IGS Summer Fellowship: Final Presentation

Lucía Vilallonga, Nils Cazemier, Dr. Yves Atchadé

Project (re-) introduction

• **Objective:** improve on numerical rainfall prediction methods using machine learning

- Particularly interested in modeling rainfall in Ghana
 - Tropical climate, where numerical methods tend to underperform
 - Availability of both satellite data and ground observation data from weather stations
 - Personal and professional connections

- Methods of interest:
 - Deep CNNs with physical constraints (de Bézenac et al, 2018)
 - Later, GNNs (Lam et al, 2022; Alet et al, 2019; Pfaff et al, 2021; and more)
 - Others: equation discovery??

$$\frac{\partial I}{\partial t} + (w.\nabla)I = D\nabla^2 I$$

$$I(x,t) = \int_{\mathbb{R}^2} k(x-w, y) I_0(y) dy$$



$$L_{t} = \sum_{x \in \Omega} \rho(\hat{I}_{t+1}(x) - I_{t+1}(x)) + \lambda_{\text{div}}(\nabla, w_{t}(x))^{2} + \lambda_{\text{magn}} \|w_{t}(x)\|^{2} + \lambda_{\text{grad}} \|\nabla w_{t}(x)\|^{2}$$

Method from <u>de Bézenac et al., 2019</u> (predicting sea surface temperature by estimating a motion or wind field)



Method from <u>de Bézenac et al., 2019</u> (predicting sea surface temperature by estimating a motion or wind field)

Original goals...

- 3 weeks determining the optimal initial condition;
- 2 weeks developing the numerical PDE solver;
- 1 week generating simulation data from the solver;
- 3 weeks on training and testing ML algorithms, starting from a CDNN and exploring other options; and
- 1 week on miscellaneous tasks, eg. drafting a manuscript write-up, the fellowship deliverable, and/or a presentation for the Institute.

Progress

- 3 weeks determining the optimal initial condition;
- 2 weeks developing the numerical PDE solver;
- 1 week generating simulation data from the solver;
- 3 weeks on training and testing ML algorithms, starting from a CDNN and exploring other options; and
- 1 week on miscellaneous tasks, eg. drafting a manuscript write-up, the fellowship deliverable, and/or a presentation for the Institute.

Progress

• 8 weeks on data preparation

• 2 weeks on implementing a CDNN model

• Some time on this presentation :)

Challenges and opportunities

- Shifting goalposts and approaches
 - Finding new methods that are a better fit for my competencies and domain knowledge

Learning about a new scientific data format

 NetCDF, HDF-5

- Reinventing the wheel during data prep stage
 - Efficient and fast open-source software exists for what I was trying to do by hand

CDO



US

Max-Planck-Institut für Meteorologie



Progress!

Next?

- Finish (de Bézenac et al, 2019) model implementation and begin training, testing, tuning
 - Experiment with optimizers and regularization terms
 - ** IGS video presentation deliverable **

• Learn about and explore Graphical Neural Networks (GNNs)

- Multiscale, heterogeneous (multiple data types) model for rainfall?
 - Incorporate both satellite image data and ground measurements





A peek ahead: GNN







Graph Element Networks (Alet et al, 2019)



A peek ahead: GNN



A peek ahead: GNN

Thank you!