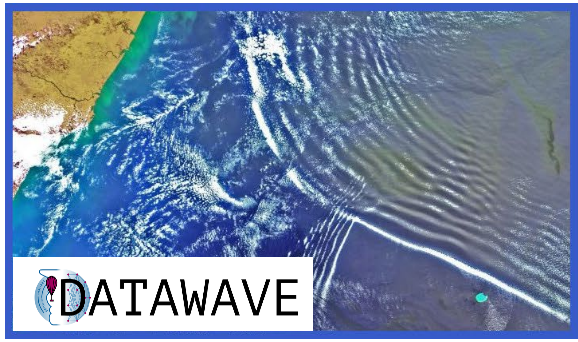
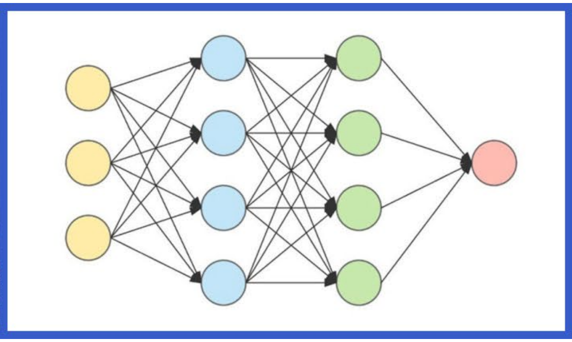
Data Imbalance, Uncertainty Quantification, and Generalization via Transfer Learning in Data-driven Parameterizations

Lessons from the Emulation of Gravity Wave Momentum Transport in WACCM





https://cssi-gws.github.io/index.html

Hamid A. Pahlavan

NorthWest Research Associates

Y. Qiang Sun, Ashesh Chattopadhyay, Pedram Hassanzadeh, Sandro W. Lubis

M. Joan Alexander, Edwin Gerber, Aditi Sheshadri, Yifei Guan

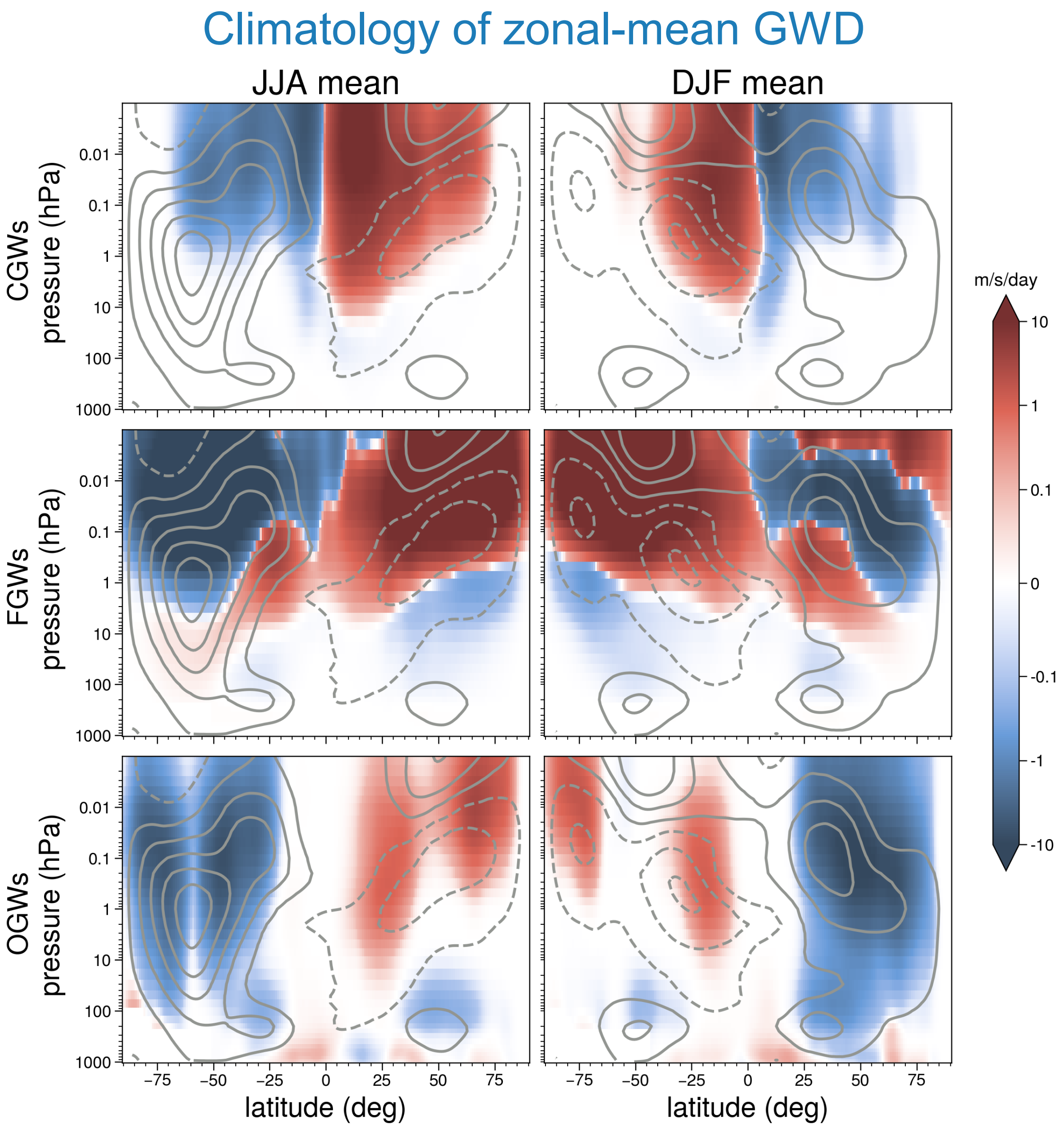
February 14, 2024

Emulating the complex physics-based GW parameterization in WACCM serves as a testbed for exploring solutions to these challenges

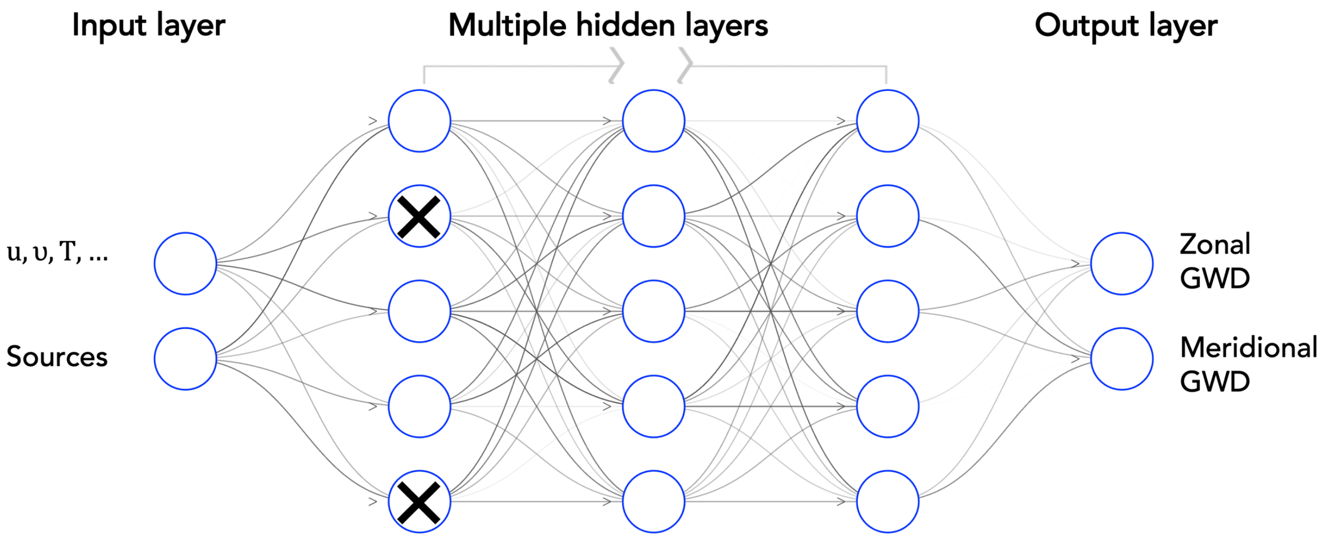
Convective GWs (CGWs)

