Supporting Information for "Non-local parameterization of atmospheric subgrid processes with neural networks"

Peidong Wang[1], Janni Yuval[1], and Paul A. O'Gorman[1]

[1]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Contents of this file

1. Text S1 to S4
2. Figures S1 to S14

Introduction

Here we describe the layer-wise relevance propagation rules that we use (Text S1), how we calculate the instantaneous precipitation rate (Text S2), how we calculate the convective available potential energy (Text S3), and how we calculate other measures for convective and symmetric instabilities (Text S4). Figure S1-S14 show supporting figures for the main text.

Text S1. Layer-wise relevance propagation rules

The general relevance propagation rule is as follows (Montavon et al., 2019):

$$R_j = \sum_k \frac{a_j \cdot (w_{jk} + \gamma w_{jk}^+)}{\epsilon + \sum_{0,j'} a'_j \cdot (w_{j'k} + \gamma w_{j'k}^+)} R_k, \tag{1}$$

where R stands for the relevance, j and k are the neurons in two consecutive layers, with j indexing a layer closer to the input layer and k indexing a layer closer to the output layer, a_j is the value propagated forward by neuron j (activated by ReLU), w_{jk} is the weight that connects neuron j to k, and w_{jk}^+ is the same as w_{jk} but only includes positive weights. The sum over j' is over all neurons in the layer indexed by j but also including an extra neuron with index 0 representing the bias. When back-propagating the relevance from the first hidden layer to the input layer, a different propagation rule is applied (Montavon et al., 2017), and the relevance is calculated as:

$$R_j = \sum_k \frac{w_{jk}^2}{\sum_{j'} w_{j'k}^2} R_k, \tag{2}$$

where j indexes the input in the input layer, and k indexes the neuron in the first hidden layer. There are two tunable parameters in LRP, namely, ϵ and γ, which reduce the noise and favors positive weights, respectively. We present the LRP results when both ϵ and γ set to zero for simplicity, but we verified that the results presented are not sensitive to different choices of the parameters ϵ and γ (Figure S11).

Text S2. Calculation of instantaneous precipitation rate

Following Yuval, O'Gorman, and Hill (2021), the instantaneous precipitation rate is estimated as:

$$P_{tot}(z=0) = \int_0^\infty \rho_0 \left(\frac{\partial q_p}{\partial t} \right)_{\text{micro}} dz, \tag{3}$$

where $P_{\text{tot}}(z = 0)$ is the total precipitation rate at the surface (including all phases) and $(\partial q_p/\partial t)_{\text{micro}}$ is the tendency of the precipitating water mixing ratio. We use this approach for the neural-network parameterization which does not make a separate prediction of the surface precipitation. We also use this approach to calculate the "true" instantaneous precipitation on the coarse grid by using the coarse-grained $(\partial q_p/\partial t)_{\text{micro}}$ from the hi-res simulation. This was necessary because only 3-hourly averaged precipitation was stored as simulation output. In both cases, we only integrate over the bottom 30 vertical levels from 0 to 13.4 km which is the height range of the neural-network outputs.

Text S3. Convective available potential energy (CAPE)

Following Muller, O'Gorman, and Back (2011), we calculate CAPE (with any convective inhibition removed) as:

$$\text{CAPE} = \int_0^{z_{LZB}} g \left[\frac{T_p - T_e}{T_e} + \left(\frac{R_v}{R_d} - 1 \right) (q_{v,p} - q_{v,e}) - (q_{c,p} - q_{c,e}) - (q_{i,p} - q_{i,e}) \right] dz, \quad (4)$$

where z_{LZB} is the level of zero buoyancy, g is the gravitational acceleration, T_e is the temperature of the environment, T_p is the temperature of the parcel following pseudo-adiabatic ascent, R_v and R_d are the gas constants for water vapor and dry air, respectively, q is the mixing ratio where subscript v stands for vapor, c for cloud liquid, i for ice, e for environment, and p for parcel.

Text S4. Quantification of symmetric and convective instabilities

To provide measures of upright convective and symmetric instability, we first calculate the saturation moist static energy (MSE^*) as:

$$\text{MSE}^* = c_p T + gz + L q_v^*, \quad (5)$$

where c_p is the specific heat of air at constant pressure, T is the absolute temperature, g is the gravitational acceleration, z is the height, q_v^* is the saturation vapor mixing ratio, and $L = L_c + L_f(1 - \omega_n)$ with L_c the latent heat of condensation, L_f the latent heat of fusion, and ω_n the phase partition function for non-precipitating water in SAM. MSE* is conserved in SAM for saturated moist-adiabatic processes in the absence of melting and freezing of cloud condensates and precipitation.

We diagnose conditional instability for upright convection at a given level when $\partial \text{MSE}^*/\partial z < 0$. Similarly, we diagnose conditional symmetric instability when the saturation moist potential vorticity MPV* is negative (Korty & Schneider, 2007). We modify MPV* to use MSE* and the anelastic reference density profile as follows:

$$\text{MPV}^* = \frac{\text{sign}(f)}{c_p \rho_0(z)} (\nabla \times \mathbf{v} + f\hat{z}) \cdot \nabla \text{MSE}^*, \tag{6}$$

where $\rho_0(z)$ is the reference density profile, $\mathbf{v}$ is the three-dimensional velocity, and f is the Coriolis parameter on the equatorial β-plane. By multiplying MPV* with the sign of f, the atmosphere is conditionally symmetrically unstable when MPV$^* < 0$ in both hemispheres. Bands of increased stability are evident near the melting lines in Figures S8 and S9, and this is related to melting and freezing of precipitation and condensate.

