Acta Tecnología - International Scientific Journal about Technologies
Volume: 11 2025 Issue: 3 Pages: 127-136 ISSN 2453-675X

https://doi.org/10.22306/atec.v11i3.290 Received: 16 May 2025; Revised: 14 July 2025; Accepted: 12 Sep. 2025

Impact of weather conditions on forecasting the number of road accidents in Poland

Piotr Gorzelanczyk
Stanislaw Staszic State University of Applied Sciences in Pila, Podchorazych 10 Street, 64-920 Pila, Poland, EU,
https://orcid.org/0000-0001-9662-400X, piotr.gorzelanczyk@ans.pila.pl

Keywords: traffic accident, pandemic, weather conditions, forecasting, exponential smoothing.

Abstract: The incidence of road traffic accidents in Poland remains significantly high. When compared to the average levels recorded across European Union (EU) member states, Poland's rate is approximately 1.3 times greater. The COVID-19 pandemic contributed to a temporary reduction in road accidents. The primary objective of this study is to forecast the number of road accidents in Poland under varying weather conditions. The analysis is based on monthly accident data for the years 2007-2021, obtained from national police records. Using this historical data, projections for the years 2022-2024 were developed. The findings suggest that, despite minor downward trends, the level of road traffic incidents is likely to remain comparable to pre-pandemic patterns. It is important to note, however, that the pandemic has introduced distortions that may affect the accuracy of the forecast. The predictive analysis was conducted in Statistica using selected time series models. Forecasts for 2022-2024 reveal that the majority of accidents are expected to occur under favorable weather conditions, with an average of 24,342 incidents. This is likely associated with increased traffic volume during such conditions. Nevertheless, extreme weather events pose heightened risks: heavy rainfall can lead to as many as 4,347 accidents, while strong winds may contribute to up to 12,880 incidents. Additionally, intense sunlight—through reduced visibility—accounts for an average of 4,150 accidents annually. Although fog and snowfall are less frequent (averaging 109 and 352 incidents respectively), they represent particularly hazardous conditions due to compromised road traction.

1 Introduction

Road accidents are incidents that result not only in injuries or fatalities among road users but also in significant property damage. According to the World Health Organization (WHO), approximately 1.3 million people lose their lives each year due to road traffic crashes. In many countries, road accidents account for nearly 3% of their gross domestic product (GDP). Furthermore, they remain the leading cause of death among children and young adults aged 5 to 29 [1]. In response to this global challenge, the United Nations General Assembly has adopted an ambitious goal: to reduce road traffic deaths and injuries by half by the year 2030.

The severity of a traffic accident is a key factor in assessing its consequences. Accurate prediction of accident severity is crucial for authorities aiming to develop effective traffic safety policies designed to prevent crashes, minimize fatalities and injuries, and reduce material losses [2,3]. Identifying critical determinants of accident severity is a fundamental step in formulating countermeasures that can mitigate or eliminate severe outcomes [4]. In this context, Yang et al. proposed a multi-node deep neural network (DNN) framework to predict varying levels of injury severity, fatalities, and property damage. This approach facilitates a comprehensive and precise assessment of accident severity [5].

Accident data are typically obtained from various sources, most commonly collected by governmental agencies. These data are gathered from police accident reports, insurance company records, and hospital databases. Once collected, the information is often compiled and analyzed at a broader scale within the transport sector [6].

Today, intelligent transportation systems (ITS) play a vital role in the collection and analysis of road accident data. Modern vehicles equipped with GPS devices generate continuous data streams [7]. Additionally, microwave vehicle detection systems installed on roadways capture information such as speed, traffic flow, and vehicle types [8]. License plate recognition technologies further expand the capacity to gather traffic data over time [9]. Social media can also serve as an unconventional data source, though the accuracy of this information is often limited due to the inexperience or unreliability of individuals posting such content [10].

In order to ensure the accuracy and reliability of traffic accident analyses, it is essential to integrate and reconcile data from multiple sources. The combination of heterogeneous datasets significantly enhances the quality of the resulting insights [11].

Vilaca et al. [12] conducted a statistical analysis aimed at evaluating accident severity and establishing relationships between crash types and the characteristics of those involved. Their study offered recommendations for improving traffic safety standards and related public policies. Similarly, Bak et al. [13] analyzed traffic safety in a selected region of Poland by examining accident counts and underlying causes. The research employed multivariate statistical methods to assess the safety behaviors of those responsible for accidents.

The selection of accident data sources for analysis largely depends on the specific traffic safety problem under consideration. Incorporating statistical models along with real-world driving data or intelligent traffic system outputs helps improve the accuracy of forecasts and supports effective accident prevention strategies [14].

A variety of forecasting methods are found in the literature. Among them, time series models are the most widely used for predicting accident counts [15,16]. However, such methods have limitations, including the inability to evaluate the quality of prior forecasts and the presence of autocorrelated residuals [17]. Procházka et al. [18] applied a multi-seasonality model for forecasting, while Sunny et al. [19] utilized the Holt-Winters exponential smoothing technique. Nevertheless, these models often lack the capacity to incorporate exogenous variables [20,21].

Vector autoregressive (VAR) models, though effective, require a large number of observations to estimate parameters reliably [22]. Other autoregressive techniques have also been used, such as the models by Monedero et al. [23] and Al-Madani [24], which were applied to fatality prediction. Regression curve-fitting models are also employed, though they typically assume simple linear relationships [25] and require that the series be stationary with an appropriate autoregressive order [26].

