Multimodel Superensembling Using Convolutional Neural Networks

Erik Larsson[1*], Andrea Littardi[2,3*], Mihalis Nicolaou[2], Leif Eriksson[1], and Waqas Qazi[3]

[1] Chalmers University of Technology, Sweden
{erik.larsson2, leif.eriksson}@chalmers.se
[2] The Cyprus Institute, Cyprus
{a.littardi, m.nicolaou}@cyi.ac.cy
[3] Offshore Monitoring Ltd., Cyprus
{andrea.littardi, waqas.qazi}@offshoremonitoring.com

Abstract. Motivated by the need for more accurate and robust weather forecasts, multimodel superensembling has become a widely researched topic. We propose an approach for multimodel superensembling using convolutional neural networks to perform merging of weather models based on local weather events. Our system is trained on medium range global wind forecasts and shows an average RMSE improvement of 7.3% compared to the ensemble mean.

Keywords: Multimodel Superensembling · NWP · Deep Learning · Convolutional Neural Networks

1 Background

Atmospheric, wave and ocean forecasts are typically produced in Numerical Weather Prediction where physical processes are simulated in mathematical models. Due to the complexity of these systems, some approximations are inevitable and can be present both in the representation of physical systems and in the initial states. Because of these approximations, some measure of errors appear and grow exponentially over time [1]. Here we are concerned mainly with errors that are systematic and inherent to specific weather models and how to mitigate them. One possible approach to this problem is by combining information from multiple different weather models in a process called multimodel superensembling.

Multimodel superensembling was first proposed in [2], where linear regression models were trained for multimodel superensembling for medium range weather, seasonal climate and hurricane tracks forecasts. Several studies have since proposed other methodologies to address the problem, focusing on single parameters or regions [3, 4].

*co-first authorship

An important aspect to consider when combining forecasts from different models, is the possibility of missing the location of extreme realizations of the variable under study. This can be devastating when e.g. combining the predictions of severe weather conditions, where misalignment of the storms may cause the merged prediction to average out the extreme values. In [5], a locally weighted learning algorithm is shown to mitigate this issue.

We propose an approach where Convolutional Neural Networks are trained on archived wind forecasts to dynamically combine the models based on current weather events.

2 Method

We trained a deep learning model for multimodel superensembling on forecasts for global surface wind fields. We observed that for all of the input models to the system, there was no clear correlation between forecast skill and location in space. The forecast skill appeared to be more associated with local weather conditions and forecast lead time. With this observation, convolutional neural networks are very fitting for the problem at hand as they are able to learn patterns in local areas of the data, regardless of spatial location, while keeping the number of parameters low.

As input to our system we use the U and V components of the global surface wind fields from three different major forecast providers[1]. The data was provided by the TIGGE research project by ECMWF [6]. The native resolution that the forecasts are produced on is different between all three of the models, but as provided the data has been resampled to the same spatial resolution. This difference in native resolution may have an influence on individual model performance. We train our system using ECMWF ERA5 [7] reanalysis as target data, with MSE as loss function.

We use forecasts for the month of February from years 2008 to 2018[2] with four forecasts per day for seven days. We used years 2013 and 2018 for validation and remaining years for training our models. To enable learning of time dependent correlations between input models, we split the seven days of forecasts into seven subsets and train seven networks, one per forecast day.

The architecture used is a purely convolutional neural network as visualized in Figure 1. We keep the spatial resolution fixed throughout the network. To allow information spreading in local neighbourhoods around all locations in space we use a sequence of dilated convolutions. To enable pattern matching on multiple scales we use skip connections such that the outputs of each of the dilated convolutions are used. We use Exponential Linear Unit (ELU) as activation function throughout the network to allow negative values. As the data is cyclic in the east-west direction but not in the north-south direction, we use circular padding in the east and west and zero padding in the north and south directions to enforce this periodicity.

[1] ECMWF ENS, UK Met Office MOGREPS-G and NCEP GEFS

[2] except 2015 due to an error in the service

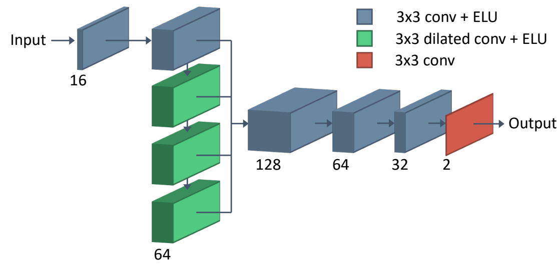

Fig. 1: A diagram of the network architecture. All convolutional kernels are of size 3x3. The dilated convolutions use dilations 2, 3 and 5 respectively. Stride is set to 1 for all layers. The amount of padding is set such that the spatial resolution is kept fixed.