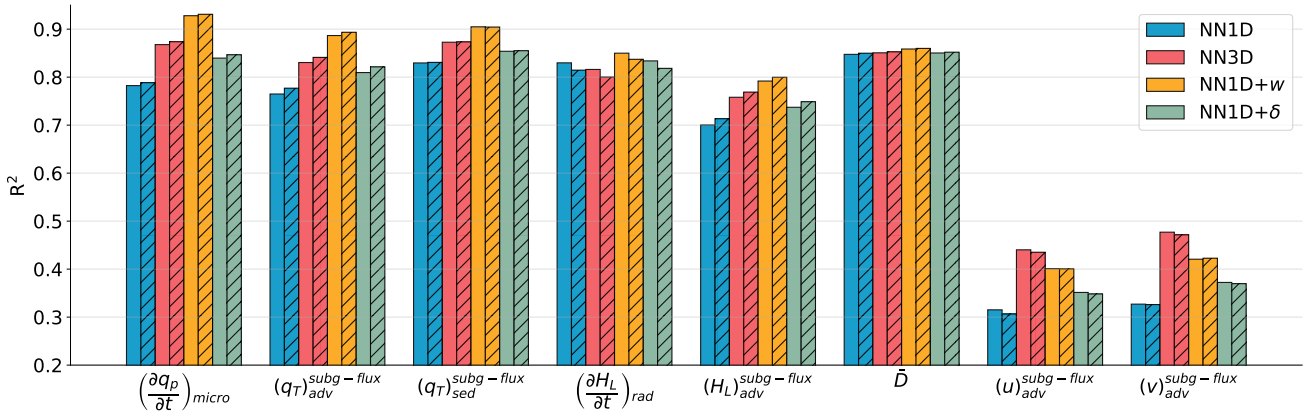

Figure S1. Similar to Figure 2, but also showing results for a different train/test split of the data. The bars without hatches are the same as the bars in Figure 2 (trained on 50% of the data and tested on 40% of the data), and the bars with hatches indicate more of the data is used for training (trained on 80% of the data and tested on 10% of the data). The results show that NN performance is not hampered by training on a smaller portion of data.

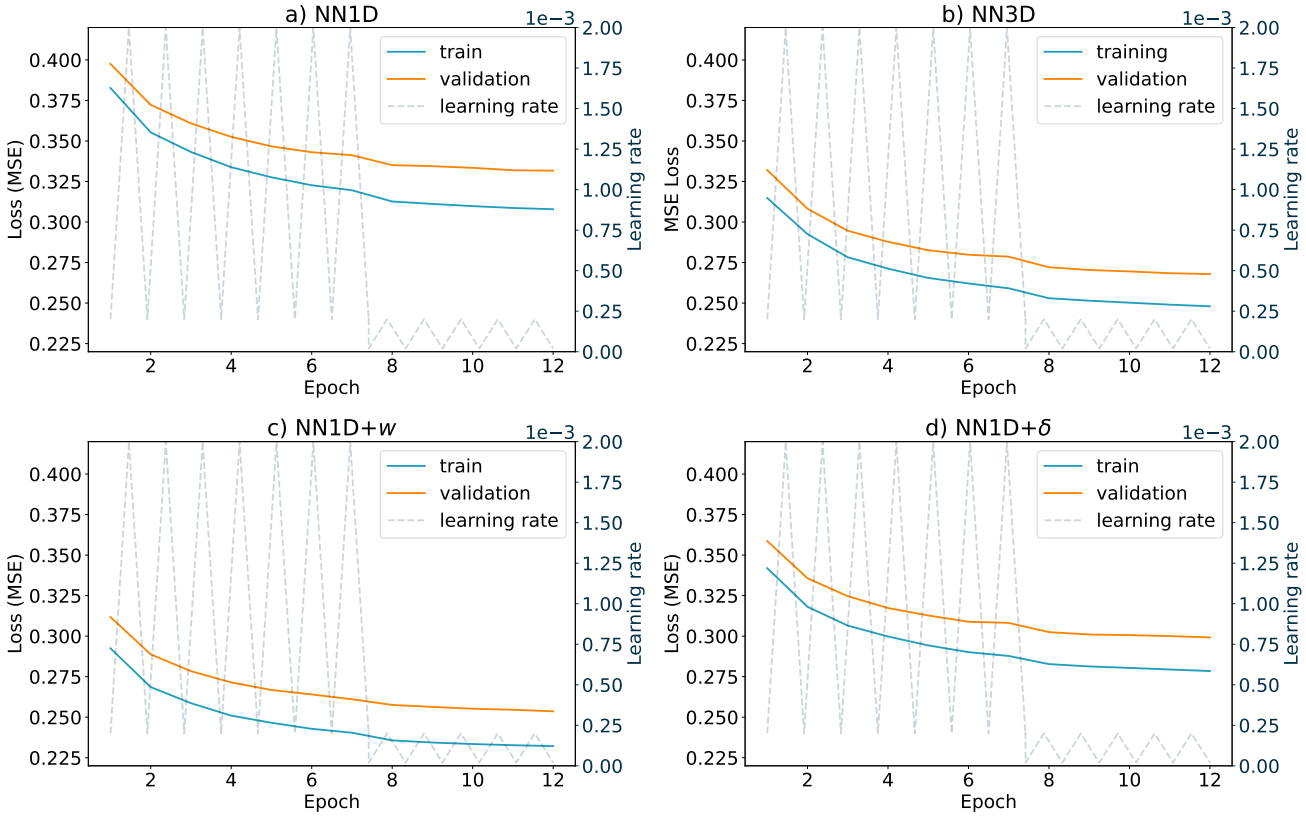

Figure S2. Loss (mean-squared error) for the training (blue) and validation (orange) sets at each epoch for the NNs used in Figure 2: **(a)** the single-column parameterization (NN1D), **(b)** the non-local parameterization (NN3D), **(c)** the single-column parameterization with local w as an additional feature (NN1D+w), and **(d)** the single-column parameterization with local horizontal wind divergence δ as an additional feature (NN1D+δ). Losses are plotted for the average of the batch at the end of each epoch. Learning rate is plotted in gray for each batch. The learning rate scheduler cycles through minimum and maximum learning rates once over each epoch. The maximum and minimum learning rates are reduced by a factor of 10 after training for 7 epochs.

