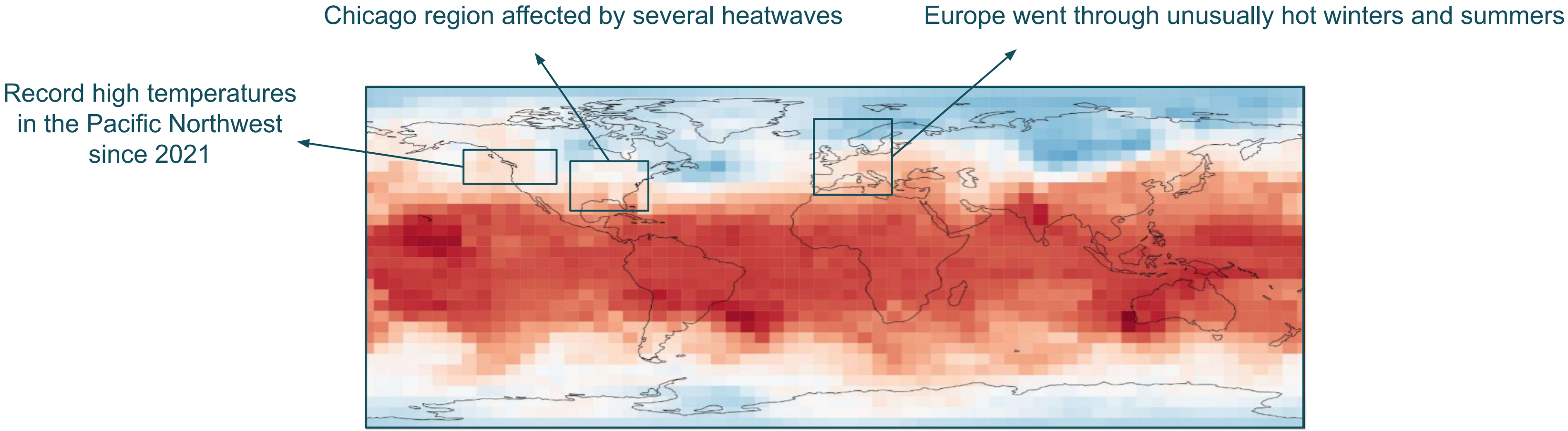
Learning Extreme Temperature Regimes

Shirin Goshtasbpour, Maxim Samarin, Michele Volpi
Swiss Data Science Center, ETH Zurich and EPFL

ICLR 2025 Workshop:
Tackling Climate Change with Machine Learning

The authors are part of SPEED2ZERO,
a Joint Initiative co-financed by the ETH Board.

Shifting Distribution of Temperature Anomalies



- Few observed extreme temperatures in collected datasets
- Expensive simulation via physics-based climate models

Leveraging generative models

- Simulation and analysis of possible extreme conditions and their impact pattern
- Stress tests and extreme scenarios for disaster preparedness
- Assist infrastructure planning and climate adaptation

Conditional Generative Modeling of Extreme Events

K-dimensional spatio-temporal climate process $u(\boldsymbol{x},t) = \big(u_1(\boldsymbol{x},t),\dots,u_K(\boldsymbol{x},t)\big)$

Time $t \in \mathbb{R}$ and location $\boldsymbol{x} \in \Omega = [-90^\circ, 90^\circ] \times [-180^\circ, 180^\circ]$, $u_k : \Omega \times \mathbb{R} \to \mathbb{R}$

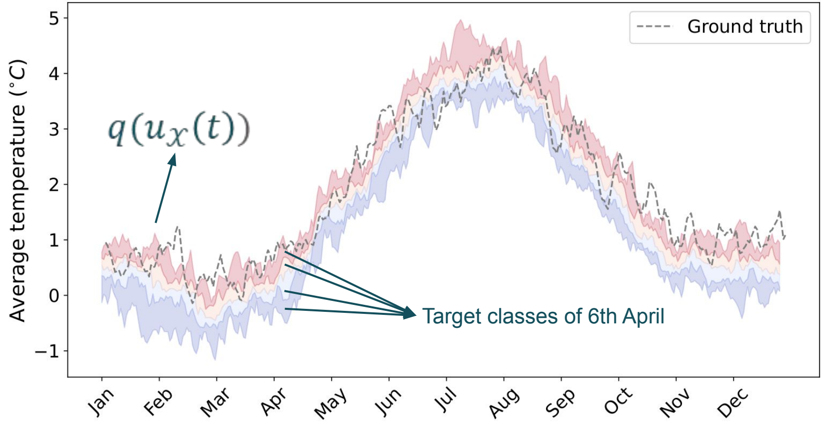
Restricted climate process $u_{\mathcal{X}}(t) = \big(u(\boldsymbol{x},t)\big)_{\boldsymbol{x}\in\mathcal{X}}$ to region $\mathcal{X}$

