



# A Novel Formulation for Inverse Distance Weighting from Weighted Linear Regression

Leonardo Ramos Emmendorfer<sup>(✉)</sup> and Graçaliz Pereira Dimuro

Center for Computational Sciences, Universidade Federal do Rio Grande,  
Rio Grande, RS 96203900, Brazil  
leonardo.emmendorfer@gmail.com, gracaliz@gmail.com

**Abstract.** Inverse Distance Weighting (IDW) is a widely adopted interpolation algorithm. This work presents a novel formulation for IDW which is derived from a weighted linear regression. The novel method is evaluated over study cases related to elevation data, climate and also on synthetic data. Relevant aspects of IDW are preserved while the novel algorithm achieves better results with statistical significance. Artifacts are alleviated in interpolated surfaces generated by the novel approach when compared to the respective surfaces from IDW.

**Keywords:** Inverse distance weighting · Interpolation · Weighted linear regression · Digital elevation map

## 1 Introduction

Most natural properties vary continuously. However, in general, we can observe at only a finite number of the infinity of possible locations [20]. Spatial interpolation is the estimation of approximate values for specific locations from known values measured at other locations. Given a set of spatial data either in the form of discrete points or for subareas, spatial interpolation aims to find the function that will best represent the whole surface and that will predict values at other points or for other subareas [14]. This general problem has long been a concern majorly in geosciences, water resources, environmental sciences, agriculture, soil sciences among other disciplines [15, 29].

Environmental data collected from field surveys are often difficult and expensive to acquire. In such cases, spatial interpolation methods provide a tool for estimating an environmental variable at unsampled sites [15]. For instance, in [11] as a result from the sparsity of observational networks the distance to the nearest station can be of the order of several hundred kilometers. As a result, the only available data may not be representative of the climatology at the desired location. Ideally, the nearest recording station would be situated such that its climatology was identical to that of the location of interest.

Point interpolation deals with data collectable at a point, such as temperature readings or elevation [14]. Several solutions are available, such as Kriging [13, 17], interpolating polynomials, splines, among others [7]. Inverse distance weighting (IDW) [25] is one of the most simple and widespread adopted [15]. The method does not require specific statistical assumptions, as the case for Kriging and other statistical interpolation methods. However, although empirical evaluations consistently show that IDW delivers inferior results when compared to other methods [15, 19, 30], the evaluation of improvements in IDW is a relevant topic of research [2, 9, 16, 22, 28].

The IDW interpolation of a value  $\hat{y}_j$  for a given location  $j$  is computed as:

$$\hat{y}_j^{IDW} = \sum_{i=1}^n w_{i,j} y_i \quad (1)$$

where each  $y_i$ ,  $i = 1, \dots, n$  is a data point available at a location  $i$ . The weights  $w_{i,j}$  for each data point are given as:

$$w_{i,j} = \frac{d_{i,j}^{-\alpha}}{\sum_{k=1}^n d_{k,j}^{-\alpha}} \quad (2)$$

where  $d_{i,j}$  is the Euclidean distance between a data point available at location  $i$  and the unknown data at location  $j$ ;  $n$  is the number of data points available;  $\alpha$  means the power, and is a control parameter. In this work, IDW is restricted to Inverse Squared-Distance Weighting since  $\alpha = 2$  is assumed, which is the most commonly adopted value.

The maximum and minimum of the estimated values from IDW are limited to the extreme data points:  $\min y_i \leq \hat{y}_j^{IDW} \leq \max y_i$ . This is considered to be an important shortcoming because, to be useful, an interpolated surface should predict accurately certain important features of the original surface, such as the locations and magnitudes of maxima and minima even when they are not included as original sample points [14].

This work aims to (i) introduce an alternative interpolation algorithm which is similar to IDW and (ii) evaluate the novel method under a variety of conditions considering diverse of sampling densities, sample spatial distributions and surface types. Those are pointed out as important factors that affect the performance of spatial interpolation methods [15].

The paper is organized as follows. The proposed approach is presented in Sect. 2. The resulting model is evaluated and compared to the original IDW in Sect. 4, following the methodology proposed in Sect. 3. Section 5 concludes the paper.

## 2 Proposed Method

Consider a variable  $Y$  which is measured at  $n$  locations. One might be interested in obtaining an estimation for the value of  $Y$  at a specific location  $j$ , where a value for  $Y$  is not available for some reason.

