University of North Carolina Asheville
Journal of Undergraduate Research
Asheville, North Carolina
May 2021

Comparing the Performances of Neural Network Architectures on Short-Range Weather Forecasts

Andy Hill
Atmospheric Sciences
The University of North Carolina Asheville
One University Heights
Asheville, NC 28804

Faculty Advisor: Dr. Christopher Hennon

Abstract

A suite of different artificial neural networks (ANN) is used to create 24-hour point forecasts for U.S. cities selected in the WxChallenge national forecast competition. Maximum and minimum temperature, maximum sustained wind speed, and total precipitation forecasts are taken across five weather forecasting models as input to the neural networks. The five models serving as input are: Global Forecast System (GFS) and North American Mesoscale (NAM) model output statistics (MOS), grid interpolated forecasts from the Hi-Res Rapid Refresh (HRRR) model, National Blend of Models (NBM), and National Weather Service point forecast matrix (PFM) forecasts. The ANNs are trained on over two years of historical model forecast data for every city. All ANNs—and weighted and unweighted ensembles of them—are compared against a multiple linear regression model and human forecasters to determine if ANNs provide an increase in skill of short-range forecasting of the given parameters. Performances are evaluated through 10 cities (80 forecasts) of the 2020–21 WxChallenge season.

1. Introduction

Within the field of meteorology, the usage of artificial neural networks in deep learning has become a common technique for improving numerical weather prediction (NWP) forecasts. Given the current array of NWP models and denominations of those models that produce forecasts for a myriad of atmospheric parameters, it is desirable to improve upon their skill due to known biases, inaccuracies and instances of poor initialization of current observations into the models leading to large errors. Neural networks fit that position ideally; certain types of artificial neural networks are able to generalize, handle trends, analyze and correct for specific biases in a model not apparent to human analysis—and apply all of the above to optimize a time-series forecast based on NWP model inputs. These optimized forecasts are particularly useful in many tiers of the meteorology sector, from public consumption to a reanalysis of model skill to further the optimization of underlying physics. Thus, it is important to seek areas where neural networks could most effectively build on NWP and highlight their applicability to consumers based on their statistical skill over time. One of the best methods for carrying out this idea is the WxChallenge, a national forecast competition which focuses on temperature, wind and precipitation—all of which directly impact daily life. This paper compares the skill—based on WxChallenge performance—of several architectures of neural networks in forecasting these valuable variables.

A neural network approach to optimizing a WxChallenge forecast must consider which types of neural networks are the best for this application, as well as what NWP models should be used as inputs to the networks. Ustaoglu et al. employed three types of neural networks of interest: the feedforward backpropagation (FFB), generalized regression (GRNN), and radial basis function (RBF) networks[1]. They found—for temperature forecasting only in two specific areas of Turkey—that the three neural network types all demonstrated considerable results over the control, which is multiple linear regression (MLR). MLR models, given a set of dynamical model outputs, improve upon those outputs using known biases found through comparison of historical data of the dynamical models' forecasts to what actually occurred. This process makes them inherently more skillful than dynamical models[2]; therefore, having a significant increase in skill over this control variable is a primary factor in the benefit of applying neural networks to temperature forecasts.

Precipitation and wind case studies using neural networks from the last decade also suggest their usefulness for forecasting for the WxChallenge. Specifically, Ghosh and Krishnamurti employ the GRNN to hurricane intensity (wind speed) forecasting due to its provenance of statistical skill[3]. Ghosh and Krishnamurti note the amount of previous work Krishnamurti produced highlighting the GRNN's performance in wind speed and precipitation forecasts in their reasoning for employing it in this study. In another study, Krasnopolsky highlights a simple nonlinear neural network in statistically significant precipitation forecasting[4]. Since precipitation is difficult to adequately forecast, a neural network's ability to generalize and correct for biases (too wet or too dry) is desirable.

Neural networks are entirely dependent on their input dataset, which, for this research, will be a set of dynamical (full physics) and dynamical-statistical models that capture all listed variables. Thorarinsdottir finds the postprocessing of dynamical model output through statistical methods to be skillful in wind speed forecasts in the Pacific Northwest[5]. This highlights the usefulness of model output statistics (MOS) products, making them good candidates for the neural networks. Building on this, Sanders focused on the application of machine learning—which has similar goals of quality and consistency—with the North American Mesoscale (NAM) model as input. With a range of many different variables, including those in the WxChallenge, Sanders found noticeably higher skill from a machine learning model based on just the NAM than even the NAM MOS[6].

