Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Fine-scale variation in the effect of national border on COVID-19 spread: A case study of the Saxon-Czech border region

Adam Mertel a,b,* , Jiří Vyskočil a,b , Lennart Schüler a,b,c , Weronika Schlechte-Welnicz a,b , Justin M. Calabrese a,b,d,e

a Center for Advanced Systems Understanding (CASUS), D-02826 Göttingen, Germany
b Helmholtz-Zentrum Dresden Rossendorf (HZDR), D-01328 Dresden, Germany
c Department of Computational Hydrosystems, Helmholtz Centre for Environmental Research (UFZ), D-04318 Leipzig, Germany
d Department of Ecological Modelling, Helmholtz Centre for Environmental Research (UFZ), D-04318 Leipzig, Germany
e Department of Biology, University of Maryland, 20742-4415 College Park, MD, USA

A R T I C L E   I N F O

Keywords:
COVID-19
Spatial epidemiological modeling
Regression model
Border effect
Spatio-temporal data analysis

A B S T R A C T

The global extent and temporally asynchronous pattern of COVID-19 spread have repeatedly highlighted the role of international borders in the fight against the pandemic. Additionally, the deluge of high resolution, spatially referenced epidemiological data generated by the pandemic provides new opportunities to study disease transmission at heretofore inaccessible scales. Existing studies of cross-border infection fluxes, for both COVID-19 and other diseases, have largely focused on characterizing overall border effects. Here, we couple fine-scale incidence data with localized regression models to quantify spatial variation in the inhibitory effect of an international border. We take as a case study the border region between the German state of Saxony and the neighboring regions in northwestern Czechia, where municipality-level COVID-19 incidence data are available on both sides of the border. Consistent with past studies, we find an overall inhibitory effect of the border, but with a clear asymmetry, where the inhibitory effect is stronger from Saxony to Czechia than vice versa. Furthermore, we identify marked spatial variation along the border in the degree to which disease spread was inhibited. In particular, the area around Löbau in Saxony appears to have been a hotspot for cross-border disease transmission. The ability to identify infection flux hotspots along international borders may help to tailor monitoring programs and response measures to more effectively limit disease spread.

1. Introduction

The COVID-19 pandemic is an event of unprecedented magnitude in modern world history and has consequently set off a flurry of research activity across the health-related sciences. These efforts have generated an equally unprecedented deluge of spatially referenced epidemiological data. High resolution epidemiological data have the potential to reveal disease transmission processes on heretofore inaccessible spatial scales, but significant challenges in data harmonization and modeling remain. The global extent and temporally asynchronous nature of the COVID-19 spread have repeatedly put the spotlight on international borders and the policies governing them. Research focusing on cross-border infection fluxes can thus deepen our understanding of COVID-19 spread dynamics while also serving to inform border-related policy decisions.

Several papers have studied COVID-19 border dynamics for different regions with a range of methodologies. For example, Grimée et al. (2021) used the Endemic-Epidemic framework (Bekker-Nielsen Dunbar and Held, 2020) to model the effect of policies on Swiss-Italian borders. Eckardt et al. (2020) applied a Bayesian spatio-temporal Poisson model to study the role of border controls in the Schengen Area, and Hossain et al. (2020) used a metapopulation model for estimating the effect of travel restrictions on the COVID-19 outbreak. Additionally, Laroze et al. (2021) constructed a multi-source spatial model on top of the pre-crisis commute to work to study the effects of borders between regions within the same country, and Han et al. (2021) explored the balance between domestic and imported cases in Chinese cities. While all of these studies suggested a significant effect of border presence and border control regime on the spread of COVID-19, all focused on determining the overall effect of the focal borders.

* Corresponding author.
E-mail address: mertel.adam@gmail.com (A. Mertel).

https://doi.org/10.1016/j.sste.2022.100560
Received 25 March 2022; Received in revised form 14 September 2022; Accepted 7 December 2022
Available online 11 December 2022
1877-5845/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).
Most of the border-focused studies mentioned in the previous paragraph are based on the spatial level of whole countries, regions, or spatially separated metropolitan areas. More generally, most spatio-temporal COVID-19 models are applied either on the level of whole countries or on medium-sized provinces and regions due to data limitations. In contrast, we developed a model of COVID-19 spread on the very detailed level of individual municipalities. While several studies consider this level of granularity for data analysis and modeling (e.g., Araujo-Card, 2021; Cole et al., 2020; Neyens et al., 2020; Schuler et al., 2021), research combining datasets on this scale from more than one country is still exceedingly rare even it may provide an unique value for studying the spatiotemporal trends in detail. Furthermore, while governing policy may be uniform across an entire border, variation in geomorphology, population distribution, transportation infrastructure, socio-economic aspects, history and other factors along a border may generate spatial variation in the inhibitory effect of the border on disease spread. To our knowledge, this type of fine-scale variation has not been studied in the context of COVID-19 spread. We therefore aim to quantify cross-border infection fluxes at the scale of individual municipalities in the Saxony-Czech border region.

