ISSN: 2618-6055 / 8, (2), 182 – 191, 2024
DOI: 10.62301/usmtd.1577839

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt: 8, Sayı: 2, (2024), 182 – 191
Araştırma Makalesi – Research Article

PREDICTING KONYA'S AIR TEMPERATURE: GENETIC PROGRAMMING, GRADIENT BOOSTING AND RANDOM FOREST APPROACHES

*Kazım KUMAŞ[1], Ali AKYÜZ[1,2]

[1]Burdur Mehmet Akif Ersoy Üniversitesi, Bucak Emin Gülmez Teknik Bilimler MYO, Burdur
[2]Burdur Mehmet Akif Ersoy Üniversitesi, Bucak Bilgisayar ve Bilişim Fakültesi, Burdur

(Geliş/Received: 01.11.2024, Kabul/Accepted: 02.12.2024, Yayınlanma/Published: 31.12.2024)

ABSTRACT

Average temperature prediction is important in many areas, such as climate change, agriculture, and energy management. It is also necessary for estimating energy demand, managing energy, and developing sustainable energy policies. In this study, using monthly average air temperature data between 1960-2017, temperature predictions were performed for Konya province using genetic programming, gradient boosting, and random forest techniques. The predicted average monthly temperature values between 2018-2021 were compared with the real values. Then, future predictions for the years 2022-2025 were also performed. Metrics such as R^2, RMSE, and MAE were used in model evaluations. R^2=0.9477, RMSE=1.950 and MAE=1.500 for the genetic programming model, R^2=0.9663, RMSE=1.564 and MAE=1.203 for the gradient boosting model, and R^2=0.9905, RMSE=0.833 and MAE=0.625 for the random forest model. The same algorithms gave good results for future prediction of the average air temperature between 2022 and 2025. In conclusion, the applied machine learning methods gave successful results in monthly average air temperature predictions for Konya province, and these findings show that machine learning techniques can be used effectively in air temperature prediction.

Keywords: Air temperature, Genetic programming, Gradient boosting, Random forest.

KONYA'NIN HAVA SICAKLIĞININ TAHMİN EDİLMESİ: GENETİK PROGRAMLAMA, GRADİENT BOOSTİNG VE RASTGELE ORMAN YAKLAŞIMLARI

ÖZ

Ortalama sıcaklık tahmini iklim değişikliği, tarımsal ve enerji yönetimi gibi birçok alanda önemlidir. Ayrıca, enerji talep tahminleri, enerji yönetimi ve sürdürülebilir enerji politikalarının geliştirilmesi için de gereklidir. Bu çalışmada, 1960-2017 yılları arasındaki aylık ortalama hava sıcaklığı verileri kullanılarak Konya ili için genetik programlama, gradient boosting ve random forest teknikleri ile sıcaklık tahminleri gerçekleştirilmiştir. 2018-2021 yılları arasındaki her ay için tahmin edilen ortalama sıcaklık değerleri gerçek değerlerle karşılaştırılmıştır. Ardından, 2022-2025 yılları için gelecek tahminleri de yapılmıştır. Model değerlendirmelerinde R^2, RMSE've MAE gibi metrikler kullanılmıştır. Genetik programlama modeli için R^2=0.9477, RMSE=1.950 ve MAE=1.5000, gradient boosting modeli için R^2=0.9663, RMSE=1.564 ve MAE=1.203, random forest modeli için ise R^2= 0.9905, RMSE=0.833 ve MAE=0.625 değerleri elde edilmiştir. 2022-2025 yılları arasındaki ortalama hava sıcaklığı içinde aynı algoritmalar gelecek tahmini iyi sonuçlar vermiştir. Sonuç olarak, uygulanan makine öğrenimi yöntemleri, Konya ili için aylık ortalama hava sıcaklığı tahminlerinde başarılı sonuçlar vermiştir ve bu bulgular, hava sıcaklığı tahmininde makine öğrenimi tekniklerinin etkili bir şekilde kullanılabileceğini göstermektedir.

Anahtar kelimeler: Hava sıcaklığı, Genetik programlama, Gradient boosting, Random forest.

