ALFA BK UNIVERSITY
FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
FACULTY O F COMPUTER SCIENCE
ALFATECH
Proceedings of Conference

UDK: 551.509:519.87
004:007]:711.4
COBISS.SR-ID 148884489
DOI: 10.5281/zenodo.12614633
Original scientific paper

TIME SERIES MODELS FOR WEATHER FORECASTING IN SMART CITIES

Milos Todorov [34]; Ninoslava Tihi[35]; Srdjan Popov[36]; Biljana Stamatovic[37]

Abstract

Weather catastrophes are highly destructive occurrences that impact numerous sectors, such as public health, agriculture, industry, and the environment. In order to prevent this catastrophic event, many prediction methodologies can be implemented in smart cities. The Seasonal ARIMA method is the most used approach for time series forecasting. This study aims to outline a technique for the preparation and evaluation of raw data, specifically focusing on its application to meteorological data sets. Applying a mathematical-informatics model for time series, using R and its library *forecast* for prediction, the following results were obtained.

Keywords: *Smart cities, mathematical-informatics model, time series, SARIMA, weather forecasting, R*

Introduction

Extreme weather events have a disproportionate impact on urban areas. Global warming adds to greater temperatures, which leads to increased evaporation and drying of water supplies. To minimize adverse weather impacts and reduce water demand while also conserving and managing limited water resources, diligent monitoring and adequate water strategies are required [1]. Forecasting urban droughts is a strategy that municipalities and water management organizations can use to prepare for and take appropriate measures in a timely manner.

One strategy used to predict urban extreme weather events is statistical forecasting, such as time series analysis. Historical meteorological data analysis can reveal trends and patterns that can then be utilized to predict urban drought factors.

[34] Milos Todorov, 1982, Assistant professor, PhD, Faculty of Mathematics and Computer Sciences, Boulevard Marsala Tolbuhina 8, Belgrade, milos.todorov@alfa.edu.rs https://orcid.org/0000-0002-5614-3057

[35] Ninoslava Tihi, 1979, lecturer, MsC, The Higher Education Technical School of Professional Studies in Novi Sad, tihi@vtsns.edu.rs https://orcid.org/0009-0004-8009-8120

[36] Srdjan Popov, 1969, Full professor, PhD, Faculty of Technical Science, Trg Dositeja Obradovica 6, Novi Sad, srdjanpopov@uns.ac.rs https://orcid.org/0000-0003-1215-3111

[37] Biljana Stamatovic, 1968, full professor, PhD, Faculty of Philosophy, University of Montenegro, biljas@ucg.ac.me https://orcid.org/0000-0002-3934-9686

ISBN: 978-86-6461-074-2

Forecasting Weather in Smart Cities

Smart cities are transforming how we work, live and engage with our surroundings. These contemporary cities use a range of technologies and data analytics to boost productivity, increase sustainability, and improve quality of life. The capacity of a smart city to forecast the weather precisely and instantly is one of its main features. Smart cities can provide accurate weather forecasts using satellites, sophisticated sensors, and machine learning algorithms. This makes it possible for residents and local authorities to plan and get ready for weather-related events more efficiently.

The Internet of Things (IoT) is the fundamental idea behind smart cities' ability to predict the weather. Massive volumes of data about atmospheric conditions, temperature, humidity, and wind patterns are gathered by Internet of Things devices, including weather stations, sensors on buildings, and even smartphone applications. After that, the data is sent to a central hub for real-time processing and analysis. The gathered data is converted into precise weather forecasts with the aid of machine learning algorithms, offering vital information that helps people, organizations, and governmental bodies make educated decisions.

The capacity to prepare for and react to extreme weather events in smart cities is the main benefit of accurate weather prediction. Climate change has made heatwaves, storms, and floods more frequent in today's world. Accurate weather forecasts allow local government agencies to plan, efficiently distribute emergency assistance, and even evacuate high-risk regions when needed. Furthermore, residents can use smartphone applications to obtain fast weather reports, which allows them to modify their schedules and take the required safety steps.

