

## Article

# Downscaling MERRA-2 Reanalysis PM<sub>2.5</sub> Series over the Arabian Gulf by Inverse Distance Weighting, Bicubic Spline Smoothing, and Spatio-Temporal Kriging

Youssef Saliba <sup>1</sup>  and Alina Bărbulescu <sup>2,\*</sup> 

<sup>1</sup> Doctoral School, Technical University of Civil Engineering of Bucharest, 122-124 Lacul Tei Av., 020396 Bucharest, Romania; [youssefsaliba@gmail.com](mailto:youssefsaliba@gmail.com)

<sup>2</sup> Department of Civil Engineering, Transilvania University of Braşov, 5 Turnului Str., 900152 Braşov, Romania

\* Correspondence: [alina.barbulescu@unitbv.ro](mailto:alina.barbulescu@unitbv.ro)

**Abstract:** This study offers a detailed analysis of the fine particulate matter (PM<sub>2.5</sub>) series in the Arabian Gulf zone, employing three interpolation models, Inverse Distance Weighting (IDW), Bicubic Spline Smoothing (BSS) and Spatio-Temporal Kriging (STK). Unique advancements include the use of complete temporal records in IDW, the management of edge effects in S with synthetic buffer points, and the application of STK to detrended data residuals. The results indicated that the BSS, particularly adept at handling boundary conditions, significantly outperformed the other methods. Compared to IDW, the Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE) decreased by 21%, 15%, and 21%, respectively, in BSS. Compared to STK, MAE, RMSE, and MAPE were lower with around 60%, 61%, and 58%, respectively in BSS. These findings underscore the efficacy of the BSS method in spatial interpolation for environmental monitoring, contributing to enhanced PM<sub>2.5</sub> analysis and public health management in the region.

**Keywords:** PM<sub>2.5</sub>; spatial interpolation; IDW; BS; STK



**Citation:** Saliba, Y.; Bărbulescu, A. Downscaling MERRA-2 Reanalysis PM<sub>2.5</sub> Series over the Arabian Gulf by Inverse Distance Weighting, Bicubic Spline Smoothing, and Spatio-Temporal Kriging. *Toxics* **2024**, *12*, 177. <https://doi.org/10.3390/toxics12030177>

Academic Editor: Jian Sun

Received: 19 January 2024

Revised: 14 February 2024

Accepted: 22 February 2024

Published: 25 February 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Fine particulate matter (PM<sub>2.5</sub>) is a mixture of solid particles and liquid droplets with diameters lower than 2.5 µm, easily inhalable, and found in the air. They contain metallic and organic compounds, elemental carbon, inorganic ions, etc. [1]. Particulate matter (PM) mainly originates from fossil fuel or wood combustion, volcanoes, forest fires, the Earth's crust degradation, dust storms, sea spray, and biological emissions [2–4]. They may also result from chemical reactions between precursor gases [5]. Due to their size, density, and atmospheric conditions, PM<sub>2.5</sub> remains suspended, polluting the air [6]. They can be inhaled and deposited on the lung surface, damaging tissues and producing inflammation. The most affected are the children and adults who already suffer from lung or heart diseases.

Exposure for up to 24 h leads to respiratory symptoms, asthma attacks, and acute and chronic bronchitis, whereas exposure for months or years results in premature death [7]. The report on the ambient air quality indicates that in California, during 2014–2016, 2800 hospitalizations/year for respiratory and cardiovascular diseases and 6700 visits/year to emergency rooms due to asthma provoked or aggravated by PM<sub>2.5</sub> were noticed [1]. Statistics show that PM<sub>2.5</sub> is responsible for about 4 million deaths from respiratory infections, chronic lung disease, cancers, heart disease, preterm births, etc. A review of the harmful effects of PM<sub>2.5</sub> on the organism is presented in [8].

The World Bank statistics indicate that in the United Arab Emirates (UAE), the mean annual exposure to PM<sub>2.5</sub> surpasses eight times the safety limits established by the World Health Organization (WHO). Even if the sandstorms contribute to the deterioration of the air quality, studies indicate that the emissions (mainly from fossil fuels) represent the

primary source of air pollution in the UAE [9–13]. In September 2023, the 30 governmental monitoring stations in the UAE indicated that the PM<sub>2.5</sub> levels were three times higher than the warning limits established in 2021 by WHO.

One-fifth of the UAE's anthropic pollution with PM<sub>2.5</sub> is due to road transport, whereas about two-thirds result from industry [14]. In 2019, the mean annual exposure to PM<sub>2.5</sub> in the UAE was 44 µg/m<sup>3</sup>, compared to the global average—46, USA—8, Brazil—13, Russia—12, China—48, Saudi Arabia—62, India—83, and Qatar—76 [15,16].

Given the impacts on human health and public health management, many studies focus on monitoring and estimating pollutant concentration [17,18]. Due to the limited number of monitoring stations, spatial interpolation is necessary to evaluate PM concentrations at places where data are not available. Still, one must consider that PM<sub>2.5</sub> can vary rapidly from one site to the other, and a minimum number of stations is necessary to produce a good mapping. Various methods have been used for this goal. They include Inverse Distance Weighting (IDW) or some of its versions for PM<sub>2.5</sub> trend evaluation [19,20], kriging [21,22], neural networks [23–26], generalized additive mixed models [27], Multiscale Geographically Weighted Regression (MGWR) [28], Bayesian Kriging, and Tensor Spline Function [29], Random-forest-spatio-temporal kriging [30], hierarchical modeling [31], hybrid models [32], data fusion [33], optimal interpolation [34], exponentially smoothing [35], etc. Each one addressed some issues of the classical interpolation techniques, intending to increase the forecast accuracy.

This article's aim is in range with the mentioned approaches. It fills a gap in the knowledge of the PM<sub>2.5</sub> series analysis carried out by other scientists (using statistical analysis [36] and PM<sub>10</sub>/PM<sub>2.5</sub> ratio [37]) by introducing a refined version of Bicubic Spline Smoothing (BSS). It enhances the estimation of PM<sub>2.5</sub> concentrations by effectively managing corner and edge effects, integrating synthetic buffer points around the spatial domain of the data with values derived from the mean of adjacent real data points, thereby ensuring the continuity and precision of the study area's periphery. Complementing this, we also present an adapted IDW technique specifically tailored for spatio-temporal analysis. Incorporating the complete temporal records from each monitoring station into the Leave-one-Out-Cross-validation (LOOCV) process and optimizing the power parameter based on spatial and temporal data. Furthermore, we explore a spatio-temporal kriging (STK) approach, underlining the critical role of detrending in spatio-temporal analysis. The STK technique first employs a Generalized Additive Mixed Model (GAMM) to isolate the residuals for subsequent analysis. It then selects the best fit from various variogram models using a Bayesian approach. This dual-phase process, combining GAMM detrending with STK, allows us to capture the nuanced spatio-temporal dynamics of PM<sub>2.5</sub> concentrations, ensuring a robust interpolation that accounts for both spatial correlations and temporal variability.

Given the high pollution recorded, the Arabian Gulf zone was chosen for this study. Despite some reports providing data on high PM<sub>2.5</sub> concentrations in the UAE, no extended study was performed. Moreover, reliable interpolation methods can provide values of the PM<sub>2.5</sub> concentrations at places where measurements are not available. The results of the study could be used by the decision factors for taking insight measures for pollution reduction (and, implicitly, reducing its impact on population health).

## 2. Materials and Methods

### 2.1. Data Series

The second Modern-Era Retrospective analysis for Research and Applications (MERRA-2) is produced by the NASA Global Modeling and Assimilation Office (GMAO) [38–40] using the Goddard Earth Observing System, version 5 (GEOS-5) and data assimilation system [41]. The GEOS-5 incorporates atmospheric composition and circulation, land and oceanic components, as well as aerosol processes based on the Goddard Chemistry, Aerosol, Radiation, and Transport (GOCART) model, and can analyze the atmosphere composition and interactions between the climate and aerosols [41–43].

Apart from meteorological parameter data assimilation [44], MERRA-2 incorporates bias-corrected 550 nm aerosol optical depth (AOD) from assimilation from Advanced Very High-Resolution Radiometer, Moderate Resolution Imaging Spectroradiometer, assimilated the non-bias-corrected 550 nm aerosol optical depth from the Multiangle Imaging Spectroradiometer, and ground-sources (Aerosol Robotic Network stations) [37,40,44,45].

Based on GEOS-5 and GOCART, CMAO simulated the following aerosol types: sulfate ( $\text{SO}_4$ ), black and organic carbon (BC and OC, respectively), sea salt (SS), and dust (Dust). The  $\text{PM}_{2.5}$  concentrations were computed in MERRA-2 using the concentrations from the GOCART aerosol module, with

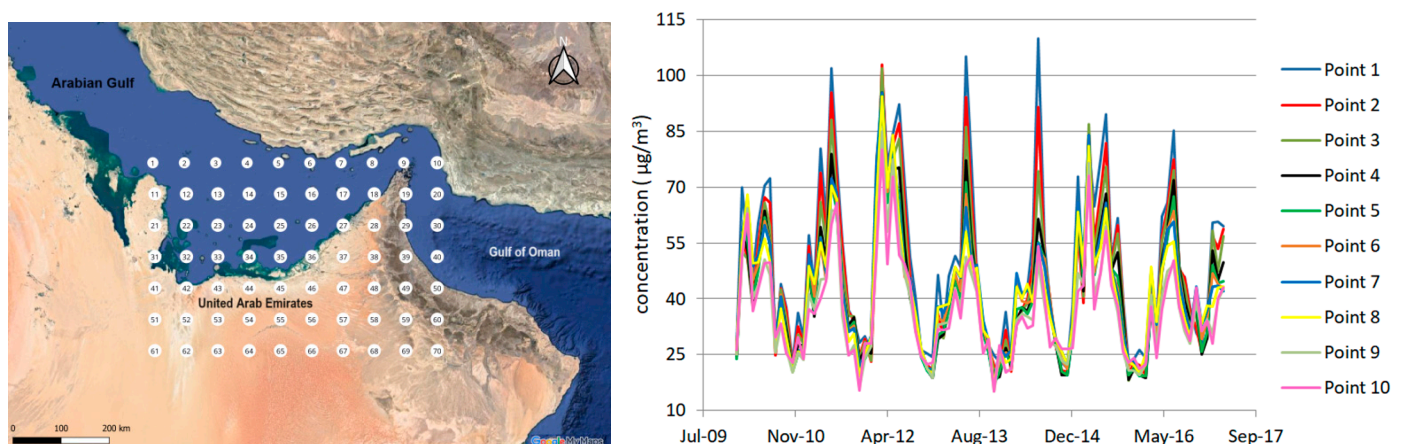
$$\text{PM}_{2.5} = 1.375 \times \text{SO}_4 + \text{BC} + 1.6 \times \text{OC} + \text{SS}_{2.5} + \text{Dust}_{2.5}, \quad (1)$$

where the index 2.5 at SS and Dust indicates that the PM diameter is less than  $2.5 \mu\text{m}$  [46].

Information on  $\text{PM}_{2.5}$ -related chemical composition is provided by satellite-derived products. Optical signals fall in this class, which must be modeled for obtaining surface mass concentrations. Due to clouds and high surface albedo, missing data (nonrandom) are inherent, raising difficulties in quickly obtaining updated  $\text{PM}_{2.5}$  components' concentrations with full spatial coverage. The reanalysis data provide the optimal estimates of these concentrations. Variational data assimilation algorithms can produce error estimates for the studied variables. Although they are concordant with the background and observation error statistics, they are challenging to obtain and limited by the algorithm's assumed errors. So, the main uncertainties are due to the algorithms used for modeling the signals to obtain surface mass concentration, 'completing' the randomly missing data, and the errors from the data assimilation algorithms [47–52].

In this research, we used the Dust Column Mass Density  $\text{PM}_{2.5}$  concentration ( $\mu\text{g}/\text{m}^3$ ) series for January 1990–April 2017 from the `tavgM_2d_aer_Nx`, which is a time-averaged two-dimensional monthly mean data collection in MERRA-2 [53].

The MERRA-2 grid points are represented on the map in Figure 1 (left), and their coordinates are given in Table A1. They are in the Arabian Gulf zone and over the United Arab Emirates. The dataset is complete (without gaps). The series from the first ten points are presented in Figure 1 (right). All exhibit a similar seasonal trend.