Frontal GWs (FGWs)

Orographic GWs (OGWs)



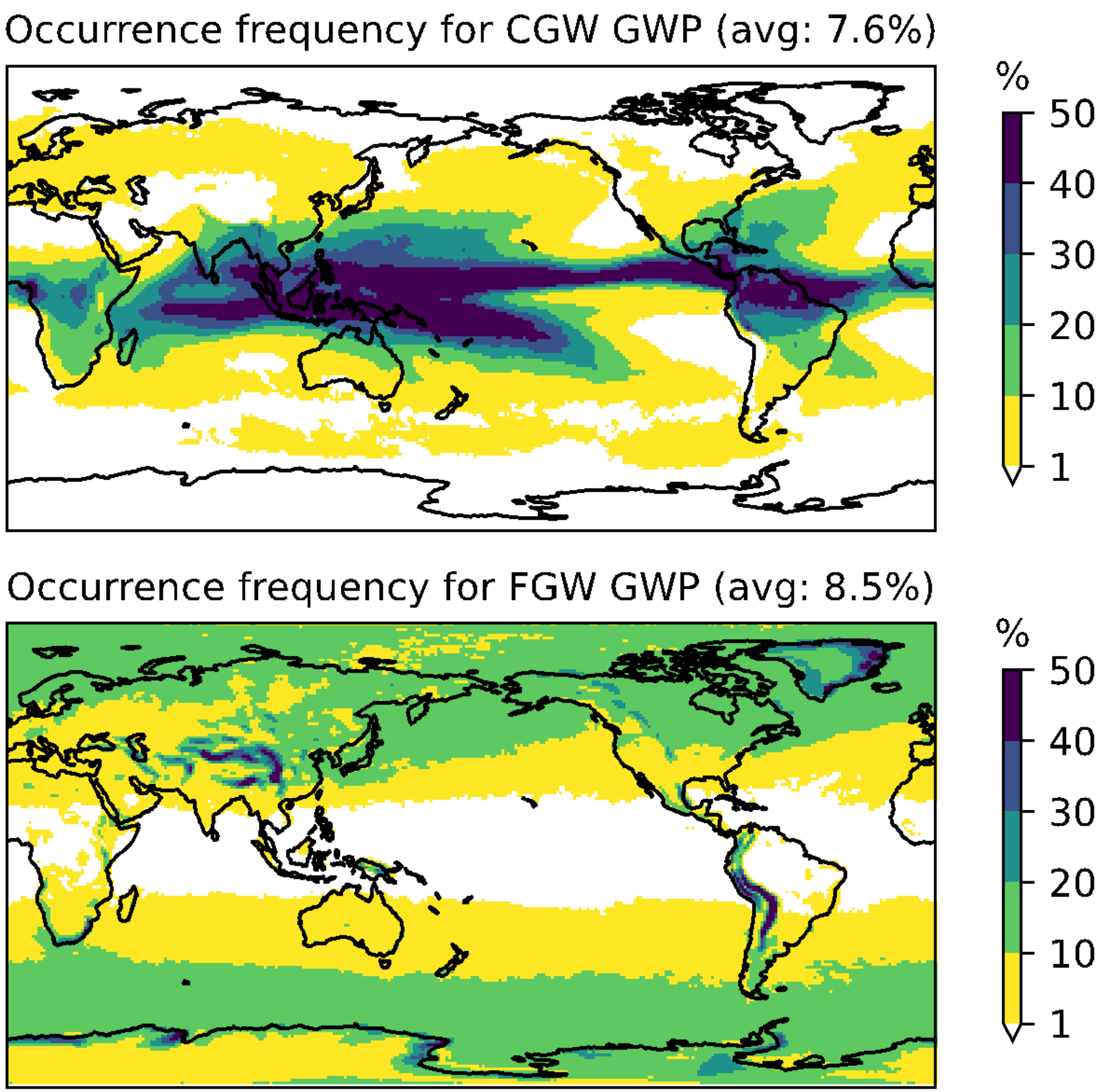
Fully connected NNs are used as emulators



GWP	Input			Output
	pressure levels	surface level	forcing	
CGWs	$u(70)$, $v(70)$, $T(70)$, $z(70)$, $\rho(71)$, Brunt–Väisälä frequency N (70), dry static energy DSE (70)	lat (1), lon (1), $P_{surface}$ (1),	diabatic heating (70)	zonal drag GWD_x (70), meridional drag GWD_y (70),
FGWs			frontogenesis function (70)	
OGWs			mxdis (16), hwdth (16), clngt (16), angll (16), anixy (16),	

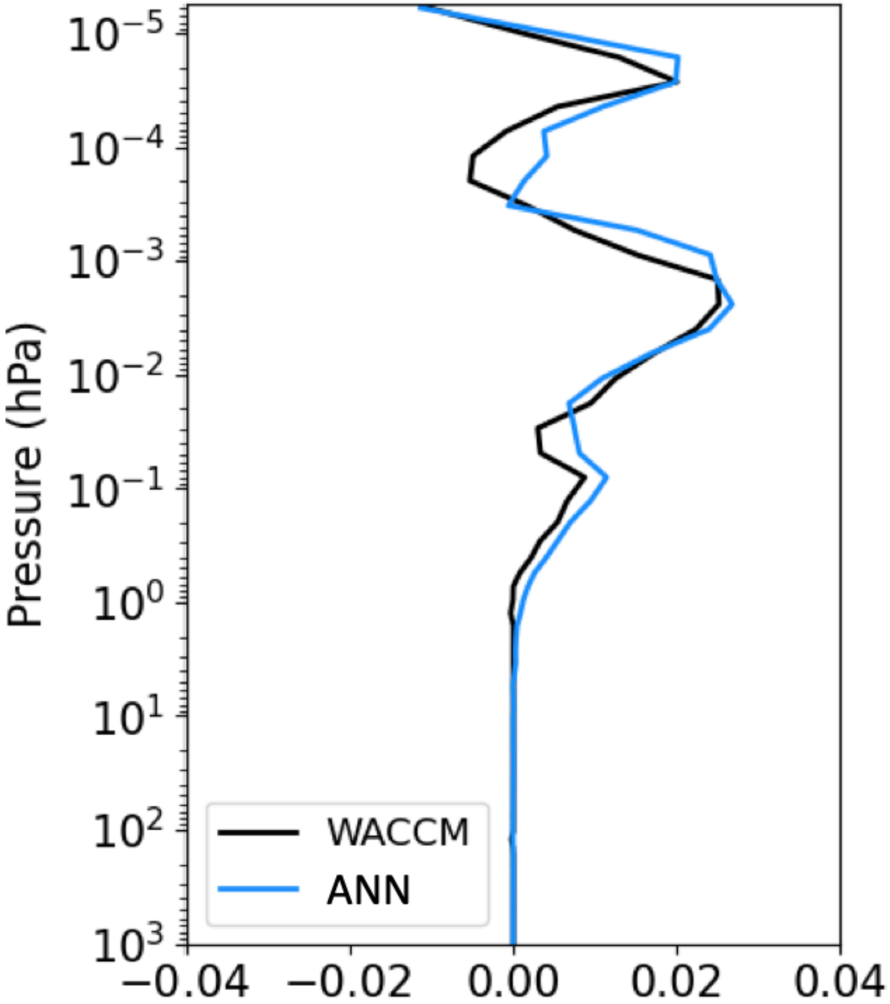
$$\mathcal{L}(\Theta) = \frac{1}{n}\sum_{i=1}^{n}\left\|\mathbf{NN}\left(x_i, \Theta\right) - y_i\right\|_2^2$$

