Parameterizing mesoscale eddy thickness/buoyancy fluxes using small neural networks

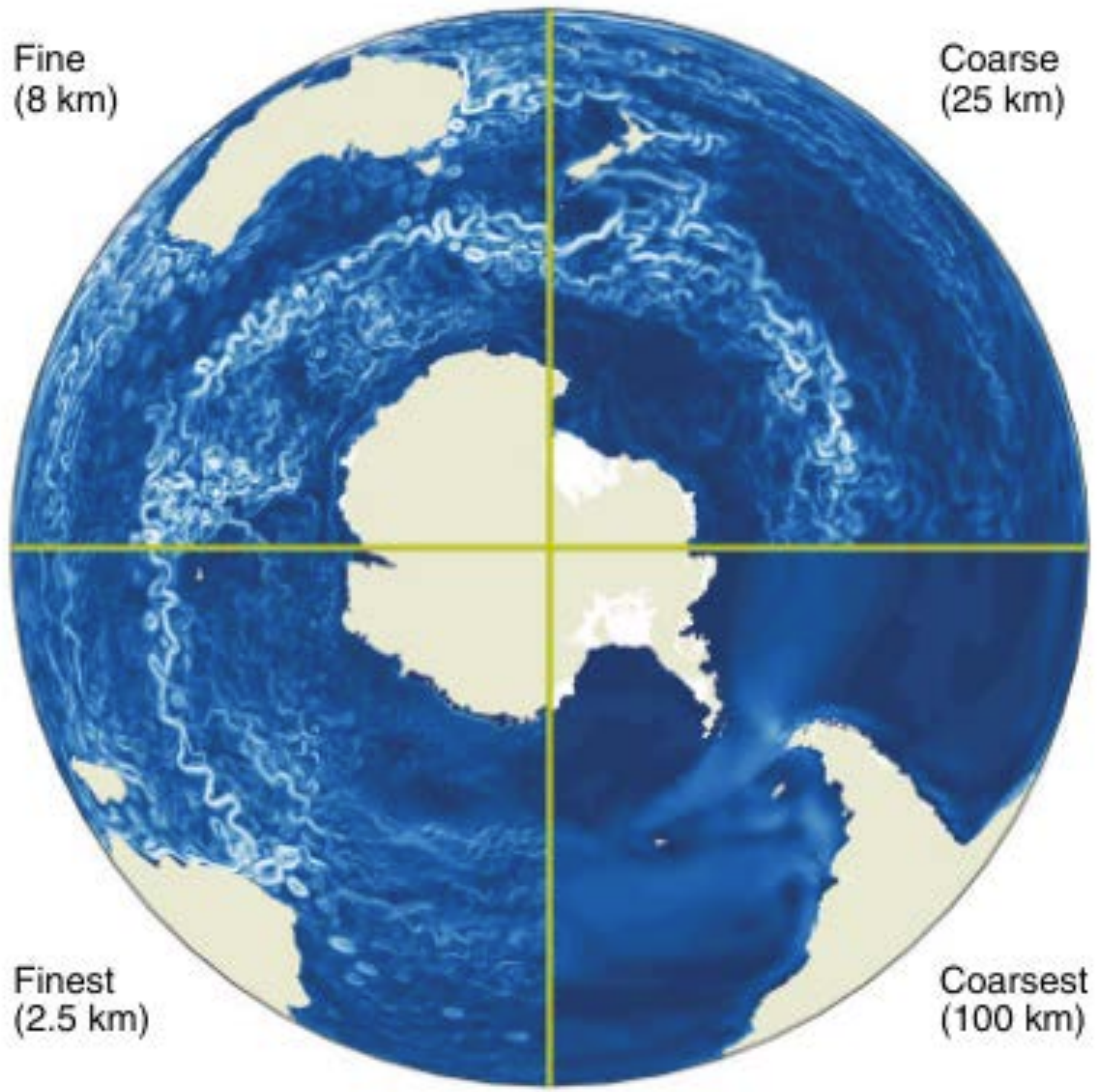
CLIVAR-OMDP Workshop, 11 September 2024

Dhruv Balwada[2], Pavel Perezhogin[1],
Ryan Abernathey[2], Alistair Adcroft[3], & Laure Zanna[1]

[1] Courant Institute of Mathematical Sciences, New York University
[2] Lamont-Doherty Earth Observatory, Columbia University
[3] Princeton University

@M2LInES
https://m2lines.github.io

Turbulence at macro-scales (1-100 km)

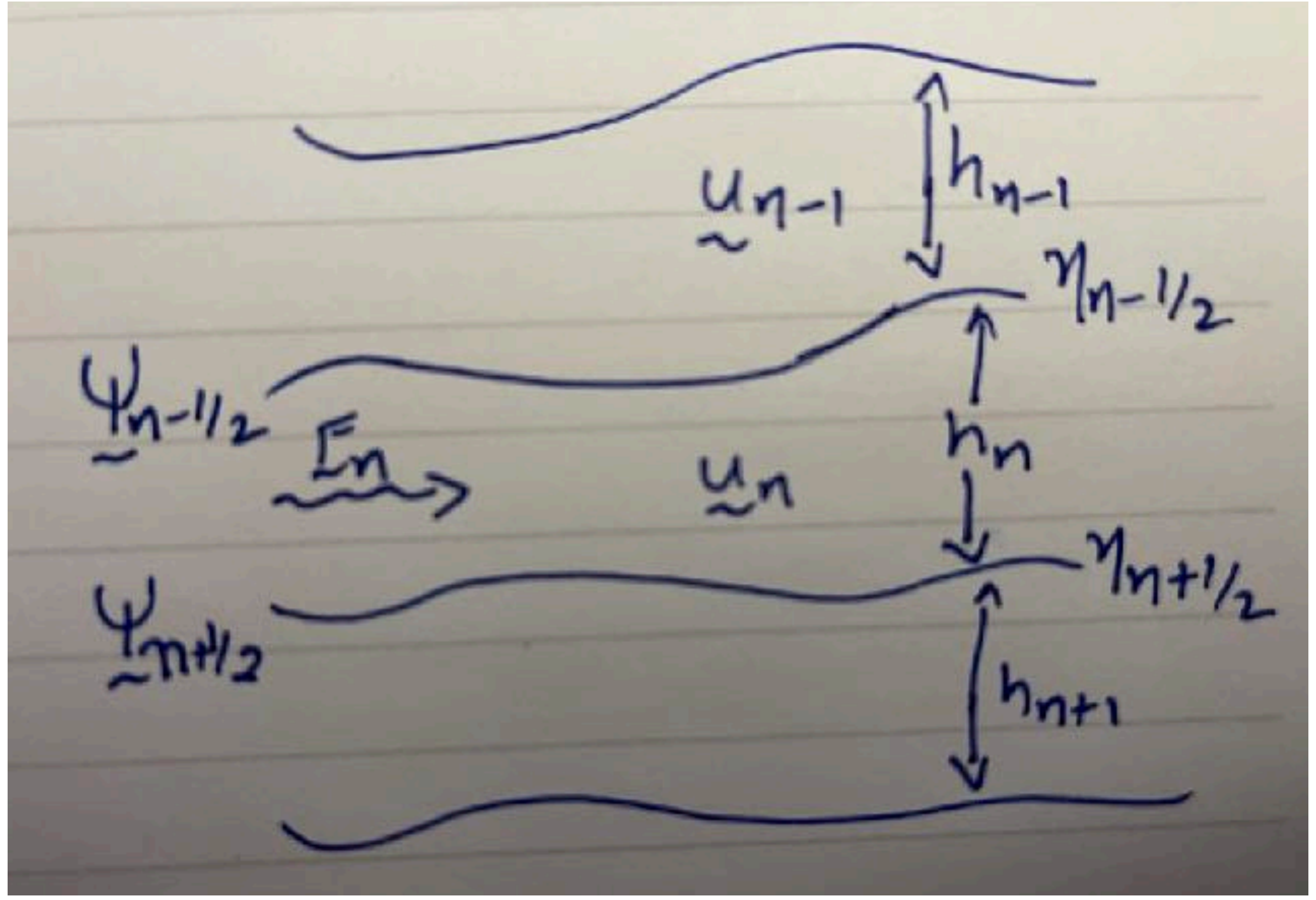


Hewitt et al 2022

- Transport of momentum, heat, salt, and other tracers.
- Modulate the patterns of large scale flow and stratification.
- Play a vital role in the ocean's mechanical energy cycle
  - Form bulk of kinetic energy in ocean.
  - Transfers kinetic energy across scales.

