
Article

Using Spatial Patterns of COVID-19 to Build a Framework for Economic Reactivation

Renato Quiliche¹, Rafel Rentería², Irineu de Brito Junior³, Ana Luna¹ and Mario Chong^{1*}

¹ Universidad del Pacífico, Lima, Perú.

² Universidad Nacional Abierta y a Distancia, Bogotá, Colombia.

³ Environmental Engineering Department, Sao Paulo State University, Sao Jose dos Campos, Brazil and Graduate Program in Logistics Systems Engineering, Sao Paulo University, Sao Paulo, Brazil.

* Correspondence: m.chong@up.edu.pe

Abstract: In this article we propose an application of humanitarian logistics theory to build a supportive framework for economic reactivation and pandemic management based on province vulnerability against COVID-19. The main research question is: which factors are related to COVID-19 mortality between Peruvian provinces? We conduct a spatial regression analysis to explore which factors determines the differences in COVID-19 cumulative mortality rates for 189 Peruvian provinces up to December 2020. The most vulnerable provinces are characterized by having low outcomes of long-run poverty and high population density. Low poverty means a high economic activity that leads to more deaths of COVID-19. There is a lack of supply of a set of relief goods defined as Pandemic Response and Recovery Supportive Goods and Services (PRRSGS). These goods must be delivered in order to mitigate the risk associated to COVID-19. A supportive framework for economic reactivation can be built based on regression results and a delivery strategy can be discussed according to the spatial patterns that we found for mortality rates.

Keywords: humanitarian logistics, pandemic, economic reactivation, spatial modelling

1. Introduction

Since 2020, the novel coronavirus has caused changes in the livelihoods of the people and the nature of work. These changes motivated a set of public policies that aims to reactivate the economies while minimizing the total number of contagions to reduce mortality rates. For the case of study of Peru, the government adopted a preventive framework including early lockdowns, curfews, prolonged mobility restrictions between regions, and prohibitions to agglomeration in closed spaces (Presidencia del Consejo de Ministros, 2020a; Presidencia del Consejo de Ministros, 2020b; Poder Ejecutivo, 2020a; Poder Ejecutivo, 2020b and Ministerio del Interior, 2020). After the lockdowns, a set of subsidiary policies to population and companies (Poder Ejecutivo, 2020a; Poder Ejecutivo, 2020b and Ministerio de Economía y Finanzas, 2020), and macroeconomic policies oriented to reactivate consumption and production were implemented in order to mitigate the impacts of COVID-19 (Montoro et al., 2020).

The COVID-19 pandemic had an impact on both the directly affected people who get infected by the virus and had to waste a lot of resources in the recovery and on the indirectly affected who suffer from the scarcity of income sources due to the short-run contractions of the economy, which includes the livelihoods of healthy people affected by lockdowns. Recent literature had the objective of generating evidence of the direct and indirect impacts of COVID-19 on variables such as macroeconomic consumption, investment, production, microeconomic poverty, health, labor, among others (Brodeur et al., 2020; Acemoglu et al., 2020; Béland et al., 2020). This branch of COVID-19 literature proposed policies oriented to mitigate the impacts of COVID-19 like optimal-duration quarantines, labor-oriented policies, among others. But what is missing in general COVID-19 literature is a humanitarian

logistics perspective, which is important because it aims to reduce human suffering in the context of pandemic (Overstreet et al., 2011; Holguín-Veras et al., 2013; Leiras et al., 2014), and allows building a supportive framework for economic reactivation based on vulnerability assessment and disaster management principles (Renteria et al., 2021).

Another branch of COVID-19 literature explores the factors affecting the rates of contagions or mortality by COVID-19 (Andersen et al., 2021; Khalatbari-Soltani et al., 2020). This branch uses statistical modeling and supporting machine learning methodologies in order to determine the relationship between a set of observable factors and the COVID-19 mortality rates (Pedrosa & de Albuquerque, 2020; Rollston & Galea, 2020). This literature contributes with empirical evidence to vulnerability assessment while most of the COVID-19 related literature is based on calibrated methodologies. Empirical methodologies still represent a gap in the COVID-19 literature.

This work focuses on supporting economic reactivation and pandemic management, in the post-lockdown context. A humanitarian logistics supportive framework is built based on provinces' vulnerability against COVID-19 (Renteria et al., 2021). This vulnerability, according to the literature review, could be directly related to economic activity (Andersen et al., 2020; Khalatbari-Soltani et al., 2020; Pedrosa & de Albuquerque, 2020), socio-demographic characteristics (Rollston & Galea, 2020; Sugg et al., 2021; White & Hébert-Dufresne, 2020), the population's health insurance (Gupta et al., 2020; Benzakoun et al., 2020; Zhang & Schwartz, 2020), and health system performance (Bruno et al., 2020; Carrillo-Larco & Castillo-Cara, 2020; Chakraborti et al., 2020; Rustam et al., 2020). The vulnerability against COVID-19 is directly related to the mortality rates, which is the variable of interest in this study since mortality focuses on symptomatic population and it represents a better proxy for the demand for relief goods. The main research question is: which factors are related to COVID-19 mortality between Peruvian provinces? And with this information how a supportive framework can be implemented for underpinning the economic reactivation. We conduct a regression analysis to explore which factors determines the differences in COVID-19 cumulative mortality rates for 189 Peruvian provinces up to December 2020. In order to contribute to the empirical literature, in this work, we also test the significance of spatial interaction effects of the COVID-19 vulnerability drivers extracted from the literature (Andersen et al., 2021; Khalatbari-Soltani et al., 2020; Bruno et al., 2020). From a theoretical perspective, there are several reasons to suspect that these vulnerability drivers have spatial interaction effects, as socioeconomic variables like wealth or employment tend to have spatial spillovers (Crandall & Weber, 2004; Wardhana et al., 2017) and these could be correlated with COVID-19 mortality. Controlling these spatial effects helps to reduce the bias on the coefficients, so post estimation analysis will be more reliable. This is an important step because model predictions, and vulnerability assessment, will be used as the basis of a supportive framework for economic reactivation.

After the regression analysis, we use the results to build support for economic reactivation, which is based on the delivery of essential goods and services to face the pandemic. The type of goods will be drawn from literature and will be defined as Pandemic Response and Recovery Supportive Goods and Services (PRRSGS). We classify the demand of these essential supplies according to quartiles of the low, medium, high, and very high demand, and this classification will be used to assess economic reactivation based on the level of vulnerability. For the case of seven provinces with zero deaths, we will use an elastic-net regression to forecast their demand classification. This methodology covers the four phases of humanitarian logistics: risk mitigation, preparedness, response, and recovery (Alexander, 2002). Risk mitigation can be achieved by implementing policies based on the vulnerability drivers (seen in regression analysis), disaster preparedness policies can be developed for the zero deaths provinces, where there has been no infection yet. Response and recovery policies imply defining PRRSGS and tracing a delivery strategy based on demand estimation.

The main contribution of this work is methodological, as a humanitarian logistics perspective is missing in COVID-19 literature. This implies the proposal of an alternative framework for economic reactivation and pandemic management, which includes supportive policies that prioritize the vulnerable. This work also contributes to the empirical literature with

the estimation of spatial dependence models, a statistical modeling approach (Smith & Menis, 2020), and elastic-net regression, a conventional supervised machine learning technique. To the best of our knowledge, both applications are missing in the empirical literature.

The remainder of this study is structured as follows. Section 2 presents the relevant literature and the theoretical framework. Section 3 describes the data collection procedures and the data processing methods to estimate provinces' vulnerability to COVID-19 and the demand for PRRSGS. Section 4 outlines the main results and discusses the hypothesis. Section 5 addresses the implications of our results for policymakers. Finally, Section 6 presents the conclusions and recommendations for future research.

2. Literature review

First of all, in this literature review, we show how humanitarian logistics serves for disaster management, specifically in the case of a pandemic situation, considered as a type of disaster. Second, we reviewed some works in which the sets of determinants of cases and numbers of deaths resulting from COVID-19 were analyzed. Third, vulnerability to COVID-19 mortality is defined based on those factors, and a contribution to the definition of vulnerability based on the effects of spatial interaction is established.

2.1. A humanitarian logistics framework for pandemic assessment

Governments around the world are applying social policies oriented to mitigate COVID-19 direct and indirect impacts as an economic catastrophe (Zeckhauser, 1996 and Pyndick & Wang, 2013). Direct impacts reside in infected people, as they have to expend time and resources in their recovery. Indirect impacts arise from the changes in the operating context, the economies are contracting themselves as a consequence of pandemic outbreaks, so are the consumption and production, furthermore, the livelihoods of the people are being affected (Brodeur et al., 2020; Yoshizaki et al., 2020). Human suffering raises (Holguin-Veras et al., 2013), because, in most of the economies, the supply could not be able to respond to rising demand regarding health care services (hospitalization and intensive care units), key health care goods (oxygen, automatic respirators), and social aid (because with lockdowns, unemployment caused hunger and raised the incidence of poverty). All the aforementioned lead to an increase in deprivation costs that need to be mitigated (Holguin Veras et al., 2013). In this work, the concept of deprivation costs is emphasized.

The central aim of the humanitarian logistics discipline is to minimize the deprivation costs, or the lack of essential goods to survive the disaster aftermath (Holguín-Veras et al., 2013; Overstreet et al., 2011; Leiras et al., 2014). Then, four phases of disaster management are followed: risk mitigation, preparedness, response, and recovery (Alexander, 2002). This discipline also tries to prioritize cases where the costs of deprivation are the highest or those places where help is required immediately (Gutjahr and Fischer, 2018). For our case of study, this happens when a community faces high COVID-19 mortality because mortality is a consequence of scarcity of essential goods and services, and in general of essential supplies to survive COVID-19 symptoms.

