GEOMETRIC DEEP LEARNING FOR MEDIUM-RANGE WEATHER PREDICTION

Icíar Lloréns Jover

Supervisors: Michaël Defferrard and Gionata Ghiggi
Professor: Pierre Vandergheynst
Expert: Peter Düben

WHAT IS MEDIUM-RANGE WEATHER PREDICTION AND WHY DO IT?

What is it?

2 days - 2 weeks forecasting

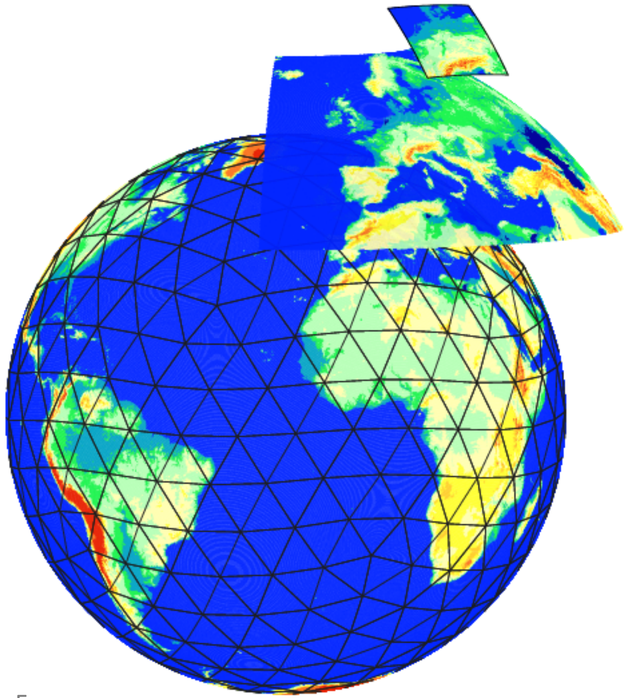
Usefulness

Economic **resource management**
→ preparation for **extreme weather events**

From https://america.cgtn.com/2018/08/17/the-heat-extreme-weather

CHALLENGES

- **Extended area** of influence
- Dependent on quality of **initial atmospheric state**
- Influence of **land and ocean**



From https://www.dwd.de/EN/research/weatherforecasting/num_modelling/06_nwp_emergency_response_system/num_weather_prediction_emergency_system.html

WHY USE DATA-DRIVEN METHODS?

- Current operational NWP: successful but needs a lot of computing power

- Data-driven methods provide:
  - **Flexible** models that can **automatically learn representations** from the data

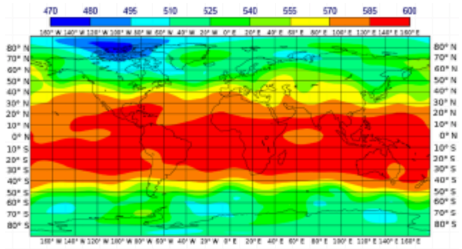
  - **Empirically good performance** when enough data

  - **Computationally cheaper** forecasts

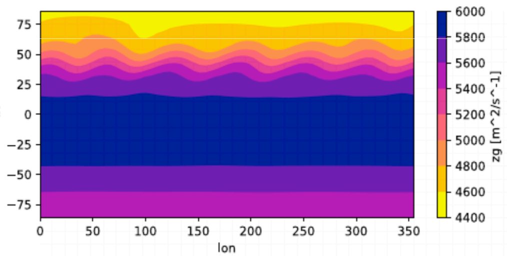
PREVIOUS WORKS

Planar projections

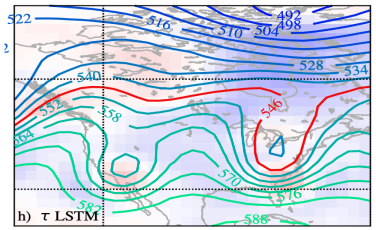
Düben and Bauer, 2018



Scher, 2018

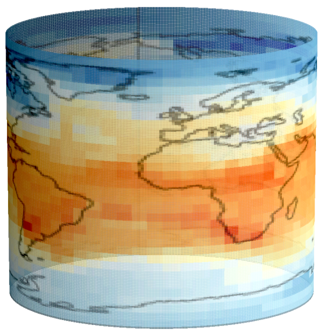


Weyn et al., 2019



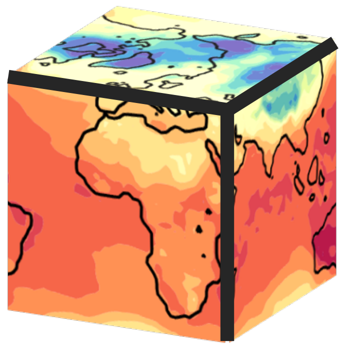
Spherical approximations

Rasp et al., 2020



Adapted from Rasp, S., Dueben, P. D., Scher S., Weyn, J. A., Mouatadid, S., and Thuerey, N. (2020). WeatherBench: A benchmark dataset for data-driven weather forecasting. arXiv.

Weyn et al., 2020

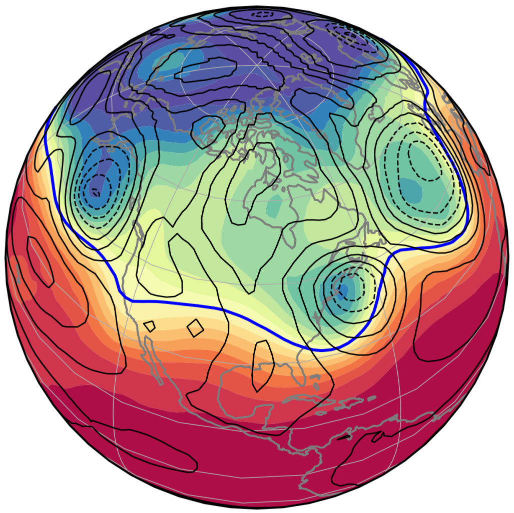


Adapted from Weyn, J . A., Durran, D. R , and Caruana, R. (2020). Improving data-driven global weather prediction using deep convolutional neural networks on a cubed sphere. JAMES.

OUR WORK

Objectives

- Use spherical domain with adapted spherical grid
- Implement computations on the sphere
- Include temporal dimension
- Informed feature selection
- Benchmark using a wide range of metrics
- Spatial evaluation using new metrics



Adapted from J. A. Weyn, D. R. Durran, and R. Caruana, 2020. Improving data-driven global weather prediction using deep convolutional neural networks on a cubed sphere. JAMES.

