

COVID-19 data sources: evaluation of map applications and analysis of behavior changes in Europe's population

VÍT PÁSZTO^{1,2}, JAROSLAV BURIAN^{1,2}, KAREL MACKŮ¹

¹ Palacký University in Olomouc, Faculty of Science, Department of Geoinformatics, Olomouc, Czechia; e-mail: vit.paszto@gmail.com, jaroslav.burian@upol.cz, karel.macku@upol.cz

² Moravian Business College Olomouc, Department of Informatics and Applied Mathematics, Olomouc, Czechia

ABSTRACT Due to the current situation, and preventive measures taken to tackle COVID-19, it is crucial to keep society well-informed. Besides media and official news, that often include tabular data, it has also become a new standard for information sources to incorporate a map application or geovisualization. This paper offers a comprehensive and systematic overview describing the most prominent and useful map applications and map visualizations. News outlets should place the same importance on data analysis and interpretation as they place on data visualization. This paper emphasizes the role of geospatial data and analysis during the COVID-19 pandemic and aims to provide insights into the topic in order to better understand the consequences caused by the disease. Specifically, the paper deals with the COVID-19 Community Mobility Reports dataset, offering unique information about changes in human activity due to the pandemic. We show how this dataset can be utilized in terms of geovisual analytics and clustering in order to reveal the spatial pattern of such changes in human behavior.

KEY WORDS COVID-19 – Community Mobility Reports – Google Location – Spatial Analyses – Clustering – GIS

PÁSZTO, V., BURIAN, J., MACKŮ, K. (2020): COVID-19 data sources: evaluation of map applications and analysis of behaviour changes in Europe's population. *Geografie*, 125, 2, 171–209.

<https://doi.org/10.37040/geografie2020125020171>

Received April 2020, accepted May 2020.

1. Introduction

In December 2019, a new coronavirus began to spread from the Chinese province of Hubei, causing Severe Acute Respiratory Syndrome (SARS) called SARS-CoV-2. The disease was later renamed to COVID-19. Initially, the disease spread in inland China with its epicentre in Wuhan City, Hubei Province (e.g., Huang et al. 2020 or Li et al. 2020b) and gradually spread to other parts of the world. After China, Europe was hardest hit by the epidemic, specifically the north of Italy became the epicentre of the spread of the disease, and Europe was declared the global epicentre in mid-March 2020 (Fisher, Wilder-Smith 2020). Although Europe had around one million confirmed cases of COVID-19 in mid-April 2020, by then the United States (around 600,000 confirmed cases) had already been identified as the current global epicentre (WHO 2020). The total cumulative number of people diagnosed with COVID-19 reached 2,406,575 worldwide on 19 April 2020, and 165,031 people had died from the disease (Worldometers.info 2020).

Given that the history of the spread of the disease is known to the general public and the aim of this article is to offer an overview of visual data sources and especially to analyze the impact of the disease on the behaviour of Europeans, the geographical spread itself and basic statistics of the COVID-19 disease will not be described in detail. However, the geographical spread of viral infectious diseases is very specific and is often associated with individual or group behaviour of vectors through geographical hot-spots, or hubs (Kissler et al. 2019), which are characterized by a high degree of connectivity (e.g. airports) or population concentrations (Ferguson et al. 2006). The issue of viral diseases is generally studied by epidemiologists using spatial tools (e.g. Dellicour, Rose, Pybus 2016; Kissler et al. 2019; Charu et al. 2017; Carroll et al. 2015; Lemey et al. 2010). A large number of articles and studies on the current COVID-19 pandemic have already been published, also thanks to the response of the scientific community and major publishers of scientific papers, who during the COVID-19 pandemic are facilitating an accelerated review process and are then collecting the resulting papers and making them freely available to the public. Examples include an initiative by Springer Nature publishers, or Elsevier's Novel Coronavirus Resource Directory. However, the vast majority of publications deal with purely epidemiological aspects of COVID-19. So far, few implicitly include the geographical aspect as a major research topic. One of the most recent examples is an article on the impact of human migration on the spread of COVID-19 (Sirkeci, Yucesahin 2020). Its authors point out the importance of data on traffic intensity and the associated migration of people with a focus on travel corridors, mentioning that the study of immigration data (arrivals) with respect to their point of origin significantly contributes to controlling the spread of the disease. Huang, Liu, and Ding (2020) studied the spread of the new coronavirus in China using spatial statistics calculated from existing data. In

their analysis, which also includes the geographical aspects of the spread of the disease, they created a model predicting the moment when the epidemic could be brought under control (which they subsequently confirmed using real data). Spatio-temporal modelling using the scan-statistics method was performed in the United States, which is currently the most dramatically affected by the COVID-19 disease, by Desjardins, Hohl, and Delmelle (2020). Using this approach, the authors detected active clusters (the states of New York, New Jersey, Pennsylvania, and Delaware on the East Coast and California, Oregon, and Washington on the West Coast) and potential clusters (the states of Colorado, Utah, and major cities in the Mideast) of contagion in the United States. It should be noted that the study was based on data up to 9 March 2020, and so the potential clusters identified by the authors may now seem obvious. Nevertheless, their research estimated the regions and cities that would become foci of the further spread of the disease fairly accurately.

Kamel Boulos and Geraghty (2020) also mention geographical tracing and mapping of the COVID-19 disease pandemic as a contribution by geographical information systems (GIS) toward addressing the current situation. The authors discuss that GIS tools are actually nothing new in the analysis and mapping of infectious diseases. Applications such as (near) real-time disease mapping, predictive disease risk modelling using traffic or travel data, or tracking the trajectories of super-carriers and their contacts have already been helping experts understand viral infectious diseases for a number of years and decades. One of the first visualizations of the relationship between place and disease (in this case the plague) dates back to 1694 in Italy (Koch 2005). In his book, Walter (2000) mentions other cases of early spatial epidemiology from the eighteenth century involving yellow fever and cholera. But it was John Snow (1813–1858) who finally laid the scientific foundations for “modern” epidemiology and emphasized the relationship between the place of infection and the way it spread (for more see Paneth 2004, or specifically Snow 1856). And it is through these cases that Boulos and Geraghty (2020) illustrate the importance of the role of GIS as a tool that can not only analyze but also effectively visualize data. According to Kamel Boulos and Geraghty (2020), the data shown on the map was the initial sources of information on the COVID-19 pandemic. Based on this, Chapter 2 describes the deployment of various forms of data visualization in greater detail.

The added value of GIS tools and “big data” is described in a review article by Zhou et al. (2020). The term first appeared in connection with visualization more than 20 years ago (Cox, Ellsworth 1997). Batty (2016) states that defining the limits of what is and is not big data is very relative, because over the course of the past 60 years there have virtually always been data sets that at the time were impossible to process. Zhou et al. (2020) demonstrate the role of spatial analyses thanks to data on the complete course of the disease in China, hence from its onset through

culmination to the stage of keeping the spread of the disease under control. The results of these authors' empirical research highlight the role of GIS primarily in providing timely and rapid information to the general public and the professional community (in the form of clear aggregation of a large number of data sources), visualization, prediction of regional disease spread, spatial segmentation of epidemiological risk and prevention, management of medical material supply and demand, or for monitoring social and emotional characteristics in order to mitigate panic in the population. The authors also note the importance of data sources that do not come from official statistics, but rather from various special-interest groups (often ad hoc), academia, or social networks. All this with the aim of using GIS to support decision-making and the creation of "rapid knowledge" (Zhou et al. 2020). It is important to note that multinational private corporations such as Microsoft, Facebook, Google, or even mobile operators, collect large amounts of data (which can be called big data due to its heterogeneity) on their users, which in the case of the COVID-19 pandemic appears to provide an opportunity to stop the uncontrolled spread of the disease. Batty (2013, 2016) considers the existence of mobile phones and the large amount of potential data associated with them to be a key aspect of big data. An example of the use of this potential is the introduction of a "smart quarantine system" and the creation of memory maps, where with the consent of the given individual their individual data from the relevant mobile operator is processed into a digital map that makes it easier to remember all the places they visited and hence makes finding all the people they have come into contact easier and more efficient. At the time of writing (April 2020), the creation of memory maps in Czechia is being supplemented by other technologies such as the eRouška application or tracing in the Mapy.cz application (MHCR 2020). Even competitors Apple and Google have adopted a common approach to detecting contacts between people using different platforms, making it easier to trace individuals infected with COVID-19 (Apple Inc. 2020). The smart quarantine and smart tracing approach has proven successful in countries such as Singapore or South Korea (COVID-19 NERC 2020), and currently also in the United States (e.g. Yasaka, Lehigh, Sahyouni 2020). Google is a good source of data in this regard. For example, it has a service called Google Trends for detecting trends in keyword searches on the internet, which can be used for scientific purposes. A study monitoring the incidence (number of new patients) of dengue virus infection in Brazil (Romero-Alvarez et al. 2020) used Google Health Trends (GHT) and official dengue fever statistics to identify places where GHT can be used to monitor the fever's dynamics. The potential of Google Trends data is also mentioned in a short study by Santangelo et al. (2019). Li et al. (2020a), who discuss the possibilities of predicting the spread of viral infection according to the keywords "coronavirus" and "pneumonia" entered into Internet search engines (including Google) dealt with the COVID-19 disease in retrospective. They found that internet keyword

searches closely correlated with official statistics on the spread of COVID-19. In addition, on 3 April 2020, Google released a new service, COVID-19 Community Mobility Reports (Google LLC 2020), in response to requests from official institutions to provide at least anonymized data and the behaviour of Location History users (this involves conscious provision of information on the position of a mobile device based on GPS and device registration in the mobile network). For the sake of completeness, it should be added that Apple has published similar data from its users (Mobility Trends Reports), which, however, focus more on navigation data and vehicle use. According to Zhou et al. (2020), mass-processed individual user data can contribute to understanding the impacts of the epidemiological situation in general, hence also in the case of COVID-19. And it is precisely the analysis of data from Google that is dealt with in Chapter 3 of this article.

The authors' key motivation for writing this article was to contribute to the current discussion in (not only) the Czech geography community, based on an article by Džúrová and Jarolímek (2020) published in this journal. The study by the above authors focuses mainly on evaluating the current state of the pandemic based on official statistics and discussing what national approaches exist in the fight against the pandemic, how much of the population has been tested, and also broaches assumptions of further development and consequences of the COVID-19 pandemic. Our article brings insight into information sources and possible spatial analysis of Google Location data that provides good information on changes in population behaviour. This approach thus complements the findings of Džúrová and Jarolímek (2020) with another dimension of analytical processing of spatial data. The article starts with an overall overview (sources of map visualizations and data) and ends with a trans regional analysis of Google Mobility Community Reports for Europe data. It is therefore not an updated assessment of the current situation in terms of statistics tracking the number of infected, deceased, and cured, but rather a demonstration of how geoinformatics and spatial analysis tools can be used effectively to understand or highlight other aspects of COVID-19 in a broader geographical context. Emphasis is therefore not placed on a detailed analysis of changes in behaviour, which would require a longer timeline, but rather on acquainting readers with the possibilities presented by analyses, which can be further processed in subsequent studies by other authors.

First, the article systematically categorizes and critically evaluates map sources of information and data. This review is based on the authors' experience and will (perhaps) help international geographic community orient themselves in the large number of reports concerning COVID-19. The second part of the article focuses on practical spatial analysis and visualizations of real data on changes in human behaviour and the related six groups of socio-economic activities at the level of European regions. We are basing ourselves on the hypothesis that the COVID-19 pandemic and subsequent restrictions to slow the spread of the disease led to

changes in human behaviour in individual regions. The aim of the second part is to provide empirical evidence of behavioural changes, and to identify and describe the types of European regions sharing similar characteristics in response to the COVID-19 pandemic. As far as we know, studies of Google data for the period since the start of the pandemic have not been published anywhere in the field of geosciences. We are therefore confident that the article will inspire the professional community to perform their own research and to gain a deeper understanding of the geographical and other aspects of COVID-19.

2. Overview of map data sources

Given the current situation, there is an effort to provide the general public with quick, clear, and up-to-date information regarding the development of the pandemic. A number of both spatial and non-spatial visualizations are available on the internet that follow the current trend of presenting data through interactive web applications. These presentations are of various origins, use various data sources, and present information using various (not always correctly chosen) methods, which can often be confusing for the average reader. The most relevant of these sources were therefore arranged in a systematic overview in order to provide information on their technical processing, scope, authorship, and key content.