- Goal: **Develop scale aware parameterizations that work over both non-eddying and eddy permitting resolutions.**

Eddies in thickness equation



In layered models: $$\partial_t h_n + \nabla\cdot(\mathbf{u}_n h_n) = 0$$

$\nabla = (\partial_x, \partial_y)$

In filtered form: $$\partial_t \overline{h}_n + \nabla\cdot(\overline{\mathbf{u}}_n \overline{h}_n) = -\nabla\cdot(\overline{\mathbf{u}_n h_n} - \overline{\mathbf{u}}_n \overline{h}_n) = -\nabla\cdot(\mathbf{F}_n)$$

Advective form: $\mathbf{F}_n = \delta_n \boldsymbol{\Psi} = \boldsymbol{\Psi}_{n-1/2} - \boldsymbol{\Psi}_{n+1/2}$

Gent-McWilliams: $\boldsymbol{\Psi}_{n-1/2} = -\kappa_{GM}\nabla\eta_{n-1/2}$

B.C.- No depth integrated flux: $\boldsymbol{\Psi}_0 = \boldsymbol{\Psi}_{N+1/2} = 0$

Our goal is to build a data driven parameterization of F:

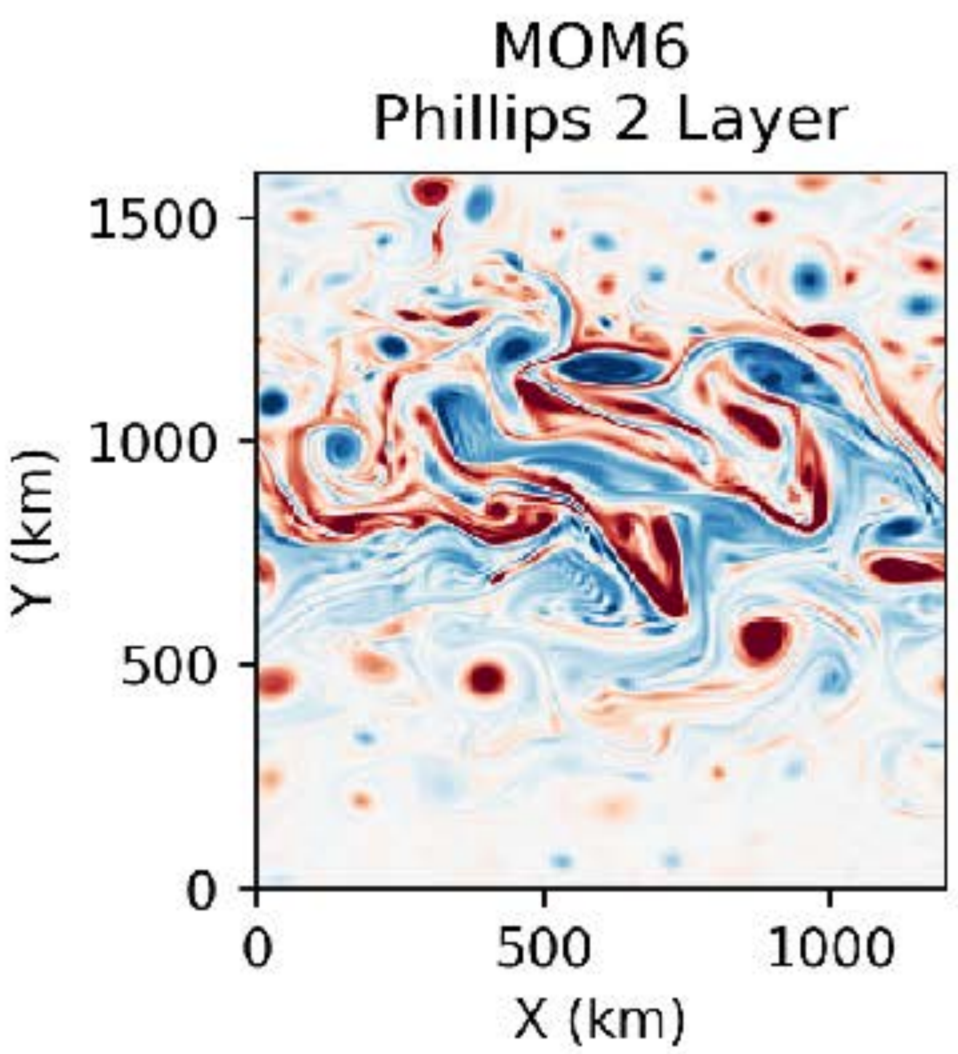
Find/build a function connecting F to large scale variables.

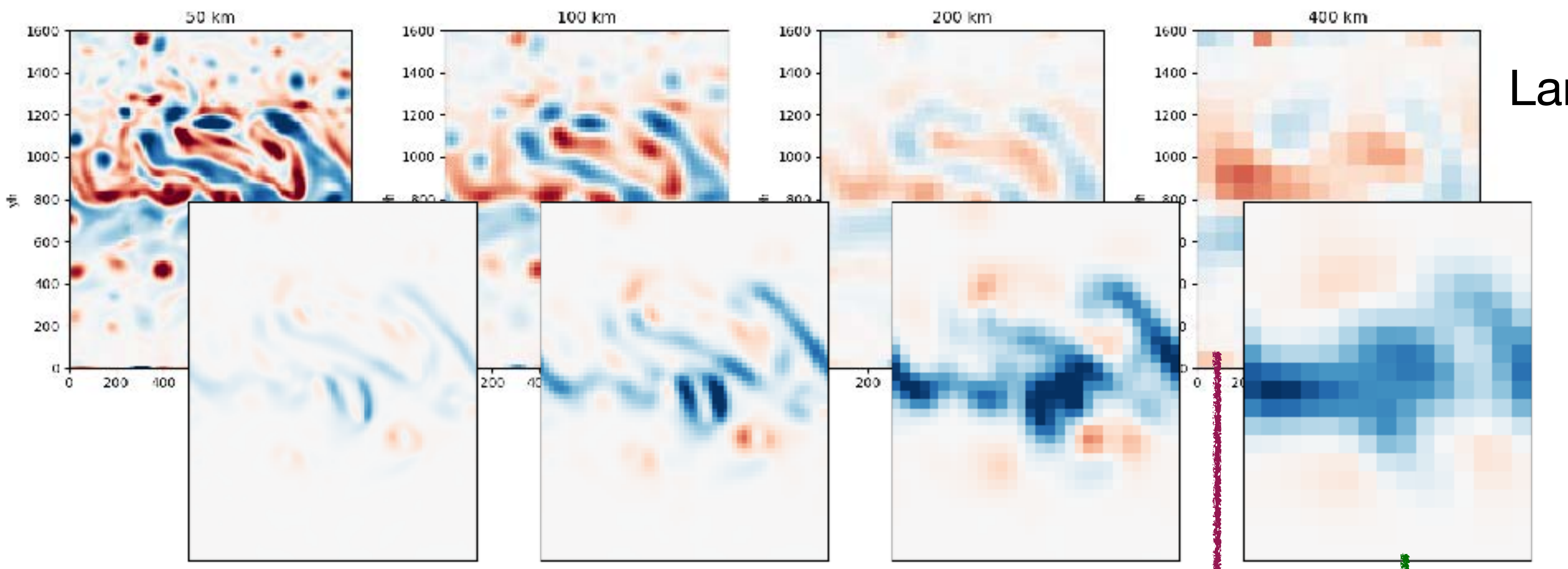
General process

General process

Filter scale (l)

Spatial filter



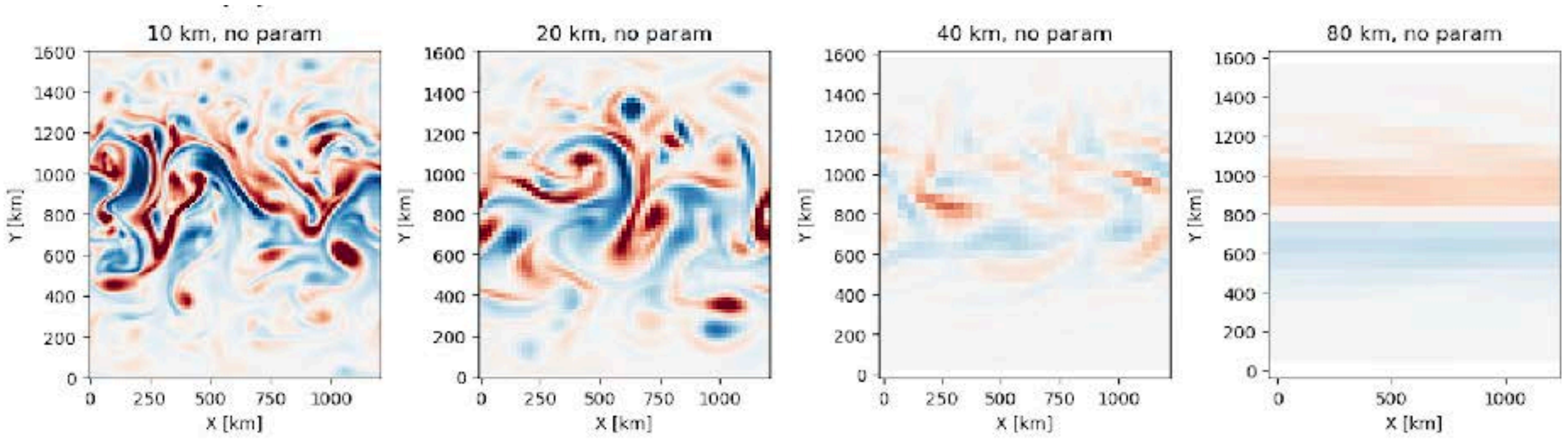


Large scale fields

Sub-grid forcing

F = f(large scale)

Coarse resolution simulations



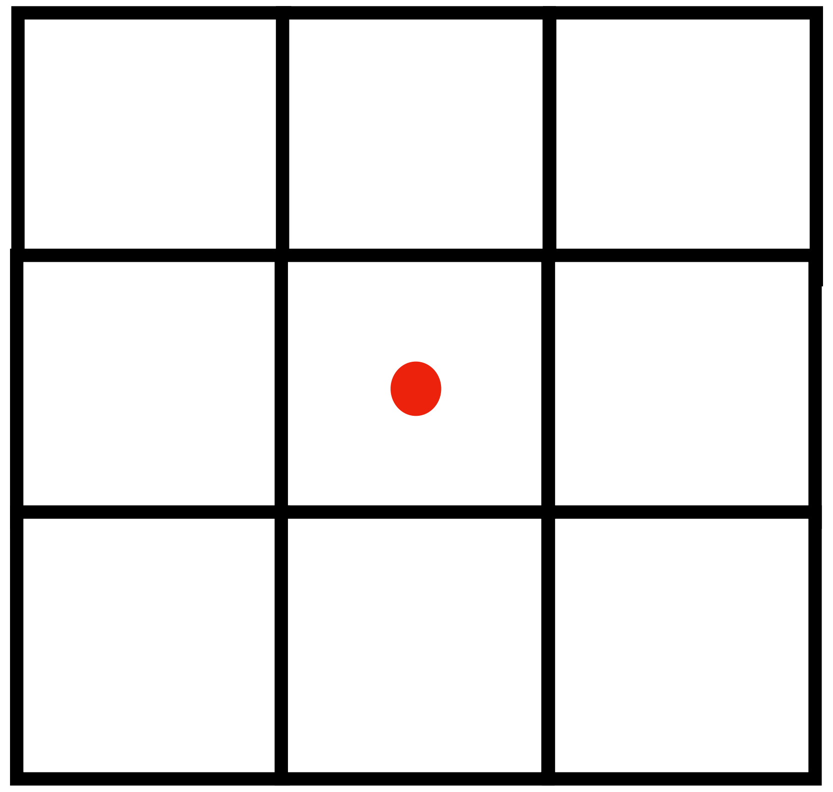
Model Resolution (dx)