Biswas et al. [27] employed a Random Forest regression model for traffic accident prediction. This method, while capable of handling groups of correlated features, often favors smaller feature subsets and is sensitive to variations in input, which may result in instability and overfitting in peak predictions [28,29]. Chudy-Laskowska and Pisula [30] tested an autoregressive quadratic trend model, a univariate periodic trend model, and an exponential smoothing model. Moving average models, though simple, suffer from limitations such as low accuracy, data loss within sequences, and poor handling of trends and seasonality [31].

Procházka and Camej [32] implemented the GARMA method, which imposes constraints on parameter space to ensure stationarity. ARMA models are commonly used for stationary series, while ARIMA and SARIMA are suited for non-stationary data [19,32-34]. These models are highly flexible but can be challenging to identify and interpret without extensive expertise [35]. Furthermore, ARIMA models are inherently linear, which limits their applicability to nonlinear phenomena [20].

Chudy-Laskowska and Pisula [36] also used ANOVA for forecasting accident counts. While informative, this method requires additional assumptions, such as sphericity, the violation of which may lead to misleading conclusions [37]. Artificial neural networks (ANNs) are increasingly utilized for forecasting in this domain. However, they demand specialized knowledge, are sensitive to initial conditions, and are often criticized as "black-box" models due to their lack of interpretability [36,38,39].

A novel approach involves using the Hadoop framework, as demonstrated by Kumar et al. [40], although this method is not well suited for small datasets [41]. Karlaftis and Vlahogianni [34] applied the GARCH model, which, while powerful, is computationally complex and difficult to calibrate [42,43]. McIlroy and colleagues used the Augmented Dickey-Fuller (ADF) test [44], although its effectiveness can be limited by low statistical power when detecting autocorrelation in the error terms [45].

Data mining techniques have also been used to forecast accident frequency [46,47], though they may produce overly general or ambiguous outputs due to the volume of unstructured data [48,49]. Some researchers, including Sebe et al. [50], propose combining multiple models to enhance predictive performance. Parametric models are another alternative, as outlined by Bloomfield [51].

Research on the impact of weather on forecasting the number of traffic accidents in Poland is important because weather significantly affects driving safety. Downpours, snow or fog increase the risk of collisions, so better forecasts will help alert drivers and services. This will reduce the cost of accidents - rescues, traffic jams and road repairs.

Today, the climate is changing - violent storms and glaze are more common in Poland, so up-to-date data is needed. Police and road managers could use such forecasts to better manage traffic, such as sending more patrols to dangerous areas.

In addition, Poland can learn from Western countries, where weather warning systems are more developed. New technologies, such as artificial intelligence, will help predict dangers more accurately. As a result, safety and traffic flow will improve.

Therefore, the purpose of this article is to develop an adaptive model for forecasting the number of traffic accidents in Poland, taking into account meteorological variables such as precipitation, temperature and visibility, in order to improve the accuracy of predictions.

2 Methodology

A substantial number of fatalities continue to occur on Polish roads. Although a gradual year-over-year decline has been observed, the overall figure remains alarmingly high. While the COVID-19 pandemic temporarily reduced the number of road accidents, the incidence remains significantly elevated. A monthly analysis of the data reveals distinct fluctuations, though the overall trend indicates a continued decrease in accident frequency. Despite this, Poland still reports a considerably higher number of accidents compared to the European Union average.

The data also show that the lowest number of accidents tends to occur during foggy conditions, likely due to more cautious driver behavior and reduced speeds. In contrast, the highest accident rates are observed during clear weather conditions, possibly due to increased traffic volume and riskier driving behavior. These findings highlight the importance of implementing targeted measures to further reduce accident rates and to identify the specific conditions under which accidents are most likely to occur (Figure 1).

The literature review and methodology for this study proceeded as follows: Based on statistical data provided by the Polish Police concerning road accident frequency under varying weather conditions, a set of forecasts was generated using 15 exponential smoothing models within the Statistica software environment. Subsequently, forecast errors were calculated for each model. The model yielding the lowest forecast error was selected as the most accurate predictor of future accident trends.

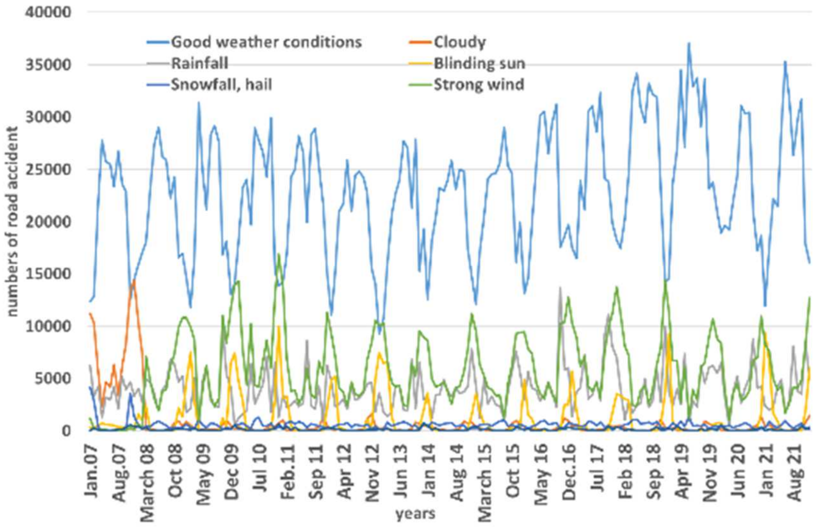

Figure 1 Number of accidents in Poland from 2007 to 2021

The variability in the number of road accidents under different weather conditions was assessed using the Kruskal-Wallis test. In this analysis, the test statistic was 133.8816 with a significance level of p = 0.000. These results provide strong evidence to reject the null hypothesis of equal mean accident levels across weather conditions. Thus, it can be concluded that the average number of road accidents significantly differs depending on the prevailing weather.

