3D visualization of sparse geophysical data representing uncertainty

Vítor Gonçalves[1,2], Paulo Dias[2,3], Fernando Almeida[4], Joaquim Madeira[2,3], Beatriz Sousa Santos[2,3]

[1]Polytechnic Institute of Castelo Branco, Portugal
[2]Department of Electronics, Telecommunications and Informatics, University of Aveiro, Portugal
[3]IEETA, University of Aveiro, Portugal
[4]Department of Geosciences, University of Aveiro, Portugal

ABSTRACT

Geophysical data are sparse and by nature difficult to analyze. Usually domain experts use "mental models" to infer missing data according to the surrounding data and their own knowledge. The main goal of this work is to explore the best way to represent uncertainty in geophysical data. Given the sparse nature of the represented data, it is important to provide a 3D volumetric representation of the whole subsoil, based on a geostatistical process. We use kriging interpolation to generate a structured grid from the original sparse data. However, the analysis of such an interpolated representation must be careful, since the uncertainty varies significantly according to the distance to real measurements. We use different representations to emphasize data uncertainty during the analysis stage. The different visualization techniques implemented in our prototype, as well as methods used to simultaneously visualize resistivity and uncertainty information, are presented.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Visualization, uncertainty visualization, geophysical data, VTK (Visualization Toolkit).

1. Introduction

Geophysical data represent subsoil structure in a particular area. Their analysis allows defining a model of the geological area that helps, among others, drilling wells for water, mining, bedrock detection, monitoring of environmental problems, as well as the degree of change in rock masses. Data acquisition is always performed in a discontinuous way at different locations in a limited area, since continuous acquisition at all locations is not technically and economically feasible. Thus, geophysical data are typically sparse, resulting in large areas where no data is available.

Most commercial systems for geophysical data visualization allow a simple representation of the sparse data extracted from the subsoil (as ArcGIS or Oasis montaj [E11, G11], which do not allow the visualization of uncertainty or interpolation in a volume). However, it is desirable to visualize data between samples, either through manual inclusion or through an interpolation method. Flexible tools allowing the interpolation of sparse data, as well as the ability to interactively visualize such data through different representations, are important for domain experts. If no such interpolation is performed, experts have to create/imagine a hypothetical formation of subsoil areas based on their knowledge, (i.e., a "mental model"). A visual and interactive tool simultaneously providing various interpolation and visualization methods, as well as the possibility to adjust some parameters, could significantly help and guide experts during data exploration [S07].

Visualization offers powerful methods to represent data. Nevertheless, when the acquired data have some associated uncertainty, the final visualization may lead to erroneous conclusions, since users may consider interpolated data as acquired data [SBB08]. To avoid this, and according to [GR02], uncertainty visualization techniques should be informative, intuitive, non distracting, and interactive.

Several visualization methods for sparse geophysical data (electric resistivity) have been described in a previous work [GDAS10]. The methods developed proved to be useful in sharing knowledge between experts (representing a visual subset of the expert's "mental model"). In this work, we integrated an interpolation model to obtain values in areas without data. To prevent an expert being misled by those areas, we have evaluated several visualization techniques for uncertainty values associated with the interpolated data.

Although the main purpose of this study is to validate the adequacy of the uncertainty visualization methods for geophysical data, we also present some interactive combination and filtering methods to visualize resistivity and uncertainty. Data used in this study came from a site on the island of Porto Santo (Portugal) and include preliminary studies for the construction of a holiday resort, and the software was developed in C++ using VTK (Visualization Toolkit) [SMML98].

Section 2 presents the type of data for which visualizations had to be developed, as well as the acquisition method. In section 3 the used interpolation method is described. Section 4 presents the representations developed to explore volumetric data, and section 5 presents the implemented techniques to visualize uncertainty resulting from data interpolation. Section 6 addresses various issues that came up during the implementation, and, finally, some conclusions and ideas for future work are presented.

2. Geophysical Data — Electrical Resistivity

Electrical resistivity data consist of a resistivity grid corresponding to a perpendicular plane to the topographic surface. Soil resistivity values are continuous scalar values that vary with the material of the subsoil. Soil and rocks are generally bad conductors (although they can have different conductivity values depending on moisture); increasing water content makes the soil a good conductor. Resistivity is measured in Ω·m (ohm · meters), ranging from close to zero for good conductors (particularly in the case of intruded sea water formations) to very high values for materials which are bad conductors [L05].

