Modeling Multiple Rock Types with Distance Functions: Methodology and Software

Daniel A. Silva and Clayton V. Deutsch

The sub division of the deposit into estimation domains that are internally consistent and externally different has a great practical importance in resource estimation. The geometric limits of the different domains have a large impact on the mineral resource estimate. It is common to create a manual 3D wireframe geological interpretation to provide the geometric limits for estimation or simulation of the grades. The distance function technique provides a fast automatic approach to support the interpretation process; however, the distance function approach is primarily developed for the binary case. An extension to more than two types of rock is proposed in this paper. A measure of uncertainty in the correct domain is measured by a coefficient that provides a measure of confidence of the different geologic domains prevailing at each location.

Introduction

All geostatistical techniques require an assumption of stationarity. The deposit is subdivided into domains, each of them modeled separately. The key problems are to assemble representative statistics within each domain and establish the spatial extents of each domain. The answer to these problems amounts to a decision of stationarity.

Cross sectional and plan view interpretations constitute an effective and reliable technique to understand the spatial extent of the different domains. Construction of 3D geological wireframes within a mining package is a critical step in most resource estimation. The different domains are digitalized on cross sectional and plan views, allowing an interpretation of the shape and volume of the different mineralized domains. This is, however, a time consuming process. Moreover, this manual interpretation process is not particularly flexible for fast updating when new information becomes available. Also, the manual interpretation permits professional judgment to be brought to bear on the results, which is an advantage; however, there is a significant amount of subjectivity and lack of reproducibility, which is a disadvantage.

The subjectivity of geological models constructed by manual interpretation affects the shape and volume of modeled domains, especially in those sectors where the information is sparse. The manual interpretation assumes fixed boundaries between domains with no assessment of uncertainty is not provided. The distance function technique is a useful approach to overcome some concerns related to the manual interpretation process. It does not account for the fact that a trained professional can often understand important trends and specific patterns leading to improved results. The distance function methodology is based on the interpolation of a distance measure. Each sample is assigned an anisotropic Euclidean distance between itself and the nearest sample belonging to another domain. Negative values represent the distance to the boundary for samples within a particular domain, and positive values are used for the distance to the boundary for samples outside the domain.

The interpolation of this distance measure aims at constructing binary domain models through the application of a zero cut off rule. Volumes of negative distance function are inside the domain (McLennan, 2008; Hosseini, 2009). Any estimation technique could be used for the distance function variable, but kriging takes into account the anisotropy and spatial continuity of the distance measure. An original approach developed by Wilde and Deutsch (2011) introduces the uncertainty concept and proposes its assessment between domains. The key idea consists of calibrating a bandwidth of uncertainty along the domain frontier with two parameters C and β. The first controls the thickness of bandwidth and the second the position and distribution of tonnage within the bandwidth. This is fully developed in the recent master's thesis of Mike Munroe (Munroe, 2012).

All distance functions techniques proposed before have been focused on geological modeling the binary case - interpolation of positive and negative values corresponding to out and in a specific domain of interest. We propose an extension to any number of domains and introduce a variance ratio that is able

to assess the uncertainty in the assignment of domain types. A FORTRAN code – dfmod– was developed to perform the multiple domain modeling.

Distance Function Binary Case

As explained by Hosseini 2009, the distance function starts by coding the samples with zero or one according to their domain, as shown in expression (1). Either positive or negative distances measure may be calculated. Commonly, negative values are assigned to those samples inside the domain and positive values outside (2).

$$i(u_\alpha) = \begin{cases} 1 & \text{if } u_\alpha \text{ is located within domain} \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where u_α corresponds the position of samples, with $\alpha = \{1, \ldots, n\}$. A classical Euclidean distance measure is considered to calculate the distance from the sample to the nearest boundary (2).

$$dF(u_\alpha) = \begin{cases} -d & \text{if } u_\alpha \text{ is located within domain} \\ +d & \text{otherwise} \end{cases} \quad (2)$$

