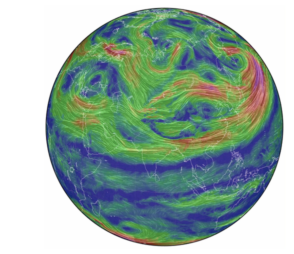
Aalto University

Towards modeling weather



ClimODE: Climate and Weather Forecasting using Physics-Informed Neural ODEs

Yogesh Verma, Markus Heinonen, Vikas Garg

Aalto University YaiYai Ltd

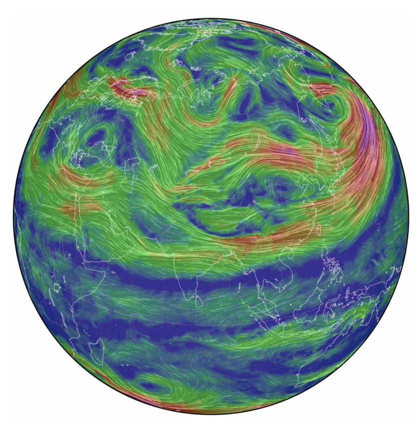
ICLR

Oral

Objective

Model and Forecast:

$$\mathbf{u}(\mathbf{x},t) = \begin{pmatrix} u_1(\mathbf{x},t) \\ \vdots \\ u_K(\mathbf{x},t) \end{pmatrix} \begin{matrix} \text{Temperature} \\ \vdots \\ \text{Precipitation} \end{matrix}$$



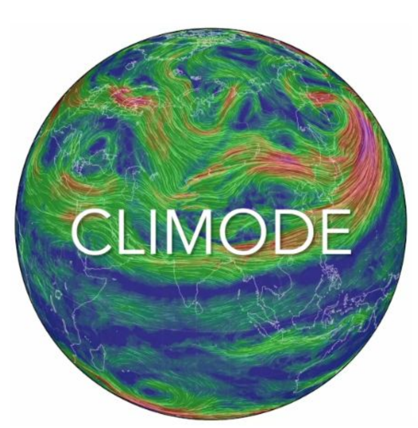
Issues with black box modeling

- Black box methods based on Transformers, UNets, etc overlook the **physical dynamics (P)** and **continuous time (CT) nature** and are not **compact (C).**
- **Free Form Neural PDEs** does not assume any physical dynamics but gives solution for continuous time

	P	CT	C
UNet	✗	✗	✗
Neural PDE	✗	✓	✗
ClimODE	✓	✓	✓

Contributions

- Develop Neural ODEs/PDEs that:
  - Follow **P**, **CT** and **C**
  - **Provides uncertainty estimates**



Method	Value-preserving	Explicit Periodicity/Seasonality	Uncertainty	Continuous-time	Parameters (M)	
FourCastNet	✗	✗	✗	✗	N/A	Pathak et al. (2022)
GraphCast	✗	✗	✗	✗	37	Lam et al. (2022)
Pangu-Weather	✗	✗	✗	✗	256	Bi et al. (2023)
ClimaX	✗	✗	✗	✗	107	Nguyen et al. (2023)
NowcastNet	✓	✗	✗	✗	N/A	Zhang et al. (2023)
ClimODE	✓	✓	✓	✓	2.8	this work

Advection PDE

Basic idea:

