A Short Note on Trend Modeling using Moving Windows

John G. Manchuk and Clayton V. Deutsch

In some geostatistical modeling applications, such as for properties deposited under a known physical or genetic phenomenon or for categorical variables, the use of trend models may be useful. Trends are incorporated into geostatistics through methods such as universal kriging, kriging with a locally varying mean, or using the trend to compute residuals. Several methods exist to model explicit trend models including kriging, regression, moving windows, and radial basis functions. This paper explores the use of moving window averaging for computing vertical, aerial and three dimensional trends. A flexible program is developed to allow users to explore moving windows for trend modeling. Flexibility in window geometry, weighting schemes, and filtering is provided along with plotting capabilities for vertical trends. It is also possible to calibrate three dimensional trends to reproduce a vertical trend, since the later is often more informed than the former.

1. Introduction

Inferring and using trends in Geostatistical modeling is a subjective area. Practitioners are required to make three decisions, those being if using a trend is necessary, how the trend will be inferred from the available data, and how the trend will be incorporated into the models. Trends may be necessary when the physics of a spatial distribution indicate one, for example, a fining-upwards trend of particle size distribution as larger particles settle prior to finer particles during deposition (Deutsch, 2002). However, data may exhibit a trend even when no physical or genetic explanation presents itself (Chiles and Delfiner, 1999). In these cases, using a trend is not necessarily justified. One case where trends are often used is for categorical variables. Categories tend to display regions of higher and lower proportions. In most cases, using trends for categorical variables is geologically rational.

This paper addresses the second decision, that is, how the trend is inferred from available data. Several possibilities exist for inferring a trend such as linear and non-linear regression, kriging, polynomial interpolation as in universal kriging, also called kriging with a trend (Journel and Rossi, 1989), calibration from exhaustive secondary information such as seismic and gravimetric surveys, and in this case using moving window averaging (moving windows). Trends modeled using moving windows are deterministic, as opposed to the random variable they are inferred from, which is stochastic. Moving windows is a useful technique when there is an abundance of sample data. In cases with sparse samples, such as a fairly young reservoir with very few wells, it may only be possible to infer a vertical trend. Once a trend model is obtained from moving windows, it can be incorporated into modeling in a few ways including: as a secondary variable in collocated co-kriging or co-kriging with the intrinsic model (Babak and Deutsch, 2009); as a locally varying mean in the simple kriging equations; and as a trend for computing residuals that are used in the kriging equations. Methods of incorporating trends are not discussed.

The method of moving windows for trend modeling is covered in this paper. The discussion is primarily targeting categorical variables; however, the results apply to continuous variables as well. To accompany this, a program for modeling trends using moving windows is developed. The program includes several options for different window shapes, weighting schemes, and filtering.

2. Moving Windows

Moving window averaging is a straightforward method for fitting a function to a set of points. Points represent samples and the function is usually computed on a grid that will be used for geostatistical modeling. The moving window average of each grid cell with centre $\mathbf{u}_k, k = 1, \dots, N$, is evaluated by finding the samples within a specified averaging volume (the window) and computing their average (Figure 1, Equation (1)), where $z(\mathbf{u}_i)$ is a sample at location $\mathbf{u}_i$, $\bar{z}(\mathbf{u}_k)$ is the average, w_i are weights, and n is the number of samples inside the window. There is no restriction on the geometry of the window, the averaging technique, or the dimensionality of the problem. The averaging method shown is a weighted arithmetic average and is used for several variables such as porosity and categorical proportions.

$$\bar{z}(\mathbf{u}_k) = \frac{\sum_{i=1}^{n} w_i z(\mathbf{u}_i)}{\sum_{i=1}^{n} w_i} \tag{1}$$

In the case of a categorical variable with m categories, the average results in a vector of proportions that sum to one, Equation (2). $\bar{z}_j(\mathbf{u}_k)$ is the proportion of category j and $\delta(z(\mathbf{u}_i), j)$ is Kronecker's delta often referred to as the indicator transform function in geostatistics defined by Equation (3).

