

Kuebart, Andreas; Stabler, Martin

**Article — Published Version**

## Waves in time, but not in space – an analysis of pandemic severity of COVID-19 in Germany

Spatial and Spatio-temporal Epidemiology

**Provided in Cooperation with:**

Leibniz Institute for Research on Society and Space (IRS)

*Suggested Citation:* Kuebart, Andreas; Stabler, Martin (2023) : Waves in time, but not in space – an analysis of pandemic severity of COVID-19 in Germany, Spatial and Spatio-temporal Epidemiology, Elsevier BV, Amsterdam, Vol. 47,  
<https://doi.org/10.1016/j.sste.2023.100605>

This Version is available at:

<https://hdl.handle.net/10419/295112>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

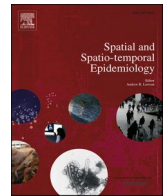
*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by-nc-nd/4.0/>



# Waves in time, but not in space – an analysis of pandemic severity of COVID-19 in Germany

Andreas Kuebart<sup>a,b,\*</sup>, Martin Stabler<sup>a,c</sup>

<sup>a</sup> Leibniz Institute for Research on Society and Space, Flakenstraße 29-31, Erkner 15537, Germany

<sup>b</sup> Brandenburg Technical University, Platz der Deutschen Einheit 1, Cottbus 03046, Germany

<sup>c</sup> Free University of Berlin, Malteser Str. 74-100, Berlin 12249, Germany

## ARTICLE INFO

### Keywords:

Sars-cov-2

Covid-19

Space-time clusters

Spatial autocorrelation

Lisa

Infectious disease diffusion

## ABSTRACT

While pandemic waves are often studied on the national scale, they typically are not distributed evenly within countries. This study presents a novel approach to analyzing the spatial-temporal dynamics of the COVID-19 pandemic in Germany. By using a composite indicator of pandemic severity and subdividing the pandemic into fifteen phases, we were able to identify similar trajectories of pandemic severity among all German counties through hierarchical clustering. Our results show that the hotspots and cold spots of the first four waves were relatively stationary in space. This highlights the importance of examining pandemic waves on a regional scale to gain a more comprehensive understanding of their dynamics. By combining spatial autocorrelation and spatial-temporal clustering of time series, we were able to identify important patterns of regional anomalies, which can help target more effective public health interventions on a regional scale.

## 1. Introduction

The COVID-19 pandemic, like any other pandemic, followed a pattern of acceleration and deceleration concerning case incidences among populations and regions. Over the course of a pandemic, this creates a temporal pattern in the form of logistic curves often described as wave(s). These waves consist of (fast) ascents, (up to several) peaks, and declines, indicating several stages or phases of a pandemic. While pandemic waves are often analyzed on the national scale, they typically are not distributed evenly within territories (Cliff et al., 2009; Teller 2021; Boterman 2022; Keeler and Emch 2018).

The spread of a disease is determined by its characteristic transmission dynamics on the one side and the measures implemented for public health control, such as vaccinations and non-pharmaceutical interventions (NPIs) on the other hand. Initially, disease diffusion of COVID-19 was significantly influenced by overdispersion and super-spreading events, especially before NPIs were introduced (Kuebart and Stabler 2020; Schmitz et al., 2023). Furthermore, environmental factors such as seasonal effects have been found to influence COVID-19 transmission rates (Coccia 2022).

As a result, the spatial patterns of infections and pandemic severity vary over time, which is why a spatio-temporal perspective is necessary

to understand the spread of infectious diseases (Ghosh and Cartone, 2020). In the case of COVID-19, several studies have found COVID-19 infections to be clustered within countries (Scarpone et al., 2020; Murgante et al., 2020; Rodríguez-Pose and Burlina 2021), within greater urban areas (Boudou et al., 2023) and even within cities or districts (Siljander et al., 2022; Lambio et al., 2023). The findings on how these clusters change over time are less consistent. While Kim et al. (2021) find wavelike patterns in space for the case of South Korea, Boterman (2022) finds no consistent patterns for the case of the Netherlands, and D'Angelo et al. (2021) find the trajectories of Italian regions to be relatively independent of each other. However, all these studies focus on the early stage of the pandemic. They consider the first months or the first year of the pandemic and give valuable insights into the complex dynamics of the diffusion of COVID-19. To reflect the context of timely changing variables of the pandemic, a longer study period may be beneficial. Similarly, in the German context studies dealing with COVID-19 from a spatial perspective explored the propagation effect of commuting patterns as channels of disease transmission (Mitze and Kosfeld 2022), spillover effects of German border regions (Chilla et al., 2022), pandemic developments in association with variants of concern (Mitze and Rode 2022), epidemiological analysis of the first wave (Küchenhoff et al., 2021) and spatio-temporal disease monitoring

\* Corresponding author at: Leibniz Institute for Research on Society and Space, Flakenstraße 29-31, Erkner 15537, Germany.

E-mail address: [andreas.kuebart@leibniz-irs.de](mailto:andreas.kuebart@leibniz-irs.de) (A. Kuebart).

<https://doi.org/10.1016/j.sste.2023.100605>

Received 31 December 2022; Received in revised form 5 June 2023; Accepted 14 July 2023

Available online 17 July 2023

1877-5845/© 2023 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

(Rohleder and Bozorgmehr 2021). Again, these studies have opted for examining the initial phase of the pandemic and restricting their analysis to concise temporal periods.

In contrast, this study contributes to the literature by taking a long-term perspective and tracing the spatio-temporal patterns of the COVID-19 pandemic in Germany throughout the years 2020 and 2021. By employing a novel approach to analyze the spatio-temporal dynamics of the COVID-19 pandemic we aim to analyze the regional variation in how the pandemic unfolded and examine the different trajectories of pandemic severity between German counties. Further, we want to highlight the importance of examining pandemic waves not only on a national scale but also on a regional scale to gain a more comprehensive understanding of their dynamics.

The first two years of the COVID-19 pandemic in Germany were analyzed by developing a novel method, for which four analytical steps were performed: First, we develop a composite index of ‘pandemic severity’, which integrates the three sub-indicators of COVID-19 case incidence, the incidence of death due to COVID-19, and the incidence of COVID-19 patients on ICU (intensive care unit). Second, a phase model of fifteen pandemic phases is developed based on a change point analysis. Each of the fifteen stages in the model is coherent in terms of the trends of pandemic dynamics (e.g., rising, decreasing, stable). Third, the spatial pattern during each phase is analyzed by considering global and local spatial autocorrelation on the level of all 400 German counties. To analyze the spatio-temporal variation of pandemic severity we thus opted to analyze the spatial patterns of pandemic severity during the whole phases instead of relying on snapshots on specific dates. Fourth, hierarchical clustering was performed to identify types of similar trajectories of pandemic severity in German counties.

