Analysis of information correlation in geological models

Miguel de la Varga & Florian Wellmann
The Aachen Institute for Advanced Study in Computational Engineering Science, RWTH Aachen University.
Schinkelstrae 2, 52062, Aachen, Germany. Email: wellmann@aices.rwth-aachen.de

Abstract

The quantification and analysis of uncertainties is important in all cases where maps and models of uncertain properties are the basis for further decisions. Once these uncertainties are identified, the logical next step is to determine how they can be reduced. Information theory provides a framework for the analysis of spatial uncertainties when different subregions are considered as random variables. In the work presented here, joint entropy, conditional entropy, and mutual information are applied for a detailed analysis of spatial uncertainty correlations. The aim is to determine (i) which areas in a spatial analysis share information, and (ii) where, and by how much, additional information would reduce uncertainties. As an illustration, a typical geological example is evaluated: the case of a subsurface layer with uncertain depth, shape and thickness. Mutual information and multivariate conditional entropies are determined based on multiple simulated model realisations. Even for this simple case, the measures not only provide a clear picture of uncertainties and their correlations but also give detailed insights into the potential reduction of uncertainties at each position, given additional information at a different location. The methods are directly applicable to other types of spatial uncertainty evaluations, especially where multiple realisations of a model simulation are analysed. In summary, the application of information theoretic measures opens up the path to a better understanding of spatial uncertainties, and their relationship to information and prior knowledge, for cases where uncertain property distributions are spatially analysed and visualised in maps and models.

Considering Geological Modeling as an Inference Step

In order to perform the analysis of uncertainty, we need to embed the 3D geological modeling process into a probabilistic framework. The foundation of the method, here presented, consists of treating the generation of a geological model as a deterministic step within a larger Bayesian network. This consideration enables the use of complex data types or hypotheses that otherwise would not be possible to relate with the prior parameters. To create the models, we use here a cokriging interpolation of an implicit potential-field method, implemented in software package GeoModeller.

We follow in this work a novel approach to consider 3-D structural geological modelling not as a purely deterministic step, but as an inference step where we aim to consider all available geological and geophysical information. In order to achieve this aim, we develop a Bayesian inversion scheme in the typical way (Eq. 1):

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta} \qquad (1)$$

From a view point of geological modeling, we can interpret the different terms as follows:

• Priors, $p(\theta)$, define the belief for the multidimensional parametric space that encapsulates all the information of the system.

• Likelihood functions, $p(y|\theta)$, reduce model uncertainties by constraining it with observations. Furthermore, this step enables us to identify parameter correlations in the model. This leads to a reduction of uncertainty, since the uncertainty space gets reduced. However, contradictory observation may lead to an increase of the overall information entropy of the system.

• The step of choosing a suitable model, $\mathcal{M}$, quickly increases in complexity in typical geological modeling applications. The use of interpolation methods in order to be able to describe certain data can be framed in a non-parametrical model. Furthermore, geological data used for interpolation is often based on empirical expert opinion. On top of this, most of the observations in geology proceed from various different sources so that more than one forward model must be chosen at the same time. Nevertheless, models are only mathematical tools to describe data based on the chosen priors and therefore do not contribute to change the information of the system.

Application to a geological model of a Greenstone Belt

We test these considerations here in an application to a geological model of the Sandstone Greenstone Belt in the Yilgarn Craton, Australia. We first consider geological parameters as independent random variables and generate stochastic realizations of input data sets, which are then used in the forward modeling step to create a geological model realization (Wellmann et al., 2010). We then discretize the resulting continuous model on a discrete mesh structure and calculate probabilities for each specific geological layer in a given cell, and finally the information entropy in each cell as a summary measure of uncertainty. The information entropy if all cells with entropy values > 0 is presented in Fig. 1A.

In a next step, we now add information in the form of geologically motivated likelihood functions. These functions can partly be applied as constraints on the input data set (e.g. as constraints on the spatial difference between two geological surface contact points that define a layer thickness). In other cases, these constraints require a full forward modeling of the geological model, for example when using observations of a geological unit at a defined location. It is important to point out here that these types of constraints can not be used directly in the interpolation step, and that they are therefore incorporated in the form of likelihood functions. With these considerations, the remaining subsurface uncertainties are clearly reduced (Fig. 2B).

Finally, we use an additional type of information that is often available in these types of geological environments: potential-field measurements of gravitational acceleration. In order to obtain a likelihood function, we couple the geological modeling step with a model of gravity on the basis of rock densities assigned to each geological unit. The calculated gravity field is then compared to a gravity measurement. This additional type of information reduces the resulting uncertainties even further (Fig. 3C).