$$\bar{z}_j\left(\mathbf{u}_k\right) = \frac{\sum_{i=1}^{n} w_i \delta\left(z(\mathbf{u}_i), j\right)}{\sum_{i=1}^{n} w_i}, j = 1,\ldots,m \tag{2}$$

$$\delta\left(z(\mathbf{u}_i), j\right) = \begin{cases} 1, & \text{if } z(\mathbf{u}_i) = j \\ 0, & \text{if } z(\mathbf{u}_i) \neq j \end{cases} \tag{3}$$

By construction, Equation (2) always yields an average vector of proportions, $\bar{z}_j(\mathbf{u}_k)$, that sum to one. Letting W represent the sum of the weights in the denominator, the sum of the proportions is given by Equation (4).

$$\sum_{j=1}^{m} \bar{z}_j\left(\mathbf{u}_k\right) = \sum_{j=1}^{m} \frac{1}{W} \sum_{i=1}^{n} w_i \delta\left(z(\mathbf{u}_i), j\right) \tag{4}$$

Rearranging the summations on the right hand side yields Equation (5). By the nature of indicators, the sum over j at any given location is always equal to 1; therefore the sum of the proportion vector is 1. This result holds as long as the same weighting scheme is applied to all indicators.

$$\begin{aligned} \sum_{j=1}^{m} \bar{z}_j\left(\mathbf{u}_k\right) &= \frac{1}{W} \sum_{i=1}^{n} w_i \sum_{j=1}^{m} \delta\left(z(\mathbf{u}_i), j\right) \\ &= \frac{1}{W} \sum_{i=1}^{n} w_i \\ &= 1 \end{aligned} \tag{5}$$

The weighted arithmetic average for moving windows has a few other nice properties for trend modeling. At any location, the computed average will not exceed the bounds of the variable if all the weights are positive. This property leads to safe trend models. For example, where a trend is modeled away from available sample data, there is no chance for extrapolation of extreme values that can occur with other trend modeling techniques such as regression. The bounded property can be defined locally as well: the trend computed at a location, $\mathbf{u}$, will not exceed the bounds of the variable observed inside the window associated with $\mathbf{u}$. Interpolation using other techniques such as splines, Lagrange polynomials, or even kriging can exceed the bounds observed in the local neighborhood.

Another property is that the computed trend is not constrained to reproduce the data. Such a trend would account for a high fraction of the total variance of a random variable and is a case of over-fitting. Using moving windows, the amount of variability in the trend is directly controlled by the window size. Using the smallest window possible, which translates into the single nearest neighbor for each average, the trend reproduces the data exactly and has maximum variance; whereas using an infinitely large window and equal weighting results in a trend that is constant and equal to the mean of the variable. It has zero variance. Altering the window geometry and weighting scheme results in a trend that exists between these extremes of variability; hence, moving window averaging for trend modeling is a flexible technique.

There are some limitations as well. In cases with limited data, it is more difficult to determine if a trend is necessary. Moving windows can result in a trend that imparts structure which does not necessarily exist or cannot be rationalized. Another limitation is extrapolation beyond the bounds of a variable. If it is known that such structure exists, then moving windows is not the appropriate choice.

3. Methodology and Program

Several variants of moving window averaging for trend modeling have been compiled into a program: trends can be modeled vertically, aerially, and in three dimensions with box shaped or ellipsoid shaped windows; windows can be anisotropic and rotated; weighting can be done using equal weights or Gaussian weights; and filtering can be applied to resulting trends. Trend models are built on regular grids defined by GSLIB conventions (Deutsch and Journel, 1998).

Both box and ellipsoid windows are defined by three search radii, $r_u,\ r_v,\ r_w$, and two angles, the azimuth and dip. The search radii define the extents of the windows in local coordinates defined by the angles (Figure 2). Azimuth is measured positive from North or the y-axis while dip is measured positive down from the xy-plane. Mapping between the global coordinate system where the grid is defined to the local coordinate system of the window is done using a rotation matrix defined in Deutsch and Journel (1998). For vertical trends, the angles are meaningless and only r_w is required to define the window, which is an xy slice with thickness r_w. For aerial trends, only the azimuth angle and r_u, r_v are required. They define a window in the xy-plane with infinite vertical extent. All radii and angles are required for building three-dimensional trends. In the trend modeling program, it is possible to use explicitly defined windows where the radii and angles define the geometry used at all grid node locations, as well as implicitly defined windows where the radii and angles define only the anisotropy and rotation, but not the size. Instead, the size is determined by an additional parameter, the number of nearest neighbors, n. For each grid cell, the window always contains n samples and the radii are determined from the furthest sample from the grid cell centre in the local coordinate system. For the explicit window definition, n may be different for every grid cell.

