

# Multi-source Sensor based Urban Habitat and Resident Health Sensing:A Case Study of Wuhan,China

Yan Zhang<sup>a</sup>, Nengcheng Chen<sup>a,d</sup>, Wenying Du<sup>a,d</sup>, Yingbing Li<sup>b</sup> and Xiang Zheng<sup>c</sup>

<sup>a</sup>*State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan 430079, China*

<sup>b</sup>*School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China*

<sup>c</sup>*School of Information Management, Wuhan University, Wuhan 430079, China*

<sup>d</sup>*National Engineering Research Center for Geographic Information System, China University of Geosciences, Wuhan 430074, China*

---

## ARTICLE INFO

### Keywords:

Social sensing  
COVID-19  
Sensor  
Street View  
PLS-SEM  
GIS

---

## ABSTRACT

The COVID-19 pandemic undoubtedly has a great impact on the world economy, especially the urban economy. It is urgent to study the environmental pathogenic factors and transmission route of it. We want to discuss the relationship between the urban living environment and the number of confirmed cases at the community scale, and examine the driving forces of community infection (e.g., environment, ecology, convenience, livability, and population density). Besides, we hope that our research will help make our cities more inclusive, safe, resilient, and sustainable. 650 communities with confirmed COVID-19 cases in Wuhan were selected as the research objects. We utilize deep learning semantic segmentation technology to calculate the Visible Green Index (VGI) and Sky View Factor (SVF) of street view and use Partial Least Squares Structural Equation Modeling (PLS-SEM) to study the driving forces of pandemic situation. Temperature and humidity information recorded by sensors was also used for urban sensing. We find that the large SVF has a certain inhibitory effect on the virus transmission, but contrary to our intuitive perception, higher VGI has a certain promotion effect. Also, the structural equation model constructed in this paper can explain the variance of 28.9% of the number of confirmed cases, and results (path coef.) demonstrate that residential density of community (0.517) is a major influencing factor for pandemic cases, whereas convenience of community living (0.234) strongly influence it. Communities with good suitability of community human settlement (e.g., construction time, price) are safer in the face of pandemic events. Does the influence of SVF and VGI on the results of the pandemic situation mean that sunlight can effectively block the spread of the virus? This spatial heterogeneity in different communities is helpful for us to explore the environmental transmission route of COVID-19.

---

## 1. Introduction

Cities and metropolitan areas are powerhouses of economic growth contributing about 60 % of global GDP. However, they also account for about 70 % of global carbon emissions and over 60 % of resource use. In the past 20 years, China has carried out large-scale infrastructure construction and made great achievements. The urbanization rate has risen from 36.22% in 2000 to 60.6% in 2019. Considering China's population base of one billion, about 400 million people have migrated from rural areas to cities. Correspondingly, a large number of houses and communities have been built in the cities (CSY, 2018). This rapid urbanization process has led to the birth of a group of super cities with a scale of tens of millions of people and a construction area of hundreds of square kilometers.

At the same time, rapid development of industrialization, urbanization and modernization in China is resulting in a growing number of slum dwellers (Wang et al., 2019a), inadequate and overburdened infrastructure and services (such as waste collection and water and sanitation systems, roads and transport), worsening air pollution and unplanned urban sprawl. Several literatures have argued that urban infrastructure, air pollution, population mobility, lifestyle, and the level of health care resource coverage during urbanization have an impact on public health systems and human health (Liu et al., 2017). In addition, high population density and the lack of green buffers between communities

\*This research was supported by the National Key R&D Program (no.2018YFB2100500), the Fundamental Research Funds for the Central Universities (no. 2042020kf0011), National Natural Science Foundation of China program (no.41890822) and Creative Research Groups of Natural Science Foundation of Hubei Province of China (no.2016CFA003).

\*\*The numerical calculations in this paper have been done on the supercomputing system in the Supercomputing Center of Wuhan University.

\*Corresponding author

ORCID(s): 0000-0002-3202-5128 (W. Du)

can contribute to the spread of epidemics (Hunter, 2017).The COVID-19 pandemic has brought enormous economic damage, especially to fragile human settlements and densely populated urban areas, where overcrowding also makes it difficult to follow recommended measures such as social distancing and self-isolation.

It is currently believed that environmental temperature and humidity have a significant effect on virus transmission. However, most studies have generally analyzed the disease on a large spatial scale (such as city and country) rather than at the community scale, and due to regional differences, even opposite conclusions may be reached (Kerr et al., 2021). In addition, the level of environmental management and air pollution can also play a key role in the spread of the pandemic (Espejo et al., 2020). Communities with sunny, easily ventilated built environments are more likely to interrupt the SARS-CoV2 pandemic virus transmission (Jayaweera et al., 2020). The Sustainable Development Goals 2030<sup>1</sup> are a call for action by all countries poor, rich and middle-income to promote prosperity and make cities inclusive, safe, resilient and sustainable while protecting the planet. It means that in the post-COVID-19 pandemic era, future community planning must pay more attention to the construction of ecological and environmental services.

We try to study the relationship between urban living environment and pandemic, using multi-source sensing data represented by urban street view (a kind of open, voluntary upload, wide coverage of geo-tagged image dataset) to discuss from the community scale. Urban street view is proved to be a potential data source covering a wide area and is a good environmental perception solution. It has been applied to neighborhood safety (Wang et al., 2019b; Rundle et al., 2011), sensing the level of urban greening (Ye et al., 2019), neighborhood planning and street design (Wu et al., 2020a), resident emotion judgment and environmental health research (Huang et al., 2020).

Urban streetscape establishes the mapping relationship between urban images and geographical entities, which provides a new opportunity to solve the large-scale derivation of urban sensing (Yao et al., 2019). It helps us understand the surrounding environment and conduct geographic computing (Liang et al., 2020a). There are a lot of studies on the impact of residential environment health, such as associating geriatric depression with street view green using street view data (Helbich et al., 2019), the positive effect of Normalized Difference Vegetation Index (NDVI) on mental wellbeing (Liu et al., 2019a), the correlation between greenness exposure and depression in developing countries (Liu et al., 2019b), studying the social inequality among different communities (Zhou et al., 2019) and urban sensing for multiple emotional dimensions (security, liveliness, beauty, wealth, boredom, depression) (Zhang et al., 2020). However, a large number of researchers have focused on the research direction of mental illness or mental state in the urban environment (Kang et al., 2020), while few have studied the difference in the degree of transmission of epidemic diseases (SARS, COVID-19, MERS) between different communities (e.g.slums vs. rich area) and the coupling effect with the human habitat (Vazquez-Prokopec et al., 2010).

The subject of our analysis is the community and the dependent variable is the number of confirmed pandemic diagnoses, aiming to reveal their heterogeneous nature. The model indicators include physical indicators, socio-economic indicators, spatial location indicators, ecological environment indicators, and population density indicators. Research data related to the natural, built and social environment based on social sensing, sensor networks and remote sensing images (Wang et al., 2019c; Liu et al., 2015).

The advantages of this paper are: in addition to the urban street view, we also utilize a variety of evaluation indicators proving to be related to virus transmission (e.g., climatic conditions, the surrounding environment of community) (Qu et al., 2020; Kraemer et al., 2020), and establish a relatively complete human centered risk exposure assessment system from the view of Geographic Information System (GIS), evaluate the safety of the community using the observable variable value. We consider the process of exposed model as an evolving coupled confirmed cases and human settlements and system (Yawson, 2020). Based on the established index evaluation system, a case study of COVID-19 was carried out by using Structural Equation Modeling (SEM).

The research questions this paper answers are as follows:

(RQ1):Are there differences in the risk resilience of communities with different physical attributes in the face of public health events?

(RQ2):What does the interaction between the different community indicators look like? What is the explanatory power of the model for COVID-19 pandemic confirmed cases?

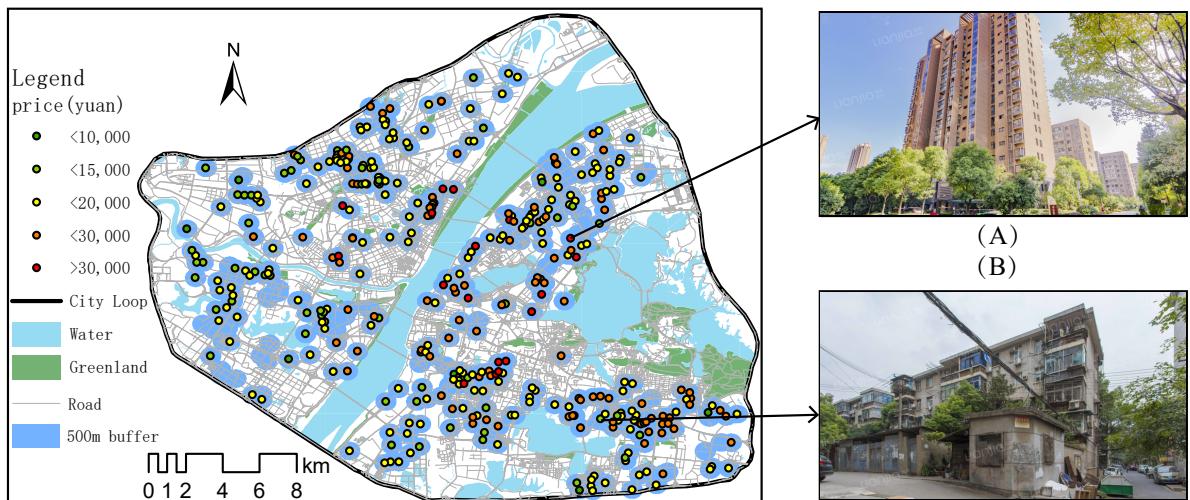
(RQ3):What are the main environmental drivers of the pandemic? Beyond that, is there a significant effect of temperature and humidity?

