https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


# Understanding European Heatwaves with Variational Autoencoders

Aytaç Paçal[1,2], Birgit Hassler[1], Katja Weigel[2,1], Miguel-Ángel Fernández-Torres[3], Gustau Camps-Valls[4], and Veronika Eyring[1,2]

[1]Deutsches Zentrum für Luft- und Raumfahrt (DLR), Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany
[2]University of Bremen, Institute of Environmental Physics (IUP), Bremen, Germany
[3]Department of Signal Theory and Communications, Universidad Carlos III de Madrid (UC3M), Leganés, Madrid, Spain
[4]Image Processing Laboratory (IPL), Universitat de València (UV), Paterna, València, Spain

**Correspondence:** Aytaç Paçal (aytac.pacal@dlr.de)

**Abstract.** Understanding the dynamics of heatwaves is critical for accurate climate risk assessment. Traditional definitions, based solely on surface temperature thresholds, often overlook the complex, multivariate nature of heatwaves. This study uses a spatiotemporal Variational Autoencoder (VAE), an unsupervised machine learning method, to identify compact representations of multivariate, year-round heatwave patterns. Focusing on key atmospheric variables (e.g., circulation, humidity, temperature, geopotential height, cloud cover, stream function, and radiation), we extract eleven-day heatwave samples from ERA5 reanalysis data over the North Atlantic, centered on near-surface temperature extremes in Western Europe. The VAE was trained on data from 1941–1990 and evaluated using 2001–2022 samples, and effectively clustered heatwave events by season, revealing known dynamical regimes such as summer blocking highs and winter omega blocks. The VAE model captures the interplay and temporal evolution between different atmospheric variables in their contributions to heatwaves over Western Europe. Notably, recent summer heatwaves form a distinct cluster within the latent space, pointing to a shift in atmospheric dynamics consistent with climate change. Composite anomaly maps further show coherent pre-onset patterns across variables. These results demonstrate the potential of VAEs to uncover meaningful structure in complex heatwave dynamics from data, and promise advances in understanding heatwaves.

## 1 Introduction

Extreme events and heatwaves significantly threaten the environment and human health (López-Bueno et al., 2021). Extreme events refer to rare and intense meteorological phenomena, while heatwaves are characterized by exceptionally high temperatures persisting over several days. Since the effects of climate change intensify with the increased amount of human-induced greenhouse gases in the atmosphere (Lynas et al., 2021), the frequency and intensity of hot (cold) extremes increase (decrease) globally (IPCC, 2021). This trend is projected to continue even at the lowest projected global warming scenario, and the intensity of extremes will increase proportionally with the amount of warming. Currently, 30% of the global population is exposed to deadly climatic conditions, which are projected to increase up to 75% under the current greenhouse gas emissions scenarios (Mora et al., 2017b). Prolonged extreme heat conditions over a specific area, i.e., heatwaves, combined with the physiological limits of the human body, can trigger 27 physiological pathways that may ultimately lead to fatal outcomes (Mora et al., 2017a).

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


Furthermore, the increase in extreme event frequency and intensity –including heatwaves, floods, storms, and droughts– combined with existing social inequalities, will disproportionately impact the poorest places, with the frequency of local climate changes projected to be more than twice that observed in the wealthiest areas (King and Harrington, 2018; Dosio et al., 2018; Ronco et al., 2023; Kotz et al., 2024).

The growing severity of extreme events also affects the environment, infrastructure, and economy (Yuan et al., 2024; IPCC, 2021). Recent heatwaves in Europe (2003, 2010, 2015, and 2018) caused damages that amounted to 0.3-0.5% of European gross domestic product (GDP) (García-León et al., 2021). Globally, the annual cost of extreme weather events is estimated at US$143 billion(Newman and Noy, 2023). Looking ahead, global cumulative GDP losses could reach 16-22% by the end of the century if the current 3°C warming trajectory continues (Burke et al., 2015; Newell et al., 2021; Kotz et al., 2024; World Economic Forum, 2024). Therefore, identifying and characterizing heatwaves is crucial to effectively mitigating and adapting to their impacts. Since extreme temperature events, including heatwaves, by definition, are rare occurrences at the tails of temperature distributions (McPhillips et al., 2018), percentile and block maxima methods have been widely used for their detection (Zhang et al., 2011; Huang et al., 2016; Alaya et al., 2020; Li et al., 2021; IPCC, 2021). Russo and Domeisen (2023) showed that historically unprecedented extreme heatwaves increased 10-fold in recent years using cumulative indices for their identification rather than indices relying on temporal averages, as cumulative indices are more reliable for comparing events with different lengths. Brunner and Voigt (2024) found that seasonal running windows commonly used to define extreme thresholds can introduce systematic biases and underestimate the frequency of extreme events. However, the definition of an extreme depends on the application, and there is no standard definition for heatwaves (Barriopedro et al., 2023; Boni et al., 2023). Many existing indices rely solely on temperature, which limits their ability to capture the full complexity of these events (Perkins, 2015; McPhillips et al., 2018; Barriopedro et al., 2023). The complex nature of heatwaves, however, requires more sophisticated analysis methods. Multiple atmospheric processes and local drivers play an essential role in the occurrence of heatwaves, and a multidimensional approach considering the tails of high-dimensional, multivariate probability distributions, whose shape and structure are themselves changing due to non-stationary climate dynamics, is required for defining and studying heatwaves (Sardeshmukh et al., 2015; Domeisen et al., 2022; Barriopedro et al., 2023; Camps-Valls et al., 2025).

Machine learning (ML) and deep learning (DL) methods are frequently used in climate science as they can discover complex relationships in multivariate and high-dimensional climate datasets (Reichstein et al., 2019; Zhu et al., 2023; Salcedo-Sanz et al., 2023; Camps-Valls et al., 2025). ML and DL techniques rely on high-quality and reliable labeled data. However, a major limitation for extreme events and heatwaves is precisely the lack of high-quality and labeled datasets (Prabhat et al., 2021; Lacombe et al., 2023). One potential solution is to use unsupervised learning methods where the model learns patterns from data autonomously. Unsupervised learning methods help identify complex data patterns without relying on predefined labels or thresholds. In Paçal et al. (2023), the Gaussian Mixture Model (GMM), an unsupervised learning approach that models data as a combination of multiple Gaussian distributions, was applied to characterize extreme temperature events. By representing daily maximum temperatures as a bimodal rather than a traditional unimodal distribution, the study found that extreme temperature events will become significantly more frequent under future global warming levels and revealed regional differences not captured by conventional methods. Another example of an unsupervised learning method is the Variational

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


Autoencoder (VAE), composed of an encoder to reduce the dimensionality of the input data to a latent space and a decoder to reconstruct the input data from that latent space (Kingma and Welling, 2019). Recent advances in artificial intelligence for climate extremes have highlighted the importance of unsupervised learning models like GMMs and VAEs for both detecting anomalies and uncovering the mechanisms behind rare and compound events, especially when labeled data is scarce or absent (Materia et al., 2024).

VAEs are predominantly used in anomaly detection in domains such as network security, risk management, health monitoring, and computer vision (Pang et al., 2021; Nassif et al., 2021; Albuquerque Filho et al., 2022). Among other applications, autoencoders and VAEs have been applied to learn spatiotemporal regularities from video data (Hasan et al., 2016; Fan et al., 2020). VAEs and clustering methods, such as GMMs, are also used to identify regimes in the latent space and analyze their dynamical behavior (Lindhe et al., 2021; Happé et al., 2024). Since clustering is also widely used in climate science to categorize and understand different types of heatwave events (Stefanon et al., 2012; Zhao et al., 2024), combining VAEs with clustering provides a promising framework for characterizing complex climate extremes.

This study aims to understand how heatwaves develop and what processes and local phenomena contribute to their evolution. To this end, we first use a spatiotemporal VAE trained on multivariate reanalysis heatwave samples, then cluster and analyze the latent space representation of heatwaves in Western Europe. Section 2 describes the data and methodology followed in this research. Section 3 provides the results for the heatwave cluster analysis. Finally, the main results are summarized and discussed in Section 4.

## 2 Data and Methods

### 2.1 Data Sources and Preprocessing

We used the ERA5 reanalysis dataset (Hersbach et al., 2017; Soci et al., 2024) to detect and learn about heatwave dynamics. We focused only on atmospheric variables, as shown in Table 1, which are known to be relevant to heatwave dynamics (Jézéquel et al., 2018; Domeisen et al., 2022; Rousi et al., 2022; Barriopedro et al., 2023; Kim and Seo, 2023; Tian et al., 2024). All data was accessed via the German Climate Computing Center (DKRZ) data catalog (Hersbach et al., 2017). We aggregated the hourly data in GRIB (General Regularly-distributed Information in Binary form) file format for each variable into daily values using variable-specific temporal operators, such as daily maxima for temperature and daily means for pressure fields, as shown in Table 1. The resulting dataset comprises daily observations from 1940 to 2022 for the North Atlantic region, extending from 75°W to 60°E, and from 30°N to 75°N on a 0.7°grid, as shown in Figure 2. We standardized each variable using the 1941-1980 climatological baseline to remove seasonal variability, since this period precedes the stronger warming trends after 1980 in Europe (Elguindi et al., 2013; Reid et al., 2016). This involved computing the daily climatological mean and standard deviation using a 15-day moving window centered on each calendar day across the baseline years.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


**Table 1.** Overview of the atmospheric variables used for heatwave characterization in this study. The temporal operator refers to the daily aggregation method applied to each variable during preprocessing to convert from hourly to daily frequency. Heatwave samples for training and analysis, constituted by all nine variables (multivariate), are extracted from the broader North Atlantic region, centered temporally around onset dates. ($^\star$) 2-meter temperature (*t2m*) is used to identify heatwave onset dates specifically from the Western Europe subregion. Regions are shown in Figure 2.

| Short Name | Long Name | Temporal operator |
|---|---|---|
| msl | Mean sea level pressure | Mean |
| r | Relative humidity | Mean |
| ssr | Surface net solar radiation | Sum |
| stream250 | Stream function at 250 hPa | Mean |
| t2m$^\star$ | 2-meter temperature | Max |
| tcc | Total cloud cover | Mean |
| u10 | 10-meter U wind component | Mean |
| v10 | 10-meter V wind component | Mean |
| z500 | Geopotential height at 500 hPa | Mean |

### 2.2 Heatwave Identification

We followed the methodology proposed by Happé et al. (2024) to identify heatwaves. The identification process consists of three main steps: (1) threshold-based detection of anomalies at the grid-point level, (2) clustering of spatially and temporally contiguous anomalous grid cells, and (3) extraction of multivariate spatiotemporal samples suitable for model training. The overall workflow for heatwave onset detection and sample extraction is illustrated in Figure 1.

The first step involves detecting anomalies at the level of individual grid cells. We focused only on land grid cells in the Western Europe subregion, extending from 10°W to 55°E and from 35°N to 55°N (see the blue shaded area in Figure 2). Then, we used daily maximum 2-meter temperature data from the ERA5 reanalysis dataset. For each grid cell, we calculated the 90th percentile threshold for each calendar day using a multiyear 15-day moving window centered on each calendar day over the historical baseline period from 1941 to 1980. We identified anomalies each day by comparing the temperature to the corresponding threshold. This approach captured relative temperature anomalies while accounting for seasonality and regional climate norms. The result was a binary mask indicating whether or not a grid cell experienced a day of anomalous heat. A value of 1 for a grid cell marks exceedance of the local threshold, while 0 represents climatologically typical conditions.

Subsequently, during the second step, we applied the Generalized Density-Based Spatial Clustering of Applications with Noise (GDBSCAN) algorithm to detect heatwave events (Sander et al., 1998), using a minimum cluster size of 21 cells across space and time, as suggested by Happé et al. (2024). This procedure identified 2565 unique heatwave onset dates, between 1941 and 2022, from the land grid cells over Western Europe, as shown in Figure 2.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


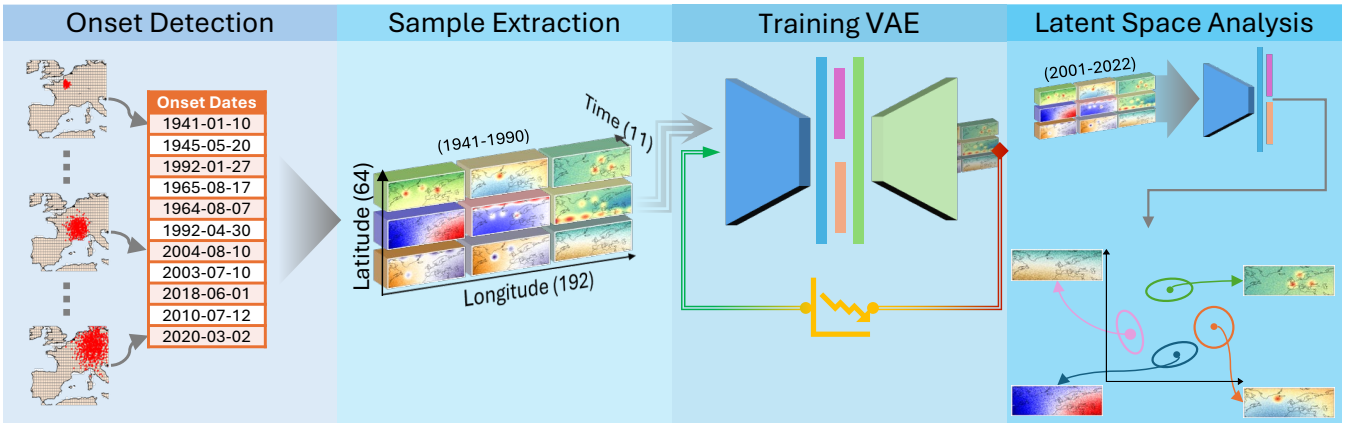


**Figure 1.** Workflow illustration of the heatwave onset detection and sample extraction process. Detected onset dates over Western Europe are used to extract eleven-day multivariate heatwave samples from the North Atlantic region, centered on each event onset. Samples from 1941–1990 are used to train a Variational Autoencoder (VAE) that learns compact representations of the spatiotemporal structure of heatwaves. Subsequently, samples from 2001–2022 are encoded into the latent space to analyze the evolution and dynamics of recent heatwave events.

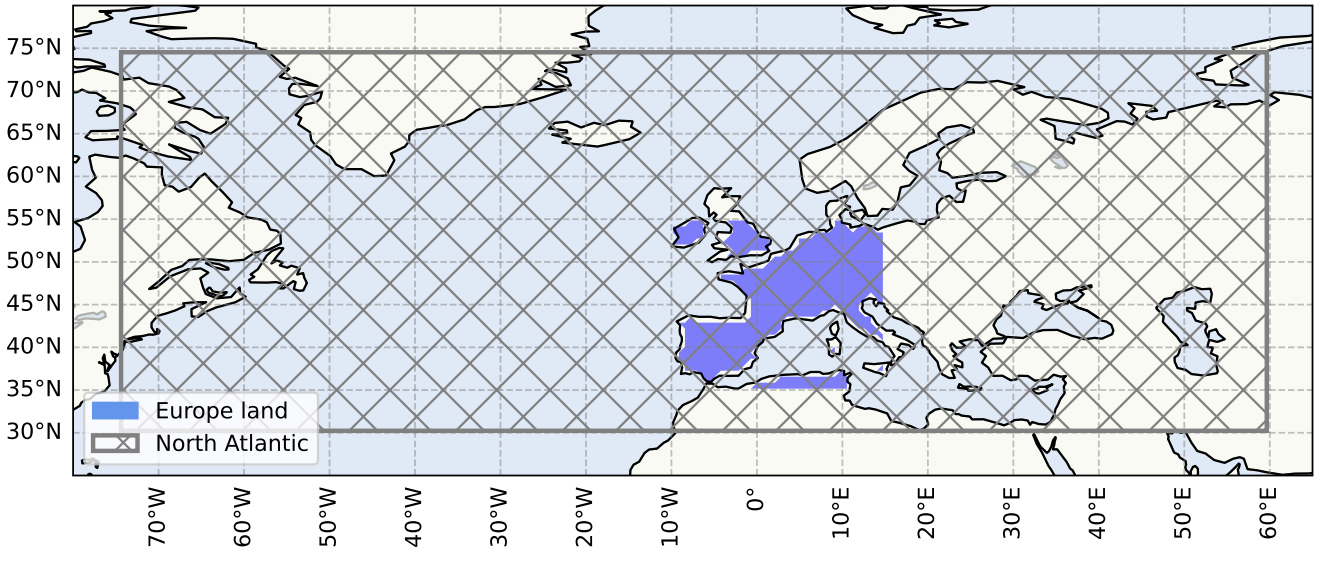


**Figure 2.** Geographical domains used to identify heatwave onset dates and train the VAE model. The shaded blue area represents the Western Europe land grid cells used to identify heatwave onset dates based on hot grid points derived from ERA5 temperature fields. These onset dates are then used to extract 11-day heatwave samples, centered on the onset dates, from the broader North Atlantic region (indicated by the hatched gray box), covering the nine variables (multivariate) listed in Table 1. These multivariate heatwave samples are used for training and analysis.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


Finally, the last step involves extracting heatwave samples using these onset dates, where we fixed 11-day windows centered on each onset. The 11-day window includes the five days before the onset, the onset day itself, and the following five days, similar to Rouges et al. (2023). We focused on nine different atmospheric variables (multivariate), as listed in Table 1. In the first step, we only used land grid cells over Western Europe to define heatwave onset dates. However, the heatwave samples used for training the model and analyses cover a broader spatial domain, namely the extended North Atlantic region, to capture large-scale atmospheric conditions as shown with the hatched gray box in Figure 2. This framing (spatiotemporal) captures both the build-up and evolution of heatwave conditions, enabling the VAE model to learn characteristic patterns associated with the heatwaves.

