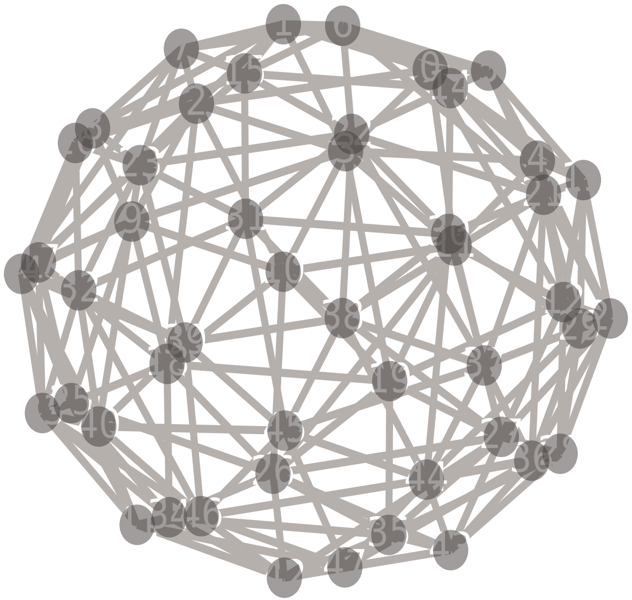


Probabilistic Deep Learning on Spheres for Weather/Climate Applications

Speakers

Wentao Feng

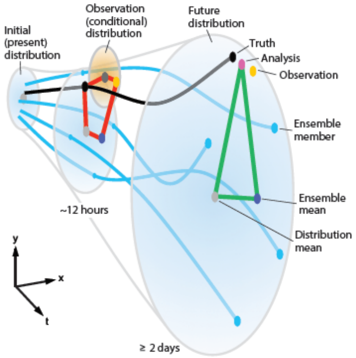
Yann Yasser Haddad

Supervisors

Michaël Defferrard

Gionata Ghiggi

16 December 2020



Introduction

Geometric projection

Spherical modeling

Accuracy and efficiency

Why go probabilistic?

Methods

Results

Conclusion and future work

Wentao Feng / Yann Yasser Haddad

Introduction

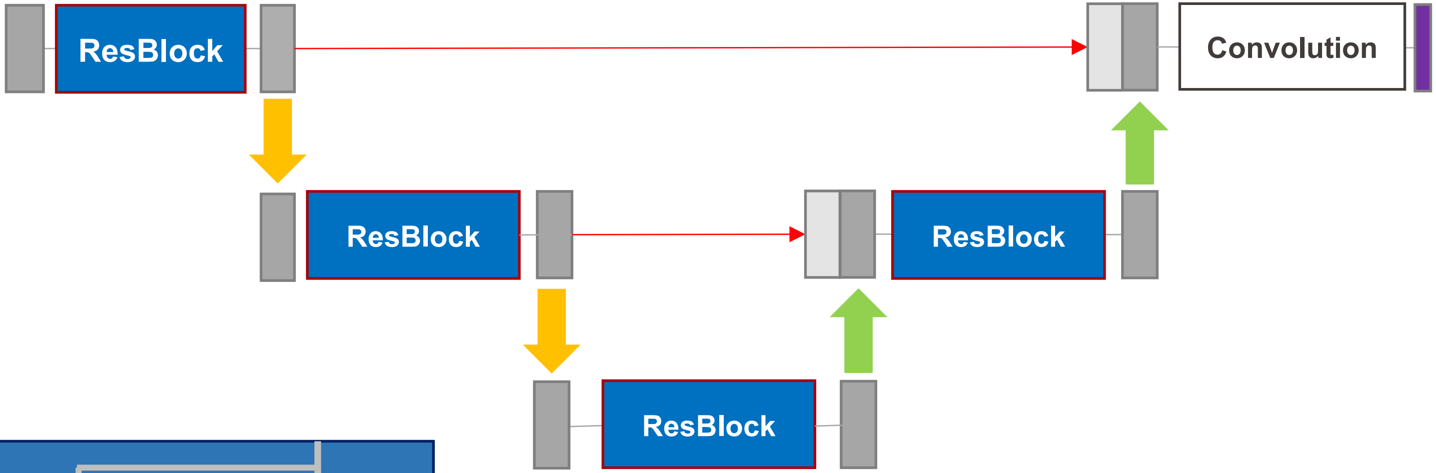
Our motivation:

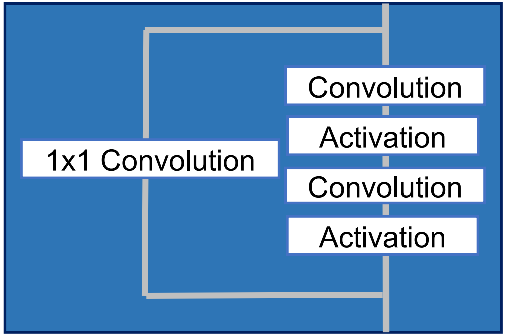
1. Deep learning method cost less computation power than NWP models
2. Can help us understand the impact of initial conditions
3. Can be used on other atmospheric problem

