Statistical Postprocessing of Local Numerical Weather Prediction Model Forecasts using Deep Learning

Master's Thesis Defense

Y. Harun KIVRIL

Dept. of Industrial Engineering
Bogazici University
Assist. Prof. Mustafa Gökçe Baydoğan (Thesis Supervisor)

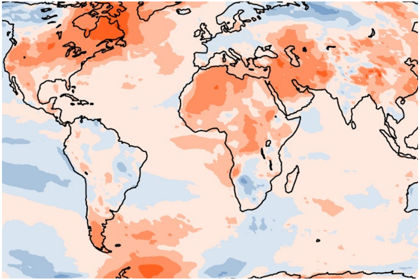
August 31, 2022

Outline

Outline

Weather Forecasts

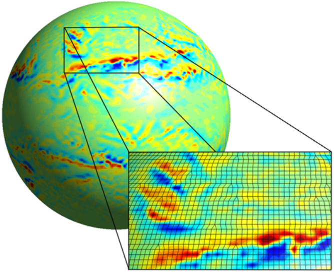
- Core element for many decision making processes (i.e. energy planning, flight scheduling, agriculture)
- Accuracy of the forecast is important to have better decisions
- Currently the forecasts are mainly generated by Numerical Weather Prediction Models



Numerical Weather Predictions (NWPs)

- Physical models developed by meteorological institues
- European Centre for Medium-Range Weather Forecasts (ECMWF), US National Oceanic and Atmospheric Administration (NOAA) etc.
- Starting from an initial condition solves the equations of atmospheric motions
- Considers the earth surface as a grid of coordinates
- Forecasts are made for multiple weather variables at multiple vertical levels for each grid point

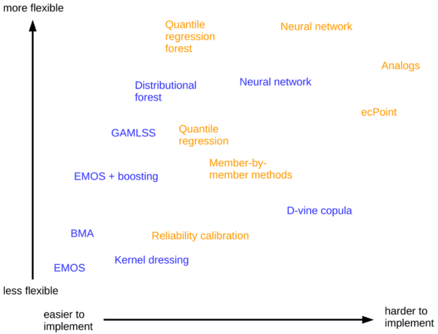
ie. Temperature of latitude 35 and longitude 36 at 850 mb for 2022.09.02 13:00



Problems of NWPs

- The models are sensitive to initial conditions, boundary conditions
- Also, the model structural errors lower the forecast accuracy
- Chaotic nature of the atmosphere makes forecasts useless after a level
- These systematical errors are needed to be corrected for better forecast performance
- Statistical postprocessing methods provides a solution to correct the systematical errors
- In this thesis, alternative neural network architectures are proposed to statistically postprocess an NWP model forecasts

Outline



Models with distributional assumptions		Models without distributional assumptions	
BMA	Raftery et al. (2005)	Analogs	Delle Monache et al. (2013)
D-vine copula	Möller et al. (2018)	ecPoint	Hewson and Pillosu (2020)
Distributional forest	Schlosser et al. (2019)	Member-by-member	Van Schaeybroeck and Vannitsem (2015)
EMOS	Gneiting et al. (2005)	Neural network	Bremnes (2020)
EMOS + boosting	Messner et al. (2016)	Quantile regression	Bremnes (2004)
GAMLSS	Lang et al. (2019)	Quantile regression forest	Taillardat et al. (2016)
Kernel dressing	Roulston and Smith (2003)	Reliability calibration	Flowerdew (2014)
Neural network	Rasp and Lerch (2018)		

Figure: Postprocessing methods based on flexibility and implementation difficulty [1].

Models with Distributional Assumptions

- An assumption on forecast distribution is made
- Learning is made to estimate the parameters of the assumed distribution

Models without Distributional Assumptions

- No assumptions on forecast distribution
- Uncertainty is reported based on estimated quantiles

Literature Review

Pros and Cons of ANNs

- Most methods focus on single variable at single level. However, ANNs flexible structure allows model multi output structures and they work well with multidimensional tensors
- ANNs are able to approximate almost any function using sequence of nonlinearities
- ANNs allow concepts like parameter sharing and spatial modeling via convolution operations
- ANNs models are complicated to train, random initialization of weights and using SGD creates reproducibilty issues
- Tuning model complexity and training parameters requires intense computational power
- ANNs require large amount of data to work well

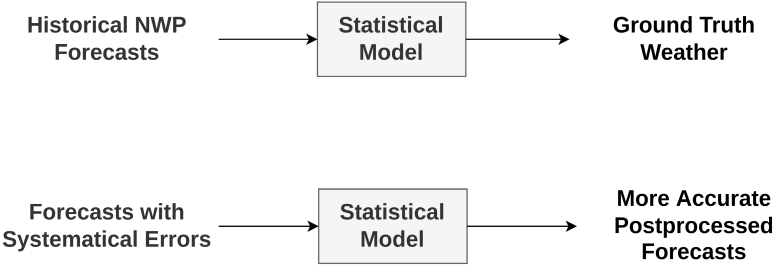
Literature Review

Contribution to Literature

- Postprocessing of Aegean Region NWP model forecasts
- Unlike many studies, multiple weather variables and pressure levels are used
- Models with extrapolation capability
- Wind power forecasting task to measure the effect on a real world application.