<sup>1</sup>[www.un.org/sustainabledevelopment](http://www.un.org/sustainabledevelopment)

## 2. Study area and Methodology

### 2.1. Study area and data

Wuhan is the city most seriously affected by the COVID-19 pandemic in China. We take the main urban area within the Third Ring Road of Wuhan City as the research area. The communities with confirmed cases and their property of house prices are shown in the Figure 1. Different communities have different living environment. The buildings were completed over a wide span of time, including modern residential areas built in recent years, and state-owned enterprise employee communities built in the last century, which show different states in the urban streetscape.



**Figure 1:** Study area profile and different community streetscape.

The study data were as follows: levels of Visible Green Index(VGI), Sky View Factor(SVF) and the number of confirmed cases in the community. We base on the application program interface (API) provided by the largest online map service provider in China. By entering the latitude and longitude with the buffer radius, users can obtain the number of various points of interest (POIs) around the address. The population density data is from the Gridded Population of the World (GPW)<sup>2</sup>. The number of people diagnosed in the community comes from the data collected by the research team from various communities and the official statistics of the local health committee. The urban temperature and humidity sensing data are from the sensors deployed by our research group. Training set of street scene segmentation is from cityscapes dataset, remote sensing data is from Landsat Image, and administrative boundary is from Open Street Map(OSM). We capture a total of 10,294 street view pictures with a size of 710px × 400px from 650 communities in Wuhan from China's largest real estate trading platform. In addition, the number of houses, the number of buildings and the cost of property in communities were also included. The data and sources used in this paper are shown in Table 1.

### 2.2. Overall Methodology

Figure 2 shows the connections among Section 2.3 and Section 2.4, street view segmentation results, POI data, sensor monitoring record, and they together drive the next part of the model analysis. At the community scale, multiple linear regression and correlation analysis were carried out on the observation variables obtained by various sensing methods, which was applied to explore the correlation between independent variables preliminarily and prepare for the determination of the structure equation. We employ path coefficient and factor loadings to study the main driving force and driving effect of environment on pandemic situation. Meanwhile, the environmental temperature & humidity and the number of patients diagnosed & rehabilitated were discussed, as the confirmation and supplement of the structural model.

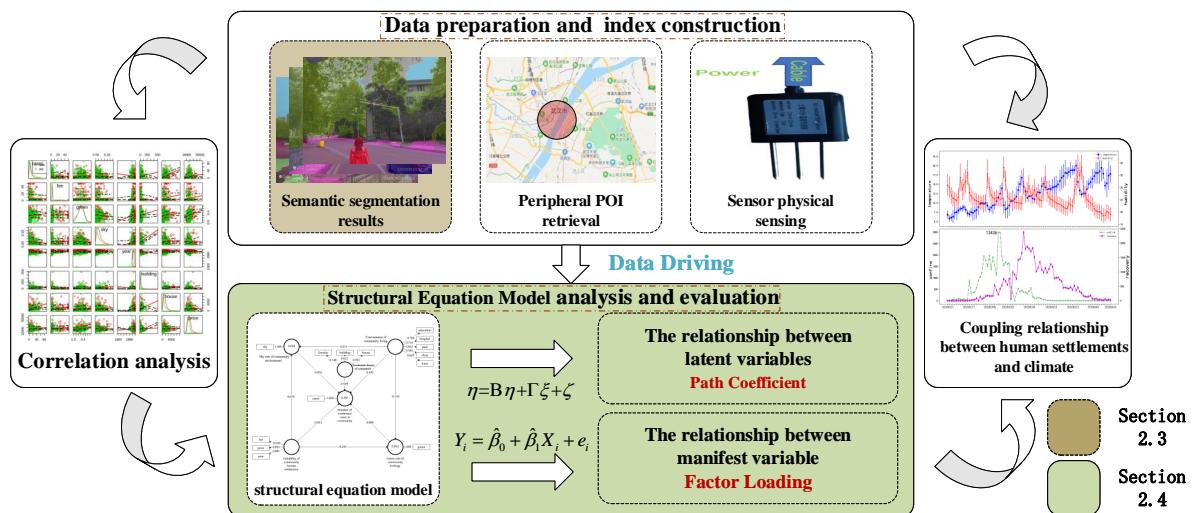
<sup>2</sup><http://sedac.ciesin.columbia.edu/gpw>

**Table 1**

Data used in this paper and data sources

Data Name		Data Source
POI data	●	Baidu Map( <a href="http://www.lbsyun.baidu.com">www.lbsyun.baidu.com</a> )
Population Density	●	Gridded Population of the World( <a href="http://sedac.ciesin.columbia.edu/gpw">sedac.ciesin.columbia.edu/gpw</a> )
Number of cases	●	health committee( <a href="http://wjw.wuhan.gov.cn">wjw.wuhan.gov.cn</a> )
Temperature and humidity	●	Sensor(sensors deployed by our research group)
cityscapes-dataset	●	Cityscapes( <a href="http://www.cityscapes-dataset.com">www.cityscapes-dataset.com</a> )
Landsat Remote sensing image	●	USGS( <a href="http://glovis.usgs.gov">glovis.usgs.gov</a> )
Administrative boundary	●	OSM( <a href="http://www.openstreetmap.org">www.openstreetmap.org</a> )
Streetscape Image and community attribute	●	<a href="http://www.wh.lianjia.com">lianjia(<a href="http://www.wh.lianjia.com">www.wh.lianjia.com</a>)</a>

Acquisition method: social sensing ● ; remote sensing ● ; sensor network ●; Other Source ●

**Figure 2:** Method introduction and chapter connection.

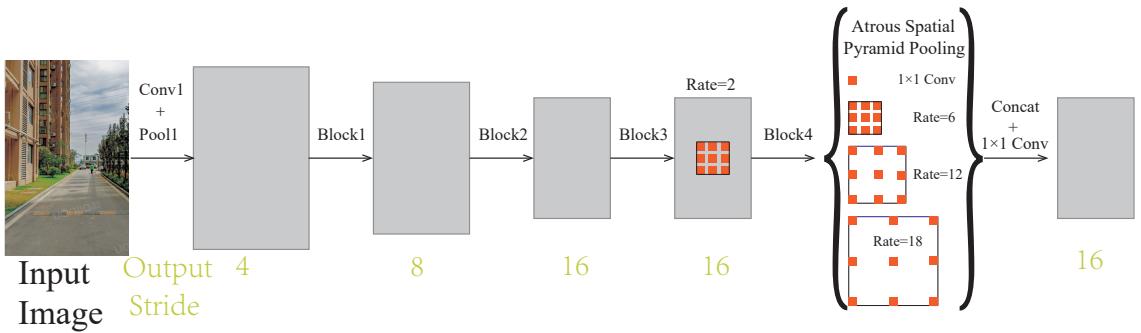
### 2.3. Deeplabv3-ResNet101 for perception computing and fusion

Facing tens of thousands of street view pictures, it is a challenge to accurately extract SVF and VGI. Plants will show different colors in different seasons and it is difficult for us to define the color range of sky and green. Besides, we will encounter all kinds of weather, plants and complex district environment. To solve this puzzle, we use deep learning semantic segmentation technology which has stronger segmentation accuracy, better robustness and can accurately identify the entity type (Gong et al., 2018).

We utilize the Deeplabv3-ResNet101 as semantic segmentation algorithm to calculate the street view pictures, which could classify each pixel of the image, and the contour of the object is accurately outlined. Compared with the traditional Fully Convolutional Networks (FCN), it has higher classification accuracy (Chen et al., 2017).

#### 2.3.1. Deeplabv3 module

Deeplabv3 framework combine Deep Convolutional Neural Networks(DCNNs) with Dense Conditional Random Fields(DenseCRFs) and the serial and parallel convolution modules with holes are designed. As we can see from Figure 3, there are a few parallel atrous convolutions with different rates on top of the feature map in Atrous Spatial Pyramid Pooling (ASPP) modular, which could expand the sampling range to obtain better field of view and effectively capture multi-scale information. Output Stride is defined as the ratio of input image spatial resolution to final out. Besides, we adapt the ImageNet-pretrained ResNet (He et al., 2016) to the semantic segmentation by applying atrous convolution to extract dense features.



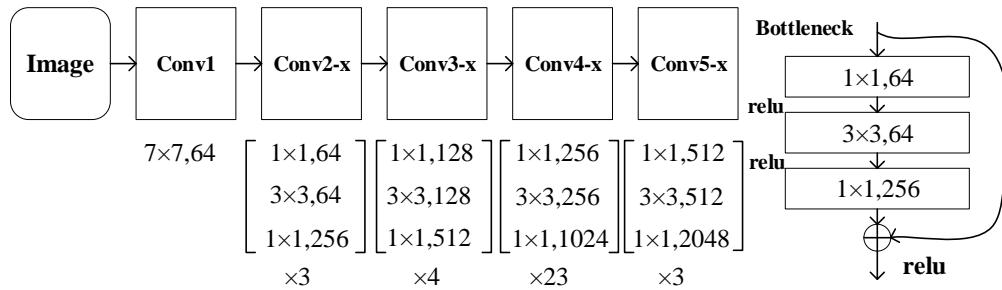
**Figure 3:** Atrous Spatial Pyramid Pooling modular.

