

# A forward look at the alteration of the weather forecast by artificial intelligence

Tianyu Wu

Villanova Preparatory School, 12096 N. Ventura Avenue, Ojai, USA, CA 93023

twu@villanovaprep.org

**Abstract.** Most modern weather forecasts are accomplished by atmospheric simulations based on physics models. This method requires a lot of processing time due to the massive amount of computing work. Meanwhile, accuracy faces a limit due to the chaotic nature of the atmosphere. As machine learning gets better and becomes more involved in society, which has been an accelerated trend in recent years, some machine learning-based weather forecast models have been developed. This paper provides a forward-looking study of the potentials of machine learning in terms of theoretical abilities, shown advantages, and possible alterations in the field of weather forecasting based on existing literature and data. Explanations and discussions of two outstanding machine learning-based models, MetNet-2 and Pangu, are provided. Based on the results of tests of these two models, machine learning has good advantages that can be largely applied in the field of weather forecasting and other atmosphere-related work. Machine learning-based methods are likely going to replace physics models in most parts, while physics knowledge is still valuable in this field as it can help with the training of machine learning models and improve performance.

**Keywords:** Weather Forecast, Atmosphere, Machine Learning, Neural Network, Physics

## 1. Introduction

The atmosphere of Earth, surrounding the planet, is a massive gas layer held by the gravity of Earth. The motions of the gas particles and the change of states of the water steam, mainly contributed by solar irradiance, form wind and precipitation, which are usually referred to as the main elements of weather. Efforts on weather prediction appeared in human history because many human activities require fitting weather, and a weather forecast is necessary for managing these activities [1]. Nevertheless, the atmosphere is a chaotic system in which any small change could cause a significant difference. As a result, though weather simulation systems based on physics-based models have appeared continuously since the invention of computers, radar, and weather satellites, the accuracy of these simulations faces a limitation [2]. Meanwhile, running these models requires a massive amount of computing work due to the complex mathematical equations and thus requires very long processing times, making the job of weather prediction difficult. In order to achieve more accurate and timely weather prediction, a new method other than a physics-based model is needed.

Some theories, starting from the late 20th century, suggest that machine-learning-based weather simulations take much less computing work and could accomplish prediction work in a significantly shorter amount of time while maintaining good accuracy compared to ordinary physics models [3]. As

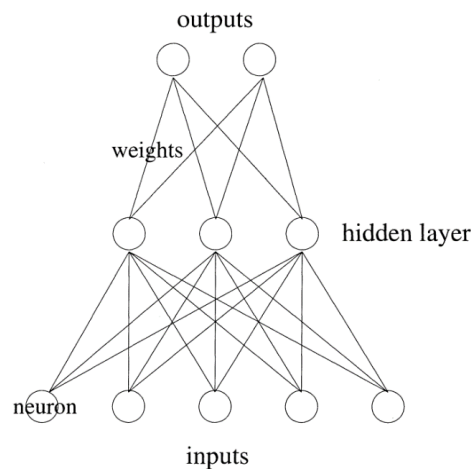
the technology of the graphic processing unit (GPU) has accelerated in recent years, machine learning appears to involve, in different forms, all parts of society. The possibility of AI-based models altering the whole society in the form of a much more advanced weather forecast system has been shown. This paper presents basic ideas about machine learning and how it is connected to weather forecasting, as well as explanations and discussions of two different existing machine-learning-based weather forecast models, in order to analyse the capabilities of machine learning in this field better.

## 2. Literature review

While physics simulation still occupies a mainstream position in the field of weather forecasting, studies about the application of machine learning in this field have appeared since the late 20th century. A number of studies have made valuable comparisons between the predictions of traditional methods and those of machine-learning methods. In 2010, a study of the applications of the Adaptive Neuro-Fuzzy Inference System (ANFIS) and Autoregressive Integrated Moving Average (ARIMA) models in weather forecasting was published [4]. ANFIS is a system combining the learning capabilities of neural networks with the interpretability and structure of fuzzy systems in order to form a powerful system dealing with problems regarding approximation, classification, and forecasting [5]. ARIMA is a statistical model used for time-series forecasting [6]. This study presented a weather forecast for Göztepe, Istanbul, Turkey, using data on daily average temperature (dry-wet), air pressure, and wind speed from 2000 to 2008 [4]. The result of this test shows that ANFIS, the neural network-based model, presents better performance than ARIMA, the traditional statistical model. Results suggest that the ANFIS model, while showing its simplicity and reliability, can capture dynamic behaviours of weather data more effectively and efficiently, resulting in a better presentation of the temporal information contained in the weather profile. Based on existing literature and data, the paper presents results of studies comparing machine learning-based weather forecasts and traditional weather forecasts that were published in recent years, finds specific traits or advantages of machine learning that are shared in different comparisons, and suggests capabilities of machine learning in this field.

