Remote Sensing Applications: Society and Environment 31 (2023) 100990

Contents lists available at ScienceDirect

Remote Sensing Applications: Society and Environment

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

Sentinel satellite data monitoring of air pollutants with interpolation methods in Guayaquil, Ecuador

Danilo Mejía C. [a, b, c, *], Hermel Alvarez [a], Rasa Zalakeviciute [d], Diana Macancela [e], Carlos Sanchez [c], Santiago Bonilla [f]

[a] Grupo CATOx, CEA de la Universidad de Cuenca, Campus Balzay, 010207, Cuenca, Ecuador
[b] Laboratorio de Ecología Acuática (LEA) General Escandon s/n, Facultad de Ciencias Químicas Universidad de Cuenca, Ecuador
[c] Carrera de Ingeniería Ambiental de la Universidad de Cuenca, Campus Balzay, 010207, Cuenca, Ecuador
[d] Grupo de Biodiversidad Medio Ambiente y Salud (BIOMAS), Universidad de Las Americas, Quito – EC, 170125, Ecuador
[e] Carrera de Ingeniería Ambienta de la Universidad Católica de Cuenca, Ecuador
[f] Research Center for the Territory and Sustainable Habitat, Universidad Tecnológica Indoamérica, Machala y Sabanilla, 170301, Quito, Ecuador

ARTICLE INFO

Keywords:
Satellite
Remote sensing
Air pollution
Interpolation
NO_2
Sentinel-5

ABSTRACT

In Ecuador, there is a limitation on air quality monitoring due to the cost of monitoring networks. Although air quality monitoring stations are instruments for air measurement, they do not cover an entire city due to their scope. Satellite remote sensing is now an effective tool to study atmospheric pollutants and has been applied to continuously assess a region and overcome the limitations of fixed stations. Despite the application of satellite data for air quality monitoring, there are some limitations, such as measurement frequency, cloud cover and wide spatial resolution, which do not allow the assessment of air pollution in cities. Therefore, downscaling, applying interpolation methods, is essential for continuous air quality monitoring at smaller scales. For this research, Nitrogen Dioxide (NO_2) data from the Sentinel-5 satellite percussor was used in the city of Guayaquil for January–December 2020, which is considered before, during and after the COVID-19 quarantine. This mid-size port city does not have a permanent monitoring network, which prevents us from knowing the air quality. Due to the limitation of pixel size, this study used satellite data to apply interpolation techniques and reduce pixels to assess air quality. Two categories of interpolation were selected: deterministic and stochastic. The empirical Bayesian kriging (EBK) interpolation obtained a R^2 of 0.9546, which was superior to the other methods applied. Therefore, the EBK method had the best accuracy for tropospheric NO_2 concentration. Finally, the method used in this research can help monitor air quality in cities lacking continuous monitoring networks, as the reduction of the pixel size gives us a better pattern of pollutants.

1. Introduction

In the last few decades, the accelerated process of urbanization and industrialization has led to increased emissions of air pollutants, such as nitrogen dioxide (NO_2), sulfur dioxide (SO_2), carbon monoxide (CO), and particulate matter (PM). These pollutants have significant adverse effects on the economy, human health, and the environment (Yang and Liu, 2018). To mitigate these risks, the key strategies are real-time monitoring and control of air pollutants, and the communication of air quality status to the public

* Corresponding author. Grupo CATOx, CEA de la Universidad de Cuenca, Campus Balzay, 010207, Cuenca, Ecuador.
E-mail addresses: danilo.mejia@ucuenca.edu.ec (D. Mejía C.), hermeldan_@hotmail.com (H. Alvarez), rasa.zalake@gmail.com (R. Zalakeviciute), dmacancelae@ucacue.edu.ec (D. Macancela), carlos.sancheza@ucuenca.edu.ec (C. Sanchez), santiagobonillab@hotmail.es (S. Bonilla).

https://doi.org/10.1016/j.rsase.2023.100990
Received 13 June 2022; Received in revised form 1 May 2023; Accepted 6 May 2023
Available online 19 May 2023

(Athira et al., 2018). However, in the developing countries high costs of technology and maintenance limit number of monitoring stations, which, in turn, hampers the implementation of effective pollution prevention, control, and monitoring programs (Ma et al., 2019). For example, in Ecuador, currently, only two cities, Quito and Cuenca, have air quality monitoring networks.

Due to their limited monitoring range, however, air quality monitoring stations cannot reflect the full picture of the real-time spatial distribution of air pollution in complex urban areas, which impedes the spatial analysis and monitoring of air pollution (Tong et al., 2015) (Zhi et al., 2017). As a result, there have been research attempts to fill the gaps between sparse ground-based data by spatial interpolation methods: kriging algorithm and inverse distance weighting (Janssen et al., 2008; Núñez-Alonso et al., 2019). More recently, machine learning modelling and neural networks, have been employed to increase the accuracy of these interpolations (Ma et al., 2019; Yi et al., 2018; Zalakeviciute et al., 2021). However, there are several drawbacks to using the models that require ground-based sensor data, such as values falling outside the range of the training data, no consideration of spatiotemporal associations, requirements for a large historical dataset, or lack of ground-based sensor data (Kn et al., 2021). Thus, an alternative approach is to use remote sensing data to monitor air quality, as high correlations ($r = 0.68–0.75$) have been found between satellite data of tropospheric NO_2 from Sentinel-5P and ground-based measurements (Ialongo et al., 2020; Muniraj et al., 2021; Vîrghileanu et al., 2020).

Satellite remote sensing has now emerged as an effective tool for studying environmental pollutants in the atmosphere, and is commonly applied to help continuously assess air quality at any spatial-scale region (Engel-Cox et al., 2004). Satellite remote sensing can provide a complete survey of a city, by identifying the main sources of pollution and its spatial distribution patterns, which then can help determine where to focus pollution abatement efforts (Engel-Cox et al., 2004) (Schneider et al., 2021). For instance, Omrani et al. (2020) used spatiotemporal data from the Sentinel-5P remote sensing platform to measure NO_2 concentration differences over mainland France. Haq et al. (2014) studied tropospheric NO_2 hotspots in Pakistan using Ozone Monitoring Instrument (OMI) on NASA's Earth orbiting satellite Aura. Kn et al. (2021) used multiple spatiotemporal datasets from the Sentinel-5P satellite to map air quality (SO_2, NO_2, CO, O_3, HCHO) in India. Muhammad et al. (2020) used Sentinel-5P and Aura data to assess NO_2 concentration changes due to the coronavirus pandemic in China, Spain, France, Italy, and the United States. Schneider et al. (2021) found that the TROPOspheric Monitoring Instrument (TROPOMI) NO_2 dataset provides sufficiently high and valuable data availability to assess NO_2 pollution in Norwegian cities.

Cooper et al. (2020) and Wu et al. (2019) developed algorithms to estimate ground-level NO_2 concentrations with high spatial resolution, applied to the TROPOMI satellite instrument, aiming to monitor concentrations in urban areas. Similarly, Cersosimo et al. (2020) proposed an area-weighted interpolation method to improve the spatial resolution of tropospheric NO_2 observations from TROPOMI. Moreover, pixel reconstruction methods have proven to be efficient in increasing the spatial resolution of satellite-measured NO_2 without compromising the quality of the original data (Kim et al., 2018).

While emission inventories of isolated coal-powered power plants can help validate the performance of NO_2 instrument on satellites (de Foy et al., 2015); in turn, tropospheric NO_2 column, measured by Ozone Monitoring Instrument (OMI), can help detect air pollution from different sources (Wang et al., 2012) (Filonchyk and Peterson, 2023). Evolution in temporal satellite NO_2 profiles can also help identify air quality deterioration due to new coal power plant construction (Wang et al., 2012) or improvements due COVID-19 pandemic (Filonchyk et al., 2021).

Despite the wide application of satellite data, there are some limitations, such as cloud cover blockage, or time-limited measurement, restricted to data acquired when the satellite passes overhead, and recorded as a concentration of contaminants within a vertical column (Berman and Ebisu, 2020). Currently, there are several satellite air quality products that have been successfully used in numerous environmental applications, but due to their wide spatial resolution, they do not allow air pollution assessments in urban areas (Shelestov et al., 2018).

In this context, it is essential to explore downscaling techniques, such as interpolation methods, to ensure continuous monitoring of air quality at the urban scale while also eliminating cloud interference. This research focuses on identifying the most suitable interpolation method using data from the Sentinel-5P satellite in cities where ground-based measurements are not available. The long-term goal is for this methodology to influence public policies and decision-making processes related to land-use planning, which would facilitate a more efficient approach to address the challenges associated with air quality and pollution in these areas. Digital image interpolation is the process by which a continuous intensity surface is generated from discrete image data samples. There are many types of interpolation methods, each of which results in a different appearance of the final image. Therefore, it is best if the quality, or visible distinctness of each pixel, is maintained throughout the interpolation function. Image interpolation is a term used for image processing, but it is often used with different terminology in the literature, such as image scaling, image resampling, and image resizing (Patel and Mistree, 2013). Interpolation is a technique that enables the transfer of images between resolutions without any loss of quality (Prajapati et al., 2012). When the resolution of an image is increased, it is known as up-sampling or up-scaling, and when decreased, it is called down-sampling or down-scaling (Prajapati et al., 2012). This technique has numerous applications, such as image resizing, zooming, enhancement, reduction, sub-pixel image registration, image decomposition, and spatial distortion correction (Prajapati et al., 2012).

The present study provides a detailed comparison of the performance of six interpolation techniques to visualize and analyze the spatial variability and temporal dynamics of NO_2 concentration in the tropospheric column from January to December 2020 consequence of COVID-19 using Sentinel-5P data, namely inverse distance weighting (IDW), local polynomial interpolation (LPI), ordinary kriging (OK), simple kriging (SK), empirical Bayesian kriging (EBK), and random forest for spatial predictions framework interpolation (RFSP). The interpolated surfaces were created using ArcGIS software within the Spatial Analyst and Geostatistical Analyst series of interpolation techniques. To compare interpolation techniques, cross-validation parameters were used. The best method was then employed to produce a city-scale monthly tropospheric NO_2 map of Guayaquil using a spatial resolution of 30 m.

