Ecological Indicators 154 (2023) 110551

Contents lists available at ScienceDirect

Ecological Indicators

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

New indices to quantify patterns of relative errors produced by spatial interpolation models – A comparative study by modelling soil properties

Urszula Bronowicka-Mielniczuk [a], Jacek Mielniczuk [a,*]

[a] Department of Applied Mathematics & Computer Science, University of Life Sciences in Lublin, Głęboka 28, Lublin 20-612, Poland

ARTICLE INFO

Keywords:
Interpolation methods
Environmental monitoring
Soil properties
Prediction errors statistics
Entropy
Quantile related inequality indices

ABSTRACT

Spatial interpolation has been applied for mapping various variables in a wide range of environmental disciplines. This study aims to develop novel tools for examining the relative performance of different interpolation methods. We shall quantify and compare the quality of interpolation models by applying, among others, some inequality indices of error distributions. Such indices can generally be classified as non-dimensional and global. The performance measures explored here provide a valuable supplement to the conventional accuracy assessment, and have so far received only scant attention in the relevant literature. Given a wide range of potential applications for the methods discussed here, the main focus of the paper will be on empirical research concerning variability of soil properties. The eight interpolation methods, i.e., Inverse Distance Weighting (IDW), Modified Shepard's Method (MS), Radial Basis Function (RBF), Natural Neighbour (NaN), Nearest Neighbour (NeN), Triangulation with Linear Interpolation (TIN), Local Polynomial (LP) and Ordinary Kriging (OK) were applied to estimate spatial distribution of soil pH, nitrogen, potassium and phosphorus content. Biplot methods were applied to visually examine the numerical results on the assessment of prediction quality. The ordinary kriging showed superior performance compared to the competing methods in majority of the cases. Significantly, predictions by kriging approaches revealed substantial improvement by considering data transformations. As concerns the other tested methods, the IDW and the LP algorithms tend to share similar characteristics. In turn, the NeN, RBF and MS algorithms scored relatively small inequality indices, when compared to the other methods. The use of new proposed measures will enable practitioners to gain more insightful and comprehensive evaluations of spatial interpolation techniques.

1. Introduction

Applications of geospatial methods are readily found within a multifold scientific disciplines including environmental science, agriculture, ecology, geology, epidemiology and socio-demographics. There are many geostatistical techniques that can be employed to explore distribution patterns of spatial phenomena (Anselin, 2006). Rich literature has provided a comparative analysis of interpolation algorithms. Despite significant developments in both geospatial analysis methods and software tools (Anselin, 2013), distance-based error statistics are still at the core of performance evaluation. Our paper can be viewed as making a complementary contribution, in particular by applying some relatively novel performance measures for the assessment of prediction errors. Special emphasis is placed on compiling information from multiple statistics to provide a comprehensive picture of the model quality.

There is a wide range of applications across various disciplines both for the geostatistical methods and performance measures discussed here. Potential applications will be demonstrated by considering a comparative study of the spatial variability of soil properties. There is a large body of recent literature on the geostatistical monitoring of soil properties (AbdelRahman et al., 2021; Ijaz et al., 2023; Tiruneh et al., 2023). As (Bogunovic et al., 2021) put it, understanding the spatial variability of soil chemical properties improves agricultural production, reduces environmental problems (e.g., soil pollution, offsite effects), and contributes to achieving sustainable agroecosystems. Spatial interpolation has a wide range of applications in precision agriculture and in the management of soil contamination (Shi et al., 2009; Ma et al., 2018; Radočaj et al., 2021). Much attention has also been paid to predicting topsoil pH (Laslett et al., 1987; Laslett and McBratney, 1990; Robinson and Metternicht, 2006; Shi et al., 2009). This characteristic is widely recognised as an important property that has a significant impact on the productivity of the soil. A number of studies have also been carried out

* Corresponding author.
E-mail addresses: urszula.bronowicka@up.lublin.pl (U. Bronowicka-Mielniczuk), jacek.mielniczuk@up.lublin.pl (J. Mielniczuk).

https://doi.org/10.1016/j.ecolind.2023.110551
Received 9 February 2023; Received in revised form 14 June 2023; Accepted 22 June 2023
Available online 11 July 2023

to evaluate the suitability of interpolating models for mapping soil organic matter. It remains a key issue for both agricultural production and environmental protection to have detailed and accurate information on the spatial distribution of soil organic matter (Zhao et al., 2016; Durdević et al., 2019; Lai et al., 2021; Barrena-González et al., 2022). Regardless of the ultimate purpose of the interpolation method, a common concern in the literature on the subject is a relative performance of interpolation algorithms. Unfortunately, there are quite a few inconsistencies in the final conclusions about the optimal method. Here is a brief review of relevant literature.

In a nutshell, most related developments usually focused on relative performance of a kriging technique versus one or two alternative deterministic methods such as inverse distance weighting (IDW), spline, radial basis functions (RBF), natural neighbour interpolation (NaN), Thiessen polygons (TP), local polynomial interpolation (LPI), global polynomial interpolation (GPI).

In most of the studies, geostatistical method was generally found to outperform deterministic competitors. Examples include (Laslett et al., 1987; Voltz and Webster, 1990; Leenaers et al., 1990; Hosseini et al., 1994; Gotway et al., 1996; Kravchenko and Bullock, 1999; Anselin and Le Gallo, 2006; Panagopoulos et al., 2006; Yasrebi et al., 2009; Reza et al., 2010; Xie et al., 2011; Mousavi et al., 2017; Long et al., 2020; Ouabo et al., 2020; Abdulmanov et al., 2021). Conversely, the case studies by (Qiao et al., 2018; Myslyva et al., 2019; Zhao et al., 2019; Benslama et al., 2020; AbdelRahman et al., 2021; Radočaj et al., 2021) revealed that deterministic methods resulted in a better performance than kriging.

Several authors including (Schloeder et al., 2001; Mueller et al., 2004; Ouabo et al., 2020) did not identify the overall best performing method in the study area.

At present, there is substantive literature reporting extensive tests involving a variety of both deterministic and stochastic interpolation techniques, including, among others, lognormal ordinary kriging (LOK), co-kriging, empirical Bayesian kriging (EBK), and regression kriging (RK). Not surprisingly, there has been a lack of agreement with the final recommendations, however, most studies exhibited relative superiority of the kriging techniques. To name a few, we mention (Bishop and McBratney, 2001; Bekele et al., 2003; Zare-Mehrjardi et al., 2010; Emadi and Baghernejad, 2014; Chen et al., 2016; Bhunia et al., 2018; Shen et al., 2019; Tunçay, 2021).

Remarkably, (Attaeian et al., 2015; Saha et al., 2022) concluded that a deterministic method results in a better performance than stochastic alternatives. Finally, the results reported by (Robinson and Metternicht, 2006; Barrena-González et al., 2022) demonstrate that there is no substantial difference between competing interpolators.

It can be noted that in the vast majority of papers the comparative evaluation of interpolation methods is performed by summarizing the leave-one-out cross-validation errors in a single metric. The choice of the appropriate metric for aggregating prediction errors is a matter of considerable debate providing conclusions being far from clear (see (Hodson, 2022). The performance of interpolation techniques is typically evaluated using the RMSE or MAE accuracy metric as the primary criterion. There is currently a substantial literature providing fairly comprehensive comparisons between these criteria. (Willmott and Matsuura, 2005 and Willmott et al., 2009) advised against the reporting of RMSE and recommended the use of MAE as it is less sensitive to extremes. However, (Chai and Draxler, 2014) presented arguments for favouring RMSE. They reported that RMSE is more appropriate than MAE when the errors are normally distributed, while the opposite is true in the case of a uniform distribution. In turn, (Karunasingha, 2022) concluded that RMSE gives better performance for platykurtic distributions, whereas MAE is preferred for leptokurtic distributions. Other useful, and recommended references include (Armstrong, 2001; Willmott et al., 2015 and Liemohn et al., 2021). For a comprehensive review of the issues involved in model validation, we refer the reader to (Bellocchi et al., 2010) and the references therein. (Eker et al., 2019) provide a comprehensive overview of the literature on validation in environmental modeling.

Given the legitimate criticisms of the standard distance-based measures, we will use multiple statistics to provide a more reliable picture of model accuracy. The literature provides quite a few alternative measures that have been reported as the overall performance scores for a comparison of multiple models. Examples include the Pearson correlation coefficient, the Nash–Sutcliffe efficiency and its modified forms, Legates and McCabe's index, relative root mean square error of Loague and Green, Willmott's indices of agreement and their refined versions, Mielke's permutation indices and the associated transformation considered by Watterson. The aforementioned criteria have received considerable attention over the years, beginning with the works on hydrologic and hydroclimatic models: (Nash and Sutcliffe, 1970; McCuen and Snyder, 1975; Willmott, 1981; Willmott, 1982; Mielke, 1984; Willmott et al., 1985; Willmott et al., 2012; Loague and Green, 1991; Legates and McCabe, 1999; Legates and McCabe, 2013; Krause et al., 2005). Some examples of recent applications include (Nas and Berktay, 2010; Meng et al., 2013; Gong et al., 2014; Arslan and Turan, 2015; Duveiller et al., 2016; Ozelkan et al., 2016; Deepika et al., 2020; Ananias et al., 2021; Yang and Xing, 2021).

A thorough report summarising commonly used error metrics is given by (Botchkarev, 2019). A wide range of information about assessment of interpolation techniques can be found in the review papers by (Li and Heap, 2011; Li and Heap, 2014). The issues concerning interpolation techniques will not be dealt with in great depth here. The reader can find required information in (Isaaks and Srivastava, 1989; Buhmann, 2003; Wackernagel, 2003 and Webster and Oliver, 2007).

2. Data set and methods

In the course of time much attention has been paid to quantitative comparison of spatial interpolation techniques. In the present paper we intend to make some contributions to this area by investigating some potentially useful concepts rooted in econometrics and information theory.

Recommended performance indices are illustrated with a comparative evaluation of interpolation models for mapping soil properties. The data upon analysis were collected using the LUCAS open-access dataset of topsoil properties available for the European Union (source of data: https://esdac.jrc.ec.europa.eu/resource-type/soil-point-data). We shall confine ourselves to dealing with 1377 sampling locations covered the region of Poland. Our interest focuses upon the following soil variables: total nitrogen (N), phosphorus (P), potassium (K), organic carbon (OC) and pH in H_2O. For more methodological details, see (Orgiazzi et al., 2018; Ballabio et al., 2016; Ballabio et al., 2019; Commission et al., 2020; Fernández-Ugalde et al., 2020).

We shall make comparative evaluation of eight interpolation methods including Triangulation with Linear Interpolation (TLI), Inverse Distance Weighting (IDW), Modified Shepard's (MS), Nearest Neighbour (NeN), Natural Neighbour (NaN), Local Polynomials (LP), Radial Basis Function (RBF) and Ordinary Kriging (OK). For the OK method, interpolations were performed with different types of variogram functions. Specifically, stable (sta), spherical (sph), pentaspherical (psph), hole effect (hoe), exponential (exp), circular (cir), Gaussian (gau) and J-Bessel (JBe) variograms were used. Calculations were carried out both on the original observations and on the transformed data. A Box-Cox and a logarithmic transformation were used. OK interpolations and variogram modelling were performed in ArcMap10.2 using built-in optimising algorithms. Interpolations for the other methods were performed in Surfer 24.2.240. Five variants of basis functions were used for the RBF method: multiquadratic (MQ), inverse multiquadratic (IMQ), natural cubic spline (NCS), thin plate spline (TPS) and multilog (Mlog). The IDW and LP algorithms were performed by setting power parameters to 1 and 2, hereafter denoted as p1 and p2, respectively.

The class of primary error metrics to be considered include the root mean square error (RMSE) and the mean absolute relative error (MARE). To give a more comprehensive description of the model accuracy, we also employed Willmott's index of agreement and the arcsine transformation of Milke's index proposed by (Watterson, 1996), the details of which will be explained in Appendix A.

With the purpose of exploring a potential use in performance assessment, some inequality curves and indices were considered in this study. We now highlight the major theoretical developments of this work. Throughout this paper use will be made of quantile versions of inequality curves, the definition of which goes back to (Prendergast and Staudte, 2016). These will be described in more detail in Section 3.2. In Section 3.3, we suggest and briefly discuss an alternative strategy for derivation of inequality measures. Overall, our proposition mirrors standard approach by reporting the ratio of total errors received by the top and bottom fractions of actual scores. Model assessment shall also be assisted by examination of asymmetry. Section 3.4 briefly touches on the quantile measure of skewness we shall report on. In Section 3.5 we introduce a biparametric extension of the generalized entropy index involving a two-parameter deformation of Tsallis' logarithm. The setting in which this has been done is that of Csiszár's divergence.