A logistics-based humanitarian framework is especially useful for pandemic outbreaks management, due to the need to improve demand response and prioritize vulnerable populations. However, for the Peruvian case study, policymakers are not considering the implementation of this framework (Presidencia del Consejo de Ministros, 2020a). Instead, the government adopted a preventive framework including early closures, curfews, prolonged restrictions on mobility between regions, and prohibitions on crowding in closed spaces (Presidencia del Consejo de Ministros, 2020b; Poder Ejecutivo, 2020a; Poder Ejecutivo, 2020b y Ministerio del Interior, 2020). Given the restrictive nature of the first measures, it is counterintuitive that the Peruvian case is one of the most affected countries by COVID-19 due to its number of cases and deaths.

Furthermore, the humanitarian logistics conceptual framework allows us to identify and prioritize the vulnerable population based on the level of deprivation costs, which are

specially related to higher COVID-19 mortality rates. Higher mortality rates can be explained by high economic activity indicators such as employment or internal mobility. Especially in informal economies, the working population which cannot make home-office, and all the livelihoods that require face-to-face interaction with other people, are especially vulnerable (Bruno et al., 2020). COVID-19 deaths are more likely if a set of vulnerability characteristics are met by provinces (see Section 2.3 for more details). It is worth mentioning that we are emphasizing the demand analysis on the deaths rather than on the contagions because we want to avoid asymptomatic outcomes population and prioritize the human suffering.

Pandemic response and recovery framework that was applied in Peru, till the end of 2020, included a set of subsidiary policies to population and companies (Poder Ejecutivo, 2020a; Poder Ejecutivo, 2020b and Ministerio de Economía y Finanzas, 2020), and macro-economic policies oriented to reactivate consumption and production (Montoro et al., 2020). In addition to this, the humanitarian logistics proposal, for response and recovery phases, while economic reactivation is being carried out, is a supportive framework oriented to help economic reactivation by mitigating COVID-19 impacts (based on the idea that maintaining low mortality rates is required to carry an economic reactivation). This framework is based on the estimation of the demand for a set of PRRSGS at the province level, which is based on province vulnerability, so that the supply can respond to it. In Table 1 we list the PRRSGS, we also consider the goods from Alibaba (2020):

Table 1. Pandemic Recovery Supportive Goods and Services

Types	Support to response and recovery from COVID-19 pandemic (Alibaba, 2020a)	COVID-19 diagnosis & treatment and disinfection & sterilization medical equipment (Alibaba, 2020b)
Goods and services	Oxygen, automatic respirators, dexamethasone, prednisone, acetaminophen, antibiotics (azithromycin, levofloxacin), anti-clotting medication, KN95 masks, gloves, intensive care beds, alcohol, other disinfection products. Health care, emergency transport, hospitals, doctors. Funerary protocols (incineration) and transport.	Stethoscope, thermometer, sphygmomanometer, oxygen flowmeter, oxygen saturation monitor, air disinfection machine, crash cart, defibrillator, monitor, micro-injection pump, sputum elimination machine, non-invasive ventilator, invasive ventilator, continuous renal replacement therapy (CRRT), extracorporeal membrane oxygenation (ECMO), designated computed tomography, polymerase chain reaction (PCR) machine, nucleic acid detector, ultraviolet disinfection machine, anesthesia machine, ventilator circuit disinfection machine, infrared thermal imager, and forehead thermometer.

The responding supply will have a significant impact on deprivation costs and will serve to mitigate injuries for the response phase and accelerate the transition to a business-as-usual scheme regarding the recovery phase. In zero contagions/deaths provinces, estimating its possible demand will serve to preparedness policies regarding the COVID-19 pandemic.

2.2. Previous works on determinants of COVID-19 cases and deaths

Since the beginning of the pandemic outbreak, a large amount of literature has been published. One branch was oriented to explore the determinants of COVID-19 cases (Andersen et al., 2021; Khalatbari-Soltani et al., 2020; Pedrosa & de Albuquerque, 2020; Rollston & Galea, 2020; Sugg et al., 2021; White & Hébert-Dufresne, 2020; Gupta et al., 2020),

and of the COVID-19 deaths (Benzakoun et al., 2020; Zhang & Schwartz, 2020). Andersen et al. (2021) found age, disability, language, race, and employment as the determinants of COVID-19 cases for the United States. Khalatbari-Soltani et al. (2020) discuss the importance of disadvantaged socioeconomic position on COVID-19 disease and mortality, concluding that high-quality data on socioeconomic characteristics is needed to generate more evidence about this relationship. For Ceará, state of Brazil, Pedrosa & de Albuquerque (2020) found that the number of intensive care beds has spatial patterns, but they do not coincide with the ones for COVID-19 cases. Rollston & Gallea (2020) lead to the conclusion that socioeconomic status is essential to staying healthy in the context of a pandemic, thus poverty may lead to more contagions in the US. Sugg et al. (2021) found per-capita income, average household size, population density, and minority composition as significant predictors of COVID-19 cases in the nursing home in the US. White & Hébert-Dufresne (2020) found for the US that states with little tolerance for deviance from enforced rules saw faster early epidemic growth, and also found population density as a significant predictor for COVID-19 cases. Gupta et al., (2020) focused on the significance of geographical factors finding population density and topographic altitude as predictors of COVID-19 cases. Benzakoun et al. (2020) show that pulmonary embolism was the main determinant of excess out-of-hospital deaths during COVID-19 in Paris, France. Finally, Zhang & Schwartz (2020) found that population density, the proportion of elderly residents, and percent population tested are key predictors for both COVID-19 cases and deaths.

Another branch of literature has suggested the adoption of data-driven ML-based frameworks for pandemic assessment, these frameworks focused on the estimation of the expected number of COVID-19 cases or deaths (Bruno et al., 2020; Carrillo-Larco & Castillo-Cara, 2020; Chakraborti et al., 2020; Rustam et al., 2020), and prioritize the humanitarian operations for the vulnerable. However, the better/worse performance of ML methods concerning traditional applied statistical analysis is still being discussed. Among the applied ML methods, some authors (Bruno et al., 2020; Carrillo-Larco & Castillo-Cara, 2020) applied unsupervised cluster-analysis algorithms, allowing them to explore the agglomeration patterns in the data to recover categories that explain the differences between reported COVID-19 cases or deaths. Inherent Risk of Contagion (Bruno et al., 2020) was applied for municipalities and country-level cluster analysis was applied by (Carrillo-Larco & Castillo-Cara, 2020). Supervised approaches were discussed in (Chakraborti et al., 2020; Rustam et al., 2020), which are useful for making short-term forecasts of the number of COVID-19 cases or deaths. Supervised approaches included Random Forest (RF), Gradient Boosted Machine (GBM) (Chakraborti et al., 2020), Linear Regression (LR), Support Vector Machine (SVM), Least Absolute Shrinkage and Selection Operator (LASSO), and Exponential Smoothing (ES) (Rustam et al., 2020) for the country-level number of COVID-19 cases or deaths. According to the results, unsupervised approaches outperformed supervised ones. Among the most important predictors it remained: economic activity, size of the population, the prevalence of chronic illness, and environmental pollution.

Spatial modeling applied to COVID-19 outcomes has been done with scarce literature. Guliyev (2020) and Sun et al. (2020) applied spatial modeling finding that there are significant spatial interaction effects that reduce the bias in the analysis of COVID-19 determinants. At this point, our paper contributes to the empirical literature. Spatial modeling is a special case of statistical modeling literature, and we want to test for significative spatial interaction effects between COVID-19 mortality and vulnerability drivers.

2.3. Vulnerability sources of Peruvian provinces regarding COVID-19 pandemic

A spatial dependence econometric model (SDM) is proposed to investigate the factors affecting the number of COVID-19 cases and deaths in 189 Peruvian provinces. This regression model allows us to test for spatial interaction effects while we conduct a regression analysis on COVID-19 vulnerability drivers. This model is used to estimate the

demand for PRRSGS based on COVID-19 mortality rates. If spatial interaction effects are significant, the predictions of the demand will be less biased compared to the case of a simple linear model. The spatial analysis contributes to the planning of PRRSGS distribution among the provinces at the country level. If there are strong COVID-19 outcomes spatial patterns, it is possible to use them to trace a delivery strategy for PRRSGS.

Based upon the concept of deprivation costs (Holguin Veras et al., 2013), we can theoretically build up a definition of vulnerability at the province level which will serve to motivate the SDMs. This definition has seven blocks, that are built on previous literature that explored COVID-19 determinants (Andersen et al., 2021; Khalatbari-Soltani et al., 2020; Pedrosa & de Albuquerque, 2020; Rollston & Galea, 2020; Sugg et al., 2021; White & Hébert-Dufresne, 2020; Gupta et al., 2020) and ML applications (Bruno et al., 2020; Carrillo-Larco & Castillo-Cara, 2020; Chakraborti et al., 2020; Rustam et al., 2020). This definition also contains the main hypothesis that will be tested in this paper:

- First: a province is vulnerable in the post-lockdown context if it has a lower incidence of poverty, unemployment, and other indicators of bad economic performance. Poverty would be inversely related to COVID-19 deaths because high economic activity, which is the counterpart of poverty, requires internal mobility and interpersonal interaction (Engle et al., 2020). This does not mean that we should give less importance to the greater poverty provinces, as they might have a lower demand for PRRSGS. The optimal policy must give the latter equal weight in the humanitarian objective function. This proposition is valid if we are measuring long-run poverty from households, short-run measures may not serve to prove this relationship.
- Second: provinces that have more proportion of the vulnerable population in terms of age (the older), sex (the males), skin color (white), chronic diseases such as hypertension, diabetes, obesity, among others. Deprivation costs are higher for people with a high probability to die by COVID-19, concerning young and healthy people. Additionally, provinces with more overcrowding households may have worse COVID-19 outcomes, so overcrowding will be considered in this block as a demographic characteristic.
- Third: a province is exposed to high deprivations if fewer people have health insurance. In consequence, COVID-19 deaths could be more likely to happen. The short-run households' income has been drastically reduced and then slowly recovered, so there will be a shortage in the available resources to face the contagion which cannot be covered, this would lead to a greater number of deaths.
- Fourth: health system performance indicators play a key role in the COVID-19 death outcome, especially when the COVID-19 cases exceed the capacity of hospitals and other health facilities. Provinces where the health system performs bad, are more vulnerable and likely to have a greater number of deaths.
- Fifth: the COVID-19 outcome of a province may be affected by neighboring provinces with a high incidence of COVID-19 cases and deaths. These are called endogenous interaction effects.
- Sixth: considering previous exploratory works on determinants of COVID-19 cases and deaths, and the previous vulnerability definition block, there is no reason not to suspect that the COVID-19 outcome in a province will be affected by the exogenous outcomes of neighboring provinces (i.e., poverty, demographics, health systems, etc.). These are called exogenous interaction effects.
- Seventh: after considering all the information above, it is possible that the spatial dependence model still has a residual that represents all the variability of COVID-19 outcomes that could not be explained by the predictors. In this residual, there may be correlated effects as is pointed by Elhorst (2010), where COVID-19 outcomes of a province are affected by unobserved similar characteristics of the neighbors.

