mean of the daily series from 10 test members reconstructed by the UNET and the analogs. (d) JJA means the daily reconstructions of one member on the black point, (e) and (f) variance ratio for JJA means of the daily series from the 10 test members reconstructed by the UNET and the analogs, and (g), (h), (i) same as (d), (e), (f) for DJF.

part of the continent, variance ratios are around 0.9. Despite overall loss of variance in the reconstruction process (variance ratios are below 1 everywhere), the overall high values for the variance ratios suggest that the UNET can reproduce a large spectrum of the temperature temporal variability, including the daily extremes. For instance, at the grid point of Toulouse, we find that the UNET's RMSEs computed respectively on the subset of the 10% hottest anomalies in summer (JJA), and the subset of the 10% coldest anomalies in winter (DJF) are 0.92°C and 0.95°C, compared with 0.80°C when all days are considered. Note that there is a slight deterioration of the scores toward the East of the domain in all evaluation metrics, especially in RMSE, with errors close to 1.7°C as opposed to 0.8°C along the Atlantic coast of Western Europe.

There is no guarantee that a correct reconstruction of daily anomalies leads to good skill on a seasonal to interannual timescale. If the daily errors are correlated one another, that is, if warm or cold biases persist over several months, then they will not average out to zero when aggregating over the season or the year, and so the seasonal mean can exhibit errors relatively large compared to its own variability. Variance ratios only peaks around 0.6 when computed from annual mean. This decrease in scores compared to daily series is also observed for seasonal averages (winter and summer). At continental scale, the variance is better captured by the UNET in winter (DJF) than in summer (JJA) (Figures 2d and 2e, 2g, 2h). Regionally per season, the errors gradually increase from west to east in winter, while in summer the highest RMSE values are found mainly over southern Europe (Figure S3 in Supporting Information S1). This spatial distribution of the skill can be related to the relative weight between influences from atmospheric dynamics and influences from other processes contributing to temperature variability, which are independent of SLP and are not taken into account by the UNET by construction. It is well established that winter variability is more controlled by heat and moisture transport by

19448007, 2025, 9, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL113540 by Cochrane France, Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

circulation than summer variability where local processes play a greater role (Hurrell et al., 2003). Similarly, the western flank of Europe is more dynamically driven by the fluctuation of the mean westerlies than in Eastern Europe where surface feedbacks tend to dominate (soil moisture, snow cover). Sea surface temperature anomalies are also expected to play a role along the coast and in the Mediterranean basin. We have also ignored the atmospheric circulation of the previous days which can introduce some persistence, each day being considered independent in our reconstruction chain.

Evidence is provided for much lower skill using analogs versus UNET (Figure 1). Best daily reconstructions are obtained in western Europe, with maximum correlations reaching 0.70 versus 0.97 with UNET. Errors range from 1.5°C on the West up to 4.2°C over the Easternmost part of the domain, namely doubled on average compared to UNET. Best variance ratios are also located in western Europe but only reach 0.4 at most. Averaging the 20 analogs results expectedly in a smoothing of the variance, hence a misrepresentation of the actual daily temperature variability, despite the large training sample size used in this study. Interestingly, similar spatial distribution of skill found in UNET for seasonal and interannual reconstructions are also obtained with the analogs, but with considerably lower scores (Figure 2). Variance is not well represented by the analogs, especially in JJA (0.07 for the spatial mean of the variance ratio) despite correlations reaching 0.8 in western/north-western Europe (Figure S3 in Supporting Information S1).

It is noticeable that we do not give UNET any information about the time of year. The only input is the daily SLP field, and since the SLP-temperature relationship is season-dependent (Cassou & Cattiaux, 2016), the performance of the UNET suggests that it may learn other information, such as the calendar day, enabling it to account for the seasonality of temperature changes to similar daily atmospheric circulation across the year. In the following, we carry out different tests on the UNET to better understand the information (a) used in the learning process, and (b) learned by the network. First, we work on the sensitivity of the reconstruction skill to the choice of spatial SLP domains as input for UNET. Then we investigate the capacity of the UNET to predict other indicators, such as the belonging to a given season and the calendar day or the year.