General process

ML Parameterization design

$$\mathbf{F}_n = f(\nabla \overline{\mathbf{u}}_n, \nabla \overline{h}_n, \triangle)$$

1. Non-dimensionalization
2. Range limited inputs
3. Separation of variables
4. Coordinate-Invariance
5. Scale-awareness
6. Lateral non-locality
7. Small ANN

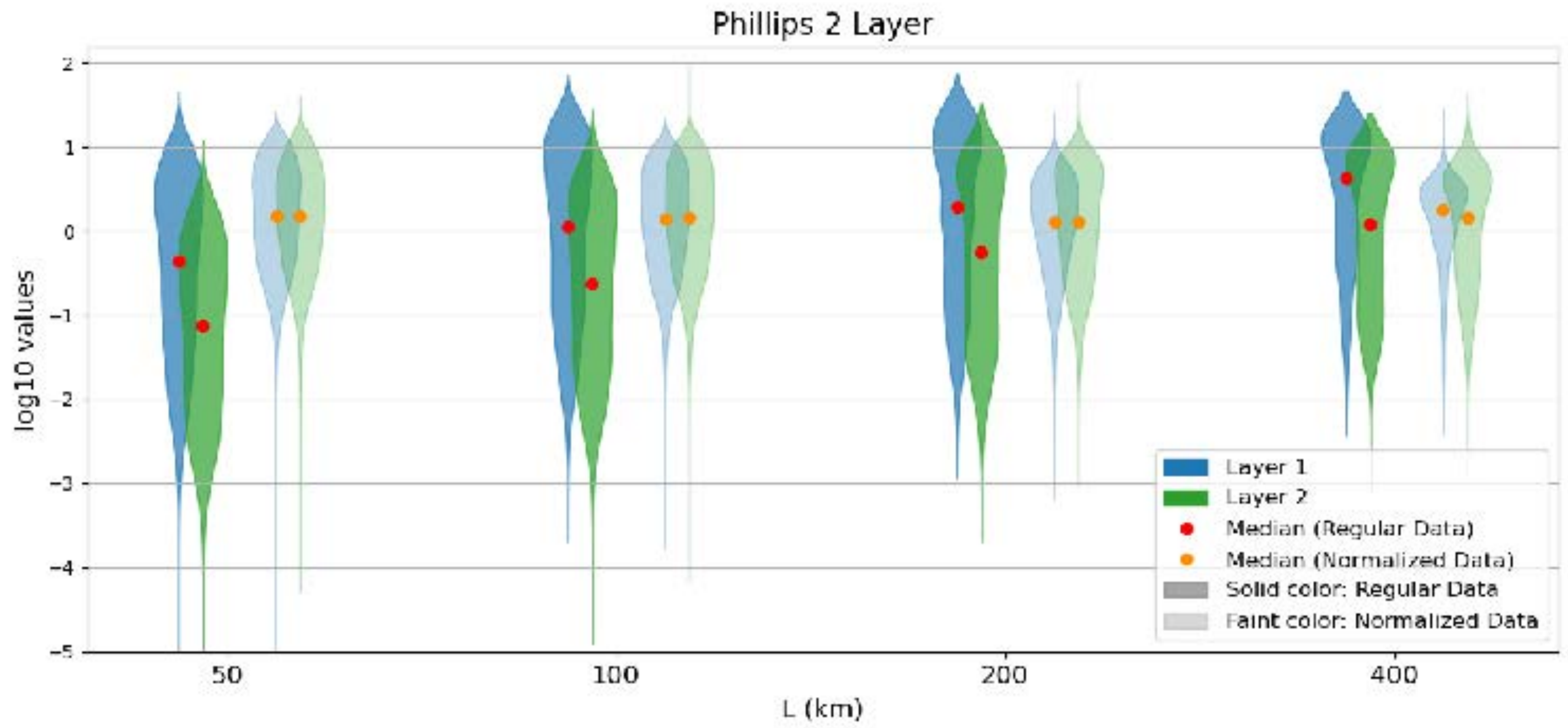


Point-wise/Local models

$$\widetilde{\mathbf{F}}_{n,(i,j)} = \triangle^2 |\nabla \overline{\mathbf{u}}_{n,(i+p,j+p)}| f\Big(\frac{\widetilde{\nabla \overline{\mathbf{u}}}_{n,(i+p,j+p)}}{|\nabla \overline{\mathbf{u}}_{n,(i+p,j+p)}|}, \frac{\widetilde{\nabla \overline{h}}_{n,(i+p,j+p)}}{|\nabla \overline{h}_{n,(i+p,j+p)}|}\Big) g(|\nabla \overline{h}_{n,(i+p,j+p)}|).$$

ML Parameterization design

$$\mathbf{F}_n = f(\nabla \overline{\mathbf{u}}_n, \nabla \overline{h}_n, \triangle)$$

