S&E

Vol. 91
SEPTEMBER 2022

Space & Environment
is a quarterly English magazine intended to disseminate research output and international activities of Korea Research Institute for Human Settlements

Post COVID-19 and Spatial Data

01 **Supporting Infectious Disease Prevention Policy through Spatiotemporal Simulation Model**

1. A global pandemic and the need for new policy simulations
2. Simulation methodology and empirical results to judge policy effectiveness
3. Implications and future challenges

02 **Pure-inTention: A Study on Deep Learning-based Purpose-driven Trip Demand Estimation Model for Post-Pandemic Countermeasures**

1. Introduction
2. Trip demand according to purpose
3. Deep learning-based trip demand identification according to purpose using observed link counts
4. Research design for Pure-inTention
5. Data integration and pre-processing for Pure-inTention
6. Experimental results and conclusion

03 **Current Status and Improvement Directions of COVID-19 Outbreak Monitoring: Based on An Analysis of New Cases by *Eup-Myeon-Dong* in 2020 and 2021**

1. Introduction
2. Current status of COVID-19 monitoring in South Korea
3. An empirical analysis of COVID-19 cases by *Eup-Myeon-Dong*
4. Conclusion: Suggestions for improved COVID-19 monitoring

ISSN 2671-9037
9 772671 903003 23

Published by Korea Research Institute for Human Settlements (KRIHS, http://www.krihs.re.kr/)
Publisher Hyun-Soo Kang
Edited by Global Development Partnership Center (GDPC, http://gdpc.kr/)
Editorial Team Sang Keon Lee (Editor), Hye Jung Park, Eunji Ju, Jeong Yun Hwang
Tel. +82-44-960-0429 Fax. +82-44-211-4772
Designed and Produced by Goryeo C&P Co., Ltd. Tel. +82-2-2277-1508~9
Copyright 2022 ©Korea Research Institute for Human Settlements All Rights Reserved. 5 Gukchaegyeonguwon-ro, Sejong-si, 30147, Korea

Korea Research Institute for Human Settlements · 국토연구원 ·

01 Supporting Infectious Disease Prevention Policy through Spatiotemporal Simulation Model

Jaesoen Son

1. A global pandemic and the need for new policy simulations

■ The rapidly spreading COVID-19 and social distancing

Unlike the existing coronaviruses, COVID-19 is characterized by mutations in the spikes that come into contact with the respiratory epithelial cells resulting in high transmission rate, rapid worsening of symptoms, and rapid spread to other organs (OpenWHO; Ahn, 2021). Although COVID-19 has spread all over the world, the health systems and quarantine policies in different countries have made a big difference in the rate of increase and decrease in the number of confirmed cases and associated deaths (Cacciapaglia et al. 2021). The initial strict lockdown policies drastically reduced people's mobility slowing the spread of COVID-19. However, the lockdown policy was lifted in stages due to various socio-economic side effects caused by the prolonged lockdown. Korea has implemented social distancing as a policy in response to COVID-19, and has suggested precautions to be observed by individuals, facilities, and in workplaces. In accordance with the standards for limiting the spread of infectious diseases and the occurrence of confirmed cases, we are implementing step-by-step preventive measures such as social distancing in advance.

■ Traditional mathematical models and latest research trends

The traditional way to predict the spread of an infectious disease uses a mathematical model. A mathematical model refers to the description of a real phenomenon with the help of a mathematical expression; for the construction and simulation of such models, the application of mathematical, statistical, and numerical calculation techniques is required. In a mathematical model, the most important indicator used to determine whether an infectious disease spread is the basic reproduction number, that is, the number of secondary infections that the first infected person can cause on an average. The SIR[01] model, which is a representative mathematical model, predicts the macroscopic changes in the number of infected people over time by applying key parameters affecting the spread of the disease, such as infection and recovery rates, to an entire country or region.

In recent times, the development of micro-models is accelerating due to the development of computer technology and the explosive increase in the available data. For example, various mobility indicators help explore spatially different changes in mobility along with time-series changes in daily life due to COVID-19. Additionally, indicators for measuring the effectiveness of quarantine policies were developed. For example, OxCGRT[02] provides a comprehensive stringency index based on the quantification of policy responses to compare various government responses to COVID-19.

■ Predicting policy effect through simulation

A general prediction of the number of confirmed cases has a limitation in that the quarantine policy cannot provide a spatial and temporal prediction that can be applied to a detailed spatial unit where an individual's real life takes place. "Social distancing" was initially applied to a wide range of regions and groups, which caused damage to the local economy and increased antipathy due to continued control; thus, increasing social stress. In the end, it is difficult to resolve the dissatisfaction of the general public subjected to social distancing, unless treatment for infected patients is generalized through the development and dissemination of therapeutic agents.

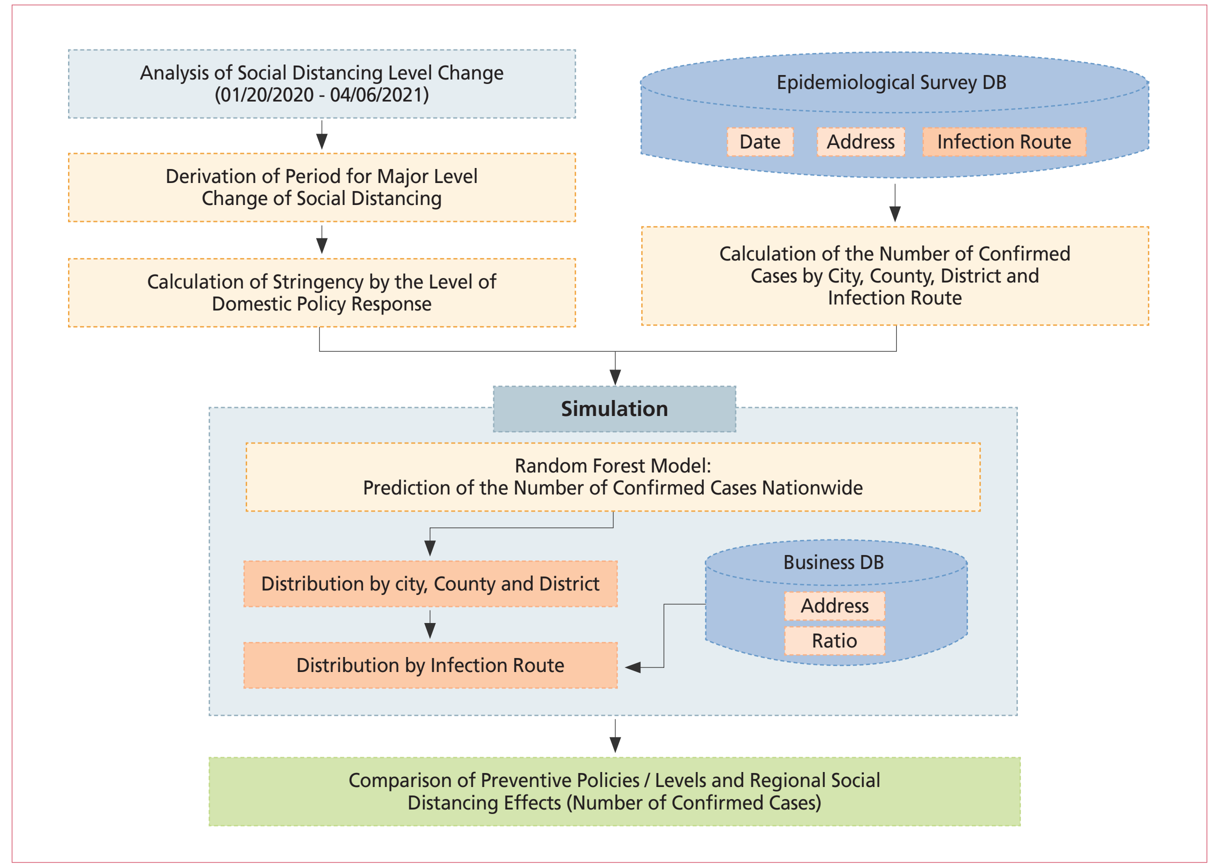
Rather than predicting that the number of confirmed cases will reach thousands next week, it is more practical to suggest what social distancing policies will work in an area of interest. If the effects of social distancing policies are presented in a detailed level in terms of time and space through scientific methodologies, it will be possible to reduce the antipathy towards social distancing and induce more active participation. Considering this, we propose a simulation method that can predict the effect of government response policies to prevent the spread of infectious diseases such as COVID-19, and derive the necessary future tasks through demonstration.

