Interpolation Methods Comparison

C. Caruso
CISE SpA, via Reggio Emilia 39,
20090 Segrate (MI), Italy

F. Quarta
ENEL SpA DTI/VDA, viale Italia 26,
20099 Sesto San Giovanni (MI), Italy

Abstract—In many fields, spatial interpolation is used to evaluate physical data in a continuous domain. The many different techniques offer different performances, according to the characteristics of initial data points. The aim of this paper is to provide help in choosing and evaluating the technique that best suits the data set: a few indices measure the quality of interpolation by different viewpoints. An extensive test-case is reported, involving four different interpolation methods. Data sets are generally connected with environmental topics or taken from literature. © 1900 Elsevier Science Ltd. All rights reserved.

1. INTRODUCTION

The generation of continuous surfaces starting from irregularly distributed data is a task for many disciplines. There is a variety of methods which can perform this task, but the difficulty lies in the choice of the one that best reproduces the actual surface.

Each method has its own advantages and drawbacks, which depend strongly on the characteristics of the set of point data: a method that fits well with some data can be unsuited for a different set of data points, or if measured in different locations of the same surface. Thus, it is important to give a few criteria to decide if the method chosen is suited for the point data set. Nevertheless, it is important to specify the aims of the interpolation, because different aims can produce different criteria of evaluation of the interpolation.

In this paper, four different methods have been applied to several test cases, respectively: **Inverse Square Distance Method, Kriging Method, Hardy's Multiquadric Method, and Tension Finite Difference Method**. Each method was tested with different parameters, so that the sensitivity of the method to its parameters can also be evaluated. Names can be misleading, but these methods will be discussed in detail in the following.

Test cases have been taken mainly from environmental studies. To compare results of different interpolation schemes, it was decided to analyze the related grids obtained at regularly spaced points. Moreover, Franke point data sets [1] were used to improve the methods' comparison.

2. INTERPOLATION METHODS

The four interpolation methods were chosen mainly for the availability of their source codes and will be briefly explained in this section. The interested reader can find specific details in the quoted references. This section is not necessary for general understanding of this paper, and can be skipped.

For the sake of clarity, the method classification quoted in [2] will be used. In this classification, spatial interpolation is divided into two main categories; **point interpolation** and **area interpolation** (we are interested only in the first). Point interpolation is divided into two subcategories; **exact methods** and **approximate methods**, respectively, whether or not they preserve the original point values.

In that sense, we have used two exact methods, Kriging and Hardy's Multiquadric, and two approximate ones, Inverse Square Distance and Tension Finite Difference. Note that the characteristics of preserving the point values depend strongly on the particular implementation of the method.

Inverse Square Distance Method

The principle behind this method [3] is to give more weight to nearby points than to distant points. Stating $f(x, y)$ the analytical expression of the surface, we have:

$$f(x,y) = \frac{\sum_{j=1...N} w(d_j) v_j}{\sum_{j=1...N} w(d_j)}, \tag{1}$$

where N is the number of data points, v_j is point j value, d_j is the Euclidean distance with point j, and $w(d)$ is the weighting function. In our example, $w(d)$ has this formula:

$$w(d) = \begin{cases} \dfrac{1}{d_{\min}^2}, & \text{if } d \le d_{\min}, \\[2mm] \dfrac{1}{d^2}, & \text{if } d_{\min} < d < d_{\max}, \\[2mm] 0, & \text{if } d > d_{\max}, \end{cases} \tag{2}$$

where $d_{\min}$ is minimum distance and $d_{\max}$ is maximum distance. Index $d_{\min}$ prevents infinite weight values for $d = 0$, while index $d_{\max}$ avoids using too distant points. If no points fall into the circle of radius $d_{\max}$, average data value is taken.

Reported here are a few considerations on the advantages and disadvantages of using this method. The main advantage is its simplicity; the second one is that this method leads to reasonable results for a wide variety of data, and, moreover, there is no problem with results exceeding the range of meaningful values.

On the other hand, there are several drawbacks. First of all, the method is very sensitive to the weighting function. Second, the method can be affected by uneven data points distribution. Finally, the method has scarce predictive characteristics: for instance, global maxima and minima are always among data points.

Kriging Method

Kriging is a geostatistical method for point interpolation. It derives its name from D.G. Krige, who introduced the use of moving averages to avoid systematic errors in interpolation [4]. The generalization of this method was developed by Matheron [5]. Kriging states the statistical surface as a regionalized variable, with a certain degree of continuity. The Kriging estimate is known as the Best Linear Unbiased Estimate (BLUE), because it is a linear combination of the weighted sample values, whose expected value for error equals zero and whose variance is a minimum. The implementation used in our paper is the simplified version of Kriging suggested by Trochu [6].

The main characteristic of Kriging is the generalized covariance $k(d)$. It was proved [7] that Kriging in the presence of a linear drift is equivalent to spline interpolation for the following generalized covariances:

$$k(d) = \begin{cases} d^3, & \text{(one-dimensional case)}, \\ d^2 \ln(d), & \text{(two-dimensional case)}, \\ d, & \text{(three-dimensional case)}. \end{cases} \quad (3)$$

In our case, we used the following formulas for the two-dimensional case [6]:

$$k(d) = \begin{cases} 1 + \dfrac{1}{c_1} \left(\dfrac{dc_0}{d_{\max}} \right)^2 \ln \left(\dfrac{dc_0}{d_{\max}} \right), & \text{if } d \leq d_{\max}, \\ 0, & \text{if } d > d_{\max}, \end{cases} \quad (4)$$

where $d_{\max}$ is the maximum correlation distance considered, $c_0 = 0.60653066$ and $c_1 = 0.18393972$; and c_0 is the value for which $k(d)$ function has its minimum, which equals $-c_1$. The statistical surface is obtained with the formula:

$$f(x, y) = a_1 + a_2 x + a_3 y + \sum_{j=1...N} b_j k(d_j), \quad (5)$$

where a_i, b_j are calculated resolving the following linear system:

$$\begin{bmatrix} k(d_{1,1}) & k(d_{1,2}) & \cdots & k(d_{1,N}) & 1 & x_1 & y_1 \\ k(d_{2,1}) & k(d_{2,2}) & \cdots & k(d_{2,N}) & 1 & x_2 & y_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ k(d_{N,1}) & k(d_{N,2}) & \cdots & k(d_{N,N}) & 1 & x_N & y_N \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 \\ x_1 & x_2 & \cdots & x_N & 0 & 0 & 0 \\ y_1 & y_2 & \cdots & y_n & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \cdots \\ b_N \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \cdots \\ v_N \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (6)$$

The first N equations express the condition that Kriging is an exact interpolator, while the further three find the coefficients of the plane that fits data points.

Hardy's Multiquadric Method

This method was first formulated by Hardy [8] and was applied in different disciplines. The surface is made of the sum of several cones, whose vertex is placed in the data points.

The formula of a cone is:

$$f(x, y) = \sqrt{((x - x_0)^2 + (y - y_0)^2) \tan^2 \alpha}, \quad (7)$$

where x_0, y_0 are the coordinates of the vertex of the cone, and α is the slope of its surface.

