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

Yogesh Verma, **Markus Heinonen**, **Vikas Garg**

(yogesh.verma@aalto.fi , markus.o.heinonen@aalto.fi , vgarg@csail.mit.edu)

Aalto University & **YaiYai Ltd**

ICLR Oral

Objective

Model and forecast weather and climate:

$$\mathbf{u}(\mathbf{x},t) = \begin{pmatrix} u_1(\mathbf{x},t) \\ \vdots \\ u_K(\mathbf{x},t) \end{pmatrix} \begin{matrix} \text{Temperature} \\ \vdots \\ \text{Precipitation} \end{matrix} \qquad \mathbf{u}(x,t) = \mathbf{u}(x,0) + \underbrace{\int_0^t \dot{\mathbf{u}}(x,t)\,dt}_{\text{unknown}}$$

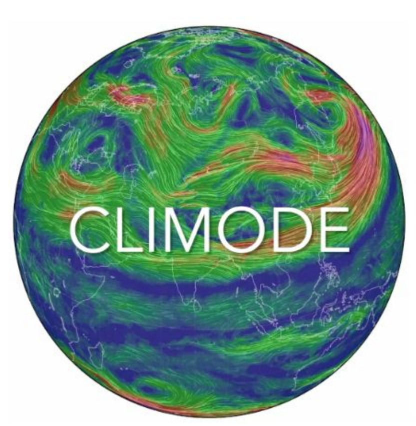
Issues with black box modeling

- 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** do not include any physical dynamics but solve for continuous time

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

Contributions

- Develop Neural ODEs/PDEs that
  - **Respect P, are CT and C**
  - **Provide uncertainty estimates**



Method	Value preserving	Explicit Periodicity/Seasonality	Uncertainty	Continuous-time	Parameters (M)
FourCastNet	✗	✗	✗	✗	N/A
GraphCast	✗	✗	✗	✗	37
Pangu-Weather	✗	✗	✗	✗	256
ClimaX	✗	✗	✗	✗	107
NowcastNet	✓	✗	✗	✗	N/A
ClimODE (ours)	✓	✓	✓	✓	**2.8**

Advection PDE

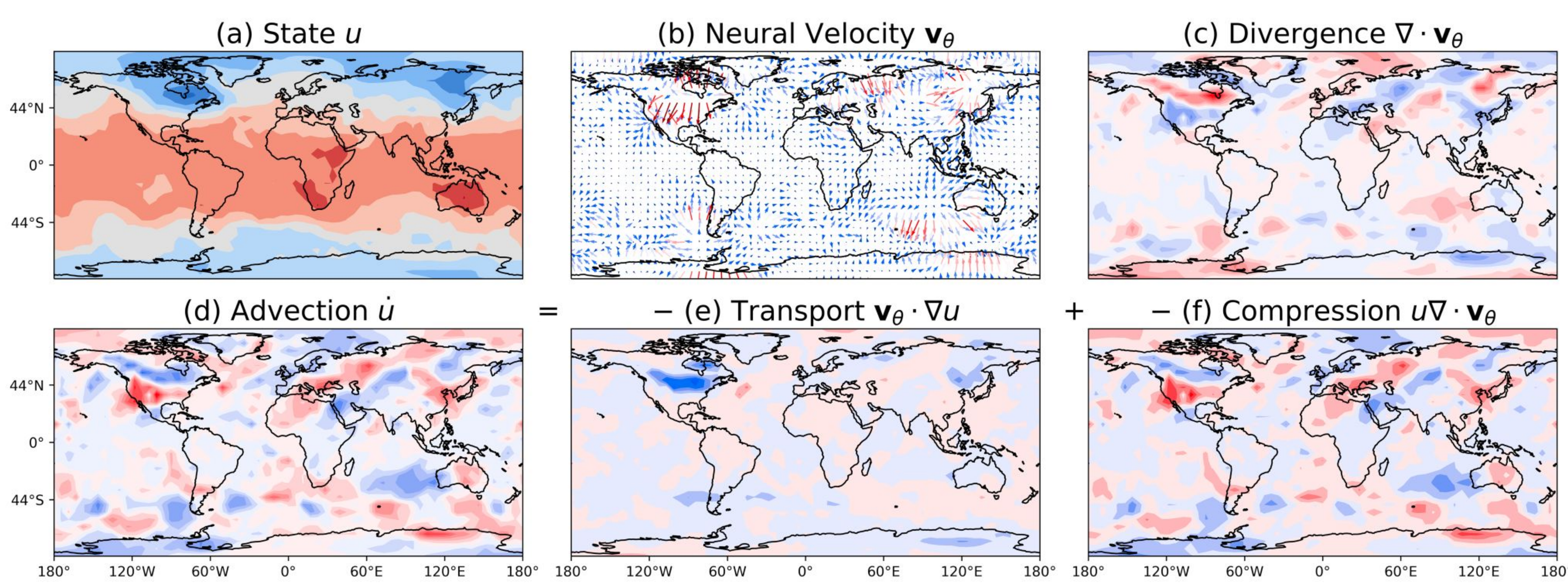
1. Movement of quantities u (scalar field) due to velocity $\mathbf{v}$ (vector field)
2. Value-conserving dynamics