Happé et al. (2024) used the KNMI-LENTIS dataset (Muntjewerf et al., 2023) to analyze heatwaves. We used here heatwave samples extracted from the ERA5 reanalysis dataset. Furthermore, we extracted the heatwave samples from the whole year, not only the summer months. This makes our approach more general and independent of an arbitrary (and ever-changing) definition of "summer" (Wang et al., 2021). Besides, the data is standardized, so we can analyze the summer heatwaves and the anomalous heat during colder months, i.e., warm spells or winter heatwaves. This approach allows, in turn, more heatwave samples from the ERA5 reanalysis dataset to train the VAE model. An additional difference between our approach and the method applied in (Happé et al., 2024) is that we used additional atmospheric variables to investigate their effect on the occurrence of a heatwave. We also incorporate timesteps before the onset of heatwaves, which allows us to identify atmospheric patterns that build up to and persist during the progression of heatwaves.

### 2.3 Variational Autoencoder

The Variational Autoencoder (VAE) is an unsupervised machine learning method that uses neural networks to generate a latent space representation of the input data to reconstruct the input data (Kingma and Welling, 2019). VAEs have been widely and successfully used in anomaly detection, image recognition, and generative modeling (Pang et al., 2021). VAEs consist of an encoder network and a decoder network. The encoder network takes an input, $x$, such as an image, and compresses it to a latent space representation. This latent space is represented by a probability distribution, generally by a Gaussian. The decoder then takes the representation from the latent space and reconstructs the input, $\hat{x}$. The main goal of VAE is to approximate the reconstructed output as closely as possible to the original input by minimizing the reconstruction loss between the output and input. A VAE summarizes, therefore, complex high-dimensional data into a low-dimensional embedding. This low-dimensional latent space compresses the input data and allows the discovery of hidden patterns within the data. It can be further analyzed using visualization techniques and clustering algorithms to identify patterns and regimes in the data (Happé et al., 2024; Lindhe et al., 2021).

We consider a 3D convolutional neural network architecture for our VAE model to learn multivariate spatiotemporal representations of heatwaves in an unsupervised manner (see Table A1 in the Appendix for the description of the VAE architecture and more details on the training process). By training the model on historical heatwave samples, the VAE learns an internal representation of typical patterns over time, capturing the complex spatial and temporal relationships between the different climate-relevant variables. This allows the model to learn a low-dimensional representation of extreme multivariate events in

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


the latent space. Each heatwave sample consisted of 9 climate variables for 11 days on a $64 \times 192$ spatial grid (North Atlantic, see hatched gray region in Figure 2), resulting in input tensors with the shape $(9, 11, 64, 192)$. We used all heatwave samples from 1941-1990 for training, 1991-2000 for validation, and 2001-2022 for testing, resulting in 1408, 320, and 928 heatwave samples, respectively.

After training the VAE model, we evaluated its ability to reconstruct the climate variables. Table 2 summarizes the $R^2$ scores obtained for each meteorological variable individually and the overall score across all variables for the test period (2001-2022). The model demonstrates good reconstruction skill, with an overall $R^2$ of 0.7235. Notably, the highest $R^2$ values are achieved for geopotential height (*z500*) and mean sea level pressure (*msl*), while relatively lower scores are observed for 2-meter temperature (*t2m*) and 10-meter wind components (*u10* and *v10*). Similarly, Jézéquel et al. (2018) showed that geopotential height (*z500*) better captures temperature anomalies associated with European heatwaves. The relatively lower $R^2$ scores for 2-meter temperature (*t2m*) and 10-meter wind components (*u10* and *v10*) can be attributed to higher spatiotemporal variability in near-surface fields, influenced by local surface conditions, which are difficult for the model to capture accurately, considering the coarse spatial resolution of the data.

**Table 2.** Overall $R^2$ score and per channel for samples in the test subset. The evaluation is based on the input data structured as 928 samples with nine channels, each covering 11 timesteps and a spatial grid of $64 \times 192$ points.

| **Channel** | **$R^2$ Score** |
|---|---|
| 2-meter temperature (*t2m*) | 0.6163 |
| 10-meter zonal wind (*u10*) | 0.6416 |
| 10-meter meridional wind (*v10*) | 0.6567 |
| Total cloud cover (*tcc*) | 0.7100 |
| Surface solar radiation (*ssr*) | 0.7139 |
| Relative humidity (*r*) | 0.7722 |
| Stream function (*stream250*) | 0.7753 |
| Mean sea level pressure (*msl*) | 0.8027 |
| Geopotential height (*z500*) | 0.8226 |
| **All variables** | **0.7235** |

Once the model was trained, we evaluated the latent space by encoding heatwave samples from the test period (2001–2022), which includes 928 heatwave events. After passing the data through the encoder, each event was mapped to a point in the 128-dimensional latent space. To uncover distinct types of heatwaves, we first reduced the dimension of the latent space to 2-dimension using t-distributed stochastic neighbor embedding (t-SNE), which is a statistical method to visualize high-dimensional data (van der Maaten and Hinton, 2008), and then applied GMM clustering to the latent vectors. GMM is an unsupervised learning method that probabilistically clusters data points into different Gaussian components. This technique identifies the latent space's underlying groups or clusters of heatwave events. Each cluster corresponds to a distinct type of heatwave event with similar atmospheric characteristics. We extracted the centroids for each GMM cluster to interpret the

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


resulting clusters. Then, we retrieved the closest 100 heatwave samples to each cluster centroid in the latent space. Using these 100 heatwave samples, we created a composite representative sample of each of the nine climate-relevant variables for each heatwave cluster. This process allows us to analyze the spatial patterns associated with each cluster and the conditions described by these variables, building up to the heatwave onset (Rouges et al., 2023).

# 3 Results

## 3.1 Identification of heatwaves

Figure 3 shows the annual number of heatwave events in Western Europe identified in the ERA5 reanalysis dataset from 1940 to 2022. We observe a gradual increase in heatwave frequency starting in the 1980s, which becomes especially pronounced after the mid-1990s. In the most recent two decades, annual heatwaves have frequently exceeded 40, culminating in a peak of 55 events in 2021, the highest value within the analyzed period. These findings are consistent with other studies, where it has been shown that heatwaves become more frequent (Coumou and Rahmstorf, 2012; Rahmstorf and Coumou, 2011; Perkins-Kirkpatrick and Lewis, 2020; Fischer et al., 2021; IPCC, 2021; Wang et al., 2024; Huntingford et al., 2024). While the overall trend indicates an increase in the frequency of heatwaves, we also find occasional declines in annual counts. These lower counts can be attributed to interannual climate variability and varying meteorological conditions for heatwave formation (Perkins-Kirkpatrick et al., 2017).

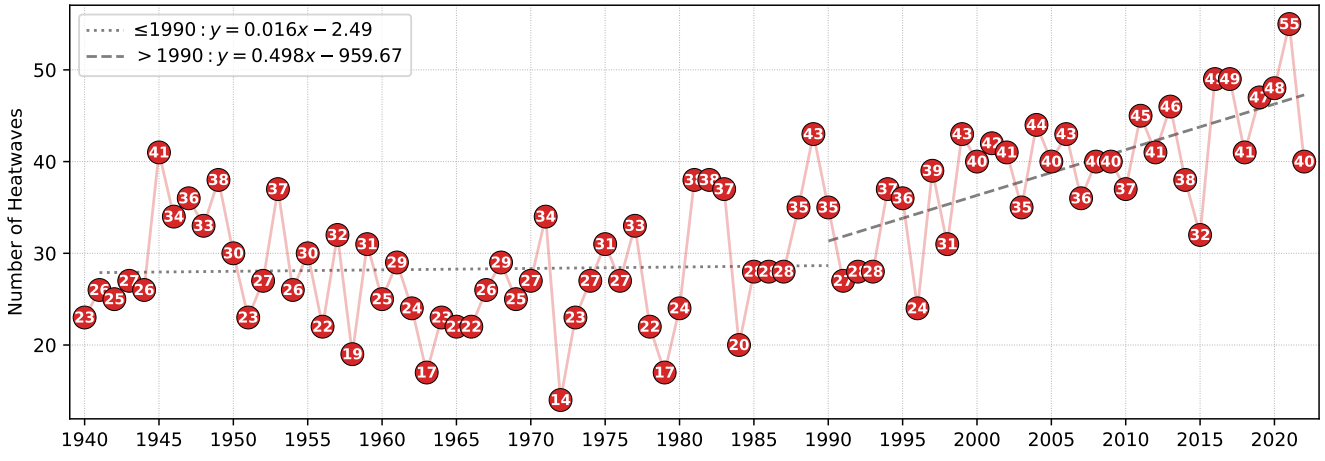


**Figure 3.** Number of detected heatwave events per year in Western Europe, from 1940 to 2022, identified by the GDBSCAN algorithm (Sander et al., 1998) in ERA5 reanalysis data using the grid cells that exceed the 90th percentile with respect to the 1941-1980 daily maximum temperature climatology. This period precedes the stronger warming trends after 1980 in Europe (Elguindi et al., 2013).

After training the model, we first analyzed the latent space representation of heatwave samples. Each multivariate heatwave sample was encoded into a 128-dimensional latent space by the VAE. As an intermediate processing step, we first used Principal

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


Component Analysis (PCA) to reduce the dimensionality to 50 components, as it is recommended to suppress noise and increase computational efficiency (van der Maaten and Hinton, 2008; Pedregosa et al., 2011). Then, we applied the t-SNE algorithm to reduce it to 2 dimensions to visualize the latent space.

Figure 4 shows the t-SNE projections for training (1941-1990), validation (1991–2000), test (2001–2022), and combined (1941-2022) samples. During training (heatwaves from 1941 to 1990), the VAE was exposed to a larger set of heatwave patterns, which were encoded into a latent space. This spread suggests that there was a broad diversity in training examples. The validation period shows a consistent distribution with the training periods, suggesting that the model generalizes well with the validation data (1991-2000). Interestingly, the latent space for the test period (2001-2022) shows a ring-like structure with fewer samples in the center and an accumulation toward the positive values of the first component (see Figure 4c). This shift compared to previous periods might imply that the nature of heatwaves changed compared to historical periods. As heatwaves become more frequent and extreme (Perkins-Kirkpatrick and Lewis, 2020; Fischer et al., 2021; IPCC, 2021; Russo et al., 2015; Lhotka and Kyselý, 2022; Paçal et al., 2023), the VAE encodes these heatwave samples, which are underrepresented in the training period, into regions of the latent space not occupied by training heatwave samples. This subtle drift is clearer in the combined panel in Figure 4d, which suggests a temporal change in heatwave characteristics.

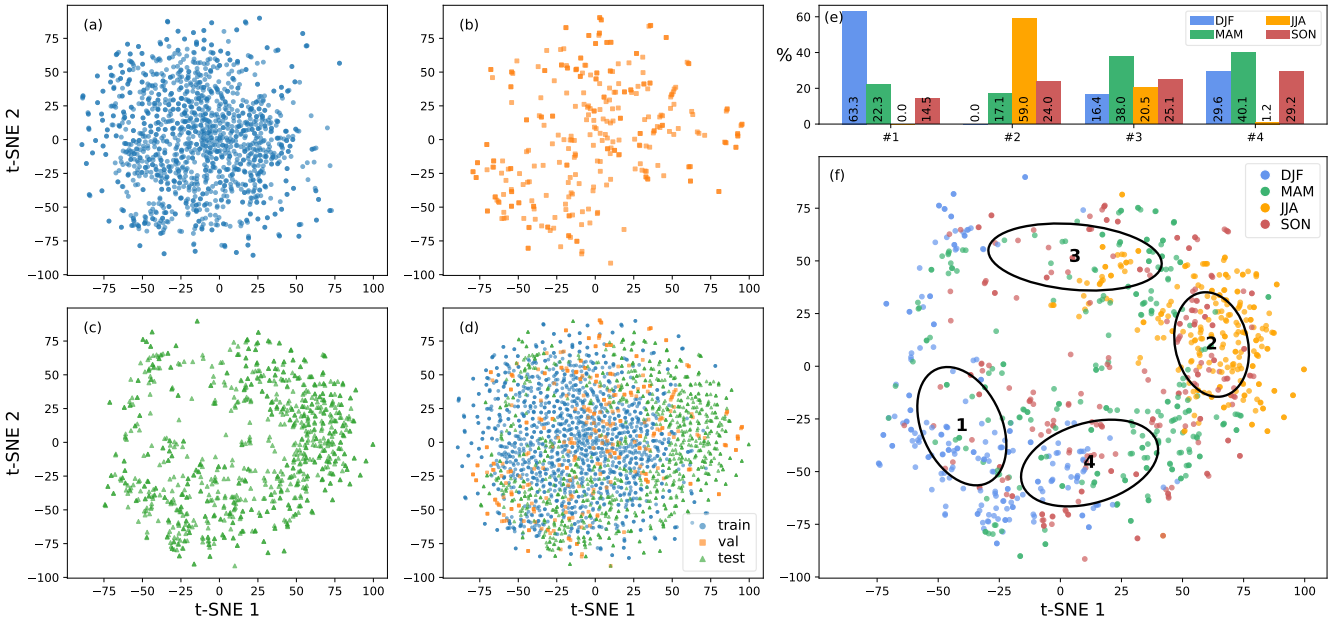


**Figure 4.** t-SNE representation of the latent space for (a) training (1941–1990), (b) validation (1991–2000), (c) test (2001–2022), and (d) the full period (1941–2022). (e) shows the seasonal distribution of samples within each cluster. (f) displays the t-SNE projection of test-period heatwave samples colored by the season of onset, with ellipses representing the four Gaussian Mixture Model components. Each point corresponds to an individual heatwave event.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


We fitted multiple GMM components to the latent space to identify different clusters of heatwaves. After testing the GMM fits with the Bayesian Information Criterion (BIC) scores and cluster stability across different random initializations, we found that four clusters provide the most robust grouping of test samples (see Figure A2 in the Appendix for BIC scores across different components for training, validation and test periods). While the BIC scores favor simpler models for training and validation sets, the test set exhibits a clear minimum score at four components. This balances model complexity and fit well, as BIC values steadily increase beyond four across all datasets. We color-coded heatwave sample representations in the latent space to interpret these clusters according to their corresponding seasons. As shown in Figure 4f, some clusters are dominated by heatwaves happening at specific seasons (e.g., Cluster *#2* with mostly June-July-August (JJA) events), indicating that the model has learned to differentiate multivariate heatwave patterns reflecting seasonal influences. This seasonal structure emerges even though the input data were deseasonalized and standardized. Clusters *#2* and *#3* in Figure 4f, respectively, correspond to nearly the same heatwave samples as those shown in the three-component clustering (see Figure A1 in the Appendix). Since the samples correspond mostly to JJA months in the latent space represented on the right side of the t-SNE plots, this also suggests that these heatwave samples were not very frequent historically, as this aggregation is clearly outside of training period samples as shown in the combined panel in Figure 4. Since four components in Figure 4f provide the possibility of a more detailed analysis for heatwave samples within cluster *#1*, i.e., those corresponding to December-January-February (DJF) winter months, while having a similar cluster *#2* and *#3* as in the three-cluster case (see, Figure A1 in the Appendix), we chose four GMM components for our further analysis. Similarly, Happé et al. (2024) found that four clusters were optimal for their latent space analysis. Our identified clusters represent then the following seasons: cluster *#1* represents DJF (Section 3.2), cluster *#2* represents JJA (Section 3.3), clusters *#3* and *#4* represent transition seasons March–April-May (MAM) or September-October–November (SON) (Section 3.4).

