Journal of Applied Research and Technology

www.jart.icat.unam.mx

Journal of Applied Research and Technology 21 (2023) 205-211

Original

Prediction of weather forecasting using artificial neural networks

A. Ajina[a*] • Jaya Christiyan K. G.[b] • Dheerej N Bhat[c] • Kanishk Saxena[c]

[a]*RV Institute of Technology and Management, Bangalore*
[b]*M S Ramaiah Institute of Technology, Bangalore*
[c]*Sir. M Visveswaraya Institute of Technology, Bangalore*

Received 07 05 2021; accepted 10 07 2022
Available 04 30 2023

Abstract: Currently, weather forecasting is the most discussed topic by social and economic activists. It is also attracting wide spread interest due to its application in various public and private sectors that include marine, agriculture, air traffic, and forestry. Recent developments have made climatic changes happen at a dramatic rate, making old methods of weather forecasting less effective, more hectic, and unreliable. Improved and efficient methods of weather prediction are needed to overcome these difficulties. This paper describes machine learning approaches using artificial neural networks to predict the weather of a particular city and compare the different weather conditions in different cities. We demonstrate empirically that artificial neural networks produce incredibly lower deviations than GDAS evaluation. Hence the prediction of nearly accurate results for weather forecasts on a daily basis.

Keywords: Artificial neural networks; AI weather forecast; Machine learning; Weather forecasting; Weather prediction

[*]Corresponding author.
E-mail address: ajinajaya@gmail.com (A.Ajina).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

1. Introduction

In the past decade, many major attempts have been made to develop cutting edge neural network models for weather forecasting using statistical and machine learning techniques which have produced promising results (Chen et. al 2000; Kwong et al., 2008; Maqsood et al., 2004). A growing body of literature has studied how climate factors like temperature, pressure, humidity affects the weather. A study by the Harvard School of Public Health (Harvard T.H. Chan School of Public Health 2020) points out that temperature and rainfall are the two vital factors that favor/oppose the spread of infectious diseases and pathogens. Infectious diseases cause immediate effects on human lives and animals whereas climate change is a slow-moving condition whose dangers may feel impersonal, and their causes spread all over the world. Weather prediction methodologies model the state of the environment and forecast the weather using data from different atmospheric properties such as temperature, pressure, and humidity, etc. The most used approaches are Kalman filters (Burgers et al. 1998; Evensen et al. 2003) and neural networks (Chow & Leung 1996; Lai et al., 2004).

Weather forecast reports enable cognitive computation to read nonlinear data and produce rules and patterns to analyze from observable data in order to predict future weather. The use of neural networks can provide more reliable results. The errors produced by these networks may or may not be absolutely eliminated in this situation. However, in contrast to previous predictions, the accuracy would increase. With the generation of a tremendous amount of data in previous years, there is immense weather-related information that can be utilized for training a machine learning algorithm to predict climatic conditions of a particular town/city or a state. Due to the unexpected COVID-19 pandemic, a study as shown by Quéré et al. (2020) shows that there was a temporary reduction of carbon-dioxide emissions daily of -17% which leads to an annual decrease of around -4.2 to -7.5% that would limit climate change by 1.5℃ warming. The principal sectors which have led to this decrease are the commercial enterprises, surface transport, various manufacturing industries, residential and aviation industry.

In today's environment, weather forecasting has become much more challenging. Natural disasters including cyclones, hurricanes, earthquakes, forest fires, and tsunamis occur almost regularly as a consequence of climate change and inflict catastrophic damage. This certainly changes how the weather forecasting models used the previous data to predict the weather for the future. Accurate weather monitoring is extremely important and beneficial in these difficult times. We begin with a literature survey in section 2, which provides a brief tour of the previous works and knowledge on the topic. In section 3, we provide a machine learning approach for weather prediction. Section 4 describes the settings for performing the experiments. Section 5 explains the evaluation results followed by section 6 which concludes the paper.

2. Related work

In recent years there has been growing interest among the communities in developing automated machines that will predict accurate weather based on the previous data available. In this section, we present some papers which have focused on weather prediction using machine learning algorithms. The National Centre for Atmospheric Research (NCAR) was one of the premier institutes that started applying machine learning and statistical techniques for accurate weather prediction. The system was called dynamic integrated forecasting (DICast 1998) system which utilized statistical data and observations for automatic inferences. The NCAR has also developed a weather research and forecasting (WRF) model (1990) which is a scientific weather forecasting system used for both meteorological analysis and forecast applications.