*Sorumlu Yazar/Corresponding Author: kkumas@mehmetakif.edu.tr

 ISSN: 2618-6055 / 8, (2), 182 – 191, 2024
DOI: 10.62301/usmtd.1577839

1. Introduction

Climate change is among the most important environmental problems of the 21st century [1]. This problem causes an increase in precipitation regime, temperature, and CO_2 concentration on the earth. It is known that global temperature is increasing due to the rise in greenhouse gases [2]. The increase in global temperature in recent years has also accelerated the hydrological cycle [3]. Temperature, one of the climate parameters, directly affects evaporation, snowmelt, and frost and indirectly affects atmospheric stability and precipitation conditions. Temperature is also one of the main atmospheric variables used in hydrological modeling [4]. Ground-level air temperatures are expected to warm land faster than oceans [5]. The global increase in near-surface air temperature since the late nineteenth century has increased the frequency, intensity, and duration of extra heat and warm weather events [6,7]. Temperature affects climate systems and the hydrological cycle directly or indirectly. It is extremely important to make possible temperature predictions to predict these effects of temperature and take precautions against disasters caused by temperature. This raises awareness and provides warnings and guidance for governments to implement appropriate policies in climate, economy, health, etc. Forecasting weather phenomena uses numerical simulations to describe climate behavior, focusing on temperature, pressure, and humidity changes. Data from current atmospheric conditions are input into a computer to program mathematical equations. Different weather models mimic the Earth's atmosphere, offering a unique interpretation of atmospheric energy [8,9]. Forecasters are familiar with the characteristics of each model and prioritize those with the highest ability to predict their specific conditions. Accurate temperature prediction is crucial for governments' climate policies, energy companies' plans to meet energy demand, agriculture, food management, forest conservation, and health. Understanding temperature change and predicting future temperature values prevent disasters such as heat waves, droughts, floods, forest fires, etc. [10]. Air temperature forecasts for Konya province are critical for many sectors, such as agriculture, animal husbandry, energy management, and urban planning. As Konya is one of Türkiye's most important agricultural regions, temperature changes directly affect crop growth, yield, and quality. Temperature forecasts provide farmers with information to plan planting and harvesting periods better and optimize irrigation and fertilization strategies, thus minimizing crop loss and increasing yields. Thus, knowing the weather forecasts in Konya enables effective use of resources by increasing efficiency and safety.

As a result, knowing the weather forecasts in Konya is critical to increasing agricultural productivity, protecting animal health, using water resources efficiently, and managing energy demand. It also ensures public safety by taking precautions against floods and heat waves and contributes to the efficient use of resources [11-13]. This study used regression-based machine learning algorithms to predict the monthly average air temperatures, which consist of real number values for Konya province. To test the prediction success of machine learning algorithms, the data set consisting of monthly average air temperatures for the years 1960-2021 was divided into a time series. Accordingly, predictions for 2019-2021 and 2022-2025 were evaluated.

2. Literature Review

Kınacı et al. analyzed one-year temperature series with bilinear models using daily average temperature data for Konya province between 1997 and 2004. The study compared the forecasting performances of linear and nonlinear models. The results show that the dual linear model is more successful than linear models in certain situations [14]. Terzi and Ersoy conducted drought analysis using precipitation data from 1971-2014 at meteorological stations in Konya. Drought categories of 3, 6, 9, and 12 months were determined by the Standardized Precipitation Index (SPI) method, and drought prediction models were developed using the artificial neural networks (ANN) method. The results showed that ANN could predict drought accurately, especially in the 12 months [15].

Various machine learning algorithms have been used to predict annual and average seasonal temperatures in Shafin, Bangladesh. Linear Regression, Polynomial Regression, Isotonic Regression, and Support Vector Regression methods were applied in the simulation experiments. While isotonic regression gave the most successful results on training data, polynomial regression, and support vector regression provided the most accurate results for future temperature predictions [16]. Süzülmüş used different artificial neural network models and multiple linear regression methods to predict the monthly average temperatures of Gaziantep based on geographical and meteorological data. Among the artificial

*Sorumlu Yazar/Corresponding Author: kkumas@mehmetakif.edu.tr

ISSN: 2618-6055 / 8, (2), 182 – 191, 2024
DOI: 10.62301/usmtd.1577839

neural networks, the Multilayer Perceptron (MLP) model was found to have the highest accuracy rate with an R^2 value of 99.90%, while the MLR model showed a lower performance. As a result, the predictions made with MLP were found to be closer to the actual data than the other models [17]. Turgut et al. recorded weather parameters such as temperature, humidity, wind direction, and intensity through sensors connected to the microcontroller and transferred them to the database in real time via an ethernet module. ANN processed the database to predict the air temperature. This method provides a faster and more practical prediction model than meteorological formulas [18]. Cifuentes et al. examined different machine-learning strategies in the literature for temperature prediction. They reported that deep learning methods are more successful in temperature prediction with lower error rates than traditional neural network architectures.

Moreover, the accuracy of the prediction methods varies depending on the data combinations, architectures, and learning algorithms used [19]. Sevinç and Kaya applied LSTM and ARIMA models for air temperature prediction using meteorological data from the Solhan district of Bingöl. In the analyses performed on the real dataset, both models provided highly accurate temperature forecasts, with an R-squared score of 0.95 for LSTM and 0.97 for ARIMA [20]. Adnan et al. evaluated the performance of LSSVM, GMDHNN, and CART in temperature prediction using monthly temperature data from Astore and Gilgit stations in Pakistan. LSSVM model provided more accurate temperature than the other models and gives lower errors, especially in RMSE values [21]. Sevinç and Kaya used LSTM and ARIMA models to forecast temperature for Diyarbakır province using air temperature data for 2014-2020. The R-square score of the LSTM model was obtained as 0.96, which indicates that the model predicts close to the actual temperature values. As a result, both models were successfully applicable in temperature forecasting [22]. Azari et al. tested various machine learning methods for temperature time series forecasting. Linear regression, k-nearest neighbor, support vector machines, ANN, Random Forest, and Adaptive Boosting methods were used to predict the temperature in Memphis, TN. As a result of the analysis, Artificial Neural Network was more successful than other methods and showed the best performance in temperature prediction [23]. Fister et al. proposed three artificial intelligence frameworks for long-term summer temperature forecasting. Using historical data for the regions of Paris (France) and Córdoba (Spain), average temperatures in August's first and second half are predicted. The proposed methods include CNN, various machine learning approaches, and Recurrence Plot-based CNN that translates temporal series into images, and successful prediction results were obtained in both regions [24]. Yılmaz et al. estimated Turkey's long-term average temperatures using three different interpolation methods: Inverse Distance Weighted Interpolation (IDW), Kriging, and Radial Basis Function (RBF). In the study, forecasts were made using monthly temperature data of 81 provinces between 1981-2020, and these methods' accuracy was tested. According to the results obtained, the Kriging method gave the best prediction at Ardahan station (COC: 30.22°C, OMH: 5.29°C, R^2: 0.988), while the IDW method gave the worst result at Aksaray station (COC: 121.94°C, OMH: 3.48°C, R^2: 0.375). The best prediction results of each method were observed at Şanlıurfa, Ardahan, and Şırnak stations, but the Kriging and RBF methods showed the lowest performance in Tunceli [25]. Coşkun analyzed the climate trends in the Salt Lake-Konya basin in the Central Anatolia Region of Turkey between 1970-2018 with data obtained from meteorological and flow stations. The analysis based on Mann-Kendall and Spearman's Rho tests showed that while there was an increase in temperatures, there was no significant trend in precipitation and runoff data. The results show that temperature increases do not significantly affect evaporation, precipitation, and runoff [26].