The paper is structured as follows: The following section describes how we proceeded and highlights methodological reasoning for the pandemic severity index, the phase model, the spatial analysis, and the hierarchical clustering. The subsequent section presents the findings for each step of the analysis. The final section discusses the results and concludes.

2. Materials

Our analysis is based on a composite indicator, which is composed of three sub-indicators: The incidence of cases (IC) tested positive for COVID-19, the incidence of patients on ICU (IICU), and the incidence of registered deaths due to COVID-19 (ID, see Table 1). All datasets for the pandemic severity index were accessed via Healthcare Daten Plattform,<sup>1</sup> Data on COVID-19 cases and deaths originate from the federal Robert Koch Institute (RKI), data on patients with COVID-19 on ICU were used as reported by the German Association of intensive care physicians (DIVI). The three indicators were available at the county level (Kreise and Kreisfreie Städte, NUTS-3 level). There are 400 counties in

Table 1  
Elements of the pandemic severity composite indicator.

Indicator	Time range	Spatial resolution
Incidence of patients tested positive for COVID-19 (IC)	2020–02–01 –2021–12–31	400 counties
Incidence of deaths of patients tested positive for COVID-19 (ID)	2020–02–01 –2021–12–31	400 counties
Incidence of patients tested positive for COVID-19 on ICU (IICU)	2020–03–04 –2021–12–31	96 planning regions

<sup>1</sup> Healthcare Daten Plattform or formerly Corona Daten Plattform is a data repository that has been sponsored by the German Federal Ministry for Economic Affairs and Energy: <https://www.healthcare-datenplattform.de/>

Germany, of which 107 are independent cities. While all three sub-indicators were gathered at the county-level, we decided to use the regional average in 96 planning regions (BBSR 2017) for IICU, because hospitals are distributed rather unevenly over German counties and especially more specialized medical infrastructure such as ICUs tend to be concentrated in larger towns and cities, which tend to be their own county.

3. Methods

3.1. A composite index for pandemic severity

Studies focusing on the spatio-temporal aspects of the COVID-19 pandemic rely on the case incidence of patients who tested positive for COVID-19 almost exclusively as an indicator (Nazia et al., 2022). This is somewhat problematic, especially since the testing regime changed throughout the pandemic and potentially also varied regionally. For example, during the first wave of COVID-19 in Germany (March through May 2020), only limited capacities of PCR testing were available, whereas later PCR testing and antigen testing were widely available in 2021. Further it is likely that the relationships between the sub-indicators vary over time. For example, Coccia (2021) discovered that in Italy, the first wave had more severe consequences in terms of deaths and intensive care if compared to the second. Following suggestions to combine indicators to add robustness to the spatio-temporal analysis of COVID-19 (Pagel and Yates 2021; Rohleder and Bozorgmehr, 2021), we decided to develop a composite indicator of pandemic severity.

The components of a composite indicator need to be scaled, weighted, and aggregated (Libório et al., 2022). The pandemic severity index was calculated as follows: First, the rolling 14-day mean of each indicator (IC, ID, IICU) was calculated using the zoo package in R (Zeileis and Grothendieck 2005) and they were normalized by using z-scores. Second, the scaling of the indicators was completed by minimum-maximum normalization values to a scale between 0 and 1 (Wickham and Seidel 2022). Thus, each of the three sub-indicators was normalized and scaled to a dimensionless value between 0 and 1. Third, the three normalized and scaled sub-indicators were aggregated by using the arithmetic mean, in which all three indicators have the same weight. Subsequently, the composite index was rounded to four decimal places. This approach represents the common procedure to calculate composite indicators (Dialga and Giang 2017). The resulting pandemic severity index indicates the pandemic burden during a given time in each location through a dimensionless value indicating the mean of the three sub-indicators<sup>2</sup>. It was calculated for each day between 2020–03–01 and 2021–12–31 for each German county, resulting in 268, 400 individual values. We focused on the first two full years of the pandemic in Germany since we aimed to take a long-term perspective and considered the available data for this period as the best fit. The dataset has been made available on the Zenodo repository (Kuebart and Stabler 2023).

3.2. Identifying change points of pandemic severity

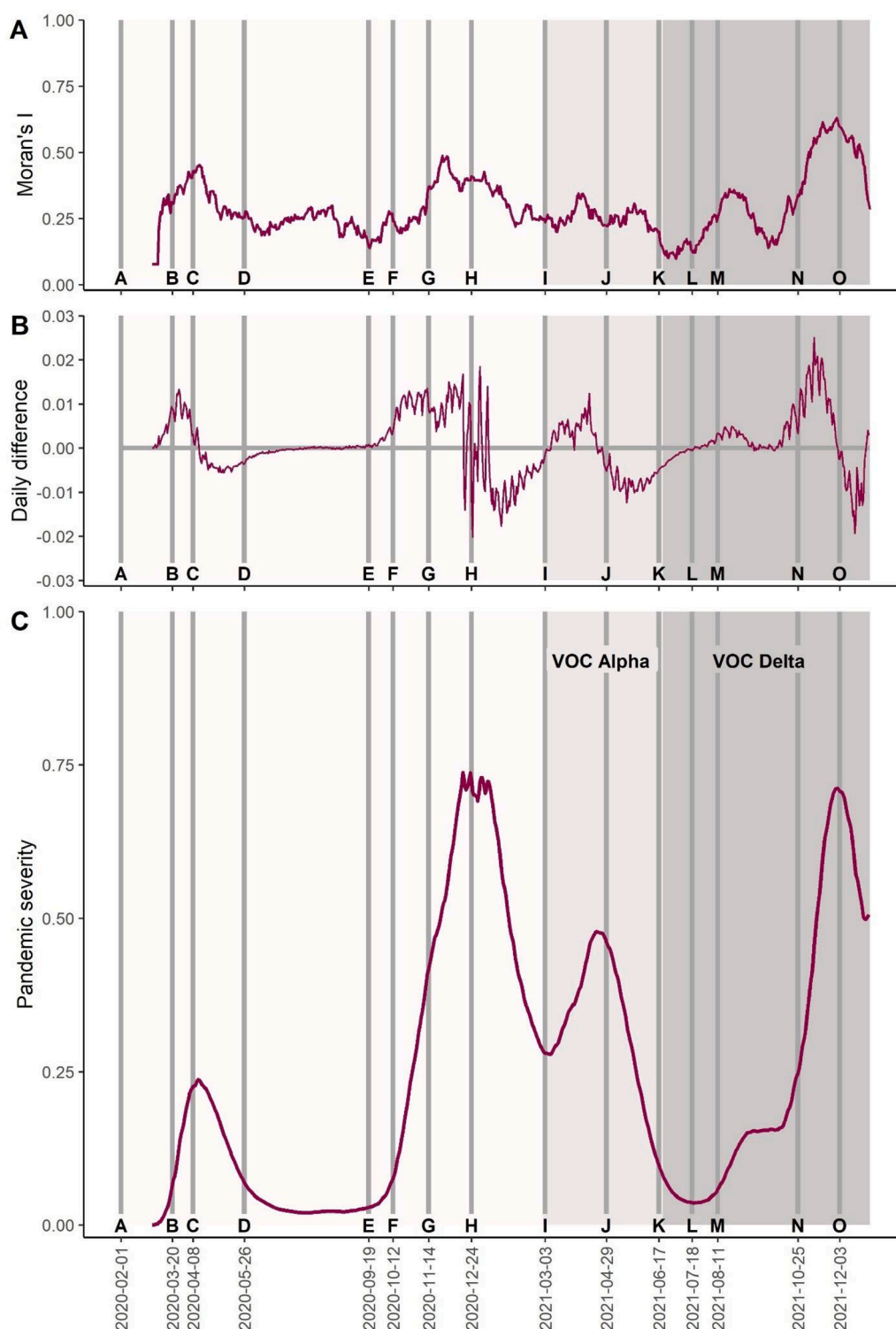
A comprehensible method for defining pandemic stages is rarely found in the literature on COVID-19. Existing phase models (Ghosh and Cartone, 2020; Benita and Gasca-Sanchez 2021; Li et al., 2021; Zawbaa et al., 2022) define the beginning of each stage relatively arbitrarily based on individual indicators, such as incidence rates, mortality rates, or the implementation of countermeasures like lockdowns and social distancing. Typical types of phases include ‘beginning’, ‘outbreak’,

