Characterizing patterns in HPC simulations using AI driven image recognition/categorization

+ Dynostar grand challenge
+ early works for AI for small-scale turbulence closure

P. Tremblin 3/10/24 NumPEx AI4HPC@Exascale

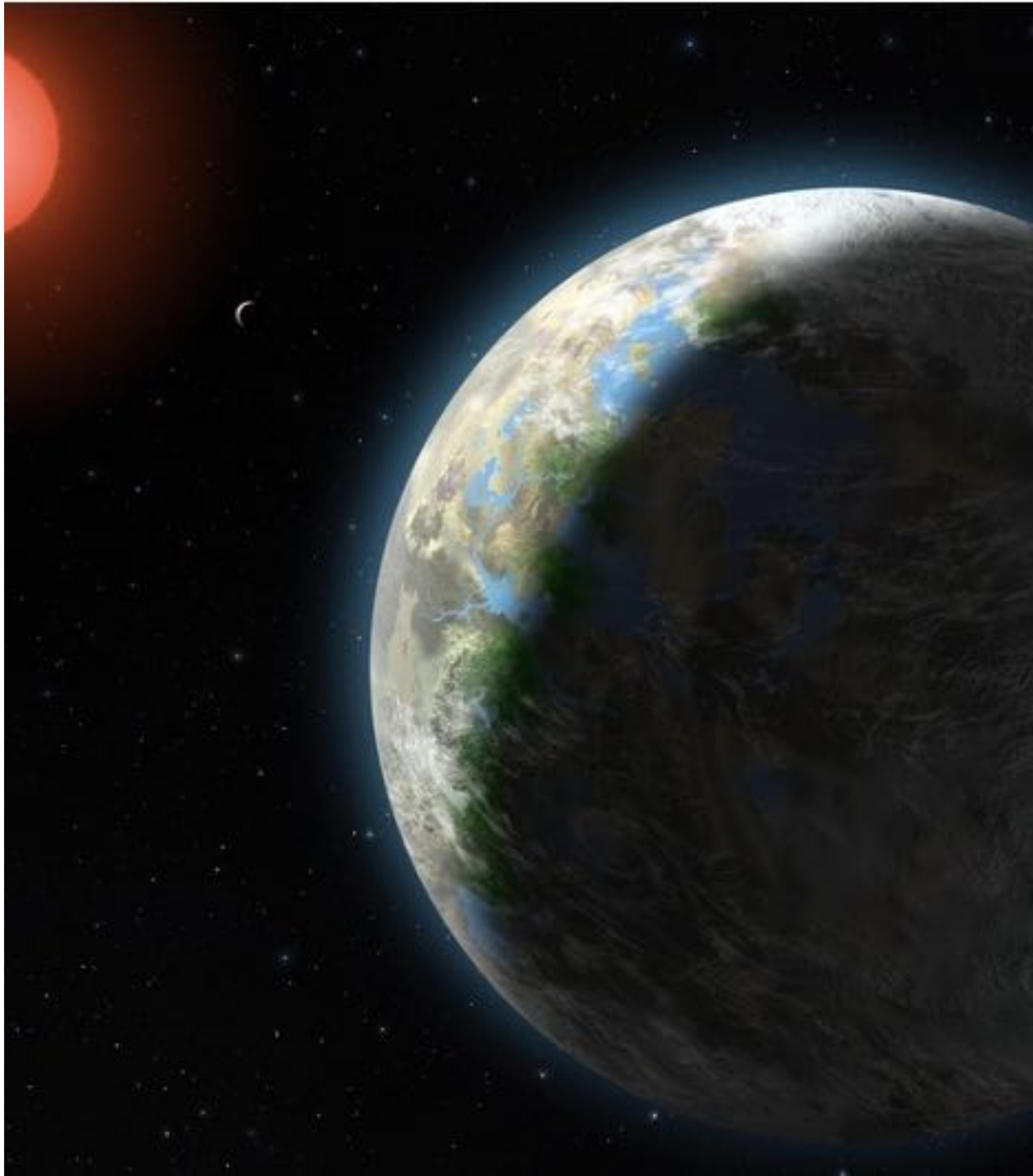
European Research Council
Established by the European Commission

MAISON DE LA SIMULATION

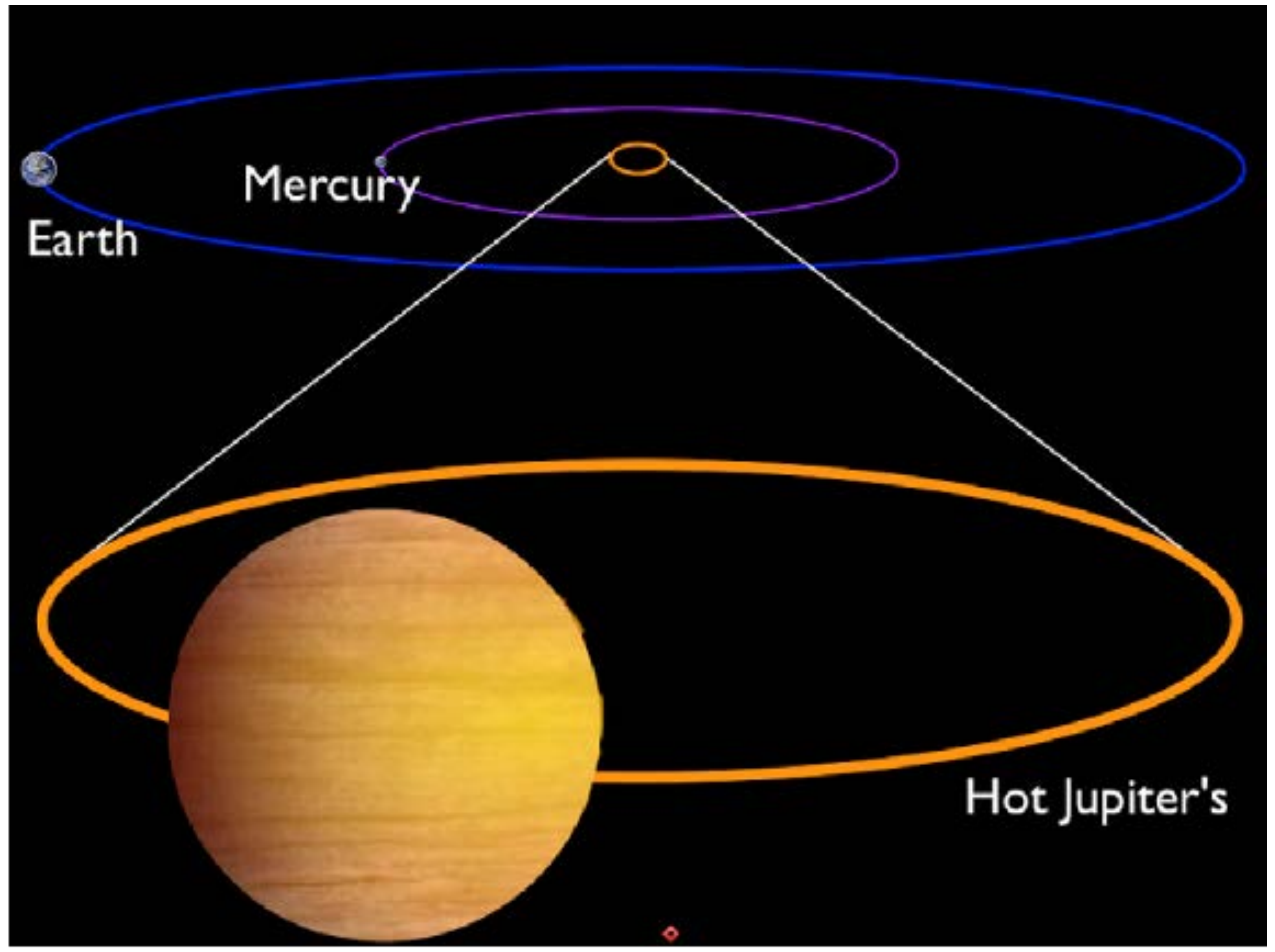
Use case

Exoplanets

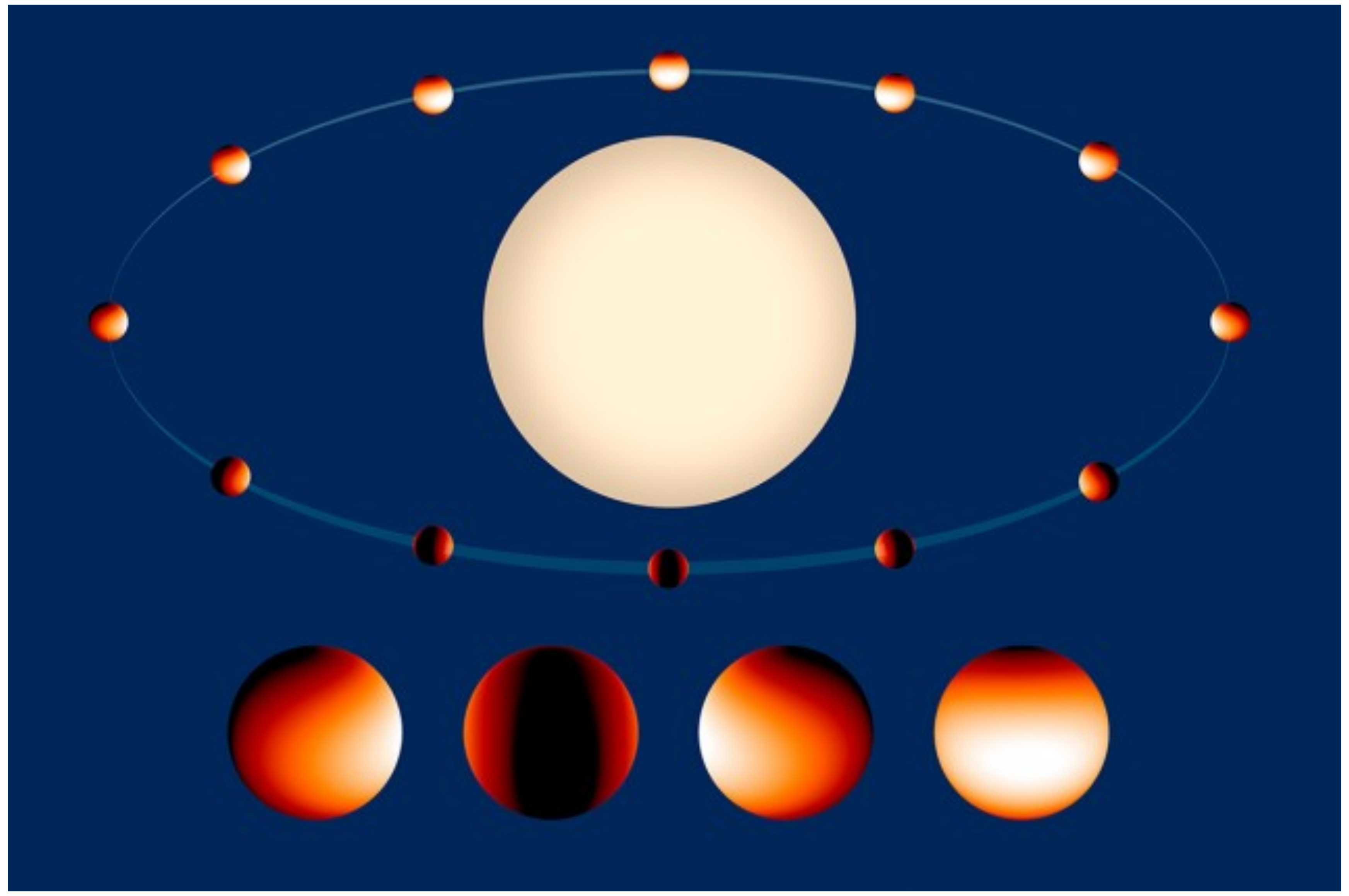
- A bit of history:
- In 1995 first observation of an exoplanet using radial velocities (Nobel prize in 2019, M. Mayor and D. Queloz)
- In 1999 first observation using the transit method (showing inflated radii)
- These exoplanets are similar to Jupiter but well inside the orbit of Mercury: Hot Jupiters



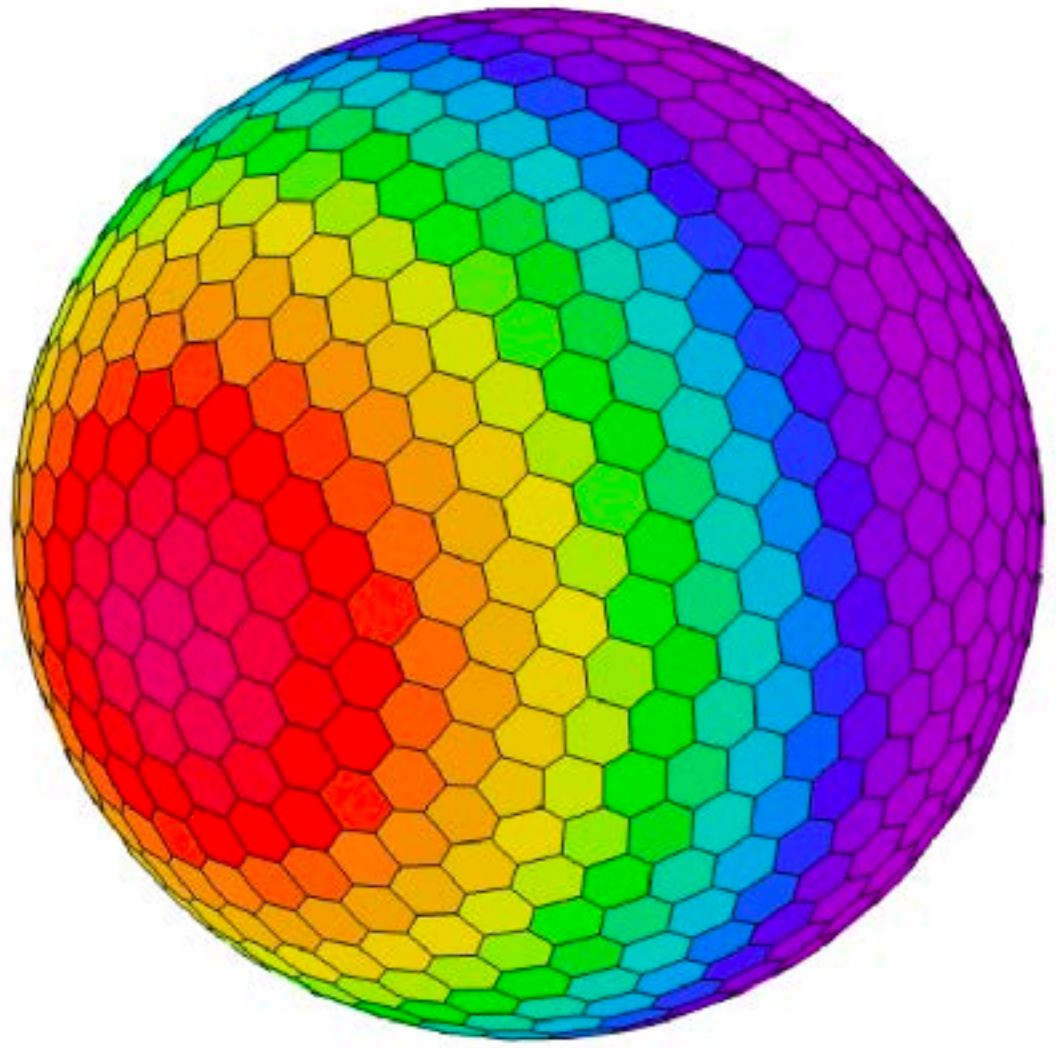
Close -> Hot



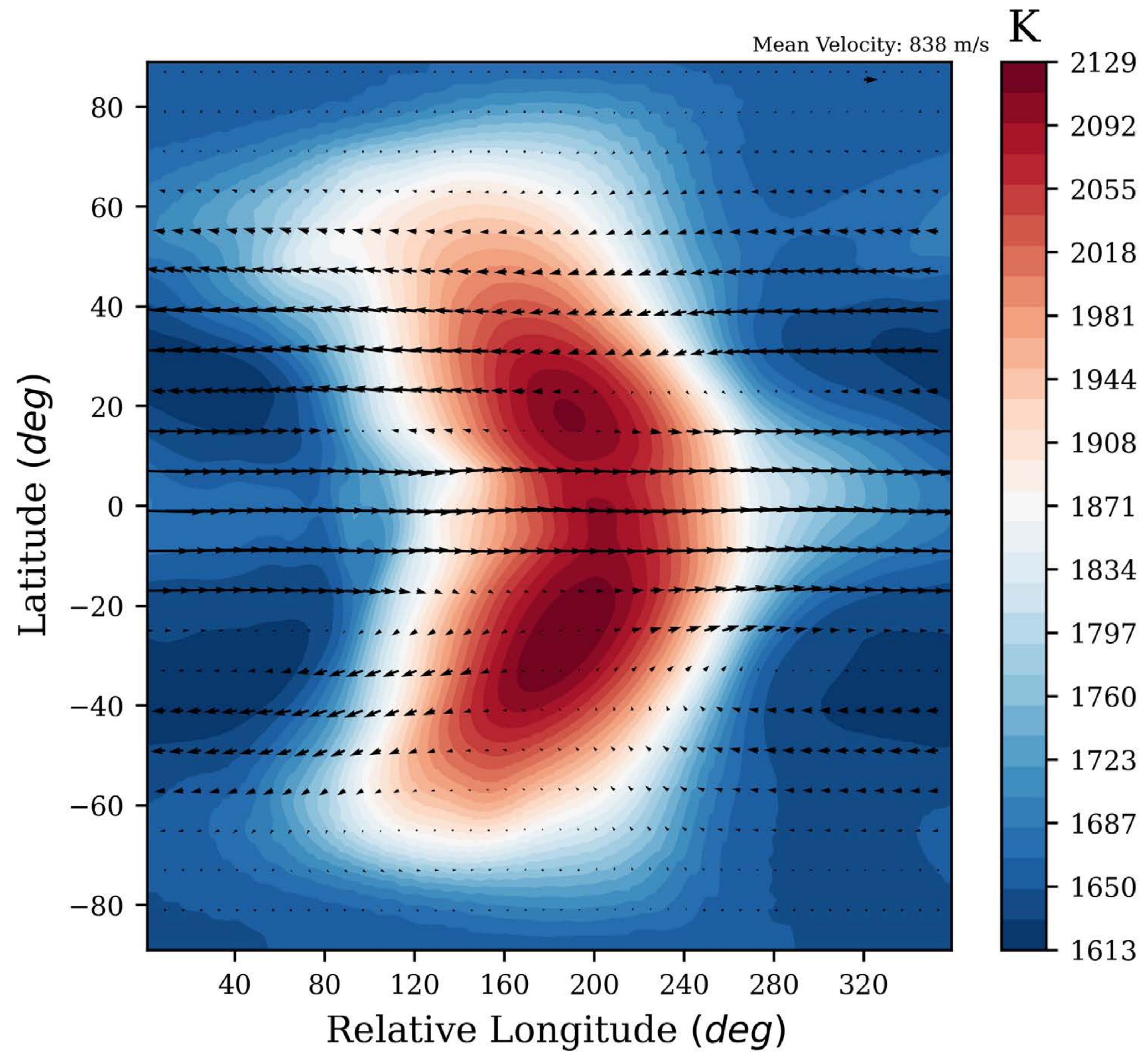
Synchronous rotation ->
Permanent hot dayside and cold nightside



