Computers & Geosciences 34 (2008) 1044–1055

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

An adaptive inverse-distance weighting spatial interpolation technique

George Y. Lu [1], David W. Wong *

Earth Systems & GeoInformation Sciences, George Mason University, 4400 University Drive, MS 6A2, Fairfax, VA 22030, USA

ARTICLE INFO

Article history:
Received 20 September 2006
Received in revised form
11 July 2007
Accepted 12 July 2007

Keywords:
Spatial interpolation
Distance-decay parameter
Cross validation
Adaptive inverse-distance weighting
Kriging

ABSTRACT

One of the most frequently used deterministic models in spatial interpolation is the inverse-distance weighting (IDW) method. It is relatively fast and easy to compute, and straightforward to interpret. Its general idea is based on the assumption that the attribute value of an unsampled point is the weighted average of known values within the neighborhood, and the weights are inversely related to the distances between the prediction location and the sampled locations. The inverse-distance weight is modified by a constant power or a distance-decay parameter to adjust the diminishing strength in relationship with increasing distance. Recognizing the potential of varying distance-decay relationships over the study area, we suggest that the value of the weighting parameter be allowed to vary according to the spatial pattern of the sampled points in the neighborhood. This adaptive approach suggests that the distance-decay parameter can be a function of the point pattern of the neighborhood. We developed an algorithm to search for "optimal" adaptive distance-decay parameters. Using cross validation to evaluate the results, we conclude that adaptive IDW performs better than the constant parameter method in most cases, and better than ordinary kriging in one of our empirical studies when the spatial structure in the data could not be modeled effectively by typical variogram functions.

1. Introduction

The inverse-distance weighting (IDW) method, a deterministic spatial interpolation model, is one of the more popular methods adopted by geoscientists and geographers partly because it has been implemented in many GIS packages. The general premise of this method is that the attribute values of any given pair of points are related to each other, but their similarity is inversely related to the distance between the two locations. However, many studies, especially in the spatial interaction literature, have revealed that the decline in spatial relationship between any two locations is not simply proportional to distance (e.g., Fotheringham and O'Kelly, 1989). As a result, a power or exponential function modifying the distance weight is often used to model spatial interaction between places. In applying the IDW method, such a function is often considered when predicting the unknown attribute values at certain locations. For example, Bekele et al. (2003) used inverse-distance weights of powers 1, 2, and 3 to map soil potassium. Ping and Green (2004) used inverse-distance weights of powers 1 through 5 to determine the spatial weights matrix for modeling autocorrelation functions (Moran's *I* and Geary's *C*) when exploring spatial dependencies in cotton yield. Lloyd (2005) used a power of 2, a frequently used value, to interpolate precipitation.

Despite its popularity, IDW suffers from several limitations, including the fact that the weighting parameters are chosen *a priori*, but not empirically determined. In addition, IDW cannot estimate the variances of predicted values in

* Corresponding author. Tel.: +1 703 993 9260; fax: +1 703 993 9299.
E-mail addresses: GeorgeLu@mapmodteam.com (G.Y. Lu), dwong2@gmu.edu (D.W. Wong).
[1] Currently works at Michael Baker Corporation.

0098-3004/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cageo.2007.07.010

unsampled locations as compared to what geostatistical methods such as kriging can provide (Burrough and McDonnell, 1998). Another limitation is that the distance-decay parameter is applied uniformly throughout the entire study area without considering the distribution of data within it. In other words, the standard application of IDW assumes that the distance-decay relationship is constant over space, though it may not be the case. A constant distance-decay value could be part of the reason that IDW provides less accurate predictions as compared to other interpolation methods (Goovaerts, 2000; Lloyd, 2005).

In this paper, we propose an adaptive inverse-distance weighting (AIDW) method to take advantage of the computational simplicity of IDW, but to provide the additional flexibility to accommodate variability in the distance-decay relationship over the study area. The general concept is similar to an adaptive kernel estimation method that adjusts the kernel size based upon the density of the observations (Bracken and Martin, 1989; Bailey and Gatrell, 1995). As with any interpolation method, the adaptive IDW method proposed here will not work well in some situations, but will perform quite well especially when local variability is relatively large. In the next section, we briefly review the IDW method and compare this method with ordinary kriging—one of the most commonly employed geostatistical methods. We also elaborate on the major weaknesses of IDW. In Section 3, we present the formulation of the adaptive IDW method. In Section 4 we use two examples to illustrate how the AIDW method can be applied and compare its performance with other methods. The paper ends with a conclusion and discussion section.

2. Spatial interpolation methods: a brief review

The IDW method is a straightforward and non-computationally intensive method. It has been regarded as one of the standard spatial interpolation procedures in geographic information science (Burrough and McDonnell, 1998; Longley et al., 2001) and has been implemented in many GIS software packages. As a result, many GIS users without much background in spatial statistics and geostatistics will use IDW as a default method to generate a surface when attribute values are available only at sampled locations. Formally, the IDW method is used to estimate the unknown value $\hat{y}(S_0)$ in location S_0, given the observed y values at sampled locations S_i in the following manner:

$$\hat{y}(S_0) = \sum_{i=1}^{n} \lambda_i y(S_i) \tag{1}$$

Essentially, the estimated value in S_0 is a linear combination of the weights (λ_i) and observed y values in S_i where λ_i is often defined as

$$\lambda_i = d_{0i}^{-\alpha} / \sum_i^n d_{0i}^{-\alpha} \tag{2}$$

with

$$\sum_i^n \lambda_i = 1. \tag{3}$$

In Eq. (2), the numerator is the inverse of distance (d_{0i}) between S_0 and S_i with a power α, and the denominator is the sum of all inverse-distance weights for all locations i so that the sum of all λ_i's for an unsampled point will be unity (Eq. (3)). The α parameter is specified as a geometric form for the weight while other specifications are possible. This specification implies that if α is larger than 1, the so-called distance-decay effect will be more than proportional to an increase in distance, and vice versa. Thus, small α tends to yield estimated values as averages of S_i in the neighborhood, while large α tends to give larger weights to the nearest points and increasingly down-weights points farther away. Therefore, when $\alpha \to 0$ and $\lambda_i = 1/n$, then

$$\begin{aligned} y(S_0) &= \sum_{i=1}^{n} \lambda_i y(S_i) \\ &= \sum_{i=1}^{n} \frac{1}{n} y(S_i) \end{aligned} \tag{4}$$

Then, the estimated value is the average of all sampled values. Let S_j be the nearest neighbor to S_0, and L_i denote the distance between S_0 and S_i, then $\min\{L_i\} = L_j$. When $\alpha \to \infty$, the weight will be defined in the following manner:

$$\lambda_i = \begin{cases} 1 & i = j(L_j = \min\{L_i\}) \\ 0 & i \neq j \end{cases} \tag{5}$$

and

$$\begin{aligned} \hat{y}(S_0) &= \sum_{i=1}^{n} \lambda_i y(S_i) \\ &= y(S_j) \end{aligned} \tag{6}$$

In this case, the estimated value will be the same as the value in the nearest sampled location j.

