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.

Transportation Research Part C 129 (2021) 103231

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Influence of transportation network on transmission heterogeneity of COVID-19 in China

Jing Lu [a], Anrong Lin [a], Changmin Jiang [b,*], Anming Zhang [c], Zhongzhen Yang [d]

[a] College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
[b] Asper School of Business, University of Manitoba, Winnipeg, MB R3T 2N2, Canada
[c] Sauder School of Business, University of British Columbia, Vancouver, BC V6T1Z2, Canada
[d] Faculty of Maritime and Transportation, Ningbo University, Ningbo 315211, China

ARTICLE INFO

Keywords:
Transportation network
COVID-19
Spatial heterogeneity
China
Deep neural network
Geographical weighted regression

ABSTRACT

In this paper, we propose a novel approach to model spatial heterogeneity for epidemic spreading, which combines the relevance of transport proximity in human movement and the excellent estimation accuracy of deep neural network. We apply this model to investigate the effects of various transportation networks on the heterogeneous propagation of COVID-19 in China. We further apply it to predict the development of COVID-19 in China in two scenarios, i.e., i) assuming that different types of traffic restriction policies are conducted and ii) assuming that the epicenter of the COVID-19 outbreak is in Beijing, so as to illustrate the potential usage of the model in generating various policy insights to help the containment of the further spread of COVID-19. We find that the most effective way to prevent the coronavirus from spreading quickly and extensively is to control the routes linked to the epicenter at the beginning of the pandemic. But if the virus has been widely spread, setting restrictions on hub cities would be much more efficient than imposing the same travel ban across the whole country. We also show that a comprehensive consideration of the epicenter location is necessary for disease control.

1. Introduction

Coronavirus disease 2019 (COVID-19) was first detected in Wuhan, Hubei province of China on December 8th, 2019 (Wuhan Municipal Health Commission, 2020). Due to the highly contagious and pathogenic nature of the disease (D'Amico et al., 2020), it is believed that many people in Wuhan became infected in the next 45 days. Before the lockdown of Wuhan on January 23rd, 2020, an estimated 5 million people had left the city, mainly to visit family for Chinese Lunar New Year (Hubei Provincial People's Government, 2020), thus causing a national epidemic affecting the life of 1.4 billion people across the country (Zou, 2020).

As of February 29th, 2020,[1] over 330 cities in China had reported up to 79,824 infectious cases, with the spatial distribution shown in Fig. 1. It is clear from the figure that the number of cases varies across regions, demonstrating a geographically differentiated epidemic progression. According to the existing research about infection propagation, such differentiation is determined by various local factors like population density, prevention policy, medical level, transportation, etc. (Lowe et al., 2014; Gulland and Fox, 1992).

* Corresponding author.
E-mail address: changmin.jiang@umanitoba.ca (C. Jiang).

[1] China's CDC (Center for Disease Control) started to announce imported infections from foreign countries in March 2020. As the statistical date in this study is decided to be February 29th, 2020, to avoid the effects of international transportation network.

https://doi.org/10.1016/j.trc.2021.103231
Received 22 August 2020; Received in revised form 26 April 2021; Accepted 17 May 2021
Available online 2 June 2021

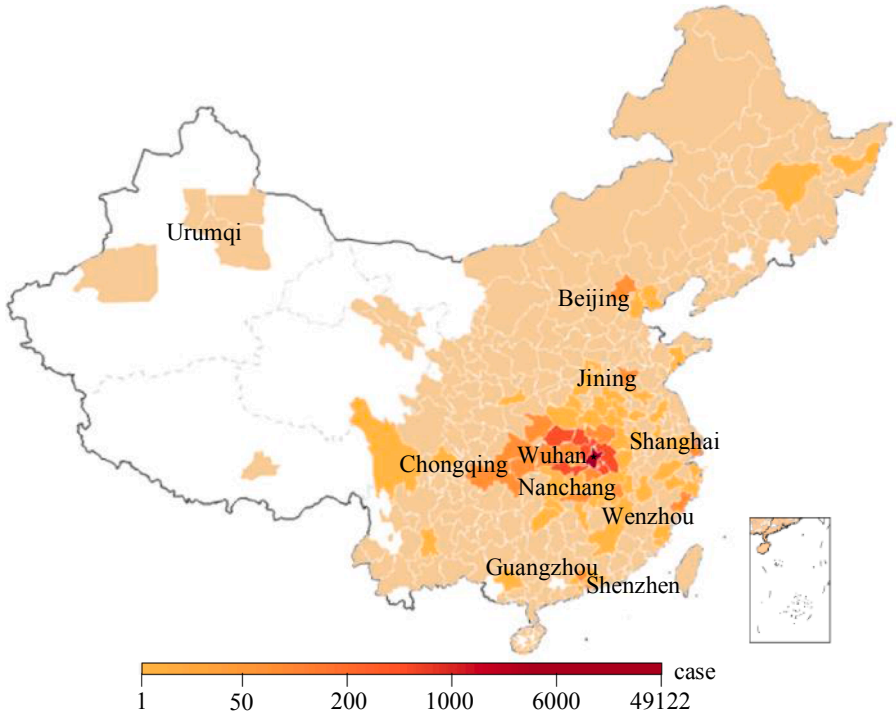


Fig. 1. The spatial distribution of COVID-19 cases until Feb. 29th, 2020.

At the same time, the relationship between infection and such local factors is also highly spatially heterogeneous (Kenneson et al., 2017; Macintyre and Ellaway, 2000).

Spatial heterogeneity is a very common feature of infection propagation (Bacchetti and Jewell, 1991; Cazelles and Hales, 2006), it describes the spatially differentiated relationship between infection and local factors. In other words, it means that the importance of a specific factor varies city by city (Gong et al., 2012). For instance, the air transport turnover may play a much more important role in the virus transmission in Beijing than in Hong Kong. However, it has not been adequately taken into account in most relevant studies (K. Wang et al., 2020). On the one hand, none of the prediction methods in the existing literature about COVID-19 propagation has considered spatial heterogeneity, leading to potentially biased forecasts at the city level. On the other hand, even for the more general infection propagation literature, there also exist some insufficiencies in the measures of spatial heterogeneity.

In particular, previous research has pointed out that the spatial heterogeneity of epidemic distribution is resulted from the heterogeneous spatial dependencies between cities (Lin and Wen, 2011). As per Tobler's First Law of Geography "*Everything is related to everything else, but near things are more related to each other*" (Tobler, 1970), the dependency is usually measured by Euclidean distance in most cases (Fotheringham et al., 1998). However, with the development of transportation, the Euclidean distance would no longer be suitable to depict how proximate one city is to another. Zhang et al. (2020) have shown that the spread of COVID-19 in China was closely correlated to the routes and the frequencies of domestic air, train and coach services,[2] and Müller et al. (2020) further figure out that the public transport system speeds up the COVID-19's transmission based on the simulation results in MATSim (Horni et al., 2016; Sun et al., 2021b).

Besides the transportation network itself, people's movement on the network driven by the social connection is also a key determinant influencing the infection propagation. For instance, the fact that about 3600 people travel from Wuhan to Wenzhou per day for family visit before the Chinese Lunar New Year causes a much higher infectious rate in Wenzhou than in first-tier cities like Beijing and Shanghai. Furthermore, the linkages to the epicenter on the transportation network should be paid more attention to when constructing the spatial dependency, because such links always bring much more infectious risks than others. Therefore, the intercity spatial dependency for measuring the heterogeneous pandemic distribution may not be simply described by a single attribute, but should be constructed incorporating all the above-mentioned attributes (e.g. Euclidean distance, travel time, frequency, people movement, and linkage risk).

Among the approaches to model spatial heterogeneity, the weighted regression model (GWR) has been widely applied thanks to its remarkable power of explanation (Fotheringham et al., 2002). The existing GWR models always use kernel functions to measure the spatial dependency, and these functions perform well in environmental or land-use studies but may not capture the complicated

[2] Several studies have also investigated the role of air transport in the spread of the COVID-19 pandemic internationally and found similar results (e.g., Christidis and Christodoulou, 2020; Sun et al., 2020).

connections between cities in our case (Yang, 2014). That is because the classical kernel function in GWR is suitable to measure the spatial dependency mainly affected by a single factor, but it is hard to take all relevant attributes into account. With this shortcoming in mind, we integrate deep neural network into GWR, which enables us to simulate the inter-city proximity considering the transportation networks, the people movement as well as the linkage risk (Hubbard et al., 2010; Lu et al., 2016). Thus, we propose a novel model called "transport proximity deep neural network weighted regression" (TPDNNWR), which integrates a deep neural network-based kernel function into a geographical weighted regression (GWR) model.

The contributions of the paper are two-fold. On the one hand, with the incorporation of spatial heterogeneity, our model would be the first, as far as we know, that can predict the spatially differentiated propagation of COVID-19 at the city level. This can facilitate our understanding towards how COVID-19 spreads through various transportation networks, and provide some much-needed policy suggestions regarding how to effectively contain the propagation of the virus by altering transportation operations. In particular, our model can accurately identify higher-risk cities, making it possible to impose heterogeneous control policy. On the other hand, we have also made a methodological contribution to the modelling of spatial heterogeneity. As a result, the effects of transportation networks in epidemic propagation can be more precisely examined.

The remainder of the article is organized as follows: Section 2 discusses the related literature; Section 3 introduces the modelling approach; Section 4 describes the collected data; Section 5 introduces the hyper-parameter tuning procedure, and Section 6 reports the results. Section 7 makes the scenario analysis and, finally, Section 8 draws conclusions.

2. Literature review

As demonstrated by Anselin (1988), the relationship between dependent and independent variables in spatial analysis always vary over space, and such variation, aka spatial heterogeneity, is quite complicated. This spatial heterogeneous relationship has been tested in the fields of geography, environment, social economy, tourism, medicine, etc. (Fuentes, 2002; Panek, 2019; Fuleky et al., 2014), which reveals the underlying mechanism behind the heterogeneity in spatial distribution (Fotheringham et al., 1996).

For the geographical distribution of infection cases, the spatial heterogeneous relationship has been found long ago (Zhou et al., 2011), and it has been proved to be affected by socio-economic, environmental and ecological variables (Jones et al., 2008). For instance, Fan and Ying (2005) found that there is a heterogeneous relationship between the number of SARS cases and the city location (environmental variable). Sirisena et al. (2017) showed the effects of climate (environmental variable) as well as population density (socio-econometric variable) on the spatial heterogeneity of the Dengue fever epidemic in Sri Lanka. Besides, other socio-economic variables such as GDP, education, income level, and number of hospital beds were all demonstrated to be important in affecting the spatial heterogeneous relationship (Ghosal et al., 2020; Lipner et al., 2017).

In addition to the above-mentioned variables, many studies have shed light on the importance of human connectivity in infection prevalence, due to the fact that *people become more mobile* with a developed transportation system (Kraemer et al., 2016). As per the results of Reimering et al. (2020) and Hsu and Babiker (2019), the prediction accuracy of infection numbers can be improved by incorporating routes and frequencies of flight and train. Furthermore, based on the statistical analysis, Zhang et al. (2019) pointed out that the cities with imported infectious cases would be attacked more seriously than the epicenter without effective prevention, meanwhile that transfer hubs might suffer higher risks than other cities. That is to say, the transportation network has significant influence on the pandemic propagation, so the spatial dependency should be constructed considering lines and nodes of transportation network.

To model the spatially heterogeneous relationship, Brunsdon et al. (1996) developed the geographical GWR. The model can give out parameters drifting across space using spatial linear regression, so it performs better than ordinary linear regression (OLR) which can only estimate parameters from a universal perspective (Brunsdon et al., 1997). Pu et al. (2017) used GWR to analyze the spatial heterogeneity in the sensitivity of parking fee in different blocks, and showed that the GWR model achieved higher prediction accuracy than the generalized linear model. As per the benefits of GWR, a number of studies have applied it to address spatial heterogeneity in epidemic distribution (Lin and Wen, 2011; Rokhman et al., 2019; Mohammadinia et al., 2017). In the modelling, the local regressive parameters of a specific region are estimated by incorporating its spatial dependency to other regions, and the effects of such dependency is measured by neighborhood weights calculated using kernel functions (Leung et al., 2000).

Various types of kernel functions have been implemented in GWR like Gaussian, Poisson and tri-cube (Song et al., 2016), and researchers have spent much effort in choosing the kernel type and deciding the bandwidth value to improve the data-fitting (Tasyurek and Celik, 2020). Reported by existing literature, the proposed kernel functions performed well in estimating the neighborhood weights when only incorporating mono-element like the Euclidean distance (Dziauddin, 2019). The Euclidean distance may accurately measure the neighboring relation in environmental or land-use study, but would not be appropriate to depict how proximate one city is to the other in our study. In the case study of the SARS and the H1N1 pandemics, Brockmann and Helbing (2013) have proved that the 'effective distance' derived from the air transportation network is more efficient than the geographic distance in predicting the epidemic arriving time. Furthermore, Jia et al. (2020) and Zhang et al. (2020) highlight the fact that COVID-19 is to a large extent brought by traffic flow from Wuhan to other cities in China through the public transportation network. In the sense, the transport proximity which has been found to be an important determinant of inter-city traffic flows (e.g., Zhang and Zhang, 2016; Zhang et al., 2018) should be implemented instead of Euclidean distance in this paper.

The transport proximity is similar to the transportation accessibility concept proposed by Koenig (1980), i.e., "the ease with which any destination can be reached from a location, using a particular transport system". Following the definition, the transport proximity can be recognized as the integrated accessibility which is determined by distance, travel time, travel fare, convenience of all available travel modes (Litman, 2016). Other than the transport proximity, the geographical dependency is also influenced by social connection

Table 1
Explanatory variables in TPDNNWR.

$x_i(k)$	Variables	Function
$x_i(1)$	Air passenger density	$m_{iA}/n_{iA}*f_{iA}$
$x_i(2)$	Rail passenger density	$m_{iH}/n_{iH}*f_{iH}$
$x_i(3)$	Road passenger density	m_{iE}/n_{iE}
$x_i(4)$	Population density	z_i

(Medina and Hepner, 2011), so the inter-city people movement which is a reflection of social connection, should be incorporated into the measurement of neighborhood weights (Southern, 2012). In addition, as we are discussing the geographical dependency in the context of an epidemic, the infectious risks on links of transportation network need to be highlighted. Therefore, the transport proximity, the social connection as well as the linkage risk will be integrated in order to measure the neighborhood weight.