2. Simulation methodology and empirical results to judge policy effectiveness

■ Policy simulation methodology and data structure

The policy simulation model (Figure 1) first analyzes the level changes in social distancing, derives the cycle of major level change, and starts by calculating the level stringency of domestic prevention policies. Using the epidemiological investigation data of the Korea Disease Control and Prevention Agency (KDCA), the number of confirmed cases by city/county/district and the infection route is calculated. Using this data, the number of confirmed cases on a nationwide scale is predicted with a random forest model, that is, a machine learning model. By inputting the expected number of confirmed cases nationwide into the distribution model, the number of confirmed cases by city/county/district and the infection route is calculated. Finally, by simulating the type and stringency of the prevention policy, the expected number of confirmed cases by region[03] and by infection route are predicted, and through this, the effect of social distancing is compared by policy type, stringency level, and region. The data of the simulation model consists of eight matrices consisting of time (x-axis), 250 municipalities (y-axis), and eight government policies (z-axis), and one matrix of the number of confirmed cases (see Figure 2).Figure 1 Figure 2

Figure 1. Structure of policy simulation methodology



Note
Epidemiological Survey database is provided by KDCA (Korea Disease Control and Prevention Agency); Business database is collected from localdata.go.kr.

Source
Son, et al. (2021, 222).

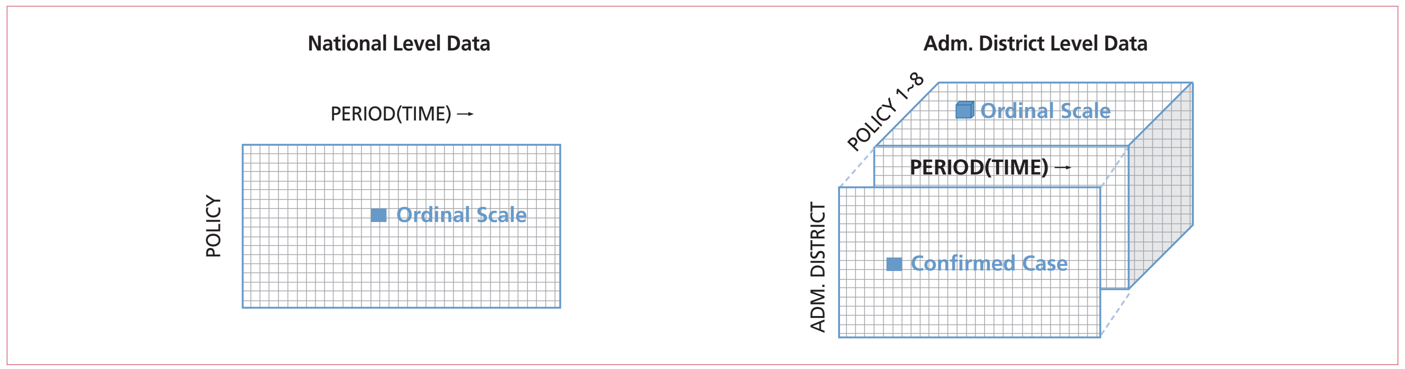
01. Susceptible-Infected-Recovered.
02. The Oxford Covid-19 Government Response Tracker, Blavatnik School of Government, University of Oxford.
03. Administrative districts at the municipal level (city/county/district).

■ Data preparation and simulation models

To demonstrate the simulation model of the COVID-19 prevention policy, first, regular briefings by the Ministry of Health and Welfare, changes in the social distancing system, and level adjustment documents for each city/county/district were collected and organized accordingly. Based on the eight detailed prevention policies from January 20, 2020 to April 6, 2021 in 250 cities/counties/districts, the daily stringency was calculated by referring to the changes in the level of social distancing and preventive policy guidelines (Table 1). The number of confirmed cases was calculated using the data on the number of daily confirmed cases, collected and managed by the KDCA during the same period.

The method of predicting the national frequency of confirmed cases starts with constructing input variables using a random forest model and identifying the correlations between variables. The random forest model adjusts the predictive power of a model while collecting and developing a set of causal relationships between the variables in an ensemble form. The number of confirmed cases by day, region, and infection route was assumed as the dependent variable, and the change in stringency of eight policies by day and region was assumed as the independent variable. The model for predicting the occurrence of confirmed cases was configured to calculate in terms of one-week units from the one-week average to the four-week average. The distribution model was constructed to be proportional to the ratio of the number of past confirmed cases in the city/county/district. In addition, the simulation process included allocating the number of confirmed cases in consideration of the density of facilities located in the relevant city/county/district related to the route of infection. Table 1

Figure 2. Schematic diagram of policy simulation data



Note For national data, we used the Korean data from the OxCGRT GitHub public data; for the city, county, and district level data, we used the combined data from the Central Quarantine Countermeasures Headquarters Information Analysis Team (2021) and the Ministry of Health and Welfare press releases on COVID-19.
Source Son, et al. (2021, 116).

Table 1. Stringency of prevention policy and social distancing levels

Government policy	Level 1	Level 1.5	Level 2	Level 2.5	Level 3
Mandatory mask-wearing	1	2	3	4	4
Restrictions on meetings and events	1	2	3	4	5
Restrictions on attending sports events	1	2	3	4	5
Restrictions on using transportation facilities	1	1	2	3	4
Restrictions on attending school	1	2	3	4	5
Restrictions on religious activities	1	2	3	4	5
Restrictions on job work	1	2	2	3	4
Restrictions on multi-use facilities	1	2	3	4	5

Note 1 During the period when social distancing was operated as a three-level system, levels 1, 2, and 3 of the five-level system were applied as it is.
Note 2 Modified from the Ministry of Health and Welfare (2020, 48-50).
Source Son, et al. (2021, 115).

■ Simulation demonstration results

The prediction of the number of confirmed cases on a nationwide scale showed a pattern very similar to the actual value. The true and predicted values were more similar to the result of the four-week average number of confirmed cases than that of the one-week average. Regarding the relative influence according to the route of infection, "workplace" and "religious facility" showed the highest relative influence. The relative influence of "correctional institute" and "unclassified" showed an increasing trend toward long-term prediction. On simulating the stringency of religious facilities and workplaces, it was found that the higher the stringency, the higher is the

