



Article A Spatial Model for Repairing of the Dam Safety Monitoring Data Combining the Variable Importance for Projection (VIP) and Cokriging Methods

Shiwan Li^{1,2}, Yanling Li^{1,2,*}, Xiang Lu^{1,2}, Zhenyu Wu^{1,2}, Liang Pei^{1,2} and Kexin Liu³

- State Key Laboratory of Hydraulics and Mountain River Engineering, Sichuan University, Chengdu 610065, China
- ² College of Water Resource & Hydropower, Sichuan University, No. 24 South Section 1, Yihuan Road, Chengdu 610065, China
- ³ Huangtu Township Government, Jiangyin 214445, China
- * Correspondence: liyanling@scu.edu.cn

Abstract: The safe operation of dams is related to the lifeline of the national economy, the safety of the people, and social stability, and dam safety monitoring plays an essential role in scientifically controlling the safety of dams. Since the effects of environmental variables were not considered in conventional monitoring data repairing methods (such as the single time series model and spatial interpolation model), a spatial model for repairing monitoring data combining the variable importance for projection (VIP) method and cokriging was put forward in this paper. In order to improve the accuracy of the model, the influence of different combinations of covariates on it was discussed, and the VIP_i value greater than 0.8 was proposed as the threshold of covariates. The engineering verification shows that the VIP-cokriging spatial model had the advantages of high precision and strong applicability compared with the inverse distance weighting (IDW) model, the ordinary kriging model, and the universal kriging model, and the overall error can be reduced by more than 60%, which could better realize the expansion of the monitoring effect variable to the whole area of the dam space. The engineering application of the PBG dam showed that the model scientifically correlated the existing monitoring points with the spatial location of the dam, and reasonably repaired the measured values of the stopping and abnormal measured points, effectively ensuring that the spatial regular of the monitoring data could truly reflect the actual safety and operational status of the dam.

Keywords: dam safety monitoring; variable importance for projection; cokriging; data repair; the spatial model

1. Introduction

Dam safety is related to the national economy and people's livelihood. With the rapid development of information technology, big data, the Internet of Things, electronic communication, and other technologies, dam operation and safety control are gradually developing towards automation and intelligence [1] while continuous and reliable dam safety monitoring data is a prerequisite for the scientific evaluation of dam operation and safety conditions. However, affected by factors such as short-term abnormalities in monitoring instruments, monitoring instrument replacement, measurement errors, and external environmental disturbances [2], dam safety monitoring data is prone to data omissions, oscillation fluctuations, response misalignment, and other data anomalies. The occurrence of these anomalies will affect the continuity and reliability of the monitoring sequence, cause misjudgment of the dam's operational state, and even endanger the safe operation of the dam [3]. Therefore, it is essential to build a high-precision model for repairing dam safety monitoring data to control the general law and development trend of dams in real time for the intelligent control of dam safety operation [4,5].



Citation: Li, S.; Li, Y.; Lu, X.; Wu, Z.; Pei, L.; Liu, K. A Spatial Model for Repairing of the Dam Safety Monitoring Data Combining the Variable Importance for Projection (VIP) and Cokriging Methods. *Appl. Sci.* 2022, *12*, 12296. https://doi.org/ 10.3390/app122312296

Academic Editors: Grzegorz Psuj and Barbara Grochowalska (Szymanik)

Received: 24 October 2022 Accepted: 28 November 2022 Published: 1 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

At present, the repair of safety monitoring data mainly starts from the dimension of time and mostly adopts linear regression analysis [6,7], principal component analysis [8], machine learning algorithm [9,10], and so on. Among them, Vazifehdan et al. [11] proposed a method of combining a naive Bayesian network with tensor decomposition to repair missing data. Stojanovic et al. [12] used linear regression and genetic algorithm to construct an adaptive system for dam behavior modeling, which significantly improves the accuracy of data repair. Du et al. [13] proposed a model for repairing data with good performance in terms of training time and model robustness based on a deep learning approach. Nevertheless, these methods realize data repair by constructing time series models, which have high repair accuracy for data sequences with good periodic regularity and fail to meet the application requirements for data sequences with long-missing measurement time and poor data regularity [14]. Moreover, the single-point monitoring model can only reflect the local characteristics of the dam, but cannot reflect the overall state. The dam as a whole structure, the sequence of monitoring points for deformation, seepage and stress-strain of the dam contains information in both the temporal and spatial dimensions, that is, there is a correlation between time and space. Therefore, constructing a multi-dimensional spatial model for repairing data from the spatial coordination and consistency of dam effect variables is undoubtedly an effective path to solving this problem.

Since the gradual development of spatial information technology in the 1960s, many scholars have conducted a lot of research on spatial models. Models such as Thiessen polygon [15], inverse distance weighting [16,17], and kriging [18,19] had been introduced into spatial model construction. At present, there have been in-depth studies on the construction methods of spatial models in many fields such as meteorology [20,21] and geostatistics [22] that achieved good results. For example, Adhikary [23] used the KGP, which combined the nonparametric variogram model based on genetic programming with kriging, as a feasible alternative technique for spatial estimation and mapping of rainfall. Seo [24] proposed a hybrid model called RKNNRK that combined regression kriging and neural network residual kriging for determining spatial precipitation distribution. In the field of water conservancy project, the research of spatial model is still in the exploratory stage. Mao et al. [25] proposed a deep neural network multi-view learning method (DNN-MVL) based on inverse distance weighting to reveal the complex non-linear spatio-temporal relationships of dam deformation effectively. Zhao et al. [26] developed a spatio-temporal monitoring model of the center of mass based on a least-squares support vector machine by introducing the initial coordinates of the center of mass. Lu et al. [27] used the kriging interpolation method to construct a high-precision horizontal space-time gradient expansion model of the core wall of the earth-rock dam, which effectively reflected the three-dimensional deformation trend of the dam. Dai et al. [28] combined Kalman filtering and kriging spatial interpolation to construct a spatio-temporal model of dam deformation, efficaciously filtering out the noise of deformation data in time and space. Yang et al. [29] proposed a geographically and temporally weighted regression (GTWR) model which improved the accuracy of data repair during the missing period of measuring points. These methods are limited by monitoring technology and analysis theory, and mainly consider the geometric location of measuring points, and cannot reasonably consider the influence of important environmental factors on the number of monitoring effects. The accuracy and physical meaning of the model need to be improved [30], and also there are relatively few application cases in dam safety monitoring data restoration.

To sum up, a cokriging spatial model for repairing data based on variable importance for the projection method was proposed in this paper. The model could identify the important environmental impact factors of different parts, and combined with the spatial layout characteristics of the same monitoring effect variables to construct a one-dimensional, two-dimensional, or three-dimensional spatial model with the measured values of the measuring points as the main variables and the important environmental impact factors as the covariates. The expansion of dam safety monitoring data to the whole spatial area of the dam effectively solves the problem of the existing spatial model being unable to take into account the influence of environmental quantities and the selection of environmental impact factors relying on experience, which can significantly improve the accuracy of repairing abnormal or missing data and ensure that safety monitoring data can truly reflect the safe operation of the dam.