Outline

Methodology



Data Sources

- **Global Ensemble Forecasting System (GEFS)** forecasts are selected as input
  - GEFS Forecast are three hourly, in 0.25x0.25 latitude and longitude resolution with 6 weather variables and 25 pressure levels between 1 mb to 1000 mb
  - GEFS has $6x25x720x1440$ ($\approx 155M$) features for every third hour for the whole globe.
  - GEFS is open source and currently operationally available. It has reforecasts from 2000 to 2020
  - Control forecasts of the ensembles are selected for postprocessing
- **ECMWF Reanalysis 5th Generation (ERA5) Reanalysis** used as ground truth
  - The values in ERA5 is the reanalysis of forecasts after collection actual measurements
  - ERA5 is hourly, in 0.25x0.25 latitude and longitude resolution with 16 weather variables and 39 pressure levels between 1 mb to 1000 mb
  - ERA5 has $16x39x720x1440$ ($\approx 647M$) features for every hour for the whole globe.

Data Preprocessing - Variable, Level, Region Selection

Table: Selected Variables, Levels and Region

Dimension	ERA5	GEFS
Variable	*tmp*, *u*, *v*, *w*, *r*	*tmp*, *u*, *v*, *w*, *q*
Level	1000, 975, 950, 925, 900, 875, 850, 825, 800 mb	1000, 975, 950, 925, 900, 850, 800 mb
Latitude	[36.5, 40.5]	[36.5, 40.5]
Longitude	[25, 29.5]	[25, 29.5]

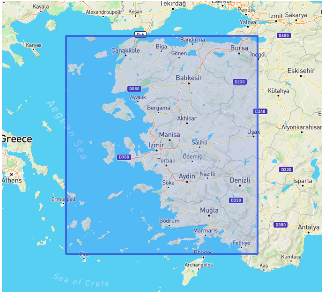


Figure: Selected Local Region

Data Preprocessing - Scaling & Tensor Shaping

Min-Max Scaling

$$\hat{\alpha}_{tijkl} = \frac{\alpha_{tijkl} - min(\alpha_{ijkl})}{max(\alpha_{ijkl}) - min(\alpha_{ijkl})}$$

GEFS (3 hourly time x variable x pressure level x latitude x longitude)
ERA5 (3 hourly time x [t, t+1, t+2] x variable x pressure level x latitude x longitude)