When we consider the whole surface, constituted by N cones with vertices in the data points, we obtain the following formula:

$$f(x, y) = \sum_{j=1...N} c_j \sqrt{(x - x_j)^2 + (y - y_j)^2}, \quad (8)$$

where c_j are obtained, solving the linear system of N equations:

$$\begin{bmatrix} d_{1,1} & d_{1,2} & \cdots & d_{1,N} \\ d_{2,1} & d_{2,2} & \cdots & d_{2,N} \\ \cdots & \cdots & \cdots & \cdots \\ d_{N,1} & d_{N,2} & \cdots & d_{N,N} \end{bmatrix}, \begin{bmatrix} c_1 \\ c_2 \\ \cdots \\ c_N \end{bmatrix} = \begin{bmatrix} v_1 \\ v_2 \\ \cdots \\ v_N \end{bmatrix}, \quad (9)$$

where v_j is data value in point j, and $d_{i,j}$ is the Euclidean distance between points i and j.

This method provides an elegant way to obtain a continuous surface by interpolation of scattered data. The disadvantages are mainly connected with the difficulty in solving the linear system: in fact, its size increases with the number N of data points.

Tension Finite Difference Method

This method finds the surface that solves a differential equation, and at the same time, fits data points. From an operational point of view, a regular grid is defined and data values are assigned to the including cell, so that this method becomes no more exact. It was proposed by Briggs [9] and developed by Smith [10], inside the General Mapping Tool system [11], a free software for scientific use. In formula:

$$(1 - T)\Delta^2 f(x, y) + T\Delta f(x, y) = 0, \tag{10}$$

where Δ is the Laplacian operator, Δ^2 is the biharmonic operator and T is a tension factor between 0 and 1. When $T = 0$, undesired oscillations and false local maxima and minima can be noted, while when $T = 1$, no maxima and minima are possible except at control data points. In [11], a value of $T \cong 0.25$ is suggested for potential field data, while a value of $T \cong 0.35$ is preferable for steep topography data.

Because the implementation considers a regularly spaced grid, the solution accuracy depends mainly on the grid dimension. In our examples, the number of data points (and hence of initialized grid cells) was always well under the number of grid cells.

The parameter T plays an important role in determining the quality of interpolation; that is, by varying its value, it is possible to have better control of results. As mentioned above, the source code for this method was taken from General Mapping Tool software.

3. COMPARISON OF DIFFERENT METHODS

The aim of this paper is to provide a criterium for choosing the best (or at least a good) interpolation for the set of data values. It is not so easy to decide which interpolation is better by looking only at the corresponding contour lines, and often interpolated surfaces can be very different among various methods.

First, there is the difficulty in defining the viewpoints according to which an interpolation can be said to be good, and different viewpoints can lead to different rankings of the methods. Here we considered two viewpoints, **prediction** and **characterization**, which will be explained in the following.

A previous analysis must have been done on point data characteristics. The spatial distribution and the variance of experimental data greatly influence the interpolation process, making it necessary to relate the quality of the interpolation to point data. In this context, some indices were defined to a give a first idea about point data characteristics in terms of spatial homogeneity and experimental surface roughness. For the sake of simplicity, spatial coordinates are normalized: this does not affect the generality of the following considerations. The characteristics considered for a point data set are the kind of spatial distribution and the level of surface roughness. For the first attribute, the uniform index u is defined, dependent on the number of point data N:

$$u(N) = \frac{\sum_{i=1...N} \sum_{j=1...N} p(d_{i,j}, s(N))}{\pi N}, \tag{11}$$

where $s(N)$

$$s(N) = \frac{1}{\sqrt{N}} \tag{12}$$

and $p(d, s)$ is

$$p(d, s) = \begin{cases} 1, & \text{if } d \leq s, \\ 0, & \text{if } d > s. \end{cases} \tag{13}$$

According to the formulas, u measures the average number of points that fall into the circles of radius $s(N)$ centred in the data points.

It can be noted that $s(N)$ decreases with N. On average, a circle of radius $s(N)$, randomly centred in the normalized square domain, contains π points (this can be obtained by comparing

circle area with domain area). As we said before, circles are always centred on data points, thus including at least the centre data point. Border effects (i.e., points whose circle overlaps the domain), decreasing for larger N, and circle positioning effects (i.e., inclusion of at least one point in the circles), increasing for larger N, must be taken into account.

First of all, border effects can be accurately evaluated. The evaluation results in a formula that describes the average number of points in the circle as a function of N (in the appendix, the derivation of this formula is reported):

$$n(N) = \pi - \frac{8}{3\sqrt{N}} + \frac{16\sqrt{2} - 17}{12N}, \qquad \text{with } N \geq 4. \tag{14}$$

Then, to avoid considering circle positioning effects, we can modify the previous formula, not considering point i when it is the centre of the circle. It is equivalent to test with random point i the remaining $N - 1$ points; globally, N circles are used for a fictitious data set of $N - 1$ points. In that case, this new formula is obtained:

$$u(N) = \frac{\sum_{i=1...N} \sum_{j=1...N}^{i \neq j} p(d_{i,j}, s(N))}{n(N-1)N}. \tag{15}$$

Index u, which ranges from 0 to $(N-1)/n(N-1)$, is about 1 for uniform and random distribution, is less than 1 for uniform, not random, distribution, while it is more than 1 for not uniform and, possibly, not random distribution. Possibly, point data sets with index $u \cong 1$ that have not a random uniform distribution can be found, but on average, u measures efficiently these characteristics of spatial distribution.

Function $r(d)$ is introduced to measure surface roughness. Here it is in formula:

$$r(d) = \frac{\sum_{i=1...N} \sum_{j=1...N}^{j \neq i, d_{i,j} < d} \left((v_i - v_j)/\sqrt{2}\sigma_v)^2 \right) w_\tau(d_{i,j})}{\sum_{i=1...N} \sum_{j=1...N}^{j \neq i, d_{i,j} < d} w_\tau(d_{i,j})}, \tag{16}$$

where $w_\tau(d)$ is

$$w_\tau(d) = e^{-d^2}, \tag{17}$$

and σ_v is the standard deviation of v_i, d is an appropriate value of distance. This index measures how variables are spatially correlated.

Let d_n and d_x be the distances between, respectively, the nearest and furthest pair of data points.

With d greater than d_x, $r(d)$ nearly compares the variance of v_i, weighted by pair distance, with the variance of v_i itself. It can be noted that, if the term that depends on $d_{i,j}$ is neglected, $r(d)$ is $N/(N-1)$ and nears 1 (see the Appendix).

Function $r(d)$ is defined only for d, not less than d_n, and it is more significant for values of d that consider a certain number of pairs of data points.

From these considerations, a few indications on the interpretation of $r(d)$ can be suggested. First of all, $r(d_x)$ measures whether spatial distribution can explain value variance. If $r(d_x) \cong 1$, on average there is no spatial correlation (i.e., near points have values that are not more correlated than distant points), or, at least, there is a correlation for pairs nearer than d_n, that cannot be found with the current data set. If $r(d_x)$ is more than 1, there is, on average, an "inverse" correlation (i.e., distant points have values more correlated than near points), or, again, there is a correlation for pairs nearer than d_n, not found with the current data points. Finally, when $r(d_x)$ is less than 1, standard deviation is partially explained by spatial correlation (i.e., near points have values that are more correlated than distant points). Value $r(s(N))$ is another important index whose meaning is similar to that of $r(d_x)$, on a different spatial scale: its advantage is that it takes into account the number N of data points. Finally, we considered d_p, greater than $s(N)$,

as the maximum value of d for which $r(d)$ is always less than 70% of $r(d_x)$. We defined d_p greater than $s(N)$ to avoid unlikely computation with a little number of data pairs. The greater d_p, the better the spatial correlation of the function.