Empirically, we test blocks fifth, sixth, and seventh, and we estimate seven models to visualize different types of spatial interaction effects. These models or specifications passed a model selection phase where we selected the one that better fits and matches the data.

The main contribution of the framework proposed in this paper to the literature is that it configures the pandemic management in the function of vulnerability regarding COVID-19 mortality, arguing that higher deprivation costs are found where there are greater mortality rates. Our proposal tries to build a supportive framework for economic reactivation-based on-demand assessment. This framework will support the policymakers to optimize the available resources and improve the disaster management policies.

3. Materials and Methods

This section depicts the data collection and data processing methods. First, the construction of variables is explained. Second, the spatial regression models are defined. Third, the model selection phase is explained, as well as the metrics used for model selection. Fourth, the methodology for the assessment of demand is defined.

3.1. Data collection methods

For this paper, data listed in table 2 were collected from different open-access sources such as Ministerio de Salud (MINSa by its spanish acronym), the XII Population, VII Household, and III Indigenous Communities Census carried out in 2017, the Encuesta Demográfica y de Salud Familiar (ENDES by its spanish acronym), the Encuesta Nacional a Hogares (ENAHo by its spanish acronym), and the Sistema Nacional de Defunciones (SINADEF by its spanish acronym). We rescaled the data at the province level to obtain indicators about poverty, employment, education, demographic characteristics, the prevalence of chronic illness such as diabetes, hypertension, the prevalence of obesity, health insurance, and health system performance indicators. The number of COVID-19 deaths was recovered from MINSa (2020) database (from March 06, 2020 to Dec 12, 2020). The deaths are reported by familiars and also confirmed by autopsy. The quality of these data to represent the real differences between provinces is not the best, however, we show that the data are reliable for the analysis when interpreting the empirical results. Table 2 describes all the variables included in this study.

Table 2. Description of variables

Name	Description	Source
Logdeaths1000	The logarithm of the cumulative number of COVID-19 deaths per 1000 inhabitants till December 12 th of 2020	MINSa (2020)
Composite_Poverty_Index	The average index of poverty estimated by the Multiple Correspondence Analysis (MCA) methods	Census* (2017)
Employed	People employed per 1000 inhabitants	S.A. ¹
Secondary_Educ	People having a complete secondary education per 1000 inhabitants.	S.A.
Vulnerable_pop	Elderly people (>65 years) per 1000 inhabitants	S.A.
LogPD_1000	The logarithm of the population density measured as 1000 inhabitants per km ² (urban area)	S.A.
Males_1000_Inhab	Males per 1000 inhabitants	S.A.
White_1000_Inhab	White people per 1000 inhabitants	S.A.
Black_1000_Inhab	Black people per 1000 inhabitants	S.A.
Assian_1000_Inhab	Asian people per 1000 inhabitants	S.A.
Overcrowding	The average proportion of households with more members than rooms	S.A.
Natural_Region1	Province is located in the Coastal region	S.A.
Natural_Region2	Province is located in the Highlands region	S.A.
Natural_Region3	Province is located in the Jungle region	S.A.
Hypert1	People with hypertension measured by results on differential pressure per 1000 inhabitants	S.A.

Name	Description	Source
Hypert2	People with diagnosed hypertension per 1000 inhabitants	S.A.
Diabetes	People with diagnosed diabetes per 1000 inhabitants	S.A.
Obesity	People with obesity by body mass index per 1000 inhabitants	S.A.
Chronic	People with a chronic illness (COPD, diabetes, hypertension, etc.) per 1000 inhabitants	ENAHO* (2019)
Health_Insurance	Number of people with health insurance per 100 inhabitants	Census* (2017)
Life_Expectancy	The average life between 2017 and 2019 before the COVID-19 pandemic	S.A.
Days_Till_Attended	The average days till medical attention	ENAHO* (2019)
SD_DTA	The standard deviation of days till medical attention	S.A.
Travel_Time	The average hours of travel time to health facility	S.A.
SD_TTtHFH	The standard deviation of hours of travel time to health facility	S.A.
Waiting_Time	The average hours of waiting time till attention in health facility	S.A.
SD_WT4AH	The standard deviation of hours of waiting time till attention in the health facility.	S.A.

¹ S.A.: same as above. Source: MINSA (2020), CENSO (2017), ENDES (2019), ENAHO (2019) and SINADEF (2019) (*) data recovered by National Institute of Statistics and Informatics (INEI in Spanish).

3.2. Data processing methods

The SDM to be used has to be selected from a family of cross-section spatial dependence models described in Elhorst (2010) and Kelejian-Prucha (1999). This family of models arises from different ways in which spatial patterns can be affecting a dependent variable outcome: the endogenous interaction effects, the exogenous interaction effects, and the correlated effects. There are a total of seven models, including the single ordinary least squares model (OLS) which is a sub-case with none of the effects listed above (these models are estimated in literature for spatial dependence models application). From the most general aspect to the particular one, these models are defined as follows:

$$SDM (1,1,1): Y_i = \beta X_{ik} + \theta W X_{ik} + \lambda W Y_i + v_i; v_i = \rho W v_i + \epsilon_i \quad \text{Manski model,} \quad (1)$$

$$SDM (1,0,1): Y_i = \beta X_{ik} + \lambda W Y_i + v_i; v_i = \rho W v_i + \epsilon_i \quad \text{Kelejian-Prucha model,} \quad (2)$$

$$SDM (1,1,0): Y_i = \beta X_{ik} + \theta W X_{ik} + \lambda W Y_i + v_i \quad \text{Spatial Durbin model,} \quad (3)$$

$$SDM (0,1,1): Y_i = \beta X_{ik} + \theta W X_{ik} + v_i; v_i = \rho W v_i + \epsilon_i \quad \text{Spatial Durbin Error model,} \quad (4)$$

$$SDM (1,0,0): Y_i = \beta X_{ik} + \lambda W Y_i + v_i \quad \text{Spatial lag model,} \quad (5)$$

$$SDM (0,0,1): Y_i = \beta X_{ik} + v_i; v_i = \rho W v_i + \epsilon_i \quad \text{Spatial error model,} \quad (6)$$

$$SDM (0,0,0): Y_i = \beta X_{ik} + v_i \quad \text{OLS model,} \quad (7)$$

In the equations from (1) to (7), Y_i represents a vector of $N \times 1$ observations on the dependent variable which is the number of deaths by COVID-19. First, on the right side X_{ik} is a matrix of N observations for the K exogenous variables of the model and β is the $K \times 1$ vector of parameters related to the K exogenous regressors. Second, WX_{ik} is the matrix of spatially lagged exogenous variables pre-multiplied by the vector θ of $K \times 1$ parameters associated with these spatially lagged variables (to account for exogenous interaction effects). Third, λWY_i represents the spatially lagged dependent variable, with the scalar λ (called the autoregressive parameter) pre-multiplying the WY_i term, which is a $N \times 1$ vector of observations on the spatially lagged dependent variable (to account for endogenous interaction effects). Finally, v_i is an error, which can be correlated with its spatial lag as it is represented by equations (1), (2), (4), and (6), ρ is a scalar parameter that accounts for the spatial correlation of the error term. Finally, ϵ_i is the idiosyncratic stochastic error, which is assumed to be a white noise that captures all the remaining variability of the dependent variable and is uncorrelated with other variables in the model.

The W matrix represents a spatial weighting matrix that is important to test if the spatial interaction effects are significant. This matrix provides information about the spatial units that in this work are the provinces. A spatial weights matrix, formally, is a $N \times N$ positive symmetric matrix which weights named w_{ij} define the spatial relations among locations and, therefore, determine the spatial autocorrelation statistics (Zhou & Lin, 2008).

The following equation (8) has been used for defining spatial weight matrix W :

$$W = \begin{pmatrix} w_{11} & \dots & w_{1n} \\ \vdots & \ddots & \vdots \\ w_{n1} & \dots & w_{nn} \end{pmatrix}, \quad (8)$$

$$\text{Where } w_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are contiguous} \\ 0 & \text{if } i \text{ and } j \text{ are not contiguous} \end{cases}$$

Nevertheless, there are several ways of defining W matrix, we list three of them that are based on Euclidian distance: Queen's criteria (contiguity), Gabriel criteria, and K-nearest neighbors' criteria (Guliyev, 2020 and Sun et al., 2020). We based our selection of W on the maximization of the value of the Moran Statistic that is estimated with a W definition (Sun et al., 2020). The Moran Statistic is a measure of spatial autocorrelation that lies in the interval $[-1,1]$, but its significance determines the level of observable spatial autocorrelation among a variable. We found that Queen's criteria maximize the significance of Moran statistic, and thus this is the optimal way to define the W matrix (See appendix 1). We will focus on the spatial patterns that can be measured with the data.

