OECONOMIA COPERNICANA

VOLUME 12 ISSUE 3 SEPTEMBER 2021

p-ISSN 2083-1277, e-ISSN 2353-1827

www.oeconomia.pl

ORIGINAL ARTICLE

Citation: Landmesser, J. (2021). The use of the dynamic time warping (DTW) method to describe the COVID-19 dynamics in Poland. *Oeconomia Copernicana, 12*(3), 539–556. doi: 10.24136/oc.2021.018

Contact: joanna_landmesser@sggw.edu.pl

Article history: Received: 30.01.2021; Accepted: 7.08.2021; Published online: 30.09.2021

Joanna Landmesser
Warsaw University of Life Sciences – SGGW, Poland
orcid.org/0000-0001-7286-8536

The use of the dynamic time warping (DTW) method to describe the COVID-19 dynamics in Poland

JEL Classification: *C32; C38; I15*

Keywords: *COVID-19; coronavirus; regional analysis; dynamic time warping; clustering*

Abstract

Research background: In recent times, the whole world has been severely affected by the COVID-19 pandemic. The influence of the epidemic on the society and the economy has caused a great deal of scientific interest. The development of the pandemic in many countries was analyzed using various models. However, the literature on the dissemination of COVID-19 lacks econometric analyzes of the development of this epidemic in Polish voivodeships.
Purpose of the article: The aim of the study is to find similarities in time series for infected with and those who died of COVID-19 in Polish voivodeships using the method of dynamic time warping.
Methods: The dynamic time warping method allows to calculate the distance between two time series of different lengths. This feature of the method is very important in our analysis because the coronavirus epidemic did not start in all voivodeships at the same time. The dynamic time warping also enables an adjustment of the timeline to find similar, but shifted, phases. Using this method, we jointly analyze the number of infected and deceased people in each province. In the next step, based on the measured similarity of the time series, the voivodeships are grouped hierarchically.
Findings & value added: We use the dynamic time warping to identify groups of voivodeships affected by the epidemic to a different extent. The classification performed may be useful as it indicates patterns of the COVID-19 disease evolution in Polish voivodeships. The results obtained

at the regional level will allow better prediction of future infections. Decision makers should formulate further recommendations for lockdowns at the local level, and in the long run, adjust the medical infrastructure in the regions accordingly. Policymakers in other countries can benefit from the findings by shaping their own regional policies accordingly.

Introduction

The first case of SARS-CoV-2 virus infection was detected in Wuhan, China in early December 2019. Within a short period of time, outbreaks of coronavirus (COVID-19) have been confirmed in 221 countries (as of January 30th, 2021). The virus is highly contagious, and thousands of new cases are being reported around the world each day. So far, the virus has affected over 102 million people, and 2.2 million of them have died from it. Due to the importance of the problem, it is receiving increasing interest from researchers in different countries (for example, scientific papers analyzing the spread of the pandemic and its economic impact are Pardal *et al.*, 2020; Zinecker *et al.*, 2021; Svabova *et al.*, 2021).

Poland noted its first coronavirus infection on March 4th, 2020 in Zielona Góra. Following the rapid spread of the disease across the whole country, the infection curve has stabilized at a relatively low level. The flattening of the curve was a consequence of the health and social restrictions imposed on the society. The number of daily cases increased considerably in October, and there was a sharp rise in the total number of reported infections in mid-November. November 25th was the deadliest day so far, with 674 deaths recorded. As of January 30th, 2021, the number of infected individuals amounted to 1,508,674, and the number of confirmed coronavirus-related fatalities was 37,082 (Worldometer, 2021).

The virus has attacked Polish provinces at different times. The earliest it appeared in Lubuskie province and the latest in Kujawsko-Pomorskie and Podlaskie (time delay of 13 days). The first death occurred in Wielkopolskie voivodeship and it took 215 days before the first person died in Lubuskie Province (the different lengths of the series for total cases and total deaths for Polish voivodeships are presented in Table 1). Therefore, the fact of uneven spread of the epidemic requires special study methods.

The highest number of infected people occurred in the Masovian province, followed by Śląskie and Wielkopolskie voivodeships. Currently, the rate of infection per one million inhabitants is the highest in Zachodniopomorskie, followed by Kujawsko-Pomorskie voivodeship. The highest number of people who died of COVID-19 per one million population recorded Opolskie voivodeship (Rogalski, 2021; Statista, 2021).

Currently, the biggest public concern in Poland relates to the economy. The epidemic has damaged the material situation of Poles, and they have suffered the negative impact of the virus on their emotional well-being. People fear deep depression, reduction on freedom, and social separation. Measures to counteract the COVID-19 epidemic have been introduced by the Minister of Health: limiting the size of public gatherings, the distance between people in public places, limiting the number of people using public transport, closing schools, universities, shopping malls, restaurants, sports centers, hotels, etc., compulsory covering of the mouth and nose in public places, inside and outside (cf., Korzeb & Niedziółka, 2020).

The purpose of this work is to analyze patterns and find similarities in the development of time series for infected with and those who died of COVID-19 in Polish voivodeships. We use the dynamic time warping (DTW) measure of similarity in order to calculate the distance between the respective time series. DTW makes it possible to compare time series of unequal length, which occurs when an epidemic breaks out at different times in different parts of the country. The DTW results are further used in the classification of Polish voivodeships affected by the pandemic to a different extent. The grouping obtained may be useful as it will indicate patterns of the COVID-19 disease evolution in Polish voivodeships.