Anisotropy can also be accounted for according to the geological shape and orientation of the domain, as shown the following equation. Where *dx*, *dy* and *dz* represent the differences between the coordinates of two samples, and *rx*, *ry* and *rz* correspond to the ranges of anisotropy. The X, Y, Z coordinates would correspond to the rotated directions of major continuity.

$$d = \sqrt{\left(\frac{dx}{rx}\right)^2 + \left(\frac{dy}{ry}\right)^2 + \left(\frac{dz}{rz}\right)^2} \quad (3)$$

An interpolation process – inverse distance or kriging – is performed using the assigned distance measure. Both inverse and kriging may generate smooth and realistic boundaries for the domain. The use of all samples in a global search strategy avoids the presence of artifacts or undesirable patterns.

Figure 1 shows the modeling of two domains based on the distance function for the binary case. First, the samples are flagged using expression (1); afterward, the distance from the sample to the nearest boundary is calculated according with (3). A global ordinary kriging is performed considering the variography and geological anisotropy from the exploratory data analysis. Finally, at all estimated location the expression (2) will determine the domain which it belong applying the zero cut off rule.

Distance Function Extension based on Multiple Domains

This work proposes an extension of the distance function technique from the binary case to multiple domains. The aim is to model multiple domains at the same time using a procedure similar to the binary case. First, we start by coding the samples *k* times, once for each domain *(k=1,…,K)*.

$$I_k(u_\alpha) = \begin{cases} k & \text{if } u_\alpha \text{ is located within domain k} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

Thus, we will have at each sample *k-1* indicators equal to zero and one equal to *k*, which is associated to the *k* domain. A function *F* maps the indicator of domain *k* with the distance measure at that sample, as shows expression (5).

$$F(I_k(u_\alpha)) = d \bullet F_k(u_\alpha) = \begin{cases} -d & \text{if } I_k(u_\alpha) = k \\ +d & \text{if } I_k(u_\alpha) = 0 \end{cases} \quad (5)$$

The *k-1* distance functions that have positive values indicate the distance between the samples outside of domain *k* and the closest sample within domain *k*. Conversely, the remaining *k* negative value represents the distance of the sample within the domain and the closest sample outside. Finally, for each *k* indicator an interpolation of distance functions is performed. Each unsampled location has *K* estimate distances.

The distance function estimates may be interpreted as the closeness to the domain boundary. In the case of multiple domains, the minimum value between all estimated distances represents the most distant domain to the boundary or the domain with less uncertainty at that location. Expression (6) summarizes the proposed solution for multiple domains, which codes each interpolated location with the indicator associated to the minimum value of the *K* estimated distances.

$$I_k^*(u_\square) = F^{-1}(\min\{dF_1^*(u_\square, I_1), dF_2^*(u_\square, I_2), ..., dF_k^*(u_\square, I_k)\}) \quad (6)$$

Let's illustrate the proposal technique with an example using four domains. Figure (2.A) represents the view on a bench of samples of each domain, each of them represented by squares, circles, diamonds and triangles. The indicator expression (4) is used to code each sample for each domain, repeating the process four times. Afterward, at each data location, the sample code is mapped to a distance function using (5), assigning negative values for those samples inside the domains and positive outside. A global ordinary kriging is performed for each domain taking into account the anisotropy and variography analyzed from the samples. The final code at all unsampled location corresponds to the minimum estimated distance measure from the four domains according with (6). Figure (2.B) shows the samples and the final coded model.

A second example is presented in Figure 3. In this case, the modeling process is performed on a cross sectional. Similar to the example above, four drillholes show that four domains have been interpreted: hanging wall (circles), the first vein structure (squares), a second vein structure (diamonds) and finally the foot wall (triangles). Figure (3.B) shows the shape and orientation of the final model of the four domains.