The research is structured into four parts: a) interpolation methods are evaluated when applied to the NO_2 variable for spatial resolution reduction, b) the results of the interpolation regarding spatial scale are analyzed, and a temporal analysis of NO_2 concentrations is incorporated, c) the findings obtained are discussed, and finally, d) the research is concluded by summarizing the results and highlighting the implications of the interpolation methods used in the study of NO_2.

2. Methods

2.1. Study area

The city of Guayaquil was selected as the study area of this work. It is located in the Guayaquil canton in the province of Guayas in the extreme west of Ecuador (Fig. 1) It is the largest, most populated, and most industrialized city in Ecuador, with a total area of 1800 km^2 and an estimated population of 2.69 million people (Jácome et al., 2019; Pacheco et al., 2020). The city is located in low and swampy areas (4.6 m above sea level), except some mountainous areas to the north (highest point at 460 m) (Delgado, 2013). Guayaquil's climate is considered humid tropical, with an average annual temperature of 26 °C, an average rainfall of 1030 mm, and prevailing winds that blow to the west. Ambient NO_2 concentrations in this area are mainly attributed to traffic, with minor contributions from the industrial sector.

2.2. Data collection

The tropospheric NO_2 column concentration data used in this paper was collected from the Copernicus Open Access Hub platform on the European Space Agency's Sentinel-5P. The tropospheric NO_2 column concentration data were produced by a TROPOspheric Monitoring Instrument (TROPOMI) instrument on a Sentinel-5 precursor (also known as Sentinel-5P) (Veefkind et al., 2012), which provides global coverage of the main pollutant gases, such as NO_2, SO_2, CO, CH_4, and O_3 (European Commission and ESA, 2015).

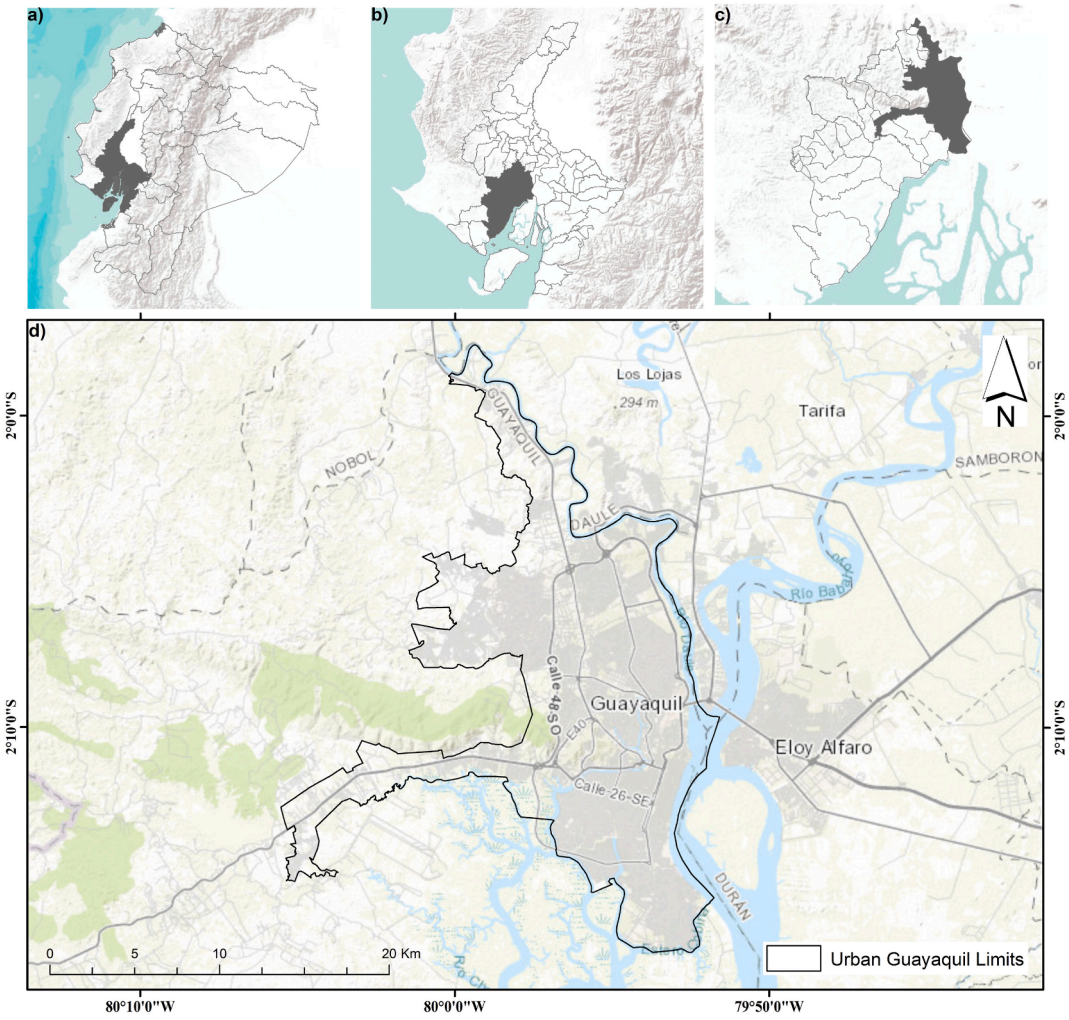

Fig. 1. (a)The map shows the province of Guayas in the gray area, (b) the canton Guayaquil in the gray area, (c) the city of Guayaquil in the gray area, and (d) the study area, the city of Guayaquil.

TROPOMI, is a passive sun backscatter imaging spectrometer for measuring ultraviolet–visible (270–500 nm), near-infrared (675–775 nm), and short-wave infrared (2305–2385 nm) spectral bands (Ialongo et al., 2020), which means it can image various air pollutants more accurately than previous systems (Zheng et al., 2019). Sentinel-5P can achieve global coverage every 24 h, providing access to dense measurements around the world. The TROPOMI tropospheric NO_2 product can retrieve good quality information resolution that is finer than any previous satellites could archive (Griffin et al., 2019). All mission measurements of atmospheric gases are "column data," meaning they cover the entire depth of the atmosphere. The tropospheric NO_2 column has been detailed by Eskes et al. (2021) and van Geffen et al. (2021).

For this research we worked with monthly average NO_2 column concentration data in the troposphere. (January–December 2020) were used in this paper. These periods were selected to identify and compare the evolution of NO_2 concentrations under normal conditions and during the coronavirus pandemic lockdown. The images were retrieved using Google Earth Engine (GEE) Level 3 products for Sentinel-5P (Gorelick et al., 2017; Zalakeviciute et al., 2020). The spatial data resolution of tropospheric NO_2 concentration ($\mu mol/m^2$) columns was 1 km^2 per pixel. The images were downloaded in Geotiff format and referenced to the World Geodetic System of 1984 (WGS1984). In the city of Guayaquil, there is no permanent monitoring network, so it was not possible to establish trends of deterioration in air quality in another way. Due to this great limitation, this study uses data from satellite platforms to apply interpolation techniques to assess air quality.

2.3. Interpolation methods

The choice of interpolation algorithm directly affects the quality of the final image, making it important to find a suitable algorithm to improve image quality. The results of the interpolation process are sensitive and dependent on the method used, the selected dataset, and the optimization processes of the interpolated values (Ali et al., 2021). Since a variety of methods are currently available in the literature to estimate unknown values of continuous variables, such as pollutant gas concentrations, this study systematically analyzed spatial interpolation methods to assess the degree to which these different methods affect the estimated values of the tropospheric NO_2 column concentration. Based on a comprehensive literature review, we selected interpolation techniques in two different categories: deterministic and stochastic (Deligiorgi and Philippopoulos, 2011).

The deterministic category is represented by two methods: IDW and LPI. The stochastic category is represented by four methods, namely OK, SK, EBK, and RFSP. The processing of monthly tropospheric NO_2 column concentration data from Sentinel-5P consisted of the transformation of the NO_2 pixels into points through the "raster to point" conversion tool of ArcGIS 10.5. For the development of the spatial interpolation methods (IDW, LPI, OK, SK, and EBK), ArcGIS software was used. The RFSP model was generated in RStudio (Hengl et al., 2018a; 2018b; Hengl and MacMillan, 2019). The data used in this study were placed in the Geostatistical Analyst of ArcGIS 10.5 with a grid resolution of 1 × 1 km. Between 408 and 422 sampling points were used, depending on the availability of tropospheric NO_2 data, to cover the entire city of Guayaquil.

The common objective of the selected interpolation methods is to estimate the unknown value of $\hat{Z}$ at the point x_i, so the methods have the same basic mathematical formulation. The interpolation relationship is calculated as:

$$\hat{Z}(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \tag{1}$$

where $Z(x_i)$ represents the measured value at point x_i, n is the total number of existing data points, λ_i is the weight function allocated to data points, and the sum of the weights is 1 to ensure that the interpolation is unbiased (Ali et al., 2021; Wong et al., 2004). The six interpolation methods differ only in their choice of sample weights. For these approaches, interpolation involves: (1) defining the search area or neighborhood around the point to be predicted; (2) locating the observed data points within this neighborhood; and (3) assigning appropriate weights to each of the observed data points (Wong et al., 2004).

2.3.1. Deterministic methods

***Inverse distance weighting (IDW)*:** IDW is one of the simplest and most widely used methods for interpolating air quality data (Fontes and Barros, 2010) and is based on the presumption that nearby sampling points should have a greater relationship or effect than more distant points. This method does not require subjective assumptions in the selection of the semi-variogram model, which gives it an advantage over other methods, such as kriging (Ali et al., 2021). With this method, values are estimated using a linear combination of the value at the sampled point weighted by an inverse function of the distance between two points. The weights are determined according to the inverse distance from sampled points to the new point following Equation (2):

$$\lambda_i^{IDW} = \frac{1/d_i^p}{\sum_{i=1}^{n} 1/d_i^p} \tag{2}$$

where d_i represents the distance between the measured point x_i and the predicted point x_0, n represents the total number of measured observations used, and p represents a power parameter that is used to highlight the spatial relationship between points. The power parameter for IDW used in our study was 2. Van Roode et al. (2019) recommends p limits between 1 and 4 for air pollution. In this work, all available sampled points were used to compute the new measurement at a point.