3 Preliminary results

Our preliminary results show we reduce the RMSE when comparing with the ensemble mean by 7.3%. It is worth noting that both the mean and our model perform better than the best performing model alone, even though our target variable comes from the same provider, possibly introducing some bias. While additional benchmarking against other approaches is still required, these results give positive early evidence. We have studied how the performance varies across time and space.

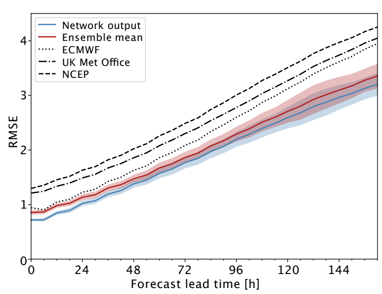

Fig. 2: RMSE score of the proposed model, the ensemble mean and the individual ensemble members. The shaded areas show the standard deviation of the RMSE for corresponding models.

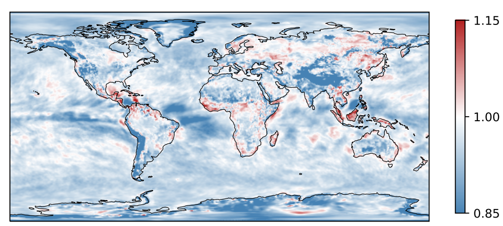

Fig. 3: Time-averaged relative RMSE compared to ensemble mean for all spatial locations. Locations with values < 1 mean our model performs better than the ensemble mean on average.

For short lead times, the ensemble mean does not offer a major improvement over the best performing ensemble member, while our model shows a relatively big improvement, as can be seen in Figure 2. For long lead times, both the ensemble mean and our model show a large improvement over the best performing model. A possible explanation could be that larger and diverging errors are being smoothed out. However, the relative improvement of our model compared to the ensemble mean is smaller in the case for long lead times. Figure 3 shows the average error relative to the ensemble mean for all locations. Our model performs better on open sea and for some mountainous regions, but is underperforming in some tropical regions. This may be due to highly varying weather conditions in these regions for the time period included in the training data, disabling the network to learn useful patterns.

4 Conclusions

A limitation of the proposed model is the focus on a 2D representation of a spherical object. The same kernel size in a convolutional layer covers different geographical regions whether applied at the equator or closer to the poles. Recent developments in geometrical deep learning [8, 9] allow for dealing with such cases, and could represent further useful additions to the model. Additionally, results are affected by the error metric selected. Utilizing a learnable loss function with Generative Adversarial Networks may be an interesting to explore development.

Acknowledgements This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 820593. The computations were performed on resources at Chalmers Centre for Computational Science and Engineering (C3SE) provided by the Swedish National Infrastructure for Computing (SNIC).

References

1. Buizza, R. "Sources of uncertainty". ECMWF (2018). https://www.ecmwf.int/en/research/projects/tigge.
2. Krishnamurti, T. N., et al. "Improved weather and seasonal climate forecasts from multimodel superensemble." Science 285.5433 (1999): 1548-1550.
3. Lenartz, et al. Super-ensemble techniques applied to wave forecast: performance and limitations. Ocean Science Discussions (OSD). 7. 10.5194/osd-7-709-2010.
4. Korner, S.. (2017). "Multi-model ensemble wake vortex prediction." DLR Deutsches Zentrum fur Luft- und Raumfahrt e.V. - Forschungsberichte. 2017. 1-155.
5. Pan, S.Q., Fan, Y.M., Chen, J.M., Kao, C.C., 2016a. Optimization of multi-model ensemble forecasting of typhoon waves. Water Sci. Eng. 9(1),52e57.http://dx.doi.org/10.1016/j.wse.2016.02.001.
6. TIGGE, https://www.ecmwf.int/en/research/projects/tigge. Last accessed 20 May 2020
7. ERA5, https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5. Last accessed 28 May 2020.
8. Chiyu Jiang, et al. Spherical CNNs on unstructured grids. In International Conference on Learning Representations (ICLR), 2019.
9. Cohen, T.S., et al: Spherical CNNs In: International Conference on Learning Representations. (2018)