However, as the kernel function of GWR is not able to take multiple attributes into account (Gastaldi et al., 2014), other available methods should be explored. Some approaches in the field of measuring accessibility should be mentioned, such as the opportunity based model (Chen et al. 2011), the potential model (Salze et al., 2011) and the utility model (Nassir et al., 2016). In particular, Sarlas and Axhausen (2015) proposed a spatial simultaneous autoregressive (SAR) model in which the regional dependency is measured on the travel time along different type of roads. It is theoretically probable to integrate the above models into GWR, but the estimation and mathematical proof would be complicated. Instead, Wu (2019) used the neural network to establish a novel "kernel function", and integrate it with OLR to build a geographically neural network weighted regression (GNNWR).

The "kernel function" constructed on neural network is more flexible than econometric models in modelling the nonlinear relation between neighborhood weights and selected attributes. Meanwhile the machine learning algorithm has the nature to help fit the data well, so the spirit of GNNWR will be adopted in our research. In addition, our "kernel function" may need a hierarchical structure, because the transport proximity, the social connection as well as linkage risk should be fed into the same level of the neural network, but the transport proximity needs to be measured by transport related attributes on the second level. In this sense, we try to update the GNNWR in Wu (2019) using deep neural network (DNN), thus creating a hierarchical "kernel function" involving a transport proximity.

DNN has received much more attention in transportation research for its high prediction accuracy (Ma et al., 2017). Duan et al. (2016) find its benefits of solving the traffic data imputation. Yi et al. (2017) apply DNN to distinguish the congestion situation on a transportation network. Wang et al. (2019) predict the traffic speed for an urban transportation network, and they prove that DNN is excellent in clarifying the interaction between traffic nodes on transportation network, so this approach will be suitable in modelling the geographical dependency in our research. More importantly, the flexible structure of DNN has advantages in depicting the hierarchical "kernel function". For instance, Wang et al. (2020b) construct a branched DNN structure to propose an alternative-specific utility function, and Sifringer et al. (2020) create a dedicated hierarchical DNN in describing the nested logit utility function. Therefore, we will use DNN to model the "kernel function" and to establish a "transport proximity deep neural network weighted regression" (TPDNNWR) model.

3. Modelling approach

In this section, three models, i.e., OLR, GWR as well as TPDNNWR, will be applied to interpret the heterogeneous relationship between the number of COVID-19 cases $\boldsymbol{Y}$ and the related explanatory variable $\boldsymbol{X}$ underlying the spatial dependency through a transportation network. Here, $\boldsymbol{Y}$ is the set of y_i which denotes the infection number of COVID-19 at the spatial point i, and $\boldsymbol{X}$ is the set of $x_i(k)$ meaning the kth local explanatory variable at the i^{th} location. According to existing researches relevant to COVID-19, the local attributes, including the passenger turnover volume, the transportation frequency and the station volume of air, rail and road transportation, have positive relations with the number of infectious cases (Hu et al., 2020; Zhang et al., 2020). Hence, we take the following variables into account: airport turnover volume per day (m_{iA}), rail station turnover volume per day (m_{iH}), express-way-based turnover volume per day (m_{iE}); airport volume (n_{iA}), train station volume (n_{iH}), inbound and outbound express way volume (n_{iE}); flight frequency per day (f_{iA}) and train frequency per day (f_{iH}).

Considering the potential strong correlation between the above variables, they are pre-processed using functions in Table 1 to generate the final explanatory variables ($x_i(k)$) related to local transportation, which are air, rail and road passenger densities, respectively. Besides, the population density (z_i) at spatial point i is also incorporated as per the existing studies about COVID-19 pandemic (Rashed et al., 2020; Sun et al., 2021a). In summary, the first three variables in the table represent the levels of business for the air, rail and road transportation on the ith point, and $x_i(4)$ shows the potential risk for local spread.

3.1. OLR and GWR

In order to help interpret the TPDNNWR, we first introduce the mechanism of OLR (ordinary linear regression) and the traditional GWR model. The OLR model is a classical econometric model, the model formulation is as Eq. (1), where β_0 is the intercept, β_k is the coefficient, and ε is the error term following the Normal distribution. In the studies of Büla et al. (1995), Liu et al. (2011), Küchenhoff et al. (2020), etc., the OLR has been applied to interpret the importance of different attributes to the infection in a global context.

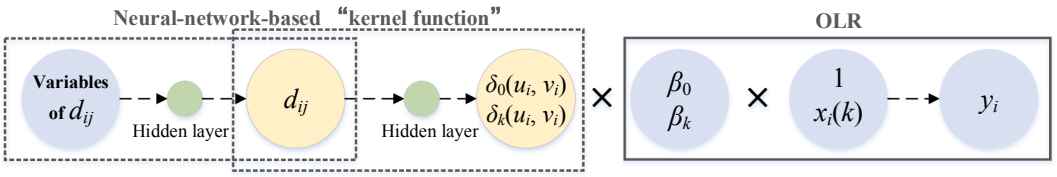


Fig. 2. Conceptualization of GNNWR.

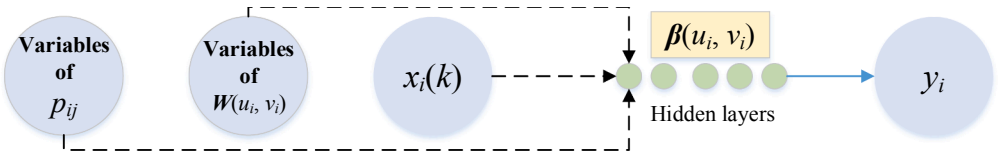


Fig. 3. Conceptualization of F-DNN.

$$y_i = \beta_0 + \sum_{k=1}^{K} \beta_k \cdot x_i(k) + \varepsilon\ , \quad i = 1,\ 2,\ \ldots, j,\ \ldots, I \tag{1}$$

However, the OLR is not able to clarify the spatial heterogeneous relationship between the number of infection cases and the local attributes, so the GWR model is employed (Brunsdon et al., 1996). In the model, the whole study area is divided into several regions, each region is treated as a spatial point *i*. The dependent variable in region *i* (y_i) can be explained by local explanatory variable $x_i(k)$ based on spatial dependency, and the local regression model is as Eq. (2),

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{K} \beta_k(u_i, v_i) \cdot x_i(k) + \varepsilon_i, \quad i = 1,\ 2,\ \ldots, j,\ \ldots, I \tag{2}$$

In the equation, (u_i, v_i) is the coordinates of spatial point *i*, $\beta_0(u_i, v_i)$, $\beta_k(u_i, v_i)$ and ε_i are the intercept, the coefficients and the error term for i^{th} location, respectively. To solve the model, the spatial weighting matrix $\boldsymbol{W}(u_i, v_i)$ composed of neighborhood weight $w_j(u_i, v_i)$ should be established as Eq. (3). $\boldsymbol{W}(u_i, v_i)$ and $\boldsymbol{\beta}(u_i, v_i)$ can be estimated as Eq. (4), in which $\boldsymbol{X}$ and $\boldsymbol{Y}$ are the matrix of explanatory and dependent variables of all spatial points.

$$\boldsymbol{W}(u_i, v_i) = \begin{bmatrix} w_1(u_i, v_i) & 0 & \cdots & 0 \\ 0 & w_j(u_i, v_i) & 0 & 0 \\ \vdots & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & w_I(u_i, v_i) \end{bmatrix} \tag{3}$$

$$\widehat{\boldsymbol{\beta}}(u_i, v_i) = \left(\boldsymbol{X}^T \boldsymbol{W}(u_i, v_i) \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{W}(u_i, v_i) \boldsymbol{Y} \tag{4}$$

In the traditional GWR, $w_j(u_i, v_i)$ is calculated by the kernel function established on the spatial proximity d_{ij}, e.g. the Euclidean distance or the travel time between *i* and *j*. For example, the bi-square kernel function is as Eq. (5), θ is the bandwidth value needs to estimate.

$$w_j(u_i, v_i) = \begin{cases} \left[1 - \left(d_{ij}/\theta\right)^2\right]^2 & d_{ij} < \theta \\ 0 & otherwise \end{cases} \tag{5}$$

3.2. Construction of TPDNNWR

Wu et al. (2020) proposed a neural network geographical weighted regression model (GNNWR) about coastal ecosystem. The model is composed of three parts, i.e., two neural networks and an OLR model, whose structure is shown in Fig. 2. The first neural network establishes the relationship between the relevant variables and the spatial proximity d_{ij}, while the second neural network further measures the weights of the regression coefficients ($\delta_0(u_i, v_i)$ and $\delta_k(u_i, v_i)$) using the output of the first neural network. $\delta_0(u_i, v_i)$ and $\delta_k(u_i, v_i)$ are later integrated into the OLR model in which β_0 and β_k have already been estimated from a universal perspective. In particular, $\beta_k(u_i, v_i)$ equals to $\delta_k(u_i, v_i) \times \beta_k$, while β_0 (u_i, v_i) equals to δ_0 $(u_i, v_i) \times \beta_0$. However, in this model, there is no mechanism to measure the spatial weighting matrix $\boldsymbol{W}(u_i, v_i)$.

TPDNNWR follows the conceptualization of Wu et al. (2020) to a certain extent, but there are some differences. First, a deep neural network (DNN) is implied instead of a normal neural network to enhance data fitting. Second, a hierarchical structure is constructed to simulate the mechanism of GWR. Third, the transport proximity p_{ij} between spatial points *i* and *j* is adopted. Besides, as the TPDNNWR is constructed based on the fully connected deep neural network (F-DNN) which has been widely applied in solving regression problems (Dobrescu et al., 2019), the results of the F-DNN (in Fig. 3) will be set as a benchmark to test the improvement of TPDNNWR.

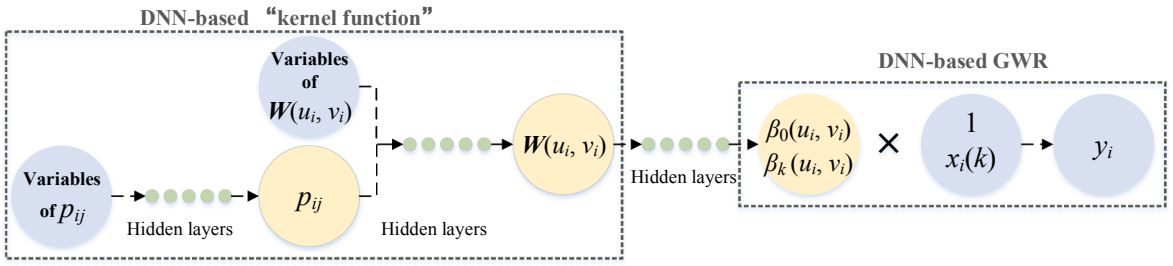


Fig. 4. Conceptualization of TPDNNWR.

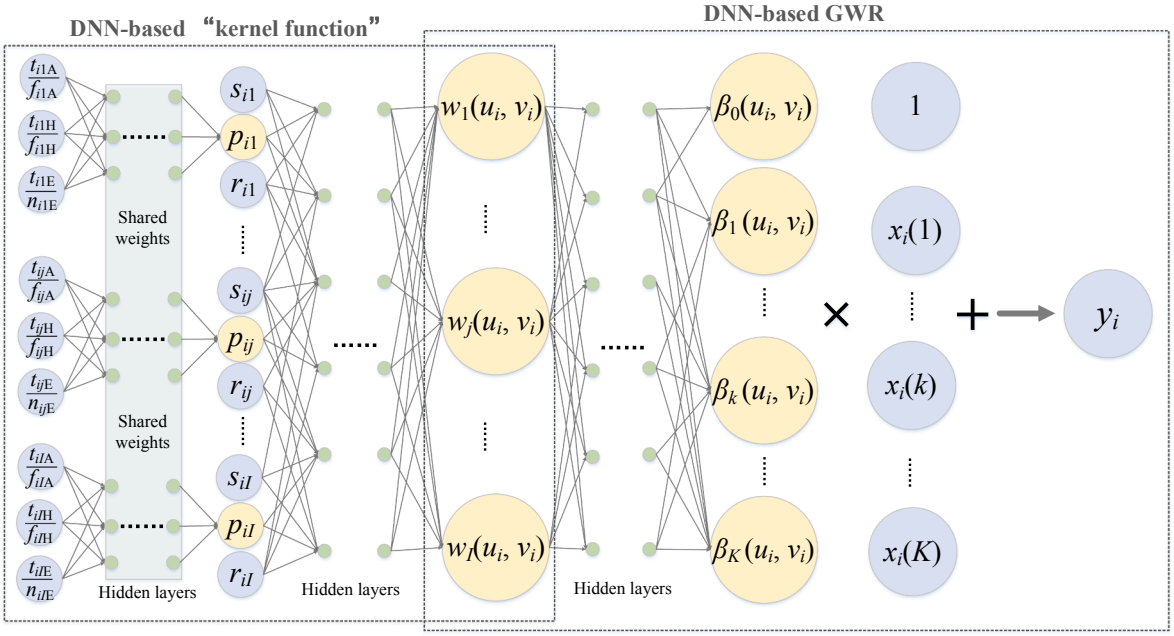


Fig. 5. Modelling structure of proposed TPDNNWR.

In the F-DNN infrastructure, $\boldsymbol{\beta}(u_i, v_i)$ will be obtained by mapping the variables of the ith spatial point to the corresponding number of infection cases y_i.

On the basis of F-DNN and GNNWR, the structure of TPDNNWR is constructed as shown in Fig. 4. The DNN-based "kernel function" updates the neural-network-based "kernel function" in Fig. 2 in order to raise the prediction accuracy and create a hierarchical structure measuring the geographical dependency $\boldsymbol{W}(u_i, v_i)$ based on the transport proximity p_{ij} and other related variables. Then, the DNN constructs a mapping from $\boldsymbol{W}(u_i, v_i)$ to $\beta_0(u_i, v_i)$ and $\beta_k(u_i, v_i)$, thus creating a structure simulating the estimation procedure of GWR.

Furthermore, we first introduce the DNN-based "kernel function", whose structure is in Fig. 5. In the first layer of the hierarchical structure, the neighborhood weight $w_j(u_i, v_i)$ is relevant to the transport proximity (p_{ij}), the social connection (s_{ij}) as well as the linkage risk (r_{ij}). The transport proximity (p_{ij}) measures the ease of reaching j from i through the transportation network. The social connection (s_{ij}) is represented by the people movement index between point i and j, which are constants collected from map.baidu.com. The linkage risk (r_{ij}) is determined by the product of the infectious risk indices on i (r_i) and on j (r_j), with the risk of Wuhan set to be 0.9 (r_{171}) and that of other cities set to be 0.1.[3] The formulation of $w_j(u_i, v_i)$ is shown in Eq. (6), where α and b represent the weight and the bias of the designed deep neural network, respectively. We need to point about that $w_j(u_i, v_i)$ is not only determined by the variables between i and j but also influenced by the variables on other links considering the transfer trips throughout the transportation network.

$$w_j(u_i, v_i) = DNN_{\alpha,b}(p_{i1}, p_{i2}, \cdots, p_{iI} \quad s_{i1}, s_{i2}, \cdots, s_{iI} \quad r_{i1}, r_{i2}, \cdots, r_{iI}) \tag{6}$$

[3] We set nine risks levels according to the WHO's coronavirus risk assessment which are: no risk (0.1), low risk (0.2–0.3), middle risk (0.4–0.6), high risk (0.7–0.8), very high risk (0.9).

