Learning Cloud Processes Across Scales Using Scientific Machine Learning

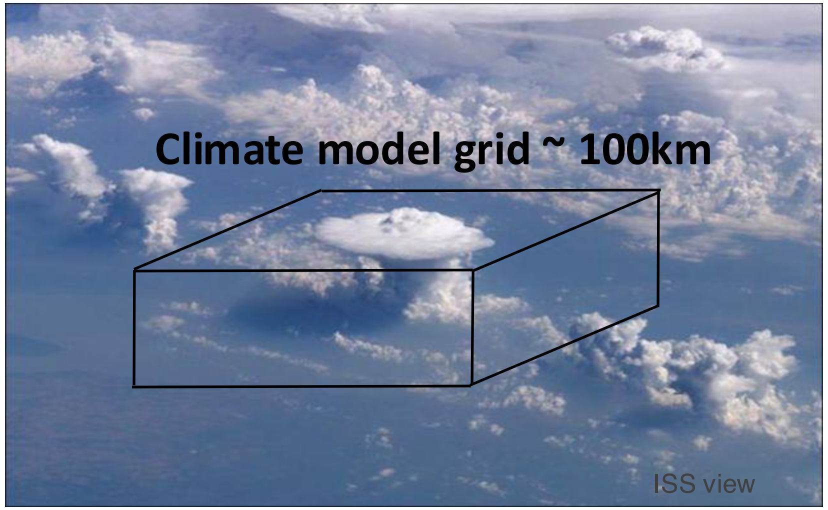
Kara D. Lamb

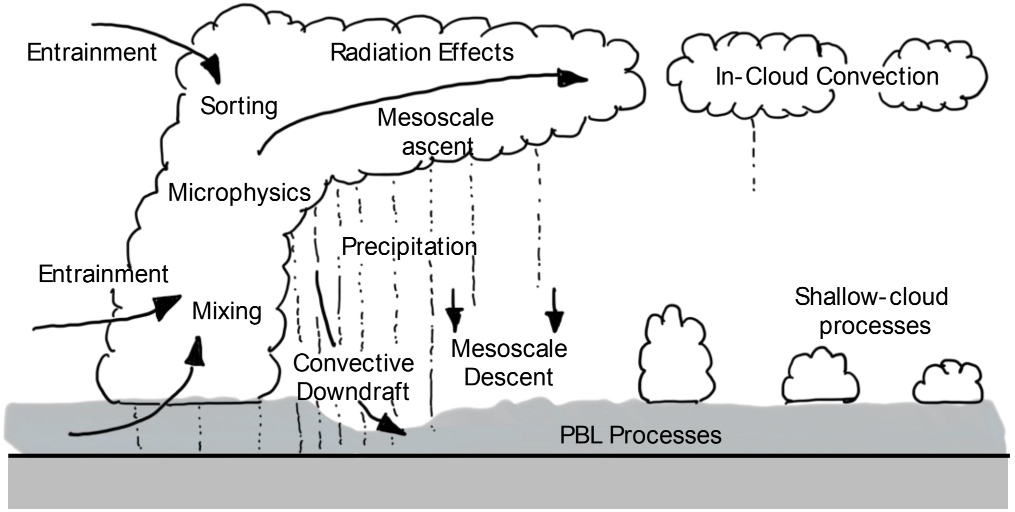
Marcus van Lier Walqui, Sean Santos, Hugh Morrison, Jerry Harrington, Joseph Ko, Obin Sturm, Al Moyle, Jonas Mikhaeil, Justus Will, Stephan Mandt, Andrea Jenney, Mike Pritchard, Colleen Kaul, Kyle Pressl, Po-Lun Ma, Sara Shamekh, Yu Huang, Pierre Gentine

ELLIIT Focus Period Symposium

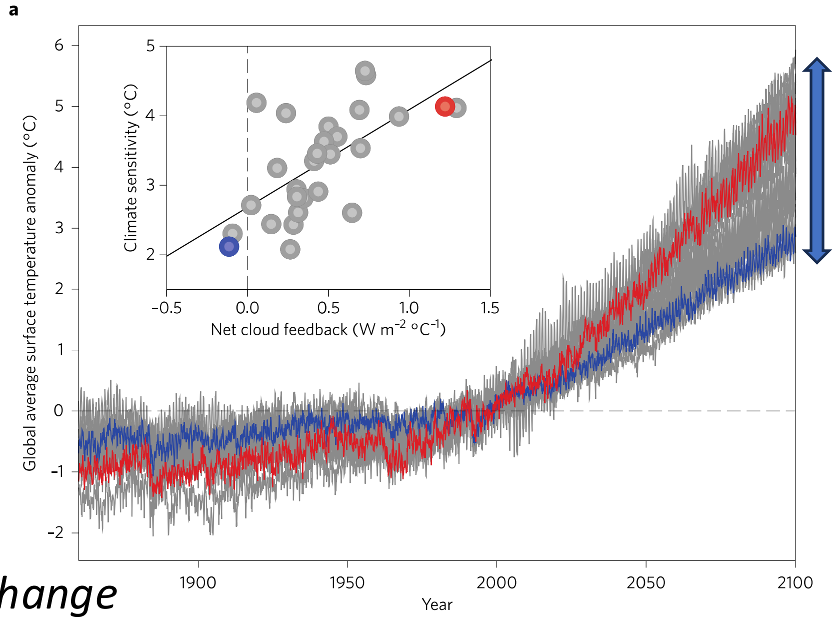
October 8, 2024

Cloud and convection parameterizations limit climate projections





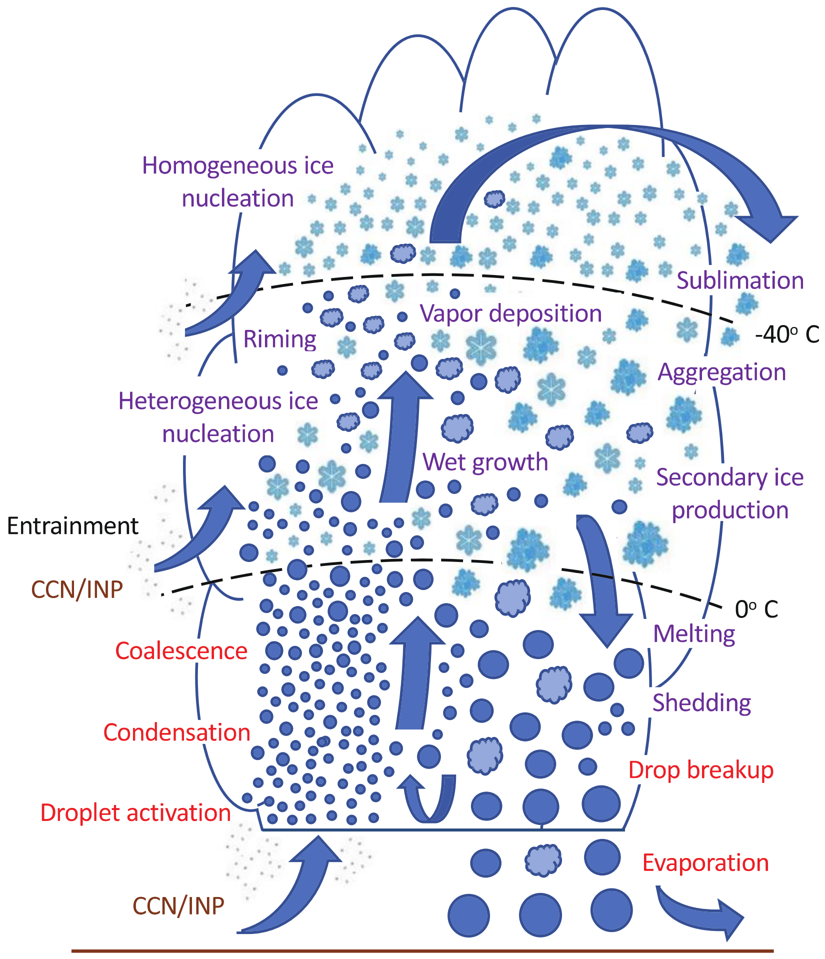
Arakawa, 2004

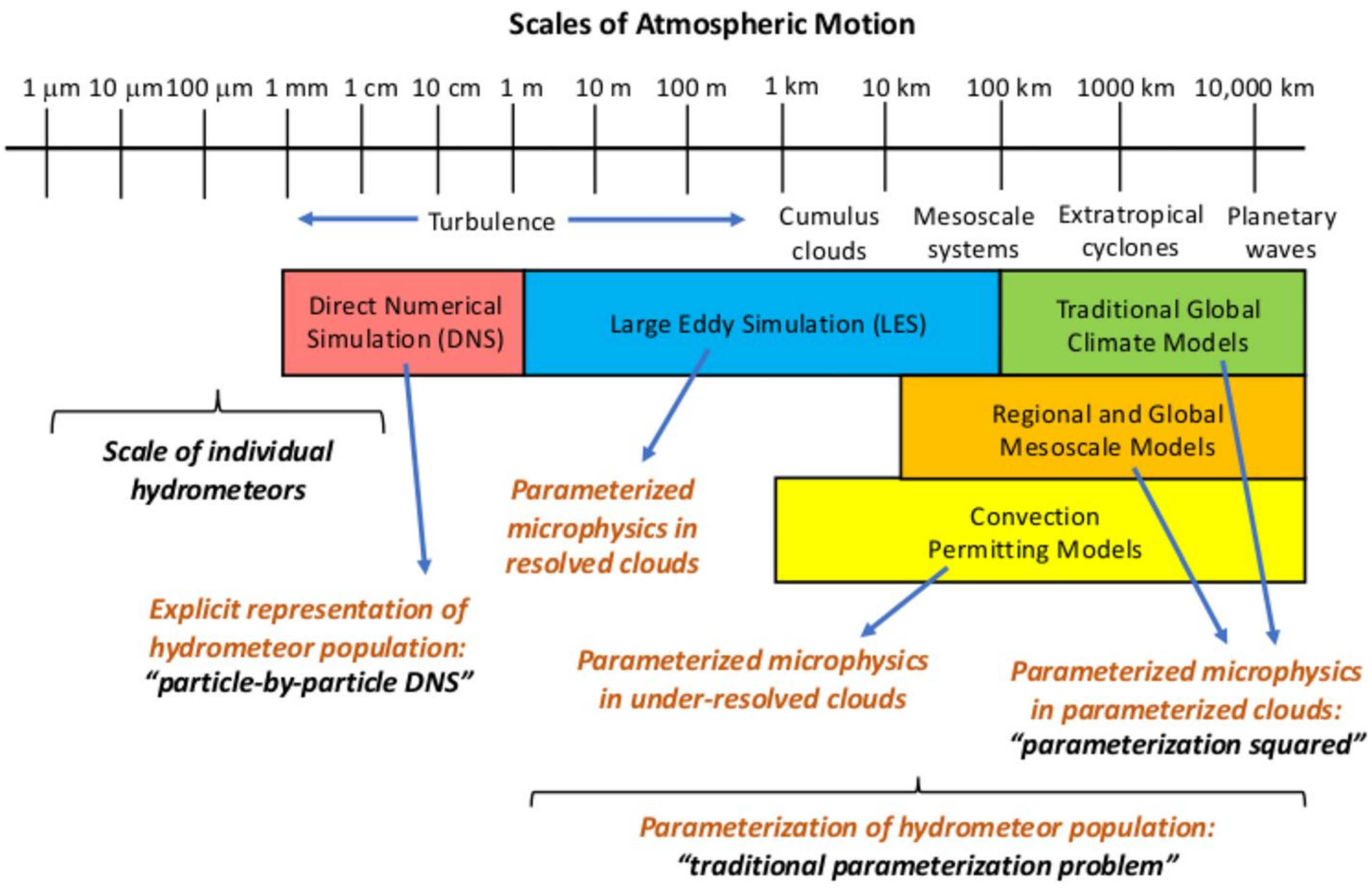


Model spread

Zelinka et al. 2017, *Nature Climate Change*

Cloud processes are complex, non-linear, and multi-scale

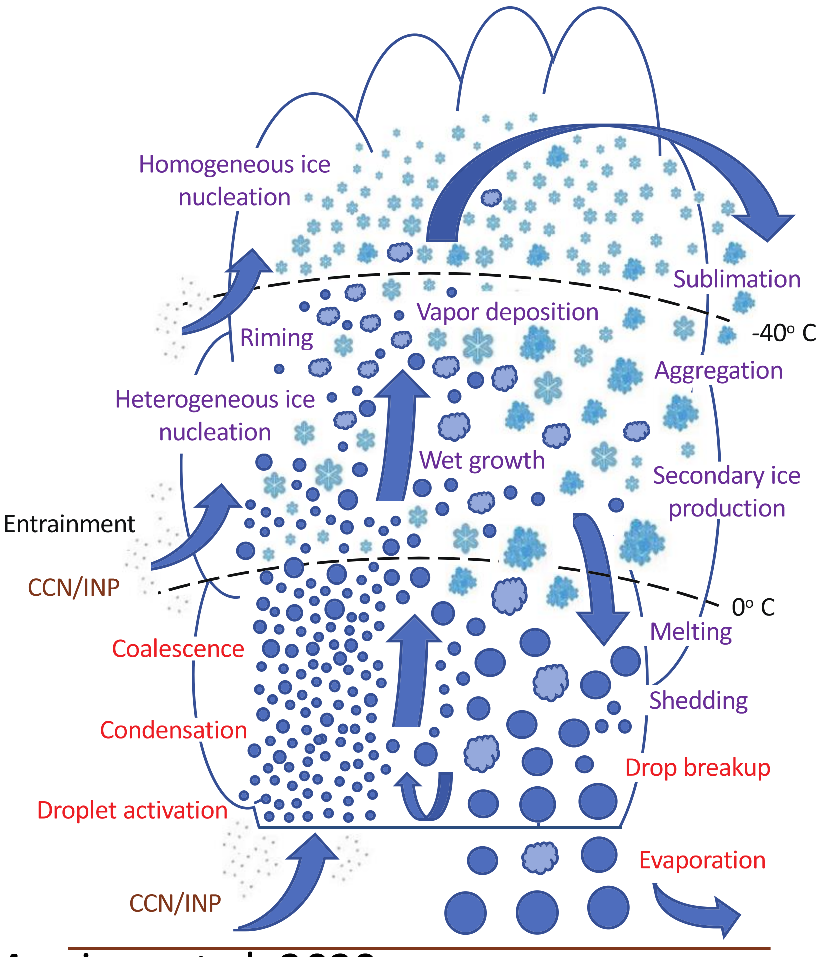




Morrison et al. 2020, JAMES

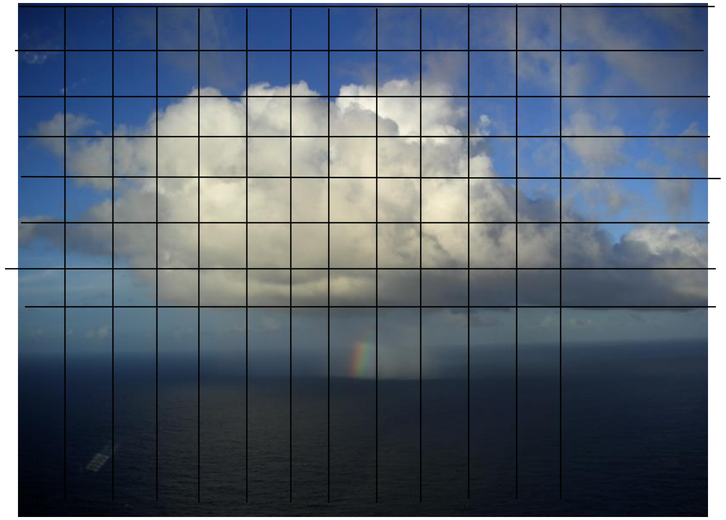
What don't we know about cloud processes?

We don't know all of the physics

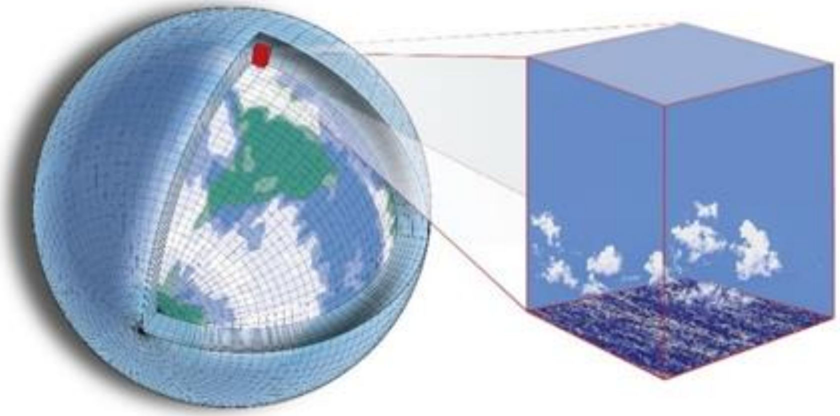


Morrison et al. 2020

We don't know the optimal way to represent cloud processes in models



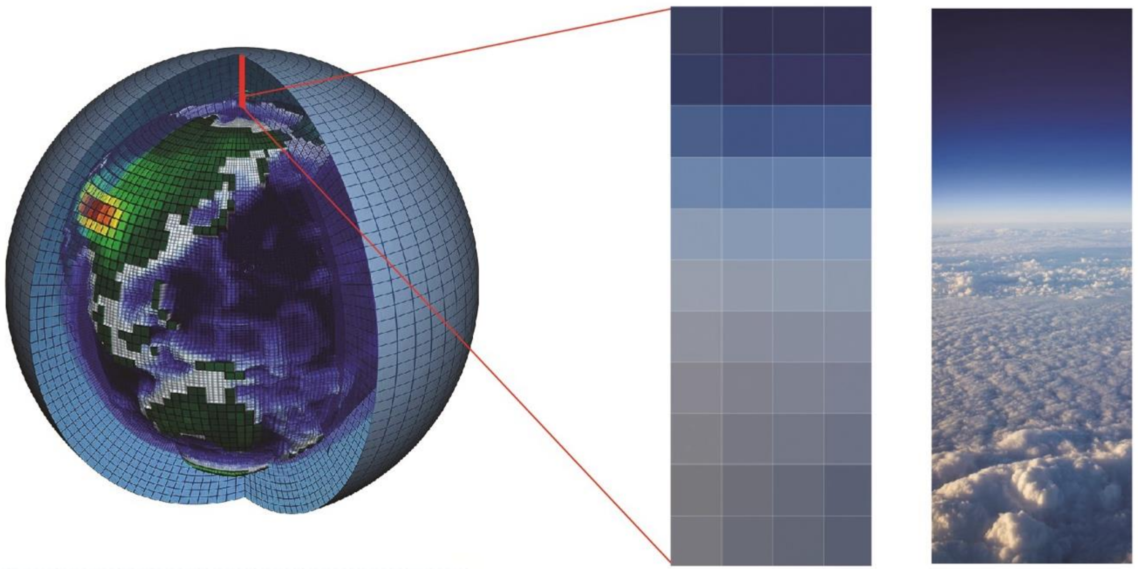
We don't know how to consistently represent these processes at different scales