Specifying the averaging windows explicitly or implicitly leads to different trends and the choice is problem dependent. Explicit windows cover the same amount of space for all locations in a grid and may be more appropriate for problems that are more stationary in terms of covariance. Explicit windows are also more appropriate in problems where the sample data are roughly equally spaced. In situations where data is clustered, explicit windows can lead to a trend model that is over-smoothed in densely sampled areas. There is also the potential for defining a window that is too small leading to areas where the trend is undefined. Implicit windows are more appropriate for clustered data. Another advantage of implicit windows is each average has constant support n. In most cases, choosing which method to use is best done by trial and error. The choice may depend on the variance the trend is to account for and by visual inspection.

Two weighting schemes are possible, equal and Gaussian. Applying equal weights to compute the average of the samples within a window simplifies Equation (1) to the arithmetic average in Equation (6).

$$\overline{z}(\mathbf{u}_k) = \frac{1}{n}\sum_{i=1}^{n} z(\mathbf{u}_i) \tag{6}$$

An advantage of using the arithmetic average is that it does not introduce information to the trend that is not necessarily present in the data. With no additional information, the assumption that all data be treated equal is unbiased. However, equal weighting has the disadvantage that it can lead to discontinuous trends, which may not be a desirable result. Discontinuities are caused when an infinitesimal shift in the position of the averaging window excludes a sample previously inside the window or includes an additional sample previously outside the window. A one dimensional example involving samples from the function $z = u^3$ with some additional noise demonstrates the effect. Discontinuities can be mitigated by filtering the result, which is also shown. Filtering is discussed below.

The second weighting scheme calculates weights using a function similar to the Gaussian distribution function, Equation (7), where R_i is the distance from $\mathbf{u}_k$ to the i^{th} datum in the window in the local coordinate system. Using Gaussian weights assumes data that are closer to $\mathbf{u}_k$ are more important that those further away, which is often the case. Note that to simplify calculations the local coordinate system is transformed to remove anisotropy; therefore, $r_u = r_v = r_w = r$ and R_i is the Euclidean distance.

$$w_i = \exp\left(\frac{-9R_i^2}{r^2}\right) \tag{7}$$

If windows are defined implicitly, r is defined as the maximum distance observed among the n points found near $\mathbf{u}_k$. With either implicit or explicit windows, the choice of r ensures a continuous trend even when a small shift in the window leads to an exchange of data. This statement is not exactly correct because at the edge of the window, the derivatives of Equation (7) are not exactly zero, which is a necessary condition to achieve continuity across the boundary of the window; however, the derivatives are extremely small and a high degree of continuity is usually achieved numerically. Using the same example as for the equal weighted scheme, the resulting trend is continuous (Figure 4). Any undesirable local variability present in the trend can be removed with filtering.

After a trend model is built, one or more filters can be applied to accomplish tasks such as removing discontinuities and smoothing local variations. In the program, filters are applied using a template of regular grid cells with equal or Gaussian weighting described previously. A template is defined by grid cell offsets, nx, ny, nz, and is centred at $\mathbf{u}_k$ (Figure 5). For example, filtering a vertical trend with $nz = 5$, the average at $\mathbf{u}_k$ is computed using the column of grid cells ranging from 2 cells below $\mathbf{u}_k$ to 2 cells above. If Gaussian weights are used, they are based on the distance from $\mathbf{u}_k$ to the other grid cell centres in the template.

The trend program is called MAKETREND and is a typically GSLIB-style executable controlled by a text-based parameter file. Parameters are described in Table 1 and explanations follow. Only those parameters that are not clearly explained in the parameter file are discussed. All input and output data files are of GeoEAS format with the first line being a title, second line the number of columns, $ncol$, the next $ncol$ lines are column titles, and the actual space delimited data follows.

Table 1: Default parameter file generated by the MAKETREND program.