<sup>2</sup> Theoretically, the range is between 0 and 1, in practice between 0 and 0.7625 because the minimum-maximum normalization was performed on the sub-indicators

‘recession’, and ‘plateau’ (Li et al., 2021). Schilling et al. (2022) used a multivariate approach by combining several variables for their phase model for Germany, although their method of delineating phases remains unclear. Küchenhoff et al. (2021) calculated change points in the early course of the pandemic from March to May 2020 using a back-projection for estimated daily infections in Germany. Their retrospective exploratory analysis identified five phases for Germany within the first wave. In contrast to existing approaches, our phase model of the

COVID-19 pandemic in Germany is based on a composite indicator and spans a longer period of several waves, during which our phases are trend-coherent to serve as a heuristic for further analytical steps.

We performed a change point analysis on a time series of the pandemic severity index on the national scale. To establish a phase model on the national scale the pandemic severity indicator was aggregated daily to the national level. To establish trend-coherent phases, we first calculated the lagged difference, which is the



**Fig. 1.** Pandemic severity in Germany and the fifteen phases of our phase model during the years 2020 and 2021 (A). The phase model is based on the change points of the daily difference charts (B). The spatial autocorrelation within Germany reaches the highest levels around the peaks of the four waves (C).

difference between the value for pandemic severity for each day and the value for the day before. Then, a change point analysis using the binary segment approach (Scott and Knott 1974) was performed on the lagged difference of the pandemic severity indicator to establish the phases. A minimum segment length of 14 days was used in the calculation. The change points were calculated using the *changepoint* package in R (Killick and Eckley 2014). The resulting phase model consists of 15 individual phases, ranging from 19 to 116 days in length.

### 3.3. Local and global autocorrelation

First, we used the global autocorrelation of pandemic severity, which was calculated for each day in the study period in form of the Moran's I metric (Moran 1950). Moran's I and the related test statistic were calculated by using the *sfdep* package in R (Parry 2022). The global Moran's I statistic describes the spatial autocorrelation over the study area with values ranging from  $-1$  to  $+1$ . Positive values indicate the clustering of similar values. Second, each of the fifteen phases (Fig. 1) was analyzed in terms of local spatial autocorrelation. Local indicators of spatial association (LISA) were calculated for each phase using localized Moran's I using the *sfdep* package in R (Parry 2022). LISA is a tool to detect geographical clustering of high and low values in a dataset by calculating local spatial autocorrelation at each county and indicating how similar an observation is to others nearby (Anselin 1995; Sokal et al., 1998). The contiguity matrix for the LISA analysis was established based on the  $k$ -nearest neighbor criterium with  $k = 6$ , which is the average number of neighboring counties in our dataset, to reflect the irregular configuration of territorial borders (Ghosh and Cartone 2020), since many German counties have just one neighbor. A simulation of 1499 iterations resulted in a pseudo  $p$ -value for the results of the LISA analysis, which were capped at  $p = 0.05$ . The LISA method has been used before to identify significant clusters of pandemic outbreaks (e.g., Scarpone et al., 2020; Siljander et al., 2022; Ghosh and Cartone 2020). However, to our knowledge, this study is the first to use this approach based on a composite indicator of pandemic severity instead of case incidence or mortality as a single indicator.

The results of a LISA analysis grouped the spatial units (here German counties) relative to their neighbors in local clusters of high values (high-high/ HH) or low values (low-low/ LL), and identified spatial outliers with (high-low/ HL) or (low-high/ LH) values. If a region was classified as HH or LL it was part of a statistically significant cluster of high or low values, respectively. Since the geography of the COVID-19 pandemic is uneven in space (Scarpone et al., 2020; Rodríguez-Pose and Burlina 2021), the pandemic severity index can also help to reveal the complex spatio-temporal patterns of how the pandemic unfolded better compared to case numbers. Supplementary Materials 1 contains the equations for Moran's I and LISA. Visualizations of LISA results were created in R using the *ggplot2* (Wickham 2016), and *sf* (Pebsma 2018) packages.

### 3.4. Hierarchical clustering of lisa sequences

To identify temporal patterns of spatial association, we further analyzed the results of the LISA analysis through hierarchical clustering. Hierarchical clustering is usually performed in a two-step procedure. First, a dissimilarity matrix between all  $n$  observations of the input dataset (e.g., regions) is calculated using a distance metric. Second, the hierarchical clustering algorithm sorts the observations according to their distance to all other observations. We followed a principle used by Bucci et al. (2022) and Mattera (2022), both of which used spatio-temporal dissimilarity of time series data between regions as input for hierarchical clustering, analog to a sequence analysis. We deviated from this by using the results of the LISA analysis for each county during each of the fifteen phases to calculate the dissimilarity matrix, which was used for the hierarchical clustering. Thus, the similarity of German counties is established in terms of how well their LISA

values are aligned. Two regions with the identical combination of LISA values over the fifteen phases are considered as most similar. To establish the dissimilarity matrix, we used Gower's distance (Gower 1971) for the dissimilarity matrix. Only counties, which were classified into one of the four LISA categories during four or more of the fifteen phases ( $n = 142$ ) to avoid a sparse dissimilarity matrix. Then hierarchical clustering was performed based on the average linkage method. Both steps were performed using the *cluster* package in R (Maechler et al., 2022). The resulting dendrogram was capped at the height of 1.0 to receive six relatively coherent clusters. The formula for Gower's distance can be found in Supplementary Materials 1.

