

# Goal-Oriented Surrogate Construction for Groundwater Problems

Steven A. Mattis, Technical University of Munich

**Key words:** uncertainty quantification, groundwater, and parameter estimation

## Introduction

Solving stochastic inverse problems for groundwater problems generally involves some form of finite sampling, which leads to a type of stochastic error in solutions. For computationally expensive models, which are common in hydrology, surrogate response surfaces are often employed to increase the number of samples used in approximating the solution. The result is generally a trade off in errors where the stochastic error is reduced at the cost of an increase in deterministic/discretization errors in the evaluation of the surrogate. Such stochastic errors pollute predictions based on the stochastic inverse. In this work, we formulate a method for adaptively creating a special class of surrogate response surfaces with this stochastic error in mind. Adjoint techniques are used to enhance the local approximation properties of the surrogate allowing the construction of a higher-level enhanced surrogate. Using these two levels of surrogates, appropriately derived local error identifiers are computed and used to guide refinement of both levels of the surrogates. Three types of refinement strategies are presented and combined in an iterative adaptive surrogate construction algorithm. Numerical examples for groundwater flow and transport demonstrate how this adaptive strategy allows for accurate predictions under uncertainty for a much smaller computational cost than uniform refinement.

## Surrogate Models and Error

For an observation map,  $Q : \mathbf{\Lambda} \rightarrow \mathcal{D}$ , let  $Q_s(\lambda)$  denote a computationally inexpensive surrogate approximation to  $Q(\lambda)$ , so

$$Q_s(\lambda) = Q(\lambda) + \epsilon_s(\lambda),$$

where  $\epsilon_s(\lambda)$  is the error in the surrogate. We use  $Q_s(\lambda)$  to more efficiently map large numbers of samples between  $\mathbf{\Lambda}$  and  $\mathcal{D}$ . However, constructing  $Q_s(\lambda)$  often requires using some particular set of samples of  $Q_h(\lambda)$  based on a specific type of sampling in  $\mathbf{\Lambda}$ , e.g., using a possibly different set of random samples or using deterministic sampling approaches such as sparse grids. We let  $Q_{s,h}(\lambda)$  denote this numerically constructed surrogate and  $\epsilon_{s,h}(\lambda)$  denote its error. We decompose the error as

$$\epsilon_{s,h}(\lambda) := \epsilon_s(\lambda) + \epsilon_h(\lambda),$$

where  $\epsilon_s(\lambda)$  is the error in the choice of surrogate due to limited approximation properties of the surrogate, and  $\epsilon_h(\lambda)$  is the error in the surrogate from numerical solution of the model.

The space of uncertain parameters  $\mathbf{\Lambda}$  can be discretized by a Voronoi tessellation simply by sampling the space. Suppose that  $\{\lambda^{(i)}\}_{i=1}^N$  is a finite set of  $N$  distinct points in  $\mathbf{\Lambda}$  that we will call “samples.” This forms an implicit tessellation, i.e. the Voronoi cells do not have to be explicitly constructed. It is generally only necessary to identify which cell  $\mathcal{V}_{i,N}$  contains a point  $\lambda$  via a nearest neighbor search amongst  $\{\lambda^{(i)}\}_{i=1}^N$ . By forming piecewise polynomial approximations in the Voronoi cells defines a cheap surrogate model. The piecewise polynomial approximations are formed by evaluating the model at the sample points and solving adjoint problems to calculate derivative information. The adjoint solutions are also used to calculate a posteriori error estimates which can be used to create an error-enhanced surrogate model. The two levels of surrogates are used to calculate local error indicators based on a given problem of prediction under uncertainty. These indicators are used to adaptively refine the surrogate.

## Adaptive surrogate construction

The calculated local error indicators can be used to refine the surrogate to increase its accuracy and the accuracy for solving the stochastic inverse problem. We consider three types of refinement. The first type of adaptivity is *p-refinement*. In this type of refinement, the local polynomial order of the surrogate on a Voronoi cell is increased. This should decrease the local effect of  $\epsilon_s(\lambda)$ , error in the approximate QoI map due to the increase in quality of the surrogate model. The second type of refinement is *level-refinement*. In this type of refinement, the model level of the surrogate on a Voronoi cell  $\mathcal{V}_{i,N_k}$  is increased. This should decrease the local effect of  $\epsilon_h(\lambda)$ , error in the approximate QoI map due to numerical error in solving the model. A higher-level solve of the model should generally decrease the error in the approximate computations of QoI. The third type of refinement is *h-refinement*. In this type of refinement, new samples (generating points for the Voronoi tessellation) are added. Adding new samples should locally decrease both  $\epsilon_s(\lambda)$  and  $\epsilon_h(\lambda)$  on the approximate QoI map. The local

Taylor approximations become better because of the decreases radius and the numerical error is extrapolated less.

