Forecasting the number of road accidents in Poland by province

Prognozowanie liczby wypadków drogowych w Polsce według województw

Piotr Gorzelanczyk

Dr inż.

Akademia Nauk Stosowanych
im. Stanisława Staszica w Pile

piotr.gorzelanczyk@ans.pila.pl
ORCID: https://orcid.org/0000-0001-9662-400X

Streszczenie: Każdego roku na polskich drogach ginie bardzo duża liczba osób. Z roku na rok wartość ta spada, ale liczba ta nadal jest bardzo wysoka. Pandemia znacznie zmniejszyła liczbę wypadków drogowych, ale wartość ta nadal jest bardzo wysoka. Z tego powodu należy dowiedzieć się, w których województwach dochodzi do największej liczby wypadków drogowych oraz poznać prognozę wypadków na najbliższe lata, aby móc zrobić wszystko, aby tę liczbę zminimalizować. Celem artykułu jest sporządzenie prognozy liczby wypadków drogowych w Polsce w podziale na województwa. W tym celu przeanalizowano miesięczne dane dotyczące liczby wypadków w Polsce w latach 2007-2021 pochodzące ze statystyk Policji oraz dokonano prognozy na lata 2022-2024. Na podstawie uzyskanych danych można stwierdzić, że pandemia spowodowała spadek liczby wypadków drogowych w Polsce średnio o 21%. Rozrzut w zależności od województwa waha się w przedziale: 10% dla województwa lubuskiego do prawie 53% dla województwa lubelskiego. Spadek jest najbardziej zauważalny w województwach lubelskim, wielkopolskim i małopolskim. Ponadto prognozy pokazują, że w obecnej sytuacji możemy spodziewać się dalszego spadku liczby wypadków drogowych w Polsce. Wyniki badania pokazują, że nadal możemy spodziewać się podobnego poziomu wypadków drogowych jak przed pandemią z minimalnym spadkiem na polskich drogach, ale panująca pandemia zniekształca uzyskane wyniki. Do prognozowania liczby wypadków drogowych wykorzystano szeregi czasowe i modele wykładnicze.

Słowa kluczowe: *Wypadek Drogowy; Pandemia; Polska; Prognozowanie; Wygładzanie Wykładnicze*

Abstract: Every year a very large number of people die on Polish roads. From year to year, the value decreases, but the number is still very high. The pandemic has significantly reduced the number of road accidents, but the value is still very high. For this reason, it is necessary to find out which provinces have the highest number of traffic accidents and to know the accident forecast for the coming years, so that we can do everything possible to minimize this number.

The purpose of the article is to make a forecast of the number of road accidents in Poland by province. For this purpose, monthly data on the number of accidents in Poland in 2007-2021 from the statistics of the Police were analyzed, and a forecast was made for 2022-2024. Based on the data obtained, it can be said that the pandemic caused a decrease in the number of road accidents in Poland by an average of 21%. The spreads depending on the province sniff in the range: 10% for Lubuskie Voivodeship to almost 53% for Lubelskie Voivodeship. The decrease is most noticeable in the Lubelskie, Wielkopolskie and Małopolskie provinces. In addition, forecasts show that in the current situation we can expect a further decrease in the number of road accidents in Poland. The results of the study show that we can still expect a similar level of road accidents as before the pandemic with a minimal decrease on Polish roads, but the prevailing pandemic distorts the results obtained. Time series and exponential models were used to forecast the number of traffic accidents.

Keywords: *Traffic Accident; Pandemic; Poland; Forecasting; Exponential Smoothing*

Introduction

Road traffic accidents are events that cause not only injuries or death to road users, but also damage to property. According to the World Health Organization (WHO), approximately 1.3 million people die each year as a result of traffic accidents. Traffic accidents account for around 3% of their GDP for most of the countries in the world. Road traffic accidents are the leading cause of death for minors and young people aged 5-29 [1]. The UN General Assembly has set an ambitious goal of halving the number of road deaths and injuries by 2030.

The extent of a traffic accident is an attribute for determining its severity. Predicting the severity of accidents is important for competent authorities when designing transport safety policies to eliminate accidents, reduce injuries, deaths and property losses [2,3]. The identification of critical factors that affect the severity of accidents is a precondition for taking countermeasures to eliminate and mitigate the severity of accidents [4]. Yang et al. proposes a DNN (Deep Neutral Network) multi-

-carbon framework to predict different levels of severity of injury, death and property loss. It allows a comprehensive and accurate analysis of the severity of traffic accidents [5].

There are several sources of accident data. They are mostly collected and analyzed by government authorities through the relevant government agencies. Data collection is carried out through police reports, insurance databases or hospital records. Partial traffic accident information is subsequently processed for the transport sector on a larger scale [6].

Intelligent transportation systems are currently the most important source of data related to the analysis and prediction of traffic accidents. The data can be processed due to the use of GPS devices in vehicles [7]. Microwave vehicle detection systems at roadsides can continuously record vehicle data (speed, traffic volume, vehicle type, etc.) [8]. The Vehicle License Plate Recognition system also makes it possible to collect large amounts of traffic data over a monitored period [9]. Another source of data for obtaining traffic and accidents information can be social media, but their relevance may be insufficient due to the incompetence of reporters [10].

For the relevance of accident data, it is necessary to work with several data sources that need to be confronted correctly. The combination of different data sources by consolidating heterogeneous traffic accident data helps to increase the accuracy of the analysis results [11].

A statistical survey aimed at assessing the severity, finding out the connection between traffic accidents and road users was performed by Vilaca et al [12]. The result of the study is a proposal to improve road safety standards and the adoption of other policies related to transport safety.

Bak et al. [13] conducted a statistical survey of traffic safety in a selected region of Poland based on the number of traffic accidents, the pace of finding out the causes of their occurrence. The survey applied a multidimensional statistical analysis to examine safety aspect of persons responsible for accidents.

The choice of the source of accident data for the analysis depends on the type of traffic problem being addressed. The combination of statistical models with other natural driving data or other data obtained through intelligent transport systems contributes to increasing the accuracy of accident forecasts and contributes to their elimination [14].

Various methods of forecasting the number of accidents can be found in the literature. Most often, time series methods are used for forecasting the number of road traffic accidents [15-16], the disadvantages of which are the impossibility of assessing the quality of forecast on the basis of expired forecasts and the often-occurring autocorrelation of the residual component [17]. Procházka et al. [18] used the multiple seasonality model for forecasting and Sunny et al. [19] used the Holt-Winters exponential smoothing method. Its limitations include the inability to introduce exogenous variables into the model [20,21].

