Advanced Industrial SCIence
Vol. 4, No. 1, pp. 49-55, 2025
ISSN 2951-2476
https://doi.org/10.23153/AI-Science.2025.4.1.049

Abnormal Weather Notification Method using SARIMA Model

Shin-Hyeong Choi[1], Choon-Soo Lee[2*]
[1]Professor, Division of Electrical, Control & Instrumentation Engineering, Kangwon National University
[2]Senior manager, Component Engineering Team, Korea Hydro-Nuclear Power Company

SARIMA 모델을 사용한 이상기후 알림 방안

최신형[1], 이춘수[2*]
[1]강원대학교 전기제어계측공학부 교수, [2]한국수력원자력(주) 기기엔지니어링부 차장

Abstract Global warming is causing enormous damage to the Earth by causing natural disasters and environmental changes that are different from the past. Among the various factors that threaten human health, smog is a representative example of respiratory disease, and recent news reports show that many people are hospitalized with respiratory diseases due to extreme harmful smog. This study proposes a system that uses the SARIMA(Seasonal ARIMA) model to predict wind speed and direction and detects abnormal weather by comparing the predicted values with past averages. To this end, we collected wind direction and wind speed data for the past 10 years in Region G, and if the predicted values differ significantly from the past, we can use this to reduce casualties and respond by sending warnings to local residents via SNS or emergency disaster text messages.

Key Words : Respiratory diseases, Wind direction, Wind speed, SARIMA model, Abnormal weather

요 약 지구온난화는 과거와 다른 자연재해 및 환경에 변화를 줌으로써 지구상에 막대한 피해를 주고 있다. 이와 같이 인간의 건강을 위협하는 여러 요인중에 호흡기 질환을 일으키는 것으로는 스모그가 대표적인 사례이며, 최근의 뉴스를 살펴보면, 극심한 유해 스모그로 인해 많은 사람이 호흡기 질환으로 입원한다는 보고가 있다. 본 연구는 SARIMA 모델을 사용하여 풍속 및 풍향을 예측하고, 예측된 값과 과거 평균값을 비교하여 이상기후를 감지하는 시스템을 제안한다. 이를 위해 G 지역의 10년간의 풍향, 풍속 데이터를 수집하였고, 예측된 값이 과거와 크게 달라질 경우, 해당 지역 주민들에게 SNS 또는 긴급재난문자를 통해 경고를 발송하여 인명 피해를 줄이고 대응할 수 있는데 활용할 수 있다.

주제어 : 호흡기 질환, 풍향, 풍속, SARIMA 모델, 이상기후

*Corresponding Author : Choon-Soo Lee(ceprlee@gmail.com)
Received December 14, 2024 Revised January 3, 2025
Accepted January 20, 2025 Published January 30, 2025

1. Introduction

As can be seen from the UN survey results that approximately 40% of the world's land has become dryland, the desertification phenomenon on Earth has been accelerating in recent years [1]. Due to global warming, climate change has been aggravated in recent years and natural disasters that were not experienced in the past have occurred. The environmental impacts caused by this pose a serious threat to human health. In particular, air pollution such as smog is a major cause of respiratory diseases. Recent news reports have reported cases of many people being hospitalized for respiratory diseases due to severe harmful smog, which clearly shows the impact of climate change on human health [2-5]. In particular, wind direction and wind speed play an important role in the spread of air pollution, and rapid changes in wind direction and wind speed can further spread environmental disasters such as smog. Therefore, it is important to establish a system that predicts climate change and detects abnormal climate resulting from it.