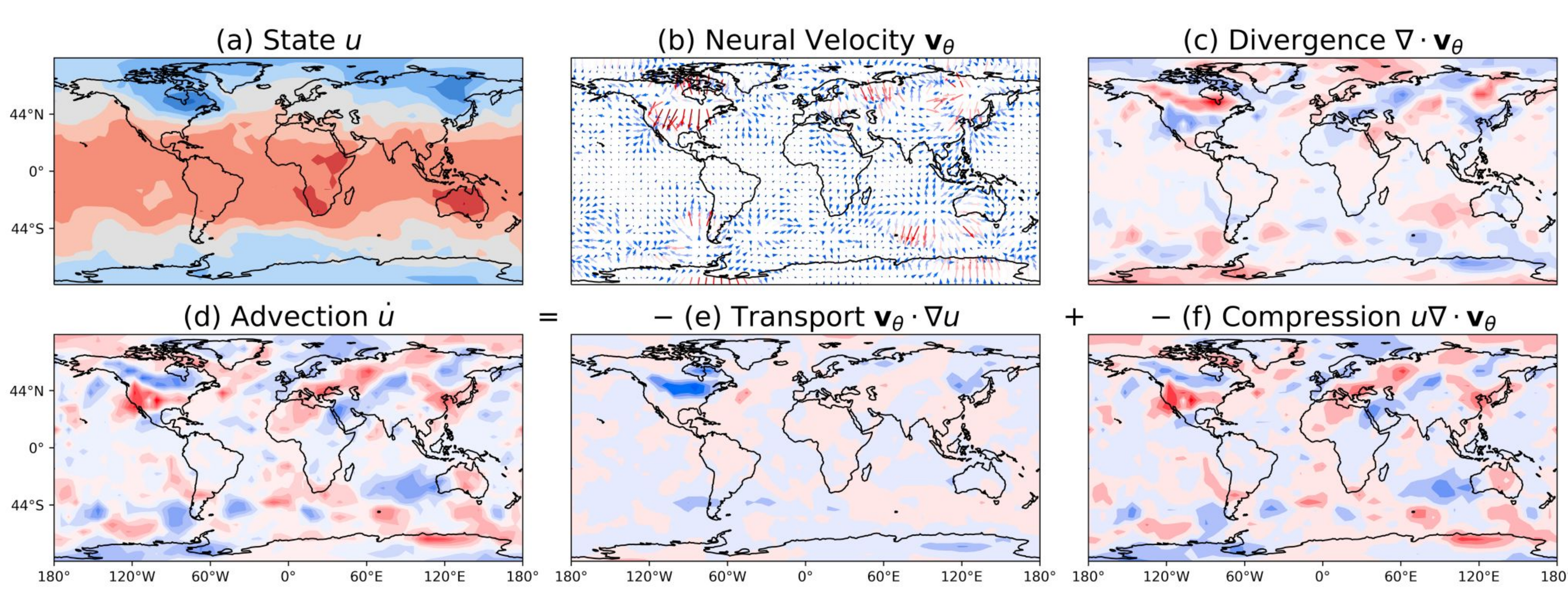
1. Movement of a substance u (scalar field) via motion $\mathbf{v}$ (vector field) as bulk motion.
2. Value conserving dynamics and require currents.

$$\underbrace{\frac{\partial u}{\partial t}}_{\text{Time evolution } \dot{u}} + \overbrace{\mathbf{v}\cdot\nabla u}^{\text{Transport}} + \overbrace{u\nabla\cdot\mathbf{v}}^{\text{Compression}} = 0$$

Neural Transport Model

We model weather/climate as a spatiotemporal process $\mathbf{u}(x,t) = (u_1(x,t),\ldots,u_K(x,t))$ of K quantities as an advection PDE,

$$\dot{u}_k(x,t) = -\mathbf{v}_k(x,t)\cdot\nabla u_k(x,t) - u_k(x,t)\nabla\cdot\mathbf{v}_k(x,t)$$



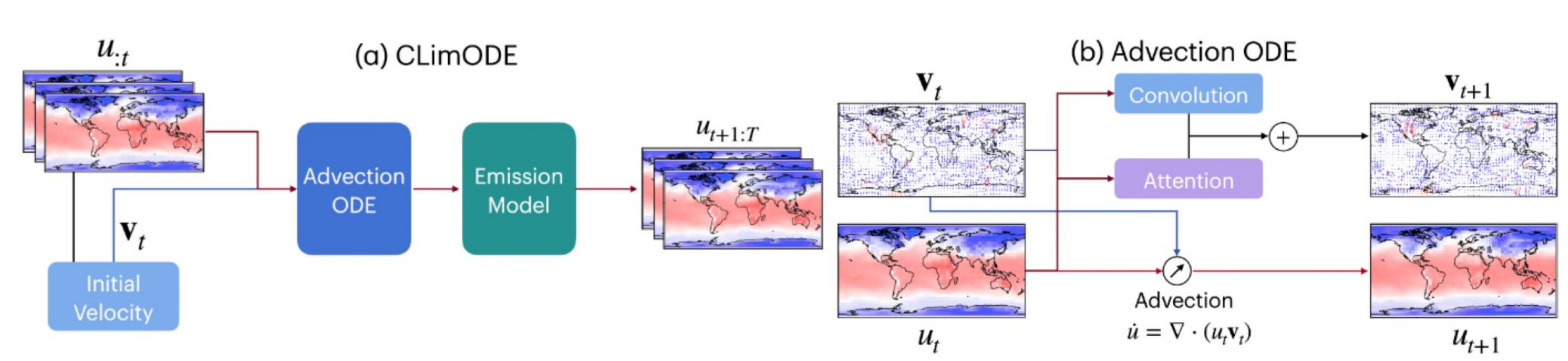
Flow velocity v

We model the flow velocity via a NN, as a function of observed values $\mathbf{u}(t)$, $\nabla\mathbf{u}$, $\mathbf{v}(t)$ and spatio-temporal embeddings ψ

$$\dot{\mathbf{v}}_k(x,t) = f_\theta\Big(\mathbf{u}(t), \nabla\mathbf{u}(t), \mathbf{v}(t), \psi\Big)$$