### 2.3.2. ResNet101 module

The 101 in ResNet101 model refers to one hundred and one layers of convolution layer and full connection layer. The basic structure of network model is five convolution parts as shown in Figure 4: Conv1 conclude a  $7 \times 7$  convolution kernel; Conv2-X consists of three groups of convolution kernels, each of which contains 64  $1 \times 1$ , 64  $3 \times 3$  and 256  $1 \times 1$  convolution kernels. Conv3-X was divided into four groups, each group consists of 128  $1 \times 1$ , 128  $3 \times 3$  and 512  $1 \times 1$  convolution kernels. Conv4-X consists of 23 sets of convolution kernels with 256  $1 \times 1$ , 256  $3 \times 3$  and 1,024  $1 \times 1$  in each group; Conv5-X consists of three groups convolution kernels, each of which contains 512  $1 \times 1$ , 512  $3 \times 3$  and 2,048  $1 \times 1$  convolution kernels. The mapping formula of convolution kernel features is as follows:

$$f_k^l = w_k^{lT} x_{i,j}^l + b_k^l \quad (1)$$

Among them,  $f_k^l$  is the feature of layer  $l$  after convolution kernel of layer  $k$ ,  $w_k^{lT}$  is the weight of layer  $k$ ,  $x_{i,j}^l$  is the features of layer  $l$  with size  $(i, j)$ , and  $b_k^l$  is the offset of layer  $k$ . In addition, ResNet101 adopts bottleneck structure to enhance learning ability, and take identity shortcut connection to realize layer skipping connection, which speeds up the convergence speed and solves the problems of model gradient disappearance and gradient explosion (Szegedy et al., 2016).

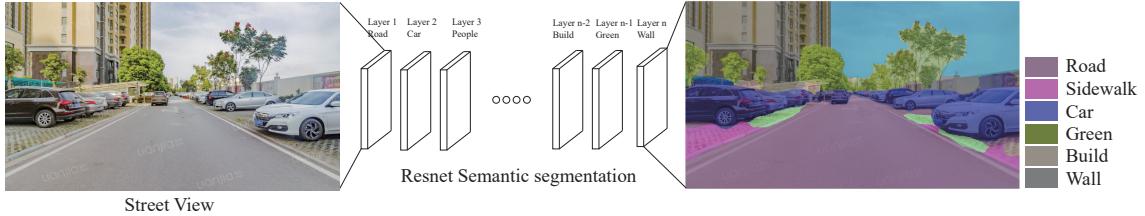


**Figure 4:** ResNet101 module structure.

### 2.3.3. Realization of segmentation technology

We take the cityscapes dataset, which is a large scale dataset that contains a diverse set of stereo video sequences recorded in street scenes from 50 different cities, with high quality pixel-level annotations of 5,000 frames in addition to a larger set of 20,000 weakly annotated frames as a training set. Cityscapes dataset divides city images into eight categories (flat, human, vehicle, construction, object, green, sky, void) and dozens of small categories. Taking into account the main factors affecting human settlements, we mainly use green and sky index to extract the VGI (the pixel ratio of plants in the street view) and the SVF (the pixel ratio of sky in the street view) (Wang et al., 2020; Zhang et al., 2018). We adopt Amazon Mxnet deep learning framework to calculate (Chen et al., 2015). After semantic

segmentation, a street picture is decomposed into various layers. We could calculate the proportion of green layer pixels and sky layer pixels in total picture pixels. Since there are many street view pictures in a community, we use the median to represent the SVF and VGI attributes of the community. The segmentation effect of street view is shown in the Figure 5. In China's complex urban environment, Deeplab-ResNet101 model could achieve satisfactory results at the community scale after the experimental test.



**Figure 5:** The segmentation effect of Deeplabv3.

## 2.4. PLS-SEM based pandemic-environment coupling structure

We first employ the Ordinary Least Squares (OLS) (Stone and Brooks, 1990) model to fit the number of confirmed cases in the community and the observation formula is as follows:

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + e_i, (i = 1, 2, \dots, T) \quad (2)$$

$\hat{\beta}$  is the unknown parameter,  $e$  is the error term,  $X_i$  is the observation value,  $Y_i$  is the forecast value. However, OLS still has some problems: multiple dependent or output variables are not allowed, the predictor is assumed to have no measurement error and multicollinearity among predictors may hinder the interpretation of the results.

SEM is a system of equations in form, including random variable and structural parameters. Random variable includes three types of indicator variables, latent variable and error variable. Indicator variable refers to the quantity that could be directly observed, also known as the manifest variable. Latent variable refers to the variable that can not be directly observed by the construction with theory or hypothesis but could be described by one or more indicator variables. For example, we can use literature, mathematics, science, geography, history and other subjects as a measure of academic performance, and employ observable indicators to describe our perceptual cognition. SEM consists of a measurement model and a structural model. The measurement model is used to indicate the connection between the manifest variable and the latent variable. It is expressed by the following formula:

$$x = \Lambda_x \xi + \delta \quad (3)$$

$$y = \Lambda_y \eta + \epsilon \quad (4)$$

Among them,  $x$  and  $y$  are exogenous and endogenous indicators (Varian and Varian, 1992).  $\delta$  and  $\epsilon$  refer to the measurement error and  $\Lambda$  refers to the relationship between the measurement index (indicator variable) and potential variables. Structural model is used to measure the relationship among latent variables.

$$\eta = B\eta + \Gamma\xi + \zeta \quad (5)$$

$\eta$  is endogenous(dependent)latent variable,  $\xi$  is exogenous(independent)latent variable,  $B$  refers to the correlation between endogenous latent variables.  $\Gamma$  refers to the exogenous latent variable affect the endogenous latent variable.  $\zeta$  is the part of the model that cannot be explained.

In order to avoid some problems of OLS method, we use Partial Least Squares Structural Equation Modeling (PLS-SEM) as a comparison (Hair et al., 2011), which is also suitable for non-normal distribution data (Hair Jr et al., 2016; Hair et al., 2011). Comparing with Covariance Based Structural Equation Modeling (CB-SEM), PL-SEM is more suitable for small sample exploratory research and have the ability to evaluate the interaction between variables effectively (Fan et al., 2016).

As shown in Table 2, this paper collates and calculates fourteen evaluation indexes of 650 communities in the study area, taking Housing price, Property fee, Construction year, Visible Green Index, Sky View Factor, population density,

**Table 2**  
Modeling indicators and explanation

Latent Variable	Manifest Variable(abbreviation)	Explanation
Number of confirmed cases in community (Orford, 2017)	Confirmed Cases(Cases)	Statistics of infected persons
Suitability of community human settlement (Wang et al., 2019c)	Housing Price(Price)	Average price per $m^2$
Residential density of community (Liang et al., 2020b)	Construction Year(Year)	Completion time of community
Convenience of community living (Diao and Ferreira Jr, 2010)	Population Density(Density)	GPW sampling
Green rate of community ecology (Houghton and Castillo-Salgado, 2020)	Community Housing (House)	Num. of household
Sky rate of community environment (Gong et al., 2018)	Community Buildings(Building)	Num. of building
	Peripheral Park POI(Park)	Num. of surrounding park
	Peripheral Shop POI(Shop)	Num. of surrounding shop
	Peripheral Transport POI(Transport)	Num. of surrounding transport
	Peripheral Hospital POI(Hospital)	Num. of surrounding hospital
	Peripheral Education POI(Education)	Num. of surrounding education
	Visible Green Index(VGI)	Here we define the proportion of the sky within the field of view
	Sky View Factor(SVF)	Here we define the proportion of the vegetation within the field of view
	Property Fee(Fee)	Property cost per $m^2$

Note1: natural condition ● ; built condition ○ ; social condition ● ; other ●

Note2: We represent this community with the median of VGI/SVF

Note3: POIs within a 500-meter buffer zone around the community are included in the statistics

Community housing, Community buildings, Confirmed cases, Number of peripheral education POI, shop POI, transport POI, hospital POI and park POI as indicator variables. According to the manifest variables, we employed four different environmental variables and two physical variables in our PLS-SEM: Suitability of community human settlement, Green rate of community ecology, Residential density of community, Number of confirmed cases in community, Sky rate of community environment and Convenience of community living.