Table 2
Data related to explanatory variables in TPDNNWR.

Data	City volume	Mean	standard deviation
Air turnover m_{iA} (10,000 people per day)	245	1.19	3.65
Airport volume n_{iA}	245	0.66	0.64
Flight frequency f_{iA}	245	67.21	186.23
Rail turnover m_{iR} (10,000 people per day)	314	2.98	5.69
Train station volume n_{iR}	314	2.53	2.29
Train frequency f_{iR}	314	150.6	179.04
Express way turnover m_{iE} (10,000 people per day)	338	10.52	13.71
Express way volume n_{iE}	338	7.96	6.64
Population density z_i (10,000 people/square km)	340	0.05	0.13

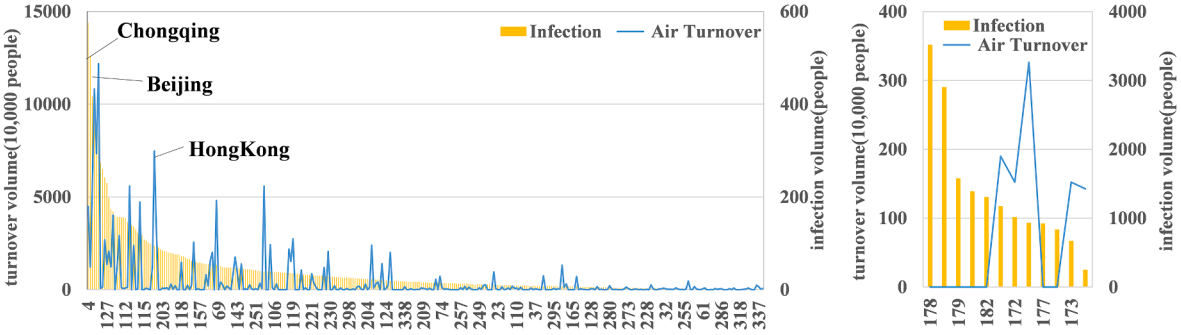


Fig. 6. The relationship between number of COVID-19 cases and airport turnover volume.

$$p_{ij} = DNN_{\alpha',b'}\left(\frac{t_{ijA}}{f_{ijA}}, \frac{t_{ijH}}{f_{ijH}}, \frac{t_{ijE}}{n_{ijE}}\right), \ (\alpha' = 0, b' = 0 | i = j) \tag{7}$$

Then, in the second layer, the transport proximity (p_{ij}) is determined by the travel time and the service frequency (Djurhuus et al., 2016; Boisjoly and El-Geneidy, 2017). Here, we take into account the travel time for air, rail and road (t_{ijA}, t_{ijR}, t_{ijE}), the travel frequencies for air and rail (f_{ijA}, f_{ijR}), as well as the road traffic volume (n_{ijE}). To simplify the structure, we generate joint variables to describe the transport proximity. In particular, p_{ij} is measured by Eq. (7), with α' and b' representing, respectively, the weight and the bias that need to be estimated. It should be noted that p_{ii} equals 0.

According to the estimation procedure of GWR in Eq. (4), $\boldsymbol{\beta}(u_i, v_i)$ is determined by the neighborhood weight $w_j(u_i, v_i)$, the explanatory variable $x_i(k)$ as well as the number of infection cases y_i. We first establish the relationship between $\boldsymbol{\beta}(u_i, v_i)$ and $w_j(u_i, v_i)$ through Eq. (8). In this equation, the mapping from $w_j(u_i, v_i)$ to $\beta_0(u_i, v_i)$ and $\beta_k(u_i, v_i)$ is constructed through DNN with the weight and the bias α'' and b'', where $w_j(u_i, v_i)$ is the result of the DNN-based "kernel function". $\boldsymbol{\beta}(u_i, v_i)$ is then integrated into the structure of classical GWR based on Eq. (9) to generate the estimated $\widehat{y}_i$ based on $x_i(k)$. The structure of DNN-based GWR is shown in Fig. 4.

$$\begin{aligned} \beta_0(u_i, v_i) &= DNN_{\alpha'',b''}\left(w_1(u_i, v_i), \cdots, w_j(u_i, v_i), \cdots, w_I(u_i, v_i)\right) \\ \beta_k(u_i, v_i) &= DNN_{\alpha'',b''}\left(w_1(u_i, v_i), \cdots, w_j(u_i, v_i), \cdots, w_I(u_i, v_i)\right) \end{aligned} \tag{8}$$

$$\widehat{y}_i = \beta_0(u_i, v_i) + \sum_{k=1}^{K} \beta_k(u_i, v_i) \cdot x_i(k), \quad i = 1, 2, \ldots, j, \ldots, I \tag{9}$$

4. Data collection and analysis

Our research covers the whole territory of China including 340 regions.[4] As we would like to confine our analysis to the spread of COVID-19 within China and the first foreign infectious case was reported on March 1st, 2020, the number of COVID-19 cases until February 29th, 2020, is chosen as the dependent variable of the model, and its distribution has been demonstrated in Fig. 1. Meanwhile, as the infection in Wuhan is to a large extent due to the local propagation, Wuhan itself will not be incorporated as a study point for prediction but the intercity connection between Wuhan and other cities will be taken into account.

Besides, the data that we collect to calculate the explanatory variables is summarized in Table 2. The air, rail and road turnover

[4] According to Chinese administrative division, 340 regions include four municipalities (Beijing, Tianjin, Shanghai, and Chongqing), Taiwan province, two special administrative regions (Hong Kong, Macau), 293 prefecture-level cities, 7 prefecture-level regions, 30 autonomous Prefectures, and three leagues.

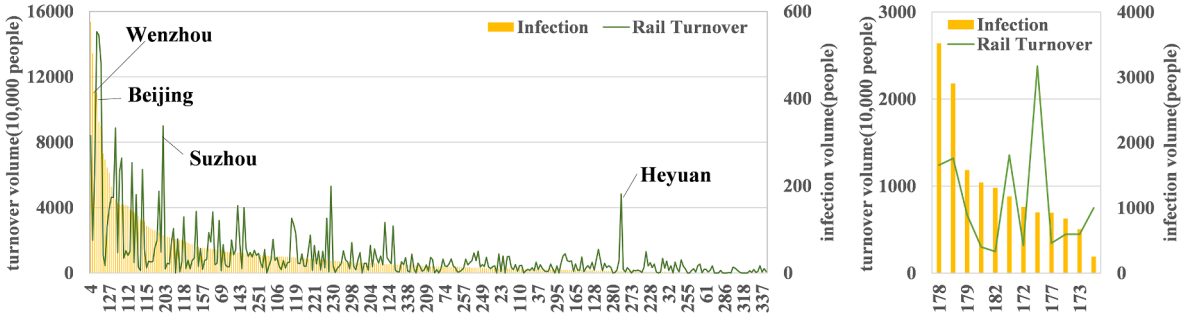


Fig. 7. The relationship between number of COVID-19 cases and rail turnover volume.

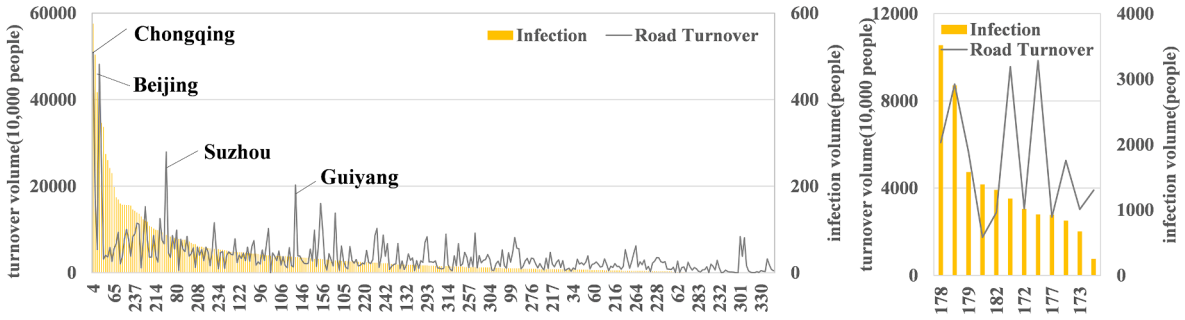


Fig. 8. The relationship between number of COVID-19 cases and road turnover volume.

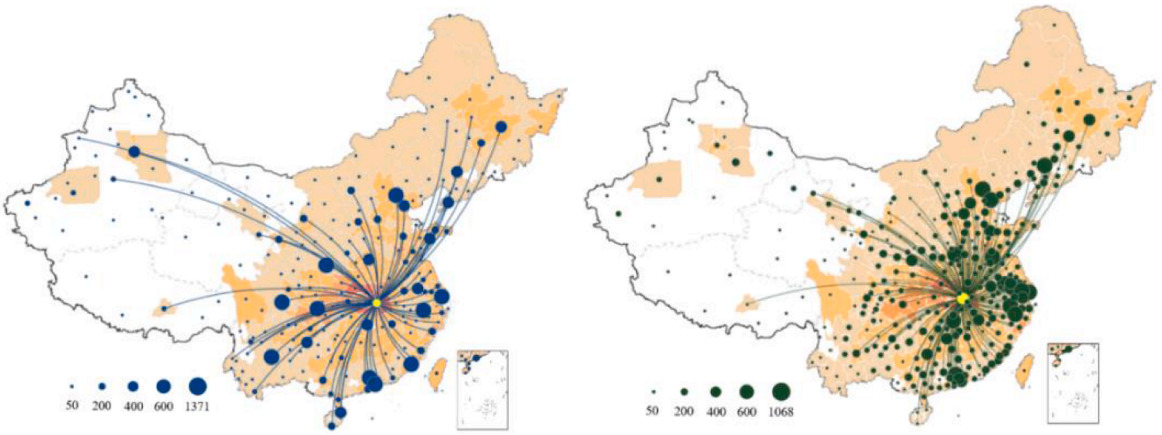


Fig. 9. Transportation network connecting Wuhan and other regions.

volumes are collected from the 2019 China Statistical Yearbook, while the airport, train station and highway traffic volumes are collected from the 2019 China City Statistical Yearbook. The flight and train frequencies are derived from the information scraped from the online travel agent Ctrip.com for the period between December 15th, 2019 and January 15th, 2020. The population density data is collected from the 2019 China Population Statistics Yearbook. In addition, the city volume in the table refers to the number of cities covered by different types of transportation networks.

To further clarify the heterogeneous relationship between number of COVID-19 cases and local transportation attributes, we add Figs. 6, 7 and 8 to demonstrate the relationship between the infection number and the turnover volume of local airport, train station and road-transportation. As the number of COVID-19 cases of Hubei province is much higher than that of other areas, we show the Hubei data separately. In addition, the city number is listed in the appendix.

It is easy to find the spatial heterogeneity between the number of COVID-19 cases and the local turnover volumes. For instance,

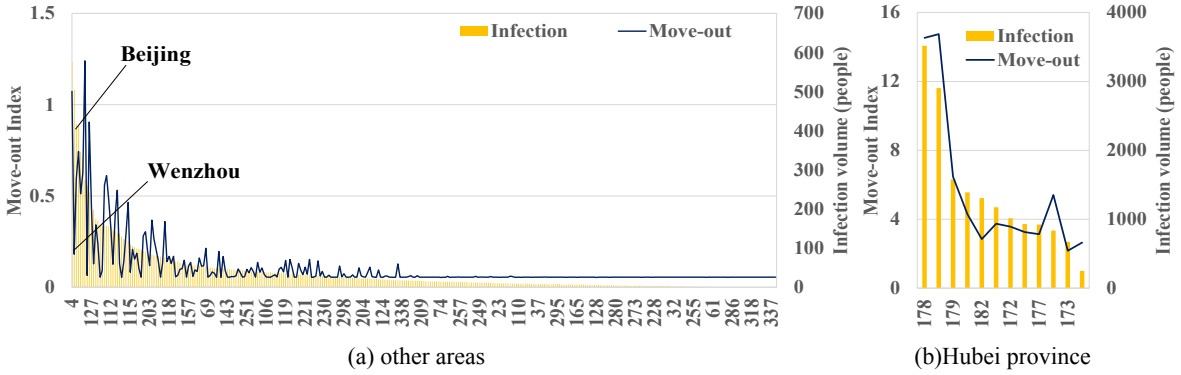


Fig. 10. The relationship between number of COVID-19 cases and move-out index from Wuhan.

