

EGU2020-1355, updated on 25 Oct 2020

<https://doi.org/10.5194/egusphere-egu2020-1355>

EGU General Assembly 2020

© Author(s) 2020. This work is distributed under the Creative Commons Attribution 4.0 License.



HER: an information theoretic alternative for geostatistics

Stephanie Thiesen¹, Diego Vieira^{2,3}, Mirko Mälicke¹, Florian Wellmann⁴, and Uwe Ehret¹

¹Karlsruhe Institute of Technology, Institute of Water Resources and River Basin Management, Germany
(stephanie.thiesen@kit.edu)

²Department for Microsystems Engineering, University of Freiburg, Freiburg, Germany

³Bernstein Center Freiburg, University of Freiburg, Freiburg, Germany

⁴Computational Geosciences and Reservoir Engineering, RWTH Aachen University, Aachen, Germany

Interpolation of spatial data has been considered in many different forms. This study proposes a stochastic, non-parametric, geostatistical estimator that combines measures of information theory with probability aggregation method. Histogram via entropy reduction (HER) can be used to analyze the data spatial correlation and for predicting distributions at unobserved locations directly based on empirical probability. The method minimizes estimation uncertainty, relaxes normality assumptions and therefore avoids the risk of adding information not available in data (or losing available information). In particular, the applied probability aggregation method provides a proper framework for uncertainty estimation that reflects both the spatial configuration of the data as well as data values, while allowing to infer (or introduce) physical properties (continuous or discontinuous characteristics) from the field under study. Three different aggregation methods were explored in terms of uncertainty, resulting in predictions ranging from conservative to more confident ones. We investigate the performance of the framework using four synthetically generated datasets from known Gaussian processes and demonstrate the efficacy of the method in ascertaining the underlying true field with varying sample sizes. By comparing the method performance to popular benchmark models, namely nearest neighbors (NN), inverse distance weighting (IDW) and ordinary kriging (OK), we were able to obtain competitive results with respect to OK, with the advantage of presenting generalization properties. The novel method brings a new perspective of spatial and uncertainty analysis to geostatistics and statistical learning, using the lens of information theory.