Variational Autoencoders for Efficient and Interpretable Climate Downscaling

Nicolas Gauthier [1,2] Darcy Bird [1] Kai Blahnik [1,3]

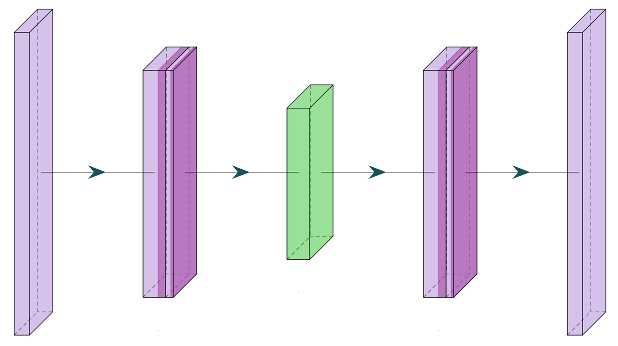
[1]Florida Museum of Natural History, University of Florida [2]Department of Geography, University of Florida [3]Department of Computer &Information Science &Engineering, University of Florida

Introduction

Climate impact assessments and adaptation planning require high-resolution climate data, but global climate models typically operate at coarse spatial scales (>10km). While traditional statistical downscaling methods like EOF analysis have proven valuable, deep learning approaches offer new possibilities for capturing complex spatial relationships and generating realistic climate fields. VAEs are generative neural networks that learn compressed representations of high-dimensional data while maintaining probabilistic properties. Unlike traditional dimensionality reduction techniques (e.g., PCA/EOF analysis), VAEs can capture non-linear relationships and generate new, physically consistent samples of climate states.

Autoencoders: Learning compressed representations with neural networks

Autoencoders are neural networks that compress high-dimensional data through a learned latent space, extending the linear dimensionality reduction of principal components analysis to capture non-linear patterns and interactions. The architecture consists of an **encoder that maps inputs to a compact latent representation** and a **decoder that reconstructs the original data**, with both components optimized jointly through gradient descent. This foundation enables more sophisticated generative approaches like variational autoencoders (VAEs).