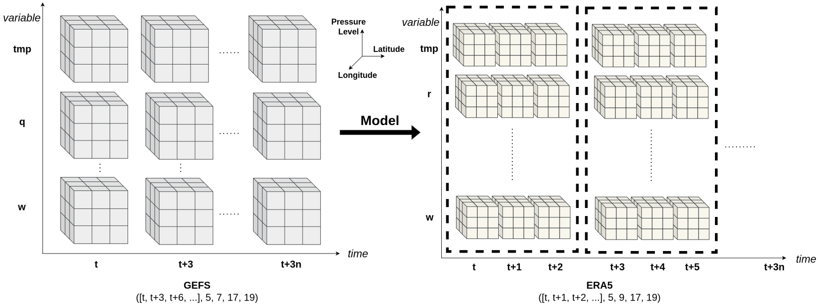


Figure: Illustration of Input and Output Schema

Proposed Architectures - Multi Layer Perceptron

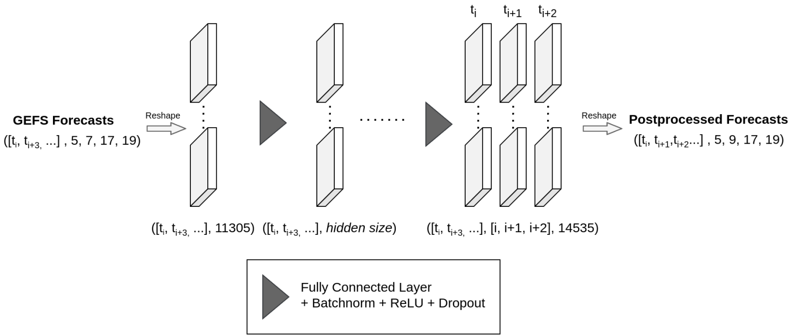


Figure: Illustration of Multi Layer Perceptron (MLP) Architecture

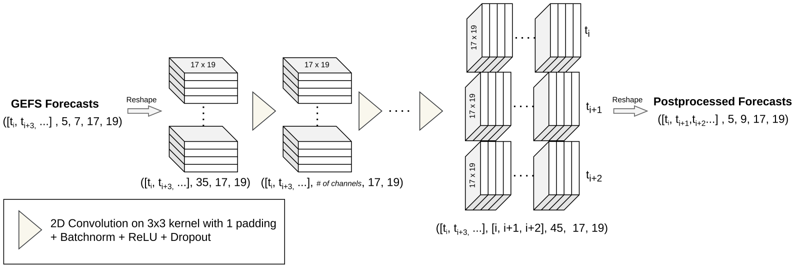


Figure: Illustration of Fully Convolutional Architecture

Proposed Architectures - U-Shaped Architecture

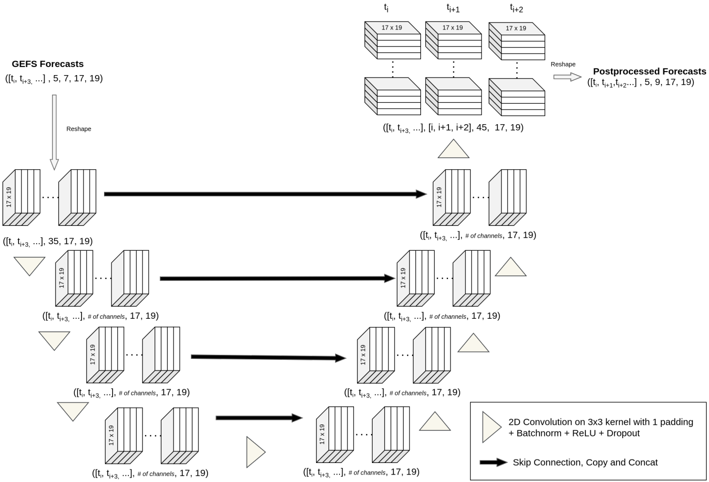


Figure: Illustration of U-Shaped Architecture

Proposed Architectures - Summary

Table: Summary of Proposed Architectures.

	Block Type	**Parameters**
MLP	Fully Connected	Dropout Ratio Batch Normalization Usage # of Blocks Hidden Size
Fully Conv. ANN	Convolution	Dropout Ratio Batch Normalization Usage Pooling Type # of Blocks # of Output Channels
U-Shaped ANN	Convolution	Dropout Ratio Batch Normalization Usage Pooling Type # of Output Channels

Outline

Modeling Pipeline

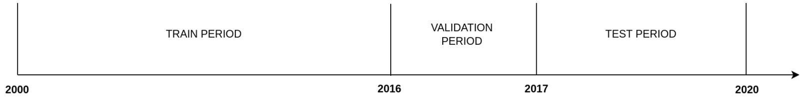


Figure: Illustration of Train, Validation and Test Periods

Table: Training Parameters

Training Parameters	Value
Batch Size	Hyperparameter
Learning Rate	Hyperparameter
Weight Decay	Hyperparameter
Max Epoch	30
Early Stop Patience	5 epochs
Early Stop Delta	0.0001

- Python 3.9.1
- Pytorch, Pytorch Lightning
- numpy, pandas, matplotlib, seaborn, netcdf4, wandb, optuna
- R GLMnet
- NVIDIA GeForce GTX 1070 Ti

Hyperparameter Tuning

48 hours of randomized search for each architecture

Table: Hyperparameter Search Space

Parameter Source	Hyperparameter	Candidate Values
Training	Batch Size	8, 16, 32, 64, 128
	Learning Rate	0.001, 0.01, 0.1
	Weight Decay	0.0001, 0.001, 0.01
MLP	Dropout Ratio	0, 0.1, 0.2
	Batch Normalization Usage	True, False
	# of Blocks	2, 3, 4, 5, 6, 7, 8
	Hidden Size	32, 64, 128, 256, 512, 1024
Fully Conv. ANN	Dropout Ratio	0, 0.1, 0.2
	Batch Normalization Usage	True, False
	Pooling Type	Average, Max
	# of Blocks	2, 3, 4, 5, 6, 7, 8
	# of Output Channels	32, 64, 128, 256, 512, 1024
U-Shaped ANN	Dropout Ration	0, 0.1, 0.2
	Batch Normalization Usage	True, False
	Pooling Type	Average, Max
	# of Output Channels	32, 64, 128, 256, 512, 1024

Wind Power Forecasting

- 19 wind farms are selected from the region
- 2017, 2018 forecasts as training, 2019 forecasts as test data
- Wind speed (ws) and ws^2 and ws^3 is used as predictors since the relation between wind speed and power generation is cubic
- GLMnet LASSO Model
- Penalization parameter tuned with 5 fold cross validation in train period
- WMAPE values reported for each farm and each weather source