Our aim:

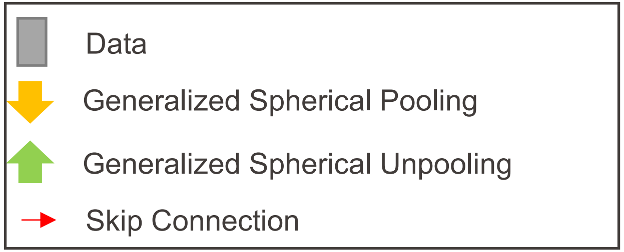
1. Design interpretable modeling method
2. Predict the weather with deep neural network

Introduction – U Net





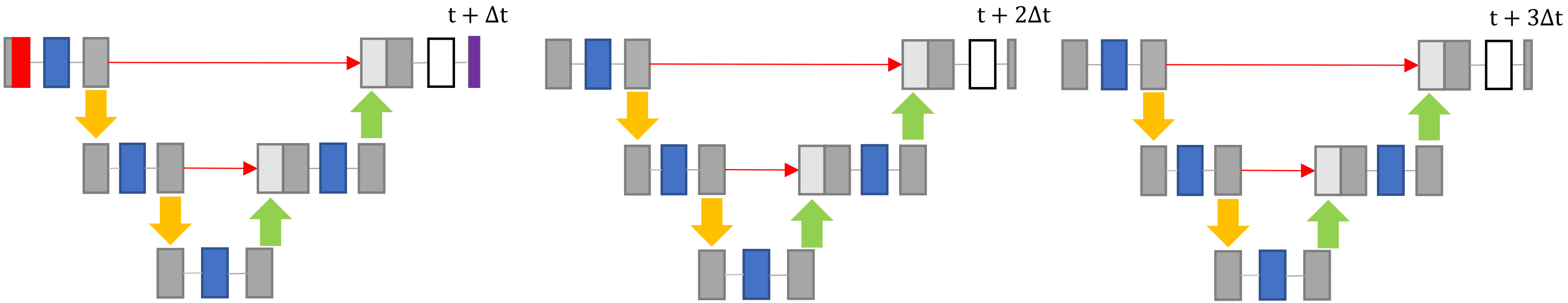
ResBlock



Wentao Feng / Yann Yasser Haddad

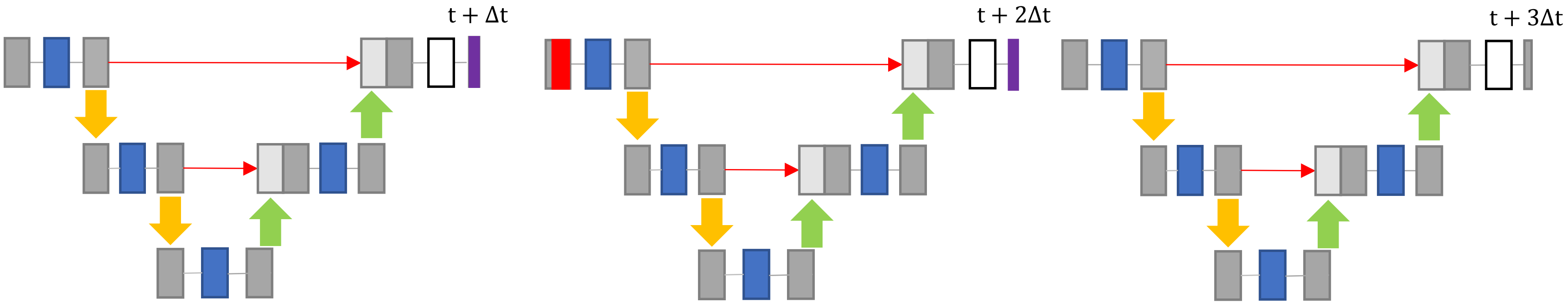
Gionata Ghiggi

Introduction – Autoregressive model



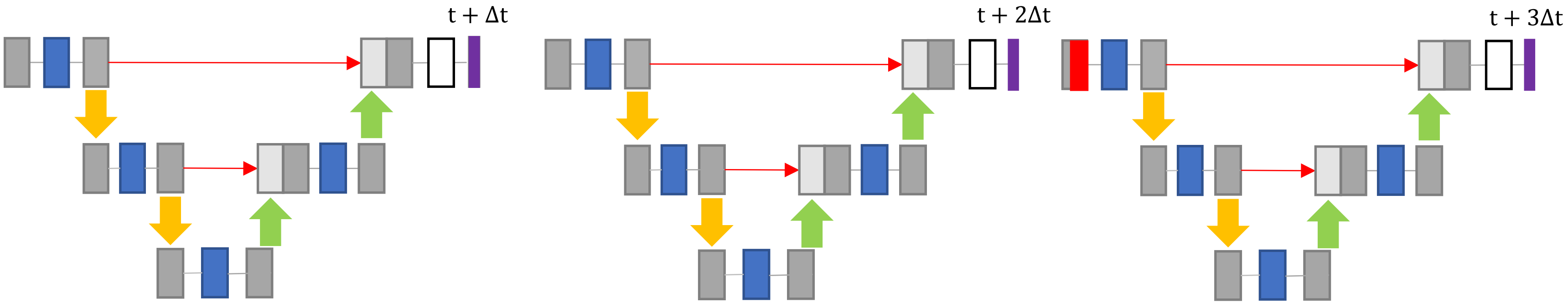
$\Delta t = 6$ hours

Introduction – Autoregressive model



$\Delta t = 6$ hours

Introduction – Autoregressive model



$\Delta t = 6$ hours

Loss function = MSE(,)

Predictions
Observations

Wentao Feng / Yann Yasser Haddad

Introduction – Input/Output

Input at time t
1. geopotential at 500 hPa (Z500)
2. temperature at 850 hPa (T850)

Additional feature at time t:
1. top-of-atmosphere solar radiation