Our approach is based on the simple assumption that the virus travels mainly with humans. Therefore, COVID-19 may be transmitted more easily between two localities with a higher intensity of mobility connecting them than between localities with lower mobility. Several approaches are widely used in epidemiology to estimate mobility. Probably the most common is the gravity model, which has many possible adaptations and extensions (Yan and Zhou, 2019) and has been a mainstay of spatial COVID-19 modeling (e.g., Chen et al., 2021; Ezzat et al., 2021; Werner PA, 2021; Zhu et al., 2021).

The basic gravity model estimates the mobility between two places primarily from two variables - their separation distance and the population size of both places. The gravity model can be further extended to include border effects. This idea was applied in the context of various topics, e.g., international to domestic trade (McCallum, 1995), the cross-border flows of immigration (Lewer and Van den Berg, 2008), air traffic (Becker et al., 2018), or mergers and acquisitions (Wong, 2008).

Even more relevant to this study is the paper of Kramer et al. (2016), in which the authors extended the gravity model with a border effect to study the international transmission of Ebola virus. This study concluded that the presence of the national border and its closure may have played a significant role in the transmission of the virus.

Here, we propose a set of local multiple regressions -inspired by the principles of the gravity model-that allow us to quantify fine-scale spatial variation in the effect of the border. Consistent with other studies, we find that the Saxon-Czech border has an overall inhibitory effect on COVID-19 spread. However, we also identify both between-country asymmetry and substantial fine-scale variation in the strength of the border effect. In particular, the border in the region around Löbau in Saxony appears to have had a weaker inhibitory effect than other areas along the Saxony-Czech border. We conclude by discussing how high resolution border studies can potentially pinpoint hotspots for disease spread where additional border monitoring and controls may be warranted.

1.1. Data

This section describes the data sources used in this study and the first preprocessing to clean, integrate, and precalculate the Czech and German data. For Germany, we work with the dataset "Gemeindegrenzen 2018 mit Einwohnerzahlen" (GeoBasis-DE et al., 2020) for the geometries and population sizes on the municipality ("Gemeinde") level. Czech population numbers on the municipality level ("obec") were taken from the Czech Statistical Office (Czech Statistical Office, 2021), while the geometries were obtained from Registry of territorial identification, addresses and real estate (Czech Office for Surveying et al., 2021). We further joined the population size numbers with their associated geometries and merged the datasets from both countries. To keep the same geometry detail on both sides of the borders, we applied the Douglas-Peucker (Douglas and Peucker, 1973) simplification algorithm implemented in the Python library TopoJSON (https://github.com/mattijn/topojson).

Due to the geographical focus of this study, the number of infections in single municipalities was obtained only for Saxony on the German side. These data were published on the official website of Saxonian government coronavirus.sachsen.de (Sächsische Staatsregierung, 2021). These numbers are taken from the reports of doctors and laboratories that are responsible for PCR tests, the assignment to the municipalities is done based on the information from patients. Unfortunately, the website does not provide any historical data directly in a tabular format or via an API service, just a table with data from the last seven days. We therefore set up a CRON job to visit the website every day and scrape the contents in the form of a .csv table. On the Czech side, we worked with the daily updated dataset from the Czech Ministry of Health (Komenda et al., 2021), which provide municipality-level values localised based on the usual presence of the patients. We also reclassified both datasets to weekly granularity to avoid problems with weekly cyclicity (i.e., pronounced decreases in cases reported on weekends, national holidays, etc.). For this study, on the Czech side, we used only data for the three regions in Czechia that share a border with Saxony - Liberec, Ústí nad Labem, and Karlovy Vary. The entire region defined for the case study is shown in Fig. 1. The preprocessed and cleaned dataset is available on the Zenodo server ((Mertel, 2022): 2).