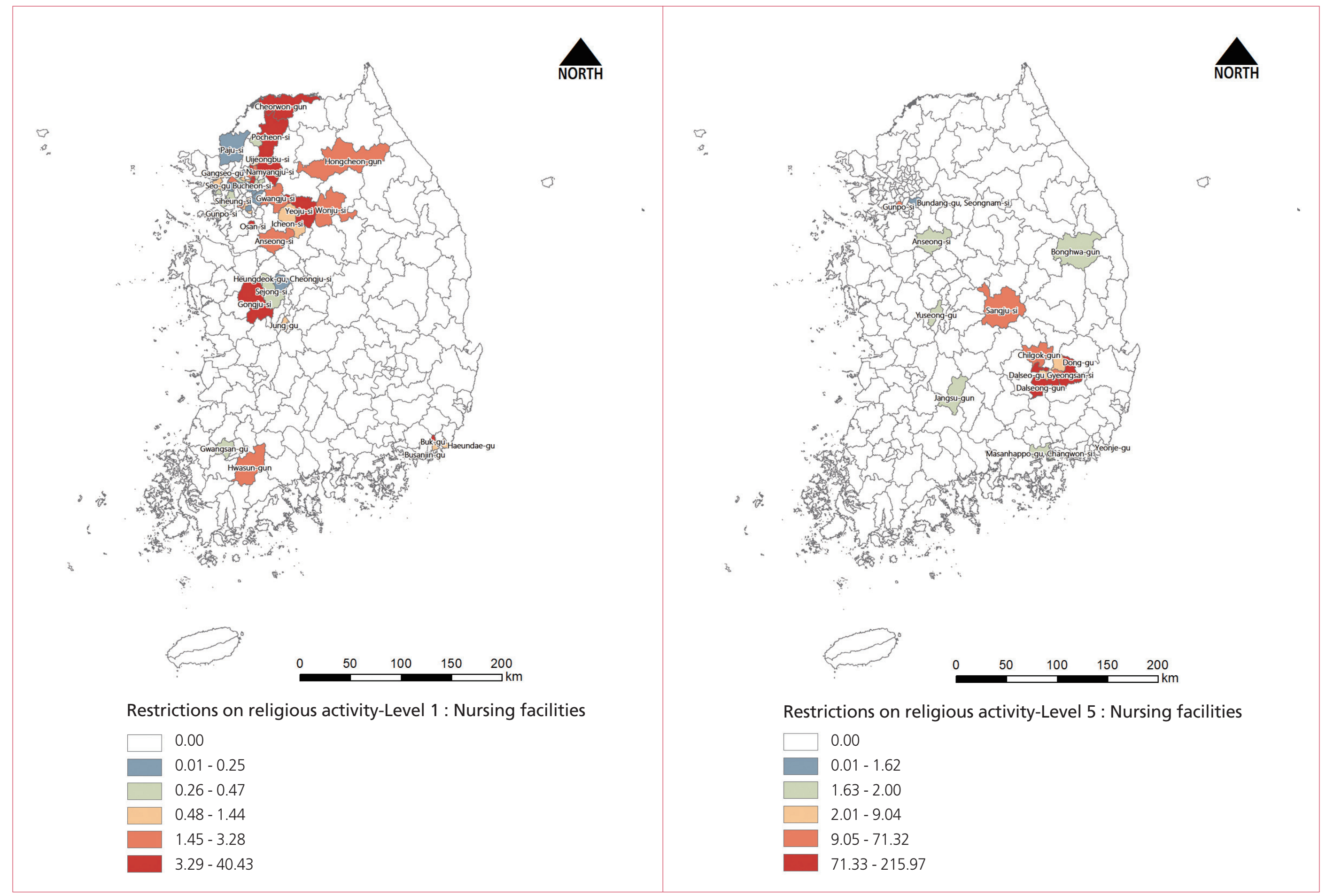
Table 2. Simulation result of the number of confirmed cases by infection route according to the level adjustment for each prevention policy

Unit: person

Simulation	Nursing facilities	Medical institutions	Religious facilities	Educational institutes	Total
Restrictions on religious activity-Level 1	224.39	203.68	418.20	173.21	1,019.48
Restrictions on religious activity-Level 5	722.87	908.72	261.44	16.97	1,910.00
Restrictions on job work-Level 1	220.34	203.79	406.05	170.89	1,001.07
Restrictions on job work-Level 4	723.19	908.48	261.29	17.05	1,910.01

Source
Son, et al. (2021, 182).

Figure 3. Examples of distribution pattern by city, county, and district as a result of simulation (Restrictions on religious activity-Level: Infection route)



Source Son, et al. (2021, xiv).

average number of confirmed cases. This is presumed to reflect the reality of policy implementation by raising the level of social distancing after the increase in the number of confirmed cases.

The scenario for the distribution model assumes the number of confirmed cases nationwide as 2,000, and raises the stringency of "restrictions on religious activity" and "restrictions on job work" to level 5 and level 4, respectively. The number of confirmed cases by infection route was predicted. For both "restrictions on religious activity" and "restrictions on job work," it was predicted that more than 95% of the total number of confirmed cases would occur intensively in the four infection routes (nursing facilities, medical institutions, religious facilities, and educational institutes) (Table 2). When the stringency was increased compared to level 1, the occurrence was concentrated in medical institutions and nursing facilities. Geographically, in the case of level 1 restrictions on religious activity, a large number of outbreaks were expected to center on the metropolitan area. However, when the "restriction on religious activity" is raised to level 5, the geographical scope of the outbreak area is limited and it is predicted that the outbreak will occur on a fairly large scale in a small number of areas (Figure 3). Table 2 Figure 3

3. Implications and future challenges

The implications that can be drawn from the policy simulation model and demonstration are as follows. First, through policy simulation, the importance of adjusting the social distancing level regardless of the type of prevention policy was presented. Second, in order to increase the effectiveness of the preventive policies, the necessity of preparing flexible prevention policies in consideration of the actual conditions of each region was confirmed by subdividing the factors that cause regional deviations. Finally, this simulation model can support the predictions of the possible outbreak areas and the routes of spread; hence, it was confirmed that this model can be used to support the transition from a response after an outbreak to a management system following a pre-emptive response.

In future research, it is necessary to improve the predictive power and logical explanations of the simulation model. Further research on applicable simulation models and expansion of policies applicable to the simulation are also needed. Considering the fact that COVID-19-related research is being conducted at various institutions, it is necessary to prepare an environment that can comprehensively collect and share infectious disease-related data, analysis methods, and results. In addition, it is necessary to enable the use of standardized information by reviewing the types of infection route information and collection and use systems. Lastly, in accordance with the Infectious Disease Prevention and Management Act, it is necessary to prepare more detailed systems to respond and manage the infectious diseases using data analysis and information and communication technology.

Jaesoen Son
Associate Research Fellow
Geospatially Enabled Society Research Division
jsson@krihs.re.kr

> "If the effects of social distancing policies are presented in a detailed level in terms of time and space through scientific methodologies, it will be possible to reduce the antipathy towards social distancing and induce more active participation."

References

- Ministry of Health and Welfare. 2020. COVID-19 Central Disaster and Safety Countermeasures Headquarters Regular Briefing, November, 1. Press reference.
- Son Jaesoen, Chang Yohan, Oh Changhwa, Im Eunsun, Hwang Myunghwa, and Lee Gyungju. 2021. *A Study on the Development of a Spatiotemporal Simulation Model to Support an Infectious Disease Prevention Policy.* Sejong: Korea Research Institute for Human Settlements.
- Ahn Kwang Seok. 2021. *COVID-19 variant vs immune defense system.* COVID-19 Scientific Report 2, January 15. Daejeon: Institute for Basic Science. https://www.ibs.re.kr/cop/bbs/BBSMSTR_000000001003/selectBoardArticle.do?nttId=19578 (accessed July 22, 2022)
- Central Quarantine Countermeasure Headquarters Data Analysis Team. 2021. COVID-19 confirmed patients' database and list of materials provided as of May 3, 2022. Unpublished manuscript.
- Cacciapaglia, G., Cot, C., and Sannino, F. 2021. Multiwave pandemic dynamics explained: how to tame the next wave of infectious diseases. *Scientific Reports,* 11(1), 6638. https://doi.org/10.1038/s41598-021-85875-2
- OpenWHO. Introduction to COVID-19: methods for detection, prevention, response and control. https://openwho.org/courses/introduction-to-ncov (accessed July 22, 2022)
- OxCGRT GitHub Data. 2022. https://github.com/OxCGRT/covid-policy-tracker/tree/master/data (accessed July 22, 2022)

02 Pure-inTention: A Study on Deep Learning-based Purpose-driven Trip Demand Estimation Model for Post-Pandemic Countermeasures

Yohan Chang

1. Introduction

The high transmission rate of new infectious respiratory diseases is characterized by rapid and contact-dependent infection. In response to the recently identified infectious respiratory diseases, the government implemented policies to prevent the spread of disease in facilities for group gatherings, which may lead to personal quarantine. These timely policies included reduced business hours for restaurants and 2-meter social distancing, which seemed to have had some effect in preventing the spread of the infectious respiratory diseases in the early stages. However, unified quarantine measures that did not consider regional characteristics led to unintended consequences such as deterioration of the local economy and domestic market. Current quarantine policies focus on controlling the density of activity spaces rather than considering the reasons for traveling to specific spaces, which have led to prolonged quarantine policies and consequences such as fatigue and balloon effects. Quarantine policy measures based on identifying the causal relationship of trip patterns and the purpose of trip rather than forcibly regulating the spatial density are expected to provide tailored policies for different regions and alleviate the unintended consequences of the current quarantine policies.