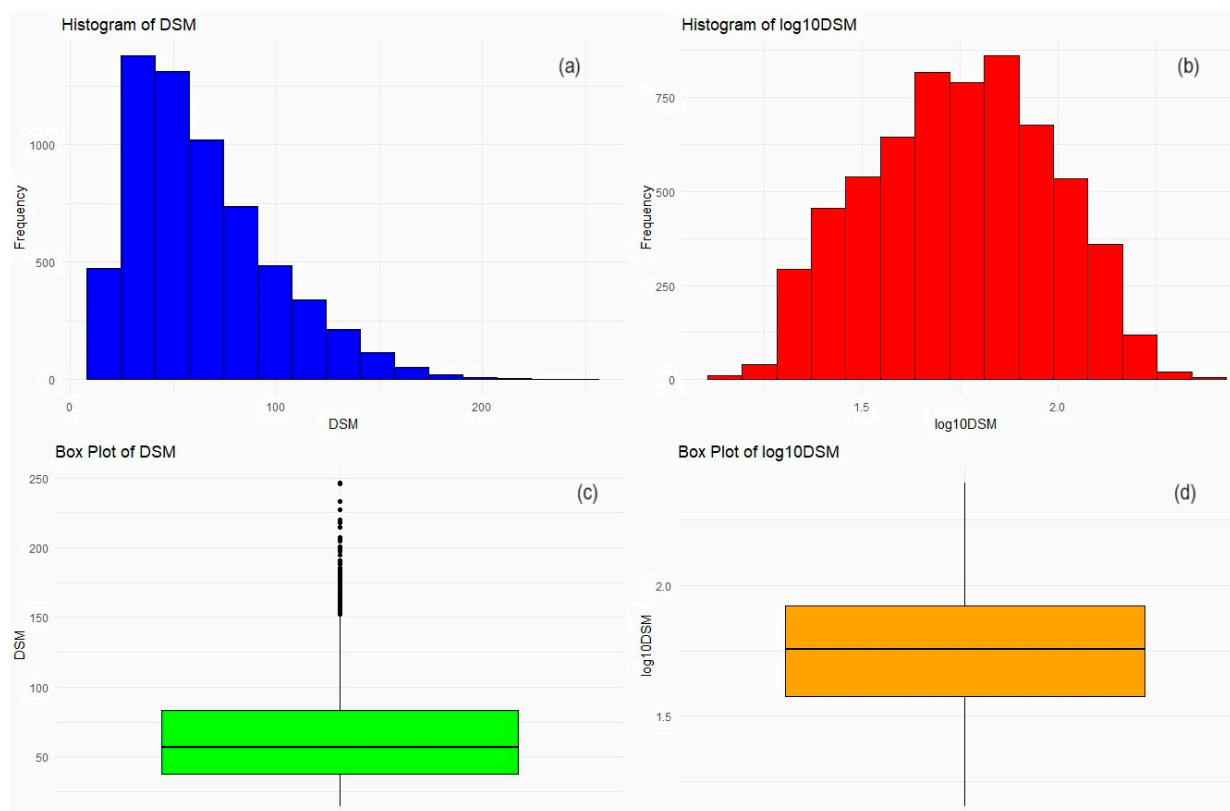
**Figure 1.** (left) Map of the region and the positions of the observation points; (right) the data series recorded at the first ten points.

The range of the basic statistics of the data series is presented in Table 1. There are significant differences between the series values, indicated, for example, by the range of the average values— $[3.65 \times 10^{-8}, 1.19 \times 10^{-7}]$ , maximum values— $[8.01 \times 10^{-8}, 2.46 \times 10^{-7}]$ , and standard deviations— $[1.31 \times 10^{-8}, 3.71 \times 10^{-8}]$ .

**Table 1.** Extreme values of the PM<sub>2.5</sub> data series' basic statistics.

	Min ( $\mu\text{g}/\text{m}^3$ )	Max ( $\mu\text{g}/\text{m}^3$ )	Average ( $\mu\text{g}/\text{m}^3$ )	Stdev ( $\mu\text{g}/\text{m}^3$ )	Cv (%)	Skew	Kurt
min	14.30	80.10	36.49	13.19	29.69	0.07	−1.10
max	51.90	246.00	119.36	37.34	46.15	1.17	1.92

The normality hypothesis was rejected by the Kolmogorov–Smirnov (K-S) test [54,55]. The hypothesis that the series originate from the same distribution was also rejected by the Kruskal–Wallis (K-W) test [56]. The set obtained considering all the series together, denoted by DSM, presented a highly positive skew and outliers (Figure 2), which can pose significant challenges in spatial modeling—Section 2.2.3.



**Figure 2.** (a) Histogram of DSM, (b) histogram of log<sub>10</sub>DSM, (c) boxplot of DSM, and (d) boxplot of log<sub>10</sub>DSM.

A log<sub>10</sub> transformation effectively reduces skewness and diminishes the outliers' impact by compressing the data's range, which can be concluded from the K–S results presenting a D statistic score of 0.028086 (suggesting a closer fit to the normal distribution). However, even after this transformation, the dataset did not fully conform to the normality assumptions, as indicated by the small *p*-value of 0.0001204. Figure 2 depicts the histograms and boxplots of DSM before and after the log<sub>10</sub> transformation.

## 2.2. Methodology

### 2.2.1. IDW Interpolation

IDW is one of the oldest spatial interpolation methods used in different research fields [57–59], with various versions aiming to improve the parameter selection [60–63].

IDW estimates values at unsampled locations using the data recorded at neighbor locations. The formula used for this purpose is

$$\hat{z}(x_0) = \left( \sum_{i=1}^n \frac{z(x_i)}{d_i^\beta} \right) / \left( \sum_{i=1}^n \frac{1}{d_i^\beta} \right), \quad (2)$$

where:

- $n$  is the number of sampling points,
- $\hat{z}(x_0)$  is the estimated value at the site  $x_0$ ,
- $z(x_i)$  is the value recorded at the site  $x_i$ ,
- $d_i$  is the distance from  $x_0$  and  $x_i$ ,
- $\beta > 1$  is a parameter [64].

According to Equation (1), the closer the location is to  $x_0$ , the higher the contribution in computing the estimated value [64]. The default  $\beta$  utilized in most applications is 2, called the inverse squared distance (ISD) interpolator [65].

The main disadvantage of this method is the arbitrary choice of  $\beta$  and low performances for clustered or unevenly distributed points. Such points located at similar distances from the target location will have approximately the same weight, and so have almost the same contribution to the estimated value [66].

In optimizing the IDW model, the focus lies on selecting the most effective power parameter  $\beta$  using LOOCV. This optimization process involves evaluating a range of power parameters (from 1 to 10) to determine their influence on the model's accuracy. During each iteration, for a specific  $\beta$ , LOOCV is conducted, wherein the model's predictive performance is assessed by temporarily omitting the entire records from a point at a time from the dataset [67].

Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE) were used to assess the model's performance. These metrics provide a comprehensive model evaluation, with RMSE as the primary criterion for selecting the optimal  $\beta$ . The power parameter that results in the lowest RMSE is identified as optimal, ensuring precision and reliability in the model's spatial predictions.

### 2.2.2. Spline Smoothing

The key idea behind spline interpolation is to fit a function consisting of local polynomials of degree  $p$  through the data points. These polynomials describe pieces of a line or a surface and can be constrained to pass through all the known data points while maintaining smoothness. Among the types of splines used in spatial interpolation, the most common is the cubic spline. It is defined over small intervals, and its cubic polynomials are stitched together at specific points, called knots. These knots are the known data points where the curve must pass through. Alternatively, knots away from the data points can be fitted using least squares or other methods to produce smoothing splines [68].

The most significant advantage of cubic spline interpolation (CSI) is the smoothness of the curve, capturing the overall data shape. The basic minimum curvature technique makes CSI more convenient for gently varying surfaces. However, we must be aware that errors can be expected in cases of uneven distribution. In such cases, the tension spline interpolation technique can be applied [68,69].

BSS is an extension of CSI for interpolating data points on a two-dimensional regular grid. The interpolated surface is defined by a piecewise polynomial function in two variables,  $x$  and  $y$ . In each rectangular cell formed by the data points, the interpolating function is a bicubic polynomial:

$$z = P(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j \quad (3)$$

where  $a_{ij}$  are the coefficients of the polynomial.

We utilized the ‘**interp()**’ function of the **akima** package in R [70] with the parameter `linear = FALSE` to implement bicubic interpolation. The method offers smooth surface fitting for our spatial dataset. Despite setting `extrap = TRUE` to enable the extrapolation of values beyond the convex hull of our data points, we encountered instances of NA results, particularly at the edge points of the rectangular grid (points 1, 3–9). Note that, in this context, ‘NA’ signifies missing interpolated values for some locations due to the insufficient nearby data points that BSS requires to compute a value. This limitation arises because the interpolation relies on data that fall within the range of observed values, and points at the periphery may lack sufficient neighboring data to guide the extrapolation process. The issue is particularly pronounced during LOOCV when edge points are omitted, leading to significant uncertainty and the potential for less accurate predictions as the algorithm stretches beyond its reliable bounds.

To address this challenge, our R code introduces a solution by creating a boundary buffer: a series of synthetic points added around the grid’s perimeter. This technique extends the convex hull to include these new points, effectively converting extrapolation scenarios into interpolation ones. The synthetic points are placed just beyond the original grid’s extent, and their associated values are derived from the mean of the nearest actual data points, ensuring a smooth gradient that aligns with the known data distribution. By integrating this boundary buffer into our LOOCV process, we facilitate the ‘**interp()**’ function’s ability to perform interpolation across the entire grid, including the previously problematic edge points. This approach mitigates the occurrence of NA results and enhances the precision of our spatial predictions, leveraging the strengths of bicubic interpolation across the augmented dataset.

### 2.2.3. Spatio-Temporal Kriging

Kriging represents a comprehensive suite of generalized least-squares regression algorithms for spatial interpolation. Apart from deterministic methodologies, this approach optimizes weights in a linear predictor to achieve the lowest possible average error in interpolation [71]. STK expands upon the foundational principles of traditional kriging, enabling the analysis of data that encompass both spatial and temporal aspects, which can be particularly advantageous. STK aims to proficiently forecast values at unobserved spatio-temporal points, utilizing the intricate patterns of spatial and temporal correlations within the dataset. STK’s strength lies in its ability to provide statistically robust interpolation, even at boundary points. It can make better-informed predictions at the edges of a grid by considering the spatial relationship of points within the dataset. However, its complexity requires a thorough understanding of the underlying statistical models and assumptions, making it more challenging to implement than simpler methods such as IDW or BSS.

At the core of STK lies the concept of Gaussian processes, which can be thought of as a generalization of the normal distribution of functions. It is a probabilistic model assuming that every point in some continuous input space is associated with a normally distributed random variable [72]. Understanding Gaussian processes helps in comprehending how STK makes predictions at new spatio-temporal points by considering both the mean and variability of the data, providing a statistical framework for our analysis [73].

Considering that a sequence of observations is sampled from a process that can be decomposed into a true spatio-temporal process (assumed Gaussian) and an observational error, the process is expressible through spatio-temporal fixed effects attributed to various covariates. The observational error is modeled as a spatio-temporal-dependent random process. In this case, the process can be modeled by

$$Z(s_{ij}, t_j) = Y(s_{ij}, t_j) + \varepsilon(s_{ij}, t_j), \quad i = 1, \dots, m_{t_j}, \quad j = 1, \dots, T, \quad (4)$$

where for each  $t \in \{t_1, \dots, t_T\}$  there are  $m_j$  observations [66].

This formulation effectively separates the observed data into a deterministic component influenced by specific covariates and a stochastic component that accounts for the randomness inherent in the observation process.

The overarching objective is to employ the dataset  $z$  (which includes time- and space-specific observations) to build a model of the random field  $Z$  [74] aiming to predict values at unobserved spatio-temporal locations or to enable simulations based on its conditional distribution. It assumes that  $Z$  maintains stationarity and exhibits spatial isotropy, allowing it to be defined by a mean function and a covariance function. With an aptly chosen covariance function, one can ascertain the covariance matrices essential for the linear predictor. By applying algebraic methods such as those used in established spatial analyses, it becomes feasible to accurately predict the values of the intended random field [75]. In this article, one covariance model was implemented from the **gstat** package, as described in [74]. The separable model assumes that the global spatio-temporal covariance could be expressed as the product of a spatial and a temporal term, with a variogram given by

$$\gamma_{sep}(h, \tau) = sill \times (\overline{\gamma}_s(h) + \overline{\gamma}_t(\tau) - \overline{\gamma}_s(h)\overline{\gamma}_t(\tau)), \quad (5)$$

where  $\overline{\gamma}_s(h)$  and  $\overline{\gamma}_t(\tau)$  are the standardized spatial and temporal variograms with separate nugget effects and (joint) sill.

For this study, we explored diverse variogram types (exponential, spherical, Gaussian, and Matérn) that best fit the data series and evaluated them in relation to the covariance model under consideration.

A direct implementation of the above procedure was not possible, given the particularities of the series set. Outliers in spatial and temporal datasets can markedly distort the predictive modeling in kriging, leading to skewed semivariograms and, thus, unreliable interpolation results. Data transformation techniques play a pivotal role in counteracting this. Transformations such as the logarithmic, square root, or more adaptive methods such as the Box–Cox transformation re-scale the data, diminishing the influence of outliers. These transformations compress the range of extreme values, reducing their leverage on the analysis. By normalizing the data distribution and minimizing the variance introduced by outliers, these transformations facilitate a more accurate representation of spatial and temporal autocorrelation, which is central to kriging.