Schneider et al. 2017

Scientific machine learning can help to solve these challenges

- Emulation of computationally expensive models
- Hybrid-physics machine learning models
- Neural ordinary differential equations
- Data-driven reduced order modeling
- Equation discovery

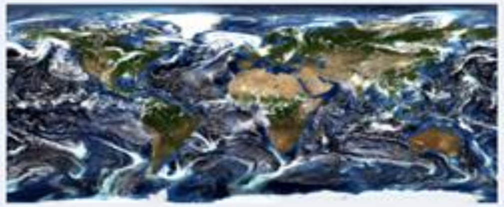


Current

Goal

1. Massive data from Earth observation

2. High-resolution cloud resolving models

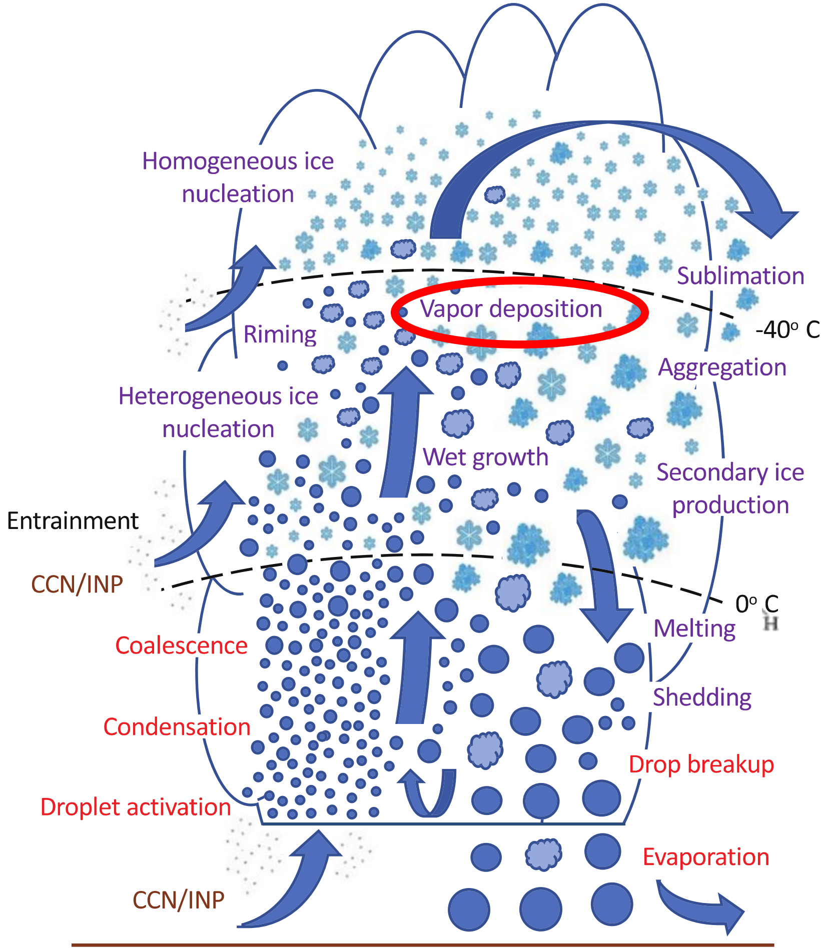


3. Progress in machine learning

Improved Earth System Understanding!

What don't we know about cloud processes?

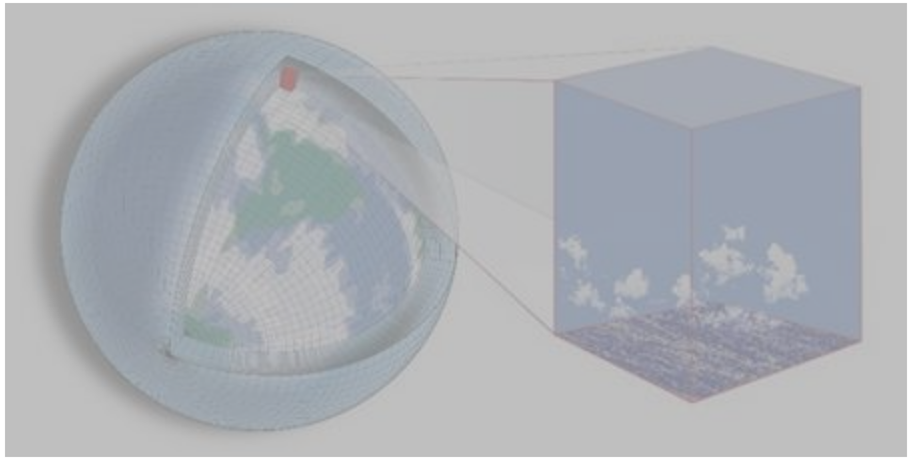
We don't know all of the physics



Morrison et al. 2020

We don't know the optimal way to represent cloud processes in models

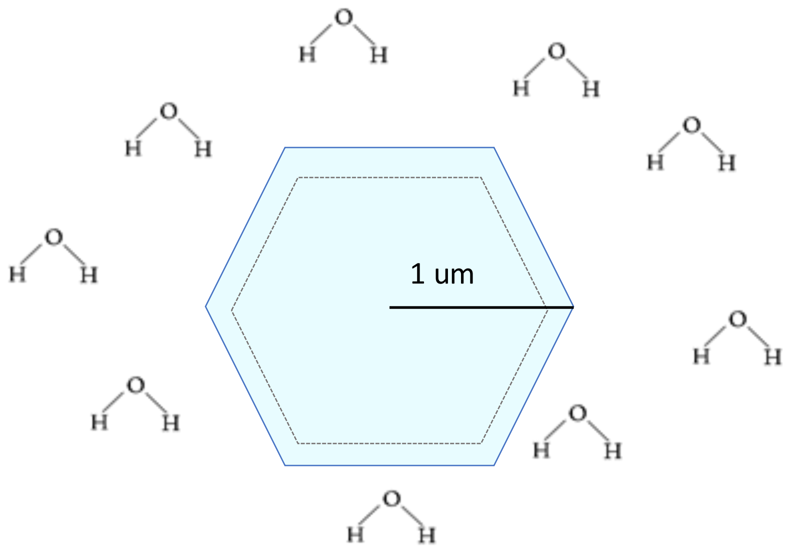
We don't know how to consistently represent these processes at different scales



Schneider et al. 2017

How do ice crystals grow in the atmosphere?

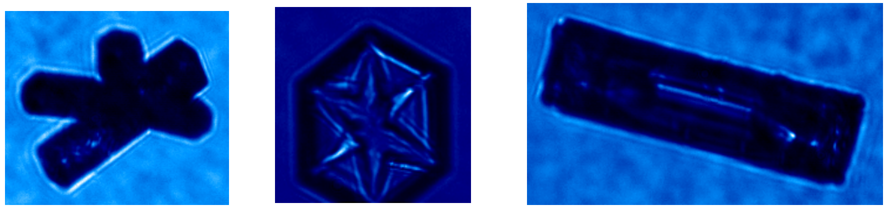
Understanding how ice crystals grow in the atmosphere is fundamental for constraining the radiative effects of clouds, cloud lifetimes, and the distribution of water vapor in the atmosphere– all of which have important climate effects.



$$\frac{dm_p}{dt} = \frac{4\pi C(S_{ice} - 1)}{\frac{RT_g}{\hat{e}_{ice}(T_g)D_w^* M_w} + LH}$$

Single crystal ice mass growth rate

$$D_w^* = \frac{D_w}{\frac{r}{(r+\Delta_v)} + \frac{D_w}{r\boxed{\alpha_D}}\left(\frac{w\pi M_w}{RT_a}\right)^{1/2}}$$



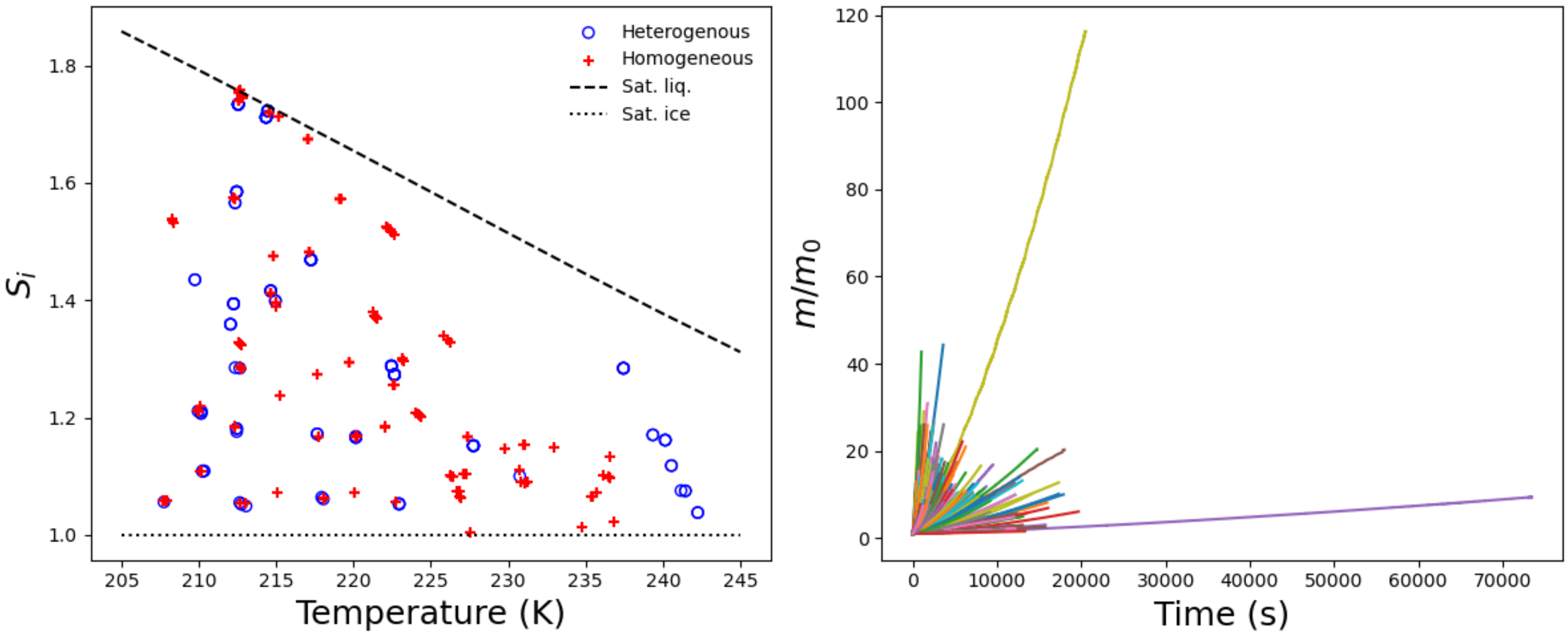
The functional dependence of $\boldsymbol{\alpha}_D$ is uncertain (typically assumed to be a constant value)

Pruppacher and Klett, 1997

Observational data sets to study depositional ice growth

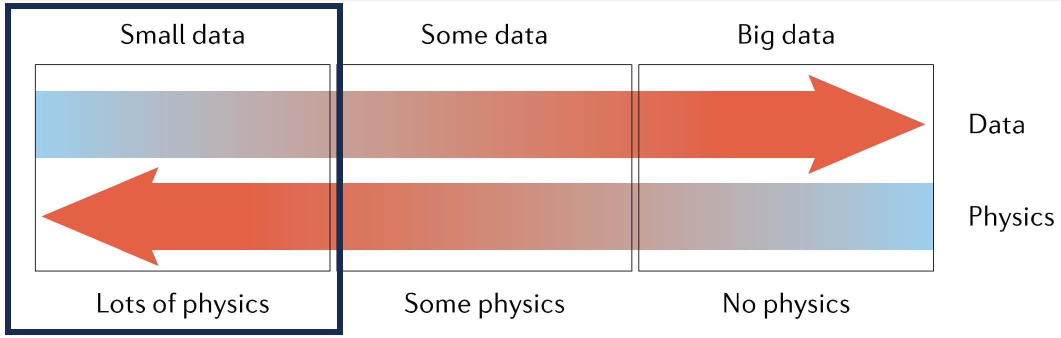
Levitation diffusion chamber experiments

- Single ice crystals, constant temperature and supersaturation
- 307 experiments [Harrison et al. 2016; Pofrika et al. 2020; 2023]



K. Lamb and J. Harrington. “Discovering How Ice Crystals Grow with Neural ODE’s and Symbolic Regression.” Submitted, NeurIPS ML4Physics Workshop.

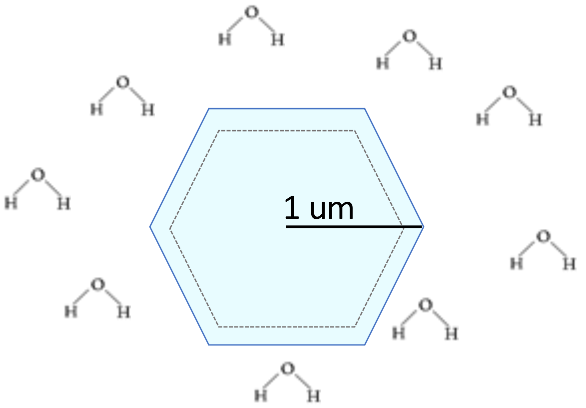
Physics-informed machine learning for depositional ice growth

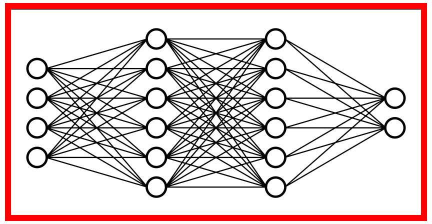


Karniadakis et al. 2021, *Nat. Rev. Phys.*

$$\frac{dm_p}{dt} = \frac{4\pi C(S_{ice} - 1)}{\frac{RT_g}{\hat{e}_{ice}(T_g) D_w^* M_w} + LH}$$

$$D_w^* = \frac{D_w}{\frac{r}{(r+\Delta_v)} + \frac{D_w}{r\boxed{\alpha_D}}\left(\frac{w\pi M_w}{RT_a}\right)^{1/2}}$$

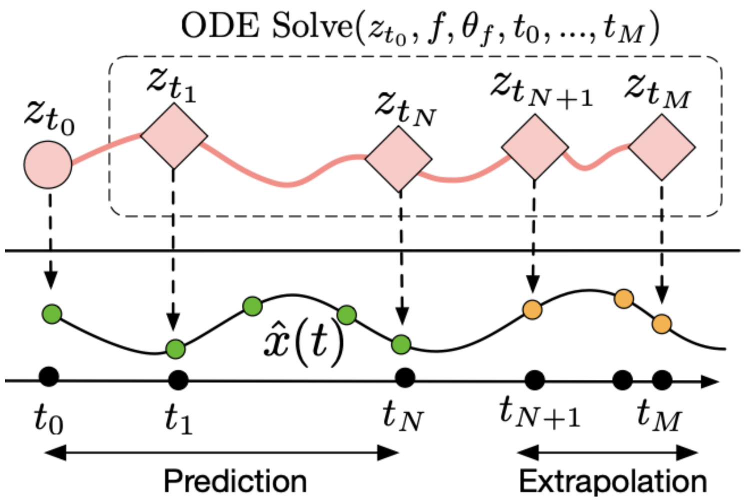




$$\alpha_D = f(S, T|\theta)$$

We want to know the structure of **single particle mass growth rate** but observations provide constraints on **ice mass**

Neural ordinary differential equations (NODEs) perform efficient backpropagation through typical numerical ODE solvers [Chen et al. 2018]



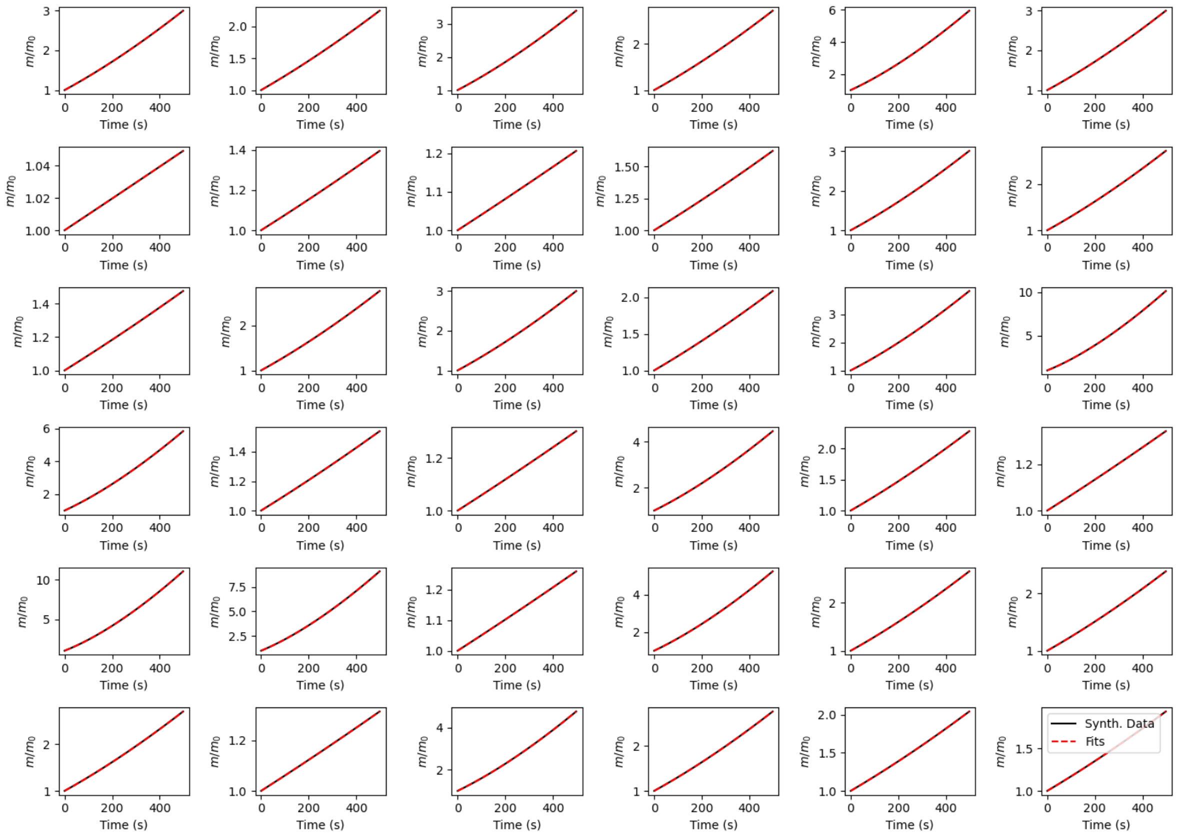
K. Lamb and J. Harrington. "Discovering How Ice Crystals Grow Using Neural ODE's and Symbolic Regression." Submitted, NeurIPS ML4Physics Workshop.