In any case, it is always useful to analyze the behaviour of the whole function to have a better idea of spatial data set correlation. Two examples of $r(d)$ function are reported in Figures 1 and 2.
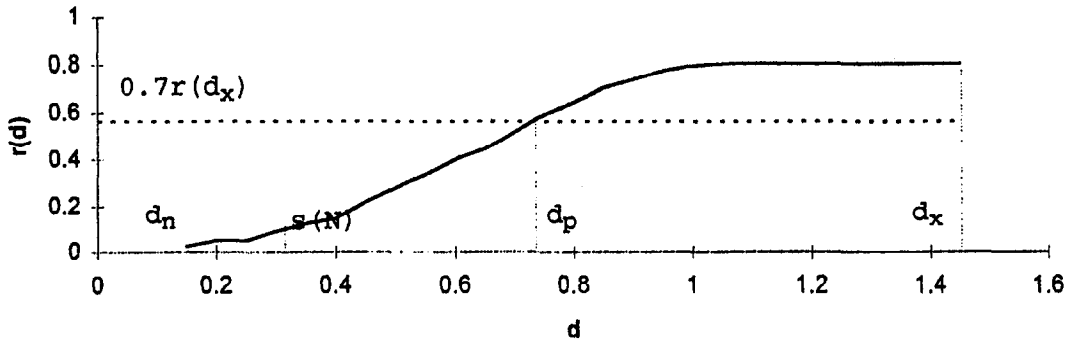
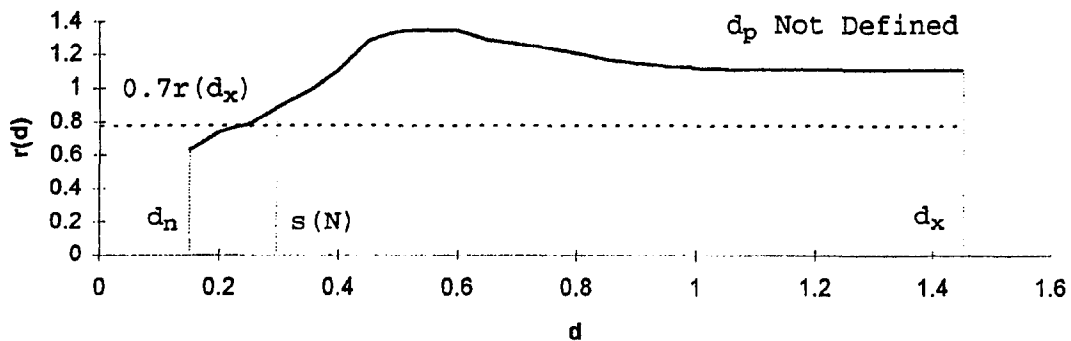Figure 1. Roughness function $r(d)$ example.

Figure 2. Roughness function $r(d)$ example.

After the evaluation of point data characteristics, it is important to decide how to measure the quality of interpolation. As we mentioned before, we chose two viewpoints. **Prediction** states that the best interpolation is the one that minimizes the prediction error in an unknown point. **Characterization** states that the resulting surface must globally look like the actual surface (i.e., a steep actual surface must lead to a resulting steep surface and so on). Often prediction and characterization can be in conflict: for instance, in very rough surfaces the best estimate can be the average of data points, thus resulting in a flat surface.

Prediction

Prediction tells that the best interpolation is the one that minimizes the prediction error in an unknown point. But, unfortunately, we have no knowledge of the true value in unsampled points. Thus, the idea is to simulate this situation. We considered each data point and executed the interpolation ignoring it; then the interpolated value for this point was compared with the true value. This procedure is called **crossvalidation** [12,13]. We thus had two sets of N data, the measured values and the calculated ones. To judge how calculated values fit measured ones, we used Theil decomposition [14] and then analyzed its terms (R^2, bias, slope, residuals). The most important factor is R^2, called expected variance, which tells the degree of prediction of the model. R^2 is by definition a number less equal to 1, and the more it nears 1, the better the fitting between measured and calculated data. If R^2 equals 0, the method has the same predictability of the average of data points, while if it is negative, it has a negative correlation.

In formula,

$$R^2 = 1 - \frac{\sum_{i=1...N} \left(v_i^c - v_i^0\right)^2}{\sum_{i=1...N}((v_i^0 - \bar{v}^0)^2)},$$ (18)

where v_0 are observed values and v_c are calculated values, and overline indicates average value.

As a final indicator of prediction performances, we used R^2 between observed and crossvalidated data, namely R_p^2.

Characterization

Characterization states that the resulting surface must globally look like the actual surface. Even in this case, unfortunately, we do not know the actual surface; we know only sampled point values. The idea is to associate some statistical characteristics, evaluated for a subset of the actual surface (the sampled set) to the whole surface: from a statistical point of view, that means that sampling was done randomly, so that statistical indices can be compared. This is not always the case, but that seemed a good approach to general problems.

We introduced an index to describe surface roughness of a set of data points. We defined the same roughness index (adapted to the interpolated surface) calculated on grid points:

$$R(d) = \frac{\sum_{i=1...N} \sum_{j=1...X*Y}^{d_{i,j}<d} \left((v_i - V_j)/\sqrt{2}\sigma_v\right)^2 w_\tau(D_{i,j})}{\sum_{i=1...N} \sum_{j=1...X*Y}^{0<D_{i,j}<d} w_\tau(D_{i,j})},$$ (19)

where V_j is grid value in j, X and Y are grid node coordinates (integer numbers), and $D_{i,j}$ is the distance between point data i and grid data j.

In this case, we considered both point and grid data sets, in a mixed formula. We recall that point data are experimental values, while grid data are obtained by interpolation on a regular grid.

The idea is that function $R(d)$ should be similar to function $r(d)$, or again, a surface with $R(d)$ similar to $r(d)$ should have a greater probability of being the exact surface that produced our set of point data. The comparison between $R(d)$ and $r(d)$ was done by sampling the functions for a few values of d, say d_i, greater than $s(N)$. The two sets of data, $r(d_i)$ and $R(d_i)$ were compared again with Theil decomposition and corresponding parameters were evaluated. Even in this case, visual analysis of these functions can help the understanding of the interpolation characteristics.

As a final indicator of characterization performances, we calculated the explained variance between $r(d_i)$ and $R(d_i)$, namely R_c^2.

To summarize the previous concepts, two indices help in evaluating prediction and characterization performances of the interpolation method used. Prediction index is R_p^2, while characterization index is R_c^2. The more they near 1, the better the interpolation. The indices u and $r(d)$ of point data set can suggest if R_p^2 and R_c^2 values can be considered good, even if they do not approach 1.

As a final indicator, the use of R_t^2, defined as the average of indices R_p^2 and R_c^2, was suggested:

$$R_t^2 = \frac{R_p^2 + R_c^2}{2}.$$ (20)

We were interested in positive values of R^2 indices, negative values indicating that the interpolation model is totally inefficient in reproducing experimental data. That is why we chose to consider -1 as a lower bound for all these indices.

4. RESULTS

The simulation was performed on 15 data sets, concerning a few environmental test cases. For the sake of simplicity, all data sets have the x, y coordinates normalized. It must be said that the same scale factor was used for both x and y direction.