Equations from (1) to (7) are estimated with maximum-likelihood methods assuming that $\epsilon_i \sim N(0, \sigma^2 I)$. This implies that the distribution of COVID-19 mortality is normal. The parameters are obtained maximizing the following equation (9):

$$\ln L(Y|\psi, \lambda, \rho, \sigma^2) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) + \ln|I - \lambda W| + \ln|I - \rho W| + \frac{1}{2\sigma^2} (\epsilon' \epsilon), \quad (9)$$

Where $\psi = (\beta, \theta)$, and errors $\epsilon = F(\psi, \lambda, \rho)$ are obtained from the reduced form of models from (1) to (7). The terms $\ln|I - \lambda W|$ and $\ln|I - \rho W|$, can appear or disappear in the expression according to the model formulation. Definition of errors and the above-mentioned terms changes according to model; for example, if we estimate equation (6) spatial error model, the errors will be $\epsilon = F(\psi, \rho)$ with $\psi = (\beta)$ and we only maintain the term $\ln|I - \rho W|$ in the likelihood equation. Equation (9) is known as the unconcentrated log-likelihood function and is discussed in Lee (2004).

A disadvantage of SDMs is that they do not support out-of-sample predictions, which will be the case for the seven provinces. Data is available for 189 provinces, out of

a total of 196 provinces in Peru. The missing provinces are considered to contribute zero cases or deaths. However, this hypothesis is questionable because, in some of these provinces, there are high indicators of economic activity (for example, Cañete with a value of 38.60 in the composite poverty index that reflects low poverty below the average of 56.81). Therefore, for the estimation of the provinces with missing data, we use Elastic-Net Regression (ENR), a machine learning tool that comes as close as possible to linear models of spatial dependency. The following equation (10) shows the parameter estimation method and the objective function for ENR:

$$\beta = \min_{\beta} (||Y - X\beta||^2 + \lambda_2 ||\beta||^2 + \lambda_1 ||\beta||) \quad (10)$$

Where λ_1 is the parameter for the term associated to Least Absolute Shrinkage and Selection Operator (LASSO) and λ_2 to the Tikhonov regularization associated with ridge regression. The parameters λ_1 and λ_2 can take any value from zero to infinity, and they measure the weight of the penalization for irrelevant or correlated regressors, respectively (Ogutu et al., 2012). The method sacrifices bias to get a lower variance estimator that has been proven to outperform low bias over-fitted estimators in the context of out-of-sample predictions. The selection of λ_1 and λ_2 penalization size and the ratio will be made by repeated k-fold cross-validation (with $k = 4$), maximizing the out-of-sample predictive power measured with mean absolute error.

In order to qualitatively estimate the demand of PRRSGS, equations (1) to (7) are estimated, then we selected the model that better fit the data based upon the following metrics (Ogutu et al., 2012) (the less-biased model, considering only the 189 non-zero provinces that were part of the estimation of equations 1 to 7):

$$MF R^2 = 1 - \frac{LL_m}{LL_0} \quad (11)$$

$$R^2 = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \quad (12)$$

$$AIC = -2\text{Log}L + 2q, \quad (13)$$

$$BIC = 2\text{Log}(L) + q\text{Log}(N), \quad (14)$$

$$WT = (Rb - r)'(RVR')^{-1}(Rb - r), Rb = \begin{pmatrix} \theta \\ \rho \end{pmatrix} \text{ and } r = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (15)$$

$$RMSE = \sqrt{\sum_{i=1}^N \frac{(\hat{y}_i - \bar{y}_i)^2}{N}} \quad (16)$$

This information is complemented with Elastic-Net Regression outcomes, this model does not pass for model selection phase, but it passes through repeated k-fold cross-validation to maximize its predictive power.

3.3. Demand assessment

Once the model selection phase is carried out, the one that better fitted the data was used to make predictions. Within the sample, predictions were used as proxies for the demand of PRRSGS, and out-of-sample predictions used with ENR tool are employed as forecasts of demand for those provinces that have zero contagions or deaths (eq. 10). Data is visualized spatially, and we report results for the best models.

Demand will be assessed by quartiles of the number of predicted COVID-19 mortality rates. This implies that demand assessment is based on vulnerability as predicted

mortality rates are a function of a set of vulnerability drives listed in section 2.3. We can distinguish four groups. This categorization of demand helps to trace a delivery strategy and to identify vulnerability patterns in the space, as mortality is ranked from low to very high categories:

- $COVID19_{Q1}$: Low demand
- $COVID19_{Q2}$: Medium demand
- $COVID19_{Q3}$: High demand
- $COVID19_{Q4}$: Very high demand

In the end, every province will have an established demand, so that a distribution strategy for the 196 provinces can be discussed. In the results and discussion section, we focus on spatial insights and their contribution to the humanitarian logistics framework. In managerial and research implications we summarize and discuss the main results, and their utility at the moment of tracing a humanitarian logistics supportive framework strategy for pandemic management.

The methodology first does a regression analysis using the SDM in equations from (1) to (7) and (10), then it passes the model selection phase using equations from (11) to (16). After the model selection phase, we use the best SDM specification for predicting the level of demand of PRRSGS and, for provinces with zero deaths, we use estimates from equation (10) to forecast demand. Finally, with the post estimation results and demand quartiles classification, we discuss a delivery strategy at the national level using provinces as the unit of analysis.

4. Results

The COVID-19 mortality rates have been log-transformed to draw a normal distribution for it. Figure 1 shows a box plot with a swarm plot and a kernel density of the log-transformed mortality rates.

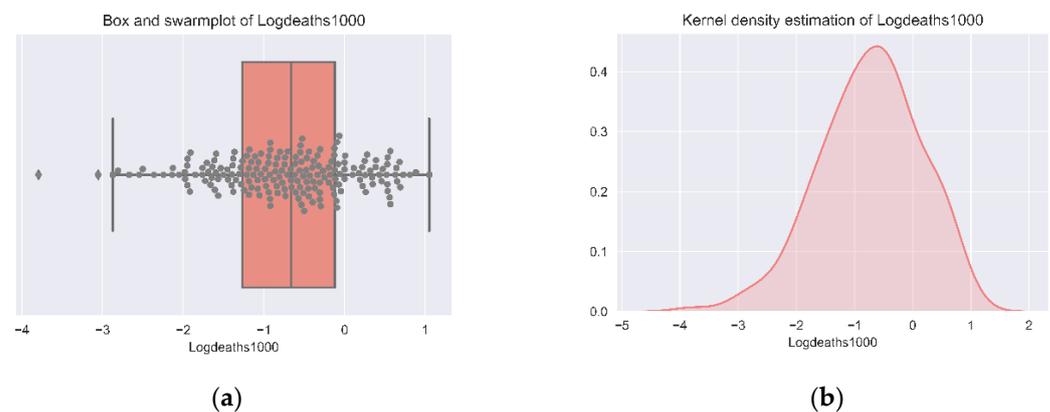


Figure 2. Distribution plots of Logdeaths1000. (a) Boxplot of Logdeaths1000. (b) Kernel density plot of Logdeaths1000.

We observe that, on average, provinces face less than one death per 1000 inhabitants. This is consistent with the mortality rates from international estimations for COVID-19. Although the mortality rates are low, they can vary a lot, and this is why it is important to understand the factors behind those differences. For this purpose, we propose a set of variables listed in Table 2 in the previous section. Figure 2 shows a plot of the correlation matrix between the variables that explain the differences between COVID-19 mortality rates.

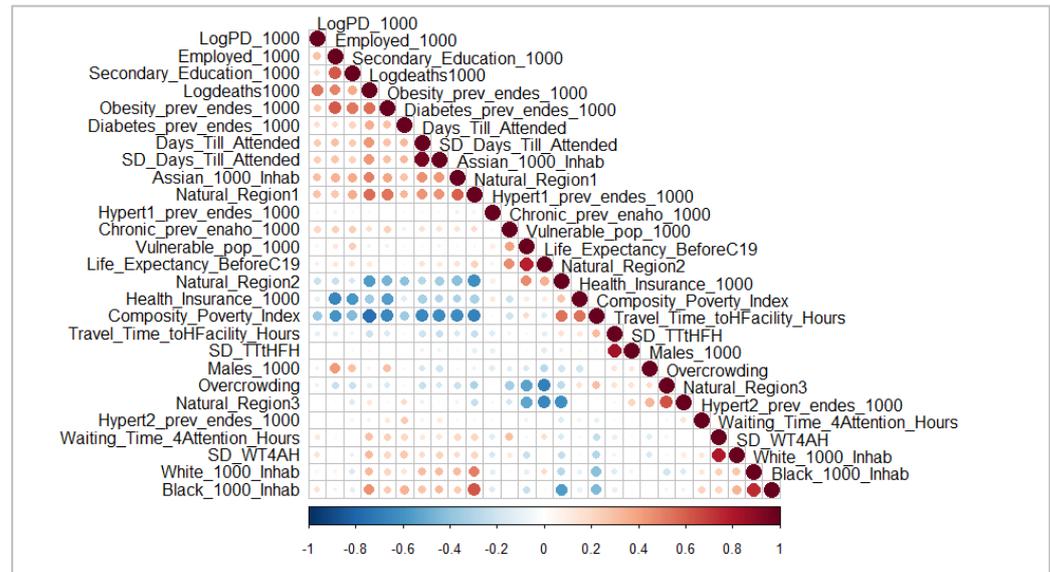


Figure 3. Correlation matrix of variables.

In Figure 2, blue circles represent the negative correlation and red ones represent the positive correlation. The size and transparency of the circles represent the magnitude of the correlation between the variables. The mortality rates (Logdeaths1000) are, on the one hand, negatively related to belonging to the highlands region (Natural_Region2), health insurance (Health_Insurance_1000), and poverty (Composite_Poverty_Index). On the other hand, they are positively related to belonging to the coastal region (Natural_Region1), presence of Asian population (Assian_1000_Inhab), presence of Black population (Black_1000_Inhab), the average days till a patient is attended (Days_Till_Attended), the inequity of the days for medical attention (SD_Days_Till_Attended), the average waiting time for attention in a health facility (Waiting_Time_4Attention_Hours), the inequity of waiting time for attention (SD_WT4AH), the prevalence of diabetes (Diabetes_prev_endes_1000) and the log population density (LogPD_1000).