In both cases, the anisotropy plays an important role in the interpretation realized by the distance function technique. Also, the understanding of directions of continuity is extremely important to regulate the extrapolation of certain domains. Some professional judgment and subjectivity is inevitably required.

Estimation Parameters Calibration and U Coefficient

Cross validation or the Jackknife could be used to evaluate the performance of the technique and improve the estimation/simulation parameters. Most cross validation/jackknife implementations in mining software packages work with continuous variables. In addition to cross validation, it would be useful to have a measure of certainty in the assignment of rock types at each location.

The proposed methodology estimates a distance measure for each domain. The classification process explained by Expression (6) shows that the code associated at an estimated location corresponds to the minimum value. We can utilize the relationship between this minimum estimated distance and the remaining estimates to introduce a U coefficient that measures how much the chosen domain dominates the others. Let's define the coefficient as the ratio between the variance of all the estimated distance values less the minimum value standardized by the variance of all estimated distances.

$$U(u_\square) = \frac{Var(dF_1(u_\square), ..., dF_{m-1}(u_\square), dF_{m+1}(u_\square), ..., dF_k(u_\square))}{Var(dF_1(u_\square), ..., dF_{m-1}(u_\square), dF_m(u_\square), dF_{m+1}(u_\square) ....., dF_k(u_\square))} \quad (7)$$

where $dF_m\ (u_\square)$ is the minimum estimated distance. This ratio takes a value between zero and one. If the value is close to zero the difference between the minimum and the remaining estimated distances is large, that is, most of total variance is due to the minimum value. This can be interpreted as the code that was assigned at that location has a high probability to belong to that domain. On the other hand, if that value is close to one the distance of the chosen domain is close to the distances of the other domains, namely, any of the other domains could be found at the location; there is large uncertainty.

An example developed in Figure (4) shows some results for a model of four different cross validation estimation parameters settings; Figure (4.A) represents the original domain configuration of the samples. The performed cross validation considers the same anisotropy and variogram ranges and structures; however, the orientation of the maximum range in the variogram changes for each cross validation. Thus, Figure (4.B) show a configuration with a dip of 135 degrees, (4.C) 70 degrees, (4.D) 0 degrees and (4.E) 35 degrees major axis orientation. The aim is to show cross validation using parameters that do not reflect the real shape and orientation, and parameters that properly fit the ore body characteristics. A specified set of samples was chosen to evaluate the cross validation performance. Each domain changes from the worst to the best configuration of parameters (bar graphic from F to I). It is possible to observe that the best parameter configuration is represented by cross validation E and Figure (4.I), a variogram with an orientation of 35 degrees reflecting the original model Figure (4.A).

The U coefficient quantifies the performance of the input parameters on the cross validation/jacknife process. Figure (5) shows a plot of the U coefficient for all samples submitted to cross

validation with different settings. The graphic shows how the line of each cross validation moves to the upper left corner because the values of U coefficient decreases from cross validation configuration B to E, which indicates the improvement of the estimation. Thus, the U coefficient allows comparing the performance between different configurations of estimation parameters using a cross validation process. Finally, Figure (6) shows the final estimated model related with each of the input parameters used. The different input parameters affect the estimated models from a poor quality Figure (6.1) to the best acceptable model Figure (6.4).

The use of the U coefficient is not limited to the evaluation of cross validation/jacknife results. The coefficient can be calculated at all estimated locations on a grid representing a measure of uncertainty. The same example utilized in Figure (2) illustrates its use. Figure (7) shows a U coefficient map where the locations near to the boundaries present values close to one, and locations clearly within a specific domain have values close to zero. There are clear zones of uncertainty near some boundaries. The proposed application methodology helps us understand the continuity and relationship between domains, but requires a reasonable knowledge of the deposit orientation. The main goal is to provide a supporting tool for geological interpretation that allows distinguishing the places where professional judgment should be considered.