3. Materials and Methods

This study uses regression-based machine learning algorithms to predict monthly average air temperatures with real number values. Genetic Programming (GP), Gradient Boosting (GB), and Random Forest (RF) were preferred as machine learning algorithms.

GP is an innovative artificial intelligence method that enables the automatic generation of computer programs inspired by biological evolution. This approach iteratively evolves randomly initialized sets of programs using the basic mechanisms of biological evolution: natural selection, crossover, and mutation. The main goal of GP is to generate programs capable of achieving a specific goal or solving a problem. In this process, the generated programs are evaluated according to their performance; the most successful solutions are selected as the parents of the next generations. The crossover process combines the characteristics of these parent programs to produce new and potentially more successful child programs, while the mutation process adds diversity to the population, avoiding possible local

ISSN: 2618-6055 / 8, (2), 182 – 191, 2024
DOI: 10.62301/usmtd.1577839

optima. The iterative nature of GP allows for the development of more effective and efficient programs over generations. While traditional programming involves writing code manually, GP creates self-optimizing and dynamically adaptive programs through a fully automated process. This is a great advantage when flexibility and creativity are required, especially in complex and changing problems. GP is used in many fields, such as data analytics, optimization, control systems, robotics, finance, etc. For example, while GP can be used to discover hidden patterns in complex data sets, it can also provide solutions for multiple objectives in optimization processes. It can also create a competitive advantage by delivering creative and innovative approaches to strategy development and decision-making problems. GP's capacity to generate automated solutions saves both time and resources while enabling customized and optimized results that are impossible with manual programming. For this reason, GP is increasingly preferred in many disciplines as an artificial intelligence approach that transcends the limits of traditional methods [27-29].

GB is a powerful method that combines weak predictive models to compensate for each other's shortcomings. It has been effectively used in machine learning to solve problems such as classification and regression. The technique follows a process whereby each model is trained by focusing on the mistakes made by previous models, producing more accurate predictions at each step. The main goal is to build a strong model by combining a set of weak learners. These weak learners are usually decision trees, but GB is a flexible method that can be applied to different learners. The basic mechanism of GB is based on reducing errors from previous predictions. Using the gradient descent algorithm, each new model minimizes these errors. The process focuses on reducing the overall error of the model, which allows for highly accurate predictions, especially for complex data sets. GB is an ideal method for large data sets and complex analysis problems. One of the advantages of GB is its capacity to improve model performance systematically.

First, a simple model is built, and its prediction errors become the learning target for the next model. This iterative process ensures that mistakes are progressively reduced and that the resulting model becomes much more powerful and competent than the initial individual models. However, the GB method requires a careful approach to hyperparameter settings. It is important to choose the right parameter choices so that the model does not tend to overlearn (overfitting). Besides classification and regression, GB has been successfully used in many other application areas, such as anomaly detection, ranking systems, and time series analysis. For example, it draws attention to its high performance in financial data analysis, health prediction models, and analysis of user behavior.

Moreover, the flexibility and accuracy of this method allow users to create a suitable solution for different data sets and problems. This wide range of applications and strong performance of Gradient Boosting explains why it is often preferred in modern machine learning projects [30-32].

RF is a powerful and flexible method widely used in machine learning to solve classification and regression problems. This approach is based on an ensemble model that is built by combining a large number of decision trees. The model trains each decision tree on a different subset of the dataset and randomly selected features. This process increases the diversity of the model, resulting in generalizable results and allowing it to make predictions with high accuracy. Each decision tree produces independent predictions depending on the subset it is trained on. In classification problems, these predictions are combined by majority voting, while in regression problems, the final result is obtained by taking the arithmetic mean of the projections. With this approach, the errors of individual decision trees are balanced within the ensemble model, and the overall performance is improved.