2. Methodology

Reasonable selection of main variables and covariates is the key to affecting the accuracy of the cokriging spatial model. In general, the main variables are usually chosen as the effect variables, while the covariates are mostly chosen as the key factors affecting the effect variables, which are the environmental impact factors for dam safety monitoring. Hence, the VIP-cokriging spatial model for repairing dam safety monitoring data was proposed in this paper, which combined the variable importance for projection method and the basic principle of cokriging. Firstly, the dimension of spatial model construction is determined according to the spatial arrangement characteristics of similar monitoring instruments. Secondly, the variable importance for the projection method is used to identify important environmental impact factors for effect variables and use them as covariates. Then, according to the principle of cokriging, it uses cross-covariance and cross-semivariance function modeling between multiple regionalized variables to characterize the correlation between them to obtain unbiased and optimal estimates of the regionalized variables. In order to calculate the covariance, the main variable covariance function, co-variable covariance function and cross-covariance function, the exponential model with better universality is selected to fit the covariance function. Finally, by fitting the covariance function, the measured values of spatially similar measurement points, covariates, and spatial location parameters are used to calculate spatially repaired measurements of data anomalies or unmeasured points. The flow is shown in Figure 1. The specific steps of the method are described as follows:



Figure 1. VIP-cokriging spatial model construction flow chart.

2.1. Decision of the Construction Dimension of Spatial Model

The spatial position of monitoring instruments is usually expressed using dam wheelbase, pile number and elevation, which are essential parameters for constructing the VIP-cokriging spatial model. Depending on the spatial arrangement characteristics of similar measurement points, it is generally considered to construct one-dimensional, twodimensional, and three-dimensional models. The one-dimensional models are usually constructed along a line of measurement, such as the horizontal displacement one-dimensional model of gravity dam crest and the uplift pressure one-dimensional model of sluice foundation, etc. The two-dimensional models are often constructed in conjunction with typical sections of dam safety monitoring, such as the internal horizontal displacement twodimensional model of earth-rock dam and the temperature two-dimensional model of gravity dam section. The three-dimensional models mainly aim at the effect variables of the whole dam or the entire dam area, such as the appearance deformation three-dimensional model of the earth-rock dam, the three-dimensional model for seepage around the dam, and so on. Due to the complexity and difficulty of implementing kriging interpolation in threedimensional space, and the covariance is only a function of the distance between points. Cokriging is only applicable to two-dimensional plane coordinate systems. Therefore, when constructing the three-dimensional model, it is necessary to transform the threedimensional spatial coordinates of the buried position of the monitoring equipment into two-dimensional plane coordinates, that is, the measuring points are projected onto the model building plane, and keep the distance between the measuring points after twodimensional transformation basically consistent with the distance under the original coordinates. In practical engineering, different two-dimensionalization methods of threedimensional coordinates should be selected according to the arrangement of instruments to ensure the application effect of the model after two-dimensional coordinates.

2.2. Selection of Covariates

The degree of correlation between the main variables and the covariates affects the accuracy of the cokriging model, so the factors with the higher degree of influence should be selected as the covariates. The influencing factors of dam deformation usually include water pressure, rainfall, temperature, aging, and so on. Taking the earth-rock dam as an example, the deformation of earth-rock dam caused by the change of water pressure is mainly formed by the compression and shear of soil and the displacement of dam foundation. The effect of temperature on the deformation of earth-rock dam is not obvious, and the change of external temperature generally only affects the surface soil. Rainfall infiltration raises water level and changes soil water content, thus affecting dam deformation. The creep of rockfill, the compression and plastic deformation of fissure joints and other weak structures at the bottom of reservoir under the action of water pressure, and the irreversible deformation of dam caused by cyclic loads such as the rise and fall of water level in front of dam also increase with the increase in dam operation life. Generally, three covariates are selected at most, and the more covariates are selected, the more complex and time-consuming the calculation is. Therefore, selecting appropriate covariates is the key to improving the accuracy and computational efficiency of the cokriging spatial model.

In this paper, the variable importance for projection method based on the partial least squares regression method [31] was used to identify the influence degree of environmental factors. It can effectively solve the problems of unstable recognition effect and lack of spatial consistency commonly found in the currently used methods such as the mathematical model method, weighted area method, and grey correlation analysis [32]. The central idea of the method is that the explanatory ability of the factor principal component t_h for the component x_i is equivalent to the correlation coefficient R² of the linear regression equation of x_i on t_h . Since the explanatory effect of x_i on y is transmitted through the principal component t_h , if the explanatory ability of t_h on y is strong and the effect of x_i on t_h is very significant, it can be assumed that the explanatory ability of x_i on y is strong. Therefore, the environmental factors such as water pressure, temperature, rainfall, and aging are regarded as independent variables x_i , and the effects variables such as dam deformation and seepage are regarded as dependent variables y. Then, the VIP_i value of the variable importance for the projection index is solved for each environmental factor separately, as seen in Equation (1). Among them, the values of water pressure, temperature and rainfall are directly measured by instruments, and the days from the observation time to the initial observation time are taken as the aging measurement values.

$$\operatorname{VIP}_{j} = \sqrt{\frac{k}{\sum\limits_{h=1}^{m} \operatorname{Rd}(y, t_{h})} \sum\limits_{h=1}^{m} \operatorname{Rd}(y, t_{h}) w_{hj}^{2}},$$
(1)

where *k* is the number of independent variables; t_h is the principal component extracted from related independent variables; $Rd(y, t_h)$ is the correlation coefficient between dependent

dent variables and principal components, which indicates the explanatory ability of the principal component to y; w_{hj} is the weight of independent variables on principal components.

The VIP_j value represents the degree of importance of the independent variables to the fit of the dependent variables, so the higher the value of VIP_j, the more significant the contribution of the independent variables to the dependent variables, and the more critical it is in the interpretation. Generally, if the VIP_j value of the independent variable is less than or equal to 0.8, its explanatory ability for *y* can be ignored [33]. In this paper, combined with the analysis of several practical engineering cases, it was found that the introduction of environmental factors with VIP_j value greater than 0.8 as covariates had a significant effect on improving the accuracy of the spatial model, while the introduction of environmental quantity factors with VIP_j value less than 0.8 had little effect, see Section 3 for details. Therefore, VIP_j = 0.8 was taken as the control threshold for introducing covariates, that was, the environmental quantities with VIP_j value greater than 0.8 were taken as covariates to participate in the construction of the cokriging spatial model.

