

Article

Not peer-reviewed version

A Non-linear Trend Function for Kriging with External Drift Using Least Squares Support Vector Regression

Kanokrat Baisad , [Nawinda Chutsagulprom](#) , [Sompop Moonchai](#) *

Posted Date: 31 October 2023

doi: [10.20944/preprints202310.1993.v1](https://doi.org/10.20944/preprints202310.1993.v1)

Keywords: geostatistics; spatial interpolation; kriging with external drift; least squares support vector regression; trend function



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

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.

Article

A Non-Linear Trend Function for Kriging with External Drift Using Least Squares Support Vector Regression

Kanokrat Baisad, Nawinda Chutsagulprom and Sompop Moonchai *

Department of Mathematics, Faculty of Science, Chiang Mai University, Chiang Mai, 50200, Thailand

* Correspondence: sompop.m@cmu.ac.th

Abstract: Spatial interpolation of meteorological data can have immense implications on risk management and climate change planning. Kriging with external drift (KED) is a spatial interpolation variant that uses auxiliary information in the estimation of target variable at unobserved locations. However, the traditional KED methods with linear trend functions may not be able to capture the complex and non-linear interdependence between target and auxiliary variables, which can lead to an inaccurate estimation. In this work, a novel KED method using least squares support vector regression (LSSVR) is proposed. This machine learning algorithm is employed to construct trend functions regardless of the type of variable interrelations being considered. To evaluate the efficiency of the proposed method (KED with LSSVR) relative to the traditional method (KED with a linear trend function), a systematic simulation study for estimating the monthly means temperature and pressure in Thailand in 2017 was conducted. The KED with LSSVR is shown to have superior performance over the KED with the linear trend function.

Keywords: geostatistics; spatial interpolation; kriging with external drift; least squares support vector regression; trend function

1. Introduction

Spatial interpolation is a fundamental technique employed in spatial data analysis to estimate the variable of interest at unobserved locations based on available data. Kriging is a geostatistical approach for spatial interpolation that provides a best linear unbiased prediction with the minimum estimation variance. An approximation of the kriging models relies on the assumption that a random process can be decomposed into a trend function and a random residual component. Ordinary kriging (OK) is a commonly used method with a constant trend, which is not suitable in the presence of a strong trend structure. On the other hand, kriging with external drift (KED) allows for the inclusion of auxiliary variables that have a strong spatial correlation with the target variable in order to increase precision in estimates [1,2]. In the KED, a trend model that is fitted to both target data points and significant auxiliary samples is first generated. The empirical variogram is thereafter derived from residuals computed from the difference between the trend estimates and measured values. A final prediction of the target variable is obtained as a weighted linear combination of observations, in which the weights are calculated through the Lagrange multiplier method. The KED method has been applied in various fields, including meteorology [3–6], geology [7–10], environmental modeling [11–13], agronomy [14–16], and hydrology [17].

The trend term for the KED is conventionally modelled by polynomial functions of degree one or two. However, in practice, a non-linear relationship often exists between influence factors and response variables where the application of such a polynomial is not adequate. Despite extensive research on the prediction using KED having been conducted, a study of non-linear trend functions for the KED remains scarce. Snepvangers et al. [18] developed a non-linear trend represented by a logarithmic function to interpolate soil water content using the KED technique with net-precipitation as an auxiliary variable. Freier and Lieres [19] introduced a novel extension to universal kriging (UK),

a specific instance of KED, aimed at handling non-linear trend patterns. They utilized a Taylor-based linearization approach in conjunction with an iterative parameter estimation procedure to construct a non-linear trend model. The method was applied to the Michaelis-Menten equation, which describes an enzymatic reaction. Freier et al. [20] subsequently employed this kriging technique to interpolate biocatalytic data with low and irregular density. Their method can be particularly of use in the presence of an explicit expression of the non-linear trend functions. Nevertheless, an interaction between design factors and system response in real-world applications is naturally described by diverse and complex behaviour, which is difficult to establish in an explicit form.

Machine learning (ML) has recently been gaining attention as a computationally efficient tool for identifying implicit relationships between variables. This allows one to generate and optimize the complex model based on the huge amount of data available for analysis. Support vector machine (SVM) is a kernel-based machine learning approach used for classification and regression. The particular use of SVMs for regression problems is called support vector regression (SVR), which was first introduced by Vapnik in 1992 [21]. The method adopts the structural risk minimization principle by minimizing the upper bound of the generalization errors. This leads to a linear decision function, which is essentially a convex quadratic programming (QP) problem. The core element of SVR is to search for the optimal hyperplane that fits the learning data while maximizing the distance between the hyperplane and the data points. In the case of non-linear problems, the SVR procedure starts by projecting input data into high-dimensional feature space through some non-linear mapping, and the SVR subsequently performs linear regression to obtain the optimal hyperplane. Apart from producing high prediction accuracy for non-linear data, SVR is also suitable for applications characterized by small datasets [22–24]. Furthermore, this SVM regression-based algorithm has the generalization ability to reduce overfitting issues by introducing regularization term into the loss function. Due to all these advantages, the technique has therefore been applied in diverse disciplines, including finance [25,26], economics [27,28], climate modelling [29,30], and healthcare [31,32]. However, SVR requires substantial computational time and significant memory usage to solve the QP problem. To overcome these limitations, Suykens and Vandewalle [33] proposed a variant of SVR known as least squares support vector regression (LSSVR). This method extends the traditional SVR by using a squared loss function rather than quadratic programming. The LSSVR results yield higher accuracy and require less computational resources compared to the reliability method relying on SVR [34].

There is a notable absence of research on the utilization of ML in geostatistical techniques. In this work, we will present a novel interpolation method in which the LSSVR method is used to compute non-linear trend functions within the context of KED. The proposed technique entails expressing the trend function in a structured form through explicit feature mapping. The purpose of our technique is to enhance the predictive capabilities of the KED model by incorporating the powerful capabilities of LSSVR for capturing non-linear relationships between variables.

The remainder of this paper is outlined as follows. Section 2 reviews the theory regarding the KED methodology and LSSVR technique. A detailed description of the KED using the LSSVR for modelling the non-linear trend functions is provided in Section 3. In Section 4, we conduct a comparative simulation study using the conventional KED model and the proposed method for temperature and pressure estimation in Thailand. Conclusion and discussion are drawn in Section 5.

2. Mathematical Background

2.1. Kriging with External Drift

Kriging is a spatial interpolation method that uses variogram analysis to predict the variable of interest at an unmeasured location based on the values of surrounding measured locations. It is the best linear unbiased estimator (BLUE) for the random function, $\{Z(s) : s \in D \subset \mathbb{R}^d\}$, where D is a

defined spatial domain and d a positive integer representing the number of dimensions in the spatial domain. The value of $Z(s)$ can be obtained through

$$Z(s) = \mu(s) + \epsilon(s), \quad (1)$$

where a deterministic component $\mu(s)$ indicates the underlying trend or drift and $\epsilon(s)$ is a stochastic residual component with a mean of zero and a variogram, which is a function of lag vector [2].

In KED, the trend is modelled by a function of auxiliary variables which can be expressed as

$$\mu(s) = \sum_{l=0}^L a_l f_l(s), \quad (2)$$

where $a_l \in \mathbb{R} \setminus \{0\}$ is coefficient to be estimated, $f_l(s)$ is prescribed function that maps from the domain D into the range \mathbb{R} , and $L + 1$ is the number of terms used in the approximation. Additionally, the function $f_0(s)$ is defined to be 1 for all s in D [2].

