Assessing local and spatial uncertainty with nonparametric geostatistics

Stephanie Thiesen (stephanie.thiesen@gmail.com)
Karlsruhe Institute of Technology Faculty of Civil Engineering Geo- and Environmental Sciences: Karlsruher Institut fur Technologie Fakultat fur Bauingenieur- Geo- und Umweltwissenschaften
https://orcid.org/0000-0001-7501-2223

Uwe Ehret
Karlsruhe Institute of Technology Faculty of Civil Engineering Geo- and Environmental Sciences: Karlsruher Institut fur Technologie Fakultat fur Bauingenieur- Geo- und Umweltwissenschaften

Research Article

Keywords: Nonparametric geostatistics, Non-Gaussian, Conditional distribution, Sequential Simulation, Uncertainty analysis, Risk mapping

Posted Date: March 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-272229/v1

Version of Record: A version of this preprint was published at Stochastic Environmental Research and Risk Assessment on July 15th, 2021. See the published version at https://doi.org/10.1007/s00477-021-02038-5.

Abstract

Uncertainty quantification is an important topic for many environmental studies, such as identifying zones where potentially toxic materials exist in the soil. In this work, the nonparametric geostatistical framework of histogram via entropy reduction (HER) is adapted to address local and spatial uncertainty in the context of risk of soil contamination. HER works with empirical probability distributions, coupling information theory and probability aggregation methods to estimate conditional distributions, which gives it the flexibility to be tailored for different data and application purposes. To explore the method adaptation for handling estimations of threshold-exceeding probabilities, it is used to map the risk of soil contamination by lead in the well-known dataset of the region of Swiss Jura. Its results are compared to indicator kriging (IK) and to an ordinary kriging (OK) model available in literature. For the analyzed dataset, IK and HER achieved the best performance and exhibited comparable accuracy and precision of their predictions. When compared to IK, HER has shown to be a unique approach for dealing with uncertainty estimation in a fine resolution, without the need of modeling multiple indicator variograms, correcting order-relation violations, or defining interpolation/extrapolation of distribution. Finally, to avoid the well-known smoothing effect when using point estimations (this is the case with kriging, but also with HER) and to provide maps that reflect the spatial fluctuation of the revealed reality, we demonstrate how HER can be used in combination with sequential simulation to assess spatial uncertainty (uncertainty jointly over several locations).

Full Text

Due to technical limitations, full-text HTML conversion of this manuscript could not be completed. However, the latest manuscript can be downloaded and accessed as a PDF.