The systematic study of using directed acyclic graphs namely Bayesian networks to model and make climate expectation was performed by Cofiño et al. (2002) which gave promising results. But the practical operative efficiency of the Bayesian system was computationally costly because of the significant measures of the various conditions being performed. In the review paper by Saima et al. (2011), the author talks about all the different types of hybrid models like autoregressive integrated moving average (ARIMA), adaptive neuro fuzzy inference system, fuzzy clustering for weather forecasting. The authors draw a conclusion to the fact that the type of model being developed depends substantially on the nature of the data i.e. linear or non-linear and achieving highly accurate prediction is more important than processing time. Another technique was demonstrated (Dey & Chakraborty, 2015) which utilized DBSCAN and convex hull algorithms for incremental clustering of new data. This instance-based learning can be successfully applied to dynamic databases where the climate data are frequently changed. Weather forecasting was modeled as a multi-view LS-SVM regression problem (Houthuys et al., 2017) were to predict the temperature of the current city, it utilizes temperatures from neighboring cities along with a coupling term that enforces the alignment of error variables when training the current view. A very short-term weather prediction system (Yonekura et al., 2018) was developed using deep neural networks that studied the data of a previous couple of hours. Two methods were proposed: a point prediction model and a tensor prediction model which make use of POTEKA weather stations for high accuracy predictions. The ENSO vector features were used in rainfall predictions (Salman et al., 2015) based on weather datasets. These experiments were conducted using recurrent neural networks whose performance evaluated using the k-cross fold validation technique yielded an accurate representation of results. The authors (Singh et al., 2019) proposed a method for weather forecasting using random forest classification. They utilized a raspberry pi board that operates on real-time humidity, temperature, and pressure sensor data for rain forecasting on any particular day.

3. Methodology

Weather forecasting is one of the most technologically demanding challenges. It is an art to project the weather with incredibly low fluctuations and to get decent results. A graphical user interface (GUI) has been developed for the application using python that can easily interact with the users to predict the results. The application developed can both predict the weather of a particular city and compare the different weather conditions in different cities. Figures 1 and 2 show the complete workflow of the proposed scheme. The dataset comprises various weather attributes like mean temperature, maximum temperature, minimum temperature, pressure, wind speed, maximum wind speed, dewpoint, visibility etc. These parameter values of previous years are collected and passed to the list of machine learning models. The current data for which the weather must be estimated is also sent for prediction. We perform a comprehensive analysis of artificial neural networks at their unit level to obtain the desired aims and outcomes in the proposed system.

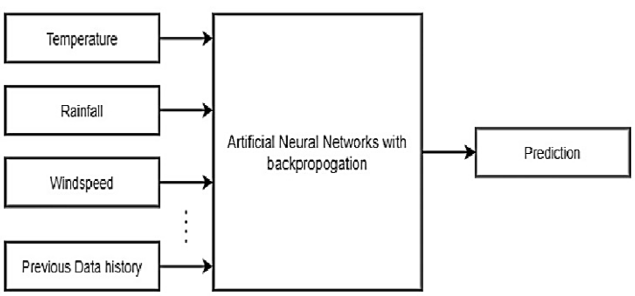

Figure 1. Prediction system.

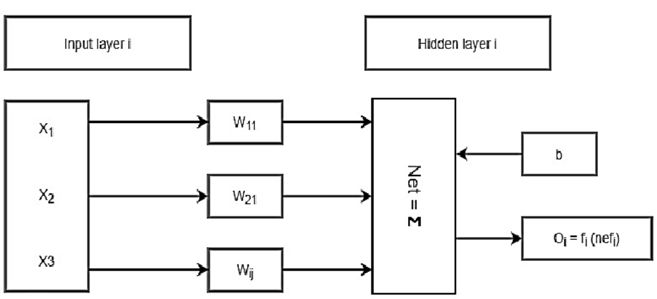

Figure 2. Perceptron model.

