
Decoding Climate Extremes: A Deep Learning Approach

Shambhavi Jilkar

Language Technology Institute
Carnegie Mellon University
Pittsburgh, PA 15213
sjilkar@andrew.cmu.edu

Tyler Ho

Language Technology Institute
Carnegie Mellon University
Pittsburgh, PA 15213
tylerho@andrew.cmu.edu

Changwook Shim

Language Technology Institute
Carnegie Mellon University
Pittsburgh, PA 15213
cshim@andrew.cmu.edu

Miguel Almeida

Language Technology Institute
Carnegie Mellon University
Pittsburgh, PA 15213
malmeida@andrew.cmu.edu

Abstract

Accurate understanding of climate patterns is critical for predicting and mitigating the impacts of extreme weather events, a problem that is especially relevant given the exacerbation of severe events in the previous few years. However, the complex relationship between large-scale atmospheric circulation patterns and other weather data remains difficult to model. In this study, we employ Autoencoder (AE) and Conditional Autoencoder (CAE) methods to identify and classify climate regimes. Our models compress high-dimensional pressure fields into low-dimensional latent representations that capture meaningful weather patterns, from which we use K-Means clustering to identify five distinct climate regimes. Compared to traditional AE approaches, the Conditional Autoencoder is particularly effective at finding better-defined and distinct clusters, offering a promising tool for deep-learning based weather understanding.

1 Overview and Context

1.1 Motivation

Understanding climate regimes and their relationship to other weather data is crucial for climate modeling, forecasting, and resource management. Deep learning approaches have shown promise in identifying these complex relationships between regime identification and precipitation events. While previous work with Variational Autoencoders (R-VAE and RMM-VAE) has established useful methodologies for computing representations of climate phenomena, there is significant potential in exploring how conditional approaches can embed not only reconstructive information but predictive information as well.

1.2 Objective

This project aims to investigate whether Conditional Autoencoders (CAE) can improve upon standard Autoencoders (AE) for atmospheric pattern analysis and short-term precipitation forecasting in the Mediterranean region. In particular, we focus on evaluating reconstruction quality, clustering coherence, and predictive skill. Our specific objectives include:

- Implementing AE and CAE models that effectively compress each timestamp’s feature vector into a meaningful lower-dimensional latent space
- Assessing the reconstruction and prediction capabilities of the CAE model versus traditional AEs
- Applying clustering techniques to the latent representations to identify distinct weather regimes

2 Related Work

2.1 Weather Regimes and Pattern Identification

Weather regimes – persistent and recurrent large-scale circulation patterns – are a common tool for characterizing low-frequency atmospheric variability and teleconnections [1]. Traditionally, such regimes are identified via clustering in a reduced-dimensional space (e.g., k-means on principal components), which captures dominant variability but may not isolate the processes most relevant for specific local impacts [1]. Recent studies leverage deep learning to improve regime identification by incorporating non-linear features and predictive information. For example, Spuler *et al.* [1] propose a Regression Mixture Model VAE (RMM-VAE) that combines a variational autoencoder for non-linear dimensionality reduction with a simultaneous prediction task and probabilistic latent clustering. Applied to Euro-Atlantic circulation patterns with local precipitation as the target, this method discovers weather regimes that are more predictive of precipitation extremes than those from two classical linear methods, while also improving input-field reconstruction. The regimes identified by RMM-VAE were additionally more robust and persistent than those from an alternative machine-learning baseline, highlighting the value of embedding a predictive objective into the unsupervised regime-finding process. We will be using the RMM-VAE model as a baseline for our own experiments.

In a complementary direction, Nielsen *et al.* [2] took a supervised approach, training a deformable convolutional neural network (deCNN) to forecast the North Atlantic-European regime category (e.g., NAO or blocking regime) up to 15 days ahead. Their deCNN model demonstrated significantly better forecasting skill than classical meteorological benchmarks and other ML classifiers. Notably, due to its wider receptive field, the deCNN outperformed a standard CNN architecture at lead times beyond approximately 6 days. Transfer learning (pretraining on a large dataset before fine-tuning) was crucial for capturing the relevant patterns from limited training data. These works show that both unsupervised and supervised deep learning can yield regimes or patterns with improved predictive power, especially when the learning objective is aligned with forecasting targets.