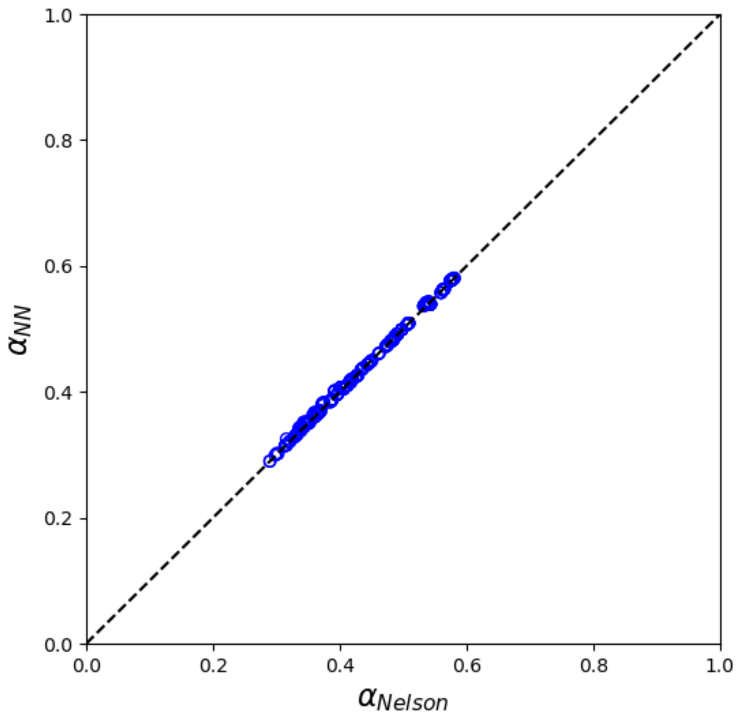
NODE optimized to ice mass growth rates (synthetic data w/ *known* functional form)

Optimize NODE model against 307 **synthetic** ice mass time series to learn the functional dependence of α_D
→ Minimize distance between model and observations

$$\frac{dm_p}{dt} = \frac{4\pi C(S_{ice} - 1)}{\frac{RT_g}{\hat{e}_{ice}(T_g) D_w^* M_w} + LH}$$

$$D_w^* = f(D_w, r, \boxed{\alpha_D})$$





K. Lamb and J. Harrington. "Discovering How Ice Crystals Grow Using Neural ODE's and Symbolic Regression." Submitted, NeurIPS ML4Physics Workshop.

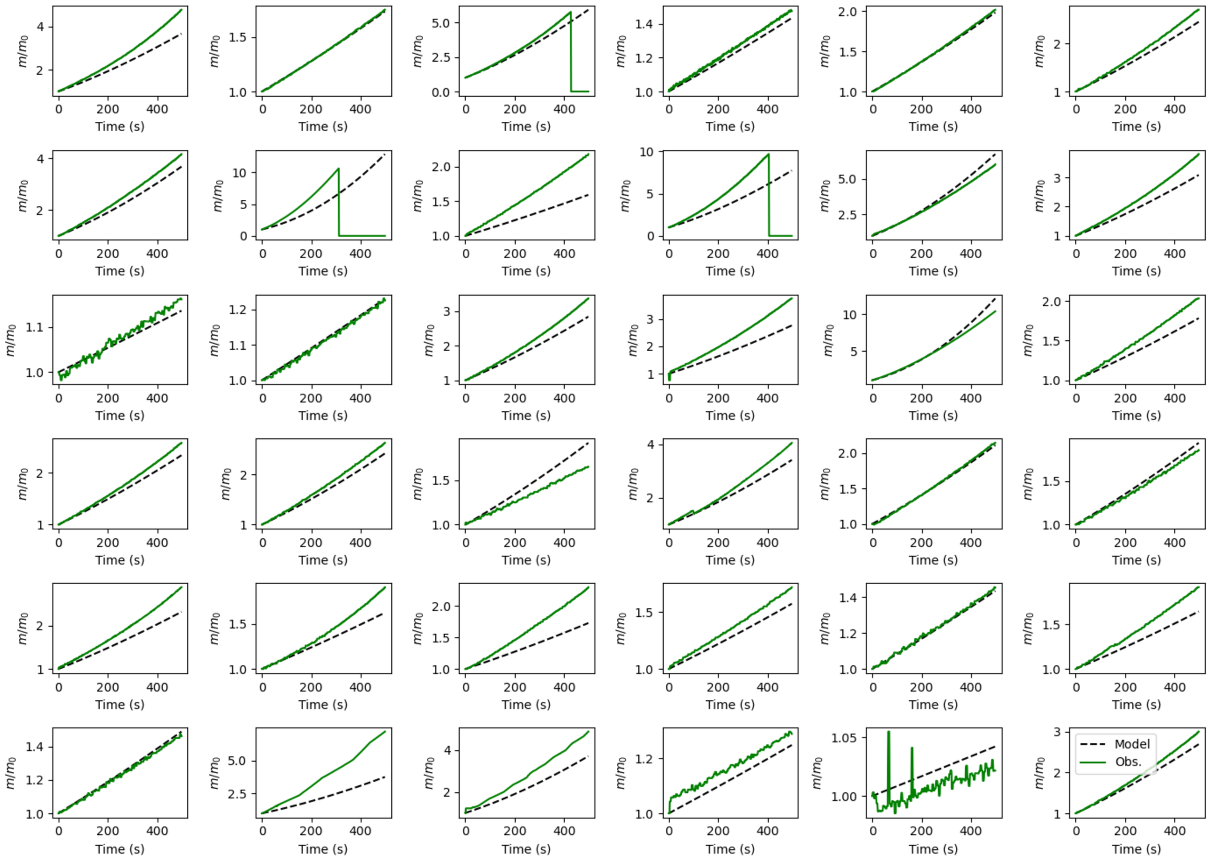
NODE optimized to ice mass growth rates (real data w/ *unknown* functional form)

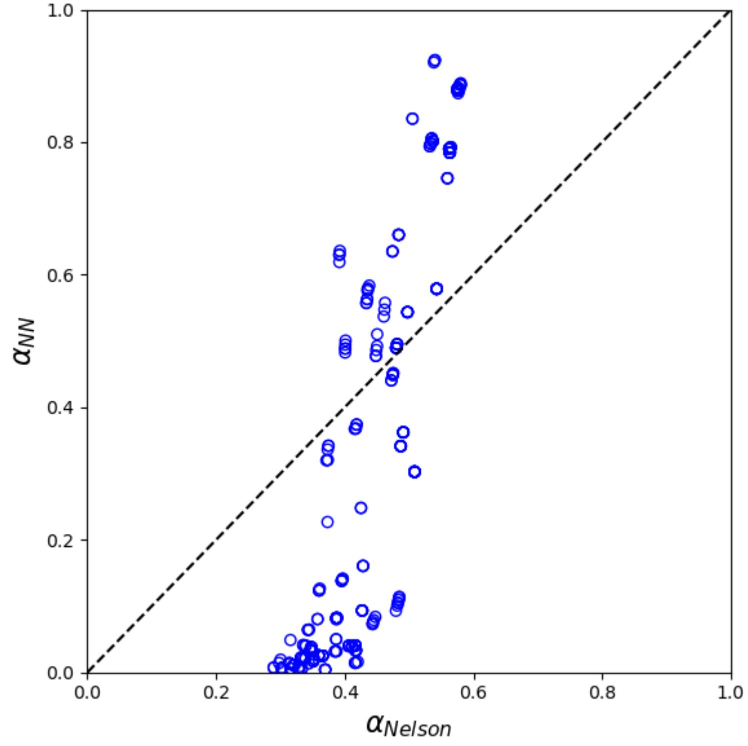
Optimize NODE model against 307 **real** ice mass time series to learn the functional dependence of $\boldsymbol{\alpha_D}$

→ Minimize distance between model and observations

$$\frac{dm_p}{dt} = \frac{4\pi C(S_{ice} - 1)}{\frac{RT_g}{\hat{e}_{ice}(T_g) D^* M} + LH}$$

$$D^* = f(D_w, r, \boxed{\alpha_D})$$





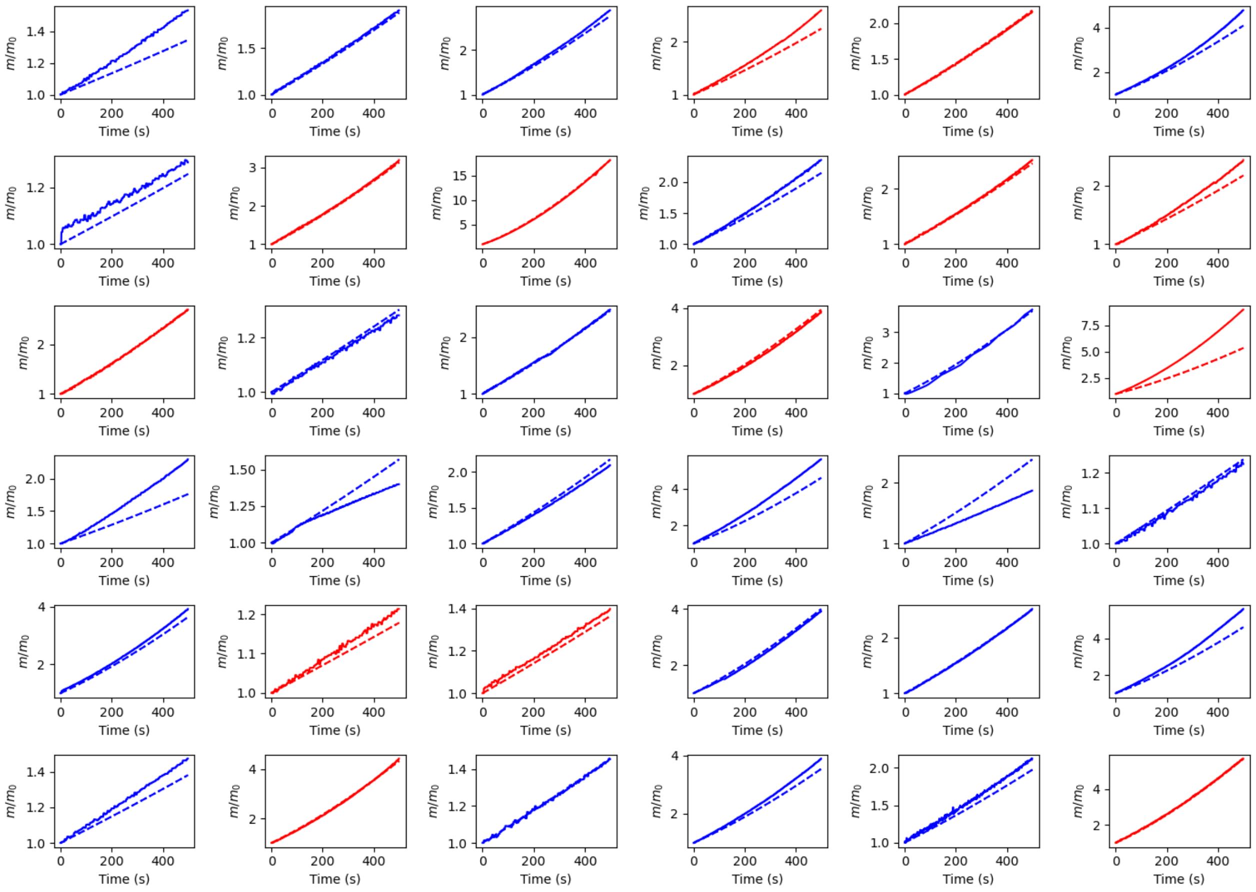
K. Lamb and J. Harrington. "Discovering How Ice Crystals Grow Using Neural ODE's and Symbolic Regression." Submitted, NeurIPS ML4Physics Workshop.

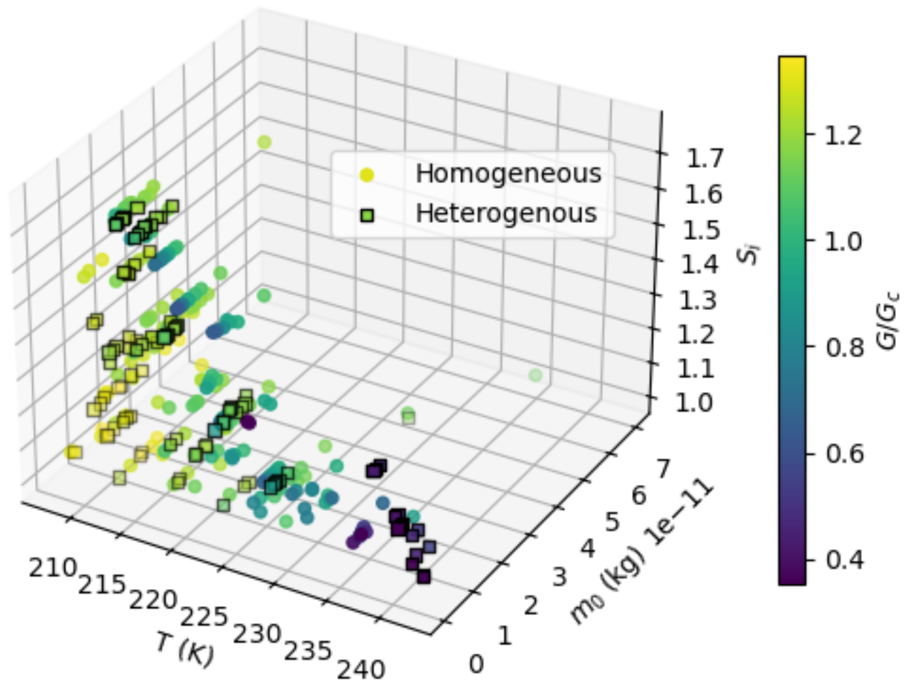
NODE optimized to ice mass growth rates (real data w/ *unknown* functional form)

Optimize NODE model against 307 **real** ice mass time series to learn the functional dependence of G/G_c
→ Minimize distance between model and observations

$$\frac{dm}{dt} = 4\pi r(S_{ice} - 1)G$$

$$G = \boxed{f(m, T, S_i|\theta)}G_c$$



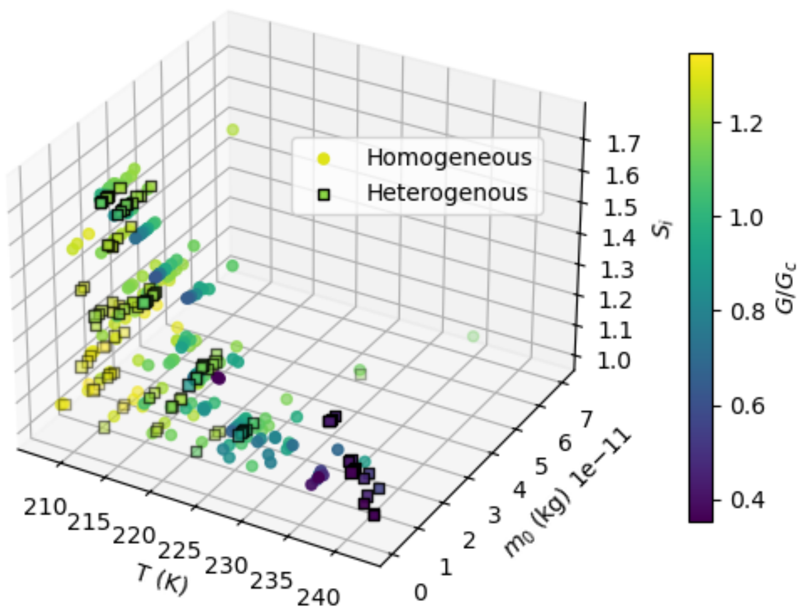


K. Lamb and J. Harrington. "Discovering How Ice Crystals Grow Using Neural ODE's and Symbolic Regression." Submitted, NeurIPS ML4Physics Workshop.