The systematic overview, whose 30 records complete with a detailed description are provided in Table 1, included information on the technology used, geographical coverage and level of detail, update frequency, data source, presentation type, author, data download availability, and licence type. The records were split into two groups – worldwide and focused only on Czechia.

Most of the examples present a global overview focused on the level of individual countries. This information can be found in all sources presented in the first part of the systematic overview, but also in most regional sources, where it serves as additional information (e.g. in [8, 9, 10, 26]). An important aspect of the web presentations under discussion is the source data used. Most global sources draw on data from the World Health Organization (WHO), European data is provided by the European Centre for Disease Prevention and Control (ECDC), and a source of more detailed data for the United States is the Centers for Disease Control and Prevention (CDC). A number of applications cite Johns Hopkins University in Maryland as the data source. It is often cited because it collects data from the aforementioned main data sources, supplements it with other available information, and publishes the aggregate data daily in the Github development environment, where it can be freely downloaded and subsequently used in any application. The Worldometer organization, an association of volunteers, scientists, and developers, works in a similar way. Applications focused on Czechia primarily

use data presented by the Ministry of Health of the Czech Republic, which is freely available through a web application [25] and is updated several times a day. It aggregates data from individual regional hygienic stations that is reported from individual laboratories performing tests.

Interesting from a technical point of view is the evaluation of how the data presentation as such is created. The systematic overview offers information on specific technological solutions – commercial solutions are dominated by the ArcGIS Dashboard platform from Esri, while open source applications use open libraries – primarily Leaflet and Mapbox – for which they develop necessary extensions. Interestingly, to address the current crisis Esri has made a large part of its products available for free for a limited time. In terms of how the data is processed, there are two main groups. The first is a basic presentation of information in the form of an interactive web map. Users can thus easily obtain simple descriptive information related to a specific region. An interactive map is for example used to present information by [8, 11, 15, 16, 24, 25, 26, 27]. The second group, involving deeper and more sophisticated processing, offers complex dashboards integrating several interconnected elements into one application (e.g. map, graph, text, infographic, table). Examples of such a solution are sources [1, 2, 3, 4]. A special group of data sources involves web applications produced by the media. As these applications are aimed at the general public, they need to be as easy as possible to read and are expected to be up-to-date and correct. Most often this involves a simple map dealing with one topic, perhaps with basic interactivity. The systematic overview shows that for these reasons, the media mainly use interactive maps to present information (sources [8, 9, 10, 11, 26]), which are clear and easy to read.

The overview of web map applications features products designed in various technological environments, published by various authors with various focus. All these aspects can play an important role in how well the information is presented, evaluated from a cartographic and geoinformatics perspective. Since these applications are used by a large number of ordinary users, it is appropriate, in addition to listing them, to briefly evaluate their positive or negative aspects.

Although the creation of such web products is largely influenced by the author's subjective feelings (choice of colours, composition of individual elements), which will not be evaluated in any way, there are a number of generally accepted rules that should not be violated. One of the main rules for correct cartographic expression of values is the use of a proportional symbols for absolute values and a choropleth map (or pseudo-choropleth) for relative values. This rule was followed in most cases, with examples of the correct use of a given method of expression being sources [1, 2, 3, 4, 7, 8, 25, 26]. Cases of an inappropriately used method were also identified [10, 18, 30], where absolute values were expressed through the intensity of the colour used to fill the given administrative unit. Some of the examined data sources do not provide the reader with sufficient tools for reading

Table 1 – Systematic overview of sources of map information on COVID-19

Title	Description	Technology	Geographical Coverage	Delimitation	Update	Data Source	Link	Category	Data Download	License	Author Group
[1] COVID-19 Global Cases	Johns Hopkins University, information on total infections, incidences, deaths, lethality. Includes numerical and graphical information.	ArcGIS Dashboards	global	states (China and Australia – provinces, states; USA – counties, cities)	several times a day	WHO, CDC, ECDC, NHC, DXY, 1point3acres, Worldometers.info, BNO, governmental institutions, media etc.	https://tinyurl.com/uwns6z5	dashboard	yes	commercial	university
[2] Robert Koch-Institut	Robert Koch Institute (Berlin, Germany) displaying information on COVID-19.	ArcGIS Dashboards	regional	states, regions	several times a day	Robert Koch-Institut	https://tinyurl.com/Koch-Institut	dashboard	no	commercial	university
[3] WHO – World	Dashboard divided into two parts – main map and graphic with various interactive visualizations according to selected categories.	Sprinklr	global	states	daily	WHO	https://who.sprinklr.com/	dashboard	yes	commercial	official institution
[4] WHO – Europe	Official WHO dashboard for the Europe.	ArcGIS Dashboards	regional	states	daily	ECDC	https://tinyurl.com/WHO-Europe	dashboard	no	commercial	official institution
[5] NPGeo Corona Hub	Geographic data and applications from the German National Platform for Geographic Data (NPGeo), an overview of national dashboards.	ArcGIS Dashboards	regional	states, European regions	unknown	various	https://npgeo-corona-npgeo-de.hub.arcgis.com/	dashboard, map application	yes	commercial	official institution
[6] European Union	European Center for Disease Prevention and Control, basic overviews of COVID-19 in the map and graphs, the possibility of comparison by users of selected countries over time.	JavaScript	global	states	unknown	ECDC	https://qap.ecdc.europa.eu/public/extensions/COVID-19/COVID-19.html	map application	no	unknown	official institution
[7] The New York Times	An interactive map with basic indicators about the disease, but also the rate of spread of the disease.	Mapbox	global, local	states	daily	Local governments, Johns Hopkins University, NHC China, WHO etc.	https://www.nytimes.com/interactive/2020/world/coronavirus-maps.html	interactive map	no	open source	media
[8] Seznamzpravy.cz	Information hub – number of tests performed, confirmed cases, deaths and number of recoveries.	mapy.cz	Czechia	regions	daily	Hygienic stations	https://www.seznam.cz/stitek/koronavirus-876681	interactive map	no	commercial	media
[9] idnes.cz	Graphs with basic statistical indicators on COVID-19. Map with absolute values of infected at the level of regions of Czechia and global overview.	JavaScript, SVG, Datawrapper	global, regional	states, Czech regions	daily	Ministry of Health, Johns Hopkins University	https://www.idnes.cz/koronavirus/statistiky	dashboard	yes	unknown	media
[10] Hospodářské noviny (Economic Newspaper)	A mixture of map visualizations (different platforms), graphs and tables.	JavaScript, Google Maps, Tableau	global	states, Czech regions	several times a day	Ministry of Health, ČTK, Worldometer, Johns Hopkins University	https://tinyurl.com/ihned	interactive map	no	commercial	media
[11] denik.cz	Overview of the number of infected, deceased and cured persons.	Infogram	global	states, regions (Italy and Spain)	several times a day	Johns Hopkins University	https://www.denik.cz/ze_sveta/koronavirus-mapa-svet.html	interactive map	no	commercial	media

Table 1 (cont.) – Systematic overview of sources of map information on COVID-19

Title	Description	Technology	Geographical Coverage	Delimitation	Update	Data Source	Link	Category	Data Download	License	Author Group
[12] Tableau COVID-19 Data Hub	Information on the total number of confirmed cases, new cases, deaths and daily changes. Offers instructions for preparing your own dashboard.	Tableau, MapBox	global	states	daily	Johns Hopkins University	https://www.tableau.com/covid-19-coronavirus-data-resources	dashboard	yes	commercial	private company
[13] COVID-19 GIS Hub	Esri's initiative to bring together map applications and datasets about the disease. Instructions on how to build own map application, including the possibility of professional support from Esri.	Esri	global	states, selected regions	daily	various (more than 50)	https://coronavirus-resources.esri.com/	dashboard, map application	yes	commercial	private company
[14] Esri dashboards overview	The Story Maps website provides an overview of dashboards created on the Esri platform.	Esri	global	states, selected regions	unknown	various	https://storymaps.arcgis.com/stories/a1746ada9bff48c09ef76e5a788b5910	dashboard	no	commercial	private company
[15] Worldometer	Key overviews by states on the number of infected, active cases, deaths, critically ill patients, the ratios of infected and deceased to the population, the number of tested. Contains both numerical and graphical information, including interactive maps and animations.	JavaScript	global	states	real-time	more than 25 sources (WHO, NHC, CDC, media, national institutions etc.)	www.worldometers.info/coronavirus	interactive map, animation	yes	open source	interest group
[16] HealthMap	HealthMap, in collaboration with universities and other volunteers, offers basic insights and animations of developments over time.	Mapbox	global (with exceptions)	states, regions, cities	daily	various	https://www.healthmap.org/covid-19/	interactive map	yes	open source	interest group
[17] Nextstrain	Data on the genetic development of the virus and its spread around the world	Leaflet, Mapbox	global	states	daily	various laboratories	https://nextstrain.org/ncov	dashboard	no	open source	interest group
[18] TheWuhanVirus	Basic overview information in map and non-spatial form.	JavaScript	global, local	selected regions	twice a day	WHO, Worldometers	https://coronavirus.thebaselab.com/	dashboard	yes	open source	individual
[19] CleverMaps	In addition to the basic indicators associated with COVID-19, the map application also shows the results of the economic impacts of the disease.	unknown	global	states	daily	Johns Hopkins University, Worldometers	https://coronavirus.clevermaps.io/	dashboard	no	commercial	private company
[20] TKP GEO COVID 19	Numbers of infected, cured and deceased by states with the possibility of displaying the trend over time.	Leaflet	global	states	unknown	Johns Hopkins University	https://www.tkpgeo.cz/amg/koronavirus/public_html/app/#	interactive map	no	open source	private company
[21] T-Mapy – World	Web application focused on data from the world, current information on the spread of the disease, visualization of trends of monitored indicators.	Corpis Maps	global	states	daily	Johns Hopkins University	https://app.corpismaps.com/maps/04510f00-e1f0-4a74-a6e8-4944c1bda9a8	dashboard	no	commercial	private company

Table 1 (cont.) – Systematic overview of sources of map information on COVID-19

Title	Description	Technology	Geographical Coverage	Delimitation	Update	Data Source	Link	Category	Data Download	License	Author Group
[22] Microsoft COVID-19 Tracker	Detail adjusted for individual countries, showing the usual indicators, supplemented by graphs and reports from individual areas.		global	states, regions	daily	CDC, WHO, ECDC	https://www.bing.com/covid	dashboard	no	commercial	private company
[23] Healthcybermap	Various sources of information, newspaper articles, data and maps about COVID-19. The site serves as a kind of "feed" of various data about the disease.	—	global	—	—	—	http://healthcybermap.org/WHO_COVID19/	webpage	—	open source	—
[24] Department of Geoinformatics in Olomouc	Basic data on COVID-19 sampling sites and closures in the Olomouc region. Locates and collects official reports and texts from the media.	Leaflet	regional	territory of municipalities with extended powers	several times a day	krajpomaha.cz, news and official announcements	https://gis.upol.cz/covid/mapa/	interactive map	no	open source	university
[25] Ministry of Health of the Czech Republic	The main portal for finding official reference data on COVID-19 for Czechia.	JavaScript	regional	Czech regions	several times a day	Ministry of Health, Hygienic stations, NHIS	https://koronavirus.mzcr.cz/	interactive map	yes	open source	official institution
[26] Portal iRozhlas	Czech Radio portal offering comprehensive information on COVID-19, including tables, graphs and maps.	JavaScript, Leaflet	global, regional	states, selected regions, Czech regions	real-time	Ministry of Health, Johns Hopkins University, WHO, CDC, ECDC, NHC, DXY, Apify	https://www.irozhlas.cz/koronavirus	interactive map	yes	open source	media
[27] Google – COVID-19 map	Map with basic graphs of the most monitored indicators (number of cases, healed, deceased).	Google	global	states, regions	several times a day	WHO, Wikipedia	https://www.google.com/covid19-map/	interactive map	yes	commercial	private company
[28] Data against COVID-19	Joint activity of Czech technology companies and IT enthusiasts. Summary information about the disease and established projects, an overview of (open) data sources.	—	Czechia	—	daily	—	https://covid19cz.cz/projekty/data	—	yes	open source	interest group
[29] ARCDATA COVID-19	Dashboard based on the Esri platform showing the situation in Czechia	ArcGIS Dashboards	regional	Czech regions	unknown	Ministry of Health	https://www.arcdata.cz/covid-19 https://arcgis.is/DfaGj	dashboard	no	commercial	private company
[30] T-Mapy Covid ČR	Focused on Czech data. Current data on infection, visualization of additional information, such as place of infection, date of infection or filtering by age.	Corpis Maps	regional	Czech regions	unknown	Ministry of Health	https://app.corpismaps.com/maps/993f6d45-cdc8-4bda-814f-8d1ba671cf04?fbclid=IwAR0kEw0y1oO-tJkhS6T209VFf4A9gwauHPOwomOwfdRP3aRf_t37CzSjKES	dashboard	no	commercial	private company

information – an example of such insufficient presentation of data are sources [9, 11, 20], where there is no legend that would identify which specific phenomenon is being expressed in the map and what the limit values of the colour scale intervals are. In the case of [11] some additional text states that it is the number of infected individuals, but without specifying whether it is an absolute number (in which case it would be a poorly chosen method of expression) or a relative number (in which case the method is appropriate, but is compromised by the absence of a legend). In the case of source [16], the legend merely breaks down the number of cases into intervals by colour, but in the map each proportional symbol showing information has a different size that is not explained anywhere.