The long-run poverty (Composite_Poverty_Index) variable is related to many other variables in the model. It is negatively related to the log population density (LogPD_1000), employment (Employed_1000), human capital with secondary education (Secondary_Education_1000), log of mortality rates (Logdeaths1000), obesity prevalence (Obesity_prev_endes_1000), the average days till a patient is attended in a health facility (Days_Till_Attended), the inequity in time for patient attention (SD_Days_Till_Attended), the Asian population (Assian_1000_Inhab), and the belonging to the coastal region (Natural_Region1), and its negative related to belonging to highlands region (Natural_Region2) and being health insured (Health_Insurance_1000). This could lead to a possible multicollinearity problem. Nevertheless, the mean-variance inflation factor across all variables after an OLS estimation is 4.96, which reflects low multicollinearity. The following Table 3 shows the descriptive statistics for all the variables considered in the paper:

Table 3. Descriptive statistics

Variable	Mean	S.D.	Min	Max	Median
Logdeaths1000	-0.74	0.87	-3.80	1.05	-0.67
Employed_100	371.12	72.03	140.31	589.99	379.94
Males_1000_Inhab	501.72	19.55	471.90	604.85	498.97
Vulnerable_pop	199.71	40.98	83.02	350.88	197.91
Health_Insurance	811.87	93.86	488.04	958.34	817.18
Secondary_Educ	313.98	52.78	186.49	436.08	314.60
Life_Expectancy	63.34	6.49	35.09	74.81	65.03
Chronic	364.96	96.99	117.03	630.40	362.31
Composite_Poverty_Index	56.62	12.17	25.71	78.74	60.63
White	32.24	25.05	0.71	119.25	27.29
Asian_1000_Inhab	0.18	0.40	0.00	3.32	0.04
Black_1000_Inhab	21.10	24.99	0.00	102.08	12.16
Hypert1	434.69	80.36	96.14	724.68	430.87
Hypert2	89.86	44.14	0.00	277.53	84.54
Diabetes	25.32	25.18	0.00	171.49	21.70
Obesity	184.44	94.21	0.00	514.62	173.56
Days_Till_Attended	1.06	1.79	0.01	12.97	0.34
SD_DTA	3.67	5.10	0.00	35.44	1.49
Travel_Time	0.65	0.45	0.20	4.37	0.53
SD_TTtHFH	1.61	2.84	0.16	30.70	0.82
Waiting_Time	0.11	0.05	0.02	0.28	0.11
SD_WT4AH	0.41	0.16	0.05	1.21	0.43
logPD_1000	-0.31	3.33	-9.85	7.70	-0.15
Overcrowding	17.24	7.85	7.54	53.62	14.71
Natural_Region1	0.19	0.39	0.00	1.00	0.00
Natural_Region2	0.62	0.49	0.00	1.00	1.00
Natural_Region3	0.19	0.39	0.00	1.00	0.00

After running an OLS regression, we test for spatial dependence in the residuals using a contiguity spatial weighting matrix using the residuals. The Moran's standard deviates statistic is 5.0581 with the lowest p-value between other spatial weighting matrices (see Table 4 for results with other spatial weighting matrices), rejecting the null hypothesis that states that the errors are independently and identically distributed with 99% of confidence and a p-value of 0.000, favoring the hypothesis of spatially correlated errors.

Table 4. Moran's standard deviate statistic and significance (p-values)

Spatial weighting matrix	Moran's I standard deviate	Significance (p-value)
Queen's contiguity criteria	5.058	4.23E-07
K-nearest neighbors (K=1)	2.864	4.18E-03
K-nearest neighbors (K=2)	3.613	3.02E-04
K-nearest neighbors (K=3)	3.648	2.64E-04
K-nearest neighbors (K=4)	3.874	1.07E-04

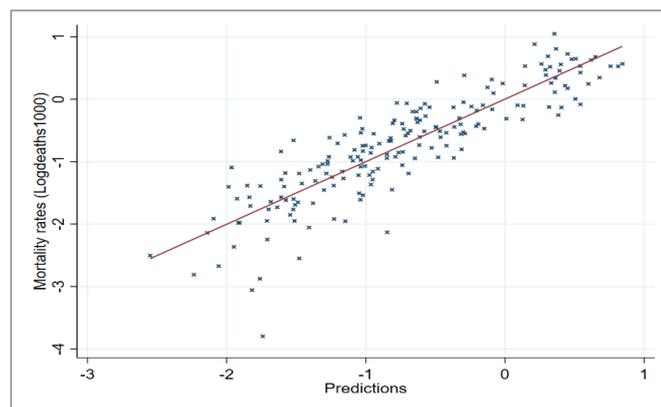
We proceed with the results of the SDMs, in Table 5 we summarize the main ones that serve for the model selection. We select the one with a lower bias for in-sample predictions. For out-of-sample prediction, we use an Elastic Net regression, which minimizes the variance, sacrificing some bias to get lower variance predictions.

Table 5. Models for COVID-19 deaths (per 1000 inhabitants)

Model	SDM (1,1,1)	SDM (1,0,1)	SDM (1,1,0)	SDM (0,1,1)	SDM (1,0,0)	SDM (0,0,1)	OLS (0,0,0)	ENR
MF R ²	0.67	0.53	0.67	0.67	0.53	0.53	0.52	N.E.
R ²	0.75	0.72	0.75	0.75	0.73	0.72	0.74	0.68
AIC	443.80	283.20	267.30	266.80	281.90	281.90	287.00	N.E.
BIC	443.80	377.20	439.1	438.60	372.70	372.70	374.60	N.E.
WT	98.63*	11.95*	88.16*	99.55*	7.138*	18.75*	N.E.	N.E.
RMSE	0.27	0.28	0.27	0.27	0.28	0.28	0.27	0.30

Where N.E.: Not estimated. First, the variables used in OLS regression produced a McFadden pseudo-R² of 0.52 and an R² of 0.74. These results support the hypothesis of the positive relationship between vulnerability against COVID-19 and the mortality rates. Even though the lowest AIC and BIC outcomes are for the SDM(1,1,0) and SDM(0,1,1), they have a higher bias than OLS, as they respectively have 0.275 and 0.273 outcomes in RMSE versus 0.269. Furthermore, RMSE is reduced in SDM(1,1,1) to 0.267. This means that SDM(1,1,1) is the best candidate for the model that better fits the data. Although SDM(1,1,1), SDM(1,1,0) and SDM(0,1,1) are equivalent in McFadden pseudo-R², R², and RMSE, we must consider that the spatial dependence in residuals is strongly significant, so instead of basing the model selection decision on AIC BIC, it's better to select the model with higher significance on spatial terms, this is measured by the Wald Test for significance of spatial terms. In consequence, the best model that better fits the data is the SDM(1,1,1).

We consider the possibility that residuals vary in a non-random way. Specifically, we believe that the dataset suffers from heteroscedasticity. In order to test it, we used the OLS model residuals. The Breusch-Pagan test chi-squared statistic, subject to one degree of freedom, is 16.69, which rejects the null hypothesis of homoscedasticity with 99% of confidence and a p-value of 0.000. The following Figure 3 plots the mortality rates versus the predictions of the SDM(1,1,1) model, it will serve to confirm the presence of heteroscedasticity:

**Figure 3.** Linear fit plot of real mortality rates versus predictions.

We observed that provinces with lower mortality rates have a higher variance in the residuals than other provinces. Furthermore, the variance for other provinces seemed to be constant. For this reason, we report robust standard errors for the individual significance of coefficients in the SDM(1,1,1) model. The following Table 6 shows the regression results:

Table 6. Regression results (coefficients)

VARIABLES	SDM(1, 1, 1)			ENR	
	Logdeaths1000	S.E.	W	S.E.	Logdeaths1000
Employed_100	0.001	(0.001)	-0.005*	(0.003)	0.000
Males_1000_Inhab	-0.001	(0.003)	-0.002	(0.005)	0.002
Vulnerable_pop	0.001	(0.002)	0.003	(0.005)	0.004
Health_Insurance	-0.001	(0.001)	0.001	(0.002)	0.001
Secondary_Educ	0.002	(0.001)	0.005	(0.003)	0.001
Life_Expectancy	0.018	(0.012)	0.006	(0.035)	0.000
Chronic	-0.000	(0.000)	-0.002*	(0.001)	0.000
Composite_Poverty_Index	-0.031***	(0.009)	0.009	(0.024)	-0.048
White_1000_Inhab	-0.006*	(0.003)	0.024***	(0.006)	-0.002
Assian_1000_Inhab	-0.076	(0.122)	-0.430	(0.387)	0.000
Black_1000_Inhab	0.004	(0.003)	-0.020**	(0.008)	0.007
Hypert1	-0.000	(0.000)	0.000	(0.001)	0.000
Hypert2	0.000	(0.001)	0.002	(0.003)	-0.001
Diabetes	-0.000	(0.002)	0.007	(0.004)	0.000
Obesity	0.000	(0.001)	0.002	(0.002)	0.001
Days_Till_Attended	-0.021	(0.061)	-0.117	(0.186)	0.000
SD_DTA	0.008	(0.020)	0.059	(0.066)	-0.003
Travel_Time	-0.244	(0.162)	-0.083	(0.528)	0.000
SD_TTtHFH	0.051**	(0.025)	0.046	(0.080)	0.000
Waiting_Time	1.498	(1.450)	-4.008	(4.174)	0.000
SD_WT4AH	0.219	(0.391)	1.434	(1.083)	0.000
LogPD_1000	0.077***	(0.012)	0.131***	(0.050)	0.053
Overcrowding	0.019**	(0.008)	-0.043***	(0.015)	0.019
Natural_Region1	(base)		(base)		0.000
Natural_Region2	-0.056	(0.187)	-0.483	(0.460)	0.000
Natural_Region3	0.577**	(0.259)	-0.238	(0.525)	0.000
Logdeaths1000			0.109	(0.100)	N.E.
e.Logdeaths1000			0.305**	(0.151)	N.E.
var(e. Logdeaths1000)			0.190***	(0.027)	N.E.
Constant	1.468	(1.524)			0.403

Robust standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

The results can be summarized as follows:

Poverty (Composite_Poverty_Index): there is a strong negative relationship between long-run poverty and COVID-19 mortality rates. While economic reactivation is being carried out, provinces that had a high level of economic activity, which is related to lower poverty and is captured in the index, are especially vulnerable against COVID-19 mortality. After the short-run shock, the economy tends to return to its long-run level, this is the main reason behind this sign. Poor provinces are vulnerable to the indirect effects of a pandemic: unemployment, shortage in available resources, among others. This insight favors the proposal of an implementation of a reactivation framework rather than a

lockdown: provinces are returning to their long-run level economic activity, and this must be done safely. There is no evidence that poverty produces exogenous interaction effects on mortality rates.