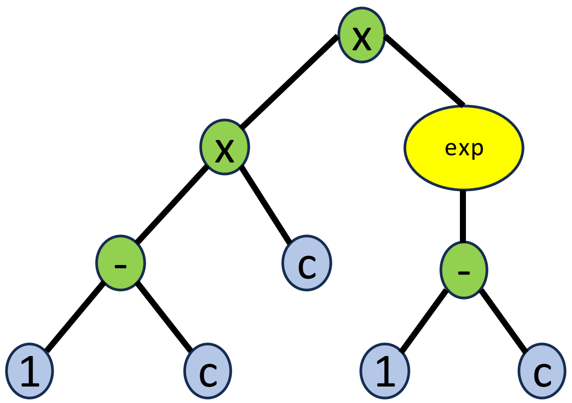
Learned functional dependence for ice growth using symbolic regression

$$\frac{dm}{dt} = 4\pi r(S_{ice} - 1)G$$

$$G = \boxed{f(m, T, S_i|\theta)}G_c$$



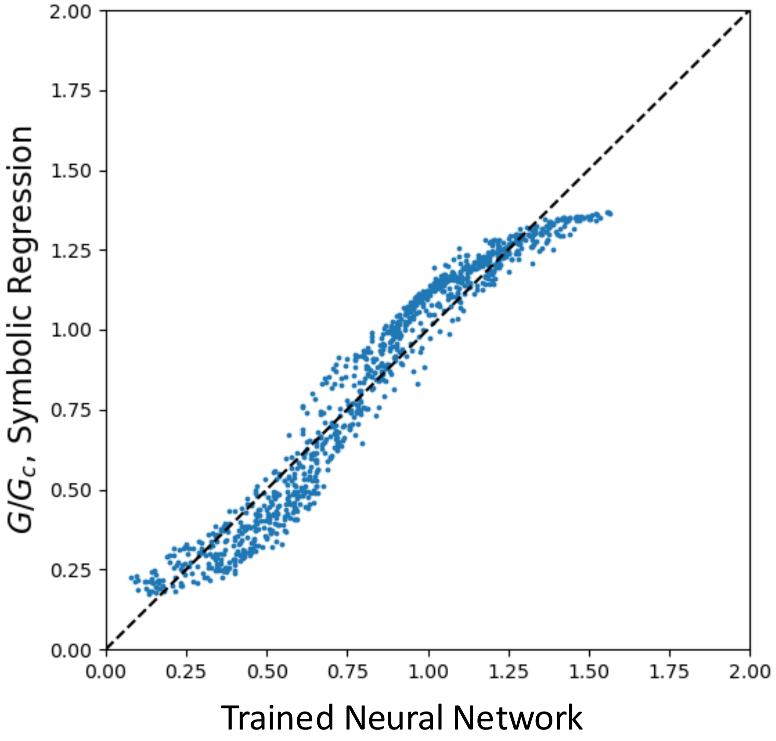
PySR library [Cranmer, 2023]



Learned functional form:

$$\boxed{\frac{G}{G_c} = |ax_0 + exp(-2x_1(x_1^3 + x_2))|}$$

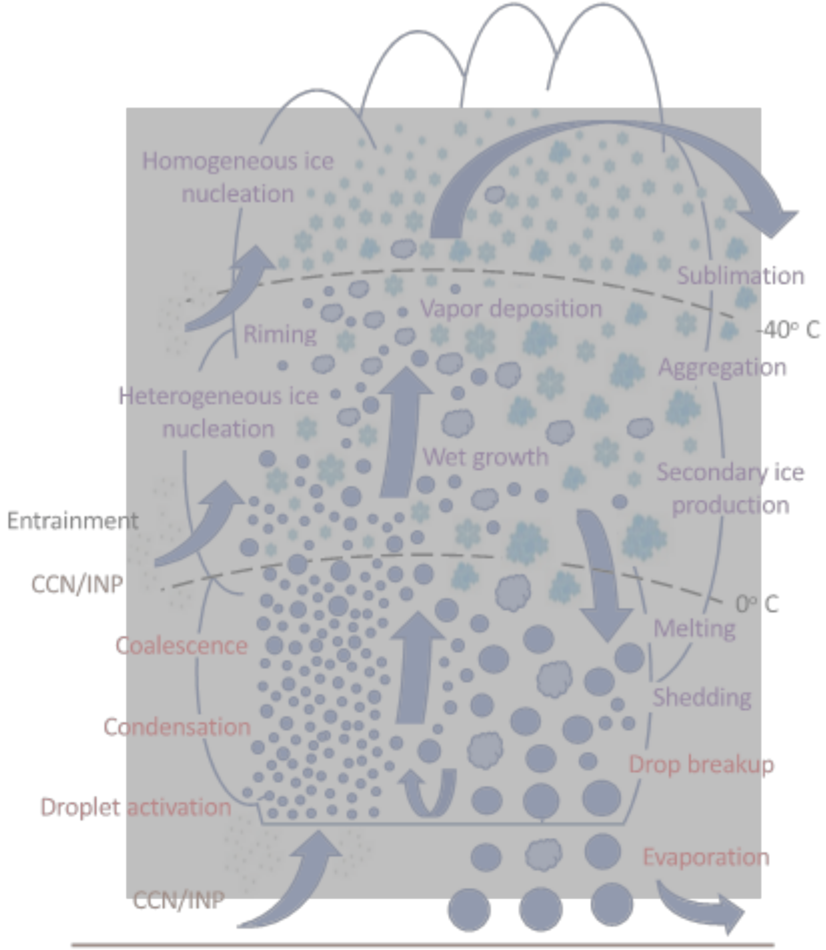
$$a = 0.591, (m \to x_0, T \to x_1, S_i \to x_2)$$



K. Lamb and J. Harrington. "Discovering How Ice Crystals Grow Using Neural ODE's and Symbolic Regression." Submitted, NeurIPS ML4Physics Workshop.

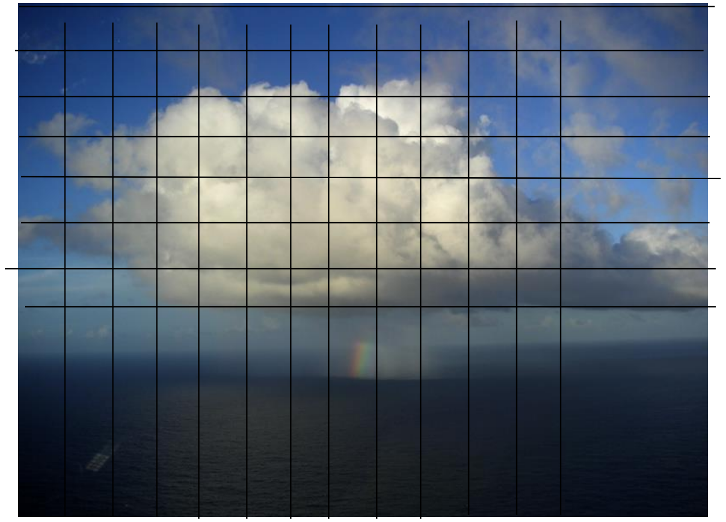
What don’t we know about cloud processes?

We don’t know all of the physics

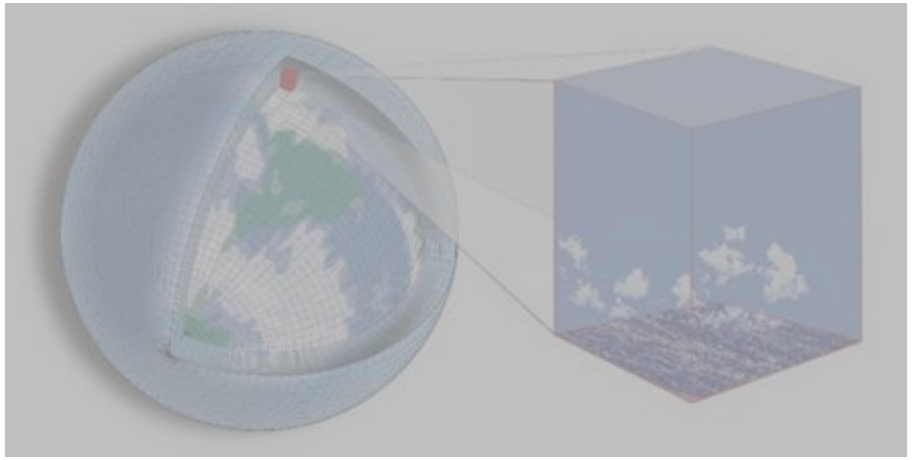


Morrison et al. 2020, JAMES

We don’t know how to represent these processes in models



We don’t know how to consistently represent these processes at different scales

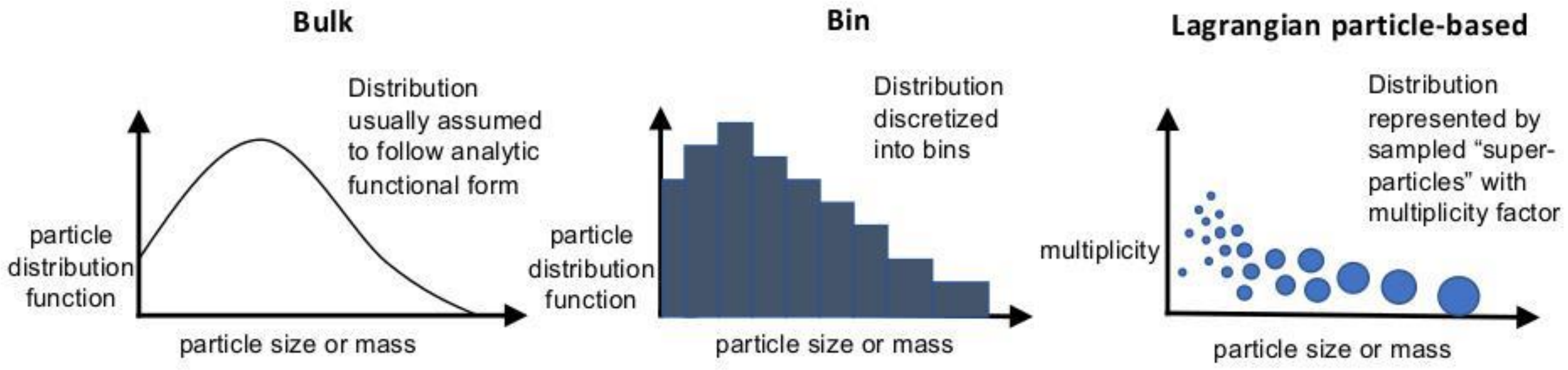


Schneider et al. 2017

How can we efficiently represent a spectrum of droplets in models?

- Droplet size distribution (DSD) – spectrum of droplets of various sizes in a cloud
- Details are important for cloud radiative effects and the initiation and timing of precipitation

4-6 parameters | ~30-70 parameters | ~100's parameters



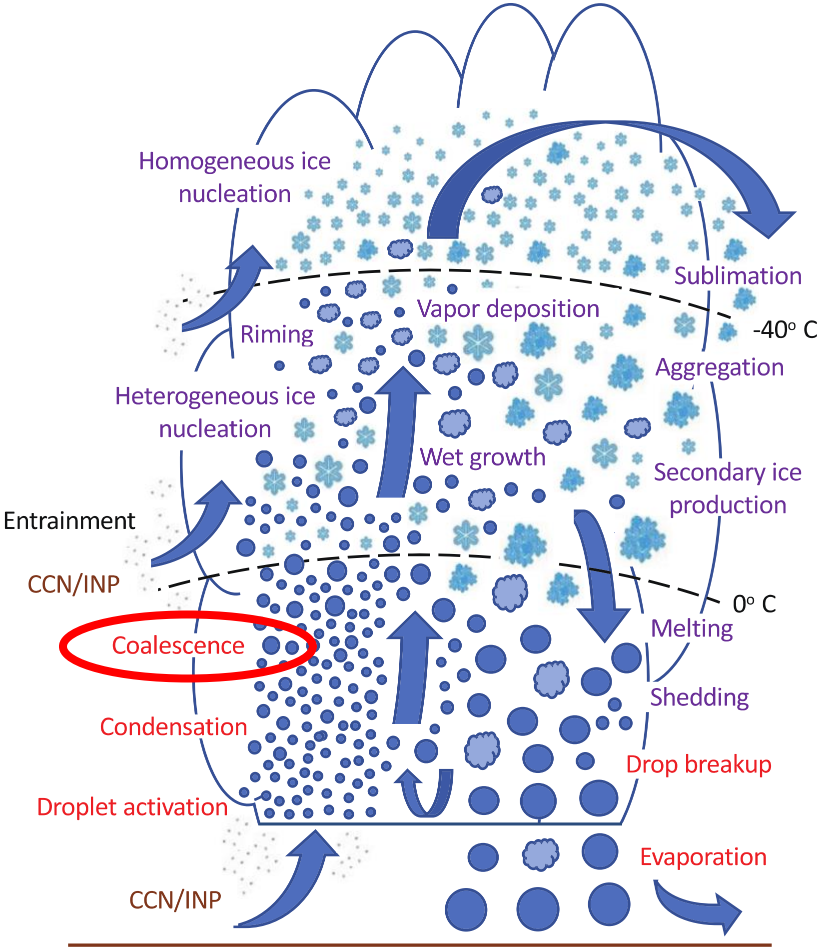
Computational expense
Fidelity

Climate models → *Process level models*

Morrison et al. 2020

Typical microphysical process rate representation in bulk schemes

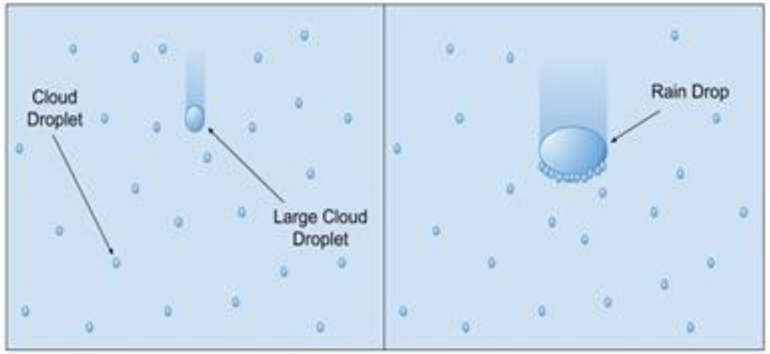
Microphysical process rates

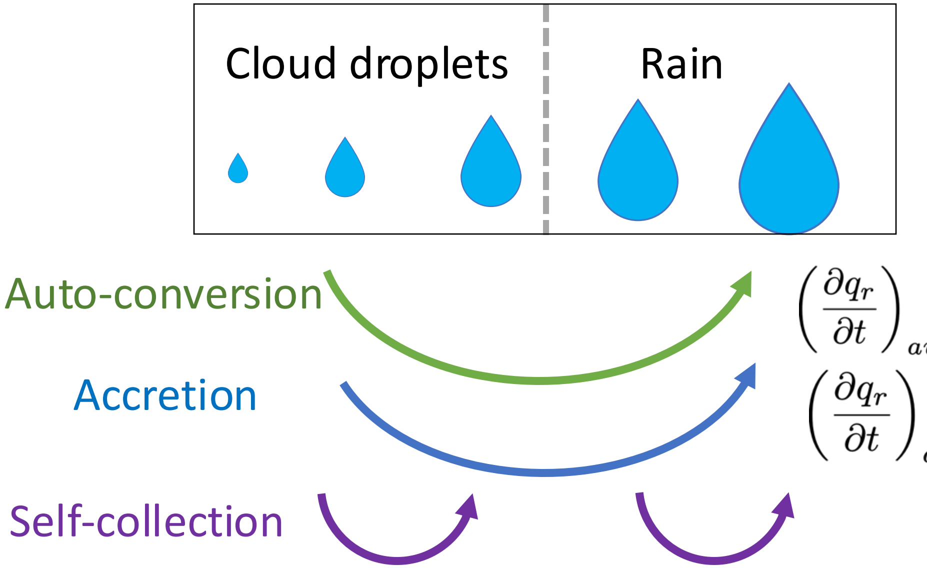


Morrison et al. 2020

Typical bulk microphysical process rates:

- power laws of cloud and rain moments





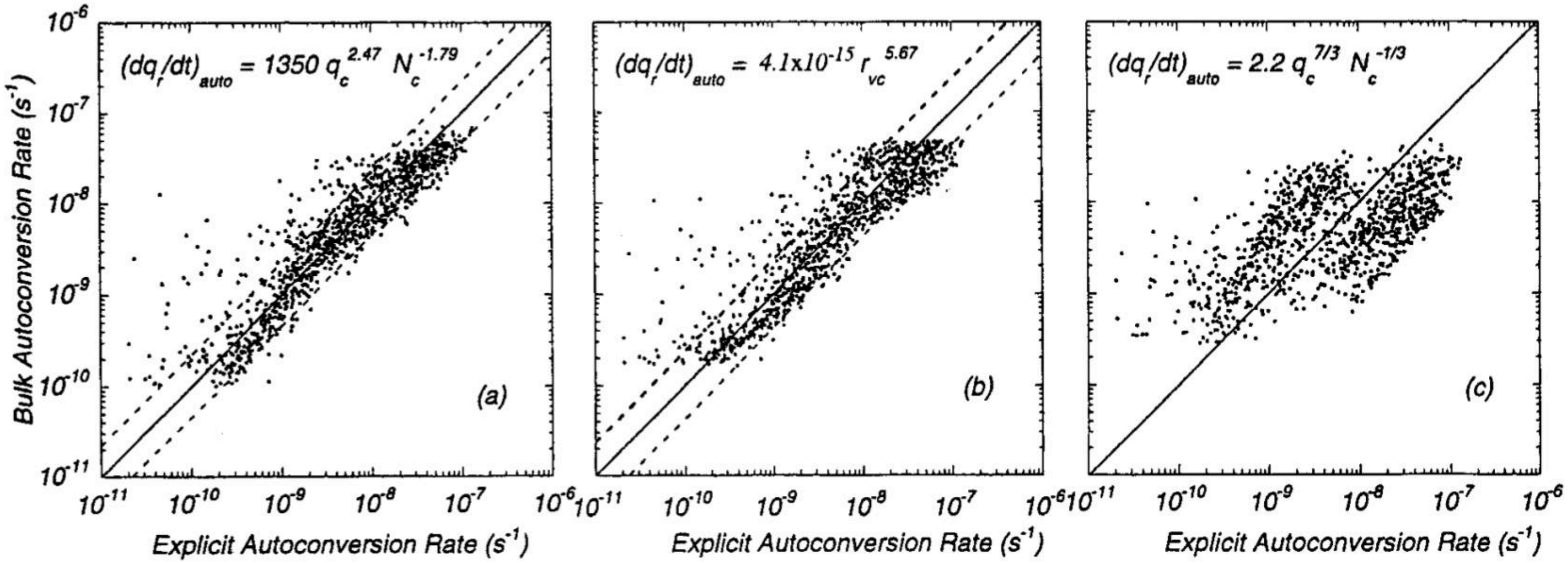
$$\left(\frac{\partial q_r}{\partial t}\right)_{auto} = 1350 q_c^{2.47} N_c^{-1.79}$$

$$\left(\frac{\partial q_r}{\partial t}\right)_{accr} = 67 (q_c q_r)^{1.15}$$

Khairoutdinov and Kogan, 2000

Typical approach to develop a bulk cloud microphysics scheme

Assume some functional form for bulk microphysical process rates, use higher fidelity (bin or superdroplet) simulations as the ground truth, fit best parameters