This raises the question of introducing more inputs into a deep learning model, versus fewer. Weyn produced a machine learning tool named "UW-MOS X," which only consists of Global Forecast System (GFS) and NAM model forecasts as inputs[7]. He found that after many tests with several models, the best forecasting skill was deduced from only keeping these two models as inputs to the machine learning system. Given this result, deep learning systems may face the same situation for optimization; therefore, the key question the MOS-X raises surrounding the input to a neural network is that of the particular set of data that produces the best forecasts over time. The research presented here addresses this question by having created two separate iterations of each neural network type that have different amounts of model inputs (two: Hi-Res Rapid Refresh (HRRR) and the National Blend of Models (NBM); and five: GFS, HRRR, NAM, NBM, and the National Weather Service Point Matrix Forecast (PFM)). Whether or not the set of fewer input model networks outperform the set of full model networks will be specifically discussed, as it is an important dimension to forecast error reduction.

With this set of dynamical models as inputs, and the three types of networks covered (FFB, GRNN, RBF), this research will investigate which architectures of these neural networks will produce the best WxChallenge forecasts. The goal is to observe how best to reach optimized forecasts from each type of neural network by tweaking their input set as well as their algorithms. The analysis of the root-mean squared error (RMSE) and bias of each network-forecasted-variable and all of their daily errors in each variable across 20 weeks (10 cities) of the WxChallenge will determine their skill. MATLAB was used for building and running the neural networks, and the WxChallenge website for tracking verification and per-variable daily errors. The results will dictate which types, if any, are most suitable for short-range forecasts of valuable meteorological variables with the goal of minimal forecast error, and will direct further optimization efforts for short-range forecast precision.

2. Data and Methodology

2.1. Model Data Aggregation

Forecast data from the GFS and NAM MOS, HRRR, NBM and the National Weather Service PFM were gathered from historical archives (National Centers for Environmental Prediction (NCEP), Iowa State Mesonet, and the HRRR Archive at the University of Utah[8]) from January 2018–current week of the WxChallenge. The archived data were selected based on the WxChallenge city's identifier (i.e., KGRR for Grand Rapids, MI) or the closest gridpoint for the HRRR model. Then, verification for those days was gathered from the National Centers for Environmental Information. This process was repeated for each new city selected for the WxChallenge during the season. With this set of historical data, each neural network will train on about 4000–6000 data points for each variable with the full suite of models, or about 1000–1500 data points with only the NBM and HRRR. The historical dataset for the NBM extends back to November 2018, while the HRRR 36-hour archive only exists from July 2019 onward.

During the WxChallenge, current forecast data for the selected city from the suite of models were taken each day from the National Weather Service site for the PFM, MOS and NBM products. The current HRRR data were available on NCEP's NOAA Operational Model Archive and Distribution System. The 12Z NAM MOS, 18Z HRRR and GFS MOS, and 19Z NBM and PFM forecasts were used in running the neural networks each day in order to use the latest model data before the 00Z WxChallenge forecast deadline. The WxChallenge forecasts are from 06Z next day to the following 06Z, so each model's forecast for that time period was retrieved for the neural networks.

All variables are taken from each model except for accumulated precipitation. The precipitation forecasts are only taken from the NBM and HRRR models, as the MOS tables have coded ranges of precipitation and the PFM table forecasts accumulations that do not line up with the 06–06Z forecast period.

2.2 Neural Network Architectures

In MATLAB, two FFB neural networks for each variable ("ncaFFB" and "ncaFNH") that handle all variables were created. The suffix "-NH" stands for the networks that only use the NBM and HRRR. They were trained every two weeks on each new city's historical forecast set versus verification. All temperature-related FFB neural networks use the damped least-squares algorithm for training, while wind speed and precipitation FFB neural networks interchangeably use the damped least-squares and Bayesian linear regression algorithms. Since wind speed and precipitation do not have as strong a signal as the diurnal component of temperature, the different algorithms present another opportunity for optimizing the neural network outputs.

For the GRNN and RBF neural networks, MATLAB poses some limitations in functionality. The structure of MATLAB's generalized regression and radial basis function backends do not allow more than one input per neural network; therefore, we created 17 total instances of GRNN and RBF networks each, one per model per variable. The two iterations of each model ("ncaGRN," "ncaGNH," "ncaRBF," and "ncaRNH") use either all 17 instances, or eight of them for only the NBM and HRRR networks. The spread of the GRNN and RBF networks is a controllable parameter that reflects the euclidean distance between nodes in the network that is considered when creating output. The higher the spread, the more generalized an output (forecast) will be in terms of the average of the overall dataset. Spreads of 3.3 for temperature were used across both networks; GRNN used a spread of 0.85 for the wind while the RBF used a spread of 9.3, and for precip, GRNN used a spread of 0.13 while RBF used 1.65. Lower spreads were found to be the best for the GRNN as it would allow for a full range of forecasts instead of becoming too smoothed, while the RBF would output unrealistic values if a small enough spread prevented it from reaching the next node. This latter point is due to the usage of an exact radial basis function network, which is desirable for its potential ability to pinpoint forecasted variables based on previous iterations of model forecasts.