Line	Parameters and Description	
1	`START OF PARAMETERS:`	
2	`samples.dat`	`-file with data`
3	`1   2   0`	`-   columns for X, Y, Z coordinates`
4	`2   3   4`	`-   number of variables, column numbers`
5	`-1.0e11     1.0e11`	`-   trimming limits`
6	`0`	`-trend type: 0 - vertical; 1 - aerial; 2 - 3D`
7	`0`	`-type of averaging volume: 0 - sphere, 1 - cube`
8	`100.0 100.0 10.0`	`-search radii`
9	`0.0 0.0`	`-azimuth and dip`
10	`25`	`-nearest neighbors to use (0 for all)`
11	`0`	`-weighting function: 0 - equal, 1 - Gaussian`
12	`0 3`	`-apply post processing filter: 1 - yes, iterations`
13	`5 5 3`	`-filter size: nx, ny, nz`
14	`0`	`-filter weighting function: 0 - equal, 1 - Gaussian`
15	`1`	`-ensure sum to 1 for compositional trends: 1 - yes`
16	`trend.out`	`-file for trend output`
17	`100  0.5  1.0`	`-nx, xmin, xsize`
18	`100  0.5  1.0`	`-ny, ymin, ysize`
19	` 10  0.1  0.2`	`-nz, zmin, zsize`
20	`1`	`-if vertical trend, generate a plot, 1 - yes`
21	`vertical.ps`	`-file for postscript output`
22	`1`	`-cumulative curves, 1 - yes`
23	`0. 1.`	`-min and max values for x axis`
24	`50`	`-scaling factor for x axis`
25	`1 2`	`-colors for curves, see codes below`

The first line is required for the program to find the start of the parameters. Trend type is fairly self explanatory; however, there is some flexibility not specified. Most geostatistical modeling applications are three dimensional and to incorporate a vertical trend, it is useful to replicate the trend to cover the whole three dimensional model. For a vertical trend, if nx and/or ny are greater than 1, the resulting trend model is repeated for all nx, ny. For aerial trends, if nz is greater than 1, the trend model is repeated for all nz. Of the two averaging volume types available on Line 7, the spatial search for cubes has not been optimized when the nearest neighbors on Line 10 is fixed and performs slower than the sphere type. When the number of nearest neighbors is unconstrained (parameter set to 0), the cube search performs faster. The search radii and angles on Lines 8 and 9 that are used depends on the trend type. For vertical trends, only the third search radius is used and the angles have no effect on results. Aerial trends only depend on the first two radii and the azimuth angle. Three dimensional trends use all search radii and angles.

Earlier, it was proven than for compositional variables the moving window averaging method used ensures the trends sum to one; however, for continuous properties when the number of samples in each variable is not equal, this result does not necessarily hold. The parameters on Line 15 will ensure the trends for compositional variables sum to one in this case.

Before explaining the plotting options, there is an additional option for three dimensional trends that is not identified in the parameter file. It is possible to calibrate 3D trends using a vertical trend. This is important since vertical trends are often more informed than aerial or 3D trends, especially in petroleum applications where most wells are drilled vertically. The calibration is done by inputting a second trend type option and a vertical search radius on Line 6, for example, Line 6 may look like:

```
2  0  15.0    -trend type...
```

The search radius is input so that a search different than that used for the 3D windows is possible. An additional parameter is also required on Line 10, which is a second number of nearest neighbors to use. The first number is for the 3D trend and the second for the vertical trend. For example, the number can be limited to 25 for 3D trends and all in the window size for vertical:

```
25   0     -nearest neighbors to use
```

Some plotting capabilities have been added to the program so that vertical trends can be easily verified. Curves are plotted with the trend value on the x-axis and the z coordinate on the y-axis. For compositional variables such as facies proportions, the curves can be plotted as is, or the cumulative proportion curves can be plotted using the parameter on Line 22. Cumulative curves are filled with color to help visualize relative proportions. On Line 24, a scaling parameter for the x-axis is provided so that the plot can be tuned to better fill the postscript page. When a plot is generated, an additional curve is added that shows the number of nearest neighbors found in the specified vertical search range. When the number of neighbors is fixed, the plotted curve provides the effective search radius that would return that number of neighbors. The maximum value of the curve is also added to the plot so that values can be scaled relative to the x-axis.

Plots generated by the MAKETREND program for vertical trends contain no annotations. They can be enhanced using various commands that are added to the parameter file (Table 2). Continuing line numbers from the last table of parameters they are:

Table 2: Additional parameters for vertical trend plotting.