4. Experiment & results

4.1. Data

We investigate the global meteorological forecasts from March to May 2020 to explore the influence of aviation on these meteorological forecasts. We see a reduction in the observations worldwide during this period. We have collected our data from ncdc.noaa.gov. The model couple's several factors like land/soil, ocean, sea, ice, and atmospheric modules for the meteorological forecasts. One of the finest metrics for meteorological evaluation is data from the sigma pressure layers that occur in the atmosphere from bottom terrains to 0.27hPa (Sela, 2009). The detailed meteorological monitoring study regarding this is given in NCEP (2018). We have included in our study up to 192 hours of forecast data. Various climatic factors such as wind speed, Carbon dioxide, humidity, atmospheric pressure, and temperature are reducing our forecast prediction. Here we highlight the reduction in accuracy for temperature which plays a major role in commercial flights as previously mentioned (Petersen, 2016) in their GDAS evaluation.

To probe the impact of COVID-19 epidemic on the meteorological forecast accuracy, we compare and contrast the meteorological prediction data from March to May 2020 (Chen, 2020) (during the lockdown period) against the average of the predicted data from March to May 2015-2019. We have carried out our experiments for various cities of India like Bangalore, Mumbai, Kolkata, Chennai and Delhi. These cities are a key contributor to the climatic changes as witnessed in the country. Moreover, we direct a similar investigation for February 2020 preceding worldwide lockdown, to show that this effect on precision is related with the COVID-19 outbreak in March to May 2020 as opposed to the climatic attributes of 2020.

4. 2. Model selection

The neural network GUI nnstart accessible in MATLAB (R202) is utilized to do the experimentation on the meteorological data utilizing feed forward ANNs with backpropagation principle. The paper (Maqsood et al., 2004) does a similar investigation of MLPN, ERNN. We started our experiments with a simple neural network model and have examined the significance of the artificial neural networks (ANNs) at its fundamental level i.e the neurons. These neurons are an integral part of the whole ANN framework and to comprehend their latent capacity has been our fundamental examination.

The questions we have endeavored to provide explanations in the paper include:

1. How does the quantity of artificial neurons in the model impact the performance of the system?
2. To what extent do the hidden layers have the potential to influence the system's performance?

The inferences obtained from the above two queries are further used for parameter optimization which can help machine learning engineers and designers to choose the right number of hidden layers and neurons in each layer for optimal performance of the system. We have experimented with multiple hidden layers and utilized grid search to find the optimum number of hidden layers and artificial neurons for the best performance. The primary goal is to design a neural network architecture that can foresee the independent climatic factors, for example maximum temperature, least temperature, wind speed and so forth for a specific station and a specific day given the climate for the earlier day (target information) and previously recorded information of that specific day (input information). Our examination has inspected only the maximum temperature on the grounds that ANNs itself has various parameters that must be varied for optimization purposes during experimentation. The model consequently designed would then have the option to foresee other climate factors likewise. We have utilized the data from various cities like Bangalore, Mumbai, Kolkata, Chennai and Delhi. The input data consists of samples which include various climatic factors like atmospheric wind speed, maximum speed, pressure, maximum temperature, minimum temperature, mean temperature, humidity, pertaining to each day of the year. For the sake of consistency and uniformity, the data belonging to February 29th (in case of Leap years) has been excluded. We have doubled and several times multiplied the amount of our dataset by simply connecting it to itself on multiple instances. Even though it doesn't give the assortment expected to better speculation, it expands the learning rate. We train the artificial neural network utilizing the Levenberg Marquardt algorithm, one of the principal algorithms. The training is completed by an early stopping method. At last, the training function produces gauge results based on MSE (Mean square Error) minimization measures. In one loop of the input training, the set of data with all the attributes is introduced to the nodes of the neural network. The output is compared against the optimal output, and an error signal is created, which changes the neuron's weights and bi-ases for every iteration. This process continues until the preferred mapping to the target variable is obtained as close as possible. The weights are saved after training the neural network. To evaluate the quality of the model, the test set of information is introduced to the trained model.

5. Results and discussion

As demonstrated in Figure 3, the precision of surface meteorology forecast from March to May 2020 declines amazingly (as for March to May 2015–2019; red tones show more terrible conjectures, blue tones show better estimates) all throughout the first to eighth day forecasts. Figure 4 shows 0.5 – 1.0°C inclination contrasted with that in March–May 2015–2019. Prior to the worldwide lockdown in February, the temperature gauge is for the most part improved by 0.5–1.5°C in 2020 against 2015–2019, with little special cases in 24–48-hour estimate (Figure 3a).