Multivariate data analysis methods were used to integrate numerical results on the relative performance of interpolation techniques. To this end, principal component analysis (PCA) was performed for each of the soil parameters analysed. The information extracted was further subjected to cluster analysis (CA). For interpolation methods, agglomerative cluster analysis was applied using Ward's algorithm with the Euclidean metric as a similarity measure. Clustering for quality indices was performed by applying the K-means technique. The results are presented graphically by means of PCA biplots. The statistics applied in this paper were implemented in the *R* computing software (R Core Team, 2021). The multivariate statistical analyses were performed using the factoextra package (Kassambara and Mundt, 2020).

3. New measures for the evaluation of interpolation algorithms

3.1. Preliminaries

We first set up some notation and terminology. Let N be a positive integer and let $\mathbb{R}_+^N$ denote the set of all real (row) vectors with N positive components. A vector $\boldsymbol{x}$ in $\mathbb{R}_+^N$ is a *scalar vector* provided $x_1 = \cdots = x_N$. We write $\overrightarrow{\boldsymbol{x}}$ for the increasing rearrangement of $\boldsymbol{x}$, i.e. $\overrightarrow{x}_1 \leqslant \overrightarrow{x}_2 \leqslant \ldots \leqslant \overrightarrow{x}_N$. Hereafter, any positive N-vector $\boldsymbol{x}$ is considered together with the corresponding probability (normalized) vector $\boldsymbol{x}^*$ obtained by dividing each component x_i by the sum of the components. We let $\mathscr{P}^N$ denote the set of all normalized N-vectors.

We shall consider a setting with the *actual* $\boldsymbol{a} = (a_1, a_2, \ldots, a_N)$ and the *forecasted* $\boldsymbol{f} = (f_1, f_2, \ldots, f_N)$ data vectors. In the sequel, we shall almost exclusively be dealing with the *relative absolute* errors $\varepsilon_i = a_i^{-1}|a_i - f_i|$. We assume throughout that all actual scores a_i being considered have positive values. This section is then concerned with discussing some measures of inequality. Reference for this subject is made to the book by (Arnold and Sarabia Alegria, 2018). One should bear in mind that everything established below applies to relative absolute errors. As already indicated, the R software environment has been used for the study reported here. The numerical methods used throughout provide a potentially useful complement to the *ineq* package (Zeileis and Kleiber, 2014). A number of indices we shall deal with call for determining sample quantiles. Here we use the *type 5* version of the *quantile* command in R software, which uses a continuous interpolation scheme. The sample quantile function corresponding to $\boldsymbol{x} \in \mathbb{R}_+^N$ will be denoted by $Q_x(u), 0 \leqslant u \leqslant 1$. When only one sample is under discussion we shall omit mention of $\boldsymbol{x}$ and write $Q(u)$. The reader is referred to (Hyndman and Fan, 1996) for a detailed discussion of quantile definitions in statistical packages.

3.2. Quantile versions of inequality curves

The Lorenz curve and Gini index are very popular in the study and quantification of inequality, particularly in economics. Since the pioneering works of (Lorenz, 1905 and Gini, 1921) these basic concepts have been subjected to several modifications and refinements. We shall make use of the quantile versions due to (Prendergast and Staudte, 2016), which we introduce below

$$L_{(1)}\left(u, \boldsymbol{x}\right) = u\frac{Q_x(0.5u)}{Q_x(0.5)}, \quad L_{(2)}\left(u, \boldsymbol{x}\right) = u\frac{Q_x(0.5u)}{Q_x(1-0.5u)}, \tag{1}$$

$$L_{(3)}\left(u, \boldsymbol{x}\right) = 2u\frac{Q_x(0.5u)}{Q_x(0.5u) + Q_x(1-0.5u)} = \frac{2}{1/u + 1/L_{(2)}(u, \boldsymbol{x})}.$$

Therefore, $L_{(1)}$ compares the median corresponding to the smallest proportion u of the population with the median for the entire population. In turn, $L_{(2)}$ captures the ratio between the median values of the top and bottom extremes of the population, whereas $L_{(3)}$ gives the median of the lower fraction, relative to the mid-range value of the middle fraction of the population. It is important to realize that $L_{(2)}$ and $L_{(3)}$ provide quantile counterparts of inequality curves introduced by (Gastwirth, 2016) and discussed in more detail by (Arnold and Sarabia Alegria, 2018)[Section 6.4.3]. To each of the $L_{(i)}$ given by (1) there corresponds a coefficient of inequality defined by

$$G_{(i)}(\boldsymbol{x}) = 2\int_0^1 \left|u - L_{(i)}(u, \boldsymbol{x})\right| du = 1 - 2\int_0^1 L_{(i)}(u, \boldsymbol{x})\, du. \tag{2}$$

3.3. Modification proposal

Hereafter, the counterparts of inequality curves are defined and an investigation is made of two possible new versions of these functions. The core concept is to report proportions of errors held by the top and bottom fractions of the actual scores $\overrightarrow{\boldsymbol{a}}$. Towards this end, it is convenient to consider an rearrangement of $\boldsymbol{\varepsilon}$ with respect to the ascending order of actual scores. Formally, for the jth value in the sequence $\overrightarrow{a}_1 \leqslant \overrightarrow{a}_2 \leqslant \ldots \leqslant \overrightarrow{a}_n$ there is an error $\overleftrightarrow{\varepsilon}_j$ in the corresponding, not necessarily monotonic, sequence $\overleftrightarrow{\boldsymbol{\varepsilon}} = \left(\overleftrightarrow{\varepsilon}_1, \overleftrightarrow{\varepsilon}_2, \ldots, \overleftrightarrow{\varepsilon}_n\right)$.

Here, we provide computational details of the new empirical curves under study. Given $\overleftrightarrow{\boldsymbol{\varepsilon}}$, the inequality curve $L_1^*(u, \boldsymbol{\varepsilon})$, where $0 \leqslant u \leqslant 1$, is obtained by the linear interpolation of

$$L_1^*\left(\frac{i}{n}\right) = \frac{i}{n}\cdot\frac{i^{-1}\sum_{k=1}^{i}\overleftrightarrow{\varepsilon}_k}{n^{-1}\sum_{k=1}^{n}\overleftrightarrow{\varepsilon}_k} = \frac{\sum_{k=1}^{i}\overleftrightarrow{\varepsilon}_k}{\sum_{k=1}^{n}\overleftrightarrow{\varepsilon}_k} \tag{3}$$

for $i = 0, 1, \ldots, n$, with the convention $L_1^*(0) = 0$. Hence, L_1^* captures the ratio of the partial sum of errors held by a lower fraction of actual scores to the overall total error. In turn,

$$L_2^*\left(\frac{i}{n}\right) = \frac{\sum_{k=1}^{i}\overleftrightarrow{\varepsilon}_k}{\sum_{k=n-i+1}^{n}\overleftrightarrow{\varepsilon}_k} \tag{4}$$

captures the gap between error totals (equivalently, means) received by the top and bottom fractions of the actual scores. We briefly comment on the three points raised by our formulation. First and foremost, the proposed curves may not lie beneath the perfect equality line $y = x$. Further, they are neither increasing nor convex in general. We also call the reader's attention to the fact that there is no upper limit for the ratio between the lower and the upper sums (means) involved in (4). As usual, once an inequality curve L_i^* is specified, we can define the corresponding inequality index G_i^* via the following integral

$$G_i^*\left(\boldsymbol{\varepsilon}\right) = 2\int_0^1 |u - L_i^*(u, \boldsymbol{\varepsilon})|du. \quad (5)$$

One easily establishes the quantile counterparts of (3) through (5) by incorporating the ideas developed by (Prendergast and Staudte, 2016). We will write the quantile versions as $L_{(i)}^*$ and $G_{(i)}^*$, respectively. Appendix C provides further methodological details and reports some specific results.

3.4. Quantile measure of skewness

For $0 < r < 0.5$, the *rth coefficient of skewness* is defined to be the ratio

$$\gamma\left(r\right) = \frac{Q(r) + Q(1-r) - 2Q(0.5)}{Q(1-r) - Q(r)}. \quad (6)$$

The case of $r = 0.25$ was originally considered by (Yule, 1911 and Bowley, 1920), whereas the general version (6) was put forward by (David and Johnson, 1956) and has subsequently been used by (Hinkley, 1975 and Hogg, 1974), among others. One readily establishes that coefficient (6) is sign equivariant and location and scale invariant. Furthermore, it holds that $-1 \leqslant \gamma(r) \leqslant 1$ with $\gamma(r) = 0$ for symmetric distribution. For more information, we refer interested readers to (Joanes and Gill, 1998; MacGillivray, 1986; Staudte, 2014) and references therein. In the present note, our evaluation of the models' performances is based upon a *weighted coefficient of skewness* γdefined as

$$\gamma = \int_0^{0.5} \gamma\left(r\right)\omega\left(r\right)dr, \quad (7)$$

where $\omega(r)$ is a weight function given by $\omega(r) = \pi\sin(2\pi r)$.

3.5. Entropy based inequality measures

We deal here in detail with a biparametric extension of the generalized entropy index. A key tool underlying our development is the class of *φ-divergences* originating in the papers of (Csiszár, 1963 and Morimoto, 1963). Consider $\varphi(\tau)$ as a fixed convex function in the interval $(0, \infty)$ being equal to zero and strictly convex at $\tau = 1$. We define the *φ-divergence* D_φ on the product monifold $\mathscr{P}^N \times \mathscr{P}^N$ by putting for any given vectors $\boldsymbol{p}, \boldsymbol{q} \in \mathscr{P}^N$

$$D_\varphi(\boldsymbol{p}, \boldsymbol{q}) = \sum_{i=1}^N q_i \varphi\left(\frac{p_i}{q_i}\right).$$

Therefore, by denoting the scalar vector in $\mathscr{P}^N$ by $\boldsymbol{u}_N^*$, we can define an inequality index of a positive vector $\boldsymbol{x}$ by putting

$$I_\varphi(\boldsymbol{x}) = D_\varphi\left(\boldsymbol{u}_N^*, \boldsymbol{x}^*\right) = \sum_{i=1}^N x_i^* \varphi\left(\frac{1}{Nx_i^*}\right). \quad (8)$$

We shall hereafter consider a special case of (8) corresponding to the deformed (k, r)-logarithm given by

$$\log_{\{k,r\}}(x) = x^r \frac{x^k - x^{-k}}{2k} = x^r \log_{\{k\}}(x), \quad (9)$$

where the parameters k and r are restricted to a two-dimensional region

$$\mathscr{R} = \left\{(k, r) : 0 < k < 1 \wedge \left|\frac{1}{2} - k\right| - \frac{1}{2} \leqslant r \leqslant \frac{1}{2} - \left|\frac{1}{2} - k\right|\right\}, \quad (10)$$

in which the extended logarithm meets the requirements of being increasing and concave. Noteworthy, $\mathscr{R}$ accommodates the supplementary range constraint $r > k - 1$ making the deformed logarithm to satisfy the property of having integrable divergence for $x \to 0^+$ (Kaniadakis et al., 2005). It is worth noting here that $x^{-1}\log_{\{k,r\}}(x)$ goes to zero as x approaches infinity. Finally, due to evenness of (9) with respect to k, we shall be concerned here with the case $k > 0$. The definition of relative entropy induced by generalized logarithm (9) goes back to (Sharma and Taneja, 1975 and Mittal, 1975) with further studies by (Kaniadakis et al., 2004; Kaniadakis et al., 2005; Scarfone, 2006; Furuichi, 2010). Let $I_{\{k,r\}}(\boldsymbol{x})$ designate a biparametric family of inequality indices derived from (8) by setting $\varphi(x) = -\log_{\{k,r\}}(x)$. For the case $r = \pm k, k < 0.5$, one recovers the generalized entropy index involving Tsallis' logarithm, which turns into the familiar Theil index if $k \to 0$. A thorough discussion of the generalized entropy index and the Theil index is contained in Section 5.4.2 of (Arnold and Sarabia Alegria, 2018). Throughout this paper, use will be made of the integral version of the new index given by

$$I_{STM}\left(\boldsymbol{x}\right) = \iint_{\mathscr{R}} I_{\{k,r\}}\left(\boldsymbol{x}\right)\omega(k, r)\,dkdr. \quad (11)$$

We take a bivariate function ω to be $\omega(k, r) = 1 - \sqrt{k^2 - r^2}$, which reflects a priori preference over region $\mathscr{R}$ to the generalized entropy index by assigning relatively higher weights to Tsallis' logarithm. As is commonly done, we performed unit range normalization which scales divergence measures (8), (11) to lie in the unit interval $[0, 1]$. Then by convention, a value of zero corresponds with perfect equality, a value of one indicates a case of extreme inequality. Some remarks are in order. A principal advantage of the approach outlined lies in the great flexibility of the mathematical framework. A potential drawback with this type of index is that it is determined assuming a specific prior distribution ω for the parameters of interest. The same concern applies to the index defined by (7) as the choice of integral aggregation to establish an overall score seems rather arbitrary. Finally, in an information-theoretic context, entropy-like measures have long been used to calculate the diversity of ecological systems (see (Magurran, 2004)). Thus, an information measure derived from (11) can provide a flexible metric to compare diversity across biological communities.