Query function q is a measure of $u_{\mathcal{X}}(t)$

- e.g. average temperature over $\mathcal{X}$

Query values consist of Q *target classes*

Each $u_{\Omega}(t)$ in the dataset is labeled y according to $q(u_{\mathcal{X}}(t))$



Using consecutive years of ERA5 dataset as independent day-of-year samples

→ Label assigned based on corresponding quartile bin index to $q(u_{\mathcal{X}}(t))$

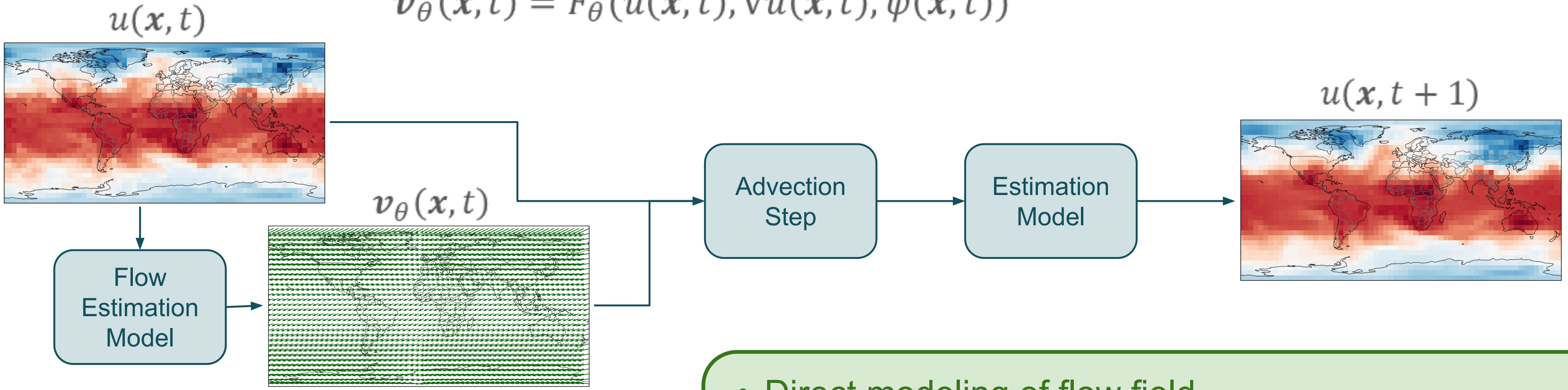
UQClimODE: Stochastic Advection Based Climate Modeling

Extension of ClimODE [Verma+24']

Advection evolution

$$du_k(\boldsymbol{x},t) = \boxed{-\nabla \cdot \big(u_k(\boldsymbol{x},t), \boldsymbol{v}_k(\boldsymbol{x},t)\big)dt} + \mu_\theta(\boldsymbol{x},y)dt + \sqrt{2\Sigma}dB_t$$

$$\boldsymbol{v}_\theta(\boldsymbol{x},t) = F_\theta(u(\boldsymbol{x},t), \nabla u(\boldsymbol{x},t), \psi(\boldsymbol{x},t))$$



- Direct modeling of flow field
- Tokenized with UNets for scalability and physical bias
- Teacher forcing with RMSE for training

Long-Term Stability and Competitive Short-Term Prediction

- ClimODE variants **struggle to stay stable** for long-term predictions
- Removing context info. and reducing observation frequency **deteriorates short-term results**
- UQClimODE is comparable to the baseline with daily observations without context
- UQClimODE **retains its performance** over 90 days

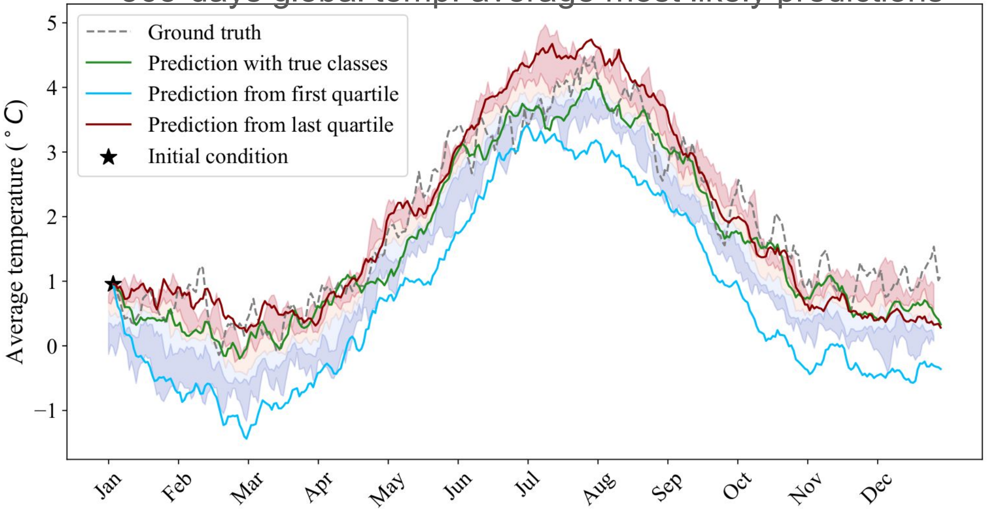
CRPS of predictions for global forecasting