Population density (LogPD_1000): there is a strong positive relationship between population density and COVID-19 mortality rates. Higher mortality is found in provinces with higher population density. In low population density provinces or rural provinces, the health system may be inexistent, but since the data used for this analysis represent the initial period of COVID-19 with the limited spread between provinces we do not observe these effects, but the opposite that is characterized by high population density provinces where infection arrived earlier. Future deeper research is needed to better explain this relationship. There is marginal evidence that population density produces positive exogenous interaction effects on mortality: increases in population density, in neighboring provinces, can lead to higher mortality rates, but the significance of this effect is not sufficient to reject the hypothesis of no effect at all.

Overcrowding in households (Overcrowding): is intuitive that there is a strong positive relationship between overcrowding in households and COVID-19 mortality rates. Overcrowding is a serious issue in less developed countries, and it is a consequence of the informal settlement, bad birth control policies, among other factors. Mortality is magnified by the fact that more individuals are living in a household than the number that is normal given the number of bedrooms. In order to improve pandemic management, the government should minimize the contagion not only outside, but inside the household. Nevertheless, there is strong evidence that overcrowding produces negative exogenous interaction effects on mortality: increases in overcrowding, in neighboring provinces, can lead to lower mortality rates. The indirect effect and the direct effect are counteracted, but the indirect effect is greater than the direct effect (see Table 4).

Population ethnic composition (White_1000_Inhab and Black_1000_Inhab) produces significant effects on mortality rates. These effects are spatial interaction effects. An increase in the white population in neighboring provinces would lead to higher mortality rates in a province, on the other side, an increase in the black population in neighboring provinces leads to lower mortality rates in a province. These effects are consistent with the literature and the spatial insight is useful for the management of the pandemic.

ENR validates effects for poverty, population density, overcrowding, vulnerable population, health insurance, white population, black population, hypertension2 (second way of measuring diagnosed hypertension), obesity, and inequity in health. However, the size of the effect for the vulnerable population, health insurance, white population, black population, hypertension2, obesity, and inequity in health is relatively small (i.e., $0.001 < \text{effect} < 0.01$). These results prove the robustness of the SDM(1,1,1). The following Table 6 shows the direct and indirect (from spatial spillovers) impacts predicted by the model:

Table 6. Direct, indirect and total effects predicted after SDM(1,1,1)

Logdeaths1000	Direct	Indirect	Total
Employed_100	0.001	-0.001	0.001
Males_1000_Inhab	-0.001	-0.007	-0.007
Vulnerable_pop	0.001	-0.001	0.001
Health_Insurance	-0.001	0.003*	0.002
Secondary_Educ	0.000	0.003	0.003
Life_Expectancy	0.013	0.001	0.014
Chronic	0.000	-0.001	0.000
Composite_Poverty_Index	-0.029***	0.008	-0.021**

Logdeaths1000	Direct	Indirect	Total
White_1000_Inhab	-0.007***	0.011*	0.004
Assian_1000_Inhab	0.038	0.247	0.285
Black_1000_Inhab	0.005*	-0.009	-0.004
Hypert1	0.000	0.001	0.001
Hypert2	0.000	-0.001	0.000
Diabetes	0.000	0.006	0.006
Obesity	0.000	0.003	0.003
Days_Till_Attended	-0.048	0.104	0.055
SD_DTA	0.017	-0.047	-0.031
Travel_Time	-0.218	-0.693	-0.911
SD_TTtHFH	0.045*	0.106	0.150*
Waiting_Time	-0.108	-1.800	-1.908
SD_WT4AH	0.672*	0.511	1.183
LogPD_1000	0.079***	0.038	0.117***
Overcrowding	0.019***	-0.031**	-0.012
Natural_Region1	-0.100	-0.290	-0.391
Natural_Region2	0.337	-0.206	0.131
Natural_Region3	0.001	-0.001	0.001

First, keeping the remaining variables constant, the effect of an increase in one unit of Composite_Poverty_Index produces a significant direct decrease of 0.029% deaths per 1000 inhabitants, and the same increase in neighboring provinces produce an insignificant increase on 0.008% deaths per 1000 inhabitants, the total effect is a significant decrease of 0.021 in the deaths per 1000 inhabitants. Second, an increase of 1% in logPD_1000, is related to a direct increase of 0.079% in deaths per 1000 inhabitants, and an insignificant indirect increase of 0.038%, the total effect is insignificant. Third, an increase in one percentage unit in Overcrowding would lead to a positive direct effect of 0.019% in deaths per 1000 inhabitants, and a negative indirect effect of 0.031%, both effects are statistically significant, the total effect is insignificant. After reviewing the marginal effects is clear that provinces vulnerable to COVID-19 mortality are those with higher economic activity, high population density, and a high proportion of overcrowded households, most of the indirect effects are insignificant except for overcrowding. However, overcrowding's direct and indirect effects are counteracted, then policies linked to this variable may not have the desired impact, we recommend future research to evaluate the impact of overcrowding in households on mortality rates.

The following provinces listed in Table 7 have zero reported deaths (if there were zero COVID-19 cases, the deaths had been replaced by zero):

Department	Province	ENR COVID19 Deaths Forecast	Demand category
Ucayali	Purús	0.35	Medium
Ancash	Asunción	0.28	Medium
Ancash	Antonio Raymondi	0.23	Low
Huánuco	Marañón	0.20	Low
Lambayeque	Ferreñafe	0.72	High
Lima	Cañete	1.28	Very high
Loreto	D.Marañón	0.31	Medium

The ENR regression model was cross-validated using a repeated k-fold algorithm, it resulted in parameters $\lambda_1 = 0.07$ and $\lambda_2 = 0.93$. It predicts demand based on vulnerability characteristics against COVID-19 mortality. The predictions vary from 0.20 to 1.28 deaths per 1000 inhabitants. And we have prediction categories from low to very high demand. According to their observable characteristics, is not very likely that these provinces have zero deaths, as is reported on the dataset. In Figure 4 we plot the spatial distribution of the mortality rates. In the first panel, we show the real mortality rates; in the second, the rates predicted by SDM(1,1,1); and in the third panel we show the predictions for ENR model. We observe that the predictions of SDM(1,1,1) are higher in the north. We have high demand predictions in coastal and jungle regions. In highlands, the predominant category is high, and in central-south highlands medium demand is predominant. Low demand is present on the south and central-north highlands.

Government should deliver to very high demand provinces first, furthermore, the central warehouses for PRRSGS must be located in coastal provinces and jungle provinces with very high demand categories. The second echelon of the delivery could be from very high-demand provinces to high, medium, and low-demand ones. The number of warehouses depends on the number of spatial clusters of very high-demand points. Across the coastal region, delivery must use the main road to reach all very high-demand provinces, and then in a second echelon, reach the other provinces. In the second echelon, the delivery must be from coastal to highlands, and from jungle to highlands. The low-demand provinces must be provided with PRRSGS at the end so that we prioritize the medium and high-demand provinces. To follow this delivery strategy, policymakers should consider the predictions of SDM(1,1,1) model complemented by the predictions of ENR model. This must be done in order to overcome data limitations and to establish a delivery strategy for PRRSGS. As pandemic outbreak evolves, more and more supplies will be needed, so establish a national-level delivery strategy is very important for successful pandemic management. In Figure 4 we also observe spatial clusters in coastal and jungle regions. Delivery of PRRSGS must be made first in those clusters of high COVID-19 mortality, and then the supplies must be delivered to medium and low demand quartiles, according to the spatial distribution. If we consider the forecasts of demand present in the third panel of Figure 4, we also found that in coastal and jungle regions we have spatial clusters with high COVID-19 mortality.