4. Results

4.1. Introduction

Multivariate data analysis methods were used to integrate numerical results on the relative performance of interpolation techniques. Biplots allow visual exploration of correlations among the performance indices and reveal the ranks of interpolation algorithms according to their values in each of the indices under study. In the following, the basic ideas for interpreting a biplot representation are explained. Objects (interpolation methods) that are close to each other in the biplot have similar scores on the variables (performance measures) displayed in the plot. The diagram also reveals the variance–covariance structure among the variables. The angle between a pair of vectors corresponds to the correlation between the associated variables. A small angle indicates a high positive correlation between the variables, whereas nearly perpendicular vectors reflect a very weak correlation. Negatively correlated variables are positioned on opposite coordinate plane quadrants. The lengths of the variable vectors along the principal axes are directly related to their contribution to the axes. The longer the vectors are, the more influence they have on the associated component. A significant merit of a biplot is that it can graphically show the which-won-where pattern among objects and variables. In a nutshell, the objects are ranked according to their orthogonal projections onto the line passing through the biplot origin and the marker for the variable. Those projecting furthest away from the biplot origin in the direction of the variable vector have the highest scores. The biplot origin represents the average value for each variable. It is worth recalling here that the top value of both Willmott's and Watterson's metrics is one indicating a perfect fit. We will not keep repeating this point throughout this paper.

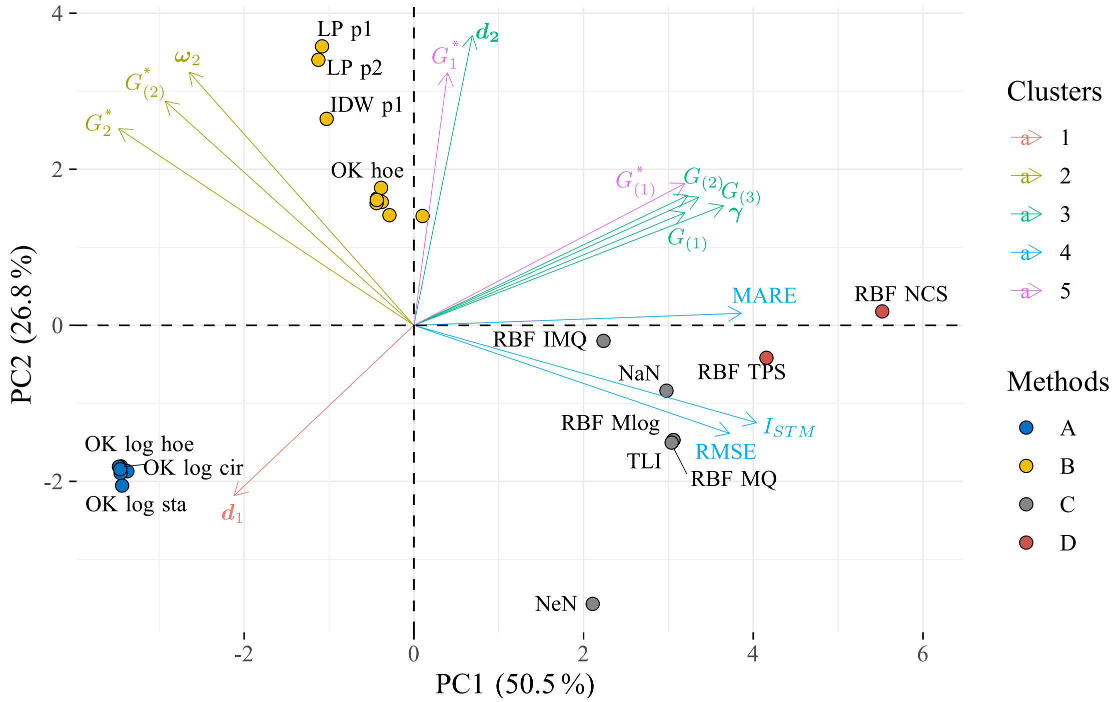

Fig. 1. PCA biplot for N.

4.2. Nitrogen

The first three principal components explain 95.3% of the total variance with 50.5% for the first component (henceforth referred to as PC1) and 26.8% for the second component (PC2). It follows from Fig. 1, that PC1 has a large positive association both with the standard accuracy metrics MARE, RMSE, and the coefficient of asymmetry. It is noteworthy that almost all inequality indices are located on the right half-plane yielding a high positive influence along PC1 ($I_{STM}, G_{(1)}, G_{(2)}, G_{(3)}, G^*_{(1)}$) and a negligible one along PC2. In turn, G^*_2 and $G^*_{(2)}$ have a strong negative influence along PC1. Concerning PC2, the Willmott measures of agreement and G^*_1 were the main contributors (towards the positive half-axis). The third component (PC3) has also large positive associations with d_1 version of Willmott's index. Clustering result for the quality measures distinguished 5 groups. As detailed in Section 4.1, the cosine of the angle between the vectors of two variables approximates the correlation coefficient between them. According to Fig. 1, there is a strong collinearity of RMSE, I_{STM} and the members of the second cluster. In addition, a near perfect correlation occurs between the quantile versions of inequality indices and the asymmetry coefficient. On the contrary, watterson-willmott correlation is quite poor. Each interpolation algorithm was classified into one of the four groups. In the first place, it is observed that the kriging variants with log transformation clustered quite tightly in group A with long negative projections onto the vectors in the right half-plane. By contrast, group A shows relatively high projections onto the vectors corresponding to the measures of agreement, thus indicating the overall best performance of this cluster. It is worth

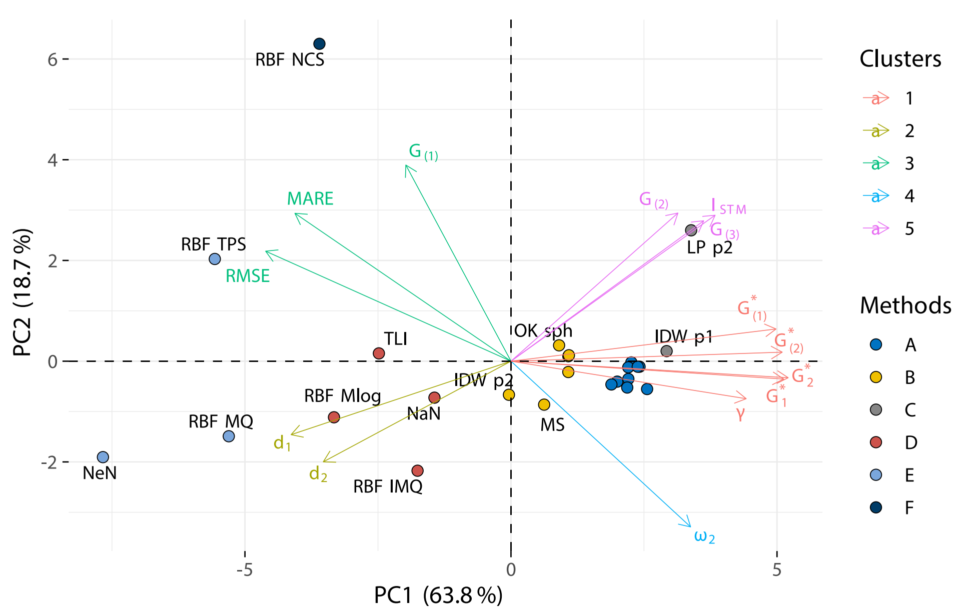

Fig. 2. PCA biplot for P.

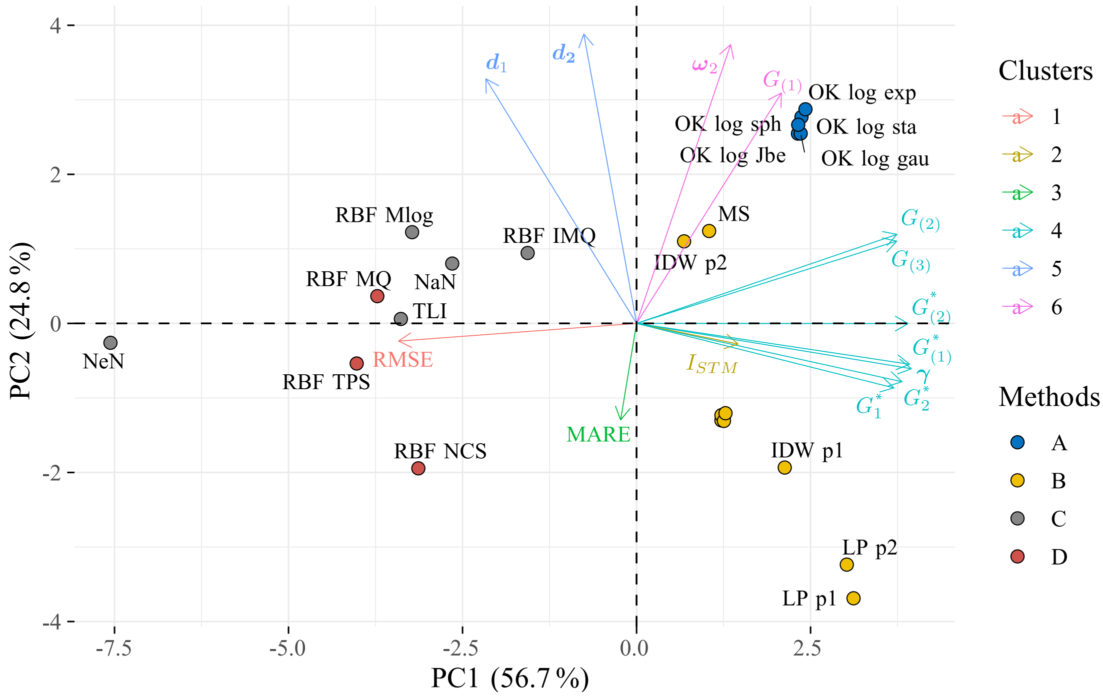

Fig. 3. PCA biplot for K.

noting that the lack of the logarithmic transformation places the kriging scores on the level of those provided by the deterministic methods in group B. To further figure out the relative performance of the kriging algorithms, a zoomed-in version of group A is provided. As shown in Fig. B.6, the kriging algorithm with hole effect variogram performs a little bit better than the other members of the cluster. Fig. 1 provides a rough indication that algorithms associated with the positive part of PC1 have the relative worst performance among the methods being compared. Finally, the IDW and the LP algorithms are perceived as having a high similarity of performance scores.

4.3. *Phosphorus*

In Fig. 2, we have the biplot representation for the phosphorus data. PCA explain 82.5% of the total variance of the original data with only the first 2 axes (PC1 = 63.8%, PC2 = 18.7%). Concerning the positive half-axis of PC1 the strongest correlations are those with G^* measures. In turn, the negative half-axis of PC1 is strongly associated with the accuracy measures, excluding the Watterson index, which gives the largest contribution to PC2. The biplot reveals five distinct clusters of quality measures. There is a close positive correlation between γ and G^*

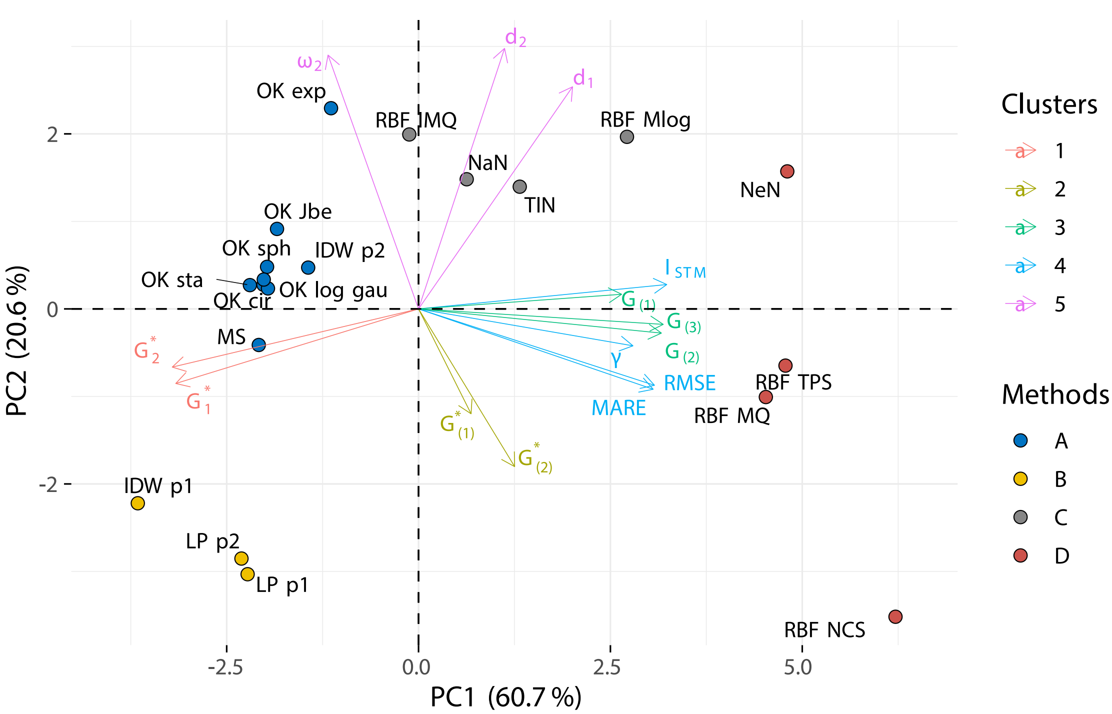

Fig. 4. PCA biplot for pH.