Moreover, pre-processing datasets, particularly the removal of spatio-temporal trends to focus on residuals, is an indispensable step that significantly bolsters the interpretability and precision of kriging outcomes. The detrending process separates the core spatial structure and inherent temporal dynamics from the observed variability, thus enabling kriging to finely tune into the intrinsic autocorrelation of the residuals without the confounding influence of broader trends. Working directly with residuals circumvents the complications introduced by non-stationary behavior attributable to deterministic trends, leading to enhanced prediction accuracy and a profound comprehension of the stochastic elements of the data. Therefore, we adopted a modeling approach using generalized mixed models (GMM) and linear mixed effects (LME) models to remove spatio-temporal trends and work with the residuals.

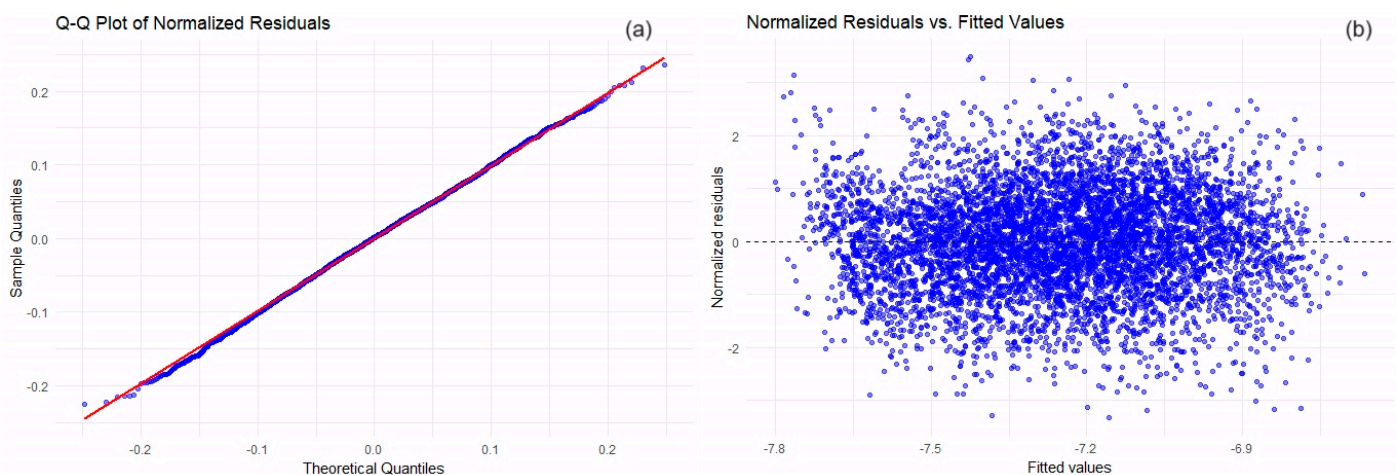
To address the normality issue that remained after the log10 transformation (Section 2.1), we initially adopted a straightforward approach using linear models by implementing the **lm()** function, which fits linear surfaces to the data. This first step involves exploring simple relationships, such as first-order (straight) or second-order (curved) surfaces, to understand the basic trends in the series set. However, the data series is collected over sea and desert areas, and these basic linear models often fail. They might not adequately capture the complex and nuanced relationships inherent in the data, such as the non-linear patterns and the unique characteristics of different sampling points. Therefore, we transitioned to more advanced statistical models such as generalized additive models (GAM), GAMM, and LME [76].

GAMs are particularly adept at uncovering complex, non-linear relationships within a data series. For instance, they can help us understand how the concentration levels might increase non-uniformly in response to changing environmental factors such as humidity or temperature. This aspect is particularly relevant when considering the diverse nature of the data from various geographical locations. Building on this, GAMM adds another layer

of sophistication by incorporating ‘random effects’. These are useful in our scenario, where each of the 70 distinct points presents a unique environmental characteristic, allowing modeling these differences while analyzing the overall data trends.

LME models come into play due to their strength in handling linear relationships and their capability to address fixed effects (common trends across all points) and random effects (unique characteristics of each point). In the context of our dataset, which exhibits a clear seasonal pattern with peaks every six months, LMEs are particularly suitable. They can effectively capture these predictable temporal trends while accounting for spatial variations across locations. By focusing on the residuals from the LME component, we can discern aspects of the seasonal variation and the site-specific differences that simpler linear models do not fully explain.

GAMM from the ‘mgcv’ package in R was utilized to model the spatial field and temporal trends. The thin plate regression splines ‘t20’ within GAMM allowed for flexible shaping of the spatial and temporal effects, capturing the complex underlying patterns in the data. The ‘`gamm()`’ function was used to fit these models, which internally uses the ‘`lme()`’ function to estimate random effects, providing a robust framework for detrending the dataset. After fitting the GAMM, we extracted the residuals representing the detrended data—the pure spatial-temporal stochastic component we aim to model with kriging. The residuals follow a normal distribution almost perfectly with an adequate, stable variance as indicated by a  $D = 0.013273$  and a  $p\text{-value} = 0.2279$  of the K–S test and by the visual verification of diagnostic plots in Figure 3.



**Figure 3.** Diagnostic plots for GAM model residuals analysis. (a) Q-Q plot of the normalized residuals; (b) normalized residuals vs. fitted values.

For LOOCV, we systematically removed each station from the dataset and used the remaining stations to predict the left-out station’s values. We performed spatio-temporal kriging on these residuals, ensuring that the predictions were based on the stochastic properties rather than any underlying deterministic trends.

The final step involved back-transforming the predictions to the original series scale to interpret the results in their natural units. It was accomplished by first predicting the trend component for each data point using the ‘`predict()`’ function on the fitted GAMM model. Then, we combined this trend component with the kriged residuals to obtain the final predictions in the log10 scale. To correct the bias introduced by the log transformation, a correction factor of  $10^{\sigma^2/2}$  was calculated, where  $\sigma^2$  is the variance of the log10-transformed predictions. The final predictions were then scaled back to the original series units by reversing the log10 transformation and applying the correction factor. This back-transformation is crucial as it allows the results to be presented on the same scale as the original measurements, making them directly comparable and interpretable in the context of the study.

The entire LOOCV process was automated in R software, where we evaluated various variogram models, including spherical, exponential, Gaussian, and Matérn, to accurately capture the dataset's complex spatio-temporal dynamics. This systematic approach involved fitting each model to the detrended data to identify the ones that best represent the spatio-temporal correlations. We kept the spatio-temporal combination of variogram models that yielded the lowest MSE.

To gauge the overall performance of the models, MAE, RMSE, and MAPE were calculated for each station and then averaged across all periods. MAE provides the evaluation of the systematic error. RMSE better explains the random errors. RMSE is valuable in situations where significant errors are particularly undesirable, as it helps identify models with occasional large deviations from the actual values. MAPE is beneficial when dealing with data where the magnitude varies significantly or when comparing the accuracy of models across different data scales. Values close to zero indicate the models' good performances.

Three more efficiency indicators were also used for comparing the models:

- Nash–Sutcliffe Efficiency (NSE) is a normalized statistic that determines the relative magnitude of the residual variance compared to the measured data variance. It is defined by

$$NSE = 1 - \frac{\sum_{i=1}^N (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^N (Q_{obs,i} - \bar{Q}_{obs})^2} \quad (6)$$

where  $Q_{obs,i}$  = observed data,  $Q_{sim,i}$  = simulated data,  $\bar{Q}_{obs}$  = mean of observed data, and  $N$  = number of time steps. The range of NSE is  $(-\infty, 1)$ .  $NSE = 1$  indicates a perfect match between the model and the observations. A value of 0 shows that the model predictions are as accurate as the average of the observed data. Still, NSE is sensitive to both high and low values.

- Kling–Gupta Efficiency (KGE) decomposes the NSE into components representing correlation, variability, and bias, providing a more comprehensive measure of model performance. The formula is

$$KGE = 1 - (r - 1)^2 + (\alpha - 1)^2 + (\beta - 1)^2 \quad (7)$$

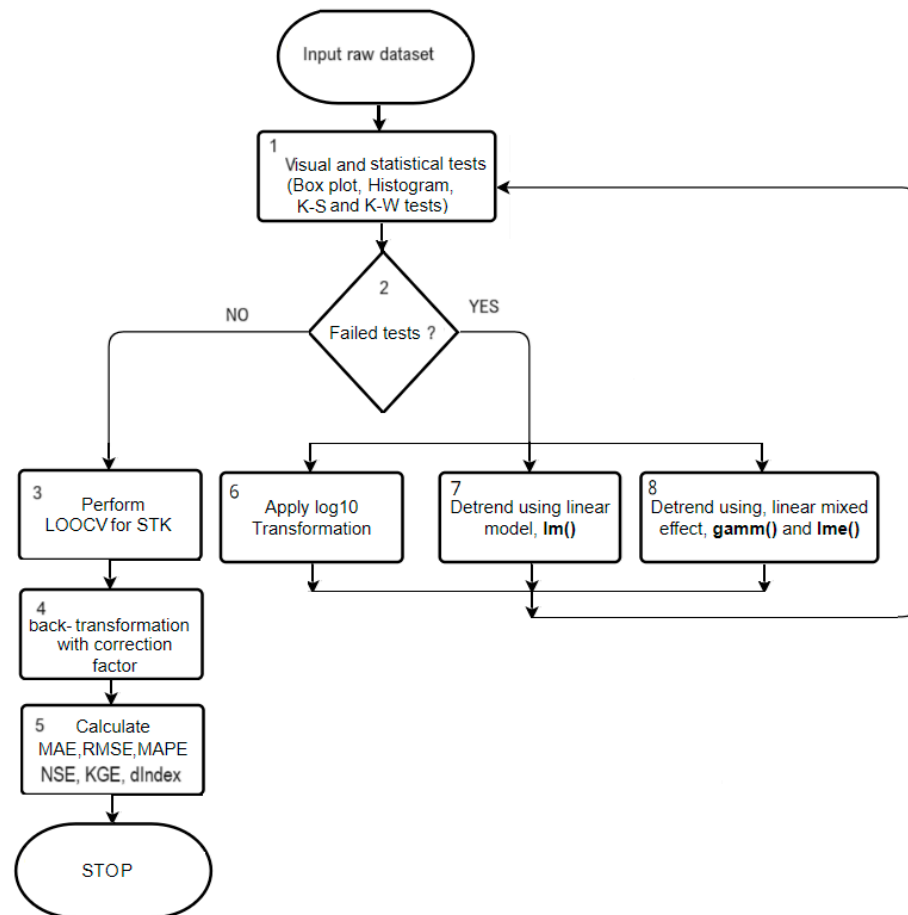
where  $r$  = correlation between observed and simulated data,  $\alpha$  = ratio of the standard deviation of simulated to observed data, and  $\beta$  = ratio of the mean of simulated to observed data. The benefits of using KGE are as follows:

- (1) The decomposition allows us to diagnose which aspect of the model is contributing to inefficiencies.
  - (2) KGE balances different performance aspects, providing a more holistic view of model performance.
  - (3) Using it with NSE to assess model performance leads to making informed decisions about model improvements.
- Index of Agreement (dIndex) is a standardized measure of model prediction error and represents the ratio of the mean square error and the potential error. It is given by

$$dIndex = 1 - \frac{\sum_{i=1}^N (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^N (|Q_{sim,i} - \bar{Q}_{obs}| + |Q_{obs,i} - \bar{Q}_{obs}|)^2}. \quad (8)$$

dIndex measures the magnitude of the error and its pattern and distribution. It is effectively an agreement measure, assessing how well the predicted data match the observed data range. Moreover, it can be used across different disciplines, including hydrology, climatology, and ecology.

Finally, we present a comprehensive workflow diagram (Figure 4) and its corresponding algorithmic description. Algorithm 1 details the step-by-step process of analyzing data series through STK, including data preprocessing, transformation, and validation stages, culminating in the prediction and error quantification phases.