### 3.2. UNET Sensitivity to Input Domain

To assess the sensitivity of the UNET skill to the spatial extension of SLP data, we use the same structure of the UNET as above, but trained with different SLP domains as input, keeping unchanged the output domain for temperature reconstruction. CTL is the control experiment described above, EUR has the same input SLP domain as the output temperature domain (28°N–72°N and 11°W–32°E), that is, Europe only. In NATL, the input SLP domain is fully disjoint from the output temperature domain and corresponds to the North Atlantic only (28°N–72°N and 54°W–11°W). Split domains are represented in Figure S1 in Supporting Information S1.

The correlations, RMSEs and variance ratios show that the UNET only needs local information to reconstruct temperatures. The scores for EUR are very close to CTL (spatial average of RMSE = 1.15°C for CTL and 1.18°C for EUR). This suggests that the network does not need to learn the information from the SLP over the Atlantic to robustly reconstruct temperatures in Europe. This is confirmed by NATL: when the input domain does not include the output domain, scores are significantly lower (but still better than the analogs approach). Table S1 in Supporting Information S1 shows the average scores (RMSE, correlation and variance ratio) of all the grid points and the scores at the specific grid point 44.12°N–1.406°E (Figure S4 in Supporting Information S1). Note though that skills from UNET are always greater than those obtained from the analogs whatever the selection of spatial domain in inputs.

Usually, a large North Atlantic domain is studied to take into account the large-scale atmospheric circulation captured by the modes of variability (NAO, blocking, etc.) and their remote influence over Europe. Our results suggest that the information of the atmospheric circulation outside the restricted European domain selected for reconstruction is not necessary for the neural network. Collocated SLP and temperature fields seems to be enough. This is a typical feature of UNET structures, which tend to prioritize the local information contained in the input data when predicting the output at a given location (González-Abad et al., 2023). In this case, it may be surprising that the UNET performs well even without direct information on the position and intensity of the centers of the North Atlantic circulation, but (a) this information may be partially contained in the pressure maps restricted to Europe, and (b) this could indicate that the temperature variations at a grid point depend mainly on the local pressure gradients around that point, especially when working with instantaneous maps, without any lag. Note that the spatial distribution of skill remains the same in all sensitivity experiments.

10.1029/2024GL113540

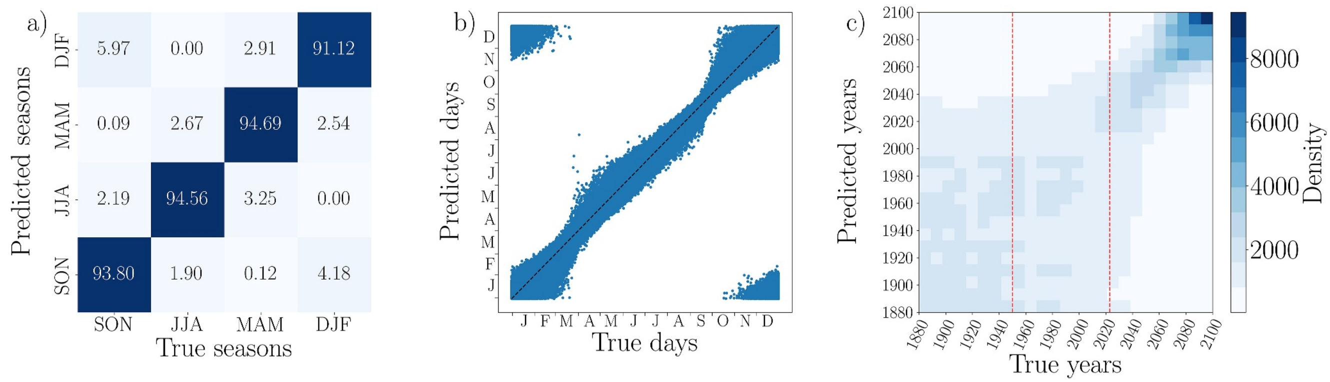


**Figure 3.** (a) Contingency table for the prediction of the seasons for test members, (b) predicted days versus true days, and (c) density of predicted years in relation to actual years. The intensity of the colors indicates the number of predictions in each bin. Red lines represent the period we are reconstructing. For the three graphs, predictions are made over the period 1880–2100.

### 3.3. UNET Interpretability: Seasonal Cycle and Timeline

The influence of the SLP variability on European temperature anomalies depends on the seasons (Cassou & Cattiaux, 2016). The high scores from the UNET obtained with all-calendar days concatenated, suggest that the network is able to learn the seasonality information during the training.