## 4. Results

### 4.1. Phase model and global spatial autocorrelation

We developed a pandemic phase model for the first two years of the pandemic using change point analysis of the pandemic severity index. Our model consists of fifteen phases, each defined by trend coherence and varying in length (Fig. 1). On average, each phase lasts about 47 days, with the longest phase lasting 116 days (summer plateau 2020) and the shortest lasting only 19 days (surge of the first wave). Pandemic waves consisted of at least two phases (increase and decrease), but more complex waves can have several intermediate phases of acceleration and deceleration. For example, the second COVID-19 wave in Germany consisted of four individual phases. The periods between phases are also considered individual phases. Of the fifteen phases during the study period, seven were characterized by increasing pandemic severity, four by decreasing severity, and four were stable. This fine-grained phase model provides a heuristic for gaining more specific insights into spatio-temporal patterns.

Moran's I was positive throughout the study period, indicating a spatial association of similar values rather than a spatial association of high and low values. The value of Moran's I ranged from 0.076 in March 2020 to 0.629 during the peak of the fourth wave (Fig. 1A), indicating the presence of spatial structure and non-random distribution of pandemic severity. There were notable variations in the indicator, with five peaks in weeks 2020-w16 (0.452), 2020-w49 (0.484), 2021-w14 (0.342), 2021-w33 (0.361), and 2021-w47 (0.629). Interestingly, these peaks roughly correlated with peaks in pandemic severity (Chart 1 A and D), while periods with low pandemic severity tend to have lower overall spatial autocorrelation. This pattern suggested that during the four waves of high pandemic severity, areas with similar pandemic severity (e.g., hotspots or cold spots) were located close to each other. We further analyzed specific patterns of local spatial autocorrelation.

### 4.2. Local spatial association during the different phases

The LISA analysis compared the level of pandemic severity in each county with that in its neighboring counties. As a result, each county is associated with high-high (HH), low-low (LL), low-high (LH), or high-low (HL) values for each phase and a level of significance for this value (Fig. 2). It should be noted that pandemic severity varied substantially between phases. For example, an HH value during a phase of high pandemic severity (such as phases B, G, or I) would indicate a dire situation with many cases, deaths, and high pressure on the local healthcare system, while an HH value during a phase with modest pandemic severity would indicate a less tense situation.

The resulting maps of LISA values show that the pattern of local spatial autocorrelation varied substantially throughout the study period, although subsequent phases usually have similar patterns. During 2020 (phases A-H), the overall picture changed several times, while during 2021 the situation was more stable. Most phases were dominated by one or more areas with several counties of HH and LL values, while HL and especially LH values were much rarer. Interestingly, the two largest counties, Berlin and Hamburg, both featured HL values relatively often,



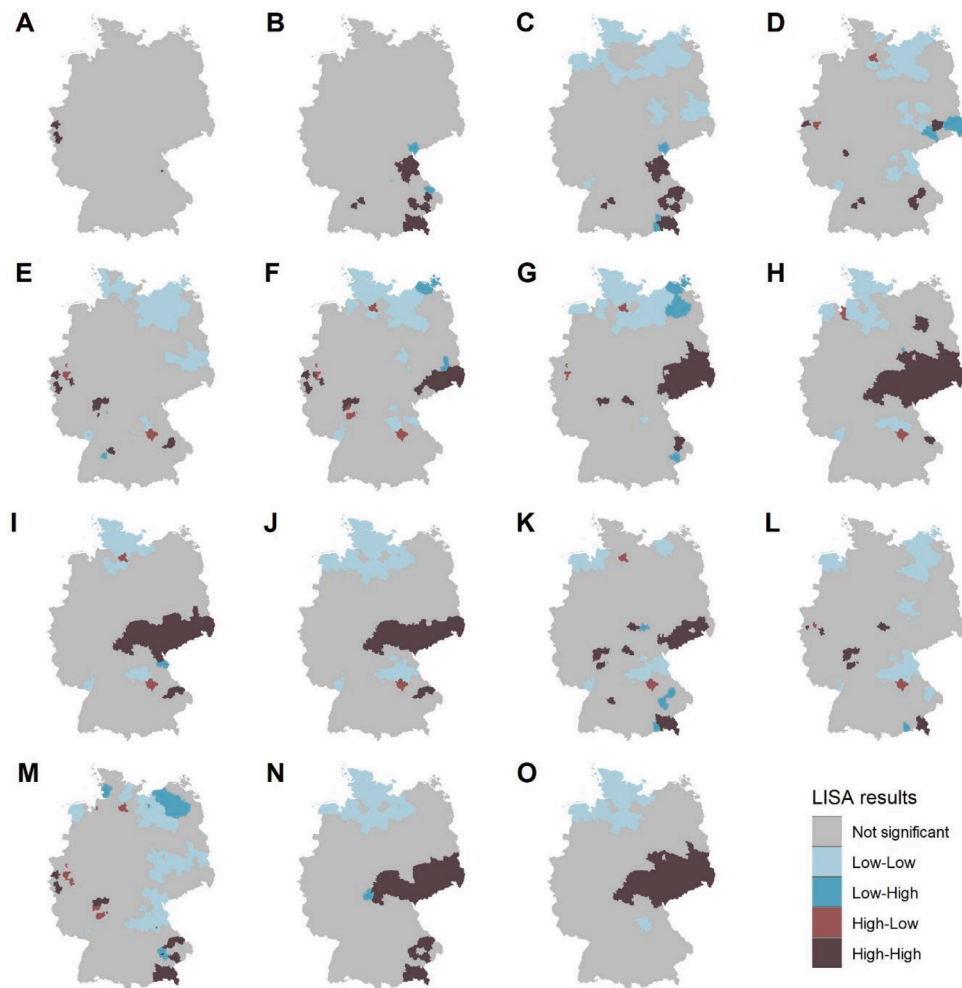


Fig. 2. Local indicators of spatial autocorrelation of pandemic severity for the fifteen phases of the pandemic during 2020 and 2021 (A–O).

indicating that they were somewhat decoupled from the development around them.

While the overall pattern showed some variation, some regions had relatively stable values over time. LL values were mostly concentrated in Northern Germany for much of the study period. On the other hand, a concentration of HH values was found in east-central Germany during phases F–K and N–O, corresponding to a repetitive pattern during waves two, three, and four of COVID-19 in Germany (Fig. 2). This area, which includes much of Saxony and Thuringia and parts of Saxony-Anhalt and Brandenburg, covers virtually all of the counties in Germany that were hardest hit by the pandemic overall. The LISA analysis results suggest the existence of stable patterns of high and low pandemic severity, respectively, some of which appear to be relatively stable in place.