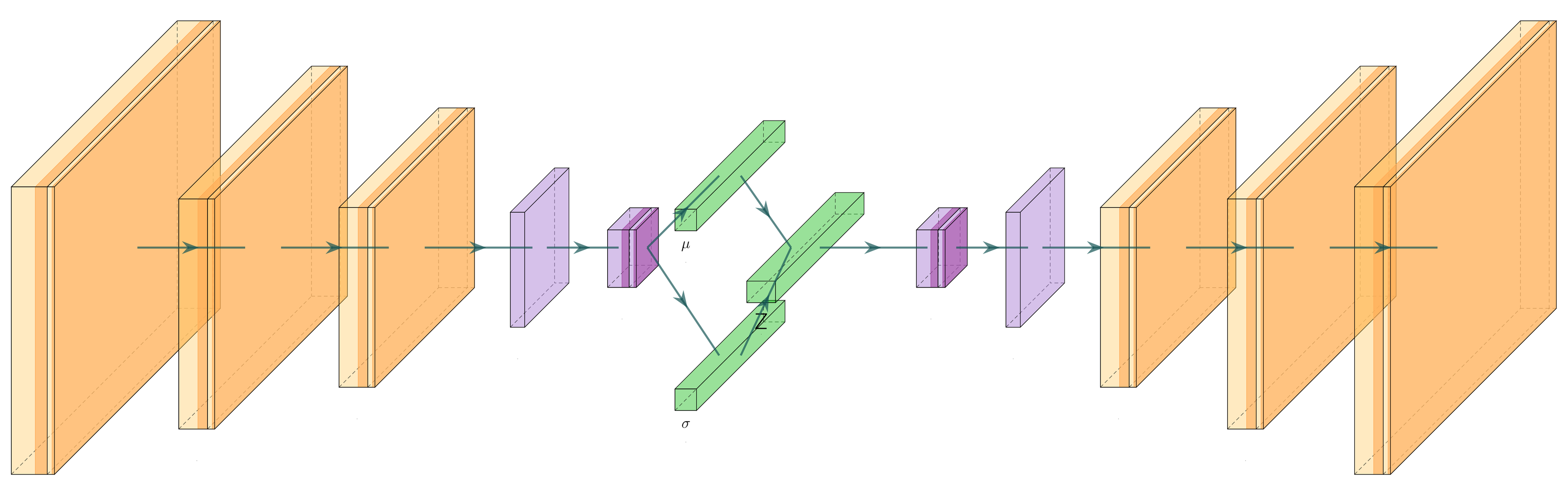


Variational Autoencoders

VAEs augment standard autoencoders with probabilistic latent representations, mapping climate fields to distributions rather than point estimates [1]. For spatial climate data, convolutional encoders and decoders efficiently capture local patterns and spatial dependencies. By constraining the latent space to approximate a standard normal distribution, VAEs learn a structured representation that enables both reconstruction and generation of new samples. This probabilistic framework makes them particularly suited for climate applications, where uncertainty quantification and ensemble generation are crucial for downstream tasks like downscaling [2].

Key components:

- Convolutional encoder maps climate fields to a probability distribution in latent space
- Structured latent space enables smooth interpolation and sampling
- Convolutional decoder generates physically consistent climate fields from latent vectors



Useful facts about VAEs

- During training, the encoder is learned before the decoder, and is much more generalizable between datasets [3].
- VAEs converge on the Principal Components in simpler, linear domains when the latent space is small, but are more parameter efficient, flexible, and modular [4, 5].
- Increasing the size of the latent space does not lead to overfitting, the VAE will learn to ignore latent dimensions that are not relevant
- There is a tradeoff between the skill at reconstruction and generalization, so that better generative models with a richer latent space tend to result in blurrier images [6]. This can be fixed with more specific loss functions.

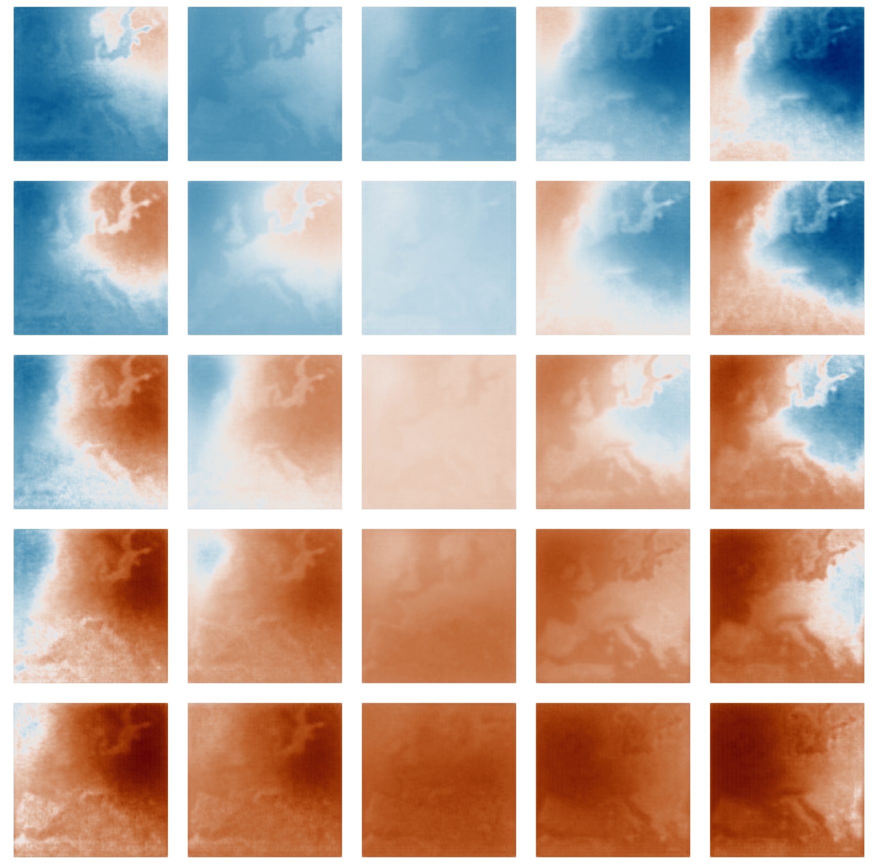
Case study: Monthly air temperatures over Europe