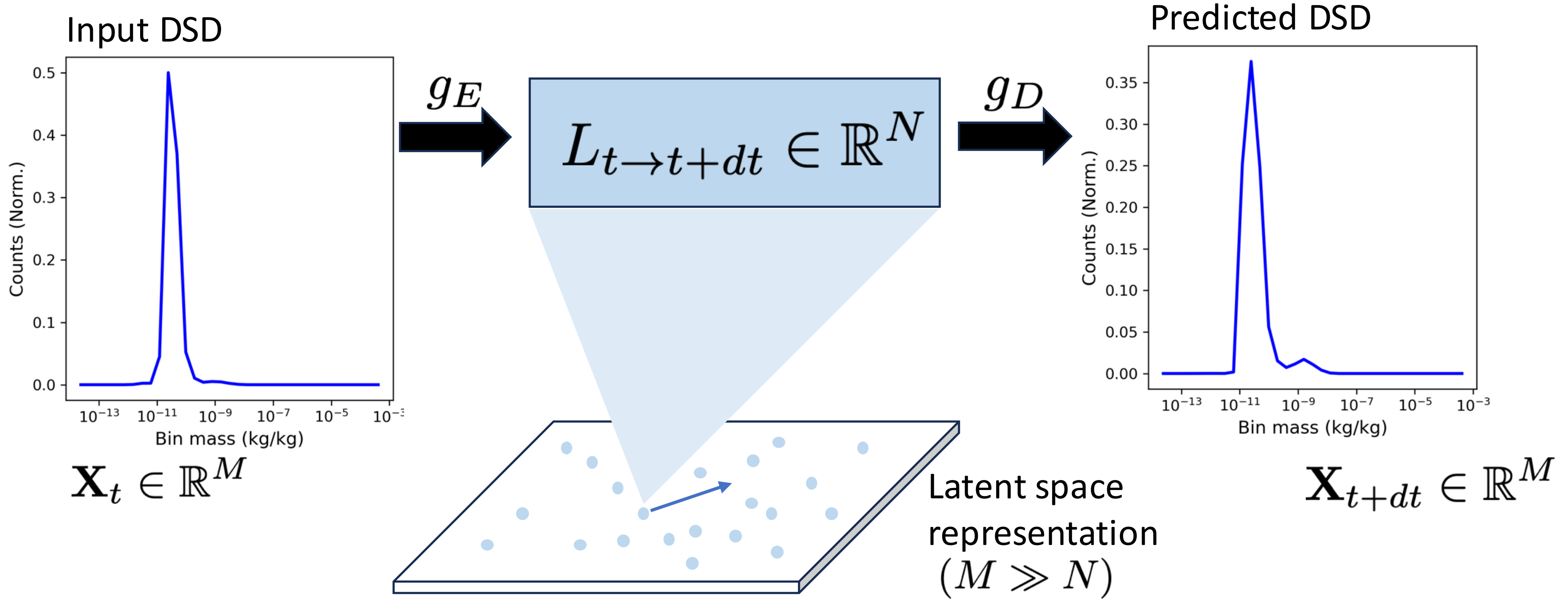
Limitations:

- *A priori* DSD representation (cloud and rain moments)
- Microphysical process rates with assumed functional forms

Structural and parametric uncertainty (artificial threshold between cloud and rain is known to be problematic)

Khairoutdinov and Kogan, 2000

Our approach: reduced order modeling to learn latent "bulk" scheme



Apply ROM to model with higher fidelity (bin or superdroplet) microphysics scheme

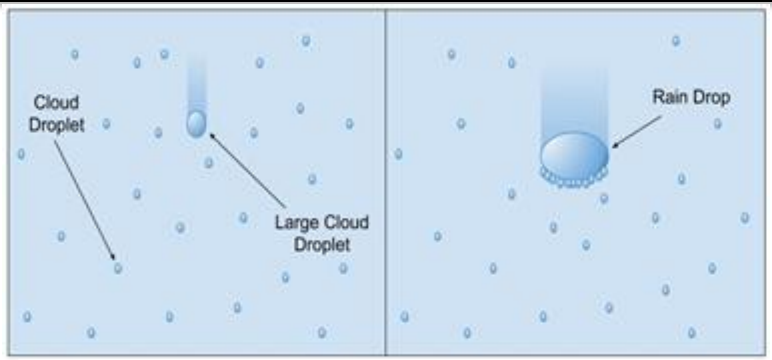
- Reduced order ("latent") representation of the DSD
- Microphysical process rates acting on latent representation of DSD ("latent dynamics")

Ground truth data: TAU bin microphysics scheme [Tzivion et al. 1987, 1989; Feingold et al. 1988].

- 1D kinematic driver model, simulating a column of air under action of sinusoidal updraft
- 16 cases with different initial conditions for aerosol concentration, updraft speeds (~231000 samples)

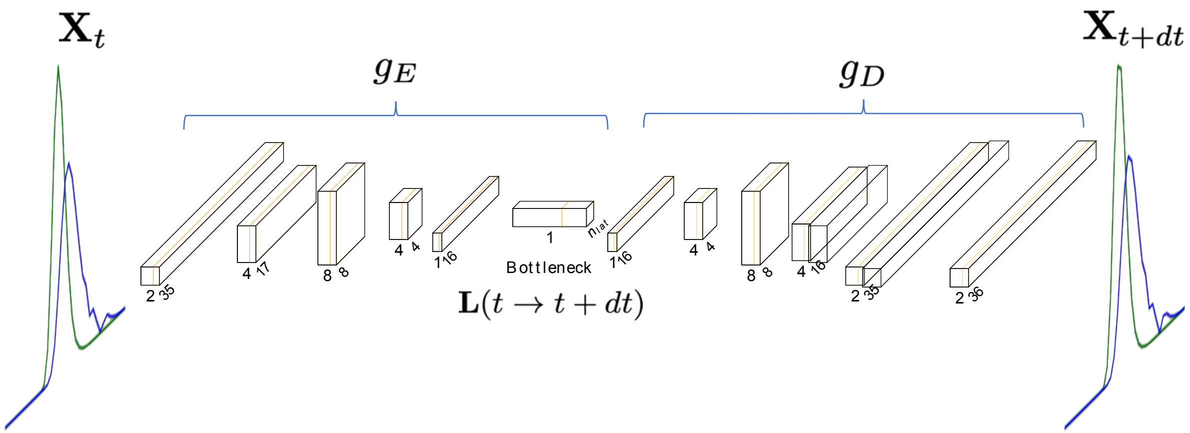
How many variables do we need to represent collision-coalescence?

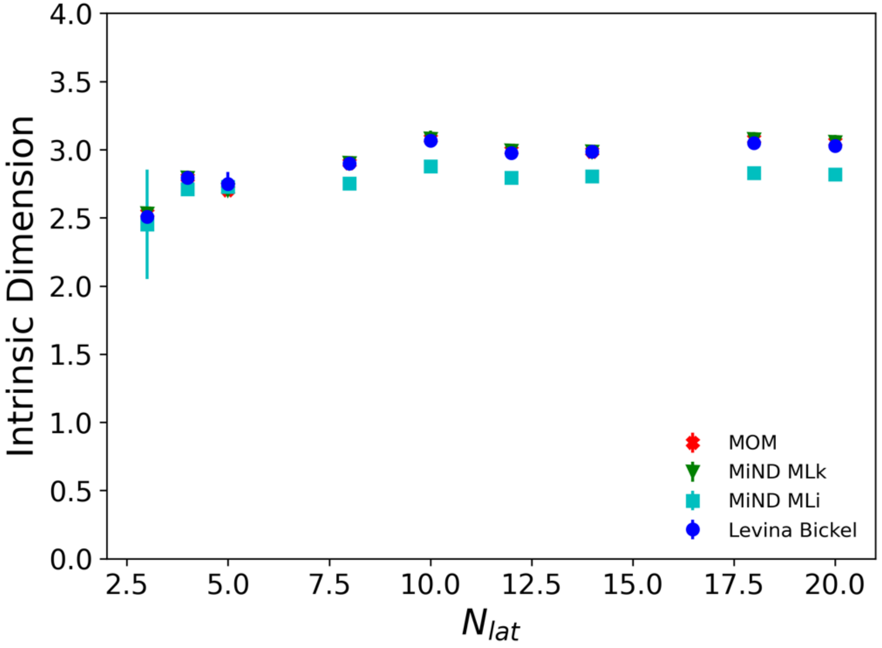
How many degrees of freedom are needed to accurately predict the future state of a physical system? [Chen et al. 2022, *Nat. Comp. Sci.*]



Intrinsic dimension = topological dimension of lower dimensional "latent" manifold approximates # of independent degrees of freedom needed

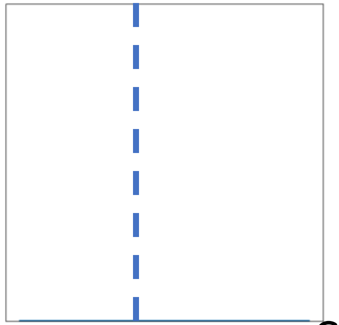
ID_{cc} = 3 (the minimum number of independent variables needed to parameterize collision coalescence)

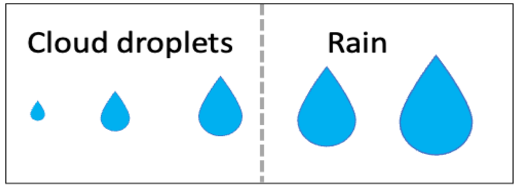




K.D. Lamb, M. van Lier Walqui, S. Santos, H. Morrison. "Reduced Order Modeling for Linearized Representations of Microphysical Process Rates, *JAMES*.

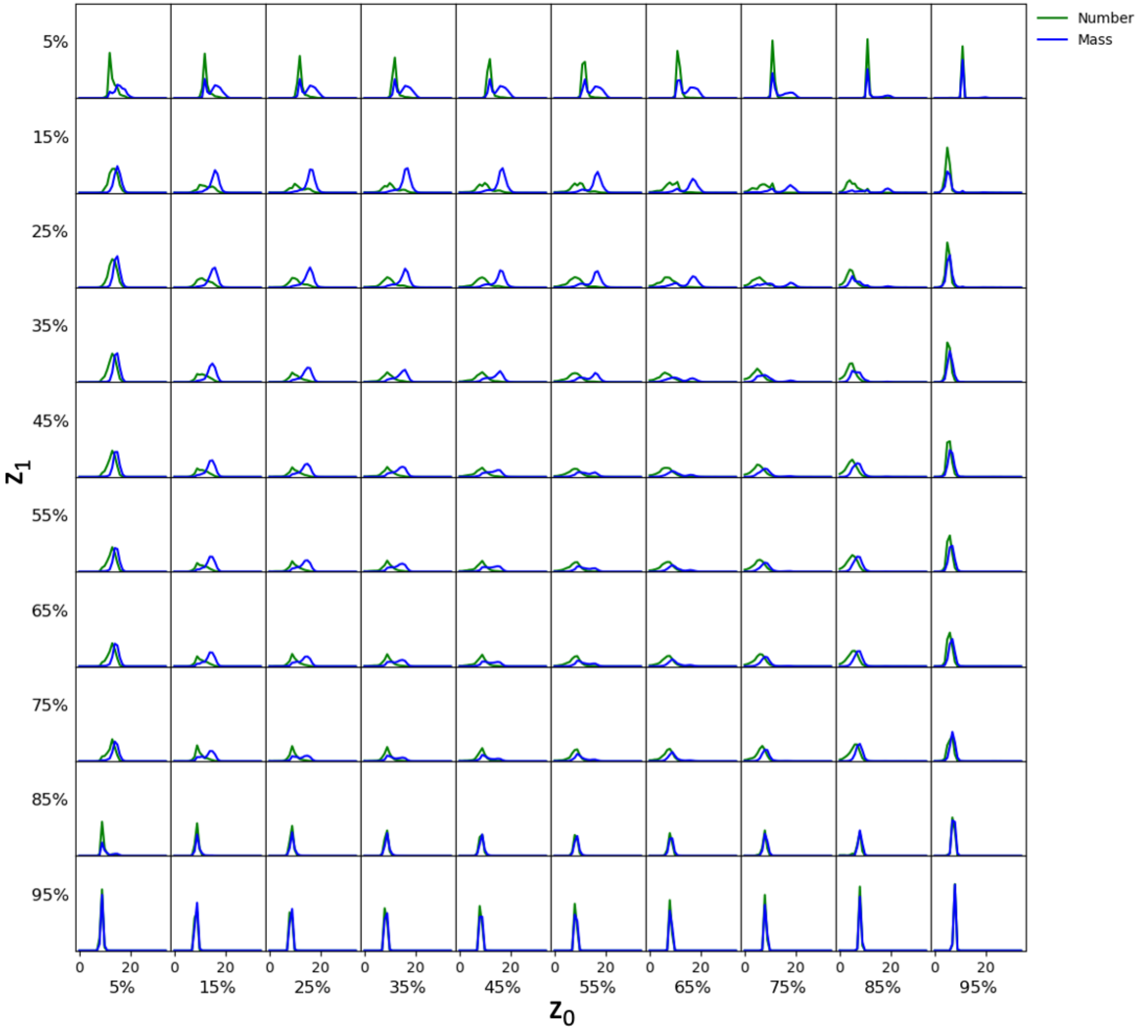
What do the latent variables represent in terms of the DSD?



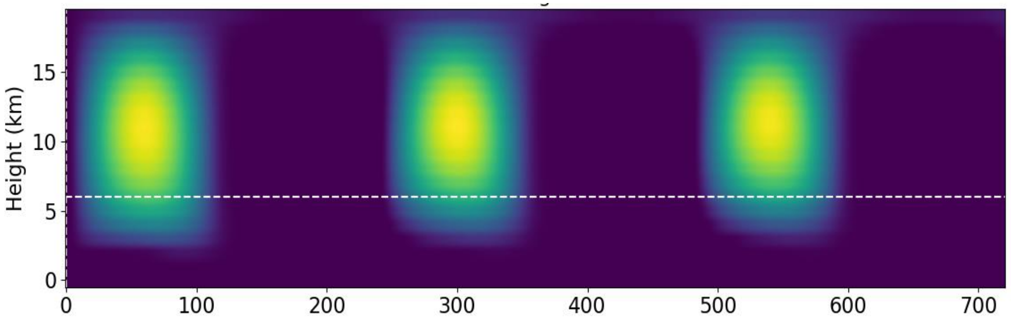


q_c, N_c $\quad$ q_r, N_r

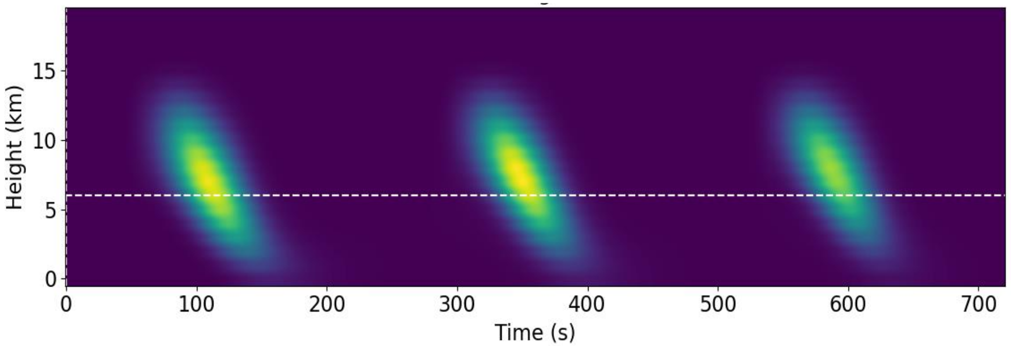
Latent variables



Cloud moments (q_c)



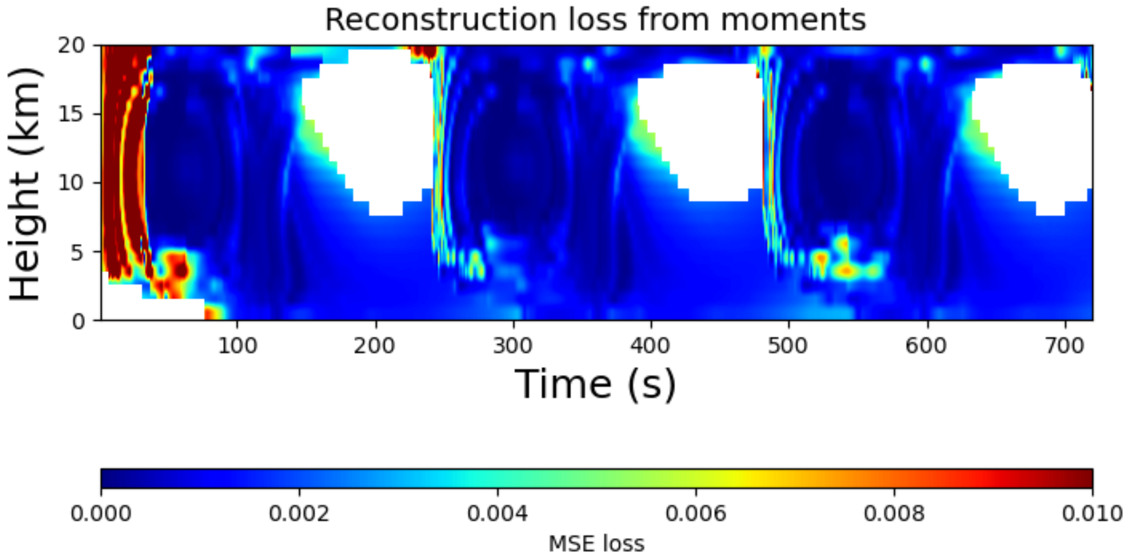
Rain moments (q_r)

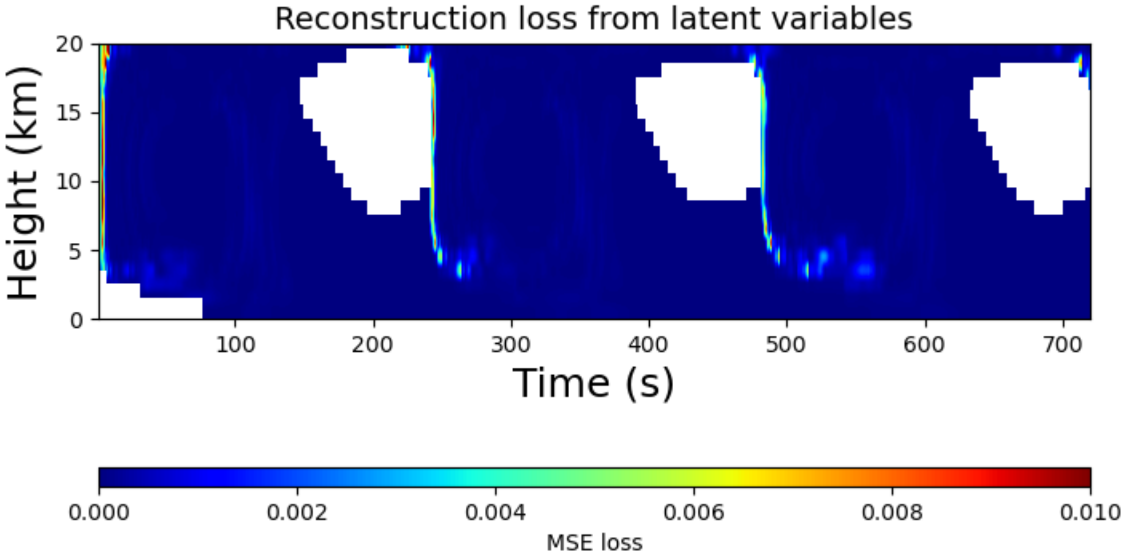


K.D. Lamb, M. van Lier Walqui, S. Santos, H. Morrison. "Reduced Order Modeling for Linearized Representations of Microphysical Process Rates, *JAMES*.