Furthermore, the analysis confirms a consistent decline in the average number of road accidents over the studied period. There is also a distinct variation in accident frequency relative to weather conditions: the highest number of accidents occurs during favorable weather, while the lowest number is observed in adverse weather conditions. This trend is likely attributable to more cautious driving behavior during poor weather, such as reduced speed and heightened attention (Figure 2).

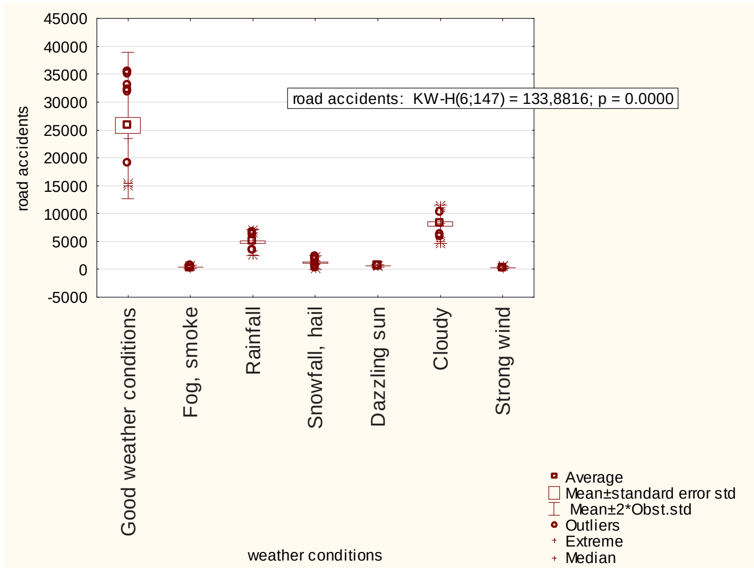

Figure 2 Comparison of the average number of road accidents in Poland by weather conditions

Based on the analysis of the number of road accidents in Poland, it can be concluded that they are seasonal in nature with a downward trend. Therefore, for further analysis, selected time series models were used to determine the projected number of road accidents in the analyzed period depending on the prevailing weather conditions. Weather conditions were divided into the following categories:

- good weather conditions,
- fog, smoke,
- precipitation,
- snowfall, hail,
- blinding sun,
- cloudy,
- strong wind.

3 Forecasting the number of road accidents

To forecast the number of traffic accidents, selected exponential smoothing models were employed. This method involves representing the time series of the forecasted variable using a weighted moving average, where the weights decline exponentially for past observations. These weights were optimally determined by the Statistica software, in which the analysis was conducted. The forecast is thus based on a weighted average of both current and historical data points. The accuracy of the forecast produced using this method is highly dependent on the selection of the specific model and its parameter settings.

The following errors of expired forecasts determined from equations (1-5) were used to calculate measures of analytical forecasting perfection:

- ME – mean error

$$ME = \frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_p) \quad (1)$$

- MAE –mean everage error

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_i - Y_p| \quad (2)$$

- MPE –mean percentage error

$$MPE = \frac{1}{n}\sum_{i=1}^{n}\frac{Y_i - Y_p}{Y_i} \quad (3)$$

- MAPE - mean absolute percentage error

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|Y_i - Y_p|}{Y_i} \quad (4)$$

- MSE – mean square error

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - Y_p)^2 \quad (5)$$

where:
n – the length of the forecast horizon,
Y – observed value of road accidents,
Y_p – forecasted value of road accidents.

In order to compare the number of accidents during a pandemic and if it did not exist, the mean absolute percentage error was minimized. To forecast the number of accidents depending on the prevailing weather conditions, data from the Polish Police from 2007-2021 was used. The forecast results for the prevailing weather conditions are shown in Figures 3-9. The different forecasting methods used in the study are coded M1, M2,......, Mn. The forecasting techniques used in the study are as follows:

M1 - moving average method 2-points,
M2 - moving average method 3-points,
M3 - moving average method 4-points,
M4 - exponential smoothing no trend seasonal component: none,
M5 - exponential smoothing no trend seasonal component: additive,
M6 - exponential smoothing no trend seasonal component: multiplicative,
M7 - exponential smoothing linear trend seasonal component: none HOLTA,
M8 - exponential smoothing linear trend seasonal component: additive,
M9 - exponential smoothing linear trend seasonal component: multiplicative WINTERSA,
M10 - exponential smoothing exponential seasonal component: none,
M11 - exponential smoothing exponential seasonal component: additive,
M12 - exponential smoothing exponential seasonal component: multiplicative,
M13 - exponential smoothing fading trend seasonal component: none,
M14 - exponential smoothing fading trend seasonal component: additive,
M15 - exponential smoothing fading trend seasonal component: multiplicative).