Moreover, precise weather forecasting in smart cities has usages beyond disaster relief. Forecasts with a high degree of accuracy are very helpful for industries like transportation, tourism, and agriculture that are highly dependent on the weather. Farmers can predict crop yields, manage irrigation schedules, and shield their harvests from inclement weather. By creating more effective routes and schedules, transportation companies may cut down on delays and raise customer satisfaction. In the interim, travellers can plan for outdoor events and activities with knowledge, avoiding any potential hassles brought on by inclement weather.

Metrics for Time Series

Forecasting time series entails projecting future values based on past data, and it is an important concept in many industries, including finance, economics, and meteorology. A time series is a collection of observations or data points on a single variable over time. Time series forecasting employs a variety of methods. These data-driven models have been shown to be extremely accurate in meteorological forecasting applications. [2].

When forecasting time series, it is essential to evaluate the data's properties, such as trends, seasonality, and any other aspects that may affect the model's accuracy. Model performance should be evaluated using statistical metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), or Mean Absolute Percentage Error (MAPE).

ARIMA and SARIMA method

ISBN: 978-86-6461-074-2

Arima forecasting, often referred to as autoregressive integrated moving average forecasting, is a widely used method for predicting the future values of a time series based on its historical values. This statistical approach employs advanced techniques such as autoregression (AR), moving average (MA) models, and differencing (I) to ensure the stability of the series.
Arima forecasting is a proficient method for prognosticating future values in a time series through the analysis of historical data. The model incorporates autoregressive, moving average, and differencing components to accurately depict the dependencies and patterns of the series.
Proficiency in Arima forecasting necessitates a comprehensive understanding of time series theory and statistical principles.
The autoregressive component of the model pertains to the current value's reliance on preceding values, with the autoregression order (p) determining the number of previous values considered. The moving average component considers the influence of previous errors on the present value, where the moving average order (q) specifies the number of past errors to be taken into consideration. To get a series to be stationary, differencing (d) is used. This is done by subtracting the previous observation from the current one and getting rid of any trend or seasonal effects.
The typical notation for the non-seasonal ARIMA model is ARIMA (p,d,q), and each parameter is identified based on the time series autocorrelation function (ACF) and partial autocorrelation function (PACF).
The moving average component shows how the current observation is related to the error left over after using a moving average model on past values [3].
Sarima forecasting, often referred to as seasonal autoregressive integrated moving average, is a valuable statistical method for predicting time series data that exhibits seasonal trends. The system employs two widely recognized forecasting techniques, ARIMA and seasonal decomposition of time series (STL), to accurately capture both the long-term trend and seasonal patterns present in the data. To comprehend Sarima, one must possess a comprehensive understanding of time series analysis as well as mathematical concepts such as autocorrelation and differencing.
The ARIMA model is adapted to forecast seasonal data. SARIMA includes three seasonal components, namely P, D, and Q, in addition to the three non-seasonal components mentioned earlier [4].

Research Methodology

The R programming language was combined with RStudio (version 2021.09.0 Build 351), a R IDE, to serve as the software tool for this study. R is a prominent statistical programming language and environment that is widely used for statistical computing, data analysis, and visualisation. It contains many packages that are available through the R archive network CRAN (http://cran.r-project.org) [5].
The data for this study was gathered from a sensor station that is part of a network known as "Proactive Networks" and is situated in the municipality of Novi Sad Station. The data was collected between January 2014 and December 2018. The sensors can be categorized into six groups: soil moisture sensors (SM1, SM2, SM3, SM4, SM5, SM6), humidity sensor (AH1), wind speed sensor (WS1), wind direction sensor (WD1), air temperature sensor (AT1), precipitation sensor (PP1), and battery power state sensor (BT1).

ISBN: 978-86-6461-074-2

Records are made by sensors every hour, resulting in a total of 288 records every day. Once captured, the data is transmitted to the online portal, where it is available for download in CSV format. *Table 1* displays four attributes present in the data set.