The results obtained will provide a better understanding of infection dynamics and more accurate predictions of the future incidence at the regional level. Such knowledge will support policymakers in making further recommendations for lockdowns, which should be considered primarily at the local level (e.g., travel restrictions, school closures only in certain regions). Our analyses could support public health officials in preparing for the next wave of infections. The results could help regions adapt their medical infrastructure to the risk of future outbreaks appropriately. In the longer term, the possibility of future pandemics and the need to adapt to them in an adequate manner should be considered. The findings can be used by policymakers in other countries to shape their own regional policies in times of pandemics accordingly.

The paper is structured as follows. We provide an overview of the literature in the next section. The research methodology section presents the method of dynamic time warping. Then, we characterize the data and provide the results of the analysis for the 16 Polish voivodeships. In the subsequent sections, we discuss the findings, and we conclude.

Literature review

The COVID-19 epidemic is attracting attention not only in medicine and biology, but also in the field of mathematics and statistics. Differential equation models and statistical prediction models are applicable.

Researchers in epidemiology use the SIR model, which is the basis for further modified predictive models. The SIR model is a system of differential equations for populations of susceptible, infected, and removed individuals (Anderson & May, 1992). Anastassopoulou *et al.* (2020) used the SIR model to predict the development of the COVID-19 pandemic in China. The Wuhan outbreak was also analyzed employing a modified SEIR model for susceptible, exposed, infective, and recovered people (Lin *et al.*, 2020). Casella has applied a SIR-based model with control strategies and delays in reporting tests (Casella, 2021). Construction of other SIR models for COVID-19 was also undertaken by Roques *et al.* (2020) and Rojas *et al.* (2020).

Statistical time series modeling methods are also applied to investigate pandemic evolution. The ARIMA models prevail within the literature on predicting the dynamics of COVID-19 cases. This kind of models has been used to estimate the COVID-19 incidence by Ceylan (2020) or Benvenuto *et al.* (2020). Kumar *et al.* (2020) applied ARIMA models with a machine learning approach in epidemic prediction. Kufel (2020) evaluated the adequacy of ARIMA modeling for forecasting pandemic dynamics in European countries, including Poland.

The analysis of the COVID-19 outbreak in Poland within the first two months of the epidemic is presented in Raciborski *et al.* (2020). Orzechowska and Bednarek (2020) used the eSIR model (the extended SIR model) to forecast the pandemic assuming various lockdown scenarios. The MCMC algorithm made it possible to guage the effectiveness of variants of social isolation in reducing the spread of COVID-19. The SIIE model for susceptible, infected, infectious, and excluded constructed Kochanczyk *et al.* (2020).

Jarynowski *et al.* (2020a) investigated the connection between the number of coronavirus cases and socio-economic variables at the Polish poviats' level. They used data visualization techniques, correlation analysis, statistical regression, and clustering. In Jarynowski *et al.* (2020b) the influence of the pandemic on society by measuring online activity was examined. Wielechowski *et al.* (2020) analyzed the changes in public transport mobility in Poland during the coronavirus crisis. The COVID-19 impact on the Polish economy, specifically on the financial market was also examined, e.g., in Czech *et al.* (2020).

To examine the evolution patterns of COVID-19 in Polish voivodeships, the dynamic time warp method was implemented. This method was developed by Bellman and Kalaba (1959) and has many applications, as an example, in speech recognition (Rabiner *et al.*, 1978; Sakoe & Chiba, 1978), in gesture recognition (Arici *et al.*, 2014), in computer animation (Müller 2007), and even in finance (Stübinger, 2019). Stübinger and Schneider (2020) used the DTW clustering of cumulative COVID-19 cases to spot the lead-lag effect among the affected countries. Rojas *et al.* (2020) employed DTW combined with clustering to predict of pandemic data. Krywyk *et al.* (2020) characterized the dynamics of the virus spread in 14 countries and grouped them into 5 clusters of "differentiated epidemic patterns" with the support of DTW. A similar analysis for EU countries was conducted by Landmesser (2021). In line with our knowledge, no research on this subject has been carried out in Poland from the regional (voivodeships) perspective.

Research methodology

Dynamic time warping is often used to compare time series of different sizes. What is also important, this method enables an adjustment of the timeline to find similar but shifted phases.

The purpose of DTW is to measure the distance between two sequences X and Y, defined by $X = (x_1, x_2, \dots, x_S)$ and $Y = (y_1, y_2, \dots, y_T)$. The similarity of sequences of the same size can be obtained by computing the Euclidean or Minkowski distance. DTW can compare sequences of different sizes. It calculates an optimal alignment between two series by behaving non-linearly, i.e., stretching or compressing along the timeline. Such a 'warping' allows finding related regions between two time series and thus the similarity between them is specified.

The time warping path is a point-to-point alignment between X and Y and is formally defined as the sequence $wp = (wp_1, \dots, wp_L)$, where $wp_l = (s_l, t_l) \in \{1, \dots, S\} \times \{1, \dots, T\}$ for $l \in \{1, \dots, L\}$ ($L \in \{\max(S, T), \dots, S + T - 1\}$), satisfying the boundary, monotonicity, and step size conditions (Keogh & Ratanamahatana, 2005).