2.2 Deep Learning for Weather and Climate Forecasting

In the past five years, data-driven weather prediction models have made remarkable progress, rivaling traditional numerical weather prediction (NWP) systems. At least eight different research groups since 2022 have reported deep learning models that produce deterministic forecasts more accurate than the ECMWF operational model on various weather variables [7]. One of the first was FourCastNet (Pathak *et al.* [3]), a global neural forecasting model using adaptive Fourier neural operators. FourCastNet matched the short-range forecast accuracy of ECMWF’s Integrated Forecasting System and even outperformed it on finer-scale phenomena like precipitation, while generating a 1-week global forecast in under 2 seconds.

Building on this success, Bi *et al.* [4] introduced Pangu-Weather, a 3D Earth-specific vision transformer model, which surpassed the ECMWF model’s skill across all tested variables. Pangu-Weather achieved about 10% lower 5-day forecast error (RMSE) than the operational ECMWF system and approximately 30% lower error than FourCastNet. Similarly, DeepMind’s GraphCast (Lam *et al.* [5]) employs a graph neural network on a sphere and delivers 10-day forecasts more accurately and much faster than ECMWF’s high-resolution system.

The rapid development of these models has been facilitated by large-scale datasets like ERA5 reanalysis and by community benchmarks such as WeatherBench [6], which standardized data and metrics for comparison. The more recent WeatherBench2 initiative confirms that several deep learning models indeed outperform traditional models on standard skill scores [7].

Despite these breakthroughs, limitations and open challenges remain. Current models are mainly evaluated on average forecast skill, and their ability to predict rare extreme events reliably is less certain. Some studies note that excellent performance on mean error metrics does not necessarily translate into equally skillful predictions for distribution tails [7]. This has led to calls for more rigorous assessment of AI models before operational deployment. Other challenges include ensuring physical consistency and generalization to changing climate conditions — active areas of research where hybrid models and continual learning approaches are being explored.

2.3 Autoencoders and Conditional Architectures

Unsupervised neural networks like autoencoders have become a popular tool in climate science for pattern extraction and dimensionality reduction. An autoencoder compresses high-dimensional climate fields into a latent representation and reconstructs them, essentially performing a non-linear form of PCA [1]. This approach has been used to identify recurrent climate patterns and regimes by clustering in the latent space.

However, a standard autoencoder trained only to reconstruct its input may prioritize the largest-variance features, which are not necessarily the most predictive for future states. To address this, researchers have begun exploring conditional and predictive autoencoder frameworks. In the RMM-VAE discussed above [1], the autoencoder’s training is conditioned on predicting a target variable (precipitation), which yielded latent features more strongly linked to important outcomes. More generally, adding a prediction objective or conditioning variable can guide the learned representation to embed dynamical information.

In our approach, we implement this idea by comparing a standard autoencoder (trained to reconstruct the input sequence $0, \dots, T - 1$) with a conditional autoencoder trained to decode the "next" timesteps ($1, \dots, T$). This setup forces the model to learn features relevant for temporal progression, effectively bridging pattern recognition and prediction. Such an architecture is analogous to a sequence-to-sequence autoencoder or a simplified video-prediction model. By evaluating both variants, we can assess how conditioning on future state information impacts the quality of learned climate patterns.

3 Methodology

3.1 Model Description

We implemented two distinct autoencoder architectures to explore different approaches to atmospheric pattern analysis:

3.1.1 Standard Autoencoder (AE)

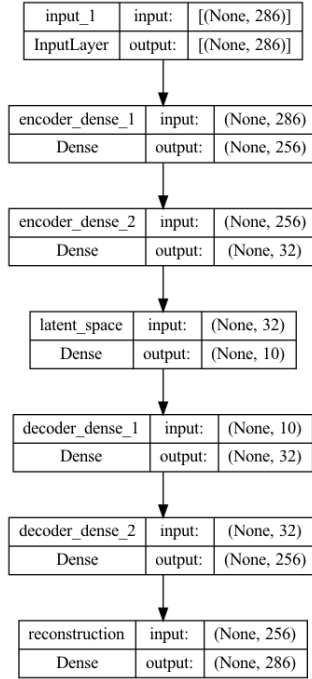
The standard autoencoder is designed to learn efficient representations of z500 fields through efficient encoding and reconstruction:

- **Architecture:** A feed-forward neural network (also shown in figure 1a) with three components:
 - *Encoder:* Input layer (286 dimensions) $\rightarrow$ Dense layer (256 units, ReLU) $\rightarrow$ Dense layer (32 units, ReLU) $\rightarrow$ Latent space (10 units, ReLU)
 - *Bottleneck:* 10-dimensional latent space with ReLU activation
 - *Decoder:* Latent space (10 units) $\rightarrow$ Dense layer (32 units, ReLU) $\rightarrow$ Dense layer (256 units, ReLU) $\rightarrow$ Output layer (286 dimensions, linear)
- **Training Objective:** Minimize the Mean Squared Error (MSE) between the ground truth z500 values and its reconstruction z500 values

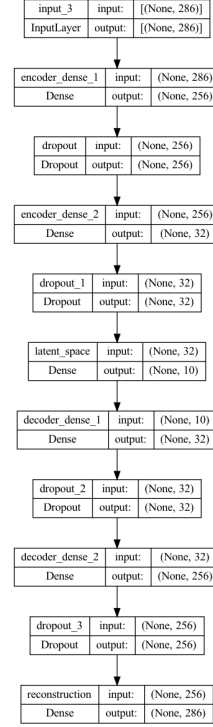
3.1.2 Conditional Autoencoder (CAE)

The conditional autoencoder extends the standard architecture to predict future atmospheric states:

- **Architecture:** A feed-forward neural network similar to the standard AE (also shown in figure 1b) with small modifications to prevent overfitting:



(a) Autoencoder model architecture diagram



(b) Conditional autoencoder model architecture diagram

Figure 1: Comparison of standard and conditional autoencoder architectures

- *Encoder*: Input layer (286 dimensions) → Dense layer (256 units, ReLU) with Dropout (0.2) → Dense layer (32 units, ReLU) with Dropout (0.2) → Latent space (10 units, ReLU)
- *Bottleneck*: 10-dimensional latent space with ReLU activation
- *Decoder*: Latent space (10 units) → Dense layer (32 units, ReLU) with Dropout (0.2) → Dense layer (256 units, ReLU) with Dropout (0.2) → Output layer (286 dimensions, linear)
- **Training Objective**: Minimize the MSE between the predicted ground truth z500 values at time $t + 1$ and the predicted z500 values for time $t + 1$
- **Input/Output Pairing**: Input data at time t is mapped to output data at time $t + 1$
- **Regularization**: Dropout layers (rate=0.2) added to prevent overfitting and improve generalization

3.2 Dataset

Our dataset consisted of a subset of the ERA5 reanalysis data: 500 hPa geopotential height (z500) values covering the North Atlantic-European region (25°N-50°N, 20°W-45°E) from 1940 to 2022

3.2.1 Data Preprocessing

The following preprocessing steps were applied:

1. **Spatial Filtering**: Selected the North Atlantic-European domain for z500 fields
2. **Anomaly Calculation**: Computed daily anomalies by subtracting the calendar day mean to remove seasonal cycles
3. **Normalization**: Standardized fields by dividing by their standard deviation

3.2.2 Training, Validation, and Testing

A temporal 80%-20% split was used for training and validation in order to ensure results generalize well to the future. Training involved 50 epochs with a batch size of 32 for both models.

3.3 Post-Processing and Analysis

After training the models, we applied several analysis techniques:

1. **Latent Space Extraction:** Used the trained encoders to project the z500 values into the 10-dimensional latent space
2. **Clustering:** Applied K-means clustering ($k=5$) to the latent representations to identify weather regimes
3. **Regime Visualization:** Decoded cluster centroids to visualize the characteristic patterns for each regime
4. **Dimensionality Visualization:** Applied t-SNE to project the 10D latent space to 2D for visualization, colored by cluster assignments
5. **Performance Evaluation:** Calculated reconstruction/prediction MSE across the spatial domain

3.4 Evaluation Metrics

We evaluated our models using several complementary metrics:

- **Mean Squared Error (MSE):** $MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$ where y_i represents the true value and $\hat{y}_i$ the predicted/reconstructed value
- **Silhouette Score:** To evaluate the quality of clusters formed in the latent space, measured as $\frac{b(i) - a(i)}{\max(a(i), b(i))}$ where $a(i)$ is the mean distance between point i and all other points in its cluster, and $b(i)$ is the mean distance to points in the nearest neighboring cluster