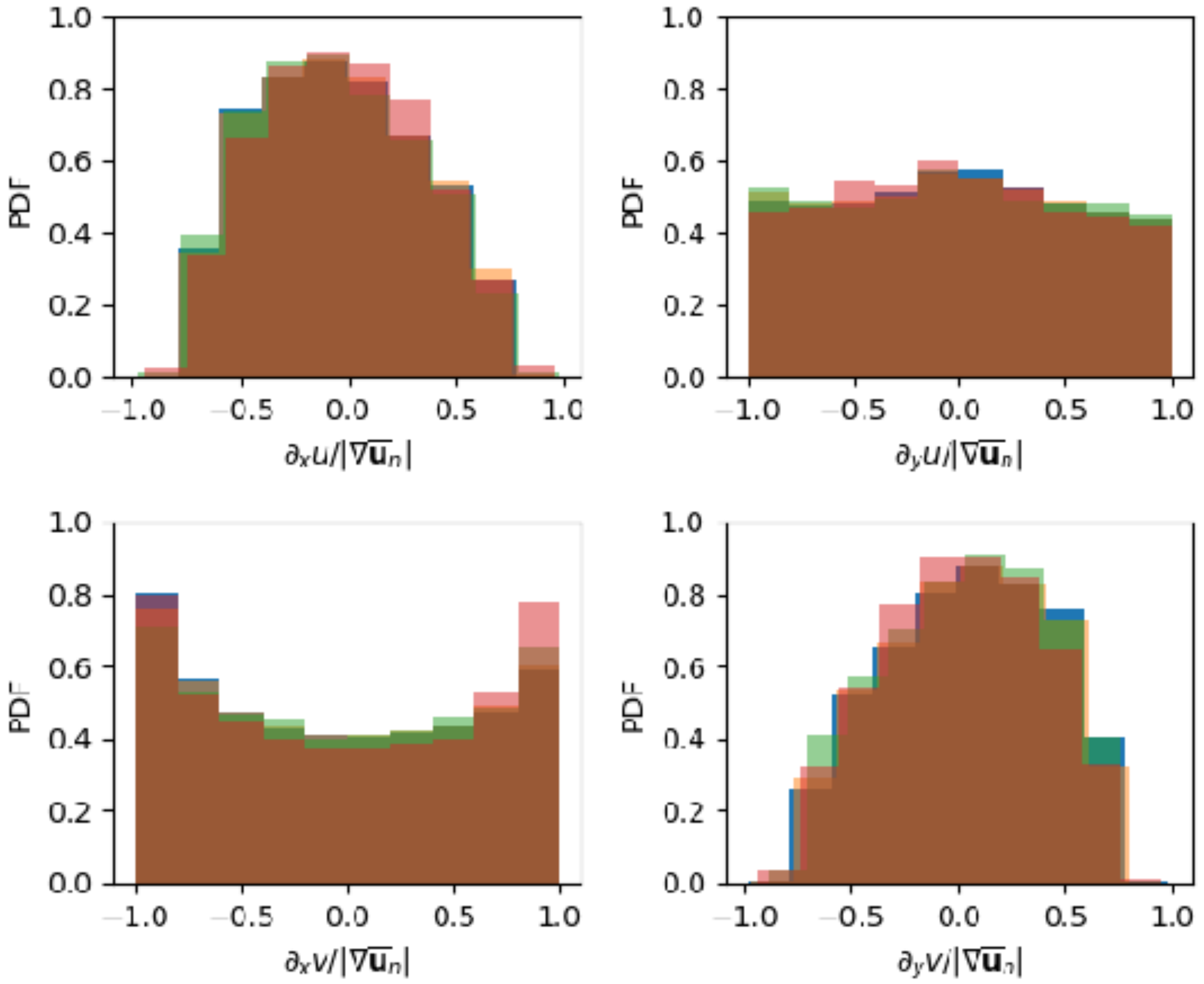


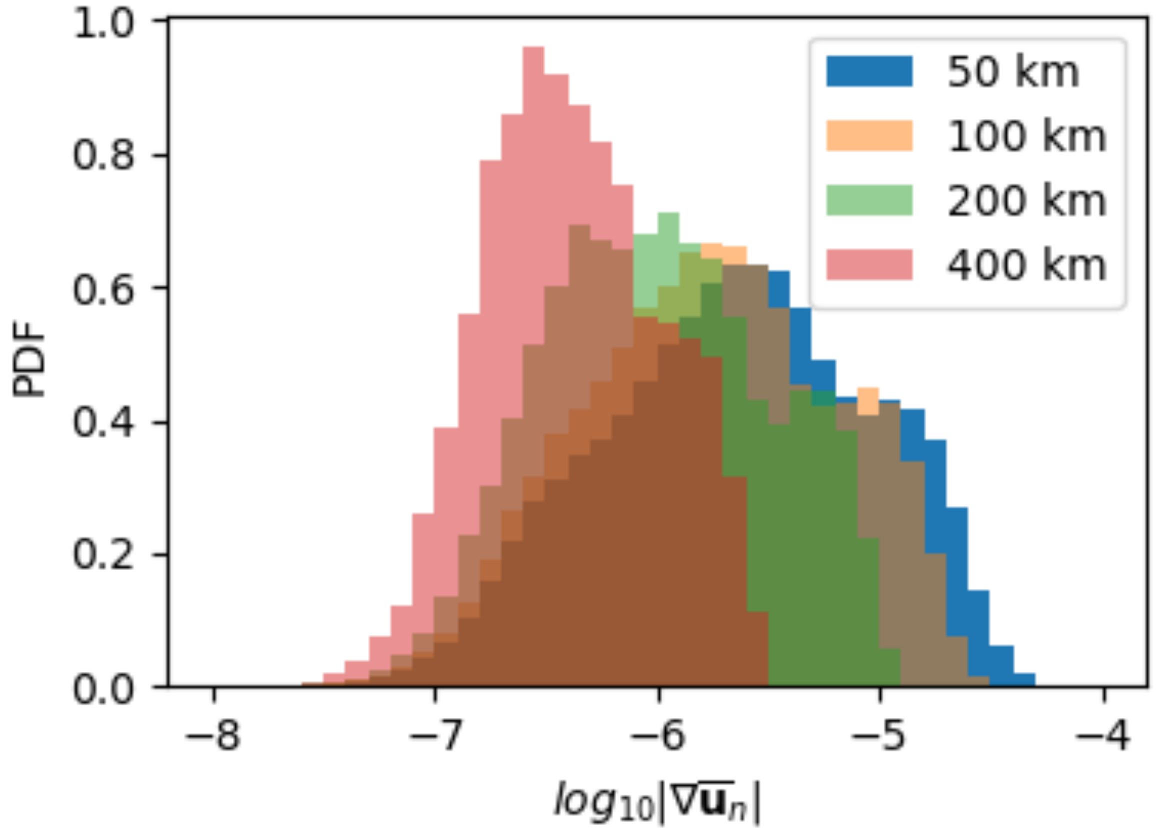
Outputs: $|\mathbf{F}_n| \rightarrow \dfrac{|\mathbf{F}_n|}{\triangle^2 |\nabla \overline{\mathbf{u}}_n|}$

1. **Non-dimensionalization**
2. Range limited inputs
3. Separation of variables
4. Coordinate-Invariance
5. Scale-awareness
6. Lateral non-locality
7. Small ANN

$$\widetilde{\mathbf{F}}_{n,(i,j)} = \triangle^2 |\nabla \overline{\mathbf{u}}_{n,(i+p,j+p)}| f\left(\frac{\widetilde{\nabla \overline{\mathbf{u}}}_{n,(i+p,j+p)}}{|\nabla \overline{\mathbf{u}}_{n,(i+p,j+p)}|}, \frac{\widetilde{\nabla \overline{h}}_{n,(i+p,j+p)}}{|\nabla \overline{h}_{n,(i+p,j+p)}|}\right) g(|\nabla \overline{h}_{n,(i+p,j+p)}|).$$

ML Parameterization design





Inputs:

$$\nabla\overline{\mathbf{u}}_n \rightarrow \left(\frac{\nabla\overline{\mathbf{u}}_n}{|\nabla\overline{\mathbf{u}}_n|}, |\nabla\overline{\mathbf{u}}_n|\right)$$

$$\nabla\overline{h}_n \rightarrow \left(\frac{\nabla\overline{h}_n}{|\nabla\overline{h}_n|}, |\nabla\overline{h}_n|\right)$$

Functions:

$$f(\nabla\overline{h}_n) \rightarrow f_1\left(\frac{\nabla\overline{h}_n}{|\nabla\overline{h}_n|}\right)f_2(|\nabla\overline{h}_n|)$$

$$\mathbf{F}_n = f(\nabla\overline{\mathbf{u}}_n, \nabla\overline{h}_n, \triangle)$$

1. Non-dimensionalization
2. **Range limited inputs**
3. **Separation of variables**
4. Coordinate-Invariance
5. Scale-awareness
6. Lateral non-locality
7. Small ANN

$$\widetilde{\mathbf{F}}_{n,(i,j)} = \triangle^2|\nabla\overline{\mathbf{u}}_{n,(i+p,j+p)}|f\left(\frac{\widetilde{\nabla\overline{\mathbf{u}}}_{n,(i+p,j+p)}}{|\nabla\overline{\mathbf{u}}_{n,(i+p,j+p)}|}, \frac{\widetilde{\nabla\overline{h}}_{n,(i+p,j+p)}}{|\nabla\overline{h}_{n,(i+p,j+p)}|}\right)g(|\nabla\overline{h}_{n,(i+p,j+p)}|).$$