***Local polynomial interpolation (LPI)*:** The LPI interpolation method uses polynomials on all surfaces and has algorithmic functions that correspond to polynomial sequences. In this study, we used the first-order polynomial. LPI conforms to a local subset de-

fined by a moving window. This window should include a reasonable number of data points; therefore, the window size must be large enough for this process (Wu et al., 2019). The surface value $\mu_0(x_i,y_i)$, at the center of this window is computed at each point. The LPI interpolation method only uses the surrounding sample. For first-order polynomials, it can be represented as:

$$\mu_0\left(x_i, y_i\right) = \beta_0 + \beta_1 x_i + \beta_2 y_i \tag{3}$$

Indices β are estimated as the center and window move through space.

2.3.2. Geo-statistical methods

***Empirical Bayesian kriging (EBK)*:** EBK is a geostatistical interpolation method that automates the difficult aspects of building a valid kriging model. Other kriging methods require a to manually adjust the model parameters, but EBK automatically calculates these parameters through a process of sub-configuration and simulations (Sağir and Kurtuluş, 2017). EBK is a robust and straightforward spatial interpolation technique that consists of two geostatistical models: (1) intrinsic random function kriging and (2) linear mixed model (Gribov and Krivoruchko, 2020). EBK considers the uncertainty in the spatial parameters. The algorithm behind EBK generates several semi-variogram models to minimize the prediction error generated by the uncertainty of the model parameters (Banerjee et al., 2018).

Simple kriging (SK): SK is a geostatistical interpolation technique that considers both the distance and degree of variation between known data points when estimating values in unknown areas (Narváez and Useche, 2020). SK is a variety of Kriging interpolations that require less computation but is a less general approach (Rebholz and Almekkawy, 2020). With simple kriging, the mean value is presumed to be known. SK has limited applications, but it is important because it has good properties that ordinary kriging does not share (Daya et al., 2018). SK is based on the kriging estimator and is approximated as follows:

$$\hat{Z}\left(x_0\right) - \mu\left(x_0\right) = \sum_{i=1}^{n} \lambda_i \left[Z\left(x_i\right) - \mu\left(x_i\right)\right] \tag{4}$$

where λ_i represents the kriging weight, n is the total number of existing data points, $Z(x_i)$ is the measured variable value at a given data point i, and μ represents a known stationary mean. μ is derived as the mean of the data and is assumed to be constant over the entire study area. The weight λ_i is calculated using a semi-variogram or covariance function. In this paper, a semi-variogram was used due to its wide applications in different interpolation studies (Ali et al., 2021) and was estimated using:

$$\gamma\left(x_i, x_0\right) = \gamma\left(h\right) = \frac{1}{2} var\left[Z\left(x_i\right) - Z\left(x_0\right)\right] \tag{5}$$

where $\gamma(h)$ represents the semi-variance and h represents the distance between the measured and predicted data points. The stable semi-variogram model was used in this study. As mentioned above in SK, the trend component is presumed to be an exactly known constant for the entire study area, which is approximated using the average value of the measured data $\mu(x_0) = \mu$, such that:

$$\hat{Z}_{SK}\left(x_0\right) = \mu + \sum_{i=1}^{n} \lambda_i^{SK}\left(x_0\right) + \left[Z\left(x_i\right) - \mu\right] \tag{6}$$

Ordinary kriging (OK): Ordinary kriging is the most general and widely used method that presumes a constant unknown mean in the local neighborhood (Xie et al., 2017). OK is an estimation technique known as the best unbiased linear estimator. It has the advantage of using semi-variogram information (Qiao et al., 2019). The semi-variogram is used to describe the randomness and structuring of variables (Qu et al., 2017). OK uses the semi-variogram to determine the weights and characterize spatial dependence. In this study, the stable model is adopted for the interpolation processes. The coefficients of the chosen model are then used to determine the weights for Kriging. OK is also described as the acronym BLUE, representing the "best linear unbiased estimator" (Ali et al., 2021). OK is then implemented by the equation below:

$$\hat{Z}_{OK}\left(x_0\right) = \sum_{i=1}^{n} \lambda_i^{OK}\left(x_0\right) Z\left(x_i\right) \quad with \sum_{i=1}^{n} \lambda_i^{OK}\left(x_0\right) = 1 \tag{7}$$

The difference between OK and SK is that in SK, it is presumed that the mean value is known, while with OK, it is presumed that there is an unknown constant mean value that is determined during the interpolation (Deligiorgi and Philippopoulos, 2011).

Random forest for spatial predictions framework (RFSP): Random Forest (Biau and Scornet, 2016; Breiman, 2001; Cutler et al., 2012; Prasad et al., 2006) is a powerful spatial and spatiotemporal prediction technique (Bonilla-Bedoya et al., 2021; Hengl et al., 2015; Nussbaum et al., 2018; Prasad et al., 2006; Vaysse and Lagacherie, 2015) that responds efficiently to linear prediction methods without requiring rigorous statistical assumptions. The RFsp package (Wright and Ziegler, 2017) is an algorithm developed by Hengl et al. (2018) that includes a spatial prediction approach so that sampling locations and the overall sampling pattern are considered in the estimation of model parameters.

However, RFSP does not restrict the handling of different types of covariates (Hengl et al., 2018a; 2018b; Hengl and MacMillan, 2019). Thus, spatial prediction of the spatial variability and temporal dynamics of the tropospheric NO_2 column concentration was performed with R Project applying RFsp (Hengl et al., 2018a; 2018b; Hengl and MacMillan, 2019). Digital Elevation Model (DEM), slope, precipitation, temperature, and aspect were used as co-variables.

"RFSP", includes a spatial approach to prediction in the sense that sampling locations and the overall sampling pattern are considered in the estimation of model parameters.

$$Y(s) = f(XG) \tag{8}$$

where XG denotes the predictors that explain the geographic proximity and spatial relationship between observations (Hengl and MacMillan, 2019; Bonilla-Bedoya et al., 2021).

$$XG = dp1, dp2, \ldots, dpN \tag{9}$$

where N is number of predictors.

2.4. Comparison of the precision and accuracy of interpolation methods

The interpolation methods were evaluated by leave-one-out cross validation (Esmaeilbeigi et al., 2020; Losser et al., 2014; Yuval et al., 2017), where the interpolation accuracy is determined by iteratively removing an observed value from the dataset and re-estimating it using the remaining values. This procedure is repeated n times, where n is equal to the number of all samples. Leave-one-out cross-validation is embedded in the ArcGIS Geostatistical Analyst extension. The error is determined by calculating the difference between the observed and the estimated value. Under this approach, different statistical cross-validation parameters were used to evaluate the accuracy of the interpolation methods.

Based on a comprehensive review of the related literature, we focused on six cross-validation methods: mean error (ME), mean absolute error (MAE), root mean square error (RMSE), relative root mean square error (RRMSE), relative mean absolute error (RMAE), and coefficient of determination (R^2) (Arétouyap et al., 2016; Bezyk et al., 2021; Fontes and Barros, 2010; Li and Heap, 2011, 2014; Li et al., 2016; Losser et al., 2014; Ma et al., 2019; Nouck, 2015; Qu et al., 2017; Sajjadi et al., 2017; Sekulić et al., 2020; Tong et al., 2015; Van Roode et al., 2019; Yi et al., 2018). Most studies have used two or three parameters to determine the most appropriate interpolation method, with R^2 and RMSE being the most widely used (Le et al., 2020; Qiao et al., 2019; Wong et al., 2004; Yu et al., 2018). The details of all cross-validation parameters are presented in Table 1.

Additionally, the performance of six interpolation techniques were compared: IDW, LPI, EBK, OK, SK, and RFSP. This study was based on the monthly average time series of atmospheric NO_2 concentrations measured at the stratospheric level for the year 2020. The time series was selected for the purpose of observing the effect of COVID-19 restriction measures on tropospheric NO_2 concentrations in the study area. Restriction due to COVID-19 reduced transport activities, which resulted in less energy consumption, lower oil demand, and limited domestic travel. These changes in transportation activities had a significant impact on air quality. This behavior of reduction of pollutant gas emissions due to COVID-19 has been registered in some regions in the world regions in the world such as China, France, Italy, USA, Spain.

The parameters ME, MAE, RMSE, RMAE, RRMSE, and R^2 were calculated based on the monthly averages for the analyzed time series. Fig. 2 shows the results of the statistical analysis by cross-validation analysis, making use of the leave-one-out principle, and Table 2 shows the average of the statistical parameters of each evaluated model. The statistical parameters and the tropospheric NO_2 column concentration were calculated for each month of 2020 to study the spatial behavior of the models. The statistical parameters were calculated using the model results and the actual concentration values measured by Sentinel-5P. In Table 2 and Fig. 2, the values of the statistical parameters of the RFSP model are not shown because the model presented the worst performance of the six models

Table 1
Details on the cross-validation parameters used.

Cross-validation methods	Parameters
The Mean Error (ME) represents the arithmetic average of all estimated errors in the interpolation. It also indicates the direction (average) of the estimated errors. Negative bias represents an underestimation of the variable, while overestimation is represented by a positive bias. This parameter is used as an indicator of precision because the positive and negative estimates counteract each other, which makes the ME comparatively lower than the actual error.	$ME = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)$
The Mean absolute error (MAE) is the mean of the absolute values of the individual prediction errors in all instances of the test set. Each prediction error is the difference between the actual value and the predicted value for the instance.	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i\rvert$
The Root mean square error (RMSE) is the most commonly used cross-validation parameter. It is the square root of the mean square of all errors indicating how concentrated the data are around the line of best fit. It has the same unit of measurement as the predicted value.	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$
These two new measures and their modified versions can be interpreted as measures of relative errors in the estimates. They allow comparison of the performance of interpolation methods with different units/scales from various studies in different disciplines (Li and Heap, 2011; Prestwich et al., 2014).	$RMAE = \frac{1}{n}\sum_{i=1}^{n}\frac{\lvert y_i-\hat{y}_i\rvert}{\hat{y}_i} * 100$ $RRMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i-\hat{y}_i}{\hat{y}_i}\right)^2} * 100$
The Coefficient of determination, also known as R^2, is perhaps the most popular measure of goodness-of-fit for linear models. It describes how well the observed results are replicated by the model. An R^2 close to 1 implies a near perfect relationship between the model and the data, while an R^2 close to 0 implies failure to accurately model the data.	$R^2 = \frac{\sum_{i=1}^{n}\left[(y_i-\bar{y}_i)\left(\hat{y}_i-\overline{\hat{y}_i}\right)\right]^2}{\sum_{i=1}^{n}(y_i-\bar{y}_i)^2\sum_{i=1}^{n}\left(\hat{y}_i-\overline{\hat{y}_i}\right)^2}$
y_i: represents observed/measured value $\hat{y}_i$: is the predicted value $\bar{y}_i$: is the average of observed/measured values $\overline{\hat{y}_i}$: is the average of predicted values n: the number of observations in a validation dataset	