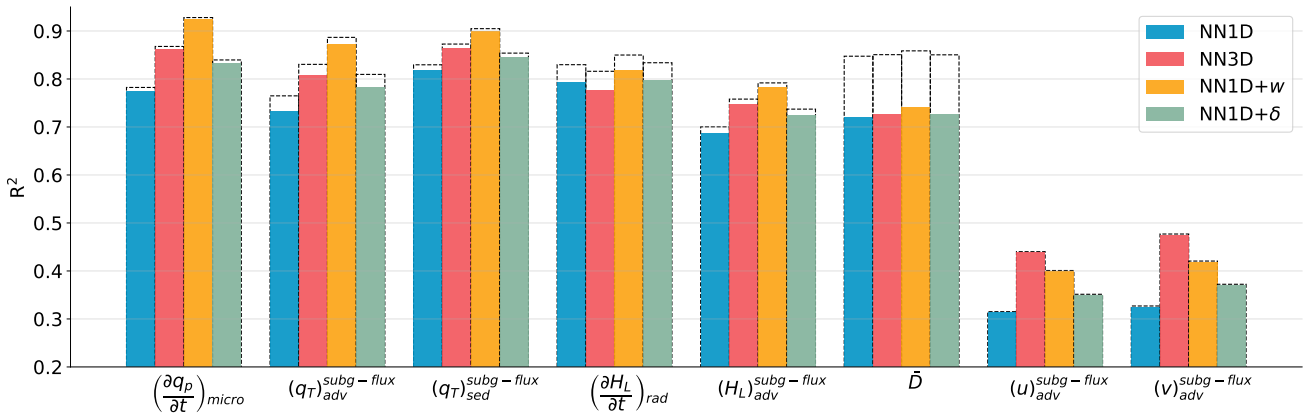

Figure S3. Similar to Figure 2, but also showing R^2 values when the mean at each vertical level is first removed. The dashed bars are the same as the bars in Figure 2, and the colored bars indicate R^2 values when the mean at each vertical level is first removed.

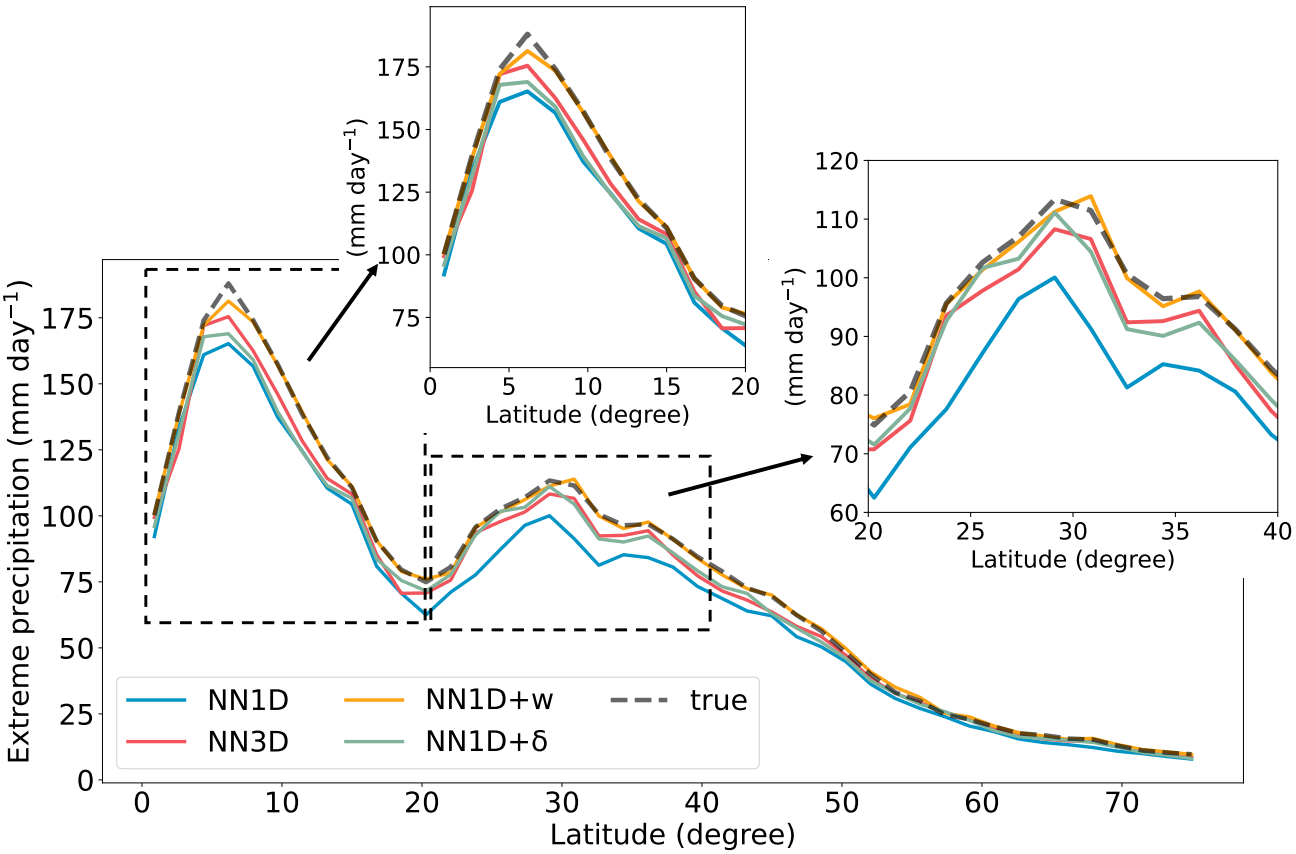

Figure S4. The 99.9$^{\text{th}}$ percentile of the instantaneous surface precipitation rate coarse-grained from the high-resolution simulation (gray dashed line), and from NNs with different input feature combinations: a single-column parameterization (NN1D; blue), a single-column parameterization with an additional 1D vertical velocity input (NN1D+w; yellow), a single-column parameterization with an additional 1D wind divergence input (NN1D+δ; green), and a non-local parameterization (NN3D; red). NN1D underestimates the extreme precipitation by up to 20% at the mid-latitudes, while NN3D substantially improves the extreme precipitation in this region, and NN3D+w improves the extreme precipitation at all latitudes.