Moreover, the RF method significantly reduces the risk of overfitting a single decision tree. The dependency of each tree on the data subset and features increases the model's ability to adapt to different data samples and makes the results more balanced. RF performs impressively, especially for large data sets and high-dimensional problems. It provides successful results even when there are complex relationships between features in the data set or when there is missing data. Moreover, the ability to determine the importance of features makes RF more advantageous in terms of interpretability than other machine learning methods. This makes it possible to decide which features impact the target variable more. Besides classification and regression, RF has many applications such as anomaly detection, data mining, bioinformatics, financial analysis and image processing. Thanks to its robust structure, flexibility, and accuracy, this method is preferred in many fields and is notable for its user-

*Sorumlu Yazar/Corresponding Author: kkumas@mehmetakif.edu.tr

ISSN: 2618-6055 / 8, (2), 182 – 191, 2024
DOI: 10.62301/usmtd.1577839

friendliness. The fact that it is less sensitive than other machine learning methods in optimizing the parameters makes RF a more easily applicable solution.

In conclusion, Random Forest is a method in which independent decision trees form a strong ensemble model, offering superior performance in terms of accuracy and generalizability. Its flexibility and robustness when working with large and complex datasets make RF an indispensable tool in modern machine-learning applications. [33-35].

The dataset is a comprehensive data set that includes monthly average temperature values and covers 696 months of data [36]. These data include long-term temperature records from 1960 to 2017 and constitute a fundamental source for temperature forecasting studies for Konya province. In Figure 1, the average temperature values between these years are visualized. In the figure, the temperature sets for each year, grouped by month, are visible and indicated by a red box as an example. This grouping allows us to understand better the temperature trends over many years and the monthly variations. The dataset recorded maximum temperature was 27.7 ºC in August 2010, while the minimum temperature was -8 ºC in January 1989. This wide temperature range provides an important reference for assessing the region's climate changes and extreme weather events. The data presented in Figure 1 provide important information, especially for analyzing temperature variations in different periods and preparing data for forecast models.

In this study, temperature data from 1960-2017 were used in the training process of machine learning algorithms. To test the accuracy of the obtained models and to evaluate future temperature forecasts, a total of 48 months of temperature data between 2018 and 2021 were tried to be predicted. These forecasts are critical to assess the performance of the models by comparing them with actual values and to test the effectiveness of different algorithms. This large data set and temperature records spanning different years constitute a rich source of information in modeling processes and provide a reliable basis for long-term forecasts.

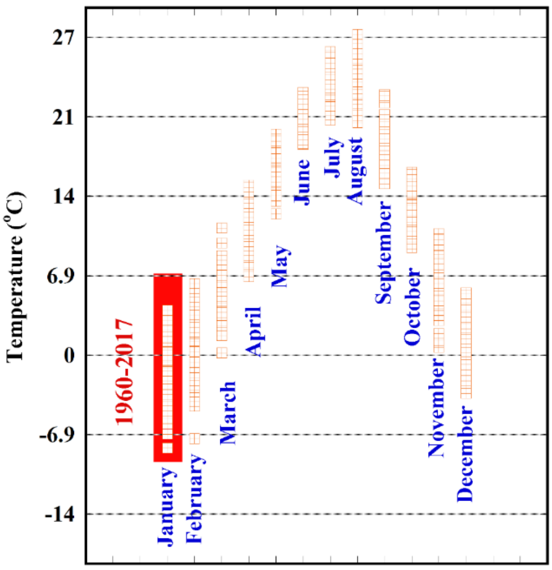

Figure 1. The data set of average air temperatures of Konya province between 1960-2017

4. Results and Discussion

The dataset for training and testing machine learning models has been treated as a time series. The training set includes data between 1960-2017, while the test set covers 2019-2021. The expected values between 2022 and 2025 are estimated based on the model's performance.To analyze the performance of machine learning models, evaluation metrics such as MAE, RMSE, and R^2, frequently preferred in regression problems, were utilized. In the literature, various metrics are used to analyze the performance of machine learning-trained models. In this study, mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R^2) metrics were used to determine the best model [31].

*Sorumlu Yazar/Corresponding Author: kkumas@mehmetakif.edu.tr

ISSN: 2618-6055 / 8, (2), 182 – 191, 2024
DOI: 10.62301/usmtd.1577839

Figure 2a compares the actual average temperature data and the data obtained with GP prediction. It can be considered that the GP predictions are consistent with the original temperature values. Figure 2b shows the residual values between GP predicted and real temperature values (2019-2021). The residual is the difference between the model-predicted and real measured values. In general, residual values for 48 months of data fluctuate between approximately 4 and -2. When the residual values are analyzed, it can be seen that the model is generally successful. In the training of the model for GP, R^2= 0.9477, RMSE= 1.950, and MAE= 1.500. In the testing phase of the model, R^2= 0.9560, RMSE= 1.7070, and MAE= 1.3790. Based on these values, it can be concluded that the model is successful. Figure 2c shows the box plot. When the box plot in Figure 2c is analyzed for GP, the first quartile (Q1) is 6.92 °C, the median (Q2) is 13.49 °C, and the third quartile (Q3) is 19.69 °C. The original temperature values are 7.70, 13.90, and 20.10 °C, respectively.