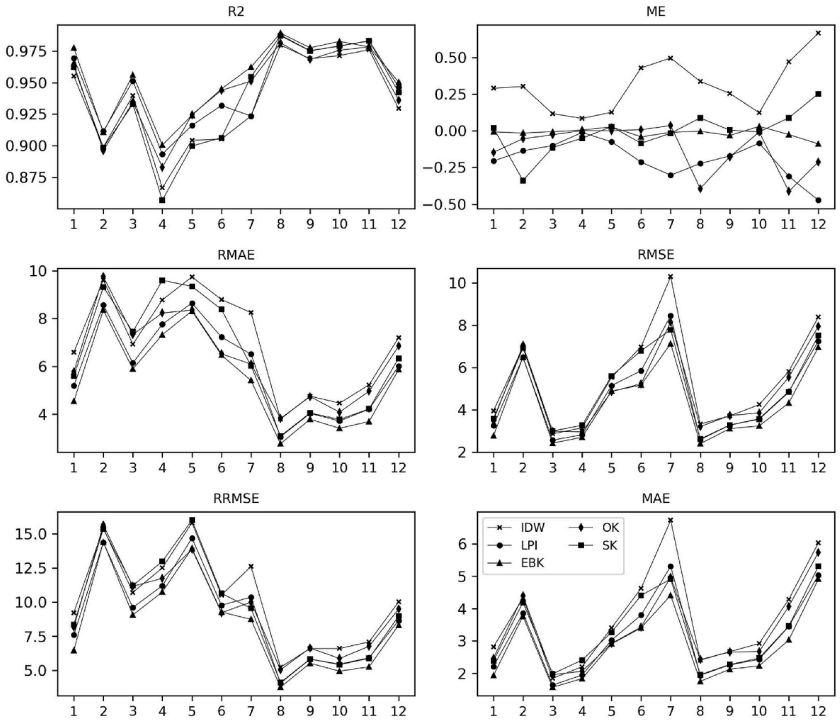

Fig. 2. The statistics of cross-validation results (R^2, ME, MAE, RMSE, RRMSE, and RMAE) of five interpolation methods (IDW, LPI, EBK, OK, and SK) for the 12 months of 2020.

Table 2
Average of the statistical results Coefficient of determination (R^2), Mean error (ME), Mean absolute error (MAE), Root mean square error (RMSE), Relative root mean square error (RRMSE), and Relative mean absolute error (RMAE) using cross-validation for the different interpolation methods for each month of 2020 pollutants in Guayaquil. The best values are in bold. Furthermore, the P value of the Pearson correlation coefficient is displayed.

Method	R^2	ME	MAE	RMSE	RRMSE	RMAE	P Value
IDW	0.9350	0.3079	3.6922	5.4295	10.1934	7.0133	<0.0001
LPI	0.9474	−0.1921	3.0792	4.6729	8.9615	5.9252	<0.0001
EBK	**0.9546**	−0.0139	**2.8270**	**4.2954**	**8.3673**	**5.4887**	<0.0001
OK	0.9449	−0.1154	3.3167	4.9140	9.4668	6.3789	<0.0001
SK	0.9399	**−0.0109**	3.2560	4.8909	9.5254	6.4330	<0.0001

analyzed with an RMSE = 16.55 and R^2 = 0.33, values with errors well above those of the other models; therefore, it was unnecessary to compare them.

Table 2 shows that the estimated values were significantly related to the observed values. EBK had the highest average R^2 value (0.9546), while IDW had the lowest value (0.9350). The order of suitability of the interpolation techniques according to the average R^2 was EBK > LPI > OK > SK > IDW. The ME, which measured bias, showed that the bias was very small (close to zero) for the SK method (−0.0109), while the IDW method (0.3079) had the most bias comparatively. The order of suitability of the interpolation techniques according to the average ME was SK > EBK > OK > LPI > IDW. RMSE and MAE are two similar measurement indicators for interpolation accuracy. The interpolation of NO_2 data resulted in an RMSE and MAE of 4.2954–5.4295 and 2.8270–3.6922, respectively. The highest accuracy was obtained with EBK (RMSE = 4.2954, MAE = 2.8270). The LPI method performed almost as well as EBK. The order of suitability of the interpolation techniques according to the average MAE and RMSE was EBK > LPI > SK > OK > IDW. RRMSE and RMAE are indicators used to evaluate the overall performance of the models and to compare them with other variables on a different unit/scale. The best performing model was EBK (RRMSE = 8.3673, RMAE = 5.4887), while IDW had the lowest performances (RRMSE = 10.1934, RMAE = 7.0133). The order of suitability of the interpolation techniques according to the average RMAE and RRMSE was EBK > LPI > OK > SK > IDW.

Among the different methods analyzed, IDW interpolation had the highest error. Overall, considering all statistical parameters, the EBK method had the best performance compared to the other methods, although the LPI method also produced equally good results. The EBK method outperformed the other models in all statistical parameters except ME. The best average results of the statistical parameters for each interpolation method are marked in bold in Table 2. Fig. 2 shows that our study, good cross-validation results suggested that EBK is a reliable interpolation method for tropospheric NO_2 in the city of Guayaquil.

As shown in Fig. 2, the interpolation methods present similar behaviors, with EBK being the method with the best statistical results during most of 2020, while IDW had the worst results. The spatial distribution of tropospheric NO_2 columns in the urban area of

Guayaquil was interpolated using IDW, LPI, EBK, OK, SK, and RFSP methods. To explore the effect of the six interpolation methods on the spatial variation of NO_2, monthly average maps were made from January to December 2020. All interpolation methods, except RFSP, reflected the general trend of tropospheric NO_2 columns (Fig. 3a and b).

The tropospheric NO_2 concentration showed a clear spatial heterogeneity in the study area. Almost all interpolation methods showed a similar general distribution and concentration pattern, indicating an increase in NO_2 concentrations from north to south in the first half of 2020. In contrast, the pattern increased from south to north in the second half of 2020. The spatial variations between the different interpolation methods were similar between the deterministic (IDW and LPI) and stochastic (EBK, OK, and SK) methods.

In addition, the findings were corroborated by Fig. 4, which presents scatter plots comparing the measured and predicted tropospheric NO_2 columns for the five methods (IDW, LPI, EBK, OK, and SK), excluding the RFSP stochastic method, which exhibited distinct scattering behaviors. Measured and estimated values were plotted for all months of 2020. The models adequately fit the actual values. On the other hand, in light of the fact that the calculated P-values, across all examined cases, are consistently lower than the pre-determined threshold of statistical significance ($\alpha = 0.05$), it can be inferred that the observed correlation in the present study is statistically significant.

In this paper, tropospheric NO_2 column concentration data from January to December 2020, obtained with the best performing EBK model, were used for monthly statistics, and the results showed that there was significant seasonal variation in Guayaquil (Table 3).

The minimum concentration of the NO_2 column in each month was approximately 7–24 μmol/m^2, with April having the lowest concentration (7.242 μmol/m^2). The highest NO_2 column concentrations in each month were significantly different and decreased from February to April, before again increasing to the highest level in July (175.086 μmol/m^2). The mean concentration of NO_2 in April was the lowest (26.951 μmol/m^2), while the mean value in December was the highest (90.639 μmol/m^2). March and April, in which the lowest mean NO_2 concentrations occurred (27.924 and 26.951, respectively), coincided with the lockdown measures taken in Ecuador due to the COVID-19 pandemic.

3. Discussion

The EBK method proved to be capable of providing higher accuracy when generating NO_2 maps with high spatial and temporal resolution, as it achieved an R^2 of 0.9546. This result contrasts with the study by Ma et al. (2021), where the EBK method was applied to improve the spatial resolution of satellite data, yielding an R^2 of 0.8230. This evidence demonstrates that the EBK method is precise when it comes to downscaling spatial resolution. On the other hand, the results observed by Banerjee et al. (2018) concluded that EBK is an ideal method for generating air pollution maps for CO, NO2, SO2, and Particulate Matter (PM). In contrast, in a study conducted by Sajjadi et al. (2017), the best interpolation method for PM was the IDW method. This contradicts what was found in this research, since IDW had a lower performance in the statistical parameters (Table 2). In our study, IDW presented an average $R^2 = 0.9350$, which was much higher than that found by Van Roode et al. (2019) when analyzing IDW to interpolate NO_2 concentrations in the Bay of Algeciras in Spain ($R = 0.72$). In the study developed by Fontes and Barros (2010), using cross validation, it was demonstrated that IDW was the best method to interpolate NO_2 data in an urban sensitive area. A possible reason for the different results could be the comparison approach or the distribution of the monitoring networks in conjunction with the study scale. For Qu et al. (2017), who studied four spatial interpolation methods, namely inverse distance weighting (IDW), radial basis function (RBF), ordinary Kriging (OK), and universal kriging (UK), the comparison of predicted values with measured values indicated that OK was the optimal method for analyzing the spatial distribution of N deposition.

In another study developed by Kumar et al. (2016), the interpolated results of air pollutants SO_2, NO_2, and suspended PM obtained better results with the IDW and kriging models. The RFSP model had the worst performance of the six evaluated methods. This is contradictory to the results for $PM_{2.5}$ obtained by Ma et al. (2019), who after comparing traditional spatial interpolation approaches (IDW and Ordinary Kriging) with machine learning models and neural networks, found that the latter obtained a higher R^2 value (0.6732) and a lower RMSE (0.0677), which are better results than those found in our study.

It is important to consider that the results obtained in our study have been compared with papers that have used monitoring networks to develop interpolation processes. This approach may be the main reason for the difference in the results since we used information from Sentinel-5P satellite images to generate the interpolation processes.