Line	Parameters and Description	
26	LEGEND 3	-Adds a color coded legend to the plot
27	3 sand	- legend color code and label
28	15 mixed	
29	10 shale	
30		
31	XLABEL (name)	-Add a label to the x-axis
32	YLABEL (name)	-Add a label to the y-axis
33	TITLE (name)	-Add a title to the plot
34	XGRID (default)	-Plot a set of vertical grid lines
35	YGRID (default)	-Plot a set of horizontal grid lines
36	AXIS (10,1,20,2)	-Set the number of labels and precision for x and y axes
37	BOX	-Add a box around the plot

On Line 26, the legend will have 3 labels, which must be input on the following lines. Each line contains the color code and label. The label cannot have spaces so for multiple words use underscores. For the axis labels and title (Lines 31, 32, 33), 'name' can be any text; however, it must be enclosed in parenthesis to be added to the plot. For XGRID and YGRID, the default option plots grid lines at the spacing defined by AXIS, and if AXIS is not present, this number is 10 for both axes. To specify a different grid, default is replaced by the number of grid lines, origin, and spacing. For example, `XGRID (10,0.5,0.25)` will plot 10 grid lines starting at 0.5 and increasing by 0.25. The AXIS command specifies the number of labels and decimal precision for each axis. For the values input on Line 36, the x-axis will have 10 labels with 1 decimal retained while the y-axis will have 15 labels with 2 decimals retained.

A simple example of the vertical trend plotting involving three analytical trend functions is given in Figure 6. The functions represent proportions and therefore sum to 1. There are 41 equally spaced samples with z coordinates ranging from -1 to 1. The three functions are $A = z^2/3$, $B = z^3/4 + 1/4$, and $C = 1 - (A + B)$, where A, B and C are variables. A vertical search radius of 0.11 was used with Gaussian weighting.

5. Conclusions

In some instances, trend models are a useful component for geostatistical modeling. A variety of methods exist to build trend models from sample data. This paper explored the use of moving window averaging primarily for constructing trend models for categorical variables, that is, trends that describe the local categorical proportions at any location. The proportions form a composition that sums to one. Using moving window averages ensures this property is maintained by the trends. A flexible program for building moving window average trends was developed.

References

Babak, O., Deutsch, C.V., 2009. Improved spatial modeling by merging multiple secondary data for intrinsic collocated cokriging. Journal of Petroleum Science and Engineering 69, 93-99.

Chiles, J.P. and Delfiner, P., 1999, Geostatistics: modeling spatial uncertainty. Wiley-Interscience, 720 pp.

Deutsch, C.V., 2002, Geostatistical reservoir modeling. Oxford University Press, 384 pp.

Journel, A.G. and Rossi, M.E., 1989, When do we need a trend model in kriging? Mathematical Geology, 21(7), 715-739

Deutsch, C.V. and Journel, A.G., 1998, GSLIB: geostatistical software library and user's guide. Oxford University Press, 384 pp.

Figures

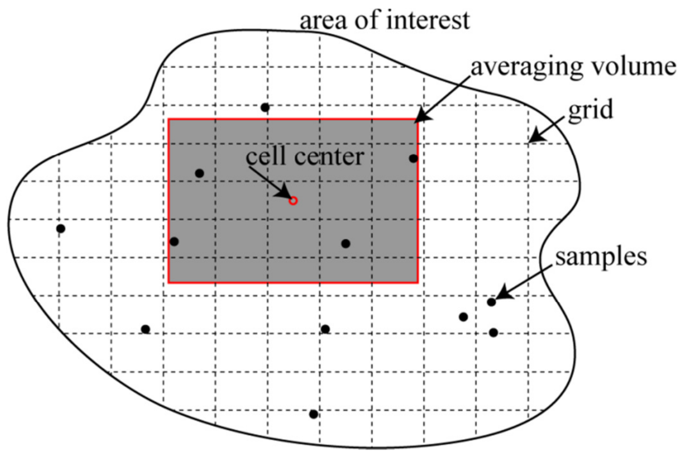

Figure 1: Elements involved in moving window averaging.