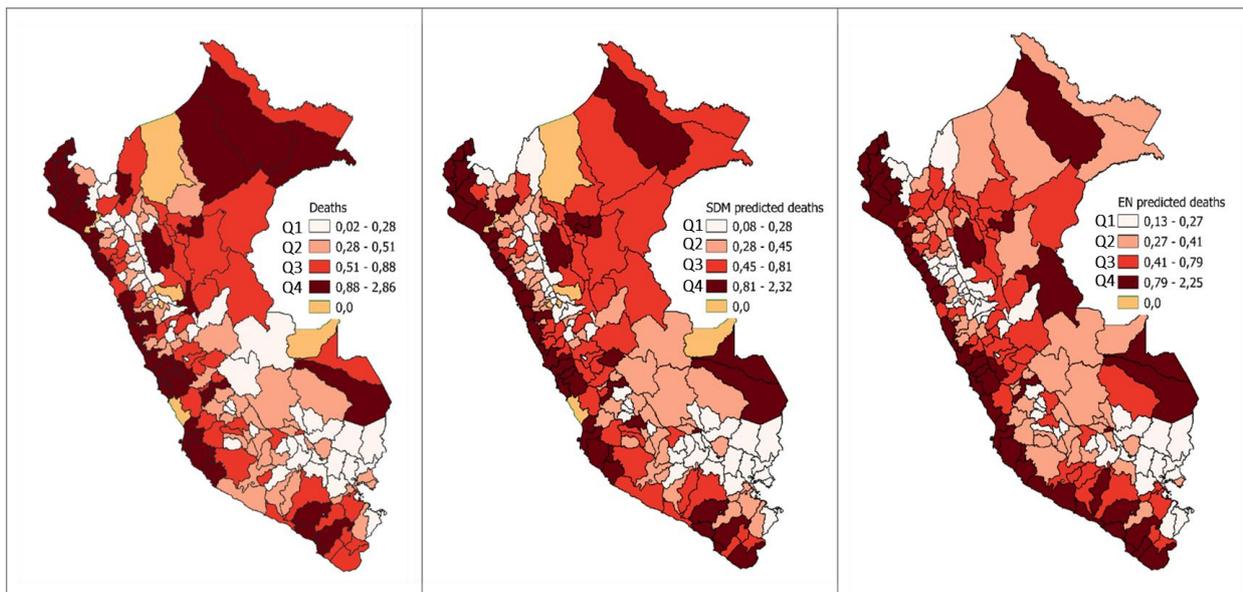


Figure 2. Number of deaths mapping by COVID-19 (per 1000 inhabitants) versus the predictions. First panel shows the actual mortality rates. Second panel shows the predicted mortality rates by SDM (1,1,1). Third panel shows the forecasts of elastic-net regression model.

4. Discussion

Based on the results of this work, we summarize the managerial and research implications:

The highly vulnerable provinces are those with lower poverty, good economic performance, and high urban population density. (Population density: total population divided by total urban area for each province). Economic activity is related to population density and agglomeration, provinces that maintain a certain level of economic activity tend to have higher mortality rates, that is the reason why we need a supporting framework for economic reactivation. High economic activity provinces are suffering the lack of supply for PRRSGS, that is the reason why we have found this effect.

In order to mitigate risks, decision-makers must pay attention to economic performance, population density, and household overcrowding. High mortality should be expected in highly vulnerable provinces with the previously defined characteristics; thus, risk mitigation policies must consider these vulnerability drivers. Population overcrowding in households was a significant predictor, but its direct effect is counteracted by its indirect spatial spillover effect, so more research is needed to clarify its impact on COVID-19 mortality.

There are significant spatial patterns in COVID-19 mortality rates (based on Moran's statistic's significance). This allowed us to trace a delivery strategy based on spatial clusters. First, we must deliver to very high demand provinces. Second, deliver PRRSGS to neighboring high-demand provinces. Third, deliver to medium and low demand provinces, this delivery will be done from the high demand points to low and medium demand points. Policymakers should treat forecasts as complementary information and include the zero deaths provinces into the general delivery strategy.

The forecasts are made based on province-level observable characteristics and suggest that we can follow the same strategy of delivery from coastal and jungle regions to highlands.

Pandemic preparedness, response, and recovery success depend on the methodology used to estimate the demand. This methodology guarantees that we can proxy demand with vulnerability against COVID-19 mortality. Deprivation costs are high where there is high mortality against COVID-19, so we propose that demand assessment must follow this criterion.

5. Conclusions

In this paper, we have argued that a supportive framework for economic reactivation can be built based on vulnerability against COVID-19 mortality. This vulnerability arises from deprivation costs that are high as a consequence of the pandemic outbreak. Following this idea, we decided to estimate SDMs where the dependent variable was the mortality rates for COVID-19 at the province level. This regression model helped to understand first the difference between mortality rates as the difference in vulnerability against COVID-19. The rest of the methodological steps included model selection and post estimation analysis that helped to find the best model and to use their predictions to trace a delivery strategy for PRRSGS. The demand for PRRSGS was proxied by the predicted COVID-19 mortality rates, which determined vulnerability against COVID-19 mortality.

The principal result for the regression analysis is that $SDM(1,1,1)$, the more general SDM is the one that better fits the data suggesting that there are endogenous, exogenous and correlated spatial interaction effects. This means that mortality rates exhibit spatial patterns. Poverty, population density, and overcrowding remained as the main predictors. However, overcrowding is direct and indirect (spatial spillovers) effects are counteracted, so further research is needed to clarify the relationship between overcrowding and mortality rates. Provinces are inevitably returning to their long-run level of economic activity, and high mortality rates exist because there is scarcity in essential supplies that are needed to save lives regarding COVID-19 illness. Population density is negatively related to mortality rates, so for provinces with low population density, policymakers should

expect higher mortality rates, and thus delivery of PRRSGS must be prioritized. The rest of vulnerability drivers related to demographics, health insurance and health system performance were not significant and empirically they were not related to COVID-19 mortality rates.

Economic reactivation must be done considering humanitarian logistics principles which imply the assessment of the demand for humanitarian aid regarding COVID-19 mortality. The applied methodology shows that data-driven techniques are especially useful to improve disaster management techniques and decision-making. This methodology can be replicated with other cases of study to optimize the policies for the pandemic response and recovery phases. However, future research is still needed to overcome the limitations of data and to validate the exploratory results in this paper.

6. Patents

Author Contributions: Conceptualization, R.R., R.Q.A, I.B. and M.C.; methodology, R.Q.A., R.R., A.L.; software, R.Q.A.; validation, R.R., A.L. and R.Q.A.; formal analysis, R.Q.A., R.R. and A.L.; investigation, I.B., M.C. and R.Q.A; resources, R.R., I.B. and M.C.; data curation, R.Q.A.; writing—original draft preparation, R.Q.A, I.B. and M.C.; writing—review and editing, A.L., I.B., M.C., R.R. and R.Q.A.; visualization, R.R., A.L. supervision, M.C.; project administration, M.C.; funding acquisition, M.C., R.R. and I.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Centro de Investigación de la Universidad del Pacífico, Universidad Nacional Abierta y a Distancia and Sao Paulo State University.

Data Availability Statement: Data as well as software program used to perform the estimation is attached in the following link: <https://github.com/renatoquiliche/COVID19-socsi>.

Acknowledgments: We acknowledge the financial support of Vicerrectorado de Investigación de la Universidad del Pacífico.

Conflicts of Interest: Authors declare no conflicts of interest.

References

- (Acemoglu et al., 2020) Acemoglu, D., Chernozhukov, V., Werning, I., & Whinston, M. D. 2020. A Multi-Risk SIR Model with Optimally Targeted Lockdown (Working Paper No. 27102; Working Paper Series.). *National Bureau of Economic Research* <https://doi.org/10.3386/w27102>
- (Alexander, 2002) Alexander, D.E. 2002. *Principles of Emergency Planning and Management*. Oxford University Press, US: Oxford.
- (Alibaba, 2020a) Alibaba. 2020a. *COVID-19 Outbreak Hospital Response Strategy*. Hangzhou, China.
- (Alibaba, 2020b) Alibaba 2020b. *Handbook of COVID-19 Prevention and Treatment*. Zhejiang University School of Medicine: China.
- (Andersen et al., 2021) Andersen, L. M., Harden, S. R., Sugg, M. M., Runkle, J. D., & Lundquist, T. E. 2021. Analyzing the spatial determinants of local Covid-19 transmission in the United States. *Science of the Total Environment*, 754, 142396. <https://doi.org/10.1016/j.scitotenv.2020.142396>
- (Béland et al., 2020) Béland, L.-P., Brodeur, A., Mikola, D., & Wright, T. 2020. The Short-Term Economic Consequences of Covid-19: Occupation Tasks and Mental Health in Canada. *Social Science Research Network*. <https://papers.ssrn.com/abstract=3602430>
- (Benzakoun et al., 2020) Benzakoun, J., Hmeydia, G., Delabarde, T., Hamza, L., Meder, J. F., Ludes, B., & Mebazaa, A. 2020. Excess out-of-hospital deaths during the COVID-19 outbreak: evidence of pulmonary embolism as a main determinant. *European Journal of Heart Failure*, 22(6), 1046–1047. <https://doi.org/10.1002/ejhf.1916>
- (Brodeur et al., 2020) Brodeur, A., Gray, D. M., Islam, A. & Bhuiyan, S. 2020. A Literature Review of the Economics of Covid-19. IZA Discussion Paper No. 13411, Available at SSRN: <https://ssrn.com/abstract=3636640>
- (Bruno et al., 2020) Bruno, G., Wemole, R., Ronsivalle, G. B., Foresti, L., & Poledda, G. 2020. A Prototype Model of Georeferencing the Inherent Risk of Contagion from COVID-19. March. <https://doi.org/10.13140/RG.2.2.30077.72163>
- (Castillo-Larco & Castillo-Cara, 2020) Carrillo-Larco, R. M., & Castillo-Cara, M. 2020. Using country-level variables to classify countries according to the number of confirmed COVID-19 cases: An unsupervised machine learning approach. *Wellcome Open Research*, 5, 1–22. <https://doi.org/10.12688/wellcomeopenres.15819.3>
- (Chakraborti et al., 2020) Chakraborti, S., Maiti, A., Pramanik, S., Sannigrahi, S., Pilla, F., Banerjee, A., & Das, D. N. 2020. Evaluating the plausible application of advanced machine learnings in exploring determinant factors of present pandemic: A case for continent specific COVID-19 analysis. *Science of the Total Environment*, 142723. <https://doi.org/10.1016/j.scitotenv.2020.142723>