On March 16, 2020, a partial blockade began in Ecuador with curfew measures mainly from 21:00–5:00 UTC, and vehicle traffic was prohibited 4 days/week. On March 25, 2020, a total blockade opted for curfew measures from 14:00–5:00 UTC, and vehicle traffic was prohibited 4 days/week. Subsequently, on April 6, 2020, the strongest blockade was implemented with curfew measures from 14:00–5:00 UTC, and vehicle traffic was prohibited 6 days/week (Pacheco et al., 2020). As shown in Table 3, NO_2 concentration started to increase significantly after May, when the closure measures became less restrictive. Zalakeviciute et al. (2020) reported on an analysis of satellite data in Ecuador before and during the quarantine, which shows differences in the spatial concentration of NO_2 in relation to the application of protective measures in the country, especially in the city of Quito. Muhammad et al. (2020) reported that NO_2 concentrations in the tropospheric column were reduced by 20–30% in Spain due to the lockout, especially in major cities such as Madrid, Barcelona, and Seville. Huang and Sun (2020) indicated that COVID-19 control measures, including lockouts and refuge-in-place standards, led to 87–90% and 62–89% reductions in tropospheric NO_2 in Wuhan and Beijing, respectively. Using real-time measurements, Berman and Ebisu (2020) found evidence of a 25.5% reduction in NO_2 in the USA during the COVID-19 pandemic. In the context of the study area, Pacheco et al. (2020) performed an analysis considering the strongest blocking measures applied in Ecuador, which led to the discovery of a 35.0% tropospheric NO_2 reduction in Guayaquil.

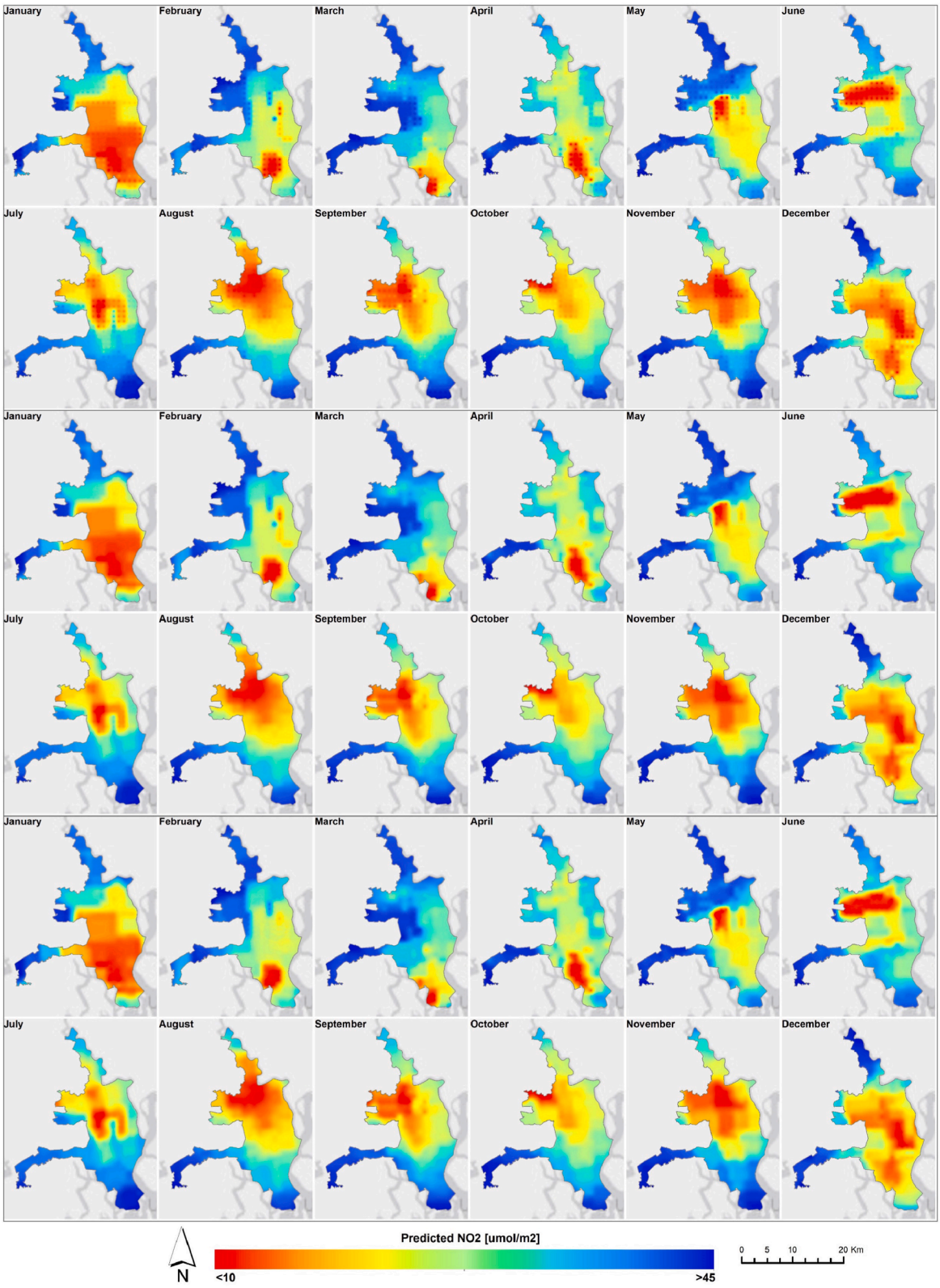

Fig. 3a. Spatial distribution of monthly tropospheric NO_2 column concentration (μmol/m^2) in Guayaquil city based on different interpolation methods (IDW, LPI and EBK).

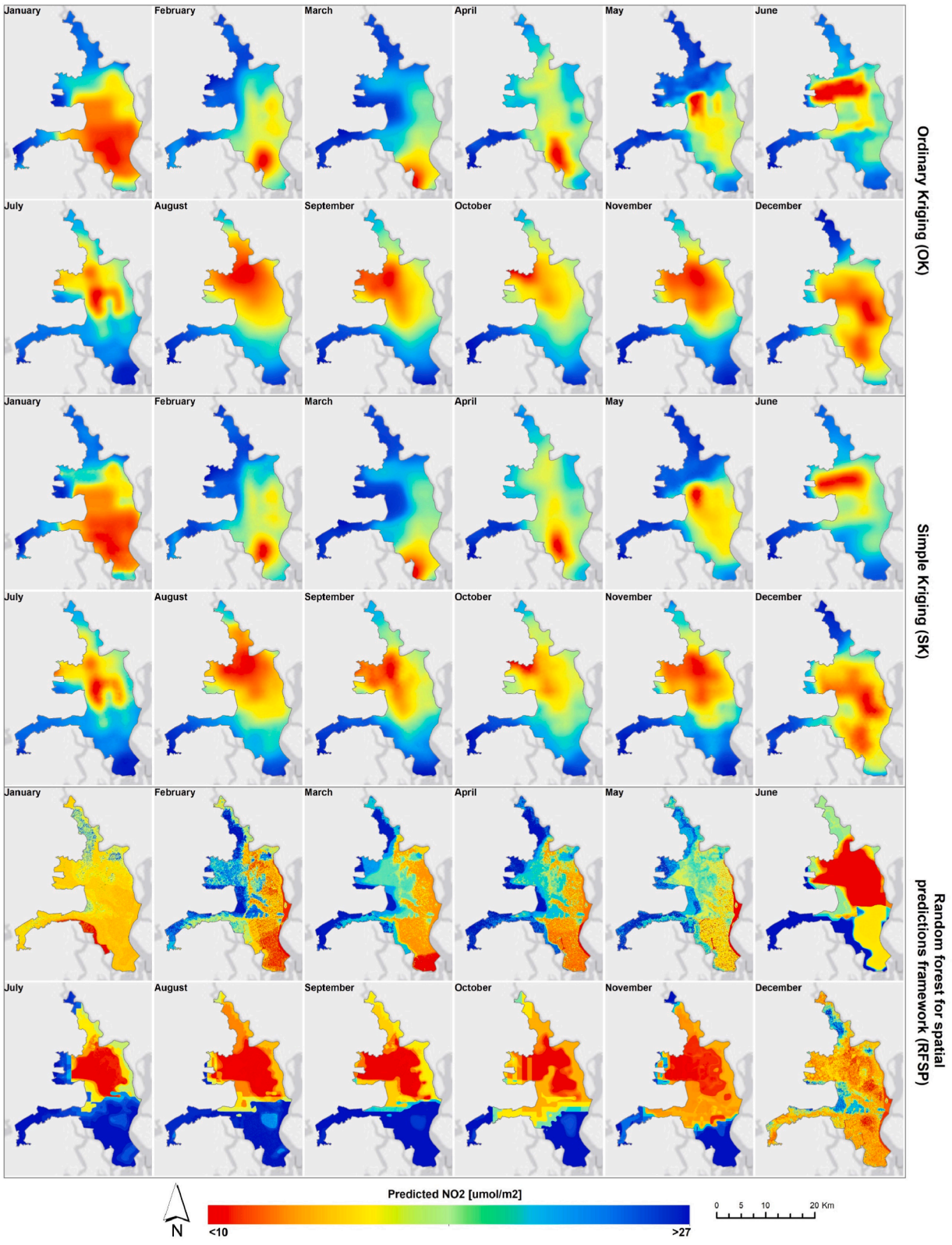

Fig. 3b. Spatial distribution of monthly tropospheric NO_2 column concentration (μmol/m^2) in Guayaquil city based on different interpolation methods (OK, SK, and RFSP).