To document this skill, we take the encoder from the UNET already trained to predict the temperature anomaly from the SLP (Section 1.4.) and add one trainable dense layer and a softmax activation function to its convolutional and max pooling layers. In machine learning, the softmax function is used for classification problems, and transforms values into probabilities. We have then built an alternative network designed to predict the probabilities of a day to belong to a season based on the sole knowledge of the SLP daily field map of that day used for training UNET. In other words, the parameters of the pre-trained encoder are thus frozen and only those of the last dense layer are trainable. This new model is trained with only one member chosen randomly among the 40 available training ones: the SLP map as input and the probability vector for the associated season as output. We checked the sensitivity of the results to the training member and the results were again very similar and did not change the conclusion of this study. For example, if the input map corresponds to a winter (DJF) day, the output vector will be (1, 0, 0, 0), if it is a spring (MAM) day the vector will be (0, 1, 0, 0), etc. In this problem the loss function is a categorical cross entropy. By testing this model on the test members, the results show that the season is predicted with remarkable accuracy (Figure 3a): more than 90% good outcome is obtained for the four seasons. Note that the UNET never confuses winter with summer seasons in the daily attribution.

A further step is to predict the day of the year. To do so, the encoder part remains unchanged and, two real periodic numbers are now predicted $\left(\cos\left(\frac{2\pi d}{365}\right), \sin\left(\frac{2\pi d}{365}\right)\right)$ with $d$ ranging from 1 to 365. The last activation function is linear and the loss function is the MSE. The calendar days are strikingly well predicted on the test members (Figure 3b). On average over all days and members, the error between the actual day and the predicted day, computed as the absolute difference between them, is about 8 days. Note that errors in attribution are even smaller in the MAM and SON seasons that can be understood as transition seasons between wintertime and summertime dynamics.

Finally, we examine the skill in predicting the year of the calendar day to be reconstructed, that is, the ability of the UNET to predict the year associated with a daily SLP map. We use the same structure as the classification model that is used to predict the season, but the size of the output vector is 221 over 1880–2100. The neural network fails to clearly identify the years of the days until the middle of the 21st century (Figure 3c) as opposed to the end of the period. This is consistent with the fact that long-term anthropogenically forced trends in the North Atlantic SLP are not detectable over the historical period (including the period we are reconstructing, shown in red in Figure 3), that is, is hardly distinguishable from internal variability (IPCC AR6 WG1 Chapter 3) (Simpson et al., 2018). This supports the decision not to detrend the SLP, assuming that the forced signal in the SLP is rather small over the historical period. The change in performance of the UNET from the mid-century is indicative that the neural network detects a change in the pressure/temperature relationship. This could be related to the emergence of a



19448007, 2025, 9, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL113540 by Cochrane France, Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

climate change signature in the mean SLP and/or variability at high warming levels in the MIROC6 ensemble (e.g., shift of the storm tracks, etc.). Further analyses would be needed to disentangle the processes at the origin of the change in UNET skill in year prediction, however this does not affect the reconstruction of temperatures over the period 1950–2022.

## 4. Conclusion and Discussion

We here provide evidence for an excellent ability of AI to reconstruct temperature variations over Europe at all timescales from the knowledge of the sole daily atmospheric circulation, in a perfect model framework. Our novel method outperforms the traditionally used analogs framework impaired by mismatch between selected analogs and targeted situations, even with the large dataset used in this study, and by the loss of variance in the reconstructed temperature anomalies. Results are best at daily timescale and degraded at seasonal and interannual timescale, possibly due to slow physical processes and variables with memory or inertia that also contribute to drive temperature variations beyond synoptic time scales and that are not captured by UNET by construction. We further find that the UNET primarily uses local SLP information to predict the temperature anomaly, and that it is able to retrieve information about the season, and even the day of the year, during its learning process. This all together contributes to explain the good scores obtained by the UNET and to build confidence in the overall reconstruction framework.

This work should be interpreted as a first attempt to use AI to isolate the contribution of atmospheric circulation to daily and yearly temperature variations in Europe. Our network seems to perform better than traditional methods at reconstructing a daily temperature anomaly from the SLP map but note that we didn't compare the results with recent reconstructed analogs methods (like Deser et al. (2016) or Terray (2021)). Note also that we have used a UNET architecture adapted from a similar application (Doury et al., 2023), which has not been specifically optimized, nor tuned to address the present question. Based on these sensitivity tests, we finally chose an architecture that worked well and gave good results for our initial problem. There is certainly room for improvement, but an exhaustive sensitivity analysis to further tune our UNET is beyond the scope of this paper.