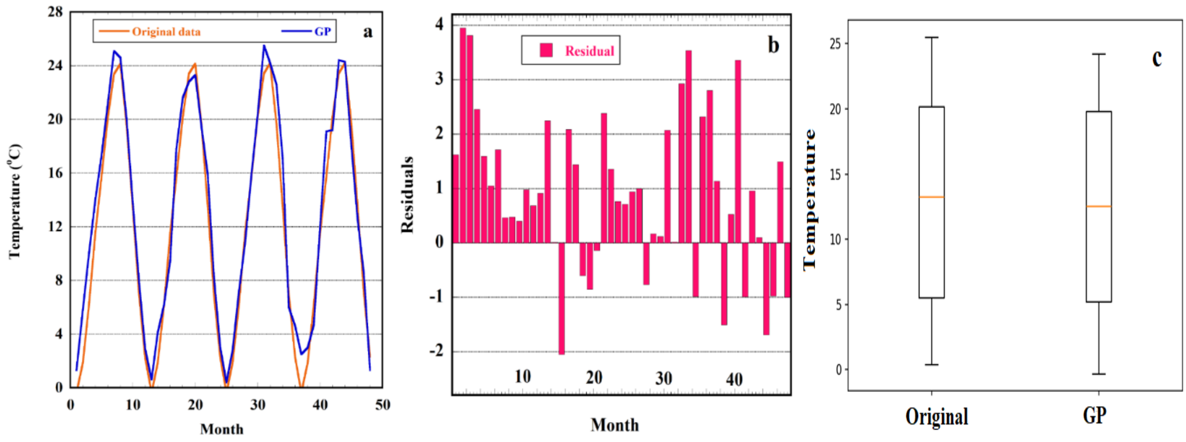

Figure 2. (a) Comparison of GP prediction results with original data for (2018-2021) **(b)** Residuals from GP prediction for (2018-2021) **(c)** Box plots of GP's predictions with original data for (2018-2021)

Figure 3a compares the GP forecast and the average temperature data. GB forecasts can be consistent. Figure 3b shows the residual values between the predicted and actual average temperature values obtained from GB (2019-2021). In general, the residual values vary between approximately 5 and -3. In the training of the model for GB, R^2= 0.9663, RMSE= 1.564, and MAE= 1.203. In the testing phase of the GB algorithm, R^2= 0.9211, RMSE= 2.285, and MAE= 1.777. It can be concluded that the model is successful. Figure 3c shows the box plot. The results are 6.50 °C for Q1, 13.45 °C for Q2 and 21.03 °C for Q3. The original temperature values are 7.70, 13.90 and 20.10 °C respectively

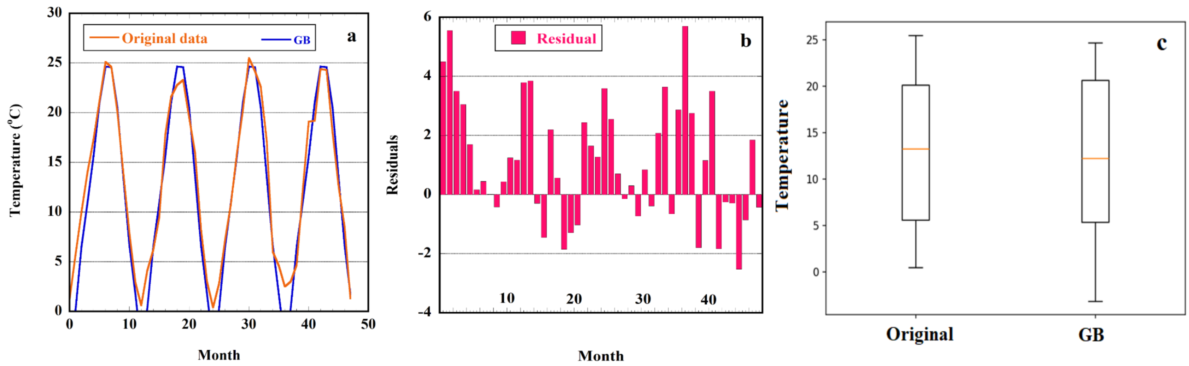

Figure 3. (a) Comparison of GB prediction results with original data for (2018-2021) **(b)** Residuals from GB prediction for (2018-2021) **(c)** Box plots of GB's predictions with original data for (2018-2021)

Figure 4a compares the original average temperature data with the data obtained with the RF prediction. RF predictions are consistent with the original temperature values. Figure 4b shows the residual values for the RF prediction, which fluctuate between about 5 and -3. It can be seen that the RF model is generally successful. In the training for RF, R^2= 0.9905, RMSE= 0.833 and MAE= 0.625. In the testing,

*Sorumlu Yazar/Corresponding Author: kkumas@mehmetakif.edu.tr

ISSN: 2618-6055 / 8, (2), 182 – 191, 2024
DOI: 10.62301/usmtd.1577839

R^2= 0.9205, RMSE= 2.294 and MAE= 1.812. The box plot shows Q1 is 6.39 ˚C, Q2 is 13.05 ˚C and Q3 is 21.35 ˚C. The original temperature values are 7.70, 13.90 and 20.10 ˚C respectively.