2. Trip demand according to purpose

Each trip from an origin to a destination has its own purpose and is achieved as the interaction between spatial service characteristics and corresponding demand reach an equilibrium. Trip makers decide to their trip for various purposes such as commuting to work or school, shopping, and socializing. Understanding the purpose of trip can provide important information not only for the current trip but also for predicting future travel patterns. Traditionally, the Korean National Household Travel Survey has been conducted detailed travel demand with comprehensive information in terms of purpose, means, and space. However, as the survey is conducted only every 5 years, it is limited with respect to immediately solving problems including real-time monitoring of the infectious diseases. Recently, studies using floating population big data from mobile phone base station information have been reported. While, these studies have critical issues such as privacy concerns in the process of data acquisition and purification and wasteful spending to purchase high-priced nationwide data.

3. Deep learning-based trip demand identification according to purpose using observed link counts

The recent development of infrastructure for collecting traffic counts on the road and the advancement of the trip demand prediction algorithm have enabled near real-time monitoring of travel demand; this led to identification of trip demand according to origin and destination. In accordance with the Intelligent Transportation System (ITS) and related policies, vehicle detectors on roads across the country produce real-time nationwide traffic volume information. The real-time trip demand estimation model using the observed traffic volume, which had limitations due to hardware performance, have been integrated with advanced methodologies such as deep learning (DL) for various ways. The DL mechanism that allows the model to learn a causal relationship pattern between an input value and output (or phenomenon) can rapidly recognize and derive reasonable results for repeated similar problems. In addition, the DL mechanism can learn and predict even minor details that cannot be detected by human intelligence. DL models that have completed learning are able to derive highly accurate results within just a few seconds even with a high volume of input data. Thus, DL models are highly regarded worldwide for identifying real-time trip demand using observed traffic volume.

4. Research design for Pure-inTention

The purpose of this study was to develop a model for predicting trip demand for each origin/destination and its purpose using real-time flow information (i.e., observed traffic volume), establish a trip demand estimation model using purpose-based DL of the developed model's pattern, and ultimately develop policy tools to seek measures for gradual recovery in the post-pandemic period. The characteristics of spaces related to trip demand according to purpose were analyzed by combining public and private big data, and the analyzed characteristics were linked using observed traffic volume-based trip demand estimation model by origin and destination. Subsequently, temporal and spatial (regional and daily) pre- and post-pandemic data patterns, such as confirmed cases, social distancing information, and real-time road traffic information were combined with a DL model and trip demand estimation model by each trip purpose. The developed model was then used to understand the causal relationship between trip patterns and the purpose of trip that have changed post-pandemic and derive preemptive post-pandemic response policies.

The Pure-inTention model presented in this study consisted of four steps (Table 1). The first step was data integration, which involved collecting various data from public and private sources into the same space units such as town, township, and neighborhood. In the second step, factor analysis of the collected spatial unit data was performed to identify local attraction factors according to each purpose of trip. In this particular step, various machine learning (ML) models such as a multidimensional data reduction algorithm and random forest were used. The third step clustered the attraction factors according to purpose to derive the key contributing factors for each cluster. The last step was the prediction step, and two models, A and B, are presented. Model A measures the effects of infectious diseases by learning the trip demand patterns according to purpose from the observed pre-pandemic traffic volume and predicts the effects using post-pandemic traffic volume by region. The model suggests differences between the predicted and actual trip patterns by region and trip purpose; regions where the predicted trip pattern according to purpose differed significantly from the actual observed trip pattern were selected for analysis of the key factors in ML fashion. Model B estimates trip sensitivity for each purpose according to changes in regional social distancing policies. After learning the changes in social distancing and number of confirmed cases in region level as well as trip demand information by purpose following the spread of infectious disease, the sensitivity of observed traffic volume with respect to changes in social distancing was estimated by region and trip purpose.[01] Table 1

Table 1. General process for Pure-inTention

Step	Descriptions		Methodologies & Models
I	Data integration into the same space unit (i.e., jurisdictional level)		Data analytic (with data fusion)
II	Factor analysis to identify trip attraction factors for each trip purpose		PCA, random forest, UMAP
III	Classify space units by similar characteristics of each space		K-means clustering, random forest
IV	Model development		GAN, DNNs, Frank-Wolf algorithm, Dijkstra's algorithm, random forest, GA
	Model A	Pandemic effect measurement	
	Model B	Trip (by purpose) sensitivity analysis during pandemic against policy deployment	

Note
PCA, principal component analysis; UMAP, uniform manifold approximation and projection; GAN, generative adversarial networks; DNN, deep neural network; GA, genetic algorithm.
Source
Prepared by the author.

01. Changes in social distancing policies by region included ① mandatory wearing of masks, ② restrictions on gatherings and events, ③ restrictions on attending and watching sports, ④ restrictions on the use of transportation facilities, ⑤ restrictions on physical school attendance, ⑥ restrictions on religious activities, ⑦ restrictions on on-site work, and ⑧ restrictions on multi-use facilities. Data based on OxCGRT by Son et al. (2021) were used. OxCGRT: COVID-19 Government Response Tracker provided by Blavatnik School of Government, University of Oxford (Son et al., 2021, 95).

5. Data integration and pre-processing for Pure-inTention

Eleven different types of data were used to predict trip demand according to purpose and the impact of the infectious disease: 1) trip demand according to purpose; 2) GIS road network; 3) traffic volume counts (in static) 4) population by sex, age, and region; 5) national business address book; 6) national corporate network data; 7) household credit data; 8) mobile phone-based floating population data; 9) dynamic traffic volume (in dynamic) 10) confirmed COVID-19 patient records; and 11) changes in social distancing. The purpose, source, update frequency, and provider of each data type are shown in Table 2. Mobile floating information data included statistical data of mobile population movement, provided by Statistics Korea in collaboration with SK Telecom. The Korea Transport Database (KTDB) provides seven purposes of trip: 1) returning home, 2) going to work, 3) working, 4) going to school, 5) shopping, 6) leisure activity, and 7) others. All data were classified in the units of town, township, and neighborhood, and analyzed using open sources such as R, Python, and postgreSQL.[02] Table 2

6. Experimental results and conclusion

Figure 1 is a schematic showing partial results calculated using the presented Pure-inTention model. Figure 1A shows the results of clustering based on the purpose of commuting to work. The results of clustering by city, county, district or town, township, and neighborhood without specific prior information on the region show various characteristics of each region in South Korea. Figure 1B is a scatter diagram of Model A after applying cluster 2, among different clusters derived from Figure 1A. The x-axis represents the actual trip rate confirmed by the floating population data, and the y-axis shows the trip rate predicted by Model A. As the dispersion value of each axis converges diagonally, there are no changes in the post-pandemic compared with the pre-pandemic trip pattern. The area of ① depicts where a high post-pandemic trip rate was observed, although a low trip rate was expected pre-pandemic, and the area of ② represents the opposite case. The sensitivity of the regions to changes in the social distancing policy were as follows. As shown in the graph for sensitivity of trip rate according to remote work policies of Model B in Figure 1C, the trip rate tended to decrease when the level of social distancing was low. However, as the remote work policy level increased from level 2 to level 3, the trip rate increased rapidly. The regions included in the relevant cluster (cluster 2) mainly