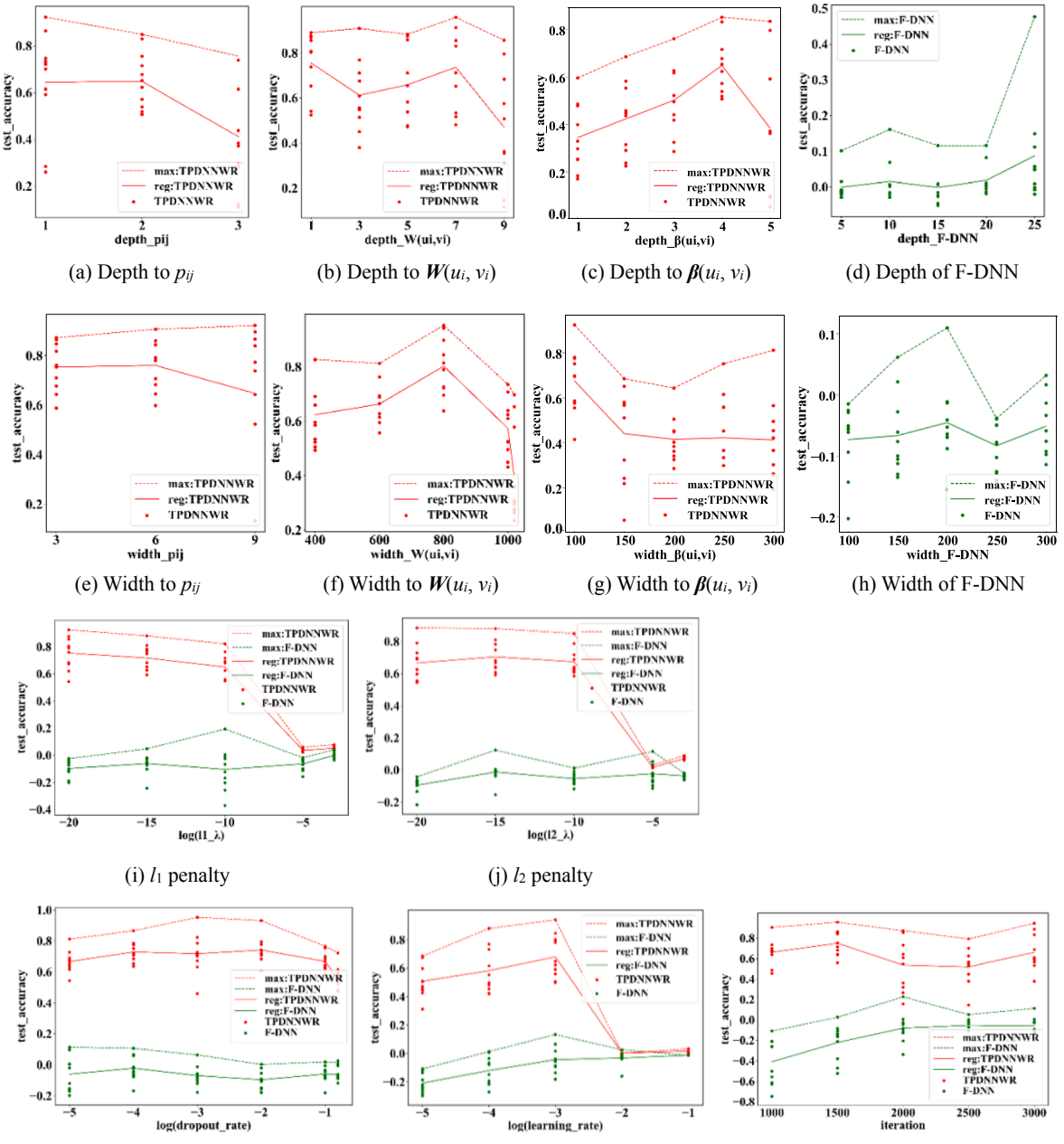
Table 3
Hyper-parameters and corresponding search space.

Hyper-parameters	TPDNNWR	F-DNN
Category 1: *invariant hyper-parameters*		
Activation function	Sigmoid	Tanh
Loss	MSE(5-folds)	MSE(5-folds)
Initialization	He Initialization	He Initialization
Category 2: *model-specific varying hyper-parameters*		
Depth	to p_{ij}: [1,2,3] to $\boldsymbol{W}(u_i, v_i)$: [1,3,5,7,9] to $\boldsymbol{\beta}(u_i, v_i)$: [1,2,3,4,5]	[5,10,15,20,25]
Width	to p_{ij}: [3,6,9] to $\boldsymbol{W}(u_i,v_i)$:[400,600,800,1000,1020] to $\boldsymbol{\beta}(u_i, v_i)$:[30,60,120,240,340]	[100,150,200,250,300]
Category 3: *general varying hyper-parameters*		
λ of l_1 penalty	$[10^{-20}, 10^{-15}, 10^{-10}, 10^{-5}, 10^{-3}]$	
λ of l_2 penalty	$[10^{-20}, 10^{-15}, 10^{-10}, 10^{-5}, 10^{-3}]$	
Dropout rate	$[10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1]$	
Learning rate	$[10^{-5}, 10^{-4}, 0.001, 0.01, 0.1]$	
Iteration	[1000, 1500, 2000, 2500, 3000]	

Chongqing has the largest number of COVID-19 cases outside Hubei province, it has lower air and rail turnover volumes but higher road turnover volume than the cities like Beijing, Shanghai, and Shenzhen. Wenzhou has the second largest number of COVID-19 cases, but its turnover volumes are rather low compared with other cites with larger number of confirmed cases. Hong Kong, Suzhou and Guiyang are the cities with a large airport, train station and coach station respectively, however, the number of COVID-19 casess of these three cities are not very high. For the cities in Hubei province, the distribution of the number of COVID-19 cases seems to be affected by rail and road transportation more than aviation. Besides, for constructing the model, the flight/train frequencies, the number of airport/train stations and express ways are also collected.

Then, in order to establish the weighting matrix, the data on air, rail and road networks is collected to measure the transport proximity. On the basis of the scraped information from Ctrip.com, 12306.com and map.baidu.com on January 15th, 2020, the transportation network of the three modes are obtained, with 22,854 flights involving 246 airports, 51,204 rail links involving 820 train stations, and 876 express ways. As it is difficult to depict all the linkages within one figure, we illustrate in Fig. 9 the air and the rail networks connected to Wuhan as examples. In this figure, the circles on the map represent the turnover volume of the airports and the train stations with direct flights and trains to Wuhan, and the circle scale refers to the level of turnover volume. In the figure, a total of 134 regions are connected to Wuhan by direct flights or trains. According to the statistics, these regions receive about 188 imported detections from Wuhan, thus inducing 5850 local infectious cases which comprise 45.39% of the total number of COVID-19 cases out of Hubei province. Besides, Beijing, Xiamen, and Chengdu are the three regions with the top frequent direct flights, and they also have high number of COVID-19 cases among the regions which are over 800 km away from Wuhan.

The social connection index (s_{ij}) in the model is represented by the people movement index collected from Baidu map. Taking the move-out index of Wuhan as an example (in Fig. 10), we could see 67% of the move-out people go to the cities in Hubei Province, proving that the high number of COVID-19 cases in these cities resulted from the close social dependency to Wuhan. Meanwhile, as a lot of people born in Wenzhou work in Wuhan, so the social dependency between Wenzhou and Wuhan is obviously higher than that between normal small cities and Wuhan, thus many infectious cases in Wuhan bring the virus to Wenzhou and cause a local pandemic. Moreover, the move-out index from Wuhan to Beijing seems not that high, which may state that the imported risks in Beijing is not only from Wuhan but from other cities due to its busy transportation system.



(a) Depth to p_{ij} (b) Depth to $\boldsymbol{W}(u_i, v_i)$ (c) Depth to $\boldsymbol{\beta}(u_i, v_i)$ (d) Depth of F-DNN

(e) Width to p_{ij} (f) Width to $\boldsymbol{W}(u_i, v_i)$ (g) Width to $\boldsymbol{\beta}(u_i, v_i)$ (h) Width of F-DNN

(i) l_1 penalty (j) l_2 penalty

(k) Dropout rate (l) Learning rate (m) Iteration

Fig. 11. Effects of hyperparamters on prediction accuracy. Note: the MSE loss is used in the coding for training the F-DNN and the TPDNNWR models, however the MSE is transferred to be the accuracy R^2 for simplifying the comparison of OLR, GWR and machine learning models.

5. Hyper-parameter tuning

5.1. Hyper-parameter space

In this section, the hyper-parameters of TPDNNWR, GNNWR as well as the F-DNN will be tuned, with TensorFlow 2.0 constructed on Python 3.6 as the coding framework. However, as the training accuracy of the GNNWR cannot converge, we will not discuss it. The hyper-parameters of the TPDNNWR and F-DNN are listed in Table 3, with specific value or search space respectively. Following Wang et al. (2020b), the hyper-parameters are classified into three categories, i.e., the invariant hyper-parameters, the model-specific varying hyper-parameters and the general varying hyper-parameters.

According to Fig. 5, TPDNNWR is composed of three segments, i.e., the segments to measure p_{ij}, $\boldsymbol{W}(u_i, v_i)$ and $\boldsymbol{\beta}(u_i, v_i)$, respectively.

Table 4
Chosen hyper-parameters of TPDNNWR.

Hyper-parameters		TPDNNWR	F-DNN	Hyper-parameters	TPDNNWR	F-DNN
Depth	to p_{ij}	2	25	λ of l_1 penalty	10^{-10}	10^{-10}
	to $\boldsymbol{W}(u_i, v_i)$	7		λ of l_2 penalty	10^{-10}	10^{-10}
	to $\boldsymbol{\beta}(u_i, v_i)$	4		Dropout rate	0.2	0.2
Width	to p_{ij}	9	200	Learning rate	10^{-3}	10^{-3}
	to $\boldsymbol{W}(u_i, v_i)$	800		Iteration	1500	3000
	to $\boldsymbol{\beta}(u_i, v_i)$	100				

Sigmoid is chosen to be the activate functions of the hidden layers for TPDNNWR, while Tanh is set as the activation function in F-DNN for quick convergence (Liu and Di, 2020). Moreover, MSE loss function is applied and the He initialization method is adopted (Zhao et al., 2017; He et al., 2015). Meanwhile, the cross validation (5-fold) is adopted to reduce the risk of overfitting (Benyamin and Mark, 2019).

In the application, the sample (339 city points in total) is divided into a testing set (10% of the data) and a training set (90% of the data), with the training set further divided into 5 parts with equal size to implement cross validation. In the cross validation, the model with the lowest MSE is the best model conditioning on the fixed hyper-parameters.

The depth and the width denote the hidden layer volume and the neuron number on each layer. In addition, l_1 and l_1 penalties are added in the loss function to eliminate the risk of overfitting (Ng, 2004), with λ being the hyper-parameter of the penalty. The dropout rate tries to keep the sparsity of DNN configuration (Dmitry et al., 2017) and the learning rate determines the convergence speed (Zeiler, 2012).

To define all the varying hyper-parameters in Table 3 to be γ', the best hyper-parameter γ'* should be searched based on Eq. (10), where γ is the potential optimal parameter, and $\lambda||\gamma||_{penalty}$ is the added regulation to the loss function. Hence, the best hyper-parameter γ'* can be found by minimizing the empirical risk caused by the difference between γ and γ' $ER(\gamma, \gamma')$. In other words, the γ'* should make the designed DNN with the highest prediction accuracy conditioning on the alternatives in the defined space in Table 3.

$$\min_{\gamma} ER(\gamma, \gamma') = \min_{\gamma} \frac{1}{I} \sum_{i=1}^{I} \left(\widehat{y}_i|_{\gamma} - \widehat{y}_i\Big|_{\gamma'} \right)^2 + \lambda \|\gamma\|_{penalty}, \quad i = 1,\ 2,\ \ldots, j,\ \ldots, I \tag{10}$$

5.2. Tuning procedure

With varying hyper-parameters, several models involving TPDNNWR and F-DNN are trained, and the efficiency of the hyper-parameter change on the variation of prediction accuracy is calculated based on the test data set, with the results being illustrated in Fig. 11. In the figures, the vertical axis denotes the predicting accuracy ($R^2 = \sum_{i=1}^{I} \left(y_i - \widehat{y}_i \right)^2 / \sum_{i=1}^{I} \left(y_i - \bar{y} \right)^2$), where $\bar{y}$ is the mean value of y_i, and the horizontal axis shows the search space. Meanwhile, the red and the green points represent the prediction accuracy of TPDNNWR and F-DNN, respectively, and the dotted lines connect the top prediction accuracy of the models conditioning on specific hyper-parameters.

The first two rows demonstrate the efficiency of the model-specific varying hyper-parameters, i.e., the depth and the width. Fig. 11 (i) and (j) show the efficiency of l_1 and l_2 penalty and the last three figures describe the variation of predicting accuracy to the change of other general varying hyper-parameters in Table 3.

Model-specific varying hyper-parameters: the first row shows the effects of depth on the prediction accuracy, and the accuracy of TPDNNWR is 70%–85% higher than that of F-DNN over the varying depth ranges. More specifically, the prediction accuracy of TPDNNWR is more sensitive to depth_$\boldsymbol{\beta}(u_i, v_i)$, and the maximized top values appear when depth_p_{ij}, depth_$\boldsymbol{W}(u_i, v_i)$ and depth_$\boldsymbol{\beta}(u_i, v_i)$ equal 1, 7 and 4. When depth_p_{ij} exceeds 2, depth_$\boldsymbol{W}(u_i, v_i)$ exceeds 7 and depth_$\boldsymbol{\beta}(u_i, v_i)$ exceeds 4, the mean prediction accuracy starts to decrease. The second row presents the efficiency of the architectural width, where TPDNNWR also shows much better prediction accuracy than F-DNN, and the top prediction accuracy of TPDNNWR appears when width_p_{ij}, width_$\boldsymbol{W}(u_i, v_i)$ and width_$\boldsymbol{\beta}(u_i, v_i)$ equal 9, 800 and 100, respectively. It is clear that the top and the mean prediction accuracy of TPDNNWR is more sensitive to the variation of width_$\boldsymbol{W}(u_i, v_i)$.

General varying hyper-parameters: Fig. 11(i) and (j) illustrate the changing prediction accuracy with the two regularizations l_1 and l_2 penalties. The prediction accuracy of TPDNNWR shows a declining trend with the increment of l_1 and l_2 penalties. Meanwhile, when the l_1 and l_2 penalties are larger than 10^{-10}, the prediction accuracy of TPDNNWR starts to drop much more rapidly and converge with the prediction accuracy of F-DNN. Hence, the l_1 and l_2 penalties are efficient on regularizing the model with the purpose of reducing the risk of over fitting, but the value of penalties should not be too large as the training set size in our study is small.

The dropout rate is another hyper-parameter that regularizes the model. The maximized top prediction accuracy of TPDNNWR is associated with the dropout rate 0.001, but the mean value of the prediction accuracy does not change significantly, so we decide to choose a larger dropout rate which is over 0.1 considering the risk of over fitting. Then, Fig. 11(l) demonstrates the efficiency of varying learning rate, and the best value of learning rate equals 0.001. Fig. 11(m) shows the variation of the prediction accuracy to the increase of iteration. TPDNNWR seems not sensitive to the iteration variation, so 1500 times of iteration should be enough for training TPDNNWR, but F-DNN may need more iterations.

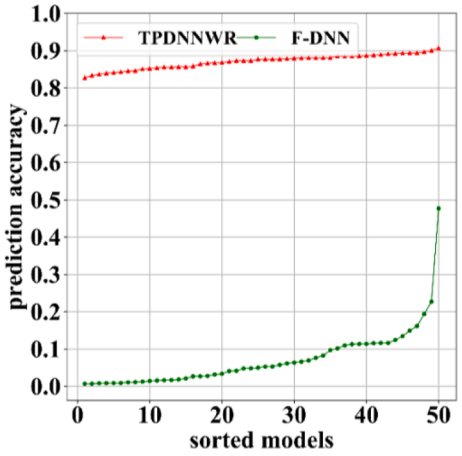


Fig. 12. Prediction accuracy of TPDNNWR and F-DNN.