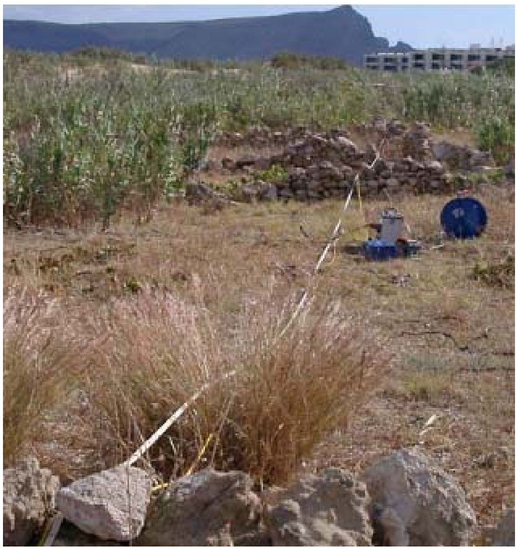

Figure 1: *Acquisition of electrical resistivity data.*

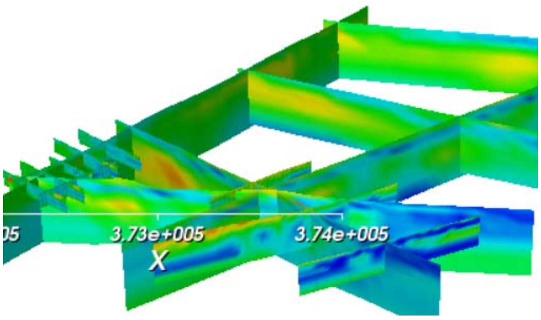

Figure 2: *Representation of electric resistivity.*

Figure 1 shows the acquisition process of electrical resistivity data, and Figure 2 is a typical representation of 16 2D grid sections. As it can be seen, the data are relatively sparse.

The adopted rainbow color scale [SSM11] is close to the one generally used in Geographic Information Systems (GIS), since it is familiar to experts and provides an intuitive interpretation: blue tones correspond to areas where there is a greater possibility of having water, yellow/brown/red tones to rocky areas, and greens correspond to intermediate zones.

3. Interpolation using Kriging

In the case of sparse scalar volumetric data based on regular or irregular grids, where some data are missing, it is useful to use interpolation to generate a final visualization without gaps. This implies the use of 3D interpolation methods.

Linear interpolation methods (e.g., trilinear interpolation) have a simpler implementation. However, they usually produce unsatisfactory results for geospatial data, since the subsoil structure does not follow a linear distribution. Also, linear interpolation works well only when the data are not sparse. More sophisticated methods based on statistical distributions produce better results, as they analyze correlations, variances, trends and other factors. Kriging, IDW (Inverse Distance Weighting), and trend estimation, among others, are examples of statistical interpolation methods [ZPRA99].

Kriging has been used to interpolate data in a regular grid [C02]. It is a regression technique mainly used in geostatistics, but also in other application areas such as medical imaging [GKD*10]. There are several types of kriging interpolation: simple kriging, ordinary kriging, cokriging, etc.

Ordinary kriging produces good results for geophysical data in particular, as well as geostatistical analysis in general, and comparatively to other techniques is one of the best for this purpose [PSAR93]. It calculates the variation between spatial points through statistical methods, since closer points have a high correlation and points far apart have no correlation. This is a method of the so called BLUE (Best Linear Unbiased Estimator) type: linear, as the estimates are weighted linear combinations, and unbiased since the average error tends to zero [GKD*10].

Ordinary kriging was chosen, as it is the most often used technique to deal with this type of data [ZPRA99], and several tests with synthetic data confirmed its adequacy to our data.

For each spatial point uncertainty was also computed using the formula:

Uncertainty = 255 · (*dist*/*maxDist*),

where *dist* is the Euclidean distance corresponding to the shortest distance between each point and a point included in acquired sections (see Figure 2) corresponding to 100% certainty. *maxDist* is the maximum distance observed in the data (which corresponds to the furthest point of all sections shown in figure 2). The values are normalized to 255 to use the full range available in the volume rendering algorithm that only accepts 8 bit images. This formula implies a linear variation of the uncertainty that produces values close to zero in points near the samples and values close to 255 in points further away. Although simple, this method is good enough as a first approximation to display the uncertainty associated with the data.

4. Visualization

Interpolation of the original sparse data results in a volumetric regular grid. Typically mesh visualization (as shown in Figure 2) is not adequate for this type of data, thus other visualization methods, more adequate to volumetric data, were explored. In this section we present the several visualization techniques we implemented.