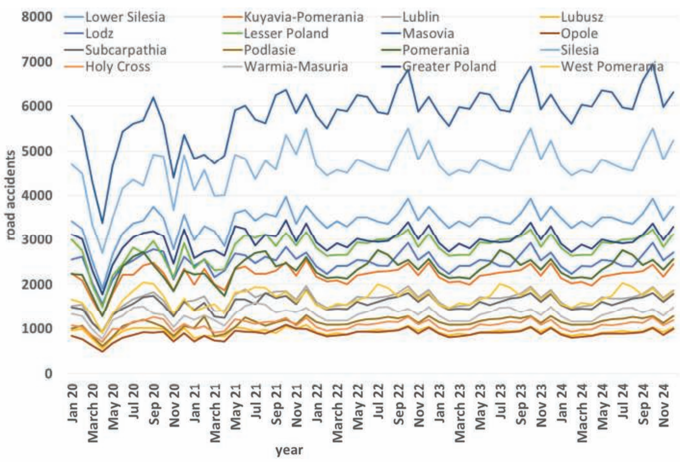
For forecasting the number of road accidents, the vector autoregression model has also been used, whose drawback is the need to have a large number of observations of the variables in order to correctly estimate their parameters [22] as well as the autoregression models of Monederoa et al. for analysing the number of fatalities [23] and Al-Madani, curve-fitting regression models [24]. These, in turn, require only simple linear relationships [25] and the order of the autoregression (assuming that the series are already stationary) [26].

Biswas et al. [27] used Random Forest regression to predict the number of road accidents. In this case, the data contain groups of correlated features with similar significance to the original

Tab. 1. *Population of Poland from 2007 to 2020 [52]*

Province	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Lower Silesia	2 877 059	2 876 627	2 917 242	2 916 577	2 914 362	2 909 997	2 908 457	2 904 207	2 903 710	2 902 547	2 901 225	2 900 163	2 891 321
Kuyavia-Pomerania	2 067 918	2 069 083	2 098 711	2 098 370	2 096 404	2 092 564	2 089 992	2 086 210	2 083 927	2 082 944	2 077 775	2 072 373	2 061 942
Lublin	2 161 832	2 157 202	2 178 611	2 171 857	2 165 651	2 156 150	2 147 746	2 139 726	2 133 340	2 126 317	2 117 619	2 108 270	2 095 258
Lubusz	1 008 962	1 010 047	1 023 215	1 023 158	1 023 317	1 021 470	1 020 307	1 018 075	1 017 376	1 016 832	1 014 548	1 011 592	1 007 145
Lodz	2 548 861	2 541 832	2 542 436	2 533 681	2 524 651	2 513 093	2 504 136	2 493 603	2 485 323	2 476 315	2 466 322	2 454 779	2 437 970
Lesser Poland	3 287 136	3 298 270	3 336 699	3 346 796	3 354 077	3 360 581	3 368 336	3 372 618	3 382 260	3 391 380	3 400 577	3 410 901	3 410 441
Masovia	5 204 495	5 222 167	5 267 072	5 285 604	5 301 760	5 316 840	5 334 511	5 349 114	5 365 898	5 384 617	5 403 412	5 423 168	5 425 028
Opole	1 033 040	1 031 097	1 017 241	1 013 950	1 010 203	1 004 416	1 000 858	996 011	993 036	990 069	986 506	982 626	976 774
Subcarpathia	2 099 495	2 101 732	2 127 948	2 128 687	2 129 951	2 129 294	2 129 187	2 127 657	2 127 656	2 129 138	2 129 015	2 127 164	2 121 229
Podlasie	1 191 470	1 189 731	1 203 448	1 200 982	1 198 690	1 194 965	1 191 918	1 188 800	1 186 625	1 184 548	1 181 533	1 178 353	1 173 286
Pomerania	2 219 512	2 230 099	2 275 494	2 283 500	2 290 070	2 295 811	2 302 077	2 307 710	2 315 611	2 324 251	2 333 523	2 343 928	2 346 671
Silesia	4 645 665	4 640 725	4 634 935	4 626 357	4 615 870	4 599 447	4 585 924	4 570 849	4 559 164	4 548 180	4 533 565	4 517 635	4 492 330
Holy Cross	1 272 784	1 270 120	1 282 546	1 278 116	1 273 995	1 268 239	1 263 176	1 257 179	1 252 900	1 247 732	1 241 546	1 233 961	1 224 626
Warmia-Masuria	1 427 073	1 427 118	1 453 782	1 452 596	1 450 697	1 446 915	1 443 967	1 439 675	1 436 367	1 433 945	1 428 983	1 422 737	1 416 495
Greater Poland	3 397 617	3 408 281	3 446 745	3 455 477	3 462 196	3 467 016	3 472 579	3 475 323	3 481 625	3 489 210	3 493 969	3 498 733	3 496 450
West Pomerania	1 692 957	1 693 198	1 723 741	1 722 739	1 721 405	1 718 861	1 715 431	1 710 482	1 708 174	1 705 533	1 701 030	1 696 193	1 688 047

1. Location of provinces in Poland [53]



2. Comparison of the number of accidents in Poland from 2007 to 2021 [54]

data, smaller groups are favoured over larger ones [28], and there is instability in the method and spike prediction [29]. Chudy-Laskowska and Pisula [30] used the autoregressive model with quadratic trend, the univariate periodic trend model and the exponential equalization model for the forecasting issue discussed. A moving mean model can also be used for forecasting the discussed issue, the disadvantages of which are low forecast accuracy, loss of data in the sequence, lack of consideration of trends and seasonal effects [31]. Prochozka and Camej [32] used the GARMA method, in which some restrictions are imposed in the parameter space to guarantee the stationarity of the process. Very often the ARMA model for a stationary process or ARIMA or SARIMA for a non-stationary process is used for forecasting [19,32-34]. These models result in very high flexibility of the discussed models, but it is also their disadvantage, as good model identification requires more experience from the researcher than, for example, regression analysis [35]. Another disadvantage is the linear nature of the ARIMA model [36].

