https://doi.org/10.52131/jcsit.2021.0201.0006

*i*RASD Journal of Computer Science and Information Technology

Volume 2, Number 1, 2021, Pages 01 - 12

Journal Homepage:
https://journals.internationalrasd.org/index.php/jcsit

iRASD
JOURNAL OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY
INTERNATIONAL RESEARCH ASSOCIATION FOR SUSTAINABLE DEVELOPMENT

Climate Change Forecasting Using Machine Learning SARIMA Model

Shanza Zia[1]

[1] MS Scholar, Department of Computer Science, The Islamia University of Bahawalpur, Pakistan.
Email: shanzazia45@gmail.com

ARTICLE INFO

Article History:

Received: August 14, 2021
Revised: October 25, 2021
Accepted: December 29, 2021
Available Online: December 31, 2021

Keywords:

Climate Change
Forecasting
Machine learning
SARIMA Model
Artificial Intelligence
Time series

ABSTRACT

Every country's population will have to deal with the effects of climate change. The meteorological department needs to implement effective forecasting methods to deal with climate changes. Accurate temperature forecasts help in protecting people and property is an essential aspect of government, business, and the general public planning. Early predictions help farmers and industrialists to make approaches and store crops more effectively. When the climate continuously changes, it is not easy to make accurate predictions for the meteorological department and government authorities. Artificial intelligence (AI) algorithms have stimulated improvements in various fields. Machine learning (ML) may find teleconnections where complicated feedbacks make it challenging to determine how proposed work from a straightforward analysis and observations. Our proposed research uses the machine learning algorithm, SARIMA Model, to comprehend and utilize existing datasets and simulations.

Corresponding Author's Email: shanzazia45@gmail.com

1. Introduction

Climate change is one of the most severe challenges to modern society. It can have an impact on the economy, community, or environment. Due to the region's dry climate, poor educational level, and individual economic restraints, the sandy portion of the area is hostile to flourish. Irrigation and residential water shortages are two of the biggest issues, and there's also a significant risk of desertification (Fang & Lahdelma, 2016). There is a need to find a technical solution to these issues if need them to deteriorate.

As the climate changes, discuss how AI is affecting environmental protection. The article continues: AI can help us better understand how climate change affects biological systems in the following section. To better understand how climate change affects biological systems, examine AI's involvement in data collection and classification, decision-making, and enforcement of management policies (Luo, Zhou, & Wei, 2013). Environmental governance is a common thread that connects all of these issues. Among the problems raised in this research work is how AI technologies can be utilized ethically and alter control associations, as demonstrated in Figure 1.2 - SARIMA Model for forecast climate change. On the other side, new technology can improve the quality of people's lives. The artificial intelligence algorithms prominence in recent years, allowing us greater control over the world around us, as shown in Figure 1 Modulating influences for climate change. Machine learning approaches ' primary uses are predicted and control climate change and environmental processes (Mao, Zhang, Yan, Cheng, & health, 2018). These sensors are impeccable because of their ability to collect data on climate change's ecological effects.

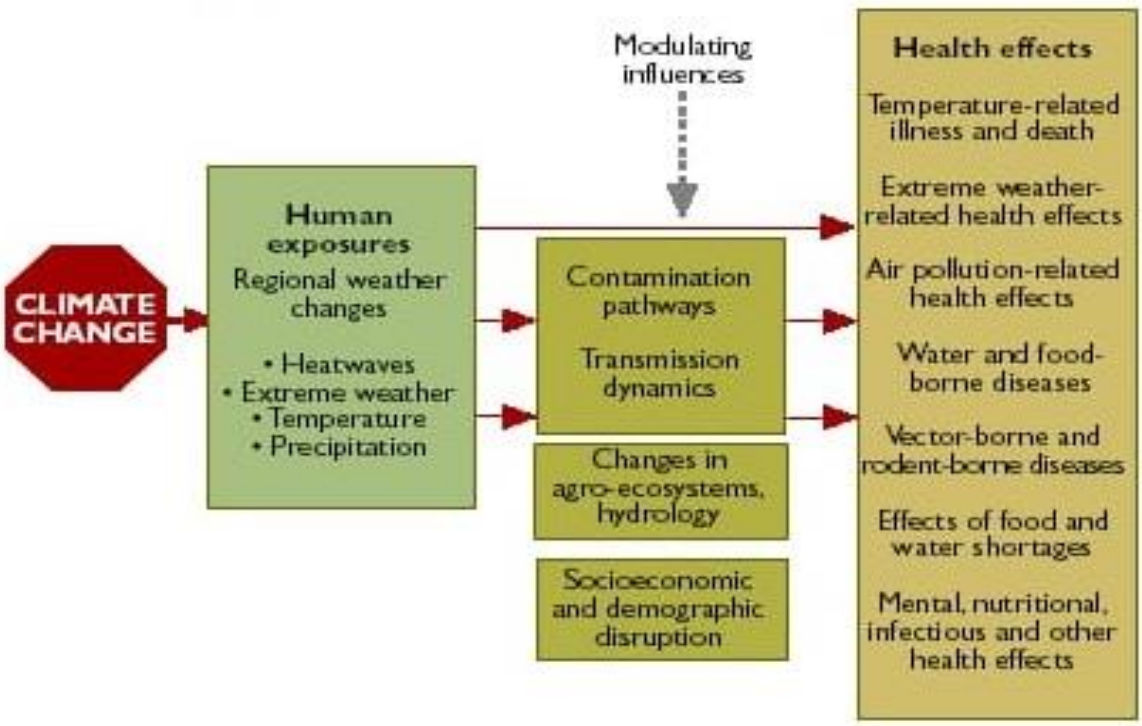

Figure 1: Modulating Influences for Climate Change

A lot has changed in the last few decades, however. There are no longer any forests, and animals and plants struggle to survive due to water. Not only that but also, the quality of water supplies and notable groundwater declined during the previous 30 years (Riley, Ben-Nun, Turtle, Bacon, & Riley, 2020). These consequences are irreversible and can't be reversed for the most part. The loss of soil, the inability to replace water, and the destruction of trees are examples, as displayed in Table 1.