General cartographic rules are also often violated in web applications due to the availability of other web tools, thanks to which the missing information can be easily obtained interactively by hovering over or clicking on a specific element. A pop-up window then displays text telling the user the exactly nature of the information, so some users may argue that such a way of presenting data is fine. However, even when using attractive web technologies, the basic rules of cartography should not be ignored and the information provided to the user should be correct and accurate at first glance. In addition, interactivity may, in extreme cases, obscure the overall inappropriate composition of the solution, as in example [12] – although it is an interactive web application, there is no tool that would allow one to zoom in and enlarge areas with great map symbol density. It can again be argued that due to interactivity this is not necessary, but on the contrary, in order to monitor conditions in a particular locality via precise markings in the map, such functionality is necessary.

Evaluating the technological sophistication of individual solutions is a relatively subjective affair, as it depends on a number of factors such as price, functionality, availability, processing speed, and how quickly the application performs. However, this factor often decides whether the application is used and becomes popular, or whether users will give preference to a different one. For example, for records processed in the ArcGIS Dashboard [1, 2, 6], a longer web application load time was observed. This depends not only on the speed of the internet connection and overall browser configuration, but also on the complexity of the application itself. Dashboard applications are often very complex, integrating several web elements and data sources (mostly automatically updated), so often there is a compromise between speed and level of detail.

Comments on the quality of the solutions used show that although current web technologies offer many interesting tools for data presentation, these tools may not always be used properly and information may be misinterpreted. On the other hand, the current situation illustrates very well that one topic (and often also an identical data set) offers a considerable number of ways to visualize and analyze data. If the user's expertise allows, the obvious approach is to use visualizations

for inspiration and to formulate one's own idea of the situation using available source data through (spatial) analysis. The acquisition of useful data related to the COVID-19 pandemic and its subsequent analytic processing is presented in the following chapter on the example of changes in human behaviour during the pandemic.

3. Changes in the behaviour of the European population during the COVID-19 pandemic

One of the ways to study changes in human behaviour is to analyze data from social networks or mobile operators. Options for processing localized data can be best demonstrated on the example of data from social networks, as well as the context of big data analysis, which can be used to monitor the mobility of people (e.g. Noulas et al. 2012; Long, Jin, Joshi 2012; Noulas et al. 2011). Data from applications for sports activities is analyzed in a similar manner to determine the behaviour of drivers and their volumes (Sun 2017). In addition to data from social networks and mobile operators, one can also use data from the Google Location service (location tracking). This data is collected by Google if a Google Maps user agrees to share their location. In Google Maps, this is a specific service called Timelines, which allows the user to keep track of their movement. In connection with COVID-19, Google has published data on changes in people's behaviour as its contribution to assessing the impact of a pandemic. As mentioned in the introduction, this involves Community Mobility Reports data, which provides a percentage change in the presence of Google Location users in six basic types of activities/locations.

The analytical part of the article initially focuses on evaluating behavioural changes in connection with the spread of COVID-19 in Europe. Each of us has personal experience with restrictions due to government measures implemented to slow down and stop the spread of COVID-19. Nevertheless, each country in Europe has implemented different restrictions at different times. The following section describes the spatial pattern of the consequences of these restrictions and other geographical contexts.

3.1. Localized mobility data from Google

Data from Google Location was used to analyze changes in human behaviour. This data is in principle very similar to data from mobile operators (used for example in smart quarantine systems), but is usually more accurate, because the location of the device (most often a mobile phone) is determined not only based

on the location of nearby BTS stations but also by connection to Wi-Fi networks (by IP address) and especially via GPS (if this option is enabled on the device). In addition, a user's location is determined from past activity on websites and Google services, including from multiple devices (e.g. mobile phone, computer, tablet). The data then contains not only information about position, but also how accurately the given point's location was determined and the probability of what type of movement it exhibited (e.g. car, train, walking, running), or whether it remained stationary.

Hence it is possible to better determine from the data in what kind of vehicle the person was moving, whether they remained stationary, and positioning accuracy. The data is mainly used by Google for internal purposes, but each user can download their own data (in JSON format) for visualization and analysis. The data can for example be used to identify most frequent routes (Löchtefeld 2019), for population movement analysis (Ruktanonchai et al. 2018), evaluation of spatial patterns of movement, or attendance analysis (Romero 2019). The issue of personal data protection also plays an important role, as it is personal information. Even so, this data is starting to be used more frequently (Calabrese et al. 2013; Kahrík et al. 2015; Novák, Temelová 2012).

COVID-19 Community Mobility Reports (Google LLC 2020) data available at <https://www.google.com/covid19/mobility/> was used to assess changes in the behaviour of Europeans following the outbreak of the COVID-19 pandemic. Initially, data was published at the end of March and start of April in PDF format (later also in CSV format) with the main information on the average decline in population activity in geographical detail, either only for the entire country or for the country and its main regions. The data is constantly being updated, but Google provides no information on the date of future dissemination of data. The data is anonymized and aggregated from Google Location users; there are different numbers (percentage of population) in different countries, which results in the (un)availability of data for some countries or their territorial units. In some cases, data for regions was provided on the basis of a very small number of users and hence was not sufficiently representative. A more detailed breakdown into regions in the case of the European countries used ranges from NUTS-1 (for example Germany), NUTS-2 (Spain), to NUTS-3 (Czechia). The United Kingdom and Ireland use a mixture of these units and the NUTS classification is not followed. Behavioural changes as such are monitored in six categories according to the location of the activity, which was determined by Google as useful for maintaining social distancing or from the point of view of the availability of basic services. The categories are as follows:

- a) Residential – places of usual residence.
- b) Transit Stations – transport hubs such as metro, bus, and train stations.
- c) Retail and Recreation – includes restaurants, cafés, shopping centres, amusement parks, museums, libraries, cinemas, etc.

- d) Grocery and Pharmacy – places where daily necessities can be purchased, such as grocery stores, wholesale grocery stores, supermarkets, farmers' markets, speciality shops, drugstores, pharmacies, etc.
- e) Parks – places for spending leisure time, such as local parks, national parks and protected areas, public beaches, harbours, dog parks, squares, public gardens, etc.
- f) Workplaces – places of employment.

Key information is the percentage decrease or increase in the number of individuals presence compared to the usual state of affairs (baseline) in these six fundamental localities. The percentage decrease/increase is calculated for each day retroactively from 15 February 2020, with the baseline being the median value of the activity at a given location derived again for each corresponding day of the week from a sample of data from 3 January to 6 February 2020. Simply put, Monday data is compared to the median of all Mondays from the reference data sample, and so on for the other days. For the purposes of this article, the first available Community Mobility Reports dataset was originally downloaded in PDF format and the activity change values were manually extracted from the document (only for selected days due to processing complexity) and connected to the GIS environment. After the data was updated in mid-April and made available in CSV format, all available days were processed and the average value from 5 March to 11 April 2020 (the last day for which data was available) was calculated from them. The date 5 March 2020 was chosen based on an examination of the beginning of changes in the general trend in the data and taking into account the history of the spread of COVID-19 in Europe. The resulting dataset, comprising 567 regions with six attribute information on activity in the above categories, was attached to spatial data regarding the administrative subdivision of Europe.

3.2. *Methods*

Behavioural change analysis was performed in several consecutive steps. From the point of view of data processing as such, these involve basic data processing without the need for GIS (Chapter 3.2.1). GIS tools were fully employed for geo-visual analysis and cartographic processing (Chapter 3.2.2) and cluster analysis (Chapter 3.2.3). For easier reproducibility of data processing, a detailed description is given in the following sections.

3.2.1. *Google Mobility Community Reports data processing*

The processing of Google Mobility Community Reports data used updated CSV (comma separated value) text format, which is available on the project website

<https://www.google.com/covid19/mobility/>. This is a sorted table where the rows give the monitored regions and the calendar date while the columns provide information on the percentage change of the six monitored categories. At the time of processing, the global dataset contained over 220,000 records. The data was then filtered in Excel using the area of interest (selected European countries and their subregions) and the selected time period. Using a contingency table, the average value of the change in behaviour in the monitored activities was calculated for each subregional unit. The modified table was then linked with geodata in the GIS environment, specifically to the corresponding administrative division, using the “Join” tool in the ArcGIS Pro. Because the Google data did not contain any unique identifier other than the name of the subregion, the nomenclature had to be unified so that join could be made without losing information. This step was time-consuming due to the manual search for of endonym and exonym equivalents in geographical nomenclature.

3.2.2. Geovisual analysis and cartographic processing of maps

After the data was edited and processed, geovisual analytics were commenced, which according to Andrienko et al. (2007) represents a scientific process of analytical reasoning in the decision-making process concerning geographic information displayed in the form of a visual interface (map). When studying geographical data, it is recommended that one should start the analysis by displaying it (more in Pászto 2020) and searching for a basic geographical context, which can be accompanied by a quantitative evaluation of similarity of the resulting maps (e.g. Dvorský, Snášel, Voženílek 2009). To understand the pattern of changes in human behaviour in Europe's regions, each of the six monitored activities was visualized in the form of a pseudo-choropleth map. The Jenks method was used to set the intervals for individual categories; this method minimizes differences within the interval and maximizes differences between them (Jenks 1967). With regard to providing information on the percentage change in behaviour, no other distribution was considered, for example by standard deviation or an otherwise standardized range of intervals. The limit values from the Jenks method were manually adjusted to meet the basic cartographic rules for setting intervals in maps (Voženílek, Kaňok et al. 2011). All map visualizations employed a Lambert azimuthal uniform-area projection designated ETRS89-LAEA, which is officially used by the European Union for its map products.

3.2.3. Cluster analysis and typology

The methodologically most demanding part of the analysis of population behaviour change data was the creation of a typology based on the results of cluster analysis,

which is why we present its very detailed description at this point. A visual analysis of the spatial representation of the categories from the Community Mobility Reports reveals certain characteristic features of the behaviour, which reflect the similarity of regions (administrative units) in terms of the monitored categories of change in occurrence. However, this is a very empirical, time-consuming, and to some extent also subjective method. Multiple variables can be analyzed simultaneously to provide groups of types of regions with common properties. Cluster analysis in n -dimensional space can be used to describe the typology of administrative units hidden in this data. This overall task of processing multi-dimensional data aims to objectively collect objects (in this case, administrative units) into groups based on their similarities and differences (Tryon 1939). Meloun and Militký (2011) state that cluster analysis belongs among methods that deal with the investigated similarity of multidimensional objects and their classification into classes or clusters. The successful use of multidimensional methods and cluster analyses for geographic data processing is described in many scientific articles (e.g. Shaffer, Zolnik 2014; Novák, Netrdová 2011; Netrdová, Nosek 2016), including those from the authors of this article (e.g. Pászto et al. 2019; Macků, Voženílek 2019; Marek, Pászto, Tuček 2015).