Model	Lead time (days) 1	3	7	14	30	90
ClimODE (24)	**0.010** $\pm$ **0.010**	**0.017** $\pm$ **0.018**	0.034 $\pm$ 0.041	NaN	NaN	NaN
daily (24)	0.017 $\pm$ 0.017	0.025 $\pm$ 0.027	0.034 $\pm$ 0.039	0.079 $\pm$ 0.072	NaN	NaN
daily (6)	0.017 $\pm$ 0.021	0.022 $\pm$ 0.026	0.027 $\pm$ 0.037	0.031 $\pm$ 0.041	0.052 $\pm$ 0.061	NaN
daily (4)	0.019 $\pm$ 0.022	0.023 $\pm$ 0.026	**0.026** $\pm$ **0.032**	**0.029** $\pm$ **0.036**	0.041 $\pm$ 0.049	NaN
nocontext (24)	0.022 $\pm$ 0.062	0.028 $\pm$ 0.064	0.035 $\pm$ 0.069	0.052 $\pm$ 0.076	0.108 $\pm$ 0.169	NaN
nocontext (6)	0.033 $\pm$ 0.076	0.036 $\pm$ 0.077	0.039 $\pm$ 0.079	0.047 $\pm$ 0.081	0.077 $\pm$ 0.089	NaN
UQClimODE	0.015 $\pm$ 0.01	0.024 $\pm$ 0.027	0.031 $\pm$ 0.034	0.034 $\pm$ 0.037	**0.035** $\pm$ **0.037**	**0.034** $\pm$ **0.036**

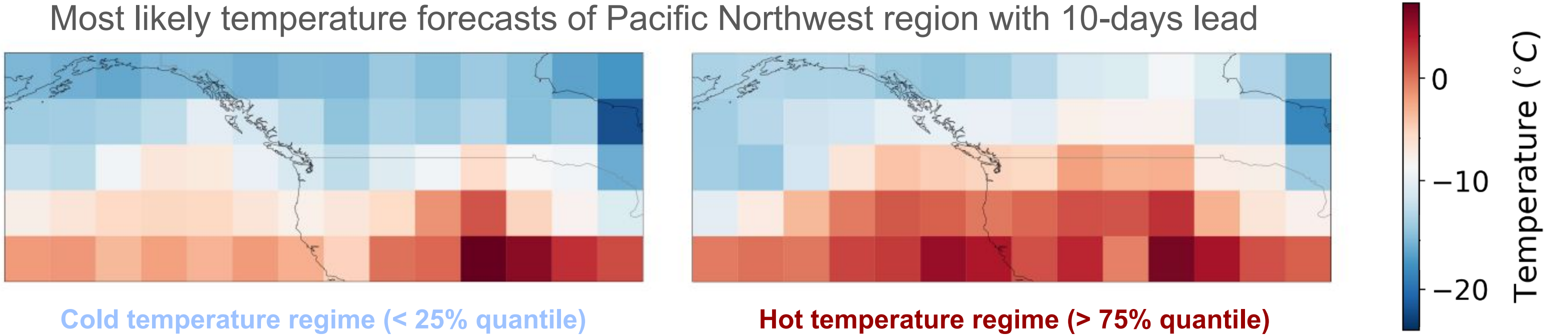
Enabling Targeted Generation

- Generation conditional on **first** (**last**) quartiles → **cold** (**hot**) temp. averages over the target region
- Temp. maps of predictions with regional target → different patterns of **cold** and **hot** values with sufficiently long lead time

365-days global temp. average most likely predictions



Most likely temperature forecasts of Pacific Northwest region with 10-days lead



Cold temperature regime (< 25% quantile)

Hot temperature regime (> 75% quantile)

Thank you!

shirin.goshtasbpour@sdsc.ethz.ch

The authors are part of SPEED2ZERO,
a Joint Initiative co-financed by the ETH Board.