**Figure 4.** Workflow diagram for STK analysis of data series.

**Algorithm 1.** The algorithm for the data series interpolation using STK has the following stages

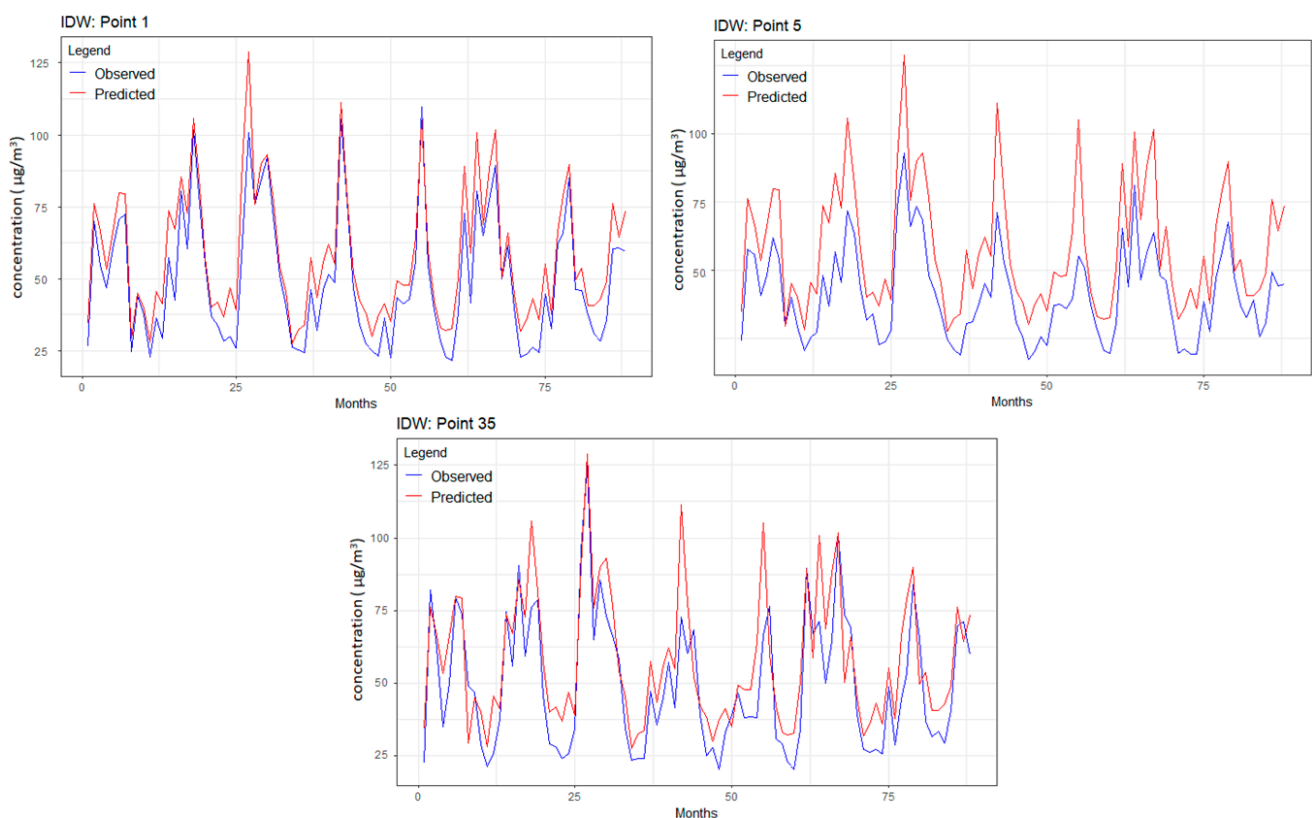
0. Input: Raw dataset
1. Perform Exploratory Data Analysis (EDA) on raw dataset
  - a. Visual inspections: Box plot, Histogram
  - b. Performs the K-S and K-W tests
2. Assess EDA Results: IF the dataset passes normality and homoscedasticity tests, THEN GOTO Step 3, ELSE, GOTO Step 6.
3. Perform LOOCV for STK
4. Back-transformation of Predicted Data:
  - a. Predict the trend for each data point (if data was detrended in Steps 7 or 8)
  - b. Apply the correction factor and reverse any log transformations
5. Calculate Error Metrics: MAE, RMSE, and MAPE END Algorithm
6. Data Transformation (IF needed after Step 2)
  - a. Apply log10 transformation to the data
  - b. GOTO Step 1 (Reapply EDA)
7. Simple Linear Detrending (IF Step 6 fails to normalize data)
  - a. Detrend data using a simple linear model with the 'lm()' function
  - b. GOTO Step 1 (Reapply EDA)
8. Complex Model Detrending (IF Step 7 fails to normalize data)
  - a. Detrend data using complex models like GAMM and LME
  - b. GOTO Step 1 (Reapply EDA)

### 3. Results and Discussion

In the presentation of the results, we strategically select points from the spatial grid, chosen to represent diverse locations: points 1, 10, 61, and 70 representing the corners, point 35 representing the central area and surrounded by a higher density of data points, and points 5, 30, 31, 65 representing an edge of the grid, allowing us to demonstrate the robustness of each model under various spatial scenarios, ranging from data-rich central areas to potentially data-sparse edges and corners. Furthermore, we globally assess errors across all stations for each method. The key error metrics—MAE, RMSE, and MAPE—are computed for these predictions.

#### 3.1. IDW Results

Figure 5 depicts the predicted and observed series values for the study period for points 1, 5, and 35. In all cases, the forecast values were higher than the data series, but the predicted series had the same shape as the raw one. The goodness of fit parameters for the nine points above are presented in Table 2, together with the minimum, maximum, and average MAEs, RMSEs, and MAPEs computed for all the sampling points.



**Figure 5.** Observed and predicted values of the series from points 1, 5, and 35 (the months are counted from 1 (January 2010) to 88 (April 2017)).

MAEs and RMSEs were extremely low, indicating the goodness of fit of IDW with the optimum parameters  $\beta$ . The highest MAPE among the points mentioned above corresponded to point 1, located above the Gulf in the northeastern corner of the grid due to uneven distribution of the neighbors. By comparison, MAPEs computed for the points in the other corners (3.115, 3.848, and 5.524) were under the average MAPE, indicating a good fit for the data series.

The series from the points in the central areas were not the best fit given the contribution of the series inside the UAE and over the Arabian Gulf (with different concentrations of  $PM_{2.5}$ ). The first incorporates the  $PM_{2.5}$  originated from industrial sources and dust storms from the desert (particularly in spring and the beginning of autumn) [10,11]. The

second reflects the influence of the dust transported over the Arabian Gulf from North Africa [77–80].

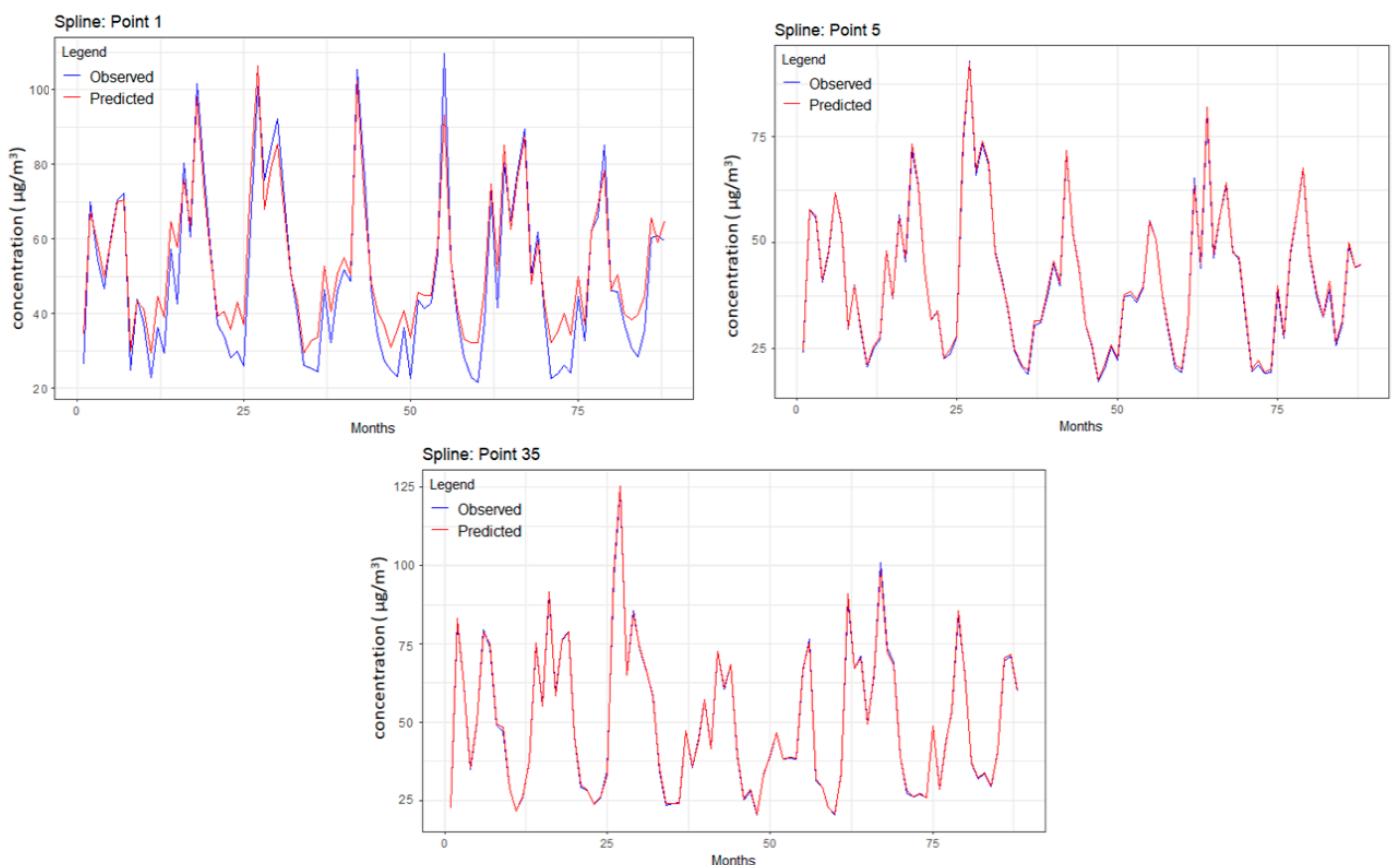
**Table 2.** MAE, RMSE, and MAPE in the IDW interpolation with the optimal  $\beta$ .

	min	max	average	Point 1	Point 5	Point 35
MAE	0.26	10.5	3.33	8.62	1.85	2.88
RMSE	0.037	11.2	4.03	10.3	2.67	3.13
MAPE	0.544	22.8	5.58	21.478	4.161	6.957
	Point 10	Point 61	Point 70	Point 30	Point 31	Point 65
MAE	1.26	4.58	4.16	2.05	7.3	0.7
RMSE	2.19	5.59	5.74	2.46	8.37	1.09
MAPE	3.115	3.848	5.524	5.863	7.955	0.737

This approach to optimizing the IDW model underscores a balance between methodological rigor in parameter selection and flexibility in model customization, ensuring robustness and adaptability in spatial data analysis.

### 3.2. Results of BSS

All predicted series showed an excellent concordance with the recorded series, as exemplified in Figure 6 for points 1, 5, and 35. The worst performances were recorded in point 1 (at the grid corner) compared to 5 and 35.



**Figure 6.** Observed and predicted concentration series for points 1, 5, and 35.

The explanation is that the corners or sides have a smaller number of neighbors, so a lower estimation quality is expected. Among the points in the grid corners, the series from 1 was also the worst fit, whereas the best fit was the series registered at point 61.

Remember that the original spline algorithm returned NA for the series from 1, and the buffering technique implemented here might introduce some errors in the series forecast. Even in these conditions, the MAE, RMSE, and MAPE corresponding to point 1 (Table 3 column 5) were not the highest among the other MAPEs, indicating that the proposed approach yielded good results in forecasting the PM<sub>2.5</sub> series. The same is true for the series from points 3–9, for which the algorithm provided outstanding results in terms of all goodness-of-fit indicators (with values lower than the corresponding averages).

**Table 3.** MAE, RMSE, and MSE in the splines interpolation after using buffering points.

	min	max	average	Point 1	Point 5	Point 35
MAE	0.38	12.69	2.96	5.55	1.22	2.51
RMSE	0.55	16.11	3.71	6.72	1.64	2.96
MAPE	0.709	23.834	5.022	15.556	2.623	5.961
	Point 10	Point 61	Point 70	Point 30	Point 31	Point 65
MAE	1.77	0.81	2.16	2.29	4.29	1.90
RMSE	2.30	1.14	2.69	2.61	4.86	2.31
MAPE	3.569	0.844	5.363	4.671	6.772	2.002

Table 3 presents the MAE, RMSE, and MAPE for the same selected points, and the maximum, minimum, and mean values of the goodness of fit parameters for all 70 points. MAPE for point 61 was close to the minimum MAPE. Moreover, MAEs and RMSEs for all the series recorded at the points on the grid edges were under the average MAE and RMSE, indicating the fit quality provided by the implemented interpolation procedure.

Other approaches can be explored to refine the BSS potential. One could use a weighted average that considers the distance of each neighboring point, giving closer points a greater influence on the value of the synthetic point.

Another might involve spatial trend analysis, where the synthetic points’ values are informed by identified spatial gradients or patterns within the data—this could be particularly useful if the variables show systematic changes across the study area.