Program Parameters

The source code and the executable file are provided with this paper.

```
1                  Parameters for DFMOD
2                  ********************
3
4  START OF PARAMETERS:
5  Data.prn                          -   file with data
6  0   1   2   3                     -   columns for DH,X,Y,Z,var
7  -1.0     1.0e21                   -   trimming limits
8  3                                 -   number of categories
9  1 2 3 4                           -   codes
10 dfmod.out                         -   file for distance function output
11 rtmod.out                         -   file for rock type output
12 0                                 -   option: 0=grid, 1=cross, 2=jackknife
13 xvk.dat                           -   file with jackknife data
14 1   2   0    3    0               -   columns for X,Y,Z,vr and sec var
15 500     0.5       1               -   nx,xmn,xsiz
16 500     0.5       1               -   ny,ymn,ysiz
17 1       0.5       1               -   nz,zmn,zsiz
18 1    1     1                      -   x,y and z block discretization
19 16   32                           -   min, max data for kriging
20 0                                 -   max per octant (0-> not used)
21 1                                 -   Code
22 1000   1000  1000                 -   maximum search radii
23  0.0    0.0   0.0                 -   angles for search ellipsoid
24 1     2.302                       -   0=SK,1=OK
25 1    0.1                          -   nst, nugget effect
26 1    0.9  135.0   0.0   0.0       -   it,cc,ang1,ang2,ang3
27           150.0  50.0  100.0      -   a_hmax, a_hmin, a_vert
28 2                                 -   Code
29 1000  1000  1000                  -   maximum search radii
30  0.0    0.0   0.0                 -   angles for search ellipsoid
31 1     2.302                       -   0=SK,1=OK
32 1    0.1                          -   nst, nugget effect
33 1    0.9  135.0   0.0   0.0       -   it,cc,ang1,ang2,ang3
34           150.0  100.0  100.0      -   a_hmax, a_hmin, a_vert
35 3                                 -   Code
36 1000  1000  1000                  -   maximum search radii
37  0.0    0.0   0.0                 -   angles for search ellipsoid
38 1     2.302                       -   0=SK,1=OK
39 1    0.1                          -   nst, nugget effect
40 1    0.9  0.0   0.0   0.0         -   it,cc,ang1,ang2,ang3
41           400.0  100.0  100.0     -   a_hmax, a_hmin, a_vert
```

The parameters are equivalent to the kt3d program. Lines 5 to 9 are self-explanatory. The output file with the calculated distance between units and the final file with the codes for each unit at all location are set up on lines 10 and 11. There exists the option on line 12 to perform the methodology as a cross-validation or jacknife tool. From lines 14 to 20 correspond to parameters for the grid, block discretization and estimation. From line 21 to 27 are established the parameters for search, variography and type of kriging, which are evident. From here forward, for each new categorical code a new set of lines with the parameters need to be written, that is the case for the codes 2 and 3.

Conclusions

The interpretation and knowledge of domains – geological units – is a crucial stage for mineral resources estimation. Nowadays, the interpretation process is mostly done manually considering the judgment and expertise of geo-professionals; however, this process is time consuming due to the amount of plan and cross sectional views that must be modeled, which depends on the complexity of the ore body.

Furthermore, there is a lot of work when new information becomes available, e.g., new drilling campaigns. The distance function technique partially overcomes these issues by supporting the geological interpretation. The proposed work extends the capacity to model more than two domains and also introduces an uncertainty U coefficient. This coefficient is able to quantify the performance of estimation parameters through a cross validation/jacknife process adjusting the parameters and improving the quality of the final models. Indeed, the U coefficient can be interpreted as a tool to assess the uncertainty at unsampled locations, especially to those near to the boundaries. This novel distance function approach supports information about patterns, location boundaries and uncertainty for geo-modeling purposes.

