Inverse Distance Weighting Interpolation

The Inverse Distance Weighting (IDW) algorithm is a moving average interpolator that is usually applied to highly variable data. For certain data types it is possible to return to the collection site and record a new value that is statistically different from the original reading but within the general trend for the area. Examples of this type of data include soil chemistry results, environmental monitoring data, and consumer behaviour observations. It is not desirable to honour local high/low values but rather to look at a moving average of nearby data points and estimate the local trends.
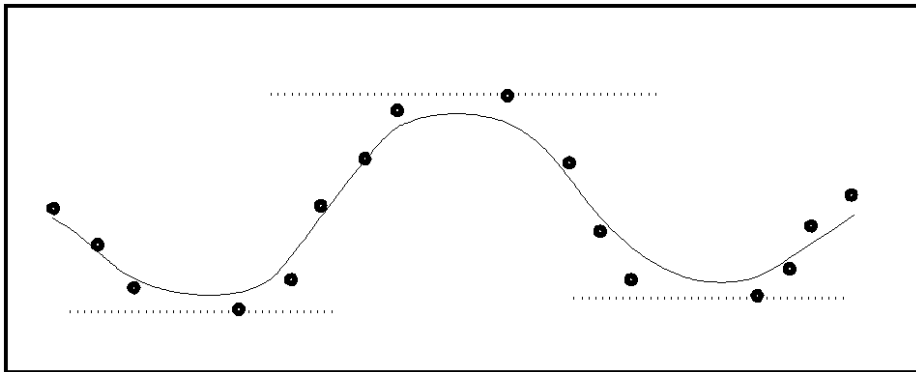

Figure 3.3. The interpolated surface, estimated using a moving average technique, is less than the local maximum value and greater than the local minimum value.

The IDW technique calculates a value for each grid node by examining surrounding data points that lie within a user-defined search radius. Some or all of the data points can be used in the interpolation process. The node value is calculated by averaging the weighted sum of all the points. Data points that lie progressively farther from the node influence the computed value far less than those lying closer to the node (Figure 3.4). For an instructional lesson on IDW interpolation refer to Lesson 3 of the *Installation and Tutorials* manual.
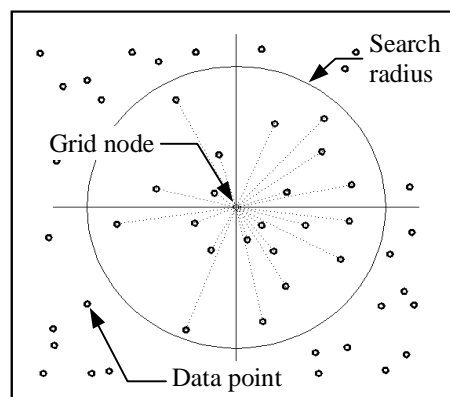

Figure 3.4 A radius is generated around each grid node from which data points are selected to be used in the calculation.

The Inverse Distance Weighted Interpolation Dialogue

As discussed, the IDW interpolation method is initiated from the first screen of the Interpolation

Wizard. After entering the appropriate table and column names in the *Select Table and Column* dialogue box, the *Next* button takes the user into the **Inverse Distance Weighted Interpolation** dialogue box.
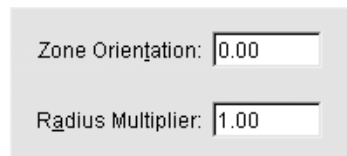
- The *Inverse Distance Weighted Interpolation* dialogue box allows a number of user-defined parameters to be set in the gridding algorithm.
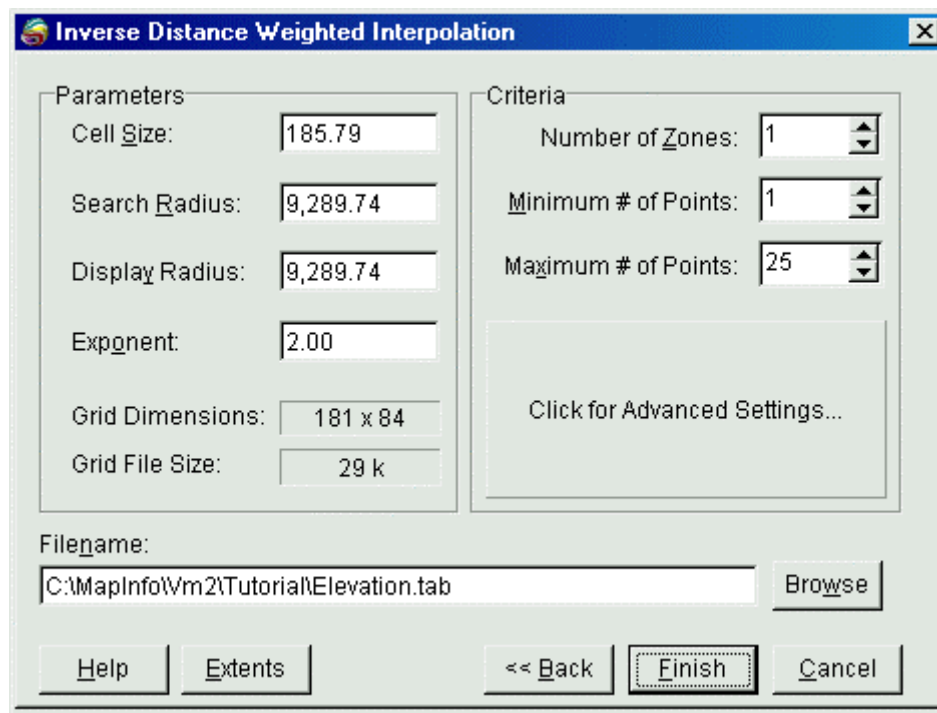
1 **Cell Size** is defined in map units. Note that the grid dimension (in cell units) varies inversely with cell size: the smaller the cell, the larger the grid file. The value chosen should be a compromise between the degree of resolution required for analysis and visualization purposes and the processing time and file size factors. The default value is calculated by dividing the diagonal extent of the point file by 200 which is considered an optimum number based on the computing power required to solve this slightly more complex algorithm.