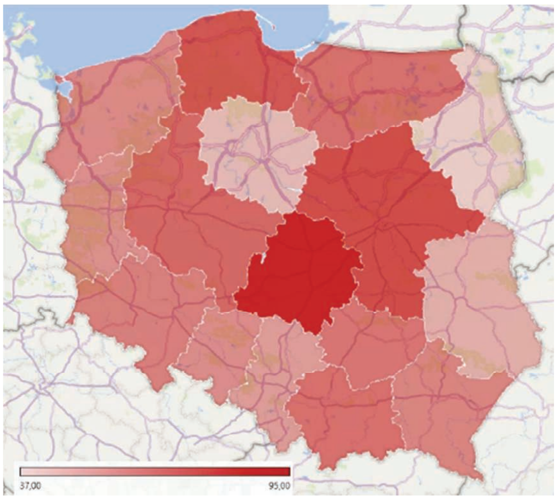
Chudy-Laskowska and Pisula in their work [37] used the ANOVA method to forecast the number of road crashes. The disadvantage of this method is the adoption of additional assumptions, especially the assumption of sphericity, the violation of which may lead to erroneous conclusions [38]. Neural network models are also used to forecast the number of road accidents. The disadvantage of ANN is the need for experience in this field [37, 39] and the dependence of the final solution on the initial conditions of the network, as well as the lack of interpretability in the traditional way since ANN is usually referred to as blackbox where you give input and the model gives output without any knowledge about the analysis [40].

Tab. 2. Area, population by province in Poland in 2020 [52]

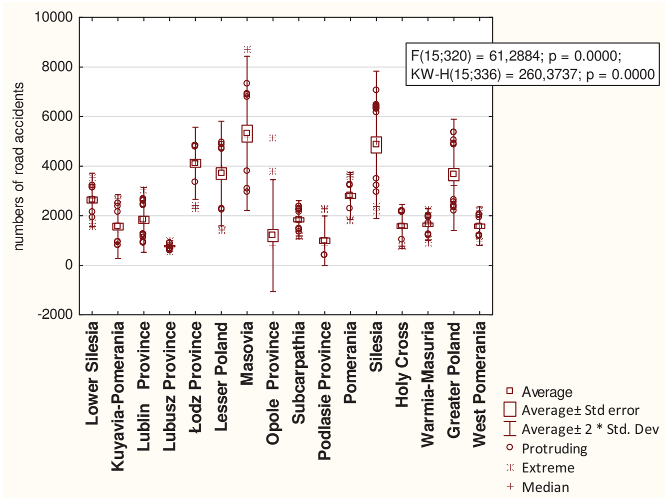
Province	Area		Population	
	in ha	in km^2	total	per 1 km^2
Poland	**31270525**	**312705**	**38265013**	**122**
Lower Silesia	1994670	19947	2891321	145
Kuyavia-Pomerania	1797134	17971	2061942	115
Lublin	2512246	25123	2095258	83
Lubusz	1398793	13988	1007145	72
Lodz	1821895	18219	2437970	134
Lesser Poland	1518279	15183	3410441	225
Masovia	3555847	35559	5425028	153
Opole	941187	9412	976774	104
Subcarpathia	1784576	17846	2121229	119
Podlasie	2018702	20187	1173286	58
Pomerania	1832368	18323	2346671	128
Silesia	1233309	12333	4492330	364
Holy Cross	1171050	11710	1224626	105
Warmia-Masuria	2417347	24173	1416495	59
Greater Poland	2982650	29826	3496450	117
West Pomerania	2290472	22905	1688047	74

A new prediction method is the use of the Hadoop model by Kumar et al. [41]. The drawback of this method is its inability to work with small data files [42]. Karlaftis and Vlahogianni [34] used the Garch model for prediction. The disadvantage of this method is its complex form and complicated model [43,44]. On the other hand, McIlroy and his team used the ADF test [45] which has the disadvantage of poor power in the case of autocorrelation of the random component [46].

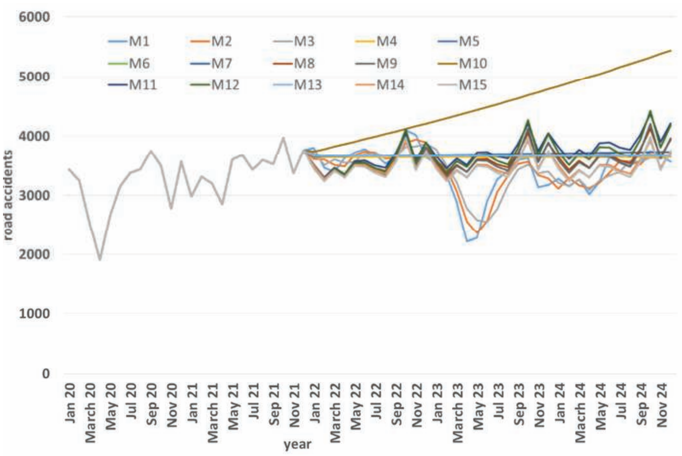
Authors of publications [47-48] have also used Data-Mining techniques for forecasting, which usually have the disadvantage of huge sets of general descriptions [49]. The combination of models proposed by Sebego et al. as a combination of different models is also encountered [50] Parametric models are also proposed in the work of Bloomfield [51]. Taking into account the above analysis, the author made forecasts of the number of road accidents in Poland by province. Selected



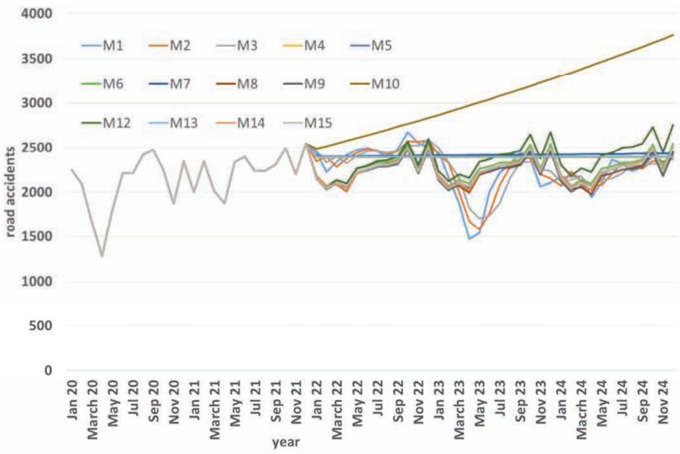
3. Accident rate per 100,000 population in 2021 [54]



4. Average number of road accidents by province from 2007 to 2021 [54]



5. Forecasting the number of road accidents in the Lower Silesian province from 2022 to 2024



6. Forecasting the number of road accidents in the Kuyavian-Pomeranian province from 2022 to 2024

time series models and exponential models were used to forecast the number of accidents.