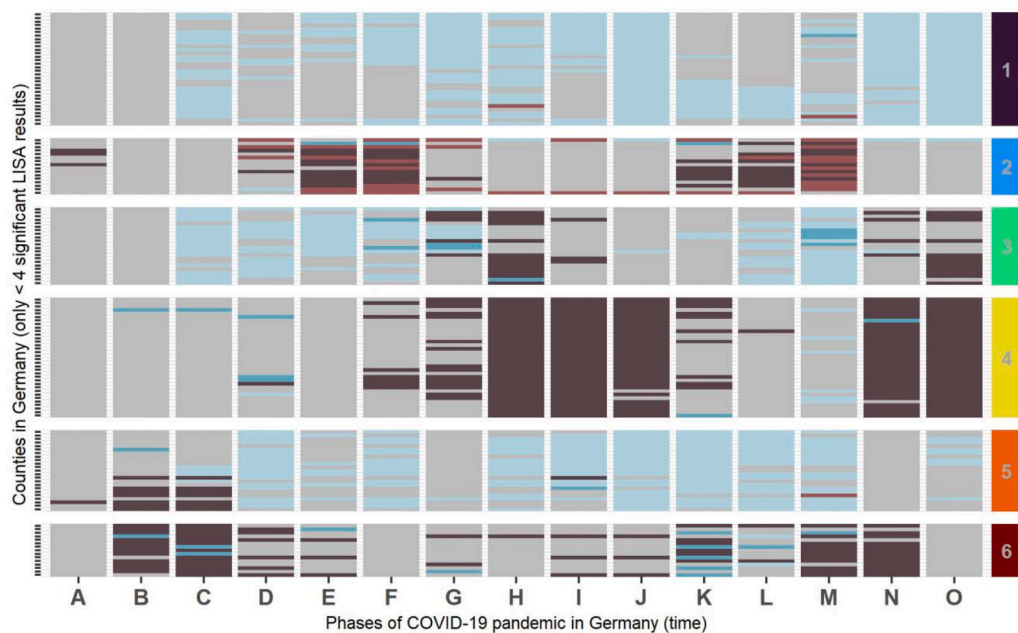
Significant values from the LISA analysis were somewhat unevenly distributed in space. Only a quarter of the possible results (fifteen phases in 400 counties) were significant, and about half of the 400 counties had significant values for two or fewer phases. Additionally, 52% of all significant results were concentrated in 25% of the counties. The proportion of significant results varied substantially among phases, with 13 out of 400 counties having received a significant result in phase A and 150 out of 400 having received a significant result in phase M. We further explored the spatio-temporal trajectories of counties by performing hierarchical clustering on the LISA analysis results to identify patterns of similar trajectories of pandemic severity.

#### 4.3. Types of regional trajectories

The results of hierarchical clustering suggest that there are several relatively stable patterns of trajectories of pandemic severity among German counties. Six types of county trajectories were identified (Fig. 3), characterized by similar temporal patterns of pandemic severity. Four types were characterized by relatively persistent values of HH (types 4, 2, and 6) or LL (type 1), while the other two types were characterized by combinations of HH and LL (types 3 and 5). Interestingly, each type also had a relatively stable spatial pattern and was dominant in specific regions of Germany (Fig. 3B).

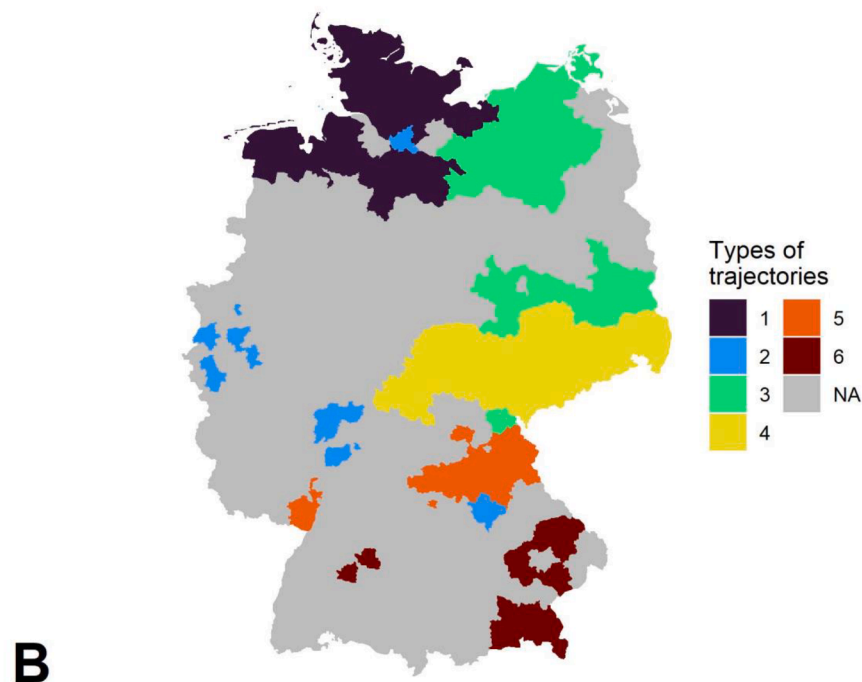
The largest type in terms of included counties was type 4, consisting of 34 counties that were classified as HH during the second (phases E–H), third (phases I–J), and fourth waves (phases M–O) of COVID-19 in Germany. However, none of the counties in this group were classified as HH during the first wave (phases A–C), which had a spatial pattern unlike the subsequent waves. Geographically, this type included all counties in Saxony and some in neighboring Thuringia and Saxony-Anhalt. This area included metropolitan areas around Dresden and Leipzig as well as rural areas.

In contrast, the second-largest type, type 1, was characterized by 32 counties that were classified as LL for most of the study period. This cluster of low pandemic severity persisted throughout all four waves of COVID-19 in 2020 and 2021 and was only absent during the first two phases. Geographically, this type was present in northwestern Germany and included large parts of Schleswig-Holstein and Lower Saxony, as well as one county in Mecklenburg-West Pomerania. While this area is



**Fig. 3.** Six types of spatio-temporal trajectories of pandemic severity among German counties during the years 2020–2021. Each county with more than four significant LISA results represents a row in Fig. 3A and is color coded by the LISA value for each of the fifteen phases. Thus, counties sharing the same colors during a phase of the pandemic (vertical; A to O) has a similar trajectory. Hierarchical clustering revealed six types of spatio-temporal trajectories. The counties in chart 3B are color coded by the type of spatio-temporal trajectory.

**A** LISA results counties Low-Low High-Low Low-High High-High



predominantly rural, it also includes parts of the Hamburg metropolitan region. Another somewhat similar type was type 3, dominated by LL values during much of 2020 and in the summer of 2021. However, most counties of this type had HH values during the peak of the second wave and during the fourth wave in late 2021. This type was present in 22 counties in northeastern Germany and south of Berlin in predominantly rural areas in Mecklenburg-West Pomerania, Brandenburg, and Saxony-Anhalt.