Information entropy as a measure of geological uncertainty

In our work, we develop methods to characterize and analyze uncertainties related to the identification of major geological features in the subsurface, such as structural geological elements (main boundaries between geological units, faults, folding patterns) or sedimentary features (channels, fans, etc.). Our knowledge about these features at a given location is typically limited as direct observations at depth in boreholes are rare and expensive. Model interpolations between these points carry furthermore many assumptions (e.g. Jessell et al., 2014). We therefore investigate how uncertainties in observations and the following modeling methods propagate into the space of the geological model.

We first consider geological parameters as independent random variables and generate stochastic realizations of input data sets, which are then used in the forward modeling step to create a geological model realization (Wellmann et al., 2010). We then discretize the resulting continuous model on a discrete mesh structure and calculate probabilities for each specific geological layer in a given cell, and finally the information entropy in each cell as a summary measure of uncertainty (Mann, 1993; Wellmann and Regenauer-Lieb, 2012).

Analysis of information correlation and uncertainty reduction

The logical next step from quantifying and visualizing uncertainties is the question where additional information would reduce uncertainties. To address this question, we apply the concept of multivariate conditional entropy: we want to know the remaining amount of uncertainty at a specific location, given the information (observations) at multiple other locations. We tested this measure before on simple systems with the expected result that spatial correlations lead to larger reductions of uncertainty when measurements are taken.
Next steps will therefore be to apply this concept to the presented Greenstone model. Furthermore, we attempt to characterize the entire uncertainty in the model space calculating the joint entropy of all cells. This step is computationally demanding, but could provide a way to consider all uncertainties, including their correlations, in the full 3-D space of a geological model.