Let us assume that variable  $Y$  is related to a function of the distance to  $j$ . This leads to a model which represents the relationship between the variable  $Y$ , which occurs at diverse locations, and a single explanatory variable which is a function of the distance from a given reference  $j$  to the location of each available measure of  $Y$ . One might assume, for instance, that squared distance from  $j$  influences  $Y$  as:

$$\mathbf{Y} = \beta_j^0 + \beta_j^1 \mathbf{D}_j + \mathbf{E}_j \tag{3}$$

where coefficients  $\beta_j^0$  and  $\beta_j^1$  are both scalars which must be obtained for each  $j$ .  $\mathbf{Y} = \{y_1, y_2, \dots, y_n\}$  is a vector with  $n$  values of the variable under consideration at diverse locations  $i = 1, 2, \dots, n$  and the corresponding vector  $\mathbf{D}_j = \{d_{1,j}^2, d_{2,j}^2, \dots, d_{n,j}^2\}$  contains the squared distances  $d_{i,j}^2$  from location  $j$  to each location  $i$  corresponding to a respective  $y_i$ .  $\mathbf{E}_j = \{\epsilon_1, \epsilon_2, \dots, \epsilon_n\}$  is the vector of residues.

The estimation of the scalars  $\beta_j^0$  and  $\beta_j^1$  from (3) can be achieved by solving a weighted linear regression, where the regression weights  $w_{1,j}^R, w_{2,j}^R, \dots, w_{n,j}^R$  for a given  $j$  are computed similarly to the IDW weights in (2) with  $\alpha = 2$ :

$$w_{i,j}^R = \frac{d_{i,j}^{-2}}{\sum_{k=1}^n d_{k,j}^{-2}} \tag{4}$$

For the sake of clarity, let us define the scalar variable  $s_j$  for a given  $j$  as:

$$s_j = \frac{1}{\sum_{k=1}^n d_{k,j}^{-2}} \tag{5}$$

Then, substituting (5) on (4):

$$w_{i,j}^R = d_{i,j}^{-2} s_j \tag{6}$$

The weighted sum of squared residuals (WSSE) of model (3) for data points  $\{y_1, y_2, \dots, y_n\}$  is given by:

$$WSSE = \sum_{i=1}^n w_{i,j}^R (y_i - \hat{y}_i)^2 = \sum_{i=1}^n w_{i,j}^R (y_i - \beta_j^0 - \beta_j^1 d_{i,j}^2)^2 \tag{7}$$

Substituting (6) into (7) leads to:

$$WSSE = \sum_{i=1}^n d_{i,j}^{-2} s_j (y_i - \beta_j^0 - \beta_j^1 d_{i,j}^2)^2$$

where the analytical solution for the minimal WSSE is:

$$\hat{\beta}_j^0 = s_j \sum_{i=1}^n y_i d_{i,j}^{-2} - n \beta_j^1 s_j \tag{8}$$

$$\hat{\beta}_j^1 = \frac{\sum_{i=1}^n y_i - n s_j \sum_{i=1}^n y_i d_{i,j}^{-2}}{\sum_{i=1}^n d_{i,j}^2 - n^2 s_j} \tag{9}$$

The estimated value for  $Y$  as a function  $\hat{f}$  of the distance  $r$  from  $j$  using the model (3) is:

$$\hat{f}(r) = \hat{\beta}_j^0 + \hat{\beta}_j^1 r^2 \tag{10}$$

Since the aim of interpolation is the estimation of a value for  $Y$  at  $j$ , therefore the distance is  $r = 0$ . Then, from (10):

$$\hat{y}_j^R = \hat{f}(0) = \hat{\beta}_j^0 + \hat{\beta}_j^1 0^2 = \hat{\beta}_j^0 \tag{11}$$

Substituting (9) into (8) and (11) leads, after simplification, to the expression for the interpolated value at a given location  $j$ , from a set of values  $\{y_1, y_2, \dots, y_n\}$  and respective distances  $\{d_{1,j}, d_{2,j}, \dots, d_{n,j}\}$  from  $j$  :

$$\hat{y}_j^R = s_j \sum_{i=1}^n y_i d_{i,j}^{-2} + n \frac{\sum_{i=1}^n y_i - n s_j \sum_{i=1}^n y_i d_{i,j}^{-2}}{n^2 - \sum_{i=1}^n d_{i,j}^{-2} \sum_{i=1}^n d_{i,j}^2} \tag{12}$$

From (1), (2) and (5) one can find out that  $s_j \sum_{i=1}^n y_i d_{i,j}^{-2} = \hat{y}_j^{IDW}$  (with  $\alpha = 2$ ), therefore (12) can be rewritten as:

$$\hat{y}_j^R = \hat{y}_j^{IDW} + n \frac{\sum_{i=1}^n y_i - n \hat{y}_j^{IDW}}{n^2 - \sum_{i=1}^n d_{i,j}^{-2} \sum_{i=1}^n d_{i,j}^2} \tag{13}$$

The resulting expression, which is derived from a weighted linear regression, results equivalent to IDW with an additional term. We call this method as Inverse Distance Weighted Regression (IDWR). More specifically, since  $\alpha$  was set to 2, this paper investigates Inverse Squared-Distance Weighted Regression.

