Supporting information of "CERA: A Framework for Improved Generalization of Machine Learning Models to Changed Climates"

Shuchang Liu[1], **Paul A. O'Gorman**[1]

[1]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology

Corresponding author: Shuchang Liu, shuchang.liu@hotmail.com

Contents of this file

1 Training details

In each iteration, the autoencoder processes 8192 input samples, with 4096 from the control climate and 4096 from the +4 K climate. The reconstruction loss is computed on the full batch, and the latent alignment loss is evaluated between the control and +4 K samples. In the mean time, the downstream prediction loss is calculated using only labeled control-climate data. The final loss is a weighted sum of reconstruction loss, latent alignment loss and prediction loss to enable joint updates of the autoencoder and predictor.

The model is optimized using the AdamW optimizer with learning rates of 3×10^{-3} and and a weight decay of 10^{-3} for both the autoencoder and the MLP predictor. A learning rate scheduler with exponential decay and 4000 linear warm up steps is applied to both learning rates. Training is run for 30 epochs.

After initial tests on generalization performance, hyperparameters are selected through a sweep aimed at balancing latent alignment and predictive accuracy. The weight on the latent alignment loss, λ_{EMD}, was chosen as the largest value that did not lead to latent space collapse, which we define as the average standard deviation of the latent representations falling below 0.1. Values of λ_{EMD} ranging from 10^{-1} to 10^{-5} were tested, and $\lambda_{\text{EMD}} = 10^{-4}$ was selected as the largest stable value. The weight on the supervised prediction loss, λ_{pred}, was progressively decreased from 1 to 0.001. Since $\mathcal{L}_{\text{total}}$ depends on this weighting, we effectively selected λ_{pred} by minimizing the unweighted combination of the underlying loss terms. For λ_{pred} values between 1 and 0.01, $\mathcal{L}_{\text{pred}}$ varies by less than 5% at the end of training, and $\mathcal{L}_{\text{reconstruction}}$ remains small ($< 10^{-4}$). Interestingly, $\text{EMD}(Z_0, Z_{+4K})$ decreases when $\mathcal{L}_{\text{reconstruction}}$ becomes smaller, presumably because better reconstructions more strongly constrain the latent space. Consequently, $\text{EMD}(Z_0, Z_{+4K})$ reaches its minimum when λ_{pred} is set to 0.01. Further decreasing λ_{pred} below 0.01 led to an increase in $\mathcal{L}_{\text{pred}}$. Therefore, the best-performing configuration used $\lambda_{\text{pred}} = 0.01$. This combination provided a stable latent alignment without compromising predictive performance on the labeled control-climate data.

For the alternative version of our analysis (results shown in Fig. 3) and to enable direct comparison with Beucler et al. (2024), we fine-tuned the λ_{EMD} parameter to account for the modified input/output configuration. We found that we could use a larger λ_{EMD} to strengthen alignment, stopping before latent space collapse (defined as std $<$ 0.1). We selected $\lambda_{\text{EMD}} = 10^{-3}$ as the largest value that preserved sufficient latent variability.

For the baseline models, they are the same as CERA without the autoencoder. We used the same learning rate (3×10^{-3} and and a weight decay of 10^{-3}) for training the baseline models.

2 Formula for instantaneous surface precipitation rate

The instantaneous surface precipitation rate for both the ML models and the high-resolution simulation is computed by vertically integrating the microphysical tendency of total condensate, q_{T_mic}, with density weighting:

$$P_{\text{tot}}(z=0) = -\int_0^\infty \rho_0 q_{T\_\text{mic}} dz. \tag{1}$$

For simplicity, we exclude the surface ice sedimentation flux, which is typically small. This formulation is similar to Equation S6 of Yuval et al. (2021), but we note that their equation omits a negative sign.

References

Beucler, T., Gentine, P., Yuval, J., Gupta, A., Peng, L., Lin, J., ... O'Gorman, P. A. (2024). Climate-invariant machine learning. *Science Advances*, *10*(6), eadj7250.

Yuval, J., O'Gorman, P. A., & Hill, C. N. (2021). Use of neural networks for stable, accurate and physically consistent parameterization of subgrid atmospheric processes with good performance at reduced precision. *Geophysical Research Letters*, *48*(6), e2020GL091363.