IDW is a deterministic method as compared to geostatistical methods such as kriging (Bailey and Gatrell, 1995). On the other hand, kriging is the preferred method in many spatial statistical analyses. It is relatively cumbersome. Though there are many types of kriging, they are very much based on similar concepts. One has to first examine the data to identify the spatial structure, which is often represented by the empirical variogram (Isaaks and Srivastava, 1989). Then given the empirical variogram that captures the spatial autocorrelation structure of the data, a mathematical function is used to fit the empirical variogram as the theoretical variogram function to model spatial autocorrelation. Based upon this function, weights are derived. The linear or non-linear combinations of these weights and the corresponding observed values are used to estimate values at unsampled locations. In reference to the IDW method as depicted in Eq. (1), kriging requires the extra steps to derive the empirical variogram and to fit a function to the variogram from which the weights λ_i are derived. One clear advantage of kriging is that kriging variance is readily available during the computational process and it offers the confidence level of the interpolation, in addition to other strengths (Zimmerman et al., 1999). For IDW,

jackknife approach will be needed to derive the confidence intervals for the estimates (Tomczak, 1998).

From a statistical perspective, kriging is a sound method. In practice, "the success of kriging maps is dependent on the suitability of the theoretical semivariogram to the data at hand." (Sen and Sahin, 2001, p. 12). Therefore, identifying the most appropriate theoretical variogram for the given data is critical. In general, we expect that the level of spatial autocorrelation diminishes as spatial lag increases, and changes in spatial autocorrelation level over different spatial lags are captured and represented by the variogram. Unfortunately, the spatial structure of the data may not conform to the general structure in many empirical studies due to various reasons, including poor data quality. Another possible reason is the presence of spatial heterogeneity. As a result, the theoretical variogram function may not provide a good description of the empirical data. Still spatial interpolation can be performed using kriging (e.g., Mulholland et al., 1998) despite the poor fit of the variogram model at the local scale. Nevertheless, the estimation results may not be satisfactory. In other words, if the variogram does not reflect the spatial structure of the data adequately, kriging may not provide good predictive results. Different kriging methods have been suggested to address the non-stationarity problem, but the treatment of this problem is somewhat ac hoc (e.g., Goovaerts, 1997, Chapters 5 and 6; Sampson et al., 2001; Rivoirard, 1994).

On the other hand, the assumption of IDW is simple. Though the method assumes that Tobler's first law of geography (Tobler, 1970) is true in the data, there is no need to identify a theoretical distribution for the data. The method does not involve computationally intensive procedures, such as inverting the covariance matrix in kriging. An obvious drawback of IDW is that the inverse-distance weights are not determined by the empirical data of the study area. The proposed adaptive IDW method is intended to modify the weights with information derived from the empirical data.

Many forms of inverse-distance weights have been proposed to alleviate the limitations of using the simplistic inverse-distance weights discussed above (Cressman, 1959; Gandin, 1970; Barnes, 1964). Still, some major drawbacks of the IDW framework persist (Sen and Sahin, 2001). Specifically, the general IDW framework does not take into account the spatial variability of the distance-decay relationship of the phenomenon under study. Instead, it assumes that the distance-decay structure in Eq. (2) is uniform throughout the entire study area, and that the weighting forms are correct depictions of the spatial correlation structure.

As a result, methods have been proposed to address some of these drawbacks of IDW. For example, Henley (1981) proposed a non-parametric approach based on the weighted median of data within a neighborhood. The weight is the product of inverse-distance weights and the de-clustered weights that include the effects of distance and clustering among spatially correlated data in the estimator. Weber and Englund (1992) proposed the use of piecewise least-square polynomial regression estimators to increase the accuracy of the predictions.

In this paper, we explore another approach, suggested by the kernel density estimation literature that takes into account the spatial variability of point density. Instead of using a constant distance-decay parameter value in Eq. (2) under the assumption that the distance-decay relationship is uniform across the surface of the study region, the parameter value is modified according to the point pattern within the neighborhood of the unsampled location. In other words, we suggest that the distance-decay parameter should be a function of the neighborhood point pattern. This local modeling approach is in line with the recent development of stressing the local in spatial analysis and statistics (Fotheringham, 1997), as many spatial statistical measures were developed to evaluate the local or neighborhood situation (e.g., Getis and Ord, 1992). The adaptive distance-decay parameter α will be assigned according to the spatial pattern of neighborhood points to reflect the regional variability of point distribution. Cross validation is used to evaluate the results, and the performance among kriging, IDW with constant distance-decay parameters, and adaptive distance-decay parameters are compared accordingly.

3. Methodology

When IDW is applied, regardless of what value that α may carry, the general assumption is that the estimated value $\hat{y}$ at location S_0 will draw more weight from closer locations than from farther locations. In practice, when two unsampled locations have different intensities of sampled points in their neighborhoods, the same distance-decay parameter will be applied in both cases. There are some clear disadvantages of applying the same distance-decay value and disregarding the spatial pattern of the sampled locations around the unsampled point. For an unsampled location with highly clustered points in its neighborhood, there is no need to use a relatively large α to estimate the unknown value, as all neighboring values will carry relatively large weights because they are close to the estimated location (Fig. 1). Instead, a small distance-decay value will be desirable so that nearest sampled values will not have an overwhelming influence on the estimated value. This idea is similar to the de-clustering notion in geostatistical analysis.

On the other hand, when the point pattern is relatively dispersed, then the more reliable source for the estimate will likely come from the closest locations (Fig. 1). Therefore, if a relatively small value is used, the contributions from the local and more reliable sources will be relatively small, and will result in less reliable estimates. Thus, when the pattern is dispersed, a larger α will be more desirable so that the local influence will be accounted for more strongly than influences from more distant locations.

In this paper, we explore how the distance-decay parameter can be adjusted according to the spatial pattern of sampled locations in the neighborhood of the unsampled points. The ultimate goal is to achieve a better predictive power using the IDW framework.

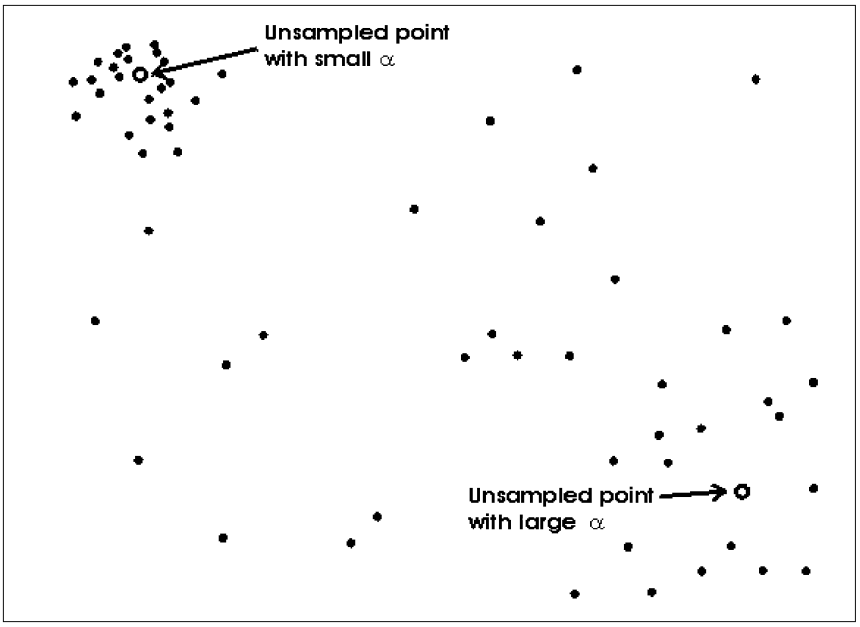

Fig. 1. Distribution of neighboring sampled locations and size of α.