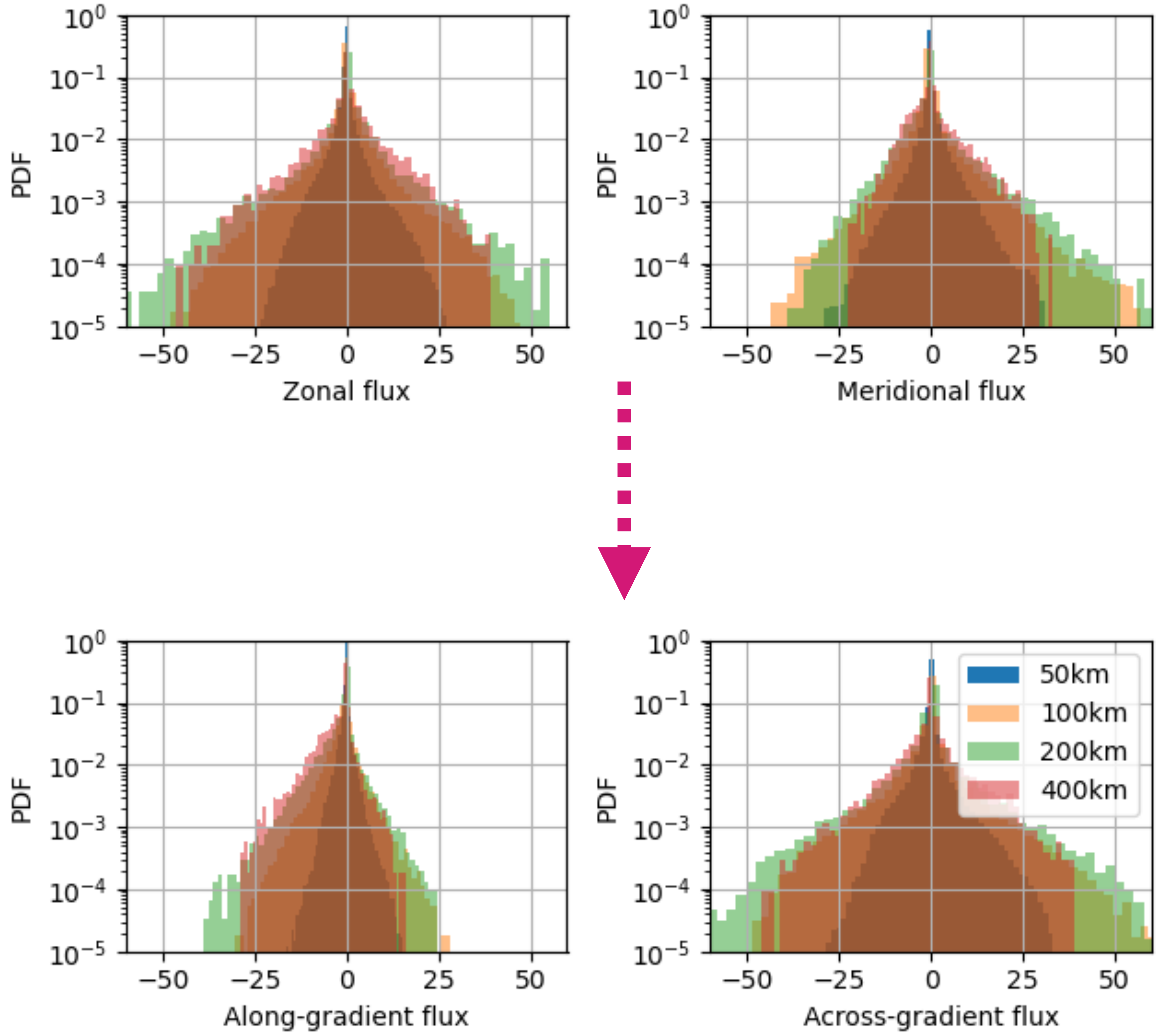
ML Parameterization design

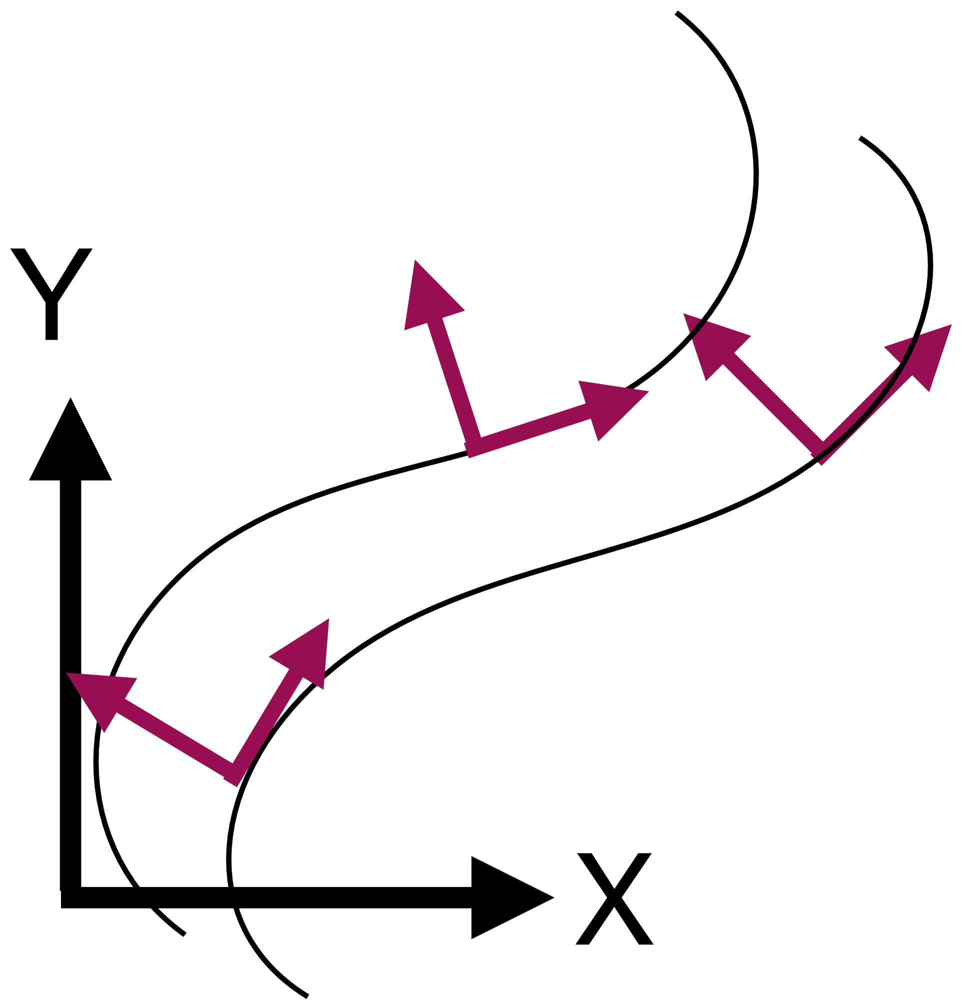
Rotate to flow dependent coordinate:

x/y -> along/across h grad

$$\widetilde{\mathbf{F}}_n = \mathbf{R}_n^T \mathbf{F}_n$$

$$\widetilde{\nabla \overline{\mathbf{u}}_n} = \mathbf{R}_n^T (\nabla \overline{\mathbf{u}}_n) \mathbf{R}$$





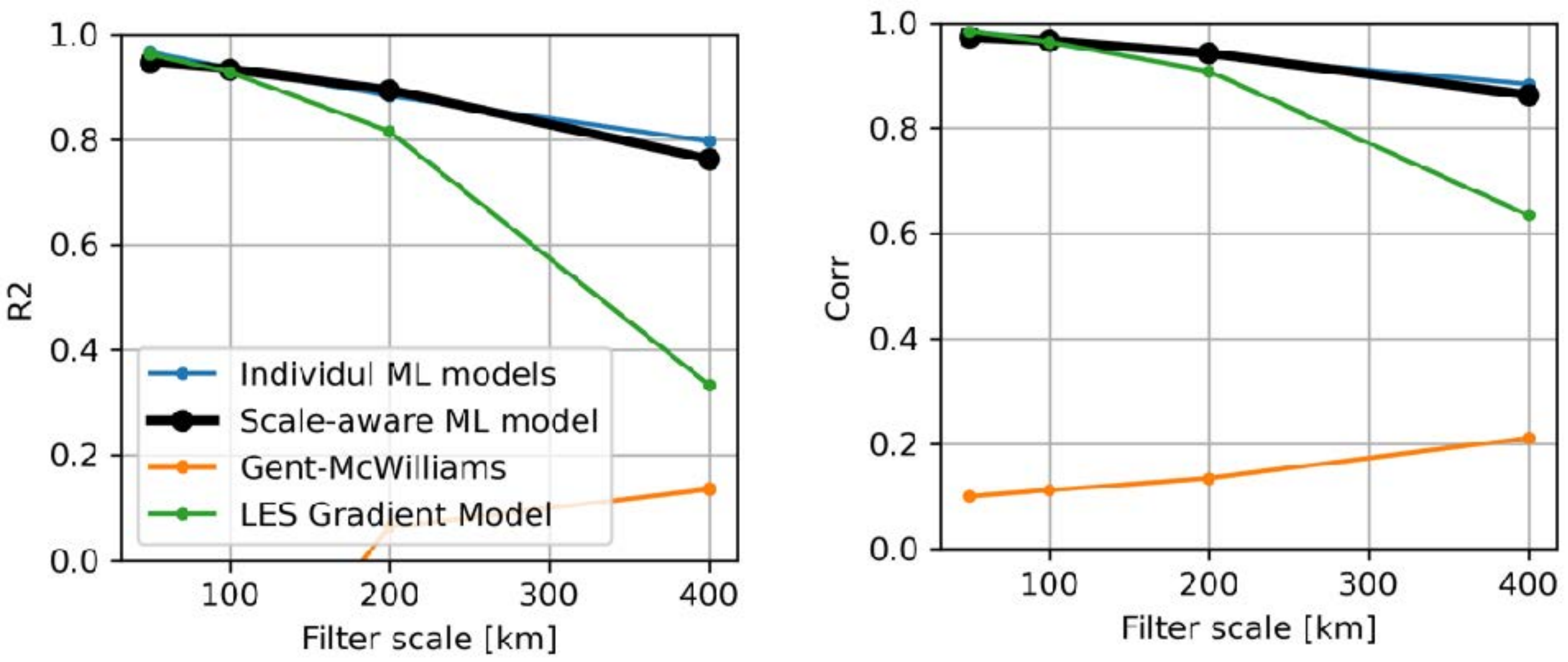
$$\mathbf{F}_n = f(\nabla \overline{\mathbf{u}}_n, \nabla \overline{h}_n, \triangle)$$

1. Non-dimensionalization
2. Range limited inputs
3. Separation of variables
4. **Coordinate-Invariance**
5. Scale-awareness
6. Lateral non-locality
7. Small ANN

$$\widetilde{\mathbf{F}}_{n,(i,j)} = \triangle^2 |\nabla \overline{\mathbf{u}}_{n,(i+p,j+p)}| f\left(\frac{\widetilde{\nabla \overline{\mathbf{u}}}_{n,(i+p,j+p)}}{|\nabla \overline{\mathbf{u}}_{n,(i+p,j+p)}|}, \frac{\widetilde{\nabla \overline{h}}_{n,(i+p,j+p)}}{|\nabla \overline{h}_{n,(i+p,j+p)}|}\right) g(|\nabla \overline{h}_{n,(i+p,j+p)}|).$$

ML Parameterization design

$$\mathbf{F}_n = f(\nabla \overline{\mathbf{u}}_n, \nabla \overline{h}_n, \triangle)$$



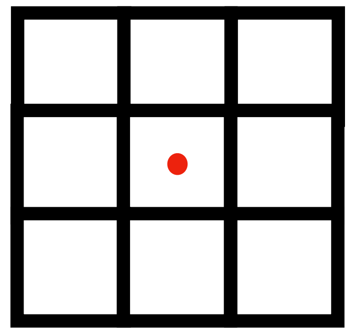
1. Non-dimensionalization
2. Range limited inputs
3. Separation of variables
4. Coordinate-Invariance
5. **Scale-awareness**
6. Lateral non-locality
7. Small ANN

$$\widetilde{\mathbf{F}}_{n,(i,j)} = \triangle^2 |\nabla \overline{\mathbf{u}}_{n,(i+p,j+p)}| f\left(\frac{\widetilde{\nabla \overline{\mathbf{u}}}_{n,(i+p,j+p)}}{|\nabla \overline{\mathbf{u}}_{n,(i+p,j+p)}|}, \frac{\widetilde{\nabla \overline{h}}_{n,(i+p,j+p)}}{|\nabla \overline{h}_{n,(i+p,j+p)}|}\right) g(|\nabla \overline{h}_{n,(i+p,j+p)}|).$$

ML Parameterization design

$$\mathbf{F}_n = f(\nabla \overline{\mathbf{u}}_n, \nabla \overline{h}_n, \triangle)$$