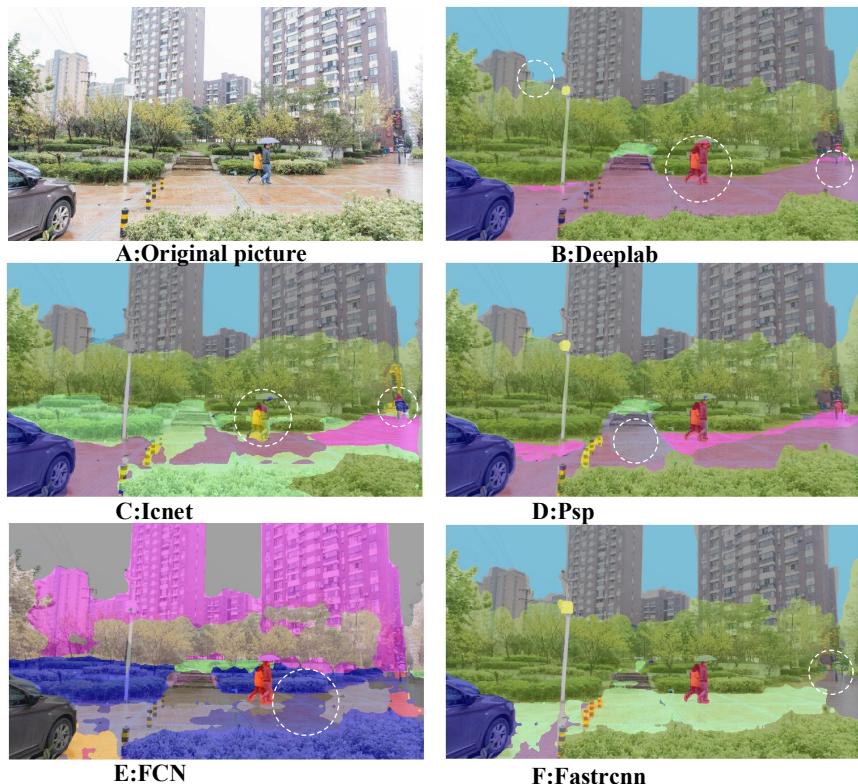
We believe that the residential density of the community has an absolute impact on the confirmed cases, and the suitability index represented by the community housing price has an impact on the VGI and SVF of the community (It is generally believed that the better the neighborhood is, the better the management and environment). Similarly, the surrounding environment of the community has an impact on the VGI/SVF of the community. We assume that all of the above latent variables have an effect on the number of confirmed cases in the community (Li et al., 2020; Guida and Carpentieri, 2020). The PLS-SEM was constructed based on this assumption according to the established evaluation indicators and we will discuss the rationality and reliability of our model in the next section.

### 3. Results and Analysis

#### 3.1. Street View Segmentation Results and Verification of satellite ground fusion

The main work in this section is the comparison of semantic segmentation algorithms, the street view computation of the study communities, and the reliability verification of the computation results based on the comparison of satellite data. We compare the Pyramid scene parsing network (Psp) (Zhao et al., 2017), ICNet (Zhao et al., 2018), Fast Region-based Convolutional Network method (FastRnn) (Girshick, 2015; Ren et al., 2015) and FCN semantic segmentation algorithms and carry on the contrast experiment. We can see from the Figure 6 that the segmentation algorithm used in this paper is better, which can effectively remove salt and pepper and noise in the community street image, and accurately extract the boundary between green plants and sky. However, Psp method misjudged the road, ICNet method and FastRnn method confused the road and green plants, and could not recognize the people in the image distance. FCN has the worst recognition effect, and the recognition result has some noise.

An example of four street image segmentation results is shown in Figure 7. For example, the model accurately separates buildings, grassland, pedestrian, sky, trees, etc. We calculate the SVF and VGI indices for all street view pictures in turn, and use the median of these two indices to classify the street view data into four categories: High SVF



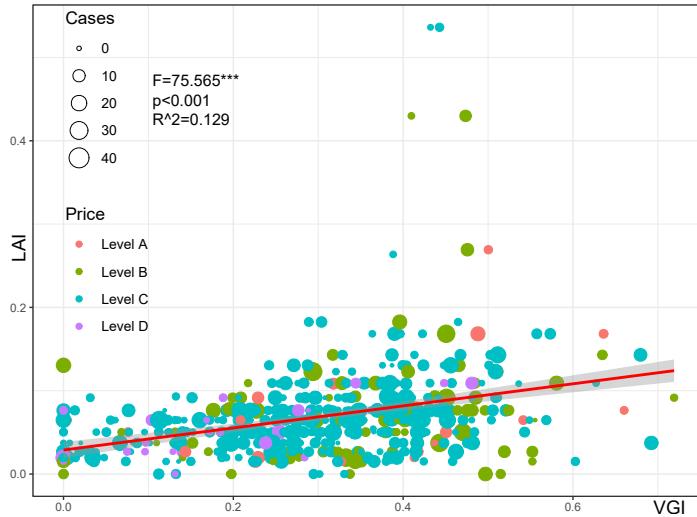
**Figure 6:** Comparison of different semantic segmentation algorithms.



**Figure 7:** Scene classification based on deep learning methods: street view image (A) and after the segmentation (B).

& High VGI (HH Class), High SVF & Low VGI (HL Class), Low SVF & High VGI (LH Class), Low SVF & Low VGI (LL Class).

To evaluate the reliability of Visible Green Index, we establish an empirical formula to calculate the leaf area index (LAI) based on the correlation between easily measured parameters such as Diameter at Breast Height (DBH), tree height, sapwood area, crown width and leaf area index. LAI refers to the multiple of the total leaf area in the unit land



**Figure 8:** Verification of satellite ground fusion (According to the house price>30,000, 20,000<price<=30,000, 10,000<price<=20,000, and price <=10,000, we construct four levels).

area. The formula is:

$$LAI = \alpha e^{\beta NDVI} \quad (6)$$

Both  $\alpha$  and  $\beta$  are empirical coefficients (Here  $\alpha$  is 0.127 and  $\beta$  is 4.104). NDVI is calculated through Landsat Remote sensing image data set.

As shown in the LAI and VGI fitting result in Figure 8, the LAI and VGI show a positive correlation, which to some extent indicates that our streetscape extraction effect is successful.

### 3.2. Correlation Analysis between Urban Residential Environment and COVID-19

#### 3.2.1. Variable Correlation Analysis

In this section, we try to explore the relationship between community environment & management level and the results of the pandemic, without considering the surrounding environment of the community. As shown in Figure 9, we examine the cases with the other seven manifest variables except the convenience of community living variable. Based on the correlation coefficient (covariance) of the number of confirmed cases and the other seven factors, it is seen that confirmed cases are significantly and positively correlated with the number of houses and buildings but negatively correlated with the SVF of community environment. There are still some interesting findings that: At the confidence level of  $p = 0.05$ , the number of confirmed cases in newly established communities is relatively few but the more expensive the community price is, the higher the number of confirmed cases is, the correlation coefficient is 0.09. However, at the confidence level of  $p = 0.01$ , the newer the community, the more houses there are, but fewer buildings, which means that new communities tend to build high-rise buildings.

What stands out in this figure is VGI is plus correlation with cases and the SVF is negatively correlated with the VGI at the confidence level of  $p = 0.05$ , which is not consistent with previous studies. A possible explanation for this might be that the high rate of greening blocks part of the sunlight. Correspondingly, the higher SVF, the more sunlight there will be and sunlight has a certain blocking effect on the spread of the pandemic. This conclusion cannot exclude some other reasons. A higher vegetation cover is usually connected with long-term established community and old trees, which may be at the old district of the city. In other words, these communities may not have necessary infrastructures and not be well-managed. Therefore, the number of cases in these communities is higher. In addition to this, there is also the interference of population density, occupation of community population, age distribution of community population, the environment around the community, temperature and humidity. However, some scholars have studied the transmission mechanism between sunlight and virus (Asyary and Veruswati, 2020; Lansiaux et al., 2020) which support our conclusion to some extent. We will analyze the interaction mechanism between human settlements and pandemic transmission in detail in the next section.



**Figure 9:** Variable correlations and interactions (\*\*Representative passed the test of  $p = 0.01$ , \*passed the test of  $p = 0.05$ ).

### 3.2.2. Variable Fitting Analysis

We first carry out simple multiple linear regression modeling. Due to the existence of some outliers in the data, the fitting effect of the model is poor. As shown in Table 3, after eliminating the communities with missing fields, we reserved 544 communities for fitting with the fitting  $R^2 = 0.158$  and the adjusted  $R^2 = 0.147$ . We use the influencePlot function of R language to plot outliers, leverage points and strong influence points, and carefully delete some community points (there are outliers with building years earlier than 1900 and building of community buildings higher than 500). We retain 513 community information, and the fitting R square of the model is 0.276, adjusted R square is 0.266, the model after excluding gross errors can explain 26.6% of the total number of community diagnosis.

We also analyze the relationship between various elements. We can see from Figure 10 that after eliminating gross errors, the distribution of points is more reasonable and the fitting effect is better. The correlation between fee and VGI changed from positive correlation to non correlation (slope and correlation coefficient are near 0), which is also verified in the factor load analysis of structural equation model; The more expensive the house price, the more super high-rise buildings in the community, and the higher VGI of the community. Besides, we use the median of the construction year to segment the community data. Red represents the newer community (Year of construction is after 2005), and the green represents the older community (Year of construction is before 2004), the SVF distribution of new communities and older communities did not pass the t-test at the confidence level of  $p = 0.01$ , which means that the distribution of these two types is significantly different. In addition to this, consistent with our common knowledge (i.e., the more households, the more confirmed cases) new communities tend to have a larger population, and a larger population poses a greater risk of confirmed cases. However, the year of construction completion showed a negative correlation with the number of confirmed cases, which may be influenced by other factors (e.g., the newer the neighborhood, the higher the SVF value), so we use SEM models to explore further in the next section. Older neighborhoods show a single-peak distribution of home prices, while newer neighborhoods show a double-peak distribution. Some newer communities in better locations are more expensive than older communities, and correspondingly, some newer communities in more remote locations have lower house prices than older communities in mature developments. The number of new communities was significantly lower around 2015 (which may be related to the cooling of the real estate market at that time).