Moreover, 18 data sets, obtained from three sets of points and six different generating functions taken from literature [1,15] were tested. This allows a more concrete possibility of comparing the study results. Even in this case, the interpolation domain ranges from 0 to 1 for both x and y coordinates. Reported here are the generating functions:

$$f_1(x,y) = 0.75e^{\left(-\left((9x-2)^2+(9y-2)^2\right)/4\right)} + 0.75e^{\left(-(9x+1)^2/49-(9y+1)/10\right)},$$
$$+ 0.5e^{\left(-\left((9x-7)^2+(9y-3)^2\right)/4\right)} - 0.2e^{\left(-(9x-4)^2-(9y-7)^2\right)},$$
$$f_2(x,y) = \frac{1}{9}\left(\tanh(9y-9x)+1\right),$$
$$f_3(x,y) = \frac{1.25+\cos(5.4y)}{6\left(1+(3x-1)^2\right)}, \tag{21}$$
$$f_4(x,y) = \frac{1}{3}e^{-(81/16)\left((x-0.5)^2+(y-0.5)^2\right)},$$
$$f_5(x,y) = \frac{1}{3}e^{-(81/4)\left((x-0.5)^2+(y-0.5)^2\right)},$$
$$f_6(x,y) = \frac{1}{9}\sqrt{64-81\left((x-0.5)^2+(y-0.5)^2\right)} - 0.5.$$

All data sets can be requested via e-mail at the following address: 1037caru@s1.cise.it. In Tables 1 and 2, point data set indices are reported for both environmental and literature test cases. For literature cases, a few points are outside the square domain; thus, to compute correctly uniform index u, x, and y coordinates were normalized, with respect to the square domain ranging from -0.1 to 1.1 in both directions, which contains all data points.

Table 1. Environmental test cases.

Case	N	Variable	u	$r(d_x)$	$r(S(N))$	d_p
BA	50	batimetry	1.1696	0.8052	0.0462	0.6813
CO	77	carbon monoxide	1.0237	1.1048	0.6600	0.1873
CS	14	sulphur dioxide	2.1681	1.0735	1.1000	N.D.
DB	30	decibel	1.4773	1.0154	0.7117	N.D.
EL	92	daytime hours	2.8281	0.9382	0.6997	N.D.
NO	107	nitrogen dioxide	1.1201	0.9453	0.4450	0.2505
O3	94	ozone	1.0379	0.9961	0.3699	0.2520
PA	39	rainfall	1.3380	1.0299	0.5654	0.2191
PG	33	rainfall	1.5125	1.0948	1.0320	N.D.
PR	168	rainfall	0.8631	0.9778	0.3651	0.2966
QU	147	quota	1.0143	0.8656	0.1091	0.5343
SA	50	salinity	1.1696	0.8403	0.0817	0.6210
SO	167	sulphur dioxide	0.7420	0.9876	0.6518	0.1612
TM	16	air temperature	1.1070	1.0730	1.4079	N.D.
TX	86	air temperature	0.8870	0.9917	0.8968	N.D.

For each data set, the following methods were applied:

1. Inverse Square Distance method (from now on WD) with $d_{\min} = 0.001$ and with five different values of $d_{\max}$ (0.10, 0.20, 0.40, 0.80, 1.50);
2. Kriging method (from now on KR) with five different values of $d_{\max}$(0.10, 0.20, 0.40, 0.80, 1.50);

Table 2. Literature test cases.

Case	N	u	$r(d_x)$	$r(S(N))$	d_p
A1	100	0.8269	0.8621	0.0871	0.5924
A2	100	0.8269	0.8621	0.0705	0.5961
A3	100	0.8269	0.9499	0.0928	0.4358
A4	100	0.8269	1.0294	0.1161	0.3802
A5	100	0.8269	1.1102	0.1775	0.3141
A6	100	0.8269	0.9332	0.1257	0.4407
B1	33	1.0835	0.8589	0.0985	0.6505
B2	33	1.0835	0.9091	0.2671	0.5804
B3	33	1.0835	0.9594	0.2588	0.4067
B4	33	1.0835	1.0188	0.4141	0.2909
B5	33	1.0835	1.1355	0.6817	0.2131
B6	33	1.0835	0.9073	0.2290	0.4076
C1	25	0.7032	0.9154	0.2762	0.4983
C2	25	0.7032	0.8867	0.2723	0.5905
C3	25	0.7032	0.9789	0.2813	0.4329
C4	25	0.7032	1.0479	0.3227	0.4466
C5	25	0.7032	1.1675	0.9561	N.D.
C6	25	0.7032	0.9262	0.1727	0.4959

Table 3. Environmental cases: average interpolation indices per test case.

Case	R_p^2	R_c^2	R_t^2
BA	0.6677	0.4006	0.5341
CO	0.8368	−0.8476	−0.0054
CS	−0.6017	−1.0000	−0.8009
DB	0.4380	−0.8816	−0.2218
EL	0.3340	−0.8405	−0.2532
NO	0.5320	−0.0618	0.2351
O3	0.6853	−0.2762	0.2046
PA	0.4676	−0.7237	−0.1281
PG	0.2195	−1.0000	−0.3903
PR	0.6561	0.5801	0.6181
QU	0.9251	0.8759	0.9005
SA	0.7589	0.7950	0.7770
SO	0.6101	0.4583	0.5342
TM	−0.2101	−1.0000	−0.6050
TX	−0.0848	−0.9707	−0.5277

Table 4. Literature cases: average interpolation indices per test case.

Case	R_s^2	R_p^2	R_c^2	R_t^2
A1	0.9561	0.9246	0.9355	0.9301
A2	0.9719	0.9441	0.9292	0.9366
A3	0.9549	0.9070	0.7567	0.8318
A4	0.9429	0.8717	0.6764	0.7740
A5	0.9140	0.8744	0.5934	0.7339
A6	0.9195	0.7444	0.4954	0.6199
B1	0.9300	0.8320	0.8529	0.8425
B2	0.9147	0.8425	0.7175	0.7800
B3	0.8082	0.7511	0.2509	0.5010
B4	0.7342	0.5704	−0.2449	0.1627
B5	0.5074	0.5624	−0.6510	−0.0443
B6	0.8138	0.2818	−0.2337	0.0241
C1	0.8608	0.5518	0.4062	0.4790
C2	0.8922	0.8041	0.6960	0.7501
C3	0.7777	0.4495	−0.3543	0.0476
C4	0.7857	0.3878	−0.4961	−0.0541
C5	0.8126	0.4044	−1.0000	−0.2978
C6	0.7277	0.1481	−0.4526	−0.1523

3. Hardy's Multiquadric method (from now on MQ), with no parameters;
4. Tension Finite Difference method (from now on TM), with five different values of T (0.00, 0.20, 0.40, 0.70, 1.00).

Globally, 16 different method versions were executed for each test case. For each execution, a regular grid of 26 × 26 values, with a step of 0.04 units in both x and y directions, was generated. Method versions will be indicated with the first letter of method, followed by three digits for the parameter value.

In Tables 3 and 4, the average value obtained by the 16 method versions for indices R_p^2, R_c^2, and R_t^2 is reported for both environmental and literature test cases. For literature cases, for which the actual surface is known, another index, R_s^2, was introduced, obtained via Theil decomposition by comparing surface values and grid values for all 26 × 26 cell values.