After data cleaning and harmonization, we obtained a dataset of 1116 municipalities (almost 6 million inhabitants), with 415 located in Saxony (4,077,937 inhabitants), and 701 in Czechia (1,810,283 inhabitants). The municipalities in Czechia are smaller on average, which reflects a slightly different administrative organization. The timeframe of the case study ranges from the 31st of January to the 13th of June 2021. This timeframe was selected due to the availability of the data on both sides of the border, but also to capture the wave of high case numbers in late winter and spring 2021. In Fig. 2, we can see the weekly case numbers in the study area in each municipality. This series of maps shows the municipalities with increased cases each week of the time series. Here we can follow the general spatio-temporal trends of the spread. At the very beginning of the timeframe, the situation was considerably worse in the western part of Czechia, particularly around the towns of Karlovy Vary and Cheb. During February, the wave of high numbers came to all Czech regions, while from March, the places with increases in case numbers are more dispersed and mostly located within the regions around Děčín, Ústí nad Labem, and Chomutov. On the German side, the first more significant and spatially concentrated wave of increased cases appeared in the middle of March in the region of western Saxony, around Plauen and Zwickau. The following week, we can see an increase also in north-western Saxony (Leipzig region), followed by the central (Dresden, Chemnitz) and then eastern (Görlitz, Bautzen, Zittau)regions of Saxony. Fig. 3 shows these values grouped based on the administrative regionalization on a timeline. This graph show a substantial difference between the trends of the German and Czech regions. While the numbers in the Czech regions tend to rise at the beginning of the timeframe (February), German regions have a wave of high numbers in the middle of the timeframe (late March, April). Even though the trends within each country are similar, considerable differences also exist among neighboring regions. Although the Czech regions are considerably smaller in population, the number of cases is similar, implying a higher incidence on the Czech side.

Most of the above-cited spatial epidemiology papers use the Euclidean distance between location centroids as the measure of distance. However, this method is problematic on the granular spatial scale of our study. Instead, the connectivity between municipalities is mostly defined by the infrastructure that is determined by social (e.g., culture, work/school commuting, local subsidies) and physical (e.g., mountainous areas, rivers) factors. To account for variation in these structural
factors, we estimate “temporal distances” between pairs of municipalities via the typical car-based driving time between them.

To calculate the temporal distances between municipalities, we firstly needed to extract the point representing the real central place of each municipality. This was done by the search engine Nominatim (https://nominatim.openstreetmap.org, © OpenStreetMap contributors), which offers a free API service to geocode places by their names. For every request, it returns the point coordinates that represent the geographical unit. For municipalities, the API returns mostly the point of the central square or the centroid of the built-up areas. This point was then used as a starting point for the calculation of isochrones. For this step, we used the Openroute API (https://api.openrouteservice.org, © openrouteservice.org by HeiGIT | Map data © OpenStreetMap contributors) that returned the areas accessible by car within a specific time limit. We set the upper value to 1 hour and worked with 10 six minute intervals. Then, the temporal distance was calculated by iterating through all pairs of municipalities and checking into which isochrone interval of the first municipality the centroid of the second municipality falls. This measure of distance between two municipalities accounts for the transportation networks connecting them and is thus more informative for our purposes than the simple Euclidean distance.

We performed several exploratory analyses before implementing our regression models. Here, we quantified the overall patterns in the temporal trends, the spatial co-occurrence of the case number increases, and the effect of the spatial and temporal autocorrelation. We conclude that the pandemic situation in each country developed somewhat independently, with indications of a significant but heterogeneous spatial effect of the national border. Case numbers in municipalities within a short driving distance (10 minutes and less) are much more strongly correlated than those separated by longer distances. Additionally, municipalities’ case numbers are, in general, more correlated at shorter time lags, with the lowest variance in correlations occurring at time lags of 1 to 2 weeks.

1.2. Model

To quantify the local effect of the national border on the spread of COVID-19, we constructed a local model, which was further applied to each municipality in the vicinity of the national border. Inspired by the gravity model, each local model quantifies the contributions of neighboring municipalities to an index of “virus import potential” for the focal municipality as a function of population sizes, inter-municipality temporal distances, and presence/absence of the border between each pair of municipalities.

1.2.1. Time gap

In the whole model, we work with the assumption that the high number of cases in municipality $b$ in time $t - x$ may predict the number of cases in the neighboring municipality $a$ in time $t$. The $x$ is then the time lag that defines the time that the virus needs to be imported between municipalities. We chose the $x$ to be one week based on the common understanding of the virus spread and the time-lagged correlation analysis, which can be seen in Fig. 4. While the median correlation values change little across various time lags, we note a difference in the distribution of values within the higher part of the scale (75th percentile). This refers to the situations when not all municipalities have similar trends in case numbers, but the correlation is present much more often when the temporal lag of zero or one week and the municipalities

Fig. 1. Overview map shows the location of the case study area and the extent of two selected bordering regions - the federal state Saxony in Germany, and the regions of Liberec, Ústí nad Labem and Karlovy Vary in Czechia. Map tiles by Stamen Design, under CC BY 3.0. Data by OpenStreetMap, under ODbL.
lie within one country. Maximum values of the correlation for each municipality show a significant decrease in the time lag of one week, while all the other time lags retain similar distributions. More interesting is the comparison of the variance in the neighborhood municipality-municipality correlations. In this case, the time lags of one and two weeks have the lowest numbers; slightly higher numbers are reached when we consider the municipalities on the other side of the border. This analysis may indicate overall spatio-temporal trends of the virus spread. While the median values do not change significantly over time, the distributions of the medians and the values of the maximum correlations reveal that the similarity within the neighborhood is higher when there is no time lag. In contrast, the variance analysis revealed the
highest consistency in correlation values when the time lag is one or two weeks. Additionally, correlations are generally higher when we consider only municipalities from one country, which again highlights the effect of the border.