2. orography
3. land-sea mask
4. latitude
5. soil type

Predict at $t + \Delta t$
1. geopotential at 500 hPa (Z500)
2. temperature at 850 hPa (T850)

Introduction – Metric

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=0}^{N}(p_n - o_n)^2}$$

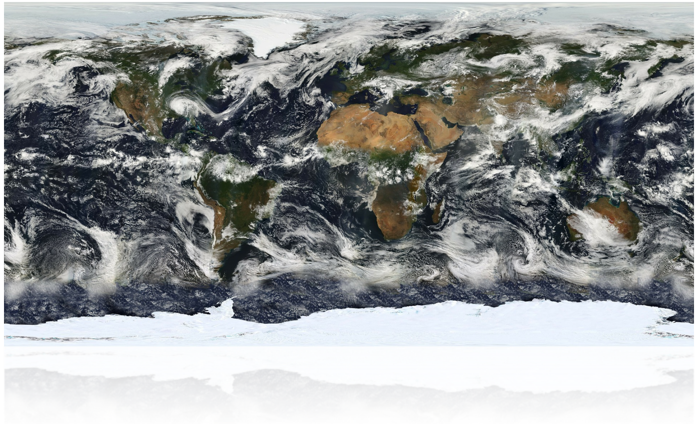
- p_n = predicted value at location n
- o_n = observed value at location n
- Measure the distance between prediction and observation
- Lower is better

Rasp et al., 2020

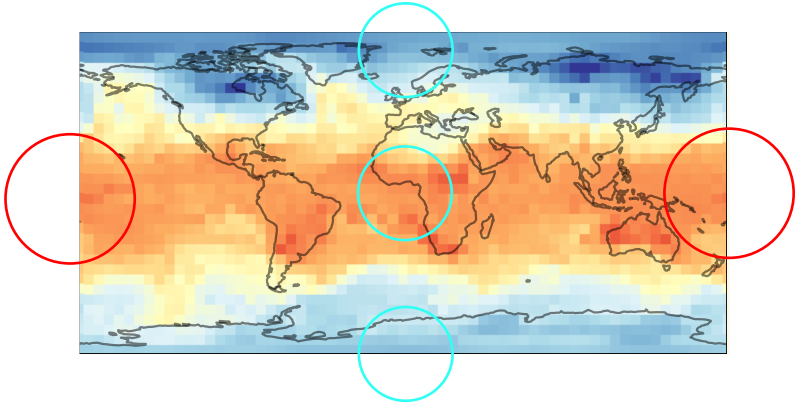
Introduction – Dataset

Results in this presentation:
- Weather Benchmark
- Region: Global
- Spatial resolution: 5.625° (approx. 600 km)
- Temporal resolution: 6h

Currently:
- ECMWF ER5



Geometric projection – Planar



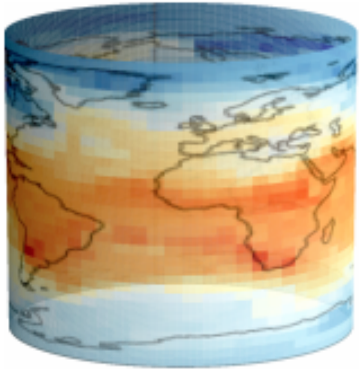
Project the sphere as planar

☹️:

1. Discontinuity on borders
2. Imbalanced projection

Figure: Rasp et al., 2020

Geometric projection – Cylinder



Project the sphere as cylinder

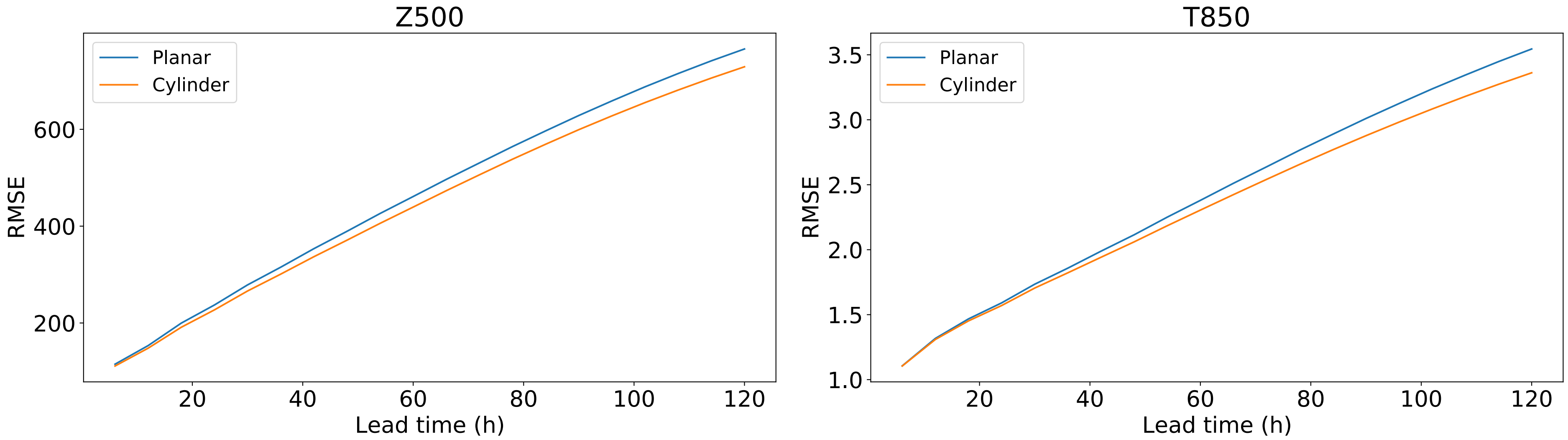
☹:

1. ~~Discontinuity on borders~~
2. Imbalanced projection

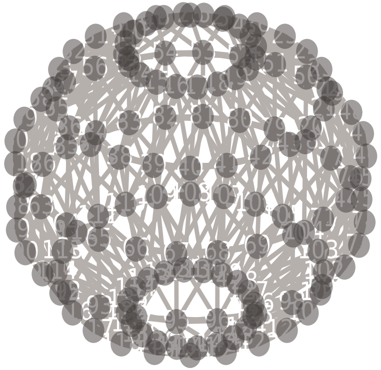
Figure: Rasp et al., 2020

Geometric projection – Cylinder

How much improvement?



How to solve imbalanced projection?