Analysis of the seasonality of road accidents

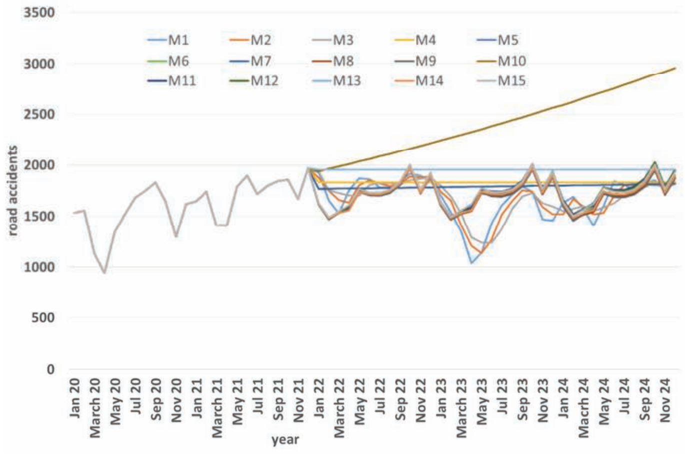
Poland has a population of over 38 million (tab.**1**). Poland covers an area of 312705 km2 and is divided into 16 provinces (tab.**2**, fig.**1**). Based on Police statistics, it can be concluded that the number of road accidents on Polish roads is decreasing from year to year. For all the analyzed provinces in the analyzed period 2007-2021, the average decrease is more than 56%. The most, this is evident in the Kujawsko-Pomorskie (70%) and Podlaskie (69%) provinces, and the least, in the Lubuskie (32%) province. The number of traffic accidents depends on the number of residents living in a particular province. In addition, some provinces have fluctuations in the number of accidents with a downward trend. Compared to the European Union, the number of accidents in Poland is still very high.

During the pandemic, there is a decrease in the number of road accidents. Compared to 2019, there were on average 21% fewer road accidents in 2020, and comparing the statistics of 2021 compared to 2019, the decrease is more than 23%. Figure **2** graphically shows the trend of road accidents in Poland from 2001 to 2021 for the analyzed provinces. The number of road accidents depends on the province. The accident rate per 100,000 residents in 2021 is shown in Figure **3**. The highest number of accidents per 100,000 residents is observed in the Łódzkie (95) and Mazowieckie (79.3) provinces, and the lowest in the Podlaskie (37) and Kujawsko-Pomorskie (38.8) provinces.

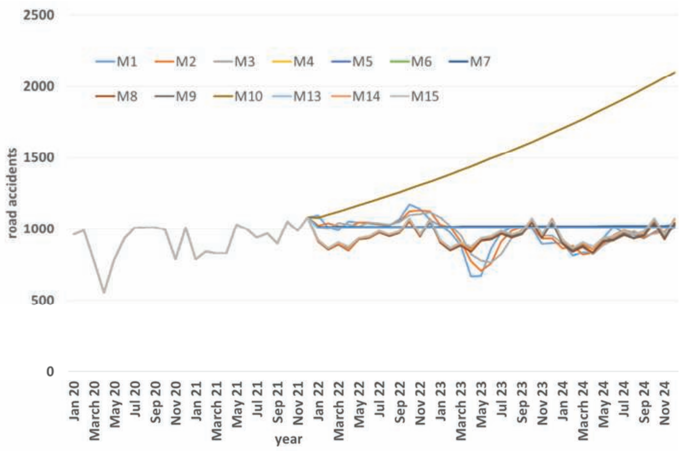
The analysis of the change in the number of traffic accidents by province was examined using the Kruskal-Wallis test. The value of the test statistic is 260.4 with a test probability of p = 0.000. The value obtained indicates the rejection of equality of the average level of road accidents. On this basis, it can be concluded that the number of traffic accidents analyzed, over the years, shows a systematic decrease in the average level of accidents. In addition, depending on the provinces analyzed, there is a clear variation in the number of accidents, as shown in Figure **1**. In Figure **4**, it is clear that the highest number of road accidents occurs in the provinces of Silesia and Lesser Poland, and these provinces should be looked at closely.

Forecasting the number of road accidents

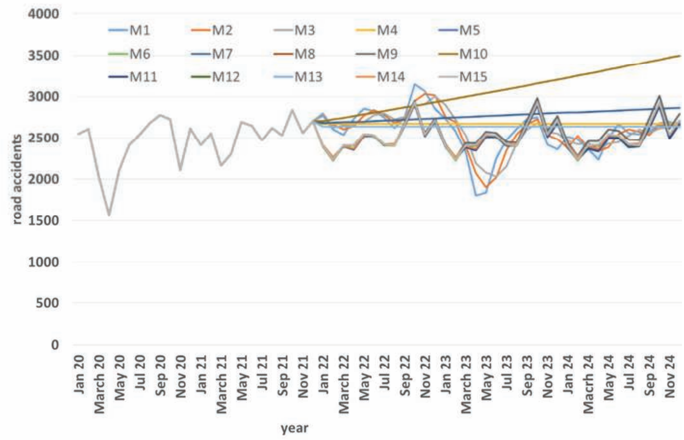
Selected exponential equalization models were used to forecast the number of traffic accidents for provin-



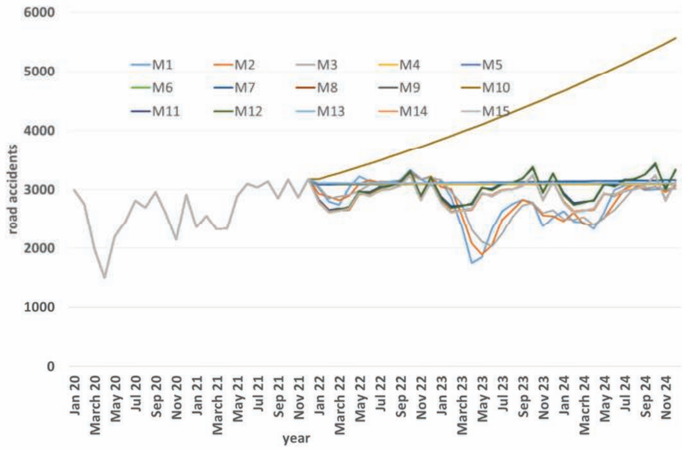
7. Forecasting the number of road accidents in Lublin Province from 2022 to 2024