Along with each of the networks' individual forecasts, their RMSEs are also calculated based on a run of their performance off of the historical forecasts compared to verification. This process is done after training (feedforward networks) or simultaneously with training (generalized regression and radial basis function networks). For the GRNN and RBF, this allows for a weighted ensemble of the instances based on their RMSEs in order to get the actual "ncaGRN," "ncaGNH," "ncaRBF" and "ncaRNH" forecasts. For the overall consensus models, these RMSE values allow for a weighted superensemble ("ncaSEN") model consisting of the six neural networks. Table 1 shows an example of the computed RMSE values for Juneau, AK (PAJN), that went into the superensemble model, as well as a comparison to the mean RMSE values.

Table 1. Juneau, AK (PAJN) Calculated RMSE Values based on Performance on Historical Dataset Forecasts Versus Verification Used in the Weighted Superensemble

PAJN RMSEs	Tmax (°F)	Tmin (°F)	Wspd (kt)	Prcp (in.)
Average	3.09	3.27	5.43	0.16
ncaFFB	2.48	2.85	5.53	0.157
ncaFNH	3.00	3.31	5.08	0.158
ncaGNH	3.67	4.03	5.46	0.162
ncaGRN	3.32	3.47	5.45	0.162
ncaRBF	3.00	3.16	5.46	0.16
ncaRNH	3.28	3.70	5.44	0.160

The control, an MLR model using all five models as input, was also coded within MATLAB in order to compare against the neural network results. Additionally, an ensemble mean of the six neural networks ("ncaENS") was created to compare to individual neural networks. The results of the MLR and ensembles will be evaluated alongside the suite of neural networks through statistical interpretations of their WxChallenge scores.

3. Results

Overall, the suite of neural networks displayed significant results in terms of WxChallenge placement, against 1395 competitors in the 2020–21 season containing 10 different cities for 80 total forecasts. Table 2 gives their normalized cumulative scores and corresponding ranks per city and overall, where lower numbers for each are better. All neural networks had a negative cumulative score, which indicates that they all outperformed the human consensus over the season. They also outperformed each of their five input models' individual forecasts. The MLR control model was the overall winner, but its lead is narrow and likely due to some instabilities in the networks that will be discussed in each city. Comparisons to the uncoupled surface layer model (USL) 22Z run known for its accurate forecasting and the UWMOS model show that the suite of networks are delivering promising forecasts across a broad range of different climates in the U.S.

Table 2. WxChallenge Placements Per City, Overall Normalized Score and Rank out of 1395 Entrants; Best Placements Highlighted

	GRR	GFL	AJN	MAF	MLB	MSY	LAX	DSM	GSO	JAN	Cumulative Score (versus consensus)	Rank
ncaENS	38	335	121	247	205	115	325	15	138	52	-1.93	112
ncaFFB	**18**	340	163	162	269	154	265	23	**2**	40	-2.22	75
ncaFNH	34	371	355	359	348	**86**	363	89	307	39	-0.73	241
ncaGNH	22	361	300	309	278	251	391	30	262	116	-0.81	236
ncaGRN	95	307	27	272	267	211	363	26	169	108	-1.62	156
ncaMLR	86	281	74	262	72	147	176	38	42	**14**	**-2.45**	**48**
ncaRBF	59	400	**10**	366	205	223	242	30	173	127	-1.58	164
ncaRNH	146	451	292	221	389	257	365	20	280	118	-0.52	254
ncaSEN	55	321	54	272	226	183	302	**6**	178	72	-1.86	126
USL	113	131	374	**21**	314	145	**85**	311	51	228	-1.57	166
UWMOS	168	**128**	397	297	**3**	225	260	235	297	308	-0.54	251

Table 2 shows clear skill in deep learning in forecasting being able to reduce short-range forecast error for valuable variables. For seven out of 10 cities, at least one model placed in the top 100 forecasters after eight forecasts. This is significant, as across a period of two weeks, highly variable weather conditions are probable and will skew model data relying on climatological normals; thus, the neural networks are adapting and improving upon known biases found in the input models. Figures 1–4 will be dissected according to intriguing case studies that happened within the 80 forecasts in the following results sections.