Figures

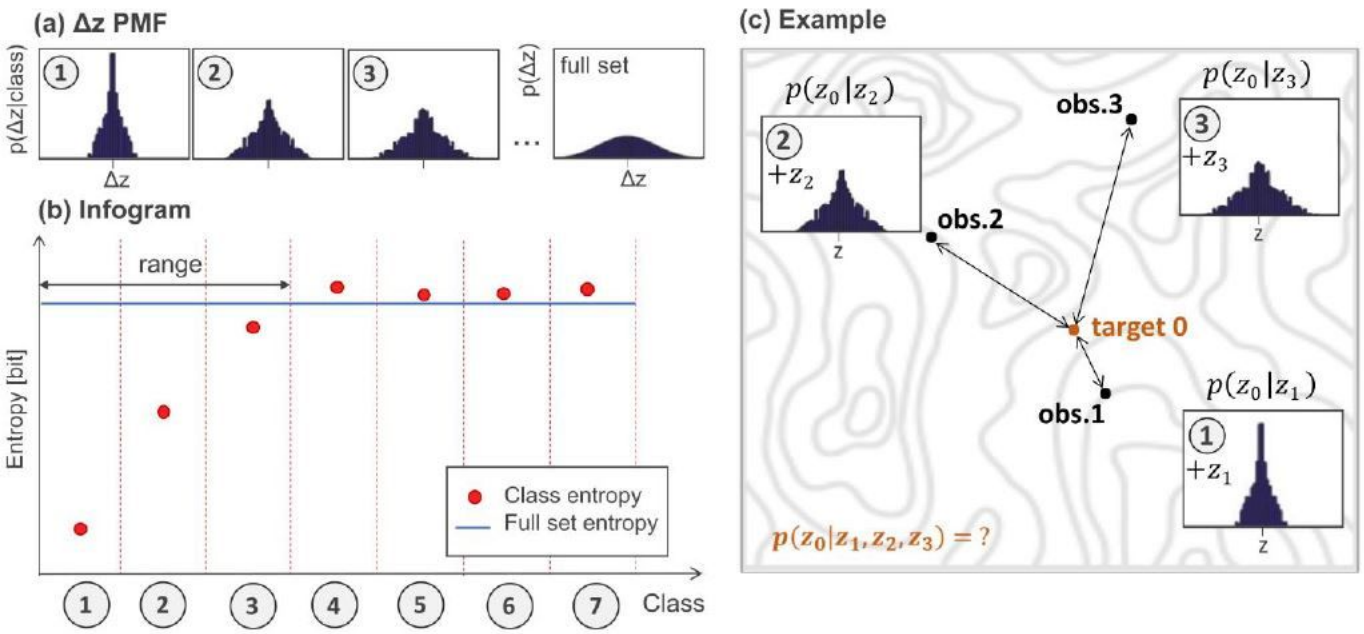

Figure 1

Characterization of spatial correlation. (a) Δz PMFs of the classes and full dataset, (b) infogram, and (c) practical example. 2.2.2 Probability aggregation

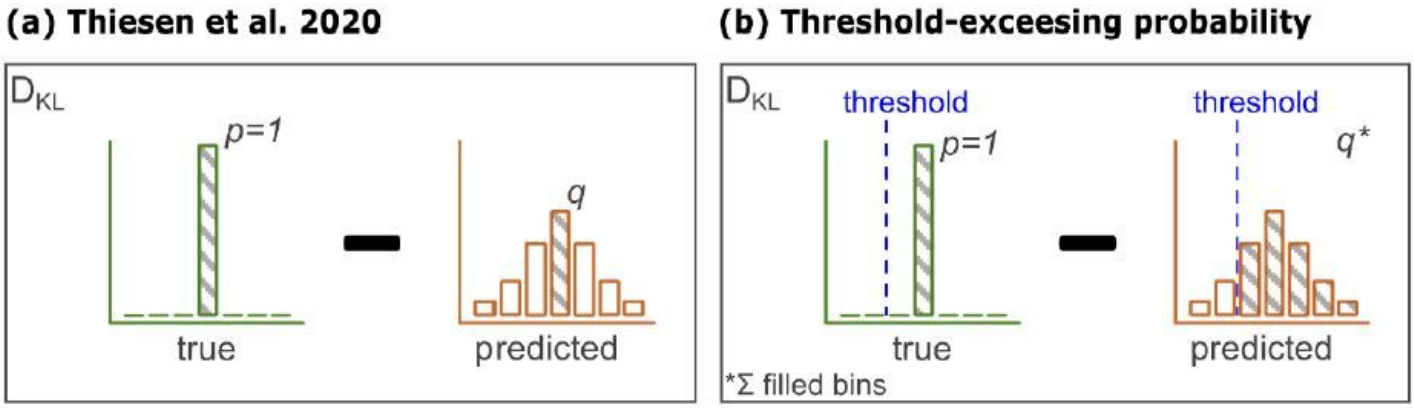

Figure 2

Optimization problem. (a) Maximizing the probability of the 'true' observation (Thiesen et al. 2020) and (b) maximizing the estimation of threshold-exceeding probability.