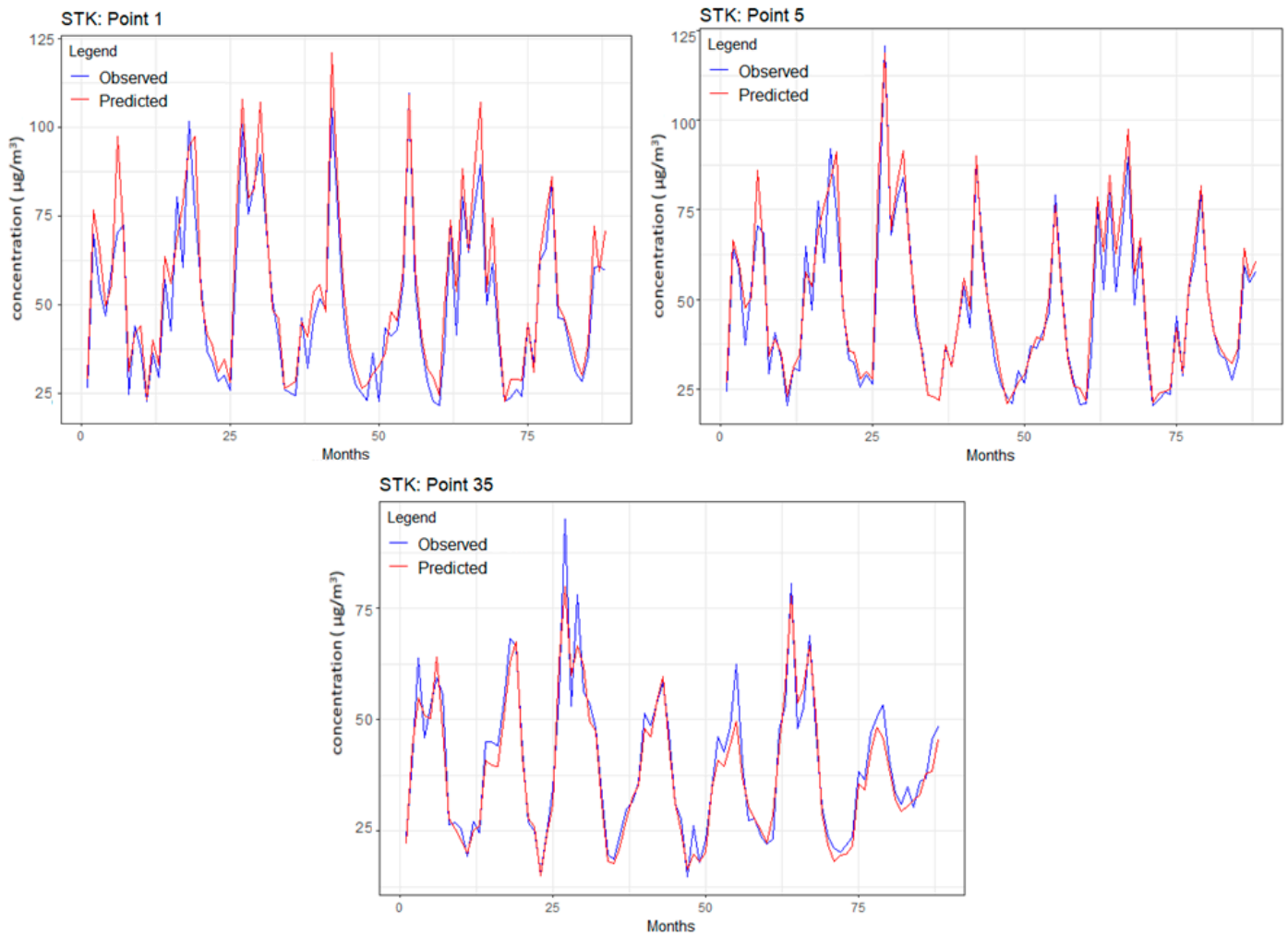
### 3.3. Results of STK

Table 4 contains the minimum, maximum and average values for the selected points analyzed in the previous sections. Among the series displayed in Table 4, the best fit was the series from point 35 (situated inside the grid), followed, unexpectedly, by that from the grid corner—point 10.

**Table 4.** MAE, RMSE, and MSE in the STK interpolation.

	min	max	average	Point 1	Point 5	Point 35
MAE	3.14	15.31	7.37	8.58	7.24	3.92
RMSE	3.86	20.62	9.64	10.65	9.85	5.29
MAPE	7.656	20.863	12.000	20.863	16.717	10.319
	Point 10	Point 61	Point 70	Point 30	Point 31	Point 65
MAE	4.17	12.13	5.35	5.23	6.20	14.40
RMSE	5.19	16.06	7.02	6.48	7.93	19.34
MAPE	9.107	14.183	12.693	10.672	10.425	13.677

Figure 7 depicts the predicted and observed values for the series from points 1, 5, and 35, respectively. A visual investigation of the chart from Figure 7 indicates that the predicted series mostly slightly overestimated the recorded ones. The series extremes are not always captured very well. These remarks are confirmed by the MAPE, which varied between 7.656 and 20.863.



**Figure 7.** Observed and predicted series for points 1, 5, and 35.

Considering the transformation performed on the data series, it was expected to obtain slightly higher errors than in the other models since each operation introduces additional errors.

### 3.4. Discussion

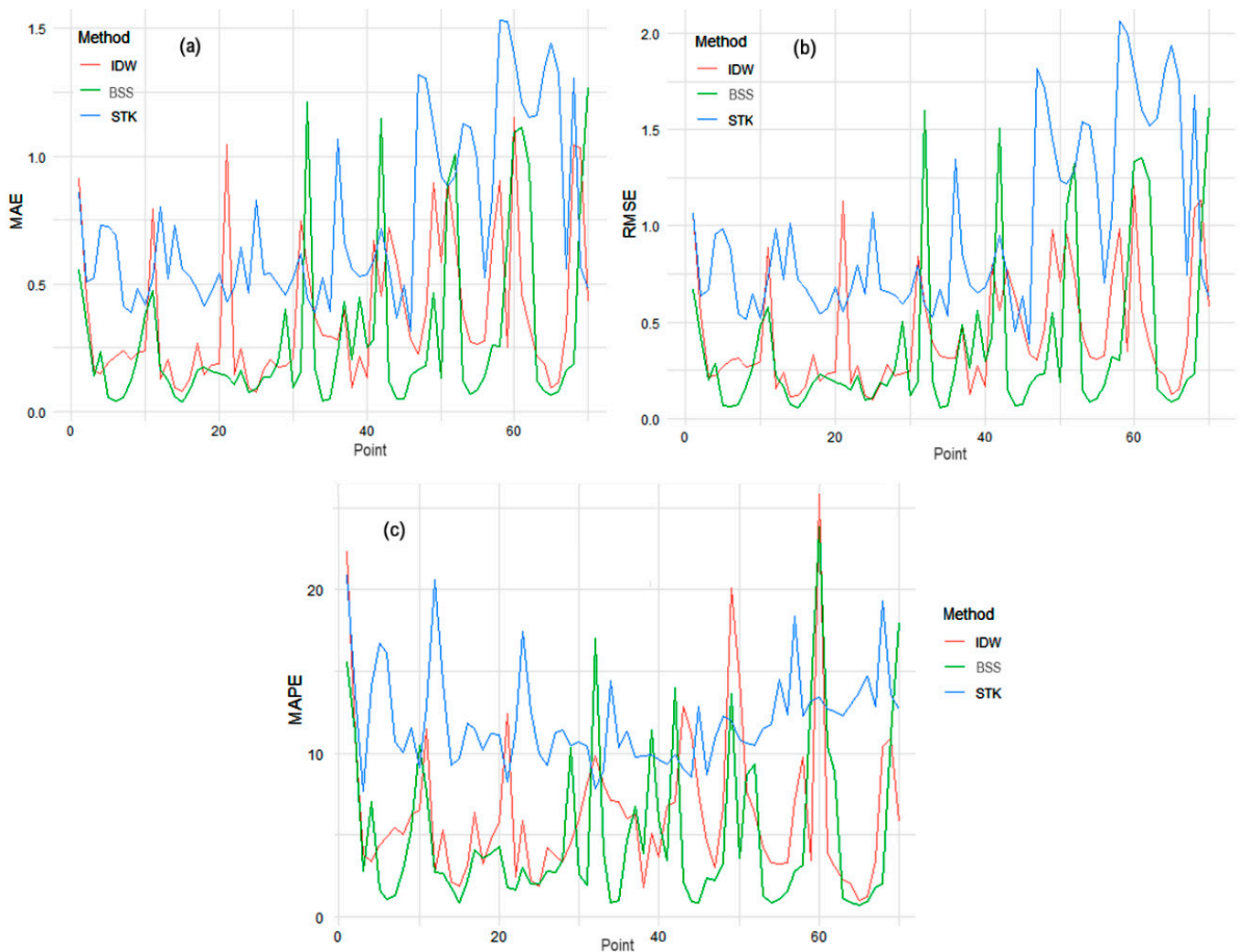
In this study, we compared three distinct interpolation methods, IDW, an adapted Spline method, and STK, to interpolate the PM<sub>2.5</sub> concentration series. Comparing the MAE, MSE, and MAPE results in the following:

- The lowest minimum and maximum MAE, RMSE were obtained using IDW.
- The lowest minimum (maximum) MAPE was obtained utilizing IDW (STK).
- The lowest average MAE, MSE, and MAPE resulted from BSS.
- BSS best fitted the series from the grid corners 1, 61, and 70, whereas the values of series from 10 were better fit by IDW.
- The series on the edges were mostly best fit (at least in terms of MAPE) by BSS.
- The most difficult to forecast were the series on the grid corners and edges due the neighbors' positions.

Using the non-parametric Friedman test [81], which is particularly suited for datasets without a normal distribution, we tested the hypothesis that there is no difference between the performance of these methods (based on MAE, RMSE, and MAPE) against the existence of some differences. The test results showed significant disparities ( $p$ -values  $< 2.2 \times 10^{-16}$  for MAE and RMSE, and  $p$ -value =  $1.085 \times 10^{-15}$  for MAPE). A Nemenyi post-hoc test [82]

further clarified these findings, indicating that our adapted spline method outperformed IDW and STK in all error metrics. Specifically, it exhibited a reduction in mean errors by approximately 21% in MAE, 15% in RMSE, and 21% in MAPE compared to IDW, and around 60% in MAE, 61% in RMSE, and 58% in MAPE compared to STK. These high-performance results of the spline method can be attributed to the specific adaptations for handling boundary and corner points, underscoring its effectiveness and precision in interpolating the PM<sub>2.5</sub> concentrations.

Figure 8 depicts the variation in the error metrics across all the 70 points.



**Figure 8.** (a) MAE, (b) RMSE, and (c) MAPE variation by method across stations.

When working with IDW, the distance from the target point to the neighbor and the degree of similarity between the data series are essential. The positive spatial autocorrelation affects the IDW's performance [83] because the similarity of the series collected in neighboring locations is more probable than that of the series situated at a higher distance from each other [84]. These findings are consistent with those from [6], which indicated that the IDW effectiveness is notable in regular grid distributions, like in this study case. Limitations were noticed at grid edges, where the absence of neighbors can alter the interpolations. This boundary effect is also a challenge in BSS.

BSS can also fail at the edge or corners of the grid (like in the initial attempt of our study) because the smoothness and continuity hypotheses on which they rely may not be satisfied at the mentioned points due to the limited number of data [58,85]. This is an important issue when the models are used for forecasting pollutant series, such as in the

present case. In this study, we overcame this limitation by creating a boundary buffer that led to the best forecast among the three considered methods.

Although STK can potentially improve prediction accuracy by integrating data across time and allowing the assessments at specific points or across entire fields during and beyond observation periods, this method must account for complex space–time variability, which is often more challenging than spatial or temporal methods alone. Dependence structures, such as variograms, vary and may require an anisotropy parameter for simplification, introducing significant assumptions about the observed processes [75]. Furthermore, the covariance function separability assumption is rarely valid for real-world processes. It implies that the temporal pattern observed at a specific location is independent and not directly influenced by the temporal patterns at other locations [71]. For this reason, one should apply more complex covariogram function formulations such as sums and products.

Figures 9–11 contain the plots of the NSE, KGE, and dIndex for IDW, BSS, and STK. In our spatial analysis, the station-specific performance of interpolation methods is evident from the NSE, KGE, and dIndex.

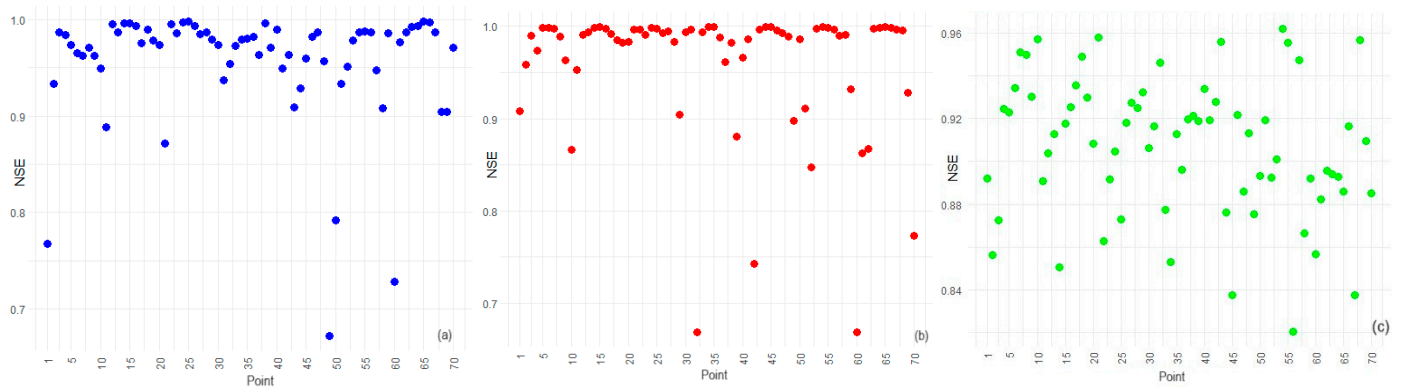


Figure 9. NSE for (a) IDW, (b) BSS, and (c) STK.