4.1 Volume rendering

Volume rendering (VR) or direct volume rendering (DVR) visualization techniques consist in representing the data as a "translucent" material with different color and opacity properties [BW01] which can be configured through color and opacity "Transfer Functions" (TFs) [HB04].

There are a number of well-known techniques for implementing DVR: ray casting, splatting, texture mapping and shear warping, among others [lXLyX06, MKCY09]. In this work ray casting was used, since it provides a good trade off between image quality and processing speed. Additionally, this method is already implemented in VTK.

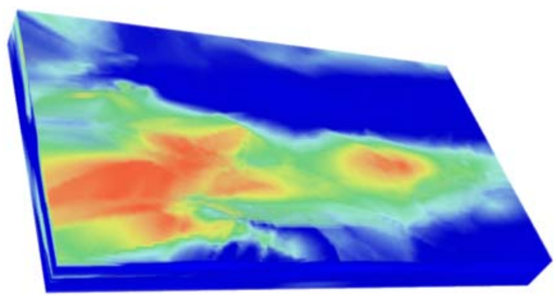

Figure 3: *Data visualization using volume rendering.*

Figure 3 shows a volume where the scalar data is mapped through color using a constant opacity of 35%.

4.2 Slicing

Slicing is useful for probing volume data sets to discover where interesting regions exist, maintaining the context. Figure 4 shows an example of slicing applied to our data.

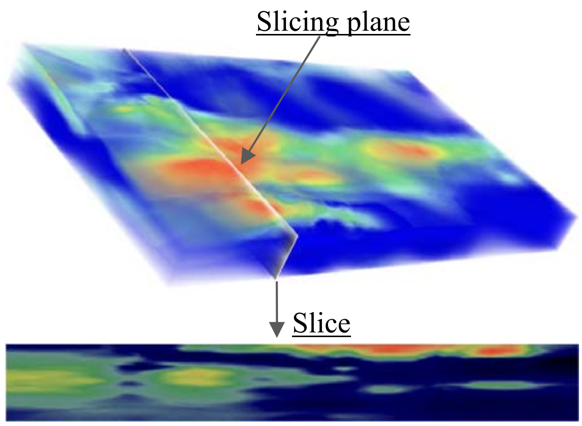

Figure 4: *Orthogonal slicing of the volume data.*

4.3 Isosurfaces

In this case the user defines a resistivity value to compute an isosurface, or two upper and lower resistivity values as well as the number of intermediate surfaces to be represented. Resistivity values are normalized so that resistivity can be considered a percentage. For example, Figure 5 provides a set of five isosurfaces corresponding to resistivity values of 80%, 85%, 90%, 95% and 100%.

Figure 5: *Five isosurfaces corresponding to resistivity values 80%, 85%, 90%, 95% and 100%.*

The representation of isosurfaces may be useful to help create a model of soil stratification.

4.4 Simultaneously filtering with resistivity and uncertainty

Isosurface filtering is useful to segment subsoil strata, when using an interactive technique, filtering is useful and visually informative. To avoid having to represent a new data volume for each operation, two scalar values are associated to each voxel, representing resistivity and uncertainty. Resistivity is mapped through color and uncertainty by opacity.

Next, in each filtering iteration, it is only necessary to modify the opacity transfer function according to the desired filtering. For example, Figure 6 represents the data set that has an uncertainty value less than 30%. In this case, the opacity transfer function is defined so that all uncertainty values higher than 30% correspond to zero opacity, and the remaining assume a constant opacity value, for example 60%. This is only possible since opacity and color are mapped independently, and it is faster since it prevents the generation of a new volume.

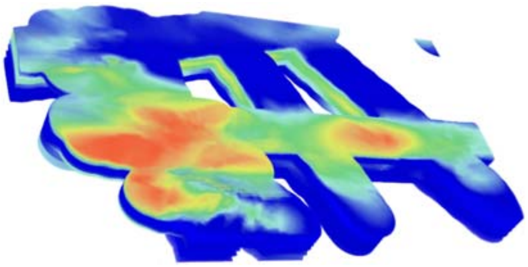

Figure 6: *Filtering visible voxels by uncertainty less than 30%.*

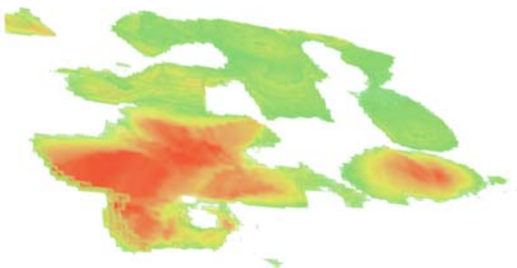

Figure 7: *Result of a similar filtering, although in this case two filters are applied simultaneously: resistivity (higher than 70%) and uncertainty (less than 30%).*

5. Uncertainty Visualization

Uncertainty representation in volumetric data is a four dimensional representation problem, with great interest in various application areas, and for which several solutions have been proposed.