To capture **local** and **global** effects, we propose a hybrid network

$$f_\theta\Big(\mathbf{u}(t), \nabla\mathbf{u}(t), \mathbf{v}(t), \psi\Big) = f_{\text{conv}}\Big(\mathbf{u}(t), \nabla\mathbf{u}(t), \mathbf{v}(t), \psi\Big) + \gamma f_{\text{att}}\Big(\mathbf{u}(t), \nabla\mathbf{u}(t), \mathbf{v}(t), \psi\Big)$$



Training Objective

Optimise the log likelihood over the observations,

$$\log p(y \mid \theta,\phi) \propto \sum_{i=1}^{N} \log \mathcal{N}\Big(y_i \mid \mathbf{u}_\theta(t_i) + \mu_\phi(t_i), \mathrm{diag}(\sigma_\phi(t_i))\Big)$$

1. Solve $\mathbf{u}(t)$ forward with neural velocity
2. Evaluate likelihood
3. Backpropagate wrt θ,ϕ

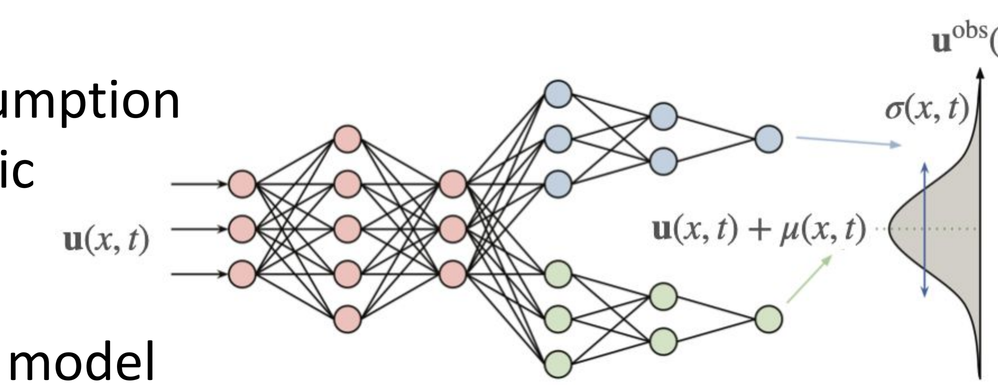
Sources and Uncertainty

Limitation:

1. Closed system assumption
2. ODE is deterministic

Idea:

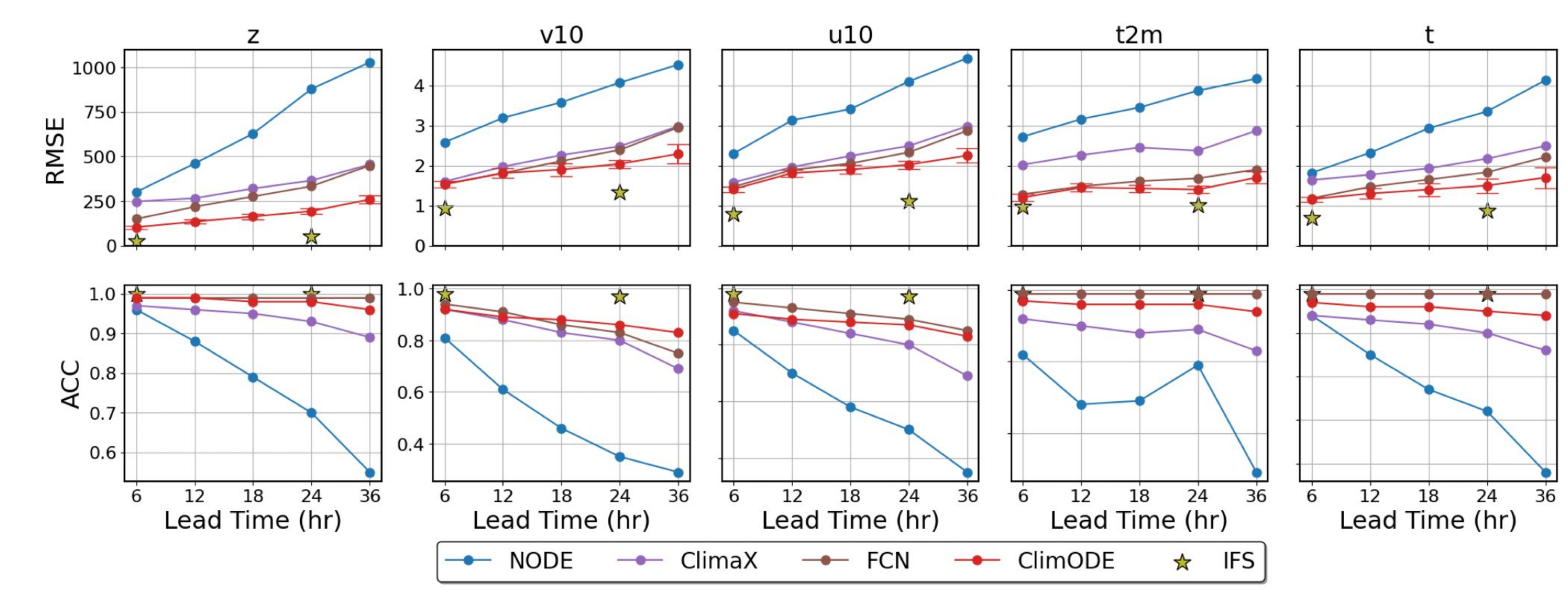
1. Gaussian Emission model
2. ODE becomes mean forecast



