

Article

# Spatiotemporal Interpolation for Environmental Modelling

Ferry Susanto <sup>1,2,\*</sup>, Paulo de Souza Jr. <sup>1</sup> and Jing He <sup>2</sup>

<sup>1</sup> Data61, CSIRO, College Road, Sandy Bay TAS 7005, Australia; Paulo.desouza@data61.csiro.au

<sup>2</sup> College of Engineering and Science, Victoria University, Footscray VIC 3011, Australia; Jing.He@vu.edu.au

\* Correspondence: Ferry.Susanto@data61.csiro.au; Tel.: +61-3-62375634

Academic Editor: Leonhard M. Reindl

Received: 22 June 2016; Accepted: 2 August 2016; Published: 6 August 2016

**Abstract:** A variation of the reduction-based approach to spatiotemporal interpolation (STI), in which time is treated independently from the spatial dimensions, is proposed in this paper. We reviewed and compared three widely-used spatial interpolation techniques: ordinary kriging, inverse distance weighting and the triangular irregular network. We also proposed a new distribution-based distance weighting (DDW) spatial interpolation method. In this study, we utilised one year of Tasmania's South Esk Hydrology model developed by CSIRO. Root mean squared error statistical methods were performed for performance evaluations. Our results show that the proposed reduction approach is superior to the extension approach to STI. However, the proposed DDW provides little benefit compared to the conventional inverse distance weighting (IDW) method. We suggest that the improved IDW technique, with the reduction approach used for the temporal dimension, is the optimal combination for large-scale spatiotemporal interpolation within environmental modelling applications.

**Keywords:** spatiotemporal interpolation; ordinary kriging; inverse distance weighting; triangular irregular network; distribution-based distance weighting

---

## 1. Introduction

Environmental sensor networks offer a significant contribution to our society, and the data collected are crucial for environmental managers to support decisions for the effective use of natural resources. Applications include environmental monitoring for hazard warning services (e.g., cyclone, flood and bushfire) [1], which introduces the need for environmental scientists to have spatially- and temporally-continuous data flows to support accurate and justified interpretations [2].

However, because of current technology limitations, we are not yet able to obtain high quality data in terms of spatial and temporal coverage [3], especially in mountainous and deep marine regions. The main restriction is the cost to set up comprehensive sensor networks in these environments and in a real-time manner. It is arguable whether it is necessary to set up such costly networks to obtain high quality data; while we could set up the network only at some critical points and apply appropriate interpolation technique to estimate the remaining unmeasured locations.

Spatial interpolation studies have been conducted in many environmental projects, involving rainfall, temperature and air pollution data. Table 1 lists some of the reviewed papers that make use of different interpolation methods for different purposes (Table 2 provides the list of interpolation abbreviations used within the table). Li and Heap [2] have provided a comprehensive review of the most commonly-used spatial interpolation techniques used in different areas: environment, hydrology, agriculture, forestry and engineering. A list of statistical error measurements to evaluate the interpolation model are provided in their work.

**Table 1.** List of previous studies that utilised different interpolation techniques for environmental studies, and the findings from each study are also provided. Comparing methods as done here cannot be generalized, as this is limited to the empirical exercises, and no rigorous theoretical demonstration is provided (the table is sorted based on the similarity of interpolation techniques adopted and compared in each study).

Process	Techniques	Findings	Ref.
Spatial rainfall mapping using IDW in the middle of Taiwan.	IDW	- IDW can be improved by the adjustment of the distance-decay parameter and the search radius. - Does not have significant interpolating ability for extreme values.	[4]
Compared to the developed CAR method with IDW for climate datasets in China.	IDW, CAR	- High-elevation data are an important factor for meteorological studies. - CAR performs slightly better than GIDW in terms of objective comparisons, especially in estimating local neighbouring patterns.	[5]
Comparing different techniques for precipitation and elevation.	IDW, AIDW, Kriging	- Varying the distance-decay parameter of IDW based on the spatial pattern can improve overall performance (AIDW). - AIDW can perform better than kriging in some cases.	[6]
Proposing regression-based IDW and comparing it with IDW and kriging.	IDW, RIDW, Kriging	- Integrating linear regression in IDW provides comparable objective evaluation to kriging and is computationally less demanding. - The confidence interval (CI) of RIDW also surpasses kriging CI.	[7]
Evaluating different interpolation techniques for air temperature data.	Spline, IDW, Kriging	- Kriging performs better overall, followed by IDW and spline.	[8]
Comparing different techniques for rainfall mapping in Sri Lanka.	IDW, TPS, OK, BK	- Interpolation results are very much dependent on the settings. - Different methods with different settings must be tested to define the suitable technique. - Bayesian kriging and splines performed best overall.	[9]
Assessing different interpolation methods to define the most suitable technique for the McArthur Forest Fire Danger Index (FFDI).	IDW, OK, RF, RFOK, RFIDW	- Combination of methods: RFOK and RFIDW shows the most promising results (least error). - Fire danger index is highly related to the behaviour of climate change, and should be considered carefully.	[1]
Investigated several techniques for depth to underground water in northwest China.	IDW, RBF, OK, SK, UK	- Simple kriging is the optimal method in terms of result consistency and the smallest prediction interval of 95%. - Depth of underground water increases significantly over the year because of excessive exploitation.	[10]
Ranking spatial interpolation techniques using GIS-based DSS.	Spline, IDW, Kriging, TP, TS	- No optimal technique that can accurately predict the rainfall. - Performance of each technique depends on the scale of the input data. - Kriging is the recommended technique, as it produces the most consistent results.	[11]
Using an interpolation method to construct a comprehensive archive of Australian climate data.	TPS, OK	- Different climate variables can be more accurately interpolated using different techniques due to the characteristic variability.	[3]

Based on the findings in Table 1, kriging has been reported as the most suitable spatial interpolation method overall [8,10,11], yet it suffers from the drawback that it is very computationally demanding. Subsequently, IDW has been extensively used and improved in many studies because of its advantages of efficiency and comparable performance with kriging [4,6,7]. Sanabria et al. have shown that using a hybrid method (RFOK and RFIDW) can enhance the quality of the interpolator [1].

None of the methods is uniquely optimal for any particular situation [12,13]. These factors are described in detail by Li and Heap [2]: (a) the sampling design and the samples' distribution; (b) the data nature and quality; (c) the correlation between primary and secondary variables. Therefore, the best technique for a specific purpose can only be obtained by comparing them and selecting the one that is most appropriate for the particular task [3,9,11].

Nevertheless, not much research has been reported on interpolation methods that include the temporal dimension: spatiotemporal interpolation (STI) techniques. Li proposed using triangular meshes to perform spatiotemporal interpolation to measure the housing prices in Lincoln, Nebraska [14]. Then, a triangulation-based STI method was applied to the air pollution study as reported in 2011 [15]. Later in 2014, Li and collaborators utilised an IDW-based STI to further improve interpolation performance by increasing computational efficiency [16].

Kilibarda and collaborators utilised a geostatistical STI method to model global surface temperature [17]. They indicate that this kind of STI method also facilitates computational efficiency by creating a single semivariogram for the entire year, instead of 365 semivariograms for each day in a year. Such an interpolator does not show abrupt jumps between days that have no physical basis.

### 1.1. Objective

In this study, we propose a new spatiotemporal interpolation (STI) technique with the following features:

1. Computationally effective: The sensor network is assumed to collect hundreds (or even thousands) of data items every second, and thus, a fast algorithm is critical for effective processing.
2. The ability to obtain close-to-reality measurements (i.e., low statistical error and visually not-abrupt: To ensure environmental managers can make confident interpretations at a later stage.

**Table 2.** Abbreviation list of the interpolation techniques used in Table 1.

Title	Abbreviations
IDW	Inverse Distance Weighting
AIDW	Angular IDW
GIDW	Gradient IDW
SK	Simple Kriging
OK	Ordinary Kriging
BK	Bayesian Kriging
UK	Universal Kriging
CAR	Clustering Assisted Regression
TPS	Thin Plate Splines
RBF	Radial Basis Function
RF	Random Forest
RFIDW	RF + IDW
RFOK	RF + OK
TS	Trend Surface
TP	Thiessen Polygon

## 2. Materials and Methods

### 2.1. Experimental Data

The dataset used in our experiment is the Tasmania's South Esk Hydrological dataset (Figure 1), developed by CSIRO [18]. It is in netCDF format, a common data format used for environmental monitoring, in which data are stored within multi-dimensional gridded arrays.

In this dataset, the map is gridded every 0.01 degree ( $\approx 1.1 \text{ km}^2$  spatial area), which results in 151 horizontal and 101 vertical grids (a total of 15,251 cells). The data also have a temporal dimension that records data hourly within the area of interest.

