
An Improved Neural Network for Shortwave Radiative Feedback Quantification

Ezra Huang

School of Computer Science
McGill University
Montréal, QC, H3A 0G4, CA
ezra.huang@mail.mcgill.ca

Han Huang

Department of Atmospheric and Oceanic Science
McGill University
Montréal, QC, H3A 0G4, CA
han.huang@mcgill.ca

Yi Huang

Department of Atmospheric and Oceanic Science
McGill University
Montréal, QC, H3A 0G4, CA
yi.huang@mcgill.ca

Abstract

The radiative kernel method has become a key tool in quantifying individual climate feedbacks in the forcing-feedback framework of climate change. This method, however, makes linearity and independence assumptions about climate feedbacks that have been observed to have been violated, especially in the Arctic, where climate change involves feedback variable changes of large magnitudes and complex patterns. While neural-network-based methods of feedback quantification, which avoid these assumptions, have been proposed, existing approaches require inputs irrelevant to radiative transfer theory or make inaccurate predictions of radiative sensitivity. In this work, we improve existing neural network methods with data augmentation and Sobolev training. Our results show that the new neural network is able to make accurate predictions of top-of-atmosphere net shortwave radiative sensitivity, and achieves better radiative closure than the kernel method in a case study of analyzing shortwave radiative feedbacks in the Arctic.

1 Introduction

In climate science, the climate feedback parameter (λ) determines the equilibrium-state ($\Delta R = 0$) global mean surface temperature change (ΔT_s) in response to a radiative forcing, F , through the relationship $\Delta R = F + \lambda \Delta T_s$ [1]. Despite its importance to understanding global warming, current estimations of λ still come with large uncertainties, with different general circulation models (GCMs) estimating values between -1.5 and $-0.6 \text{ W m}^{-2} \text{ K}^{-1}$ [2]. To understand the causes of this uncertainty, λ is often decomposed as a sum of univariate feedbacks corresponding to different climate processes [2, 3], each of which is typically quantified by the partial radiation flux change ΔR_i induced by a feedback variable Δx_i , normalized by change the in surface temperature ΔT_s [2].

While these partial radiation flux changes can be directly computed with the partial radiative perturbation (PRP) method [4], its computational cost has led to widespread use of the cheaper radiative kernel method [5, 6]. This method, however, assumes that individual radiative feedbacks are both linear and independent of one another. Nonlinear effects, however, are known to cause quantitative biases [7], which recent works have attempted to address; for example, [8] documented nonlinearity in the shortwave (SW) albedo feedback and nonlinear coupling between the SW cloud and albedo feedbacks, and [9] found nonlinear coupling between longwave (LW) cloud and water vapour feedbacks.

Neural networks (NNs), which can avoid these assumptions, have been proposed as an alternative method to quantify radiative feedbacks, but development to date has not offered a practical tool that can replace the kernel method. The model of [10] requires inputs (latitude, longitude) outside of radiative transfer theory, while the model of [11] struggles to quantify albedo radiative sensitivity.

In this work, we make two modifications to improve these models. First, we add clear-sky samples to the training set, since clear-sky cases are often used as a diagnostic for climate models, and to encourage the model to learn simpler clear-sky patterns as a base for further generalization. Second, we use Sobolev learning [12] to promote learning of radiative sensitivities that are key to the quantification of univariate feedbacks (i.e., radiation change attributable to individual feedback variables), which is missed in previous datasets that only train on multivariate changes of feedback variables.

2 Materials and Methods

2.1 Datasets

The main dataset we use is ERA5 monthly-mean data on single levels [13] regridded to a 1° resolution. The years 2011 to 2014 are used for training, and 2015 is used for validation. For training and validation, we use the months of March, June, September, and December, which are chosen to capture a full range of climatological conditions. Each ERA5 point after regridding yields both an all-sky and clear-sky sample, so the training dataset forms a 4-dimensional (latitude, longitude, date, sky) array of shape $180 \times 360 \times 16 \times 2$. For the clear-sky sample, the all-sky inputs listed in Table 1 are set to zero and the ERA5 target variable is tsr; otherwise all inputs are used and the target variable is tsr. We also use the additional derived input features of effective cloud optical depth (ecod) and effective cloud optical depth multiplied by forecast albedo (ecod $\times$ fal), motivated by their presence in radiative transfer theory equations [14]. Details of their calculation is available in Appendix A. Before training, all inputs and outputs are linearly rescaled to the range of $[-1, 1]$.

2.2 Model Architecture

Our model has the same structure as [11], which is a feedforward neural network with 12 radiative-transfer-relevant inputs (Table 1), one 11-dimensional hidden layer, and a 1-dimensional output of TOA net SW radiation, using hyperbolic tangent activation functions for the hidden layer.

To train the model, we use Sobolev training, which has been shown to improve data-efficiency and generalization ability by constructing the loss as a metric in Sobolev space [12]; in other words, the

Table 1: Model inputs and output for clear-sky and all-sky conditions. Variables with no asterisk in the ERA5 column are direct model inputs, otherwise they are derived from the indicated fields.

	Variable	ERA5 source
Inputs (clear-sky)	TOA incident SW radiation	tsr
	forecast albedo, surface pressure	fal, sp
	total column {water vapour, ozone}	tcwv, tco3
Inputs (all-sky, additional)	{high, medium, low} cloud cover	hcc, mcc, lcc
	{ice, liquid} cloud water content	tcw, tclw
	effective cloud optical depth	tcc*, tciw*, tclw*
	effective cloud optical depth $\times$ forecast albedo	fal*, tcc*, tciw*, tclw*
Output	TOA net SW radiation (clear-sky)	tsr(c)

prediction error depends on both the model output and its derivatives. We use first-order Sobolev training with albedo-kernel gradient data, giving a per-sample loss of

$$\mathcal{L}(x) = (R(x) - N(x))^2 + \lambda (\partial_{\text{albedo}} R(x) - \partial_{\text{albedo}} N(x))^2 \quad (1)$$

where R is the ERA5 target field, N is the neural network, $\partial_{\text{albedo}} R(x)$ is a PRP-approximation of TOA SW albedo sensitivity, and $\lambda = 0.5$ weights the Sobolev loss component. The PRP-approximation of $\partial_{\text{albedo}} R$ is derived from data generated for computing the ERA5 radiative kernels in [15]: RRTM [16] is run on both a baseline ERA5 profile and the profile with a small (1%) surface albedo perturbation, and the resulting radiation difference is divided by the perturbation to yield the approximate gradient.

The model is trained in PyTorch [17] with the Adam optimizer with default hyperparameters, with a learning rate of 0.001, batch size of 256, and stopping once validation error did not decrease for 10 epochs.

Similar to [8], we can use the neural network to quantify univariate feedbacks as $\Delta R_i := N(x + \Delta x_i) - N(x)$, multivariate feedbacks as $\Delta R_{ij} := N(x + \Delta x_i + \Delta x_j) - N(x) - \Delta R_i - \Delta R_j$, and the net feedback (including all cross terms) as $\Delta R_{\text{allcross}} := N(x + \Delta x) - N(x)$, where x is a climate state and $i, j \in \{a, c, q\}$ index the albedo, cloud, and water vapour components of the climate state.