Looking at Tables 1 and 2, the test cases that seem easier for interpolation (u almost 1, $r(d_x)$ and $r(S(N))$ low, d_p high) are BA, SA, QU, A1, B1, A2, C2, and A3. Conversely, the cases that appear worse for interpolation (u possibly not very close to 1, $r(d_x)$ and $r(S(N))$ high, d_p low or not defined) are CS, DB, SO, TX, PG, TM, B5, and C5. On average, these considerations are confirmed by indices of Tables 3 and 4.

In Tables 5 and 6, average methods and versions performances are reported for both environmental and literature test cases. Average method performances were calculated by averaging interpolation indices for a single method, varying its parameter, on all experimental cases. On the contrary, version performances were calculated by averaging interpolation indices on all test cases, fixing the method and its parameter. Even in this case, R_s^2 is also reported, for literature data. Method ranking changes meaningfully in the two classes of test cases. This happens because some methods work better with difficult test cases, but have worse performances with simpler data sets. Literature data are on average simple cases, while environmental data can have difficult situations to interpolate.

Table 5. Environmental cases: average interpolation indices per method version.

Method	R_p^2	R_c^2	R_t^2
K010	0.3686	−0.5169	−0.1483
K020	0.3996	−0.0879	0.3117
K040	0.3114	0.0026	0.3140
K080	0.3917	0.0859	0.4776
K150	0.4074	0.0568	0.4642
M000	0.3897	−0.3854	0.0043
T000	0.4494	−0.3122	0.1373
T020	0.4555	−0.3335	0.1220
T040	0.4610	−0.3662	0.0949
T070	0.4666	−0.4549	0.0117
T100	0.4570	−0.4393	−0.0823
W010	0.4624	−0.5458	−0.0834
W020	0.4646	−0.5832	−0.1186
W040	0.4463	−0.6073	−0.1610
W080	0.4352	−0.6238	−0.1886
W150	0.4325	−0.6283	−0.1958
KR	0.3758	−0.0919	0.2838
MQ	0.3897	−0.3854	0.0043
TM	0.4579	−0.4012	0.0567
WD	0.4482	−0.5977	−0.1495

Table 6. Literature cases: average interpolation indices per method version.

Method	R_s^2	R_p^2	R_c^2	R_t^2
K010	0.5065	0.2674	−0.2926	−0.0126
K020	0.7780	0.4919	0.0516	0.2717
K040	0.9635	0.7929	0.5910	0.6920
K080	0.9769	0.9013	0.6259	0.7636
K150	0.9804	0.9314	0.6205	0.7636
M000	0.9582	0.8803	0.6311	0.7557
T000	0.8869	0.6729	0.5104	0.5916
W010	0.8916	0.7068	0.2493	0.4780
W020	0.8545	0.6699	0.0406	0.3552
W040	0.8128	0.6111	−0.1472	0.2320
W080	0.7905	0.5603	−0.2294	0.1654
W150	0.7882	0.5620	−0.2454	0.1583
KR	0.8410	0.6770	0.3193	0.4981
MQ	0.9582	0.8803	0.6311	0.7557
TM	0.8463	0.6320	0.3103	0.4711
WD	0.8275	0.6220	−0.0664	0.2778

Table 7. Environmental cases: best/worst interpolations.

Method	Best Cases			Worst Cases		
	R_p^2	R_c^2	R_t^2	R_p^2	R_c^2	R_t^2
K010	0	0	0	6	2	5
K020	1	2	2	1	0	0
K040	0	1	1	4	1	4
K080	3	4	6	0	0	0
K150	4	2	1	0	0	0
M000	0	4	0	2	9	2
T000	2	0	0	0	0	0
T020	0	0	0	0	0	0
T070	1	0	0	0	0	0
T100	1	0	1	1	0	0
W010	1	0	1	0	0	0
W020	1	1	2	0	1	0
W040	0	1	0	0	0	0
W080	0	0	0	0	0	1
W150	1	0	1	1	2	3
KR	8	9	10	11	3	9
MQ	0	4	0	2	9	2
TM	4	0	1	1	0	0
WD	3	2	4	1	3	4

In Tables 7 and 8, the number of best and worst cases for methods and versions are reported for both environmental and literature test cases. Again, for literature data, this information is also reported for R_s^2. Method ranking is less different than in Tables 5 and 6 for the two classes of test cases. That means that parameter tuning can have great influence on method performances.

Thus, a few considerations can be obtained for methods WD, KR, and MT, the ones that depend on a parameter value. WD works well with low values of d_{max}, because the effect of smoothing is reduced. KR works well with high values of d_{max}, because all spatial correlation is considered. MT, on average, works well with low values of T, even if best cases are more often obtained for high values of T.

In any case, all the remarks obtained from analysis of Tables 5–8 are not meant to suggest to the reader a general methods ranking, but only a means of interpretation of proposed indices.

At the end of this section, it seems useful to present the analysis of a single test case. Case B1 was chosen for two reasons. First, the knowledge of actual surface can help evaluating the results, second, case B1 shows different performances when different evaluation criteria are adopted. In Table 9, the interpolation indices for all 16 versions are reported. In Figures 3–6, the resulting surface, obtained by four interpolation method versions, compared with the actual surface, are shown. In these figures, shaded colours indicate the actual surface, contour lines indicate the interpolated surface, and markers indicate sampled points.

5. CONCLUSIONS

Spatial interpolation is a technique used to evaluate physical data in a continuous domain. In this paper, a comparison of interpolation methods was proposed.

First of all, the point data set is analyzed in order to obtain useful information for the interpolation process.A few indices can suggest if methods are likely to produce a good interpolated

Table 8. B1 Literature cases: best/worst interpolations.

Method	Best Cases				Worst Cases			
	R_s^2	R_p^2	R_c^2	R_t^2	R_s^2	R_p^2	R_c^2	R_t^2
K010	0	0	0	0	16	15	12	13
K020	0	0	1	0	0	0	0	0
K040	3	2	2	1	0	0	0	0
K080	6	5	9	7	0	0	0	0
K150	9	11	1	7	0	0	0	0
M000	0	0	2	2	0	0	1	0
T000	0	0	1	1	1	1	0	0
T020	0	0	0	0	0	0	0	0
T070	0	0	0	0	0	0	0	0
T100	0	0	1	0	0	0	0	0
W010	0	0	1	0	0	0	0	0
W020	0	0	0	0	0	0	0	0
W040	0	0	0	0	0	0	0	0
W080	0	0	0	0	0	0	0	0
W150	0	0	0	0	1	2	5	5
KR	18	18	13	15	16	15	12	13
MQ	0	0	2	2	0	0	1	0
TM	0	0	2	1	1	1	0	0
WD	0	0	1	0	1	2	5	5

Table 9. B1 literature case interpolation indices.