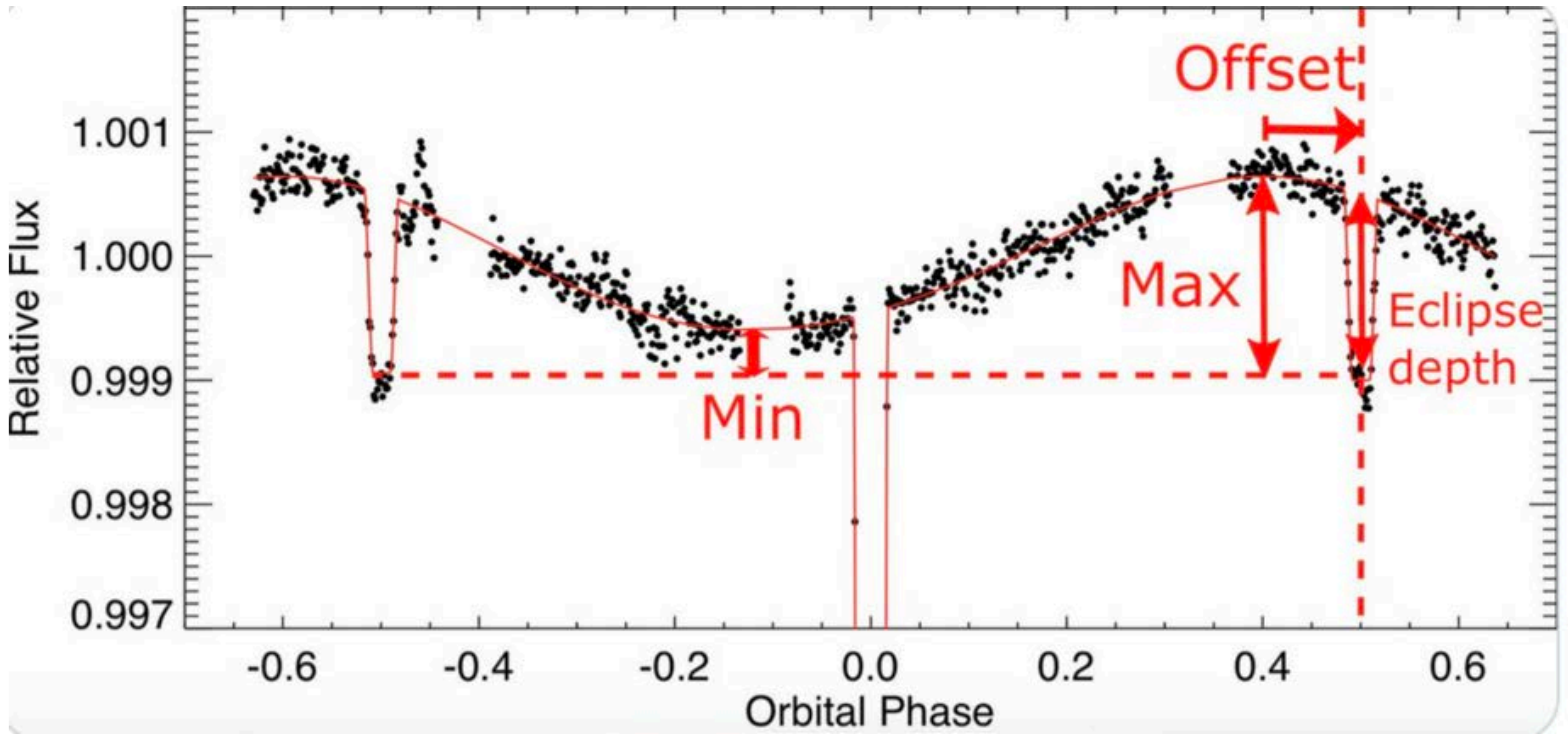
-> atmospheric circulation



3D Global Circulation Models (GCM)
DYNAMICO (Dubos et al. 2015)



Hotspot with a zonal jet -> offset



Observed with phase curves (e.g. James Webb Space Telescope)

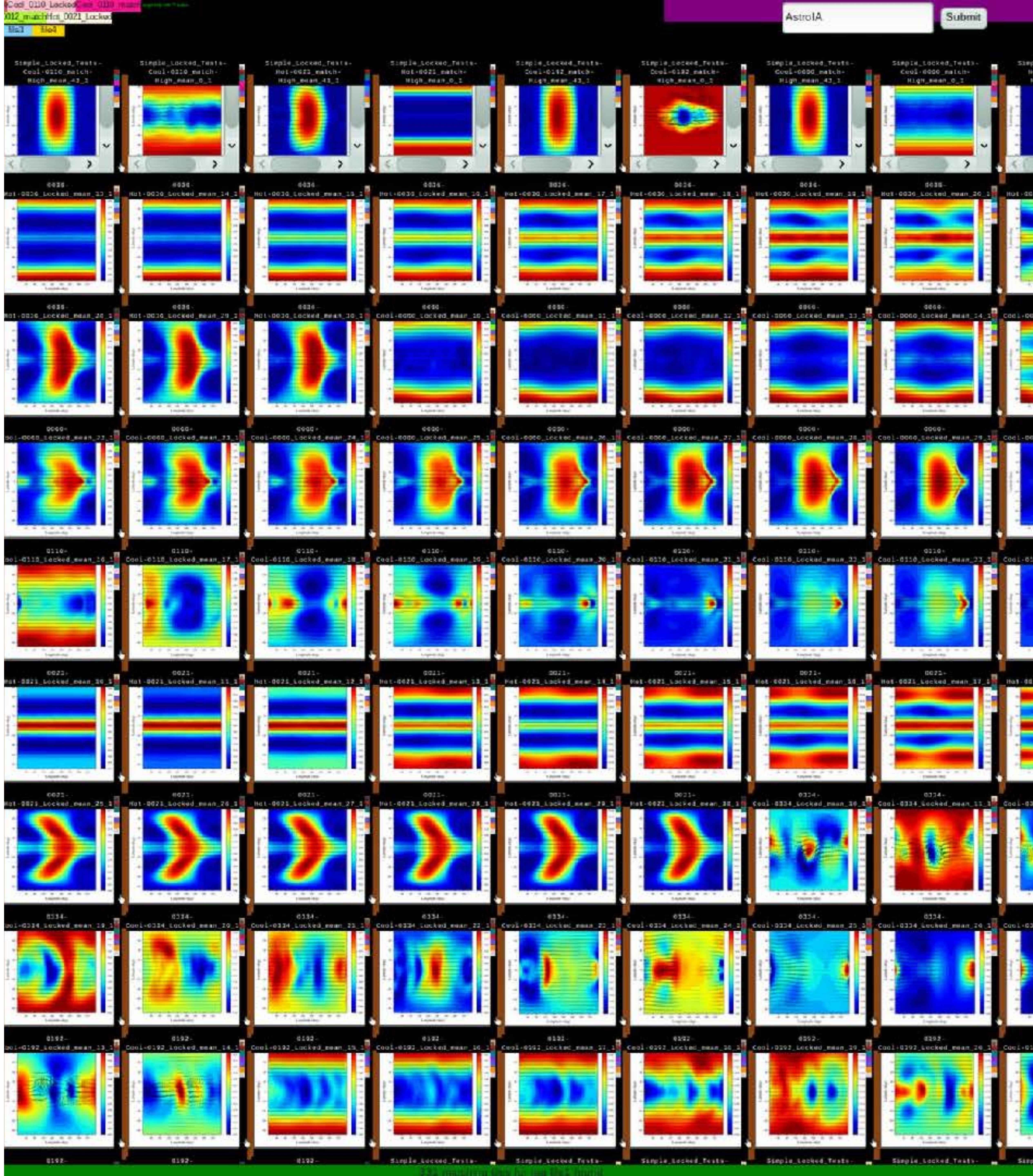
Simulation dataset

3D GCM simulations

- Initial setup of Sainsbury-Martinez et al. 2019 to provide a solution for the inflated radii of Hot Jupiters

- Parametric study to explore the effect of rotation rate on the deep circulation

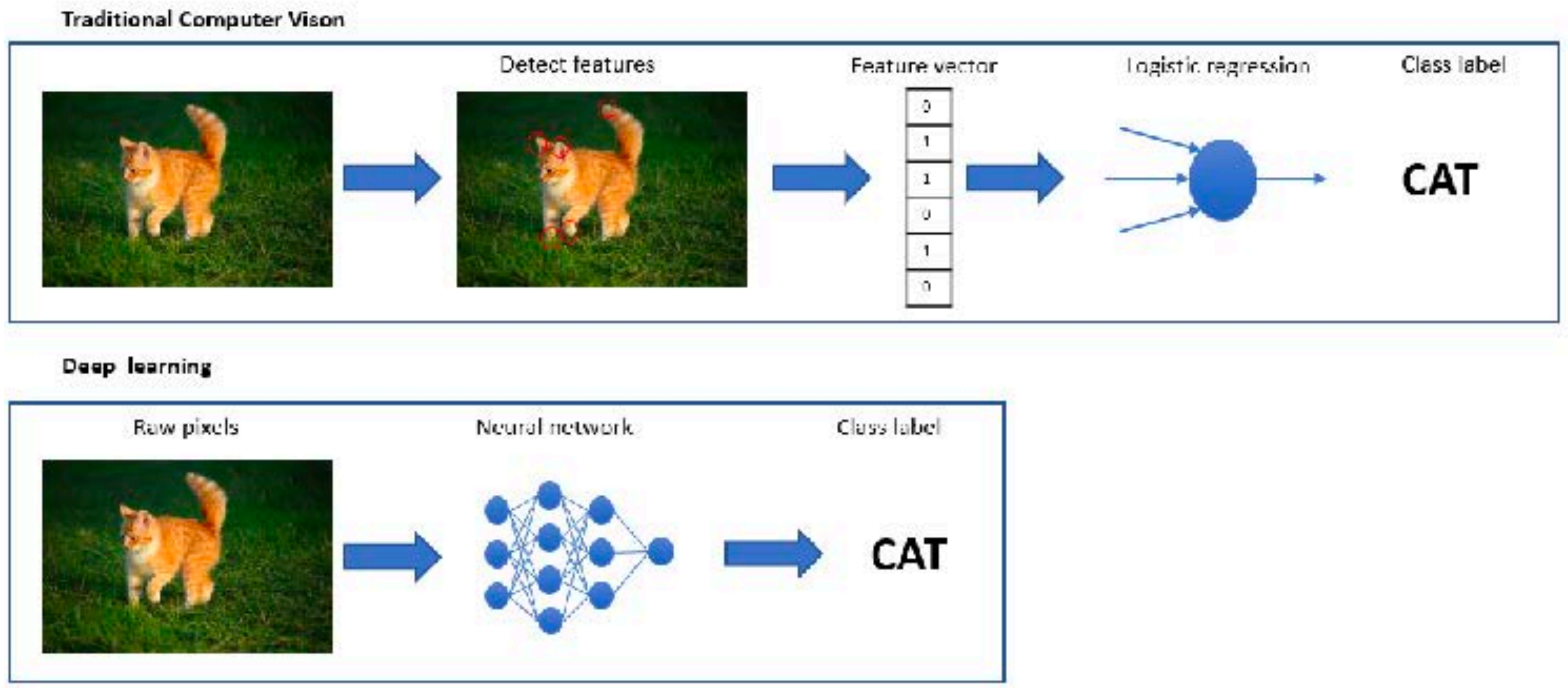
- Create a dataset of 128x128 2D images at a given time and pressure level in the atmosphere