WEATHERBENCH : DATASET *(Rasp et al., 2020)*

	The ERA5 data	WeatherBench
• **Time span:**	**1979 to present**	**1979 to 2018**
• **Temporal resolution**:	**1 hour**	**1 hour**
• **Spherical resolution (lat/lon grid)**:	**0.25°**	***5.625°***, 2.8125° and 1.40525°
• **Vertical resolution:**	**37** pressure levels	**10** pressure levels
• **Atmospheric fields**:	**344**	**19**

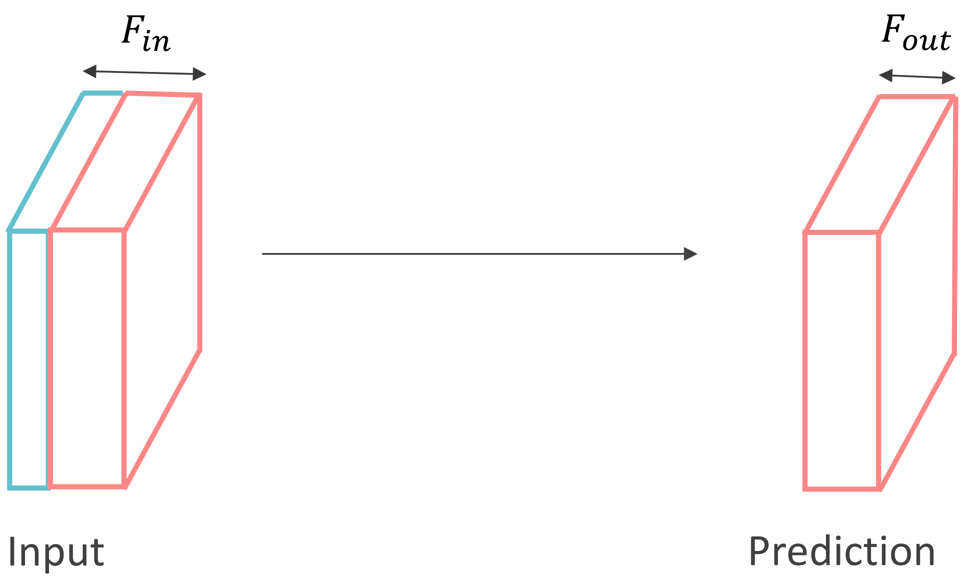
WEATHERBENCH : FEATURES

Static

- Constants
  - Soil type
  - Orography
  - Latitudes / longitudes
  - Land-sea mask
- Radiation

Dynamic

- Temperature
- Geopotential height
- Wind
- Humidity
- Vorticity



WEATHERBENCH: BENCHMARKING

Features: **Z500** and **T850**

Metrics (p_n, o_n are prediction and observation respectively):

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_n (p_n - o_n)^2}$$

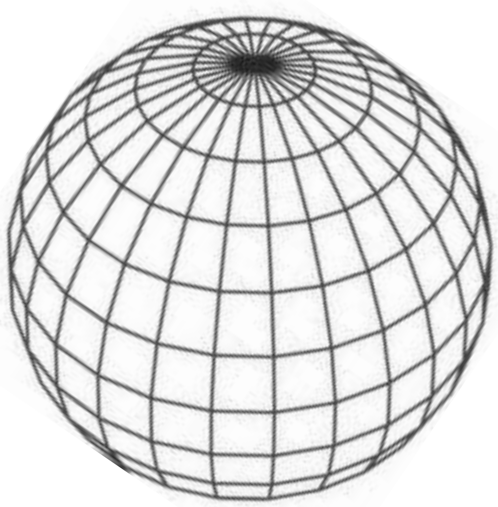
$$\text{MAE} = \frac{1}{N}\sum_n |p_n - o_n|$$

$$\text{ACC} = \frac{\sum_n p'_n o'_n}{\sum_n p'^2_n \ \sum_n o'^2_n}, \quad \text{with } x'_n = x_n - \frac{1}{N}\sum_n x_n$$

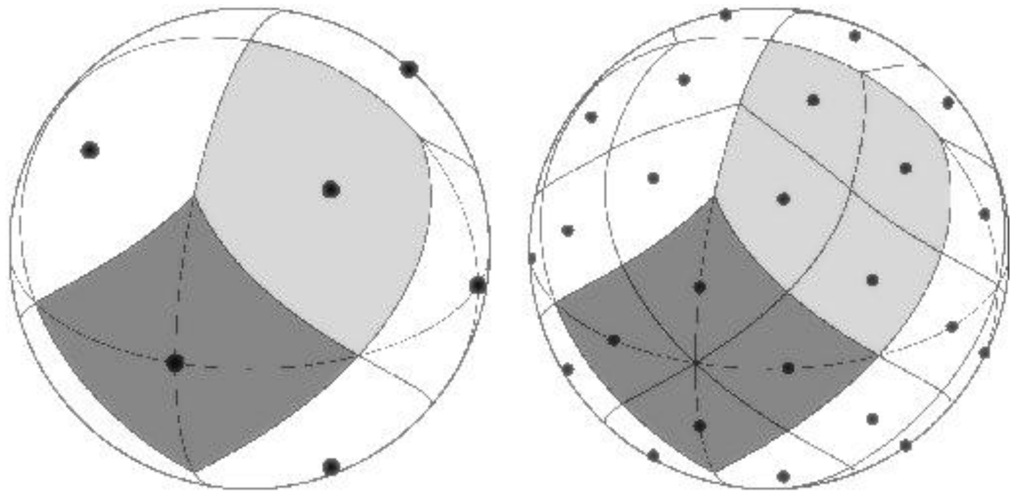
METHODOLOGY - OUTLINE

- Sphere discretization
- Spherical convolutions
- Temporal dimension
- Network architecture
- Training

SPHERE DISCRETIZATION



Equiangular (*Driscoll & Healy, 1994*)

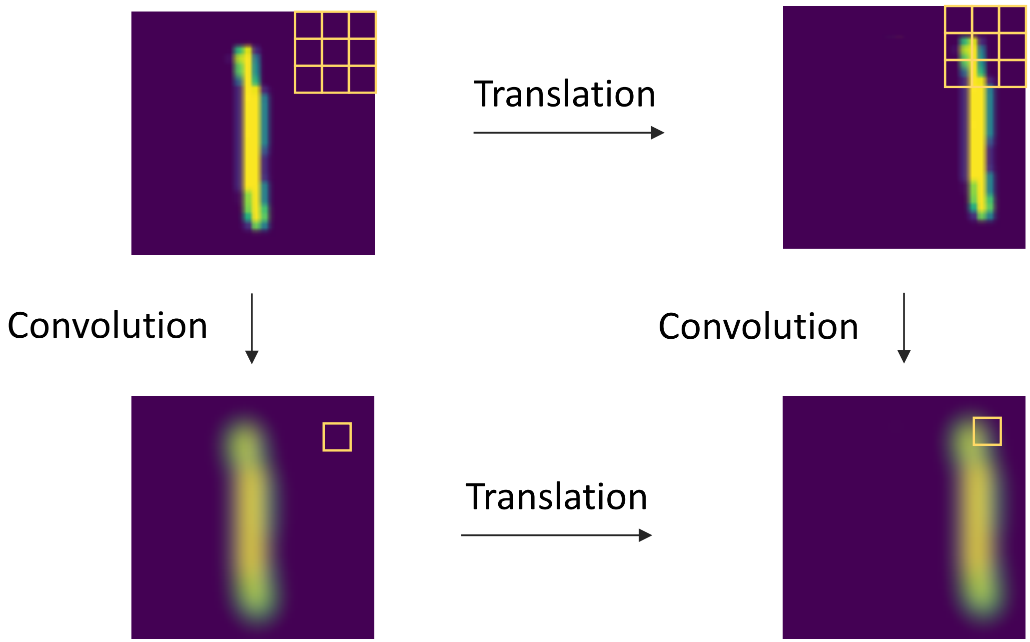


HEALPix (*Górski et al., 2005*)