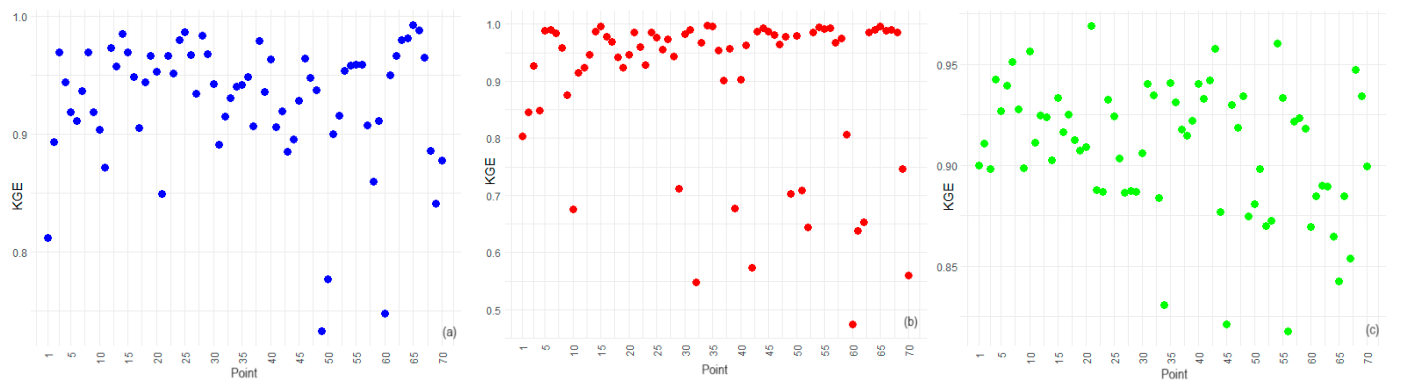
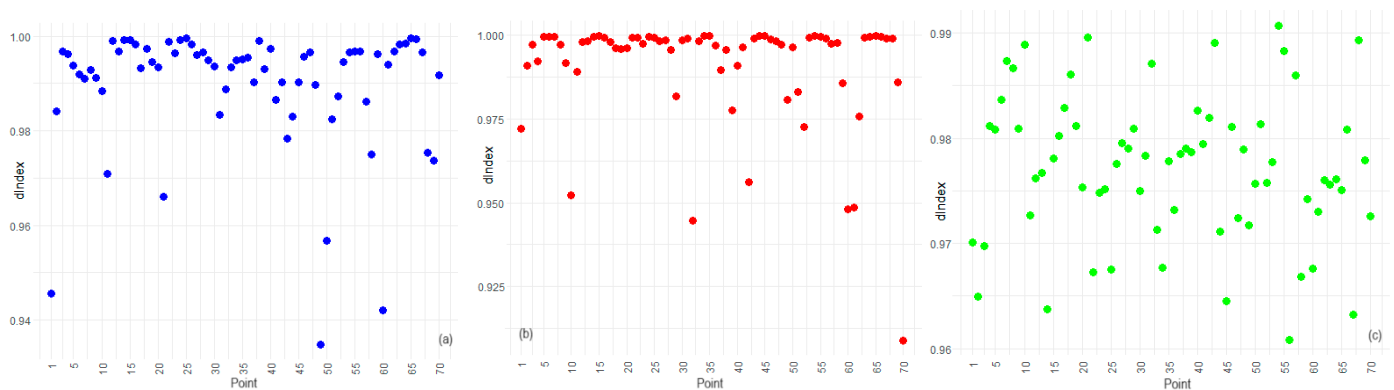


Figure 10. KGE for (a) IDW, (b) BSS, and (c) STK.

Edge and corner stations, such as points 1 and 70 at the periphery and points 10 and 11 at coastal boundaries, are particularly informative in understanding the spatial limitations of these methods.

The BSS method’s high NSE and dIndex values across the board, including edge stations such as point 1 and coastal corners such as point 10, demonstrate its robustness and spatial consistency. This result aligns with findings from the Friedman and Nemenyi tests, which identified BSS as the superior method in terms of overall error metrics. The method’s strength is not compromised by the spatial position of the stations, indicating its suitability for varied terrain and its resilience to edge effects.



**Figure 11.** dIndex for (a) IDW, (b) BSS, and (c) STK.

For IDW, lower KGE values at points on the periphery suggest that the spatial distribution of points somewhat influences its performance. Since IDW relies on the proximity and weighting of nearby points, stations with fewer neighbors, such as point 70, may experience reduced accuracy, particularly in capturing variability and bias. These findings highlight the method's spatial sensitivity, necessitating careful application or potential methodological adjustments in areas where data points are sparse or unevenly distributed.

STK's KGE and dIndex results, particularly at edge stations such as point 21 and corner locations such as point 11, imply that this method's assumptions may not uniformly accommodate the unique spatial-temporal dynamics present at these stations. The broader spread in efficiency scores for STK across different stations underscores the method's nuanced application, warranting a closer look at the spatial-temporal structuring of the data in these areas.

Moran's I [86] and Geary's C [87] tests, which showed no significant spatial autocorrelation in residuals, reinforce the interpretation that the observed variations in efficiency and agreement are not due to spatial patterning in the data but rather reflect the interaction between the methods' inherent characteristics and the environmental context of each station.

While BSS demonstrates versatility and reliability across the spatial domain, IDW and STK may require station-specific considerations, particularly at spatially constrained edge and corner locations. Kriging or other geostatistical methods could be applied to estimate the values of the boundary points by modeling the spatial autocorrelation of the data. This alternative could provide another way to create the boundary buffer, potentially offering improvements in the interpolation accuracy and the robustness of the overall spatial model. This holistic assessment, integrating both statistical performance and spatial considerations, ensures a well-informed approach to PM<sub>2.5</sub> variable modeling, addressing the intricacies of the spatial and temporal variability inherent in PM<sub>2.5</sub> data.

The methods proposed here can be applied for spatial interpolation of any spatially distributed data series.

The literature search returned studies providing mostly IDW applications in different fields and comparisons with kriging and other methods but not STK and BSS. Li and Heap [58] compared 72 methods provided by 53 articles. They found that IDW, ordinary kriging (OK), and ordinary co-kriging (OCK) are the most utilized approaches. IDW outperformed some methods in a few cases [88,89], has similar performance [90,91], or is worse than OK in other cases [63,92]. For example, IDW, the principal component regression with the residual correction (PCRR), and the multiple linear regression (MLR) were employed to compare interpolated precipitation at different scales and the spatial distribution of rainfall in a catchment in China [93]. The results showed that the PCRR method was the best. IDW, OK, OCK Nearest Neighbor (NN), Trend Surfaces and Regression (TSR), TSA-OK, and Bayesian Model Averaging (BMA) were compared in the study [94], showing that TSA-OK and BMA performed best.

In its study, Dubrule [95] indicated that splines and kriging should be used alternately, depending on the purposes, but from the accuracy viewpoint, kriging is better. A comparison of spatio-temporal interpolation of geographic data using reduction and extension methods with spatio-temporal interpolation methods based on IDW and OK indicated the worst results among the last two methods [96].

Different approaches have been used for the spatio-temporal interpolation, but only a few applied BSS and STK, without providing comparisons [97–105]. The least studied is BSS interpolation [85,106]. Therefore, deeper investigations must be carried out on various data series to assess the performances of BSS and STK.

#### 4. Conclusions

This study concludes that the BSS model stands out among the three models evaluated for the PM<sub>2.5</sub> concentration in the Arabian Gulf due to its unique approach to boundary conditions. As a first contribution novelty of the study, implementing buffer points significantly reduces error metrics, establishing the BSS method's superiority in accuracy over the IDW and STK methods. Therefore, it is recommended for forecasting the PM<sub>2.5</sub> in locations where records are unavailable.

While the IDW model, in its basic form with an optimized power parameter  $\beta$  for the complete temporal records, is promising, it could benefit from enhancements in boundary effect handling. Such improvements could potentially elevate its accuracy, especially at grid edges. This approach could involve integrating strategies similar to those used in the BSS model, such as introducing buffer points or advanced weighting algorithms near boundaries.

STK, known for its robust integration of spatial and temporal data, suggests potential for refinement in future research. Exploring different covariogram functions and optimizing variogram parameters could lead to a more accurate representation of the complex spatio-temporal characteristics of the PM<sub>2.5</sub> series. Such advancements would not only enhance the model's accuracy but also its applicability in environmental monitoring. These findings underscore the efficacy of each method in spatio-temporal interpolation and the potential for further optimization, contributing to our understanding of PM<sub>2.5</sub> distribution and serving as a model for other regions facing similar environmental challenges.

Apart from the originality of using a GAMM-STK approach for non-normally distributed series and proposing the buffering for BSS, after a deep literature research, this study was evidenced as unique by comparing the proposed approaches on spatial-temporally distributed data series. Moreover, the methods can be used for any data series and region. Future research will explore other optimizations and their implications for spatio-temporal data analysis in various environmental contexts.

The article fills in a gap related to the PM<sub>2.5</sub> distribution in the UAE, providing results that can be the background for studies on the effects of pollution on the population. These results can also be used by policymakers interested in reducing PM<sub>2.5</sub> pollution in the UAE.

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

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data are freely available at: [https://disc.gsfc.nasa.gov/datasets/M2T1\\_NXLND\\_5.12.4/summary](https://disc.gsfc.nasa.gov/datasets/M2T1_NXLND_5.12.4/summary) (accessed on 10 May 2022).

**Conflicts of Interest:** The present research work carries no conflicts of interest.

## Appendix A

**Table A1.** Coordinates of the grid points.

ID	X	Y	ID	X	Y	ID	X	Y	ID	X	Y	ID	X	Y
1	51.2500	26.3125	15	53.7500	25.6875	29	56.2500	25.0625	43	52.5000	23.8125	57	55.0000	23.1875
2	51.8750	26.3125	16	54.3750	25.6875	30	56.8750	25.0625	44	53.1250	23.8125	58	55.6250	23.1875
3	52.5000	26.3125	17	55.0000	25.6875	31	51.2500	24.4375	45	53.7500	23.8125	59	56.2500	23.1875
4	53.1250	26.3125	18	55.6250	25.6875	32	51.8750	24.4375	46	54.3750	23.8125	60	56.8750	23.1875
5	53.7500	26.3125	19	56.2500	25.6875	33	52.5000	24.4375	47	55.0000	23.8125	61	51.2500	22.5625
6	54.3750	26.3125	20	56.8750	25.6875	34	53.1250	24.4375	48	55.6250	23.8125	62	51.8750	22.5625
7	55.0000	26.3125	21	51.2500	25.0625	35	53.7500	24.4375	49	56.2500	23.8125	63	52.5000	22.5625
8	55.6250	26.3125	22	51.8750	25.0625	36	54.3750	24.4375	50	56.8750	23.8125	64	53.1250	22.5625
9	56.2500	26.3125	23	52.5000	25.0625	37	55.0000	24.4375	51	51.2500	23.1875	65	53.7500	22.5625
10	56.8750	26.3125	24	53.1250	25.0625	38	55.6250	24.4375	52	51.8750	23.1875	66	54.3750	22.5625
11	51.2500	25.6875	25	53.7500	25.0625	39	56.2500	24.4375	53	52.5000	23.1875	67	55.0000	22.5625
12	51.8750	25.6875	26	54.3750	25.0625	40	56.8750	24.4375	54	53.1250	23.1875	68	55.6250	22.5625
13	52.5000	25.6875	27	55.0000	25.0625	41	51.2500	23.8125	55	53.7500	23.1875	69	56.2500	22.5625
14	53.1250	25.6875	28	55.6250	25.0625	42	51.8750	23.8125	56	54.3750	23.1875	70	56.8750	22.5625