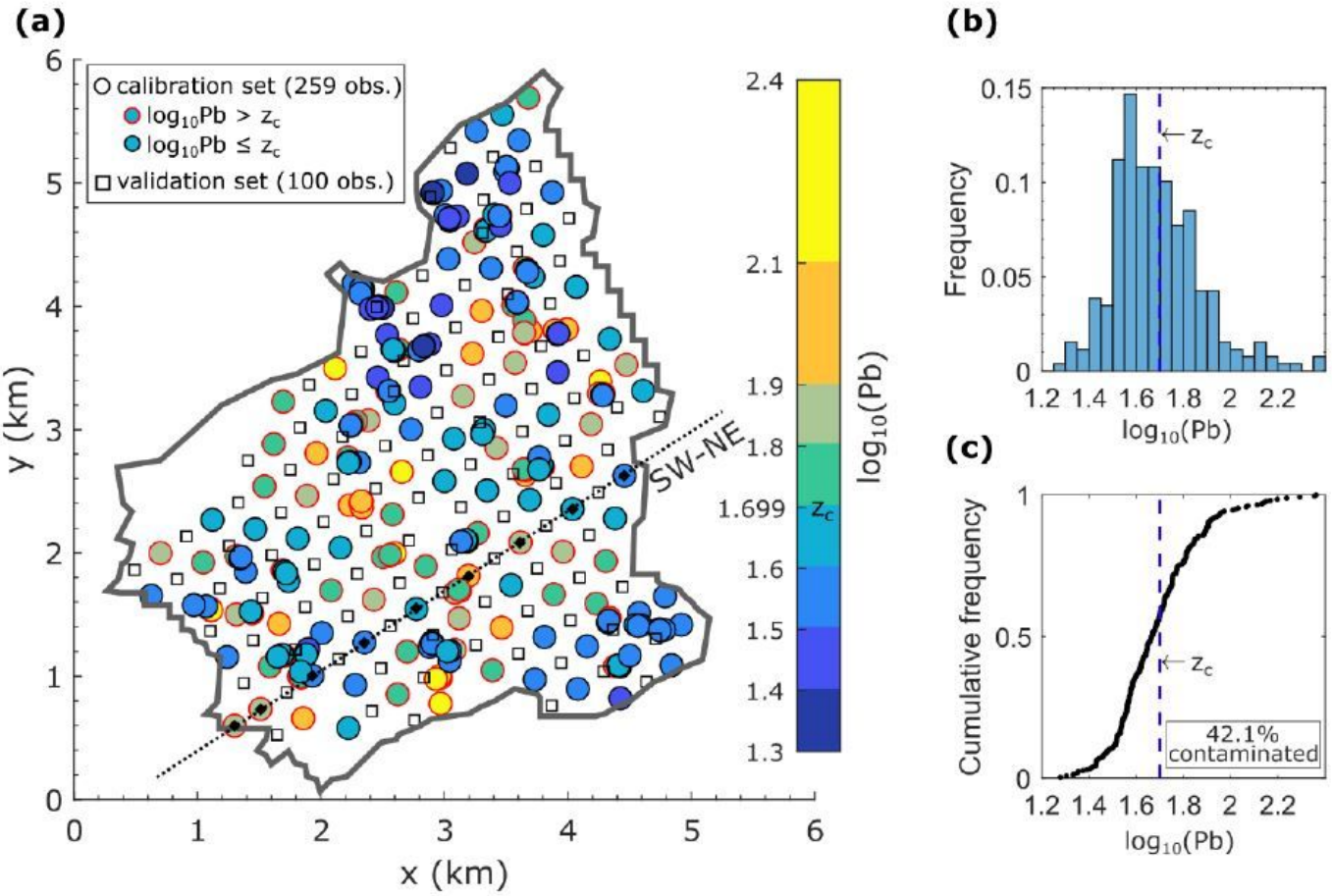

Figure 3

Logarithm of Pb of the calibration set. (a) Concentration, (b) histogram, and (c) cumulative distribution.