- **Avoids oversampling** at the poles
- **Not** region **dependent**

EQUIVARIANT CONVOLUTIONS

On images

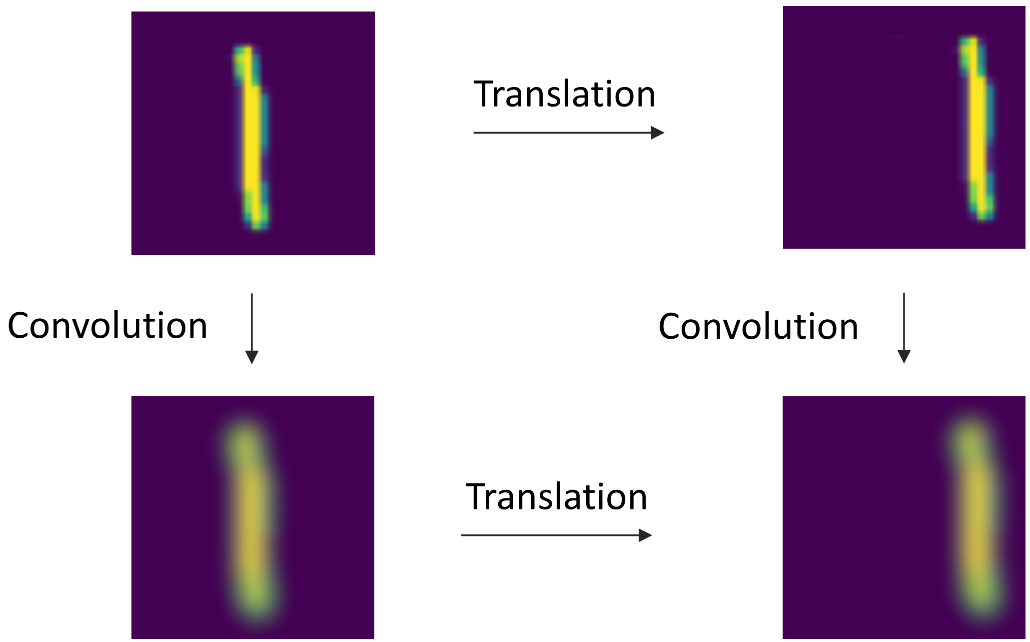


➡ 2D shift makes convolutions **translation equivariant**:

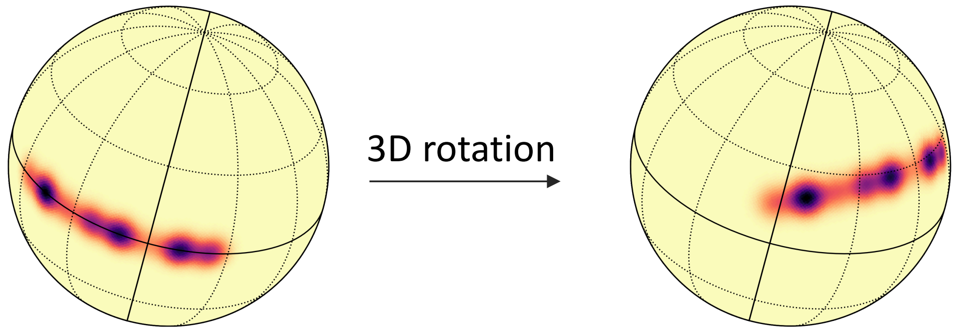
- Fewer parameters
- Not dependent on location
  - More robustness
  - No data augmentation

EQUIVARIANT CONVOLUTIONS

On images



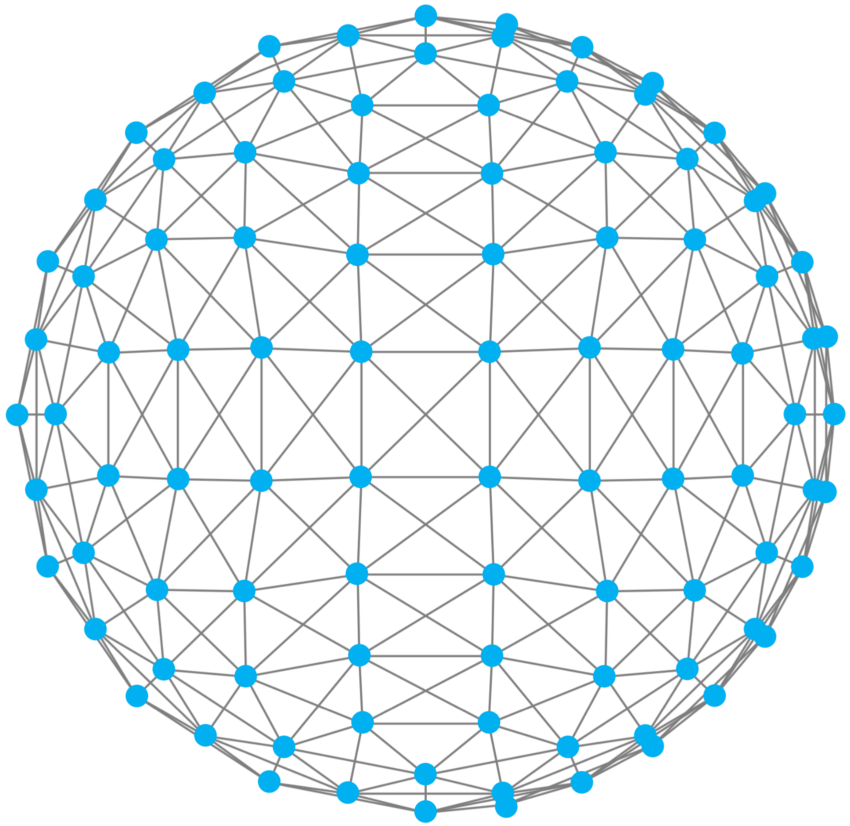
On the sphere



From Defferrard, M (2020). DeepSphere: a graph-based spherical CNN. ICLR'20 spotlight. https://zenodo.org/record/3777976#.XvtpzS2Q3ys

➡ Goal: **replace 2D** translation convolutions **by SO(3)** rotation convolutions

GRAPH SPHERICAL CONVOLUTIONS



Discrete domain: graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$

- $\mathcal{V}$: vertices
- $\mathcal{E}$: edges (weighted according to distance)
- A: adjacency (edge weights)
- D: degree matrix (neighbors)

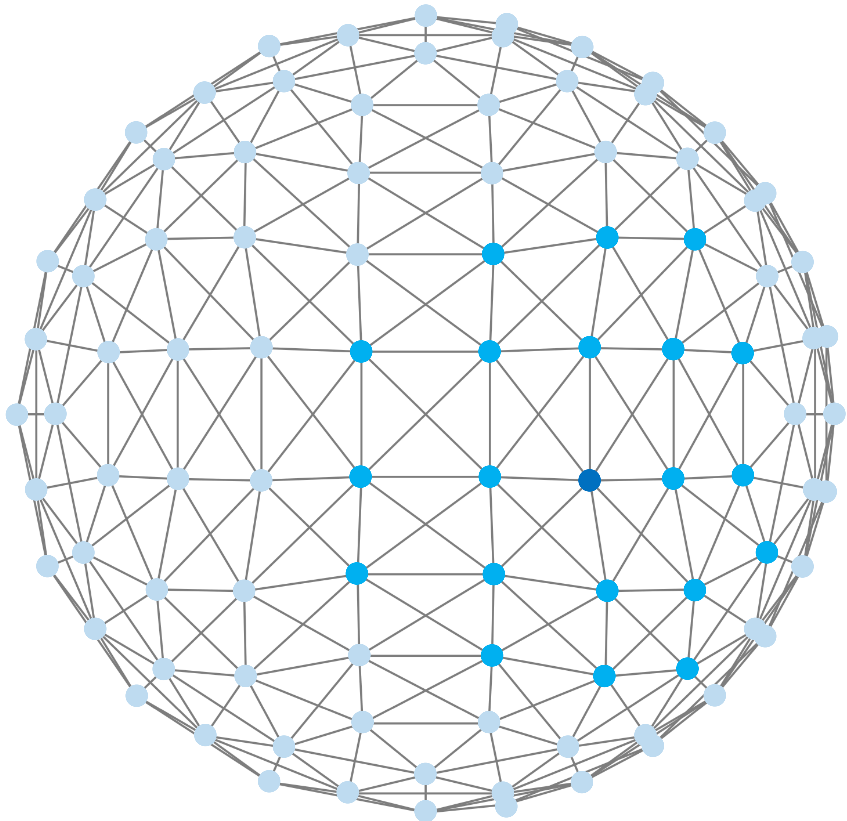
Spherical convolutions (Defferrard et al., 2016)