3.5 Loss Function

The loss function for both models was defined as the Mean Squared Error (MSE) between the input and output, with the following formulation:

$$L(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

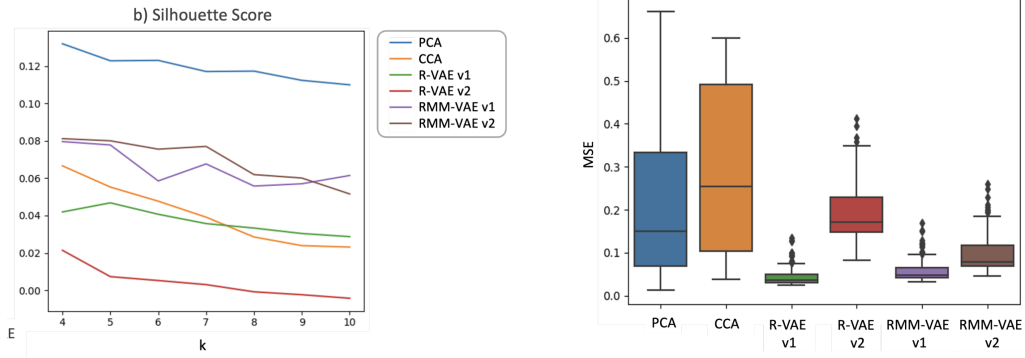
where y is the ground truth z500 field, $\hat{y}$ is the reconstructed or predicted output, and n is the number of spatial points in the field.

4 Baseline and Extensions

4.1 Baseline Selection & Evaluation

Prior work has used regression-augmented variational autoencoders (VAEs) to perform targeted regime identification. The regression-VAE (R-VAE) incorporates a scalar target (e.g., precipitation) into the latent representation by requiring the encoder to jointly predict the target and reconstruct the input (9). Spuler et al. (6) extended this by introducing RMM-VAE, which fits a Gaussian mixture model on the latent space to generate explicit probabilistic weather regimes conditioned on the target variable.

We evaluate these VAE-based models using two standard metrics: silhouette score to assess regime separation, and mean squared error (MSE) to assess reconstruction quality (see Fig. 2). While these models achieve high-fidelity reconstructions and yield regimes predictive of precipitation, they come with significant overhead. Both architectures require large encoder/decoder networks, involve multi-objective training (balancing reconstruction, KL divergence, regression, and clustering losses), and optimize for reconstruction rather than prediction. Consequently, they capture spatial structure well but do not explicitly model how regimes evolve temporally.



(a) Silhouette scores across varying cluster sizes (k). (b) Reconstruction error (MSE)

Figure 2: Comparison of models on (a) cluster quality and (b) reconstruction accuracy.

4.2 Implemented Extensions

To address these limitations, we propose two streamlined alternatives. The first is a standard autoencoder (AE) with a deterministic encoder–decoder structure that compresses each atmospheric field into a compact latent code. By eliminating stochastic sampling and mixture clustering, the AE significantly reduces parameter count and training complexity, while still achieving reconstructions comparable to R-VAE baselines.

The second is a conditional autoencoder (CAE) designed to model temporal evolution. Rather than reconstructing the input state, the CAE is trained in a time-lagged fashion to predict the atmospheric state at time $t + \Delta$ from the state at time t . This encourages the model to capture features that drive regime transitions and dynamics. The CAE maintains the AE’s architectural simplicity, but adds dropout for regularization and ensures strict chronological training to preserve temporal consistency. By aligning the objective with prediction rather than reconstruction, the CAE yields interpretable, temporally-aware regime representations without the overhead of variational mixture models.

5 Results and Analysis

5.1 Model Training Performance

Both models demonstrated successful convergence during training:

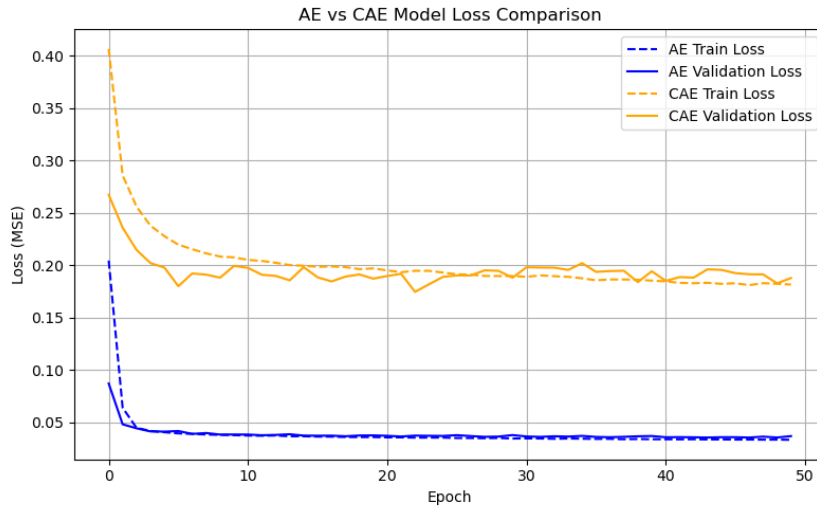


Figure 3: Loss

- **Standard AE:** Achieved a training and validation loss (MSE) of approximately 0.03 after 50 epochs, indicating good reconstruction capabilities without significant overfitting.
- **Conditional AE:** Reached slightly higher final losses with training and validation MSE of approximately 0.18, which is expected given the greater difficulty of prediction compared to reconstruction.

The training curves for the standard AE showed more rapid convergence than the CAE. Additionally, the CAE exhibited slightly higher variability between epochs. Both of these, along with the higher final loss in the CAE, reflect the inherent complexity of predicting future states rather than reconstructing current ones.

5.2 Reconstruction and Prediction Performance

Spatial analysis of reconstruction/prediction errors revealed:

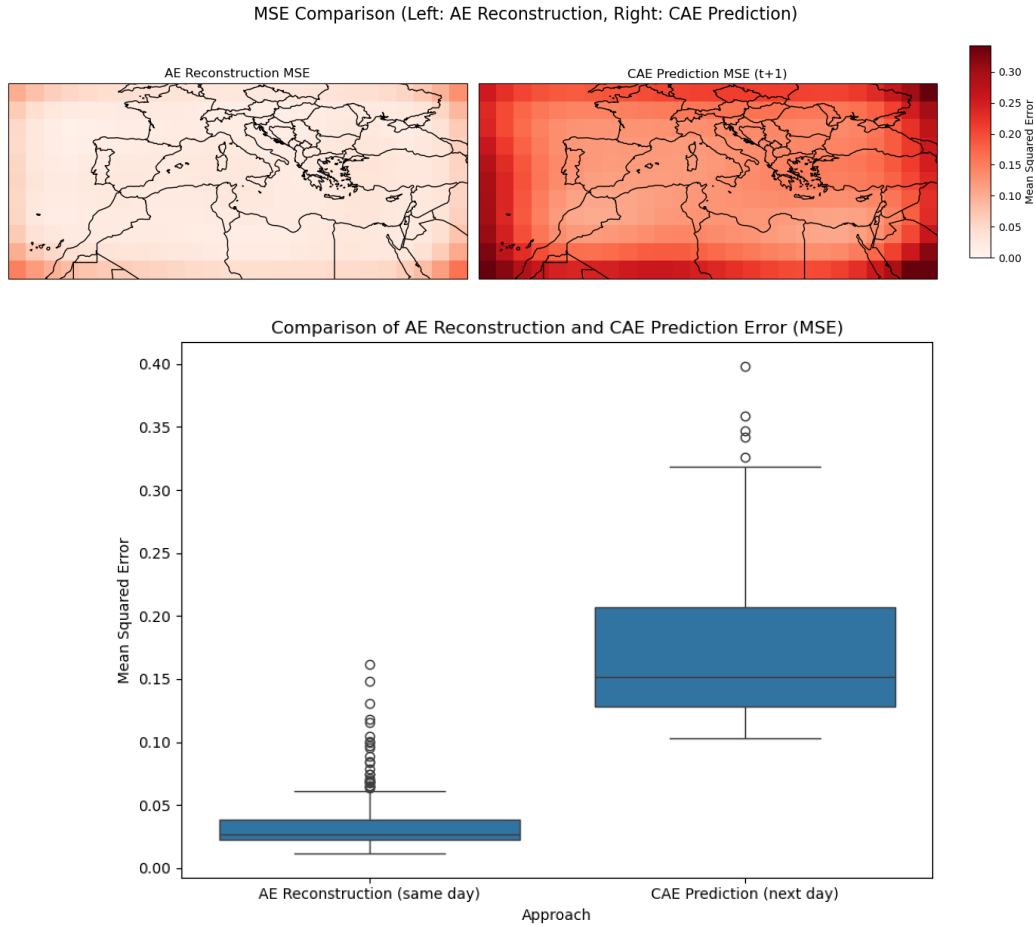


Figure 4: MSE Comparison

- **AE Reconstruction:** Achieved a median MSE of approximately 0.03 across the domain.
- **CAE Prediction:** Achieved a median MSE of approximately 0.15 for next-day prediction.