## 3. Weather system and machine learning

### 3.1. Basic ideas about machine learning



**Figure 1.** Structure of neural network [7]

Machine learning, as a type of AI, uses artificial neural networks to accomplish the prediction of results based on input. The common structure of a neural network contains an input layer, an output layer, and a hidden layer between them (Figure 1):

Each neuron in a layer has an individual input signal and an output signal. The input signal, despite the very first layer having given inputs, is the output of all neurons at the previous layer.

Values from each neuron in the previous layer will be multiplied by a weight, a number specific to each previous neuron, before being added together to form one function. It then adds a bias, a number specific to a neuron processing all these inputs. This process is referred to as a linear operation. Afterward, the value gets applied to an activation function, which is a nonlinear function, forming a new nonlinear function. This process is referred to as a nonlinear operation. After all these processes, this new function will become the output value for this neuron and be sent to all neurons at the next layer. If this layer is the final output layer of the network, then the outputs of neurons at this layer are the final output results of this neural network [8].

The essential process a neural network does is to conjecture an unknown function by processing the weighted and biased functions in each neuron. Since the conjecturing process depends on all the neurons, and the trait of each neuron depends on its weight and bias, this neural network's trait can be adjusted by simply adjusting all the weights and biases of each neuron. Thus, the entire process of training a neural network is a repetitive process of generating weights and biases, comparing output results with the expected results, and then adjust weights and biases, trying to make outputs more accurate, meaning more similar to expectations. Once this neural network can achieve a specific accuracy, the final recorded weights and biases could be used to predict the same problems with different inputs.

### 3.2. Advantages of machine learning

The theory of applying machine learning to weather forecasting appeared in the late 20th century. The process of weather prediction needs to deal with massive amounts of data with high dimensions: the Earth has a surface area of more than five hundred million square kilometres covered by air of different temperatures, air pressures, humidities, wind speeds, etc. All these values would affect the status of nearby air. The atmospheric weather system is an extremely complicated nonlinear system. Using physics knowledge and mathematical equations to conjecture this system is, therefore, a difficult job.

Nevertheless, promoted by studies done by computer scientists in recent decades, in which computer technology has been accelerating, machine learning has shown much capacity for dealing with this job because of its following four traits: 1) Machine learning shows effectiveness in capturing the nonlinear mappings and modelling the nonlinear systems, which is contributed by the usage of activation functions [9]. 2) Machine learning can apply dimensionality reduction operations on the data, transferring high-dimensional input into low-dimensional input, making the process significantly less complex [10]. 3) Machine learning can sensitively capture temporal dependencies, which is helpful since the weather system is entirely a temporal system [11]. 4) Machine learning can quantify estimates of uncertainty. It is valuable for modelling weather systems in which problems regarding uncertainty appear.

## 4. Current existing machine-learning-based weather forecast systems

### 4.1. MetNet-2: a precipitation prediction model

MetNet-2 is a neural network weather-predicting model developed by Google. This model is capable of providing weather event predictions in a given area in the form of the probability of a specific weather event, for instance, a specific amount of precipitation. Its predictions are made by learning transformations from the recorded weather data facts collected by radars and satellites [12].

**Table 1.** Comparison between MetNet-2 and NWP

Model/Hours	1	2	3	4	5	6	7	8	9	10	11	12
<b>HRRR</b>	0.32	0.24	0.23	0.23	0.22	0.20	0.20	0.19	0.18	0.16	0.16	0.16
<b>HREF</b>	0.29	0.28	0.26	0.28	0.25	0.22	0.25	0.23	0.20	0.22	0.20	0.20
<b>MetNet-2</b>	0.64	0.50	0.43	0.36	0.33	0.29	0.26	0.23	0.22	0.20	0.18	0.17

**Table 2.** Comparison between performances of MetNet-2, HRRR(NWP), and MetNet-2 in Hybrid mode

Model/Hours	1	2	3	4	5	6	7	8	9	10	11	12
<b>HRRR</b>	0.17	0.16	0.16	0.15	0.14	0.14	0.15	0.15	0.15	0.14	0.14	0.15
<b>MetNet-2</b>	0.44	0.36	0.32	0.29	0.26	0.25	0.24	0.23	0.23	0.22	0.22	0.21
<b>MetNet-2 Postprocess</b>	0.30	0.26	0.25	0.24	0.23	0.23	0.24	0.24	0.24	0.24	0.24	0.24
<b>MetNet-2 Hybrid</b>	0.44	0.36	0.32	0.29	0.28	0.27	0.26	0.27	0.26	0.26	0.25	0.25