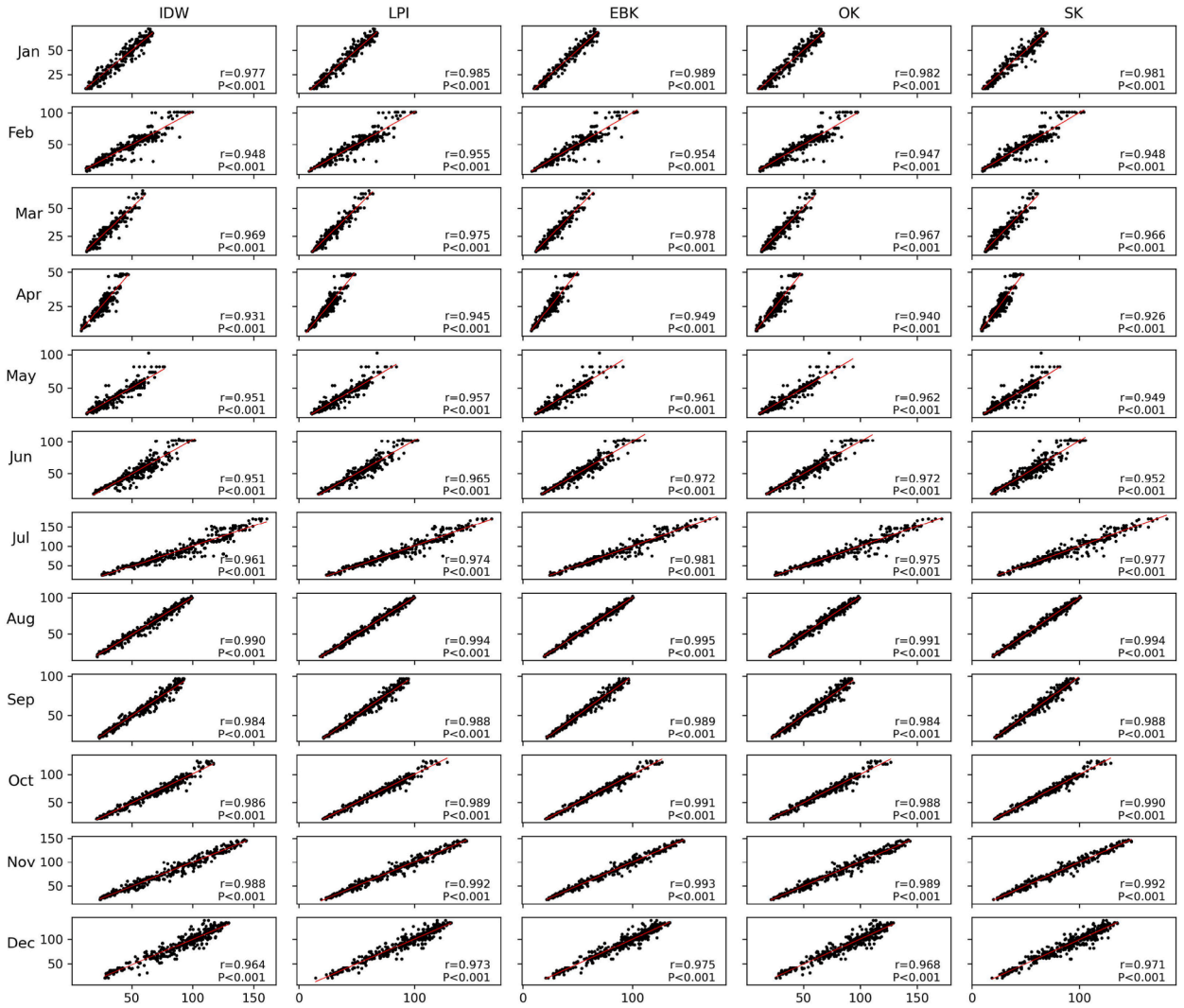

Fig. 4. Scatterplots of measured (y-axis) vs. predicted (x-axis) tropospheric NO_2 columns from IDW, LPI, EBK, OK and SK. The trendline, the Pearson coefficient (r) between the measured and estimated, and P value are shown.

Table 3
The monthly statistical results of tropospheric NO_2 column concentration for EBK interpolation.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Min	10.180	10.236	11.267	7.242	11.235	17.264	24.217	18.875	20.773	20.551	21.913	21.272
Max	70.839	103.461	64.346	49.447	100.231	107.276	175.086	101.717	97.366	123.850	147.040	135.358
Mean	45.528	48.597	27.924	26.951	38.092	57.008	86.453	66.541	58.823	68.046	87.274	90.639
Std	17.684	20.626	10.908	8.032	17.668	21.999	37.032	22.594	20.147	22.948	37.317	28.698

Despite the existence of national policies and regulations, the lack of effective implementation and limited supervision have resulted in a reduced impact on improving air quality. In 2012, the Air Quality Management Plan was published, containing actions and policies aimed at improving air quality and reducing atmospheric pollution levels in the country (Ministerio del Ambiente del Ecuador, 2012). However, in the study area, the lack of resources and political will has led to an absence of equipment for monitoring air quality. Therefore, Sentinel-5P satellite data with spatial downscaling could be useful for continuous monitoring of NO_2, thus strengthening public awareness of the health risks associated with air pollution (Anjum et al., 2021).

4. Conclusion

This research addresses the limitation of access to pollutant data in cities lacking continuous monitoring networks such as Guayaquil. The application of interpolation methods allows the measurement of air pollutant concentrations at the urban scale, which

was another limitation of the research. This will allow future research on the impact of air pollution on public health and the economy. This study assessed various interpolation methods for estimating the monthly tropospheric NO_2 column concentration in Guayaquil from January to December 2020. To achieve this goal, NO_2 data from Sentinel-5P was transformed into a centered point coordinate system. Six interpolation methods, IDW, LPI, EBK, OK, SK, and RFSP, were selected and tested. The different interpolation techniques were evaluated by employing leave-one-out cross validation and statistical criteria (R^2, ME, MAE, RMSE, RRMSE, and RMAE).

The interpolation results showed that EBK had the best ability to accurately produce the interpolated surface of the tropospheric NO_2 column concentration. Based on the results of the cross-validation technique, the EBK interpolation method generated the smallest mean error, except for ME (RMSE = 4.2954, MAE = 2.8270, RRMSE = 8.3673 and RMAE = 5.4887), and the highest coefficient of determination (R^2 = 0.9546) compared to the other five interpolation methods. LPI and EBK had similar results, suggesting that while using spatial correlation patterns like EBK makes intuitive sense, simpler methods like LPI work just as well. The best results in terms of accuracy were obtained by the EBK application. In addition, the study allowed us to identify that IDW could not generate interpolation surfaces with high performance for tropospheric NO_2 columns.

The results of the tropospheric NO_2 surface showed that the spatial distribution of NO_2 in Guayaquil increased in NO_2 concentration from north to south in the first half of 2020, while it increased from south to north in the second half of 2020. The spatial variations between the different interpolation methods were similar between the deterministic (IDW and LPI) and stochastic (EBK, OK, and SK) methods, except the stochastic RFSP method, which showed unusual behavior compared to the rest of the models.

Given that there are currently no monitoring stations in the city of Guayaquil that allow direct validation of the results obtained in this study, more studies should be performed in other urban areas to determine if similar patterns can be found. Furthermore, it is necessary to evaluate the effect of other variables, such as wind and atmospheric pressure, to improve the accuracy of the results. In conclusion, EBK can be considered a better interpolation method to interpolate tropospheric NO_2 column concentration data in densely populated areas. However, LPI proved its applicability as an alternative method.

Finally, while traditional interpolation methods are used in this research for NO_2 satellite data resampling, the natural continuation of the research guides us towards exploring new techniques such as Machine Learning and Deep Learning that can further improve the accuracy of the results. On the one hand, Deep Learning methods are increasingly popular in the field of artificial intelligence and Machine Learning that can be particularly useful for analyzing satellite data of atmospheric pollutants. On the other hand, machine learning algorithms could be used in future research because they are effective in different scenarios.

Credit author statement

Danilo Mejía: Writing, editing, Conceptualization, Methodology, Software, Validation Hermel Alvarez.: Data curation, Writing-Original draft preparation, Methodology. Rasa Zalakeviciute: Visualization, Investigation. Review. Diana Macancela: Reviewing and Editing. Carlos Sanchez: Reviewing and Editing. Santiago Bonilla: Writing- Reviewing and Editing.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Danilo Mejia reports financial support was provided by Corporación Ecuatoriana para el Desarrollo de la Investigación y la Academia.

Data availability

https://1drv.ms/u/s!AsnJ9fikkABsh6R3LzBszG0KDQyrvg?e=9n1WOv

Acknowledgement and funding

The authors would like to thank to Corporación Ecuatoriana para el Desarrollo de la Investigación y Academia - CEDIA for the financial support given to the present research, development, and innovation work through its CEPRA program, especially for the “Resiliencia ambiental en sistemas socio-ecológicos urbanos de ciudades neotropicales, casos Quito y Guayaquil” fund.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rsase.2023.100990.

References

Ali, G., Sajjad, M., Kanwal, S., Xiao, T., Khalid, S., Shoaib, F., Gul, H.N., 2021. Spatial–temporal characterization of rainfall in Pakistan during the past half-century (1961–2020), Scientific Reports. Nat. Publ. Group UK 11 (1), 1–15. https://doi.org/10.1038/s41598-021-86412-x.

Anjum, M.S., Ali, S.M., Imad-ud-din, M., Subhani, M.A., Anwar, M.N., Nizami, A.S., Ashraf, U., Khokhar, M.F., 2021. An emerged challenge of air pollution and ever-increasing particulate matter in Pakistan; A critical review. J. Hazard Mater. 402. https://doi.org/10.1016/j.jhazmat.2020.123943.

Arétouyap, Z., Njandjock Nouck, P., Nouayou, R., et al., 2016. Lessening the adverse effect of the semivariogram model selection on an interpolative survey using kriging technique. SpringerPlus 5, 549. https://doi.org/10.1186/s40064-016-2142-4.

Athira, V., Geetha, P., Vinayakumar, R., Soman, K.P., 2018. DeepAirNet: applying recurrent networks for air quality prediction. Proc. Comput. Sci. 132, 1394–1403. https://doi.org/10.1016/j.procs.2018.05.068. Elsevier B.V.

Banerjee, P., Ghose, M.K., Pradhan, R., 2018. AHP-Based spatial air quality impact assessment model of vehicular traffic change due to highway broadening in Sikkim himalaya. Ann. GIS. Taylor & Francis 24 (4), 287–302. https://doi.org/10.1080/19475683.2018.1534889.