3 Results

First, we test different models' ability to estimate TOA net SW radiative sensitivity to surface albedo. For all months of 2015 excluding March, June, September, and December, we estimate clear-sky and all-sky radiative sensitivity to forecast albedo with the neural network and compare it with PRP-computed values by calculating the global (area-weighted) MBE and RMSE, finally taking a mean over the months to get a scalar value. As we see in Table 2, when compared with a plain NN baseline, the all/clear-sky MBE (resp. RMSE) values decrease from -0.493 (0.708) / -0.302 (0.511) to 0.010 (0.096) / 0.010 (0.051) $\text{W m}^{-2} \%^{-1}$, confirming that our modifications significantly improve the model's estimation of radiative sensitivity to forecast albedo at a point after considering that the global mean over the same dates is $1.73 \text{ W m}^{-2} \%^{-1}$.

Next, we test the new model's ability to quantify radiative feedbacks using a case study of interannual variability. We use it to estimate TOA SW radiative feedbacks in the Arctic ($>65^\circ\text{N}$) between September 1992 and September 2020, then compare the results against the kernel method, using the ERA5-derived kernels of [15] and the cloud radiative effect (CRE) method of [5] for cloud effects. September 1992 is a typical year, while September 2020 is a record-low year for Arctic sea ice extent, causing large surface albedo changes. When comparing the NN against the kernel method, the MBE (resp. RMSE) decreases to 0.28 (3.45) from 1.06 (5.95) W m^{-2} (Fig. 2a,c) which is a considerable improvement considering that the ERA5 regionally-averaged mean change is 5.95 W m^{-2} (Fig. 1a).

Analyzing the spacial distribution of the NN and kernel predictions, this improvement can be attributed to NN's quantification of nonlinear albedo and joint albedo-cloud feedback terms. The kernel method assumes a linear approximation of the concave dependence of ΔR on surface albedo ([14, 8]), which, in regions of high surface albedo change (Fig. 2g) leads to it overestimating the albedo feedback (Fig. 1b) when compared to the neural network (Fig. 1c), demonstrating its ability to capture the nonlinear albedo effect. Moreover, the majority of the discrepancy between the NN's estimated

Table 2: Global mean bias error (MBE) and root-mean-square error (RMSE) for different models' predictions of TOA SW radiative sensitivity to surface albedo in clear-sky and all-sky conditions. Values are computed for each month of 2015, excluding March, June, September, and December, and averaged over all included dates.

Model	MBE ($\text{W m}^{-2} \%^{-1}$)		RMSE ($\text{W m}^{-2} \%^{-1}$)	
	All-sky	Clear-sky	All-sky	Clear-sky
NN	-0.493	-0.302	0.708	0.511
NN+clearsky	-0.465	-0.184	0.716	0.436
NN+clearsky+Sob.	0.010	0.010	0.096	0.051

net feedback (Fig. 2a) and the sum of univariate feedbacks (Fig. 2b) can be explained by the NN's prediction of the joint albedo-cloud feedback (Fig. 2d), which is known to dominate the Arctic nonlinear shortwave coupling effects [8]. Indeed, in regions such as the East Siberian Sea (albedo ↓, cloud ↑), the univariate sum overestimates the net feedback, while in regions such as Iceland and the Denmark Strait (albedo ↓, cloud ↑), it underestimates the net feedback. This is consistent with radiative transfer theory, which predicts that increased cloud optical depth reduces atmospheric transmittance, reducing radiative sensitivity to surface albedo [8].

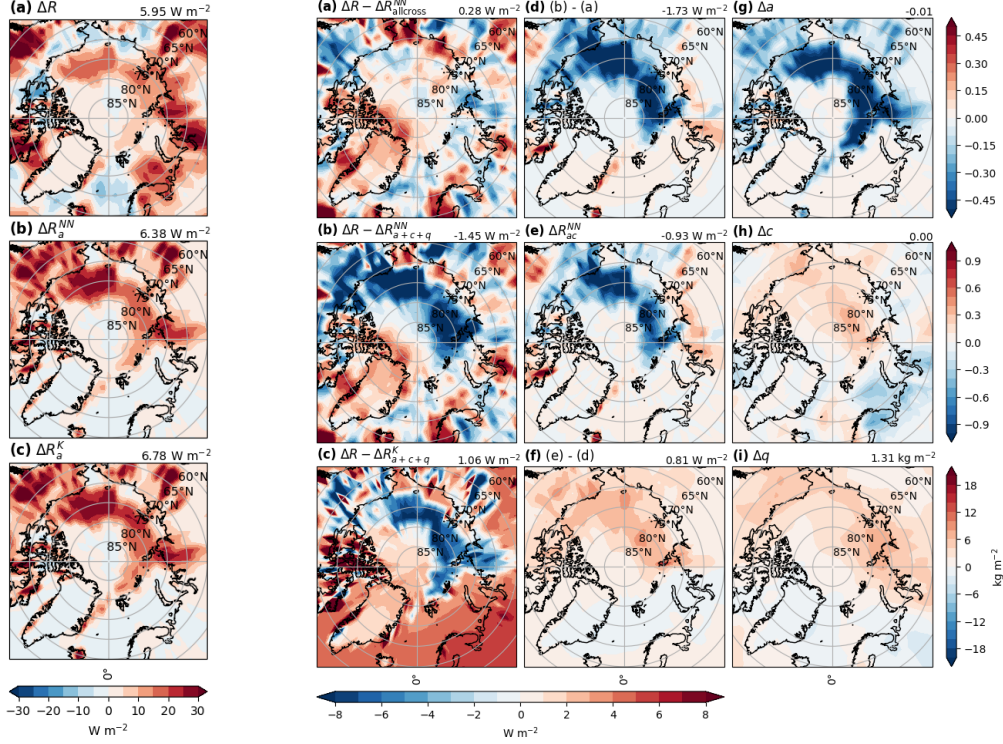


Figure 1: (a) All-sky ERA5 Sep. Arctic TOA net SW flux change (from 1992 to 2020) compared with the (b) NN-computed univariate albedo feedback and (c) kernel-computed univariate albedo feedback. Numbers in the top right of each plot are area-weighted means.

Figure 2: (a-c) Residuals of net radiative feedback estimates against the ERA5 value (ΔR) for the NN using all cross terms ($\Delta R^{\text{NN}}_{\text{allcross}}$), the NN using a univariate sum of feedbacks ($\Delta R^{\text{NN}}_{a+c+q}$), and the kernel method (R^K_{a+c+q}). (d) Difference between the allcross and univariate sum NN feedback estimates. (e) NN-estimate of the joint albedo-cloud feedback. (f) Difference between the allcross and univariate feedback sum after removing the joint albedo-cloud contribution. (g-i) ERA5 changes in forecast albedo (Δa), effective cloud optical depth (Δc), and total column water vapour (Δq). Numbers in the top right of each plot are area-weighted means. Plots (a-f) share the bottom colorbar, while plots (g-i) each have their own vertical colourbar.

4 Discussion and Conclusion

In this work, we developed a new neural network to predict TOA SW net radiation using the ERA5 dataset and community-generated PRP data. First, we showed that this neural network is better at predicting radiative sensitivity to surface albedo than past neural networks. Then, using a case study of quantifying radiative feedbacks in interannual variability, we showed that our neural network achieves better radiative closure than radiative kernel method, mainly due to the neural network's ability to capture nonlinearities in radiative feedbacks.

In future work, it will be useful to compare the NN-computed nonlinear and joint feedback estimates with RRTM experiments for more quantitative validation of the model. We also aim to incorporate more data from community PRP experiments for model training, and to expand to additionally predict

longwave and surface feedbacks. Ultimately, this work shows that neural networks show promise in enabling more accurate estimations of radiative feedbacks, allowing for more accurate estimations of climate feedbacks to reduce uncertainties in ECS estimates and better predict the magnitude of global warming.