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

Neural Transport Model

We model weather/climate as a spatiotemporal process **u**(x,t) = (u_1(x,t),…,u_K(x,t)) of K quantities as an advection PDE:

$$\dot{u}_k(x,t) = -\mathbf{v}_k^\top(x,t)\nabla u_k(x,t) - u_k(x,t)\mathrm{tr}(\nabla \mathbf{v}_k(x,t))$$



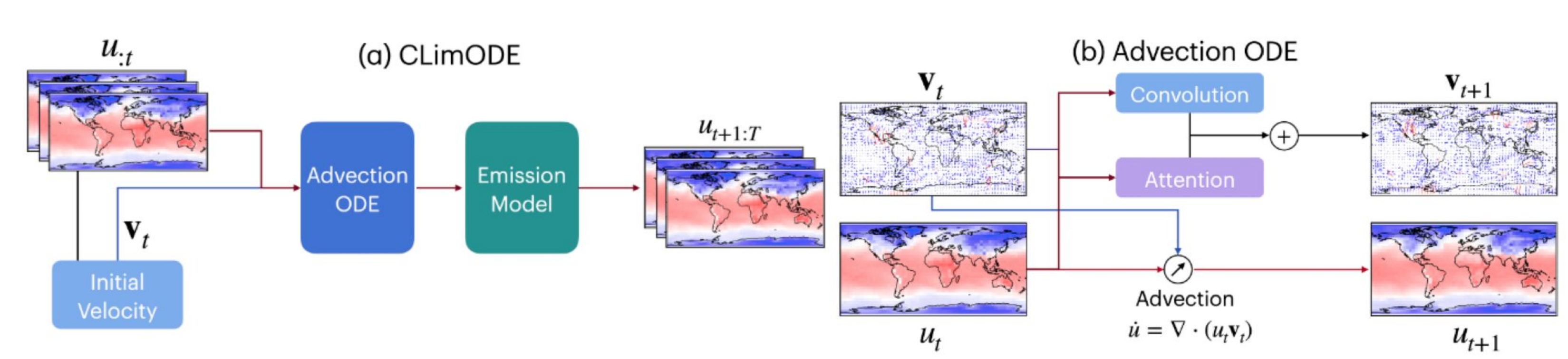
Flow velocity v

We model change in the velocity with a neural flow, as a function of observed values **u**(t), ∇**u** , **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 **u**(t) forward
2. Evaluate likelihood
3. Backpropagate wrt θ,ϕ

$$\begin{bmatrix}\mathbf{u}(t)\\ \mathbf{v}(t)\end{bmatrix} = \begin{bmatrix}\mathbf{u}(t_0)\\ \mathbf{v}(t_0)\end{bmatrix} + \int_{t_0}^{t}\begin{bmatrix}\dot{\mathbf{u}}(\tau)\\ \dot{\mathbf{v}}(\tau)\end{bmatrix} d\tau$$