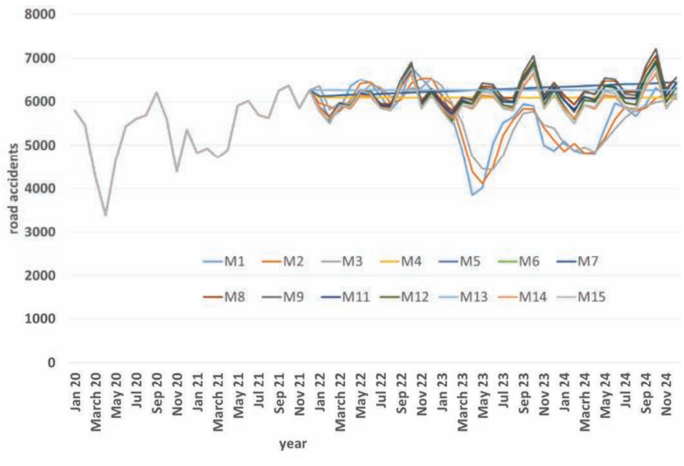
8. Forecasting the number of road accidents in Lubuskie Province from 2022 to 2024



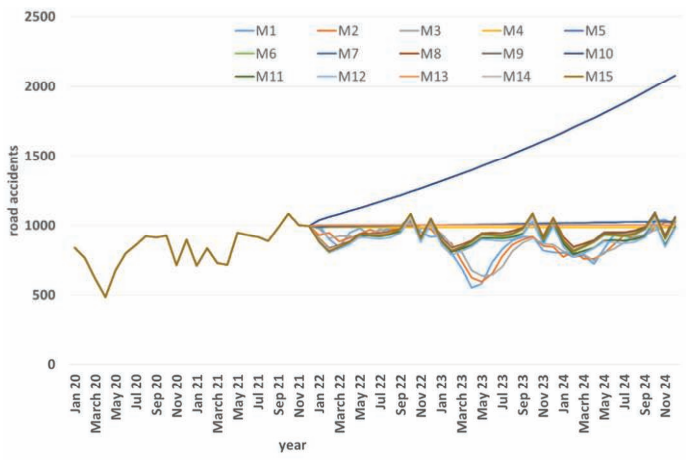
9. Forecasting the number of road accidents in Lodz province in 2022-2024



10. Forecasting the number of road accidents in Małopolska province from 2022 to 2024



11. Forecasting the number of road accidents in the Mazowieckie province from 2022 to 2024



12. Forecasting the number of road accidents in Opole Province from 2022 to 2024

ces. The essence of these methods is that the time series of the forecast variable is pronounced with a weighted moving average, and the weights are determined according to the exponential function. These weights were optimally selected by Statistica software, in which the applied analyses were carried out.

The forecast of the number of accidents for the analyzed provinces was based on a weighted average of the current and historical series. The forecast results obtained using these methods depend on the choice of the model and its optimal parameter values. Forecasting the number of accidents in Poland by occurring provinces, was carried out using selected time series models.

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

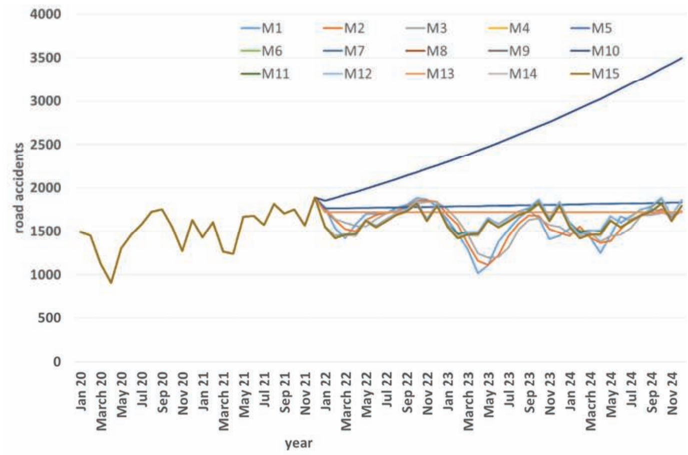
ME – mean error

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

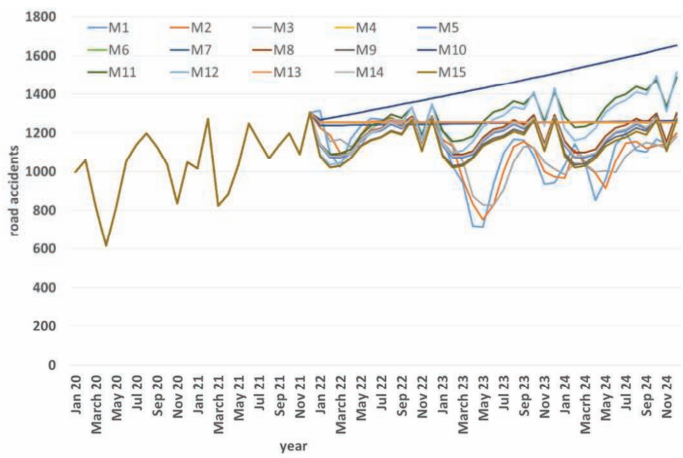
MAE –mean everage error

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

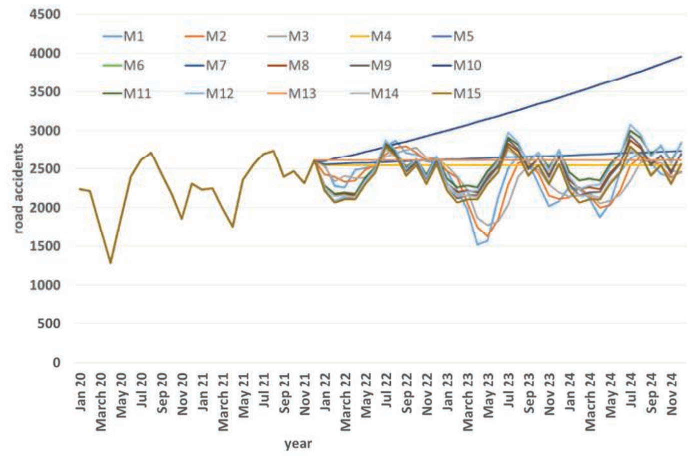
MPE –mean percentage error



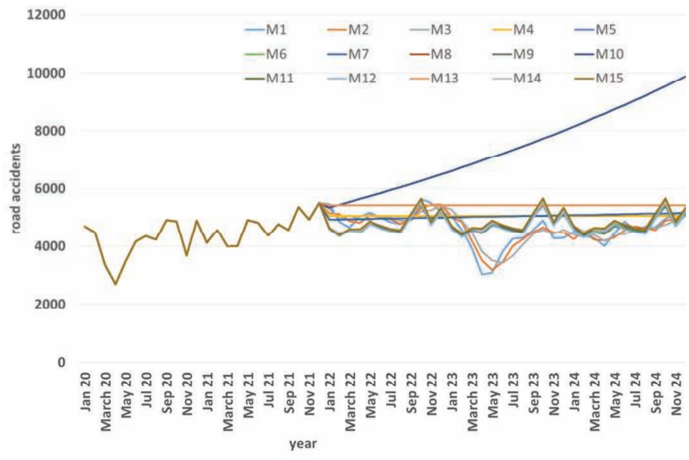
***13**. Forecasting the number of road accidents in Podkarpackie Province from 2022 to 2024*