The heterogeneous and intermittent nature of GW sources leads to a significantly imbalanced dataset

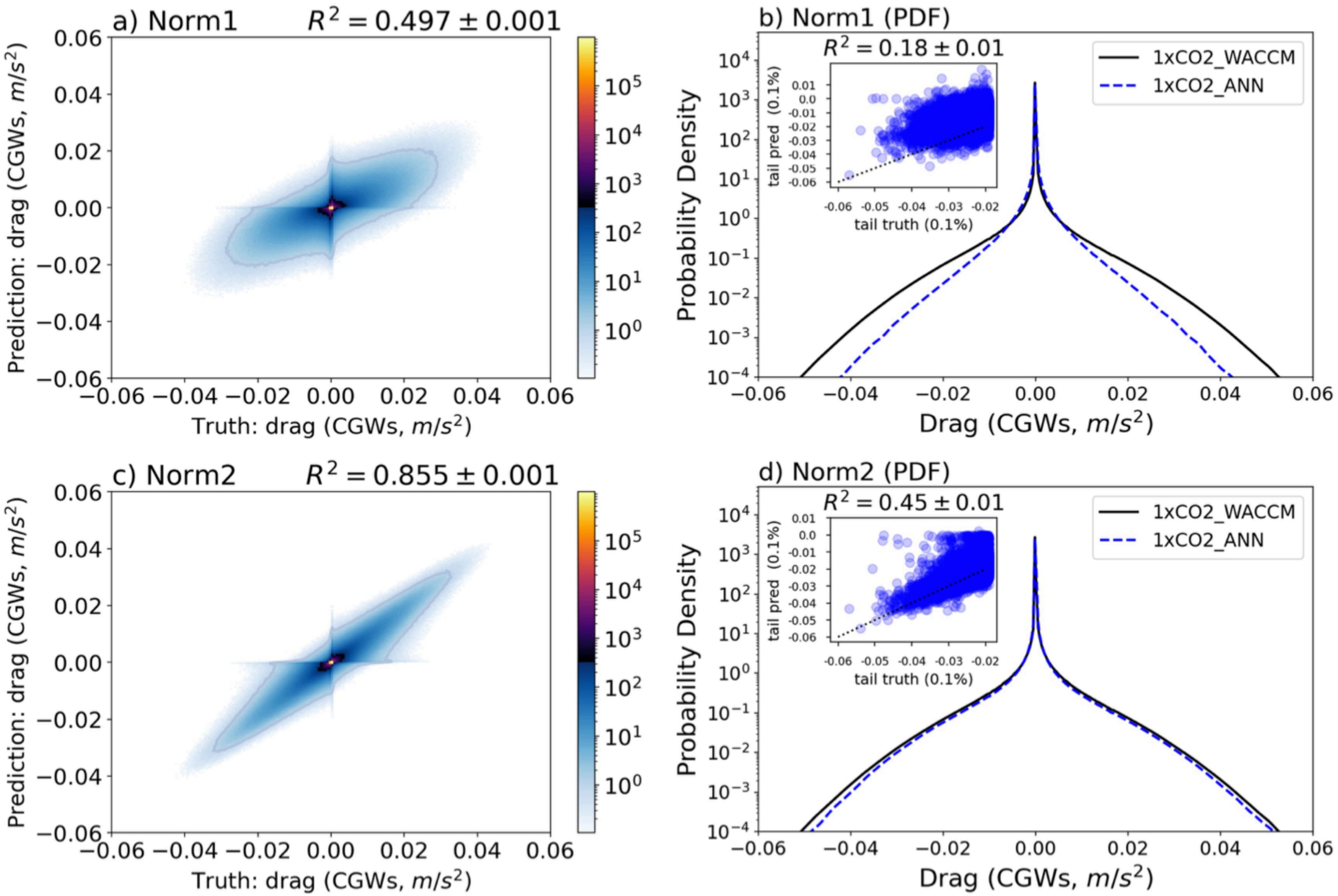


GW drags concentrate primarily at critical levels, resulting in non-smooth profiles with numerous levels exhibiting zero GW drag

A sample profile of CGWs



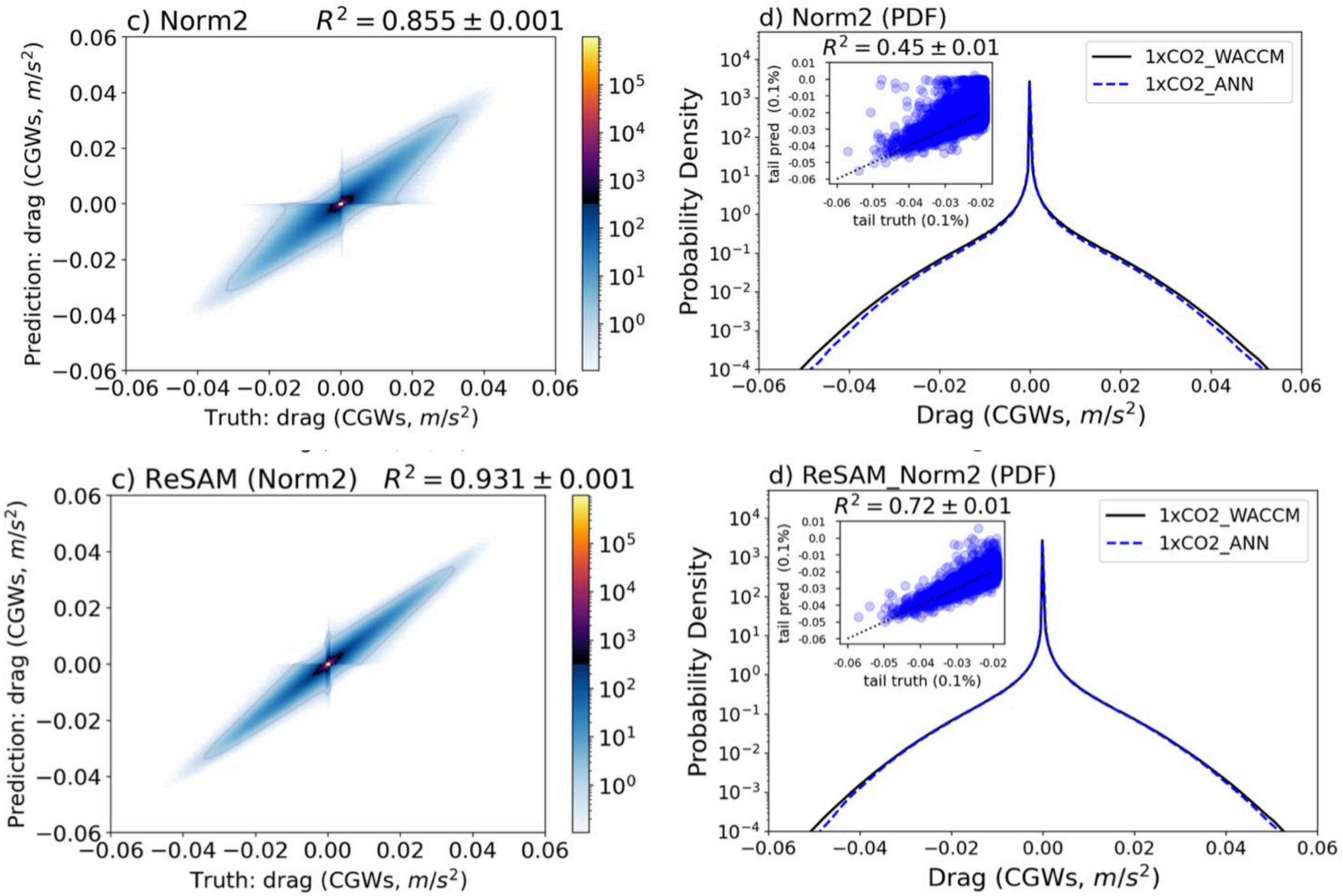
Normalizing the data while preserving the original wind and GWD profile structure enhances the emulator's performance