- Laplacian: $L = D - A$
- L is diagonalized by the Fourier basis
- Convolutions are multiplications in Fourier space
- For a signal x and a kernel g_α:

$$y = g_\alpha(L)\, x$$

➡ $\mathcal{O}(n^2)$ operations in general

GRAPH SPHERICAL CONVOLUTIONS



Discrete domain: graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, A)$

- $\mathcal{V}$: vertices
- $\mathcal{E}$: edges (weighted according to distance)
- A: adjacency (edge weights)
- D: degree matrix (neighbors)

Spherical convolutions (Defferrard et al., 2016)

- Laplacian: $L = D - A$
- L is diagonalized by the Fourier basis
- Convolutions are multiplications in Fourier space
- For a signal x and a kernel g_α:

$$y = g_\alpha(L)\, x$$

➡ $\mathcal{O}(n^2)$ operations in general

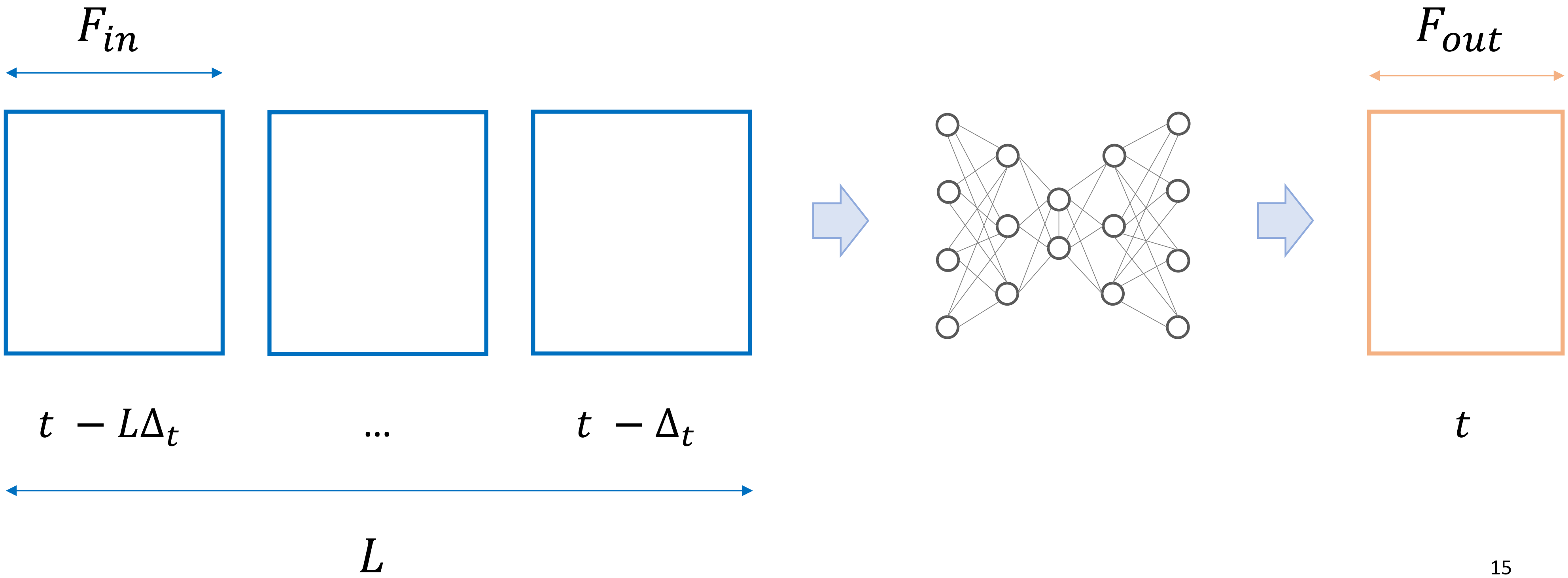
Rotation equivariance – Cost compromise: L is sparse → less neighbors