2.3. Construction of Model

Cokriging interpolation greatly improves the accuracy of the interpolation by introducing one or more covariates (also referred to as auxiliary variables) that are closely related to the main variables, which is an improved method of kriging interpolation. This method is based on the theory of co-regionalized variables and carries out spatial interpolation according to it. The cokriging interpolation formula is as in Equation (2).

$$Z_0^* = \sum_{i=1}^n \lambda_i Z_i + \sum_{j=1}^k \sum_{i=1}^n \mu_{ji} Y_{ji},$$
(2)

where $Z_1, Z_2, ..., Z_n$ and $Y_{j1}, Y_{j2}, ..., Y_{jn}$ are *n* sample data for the main variables and covariates, respectively; *j* is the number of covariates; $\lambda_1, \lambda_2, ..., \lambda_n$ and $\mu_{j1}, \mu_{j2}, ..., \mu_{jn}$ are the cokriging weighting coefficients to be determined; Z_0^* is the repaired value of the random variable at 0.

Taking the second order as an example, if there are two regionalized values $Z_i(x)$ and $Y_i(x)$ (i = 1, 2, ..., n), which are related to the attribute in a region, then $Z_0^* = \sum_{i=1}^n \lambda_i Z_i + \sum_{i=1}^n \mu_i Y_i$. In order to satisfy the unbiased condition of the estimator, the sum of weights of main variables should be equal to 1, and the sum of weights of covariates should be equal to 0, as seen in Equation (3).

$$\begin{cases} \sum_{i=1}^{n} \lambda_i = 1\\ \sum_{i=1}^{n} \mu_i = 0 \end{cases}$$
(3)

According to the irreversible bias of the estimated value and the least square method, the Lagrange multiplication is used to solve the cokriging equation, as seen in Equation (4).

$$\begin{cases} \sum_{j=1}^{n} \lambda_j C_{Z_i Z_j} + \sum_{j=1}^{m} \mu_j C_{Z_i Y} - \eta_1 = C_{Z_0 Z_i} \\ \sum_{i=1}^{n} \mu_i C_{Y_i Y_j} + \sum_{i=1}^{m} \lambda_i C_{Z_i Y_j} - \eta_2 = C_{Z_0 Y_j} \\ & \sum_{i=1}^{n} \lambda_i = 1 \\ & \sum_{i=1}^{n} \mu_i = 0 \end{cases}$$
(4)

 $\begin{pmatrix} C_{Z_{0}Y_{m}} & \cdots & C_{Z_{1}Z_{n}} & C_{Z_{1}Y} & \cdots & C_{Z_{1}Y_{m}} & 1 & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ C_{Z_{n}Z_{1}} & \cdots & C_{Z_{n}Z_{n}} & C_{Z_{n}Y_{1}} & \cdots & C_{Z_{n}Y_{m}} & 1 & 0 \\ C_{Y_{1}Z_{1}} & \cdots & C_{Y_{1}Z_{n}} & C_{Y_{1}Y_{1}} & \cdots & C_{Y_{1}Y_{m}} & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ C_{Y_{m}Z_{1}} & \cdots & C_{Y_{m}Z_{n}} & C_{Y_{m}Y_{1}} & \cdots & C_{Y_{m}Y_{m}} & 0 & 0 \\ 1 & \cdots & 1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 1 & 1 & 1 \end{pmatrix} \begin{bmatrix} \lambda_{1} \\ \vdots \\ \lambda_{n} \\ \mu_{1} \\ \vdots \\ \mu_{m} \\ -\xi_{1} \\ -\xi_{2} \end{bmatrix} = \begin{bmatrix} C_{Z_{0}Z_{1}} \\ \vdots \\ C_{Z_{0}Y_{1}} \\ \vdots \\ C_{Z_{0}Y_{m}} \\ 1 \\ 0 \end{bmatrix} .$ (5)

The cokriging method is based on the theory of co-regionalized variables. It uses cross-covariance and cross-semi-variance function modeling between multiple regionalized variables to characterize the correlation between them to obtain unbiased and optimal estimates of the regionalized variables. Because each environmental component has different effects on the effect variables of different parts, the environmental measurement value cannot be directly used as covariates to participate in the construction of the model. In this paper, by using the duration sequence of the measured value of each measuring point and the water pressure, temperature, rainfall, and aging, the non-linear fitting method was adopted to search for the most suitable non-linear expression of the measured value of each measured value of each measured value of the environmental variables at the moment to be calculated was brought into the function equation between them, and the change of the effect variable caused by the environmental variable was inferred, which was taken as a covariate to participate in the calculation, as seen in Equation (6).

$$\begin{bmatrix} f_{1}(h) & f_{1}(t) & f_{1}(p) & f_{1}(\theta) \\ f_{2}(h) & f_{2}(t) & f_{2}(p) & f_{2}(\theta) \\ \vdots & \vdots & \vdots & \vdots \\ f_{i}(h) & f_{i}(t) & f_{i}(p) & f_{i}(\theta) \\ \vdots & \vdots & \vdots & \vdots \\ f_{n}(h) & f_{n}(t) & f_{n}(p) & f_{n}(\theta) \end{bmatrix} \xrightarrow{h = h_{k+1}, t = t_{k+1}} \begin{bmatrix} Y_{1}(h) & Y_{1}(t) & Y_{1}(p) & Y_{1}(\theta) \\ Y_{2}(h) & Y_{2}(t) & Y_{2}(p) & Y_{2}(\theta) \\ \vdots & \vdots & \vdots & \vdots \\ Y_{i}(h) & Y_{i}(t) & Y_{i}(p) & Y_{i}(\theta) \\ \vdots & \vdots & \vdots & \vdots \\ Y_{n}(h) & Y_{n}(t) & Y_{n}(p) & Y_{n}(\theta) \end{bmatrix},$$
(6)

where *h*, *t*, *p*, and θ are the values of water pressure, temperature, rainfall, and aging; $f_i(h)$ is the most suitable nonlinear expression for the measured values of the duration series of measuring point *i* with respect to the measured values of the duration series of water pressure, same for $f_i(h)$, $f_i(p)$ and $f_i(\theta)$; h_{k+1} is the value of the water pressure at moment k + 1, same for t_{k+1} , p_{k+1} and θ_{k+1} ; $Y_i(h)$ is the water level covariate value of measuring point *i* at moment k + 1, same for $Y_i(t)$, $Y_i(p)$ and $Y_i(\theta)$.

In order to calculate the covariance, the main variable covariance function, co-variable covariance function, and cross-covariance function need to be fitted by selecting appropriate function models. The covariance function and the parameters to be fitted are shown in Figure 2.

Covariance function models commonly used in kriging theory include the exponential model, the spherical model, and the gaussian model [34]. As the exponential model has better generalizability compared to the spherical model, the Gaussian model may not be interpolable at certain locations. With overall consideration, the exponential model was chosen to fit the covariance function, as seen in Equation (7).

$$\gamma(h) = C_0 + C_1(1 - e^{-\frac{n}{a}}),\tag{7}$$

1.

where C_0 is nugget; $C_0 + C$ is sill; C is partial sill; a is not the range here, because when h = 3a there is $1 - e^3 \approx 0.95 \approx 1$, so $\gamma(h) \approx C_0 + C$, and the range is 3a. When $C_0 = 0$ and C = 1, it is called a standard exponential model.



Figure 2. Graph of covariance function.