Acknowledgments and Disclosure of Funding

E. H. acknowledges funding from the Undergraduate Student Research Award (no. 616310-2026) program of the Natural Sciences and Engineering Research Council of Canada (nserc-crsng.canada.ca). This research was enabled in part by support provided by Calcul Québec (www.calculquebec.ca) and the Digital Research Alliance of Canada (alliancecan.ca).

References

- [1] Yi Huang. On the longwave climate feedbacks. *Journal of Climate*, 26(19):7603–7610, October 2013.
- [2] Mark D. Zelinka, Timothy A. Myers, Daniel T. McCoy, Stephen Po-Chedley, Peter M. Caldwell, Paulo Ceppi, Stephen A. Klein, and Karl E. Taylor. Causes of higher climate sensitivity in cmip6 models. *Geophysical Research Letters*, 47(1):e2019GL085782, 2020.
- [3] S. C. Sherwood, M. J. Webb, J. D. Annan, K. C. Armour, P. M. Forster, J. C. Hargreaves, G. Hegerl, S. A. Klein, K. D. Marvel, E. J. Rohling, M. Watanabe, T. Andrews, P. Braconnot, C. S. Bretherton, G. L. Foster, Z. Hausfather, A. S. von der Heydt, R. Knutti, T. Mauritsen, J. R. Norris, C. Proistosescu, M. Rugenstein, G. A. Schmidt, K. B. Tokarska, and M. D. Zelinka. An assessment of earth’s climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, 58(4):e2019RG000678, 2020.
- [4] R. T. Wetherald and S. Manabe. Cloud feedback processes in a general circulation model. *Journal of the Atmospheric Sciences*, 45(8):1397–1416, April 1988.
- [5] Karen M. Shell, Jeffrey T. Kiehl, and Christine A. Shields. Using the radiative kernel technique to calculate climate feedbacks in ncar’s community atmospheric model. *Journal of Climate*, 21(10):2269–2282, May 2008.
- [6] Brian J. Soden, Isaac M. Held, Robert Colman, Karen M. Shell, Jeffrey T. Kiehl, and Christine A. Shields. Quantifying climate feedbacks using radiative kernels. *Journal of Climate*, 21(14):3504–3520, July 2008.
- [7] R. A. Colman, S. B. Power, and B. J. McAvaney. Non-linear climate feedback analysis in an atmospheric general circulation model. *Climate Dynamics*, 13(10):717–731, October 1997.
- [8] Yi Huang, Han Huang, and Aliia Shakirova. The nonlinear radiative feedback effects in the arctic warming. *Frontiers in Earth Science*, 9, August 2021.
- [9] Han Huang and Yi Huang. Nonlinear coupling between longwave radiative climate feedbacks. *Journal of Geophysical Research: Atmospheres*, 126(8):e2020JD033995, 2021.
- [10] Tingting Zhu, Yi Huang, and Haikun Wei. Estimating climate feedbacks using a neural network. *Journal of Geophysical Research: Atmospheres*, 124(6):3246–3258, 2019.
- [11] Diana Laura Diaz Garcia and Yi Huang. Quantification of shortwave surface albedo feedback using a neural network approach. *Atmosphere*, 15(2):150, February 2024.
- [12] Wojciech M. Czarnecki, Simon Osindero, Max Jaderberg, Grzegorz Swirszcz, and Razvan Pascanu. Sobolev training for neural networks. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.
- [13] Hans Hersbach, Bill Bell, Paul Berrisford, Shoji Hirahara, András Horányi, Joaquín Muñoz-Sabater, Julien Nicolas, Carole Peubey, Raluca Radu, Dinand Schepers, Adrian Simmons, Cornel Soci, Saleh Abdalla, Xavier Abellan, Gianpaolo Balsamo, Peter Bechtold, Gionata Biavati, Jean Bidlot, Massimo Bonavita, Giovanna De Chiara, Per Dahlgren, Dick Dee, Michail Diamantakis, Rossana Dragani, Johannes Flemming, Richard Forbes, Manuel Fuentes, Alan Geer, Leo Haimberger, Sean Healy, Robin J. Hogan, Elías Hólm, Marta Janisková, Sarah Keeley, Patrick Laloyaux, Philippe Lopez, Cristina Lupu, Gabor Radnoti, Patricia de Rosnay, Iryna Rozum, Freja Vamborg, Sebastien Villaume, and Jean-Noël Thépaut. The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730):1999–2049, 2020.

- [14] Graeme L. Stephens, Denis O'Brien, Peter J. Webster, Peter Pilewski, Seiji Kato, and Jui-lin Li. The albedo of earth. *Reviews of Geophysics*, 53(1):141–163, 2015.
- [15] Han Huang and Yi Huang. Radiative sensitivity quantified by a new set of radiation flux kernels based on the ecmwf reanalysis v5 (era5). *Earth System Science Data*, 15(7):3001–3021, July 2023.
- [16] Michael J. Iacono, Jennifer S. Delamere, Eli J. Mlawer, Mark W. Shephard, Shepard A. Clough, and William D. Collins. Radiative forcing by long-lived greenhouse gases: Calculations with the aer radiative transfer models. *Journal of Geophysical Research: Atmospheres*, 113(D13), 2008.
- [17] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.
- [18] A. Slingo. A gcm parameterization for the shortwave radiative properties of water clouds. *Journal of the Atmospheric Sciences*, 46(10):1419–1427, May 1989.
- [19] Qiang Fu. An accurate parameterization of the solar radiative properties of cirrus clouds for climate models. *Journal of Climate*, 9(9):2058–2082, September 1996.

A Effective Cloud Optical Depth Calculation

To compute the effective optical cloud depth (ecod) input feature displayed in Table 1, we multiply total cloud cover (tcc) by cloud optical depth (cod), which is the sum of ice (τ_i) and liquid (τ_l) cloud paths, i.e.

$$\text{ecod} = \text{cod} \cdot (\tau_l + \tau_i) \quad (2)$$

To compute the liquid water path, we follow [18], using coefficients for the 0.57 to 0.64 μm spectral band and assuming an equivalent radius of drop size distribution of 10 μm . Effectively, the equation becomes:

$$\tau_l = \text{tclw} * (a + b/r_e) \quad (3)$$

where tclw is the ERA5 field of vertically integrated cloud liquid water content, $a = 2.831 \times 10^{-2} \text{ m}^2 \text{ g}^{-1}$, $b = 1.317 \mu\text{m m}^2 \text{ g}^{-1}$, and $r_e = 10 \mu\text{m}$.

To compute the ice water path, we follow [19], with the same spectral band and assuming a generalized effective size of 30 μm . Effectively, the equation is:

$$\tau_i = \text{tcIW} * (a_0 + a_1/D_{\text{ge}}) \quad (4)$$

where tcIW is the ERA5 field of vertically integrated cloud ice water content, $a_0 = 0.982244 \times 10^{-4} \text{ m}^2 \text{ g}^{-1}$, $a_1 = 0.250875 \times 10^1 \text{ m}^2 \mu\text{m g}^{-1}$, and D_{ge} is given by:

$$D_{\text{ge}} = \frac{8 \cdot r_e}{3(3)^{1/2}} \quad (5)$$

for $r_e = 30 \mu\text{m}$.