Spherical modeling

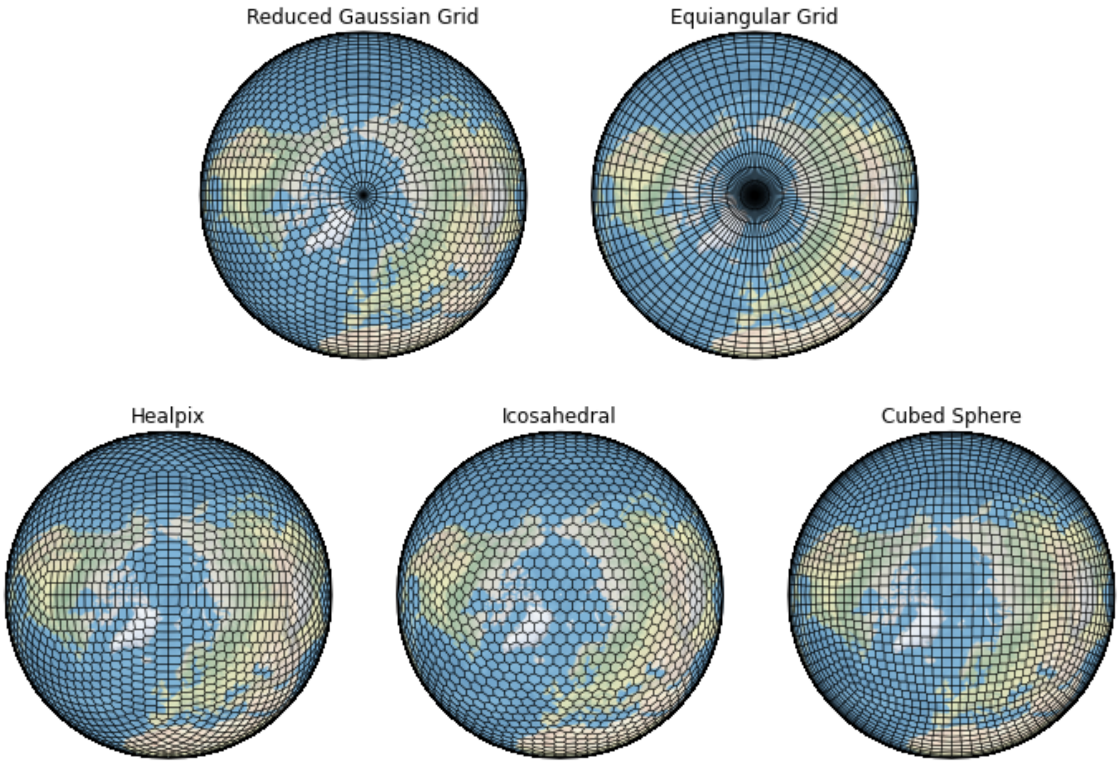
📓

- Directly modeling on spherical sampling without projection
- Using graph to represent the sphere
- Graph nodes given by sampling points

✅

- Can approximate any 3D object
- Approximation quality is proportional to number of nodes.
- Solving imbalance by using an even sampling (e.g., HEALPix)
- Adaptive to any levels of resolution

Spherical modeling – Sampling methods



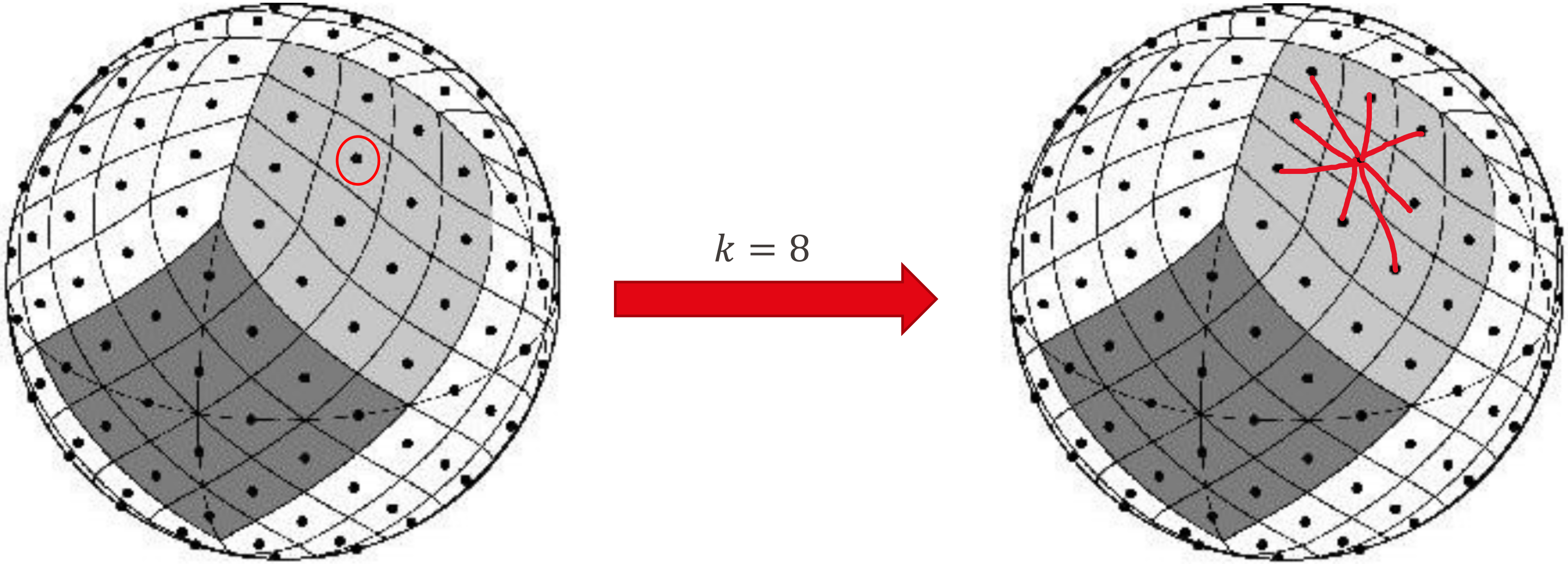
Wentao Feng / Yann Yasser Haddad

Gionata Ghiggi

Spherical modeling – Build the graph

k Nearest Neighbors (kNN) graph

- Each nodes is connected to k nearest neighbors undirectedly.
- If we have N nodes, then we have kN edges.