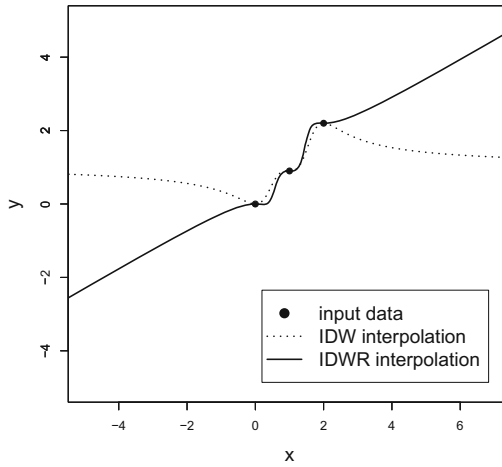
**2.1 Analysis of IDWR**

Similarly to IDW, IDWR is also a deterministic, nonstatistical interpolation method, defined by a simple expression (13). The computational complexity for interpolating a single location  $j$  for IDWR is  $O(n)$ , linear in the number of data points  $n$ , which is the same for IDW.

This section presents an initial analysis of some relevant situations. Initially, the form of expression 13 raises some concerns as the denominator might be equal to or near zero. For instance, when all data points are or tend to be at the same distance  $r$  from location  $j$ , the denominator is or tends to be equal to  $n^2 - \sum_{i=1}^n r^{-2} \sum_{i=1}^n r^2 = n^2 - n r^{-2} n r^2 = 0$ . While this situation would not be expected in most real-world applications, even when input data is distributed on a bidimensional regular grid, this feature of IDWR must be carefully taken into account before using the method. Also, one can realize that as the distance  $r \rightarrow \infty$  additional numerical concerns might arise since  $\sum_{i=1}^n r^{-2} \sum_{i=1}^n r^2 \rightarrow 0 \times \infty$ . This differs from IDW, which tends to  $\frac{1}{n} \sum_{i=1}^n y_i$  as  $r \rightarrow \infty$ .

The behavior of IDWR at the neighborhood of any given data point is also analysed. We are interested in the value of  $\hat{y}_j^R$  as  $d_{lj} \rightarrow 0$  for a given data point at location  $l$ , with  $d_{ij} \neq 0$  for the remaining data points  $i \neq l$ . Since  $d_{lj} \rightarrow 0$ ,

then  $\sum_{i=1}^n d_{ij}^{-2} \rightarrow \infty$  in expression 13 and  $\sum_{i=1}^n d_{ij}^2 \rightarrow c$  where  $c = \sum_{i \neq l} d_{ij}^2$  is a constant. This results  $\hat{y}_j^R \rightarrow \hat{y}_j^{IDW}$  in expression 13, since the denominator tends to  $-\infty$ , under the condition that the numerator should be finite. As a result IDW and IDWR will tend to compute similar values for locations which are nearby any given data point. IDWR is an exact interpolator since  $\hat{y}_j^R = \hat{y}_j^{IDW} = y_i$  for  $j = i$ . At other locations, IDWR might be able to provide useful extrapolation, since  $-\infty \leq \hat{y}_j^R \leq +\infty$ , differently from IDW which is restricted to the interval  $\min y_i \leq \hat{y}_j^{IDW} \leq \max y_i$ . From the discussion above, any differences between both methods might occur at locations that are not too close to any data point.



**Fig. 1.** The behavior of IDW and IDWR for the interpolation from a dataset with  $n = 3$  data points.

Figure 1 illustrates some of the properties discussed here using a synthetic one-dimensional dataset with three data points that follow a linear trend ( $R^2 > 0.99$ ).

### 3 Empirical Evaluation

Two types of experiments were performed, which allow one to compare the effectiveness of both algorithms considered. The first evaluation involves the interpolation of points from real functions of two variables. The functions were selected from the optimization literature, as representatives of varying roughness of surfaces, so as to impose different levels of difficulty for the interpolation methods. While those functions would not perfectly mimic real-world situations, this evaluation is still useful for the purpose of this work since it provides a scalable comparison between the two methods, through a controlled variation on the number of samples. In this first experiment sample size was set to four