For each pair of elements in X and Y we have to calculate the local cost matrix using the formula $c(x_i, y_j) = |x_i - y_j|, i = 1, \dots, S, j = 1, \dots, T$. The total cost of a warping path *wp* is defined as:

$$c_{wp}(X, Y) = \sum_{l=1}^{L} c(x_{s_l}, y_{t_l}) = \sum_{l=1}^{L} \left| x_{s_l} - y_{t_l} \right|. \tag{1}$$

Then, the optimal warping path between X and Y is indicated by the match that has the lowest total cost:

$$\mathrm{DTW}(X,Y) = c_{wp*}(X,Y) = \min\{c_{wp}(X,Y) | wp \in WP\}. \tag{2}$$

The dynamic programming algorithm finds the optimal warping path by iterating over the local cost matrix and aggregating the costs (a recursion scheme can be found in Landmesser (2021)). The computed accumulated cost matrix D contains the value of the minimum distance between sequences *X* and *Y*:

$$\mathrm{DTW}(X,Y) = D(S,T). \tag{3}$$

The disadvantage of DTW is the heavy computational burden in finding the optimal alignment path. The DTW algorithm has the quadratic complexity O(n2). Several constraints are introduced to speed up the computation (such as upper and lower envelopes for the maximum warping allowed), which reduces the complexity to O(n). Time constraints can also be imposed on the length of the DTW warping window. All this improves both speed and precision, as it avoids pathological matching.

Following Rojas *et al.* (2020), we consider both the time series of total cases (*TC*) and the time series of total deaths (*TD*) for two voivodeships *A* and *B*:

$$\mathrm{DTW}_{AB} = 0.5 \cdot \mathrm{DTW}(TC_A, TC_B) + 0.5 \cdot \mathrm{DTW}(TD_A, TD_B). \tag{4}$$

In this way, the information about the number of infected has the same relevance as the information about the number of deaths. Typically, deaths are derived from the number of infections and usually follow the trend of infected with a lag of 14–21, but preliminary analysis of data for Polish provinces has shown that this is not the case (see Lubuskie province). Potential correlations between the series may be influenced by, for example, the different state of the medical infrastructure in the regions. Therefore, the decision was made to treat both series equally, realizing the subjectivity of this solution.

Based on the DTW similarity matrix for Polish voivodeships, provinces were grouped with the use of hierarchical clustering algorithm. The average-linkage method with the quadratic Euclidean distance was applied. The optimal cluster number was identified using Dunn, Silhouette, and Caliński-Harabasz indices. The outcomes were compared between the formed clusters of voivodeships.

Results

In our analysis, we rely on daily data on the COVID-19 pandemic in Polish voivodeships available on the website http://bit.ly/covid19-poland (Rogalski, 2021). The methodology outlined above was used for the analysis of coronavirus cases from March 3, 2020, until January 15, 2021. Below we focus on the following time series: *NC* — new cases per day (smoothed), *TC* — total cases, *TD* — total deaths. All variables are expressed in units per million inhabitants. Time series for new cases per day are shown additionally and are for illustrative purposes only.

First, the DTW distance matrix between the time series for voivodeships in pairs was calculated (cf., Eq. (3) and Eq. (4)). The alignments were calculated employing the dtw package for R (Giorgino, 2009). Fig. 1 presents the results of time series alignments for the Mazowieckie and Zachodniopomorskie voivodships.

Fig. 2 shows the optimal warping paths corresponding to Fig. 1 for Mazowieckie and Zachodniopomorskie voivodeships. In these three-way-plots, the query time series are shown in small lower panels, and the reference time series are shown on the left; the warping curves are shown in the larger inner panels. The shapes of these curves give us information about the corresponding pairs of points in time. For example, from the right panel for the sequence *TD* it can be read that the time series for total deaths in the Mazowieckie voivodship leads the time series for total deaths in the Zachodniopomorskie voivodship (this is because the differences in the indices of wp^* are mostly negative).

Fig. 3 illustrates the local costs and the identified optimal warping paths wp^* given the adequate series. These paths present the point-to-point matching between the X and Y series indices. The figure shows the indices for the two selected provinces (Mazowieckie and the Zachodniopomorskie). These are the so-called query and reference indices. In the panels, the optimal paths wp^* run in the “valleys” of low cost (dark regions) and avoid the “mountains” of high cost (bright areas).

The DTW distances calculated for each pair of provinces (according to Eq. (4)) have an application in hierarchical clustering (Sardá-Espinosa, 2019). In this contribution, after measuring the similarities between the 16 voivodeships according to the COVID-19 spread, the agglomerative hierarchical clustering was carried out, mainly due to its high visualization power. The average linkage was used with the squared Euclidean distance. The grouping resulted in a nested hierarchy of similar time series. The previously determined values of the DTW similarity measure were essential in this process. Fig. 4 shows the cluster dendrogram. For determining the number of clusters, we used the Dunn, Silhouette, and Caliński-Harabasz indices. The higher the values of these indices, the better the data division. The optimum was found by varying the number of clusters from 3 to 8. For example, by selecting the solution maximizing the Caliński-Harabasz index the subsequent values of this index were: 60.8; 57.7; 69.8; 67.1; 59. The chosen solution, with a Caliński-Harabasz score of 69.8, segmented the voivodeships into five groups:

- group 1: Podkarpackie, Świętokrzyskie,
- group 2: Dolnośląskie, Mazowieckie, Podlaskie, Małopolskie, Lubelskie, Lubuskie,
- group 3: Łódzkie, Śląskie,
- group 4: Warmińsko-Mazurskie, Zachodniopomorskie, Opolskie, Pomorskie, Wielkopolskie,
- group 5: Kujawsko-Pomorskie.