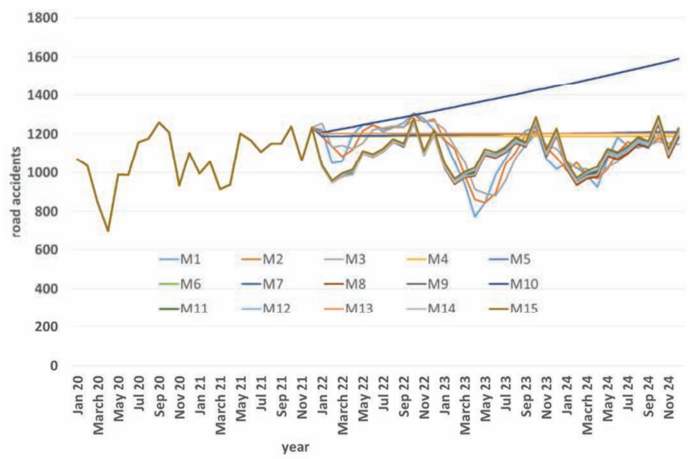
***14**. Forecasting the number of road accidents in Podlaskie Province from 2022 to 2024*



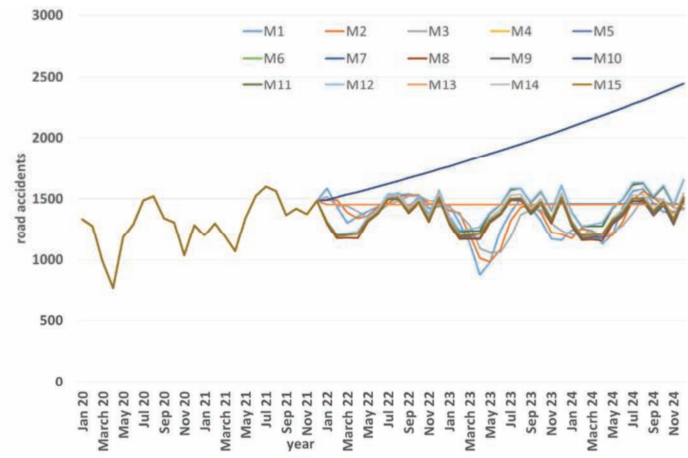
***15**. Forecasting the number of road accidents in Pomeranian Province from 2022 to 2024*



***16**. Forecasting the number of road accidents in Silesia Province from 2022 to 2024*



***17**. Forecasting the number of road accidents in the Świętokrzyskie province from 2022 to 2024*



***18**. Forecasting the number of road accidents in the Warmian-Masurian Voivodeship from 2022 to 2024*

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

MAPE - mean absolute percentage error

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

MSE – mean square error

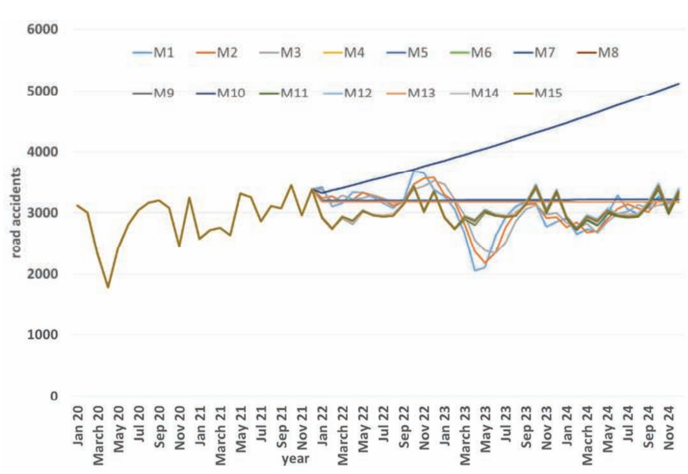
$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - Y_p\right)^2 \qquad (5)$$

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

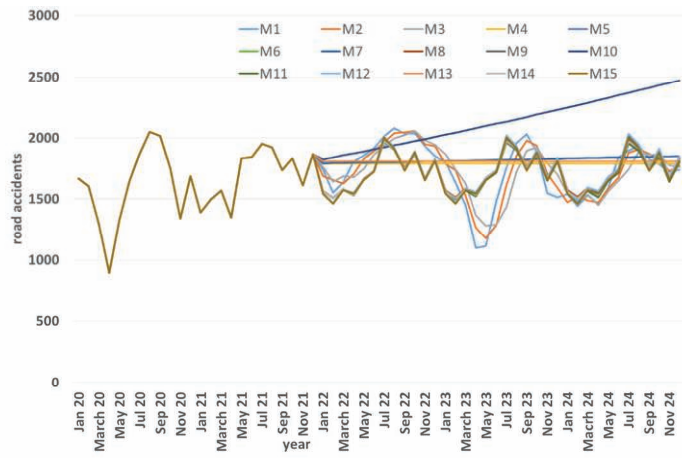
In order to compare the number of accidents during a pandemic and if it did not exist, the mean absolute percentage error was minimized

Forecasting the number of road accidents in Poland

Data from the Polish Police from 2007-2021 were used to forecast the number of accidents in the analyzed provinces. The forecast results for the provinces are shown in Figures **3**-**9**. The different forecasting methods used in the study



***19**. Forecasting the number of road accidents in the Wielkopolska province from 2022 to 2024*



***20**. Forecasting the number of road accidents in the West Pomeranian province from 2022 to 2024*

are coded M1, M2,......, Mn. The forecasting techniques used in the study are:

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,
M10exponential smoothing exponential seasonal component: none,
M11exponential smoothing exponential seasonal component: additive,
M12exponential smoothing exponential seasonal component: multiplicative,
M13exponential smoothing fading trend seasonal component: none,
M14exponential smoothing fading trend seasonal component: additive,
M15exponential smoothing fading trend seasonal component: multiplicative).