There are a number of environmental parameters available within the dataset, such as temperature, relative humidity, solar radiation, surface height, and so on. In this paper, we used the surface height data to select the near optimal location of sensor nodes for determining the most appropriate spatial interpolation technique, and such a selection will be justified in Section 3.1. Accordingly, we adopt the air temperature parameter data for the whole year of 2013, from 5 January to 27 December (inclusive) for the rest of the simulation. However, the dataset still contains some gaps, and these will not be included in the simulation process to ensure the integrity of the final results.



**Figure 1.** Location of South Esk Hydrological model, northeast of Tasmania, Australia. The exact coordinate locations of the red bounded region are:  $42^{\circ}00'S$ ,  $147^{\circ}00'E$  (bottom, left) and  $41^{\circ}00'S$ ,  $148^{\circ}30'E$  (top, right), which covers approximately  $18,588 \text{ km}^2$  (image adapted from Google Maps).

### 2.2. Interpolation Techniques

#### 2.2.1. Ordinary Kriging

In this paper, we utilised OK as the kriging-based interpolator because it is the most-widely used kriging method for interpolation. Furthermore, based on the fact that we will compare techniques based on geostatistical and non-geostatistical methods, for the case of fairness, we only consider kriging-based methods that are “univariate” [2].

OK, also called the OK estimator, is categorised as a geostatistical method; it takes spatial patterns and the uncertainty of the surface into account during the interpolation process [13]. A semivariogram is calculated to show how much a point in space is related to the points around it within a particular distance (neighbouring values) using the following equation:

$$\hat{\gamma}(h) = \frac{1}{2n} \sum_{i=1}^n [z(x_i) - z(x_i + h)]^2 \quad (1)$$

where  $n$  is the number of pairs of observed points  $z$  that are separated by distance  $h$ . Subsequently, the  $\hat{\gamma}(h)$  is plotted against  $h$  and is called the experimental variogram. The variogram curve needs to be fitted for the OK estimator; the spherical model is the most commonly used and it will also be utilised in this work. The conventional kriging estimator is of the form:

$$\hat{z}(x_0) = \sum_{i=1}^n \lambda_i [z(x_i)] \quad (2)$$

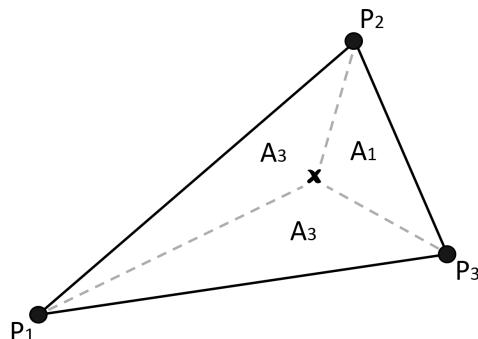
where  $\hat{z}(x_0)$  is the interpolation value at any unmeasured location;  $\lambda_i$  is the kriging weight measured at the  $i$ -th point with  $\sum_i \lambda_i = 1$ ; with the calculation based on minimising the variance of  $\hat{z}$ :

$$\sigma^2 = \sum_{i=1}^n \lambda_i \gamma(x_i, x_0) + \phi \quad (3)$$

where  $\phi$  is the Lagrange multiplier, which is required for the minimisation.

## 2.2.2. Shape Function: Triangular Irregular Network

TIN is a digital means of representing surface morphology and has been extensively used in the GIS community. It consists of connecting edges between vertices that eventually form a network of triangles, which is normally constructed using the Delaunay triangulation algorithm.



**Figure 2.** An illustration of the TIN interpolation technique at point  $x$ . In this case, point  $w$  lies within a triangle formed by points  $P_1$ ,  $P_2$  and  $P_3$ . The weight of each point is calculated based on the corresponding area. For example,  $P_1$  has the weight of  $\frac{A_1}{\sum A_i}$ , and so on.

This surface analysis technique can be further extended as a linear approximation interpolation algorithm as proposed by Peuker and co-workers in 1978 for digital elevation modelling [12]. It was described in the work done by Li in 2003, which uses area divisions for the weighting mechanism [14] (Figure 2). As it is based on triangle meshes, the total number of included observed points is 3. It is in the form of:

$$\hat{z}(x_0) = \sum_{i=1}^{N=3} \lambda_i \cdot z(x_i) \quad (4)$$

$$\lambda_i = \frac{A_i}{\sum_{i=1}^{N=3} A_i} \quad (5)$$

$$A_i = \frac{1}{2} \det \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} \quad (6)$$

where  $A$  is the area of the entire triangle and  $A_i$  is the  $i$ -th sub-triangle formed by the point to be interpolated ( $x$  in Figure 2).

### 2.2.3. Classic Inverse Distance Weighting

The traditional IDW interpolation technique proposed by Shepard in 1968 [19] uses the following equations:

$$\hat{z}(x_0) = \sum_{i=1}^N \lambda(d_{s_i})_i \cdot v_i \quad (7)$$

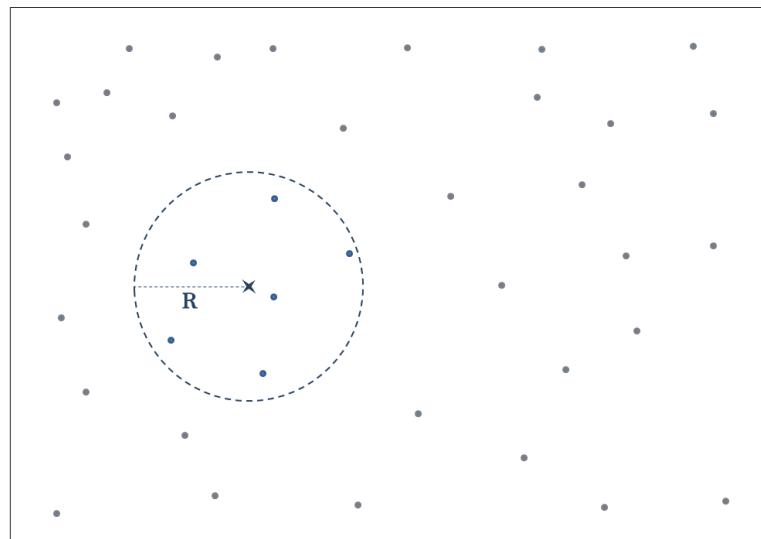
$$\lambda(d_{s_i})_i = \frac{d_{s_i}^{-u_s}}{\sum_i^N d_{s_i}^{-u_s}} \quad (8)$$

$$d_{s_i} = \sqrt{(x_i - x)^2 + (y_i - y)^2} \quad (9)$$

where  $d_{s_i}$  is the 2D-spatial Euclidean distance between the  $i$ -th known data point and the unmeasured location point;  $\lambda(d_{s_i})$  and  $v_i$  are the weights and values assigned to each of the observed points; and  $u_s$  is the spatial distance factor parameter specified by the user.

### 2.2.4. Improved Inverse Distance Weighting

IDW is based on the notion that nearer data points will have more influence than those further away, and so, including every data point throughout the map to interpolate a single point is unnecessary. This is because as the distance is further away; the particular point has very little influence on the final value. An illustration of the improved version of IDW can be seen in Figure 3.



**Figure 3.** A demonstration of the improved IDW. The dots are the sample data points, and the  $x$  is the point location to be interpolated.  $R$  is a user-defined radius parameter that indicates the farthest distance to be included from the points  $x$ . In this case, only a total of six sample points will have an influence on the interpolation process. However, no empirical approach has been developed to obtain the optimal value for the parameter  $R$ .

On the other hand, de Mesnard has discussed the problem of 0 distance in IDW (i.e.,  $d_{s_i} = 0$  as in Equation (9)), and it is divided into two categories, such that there is a discrete and a continuous case [20]. In this paper, since our dataset is based on the gridded system (Section 2.1), it implies that the value within a single cell is actually the mean value of the entire  $1.1 \text{ km}^2$  area. Therefore, we can justify that our dataset lies within the discrete category and that we can simply obtain the value at the sample point when the algorithm tries to interpolate a reference point.

The improved IDW comes with two major advantages in terms of computational efficiency: (a) the processing time does not increase as the number of known points increase; (b) we can further improve the performance by applying the kd-tree data structure algorithm, reducing the computational time from  $O(n)$  to  $O(\log n)$  [16]. Furthermore, this feature is crucial because the basic computational time required for the STI algorithm is  $O(t * n)$ , instead of  $O(n)$ , where  $t$  is the number of user-defined window lengths to be included in the process.