3.1. Spatial pattern determination

Theoretically, one may use all sampled points in a study region to derive the estimated value in S_0 when using IDW, though values in closer locations will contribute more to the estimate of $\hat{y}$. However, in practice, in order to come up with estimates efficiently, only a subset of observed points will be needed and analysts or users will need to pre-determine the number of neighbors to be included in the estimation process. In some cases, instead of fixing the number of neighbors to be included, a predetermined search distance (buffer) around the unsampled locations can be used. However, too small a distance may end up with no sampled points in the neighborhood. For illustration purposes, we will use the five nearest points in the estimation process. Other values can be used, but the actual number of search neighbors or how the search neighbors are defined will not affect the process we discuss here. In other words, a neighborhood in this illustration is the area where the five closest neighbors are located.

In order to determine how the distance-decay parameter should be adjusted according to the spatial pattern of the sampled locations around the unsampled point, we have to quantify the spatial pattern of the sampled locations first. Borrowing well-established concepts in point pattern analysis (Boots and Getis, 1988), the nearest neighbor statistic can assist in this process. The general idea of the nearest neighbor statistic is based on comparing the observed average nearest neighbor distance with the expected nearest neighbor distance, which is the average nearest neighbor distance of a random point pattern. If the observed average nearest distance is greater than the expected, then we can say that the observed point pattern is more dispersed than a random pattern. If the opposite is true, then we may claim that there is a clustered pattern. The observed average nearest neighbor distance, r_{obs}, can be calculated by taking the average of the nearest neighbor distances for all points. The expected nearest neighbor distance for a random pattern can be derived from

$$r_{exp} = 1/(2(n/A)^{0.5}) \tag{7}$$

where n is the number of points in the study area and A is the area of the study region. With both the observed and the expected nearest neighbor distances, the nearest neighbor statistic, R, can be expressed as

$$R = r_{obs}/r_{exp} \tag{8}$$

The smaller the R, the more clustered is the pattern, and vice versa.

In the context of deriving an adaptive IDW method, r_{obs} is calculated using only the five points closest to the unsampled location. For the expected nearest neighbor distance, we still use the one based upon a random pattern (r_{exp}) for the entire area. Therefore, the R values for the neighborhoods are governed by the local observed average nearest neighbor distances, r_{obs}. For each unsampled point, the five nearest sampled points are first selected. The average nearest neighbor distance is computed for these five selected points. In this case, this specific nearest neighbor statistic can be labeled as $R(S_0)$, denoting the nearest neighbor statistic for the neighborhood of S_0. The purpose of deriving this local statistic is not to test for point pattern, but to provide a simple descriptive summary measure of the point pattern in the neighborhood. It is true that r_{exp} is a constant for different locations of S_0 and therefore $R(S_0)$ is a function of the numerator only. But for easy interpretation, we still use $R(S_0)$ instead of the numerator.

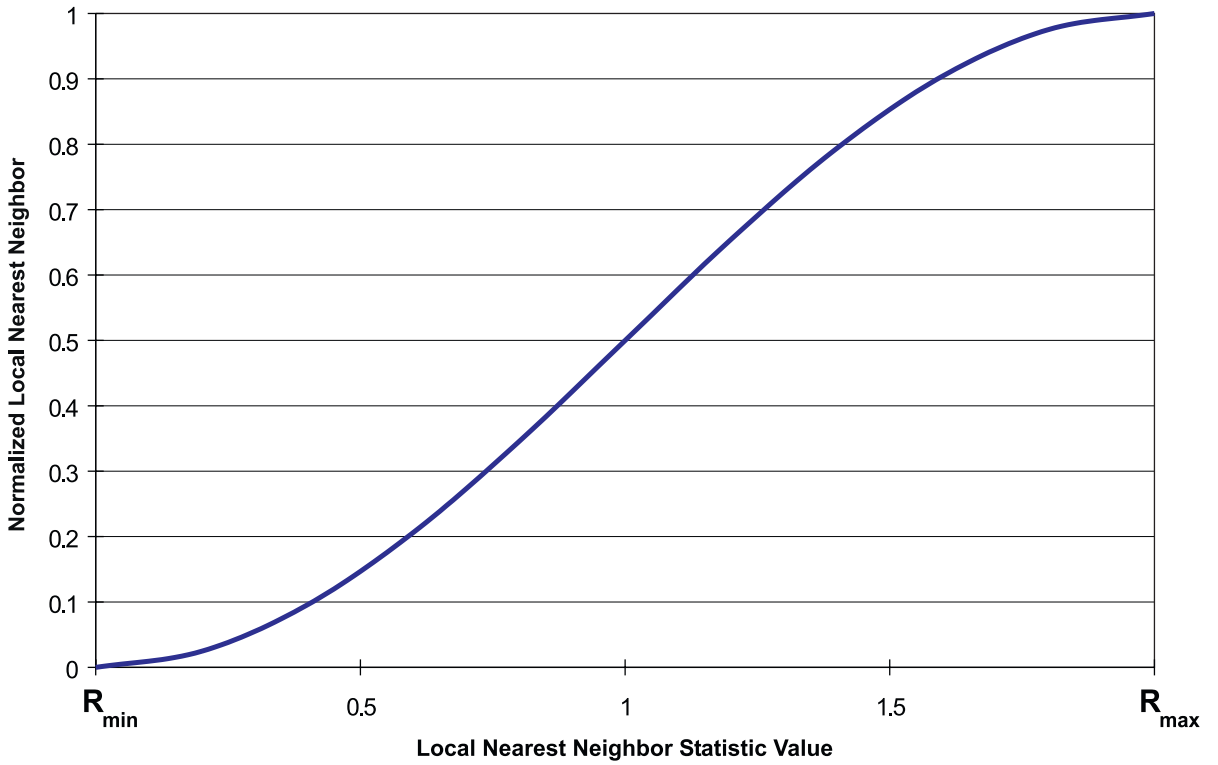

Fig. 2. Normalization function.