References

Hosseini, A.H., 2009. Probabilistic Modeling of Natural Attenuation of Petroleum Hydrocarbons, *Ph.D. Thesis*, University of Alberta, 359p.

Munroe, M.J. and Deutsch, C.V., 2008a. A methodology for modeling vein type deposit tonnage uncertainty. Center for Computational Geostatistics Annual Report 10. University of Alberta. 10p.

Munroe, M.J. and Deutsch, C.V., 2008b. Full calibration of C and β in the framework of vein-type deposit tonnage uncertainty. Center of Computational Geostatistics Annual Report 10. University of Alberta. 16p.

Munroe, M.J., 2012. A Methodology for Calculating Tonnage Uncertainty in Vein-Type Deposits, *M.Sc. Thesis*. University of Alberta. 188p.

Wilde, B.J. and Deutsch, C.V., 2011. A new way to calibrate distance function uncertainty. Center of Computational Geostatistics Annual Report 2011. University of Alberta. 8p.

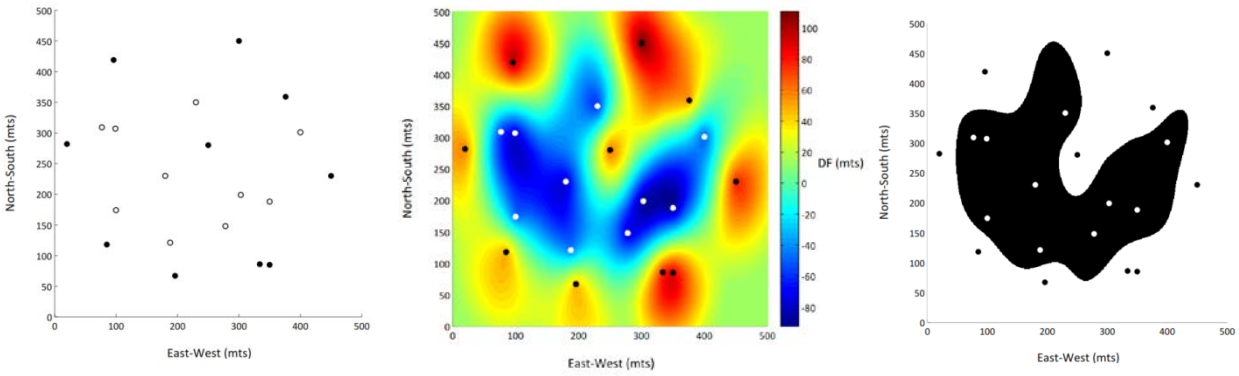

Figure 1: Samples on a plan view within domain – white bullets- and outside domain black bullets (upper left); negative and positive interpolated distances (upper right), and location flagged as domain – black – and outside – white- (bottom).

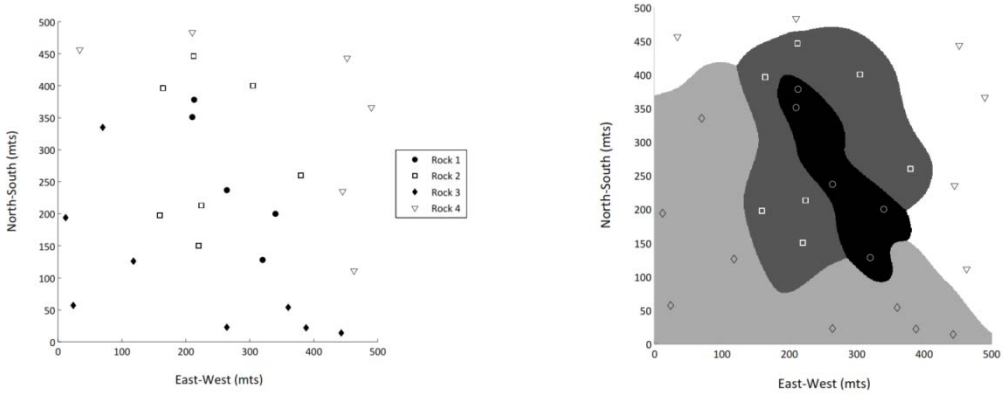

Figure 2: Samples on a plan view for four different domains (left), and the final domains model determined from distance functions interpolation (right).