#### 2.2.5. STI Technique : Extension Approach

The extension approach to the IDW spatiotemporal interpolation method is described in [16]:

$$\hat{z}(x_0, t) = \sum_{t_{start}}^{t_{end}} \sum_{i=1}^N \lambda(d_{st_i})_{t,i} \cdot v_{t,i} \quad (10)$$

$$\lambda(d_{st_i})_{t,i} = \frac{d_{st_i}^{-u_s}}{\sum_i^N d_{st_i}^{-u_s}} \quad (11)$$

$$d_{st_i} = \sqrt{(x_i - x)^2 + (y_i - y)^2 + c^2(t_i - t)^2} \quad (12)$$

where  $d_{st_i}$  is the spatiotemporal distance between the measured  $(x_i, y_i, ct_i)$  and unmeasured  $(x, y, ct)$  location point, and  $c$  is the user-defined temporal factor. However, there is still no empirical information on how to justify the temporal factor ( $c$ ), and a naive choice does not yield optimal results [16].

### 2.3. Proposed Methodology

#### 2.3.1. STI Technique: Reduction Approach

The main contribution in this paper is to propose another variation of the IDW-based STI algorithm, namely the reduction approach, which treats time independently from the spatial distance dimension [14]. Based on the conventional IDW technique, this practice calculates the weight in two steps: firstly using the inverse of 2D-spatial distance, followed by the inverse of the 1D-temporal distance. As below:

$$\hat{z}(x_0, t) = \sum_{t_{start}}^{t_{end}} \sum_{i=1}^N \lambda(d_{s_i}, d_{t_i})_{t,i} \cdot v_{t,i} \quad (13)$$

$$\lambda(d_{s_i}, d_{t_i})_{t,i} = \frac{d_{s_i}^{-u_s} \cdot d_{t_i}^{-u_t}}{\sum_i^N d_{s_i}^{-u_s} d_{t_i}^{-u_t}} \quad (14)$$

where  $d_{t_i}$  is the temporal distance and  $u_t$  is the user-defined temporal distance-decay factor.

However, based on the weighting mechanism (Equation (14)), we can see that the function will introduce an error when either the spatial ( $d_{s_i}$ ) or the temporal ( $d_{t_i}$ ) distance is 0 (divided by zero error). Therefore, we need to specify the smallest distance possible in both the spatial and temporal dimension. For our experiment, we set the smallest spatial ( $d_{s_i}$ ) and temporal ( $d_{t_i}$ ) distance at 1. These settings are artificially chosen by the user based on the fact that the lowest possible spatial and temporal distance (excluding 0) in this dataset is 1.

### 2.3.2. Distribution-Based Distance Weighting

In this paper, we also introduce the so-called distribution-based distance weighting (DDW) technique. It considers nearby data variations (similar to improved IDW, as shown in Figure 3) to produce the distribution for the weight calculations. Below are the distribution equations that are used and compared in this work:

$$\text{Gaussian}(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x_i-\mu)^2}{2\sigma^2}} \quad (15)$$

$$\text{Lorentzian}(x, \mu, \Gamma) = \frac{\Gamma}{2\pi} \cdot \frac{1}{x^2 + (\frac{\Gamma}{2})^2} \quad (16)$$

$$\text{Laplacian}(x, \mu, b) = \frac{1}{2b} \cdot e^{-\frac{|\mu-b|}{b}} \quad (17)$$

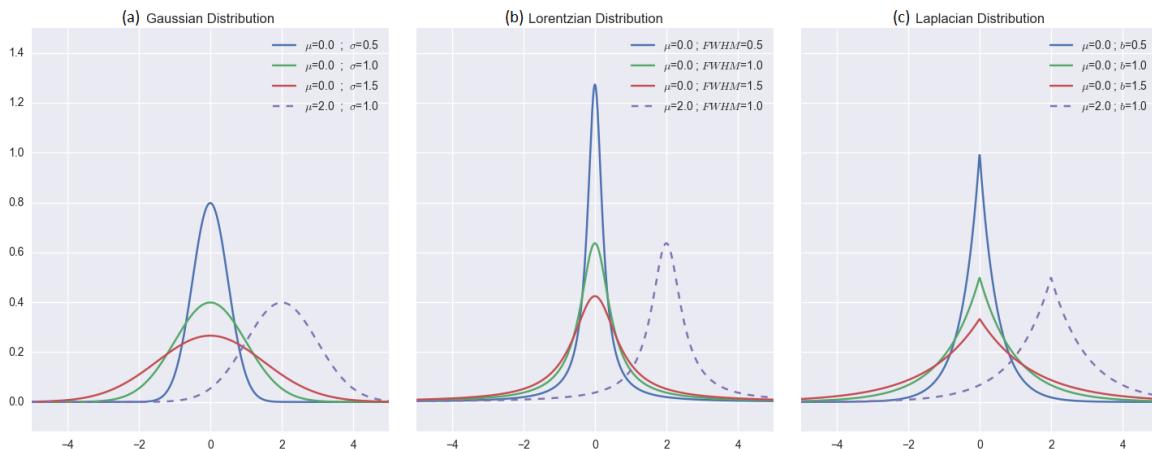
where  $\mu$  is the mean of the data  $D = \{v_1, v_2, v_3, \dots, v_i, \dots, v_N\}$ . For each of the distributions:  $\sigma$  is the standard deviation for the Gaussian;  $\Gamma$  is the full width at half maxima (FWHM) for the Lorentzian; and  $b$  is the “diversity” (or scale parameter) for the Laplacian distribution (Figure 4). However, for the purpose of this work, we will regard  $\Gamma$  and  $b$  as equivalent to the standard deviation  $\sigma$ . The calculation for DDW is in the form:

$$\hat{z}(x_0) = \sum_{i=1}^N \lambda(d_{s_i})_i \cdot v_i \quad (18)$$

$$\lambda(d_{s_i})_i = \text{Distribution}(d_{s_i}, \mu_R, \sigma_R) \quad (19)$$

$$\sigma_R = R \cdot \frac{\sigma}{\max\{|min(D) - \mu_D|, |max(D) - \mu_D|\}} \quad (20)$$

where  $\text{Distribution}(d_{s_i}, \mu_R, \sigma_R)$  can be either a Gaussian, Lorentzian or Laplacian distribution with  $\mu_R = 0$  (the middle position).



**Figure 4.** Distribution methods that are proposed in this work for comparisons: (a) Gaussian; (b) Lorentzian and (c) Laplacian distributions. Within the proposed DDW, the horizontal axis represents the distance, and the vertical axis is the weight values.

The idea of this approach is that when the data  $D$  variations are very small (i.e., low  $\sigma$ ), the distribution will have a fairly sharp peak and will cause the weighting to be more sensitive to the distance ( $d_{s_i}$ ). On the other hand, if the  $\sigma$  is high, this means that the values included in the calculation are more spread out. This makes the width of the “bell” very wide, which results in the  $d_{s_i}$  having less impact on the weight calculation. This method is capable of reducing the “bull’s eye” effect in the final results, which is a major problem in the conventional IDW technique [1]. However,

it is notable that our proposed conventional reduction approach of the STI technique is not designed for this purpose. This is because in some cases, a very smooth surface for geospatial analysis can be very unrealistic.

### 3. Simulation Procedure and Performance Assessment

#### 3.1. Sensor Networks Spatial Samplings

This methodology will be used to obtain the near optimal placement of sensor nodes (Section 3.2) that would best represent (for a specific environmental parameter) the whole region of interest (ROI), based on a given fixed number of nodes ( $N$ ).

For this purpose, we adopt Python's `scipy` built-in module, the `differential_evolution` (DE) algorithm [21], to select the locations of observed points, instead of just randomising them. This procedure is implemented based on the fact that if we entirely randomise them, there is a possibility that the stochastically-picked location points would be clustered in a specific region, and no interpolation method will produce acceptable results, especially when the sample size is very small. The results will be demonstrated in Section 4.2.

Since we are using the South Esk hydrological modelled data, we can get access to significant environmental information about the area. However, for the purpose of this simulation, we will utilise the ROI's elevation data from the dataset to calculate the fitness (or representativeness). Such a selection was made because a high-elevation sample is crucial in the dataset for meteorological studies [5].