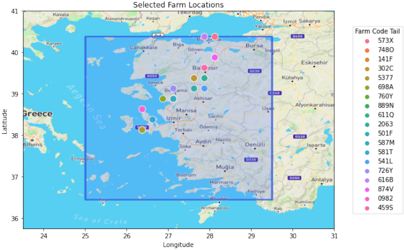


Figure: Locations of the selected farms

$$WMAPE = \frac{\sum_{i=1}^{N} |y_i - \hat{y}_i|}{\sum_{i=1}^{N} |y_i|}$$

Results - Architecture Comparison

Table: MSE Values of Variables for Three Hourly Weather (Postprocessed)

Model Name	tmp	u	v	w
GEFS	1.2175	2.6998	2.9829	0.0698
MLP	1.3164	2.3715	2.7288	0.0302
Fully Convolutional	0.7415	1.8493	1.9789	0.0421
U-Shaped	0.8070	1.8789	1.9112	0.0315

Table: MSE Values of Variables for the Remaining Two Hours (Extraploated)

Model Name	tmp	u	v	w
MLP	1.3240	2.5305	2.9515	0.0317
Fully Convolutional	0.8104	2.0983	2.2935	0.0434
U-Shaped	0.8687	2.0383	2.1870	0.0354

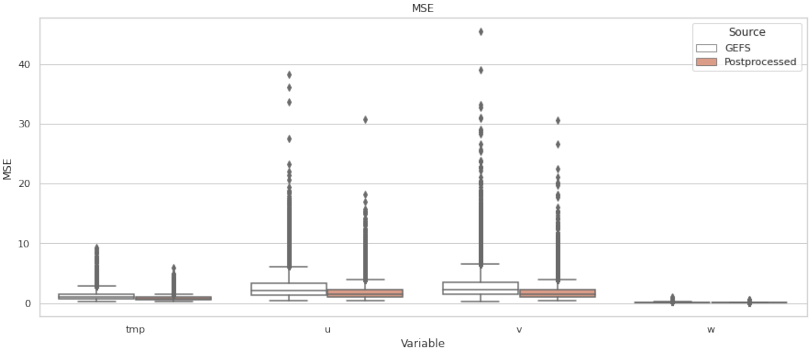
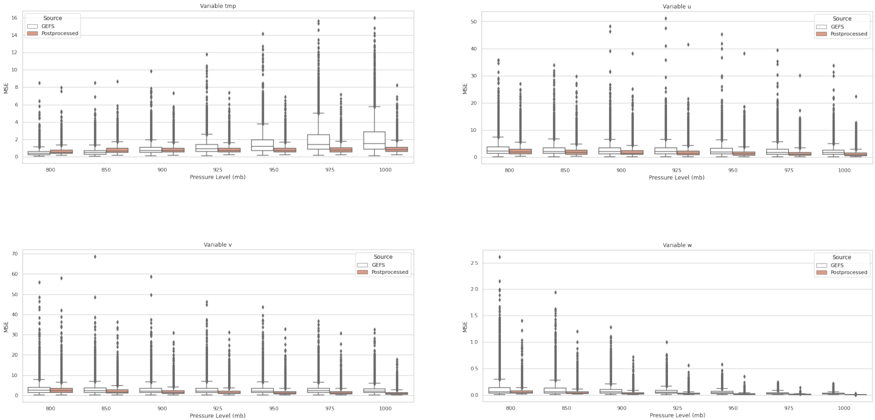


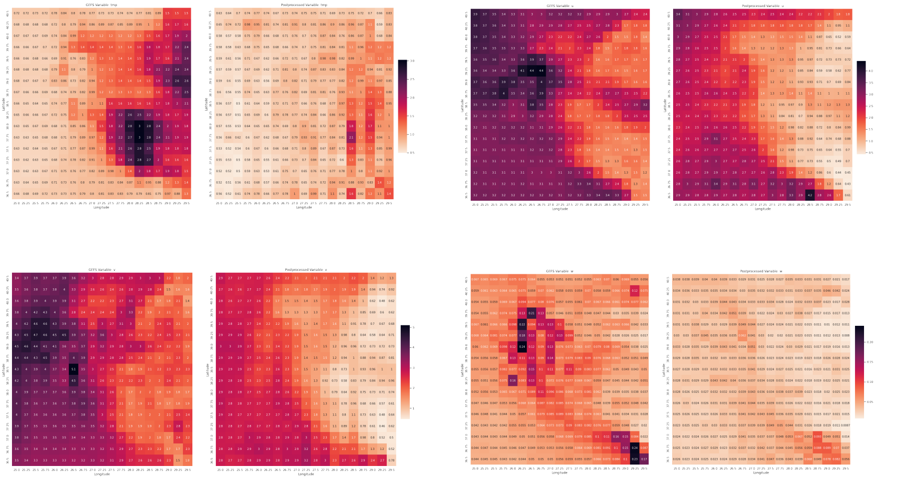
Figure: MSE Performance of U-Shaped Model for Each Variable