To analyze the atmospheric patterns before and during heatwaves, we constructed composite representations of heatwave samples for each cluster by selecting the closest 100 samples (in latent space) to each GMM cluster centroid, similar to the procedures described by Rouges et al. (2023) and Happé et al. (2024). Since the entire process, i.e., the encoding of heatwave samples to the latent space by the VAE, the dimensionality reduction via t-SNE, and the GMM clustering, contains randomness, a single sample may be sensitive to random fluctuations. The composite maps provide a more stable representation of atmospheric conditions for each cluster by averaging over the closest samples in the latent space. The anomalies of each variable are displayed for the North Atlantic region (as shown in Figure 2) for each of the eleven analyzed days (five days prior to the onset of the heatwave; the onset day of the heatwave, indicated with a broader frame around the graphs; and five days after the onset of the heatwave), and all four identified clusters. The results are shown in Figures 6 to 14.

To assess the physical relevance of our VAE-based clustering, we selected several major European heatwaves that have been extensively studied in the literature as case studies. Russo et al. (2015) identified and ranked the top ten most severe European heatwaves between 1950 and 2014, highlighting events such as the 2003 Western European and 2010 Russian heatwaves as the most severe. Lhotka and Kyselý (2022) expanded on this work by providing an updated catalog of major heatwaves up to 2021. Following these studies, we selected the 2003, 2010, and 2018 heatwaves as case studies to examine whether the VAE latent space captures the known dynamical features of these benchmark events. Due to differences in detection methods, the

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


exact onset dates of heatwaves in our dataset do not always align with those reported in the literature. Therefore, for each target heatwave date defined by Lhotka and Kyselý (2022) ("target date" in Figure 5), we identified the temporally closest matching event in our dataset ("closest date" in Figure 5). We then extracted the ten nearest neighbors in the latent space around each closest date to construct composite anomaly maps. This approach allows us to test whether similar atmospheric patterns are clustered in the same region of the latent space. As shown in Figure 5, the resulting geopotential height at 500 hPa (*z500*) anomaly maps exhibit similar patterns to those reported by Lhotka and Kyselý (2022). The 2010 and 2003 heatwaves showed positive geopotential height at 500 hPa (*z500*) anomalies over Russia and ridges over Central Europe, respectively, while the 2018 heatwave showed positive anomalies over Scandinavia, as can be seen in Figure 5 here and Figure 3 from Lhotka and Kyselý (2022).

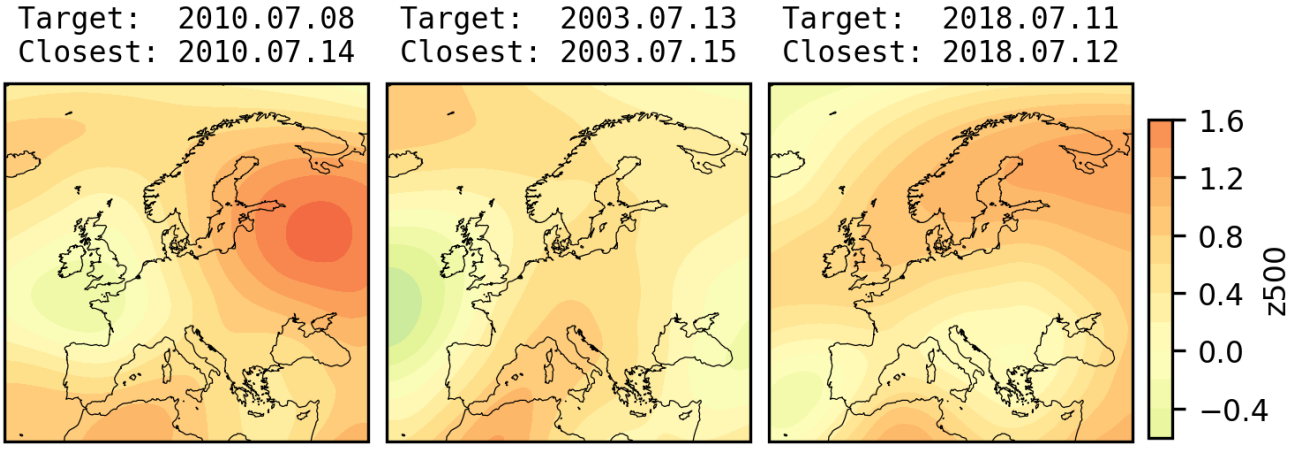


**Figure 5.** Geopotential height at 500 hPa (*z500*) anomaly maps for three historically significant European heatwaves: 2010, 2003, and 2018. The *target dates* correspond to the heatwave onset dates defined by Lhotka and Kyselý (2022), while the *closest dates* refer to the temporally nearest matching events detected in our dataset using the GDBSCAN algorithm. For each event, composite maps are constructed by averaging the ten latent space neighbors closest to the selected onset date. The spatial domain in this figure is cropped to match that used by Lhotka and Kyselý (2022), for consistency, although other figures in this study use a broader domain.

### 3.2 Cluster *#1*

Using the composite maps, we found that four heatwave clusters provide distinct atmospheric patterns for heatwaves in different seasons. In Figures 6 and 7, cluster *#1* shows a similar pattern to the UK High cluster in Happé et al. (2024) and the omega block in Rouges et al. (2023) for the onset day of the heatwaves. While Happé et al. (2024) found this UK high pattern in summer (JJA) months and Rouges et al. (2023) found this pattern in extended summer (May to September) months, we detected that heatwaves, not just in summer, but also in winter months, show this or a similar pattern.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


The positive anomalies over the British Isles and Western Europe, characteristic of this pattern, are persistent $t-5$ days before the heatwave and slowly spread into the Mediterranean. Similarly, the geopotential height at 500 hPa (*z500*) exhibits positive anomaly values centered over Western Europe before the onset of the heatwave, as shown in Figure 14, in line with findings from both Tomczyk et al. (2019) and Leach et al. (2021), who associated winter heatwaves over Europe with positive anomalies of geopotential height. Rouges et al. (2023) classified European heatwaves into five distinct types based on daily mean 2-meter temperature anomalies and geopotential height at 500 hPa anomalies, characterizing the western Europe cluster as a temperature anomaly spread across France and the British Isles, with its circulation pattern similar to omega-type blocking. This pattern is analogous to our identified cluster *#1*. As the extreme event progresses (the days after the onset), these geopotential height at 500 hPa (*z500*) anomalies begin to disperse eastward across the North Atlantic, as shown in cluster *#1* in Figure 14. Tomczyk et al. (2019) also showed this in their study: the anomalies in geopotential height begin to develop over the North Atlantic and move towards Europe, where they strengthen as they move over the continent. This pattern of dispersion is also evident in Figure 10: the positive maximum temperature anomalies associated with winter heatwaves are initially concentrated over Western Europe, gradually weakening and dissipating similarly, following the onset of the heatwave events in cluster *#1*. These patterns are also consistent with findings from previous studies where heatwaves in north-west Europe are typically found to be related to blockings (Carril et al., 2008; Zschenderlein et al., 2020). This blocking is also accompanied by a positive (negative) anomaly in the zonal wind (*u10*) component at the north (south), and a dipole structure with positive (negative) anomalies to the west (east) in the meridional wind (*v10*) (see cluster *#1* in Figures 12 and 13). Together, these features point to a blocking structure with enhanced atmospheric stagnation, consistent with an anti-cyclonic flow (Kautz et al., 2022). This blocking system helps to create and maintain unusually warm weather as it is associated with clear skies, as shown in cluster *#1* in Figures 9 and 11. The lack of cloud cover, indicated by negative anomalies in the total cloud cover (*tcc*) variable, for cluster *#1* over Southern Europe and the Mediterranean in Figure 11, allows more solar radiation to reach the Earth's surface, especially during the build-up of the heatwaves (days $t-5$ to day 0), where these negative anomalies are more pronounced. The positive anomalies in surface net solar radiation (*ssr*), as shown in Figure 9, become more distinct especially after the heatwave event has started, creating a feedback loop that brings further warm and dry air into the region, as shown with negative relative humidity ($r$) anomalies located over Europe in Figure 8 during the onset. During the days after the onset, the negative anomaly for relative humidity tends to move eastward towards the Middle East.

Notably, while previous studies have linked these atmospheric circulation patterns to summer heatwaves, our clustering, which encompasses year-round heatwave data, reveals that similar patterns also manifest during winter events.

### 3.3 Cluster *#2*

Cluster *#2*, which corresponds to mostly summer heatwaves, as shown in Figure 4f, exhibits a similar atmospheric pattern to the southern European cluster (SE) in Rouges et al. (2023). As shown in Figure 6, stream function at 250 hPa (*stream250*) shows a broad negative anomaly over the North Atlantic, expanding from Greenland to Scandinavia, while a larger positive anomaly block is observed from the central North Atlantic to Russia from $t-5$ days before the heatwave onset. Similarly, a continuous ridge pattern in geopotential height at 500 hPa (*z500*) is detectable. It is accompanied by a positive anomaly in

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


zonal wind component, as shown in Figure 12, up to $t+3$ days, and positive anomalies in the meridional wind component, especially from $t-1$ days in the build-up period of heatwaves, as shown in Figure 13. These anomalies point towards a larger heat transfer from lower latitudes to Europe. Positive anomalies in geopotential height at 500 hPa (*z500*) over the Subtropical North Atlantic region start $t-5$ days before the onset, and later sweep into Central Europe during the progression of heatwaves. This is similar to findings from Tomczyk et al. (2019), where they showed that the occurrence of warm spells in winter months is associated with positive anomalies of geopotential height at 500 hPa (*z500*) over Central Europe. After the onset, the positive anomaly ridge in stream function at 250 hPa (*stream250*), stretching from central North Atlantic to Russia, surrounds the North Atlantic trough, and extends to North America through to the Arctic during the progression of heatwaves, as seen in cluster *#2*, from $t=0$ to $t+5$, in Figure 6, while the negative anomalies over North Atlantic dissipates. A persistent negative mean sea level pressure (*msl*) anomaly occurs over the British Isles during the build-up period of these heatwaves, where it later gives place to a positive anomaly over Russia up to $t+3$ days, as shown in Figure 7. As these persistent patterns start to dissipate after the heatwave onset, cloud-free skies and a stagnant atmosphere occur. This is indicated by expanding negative anomalies in total cloud cover (*tcc*) over Europe and Russia, and near-zero wind speed anomalies, as shown in Figures 11 and 12-13, respectively. The positive anomalies in relative humidity (*r*) over the North Atlantic during the build-up period of heatwaves slowly sweep into Western Europe, as the zonal wind component shows positive anomalies during the same time. This brings more humid and warm air to Europe, creating the positive anomalies in relative humidity (*r*) over Western Europe on the onset date, and merging with the anomalies over the Middle East during the heatwave, as shown in Figure 8. Simultaneously, positive temperature anomalies propagate northeast towards central Europe as the heatwave progresses, as shown in Figure 10. On the other hand, a clear negative anomaly pattern over the North Atlantic in 2-meter temperatures (*t2m*) during the build-up and progression of the heatwaves is observed. This negative anomaly pattern over the North Atlantic is consistent with results described by Krüger et al. (2023), Bischof et al. (2023), and Lipfert et al. (2024), where they found that when the sea surface temperatures (SSTs) in the North Atlantic are colder than usual, they show a negative tendency. This leads to increased heatwave intensity and duration across Europe, as well as a shift toward positive anomalies in European summer temperatures. These different atmospheric mechanisms help to explain why the latent space representations of summer heatwave samples (see Figure 4) are clustered distinctly separate from the rest of the dataset. As heatwaves intensify and high temperatures rise more rapidly in Europe and other mid-latitude regions (Perkins-Kirkpatrick and Lewis, 2020; Bischof et al., 2023; Paçal et al., 2023), these events diverge from the historical distribution learned by the VAE. Since the test data is unseen during training, such extreme events are encoded differently in the latent space.

### 3.4 Clusters *#3* and *#4*

Clusters *#3* and *#4* primarily correspond to heatwaves in the transition seasons, occurring mainly during MAM or SON. Even though it is not possible to assign these clusters to certain seasons as we could do with the first two clusters, cluster *#4* shows some similarities in anomaly patterns in certain variables with cluster *#1*, such as (1) 2-meter temperature (*t2m*), where both show a negative–positive–negative anomaly tripole across the North Atlantic–Europe–Russia domain; (2) mean sea level pressure (*msl*), where both show negative anomalies closer to Arctic while having positive anomalies over Europe,

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


(3) stream function at 250 hPa (*stream250*), where both show negative anomalies over Greenland with positive anomalies over Europe. These similarities could explain why the GMMs in Figures A1 and 4f, with three or four components, consistently identify the same seasonal clusters (*#2* and *#3*), while clusters *#1* and *#4* are either merged or separated depending on the number of components. Four components are used in our study to better understand the differences between the characteristics of heatwave clusters.

Although clusters *#3* and *#4* are both associated with transition seasons, they exhibit nearly opposite atmospheric anomaly patterns in several key climate variables. For instance, in Figure 14, cluster *#3* displays a positive geopotential height at 500 hPa (*z500*) anomaly over Greenland and negative anomalies over the North Atlantic extending into Scandinavia, while cluster *#4* shows the opposite, i.e., negative anomalies over Greenland and positive anomalies over the North Atlantic–Scandinavia sector. This creates a contrasting tripole structure during the heatwave build-up phase, which then evolves into opposite-sign anomalies over Scandinavia as the events progress. Likewise, the zonal (*u10*) and meridional (*v10*) wind component anomalies show opposing dipole structures between the two clusters, as shown in Figure 12 and 13. Cluster *#3* shows a positive meridional wind (*v10*) anomaly over the Canaries, $t-5$ days before the onset, and slowly becomes more prominent as the anomaly expands towards Europe during the event's progress. Since this positive anomaly of wind becomes more prominent as the onset date approaches, increasing wind anomalies enhance the advection of warm air from the south, as shown in Figure 10.

On the other hand, the wind components for cluster *#4* exhibit anti-cyclonic flow with positive (negative) anomalies over the western North Atlantic (Canaries) during the build-up shown in Figures 12 and 13, which carries warm air through the North Atlantic, as shown in temperature anomalies in Figure 10. Cluster *#3* is characterized by an elongated trough in stream function at 250 hPa (*stream250*) extending from the east coast of Canada towards Scandinavia, with a ridge over Greenland. Conversely, cluster *#4* exhibits an inverse pattern with a trough over Greenland, and two ridges are located over the central North Atlantic and Scandinavia, as shown in Figure 6, showing similarities with the North Atlantic Low cluster in Happé et al. (2024). This elongated trough in cluster *#3* persists across all analyzed heatwave time steps, while the ridge over the central North Atlantic in cluster *#4* dissolves as the event progresses. Furthermore, while heatwave cluster *#3* is accompanied by positive relative humidity (*r*) anomalies over the Eastern subtropical North Atlantic during the build-up and over Europe following the onset, cluster *#4* exhibits negative relative humidity (*r*) anomalies over Central Europe and Russia (see Figure 8). Recent studies also highlight that the accelerated decline of Arctic sea ice since the 2000s has contributed to a shift toward a dipole-like atmospheric circulation mode, bringing more heatwaves to Europe (Zhang et al., 2020; Lee et al., 2025). These enhance warm air advection from the North Atlantic and favor soil drying over Europe, supporting the contrasting atmospheric variable anomalies observed between clusters *#3* and *#4*.