To determine the unknown coefficients in equation (2), we can use the ordinary least squares (OLS) estimator or its extension, the generalized least squares (GLS) estimator, which accounts for the spatial correlation between individual observations [35].

Given n observed values, $Z(s_1), \dots, Z(s_n)$, at sample points, s_1, s_2, \dots, s_n . The attribute $Z(s_0)$ at an ungauged site s_0 is estimated as a linear combination of observed values so that

$$Z^*(s_0) = \sum_{i=1}^n \omega_i Z(s_i), \quad (3)$$

where ω_i is the kriging weight assigned to $Z(s_i)$. The weights ω_i are computed by minimizing the estimation error variance subject to the unbiased constraint. This results in the following optimization problem:

$$\begin{aligned} & \text{minimum of} && \text{Var}[Z^*(s_0) - Z(s_0)], \\ & \text{subject to} && \mathbb{E}[Z^*(s_0) - Z(s_0)] = 0. \end{aligned} \quad (4)$$

The optimal weights of the system (4) for the KED model can be solved by using the Lagrange multiplier method which leads to

$$\left\{ \begin{array}{l} \sum_{j=1}^n \omega_j \gamma_\epsilon(s_i - s_j) + \sum_{l=0}^L \lambda_l f_l(s_i) = \gamma_\epsilon(s_i - s_0), \quad \forall i = 1, \dots, n, \\ \sum_{j=1}^n \omega_j f_l(s_i) = f_l(s_0), \quad \forall l = 0, 1, \dots, L, \end{array} \right. \quad (5)$$

where $\gamma_\epsilon(\cdot)$ denotes the residual variogram function of $Z(s)$ and $\lambda_l \in \mathbb{R}$ is a Lagrange multiplier.

The variogram is a fundamental and important tool that quantifies the spatial correlation structure of the sample points. The variogram model is a smooth function that is reasonably well fitted to the empirical variogram estimated from the data. In the present study, we use the empirical variogram estimator introduced by Matheron [36], and the parametric variogram is represented by an exponential model [35].

In general, both linear and quadratic functions are usually treated as a trend representation [9,17, 37,38]. However, in certain scenarios, the relationship between the target and auxiliary variables is too complex to be captured by simple polynomial functions. In this work, least squares support vector regression (LSSVR) is used to model a non-linear trend function within the KED framework.

2.2. Least Squares Support Vector Regression

Given a dataset $\{Y_i, Z_i\}_{i=1}^n$, where $Y_i \in \mathbb{R}^\eta$ is a η -dimensional training data point and $Z_i \in \mathbb{R}$ represents a target output. The objective of least squares support vector regression (LSSVR) is to find

a function that minimizes the square error between the predicted values and the actual values. In LSSVR, the input data Y_i are mapped into a higher-dimensional feature space \mathbb{R}^{η_h} , in which a linear model is adopted, so that a model function μ is formulated as

$$\mu(Y) = \mathbf{a}^T \phi(Y) + b, \quad (6)$$

where $\phi(Y)$ is an η_h -dimensional feature mapping, \mathbf{a} is an $\eta_h \times 1$ weight vector, and $b \in \mathbb{R}$ indicates a bias term.

In equation (6), the unknown vector \mathbf{a} and parameter b can be calculated by solving the following optimization problem:

$$\begin{aligned} \text{minimum of} \quad & \frac{1}{2} \mathbf{a}^T \mathbf{a} + \frac{\nu}{2} \sum_{i=1}^n \zeta_i^2, \\ \text{subject to} \quad & Z_i = \mathbf{a}^T \phi(Y_i) + b + \zeta_i, \quad i = 1, \dots, n, \end{aligned} \quad (7)$$

where ν is a regularization constant that constitutes a trade-off between the model complexity and the empirical error, and ζ_i is a regression error.

The problem (7) can be reformulated as an unconstrained problem through the Lagrange multiplier method [39]. A set of linear equations corresponding to optimality conditions is consequently obtained, which provides an expression of the weight vector \mathbf{a} :

$$\mathbf{a} = \sum_{i=1}^n \alpha_i \phi(Y_i), \quad (8)$$

where α_i is a Lagrange multiplier. This system of equations can be reduced to the following form:

$$\begin{bmatrix} 0 & \mathbf{1}_n^T \\ \mathbf{1}_n & \mathbf{\Omega} + \nu^{-1} \mathbf{I}_n \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{Z} \end{bmatrix}, \quad (9)$$

where $\mathbf{\Omega}$ is referred to as the kernel matrix whose element is $\phi^T(Y_i) \phi(Y_j)$ for $i, j = 1, \dots, n$ and \mathbf{I}_n is the identity matrix of size n . The matrix $\mathbf{1}_n$ is an $n \times 1$ unit matrix and $\mathbf{Z} = [Z_1, \dots, Z_n]^T$ is the $n \times 1$ matrix of observed values together with the $n \times 1$ matrix of Lagrange multipliers, $\alpha = [\alpha_1, \dots, \alpha_n]^T$.

The solutions of equation (9) are

$$b = \frac{\mathbf{1}_n^T \mathbf{A}^{-1} \mathbf{Z}}{\mathbf{1}_n^T \mathbf{A}^{-1} \mathbf{1}_n}, \quad (10)$$

$$\alpha = \mathbf{A}^{-1} (\mathbf{Z} - b \mathbf{1}_n), \quad (11)$$

where $\mathbf{A} = \mathbf{\Omega} + \nu^{-1} \mathbf{I}_n$, which is a symmetric and positive semi-definite matrix, thereby ascertaining the existence of its inverse, denoted as \mathbf{A}^{-1} .

By replacing equation (8) with equation (6), the model for the LSSVR function therefore becomes

$$\begin{aligned} \mu(Y) &= \sum_{i=1}^n \alpha_i \phi^T(Y_i) \phi(Y) + b, \\ &= \sum_{i=1}^n \alpha_i K(Y_i, Y) + b, \end{aligned} \quad (12)$$

where $K(\cdot, \cdot)$ is the kernel associated with the feature mapping ϕ , and it is defined as [40]

$$K(Y_i, Y) = \phi^T(Y_i) \phi(Y). \quad (13)$$

Numerous kernel functions are available for the construction of various models, such as:

Linear kernel: $K(Y_i, Y) = Y_i^T Y$.

Polynomial kernel: $K(Y_i, Y) = (1 + Y_i^T Y)^p$, $k > 0$ and $p \in \mathbb{N}$.

Radial basis function kernel: $K(Y_i, Y) = \exp(-g\|Y - Y_i\|^2)$, $g > 0$ and $\|\cdot\|$ means the Euclidean norm.

The LSSVR regarding the KED scheme is used to characterize the underlying trend. The notation for the dataset at the sample point $s_i \in \mathbb{R}^d$ is represented by $\{X(s_i), Z(s_i)\}_{i=1}^n$, where $X(s_i) = [X_1(s_i), \dots, X_\eta(s_i)]^T \in \mathbb{R}^\eta$ is the vector of η auxiliary variables and $Z(s_i)$ denotes the observation value.