Based on the above data, it can be concluded that not all the methods used are effective in the case studied. In the next step, the best forecasting method in which the MAPE error was minimal was presented. The following methods were selected as the best forecasting methods for each province:

- Lower Silesia - M5
- Kujawsko-pomorskie - M8

***Tab. 3.** Forecast errors*

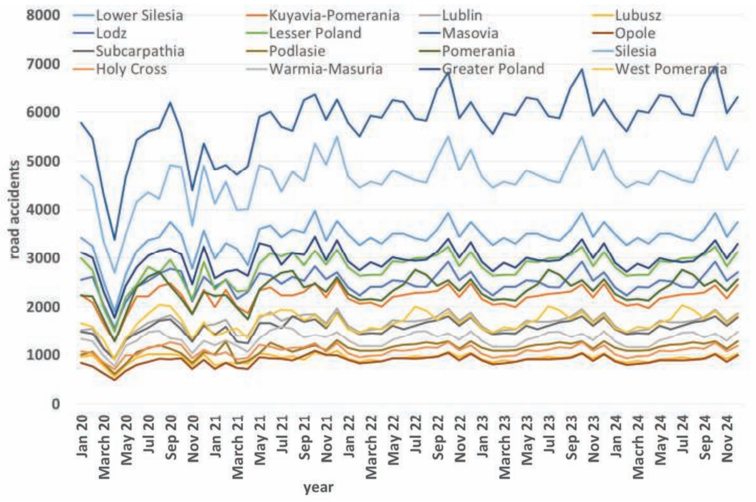
forecast error/ province	ME	MPE	MSE	MAPE [%]	MAE [%]
Lower Silesia	3.208997514	188.4726684	68878.1878	0.450566933	6.441275887
Kuyavia-Pomerania	2.346605332	134.2447066	32446.81837	0.447358531	5.997117204
Lublin	3.38357111	92.95928945	15312.7583	0.323852721	5.703924675
Lubusz	5.954995074	63.73415879	6927.33137	0.06565246	6.668295983
Lodz	12.89057225	125.2129324	30256.14262	0.077675164	5.250496692
Lesser Poland	3.738428028	137.239586	36230.34961	0.337291561	5.205621655
Masovia	33.15728838	312.2327679	211883.9746	0.013442517	6.493443012
Opole	6.120110928	55.08085009	4988.750659	0.106896808	6.999187736
Subcarpathia	3.877532058	82.77703037	13200.35334	0.294025739	5.553336533
Podlasie	0.374843299	73.18912197	12945.62756	0.886450432	7.158647302
Pomerania	12.07251273	109.163129	21988.27088	0.168959493	5.621077124
Silesia	13.92971096	228.1574737	97951.27438	0.183521704	5.424355374
Holy Cross	1.62737223	55.49788666	5278.434034	0.235158676	5.143998883
Warmia-Masuria	2.543996469	76.6130496	9890.936403	0.265121154	5.612666064
Greater Poland	6.07939686	162.031501	47276.38541	0.301951406	5.63669225
West Pomerania	9.094299176	86.47579672	15880.93027	0.025641961	5.823004188

- Lubelskie - M8
- Lubuskie - M12
- Łódź - M12
- Malopolska - M14
- Mazovia - M12
- Opolskie - M11
- Subcarpathian - M8
- Podlasie - M14
- Pomerania - M14
- Silesia - M5
- Holy Cross - M8
- Warmia and Mazury - M8
- Greater Poland - M11
- West Pomerania - M12

Based on the data obtained, it can be concluded that the choice of the method used depends on the province being analyzed. The minimum MAPE error was most common when using exponential and linear trend methods. On this basis, the forecast of the number of traffic accidents by province was determined, shown in Figure **21**, and the obtained forecast errors are presented in Table **3**. The results indicate that we can still expect a similar level of traffic accidents as before the pandemic with a minimal decrease. It should be noted that the pandemic distorted the obtained results [55-57]. An error value of at most 6% demonstrates the choice of an effective forecasting method.

Conclusions

The forecast of the number of acci-



***21**. Optimal forecast number of traffic accidents depending on the analyzed province in 2022-2024*

dents in Poland was determined by exponential equalization methods using the Statistica program. The weights used were estimated by the program to minimize the mean absolute error and mean absolute percentage error.

The results show that we can still expect a similar level of traffic accidents as before the pandemic with a minimal decrease. It should be noted that the pandemic distorted the results obtained. The error value of a maximum of 6%, with one exception, can testify to the choice of an effective forecasting method.

The forecast of the number of traffic accidents obtained in the article, can be used in the future to formulate further measures to minimize the number of accidents in the countries analyzed. These measures may include, for example, the introduction of higher fines for traffic offenses on Polish roads from January 1, 2022.

In their further research, the authors plan to take into account more factors influencing accident rates in Poland. We can include traffic volume, day of the week or age of the accident perpetrator, among others.

Declaration of Competing Interest
None ◀

References

[1] WHO. The Global status on road safety. 2018, 403 pp. ISBN: 9789241565684. Available from: https://www.who.int/publications/i/item/9789241565684

[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 J. Artificial Intell. Soft Comput. Res., 4 (1) (2014), pp. 31-42

[3] Zhu L., Lu L., Zhang W., Zhao Y., Song M. Analysis of accident severity for curved roadways based on bayesian networks. Sustainability, 11 (8) (2019), p. 2223

[4] Arteaga C., Paz A., Park J. Injury severity on traffic crashes: A text mining with an interpretable machine-learning approach Saf. Sci., 132 (2020), Article 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. Volume 146,2022, 105522, ISSN 0925-7535, 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. 2020, 107, 33-52. ISSN: 0209-3324. DOI: https://doi.org/10.20858/sjsutst.2020.107.3.