Attribute name	Description
Time	Time and date of the measured parameter value
Device	Device identification number where the value is measured
ID value	Sensor identification number
Value	Measured value of a parameter

Table 1 − Attribute description

Upon retrieval from the web portal, the data undergoes conversion and is thereafter saved in a data frame. A data frame is a fundamental data structure utilized for storing and arranging tabular data in the R programming language. Upon storage, the data undergoes filtration to extract only pertinent parameters, such as air temperature, precipitation, and soil moisture, from the dataset. The air temperature and soil moisture time series are aggregated using the "mean" function to calculate the average values. The "sum" function is utilized to calculate the cumulative values for precipitation time series. Subsequently, the values are organized in ascending order based on months and years and subsequently transformed into time-series entities. Ultimately, the acquired dataset must be split into two subsets, with one subset being utilized for constructing a model and the other subset for model verification [6].
Data is utilized for constructing predictive models after undergoing partitioning. Application of the Seasonal ARIMA technique. This approach will exclusively be employed for temperature time series, serving as an illustrative example. The function "Arima()" is utilized to construct a model using the seasonal Arima approach.

Results and Discussion

Graphical presentation of the average temperatures is shown in *Figure 1*. Y-axis displays average values of temperatures from January 2014 to December 2019. Temperature values varies and oscillations are obvious based on season. The lowest temperature was in January 2017.

ISBN: 978-86-6461-074-2

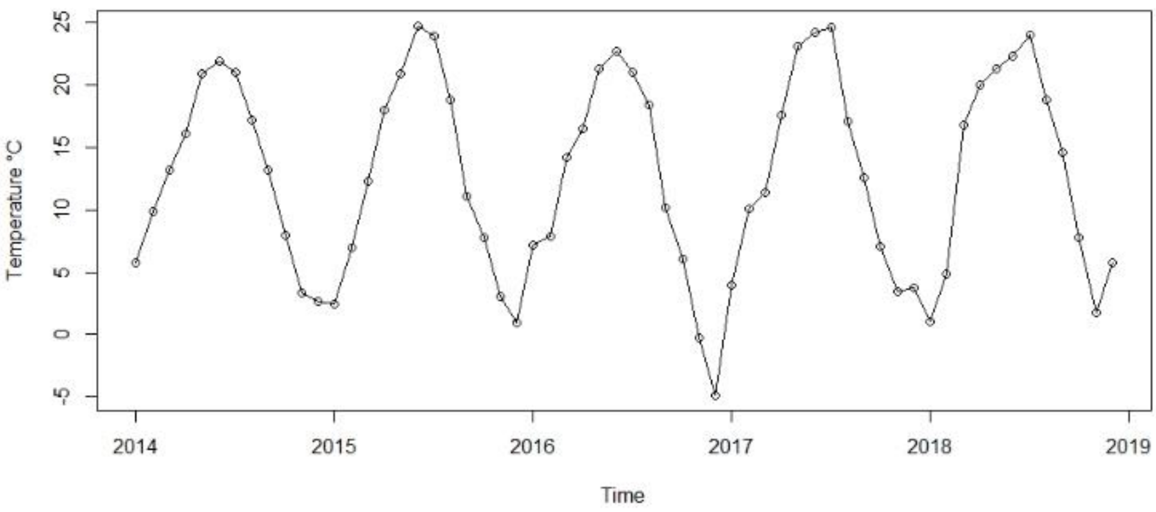

Figure 1 − Graphical representation of average temperatures

SARIMA method results

Selecting optimal parameters for the model can be very challenging and often requires expertise. Therefore, *forecast* package and its function "*auto.arima*", for automatically selecting the appropriate values for seasonal ARIMA parameters, was used to obtain the best model. This function automates the selection process by searching through a range of possible combinations of ordinary parameters p,d and q and seasonal parameters P, D and Q values and selects the model that minimizes a chosen information criterion, which measures the balance between model fit and complexity. This functions also simplifies the task of selecting the best SARIMA model for time series analysis and forecasting by automatically determining the optimal parameters based on statistical criteria. It also saves a lot of precious time because it can find a suitable ARIMA model without needing to manually test multiple combinations of parameters [7].

Applying a mathematical-informatics model for time series, using R and its package *forecast* for prediction, the following results were obtained. After applying *auto.arima* function from the forementioned package, it is found that the best selected model for predicting the air temperature is ARIMA $(1,0,1)(1,1,0)_{[12]}$. The coefficients and values of AIC (**A**kaike **I**nformation **C**riteria), AICc (**A**kaike **I**nformation **C**riteria **c**orrected) and BIC (**B**ayesian **I**nformation **C**riteria) information criteria are shown in Table 2.