https://healpix.sourceforge.io/

Spherical modeling – Computation on the graph

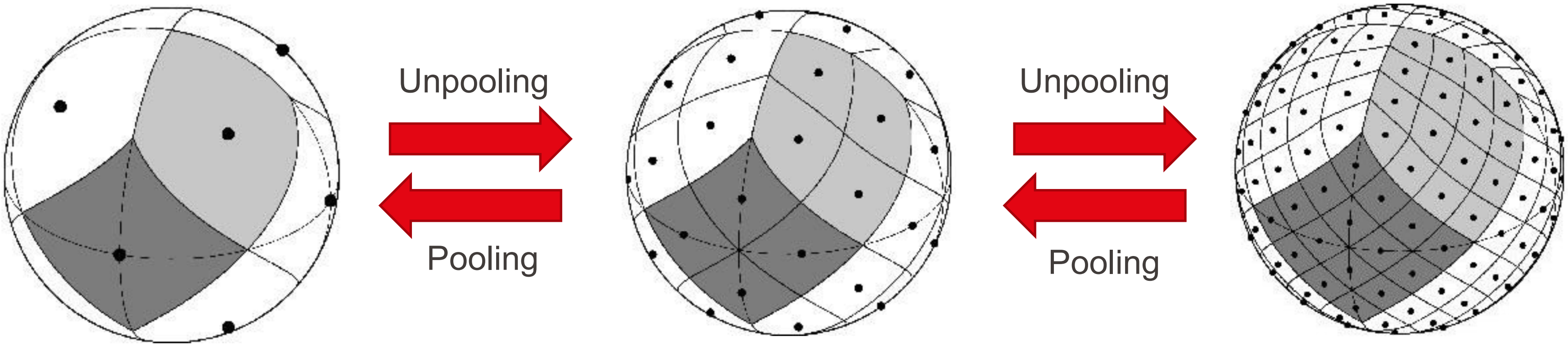
Graph convolution

1. Fourier transform to frequency domain
2. Multiply with (learnable) kernel function
3. Inverse Fourier transform

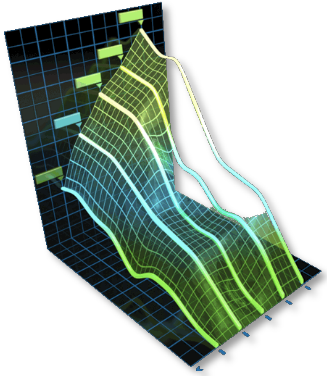
Graph pooling / unpooling

1. Max / Average aggregation
- Not applicable to all spherical sampling
- Easy to implement

2. General interpolation
- Applicable to all spherical sampling

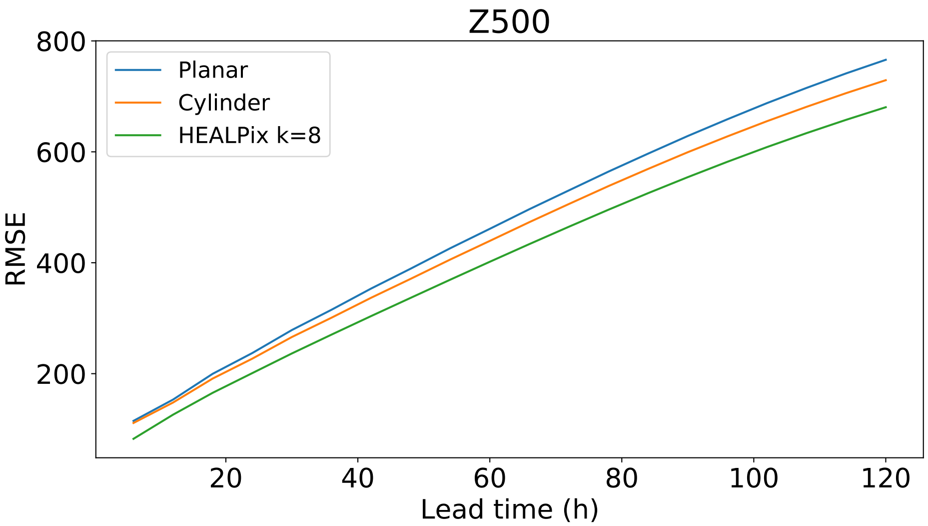


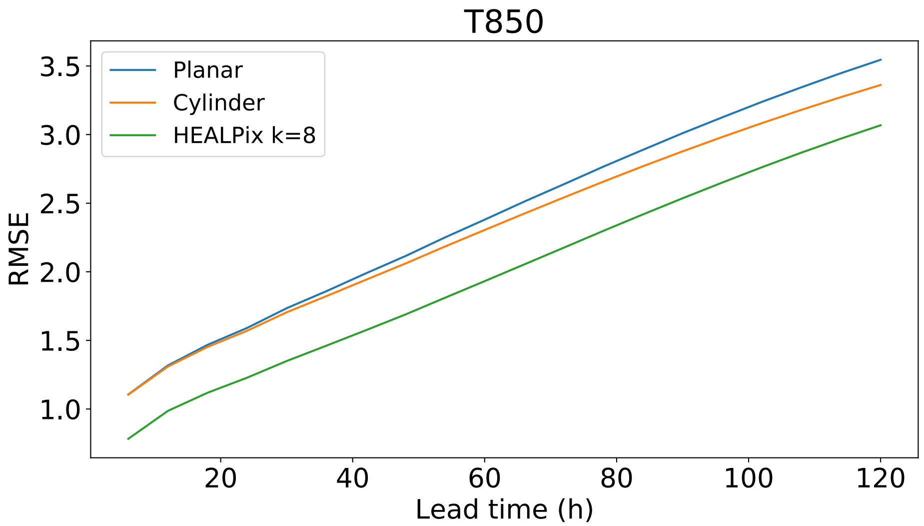
https://healpix.sourceforge.io/



Wentao Feng / Yann Yasser Haddad