1.2.2. Neighborhood

The computation of the model starts by iterating through the list of all municipalities, and for each focal municipality $a$, we define a neighborhood of municipalities $(b, c, ..., f)$, which are reachable from $a$ in less than forty minutes of driving time in a car. Further, this neighborhood of municipality $a$ is noted as $\Theta_a$. This value is supported by additional travel data for this geographical region (Bundesagentur für Arbeit, 2020; Centrum dopravního výzkumu, 2019; ifas Institut für angewandte and Sozialwissenschaft GmbH, 2017). These sources show that the largest distance people usually travel to work, school, or shopping is not more than 20 kilometers / 30 minutes. But there are types of travel that often occur over longer distances and times, such as tourism, family visits, and business-related journeys that may take even more than one hour / 100 kilometers. We therefore chose forty minutes as a conservative value that would comfortably accommodate most normal travel. An analysis of the effect of different travel distances to the correlation of new case values is displayed in Fig. 5. Here we notice a strong similarity of case numbers across municipalities within the same country. Such correlations tend to be significant when the municipalities are within 20 minutes of each other, implying strong spatial autocorrelation. A similar relationship of autocorrelation in case numbers at short temporal distances is not present for pairs of municipalities separated by the national border. This finding suggests a strong effect of the national border on the spread of the virus.

1.2.3. Virus import potential

The 'virus import potential' is a number that represents the potential that a virus is imported from a neighborhood municipalities $\Theta_a$. While the spatial covariances may differ in various periods of the pandemic (Schuler et al., 2021), we focused on the weeks when the number of cases in the municipality was increasing because. For each of the increasing weeks $t$, we looked at the situation in the neighborhood to identify the places, where the increase may have been spread from. We looked at the number of cases within the municipalities in the $\Theta_a$ at $t-1$, one week before the increase in $a$. Then, we relativized the numbers for

---

**Fig. 4.** Correlation values of the municipality-municipality temporal trends considering the neighborhood defined by one hour of driving distance. The dark blue color stands for all combinations of municipalities in one country, the light blue color for all municipality-municipality combinations.

**Fig. 5.** Spatial autocorrelation graph. The dependence of the municipality-municipality correlations in the case numbers on the temporal distance. The chart highlights the similarity between two municipalities (quantified by the correlation number) on their temporal distance.
\( \Theta a \) to the sum of all cases in the neighborhood, so we ended with virus import potential values scaled to \([0, 1]\). Formally noted, the import potential from \( b \) to \( a \), in time \( t \), \( Iab(t) \), is:

\[
Iab(t) = \frac{Nb(t - 1)}{\sum_{i \in \Theta a} N_i(t - 1)}
\]

where \( Nb(t) \) is the number of cases in \( b \) in time \( t \).

We calculate the import potential from \( a \) to \( b \) over the whole study period, \( Iab \), as the average of all weeks, when the number of cases in \( a \) increased. While not every lagged co-variate in cases increase may indicate a high potential of import, we use weights to incorporate principles of virus transmission dynamics. This can be written as:

\[
Iab = \frac{\sum_{i=1}^{m} wa(t)^*Iab(t)}{\sum_{i=1}^{m} wa(t)}
\]

Where \( np \) is the number of weeks, when the case numbers in \( a \) are increasing. The weight of import potential for municipality \( a \) in time \( t \), \( wa(t) \) has three components:

\[
wa(t) = w\Theta a(t)^* wR a(t)^* w\Delta a(t)
\]

(i) The first weight, neighborhood effect in municipality \( a \) in week \( t \), \( w\Theta a(t) \), compares the epidemiological situation in the municipality \( a \) to the situation in \( \Theta a \). We assume that a high relative number of cases in the \( \Theta a \) compared to the \( a \) may indicate a higher potential of the import from these municipalities. The \( w\Theta a(t) \) is calculated as:

\[
w\Theta a(t) = 1 - \frac{(3\Delta a + Z(t) / 6 \Theta a) - (3\Delta a + Z(t) / 6 \Theta a)}{2}
\]

where \( Z(t) \) is the z-score of the case number in \( a \) in week \( t \). \( n \) indicates the sum of cases in \( \Theta a \), \( Z(t) \) is then the z-score of this value in week \( t \), and \( ea/\Theta a \) is the standard deviation of the case values in \( a \), respectively \( \Theta a \). The resulting values will be in between 0 and 2. Value 2 indicates that the import potential in week \( t \) is high, value 0 represents the situation that \( a \) is, compared to \( \Theta a \), completely saturated and there is a low possibility of import.