Therefore, this study proposes a system that predicts wind direction and wind speed in a specific area using the SARIMA(Seasonal ARIMA) model, compares the predicted values with past average values, predicts future wind direction and wind speed, and detects abnormal weather in real time. This system can be used to reduce human casualties and respond by sending a warning to the local residents through SNS or emergency disaster text messages when the predicted values are significantly different from the past. This paper is organized as follows. Chapter 2 reviews existing studies and related literature, and analyzes studies related to wind speed and wind direction prediction. Chapter 3 explains the data and analysis methods used in the study, and explains the prediction process using the SARIMA model and the abnormal climate detection method. Chapter 4 deals with experiments and results, and provides the results of comparing the predicted wind speed and wind direction with the actual values. Chapter 5 is the conclusion.

2. Related Works

Research on wind direction and wind speed prediction has been conducted continuously from the past [6, 7]. In general, research has been conducted in relation to weather forecasting and wind power generation [8], and it is also used in the feasibility analysis of wind power plants, which has recently become more popular due to the transition to renewable energy [9]. In this chapter, we will analyze existing research data and examine research trends.

Saurabh S. Soman et al. [10] studied various forecasting techniques for wind power and wind speed, categorizing them into very short-term, short-term, medium-term, and long-term forecasting based on the concept of time. In very short-term forecasting, techniques such as Adaptive Neuro-Fuzzy Inference System(ANFIS) and Autoregressive Integrated Moving Average Model(ARIMA) are used, which are useful for forecasting within minutes, often using data from seconds to minutes ago. In short-term forecasting, techniques such as Bayesian approach, ARIMA [11-13], and hybrid models are used to identify weekly patterns and short-term weather phenomena. In medium-term forecasting, techniques such as Numerical Weather Prediction(NWP) models and integration with machine learning techniques are used to forecast up to a week ahead. In long-term forecasting, climate models and statistical methods are used to forecast from weeks to years ahead, which is used for planning and policy making of wind energy production. Hybrid approaches combine different forecasting methods to improve accuracy, and can leverage the strengths of both physical and statistical approaches, as well as short-term and long-term models. Aditi Choudhary et al [14] review various deep learning methods for wind power prediction with intermittent characteristics, emphasizing the importance of renewable energy due to climate change and global warming. In this paper, deep learning models such as Long Short-Term Memory(LSTM) and Gated Recurrent Unit(GRU) combined with machine learning are shown to improve prediction accuracy to overcome the limitations of conventional statistical models and machine learning methods that cannot capture complex and nonlinear relationships in wind power data. Xiaolei Liua et al [15] examine the effectiveness of various prediction models for predicting offshore wind speed. In particular, this study uses seasonal autoregressive integrated moving average(SARIMA) to predict hourly wind speed in coastal areas and compares it with previously discussed deep learning models, GRU and LSTM. The results show that the SARIMA model outperforms GRU and LSTM in terms of accuracy and robustness of wind speed prediction.

3. Data Analysis

3.1 Data Collection

In this study, we use 10-minute wind direction and wind speed data for 10 years from 2014 to 2023 in Region G. The data is composed of CSV files consisting of year, month, day, hour, minute, wind direction at 10M above ground, wind speed at 10M above ground, and maximum wind speed at 10M above ground. The data are collected from the automatic weather observation equipment installed in the institution. The collected data goes through preprocessing processes such as missing value processing, outlier removal, and data normalization. In the missing value processing step, missing values of wind speed and wind direction data are supplemented through linear interpolation using data from adjacent times, and in the outlier removal step, outliers outside a certain range are removed or adjusted through statistical methods. Finally, in the data normalization step, wind speed data is converted to values between 0 and 1 through Min-Max normalization to increase the efficiency of model learning.