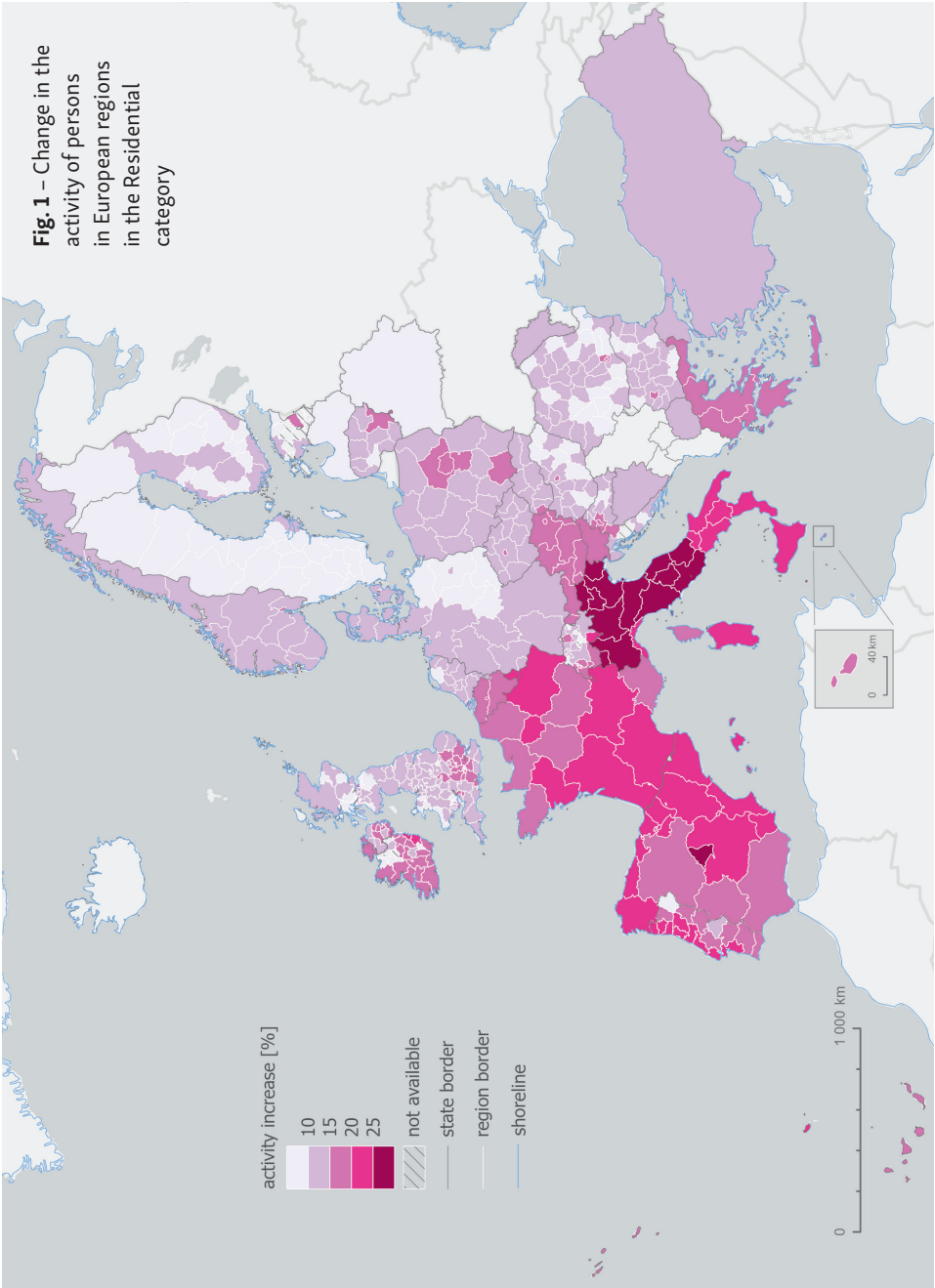
The k -medoid method was chosen for processing data on human activity type due to its lower sensitivity to outliers (Ng, Jiawei Han 2002). An important step in cluster analysis is to select an optimal number of clusters. By calculating auxiliary numerical statistics (Pseudo F-index, gap statistics, silhouette method), the optimum number of clusters was determined to be two. However, splitting the area of interest into two (or three) groups does not provide a significant understanding of the patterns of similarity associated with COVID-19. Therefore, other subdivision options were investigated based on the evolution of Pseudo F-index values for more than three clusters. The pseudo F-index describes the ratio of between-cluster variance to intra-cluster variance (Caliński, Harabasz 1974), and thus measures cluster quality for different numbers of clusters. In our case a significant decrease occurred between three and four clusters (the value of the Pseudo F-index for three clusters is 388.6 and for four clusters is 322.9 – a decrease of 60). Due to the previous exclusion of clustering into three types of regions and taking into account that the range from four to seven clusters did not exhibit a further significant decrease in the Pseudo F-index value (from 322.9 to 262.6 with a progressive decrease between clusters of only about 20), following an expert assessment five clusters were chosen. This also gives readers a finer breakdown of European regions, which more appropriately shows the diversity/similarity of individual types of regions (especially in the case of areas adjacent to the countries most severely afflicted by COVID-19 – for example Portugal, southern Austria, Slovenia, and others). Cluster analysis was calculated using ArcGIS Pro (the Multivariate Clustering tool with a random seed location initiation method).

3.3. Results

3.3.1. Geovisual analysis of behavioural changes

It is very interesting to examine the differences in the general population behaviour trend between individual regions and states, knowing what restrictions have been put in place at national levels. It must be acknowledged that European Union institutions, as supranational umbrella institutions covering most European countries, have failed to synchronize autonomous steps taken by individual governments of the Member States to stop the uncontrollable spread of COVID-19. However, it is important to note that the area of healthcare (epidemiology) in Member States is not covered by any central apparatus of the European Union (or not as significantly as in the case of legislation or monetary policy). Nevertheless, each of the European countries introduced restrictions of different intensity and with different timing according to the severity of the situation or future expectations, which in turn is reflected in the intensity of use of public and market services by residents of individual countries. Segmentation of the decline and increase of activities on a pan-European scale with detailed focus on the regions is therefore a very appropriate way to describe and understand the effects of the pandemic on “normal” public life and behaviour.

In the following text, all activity categories are shown and commented briefly. The *Residential* category (Fig. 1) showed an increase only in activity in the place of usual residence and was only category where there was no decrease in activity (i.e. people did not spend significantly more time outside their place of residence). This is quite understandable, especially in view of the restrictions on the free movement of people that were introduced sooner or later by all analyzed countries. Nevertheless, it is interesting to observe the spatial distribution of percentages. The map shows that the most affected regions of northern and central Italy and the Madrid region in Spain show an increase in activity of more than 25 percent. Restrictions also had varying degrees of impact on other regions in this respect, namely southern Italy, most regions in France, and regions in northern and eastern Spain. A markedly contrasting situation exists in Sweden and Belarus, where the approach to COVID-19 is far more moderate than in the rest of the other countries. In the case of Sweden, the impact on the behaviour of the population may be due to socio-cultural factors, where the generally greater confidence of Swedes in state institutions has led to moderate restrictions. Like in other Nordic countries, close social contact with multiple people is less common in Sweden. In addition, according to Eurostat, Sweden has around 52 % single-person households (Eurostat 2020), the highest in Europe. The approach to spending the maximum amount of time in one's place of residence split Germany in a very interesting manner, with the country's former division into West and East Germany being evident. This can