➡$\mathcal{O}(n)$ operations

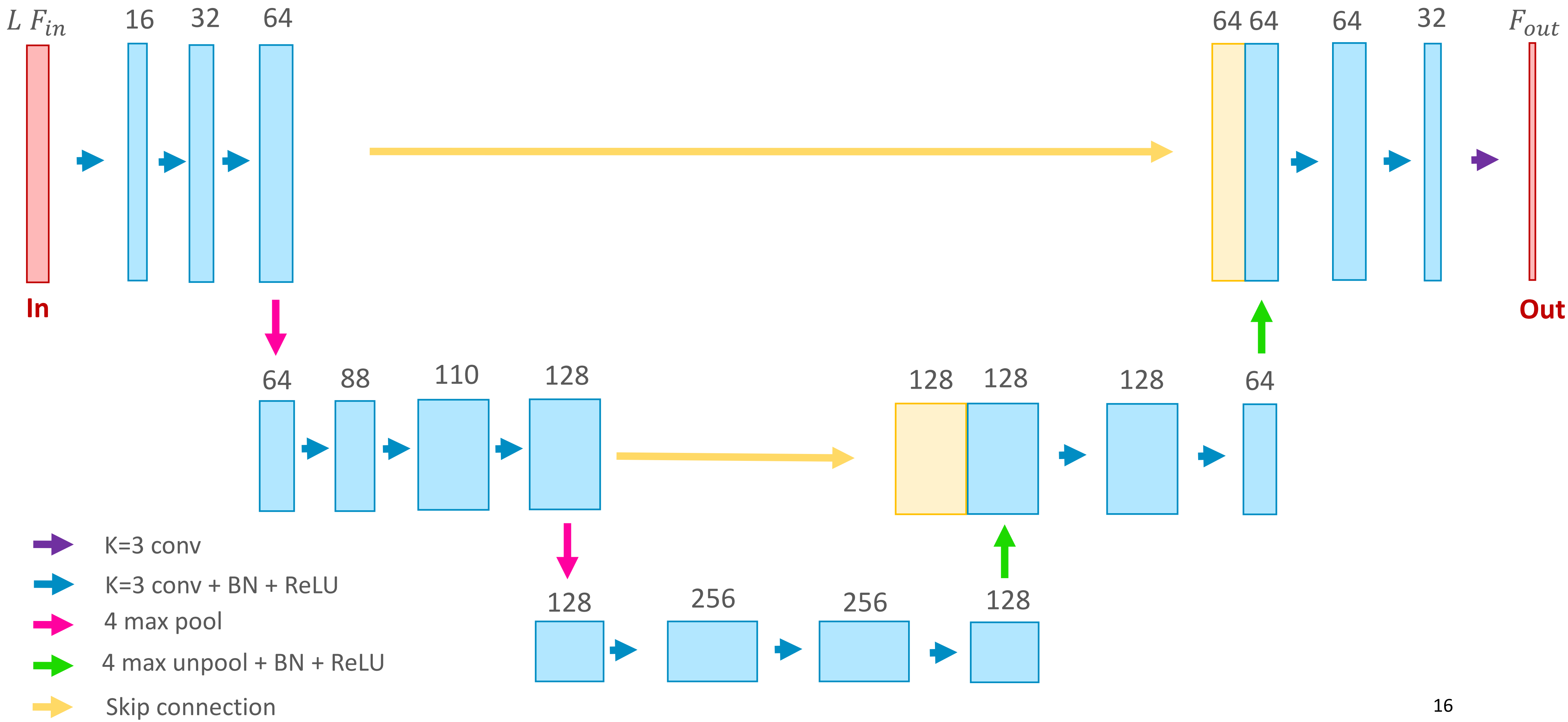
TEMPORAL DIMENSION

L : **sequence length**

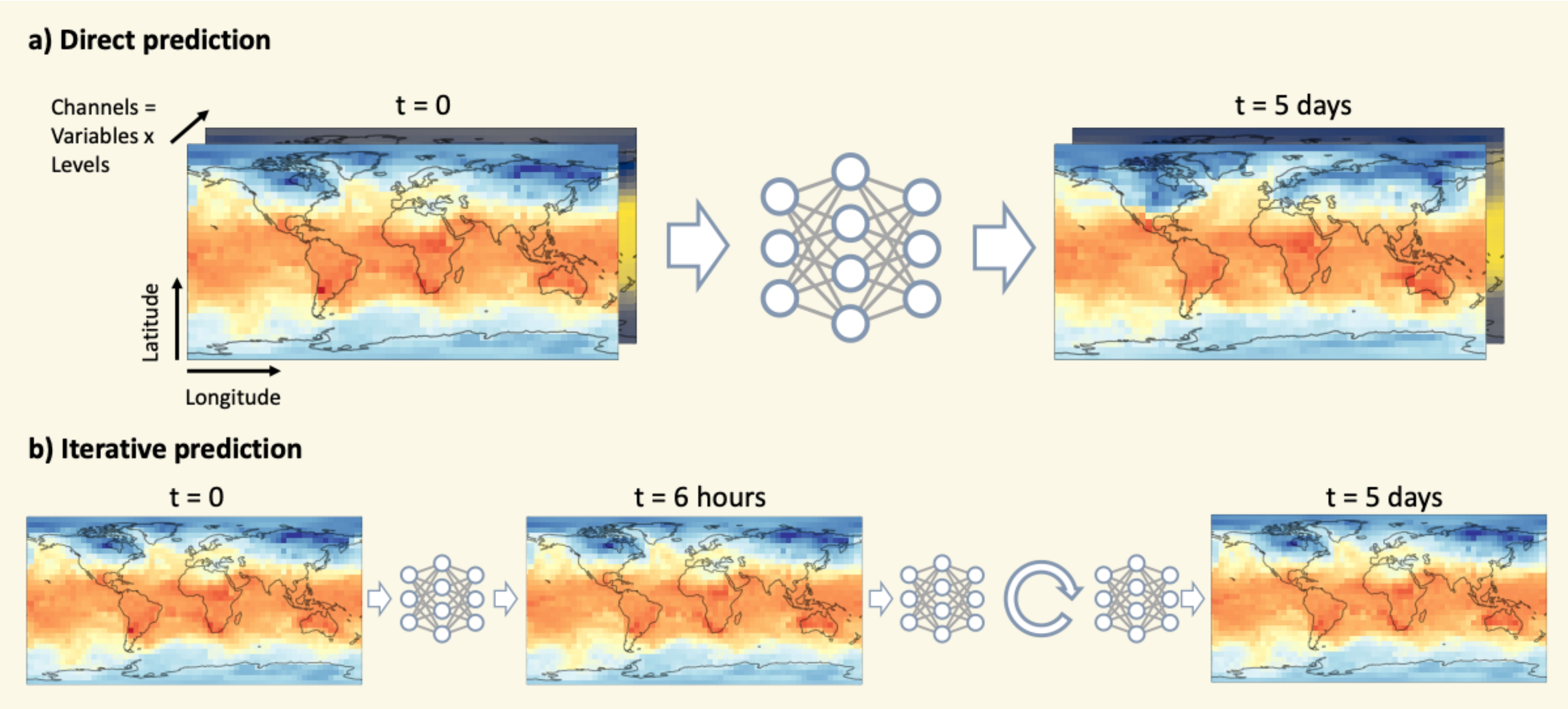
Δ_t : **temporal discretization** and **forecast lead time**