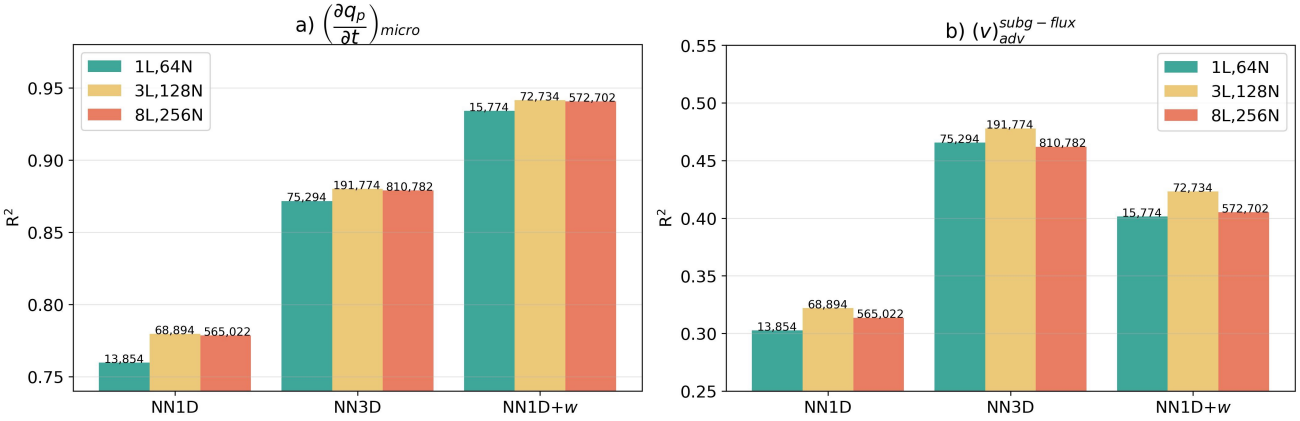

Figure S5. The coefficient of determination (R^2) of the **(a)** tendency of the precipitating water mixing ratio due to microphysical processes ($(\partial q_p/\partial t)_{\text{micro}}$) and **(b)** meridional momentum flux due to vertical advection ($(v)_{\text{adv}}^{\text{subg}-\text{flux}}$), from single-column parameterizations (NN1D), non-local parameterizations (NN3D) and single-column parameterizations with w as additional input (NN1D+w). Different architectures are considered: 1 hidden layer and 64 neurons per layer (1L,64N); 3 hidden layers and 128 neurons per layer (3L,128N); 8 hidden layers and 256 neurons per layer (8L,256N). The number of trainable parameters associated with each neural network architecture is shown on top of each bar. NN3D and NN1D+w with lower complexity (1L,64N) can significantly outperform NN1D with higher complexity (8L,256N), while using 87–97 % fewer trainable parameters.

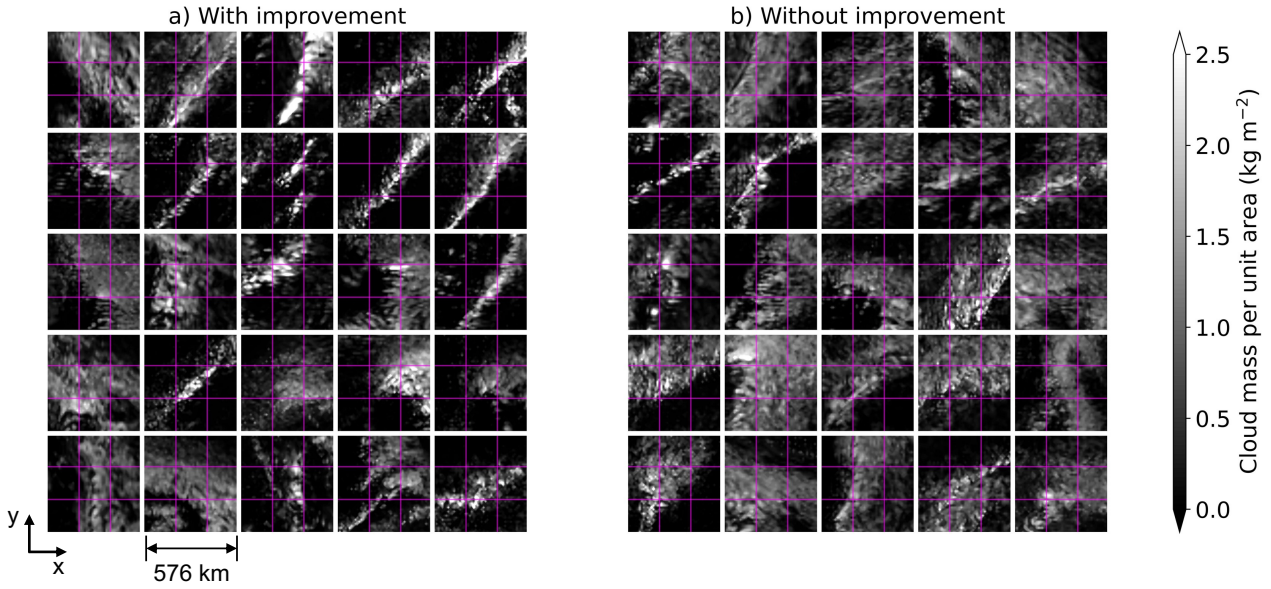

Figure S6. Similar to Figure 4, but based on improvements in predicting $(v)_{\mathrm{adv}}^{\mathrm{subg-flux}}$.

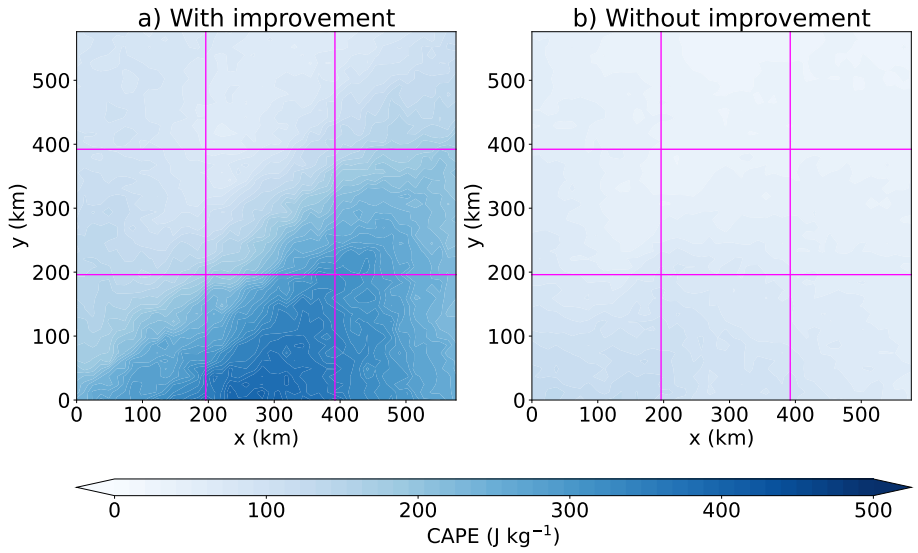

Figure S7. Similar to Figure 5, but based on the improvement in predicting $(v)_{\mathrm{adv}}^{\mathrm{subg-flux}}$.