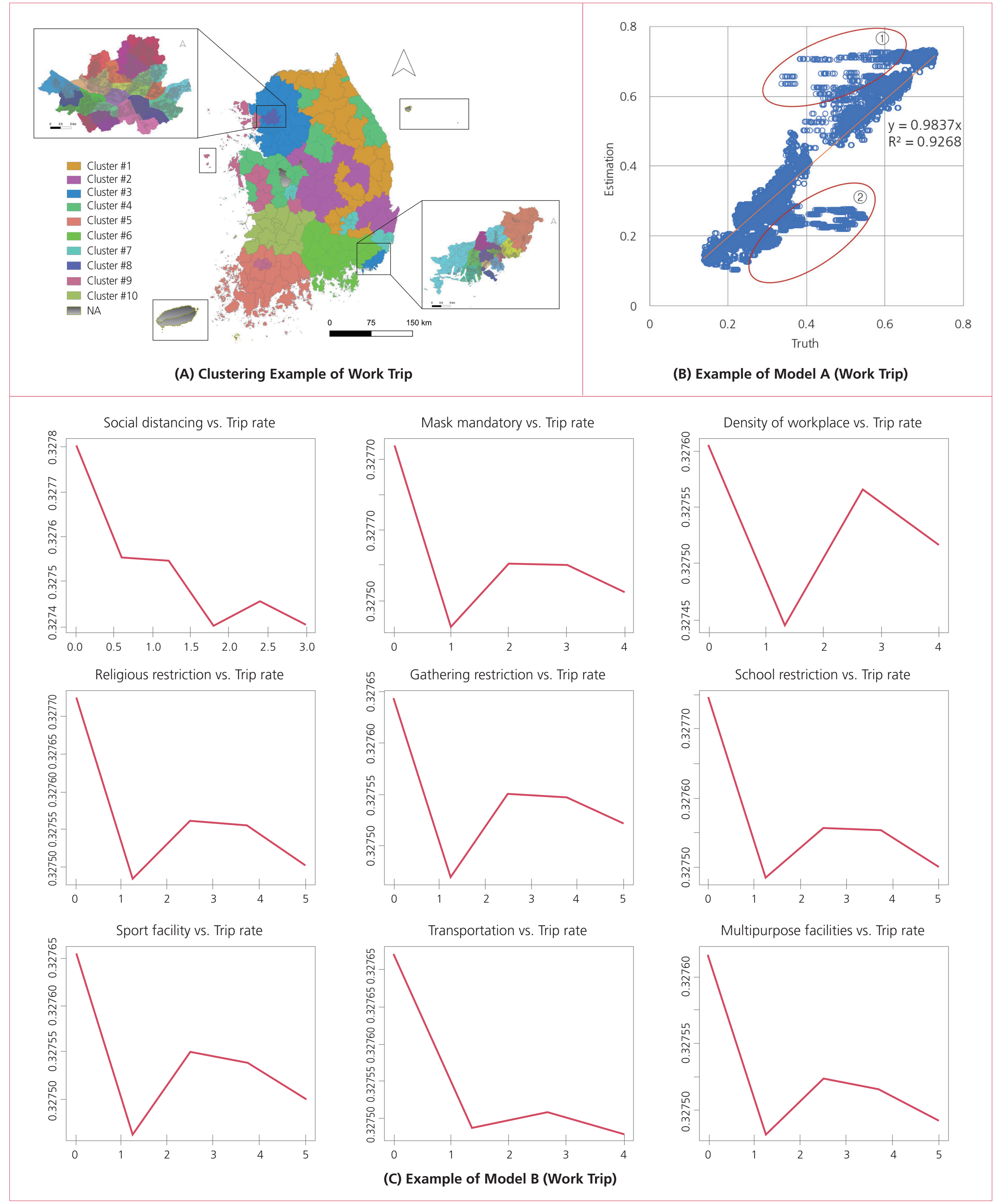
Table 2. Overview of study data

No.	Data	Purpose	Source	Data update frequency	Provider
1	Trip by purpose (NHTS)	Trip purpose (general)	KTDB	Every 5 years	Public
2	GIS network	Road network			
3	AADT	Average traffic counts	TMS	Annual	
4	Demographic	Geo-factor analysis	Statistics Korea		
5	Business information		DBRIA	Daily	Private
6	Industry information		KED		
7	Credit information		KCB	Every 3 months	
8	Movement observation from mobile data	Trip purpose (dynamic)	Statistics Korea	Hourly	Public
9	Dynamic traffic observation	Traffic counts (dynamic)	TMS	Every 30 minutes	
10	COVID-19 confirmed cases	Effect of COVID-19	KDCA	Daily	
11	Social distancing information	Policy	HoHW		

Note
NHTS, National Household Travel Survey; AADT, annual average daily traffic; KTDB, Korea Transport Database; TMS, Traffic Monitoring System; KED, Korea Enterprise Data; KCB, Korea Credit Bureau; KDCA, Korea Disease Control and Prevention Agency; MoHW, Ministry of Health and Welfare of South Korea.
Source
Prepared by the author.

02. Mobile communication population movement statistics data (https://data.kostat.go.kr, accessed February 21, 2022).

Figure 1. An example of Pure-inTention



Note (A) Clustering based on commute to work. (B) The x-axis represents the ground truth of the trip rate, and the y-axis represents the estimated trip rate according to Model A (cluster 2); (C) The x-axis represents the level of social distancing for each policy, and the y-axis represents the trip rate (cluster 2).
Source (A) Prepared by the author using QGIS; (B) Prepared by the author; (C) Prepared by the author using R.

had high numbers of wholesale and retail businesses based on corporate network information, information on the distribution of the number of trading companies in the region, and the average distance between companies in the manufacturing trading network.[03]

In this study, the Pure-inTention model presented a framework for evaluating changes in pre- versus post-pandemic preferences through deep learning and data, which were previously investigated through surveys. The Pure-inTention model also suggested the possibility of real-time nationwide monitoring. In particular, the model analyzed changes in pre- and post-pandemic trip patterns by region throughout the country, which were limited in previous studies. Figure 1

Yohan Chang
Associate Research Fellow | Manager
Geospatial Analysis & Monitoring Center | KRIHS Data Lab
ycanns@krihs.re.kr

03. The mean square error of Model A was 8.0E-06, suggesting that the model had been precisely trained.

References

- Chang Yohan, Lee Youngmin, Lim Si-yong, Park Seo hyeon, Kim Hyun kyung 2022. A Study on Deep Learning-based Purpose-driven Trip Demand Estimation Model for Post-Pandemic. Korea Research Institute for Human Settlements.
- Son Jaesoen, Chang Yohan, Oh Changwha, Im Eunsun, Hwang Myunghwa, Lee Gyoung Ju. 2021. A Study on the Development of a Spatiotemporal Simulation Model to Support an Infectious Disease Prevention Policy. Korea Research Institute for Human Settlements.
- Sheffi, Yosef. 1985. Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. MIT Press.
- Traffic Monitoring System (http://www.road.re.kr/main/main.asp, accessed July 7, 2021)
- Korea Transport Database (https://www.ktdb.go.kr/www/contents.do?key=23, accessed July 6, 2021)
- Korea Credit Bureau (http://www.koreacb.com, accessed November 11, 2021)
- Korea Enterprise Data (http://www.kedkorea.com, accessed November 11, 2021)
- Central Accident Response Headquarters. 2021. Social distancing basic quarantine measures.
- Ministry of Health and Welfare of South Korea. 2021. Revision of social distancing system.

03 Current Status and Improvement Directions of COVID-19 Outbreak Monitoring: Based on An Analysis of New Cases by *Eup-Myeon-Dong* in 2020 and 2021[01]

Myunghwa Hwang

1. Introduction

■ Changes needed in the current system of COVID-19 outbreak monitoring

Following the first confirmed case of coronavirus disease 2019 (COVID-19), the infection spread rapidly and was declared a global pandemic by the World Health Organization (WHO) in March 2020. After several waves of COVID-19 and a prolonged pandemic in South Korea, "Living with COVID-19" was announced in November 2021, increasing the autonomous authority and responsibility of local governments and individuals as well as the national government, in responding to the infectious disease. However, the current system for COVID-19 outbreak monitoring (hereafter, 'COVID-19 monitoring') still focuses on aggregate statistics of cases and deaths at the provincial and municipal levels. Thus, local governments, in charge of regional prevention policies, and the general public face difficulties in understanding the latest COVID-19 developments and risk levels in their local communities and neighborhoods.