2 **Search Radius** defines the maximum size, in map units, of a circular zone centred around each grid node within which point values from the original data set are averaged and weighted according to their distance from the node. The minimum and maximum number of data points averaged within each zone is also defined by the user (see *Search Criteria* settings below). The default setting is calculated as a percentage of the total extent of the map area and is appropriate for most data sets.

1.

3 **Display Radius** defines the size, in map units, of a circular zone centred around each grid node within which an original data point must lie in order for that grid cell to be assigned a value and be displayed. This setting is used to control the extent of grid creation in areas of the map that do not contain data, such as the outer margin of the map area. The default setting that appears is the same as the search radius which is appropriate for most data sets. However, if the original data points are characterized by very patchy distribution with occasional large gaps containing no sample sites, a smaller display radius should be

considered in order to avoid interpolation of grid values within areas that do not contain data.

4   *Exponent* is the variable that defines the exponential rate of decay of influence of neighbouring points the farther they lie from the grid node. Increasing the exponent will decrease the relative influence of more distant neighbours. Values range between one and ten with a value of two being appropriate for most data sets.

5   The search *Criteria* relate to a number of settings that define the manner in which data points are averaged for each estimation calculation. The *Number of Zones* defines the number of equally sized partitions by which each search area may be divided. This may be useful if the data set contains points that are clustered along a particular orientation such as data collected along widely separated survey lines. An equal number of points are averaged from each half, quadrant, or octant of the search area. It must be noted that in the majority of point files the most appropriate number of zones is **one**. Choosing more than one should be done only if the user is confident that the data is obviously clustered in a linear orientation. The *Minimum # of Points* per zone and the *Maximum # of Points* per zone can also be set. This refers to the minimum and maximum number of data points searched for and averaged to calculate a single grid node value.

-

- The *Click for Advanced Settings* section reveals more specialized settings to be made in the weighted average grid calculation.

Zone Orientation: 0.00

Radius Multiplier: 1.00

1   *Zone Orientation*, in azimuth degrees, allows the user to define unique orientation parameters in the creation of search area zones which, again, is appropriate only if data points are preferentially clustered in linear zones.

2   *Radius Multiplier* is a weighting factor that defines the degree of influence of neighbouring points that lie successively farther from the grid node. A radius multiplier of 1 means a point located at the edge of the search radius has a weight of zero in the averaging calculation. A radius multiplier of 2 has the effect of doubling the distance used for weighting the points.

Tip: The default values for the *Search Criteria* and *Advanced* options are appropriate for a majority of point data sets. However, it is suggested that the user experiment with various settings in order to evaluate which settings generate the most appropriate representation of their particular data sets.

1   The *Extents* button opens an information box that summarizes the geographic size and the Z-value range of the original point data base, the density of the points, as well as the data value units.

- The user may either select the *Finish* button to complete the gridding process or, if modifications to the previous dialogue are required, select the << *Back* button to return to one or more dialogues back. Once the grid is created, it appears in a MapInfo Map window with a default colour palette applied. The user may change the colour range assigned to the grid file using commands discussed in Chapter 7.