3.2. Membership function determination

Theoretically, $R(S_0)$ reflects the relative dispersion or clustering of points within the neighborhood of S_0. However, we would like to normalize the $R(S_0)$ measure such that it is bounded by 0 and 1. With this normalized measure for local point patterns, we can assign a desirable α value. A common practice to normalize the $R(S_0)$ values is to use its empirical minimum and maximum values for the entire region to scale $R(S_0)$ in a linear fashion. But in this demonstration, we chose a fuzzy membership function to normalize the values of $R(S_0)$ because it can provide more flexibility than the standard linear normalization (Kantardzic, 2003). The normalized local nearest neighbor statistic, μ_R, can be depicted in the following manner:

$$\mu_R = \begin{cases} 0, & R(S_0) < R_{\min} \\ 0.5 + 0.5\ \sin(\pi/R_{\max}) & \\ \quad (R(S_0) - R_{\min}), & R_{\min} \leqslant R(S_0) \leqslant R_{\max} \\ 1, & R_{\max} < R(S_0) \end{cases} \quad (9)$$

where $R_{\min}$ refers to a local nearest neighbor statistic value below which μ_R will be assigned a value of zero and $R_{\max}$ refers to a local nearest neighbor statistic value above which μ_R will be assigned a value of one. Empirically, $R_{\min}$ and $R_{\max}$ can be set to the empirical minimum and maximum, respectively, of $R(S_0)$ for all locations of S_0, or any other values. This framework provides a greater flexibility to determine what is a dispersed or a clustered pattern. A diagram of the membership function μ_R verses $R(S_0)$ is shown in Fig. 2.

3.3. An adaptive distance-decay parameter algorithm

The main idea of the AIDW method is to develop an algorithm to select the weights for the distance-decay function for different unsampled locations so that using these weights can increase the accuracy of spatial prediction. That is, for clustered situations, the unsampled points shall be interpolated using averages of neighboring values as the clustered points should capture similar information (values) around the neighborhood. For dispersed situations, larger weights should be used to interpolate values in unsampled points such that values for closer points are weighted more. Therefore, smaller α values should be used in clustering patterns when the averaged nearest neighbor distances are relatively small. Larger α values should be used in dispersed patterns when the averaged nearest neighbor distances are relatively large. Therefore the general rule is to assign smaller distance-decay values to smaller μ_R and larger distance-decay values to larger μ_R.

Given the normalized local nearest neighbor statistic (μ_R) for an unsampled location, we have to determine the appropriate distance-decay parameter α. As the μ_R values are normalized, we may map the μ_R values to a range of α. There are many existing methods to support such a mapping. The simplest is a linear transformation if one can decide the minimum and the maximum values of

α that are appropriate for the phenomenon under investigation. Or a non-linear mapping can be employed. But most of these mapping methods are rather deterministic, assuming a one-to-one relationship between a given μ_R and α. In order to accommodate some degree of uncertainty in determining the locations of sampled points, omitting sampled locations or other unaccounted factors of uncertainty, we adopt a triangular membership function for this mapping process (Kantardzic, 2003).

The triangular membership function has been applied in many research areas, including estimating rainfall and runoff (Akkurt et al., 2004; Tayfur, et al., 2003). It is based upon the fuzzy set concept to map an input space to an output space, and the primary mechanism for doing this is a list of if–then statements or rules. Rules for non-fuzzy logic are zero-sum rules represented in the following manner:

$$\mu_A(x) = \begin{cases} 1 & \text{iff } x \in A \\ 0 & \text{iff } x \notin A \end{cases} \quad (10)$$

In fuzzy logic, it allows a membership value to be between zero and one.

$$\mu_A(x) = \begin{cases} 1 & \text{iff } x \in A \\ 0 & \text{iff } x \notin A \\ p,\ 0<p<1 & \text{if } x \text{ partially belongs to } A \end{cases} \quad (11)$$

Instead of mapping each normalized local nearest neighbor statistic (μ_R) to a unique distance-decay value (α) as in the case of non-fuzzy logic, the neighbor statistic carries a fuzzy membership that belongs to certain levels or categories of distance-decay value. The resultant distance-decay value is a combination of any two levels or categories of distance-decay values assigned by the normalized local nearest neighbor statistic (μ_R). In the extreme cases of dispersing and clustering (μ_R equals one and zero), a large value and zero will be assigned to α, respectively. This concept is illustrated in Fig. 3.

In this triangular membership function, the x-axis is the value of μ_R obtained from normalizing the observed local nearest neighbor statistic. Five levels or categories are used (I, II, III, IV, and V as in Fig. 3) in this illustration. For each category, we assign a distance-decay value. Given a value for μ_R, we can use the triangular membership function in Fig. 3 to determine the appropriate distance-decay parameter. The determination is based upon linear interpolation. For example, if $R(S_0)$ for an unsampled location is 0.8, the corresponding μ_R will be 0.35 (based upon the normalization function in Fig. 2 or Eq. (9)). According to the triangular membership function in Fig. 3, a μ_R of 0.35 corresponds to two points in the membership degree (0.3 for category III and 0.7 for category II). These membership degrees are used as the weights applied to the corresponding α values (0.5 for category II and 1 for category III) assigned to the categories to derive the final α value. As a result, the final α will be $(0.7 \times 0.5 + 0.3 \times 1) = 0.65$. Note that the sum of the two membership degrees for any μ_R value is always 1.

The above explanation is based upon one of many possible settings of the triangular membership function. In other applications, three categories may be used with

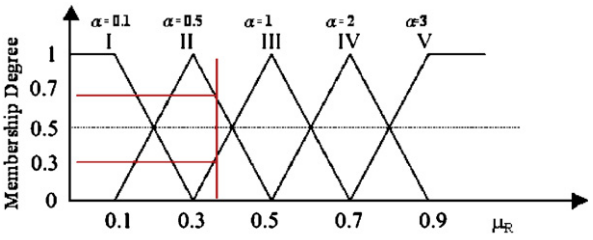

Fig. 3. Triangular membership function for different degrees of the adaptive distance-decay parameter.

labels as low, medium, and high, or any meaningful categories. Also, different sets of α values can be assigned to the categories or cases according to the phenomena under investigation.

In the above illustration, a set of α values was assigned to five cases, covering the distance-decay parameter values from 0.1 to 3, with the distribution of the values as described in Fig. 3. One can argue that the set of α values is only one of many possible sets, and it may not cover the range of α pertinent to the phenomenon under investigation and yield the best results. In our experiment here, we have constructed five distributions of α values, trying to cover the reasonable range of α with different degrees of coverage over the range (Fig. 4). On the other hand, an analyst may construct additional distributions to cover other possible situations to ensure that better or more optimal α values can be identified.

4. Application examples

To demonstrate the utilities of the AIDW method, we applied it in two examples. In the first example, we will use rain gauge data for the central portion of Taiwan to estimate a precipitation surface. In the second example, we will use selected data points from a digital elevation model (DEM) for a Virginia county to estimate an elevation surface.

4.1. Rain gauge data and precipitation

In Earth science and physical geography, precipitation is often needed as an input to various environmental models. However, precipitation data are mostly available in the form of rain gauge data that are treated as sampling points. But for many environmental models, precipitation inputs are required to cover the entire study area in a spatially continuous manner. Therefore, it is desirable to generate a precipitation surface to serve such purposes. In this example, we obtained data for the central region of Taiwan collected at 35 rain gauges (Fig. 5). This region of Taiwan has a relatively rugged terrain in the east. In this illustration using the AIDW method, we randomly selected only 15 gauges for the spatial interpolation process and reserved the remaining 20 points for validation.