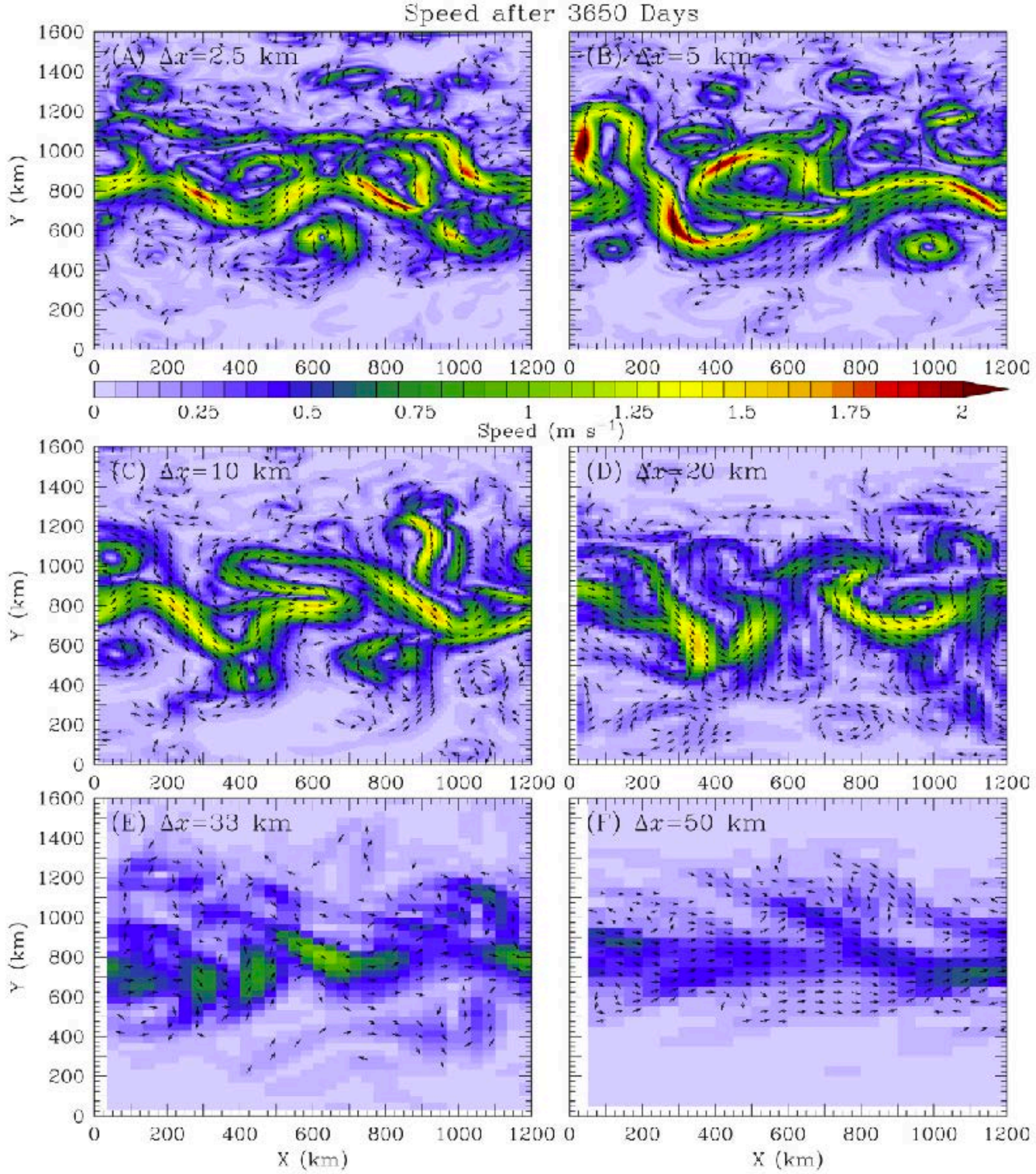
Col:Scale Row:Stencil Width	50km	100km	200km	400km
1X1	0.90	0.91	0.85	0.77
3X3	0.96	0.96	0.93	0.90
5X5	0.96	0.96	0.93	0.91

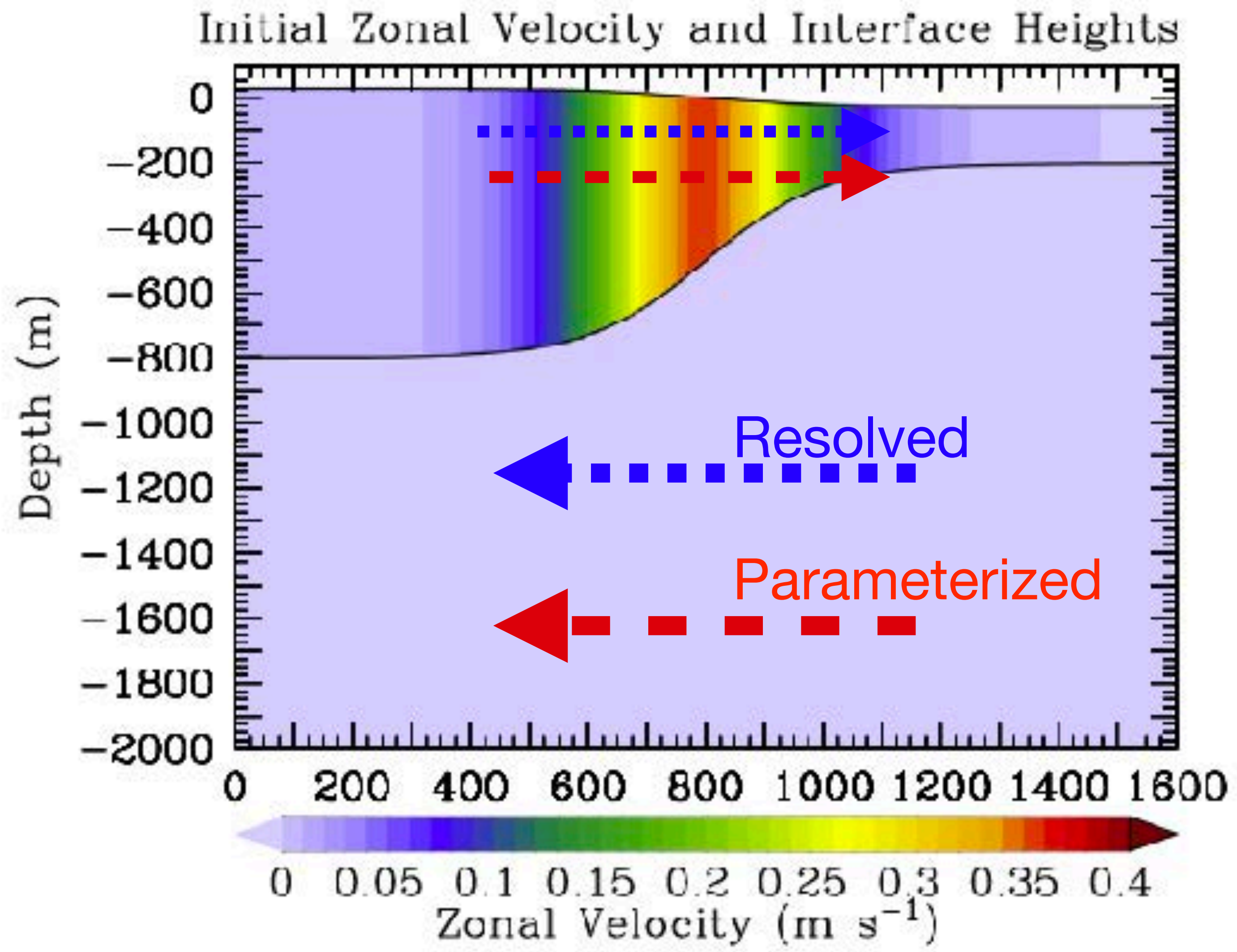


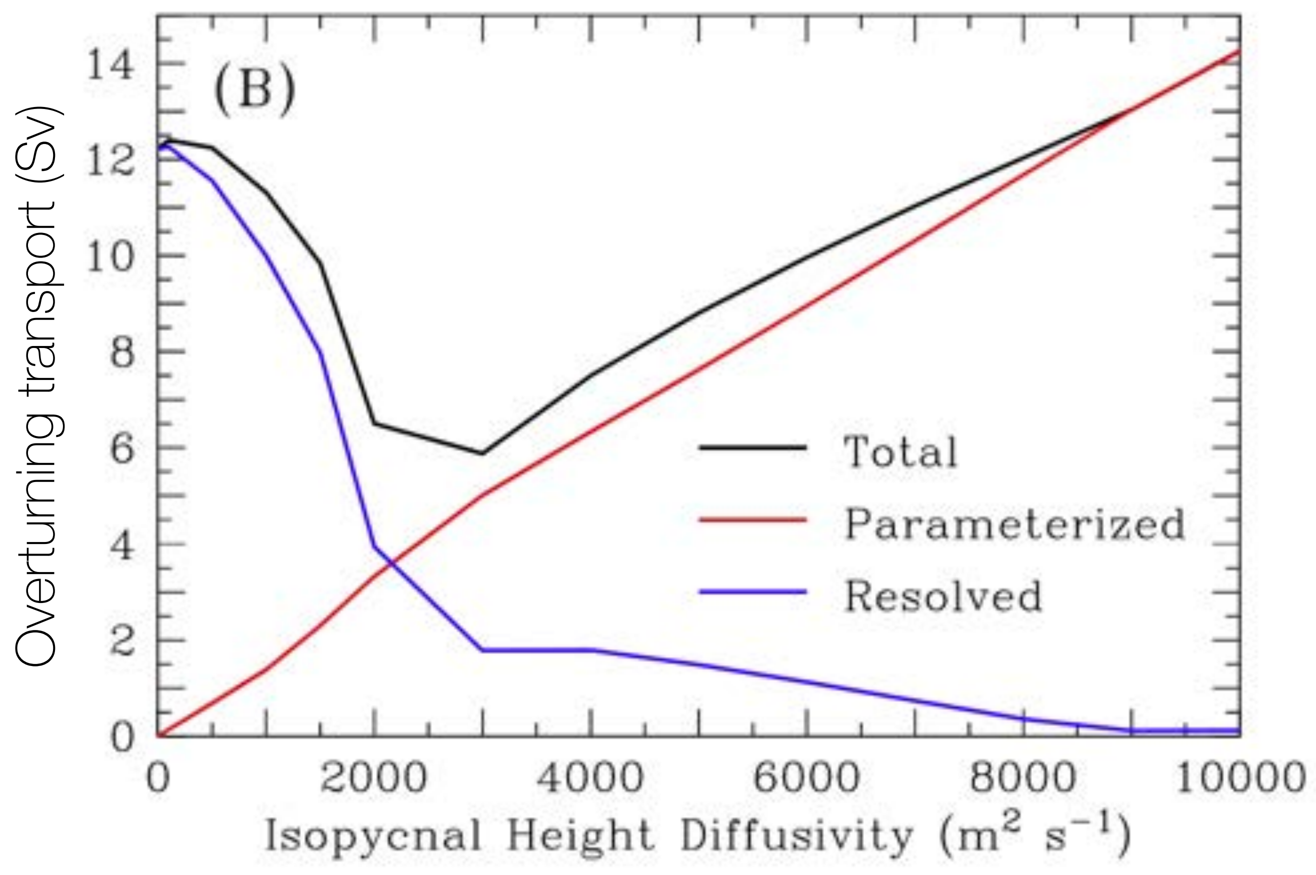
1. Non-dimensionalization
2. Range limited inputs
3. Separation of variables
4. Coordinate-Invariance
5. Scale-awareness
6. **Lateral non-locality**
7. Small ANN

$$\widetilde{\mathbf{F}}_{n,(i,j)} = \triangle^2 |\nabla \overline{\mathbf{u}}_{n,(i+p,j+p)}| f\left(\frac{\widetilde{\nabla \overline{\mathbf{u}}}_{n,(i+p,j+p)}}{|\nabla \overline{\mathbf{u}}_{n,(i+p,j+p)}|}, \frac{\widetilde{\nabla \overline{h}}_{n,(i+p,j+p)}}{|\nabla \overline{h}_{n,(i+p,j+p)}|}\right) g(|\nabla \overline{h}_{n,(i+p,j+p)}|).$$