This work is focusing on interpolation, which means that we have to manually pick 4 location points so that we could actually 'interpolate' the entire map; these are: top-left, top-right, bottom-left and bottom-right. Then, the rest of location nodes ( $N - 4$ ) will become the individuals of the DE algorithm, represented as  $X = \{x_1, x_2, \dots, x_N\}$ . The objective of the optimisation is to maximise the leave-one-out cross-validation (LOOCV) error, so that extreme values (low and high elevation sample) are obtained [22]:

$$\text{LOOCV}_{\text{SSE}}(X) = \sum_{i=1}^N [ \text{IDW}(x_i, X_i) - o_i ]^2 \quad (21)$$

where  $\text{SSE}$  denotes sum-squared-error;  $\text{IDW}(x_i, X_i)$  represents the estimated value using IDW interpolator at location  $x_i$  using the remaining nodes (such that  $X_i = X \setminus \{x_i\}$ ); and  $o_i$  is the observed value at the  $i$ -th location. In addition, sparsity is an important aspect in sensor network deployment in a way that well-distributed nodes' locations are preferred:

$$\text{sparsity}(X) = \sigma(X) / \mu(X) \quad (22)$$

where  $\sigma$  and  $\mu$  are the calculated standard deviation and the mean value from the pair-wise distances in  $X$ . Finally, the optimisation process is designed to minimise the equation below to comply with the objectives:

$$\text{fitness} = \text{LOOCV}_{\text{SSE}}(X)^{-1} \cdot \text{sparsity}(X) \quad (23)$$

#### 3.2. Choosing a "Balanced" 2D Spatial Interpolation Method

First of all, we compare the computational efficiency of currently-available 2D spatial interpolation algorithms. Based on the assumption that we will collect a huge amount of data (up to thousands of items) with a very high temporal frequency (every second), the computational efficiency is a crucial factor in the algorithm selection. Therefore, we must select a 'balanced' algorithm that produces an acceptable result in a timely manner.

In this section, we will compare four methods: kriging, basic IDW, improved IDW and TIN. The comparison will be divided into two steps: (1) evaluate the total elapsed time of each method,

and discard the method that requires an unacceptable time frame to process; (2) objectively compare the quality of the final output using root mean squared error (RMSE). RMSE is a frequently-used error measurement tool based on the difference between the estimated value produced by the interpolator, and the observed value; and it has been recognised as the best overall method for spatial analysis comparison [12]. It is in the form of:

$$RMSE = \left[ \frac{1}{N} \sum_{i=1}^N (p_i - o_i)^2 \right]^{\frac{1}{2}} \quad (24)$$

where  $N$  is the total number of cells and  $p_i$  and  $o_i$  are the  $i$ -th predicted and observed points, respectively. It should be also noted that the RMSE largely favours the largest error (i.e.,  $p_i - o_i$  in Equation (24)) to the detriment of the lowest terms; in addition, a different/larger exponent could also be chosen to exaggerate such phenomena.

Furthermore, in order to remove the "bias" of RMSE, where it favours more larger errors, we also utilised mean absolute error (MAE) and the corresponding standard error (SE) using:

$$MAE = \frac{1}{N} \sum_{i=1}^N |p_i - o_i| \quad (25)$$

$$SE = \frac{\sigma}{\sqrt{N}} \quad (26)$$

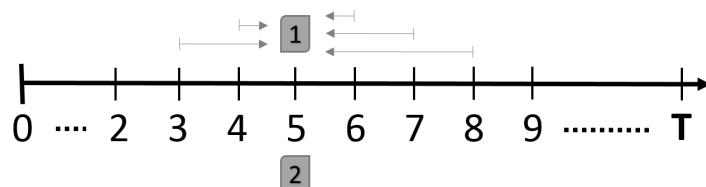
where  $\sigma$  represents the standard deviation of the list of error (i.e., calculated by  $|p_i - o_i|$  as in Equation (25)).

In this experiment, we will compare the statistical quality and the computational efficiency of each method to choose a "balanced" spatial interpolation method for our STI algorithm to be developed at a later stage. The chosen method will be used for further comparisons in the following experiments.

### 3.3. Comparing Spatiotemporal Approaches and Evaluating the DDW

The fundamental idea of the STI method is to develop an algorithm to measure the environmental situation at an unmeasured time slice and to show how well the algorithm can reconstruct the "missing" time slices. This experiment is designed mainly for that purpose, as well as to define which temporal approach (extension or reduction) will produce better results.

Our South Esk Hydrological model holds data in one-hour time intervals. For each time index, we reconstruct the map value using the proposed STI method using different schemes. The "scheme" refers to the previous and next window length to be included in the interpolation process. For example, we would like to interpolate Time Index 5 (Figure 5), and the scheme is (2, 3). This means that Time Indices 3, 4, 6, 7, 8 are used to estimate the map at Time Index 5. However, we skip the iterations that would result in an invalid time index value.

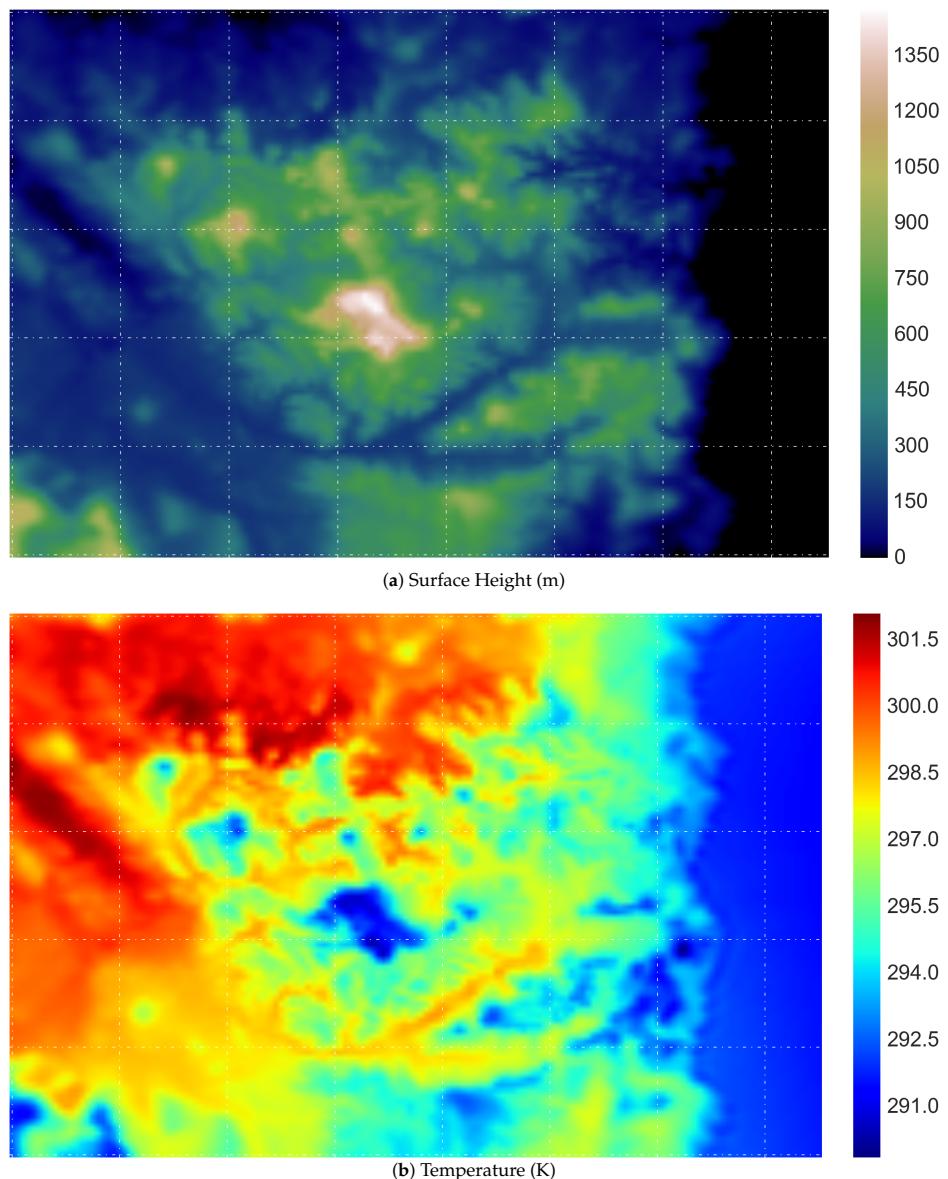


**Figure 5.** Demonstration of the entire simulation setup for each time slice using a single scheme. The example above interpolates Time Index 5, using scheme (2, 3). Map 1 (top) is the "estimated" map generated using the STI method; Map 2 (bottom) is the "observed" map based on the original data. The objective performance of the algorithm is measured using RMSE.