First, we examine the spatial correlation structure of the data from the 15 gauges. A variogram was constructed (Fig. 6). The variogram indicates that the

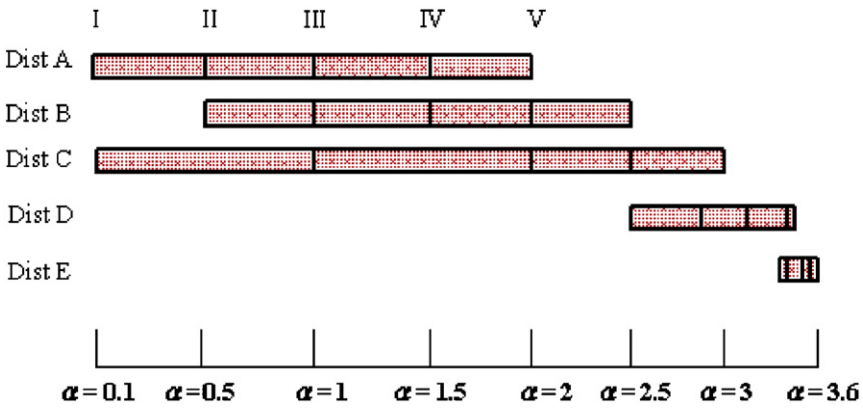

Fig. 4. Five pre-constructed distributions for α search.

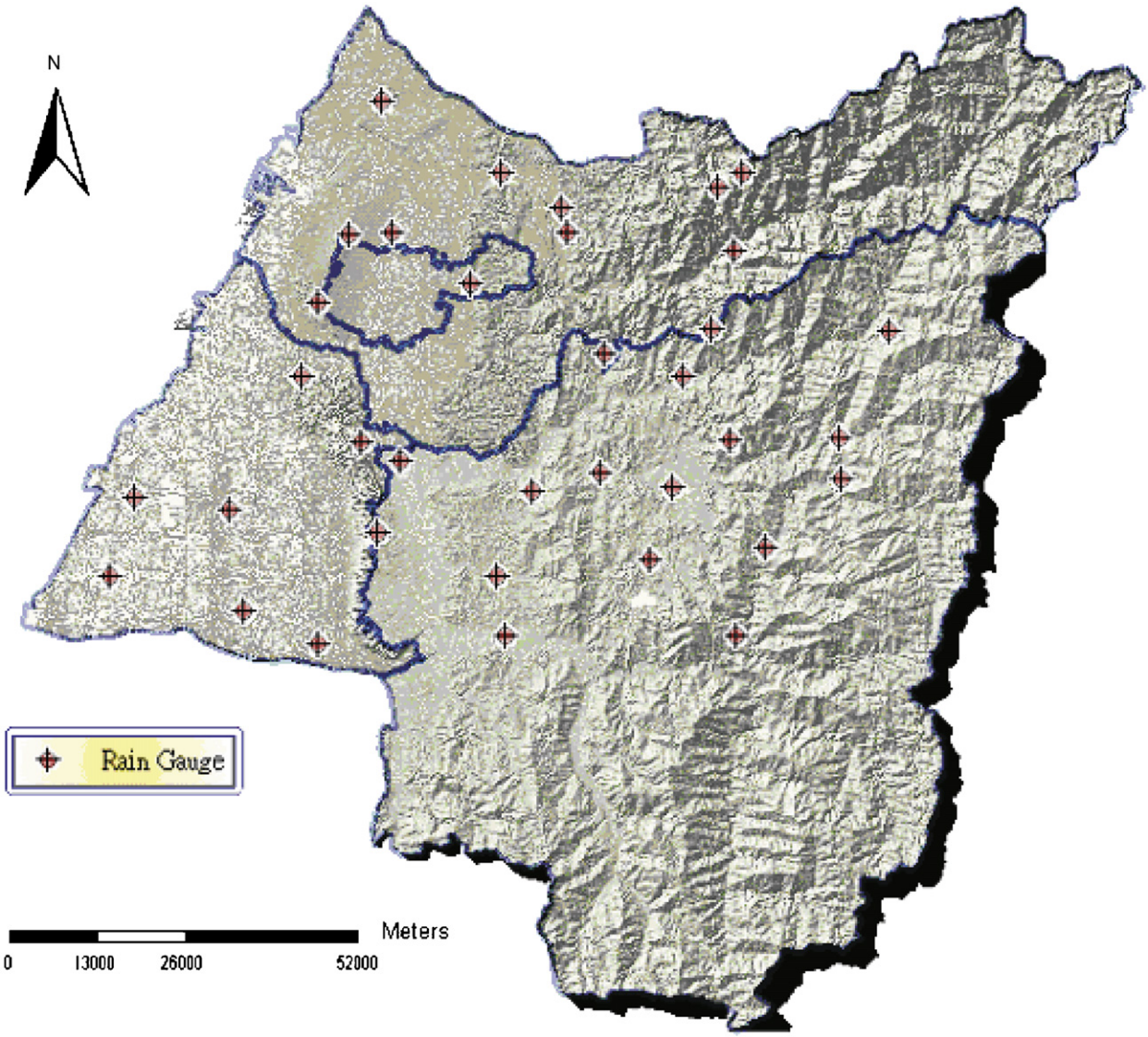

Fig. 5. Distribution of 35 rain gauges in the central portion of Taiwan.

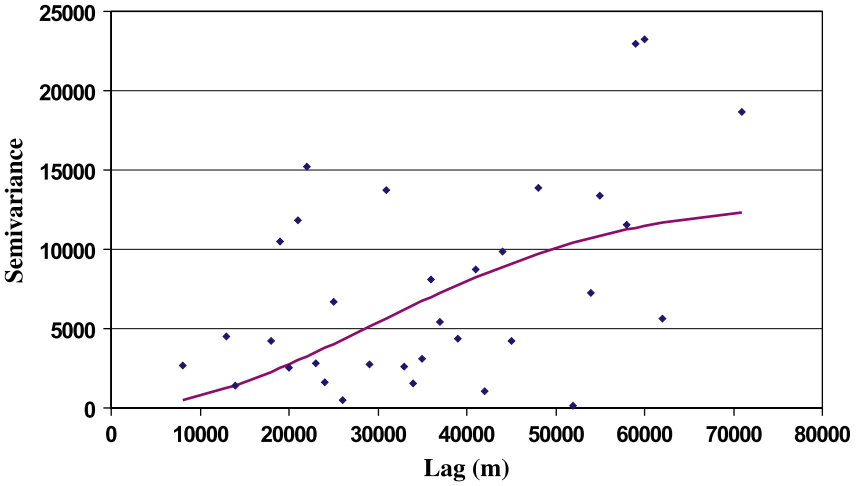

Fig. 6. Variogram of data for 15 rain gauges.