2.4. Calculation of Unrepaired Value

The optimal main variable covariance function, co-variable covariance function, and cross-covariance function are fitted by the exponential model, and the values of main variable covariance, covariate covariance, and cross-covariance are calculated by Equation (8) according to the known distance between any two points. Then, the calculated results are brought into Equation (4), and the weights of each measuring point and each covariate are obtained by solving

$$\gamma(h) = \sigma^2 - C(h), \tag{8}$$

where σ^2 is the variance in the region *I*.

Finally, using the measured values $Z_1, Z_2, ..., Z_n$ and their weights $\lambda_1, \lambda_2, ..., \lambda_n$, the covariate values $Y_{j1}, Y_{j2}, ..., Y_{jn}$, and their weights $\mu_{j1}, \mu_{j2}, ..., \mu_{jn}$ for spatially similar intact measuring points, the repair values of data abnormal or missing measuring points can be calculated using Equation (2).

3. Discussion on the Threshold of Covariate Introduction

Reasonable selection of covariates affects the model's accuracy and computational efficiency, which is the key to constructing a cokriging spatial model. In order to ensure universality, the one-dimensional and the two-dimensional cokriging spatial model for different projects were constructed in this section. At the same time, the influence of different covariate combinations on the accuracy of the model was deeply analyzed, and the rationality of setting the VIP_i control threshold of 0.8 was demonstrated.

The error analysis of a single measuring point adopted the most commonly used cross-verification method [35]. *N* spatial models were obtained by sequentially leaving the measured values of each measuring point in a measurement time (*n* measuring points in total) as the verification set, and the measured value samples of the other n - 1 measuring points as the training set. The cross-validation results of all measuring points in the monitoring effect quantity were thus obtained, and the error analysis of each measuring point was carried out.

Overall error analysis represented the cross-validation results of all measuring points in each measurement by calculating the mean absolute error (MAE). It analyzed the error sequence of each measurement as a whole, as seen in Equation (9). The lower the MAE, the higher the overall accuracy.

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |Z_i - Z_i^*|,$$
(9)

where Z_i is the measured value of the measuring point *i*; Z_i^* is the repaired value of measuring point *i*.

3.1. Analysis of the Influence of Covariates Selection on the Accuracy of the One-Dimensional Spatial Model

A survey line was arranged at the elevation of 476.5 m upstream of a concrete face rockfill dam to monitor the vertical deformation of the surface, with a total of 10 measuring points, numbered SA6-1 to SA6-10. All measuring points are intact, and their arrangement and duration hydrograph are shown in Figures 3 and 4. The VIP method was used to identify the influencing factors of the effect variables of each measuring point. It was found that the important influencing factors of each measuring point were the same, aging and temperature had a great influence on vertical deformation, VIP_j value exceeded 0.8 followed by water pressure, the VIP_j value ranged from 0.25 to 0.64, and rainfall had the least influence; the VIP_j value was less than 0.3, as shown in Figure 5.



Figure 3. Layout of vertical deformation measuring points.

For the convenience of comparison and analysis, the aging single-factor, aging temperature two-factor, and aging temperature water pressure three-factor spatial cokriging models were constructed by using the measured values of each measuring point. The single-point error analysis was conducted for the measured values on 24 December 2020, as shown in Table 1. In the environmental variables, the aging is a monotonic increasing function, and the water level, temperature and rainfall are all periodic functions with years as the unit. Therefore, the overall error analysis was conducted for the measured values from January 2020 to December 2020 (due to COVID-19 pandemic, there was no measured value in February 2020), as shown in Figure 6. It could be seen that the repair accuracy of other measuring points was ideal, with absolute error within 5 mm and relative error within 3%, except for SA6-1 and SA6-10. Depending on the relative position of the points to be repaired in the group of known points, it can be divided into interpolation and extrapolation. Interpolation refers to repairing the unknown measuring points within the range, while extrapolation refers to repairing the unknown measuring points outside the range. The two measuring points SA6-1 and SA6-10 were located on the leftmost and rightmost banks in terms of spatial location. Hence, when cross-validation was carried out, the values at this point needed to be repaired by other deformation measurement points, and there were fewer measurement points around to refer to, which would lead to larger errors. On the contrary, SA6-3, SA6-7, and SA6-9 were located in the middle of the measuring points group, and the repair error was tiny. According to the error analysis of a single measuring point, the accuracy of the two-factor model of aging temperature was obviously better than that of the single-factor model of aging, and the absolute error of each measuring point could be reduced by more than 35%. The accuracy of the three-factor model of aging temperature water pressure was slightly lower than that of the two-factor model of aging temperature, and the repair accuracy of the three-factor model of about

60% of the measuring points was lower than that of the two-factor model. The overall error analysis showed that the mean absolute error of the two-factor model was $3.85 \sim 4.08$ mm, which was on average 44% lower than the mean absolute error of the single-factor model ($6.59 \sim 7.37$ mm). The mean absolute error of the three-factor model was slightly higher than that of the two-factor model, ranging from 3.90 mm to 4.20 mm. Therefore, introducing a temperature factor with the VIP_j greater than 0.8 as a covariate could significantly improve the accuracy of the model. Nevertheless, if the water level factor with the VIP_j less than 0.8 was added on this basis, only the measuring points SA6-2, SA6-4 and SA6-8 had a small improvement in model accuracy, while the other measuring points showed a decline in model accuracy.



Figure 4. Comparison between measured value and repaired value of vertical deformation measuring points.

In order to verify the stability of the temperature aging two-factor VIP-cokriging model, according to the monitoring sequence, one measurement was selected every year for cross-verification. The calculation results are shown in Figure 4. From the figure, we could see that except the edge measuring points SA6-1 and SA6-10, the repair accuracy of other measuring points was ideal, and the relative error was basically controlled within 5%. It further proved that the accuracy of the model was relatively stable. There is no doubt that the framework cannot provide good predictions at one time interval and poor results at another.



Figure 5. Results of the variable importance for projection of vertical deformation.





3.2. Analysis of the Influence of Covariates Selection on the Accuracy of Two-Dimensional Spatial Model

One internal horizontal displacement measuring line was set at the elevations of 346.00 m, 379.00 m, 404.00 m, and 445.00 m on the left 0 + 008.20 m section of a concrete face rockfill dam, with a total of 20 measuring points, numbered EXa1-1~EXa1-2, EXa2-1~EXa2-4, EXa3-1~EXa3-6, and EXa4-1~EXa4-8, as shown in Figure 7.



Figure 7. Layout of horizontal deformation measuring points.

Table 1. Spatial repair results with selection of different covariates of the one-dimensional model.