Norm1: $u_{norm1} = \frac{u - \bar{u}}{std(u)}$

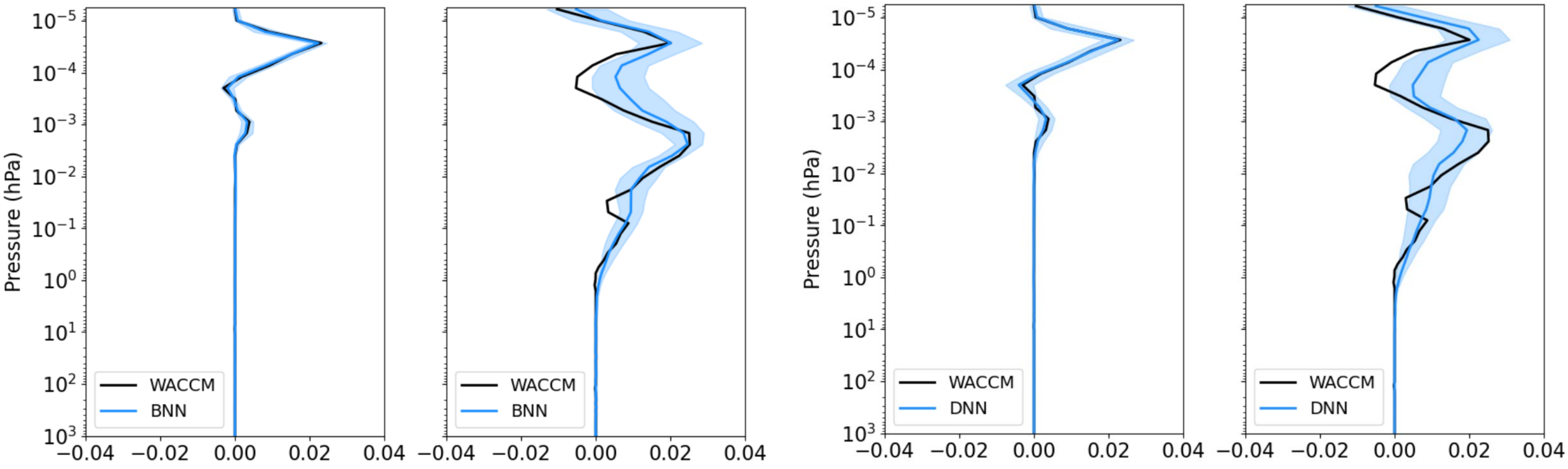
Norm2: $u_{norm2} = \frac{u}{\max(std(u))}$

Resampling the data (ReSAM): limiting the number of sample pairs with zero GWD to match the number of samples with non-zero GWD

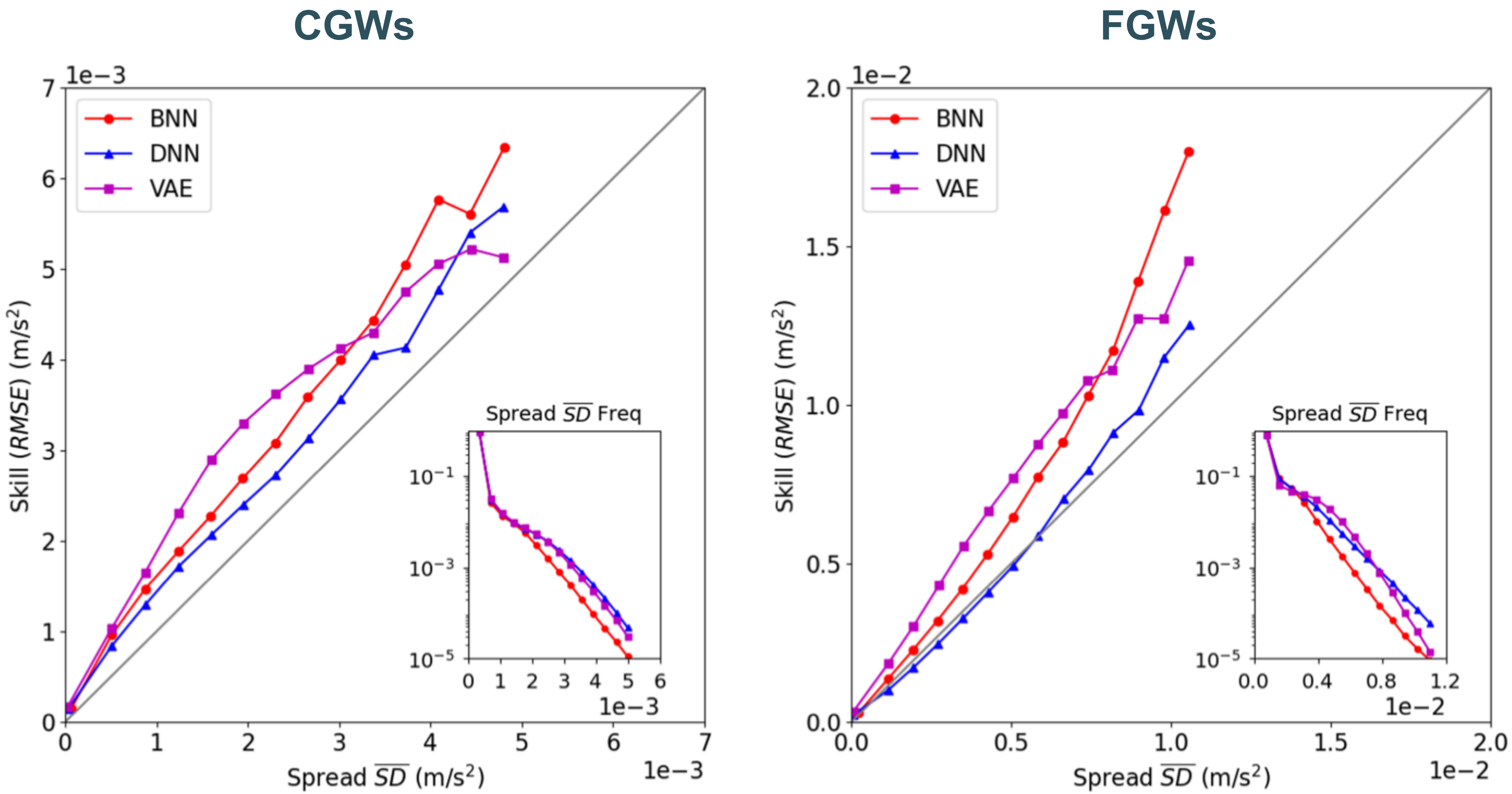


Uncertainty quantification (UQ) provides a credible confidence level for each prediction, serving as a reliable indicator of its accuracy