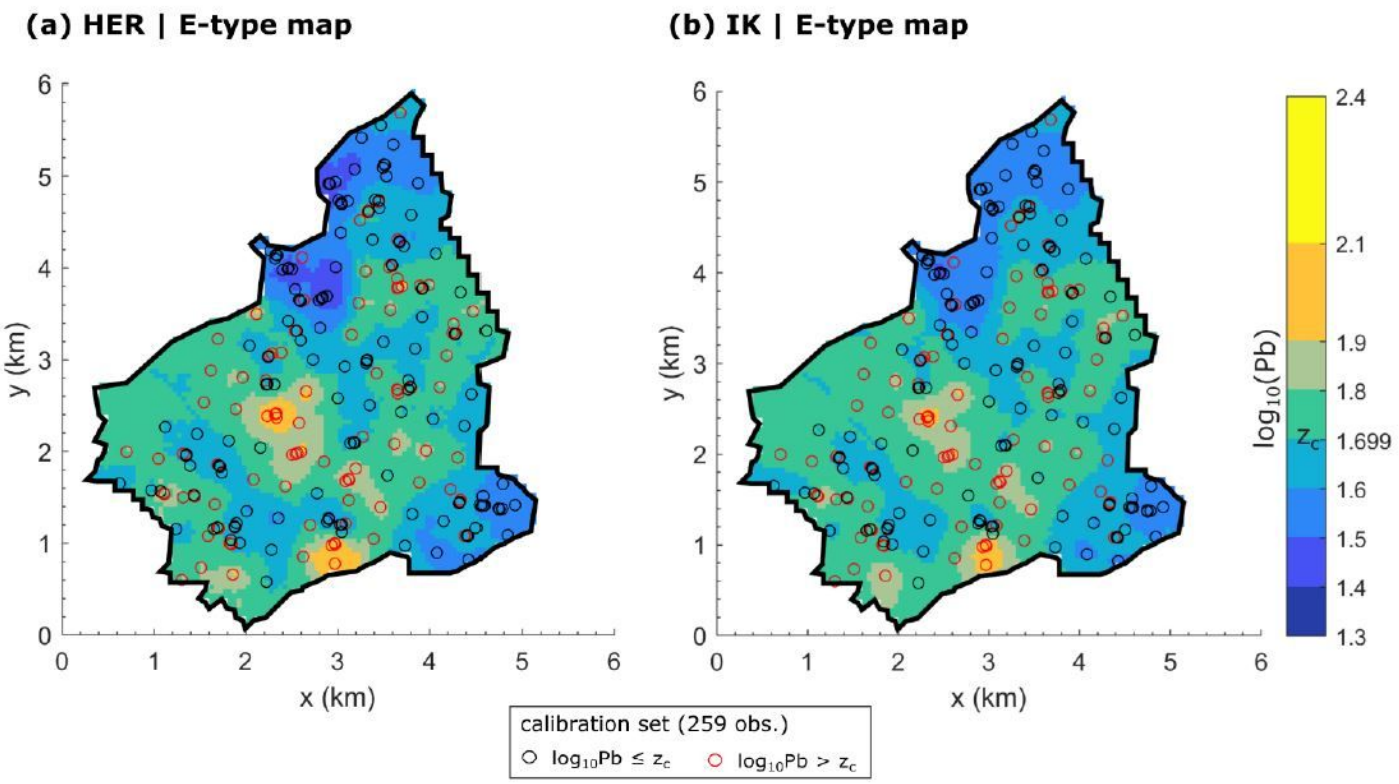

Figure 4

E-type map. (a) HER method, and (b) IK method.