## References

- Inhalable Particulate Matter and Health (PM<sub>2.5</sub> and PM<sub>10</sub>). Available online: <https://ww2.arb.ca.gov/resources/inhalable-particulate-matter-and-health> (accessed on 15 January 2024).
- Ryou, H.G.; Heo, J.; Kim, S.Y. Source apportionment of PM<sub>10</sub> and PM<sub>2.5</sub> air pollution, and possible impacts of study characteristics in South Korea. *Environ. Pollut.* **2018**, *240*, 963–972. [[CrossRef](#)]
- Dumitru, A.; Olaru, E.-A.; Dumitru, M.; Iorga, G. Assessment of air pollution by aerosols over a coal open-mine influenced region in southwestern Romania. *Rom. J. Phys.* **2024**, *69*, 801.
- Chiritescu, R.-V.; Luca, E.; Iorga, G. Observational study of major air pollutants over urban Romania in 2020 in comparison with 2019. *Rom. Rep. Phys.* **2024**, *76*, 702.
- Schlesinger, R.B. The health impact of common inorganic components of fine particulate matter (PM<sub>2.5</sub>) in ambient air: A critical review. *Inhal. Toxicol.* **2007**, *19*, 811–832. [[CrossRef](#)] [[PubMed](#)]
- Arias-Pérez, R.D.; Taborda, N.A.; Gómez, D.M.; Narvaez, J.F.; Porras, J.; Hernandez, J.C. Inflammatory effects of particulate matter air pollution. *Environ. Sci. Pollut. Res.* **2020**, *27*, 42390–42404. [[CrossRef](#)] [[PubMed](#)]
- Estimate of Premature Deaths Associated with Fine Particle Pollution (PM<sub>2.5</sub>) in California Using a U.S. Environmental Protection Agency Methodology. Available online: [https://archive.epa.gov/region9/mediacenter/web/pdf/pm-report\\_2010.pdf](https://archive.epa.gov/region9/mediacenter/web/pdf/pm-report_2010.pdf) (accessed on 12 January 2024).
- Thangavel, P.; Park, D.; LeE, Y.C. Recent Insights into Particulate Matter (PM<sub>2.5</sub>)-Mediated Toxicity in Humans: An Overview. *Int. J. Environ. Res. Public Health* **2022**, *19*, 7511. [[CrossRef](#)] [[PubMed](#)]
- You Can Smell Petrol in the Air. 2023. Available online: <https://www.hrw.org/report/2023/12/04/you-can-smell-petrol-air/uae-fossil-fuels-feed-toxic-pollution#:~:text=The%20UAE%20has%20dangerously%20high,considers%20safe%20for%20human%20health> (accessed on 15 January 2024).
- Bărbulescu, A.; Nazzal, Y.; Howari, F. Statistical analysis and estimation of the regional trend of aerosol size over the Arabian Gulf Region during 2002–2016. *Sci. Rep.* **2018**, *8*, 9571. [[CrossRef](#)] [[PubMed](#)]
- Nazzal, Y.; Bărbulescu, A.; Howari, F.M.; Yousef, A.; Al-Taani, A.A.; Al Aydaroo, F.; Naseem, M. New insight to dust storm from historical records, UAE. *Arab. J. Geosci.* **2019**, *12*, 396. [[CrossRef](#)]
- Nazzal, Y.; Bărbulescu, A. Statistical analysis of the dust storms in the United Arab Emirates. *Atmos. Resear.* **2020**, *231*, 104669.
- Nazzal, Y.; Bou Orm, N.; Bărbulescu, A.; Howari, F.; Sharma, M.; Badawi, A.; Al-Taani, A.A.; Iqbal, J.; El Ktaibi, F.; Xavier, C.M.; et al. Study of atmospheric pollution and health risk assessment. A case study for the Sharjah and Ajman Emirates (UAE). *Atmosphere* **2021**, *12*, 1442. [[CrossRef](#)]
- How Bad Is Our Air Pollution—And How Do We Tackle It? Available online: <https://www.thenationalnews.com/uae/environment/2022/09/20/explained-how-much-of-a-problem-is-air-pollution-in-the-uae/> (accessed on 15 January 2024).
- Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019 (GBD 2019) Air Pollution Exposure Estimates 1990–2019. Data Resources. 2021. Available online: <https://data.worldbank.org/indicator/EN.ATM.PM25.MC.M3?end=2017&locations=AE&start=2017&view=map> (accessed on 15 January 2024).
- Usmani, R.S.A.; Saeed, A.; Abdullahi, A.M.; Pillai, T.R.; Jhanjhi, N.Z.; Hashem, I.A.T. Air pollution and its health impacts in Malaysia: A review. *Air Qual. Atmos. Health* **2020**, *13*, 1093–1118. [[CrossRef](#)]
- Calotă, R.; Antonescu, N.N.; Stănescu, D.-P.; Năstase, I. The Direct Effect of Enriching the Gaseous Combustible with 23% Hydrogen in Condensing Boilers' Operation. *Energies* **2022**, *15*, 93733. [[CrossRef](#)]
- Antonescu, N.N.; Stănescu, D.-P.; Calotă, R. CO<sub>2</sub> Emissions Reduction through Increasing H<sub>2</sub> Participation in Gaseous Combustible—Condensing Boilers Functional Response. *Appl. Sci.* **2022**, *12*, 3831. [[CrossRef](#)]
- Li, L.; Losser, T.; Yorke, C.; Piltner, R. Fast inverse distance weighting-based spatio-temporal interpolation: A web-based application of interpolating daily fine particulate matter PM<sub>2.5</sub> in the contiguous U.S. using parallel programming and k-d tree. *Int. J. Environ. Res. Public Health* **2014**, *11*, 9101–9141. [[CrossRef](#)]

20. Choi, K.; Chong, K. Modified Inverse Distance Weighting Interpolation for Particulate Matter Estimation and Mapping. *Atmosphere* **2022**, *13*, 846. [CrossRef]
21. Deng, L. Estimation of PM<sub>2.5</sub> Spatial Distribution Based on Kriging Interpolation. In Proceedings of the First International Conference on Information Sciences, Machinery, Materials and Energy, Chongqing, China, 11–13 April 2015; pp. 1791–1794.
22. Lee, S.J.; Serre, M.L.; van Donkelaar, A.; Martin, R.V.; Burnett, R.T.; Jerrett, M. Comparison of geostatistical interpolation and remote sensing techniques for estimating long-term exposure to ambient PM<sub>2.5</sub> concentrations across the continental United States. *Environ. Health Perspect.* **2012**, *120*, 1727–1732. [CrossRef]
23. Chae, S.; Shin, J.; Kwon, S. PM<sub>10</sub> and PM<sub>2.5</sub> real-time prediction models using an interpolated convolutional neural network. *Sci. Rep.* **2021**, *11*, 11952. [CrossRef]
24. Goudarzi, G.; Hopke, P.H.; Yazdani, M. Forecasting PM<sub>2.5</sub> concentration using artificial neural network and its health effects in Ahvaz, Iran. *Chemosphere* **2021**, *283*, 131285. [CrossRef]
25. Xiao, F.; Yang, M.; Fan, H.; Fan, G. An improved deep learning model for predicting daily PM<sub>2.5</sub> concentration. *Sci. Rep.* **2020**, *10*, 20988. [CrossRef]
26. Ma, J.; Ding, Y.; Cheng, J.C.P.; Jiang, F.; Wan, Z. A temporal-spatial interpolation and extrapolation method based on geographic Long Short-Term Memory neural network for PM<sub>2.5</sub>. *J. Clean. Prod.* **2019**, *237*, 117729. [CrossRef]
27. Yanosky, J.D.; Paciorek, C.J.; Suh, H.H. Predicting chronic fine and coarse particulate exposures using spatio-temporal models for the Northeastern and Midwestern United States. *Environ. Health Perspect.* **2009**, *117*, 522–529. [CrossRef] [PubMed]
28. Oshan, T.M.; Li, Z.; Kang, W.; Wolf, L.J.; Fotheringham, A.S. Mgw: A Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 269. [CrossRef]
29. Wei, P.; Xie, S.; Huang, L.; Liu, L.; Tang, Y.; Zhang, Y.; Wu, H.; Xue, Z.; Ren, D. Spatial interpolation of PM<sub>2.5</sub> concentrations during holidays in south-central China considering multiple factors. *Sci. Total Environ.* **2020**, *740*, 139761. [CrossRef]
30. Shao, Y.; Ma, Z.; Wang, J.; Bi, J. Estimating daily ground-level PM<sub>2.5</sub> in China with random-forest-based spatio-temporal kriging. *Sci. Total Environ.* **2020**, *740*, 13761. [CrossRef] [PubMed]
31. Sampson, P.D.; Szpiro, A.A.; Sheppard, L.; Lindström, J.; Kaufman, J.D. Pragmatic estimation of a spatio-temporal air quality model with irregular monitoring data. *Atmos. Environ.* **2011**, *45*, 6593–6606. [CrossRef]
32. Liu, H.; Chen, C. Prediction of outdoor PM<sub>2.5</sub> concentrations based on a three-stage hybrid neural network model. *Atmos. Pollut. Res.* **2020**, *11*, 469–481. [CrossRef]
33. Valencia, A.; Serre, M.; Arunachalam, S. A hyperlocal hybrid data fusion near-road PM<sub>2.5</sub> and NO<sub>2</sub> annual risk and environmental justice assessment across the United States. *PLoS ONE* **2023**, *18*, e0286406. [CrossRef]
34. Tang, Y.; Chai, T.; Pan, L.; Lee, P.; Tong, D.; Kim, H.-C.; Chen, W. Using optimal interpolation to assimilate surface measurements and satellite AOD for ozone and PM<sub>2.5</sub>: A case study for July 2011. *J. Air Waste Manag. Assoc.* **2015**, *65*, 1206–1216. [CrossRef]
35. Mahajan, S.; Chen, L.-J.; Tsai, T.-C. Short-Term PM<sub>2.5</sub> Forecasting Using Exponential Smoothing Method: A Comparative Analysis. *Sensors* **2018**, *18*, 3223. [CrossRef]
36. Abuelgasim, A.; Farahat, A. Investigations on PM<sub>10</sub>, PM<sub>2.5</sub>, and Their Ratio over the Emirate of Abu Dhabi, United Arab Emirates. *Earth Environ.* **2020**, *4*, 3. [CrossRef]
37. Al-Taani, A.A.; Nazzal, Y.; Howari, F.M.; Yousef, A. Long-term trends in ambient fine particulate matter from 1980 to 2016 in United Arab Emirates. *Environ. Monit. Assess.* **2019**, *191*, 143. [CrossRef]
38. GMAO (Global Modeling and Assimilation Office). Available online: <https://gmao.gsfc.nasa.gov/reanalysis/MERRA-2/docs/> (accessed on 19 February 2024).
39. Gelaro, R.; McCarty, W.; Suárez, M.J.; Todling, R.; Molod, A.; Takacs, L.; Randles, C.A.; Darmenov, A.; Bosilovich, M.G.; Reichle, R.; et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *J. Clim.* **2017**, *30*, 5419–5454. [CrossRef]
40. Randles, C.A.; Da Silva, A.M.; Buchard, V.; Colarco, P.R.; Darmenov, A.; Govindaraju, R.; Smirnov, A.; Holben, B.; Ferrare, R.; Hair, J.; et al. The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System Description and Data Assimilation Evaluation. *J. Clim.* **2017**, *30*, 6823–6850. [CrossRef] [PubMed]
41. Chin, M.; Ginoux, P.; Kinne, S.; Torres, O.; Holben, B.; Duncan, B.; Martin, R.; Logan, J.; Higurashi, A.; Nakajima, T. Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and sun photometer measurements. *J. Atmos. Sci.* **2002**, *59*, 461–483. [CrossRef]
42. Colarco, P.; daSilva, A.; Chin, M.; Diehl, T. Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth. *J. Geophys. Res.* **2010**, *115*, D14207. [CrossRef]
43. Molod, A.M.; Takacs, L.L.; Suarez, M.J.; Bacmeister, J. Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA2. *Geosci. Model Develop.* **2015**, *8*, 1339–1356. [CrossRef]
44. Rienecker, M.M.; Suarez, M.J.; Gelaro, R.; Todling, R.; Bacmeister, J.; Liu, E.; Bosilovich, M.G.; Schubert, S.D.; Takacs, L.; Kim, G.-K.; et al. MERRA: NASA’s Modern-Era Retrospective Analysis for Research and Applications. *J. Clim.* **2011**, *24*, 3624–3648. [CrossRef]
45. Buchard, V.; da Silva, A.M.; Randles, C.A.; Colarco, P.; Ferrare, R.; Hair, J.; Hostetler, C.; Tackett, J.; Winker, D. Evaluation of the surface PM<sub>2.5</sub> in Version 1 of the NASA MERRA Aerosol Reanalysis over the United States. *Atmos. Environ.* **2016**, *125*, 100–111. [CrossRef]