[7] Chen C. Analysis and forecast of traffic accident big data ITM Web Conf., 12 (2017), p. 04029, 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 Electron, 8 (2019), 10.3390/electronics8080896

[9] Rajput H., T. Som, S. Kar. An automated vehicle license plate recognition system Computer (Long Beach. Calif)., 48 (2015), pp. 56-61, 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., 13 (2018), 10.1371/journal.pone.0201531

[11] Abdullah E., Emam A., Traffic accidents analyzer using big data, in: Proc. - 2015 Int. Conf. Comput. Sci. Comput. Intell. CSCI 2015, Institute of Electrical and Electronics Engineers Inc., 2016: pp. 392–397. 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. Transp. Res. Procedia., 27 (2017), pp. 1113-1120, 10.1016/j.trpro.2017.12.113

[13] Bąk. I., Cheba K., Szczecińska B. The statistical analysis of road traffic in cities of Poland. Transp. Res. Procedia., 39 (2019), pp. 14-23, 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, Materials Today: Proceedings. Volume 47, Part 15, 2021. Pages 5135-5141. ISSN 2214-7853, https://doi.org/10.1016/j.matpr.2021.05.415.

[15] Helgason A., Fractional integration methods and short Time series: evidence from asimulation study, Polit. Anal. 24 (1) (2016) 59–68http://www.jstor.org/stable/24573204.

[16] Lavrenz S., Vlahogianni E., Gkritza K., Ke Y., Time series modeling in traffic safetyresearch, Accid. Anal. Prev. 117 (2018) 368–380

[17] Forecasting based on time series

(2022) http://pis.rezolwenta.eu-.org/Materialy/PiS-W-5.pdf
[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). DOI:10.1109/ICDSE.2018.8527825
[20] 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, No. 38, 2009
[21] Dudek, G. (2013). 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. https://doi.org/10.1007/978-3-642-38658-9_5
[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. Vol. 193.2014.
[23] Monederoa B.D., Gil-Alanaa L.A., Martínezaa M.C.V. Road accidents in Spain: Are they persistent?. IATSS Research. Vol. 45, Issue 3, October 2021, Pages 317-325. 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, Accid. Anal. Prev. 111 (2018) 297–310,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? https://miroslawmamczur.pl/jak-dziala-regresja-liniowa-i-czy-warto-ja-stosowac/
[26] Piłatowska M. The choice of the order of autoregression depending on the parameters of the generating model. Econometrics 4(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)
[28] Random forest (2022) https://pl.wikipedia.org/wiki/Las_losowy
[29] Fijorek K., Mróz K., Niedziela K., Fijorek D. Forecasting electricity prices on the day-ahead market using data mining methods Energy Market "- 12/2010
[30] Chudy-Laskowska K., Pisula T. Forecast of the number of road accidents in Poland. Logistics 6/2014.
[31] Kashpruk N. Comparative research of statistical models and soft computing for identification of time series and forecasting. Opole University of Technology 2010.
[32] Procházka J., Flimmel S., Čamaj M., Bašta M. Modelling the Number of Road Accidents. Publishing house of the University of Economics in Wrocław. Wrocław 2017. doi: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. Int. J. Agricult. Stat. Sci. Vol. 16, No. 2, pp. 607-615, 2020
[34] Karlaftis M., Vlahogianni E. Memory properties and fractional integration in trans-portation time-series, Transp. Res. C 17 (2009) 444–453.
[35] Łobejko S. Time series analysis and forecasting with SAS. Main business school in Warsaw. Warsaw 2015
[36] Dudek G. Exponential smoothing models for short-term power system load forecasting. Energy market No. 3 (106) – 2013.
[37] Chudy-Laskowska K. Pisula T. Prognozowanie liczby wypadków drogowych na Podkarpaciu. Logistics. N0 4/ 2015
[38] Road safety assessment handbook https://www.iung.pl/PJA/wydane/11/PJA11_3.pdf
[39] Wrobel M.S. Application of neural fuzzy systems in chemistry. PhD thesis. Katowice: University of Silesia. 2017.
[40] Data mining techniques StatSoft https://www.statsoft.pl/textbook/stathome_stat.html?https%3A-%2F%2Fwww.statsoft.pl%2Ftextbook%2Fstdatmin.html
[41] Kumar S., Viswanadham V.,Bharathi B. (2019) Analysis of road accident. IOP Conference Series Materials Science and Engineering 590(1):012029. DOI: 10.1088/1757-899X/590/1/012029
[42] Top Advantages and Disadvantages of Hadoop 3 DataFlair https://data-flair.training/blogs/advantages-and-disadvantages-of-hadoop/
[43] Perczak G., Fiszeder P. GARCH model - using additional information on minimum and maximum prices. Bank and Credit. Number 2. 2014
[44] Fiszeder P. GARCH class models in empirical financial research. Scientific Publishers of the Nicolaus Copernicus University. Torun 2009.
[45] 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, Accid. Anal. Prev. 122 (2019) 8–18. https://doi.org/10.1016/j.aap.2018.09.011
[46] Muck J. Econometrics. Modeling of time series. Stationary. Unit root tests. ARDL models. Co-integration. http://web.sgh.waw.pl/~j-muck/Ekonometria/Ekonometria-Prezentacja5.pdf
[47] Shetty P., Sachin P.C., Kashyap V.K., Madi V. Analysis of road accidents using data mining techniques. Inernational Research Journal of Engineering and Technology.

04/2017

[48] 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), 2017, pp. 363-370, doi: 10.1109/SERA.2017.7965753.

[49] 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, 2015 http://www.wbc.poznan.pl/Content/373785/index.pdf

[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:atime-series analysis, Accid. Anal. Prev. 70 (2008) 33–39, 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 60 (1973) 217–226 https://www.jstor.org/stable/2334533.

[52] Central Statistical Office, www.gus.pl (17.04.2022)

[53] Province names in English, https://www.facebook.com/Pogoda-Na-Przygod%C4%99-687347091288220/pho toa.1077421475614111.1073741838.687347091288220/1077421502280775 (17.04.2022)

[54] Statistic Road Accident https://statystyka.policja.pl/

[55] Gorzelańczyk P., Jurkovič M., Kalina T., Mohanty M. 2022. Forecasting the road accident rate and the impact of the Covid 19 on its frequency in the Polish provinces. Communications 24(4)A 216-A 231. https://doi.org/10.26552/com.C.2022.4.A216-A231

[56] Gorzelańczyk, P. (2022). Change in the Mobility of Polish Residents during the Covid-19 Pandemic. Communications - Scientific letters of the University of Zilina, 24(3), A100-111. doi: 10.26552/com.C.2022.3.A100-A111

[57] Jurkovic, M., Gorzelanczyk, P., Kalina, T., Jaros, J. & Mohanty, M. (2022). Impact of the COVID-19 pandemic on road traffic accident forecasting in Poland and Slovakia. Open Engineering, 12(1), 578-589. https://doi.org/10.1515/eng-2022-0370