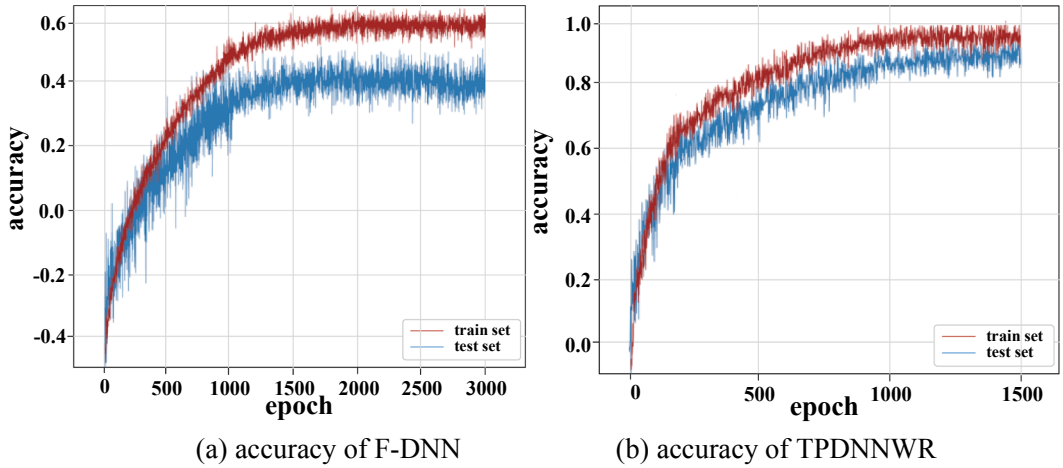


Fig. 13. Training and testing curves of F-DNN and TPDNNWR.

5.3. Prediction accuracy

Based on the efficiency of hyper-parameters on improving the prediction accuracy, we set the hyper-parameters of the two models in Table 4, 200 TPDNNWR and 150 F-DNN models are trained.

The top 50 results of TPDNNWR and F-DNN conditioning on the hyper-parameters in Table 4 are plotted in Fig. 12, the red and the green points represent the sorted prediction accuracy of the TPDNNWR and F-DNN from low to high, respectively. It is obvious that TPDNNWR consistently provides higher prediction accuracy than F-DNN does, indicating that the designed structure of TPDNNWR successfully establish a mapping from the transportation network, the social connection and linkage risk to the epidemic distribution. The prediction accuracies (R^2) of the top F-DNN model and the TPDNNWR model in Fig. 12 are 47.7% and 90.7%, respectively; and the parameter volumes of the two models are 31,901,491 and 4,094,361, respectively. The two final models will be applied to be compared with OLR and GWR in the following section.

Moreover, Fig. 13 demonstrates the training and the testing curves of F-DNN as well as TPDNNWR. It can be found that the low prediction accuracy of F-DNN may be mainly due to the F-DNN model structure because the training accuracy cannot go higher after the 1000th iteration. The reason is that the structure of F-DNN is similar to that of an OLR model in which every city point shares the same $\beta(u_i, v_i)$. Meanwhile, the testing loss of TPDNNWR goes down continuously with the training loss, but there is a difference of about 10% between the testing accuracy and the training one. Hence, to some extent, there exists overfitting limitation in the TPDNNWR model, which may be mainly due to the limited sample size. Furthermore, the robustness of F-DNN and TPDNNWR is tested and the results can be found in Appendix C (Wang et al., 2021; Goodfellow et al., 2014; Kurakin et al., 2016)

Table 5
Estimation results of OLR and GWR models.

		Intercept	Air passenger density	Rail passenger density	Road passenger density	Population density	Accuracy R^2
OLR	Coefficient	0.004	−0.040	0.027	0.126	−0.027	0.041
	P value	0.02*	0.96	0.46	0.11	0.85	
GWR1	Coefficient	0.0036~	−0.0406~	0.0269~	0.1250~	−0.0279~	0.072
Euclidean distance		0.0042	−0.0302	0.0288	0.1280	−0.0249	
	P value	0.0621*~	0.8933~	0.6938~	0.3514~	0.8021~	
		0.0819*	0.9014	0.7122	0.3530	0.8118	
GWR2	Coefficient	0.0032~	−0.0430~	0.0259~	0.1250~	−0.0266~	0.091
generalized travel		0.0045	−0.0357	0.0299	0.1282	−0.0274	
time	P value	0.0861*~	0.9920~	0.5661~	0.5884~	0.9554~	
		0.0862*	0.9999	0.5994	0.5894	0.9555	
F-DNN	Coefficient	–	–	–	–	–	Train: 0.613 Test: 0.477
TPDNNWR	Coefficient	0.0016~	0.0036~	0.0021~	0.0048~	0.0048~	Train: 0.996
		0.7692	0.1045	0.1405	0.6244	0.2232	Test: 0.907

***Significant at the 1% level; **Significant at the 5% level; *Significant at the 10% level.

6. Comparison between OLR, GWR and TPDNNWR

The correlation between the explanatory variables has been examined (Appendix B) before the estimation. OLR is estimated on the basis of Eq. (1). The Euclidean distance and the generalized travel time (air, rail and road travel time are assumed with equal importance) between the study points are set as the variables of the kernel function in estimating GWR, and the package "spgwr" in R 3.6.3 (Bivand and Yu, 2017) with Gaussian kernel function has been applied. The hyper-parameters of TPDNNWR are set as in Table 4. Then, the results of OLR, GWR and TPDNNWR are summarized in Table 5.

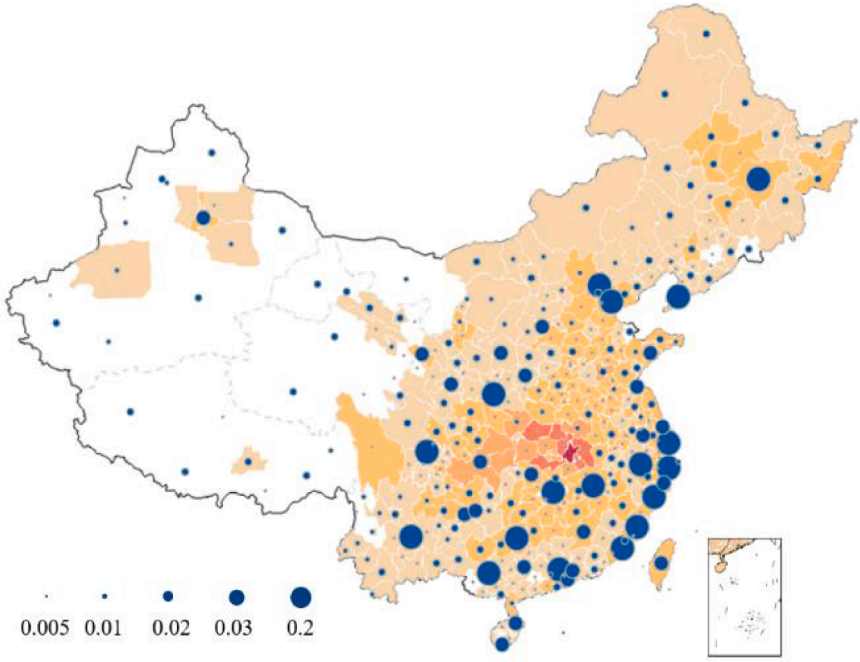
The last column of Table 5 illustrates the accuracy (R^2) of OLR, GWR, top F-DNN and top TPDNNWR (in Fig. 12). It is obvious that TPDNNWR (the top one in Fig. 12) has the highest prediction accuracy, the F-DNN has the second high accuracy, while the GWR models have higher accuracy than OLR but the difference is not significant. Specifically, the estimated R^2 of OLR only equals 0.041, indicating that the OLR results can not fit the data well. For the results of GWR, when applying the Euclidean distance, the optimized bandwidth value is 72.43, and the R^2 is 0.072; when using the generalized travel time, the optimized bandwidth value is 4.47, and the R^2 equals 0.091. Comparing the two GWR models, the implementation of travel time seems insufficient in improving the estimation accuracy.

Although OLR and GWR do not provide high prediction accuracy, some statistical results can be discovered from the results, e.g. the 'Road passenger density' plays the most important role in affecting the local infection, and the 'Air passenger Density' seems to have the lowest impacts. The results indicate that the road transportation may be the most important mode of transmitting the virus in China. Meanwhile, when incorporating the travel time into the kernel function of GWR, the variation ranges of the coefficients become larger, indicating that the second GWR model has advantages in depicting the spatial heterogeneity. However, the estimated coefficients of most OLR and GWR variables are not significant, that is to say, the current OLR and GWR models may not explain the spread of the COVID-19 pandemic very well. For instance, the coefficients of air passenger density and population density are less than 0, which may be inconsistent with the common sense.

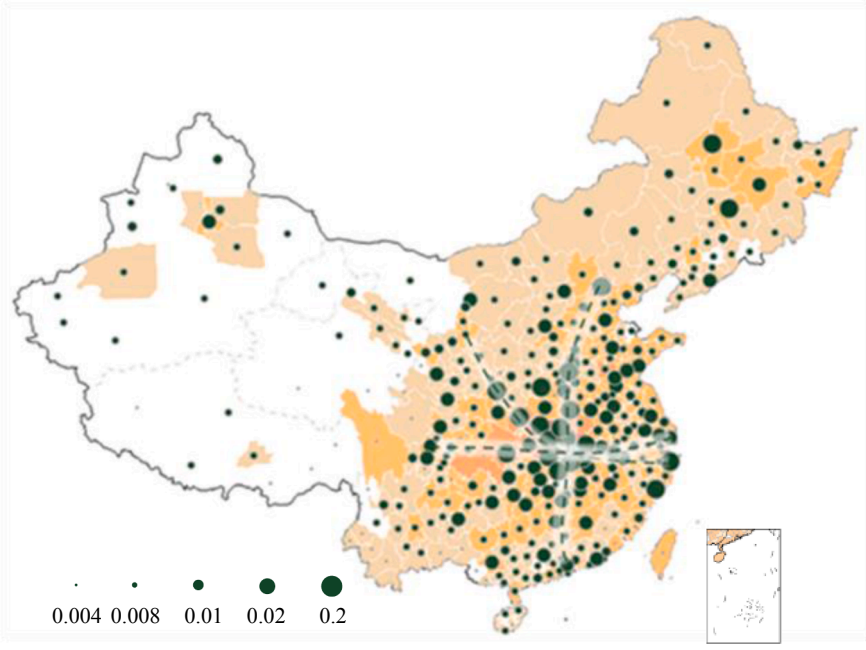
There are a few reasons behind the low prediction accuracy of GWR model: i) the importance of different modes in GWR is assumed to be equalized (due to the limitation of data collection), which may lead to inaccurate measurement; ii) some attributes relevant to the transportation network such as flight/train frequency are ignored in GWR, only the travel time may not depict the transport proximity well; iii) the lines linked to Wuhan are recognized as the same connecting routes as others in the current GWR models, thus the linkages with different level of infectious risks cannot be distinguished.

Table 5 also illustrates the coefficients trained/predicted by TPDNNWR, which are averaged from the value of coefficients ($\beta_0(u_i, v_i)$ and $\beta_k(u_i, v_i)$) of the first 10 TPDNNWR models with the highest prediction accuracy levels in Fig. 12. Generally, the variation range of the coefficients of TPDNNWR is larger than that of the coefficients of GWR2, suggesting that TPDNNWR is better at capturing the spatial heterogeneous relationship between the number of infection cases and the local variables. Besides, the mean values of coefficients are: 0.0163 for the intercept, 0.0074 for the air passenger density, 0.0105 for the rail passenger density, 0.0355 for the road passenger density, and 0.0318 for the local population density. It shows that the importance of the variables measured by TPDNNWR is consistent with that estimated by GWR, specifically, the importance of the road, the rail and the air passenger densities to the local infection are from high to low. In other words, the rail and the road transportation play more important roles than the aviation does during the pandemic. As the coefficients demonstrate obvious heterogeneity across cities, their geographical distributions are plotted on the map in Fig. 14, with the circle scale set according to the coefficient value.

Fig. 14(a) demonstrates the coefficients regarding the air passenger density. There are 24 cities associated with large coefficients (higher than 0.02) and the mean distance between these cities and Wuhan is about 800 km. The result follows the common sense that aviation brings virus to the cities far away from the epicenter. Moreover, the distribution of the cities with large coefficients is very



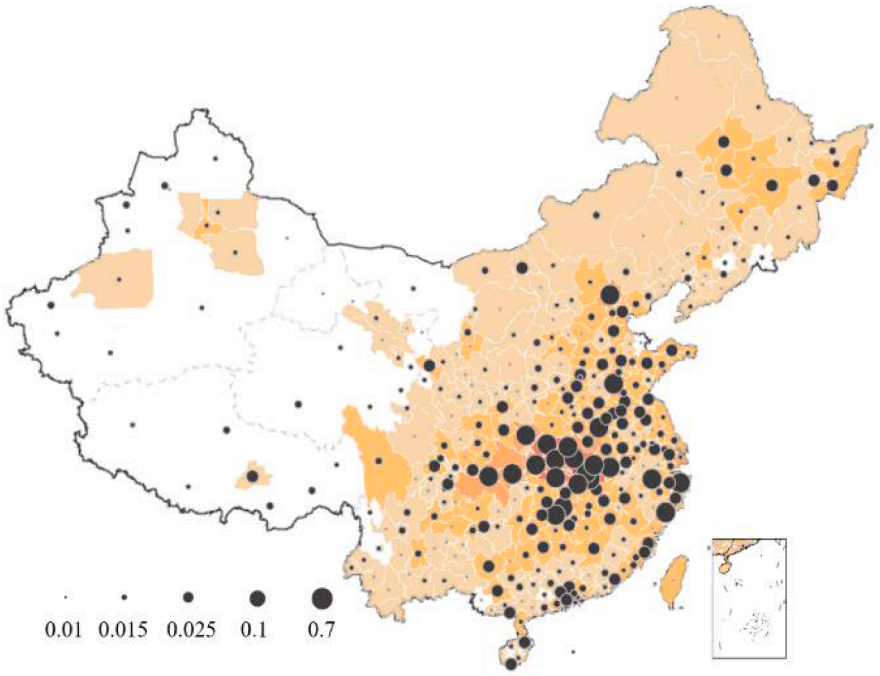
(a) Air passenger density



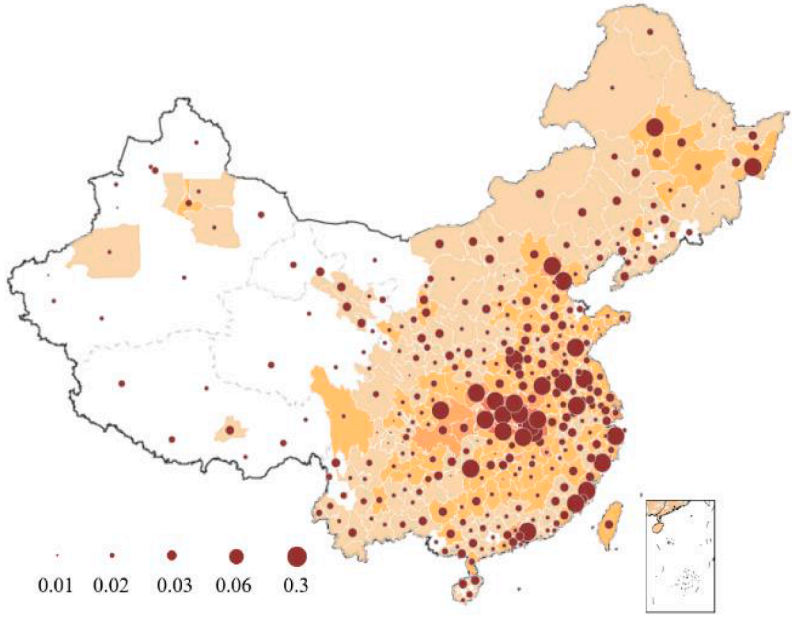
(b) Rail passenger density

Fig. 14. The coefficients of variables of TPDNNWR.

scattered, suggesting that the cities suffering from serious virus attack on the air transportation network show a punctate distribution. Among the 24 cities, 16 are linked to Wuhan with direct air routes and 8 (including Beijing, Hangzhou, Chengdu, Ha'erbin, Qingdao) have over 10 return flights per day. However, the other 8 cities which are also highly affected by the air passenger density do not have



(c) Road passenger density



(d) Population density

Fig. 14. (*continued*).

direct flights to Wuhan. These are the regional hubs on the air transportation network, e.g. Nanjing, Hefei, Changsha, etc. Hence, the cities with hub airports may receive high virus attack even without direct flights to the epicenter. That is because there is a higher probability for hubs to receive infection cases arriving from transfer routes. In this sense, a point-to-point network may perform better than a hub-and-spoke one in epidemic prevention, and cutting off the flights of hub airports should be an effective way of pandemic containment.