3.2 Data Analysis

Among the existing research models examined in Chapter 2, the SARIMA model, which shows superiority in seasonal data, is effective in analyzing time series data including seasonal patterns, so this study uses the SARIMA model to predict short-term wind speed and wind direction. In addition, as a difference from existing studies, it is designed to have a function to detect abnormal weather based on the difference between the predicted value and the past average value. The SARIMA model can be expressed as SARIMA(p,d,q)(P,D,Q,s), where p is the autoregressive(AR) order, d is the integrated order, q is the non-seasonal order indicating the moving average(MA) order, and P, D, Q, and s represent seasonal components. The model selection selects the optimal parameters by minimizing the AIC(Akaike Information Criterion) value. In this study, in order to differentiate it from other previous studies, the concept of abnormal climate is introduced. If the predicted wind speed value is significantly different from the average of the past 10 years, it is considered an abnormal climate. When the predicted value deviates from ±2 standard deviations of the past average, it is determined to be an abnormal climate. This is a method commonly used in climate analysis using statistical techniques.

4. Experiments and Results

4.1 Experimental Environment

In this study, we use 10-minute wind direction and wind speed data from 2014 to 2023 in Region G to predict wind speed and direction based on the SARIMA model, and develop a system that detects abnormal weather and sends warning notifications. The environment is configured as follows. The hardware specifications include a desktop PC with Intel(R) Core(TM) i7 2.10 GHz, 32 GB RAM, and Windows 11 OS. Python 3.12.4 is used as the programming language, and libraries such as pandas, numpy, statsmodels, matplotlib, and smtplib are installed for data preprocessing and analysis.

4.2 Experimental Results

In order to use the wind direction and wind speed data for 10 years at 10-minute intervals, the meteorological data saved as CSV files are preprocessed to handle missing values, remove outliers, and normalize data. Figure 1 shows the code and execution results for checking the 10-year data presented above.

```
file_path = "wind_data.csv"
data = pd.read_csv(file_path)
print(data.head())
```

```
0  2014  1  1  0  10   235   2    5.2
1  2014  1  1  0  20   231   1.8  4.0
2  2014  1  1  0  30   237   2.2  4.6
3  2014  1  1  0  40   243   2.3  6.0
4  2014  1  1  0  50   245   2.9  6.8
```

Fig. 1. Data Verification Code

Looking at the execution results above, you can see that the year, month, day, hour, minute, wind direction at 10M above ground, wind speed at 10M above ground, and maximum wind speed at 10M above ground are normally retrieved. Figure 2 shows the code and execution results for checking the number and type of data retrieved. Looking at the execution results above, you can see that the total number of records for wind direction and wind speed data for 10-minute periods from 2014 to 2023 is 524469, and that data on wind direction, wind speed, and maximum wind speed at 10M above ground are collected by year, month, day, hour, and minute. In addition, you can check the data type and the head value of the corresponding data for each column.

```
data['날짜'] = pd.to_datetime(
    data[['year', 'month', 'day']].rename(columns=
{'year': 'year', 'month': 'month', 'day': 'day'}))
print(data.info())
print(data.head())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 524469 entries, 0 to 524468
Data columns (total 9 columns):
 #   Column          Non-Null Count   Dtype
---  ------          --------------   -----
 0   year            524469 non-null  int64
 1   month           524469 non-null  int64
 2   day             524469 non-null  int64
 3   시              524469 non-null  int64
 4   분              524469 non-null  int64
 5   지상10M 풍향    524469 non-null  object
 6   지상10M 풍속    524469 non-null  object
 7   지상10M 최대풍속 524469 non-null  float64
 8   날짜            524469 non-null  datetime64[ns]
dtypes: datetime64[ns](1), float64(1), int64(5), object(2)
None
0  2014  1  1  0  10  235  2    5.2 2014-01-01
1  2014  1  1  0  20  231  1.8  4.0 2014-01-01
2  2014  1  1  0  30  237  2.2  4.6 2014-01-01
3  2014  1  1  0  40  243  2.3  6.0 2014-01-01
4  2014  1  1  0  50  245  2.9  6.8 2014-01-01
```

Fig. 2. Code to check the number and type of data

```
data =
data.groupby(data.index).mean(numeric_only=True)
data = data.asfreq('D')
wind_speed = data['지상10M 풍속']
model = SARIMAX(wind_speed, order=(1, 1, 1),
seasonal_order=(1, 1, 1, 12))
results = model.fit()
forecast = results.forecast(steps=30)
print(forecast)
```

Fig. 3. Code to calculate predictions for wind speed

Fig. 3 is a code that calculates the predicted value for the wind speed at 10M above ground by applying the SARIMA model that considers seasonal factors to predict the wind speed for the next 30 days.