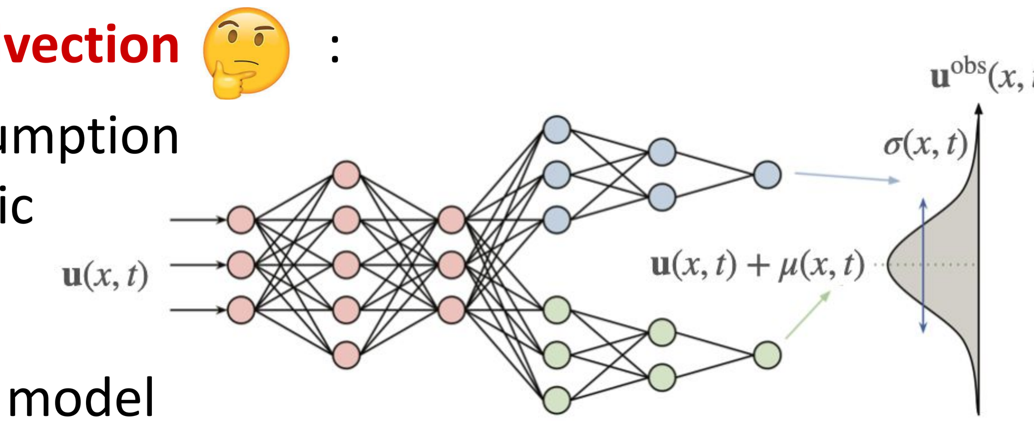
Sources and Uncertainty

Limitation of vanilla advection 🤔 :

1. Closed system assumption
2. ODE is deterministic

Idea 💡 :

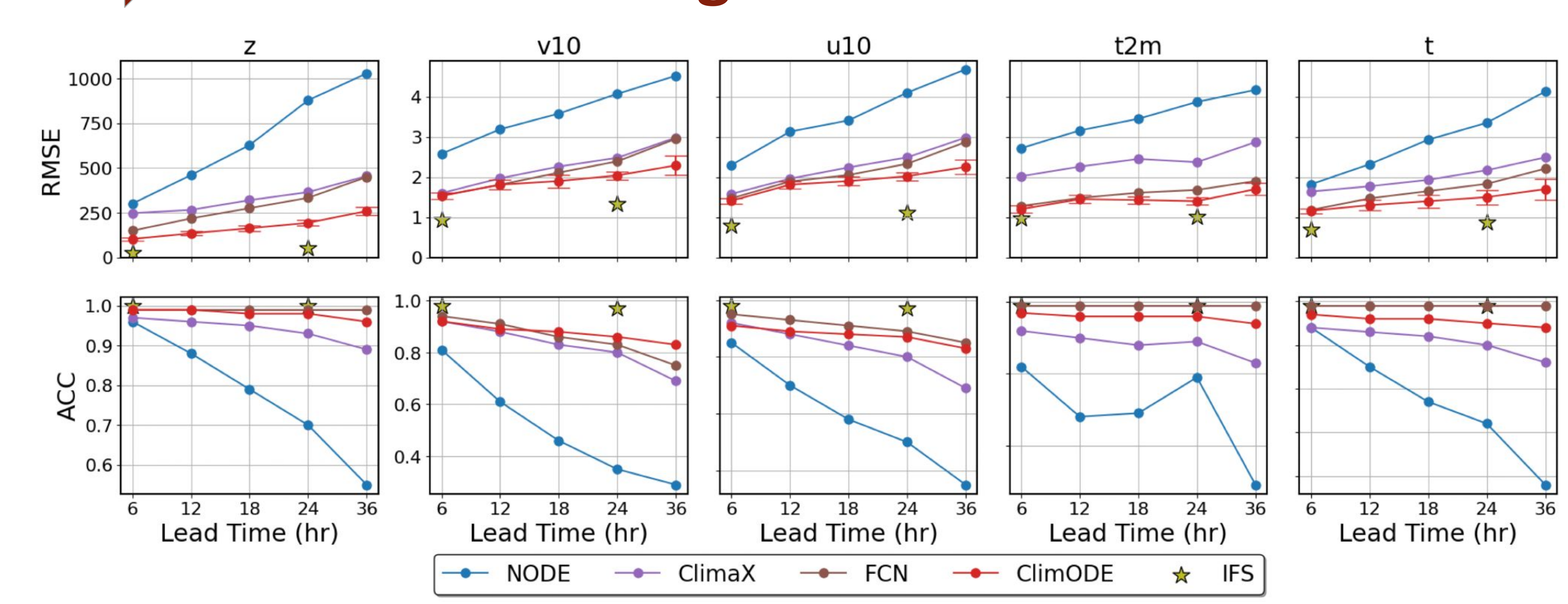
1. Gaussian emission model
2. Model predicts mean forecasts and uncertainty