Table 2 – Best selected SARIMA model and its coefficients and criteria values

Model	**Coefficients**			**Criteria**		
	AR1	**MA1**	**SAR1**	**AIC**	**AICc**	**BIC**
ARIMA $(1,0,1)(1,1,0)_{[12]}$	-0.6036	0.9209	-0.4079	228.59	229.52	236.08

ISBN: 978-86-6461-074-2

Forecast data obtained from the best fitted SARIMA model, along with historical and validation data is shown in *Figure 2*. The original historical data is shown by green line and validation data is shown by the blue line. The red line shows the forecasted values for twelve months ahead which seem quite accurate. There is a room for improvement of this model which can be performed in future work of this research.

Figure 2 − Forecast from ARIMA (1,0,1)(1,1,0)[12]

Conclusion

In conclusion, the resilience, sustainability, and general efficiency of urban environments are improved by the incorporation of weather prediction skills into smart cities. Smart cities can provide accurate and real-time weather forecasts by leveraging IoT devices, data analytics, and machine learning algorithms. This enables individuals, organizations, and government agencies to better plan for and respond to extreme weather disasters, as well as optimize various weather-dependent industries. As technology advances, the future of smart cities and weather prediction holds enormous promise for building safer and more habitable urban environments.
The Arima method is the most used for forecasting time series. The goal of this research was to demonstrate the process for selecting the optimal forecasting model based on metrics such as RMSE, MAPE, and MAE. The methodology was illustrated using a meteorological data set with average temperature values. During the investigation, the best model obtained using function auto.arima was: ARIMA (1,0,1)(1,1,0)[12]. Further work would include using different time series methods and comparing them with ARIMA results. Another way to extend this research could be by collecting more data and observing how the models behave with a different number of observations included.

References

[1] Toreti, A.; Bavera, D.; Acosta Navarro, J.; Arias Munoz, C.; Avanzi, F.; Barbosa, P.; de Jager, A.; Di Ciollo, C.; Ferraris, L.; Fioravanti, G.; Gabellani, S.; Grimaldi, S.; Hrast Essenfelder, A.; Isabellon, M.; Jonas, T.; Maetens, W.; Magni, D.; Masante, D.; Mazzeschi, M.; McCormick, N.; Meroni, M.; Rossi, L.; Salamon, P.; Spinoni, J.: Drought in Europe March 2023, *Publications Office of the European Union*, Luxemburg, 2023. doi:10.2760/998985, JRC133025

ISBN: 978-86-6461-074-2

[2] Zhu, S.; Luo, X.; Xu, Z.; Ye, L.: Seasonal Streamflow Forecasts using Mixture-kernel GPR and Advanced Methods of Input Variable Selection, *Hydrology Research*, *Vol. 50*, (2019), Issue:1, pp. 200–214, ISSN: 1998-9563, EISSN: 2224-7955.

[3] Stellwagen, E.; Tashman, L.: ARIMA:The Models of Box and Jenkins, *Foresight: The International Journal of Applied Forecasting*, *Vol. 30*, (2013), pp. 28-33.

[4] Mills, T.: *Applied Time Series Analysis-A Practical Guide to Modeling and Forecasting*, Elsevier Inc., Loughborough, UK, 2019

[5] https://www.r-project.org/about.html, accessed in April 2023

[6] Tihi, N., Popov, S., Bondžić, J., Dujović, M.: Visualization of Big Data as Urban Drought Monitoring Support in Smart Cities, *Fresenius Environmental Bulletin*, *Vol. 30*, No. 01, 2021, pp. 716-722, ISSN: 10184619

[7] Tihi, N., Popov, S.: Selection of the Best ARIMA Models for Urban Drought Prediction, Fresenius Environmental Bulletin (FEB), Vol. 32, No. 06, 2023, pp. 2564-2572

[8] https://posit.co/download/rstudio-desktop/, accessed in February 2024

ISBN: 978-86-6461-074-2