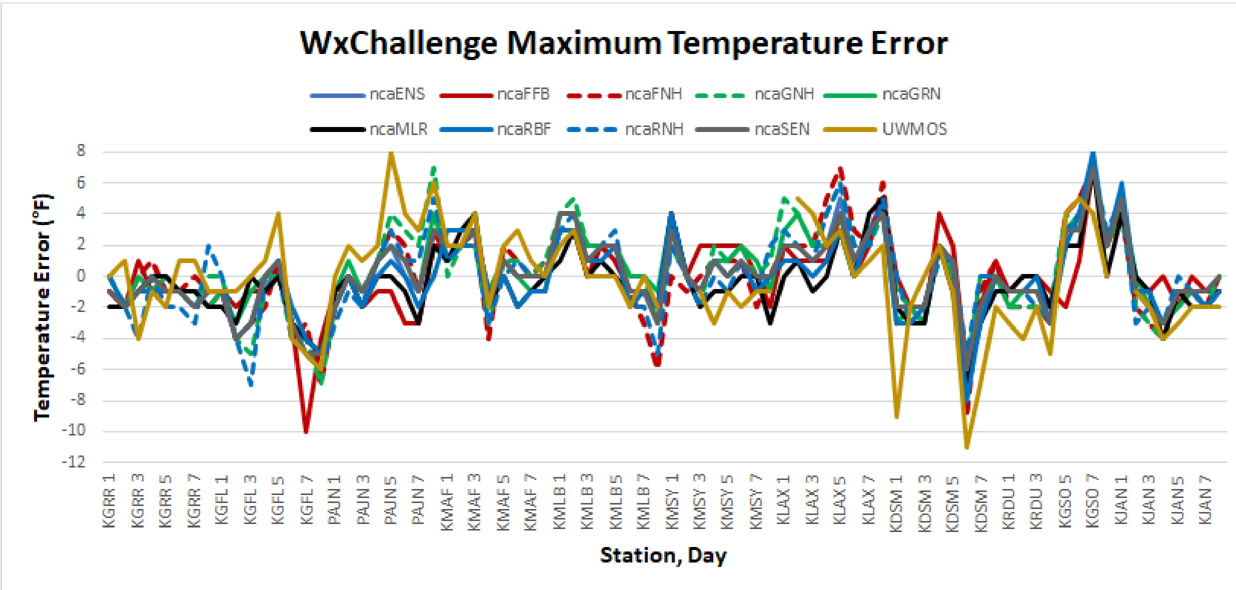

Figure 1. Daily Maximum Temperature Error (°F) for All Neural Networks (-NH Variants in Dashed Lines), MLR and UWMOS.

3.1 Maximum Temperature Analysis and Case Studies

Surface temperatures are the most important variables to consider in achieving minimal forecasting error, as certain thresholds can be critical (i.e., freezing level for vegetation, wind chill calculations for frostbite and heat index for those more at risk of heat stroke). Generally, the suite of networks outperformed every input model in temperature forecasts in each city and across all 10 cities; however, there are a few cases that are evident in Figure 1 of maximum temperature errors that are necessary to address. Coastal cities and those near bodies of water (i.e., KGRR, Grand Rapids, MI; KMLB, Melbourne, FL; KMSY, New Orleans, LA) exhibited less intense maximum temperature errors on average than more continental stations (i.e., KDSM, Des Moines, IA and KGSO, Raleigh, NC).

KGFL (Glens Falls, NY) sits in a bowl, or basin, topographically. This alone can lead to compounding model forecast errors, as a downslope wind on the nearby slopes would warm air temperatures and lead to cloud clearing. On day 7 of KGFL, the NAM MOS resolved no downslope winds, and thus kept a dense overcast deck. The model had a max temperature 11 °F cooler than the GFS MOS, and the temperature verified 12 °F above the NAM MOS forecast. The -NH networks were not affected by this factor, but the FFB (solid red line) held on to the single solution of the NAM's incorrectly resolved overcast deck. This day inspired the creation of an ensemble of FFB runs with different algorithms (i.e., least-squares, Bayesian regression, and scaled conjugate gradient) and a number of members running on each algorithm, as an effort to eliminate single solution errors. The feedforward ensemble network began running during the latter five cities, so it will be discussed in future articles.