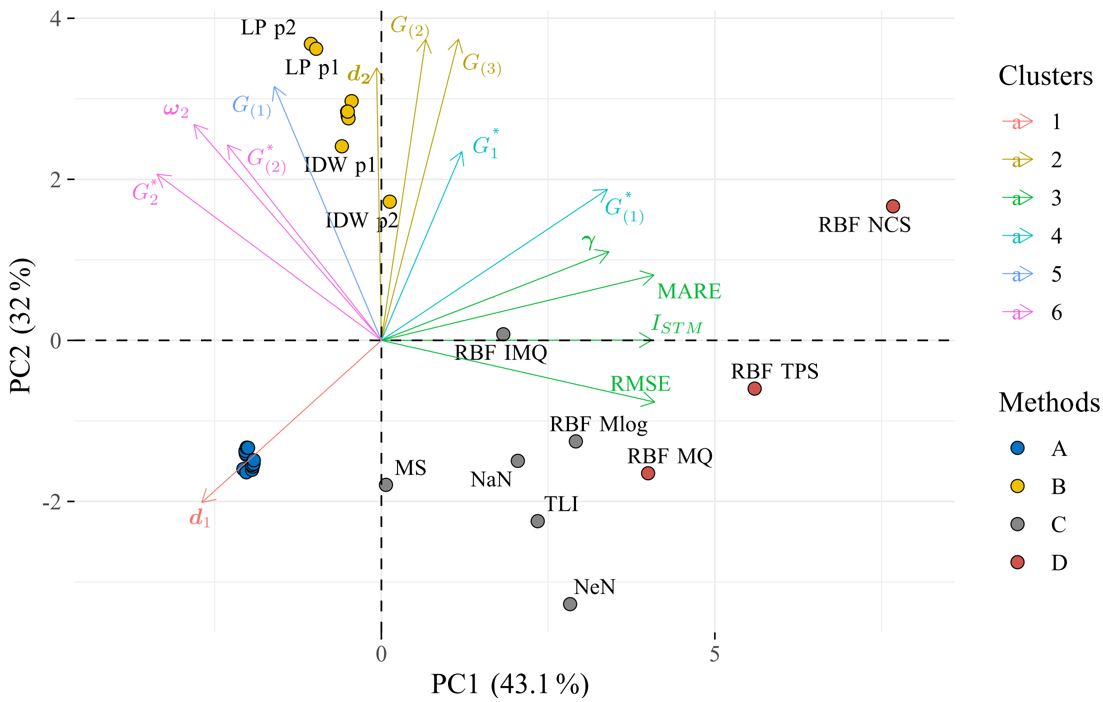

Fig. 5. PCA biplot for OC.

measures. Furthermore, let us note that the classical accuracy measures (MARE, RMSE) are strongly positively correlated and were gathered in a group, which, in turn, is strongly negatively correlated with γ and Waterrson's index, and additionally has a near-zero correlation with the fifth cluster. The clustering of interpolation algorithms resulted in the identification of six groups. In the corresponding biplot a slight pattern can be observed. Overall, the kriging methods with log data transformation (group A) received the highest accuracy ratings in terms of RMSE and MARE. Additionally, Fig. B.7, the zoomed region of Fig. 2, confirms the slight preference for the ordinary kriging with log transformation and stable variogram. Note, however, that the methods located on the right-hand side of the graph are characterised by greater asymmetry and above-average values of inequality indices gathered in the first and the fifth cluster. It is noteworthy that the algorithms in group B are clustered close to the average values for performance indices. The most striking finding to emerge from this analysis is that cluster E, characterized by the presence of RBF MQ, RBF TPS and NeN algorithms, reveals the best performance when models are ranked according to inequality criteria.

4.4. Potassium

Fig. 3 shows the PCA biplot for the results on potassium (K). The first three PCs accounted for 97.4% of the total variance with 56.7% for PC1 and 24.8% for PC2. The inequality and asymmetry indices exhibited the highest contribution to the positive half-axis of PC1, whereas RMSE scored a strong negative influence along PC1. Concerning PC2, the Watterson and Willmott measures of agreement were the main contributors. As for PC3, MARE and I_{STM} accounted for the highest contributions with positive and negative correlation, respectively. The clustering procedure for quality measures distinguished 6 groups. It is quite interesting that the fourth cluster gathered most of the investigated inequality indices and the asymmetry coefficient. We observe a near-perfect negative correlation between RMSE and the members of the fourth cluster. Let us also note that, unlike for nitrogen (N) and phosphorus (P), the conventional accuracy measures are weakly correlated. In Fig. 3, the interpolation techniques fell into four sectors. As a conclusion, the kriging methods with log transformation (located in the upper right-hand quadrant) performs much better than the rest of competitors, when compared with respect to accuracy scores. Again, the ordinary kriging with log transformation and stable variogram demonstrates the relative best behaviour among the members of the cluster (see Fig. B.8). The layout of other quality parameters reveals the poor performance of kriging methods with respect to symmetry and equality of relative errors. Interestingly, there is decisive evidence in favor of RBF and NeN relative to the other methods if one puts the inequality scores into consideration.

4.5. Potential of hydrogen

In the case of soil pH, the first three components explain 97.2% of the total variation with 60.7% and 20.6% for PC1 and PC2, respectively (Fig. 4). More specifically, G_1^* and G_2^* have a strong negative influence along PC1, whereas $I_{STM}, G_{(2)}, G_{(3)}$, RMSE and MARE scored the highest positive influence. Once again, the Watterson and Willmott metrics contribute the most for PC2. For PC3, it explains 15.9% of the total variance with $G^*_{(1)}$ and $G^*_{(2)}$ as the major (positively correlated) contributors. PCA yielded five distinguishable clusters of performance indices with a high degree of similarity. Interestingly, there is a very strong positive correlation in all but the last of the resulting classification groups. A remarkable character of the biplot is a high positive association between the third and the fourth clusters, which in turn, have a strong negative correlation with the first one. Surprisingly, there is no correlation between the first and the second cluster.

The clustering of interpolation techniques yields four separate groups. In short, the most accurate interpolation methods are located in the top-left corner of the biplot. It is worth making a point that different variants of kriging gathered in cluster A, have an additional advantage of scoring relatively small inequality indices. Hence one may conclude that kriging techniques stand out as superior to the other methods. Fig. B.9 reveals numerical results supporting the use of the kriging with stable variogram as the best performing method overall.

4.6. Organic carbon

It is important to note at the outset that the biplots for N (Fig. 1) and OC (Fig. 5) are strongly similar. PCA for OC demonstrated that the first three components explained 88.9% of the total variance (PC1 = 43.1%, PC2 = 32% and PC3 = 13.8%). Both classical indicators as well as I_{STM}, γ, and $G_2^*, G_{(1)}^*, G_{(2)}^*$ indicators highly contributed to PC1. The Watterson and Willmott measures of agreement and all $G_{(\cdot)}$ indexes presented strong associations with PC2. In turn, PC3 has also large positive association only with G_1^*. In Fig. 5, the performance measures fell into six sectors. In general, a strong correlation can be observed within the second and the third cluster. By contrast, the relatively poor correlations do occur between the members of the fifth cluster, which in turn is very weakly correlated with the second cluster. Finally, there is near-perfect correlation between $G_{(1)}^*$ and d_1.

Two additional features of the bibplot are worth noting here. There are relatively low and nearly equal proportions of variation accounted for the first two principal components. Thus the resulting biplot may not be sufficient to effectively explain the data patterns. It is also clear from the vector distribution that different metrics led to quite different rankings. In particular, the MS algorithm has scored the lowest value for each of the inequality measures gathered in the first or the second cluster. In turn, the class of kriging techniques in group B exhibits the best performance when reference is made to d_2. Numerical results for the third cluster reveal that the lowest of the G_2^* values are received from the RBF MQ, NeN and RBF TPS models, whereas the NeN, TLI and NaN models were found to have the best performance with respect to $G_{(2)}^*$. Regarding Watterson's index, the MS and the LP algorithms achieved superior results. In turn, NeN, OK BC stable and OK LOG stable were the top ranking methods with respect to d_1. As for the fifth cluster of metrics, the best results according to RMSE are provided by MS and LP, whereas the lowest MARE values are received from the OK LOG algorithms. Contrarily, OK BC outranks the competitors with respect to I_{STM} and γ. Regarding the metrics in the sixth cluster, the MS algorithm has been again recognized as the best interpolator. By scrutinizing the results from Fig. B.10, it is recognizable that the Box-Cox transformation produces slightly better results than its logarithmic alternative. Overall, applying transformations results in a quantitative improvement of the kriging predictions.

5. Discussion

Biplot methods have been applied to visually examine the numerical results on the assessment of prediction quality. The biplot analysis revealed the correlation patterns among the quality indices and provided substantial information on the relative performance of interpolation techniques. Here is a brief summary of our findings reported in Section 4. As was seen for majority of the cases regarded in this paper, the ordinary kriging beats the competing methods in accuracy. Significantly, predictions by kriging approaches revealed substantial improvement by considering data transformations. As concerns the other tested methods, the IDW and the LP algorithms tend to share similar characteristics. In turn, the NeN, RBF and MS algorithms scored relatively small inequality indices, when compared to the other methods. Interestingly, significant correlations were detected between I_{STM} and both MARE and RMSE accuracy metrics. Regarding MARE, the highest Pearson's correlations were .793 (N), .895 (pH) and .868 (OC), whereas .898 (N), .902 (pH) and .877 (OC) were the top scores for RMSE. Although it is difficult to draw firm conclusions from this survey, in light of statistics, as well as characteristics derived through the visual inspection of the PCA biplots, the kriging appears to have the overall best quality among all the interpolation methods evaluated. It was found that ordinary kriging with a hole effect variogram was the best interpolation technique for nitrogen (N). For the other parameters (pH, P, K, OC), ordinary kriging with a stable variogram showed the best performance. Remarkably, a logarithmic transformation of the data produced the best results when applied to potassium and phosphorus contents, while the Box-Cox transformation proved to be the most suitable for organic carbon. Recommendations on the choice of interpolation method can help to identify the location of potential sources of nutrient dispersion in agroecosystems and implement practices to prevent this from occurring. They can also be used to optimise fertilisation in environmental and economic terms.

6. Conclusions

Recognition of spatial patterns of soil fertility parameters is vitally important for optimizing crop management practices. Quality assessment of interpolation techniques for predicting variability of soil properties has been a long-standing topic in the literature. Accuracy is widely recognised as being the most important criterion for evaluating interpolation techniques. Conventional approaches often involve analysing multiple interpolation models and then selecting a single one that reduces a given error metric. The vast majority of studies rely on RMSE and MAE values to draw conclusions about the effectiveness of interpolation algorithms. The widespread use of these relatively standard metrics is largely due to their overall availability in popular software for geostatistical computations. Yet both of these approaches have been subject to considerable criticism in the literature. We therefore follow the widely used method of incorporating multiple statistics to provide a reliable assessment of model performance. We expect decision makers to consider the benefits of using a few relevant statistics on the distributional properties of prediction errors. The methods proposed here aim to provide a deeper insight into the issue of quality assessment by using some new inequality curves and indices. The indices under review have not yet received the attention they deserve in the relevant literature.

This paper presents the use of a family of readily interpretable dimensionless measures to quantify and compare competing predictions. The study reported here is far from exhaustive on the statistical properties of such indices. While the indices empolyed in this paper are useful in their own right, their statistical properties are still in need of further theoretical investigation. Aware of the non-exhaustive nature of our presentation, we attempt to give the reader a fair idea of the applicability of the new performance measures in testing interpolation models.

Declaration of Competing Interest

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

Data availability

The authors do not have permission to share data.