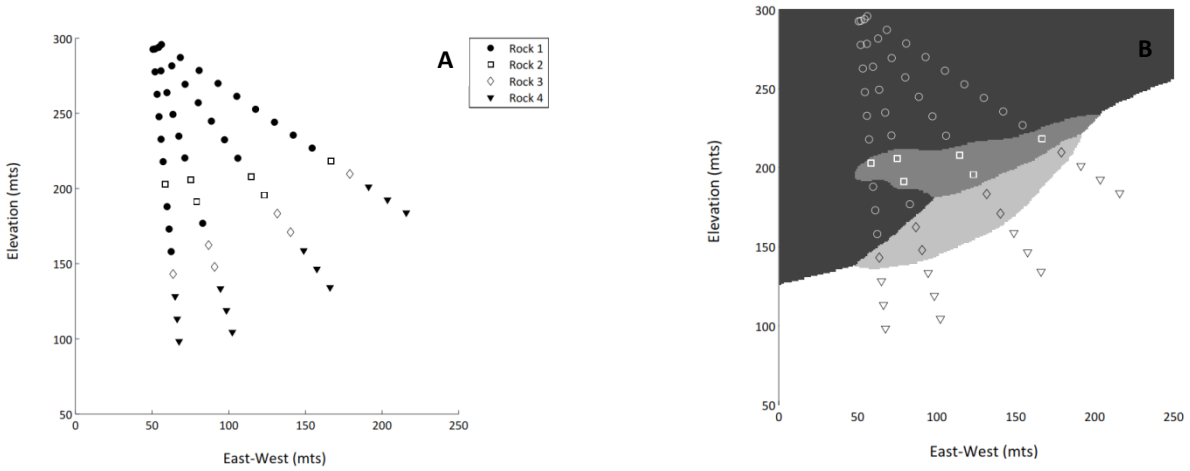

Figure 3: Four different domains samples on a cross section (left), and the final domains model determined from distance functions interpolation (right).