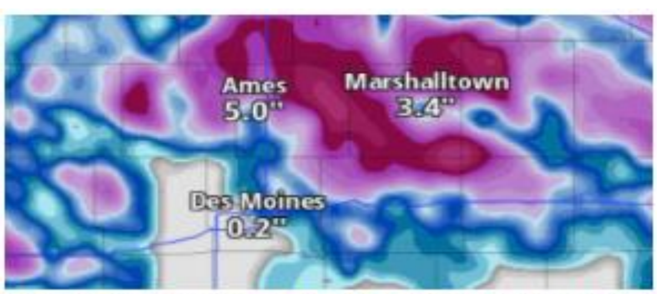

Figure 2. 25 February 2021 18Z HRRR Modeled Snow Depth from Pivotalweather.

Forecasts for KDSM were done in February 2021, when snow depth differences between urban and rural areas were large. These snow depths were not well resolved by any input model. Despite KDSM reporting a snow depth of 5” on February 25, the HRRR modeled snow depth for that day—shown in Figure 2—indicates none at the airport and overestimates 18” to the northeast. For the max temperature, this meant that the HRRR would have too warm of temperatures except on days with plenty of solar radiation and weak winds, such as KDSM day 6. On this day, every other input model was 10 °F too cool due to the snowpack causing smoothed out cold temperatures in the area. These temperatures led to unrealistic saturation with the given amount of moisture in the area during that time, making cloud decks that were far too dense and limiting the max temperature substantially. The same problem can be seen with the UWMOS on day 1 (gold line), a model which relies on dynamical GFS and NAM as opposed to MOS products. On that day, the dynamical models had substantial unrealistic cloud cover originating from smoothed snowpack depth, whereas the set of input models for the neural network suite was closer to reality.

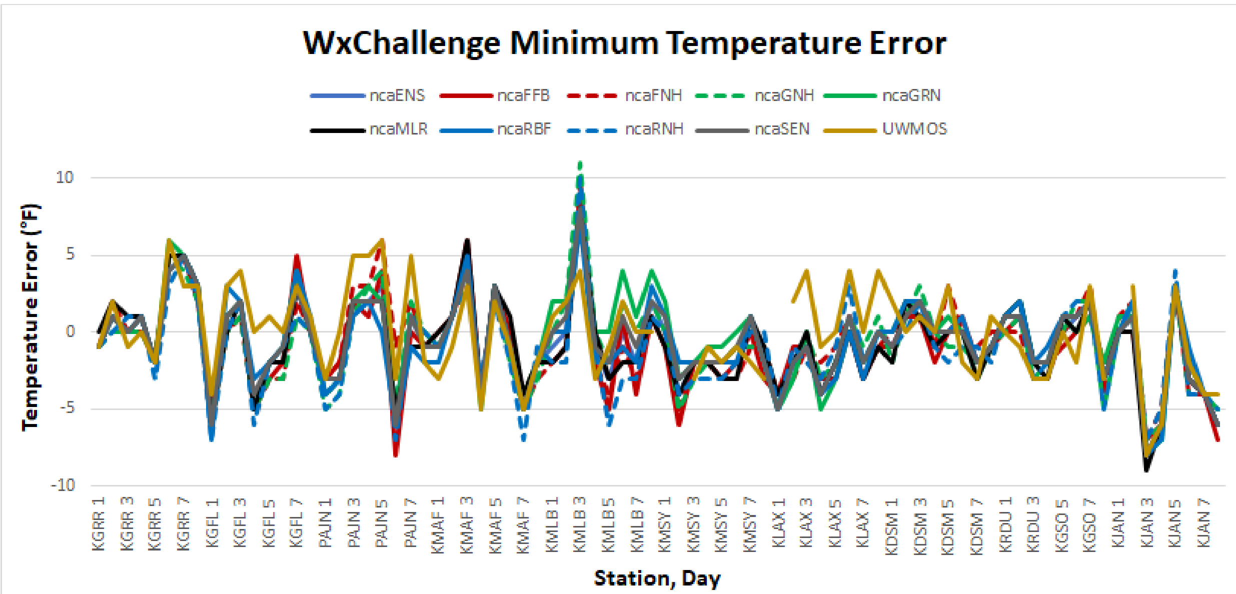

Figure 3. Daily Minimum Temperature Error (°F) for All Neural Networks (-NH Variants in Dashed Lines), MLR and UWMOS.