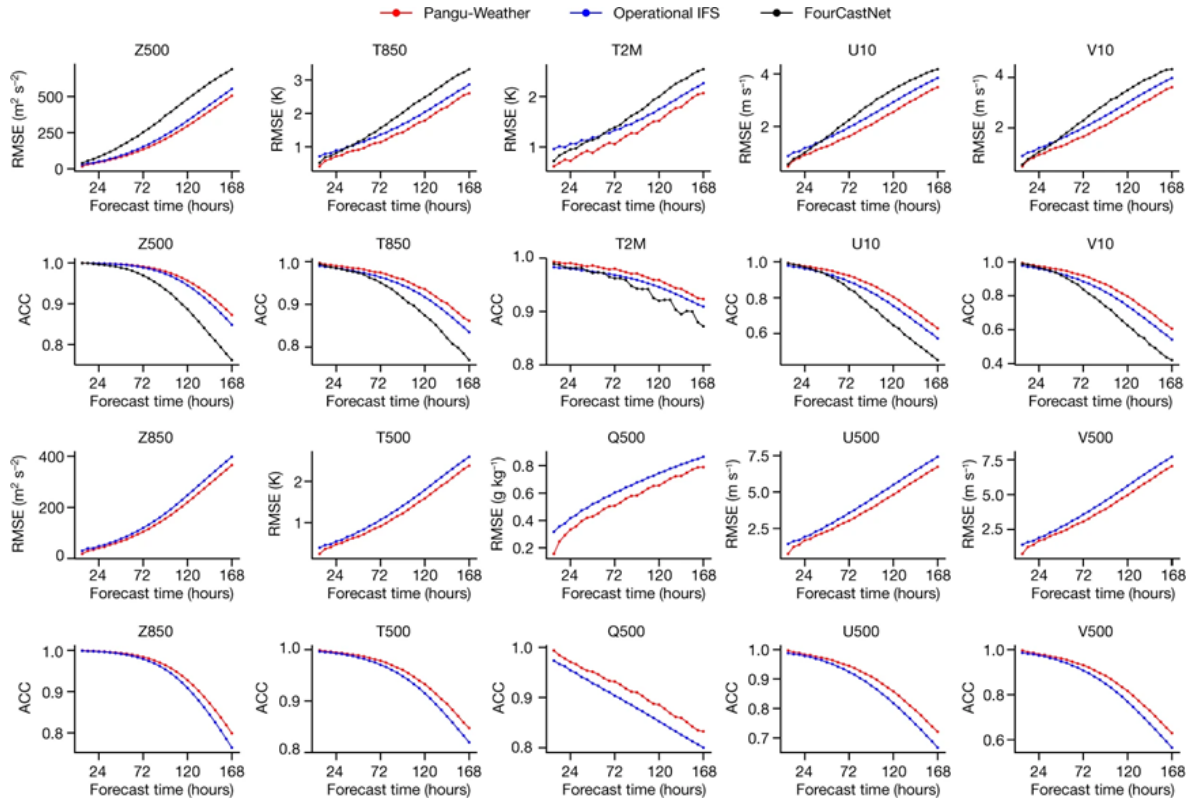
In a test of MetNet-2, the MetNet-2 model was trained to forecast the instantaneous precipitation in 12 hours for a 7000 km × 2500 km region located in the United States, compared to the predictions of the High-Resolution Rapid Refresh (HRRR) and High-Resolution Ensemble Forecast (HREF), two ordinary numerical weather prediction (NWP) models run on the same supercomputer. As shown in Table 1, which represents the comparison of the Critical Success Index score (CSI) of the prediction results in each hour, in the first nine hours, the CSI of the results of MetNet-2 is mostly higher (and partially equal to) than the CIS of the results of both NWPs. This outperformance is particularly noticeable in the first three hours: the CSI of MetNet-2 in these three hours is all greater than or equal to 0.43 and has an average of higher than 0.5, but the CSI of the NWPs in these hours is all less than or equal to 0.32 and has averages less than 0.28. In the following six hours, the outperformance of the CSI of MetNet-2 decreases continuously. In the tenth hour, the CSI of the results of HREF is higher than the CSI of MetNet-2 [12].