Fig. 4 is a code that calculates the predicted value for the wind direction at 10M above ground by applying the SARIMA model that considers seasonal factors to predict the wind direction for the next 30 days.

```
data =
data.groupby(data.index).mean(numeric_only=True)
data = data.asfreq('D')
wind_direction = data['지상10M 풍향']
model = SARIMAX(wind_direction, order=(1, 1, 1),
seasonal_order=(1, 1, 1, 12))
results = model.fit()
forecast = results.forecast(steps=30)
print(forecast)
```

Fig. 4. Code to calculate predictions for wind direction

Fig. 5 shows the results of the SARIMA model-based prediction of wind speed at 10M above ground for the next 30 days, based on data from the past 10 years. The wind speed predicted using the SARIMA model varies within the range of 2.2 to 2.5 m/s, and can be said to be a type that appears more in coastal areas than inland areas.

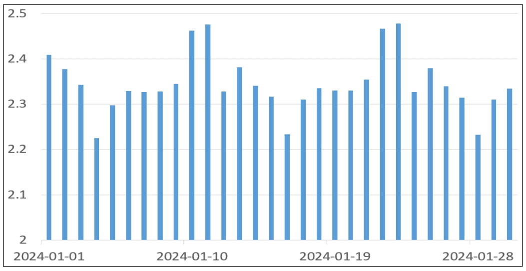

Fig. 5. Forecast of wind speed over the next 30 days

Fig. 6 shows the results of the wind direction at 10M above ground for the next 30 days predicted using the SARIMA model based on data from the past 10 years. The wind direction predicted using the SARIMA model is distributed between 270° and 290°, and the wind direction in the region is mainly from the west to northwest. It shows a tendency to be maintained stably without major fluctuations, and although there are some changes on certain days through the predicted values, the overall direction is consistent.

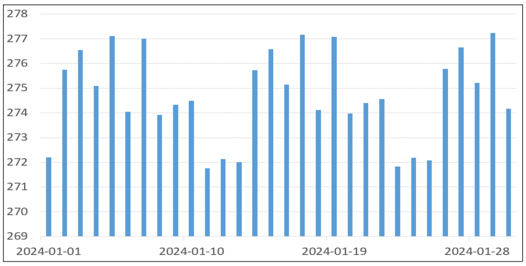

Fig. 6. Forecast of wind direction for the next 30 days

5. Conclusions

The wind speed and wind direction prediction results using the SARIMA-based model proposed in this study by collecting 10 years of wind direction and wind speed data in the G region showed that the wind speed prediction value ranged from about 2.2 to 2.5 m/s, and the stability was confirmed to be well-matched with seasonal changes. The wind direction prediction value was concentrated at

270° to 290°, and appeared in the west to northwest. This can be used as basic data to evaluate the general weather conditions in the region and analyze the difference between the normal pattern and abnormal climate. Through comparative analysis with the normal weather conditions presented in this study, if the prediction value deviates from the normal range or a sudden change in the wind direction and wind speed pattern is detected, an abnormal climate warning can be issued, and damage caused by abnormal climate can be minimized by notifying the surrounding area.

REFERENCES

[1] COP16 Overview: Biodiversity Protection and the Oceans. Name of Site or Board. https://globalgreens.org/kor/cbd-brief-oceans/

[2] Park. H. K. (2018). Air pollution and climate change: Effects on asthmatic patients. *Allergy Asthma Respir Dis, 6(2)*, 79-84. DOI : 10.4168/aard.2018.6.2.79

[3] Jung. Y. J. (2023). A Study on Changes in Disease Patterns and Countermeasures due to Global Warming. *Jour. of KoCon.a, 23(8)*, 426-438. DOI : 10.5392/JKCA.2023.23.08.426

[4] Jang. J. Y. (2009). Climate change, global warming's impact on human health. *Horizon of knowledge, 6*, 159-175.