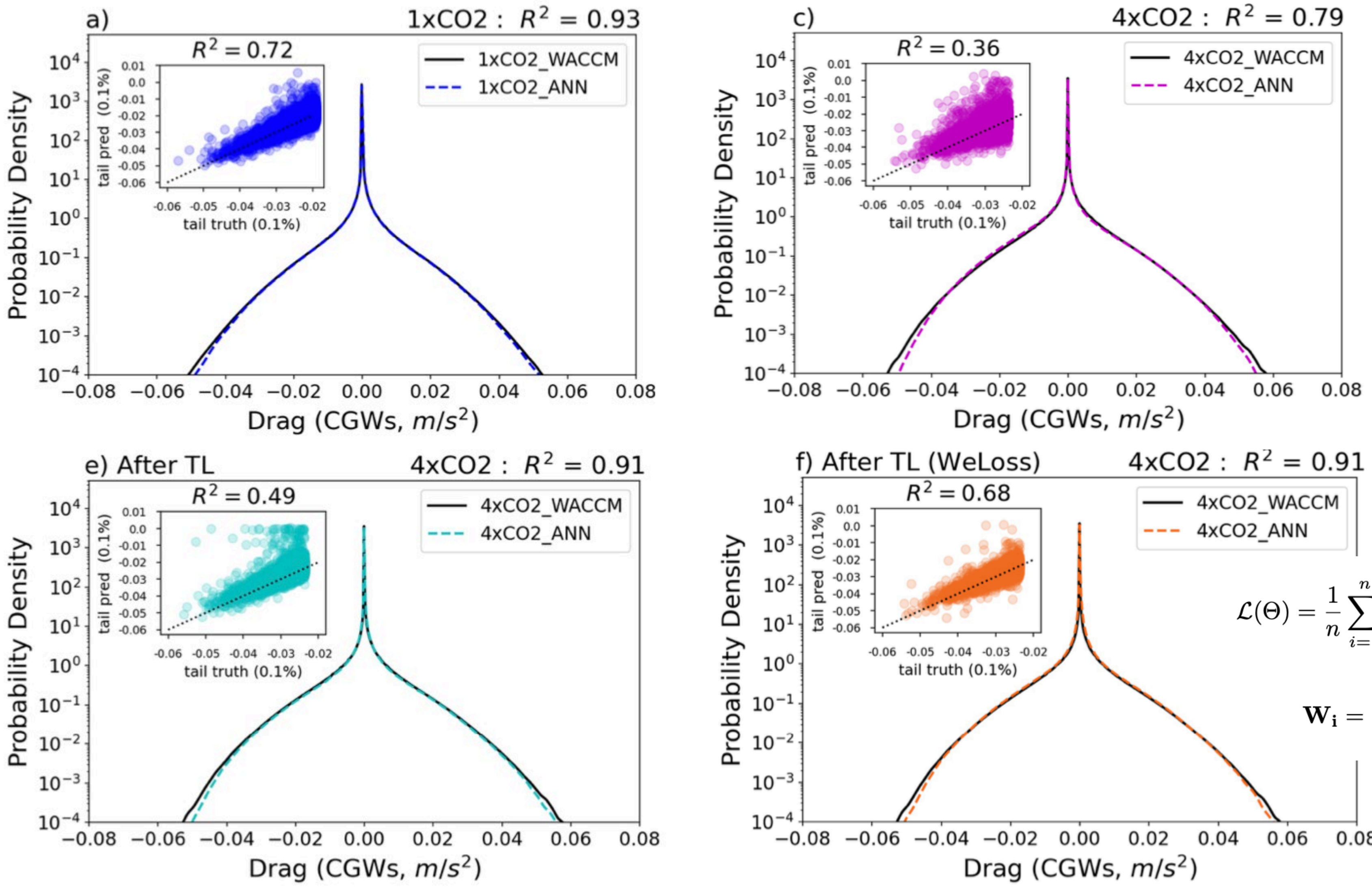
- Bayesian Neural Network (BNN)
- Dropout Neural Network (DNN)
- Variational Auto-Encoder (VAE)



All three UQ methods produce reasonably informative uncertainty estimates, as their curves closely align with the 1-to-1 line



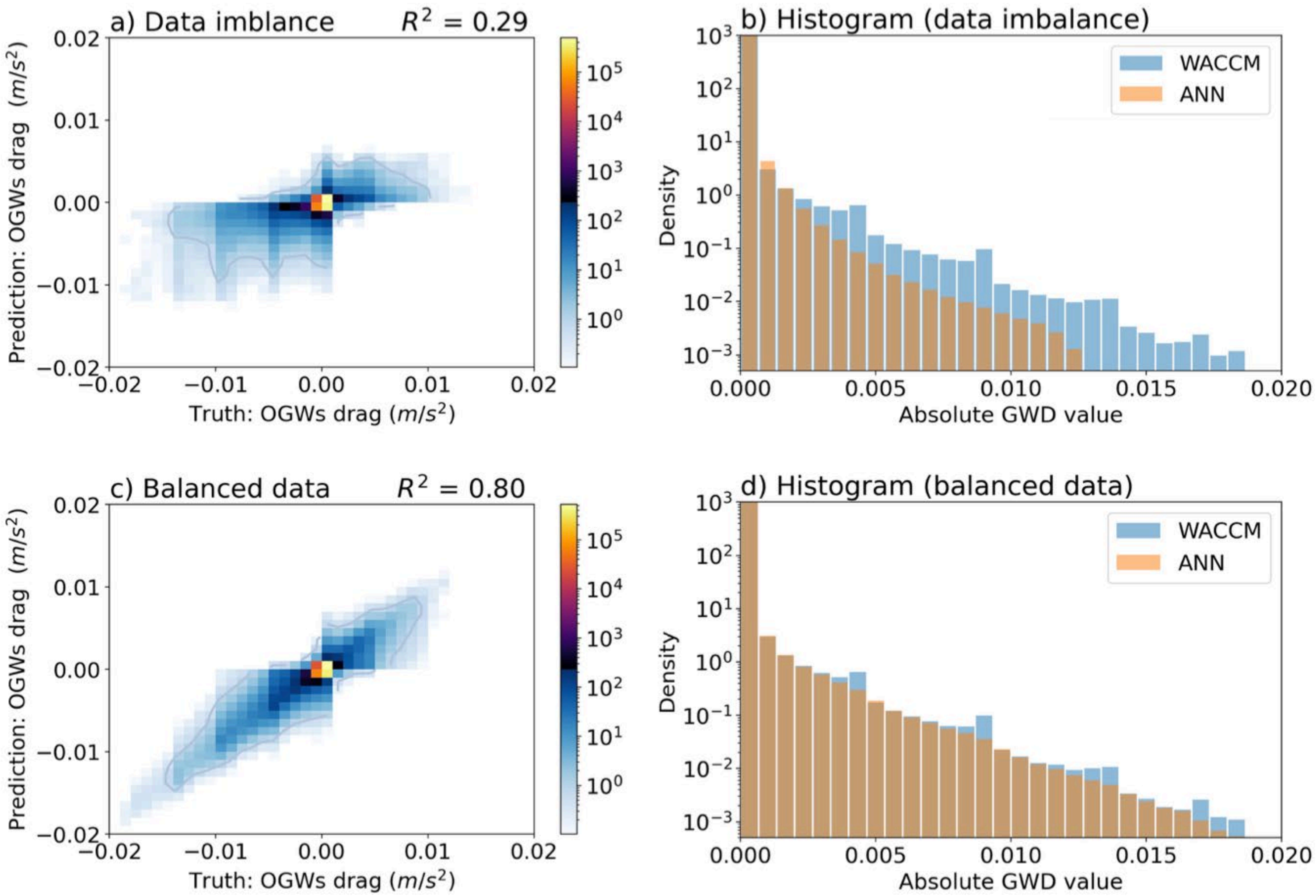
Transfer learning improves out-of-distribution generalization of the NNs under 4×CO2 forcing



$$\mathcal{L}(\Theta) = \frac{1}{n}\sum_{i=1}^{n} \left\| \mathbf{W_i}\{\mathbf{NN}(x_i, \Theta) - y_i\} \right\|_2^2$$

$$\mathbf{W_i} = \frac{1}{PDF(max(|y_i(z)|))}$$

The data imbalance issue is particularly pronounced for the OGWs



Take-home points

- WACCM's orographic, convective, and frontal GWP are emulated using NNs.
- **Data imbalance** is addressed via **resampling** and **weighted loss**.
- **Uncertainty quantification** is addressed via **Bayesian**, **dropout**, and **variational** methods.
- **Out-of-distribution generalization** under 4×CO2 forcing is enabled via **transfer learning**.
- These findings apply to the data-driven parameterizations of other climate processes.