Results - Variable-Pressure Level Based



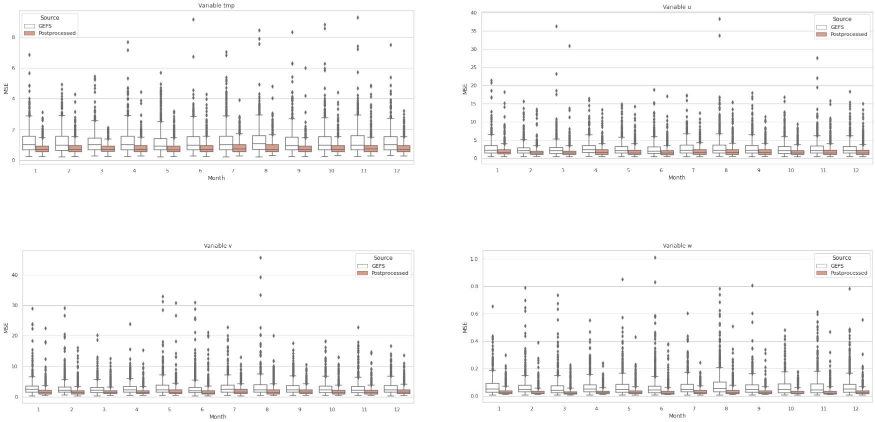
MSE Performance of U-Shaped Model for Each Variable at Each Pressure Level

Results - Variable-Location Based



MSE Performance of U-Shaped Model for Each Variable at Each Grid Point

Results - Variable-Month Based



MSE Performance of U-Shaped Model for Each Variable at Each Month of the Year

Results - Extrapolation Capability

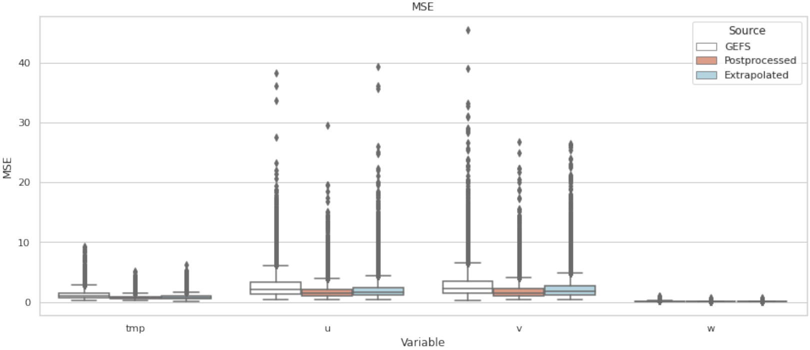


Figure: Distribution Comparison of GEFS, Postprocessed Values and Extrapolated Values

Results - Wind Power Forecasting

Table: WMAPE Values for Different Weather Sources

Farm Code	GEFS	ERA5	U-Shaped	Fully Conv.	MLP
40W00000000573X	0.3425	0.3080	**0.3307**	0.3496	0.3551
40W00000000587M	0.3466	0.2992	**0.2862**	0.3014	0.2891
40W00000000726Y	0.3911	0.3713	0.3882	**0.3747**	0.3945
40W00000000748O	0.3942	0.3684	0.3852	**0.3795**	0.3799
40W00000000760Y	0.2997	0.2569	**0.2658**	0.2705	0.2804
40W00000001581T	0.3561	0.3316	**0.3458**	0.3485	0.3460
40W00000002141F	0.3685	0.3123	**0.2982**	0.3159	0.3073
40W00000003302C	0.5182	0.4950	**0.5021**	0.5047	0.5061
40W000000042063	0.3064	0.2779	0.2899	**0.2880**	0.3048
40W00000004889N	0.4336	0.3687	**0.3525**	0.3690	0.3721
40W00000005541L	0.2976	0.2796	0.2849	**0.2755**	0.2851
40W00000005611Q	0.3524	0.2972	0.3145	**0.3135**	0.3299
40W00000005874V	0.3779	0.3446	**0.3597**	0.3611	0.3765
40W000000065377	0.3718	0.3403	0.3611	0.3544	**0.3536**
40W00000006616B	0.3468	0.3289	**0.3723**	0.3837	0.3922
40W000000070982	0.3816	0.3637	0.3864	0.3930	**0.3795**
40W00000008459S	0.3213	0.2939	**0.3091**	0.3116	0.3247
40W00000008698A	0.4021	0.3640	0.3659	**0.3608**	0.3707
40W00000010501F	0.3283	0.2941	**0.2919**	0.2991	0.2992

Outline

Conclusion

- Accurate weather forecasts are crucial for many decision making process
- NWP models are the main source of the forecast and they introduce systematical errors.
- Three alternative ANN architectures to statistically postprocess GEFS forecasts of Aegean Region for multiple weather variables at multiple pressure levels
- Fully convolutional architectures give promising results in all aspects
- U-Shaped Model has better error distributions over others
- The models have extrapolation capability due to input output structure and comparison of error distributions validates it
- Wind power forecasting case with 19 different farms demonstrated the effectiveness of the method on a real world application

Future Work

- Effect of variable level selection can be investigated further
- Modeling temporal relation between forecasts
- 3D convolutions that adds pressure levels as the third dimension to latitude and longitude
- Physics informed learning
- Ensemble members of GEFS
- Creating ensembles from the postprossesing model

Thanks for your attention!