■ Efforts needed to improve the detailedness of COVID-19 monitoring

To accept that COVID-19 is a pandemic that can continue to manifest itself—rather than regard it as a temporary infectious disease—and respond in more sustainable ways, people should be able to access and communicate each other with authorized information from one central source of the risk and current status of COVID-19 outbreaks. And this information should be detailed enough for local governments and residents, the direct actors in responding to the infectious disease, to figure out the real danger of disease transmission in their local communities. In this article we argue that improving the spatial precision or detailedness of the COVID-19 monitoring system could be the first step in initiating community-level communications of COVID-19 risks. To justify this argument, we refined the spatial unit of COVID-19 monitoring from the current *Si-Gun-Gu* (province-county) level to the *Eup-Myeon-Dong* level. We then analyzed the spatio-temporal patterns of COVID-19 cases at the *Eup-Myeon-Dong* level to show the benefits and policy implications of such attempts in COVID-19 monitoring.

2. Current status of COVID-19 monitoring in South Korea

■ From online maps and disclosure of contact tracing information of confirmed cases to sharing and text notification of case statistics by *Si-Gun-Gu*

In the early stages of the pandemic, the COVID-19 monitoring system was centered around online COVID-19 maps voluntarily managed by citizens and disclosure of contact tracing information of confirmed cases. As COVID-19 began to spread more rapidly, the national government posted outbreak trends on the COVID-19 website (http://ncov.mohw.go.kr/). This website was linked with COVID-19 websites of local governments to provide statistics on the number of confirmed

01. This article was based on "An analysis of spatio-temporal patterns of COVID-19 cases to better inform public responses to infectious diseases," a collaborative research project between KRIHS and Korea Disease Control and Prevention Agency (Hwang Myunghwa, Son Jaeseon, Lee Kunkak, 2022). The project was carried out to produce policy evidence from a geospatial perspective of better COVID-19 monitoring from September 2021 to early February 2022. Some contents of this article may not reflect the most up-to-date situations of COVID-19 and prevention policies that have changed since February 2022.

cases at the level of *Si-Do* (metropolitan cities and province). This information was also made available on Internet portals and mobile applications, such as Naver and Daum. However, the disclosure of confirmed case contact tracing posted on local government websites caused various side issues and was soon discontinued. Instead, the local government notified citizens of the statistics regarding confirmed cases in the district via disaster text messages. When the national government announced "Living with COVID-19," the COVID-19 website was revised and the COVID-19 monitoring system was changed once more. Confirmed case alert texts, which were previously sent to residents multiple times per day depending on the local government, were replaced with a daily text on confirmed cases within the district. In addition, statistics of confirmed cases by *Si-Do* began to be provided also in the spatial units of *Si-Gun-Gu*(county and district).

■ Difficult to know the infection status within local communities with the statistics of confirmed cases by *Si-Do* and *Si-Gun-Gu*

In the second half of 2021, the statistics of confirmed COVID-19 cases by *Si-Do* started to be provided in spatial units of *Si-Gun-Gu*. However, considering that the main actors of disease response in local communities were local governments and residents, they faced difficulties identifying and adequately responding to infectious disease outbreaks in local communities based on statistics by *Si-Gun-Gu*.

3. An empirical analysis of COVID-19 cases by *Eup-Myeon-Dong*

■ Analyzing the COVID-19 outbreak distribution pattern by *Eup-Myeon-Dong*

In a study by Hwang et al. (2022), which is the basis of this article, the data of 496,585 confirmed COVID-19 cases—from January 19, 2020, to December 9, 2021—were acquired with the cooperation of the Korea Disease Control and Prevention Agency to analyze the distribution of the cumulative COVID-19 crude rate in the level of *Eup-Myeon-Dong*. From the source data, the records of 398,478 confirmed cases[02] were used if each record had a valid COVID-19 diagnosis date, the reported status as a domestic infection, and a valid address enough to identify which *Eup, Myeon,* or *Dong* it belonged to. They were then aggregated by the administrative units of *Eup-Myeon-Dong* in 2020.

To analyze the pattern of the COVID-19 outbreak by the spatial units of *Eup, Myeon,* and *Dong,* the Ministry of the Interior and Safety's resident registration populations for each unit as of November 30, 2021, were used as the population at risk, and the cumulative crude incidence rate (hereafter, incidence rate)—the cumulative number of confirmed cases per 100,000 people—was calculated. During the study period, the national incidence rate was 780 cases per 100,000 people, and the mean incidence rate by *Eup-Myeon-Dong* was 629 cases per 100,000 people. In particular, 1 *Dong* in Danwon-Gu, Ansan-Si, Gyeonggi-Do; 1 *Dong* in Guro-Gu, Seoul; 1 *Dong* in Gwangsan-Gu, Gwangju; 1 *Dong* in Gangdong-Gu, Seoul; and 1 *Dong* in Dongducheon-Si, Gyeonggi-Do, had a mean incidence rate of over 5,000 cases.

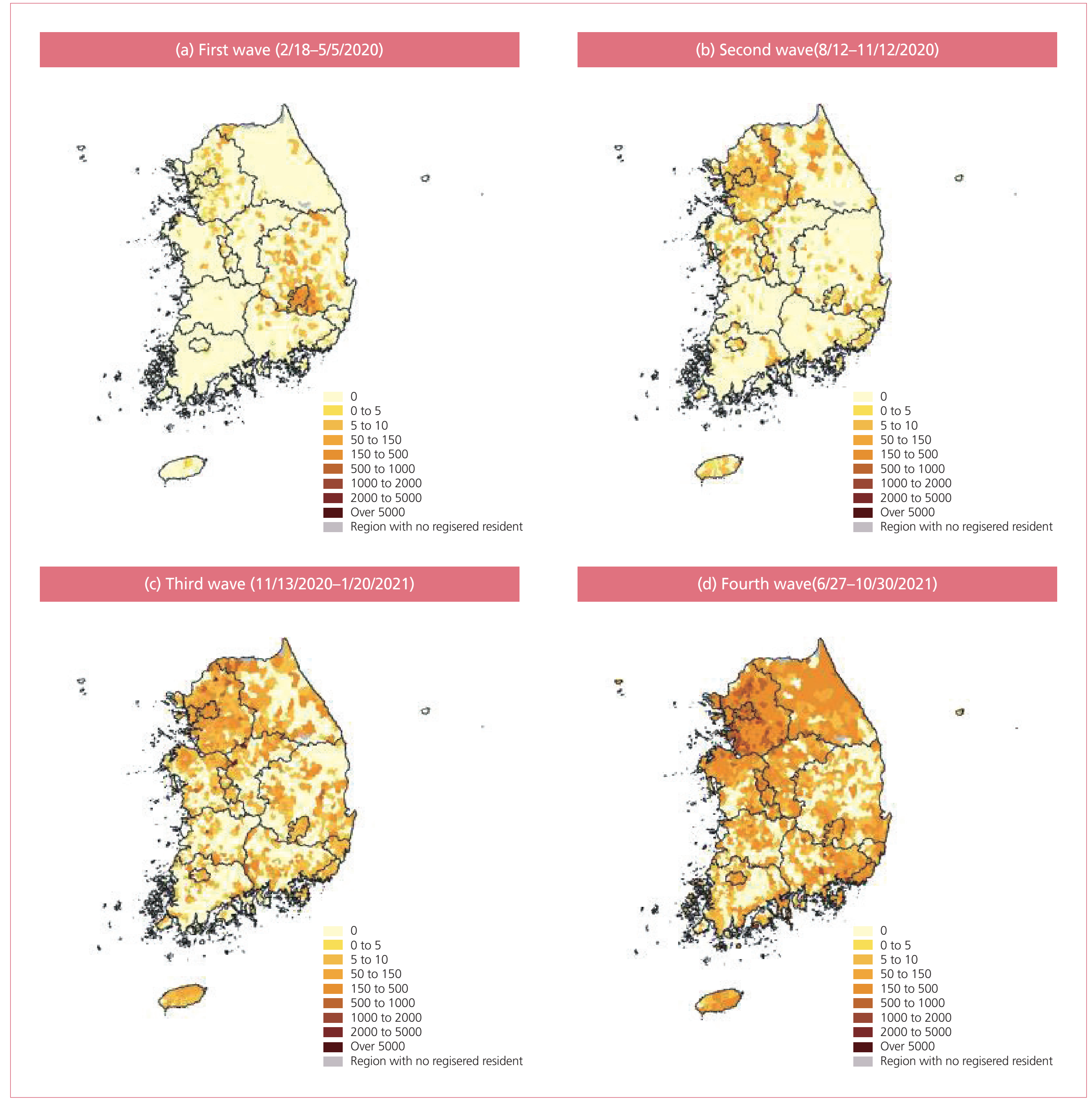
We also analyzed the distribution of confirmed cases for each wave of COVID-19 based on reports by Jang et al. (2021), Kim (2021), and other media articles. The incidence rates in the first wave were relatively high in Nam-Gu/Seo-Gu in Daegu and Cheongdo-Gun, Gyeongsangbuk-Do, due to the so-called Daegu-Si Shincheonji incidence. In the second wave, the incidence rates were high in Pocheon-Si and Yeoju-Si in Gyeonggi-Do and Suncheon-Si in Jeollanam-Do. In the third wave, the incidence rates were relatively high in Gwangsan-Gu in Gwangju, Jincheon-Gun, and Eumseong-Gun in Chungchungbuk-Do; Yeoncheon-Gun in Gyeonggi-Do; and Gimje-Si in Jeollabuk-Do. In the fourth wave, the incidence rates were high in many cities, counties, and districts within the metropolitan area (Figure 1, Table 1a). Furthermore, in the analysis of incidence rates by wave in groups of patients under the age of 10 who were exempt from mandatory vaccination in the early stages of vaccinations in 2021 (Table 1b), areas of high incidence rates were different from those for the entire population. These results indicate that high-risk areas at the level of *Eup-Myeon-Dong* varied by waves as well as by age groups. Figure 1 Table 1