NETWORK ARCHITECTURE

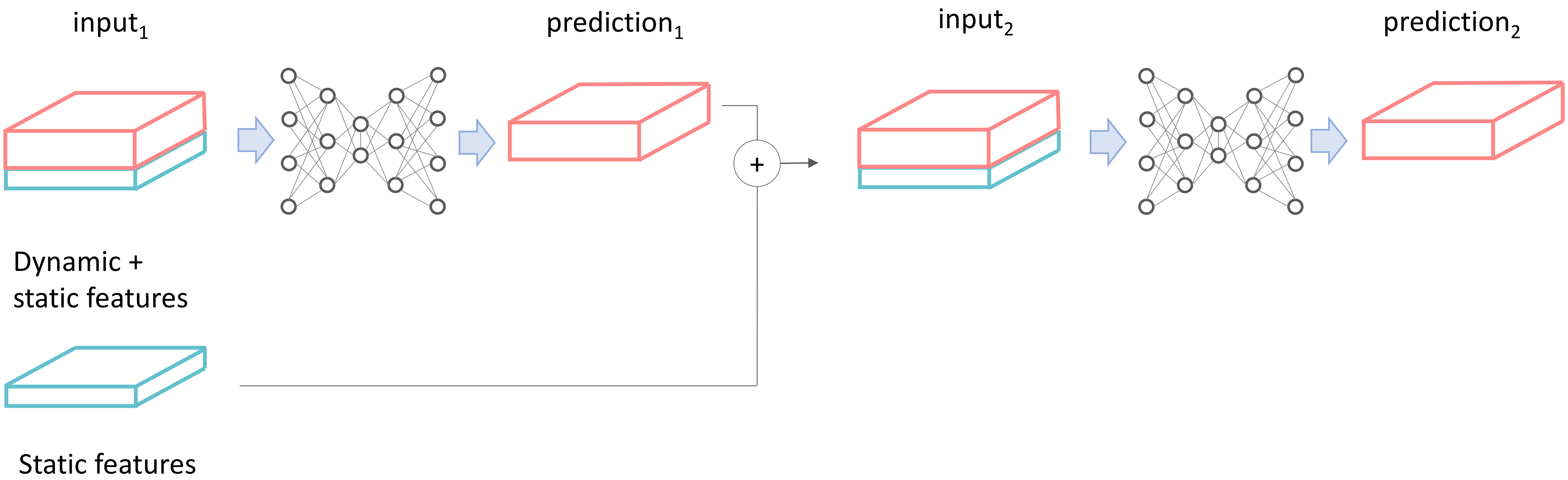


ITERATIVE PREDICTIONS



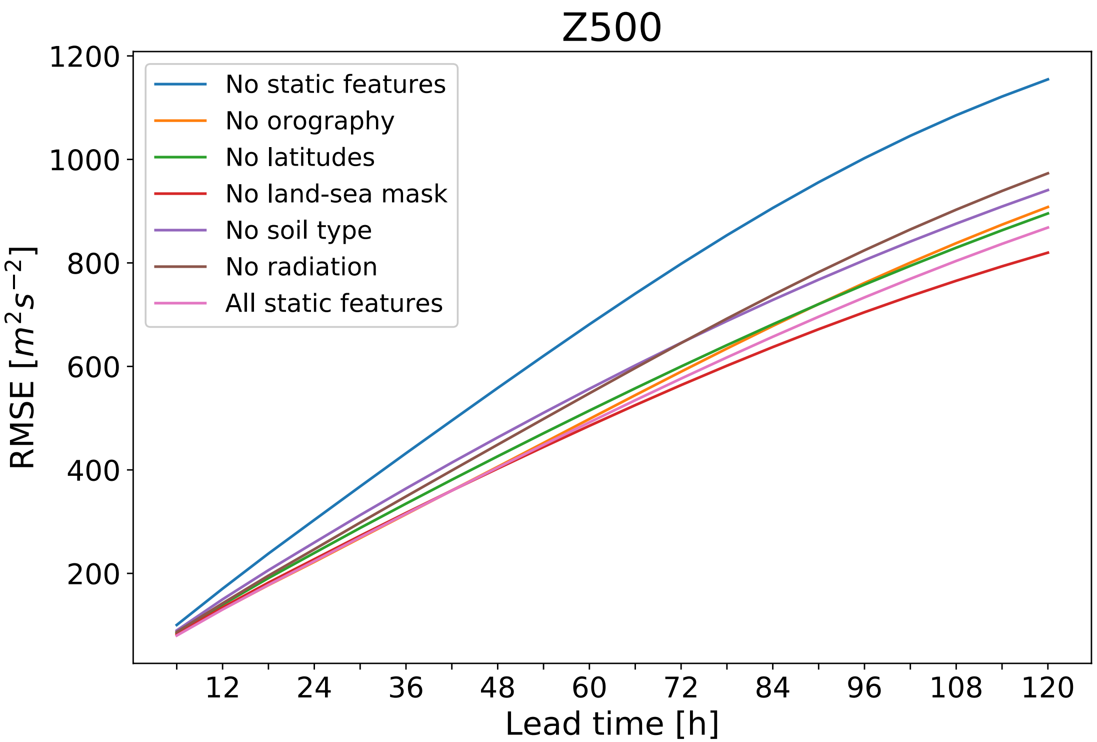
From Stephan Rasp, Peter D. Dueben, Sebastian Scher, Jonathan A. Weyn, Soukayna Mouatadid, and Nils Thuerey, 2020. WeatherBench: A benchmark dataset for data-driven weather forecasting. arXiv.

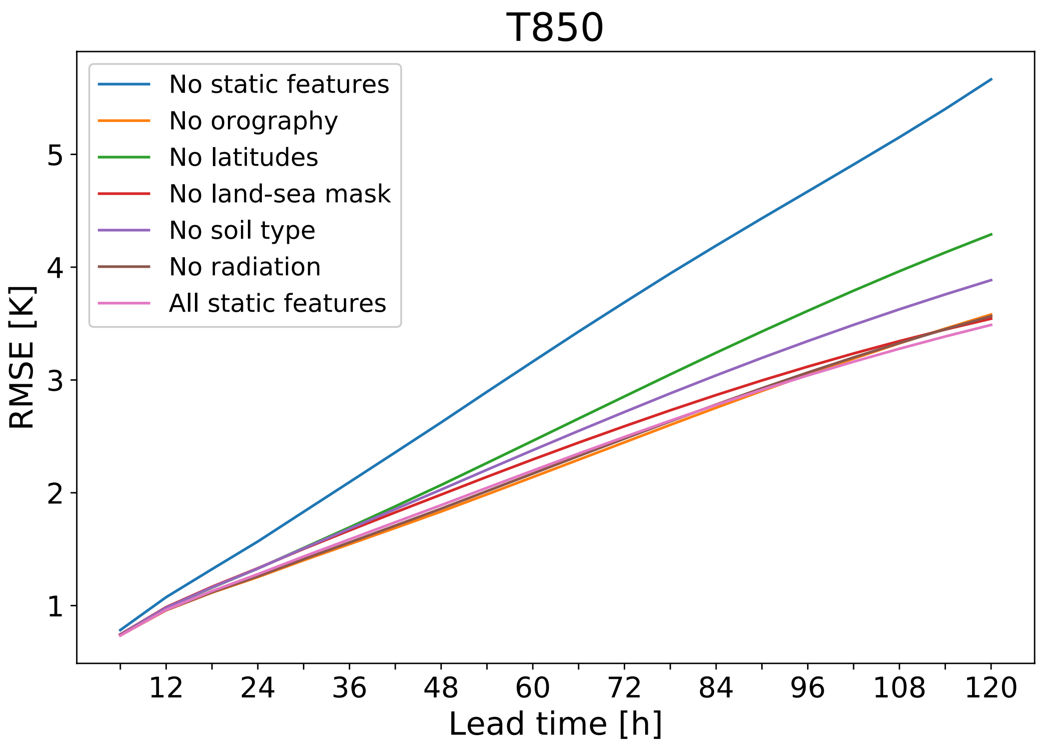
ITERATIVE TRAINING



$$Loss = \alpha_1 \, \mathrm{MSE}(pred_1,\ obs_1) \ + \ \alpha_2 \, \mathrm{MSE}(pred_2,\ obs_2)$$

STATIC FEATURE IMPORTANCE

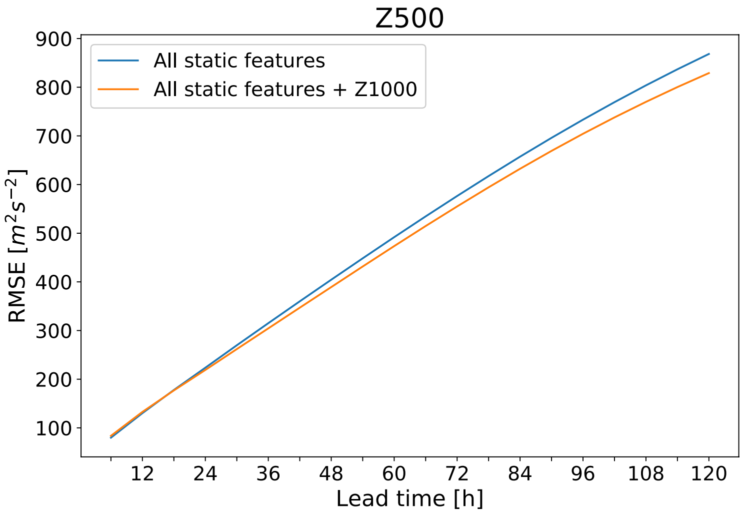


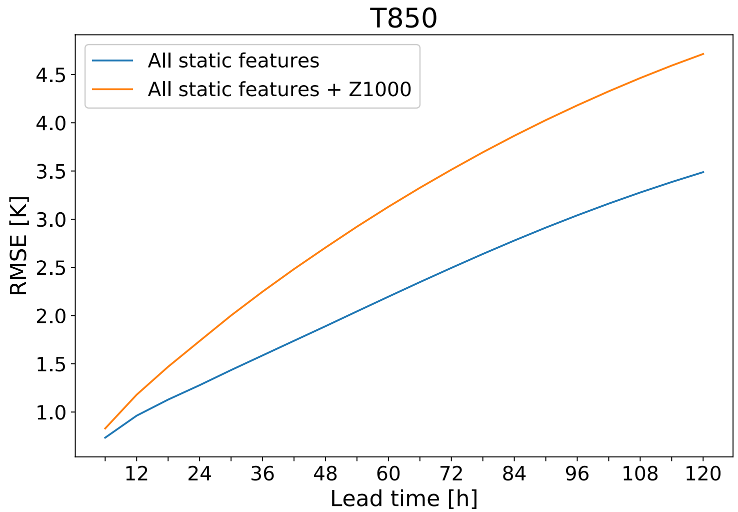


➡ Static features help with first direct prediction and decrease the error slope for subsequent iterations
➡ Z500 and T850 benefit from different static features