Our finding opens up a wide range of possible applications to understand past observed trends or future expected climate over Europe. For instance, the question of the role of the atmospheric circulation in the particularly rapid warming observed in Europe over the last two decades could be tackled through our framework. One challenge will be that there is much less data in the observed time series than in our perfect model framework (basically one 74-year realization from 1950 to 2024 instead of fifty 221-year ones). The scores will certainly be affected by this large reduction in sample size: the non-stationary normal of the temperature used to derive the anomalies will also be more difficult to accurately estimate, and it will be harder to train and test the UNET on a single member when reasoning in ensemblist approach. An alternative option would be to train a UNET on a large collection of simulations (from one or several models) and apply it to the observed data, similarly to Bodnar et al. (2024) for daily forecasts.

Further prospects include the application of this technique to isolate the contribution of various atmospheric modes of variability to surface temperature variations and trends in other areas of the globe, or at the global scale. This requires the identification of the relevant predictors (variables and spatial domains). Here we have arbitrarily chosen the SLP and the North Atlantic basin, but this is clearly another area for improvement. Incorporating a more systematic use of interpretability methods, such as proposed by González-Abad et al. (2023), in the UNET construction process, could help for this purpose.

## Data Availability Statement

Python scripts for computing the analogues, for training and testing the CNNs and for visualising the results are available on GitHub: https://github.com/carioue/UNET-vs-Analogues. The notebook used to visualise the figures has been adapted to plot the results of only one reconstruction (rather than 10 as described in the paper) and the data required are available on the Zenodo archive https://zenodo.org/uploads/14046994 (Cariou, 2024). Non-stationary normals are computed using routines from https://gitlab.com/ribesaurelien/france_study.



19448007, 2025, 9, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL113540 by Cochrane France, Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

**Acknowledgments**

The authors acknowledge climate modeling groups involved in CMIP6 for producting and making available their simulations, especially for this study, MIROC6 simulations. Analyses and figures are produced with Python.

## References

Bodnar, C., Bruinsma, W. P., Lucic, A., Stanley, M., Brandstetter, J., Garvan, P., et al. (2024). Aurora: A foundation model of the atmosphere. *No. arXiv:2405.13063*. arXiv. Retrieved from http://arxiv.org/abs/2405.13063

Bône, C., Gastineau, G., Thiria, S., Gallinari, P., & Mejia, C. (2024). Separation of internal and forced variability of climate using a u-net. *Journal of Advances in Modeling Earth Systems*, *16*(6), e2023MS003964. Retrieved from https://doi.org/10.1029/2023MS003964

Cariou, E. (2024). UNET-vs-analogues_data [Dataset]. *Zenodo*. Retrieved from https://doi.org/10.5281/zenodo.14046994

Cassou, C., & Cattiaux, J. (2016). Disruption of the european climate seasonal clock in a warming world. *Nature Climate Change*, *6*(6), 589–594. Retrieved from https://doi.org/10.1038/nclimate2969

Cattiaux, J., Vautard, R., Cassou, C., Yiou, P., Masson-Delmotte, V., & Codron, F. (2010). Winter 2010 in europe: A cold extreme in a warming climate. *Geophysical Research Letters*, *37*(20). Retrieved from https://doi.org/10.1029/2010GL044613

Cattiaux, J., Yiou, P., & Vautard, R. (2012). Dynamics of future seasonal temperature trends and extremes in europe: A multi-model analysis from CMIP3. *Climate Dynamics*, *38*(9), 1949–1964. Retrieved from https://doi.org/10.1007/s00382-011-1211-1

Deser, C., Terray, L., & Phillips, A. S. (2016). Forced and internal components of winter air temperature trends over north America during the past 50 years: Mechanisms and implications. *Journal of Climate*, *29*(6), 2237–2258. Retrieved from https://doi.org/10.1175/JCLI-D-15-0304.1

Doury, A., Somot, S., Gadat, S., Ribes, A., & Corre, L. (2023). Regional climate model emulator based on deep learning: Concept and first evaluation of a novel hybrid downscaling approach. *Climate Dynamics*, *60*(5), 1751–1779. Retrieved from https://doi.org/10.1007/s00382-022-06343-9

González-Abad, J., Baño-Medina, J., & Gutiérrez, J. M. (2023). Using explainability to inform statistical downscaling based on deep learning beyond standard validation approaches. *Journal of Advances in Modeling Earth Systems*, *15*(11), e2023MS003641. Retrieved from https://doi.org/10.1029/2023MS003641