02. Information on confirmed COVID-19 cases in South Korea was manually collected before home treatment was implemented and information on confirmed cases was collected through a mobile app. As a result, only approximately 79.9% of the source data could be identified by *Eup, Myeon, Dong* through address-refining and geocoding (process of converting addresses into coordinate values).

■ Identifying hot spots of concern in terms of disease spread in local communities

In addition to analyzing the incidence distribution pattern by *Eup-Myeon-Dong*, we extracted hot spots where regions of high incidence rates were spatially adjacent and the local spread of COVID-19 could be of concern for each COVID-19 wave.[03] We also examined emerging regions that showed signs of becoming new hot spots with high numbers of COVID-19

Figure 1. Distribution of cumulative crude incident rates (all age groups) of COVID-19 by *Eup-Myeon-Dong*



Source Created by the author.

03. We used the Getis-Ord Gi* statistics, a method for spatial cluster analysis.

cases immediately before each wave started, and compared the results to those identified from the hot spot analysis.[04] If an emerging region was identified as a hot spot, that region could be considered an early signal that the disease might spread locally around its nearby areas.

Our analysis showed that some emerging hot spots by *Eup-Myeon-Dong* were indicative of actual regions with high risks and concern of a local COVID-19 spread, and that a significant number of hot spots were located across several *Si*s, *Gun*s, and *Gu*s. At the beginning of the first wave, emerging hot spots were detected in Daegu and neighboring regions, as well as in border regions of Gyeongsangbuk-Do and Chungcheongbuk-Do. Some of these regions were found to be actually high-risk hot spots with high numbers of COVID-19 cases during the first wave. Emerging hot spots at the beginning of the second wave included regions in Seoul, northern Gyeonggi-Do, and Gwangju. Of these, areas in Seoul and northern and central Gyeonggi-Do were identified as belonging to the hot spots in the second wave. Additionally, the area of Seoul, Incheon, and parts of northern Gyeonggi-Do as well as the Gyeonggi-Chungnam-Gangwon border region were detected as both emerging hot spots just before the beginning of the third wave and high-risk hotspots during the same period, showing some spatial overlap (Figure 2). Although the detection of emerging hot spots by wave was not completely accurate, this method may be developed into a tool for early detection of areas with concern of local spread of infectious diseases if further efforts are made for the tuning of parameters such as the analysis period, hot spot requirements, and other analysis options. Figure 2

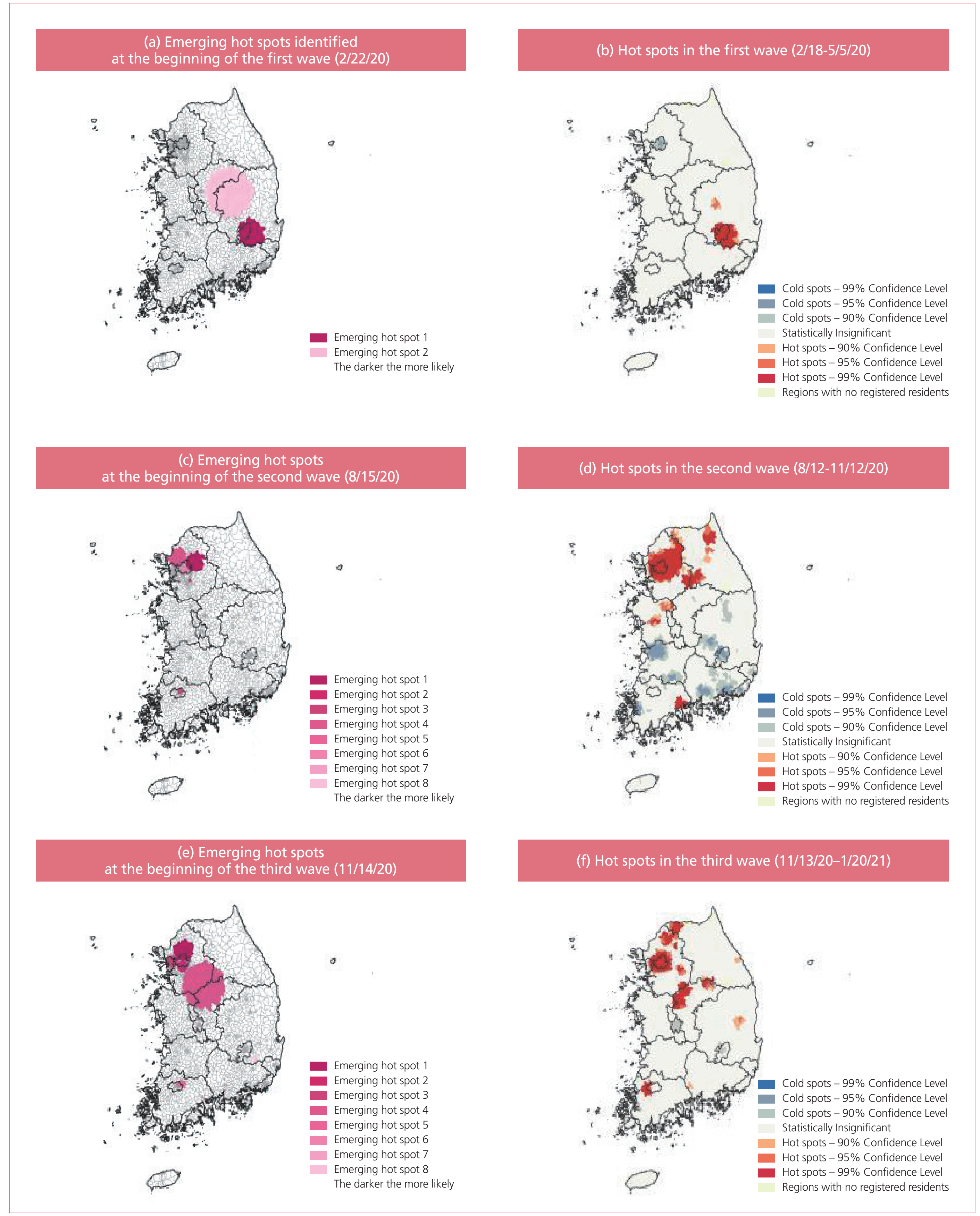
Table 1. *Eup-Myeon-Dong* units with the highest incidence rates of COVID-19