## 4 Discussion & Conclusions

The study aims to explore the interplay and temporal evolution between different atmospheric variables in their contributions to heatwaves over Western Europe. To do this, we applied a Variational Autoencoder (VAE) and interpreted different data clusters, identified by a Gaussian Mixture Model (GMM), in the latent space. Although our clustering approach is based on Happé et al.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


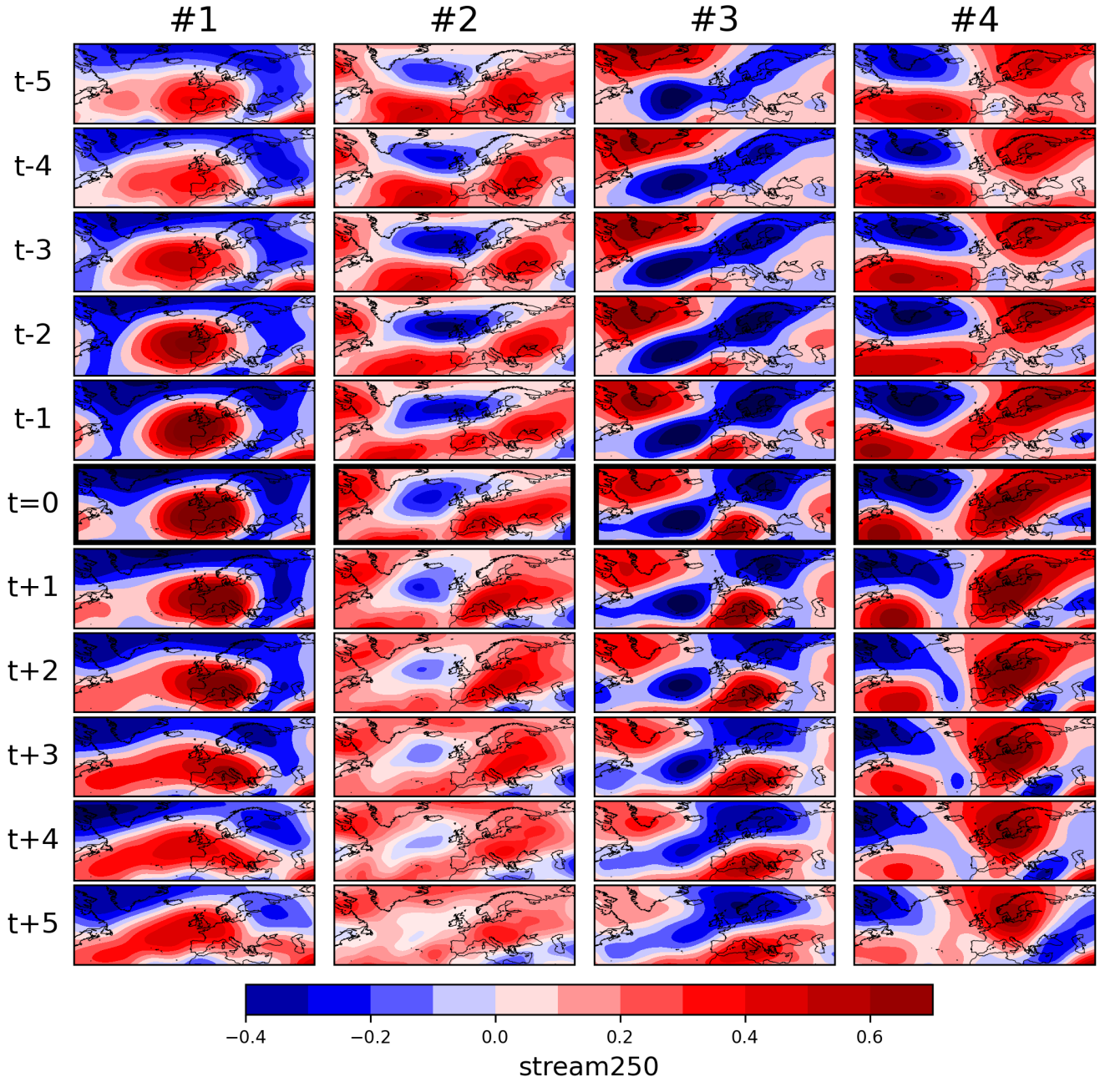


**Figure 6.** Composite map for standardized stream function anomaly at 250 hPa (*stream250*) across $t \pm 5$ days for each GMM cluster identified in Figure 4f. Rows show cluster *#1* to *#4*, and each row corresponds to the cluster average of the n=100 heatwave samples closest to the respective GMM centroids. Time progresses from top to bottom, centered on the heatwave onset date (t = 0) shown in the sixth row, emphasized with a broader frame.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


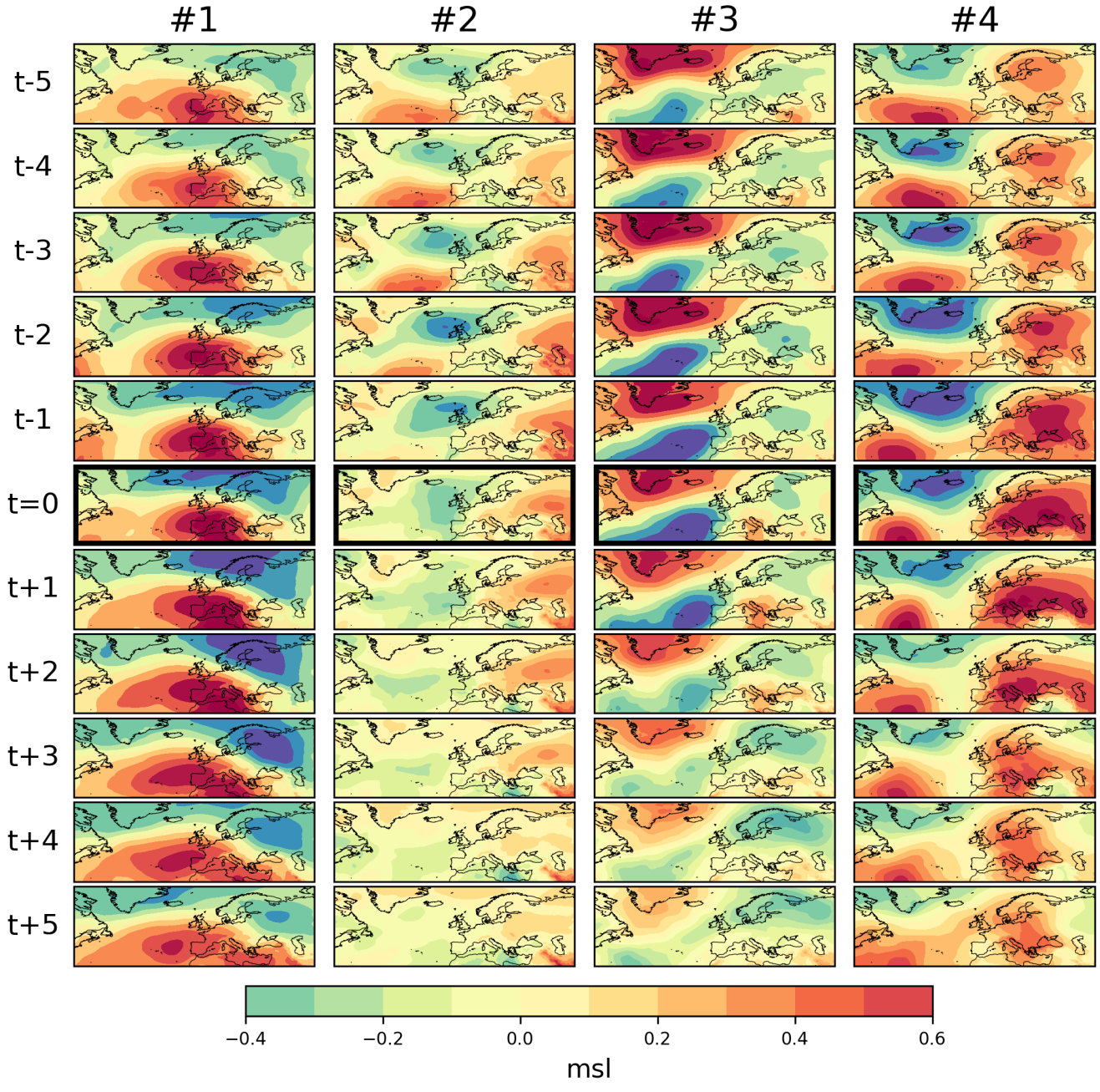


**Figure 7.** Same as Figure 6, but for standardized mean sea level pressure (*msl*) anomaly.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


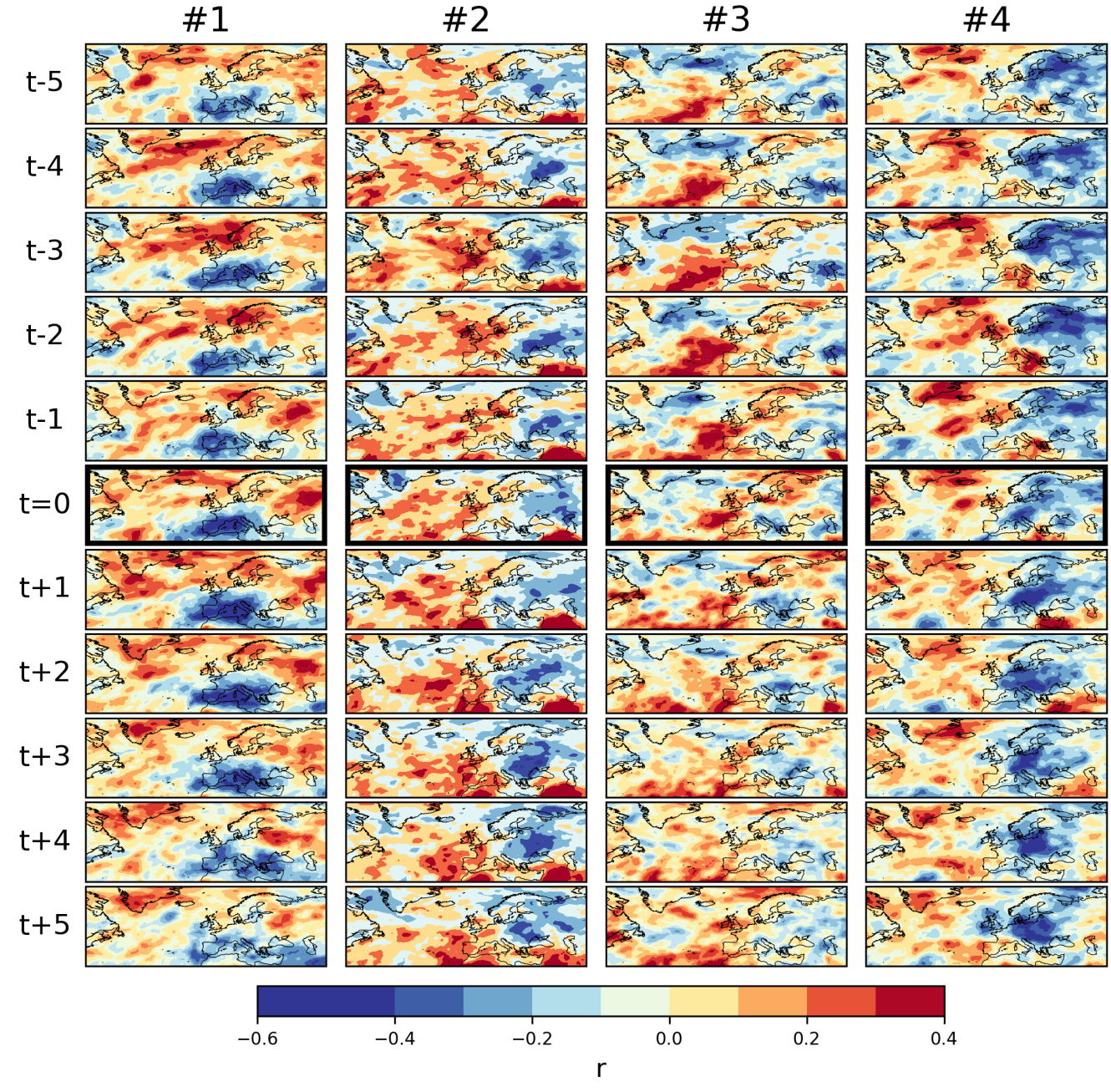


**Figure 8.** Same as Figure 6, but for standardized relative humidity ($r$) anomaly.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


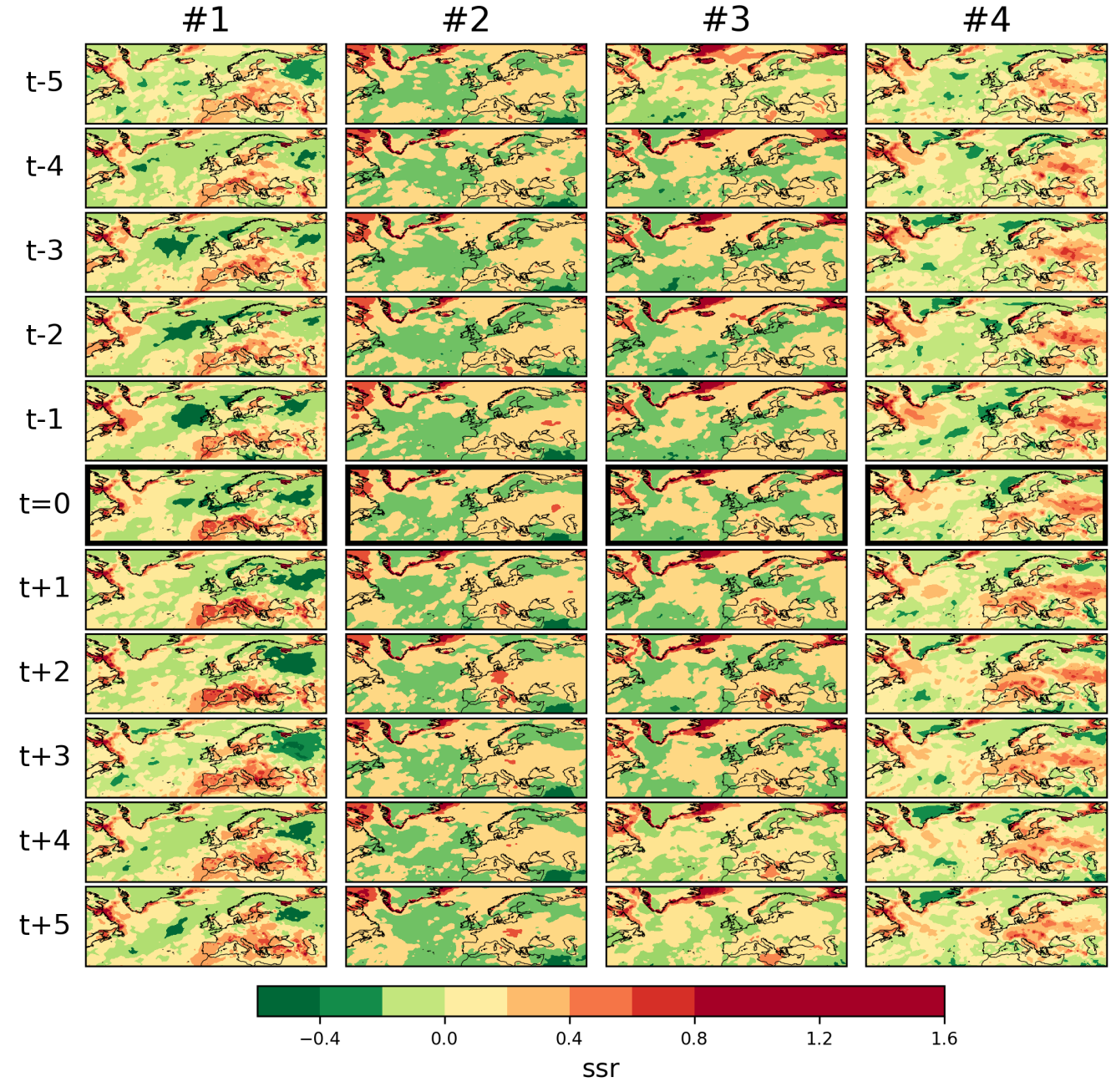


**Figure 9.** Same as Figure 6, but for standardized sum surface net solar radiation (*ssr*) anomaly.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