Measuring Points	Measured Value (mm)	Aging		Aging + Ter	mperature	Aging + Temperature + Water Pressure	
		Repaired Value (mm)	Relative Error	Repaired Value (mm)	Relative Error	Repaired Value (mm)	Relative Error
SA6-1	39.09	58.28	49.09%	50.51	29.23%	51.25	31.11%
SA6-2	85.12	90.10	5.85%	87.74	3.08%	87.26	2.52%
SA6-3	128.41	126.78	1.27%	127.85	0.44%	126.18	1.74%
SA6-4	176.22	170.73	3.11%	173.15	1.74%	174.14	1.18%
SA6-5	194.3	186.64	3.94%	190.19	2.12%	191.80	1.29%
SA6-6	198.11	190.13	4.03%	193.97	2.09%	193.63	2.26%
SA6-7	187.09	180.78	3.37%	183.60	1.86%	182.30	2.56%
SA6-8	150.56	146.17	2.91%	147.58	1.98%	149.08	0.98%
SA6-9	93.09	95.42	2.50%	93.50	0.44%	92.33	0.81%
SA6-10	71.71	85.51	19.24%	79.68	11.11%	81.12	13.12%
MAE	MAE (mm) 7.37		4.0	18	4.20		

The variable importance for the projection method was used to identify the influence factors of effect variables at each measuring point. It was found that the important influence factors at each measuring point were consistent. The effects of aging and water pressure were more remarkable, with the VIP_j values being all greater than 1, while the effects of temperature and rainfall were relatively small, with the VIP_j values being basically between 0.4 and 0.6, as shown in Figure 8.

As above, the spatial cokriging models of aging single-factor, aging water pressure two-factor, and aging water pressure rainfall three-factor were constructed by using the measured values of each measuring point. The single-point error analysis was conducted for the measured values on 11 June 2020, as shown in Table 2. The overall error analysis was conducted for the measured values from January 2020 to December 2020, as shown in Figure 9. It could be seen that the maximum relative error appeared at EXa3-4, which was analyzed to be caused by the small measured value. The analysis combined interpolation and extrapolation and found that although measuring points such as EXa1-1 and EXa4-1 were marginal points, there were more intact measuring points around them, so the

extrapolation of the two-dimensional model was better than that of the one-dimensional model. According to the single measurement point error analysis, the accuracy of the two-factor cokriging model is significantly improved compared to the accuracy of the single-factor cokriging model, which is difficult to ignore, and the absolute error of each measurement point was reduced by more than 80%. The accuracy of the three-factor cokriging model is comparable to that of the two-factor cokriging model, the relative error reduction is not significant, and the accuracy of the two-factor cokriging model is slightly better than that of the three-factor cokriging model for the measuring points EXa2-1 and EXa3-1. The overall error analysis showed that the mean absolute error of the two-factor model (0.07~0.23 mm) was significantly lower than the mean absolute error of the one-factor model (3.95~4.22 mm), while the mean absolute error of the three-factor model (0.06~0.23 mm) was less different from that of the two-factor model. Thus, by analyzing the repair effects of introducing different covariates in the one-dimensional and two-dimensional models, the results obtained for both were consistent.



Measuring point



Figure 8. Results of the variable importance for projection of horizontal deformation.

Figure 9. Comparison of MAE with selection of different covariates of the two-dimensional model.

Measuring	Measured Value (mm)	Agi	ng	Aging + Wat	er Pressure	Aging + Water Pressure + Rainfall		
Points		Repaired Value (mm)	Relative Error	Repaired Value (mm)	Relative Error	Repaired Value (mm)	Relative Error	
EXa1-1	15.45	11.18	27.67%	15.57	0.80%	15.57	0.77%	
EXa1-2	11.68	9.42	19.35%	11.85	1.47%	11.85	1.43%	
EXa2-1	15.93	8.00	49.78%	15.76	1.06%	15.75	1.13%	
EXa2-2	12.47	7.89	36.75%	12.41	0.45%	12.42	0.39%	
EXa2-3	17.27	11.30	34.55%	17.11	0.91%	17.12	0.87%	
EXa2-4	12.93	9.34	27.76%	12.91	0.17%	12.91	0.18%	
EXa3-1	-10.82	-3.37	68.85%	-10.86	0.41%	-10.87	0.48%	
EXa3-2	-9.86	-1.89	80.80%	-9.62	2.47%	-9.59	2.69%	
EXa3-3	-6.13	-0.38	93.79%	-6.04	1.43%	-6.05	1.25%	
EXa3-4	1.18	4.02	240.95%	1.47	24.68%	1.46	23.76%	
EXa3-5	10.87	7.92	27.18%	10.86	0.12%	10.86	0.13%	
EXa3-6	13.11	8.53	34.94%	13.04	0.56%	13.04	0.56%	
EXa4-1	-2.57	0.69	126.76%	-2.53	1.54%	-2.53	1.56%	
EXa4-4	-7.93	-3.19	59.81%	-7.93	0.03%	-7.93	0.02%	
EXa4-5	-6.09	-1.28	79.06%	-6.17	1.36%	-6.17	1.34%	
EXa4-6	3.37	4.04	19.86%	3.50	3.84%	3.50	3.84%	
EXa4-7	3.59	4.58	27.47%	3.67	2.27%	3.67	2.27%	
EXa4-8	5.86	4.56	22.15%	5.79	1.27%	5.79	1.23%	
MAE (mm)		4.22		0.1	.0	0.10		

Table 2. Spatial repair results with selection of different covariates of the two-dimensional model.

Based on the above, taking all environmental variables with the VIP_j greater than 0.8 as covariates to construct a cokriging spatial model could greatly improve the accuracy of the model. However, if the factor with the VIP_j value less than 0.8 was introduced as a covariate, there was no obvious effect on improving the model's accuracy, and it might even reduce the accuracy of the model at some measuring points, and it would significantly increase the calculation amount and reduce the work efficiency. Therefore, considering the accuracy and technical efficiency of the model, it was reasonable to introduce the VIP_j value greater than 0.8 as the environmental factor of the cokriging spatial model.

4. Verification and Analysis of Model Accuracy

To further analyze and validate the rationality and validity of the VIP-cokriging spatial model, the one-dimensional model case from Section 3.1 was used as an example in this section. The aging temperature two-factor cokriging spatial model was compared with the inverse distance weighting model [36], ordinary kriging model [37], and universal kriging model [38], which were more commonly used in the construction of spatial models for dam safety monitoring at present.

In this paper, the data of seven measurements from June 2020 to December 2020 of this measuring line were cross-repaired. The error statistics of the single point at typical moments are shown in Table 3, and the overall error is shown in Figure 10. As can be seen from the graph, the inverse distance weighting model was the worst, with the maximum relative error of edge measuring points reaching 178.73%, and the relative error of non-edge measuring points basically exceeding 10%, with the mean absolute error of 22.09 to 23.54 mm. The analysis of the reason should be related to the fact that the method only considered the influence of distance on unknown points while ignoring the

correlation between measuring points. The ordinary kriging model took into account the autocorrelation between each measuring point in space, and its accuracy was significantly improved compared with the inverse distance weighting model. The relative errors of its non-edge measurement points were controlled within 10%, and the mean absolute error was 16.74~17.79 mm. This method considered that the variables were stationary in the region, that is, there was no trend quantity, whereas the vertical displacement of dam appearance generally had a growing trend. Therefore, the universal kriging model, which introduced the trend variables over the whole spatial distribution, was relatively good, with the mean absolute error of 11.46 to 12.68 mm.