We evaluate our approach using monthly 2-meter air temperature anomalies from three complementary datasets of varying resolution and time spans:

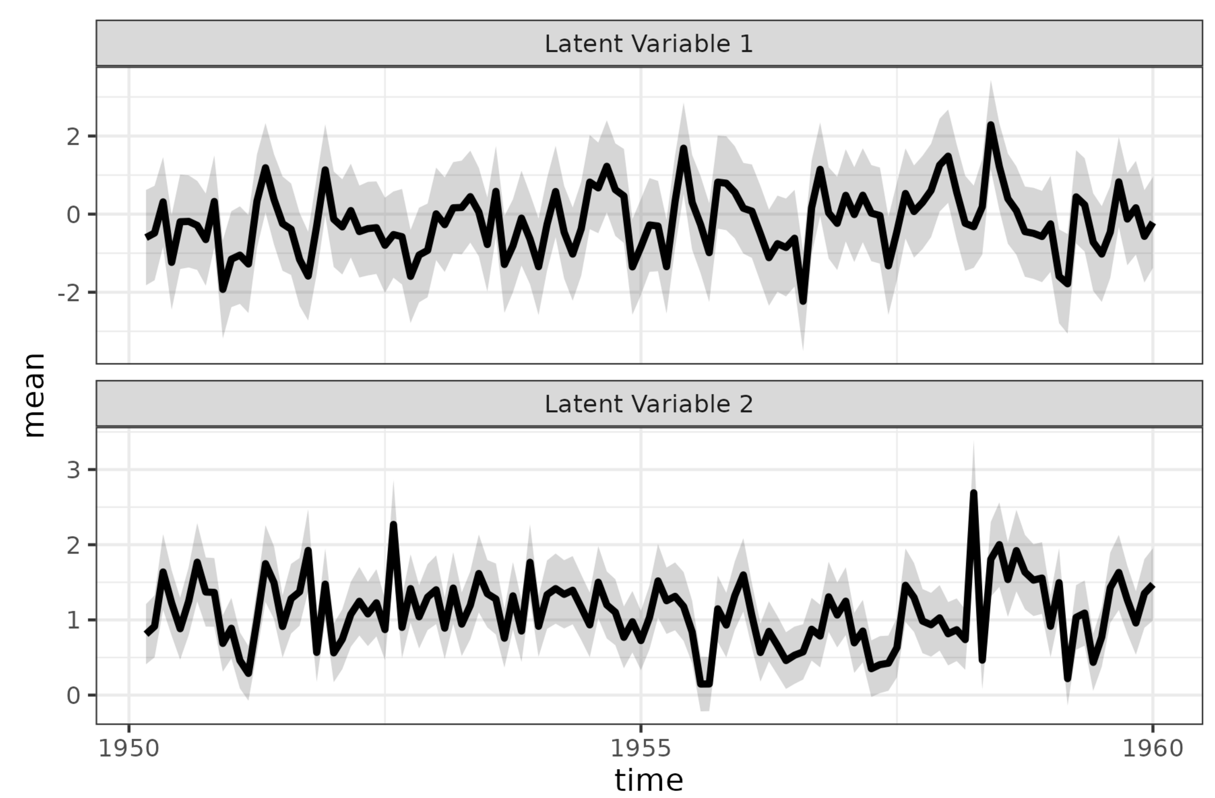
- CESM Last Millennium Ensemble (2° resolution): Extensive sampling of internal variability from 850-2005 provides rich pretraining data [7]
- ClimEX 50-member ensemble (12km): High-resolution climate model simulations over Europe from 1950-2100 [8]
- ERA5 Reanalysis (0.25° resolution): Observational benchmark for evaluating performance in observation period (1950-2024) [9]