Table 1
Surface Coverage Attributes Values (adapted oke, 2006)

Surface coverage	Attributes	Albedo (reflection)
Soil	Dark & Wet	0.05
	Light & Dry	0.40
Sand	Long	0.15 – 0.48
Agriculture	Long	0.16
Grass	Short	0.26
Clouds	Deciduous	0.15 – 0.25
	Coniferous	0.03 – 0.10
Snow	Old	0.40
	Fresh	0.95
Water	Small Zenith Angle (<45°)	0.10 – 0.15
	Large Zenith Angle (>45°)	0.25 – 0.13

In general, desertification harms a region's capacity to produce food. It's been a considerable shift in the last couple of decades. There are no longer any forests, and animals and plants struggle to survive due to water (Rusyana & Flancia, 2016). Over the past three decades, the quality of water resources, particularly groundwater, has deteriorated significantly. Irreversible and unchangeable: It is impossible to recover water resources and forest areas that have been lost due to soil degradation.

Technology can assist farmers in coping with the impact of a changing climate by providing real-time data on where and how sand moves across specific regions. Some of the subjects covered in this course include environmental education, water conservation, and the responsibilities of provincial governments to stop deforestation (Etuk & Modeling, 2013).

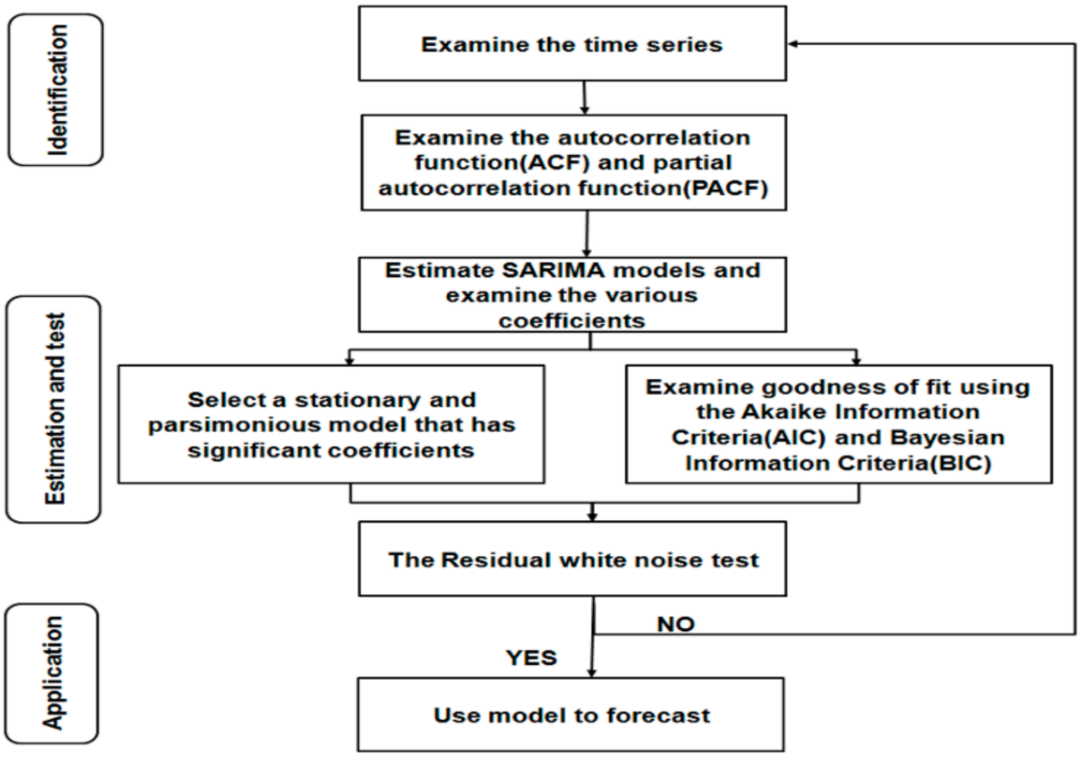

Figure 2: SARIMA Model for Forecast Climate Change

2. Objective and Motivation

Integrated urban planning approaches and consideration of climate factors are required for early prediction. Consider how the climate will change and how you can adapt to it. An all-encompassing AI-based approach to looking at and planning for changes is what this initiative is aiming to achieve (Dikshit, Pradhan, & Alamri, 2020; Mombeni, Rezaei, Nadarajah, Emami, & Assessment, 2013).

Our city's climate is continuously changing, and artificial intelligence helps us keep track of these variations. There will be models for the extension that can work, scale, and grow. In light of these new findings, it is also feasible to better understand how intertwined planning methodologies and environmental concerns are. The results can also be helpful. It's crucial to pay attention to the following characteristics (thermal and dynamic). A database including information on topography, land use, and climate is needed to help in analyzing climate factors

- Ability to store heat in the form of mass
- The sun's rays travel around the world.
- Various ways in which air can flow
- Nighttime cold air generation and movement
- Discuss the current climate change tendencies (temperature, precipitation, and number of summer days)

3. Related Work

Rolnick et al. (2022) demonstrated in their research that Machine learning is exploited to reduce greenhouse gas emissions and help people cope with a rapidly changing climate. Essential to explore to investigate are exploration queries and commercial prospects. Machine learning will contribute to combating climate change around the globe.

Cifuentes, Marulanda, Bello, and Reneses (2020) presented that change has influenced the planet and its inhabitants over the last decade. Predicted changes in air temperature have played a significant role in climate change research. Agriculture, ecology, the environment, and industry are examples of this. According to this survey, an accurate temperature prediction can be made using Machine Learning techniques. Features included the previous temperature and humidity levels and other weather data such as solar radiation and wind speed. The review indicated that for one step forward at a regional scale for one measure, Deep Learning techniques made more minor mistakes (Mean Square Error = 0.0017 °K) than typical Artificial Neural Networks architectures. Support Vector Machines (SVMs) are the most often used worldwide because are accurate and straightforward.

Veenadhari, Misra, and Singh (2014) suggested that crops in India have been adversely affected by climate change over the past two decades regarding performance. Businesses in the same field can use these forecasts to organize their supply chains better. Crop yields have been predicted and modeled in various methods, but none takes the weather into account and is primarily dependent on conjecture. Software dubbed "Crop Advisor" has been developed in this study to assist farmers in determining how climate influences crop production. Farmers in Madhya Pradesh use an algorithm known as C4.5 to determine which climatic element has the most significant impact on agricultural yields across the state.