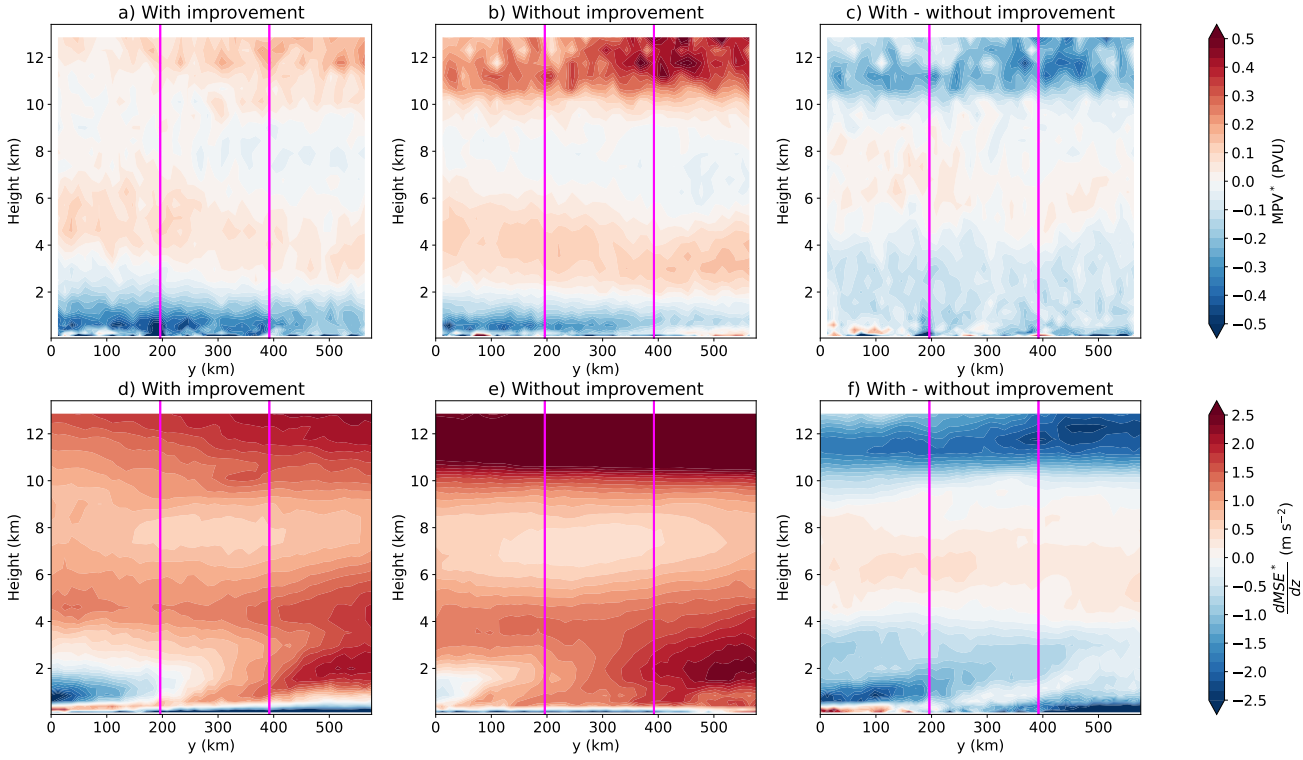

Figure S8. Different measures of instability based on the cases with largest/little improvement when predicting $(\partial q_p/\partial t)_{\text{micro}}$ using non-local parameterization compared to local parameterization. **(a-c)** The saturation potential vorticity (which is a measure of conditional symmetric instability) and **(d-f)** the vertical derivative of the moist static energy (which is a measure of conditional instability to upright convection) as a function of height and meridional distance. These measures of instability are calculated from the high resolution data for the 300 mid-latitude cases that **(a,d)** have the largest improvements when the non-local parameterization (NN3D) is used, **(b,e)** have little improvement when NN3D is used, and **(c,f)** for the difference between the cases with largest and little improvement. See Section 3.2 for details of how the groups of cases are chosen. Averages are taken in longitude over three coarse columns and over the cases in question. Magenta lines show the edges of the coarse columns. An increased band of static stability is evident near 4 – 5 km due to the melting level. Both instability measures suggest that the atmosphere is overall less stable in cases with largest improvements, especially below 4 km.