Some online evaluation - Phillips 2 Layer



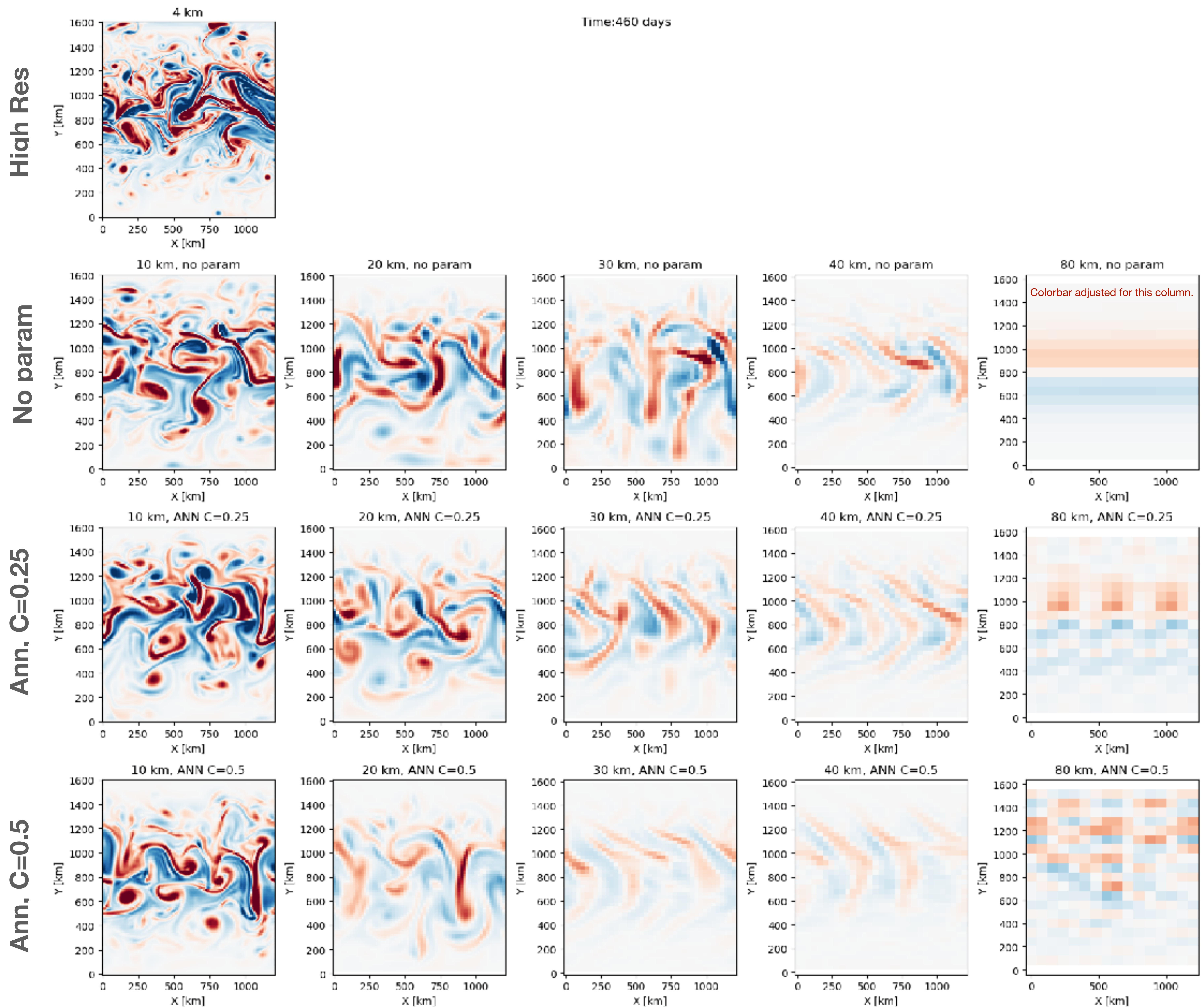




Total overturning responds non-linearly to Gent- McWilliams parameterization.

Hallberg 2013

Some online evaluation - visuals, **stable** 😰



Stable over large range of resolutions and coefficients! (No filters added).

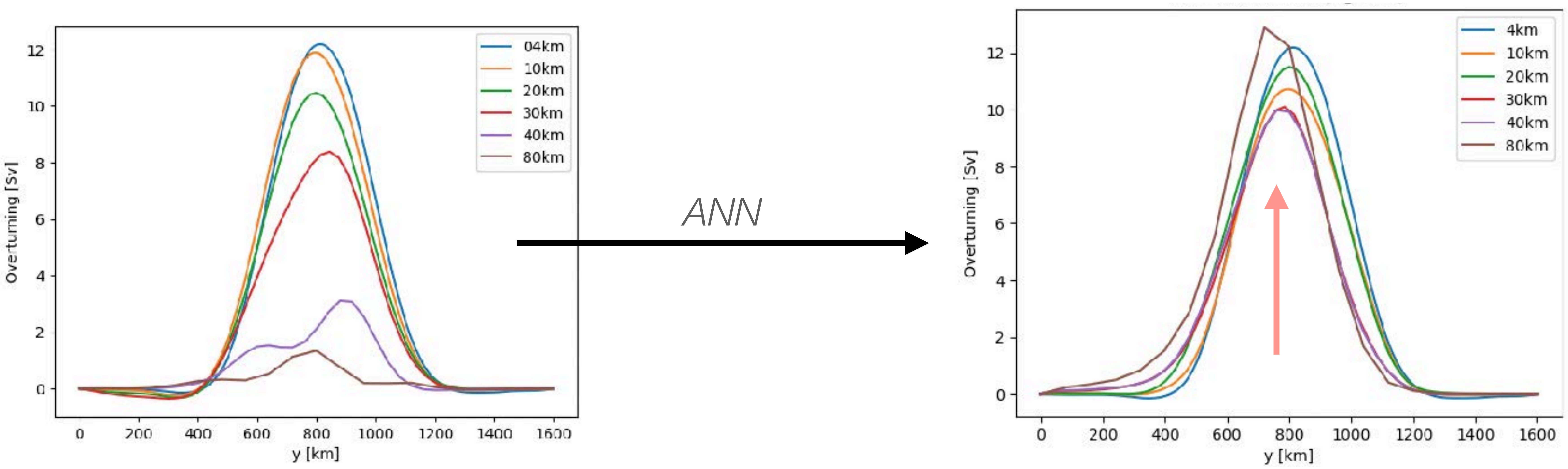
Resolved variability is only slightly damped.

Even has some skill when absolutely no eddies are resolved

Some online evaluation - Bulk

ML parameterization increases overturning.