3.2 Minimum Temperature Analysis and Case Studies

The first instances of notable generalization—and thus lack of uncertainty—from the networks happened in the minimum temperature forecasts for KGRR. Days 1–5 overnight were breezy or cloudy, or both, so the networks did well. Days 6–8 had good radiational cooling conditions overnight, so the input models and networks were too warm. This same concept would remain evident across all 10 cities. In the case of KGRR, the station’s proximity to Lake Michigan meant that the historical dataset had a bias closer to zero for the minimum temperature, as strong radiational cooling occurred less frequently. Thus, the networks would forecast minimum temperatures at KGRR much closer to the input model forecasts compared to a station that experienced strong radiational cooling more frequently, such as KGFL. Network forecasts for KGFL minimum temperatures were always at least 3 °F cooler than input model forecasts, which was a beneficial factor when the skies were clear and winds were calm. Any day during KGFL where the networks had negative temperature errors—seen in Figure 3—had clouds overnight that prevented plummeting lows.

KMLB day 3 was a special case where a decently strong thermally direct seabreeze during the day was forecast to stay throughout the night, not allowing the land breeze to advect cooler inland temperatures to the coast. The seabreeze ended up subsiding in the middle of the night right in the middle of the city of Melbourne, so all neural network and input model minimum temperature forecasts were at least 6 °F off of verification for the station. The immediate coastline remained above the expected seabreeze-influenced minimum temperatures at Cocoa Beach, FL. This was an instance of network dependency on input models’ resolutions, which were too coarse to resolve the sea breeze boundary’s subtle movement.

For KJAN (Jackson, MS) days 3–4, a variety of factors played a role in keeping temperatures way warmer than expected. There was a period of heavy rain during day 3, so it is very likely that a moist ground kept near-surface moisture higher than what was modeled. This alone would prevent dropping temperatures in the area. Other factors are less clear in this case: there is a reservoir to the north of the station, but winds were northwesterly. It's possible that mesoscale pockets of relatively warmer air tied to more moist areas that experienced heavy rainfall were continuously advected into the area throughout the night.

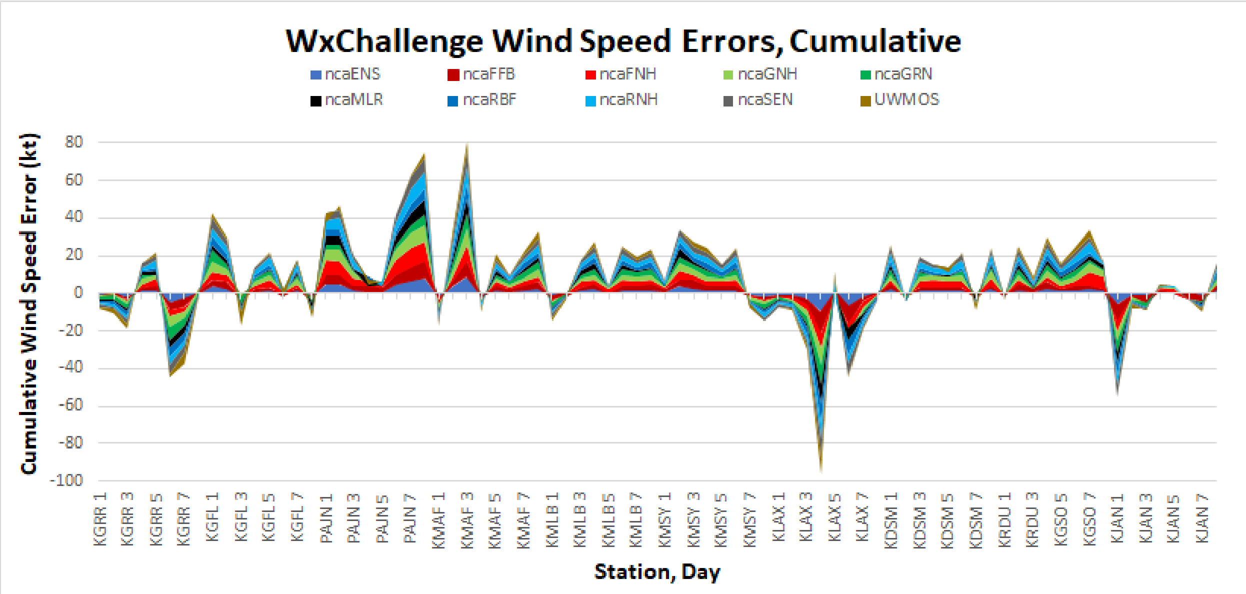

Figure 4. Daily Cumulative Wind Speed Error (kt) across All Neural Networks, MLR and UWMOS.

3.3 Wind Speed Analysis and Case Studies

Cumulative wind speed errors shown in Figure 4 demonstrate the generalized bias correction the networks create when training on historical data versus verification. The input models forecast wind in averaged hourly or 3-hourly intervals, so their bias versus verification is always largely negative, as verification reflects sustained 2-minute wind speeds. The networks attempt to correct for this negative bias continuously, so any day where winds are more constant and do not gust will typically lead to over-forecasted wind speeds.