3. A Novel Trend Function of KED based on LSSVR

This section introduces a method for constructing the trend function in KED using the LSSVR method. The approach involves identifying the fundamental functions of the trend through explicit feature mapping. Examples of explicit feature mappings derived from the corresponding kernel functions are also demonstrated.

3.1. Construction of the Trend Function

Let $\phi(X(s))$ be an M -dimensional feature mapping such that

$$\phi(X(s)) = [\phi_1(X(s)), \dots, \phi_M(X(s))]^T, \quad (14)$$

where $\phi_m(X(s))$ is the m th component of the feature mapping.

According to equation (13), the kernel function is in the form:

$$K(X(s_i), X(s)) = [\phi_1(X(s_i)), \dots, \phi_M(X(s_i))] \begin{bmatrix} \phi_1(X(s)) \\ \vdots \\ \phi_M(X(s)) \end{bmatrix}. \quad (15)$$

By substituting equation (15) into equation (12), $\mu(s)$ can be rewritten as

$$\begin{aligned} \mu(s) &= \mu(X(s)) = \sum_{i=1}^n \alpha_i \left([\phi_1(X(s_i)), \dots, \phi_M(X(s_i))] \begin{bmatrix} \phi_1(X(s)) \\ \vdots \\ \phi_M(X(s)) \end{bmatrix} \right) + b, \\ &= \sum_{i=1}^n \alpha_i \left(\sum_{m=1}^M \phi_m(X(s_i)) \phi_m(X(s)) \right) + b, \\ &= \sum_{m=1}^M \sum_{i=1}^n \alpha_i \phi_m(X(s_i)) \phi_m(X(s)) + b, \\ &= \sum_{m=1}^M \left(\sum_{i=1}^n \alpha_i \phi_m(X(s_i)) \right) \phi_m(X(s)) + b. \end{aligned} \quad (16)$$

The trend function can hence be written in the following form:

$$\mu(s) = \sum_{m=0}^M \tilde{a}_m \phi_m(X(s)), \quad (17)$$

where the coefficient $\tilde{a}_m = \sum_{i=1}^n \alpha_i \phi_m(X(s_i))$ and $\phi_m(X(s))$ is the known function for $m = 1, \dots, M$ in which $\tilde{a}_0 = b$ and $\phi_0(X(s)) = 1$.

Equation (17) possesses a similar form to equation (2) with \tilde{a}_m being treated as a_l and ϕ_m being f_l . This verifies the use of the kernel function as a non-linear trend model for the KED. The process of the KED based on the LSSVR method is provided according to the flowchart shown in Figure 1.

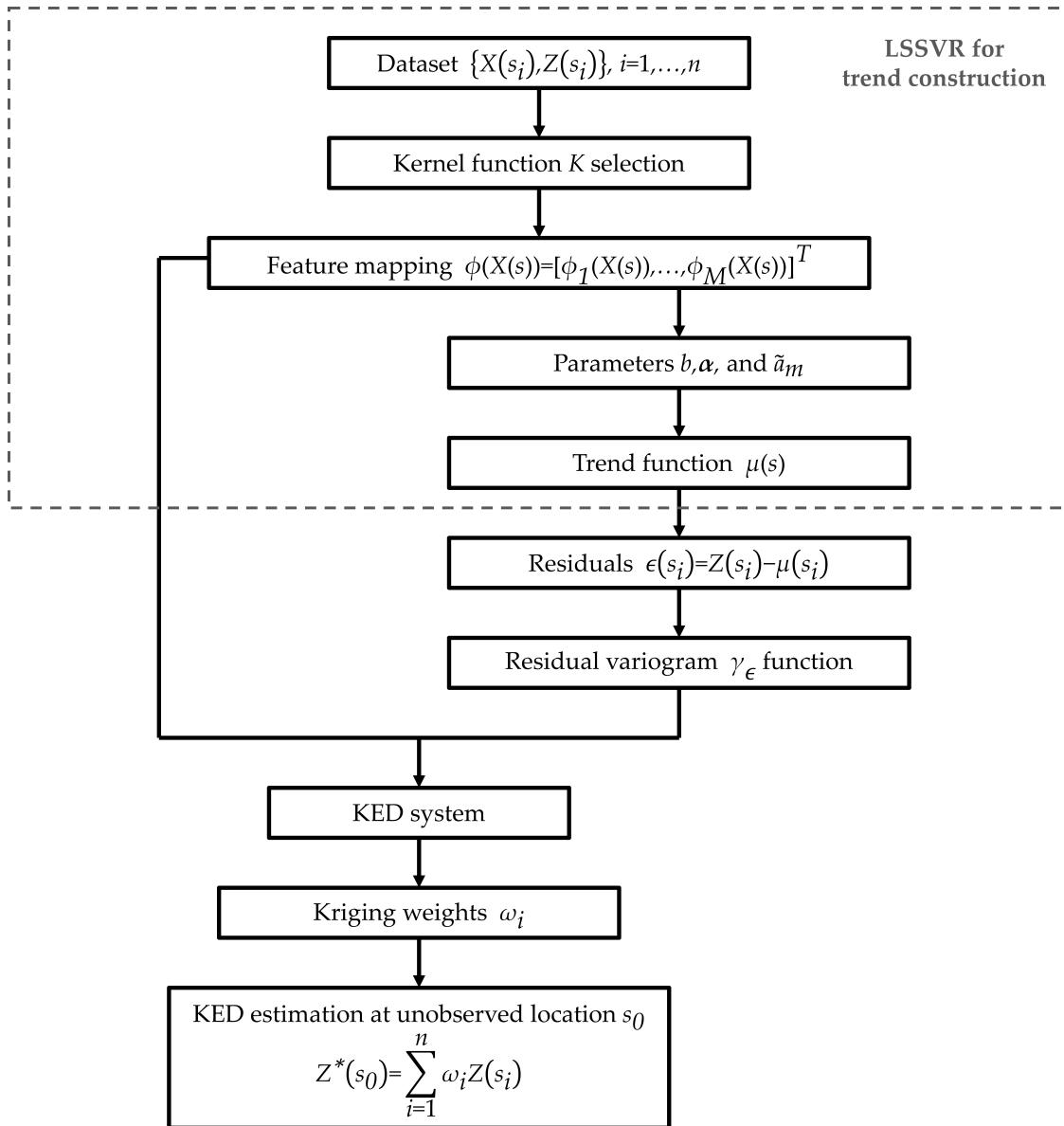


Figure 1. Flowchart of interpolation using KED method with the proposed trend function

3.2. Examples of Explicitly Feature Mapping

There have been various kernels for the LSSVR method, namely linear, polynomial, and radial basis function kernels [41–43]. This section presents the last two kernels, as they are widely used and relatively easy to tune. They will also be applied in our model to formulate the trend component.

3.2.1. Polynomial Kernel

The polynomial kernel function is defined as

$$K(X(s_i), X(s)) = (k + X^T(s_i)X(s))^p, \quad (18)$$

where $k > 0$ and $p \in \mathbb{N}$ is the degree of polynomial.