Table 6
Predicted number of COVID-19 cases under different traffic restrictions.

Category	Level	Restricted flights	Restricted trains	Restricted roads	Detection reduction percentage	Efficiency
One	20%	4,571	10,241	438	37.2%	40,995
	50%	11,427	25,602	438	49.5%	75,690
	70%	15,998	35,843	438	**59.8%**	87,423
Two	20%	3,565	3,778	212	36.2%	20,870
	50%	8,912	9,445	212	47.3%	39,257
	70%	12,477	13,223	212	56.5%	45,862
Three	20%	73	121	17	36.1%	**584**
	50%	182	303	17	46.6%	1,077
	70%	255	424	17	57.3%	1,214

Efficiency= (Restricted flights + Restricted trains + Restricted roads)/ Detection Reduction percentage.

In Fig. 14(b), the cities with high coefficients of rail passenger density (larger than 0.02) seem to concentrate more than the ones in Fig. 14(a), and they show a banded distribution. In particular, the first band goes along the Yangtze river and involve the high-speed railways connecting southeastern cites and southwestern ones, including the "Shanghai-Chengdu" line and the "Ningbo-Chengdu" line. The second band is parallel with the "Wuhan-Xi'an" highspeed railway connecting central cities and Northwest cities. The third band is a vertical one, which connects Beijing and Guangzhou via the Hubei Province (where Wuhan is located). Besides, because of the higher train frequency and the closer social connection, cities on the first band suffers from higher infectious risks than the ones on the other two bands. Different from the starlike pandemic distribution on air transportation network, the rail transportation network leads to a clear transmission trajectory through the epicenter.

Fig. 14(c) is about the coefficients of road passenger density, where the 9 cities with the highest coefficients (larger than 0.3) are all located in Hubei province, suggesting that the road transportation network is the key factor resulting in explosive regional pandemics around Wuhan during the COVID-19 pandemic. Other points highly affected by the road transportation (with coefficients larger than 0.1) also show a group-based pattern, such as the city groups centered around Beijing, Chongqing, Changsha and Hangzhou. From the figure, we can see that the transmission radius of road transportation is limited, suggesting that this network may not lead to wide epidemic propagation.

From the analysis, we can conclude that two types of cities face a high probability of coronavirus attack, i.e., the transportation hubs and the cities which are closely linked to Wuhan, the epicenter. For example, Wenzhou is probably the best example for the second type, as although it does not have a busy transportation system, it has close social dependency with Wuhan and other cities of Hubei province. Therefore, its coefficients of the air-, rail- and road- passenger densities are significantly higher than those of cities with similar sizes, which shows that the social dependency through the transportation network is as important as the other relevant factors. Moreover, the population density is also a key factor affecting the number of COVID-19 cases, especially for the large-scale cities located along the east coastline of China.

7. Discussion

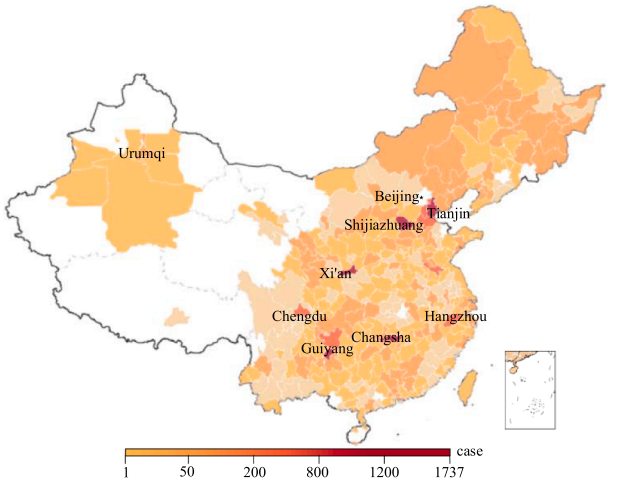
In the previous sections, we have proven that the TPDNNWR model has a very high predictability accuracy. In other words, it can be applied to generate various policy insights to help the containment of the further spread of COVID-19. In this section, we try to predict the epidemic distribution of COVID-19 in China in two scenarios using the top model of TPDNNWR in Fig. 12, so as to illustrate the potential usage of the model. The two scenarios are: i) assuming that different levels of traffic restriction policies are conducted and ii) assuming that the COVID-19 is firstly discovered in Beijing.

7.1. Prediction results under different restriction policies

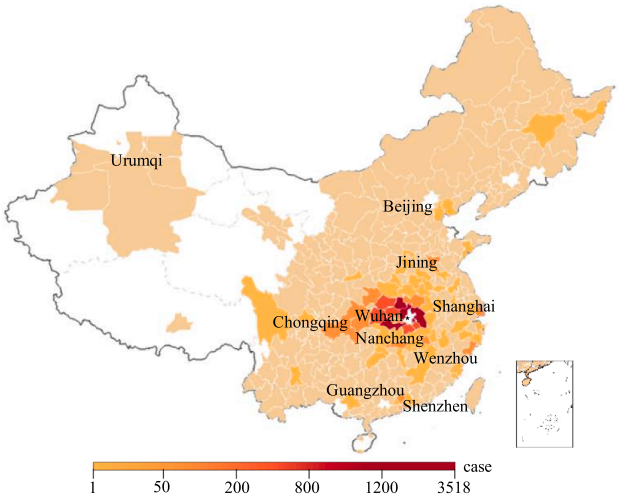
Three categories of traffic restriction are assumed: 1) setting the same traffic restriction policy for all the cities in China, 2) setting traffic restriction only for hubs on the transportation networks, and 3) setting traffic restriction only on the routes connecting the epicenter.

In the first category, the traffic of flights, trains and road transport in China will be cut down by 20%, 50% and 70%, respectively. In other words, the values of air travel frequency ($f_{ij\mathrm{A}}$), rail travel frequency ($f_{ij\mathrm{R}}$) and express way traffic volume ($n_{ij\mathrm{E}}$) between cities will be reduced, and the social connection (s_{ij}), the turnover volume ($m_{i\mathrm{A}}$, $m_{i\mathrm{E}}$, $m_{i\mathrm{R}}$), the frequency ($f_{i\mathrm{A}}$, $f_{i\mathrm{R}}$), and the express way volume ($n_{i\mathrm{E}}$) should also decline accordingly. The second category of restriction relaxes the traffic restrictions on small-scale cities, but plans to reduce 20%, 50% and 70% of the flights, trains and road transport traffic departing from and arriving at the hub cities. The hubs are defined to be the top 50 cities in the rank of total turnover volume in 2019 (the sum of air, rail and road turnover volume). The third category of restriction is designed to reduce 20%, 50% and 70% of the flights, trains and road transport traffic to the epicenter Wuhan. The policy efficiency of containing the virus pandemic is shown in Table 6. In the table, column 6 illustrates the predicted drop ratio of the number of COVID-19 cases under the restriction policies, and the efficiency is the volume of transportation links that should be restricted when we want to reduce the total number of COVID-19 cases by 1%.

Table 6 suggests that the policies in the first category have the best performance in reducing the total number of COVID-19 cases. However, the efficiency regarding this type of policy is the lowest, because the importance of the transport links to the pandemic is not



(a) Epicenter in Beijing



(b) Epicenter in Wuhan

Fig. 15. Pandemic distribution with epicenter switch.

distinguished, and the travel ban across the country will result in a sudden decrease of intercity mobility and may cause social and economic problems. Compared to the policies in category one, those in category two have almost the same effects on reducing the number of COVID-19 cases, but the number of restricted links decreases by about 50%, so the efficiency of category two policies is significantly higher. It indicates that the hubs on transportation networks play critical roles in virus transmission because the access to and the egress from hubs lead to substantial gatherings of people. As a result, 'affected' trunk lines between hubs will further bring the risks to branch routes, which may lead to a wider spread of the pandemic. In this sense, setting restrictions on hubs of transportation networks should be an efficient and practical policy of disease containment.

Although the drop ratios in category three are slightly lower than those in the other two categories, the efficiency in this category is the highest. Hence, in the context of COVID-19 transmission in China, restricting the routes connecting the epicenter seem to be the most effective way of containing the pandemic. However, we should point out that there are some special preconditions to guarantee the success of policies in category three. First, they need to be applied before the coronavirus was widely transmitted in China. Therefore, such policies should be implemented at the beginning of the infection right after the epicenter has been discovered. Second, we should also be aware of the fact that the original epicenter of COVID-19, Wuhan, is in itself an important hub located in Central China. In other words, the travel bans in Wuhan as well as in the cities of Hubei province also reduce the connectivity between other regions of China, especially on the transportation networks of rail and road. Therefore, the superior performance of the third category of policies may not necessarily persist if the epicenter is in another location.

7.2. Prediction results of changing epicenter

In the second scenario, we assume that COVID-19 is first discovered in Beijing, and discuss whether epicenter location will influence the distribution of the pandemic. In the analysis, the infectious risk value of the linkages connecting Beijing and other cities (r_{ij}) is increased to 0.9, while the one on the routes connecting Wuhan is reduced to 0.1. The predicted distribution of the number of COVID-19 cases with the assumed epicenter Beijing is illustrated in Fig. 15, and the real pandemic distribution with the real epicenter Wuhan is also listed for comparison.

As per the prediction results, the total number of COVID-19 cases increases from 30,848 (excluding Wuhan) to 52,907 (excluding Beijing and Wuhan) if the epicenter switches from Wuhan to Beijing and the coronavirus has been transmitted across China without control for about 2 months. The increase is due to the fact that Beijing obtains a more developed transportation system which transports about 7.4%, 4.1% and 3.7% of passengers in China in 2019 through air, rail and road links, respectively. Besides the increment of the total number of COVID-19 cases, the distribution of the pandemic also changes significantly. In general, the distribution of the seriously affected cities in Fig. 15(a) becomes more divergent, which may be resulted from the larger amount of direct flights and trains departing from and arriving at Beijing compared with Wuhan.

In particular, the 11 cities in Hebei province surrounding Beijing suffer the highest attack because of the regional dependency linked by busy road transportation. Tianjin and Shijiazhuang would be the cities with the highest number of COVID-19 cases among the 11 cities. Meanwhile, Xi'an, Chengdu, Guiyang, Changsha, Chongqing, and Hangzhou are the cities with the worst situation outside of Hebei, which is most likely driven by their close economic and social connections with Beijing through air and rail. Moreover, cities of Inner Mongolia and Shandong province as well as the three provinces in northeast China also have large amounts of infectious cases because over 40% of the workers in Beijing are from these regions. In summary, the epicenter switch from Wuhan to Beijing may cause a wider virus transmission and a much more serious pandemic. In other words, the location of the epicenter has significant effects on the pandemic distribution, so we can conclude that stricter restrictions need to be applied if the epicenter has higher accessibility on the transportation networks, especially on the air and rail ones which spreads the virus quickly and widely.

8. Conclusion

In this paper, we propose a novel model called transport proximity deep neural network weighted regression (TPDNNWR). This model is an improvement over the existing approaches in modelling spatial heterogeneity for disease spreading, as it incorporates a specific consideration of the inter-city transport connection in the deep neural network algorithm. In other words, the model combines the relevance of transport proximity in human movement and the excellent estimation accuracy of deep neural network. The TPDNNWR has been proved to have higher prediction accuracy than the classical GWR, and its structure is more suitable to model the heterogeneous epidemic propagation than simple F-DNN.

We further apply this model to investigate the effects of the transportation network on the heterogenous propagation of COVID-19 in China. We find that the networks of different transport modes indeed significantly affect the transmission and the distribution of the pandemic. In particular, the virus propagation through the air transportation network shows punctate distribution, and the susceptibility of hub airport cities is particularly high, even if they do not have direct flight connections with the epicenter. Meanwhile, the transmission pattern through the rail network is less divergent and only prevalent via the busy routes going through the epicenter. Further down, the impact of the road network is the most localized with the shortest transmission radius.

Other than theoretical contributions, this model can also be useful in facilitating policy making and the subsequent pandemic containment. We have analyzed two scenarios to illustrate the usefulness of the model in helping the containment of the further spread of COVID-19. First, our analysis points out that the most effective way to prevent the virus from spreading quickly and extensively would be to control the routes linked to the epicenter at the beginning of the pandemic. But if the virus has been widely spread, setting restrictions on hub cities would be much more efficient than imposing the same travel ban across the whole country. Bearing this in mind, the model that we propose in this paper has the potential to be utilized to help reduce the cost of controlling the spread of coronavirus and future epidemic/pandemic. Second, we have also shown that a comprehensive consideration of the epicenter location is necessary and helpful for disease control, suggesting that the restriction level on the epicenter should be proposed considering its importance on the transportation networks. It should be noted that the model can be directly applied to run policy discussion related to the COVID-19 pandemic as long as the characteristics of the coronavirus stays unchanged. However, its contribution is beyond the COVID-19 pandemic. In particular, the methodology can be applied to analyze and predict the spread of future epidemics with relevant data being fed into the model.