Accuracy and efficiency





For fair comparison, the HEALPix data are interpolated from planar/cylinder data

Accuracy and efficiency - Scalability

Scalability matters:

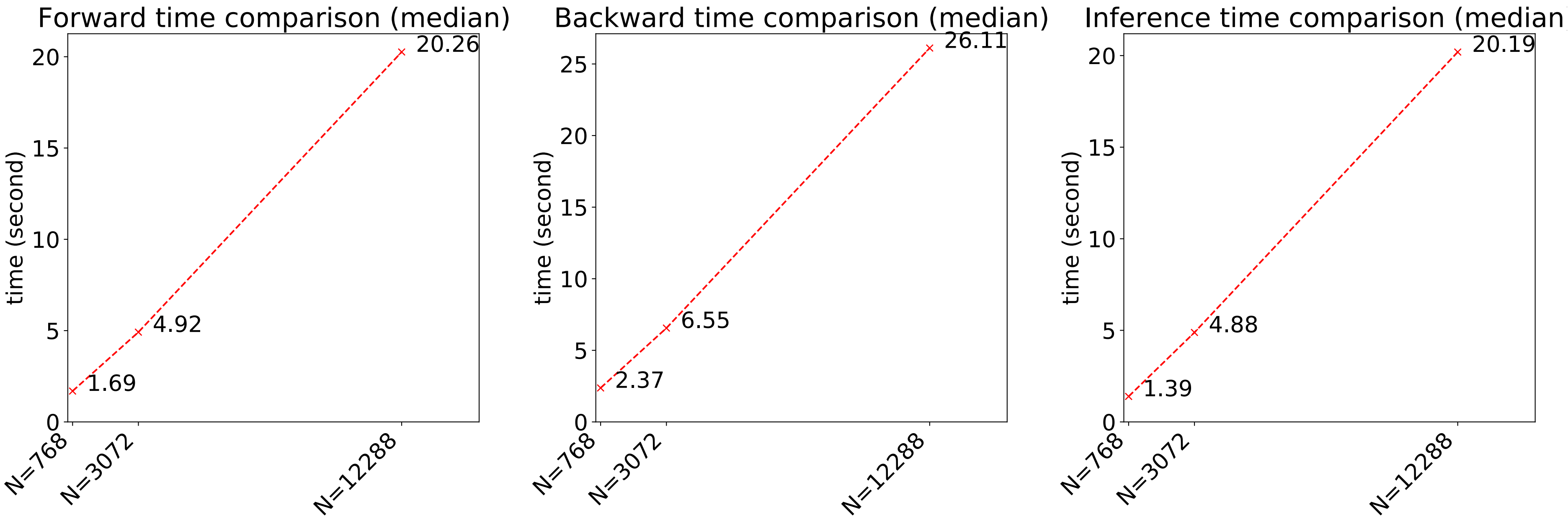
- In real life, spatial resolution can be less than 50 km (Large graph)
- The model need to handle real case efficiently (One of our motivation)

Accuracy and efficiency - Scalability on N

Wentao Feng / Yann Yasser Haddad

Scalability matters:

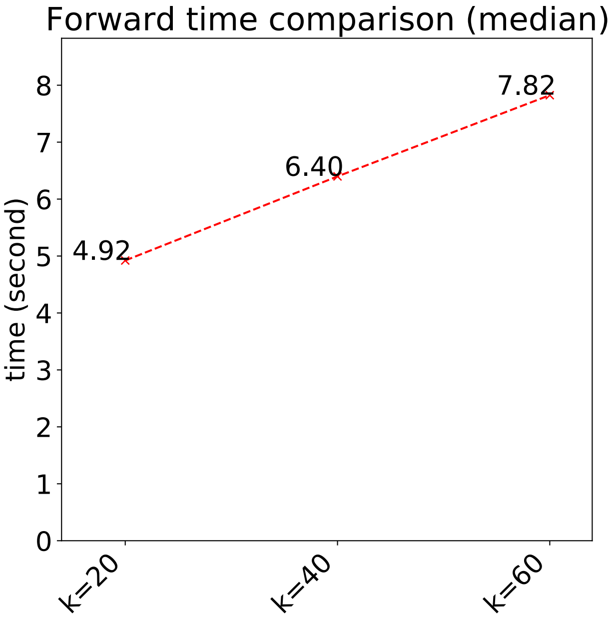
- In real life, spatial resolution can be less than 50 km (Large graph)
- The model need to handle real case efficiently (One of our motivation)