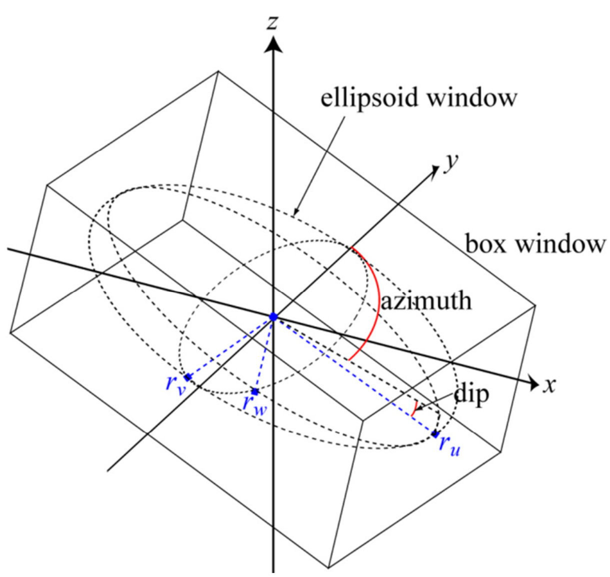

Figure 2: Box and ellipsoid averaging windows and associated parameters.

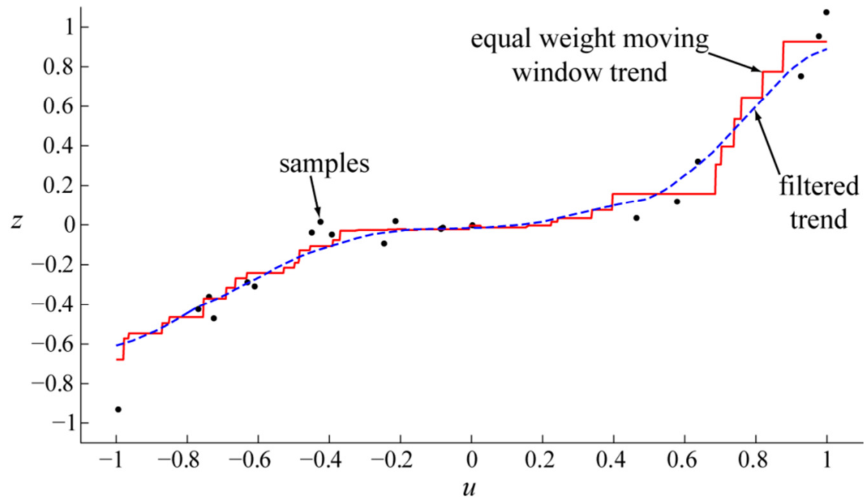

Figure 3: Example of a discontinuous trend resulting from equal weighted moving window averaging. A window radius of 0.24 units was used. Discontinuities can be filtered out.

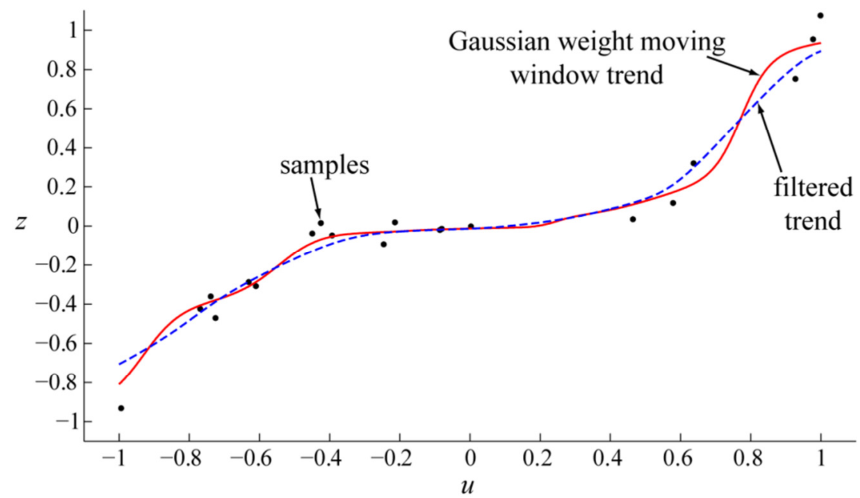

Figure 4: Example of a continuous trend resulting from Gaussian weighted moving window averaging. A window radius of 0.5 units was used. Filtering can be used to smooth the result.