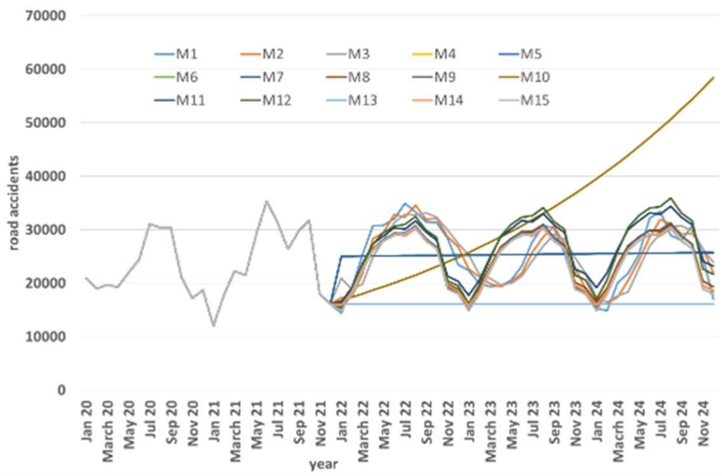

Figure 3 Forecasting the number of traffic accidents during good weather conditions from 2022 to 2024

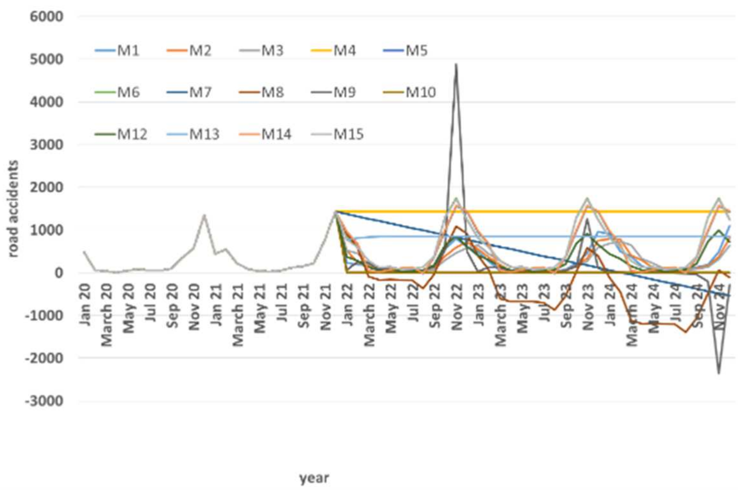

Figure 4 Forecasting the number of traffic accidents during cloud cover from 2022 to 2024

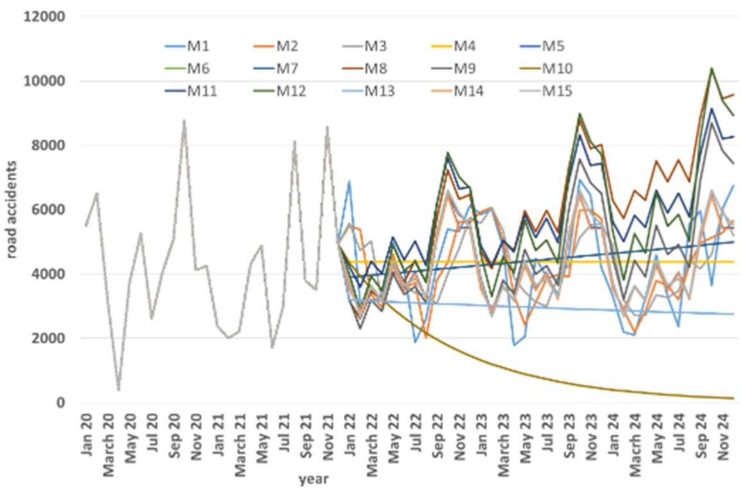

Figure 5 Forecasting the number of accidents during rainfall from 2022 to 2024

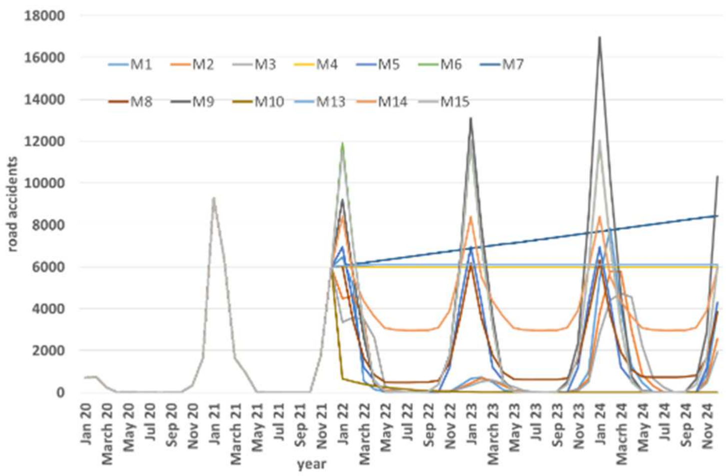

Figure 6 Forecasting the number of traffic accidents during solar glare from 2022 to 2024

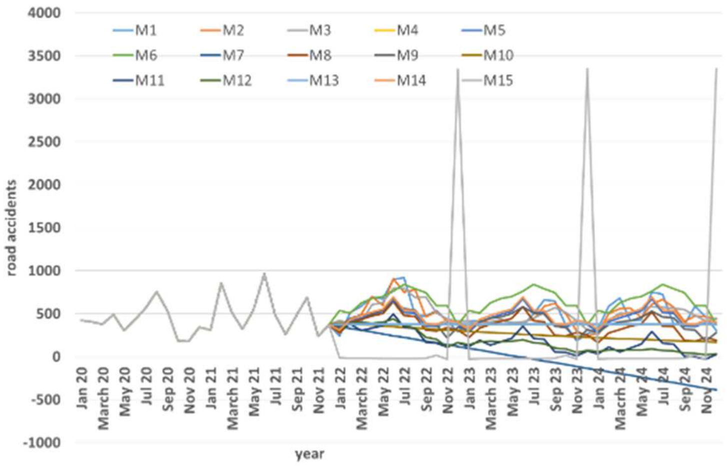

Figure 7 Forecasting the number of road accidents during snow and hailstorms from 2022 to 2024