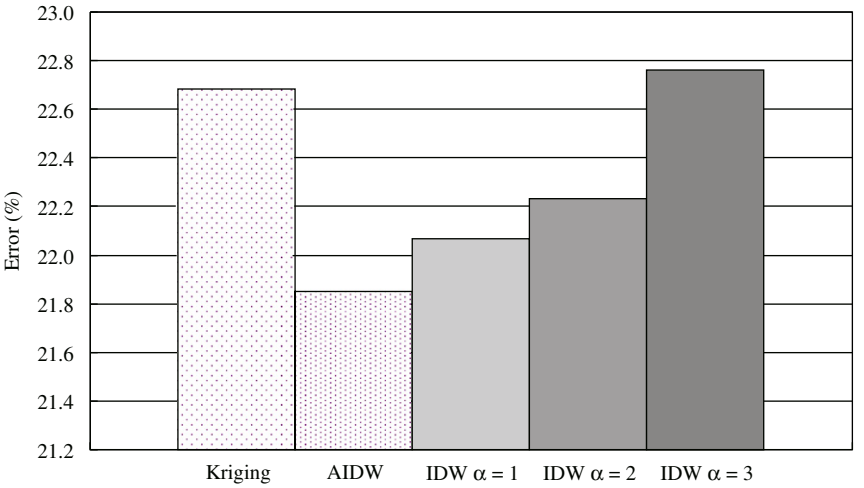

Fig. 7. Interpolation errors with different interpolation methods for precipitation example.

spatial correlation of precipitation quantities among these gauges is not very strong. The theoretical variogram selected was a Gaussian distribution with the following specification: $\gamma(h) = 1 - \exp(-3 \times (h/7100)^2) \times 13,000$ with the goodness of fit (R^2) of 0.184. A general trend of increasing variation with increasing distance is observed. It is true that the sample size of 15 may be too small to develop a reasonably stable empirical variogram, and ideally 50–100 sample points are needed (Burrough and McDonnell, 1998). But in reality, we are constrained by limited data such that we may not be able to extract the spatial correlation structure from the empirical data even if there is one. In this case, there were only 15 gauges in the region and the data are the best that we can obtain. Nevertheless, we can proceed with ordinary kriging, as many empirical studies have used kriging despite the absence of a strong spatial correlation structure in the data (e.g., Mulholland et al., 1998). Then cross validation is performed to evaluate the kriging results. Cross validation is also used to evaluate the results for both IDW and AIDW to assess the bias in the prediction results. The cross validation method is based on percent error or PE (%), which is defined as

$$\text{PE}\,(\%) = \frac{\text{RMSE}}{(1/N)\sum_{i=1}^{N} P_i^*} \times 100\,(\%) \tag{12}$$

where

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - P_i^*)^2} \tag{13}$$

and RMSE (root mean square error) is the mean of the squared difference between the observed value (P_i^*) and the predicted value (P_i), and N is the number of observations.

Maybe partly because of small sample size, a reasonably good theoretical variogram was difficult to obtain. An alternative method to estimate the precipitation amount in unsampled locations is the IDW method. In this case, a constant α is used to interpolate the values at all

Table 1
Different α values, average α, and percent error

Case	Distribution of α	Average α	Error (%)
A	$\alpha = 0.1$ for I $\alpha = 0.5$ for II $\alpha = 1.0$ for III $\alpha = 1.5$ for IV $\alpha = 2.0$ for V	0.765	21.8485
B	$\alpha = 0.5$ for I $\alpha = 1.0$ for II $\alpha = 1.5$ for III $\alpha = 2.0$ for IV $\alpha = 2.5$ for V	1.228	21.9254
C	$\alpha = 1.0$ for I $\alpha = 1.5$ for II $\alpha = 2.0$ for III $\alpha = 2.5$ for IV $\alpha = 3.0$ for V	1.728	22.1947
D	$\alpha = 2.5$ for I $\alpha = 2.8$ for II $\alpha = 3.1$ for III $\alpha = 3.3$ for IV $\alpha = \alpha = 3.5$ for V	2.863	22.7003
E	$\alpha = 3.2$ for I $\alpha = 3.3$ for II $\alpha = 3.4$ for III $\alpha = 3.5$ for IV $\alpha = 3.6$ for V	3.295	23.2358

unsampled locations, but multiple interpolated surfaces were generated using a range of α values (1, 2, and 3). The percentages of total errors using kriging and the IDW method for different α values are shown in Fig. 7. In this specific example, kriging offered less accurate predictions. Based upon Fig. 7, the performance of IDW is somewhat dependent upon the distance-decay value, α. IDW with larger α values (3) perform worse than kriging, while those with smaller α values are better. However, the AIDW provided more accurate predictions than kriging and IDW for all selected α values. The detailed results from the AIDW method are reported in Table 1. The five distributions of α that guide the search for "optimal" α are described in Fig. 4. Table 1 reports the average α value in each case, and the corresponding error. Distribution 1 was the best search scenario because it offers the lowest error (21.84%), which is also shown in the bar graph for AIDW in Fig. 7.

One may suspect that AIDW offers better predictive values than kriging in this case was a coincidence. Therefore, we performed a Monte Carlo simulation experiment 5000 times for all prediction methods concerned in this paper. The results are reported in Table 2. Though the results are slightly different from the previous results reported in Fig. 7, they do support the conclusion that AIDW performs better than others, including kriging in this situation. Somewhat surprising from the simulation results is that even IDW with different distance-decay parameters performed slightly better than kriging in this specific case study.

Table 2
RMS and error (%) from Monte Carlo simulation (5000 runs) for different methods

	Kriging	AIDW	IDW_1	IDW_2	IDW_3
RMS	71.98	69.05	71.63	71.51	71.80
Error (%)	22.51	21.59	22.40	22.36	22.45

4.2. Interpolation of elevation

In this second example, a raster DEM for Loudoun County, VA, was used (Fig. 8). The original contour data with relatively high vertical resolutions (5 feet) were developed by the County government. The contour data were then rasterized to create the DEM for this project. A total of 256 random points, as shown in Fig. 8, were selected from an area that has a moderate terrain. The area is not too rugged in general. We treated the 256 points as our sampled locations. Using all these sampled locations, we generated the elevation surface using different methods, including the AIDW method proposed here. One may be concerned that the DEM was created by interpolating contour lines and thus the DEM was a product of the interpolation method. However, only values in locations between contours were interpolated. The spatial structure of the DEM is still constrained by the original contours. Also, given the purpose of using this example is to demonstrate the utilities of AIDW, the underlying process in creating the spatial structure should not be focused.

As in the previous example, we first examine the spatial correlation structure of the data. The variogram in Fig. 9 depicts the spatial structure of these 256 points. It shows that the correlation structure is reasonably strong with increasing variance when spatial lag increases. The theoretical Gaussian variogram was chosen with the following specification: $\gamma(h) = 1 - \exp(-3 \times (h/13,100)^2) \times 5000$ and with a goodness of fit (R^2) of 0.994. We proceeded with ordinary kriging to generate an elevation surface. As we did before, we also used the standard IDW method with a range of α values and the AIDW method with different α distributions. We performed the cross validation again to compare the predicted with the original values. The results are recorded in Fig. 10 showing that kriging outperforms both the IDW and AIDW by substantial margins. This result should not be too surprising as the variogram for the elevation data shows a strong sign of spatial autocorrelation and thus kriging should be a more appropriate method. But as we compare the results from IDW and AIDW (Fig. 11), regardless of the α value used for IDW, all AIDW results with different distributions of α outperform IDW. As shown in Fig. 11, the best results were obtained from distribution C with α values equal to 0.1, 1.0, 2.0, 2.5, and 3.0, with a percent error of 2.1879.