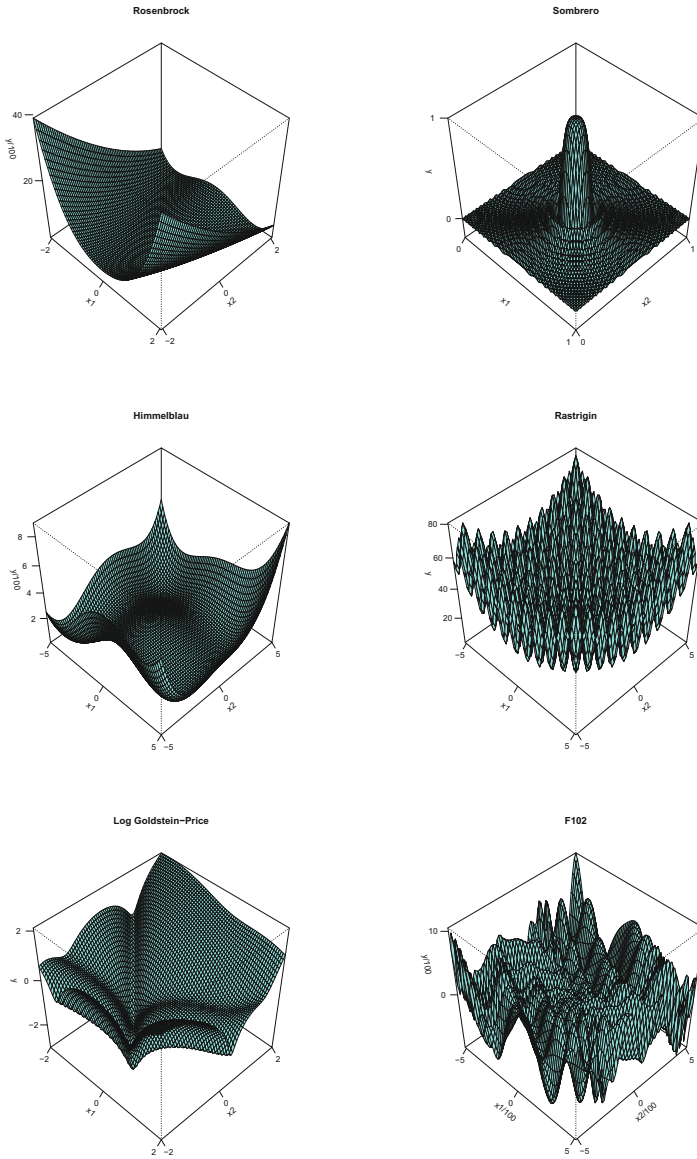
**Table 1.** Functions of two real variables  $(x_1, x_2) = \mathbf{x}$ , adopted in empirical evaluation

Function	Expression	Interval for $x_1$ and $x_2$
Rosenbrock	$y(\mathbf{x}) = 100(x_2 - x_1^2)^2 + (x_1 - 1)^2$	$[-2.048, 2.048]$
Sombbrero	$y(\mathbf{x}) = \begin{cases} \frac{\sin((16(x_1-0.5))^2 + (16(x_2-0.5))^2)}{16(x_1-0.5)^2 + (16(x_2-0.5))^2} & \text{if } x_1 \neq 0.5 \text{ and } x_2 \neq 0.5; \\ 1 & \text{otherwise} \end{cases}$	$[0, 1]$
Himmelblau	$y(\mathbf{x}) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$	$[-5, 5]$
Rastrigin	$y(\mathbf{x}) = 20 + (x_1^2 - 10 \cos(2\pi x_1)) + (x_2^2 - 10 \cos(2\pi x_2))$	$[-5.12, 5.12]$
Log Goldstein-Price	$y(\mathbf{x}) = \frac{1}{2.427} (\log((1 + (x_1 + x_2 + 1)^2 \times (19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2)) \times (30 + (2x_1 - 3x_2)^2 \times (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)))) - 8.693)$	$[-2, 2]$
F102	$y(\mathbf{x}) = -(x_2 + 47) \sin \sqrt{ x_2 + \frac{x_1}{2} + 47 } - x_1 \sin \sqrt{ x_1 - (x_2 + 47) }$	$[-512, 512]$

values:  $N = 100, 200, 300, 400$ . The variation on sample size is motivated by the need for capturing spatial changes, thus to improve the performance of the spatial interpolation methods [15].

**Table 2.** Average RMSE and standard deviation computed with leave-one-out cross-validation (LOOCV) for IDW and IDWR applied to 6 benchmark functions, after 30 replications with randomly generated sample points for each benchmark function. The number of sample points for all functions is  $N = 300$  at each replication. P-values refer to the result of two-tailed t-tests considering the null hypothesis that algorithms are equivalent in terms of average RMSE

Function	Avg. IDW LOOCV RMSE	$\sigma_{IDW}$	Avg. IDWR LOOCV RMSE	$\sigma_{IDWR}$	Relative Reduction	p-value
Rosenbrock	307.65	29.02	222.52	27.48	-27.67%	<2.2e-16
Sombbrero	0.083277	0.0089	0.0806	0.0087	-3.20%	1.016e-06
Himmelblau	76.61	3.96	64.75	3.90	-15.48%	<2.2e-16
Rastrigin	16.51	0.68	16.25	0.74	-1.59%	1.728e-07
Log Goldstein-Price	0.6036	0.02698	0.4378	0.02439	-27.47%	<2.2e-16
F102	391.11	16.93	388.39	16.89	-0.70%	3.379e-09



**Fig. 2.** Perspective visualization of the 6 functions used for the evaluation of the proposed algorithm.

**Table 3.** RMSE computed with leave-one-out cross-validation (LOOCV) for IDW and IDWR applied to 2 benchmark datasets from the literature