Exploring the latent climate space

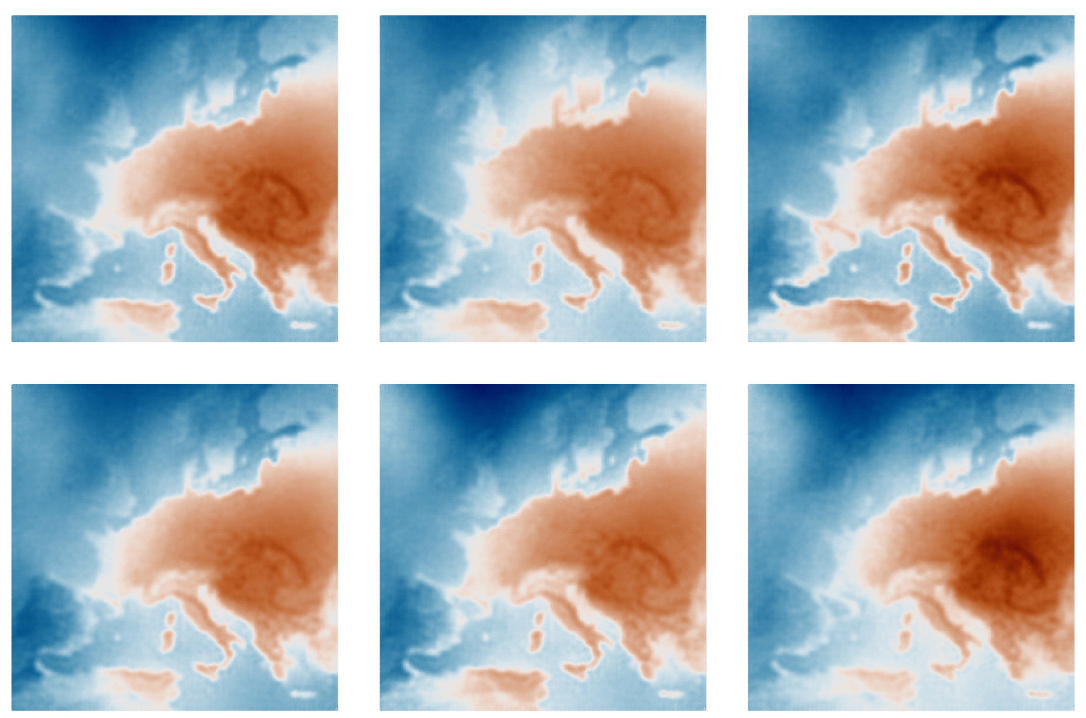
The VAE learns a compact representation where nearby points share coherent meteorological features, revealing distinct clusters of weather patterns with physically meaningful transitions between climate regimes. This structured latent space enables both analysis of atmospheric dynamics and generation of new climate states.



By tracking the evolution of latent variables over time, we can characterize the uncertainty in the climate state's trajectory and identify key transitions between weather patterns.



The probabilistic nature of the latent space enables generation of physically consistent ensemble members, capturing natural variability while maintaining large-scale spatial correlations.



What does this mean for climate downscaling?

Traditional downscaling approaches require time-matched pairs of low- and high-resolution data, limiting the amount of data available for training. Instead, we propose a more flexible approach using VAEs and **transfer learning**: first pretraining separate VAEs on extensive low-resolution GCM output and limited high-resolution regional simulations. These pretrained components can then be flexibly combined and fine-tuned for specific downscaling tasks, with several possible strategies:

1. **Feature Extraction:** Use the frozen low-resolution encoder as a general-purpose feature extractor, learning only a simple mapping from its latent space to high-resolution outputs
2. **Latent Space Translation:** Connect pretrained low- and high-resolution VAEs through a learnable adaptation layer between their latent spaces
3. **Architecture Integration:** Use pretrained components to initialize more complex architectures like U-Net, then fine-tune the entire system end-to-end on observational data.