5. Conclusion and discussion

Since a focus of recent research in spatial interpolation is to increase the prediction accuracy (Atkinson, 2005),

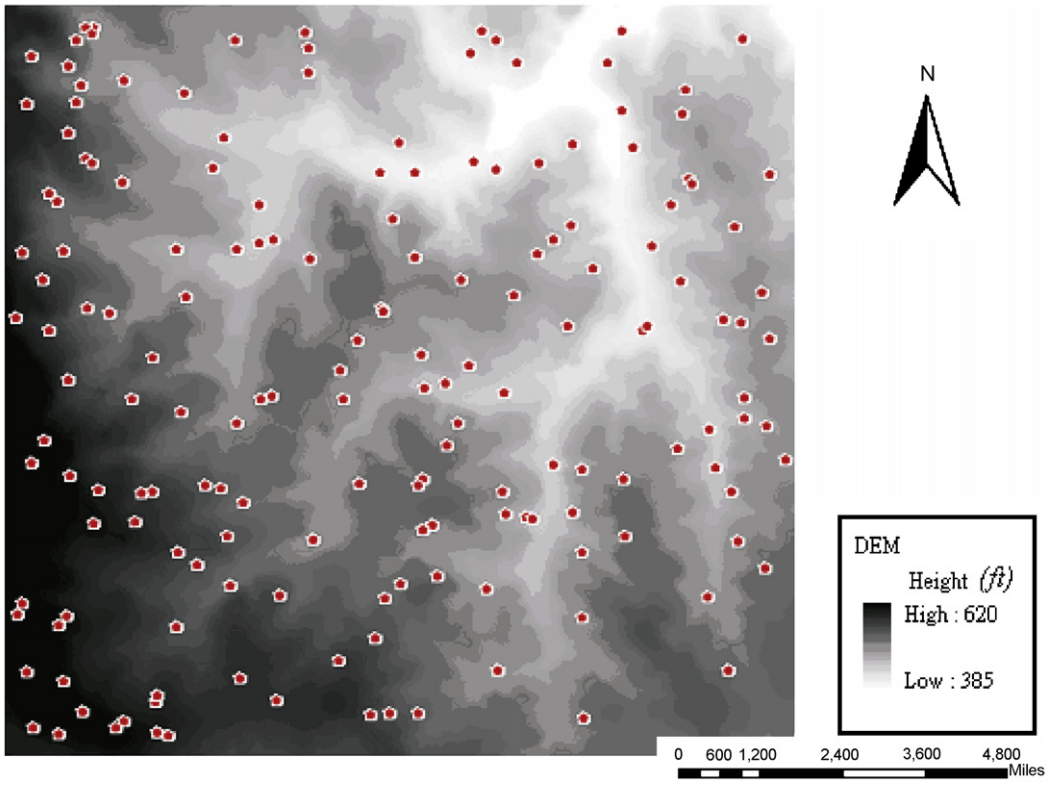

Fig. 8. Distribution of 256 random points selected from Loudoun County, Virginia.

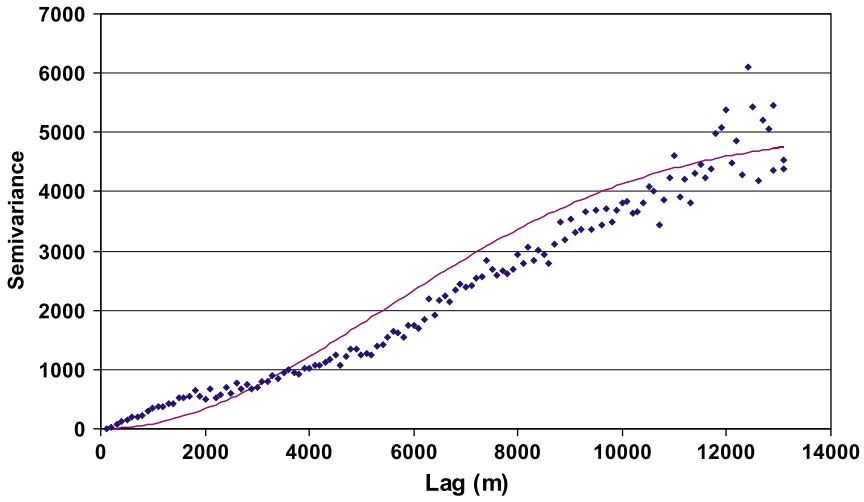

Fig. 9. Variogram of 256 random points.

the main objective of this paper is to develop an enhanced version of the IDW method to increase the prediction power. We introduced a modified IDW technique that allows the distance-decay parameter to be adjusted according to the density characteristics of the sampled point distribution. By allowing the parameter to vary by locations according to the density of local sampling points, it offers greater flexibility to model the possibly heterogeneous distance-decay relationship over space. The discussion in this paper uses exclusively the power function in specifying the distance-decay relationship (Eq. (2)), but other functional forms, such as the exponential function, are clearly applicable.

Results from the two application examples provide some insights in terms of the strengths and weaknesses, and the applicability of the AIDW method. We demonstrated that in

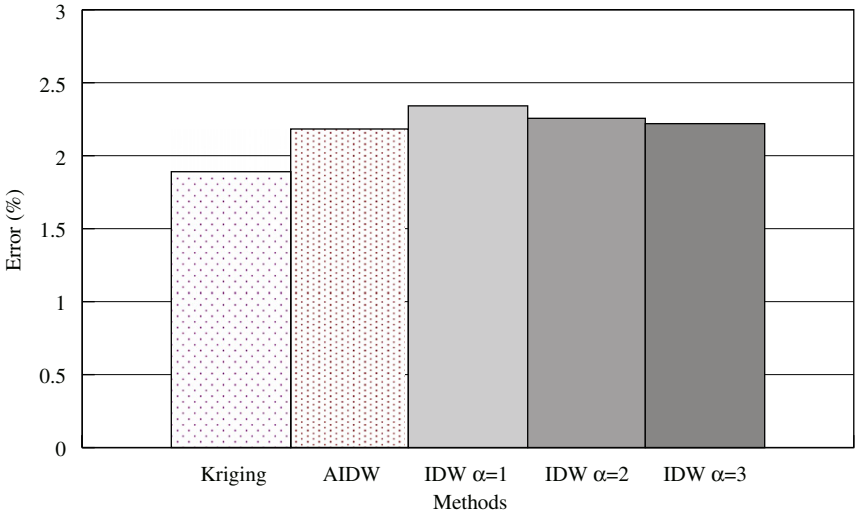

Fig. 10. Interpolation errors with different interpolation methods, the case of elevation.