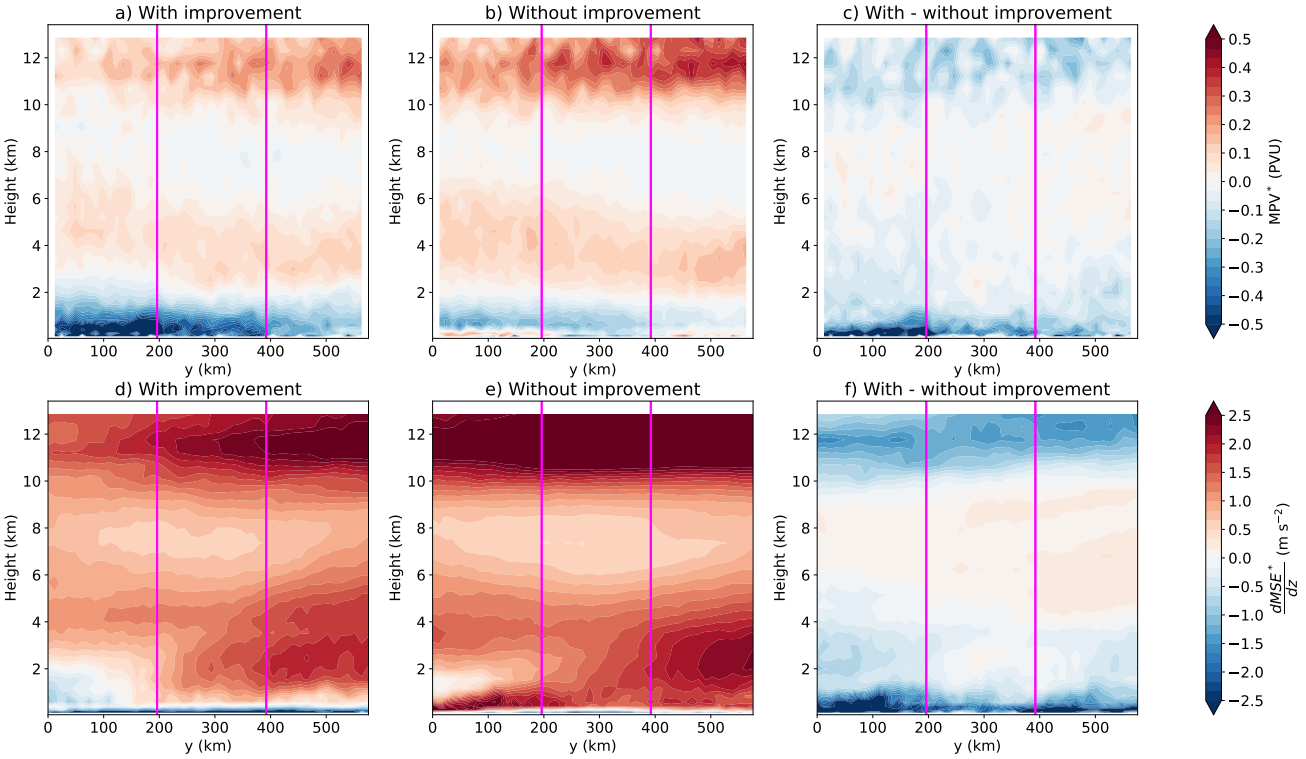

Figure S9. Similar to Figure S8, but based on the improvement in predicting $(v)_{\mathrm{adv}}^{\mathrm{subg-flux}}$.

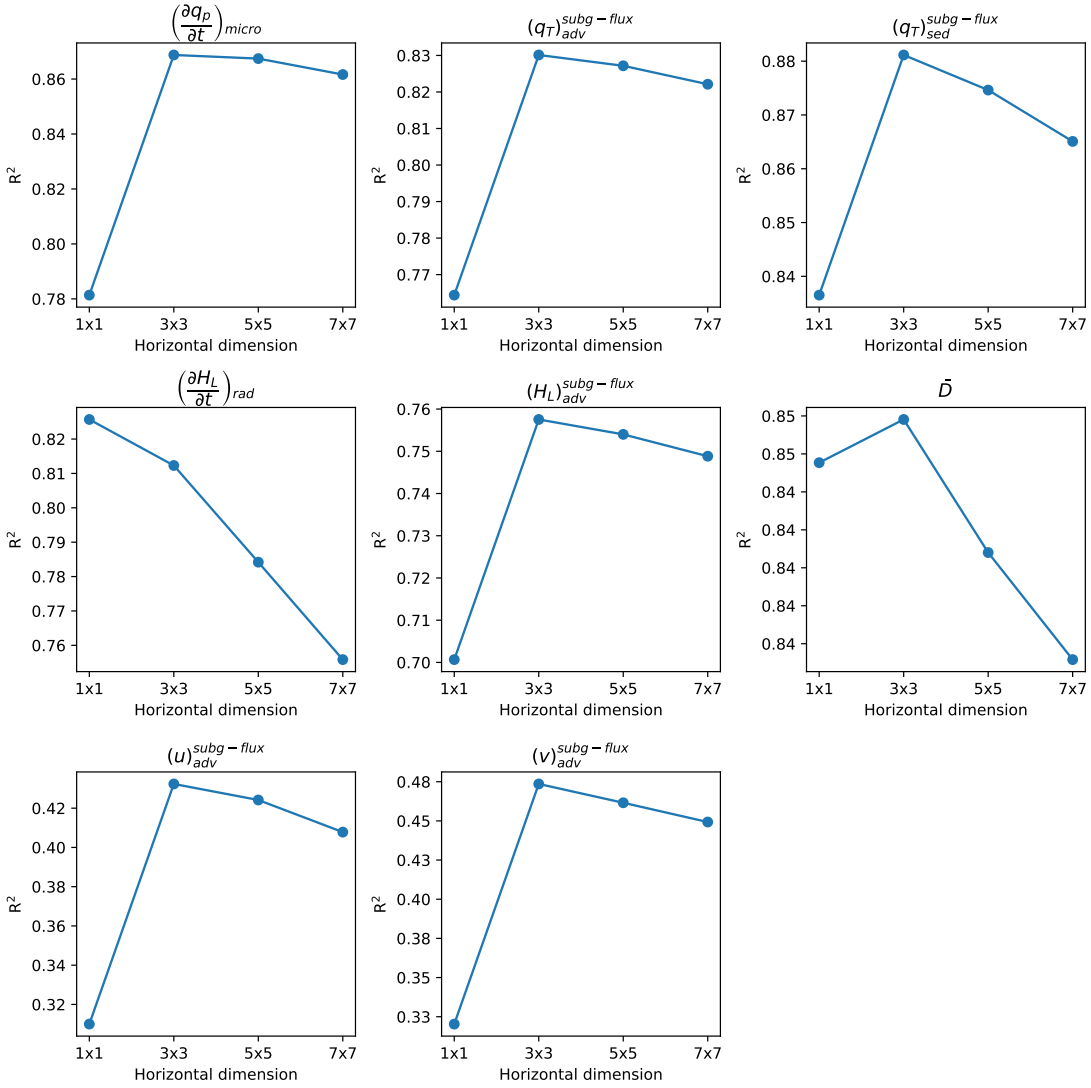

Figure S10. Global R^2 values for NNs using input features of T, q_T, u, v versus the number of non-local grid boxes used as inputs, ranging from 1×1 (local parameterization) up to 7×7 non-local grid boxes. To be consistent in the number of training/testing samples between NNs with different level of non-locality, we exclude the three outermost columns of data in each snapshot. Therefore, each snapshot has 2,520 samples instead of 2,752 samples as was used for results in other figures.