Urban street view images enrich our social sensing means. This section studies the relationship between the manifest variables and constructs the regression equation between the manifest variables and the cases. In the next section, we focus on the path relationship between the late variables and analyze the impact of human settlements on the spread of the pandemic.

**Table 3**  
OLS fitting effect before and after rejecting outliers

	<i>Reject outlier Before:</i>	<i>Reject outlier After:</i>
	cases	cases
fee	-0.087 (0.053)	-0.063* (0.038)
VGI	6.277* (3.702)	5.641** (2.631)
SVF	-23.176** (11.447)	-7.782 (8.008)
year	-0.198*** (0.061)	-0.294*** (0.057)
building	0.001 (0.018)	0.001 (0.017)
house	0.003*** (0.0004)	0.004*** (0.0003)
price	0.0002** (0.0001)	0.0002** (0.0001)
Constant	398.964*** (122.809)	590.141*** (114.155)
Observations	544	513
R <sup>2</sup>	0.158	0.276
Adjusted R <sup>2</sup>	0.147	0.266
Residual Std. Error	12.455 (df = 536)	8.517 (df = 505)
F Statistic	14.397*** (df = 7; 536)	27.476*** (df = 7; 505)

*Note:*

\*p<0.1; \*\*p<0.05; \*\*\*p<0.01

### 3.3. Main driving force and driving effect of environment on pandemic

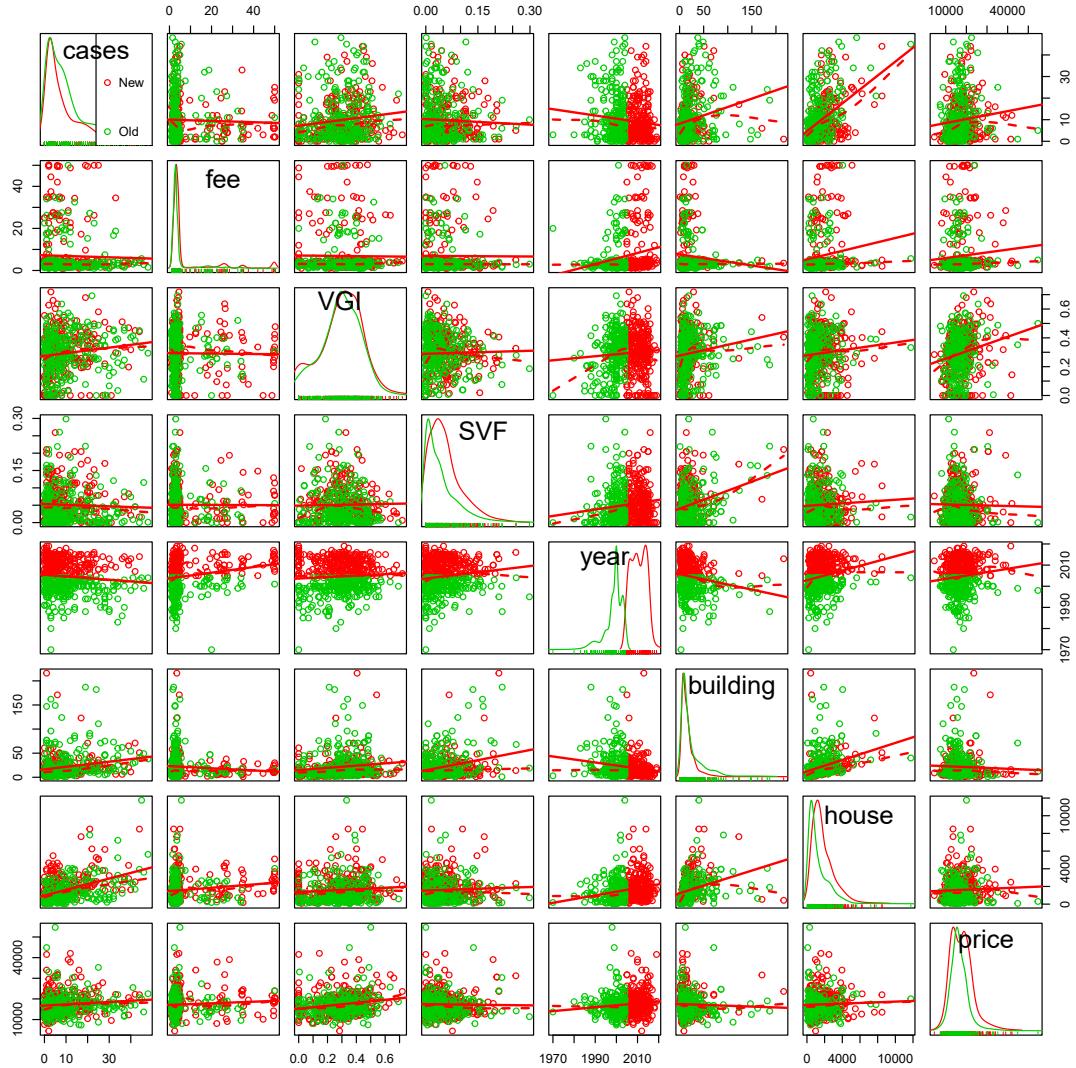
#### 3.3.1. Distribution statistics of evaluation index

In this section, we add the number of POI around the community to measure the convenience of community living index, and divide the community population density, the number of community buildings, and the number of community houses, which are statistically positively correlated with the pandemic results, into residential density of community index. We use Variance Inflation Factor (VIF) to detect 14 normalized manifest variables, and the VIF values of all observed elements are between 1 and 1.505 which means passing the collinearity test and there is no serious collinearity between all elements.

Table 4 summarizes the characteristics of the study area. The average VGI score of the remaining 513 communities is 0.293 and the SVF score is 0.051. Overall, the average price of houses in Wuhan is 17,343 Yuan, with an average of 1.897 park facilities and 4.840, 2.337 hospitals and shops respectively. The average construction time is mid-late 2004, and the community has an average of 21 buildings and 1,600 households. Descriptive statistics per index are given in Table 4.

#### 3.3.2. SEM model analysis

We believe that the better and more upscale the community, the better the environmental conditions, i.e. more green cover, lower floor area ratio and better lighting. In addition, the neighborhood's surrounding environment also



**Figure 10:** Two-by-two fitting relationship between variables (after reject outlier).

affects the neighborhood's environment. More commercial facilities around the neighborhood means more skyscrapers, which affects the neighborhood's lighting and greenery (compared to suburban neighborhoods). In addition, a more convenient neighborhood means being in the center of the city, which brings frequent population movements and noise disturbances. Based on these assumptions, we constructed a structural equation model.

We evaluate the SEM model through reliability and validity. All Factor loading (the influence of manifest variable on latent variable) of all indicators is between 0.578 with 1 in Table 5. It means that all Factor loadings in this study meet the structural validity requirements and could explain the latent variables reasonably. In addition, the convergent validity and discriminant validity of the model are based on Average Variance Extracted (AVE), which requires AVE greater than 0.5. In exploratory studies, the composite reliability(CR) value is generally higher than 0.7 and not lower than 0.6.

The study data largely satisfy the above conditions and meet part of the Fornell & Larcker criterion, indicating a good linear equivalence between the measured variables and the latent variables. In summary, the model fitting indicators are basically within the ideal values, which proves that the built environment evaluation model is basically

**Table 4**  
Distribution characteristics of each index

index	mean	median	min	max	standard deviation	excess kurtosis	skewness
cases	9.803	6	1	48	9.93	1.93	1.523
fee	6.895	3.25	1	50.25	10.254	7.327	2.79
VGI	0.293	0.303	0	0.719	0.148	-0.263	-0.202
SVF	0.051	0.038	0	0.298	0.05	2.645	1.486
year	2004.653	2005	1970	2019	7.278	0.559	-0.501
building	21.634	13	1	216	25.992	17.494	3.625
house	1600.852	1218	28	11749	1398.527	8.667	2.241
transport	15.117	17	0	20	5.521	-0.729	-0.798
education	13.881	16	0	20	6.662	-1.174	-0.578
shop	2.337	1	0	20	3.438	9.76	2.753
density	92.131	22.728	0	972.89	144.65	6.254	2.276
price	17343.936	16454	4471	54660	5593.153	6.477	1.75
park	1.897	1	0	20	2.746	13.241	3.07
hospital	4.84	3	0	20	5.455	1.326	1.471
lat	30.561	30.566	30.349	30.894	0.088	1.645	0.572
lng	114.299	114.312	114.009	114.534	0.094	-0.026	-0.512

**Table 5**  
Model reliability and validity evaluation

Latent Variable	CR	AVE	Manifest Variable	Factor Loading
Convenience of community living	0.805	0.455	Education	0.714
			Hospital	0.783
			Park	0.661
			Shop	0.578
			Trans	0.619
Residential density of community	0.771	0.637	Density	Delete
			Building	0.627
			House	0.938
Sky rate of community environment	0.657	0.498	Fee	0.561
			SVF	0.825
Suitability of community human settlement	1	1	Price	Delete
			Year	1
Green rate of community ecology	1	1	VGI	1
Number of confirmed cases in community	1	1	Cases	1

Note: Composite Reliability(CR)>0.65; Average Variance Extracted(AVE) >0.45; Factor loadings>0.55

reasonable and reliable (Sánchez-Medina et al., 2014; Hair et al., 2011; Park et al., 2017).