A general classification of uncertainty encoding methods into two groups is proposed in [P01]: mapping uncertainty by an additional piece of data or integrating it in volumetric data (through color, transparency, etc.).

One of the first proposals for uncertainty representation involves the application of high levels of transparency in places where the data have higher uncertainty, and greater opacity where data have greater certainty [ZC06]. To implement this method in VTK each voxel is assigned two scalar values, representing data under analysis (resistivity in this case) and uncertainty, as already mentioned. This technique is intuitive for the user, since more emphasis is given to data of greater certainty. Nevertheless, it has two major disadvantages [ZC06, DKLP01]: if the user cannot interactively activate and deactivate the visualization, it can lead to the non-visualization of data with little certainty that may be relevant, as well as lead to ambiguities when data are represented by color, since when applying opacity to a color it will be represented in a lighter shade (on a light background) [ZC06, DKLP01].

To overcome those problems, Djurcilov, et al. [DKLP01] proposed alternative methods to represent uncertainty:

- inserting speckles/holes — It consists in placing small speckles (glyphs) on the data representation, adding more speckles or larger speckles in locations where there is higher uncertainty [DKLP01, DKLP02];
- adding noise — It consists in introducing noise in the data according to the uncertainty associated with each voxel, by adding or subtracting a random value proportional to the uncertainty [DKLP01, DKLP02];

Three uncertainty visualization methods using opacity, noise and speckles were implemented to represent the uncertainty of the data interpolated by kriging. Additional details are presented in the following sections.

5.1 Varying opacity

Uncertainty visualization through opacity consists in mapping each voxel's opacity according to the uncertainty associated with that particular voxel. To implement this method in VTK, each voxel is given two scalar values as mentioned above.

This representation is very intuitive and there are situations where it is very useful, although it must be used with care, since it can lead to misinterpretation, given that a higher transparency can imply the perception of a lighter color. It is especially useful in an initial exploratory stage, giving a quick notion of the subsoil structure.

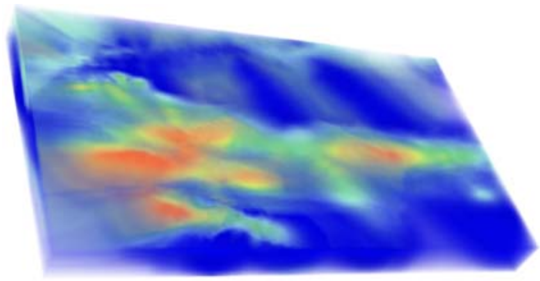

Figure 8: *Uncertainty visualization through opacity.*

5.2 Inserting speckles/holes

Speckles/holes are inserted when uncertainty is larger than a given value (e.g., 30%) and speckle size is proportional to the uncertainty associated with the data. Speckles are represented by a 2D polygonal mesh that overlaps the volume surface.

A 2D representation for uncertainty was used, since the original data are defined by planes perpendicular to the surface. Thus a 3D representation along *z* would not bring any additional information, since uncertainty is computed based on the distance to those original perpendicular planes. In addition, a 2D representation introduces less visual structures in the uncertainty representation and facilitates the data visualization, as shown in Figure 9. Squares were used; however, other shapes could have been used.

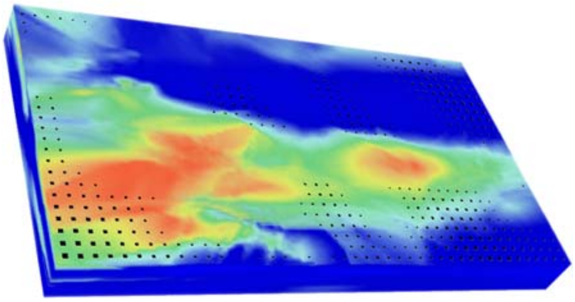

Figure 9: *Uncertainty visualization: inserting speckles on a data volume.*

5.3 Adding noise

This method consists in adding noise to the original data resulting in a different mapping for uncertain areas (small color changes, blurred areas) which leads to less precision in these locations. Noise is gradually applied according to uncertainty, ranging from no noise (uncertainty below 30%), ±10% noise level (uncertainty between 30%-44%) and so on using five levels as shown in Table 1. Only five levels were used as prior tests revealed that beyond this value it is much more difficult to perceive discontinuities. Figure 10 shows the result of applying this method to our data.