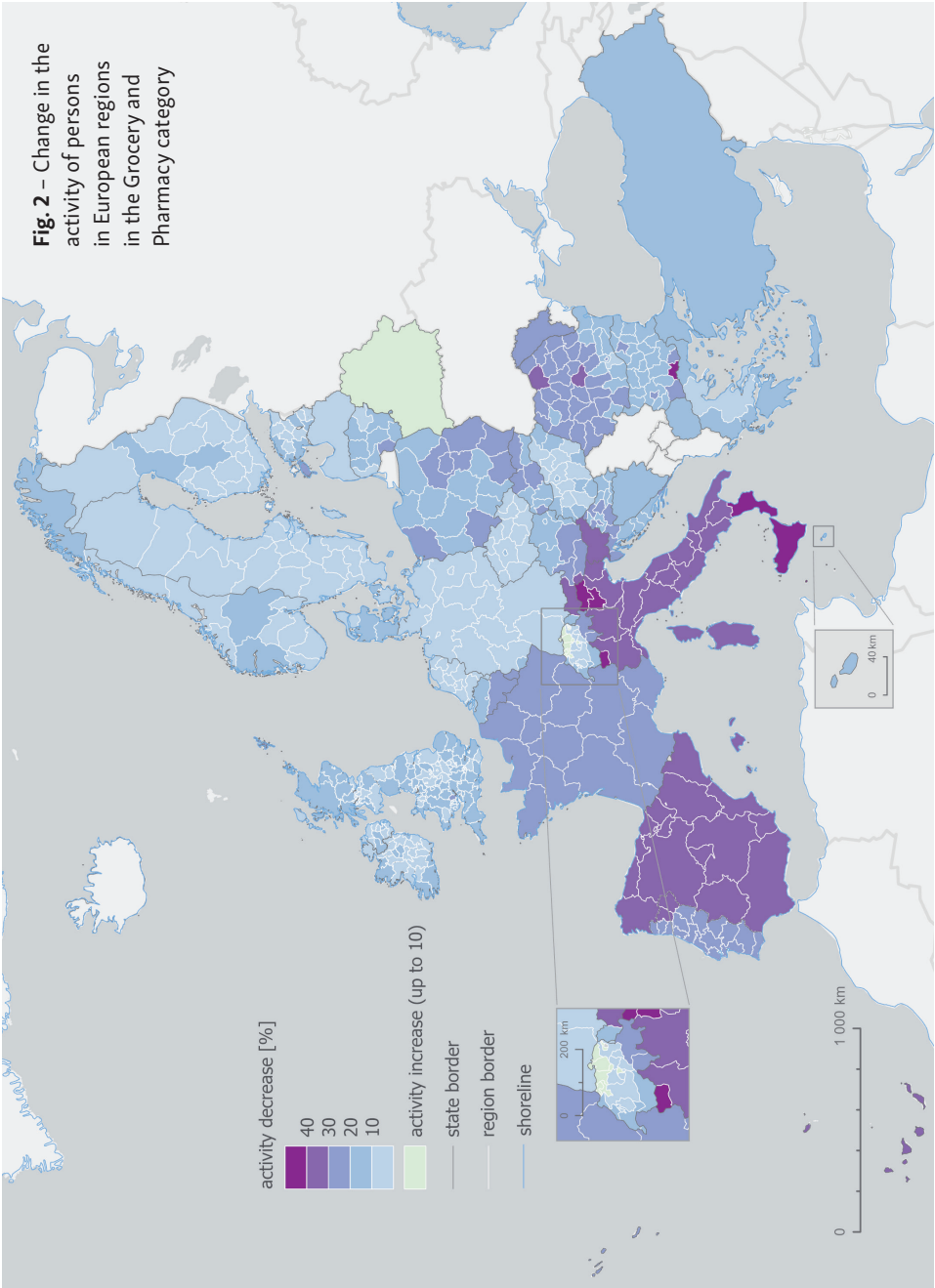


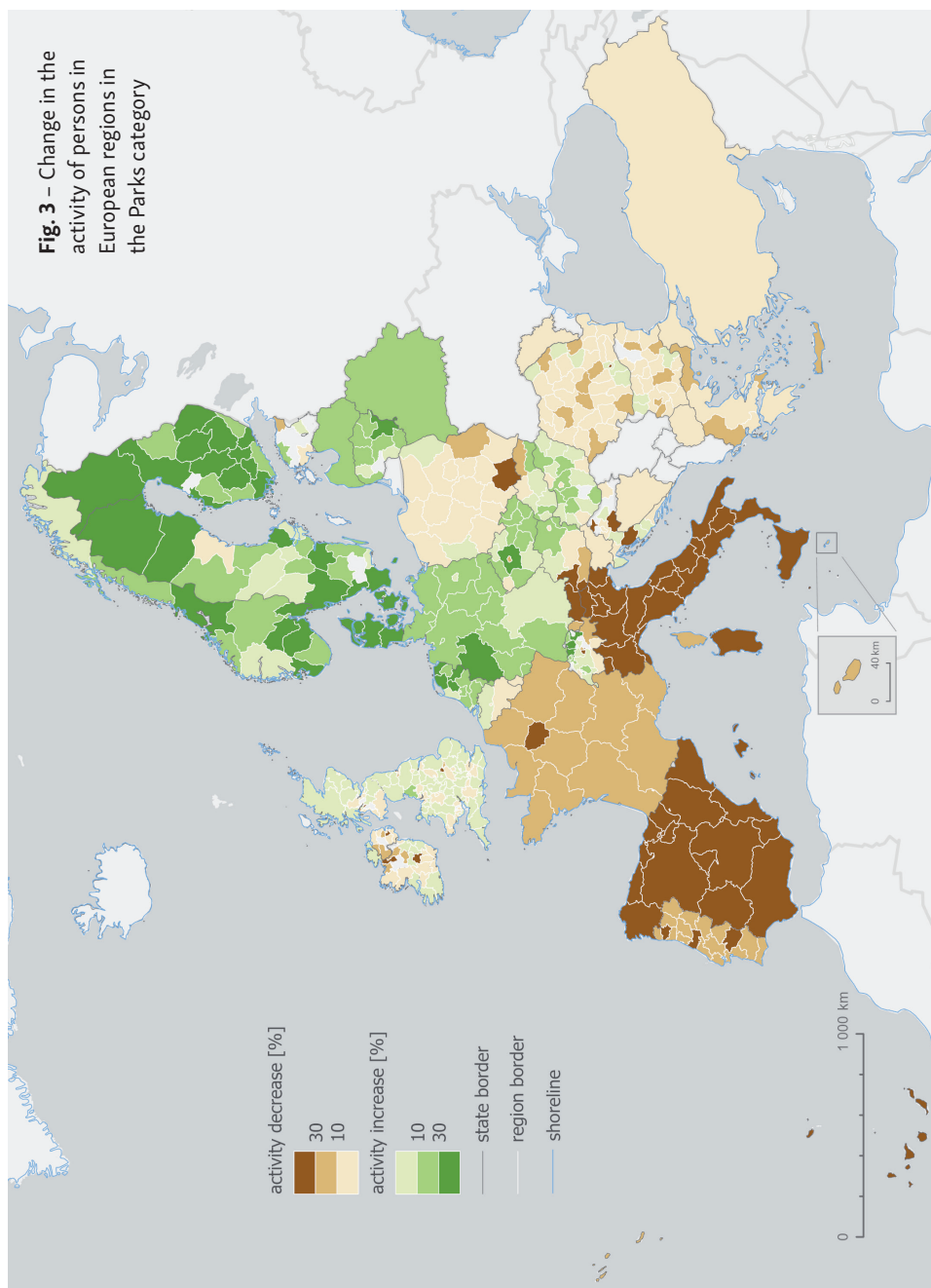
be attributed to persisting behavioural stereotypes of the former East German population, whose different sociodemographic characteristics can still be observed in the east part of Germany (Klüsener, Goldstein 2016). A possible relationship between widespread vaccination of the population against tuberculosis (with higher resistance to coronavirus as a possible side effect) and the impact of COVID-19 are noted by preprints of medical articles (Shet et al. 2020), where compulsory vaccination ceased later in low-income countries than in high-income countries.

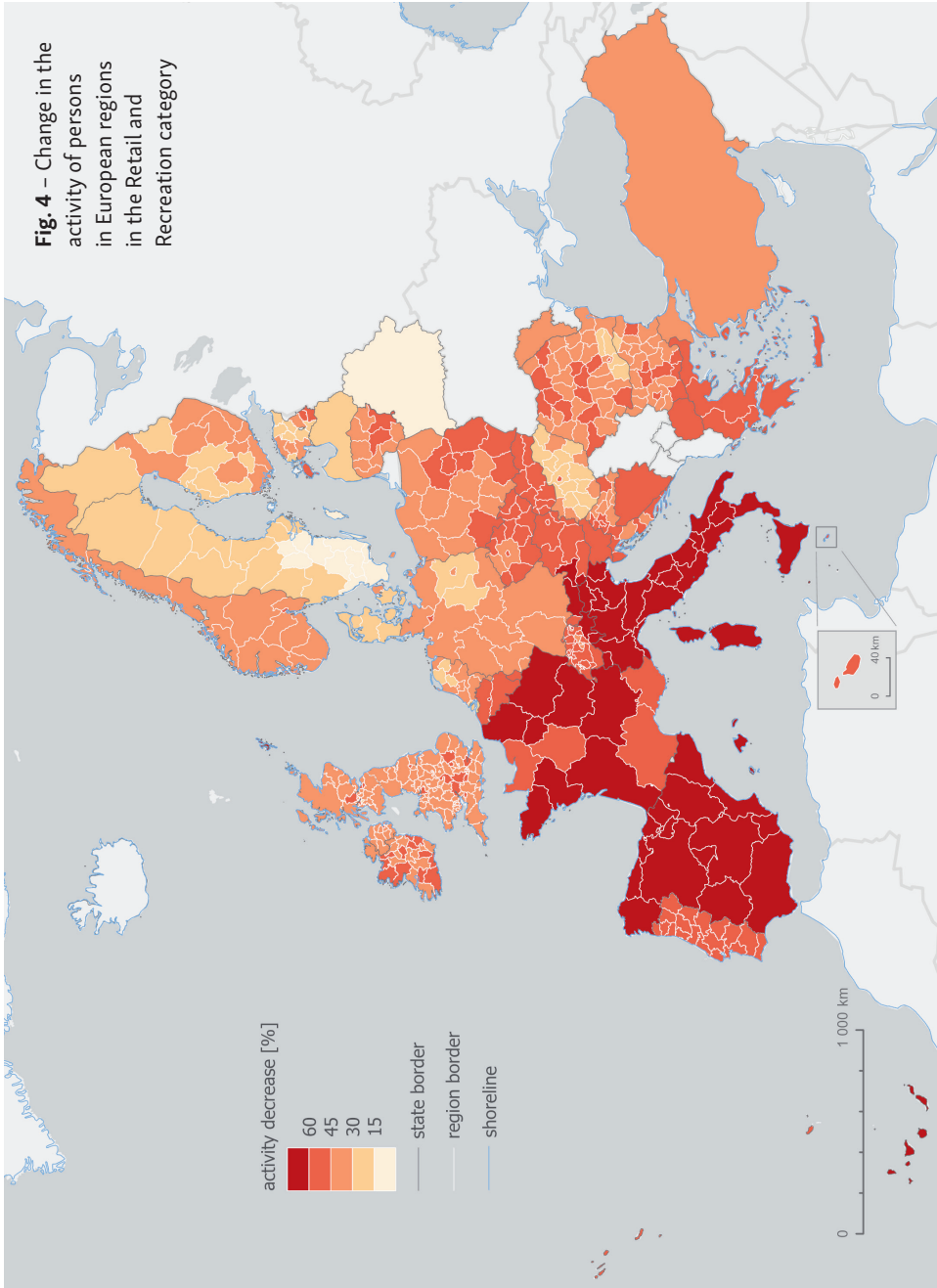
Another category in which there was an increase in activity in at least some regions during the monitored period of the COVID-19 pandemic is *Grocery and pharmacy* (Fig. 2). By the way, Belarus, which is an exception in all categories (more precisely, no significant behavioural changes occurred compared to the baseline), and the cantons of northern Switzerland achieved growth rates of up to ten percent, which in the case of Belarus is clearly related to the regime's approach. As far as the aforementioned Swiss cantons are concerned, apart from Zürich these are mostly regions less affected by COVID-19. Switzerland is an outlier in other categories as well, which may be due to the fact that the federal government did not introduce restrictions on the movement of persons even at the height of the COVID-19 pandemic. However, in other European countries, a decrease in the frequency of visits is also seen in shops with staple goods, with a decrease of more than 40% in the worst-hit areas of northern Italy. However, the dramatic decline in Sicily and Calabria, or in Bulgaria's mountainous region of Smolyan on the border with Greece require further interpretation.

The *Parks* category (Fig. 3) is the most dichotomous group of activities. Here, there was a dramatic decline in the use of leisure and green spaces, as well as a significant increase. In countries where people had no choice but to be at home and at the same time were allowed to spend time outside under certain conditions, there was an increase (for example, Germany, Czechia, Slovakia, Lithuania, etc.). The *Parks* category is an interesting category for noting activity values in Great Britain. No spatial pattern of human behaviour can be recognized, and increases and decreases were not significant (up to ten percent). One possible reason why nothing changed from the baseline is the government's long-term hesitation to introduce stricter measures against the spread of the disease.

A relatively significant change in the behaviour of the population in the area of "non-essential" services is evident in the *Retail and Recreation* category (Fig. 4). This is logical given that in places such as shopping malls, restaurants, bars and the like, the greatest and closest contact among persons occurs. And these were the activities that were for the most part limited first of all in the spirit of social distancing. In most regions of Europe, except for parts of eastern Germany, Denmark, Sweden, Finland, Latvia, as well as parts of Hungary and Estonia, there was an overall decrease in activities of more than thirty percent. All of Spain, Italy, and half of France achieved values greater than sixty percent.