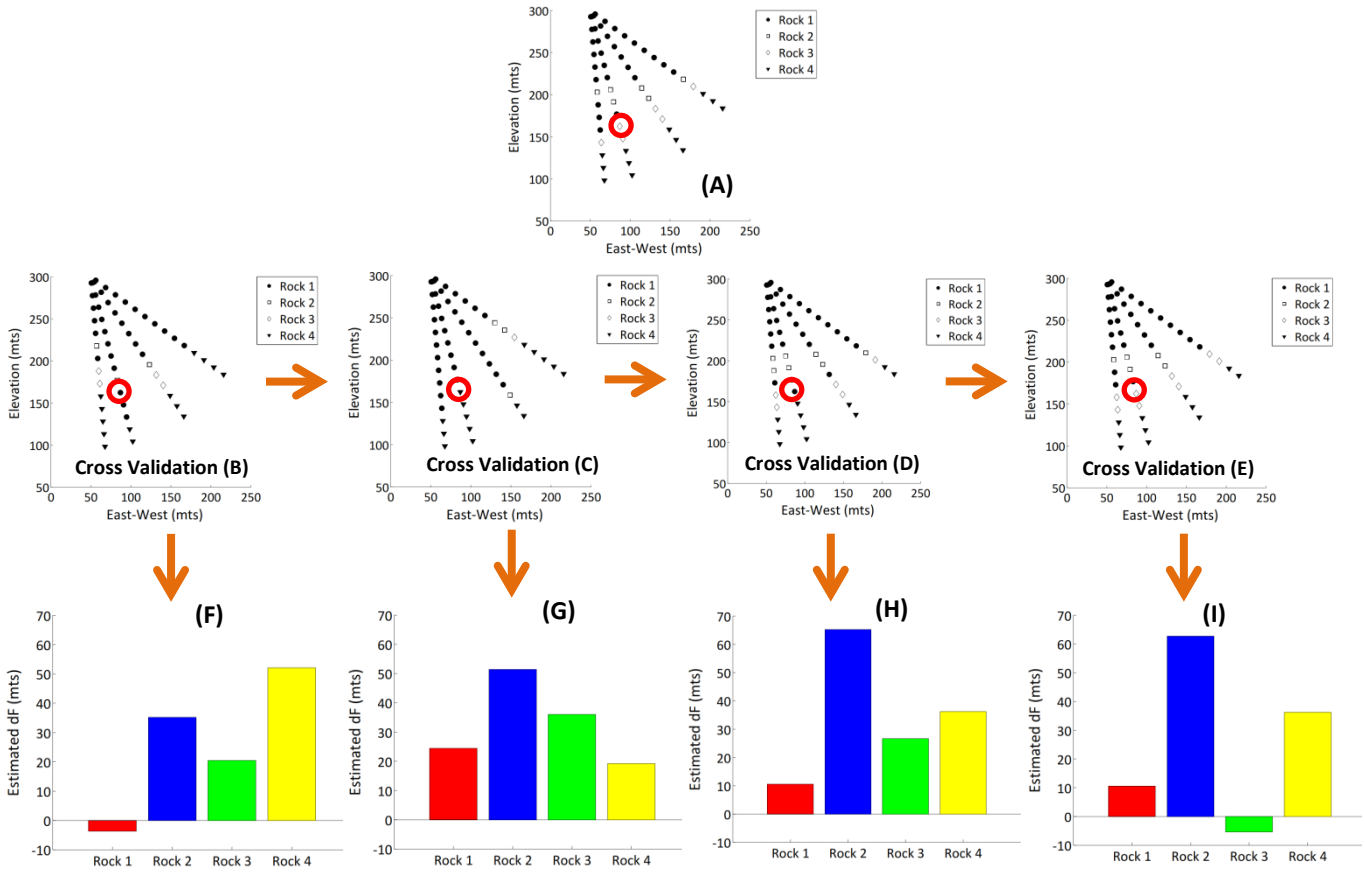

Figure 4: Cross validation of samples with four different parameters set up (above). Bar graphic with estimated distances for a particular sample (below).

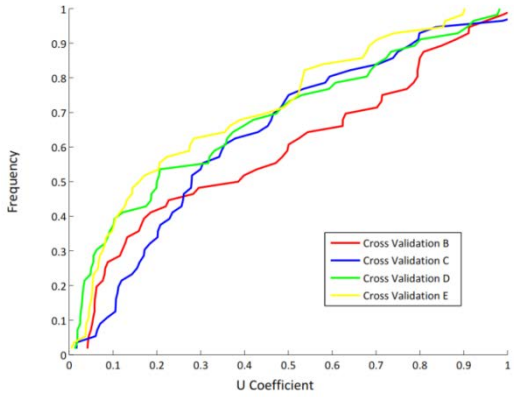

Figure 5: U coefficient plotting for each sample according with each cross validation parameters configuration