Pahlavan@nwra.com

arXiv:2311.17078

A library of high-resolution simulations with regional WRF model

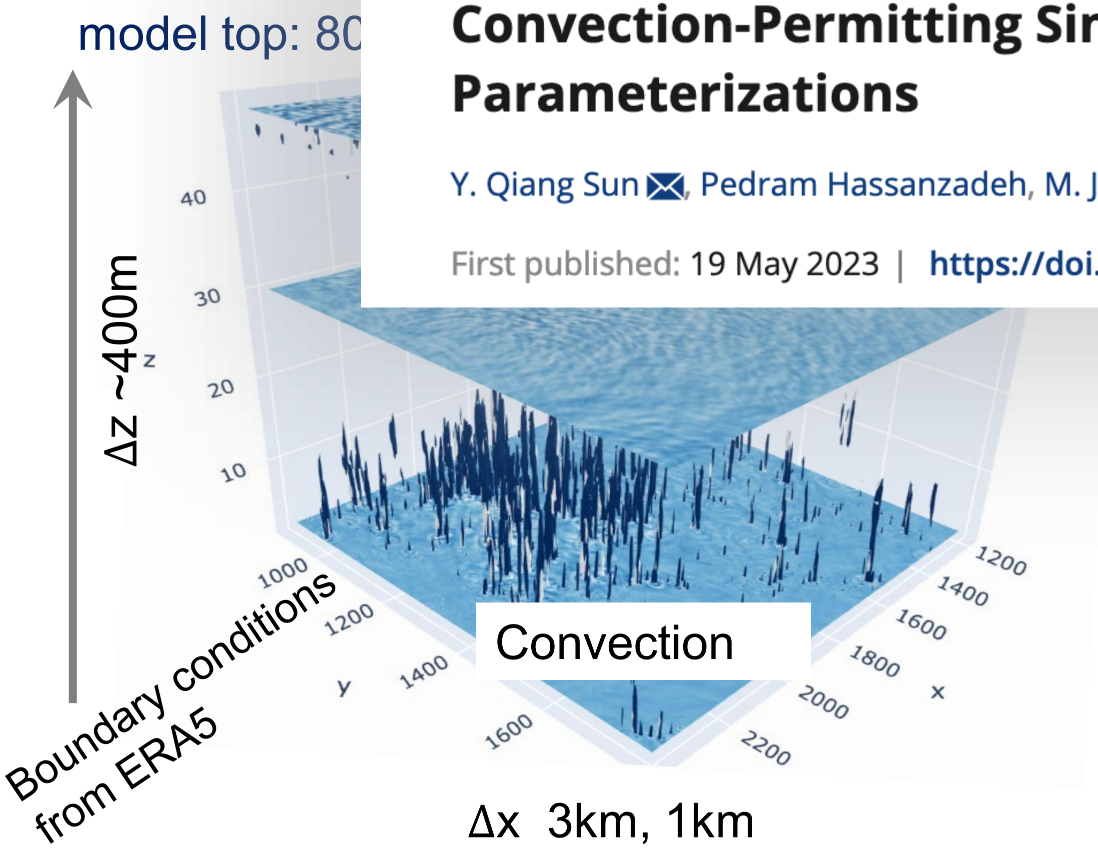
- Well-vetted physics at convection-permitting resolutions
- Constrained by reanalysis on the boundaries (no model drift)
- Sampling ... ming)