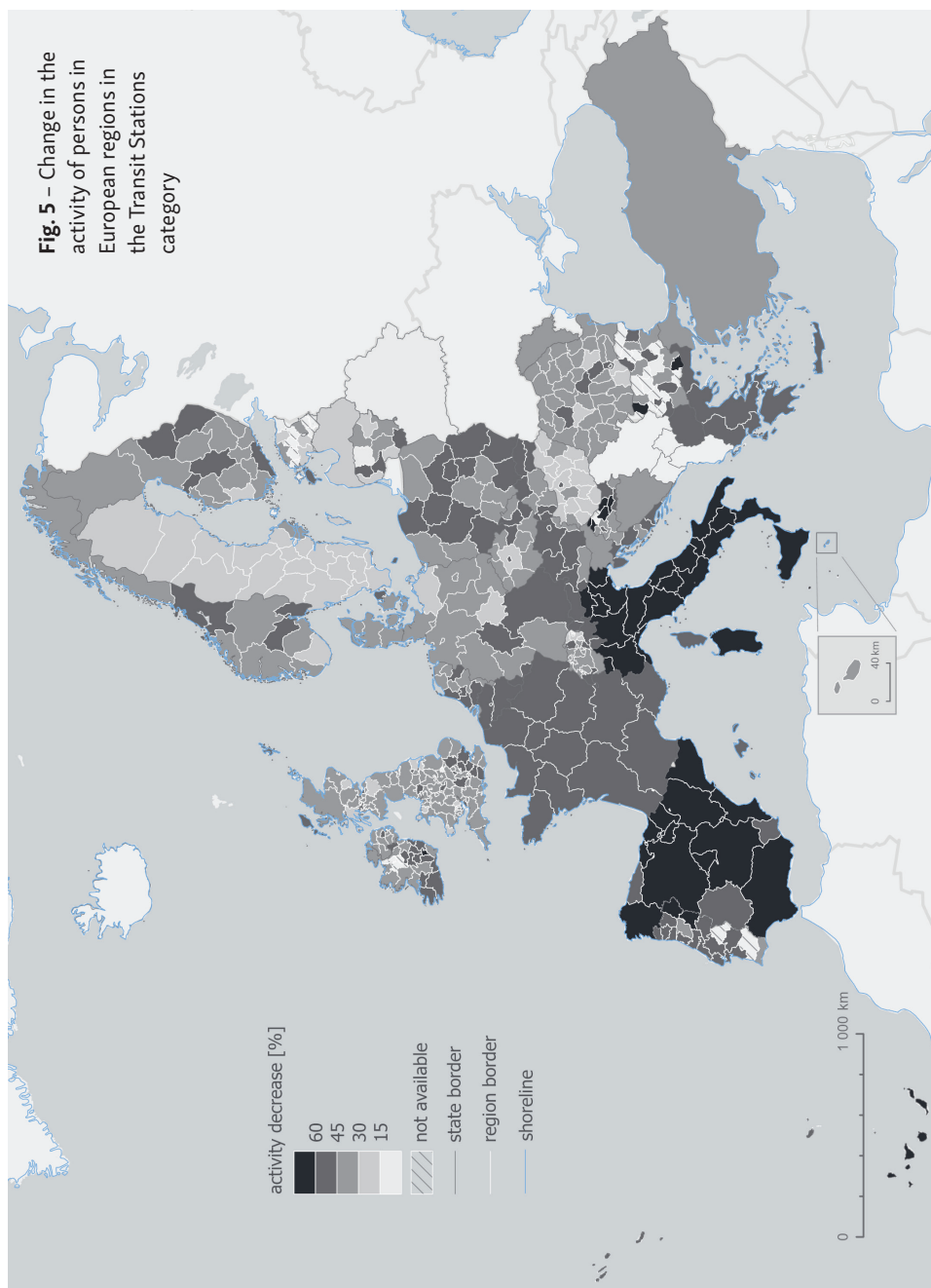


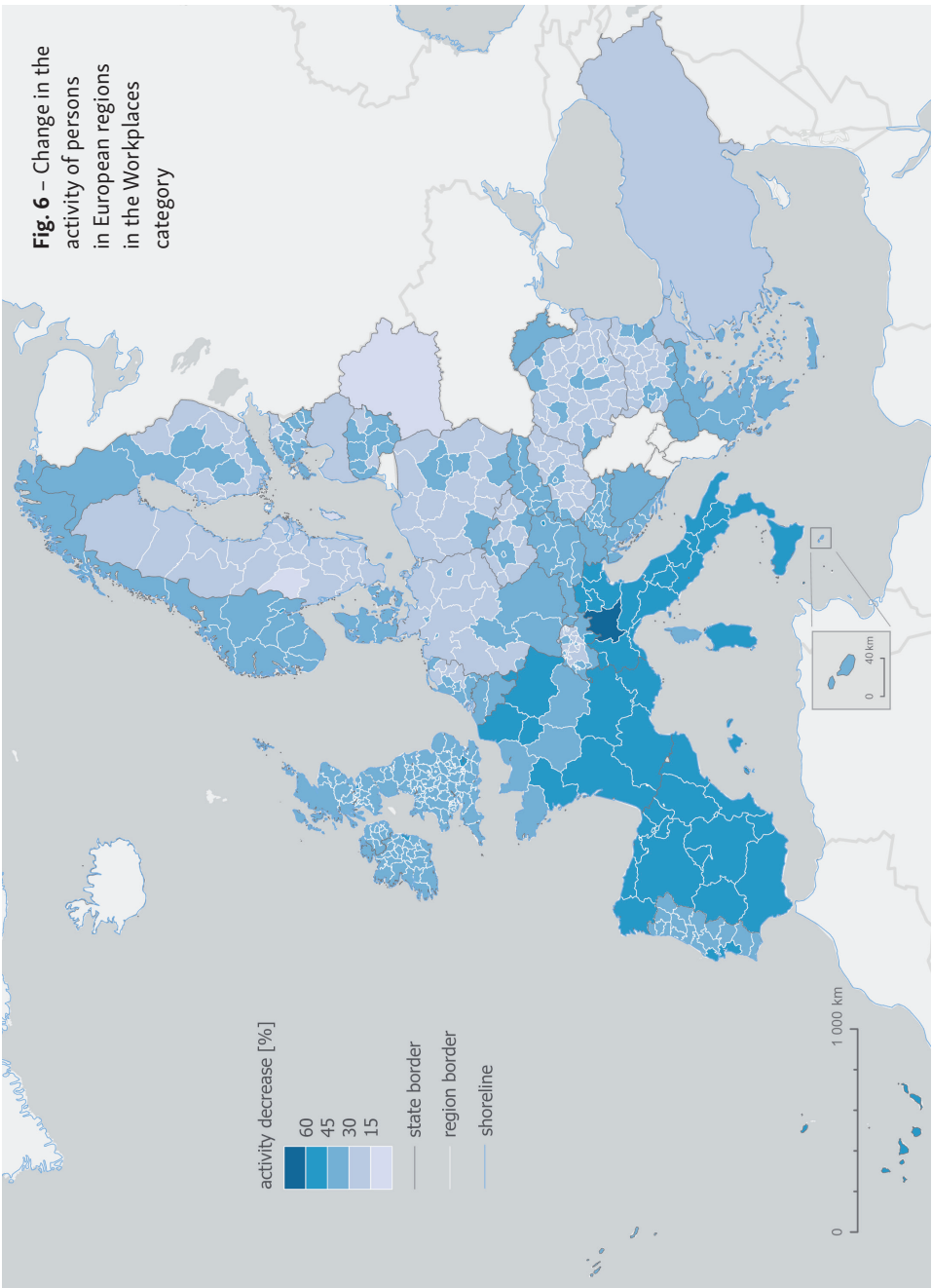
A similar spatial pattern of decline in activities can also be seen in the other two categories – *Transit Stations* (Fig. 5) and *Workplaces* (Fig. 6). In *Transit Stations* it is clear that restrictions at the level of regional (national) transport or even urban public transport (Spain and Italy) strongly affected the presence of people at important transport hubs. Relatively low declines in people's mobility can be seen in Sweden and Belarus, but also, for example, in most of Hungary or in the Central Bohemian Region (Czechia). The last category is *Workplaces*. Hypothetically, the rate of decline in activity at workplaces should correspond to the regions of increase in activity in places of usual residence. Visually, this is indeed the case – again, the worst-hit regions of Italy, Spain, and France follow this trend. A kind of buffer zone exists around these regions in areas spanning from Belgium, southwestern Germany, and Austria to the Balkans. Further to the northeast, the relative decline in the *Workplaces* category is not as great, and when looking at the activities associated with *Residential* (Fig. 1), one can identify a similar spatial pattern. Greece, which does not especially deviate in most categories, is worthy of attention. The government's emphasis on Greek behaviour was taken so seriously by the population that Greece avoided the Italian or Spanish scenario. Yet according to Eurobarometer results (European Commission 2020), Greek trust in institutions is the same or worse than in the two aforementioned countries. Greek confidence in their national government (26%) is similar to that in Italy (25%) and Spain (21%), nevertheless in the category of regional/local authorities it is 27% (compared to 30% in Italy and 40% in Spain) and in the category of public administration, the confidence of Greeks again ranks last, at 23% compared to Italy (26%) and Spain (38%).

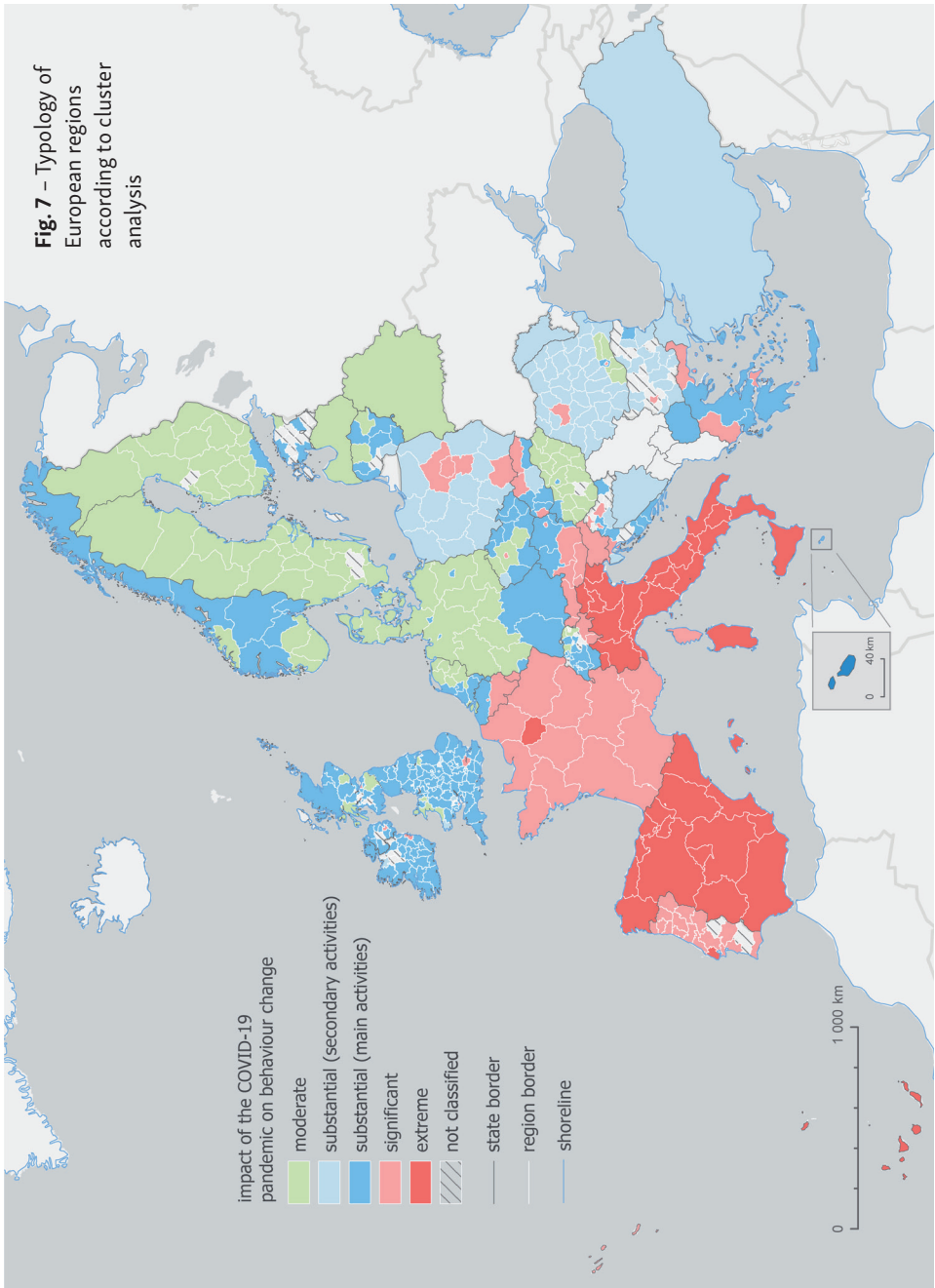
3.3.2. Typology of European regions using cluster analysis

The resulting typology categorizes the relevant administrative units into five clusters (types) in the space defined by the change in behaviour in the six Community Mobility Reports data categories (Fig. 7). These clusters can also be named according to the impact of the COVID-19 pandemic on behaviour, also taking into account the median values in Figure 8:

- Moderate – regions with a relatively small change in population activity. For this type, it must especially be emphasized that this involves a mild impact of the pandemic compared to other types and that COVID-19 and related restrictions have had a visible effect on human behaviour in these regions as well.
- Substantial – secondary activities – regions with a generally statistically average change in population activity with a more significant change in secondary activities (*Grocery and Pharmacy, Parks*).
- Substantial – main activities – regions with a generally statistically average change in population activity with a more significant change in main activities (*Residential, Transit Stations, Workplaces*).







- Significant – regions with a statistically significant change in behaviour (values in the fourth quartile outside the range of the boxplot).
- Extreme – regions with an unprecedented change in behaviour in outliers or near their borders

The least distinguishable types are regions with a substantial impact of the COVID-19 pandemic. Therefore, two summary activity categories were created to more accurately characterize both types. *Residential, Transit Stations, and Workplaces* were defined as main activities. The authors rely on the logical sequence of these activities, which distinguishes the third type from the second, in which the values of changes in human behaviour in these main activities are closer to the values of regions of moderate type. Ancillary to the main activities are secondary activities, i.e. the *Retail and Recreation, Grocery and Pharmacy, and Parks* categories.

The median of the standardized values of the input data for individual clusters (Fig. 8) can be used to interpret these clusters. Figure 9 provides a more detailed examination of the variability of input data within each cluster and category of activity.