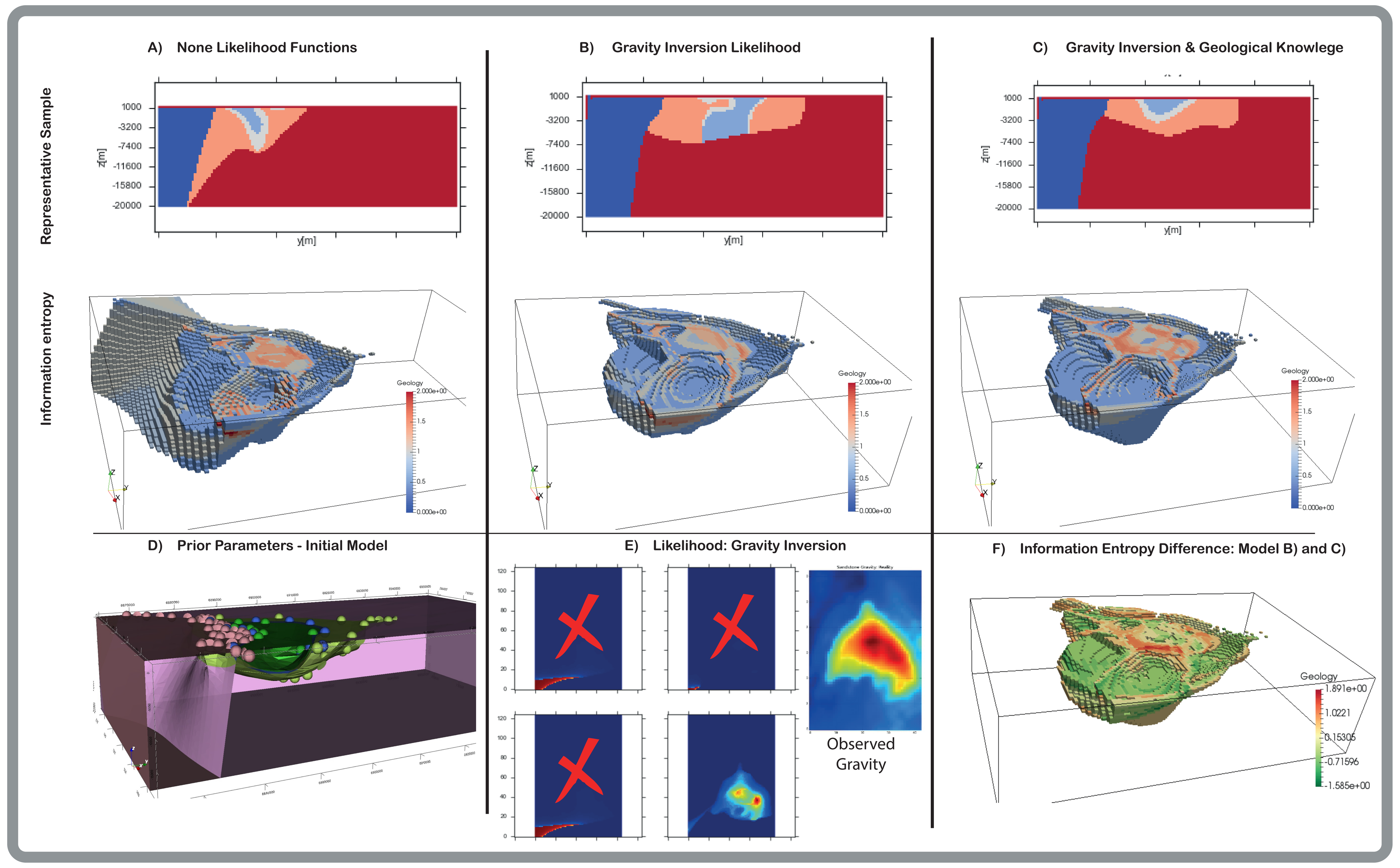

Figure 1: Uncertainty analysis of the Greenstone Belt geological model. A), B), C), show a representative section of one stochastic model as well as the information entropy of the model, for different observations. D) displays the initial 3D model and its prior parameters. E) describes the rejection of forward gravity inversion models making use of a likelihood function. F) shows the difference on information entropy increasing the number of observations.

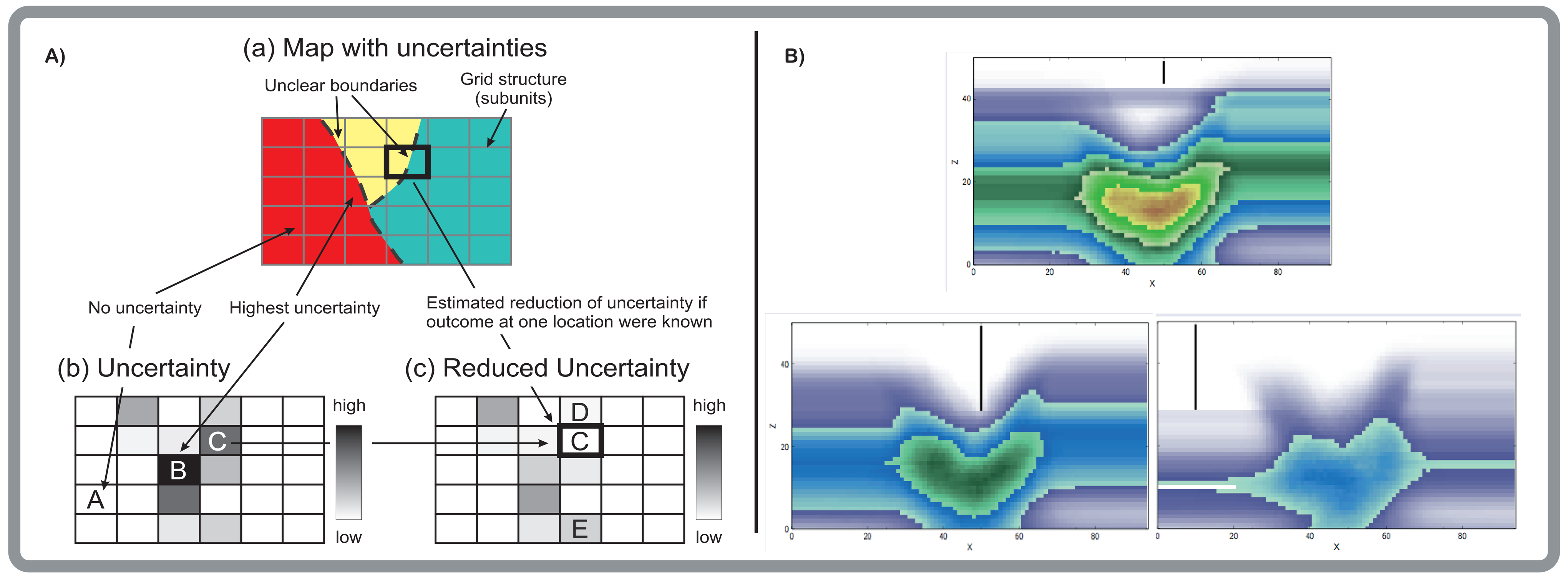

Figure 2 A: Theoretical illustration of information entropy variation in correlated cells. Figure 2 B: Evolution of information entropy of a model as we reduce uncertainty in specific regions, simulating new boreholes.

References

• Jessell, M., Ailleres, L., de Kemp, E., Lindsay, M., Wellmann, J. F., Hillier, M., et al. (2014). Next Generation Three-Dimensional Geologic Modeling and Inversion. Societyof Economic Geologists Special Publication 18, 18, 261–272.
• Mann, J. C. (1993). Uncertainty in geology. In Computers in geology---25 years of progress (pp. 241–254). Oxford University Press, Inc.
• Wellmann, J. F., Horowitz, F. G., Schill, E., & Regenauer-Lieb, K. (2010). Towards incorporating uncertainty of structural data in 3D geological inversion. Tectonophysics 490(3-4), 141–151. http://doi.org/doi: 10.1016/j.tecto.2010.04.022
• Wellmann, J. F., & Regenauer-Lieb, K. (2012). Uncertainties have a meaning: Information entropy as a quality measure for 3-D geological models. Tectonophysics, 526-529, 207–216.