Berman, J.D., Ebisu, K., 2020. Science of the total environment changes in U . S . air pollution during the COVID-19 pandemic. Sci. Total Environ. 739, 139864. https://doi.org/10.1016/j.scitotenv.2020.139864. Elsevier B.V.
Bezyk, Y., Sówka, I., Górka, M., Blachowski, J., 2021. Gis-based approach to spatio-temporal interpolation of atmospheric CO2 concentrations in limited monitoring dataset. Atmosphere 12 (3), 1–25. https://doi.org/10.3390/atmos12030384.
Biau, G., Scornet, E., 2016. A random forest guided tour. Test 25 (2), 197–227. https://doi.org/10.1007/s11749-016-0481-7.
Bonilla-Bedoya, S., López-Ulloa, M., Mora-Garcés, A., Macedo-Pezzopane, J.E., Salazar, L., Herrera, M.Á., 2021. Urban soils as a spatial indicator of quality for urban socio-ecological systems. J. Environ. Manag. 300 (December). https://doi.org/10.1016/j.jenvman.2021.113556.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1201/9780429469275-8.
Cersosimo, A., Serio, C., Masiello, G., 2020. TROPOMI NO2 tropospheric column data: regridding to 1 km grid-resolution and assessment of their consistency with in situ surface observations. Rem. Sens. 12 (14). https://doi.org/10.3390/rs12142212.
Cooper, M.J., Martin, R.V., McLinden, C.A., Brook, J.R., 2020. Inferring ground-level nitrogen dioxide concentrations at fine spatial resolution applied to the TROPOMI satellite instrument. Environ. Res. Lett. 15 (10), 104013. https://doi.org/10.1088/1748-9326/aba3a5.
Cutler, A., Cutler, D.R., Stevens, J.R., 2012. Random Forests. In: Zhang, C., Ma, Y. (Eds.), Ensemble Machine Learning. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-9326-7_5.
Daya, S.B.S., Cheng, Q., Agterberg, F., 2018. Handbook of Mathematical Geosciences: Fifty Years of IAMG, Handbook of Mathematical Geosciences: Fifty Years of IAMG. Springer International Publishing. https://doi.org/10.1007/978-3-319-78999-6.
de Foy, B., Lu, Z., Streets, D.G., Lamsal, L.N., Duncan, B.N., 2015. Estimates of power plant NOx emissions and lifetimes from OMI NO2 satellite retrievals. Atmos. Environ. 116, 1–11. https://doi.org/10.1016/j.atmosenv.2015.05.056.
Delgado, A., 2013. Guayaquil, Cities 31, 515–532. https://doi.org/10.1016/j.cities.2011.11.001.
Deligiorgi, D., Philippopoulos, K., 2011. Spatial interpolation methodologies in urban air pollution modeling: application for the greater area of metropolitan athens, Greece. Adv. Air Pollut. https://doi.org/10.5772/17734.
Engel-Cox, J.A., Hoff, R.M., Haymet, A.D.J., 2004. Recommendations on the use of satellite remote-sensing data for urban air quality. J. Air Waste Manag. Assoc. 54 (11), 1360–1371. https://doi.org/10.1080/10473289.2004.10471005.
Eskes, H., van Geffen, J., Boersma, F., Eichmann, K.-U., A, A., P, M., S, M., Veefkind, J.P., Loyola, D., 2021. S5P MPC product readme nitrogen dioxide. Tech. Rep. https://doi.org/10.5270/S5P-9bnp8q8. Report S5P-KNMI-L2-0021-MA, Issue.
Esmaeilbeigi, M., Chatrabgoun, O., Hosseinian-Far, A., Montasari, R., Daneshkhah, A., 2020. A low cost and highly accurate technique for big data spatial-temporal interpolation. Appl. Numer. Math. 153, 492–502. https://doi.org/10.1016/j.apnum.2020.03.009. Elsevier B.V.
European Commission and ESA, 2015. Sentinel-5 Precursor: DATA ACCESS AND PRODUCTS, (2), pp. 1–2. Available at: https://sentinel.esa.int/web/sentinel/sentinel-5p-tropomi-wiki.
Filonchyk, M., Peterson, M.P., 2023. An integrated analysis of air pollution from US coal-fired power plants. Geosci. Front. 14 (2), 101498. https://doi.org/10.1016/j.gsf.2022.101498.
Filonchyk, M., Hurynovich, V., Yan, H., 2021. Impact of covid-19 lockdown on air quality in the Poland, eastern europe. Environ. Res. 198, 110454. https://doi.org/10.1016/j.envres.2020.110454.
Fontes, T., Barros, N., 2010. Interpolation of Air Quality Monitoring Data in an Urban Sensitive Area, vol. 18. pp. 6–18.
Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone, Remote Sensing of Environment, pp. 18–27. https://doi.org/10.1016/j.rse.2017.06.031. The Author(s), 202.
Gribov, A., Krivoruchko, K., 2020. Empirical Bayesian kriging implementation and usage. Sci. Total Environ. 722, 137290. https://doi.org/10.1016/j.scitotenv.2020.137290. Elsevier B.V.
Griffin, D., Zhao, X., McLinden, C.A., Boersma, F., Bourassa, A., Dammers, E., et al., 2019. High-resolution mapping of nitrogendioxide with TROPOMI: first resultsand validation over the Canadianoil sands. Geophys. Res. Lett. 46, 1049–1060. https://doi.org/10.1029/2018GL081095Received. 7 NOV 2018Accepted 22 DEC 2018Accepted article online 28 DEC2018Published online 23 JAN 2019©2018.
Hengl, T., Macmillan, R.A., 2019. Predictive soil mapping with R. In: Predictive Soil Mapping with R. www.soilmapper.org.
Hengl, T., Heuvelink, G.B.M., Kempen, B., Leenaars, J.G.B., Walsh, M.G., Shepherd, K.D., Sila, A., MacMillan, R.A., De Jesus, J.M., Tamene, L., Tondoh, J.E., 2015. Mapping soil properties of Africa at 250 m resolution: random forests significantly improve current predictions. PLoS One 10 (6), 1–26. https://doi.org/10.1371/journal. pone.0125814.
Hengl, T., Nussbaum, M., Wright, M.N., Heuvelink, G.B.M., Gräler, B., 2018. Random forest as a generic framework for predictive modeling of spatial and spatio-temporal variables. PeerJ 2018 (8). https://doi.org/10.7717/peerj.5518.
Huang, G., Sun, K., 2020. Non-negligible impacts of clean air regulations on the reduction of tropospheric NO2 over East China during the COVID-19 pandemic observed by OMI and TROPOMI. Sci. Total Environ. 745, 141023. https://doi.org/10.1016/j.scitotenv.2020.141023. Elsevier B.V.
Ialongo, I., Virta, H., Eskes, H., Hovila, J., Douros, J., 2020. Comparison of TROPOMI/Sentinel-5 Precursor NO2 observations with ground-based measurements in Helsinki. Atmos. Meas. Tech. 13 (1), 205–218. https://doi.org/10.5194/amt-13-205-2020.
Jácome, G., Vilela, P., Yoo, C.K., 2019. Social-ecological modelling of the spatial distribution of dengue fever and its temporal dynamics in Guayaquil, Ecuador for climate change adaption. Ecol. Inf. 49, 1–12. https://doi.org/10.1016/j.ecoinf.2018.11.001. Elsevier B.V.
Janssen, S., Dumont, G., Fierens, F., Mensink, C., 2008. Spatial interpolation of air pollution measurements using CORINE land cover data. Atmos. Environ. 42 (20), 4884–4903. https://doi.org/10.1016/j.atmosenv.2008.02.043.
Kim, H.C., Lee, S.M., Chai, T., Ngan, F., Pan, L., Lee, P., 2018. A conservative downscaling of satellite-detected chemical compositions: NO2 column densities of OMI, GOME-2, and CMAQ. Rem. Sens. 10 (7). https://doi.org/10.3390/rs10071001.
Kn, S., Deka, L., Gupta, M., 2021. Use of Remote Sensing Data to Identify Air Pollution Signatures in India. pp. 1–17.
Le, V.D., Bui, T.C., Cha, S.K., 2020. Spatiotemporal deep learning model for citywide air pollution interpolation and prediction. In: Proceedings - 2020 IEEE International Conference on Big Data and Smart Computing. BigComp, pp. 55–62. https://doi.org/10.1109/BigComp48618.2020.00-992020.
Li, J., Heap, A.D., 2011. A review of comparative studies of spatial interpolation methods in environmental sciences: performance and impact factors. Ecol. Inf. 6 (3–4), 228–241. https://doi.org/10.1016/j.ecoinf.2010.12.003. Elsevier B.V.
Li, J., Heap, A.D., 2014. Spatial interpolation methods applied in the environmental sciences: a review. Environ. Model. Software 53, 173–189. https://doi.org/10.1016/j.envsoft.2013.12.008. Elsevier Ltd.
Li, L., Zhou, X., Kalo, M., Piltner, R.Li, L., et al., 2016. Spatiotemporal interpolation methods for the application of estimating population exposure to fine particulate matter in the contiguous U.S. and a real-time web application. Int. J. Environ. Res. Publ. Health 13 (8). https://doi.org/10.3390/ijerph13080749.
Losser, T., Li, L., Piltner, R., 2014. A spatiotemporal interpolation method using radial basis functions for geospatiotemporal big data. In: Proceedings - 5th International Conference on Computing for Geospatial Research and Application. COM, Geo, pp. 17–24. https://doi.org/10.1109/COM.Geo.2014.152014.
Ma, J., Ding, Y., Cheng, J.C., Jiang, F., Wan, Z., 2019. A temporal-spatial interpolation and extrapolation method based on geographic Long Short-Term Memory neural network for PM2.5. J. Clean. Prod. 237, 117729. https://doi.org/10.1016/j.jclepro.2019.117729. Elsevier Ltd.
Ma, X., Zhang, H., Han, G., Mao, F., Xu, H., Shi, T., Hu, H., Sun, T., Gong, W., 2021. A regional spatiotemporal downscaling method for CO2Columns. IEEE Trans. Geosci. Rem. Sens. 59 (10), 8084–8093. https://doi.org/10.1109/TGRS.2021.3052215.
Ministerio del Ambiente del Ecuador, 2012. Plan Nacional de Calidad del Aire del Ecuador.
Muhammad, S., Long, X., Salman, M., 2020. COVID-19 pandemic and environmental pollution: a blessing in disguise? Sci. Total Environ. 728, 138820. https://doi.org/10.1016/j.scitotenv.2020.138820. Elsevier B.V.
Muniraj, K., Panneerselvam, B., Devaraj, S., Jesudhas, C.J., Sudalaimuthu, K., 2021. Evaluating the effectiveness of emissions reduction measures and ambient air quality variability through ground-based and Sentinel-5P observations under the auspices of COVID pandemic lockdown in Tamil Nadu, India. Int. J. Environ. Anal. Chem. 0 (0), 1–12. https://doi.org/10.1080/03067319.2021.1902997. Taylor & Francis.
Narváez, A., Useche, J., 2020. The Kriging integration method applied to the boundary element analysis of Poisson problems. Eng. Anal. Bound. Elem. 121 (August), 1–20. https://doi.org/10.1016/j.enganabound.2020.09.001.