Appendix A. Non-dimensional measures of agreement

We present here a set of performance statistics which are commonly invoked in the assessement of spatial predictive models. There is a routine practice for using Willmott's index of agreement to evaluate the performance of forecast models. Throughout the paper use will be made of the two main versions of Willmott's index which will be discussed in more detail below. For any given positive integer k one defines the *Willmott index of*

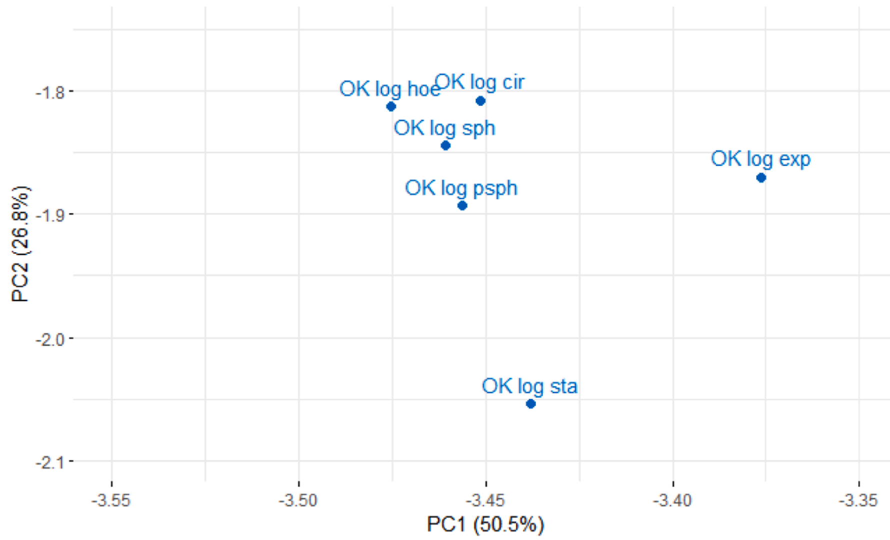

Fig. B.6. PCA biplot for N.

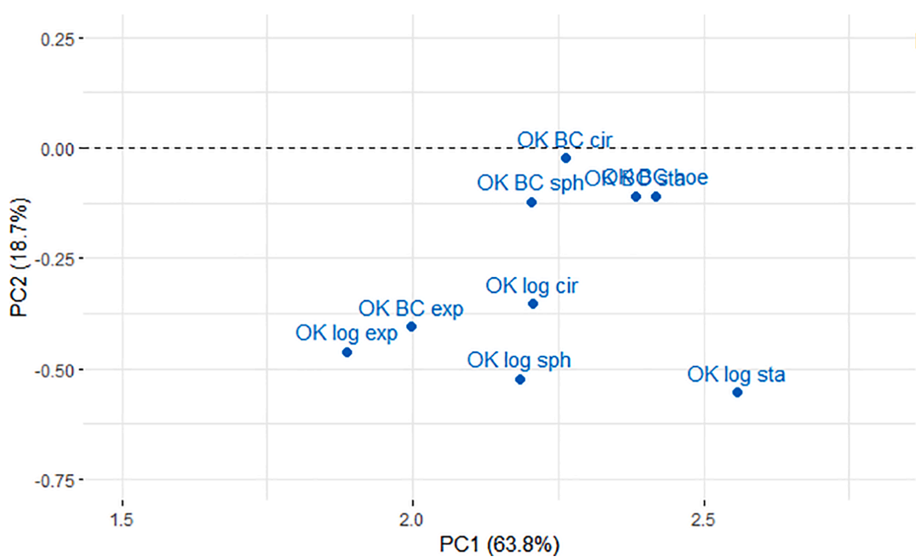

Fig. B.7. PCA biplot for P.

agreement as

$$d_k = 1 - \frac{\sum (f_i - a_i)^k}{\sum (|f_i - \overline{a}| + |a_i - \overline{a}|)^k}, \tag{A.1}$$

where $\overline{a}$ and $\overline{f}$ denote the means of the actual values a_i and the forecasts f_i. The original version for $k = 2$ was proposed by (Willmott, 1981; Willmott, 1982 and Willmott and Wicks, 1980), whereas a modification involving $k = 1$ was put forward by (Willmott et al., 1985). It is well known that both forms of the measure are bounded from below by zero, with the maximum value of one meaning a perfect match. The lower bound is reached if and only if $f_i - \overline{a} = \overline{a} - a_i$ for all i. Another concept referred to is Milke's permutation index introduced in (Mielke, 1984). By convenience of notation, put $s_2^* = \sum (f_i - a_i)^2 + 2\sum (f_i - \overline{f})(a_i - \overline{a})$.

The *Mielke permution index of the second order* may be written as

$$\mu_2 = 1 - \frac{\sum (f_i - a_i)^2}{s_2^*}. \tag{A.2}$$

The value of μ_2 ranges from -1 to 1. It is readily seen that $\mu_2 = 1$ is equivalent to complete agreement between the actual and model-predicted values, whereas $\mu_2 = -1$ if and only if $f_i - \overline{a} = \overline{a} - a_i$ for all i. The form of Mielke's index that we examine here is

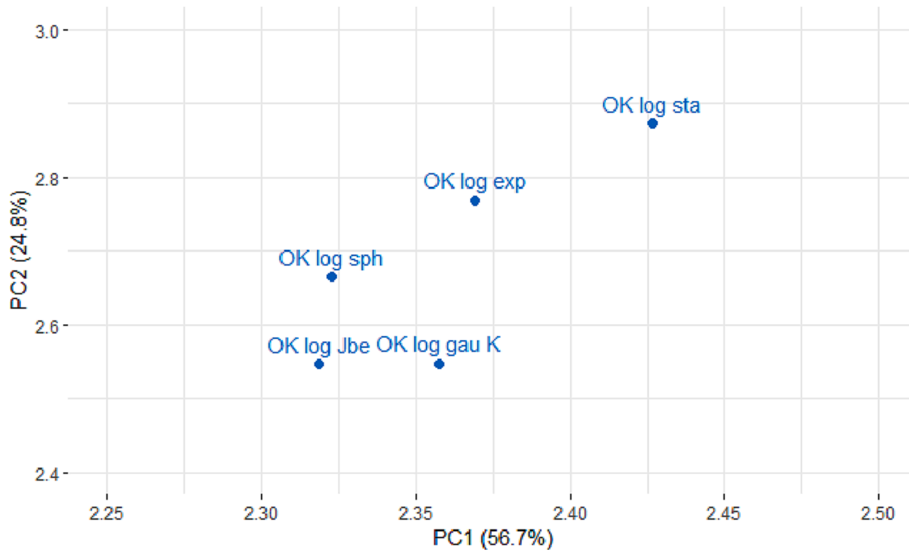

Fig. B.8. PCA biplot for K.

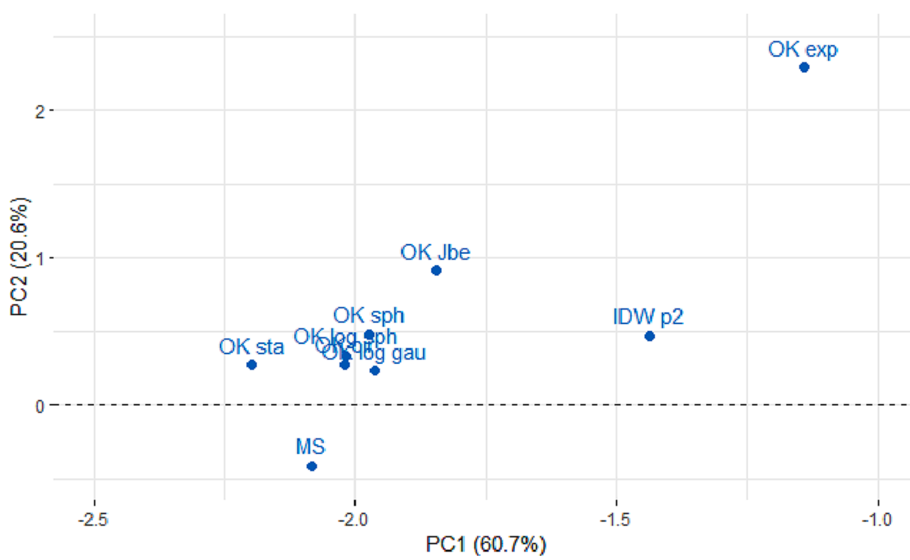

Fig. B.9. PCA biplot for pH.

$$\omega_2 = \frac{2}{\pi}\sin^{-1}\mu_2, \tag{A.3}$$

as originally considered by (Watterson, 1996). For additional information on these measures, we refer to (Duveiller et al., 2016; Willmott et al., 2012; Willmott et al., 2015).

Apart from these, the root mean square error (RMSE) and the mean absolute relative error (MARE) were calculated for evaluating model accuracy. The RMSE and MARE are defined as follows

$$RMSE = \sqrt{N^{-1}\sum(a_i - f_i)^2},$$

$$MARE = N^{-1}\sum\left|\frac{a_i - f_i}{a_i}\right|.$$

A comprehensive review and discussion of variety of methods was compiled by (Li, 2017).

Appendix B. Figures

Figs. B.6,B.7,B.8,B.9,B.10.

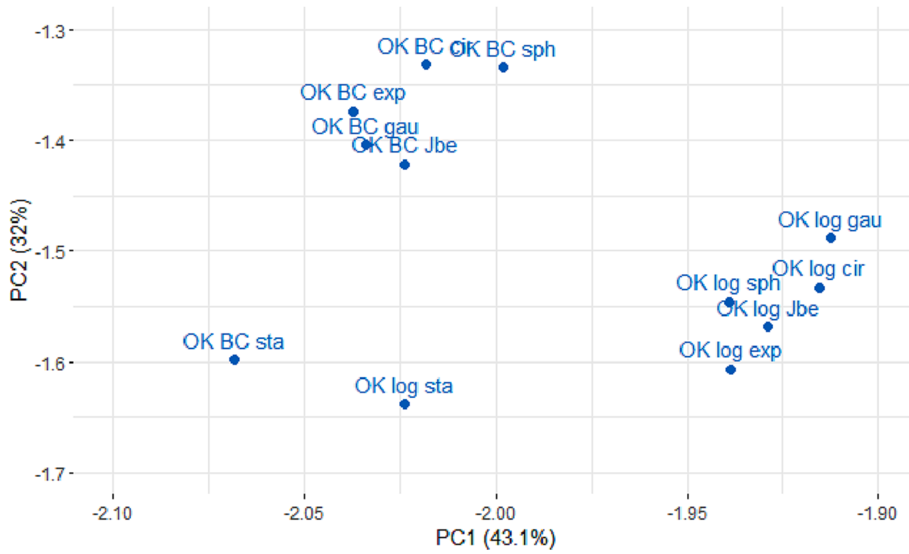

Fig. B.10. PCA biplot for OC.

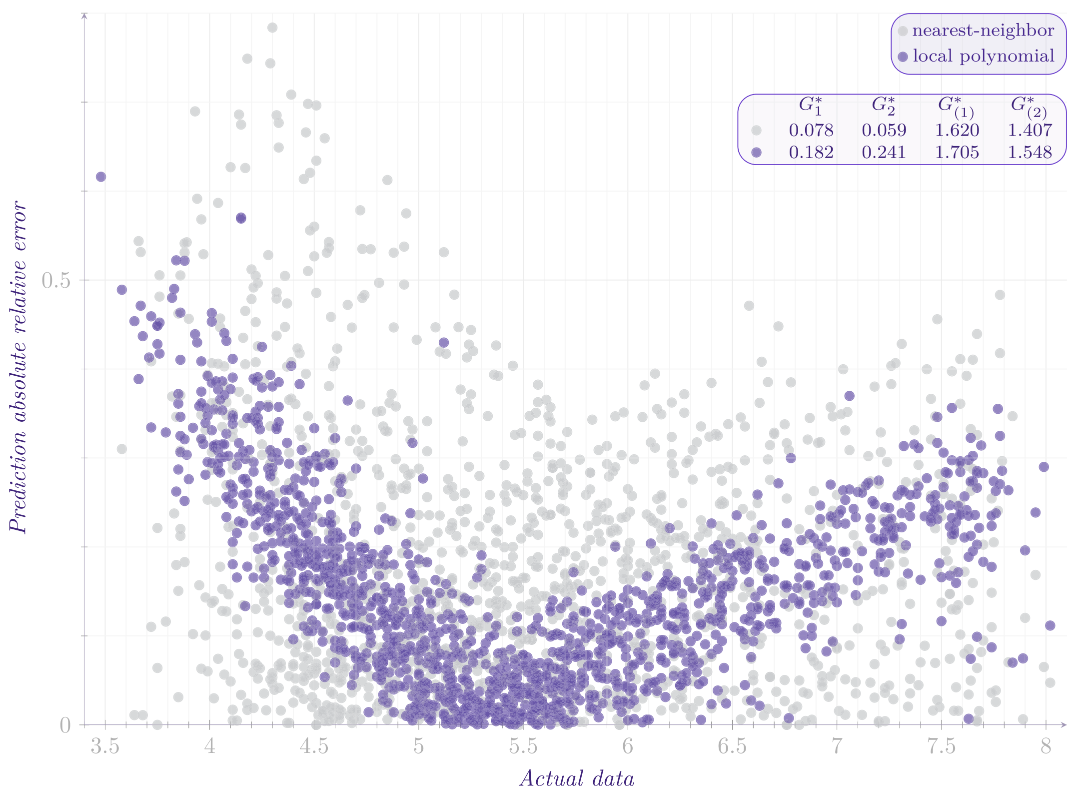

Fig. C.11. Relative error distribution for the potential of hydrogen.