## 4. Results

### 4.1. Original Data Visualisation

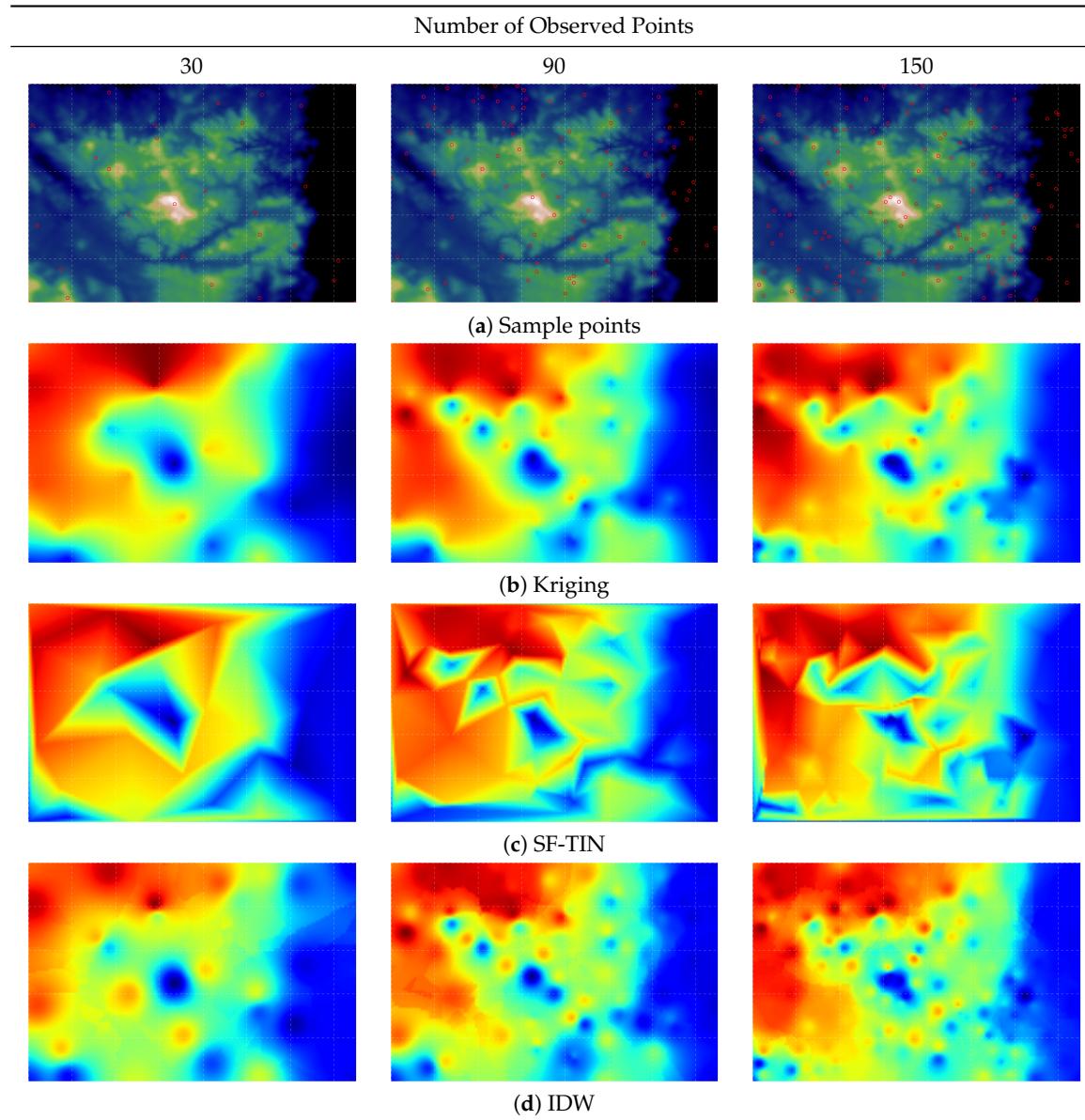
Figure 6 demonstrates a visualisation of the original data used in this work. The surface height data are mainly utilised for the purpose of selecting sensor location using an evolutionary algorithm (differential evolution, as described in Section 3.1); whilst the air temperature data are used for the rest of the simulations.



**Figure 6.** Visualisation of the South Esk hydrological modelled dataset used throughout the simulation: (a) the elevation map, in meters; and (b) the air temperature data, in Kelvin.

### 4.2. Comparing Spatial Interpolation Methods

Figure 7 is a demonstration of the maps that are generated by the interpolation techniques. We can see that kriging creates the smoothest map overall. TIN is visually abrupt, whereas IDW creates gradual visual results with an observable “bull’s eye” effect. The following experiment is designed to evaluate the computational efficiency of each method.

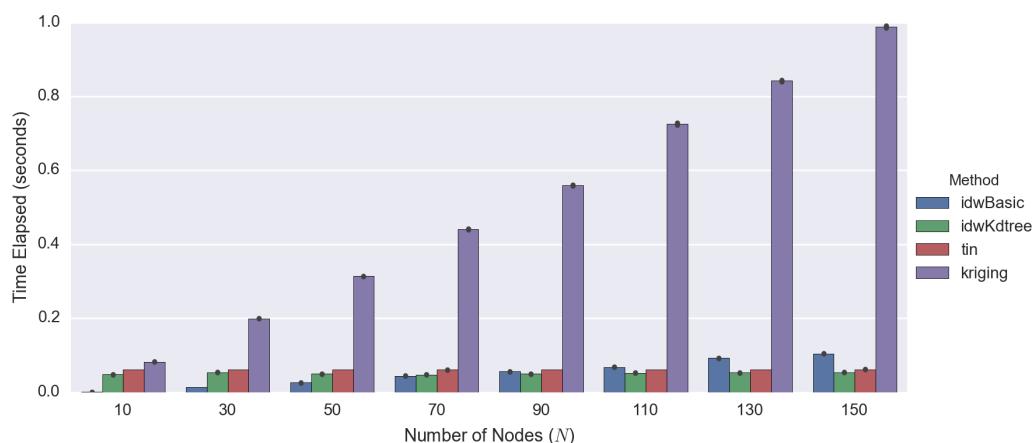


**Figure 7.** Visualisation output for: the selected sample point locations (a) using the optimisation technique (DE); and three widely-used different spatial interpolation techniques (a) ordinary kriging, (b) triangular irregular network and (c) inverse distance weighting], compared to different numbers of sample points ( $N = 30, 90, 150$ ).

**Table 3.** Total normalised time elapsed for each method (0: fastest, 1: slowest).

Number of Observed Points	Total Time Elapsed			
	Basic IDW	Improved IDW	TIN	Kriging
10	0.0024	0.0252	0.0289	0.3034
30	0.0096	0.0278	0.0281	0.3852
50	0.0214	0.0268	0.0291	0.4706
70	0.0355	0.0253	0.0291	0.6582
90	0.0447	0.0040	0.0286	0.7808
110	0.0545	0.0283	0.0288	0.8633
130	0.0596	0.0322	0.0283	0.9720
150	0.0596	0.0322	0.0283	0.9720

Based on Figure 8, kriging is the most computationally-intensive method compared to other techniques, irrespective of the number of observed points. The time required for kriging increases substantially as the number of observed values grows. Since computational efficiency is a crucial factor in this paper (Section 1.1) as previously mentioned, kriging-based spatiotemporal algorithms are not a suitable technique for our purpose and will not be considered and compared.



**Figure 8.** This graph demonstrates a comparison of computational efficiency for different 2D spatial interpolation techniques, with the data based on Table 3.

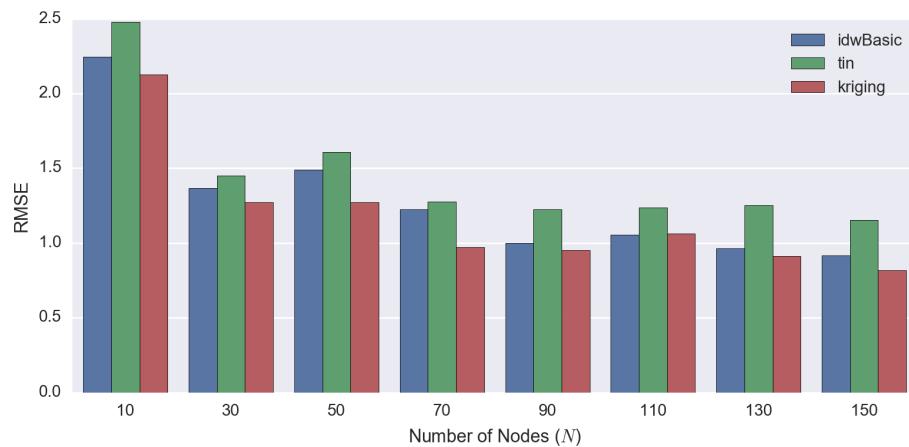
On the other hand, among the three different techniques compared (kriging, TIN and IDW), IDW performs best overall. If the number of observed points is small, basic IDW performs well. However, as the number of observed points increases, the processing time for basic IDW also increases linearly. It will reach a point where the improved IDW performs better than basic IDW (i.e.,  $N = 90$  in Figure 8).