$$u_k^{\text{obs}}(x,t) \sim \mathcal{N}\Big(u_k(x,t) + \mu_k(x,t), \sigma_k^2(x,t)\Big), \qquad \mu_k(x,t), \sigma_k(x,t) = g_\phi\big(\mathbf{u}(x,t), \psi\big).$$

Experiments

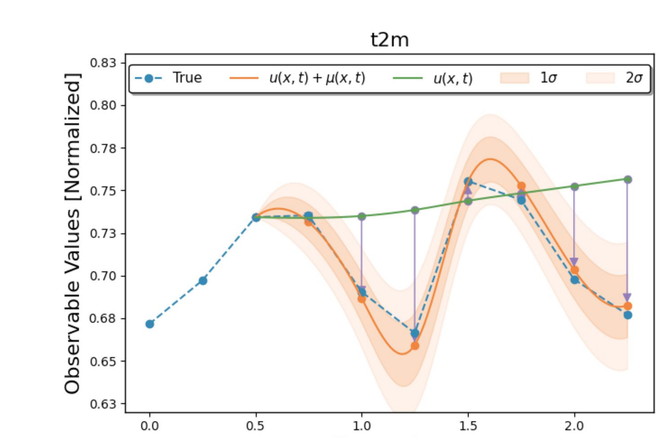
Global Forecasting

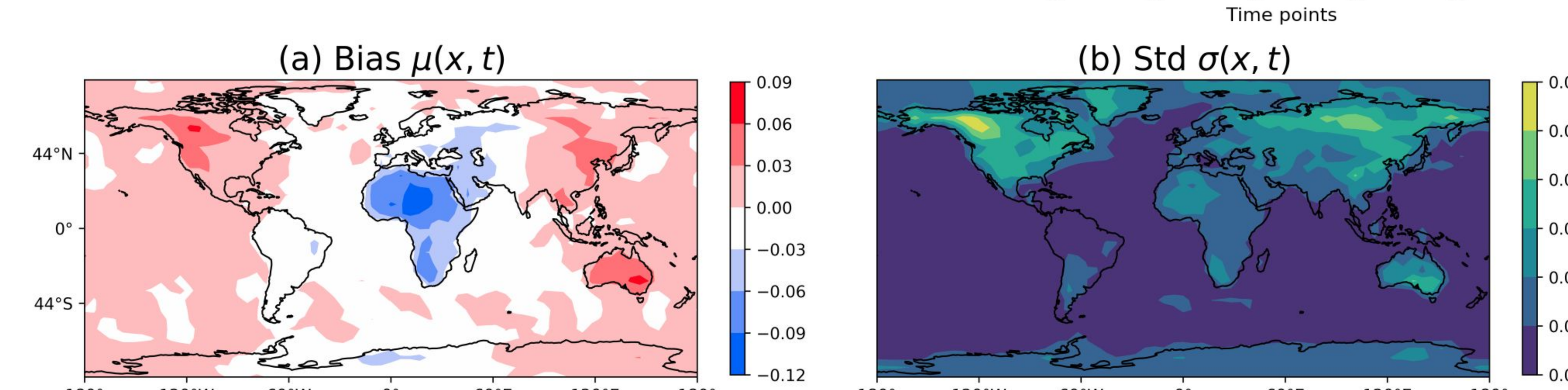


Interpretability

Bias: Explains day-night cycle

Uncertainty: Highest on land and in north according to diurnal cycle





References

1. Climax: A foundational model for weather and climate, ICML 2023
2. FourCastNet: A Global Data-driven High resolution weather model, PASC 2023
3. Learning skillful medium-range global weather forecasting, Science, 2023
4. ECMWF, IFS Documentation CY48R1
5. Neural Ordinary Differential Equations, NeurIPS 2018