References

1. Vannitsem, S., J. B. Bremnes, J. Demaeyer, G. R. Evans, J. Flowerdew, S. Hemri, S. Lerch, N. Roberts, S. Theis, A. Atencia et al., “Statistical postprocessing for weather forecasts: Review, challenges, and avenues in a big data world”, Bulletin of the American Meteorological Society, Vol. 102, No. 3, pp. E681–E699, 2021.

2. Rasp, S. and S. Lerch, “Neural networks for postprocessing ensemble weather fore- casts”, Monthly Weather Review , Vol. 146, No. 11, pp. 3885–3900, 2018.

3. Goodfellow, I., Y. Bengio and A. Courville, Deep learning, MIT press, 2016.

4. Ronneberger, O., P. Fischer and T. Brox, “U-net: Convolutional networks for biomedical image segmentation”, International Conference on Medical image com- puting and computer-assisted intervention, pp. 234–241, Springer, 2015.

5. Grönquist, P., C. Yao, T. Ben-Nun, N. Dryden, P. Dueben, S. Li and T. Hoefler, “Deep learning for post-processing ensemble weather forecasts”, Philosophical Transactions of the Royal Society A, Vol. 379, No. 2194, p. 20200092, 2021.

6. Pathak, J., S. Subramanian, P. Harrington, S. Raja, A. Chattopadhyay, M. Mar- dani, T. Kurth, D. Hall, Z. Li, K. Azizzadenesheli et al., “Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural opera- tors”, arXiv preprint arXiv:2202.11214 , 2022.

7. Schlosser, L., T. Hothorn, R. Stauffer and A. Zeileis, “Distributional regression forests for probabilistic precipitation forecasting in complex terrain”, The Annals of Applied Statistics, Vol. 13, No. 3, pp. 1564–1589, 2019.

8. Gneiting, T., A. E. Raftery, A. H. Westveld III and T. Goldman, “Calibrated prob- abilistic forecasting using ensemble model output statistics and minimum CRPS estimation”, Monthly Weather Review , Vol. 133, No. 5, pp. 1098–1118, 2005.

9. Messner, J. W., G. J. Mayr and A. Zeileis, “Non-homogeneous boosting for pre- dictor selection in ensemble post-processing”, Monthly Weather Review , Vol. 145, pp. 137–147, 2017.

Appendix - MLP Model Variable Based Results

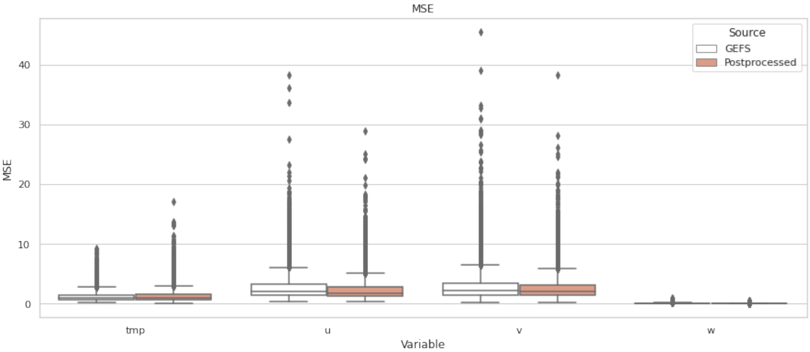
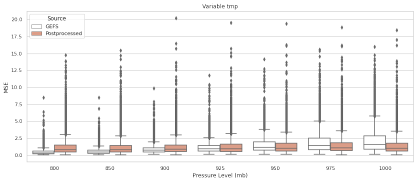
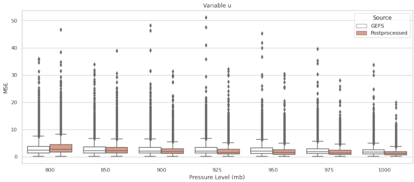
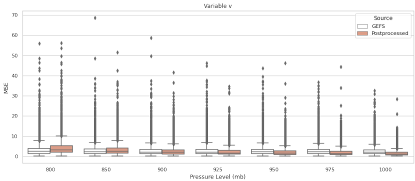