Both models showed positively skewed MSE distributions across grid points, with higher errors concentrated at the periphery of the domain and lower errors in the central region. The CAE showed a wider spread and higher median error. The AE had a lower MSE than every R-VAE and RMM-VAE model, while the CAE had a MSE comparable superior with R-VAE v2, but inferior to the rest. These results were all expected for a predictive model versus reconstructive models.

5.3 Clustering and Weather Regime Identification

We identified five distinct weather regimes from both the AE and CAE latent spaces:

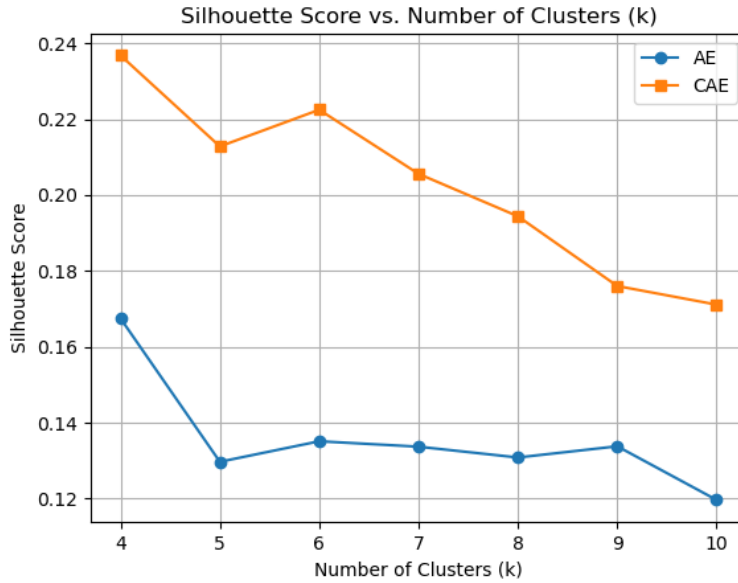
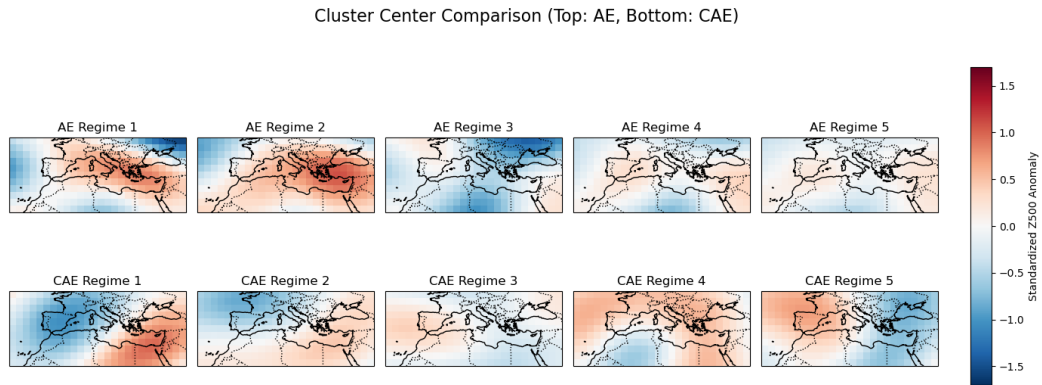


Figure 5: Regime Clustering Comparison

Standard AE: The standard AE found the following regimes:

- **Regime 1:** Strong positive anomaly over the Eastern Mediterranean; strong negative anomaly elsewhere.
- **Regime 2:** Strong positive anomaly over the Eastern Mediterranean.
- **Regime 3:** Strong negative anomaly throughout
- **Regime 4:** Neutral anomaly.
- **Regime 5:** Neutral anomaly.

There is a distinct lack of variation with these regime classifications, meaning the regimes are not very expressive, and likely aren't very useful in understanding climate patterns.

CAE: The CAE found the following regimes:

- **Regime 1:** Strong positive anomaly over the Southeast Mediterranean; strong negative anomaly over the Northwest.

- **Regime 2:** Weak positive anomaly over the Southeast Mediterranean; negative anomaly over the Northwest.
- **Regime 3:** Neutral anomaly.
- **Regime 4:** Weak positive anomaly over Europe and Northeast Africa; weak negative anomaly over Northwest Africa.
- **Regime 5:** Positive anomaly over the West Mediterranean; negative anomaly over the East Mediterranean.