An iterative algorithm for performing these types of refinements to create improved surrogates is presented. The method is applied to several hydrological applications. These applications include parameter estimation and prediction under uncertainty for groundwater flow and contaminant transport in an aquifer. These problems involve computationally expensive forward models and possibly high-dimensional parameter spaces. The method is shown to be effective for such cases.

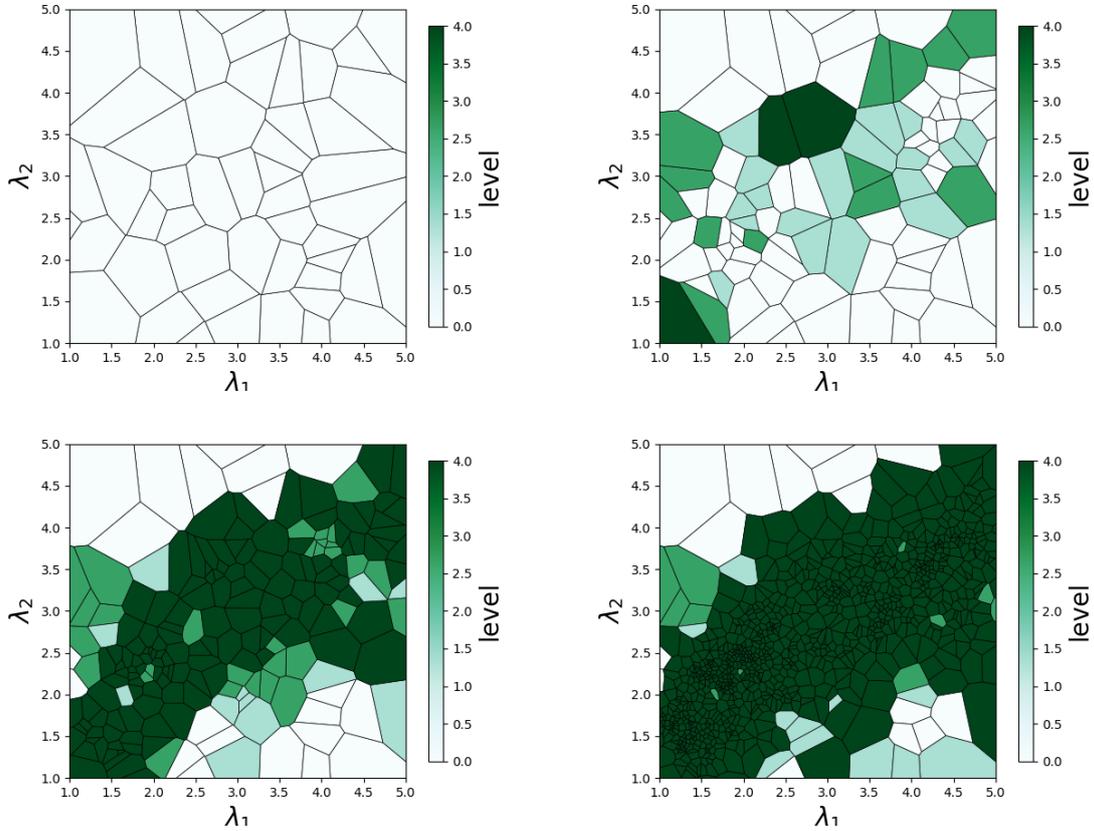


Figure 1: Four iterations of the adaptive method for an idealized 2D parameter space.

## References

- [1] Ahmad A. Rushdi, Laura P. Swiler, Eric T. Phipps, Marta D’Elia and Mohamed S. Ebeida. VPS: Voronoi Piecewise Surrogate models for high-dimensional data fitting. *International Journal for Uncertainty Quantification.*, **7:1**, 1–21, (2017)
- [2] Steven A. Mattis, Troy D. Butler, Clint N. Dawson, Don Estep and Velimir V. Vessilinov. Parameter estimation and prediction for groundwater contamination based on measure theory, *Water Resources Research.* **51**, 7608–7629, (2015)