The feature mapping for polynomial kernel degree- p is given by [44]

$$\phi(X(s)) = \left[\sqrt{\frac{p!}{j_1! \cdots j_{\eta+1}!}} X_1^{j_1}(s) \cdots X_{\eta}^{j_{\eta}}(s) \sqrt{k^{\eta+1}} \mid j_i \geq 0 \text{ with } \sum_{i=1}^{\eta+1} j_i = p \right], \quad (19)$$

where the dimensionality of $\phi(X(s))$ is $\frac{(\eta+p)(\eta+p-1)\cdots(\eta+1)}{p!}$ [45]. For example, when the degree of the polynomial kernel and the number of auxiliary variables are both equal to 2, with k being 1, then

$$\phi(X(s)) = \left[1, \sqrt{2}X_1(s), \sqrt{2}X_2(s), X_1^2(s), X_2^2(s), \sqrt{2}X_1(s)X_2(s) \right]^T. \quad (20)$$

Comparing with equation (14), components of the feature mapping are as follows:

$$\begin{aligned} \phi_1(X(s)) &= 1, \phi_2(X(s)) = \sqrt{2}X_1(s), \phi_3(X(s)) = \sqrt{2}X_2(s), \\ \phi_4(X(s)) &= X_1^2(s), \phi_5(X(s)) = X_2^2(s), \phi_6(X(s)) = \sqrt{2}X_1(s)X_2(s). \end{aligned}$$

3.2.2. Radial Basis Function Kernel

The implicit kernel function, exemplified by the radial basis function (RBF) kernel, assumes the following form:

$$K(X(s_i), X(s)) = \exp(-g\|X(s) - X(s_i)\|^2), \quad (21)$$

where $g > 0$ is a RBF kernel parameter.

The feature mapping for RBF kernel function can be formulated as

$$\phi(X(s)) = \exp(-g\|X(s)\|^2) \left[\sqrt{\frac{(2g)^r}{r!}} \phi_r(X(s)) \mid r = 0, \dots, \infty \right], \quad (22)$$

where

$$\phi_r(X(s)) = \left[\sqrt{\frac{r!}{j_1! \cdots j_{\eta}!}} X_1^{j_1}(s) \cdots X_{\eta}^{j_{\eta}}(s) \mid j_i \geq 0 \text{ with } \sum_{i=1}^{\eta} j_i = r \right], \quad (23)$$

which is described in more detail in [46]. The RBF kernel function, which maps the auxiliary data to an infinite-dimensional space, can be approximated by Taylor Polynomial-based Monomial feature mapping (TPM feature mapping). In the work of [46], a finite-dimensional approximated feature mapping of the RBF function is obtained as follows:

$$\phi(X(s)) = \exp(-g\|X(s)\|^2) \left[\sqrt{\frac{(2g)^r}{r!}} \phi_r(X(s)) \mid r = 0, \dots, r_u \right], \quad (24)$$

where r_u is a selected approximation degree and the TPM feature mapping degree- r_u has $\frac{(\eta+r_u)(\eta+r_u-1)\cdots(\eta+1)}{r_u!}$ dimensions.

Although an increase in r_u leads to an improvement in estimation as $\phi(X(s))$ approaches the true function, it is however sufficient enough to use the TPM feature mapping with low dimensions [46]. An example of TPM feature mapping with degree-two and 2 auxiliary variables is

$$\phi(X(s)) = \exp(-g\|X(s)\|^2) \left[1, \sqrt{2g}X_1(s), \sqrt{2g}X_2(s), \sqrt{2g}X_1^2(s), \sqrt{2g}X_2^2(s), 2gX_1(s)X_2(s) \right]^T. \quad (25)$$

By comparing equation (25) with equation (14), this results in
 $\phi_1(X(s)) = \exp(-g\|X(s)\|^2), \phi_2(X(s)) = \exp(-g\|X(s)\|^2)\sqrt{2g}X_1(s),$

$$\begin{aligned}\phi_3(X(s)) &= \exp(-g\|X(s)\|^2)\sqrt{2g}X_2(s), \phi_4(X(s)) = \exp(-g\|X(s)\|^2)\sqrt{2g}X_1^2(s), \\ \phi_5(X(s)) &= \exp(-g\|X(s)\|^2)\sqrt{2g}X_2^2(s), \phi_6(X(s)) = \exp(-g\|X(s)\|^2)2gX_1(s)X_2(s).\end{aligned}$$

4. Case Study: Estimations of Temperature and Pressure in Thailand

4.1. Study Area

The evaluation of the efficiency and accuracy of the proposed techniques was carried out to interpolate temperature and pressure in Thailand. The country is located between $5^{\circ}37'N$ and $20^{\circ}27'N$ latitude and $97^{\circ}22'E$ and $105^{\circ}37'E$ longitude, with a total area of $513,115\text{ km}^2$ and a coastline of $3,219\text{ km}$ [47,48]. The data used in this study consist of monthly averages of temperature, pressure, relative humidity, digital elevation model (DEM), and geographic locations (coordinates) spanning from January 2017 to December 2017. These data were acquired from the National Hydroinformatics and Climate Data Center (NHC), developed by Hydro-Informatics Institute (HII) [49]. Figure 2 displays 213 meteorological stations after data preparation and cleaning.

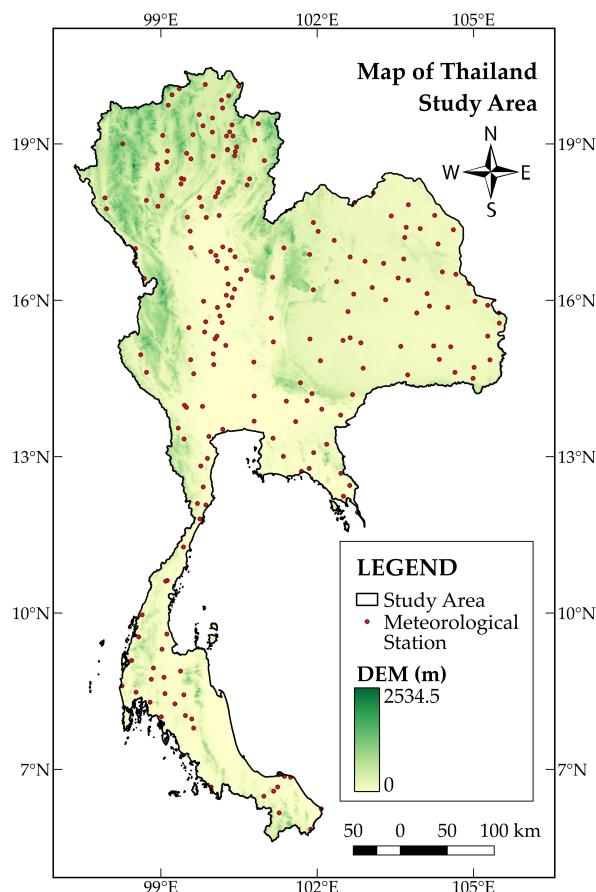


Figure 2. Spatial distributions of meteorological stations in the study area in 2017

4.2. Evaluation of Model Accuracy

In this study, we compare the accuracy of KED with three different trend functions: the linear trend function estimated using the GLS estimator (KED–GLS), the non-linear trend function based on LSSVR with polynomial feature mapping of degree one and two (KED–Poly1 and KED–Poly2), and the non-linear trend function based on LSSVR with TPM feature mapping of degree one and two (KED–TPM1 and KED–TPM2).