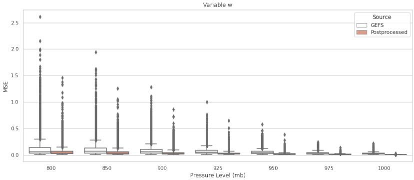
Figure: MSE Performance of MLP Model for Each Variable

Appendix - MLP Variable-Pressure Level Based Results



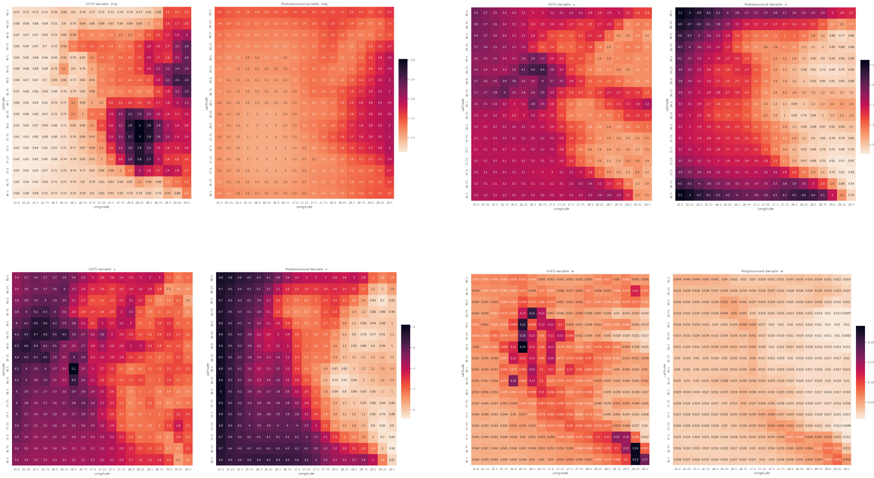






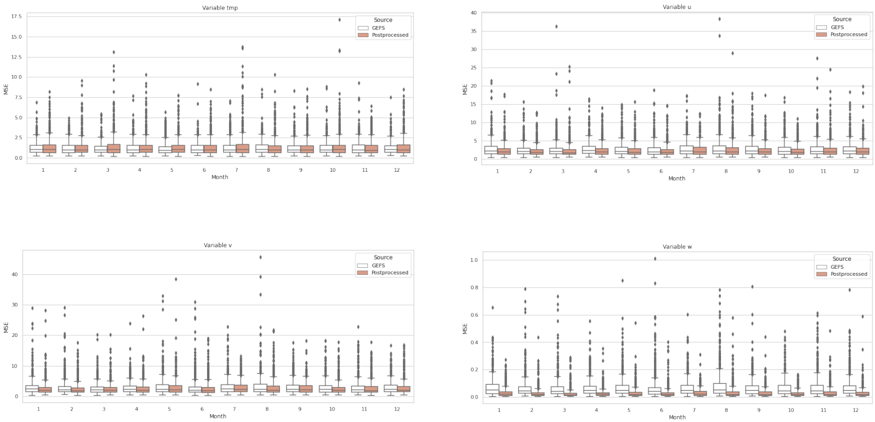
MSE Performance of MLP Model for Each Variable at Each Pressure Level

Appendix - MLP Variable-Location Based Results



MSE Performance of MLP Model for Each Variable at Each Grid Point

Appendix - MLP Variable-Month Based Results



MSE Performance of MLP Model for Each Variable at Each Month of the Year

Appendix - Fully Conv. Model Variable Based Results

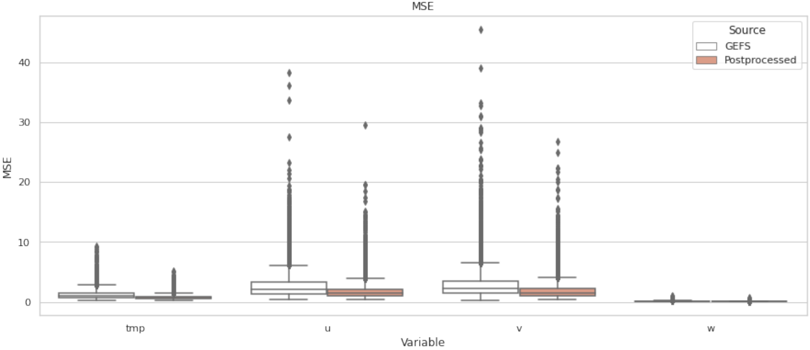
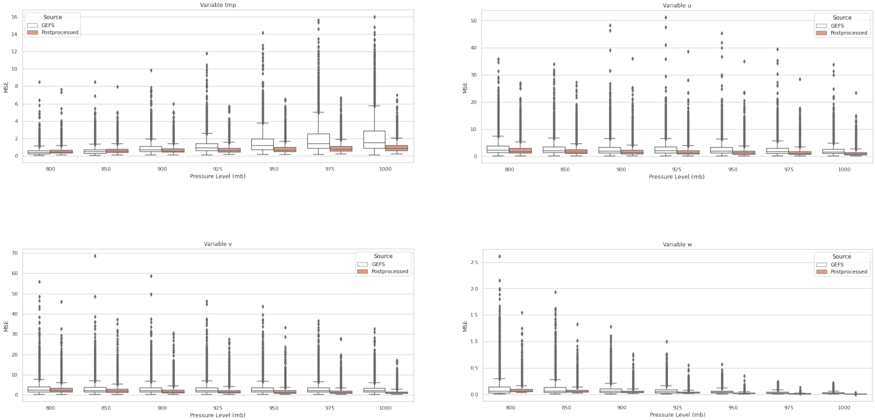