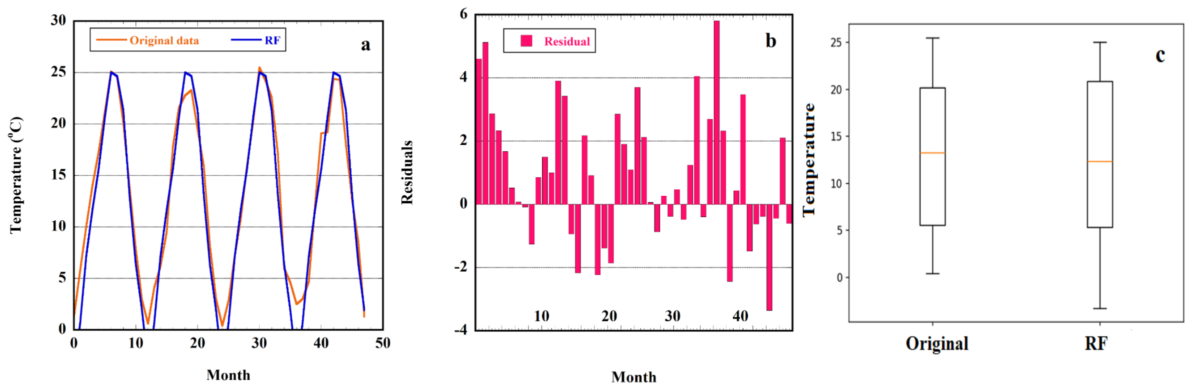

Figure 4. (a) Comparison of RF prediction results with original data for (2018-2021) **(b)** Residuals from RF prediction for (2018-2021) **(c)** Box plots of RF's predictions with original data for (2018-2021)

For the future predictions of the average air temperature for 2022-2025, GP, GB and RF algorithms were considered to give successful results. Figure 5a shows the prediction of the GP algorithm. The GP method predicts a minimum of-1.50 ˚C, a maximum of 25.12 ˚C , and an average of 12.39 ˚C for the average temperature between these years. Figure 5b shows the prediction of the GB algorithm. In this algorithm, the mean temperature's minimum, maximum and mean values are 0.89 ˚C, 24.55 ˚C, and 12.39 ˚C, respectively. For the RF algorithm, these values were estimated as 1.74, 24.59, and 12.63 °C, respectively (Figure 5c).

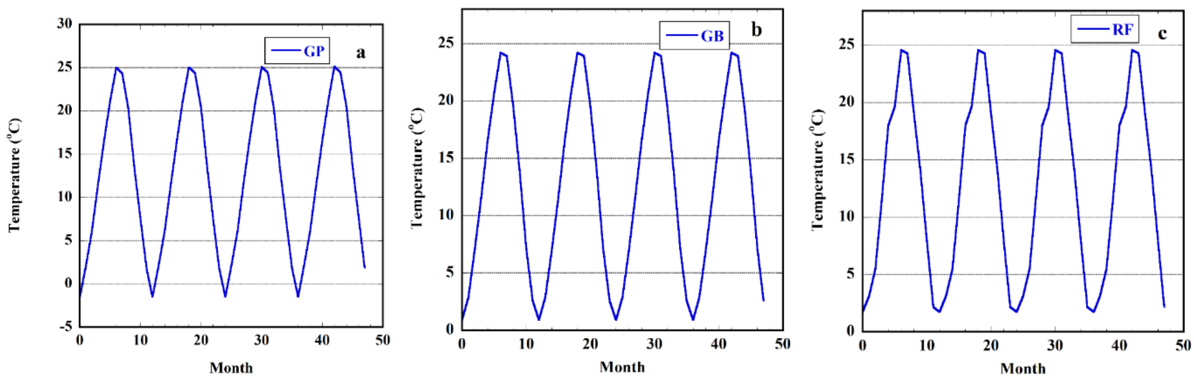

Figure 5. (a) GP predictions for average temperature **(b)** GB predictions for average temperature **(c)** RF predictions for average temperature (2022-2025).

Predicting monthly average temperature is critical in many areas, such as climate change, agricultural productivity, energy consumption, and health. Temperature data plays an important role in planning agricultural activities and crop selection. In addition, energy demand predictions are essential for energy management and the development of sustainable energy policies. Weather changes can also impact human health; accurate forecasts help mitigate adverse impacts. In this context, monthly average temperature predictions for Konya province will contribute to a better understanding and management of these important variables. In this study, GP, GB, and RF techniques were used to predict the temperature of Konya province using monthly average air temperature data between 1960 and 2017. First, the expected average monthly temperature values between 2018-2021 (48 months total) were compared with the actual values. Then, future predictions for the years 2022-2025 were also performed. Metrics used in model evaluations include MAE, RMSE, and coefficient of determination R^2. Successful results such as R^2= 0.9477, RMSE= 1.950, and MAE= 1.5000 were obtained for the genetic programming model. In the test phase of the model, R^2= 0.9560, RMSE= 1.7070, and MAE= 1.3790 values were recorded. These findings show that the GP model is successful. In the training of the GB

ISSN: 2618-6055 / 8, (2), 182 – 191, 2024
DOI: 10.62301/usmtd.1577839

model, R^2= 0.9663, RMSE= 1.564, and MAE= 1.203 values were obtained. The results in the test phase were R^2= 0.9211, RMSE= 2.285 and MAE= 1.777. These results also reveal the effectiveness of the model. In the training of the RF model, very good results such as R^2= 0.9905, RMSE= 0.833, and MAE= 0.625 were obtained. In the testing phase, R^2= 0.9205, RMSE= 2.294, and MAE= 1.812. The GP, GB, and RF algorithms for future forecasts produced different results for the average air temperature between 2022 and 2025. The GP method predicted the average temperature between -1.50 °C and 25.112 °C, while the GB algorithm predicted between 0.89 ºC and 24.0 ºC, and the RF algorithm predicted between 1.74 °C and 24.59 °C. As a result, the applied machine learning methods yielded successful results in monthly average air temperature predictions for Konya province. These findings show that machine learning techniques can effectively predict air temperature.