Not much depreciation in values corresponding to the surface pressure and wind speed is seen in 24 – 96-hour gauges (Figures 4c, 4d). Anyway, there is a negligible development in February 2020 (Figures 3c, 3d). The mistakes in the forecasting are created as the gauges are extended. The wind speed conjecture of February 2020 shows a development of 0.2 -- 0.5 m/s against February 2015–2019, all through the estimating intervals of 24–192 hours (Figure 3d). The absolute precipitation gauges during March-May 2020 as observed using information collected has contrasted the precision with that of March - May in 2015–2019 (Figure 5). Past examinations show that aircraft observations assume a significant part in the estimates of humidity, wind and temperature (James & Benjamin, 2017; Ota et al., 2013; Petersen, 2016); Some of the precipitation conjectures (James & Benjamin, 2017) are not completely removed during the worldwide lockdown. Figure 6 shows the results of the weather forecasting over the last five years. The results in Figure 7 correspond to the Actual and predicted results. There are only a few deviations when compared to actual forecasts, but the model predicts accurately the rest of the time.

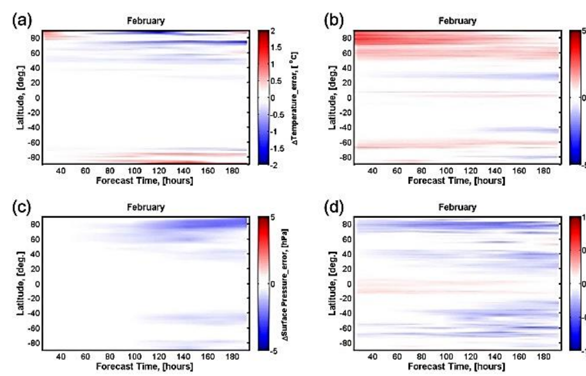

Figure 3. Forecasts in February.

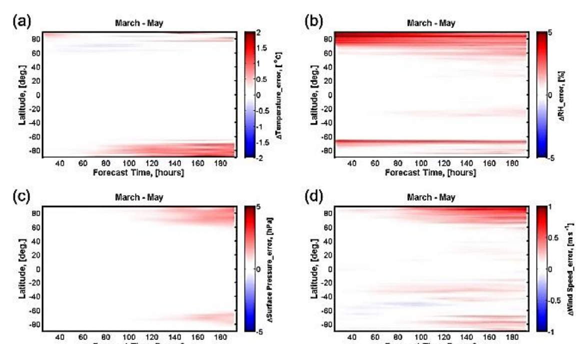

Figure 4. Forecasts in March-May.

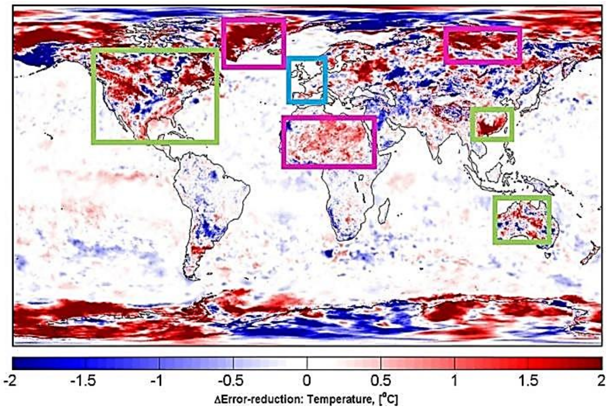

Figure 5. Difference in error reduction (March to May)-Feb.

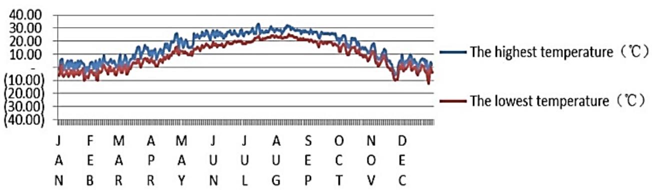

Figure 6. Weather forecast over the last 5 years.

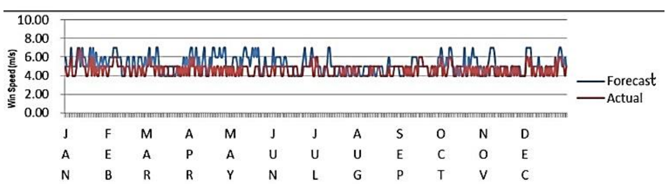

Figure 7. Actual vs predicted comparison.