[5] Cho H. M. (2012). Climate change and air pollution effect on respiratory and allergic disease in Korea. Korea Centers for Disease Control and Prevention, 1-7.

[6] Erasmo. C. & Wilfrido. R. (2007). Wind speed forecasting in the South Coast of Oaxaca, Mexico. *Renewable Energy, 32(12)*, 2116-2128. DOI : 10.1016/j.renene.2006.10.005

[7] Cadenas, E., Jaramillo, O. A., & Rivera, W. (2010). Analysis and forecasting of wind velocity in chetumal, quintana roo, using the single exponential smoothing method. *Renewable Energy, 35(5),* 925-930. DOI : 10.1016/j.renene.2009.10.037

[8] Jeong. C. S., Kim. M. M. & Heo. J. H. (2008). Analysis of Climate Change using Stochastical Methods (based on precipitation data). *2008 Korea Water Resources Association,* 1001-1006.

[9] Song. K. Y., Bang. C. H., Park. Y. S. & Choi. Y. J. (2012). Research and Analysis for Developing of Evaluation on the Site Selection of Wind Farm. *Journal of the Wind Engineering Institute of Korea, 16(1),* 3-12.

[10] Saurabh. S. S., Hamidreza. Z., Om. M. & Paras. M. (2010). A review of wind power and wind speed forecasting methods with different time horizons. *North American Power Symposium 2010,* 1-8. DOI : 10.1109/NAPS.2010.5619586.

[11] Choi. S. H. (2024). A Study on Trend Using Time Series Data. *Advanced Industrial SCIence, 3(1).* 17-22. DOI : 10.23153/AI-Science.2024.3.1.017

[12] Han, K. H., & Na, W. S. (2024). Forecasting Prices of Major Agricultural Products by Temperature and Precipitation. *Journal of Advanced Technology Convergence, 3(2),* 17-23. DOI : 10.23152/JATC.2024.03.02.017

[13] Lee, J., Han, H., & Yoon, S. (2020). Air passenger demand forecasting for the Incheon airport using time series models. *Journal of Digital Convergence, 18(12),* 87-95. DOI : 10.14400/JDC.2020.18.12.087

[14] Choudhary, A., Jain, P., & Prajesh, A. (2023, February). Wind Power Forecasting Using Deep Learning Method: A Review. *In 2023 1st International Conference on Intelligent Computing and Research Trends (ICRT)* (pp. 1-6). IEEE. DOI : 10.1109/ICRT57042.2023.10146688

[15] Liu, X., Lin, Z., & Feng, Z. (2021). Short-term offshore wind speed forecast by seasonal ARIMA-A comparison against GRU and LSTM. *Energy, 227,* 120492. DOI : 10.1016/j.energy.2021.120492

최 신 형(Shin-Hyeong Choi) [종신회원]

· 2002년 8월 : 경남대학교 컴퓨터공학과(공학박사)
· 2003년 9월 ~ 현재 : 강원대학교 전기제어계측공학부 교수
· 관심분야 : 임베디드시스템, 사물인터넷, 정보보안, 모바일컴퓨팅
· E-Mail : cshinh@kangwon.ac.kr

이 춘 수(Lee-Choon Soo) [정회원]

· 2023년 2월 : 강원대학교 방재시스템전공(공학박사)
· 2007년 5월 ~ 현재 : 한국수력원자력(주) 차장
· 관심분야 : 모바일컴퓨팅. 방사능방재, 사물인터넷
· E-Mail : choonsoo@khnp.co.kr