Appendix C. Example

Here is an illustrative example including a comparison between the LP p1 and NeN models for predicting pH values. Their results provide an interesting contrast in terms of error distributions. Fig. C.11 shows the relationship between the distributions by putting the sorted values of the actual scores versus the associated errors. As Fig. C.11 indicates, NeN exhibits a notably greater uniformity of errors and clearly outperforms its competitor for all inequality statistics. On the other hand, the results of Fig. C.11 clearly demonstrates that NeN is susceptible of larger errors than LP p1 upon considering the middle fraction of the actual scores. Fig. C.12 complements the graphical assessment of the models by depicting the inequality curves L_2^*. It allows one to capture and quantify departures from uniformity upon considering the area between the diagonal line of perfect equality and the actual distribution curve. Basically, the smaller the area the more even the distribution of errors. This enables to assess the better performance of the NeN algorithm in terms of G_2^*. For comparison, Fig. C.11 reports the scores of relevant statistics quantifying distribution uniformity.

The analysis presented here draws on the complete information provided by a joint (bivariate) distribution of the actual scores and the

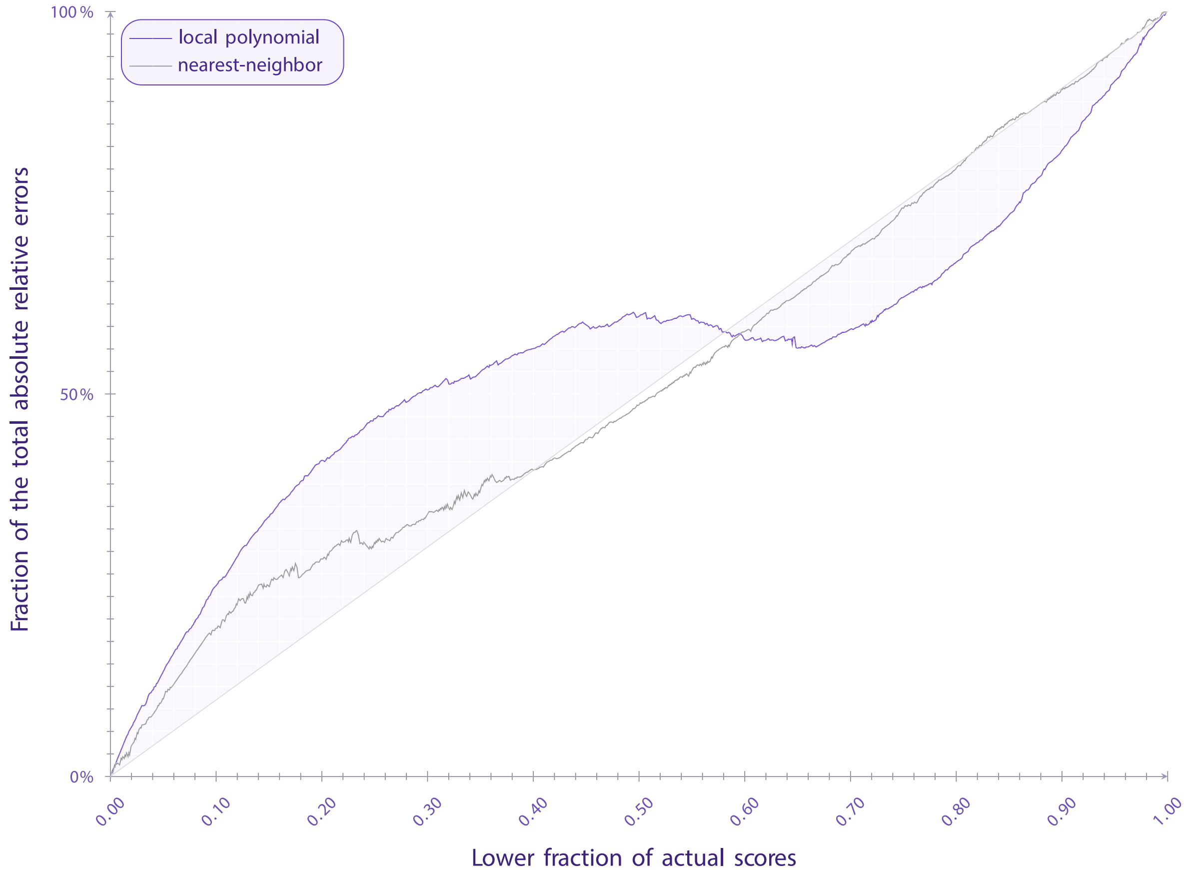

Fig. C.12. Inequality curves L_2^* for relative error distributions.

corresponding errors. On the contrary, standard inequality measures examine the distribution of a single variable whose data are arranged in a monotonic order. Our work illustrates the potential utility of the new indices developed as alternatives to commonly used inequality measures. We recommend widespread use of the novel tools presented in the paper on the assessment and ranking of prediction methods.

References

AbdelRahman, M.A.E., Zakarya, Y.M., Metwaly, M.M., Koubouris, G., 2021. Deciphering soil spatial variability through geostatistics and interpolation techniques. Sustainability 13, 194. https://doi.org/10.3390/su13010194. URL: https://www.mdpi.com/2071-1050/13/1/194.

Abdulmanov, R., Miftakhov, I., Ishbulatov, M., Galeev, E., Shafeeva, E., 2021. Comparison of the effectiveness of GIS-based interpolation methods for estimating the spatial distribution of agrochemical soil properties. Environ. Technol. Innovat. 24, 101970 https://doi.org/10.1016/j.eti.2021.101970. URL: https://www.sciencedirect.com/science/article/pii/S2352186421006180.

Ananias, D.R.S., Liska, G.R., Beijo, L.A., Liska, G.J.R., de Menezes, F.S., 2021. The assessment of annual rainfall field by applying different interpolation methods in the state of Rio Grande do Sul, Brazil. SN. Appl. Sci. 3, 687. https://doi.org/10.1007/s42452-021-04679-1. URL: https://link.springer.com/article/10.1007/s42452-021-04679-1.

Anselin, L., 2006. How (not) to lie with spatial statistics. Am. J. Prev. Med. 30, S3–S6. https://doi.org/10.1016/j.amepre.2005.09.015.

Anselin, L., 2013. From SpaceStat to CyberGIS: twenty years of spatial data analysis software., in: Garrocho, C. (Ed.), Advances in Commercial Geography: Prospects, Methods and Applications. El Colegio Mexiquense, A.C., Zinacantepec 51350, México. chapter 1, pp. 143–174. doi: 10.1177/0160017612438615.

Anselin, L., Le Gallo, J., 2006. Interpolation of air quality measures in hedonic house price models: spatial aspects. Spatial Economic Analysis 1, 31–52. URL: https://www.tandfonline.com/doi/full/10.1080/17421770600661337, doi: 10.1080/17421770600661337.

Armstrong, J.S., 2001. Evaluating forecasting methods, in: Armstrong, J.S. (Ed.), Principles of Forecasting: A Handbook for Researchers and Practitioners. Springer US, Boston, MA. volume 30 of International Series in Operations Research & Management Science, pp. 443–472. URL: https://link.springer.com/book/10.1007/978-0-306-47630-3, doi: 10.1007/978-0-306-47630-3_20.

Arnold, B.C., Sarabia Alegria, J.M., 2018. Majorization and the Lorenz Order with Applications in Applied Mathematics and Economics. 1 ed., Springer International Publishing, New York. URL:https://doi.org/10.1007/978-3-319-93773-1, doi: 10.1007/978-3-319-93773-1.

Arslan, H., Turan, N.A., 2015. Estimation of spatial distribution of heavy metals in groundwater using interpolation methods and multivariate statistical techniques; its suitability for drinking and irrigation purposes in the Middle Black Sea Region of Turkey. Environ. Monit. Assess. 187, 516. https://doi.org/10.1007/s10661-015-4725-x. URL: http://link.springer.com/10.1007/s10661-015-4725-x.

Attaeian, B., Farokhzadeh, B., Akhzari, D., Artimani, M.M., Souri, M., 2015. Comparing interpolation methods for estimating spatial distribution of topsoil pH and EC. ECOPERSIA 3, 1145–1159. URL: https://ecopersia.modares.ac.ir/article-24-1642-en.html.

Ballabio, C., Lugato, E., Fernández-Ugalde, O., Orgiazzi, A., Jones, A., Borrelli, P., Montanarella, L., Panagos, P., 2019. Mapping LUCAS topsoil chemical properties at european scale using Gaussian process regression. Geoderma 355, 113912. https://doi.org/10.1016/j.geoderma.2019.113912. URL: https://www.sciencedirect.com/science/article/pii/S0016706119304768.

Ballabio, C., Panagos, P., Monatanarella, L., 2016. Mapping topsoil physical properties at European scale using the LUCAS database. Geoderma 261, 110–123. https://doi.org/10.1016/j.geoderma.2015.07.006. URL: https://www.sciencedirect.com/science/article/pii/S0016706115300173.

Barrena-González, J., Lavado Contador, J.F., Pulido Fernández, M., 2022. Mapping soil properties at a regional scale: assessing deterministic vs. geostatistical interpolation methods at different soil depths. Sustainability 14, 10049. https://doi.org/10.3390/su141610049. URL: https://www.mdpi.com/2071-1050/14/16/10049.

Bekele, A., Downer, R.G., Wolcott, M.C., Hudnall, W.H., Moore, S.H., 2003. Comparative evaluation of spatial prediction methods in a field experiment for mapping soil potassium. Soil Sci. 168, 15–28. https://doi.org/10.1097/00010694-200301000-00003. URL: https://journals.lww.com/soilsci/Fulltext/2003/01000/COMPARATIVE_EVALUATION_OF_SPATIAL_PREDICTION.3.aspx.

Bellocchi, G., Rivington, M., Donatelli, M., Matthews, K., 2010. Validation of biophysical models: issues and methodologies. A review. Agronomy Sustain. Develop. 30, 109–130. https://doi.org/10.1051/agro/2009001. URL: https://link.springer.com/article/10.1051/agro/2009001.

Benslama, A., Khanchoul, K., Benbrahim, F., Boubehziz, S., Chikhi, F., Navarro-Pedreño, J., 2020. Monitoring the variations of soil salinity in a palm grove in Southern Algeria. Sustainability (Switzerland) 12, 6117. https://doi.org/10.3390/su12156117. URL: https://www.mdpi.com/2071-1050/12/15/6117.

Bhunia, G.S., Shit, P.K., Maiti, R., 2018. Comparison of GIS-based interpolation methods for spatial distribution of soil organic carbon (SOC). J. Saudi Soc. Agricul. Sci. 17, 114–126. https://doi.org/10.1016/j.jssas.2016.02.001. URL: https://www.sciencedirect.com/science/article/pii/S1658077X15300825.

Bishop, T.F., McBratney, A.B., 2001. A comparison of prediction methods for the creation of field-extent soil property maps. Geoderma 103, 149–160. https://doi.org/10.1016/S0016-7061(01)00074-X. URL: https://www.sciencedirect.com/science/article/pii/S001670610100074X.

Bogunovic, I., Filipovic, L., Filipovic, V., Pereira, P., 2021. Spatial mapping of soil chemical properties using multivariate geostatistics. A study from cropland in eastern Croatia. J. Central Europ. Agricult. 22, 201–210. https://doi.org/10.5513/JCEA01/22.1.3011.

Botchkarev, A., 2019. A new typology design of performance metrics to measure errors in machine learning regression algorithms. Interdisc. J. Inform., Knowl., Manage. 14, 45–76. https://doi.org/10.28945/4184. URL: https://www.informingscience.org/Publications/4184.

Bowley, A.L., 1920. Elements of Statistics. 4 ed., Charles Scribner's Sons, New York. 1st edn 1901.

Buhmann, M.D., 2003. Radial Basis Functions: Theory and Implementations. Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge. URL: https://doi.org/10.1017/CBO9780511543241, doi: 10.1017/CBO9780511543241.

Chai, T., Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature. Geosci. Model Develop. 7, 1247–1250. https://doi.org/10.5194/GMD-7-1247-2014. URL: https://gmd.copernicus.org/articles/7/1247/2014/.

Chen, T., Chang, Q., Liu, J., Clevers, J.G., Kooistra, L., 2016. Identification of soil heavy metal sources and improvement in spatial mapping based on soil spectral information: A case study in northwest China. Sci. Total Environ. 565, 155–164. https://doi.org/10.1016/j.scitotenv.2016.04.163. URL: https://www.sciencedirect.com/science/article/pii/S0048969716308658.

Csiszár, I., 1963. Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizitat von Markoffschen Ketten. A Magyar Tudományos Akadémia. Matematikai Kutató Intézetének Közleményei 8, 85–108.