5. Impacts in different region

We notice that the degradation of the climate estimate has a significant impact in the northern hemisphere compared to that in the southern hemisphere. This is on the grounds that there are a lot of aircraft observations in this region to constrain the initial conditions of the forecast model. Figure 4 shows the 168-hour forecast of the error reduction in temperature. North America, southeast China, and Australia are areas with an enormous number of aircraft observations under typical conditions (Ota et al., 2013; Petersen, 2016).

Western Europe likewise has a lot of aircraft observations, which decreased significantly during the COVID-19 epidemic with lockdown over most of the European nations. Almost no effect on a superficial level temperature estimate is noticed. This is on the grounds that there is a dense network of meteorological stations over western Europe contrasting to different locales (WMO), giving a decent limitation on the underlying conditions of the estimated model and thus a solid forecast. This mitigates the influence of the COVID-19 outbreak on weather predictions in European countries. The COVID-19 outbreak and its effects on weather forecasting in different countries has also been predicted using different approaches like LSTM (Kafieh et al., 2021), linear machine learning models, random forest, gradient boosting methods, support vector machines, KNN and decision trees (Malki et al., 2020), bibliometric analysis (Abd-Alrazaq et al., 2021)

6. Conclusion

Weather predictions are critical in everyday life, agriculture and industrial operations, and forecasting precise outcomes is difficult. Of all the parameters, the temperature plays a significant role in accurate prediction of the climatic conditions. The COVID-19 epidemic triggered a nationwide lockdown, drastically reducing airplane transportation services and aircraft related monitoring from March to May 2020. The predicted accuracy from March to May 2020 differs from that of the average accuracy from March to May 2015 – 2019. This result in the forecast deterioration is the result of global pandemic which further can impact the long-term weather forecasts in the coming future.

Conflict of interest

The authors declare that they have no competing interests.

Acknowledgments

We must express our appreciation to the Management of the Institutions for their encouragement to carry out research work.

Funding

No funding was received from any of the organizations.

References

Abd-Alrazaq, A., Schneider, J., Mifsud, B., Alam, T., Househ, M., Hamdi, M., & Shah, Z. (2021). A comprehensive overview of the COVID-19 literature: machine learning–based bibliometric analysis. *Journal of medical Internet research*, *23*(3), e23703. https://doi.org/10.2196/23703

Burgers, G., Van Leeuwen, P. J., & Evensen, G. (1998). Analysis scheme in the ensemble Kalman filter. *Monthly weather review*, *126*(6), 1719-1724. https://doi.org/10.1175/15200493(1998)126<1719:ASITEK>2.0.CO;2

Dey, R., & Chakraborty, S. (2015, October). Convex-hull & DBSCAN clustering to predict future weather. In *2015 International Conference and Workshop on Computing and Communication (IEMCON)* (pp. 1-8). IEEE. https://doi.org/10.1109/IEMCON.2015.7344438

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics*, *53*, 343-367. https://doi.org/10.1007/s10236-003-0036-9

Chen, Y. (2020). COVID‐19 pandemic imperils weather forecast. *Geophysical research letters*, *47*(15), e2020GL088613. https://doi.org/10.1029/2020GL088613

Chen, S. M., & Hwang, J. R. (2000). Temperature prediction using fuzzy time series. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *30*(2), 263-275. https://doi.org/10.1109/3477.836375

Chow, T. W., & Leung, C. T. (1996). Neural network based short-term load forecasting using weather compensation. *IEEE transactions on power systems*, *11*(4), 1736-1742. https://doi.org/10.1109/59.544636

Cofiño, A. S., Cano, R., Sordo, C., & Gutierrez, J. M. (2002). Bayesian networks for probabilistic weather prediction. In *15th Eureopean Conference on Artificial Intelligence (ECAI)*.

Harvard T.H. Chan School of Public Health. (2020). *Coronavirus and climate change*. C-CHANGE | https://www.hsph.harvard.edu/cchange/subtopics/coronavirus-and-climate-change/

Houthuys, L., Karevan, Z., & Suykens, J. A. (2017). Multi-view LS-SVM regression for black-box temperature prediction in weather forecasting. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 1102-1108). IEEE. https://doi.org/10.1109/IJCNN.2017.7965975

James, E. P., & Benjamin, S. G. (2017). Observation system experiments with the hourly updating Rapid Refresh model using GSI hybrid ensemble–variational data assimilation. *Monthly Weather Review*, *145*(8), 2897-2918. https://doi.org/10.1175/MWR-D-16-0398.1