Núñez-Alonso, D., Pérez-Arribas, L.V., Manzoor, S., Cáceres, J.O., 2019. Statistical tools for air pollution assessment : multivariate and spatial analysis studies in the Madrid region. Anal. Methods Chem. 1–9. 2019.

Omrani, H., Omrani, B., Parmentier, B., 2020. Spatio-temporal data on the air pollutant nitrogen dioxide derived from Sentinel satellite for France. Data Brief 28, 105089. https://doi.org/10.1016/j.dib.2019.105089. Elsevier Ltd.

Pacheco, H., Díaz-López, S., Jarre, E., Pacheco, H., Méndez, W., Zamora-Ledezma, E., 2020. NO2 levels after the COVID-19 lockdown in Ecuador: a trade-off between environment and human health. Urban Clim. 34 (2), 100674. https://doi.org/10.1016/j.uclim.2020.100674. Elsevier.

Patel, V., Mistree, P.K., 2013. A Review on Different Image Interpolation Techniques for Image Enhancement, 3(12). pp. 129–133.

Prajapati, A., Naik, S., Mehta, S., 2012. Evaluation of different image interpolation algorithms. Int. J. Comput. Appl. 58 (12), 6–12. https://doi.org/10.5120/9332-3638.

Prasad, A.M., Iverson, L.R., Liaw, A., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. Ecosystems 9 (2), 181–199. https://doi.org/10.1007/s10021-005-0054-1.

Prestwich, S., Rossi, R., Armagan Tarim, S., Hnich, B., 2014. Mean-based error measures for intermittent demand forecasting. Int. J. Prod. Res. 52 (22), 6782–6791. https://doi.org/10.1080/00207543.2014.917771.

Qiao, P., Li, P., Cheng, Y., Wei, W., Yang, S., Lei, M., Chen, T., 2019. Comparison of common spatial interpolation methods for analyzing pollutant spatial distributions at contaminated sites. Environ. Geochem. Health 41 (6), 2709–2730. https://doi.org/10.1007/s10653-019-00328-0. Springer Netherlands.

Qu, L., Xiao, H., Zheng, N., Zhang, Z., Xu, Y., 2017. Comparison of four methods for spatial interpolation of estimated atmospheric nitrogen deposition in South China. Environ. Sci. Pollut. Control Ser. 24 (3), 2578–2588. https://doi.org/10.1007/s11356-016-7995-0.

Rebholz, B., Almekkawy, M., 2020. Efficacy of kriging interpolation in ultrasound imaging; subsample displacement estimation. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society. EMBS, pp. 2137–2141. https://doi.org/10.1109/EMBC44109.2020.9175457 2020-July.

Sağir, Ç., Kurtuluş, B., 2017. Hydraulic head and groundwater 111cd content interpolations using empirical bayesian kriging (Ebk) and geo-adaptive neuro-fuzzy inference system. geo-ANFIS), Water SA 43 (3), 509–519. https://doi.org/10.4314/wsa.v43i3.16.

Sajjadi, S.A., Zolfaghari, G., Adab, H., Allahabadi, A., Delsouz, M., 2017. Measurement and Modeling of Particulate Matter Concentrations: Applying Spatial Analysis and Regression Techniques to Assess Air Quality, vol. 4. MethodsX. Elsevier B.V., pp. 372–390. https://doi.org/10.1016/j.mex.2017.09.006.

Schneider, M., Ertl, B., Diekmann, C., Khosrawi, F., Röhling, A. N., Hase, F., ... & Pollard, D. F. 2021. Synergetic use of IASI and TROPOMI space borne sensors for generating a tropospheric methane profile product.

Sekulić, A., Kilibarda, M., Heuvelink, G.B., Nikolić, M., Bajat, B., 2020. Random forest spatial interpolation. Rem. Sens. 12 (10), 1–29. https://doi.org/10.3390/rs12101687.

Shelestov, A., Kolotii, A., Lavreniuk, M., Medyanovskyi, K., Vasiliev, V., Bulanaya, T., Gomilko, I., 2018. Air quality monitoring in urban areas using in-situ and satellite data within era-planet project. In: Ukraine Space Research Insti', IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium. Igor Sikorsky Kyiv Polytechnic Institute, Kyiv, pp. 1668–1671. IEEE.

Tong, Y., Yu, Y., Hu, X., He, L., 2015. Performance Analysis of Different Kriging Interpolation Methods Based on Air Quality Index in Wuhan, 2015 Sixth International Conference on Intelligent Control and Information Processing (ICICIP), vol. 5. pp. 331–335. Available at: https://ieeexplore.ieee.org/document/7388192/citations?tabFilter=papers.

Van Geffen, J., Eskes, H., Compernolle, S., Pinardi, G., Verhoelst, T., Lambert, J.-C., Sneep, M., Linden, M.Ter, Ludewig, A., Folkert Boersma, K., Pepijn Veefkind, J., 2021. Sentinel-5P TROPOMI NO 2 retrieval: impact of version v2.2 improvements and comparisons with OMI and ground-based data. Atmos. Meas. Tech. https://doi.org/10.5194/amt-2021-329.

Van Roode, S., Ruiz-Aguilar, J.J., González-Enrique, J., Turias, I.J., 2019. An artificial neural network ensemble approach to generate air pollution maps. Environ. Monitor. Assess. 191 (12). https://doi.org/10.1007/s10661-019-7901-6.

Vaysse, K., Lagacherie, P., 2015. Evaluating Digital Soil Mapping approaches for mapping GlobalSoilMap soil properties from legacy data in Languedoc-Roussillon (France). Geoderma Regional 4, 20–30. https://doi.org/10.1016/j. geodrs.2014.11.003.

Veefkind, J.P., Aben, I., McMullan, K., Förster, H., De Vries, J., Otter, G., Claas, J., Eskes, H.J., de Haan, J.F., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, H., Levelt, P.F., 2012. TROPOMI on the ESA Sentinel-5 Precursor: a GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. Remote Sens. Environ. 70–83. https://doi.org/10.1016/j.rse.2011.09.027. Elsevier Inc., 120(2012).

Vîrghileanu, M., Săvulescu, I., Mihai, B.A., Nistor, C., Dobre, R., 2020. Nitrogen Dioxide (NO2) Pollution monitoring with Sentinel-5P satellite imagery over Europe during the coronavirus pandemic outbreak. Rem. Sens. 12 (21), 3575. https://doi.org/10.3390/rs12213575.

Wang, S.W., Zhang, Q., Streets, D.G., He, K.B., Martin, R.V., Lamsal, L.N., Chen, D., Lei, Y., Lu, Z., 2012. Growth in NOx emissions from power plants in China: bottom-up estimates and satellite observations. Atmos. Chem. Phys. 12, 4429–4447. https://doi.org/10.5194/acp-12-4429-2012.

Wong, D.W., Yuan, L., Perlin, S.A., 2004. Comparison of spatial interpolation methods for the estimation of air quality data. J. Expo. Anal. Environ. Epidemiol. 14 (5), 404–415. https://doi.org/10.1038/sj.jea.7500338.

Wright, M.N., Ziegler, A., 2017. Ranger: a fast implementation of random forests for high dimensional data in C++ and R. J. Stat. Software 77 (1), 17. https://doi.org/10.18637/jss.v077.i01.

Wu, C.Y., Mossa, J., Mao, L., Almulla, M., 2019. Comparison of different spatial interpolation methods for historical hydrographic data of the lowermost Mississippi River. Spatial Sci. 25 (2), 133–151. https://doi.org/10.1080/19475683.2019.1588781. Taylor & Francis.

Xie, X., Semanjski, I., Gautama, S., Tsiligianni, E., Deligiannis, N., Rajan, R.T., et al., 2017. A review of urban air pollution monitoring and exposure assessment methods. ISPRS Int. J. Geo-Inf. 6 (12), 1–21. https://doi.org/10.3390/ijgi6120389.

Yang, T., Liu, W., 2018. Does air pollution affect public health and health inequality? Empirical evidence from China. J. Clean. Prod. 203, 43–52. https://doi.org/10.1016/j.jclepro.2018.08.242. Elsevier Ltd.

Yi, X., Zhang, J., Wang, Z., Li, T., Zheng, Y., 2018. Deep distributed fusion network for air quality prediction. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 965–973. https://doi.org/10.1145/3219819.3219822.

Yu, H., Russell, A., Mulholland, J., Odman, T., Hu, Y., Chang, H.H., Kumar, N., 2018. Cross-comparison and evaluation of air pollution field estimation methods. Atmos. Environ. 179, 49–60. https://doi.org/10.1016/j.atmosenv.2018.01.045. Elsevier Ltd.

Yuval, Levy, I., Broday, D.M., 2017. Improving modeled air pollution concentration maps by residual interpolation. Sci. Total Environ. 598, 780–788. https://doi.org/10.1016/j.scitotenv.2017.04.117. Elsevier B.V.

Zalakeviciute, R., Vasquez, R., Bayas, D., Buenano, A., Mejia, D., Zegarra, R., Diaz, A., Lamb, B., 2020. Drastic improvements in air quality in Ecuador during the COVID-19 outbreak. Aerosol Air Qual. Res. 20 (8), 1783–1792. https://doi.org/10.4209/aaqr.2020.05.0254.

Zalakeviciute, R., Alexandrino, K., Mejia, D., Bastidas, M.G., Oleas, N.H., Gabela, D., Rybarczyk, Y., 2021. The effect of national protest in Ecuador on PM pollution'. Sci. Rep. 11 (1), 1–12. https://doi.org/10.1038/s41598-021-96868-6. Nature Publishing Group UK.

Zheng, Z., Yang, Z., Wu, Z., Marinello, F., 2019. Spatial variation of NO2 and its impact factors in China: an application of sentinel-5P products. Rem. Sens. 11 (16), 1–24. https://doi.org/10.3390/rs11161939.

Zhi, S., Wei, Y., Cao, Z., Hou, C., 2017. Intelligent Controlling of Indoor Air Quality Based on Remote Monitoring Platform by Considering Building Environment. Icsai), pp. 627–631.