Type 5 was somewhat similar, characterized by counties that were hotspots during the first wave in early 2020 but were classified as LL for subsequent phases. This type was present in northern Bavaria and parts of Rhineland-Palatinate. These 23 counties are mostly rural areas and

small towns but also included the Nuremberg metropolitan region. Type 5 was also characterized by counties that showed HH values during the first wave. However, unlike the previous type, these counties also had HH values during the fourth wave in late 2021, although the phases in between were largely insignificant. Geographically, this type was found in southeast Bavaria and two counties in Baden-Württemberg. The 15 counties of this type are mostly rural but relatively dense areas at the edge of the Munich and Stuttgart metropolitan regions.

Finally, type 2 arguably deviated most both in terms of the temporal pattern of pandemic severity and geographically. While all other types were either patterns of HH or LL values or a combination of both, the 15 counties of type two featured dominantly HL values, indicating that they

deviated from their immediate surroundings. Interestingly, this type was almost an inverse of type 4, with HL values concentrated not during overall peaks of pandemic severity but right in between, during periods of lower pandemic severity. Type 2 differed geographically from other types as it was the most dispersed of all six types. It included mostly urban and sub-urban counties in the Rhine-Ruhr metropolitan region around Düsseldorf and in the Rhine-Main metropolitan region around Frankfurt, as well as the city of Hamburg. Generally, temporal patterns of pandemic severity align with spatial patterns in so far as the spatial patterns repeatedly occurred over several waves. In the following section, these results are further discussed.

## 5. Discussion

In this paper, we offered a comprehensive analysis of the first two years of the COVID-19 epidemic in Germany from a spatio-temporal perspective. To do so, we proceeded in four analytical steps: First, a novel composite index of pandemic severity was developed that enabled us to differentiate the dynamics of the pandemic temporally and spatially. Second, a phase model with fifteen temporally coherent phases was established through change point analysis to serve as a heuristic for spatial analysis. Third, measures of global and local autocorrelation were performed to determine regional clusters of pandemic severity for German counties.

Regions with spatial concentrations of hotspots and cold spots respectively were identified via LISA analysis. Fourth, six types of similar regional trajectories of pandemic severity were identified through hierarchical clustering among German counties. The six types of trajectories offer insights into the spatio-temporal dynamics of COVID-19 in Germany during the years 2020 and 2021. In this section, we discuss our findings by presenting empirical and methodological contributions before concluding.

Empirically, three observations are especially noteworthy from our result. First, our analysis shows that the spatial patterns of pandemic severity of COVID-19 varied substantially between the different phases of the pandemic. However, a certain pattern is visible, in that the first wave of COVID-19 in Germany (spring 2020) displayed a different pattern than the subsequent three waves, while waves two, three, and four were much more similar to each other. This aligns with previous findings that the first wave was mostly driven by relocation diffusion through tourist returnees (Kuebart and Stabler 2020). The three subsequent waves emerged from a higher level of pandemic severity, which might somewhat explain the different spatial patterns. Second, it is remarkable that both hotspots and cold spots remained almost stationary from late 2020 onwards during the three most severe waves of COVID-19 in Germany. The persistent hotspot areas are almost exclusively located within an area in central Germany, including parts of the states of Saxony, Thuringia, and Saxony-Anhalt. The persistent cold spot areas on the other hand are mostly located in Northern Germany, in the states of Schleswig-Holstein, Lower Saxony, and Mecklenburg-West Pomerania. Neither hotspots nor cold spots followed patterns that would suggest obvious structural differences in terms of urbanity versus rurality or population density. While the high level of pandemic severity in Saxony has been noted before (Chilla et al., 2022), it is still remarkable how stable this area returned as an area of hotspots of pandemic severity in each of three subsequent waves. Third, it is also noteworthy from a spatio-temporal perspective that the epicenter of each wave was within this region, although waves three and four have been driven by the novel alpha and delta variants of SARS-CoV-2, respectively so that relocation diffusion should have been an important factor. Indeed, the first hotspots of variant alpha transmission were found to be in western and northern Germany (Mitze and Rode 2022), far from the region in central Germany that was the region with the highest concentration of pandemic severity due to variant alpha only six weeks later.

Taken together, these three empirical findings imply that the COVID-19 pandemic in Germany progressed wavelike just in temporal terms,

but not spatially. This is in line with the findings of D'Angelo et al. (2021), who found the case evolution over time in Italian regions to be relatively independent of each other. As in our case, this implies that region-specific aspects trump expansion diffusion dynamics, although this does not seem to be a universal phenomenon, since Kim et al. (2021) find wavelike patterns in space for the case of South Korea and Boterman (2022) finds no consistent patterns for the case of the Netherlands. The consistent regional patterns of pandemic severity identified in this study contrast these findings for the German context and contribute a long-term perspective to the debate. Therefore, we conclude that for the German case with its strict non-pharmaceutical interventions and largely successful vaccination campaign during the year 2021, pandemic severity is indeed less related to relocation diffusion. Instead, regional factors such as cultural or political aspects determined the effectiveness of counter-pandemic measures and thus the local pandemic diffusion seems to be more important. Further, not just the specific regional conditions but also the timing in the pandemic process should be considered when analyzing the spatio-temporal dynamics of infectious diseases.

Methodologically, also three conclusions can be drawn from the approach presented in this paper. First, the spatio-temporal clustering of time series of pandemic severity is a valuable tool to explore the dynamics of infectious outbreaks. While this approach has been successfully applied before (Bucci et al., 2022; Mattera 2022), we combined this approach with LISA analysis (Anselin 1995). In our opinion, this has the advantage that only those regions that deviate significantly from the national average are considered (e.g., hotspots and cold spots of pandemic severity), which allows focusing on the most relevant processes or regions for each phase of the pandemic. Second, a phase model based on change point analysis is a valuable heuristic to further analyze spatio-temporal dynamics of COVID-19. The choice of which points in time or periods to compare is crucial for spatio-temporal analysis. Most studies either use relatively arbitrary units such as weekly or bi-weekly (Siljander et al., 2022) intervals. In contrast, we decided to develop a phase model that distinguished different levels of pandemic severity, which proved valuable to contextualize spatial differences. Third, the use of a composite indicator as the base for subsequent analytical steps is useful to add robustness to the analysis. To follow the calls for combining different indicators when analyzing the spread of COVID-19 (Pagel and Yates 2021; Rohleder and Bozorgmehr, 2021), we developed an index of pandemic severity that incorporated the incidence of patients tested positive for COVID-19, the incidence of patients with COVID-19 on ICU, and the incidence of registered deaths due to COVID-19. Although the timespan analyzed in this paper included only 21 months, the conditions under which data were collected changed drastically in several regards. Factors that varied over time included the testing regime, the chain of reporting itself, new variants of COVID-19, and increasing levels of immunity within the population due to infections and vaccinations. Therefore, we argue that combining indicators enhances comparability over longer timespans and adds robustness to the analysis. However, this is certainly not limited to the sub-indicators used here. Other indicators could and should be included in future attempts, for example, data originating from wastewater monitoring.