(ii) \( wR a(t) \) is the second weight used for the import potential calculation and represents the impact of instantaneous reproduction number \( Ra(t) \) of the municipality \( a \) in week \( t \). Including this weight, we incorporate the factor of local transmission dynamics in \( a \) itself. We assume that if \( Ra(t) \) is high, then the dynamics in municipality \( a \) may explain the increase more probably that the import from \( \Theta a \). \( Ra(t) \) was estimated using the epyestim Python package (Cori et al., 2013). For the calculation, we interpolated daily values of the municipalities case numbers, and used the window of 14 days to ensure a smooth result in the smaller municipalities. \( wRa(t) \) can be then easily derived from \( Ra(t) \) as:

\[
wRa(t) = 1 / Ra(t)
\]

(iii) The third weight is the coefficient of the increase magnitude \( w\Delta a(t) \). For any week \( t \), this value is calculated as the increase in week \( t \) divided by a mean case increase in municipality \( a \) for the study period, \( \langle \Delta a(t) \rangle \).

\[
w\Delta a(t) = \Delta a(t) / \langle \Delta a(t) \rangle
\]

1.2.4. Regression model

For each municipality \( a \), we constructed a local beta regression model that is based on the characteristics of the relations of \( a \) to the municipalities in the neighborhood \( \Theta a \). As a dependent value for the model, we calculated the possibility of importing the virus from municipalities in \( \Theta a \) to \( a \) so that they could be used as the dependent variable in a beta (multiple) regression model. The linear predictor of the mean of our beta regression model, which is linked to the \([0, 1]\) index of virus import via the logit link function, can be written as

\[
Iab = A + wCa * Cab + wPa * Pb + wBa * Bab,
\]

where \( Iab \) stands for the virus import potential from municipality \( b \) to municipality \( a \), and \( A \) is an intercept. \( Ca, Pa \), and \( Pa \) are the three predictors of the model, and \( wCa, wB, \) and \( wP \) are the regression coefficients for the municipality \( a \), which we try to estimate with this model. In this study, we assume that both the predictors \( Ca, Pa, Pa \), and the regression coefficients \( wC, wB, \) and \( wP \) of a single municipality are constant over the time span.

\( C \) stands for closeness, which was calculated as the inverse value of the relative driving distance. The higher the closeness value, the faster it was possible to reach the focal municipality from the second municipality by car. The \( Cab \) can be calculated as:

\[
Cab = Dab / D\Theta a
\]

where \( Dab \) is the temporal distance between \( a \) and \( b \), \( D\Theta a \) is the temporal distance that we assigned for defining the neighborhood, which is for 40 minutes in this study.

\( P \) is the factor of the municipality’s population size, that compares the population in \( b \) to the sum of populations in \( \Theta a \). The calculation of this predictor can be written as:

\[
Pab = Pb / \sum_{i \in \Theta a} Pi
\]

\( B \) is the third predictor variable that represents the presence of the border. Here, the \( Bab \) value is 0 if both \( a \) and \( b \) are in the same country, value of 1 then \( a \). For example, when \( a \) was in Saxony, we assigned 0 to all the Saxonian municipalities and 1 to all Czech municipalities, and vice versa when \( a \) was in Czechia.

For each municipality \( a \), we calculated a model consisting of the dependent variable and all three predictor variables for each municipality in the neighborhood \( \Theta a \) for the mean of the response, and a separate model consisting of just the border term for the precision of the response. We first standardized predictors of relative municipality size and closeness variables via z-scores, and then estimated the beta models via maximum likelihood. These analyses were implemented in R (R Core Team, 2000) and the betareg package (Zeileis et al., 2016). For a visual guide of the whole procedure, please refer to Fig. 6.

The spatial pattern of \( wB \) is of particular interest to us. Specifically, spatial variation in this coefficient allows us to quantify and visualize spatial variation in the border effect. We then modeled the precision parameter of the beta regression as a linear function (identity link) of border presence, \( wBa * Bab \), which improved both model fit and convergence relative to a constant precision model.