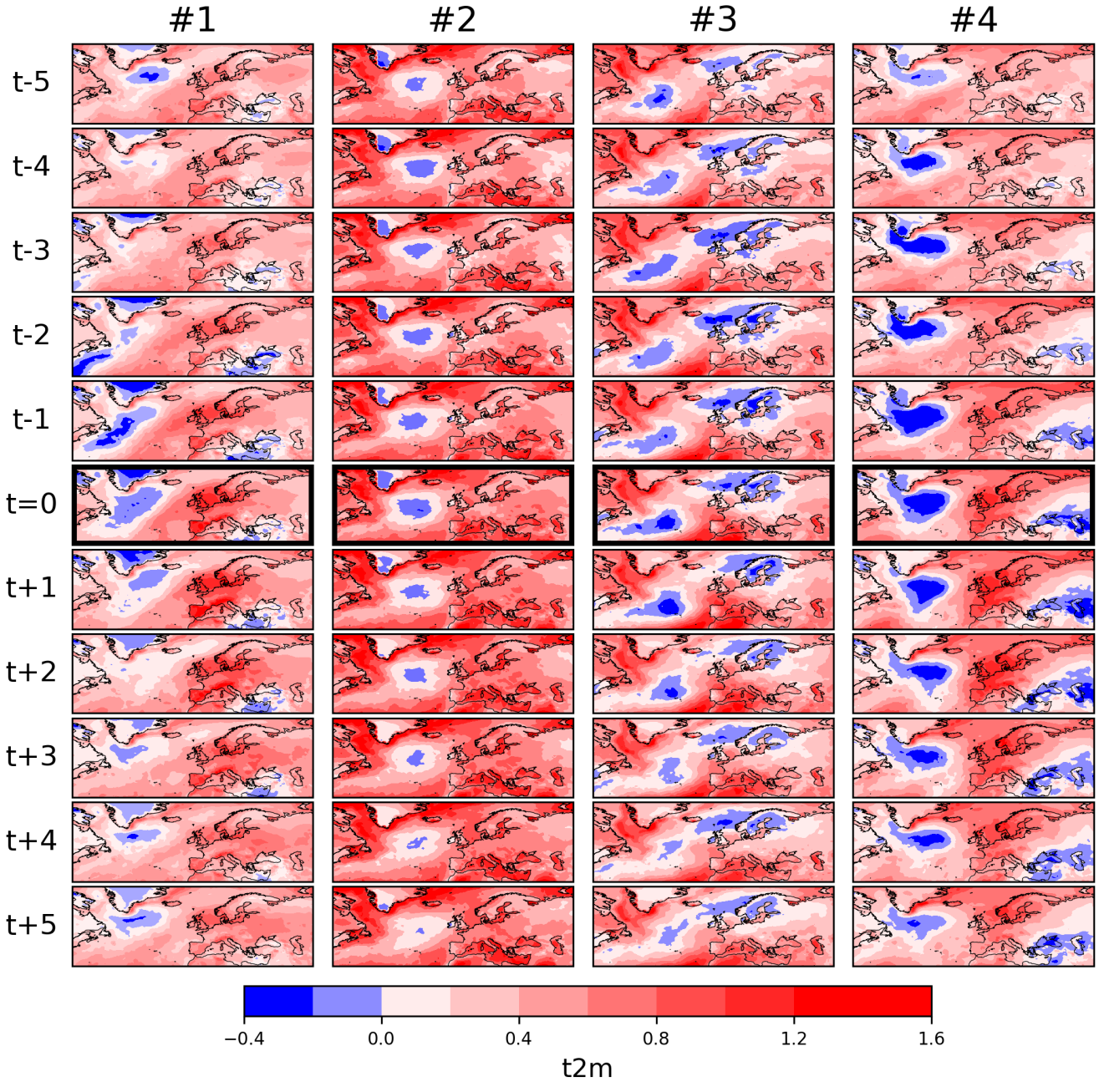


**Figure 10.** Same as Figure 6, but for standardized maximum 2-meter temperature (*t2m*) anomaly.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


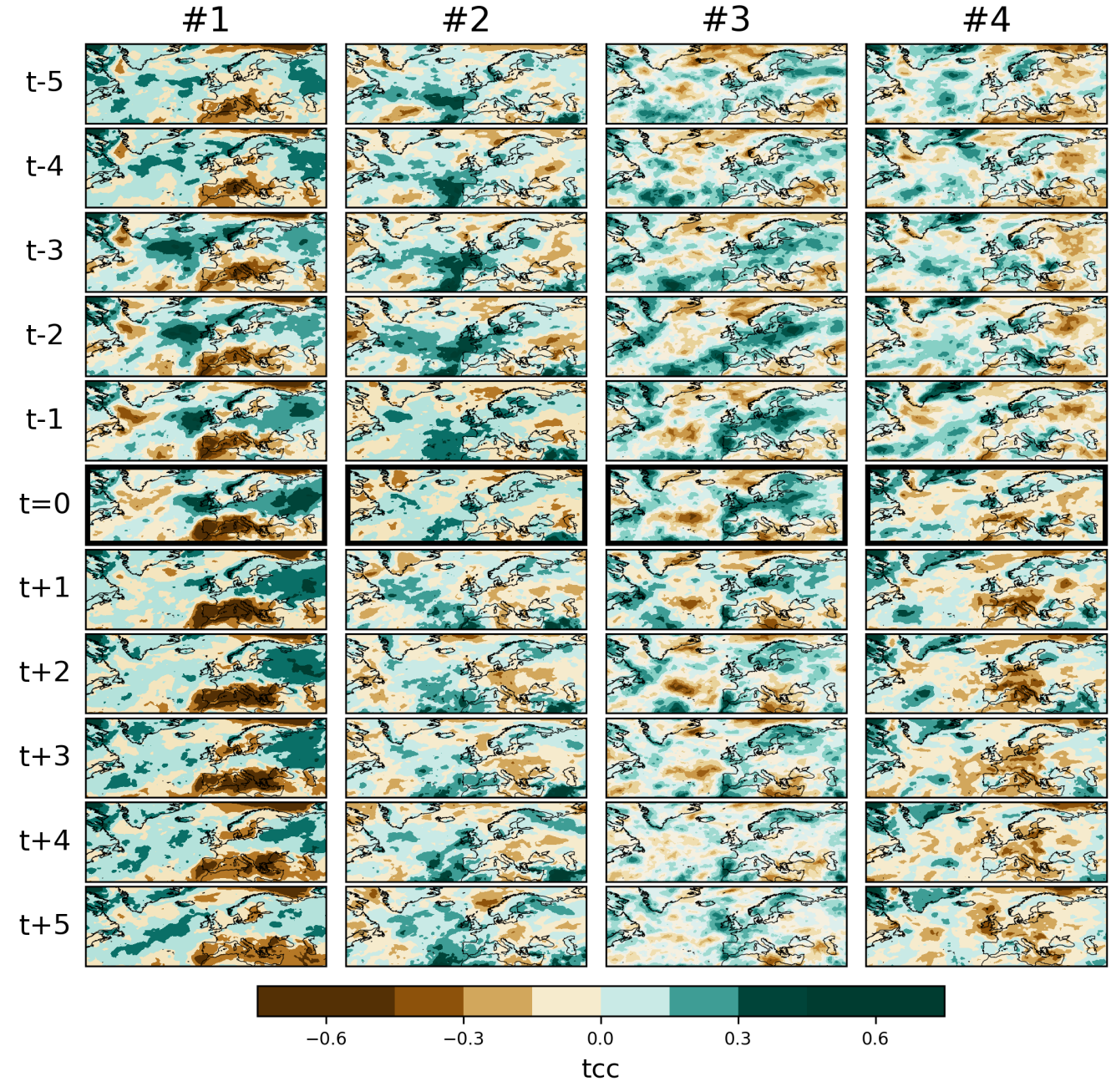


**Figure 11.** Same as Figure 6, but for standardized total cloud cover (*tcc*) anomaly.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


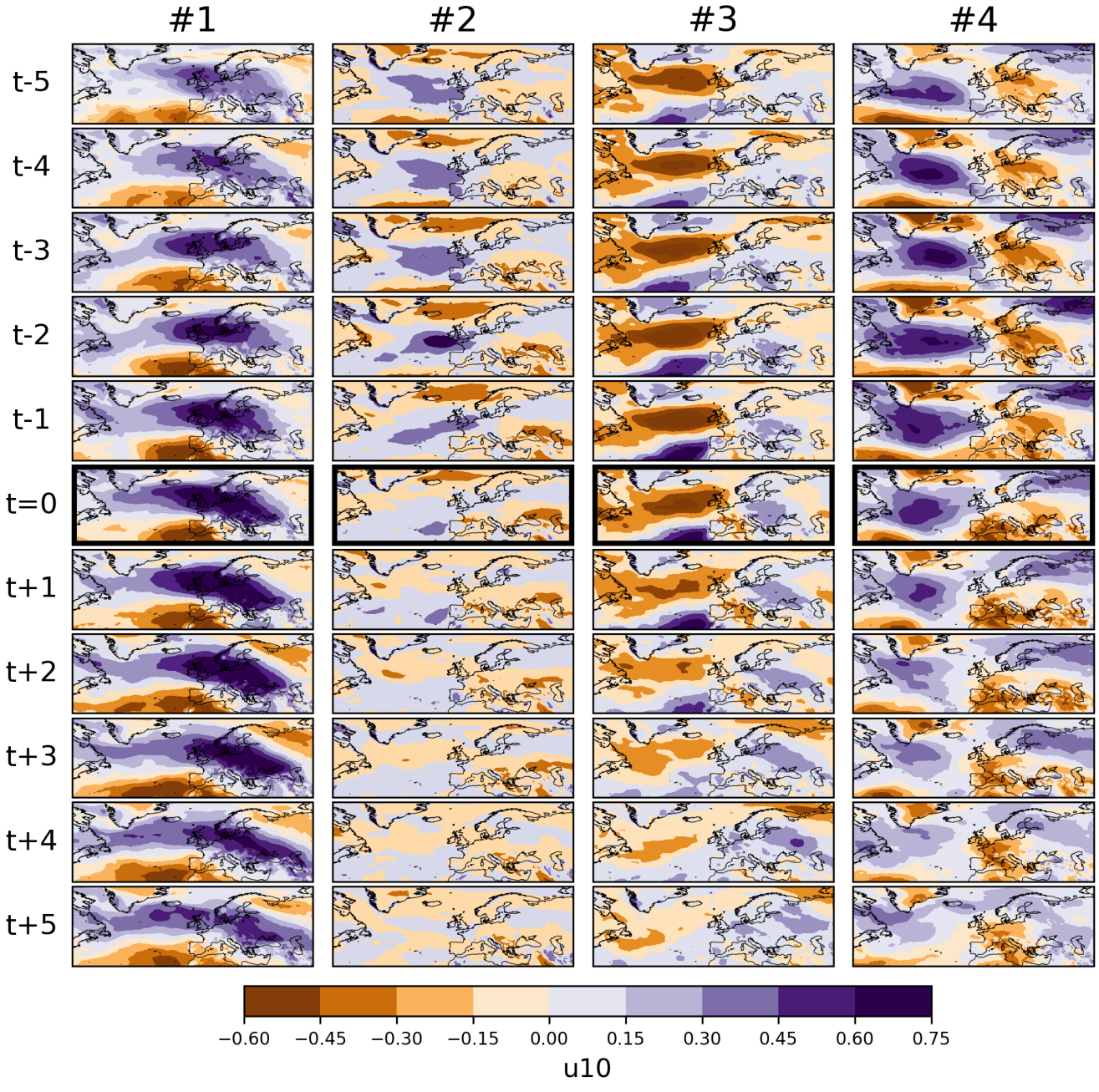


**Figure 12.** Same as Figure 6, but for standardized 10-meter U wind component (*u10*) anomaly.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


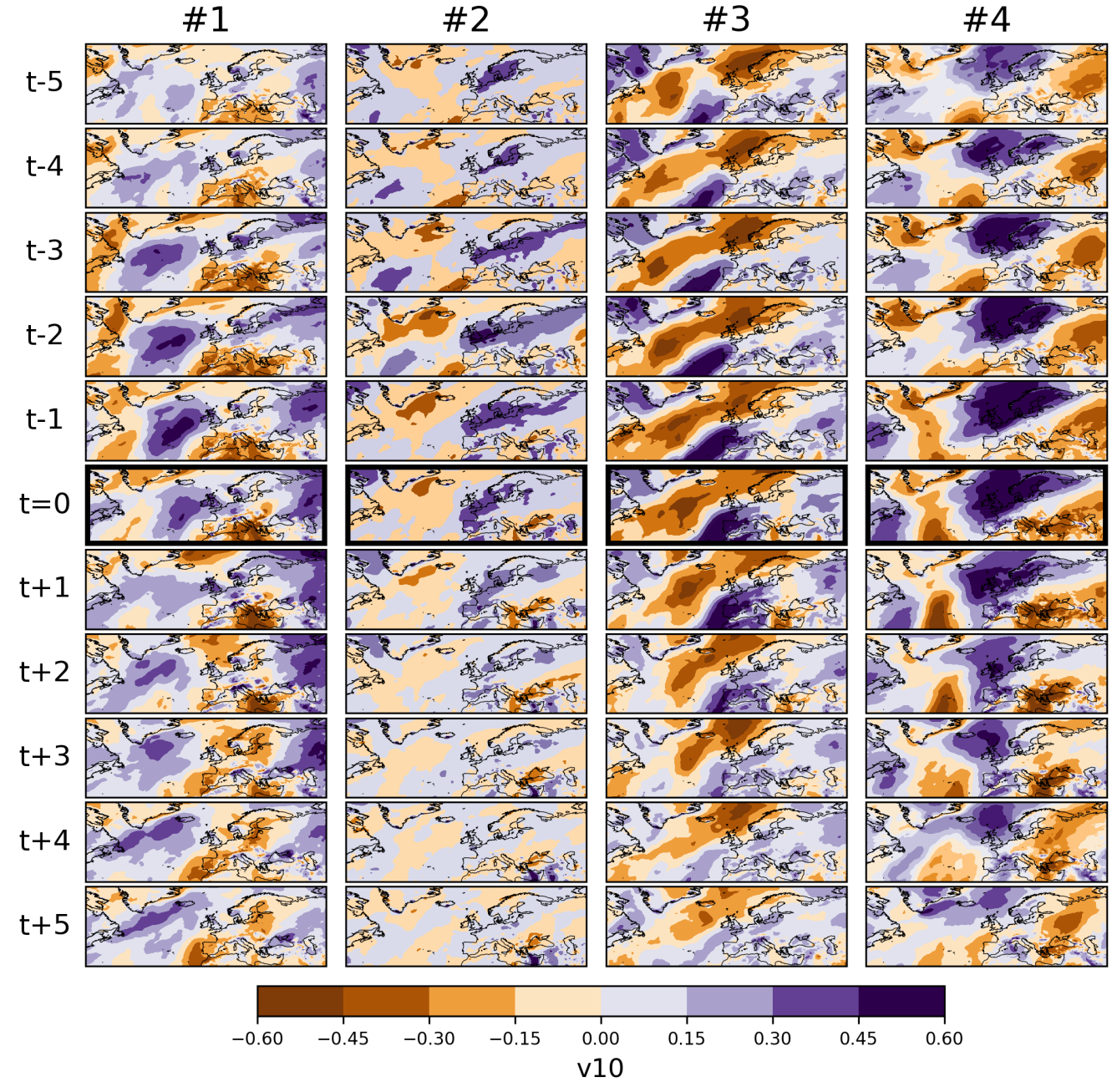


**Figure 13.** Same as Figure 6, but for standardized 10-meter V wind component (*v10*) anomaly.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


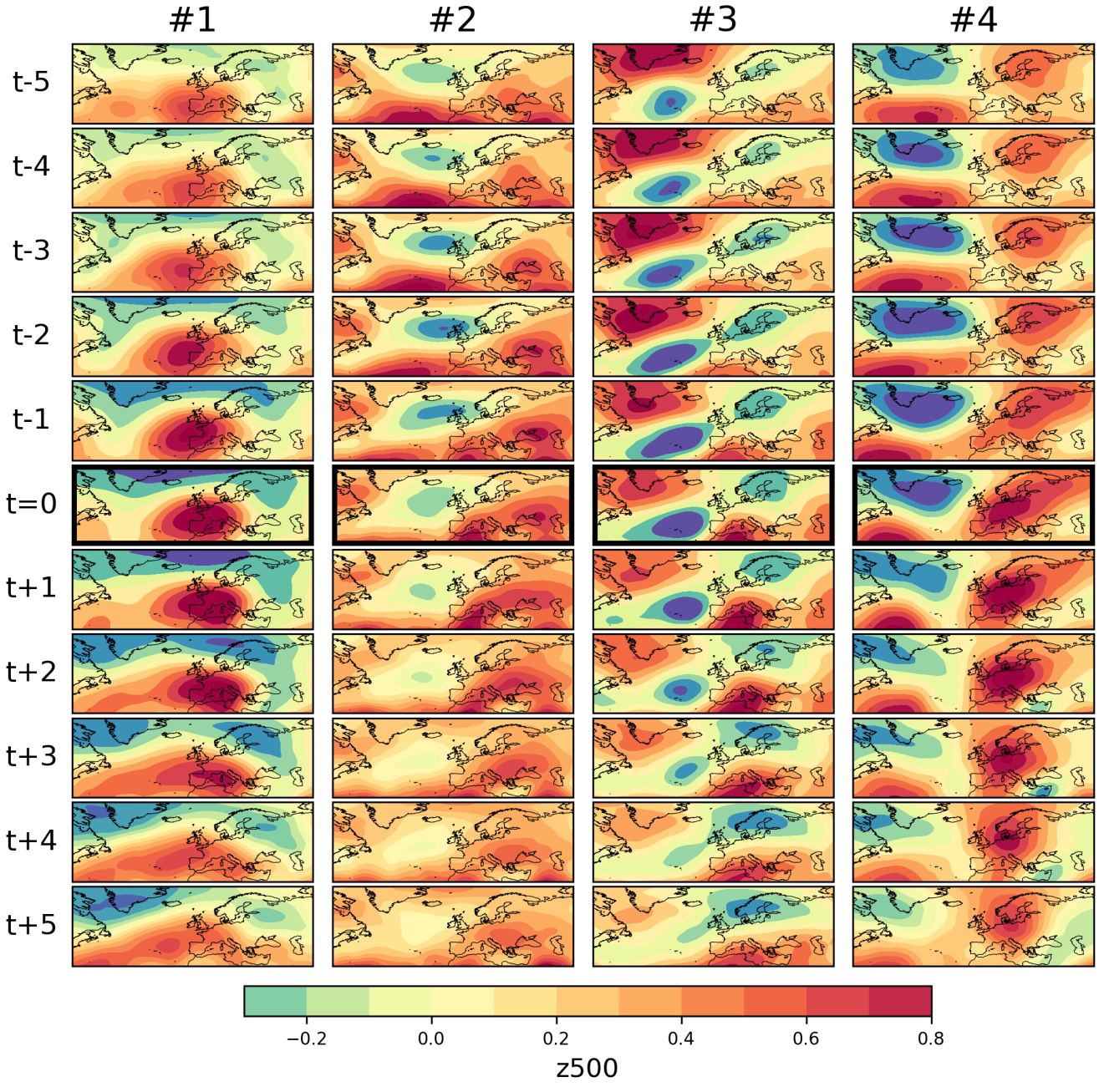


**Figure 14.** Same as Figure 6, but for standardized geopotential height at 500 hPa (*z500*) anomaly.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