Benefits of transfer learning:

- Self-supervised pretraining on extensive GCM data reduces reliance on paired training samples
- Modular architecture enables flexible combination of pretrained components
- Probabilistic framework enables uncertainty quantification and ensemble generation
- Interpretable latent space maintains physical consistency and enables analysis of climate patterns
- Particularly valuable for seasonal-to-interannual downscaling where paired samples are limited

Next Steps

- Explore asymmetric architectures for encoder and decoder networks drawing on state-of-the-art downscaling tools
  - More sophisticated convolutional architectures used in image super-resolution tasks [10]
  - Diffusion models for enhanced sample quality [11]
  - Vision transformers for capturing long-range dependencies [12]
- Investigate additional loss functions including adversarial losses to balance reconstruction fidelity with generation diversity or to enforce physical constraints [13]
- Expand pretraining to diverse climate datasets to improve feature generalization
- Leverage structured latent space for data assimilation applications [2]
- Study robustness of learned representations across climate regimes and extreme events

Acknowledgments

This research was supported by a grant from the **Burroughs Wellcome Fund** (BWF1057896) and the US **National Science Foundation** (NSF2409068).

References

[1] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," 2022.

[2] P. Brohan, "Machine learning for data asimilation," 4 2022. Documentation available at http://brohan.org/Proxy_20CR/.

[3] L. Bonheme and M. Grzes, "How do Variational Autoencoders Learn? Insights from Representational Similarity," Sept. 2022.

[4] M. Rolinek, D. Zietlow, and G. Martius, "Variational Autoencoders Pursue PCA Directions (by Accident)," Apr. 2019.

[5] Z. Li and H. Liu, "Beta-vae has 2 behaviors: Pca or ica?," 2023.

[6] I. Higgins, L. Matthey, A. Pal, C. P. Burgess, X. Glorot, M. M. Botvinick, S. Mohamed, and A. Lerchner, "beta-vae: Learning basic visual concepts with a constrained variational framework.," *ICLR (Poster)*, vol. 3, 2017.

[7] B. L. Otto-Bliesner, E. C. Brady, J. Fasullo, A. Jahn, L. Landrum, S. Stevenson, N. Rosenbloom, A. Mai, and G. Strand, "Climate Variability and Change since 850 C.E.: An Ensemble Approach with the Community Earth System Model (CESM)," *Bulletin of the American Meteorological Society*, vol. 97, no. 5, pp. 787–801, 2016.

[8] M. Leduc, A. Mailhot, A. Frigon, J.-L. Martel, R. Ludwig, G. B. Brietzke, M. Giguère, F. Brissette, R. Turcotte, M. Braun, *et al.*, "The climex project: A 50-member ensemble of climate change projections at 12-km resolution over europe and northeastern north america with the canadian regional climate model (crcm5)," *Journal of Applied Meteorology and Climatology*, vol. 58, no. 4, pp. 663–693, 2019.

[9] H. Hersbach, B. Bell, *et al.*, "The era5 global reanalysis," *Quarterly Journal of the Royal Meteorological Society*, vol. 146, no. 730, pp. 1999–2049, 2020.

[10] C. A. G. Gonzalez, "DL4DS – Deep Learning for empirical DownScaling," May 2022.

[11] N. Shidqi, C. Jeong, S. Park, E. Zeller, A. B. Nellikkattil, and K. Singh, "Generating High-Resolution Regional Precipitation Using Conditional Diffusion Model," Dec. 2023.

[12] T. Nguyen, J. Brandstetter, A. Kapoor, J. K. Gupta, and A. Grover, "ClimaX: A foundation model for weather and climate," July 2023.

[13] P. Harder, V. Ramesh, A. Hernandez-Garcia, Q. Yang, P. Sattigeri, D. Szwarcman, C. Watson, and D. Rolnick, "Physics-Constrained Deep Learning for Climate Downscaling," Mar. 2023.