The time required for both TIN and the improved IDW is  $O(\log n)$ , which means that they are suitable for use with a large-scale problem. Then, we compare these techniques based on interpolation performance.

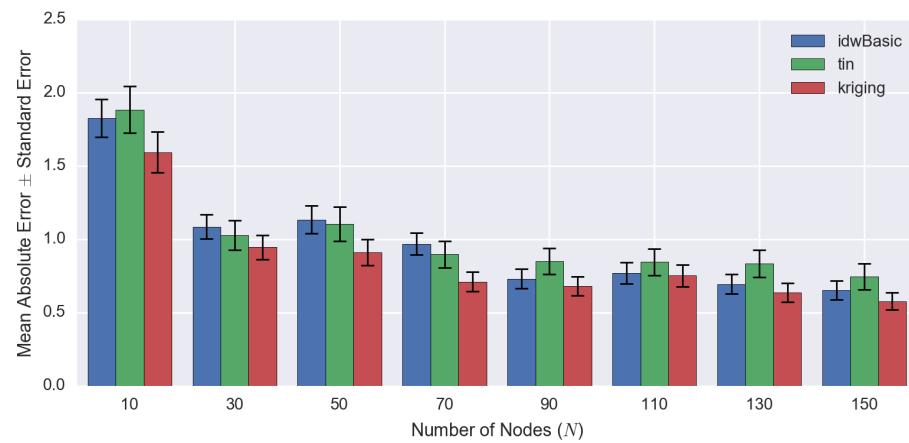
Based on Figures 9 and 10, the statistical error decreases gradually as the number of observed values ( $N$ ) increases. It is obvious that kriging produces the best objective results amongst the spatial interpolation techniques (lowest RMSE) regardless of  $N$ , but it is very computationally intensive. On the other hand, IDW and TIN have comparable statistical error. However, it is interesting to see that for some cases (i.e., when  $N = \{30, 50, 70\}$ ), the MAE of IDW is higher than TIN, but its RMSE is lower. The occurrence of this phenomenon indicates that IDW generates more error than TIN on average, but it also signifies that the error being produced by TIN is more extreme resulting from the ‘bias’ of RMSE (as previously discussed in Section 3.2).

**Table 4.** Statistical error (RMSE and MAE with its corresponding SE) for each method. In this graph, we only show one ‘IDW’, because the errors of both versions of IDW (basic and improved, based on Table 3) are the same; the only difference between the two being the processing time.

Number of Observed Points	RMSE			MAE $\pm$ SE		
	IDW	TIN	Kriging	IDW	TIN	Kriging
10	2.2461	2.4767	2.1257	1.83 $\pm$ 0.13	1.89 $\pm$ 0.16	1.60 $\pm$ 0.14
30	1.3667	1.4503	1.2720	1.09 $\pm$ 0.08	1.03 $\pm$ 0.10	0.95 $\pm$ 0.08
50	1.4897	1.6080	1.2718	1.14 $\pm$ 0.10	1.11 $\pm$ 0.12	0.91 $\pm$ 0.09
70	1.2259	1.2755	0.9727	0.97 $\pm$ 0.07	0.90 $\pm$ 0.09	0.71 $\pm$ 0.07
90	0.9985	1.2243	0.9517	0.73 $\pm$ 0.07	0.85 $\pm$ 0.09	0.68 $\pm$ 0.07
110	1.0565	1.2378	1.0640	0.77 $\pm$ 0.07	0.85 $\pm$ 0.09	0.75 $\pm$ 0.07
130	0.9645	1.2533	0.9136	0.70 $\pm$ 0.07	0.84 $\pm$ 0.09	0.64 $\pm$ 0.07
150	0.9146	1.1537	0.8168	0.65 $\pm$ 0.06	0.75 $\pm$ 0.09	0.58 $\pm$ 0.06



**Figure 9.** This graph demonstrates the quantitative comparisons (RMSE) between the different 2D spatial interpolation techniques, using the data as shown in Table 4.



**Figure 10.** This graph demonstrates the MAE with the corresponding SE between the different 2D spatial interpolation techniques, using the data as shown in Table 4.

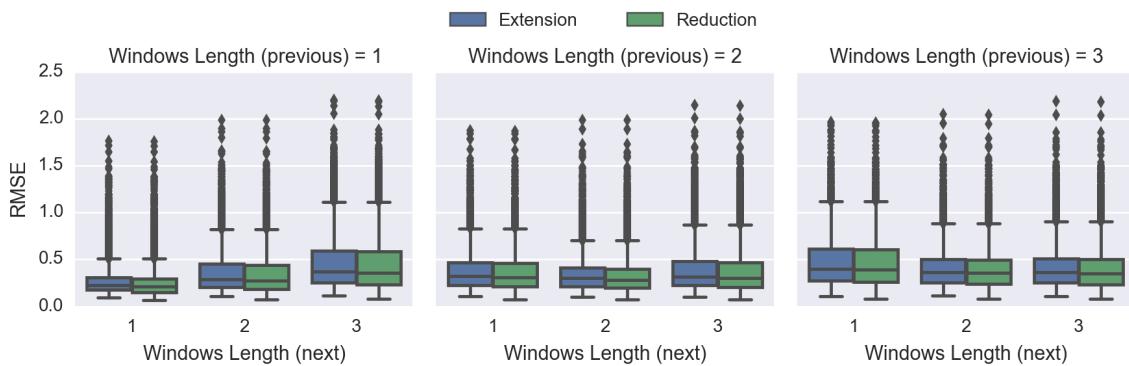
#### 4.3. Comparing Spatiotemporal Approaches

In this experiment, we will compare the performance of two approaches to temporal interpolation: extension and reduction [14]; and we will utilise the improved IDW method, which is the most suitable spatial interpolation technique according to the previous results (Section 4.2, Comparing Spatial Interpolation Methods).

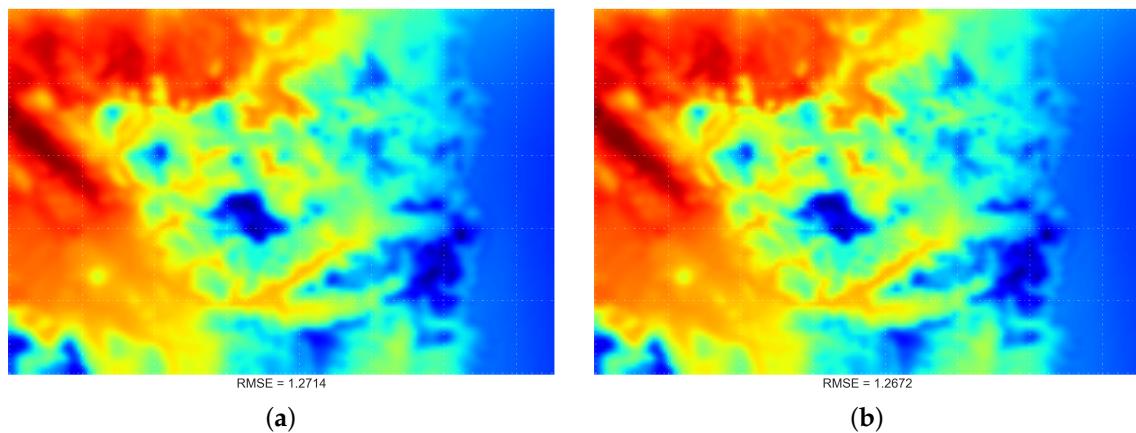
Figure 11 shows that whenever the scheme (either previous or next) is zero, it leads to very poor performance. However, this kind of scheme is actually extrapolating instead of interpolating, which is outside of the scope of this work. Therefore, we will not make any conclusions based on these results.

For each of the schemes, the extension approach always performs more poorly than the reduction method (i.e., higher RMSE). These results confirm that our proposed reduction approach is superior to the current extension approach and is more suitable as the STI technique.

Lastly, Figure 11 indicates that the scheme of including one “previous” and one “next” time index will produce the lowest RMSE overall, and a visualisation demonstration is provided in Figure 12. Such a configuration will be used in the following experiments.