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

Experiments

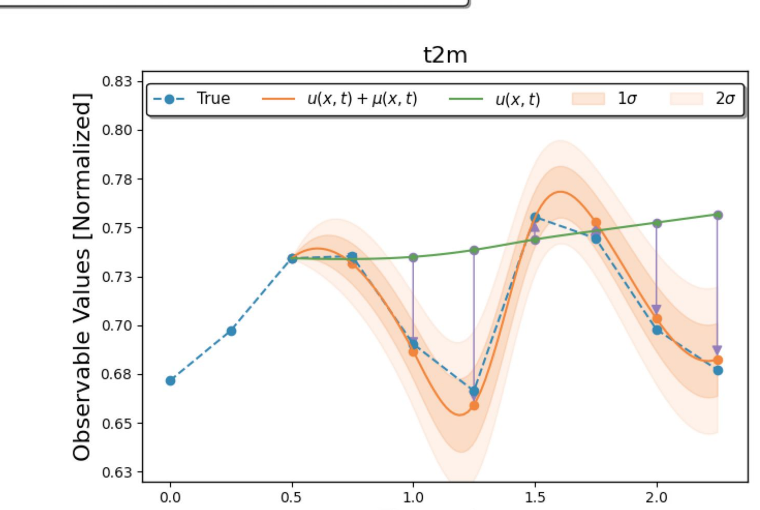
Global Forecasting

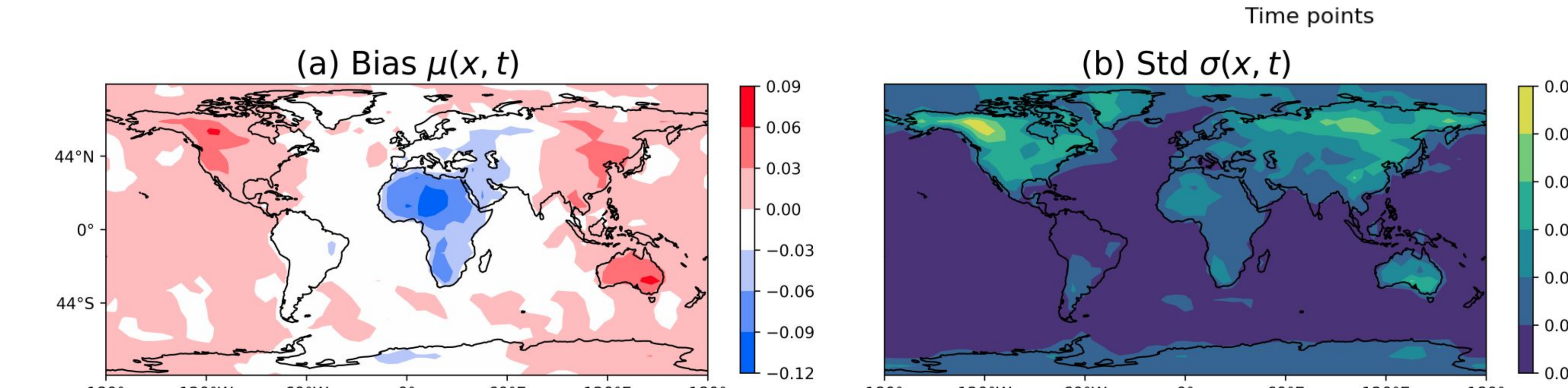


Interpretability

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