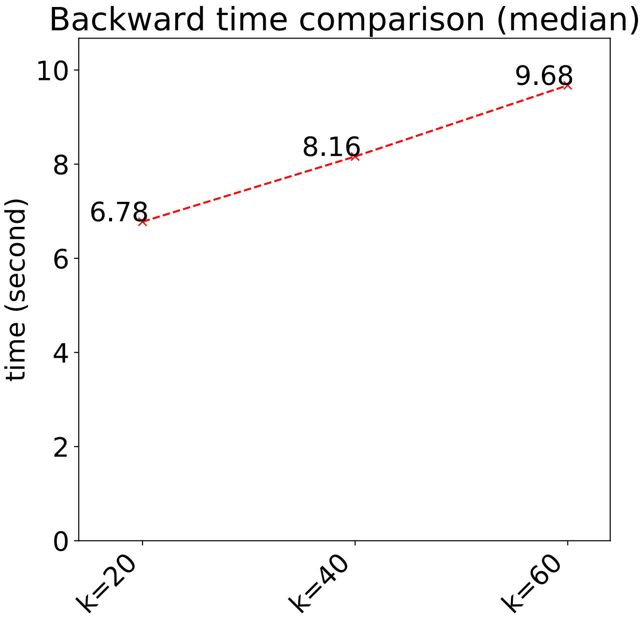


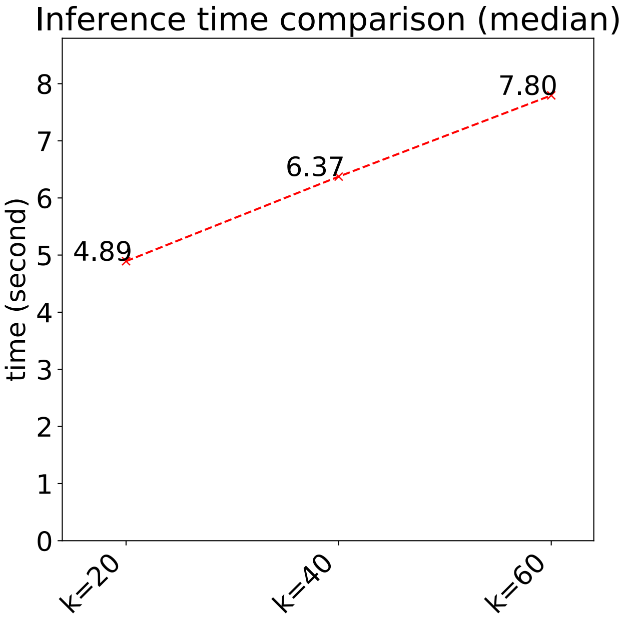
Accuracy and efficiency - Scalability on k

Wentao Feng / Yann Yasser Haddad

Scalability matters:
- In real life, spatial resolution can be less than 50 km (Large graph)
- The model need to handle real case efficiently (One of our motivation)







Accuracy and efficiency - Costs

What kind of infrastructure we need to run the following model?

- $N = 12288$ (approx. 200km)
- $k = 20$
- Process 15 graphs in parallel

?

or

- $N = 196608$ (approx. 50km)
- $k = 20$
- Process 1 graph

Answer: Single GTX 1080Ti with 8GB free graphical memory

😉

Multiple GPUs can

- Process larger graph
- Process larger batch
- Accelerate the learning

Conclusion until now

1. Spherical modeling better approximates the earth
   - No information lost on borders
   - Graph evenly represents balanced sampling
2. The approximation quality can be adjusted by the number of sampling points
   - More sampling points mean smaller polyhedrons to approximate sphere
3. The computation complexity increases linearly
   - The ability to handle high resolution data
4. Lower requirements for infrastructures than classical NWP.
   - Single GPU can run the experiments efficiently
   - Multiple GPUs are capable of high-resolution simulation.

Reference

1. Images sources:
   - https://developer.nvidia.com/
   - https://github.com/epfl-lts2/pygsp
   - https://pixabay.com/
2. Gionata Ghiggi, slides about DeepSphere
3. Michaël Defferrard, slides about DeepSphere and the Earth
4. Icíar Lloréns Jover, Geometric deep learning for medium-range weather prediction

Wentao Feng / Yann Yasser Haddad