Kafieh, R., Arian, R., Saeedizadeh, N., Amini, Z., Serej, N. D., Minaee, S., ... & Haghjooy Javanmard, S. (2021). COVID-19 in Iran: forecasting pandemic using deep learning. *Computational and mathematical methods in medicine*, *2021*. https://doi.org/10.1155/2021/6927985

Kwong, K. M., Liu, J. N., Chan, P. W., & Lee, R. (2008). Using LIDAR Doppler velocity data and chaotic oscillatory-based neural network for the forecast of meso-scale wind field. In *2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)* (pp. 2012-2019). IEEE. https://doi.org/10.1109/CEC.2008.4631064

Le Quéré, C., Jackson, R. B., Jones, M. W., Smith, A. J., Abernethy, S., Andrew, R. M., ... & Peters, G. P. (2020). Temporary reduction in daily global CO2 emissions during the COVID-19 forced confinement. *Nature climate change*, *10*(7), 647-653. https://doi.org/10.1038/s41558-020-0797-x

Lai, L. L., Braun, H., Zhang, Q. P., Wu, Q., Ma, Y. N., Sun, W. C., & Yang, L. (2004, August). Intelligent weather forecast. In *Proceedings of 2004 International Conference on Machine Learning and Cybernetics (IEEE Cat. No. 04EX826)* (Vol. 7, pp. 4216-4221). IEEE. https://doi.org/10.1109/ICMLC.2004.1384579

Maqsood, I., Khan, M. R., & Abraham, A. (2004). An ensemble of neural networks for weather forecasting. *Neural Computing & Applications*, *13*, 112-122. https://doi.org/10.1007/s00521-004-0413-4

Malki, Z., Atlam, E. S., Hassanien, A. E., Dagnew, G., Elhosseini, M. A., & Gad, I. (2020). Association between weather data and COVID-19 pandemic predicting mortality rate: Machine learning approaches. *Chaos, Solitons & Fractals*, *138*, 110137. https://doi.org/10.1016/j.chaos.2020.110137

Matlab (R2021). The Math Works Inc., Product Help- Neural Network Toolbox.

NCEP (2018). The development and success of NCEP's global forecast system.

Ota, Y., Derber, J. C., Kalnay, E., & Miyoshi, T. (2013). Ensemble-based observation impact estimates using the NCEP GFS. *Tellus A: Dynamic Meteorology and Oceanography*, *65*(1), 20038. https://doi.org/10.3402/tellusa.v65i0.20038

Petersen, R. A. (2016). On the impact and benefits of AMDAR observations in operational forecasting—Part I: A review of the impact of automated aircraft wind and temperature reports. *Bulletin of the American Meteorological Society*, *97*(4), 585-602. https://doi.org/10.1175/BAMS-D-14-00055.1

Saima, H., Jaafar, J., Belhaouari, S., & Jillani, T. A. (2011). Intelligent methods for weather forecasting: A review. In *2011 National Postgraduate Conference* (pp. 1-6). IEEE. https://doi.org/10.1109/NatPC.2011.6136289

Salman, A. G., Kanigoro, B., & Heryadi, Y. (2015). Weather forecasting using deep learning techniques. In *2015 international conference on advanced computer science and information systems (ICACSIS)* (pp. 281-285). Ieee. https://doi.org/10.1109/ICACSIS.2015.7415154

Singh, N., Chaturvedi, S., & Akhter, S. (2019). Weather forecasting using machine learning algorithm. In *2019 International Conference on Signal Processing and Communication (ICSC)* (pp. 171-174). IEEE. https://doi.org/10.1109/ICSC45622.2019.8938211

Sela, J. G. (2009). The implementation of the sigma pressure hybrid coordinate into the GFS. Office note (National Centers for Environmental Prediction (U.S.));461. https://repository.library.noaa.gov/view/noaa/11401

Weather Research & Forecasting Model (WRF) | Mesoscale & Microscale Meteorology Laboratory. (n.d.). NCAR. https://www.mmm.ucar.edu/models/wrf

Yonekura, K., Hattori, H., & Suzuki, T. (2018, December). Short-term local weather forecast using dense weather station by deep neural network. In *2018 IEEE international conference on big data (big data)* (pp. 1683-1690). IEEE. https://doi.org/10.1109/BigData.2018.8622195