<i>Dataset</i>	<i>N</i>	<i>IDW LOOCV RMSE</i>	<i>IDWR LOOCV RMSE</i>	<i>Relative Reduction</i>
Calabria	48	67.16	65.72	-2.14%
Texas	18	11.09	8.63	-28.51%

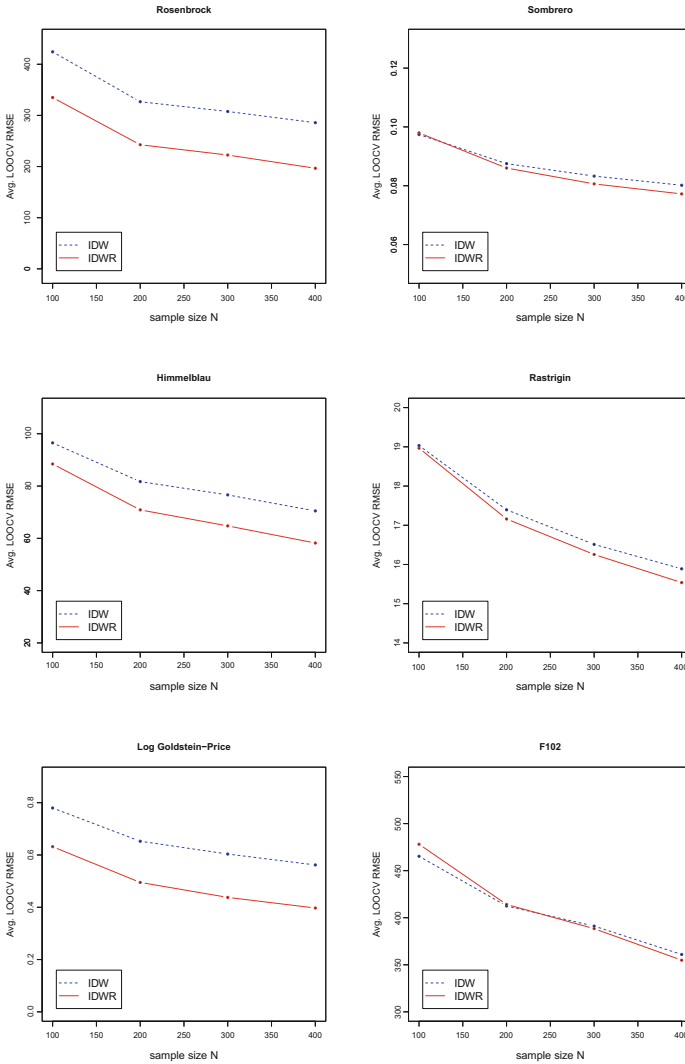
Table 1 summarizes the definitions of the functions adopted. Figure 2 provides a perspective visualization of the topology of those functions. Himmelblau [10], Rosenbrock [24] and Rastrigin [23] are non-linear, non-convex functions widely used to test the performance of optimization algorithms. The 2-dimensional version of Rastrigin is used here. Log Goldstein-Price is an adjusted version of the Goldstein-Price function [8] proposed by [21]. The function F102 [1] was also called *Egg Holder* in [27] and in other works. It is considered as a difficult function due to its high multimodality. The Sombrero function was also included in our evaluation since it was already adopted as a benchmark for evaluation of IDW, in [30].

In a second type of evaluation two datasets representing real-world situations from the literature are considered. The *Calabria* dataset, adapted from [5], is a raster low-resolution (100m) digital elevation map containing 48 elevations which vary from 760 m to 936 m. The sample area from a location in Calabria is 610 m by 810 m in size, which corresponds to a portion of sample area 1 in [5]. The *Texas* dataset contains normal annual precipitation (1941–1970) for 18 locations in Texas, which is the full list of locations from [3]. The lowest annual precipitation (7.7in) occurs in El Paso, near the western extreme of the state, while the highest precipitation is assigned to Beaumont-Port Arthur, near the eastern extreme (55.07in).

In order to allow the comparison between the interpolation methods, leave-one-out cross-validation (LOOCV) [12] was adopted. In LOOCV, a single data point  $y_i$  is used for the estimation of the squared error of the interpolation  $(y_i - \hat{y}_i)^2$  from a model built from all remaining points  $N - 1$  points. The process is repeated for all data points, and the root mean square error (RMSE) is computed, for both interpolation methods considered.

Since the computation of the RMSE for the evaluation of the interpolation of real functions is dependent on the specific sample of data points, 30 replications of leave-one-out cross-validation are performed for each algorithm on each function, in order to estimate the average RMSE for a number of  $N$  data points. Those data points are randomly generated from uniform distributions delimited by the specified real intervals for each variable.





**Fig. 3.** Average RMSE and standard deviation computed with leave-one-out cross-validation (LOOCV) for IDW and IDWR applied to 6 benchmark functions, after 30 replications with randomly generated sample points for each benchmark function. The number of sample points for all functions was set to  $N = 100, 200, 300, 400$  at each replication.