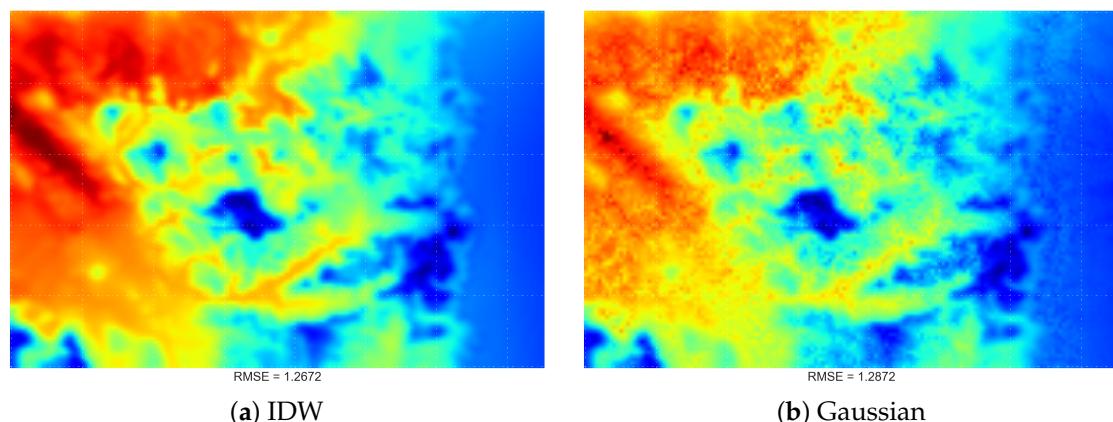
**Figure 11.** Performance comparisons between different temporal interpolation approaches using the improved IDW: extension (existing) and reduction (proposed) approaches. The vertical axis is the RMSE of the respective (previous and next) windows' length.



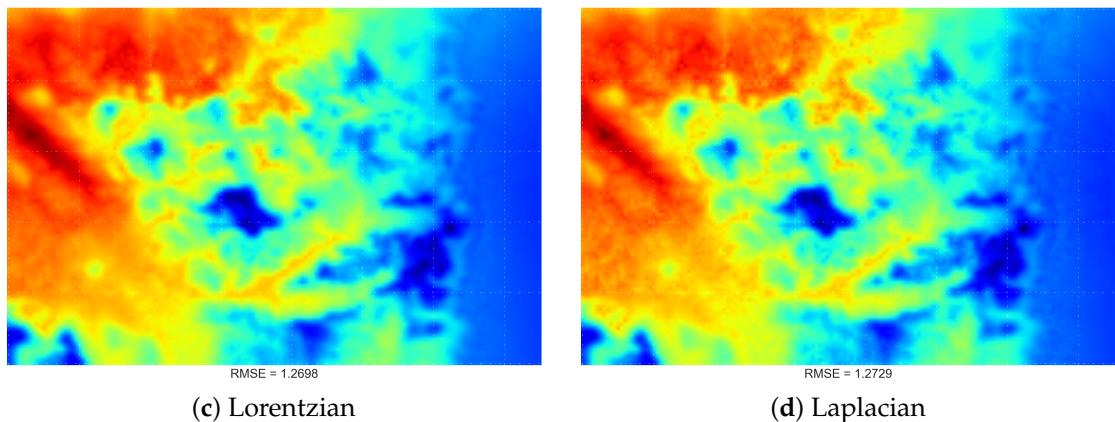
**Figure 12.** Visualisation output for the comparison results between the extension- and reduction-based STI methods. In this output, we used the scheme involving plus-and-minus one time index. This selection is based on the results obtained from Figure 11, which shows such a configuration produces the best results overall (lowest mean RMSE). (a) Extension approach; (b) reduction approach.

#### 4.4. Comparing the Proposed DDW Methods

In this experiment, we compared the IDW method with the proposed DDW method. In Figure 13, there is little visual difference between STI techniques using either IDW or the proposed DDW methods. However, in general, DDW methods are more likely to produce smoother surfaces than the IDW technique.

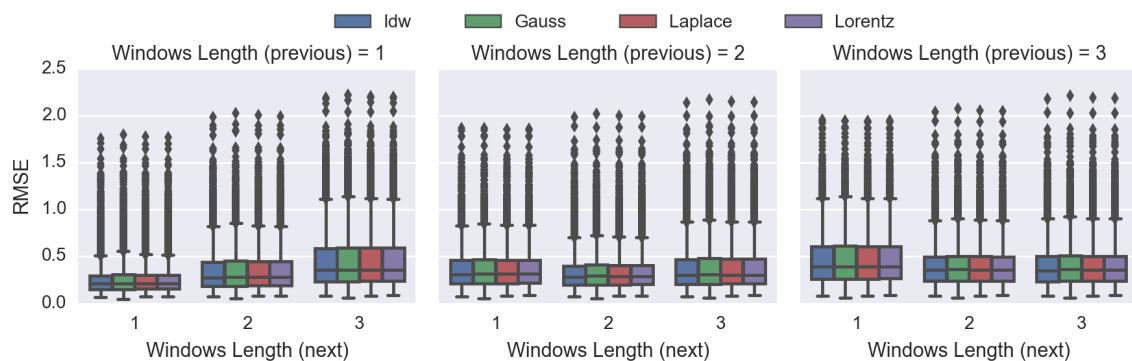


**Figure 13. Cont.**



**Figure 13.** Visualisation output of IDW (a) and the proposed DDW: (b) Gaussian; (c) Lorentzian; and (d) Laplacian; using the proposed spatiotemporal interpolation technique.

The results in Figure 14 show that IDW is superior to DDW. On average, the RMSE of IDW is approximately 38% lower than that of the DDW techniques, which is a significant difference. In this experiment, the graph also suggests that using the scheme of plus-and-minus one time index will produce the lowest error overall.



**Figure 14.** Objective comparisons between the IDW method and the proposed DDW techniques: using Gaussian, Lorentzian and Laplacian distributions. This graph demonstrates the RMSE comparison between the respective (previous and next) windows' length.

## 5. Discussion

We summarise the strengths and weaknesses of the 2D spatial interpolation techniques in Table 5. The kriging interpolation technique performs quantitatively best overall (Table 4 and Figure 9). However, the time required for processing increases dramatically as the number of observed points grows (Figure 8). Thus, because of the fact that computational efficiency is a crucial factor in our work, kriging is the least preferred method.

The improved IDW technique that utilises the kd-tree algorithm and the TIN technique using Delaunay triangulation have comparable computational efficiency  $O(\log n)$  and similar objective performance. Nevertheless, from an environmental manager's perspective, visual continuity is an important factor in making a confident interpretation. Because the TIN method is visually abrupt (Figure 7), it is less suitable for environmental applications [2], and so, IDW is considered the most suitable spatial interpolation method overall.

**Table 5.** Summary of the strengths and weaknesses of the spatial interpolation methods evaluated in this paper: ordinary kriging (OK), inverse distance weighting (IDW) and triangle irregular network (TIN). These evaluations are based on three criteria: objective performance (RMSE), computational efficiency and the final visual output smoothness.

	<b>Advantages</b>	<b>Disadvantages</b>
OK	<ul style="list-style-type: none"> <li>- Geostatistical method that considers spatial pattern.</li> <li>- Best objective comparison overall.</li> <li>- Creates very smooth surface.</li> </ul>	<ul style="list-style-type: none"> <li>- Computationally very expensive.</li> <li>- Unsuitable for very large datasets.</li> </ul>
IDW	<ul style="list-style-type: none"> <li>- Computationally inexpensive.</li> <li>- Can be improved using kd-tree to increase computational efficiency.</li> <li>- Creates smooth surface.</li> </ul>	<ul style="list-style-type: none"> <li>- Produces “bull’s eye effect” for extreme values.</li> <li>- Unsuitable for large scale problems without using kd-tree algorithm.</li> </ul>
TIN	<ul style="list-style-type: none"> <li>- Computationally inexpensive.</li> </ul>	<ul style="list-style-type: none"> <li>- Produces abrupt surfaces.</li> </ul>

Figures 11 and 14 demonstrate that whenever the previous or the next time index to be included is zero, the RMSE goes up significantly, leading to a poor performance. The reason is that when we do not include any of the previous or the next time index, we are actually extrapolating rather than interpolating that particular time frame, which is outside the scope of this work.

The results in Figure 11 indicate that our proposed IDW-based reduction spatiotemporal interpolation approach is superior to the currently-available extension temporal method. Consequently, the reduction temporal approach will be utilised. Furthermore, the optimal time frame to be included in the STI is plus-and-minus one time index, and this will yield the best results overall (Figures 11 and 14).