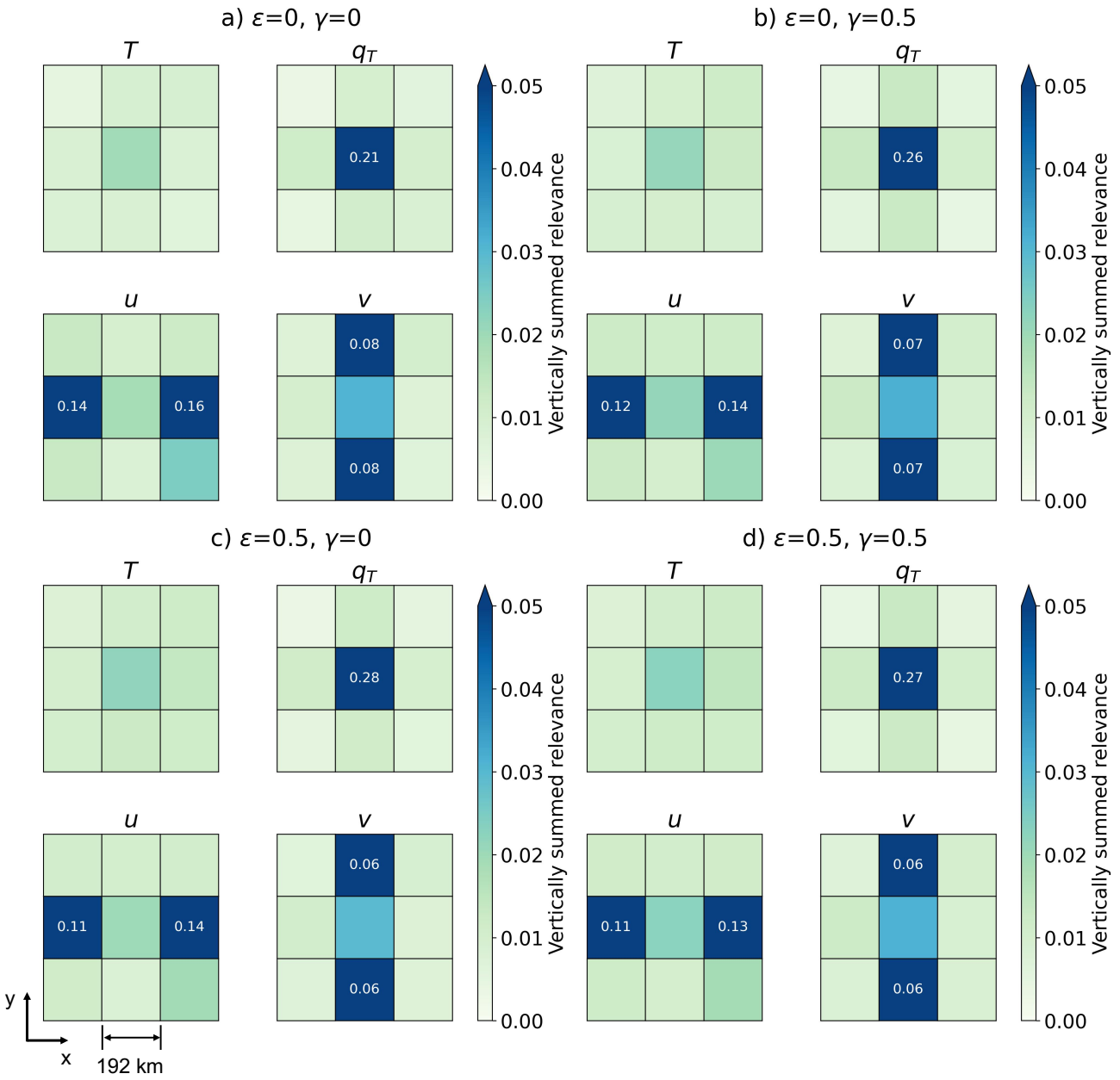

Figure S11. Similar to Figure 6a, but for different ϵ and γ values used to calculate relevance. Equation 1 in the supplement shows how ϵ and γ enter into the calculation of the relevance. Larger ϵ values tend to reduce noise in relevance, and larger γ values favor positive contributions from neurons. This figure shows that the exact relevance score can be slightly different using various parameters, but the overall pattern remains the same.

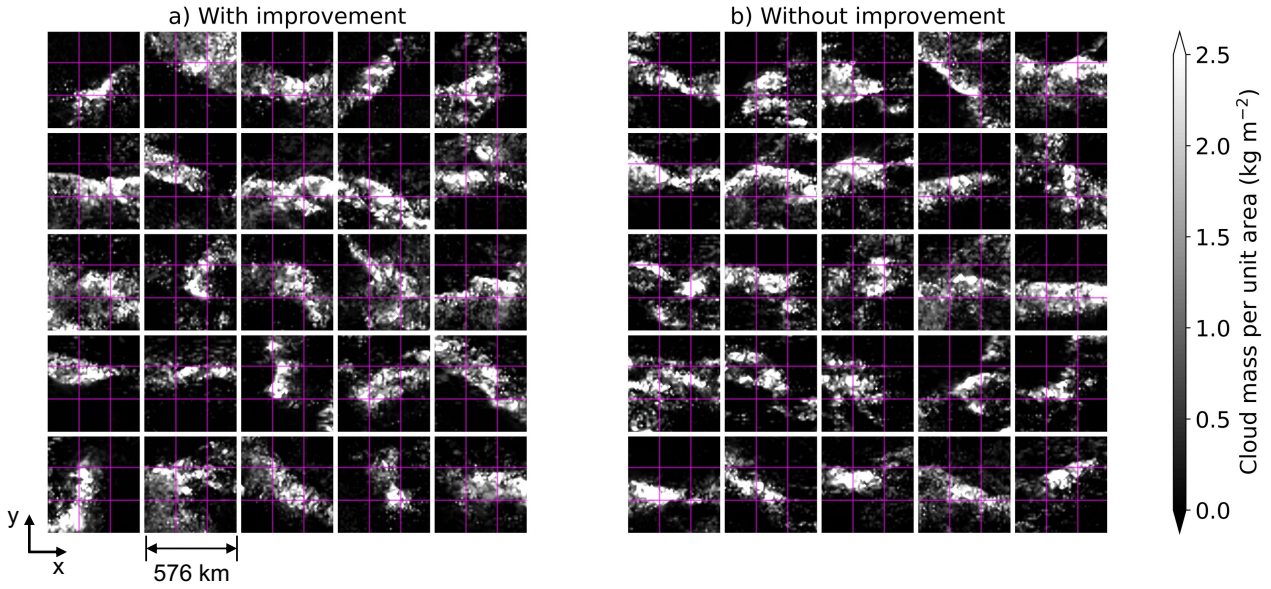

Figure S12. Similar to Figure 4 but for the tropical cases when predicting $(\partial q_p/\partial t)_{\text{micro}}$. Unlike the mid-latitude cases, cloud shapes in the cases with and without improvement do not obviously differ.