5. Conclusion

The findings of this study reveal the effectiveness of machine learning techniques such as GP, GB and RF in predicting monthly average air temperatures for Konya province. The comparison of the forecasts for the period 2018-2021 with the actual values and the future projections for the years 2022-2025 revealed the reliability and accuracy of the models. The GP performed well with high R^2 values and low error metrics, while the GB and RF models produced effective results on complex data sets. The small differences in the future predictions of the models emphasize the importance of evaluating more than one method to improve prediction accuracy. The findings suggest that machine learning methods can provide useful insights in agriculture, energy planning, and public health. In future studies, some improvements can be made to increase the accuracy and generalizability of the forecast models. First, more comprehensive models can be developed by integrating air temperature and other meteorological parameters such as humidity, wind speed, and solar radiation into the forecasting process. Such multivariate approaches can provide higher accuracy, especially in agriculture, energy, and climate change scenarios.

Furthermore, using datasets covering a wider range of dates can help improve the reliability of long-term projections. Another recommendation is the use of hybrid modeling methods. Hybrid models that utilize the strengths of various algorithms can potentially reduce error rates. Instead of limiting this study to Konya province, it is important to apply it to other regions with different climatic conditions to test the generalizability of the models and provide solutions suitable for different geographical conditions. Finally, scenarios can be created that assess the applicable impacts of the predictions obtained on agricultural planning, energy demand management, and public health. Such studies will contribute to the scientific literature and policy development processes at local and national levels.

6. Acknowledgments

The authors thank the Turkish State Meteorological Service for the data used in the study.

7. References

[1] T.H. Abebe, Time Series Analysis of Monthly Average Temperature and Rainfall Using Seasonal ARIMA Model (in Case of Ambo Area, Ethiopia), Int. J. Theor. Appl. Math. 6 (5) (2020) 76-87.

[2] J.Sillmann, T. Thorarinsdottir, N. Keenlyside, N. Schaller, L.V. Alexander, G. Hegerl, S.I. Seneviratne, R. Vautard, X. Zhang, F.W. Zwiers, Understanding, modeling and predicting weather and climate extremes: Challenges and opportunities, Weather and Climate Extremes 18 (2017) 65-74.

[3] E. Olmedo, A. Turiel, V. Gonzalez-Gambau, C. Gonzalez-Haro, A. Garcia-Espriu, C. Gabarro, M. Portabella, I. Corbella, M. Martin-Neira, M. Arias, R. Catany, R. Sabia, R. Oliva, K. Scipal, Increasing stratification as observed by satellite sea surface salinity measurements, Scientific Reports 12 (1) (2022)1-9.

[4] X. Liu, P. Coulibaly, Downscaling ensemble weather predictions for improved week-2 hydrologic forecasting, Journal of Hydrometeorology 12 (6) (2011)1564-1580.

[5] E.S. El-Mallah, S.G. Elsharkawy, Time-Series Modeling and Short Term Prediction of Annual Temperature Trend on Coast Libya Using the BoxJenkins ARIMA Model, Advances in Research 6 (5) (2016)1-11,

*Sorumlu Yazar/Corresponding Author: kkumas@mehmetakif.edu.tr

ISSN: 2618-6055 / 8, (2), 182 – 191, 2024
DOI: 10.62301/usmtd.1577839

[6] S. E. Perkins-Kirkpatrick, S.C. Lewis, Increasing trends in regional heatwaves, Nature Communications 11 (1) (2020)1-8.

[7] A. Sulikowska, A. Wypych, Seasonal Variability of Trends in Regional Hot and Warm Temperature Extremes in Europe, Atmosphere 12 (5) (2021)1- 21.

[8] S. Al-Yahyai, Y. Charabi, A. Gastli, Review of the use of numerical weather prediction (NWP) models for wind energy assessment, Renewable and Sustainable Energy Reviews 14(9) (2010) 3192-3198.

[9] P. Bauer, A. Thorpe, G. Brunet, The quiet revolution of numerical weather prediction. Nature 525 (2015) 47–55.

[10] A. Dai, Increasing drought under global warming in observations and models, Nature Climate Change 3 (2013) 52-58.

[11] B. Özgür, The assessment of socio-economic impacts of climate change in rural areas: the case of Konya, Thesis (M.S.) Graduate School of Natural and Applied Sciences, City and Regional Planning. Middle East Technical University, Ankara, Türkiye, 2019.

[12] J.W. Jones, J.W. Hansen, F.S. Royce, C. D. Messina, Potential benefits of climate forecasting to agriculture. Agriculture, Ecosystems & Environment *82*(1-3) (2000)169-184.

[13] C. Oğuz, A. Y. Ögür, Climate Change and Agriculture: The Case of Konya Province in Recent Academic Studies in Sciences, B Ranguelov R. Efe, M.S. DINU, E. Atasoy (Eds), Sofia: St. Kliment Ohridski University Press, (2021)1-20.

[14] İ. Kınacı, Konya İli Sıcaklık Verilerinin Çift Doğrusal Zaman Serisi Modeli İle Modellenmesi, 3. Yenilenebilir Enerji Kaynakları Sempozyumu, Türkiye, Mersin, 2005

[15] Ö. Terzi, T. Ersoy, Yapay Sinir Ağları ile Konya İli Kuraklık Tahmini. DSI Technical Bulletin/DSI Teknik Bülteni, (2018) (127).