The k –fold cross-validation technique was applied to examine the performance of the models. The data were randomly divided into 10 folds. For each iteration, each fold was used as a testing dataset

for the model built upon the remaining nine folds. After 10 iterations in which each fold was once selected as testing data, the overall estimation accuracy is an average of the accuracy scores calculated from each iteration [50]. The root-mean-square error (RMSE) [51] and the mean-absolute-percentage error (MAPE) [52] were the model performance indicators, which are formulated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (Z(s_i) - Z^*(s_i))^2}, \quad (26)$$

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|Z(s_i) - Z^*(s_i)|}{Z(s_i)} \times 100, \quad (27)$$

where N is the number of observations, $Z(s_i)$ and $Z^*(s_i)$ denote the observed data and the estimated value at coordinate s_i , respectively.

4.3. Results

Before proceeding to the KED simulation, a selection of auxiliary factors is required. Table 1 presents the statistical analysis of the interdependence between each selected variable and the target variables through the Pearson and Spearman correlation coefficients [53]. The results indicate a significant positive correlation between temperature and pressure (correlation coefficients greater than 0.5). While the pressure is negatively correlated to both DEM and latitude with the correlation coefficients less than -0.5. This suggests that pressure can be chosen as an auxiliary variable for temperature estimation and vice versa. On the other hand, both DEM and latitude are additionally included as auxiliary factors for interpolating pressure.

Table 1. Correlation coefficients between the target and auxiliary variables

Auxiliary variables	Temperature		Pressure	
	Pearson	Spearman	Pearson	Spearman
Temperature	1.0000	1.0000	0.5537	0.5092
Pressure	0.5537	0.5092	1.0000	1.0000
Relative humidity	-0.2474	-0.2389	0.1540	0.1869
Latitude	-0.1322	-0.2425	-0.5338	-0.5867
Longitude	0.0198	0.0345	0.0283	-0.0399
DEM	-0.4501	-0.4421	-0.7350	-0.7470

Table 2 reports the estimation efficiency of the KED model with three different trend functions via MAPE and RMSE measures. According to the accuracy statistics, the KED with a non-linear trend function based on LSSVR has a superior estimation performance to that of the KED with a linear trend for both temperature and pressure. Specifically, the prediction errors for temperature generated by the KED–TPM2 are smaller than those of all other methods, with RMSE being 0.8123 and MAPE being 2.2888. These are respectively equivalent to 1.5633% and 2.1755% improvement compared to the KED–GLS method. While the optimal pressure estimates are achieved by using the KED–Poly2 with the RMSE and MAPE equal to 7.7541 and 0.5466 respectively. The KED–Poly2 reduces both MAPE and RMSE values by over 10% with respect to the KED–GLS approach.

Table 2. Prediction errors of kriging with external drift and three different trend functions for temperature and pressure data in 2017

Target variables	Auxiliary variables	Errors	KED with linear trend	KED with non-linear trend based on LSSVR			
				Polynomial Feature Mapping		TPM Feature Mapping	
			KED-GLS	KED-Poly1	KED-Poly2	KED-TPM1	KED-TPM2
Temperature	Pressure	RMSE	0.8252	0.8275	0.8232	0.8512	0.8123
		MAPE	2.3397	2.3486	2.3439	2.4041	2.2888
Pressure	DEM Latitude Temperature	RMSE	8.6212	8.6111	7.7541	9.3854	8.7372
		MAPE	0.6170	0.6181	0.5466	0.6698	0.6329

To further compare the estimation performance of all methods, visual spatial distribution patterns of the monthly averages of temperature and pressure in Thailand in March, July, and November 2017 are presented. These maps were created using QGIS (Quantum Geographic Information System) software and the study area was partitioned into a grid of square cells of 0.05 degree per side.

Figure 3 shows the spatial distribution patterns of the monthly mean temperature. The panels on the left column depict the results generated by KED–GLS, whereas panels on the right column illustrate the results obtained from KED–TPM2. Both KED–TPM2 and KED–GLS produce a roughly similar distribution pattern for the average July temperature. This may be due to the fact that there is a small variation in temperature across the country during the rainy season (July–October). The discrepancy in temperature derived from these two models is therefore not significant. On the contrary, clear differences can be observed in March and November, in which the area of high temperature is more broadly distributed in the central part of the country for the KED–TPM2. The model also generates an overall lower temperature level concentrated in the northern region in November. Figure 4 displays spatial distribution maps of monthly mean pressure where the left column again corresponds to the estimates attained from the KED–GLS while those in the right column are derived from the KED–Poly2. The results show a distinct difference between these two methods. In particular, lower pressure values estimated by KED–Poly2 are clearly marked in the northern and western parts of the study area.

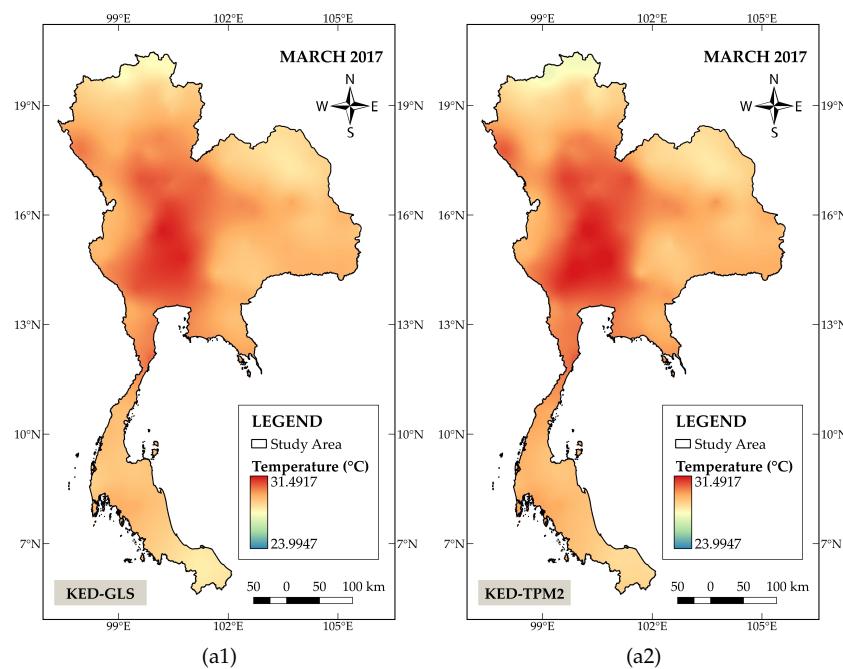


Figure 3. Cont.