Measuring Points	Measured ⁻ Value (mm)	IDW		Ordinary Kriging		Universal Kriging		VIP-Cokriging	
		Repaired Value (mm)	Relative Error	Repaired Value (mm)	Relative Error	Repaired Value (mm)	Relative Error	Repaired Value (mm)	Relative Error
SA6-1	39.09	108.96	178.73%	108.90	178.58%	52.61	34.59%	50.51	29.23%
SA6-2	85.12	104.76	23.07%	89.91	5.63%	90.03	5.77%	87.74	3.08%
SA6-3	128.41	121.42	5.45%	124.27	3.23%	115.36	10.17%	127.85	0.44%
SA6-4	176.22	163.89	7.00%	164.10	6.88%	168.71	4.26%	173.15	1.74%
SA6-5	194.3	167.87	13.60%	178.28	8.25%	182.82	5.91%	190.19	2.12%
SA6-6	198.11	173.32	12.51%	183.11	7.57%	159.29	19.60%	193.97	2.09%
SA6-7	187.09	170.71	8.76%	175.31	6.30%	177.96	4.88%	183.60	1.86%
SA6-8	150.56	130.33	13.43%	137.23	8.86%	136.33	9.45%	147.58	1.98%
SA6-9	93.09	80.71	13.30%	89.98	3.34%	95.58	2.67%	93.50	0.44%
SA6-10	71.71	98.06	36.74%	99.46	38.70%	83.40	16.30%	79.68	11.11%

Table 3. Comparison between measured values and repaired values of different models (24 December 2020).



Figure 10. Comparison of MAE of different models.

Overall, the VIP-cokriging model had the highest accuracy. The relative error of non-edge measuring points was mostly controlled within 3%, and the relative error of the edge measuring points was reduced compared with the other three models. Its mean absolute error was 3.85 to 4.08 mm, which was 83%, 77%, and 67% lower than that of the inverse distance weighting model, the ordinary kriging model, and the universal kriging model. The main reason was that cokriging introduced one or more auxiliary environmental variables that had a strong correlation with the effect variables, and using the cross-correlation between them could effectively improve the spatial expansion accuracy of the main variable at the unknown points, which could solve the problem of the insufficient number of measuring points.

5. Engineering Application

The PBG gravel soil earth core wall rockfill dam with a maximum height of 186 m was completed and stored in water for power generation in 2009. Its internal vertical displacement was monitored by the hydraulic overflow settlement gauge, and the measuring points of the typical profile (profile 0 + 240 m) were arranged, as shown in Figure 11. Due to the impact of temporary dam top removal and permanent dam top construction in 2012, there were many missing measurements between 2012 and 2013. Since completing the automation transformation in August 2013, the continuity of measured values was good. However, there were still some phenomena such as unstable measured values at some measuring points and blockage of instrument pipelines. For example, the CH10 and CH13 at the elevation of 725 m were damaged and measuring stopped in 2015. The stability of the CH14 measured value was poor, with an anomalous steep step from July 2013 to July 2014. At present, the measured value is close to the measuring range of the instrument, and the reliability of the measured value is poor, as shown in Figure 12.



Figure 11. Layout of the hydraulic overflow settlement gauge at profile 0 + 240 m.

To control the overall distribution characteristics of vertical displacement at profile 0 + 240 m, the original monitoring data of all measuring points are used to draw an isoline map of settlement deformation at a typical high water level moment (12 October 2017), as shown in Figure 13a. It could be seen from the figure that the settlement deformation of the profile showed an increasing downward trend from the middle of the dam body, and the maximum settlement deformation occurred above the dam foundation. It was inconsistent with the law that the maximum settlement deformation of gravel soil core wall

rockfill dam generally occurs in the area of $1/3\sim 2/3$ dam height. Analysis of the reasons should be related to the lack of measuring points at 731 m elevation and the unreliable measurement of CH14. To avoid misjudgment of dam operation behavior, a VIP-cokriging two-dimensional spatial model of internal vertical displacement at profile 0 + 240 m was constructed in this paper. The model was used to supplement the missing measured values of CH10 and CH13, and to repair the measured values of CH14.



Figure 12. Duration hydrograph of the hydraulic overflow settlement gauge at 725 m elevation at profile 0 + 240 m.

By analyzing the variable importance for projection, as shown in Figure 14, the cokriging spatial model is constructed by selecting aging and water pressure as covariates. To further verify the accuracy and stability of the model, cross-validation was carried out by selecting the moments when all measurement points were intact. Except for the points CH8 and CH16, the relative errors of the remaining measurement points ranged from 0.24% to 1.90%, especially the relative errors of CH10, CH13, and CH14 measurement points were minor and controlled within 2%; the calculation results are shown in Figure 15 and Table 4. The reasons for it should be related to the fact that the points CH13 and CH14 were internal measuring points and that the CH10 point was close to the known measuring points and had more known measuring points around it. Therefore, it was feasible to use the VIP-cokriging spatial model to repair the data of measuring points.



Figure 13. Isoline map of the internal settlement deformation of the dam at profile 0 + 240 m (12 October 2017).



Figure 14. Results of the variable importance for projection of horizontal deformation at profile 0 + 240 m.



Figure 15. Cross-validation results of vertical displacement.

The duration hydrograph of the repaired measuring points is shown in Figure 12, and the isoline map of the vertical displacement of the profile is shown in Figure 13b. The pattern of changes in the measured values of the repaired points was the same as that before 2012, with an overall convergence and stabilization and a high degree of restoration accuracy. The overall pattern of vertical deformation of the profile showed that the internal settlement deformation was greater than the surface at the same elevation. The settlement was mainly controlled by the section thickness; the closer to the core wall, the greater the settlement relative to other locations at the same elevation. The maximum settlement deformation of the dam occurred in the middle and lower part of the dam, about 1/3 to 2/3 of the dam height area, which was consistent with the general settlement deformation law for gravel soil core wall rockfill dams.

		CH10		CH13			CH14		
Date	Measured Value(mm)	Repaired Value(mm)	Relative Error	Measured Value(mm)	Repaired Value(mm)	Relative Error	Measured Value(mm)	Repaired Value(mm)	Relative Error
29 January 2010	1789	1785.42	0.20%	2131	2124.25	0.32%	2011	2010.56	0.02%
27 March 2010	1802	1797.24	0.26%	2146	2153.75	0.36%	2025	2025.65	0.03%
31 May 2010	1813	1816.59	0.20%	2162	2164.88	0.13%	2042	2042.52	0.03%
31 July 2010	1829	1820.88	0.44%	2191	2190.90	0.00%	2072	2071.50	0.02%
30 September 2010	1835	1834.23	0.04%	2240	2236.86	0.14%	2112	2112.92	0.04%
30 November 2010	1857	1860.65	0.20%	2288	2288.13	0.01%	2149	2150.07	0.05%
31 January 2011	1871	1874.07	0.16%	2313	2304.11	0.38%	2168	2171.76	0.17%
30 March 2011	1886	1907.13	1.12%	2316	2296.84	0.83%	2171	2172.47	0.07%
31 May 2011	1890	1921.36	1.66%	2322	2304.37	0.76%	2180	2172.82	0.33%
29 July 2011	1890	1921.11	1.65%	2327	2320.42	0.28%	2181	2184.31	0.15%
29 September 2011	1916	1931.27	0.80%	2359	2332.46	1.13%	2203	2202.86	0.01%
25 November 2011	1925	1930.84	0.30%	2390	2362.19	1.16%	2221	2219.99	0.05%

Table 4. Repair accuracy of measuring points.