Discussion

The analysis of the epidemic evolution in Poland has identified a general pattern. The development of COVID-19 shows three distinct phases: a phase of rapid disease progression, until it reaches its peak, where it begins to decline, until a point of stabilization. Although this model is quite common, there are differences in the dynamic behavior of the pandemic between provinces. The obtained results were analyzed and compared in identified groups of voivodeships, using the recorded time series for new cases per day (*NC*), total cases (*TC*), and total deaths (*TD*), as shown in Figures 5–9.

Group 1 consists of two voivodeships: Podkarpackie and Świętokrzyskie. Both provinces demonstrate a very strong similarity in their epidemic dynamics. These are the provinces that slightly suffered from the first wave of the epidemic and now have the pandemic almost under control. They have a relatively low total number of cases per million inhabitants until January, 15 (below 30,000 cases), but the values of total deaths per million are among the highest in the country (from about 800 to 950) (see Fig. 5). The curve for the number of new cases per million has a very symmetrical rise and fall phases and a short time to flat out. Although some researchers argue that population density or age structure is losing predictive power concerning the infection dynamics, it seems that COVID-19 has least hit those regions of Poland which are poor and with less communication (e.g., Podkarpackie and Świętokrzyskie voivodeships).

Group 2, the largest group, brings together 6 voivodeships: Dolnośląskie, Mazowieckie, Podlaskie, Małopolskie, Lubelskie, and Lubuskie, with a higher peak in the number of daily new cases and a decreasing phase not as intense as in Podkarpackie and Świętokrzyskie. However, the provinces in group 2 have kept their daily death count low. As a result, group 2 has a slightly higher number of total cases and a lower number of deaths (about 33,000 and 750 per million, respectively) compared to group 1 (Fig. 6). COVID-19 has hardly hit not only less-connected and poor regions of Poland (Lubelskie and Lubuskie), but also those that are the most connected and richest regions (e.g., Dolnośląskie, Mazowieckie, Małopolskie provinces). It is believed that regions of high social capital are more susceptible to the high numbers of COVID-19 notifications. This is not the case in Poland. It turns out that the infection dynamics is less driven by variables such as population size, population density, and income (compare Aldridge *et al.*, 2020). It is also possible that large and rich cities (Warszawa, Kraków, or Wrocław) with high population density are more able to combat COVID-19 than poor regions.

Group 3, made up of Łódzkie and Śląskie voivodeships, includes provinces that suffered a moderate to a medium outbreak of the epidemic. The amount of total cases is higher than in the preceding groups (approximately 37,000 per million) and the number of deaths is 900 per million (cf. Fig. 7). Although Łódzkie and Śląskie are close together, Śląskie recorded an increased number of COVID-19 cases in May 2020. After this moderate first wave, a steep second wave began in early October with a peak in mid-November. Śląskie voivodeship, which is a highly industrialized region, has become a large hotspot during the epidemic (it can be linked with workplace-associated infections). Many researchers confirmed that employment in industry correlates with the number of cases (Jarynowski *et al.*, 2020a).

Group 4 includes 5 provinces: Warmińsko-Mazurskie, Zachodniopomorskie, Opolskie, Pomorskie, and Wielkopolskie. It is marked by a downward tendency of unequal speed. The cluster records higher total incidence rates (about 40,000 per million), while total deaths range between 800 and 1,200 per million (see Fig. 8). Certainly, a high number of deaths per million inhabitants was observed in Opolskie province. The epidemic scheme comprises of a heavy growth period, up to a "peak", followed by a gradual decline in the number of deaths. The provinces have an outflow character, and the big amount of emigrants has speeded the emergence of COVID-19 and the outbreak of the epidemic. Jarynowski *et al.* (2020a) noted that cases from abroad represent the most common group of confirmed cases in the first weeks after the COVID-19 was introduced into Poland.

Group 5, exclusively made up of Kujawsko-Pomorskie, presents the province where the pandemic is still "out-of-control". The number of COVID-19 cases is the largest in Poland (47,000 per million) (Fig. 9). The cumulative number of deaths exceeds 900 per million. This voivodeship is characterized by a very long epidemic wave (a more spread out plateau at the top) and thus represents a less advanced stage of the pandemic. Immigration and mobility are argued to be the best predictors of the occurrence of infections, while emigration and industrialization explain most of the scale of the epidemic (Jarynowski *et al.*, 2020a). Probably the outflow nature of the Kujawsko-Pomorskie voivodeship with a large number of emigrants (outflow of workers abroad and related returns to Poland) accelerated the outbreak of the epidemic here.