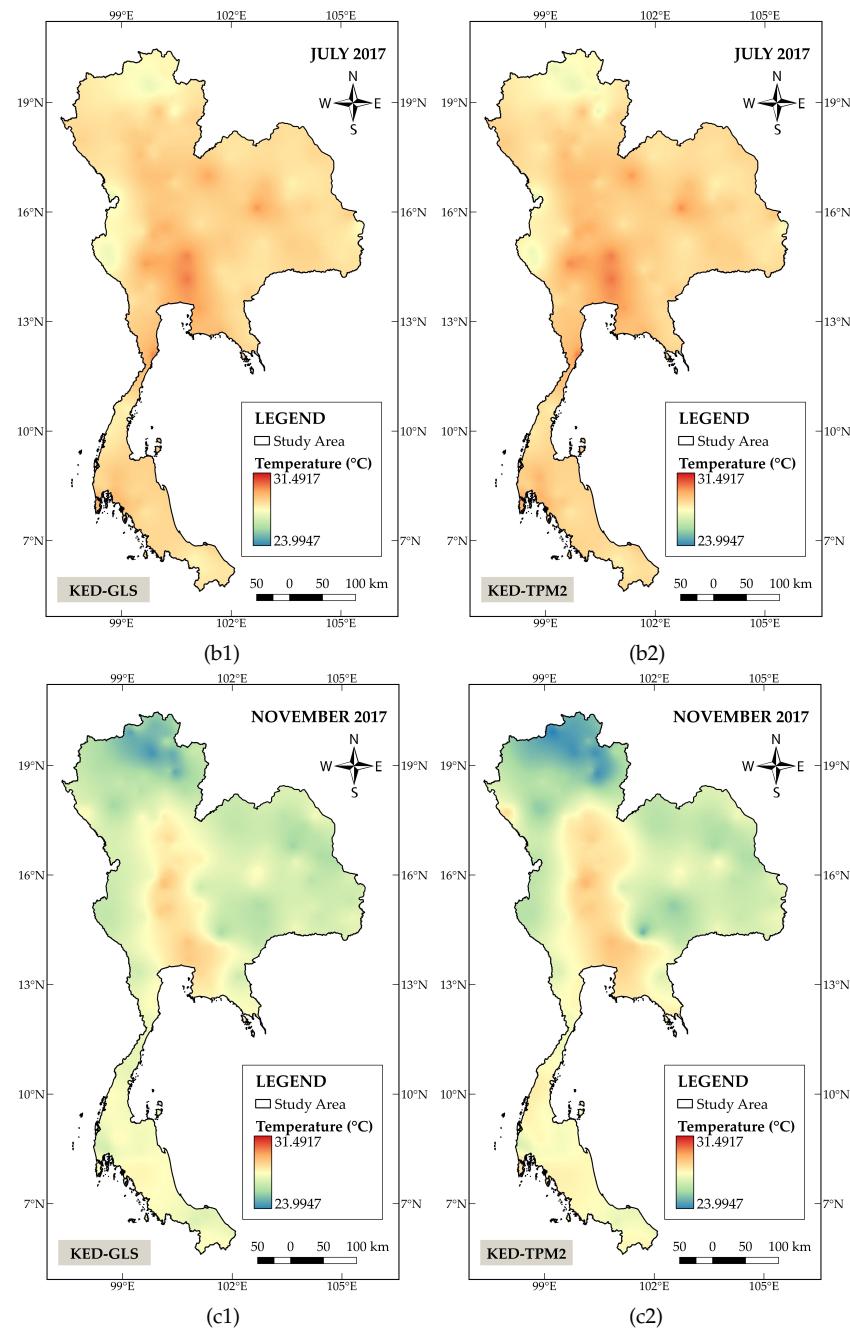


Figure 3. Spatial distribution of temperature in Thailand in March, July, and November 2017, interpolated using: (a1), (b1), and (c1) KED-GLS (left panels); (a2), (b2), and (c2) KED-TPM2 (right panels)

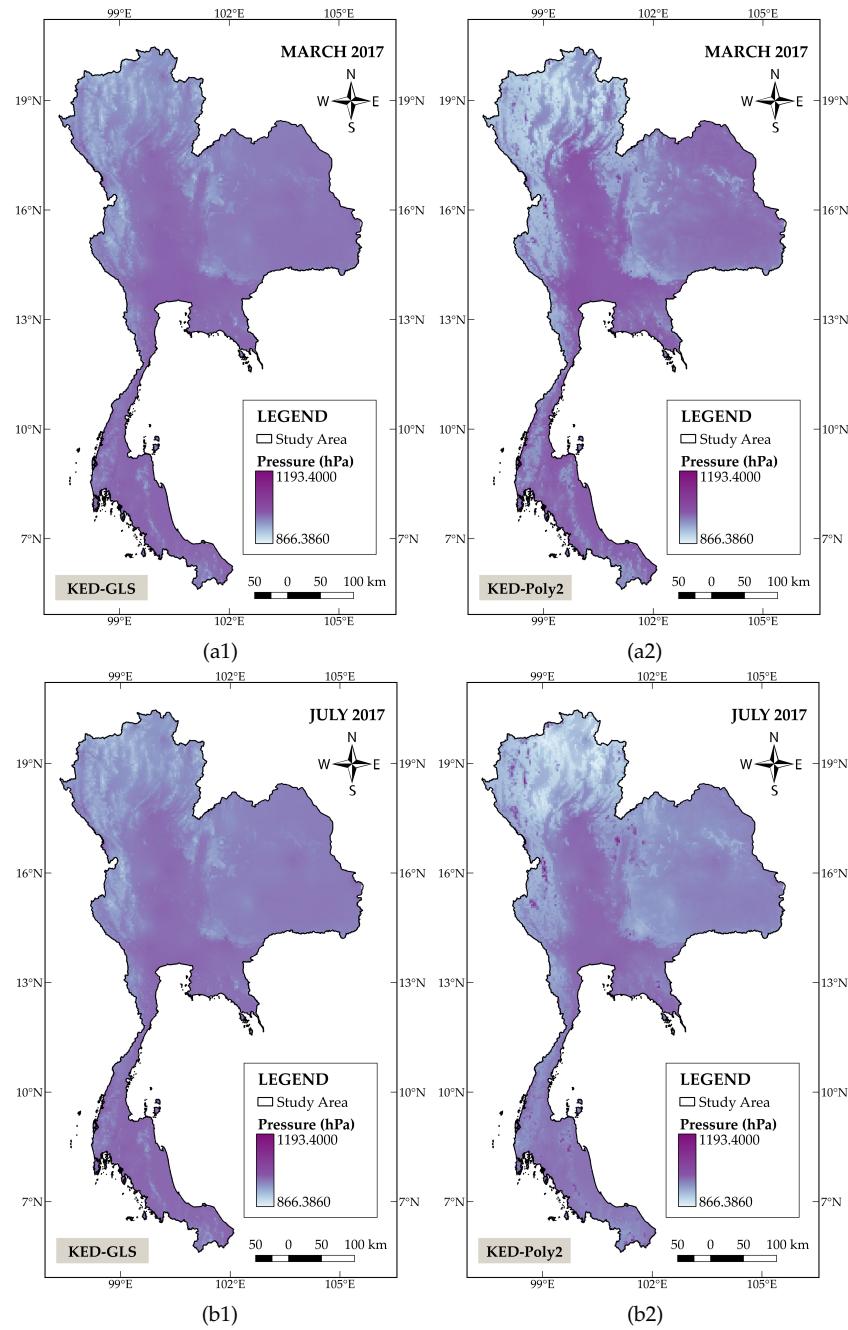


Figure 4. Cont.