(2024), the key difference in their study lies in the use of a VAE trained on 11-day multivariate heat wave samples from ERA5 reanalysis data to reveal interpretable heatwave regimes. These patterns were subsequently analyzed to gain a deeper understanding of the development and progression of heatwaves across all seasons, including summer, winter, and transition periods. By training the VAE on year-round heatwave samples, we were able to capture a comprehensive view of atmospheric dynamics associated with various heatwave events. Our VAE analysis identified several distinct heatwave regimes from the ERA5 dataset, each associated with characteristic atmospheric patterns and seasonal preferences. Understanding the dynamics of heatwaves is important as they are projected to intensify and become more frequent (Perkins-Kirkpatrick and Lewis, 2020; Russo and Domeisen, 2023; Schielicke and Pfahl, 2022; Paçal et al., 2023). This is particularly relevant for Western Europe, where temperature extremes have increased faster than the projections of the climate model (Vautard et al., 2023; Patterson, 2023; McKinnon et al., 2024; Kornhuber et al., 2024). However, these dynamics are not yet fully understood due to the complex interactions between multivariate spatiotemporal variables and their responses to climate change (Van Oldenborgh et al., 2022; Barriopedro et al., 2023). Since finding labeled and high-quality data for anomaly detection is challenging, unsupervised learning methods allow for automatically distinguishing heatwaves (Ruff et al., 2021).

We used GDBSCAN to detect heatwave events and a VAE to analyze and cluster the detected events based on their latent space representations (Sander et al., 1998; Happé et al., 2024). VAEs compress high-dimensional multivariate input data into a low-dimensional latent space while preserving essential features, allowing for efficient reconstruction of the input data (Kingma and Welling, 2019; Lindhe et al., 2021). The reconstruction produced by the VAE can be interpreted as a learned climatological baseline or 'normal state,' against which deviations can be studied. We focused on using latent space to understand the atmospheric patterns underlying different heatwave types. Similar latent space approaches have been successfully applied in climate research (Behrens et al., 2022; Oliveira et al., 2022; Shamekh et al., 2023; Mooers et al., 2023; Camps-Valls et al., 2025).

While we used year-round heatwave samples, the composite map of heatwave clusters exhibits atmospheric patterns similar to previous studies, where only summer months were commonly used (Carril et al., 2008; Horton et al., 2015; Rouges et al., 2023; Krüger et al., 2023; Bischof et al., 2023; Happé et al., 2024). As explained before, our choice of four Gaussian components for our analyses enables a more nuanced analysis of heatwave regimes and enhances the interpretability of the latent space. Each of the four identified groups captures distinct atmospheric conditions in different variables analyzed, with well-defined patterns leading to the onset of the heatwave and their subsequent progression, as illustrated in the composite anomaly maps in Figures 6-14. For example, anomalies in relative humidity ($r$) in Figure 8 highlight contrasting patterns between seasonal heatwaves: Cluster *#1*, associated with winter heatwave samples, exhibits a pronounced negative relative humidity ($r$) anomaly over Western Europe and the Mediterranean, whereas cluster *#2*, dominated by summer heatwaves, shows a positive relative humidity ($r$) anomaly following the onset. This distinction is consistent with previous studies, indicating that humid summer heatwaves are highly dangerous and are shown to be intensifying under climate change (Russo et al., 2017; Dong et al., 2024; Wang et al., 2024). The total cloud cover (*tcc*) shows a distinct negative anomaly starting from $t-5$ days and strengthening during the progression of winter heatwaves over Southern Europe, causing a positive anomaly in surface net solar radiation (*ssr*) a couple of days later, which brings more heat to Western Europe, as seen in cluster *#1* (Figure 9). Meanwhile, positive

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


anomalies in total cloud cover (*tcc*) are observed before the onset of summer heatwaves in cluster *#2*, but they gradually transition into negative anomalies in total cloud cover (*tcc*) over Central Europe as the heatwave evolves (Figure 11). These findings demonstrate the usefulness of VAE latent space analysis. Despite the data being standardized, the VAE organizes heatwave samples into clusters that align with seasonal differences, suggesting seasonal signals embedded in atmospheric patterns of the test period.. Heatwave samples corresponding to summer months are separated in the learned representation, reflecting distinct atmospheric conditions unprecedented in the training period.. This hints that parts of the changes in latent space are caused by a changing seasonal cycle of the atmospheric variables related to heatwaves. This agrees with previous studies, which found that climate change affects the seasonal cycles (Wang et al., 2021; Paçal et al., 2023).

Despite these promising results for identifying heatwaves and their underlying atmospheric conditions, several limitations of the current approach must be acknowledged. The VAE approach, an unsupervised machine learning method, introduces some uncertainty due to the choices in model architecture. Key hyperparameters (such as the dimensionality of the latent space, learning rate, batch size, and the number of layers) can significantly influence training outcomes and the structure of the latent space. In addition, dimensionality reduction using PCA and t-SNE, followed by clustering with GMM, further affects how heatwave events are grouped, as all of these steps involve stochastic elements. To address these, we fixed the random seed to 42 (Adams, 1979) for PCA, t-SNE, and GMM steps to ensure consistency across visualization runs. Finally, by constructing composite maps of atmospheric patterns by averaging the 100 latent samples closest to each GMM cluster centroid, we reduced randomness and provided a more stable and representative depiction of the atmospheric conditions associated with each cluster.

This study highlighted the potential of machine learning methods to improve our understanding of heatwaves. By using a VAE, we uncovered a data-driven classification of heatwave events that not only aligns with known atmospheric circulation regimes from previous studies but also reveals their seasonal and dynamical characteristics from several days before the onset of the heatwave until several days after.

Our results showed that the VAE's latent space effectively captures key atmospheric patterns that have previously been linked to extreme heat events, such as blocking highs, omega blocks, and persistent ridges. Furthermore, we observed that recent summer heatwaves form a distinct cluster separated from the historical distribution of events. This separation is consistent with the increasing intensity and changing nature of heatwaves under anthropogenic climate change (IPCC, 2021). The VAE's ability to detect these shifts in the latent space without supervision suggests that unsupervised learning techniques can be a powerful tool for monitoring climate change (Behrens et al., 2022; Shamekh et al., 2023; Mooers et al., 2023; Camps-Valls et al., 2025).

Future work could further improve the model's ability to analyze different types of extreme events by using additional variables like soil moisture, vegetation indices, or extreme indices. Furthermore, applying the VAE model to CMIP6 climate simulations (Eyring et al., 2016) to evaluate how well climate models capture historical heatwave patterns compared to reanalysis data could provide valuable insights on model biases and help to bridge the gap between observational and model-based understanding of extreme events in a warming climate (Domeisen et al., 2022; Kornhuber et al., 2024; Barriopedro et al., 2023; Brunner and Voigt, 2024). These insights are especially important for climate change adaptation and mitigation efforts (IPCC, 2021), as heatwaves pose serious risks to both environmental systems and human health (Mora et al., 2017b).

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


*Code and data availability.* Python and bash scripts to process and extract heatwave samples from ERA5 reanalysis data, to construct, train, and test the VAE model, and to produce all figures of this manuscript are accessible in the following GitHub repository: https://github.com/EyringMLClimateGroup/pacal25esd_UnderstandingHeatwaves_VAE. ERA5 data are provided by the ECMWF and accessed from DKRZ (Hersbach et al., 2017).

*Author contributions.* AP conceptualized the study with the help of BH, KW, and VE. AP developed the code, conducted the analysis, and generated all figures. MAFT and GCV supported the development of the machine learning approach and provided feedback. AP drafted the manuscript with the help of BH and KW. All authors contributed to the interpretation of results and the writing of the manuscript.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* Funding for this study was provided by the European Research Council (ERC) Synergy Grant "Understanding and Modelling the Earth System with Machine Learning" (USMILE) under the Horizon 2020 research and innovation programme (Grant Agreement No. 855187). M.-A.F-T. and G.C-V also acknowledge funding from the project "eXtreme events: Artificial Intelligence for Detection and Attribution" (XAIDA) under the H2020 programme (Grant Agreement No. 101003469). K.W. acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through the Gottfried Wilhelm Leibniz Prize awarded to Veronika Eyring (Reference number EY 22/2-1). This work used the Deutsches Klimarechenzentrum (DKRZ) resources granted by its Scientific Steering Committee (WLA) under project ID bd1083. The ERA5 reanalysis data was accessed from DKRZ (Hersbach et al., 2017) The results contain modified Copernicus Climate Change Service information for 2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. We would like to thank Dr. Gunnar Behrens for his valuable comments and suggestions to improve the manuscript. ChatGPT was used to improve the clarity and grammar of an earlier version of the manuscript.

## Appendix: Details for the Variational Autoencoder

VAEs, like other deep learning models, require a number of design choices to be made before training. These hyperparameters determine how the model performs, such as how the model adjusts its parameters during each training iteration (learning rate), how many samples it uses in each iteration (batch size), and how complex the internal representations are (hidden layer size and latent space size). Since these settings strongly affect the performance of the model reconstruction, we conducted a grid search over a range of hyperparameters, including learning rates ([0.0001, 0.001, 0.01]), batch sizes ([8, 16, 32, 64]), hidden layer dimensions ([64, 128, 256]), and latent space sizes ([64, 128, 256]). Based on the model's performance on the unseen validation data, we selected a learning rate of 0.001, a batch size of 32, a hidden layer size of 256, and a latent space size of 128. These choices provided a good balance between reconstruction accuracy and computational efficiency. The encoder processes the input with four layers of 3D convolutions with Max Pooling layers across the variable, time, and spatial

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


dimensions (See Table A1). Likewise, Happé et al. (2024) used a similar architecture and latent space size. By using 3D convolutions, the model captures interactions across time and space simultaneously. It gradually reduces the spatiotemporal heatwave data in the encoder. From the encoded representation, the model estimates the parameters of a latent distribution using the reparameterization trick (Kingma and Welling, 2019). Sampling from this distribution yields a 128-dimensional latent vector that serves as the compressed representation of each heatwave event. The decoder, consisting of four layers of 3D transposed convolutions with corresponding unpooling operations to recover the original input dimensions, reconstructs the heatwave sample from its latent space representation. By minimizing the reconstruction loss between the original input and the decoder output, the model learns to optimize its parameters. The encoder-decoder architecture is symmetric and fully convolutional, preserving spatiotemporal locality throughout the encoding and decoding process, as described in Table A1.

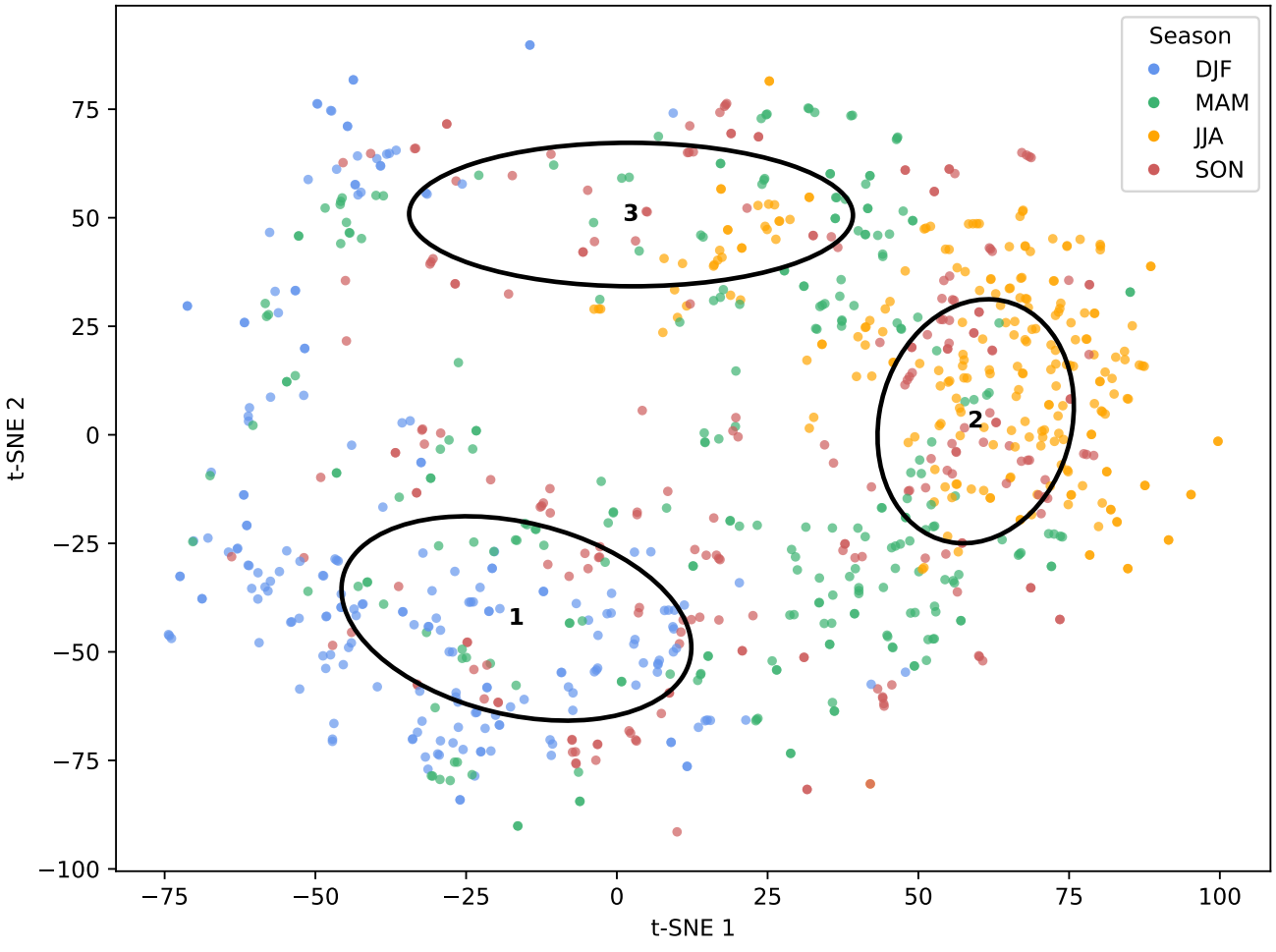


**Figure A1.** t-SNE representation of the latent space with three corresponding GMM components. Each point represents a heatwave sample. Colors represent the seasons in which the sample onset dates are located.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


**Table A1.** Architecture of the VAE model. Each row shows the layer name, the output tensor size after the operation, the number of input and output filters (where applicable), the convolution or transpose convolution kernel size, the applied padding, and the stride. The output has dimensions corresponding to variables (9), days (11), latitude (64), and longitude (192). The encoder compresses the input, and fully connected (FC) layers map the flattened encoder output to the latent space. In VAEs, the latent space is sampled from a Gaussian distribution. Two separate FC layers generate the mean ($FC_\mu$) and standard deviation ($FC_\sigma$) vectors that parameterize the latent Gaussian distribution. After sampling from the latent space, the data is reshaped and passed into the decoder to reconstruct the input.