Using these supporting visualizations, the following typological characteristics were identified. All monitored activities except for the *Residential* category follow

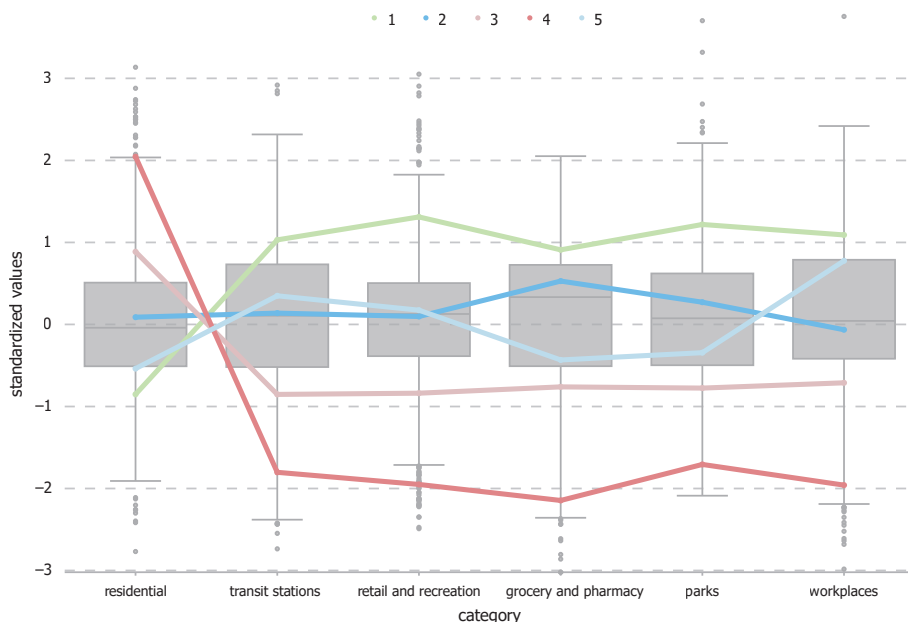


Fig. 8 – Median of standardized values of input data for individual types of regions: 1 – moderate, 2 – substantial (main activities), 3 – significant, 4 – extreme, 5 – substantial (secondary activities)

a similar trend, which organizes individual clusters according to the decrease/increase of values. *Residential* follows a trend that is opposite to all the others. The largest decrease in incidence in all public activities is exhibited by the extreme type of regions, followed by regions with significant impacts of the pandemic. On the contrary, in the *Residential* category, these two types of regions have the most significant increase. The spatial pattern of the regions in these two types involves the south-western regions of Europe, covering the territories of Portugal, Spain, France, Italy, Slovenia, and selected regions of Austria, Switzerland, and Belgium. In the context of past events, these countries are characterized by very strict restrictions (Italy, Spain, and France) or are in the buffer zone around these countries (southern Belgium, Austria, and Slovenia). The other, similar subgroup consists of the type of regions with a substantial impact from the pandemic. The observed changes are not so intense here (the median of the standardized values of these clusters is within the second and third quartiles, which confirms the low variability of the percentage values of human activities). Nevertheless, the changes in the main activities (*Residential*, *Transit Stations*, *Workplaces*) are not as markedly evident in the type with a substantial impact of the pandemic on secondary activities. The values in these categories even approximate those in regions with a moderate impact from the COVID-19 pandemic. It can be concluded that

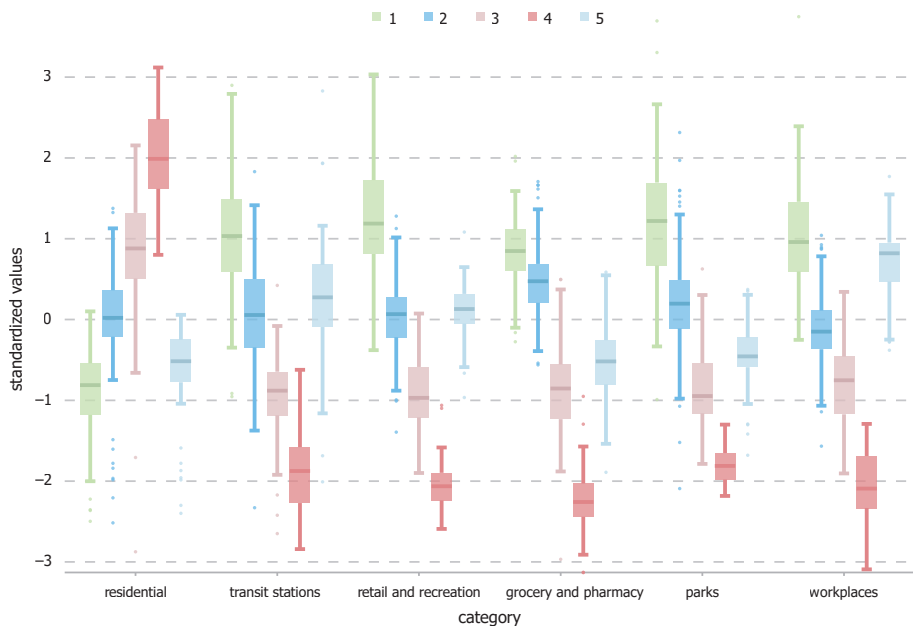


Fig. 9 – Variability of values of the resulting types of regions in monitored activities: 1 – moderate, 2 – substantial (main activities), 3 – significant, 4 – extreme, 5 – substantial (secondary activities)

all three of these activities are closely related – a relatively small change in time spent at work implies smaller increase of time spent in the place of residence. This is connected with traffic between the two places being preserved. Geographically, this type covers Poland, Romania, Bulgaria, Bosnia and Herzegovina, and Turkey. On the other hand, the type with a substantial impact on the main activities is characterized by a median increase in activities in the *Residential*, *Transit Stations*, *Retail and Recreation*, and *Workplaces* categories compared to the previous type, but “merely” a slight decrease in activities in the *Grocery and Pharmacy* and *Parks* category. Geographically, the administrative units of this type of region exhibit no obvious pattern and are spatially fragmented. They cover almost the whole of Great Britain, Norway, and Greece. A cross-border whole formed in Central Europe consisting of a belt stretching from Switzerland through the regions of southern Germany and northern Austria to the border areas of Czechia and Slovakia. In the context of the COVID-19 pandemic, these are areas cannot be clearly characterized in terms of established restrictions and the severity of the spread of COVID-19. With a few exceptions, it can be said that these are regions affected by the pandemic to an average degree with relatively strict restrictions on human activity. Activity increase or decrease was affected by all events to an average degree in comparison with other types (moderate, significant, and extreme). The fact that the United Kingdom belongs to the “average” type can be explained by the late introduction of restrictions, and thus by the smoothing of the trend in the input data (more significant and deeper decreases in the values of activities occur starting about 18 March 2020). Finally, regions with a moderate impact on behaviour change stand out against all others with only a slight decrease in the activities of persons in five of the six monitored categories and a slight increase in the *Residential* category. This makes it completely different from all other types of clusters. In some cases, this corresponds to a gradual tightening of the measures in place (e.g. the Netherlands, Sweden, individual Länder in Germany). In general, three key core areas can be identified for this type: (1) the belt of regions stretching from the Netherlands through the central part of Germany north to Denmark, Sweden, and Finland, (2) Baltic regions and Belarus, and (3) Hungary.

4. Discussion

The timely acquisition of verified information is crucial for individual and collective decision-making at this time of crisis. As stated in the first part of the article, all data sources used to deal with the current situation come from official or trustworthy entities. Thanks to the technical development of GIS and related technologies, it is now possible to use this data to create various map visualizations and applications. The professional community understands the added value of graphic

information in the form of a map, but it is also necessary to maintain perspective and admit that the widespread use of GIS and the ability to use visualizations has its practical limits. The first limit is the availability of tools for the creation of maps to the broader public. Incorrectly processed data and incorrect visualizations generated as a result (albeit unwittingly and in good faith) can lead to false conclusions. In most cases, the use of GIS ends with analysis and visualization – but GIS does not explain the observed phenomena, but rather only helps with their interpretation. A concrete example of this could be the dashboard from Johns Hopkins University listed in the overview of map data sources in this article. This probably most frequently used dashboard merely displays data and implicitly fails to provide an explanation of the terms. In contrast, the Czech website providing information on COVID-19, though not primarily a map application, supplements these statistics in a very appropriate manner with explanations.

The second limit is data availability and validity. The analytical part presented in this article is based on user data collected by Google. Data for the analysis of European regions (Chapter 3.1), i.e. data on the incidence of people in certain locations, is influenced by the number of people sharing their location. This number may be different in each region. Google has not disclosed basic information on the absolute number of unique users of its location services that it used to create the dataset, and the authors were unable to find more details in this regard. Google itself mentions those limits and limitations in the description of the Community Mobility Reports data. The very definition of six categories of monitored places is open to debate – for example, they do not include educational institutions or government bureaus.

In developed countries, having a mobile phone is taken for granted, and therefore mobile operator data can be used as the best way to track the true location of the population over time. The problem with this data, however, is its unavailability, whether due to price, incompleteness, or other objective criteria. However, if we look at mobile phone ownership from the opposite perspective, a limiting factor can be that we do not actually have data on people who do not own a mobile phone (and who can often be part of marginalized groups such as the homeless or disadvantaged). In connection with the thematic context of the article, COVID-19, other research questions may arise regarding the effects of the pandemic on the lives of these societal groups.

On the other hand, Google Community Mobility Reports is a completely unique dataset that provides a valuable insight into how the COVID-19 pandemic impacts human behavioural patterns. If we look past the data itself, we must for example not forget the practical factors that influence human behaviour before and during a pandemic. Investigating the influence of the weather on individual categories (for example, in the *Parks* category) or evaluating the increase/decrease in activity during the pandemic due to the arrival of spring could bring interesting results.

In addition, historical and cultural customs in given countries (with respect to important Catholic holidays and customs) can affect human behaviour. In the text, we (intentionally) ignored the now broadly discussed topic of influence of the economy on the pandemic, and vice versa. This may be another reason why some regions do not exhibit a very significant decline in the *Workplaces* category while others do.

It is also necessary to mention the methods that can be used to process data from Community Mobility Reports. European data from Community Mobility Reports is first evaluated using geovisual analysis, which is based primarily on maps. It is generally known that by, for example, splitting values into different intervals, very different forms of visualization can be achieved. The use of a specific clustering method and choosing the number of target clusters (types) can also be discussed. Although these are formalized methods based on statistical calculations, it is always up to the user how they set the individual parameters. That is also why we have described all the necessary settings in detail so that the procedure is easily reproducible.

The use of Google Location data to analyze and evaluate user behaviour in the COVID-19 pandemic crisis is a very unique approach that has not yet been published anywhere. Though aggregation of data into higher territorial units and a uniform approach to the time period analyzed may in some cases reduce the validity of conclusions, the results can still be considered remarkable. The current global situation clearly shows that positional and individual data can also be useful in the case of pandemics or any other situations involving the safety and health of the population. At the same time, however, it is necessary to approach this data in a very sensitive manner and to take scrupulous care to protect personal data. This makes it all the more important to work with this data and to point out the possibilities it offers as well as the limits of its processing and subsequent interpretation. As we indicated in the introduction, we hope that our text will inspire other specialists in various disciplines to engage in their own research.

5. Conclusion

The current COVID-19 pandemic is, above all, a major problem facing humanity throughout virtually the entire world. For scientists and geographers, including those working in the area of geoinformatics, it is, however, also a significant opportunity in the current situation to help better understand its consequences, how to overcome it, and how to prevent it in the future. Detailed analyzes of spatial and statistical data are proving to play an important role, especially in activity to help stop the spread of the pandemic, for example through smart quarantine systems and the tracing of infected individuals in order to identify potential new infections.

For these reasons, this article focuses primarily on geographic data and possibilities for its visualization and analysis. The intent of this article's authors was to point out existing sources of data and information that are directly related to the current COVID-19 pandemic. The introductory part of the article showed and evaluated selected sources of data and information in the form of maps and map applications that monitor the current situation, are freely available to the general public, and that the authors consider to be the most relevant. Due to the open nature of their source data, some of them directly support the creation of other outputs, which can be presented anywhere in the world thanks to distributed environment technologies.

The second part of the article focused on processing data from Google in order to describe changes in the behaviour of Europeans. An analysis of anonymized Community Mobility Reports data was used to show and describe the spatial arrangement of European regions according to changes in human behaviour in the six monitored categories. Five types of regions were identified according to the impact of COVID-19 and related restrictions – from the type moderately impacted by COVID-19 (e.g. Sweden, Latvia, Hungary), to those impacted by the pandemic in a substantial (Ireland, regions of Greece, Czechia, Norway, Switzerland), significant (e.g. regions of France, Belgium, Austria, Slovenia, Portugal), and extreme manner (Spain, Italy, and the Paris region).

Currently, COVID-19 is a priority area of research. Leaving aside research in medicine and chemistry, epidemiological and geographical studies are next in line. However, most articles published so far focus chiefly on analyses and predictions of the spread of the disease, and there are still very few other studies describing changes in human behaviour. This is why the authors consider this article to be important, especially for the geographers and humanities-oriented scientists, who can the data presented in the article for further follow-up studies.

References

- ANDRIENKO, G., ANDRIENKO, N., JANKOWSKI, P., KEIM, D., KRAAK, M.J., MACEACHREN, A., WROBEL, S. (2007): Geovisual analytics for spatial decision support: Setting the research agenda. *International Journal of Geographical Information Science* 21, 8, 839–857.
- Apple Inc. (2020): Apple and Google partner on COVID-19 contact tracing technology, <https://www.apple.com/cz/newsroom/2020/04/apple-and-google-partner-on-covid-19-contact-tracing-technology/> (8.5.2020).
- BATTY, M. (2013): Big data, smart cities and city planning. *Dialogues in Human Geography*, 3, 274–279.
- BATTY, M. (2016): Big data and the city. *Built Environment*, 42, 3, 321–337.
- CALABRESE, F., DIAO, M., DI LORENZO, G., FERREIRA, J., RATTI, C. (2013): Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation Research Part C: Emerging Technologies*, 26, 301–313.