There are some limitations of our study, mainly associated with the scale and quality of the data. Because of the limited data size, there may exist overfitting in the model training although we have tried our best to eliminate it. The accuracy of the model prediction is expected to improve if a more detailed dataset (e.g., with information regarding the specific modal split for each city) is available. Moreover, the better the quality of the dataset, the more specific and relevant the policy suggestions would be. Furthermore, considering the shortcomings of machine learning methods in explanation, a comparison between GWR and the spatial regression model can be a future research direction. Meanwhile, in order to predict the temporal-spatial pandemic distribution, our future work should further incorporate the time dimension into modelling.

CRediT authorship contribution statement

Jing Lu: Conceptualization, Methodology, Formal analysis, Writing - original draft, Writing - review & editing. **Anrong Lin:** Formal analysis, Writing - original draft, Writing - review & editing. **Changmin Jiang:** Writing - original draft, Writing - review & editing, Validation, Project administration. **Anming Zhang:** Writing - original draft, Writing - review & editing, Supervision. **Zhongzhen Yang:** Writing - original draft, Writing - review & editing, Supervision.

Acknowledgement

Financial supports from the National Natural Science Foundation of China (Grants: 71701094, 71803131) and the Social Sciences and Humanities Research Council of Canada (SSHRC 435-2017-0728, 430-2019-00725) are gratefully acknowledged.

Appendix A

City name in Figs. 6, 7, 8 and 10:

1. Beijing 2. Shanghai 3. Tianjin 4. Chongqing 5. Taiwan 6. Hongkong 7. Macao 8. Shijiazhuang 9. Tangshan 10. Qinhungdao 11. Handan 12. Xingtai 13. Baoding 14. Zhangjiakou 15. Chengde 16. Cangzhou 17. Langfang 18. Hengshui 19. Taiyuan 20. Datong 21. Yangquan 22. Changzhi 23. Jinchen 24. Shuozhou 25. Jinzhong 26. Yuncheng 27. Xinzhou 28. Linfen 29. Lvliang 30. Huhehot 31. Baotou 32. Wuhai 33. Chifeng 34. Tongliao 35. Erdos 36. HulunBuir 37. BayanNur 38. Ulanqab 39. Hinggan 40. Xilingol 41. Alxa 42. Shenyang 43. Dalian 44. Anshan 45. Fushun 46. Benxi 47. Dandong 48. Jinzhou 49. Yingkou 50. Fuxin 51. Liaoyang 52. Panjin 53. Tieling 54. Chaoyang 55. Huludao 56. Changchun 57. Jilin 58. Siping 59. Liaoyuan 60. Tonghua 61. Baishan 62. Songyuan 63. Baicheng 64. Yanbian 65. Harbin 66. Qiqihar 67. Jixi 68. Hegang 69. Shuangyashan 70. Daqing 71. Yichun 72. Kiamusze 73. Qitaihe 74. Mudanjiang 75. Heihe 76. Suihua 77. Greater Khingan 78. Nanjing 79. Wuxi 80. Xuzhou 81. Changzhou 82. Suzhou 83. Nantong 84. Lianyungang 85. Huaian 86. Yancheng 87. Yangzhou 88. Zhenjiang 89. Taizhou 90. Suqian 91. Hangzhou 92. Ningbo 93. Wenzhou 94. Jiaxing 95. Huzhou 96. Shaoxing 97. Jinhua 98. Quzhou 99. Zhoushan 100. Taizhou 101. Lishui 102. Hefei 103. Wuhu 104. Bengbu 105. Huainan 106. Maanshan 107. Huaibei 108. Tongling 109. Anqing 110. Huangshan 111. Chuzhou 112. Fuyang 113. Suzhou 114. LuAn 115. Bozhou 116. Chizhou 117. Xuancheng 118. Fuzhou 119. Xiamen 120. Putian 121. Sanming 122. Quanzhou 123. Zhangzhou 124. Nanping 125. Longyan 126. Ningde 127. Nanchang 128. Jingdezhen 129. Pingxiang 130. Jiujiang 131. Xinyu 132. Yingtan 133. Ganzhou 134. Jian 135. Yichun 136. Fuzhou 137. Shangrao 138. Jinan 139. Qingdao 140. Zibo 141. Zaozhuang 142. Dongying 143. Yantai 144. Weifang 145. Jining 146. Taian 147. Weihai 148. Rizhao 149. Linyi 150. Dezhou 151. Liaocheng 152. Bingzhou 153. Heze 154. Zhengzhou 155. Kaifeng 156. Luoyang 157. Pingdingshan 158. Anyang 159. Hebi 160. Xinxiang 161. Jiaozuo 162. Puyang 163. Xuchang 164. Luohe 165. Sanmenxia 166. Nanyang 167. Shangqiu 168. Xinyang 169. Zhoukou 170. Zhumadian 171. Wuhan 172. Huangshan 173. Shiyan 174. Yichang 175. Xiangyang 176. Ezhou 177. Jingmen 178. Xiaogan 179. Jingzhou 180. Huanggang 181. Xianning 182. Suizhou 183. Enshi 184. Changsha 185. Zhuzhou 186. Xiangtan 187. Hengyang 188. Shaoyang 189. Yueyang 190. Changde 191. Zhangjiajie 192. Yiyang 193. Chenzhou 194. Yongzhou 195. Huaihua 196. Loudi 197. Xiangxi 198. Guangzhou 199. Shaoguan 200. Shenzhen 201. Zhuhai 202. Shantou 203. Foshan 204. Jiangmen 205. Zhanjiang 206. Maoming 207. Zhaoqing 208. Huizhou 209. Meizhou 210. Shanwei 211. Heyuan 212. Yangjiang 213. Qingyuan 214. Dongguan 215. Zhongshan 216. Chaozhou 217. Jieyang 218. Yunfu 219. Nanning 220. Liuzhou 221. Guizhou 222. Wuzhou 223. Beihai 224. Fangchenggang 225. Qinzhou 226. Guigang 227. Yulin 228. Baise 229. Hezhou 230. Hechi 231. Laibin 232. Chongzuo 233. Haikou 234. Sanya 235. Sansha 236. Danzhou 237. Chengdu 238. Zigong 239. Panzhihua 240. Luzhou 241. Deyang 242. Mianyang 243. Guangyuan 244. Suining 245. Neijiang 246. Leshan 247. Nanchong 248. Meishan 249. Yibin 250. Guangan 251. Dazhou 252. Yaan 253. Bazhong 254. Ziyang 255. Ngawa 256. Ganzi 257. Liangshan 258. Guiyang 259. Liupanshui 260. Zunyi 261. Anshun 262. Bijie 263. Tongren 264. Qinxinan 265. Qindongnan 266. Qinnan 267. Kunming 268. Qujing 269. Yuxi 270. Baoshan 271. Zhaotong 272. Lijiang 273. Puer 274. Lincang 275. Chuxiong 276. Honghe 277. Wenshan 278. Xishuangbanna 279. Dali 280. Dehong 281. Nujiang 282. Diqing 283. Lhasa 284. Shigatse 285. Chamdo 286. Nyingchi 287. Shannan 288. Naqu 289. Ali 290. Xi'an 291. Tongchuan 292. Baoji 293. Xianyang 294. Weinan 295. Yan'an 296. Hanzhong 297. Yulin 298. Ankang 299. Shangluo 300. Lanzhou 301. Jiayuguan 302. Jinchang 303. Baiyin 304. Tianshui 305. Wuwei 306. Zhangye 307. Pingliang 308. Jiuquan 309. Qingyang 310. Dingxi 311. Longnan 312. Linxia 313. Gannan 314. Xining 315. Haidong 316. Haibei 317. Huangnan 318. Hainan 319. Guoluo 320. Yushu 321. Haixi 322. Yinchuan 323. Shizuishan 324. Wuzhong 325. Guyuan 326. Zhongwei 327. Urumqi 328. Karamay 329. Turpan 330. Hami 331. Changji 332. Bortala 333. Bayingol 334. Akesu 335. Kizilsu 336. Kashgar 337. Hotan 338. Ili 339. Tacheng 340. Aletai

Appendix B

The correlation between the independent variables is tested and the results are listed in Table i. From the table, most correlation coefficients pass the examination, but the correlation between population density and air passenger density as well as road passenger density is higher than 0.1. Taking previous research into account, population density is a very important variable affecting local transmission, so we still incorporate it into the model. Besides, the correlation between flight frequency, train frequency and road distance is tested, and all the correlation coefficients pass the examination.

Table i
Correlation examination between independent variables.

	Air passenger Density	Rail passenger Density	Road passenger density	Population density
Air passenger Density	1.00	0.08	0.10	**0.14**
	–	(0.12)	(0.06)	(0.05)
Rail passenger Density	0.08	1.00	0.10	0.01
	(0.12)	–	(0.06)	(0.81)
Road passenger density	0.10	0.10	1.00	**0.12**
	(0.06)	(0.06)	–	(0.05)
Population density	**0.14**	0.01	**0.12**	1.00
	(0.05)	(0.81)	(0.05)	–

Appendix C

The model robustness is tested here to analyze whether our constructed models follows the regularity of spatial heterogeneity, and the robustness is always measured by the change of predictive results led by tiny perturbations of inputs. Two kinds of perturbations are adopted, the random noise and the adversarial attack, the Gaussian noise (GN) ($x_{adv} = x + \sigma \times sign(\nabla_x L(y, \hat{y}))$) is used to simulate the random noise, and the gradient sign attack method (FGSM) ($x_{adv} = x + \sigma \times sign(\nabla_x L(y, \hat{y}))$) is applied to generate adversarial attack, σ is the coefficient controlling the strength of perturbations.

The results of robustness test are shown in Fig. i, as the prediction accuracy (R^2) of GWR is quite low in the research, we did not incorporate the GWR into the robustness test. When using the perturbation of GN, the variation of prediction accuracy of TPDNNWR is within the range of 0.006 which is much lower than that of F-DNN. However, under the condition of applying the perturbation of FGSM, the robustness pattern of the two models changes a lot, the TPDNNWR firstly provides smaller change of prediction accuracy than F-DNN does when the coefficient σ is less than 6, then the F-DNN is more robust than TPDNNWR when the coefficient σ is larger than 6. In the sense, the TPDNNWR can be recognized to be regular when the perturbation is small, so generally it is more robust than F-DNN. The reason may be due to the fact that the model structure of TPDNNWR can better explain the spatial heterogeneity, because TPDNNWR is constructed based on the GWR model.

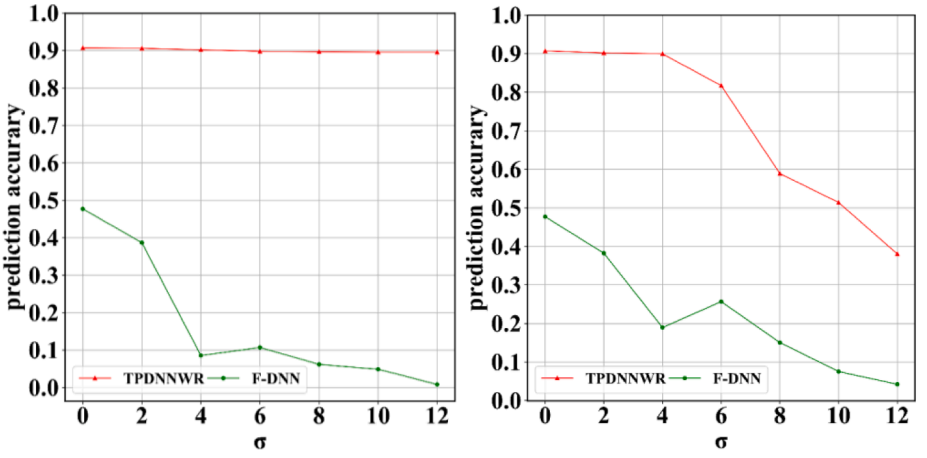


Fig. i. Prediction accuracy with two perturbations.

References

Anselin, L., 1988. Spatial econometrics: methods and models. J. Am. Stat. Assoc. 85 (411), 905–907.
Bacchetti, P., Jewell, N.P., 1991. Nonparametric estimation of the incubation period of aids based on a prevalent cohort with unknown infection times. Biometrics 47 (3), 947–960.
Benyamin, G., Mark, C., 2019. The theory behind overfitting, cross validation, regularization, bagging, and boosting: tutorial. arXiv preprint arXiv: 1905.12787.
Bivand, R., Yu, D.L., 2017. Spgwr: Geographically weighted regression. R package version 0.6-33. https://CRAN.R-project.org/package=spgwr.
Boisjoly, G., El-Geneidy, A.M., 2017. How to get there? A critical assessment of accessibility objectives and indicators in metropolitan transportation plans. Transp. Policy 55, 38–50.
Brunsdon, C., Fotheringham, A.S., Charlton, M.E., 1996. Geographically weighted regression: A method for exploring spatial nonstationarity. Geogr. Anal. 28 (4), 281–298.
Brunsdon, C., Fotheringham, A.S., Charlton, M.E., 1997. Geographical instability in linear regression modelling - a preliminary investigation. New Tech. Technol. Statist. II. 149–158.