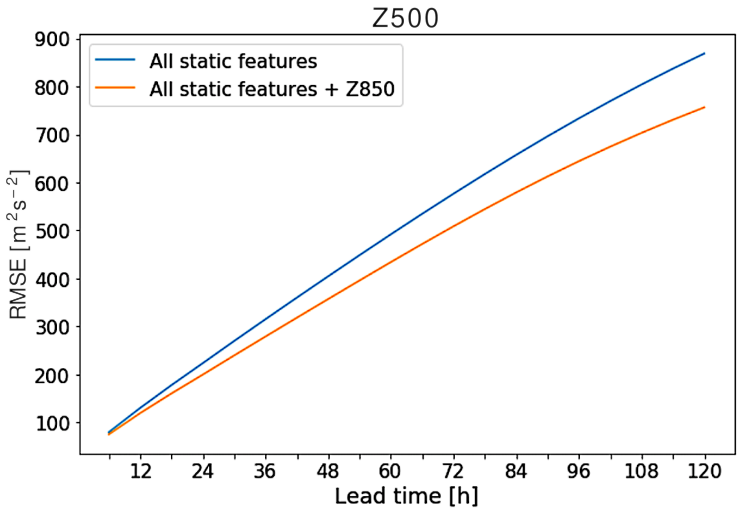
DYNAMIC FEATURES

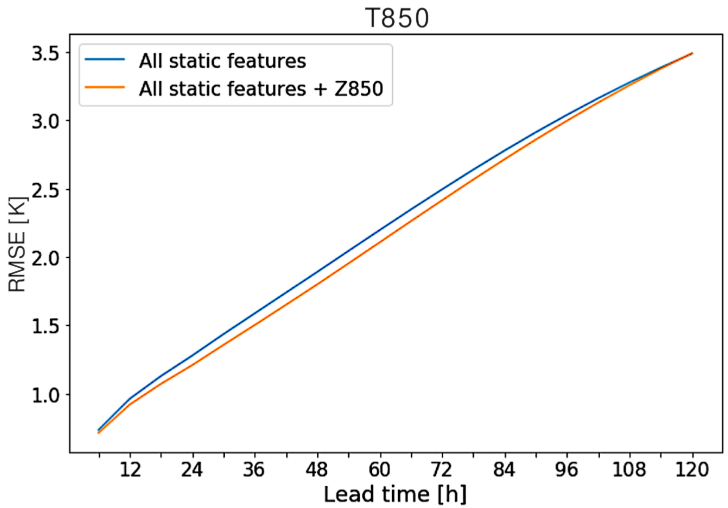
Z1000





Z850

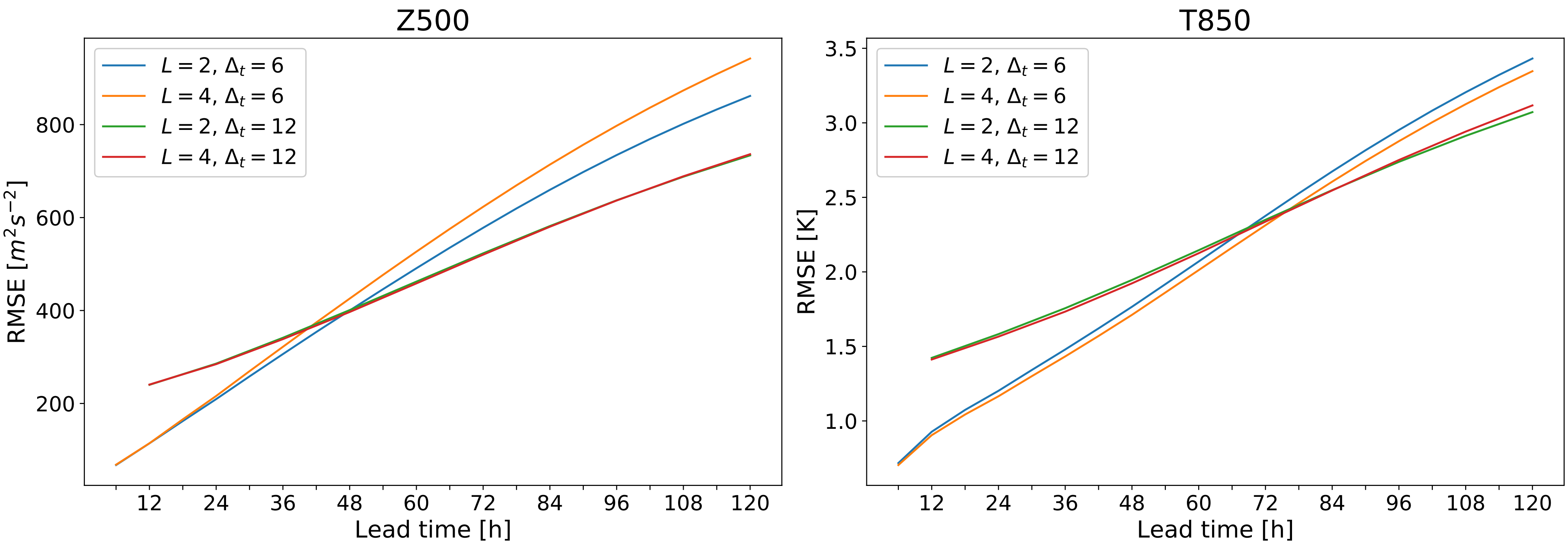




SEQUENCE LENGTH AND TIME RESOLUTION

L : sequence length

Δ_t : temporal resolution

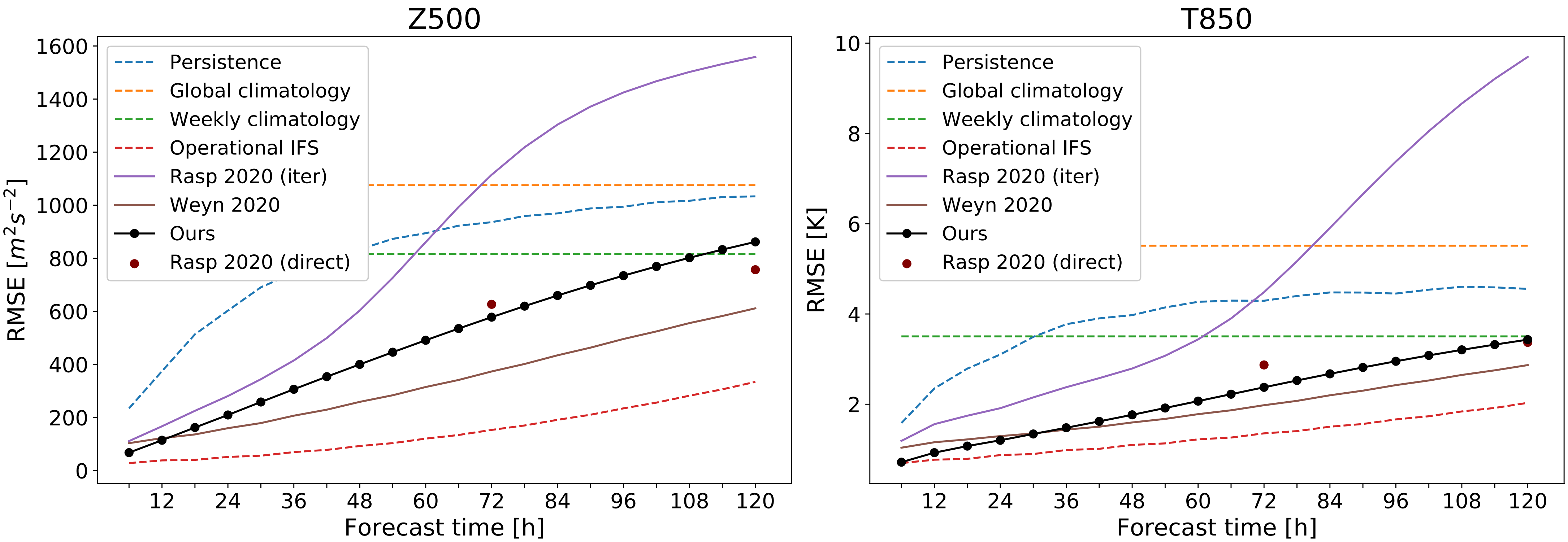


➡ No clear benefit from using a large L

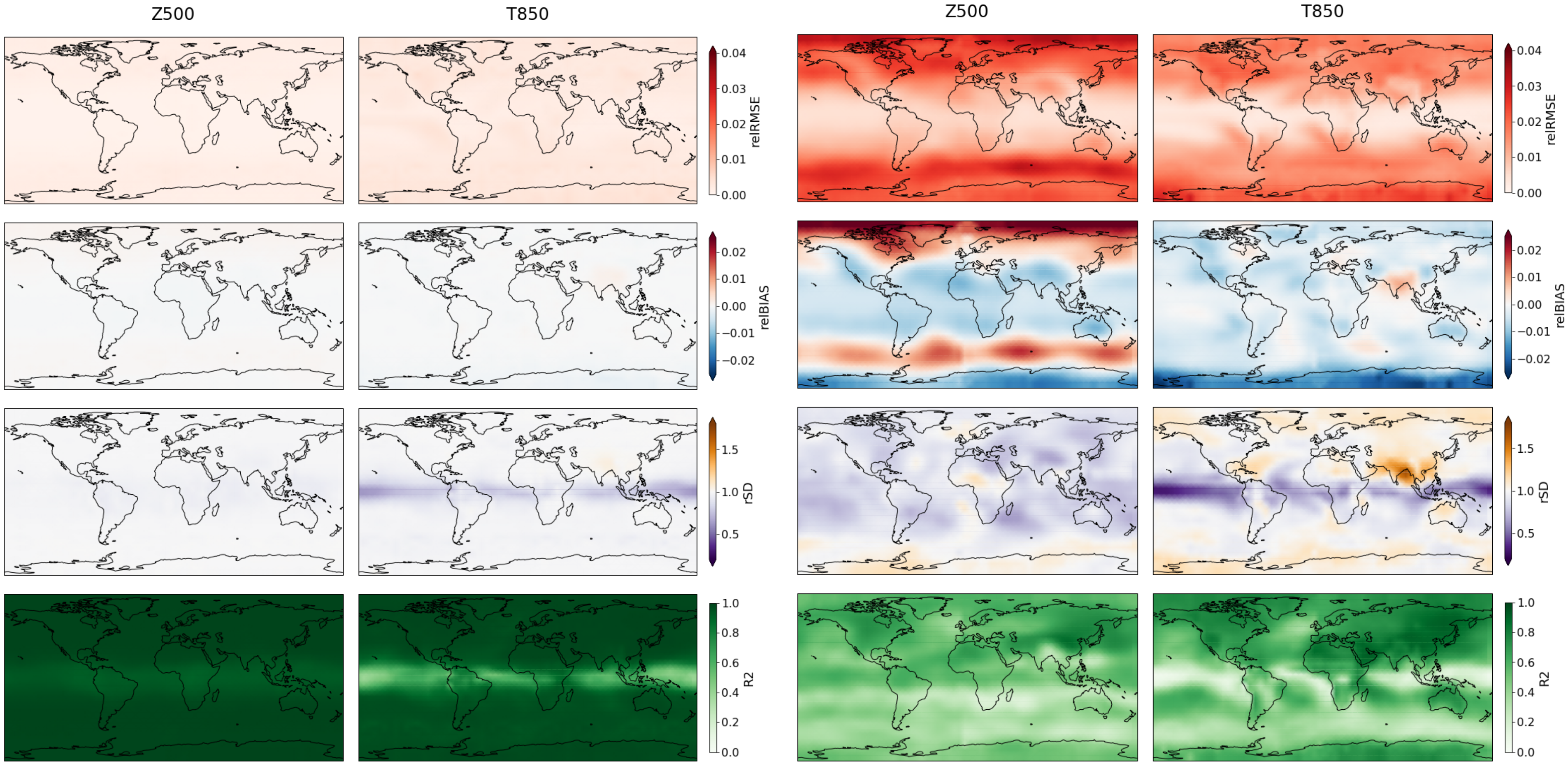
➡ Short term predictions require small Δ_t, error stabilizes in the long term with larger Δ_t.

BENCHMARK - RMSE

Best model: all static features, $L = 2$, $\Delta_t = 6\text{h}$



SKILL SPATIAL DISTRIBUTION



Lead forecast time: 6h

Lead forecast time: 120h

FUTURE RESEARCH LINES

→ Improve predictive skill

- Better usage of the space
  - **Atmosphere depth**: systematic integration of several pressure levels
- **Multi-scale**
  - **Temporal** dimension
  - **Vertical** dimension
- More sofisticated **architecture**
  - ResNets

→ Add a **notion of confidence** to the predictions

THANKS FOR THE INVOLVEMENT

Supervisors: Gionata Ghiggi and Michaël Defferrard

Professor: Pierre Vandergheynst

Expert: Peter Düben

REFERENCES

Bauer, P., Thorpe, A., and Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, 525:47–55.

Defferrard, M., Milani, M., Gusset, F., and Perraudin, N. (2020). DeepSphere: a graph-based sphericalCNN. InInternational Conference on Learning Representations.

Defferrard, M., Bresson, X., andVandergheynst, P. (2016). Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. Advances in Neural Information Processing Systems (NIPS).