| Method	R_s^2	R_p^2	R_c^2	R_t^2
M000	0.9635	0.9306	0.8863	0.9085
K010	0.7920	0.6367	0.6879	0.6623
K020	0.9308	0.7827	0.8221	0.8024
K040	0.9905	0.9491	0.8971	0.9231
K080	0.9686	0.9440	0.9226	0.9333
K150	0.9785	0.9710	0.9147	0.9429
W010	0.9089	0.8778	0.9491	0.9134
W020	0.9238	0.8578	0.8951	0.8765
W040	0.9337	0.8581	0.8142	0.8362
W080	0.9316	0.8381	0.7405	0.7893
W150	0.9305	0.8359	0.7210	0.7785
T000	0.9332	0.7670	0.9185	0.8427
T020	0.9372	0.7821	0.9039	0.8430
T040	0.9385	0.7866	0.8872	0.8369
T070	0.9309	0.7694	0.8578	0.8136
T100	0.8881	0.7258	0.8284	0.7771
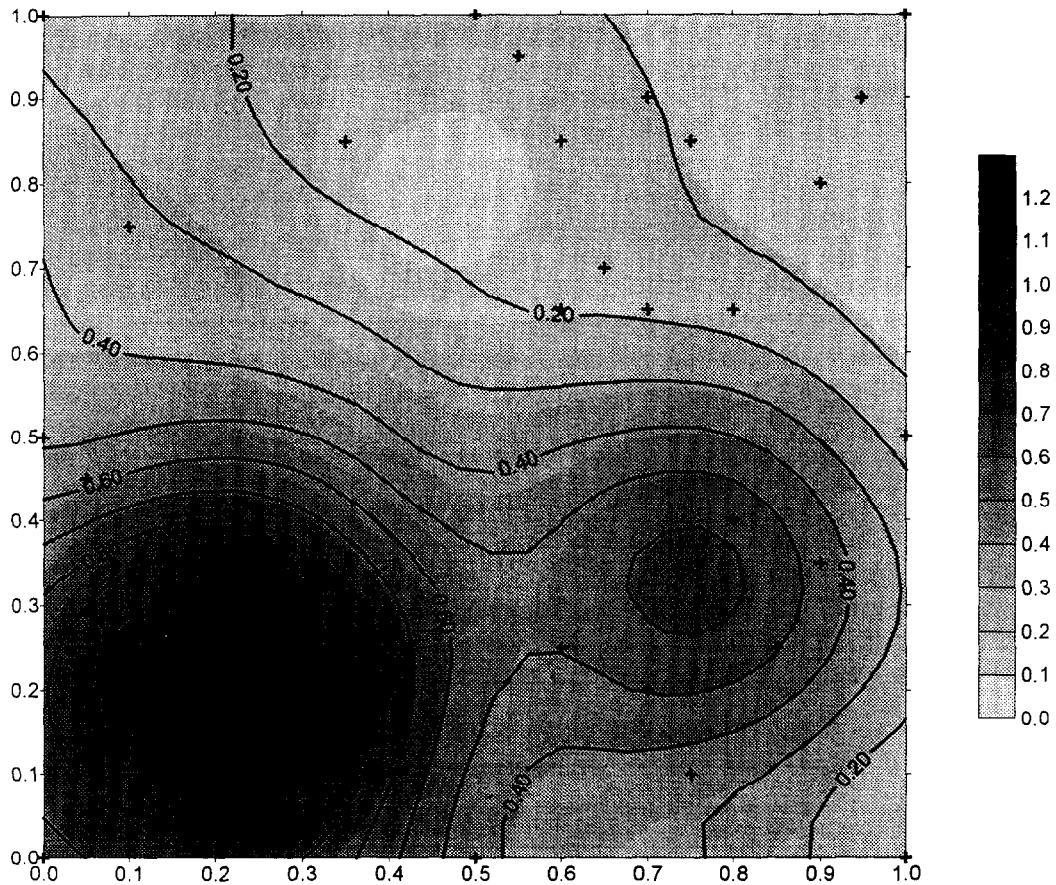
Figure 3. B1 literature case interpolation by Kriging method ($d_{max} = 0.4$).

surface, or if data do not show a significant spatial correlation. In any case, these indices can also suggest which kind of method would probably have better performances.

After creating the grid, other indices can be used to judge the interpolation performances, evaluated from different viewpoints. Sometimes, different purposes can result in different method ranking.

This approach seems to give a means of evaluating the whole interpolation process in order to choose a good method. In any case, this approach is not meant as a substitution for the judgment of experts in the fields investigated, who can use different and often more reliable criteria in evaluating a specific case.

APPENDIX

The derivation of formulas for calculation of uniform index u and for roughness index $r(d)$ are here reported.

Uniform Index u

The average number of points that fall into a circle of radius $s(N)$ in the normalized square domain is obtained comparing the circle area, $\pi s^2(N)$, with the domain area, 1. But the area that overlaps the domain can be less than $\pi s^2(N)$, because of border effects. We introduced the function $g(x, y, N)$ that measures the circle area inside the domain, as a function of x, y and of the number of data points N. It is defined differently according to the subregion in which it is considered. For convenience, the square domain is divided into four different subregions, as in Figure 7. To avoid a misleading definition of subregions, $s(N)$ should be less than 0.5, thus resulting in $N \geq 4$.
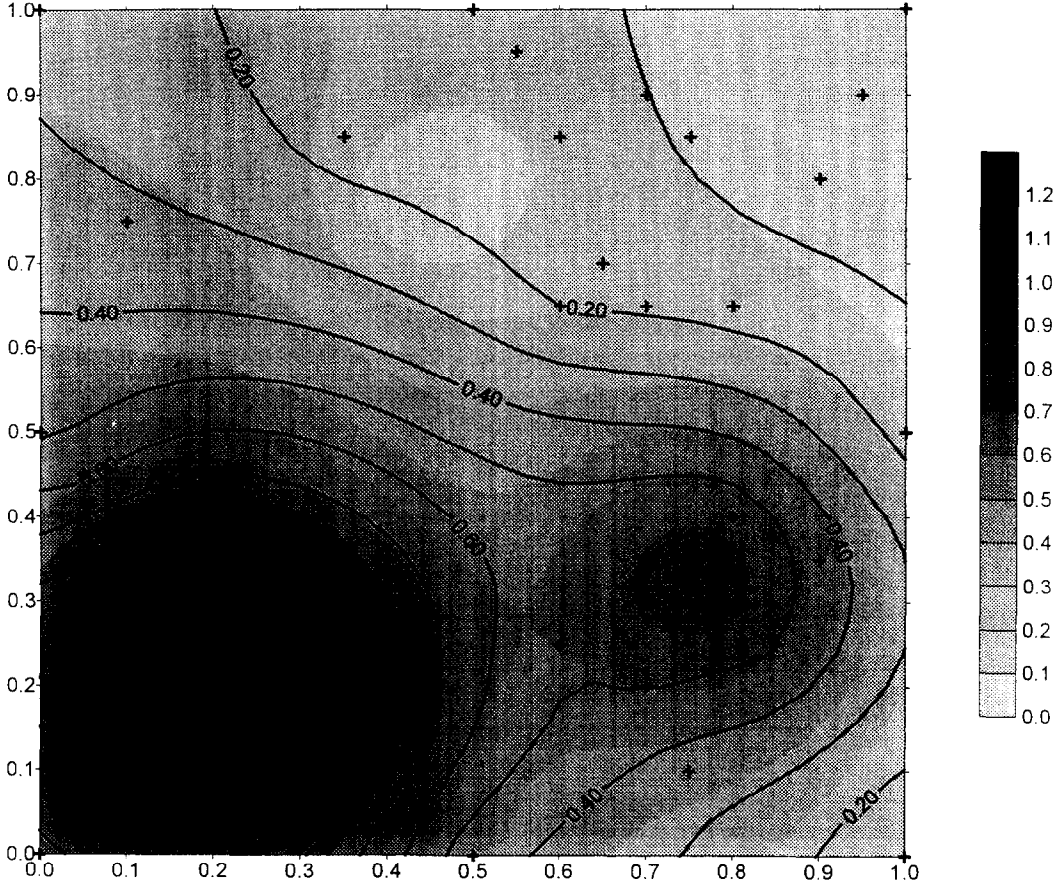
Figure 4. B1 literature case interpolation by Hardy's multiquadric method.