Two wind speed case studies are of interest: PAJN and KLAX (Los Angeles, CA). The input models' biases for PAJN wind speeds were the most negative of any of the 10 stations. This stems from the low resolution most models have of Alaska regions, in addition to the complex terrain around Juneau. PAJN sits right on the Pacific Ocean, but is also in the narrow Gastineau Strait and is surrounded by snow covered tall mountains on the order of 3000–6000 feet in elevation. Nearly every day, wind directions were forecast to be almost synonymous with synoptic flow, whereas they verified in line with the orientation of the strait. The strait tends to accelerate winds climatologically, so very light wind forecasts from the input models were still heavily corrected for in the networks' forecasts. Thus, when winds were light and variable, the networks were still forecasting breezy 15 kt days, which led to significant over-forecasting that is evident in Figure 4.

KLAX sits very close to the Pacific Ocean as well, but has more drastic temperature gradients in the area due to the hot inland temperatures and cool, upwelled ocean surface temperatures. Sea breezes here are critical in both temperature and wind speed forecasts. Greater temperature gradients onshore and offshore correspond to greater wind speed verifications. Input models that do not have fine enough resolution cannot adequately forecast the strength of the temperature gradient and thus the wind speed, as inland temperatures become blended with ocean temperatures at the coast. This was the case during day 4 of KLAX, where downslope winds from the ridgeline to the north introduced warmer temperatures to the area, which were subsequently fought off by a stronger sea breeze.

3.4 Precipitation Analysis

Out of the given atmospheric variables, precipitation is the most difficult and chaotic to accurately forecast. Cumulative errors on specific days shown in Figure 5 were all similarly aligned due to the dependence of all networks on two models for precipitation forecasts out of the set of all input models. As a result, there was little

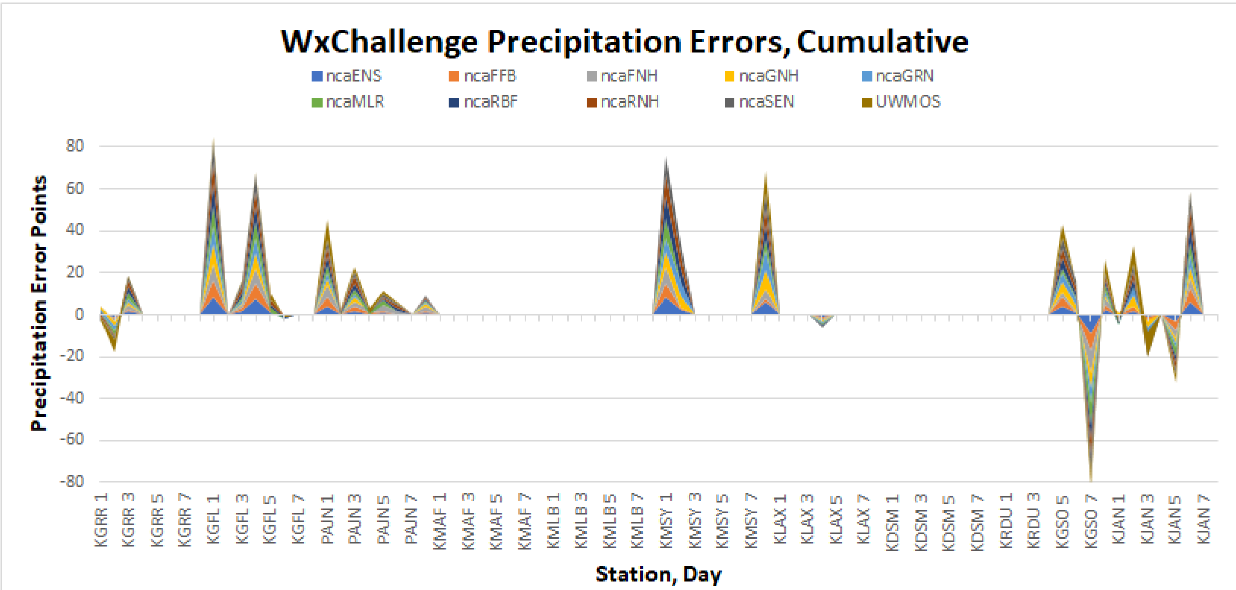

Figure 5. Daily Cumulative Precipitation Error (in WxChallenge error points per hundredth inch: 0.4 for 0.01–0.10”, 0.3 for 0.11–0.25”, 0.2 for 0.26–0.50”, 0.1 for 0.51+”) across All Neural Networks, MLR and UWMOS.

variance in the precipitation forecasts from all networks—they would all perform poorly in the same manner if verification was far off.