B TOA SW Radiation Prediction

As a sanity check, we use the neural network to predict top-of-atmosphere net shortwave radiative fluxes in both all-sky and clear-sky conditions, and compare it against the same model without clear-sky training data and a plain neural network baseline. Using all months of 1990-2020, the neural network is used to predict the monthly mean TOA shortwave radiation in clear- and all-sky conditions for each month from ERA5 climates states, and the mean bias error (MBE) and root-mean-square error (RMSE) are calculated for each (latitude, longitude) location. The values reported in Table 3 are means over all dates.

Table 3: Global mean bias error (MBE) and root-mean-square error (RMSE) of model predictions of TOA net SW radiation (tsr(c)), in clear- and all-sky conditions, averaged over each month from 1990 to 2020.

Model	MBE (W m^{-2})		RMSE (W m^{-2})	
	All-sky	Clear-sky	All-sky	Clear-sky
NN	0.250	8.774	5.192	10.621
NN+clearsky	0.204	0.712	5.116	4.708
NN+clearsky+Sob.	-0.002	0.298	5.461	5.042

C Pointwise Radiative Sensitivity Estimate

To gain an understanding of the spacial distribution of the accuracy of each neural network’s prediction of radiative sensitivity, it is useful to compare maps of their predictions against the RRTM-generated truth value. Below, we plot each prediction for the date of July 2015, which has strong hemispheric asymmetry and captures a wide range of conditions.

D Kernel sensitivity test

In this test, we validate our NN’s physical consistency by decomposing the kernel difference between two climate states using RRTM and the NN. We use August 2012 as a base state and

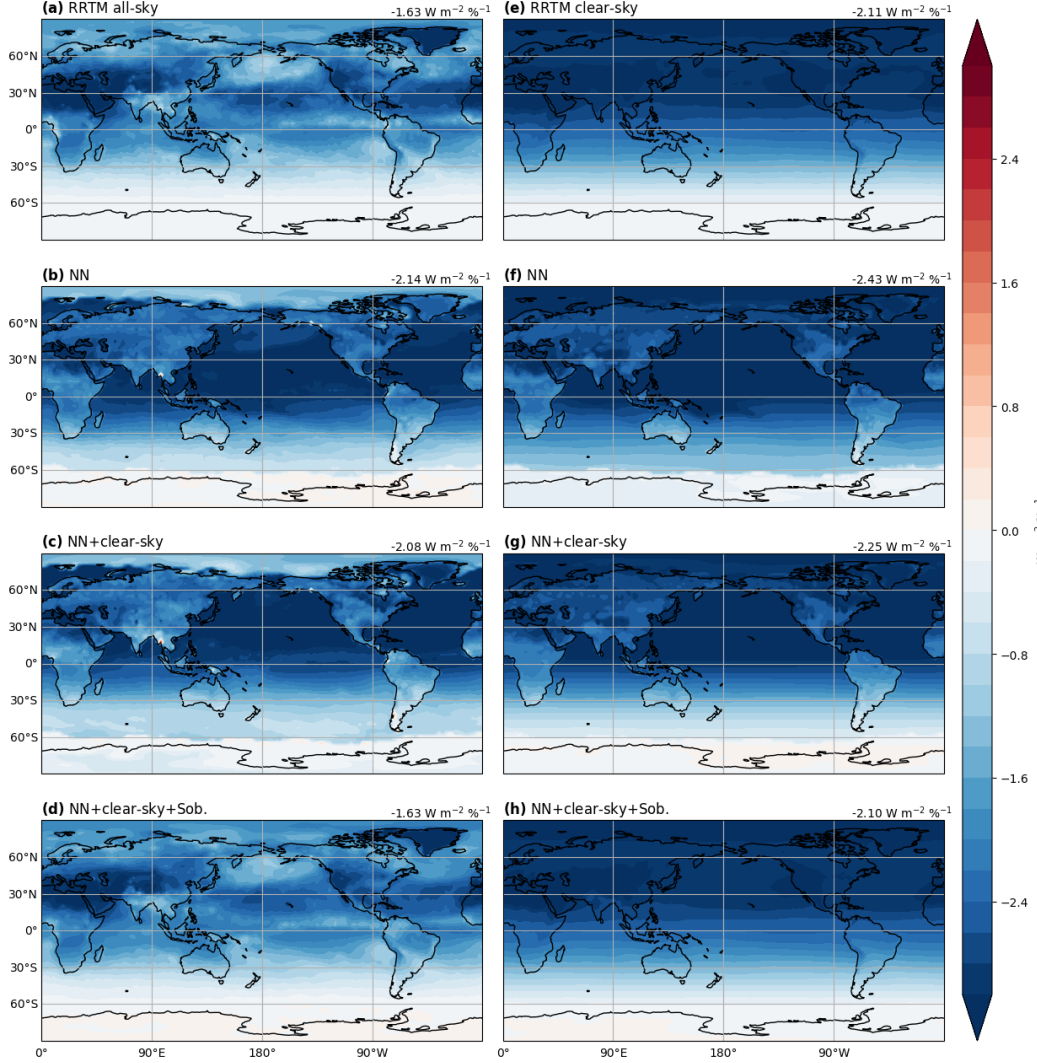


Figure 3: **(a-d)** RRTM (a), baseline NN (b), NN-with-clear-sky-training-samples (c), and our-NN (d) -quantified radiative sensitivity to forecast albedo for the month of July 2015 in all-sky conditions. **(e-h)** The same as plots on the left, but for clear-sky conditions. Numbers in the top right of each plot are area-weighted means.

August 2015 as a perturbed state (denoted by an apostrophe), chosen for their high solar insolation and large differences in surface albedo. We estimate the Arctic surface albedo sensitivities with both RRTM (K^{RRTM}) and the neural network (K^{NN}) and compare their sensitivity differences ($\Delta K = K' - K$). Then, we use the NN to decompose the change in radiative sensitivity into albedo, cloud, and water vapour contributions, and compare each contribution with the change in respective field. This decomposition is done analogously to computing univariate radiative feedbacks, where the relevant climate variables in the base state are replaced by the same variables in the perturbed state, but the predicted variable is radiative sensitivity to surface albedo instead of radiation.

In Figure 4, we observe, in general, good agreement between the kernel predictions between RRTM and the NN, with positive bias in the neural network prediction above 80° , reconfirming that the neural network makes good prediction of radiative sensitivity. Significantly, in Figure 5, we observe decreases in surface albedo are associated with more negative predicted radiative sensitivity. This matches existing literature, which predicts that radiative sensitivity to surface albedo becomes more negative as surface albedo increases, due to multiple reflection and scattering effects [8, 14]. Moreover, we observe that increases in effective cloud optical depth are associated with the less

negative predicted radiative sensitivity to surface albedo. This is also consistent with radiative transfer theory, since increased cloud cover decreases atmospheric transmittance [8, 14], decreasing radiative sensitivity to surface albedo.

Furthermore, as we can see in Figures 6 and 7 this property is unique to this neural network. Indeed, these other neural networks predict the wrong physical relationship, associating increases in surface albedo with less negative radiative sensitivity to surface albedo, and increases in effective cloud optical depth with more negative radiative sensitivity. This confirms that the Sobolev training has helped the neural network learn the correct underlying radiative physics.