Temperature such as tree rings can construct time series profiles to illuminate chronological climates by (Abbot & Marohasy, 2017). An artificial neural network (ANN) trained using the sine waves from the six datasets subjected to signal analysis. There is a good match in temperature profiles between the late Holocene times and 1830 CE by altering the original shapes' amplitude, frequency, and phase of sine waves. After that, the ANN models were used to forecast temperature changes over the twentieth century. In six distinct locations, the ANN predictions differed by an average of 0.2 degrees Celsius from the actual temperatures. There was an Equilibrium Climate Sensitivity (ECS) of around 0.06°C as a result of this. To put it another way, this is a lot less than the IPCC's general circulation models suggest (Hong et al., 2020).

Anaraki, Farzin, Mousavi, and Karami (2021) illustrated that the MARS model tree outperformed the M5 model tree in the tests. ANN and LSSVM WOA are other machine learning techniques in downscaling. In Discharge Simulations, the LSSVM WOA WT method outperforms the LSSVM WOA WT algorithm (NSE = 0.911). Observing all of the scenarios except CanESM2 RCP2.6, the 200-year release is expected to decrease shortly. In the short and long run, it's evident that hydrological models are a significant source of uncertainty when looking at ANOVA analysis of uncertainty.

Ardabili, Mosavi, Dehghani, and Várkonyi-Kóczy (2019) analyze and anticipate hydrological processes, climate change, and Earth's systems more accurately. Several strategies improve accuracy, resilience, efficiency, computing cost, and overall model performance. It also discusses the present state of affairs and possible future developments. The research into deep understanding is still ongoing, according to the article. On the other hand, these methods are already for machine learning. Ensemble and hybrid strategies are to generate newer, more effective ways. Constructions interpretation for over half of total energy use and a third of global greenhouse gas emissions in developed countries like the United States. Several factors contribute to the amount of energy a structure consumes, including physically constructed and used. In addition to physics-based building energy modeling, machine learning techniques can produce faster and more accurate estimations based on the number of previously used energy buildings. City and community managers can predict how facilities utilize energy to enhance their future energy needs plans.

Fathi, Srinivasan, Fenner, Fathi, and Reviews (2020) demonstrated that Machine learning and future weather scenarios don't examine how much space there is for various urban buildings to climate change. There is no single method for comparing items worldwide to get the most accurate machine learning-based forecast. The peculiarities of predicting difficulties have a significant impact on accuracy levels. The purpose of this study is to highlight how the use of machine learning in urban building energy performance forecasts has evolved over the course of three years. However, artificial intelligence models can be used to address this issue. Physical factors can arise as features in these models to predict and evaluate irrigation water quality indexes in aquifer systems; data included Adaptive Boosting, Random Forest, Artificial Neural Network, and Support Vector Regression (SVR) models. Results reveal that Adaboost and RF models outperform SVR and ANN when predicting the future. Although these studies claim that ANN and SVR models are more tolerant of input factors than Adaboost or RF, these models are more sensitive to input variables than the other models (Yuan, San, & Leong, 2020). The models developed around the world are excellent in predicting irrigation water quality. Farmers and other stakeholders could benefit from this information. Using biological data as input variables, the approaches in this study effectively predicted groundwater quality at low costs and in real-time.

Damage to agricultural and water resources can result from droughts, costing much money. An essential component of drought management is to develop techniques that can predict lack events, which might implement mitigating measures. Climate change appears to be the primary cause of droughts around the world. The Australian state of New South Wales (NSW) has experienced numerous shortages in recent years. A Climate Research Unit (CRU) dataset was employed at various periods (one, three, six, and 12 months) to calculate the drought index by (El Bilali, Taleb, & Brouziyne, 2021). Eight factors were climatic drivers and sea surface temperature indicators, while the other three were various meteorological variables used to estimate when the lack index would be at its top.

Predictions were made using an artificial neural network (ANN) and support vector regression (SVR) (SVR) by (Dikshit et al., 2020). Model performance was evaluated using the coefficient of determination (COD), root means square error (RMSE), and represent absolute error (MAE). After training from 1901 to 2010, the model went through its paces for nine years of testing. (MAE). The results show that the ANN technique outperformed the SVR method to predict long-term drought patterns. In terms of R2 values, the ANN approach had the highest at 0.86, while the SVR method had a value of 0.75. Sea surface temperatures and the Pacific Decadal Oscillation (PDO) do not significantly impact the temporal dry period. As a preliminary phase, this study looks at how droughts in the New South Wales region are affected by climatological variables and patterns.

Rasel, Sultana, and Meesad (2017) proposed in their research that when the weather and climate continuously change, it's difficult to make accurate forecasts for the coming days' weather and climate. Humidity, wind speed, sea level, and air density are just a few of the many variables that play a role in climate change. The Bangladesh Meteorological Department provided a six-year rainfall and temperature data archive for the Chittagong metropolitan region to conduct the tests (BMD). There are several businesses where a sound forecasting system may be quite beneficial. Industries, agriculture, tourism, transportation, and construction are all included in this list. Support Vector Regression (SVR) and Artificial Neural Networks can produce more accurate weather predictions (Shen, Valagolam, & McCalla, 2020; Yildiz, Bilbao, Sproul, & Reviews, 2017). According to this research, the SVR is better at predicting rain than the ANN, and the ANN is better than the SVR (Derbentsev, Datsenko, Stepanenko, & Bezkorovainyi, 2019).

4. Information About Artificial Intelligence (AI) and Climate Change Adaptation Techniques

There has been an evolution in the understanding of climate change over time. Policymakers have access to an overabundance of data on climate change. Climate change necessitated an increase in the availability and standardization of local weather data worldwide and the standardization of historical data used to make sense of global climate. Compiling data across time and space and between physical systems and various data formats can be done using computer models. That can also be simulated and utilized to predict the future climate (Rizwan, Raj, & Vasudev, 2017). Recent developments in computing are transforming climate science, and policymakers reviewed what to do about it (Duerr et al., 2018; Rizwan et al., 2017). To commence, more data and greater processing power are now available for analysis.

Several novel machine learning (AI) approaches have been implemented in the current years. There have been new understandings into the intricate interconnections of the climate change prediction system, resulting in more realistic climate models. Extreme weather forecasts have acquired better as a result.