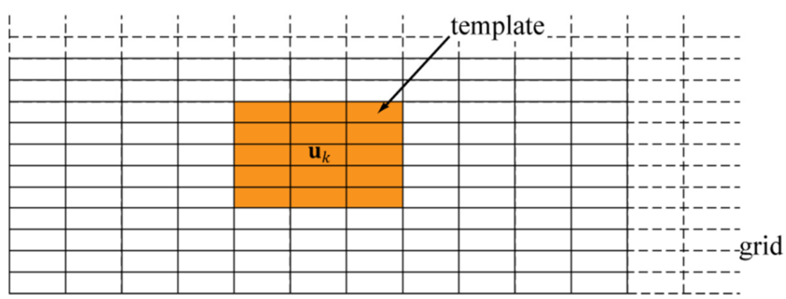

Figure 5: Example two dimensional filter template with $nx = 3$ and $ny = 5$. All shaded grid cells are used to compute the average at $\mathbf{u}_k$.

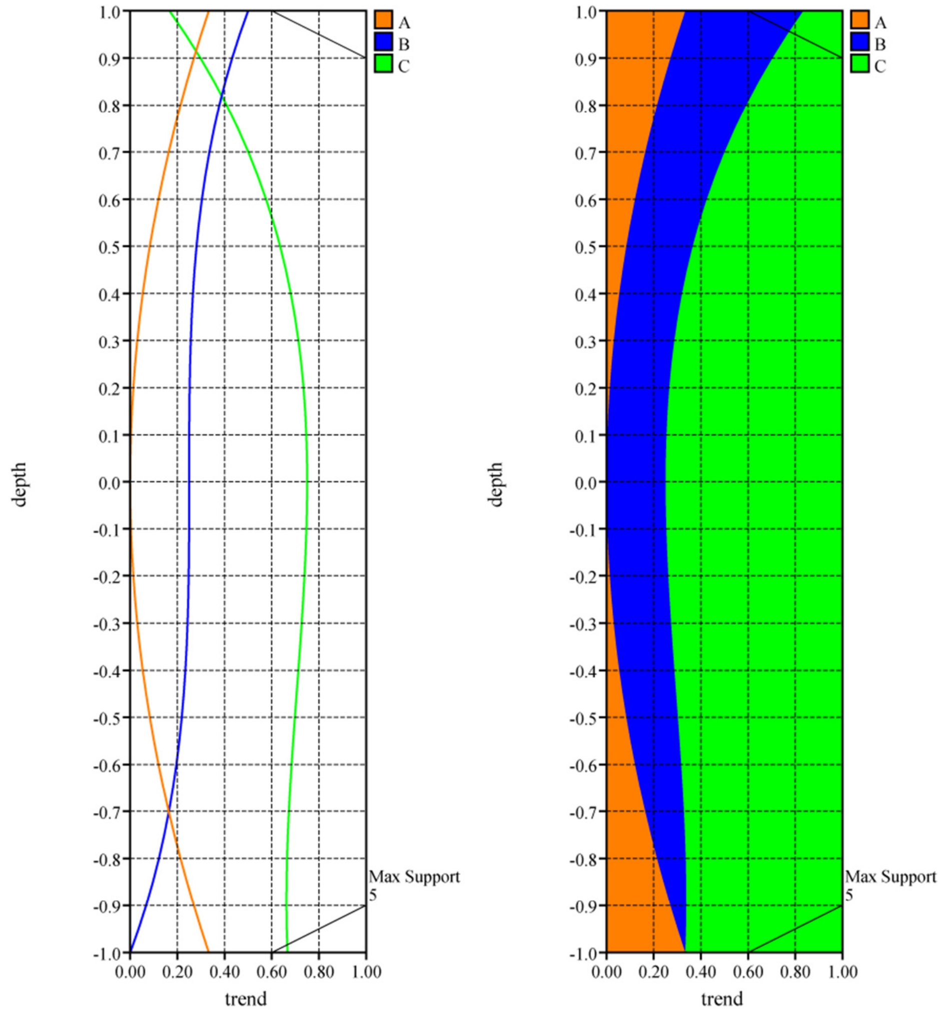

Figure 6: Simple vertical trend models showing proportion (left) and cumulative proportion (right) trend plots generated by MAKETREND using a search radius of 0.11 and Gaussian weights. Filtering was not applied.