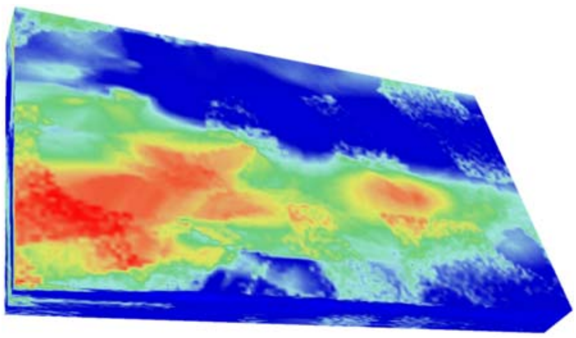

Figure 10: *Uncertainty visualization: adding noise to the original data.*

Uncertainty value	Noise level
[30% to 44%[	[±10%]
[44% to 58%[	[±20%]
[58% to 72%[	[±30%]
[72% to 86%[	[±40%]
[86% to 100%]	[±50%]

Table 1: *Noise levels used to represent uncertainty values.*

6. Discussion

The proposed visualizations have been evaluated and validated by a small group of experts and integrated in a prototype. Compared to existing tools, our prototype enables visualization without discontinuities in data, having the advantage of integrating multiple visualizations of the same data interactively. As a result should be able to explore the data more efficiently and quickly.

A preliminary study comparing the three uncertainty representations was performed with the help of two domain experts. According to experts, uncertainty representation through speckles was superior, enabling more accurate visualizations, as it is a mapping independent of the data volume. Moreover, compared with uncertainty representation through noise, speckles are especially advantageous when observing regions of higher uncertainty, since uncertainty information is shown without data distortion. Compared with the opacity method, speckles have the advantage of avoiding misleading color perception effects created by opacity variations.

Still according to the domain experts, visualization of uncertainty through opacity and noise seem more appropriate for an initial and global data exploration. Uncertainty visualization through speckles proved to be the most suitable method for more intensive explorations of the data, since it preserves all data attributes. However, to avoid unwanted data occlusion, it is necessary to allow interactive activation/deactivation of this feature.

7. Conclusions and Future Work

Often, in data visualization, uncertainty is discarded to avoid complex representations. However, such information can be relevant to guide final decisions, specifically in sparse datasets. Thus, the representation of uncertainty is important. In this paper several methods were used to visualize with uncertainty a type of sparse geophysical data (electric resistivity). According to a preliminary study using noise and opacity has proved particularly useful for a global data exploration. Representation through speckles is also useful in a global exploration; however, it seems particularly interesting for a more detailed and intensive analysis of the data.

It is fundamental to conduct further user studies to evaluate the developed methods with larger group of users. Also, other types of speckles/glyphs should be considered.

There are also other possible improvements: i) extracting uncertainty values from the actual kriging variograms [C02]; ii) allowing manual changes in interpolated regions making the tool more flexible; iii) implementing an automatic learning system based on such corrections.

A further step in our work will be the implementation of “focus + context techniques” to explore the interior of the volumetric data, for example, through “magic lenses” or deformation [RH06, MTB03]. Another interesting technique might be volume pseudo-haptics, which involves the application of force feedback in volume rendering that allows the exploration of our sense of touch [LCK*00].

References

[BW01] BRODLIE K., WOOD J.: Recent Advances in Volume Visualization. *Computer Graphics Forum*, 2001. 20(2): p. 125.

[C02] CRESSIE N.: Statistics for Spatial Data. 1993, New York: John Wiley & Sons, Inc.

[DKLP01] DJURCILOV S., KIM K., LERMUSIAUX P. F. J., PANG A.: Volume rendering data with uncertainty information, in Data Visualization 2001, D. Ebert, J.M. Favre, and R. Peikert, Editors. 2001, *Springer-Verlag Wien: Vienna*.

[DKLP02] DJURCILOV S., KIM K., LERMUSIAUX P., PANG A.: Visualizing scalar volumetric data with uncertainty. *Computers & Graphics*, 2002. 26(2): p. 239–248.