5. Material and Methods

5.1. Data Collection

The quality of the magnitude of the climate change dataset could have influenced the data acquired in each occurrence. The National Center for Hydro-Meteorological Forecasting (NCHMF), a government institution, meteorological departments, and climate websites dataset utilized. Prevent variances by climate change, regions in every administrative area considered for presence.

5.2. SARIMA Model (p, d, q) (P, D, Q, S)

SARIMA stands for Seasonal Auto-Regressive Integrated Moving Average. Figure 3 Proposed approach for climate change using Machine Learning represents the methodology for the anticipated method.

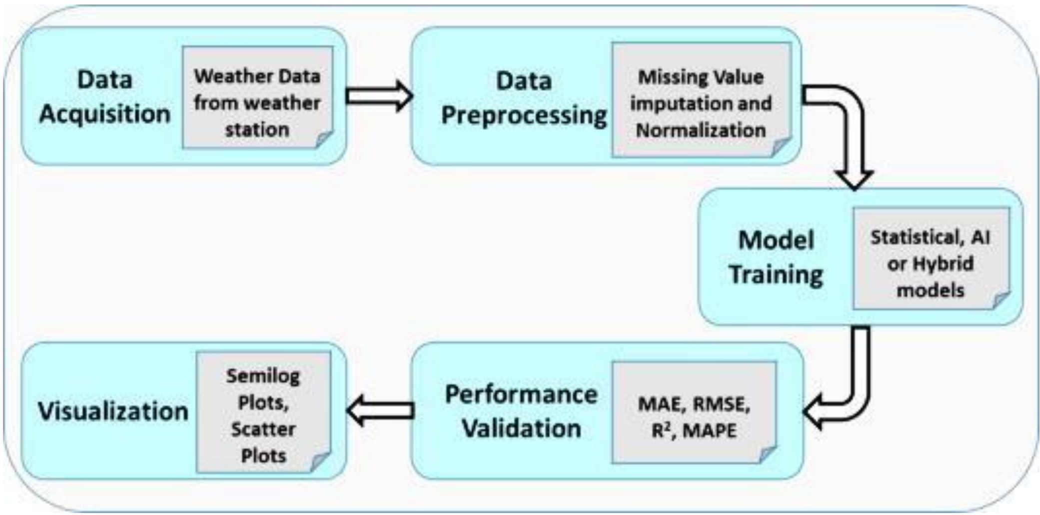

Figure 3: Proposed approaches for climate change using Machine Learning

5.3. Non-seasonal ARIMA

Split the Arima term into three terms, AR, I, MA: AR (p) stands for the autoregressive model; the p parameter is an integer that confirms how many lagged series is to be used to forecast periods, for example, The average temperature of yesterday correlates with today's temperature, so use AR(1) parameter to predict future temperatures. The formula for the AR (p) model is: y^t=μ+θ1Yt−1+...+θp Yt−py^t = μ+θ1Yt−1+...+θpYt−p Where μμ is the constant term, the p is the periods to be used in the regression, and θ is the parameter fitted to the data.

I (d) the differencing part and the d parameter show how many differencing orders will be used. It tries to make the series stationary, for example: If d = 1: yt=Yt−Yt−1yt= Yt− Yt – 1 where type is the differenced series, and Yt−periodic−period is the original series If d = 2: yt = (Yt−Yt−1)−(Yt−1−Yt−2)= Yt−2Yt−1+Yt−2yt = (Yt−Yt−1) − (Yt−1−Yt−2) = Yt – 2 Yt−1+Yt−2 Note that the second difference is a change-in-change, which measures the local "acceleration" rather than the trend.

MA (q) stands for *moving average model*, the q is the number of lagged forecast errors terms in the prediction equation, example: It's strange, but this MA term takes a percentage of the errors between the predicted value against the real. It assumes that the past mistakes will be similar in future events. The formula for the MA (p) model is: y^t= μ−Θ1et−1+ ...+Θqet−qy^t =μ−Θ1et−1+...+Θqet−q Where μμ is the constant term, q is the period to be used on the ee term, and ΘΘ is the parameter fitted to the errors

$$ET = Yt-1-y\text{^}t-1e\ t = Yt-1-y\text{^}t-1 \quad (1)$$

5.3. Seasonal ARIMA

The p, d, q parameters differ from the non-seasonal parameters. SAR (P) is the seasonal autoregression of the series. The formula for the SAR(P) model is y^t=μ+θ1Yt−sy^t= μ+θ1 Yt − s Where P is the number of autoregression terms to added, usually, no more than one term, s is how many periods ago to be used as the base, and θθ is the parameter fitted to the data.

Usually, when the subject is weather forecasting, 12 months ago, have some information to contribute to the current period. Setting P=1 (i.e., SAR (1)) adds a multiple of Yt−sYt−s to the forecast for yet Figure 4 represents Temperature Variation in Rio de Janeiro I(D) the seasonal difference MUST be used when you have a solid and stable pattern. If d = 0 and D = 1: yt=Yt−Yt−syt=Yt−Yt−s, ytyt is the differenced series, and Yt−Yt−s is the original seasonal interval. If d =1 and D = 1: yt=(Yt−Yt−1)−(Yt−s−Yt−s−1)= Yt−Yt−1− Yt− s+ Yt −s − 1yt =(Yt−Yt−1)−(Yt−s−Yt−s−1) = Yt−Yt−1−Yt−s+Yt−s−1

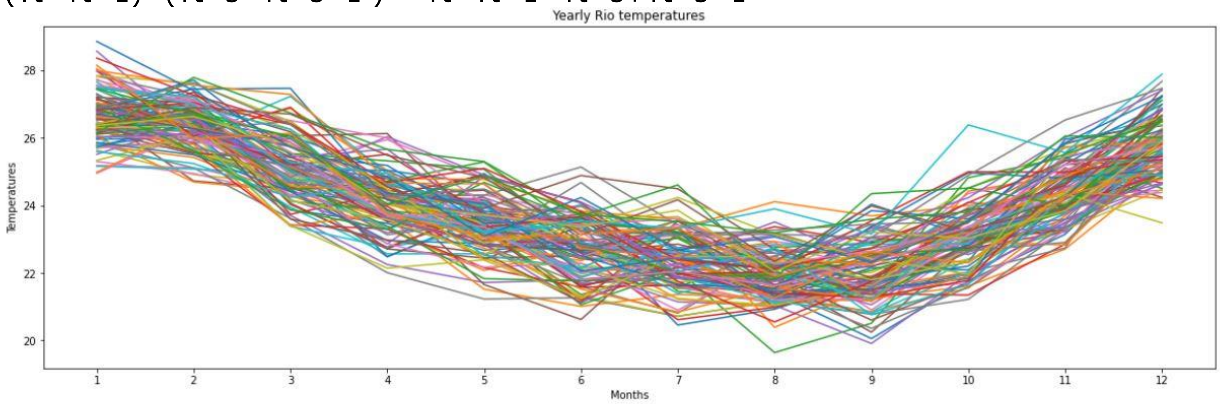

Figure 4: Temperature Variation in Rio de Janeiro

The highest temperatures occur in November and February, while the lowest temperatures occur from July through September. Averaging the monthly levels of each of these lines, create a single line Figure 5 represents Monthly Variation in Rio de Janeiro