Some limitations of the research presented here should be considered. First, we did not include demographic factors in our pandemic severity index. Spatial variation in factors such as age distribution or prevalence of chronic diseases could influence the spatial distribution of deaths by COVID-19 and would thus somewhat influence the spatial pattern of pandemic severity per our index. Relatedly, data that describes the conditions of local infections would be helpful to contextualize local pandemic severity. Especially regional data on outbreaks and imported infections would be helpful but were not made available by the German authorities. Second, the use of data that is only reported for territorial administrative areas presented a limitation, since it might obscure patterns on a smaller scale, for example in border regions



(Scarpone et al., 2020; Chilla et al., 2022). On the other hand, an analysis of spatial patterns on an even larger scale (e.g., NUTS 2 regions) might be suitable to reduce noise for the hierarchical clustering. Third, we did not consider spatial variation in vaccinations and non-pharmaceutical interventions. Vaccinations became an important factor with the third wave in spring 2021, while non-pharmaceutical interventions were relatively evenly distributed during most phases. However, there might be regional compliance gaps that are hard to grasp with data (Stabler and Kuebart 2023). In conclusion, both more fine-grained analysis of hotspot regions and analysis on a higher spatial order can be helpful steps to gain a better understanding of the dynamics of the pandemic.

## 6. Conclusion

In conclusion, the spatio-temporal dynamics of COVID-19 offered fascinating insights into how pathogens spread in contemporary societies. While it would have been misleading to put too much emphasis on the territorial dimension of space in the era of "post-Westphalian pathogens" (Fidler 2003), the "territorial immune system" of non-pharmaceutical interventions visible during the COVID-19 pandemic proved the relevance of territories and their impact on the pandemic process. Further, the differences in regional trajectory patterns found here might implicate regional anomalies in compliance to public health control measures that are hard to predict. Therefore, public health policies should consider the possibility of unforeseen spatial variations in effectiveness of interventions.

## Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Andreas Kuebart reports that this research has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 492338717.

## Data availability

Data has been made available in repository, please see 10.5281/zenodo.8004580

## Acknowledgements

We acknowledge the support of Malte Wünsche for data management, as well as the support of Anne Kuebart (Heinrich-Heine-Universität Düsseldorf) for commenting on the manuscript. We thank to anonymous reviewers for their efforts in reviewing the manuscript.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.sste.2023.100605](https://doi.org/10.1016/j.sste.2023.100605).