4. Discussion and Conclusion

A suite of neural networks, an MLR model, and two ensemble members were developed and showed promising forecast precision in short-range periods for ten different U.S. cities of various climates in the WxChallenge. Across those cities, some compound errors are evident. This suggests that the neural networks are not capable of forecasting uncertainty. They improve upon notable biases in models versus verification over the years of historical data, but do not necessarily capture the perfect forecasts when all elements of radiational cooling are in place, or if a cloudy day moderates the maximum temperature by a few degrees only for the sun to peek through a break in the clouds and warm the surface. The neural networks rely on the set of inputs from the chosen dynamical and statistical models and improve upon their biases in an overall scheme, rather than per specific scenario.

This opens up the idea for a physics-aware neural network to forecast the given atmospheric variables in this study. With parameters of wind direction and cloud cover included, a physics-aware neural network could gather a much better idea of the model biases specific to scenarios involving these additional variables. Such a network would improve low temperature errors by large amounts on nights with overcast conditions or winds that bring air from over a nearby body of water to the station. The physics-aware network would account for input model low temperature biases only in these scenarios and use this refined set of data points to produce a more accurate forecast.

There is a strong indication that a greater number of skillful models in the set of inputs for the neural networks will result in higher forecast skill from the networks. These are promising results that provide useful information in further refining the neural networks in the future. This is evident in the precipitation results, as reliance upon only two models for determining precipitation forecasts gave way to large errors when both models or either one was a daily outlier. The -NH variants, therefore, suffer equally in other variables due to sensitivity in only having two models serving as input. Temperature and wind errors were not as evident, as they are much less random and more predictable than precipitation, which is one of the most difficult meteorological variables to correctly forecast. Temperature has a strong diurnal component and therefore is easier for models to accurately predict within reason, while wind speed is directly correlated to the strength of the pressure gradient, giving models a decent shot at wind speed estimation. With these factors in mind, the -NH networks' corresponding lower cumulative scores and ranks point away from fewer models in an input set for short-range weather forecasting neural network architecture.

5. Acknowledgements

The author is grateful to Dr. Brian Blaylock at the University of Utah for their ample assistance in accessing the HRRR archive and in creating Python scripts to work with the GRIB2 files for WxChallenge times, variables, and gridpoints.

The author appreciates Yeechian Low at McGill University for assisting in conversations about the structure and mechanisms of the UW MOS-X, which served as a basis for testing different architectures of neural networks for skill performances.

6. References

1. Ustaoglu, B., H. K. Cigizoglu, and M. Karaca, 2008: Forecast of daily mean, maximum and minimum temperature time series by three artificial neural network methods. Meteor. Appl., 15, 431–445, https://doi.org/10.1002/met.83.

2. Novak, D. R., C. Bailey, K. F. Brill, P. Burke, W. A. Hogsett, R. Rausch, and M. Schichtel, 2014: Precipitation and temperature forecast performance at the Weather Prediction Center. Wea. Forecasting, 29, 489–504, https://doi.org/10.1175/WAF-D-13-00066.1.

3. Ghosh, T., and T. N. Krishnamurti, 2018: Improvements in hurricane intensity forecasts from a multimodel superensemble utilizing a generalized neural network technique. Wea. Forecasting, 33, 873–885, https://doi.org/10.1175/WAF-D-17-0006.

4. Krasnopolsky, V. M., and Y. Lin, 2012: A neural network nonlinear multimodel ensemble to improve precipitation forecasts over continental US. *Advances in Meteor.*, **2012**, https://doi.org/10.1155/2012/649450.

5. Thorarinsdottir, T. L., and Gneiting, T., 2009: Probabilistic forecasts of wind speed: ensemble model output statistics by using heteroscedastic censored regression. *J. Roy. Statist. Soc.*, **173**, 371–388, https://doi.org/10.1111/j.1467-985x.2009.00616.x.

6. Sanders, W. M., 2017: Machine learning techniques for weather forecasting. M.S. thesis, University of Georgia Athens, 77 pp.

7. Weyn, J. A., 2018: UW-MOS X. Accessed 28 September 2020, https://atmos.uw.edu/~jweyn/MOS-X.

8. Blaylock, B., and J. Horel., 2019: Archive of the High Resolution Rapid Refresh model. University of Utah Center for High Performance Computing, https://doi.org/10.7278/S5JQ0Z5B.