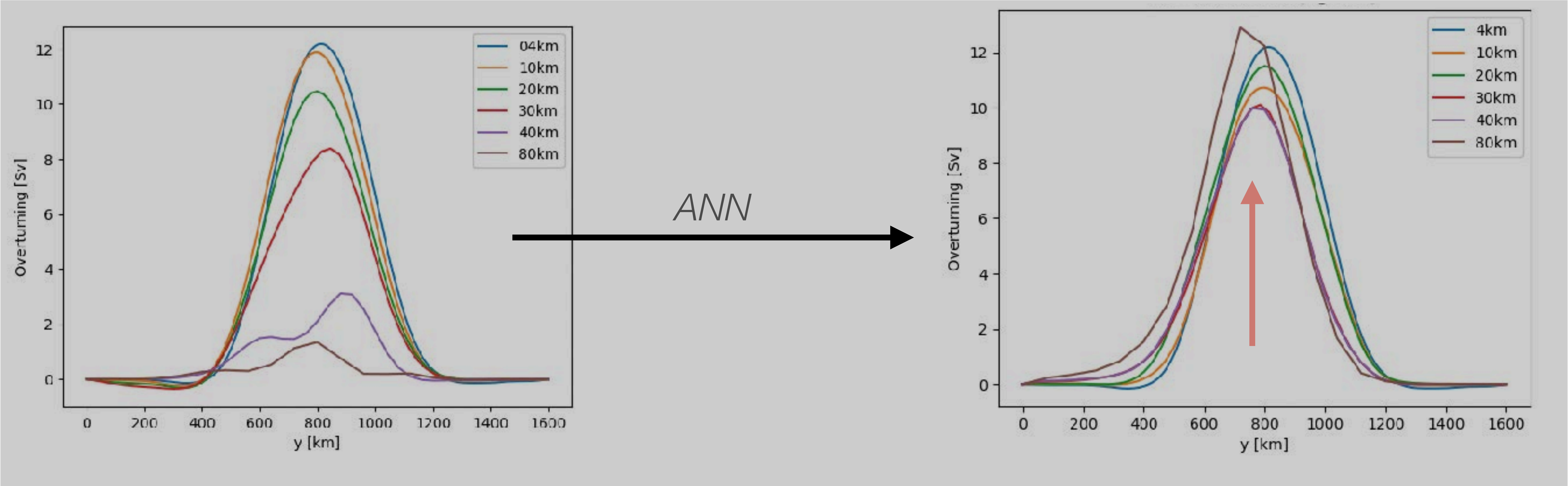
😅



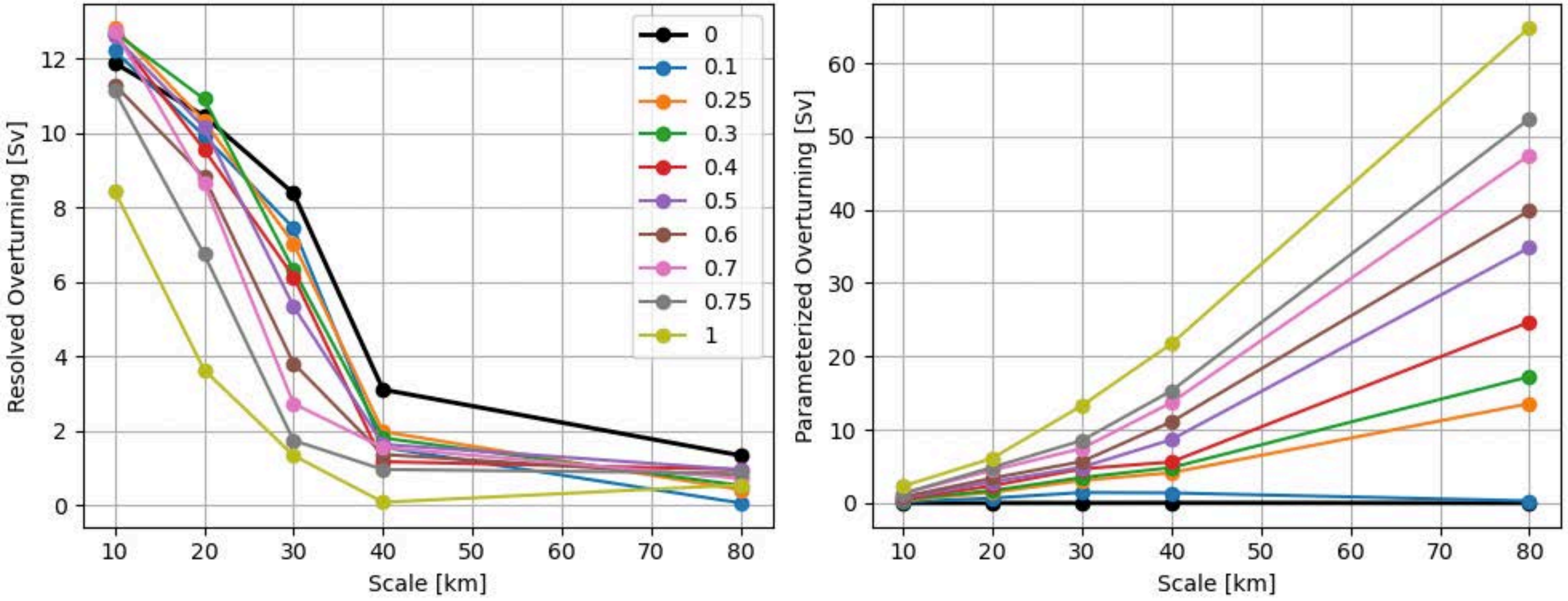
Some online evaluation - Bulk

ML parameterization increases overturning.

😅



- Parameterized flow changes ~non-linearly with tuning coefficient.
- Response of resolved flow to tuning coefficient is
  - Resolution dependent
  - Non-linear and non monotonic
- Resolved eddies are damped, but not as bad as if we used Gent-McWilliams.



Conclusions

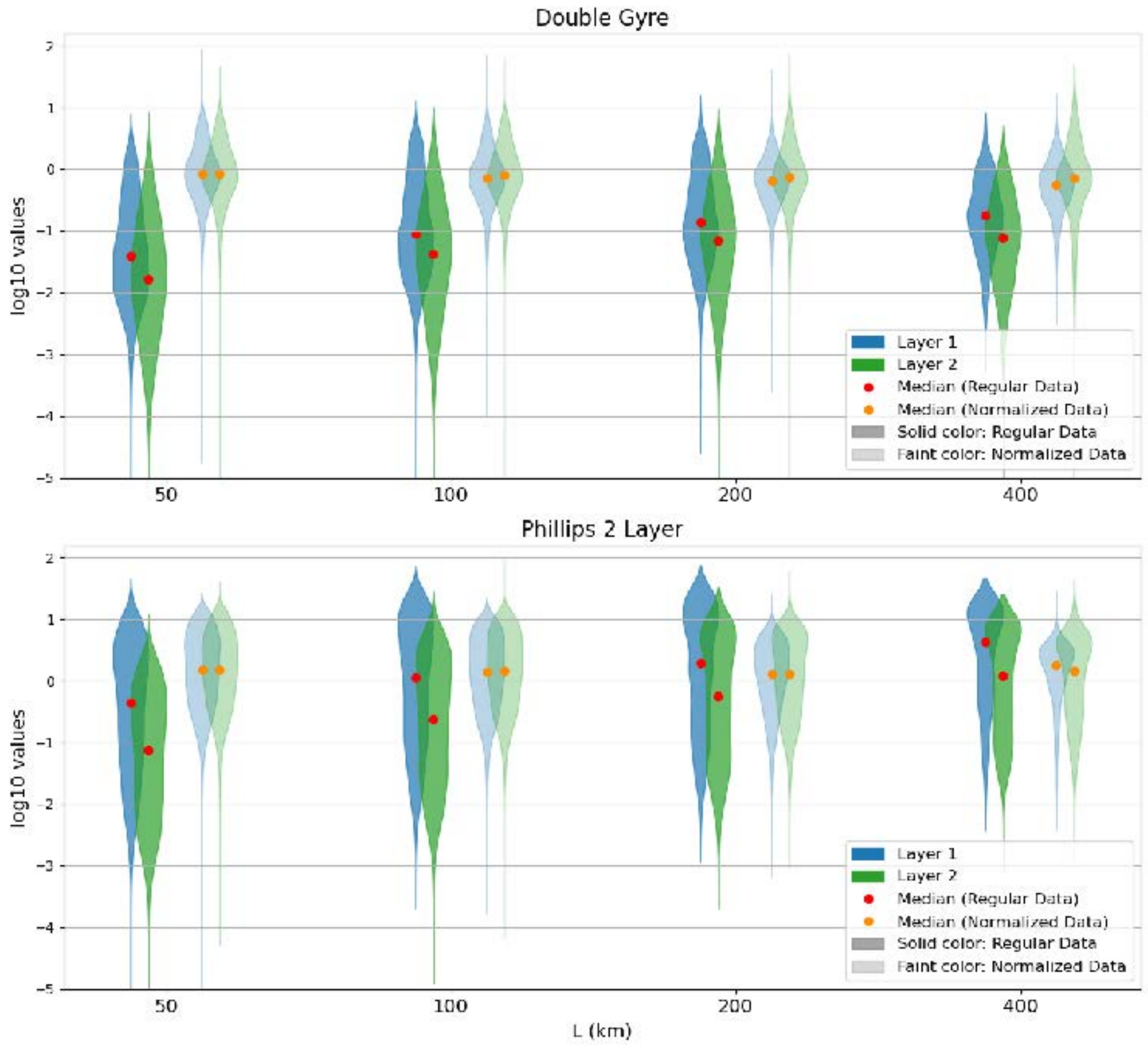
- Developed a parameterization based on ML and intuition/physics.
- Behavior of the function, and the response of the system to it, need to be studied empirically.
- Response of the resolved flow is stable and favorable (less dissipative than GM), but tuning is required.
- Some generalization when some eddies are resolved, but not for the case non-eddying resolutions.

Future Work

- Does training on larger/more diverse datasets help generalization?
- Merging thickness and momentum parameterizations (**Kelsey Everard next talk!**)
- How to ensure that ML model learn bulk response (e.g. APE reduction) to bulk parameters (Rd, f, beta, etc)?
- Discretization/Numerics? Theory? …

End

ML Parameterization design



$$|\mathbf{F}_n| \rightarrow \frac{|\mathbf{F}_n|}{\triangle^2 |\nabla \overline{\mathbf{u}}_n|}$$

$$\mathbf{F}_n = f(\nabla \overline{\mathbf{u}}_n, \nabla \overline{h}_n, \triangle)$$

1. **Non-dimensionalization**
2. Range limited inputs
3. Separation of variables
4. Coordinate-Invariance
5. Scale-awareness
6. Lateral non-locality
7. Small ANN

$$\widetilde{\mathbf{F}}_{n,(i,j)} = \triangle^2 |\nabla \overline{\mathbf{u}}_{n,(i+p,j+p)}| f\left(\frac{\widetilde{\nabla \overline{\mathbf{u}}}_{n,(i+p,j+p)}}{|\nabla \overline{\mathbf{u}}_{n,(i+p,j+p)}|}, \frac{\widetilde{\nabla \overline{h}}_{n,(i+p,j+p)}}{|\nabla \overline{h}_{n,(i+p,j+p)}|}\right) g(|\nabla \overline{h}_{n,(i+p,j+p)}|).$$

Does velocity grad magnitude scale with sub-grid KE?

•

Dataset design aspects

- GCM filters for spatial filtering
- FGR ~ 5 (maybe this helps with learning)
-

ML Design Aspects

- Sub-sample the grid to keep data uniform.
- Architecture:
  - Simple ANN, JAX, FLAX
  - 2 layers
  - N neurons per layer.
- Optimizer:
  -

Topography or interfaces

-