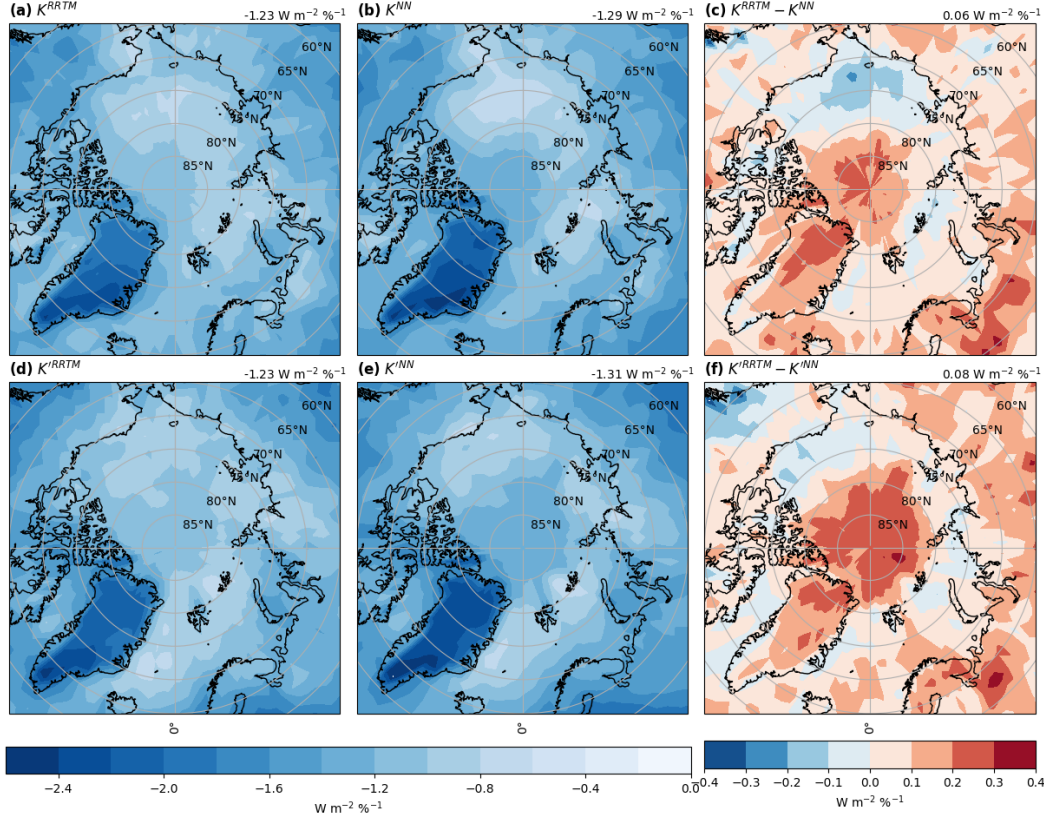


Figure 4: **(a-b)** TOA net SW radiative sensitivity to surface albedo for August 2012 (K), predicted by RRTM and the NN, respectively. **(c)** Difference between the RRTM and NN predictions of sensitivity. **(d-f)** The same for the above plots, but for August 2015 (K'). Numbers in the top right of each plot are area-weighted means.

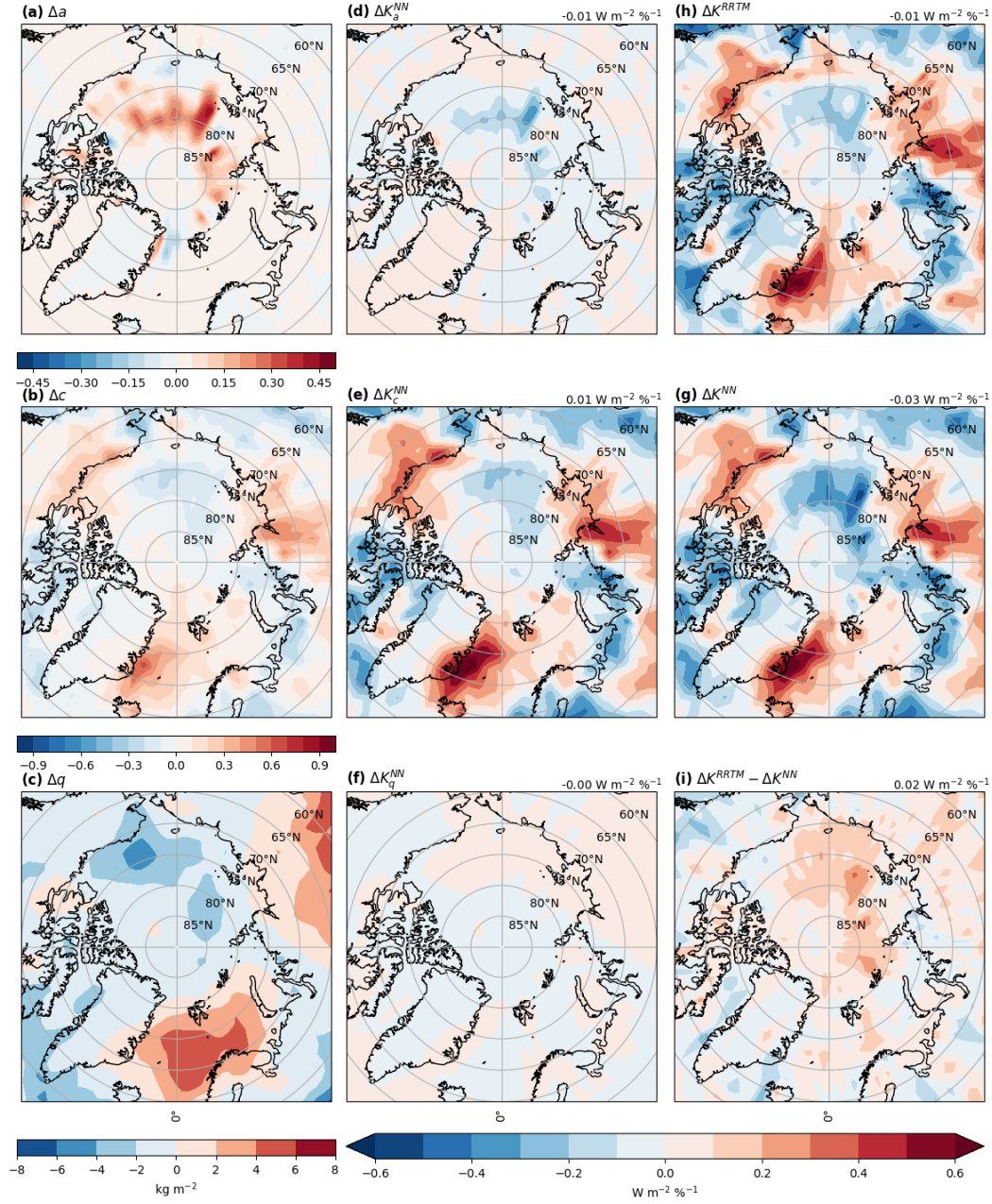


Figure 5: **(a-c)** ERA-5 derived changes in forecast albedo (Δa), effective cloud optical depth (Δc), and total column water vapour (Δq) from August 2012 to August 2015. **(d-f)** NN-computed sensitivity change from August 2012 to August 2015 attributed to the albedo, cloud, and water vapour components of the Arctic climate. **(g,h)** RRTM- and NN-estimated change in radiative sensitivity to forecast albedo from August 2012 to August 2015. **(i)** Difference between the RRTM and NN predicted change in radiative sensitivity. Numbers in the top right of each plot are area-weighted means. Plots (a-c) have their own individual colourbar, while plots (d-i) share a common colourbar.