The path analysis results of SEM are shown in Figure 11 and Table 6, the observable variables are represented by rectangles, latent variables are represented by circles, and the relationship between variables is represented by arrows. The factor loadings of fee and density were close to zero, indicating that the setting of this variable had little effect on the model structure. We adjusted the model by removing these two variables to make it more reasonable.

PLS-SEM model explained the variance of 28.9% of the diagnosis results. Under the background of the pandemic situation, the number of hospitals, education and parks in the buffer zone has the factor loading of 0.783, 0.714 and 0.661 respectively. Besides, the residential density of community can best be explained under two observations: house (0.938) and building (0.627).

There are different paths between latent variables: under the confidence level of  $p = 0.01$ , the convenience around the community has a positive relationship (path coefficient =0.234, $p = 0.01$ ) with the number of confirmed cases, which was only lower than the resident density index. The path coefficient between the convenience and SVF/VGI (-

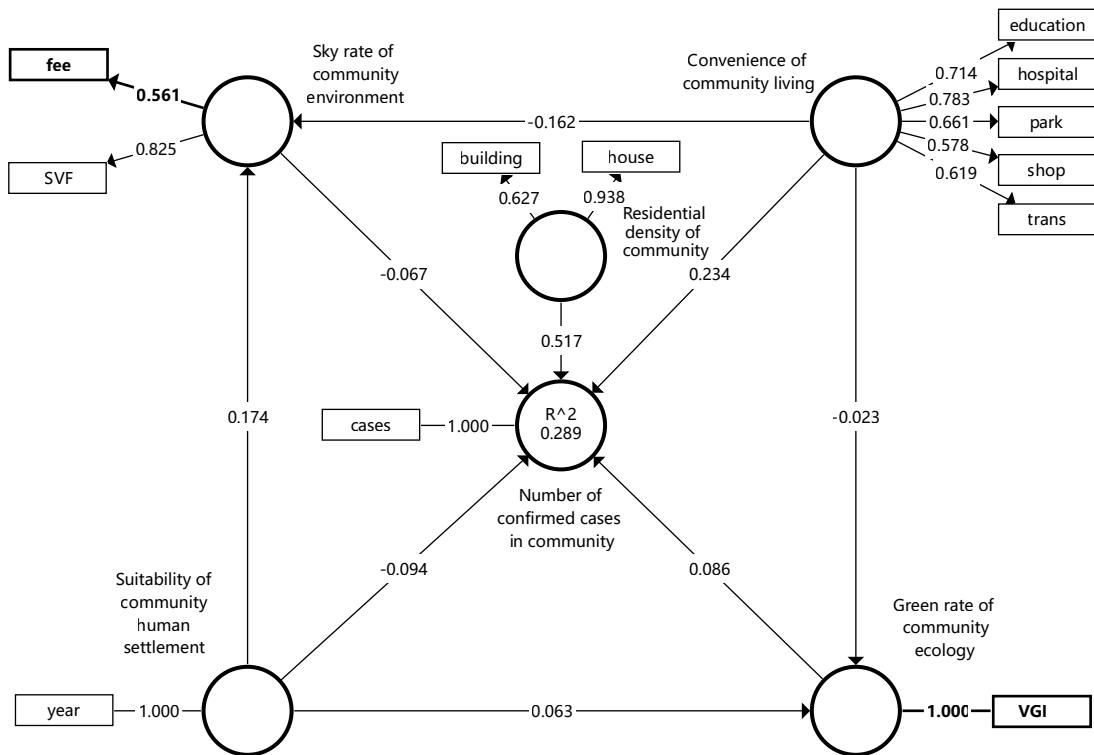


Figure 11: Structural equation model structure, path coefficients and factor loadings.

**Table 6**  
Analysis results of path coefficient

Path	Original sample estimate	Mean of resamples	Standard deviation	T-statistics	
Convenience of community living -> Green rate of community ecology	-0.023	●	-0.023	0.059	0.381
Convenience of community living -> Number of confirmed cases in community	0.234***	●	0.232	0.038	6.213
Convenience of community living -> Sky rate of community environment	-0.162**	○	-0.17	0.064	2.517
Green rate of community ecology -> Number of confirmed cases in community	0.086**	○	0.088	0.035	2.467
Suitability of community human settlement -> Number of confirmed cases in community	-0.067*	●	-0.064	0.036	1.859
Suitability of community human settlement -> Green rate of community ecology	0.063	●	0.061	0.05	1.256
Suitability of community human settlement -> Number of confirmed cases in community	-0.094**	○	-0.093	0.039	2.41
Suitability of community human settlement -> Sky rate of community environment	0.174***	●	0.166	0.064	2.709
Residential density of community -> Number of confirmed cases in community	0.517***	●	0.512	0.041	12.697

Note1: \* $p<0.1$  ● ; \*\* $p<0.05$  ○ ; \*\*\* $p<0.01$  ● ; Not significant ●

Note2: We evaluate the reliability using the bootstrap method and calculate the p-value.

0.162/-0.023) indicates that massive POIs have a negative impact on the community landscape. More POIs mean that the community is more prosperous and has more skyscrapers, which reduce affects the VGI and SVF of the community. On the other hand, the population flow will be more frequent in the areas with shopping malls and traffic stations around the community, which will also increase the risk of virus exposure among residents.

The good news is, Suitability and sky rate have some positive impacts, such as adequate property management level, enough building spacing and sufficient sunlight, leading to decreasing confirmed cases for these communities under the confidence condition of  $p = 0.05$ (Dietz et al., 2020). Differences in vegetation cover cannot cause differences in the survival of this virus, but can cause differences in how the virus is transmitted (easier or more difficult to transmit) and further cause differences in the number of covid-19 cases in different communities(Poirier et al., 2020; Notari, 2021; SanJuan-Reyes et al., 2020). The suitability of community human settlement (newer building age) will positively affect the sky rate of community environment (0.174) and the green rate of community ecology (0.063). In addition, as China's urbanization slows down, the focus has shifted from the size of the house to the quality of the community. People are willing to pay higher prices for better greenery and lower community plot ratios. Obviously, Wuhan also follows this trend. The newly built community has a better SVF and VGI as shown in Figure 9.

Driving forces of pandemic situation is a complex result under the joint action of various processes, we employed (PLS-SEM) to quantify the influence of community livability, community population density and community ecological environment, et. on community infection.

### **3.3.3. Impact analysis of temperature and humidity**

As mentioned above, SVF and VGI have strong ties to sunlight, which may have an effect on the pandemic situation. What can not be ignored is that the temperature and humidity have changed greatly during the pandemic in Wuhan. As the direct sunlight point gradually moves northward during the study time, we think that the sunlight (cloudy and sunny) also has a complex interaction connect with confirm cases, recover cases, humidity and temperature. Based on this hypothesis, we selected the period of nearly three months from January 15 to April 14, which is close to zero since the diagnosis of cases in Wuhan, to study the relationship between temperature and humidity and the number of confirmed and recovery cases in Wuhan.

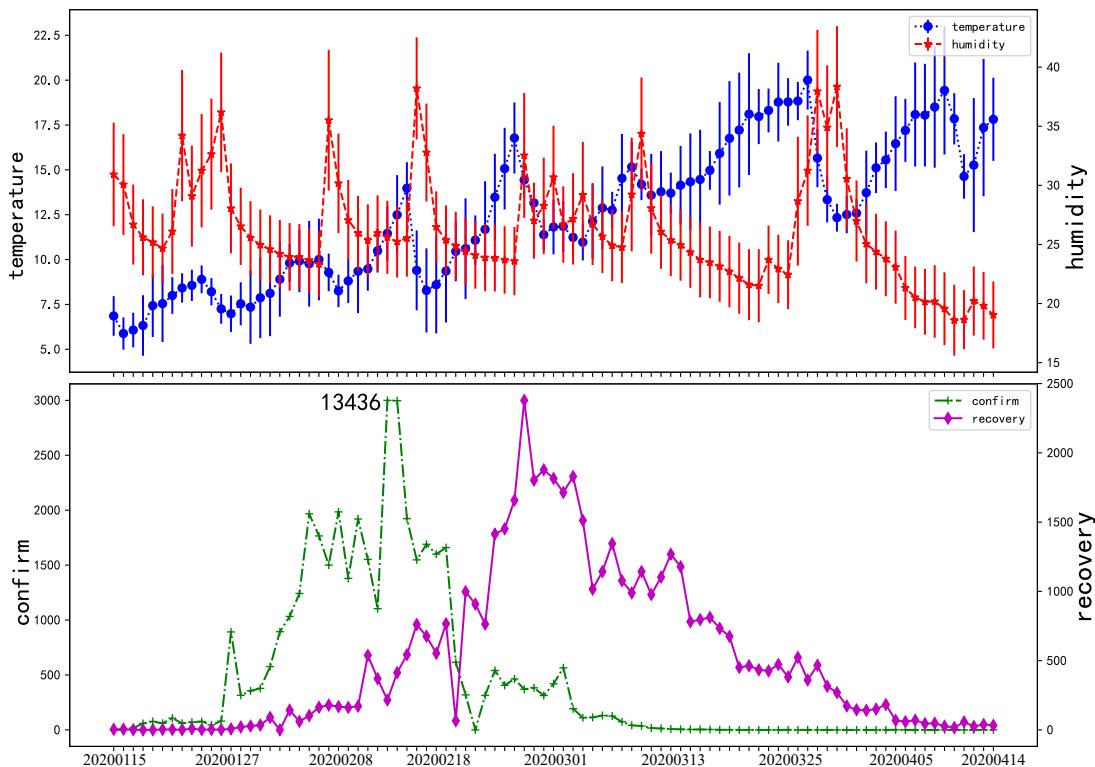
We have deployed 25 GEMHO temperature and humidity sensors <sup>3</sup> in the Baoxie region and Optics Valley areas in Wuhan. The sensor works continuously, and the recorded environmental data is uploaded to the data server. The specific parameters and layout of the sensor are displayed in the supplementary materials. We use the 1 standard deviation as an error bar and draw the daily temperature and humidity data and the daily diagnosis and rehabilitation cases reported by the health commission in Figure 12.