## 4 Results

Table 2 shows the results from the first set of experiments, where interpolation is performed from points sampled from functions defined over the bidimensional domain. Average RMSE and respective standard deviation  $\sigma$  are computed for 30

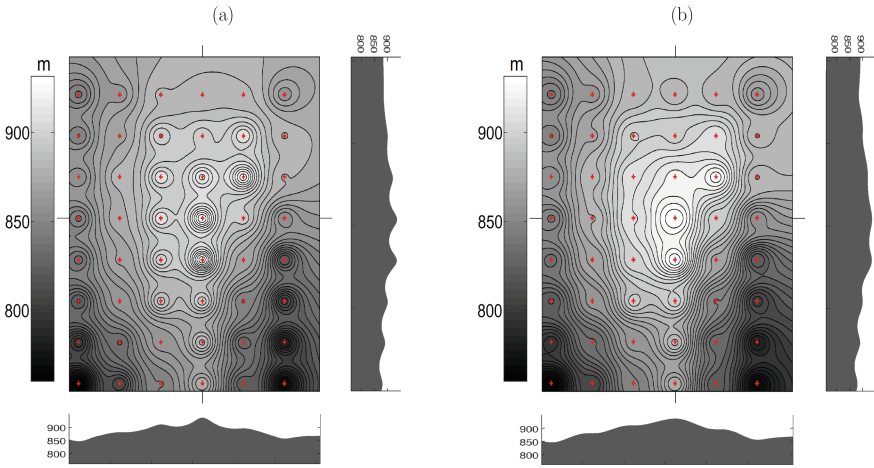
replications of leave-one-out cross-validation on the interpolation of data points from 6 functions for both algorithms considered. The number of data points in each replication was set to  $N = 300$ . The relative reductions on the values of the average RMSE for IDWR when compared to IDW are also shown. Resulting reductions range from 0.70% (F102) to 27.67% (Rosenbrock). All differences between the mean RMSE values are statistically significant at a 95% confidence level, considering paired two-tailed t-tests under the null hypothesis that both methods are equivalent.

The effect of sample size is illustrated in Fig. 3. For all functions 4 sample sizes were considered:  $N = 100, 200, 300, 400$ . RMSE is lower for IDWR when compared to IDW for all functions with all  $N$  considered, except for  $N = 100$  and  $N = 200$  where the best RMSE for the F102 function is achieved with IDW. For  $N > 200$  IDWR is superior for all functions. The tendency from the graphs in Fig. 3 is also favorable to IDWR for  $N > 400$ .

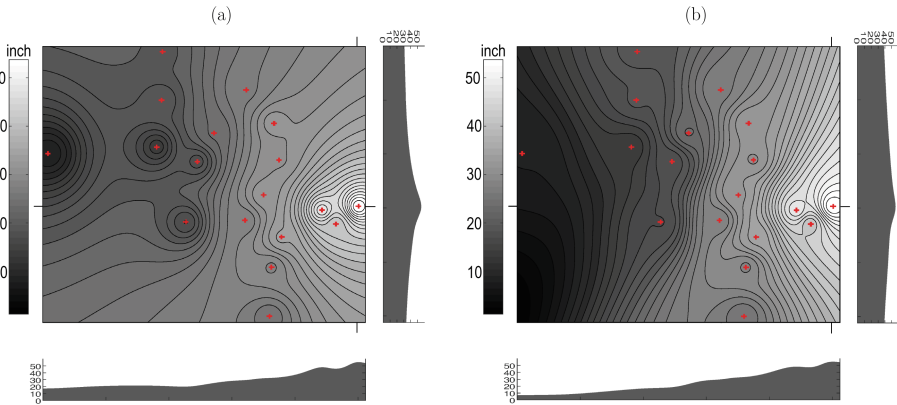
In Table 3 the values of LOOCV RMSE for both algorithms applied to two datasets considered are shown. Under this evaluation, IDWR is superior to IDW for both datasets. The error for *Calabria* dataset is 2.14% lower when compared to IDW. A higher difference was reached for the *Texas* dataset, where IDWR achieved a 28.51% reduction in the LOOCV RMSE when compared to the value obtained with IDW for the same dataset.

In order to allow a better understanding of the behavior of each algorithm, interpolated surfaces were generated for the sample areas related to each both datasets considered. For *Calabria*, two digital elevation maps with a 1 m resolution were obtained representing the interpolated surfaces obtained using both algorithms for the input data, which consists of a digital map with elevations from 48 locations regularly distributed with a resolution of 100m. This high difference between input and output resolution might not be recommended. However, for the purpose of this evaluation, the approach allows a better visual comparison between the results obtained by both methods. Figure 4 shows the resulting maps for the region on the *Calabria* dataset using both IDW and IDWR (Figs. 4(a) and 4(b) respectively).