- CALIŇSKI, T., HARABASZ, J. (1974): A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods*, 3, 1, 1–27.
- CARROLL, M.W., MATTHEWS, D.A., HISCOX, J.A., ELMORE, M.J., POLLAKIS, G., RAMBAUT, A., ABDELLATI, S. (2015): Temporal and spatial analysis of the 2014–2015 Ebola virus outbreak in West Africa. *Nature*, 524, 7563, 97–101.
- CHARU, V., ZEGER, S., GOG, J., BJØRNSTAD, O.N., KISSLER, S., SIMONSEN, L., VIBOUD, C. (2017): Human mobility and the spatial transmission of influenza in the United States. *PLoS computational biology*, 13, 2, e1005382.
- COVID-19 NERC (National Emergency Response Center), Epidemiology, Case Management Team, Korea Centers for Disease Control, Prevention (2020): Contact Transmission of Covid-19 in South Korea. Novel investigation techniques for tracing contacts. *Osong Public Health Research Perspectives*, 11, 1, 60–63.
- COX, M., ELLSWORTH, D. (1997): Application-controlled demand paging for out-of-core visualization. In *IEEE Proceedings – Visualization'97*, 235–244.
- DELLICOUR, S., ROSE, R., PYBUS, O.G. (2016): Explaining the geographic spread of emerging epidemics: a framework for comparing viral phylogenies and environmental landscape data. *BMC bioinformatics*, 17, 1, 82.
- DVORSKÝ, J., SNÁŠEL, V., VOŽENÍLEK, V. (2009): Map similarity testing using matrix decomposition. In *2009 International Conference on Intelligent Networking and Collaborative Systems*. 290–294.
- DZÚROVÁ, D., JAROLÍMEK, J. (2020): Šíření pandemie COVID-19 napříč geografickými i sociálními hranicemi: dokážeme mu čelit? *Geografie*, 125, 1, 1–20.
- DESJARDINS, M.R., HOHL, A., DELMELLE, E.M. (2020): Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. *Applied Geography*, 102202.
- Eurostat (2020): Over half of Sweden's households made up of one person, Eurostat, <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/DDN-20170905-1?inheritRedirect=true> (8.5.2020).
- European Commission (2020): Public Opinion, European Commission, <https://ec.europa.eu/commfrontoffice/publicopinion/index.cfm/General/index> (8.5.2020).
- FERGUSON, N.M., CUMMINGS, D.A., FRASER, C., CAJKA, J.C., COOLEY, P.C., BURKE, D. S. (2006): Strategies for mitigating an influenza pandemic. *Nature*, 442, 7101, 448–452.
- FISHER, D., WILDER-SMITH, A. (2020): The global community needs to swiftly ramp up the response to contain COVID-19. *Lancet*. S0140-6736(20)30679-6.
- Google LLC (2020): Google COVID-19 Community Mobility Reports, Google LLC. <https://www.google.com/covid19/mobility/> (22.5.2020).
- HUANG, R., LIU, M., DING, Y. (2020): Spatial-temporal distribution of COVID-19 in China and its prediction: A data-driven modeling analysis. *The Journal of Infection in Developing Countries*, 14, 3, 246–253.
- HUANG, C., WANG, Y., LI, X., REN, L., ZHAO, J., HU, Y., CHENG, Z. (2020): Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*, 395, 10223, 497–506.
- JENKS, G.F. (1967): The data model concept in statistical mapping. *International yearbook of cartography*, 7, 186–190.
- KAHRIK, A., NOVÁK, J., TEMELOVÁ, J., KADARIK, K., TAMMARU, T. (2015): Patterns and Drivers of Inner City Social Differentiation in Prague and Tallinn. *Geografie*, 120, 2, 275–295.

- KAMEL BOULOS, M.N., GERAGHTY, E.M. (2020): Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic and associated events around the world: how 21st century GIS technologies are supporting the global fight against outbreaks and epidemics. *International Journal of Health Geographics*, 19, 8.
- KISSLER, S.M., GOG, J.R., VIBOUD, C., CHARU, V., BJØRNSTAD, O.N., SIMONSEN, L., GRENFELL, B.T. (2019): Geographic transmission hubs of the 2009 influenza pandemic in the United States. *Epidemics*, 26, 86–94.
- KLÜSENER, S., GOLDSTEIN, J.R. (2016): A long-standing demographic East–West divide in Germany. *Population, Space and Place*, 22, 1, 5–22.
- KOCH, T. (2005): *Cartographies of disease: maps, mapping, and medicine*. Redlands, CA (USA): Esri Press.
- LI, C., CHEN, L.J., CHEN, X., ZHANG, M., PANG, C.P., CHEN, H. (2020a): Retrospective analysis of the possibility of predicting the COVID-19 outbreak from Internet searches and social media data, China, 2020. *Eurosurveillance*, 25, 10, 2000199.
- LI, Q., GUAN, X., WU, P., WANG, X., ZHOU, L., TONG, Y., XING, X. (2020b): Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New England Journal of Medicine*.
- LEMEY, P., RAMBAUT, A., WELCH, J.J., SUCHARD, M.A. (2010): Phylogeography takes a relaxed random walk in continuous space and time. *Molecular biology and evolution*, 27, 8, 1877–1885.
- LONG, X., JIN, L., JOSHI, J. (2012): Exploring trajectory-driven local geographic topics in four-square In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing – UbiComp '12*.
- LÖCHTEFELD, M. (2019): Detournavigator – Using google location history to generate unfamiliar personal routes *Conference on Human Factors in Computing Systems – Proceedings*, 1–6.
- MACKŮ, K., VOŽENÍLEK, V. (2019): Statistical synthesis of quality of life indicators – design of index construction in European regions. *Geographia Cassoviensis*, 13, 2, 196–209.
- MAREK, L., PÁSZTO, V., TUČEK, P. (2015): Using clustering in geosciences: examples and case studies. *15th International Multidisciplinary Scientific Geoconference, Informatics, Geoinformatics and Remote Sensing*, 2, 1207–1214.
- MHCR (2020): Chytrá karanténa. Ministry of Health of the Czech Republic, <https://koronavirus.mzcr.cz/chytra-karantena/> (18.4.2020).
- MELOUN, M., MILITKÝ, J. (2011): *Statistical Data Analysis A Practical Guide*. In: Meloun, M., Militký, A. (eds.): New Delhi: Woodhead Publishing India.
- NETRDOVÁ, P., NOSEK, V. (2016): Spatial patterns of unemployment in Central Europe: Emerging development axes beyond the Blue Banana. *Journal of Maps*, 12, 4, 701–706
- NG, R.T., JIAWEI HAN (2002): CLARANS: a method for clustering objects for spatial data mining. *IEEE Transactions on Knowledge and Data Engineering*, 14, 5, 1003–1016.
- NOULAS, A., SCELLATO, S., LAMBIOTTE, R., PONTIL, M., MASCOLO, C. (2012): A tale of many cities: Universal patterns in human urban mobility. *PLoS ONE*, 7, 5.
- NOULAS, A., SCELLATO, S., MASCOLO, C., PONTIL, M. (2011): An Empirical Study of Geographic User Activity Patterns in Foursquare *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, (January), 570–573.
- NOVÁK, J., NETRDOVÁ, P. (2011): Spatial patterns of socioeconomic differentiation in the Czech Republic at the level of municipalities. *Sociologický časopis / Czech Sociological Review*, 47, 4, 717–744.

- NOVÁK, J., TEMELOVÁ, J. (2012): Everyday Life and Spatial Mobility of Young People in Prague: A Pilot Study Using Mobile Phone Location Data. *Sociologický časopis / Czech Sociological Review*, 48, 5, 911–938.
- PANETH, N. (2004): Assessing the contributions of John Snow to epidemiology: 150 years after removal of the broad street pump handle. *Epidemiology*, 15, 5, 514–516.
- PÁSZTO, V. (2020): Introduction to Spatial Exploration of Economic Data. In: Pászto, V., Jürgens, C., Tominc, P., Burian, J. (eds.): *Spatonomy – Spatial Exploration of Economic Data and Methods of Interdisciplinary Analytics*. Springer, Cham.
- PÁSZTO, V., MACKŮ, K., BURIAN, J., PÁNEK, J., TUČEK, P. (2019): Capturing cross-border continuity: The case of the Czech-Polish borderland. *Moravian Geographical Reports*, 27, 2, 122–138.
- ROMERO, F.M. (2019): Creation of images in virtual map location services. The case of Google Maps Hipertext.Net: *Revista Académica Sobre Documentación Digital y Comunicación Interactiva*, 0, 18, 66–76.
- ROMERO-ALVAREZ, D., PARIKH, N., OSTHUS, D., MARTINEZ, K., GENEROUS, N., DEL VALLE, S., MANORE, C. A. (2020): Google Health Trends performance reflecting dengue incidence for the Brazilian states. *BMC infectious diseases*, 20, 1–15.
- RUKTANONCHAI, N.W., RUKTANONCHAI, C.W., FLOYD, J.R., TATEM, A.J. (2018): Using Google Location History data to quantify fine-scale human mobility. *International Journal of Health Geographics*, 17, 1, 1–13.
- SANTANGELO, O.E., PROVENZANO, S., PIAZZA, D., GIORDANO, D., CALAMUSA, G., FIRENZE, A. (2019): SHORT PAPER Digital epidemiology: assessment of measles infection through Google Trends mechanism in Italy. *Annali di igiene: medicina preventiva e di comunità*, 31, 385–391.
- SHAFFER, R., ZOLNIK, E. (2014): The geographic distribution of physician assistants in the US: Clustering analysis and changes from 2001 to 2008. *Applied Geography*, 53, 323–331.
- SHET, A., RAY, D., MALAVIGE, N., SANTOSHAM, M., BAR-ZEEV, N. (2020). Differential COVID-19-attributable mortality and BCG vaccine use in countries. medRxiv.
- SIRKECI, I., YUCESAHIN, M.M. (2020): Coronavirus and Migration: Analysis of Human Mobility and the Spread of Covid-19. *Migration Letters*, 17, 2, 379–398.
- SNOW, J. (1856): Cholera and the water supply in the south districts of London in 1854. *Journal of Public Health, and Sanitary Review*, 2, 7, 239.
- SUN, Y. (2017): Exploring potential of crowdsourced geographic information in studies of active travel and health: Strava data and cycling behaviour In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences – ISPRS Archives*
- TRYON, R.C. (1939): Cluster analysis: Correlation profile and orthometric (factor) analysis for the isolation of unities in mind and personality. Ann Arbor: Edwards Brothers.
- VOŽENÍLEK, V., KAŇOK, J. et al. (2011): *Metody tematické kartografie: vizualizace prostorových jevů*. Palacký University Olomouc for the Department of Geoinformatics.
- WHO (2020): Coronavirus disease 2019 (COVID-19) Situation Report 87. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>, (20 April 2020).
- Worldometers.info (2020): COVID-19 Coronavirus Pandemic. Dover, Delaware (USA), <https://www.worldometers.info/coronavirus/> (19.4.2020).
- WALTER, S.D. (2000): Disease mapping: a historical perspective. In: Elliott P., Wakefield J., Best, N., Briggs D.J. (eds.): *Spatial Epidemiology: Methods and Applications*. Oxford University Press, Oxford, 223–252.

- YASAKA, T.M., LEHRICH, B.M., SAHYOUNI, R. (2020): Peer-to-Peer Contact Tracing: Development of a Privacy-Preserving Smartphone App. *JMIR mHealth and uHealth*, 8, 4, e18936.
- ZHOU, C., SU, F., PEI, T., ZHANG, A., DU, Y., LUO, B., SONG, C. (2020): COVID-19: challenges to GIS with big data. *Geography and Sustainability*. In press.

ACKNOWLEDGEMENTS

This article received support from project IGA_PrF_2020_027 and project ERASMUS+ 2019-1-CZ01-KA203-061374. The authors would like to thank Lukáš Příleský for his help with data collection.

ORCID

VÍT PÁSZTO

<https://orcid.org/0000-0003-3432-2473>

JAROSLAV BURIAN

<https://orcid.org/0000-0003-0729-9757>

KAREL MACKŮ

<https://orcid.org/0000-0002-5899-2626>