

VIII. CONCLUSIONS

In this work, I developed several efficient numerical implementations of stochastic analysis methods for the spatiotemporal mapping of environmental processes, with a particular interest on applications related to water resources, ground water modelling and environmental health.

Environmental processes are challenging to analyze and model because they exhibit a high randomness in space and time, and they are characterized by a wide source of information including natural laws, empirical models and uncertain information. Traditional kriging methods lack the flexibility and mathematical rigor to account for this information, and the classic approach of solving sets of PDEs describing physical laws by means of the Finite Difference Methods (FDM), or Finite Element Methods (FEM) are not practical for the large scale and high heterogeneity of natural processes. Hence, there is a need to investigate new approaches, exploring both new theoretical formulation and efficient numerical implementation.

When this work was started, the theoretical framework for two promising methods had been developed, but, to the best of my knowledge, none had been successfully implemented numerically. The first method, the Bayesian Maximum Entropy (BME) approach, offered a general framework for spatiotemporal mapping, using a wide variety of information that the traditional kriging methods cannot not account for. Additionally the Space Transformation theory had been proposed as a novel approach to solve Partial Differential Equations (PDEs), which would allow to model physical laws for the complex subsurface flow problems encountered in natural heterogeneous systems.

A part of this work was spent on the numerical implementation of the Space Transformation (ST) method to solve the three-dimensional flow equation in heterogeneous porous media. The ST approach allows to transform the complicated set of PDEs describing three-dimensional flow into a set of one-dimensional equations (i.e. Ordinary Differential Equations, or ODEs) which are much easier to solve. I implemented the method and tested it on a cubic flow domain with heterogeneous porous media. Simulated case studies showed that the method calculates accurate solutions that honor the boundary conditions, and lends itself for parallel implementation because the computational work is divided into independent tasks (the transformed one-dimensional equations) that may be solved on different processors. This successful implementation of the ST represents working progress that will be further improved with additional numerical and conceptual improvements.

In the context of stochastic analysis of flow in porous media, the ST approach allows to efficiently calculate the flow solution corresponding to several realizations of the porous media. The statistical description of the flow properties is obtained by taking statistical averages of several flow solutions --the so-called Monte Carlo approach. Hence, the ST method covers the traditional approach for incorporating physical laws into the larger problem of mapping and analysis of space/time processes in the Environmental Sciences. The main part of the mapping analysis is then performed using a stochastic estimation method which incorporates the statistical description about the process, together with specific measurements and uncertain information, to produce maps of estimated values describing the natural process. As already mentioned, traditional stochastic estimation methods lack the flexibility to incorporate uncertain knowledge, so that there is a need to pursue new methods. Hence, I devoted the main part of this work to the development and

implementation of the BME approach, a novel spatiotemporal estimation method of modern Geostatistics.

The BME method offers a rigorous framework for spatiotemporal analysis and mapping. Due to its epistemological background and mathematical rigor, the BME approach offers a considerable flexibility in incorporating various sources of physical knowledge. However the flexibility to include a wide variety of physical knowledge results in a method with a high numerical complexity. In this work I derived formulations of the BME approach with reduced numerical complexity, and corresponding to mapping situations of practical interest. The formulations lead to efficient numerical implementations that are tested using simulated case studies, and applied to several field studies.

The first mapping situation considered corresponds to a general knowledge consisting of the spatiotemporal mean and covariance function, and the specificatory knowledge includes both hard and soft data of the interval type. This is a case of considerable interest in spatiotemporal mapping applications where uncertain information may be described by means of intervals for the measured attribute. An efficient formulation is proposed and implemented in the program `BMEintEst`. Simulated case studies show that the method provides more accurate estimation than traditional kriging methods at a numerical cost that is surprisingly small on modern days computers.

In the second mapping situation presented, soft data of probabilistic type is considered. This type of soft knowledge is very general, allowing to incorporate any form of uncertain information in the spatiotemporal mapping analysis. The resulting implementation is referred to as `BMEpdfEst`, and simulated case studies show that when using a combination of hard and soft data this method gives results that are far superior

than any traditional kriging methods, at a numerical cost that is acceptable on modern days computers. The power of the numerical code developed is demonstrated in the Equus Beds case study, where uncertain information about water table elevations were incorporated in a rigorous mapping analysis, providing useful maps describing the depletion of ground water in the Equus Beds aquifer, which is an important source of water resources for the city of Wichita, Kansas.