- (Crandall & Weber, 2004) Crandall, M. S., & Weber, B. A. 2004. Local Social and Economic Conditions, Spatial Concentrations of Poverty, and Poverty Dynamics. *American Journal of Agricultural Economics*, 86(5), 1276–1281. <https://doi.org/10.1111/j.0002-9092.2004.00677.x>
- (Elhorst, 2010) Elhorst, J. P. 2010. Applied Spatial Econometrics: Raising the Bar. *Spatial Economic Analysis*, 5(1), 9–28. <https://doi.org/10.1080/17421770903541772>
- (Engle et al., 2020) Engle, S., Stromme, J., & Zhou, A. 2020. Staying at home: mobility effects of covid-19. Available at SSRN.
- (Guliyev, 2020) Guliyev, H. 2020. Determining the spatial effects of COVID-19 using the spatial panel data model. *Spatial Statistics*, 38(January), 100443. <https://doi.org/10.1016/j.spasta.2020.100443>
- (Gupta et al., 2020) Gupta, A., Banerjee, S., & Das, S. 2020. Significance of geographical factors to the COVID-19 outbreak in India. *Modeling Earth Systems and Environment*, 6(4), 2645–2653. <https://doi.org/10.1007/s40808-020-00838-2>
- (Gutjahr & Fisher, 2018) Gutjahr, W. J., & Fischer, S. 2018. Equity and deprivation costs in humanitarian logistics. *European Journal of Operational Research*, 270(1), 185–197. <https://doi.org/10.1016/j.ejor.2018.03.019>
- (Holguín-Veras et al., 2013) Holguín-Veras, J., Pérez, N., Jaller, M., Van Wassenhove, L. N., & Aros-Vera, F. 2013. On the appropriate objective function for post-disaster humanitarian logistics models. *Journal of Operations Management*, 31(5), 262–280. <https://doi.org/10.1016/j.jom.2013.06.002>
- (Kelejian & Prucha, 1999) Kelejian, H. H., & Prucha, I. R. 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review*, 40(2), 509–533. <https://doi.org/10.1111/1468-2354.00027>
- (Khalatbari-Soltani et al., 2020) Khalatbari-Soltani, S., Cumming, R. C., Delpierre, C., & Kelly-Irving, M. 2020. Importance of collecting data on socioeconomic determinants from the early stage of the COVID-19 outbreak onwards. *Journal of Epidemiology and Community Health*, 74(8), 620–623. <https://doi.org/10.1136/jech-2020-214297>
- (Leiras et al., 2014) Leiras, A., de Brito Jr, I., Queiroz Peres, E., Rejane Bertazzo, T., & Tsugunobu Yoshida Yoshizaki, H. 2014. Literature review of humanitarian logistics research: trends and challenges. *Journal of Humanitarian Logistics and Supply Chain Management*, 4(1), 95–130. <https://doi.org/10.1108/jhlscm-04-2012-0008>
- (Ministerio de Economía y Finanzas, 2020) Ministerio de Economía y Finanzas. 2020. Aprueban el Reglamento Operativo del Fondo de Apoyo Empresarial a la MYPE (FAE-MYPE). Resolución Ministerial N° 124-2020-EF/15.
- (Ministerio del interior, 2020) Ministerio del Interior. 2020. Aprueban protocolo para la implementación de las medidas que garantizan el ejercicio excepcional del derecho a la libertad de tránsito en el marco del Estado de Emergencia Nacional declarado mediante D.S. N° 044-2020-PCM. Resolución Ministerial N° 304-2020-IN.
- (Montoro et al., 2020) Montoro, C., Pérez, F., & Herrada, R. 2020. Medidas del BCRP frente a la pandemia del nuevo coronavirus. *MONEDA* 182: 10-18.
- (Ogutu et al., 2012) Ogutu, J. O., Schulz-Streeck, T., & Piepho, H. P. 2012. Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. In *BMC proceedings* (Vol. 6, No. 2, pp. 1-6). BioMed Central.
- (Overstreet et al., 2011) Overstreet, R. E., Hall, D., Hanna, J. B., & Kelly Rainer, R. 2011. Research in humanitarian logistics. *Journal of Humanitarian Logistics and Supply Chain Management*, 1(2), 114–131. <https://doi.org/10.1108/20426741111158421>
- (Pedrosa & de Albuquerque, 2020) Pedrosa, N. L., & de Albuquerque, N. L. S. 2020. Spatial analysis of COVID-19 cases and intensive care beds in the state of Ceará, Brazil. *Ciencia e Saude Coletiva*, 25, 2461–2468. <https://doi.org/10.1590/1413-81232020256.1.10952020>
- (Pyndick & Wang, 2013) Pindyck, R. S., & Wang, N. 2013. The Economic and Policy Consequences of Catastrophes. *American Economic Journal: Economic Policy*, 5(4), 306–339. <https://doi.org/10.1257/pol.5.4.306>
- (Poder Ejecutivo, 2020a) Poder Ejecutivo. 2020a. Decreto de urgencia que establece el retiro extraordinario del fondo de pensiones en el sistema privado de pensiones como medida para mitigar efectos económicos del aislamiento social obligatorio y otras medidas. Decreto de Urgencia N° 034-2020.
- (Poder Ejecutivo, 2020b) Poder Ejecutivo. 2020b. Decreto de urgencia que establece medidas para reducir el impacto en la economía peruana, de las disposiciones de prevención establecidas en la declaratoria de estado de emergencia nacional ante los riesgos de propagación del COVID-19. Decreto de Urgencia N° 033-2020.
- (Presidencia del Consejo de Ministros, 2020a) Presidencia del Consejo de Ministros. 2020a. Decreto Supremo que precisa los alcances del artículo 8 del Decreto Supremo N 044-2020-PCM, que declara el estado de emergencia nacional por las graves circunstancias que afectan la vida de la nación a consecuencia del brote del COVID-19. Decreto Supremo N° 045-2020-PCM
- (Presidencia del Consejo de Ministros, 2020b) Presidencia del Consejo de Ministros. 2020b. Decreto Supremo que modifica el Decreto Supremo N° 116-2020-PCM, Decreto Supremo que establece las medidas que debe seguir la ciudadanía en la nueva convivencia social y prorroga el Estado de Emergencia Nacional por las graves circunstancias que afectan la vida de la Nación a consecuencia del COVID-19, modificado por los Decretos Supremos N° 129-2020-PCM, N° 135-2020-PCM, N° 139-2020-PCM, N° 146-2020-PCM, N° 151-2020-PCM, N° 156-2020-PCM, N° 162-2020-PCM y N° 165-2020-PCM. Decreto Supremo N° 170-2020-PCM.
- (Renteria et al., 2021) Renteria, R., Chong, M., de Brito Junior, I., Luna, A., & Quiliche, R. 2021. An entropy-based approach for disaster risk assessment and humanitarian logistics operations planning in Colombia. *Journal of Humanitarian Logistics and Supply Chain Management*. <https://doi.org/10.1108/jhlscm-03-2020-0018>
- (Rollston & Gallea, 2020) Rollston, R., & Galea, S. 2020. COVID-19 and the Social Determinants of Health. *American Journal of Health Promotion*, 34(6), 687–689. <https://doi.org/10.1177/0890117120930536b>
- (Rustam et al., 2020) Rustam, F., Reshi, A. A., Mehmood, A., Ullah, S., On, B. W., Aslam, W., & Choi, G. S. 2020. COVID-19 Future Forecasting Using Supervised Machine Learning Models. *IEEE Access*, 8, 101489–101499. <https://doi.org/10.1109/ACCESS.2020.2997311>

-
- (Smith & Mennis, 2020) Smith, C. D., & Mennis, J. 2020. Incorporating geographic information science and technology in response to the COVID-19 pandemic. *Preventing Chronic Disease*, 17(17), 1–7. <https://doi.org/10.5888/PCD17.200246>
- (Sugg et al., 2021) Sugg, M. M., Spaulding, T. J., Lane, S. J., Runkle, J. D., Harden, S. R., Hege, A., & Iyer, L. S. 2021. Mapping community-level determinants of COVID-19 transmission in nursing homes: A multi-scale approach. *Science of the Total Environment*, 752, 141946. <https://doi.org/10.1016/j.scitotenv.2020.141946>
- (Sun et al., 2020) Sun, F., Matthews, S. A., Yang, T. C., & Hu, M. H. 2020. A spatial analysis of the COVID-19 period prevalence in U.S. counties through June 28, 2020: where geography matters? *Annals of Epidemiology*, 52, 54-59.e1. <https://doi.org/10.1016/j.annepidem.2020.07.014>
- (Wardhana et al., 2017). Wardhana, D., Ihle, R., & Heijman, W. 2017. Agro-clusters and Rural Poverty: A Spatial Perspective for West Java. *Bulletin of Indonesian Economic Studies*, 53(2), 161–186. <https://doi.org/10.1080/00074918.2017.1298722>
- (White & Hébert-Dufresne, 2020). White, E. R., & Hébert-Dufresne, L. 2020. State-level variation of initial COVID-19 dynamics in the United States. *PLoS ONE*, 15(10 OCTOBER), 1–13. <https://doi.org/10.1371/journal.pone.0240648>
- (Yoshizaki et al., 2020) Yoshizaki, H. T. Y., de Brito Junior, I., Hino, C. M., Aguiar, L. L., & Pinheiro, M. C. R. 2020. Relationship between Panic Buying and Per Capita Income during COVID-19. *Sustainability*, 12(23), 9968. <https://doi.org/10.3390/su12239968>
- (Zeckhauser, 1996) Zeckhauser, R. 1996. The economics of catastrophes. *Journal of Risk and Uncertainty*, 12(2–3), 113–140. <https://doi.org/10.1007/bf00055789>
- (Zhang & Schwartz, 2020) Zhang, C. H., & Schwartz, G. G. 2020. Spatial Disparities in Coronavirus Incidence and Mortality in the United States: An Ecological Analysis as of May 2020. *Journal of Rural Health*, 36(3), 433–445. <https://doi.org/10.1111/jrh.12476>
- (Zhou & Lin, 2008) Zhou, X., & Lin, H. 2008. Spatial Weights Matrix. In *Encyclopedia of GIS* (pp. 1113–1113). Springer US. https://doi.org/10.1007/978-0-387-35973-1_1307