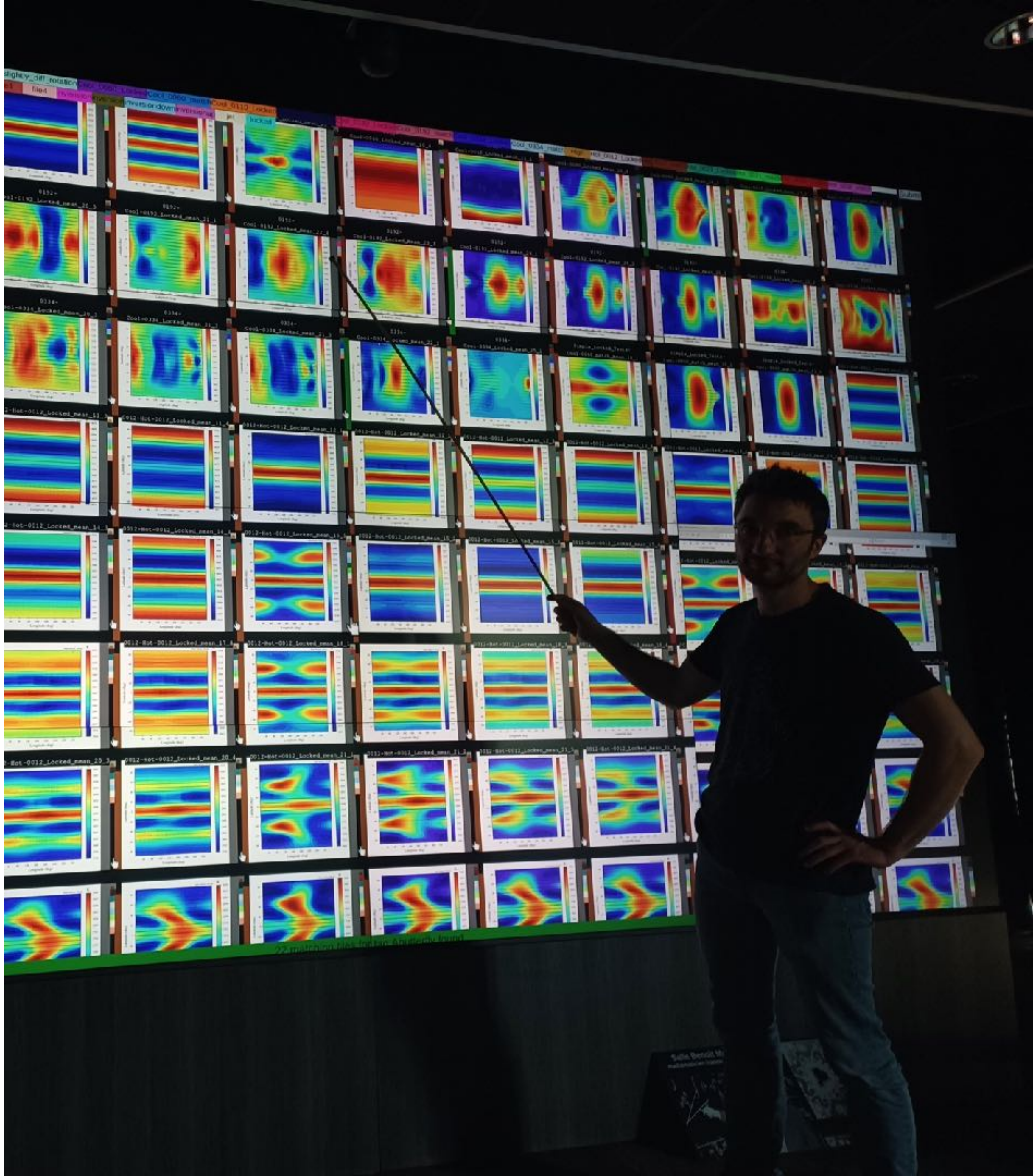


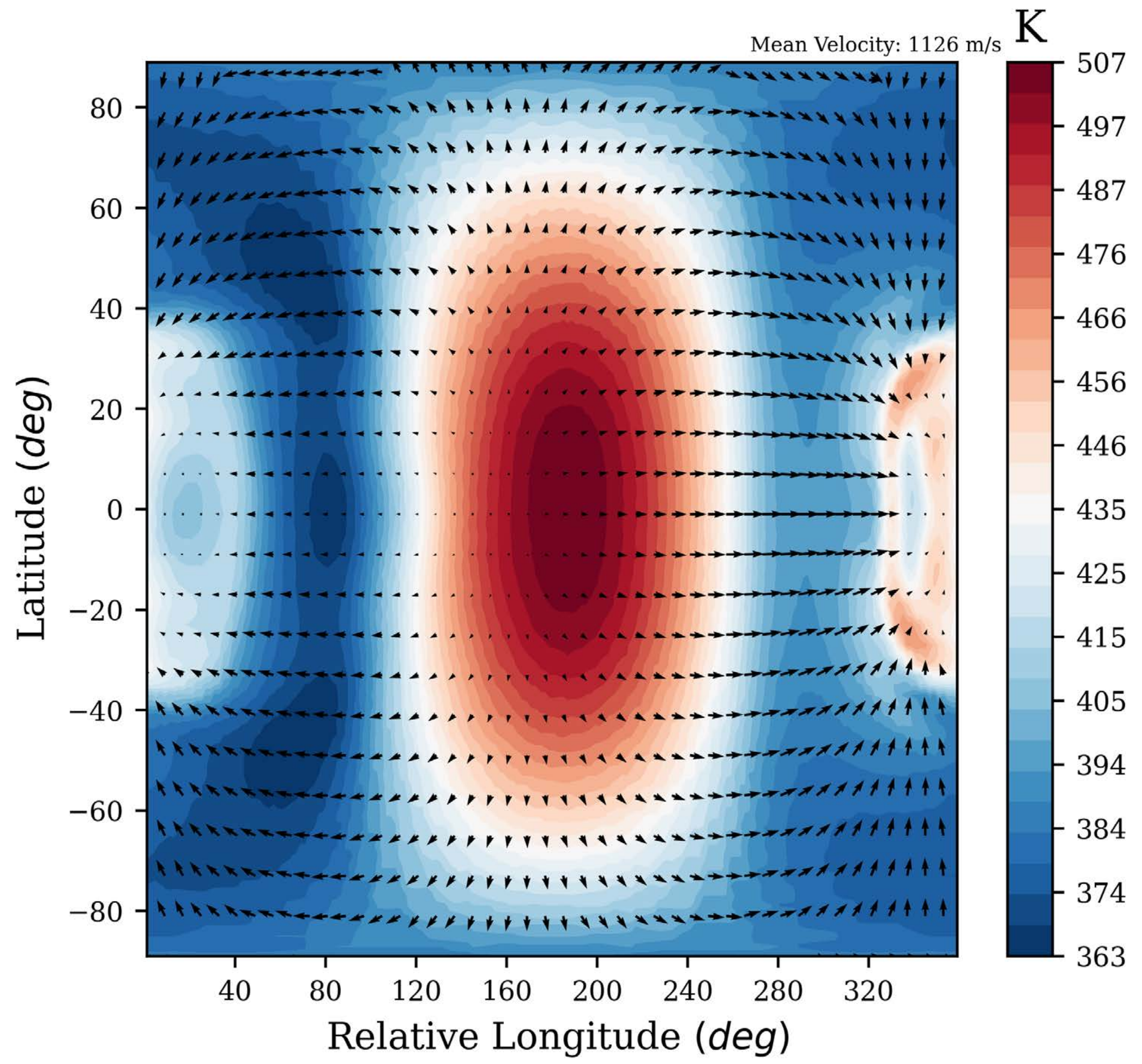
Supervised classification

Tiled visualization for display walls

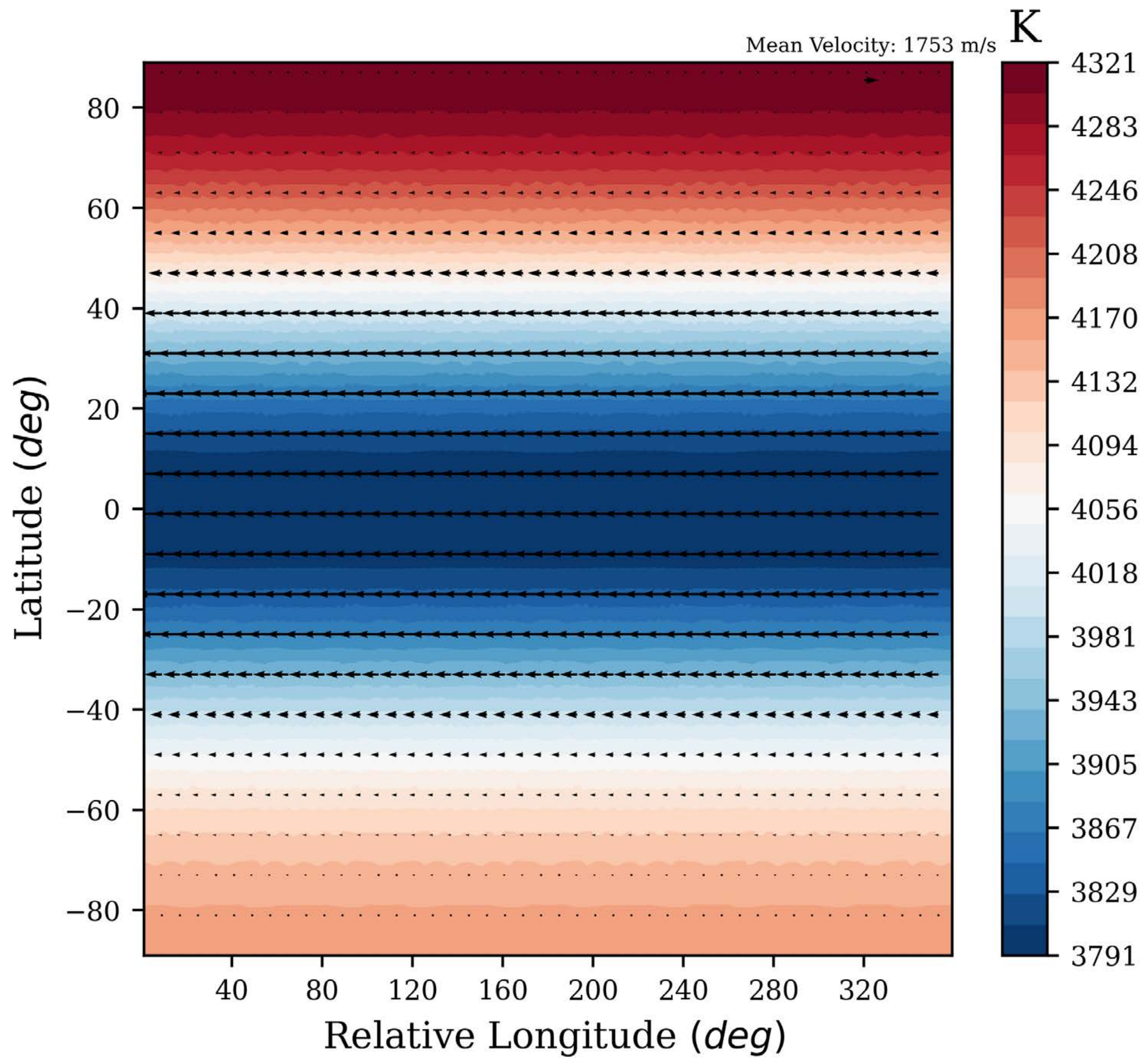
- Classification relying on AI-driven computer vision (CNN)
- Supervised learning using TiledViz tool developed at MdlS by M. Mancip
- 4 categories, ~500 images in the training dataset



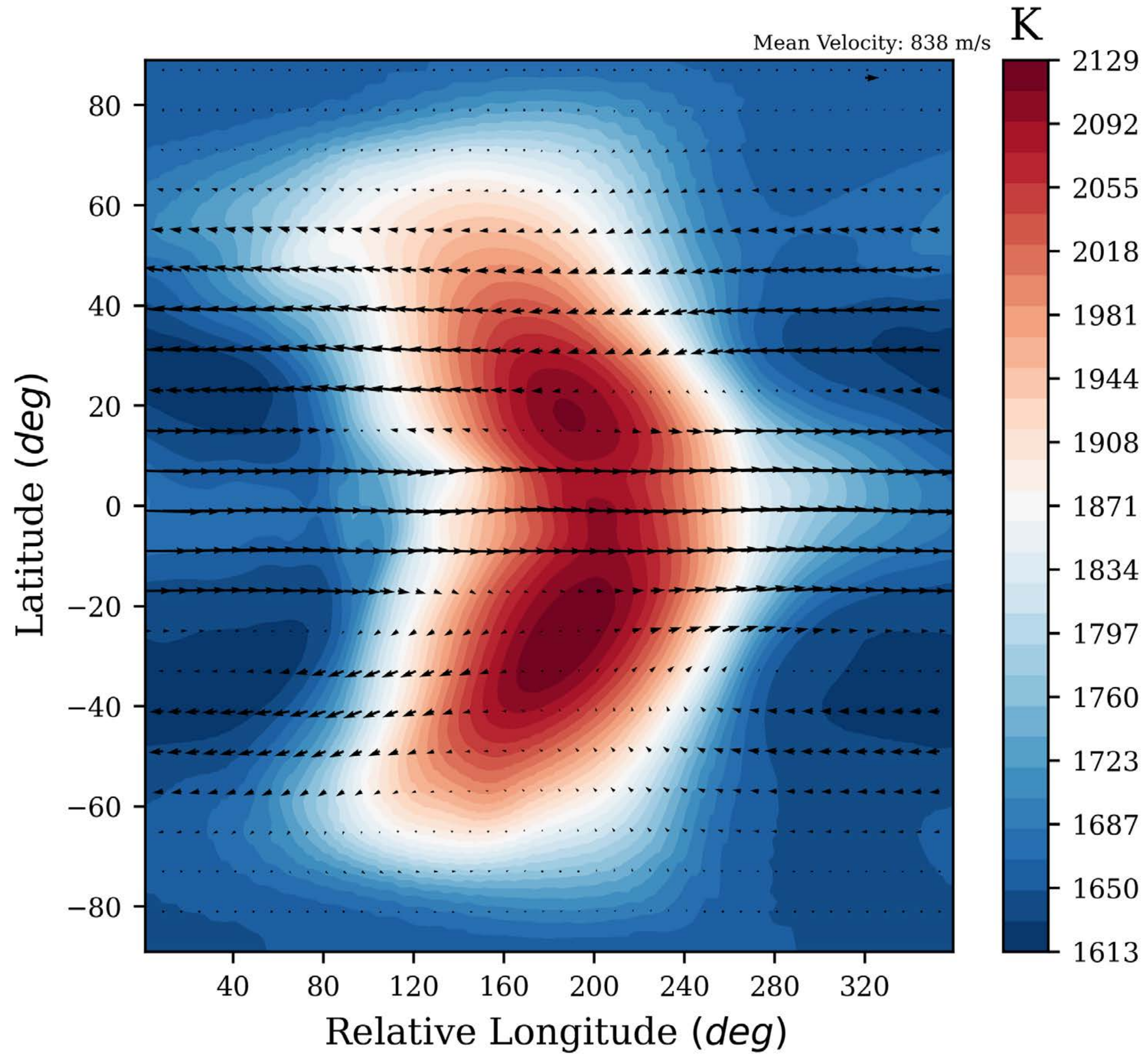




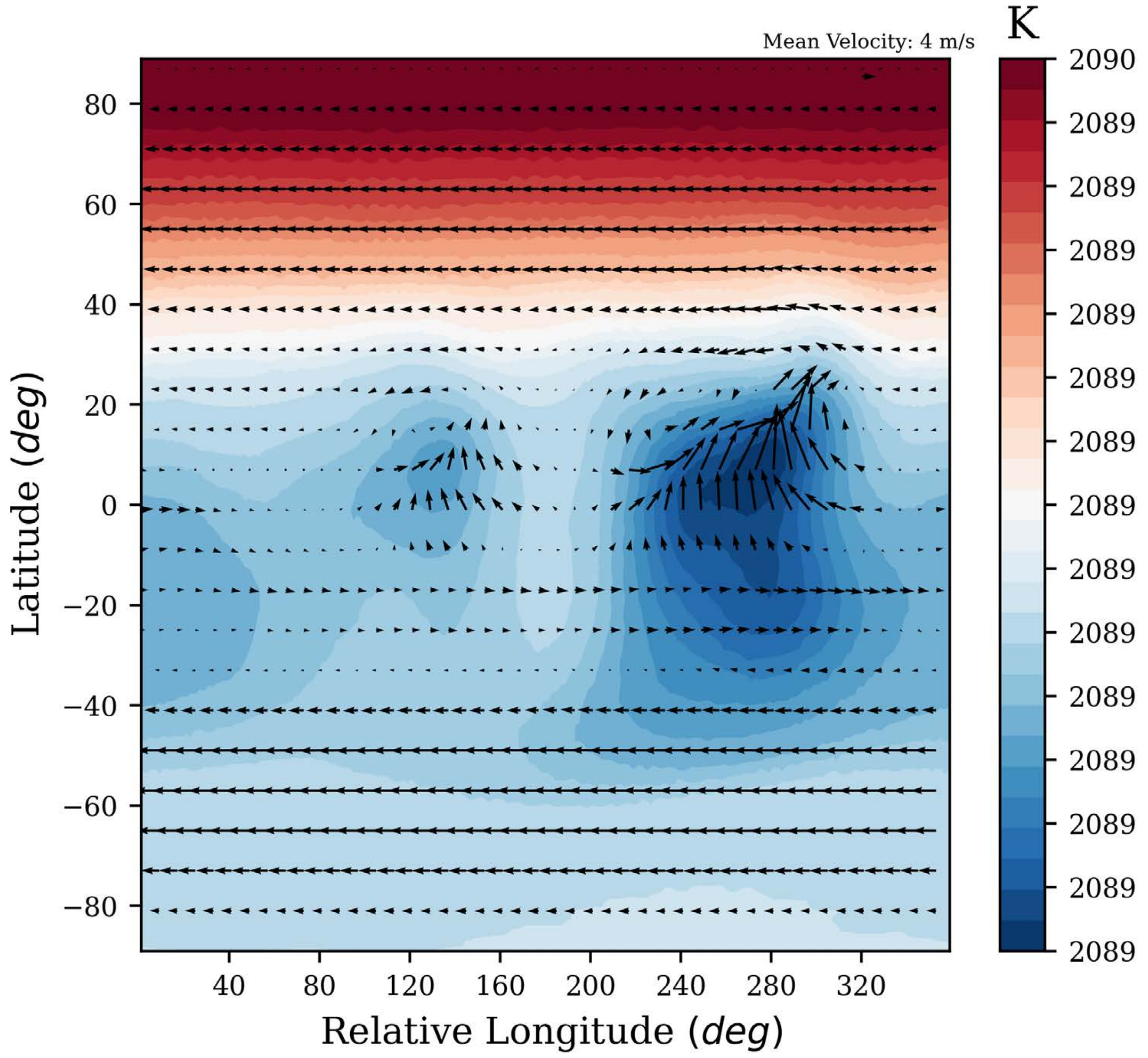
« Locked »



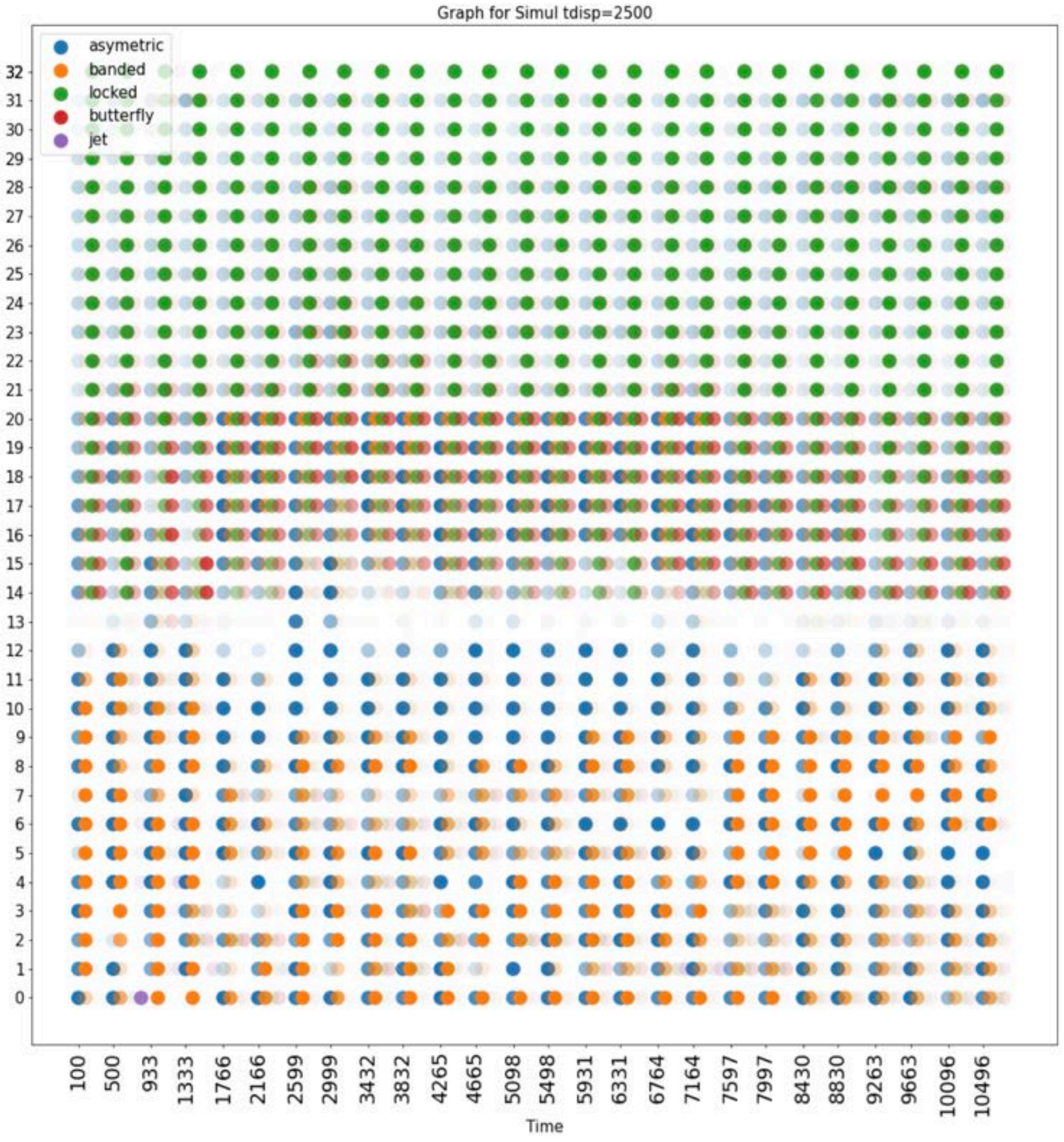
« Banded »