COVID-19 Waves	(a) All age groups	(b) Under the age of 10
First wave (2/18–5/5/20)	2 Dongs in Nam-Gu and 1 Dong in Seo-Gu, Daegu 1 Eup in Cheongdo-Gun, Gyeongsanbuk-Do (>1,000 cases per 100,000)	1 Dong in Nam-Gu, Daegu 1 Myeon in Yecheon-Gun, Gyeongsangbuk-do (>2,000 cases per 100,000)
Second wave (8/12–11/12/20)	1 Myeon in Pocheon-Si, Gyeonggi-Do 1 Myeon in Yeoju-Si, Gyeonggi-Do 1 Myeon in Suncheon-si, Jeollanam-Do (>500 cases per 100,000)	1 Myeon in Jeongeup-Si, Jeollabuk-Do 1 Myeon in Pocheon-Si, Gyeonggi-Do 1 Myeon in Suncheon-Si, Jeollanam-Do 1 Myeon in Gwangju-Si, Gyeonggi-Do (>1,000 cases per 100,000)
Third wave (11/13/20–1/20/21)	1 Dong in Gwangsan-Gu, Gwangju 1 Myeon in Jincheon-Gun, Chungcheongbuk-Do 1 Myeon in Eumseong-Gun, Chungcheongbuk-Do 1 Myeon in Yeoncheon-Gun, Gyeonggi-Do 1 Myeon in Gimje-si, Jeonllabuk-Do (>2,000 cases per 100,000)	1 Myeon in Yeoncheon-Gun, Gyeonggi-Do 1 Myeon in Pocheon-Si, Gyeonggi-Do 1 Myeon in Yeongwol-Gun, Gangwon-Do 1 Myeon in Hamyang-Gun, Gyeongsangnam-Do (>2,000 cases per 100,000)
Fourth wave (6/27–10/30/21)	1 Dong in Danwon-Gu, Ansan-Si, Gyeonggi-Do 2 Dongs in Guro-Gu, Seoul 1 Myeon in Ongjin-Gun, Incheon 1 Myeon in Sunchang-Gun, Jeollabuk-Do 1 Dong in Yeongdeungpo-Gu, Seoul 1 Myeon in Gwangju-Si, Gyeonggi-Do 1 Dong in Gangnam-Gu, Seoul 1 Dong in Pyeongtaek-Si, Gyeonggi-Do 1 Dong in Jongno-Gu, Seoul 2 Dongs in Siheung-Si, Gyeonggi-Do 1 Dong in Yeonsu-Gu, Incheon 1 Dong in Gangdong-Gu, Seoul (>2,000 cases per 100,000)	1 Myeon in Sunchang-Gun, Jeollabuk-Do 1 Myeon in Ongjin-Gun, Incheon 1 Dong in Danwon-Gu, Ansan-Si, Gyeonggi-Do 1 Dong in Jung-Gu, Busan 1 Dong in Guro-Gu, Seoul 1 Myeon in Gwangju-Si, Gyeonggi-Do 1 Myeon in Goseong-Gun, Gyeongsangnam-Do 1 Dong in Jongno-Gu, Seoul 1 Dong in Yeosu-Si, Jeollanam-Do (>4,000 cases per 100,000)

Source
Created by the author.

04. We used the prospective space-time scan statistic, a method for conducting a spatiotemporal cluster analysis.

Figure 2. COVID-19 hot spots in the first to third waves (all age groups)



Source Created by the author.

■ Detailed monitoring by *Eup-Myeon-Dong* helps identify areas of high-risk as well as areas of concern for local disease spread

Our findings show that increasing the spatial precision of COVID-19 monitoring to the units of *Eup-Myeon-Dong*, from the current units of *Si-Gun-Gu* enables the identification of high-risk regions within *Si-Gun-Gu*. In addition, they indicate that the new system may allow timely and early detection of areas with signs of disease spread within local communities and beyond the boundaries of local governments. As substantial efforts are ongoing to reduce the burden of disease prevention due to the prolonged COVID-19 pandemic, the new monitoring system that can detect regions of high risks or with symptoms of local disease spread would strengthen such efforts by providing useful data evidence for decision making related to allocation of quarantine resources and establishment of collaborative and cross-regional measures for disease response among local governments. Residents can also benefit from such system by referring to its results when making personal decisions with regard to taking preventive measures, such as wearing masks, or planning where to visit.

4. Conclusion: Suggestions for improved COVID-19 monitoring

■ Need community-level disease monitoring and data infrastructure for it

Experiencing the emergence and spread of COVID-19, many experts say that new infectious diseases would appear more frequently and continuously in the future. For a more effective response against the constant danger of new infectious diseases and the facilitation of social communication regarding the risk of these diseases, it is essential to improve the current monitoring system that focuses on aggregate statistics in the level of *Si-Gun-Gu*.

As discussed above, refining the spatial unit of disease monitoring to the level of *Eup-Myeon-Dong* may be an effective alternative for improving COVID-19 monitoring. This is because the main actors of disease prevention such local governments and residents concern and can affect their local communities, i.e., *Eup-Myeon-Dong*s or smaller areas, and sharing information on the spread of infectious disease in the level of local communities can help such actors monitor areas of high risks or with early signs of local disease spread and identify areas where cross-regional, collaborative measures should be taken for effective diseases response.

Detailed or precise monitoring of infectious disease outbreaks may be a cornerstone for a scientific and communication-oriented disease prevention system. Such system, however, cannot exist without an infrastructure that can collect and share data rapidly, in a standardized fashion and with location information. As address information of confirmed cases before home treatment was not collected in a standardized format, monitoring local COVID-19 cases could not be performed in a timely manner. In addition, although the local governments sought to cooperate each other for quarantine measures, sharing information with other local governments was difficult. For precise monitoring of infectious diseases in the future, the national government must preemptively establish an infrastructure that can enable the quick collection, linkage, integration, and analysis of standardized data on confirmed cases, tested individuals, and vaccinated individuals.

Myunghwa Hwang
Research Fellow
KRIHS Data Lab
mhhwang@krihs.re.kr

References

- Hwang Myunghwa, Son Jaeseon, Lee Kunhak. 2022. *An analysis of spatio-temporal patterns of COVID-19 cases to better inform public responses to infectious diseases*. Sejong: Korea Research Institute for Human Settlements.
- Hwang Myunghwa. 2022. Spatiotemporal patterns and implications of COVID-19 outbreaks by *Eup-Myeon-Dong* in 2020-2021. *KRIHS Policy Brief*. Unpublished.
- Kim Jihun. 2021. [News AS] The 5th wave is coming... what is the criteria for 'nth wave'? Hankyoreh, November 13. https://www.hani.co.kr/arti/society/health/1019161.html (accessed May 9, 2022).
- Jang Jinhwa, Kim Younghwa, Kim Yooyeon, Yeom Hansol, Hwang Insob, Park Kwangsuk, Park Youngjoon, Lee Sangwon, and Kwon Donghyok. 2021. Coronavirus Disease-19 (COVID-19) one-year outbreak major cluster infection report as of January 19, 2021, in the Republic of Korea. Public Health Weekly Report (PHWR) Vol.14, No.9:482-495.

KRIHS

KRIHS (Korea Research Institute for Human Settlements) was established in 1978 in order to contribute to the balanced development of national territory and the improvement of the quality of life of people by conducting comprehensive policy-oriented research in the efficient use, development, and conservation of territorial resources.