Up until now, we have concluded that a combination of the improved IDW technique using the reduction approach for STI would lead to a balanced result overall, in terms of computational efficiency and objective performance. In this work, we evaluated the performance of our proposed DDW-based STI (Figure 14). However, the proposed DDW showed very little benefit over IDW in terms of RMSE.

Finally, one major advantage of our proposed reduction approach over the extension approach is that the user does not need to justify the temporal factor ( $c$ ) as in Equation 12, which is always done manually and is highly dependent on the user's expertise.

## 6. Conclusions and Future Work

Our study shows that the improved IDW spatial interpolation algorithm, combined with the reduction approach for temporal interpolation, was found to be the most effective technique overall for spatio-temporal interpolation (STI) within environmental modelling applications. This technique provided the following advantages over the alternative STI techniques investigated:

1. Computational time does not increase as the number of sample points grows significantly;
2. The objective results are comparable with other techniques;
3. The technique does not create abrupt visual results.

The results obtained from this work are only applicable for gap filling rather than interpolating over a longer term basis where climate variation patterns need to be considered, e.g., interpolating daily temperature data for consecutive days. This issue will be addressed as a further research direction.

The distribution-based distance weighted (DDW) STI technique, also proposed in this paper, did not perform as well as the inverse distance-weighted (IDW) method and so is not recommended for environmental modelling applications.

The proposed method in this paper has a broad potential of applicability within environmental research, such as oceanography, geography, hydrology, and so on. Since there is no optimal STI interpolation method that is suitable for every kind of problem, our proposed method could perform better in some cases.

The current work relies in its results and conclusions on the demonstration conducted on real data. Therefore, the findings are limited to this set of data and cannot be generalized unless a more rigorous approach is conducted. It is recommended as future work that a theoretical demonstration be performed.

While kriging has not been applied for our spatio-temporal interpolation, as the focus on the current work was computational efficiency, we strongly encourage the use of this technique when accuracy is a major concern ([23] and the references therein).

**Acknowledgments:** Ferry Susanto would like to thank Victoria University for waiving his PhD tuition fees, Vale Institute of Technology for the award of a postgraduate scholarship and CSIRO for a top-up scholarship. We thank CSIRO for providing us with the data from Tasmania's South Esk hydrological model for this study. We also want to express our appreciation to our Swarm Sensing team members, especially Ray Williams, for providing useful comments and suggestions to help improve the quality and presentation of this work.

**Author Contributions:** Paulo de Souza and Ferry Susanto conceived of and designed the experiments. Ferry Susanto performed the experiments. Paulo de Souza, Ferry Susanto and Jing He analysed the data. Paulo de Souza, Ferry Susanto and Jing He contributed analysis tools. Ferry Susanto and Paulo de Souza wrote the paper.

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

## Abbreviations

The following abbreviations are used in this manuscript:

ROI	Region of interest
STI	Spatio temporal interpolation
DDW	Distribution-based distance weighting
IDW	Inverse distance weighting
AIDW	Angular IDW
GIDW	Gradient IDW
SK	Simple kriging
OK	Ordinary kriging
BK	Bayesian kriging
UK	Universal kriging
CAR	Clustering assisted regression
TPS	Thin plate splines
RBF	Radial basis function
RF	Random forest
RFIDW	RF + IDW
RFOK	RF + OK
TS	Trend surface
TP	Thiessen polygon
RMSE	Root mean squared error
MAE	Mean absolute error
SE	Standard error

## References

1. Sanabria, L.A.; Qin, X.; Li, J.; Cechet, R.P.; Lucas, C. Spatial interpolation of McArthur's Forest Fire Danger Index across Australia: Observational study. *Environ. Model. Softw.* **2013**, *50*, 37–50.
2. Li, J.; Heap, A.D. Spatial interpolation methods applied in the environmental sciences: A review. *Environ. Model. Softw.* **2014**, *53*, 173–189.
3. Jeffrey, S.J.; Carter, J.O.; Moodie, K.B.; Beswick, A.R. Using spatial interpolation to construct a comprehensive archive of Australian climate data. *Environ. Model. Softw.* **2001**, *16*, 309–330.
4. Chen, F.; Liu, C. Estimation of the spatial rainfall distribution using inverse distance weighting (IDW) in the middle of Taiwan. *Paddy Water Environ.* **2012**, *10*, 209–222.
5. Tang, L.; Su, X.; Shao, G.; Zhang, H.; Zhao, J. A Clustering-Assisted Regression (CAR) approach for developing spatial climate data sets in China. *Environ. Model. Softw.* **2012**, *38*, 122–128.
6. Lu, G.Y.; Wong, D.W. An adaptive inverse-distance weighting spatial interpolation technique. *Comput. Geosci.* **2008**, *34*, 1044–1055.
7. Joseph, V.R.; Kang, L. Regression-Based Inverse Distance Weighting with Applications to Computer Experiments. *Technometrics* **2011**, *53*, 254–265.
8. Anderson, S. An Evaluation of Spatial Interpolation Methods on Air Temperature in Phoenix. Available online: <http://www.cobblestoneconcepts.com/ucgis2summer/anderson/anderson.htm> (accessed on 5 August 2016).
9. Plouffe, C.C.F.; Robertson, C.; Chandrapala, L. Comparing interpolation techniques for monthly rainfall mapping using multiple evaluation criteria and auxiliary data sources: A case study of Sri Lanka. *Environ. Model. Softw.* **2015**, *67*, 57–71.
10. Sun, Y.; Kang, S.; Li, F.; Zhang, L. Comparison of interpolation methods for depth to groundwater and its temporal and spatial variations in the Minqin oasis of northwest China. *Environ. Model. Softw.* **2009**, *24*, 1163–1170.
11. Naoum, S.; Tsanis, I. Ranking spatial interpolation techniques using a GIS-based DSS. *Glob. Nest* **2004**, *6*, 1–20.
12. Li, J. *A Review of Spatial Interpolation Methods for Environmental Scientists/Jin Li and Andrew D. Heap*; Geoscience Australia: Canberra, Australia, 2008.
13. Burrough, P.A.; McDonnell, R.A. *Principles of Geographical Information Systems*; Oxford University Press: New York, NY, USA, 1998.

14. Li, L. Spatiotemporal Interpolation Methods in GIS. Ph.D. Thesis, The University of Nebraska, Lincoln, NE, USA, 2003.
15. Li, L.; Zhang, X.; Holt, J.B.; Tian, J.; Piltner, R. Spatiotemporal Interpolation Methods for Air Pollution Exposure. In Proceedings of the Ninth Symposium on Abstraction, Reformulation and Approximation (AAAI), Catalonia, Spain, 17–18 July 2011.
16. Li, L.; Losser, T.; Yorke, C.; Piltner, R. Fast Inverse Distance Weighting-Based Spatiotemporal Interpolation: A Web-Based Application of Interpolating Daily Fine Particulate Matter PM2.5 in the Contiguous U.S. Using Parallel Programming and k-d Tree. *Int. J. Environ. Res. Public Health* **2014**, *11*, 9101–9141.
17. Kilibarda, M.; Hengl, T.; Heuvelink, G.B.M.; Gräler, B.; Pebesma, E.; Perćec Tadić, M.; Bajat, B. Spatio-temporal interpolation of daily temperatures for global land areas at 1 km resolution. *J. Geophys. Res. Atmos.* **2014**, *119*, 2294–2313.
18. Katzfey, J.; Thatcher, M. Ensemble One-Kilometre Forecasts for the South Esk Hydrological Sensor Web. In Proceedings of the 19th International Congress on Modelling and Simulation, Perth, Australia, 12–16 December 2011; pp. 3511–3517.
19. Shepard, D. A two-dimensional interpolation function for irregularly-spaced data. In *ACM '68 Proceedings of the 1968 23rd ACM National Conference*; ACM: New York, NY, USA, 1968.
20. De Mesnard, L. Pollution models and inverse distance weighting: Some critical remarks. *Comput. Geosci.* **2013**, *52*, 459–469.
21. Scipy Community. *Scipy V0.15.1 Reference Guide—Differential Evolution*. Available online: [http://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.differential\\_evolution.html](http://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.differential_evolution.html) (accessed on 8 August 2015).
22. Susanto, F.; Budi, S.; Souza, P.D.; Engelke, U.; He, J. Design of Environmental Sensor Networks Using Evolutionary Algorithms. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 575–579.
23. Montero, J.; Fernandez-Aviles, G.; Mateu, J. *Spatial and Spatio-Temporal Geostatistical Modeling and Kriging*; Wiley: West Sussex, UK, 2015.



© 2016 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 (<http://creativecommons.org/licenses/by/4.0/>).