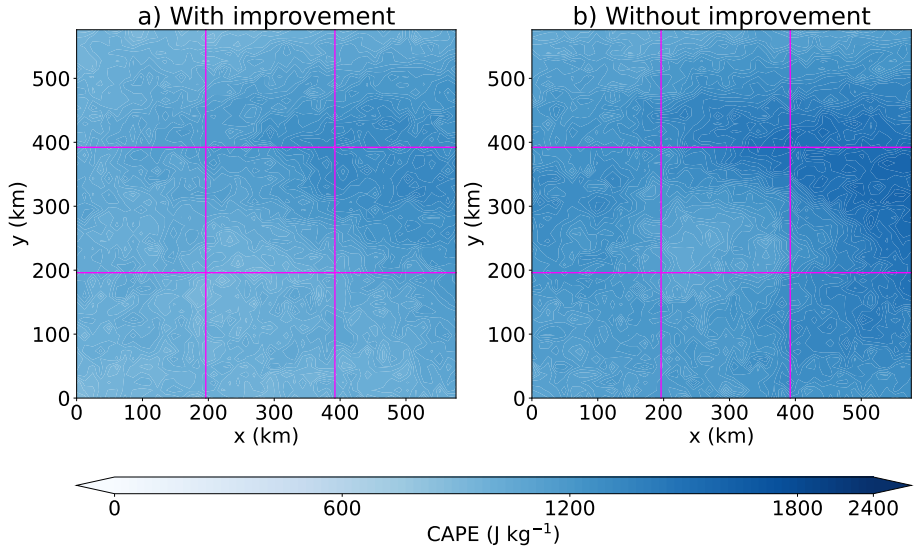

Figure S13. Similar to Figure 5, but for tropical cases when predicting $(\partial q_p/\partial t)_{\text{micro}}$. Cases with improvement have slightly lower CAPE than cases without improvement, suggesting the atmosphere is more stable to convection, which is opposite compared to the mid-latitudes.

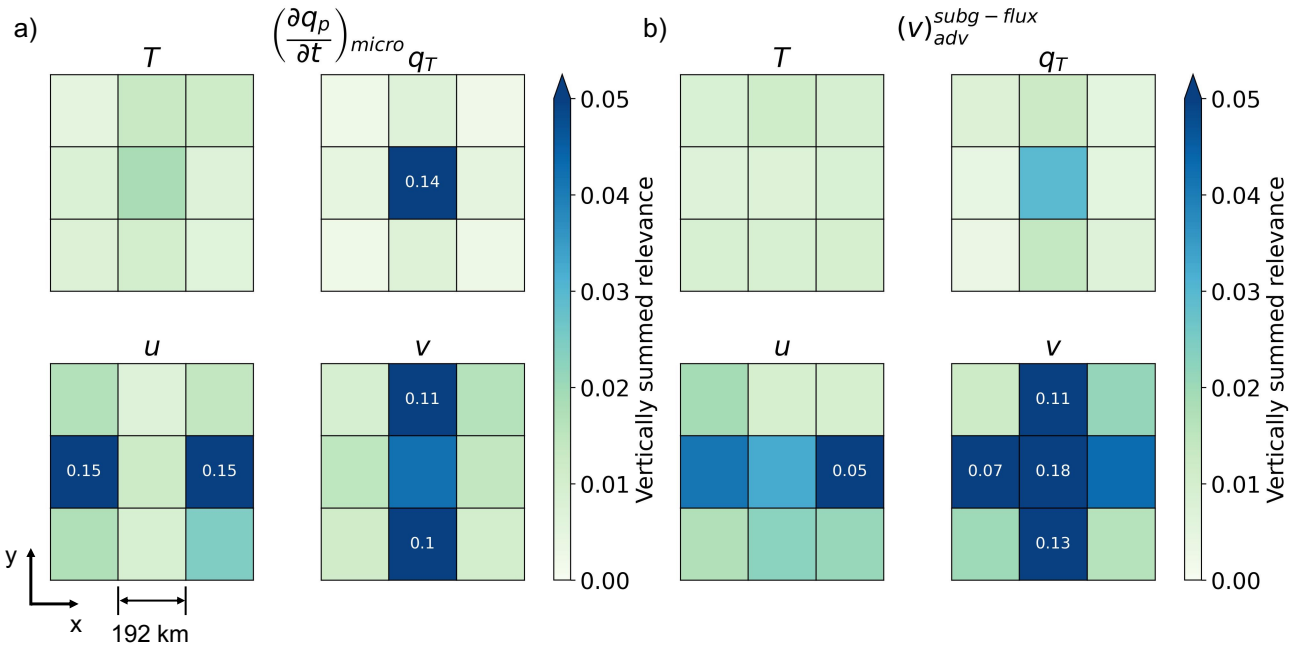

Figure S14. Similar to Figure 6, but for tropical cases. The exact relevance values are different, but the overall patterns are similar, suggesting that similar information is being used.

References

Korty, R. L., & Schneider, T. (2007). A climatology of the tropospheric thermal stratification using saturation potential vorticity. *Journal of Climate*, *20*(24), 5977–5991.

Montavon, G., Binder, A., Lapuschkin, S., Samek, W., & Müller, K.-R. (2019). Layer-wise relevance propagation: An overview. In W. Samek, G. Montavon, A. Vedaldi, L. K. Hansen, & K.-R. Müller (Eds.), *Explainable AI: Interpreting, explaining and visualizing deep learning* (pp. 193–209). Cham: Springer International Publishing.

Montavon, G., Lapuschkin, S., Binder, A., Samek, W., & Müller, K.-R. (2017, May). Explaining nonlinear classification decisions with deep Taylor decomposition. *Pattern Recognition*, *65*, 211–222.

Muller, C. J., O'Gorman, P. A., & Back, L. E. (2011). Intensification of precipitation extremes with warming in a cloud-resolving model. *Journal of Climate*, *24*(11), 2784 - 2800.

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).