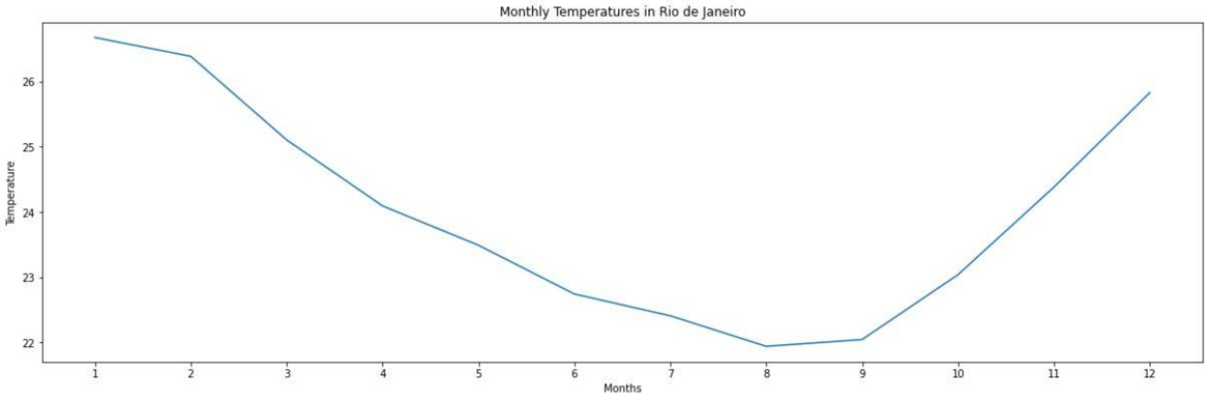

Figure 5: Monthly Variation in Rio de Janeiro

Here are some stats from this series to see a pattern over the years. The average temperature rose by 4.25 percent for more than a century, rising from 23.5 to 24.5 degrees Fahrenheit. First, divide the data into training, validation, and test sets to see how it all works together. Step ahead Figure 6 displayed Yearly Variation in Rio de Janeiro: a confirmation forward for 48 months, then extrapolate the future for another year to compare to the test set.

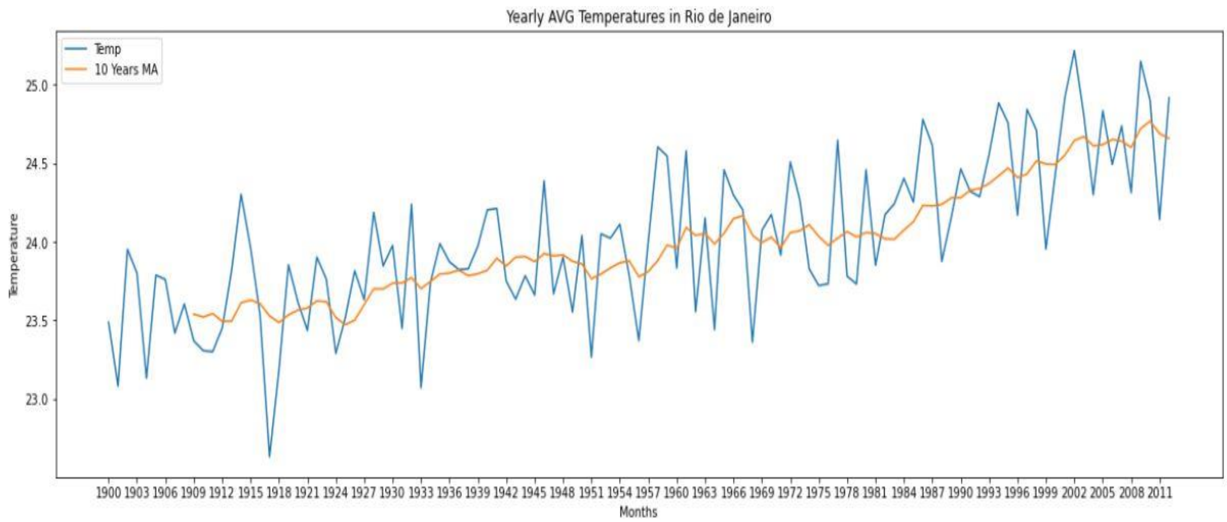

Figure 6: Yearly Variation in Rio de Janeiro

The baseline's RMSE temperature is 1.3282 Celsius. An upward trend is displayed in the data, and seasonality appears, with higher temperatures at the commencement and end of each season and lower ones in the middle. To create a time series forecast (constant mean, variance, and autocorrelation). The occupied function can determine if the series is static. It is safe to build your model if the series has less than 5% P-Value. While there are many ways to manipulate the data, there are many more if the string isn't static. This graph shows that the entire series Autocorrelation function (ACF) charts the evolution of a set of data across time.

There's a graph here showing how recently this temperature was this high. Figure 7 demonstrates Autocorrelation, Partial Correlation, and Distributed chart. The partial autocorrelation function, or PACF, is used to describe this. Without accounting for the impact of prior lags, it displays the correlation between the current temperature and the covered type. For example, in the case of temperature, it only indicates the impact of lag three without considering the effect of intervals 1 and 2.