In the third mapping situation considered, the formulations presented for interval and probabilistic soft data were extended to the case of vectorial mapping. Vectorial mapping considers information from a primary random field as well as any secondary random fields that are related. This mapping situation is addressed by including the cross-covariance between related random fields in the general knowledge. The formulation obtained for vector BME was implemented numerically, resulting in an efficient code for computational BME mapping. Computational BME mapping provides a useful tool in environmental health, because it allows to analyze the association between disease and exposure using soft (uncertain) information. This was illustrated in the mortality/temperature case study for North Carolina.

The computational BME formulations presented not only possess considerable flexibility regarding the appropriate choice of an estimate (depending on the mapping application considered), but also offer a rigorous and complete assessment of the estimation error (uncertainty analysis), as well. Indeed, instead of a simple measure of estimation error (e.g., the estimation error variance of geostatistics), BME offers a more complete uncertainty assessment in terms of BME confidence sets. This was illustrated in the fourth implementation of computational BME presented, which involved multi-point mapping. As

the numerical case studies illustrated, BME confidence sets are more informative and accurate than their traditional counter-part.

Finally, an interesting aspect of the formulation presented is that, under certain restrictions (only hard data are used, etc.) it reduces to well-known kriging methods. When these restrictions are relaxed (e.g., other sources of physical knowledge are considered, in addition to the hard data), the BME formulation provides a more general estimator that can be used in situations in which the kriging methods cannot. These BME features are of significant practical importance (e.g., they ensure that the classical geostatistics results are preserved as limiting cases of BME analysis at essentially the same numerical work; more informative and accurate estimation is achieved by incorporating additional sources of physical knowledge).

8.1. Suggestions for Future Work

The work presented has lead to numerical implementations that are of practical interest in stochastic analysis and mapping of environmental processes. Additionally this work has opened up exiting new prospects. Some of these important issues are listed below.

8.1.1. Space Transformation Theory for Bounded Flow Domain

In this work the Space Transformation (ST) method was successfully implemented to solve the three-dimensional flow equation in heterogeneous porous media. This successful implementation demonstrate that the method calculates accurate solutions and points to its promising use for parallel machines. The simplified formulation used in this work can further be improved by explicitly accounting for the boundary flow condition in the ST

theory. This will lead to additional terms that were neglected in the current implementation, and should lead to improvements in the accuracy and speed of the method.

8.1.2. BME Mapping of Data with Known Distribution

In many mapping situations the natural variable has a known distribution, usually obtained by plotting the histogram of available measurements. An approach that is often used in this case is to transform the data available such that the transformed data has a normal distribution. This approach may easily be implemented with the current BME formulation, by just adding a pre-processing stage where the data is transformed into a normally distributed variable, and a post-processing stage that would back-transform the posterior pdf (of the transformed variable) to the posterior of the original variable. The ability of back-transform the posterior pdf is a unique feature of the BME approach, so that the data transformation approach should lead to better results than with traditional kriging methods.

8.1.3. BME Mapping Using Statistical Moments of High Orders

In the computational BME formulation presented in this work, the general knowledge includes the mean and the covariance of the random field. Higher statistical moments may also be included in the analysis. The problem is then to find a prior pdf which maximizes entropy and verifies the high order statistical moments imposed. My experience from this work is that this is a difficult numerical problem, however the solution exists (under some conditions), and would be of great interest for many mapping situations.

8.1.4. BME Mapping Using Physical Laws

One of the most exciting areas that this work opened up is the computational implementation of a BME approach that would include knowledge about a physical law in the mapping analysis. This would be a tremendous interest in many scientific fields, and in particular in

stochastic ground water flow analysis, which was one of the application targeted by this work. As shown in Chapter III and Chapter VIII, the BME framework allows to incorporate information about a physical law by means of constraints that are added in the general knowledge. The work presented here was a necessary first step for any computational BME implementation, and it will offer a useful starting point for a BME approach including the knowledge of a physical law.