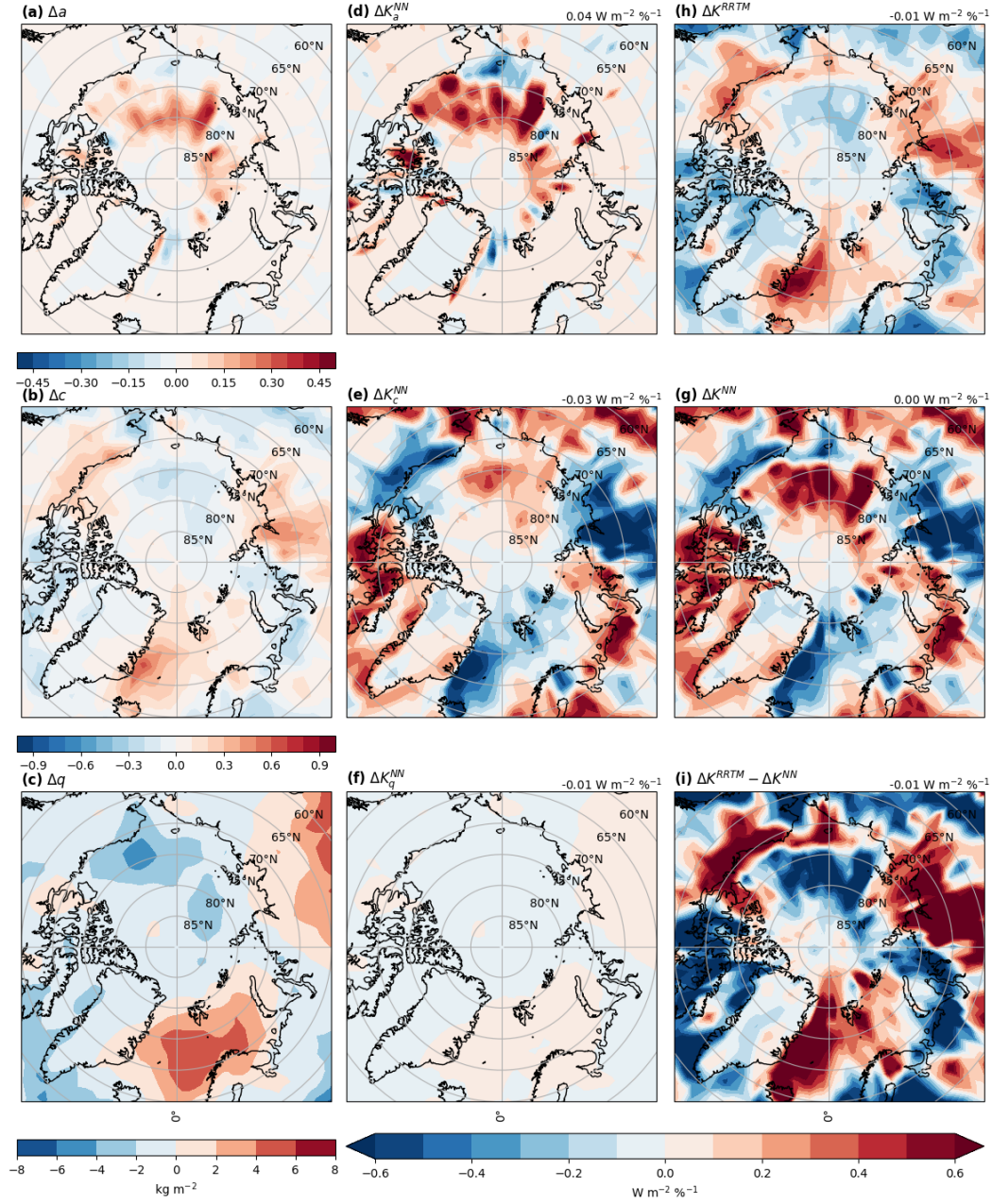


Figure 6: The same as Figure 5, but feedbacks are quantified using the baseline neural network.

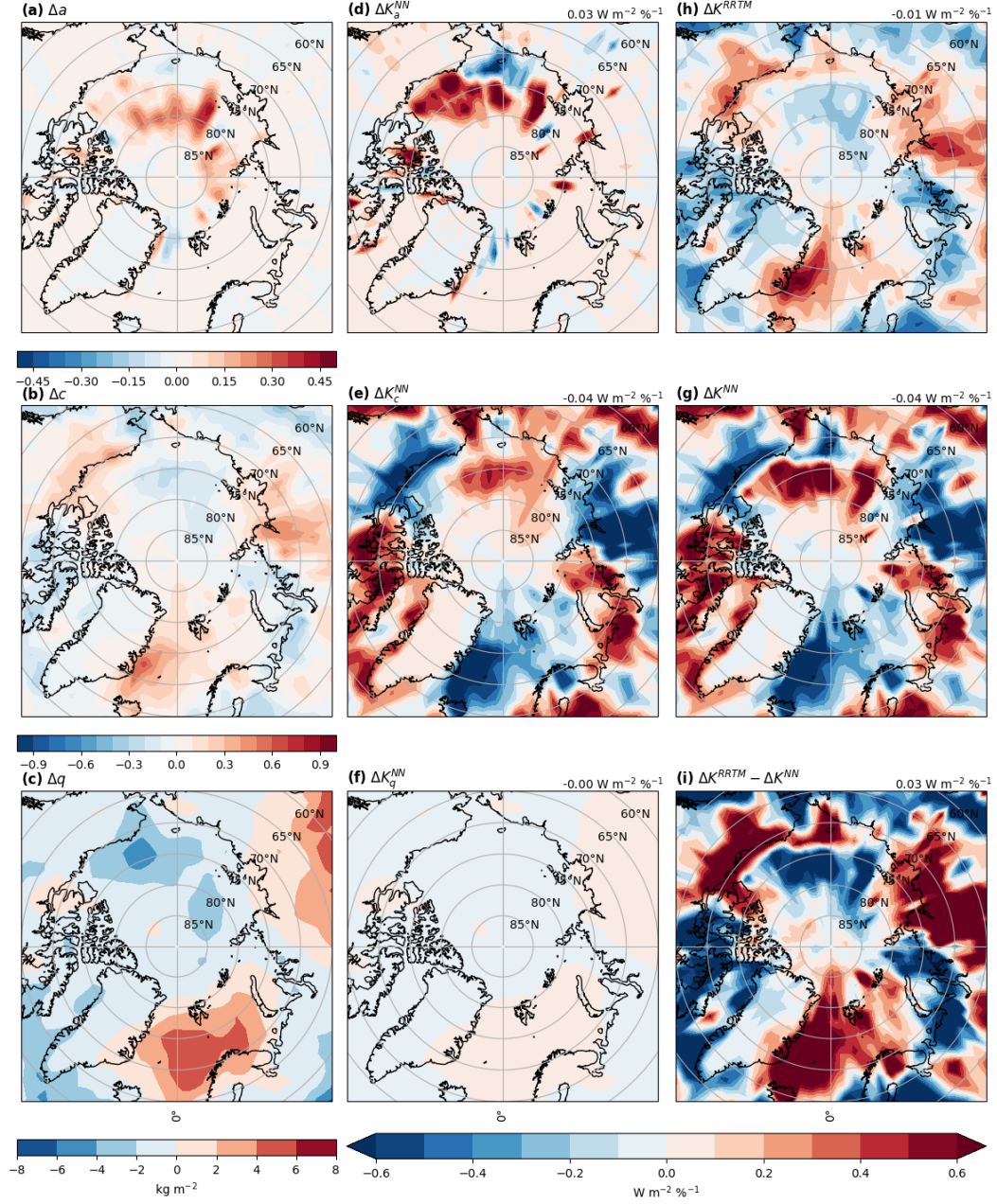


Figure 7: The same as Figure 5, but but feedbacks are quantified using the baseline neural network trained with clear-sky training samples.