The highest elevation in *Calabria* dataset is located near the center of the maps, as indicated. It also corresponds to the maximal value obtained from IDWR and also from IDW. The same occurs for the lowest elevation, which occurs at a location near the right bottom extreme of the map. Therefore, IDWR did not exceed the IDW limitations  $\min y_i$  and  $\max y_i$  for this case. Although both maps from *Calabria* are similar, qualitative differences in the behavior of the algorithms occur. The surface generated by IDWR is smoother, with smaller variations on the curvature over the space. As a result, the interpolated surface from IDWR appears as more conceivable when compared to the result from IDW. The surface generated with IDWR is smoother since artificial bumps generated between sample points are less evident. However, undesirable artifacts exist since both algorithms produce unrealistic landscape, with a terraced aspect. Elevation profiles below and beside both maps (a) and (b) in Fig. 4 provide a better illustration for this feature.



**Fig. 4.** High resolution interpolated elevation maps generated by IDW (a) and IDWR (b), for the area of *Calabria* dataset. 48 regularly distributed sample points are shown in red and elevation values are represented in grayscale levels discretized into 40 intervals with increments of  $\approx 4.4$  m. For each map, two elevation profiles (bottom and right) are shown, each parallel to a coordinate axis and both passing through the coordinates corresponding to the highest elevation in the dataset, indicated at the border of the maps.



**Fig. 5.** High resolution interpolated precipitation maps generated by IDW (a) and IDWR (b), for the area of *Texas* dataset. Sample points are shown in red and elevation values are represented in grayscale levels discretized into 40 intervals with increments of  $\approx 1.35$  in. For each map, two elevation profiles (bottom and right) are shown, each parallel to a coordinate axis and both passing through the location corresponding to the highest precipitation in the dataset, indicated at the border of the maps (Beaumont-Port Arthur). (Color figure online)

The dataset *Texas* represents a situation where a low amount of data points is available which leads to the absence of data points in some areas since large regions outside the territory of Texas are represented in the interpolated maps. Figures 5(a) and 5(b) both represent an area of size  $1258 \text{ km} \times 1060 \text{ km}$  with a resolution of 2 km. The resulting map from IDWR provides a better model for the expected behavior of precipitation from given data. Precipitation decreases roughly towards the west or south-west, reaching predicted values as low as 1.139in at where would correspond to the territory of Mexico, which is below the minimal precipitation from the dataset (7.7in).

## 5 Conclusion and Further Work

The selection of an appropriate interpolation model depends largely on the type of data, the degree of accuracy desired, and the amount of computational effort afforded [14]. Each method has its advantages and drawbacks, which depend strongly on the characteristics of the data: a method that fits well with some data can be unsuited for a different set of data points [6]. This also motivates the improvement of existing methods and search for novel alternatives.

Variations and extensions from the basic IDW method have been proposed in the literature. In [2] an improvement is presented which is based on a geometric criterion that automatically selects a subset of the original set of control points. In [22] data normalization is shown to improve the results of interpolation. In [9] weighted median of data within a neighborhood is proposed. A distance-decay parameter is explored in [16] which is adjusted according to the spatial pattern of sampled locations in the neighborhood.

This paper followed a diverse path by presenting a novel formulation that is derived from a weighted regression model where squared distance from the location of interest is assumed to influence a geographically localized variable. Resulting expression (13) is similar to IDW method while retaining its simplicity and low computational complexity. Squared distance was arbitrarily chosen, and other formats for that relationship might be explored further.

Regression is already widely adopted for problems involving spatial data. Geographically Weighted Regression (GWR), as proposed by [4], adopts weighted regression in the spatial context by extending the usual regression model. The regression coefficients are dependent on individual location and the parameters in GWR are therefore locally estimated by weighted least squares approach where the weight is higher for observations that are closer to the location considered. That premise of a higher local relationship [26] which is straightforwardly implemented by IDWR and IDW is already widely exploited [18].

Empirical evaluation of the proposed method adopted leave-one-out cross-validation using datasets from the literature and synthetic data from benchmark functions, with varying sample densities on diverse surface types and sample distributions. Study cases emphasized applications on digital elevation data and climate.

IDWR was able to attain better results when compared to IDW by obtaining lower RMSE with statistical significance for benchmark functions. Qualitatively,

the novel method delivered smoother curvatures between sample points when compared to the maps generated by IDW. Observable artifacts are alleviated in the surfaces generated by IDWR.

Further empirical and theoretical investigation should be proposed to better delineate the limitations of the novel method. It might also be studied whether the proposed method actually produces useful extrapolation. In that case, wider applicability would be reached when compared to IDW. This, however, must be carefully considered since the asymptotic behavior of IDWR is much diverse from IDW, according to the discussion in Sect. 2. A comparison to other interpolation methods could also be performed, covering a wider variety of applications.