In another test, MetNet-2 applied in a hybrid mode, using physics-based forecasts as additional input, was used in the prediction of another 12 hours: MetNet-2 Postprocess maps the forecast of HRRR to a probabilistic one; MetNet-2 Hybrid maps both the forecast of HRRR and the default input of MetNet-2. Table 2 represents the CSI comparison between the results of HREE (NWP), ordinary MetNet-2, and MetNet-2 in the two hybrid modes. The data table shows that in the first four hours, the CSIs of the results of ordinary MetNet-2 and MetNet-2 Hybrid are the same, and both are higher than HRRR (NWP) and MetNet-2 Postprocess. Starting from the fifth hour, similar to the results of the first test, the outperformance of CSI of MetNet-2 decreases continuously. However, the CSI of MetNet-2 Hybrid outperforms the results of all other models starting from the fifth hour and maintains its outperformance in all 12 hours, showing that mixing NWP (HRRR) and Metnet-2 brings a significant increase to the performance of MetNet-2, fixing its disadvantage in the performance of predictions longer than three to four hours [12].

These two tests show that MetNet-2 has good performance in weather prediction. The neural network of MetNet-2 successfully captures the nonlinear mappings of weather variables by learning from the variations among the data and thus can predict uncertainty with remarkable accuracy. On the other hand, it can capture additional mappings and strengthen its performance by learning from additional observations, which in this case are the HRRR model [12]. Additionally, MetNet-2 shows good time efficiency: MetNet-2 requires about 1 second to finish core computations for a 12-hour forecast, while HREF and other NWPs require about 1 hour in the same hardware condition.

#### 4.2. Pangu: a 3D-neural-network weather forecast model

Pangu is a medium-range global weather forecast model based on a 3D neural network developed by Huawei. It identifies height as a new dimension to develop a three-dimensional Earth-specific transformer (3DEST), an architecture enabling analysis of both the upper and surface atmospheres. This system captures the atmospheric state at different air pressures and thus improves the accuracy of predictions. This model is trained and evaluated using the fifth generation of ECMWF reanalysis (ERA5), a reliable global atmospheric reanalysis dataset [13]. The model is trained using weather data from 341,880 time points (from 1979 to 2017). It provides a comprehensive forecast with ten variables: geopotential, temperature, specific humidity, and wind speed at different air pressure conditions [14].



**Figure 2.** Comparison between Pangu, FourCastNet (AI), and Operational IFS (NWP) [14]

In a deterministic forecast test, Pangu predictions are compared to predictions of integrated forecasting system (IFS) and FourCastNet. IFS is currently one of the most accurate NWP models [15]; FourCastNet is one of the most accurate AI-based models [16]. The three models have the same starting-state data by ERA5 and predict weather variables in a range of 168 hours.

Ten variables were compared in this test where FourCastNet only provides the first five (all at 500 hPa): geopotential (Z500), temperature (T500), specific humidity (Q500), and the u-component and v-component of wind speed (U500 and V500). The last five are: geopotential and temperature at 850 hPa (Z850 and T850), 2-m temperature (T2M), and the u-component and v-component of 10-m wind speed (U10 and V10), respectively. Variables are evaluated in terms of Root Mean Square Error (RMSE) (lower means better) and accuracy (ACC) (higher means better) [14].

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(Predicted\ i - Actual\ i)^2}{n}}$$

$$ACC = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions}$$

The result evaluations, shown in Figure 2, show that Pangu outperforms the other two models in predictions of all ten variables at all time ranges. The accuracy advantage of Pangu predictions tends to be clearer when the time range gets longer. This astoundingly good performance shows the strong capacity of Pangu for medium time range weather predictions. Meanwhile, the processing speed of Pangu is comparable to IFS and much faster than FourCastNet and other NWP models: Pangu needs 1.4 seconds to compute a 24-hour forecast on a Tesla-A100 GPU. Using the same hardware condition, IFS

requires 0.28 seconds, and FourCastNet and other NWP models need several hours. Considering the amount of data Pangu processes, it shows a huge advantage of timely efficiency.