Where is microphysical information lost in the cloud?

- Train decoder to reconstruct DSD from typical bulk representations of the DSD (cloud and rain moments)
- Compare with reconstructions for latent variables from VAE
- Latent variables capture early stages of rain formation better than moments
- Latent variables more accurately reconstruct the DSD than moments





Average metrics for 10 random initializations of models evaluated on test data sets:

We don't necessarily need more complexity in our models!

Representation	Variables	MSE reconstruction loss
Latent variables	L_0, L_1, L_2	**2.6e-4 ± 5.1e-4**
Typical bulk variables	q_c, q_r, N_c, N_r	6.0e-4 ± 13.5e-4

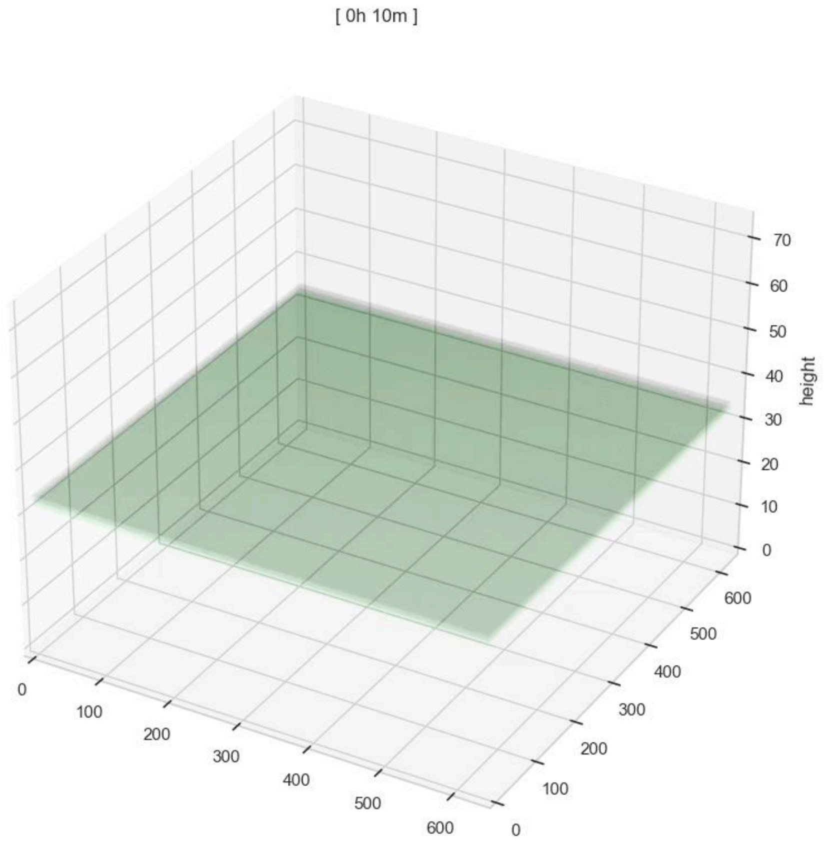
K.D. Lamb, M. van Lier Walqui, S. Santos, H. Morrison. "Reduced Order Modeling for Linearized Representations of Microphysical Process Rates, *JAMES*.

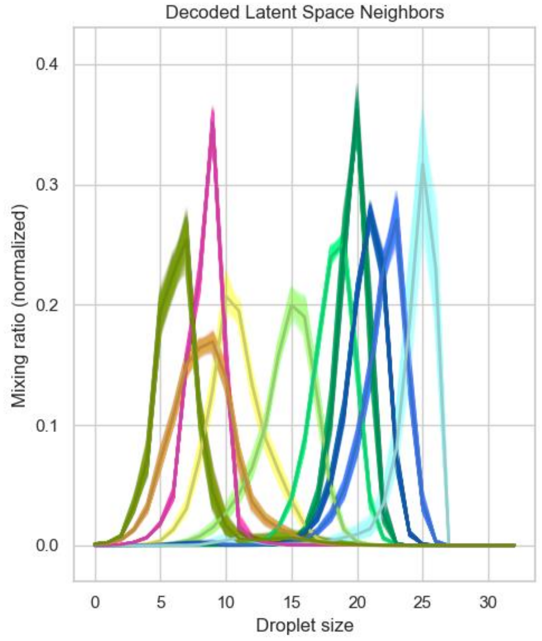
What can dimensionality reduction tell us about the evolution of clouds?

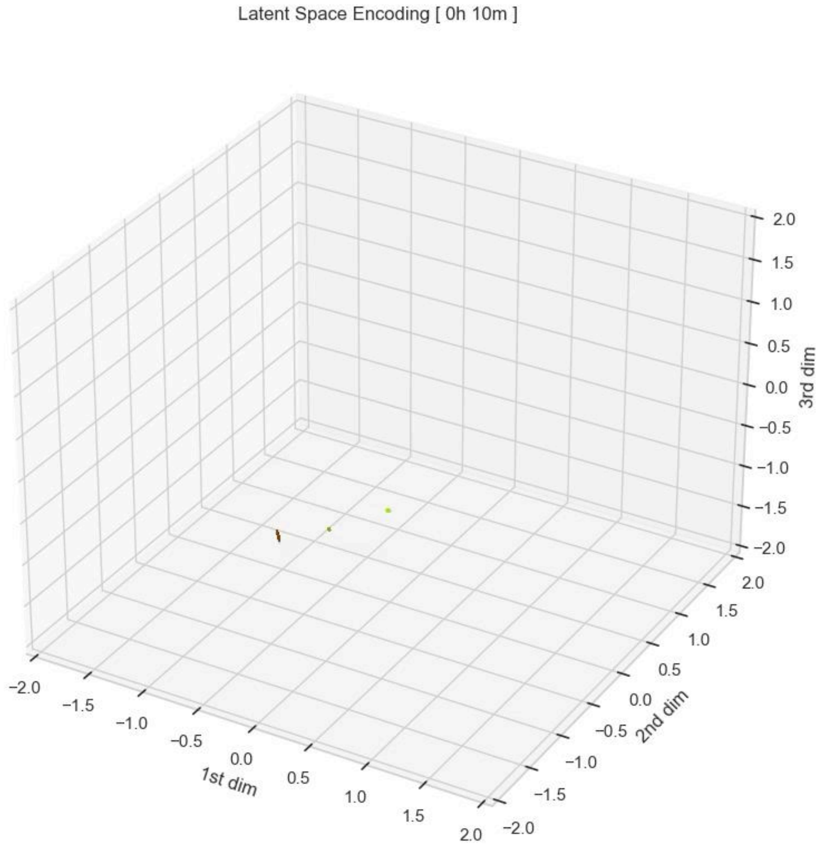
In real clouds, local droplet size distributions differ substantially from mean across the cloud [Allwayin et al. 2024]

Eagles LES simulations of stratocumulus cases with bin microphysics scheme

Color visualizes 3 latent variables projected onto RGB



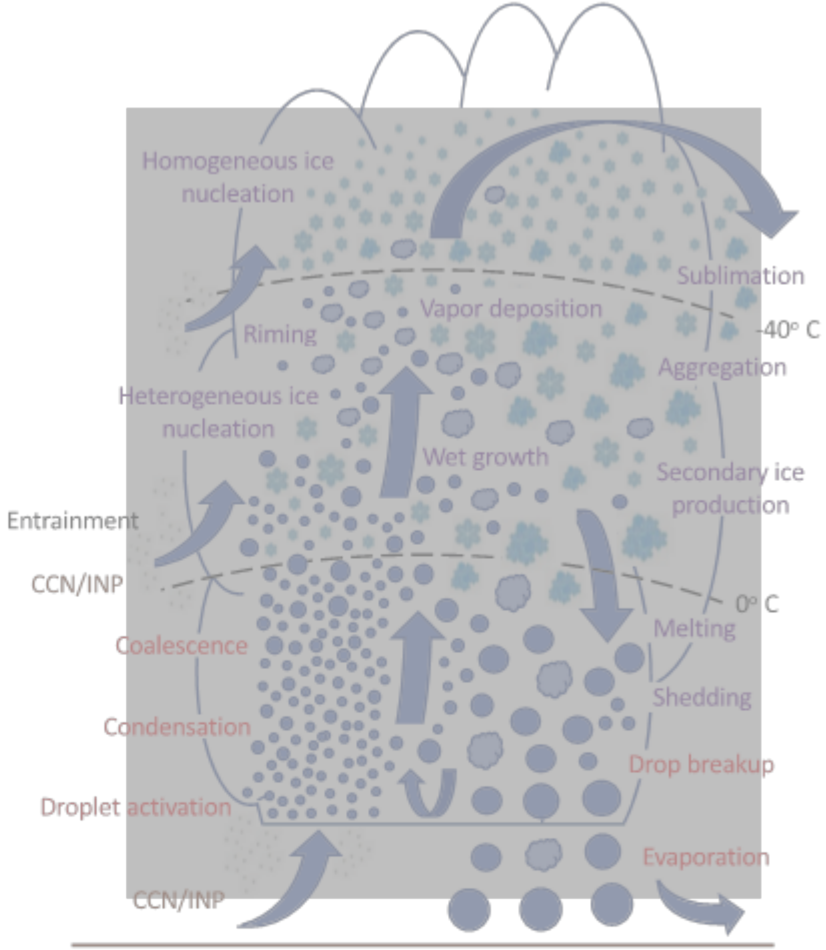




J. Will, A. Jenney, K.D. Lamb, M.S. Pritchard, C. Kaul, P.-L. Ma, K. Pressel, J. Shpund, M. van Lier-Walqui, S. Mandt, "Understanding and Visualizing Droplet Distributions in Simulations of Shallow Clouds", Neurips ML4Physics Workshop, 2023.

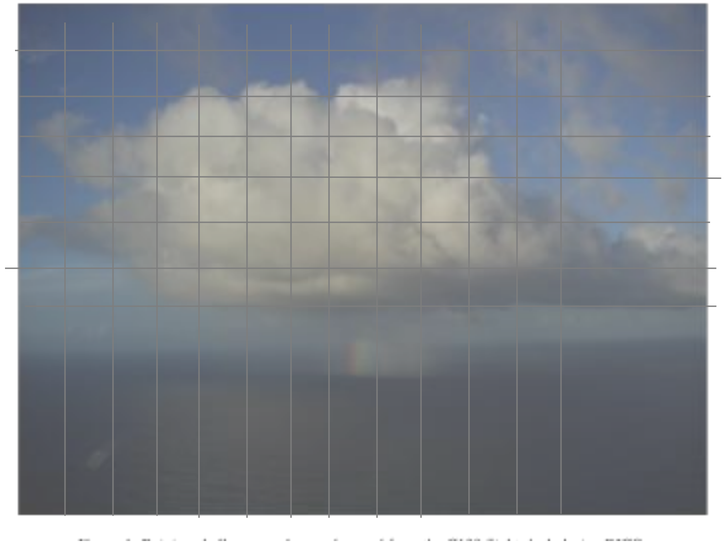
What don’t we know about cloud processes?

We don’t know all of the physics

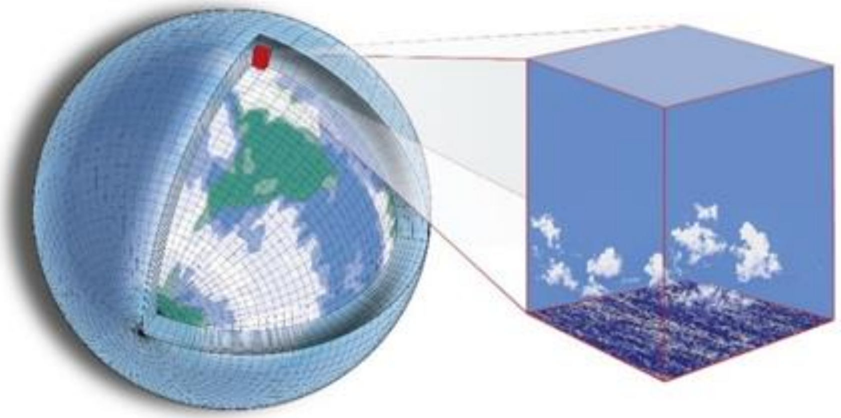


Morrison et al. 2020, JAMES

We don’t know the optimal way to represent cloud processes in models

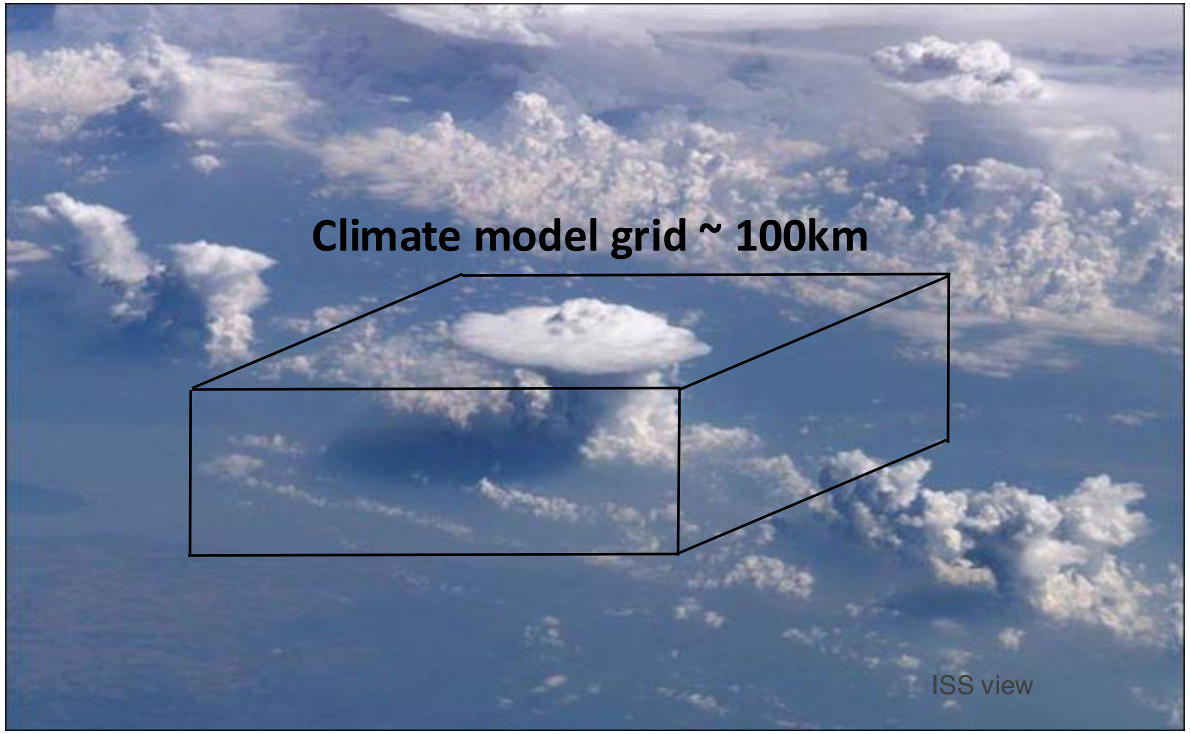


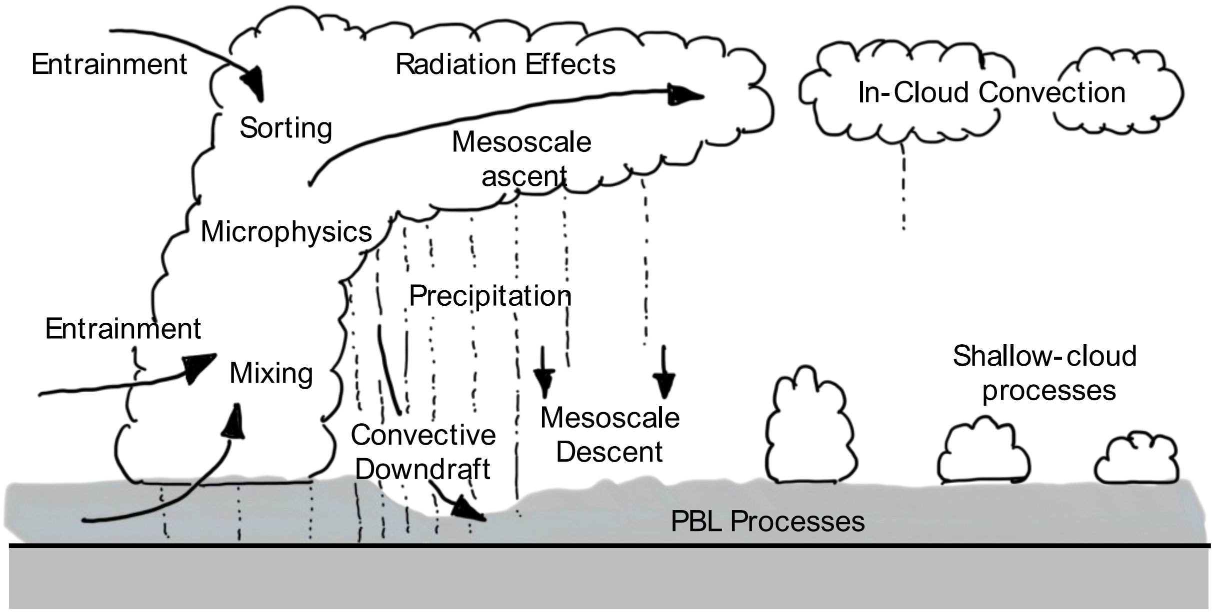
We don’t know how to consistently represent these processes at different scales