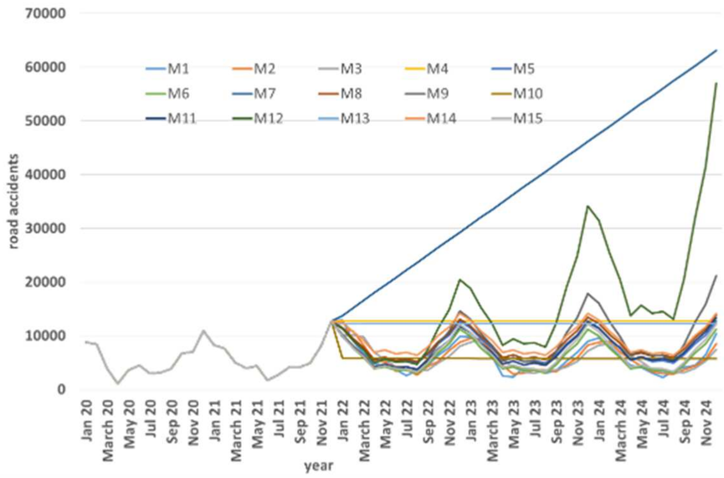

Figure 8 Forecasting the number of road accidents during high winds from 2022 to 2024

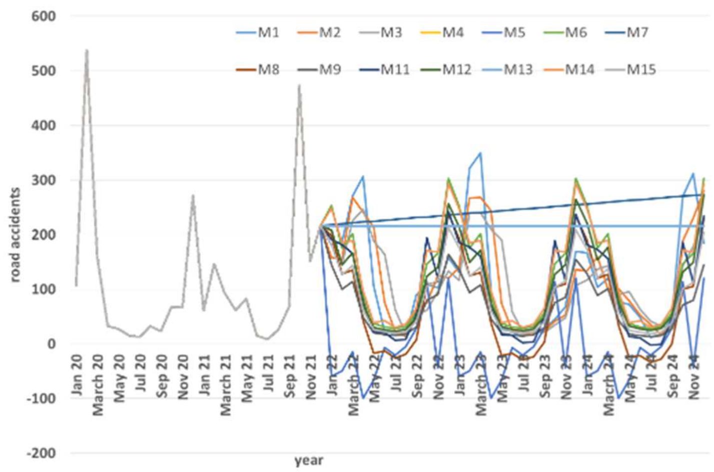

Figure 9 Forecasting the number of traffic accidents during fog, smoke in 2022-2024

Based on the results obtained, it can be concluded that not all of the applied forecasting methods proved effective for the case under study. In the subsequent step, the forecasting model with the lowest Mean Absolute Percentage Error (MAPE) was identified and presented as the most accurate method. The following methods were selected as the best forecasting methods for each day of the week:

- Good weather conditions - M5
- Overcast - M12
- Rainfall - M10
- Dazzling sunshine - M14

- Snowfall, hail - M8
- Strong wind - M1
- Fog, smoke - M12

The analysis of the obtained data suggests that the effectiveness of the forecasting method depends on the prevailing weather conditions. In most cases, the lowest MAPE values were achieved using exponential smoothing techniques. Based on these results, a forecast of the number of traffic accidents under varying weather conditions was generated and is presented in Figure 10 and Table 1. The corresponding forecast errors are summarized in Table 2.

The results indicate that traffic accident levels are expected to remain similar to those observed prior to the COVID-19 pandemic, with only a slight decline. However, it is important to acknowledge that the pandemic has introduced anomalies that may have affected the accuracy of the forecasting models. A MAPE value of 10% or lower is considered indicative of an effective forecasting method. An exception was observed in the case of sun glare, for which all methods produced relatively high error values, likely due to greater variability and unpredictability in such conditions.

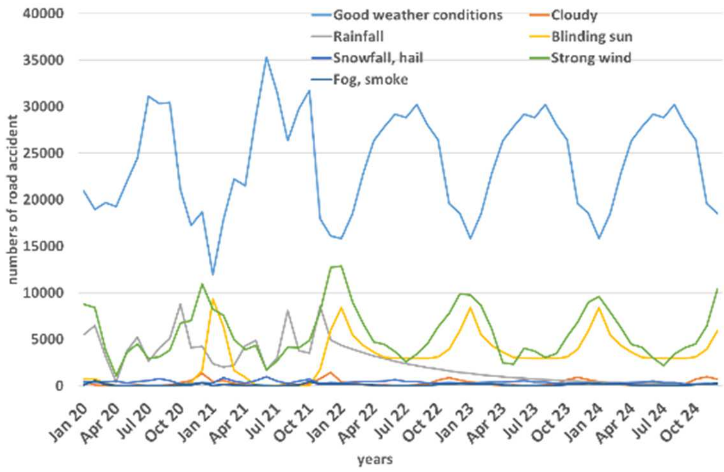

Figure 10 Optimal forecast number of road accidents depending on weather conditions in 2022-2024

Table 1 Forecast values

Road accidents	Good weather conditions	Cloudy	Rainfall	Blinding sun	Snowfall. hail	Strong wind	Fog. smoke
max	30224.98	990.85	4346.5	8417.73	639.59	12879.5	272.68
min	15830.15	36.28	134.9	2960.12	164.23	2176.5	23.6
Average	24342.424	300.5264	1242.331	4150.047	351.6436	5746.472	108.9183

Table 2 Forecast errors

forecast error/ weather conditions	Good weather conditions	Cloudy	Rainfall	Blinding sun	Snowfall. hail	Strong wind	Fog. smoke
ME	291.44	311.48	1727.71	24.32	1.11	258.47	51.01
MPE	2461.46	434.77	2113.51	935.34	290.30	2271.72	71.61
MSE	10166119.83	1217794.35	7987983.38	2554231.66	243922.08	11503810.33	15540.12
MAPE [%]	0.37	3.36	20.40	155.35	6.60	10.24	0.02
MAE [%]	11.58	78.16	57.05	952.11	110.42	42.42	60.37

The road accident forecast for 2022-2024 shows that most incidents occur in good weather conditions (24,342 on average), due to higher traffic volumes. However, extremes are dangerous: during a downpour (max. 4,347 accidents) and strong winds (max. 12,880), the risk rises sharply. The blinding sun causes an average of 4,150 accidents, mainly through reduced visibility. Fog and snow, although less frequent (109 and 352 accidents on average), are particularly dangerous due to low grip.