Ham, Y.-G., Kim, J.-H., & Luo, J.-J. (2019). Deep learning for multi-year ENSO forecasts. *Nature*, *573*(7775), 568–572. Retrieved from https://doi.org/10.1038/s41586-019-1559-7

Hurrell, J. W., Kushnir, Y., Ottersen, G., & Visbeck, M. (2003). An overview of the north atlantic oscillation. In J. W. Hurrell, Y. Kushnir, G. Ottersen, & M. Visbeck (Eds.), *Geophysical monograph series* (Vol. 134, pp. 1–35). American Geophysical Union. Retrieved from https://doi.org/10.1029/134GM01

Lorenz, E. N. (1969). Atmospheric predictability as revealed by naturally occurring analogues. *Journal of the Atmospheric Sciences*, *26*(4), 636–646. Retrieved from https://doi.org/10.1175/1520-0469(1969)26⟨636:APARBN⟩2.0.CO;2

Murphy, K. P. (2022). *Probabilistic machine learning: An introduction*. MIT Press.

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., et al. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. Retrieved from https://doi.org/10.5194/gmd-9-3461-2016

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., & Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven. *Earth System Science*, *566*(7743), 195–204. Retrieved from https://doi.org/10.1038/s41586-019-0912-1

Rigal, A., Azaïs, J.-M., & Ribes, A. (2019). Estimating daily climatological normals in a changing climate. *Climate Dynamics*, *53*(1), 275–286. Retrieved from https://doi.org/10.1007/s00382-018-4584-6

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation (No. arXiv:1505.04597). *arXiv*. Retrieved from http://arxiv.org/abs/1505.04597

Schumacher, D. L., Singh, J., Hauser, M., Fischer, E. M., Wild, M., & Seneviratne, S. I. (2024). Exacerbated summer european warming not captured by climate models neglecting long-term aerosol changes. *Communications Earth and Environment*, *5*(1), 1–14. Retrieved from https://doi.org/10.1038/s43247-024-01332-8

Shiogama, H., Tatebe, H., Hayashi, M., Abe, M., Arai, M., Koyama, H., et al. (2023). MIROC6 large ensemble (MIROC6-LE): Experimental design and initial analyses. *Earth System Dynamics*, *14*(6), 1107–1124. Retrieved from https://doi.org/10.5194/esd-14-1107-2023

Simpson, I. R., Deser, C., McKinnon, K. A., & Barnes, E. A. (2018). Modeled and observed multidecadal variability in the north atlantic jet stream and its connection to sea surface temperatures. *Journal of Climate*, *31*(20), 8313–8338. Retrieved from https://doi.org/10.1175/JCLI-D-18-0168.1

Sippel, S., Meinshausen, N., Merrifield, A., Lehner, F., Pendergrass, A. G., Fischer, E., & Knutti, R. (2019). Uncovering the forced climate response from a single ensemble member using statistical learning. *Journal of Climate*, *32*(17), 5677–5699. Retrieved from https://doi.org/10.1175/JCLI-D-18-0882.1

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., et al. (2019). Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geoscientific Model Development*, *12*(7), 2727–2765. Retrieved from https://doi.org/10.5194/gmd-12-2727-2019

Terray, L. (2021). A dynamical adjustment perspective on extreme event attribution. *Weather and Climate Dynamics*, *2*(4), 971–989. Retrieved from https://doi.org/10.5194/wcd-2-971-2021

Trok, J. T., Barnes, E. A., Davenport, F. V., & Diffenbaugh, N. S. (2024). Machine learning–based extreme event attribution. *Science Advances*, *10*(34), eadl3242. Retrieved from https://doi.org/10.1126/sciadv.adl3242

Vautard, R., Cattiaux, J., Happé, T., Singh, J., Bonnet, R., Cassou, C., et al. (2023). Heat extremes in western europe increasing faster than simulated due to atmospheric circulation. *Trends*, *14*(1), 6803. Retrieved from https://doi.org/10.1038/s41467-023-42143-3

Vautard, R., & Yiou, P. (2009). Control of recent european surface climate change by atmospheric flow. *Geophysical Research Letters*, *36*(22). Retrieved from https://doi.org/10.1029/2009GL040480



19448007, 2025, 9, Downloaded from https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2024GL113540 by Cochrane France, Wiley Online Library on [24/11/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License