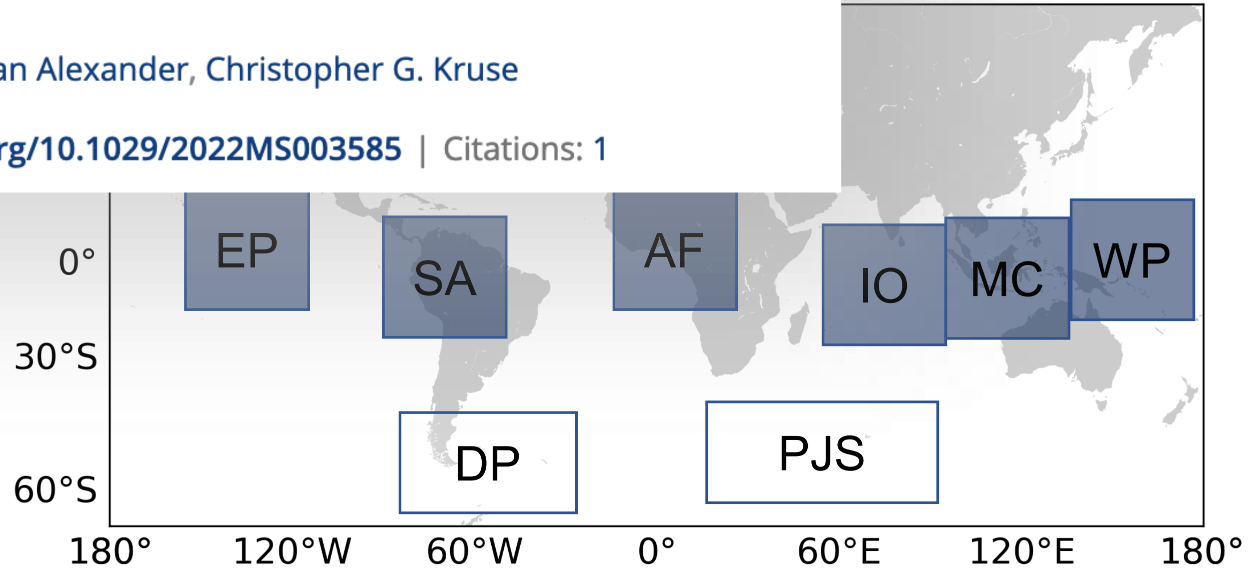
Research Article | Open Access

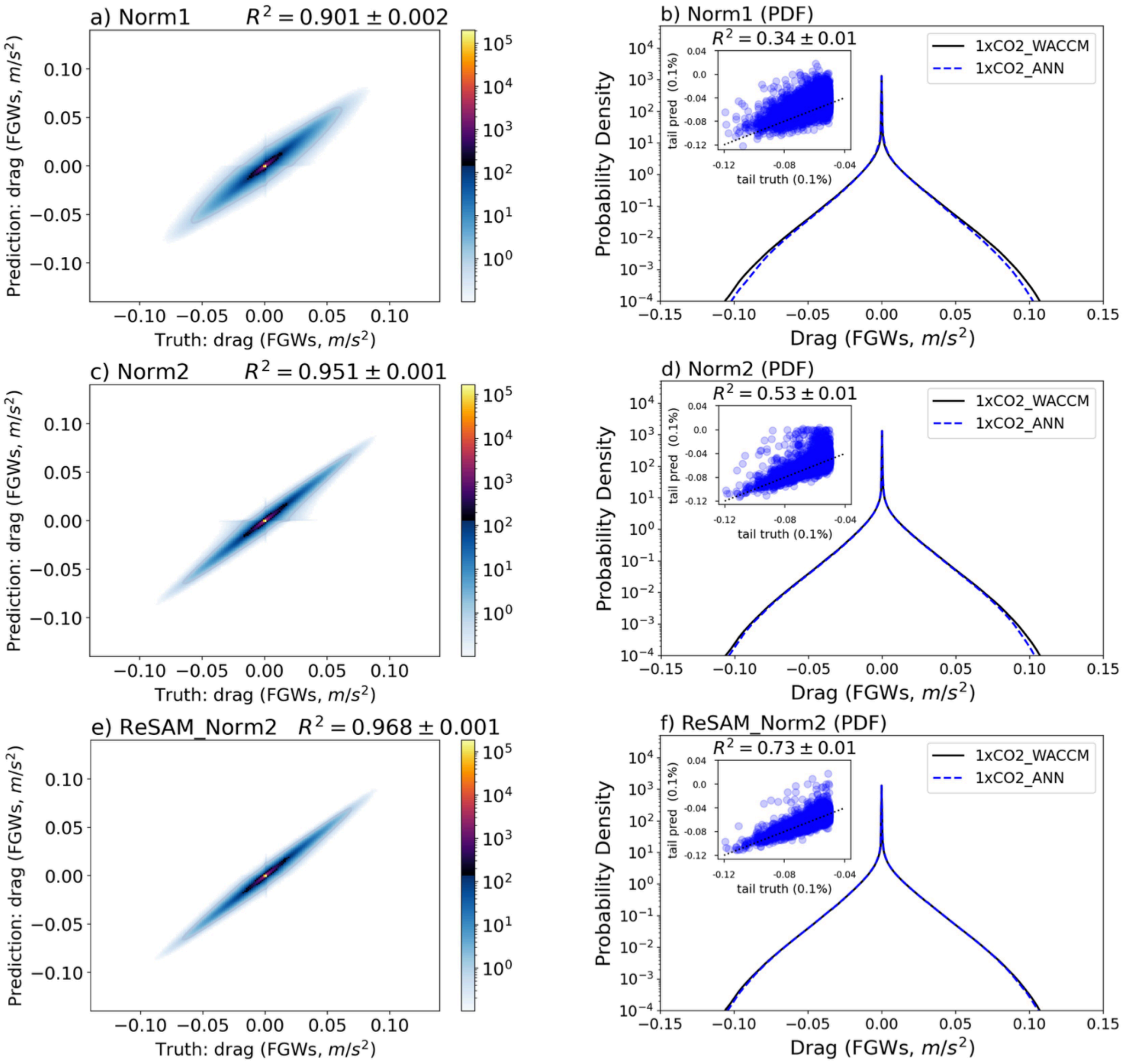
Quantifying 3D Gravity Wave Drag in a Library of Tropical Convection-Permitting Simulations for Data-Driven Parameterizations

Y. Qiang Sun, Pedram Hassanzadeh, M. Joan Alexander, Christopher G. Kruse

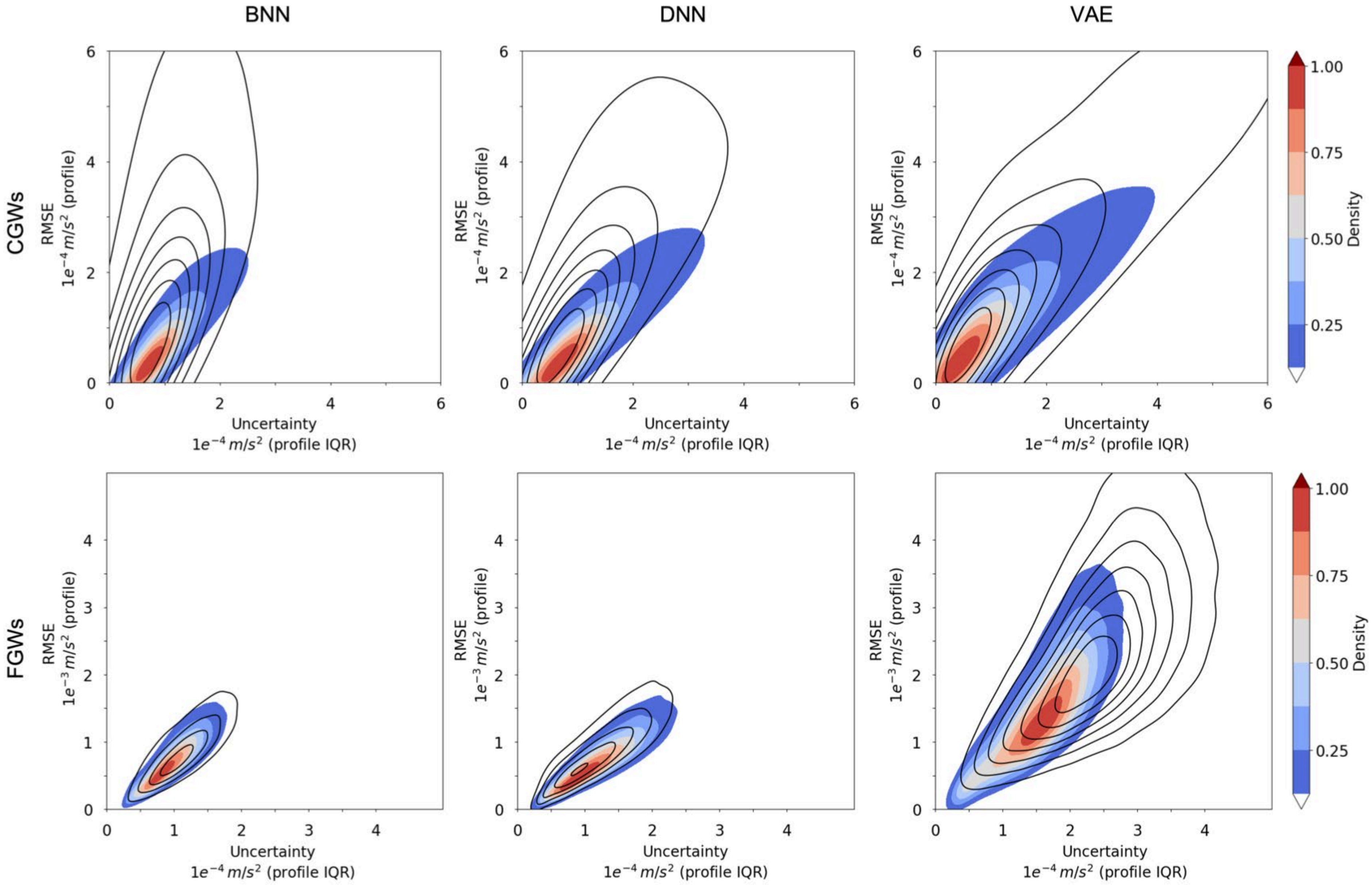
First published: 19 May 2023 | https://doi.org/10.1029/2022MS003585 | Citations: 1







a) Norm1 $R^2 = 0.901 \pm 0.002$
b) Norm1 (PDF)
$R^2 = 0.34 \pm 0.01$
c) Norm2 $R^2 = 0.951 \pm 0.001$
d) Norm2 (PDF)
$R^2 = 0.53 \pm 0.01$
e) ReSAM_Norm2 $R^2 = 0.968 \pm 0.001$
f) ReSAM_Norm2 (PDF)
$R^2 = 0.73 \pm 0.01$
Prediction: drag (FGWs, m/s^2)
Truth: drag (FGWs, m/s^2)
Probability Density
Drag (FGWs, m/s^2)
tail pred (0.1%)
tail truth (0.1%)
1xCO2_WACCM
1xCO2_ANN



BNN
DNN
VAE
CGWs
FGWs
RMSE
$1e^{-4}\,m/s^2$ (profile)
$1e^{-3}\,m/s^2$ (profile)
Uncertainty
$1e^{-4}\,m/s^2$ (profile IQR)
Density
1.00
0.75
0.50
0.25