## References

- Anselin, L., 1995. Local Indicators of Spatial Association—LISA. *Geographical Anal.* 27 (2), 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>.
- Benita, F., Gasca-Sanchez, F., 2021. The main factors influencing COVID-19 spread and deaths in Mexico: a comparison between phases I and II. *Appl. Geogr.* 134, 102523. <https://doi.org/10.1016/j.apgeog.2021.102523>.
- Boterman, W., 2022. Population density and SARS-CoV-2 pandemic: comparing the geography of different waves in the Netherlands. *Urban Stud.*, 004209802210871. <https://doi.org/10.1177/00420980221087165>.
- Boudou, M., Khandelwal, S., ÖhAiseadha, C., Garvey, P., O'Dwyer, J., Hynds, P., 2023. Spatio-temporal evolution of COVID-19 in the Republic of Ireland and the Greater Dublin Area (March to November 2020): a space-time cluster frequency approach. *Spat. Spatio Temporal Epidemiol.* 45 <https://doi.org/10.1016/j.sste.2023.100565>.
- Bucci, A., Ippoliti, L., Valentini, P., Fontanella, S., 2022. Clustering spatio-temporal series of confirmed COVID-19 deaths in Europe. *Spat. Stat.* 49, 100543. <https://doi.org/10.1016/j.sspata.2021.100543>.
- Chilla, T., Große, T., Hippe, S., Walker, B.B., 2022. COVID-19 incidence in border regions: spatiotemporal patterns and border control measures. *Public Health* 202, 80–83. <https://doi.org/10.1016/j.puhe.2021.11.006>.
- Cliff, A.D., Smallman-Raynor, M., Haggett, P., Stroup, D., Thacker, S., 2009. *Infectious Diseases: A Geographical Analysis: Emergence and Re-emergence*. Oxford University Press, Oxford.
- Coccia, M., 2022. COVID-19 pandemic over 2020 (with lockdowns) and 2021 (with vaccinations): similar effects for seasonality and environmental factors. *Environ. Res.* 208 <https://doi.org/10.1016/j.envres.2022.112711>.
- Coccia, M., 2021. The impact of first and second wave of the COVID-19 pandemic in society: comparative analysis to support control measures to cope with negative effects of future infectious diseases. *Environ. Res.* 197 <https://doi.org/10.1016/j.envres.2021.111099>.
- D'angelo, N., Abbruzzo, A., Adelfio, G., 2021. Spatio-temporal spread pattern of covid-19 in Italy. *Mathematics* 9 (19). <https://doi.org/10.3390/math9192454>.
- Dialga, I., Thi Hang Giang, L., 2017. Highlighting methodological limitations in the steps of composite indicators construction. *Soc. Indic. Res.* 131 (2), 441–465. <https://doi.org/10.1007/s11205-016-1263-z>.
- Fidler, D.P., 2003. SARS: political Pathology of the first post-westphalian pathogen. *J. Law Med. Ethics* 31, 485–505.
- Ghosh, P., Cartone, A., 2020. A Spatio-temporal analysis of COVID-19 outbreak in Italy. *Reg. Sci. Policy Pract.* 12 (6), 1047–1062. <https://doi.org/10.1111/rsp.3.12376>.
- Gower, J.C., 1971. A general coefficient of similarity and some of its properties. *Biometrics* 857–871.
- Keeler, C., Emch, M., 2018. *Infectious-disease geography*. Routledge Handbook of Health Geography. Routledge, Abingdon, pp. 45–51.
- Killick, R., Eckley, I.A., 2014. {changepoint}: an R package for changepoint analysis. *J. Stat. Softw.* 58 (3), 1–19.
- Kim, S., Kim, M., Lee, S., Lee, Y.J., 2021. Discovering spatiotemporal patterns of COVID-19 pandemic in South Korea. *Sci. Rep.* 11 (1) <https://doi.org/10.1038/s41598-021-03487-2>.
- Küchenhoff, H., Günther, F., Höhle, M., Bender, A., 2021. Analysis of the early COVID-19 epidemic curve in Germany by regression models with change points. *Epidemiol. Infect.* <https://doi.org/10.1017/S0950268821000558>.
- Kuebart, A., Stabler, M., 2020. Infectious diseases as socio-spatial processes: the COVID-19 outbreak in Germany. *Tijdschr. Econ. Soc. Geogr.* 111 (3), 482–496. <https://doi.org/10.1111/tesg.12429>.
- Kuebart, A., Stabler, M., 2023. Pandemic severity composite indicator for COVID-19 in Germany dataset. Zenodo (2022). <https://doi.org/10.5281/zenodo.8004580>.
- Lambio, C., Schmitz, T., Elson, R., Butler, J., Roth, A., Feller, S., Savaskan, N., Lakes, T., 2023. Exploring the Spatial Relative Risk of COVID-19 in Berlin-Neukölln. *Int. J. Environ. Res. Public Health* 20 (10). <https://doi.org/10.3390/ijerph20105830>.
- Li, M., Guo, X., Wang, X., 2021. Retrospective prediction of the epidemic trend of COVID-19 in Wuhan at four phases. *J. Med. Virol.* 93 (4), 2493–2498. <https://doi.org/10.1002/jmv.26781>.
- Libório, M.P., Ekel, P.Y., de Abreu, J.F., Laudaes, S., 2022. Factors that most expose countries to COVID-19: a composite indicators-based approach. *GeoJournal* 87 (6), 5435–5449. <https://doi.org/10.1007/s10708-021-10557-5>.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K., 2022. *cluster: Cluster Analysis Basics and Extensions*. R package, 2.1.4. R package.
- Mattera, R., 2022. A weighted approach for spatio-temporal clustering of COVID-19 spread in Italy. *Spat. Spatio-Temporal Epidemiol.* 41 <https://doi.org/10.1016/J.SSTE.2022.100500>.
- Mitze, T., Kosfeld, R., 2022. The propagation effect of commuting to work in the spatial transmission of COVID-19. *J. Geogr. Syst.* 24 (1), 5–31. <https://doi.org/10.1007/s10109-021-00349-3>.
- Mitze, T., Rode, J., 2022. Early-stage spatial disease surveillance of novel SARS-CoV-2 variants of concern in Germany with crowdsourced data. *Sci. Rep.* 12 (1) <https://doi.org/10.1038/s41598-021-04573-1>.
- Moran, P., 1950. Notes on continuous stochastic phenomena. *Biometrika* 37 (1–2), 17–23. <https://doi.org/10.1093/biomet/37.1.2.17>.
- Murgante, B., Borruso, G., Balletto, G., Castiglia, P., Dettori, M., 2020. Why Italy first? health, geographical and planning aspects of the COVID-19 outbreak. *Sustainability* 12 (12). <https://doi.org/10.3390/su12125064> (Switzerland).
- Nazia, N., Butt, Z.A., Bedard, M.L., Tang, W.C., Sehara, H., Law, J., 2022. Methods used in the spatial and spatiotemporal analysis of COVID-19 epidemiology: a systematic review. *Int. J. Environ. Res. Public Health* 19 (14), 8267. <https://doi.org/10.3390/ijerph19148267>.
- Pagel, C., Yates, C.A., 2021. Tackling the pandemic with (biased) data. *Science* 374 (6566), 403–404. <https://doi.org/10.1126/science.abi6602>.
- Parry, J., 2022. sfdep: spatial dependence for simple features. R package version 0.2.3.
- Pebesma, E., 2018. Simple features for R: standardized support for spatial vector data. *R J.* 10 (1), 439–446. <https://doi.org/10.32614/RJ-2018-009>.
- Rodríguez-Pose, A., Burlina, C., 2021. Institutions and the uneven geography of the first wave of the COVID-19 pandemic. *J. Reg. Sci.* 61 (4), 728–752. <https://doi.org/10.1111/jors.12541>.
- Rohleder, S., Bozorgmehr, K., 2021. Monitoring the spatiotemporal epidemiology of Covid-19 incidence and mortality: a small-area analysis in Germany. *Spat. Spatio Temporal Epidemiol.* 38 <https://doi.org/10.1016/j.sste.2021.100433>.
- Scarpone, C., Brinkmann, S.T., Große, T., Sonnenwald, D., Fuchs, M., Walker, B.B., 2020. A multimethod approach for county-scale geospatial analysis of emerging infectious diseases: a cross-sectional case study of COVID-19 incidence in Germany. *Int. J. Health Geogr.* 19 (1) <https://doi.org/10.1186/s12942-020-00225-1>.

- Schilling, J., Buda, S., Tolksdorf, K., 2022. Zweite Aktualisierung der „Retrospektiven Phaseneinteilung der COVID-19- Pandemie in Deutschland. *Epidemiologisches Bull.* 10, 3–5. <https://doi.org/10.25646/9787>.
- Schmitz, T., Lakes, T., Manafa, G., Lambio, C., Butler, J., Roth, A., Savaskan, N., 2023. Exploration of the COVID-19 pandemic at the neighborhood level in an intra-urban setting. *Front. Public Health* 11. <https://doi.org/10.3389/fpubh.2023.1128452>.
- Scott, A.J., Knott, M., 1974. A cluster analysis method for grouping means in the analysis of variance. *Biometrics* 30 (3), 507. <https://doi.org/10.2307/2529204>.
- Siljander, M., Uusitalo, R., Pellikka, P., Isosomppi, S., Vapalahti, O., 2022. Spatiotemporal clustering patterns and sociodemographic determinants of COVID-19 (SARS-CoV-2) infections in Helsinki. *Finl. Spat. Spatio Temporal Epidemiol.* 41 <https://doi.org/10.1016/j.sste.2022.100493>.
- Sokal, R.R., Oden, N.L., Thomson, B.A., 1998. Local spatial autocorrelation in a biological model. *Geogr. Anal.* 30 (4), 331–354. <https://doi.org/10.1111/j.1538-4632.1998.tb00406.x>.
- Stabler, M., Kuebart, A., 2023. Tempo-spatial dynamics of COVID-19 in Germany – A phase model based on a pandemic severity indicator. *medRxiv*. <https://doi.org/10.1101/2023.02.17.23286084>.
- Teller, J., 2021. Urban density and Covid-19: towards an adaptive approach. *Build. Cities* 2 (1), 150–165. <https://doi.org/10.5334/bc.89>.
- Wickham, H., 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.
- Wickham, H., Seidel, D., 2022. *scales: scale functions for visualization*. R package version 1.2.1.
- Zawbaa, H.M., Osama, H., El-Gendy, A., Saeed, H., Harb, H.S., Madney, Y.M., Abdelrahim, M.E.A., 2022. Effect of mutation and vaccination on spread, severity, and mortality of COVID-19 disease. *J. Med. Virol.* 94 (1), 197–204. <https://doi.org/10.1002/jmv.27293>.
- Zeileis, A., Grothendieck, G., 2005. Zoo: S3 infrastructure for regular and irregular time series. *J. Stat. Softw.* 14 (6), 1–27. <https://doi.org/10.18637/jss.v014.i06>.