6. Conclusions

In this paper, aiming at the problems of the spatial models only considering the geometric position of measuring points, or the data repair mostly adopting time series models which may lead to poor accuracy, a spatial model for repairing dam safety monitoring data based on VIP-cokriging was proposed. The construction of the model included the steps of the decision of the construction dimension of the spatial model, selection of covariates, fitting of covariance function, and calculation of unrepaired value. Finally, the accuracy and applicability of the model were verified with engineering examples, and the following conclusions can be drawn:

(1) Aiming at the problem of the spatial model for repairing dam safety monitoring data not considering the correlation with environmental variables, a cokriging spatial model based on the variable importance for projection was proposed. Firstly, the spatial model construction dimension was determined according to the spatial arrangement characteristics of similar monitoring instruments. Secondly, based on partial least squares regression, principal component analysis and typical correlation analysis were introduced to decompose and filter the information, and important environmental impact factors were identified and calculated as covariates. Finally, the weights of the same kind of intact measuring points and the weights of the covariates calculated by fitting the covariance function were brought into the model to calculate the repair value of the abnormal or missing measuring points.

(2) Based on engineering cases of the one-dimensional and the two-dimensional models, the relationship between the accuracy of the VIP-cokriging model and the selection of covariates was analyzed. By comparing the repair accuracy of different covariate schemes chosen, the reasonableness of the VIP_j control threshold setting of 0.8 was demonstrated. Choosing the environmental variable with the VIP_j value greater than 0.8 to participate in the calculation could significantly improve the model's accuracy. Otherwise, it might increase the calculation amount and reduce the work efficiency without improving the model's accuracy. Moreover, the VIP-cokriging model has strong applicability and high accuracy at any time.

(3) The VIP-cokriging spatial repair model can greatly improve the data repair effect and has the advantages of high precision and strong applicability. The application showed that VIP-cokriging effectively improved the accuracy of the spatial expansion of the main variables at the unknown point and reduced the error range at repair points by taking into account the influence of important environmental variables and utilizing the crosscorrelation between environmental variables and effect variables. Compared with the inverse distance weight model, the ordinary kriging model, and the universal kriging model, the repair error was reduced by 83%, 77%, and 67%.

(4) Aiming at the problems of the damage of the monitoring points and the instability of the measured values at the profile 0 + 240 m of the PBG dam, a two-dimensional VIP-cokriging spatial model was adopted to extend the monitoring effect variables to the whole area of the dam space. The scientific correlation between existing monitoring points and the unknown spatial location of the dam was achieved, and the problem of repairing missing or abnormal data at key measuring points was solved. The construction of the model was beneficial to reasonably grasping the overall distribution of dam deformation, effectively avoiding the misjudgment of the safe operation of the dam, and had good application value in engineering.

(5) The construction of three-dimensional model needs three-dimensional coordinates of monitoring equipment position information, but the cokriging method is only suitable for two-dimensional plane coordinate system, and kriging interpolation in three-dimensional space is very complex and difficult to realize. The covariance of the kriging interpolation method is a function based on the distance between measuring points. When constructing the three-dimensional model of dam safety monitoring, the three-dimensional coordinates can be transformed into two-dimensional coordinates, but the two-dimensional coordinates will possibly lose the anisotropy in the monitoring effect variables, which would lead to the calculation results that cannot reflect the real state of the dam. Therefore, in the follow-up study, the cokriging method should be expanded in three-dimensional space to build a more perfect three-dimensional repair model.

Author Contributions: Conceptualization, S.L. and K.L.; methodology, S.L. and K.L.; software, X.L.; validation, S.L.; formal analysis, Y.L.; investigation, Z.W.; resources, L.P.; data curation, S.L.; writing—original draft preparation, S.L., Y.L. and K.L.; writing—review and editing, X.L., Z.W. and L.P.; supervision, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: Supported by the Fundamental Research Funds for the Central Universities (YJ202287).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

- Salehi, H.; Burgueño, R. Emerging artificial intelligence methods in structural engineering. *Eng. Struct.* 2018, 171, 170–189. [CrossRef]
- Li, Y.; Bao, T.; Chen, H.; Zhang, K.; Shu, X.; Chen, Z.; Hu, Y. A large-scale sensor missing data imputation framework for dams using deep learning and transfer learning strategy. *Measurement* 2021, 178, 109377. [CrossRef]
- 3. Chen, B.; Huang, Z.; Liu, C.; Wu, Z. Spatio-temporal data mining method for joint cracks in concrete dam based on association rules. *Struct. Control Health Monit.* **2021**, *29*, e2848. [CrossRef]
- Salazar, F.; Morán, R.; Toledo, M.Á.; Oñate, E. Data-Based Models for the Prediction of Dam Behaviour: A Review and Some Methodological Considerations. Arch. Comput. Methods Eng. 2015, 24, 1–21. [CrossRef]
- Su, H.; Li, X.; Yang, B.; Wen, Z. Wavelet support vector machine-based prediction model of dam deformation. *Mech. Syst. Signal Process.* 2018, 110, 412–427. [CrossRef]
- Yan, J.; Li, H.; Bai, Y.; Lin, Y. Spatial-Temporal Traffic Flow Data Restoration and Prediction Method Based on the Tensor Decomposition. *Appl. Sci.* 2021, 11, 9220. [CrossRef]
- Lin, C.; Li, T.; Chen, S.; Liu, X.; Lin, C.; Liang, S. Gaussian process regression-based forecasting model of dam deformation. *Neural Comput. Appl.* 2019, *31*, 8503–8518. [CrossRef]
- Lee, Y.; Choi, S.-I. Data Restoration by Linear Estimation of the Principal Components from Lossy Data. *IEEE Access* 2020, 8, 172244–172251. [CrossRef]
- 9. Salazar, F.; Conde, A.; Irazábal, J.; Vicente, D.J. Anomaly Detection in Dam Behaviour with Machine Learning Classification Models. *Water* 2021, *13*, 2387. [CrossRef]