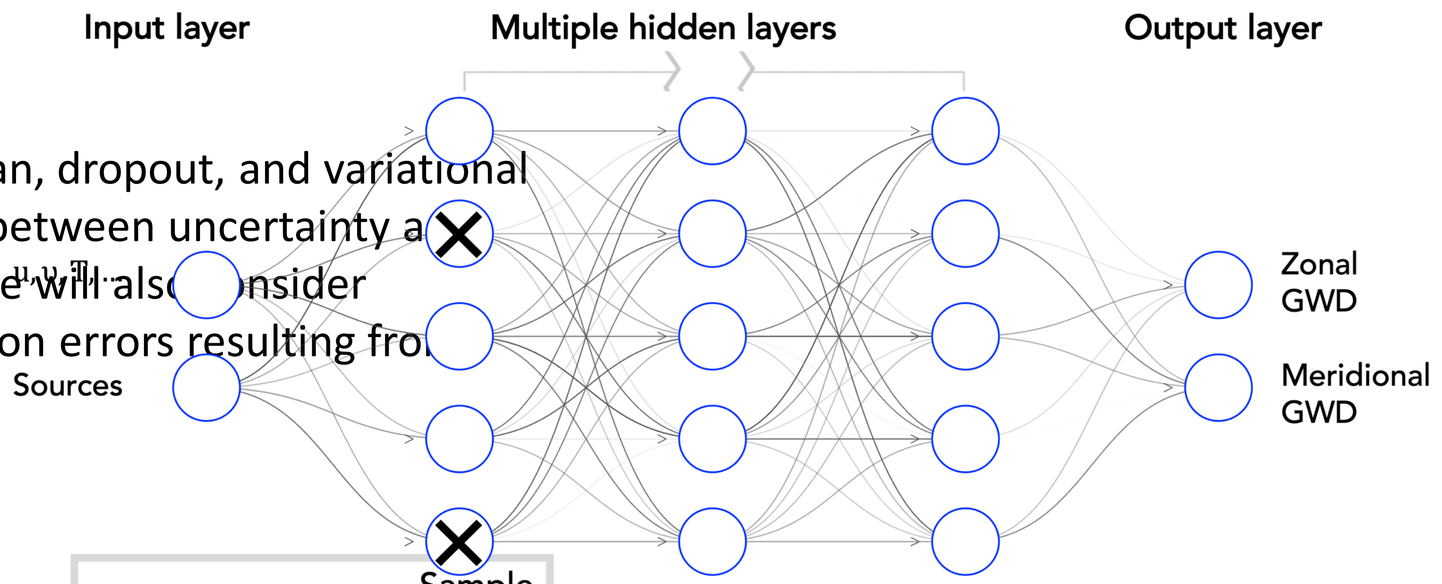
Table 2. Change of Mahalanobis distance based on the ratio of the average distance of the points that are more than 3 standard deviations away from the mean. The choice of the variables here is based on Appendix A, showing u, v, T, and source function contain most of the information needed for the NN.

Variables	u	v	T	Source (diabatic heating for CGWs, frontogenesis for FGWs)	Zonal drag	Meridional drag
Distance (Convection)	1.03	1.00	1.19	3.62	1.42	1.44
Distance (Front)	1.03	0.96	1.50	1.10	1.00	1.00

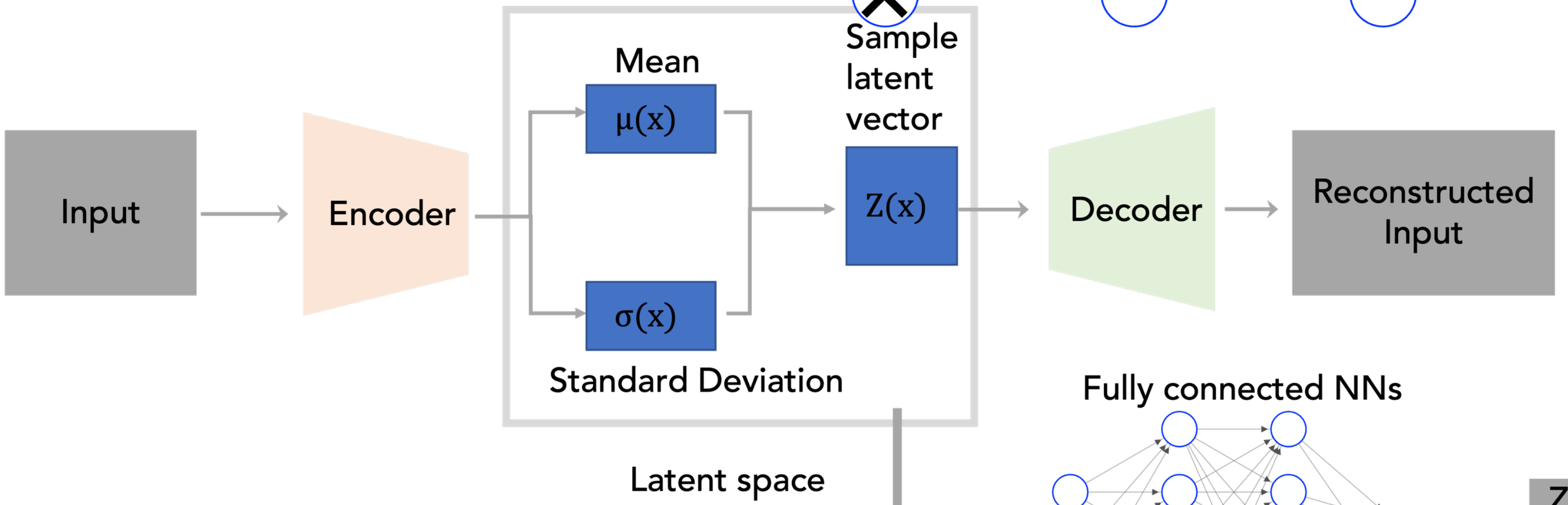
OOD generalization

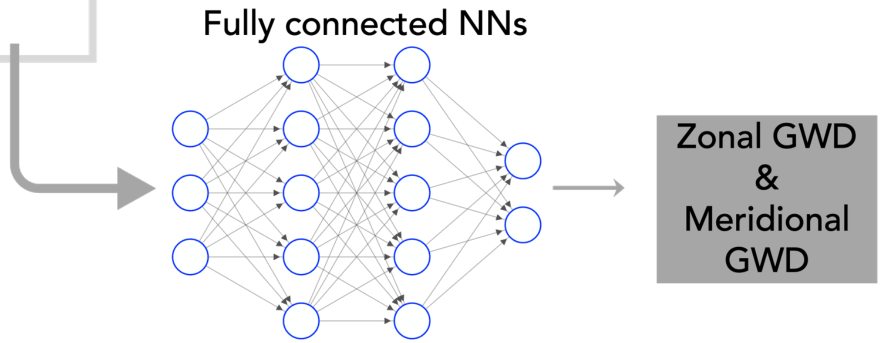
(extrapolation to a test data distri- bution different from that of the training set) is a major challenge for applications involving non-stationarity, like a changing climate
A general and powerful method for im- proving the OOD generalization capability of NNs is transfer learning (TL), which involves re-training a few or all of the layers of a NN using a small amount of data from the new system

UQ



three common UQ methods (Bayesian, dropout, and variational NNs) by analyzing the relationships between uncertainty a accuracy during inference testing. We will als onsider scenarios involving OOD generalization errors resulting fro global warming.





Bayesian neural network (BNN), dropout neural network (DNN), and variational auto-encoder (VAE)

Take-home points

- WACCM's orographic, convective, and frontal GWP are emulated using NNs.
- Data imbalance is addressed via resampling and weighted loss.
- Uncertainty quantification is addressed via Bayesian, dropout, and variational methods.
- Out-of-distribution generalization of the NNs under 4×CO2 forcing is enabled via transfer learning.
- These findings apply to the data-driven parameterizations of other climate processes.

Pahlavan@nwra.com

arXiv:2311.17078

The effect of normalization method