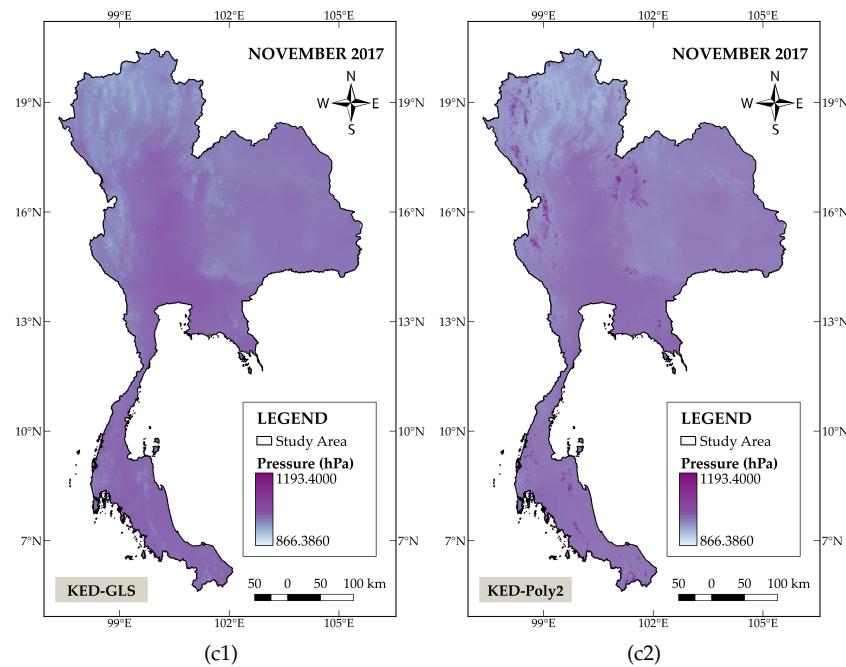


Figure 4. Spatial distribution of pressure in Thailand in March, July, and November 2017, interpolated using: (a1), (b1), and (c1) KED–GLS (left panels); (a2), (b2), and (c2) KED–Poly2 (right panels)

5. Conclusion and Discussion

This paper presents the novel KED method that applies the LSSVR technique to improve spatial interpolation accuracy in the presence of non-linear trends. The method involves determining the drift component through explicit feature mapping which is expressed in terms of kernel functions. A comparison between our proposed method and the KED with the linear trend is demonstrated in the case of the temperature and pressure estimation in Thailand in 2017. The results show that the KED with LSSVR outperforms the KED approach with a linear trend function regarding estimation accuracy.

The advantage of the KED with LSSVR can be attributed to its ability to extract implicit non-linear relationships between the target and auxiliary variables. This gives rise to more accurate interpolation results. Furthermore, the LSSVR is a powerful machine learning algorithm that has been proven effective in a variety of regression tasks. This allows our method to adapt to various data types. However, the choice of kernel function in LSSVR can have a significant impact on the estimation accuracy. Although a higher-degree polynomial kernel or a higher-degree TPM feature mapping can model more complex relationships in the data, it can also result in a number of equations in the kriging system. This can lead to more time-intensive computation and an increase in the likelihood of model overfitting issues.

Author Contributions: Conceptualization, K.B., N.C. and S.M.; methodology, K.B., N.C. and S.M.; software, K.B. and S.M.; validation, K.B., N.C. and S.M.; formal analysis, K.B., N.C., and S.M.; investigation, K.B., N.C. and S.M.; resources, K.B., N.C. and S.M.; data curation, K.B. and S.M.; writing-original draft preparation, K.B., N.C. and S.M.; writing-review and editing, K.B., N.C. and S.M.; visualization, K.B.; supervision, N.C. and S.M. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: All data were acquired from the National Hydroinformatics and Climate Data Center (NHC), developed by Hydro-Informatics Institute (HII) [49].

Acknowledgments: This work was supported by (i) Chiang Mai University (CMU) and (ii) a Fundamental Fund provided by Thailand Science Research and Innovation (TSRI) and National Science, Research and Innovation Fund (NSRF).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wackernagel, H. *Multivariate geostatistics: an introduction with applications*; Springer Science & Business Media, 2003.
2. Webster, R.; Oliver, M.A. *Geostatistics for environmental scientists*; John Wiley & Sons, 2007.
3. Hudson, G.; Wackernagel, H. Mapping temperature using kriging with external drift: theory and an example from Scotland. *International journal of Climatology* **1994**, *14*, 77–91.
4. Bostan, P.; Heuvelink, G.B.; Akyurek, S. Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey. *International Journal of Applied Earth Observation and Geoinformation* **2012**, *19*, 115–126.
5. Varentsov, M.; Esau, I.; Wolf, T. High-resolution temperature mapping by geostatistical kriging with external drift from large-eddy simulations. *Monthly Weather Review* **2020**, *148*, 1029–1048.
6. Cantet, P. Mapping the mean monthly precipitation of a small island using kriging with external drifts. *Theoretical and Applied Climatology* **2017**, *127*, 31–44.
7. Bourennane, H.; King, D.; Chery, P.; Bruand, A. Improving the kriging of a soil variable using slope gradient as external drift. *European Journal of Soil Science* **1996**, *47*, 473–483.
8. Bourennane, H.; King, D.; Couturier, A. Comparison of kriging with external drift and simple linear regression for predicting soil horizon thickness with different sample densities. *Geoderma* **2000**, *97*, 255–271.
9. Bourennane, H.; King, D. Using multiple external drifts to estimate a soil variable. *Geoderma* **2003**, *114*, 1–18.
10. Béjar-Pizarro, M.; Guardiola-Albert, C.; García-Cárdenes, R.P.; Herrera, G.; Barra, A.; López Molina, A.; Tessitore, S.; Staller, A.; Ortega-Becerril, J.A.; García-García, R.P. Interpolation of GPS and geological data using InSAR deformation maps: Method and application to land subsidence in the alto guadalentín aquifer (SE Spain). *Remote Sensing* **2016**, *8*, 965.
11. Beauchamp, M.; de Fouquet, C.; Malherbe, L. Dealing with non-stationarity through explanatory variables in kriging-based air quality maps. *Spatial statistics* **2017**, *22*, 18–46.
12. Beauchamp, M.; Malherbe, L.; de Fouquet, C.; Létinois, L.; Tognet, F. A polynomial approximation of the traffic contributions for kriging-based interpolation of urban air quality model. *Environmental Modelling & Software* **2018**, *105*, 132–152.
13. Troisi, S.; Fallico, C.; Straface, S.; Migliari, E. Application of kriging with external drift to estimate hydraulic conductivity from electrical-resistivity data in unconsolidated deposits near Montalto Uffugo, Italy. *Hydrogeology Journal* **2000**, *8*, 356–367.
14. Garcia-Papani, F.; Leiva, V.; Ruggeri, F.; Uribe-Opazo, M.A. Kriging with external drift in a Birnbaum-Saunders geostatistical model. *Stochastic Environmental Research and Risk Assessment* **2018**, *32*, 1517–1530.
15. Cafarelli, B.; Castrignanò, A. The use of geoadditive models to estimate the spatial distribution of grain weight in an agronomic field: a comparison with kriging with external drift. *Environmetrics* **2011**, *22*, 769–780.
16. Anand, A.; Singh, P.; Srivastava, P.K.; Gupta, M. GIS-based analysis for soil moisture estimation via kriging with external drift. In *Agricultural water management*; Elsevier, 2021; pp. 391–408.
17. Rivest, M.; Marcotte, D.; Pasquier, P. Hydraulic head field estimation using kriging with an external drift: A way to consider conceptual model information. *Journal of Hydrology* **2008**, *361*, 349–361.
18. Snepvangers, J.; Heuvelink, G.; Huisman, J. Soil water content interpolation using spatio-temporal kriging with external drift. *Geoderma* **2003**, *112*, 253–271.
19. Freier, L.; von Lieres, E. Kriging based iterative parameter estimation procedure for biotechnology applications with nonlinear trend functions. *IFAC-PapersOnLine* **2015**, *48*, 574–579.
20. Freier, L.; Wiechert, W.; von Lieres, E. Kriging with trend functions nonlinear in their parameters: Theory and application in enzyme kinetics. *Engineering in life sciences* **2017**, *17*, 916–922.
21. Mozer, M.C.; Jordan, M.I.; Petsche, T. *Advances in Neural Information Processing Systems 9: Proceedings of the 1996 Conference*; Vol. 9, Mit Press, 1997.
22. Al-Anazi, A.F.; Gates, I.D. Support vector regression to predict porosity and permeability: Effect of sample size. *Computers & geosciences* **2012**, *39*, 64–76.
23. Wiering, M.A.; Van der Ree, M.H.; Embrechts, M.; Stollenga, M.; Meijster, A.; Nolte, A.; Schomaker, L. The neural support vector machine. BNAIC 2013: Proceedings of the 25th Benelux Conference on Artificial