E Plots of Univariate Radiative Feedbacks

In Figures 8 and 9, we compare our NN and the kernel methods' quantification of univariate radiative feedbacks in the interannual variability case study in all and clear-sky conditions in the Arctic. In Figures 10 and 11, we show the results obtained by applying the same methods of feedback quantification when applied globally.

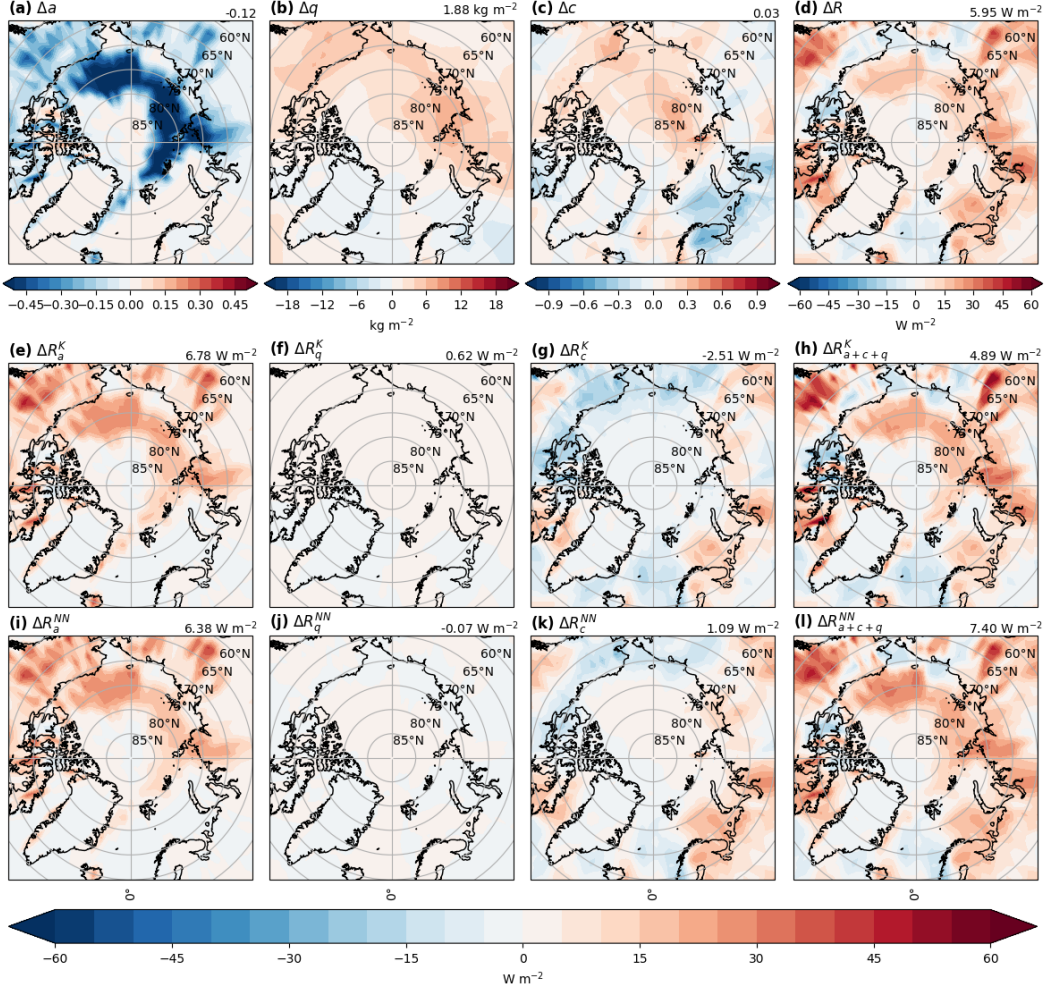


Figure 8: All-sky Arctic TOA net SW flux change (from September 1992 to September 2020) compared with the sum of univariate albedo (a), water vapour (q), and cloud (c) feedbacks quantified by the kernel (K) and neural network (NN) methods. **(a-d)** ERA5-derived changes in albedo, water vapour, effective cloud optical depth, and TOA SW flux. **(e-h)** Kernel-method-computed feedbacks: albedo, water vapour, cloud, and their sum. **(i-l)** NN-method-computed feedbacks: albedo, water vapour, cloud, and their sum. Values above each panel are area-weighted means. Plots (a-d) have their own individual colourbar, while plots (e-l) share a common colourbar.

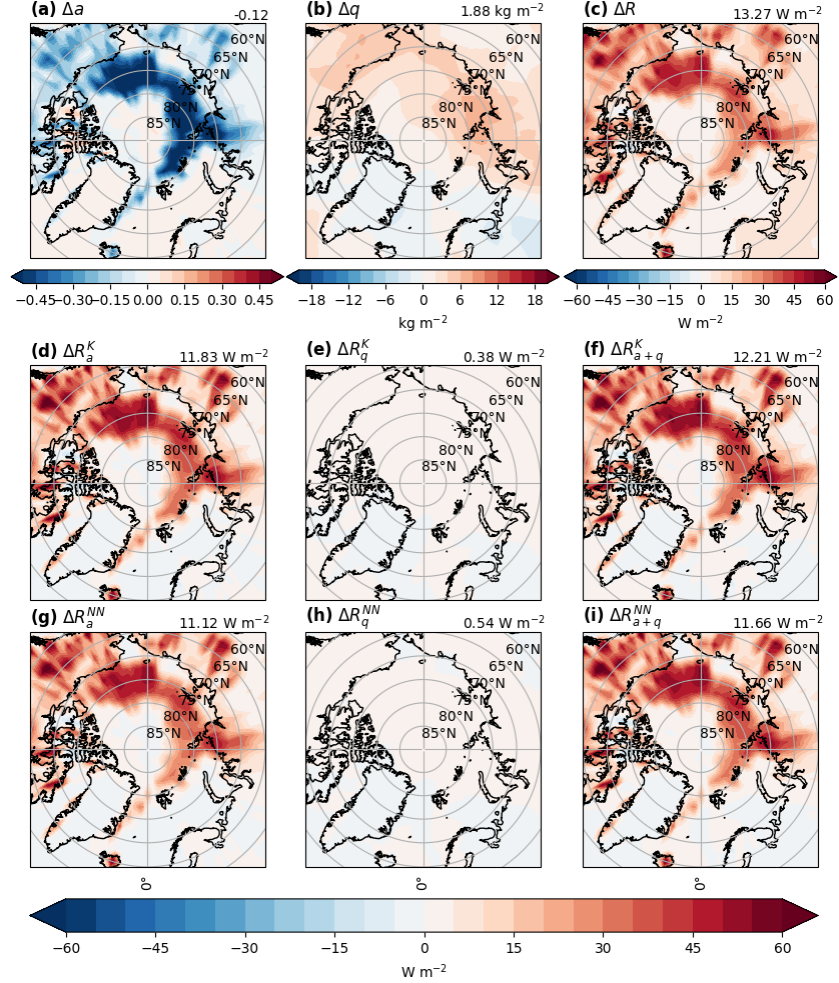


Figure 9: Clear-sky Arctic TOA net SW flux change (from September 1992 to September 2020) compared with the sum of univariate albedo (a) and water vapour (q) feedbacks, quantified by the kernel (K) and neural network (NN) methods. (a-c) ERA5-derived changes in albedo, water vapour, and clear-sky TOA SW flux. (d-f) Kernel-method feedbacks: albedo, water vapour, and their sum. (g-i) NN-method feedbacks: albedo, water vapour, and their sum. Values above each panel are area-weighted means.