2. Results

The model coefficients (\( wD, wB, \) and \( wP \)) are estimated based on the three predictor variables for each municipality, from which the \( wB \), the effect of the border, was the focal point of our study. The model summary is displayed in Fig. 7, a more detailed picture of spatial distribution of the border effect weight is displayed on map in Fig. 8. This map shows that within most of the municipalities in our study area, the border strongly inhibits virus spread (pink to magenta colors), but with pronounced the spatial variation in the strength of this inhibitory effect. In particular, the national border appears to have a less inhibitory effect (greener colors) in the area around the town of Lůbau in eastern Saxony.

In total, 365 local models were constructed, one for each municipality that fulfilled the condition to have at least five municipalities in the neighborhood that lay on the other side of the border. Only three of those models did not converge, and was therefore removed from the analysis. Across all the remaining 362 local models, the median \( r^2 \) value is \( \sim 0.42 \). Our methodology was not capable of identifying statistically significant results (\( p < 0.05 \)) mostly for the municipalities located on the
outskirts of the study area, which may have been influenced by regions not considered in this study. From the summary numbers in Fig. 7, we can conclude that the population size has the highest impact on the import potential. This result was expected, while compared to other studies, we did not use relativized incidence numbers, mainly because of the high diversity of municipalities, including very small ones. On the other hand, the closeness predictor has only a small impact on the model results, and the p-value is lower than 0.05 only in the case of 81 municipalities out of 362. While the analysis in Fig. 5 suggested the impact of the autocorrelation on the temporal distances between municipalities,
we assume that the regions close to the national border may have different dynamics. Results of the border predictor highlight marked asymmetry on opposite sides of the national border, with a higher potential of virus import from Czechia to Saxony than vice versa. This may be explained by the overall spatio-temporal trend of the spread over the timeframe of our study, in which the Northern Czech Republic (and specifically the province of Cheb) was reported as a European hotspot at the very beginning of the year 2021, followed by high incidence values in Saxony in the following months (March, April). Another explanation may be the effect of a different quality of COVID-19 tracking, healthcare quality, and different internal regulations in both countries.

The border region near the towns of Löbau and Bautzen in eastern Saxony stands out as a hotspot where increases in focal municipalities are more strongly linked to case number increases on the Czech side of the border. There are also single municipalities in the other parts of both Saxony and Czechia that show weaker inhibition (e.g., Marienberg, Johanngeorgenstadt, Neuhausen, and Dubí located in the central part of the study area). Outside of these exceptions, the overall effect of the border on disease transmission is mixed or strongly inhibitory.

To validate the results of the model, we did two post-hoc analyses. The first check is based on a comparison of the average of all weeks when cases in the focal municipality increased weighted by the magnitudes of the increases (referred to as central increase week further in the study). This number represents the differential timing of the waves in the two countries as an alternative way to look at the degree of cross-border coupling. The spatial distribution of the central increase week is shown on the left map of Fig. 9. In this picture, we can identify the spatial discontinuity created by the national border. This effect is least pronounced at two places - around the town of Šluknov (including the German neighborhoods of Löbau and Sebnitz) where the COVID-19 wave came later (during mid-March) compared to other Czech regions, and the westernmost part of Saxony, where the case numbers

Fig. 8. Map of municipalities in the neighborhood of the Saxon-Czech border. The circle size represents the population, and the color depicts the estimated \( w_B \), which represents the potential of the border to inhibit the spread of COVID-19. The green color stands for a less inhibitory effect of the border, while magenta tones represent municipalities where the border strongly inhibited disease spread. The width of the circle stroke represents the p-value of the border parameter of the municipality. The location of the municipalities Löbau and Šluknov is highlighted. Map tiles by Stamen Design, under CC BY 3.0. Data by Open StreetMap, under ODbL.

Fig. 9. Weighted central increase week (left) and the neighborhood variance in central increase week (right).
peaked relatively early compared to the rest of Saxony. The right map then shows the neighborhood variance (within 40 minutes of driving distance) of this value. Because the central increase week values form, in general, two spatial clusters - one for each country, the places with high values of the variance are then concentrated around the national border. One exception to this trend is located around the towns of Löbau (Germany) and Sluknov (Czechia), which again indicates a less inhibitory effect of the national border on the virus spread in this area. This result is consistent with the outcome of the beta regression analysis, which shows the highest potential for cross-border virus import exactly in this area.

The second post-analysis to validate the results of the model is based on the idea of the Granger causality metric (Granger, 1969; Romero García et al., 2021), which tests the causation of one time series by another time series considering a time lag. For each municipality \( m \), we constructed two microregions. The first one consisted of the municipalities falling into a driving distance limit of 20 minutes (half of the threshold used for the model) from \( m \) and the same country as \( m \), and the second consisted of the municipalities in the driving distance of 40 minutes from \( m \) but located on the other side of the border. We summed weekly case numbers within both microregions and calculated the change. These values were further used to perform the Granger causality test. Fig. 10 shows the map of Granger causality p-values for the model of border effect weight (Fig. 8). This map shows the Löbau as one of the hotspots of a proved granger causality, which is in accordance with the results of the model.