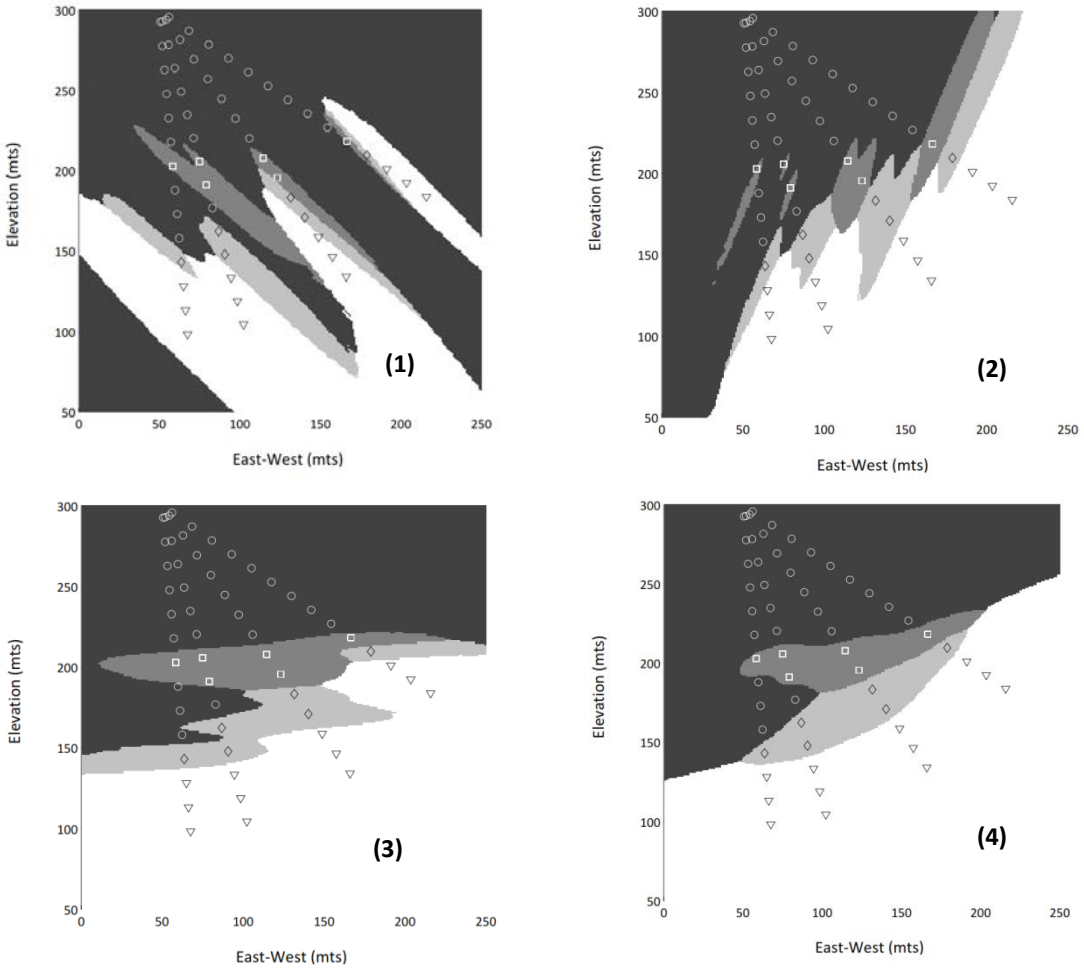

Figure 6: Four different final models for different configuration of input parameters. Each of the figure corresponds to the parameters set up for the cross validation 1 to 4.

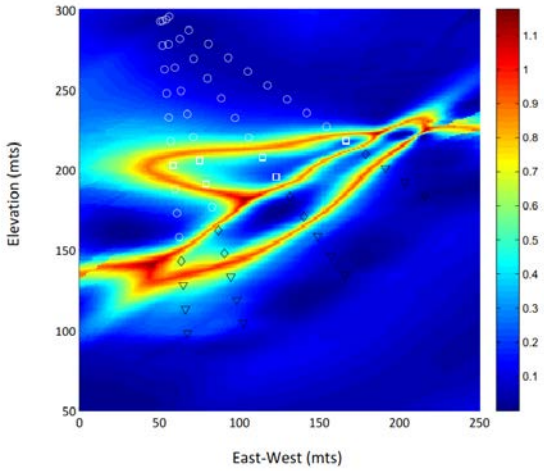

Figure 7: U coefficient map for assessment of uncertainty