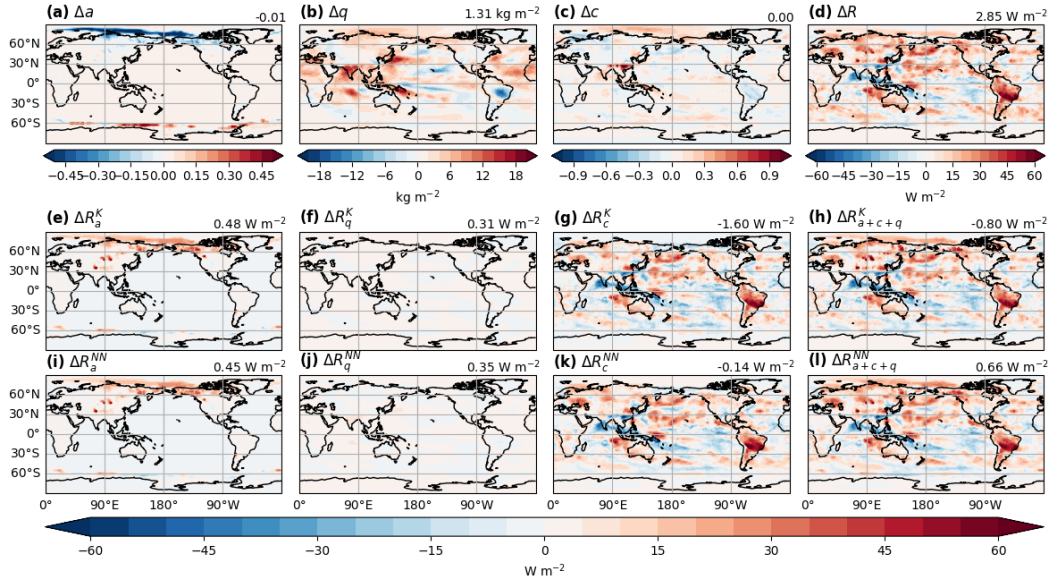


Figure 10: Like Fig. 8, but for the whole earth.

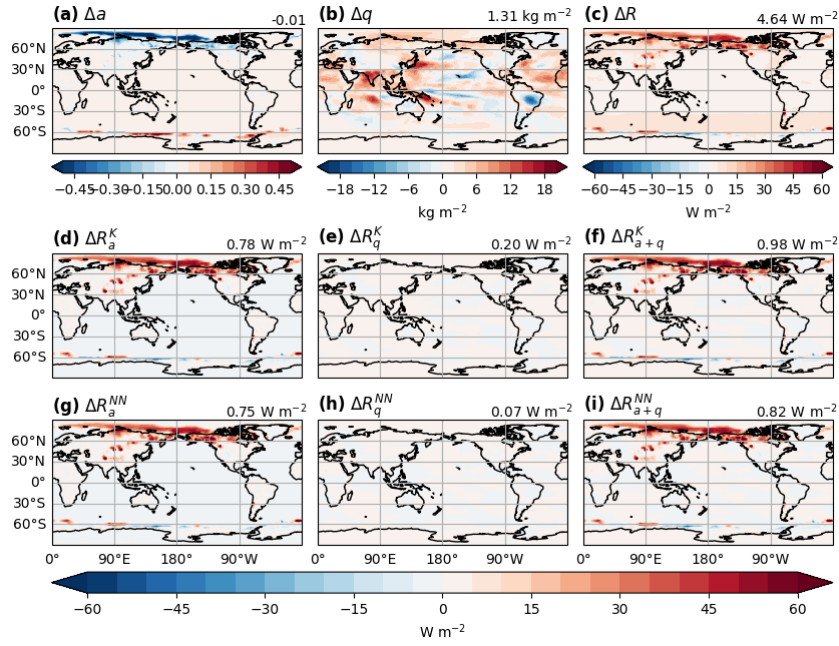


Figure 11: Like Fig. 9, but the whole earth.

F Detailed Closure Test Metrics

To ablate and comprehensively validate the model on the case study test, we report the MBE and RMSE obtained both globally and for the Arctic, comparing the predictions of ΔR with the NN-predicted feedback considering all cross terms, the NN-predicted feedback as a sum of univariate feedbacks, and the kernel method. Note that the MBE and RMSE of the all-sky and clear-sky cases for the kernel method are equal by assumption of the cloud radiative effect method, causing artificially small biases and errors in the all-sky case.

Table 4: Global performance of the different models and residual formulations. Values are reported as mean $\pm$ standard deviation, for a set of five models trained with different starting seeds.

Model	Prediction	MBE		RMSE	
		All-sky	Clear-sky	All-sky	Clear-sky
NN	ΔR_{a+c+q}^{NN}	2.379 ± 0.079	3.828 ± 0.136	5.635 ± 0.020	4.711 ± 0.170
	$\Delta R_{\text{allcross}}^{NN}$	2.599 ± 0.107	4.305 ± 0.161	5.669 ± 0.079	5.062 ± 0.173
NN+clear-sky	ΔR_{a+c+q}^{NN}	2.266 ± 0.028	3.785 ± 0.019	5.650 ± 0.060	4.400 ± 0.027
	$\Delta R_{\text{allcross}}^{NN}$	2.472 ± 0.067	4.244 ± 0.038	5.693 ± 0.068	4.748 ± 0.027
NN+clear-sky+Sob.	ΔR_{a+c+q}^{NN}	2.320 ± 0.060	3.829 ± 0.024	5.686 ± 0.066	4.458 ± 0.015
	$\Delta R_{\text{allcross}}^{NN}$	2.604 ± 0.040	4.270 ± 0.076	5.754 ± 0.041	4.790 ± 0.068
Kernel	ΔR_{a+c+q}^K	3.658	3.658	4.370	4.370

Table 5: Arctic performance of the different models and residual formulations. Values are reported as mean $\pm$ standard deviation.

Model	Prediction	MBE		RMSE	
		All-sky	Clear-sky	All-sky	Clear-sky
NN	ΔR_{a+c+q}^{NN}	-0.373 ± 0.222	0.379 ± 0.841	4.042 ± 0.374	7.524 ± 0.984
	$\Delta R_{\text{allcross}}^{NN}$	1.081 ± 0.184	1.840 ± 0.550	3.391 ± 0.061	7.242 ± 0.616
NN+clear-sky	ΔR_{a+c+q}^{NN}	-0.325 ± 0.491	1.646 ± 0.386	3.947 ± 0.313	5.786 ± 0.187
	$\Delta R_{\text{allcross}}^{NN}$	0.695 ± 0.186	3.088 ± 0.213	3.297 ± 0.098	5.840 ± 0.090
NN+clear-sky+Sob.	ΔR_{a+c+q}^{NN}	-0.861 ± 0.176	1.367 ± 0.138	4.332 ± 0.317	5.932 ± 0.065
	$\Delta R_{\text{allcross}}^{NN}$	0.306 ± 0.200	2.557 ± 0.205	3.388 ± 0.053	5.947 ± 0.097
Kernel	ΔR_{a+c+q}^K	1.065	1.065	5.952	5.952