## 5. Discussion

The potential abilities of capturing nonlinear mappings and the efficiency brought by dimensional reduction are especially represented by these two models: MetNet-2 and Pangu. It is shown in the results of both studies that machine learning models have the special advantage of capturing relationships between variables in a changing nonlinear system more efficiently and effectively than traditional methods, further proving the value of machine learning in solving problems regarding complex chaotic systems. Meanwhile, values of hybrid mode models, meaning models combining machine and traditional physics knowledge, were shown in the study of MetNet-2, in which the combination of the HRRR model significantly increased the performance of forecasts, especially at a longer time range, making its accuracy higher than other testing models at all hours. But related tests are not presented in the study of Pangu. Since Pangu is a more compressive model, more types of weather variables, physics knowledge, rules, and models that have already been discovered by humans could have a more significant effect if combined with Pangu. Thus, future studies of this topic could focus on discovering the value of hybrid models rather than models that process only based on data by combining more known physics knowledge, such as models of fluid physics and thermodynamics, with existing machine learning-based models.

## 6. Conclusion

The two machine-learning-based models have differences in their designed objectives: MetNet-2 is a 2D-trained model designed for short-term forecasts and focuses specifically on precipitation; Pangu, as a 3D-trained model, is designed for long-term forecasts and involves predictions of multiple variables. Nevertheless, both models showed remarkable performance in accuracy and time efficiency compared to NWP models. Thus, machine learning is likely to change the entire field of weather forecasting and replace most ordinary methods. Future weather forecasting and other aerology-related works are likely to rely largely on machine learning in consideration of its advantages. Nevertheless, physics knowledge and mathematical simulations are still going to be valuable because they can be combined with AI as additional input training materials and improve the effectiveness of AI model training. The current research is mainly based on the results of existing literature research, and future research will gather specific data and models for in-depth analysis.

## Acknowledgement

Firstly, I would like to show sincere gratitude to Professor Mario Campanelli and Professor Chen, who provided me with teachings about machine learning and physics. Without them, the paper cannot be completed. Secondly, I would like to thank my teaching assistants who provided guidance for my paper. Further, I want to thank the scientists at Google and Huawei who provided valuable references for this thesis.

## References

- [1] Bos, G., & Burnett, C. (2000). *Scientific Weather Forecasting In The Middle Ages: The Writings of Al-Kindi* (1st ed.). Routledge.
- [2] Webb, R. (2008). Chaotic beginnings. *Nature Geoscience*, 1(6), 346-346.
- [3] Dewitte S, Cornelis J P, Müller R, et al. 2021. Artificial intelligence revolutionises weather forecast, climate monitoring and decadal prediction[J]. *Remote Sensing*, 13(16): 3209.
- [4] Tektaş, Mehmet. 2010. "Weather forecasting using ANFIS and ARIMA models." *Environmental Research, Engineering and Management* 51.1: 5-10.
- [5] Jang, J-SR. 1993. "ANFIS: adaptive-network-based fuzzy inference system." *IEEE transactions on systems, man, and cybernetics* 23.3: 665-685.

- [6] Newbold, Paul. 1983. "ARIMA model building and the time series analysis approach to forecasting." *Journal of forecasting* 2.1: 23-35.
- [7] Wang S C, Wang S C. 2003. Artificial neural network [J]. *Interdisciplinary computing in java programming*, : 81-100.
- [8] Sharma S, Sharma S, Athaiya A. 2017. Activation functions in neural networks [J]. *Towards Data Sci*, 6(12): 310-316.
- [9] Lapedes A, Farber R. 1987. Nonlinear signal processing using neural networks: Prediction and system modelling[R].
- [10] Reddy G T, Reddy M P K, Lakshmana K, et al. 2020. Analysis of dimensionality reduction techniques on big data[J]. *Ieee Access*, 8: 54776-54788.
- [11] Lai G, Chang W C, Yang Y, et al. 2018. Modeling long-and short-term temporal patterns with deep neural networks[C]//The 41st international ACM SIGIR conference on research & development in information retrieval. 95-104.
- [12] Espeholt, L., Agrawal, S., Sønderby, C. et al. 2022. Deep learning for twelve hour precipitation forecasts. *Nat Commun* 13, 514. <https://doi.org/10.1038/s41467-022-32483-x>
- [13] Hersbach, Hans, et al. 2020. "The ERA5 global reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146.730: 1999-2049.
- [14] Bi, K., Xie, L., Zhang, H. et al. 2023. Accurate medium-range global weather forecasting with 3D neural networks. *Nature* 619, 533–538. <https://doi.org/10.1038/s41586-023-06185-3>
- [15] Déqué, Michel, et al. 1994. "The ARPEGE/IFS atmosphere model: a contribution to the French community climate modelling." *Climate Dynamics* 10: 249-266.
- [16] Pathak, Jaideep, et al. 2022. "Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators." *arXiv preprint arXiv:2202.11214*.