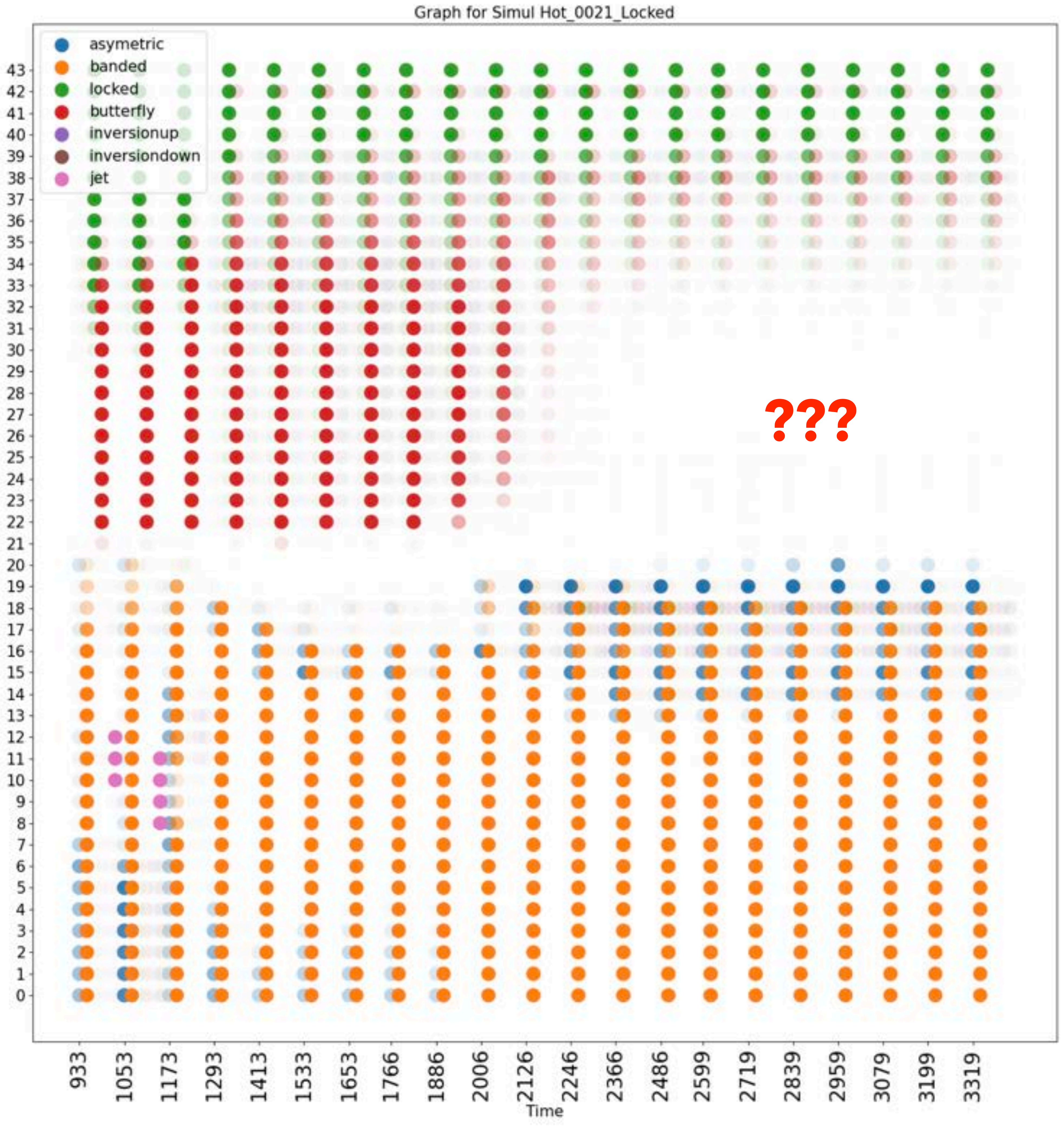
« Butterfly »



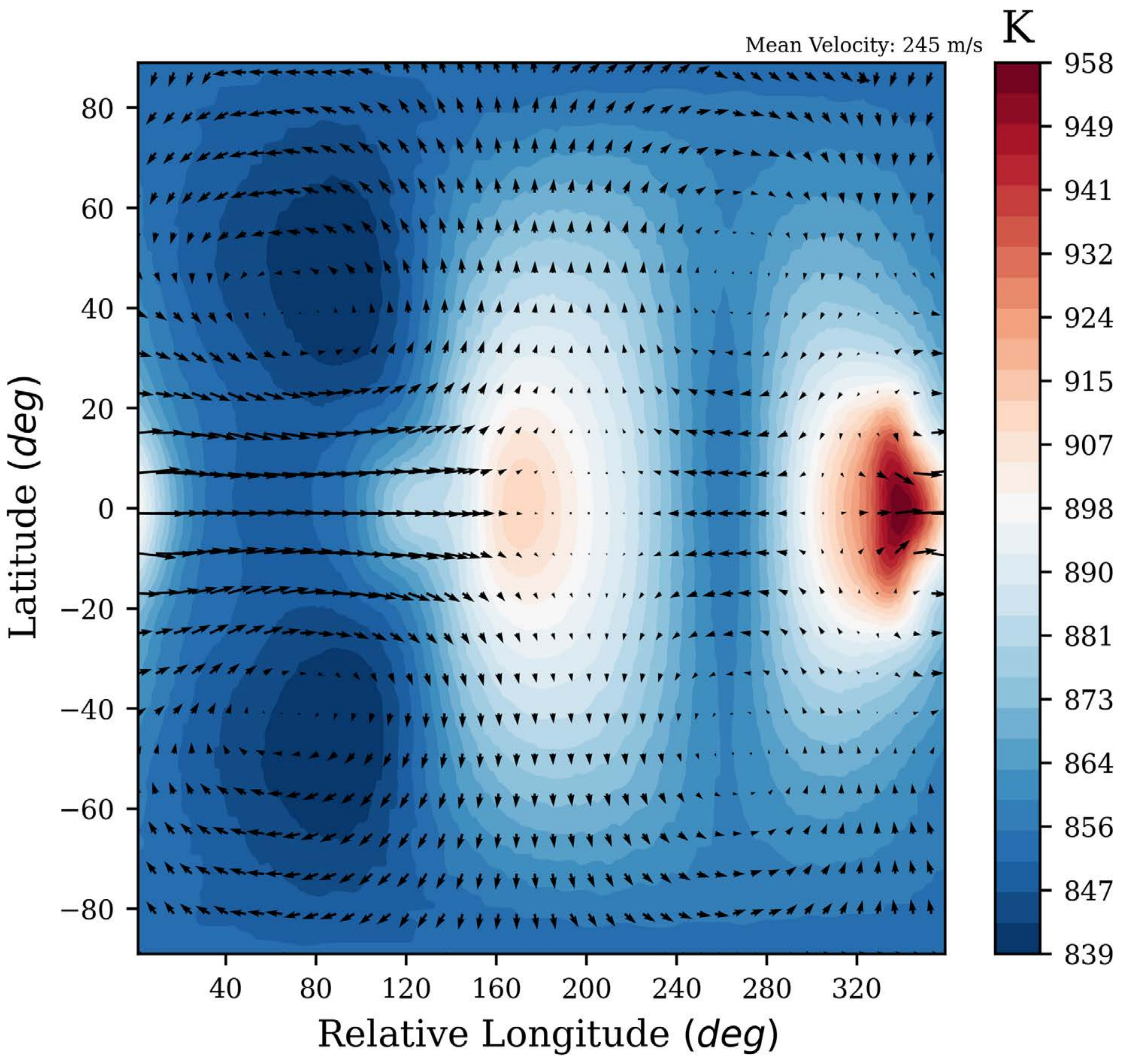
« Asymmetric »



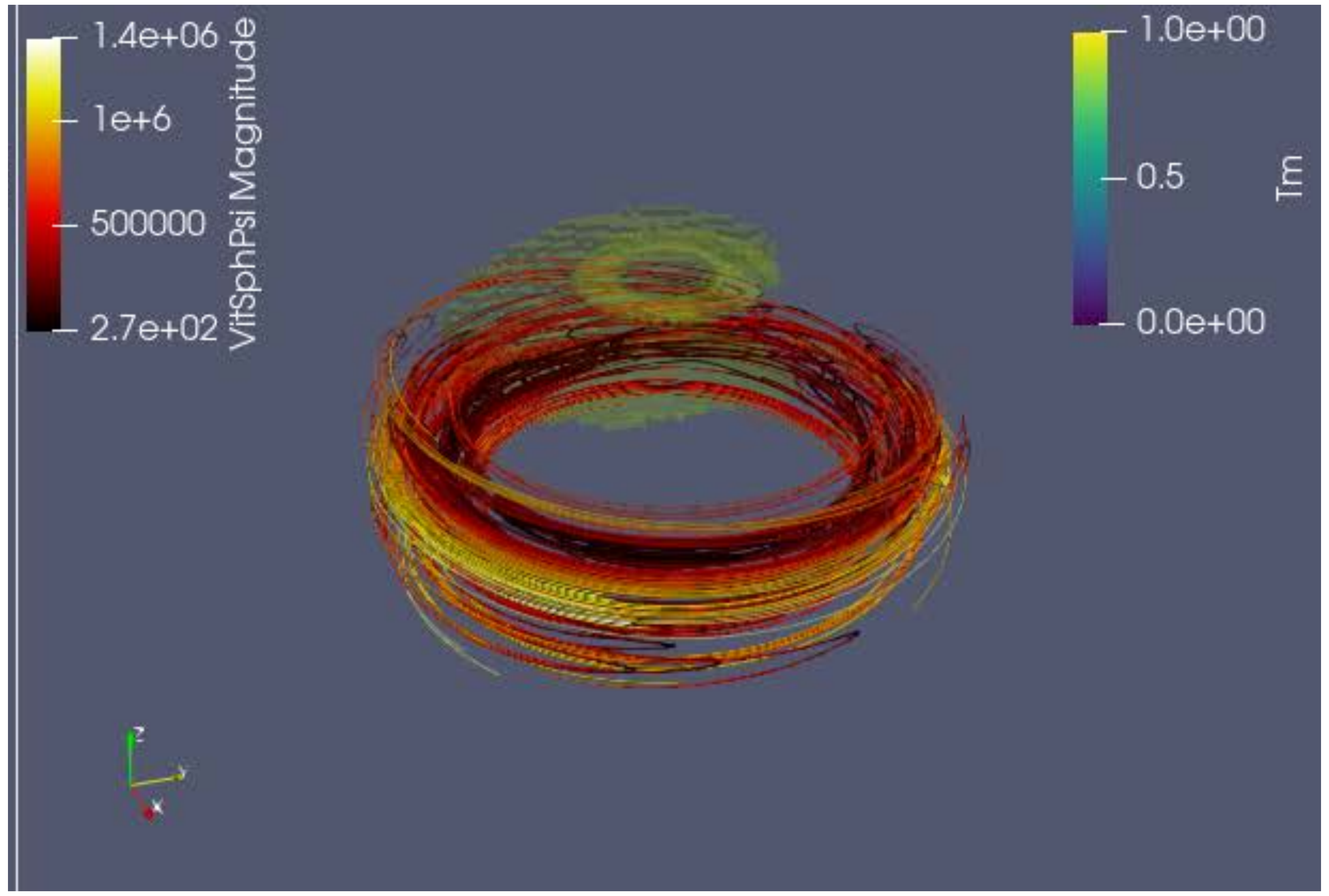
What we expect 🙂 (but boring)



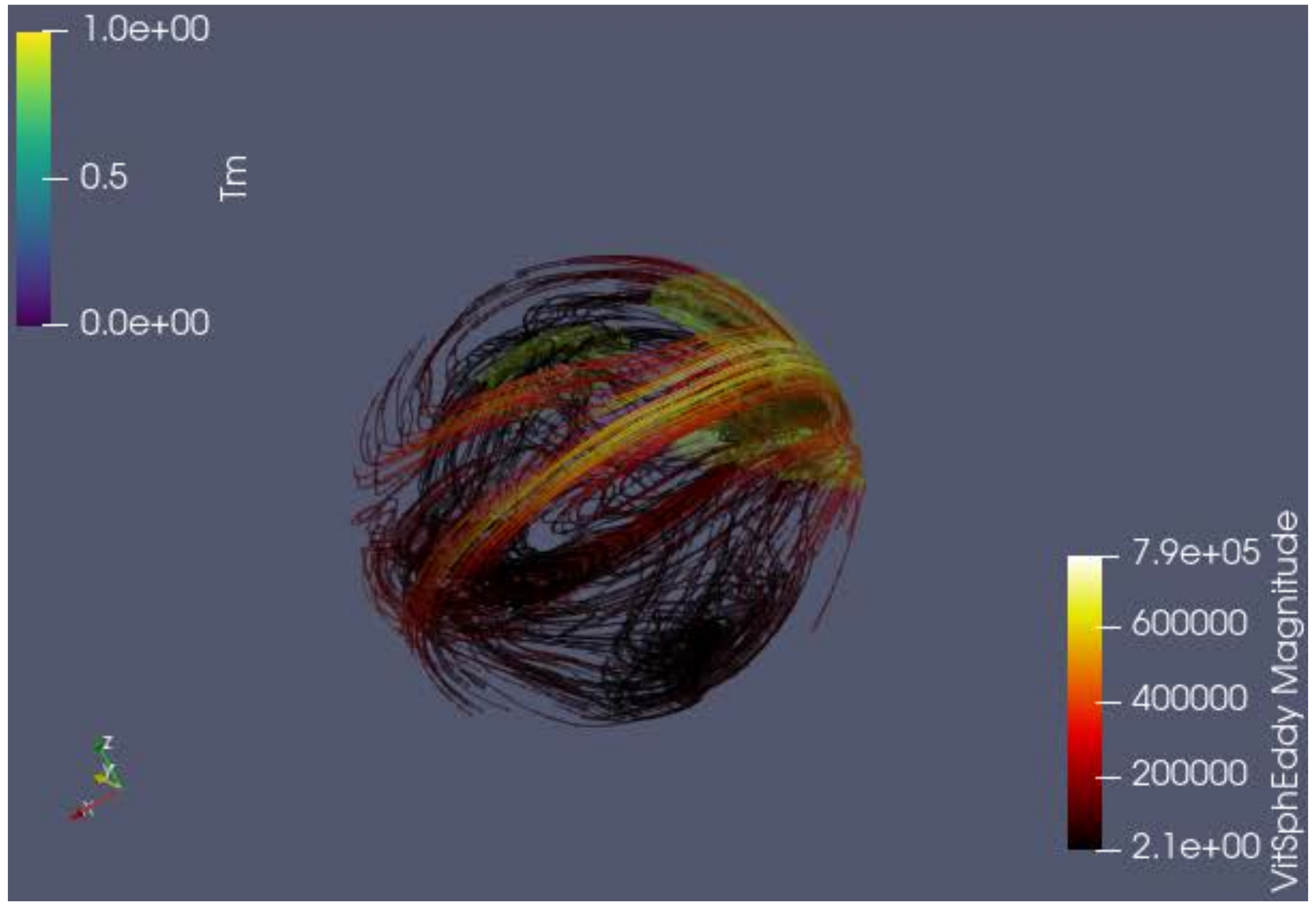
What we did not expect 🤔 (but interesting!)