[16] A. A. Shafin, Machine learning approach to forecast average weather temperature of Bangladesh, Global Journal of Computer Science and Technology 19(3) (2019) 39-48.

[17] S. Suzulmus, Prediction of average temperatures using artificial neural network methods: the case of Gaziantep Provnice, Turkey. Fresenius Environmental Bulletin 28(2A) (2019) 1494-1502.

[18] A. Turgut, A. Temir, B. Aksoy, K. Özsoy, Yapay Zekâ Yöntemleri ile Hava Sıcaklığı Tahmini İçin Sistem Tasarımı ve Uygulaması, International Journal of 3D Printing Technologies and Digital Industry *3*(3) (2019) 244-253.

[19] J. Cifuentes, G. Marulanda, A. Bello, J. Reneses, Air temperature forecasting using machine learning techniques: a review, Energies 13(16) (2020) 4215.

[20] A. Sevinç, B. Kaya, Derin öğrenme ve istatistiksel modelleme yöntemiyle sıcaklık tahmini ve karşılaştırılması. Avrupa Bilim ve Teknoloji Dergisi (28) (2021) 1222-1228.

[21] R. M. Adnan, Z. Liang, A. Kuriqi, O. Kisi, A. Malik, B. Li, F. Mortazavizadeh, Air temperature prediction using different machine learning models, Indonesian Journal of Electrical Engineering and Computer Science 22(1) (2021) 534-541.

[22] A. Sevinç, B. Kaya, Derin Öğrenme Yöntemleri ile Sıcaklık Tahmini: Diyarbakır İli Örneği, Computer Science Special (2021) 217-225.

[23] B. Azari, K. Hassan, J. Pierce, S. Ebrahimi, Evaluation of machine learning methods application in temperature prediction, Computational Research Progress in Applied Science & Engineerıng (CRPASE) 8(1) (2022) 1-12.

***Sorumlu Yazar/Corresponding Author:** kkumas@mehmetakif.edu.tr

ISSN: 2618-6055 / 8, (2), 182 – 191, 2024
DOI: 10.62301/usmtd.1577839

[24] D. Fister, J. Pérez-Aracil, C. Peláez-Rodríguez, J. Del Ser, S. Salcedo-Sanz, Accurate long-term air temperature prediction with machine learning models and data reduction techniques, Applied Soft Computing 136 (2023) 110118.

[25] C. B. Yılmaz, H. Bodu, E. S. Yüce, V. Demir, M. F. Sevimli, Türkiye'nin uzun dönem ortalama sıcaklık (°C) değerlerinin üç farklı enterpolasyon yöntemi ile tahmini, Geomatik 8(1) (2023) 9-17.

[26] S. Coşkun, tuz gölü-konya kapalı havzalarının yaz mevsimi ortalama sıcaklık, yağış, buharlaşma ve akım verilerindeki değişimlerin karşılaştırmalı trend analizi, The Journal of Social Sciences (46) (2024). 123-138.

[27] C. Gagné, M. Parizeau, Genericity in evolutionary computation software tools: Principles and case-study, International Journal on Artificial Intelligence Tools 15(02) (2006) 173-194.

[28] A. N. Sloss, S. Gustafson, 2019 evolutionary algorithms review, Genetic programming theory and practice XVII, (2020) 307-344.

[29] A. Tripathi, R. Singh, A. K. Singh, P. Gupta, S. Vats, M. Singhal, Significance of Evolutionary Artificial Intelligence: A Detailed Overview of the Concepts, Techniques, and Applications, Artificial Intelligence, Machine Learning and User Interface Design 27(53) (2024) 27-53.

[30] A. Natekin, A. Knoll, Gradient boosting machines, a tutorial, Frontiers in neurorobotics 7 (2013) 21.

[31] I. Pence, R. Yıldırım, M. S. Cesmeli, A. Güngör, A. Akyüz, Evaluation of machine learning approaches for estimating thermodynamic properties of new generation refrigerant R513A. Sustainable Energy Technologies and Assessments 55 (2023) 102973.

[32] S. Babu Nuthalapati, A. Nuthalapati, Accurate weather forecasting with dominant gradient boosting using machine learning, Int. J. Sci. Res. Arch 12(2) (2024) 408-422.

[33] S. J. Rigatti, Random forest, Journal of Insurance Medicine 47(1) (2017) 31-39.

[34] A. Parmar, R. Katariya, V. Patel, A review on random forest: An ensemble classifier. In International conference on intelligent data communication technologies and internet of things Springer International Publishing (ICICI) 2018 (2019) (pp. 758-763)

[35] I. Pence, K. Kumas, M. S. Cesmeli, A. Akyüz, Animal-based CO_2, CH_4, and N_2O emissions analysis: Machine learning predictions by agricultural regions and climate dynamics in varied scenarios, Computers and Electronics in Agriculture 226 (2024)109423.

[36] Turkish State Meteorological Service (TSMS) Turkish State Meteorological Service https://www.mgm.gov.tr/tahmin/turkiye.aspx (accessed 6.22.23) Ankara Turkey

*Sorumlu Yazar/Corresponding Author: kkumas@mehmetakif.edu.tr