| **Layer name** | **Output size** | **Filters** | **Kernel size** | **Padding** | **Stride** |
|---|---|---|---|---|---|
| Input | $9 \times 11 \times 64 \times 192$ | - | | | |
| $Conv3D_1$ | $32 \times 11 \times 64 \times 192$ | $9 \rightarrow 32$ | $3 \times 3 \times 3$ | $1 \times 1 \times 1$ | $1 \times 1 \times 1$ |
| $MaxPool3D_1$ | $32 \times 9 \times 32 \times 96$ | | $3 \times 2 \times 2$ | $0 \times 0 \times 0$ | $1 \times 2 \times 2$ |
| $Conv3D_2$ | $64 \times 9 \times 32 \times 96$ | $32 \rightarrow 64$ | $3 \times 3 \times 3$ | $1 \times 1 \times 1$ | $1 \times 1 \times 1$ |
| $MaxPool3D_2$ | $64 \times 7 \times 16 \times 48$ | | $3 \times 2 \times 2$ | $0 \times 0 \times 0$ | $1 \times 2 \times 2$ |
| $Conv3D_3$ | $128 \times 7 \times 16 \times 48$ | $64 \rightarrow 128$ | $3 \times 3 \times 3$ | $1 \times 1 \times 1$ | $1 \times 1 \times 1$ |
| $MaxPool3D_3$ | $128 \times 5 \times 8 \times 24$ | | $3 \times 2 \times 2$ | $0 \times 0 \times 0$ | $1 \times 2 \times 2$ |
| $Conv3D_4$ | $256 \times 5 \times 8 \times 24$ | $128 \rightarrow 256$ | $3 \times 3 \times 3$ | $1 \times 1 \times 1$ | $1 \times 1 \times 1$ |
| $MaxPool3D_4$ | $256 \times 3 \times 4 \times 12$ | | $3 \times 2 \times 2$ | $0 \times 0 \times 0$ | $1 \times 2 \times 2$ |
| Flatten | 36864 | | | | |
| $FC_\mu$ | 128 | $36864 \rightarrow 128$ | | | |
| $FC_\sigma$ | 128 | $36864 \rightarrow 128$ | | | |
| Unflatten | 36864 | $128 \rightarrow 36864$ | | | |
| Reshape | $256 \times 3 \times 4 \times 12$ | | | | |
| $MaxUnpool3D_4$ | $256 \times 5 \times 8 \times 24$ | | $3 \times 2 \times 2$ | $0 \times 0 \times 0$ | $1 \times 2 \times 2$ |
| $ConvTranspose3D_4$ | $128 \times 5 \times 8 \times 24$ | $256 \rightarrow 128$ | $3 \times 3 \times 3$ | $1 \times 1 \times 1$ | $1 \times 1 \times 1$ |
| $MaxUnpool3D_3$ | $128 \times 7 \times 16 \times 48$ | | $3 \times 2 \times 2$ | $0 \times 0 \times 0$ | $1 \times 2 \times 2$ |
| $ConvTranspose3D_3$ | $64 \times 7 \times 16 \times 48$ | $128 \rightarrow 64$ | $3 \times 3 \times 3$ | $1 \times 1 \times 1$ | $1 \times 1 \times 1$ |
| $MaxUnpool3D_2$ | $64 \times 9 \times 32 \times 96$ | | $3 \times 2 \times 2$ | $0 \times 0 \times 0$ | $1 \times 2 \times 2$ |
| $ConvTranspose3D_2$ | $32 \times 9 \times 32 \times 96$ | $64 \rightarrow 32$ | $3 \times 3 \times 3$ | $1 \times 1 \times 1$ | $1 \times 1 \times 1$ |
| $MaxUnpool3D_1$ | $32 \times 11 \times 64 \times 192$ | | $3 \times 2 \times 2$ | $0 \times 0 \times 0$ | $1 \times 2 \times 2$ |
| $ConvTranspose3D_1$ | $9 \times 11 \times 64 \times 192$ | $32 \rightarrow 9$ | $3 \times 3 \times 3$ | $1 \times 1 \times 1$ | $1 \times 1 \times 1$ |

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


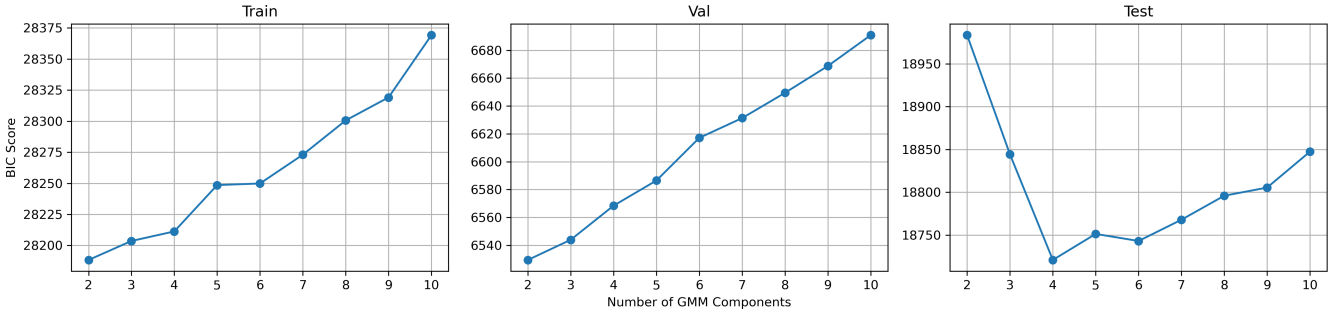


**Figure A2.** Bayesian Information Criterion (BIC) scores for GMM clustering across 2 to 10 components. Each panel shows the BIC scores for a different dataset split: (left) training period (1941–1990), (middle) validation period (1991–2000), and (right) testing period (2001-2022). The test set shows a clear minimum BIC score at four components, suggesting this is the optimal number of clusters under the BIC criterion. In contrast, the training and validation sets show steadily increasing BIC scores with additional components, favoring simpler models. Despite these differences, the four-component model provides a balance between underfitting and overfitting across all periods.

## References

Adams, D.: The Hitch Hiker's guide to the Galaxy, Pan original, Pan Books, London, ISBN 978-0-330-25864-7, 1979.

Alaya, M. A. B., Zwiers, F., and Zhang, X.: An Evaluation of Block-Maximum-Based Estimation of Very Long Return Period Precipitation Extremes with a Large Ensemble Climate Simulation, Journal of Climate, 33, 6957–6970, https://doi.org/10.1175/JCLI-D-19-0011.1, 2020.

Albuquerque Filho, J. E. D., Brandão, L. C. P., Fernandes, B. J. T., and Maciel, A. M. A.: A Review of Neural Networks for Anomaly Detection, IEEE Access, 10, 112 342–112 367, https://doi.org/10.1109/ACCESS.2022.3216007, 2022.

Barriopedro, D., García-Herrera, R., Ordóñez, C., Miralles, D. G., and Salcedo-Sanz, S.: Heat Waves: Physical Understanding and Scientific Challenges, Reviews of Geophysics, 61, e2022RG000 780, https://doi.org/https://doi.org/10.1029/2022RG000780, 2023.

Behrens, G., Beucler, T., Gentine, P., Iglesias-Suarez, F., Pritchard, M., and Eyring, V.: Non-Linear Dimensionality Reduction With a Variational Encoder Decoder to Understand Convective Processes in Climate Models, Journal of Advances in Modeling Earth Systems, 14, e2022MS003 130, https://doi.org/https://doi.org/10.1029/2022MS003130, 2022.

Bischof, S., Pilch Kedzierski, R., Hänsch, M., Wahl, S., and Matthes, K.: The Role of the North Atlantic for Heat Wave Characteristics in Europe, an ECHAM6 Study, Geophysical Research Letters, 50, e2023GL105 280, https://doi.org/https://doi.org/10.1029/2023GL105280, 2023.

Boni, Z., Bieńkowska, Z., Chwałczyk, F., Jancewicz, B., Marginean, I., and Serrano, P. Y.: What is a heat(wave)? An interdisciplinary perspective, Climatic Change, 176, 129, https://doi.org/10.1007/s10584-023-03592-3, 2023.

Brunner, L. and Voigt, A.: Pitfalls in diagnosing temperature extremes, Nature Communications, 15, 2087, https://doi.org/10.1038/s41467-024-46349-x, 2024.

Burke, M., Hsiang, S. M., and Miguel, E.: Global non-linear effect of temperature on economic production, Nature, 527, 235–239, https://doi.org/10.1038/nature15725, 2015.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


Camps-Valls, G., Fernandez-Torres, M. Á., Cohrs, K.-H., Höhl, A., Castelletti, A., Pacal, A., Robin, C., Martinuzzi, F., Papoutsis, I., Prapas, I., Pérez-Aracil, J., Weigel, K., Gonzalez-Calabuig, M., Reichstein, M., Rabel, M., Giuliani, M., Mahecha, M. D., Popescu, O.-I., Pellicer-Valero, O. J., Ouala, S., Salcedo-Sanz, S., Sippel, S., Kondylatos, S., Happé, T., and Williams, T.: Artificial intelligence for modeling and understanding extreme weather and climate events, Nature Communications, 16, 1919, https://doi.org/10.1038/s41467-025-56573-8, 2025.

Carril, A. F., Gualdi, S., Cherchi, A., and Navarra, A.: Heatwaves in Europe: areas of homogeneous variability and links with the regional to large-scale atmospheric and SSTs anomalies, Climate Dynamics, 30, 77–98, https://doi.org/10.1007/s00382-007-0274-5, 2008.

Coumou, D. and Rahmstorf, S.: A decade of weather extremes, Nature Climate Change, 2, 491–496, https://doi.org/10.1038/nclimate1452, 2012.

Domeisen, D. I. V., Eltahir, E. A. B., Fischer, E. M., Knutti, R., Perkins-Kirkpatrick, S. E., Schär, C., Seneviratne, S. I., Weisheimer, A., and Wernli, H.: Prediction and projection of heatwaves, Nature Reviews Earth & Environment, https://doi.org/10.1038/s43017-022-00371-z, 2022.

Dong, J., Brönnimann, S., Hu, T., Cheng, X., Liu, Y., and Peng, J.: Trends of the Intra-Annual Onset and End of Humid Heatwaves in the Northern Hemisphere, Earth's Future, 12, e2024EF005 163, https://doi.org/https://doi.org/10.1029/2024EF005163, 2024.

Dosio, A., Mentaschi, L., Fischer, E. M., and Wyser, K.: Extreme heat waves under 1.5 °C and 2 °C global warming, Environmental Research Letters, 13, 054 006, https://doi.org/10.1088/1748-9326/aab827, 2018.

Elguindi, N., Rauscher, S. A., and Giorgi, F.: Historical and future changes in maximum and minimum temperature records over Europe, Climatic Change, 117, 415–431, https://doi.org/10.1007/s10584-012-0528-z, 2013.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fan, Y., Wen, G., Li, D., Qiu, S., Levine, M. D., and Xiao, F.: Video anomaly detection and localization via Gaussian Mixture Fully Convolutional Variational Autoencoder, Computer Vision and Image Understanding, 195, 102 920, https://doi.org/10.1016/j.cviu.2020.102920, 2020.

Fischer, E. M., Sippel, S., and Knutti, R.: Increasing probability of record-shattering climate extremes, Nature Climate Change, 11, 689–695, https://doi.org/10.1038/s41558-021-01092-9, 2021.

García-León, D., Casanueva, A., Standardi, G., Burgstall, A., Flouris, A. D., and Nybo, L.: Current and projected regional economic impacts of heatwaves in Europe, Nature Communications, 12, 5807, https://doi.org/10.1038/s41467-021-26050-z, 2021.

Happé, T., Wijnands, J. S., Fernández-Torres, M. Á., Scussolini, P., Muntjewerf, L., and Coumou, D.: Detecting Spatiotemporal Dynamics of Western European Heatwaves Using Deep Learning, Artificial Intelligence for the Earth Systems, 3, e230 107, https://doi.org/10.1175/AIES-D-23-0107.1, 2024.

Hasan, M., Choi, J., Neumann, J., Roy-Chowdhury, A. K., and Davis, L. S.: Learning Temporal Regularity in Video Sequences, 2016.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., de Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.-N.: Complete ERA5 from 1940: Fifth generation of ECMWF atmospheric reanalyses of the global climate, Copernicus Climate Change

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


Service (C3S) Climate Data Store (CDS), https://doi.org/10.24381/cds.143582cf, accessed on 15-JUN-2023. Data distribution by the German Climate Computing Center (DKRZ), 2017.

Horton, D. E., Johnson, N. C., Singh, D., Swain, D. L., Rajaratnam, B., and Diffenbaugh, N. S.: Contribution of changes in atmospheric circulation patterns to extreme temperature trends, Nature, 522, 465–469, https://doi.org/10.1038/nature14550, 2015.

Huang, W. K., Stein, M. L., McInerney, D. J., Sun, S., and Moyer, E. J.: Estimating changes in temperature extremes from millennial-scale climate simulations using generalized extreme value (GEV) distributions, Advances in Statistical Climatology, Meteorology and Oceanography, 2, 79–103, https://doi.org/10.5194/ascmo-2-79-2016, 2016.

Huntingford, C., Cox, P. M., Ritchie, P. D. L., Clarke, J. J., Parry, I. M., and Williamson, M. S.: Acceleration of daily land temperature extremes and correlations with surface energy fluxes, npj Climate and Atmospheric Science, 7, 1–10, https://doi.org/10.1038/s41612-024-00626-0, 2024.

IPCC: Climate Change 2021: The Physical Science Basis, Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press (In Press), edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou, 2021.

Jézéquel, A., Yiou, P., and Radanovics, S.: Role of circulation in European heatwaves using flow analogues, Climate Dynamics, 50, 1145–1159, https://doi.org/10.1007/s00382-017-3667-0, 2018.

Kautz, L.-A., Martius, O., Pfahl, S., Pinto, J. G., Ramos, A. M., Sousa, P. M., and Woollings, T.: Atmospheric blocking and weather extremes over the Euro-Atlantic sector – a review, Weather and Climate Dynamics, 3, 305–336, https://doi.org/10.5194/wcd-3-305-2022, 2022.

Kim, J.-Y. and Seo, K.-H.: Physical mechanisms for the dominant summertime high-latitude atmospheric teleconnection pattern and the related Northern Eurasian climates, Environmental Research Letters, 18, 104 022, https://doi.org/10.1088/1748-9326/acfa13, 2023.

King, A. D. and Harrington, L. J.: The Inequality of Climate Change From 1.5 to 2°C of Global Warming, Geophysical Research Letters, 45, 5030–5033, https://doi.org/https://doi.org/10.1029/2018GL078430, 2018.

Kingma, D. P. and Welling, M.: An Introduction to Variational Autoencoders, Foundations and Trends® in Machine Learning, 12, 307–392, https://doi.org/10.1561/2200000056, 2019.

Kornhuber, K., Bartusek, S., Seager, R., Schellnhuber, H. J., and Ting, M.: Global emergence of regional heatwave hotspots outpaces climate model simulations, Proceedings of the National Academy of Sciences, 121, e2411258 121, https://doi.org/10.1073/pnas.2411258121, 2024.

Kotz, M., Levermann, A., and Wenz, L.: The economic commitment of climate change, Nature, 628, 551–557, https://doi.org/10.1038/s41586-024-07219-0, 2024.

Krüger, J., Kjellsson, J., Kedzierski, R. P., and Claus, M.: Connecting North Atlantic SST Variability to European Heat Events over the Past Decades, Tellus A: Dynamic Meteorology and Oceanography, 75, https://doi.org/10.16993/tellusa.3235, 2023.

Lacombe, R., Grossman, H., Hendren, L., and Lüdeke, D.: Improving extreme weather events detection with light-weight neural networks, https://doi.org/10.48550/arXiv.2304.00176, 2023.

Leach, N. J., Weisheimer, A., Allen, M. R., and Palmer, T.: Forecast-based attribution of a winter heatwave within the limit of predictability, Proceedings of the National Academy of Sciences, 118, e2112087 118, https://doi.org/10.1073/pnas.2112087118, 2021.

Lee, D. Y., Yeh, S.-W., Lee, Y.-H., Cai, W., Wang, G., and Yang, Y.-M.: The emergence of a dipole-like mode in Arctic atmospheric circulation conducive to European heat waves, Communications Earth & Environment, 6, 1–9, https://doi.org/10.1038/s43247-025-02020-x, 2025.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


Lhotka, O. and Kyselý, J.: The 2021 European Heat Wave in the Context of Past Major Heat Waves, Earth and Space Science, 9, e2022EA002 567, https://doi.org/https://doi.org/10.1029/2022EA002567, 2022.