Jonathan A. Weyn, Dale R. Durran, and Rich Caruana, 2020. Improving data-driven global weather prediction using deep convolutional neural networks on a cubed sphere. Journal of Advances in Modeling Earth Systems (JAMES).

Düben, P. and Bauer, P. (2018). Challenges and design choices for global weather and climate models based on machine learning. Geoscientific Model Development, 11:3999–4009

Haupt S.E. (2018) Short-Range Forecasting for Energy. In: Troccoli A. (eds) Weather & Climate Services for the Energy Industry. Palgrave Macmillan, Cham

Rasp, S., Dueben, P. D., Scher, S., Weyn, J. A., Mouatadid, S., and Thuerey, N. (2020). WeatherBench: A benchmark dataset for data-driven weather forecasting. arXiv.

Scher, S. (2018). Toward data-driven weather and climate forecasting: Approximating a simple general circulation model with deep learning. Geophysical Research Letters, 45.

Scher, S. and Messori, G. (2019). Weather and climate forecasting with neural networks: using generalcirculation models (GCMs) with different complexity as a study ground. Geoscientific Model Development,12:2797–2809

Weyn, J., Durran, D. R., and Caruana, R. (2019). Can machines learn to predict weather? Usingdeep learning to predict gridded 500-hpa geopotential height from historical weather data.Journal ofAdvances in Modeling Earth Systems, 11(8):2680–2693.

DATASET DIVISION

- **Training**: 1979 - 2012
- **Validation**: 2013 - 2016
- **Testing**: 2017 - 2018