Based on the study, it can be concluded that the largest errors are observed in rainfall and strong winds, indicating that the model has difficulty predicting accidents in extreme conditions. The smallest errors are observed in good weather conditions and fog/smoke, suggesting that the model is reasonably accurate under stable conditions.

Detailed analysis by weather conditions:

a) Good weather conditions

Low MAPE (0.37%) and MAE (11.58%) - the model performs very well.

High MPE (2461.46) - may indicate systematic underestimation of accidents.

b) Overcast

Moderate MAPE (3.36%), but high MAE (78.16%) - model performs well under overcast, but absolute errors are significant.

c) Rainfall

Very high MSE (7987983.38) and MAPE (20.40%) - model has great difficulty predicting accidents during rain.

High ME (1,727.71) - strong overestimation of the number of accidents.

(d) Blinding sun

Extremely high MAPE (155.35%) and MAE (952.11%) - the model completely fails under such conditions.

Low ME (24.32) - errors may be tolerable, but the spread is huge (high MSE).

(e) Snow and hail

Relatively low MSE (243922.08) and MAPE (6.60%), but high MAE (110.42%) - the model partially copes, but absolute errors are large.

(f) Strong winds

High MSE (11503810.33) and MAPE (10.24%) - as with rain, the model has trouble forecasting.

High MPE (2271.72) - strong overestimation.

g) Fog and smoke

Lowest MAPE (0.02%)

Taking into account the problems that occur during forecasting, the following conclusions can be drawn: Extreme conditions (blinding sun, rainfall, strong wind) lead to large forecast error and High MPE and MSE indicate the need for model calibration, especially for precipitation. Analyzed the model worked well during good weather, fog.

The model performs best under stable weather conditions, while the largest errors occur under precipitation, strong winds and blinding sun. It is necessary to optimize the algorithms for these phenomena to improve forecast accuracy.

4 Conclusion

The forecast of the number of road accidents in Poland was developed using exponential smoothing methods implemented in the Statistica software. The weighting parameters were optimized by the program to minimize both the Mean Absolute Error (MAE) and the Mean Absolute Percentage Error (MAPE).

The findings indicate that the number of traffic accidents is likely to remain at a level similar to that observed before the COVID-19 pandemic, with a slight downward trend. However, it should be emphasized that the pandemic introduced irregularities that may have affected the reliability of the results. In most cases, a forecast error of no more than 10% confirms the effectiveness of the selected forecasting methods—except for one outlier.

The traffic accident forecast presented in this study can serve as a foundation for future policy development aimed at reducing accident rates in the analyzed regions. An example of such a measure is the introduction of stricter penalties for traffic violations in Poland, which came into effect on January 1, 2022.

In their future research, the authors plan to incorporate additional variables that influence accident frequency in Poland. These factors may include traffic volume, day of the week, and the age of the accident perpetrator, among others.

References

[1] WHO: The Global status on road safety, [Online], Available: https://www.who.int/publications/i/item/9789241565684 [24 Aug 2024], 2018.

[2] TAMBOURATZIS, T., SOULIOU, D., CHALIKIAS, M., GREGORIADES, A.: Maximising accuracy and efficiency of traffic accident prediction combining information mining with computational intelligence approaches and decision trees, *Journal of Artificial Intelligence and Soft Computing*, Vol. 4, No. 1, pp. 31-42, 2014.

[3] ZHU, L., LU, L., ZHANG, W., ZHAO, Y., SONG, M.: Analysis of Accident Severity for Curved Roadways Based on Bayesian Networks, *Sustainability*, Vol. 11, No. 8, 2223, pp. 1-17, 2019. https://doi.org/10.3390/su11082223

[4] ARTEAGA, C., PAZ, A., PARK, J.: Injury severity on traffic crashes: A text mining with an interpretable machine-learning approach, *Safety Science*, Vol. 132, 104988, pp. 1-12, 2020. https://doi.org/10.1016/j.ssci.2020.104988

[5] YANG, Z., ZHANG, W., FENG, J.: Predicting multiple types of traffic accident severity with explanations: A multi-task deep learning framework, *Safety Science*, Vol. 146, 105522, pp. 1-13, 2022. https://doi.org/10.1016/j.ssci.2021.105522.