It should be emphasized that the only previous research to date for Poland at the regional level, which can be compared with the results presented here, was that of Jarynowski *et al.* (2020a). Since such studies for regions already exist for other countries (see Chu, 2021, for Italy and Spain), it was necessary to fill the gap for Poland. This circumstance demonstrates the novelty and value added to the field in the case of the current work.

Conclusions

The purpose of this work was to group the Polish voivodeships and, within the created clusters, to compare the COVID-19 pandemic evolution. Techniques dedicated to time series analysis were applied. The dynamic time warping method was used to measure the similarity of time series corresponding to 16 Polish provinces, taking into account the amount of infected and died of COVID-19. The number of groups found was five. Within the clusters, we compared the evolution of the epidemic.

Pandemic dynamics has the general pattern in Polish provinces. There are three characteristic phases: a phase of strong epidemic growth until a peak is reached, where it begins to decline, until a low point of stabilization. This general pattern was observed more or less for all the voivodeships under consideration.

The time series for the daily infected and dead in Polish voivodeships differ from each other in their magnitude. The reasons for this situation are diverse. We tried to understand the forces driving the dynamics of COVID-19. Mobility, migration, and industrialization may be related to infection dynamics of infections in particular regions. The size of the pandemic in

Poland seems much less associated with demographic factors such as population density, age structure, or income.

The clustering of Polish provinces by COVID-19 cases may facilitate a better understanding of past infection dynamics and more accurate forecasting of future cases at the local level. Our analyses could support policymakers and public health officials in preparation for the next wave of infections. The knowledge gained can also be helpful in preparing for trips traveling to particular regions.

International readers can also benefit from the study conducted. It would encourage them to make similar analyses for their own countries, which could result in practical applications. However, this would require further interest from the decision makers.

To summarize, the methodology used is able to uncover differences between Polish voivodeships in terms of epidemic development, as well as identify some regularities in the underlying data. The previously mentioned technical limitations of the DTW method are not a major obstacle to the research. The concern may be related to the limited applicability of the obtained results in practice. It is important to remember that these results are place- and time-specific, and the subsequent course of the pandemic requires continuous analyses.

In future research, the observed differences should be confronted with the question: to what extent do epidemic management strategies contribute to this pattern?

References

Aldridge, R. W., Lewer, D., Katikireddi, S. V., Mathur, R., Pathak, N., Burns, R., Fragaszy, E. B., Johnson, A. M., Devakumar, D., Abubakar, I., & Hayward, A. (2020). Black, Asian and minority ethnic groups in England are at increased risk of death from Covid-19: indirect standardisation of NHS mortality data. *Wellcome Open Research*, *5*(88). doi: 10.12688/wellcomeopenres.15922.2.

Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLOS ONE*, 15:e0230405. doi: 10.1371/journal.pone.0230405.

Anderson, R. M., & May, R. M. (1992). *Infectious diseases of humans: dynamics and control*. Oxford: Oxford University Press. doi:10.1002/hep.1840150131.

Arici, T., Celebi, S., Aydin, A. S., & Temiz, T. T. (2013). Robust gesture recognition using feature pre-processing and weighted dynamic time warping. *Multimedia Tools and Applications*, *72*, 3045–3062. doi: 10.1007/s11042-013-1591-9.

Bellman, R., & Kalaba, R. (1959). On adaptive control processes. *IRE Transactions on Automatic Control*, *4*(2), 1–9. doi: 10.1109/TAC.1959.1104847.

Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., & Ciccozzi, M. (2020). Application of the ARIMA model on the COVID-2019 epidemic dataset. *Data in Brief*, *29*, 105340. doi: 10.1016/j.dib.2020.105340.

Casella, F. (2021). Can the COVID-19 epidemic be controlled on the basis of daily test reports? *IEEE Control Systems Letters*, *5*(3), 1079–1084. doi: 10.1109/LCSYS.2020.3009912.

Ceylan, Z. (2020). Estimation of COVID-19 prevalence in Italy, Spain, and France. *Science of The Total Environment*, *729*, 138817. doi: 10.1016/j.scitotenv.2020.138817.

Chu, J. (2021). A statistical analysis of the novel coronavirus (COVID-19) in Italy and Spain. *PLoS ONE*, *16*(3): e0249037. doi: 10.1371/journal.pone.0249037.

Czech, K., Wielechowski, M., Kotyza, P., Benešová, I., & Laputková, A. (2020). Shaking stability: COVID-19 impact on the Visegrad Group countries' financial markets. *Sustainability*,*12*, 6282. doi: 10.3390/su12156282.

Giorgino, T. (2009). Computing and visualizing dynamic time warping alignments in R: the dtw package. *Journal of Statistical Software*, *31*(7), 1–24. doi: 10.18637/jss.v031.i07.

Jarynowski, A., Wójta-Kempa, M., & Krzowski, Ł. (2020a). An attempt to optimize human resources allocation based on spatial diversity of COVID-19 cases in Poland. *medRxiv preprint*. doi: 10.1101/2020.10.14.20090985.

Jarynowski, A., Wójta-Kempa, M., Płatek, D., & Czopek, K. (2020b). Attempt to understand public-health relevant social dimensions of COVID-19 outbreak in Poland. *Society Register*, *4*(3), 7–44. doi: 10.14746/sr.2020.4.3.01.