[E11] ESRI. ArcGIS: A Complete Integrated System.

http://www.esri.com/ software/arcgis

[accessed: February 5, 2011].

[G11] GEOSOFT. Oasis montaj Features.

http://www.geosoft.com/pinfo/oasis montaj/ keyfeatures.asp.

[accessed: February 5, 2011].

[GDAS10] GONÇALVES V., DIAS P., ALMEIDA F., SANTOS B. S.: Exploring new ways of integration, visualization and interaction with Geotechnical and Geophysical Data. *Proc. IV10: Applications of Information Visualization, IEEE Computer Society*, London, July, 2010, pp. 181–185.

[GKD*10] GUOLIANG Z., KELEI X., DONGMEI H., CHENG S., JIE S.: The comparison and study of small sample data spatial interpolation accuracy. *Natural Computation (ICNC)*, 2010 Sixth International Conference on. 2010.

[GR02] GRIGORYAN G., RHEINGANS P.: Probabilistic surfaces: point based primitives to show surface uncertainty, Proceedings of the conference on Visualization '02. 2002, IEEE.

[HB04] HEARN D., BAKER M. P., *Computer Graphics with OpenGL*. Third Edition ed. 2004: Pearson Education, Inc.

[LCK*00] LECUYER A., COQUILLART S., KHEDDAR A., RICHARD P., COIFFET P.: *Pseudo-Haptic Feedback:* Can Isometric Input Devices Simulate Force Feedback?. *Proceedings of the IEEE Virtual Reality 2000 Conference*. 2000, IEEE Computer Society. p. 83.

[lXLyX06] LIN XIAO M., LI J., YU XIU X.: 3D Visualization technology Based on BCC-Grid Shear-Warp Algorithm. in Computational Engineering in Systems Applications, IMACS Multiconference on. 2006.

[L05] LUÍS J.: Geofísica Ambiental, Método Eléctricos, Engenharia do Ambiente, Universidade do Algarve (Portugal), Capítulo 3, 25. 2005: pp. 1–4.

[MKCY09] MANAGULI R., KARADAYI K., CANXING X., YONGMIN K.: Volume rendering algorithms for three-dimensional ultrasound imaging: Image quality and real-time performance analysis. *Ultrasonics Symposium (IUS)*, 2009.

[MTB03] MCGUFFIN M. J., TANCAU L., BALAKRISHNAN R.: Using deformations for browsing volumetric data. *Proceedings of the 14th IEEE Visualization* 2003 VIS'03). 2003, IEEE Computer Society. pp. 401–408.

[P01] PANG A.: Visualizing Uncertainty in Geo-spatial Data. *In Proceedings of the Workshop on the Intersections between Geospatial Information and Information Technology*, 2001.

[PSAR93] PARROTT R. W., STYTZ M. R., AMBURN P., ROBINSON D.: Towards statistically optimal interpolation for 3D medical imaging. *Engineering in Medicine and Biology Magazine, IEEE*, 1993. 12(3): p. 49–59.

[RH06] ROPINSKI T., HINRICHS K.: Interactive Volume Visualization Techniques for Subsurface Data. *Visual Information and Information Systems*, S. Bres and R. Laurini, Editors. 2006.

[S07] SPENCER R.: Information Visualization: Design for Interaction. 2nd ed. 2007, London: Prentice Hall.

[SBB08] STREIT A., BINH P., BROWN R.: A Spreadsheet Approach to Facilitate Visualization of Uncertainty in Information. *Visualization and Computer Graphics*, IEEE Transactions, 2008.

[SMML98] SCHROEDER W., MARTIN K., MARTIN K., LORENSEN B.: The visualization toolkit, *An Object Oriented Approach to 3d Graphics*, 2nd ed. 1998: Prentice Hall.

[SSM11] SILVA S., SANTOS B. S., MADEIRA J.: Using color in visualization: A survey. *Computers & Graphics*, 2011. 35(2): p. 320–333.

[ZC06] ZUK T., CARPENDALE S.: Theoretical analysis of uncertainty visualizations. *Proc. of SPIE-IS&T Electronic Imaging*, SPIE Vol. 6060, 2006.

[ZPRA99] ZIMMERMAN D., PAVLIK C., RUGGLES A., ARMSTRONG M.: An Experimental Comparison of Ordinary and Universal Kriging and Inverse Distance Weighting. *Mathematical Geology*, 1999. 31(4): p. 375–390