A nighside hot spot! (Important for phase curve observations)



Zonal jet with dayside hotspot



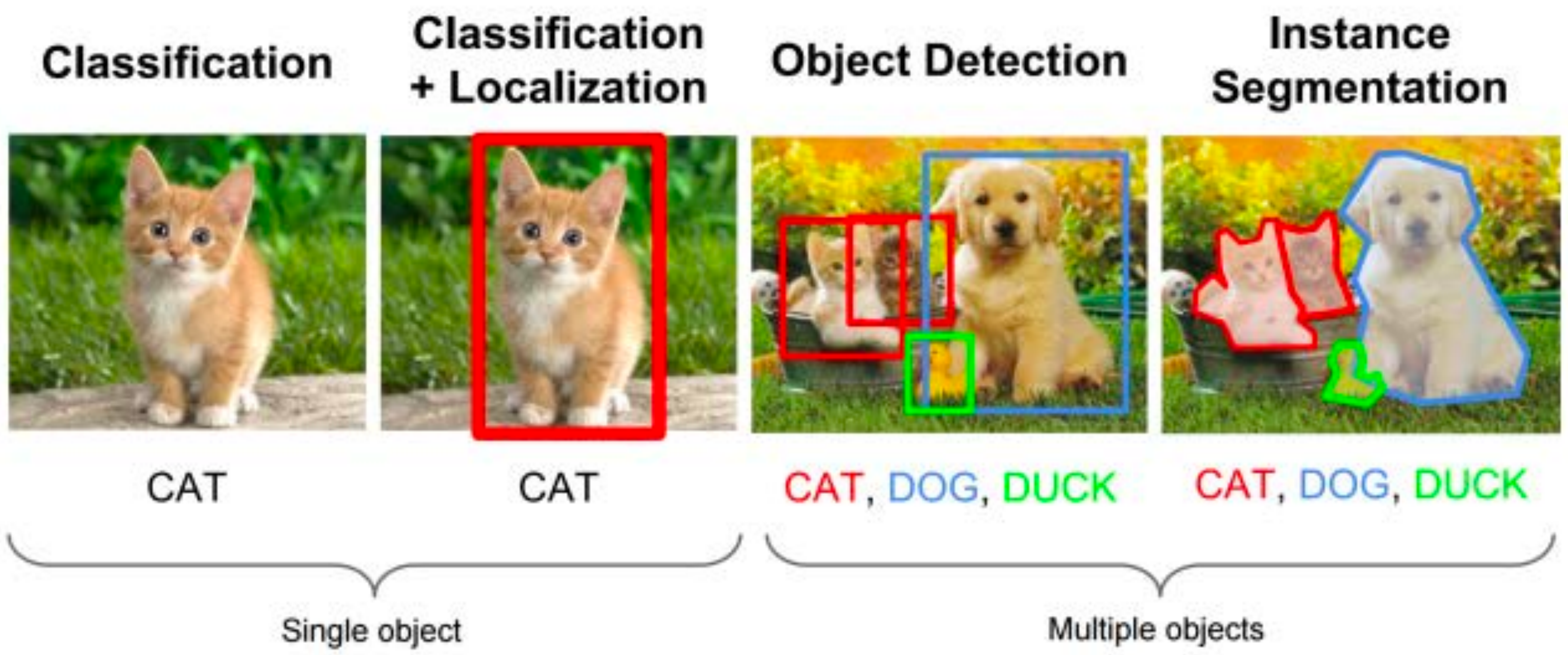
Polar recirculation with nightside hotspot

Conclusion:

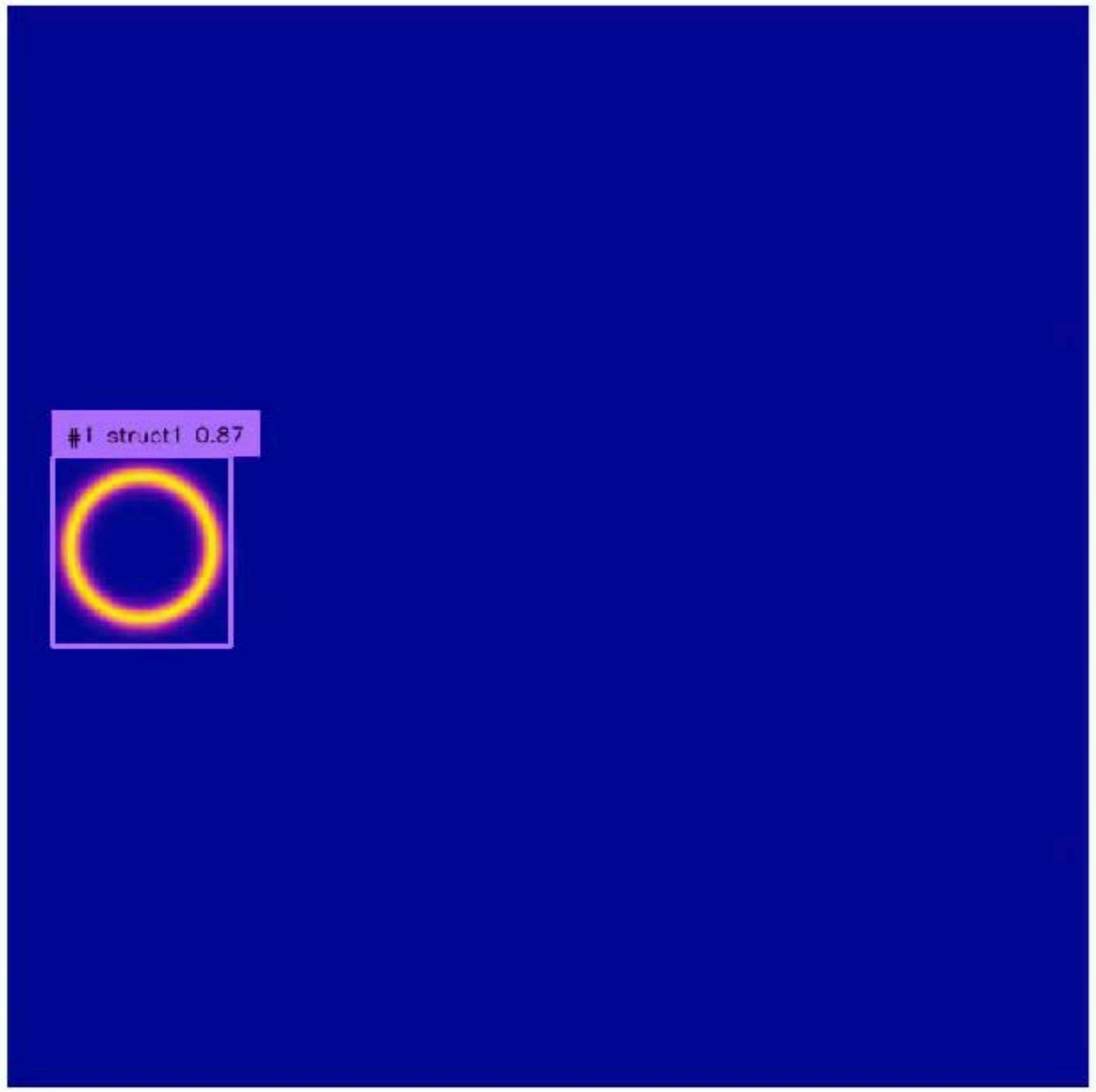
AI is often used to find something we know in noisy data (event/pattern detection). AI-driven classification can be used to remove what we know a-priori and explore what we did not expect!

Sainsbury-Martinez F., Tremblin P., Mancip M., Donfack S., Honore E., Bourenane M.
ApJ 958 68 2023

Future work:



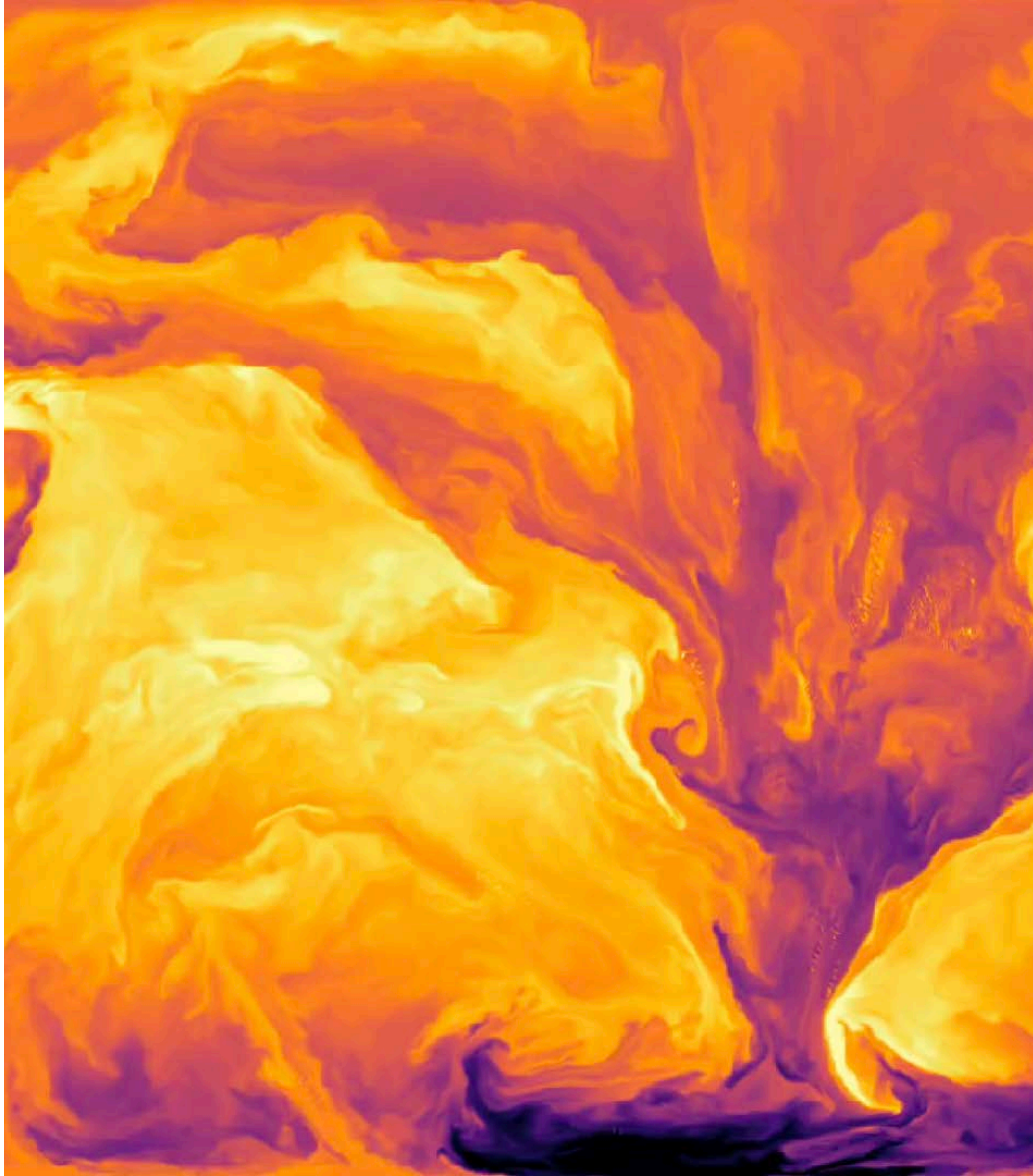
PTC ASSIST H. Taher, M. Lobet, M. Mancip



Dynostar

Pre-exascale MHD dynamo

- PhD Rémi Bourgeois @MdlS -> DES/SGLS
- MPI+Kokkos 3D finite volume simulations of MHD dynamo up to 4096^3 (126 kh GPU on ~1000 GPUs Adastra@CINES)
- PDI+DEISA in-situ data analytics (exa-DoST WP1 & WP2)
- Compute power spectra with in-situ data analytics to check the convergence before upscaling of the simulation



A side note on performances

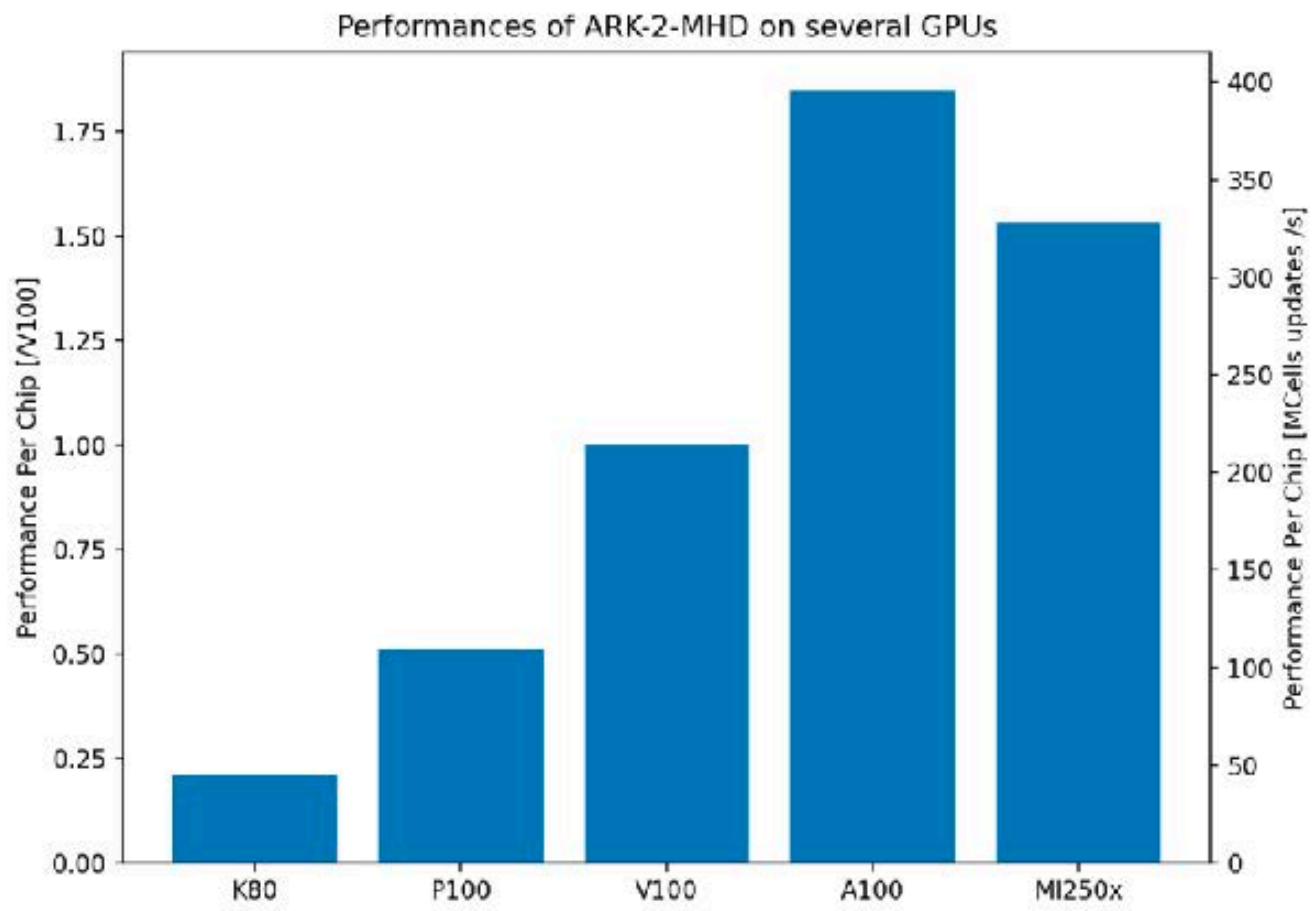


Figure 4.2: Performance comparison with several GPUs

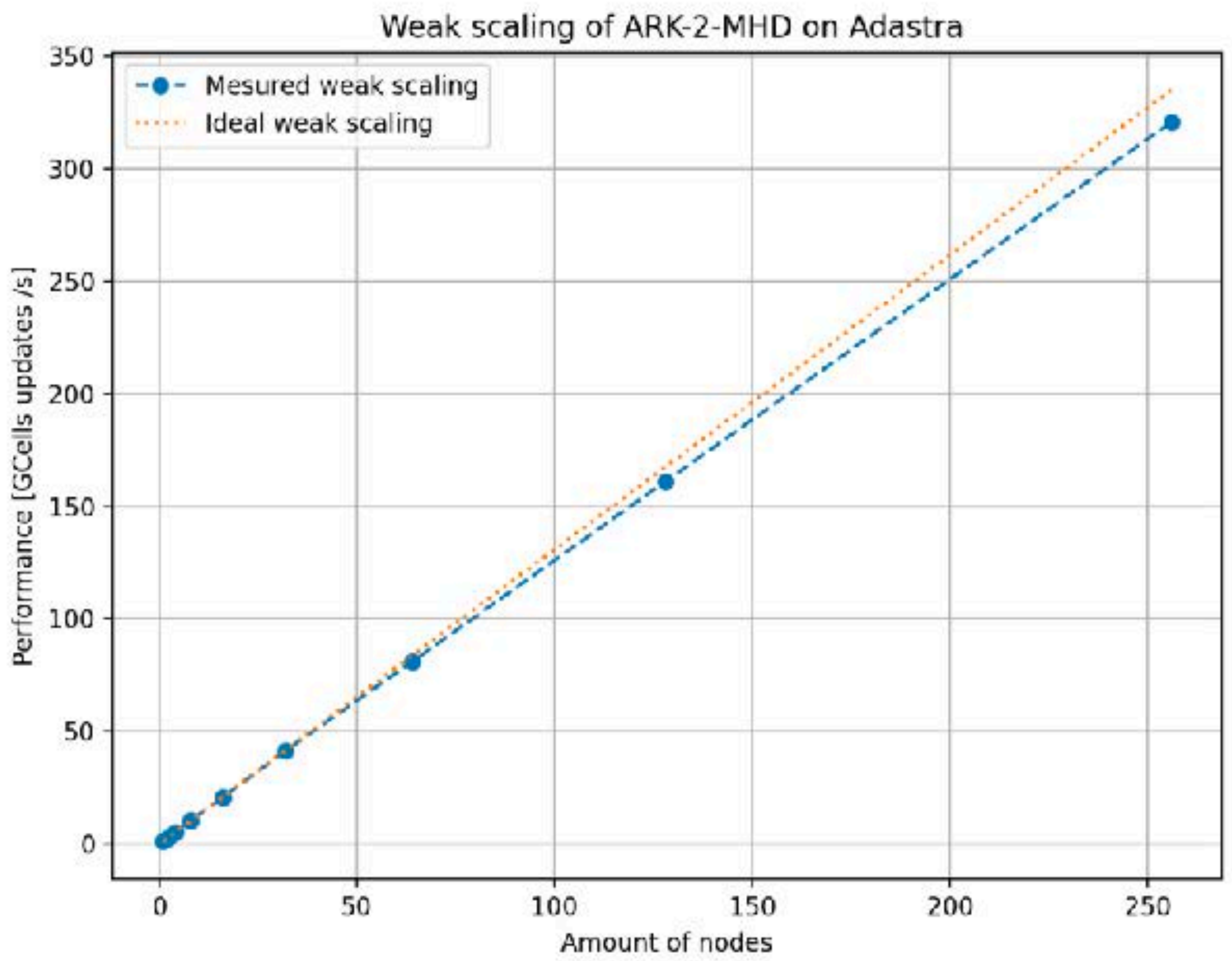


Figure 4.3: Weak scaling on *Adastra*

- For comparisons (not exactly the same models and solvers but relatively similar), JAX-Fluid (https://arxiv.org/abs/2402.05193) seems to be at 2 Mcells/s on A100 i.e. a factor x200 in favor of MPI+Kokkos

- It is very easy to get good weak scaling in HPC when the base performance is low. But this is maybe a lot of Watts that may not be efficiently used.