David, F.N., Johnson, N.L., 1956. Some tests of significance with ordered variables. J. R. Stat. Soc. Ser. B (Methodological) 18, 1–31. https://doi.org/10.1111/j.2517-6161.1956.tb00205.x. URL: https://rss.onlinelibrary.wiley.com/doi/10.1111/j.2517-6161.1956.tb00205.x.

Deepika, B.V., Ramakrishnaiah, C.R., Naganna, S.R., 2020. Spatial variability of ground water quality: a case study of Udupi district, Karnataka State, India. J. Earth Syst. Sci. 129, 221. https://doi.org/10.1007/s12040-020-01471-4. URL: https://link.springer.com/article/10.1007%2Fs12040-020-01471-4.

Durdević, B., Jug, I., Jug, D., Bogunović, I., Vukadinović, V., Stipešević, B., Brozović, B., 2019. Spatial variability of soil organic matter content in Eastern Croatia assessed using different interpolation methods. Int. Agrophys. 33, 31–39. https://doi.org/10.31545/intagr/104372.

Duveiller, G., Fasbender, D., Meroni, M., 2016. Revisiting the concept of a symmetric index of agreement for continuous datasets. Sci. Rep. 6, 19401. https://doi.org/10.1038/srep19401. URL: https://www.nature.com/articles/srep19401.

Eker, S., Rovenskaya, E., Langan, S., Obersteiner, M., 2019. Model validation: A bibliometric analysis of the literature. Environ. Modell. Software 117, 43–54. https://doi.org/10.1016/j.envsoft.2019.03.009. URL: https://www.sciencedirect.com/science/article/pii/S1364815218312763.

Emadi, M., Baghernejad, M., 2014. Comparison of spatial interpolation techniques for mapping soil pH and salinity in agricultural coastal areas, northern Iran. Arch. Agronomy Soil Sci. 60, 1315–1327. https://doi.org/10.1080/03650340.2014.880837. URL: https://www.tandfonline.com/doi/abs/10.1080/03650340.2014.880837.

European Commission, Centre, J.R., Jones, A., Fernández-Ugalde, O., Scarpa, S., 2020. LUCAS 2015 topsoil survey: presentation of dataset and results. Publications Office. URL: https://data.europa.eu/doi/10.2760/616084, doi: doi/10.2760/616084.

Fernández-Ugalde, O., Jones, A., Meuli, R.G., 2020. Comparison of sampling with a spade and gouge auger for topsoil monitoring at the continental scale. Eur. J. Soil Sci. 71, 137–150. https://doi.org/10.1111/ejss.12862.

Furuichi, S., 2010. An axiomatic characterization of a two-parameter extended relative entropy. J. Math. Phys. 51, 123302 https://doi.org/10.1063/1.3525917.

Gastwirth, J.L., 2016. Measures of economic inequality focusing on the status of the lower and middle income groups. Stat. Public Policy 3, 1–9. https://doi.org/10.1080/2330443X.2016.1213148.

Gini, C., 1921. Measurement of inequality of incomes. Econ. J. 31, 124–126. https://doi.org/10.2307/2223319. URL: https://www.jstor.org/stable/2223319.

Gong, G., Mattevada, S., O'Bryant, S.E., 2014. Comparison of the accuracy of kriging and IDW interpolations in estimating groundwater arsenic concentrations in Texas. Environ. Res. 130, 59–69. https://doi.org/10.1016/j.envres.2013.12.005. URL: https://www.sciencedirect.com/science/article/pii/S0013935113002065.

Gotway, C.A., Ferguson, R.B., Hergert, G.W., Peterson, T.A., 1996. Comparison of kriging and inverse-distance methods for mapping soil parameters. Soil Sci. Soc. Am. J. 60, 1237–1247. https://doi.org/10.2136/sssaj1996.03615995006000040040x. URL: https://acsess.onlinelibrary.wiley.com/doi/abs/10.2136/sssaj1996.03615995006000040040x.

Hinkley, D.V., 1975. On power transformations to symmetry. Biometrika 62, 101–111. https://doi.org/10.2307/2334491. URL: https://academic.oup.com/biomet/article-abstract/62/1/101/220136.

Hodson, T.O., 2022. Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. Geosci. Model Develop. 15, 5481–5487. https://doi.org/10.5194/gmd-15-5481-2022. URL: https://gmd.copernicus.org/articles/15/5481/2022/.

Hogg, R.V., 1974. Adaptive robust procedures: A partial review and some suggestions for future applications and theory. J. Am. Stat. Assoc. 69, 909–923. https://doi.org/10.1080/01621459.1974.10480225. URL: https://www.tandfonline.com/doi/abs/10.1080/01621459.1974.10480225.

Hosseini, E., Gallichand, J., Marcotte, D., 1994. Theoretical and experimental performance of spatial interpolation methods for soil salinity analysis. Trans. ASAE 37, 1799–1807. https://doi.org/10.13031/2013.28269.

Hyndman, R.J., Fan, Y., 1996. Sample quantiles in statistical packages. Am. Statist. 50, 361–365. https://doi.org/10.2307/2684934. URL: http://www.jstor.org/stable/2684934.

Ijaz, Z., Zhao, C., Ijaz, N., ur Rehman, Z., Ijaz, A., 2023. Development and optimization of geotechnical soil maps using various geostatistical and spatial interpolation techniques: a comprehensive study. Bull. Eng. Geol. Environ. 82, 215. URL: https://link.springer.com/article/10.1007/s10064-023-03244-x, doi: 10.1007/s10064-023-03244-x.

Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics, 1 ed. Oxford University Press, New York.

Joanes, D.N., Gill, C.A., 1998. Comparing measures of sample skewness and kurtosis. J. R. Stat. Soc.: Ser. D (The Statistician) 47, 183–189. https://doi.org/10.1111/1467-9884.00122. URL: https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-9884.00122.

Kaniadakis, G., Lissia, M., Scarfone, A.M., 2004. Deformed logarithms and entropies. Physica A 340, 41–49. https://doi.org/10.1016/j.physa.2004.03.075. URL: https://www.sciencedirect.com/science/article/abs/pii/S0378437104003954.

Kaniadakis, G., Lissia, M., Scarfone, A.M., 2005. Two-parameter deformations of logarithm, exponential, and entropy: A consistent framework for generalized statistical mechanics. Phys. Rev. E 71, 046128. https://doi.org/10.1103/PhysRevE.71.046128. URL: https://journals.aps.org/pre/abstract/10.1103/PhysRevE.71.046128.

Karunasingha, D.S.K., 2022. Root mean square error or mean absolute error? use their ratio as well. Inf. Sci. 585, 609–629. https://doi.org/10.1016/j.ins.2021.11.036. URL: https://www.sciencedirect.com/science/article/pii/S0020025521011567.

Kassambara, A., Mundt, F., 2020. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. URL: https://CRAN.R-project.org/package=factoextra. r package version 1.0.7.

Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Adv. Geosci. 5, 89–97. URL: https://doi.org/10.5194/adgeo-5-89-2005.

Kravchenko, A.N., Bullock, D.G., 1999. A comparative study of interpolation methods for mapping soil properties. Agronomy J. 91, 393–400. https://doi.org/10.2134/agronj1999.00021962009100030007x. URL: https://acsess.onlinelibrary.wiley.com/doi/abs/10.2134/agronj1999.00021962009100030007x.

Lai, Y.Q., Wang, H.L., Sun, X.L., 2021. A comparison of importance of modelling method and sample size for mapping soil organic matter in Guangdong, China. Ecol. Ind. 126, 107618 https://doi.org/10.1016/j.ecolind.2021.107618. URL: https://www.sciencedirect.com/science/article/pii/S1470160X21002831.

Laslett, G.M., McBratney, A.B., 1990. Further comparison of spatial methods for predicting soil pH. Soil Sci. Soc. Am. J. 54, 1553–1558. https://doi.org/10.2136/sssaj1990.03615995005400060007x. URL: https://acsess.onlinelibrary.wiley.com/doi/abs/10.2136/sssaj1990.03615995005400060007x.

Laslett, G.M., McBratney, A.B., Pahl, P.J., Hutchinson, M.F., 1987. Comparison of several spatial prediction methods for soil pH. Eur. J. Soil Sci. 38, 325–341. https://doi.org/10.1111/j.1365-2389.1987.tb02148.x. URL: https://onlinelibrary.wiley.com/doi/10.1111/j.1365-2389.1987.tb02148.x.

Leenaers, H., Okx, J.P., Burrough, P.A., 1990. Comparison of spatial prediction methods for mapping floodplain soil pollution. CATENA 17, 535–550. https://doi.org/10.1016/0341-8162(90)90028-C. URL: https://www.sciencedirect.com/science/article/pii/034181629090028C.

Legates, D.R., McCabe Jr., G.J., 1999. Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. Surface Water Climate 35, 233–241. https://doi.org/10.1029/1998WR900018. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1998WR900018.

Legates, D.R., McCabe Jr., G.J., 2013. A refined index of model performance: a rejoinder. Int. J. Climatol. 33, 1053–1056. https://doi.org/10.1002/joc.3487. URL: https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.3487.

Li, J., 2017. Assessing the accuracy of predictive models for numerical data: Not r nor r^2, why not? Then what? Plos ONE 12, e0183250. URL: https://doi.org/10.1371/journal.pone.0183250.

Li, J., Heap, A.D., 2011. A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. Ecological Informatics 6, 228–241. https://doi.org/10.1016/j.ecoinf.2010.12.003. URL: https://www.sciencedirect.com/science/article/pii/S1574954110001147.

Li, J., Heap, A.D., 2014. Spatial interpolation methods applied in the environmental sciences: A review. Environ. Modell. Software 53, 173–189. https://doi.org/10.1016/j.envsoft.2013.12.008. URL: https://www.sciencedirect.com/science/article/pii/S1364815213003113.

Liemohn, M.W., Shane, A.D., Azari, A.R., Petersen, A.K., Swiger, B.M., Mukhopadhyay, A., 2021. RMSE is not enough: Guidelines to robust data-model comparisons for magnetospheric physics. J. Atmos. Solar Terr. Phys. 218, 105624