3. Discussion

In this paper, we proposed a regression-based methodology to estimate spatial variation in the effect of a national border on disease spread. This calculation does not directly measure the real import of the virus. Instead, it is a reasonable approximation of the possibility to explain an increase in the municipality \( a \) with the number of cases in municipality \( b \), one week before this increase. Our approach leverages fine-scale, spatially-referenced case count data to estimate a location-specific index of virus import potential. We then demonstrated the power of our technique on a municipality-level case study of COVID-19 spread in the Saxon-Czech border region. Consistent with other studies of border effects, we found an overall inhibitory effect of the national border on disease spread between the two regions. Further, the border inhibition is stronger in the direction of Saxony to Czechia than vice versa, which may have been a result of the different epidemiological dynamics in the two countries. While the incidence in Czech municipalities peaked earlier than in Saxony, a cross-border commuter from Czechia would be more likely to get infected in Saxony and carry it over to Saxony than the other way around. Importantly, our approach also highlighted pronounced spatial variation in the inhibitory effect of the border, particularly in the region around Löbau in Saxony. Specifically, the border provided much weaker inhibition of disease spread in this region from Czechia to Saxony compared to most other areas along the border. This hotspot was clearly identified by our local regression models, and this central result was also verified by post-hoc analysis focusing on differences in the timing of the outbreaks on either side of the border.

In the hotspot around Löbau, the timing of the localized case count peak was much closer to that of the neighboring region in Czechia (neighborhood of town Sluknov) than it was to other accessible municipalities in Saxony. This finding suggests a pronounced, cross-border coupling of epidemic dynamics in this region, which our regression models clearly identified. In contrast, the central region of the border displayed a stronger inhibitory effect, with the disease dynamics between neighboring regions of Saxony and Czechia more clearly decoupled.

We suspect that these localized differences in border effect, particularly those between the eastern (Löbau) and central border regions, are driven by differential cross-border traffic patterns. The restrictions for the travel between two countries during the study period were regularly updated, mainly in the Czechia - Saxony direction. But during the whole time period, travel was highly restricted with the number of exceptions valid for crossing in one or both directions. First of all, commuters traveling for work from the Czech Republic to Saxony, on a daily or weekly basis, are essential to several sectors of the Saxonian economy, including healthcare, tourism, transport, and manufacturing (Sujata et al., 2020). Some of these workers had exceptions even during the hardest lockdown, and some were pushed to minimize border crossings (Sachsen.de 2021a; Sachsen.de 2021b). The state authorities announced mandatory tests in January 2021 for all cross-border workers (Sachsen.de 2021c), which increased the case numbers on the German side. On the other hand, people from Saxony used to go shopping in Czechia regularly due to the lower prices of some commodities (e.g., gas, groceries, and some services), and the closure of the border would sharply reduce such traffic. Physical conditions also modulate cross-border mobility. For example, the Ore Mountains (Erzgebirge / Krušné Hory) along the central part of the Saxon-Czech border form a natural barrier for many cross-border activities. The distribution of cities and areas of higher population density between sides of the borders can also have an effect. Specifically, the central region of the border features larger cities on the Czech side than on the German side. When coupled with the mountainous terrain along the border and sparser road networks, this difference in population density may reduce border crossings both for Czechs (less need to commute for work) and for Germans (less convenient to commute for commodities and leisure). In contrast, the opposite factors would likely contribute to increased border traffic in the region around Löbau. While our results are consistent with this hypothesis, explicit data on cross-border movements, which were unavailable to us, would be required to confirm this interpretation fully.

A limitation of our study is the lack of direct consideration of some of the surrounding regions, which may have influenced COVID-19 dynamics near the Saxon-Czech border, notably Poland to the far east and the German states of Bavaria and Thuringia to the west. The unavailability of municipality-scale data from these regions precluded their consideration in our analysis. This issue manifests itself in the lower pseudo \( R^2 \) values for the models and higher p-values on the border term that tend to occur on the outer edges of our study area. This is specifically relevant for the regions of Görlitz in the east and Cheb / Plauen in
the west. Outside of these outlying regions, we are confident that our analysis accurately quantifies disease import potential from surrounding regions and provides reliable results. This robustness is further evidenced by the agreement between our core model-based analysis and the confirmatory analyses based on timing differences and on Granger causality.