Keogh, E., & Ratanamahatana, C. A. (2005). Exact indexing of dynamic time warping. *Knowledge and Information Systems*, *7*, 358–386. doi: 10.1007/s10115-004-0154-9.

Kochanczyk, M., Grabowski, F., & Lipniacki, T. (2020). Dynamics of COVID-19 pandemic at constant and time-dependent contact rates. *Mathematical Modelling of Natural Phenomena*, *15*(28). doi: 10.1051/mmnp/2020011.

Korzeb, Z., & Niedziółka, P. (2020). Resistance of commercial banks to the crisis caused by the COVID-19 pandemic: the case of Poland. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *15*(2), 205–234. doi: 10.24136/eq.2020.010.

Krywyk, J., Oettgen, W., Messier, M., Mulot, M., Ugon, A., & Toubiana, L. (2020). Dynamics of the COVID-19 pandemics: global pattern and between countries variations. *medRxiv preprint*. doi: 10.1101/2020.07.20.20155390.

Kufel, T. (2020). ARIMA-based forecasting of the dynamics of confirmed COVID-19 cases for selected European countries. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *15*(2), 181–204. doi: 10.24136/eq.2020.009.

Kumar, P., Kalita, H., Patairiya, S., Sharma, Y. D., Nanda, C., Rani, M., Rahmani, J., & Bhagavathula, A. S. (2020). Forecasting the dynamics of COVID-19 pandemic in top 15 countries in April 2020: ARIMA model with machine learning approach. *medRxiv preprint*. doi: 10.1101/2020. 03.30.20046227.

Landmesser, J. (2021). Analysis of COVID-19 dynamics in EU countries using the dynamic time warping method and ARIMA Models. In K. Jajuga, K. Najman, & M. Walesiak (Eds.). *Data analysis and classification. Methods and applications*. Springer International Publishing. 337–352. doi:10.1007/978-3-030-75190-6_19.

Lin, Q., Zhao, S., Gao, D., Lou, Y., Yang, S., Musa, S. S., Wang, M. H., Cai, Y., Wang, W., Yang, L., & He, D. (2020). A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action. *International Journal of Infectious Diseases*, *93*, 211–216. doi: 10.1016/j.ijid.2020.02.058.

Müller, M. (2007). *Information retrieval for music and motion*. Springer-Verlag Berlin Heidelberg. doi: 10.1007/978-3-540-74048-3.

Orzechowska, M., & Bednarek, A. K. (2020). Forecasting COVID-19 pandemic in Poland according to government regulations and people behavior. *medRxiv preprint*. doi: 10.1101/2020.05.26.20112458.

Pardal, P., Dias, R., Šuleř, P., Teixeira, N., & Krulický, T. (2020). Integration in Central European capital markets in the context of the global COVID-19 pandemic. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *15*(4), 627–650. doi: 10.24136/eq.2020.027.

Rabiner, L., Rosenberg, A., & Levinson, S. (1978). Considerations in dynamic time warping algorithms for discrete word recognition. *IEEE Transactions Acoustic Speech Signal Process*, *26*(6), 575–582. doi: 10.1109/tassp.1978.116 3164.

Raciborski, F., Pinkas, J., Jankowski, M., Sierpiński, R. A., Zgliczyński, W. S., Szumowski, Ł., Rakocy, K., Wierzba, W., & Gujski, M. (2020). Dynamics of the coronavirus disease 2019 outbreak in Poland: an epidemiological analysis of the first 2 months of the epidemic. *Polish Archives of Internal Medicine*, *130*(7-8), 615–621. doi: 10.20452/pamw.15430.

Rogalski, M. (2021). COVID-19 w Polsce. Retrieved from http://bit.ly/covid19-poland (15.01.2021).

Rojas, F., Valenzuela, O., & Rojas, I. (2020). Estimation of COVID-19 dynamics in the different states of the United States using time-series clustering. *medRxiv preprint*. doi: 10.1101/2020.06.29.20142364.

Roques, L., Klein, E. K., Papaix, J., Sar, A., & Soubeyrand, S. (2020). Using early data to estimate the actual infection fatality ratio from COVID-19 in France. *Biology*, *9*(5), 97. doi: 10.3390/biology9050097.

Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *26*(1), 43–49. doi: 10.1109/tassp.1978.1163055.

Sardá-Espinosa, A. (2019). Time-series clustering in R using the dtwclust package. *R Journal*, *11*(1), 22–43. doi: 10.32614/RJ-2019-023.

Statista (2021). Coronavirus (COVID-19) in Poland. Retrieved from https://www.statista.com/study/71486/coronavirus-covid-19-in-poland/ (15.01.2021).

Stübinger, J. (2019). Statistical arbitrage with optimal causal paths on high-frequency data of the S&P 500. *Quantitative Finance*, *19*, 921–935. doi: 10.1080/14697688.2018.1537503.

Stübinger, J., & Schneider, L. (2020). Epidemiology of coronavirus COVID-19: forecasting the future incidence in different countries. *Healthcare*, *8*(2), 99. doi: 10.3390/healthcare8020099.

Svabova, L., Tesarova, E. N., Durica, M., & Strakova, L. (2021). Evaluation of the impacts of the COVID-19 pandemic on the development of the unemployment rate in Slovakia: counterfactual before-after comparison. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, *16*(2), 261–284. doi: 10.24136/eq.2021.010.