Li, C., Zwiers, F., Zhang, X., Li, G., Sun, Y., and Wehner, M.: Changes in Annual Extremes of Daily Temperature and Precipitation in CMIP6 Models, Journal of Climate, 34, 3441–3460, https://doi.org/10.1175/JCLI-D-19-1013.1, 2021.

Lindhe, A., Ringqvist, C., and Hult, H.: Variational Auto Encoder Gradient Clustering, https://doi.org/10.48550/arXiv.2105.06246, 2021.

Lipfert, L., Hand, R., and Brönnimann, S.: A Global Assessment of Heatwaves Since 1850 in Different Observational and Model Data Sets, Geophysical Research Letters, 51, e2023GL106 212, https://doi.org/https://doi.org/10.1029/2023GL106212, 2024.

Lynas, M., Houlton, B. Z., and Perry, S.: Greater than 99% consensus on human caused climate change in the peer-reviewed scientific literature, Environmental Research Letters, 16, 114 005, https://doi.org/10.1088/1748-9326/ac2966, 2021.

López-Bueno, J. A., Navas-Martín, M. A., Linares, C., Mirón, I. J., Luna, M. Y., Sánchez-Martínez, G., Culqui, D., and Díaz, J.: Analysis of the impact of heat waves on daily mortality in urban and rural areas in Madrid, Environmental Research, 195, 110 892, https://doi.org/10.1016/j.envres.2021.110892, 2021.

Materia, S., García, L. P., van Straaten, C., O, S., Mamalakis, A., Cavicchia, L., Coumou, D., de Luca, P., Kretschmer, M., and Donat, M.: Artificial intelligence for climate prediction of extremes: State of the art, challenges, and future perspectives, WIREs Climate Change, 15, e914, https://doi.org/10.1002/wcc.914, 2024.

McKinnon, K. A., Simpson, I. R., and Williams, A. P.: The pace of change of summertime temperature extremes, Proceedings of the National Academy of Sciences, 121, e2406143 121, https://doi.org/10.1073/pnas.2406143121, 2024.

McPhillips, L. E., Chang, H., Chester, M. V., Depietri, Y., Friedman, E., Grimm, N. B., Kominoski, J. S., McPhearson, T., Méndez-Lázaro, P., Rosi, E. J., and Shafiei Shiva, J.: Defining Extreme Events: A Cross-Disciplinary Review, Earth's Future, 6, 441–455, https://doi.org/https://doi.org/10.1002/2017EF000686, 2018.

Mooers, G., Pritchard, M., Beucler, T., Srivastava, P., Mangipudi, H., Peng, L., Gentine, P., and Mandt, S.: Comparing storm resolving models and climates via unsupervised machine learning, Scientific Reports, 13, 22 365, https://doi.org/10.1038/s41598-023-49455-w, 2023.

Mora, C., Counsell, C. W., Bielecki, C. R., and Louis, L. V.: Twenty-Seven Ways a Heat Wave Can Kill You, Circulation: Cardiovascular Quality and Outcomes, 10, e004 233, https://doi.org/10.1161/CIRCOUTCOMES.117.004233, 2017a.

Mora, C., Dousset, B., Caldwell, I. R., Powell, F. E., Geronimo, R. C., Bielecki, C. R., Counsell, C. W. W., Dietrich, B. S., Johnston, E. T., Louis, L. V., Lucas, M. P., McKenzie, M. M., Shea, A. G., Tseng, H., Giambelluca, T. W., Leon, L. R., Hawkins, E., and Trauernicht, C.: Global risk of deadly heat, Nature Climate Change, 7, 501–506, https://doi.org/10.1038/nclimate3322, 2017b.

Muntjewerf, L., Bintanja, R., Reerink, T., and van der Wiel, K.: The KNMI Large Ensemble Time Slice (KNMI–LENTIS), Geoscientific Model Development, 16, 4581–4597, https://doi.org/10.5194/gmd-16-4581-2023, 2023.

Nassif, A. B., Talib, M. A., Nasir, Q., and Dakalbab, F. M.: Machine Learning for Anomaly Detection: A Systematic Review, IEEE Access, 9, 78 658–78 700, https://doi.org/10.1109/ACCESS.2021.3083060, 2021.

Newell, R. G., Prest, B. C., and Sexton, S. E.: The GDP-Temperature relationship: Implications for climate change damages, Journal of Environmental Economics and Management, 108, 102 445, https://doi.org/10.1016/j.jeem.2021.102445, 2021.

Newman, R. and Noy, I.: The global costs of extreme weather that are attributable to climate change, Nature Communications, 14, 6103, https://doi.org/10.1038/s41467-023-41888-1, 2023.

Oliveira, D. A. B., Diaz, J. G., Zadrozny, B., Watson, C. D., and Zhu, X. X.: Controlling Weather Field Synthesis Using Variational Autoencoders, in: IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, pp. 5027–5030, https://doi.org/10.1109/IGARSS46834.2022.9884668, 2022.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


Pang, G., Shen, C., Cao, L., and Hengel, A. V. D.: Deep Learning for Anomaly Detection: A Review, ACM Computing Surveys, 54, 1–38, https://doi.org/10.1145/3439950, 2021.

Patterson, M.: North-West Europe Hottest Days Are Warming Twice as Fast as Mean Summer Days, Geophysical Research Letters, 50, e2023GL102 757, https://doi.org/https://doi.org/10.1029/2023GL102757, 2023.

Paçal, A., Hassler, B., Weigel, K., Kurnaz, M. L., Wehner, M. F., and Eyring, V.: Detecting Extreme Temperature Events Using Gaussian Mixture Models, Journal of Geophysical Research: Atmospheres, 128, e2023JD038 906, https://doi.org/https://doi.org/10.1029/2023JD038906, 2023.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Perkins, S. E.: A review on the scientific understanding of heatwaves—Their measurement, driving mechanisms, and changes at the global scale, Atmospheric Research, 164-165, 242–267, https://doi.org/10.1016/j.atmosres.2015.05.014, 2015.

Perkins-Kirkpatrick, S. E. and Lewis, S. C.: Increasing trends in regional heatwaves, Nature Communications, 11, 3357, https://doi.org/10.1038/s41467-020-16970-7, 2020.

Perkins-Kirkpatrick, S. E., Fischer, E. M., Angélil, O., and Gibson, P. B.: The influence of internal climate variability on heatwave frequency trends, Environmental Research Letters, 12, 044 005, https://doi.org/10.1088/1748-9326/aa63fe, 2017.

Prabhat, Kashinath, K., Mudigonda, M., Kim, S., Kapp-Schwoerer, L., Graubner, A., Karaismailoglu, E., von Kleist, L., Kurth, T., Greiner, A., Mahesh, A., Yang, K., Lewis, C., Chen, J., Lou, A., Chandran, S., Toms, B., Chapman, W., Dagon, K., Shields, C. A., O'Brien, T., Wehner, M., and Collins, W.: ClimateNet: an expert-labeled open dataset and deep learning architecture for enabling high-precision analyses of extreme weather, Geoscientific Model Development, 14, 107–124, https://doi.org/10.5194/gmd-14-107-2021, 2021.

Rahmstorf, S. and Coumou, D.: Increase of extreme events in a warming world, Proceedings of the National Academy of Sciences, 108, 17 905–17 909, https://doi.org/10.1073/pnas.1101766108, 2011.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195–204, https://doi.org/10.1038/s41586-019-0912-1, 2019.

Reid, P. C., Hari, R. E., Beaugrand, G., Livingstone, D. M., Marty, C., Straile, D., Barichivich, J., Goberville, E., Adrian, R., Aono, Y., Brown, R., Foster, J., Groisman, P., Hélaouët, P., Hsu, H.-H., Kirby, R., Knight, J., Kraberg, A., Li, J., Lo, T.-T., Myneni, R. B., North, R. P., Pounds, J. A., Sparks, T., Stübi, R., Tian, Y., Wiltshire, K. H., Xiao, D., and Zhu, Z.: Global impacts of the 1980s regime shift, Global Change Biology, 22, 682–703, https://doi.org/https://doi.org/10.1111/gcb.13106, 2016.

Ronco, M., Tárraga, J. M., Muñoz, J., Piles, M., Marco, E. S., Wang, Q., Espinosa, M. T. M., Ponserre, S., and Camps-Valls, G.: Exploring interactions between socioeconomic context and natural hazards on human population displacement, Nature Communications, 14, 8004, https://doi.org/10.1038/s41467-023-43809-8, 2023.

Rouges, E., Ferranti, L., Kantz, H., and Pappenberger, F.: European heatwaves: Link to large-scale circulation patterns and intraseasonal drivers, International Journal of Climatology, 43, 3189–3209, https://doi.org/https://doi.org/10.1002/joc.8024, 2023.

Rousi, E., Kornhuber, K., Beobide-Arsuaga, G., Luo, F., and Coumou, D.: Accelerated western European heatwave trends linked to more-persistent double jets over Eurasia, Nature Communications, 13, 3851, https://doi.org/10.1038/s41467-022-31432-y, 2022.

Ruff, L., Kauffmann, J. R., Vandermeulen, R. A., Montavon, G., Samek, W., Kloft, M., Dietterich, T. G., and Müller, K.-R.: A Unifying Review of Deep and Shallow Anomaly Detection, Proceedings of the IEEE, 109, 756–795, https://doi.org/10.1109/JPROC.2021.3052449, 2021.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


Russo, E. and Domeisen, D. I. V.: Increasing Intensity of Extreme Heatwaves: The Crucial Role of Metrics, Geophysical Research Letters, 50, e2023GL103 540, https://doi.org/10.1029/2023GL103540, 2023.

Russo, S., Sillmann, J., and Fischer, E. M.: Top ten European heatwaves since 1950 and their occurrence in the coming decades, Environmental Research Letters, 10, 124 003, https://doi.org/10.1088/1748-9326/10/12/124003, 2015.

Russo, S., Sillmann, J., and Sterl, A.: Humid heat waves at different warming levels, Scientific Reports, 7, 7477, https://doi.org/10.1038/s41598-017-07536-7, 2017.

Salcedo-Sanz, S., Pérez-Aracil, J., Ascenso, G., Del Ser, J., Casillas-Pérez, D., Kadow, C., Fister, D., Barriopedro, D., García-Herrera, R., Giuliani, M., and Castelletti, A.: Analysis, characterization, prediction, and attribution of extreme atmospheric events with machine learning and deep learning techniques: a review, Theoretical and Applied Climatology, https://doi.org/10.1007/s00704-023-04571-5, 2023.

Sander, J., Ester, M., Kriegel, H.-P., and Xu, X.: Density-Based Clustering in Spatial Databases: The Algorithm GDBSCAN and Its Applications, Data Mining and Knowledge Discovery, 2, 169–194, https://doi.org/10.1023/A:1009745219419, 1998.

Sardeshmukh, P. D., Compo, G. P., and Penland, C.: Need for Caution in Interpreting Extreme Weather Statistics, Journal of Climate, 28, 9166–9187, https://doi.org/10.1175/JCLI-D-15-0020.1, 2015.

Schielicke, L. and Pfahl, S.: European heatwaves in present and future climate simulations: a Lagrangian analysis, Weather and Climate Dynamics, 3, 1439–1459, https://doi.org/10.5194/wcd-3-1439-2022, 2022.

Shamekh, S., Lamb, K. D., Huang, Y., and Gentine, P.: Implicit learning of convective organization explains precipitation stochasticity, Proceedings of the National Academy of Sciences, 120, e2216158 120, https://doi.org/10.1073/pnas.2216158120, 2023.

Soci, C., Hersbach, H., Simmons, A., Poli, P., Bell, B., Berrisford, P., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Radu, R., Schepers, D., Villaume, S., Haimberger, L., Woollen, J., Buontempo, C., and Thépaut, J.: The ERA5 global reanalysis from 1940 to 2022, Quarterly Journal of the Royal Meteorological Society, https://doi.org/10.1002/qj.4803, 2024.

Stefanon, M., D'Andrea, F., and Drobinski, P.: Heatwave classification over Europe and the Mediterranean region, Environmental Research Letters, 7, 014 023, https://doi.org/10.1088/1748-9326/7/1/014023, 2012.

Tian, Y., Kleidon, A., Lesk, C., Zhou, S., Luo, X., Ghausi, S. A., Wang, G., Zhong, D., and Zscheischler, J.: Characterizing heatwaves based on land surface energy budget, Communications Earth & Environment, 5, 1–9, https://doi.org/10.1038/s43247-024-01784-y, 2024.

Tomczyk, A. M., Sulikowska, A., Bednorz, E., and Półrolniczak, M.: Atmospheric circulation conditions during winter warm spells in Central Europe, Natural Hazards, 96, 1413–1428, https://doi.org/10.1007/s11069-019-03621-4, 2019.

van der Maaten, L. and Hinton, G.: Visualizing Data using t-SNE, Journal of Machine Learning Research, 9, 2579–2605, http://jmlr.org/papers/v9/vandermaaten08a.html, 2008.

Van Oldenborgh, G. J., Wehner, M. F., Vautard, R., Otto, F. E. L., Seneviratne, S. I., Stott, P. A., Hegerl, G. C., Philip, S. Y., and Kew, S. F.: Attributing and Projecting Heatwaves Is Hard: We Can Do Better, Earth's Future, 10, e2021EF002 271, https://doi.org/https://doi.org/10.1029/2021EF002271, 2022.

Vautard, R., Cattiaux, J., Happé, T., Singh, J., Bonnet, R., Cassou, C., Coumou, D., D'Andrea, F., Faranda, D., Fischer, E., Ribes, A., Sippel, S., and Yiou, P.: Heat extremes in Western Europe increasing faster than simulated due to atmospheric circulation trends, Nature Communications, 14, 6803, https://doi.org/10.1038/s41467-023-42143-3, 2023.

Wang, C., Li, Z., Chen, Y., Li, Y., Ouyang, L., Zhu, J., Sun, F., Song, S., and Li, H.: Changes in Global Heatwave Risk and Its Drivers Over One Century, Earth's Future, 12, e2024EF004 430, https://doi.org/https://doi.org/10.1029/2024EF004430, 2024.

Wang, J., Guan, Y., Wu, L., Guan, X., Cai, W., Huang, J., Dong, W., and Zhang, B.: Changing Lengths of the Four Seasons by Global Warming, Geophysical Research Letters, 48, e2020GL091 753, https://doi.org/https://doi.org/10.1029/2020GL091753, 2021.

https://doi.org/10.5194/egusphere-2025-2460
Preprint. Discussion started: 16 June 2025


World Economic Forum: The Cost of Inaction: A CEO Guide to Navigating Climate Risk, https://www.weforum.org/publications/the-cost-of-inaction-a-ceo-guide-to-navigating-climate-risk/, in collaboration with Boston Consulting Group (BCG), 2024.

Yuan, X., Chen, X., Ochege, F. U., Hamdi, R., Tabari, H., Li, B., He, B., Zhang, C., De Maeyer, P., and Luo, G.: Weakening of global terrestrial carbon sequestration capacity under increasing intensity of warm extremes, Nature Ecology & Evolution, pp. 1–10, https://doi.org/10.1038/s41559-024-02576-5, 2024.

Zhang, R., Sun, C., Zhu, J., Zhang, R., and Li, W.: Increased European heat waves in recent decades in response to shrinking Arctic sea ice and Eurasian snow cover, npj Climate and Atmospheric Science, 3, 1–9, https://doi.org/10.1038/s41612-020-0110-8, 2020.

Zhang, X., Alexander, L., Hegerl, G. C., Jones, P., Tank, A. K., Peterson, T. C., Trewin, B., and Zwiers, F. W.: Indices for monitoring changes in extremes based on daily temperature and precipitation data, WIREs Climate Change, 2, 851–870, https://doi.org/10.1002/wcc.147, 2011.

Zhao, Y., Strebel, D., Derome, D., Esau, I., Li, Q., and Carmeliet, J.: Using clustering to understand intra-city warming in heatwaves: insights into Paris, Montreal, and Zurich, Environmental Research Letters, 19, 064 002, https://doi.org/10.1088/1748-9326/ad456f, 2024.

Zhu, J.-J., Yang, M., and Ren, Z. J.: Machine Learning in Environmental Research: Common Pitfalls and Best Practices, Environmental Science & Technology, 57, 17 671–17 689, https://doi.org/10.1021/acs.est.3c00026, 2023.

Zschenderlein, P., Pfahl, S., Wernli, H., and Fink, A. H.: A Lagrangian analysis of upper-tropospheric anticyclones associated with heat waves in Europe, Weather and Climate Dynamics, 1, 191–206, https://doi.org/10.5194/wcd-1-191-2020, 2020.