Compared with Standard AE, CAE produces a higher variation in regimes. These regimes are therefore much more expressive and likely to provide useful information when studying climate patterns.

Silhouette Scores: The CAE consistently achieved a Silhouette Score between 0.05-0.10 than the Standard AE. This indicates that the CAE found better-defined and more distinct clusters. Both models had a higher silhouette score than every baseline model.

These results point to a key insight from our analysis: although regime identification is an unsupervised learning task, it benefits from clustering on the latent representations produced by models that have predictive power.

6 Future Directions

Our Conditional Autoencoder (CAE) learns to predict the next day’s atmospheric state from previous time steps, capturing dynamic features that help identify weather regimes. To strengthen this approach, future work can explore using longer input sequences (e.g., 10–30 days) so the model can recognize slowly evolving precursors of regime shifts. Adding more meteorological variables such as temperature, humidity, or wind could help it detect complex relationships.

Transformer models, which attend across all points in space and time, could identify long-range dependencies that lead to regime shifts. For example, a Temporal Transformer or a Vision Transformer adapted to meteorological data could ingest a sequence of atmospheric maps and learn which regions/levels at earlier times are most predictive of the circulation at time t . Recurrent sequence models excel at capturing temporal dependencies, which is crucial for weather regimes that evolve on multi-day timescales. By providing the model with a memory of past states, an LSTM could learn regime dynamics over multiple days or weeks, potentially improving the identification of regime onset and persistence. GANs could enhance realism in regime prediction by generating physically plausible atmospheric states from latent representations. A conditional GAN, trained to produce next-day fields given past circulation, could sharpen regime-specific structures that standard decoders might blur.

Beyond model improvements, regime-based learning offers exciting possibilities for climate science. Because weather regimes reflect the large-scale organization of the atmosphere, tracking their frequency and transitions over time can reveal how climate change may be reshaping patterns that drive extreme events. On the forecasting side, regime-aware models like ours could be integrated into early warning systems, especially for droughts or floods, by recognizing circulation patterns that historically preceded such events. Longer-term, this approach could serve as a bridge between data-driven and physical models — combining the speed and pattern recognition strength of deep learning with the reliability of numerical simulations. Overall, by extending predictive autoencoding with deeper architectures, probabilistic modeling, and operational integration, regime-focused deep learning stands poised to significantly advance both weather forecasting and climate science.

7 Conclusion

This study extended previous work on variational autoencoders for atmospheric pattern recognition by implementing and evaluating two complementary autoencoder approaches: a standard Autoencoder (AE) for efficient data reconstruction and a Conditional Autoencoder (CAE) for next-day atmospheric state prediction.

Our work demonstrates that complex probabilistic frameworks are not always necessary for effective atmospheric pattern recognition. The simpler, more efficient AE and CAE architectures not only match the performance of VAE-based approaches in many respects but also offer unique advantages in terms of training stability, interpretability, and predictive capabilities. The temporal extension provided by the CAE is particularly valuable for applications in weather forecasting and early warning systems.

In summary, this study advances our understanding of atmospheric pattern recognition methodologies and provides practical tools for identifying weather regimes associated with extreme precipitation in drought-prone regions. The models developed here could serve as foundational components for improved precipitation forecasting systems and climate change impact assessments, with potential applications extending beyond Morocco to other regions where understanding the relationship between large-scale circulation and local precipitation is critical for water resource management and disaster preparedness.