Wielechowski, M., Czech, K., & Grzęda, Ł. (2020). Decline in mobility: public transport in Poland in the time of the COVID-19 pandemic. *Economies*, *8*(4), 78. doi: 10.3390/economies8040078.

Worldometer (2021). COVID-19 coronavirus pandemic. Retrieved from https://www.worldometers.info/coronavirus/ (30.01.2021).

Zinecker, M., Doubravský, K., Balcerzak, A. P., Pietrzak, M. B., & Dohnal, M. (2021). The Covid-19 disease and policy response to mitigate the economic impact in the EU: an exploratory study based on qualitative trend analysis. *Technological and Economic Development of Economy, 27*(3), 742–762. doi: 10.3846/tede.2021.14585.

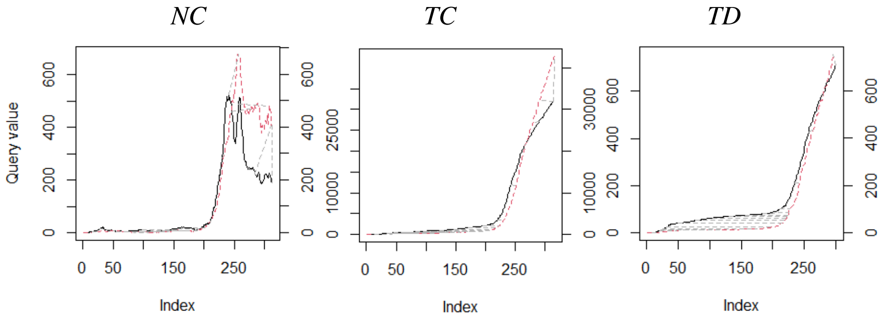
Annex

Table 1. The length of the time series for total cases (*TC*) and total deaths (*TD*) in Polish voivodeships (in days)

Voivodeship	*TC*	*TD*	Voivodeship	*TC*	*TD*
Dolnośląskie	316	309	Podkarpackie	311	306
Kujawsko-Pomorskie	305	288	Podlaskie	305	286
Lubelskie	312	308	Pomorskie	308	283
Lubuskie	318	185	Śląskie	314	299
Łódzkie	311	290	Świętokrzyskie	308	297
Małopolskie	313	293	Warmińsko-Mazurskie	316	287
Mazowieckie	314	300	Wielkopolskie	313	310
Opolskie	311	289	Zachodniopomorskie	316	296

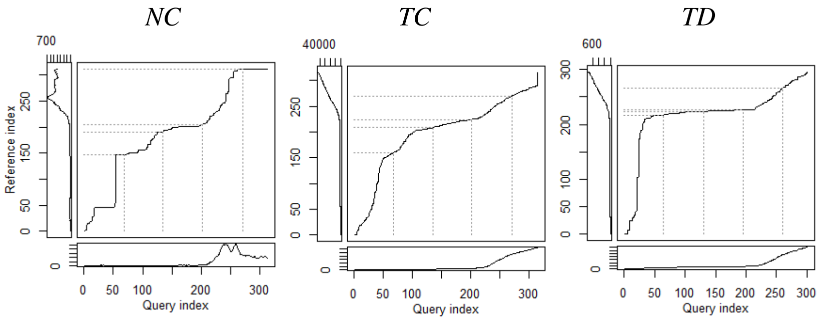
Source: own elaboration based on Rogalski (2021).

Figure 1. The alignments between sequences of *NC*, *TC*, and *TD* in Mazowieckie (solid lines) and Zachodniopomorskie (dashed lines) voivodeships



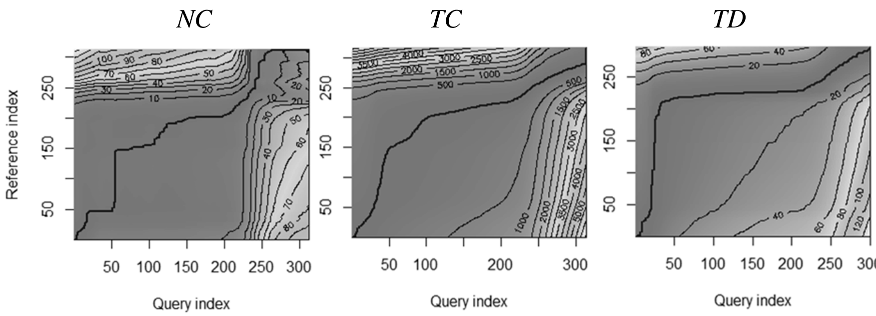
Source: own elaboration using the dtw package for R.