46. He, L.; Lin, A.; Chen, X.; Zhou, H.; Zhou, Z.; He, P. Assessment of MERRA-2 Surface PM<sub>2.5</sub> over the Yangtze River Basin: Ground-based Verification, Spatio-temporal Distribution and Meteorological Dependence. *Remote Sens.* **2019**, *11*, 460. [CrossRef]
47. Liu, C.; Chung, C.E.; Yin, Y.; Schnaiter, M. The Absorption Ångström Exponent of Black Carbon: From Numerical Aspects. *Atmos. Chem. Phys.* **2018**, *18*, 6259–6273. [CrossRef]
48. Liu, S.; Geng, G.; Xiao, Q.; Zheng, Y.; Liu, X.; Cheng, J.; Zhang, Q. Tracking Daily Concentrations of PM<sub>2.5</sub> Chemical Composition in China since 2000. *Environ. Sci. Technol.* **2022**, *56*, 16517–16527. [CrossRef]
49. Liu, Y.; Koutrakis, P.; Kahn, R.; Turquety, S.; Yantosca, R.M. Estimating Fine Particulate Matter Component Concentrations and Size Distributions Using Satellite-Retrieved Fractional Aerosol Optical Depth: Part 2—A Case Study. *J. Air Waste Manag. Assoc.* **2007**, *57*, 1360–1369. [CrossRef]
50. Meng, X.; Garay, M.J.; Diner, D.J.; Kalashnikova, O.V.; Xu, J.; Liu, Y. Estimating PM<sub>2.5</sub> Speciation Concentrations Using Prototype 4.4 km-Resolution MISR Aerosol Properties over Southern California. *Atmos. Environ.* **2018**, *181*, 70–81. [CrossRef] [PubMed]
51. Meng, X.; Hand, J.L.; Schichtel, B.A.; Liu, Y. Space-Time Trends of PM<sub>2.5</sub> Constituents in the Conterminous United States Estimated by a Machine Learning Approach, 2005–2015. *Environ. Int.* **2018**, *121*, 1137–1147. [CrossRef]
52. Wamsley, P.R.; Weimer, C.S.; Applegate, J.T.; Hunt, B. CALIPSO: Polarization Performance of a Space-Based, Backscatter LIDAR. In Proceedings of the Frontiers in Optics 2007/Laser Science XXIII/Organic Materials and Devices for Displays and Energy Conversion, San Jose, CA, USA, 16–20 September 2007; OSA Technical Digest (CD) paper LTuK4. Optica Publishing Group: Washington, DC, USA, 2007.
53. MERRA-2 tavgM\_2d\_aer\_Nx: 2d, Monthly Mean, Time-Averaged, Single-Level, Assimilation, Aerosol Diagnostics V5.12.4 (M2TMNXAER). Available online: [https://disc.gsfc.nasa.gov/datasets/M2TMNXAER\\_5.12.4/summary#citation](https://disc.gsfc.nasa.gov/datasets/M2TMNXAER_5.12.4/summary#citation) (accessed on 10 November 2023).
54. Kolmogorov, A. Sulla determinazione empirica di una legge di distribuzione. *G. Ist. Ital. Attuari* **1933**, *4*, 83–91.
55. Smirnov, N. Table for estimating the goodness of fit of empirical distributions. *Ann. Math. Stat.* **1948**, *19*, 279–281. [CrossRef]
56. Kruskal, W.; Wallis, A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583–621. [CrossRef]
57. Ly, S.; Charles, C.; Degre, A. Different methods for spatial interpolation of rainfall data for operational hydrology and hydrological modeling at watershed scale: A review. *Biotechnol. Agron. Soc. Environ.* **2013**, *17*, 67–82.
58. Li, J.; Heap, A.D. Spatial interpolation methods applied in the environmental sciences: A review. *Environ. Modell. Softw.* **2014**, *53*, 173–189. [CrossRef]
59. Saffari, M.; Yasrebi, J.; Fathi, H.; Karimian, N.; Moazallahi, M.; Gazni, R. Evaluation and comparison of ordinary kriging and inverse distance weighting methods for prediction of spatial variability of some soil chemical parameters. *Res. J. Biol. Sci.* **2009**, *4*, 93–102.
60. Chang, C.L.; Lo, S.L.; Yu, S.L. The parameter optimization in the inverse distance method by genetic algorithm for estimating precipitation. *Environ. Monit. Assess.* **2006**, *117*, 145–155. [CrossRef]
61. Golkhatmi, N.S.; Sanaeinejad, S.H.; Ghahraman, B.; Pazhand, H.R. Extended modified inverse distance method for interpolation rainfall. *Int. J. Eng. Invent.* **2012**, *1*, 57–65.
62. Gholipour, Y.; Shahbazi, M.M.; Behnia, A. An improved version of Inverse Distance Weighting metamodel assisted Harmony Search algorithm for truss design optimization. *Latin Am. J. Solids Struct.* **2013**, *10*, 283–300. [CrossRef]
63. Bărbulescu, A.; Băutu, A.; Băutu, E. Optimizing Inverse Distance Weighting with Particle Swarm Optimization. *Appl. Sci.* **2020**, *10*, 2054. [CrossRef]
64. Webster, R.; Oliver, M.A. *Geostatistics for Environmental Scientists*; Wiley: Chichester, NH, USA, 2007.
65. Diodato, N.; Ceccarelli, M. Processes using multivariate geostatistics for mapping interpolation of climatological precipitation mean in the Sannio Mountains (Southern Italy). *Earth Surf. Proc. Landf.* **2005**, *30*, 259–268. [CrossRef]
66. Delfiner, P. Linear estimation of non-stationary spatial phenomena. In *Advanced Geostatistics in the Mining Industry*; Guarascio, M., David, M., Huijbregts, C., Eds.; Springer: Dordrecht, The Netherlands, 1976; pp. 49–68.
67. Mueller, T.; Pusuluri, N.; Mathias, K.; Cornelius, P.; Barnhisel, R.; Shearer, S. Map Quality for Ordinary Kriging and Inverse Distance Weighted Interpolation. *Soil Sci. Soc. Am. J.* **2004**, *68*, 2042–2047. [CrossRef]
68. Paramasivam, C.; Venkatramanan, S. Chapter 3—An Introduction to Various Spatial Analysis Techniques. In *GIS and Geostatistical Techniques for Groundwater Science*; Venkatramanan, S., Prasanna, M.V., Chung, S.Y., Eds.; Elsevier: Amsterdam, The Netherlands, 2019; pp. 23–30.
69. Kamada, M.; Enkhbat, R. Spline Interpolation in Piecewise Constant Tension. In Proceedings of the SAMPTA'09, Marseille, France, 18–22 May 2009. pp. poster session, fhal-00453546f.
70. Package 'Akima'. 2022. Available online: <https://cran.r-project.org/web/packages/akima/akima.pdf> (accessed on 20 December 2023).
71. Wikle, C.K.; Zammit-Mangion, A.; Cressie, N.A.C. *Spatio-Temporal Statistics with R*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2019.
72. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006.
73. Duvenaud, D. Automatic Model Construction with Gaussian Processes. Ph.D. Thesis, University of Cambridge, Cambridge, UK, 2014.

74. Graeler, B.; Pebesma, E. Spatio-Temporal Interpolation using gstat. *R J.* **2016**, *8*, 204–218. [[CrossRef](#)]
75. Biondi, F. Space-time kriging extension of precipitation variability at 12 km spacing from tree-ring chronologies and its implications for drought analysis. *Hydrol. Earth Syst. Sci. Discuss.* **2013**, *10*, 4301–4335.
76. Wood, S.N. *Generalized Additive Models. An Introduction with R*, 2nd ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2017.
77. Tegen, I.; Lacis, A.A. Modeling of particle size distribution and its influence on the radiative properties of mineral dust aerosol. *J. Geophys. Res.* **1996**, *101*, 19237–19244. [[CrossRef](#)]
78. Astitha, M.; Kallos, G.; Spyrou, C.; O’Hirok, W.; Lelieveld, J.; Denier van der Gon, H.A.C. Modelling the chemically aged and mixed aerosols over the eastern central Atlantic Ocean—Potential impacts. *Atmos. Chem. Phys.* **2010**, *10*, 5797–5822. [[CrossRef](#)]
79. Spyrou, C.; Kallos, G.; Mitsakou, C.; Athanasiadis, P.; Kalogeri, C.; Iacono, M.J. Modeling the radiative effects of desert dust on weather and regional climate. *Atmos. Chem. Phys.* **2013**, *13*, 5489–5504. [[CrossRef](#)]
80. Bărbulescu, A.; Dumitriu, C.S.; Popescu-Bodorin, N. On the aerosol optical depth series in the Arabian Gulf region. *Rom. J. Phys.* **2022**, *67*, 814.
81. Friedman, M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Am. Stat. Assoc.* **1938**, *32*, 675–701. [[CrossRef](#)]
82. Hollander, M.; Wolfe, D.A.; Chicken, E. *Nonparametric Statistical Methods*; John Wiley & Sons: Hoboken, NJ, USA, 2013.
83. Vasiliev, I. *Visualization of Spatial Dependence: An Elementary View of Spatial Autocorrelation*; CRC Press: Boca Raton, FL, USA, 1996.
84. Tobler, W. A computer movie simulating urban growth in the detroit region. *Econ. Geogr.* **1970**, *46*, 234–240. [[CrossRef](#)]
85. Agrapart, Q.; Batailly, A. Cubic and Bicubic Spline Interpolation in Python. École Polytechnique de Montréal. 2020. Available online: <http://hal.science/hal-03017566/document> (accessed on 20 December 2023).
86. Moran, P.A.P. Notes on Continuous Stochastic Phenomena. *Biometrika* **1950**, *37*, 17–23. [[CrossRef](#)] [[PubMed](#)]
87. Geary, R.C. The Contiguity Ratio and Statistical Mapping. *Incorp. Stat.* **1954**, *5*, 115–145. [[CrossRef](#)]
88. Bărbulescu, A. A new method for estimation the regional precipitation. *Water Resour. Manag.* **2016**, *30*, 33–42. [[CrossRef](#)]
89. Susanto, F.; de Souza, P.; He, J. Spatio-temporal Interpolation for Environmental Modelling. *Sensors* **2016**, *16*, 1245. [[CrossRef](#)]
90. Schloeder, C.A.; Zimmerman, N.E.; Jacobs, M.J. Comparison of methods for interpolating soil properties using limited data. *Soil Sci. Soc. Am. J.* **2001**, *65*, 470–479. [[CrossRef](#)]
91. Ruddick, R. *Data Interpolation Methods in the Geoscience Australia Seascapes Maps*; Geoscience Australia: Canberra, Australia, 2007.
92. Gotway, C.A.; Ferguson, R.B.; Hergert, G.W.; Peterson, T.A. Comparison of kriging and inverse-distance methods for mapping parameters. *Soil Sci. Soc. Am. J.* **1996**, *60*, 1237–1247. [[CrossRef](#)]
93. Chen, T.; Ren, L.; Yuan, F.; Yang, X.; Jiang, S.; Tang, T.; Liu, Y.; Zhao, C.; Zhang, L. Comparison of Spatial Interpolation Schemes for Rainfall Data and Application in Hydrological Modeling. *Water* **2017**, *9*, 342. [[CrossRef](#)]
94. Liu, D.; Zhao, Q.; Fu, D.; Guo, S.; Liu, P.; Zeng, Y. Comparison of spatial interpolation methods for the estimation of precipitation patterns at different time scales to improve the accuracy of discharge simulations. *Hydrol. Res.* **2020**, *5*, 583–601. [[CrossRef](#)]
95. Dubrule, O. Comparing splines and kriging. *Comp. Geosci.* **1984**, *10*, 327–338. [[CrossRef](#)]
96. Li, L.; Revesz, P. Interpolation methods for spatio-temporal geographic data. *Comput. Environ. Urban Syst.* **2004**, *28*, 201–227. [[CrossRef](#)]
97. Graeler, B.; Gerharz, L.; Pebesma, E. *Spatio-Temporal Analysis and Interpolation of PM<sub>10</sub> Measurements in Europe*; ETC/ACM Technical Paper 2011/10; The European Topic Centre on Air Pollution and Climate Change Mitigation: Bilthoven, The Netherlands, 2011.
98. Graeler, B.; Rehr, M.; Gerharz, L.; Pebesma, E. *Spatio-Temporal Analysis and Interpolation of PM<sub>10</sub> Measurements in Europe for 2009*. Technical Paper 2012/8; The European Topic Centre on Air Pollution and Climate Change Mitigation: Bilthoven, The Netherlands, 2013.
99. Ikechukwu, M.; Ebinne, E.; Idorenyin, U.; Raphael, N. Accuracy Assessment and Comparative Analysis of IDW, Spline and Kriging in Spatial Interpolation of Landform (Topography): An Experimental Study. *J. Geogr. Inform. Syst.* **2017**, *9*, 354–371. [[CrossRef](#)]
100. Shope, C.L.; Ram, M.G. Modeling Spatio-temporal Precipitation: Effects of Density, Interpolation, and Land Use Distribution. *Adv. Meteorol.* **2015**, *2015*, 174196. [[CrossRef](#)]
101. Xiao, H.; Zhans, Z.; Chen, L.; He, Q. An Improved Spatio-Temporal Kriging Interpolation Algorithm and Its Application in Slope. *IEEE Access* **2020**, *8*, 90718–90729. [[CrossRef](#)]
102. Akima, H. Algorithm 761: Scattered-data surface fitting that has the accuracy of a cubic polynomial. *ACM Trans. Math. Softw.* **1996**, *22*, 362–371. [[CrossRef](#)]
103. Akima, H. Rectangular-Grid-Data Surface Fitting that Has the Accuracy of a Bicubic Polynomial. *J. ACM* **1996**, *22*, 357–361. [[CrossRef](#)]
104. Tsai, D.-R.; Jhuang, J.-R.; Su, S.-Y.; Chiang, C.-J.; Yang, Y.-W.; Lee, W.-C. A stabilized spatio-temporal kriging method for disease mapping and application to male oral cancer and female breast cancer in Taiwan. *BMC Med. Res. Methodol.* **2022**, *22*, 270. [[CrossRef](#)]

105. Tan, Q.; Xu, X. Comparative Analysis of Spatial Interpolation Methods: An Experimental Study. *Sens. Transducers* **2014**, *165*, 155–163.
106. Abdullah, D.; Fajriana, F.; Maryana, M.; Rosnita, L.; Siahaan, A.P.U.; Rahim, R.; Harliana, P.; Harmayani, H.; Ginting, Z.; Erliana, C.I.; et al. Application of Interpolation Image by using BiCubic Algorithm. *J. Phys. Conf. Ser.* **2018**, *1114*, 012066. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.