- 10. Kang, F.; Liu, J.; Li, J.; Li, S. Concrete dam deformation prediction model for health monitoring based on extreme learning machine. *Struct. Control Health Monit.* **2017**, 24, e1997. [CrossRef]
- 11. Vazifehdan, M.; Moattar, M.H.; Jalali, M. A hybrid Bayesian network and tensor factorization approach for missing value imputation to improve breast cancer recurrence prediction. *J. King Saud Univ.—Comput. Inf. Sci.* **2019**, *31*, 175–184. [CrossRef]
- 12. Stojanovic, B.; Milivojevic, M.; Ivanovic, M.; Milivojevic, N.; Divac, D. Adaptive system for dam behavior modeling based on linear regression and genetic algorithms. *Adv. Eng. Softw.* **2013**, *65*, 182–190. [CrossRef]
- 13. Du, J.; Chen, H.; Zhang, W. A deep learning method for data recovery in sensor networks using effective spatio-temporal correlation data. *Sens. Rev.* **2019**, *39*, 208–217. [CrossRef]
- 14. Li, Y.; Min, K.; Zhang, Y.; Wen, L. Prediction of the failure point settlement in rockfill dams based on spatial-temporal data and multiple-monitoring-point models. *Eng. Struct.* **2021**, *243*, 112658. [CrossRef]
- Liu, Y.-W.; Wang, W.; Hu, Y.-M.; Liang, Z.-M. Drought assessment and uncertainty analysis for Dapoling basin. *Nat. Hazards* 2014, 74, 1613–1627. [CrossRef]
- Lu, G.Y.; Wong, D.W. An adaptive inverse-distance weighting spatial interpolation technique. *Comput. Geosci.* 2008, 34, 1044–1055. [CrossRef]
- Ballarin, F.; D'Amario, A.; Perotto, S.; Rozza, G. A POD-selective inverse distance weighting method for fast parametrized shape morphing. *Int. J. Numer. Methods Eng.* 2019, 117, 860–884. [CrossRef]
- 18. Bostan, P.A.; Heuvelink, G.B.M.; Akyurek, S.Z. Comparison of regression and kriging techniques for mapping the average annual precipitation of Turkey. *Int. J. Appl. Earth Obs. Geoinf.* **2012**, *19*, 115–126. [CrossRef]
- 19. Liu, D.; Li, Z.; Lian, Z. Compaction quality assessment of earth-rock dam materials using roller-integrated compaction monitoring technology. *Autom. Constr.* **2014**, *44*, 234–246. [CrossRef]
- Michael, R.; O'Lenick, C.R.; Monaghan, A.; Wilhelmi, O.; Wiedinmyer, C.; Hayden, M.; Estes, M. Application of geostatistical approaches to predict the spatio-temporal distribution of summer ozone in Houston, Texas. *J. Expo. Sci. Environ. Epidemiol.* 2019, 29, 806–820. [CrossRef]
- Buttafuoco, G.; Conforti, M. Improving Mean Annual Precipitation Prediction Incorporating Elevation and Taking into Account Support Size. Water 2021, 13, 830. [CrossRef]
- AbdelRahman, M.A.E.; Zakarya, Y.M.; Metwaly, M.M.; Koubouris, G. Deciphering Soil Spatial Variability through Geostatistics and Interpolation Techniques. *Sustainability* 2021, 13, 194. [CrossRef]
- Sajal Kumar Adhikary, N.M. Ordinary kriging and genetic programming for spatial estimation of rainfall in the Middle Yarra River catchment, Australia. *Hydrol. Res.* 2016, 47, 1182–1197. [CrossRef]
- 24. Seo, Y.; Kim, S.; Singh, V.P. Estimating Spatial Precipitation Using Regression Kriging and Artificial Neural Network Residual Kriging (RKNNRK) Hybrid Approach. *Water Resour. Manag.* **2015**, *29*, 2189–2204. [CrossRef]
- Mao, Y.; Zhang, J.; Qi, H.; Wang, L. DNN-MVL: DNN-Multi-View-Learning-Based Recover Block Missing Data in a Dam Safety Monitoring System. Sensors 2019, 19, 2895. [CrossRef]
- Zhao, E.; Wu, C. Centroid deformation-based nonlinear safety monitoring model for arch dam performance evaluation. *Eng. Struct.* 2021, 243, 112652. [CrossRef]
- Lu, X.; Wu, Z.; Zhou, Z.; Chen, J. Research on the Prediction Model of Deformation of High Core Rockfill Dam During Construction Period. *Adv. Eng. Sci.* 2017, 49, 61–69. Available online: http://jsuese.cnjournals.com/html/2017/4/201601342.html (accessed on 4 June 2021). (In Chinese).
- Dai, W.; Liu, N.; Santerre, R.; Pan, J. Dam Deformation Monitoring Data Analysis Using Space-Time Kalman Filter. *ISPRS Int. J. Geo-Inf.* 2016, 5, 236. [CrossRef]
- 29. Yang, Z.; Dai, W.; Santerre, R.; Kuang, C.; Shi, Q. A Spatiotemporal Deformation Modelling Method Based on Geographically and Temporally Weighted Regression. *Math. Probl. Eng.* **2019**, 2019, 1–11. [CrossRef]
- 30. Ren, Q.; Li, M.; Song, L.; Liu, H. An optimized combination prediction model for concrete dam deformation considering quantitative evaluation and hysteresis correction. *Adv. Eng. Inform.* **2020**, *46*, 101154. [CrossRef]
- 31. Farrés, M.; Platikanov, S.; Tsakovski, S.; Tauler, R. Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. *J. Chemom.* **2015**, *29*, 528–536. [CrossRef]
- Chen, X.; Huang, J.; Yi, M. Cost estimation for general aviation aircrafts using regression models and variable importance in projection analysis. J. Clean. Prod. 2020, 256, 120648. [CrossRef]
- Neale, C.M.U.; Maltese, A.; He, P.; Xu, X.; Zhang, B.; Li, Z.; Feng, H.; Yang, G.; Zhang, Y. Estimation of leaf chlorophyll content in winter wheat using variable importance for projection (VIP) with hyperspectral data. *Remote Sens.* 2015, 9637, 963708. [CrossRef]
- 34. Huang, C.; Zhang, H.; Robeson, S.M. On the Validity of Commonly Used Covariance and Variogram Functions on the Sphere. *Math. Geosci.* 2011, 43, 721–733. [CrossRef]
- Kim, H.-S.; Chung, C.-K.; Kim, H.-K. Geo-spatial data integration for subsurface stratification of dam site with outlier analyses. Environ. Earth Sci. 2016, 75, 168. [CrossRef]
- 36. Ozelkan, E.; Bagis, S.; Ozelkan, E.C.; Ustundag, B.B.; Yucel, M.; Ormeci, C. Spatial interpolation of climatic variables using land surface temperature and modified inverse distance weighting. *Int. J. Remote Sens.* **2015**, *36*, 1000–1025. [CrossRef]

- 37. Dai, F.; Zhou, Q.; Lv, Z.; Wang, X.; Liu, G. Spatial prediction of soil organic matter content integrating artificial neural network and ordinary kriging in Tibetan Plateau. *Ecol. Indic.* **2014**, *45*, 184–194. [CrossRef]
- 38. Brus, D.J.; Heuvelink, G.B.M. Optimization of sample patterns for universal kriging of environmental variables. *Geoderma* **2007**, 138, 86–95. [CrossRef]