PDI & DEISA

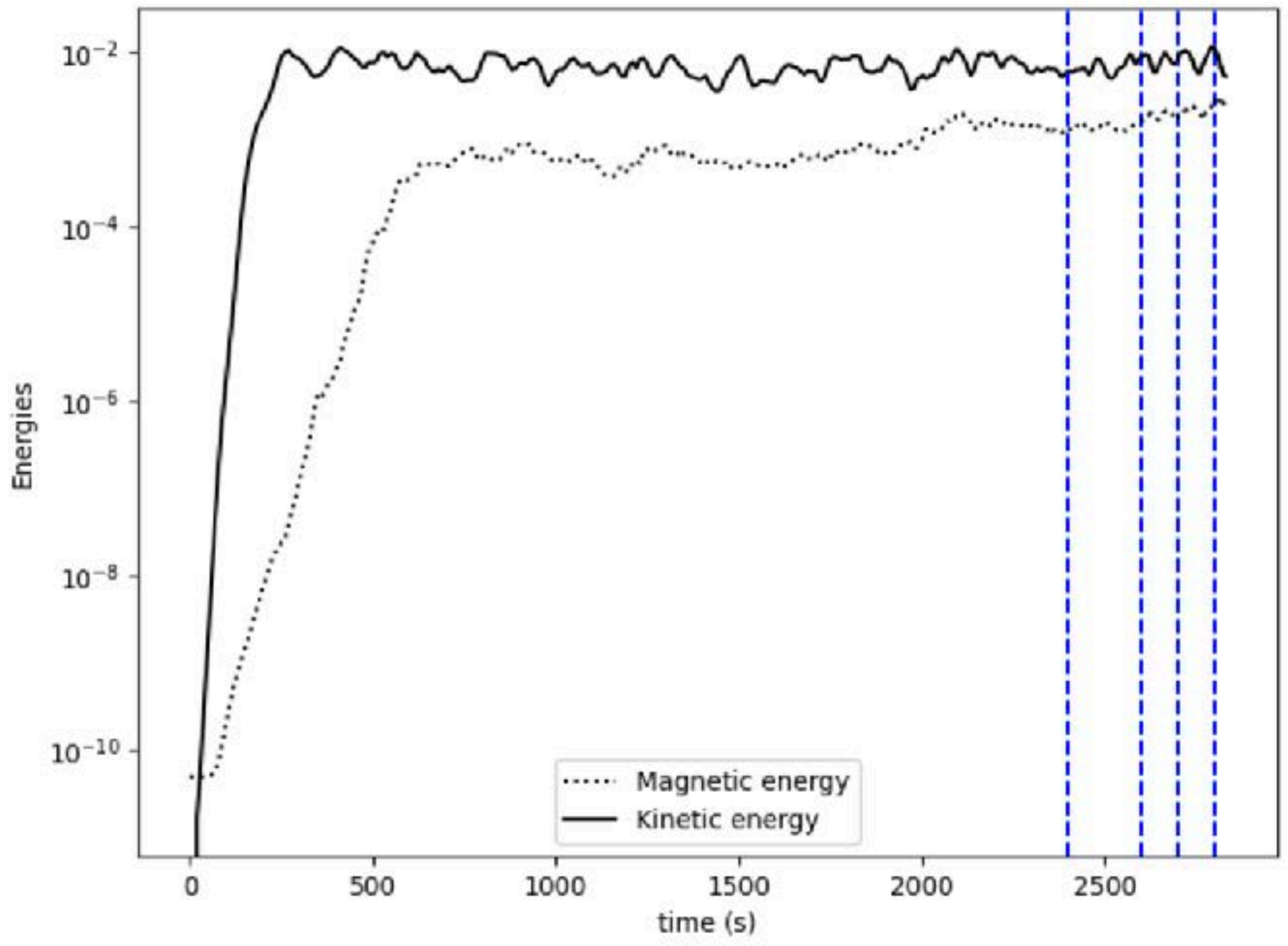
PDI Data Interface

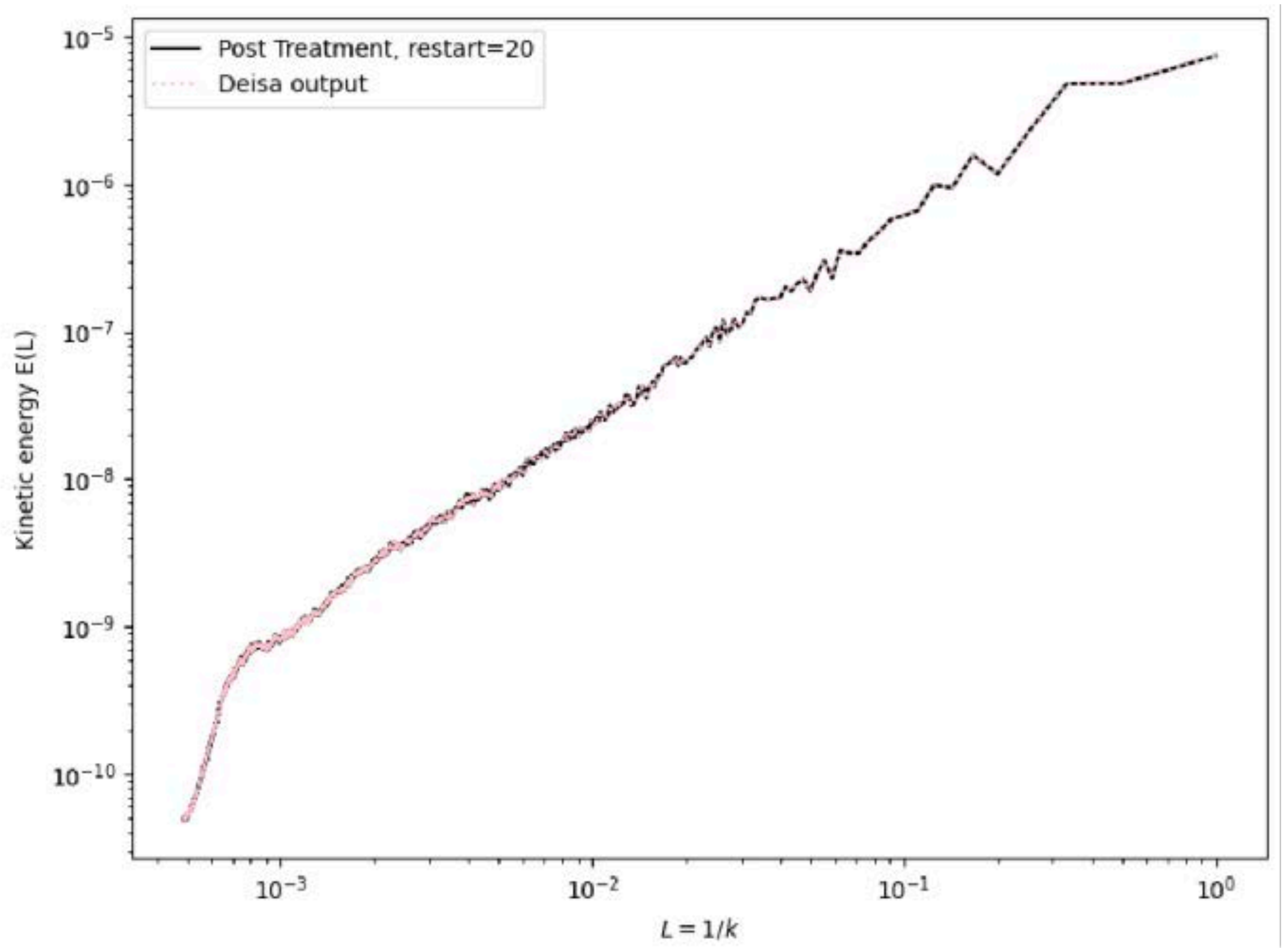
- Initially developed @MdlS, soon part of the High Performance Software Foundation and part of exa-DoST WP1 developments
- PDI supports loose coupling of simulation codes with data handling libraries. The simulation code is annotated in a library-agnostic way, libraries are used from the specification tree.
- This approach works well for a number of concerns including: parameters reading, data initialization, post-processing, result storage to disk, visualization, fault tolerance, logging, inclusion as part of code-coupling, inclusion as part of an ensemble run, etc.

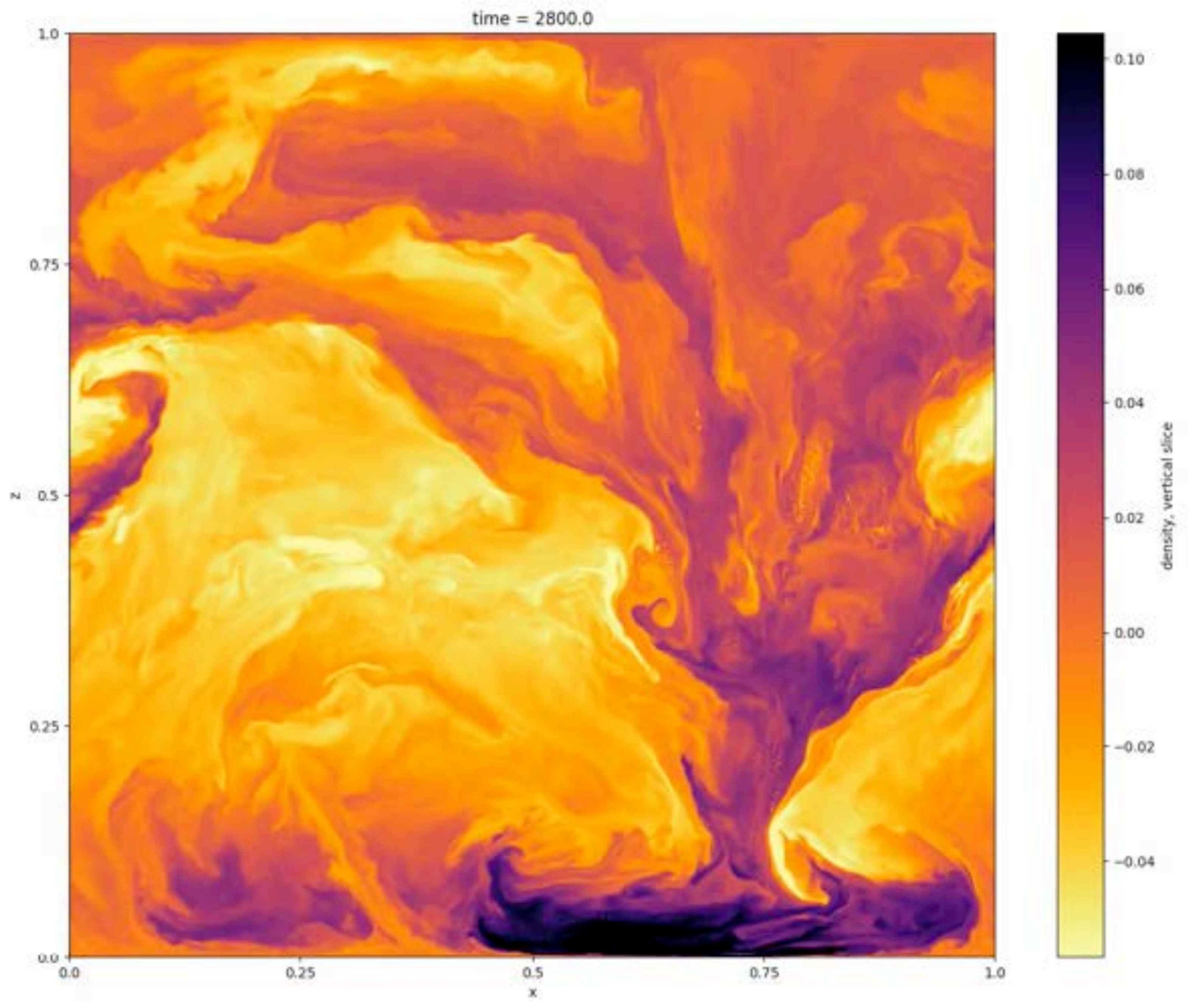
Dask-Enabled In-Situ Analytics (DEISA)

- Initially developped during the PhD of Amal Gueroudji @MdlS & INRIA-Grenoble and part of exa-DosT WP2 developments
- DEISA is a library that ensures coupling MPI simulation codes with Dask analytics
- A simulation can be instrumented with PDI to make its internal data available for DEISA thus Dask. At the beginning each simulation process reads the yaml configuration file and loads the DEISA and the MPI plugins of PDI.

- Through PDI+DEISA:
  - Monitoring of total magnetic and kinetic energies (high time frequencies)
  - Monitoring of the power spectra in the mid plane (MPI reconstruction of the plane)
  - High frequency outputs of xy xz yz planes for 2D movies



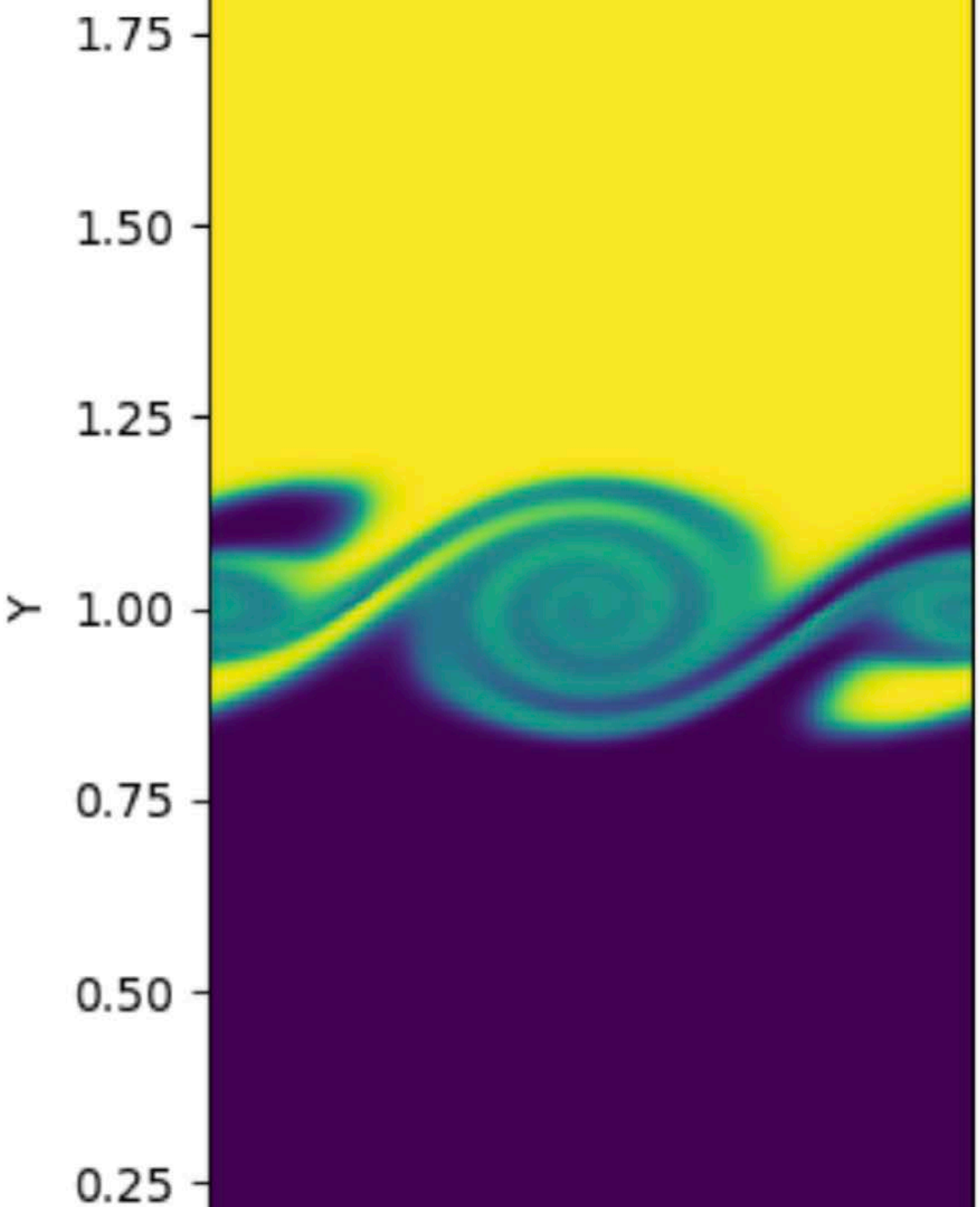




AI small-scale closure

Very early work

- Internship of Jona Nagerl @MdlS, PhD of T. Antoun @MdlS+IRFU (12/24)
- Finite volume closures for HD/MHD small scale turbulence
- Based on statistical mechanics through microscopic closure terms (Fick, stress tensor, conduction)