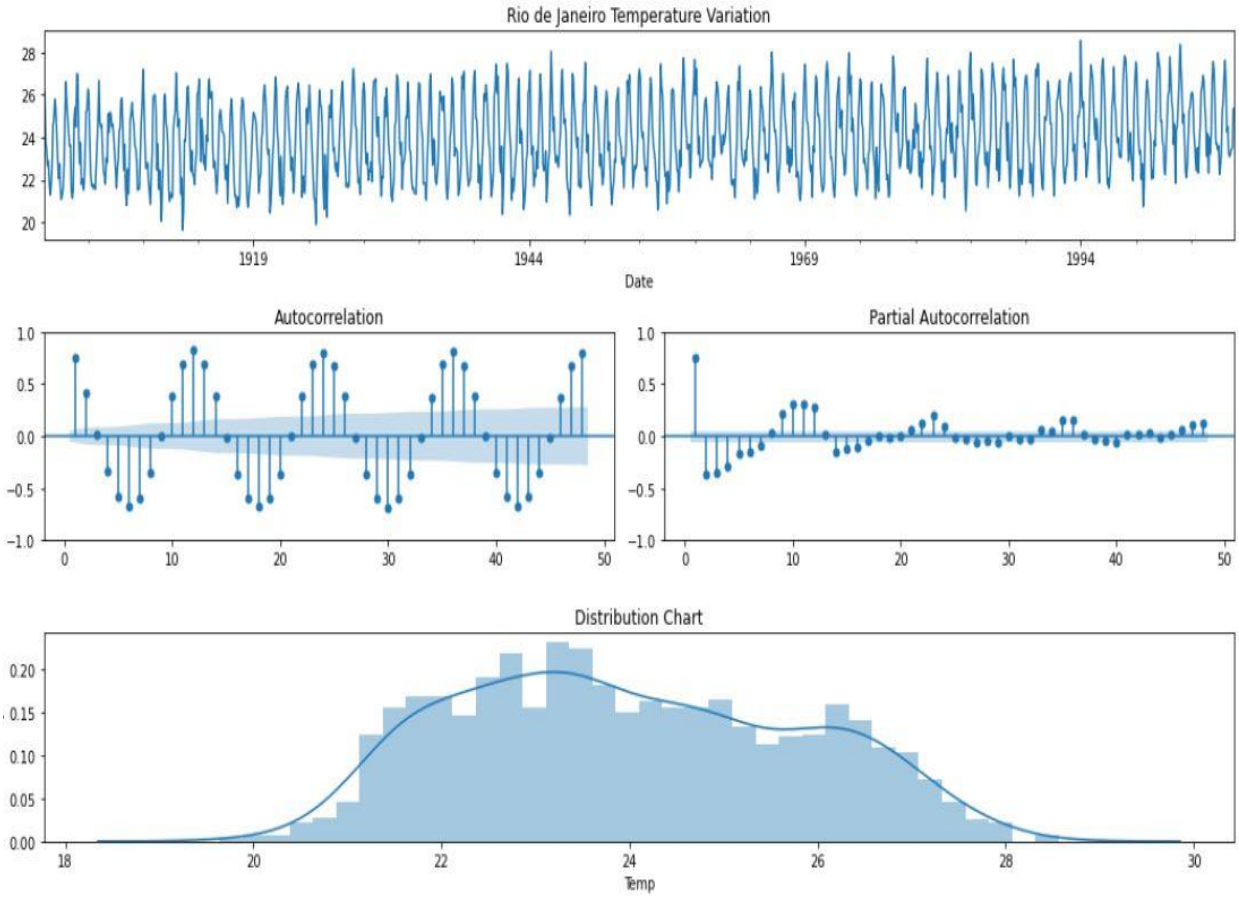

Figure 7: Autocorrelation, Partial Correlation, and Distributed chart

Expect a lower error level in our simulation than this one. Next month will use the previous month's forecast as a starting point. A negative autocorrelation begins at lag six and occurs once every 12 months. Different seasons have a role to play in this phenomenon. There is a negative autocorrelation if it is winter now and milder in six months. Figure 8 demonstrates Autocorrelation, Partial Correlation, and Distributed chart. Two temperatures usually change in different directions.

As a result, a solid positive autocorrelation is evident, starting at lag 12 and continuing for a further 12 lags. Late intervals exhibit a negative PACF. Initially, the PACF shows a positive jump and then declines to a negative PACF. The ACF and PACF charts are identical in this situation. An AR (1) model and a first seasonal difference can be contingent on this (YtYt12). SAR (P) or SMA (Q) parameters may be required; therefore, plot the static function again with the first seasonal difference to confirm.

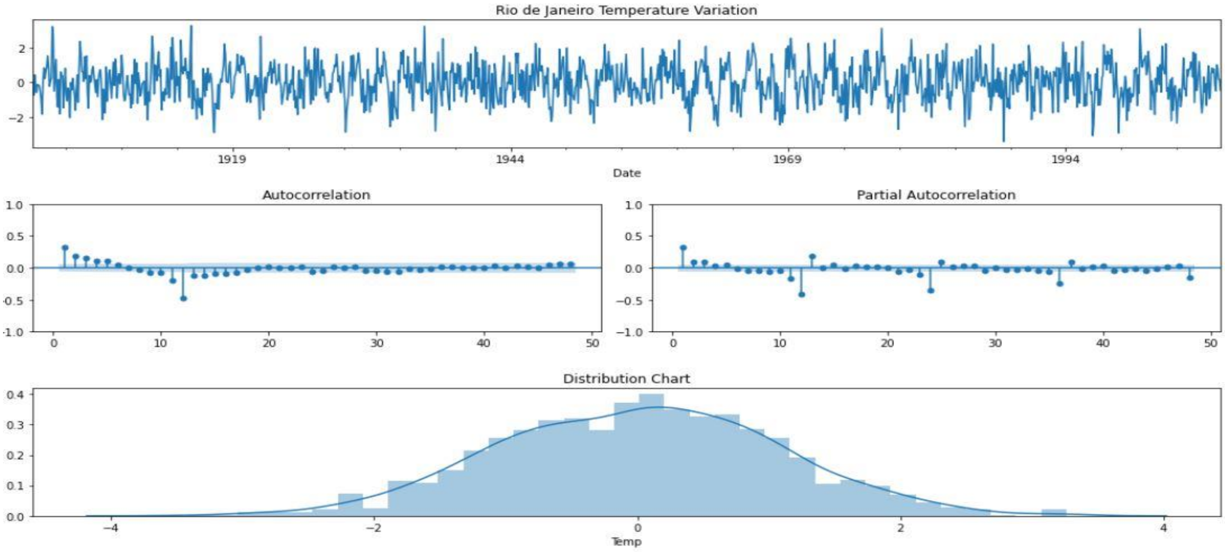

Figure 8: Autocorrelation, Partial Correlation and Distributed chart climate change

These are AR (3) models, defined as having three parameters. The figures above show that the first ACF lags are gradually decreasing. The PACF falls below the confidence interval after the third interval. Figure 9 shows the Current, predicted, and error values. The ACF and PACF both showed significant declines by the 12th interval. Put another way, this is a SAR (1) with a first difference because the SMA signature includes a parameter of 1 interval. Orders 3 and 0 in the model; orders 0 and 1 because the series has a clear uptrend, and orders 1 and 2 because the series has a clear downtrend. First, create a function that uses one-step forecasts from the entire validation set to determine how far off it is.