Administrative Details

References

- [1] Hersbach, H., et al. “The ERA5 global reanalysis.” *Q. J. R. Meteorol. Soc.* 146, 1999–2049 (2020).
- [2] Rasp, S., Dueben, P. D., Scher, S., Weyn, J. A., Mouatadid, S., & Thuerey, N. “WeatherBench: A benchmark data set for data-driven weather forecasting.” *J. Adv. Model. Earth Syst.* 12, e2020MS002203 (2020).
- [3] Weyn, J. A., Durran, D. R., & Caruana, R. “Improving data-driven global weather prediction using deep convolutional neural networks on a cubed sphere.” *J. Adv. Model. Earth Syst.* 12, e2020MS002109 (2020).
- [4] Mastrantonas, N., et al. “Weather regimes – Moroccan precipitation link in a regional climate simulation: a seasonal perspective.” *Climate Dynamics* 55, 137–153 (2020).
- [5] Falkena, S. K., de Wiljes, J., Weisheimer, A., & Shepherd, T. G. “A probabilistic view on weather regimes.” *J. Climate* 33, 7665–7684 (2020).
- [6] Spuler, F. R., Alonso Balmaseda, M., & Shepherd, T. G. “Identifying probabilistic weather regimes targeted to a local-scale impact variable.” *Environ. Data Sci.* 3, e25 (2024).
- [7] Springer, S., Laio, A., Galfi, V. M., & Lucarini, V. “Unsupervised detection of large-scale weather patterns in the Northern Hemisphere via Markov State Modeling: from blockings to teleconnections.” *npj Climate Atmos. Sci.* 7, 105 (2024).
- [8] Zhao, Q., Honnorat, N., Adeli, E., & Pohl, K. M. “Variational autoencoder with a truncated mixture of Gaussians for functional connectivity analysis.” *IPMI* (2019).
- [9] Zhao, Q., et al. “Regression Variational Autoencoder for pattern–target association.” (2019).
- [10] Ye, J., & Bors, A. G. “Mixtures of variational autoencoders.” *NeurIPS Workshop* (2020).
- [11] Weyn, J. A., Durran, D. R., & Caruana, R. “Can machines learn to predict weather? Using deep learning to predict gridded 500-hPa geopotential height from historical weather data.” *J. Adv. Model. Earth Syst.* 11, 2680–2693 (2019).
- [12] Pathak, J., et al. “FourCastNet: A global data-driven high-resolution weather model using adaptive Fourier neural operators.” *ArXiv Preprint arXiv:2202.11214* (2022).
- [13] Bi, K., et al. “Accurate medium-range global weather forecasting with 3D neural networks.” *Nature* 619, 533–538 (2023).
- [14] Lam, R., et al. “Learning skillful medium-range global weather forecasting.” *Science* 382, 1416–1421 (2023).

- [15] Alet, F., et al. “Probabilistic weather forecasting with machine learning.” *Nature* 637, 84–90 (2025).
- [16] Mayer, K., & Barnes, E. “Subseasonal forecasts with regime classification.” *Geophys. Res. Lett.* 48, e2021GL094151 (2021).
- [17] Domeisen, D. I., Grams, C. M., & Papritz, L. “The role of the stratosphere in North Atlantic weather regime predictability.” *Weather Clim. Dyn.* 1, 337–357 (2020).
- [18] Hochman, A., et al. “Challenges in quantifying and understanding rainfall extremes in the Mediterranean.” *Hydrology and Earth System Sciences* 25, 3885–3917 (2021).
- [19] Chen, S., et al. “ClimaX: A foundation model for weather and climate.” *ICLR Climate Change AI Workshop* (2023).
- [20] Stengel, K., et al. “Adversarial super-resolution of climate simulations.” *Proc. Natl. Acad. Sci. USA* 117, 16805–16815 (2020).
- [21] Ham, Y.-G., Kim, J.-H., & Luo, J.-J. “Deep learning for multi-year ENSO forecasts.” *Nature* 573, 568–572 (2019).
- [22] Kashinath, K., Mustafa, M., Albert, A., et al. “Physics-informed machine learning: case studies for weather and climate modelling.” *Phil. Trans. R. Soc. A* 379, 2194 (2021).
- [23] Dueben, P., Weyn, J., Rasp, S., Scher, S., & Thuerey, N. “The rise of data-driven weather forecasting.” *Bull. Amer. Meteorol. Soc.* 103, E2333–E2340 (2022).
- [24] Baker, A., Barnes, E., & Deser, C. “Predicting weather regime transitions in sub-seasonal forecasts using machine learning.” *J. Geophys. Res. Atmos.* 127, e2022JD036640 (2022).
- [25] Price, D., & Rasp, S. “Leveraging diffusion models for generative weather prediction.” *ArXiv Preprint arXiv:2212.14714* (2022).

Team Contributions

- Shambhavi Jilkar: Conceived the project idea. Conducted an in-depth literature review and proposed the modeling approach. Also contributed to the experimentation and analysis.
- Tyler Ho: Played a key role in implementing the autoencoder architectures, tuning hyperparameters, and running experiments. Worked closely on training, debugging, and interpreting model performance.
- Changwook Shim: Actively contributed to model development and evaluation. Focused on refining the training pipeline, analyzing output metrics, and validating the robustness of results across model variants.
- Miguel Almeida: Contributed to synthesizing results, performed data feature analysis, and led the design and delivery of the presentation.

GitHub Repository

<https://github.com/Tyrest/Neural-Netbusters>