Underlying Equations

2D conserved Flow - Euler Equation

$$\frac{\partial \mathbf{y}}{\partial t} + \nabla \cdot F(\mathbf{y}) = 0$$

Where $\mathbf{y} = (\rho, \mathbf{j}, E, \rho x)^T$ comprising the fluid density ρ, momentum density $\mathbf{j}$. total energy E , passive scalar x and

$$F(\mathbf{j}) \equiv \begin{pmatrix} \mathbf{j} \\ \frac{1}{\rho}(\mathbf{j} \otimes \mathbf{j}) + p\mathbb{I} \\ (E^t + p)\frac{\mathbf{j}}{\rho} \\ x \cdot \mathbf{j} \end{pmatrix}$$

Where $\otimes$ denotes tensor product and p is pressure

Closure Terms (1/3)

Defining Closure

Taking the average means $\overline{\mu}(y) = \int_0^1 \mu(x, y) dx'$ and assuming periodicity in x:

- **For ρ :**
  The averages $\bar{\mathbf{j}}$ are directly accessible $\Rightarrow$ so there is no closure

- **For $\mathbf{j}$:**
  $\partial_t \bar{\mathbf{j}} + \nabla \left(\frac{1}{\rho} \bar{\mathbf{j}} \otimes \bar{\mathbf{j}} \right) + \nabla \underline{\underline{\sigma}} = 0$ with

  $\underline{\underline{\sigma}} = \begin{bmatrix} \overline{\rho u^2} + \overline{p} - \overline{\rho u} \cdot \overline{u} & \overline{\rho u v} - \overline{\rho u} \cdot \overline{v} \\ \overline{\rho u v} - \overline{\rho v} \cdot \overline{u} & \overline{\rho v^2} + \overline{p} - \overline{\rho v} \cdot \overline{v} \end{bmatrix}$ where $\mathbf{j} = (\rho u, \rho v)$

Closure Terms (2/3)

Closure respecting thermodynamics

To find a closure that is thermodynamical interpretable we define the closure terms:

- **For** ρE **:**
  $\partial_t \overline{\rho E} + \nabla \left((\overline{\rho E} + \underline{\underline{\sigma}}) \cdot \bar{\mathbf{j}} / \rho \right) + \nabla \mathbf{K} = 0$ with
  $\mathbf{K} = \overline{(\rho E + p) \cdot \mathbf{j} / \rho} - (\overline{\rho E} + \underline{\underline{\sigma}}) \cdot \bar{\mathbf{j}} / \rho$
- **For** ρx **:**
  $\partial_t \overline{\rho x} + \nabla \left(\bar{x} \cdot \bar{\mathbf{j}} \right) + \nabla \mathbf{J} = 0$ with $\mathbf{J} = \overline{x \mathbf{j}} - \bar{x} \cdot \bar{\mathbf{j}}$

Closure Terms (3/3)

Skipping some calculations

We can calculate with the evolution of the internal energy $\rho e_{\text{eff}} = \rho E - \dfrac{\bar{\mathbf{j}}^2}{2\rho}$

$$\rho_0 \frac{de_{\text{eff}}}{dt} = - \underline{\underline{\sigma}} : (\nabla \otimes \bar{\mathbf{u}}) - \nabla(\mathbf{K})$$

by assuming $\sigma_{11} \approx \sigma_{22} \approx p$ and periodicity we can further simplify

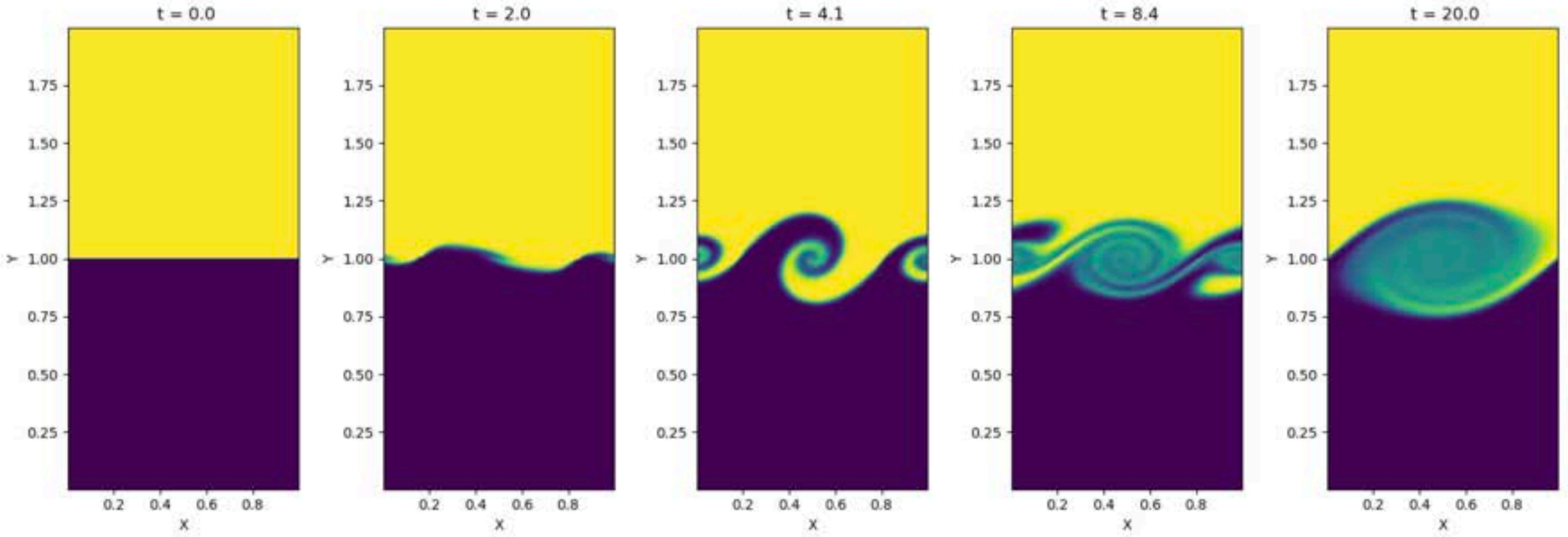
$$\underbrace{\rho \frac{de}{dt} + p \frac{d\tau}{dt}}_{=0} = - \frac{\partial_y K_y}{\rho} - \frac{\sigma_{2,1}}{\rho} \partial_y \bar{u}$$

$$\Rightarrow \quad \partial_y K_y = - \sigma_{2,1} \partial_y \bar{u}$$

We can associate the closure term of the energy with the closure of the velocity

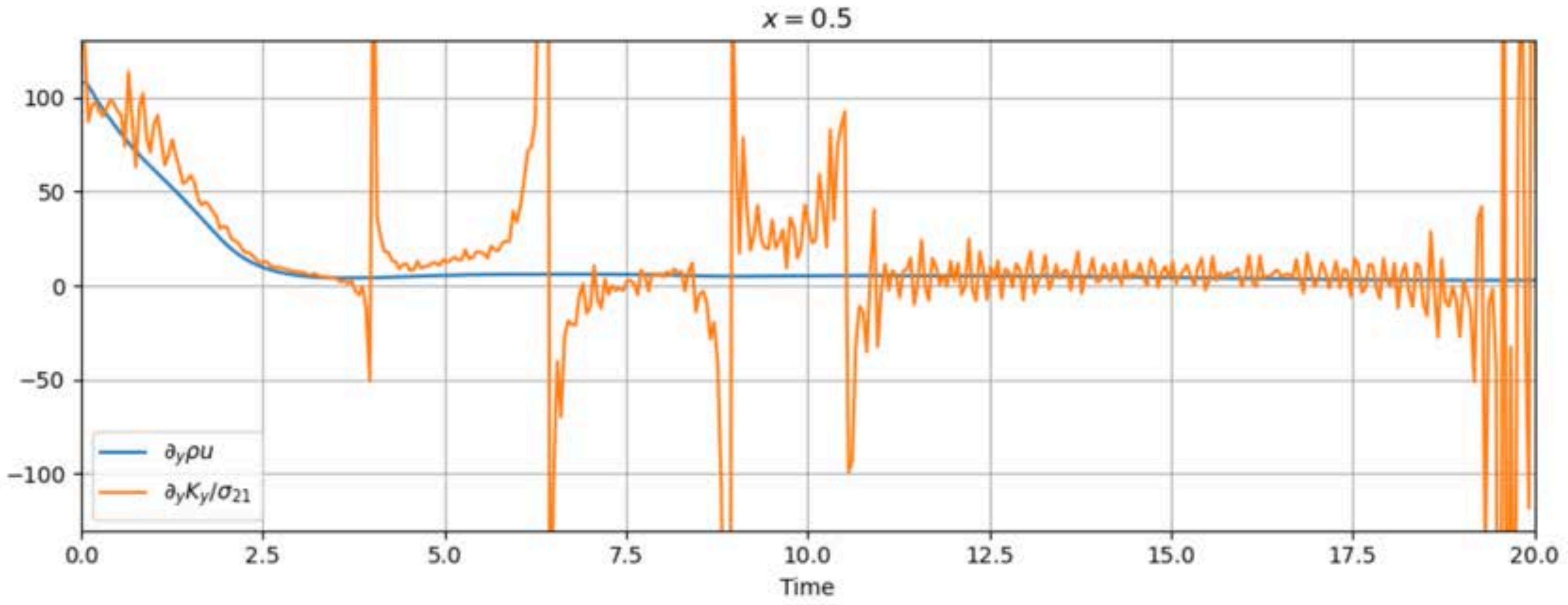
Evolution KH-Instability in 2D

Plots of the passive scalar ρx (to observe the system in 2D)



- Instability grows quickly
- Resolution allows just one mode
- In long term numerical diffusion leads to mixing

Link Kinetic Energy & Internal Energy

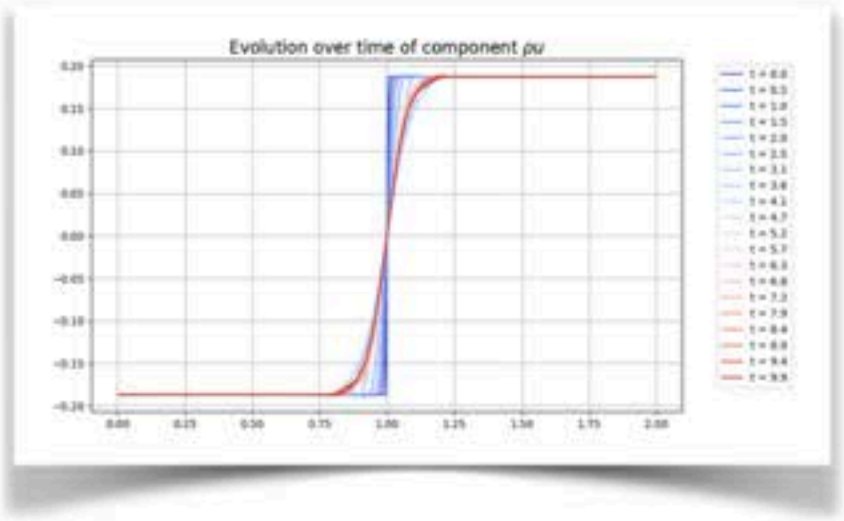


Observations:

- Connection is very visible
- Deviations relate to $\sigma_{21} \to 0$

Evolution of $\overline{\rho u}$

Modelling Evolution with Diffusion Eq.



Observation:

- Evolution looks in the beginning like diffusion (parabolic)
- Until non-linear dynamics lead to stabilisation (hyperbolic)

First model:

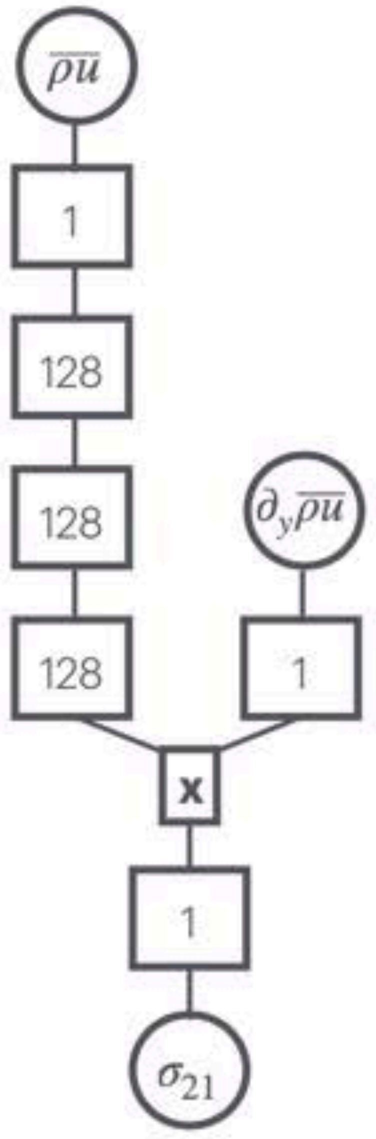
Model beginning with adaptive diffusion coefficient

$$\partial_t \overline{\rho u} + D \partial_y^2 \overline{\rho u} = 0$$

$$\partial_y \sigma_{21} = - D \partial_y^2 \overline{\rho u} \quad \Rightarrow \quad \sigma_{21} = - D \partial_y \overline{\rho u}$$

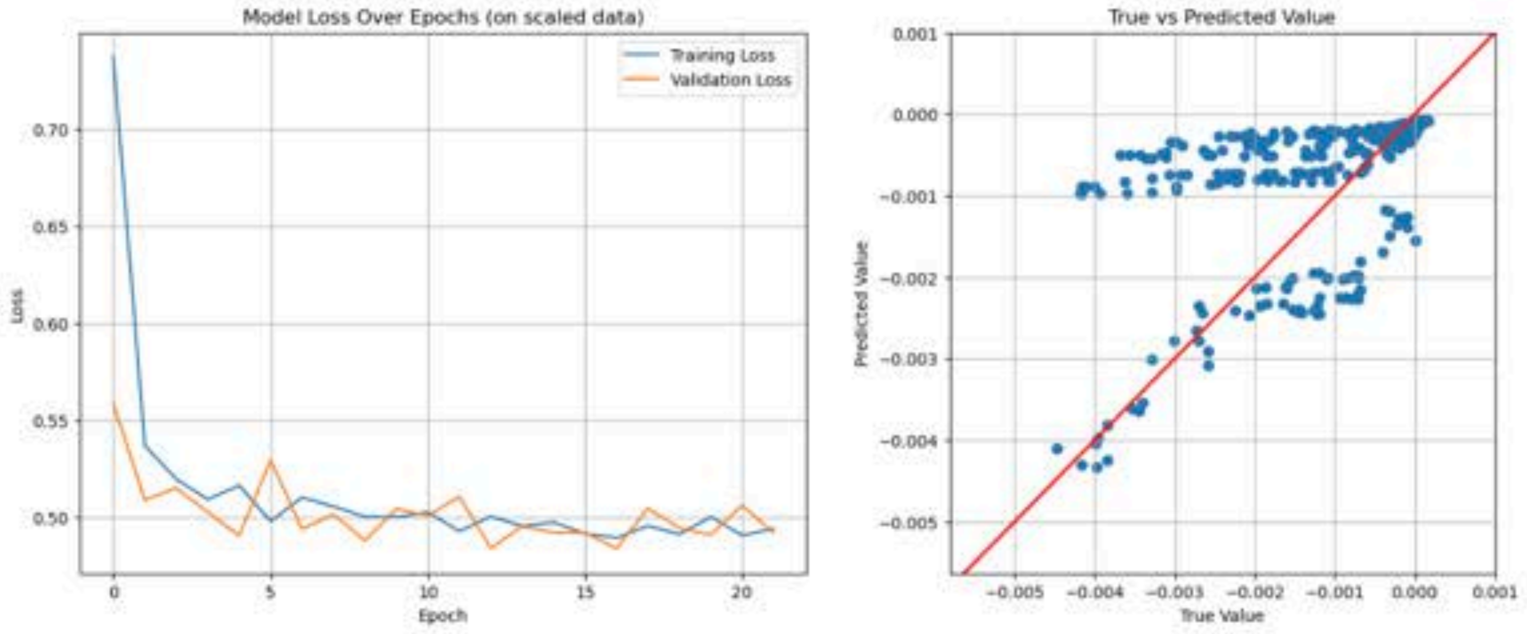
Neural Network Closure:

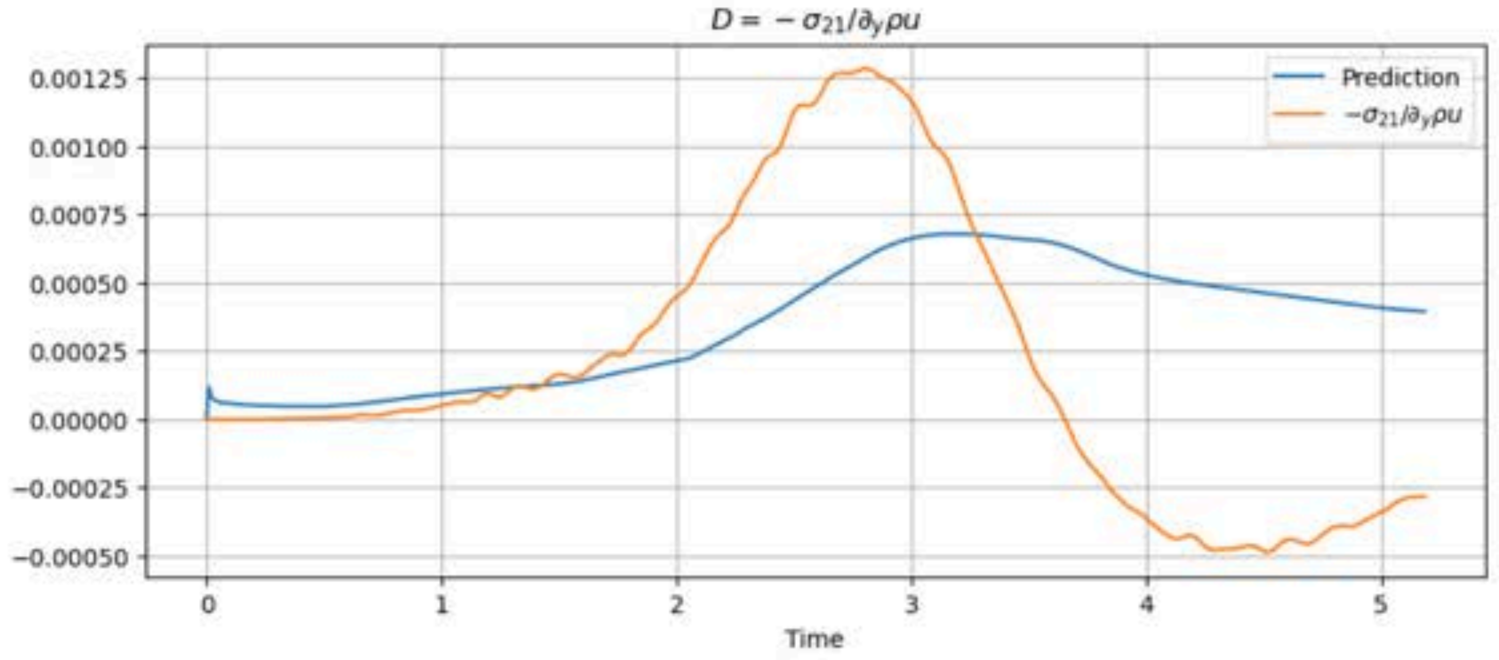
$NN(\overline{\rho u}) \cdot \partial_y \overline{\rho u} \rightarrow \sigma_{21}$ to predict the diffusion coefficient D



Exemplary Structure of a Neural Network

First Closure Model

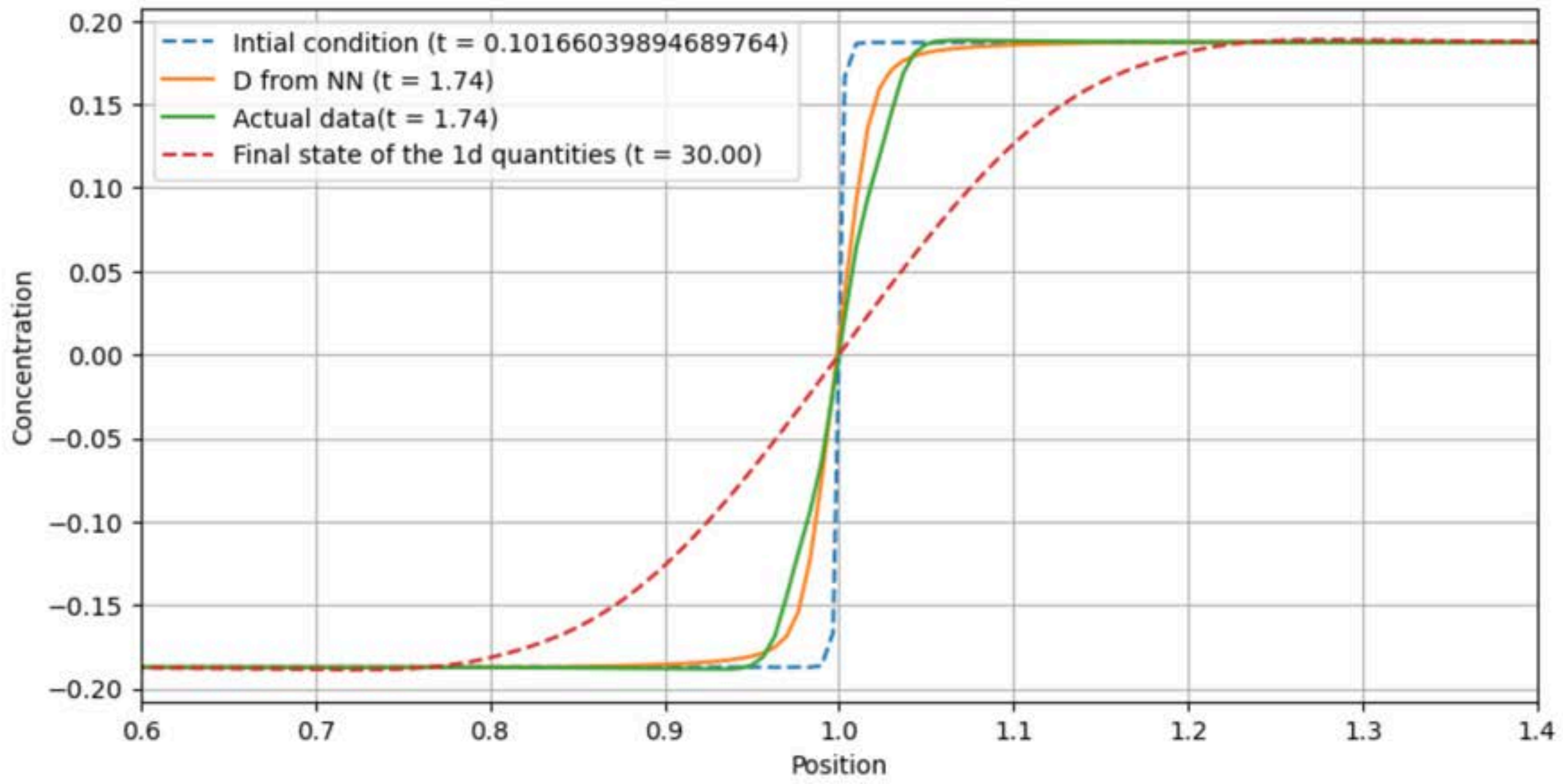




Observation:

Decent model, with room for improvement

Simulation with NN closure (1)



Observation:

NN closure captures qualitatively the start of the evolution

Improved Closure

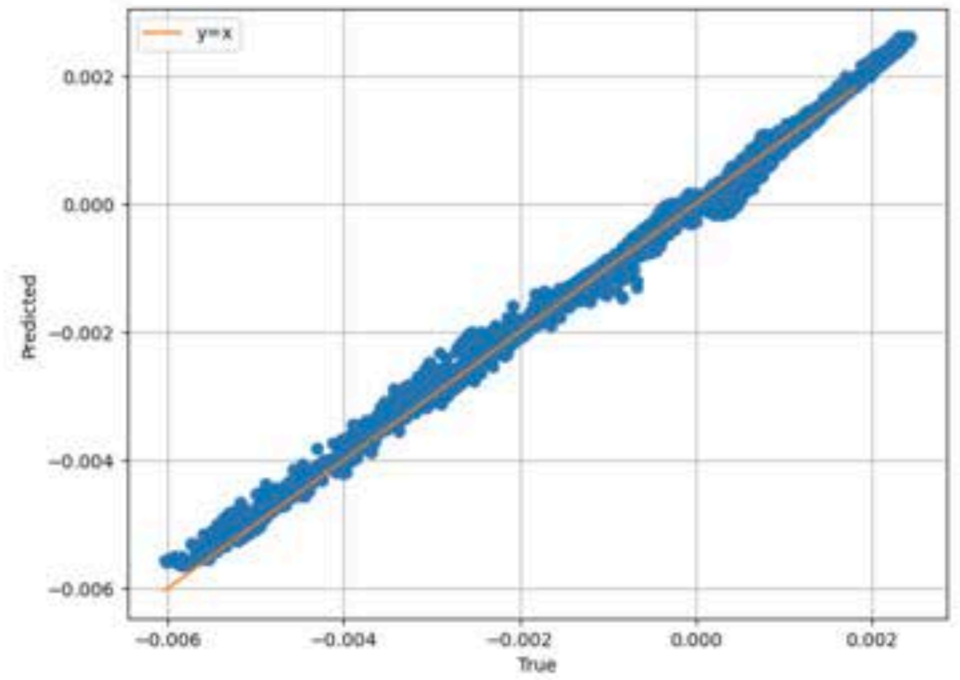
Incorporating global behaviour

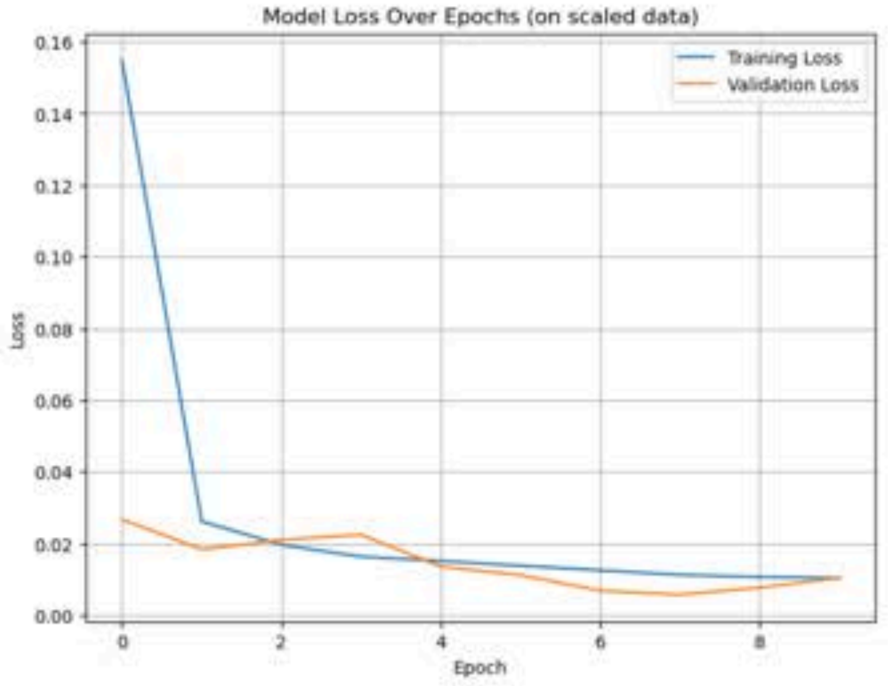
Idea:

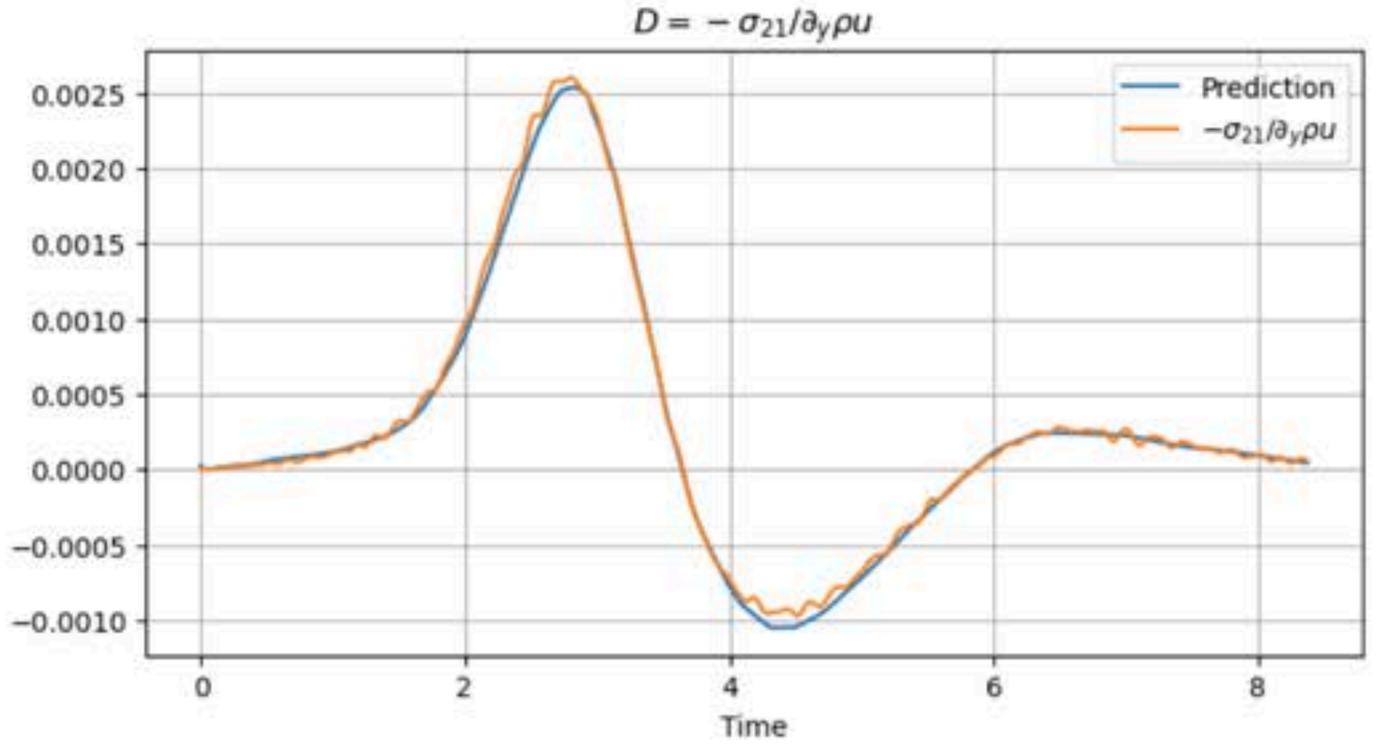
Channel size is degree of freedom which is lost by averaging, but is remnant in temporal order of scale

$\Rightarrow$ incorporate $\hat{t}$ time since onset of the instability

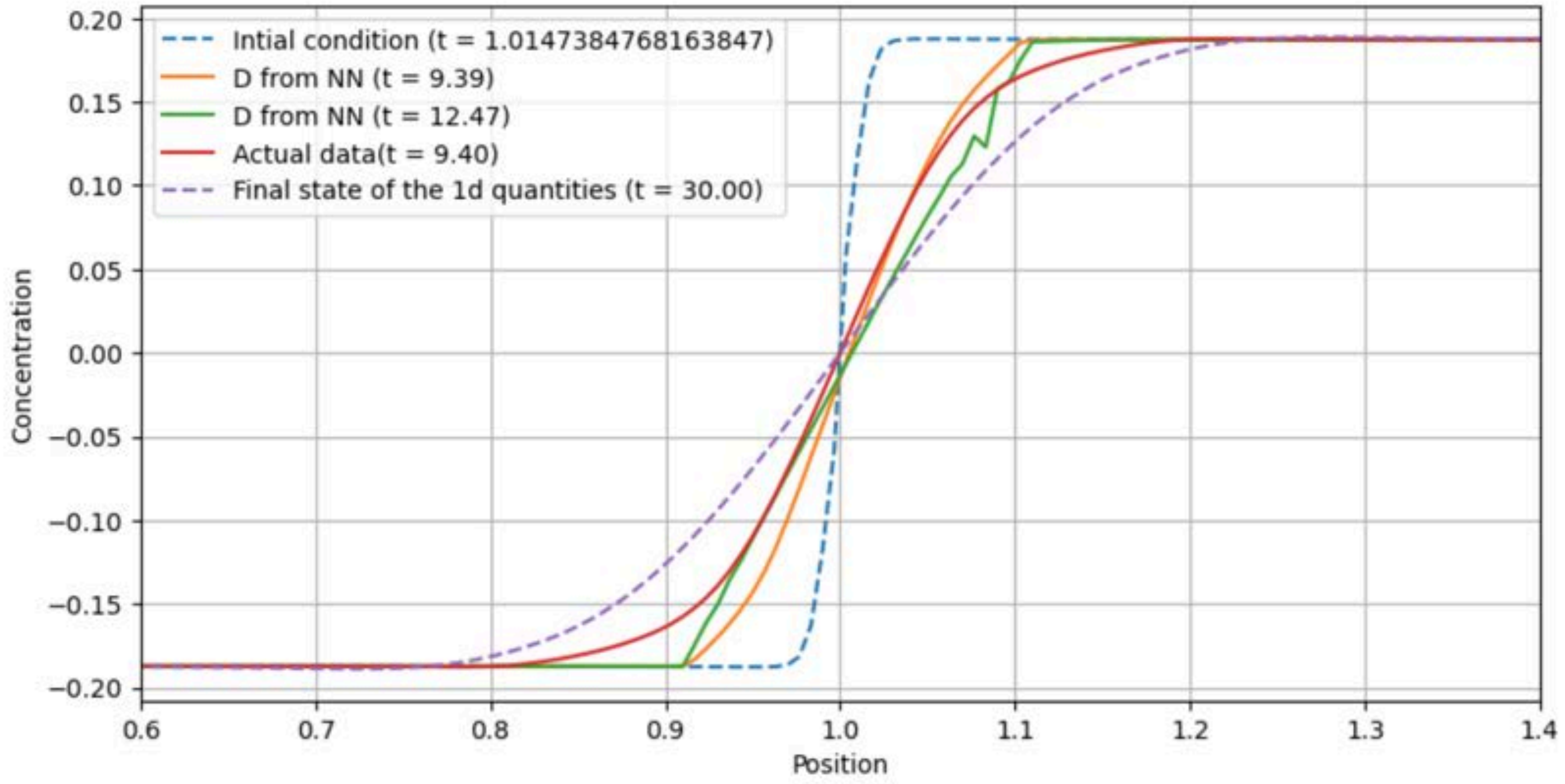
$NN(\overline{\rho u}, \hat{t}) \cdot \partial_y \overline{\rho u} \rightarrow \sigma_{21}$







Simulation with NN closure (2)



Observation:

NN closure captures long term behaviour

Transferability may be weak point