In subregion a, the circle is completely inside the domain; in b, the circle has a sector outside the domain; in c, the circle contains a vertex of the domain, and in subregion d, the circle has two sectors outside the domain. The following formula is computed for the shaded subregions (the others are symmetric):

$$g(x, y, N) = \begin{cases} \frac{\pi}{N}, & \text{with } (x, y) \in a, \\ \frac{\pi + x\sqrt{1-x^2} - \arccos(x)}{N}, & \text{with } (x, y) \in b, \\ \frac{3\pi/4 + xy + x\sqrt{1-x^2}/2 + y\sqrt{1-y^2} - (\arccos(x)/2) - (\arccos(y)/2)}{N}, & \text{with } (x, y) \in c, \\ \frac{\pi + x\sqrt{1-x^2} + y\sqrt{1-y^2} - \arccos(x) - \arccos(y)}{N}, & \text{with } (x, y) \in d. \end{cases} \tag{22}$$

These equations are obtained with trigonometric formulas. Thus, integrating $g(x, y, N)$ over the domain, we obtain:

$$n(N) = N \int_0^1 \int_0^1 g(x, y, N) \, dx \, dy. \tag{23}$$

The tedious calculations to obtain the simplified expression of $n(N)$ presented in the paper are not reported here, but can be verified by the interested reader.

Roughness Index $r(d)$ for $w_r(d) = 1$

The index $r(d)$ is here evaluated for $d > d_{\max}$ and $w_r(d) \equiv 1$. In this case, we obtain the simplified formula:

$$r(d) = \frac{\sum_{i=1...N} \sum_{j=1...N}^{j \neq i} (v_i - v_j)^2}{2N(N-1)\sigma_v^2}. \tag{24}$$
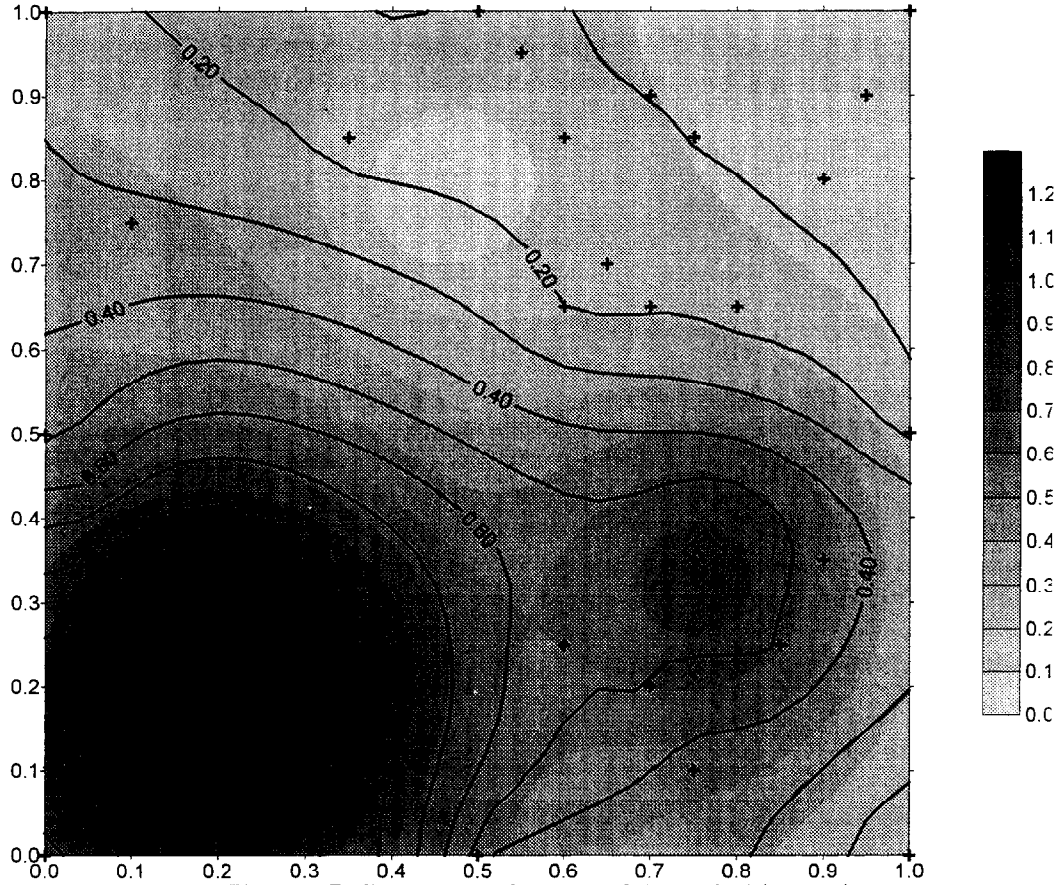
Figure 5. B1 literature case by tension finite method $(T = 0.4)$.

Introducing the obvious identity:

$$(v_i - v_j) \equiv ((v_i - \bar{v}) - (v_j - \bar{v}))$$ (25)

we obtain a new formulation for the expression:

$r(d) =$

$$\frac{\sum_{i=1...N} \sum_{j=i...N}^{j \neq i} (v_i - \bar{v})^2 + \sum_{i=1...N} \sum_{j=1...N}^{j \neq i} (v_j - \bar{v})^2 - 2 \sum_{i=1...N} \sum_{j=i...N}^{j \neq i} (v_i - \bar{v})(v_j - \bar{v})}{2N(N-1)\sigma_v^2}.$$
                                                                          (26)

First and second terms of the numerator can be simplified in this way:

$$\sum_{i=1...N} \sum_{j=i...N}^{j \neq i} (v_i - \bar{v})^2 = \sum_{i=1...N} (v_i - \bar{v})^2 \sum_{j=i...N}^{j \neq i} 1 = N(N-1)\sigma_v^2$$ (27)

while the third term is

$$-2 \sum_{i=1...N} \sum_{j=i...N}^{j \neq i} (v_i - \bar{v})(v_j - \bar{v}) = -2 \sum_{i=1...N} (v_i - \bar{v}) \sum_{j=1...N}^{j \neq i} (v_j - \bar{v})$$

$$= -2 \sum_{i=1...N} (v_i - \bar{v}) \left(\sum_{j=1...N} (v_j - \bar{v}) - (v_i - \bar{v}) \right) = -2 \sum_{i=1...N} (v_i - \bar{v}) (0 - (v_i - \bar{v}))$$ (28)