The temperature in Wuhan is likely to fluctuate, rising from 5 centigrade to 20 centigrade during the study period. There seems to be a lag negative correlation between humidity and temperature, but this is not the focus of this paper. There is a half month lag time between the peak value of recovery and the peak value of confirmed, which indicates that patients can recover in more than half a month under a good treatment environment. More importantly, the Pearson coefficients of the number of confirmed cases and median temperature recorded by sensor showed associations of -0.372 ( $p = 0.000$ ), and there is a significant positive related coefficient (0.183) between ambient temperature and the number of people recovered ( $p = 0.182$ ). In the meantime, the Pearson correlation index of humidity and confirmed cases is 0.102 ( $p = 0.332$ ), which means that the virus has a preference for damp and cool conditions to some extent.

We divided the data into six equally spaced groups based on temperature and humidity respectively. Each group recorded 15 days of observations and we performed Analysis of Variance (ANOVA) analysis on the data. As shown in Table 7, there is a significant difference in the number of confirmed diagnoses between the groups at the level of  $F = 3.391$  and  $p = 0.000$ . The number of confirmed diagnoses at different temperatures ( $p = 0.000$ ), the number of confirmed diagnoses at different humidity levels ( $p = 0.022$ ) were statistically significant.

This section makes a further exploration from the perspective of temperature and humidity on whether sunlight can inhibit the spread of the COVID-19 pandemic. The results show that the increase in temperature can significantly reduce the number of confirmed cases(Wu et al., 2020b). There are several possible explanations for this result. Higher temperatures may allow buildings to receive more sunlight for longer periods of time and it is also closely related to the occupation, age distribution of community residents, air quality and the concentration of airborne particulate matter (Organization et al., 2020). It is necessary to further study the complex mechanism.

<sup>3</sup>[www.minestar.cn](http://www.minestar.cn)



**Figure 12:** Temperature and humidity and the development of pandemic time series (Wuhan city adjusted the diagnostic criteria on February 12, 2020, and we compressed the data point on this day).

**Table 7**  
Tests of Between-Subjects Effects

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	26455718.788a	27	979841.437	3.391	0
Intercept	15163101.93	1	15163101.93	52.471	0
humidity	4123561.186	5	824712.237	2.854	0.022
temperature	8551959.67	5	1710391.934	5.919	0
humidity × temperature	6874415.607	17	404377.389	1.399	0.167
Error	18205904.25	63	288982.607		
Total	61830675	91			
Corrected Total	44661623.03	90			

*Note:*  $R^2 = .592$  (Adjusted  $R^2 = .418$ ) ; Dependent Variable: confirm cases

#### 4. Conclusion

In response to RQ1, we believe that the risk of virus propagation is different for communities under different conditions. The community with more High SVF & Low VGI scenarios (HL Class, usually newer, more expensive communities) is much more secure than LH Class.

As for RQ2, there is evidence for a relationship between the quantity of green space in the living environment and general health and mortality (Van den Berg et al., 2015). There is a positive influence of green space exposure on health outcomes (Shin et al., 2020). On the contrary, the main findings of this article are as follows: If a community has a low VGI and a high SVF, they have decreased confirmed cases. The mixing of environmental variables and physical variables exerts an effect upon the COVID-19 pandemic, which could explain the variance of 28.9%.

With the direct sunlight point moving north, the pandemic in Wuhan was under control in April. Based on the above two RQ findings, we hypothesize that communities with higher SVF have more access to sunlight, which has a significant inhibitory effect on the spread of the pandemic (Asyary and Veruswati, 2020; Lansiaux et al., 2020). At the same time, we verify the "moisture and cold tolerance" of the virus using the deployed sensors.

The spread of COVID-19 is influenced by a combination of factors: governmental measures to prevent the pandemic, buildings and environment, adequacy of access to medical supplies, temperature, humidity and climate, and the proportion of elderly people in the population, etc. This paper provides the first analysis of pandemic at the community scale in terms of economic and environmental heterogeneity. Using statistical methods, we found that the number of buildings and the density of community (residential density of community) had the most direct effect on the confirm cases, followed by the surrounding environment (convenience of community living). This implies that it is necessary to reduce the movement of outsiders in the community and to implement community closure/semi-closure in areas with severe pandemic. According to our findings in RQ2, good community management, as well as a clean built environment with more sunlight, can lead to a better ability of human communities to interrupt virus transmission.

Controlling the source of infection, cutting off the route of transmission and protecting the susceptible population are the three most important measures in the prevention of infectious diseases. We hope that our research will help to reveal the mechanism of sunlight in the transmission of viruses and defeat the COVID-19 as soon as possible. The model only explain a relatively small portion of the variance of the confirmed cases. Some important variables (e.g., social distancing, isolation of febrile patients, acceptance of wearing masks and school closing) were not considered. We will collect more relevant data and improve it in the future research.

## A. Declaration of competing interest

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

## B. Acknowledgments

The numerical calculations in this paper have been done on the supercomputing system in the Supercomputing Center of Wuhan University. Thanks to Lin Xin and Cheng Bowen for their help in sensor data. Dr.Chen Dong provided specific parameters of the sensor, Dr. Ma Hongliang provided discussion on the relationship between humidity and air quality, and Dr. Wang Peixiao supported the pandemic data in Wuhan.

## References

- CSY, . China Statistical Yearbook. China Statistical Publishing House; 2018.
- Wang, F., Zheng, P., Dai, J., Wang, H., Wang, R.. Fault tree analysis of the causes of urban smog events associated with vehicle exhaust emissions: A case study in jinan, china. *Science of the total environment* 2019a;668:245–253.
- Liu, M., Liu, X., Huang, Y., Ma, Z., Bi, J.. Epidemic transition of environmental health risk during china's urbanization. *Science bulletin* 2017;62(2):92–98.
- Hunter, P.. More science in urban development: Scientific evidence is crucial for managing health and infrastructures of megacities. *EMBO reports* 2017;18(2):201–204.
- Kerr, G.H., Badr, H.S., Gardner, L.M., Perez-Saez, J., Zaitchik, B.F.. Associations between meteorology and covid-19 in early studies: Inconsistencies, uncertainties, and recommendations. *One Health* 2021;:100225.
- Espejo, W., Celis, J.E., Chiang, G., Bahamonde, P.. Environment and covid-19: Pollutants, impacts, dissemination, management and recommendations for facing future epidemic threats. *Science Of The Total Environment* 2020;747:141314.
- Jayaweera, M., Perera, H., Gunawardana, B., Manatunge, J.. Transmission of covid-19 virus by droplets and aerosols: A critical review on the unresolved dichotomy. *Environmental research* 2020;:109819.
- Wang, R., Yuan, Y., Liu, Y., Zhang, J., Liu, P., Lu, Y., et al. Using street view data and machine learning to assess how perception of neighborhood safety influences urban residents' mental health. *Health & place* 2019b;59:102186.
- Rundle, A.G., Bader, M.D., Richards, C.A., Neckerman, K.M., Teitler, J.O.. Using google street view to audit neighborhood environments. *American journal of preventive medicine* 2011;40(1):94–100.
- Ye, Y., Richards, D., Lu, Y., Song, X., Zhuang, Y., Zeng, W., et al. Measuring daily accessed street greenery: A human-scale approach for informing better urban planning practices. *Landscape and Urban Planning* 2019;191:103434.
- Wu, C., Peng, N., Ma, X., Li, S., Rao, J.. Assessing multiscale visual appearance characteristics of neighbourhoods using geographically weighted principal component analysis in shenzhen, china. *Computers, Environment and Urban Systems* 2020a;84:101547.
- Huang, Y., Fei, T., Kwan, M.P., Kang, Y., Li, J., Li, Y., et al. Gis-based emotional computing: A review of quantitative approaches to measure the emotion layer of human–environment relationships. *ISPRS International Journal of Geo-Information* 2020;9(9):551.