Schneider et al. 2017

Parameterizing sub-grid-scale clouds and convection is challenging



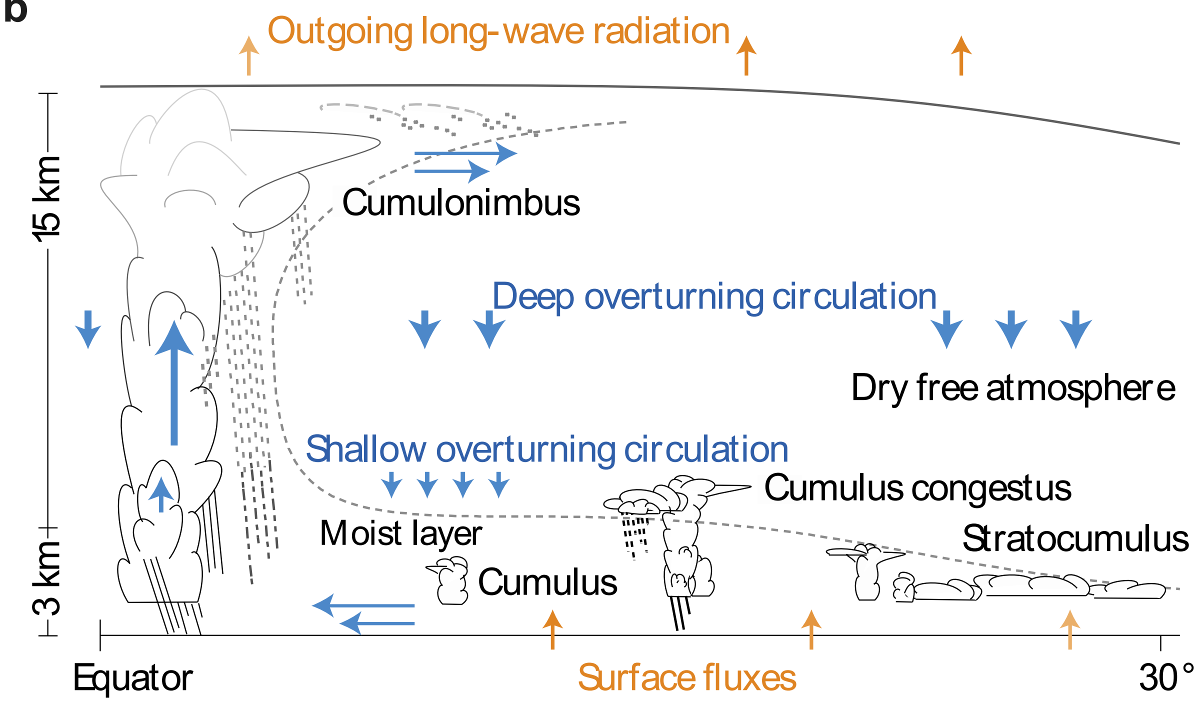


- Quasi-equilibrium hypothesis assumes there is a separation of scales between the large (grid scale) and small (sub-grid-scale) [Arakawa and Schubert, 1974]
- Most climate models rain too little and too often [Stevens et al. 2010]
- Do we need stochastic parameterizations of deep convection to represent unresolved physics? [Majda et al. 2002; Khouider et al. 2003]

Arakawa, 2004

Convection typically becomes more organized over time

a



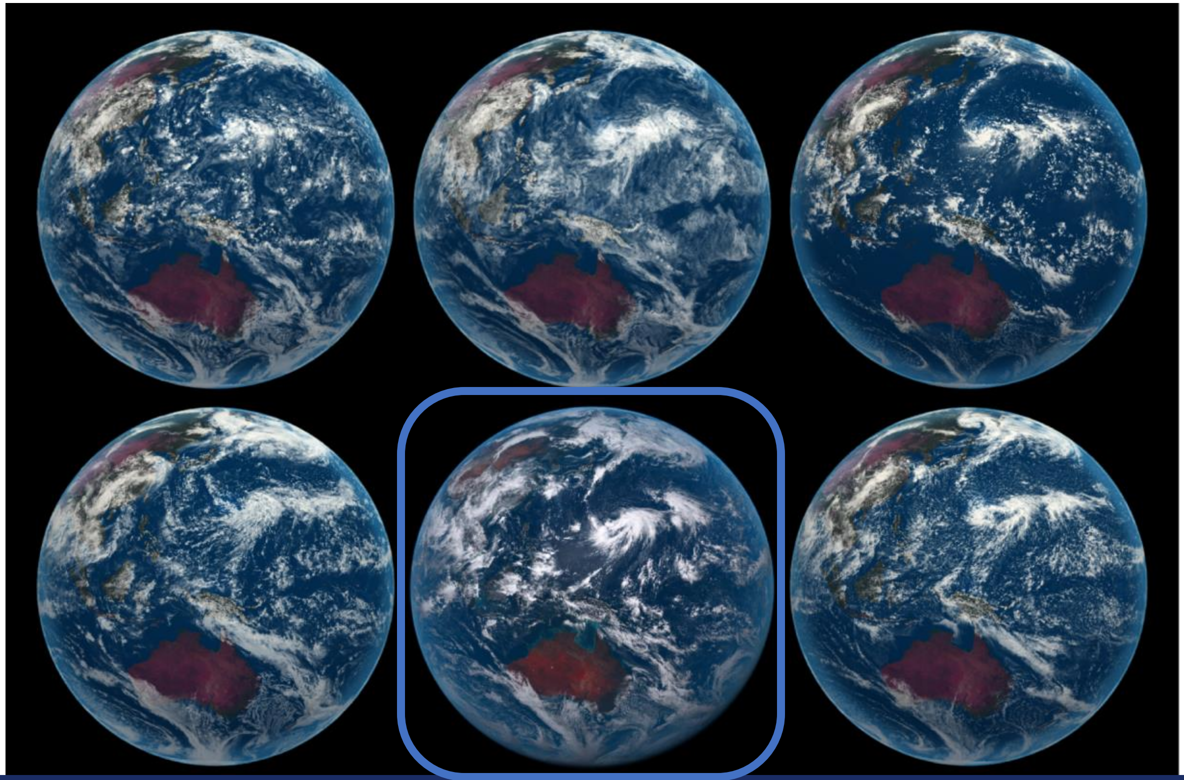
- Convection organizes at different scales and becomes more organized over time
- Should climate model deep convection schemes include a parameter for cloud organization at the sub-grid scale [Mapes and Neale, 2011]?

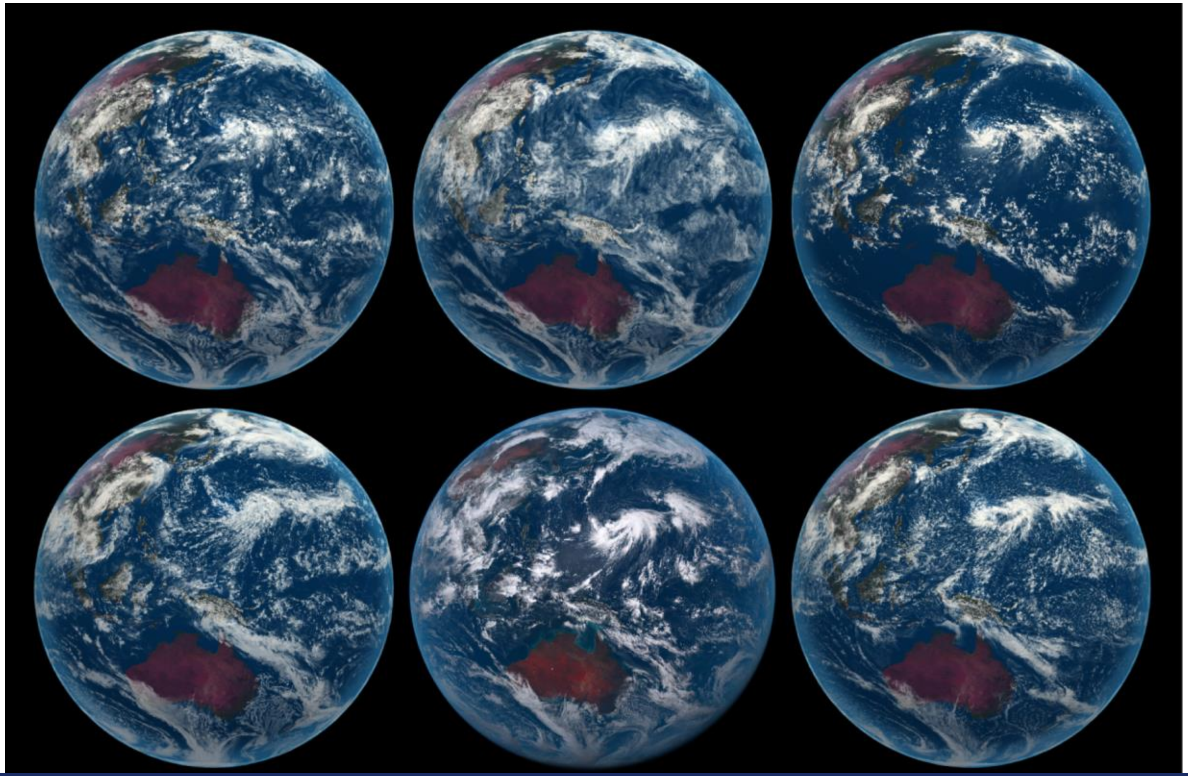
Bony et al. 2015

DYAMOND & DYAMOND 2 Intercomparison Projects

- **DY**namics of the **A**tmospheric general circulation **M**odeled on **N**on-hydrostatic **D**omains [Stevens et al. 2019]
- Model intercomparison project for global storm-resolving models
- Deep convection does not need to be parametrized (grid scale ~ 2-4 km)
- Winter: Simulations of 40 days & nights initialized on 20 January 2020
- Specified sea-surface temperature (atmospheric experiment only)
- Global SAM [Khairoutdinov, Blossey, and Bretherton, 2022]

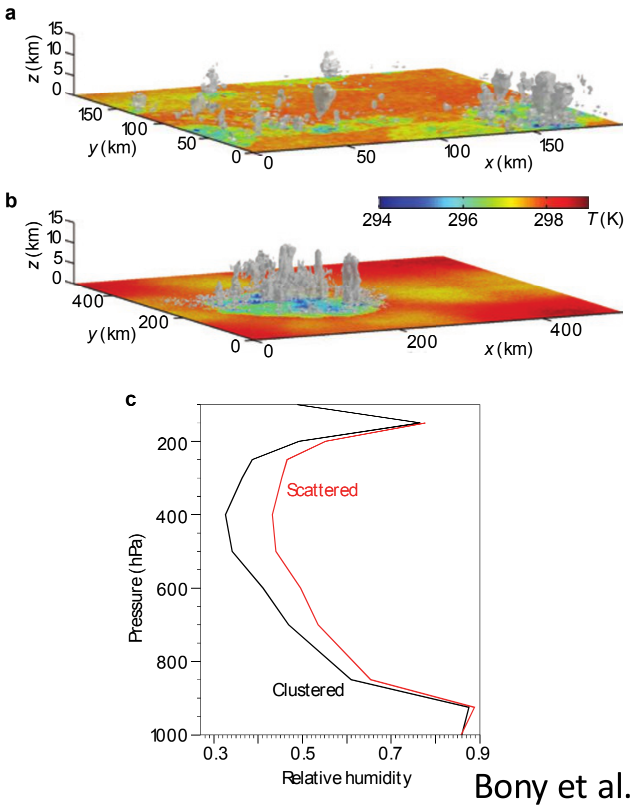
Stevens et al. 2019





There are <u>many</u> different metrics for cloud spatial organization [Janssens et al. 2021]

Which one should we use?

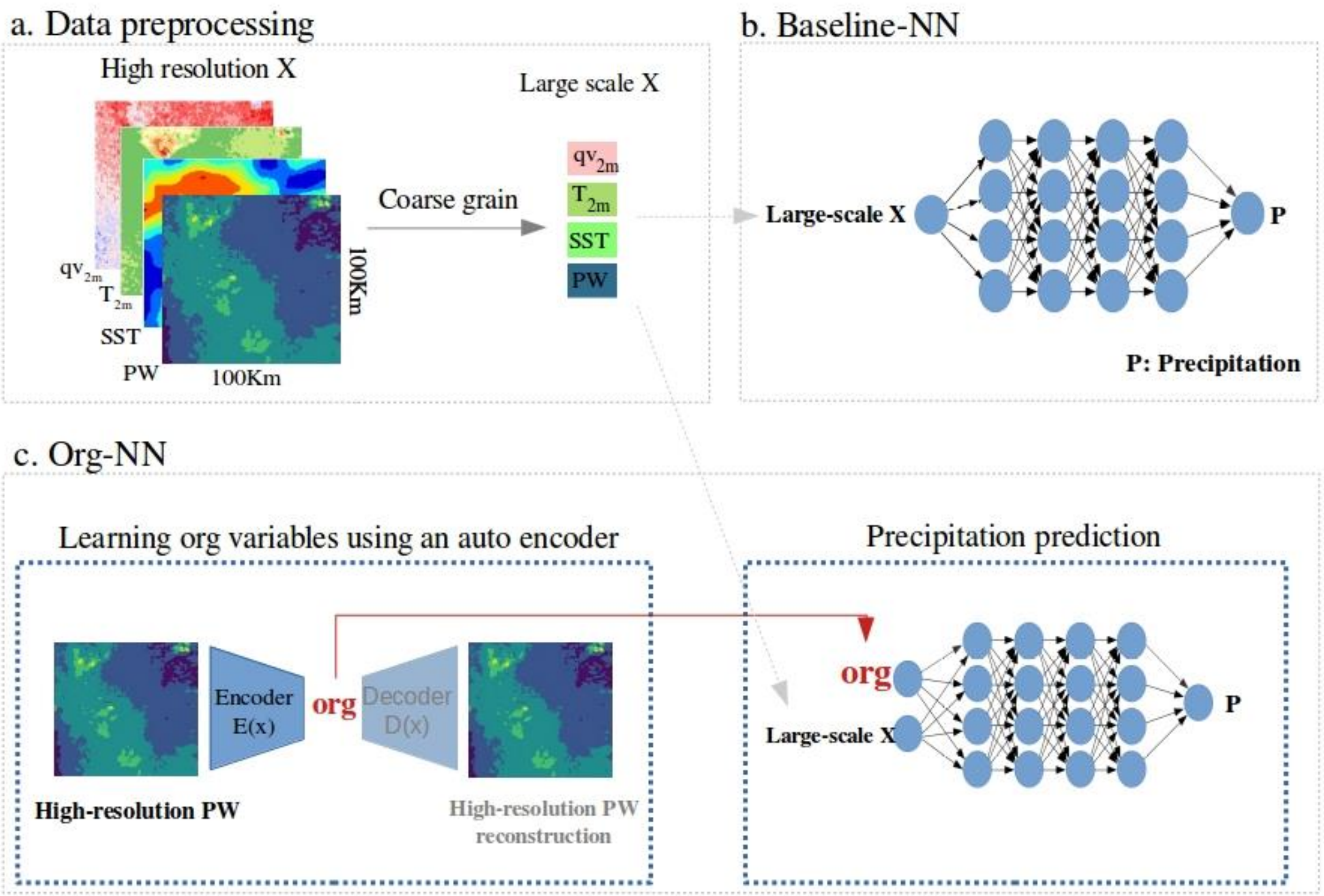


Bony et al. 2015

	PR	CLF	PW_{mean}	PW_{std}	$CLW_{var.ratio}$	N_{cloud}	OS_{mean}	I_{mean}	I_{max}	$Iorg_P$	SCAI	COP	RDF_{max}	SizeExp	FracDim	SpecL
PR	1.00	0.40	0.38	0.17	0.21	0.56	-0.39	0.52	0.59	-0.34	0.40	-0.00	-0.39	-0.05	0.47	0.47
CLF	0.40	1.00	0.34	0.11	0.17	0.43	-0.49	0.24	0.34	-0.25	0.52	-0.07	-0.29	0.01	0.60	0.09
PW_{mean}	0.38	0.34	1.00	-0.09	0.04	0.35	-0.27	0.17	0.24	-0.27	0.22	-0.13	-0.24	-0.09	0.24	0.21
PW_{std}	0.17	0.11	-0.09	1.00	0.27	0.09	0.07	0.23	0.21	0.04	-0.03	0.13	-0.07	0.02	0.01	0.21
$CLW_{var.ratio}$	0.21	0.17	0.04	0.27	1.00	0.03	0.21	0.67	0.60	0.15	-0.17	0.43	-0.01	0.20	-0.10	0.53
N_{cloud}	0.56	0.43	0.35	0.09	0.03	1.00	-0.59	0.06	0.29	-0.48	0.65	-0.40	-0.60	-0.32	0.66	0.28
OS_{mean}	-0.39	-0.49	-0.27	0.07	0.21	-0.59	1.00	0.00	-0.18	0.50	-0.84	0.45	0.50	0.20	-0.89	0.04
I_{mean}	0.52	0.24	0.17	0.23	0.67	0.06	0.00	1.00	0.83	-0.05	0.01	0.49	-0.11	0.20	0.12	0.59
I_{max}	0.59	0.34	0.24	0.21	0.60	0.29	-0.18	0.83	1.00	-0.14	0.21	0.22	-0.22	0.04	0.30	0.53
$Iorg_P$	-0.34	-0.25	-0.27	0.04	0.15	-0.48	0.50	-0.05	-0.14	1.00	-0.56	0.52	0.54	0.13	-0.48	-0.11
SCAI	0.40	0.52	0.22	-0.03	-0.17	0.65	-0.84	0.01	0.21	-0.56	1.00	-0.50	-0.51	-0.19	0.90	-0.08
COP	-0.00	-0.07	-0.13	0.13	0.43	-0.40	0.45	0.49	0.22	0.52	-0.50	1.00	0.52	0.33	-0.33	0.20
RDF_{max}	-0.39	-0.29	-0.24	-0.07	-0.01	-0.60	0.50	-0.11	-0.22	0.54	-0.51	0.52	1.00	0.23	-0.51	-0.23
SizeExp	-0.05	0.01	-0.09	0.02	0.20	-0.32	0.20	0.20	0.04	0.13	-0.19	0.33	0.23	0.99	-0.15	0.04
FracDim	0.47	0.60	0.24	0.01	-0.10	0.66	-0.89	0.12	0.30	-0.48	0.90	-0.33	-0.51	-0.15	1.00	-0.02
SpecL	0.47	0.09	0.21	0.21	0.53	0.28	0.04	0.59	0.53	-0.11	-0.08	0.20	-0.23	0.04	-0.02	1.00

S. Shamekh, K. Lamb, Y. Huang, P. Gentine. Implicit Learning of convective organization explains precipitation stochasticity. *PNAS*, 2023