$$= +2 \sum_{i=1...N} (v_i - \bar{v})^2 = +2N\sigma_v^2.$$

Thus, adding the three terms, the formula is

$$r(d) = \frac{N(N-1)\sigma_v^2 + N(N-1)\sigma_v^2 + 2N\sigma_v^2}{2N(N-1)\sigma_v^2} = \frac{2N^2\sigma_v^2}{2N(N-1)\sigma_v^2} = \frac{N}{N-1}.$$ (29)
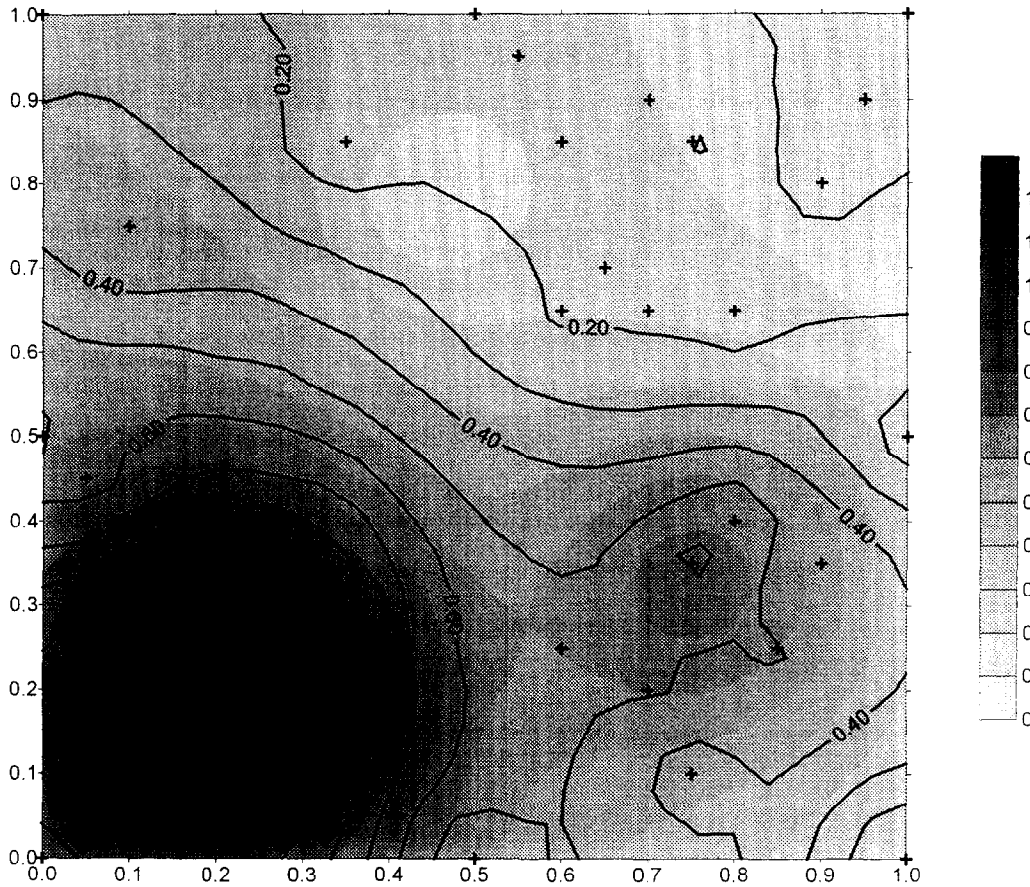
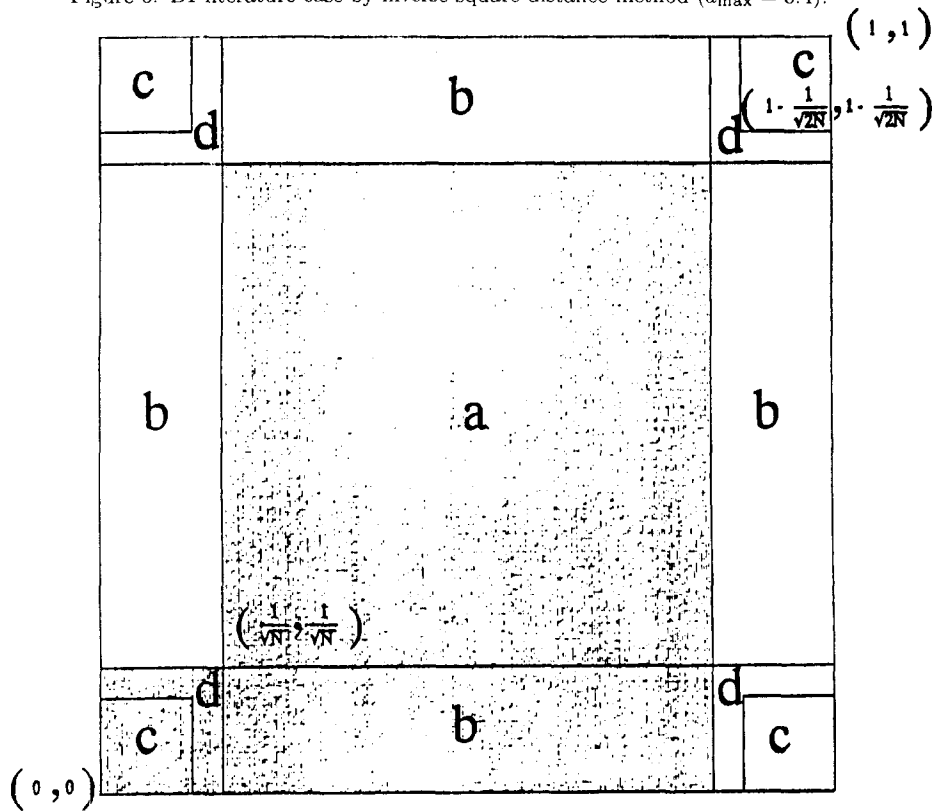Figure 6. B1 literature case by inverse square distance method ($d_{max} = 0.4$).

Figure 7. Subregions of the normalized square domain.

REFERENCES

1. R. Franke, A Critical Comparison of Some Methods for Interpolation of Scattered Data, TR No. NPS-53-79-003, Naval Postgraduate School.
2. N.S. Lam, Spatial interpolation methods: A review, *The American Cartographer* 10 (2), 129–149, (1983).
3. I.K. Crain and B.K. Bhattacharyya, Treatment of non-equispaced two-dimensional data with a digital computer, *Geoexploration* 5 (4), 173–194, (1967).
4. D.G. Krige, A review of the development of geostatistics in South Africa, In *Advanced Geostatistics in the Mining Industry*, (Edited by M. Guarascio, M. David and C. Huijbregts) pp. 279–293, Holland, (1976).
5. G. Matheron, The theory of regionalized variables and its application, *Les Cahiers du Centre de Morphologie Mathematique de Fontainebleau* 5, 221.
6. F. Trochu, A contouring program based on dual Kriging interpolation, *Engineering with Computers* 9 (3), 160–177, (1993).
7. G. Matheron, Splines et Krigeage: leur equivalence formelle, In Rapport N-667, Centre de Geostatistique, Ecole des Mines de Paris, (1980).
8. R.L. Hardy, Multiquadric equations of topography and other irregular surfaces, *Journal of Geophysics Resource* 76 (8), 1905–1915, (1971).
9. I.C. Briggs, Machine contouring using minimum curvature, *Geophysics* 39, 39–48, (1974).
10. W.H.F. Smith and P. Wessel, Gridding with continuous curvature splines in tension, *Geophysics* 55 (3), 293–305, (1990).
11. P. Wessel and W.H.F. Smith, *The GMT System v.2.1.4 Technical Reference & Cookbook*, SOEST/NOAA, (1993).
12. G. Bastin and M. Gevers, Identification and optimal estimation of random fields from scattered pointwise data, *Automatica* 21 (2), 139–155, (1985).
13. G.Q. Tabios and J.D. Salas, A comparative analysis of techniques for spatial interpolation of precipitation, *Water Resources Bulletin* 21 (3), 365–380, (1985).
14. H. Theil, *Applied Economic Forecasting*, North-Holland, (1963).
15. R. Franke, Scattered data interpolation: Tests of some methods, *Mathematics of Computation* 38 (157), 181–200, (1982).