## References

1. Evaluating evolutionary algorithms: *Artif. Intell.* **85**(1), 245–276 (1996)
2. Ballarin, F., D’Amario, A., Perotto, S., Rozza, G.: A POD-selective inverse distance weighting method for fast parametrized shape morphing. *Int. J. Numer. Methods Eng.* **117**(8), 860–884 (2019)
3. Bomar, G.W.: A climatological summary of Texas weather in 1977. Technical report (1978)
4. Brunsdon, C., Fotheringham, A.S., Charlton, M.E.: Geographically weighted regression: a method for exploring spatial nonstationarity. *Geograph. Anal.* **28**(4), 281–298 (1996)
5. Carrara, A., Bitelli, G., Carla, R.: Comparison of techniques for generating digital terrain models from contour lines. *Int. J. Geograph. Inf. Sci.* **11**(5), 451–473 (1997)
6. Caruso, C., Quarta, F.: Interpolation methods comparison. *Comput. Math. Appl.* **35**(12), 109–126 (1998)
7. Cressie, N.: *Statistics for Spatial Data*. Wiley, Hoboken (1993)
8. Dixon, L.C.W., Szegő, G.P.: The global optimization problem: an introduction. In: *Towards Global Optimization*, vol. 2. North-Holland, Amsterdam (1978)
9. Henley, S.: *Nonparametric Geostatistics*. Springer, Heidelberg (2012)
10. Himmelblau, D.: *Applied Nonlinear Programming*. McGraw-Hill, New York (1972)
11. 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.* **16**(4), 309–330 (2001)
12. Kohavi, R., et al.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Proceedings of the International Joint Conference on Artificial Intelligence*, Montreal, Canada, vol. 14, pp. 1137–1145 (1995)
13. Krige, D.: A review of the development of geostatistics in South Africa. In: Guarascio, M., David, M., Huijbregts, C. (eds.) *Advanced Geostatistics in the Mining Industry*, vol. 24, pp. 279–293. Springer, Heidelberg (1976). <https://doi.org/10.1007/978-94-010-1470-0.17>
14. Lam, N.S.N.: Spatial interpolation methods: a review. *Am. Cartographer* **10**(2), 129–150 (1983)
15. Li, J., Heap, A.D.: A review of comparative studies of spatial interpolation methods in environmental sciences: performance and impact factors. *Ecol. Inform.* **6**(3–4), 228–241 (2011)
16. Lu, G.Y., Wong, D.W.: An adaptive inverse-distance weighting spatial interpolation technique. *Comput. Geosci.* **34**(9), 1044–1055 (2008)

17. Matheron, G.: The theory of regionalised variables and its applications. *Les Cahiers du Centre de Morphologie Mathématique* **5**, 212 (1971)
18. Miller, H.J.: Tobler's first law and spatial analysis. *Ann. Assoc. Am. Geograph.* **94**(2), 284–289 (2004)
19. Murphy, R.R., Curriero, F.C., Ball, W.P.: Comparison of spatial interpolation methods for water quality evaluation in the Chesapeake Bay. *J. Environ. Eng.* **136**(2), 160–171 (2010)
20. Oliver, M.A., Webster, R.: Kriging: a method of interpolation for geographical information systems. *Int. J. Geograph. Inf. Syst.* **4**(3), 313–332 (1990)
21. Picheny, V., Wagner, T., Ginsbourger, D.: A benchmark of Kriging-based infill criteria for noisy optimization. *Struct. Multidisc. Optim.* **48**(3), 607–626 (2013)
22. Qu, R., et al.: Predicting the hormesis and toxicological interaction of mixtures by an improved inverse distance weighted interpolation. *Environ. Int.* **130**, 104892 (2019)
23. Rastrigin, L.A.: Extremal control systems. *Theoretical Foundations of Engineering Cybernetics Series*. Nauka, Moscow (1974). (in Russian)
24. Rosenbrock, H.H.: An automatic method for finding the greatest or least value of a function. *Comput. J.* **3**(3), 175–184 (1960)
25. Shepard, D.: A two-dimensional interpolation function for irregularly-spaced data. In: *Proceedings of the 1968 ACM National Conference*, pp. 517–524 (1968)
26. Tobler, W.R.: A computer movie simulating urban growth in the detroit region. *Econ. Geogr.* **46**(sup1), 234–240 (1970)
27. Vanaret, C., Gotteland, J.B., Durand, N., Alliot, J.M.: Certified global minima for a benchmark of difficult optimization problems (2014). <https://hal-enac.archives-ouvertes.fr/hal-00996713>
28. Weber, D., Englund, E.: Evaluation and comparison of spatial interpolators. *Math. Geol.* **24**(4), 381–391 (1992)
29. Zhou, F., Guo, H.C., Ho, Y.S., Wu, C.Z.: Scientometric analysis of geostatistics using multivariate methods. *Scientometrics* **73**(3), 265–279 (2007)
30. Zimmerman, D., Pavlik, C., Ruggles, A., Armstrong, M.P.: An experimental comparison of ordinary and universal Kriging and inverse distance weighting. *Math. Geol.* **31**(4), 375–390 (1999)