Brockmann, D., Helbing, D., 2013. The hidden geometry of complex, network-driven contagion phenomena. Science 342 (6164), 1337–1342.
Büla, C.J., Bille, J., Glauser, M.P., 1995. An epidemic of food-borne listeriosis in western Switzerland: description of 57 cases involving adults. Clin. Infect. Dis. 20 (1), 66–72.
Cazelles, B., Hales, S., 2006. Infectious diseases, climate influences, and nonstationarity. PLoS Med. 3 (8), 33–45.
Chen, Y., Ravulaparthy, S., Deutsch, K., Dalal, P., Yoon, S.Y., Lei, T., Goulias, K.G., Pendyala, R.M., Bhat, C.R., Hu, H.-H., 2011. Development of indicators of opportunity-based accessibility. Transp. Res. Rec. 2255 (1), 58–68.
Christidis, P., Christodoulou, A., 2020. The predictive capacity of air travel patterns during the global spread of the COVID-19 pandemic: Risk, uncertainty and randomness. Int. J. Environ. Res. Public Health 17, 3356.
D'Amico, F., Baumgart, D.C., Danese, S., Peyrin-Biroulet, L., 2020. Diarrhea during COVID-19 infection: Pathogenesis, epidemiology, prevention and management. Clin. Gastroenterol. Hepatol. 18 (8), 1663–1672.
Djurhuus, S., Sten Hansen, H., Aadahl, M., Glümer, C., 2016. Building a multimodal network and determining individual accessibility by public transportation. Environ. Plan. B: Plan. Des. 43 (1), 210–227.
Dmitry, M., Arsenii, A., Dmitry, V., 2017. Variational dropout sparsifies deep neural networks. In: Proceedings of the 34th International Conference on Machine Learning, vol. 70, no. 8, pp. 2498–2507.
Dobrescu, A., Giuffrida, M.V., Tsaftaris, S.A., 2019. Understanding deep neural networks for regression in leaf counting. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 2600–2608.
Duan, Y., Lv, Y., Liu, Y., Wang, F., 2016. An efficient realization of deep learning for traffic data imputation. Transp. Res. Part C: Emerg. Technol. 72 (11), 168–181.
Dziauddin, M.F., 2019. Estimating land value uplift around light rail transit stations in Greater Kuala Lumpur: An empirical study based on geographically weighted regression (GWR). Res. Transp. Econ. 74 (5), 10–20.
Fan, X., Ying, L., 2005. Exploratory spatial data analysis of SARS epidemic in China. Adv. Earth Sci. 20 (3), 282–291.
Fotheringham, A.S., Brunsdon, C., Charlton, M., 2002. Geographically weighted regression: The analysis of spatially varying relationships. Geogr. Anal. 35 (3), 272–275.
Fotheringham, A.S., Charlton, M., Brunsdon, C., 1996. The geography of parameter space: An investigation of spatial non-stationarity. Int. J. Geogr. Inform. Syst. 10 (5), 605–627.
Fotheringham, A.S., Charlton, M., Brunsdon, C., 1998. Geographically weighted regression: A natural evolution of the expansion method for spatial data analysis. Environ. Plan. A 30 (11), 1905–1927.
Fuentes, M., 2002. Interpolation of nonstationary air pollution processes: A spatial spectral approach. Statist. Model. 2 (4), 281–298.
Fuleky, P., Zhao, Q., Bonham, C., 2014. Estimating demand elasticities in non-stationary panels: the case of Hawaii tourism. Ann. Tourism Res. 44 (1), 131–142.
Gastaldi, M., Gecchele, G., Rossi, R., 2014. Estimation of annual average daily traffic from one-week traffic counts. A combined ANN-Fuzzy approach. Transp. Res. Part C 47 (1), 86–99.
Ghosal, S., Sinha, B., Sengupta, S., Majumder, M., 2020. Frequency of testing for COVID 19 infection and the presence of higher number of available beds per country predict outcomes with the infection, not the GDP of the country - A descriptive statistical analysis. medRxiv preprint.
Gong, P., Liang, S., Carlton, E.J., Jiang, Q., Wu, J., Wang, L., Remais, J.V., 2012. Urbanisation and health in China. Lancet 379 (9818), 843–852.
Goodfellow, I.J., Shlens, J., Szegedy, C., 2014. Explaining and Harnessing Adversarial Examples. arXiv preprint arXiv: 1412.6572.
Gulland, F., Fox, M., 1992. Epidemiology of nematode infections of Soay sheep (Ovis aries L.) on St Kilda. Parasitology 105 (3), 481–492.
He, K., Zhang, X., Ren, S., Sun, J., 2015. Delving deep into rectifiers: surpassing human-level performance on imagenet classification. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), pp. 1026–1034.
Horni, A., Nagel, K., Axhausen, K.W. (Eds.). (2016). The multi-agent transport simulation MATSim. Ubiquity Press.
Hsu, D., Babiker, Z.O.E., 2019. Geographical pattern of infectious diseases and infection prevention for travellers. Tutorial Top. Infect. Combin. Infect. Train. Program. 447–459.
Hu, M., Lin, H., Wang, J., Xu, C., Lai, S., 2020. The risk of covid-19 transmission in train passengers: an epidemiological and modelling study. Clin. Infect. Dis. 7, 1–20.
Hubbard, D.P., Kitchin, D.R., Valentine, G., 2010. Key thinkers on space and place. SAGE, USA.
Hubei Provincial People's Government, 2020. Available from http://www.hubei.gov.cn/.
Jia, J., Lu, X., Yuan, Y., Xu, G., Christakis, N., 2020. Population flow drives spatio-temporal distribution of COVID-19 in China. Nature 582, 389–394.
Jones, K.E., Patel, N., Levy, M., Storeygard, A., Balk, D., Gittleman, J.L., Daszak, P., 2008. Global trends in emerging infectious diseases. Nature 451 (7181), 990–993.
Kenneson, A., Beltranayala, E., Borborcordova, M.J., Polhemus, M.E., Ryan, S.J., Endy, T.P., Stewartibarra, A.M., 2017. Social-ecological factors and preventive actions decrease the risk of dengue infection at the household-level: Results from a prospective dengue surveillance study in Machala, Ecuador. PLOS Neglected Trop. Dis. 11 (12), 1–19.
Koenig, J.G., 1980. Indicators of urban accessibility: theory and application. Transportation 9 (2), 145–172.
Kraemer, M.U., Hay, S.I., Pigott, D.M., Smith, D.L., Wint, G.R., Golding, N., 2016. Progress and challenges in infectious disease cartography. Trends Parasitol. 32 (1), 19–29.
Kurakin, A., Goodfellow, I., Bengio, S., 2016. Adversarial machine learning at scale. arXiv preprint arXiv:1611.01236.
Küchenhoff, H., Guenther, F., Höhle, M., Bender, A., 2020. Analysis of the early COVID-19 epidemic curve in Germany by regression models with change points. medRxiv preprint.
Leung, Y., Mei, C., Zhang, W., 2000. Statistical tests for spatial nonstationarity based on the geographically weighted regression model. Environ. Plan. A 32 (1), 9–32.
Lin, C., Wen, T., 2011. Using geographically weighted regression (GWR) to explore spatial varying relationships of immature mosquitoes and human densities with the incidence of dengue. Int. J. Environ. Res. Public Health 8 (7), 2798–2815.
Lipner, E.M., Knox, D., French, J.P., Rudman, J., Strong, M.J., Crooks, J., 2017. A geospatial epidemiologic analysis of nontuberculous mycobacterial infection: An ecological study in Colorado. Ann. Am. Thoracic Soc. 14 (10), 1523–1532.
Litman, T., 2016. Evaluating accessibility for transportation planning. Mobility 6, 2–64.
Liu, X., Di, X., 2020. Tanhexp: a smooth activation function with high convergence speed for lightweight neural networks. IET Comput. Vision 1–7.
Liu, X., Jiang, B., Gu, W., Liu, Q., 2011. Temporal trend and climate factors of hemorrhagic fever with renal syndrome epidemic in Shenyang city, China. BMC Infectious Dis. 11 (1), 1–6.
Lowe, J., Gauger, P., Harmon, K., Zhang, J., Connor, J., Yeske, P., Loula, T., Levis, I., Dufresne, L., Main, R., 2014. Role of transportation in spread of porcine epidemic diarrhea virus infection, United States. Emerg. Infectious Dis. 20 (5), 872–874.
Lu, B., Charlton, M., Brunsdon, C., Harris, P., 2016. The Minkowski approach for choosing the distance metric in geographically weighted regression. Int. J. Geogr. Inform. Sci. 30 (2), 351–368.
Ma, X., Dai, Z., He, Z., Ma, J., Wang, Y., Wang, Y., 2017. Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction. Sensors 17 (4), 1–16.
Macintyre, S., Ellaway, A., 2000. Ecological approaches: rediscovering the role of the physical and social environment, fifth ed. Oxford University Press, New York.
Medina, R.M., Hepner, G.F., 2011. Advancing the understanding of sociospatial dependencies in terrorist networks. Trans. GIS 15 (5), 577–597.
Mohammadinia, A., Alimohammadi, A., Saeidian, B., 2017. Efficiency of geographically weighted regression in modeling human leptospirosis based on environmental factors in Gilan province, Iran. Geosciences 7 (12), 1–12.
Müller, S.A., Balmer, M., Neumann, A., Nagel, K., 2020. Mobility traces and spreading of COVID-19. medRxiv preprint.
Nassir, N., Hickman, M., Malekzadeh, A., Irannezhad, E., 2016. A utility-based travel impedance measure for public transit network accessibility. Transp. Res. Part A: Policy Pract. 88, 26–39.
Ng, A.Y., 2004. Feature selection, L 1 vs. L 2 regularization, and rotational invariance. In: Proceedings of the twenty-first international conference on Machine learning, p. 78..

Panek, E., 2019. Non-stationary gale economy with limited technology and multilane turnpike. “weak”, “strong” and “very strong” turnpike theorem. Econ. Bus. Rev. 5 (2), 3–23.
Pu, Z., Li, Z., Ash, J., Zhu, W., Wang, Y., 2017. Evaluation of spatial heterogeneity in the sensitivity of on-street parking occupancy to price change. Transp. Res. Part C-Emerg. Technol. 77 (4), 67–79.
Rashed, E.A., Kodera, S., Gomez-Tames, J., Hirata, A., 2020. Influence of absolute humidity, temperature and population density on COVID-19 spread and decay durations: Multi-prefecture study in Japan. Int. J. Environ. Res. Public Health 17 (15), 5354–5359.
Reimering, S., Munoz, S., Mchardy, A.C., 2020. Phylogeographic reconstruction using air transportation data and its application to the 2009 H1N1 influenza a pandemic. PLoS Comput. Biol. 16 (2), 16(2), 1–16.
Rokhman, N., Nuryati, Salim, M.F., Syairaji, M., Lubis, I.K., 2019. The implementation of spatial weighted regression on detecting the risk factors of malaria incidences in Kulonprogo district. KnE Soc. Sci. 3(23), 148–161.
Sarlas, G., Axhausen, K.W., 2015. Localized speed prediction with the use of spatial simultaneous autoregressive models. In: Transportation Research Board 94th Annual Meeting, Washington, D.C.
Salze, P., Banos, A., Oppert, J.M., Charreire, H., Casey, R., Simon, C., Weber, C., 2011. Estimating spatial accessibility to facilities on the regional scale: an extended commuting-based interaction potential model. Int. J. Health Geographics 10 (1), 2–5.
Sifringer, B., Lurkin, V., Alahi, A., 2020. Enhancing discrete choice models with representation learning. Transp. Res. Part B: Methodol. 140, 236–261.
Sirisena, P., Noordeen, F., Kurukulasuriya, H., Romesh, T.A., Fernando, L., 2017. Effect of climatic factors and population density on the distribution of dengue in Sri Lanka: A GIS based evaluation for prediction of outbreaks. PLoS ONE 12 (1), 1–14.
Song, W., Li, Y., Hao, Z., Li, H., Wang, W., 2016. Public health in China: An environmental and socio-economic perspective. Atmos. Environ. 129 (3), 9–17.
Southern, J., 2012. Comobility: how proximity and distance travel together in locative media. Can. J. Commun. 37 (1), 34–56.
Sun, X., Wandelt, S., Zhang, A., 2021a. On the degree of synchronization between air transport connectivity and COVID-19 cases at worldwide level. Transp. Policy 105, 115–123.
Sun, X., Wandelt, S., Zheng, C., Zhang, A., 2021b. COVID-19 pandemic and air transportation: Successfully navigating the paper hurricane. J. Air Transp. Manage. 94, 102062.
Sun, Z., Zhang, H., Yang, Y., Wan, H., Wang, Y., 2020. Impacts of geographic factors and population density on the COVID-19 spreading under the lockdown policies of China. Sci. Total Environ. 746, 141–347.
Tasyurek, M., Celik, M., 2020. RNN-GWR: a geographically weighted regression approach for frequently updated data. Neurocomputing 399 (25), 258–270.
Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. Econ. Geogr. 46 (6), 234–240.
Wang, J., Chen, R., He, Z., 2019. Traffic speed prediction for urban transportation network: A path based deep learning approach. Transp. Res. Part C: Emerg. Technol. 100 (3), 372–385.
Wang, K., Jiang, C., Ng, A.K., Zhu, Z., 2020a. Air and rail connectivity patterns of major city clusters in China. Transp. Res. Part A: Policy Pract. 139, 35–53.
Wang, S., Mo, B., Zhao, J., 2020b. Deep neural networks for choice analysis: architecture design with alternative-specific utility functions. Transp. Res. Part C: Emerg. Technol. 112, 234–251.
Wang, S., Mo, B., Zhao, J., 2021. Theory-based residual neural networks: a synergy of discrete choice models and deep neural networks. Transp. Res. Part B Methodol. 146 (2082), 333–358.
Wu, S., 2019. The theory and method of geographically and temporally neural network weighted regression, Ph.D. Dissertation. Zhejiang University, Hangzhou, China.
Wu, S., Du, Z., Wang, Y., Lin, T., Zhang, F., Liu, R., 2020. Modeling spatially anisotropic nonstationary processes in coastal environments based on a directional geographically neural network weighted regression. Sci. Total Environ. 709 (3), 2–17.
Wuhan Municipal Health Commission, 2020. Available from http://wjw.wuhan.gov.cn/.
Yang, W., 2014. An extension of geographically weighted regression with flexible bandwidths. Ph.D. Dissertation. University of St Andrews, Scotland, UK.
Yi, H., Jung, H. J., Bae, S., 2017. Deep neural networks for traffic flow prediction. In: 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), Jeju, South Korea, pp. 328–331.
Zeiler, M.D., 2012. Adadelta: an adaptive learning rate method. arXiv preprint arXiv: 1212.5701.
Zhang, N., Zhao, P., Li, Y., 2019. Increased infection severity in downstream cities in infectious disease transmission and tourists surveillance analysis. J. Theor. Biol. 470 (3), 20–29.
Zhang, Y., Lin, F., Zhang, A., 2018. Gravity models in air transport research: A survey and an application. In: Blonigen, B.A., Wilson, W.W. (Eds.), Handbook of International Trade and Transportation, vol. 4. Edward Elgar, Cheltenham, pp. 141–158 (Chapter 4).
Zhang, Y., Zhang, A., 2016. Determinants of air passenger flows in China and gravity model: Deregulation, LCCs, and high-speed rail. J. Transp. Econ. Policy 50 (3), 287–303.
Zhang, Y., Zhang, A., Wang, J., 2020. Exploring the roles of high-speed train, air and coach services in the spread of COVID-19 in China. Transp. Policy 94, 34–42.
Zhao, H., Gallo, O., Frosio, I., Kautz, J., 2017. Loss functions for image restoration with neural networks. IEEE Trans. Comput. Imaging 3 (1), 47–57.
Zhou, X.N., Yang, G.J., Yang, K., 2011. The development course and development trend of Chinese space epidemiology. Chin. J. Epidemiol. 32 (9), 854–858.
Zou, H., 2020. Routine press conference on prevention and control of COVID-19. Retrieved August 3, 2020, from http://www.chinanews.com.