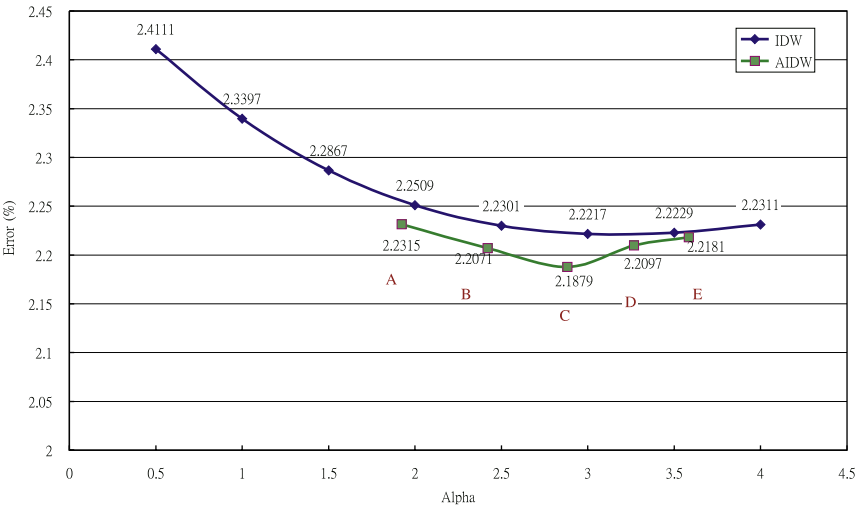

Fig. 11. Percent error with different α values for IDW and AIDW methods with different α distributions.

the interpolation of rain gauge data, if spatial correlation structure is not prominent or the data are not adequate to support kriging, results from kriging may provide less accurate predictions than other methods, such as the AIDW. When traditional IDW can provide reasonable prediction results, AIDW should perform better. On the other hand, kriging will do quite well with less spatially heterogeneous phenomena and when the spatial correlation structure is strong, and when the variogram can depict the spatial structure reasonably well. In these situations, using AIDW will not improve the prediction results substantially. In other words, AIDW is not appropriate in all cases, especially when the spatial correlation structure is strong and the variogram can model the spatial relationship effectively.

As compared to IDW, the proposed AIDW method involves extra effort and a significant level of heuristics. The analyst has to formulate the distribution of the distance-decay parameter α for the searching process to identify the "optimal" set of parameter values. One may have to narrow the search in an iterative manner to get closer to the optimal values. An obvious alternative is to automate the search across the range of α values, but this computational intensive process may increase the predictive power only marginally so that the extra effort may not be justifiable.

As stated earlier that kriging and IDW adopt different approaches, the intention of proposing the AIDW method is by no means to replace kriging. AIDW will likely offer a better predictive power than the IDW method. But when the spatial correlation structure of the data is not strong, or when the data are too limited to support kriging, as in the case of the precipitation data for Taiwan, IDW is a logical alternative to kriging, but AIDW offers a better alternative.

References

Akkurt, S.G., Tayfur, G., Can, S., 2004. Fuzzy logic model for the prediction of cement compressive strength. Cement and Concrete Research 34, 1429–1433.

Atkinson, P.M., 2005. Spatial prediction and surface modeling. Geographical Analysis 37, 113–123.

Bailey, T.C., Gatrell, A.C., 1995. Interactive Spatial Data Analysis. Longman, Essex, England, 413pp.

Barnes, S.L., 1964. A technique for maximizing details in numerical weather map analysis. Journal of Applied Meteorology 3, 396–409.

Bekele, A., Downer, R.G., Wolcott, M.C., Hudnall, W.H., Moore, S.H., 2003. Comparative evaluation of spatial prediction methods in a field experiment for mapping soil potassium. Soil Science 168 (1), 15–28.

Boots, B.N., Getis, A., 1988. Point Pattern Analysis. Sage Publications, Newbury Park, CA, 93pp.

Bracken, I., Martin, D., 1989. The generation of spatial population distributions from census centroid data. Environment and Planning A 21, 537–543.

Burrough, P.A., McDonnell, R.A., 1998. Principles of Geographical Information Systems. Oxford University Press, Oxford, 333pp.

Cressman, G.D., 1959. An operational objective analysis system. Monthly Weather Review 87, 367.

Fotheringham, A.S., 1997. Trends in quantitative methods I: stressing the local. Progress in Human Geography 21, 58–96.

Fotheringham, A.S., O'Kelly, M.E., 1989. Spatial Interaction Models: Formulations and Applications. Kluwer, Boston, 224pp.

Gandin, L.S., 1970. The planning of meteorological station networks. Technical Note, vol. 111. World Meteorological Organization, Geneva, WMO no. 265, 35pp.

Getis, A., Ord, J.K., 1992. The analysis of spatial association by use of distance statistics. Geographical Analysis 24, 189–206.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York, 483pp.

Goovaerts, P., 2000. Geostatistical approaches for incorporating elevation into the spatial interpolation of rainfall. Journal of Hydrology 228, 113–129.

Henley, S., 1981. Nonparametric Geostatistics. Wiley, New York, 145pp.

Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford University Press, Oxford, 561pp.

Kantardzic, M., 2003. Data Mining: Concepts, Models, Methods, and Algorithms. Wiley, Hoboken, NJ, 345pp.

Lloyd, C.D., 2005. Assessing the effect of integrating elevation data into the estimation of monthly precipitation in Great Britain. Journal of Hydrology 308, 128–150.

Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W., 2001. Geographic Information Systems and Science. Wiley, Chichester, NY, 454pp.

Mulholland, J.A., Butler, A.J., Wilkinson, J.G., Russell, A.G., Tolbert, P.E., 1998. Temporal and spatial distributions of ozone in Atlanta: regulatory and epidemiologic implications. Journal of the Air and Waste Management Association 48, 418–426.

Ping, J.L., Green, C.J., Zartman, R.E., Bronson, K.F., 2004. Exploring spatial dependence of cotton yield using global and local autocorrelation statistics. Field Crop Research 89 (2–3), 219–236.

Rivoirard, J., 1994. Introduction to Disjunctive Kriging and Non-Linear Geostatistics. Oxford University Press, New York, 180pp.

Sampson, P.D., Damian, D., Guttorp, P., 2001. Advances in modeling and inference for environmental processes with nonstationary spatial covariance. In: Monestiez, P., Allard, D., Froidevaux, R. (Eds.), GeoENV 2000: Geostatistics for Environmental Applications. Kluwer, Dordrecht, pp. 17–32.

Sen, Z., Sahin, A.D., 2001. Spatial interpolation and estimation of solar irradiation by cumulative semivariograms. Solar Energy 71 (1), 11–21.

Tayfur, G., Ozdemir, S., Singh, V.P., 2003. Fuzzy logic algorithm for runoff-induced sediment transport from bare soil surfaces. Advances in Water Resources 26, 1249–1256.

Tobler, W., 1970. A computer movie simulating urban growth in the Detroit region. Economic Geography 46 (2), 234–240.

Tomczak, M., 1998. Spatial interpolation and its uncertainty using automated anisotropic inverse distance weighting (IDW)-cross-validation/jackknife approach. Journal of Geographic Information and Decision Analysis 2 (2), 18–30.

Weber, D., Englund, W., 1992. Evaluation and comparison of spatial interpolators. Mathematical Geology 24, 381–391.

Zimmerman, D., Pavlik, C., Ruggles, A., Armstrong, M.P., 1999. An experimental comparison of ordinary and universal kriging and inverse distance weighting. Mathematical Geology 31 (4), 375–390.