Intelligence, Delft, The Netherlands, November 7–8, 2013. Delft University of Technology (TU Delft); under the auspices of the Benelux . . . , 2013.

- 24. Zhong, H.; Wang, J.; Jia, H.; Mu, Y.; Lv, S. Vector field-based support vector regression for building energy consumption prediction. *Applied Energy* **2019**, *242*, 403–414.
- 25. Henrique, B.M.; Sobreiro, V.A.; Kimura, H. Stock price prediction using support vector regression on daily and up to the minute prices. *The Journal of finance and data science* **2018**, *4*, 183–201.
- 26. Zhang, J.; Teng, Y.F.; Chen, W. Support vector regression with modified firefly algorithm for stock price forecasting. *Applied Intelligence* **2019**, *49*, 1658–1674.
- 27. Mishra, S.; Padhy, S. An efficient portfolio construction model using stock price predicted by support vector regression. *The North American Journal of Economics and Finance* **2019**, *50*, 101027.
- 28. Fan, G.F.; Yu, M.; Dong, S.Q.; Yeh, Y.H.; Hong, W.C. Forecasting short-term electricity load using hybrid support vector regression with grey catastrophe and random forest modeling. *Utilities Policy* **2021**, *73*, 101294.
- 29. Arulmozhi, E.; Basak, J.K.; Sihalath, T.; Park, J.; Kim, H.T.; Moon, B.E. Machine learning-based microclimate model for indoor air temperature and relative humidity prediction in a swine building. *Animals* **2021**, *11*, 222.
- 30. Quan, Q.; Hao, Z.; Xifeng, H.; Jingchun, L. Research on water temperature prediction based on improved support vector regression. *Neural Computing and Applications* **2022**, pp. 1–10.
- 31. Jaiswal, P.; Gaikwad, M.; Gaikwad, N. Analysis of AI techniques for healthcare data with implementation of a classification model using support vector machine. *Journal of Physics: Conference Series*. IOP Publishing, 2021, Vol. 1913, p. 012136.
- 32. Al-Manaseer, H.; Abualigah, L.; Alsoud, A.R.; Zitar, R.A.; Ezugwu, A.E.; Jia, H. A novel big data classification technique for healthcare application using support vector machine, random forest and J48. In *Classification applications with deep learning and machine learning technologies*; Springer, 2022; pp. 205–215.
- 33. Suykens, J.A.; Vandewalle, J. Least squares support vector machine classifiers. *Neural processing letters* **1999**, *9*, 293–300.
- 34. Guo, Z.; Bai, G. Application of least squares support vector machine for regression to reliability analysis. *Chinese Journal of Aeronautics* **2009**, *22*, 160–166.
- 35. Cressie, N. *Statistics for spatial data*; John Wiley & Sons, 2015.
- 36. Vallejos, R.; Osorio, F.; Bevilacqua, M. *Spatial relationships between two georeferenced variables: With applications in R*; Springer Nature, 2020.
- 37. Ly, S.; Charles, C.; Degre, A. Geostatistical interpolation of daily rainfall at catchment scale: the use of several variogram models in the Ourthe and Ambleve catchments, Belgium. *Hydrology and Earth System Sciences* **2011**, *15*, 2259–2274.
- 38. Amini, M.A.; Torkan, G.; Eslamian, S.; Zareian, M.J.; Adamowski, J.F. Analysis of deterministic and geostatistical interpolation techniques for mapping meteorological variables at large watershed scales. *Acta Geophysica* **2019**, *67*, 191–203.
- 39. Huang, P.; Yu, H.; Wang, T. A Study Using Optimized LSSVR for Real-Time Fault Detection of Liquid Rocket Engine. *Processes* **2022**, *10*, 1643.
- 40. Yeh, W.C.; Zhu, W. Forecasting by Combining Chaotic PSO and Automated LSSVR. *Technologies* **2023**, *11*, 50.
- 41. Xie, G.; Wang, S.; Zhao, Y.; Lai, K.K. Hybrid approaches based on LSSVR model for container throughput forecasting: A comparative study. *Applied Soft Computing* **2013**, *13*, 2232–2241.
- 42. Hongzhe, M.; Wei, Z.; Rongrong, W. Prediction of dissolved gases in power transformer oil based on RBF-LSSVM regression and imperialist competition algorithm. 2017 2nd International Conference on Power and Renewable Energy (ICPRE). IEEE, 2017, pp. 291–295.
- 43. Wang, X.; Wang, G.; Zhang, X. Prediction of Chlorophyll-a content using hybrid model of least squares support vector regression and radial basis function neural networks. 2016 Sixth International Conference on Information Science and Technology (ICIST). IEEE, 2016, pp. 366–371.
- 44. Shashua, A. Introduction to machine learning: Class notes 67577. *arXiv preprint arXiv:0904.3664* **2009**.
- 45. Chang, Y.W.; Hsieh, C.J.; Chang, K.W.; Ringgaard, M.; Lin, C.J. Training and testing low-degree polynomial data mappings via linear SVM. *Journal of Machine Learning Research* **2010**, *11*.
- 46. Lin, K.P.; Chen, M.S. Efficient kernel approximation for large-scale support vector machine classification. *Proceedings of the 2011 SIAM International Conference on Data Mining*. SIAM, 2011, pp. 211–222.
- 47. Chariyaphan, R. Thailand's country profile 2012. *Department of Disaster Prevention and Mitigation, Ministry of Interior, Thailand* **2012**.

48. Laonamsai, J.; Ichiyanagi, K.; Kamdee, K. Geographic effects on stable isotopic composition of precipitation across Thailand. *Isotopes in Environmental and Health Studies* **2020**, *56*, 111–121.
49. OpenData. Available online: URL <https://data.hii.or.th/#/>. (accessed on 14 October 2020).
50. Du, K.L.; Swamy, M.N. *Neural networks and statistical learning*; Springer Science & Business Media, 2013.
51. Li, J.; Heap, A.D. A review of spatial interpolation methods for environmental scientists **2008**.
52. Bae, B.; Kim, H.; Lim, H.; Liu, Y.; Han, L.D.; Freeze, P.B. Missing data imputation for traffic flow speed using spatio-temporal cokriging. *Transportation Research Part C: Emerging Technologies* **2018**, *88*, 124–139.
53. Akoglu, H. User's guide to correlation coefficients. *Turkish journal of emergency medicine* **2018**, *18*, 91–93.

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.