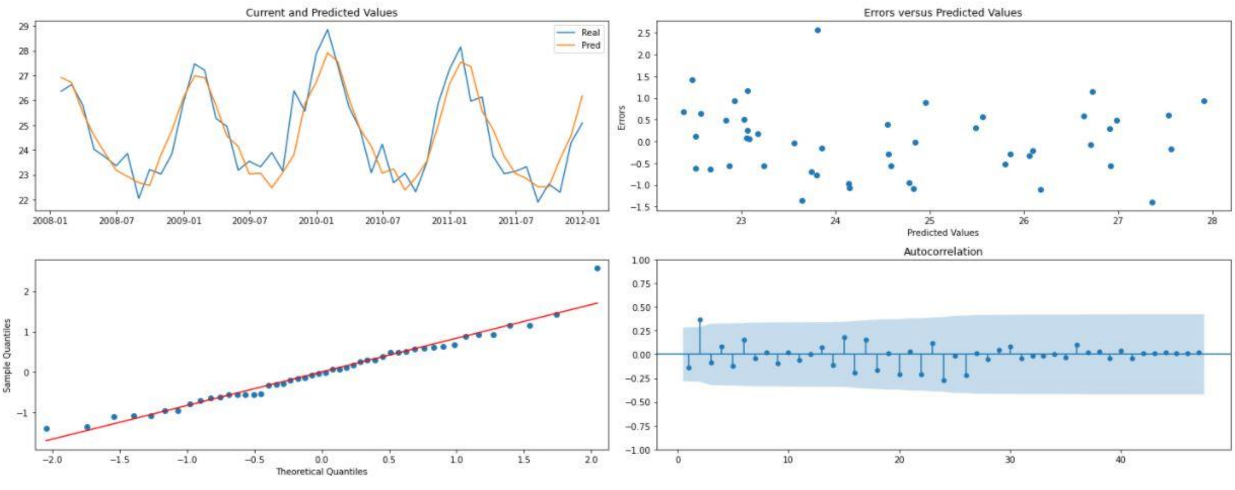

Figure 9: Current predicted and error values

The residuals will be easier to see with the help of a function. The graphs below show historical data as well as projections for the future. An example of a scatter plot presents the difference between expected and actual values. An autocorrelation plot of the residuals can be used to determine whether or not there is still some association.

This page has graphs at the top that the forecasts match the present values pretty well. The Error vs. Predicted values has a line in them. The mistakes increase from -1.5 to +1.5 as the temperature rises, as represented in Figure 5.8 current values compared to the Extrapolated ones. However, an autocorrelation plot indicates a positive spike just above lag 2 in the confidence interval across some outliers exhibited in the QQ Plot. The field doesn't need any more tweaks, in my opinion. To assess how accurate the test set's prediction it's time to look back at 12 months.

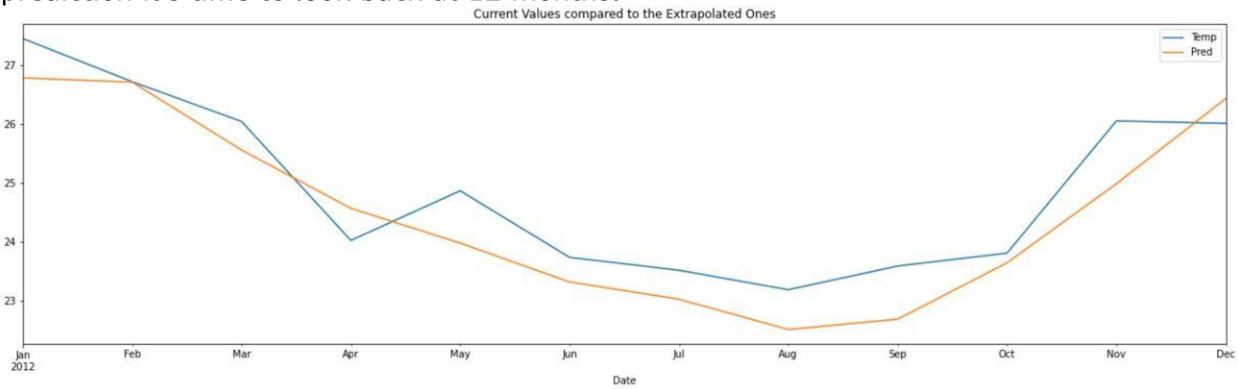

Figure 10: Current values compared to the Extrapolated ones

The SARIMA parameters appear to be well-fitted from this result. These are the predicted values, and the seasonal pattern matches the actual values and the SARIMA parameters. With the test set (baseline vs. extrapolation) in place, RMSE measures the model's standard deviation when it emanates to testing.

6. Conclusion and Discussion

Real-world climate and weather changes are difficult to anticipate. As a result, climate and weather predictions are based on factors unique to each location and time, making it difficult to predict the future. According to this study, climate change and excessive sand migration harm some places and provinces around the globe. The study's main objective was to provide relevant data to the government and the general public about the current level of climate change. Design, applicability, efficiency, and economy are all terms.

There are three pieces to the system: a web application that lets you control it, sensors that collect data from each node, and linear regression in ML that lets you look at the data. Initially, a mechanical construction was used to obtain sensor data from a sensor node. A straightforward machine learning technique was proposed for the analysis to determine early prediction before climate changes. Multiple linear regressions and the SARIMA Model are used to examine various factors to improve and develop the system.