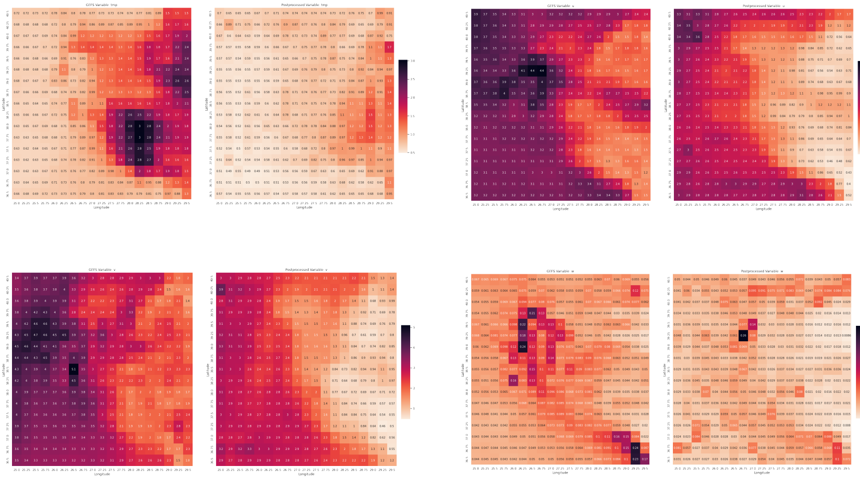
Figure: MSE Performance of Fully Conv. Model for Each Variable

Appendix - Fully Conv. Variable-Pressure Level Based Results

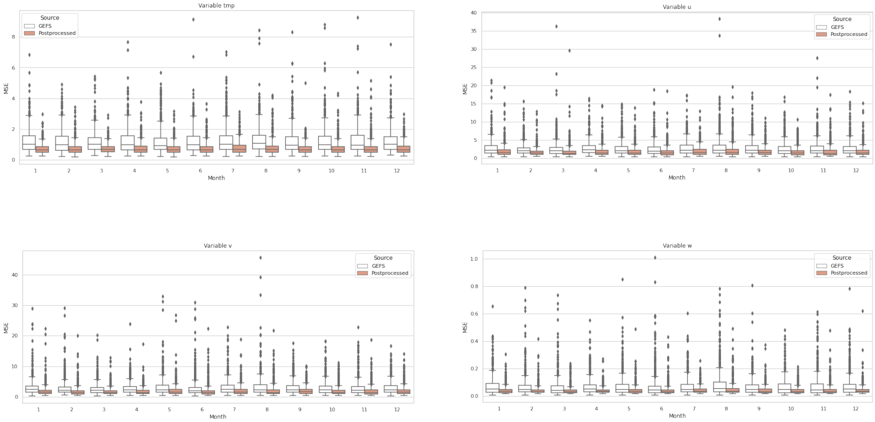


MSE Performance of Fully Conv. Model for Each Variable at Each Pressure Level

Appendix - Fully Conv. Variable-Location Based Results



MSE Performance of Fully Conv. Model for Each Variable at Each Grid Point



MSE Performance of Fully Conv. Model for Each Variable at Each Month of the Year

Appendix - Selected Hyperparameters

Table: MLP Best Hyperparameters

Hyperparameter	Value
Batch Size	32
Learning Rate	0.001
Weight Decay	0
# of Blocks	4
Hidden Size	1024
Dropout Ratio	0.1
Batchnorm Usage	True

Table: Fully Convolutional Architecture Best Hyperparameters

Hyperparameter	Value
Batch Size	8
Learning Rate	0.001
Weight Decay	0
# Number of Blocks	4
# of Channels	512
Dropout Ratio	0
Batchnorm Usage	True
Pooling Function	Max Pooling

Table: U-Shaped Convolutional Architecture Best Hyperparameters

Hyperparameter	Value
Batch Size	16
Learning Rate	0.001
Weight Decay	0
# of Channels	512
Dropout Ratio	0
Batchnorm Usage	True
Pooling Function	Average Pooling