- Yao, Y., Liang, Z., Yuan, Z., Liu, P., Bie, Y., Zhang, J., et al. A human-machine adversarial scoring framework for urban perception assessment using street-view images. *International Journal of Geographical Information Science* 2019;33(12):2363–2384.
- Liang, J., Gong, J., Zhang, J., Li, Y., Wu, D., Zhang, G.. Gsv2svf-an interactive gis tool for sky, tree and building view factor estimation from street view photographs. *Building and Environment* 2020a;168:106475.
- Helbich, M., Yao, Y., Liu, Y., Zhang, J., Liu, P., Wang, R.. Using deep learning to examine street view green and blue spaces and their associations with geriatric depression in beijing, china. *Environment international* 2019;126:107–117.
- Liu, Y., Wang, R., Grekousis, G., Liu, Y., Yuan, Y., Li, Z.. Neighbourhood greenness and mental wellbeing in guangzhou, china: What are the pathways? *Landscape and urban planning* 2019a;190:103602.
- Liu, Y., Wang, R., Xiao, Y., Huang, B., Chen, H., Li, Z.. Exploring the linkage between greenness exposure and depression among chinese people: Mediating roles of physical activity, stress and social cohesion and moderating role of urbanicity. *Health & place* 2019b;58:102168.
- Zhou, H., He, S., Cai, Y., Wang, M., Su, S.. Social inequalities in neighborhood visual walkability: Using street view imagery and deep learning technologies to facilitate healthy city planning. *Sustainable Cities and Society* 2019;50:101605.
- Zhang, F., Zu, J., Hu, M., Zhu, D., Kang, Y., Gao, S., et al. Uncovering inconspicuous places using social media check-ins and street view images. *Computers, Environment and Urban Systems* 2020;81:101478. URL: <http://www.sciencedirect.com/science/article/pii/S0198971519306003>. doi:<https://doi.org/10.1016/j.compenvurbsys.2020.101478>.
- Kang, Y., Zhang, F., Gao, S., Lin, H., Liu, Y.. A review of urban physical environment sensing using street view imagery in public health studies. *Annals of GIS* 2020;26(3):261–275.
- Vazquez-Prokopec, G.M., Kitron, U., Montgomery, B., Horne, P., Ritchie, S.A.. Quantifying the spatial dimension of dengue virus epidemic spread within a tropical urban environment. *PLoS Negl Trop Dis* 2010;4(12):e920.
- Wang, R., Liu, Y., Xue, D., Helbich, M.. Depressive symptoms among chinese residents: how are the natural, built, and social environments correlated? *BMC public health* 2019c;19(1):887.
- Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., et al. Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers* 2015;105(3):512–530.
- Qu, S., Hu, S., Li, W., Zhang, C., Li, Q., Wang, H.. Temporal variation in the effects of impact factors on residential land prices. *Applied Geography* 2020;114:102124.
- Kraemer, M.U., Yang, C.H., Gutierrez, B., Wu, C.H., Klein, B., Pigott, D.M., et al. The effect of human mobility and control measures on the covid-19 epidemic in china. *Science* 2020;368(6490):493–497.
- Yawson, D.O.. Quantifying perceived landscape desirability in human settlements: the case of four communities in cape coast, ghana. *Ghana Journal of Geography* 2020;12(1):74–98.
- Gong, F.Y., Zeng, Z.C., Zhang, F., Li, X., Ng, E., Norford, L.K.. Mapping sky, tree, and building view factors of street canyons in a high-density urban environment. *Building and Environment* 2018;134:155–167.
- Chen, L.C., Papandreou, G., Schroff, F., Adam, H.. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:170605587* 2017;.
- He, K., Zhang, X., Ren, S., Sun, J.. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, p. 770–778.
- Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A.. Inception-v4, inception-resnet and the impact of residual connections on learning. *arXiv preprint arXiv:160207261* 2016;.
- Wang, R., Lu, Y., Wu, X., Liu, Y., Yao, Y.. Relationship between eye-level greenness and cycling frequency around metro stations in shenzhen, china: A big data approach. *Sustainable Cities and Society* 2020;:102201.
- Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H.H., Lin, H., et al. Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning* 2018;180:148–160.
- Chen, T., Li, M., Li, Y., Lin, M., Wang, N., Wang, M., et al. Mxnet: A flexible and efficient machine learning library for heterogeneous distributed systems. *arXiv preprint arXiv:151201274* 2015;.
- Stone, M., Brooks, R.J.. Continuum regression: cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression. *Journal of the Royal Statistical Society: Series B (Methodological)* 1990;52(2):237–258.
- Varian, H.R., Varian, H.R.. *Microeconomic analysis*; vol. 3. Norton New York; 1992.
- Orford, S.. Valuing the built environment: GIS and house price analysis. Routledge; 2017.
- Liang, X., Liu, Y., Qiu, T.. Livability assessment of urban communities considering the preferences of different age groups. *Complexity* 2020b;2020.
- Houghton, A., Castillo-Salgado, C.. Analysis of correlations between neighborhood-level vulnerability to climate change and protective green building design strategies: A spatial and ecological analysis. *Building and Environment* 2020;168:106523.
- Diao, M., Ferreira Jr, J.. Residential property values and the built environment: Empirical study in the boston, massachusetts, metropolitan area. *Transportation research record* 2010;2174(1):138–147.
- Hair, J.F., Ringle, C.M., Sarstedt, M.. Pls-sem: Indeed a silver bullet. *Journal of Marketing theory and Practice* 2011;19(2):139–152.
- Hair Jr, J.F., Hult, G.T.M., Ringle, C., Sarstedt, M.. A primer on partial least squares structural equation modeling (PLS-SEM). Sage publications; 2016.
- Fan, Y., Chen, J., Shirkey, G., John, R., Wu, S.R., Park, H., et al. Applications of structural equation modeling (sem) in ecological studies: an updated review. *Ecological Processes* 2016;5(1):19.
- Li, X., Zhou, L., Jia, T., Wu, H., Zhou, Y., Qin, K.. Influence of urban factors on the covid-19 epidemic: A case study of wuhan city. *Wuhan Daxue Xuebao (Xinxix Kexue Ban)/Geomatics and Information Science of Wuhan University* 2020;45(6).
- Guida, C., Carpentieri, G.. Quality of life in the urban environment and primary health services for the elderly during the covid-19 pandemic: An application to the city of milan (italy). *Cities* 2020;110:103038.
- Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J.. Pyramid scene parsing network. In: *Proceedings of the IEEE conference on computer vision and*

- pattern recognition. 2017, p. 2881–2890.
- Zhao, H., Qi, X., Shen, X., Shi, J., Jia, J.. Icnet for real-time semantic segmentation on high-resolution images. In: Proceedings of the European Conference on Computer Vision (ECCV). 2018, p. 405–420.
- Girshick, R.. Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision. 2015, p. 1440–1448.
- Ren, S., He, K., Girshick, R., Sun, J.. Faster r-cnn: Towards real-time object detection with region proposal networks. In: Advances in neural information processing systems. 2015, p. 91–99.
- Asyary, A., Veruswati, M.. Sunlight exposure increased covid-19 recovery rates: A study in the central pandemic area of indonesia. *Science of The Total Environment* 2020;;139016.
- Lansiaux, É., Pébaÿ, P.P., Picard, J.L., Forget, J.. Covid-19 and vit-d: Disease mortality negatively correlates with sunlight exposure. *Spatial and Spatio-temporal Epidemiology* 2020;35:100362.
- Sánchez-Medina, A.J., Romero-Quintero, L., Sosa-Cabrera, S.. Environmental management in small and medium-sized companies: an analysis from the perspective of the theory of planned behavior. *PloS one* 2014;9(2):e88504.
- Park, H., Fan, P., John, R., Chen, J.. Urbanization on the mongolian plateau after economic reform: Changes and causes. *Applied Geography* 2017;86:118–127.
- Dietz, L., Horve, P.F., Coil, D., Fretz, M., Eisen, J., Van Den Wymelenberg, K.. 2019 novel coronavirus (covid-19) outbreak: A review of the current literature and built environment (be) considerations to reduce transmission 2020;.
- Poirier, C., Luo, W., Majumder, M.S., Liu, D., Mandl, K.D., Mooring, T.A., et al. The role of environmental factors on transmission rates of the covid-19 outbreak: an initial assessment in two spatial scales. *Scientific reports* 2020;10(1):1–11.
- Notari, A.. Temperature dependence of covid-19 transmission. *Science of The Total Environment* 2021;763:144390.
- SanJuan-Reyes, S., Gómez-Oliván, L.M., Islas-Flores, H.. Covid-19 in the environment. *Chemosphere* 2020;;127973.
- Wu, Y., Jing, W., Liu, J., Ma, Q., Yuan, J., Wang, Y., et al. Effects of temperature and humidity on the daily new cases and new deaths of covid-19 in 166 countries. *Science of the Total Environment* 2020b;729:139051.
- Organization, W.H., et al. Population-based age-stratified seroepidemiological investigation protocol for coronavirus 2019 (covid-19) infection, 26 may 2020. Tech. Rep.; World Health Organization; 2020.
- Van den Berg, M., Wendel-Vos, W., van Poppel, M., Kemper, H., van Mechelen, W., Maas, J.. Health benefits of green spaces in the living environment: A systematic review of epidemiological studies. *Urban Forestry & Urban Greening* 2015;14(4):806–816.
- Shin, J.C., Kwan, M.P., Grigsby-Toussaint, D.S.. Do spatial boundaries matter for exploring the impact of community green spaces on health? *International Journal of Environmental Research and Public Health* 2020;17(20):7529.