[6] GORZELANCZYK, P., PYSZEWSKA, D., KALINA, T., JURKOVIC, M.: Analysis of road traffic safety in the Pila poviat, *Scientific Journal of Silesian University of Technology. Series Transport*, Vol. 107, pp. 33-52, 2020. https://doi.org/10.20858/sjsutst.2020.107.3

[7] CHEN, C.: *Analysis and forecast of traffic accident big data*, ITM Web Conf., Vol. 12, 04029, pp. 1-6, 2017. https://doi.org/10.1051/itmconf/20171204029

[8] KHALIQ, K.A., CHUGHTAI, O., SHAHWANI, A., QAYYUM, A., PANNEK, J.: Road accidents detection, data collection and data analysis using V2X communication and edge/cloud computing, *electronics*, Vol. 8, No. 8, 896, pp. 1-28, 2019. https://doi.org/10.3390/electronics8080896

[9] RAJPUT, H., SOM, T., KAR, S.: An automated vehicle license plate recognition system, *Computer* (*Long Beach. Calif*), Vol. 48, No. 8, pp. 56-61, 2015. https://doi.org/10.1109/MC.2015.244

[10] ZHENG, Z., WANG, C., WANG, P., XIONG, Y., ZHANG, F., LV, Y.: Framework for fusing traffic information from social and physical transportation data, *PLoS One*, Vol. 13, No. 8, e0201531, 2018. https://doi.org/10.1371/journal.pone.0201531

[11] ABDULLAH, E., EMAM, A.: *Traffic accidents analyzer using big data*, 2015 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, pp. 392-397, 2015. https://doi.org/10.1109/CSCI.2015.187

[12] VILAÇA, M., SILVA, N., COELHO, M.C.: Statistical analysis of the occurrence and severity of crashes involving vulnerable road users, *Transportation Research Procedia*, Vol. 27, pp. 1113-1120, 2017. https://doi.org/10.1016/j.trpro.2017.12.113

[13] BĄK, I., CHEBA, K., SZCZECIŃSKA, B.: The statistical analysis of road traffic in cities of Poland, *Transportation Research Procedia*, Vol. 39, pp. 14-23, 2019. https://doi.org/10.1016/j.trpro.2019.06.003

[14] CHAND, A., JAYESH, S., BHASI, A.B.: Road traffic accidents: An overview of data sources, analysis techniques and contributing factors, *materialstoday: Proceedings*, Vol. 47, Part 15, pp. 5135-5141, 2021. https://doi.org/10.1016/j.matpr.2021.05.415

[15] HELGASON, A.: Fractional integration methods and short Time series: evidence from a simulation study, *Political Analysis*, Vol. 24, No. 1, pp. 59-68, 2016.

[16] LAVRENZ, S., VLAHOGIANNI, E., GKRITZA, K., KE, Y.: Time series modeling in traffic safetyresearch, *Accident Analysis & Prevention*, Vol. 117, pp. 368-380, 2018.

[17] Forecasting based on time series, [Online], Available: http://pis.rezolwenta.eu.org/Materialy/PiS-W-5.pdf [24 Aug 2024], 2024.

[18] PROCHAZKA, J., CAMAJ, M.: *Modelling the number of road accidents of uninsured drivers and their severity*, Proceedings of International Academic Conferences 5408040, International Institute of Social and Economic Sciences, 2017.

[19] SUNNY, C.M., NITHYA, S., SINSHI, K.S., VINODINI, V.M.D., LAKSHMI, A.K.G., ANJANA, S., MANOJKUMAR, T.K.: *Forecasting of Road Accident in Kerala: A Case Study*, Environmental Science, Computer Science, 2018 International Conference on Data Science and Engineering (ICDSE), Kochi, India, pp. 1-5, 2018. https://doi.org/10.1109/ICDSE.2018.8527825

[20] DUDEK, G.: *Forecasting Time Series with Multiple Seasonal Cycles Using Neural Networks with Local Learning*, In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds) Artificial Intelligence and Soft Computing, ICAISC 2013, Lecture Notes in Computer Science(), Vol. 7894, Springer, Berlin, Heidelberg, pp. 52-63, 2013. https://doi.org/10.1007/978-3-642-38658-9_5

[21] SZMUKSTA-ZAWADZKA, M., ZAWADZKI, J.: *Forecasting on the basis of Holt-Winters models for complete and incomplete data*, Research papers of the Wrocław University of Economics, Wroclaw, Poland, 2009.

[22] WÓJCIK, A.: *Autoregressive vector models as a response to the critique of multi-equation structural econometric models,* Publishing house of the University of Economics in Katowice, Katowice, Poland, 2014.

[23] MONEDEROA, B.D., GIL-ALANAA, L.A., MARTÍNEZAA, M.C.V.: Road accidents in Spain: Are they persistent?, *IATSS Research*, Vol. 45, No. 3, pp. 317-325, 2021. https://doi.org/10.1016/j.iatssr.2021.01.002

[24] AL-MADANI, H.: Global road fatality trends' estimations based on country-wise microlevel data, *Accident Analysis & Prevention*, Vol. 111, pp. 297-310, 2018. https://doi.org/10.1016/j.aap.2017.11.035

[25] MAMCZUR, M.: Machine learning How does linear regression work? And is it worth using?, [Online], Available: https://miroslawmamczur.pl/jak-dziala-regresja-liniowa-i-czy-warto-ja-stosowac/ [10 May 2025], 2022.

[26] PIŁATOWSKA, M.: The choice of the order of autoregression depending on the parameters of the generating model, *Econometrics*, Vol. 4, No. 38, 2012.

[27] BISWAS, A.A., MIA, J., MAJUMDER, A.: *Forecasting the Number of Road Accidents and Casualties using Random Forest Regression in the Context of Bangladesh*, July 2019 Economics 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2019.

[28] Random forest, [Online], Available: https://pl.wikipedia.org/wiki/Las_losowy [10 May 2025], 2024.

[29] FIJOREK, K., MRÓZ, K., NIEDZIELA, K., FIJOREK, D.: Forecasting electricity prices on the day-ahead market using data mining methods, *Energy Market*, Vol. 12, 2010.