Efficient governments and farmers can both benefit from the research's outcomes. Proper estimates of climate changes save costs and increase output in agriculture in the future. These digital technology applications can predict how the climate will change. The project's success in the pilot stage allows more people to take advantage of its benefits. Using SARIMA Model and ML technology, this study could monitor and predict weather conditions. Reviewing current conditions and forecasting future climate change were two aspects of research. Extreme weather and climate change are linked using the scientific method.

This paper presents artificial intelligence and machine learning (AI/ML) to improve crop yields, human well-being, and economics. In addition to helping the government, this research could also benefit the public by giving them the right idea about their investment costs. Problems are overcome through collaborative efforts between the government and the general public.

References

Abbot, J., & Marohasy, J. J. G. (2017). The application of machine learning for evaluating anthropogenic versus natural climate change. *14*, 36-46. doi:10.1016/j.grj.2017.08.001

Anaraki, M. V., Farzin, S., Mousavi, S.-F., & Karami, H. J. W. R. M. (2021). Uncertainty analysis of climate change impacts on flood frequency by using hybrid machine learning methods. *35*(1), 199-223. doi:10.1007/s12040-014-0497-x

Ardabili, S., Mosavi, A., Dehghani, M., & Várkonyi-Kóczy, A. R. (2019). *Deep learning and machine learning in hydrological processes climate change and earth systems a systematic review.* Paper presented at the International conference on global research and education.

Cifuentes, J., Marulanda, G., Bello, A., & Reneses, J. J. E. (2020). Air temperature forecasting using machine learning techniques: a review. *13*(16), 4215. doi:10.3390/en13164215

Derbentsev, V., Datsenko, N., Stepanenko, O., & Bezkorovainyi, V. (2019). *Forecasting cryptocurrency prices time series using machine learning approach.* Paper presented at the SHS Web of Conferences.

Dikshit, A., Pradhan, B., & Alamri, A. M. J. A. (2020). Temporal hydrological drought index forecasting for New South Wales, Australia using machine learning approaches. *11*(6), 585. doi:10.3390/atmos11060585

Duerr, I., Merrill, H. R., Wang, C., Bai, R., Boyer, M., Dukes, M. D., . . . Software. (2018). Forecasting urban household water demand with statistical and machine learning methods using large space-time data: A Comparative study. *102*, 29-38. doi:10.1016/j.envsoft.2018.01.002

El Bilali, A., Taleb, A., & Brouziyne, Y. J. A. W. M. (2021). Groundwater quality forecasting using machine learning algorithms for irrigation purposes. *245*, 106625. doi:10.1016/j.agwat.2020.106625

Etuk, E. H. J. M. T., & Modeling. (2013). The fitting of a SARIMA model to monthly Naira-Euro Exchange Rates. *3*(1), 17-26.

Fang, T., & Lahdelma, R. J. A. e. (2016). Evaluation of a multiple linear regression model and SARIMA model in forecasting heat demand for district heating system. *179*, 544-552. doi:10.1016/j.apenergy.2016.06.133

Fathi, S., Srinivasan, R., Fenner, A., Fathi, S. J. R., & Reviews, S. E. (2020). Machine learning applications in urban building energy performance forecasting: A systematic review. *133*, 110287. doi:10.1016/j.rser.2020.110287

Hong, J., Lee, S., Bae, J. H., Lee, J., Park, W. J., Lee, D., . . . Lim, K. J. J. W. (2020). Development and evaluation of the combined machine learning models for the prediction of dam inflow. *12*(10), 2927. doi:10.3390/w12102927

Luo, C. S., Zhou, L. Y., & Wei, Q. F. (2013). *Application of SARIMA model in cucumber price forecast.* Paper presented at the Applied Mechanics and Materials.

Mao, Q., Zhang, K., Yan, W., Cheng, C. J. J. o. i., & health, p. (2018). Forecasting the incidence of tuberculosis in China using the seasonal auto-regressive integrated moving average (SARIMA) model. *11*(5), 707-712. doi:10.1016/j.jiph.2018.04.009

Mombeni, H. A., Rezaei, S., Nadarajah, S., Emami, M. J. E. M., & Assessment. (2013). Estimation of water demand in Iran based on SARIMA models. *18*(5), 559-565. doi:10.1007/s10666-013-9364-4

Rasel, R. I., Sultana, N., & Meesad, P. (2017). *An application of data mining and machine learning for weather forecasting.* Paper presented at the International conference on computing and information technology.

Riley, P., Ben-Nun, M., Turtle, J., Bacon, D., & Riley, S. J. m. (2020). Sarima forecasts of dengue incidence in brazil, mexico, singapore, sri lanka, and thailand: Model performance and the significance of reporting delays. doi:10.1101/2020.06.26.20141093

Rizwan, M., Raj, R. J. R., & Vasudev, M. (2017). *A novel approach for time series data forecasting based on ARIMA model for marine fishes.* Paper presented at the 2017 International Conference on Algorithms, Methodology, Models and Applications in Emerging Technologies (ICAMMAET).

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., . . . Waldman-Brown, A. J. A. C. S. (2022). Tackling climate change with machine learning. *55*(2), 1-96.

Rusyana, A., & Flancia, M. (2016). *SARIMA model for forecasting foreign tourists at the Kualanamu International Airport.* Paper presented at the 2016 12th International Conference on Mathematics, Statistics, and Their Applications (ICMSA).

Shen, J., Valagolam, D., & McCalla, S. J. P. (2020). Prophet forecasting model: a machine learning approach to predict the concentration of air pollutants (PM2. 5, PM10, O3, NO2, SO2, CO) in Seoul, South Korea. *8*, e9961. doi:10.7717/peerj.9961

Veenadhari, S., Misra, B., & Singh, C. (2014). *Machine learning approach for forecasting crop yield based on climatic parameters.* Paper presented at the 2014 International Conference on Computer Communication and Informatics.

Yildiz, B., Bilbao, J. I., Sproul, A. B. J. R., & Reviews, S. E. (2017). A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *73*, 1104-1122. doi:10.1016/j.rser.2017.02.023

Yuan, C. Z., San, W. W., & Leong, T. W. (2020). *Determining Optimal Lag Time Selection Function with Novel Machine Learning Strategies for Better Agricultural Commodity Prices Forecasting in Malaysia.* Paper presented at the Proceedings of the 2020 2nd International Conference on Information Technology and Computer Communications.