Figure 2. Three-way plots showing the optimal warping paths (inner panels) for sequences of *NC*, *TC*, and *TD* in Mazowieckie (query indices) and Zachodniopomorskie (reference indices) voivodeships



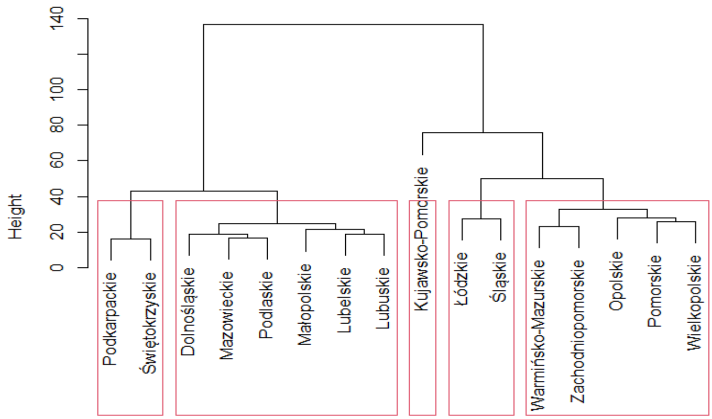
Source: own elaboration using the dtw package for R.

Figure 3. Local costs and the optimal warping paths (solid lines) for the alignments between sequences of *NC*, *TC*, and *TD* for Mazowieckie (query indices) and Zachodniopomorskie (reference indices) voivodeships.



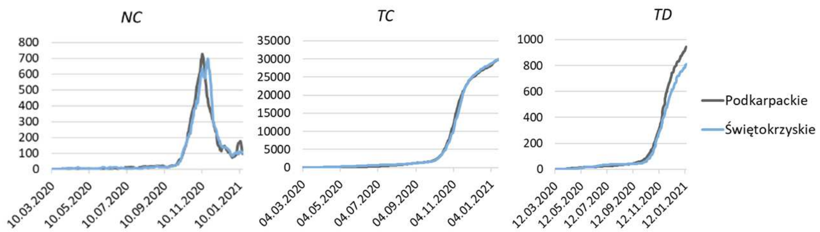
Source: own elaboration using the dtw package for R.

Figure 4. Cluster dendrogram showing 5 groups of voivodeships



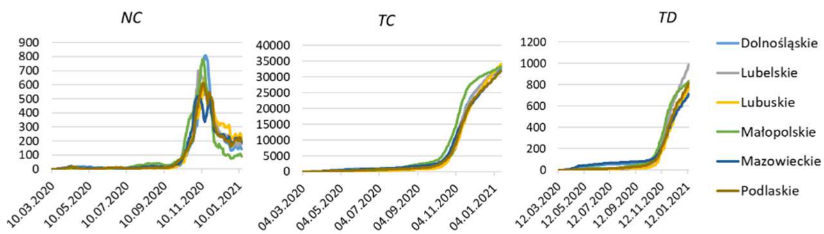
Source: own elaboration using the hclust function in R.

Figure 5. Group 1: daily cases (*NC*), total cases (*TC*), total deaths (*TD*) per million



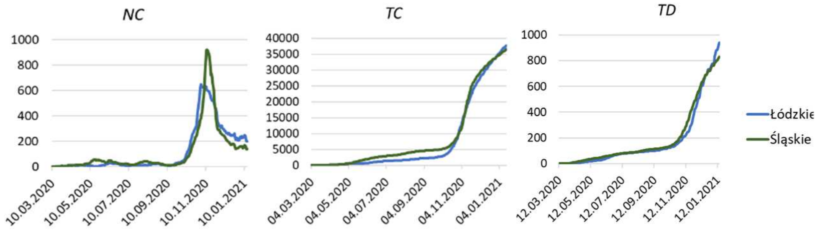
Source: own elaboration based on Rogalski (2021).

Figure 6. Group 2: daily cases (*NC*), total cases (*TC*), total deaths (*TD*) per million



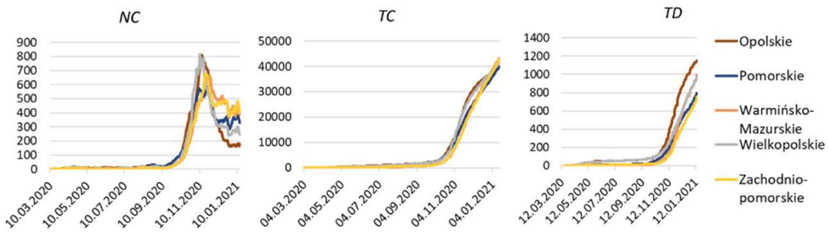
Source: own elaboration based on Rogalski (2021).

Figure 7. Group 3: daily cases (*NC*), total cases (*TD*), total deaths (*TD*) per million



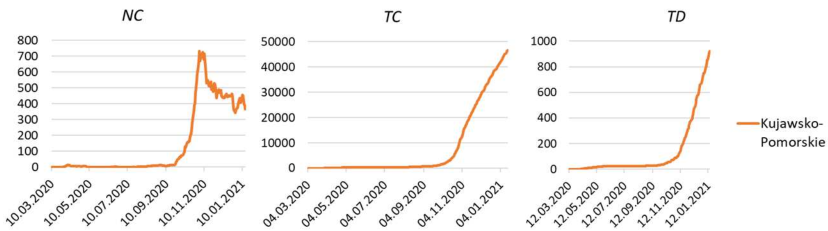
Source: own elaboration based on Rogalski (2021).

Figure 8. Group 4: daily cases (*NC*), total cases (*TC*), total deaths (*TD*) per million



Source: own elaboration based on Rogalski (2021).

Figure 9. Group 5: daily cases (*NC*), total cases (*TC*), total deaths (*TD*) per million



Source: own elaboration based on Rogalski (2021).