[30] CHUDY-LASKOWSKA, K., PISULA, T.: Forecast of the number of road accidents in Poland, *Logistics*, Vol. 6, 2014.

[31] KASHPRUK, N.: *Comparative research of statistical models and soft computing for identification of time series and forecasting*, Opole University of Technology, Opole, Poland, 2010.

[32] PROCHÁZKA, J., FLIMMEL, S., ČAMAJ, M., BAŠTA, M.: *Modelling the Number of Road Accidents*, 20-th AMSE, Applications of Mathematics and Statistics in Economics, International Scientific Conference: Szklarska Poręba, 30 August-3 September 2017, Conference Proceedings Full Text Papers, pp. 355-364, Wrocław, Poland, 2017. https://doi.org/10.15611/amse.2017.20.29

[33] DUTTA, B., BARMAN, M.P., PATOWARY, A.N.: Application of Arima model for forecasting road accident deaths in India, *International Journal of Agricultural and Statistical Sciences*, Vol. 16, No. 2, pp. 607-615, 2020.

[34] KARLAFTIS, M., VLAHOGIANNI, E.: Memory properties and fractional integration in transportation time-series, *Transportation Research*, Vol. 17, No. 4, pp. 444-453, 2009. https://doi.org/10.1016/j.trc.2009.03.001

[35] ŁOBEJKO, S.: *Time series analysis and forecasting with SAS*, Main business school in Warsaw, Warsaw, Poland, 2015.

[36] CHUDY-LASKOWSKA, K. PISULA, T.: Forecasting the number of road accidents in Subcarpathia, *Logistics*, Vol. 2015, No. 4, 2015.

[37] Road safety assessment handbook, [Online], Available: https://www.iung.pl/PJA/wydane/11/PJA 11_3.pdf [11 May 2025], 2024.

[38] WROBEL, M.S.: *Application of neural fuzzy systems in chemistry*, PhD thesis, Katowice, University of Silesia, Katowice, Poland, 2017.

[39] Data mining techniques StatSoft, [Online], Available: https://www.statsoft.pl/textbook/stathome_stat.html?https%3A%2F%2Fwww.statsoft.pl%2Ftextbook%2Fstdatmin.html [11 May 2025], 2024.

[40] KUMAR, S., VISWANADHAM, V., BHARATHI, B.: Analysis of road accident, *IOP Conference Series Materials Science and Engineering*, Vol. 59, No. 1, 012029, pp. 1-6, 2019. https://doi.org/10.1088/1757-899X/590/1/012029

[41] Top Advantages and Disadvantages of Hadoop 3 DataFlair, [Online], Available: https://data-flair.training/blogs/advantages-and-disadvantages-of-hadoop/ [11 May 2025], 2024.

[42] PERCZAK, G., FISZEDER, P.: GARCH model - using additional information on minimum and maximum prices, *Bank and Credit*, Vol. 45, No. 2, pp. 105-132, 2014.

[43] FISZEDER, P.: *GARCH class models in empirical financial research*, Scientific Publishers of the Nicolaus Copernicus University, Torun, Poland, 2009.

[44] MCILROY, R.C., PLANT, K.A., HOQUE, M.S., WU, J., KOKWARO, G.O., NAM, V.H., STANTON, N.A.: Who is responsible for global road safety? A cross-cultural comparison ofactor maps, *Accident Analysis & Prevention*, Vol. 122, pp. 8-18, 2019. https://doi.org/10.1016/j.aap.2018.09.011

[45] MUCK, J.: Econometrics, Modeling of time series, Stationary, Unit root tests, ARDL models, Co-integration, [Online], Available: http://web.sgh.waw.pl/~jmuck/Ekonometria/EkonometriaPrezentacja5.pdf [11 May 2025], 2024.

[46] SHETTY, P., SACHIN, P.C., KASHYAP, V.K., MADI, V.: Analysis of road accidents using data mining techniques, *International Research Journal of Engineering and Technology*, Vol. 04, No. 4, pp. 1494-1496, 2017.

[47] LI, L., SHRESTHA, S., HU, G.: *Analysis of road traffic fatal accidents using data mining techniques*, 2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA), pp. 363-370, 2017. https://doi.org/10.1109/SERA.2017.7965753

[48] MARCINKOWSKA, J.: *Statistical methods and data mining in assessing the occurrence of syncope in the group of narrow-QRS tachycardia (AVNRT and AVRT)*, Medical University of Karol Marcinkowski in Poznań, Poznań, Poland, 2015.

[49] BOCZEK, M., HOVANA, A., HUTNÍK, O.: General Chebyshev type inequality for seminormed fuzzy integral, In: Torra, V., Narukawa, Y., Pasi, G., Viviani, M. (eds) Modeling Decisions for Artificial Intelligence, MDAI 2019, Lecture Notes in Computer Science(), Vol. 11676, Springer, Cham., pp. 3-16, 2019. https://doi.org/10.1007/978-3-030-26773-5_1

[50] SEBEGO, M., NAUMANN, R.B., RUDD, R.A., VOETSCH, K., DELLINGER, A.M., NDLOVU, C.: The impact of alcohol and road traffic policies on crash rates in Botswana, 2004–2011: A time-series analysis, *Accident Analysis & Prevention*, Vol. 70, pp. 33-39, 2014. https://doi.org/10.1016/j.aap.2014.02.017

[51] BLOOMFIELD, P.: An exponential model in the spectrum of a scalar time series, *Biometrika*, Vol. 60, No. 2, pp. 217-226, 1973. https://doi.org/10.2307/2334533

Review process

Single-blind peer review process.