https://doi.org/10.1016/J.JASTP.2021.105624. URL: https://www.sciencedirect.com/science/article/pii/S1364682621000857.
Loague, K., Green, R.E., 1991. Statistical and graphical methods for evaluating solute transport models: Overview and application. J. Contam. Hydrol. 7, 51–73. https://doi.org/10.1016/0169-7722(91)90038-3. URL: https://www.sciencedirect.com/science/article/pii/0169772291900383.
Long, J., Liu, Y., Xing, S., Zhang, L., Qu, M., Qiu, L., Huang, Q., Zhou, B., Shen, J., 2020. Optimal interpolation methods for farmland soil organic matter in various landforms of a complex topography. Ecol. Ind. 110, 105926 https://doi.org/10.1016/j.ecolind.2019.105926. URL: https://www.sciencedirect.com/science/article/pii/S1470160X19309215.
Lorenz, M.O., 1905. Methods of measuring the concentration of wealth. Publications of the American Statistical Association 9, 209–219. https://doi.org/10.2307/2276207. URL: http://www.jstor.org/stable/2276207.
Ma, H.H., Yu, T., Yang, Z.F., Hou, Q.Y., Zeng, Q.L., Wang, R., 2018. Spatial interpolation methods and pollution assessment of heavy metals of soil in typical areas. Huanjing Kexue/Environ. Sci. 39, 4684–4693. https://doi.org/10.13227/j.hjkx.201712185.
MacGillivray, H.L., 1986. Skewness and asymmetry: measures and orderings. Ann. Stat. 14, 994–1011. https://doi.org/10.1214/aos/1176350046.
Magurran, A., 2004. Measuring Biological Diversity. Blackwell Publishing, Oxford.
McCuen, R.H., Snyder, W.M., 1975. A proposed index for comparing hydrographs. Water Resour. Res. 11, 1021–1024. URL: https://doi.org/10.1029/WR011i006p01021.
Meng, Q., Liu, Z., Borders, B.E., 2013. Assessment of regression kriging for spatial interpolation - comparisons of seven GIS interpolation methods. Cartogr. Geograph. Inform. Sci. 40, 28–39. https://doi.org/10.1080/15230406.2013.762138.
Mielke, P.W., 1984. Meteorological applications of permutation techniques based on distance functions, in: Sen, P.K.&P. (Ed.), Handbook of Statistics. North-Holland: Elsevier Science Publishers, Amsterdam. volume 4, pp. 813–830.
Mittal, D.P., 1975. On some functional equations concerning entropy, directed divergence and inaccuracy. Metrika 22, 35–46. https://doi.org/10.1007/BF01899712.
Morimoto, T., 1963. Markov processes and the H-theorem. J. Phys. Soc. Jpn. 18, 328–331. https://doi.org/10.1143/JPSJ.18.328.
Mousavi, S.R., Sarmadian, F., Dehghani, S., Sadikhani, M.R., Taati, A., 2017. Evaluating inverse distance weighting and kriging methods in estimation of some physical and chemical properties of soil in Qazvin Plain. Eurasian. Soil Sci. 6, 327–336. https://doi.org/10.18393/ejss.311210.
Mueller, T.G., Pusuluri, N.B., Mathias, K.K., Cornelius, P.L., Barnhisel, R.I., Shearer, S.A., 2004. Map quality for ordinary kriging and inverse distance weighted interpolation. Soil Sci. Soc. Am. J. 68, 2042–2047. https://doi.org/10.2136/sssaj2004.2042. URL: https://acsess.onlinelibrary.wiley.com/doi/full/10.2136/sssaj2004.2042.
Myslyva, T., Kutsaeva, O., Krundzikava, N., 2019. Efficiency of interpolation methods based on GIS for for estimating of spatial distribution of pH in soil. Baltic Surveying 11, 53–59. https://doi.org/10.22616/j.balticsurveying.2019.017. URL: https://llufb.llu.lv/Raksti/Journal_Baltic_Surveying/2019/Journal_Baltic_Surveying_Vol_11_2019.pdf.
Nas, B., Berktay, A., 2010. Groundwater quality mapping in urban groundwater using GIS. Environ. Monit. Assess. 160, 215–227. https://doi.org/10.1007/s10661-008-0689-4. URL: http://link.springer.com/10.1007/s10661-008-0689-4.
Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I – A discussion of principles. J. Hydrol. 10, 282–290. https://doi.org/10.1016/0022-1694(70)90255-6. URL: https://www.sciencedirect.com/science/article/pii/0022169470902556.
Orgiazzi, A., Ballabio, C., Panagos, P., Jones, A., Fernández-Ugalde, O., 2018. LUCAS Soil, the largest expandable soil dataset for Europe: a review. Eur. J. Soil Sci. 69, 140–153. https://doi.org/10.1111/ejss.12499.
Ouabo, R.E., Sangodoyin, A.Y., Ogundiran, M.B., 2020. Assessment of ordinary kriging and inverse distance weighting methods for modeling chromium and cadmium soil pollution in e-waste sites in Douala, Cameroon. J. Health Pollution 10, 200605. https://doi.org/10.5696/2156-9614-10.26.200605. URL: https://meridian.allenpress.com/jhp/article/10/26/200605/445417/Assessment-of-Ordinary-Kriging-and-Inverse.
Ozelkan, E., Chen, G., Ustundag, B.B., 2016. Spatial estimation of wind speed: a new integrative model using inverse distance weighting and power law. Int. J. Digital Earth 9, 733–747. https://doi.org/10.1080/17538947.2015.1127437. URL: https://www.tandfonline.com/doi/abs/10.1080/17538947.2015.1127437.
Panagopoulos, T., Jesus, J., Antunes, M.D.C., Beltrão, J.G.T., 2006. Analysis of spatial interpolation for optimising management of a salinized field cultivated with lettuce. Eur. J. Agron. 24, 1–10. https://doi.org/10.1016/j.eja.2005.03.001. URL: https://www.sciencedirect.com/science/article/pii/S1161030105000389.
Prendergast, L.A., Staudte, R.G., 2016. Quantile versions of the Lorenz curve. Electron. J. Statist. 10, 1896–1926. https://doi.org/10.1214/16-EJS1154.
Qiao, P., Lei, M., Yang, S., Yang, J., Guo, G., Zhou, X., 2018. Comparing ordinary kriging and inverse distance weighting for soil as pollution in Beijing. Environ. Sci. Pollut. Res. 25, 15597–15608. https://doi.org/10.1007/s11356-018-1552-y. URL: https://link.springer.com/article/10.1007%2Fs11356-018-1552-y.
R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.
Radočaj, D., Jurišić, M., Antonić, O., 2021. Determination of soil C: N suitability zones for organic farming using an unsupervised classification in eastern Croatia. Ecol. Ind. 123, 107382 https://doi.org/10.1016/j.ecolind.2021.107382. URL: https://www.sciencedirect.com/science/article/pii/S1470160X21000479.
Reza, S., Sarkar, D., Baruah, U., Das, T.H., 2010. Evaluation and comparison of ordinary kriging and inverse distance weighting methods for prediction of spatial variability of some chemical parameters of Dhalai district, Tripura. Agropedology 20, 38–48.
Robinson, T.P., Metternicht, G.I., 2006. Testing the performance of spatial interpolation techniques for mapping soil properties. Comput. Electron. Agricult. 50, 97–108. https://doi.org/10.1016/j.compag.2005.07.003. URL: https://www.sciencedirect.com/science/article/pii/S0168169905001523.
Saha, A., Gupta, B.S., Patidar, S., Martínez-Villegas, N., 2022. Spatial distribution based on optimal interpolation techniques and assessment of contamination risk for toxic metals in the surface soil. J. S. Am. Earth Sci. 115, 103763 https://doi.org/10.1016/j.jsames.2022.103763. URL: https://www.sciencedirect.com/science/article/pii/S0895981122000542.
Scarfone, A.M., 2006. Legendre structure of the thermostatistics theory based on the Sharma-Taneja-Mittal entropy. Physica A 365, 63–70. https://doi.org/10.1016/j.physa.2006.01.023. URL: https://www.sciencedirect.com/science/article/pii/S0378437106000677.
Schloeder, C.A., Zimmermann, N.E., Jacobs, M., 2001. Comparison of methods for interpolating soil properties using limited data. Soil Sci. Soc. Am. J. 65, 470–479. https://doi.org/10.2136/sssaj2001.652470x. URL: https://acsess.onlinelibrary.wiley.com/doi/abs/10.2136/sssaj2001.652470x.
Sharma, B.D., Taneja, I.J., 1975. Entropy of type (α, β) and other generalized measures in information theory. Metrika 22, 205–215. https://doi.org/10.1007/BF01899728. URL: https://link.springer.com/article/10.1007%2FBF01899728.
Shen, Q., Wang, Y., Wang, X., Liu, X., Zhang, X., Zhang, S., 2019. Comparing interpolation methods to predict soil total phosphorus in the Mollisol area of Northeast China. CATENA 174, 59–72. https://doi.org/10.1016/j.catena.2018.10.052. URL: https://www.sciencedirect.com/science/article/pii/S0341816218304879.
Shi, W., Liu, J., Du, Z., Song, Y., Chen, C., Yue, T., 2009. Surface modelling of soil pH. Geoderma 150, 113–119. https://doi.org/10.1016/j.geoderma.2009.01.020. URL: https://www.sciencedirect.com/science/article/pii/S0016706109000287.
Staudte, R.G., 2014. Inference for quantile measures of skewness. TEST 23, 751–768. https://doi.org/10.1007/s11749-014-0391-5. URL: https://link.springer.com/article/10.1007/s11749-014-0391-5.
Tiruneh, G.A., Alemayehu, T.Y., Meshesha, D.T., Adgo, E., Tiruneh, G.A., Reichert, J.M., 2023. Variability modeling and mapping of soil properties for improved management in Ethiopia. Agrosyst., Geosci. Environ. 6, e20357 https://doi.org/10.1002/agg2.20357. URL: https://acsess.onlinelibrary.wiley.com/doi/abs/10.1002/agg2.20357.
Tunçay, T., 2021. Comparison quality of interpolation methods to estimate spatial distribution of soil moisture content. Commun. Soil Sci. Plant Anal. 52, 353–374. https://doi.org/10.1080/00103624.2020.1854283. URL: https://www.tandfonline.com/doi/abs/10.1080/00103624.2020.1854283.
Voltz, M., Webster, R., 1990. A comparison of kriging, cubic splines and classification for predicting soil properties from sample information. J. Soil Sci. 41, 473–490. https://doi.org/10.1111/j.1365-2389.1990.tb00080.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2389.1990.tb00080.x.
Wackernagel, H., 2003. Multivariate Geostatistics, 3 ed.,. Springer-Verlag, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-05294-5. URL: https://link.springer.com/book/10.1007/978-3-662-05294-5.
Watterson, I.G., 1996. Non-dimensional measures of climate model performance. Int. J. Climatol. 16, 379–391. URL: https://doi.org/10.1002/(SICI)1097-0088(199604)16:4<379::AID-JOC18>3.0.CO;2-U.
Webster, R., Oliver, M.A., 2007. Geostatistics for Environmental Scientists, 2 ed.,. John Wiley & Sons, Chichester. https://doi.org/10.1002/9780470517277.
Willmott, C.J., 1981. On the validation of models. Phys. Geogr. 2, 184–194. https://doi.org/10.1080/02723646.1981.10642213. URL: https://www.tandfonline.com/doi/abs/10.1080/02723646.1981.10642213.
Willmott, C.J., 1982. Some comments on the evaluation of model performance. Bull. Am. Meteorol. Soc. 63, 1309–1313. URL: https://www.jstor.org/stable/26222954, doi: 10.1175/1520-0477(1982)063<1309:SCOTEO>2.0.CO;2.
Willmott, C.J., Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J., Rowe, C.M., 1985. Statistics for the evaluation and comparison of models. J. Geophys. Res.: Oceans 90, 8995–9005. https://doi.org/10.1029/JC090iC05p08995.
Willmott, C.J., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim. Res. 30, 79–82. https://doi.org/10.3354/cr030079. URL: http://www.int-res.com/abstracts/cr/v30/n1/p79-82/.
Willmott, C.J., Matsuura, K., Robeson, S.M., 2009. Ambiguities inherent in sums-of-squares-based error statistics. Atmos. Environ. 43, 749–752. https://doi.org/10.1016/j.atmosenv.2008.10.005. URL: https://www.sciencedirect.com/science/article/pii/S1352231008009564.
Willmott, C.J., Robeson, S.M., Matsuura, K., 2012. A refined index of model performance. Int. J. Climatol. 32, 2088–2094. https://doi.org/10.1002/joc.2419.
Willmott, C.J., Robeson, S.M., Matsuura, K., Ficklin, D.L., 2015. Assessment of three dimensionless measures of model performance. Environ. Modell. Software 73, 167–174. https://doi.org/10.1016/j.envsoft.2015.08.012. URL: https://www.sciencedirect.com/science/article/pii/S1364815215300384.
Willmott, C.J., Wicks, D.E., 1980. An empirical method for the spatial interpolation of monthly precipitation within California. Phys. Geogr. 1, 59–73. https://doi.org/10.1080/02723646.1980.10642189. URL: https://doi.org/10.1080/02723646.1980.10642189.
Xie, Y., bin Chen, T., Lei, M., Yang, J., jun Guo, Q., Song, B., yong Zhou, X., 2011. Spatial distribution of soil heavy metal pollution estimated by different interpolation methods: Accuracy uncertainty analysis. Chemosphere 82, 468–476. URL: https://www.sciencedirect.com/science/article/pii/S0045653510010805, doi: 10.1016/j.chemosphere.2010.09.053.

Yang, R., Xing, B., 2021. A comparison of the performance of different interpolation methods in replicating rainfall magnitudes under different climatic conditions in Chongqing province (China). Atmosphere 12, 1318. https://doi.org/10.3390/atmos12101318. URL: https://www.mdpi.com/2073-4433/12/10/1318.

Yasrebi, J., Saffari, M., Fathi, H., Karimian, N., Moazallahi, M., Gazni, R., 2009. Evaluation and comparison of ordinary kriging and inverse distance weighting methods for prediction of spatial variability of some soil chemical parameters. Res. J. Biolog. Sci. 4, 93–102. URL: https://www.medwelljournals.com/abstract/?doi=rjbsci.2009.93.102, doi: rjbsci.2009.93.102.

Yule, G.U., 1911. An Introduction to the Theory of Statistics, 1 ed. Charles Griffin & Company, London.

Zare-Mehrjardi, M., Taghizadeh-Mehrjardi, R., Akbarzadeh, A., 2010. Evaluation of geostatistical techniques for mapping spatial distribution of soil PH, salinity and plant cover affected by environmental factors in Southern Iran. Notulae Scientia Biologicae 2, 92–103. https://doi.org/10.15835/nsb244997.

Zeileis, A., Kleiber, C., 2014. ineq: Measuring Inequality, Concentration, and Poverty. URL: https://CRAN.R-project.org/package=ineq. r package version 0.2-13.

Zhao, W., Cao, T., Li, Z., Sheng, J., 2019. Comparison of IDW, cokriging and ARMA for predicting spatiotemporal variability of soil salinity in a gravel-sand mulched jujube orchard. Environ. Monit. Assess. 191, 376. https://doi.org/10.1007/s10661-019-7499-8. URL: https://link.springer.com/article/10.1007/s10661-019-7499-8.

Zhao, Y., Xu, X., Tian, K., Huang, B., Hai, N., 2016. Comparison of sampling schemes for the spatial prediction of soil organic matter in a typical black soil region in China. Environ. Earth Sci. 75, 4. https://doi.org/10.1007/s12665-015-4895-4. URL: https://link.springer.com/article/10.1007/s12665-015-4895-4.