Our index of virus import is based on a comparison of the case increase in a focal municipality with the relative number of cases in the neighboring municipalities one week earlier. Calculation of this index therefore requires determining appropriate values for neighborhood size, time lag, and relative weightings of contributions with the neighborhood. While our selection of these parameters was guided by literature review and knowledge gained via exploratory analysis, we tested the robustness of our results by various driving distance thresholds, temporal lags, and weighting techniques. Similarly, we considered a range of options for dealing with exact 0’s and 1’s in our index of virus import. Collectively, these alternative modeling assumptions do slightly affect our numerical results; however, these variations changed neither our qualitative results nor the inferences we drew from our analysis.

We have developed a method for quantifying fine-scale variation in the effect of a border on between-country (or between-region) disease transmission potential. A key advantage of our approach is that it does not require data on cross-border movements of individuals, which are generally difficult to obtain and come with substantial privacy concerns. Furthermore, our approach is able to detect potential hotspots for cross-border disease transmission, which might then be flagged for additional monitoring efforts or stricter border measures in the future. While this type of analysis is retrospective, it could potentially be automated such that, as new data become available, it provides near real-time results on patterns of cross-border disease transmission. Such an automated system could help authorities decrease the time needed to adjust border policies to changing epidemic conditions. A related issue concerns determining the extent to which such cross-border transmission hot spots are consistent over time. The multiple waves of COVID-19 that have occurred worldwide should provide ample opportunities to address this and related questions in future studies.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

This work was partially funded by the Center of Advanced Systems Understanding (CASUS), which is financed by Germany’s Federal Ministry of Education and Research (BMBF) and by the Saxony Ministry for Science, Culture, and Tourism (SMWK) with tax funds on the basis of the budget approved by the Saxony State Parliament.

References

Araujo-Card, J.M., 2021. A first insight about spatial dimension of COVID-19: analysis at municipality level. J. Public Health 43 (1), 98–106. https://doi.org/10.1093/pubmed/fdb146.
Romero García, C., Iftimi, A., Briz-Redón, A., et al., 2021. Trends in incidence and transmission patterns of COVID-19 in Valencia, Spain. JAMA Netw. Open 4 (6), e2113818. https://doi.org/10.1001/jamanetworkopen.2021.13818.

Sachsen.de, 2021a. Ausweisung der Übernachtungszuschüsse für tschechische Pendler aller Branchen. Available at: https://www.medienservice.sachsen.de/medien/news/247147 (accessed 20 February 2022).

Sachsen.de, 2021b. Sachsen passt Quarantäne-Verordnung an: Ausnahmen für Berufs-Pendler aus Mutationsgebieten festgelegt. Available at: https://www.medienservice.sachsen.de/medien/news/247142 (accessed 20 February 2022).

Sachsen.de, 2021c. Testpflicht für Grenzpendler ab 18. Januar 2021. Available at: https://www.medienservice.sachsen.de/medien/news/245443 (accessed 20 February 2022).

Sächsische Staatsregierung, 2021. Infektionsfälle in Sachsen - sachsen.de. Available at: https://www.coronavirus.sachsen.de/infektionsfaelle-in-sachsen-4151.html (accessed 18 December 2021).

Schuler L., Calabrese J.M. and Attinger S., 2021. Data driven high resolution modeling and spatial analyses of the COVID-19 pandemic in Germany. medRxiv: 21.

Sujata, U., Weyh, A., Zillmann, M., 2020. Kurzstudie zur Bedeutung von Grenzpendlern für den sächsischen Arbeitsmarkt. IAB-Regional Sachsen.

Werner PA, 2021. Tracing and modeling of the COVID-19 pandemic infections in Poland using spatial interactions models. Computational Science and Its Applications – ICCSA 2021. Springer International Publishing, Cham, pp. 641–657. https://doi.org/10.1007/978-3-030-86979-3_45. O Gervasi, B Murgante, S Misra2021Lecture Notes in Computer Science.

Wong, W.K., 2008. Comparing the fit of the gravity model for different cross-border flows. Econ. Lett. 99 (3), 474–477. https://doi.org/10.1016/j.econlet.2007.09.018.

Yan, X.Y., Zhou, T., 2019. Destination choice game: a spatial interaction theory on human mobility. Sci. Rep. 9 (1), 9466. https://doi.org/10.1038/s41598-019-46026-w.

Zeileis, A., Cribari-Neto, F., Gruen, B., et al., 2016. Package ‘betareg’. R package 3 (2).

Zhu, D., Ye, X., Manson, S., 2021. Revealing the spatial shifting pattern of COVID-19 pandemic in the United States. Sci. Rep. 11 (1), 8396. https://doi.org/10.1038/s41598-021-87902-8.