Use neural networks to investigate variability in predicted precipitation



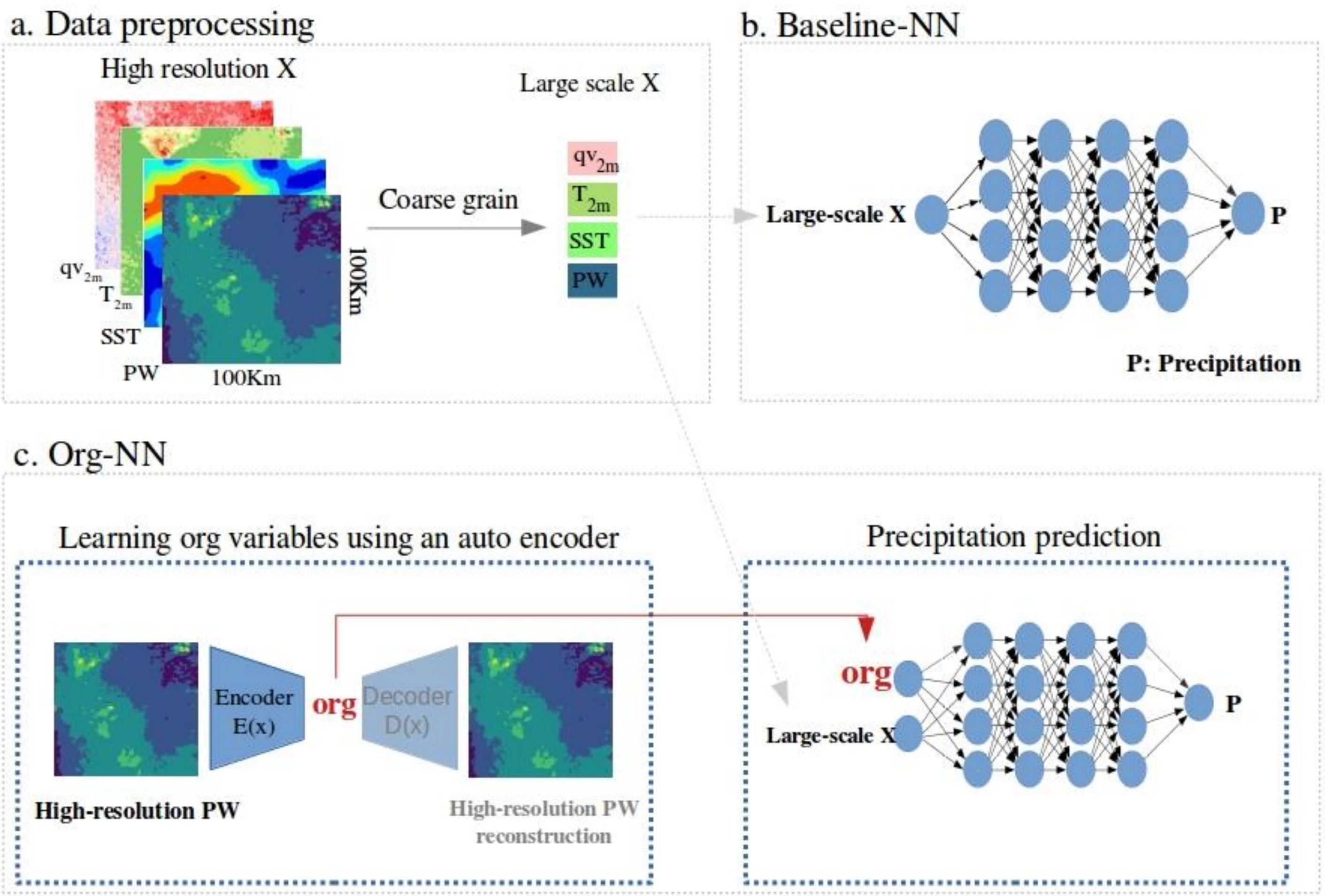
Is precipitation at the climate grid scale predictable from large scale variables only?

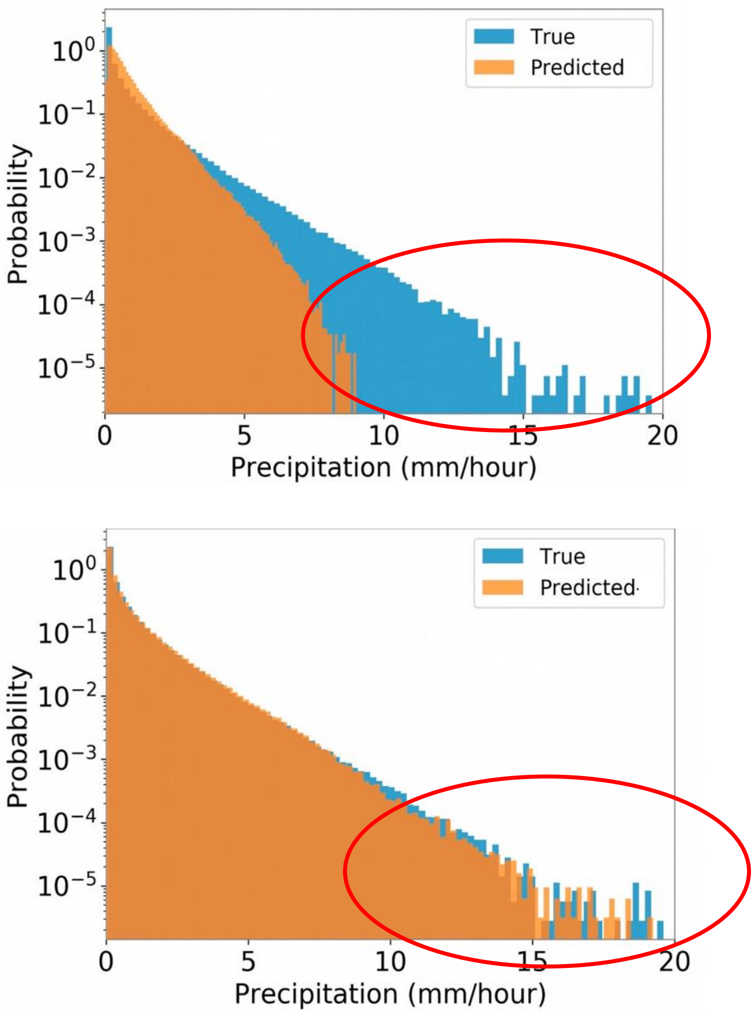
Or do we also need to include some additional information from the sub-grid-scale?

If so, what information do we need?

S. Shamekh, K. Lamb, Y. Huang, P. Gentine. Implicit Learning of convective organization explains precipitation stochasticity. *PNAS*, 2023

Org. parameter improves prediction of precipitation extremes

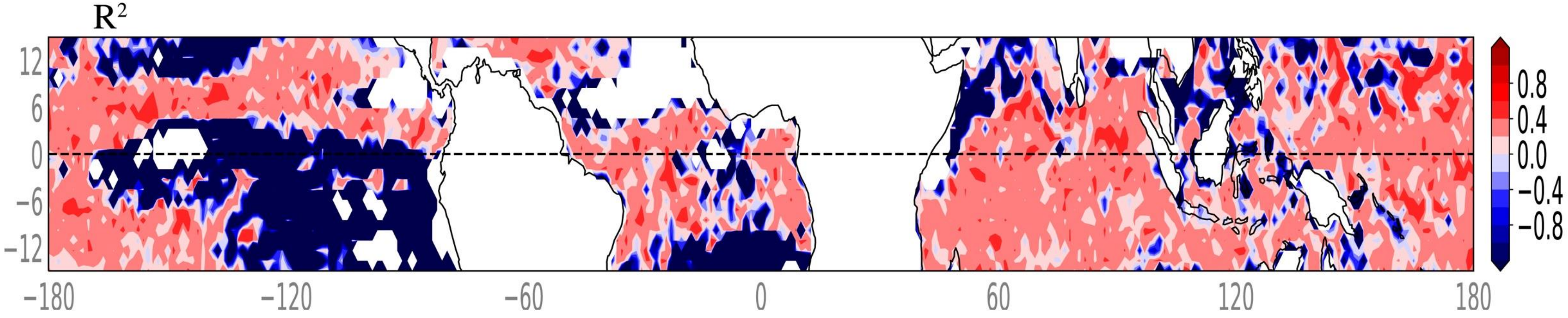




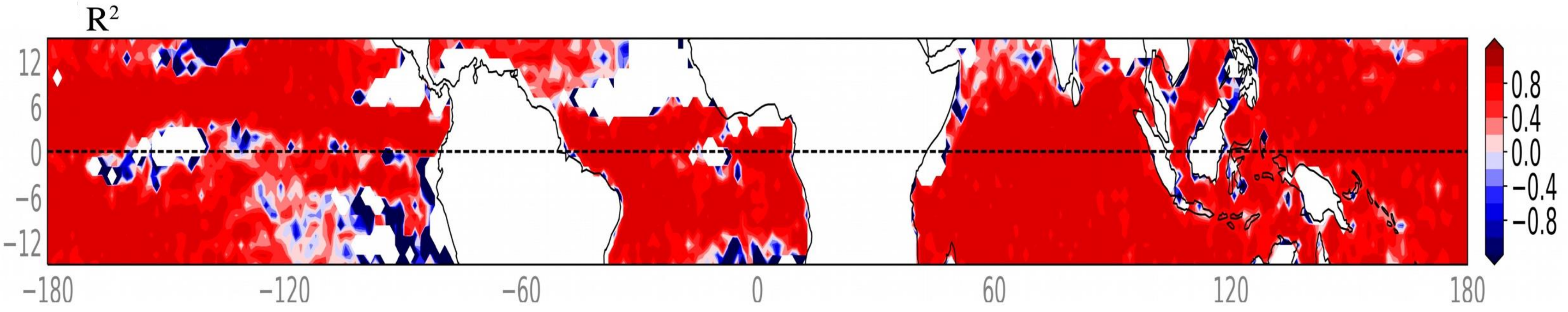
S. Shamekh, K. Lamb, Y. Huang, P. Gentine. Implicit Learning of convective organization explains precipitation stochasticity. *PNAS*, 2023

Org. parameter also improves spatial correlation of precipitation

Baseline-NN

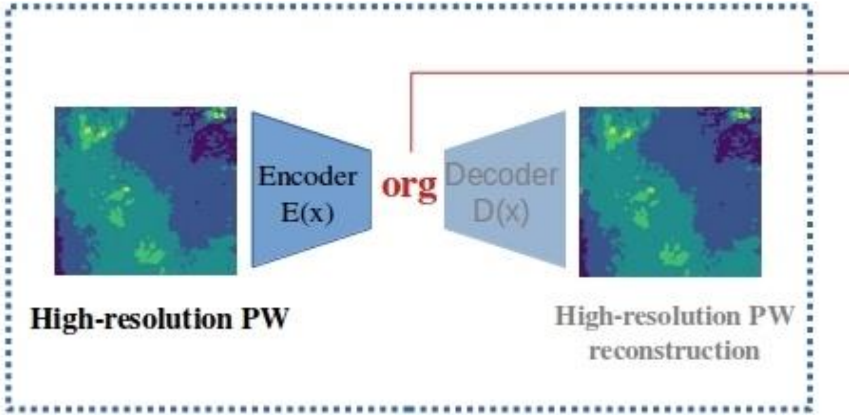


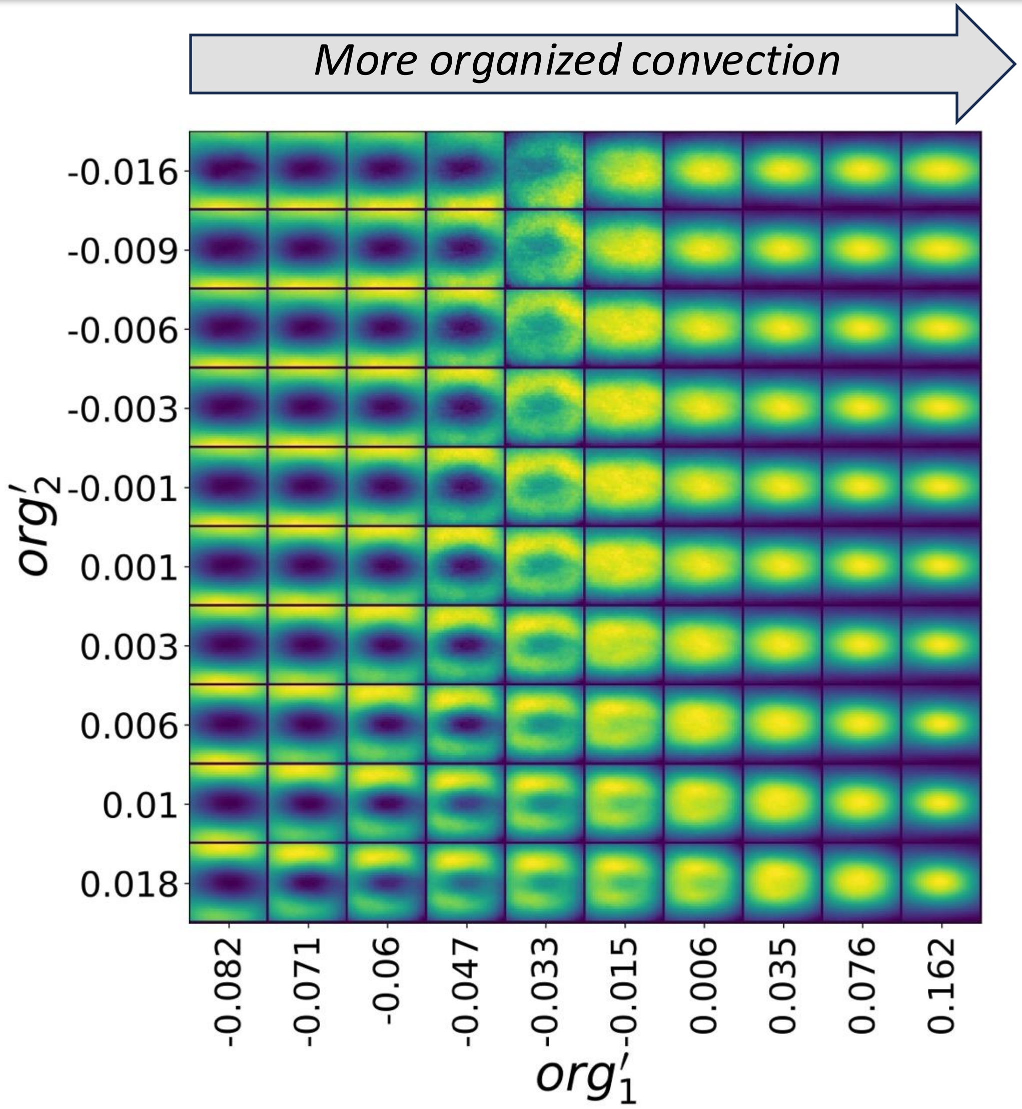
Org.-NN



S. Shamekh, K. Lamb, Y. Huang, P. Gentine. Implicit Learning of convective organization explains precipitation stochasticity. *PNAS*, 2023

Org. parameter is related organization of sub-grid-scale convection





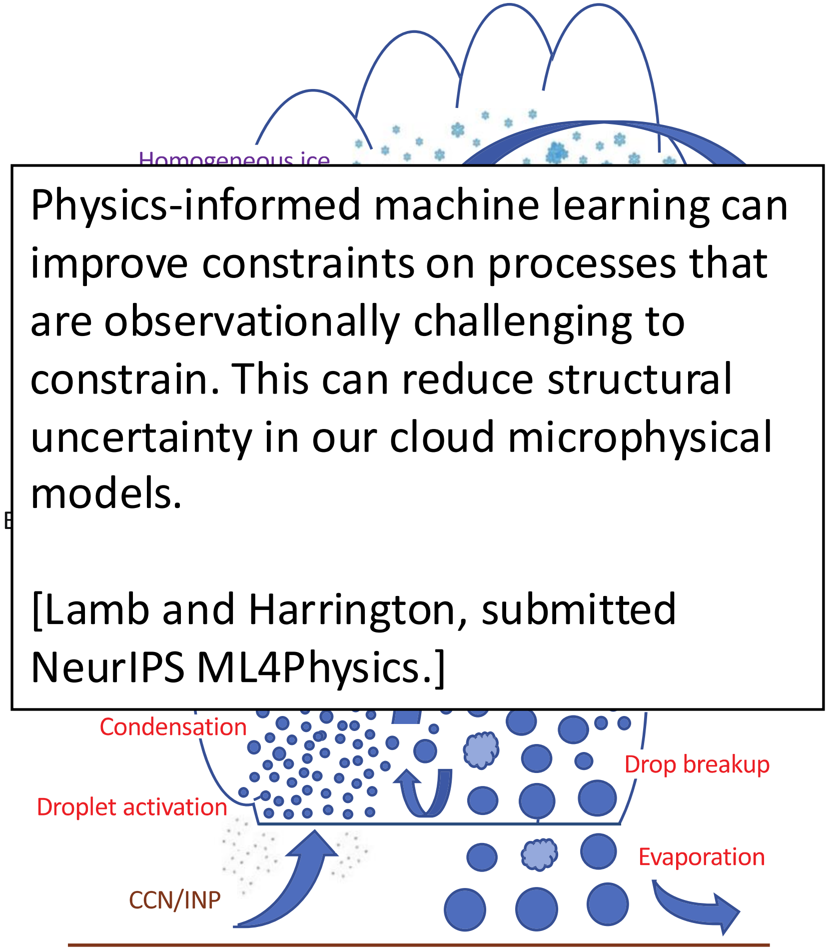
Org. parameters can be predicted from their own history

This suggests we don't need stochastic parameterizations of deep convection

S. Shamekh, K. Lamb, Y. Huang, P. Gentine. Implicit Learning of convective organization explains precipitation stochasticity. *PNAS*, 2023

What don’t we know about cloud processes?

We don’t know all of the physics



Physics-informed machine learning can improve constraints on processes that are observationally challenging to constrain. This can reduce structural uncertainty in our cloud microphysical models.

[Lamb and Harrington, submitted NeurIPS ML4Physics.]

Morrison et al. 2020, JAMES

We don’t know the optimal way to represent cloud processes in models

Reduced-order modeling of the cloud microphysical state is a promising method for learning simplified bulk microphysical models.

It can be used to connect the microscopic evolution of the cloud to its macroscopic evolution.

[Lamb et al. 2024; Will et al. 2023; Sturm et al. in prep.]

We don’t know how to consistently represent these processes at different scales

We can learn sufficient information from the sub-grid-scale to predict precipitation at the climate scale using machine learning.

This suggests that we don’t need stochastic parameterizations of deep convection.

[Shamekh, Lamb, Yu, Gentine, 2023]

Schneider et al. 2017

Links to publications

K.D. Lamb & J.Y. Harrington “Discovering How Ice Crystals Grow using Neural ODE’s and Symbolic Regression”, submitted.

K.D. Lamb, M. van Lier Walqui, S. Santos, H. Morrison. “Reduced Order Modeling for Linearized Representations of Microphysical Process Rates, *JAMES*.

J. Will, A. Jenney, K.D. Lamb, M.S. Pritchard, C. Kaul, P.-L. Ma, K. Pressel, J. Shpund, M. van Lier-Walqui, S. Mandt, “Understanding and Visualizing Droplet Distributions in Simulations of Shallow Clouds”, Neurips ML4Physics Workshop, 2023.

S. Shamekh, K. Lamb, Y. Huang, P. Gentine. Implicit Learning of convective organization explains precipitation stochasticity. *PNAS*, 2023

ASR Atmospheric System Research

LEAP

Contact: kl3231@columbia.edu