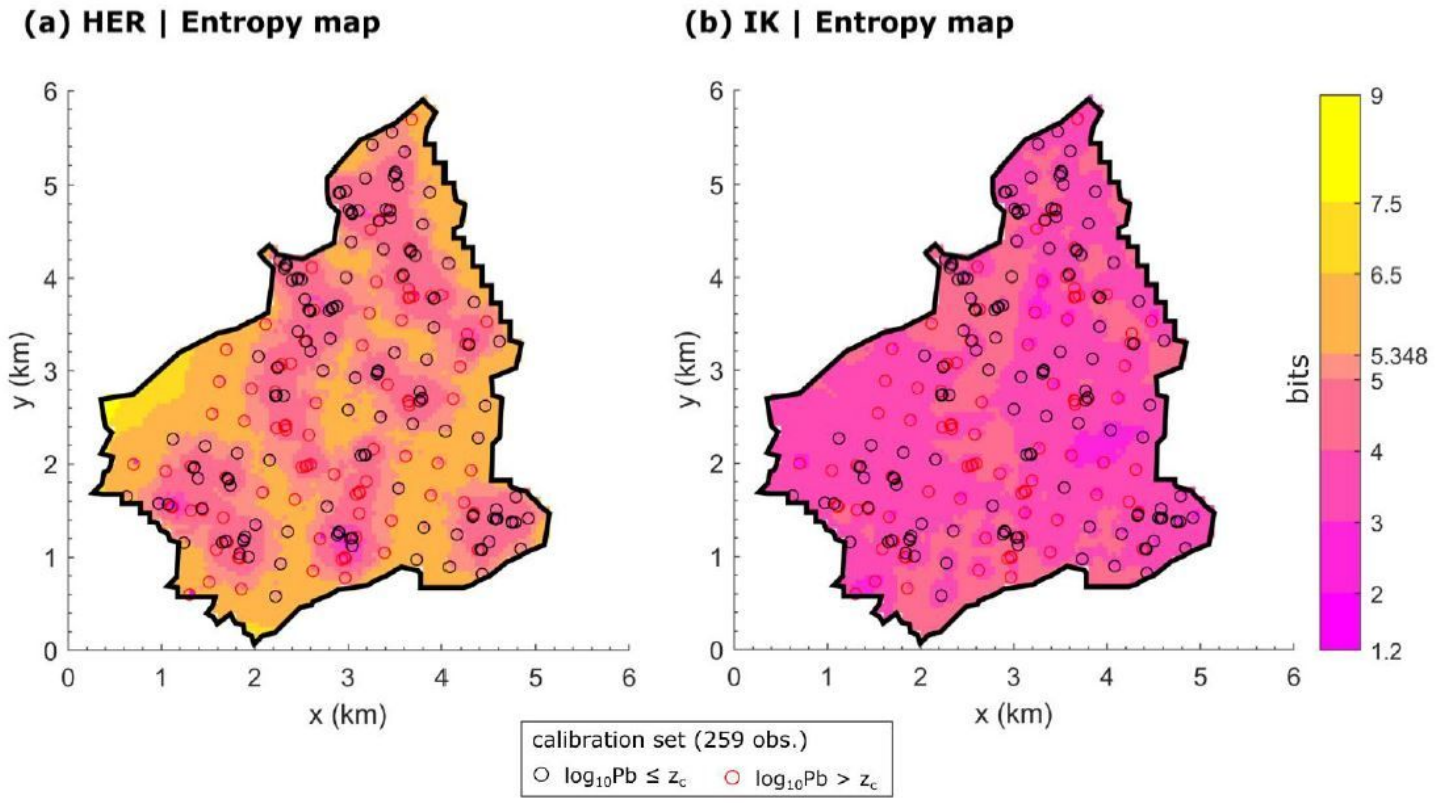

Figure 5

Local uncertainty via Entropy map. (a) HER method, and (b) IK method.

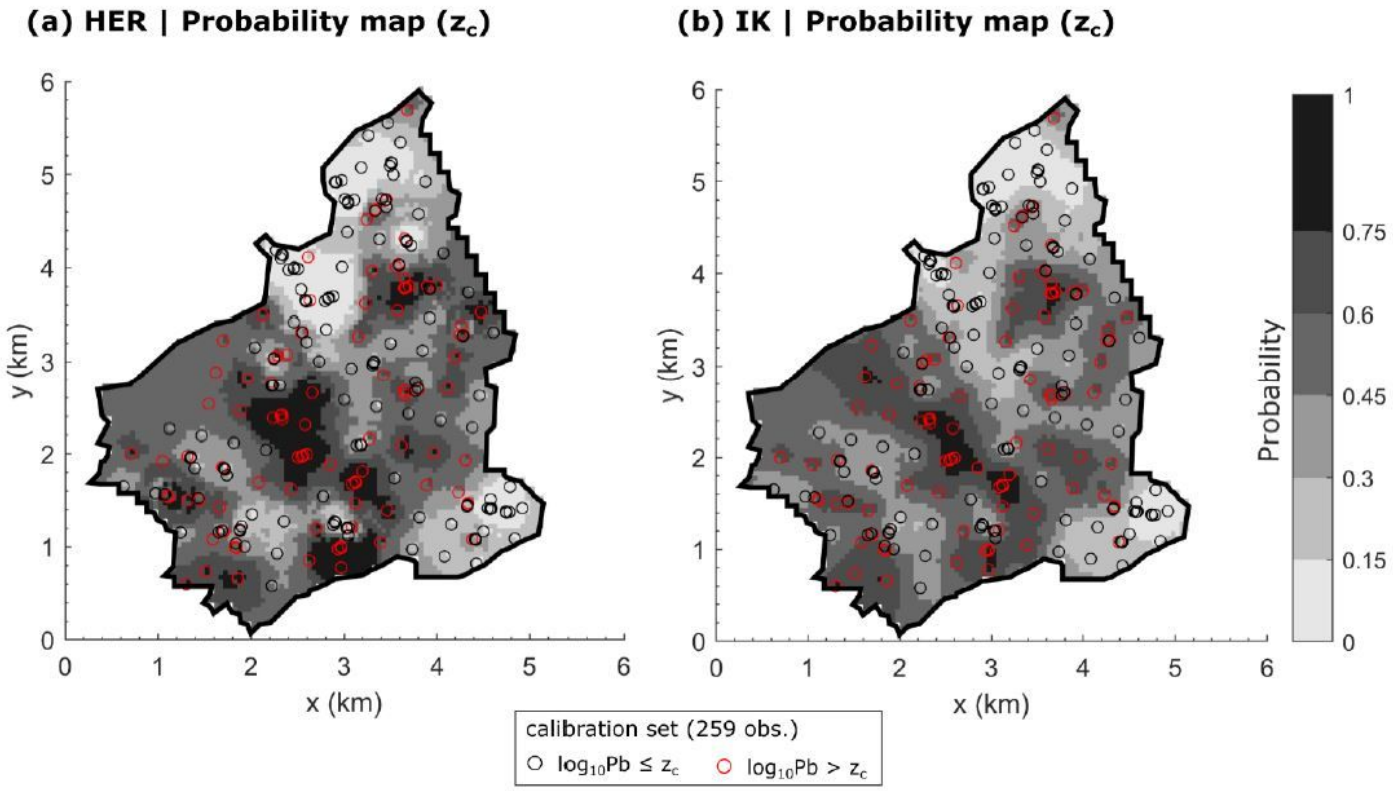

Figure 6

Probability of exceeding the critical threshold ($\boldsymbol{zc}$=$\mathbf{1.699}$). (a) HER method, and (b) IK method.

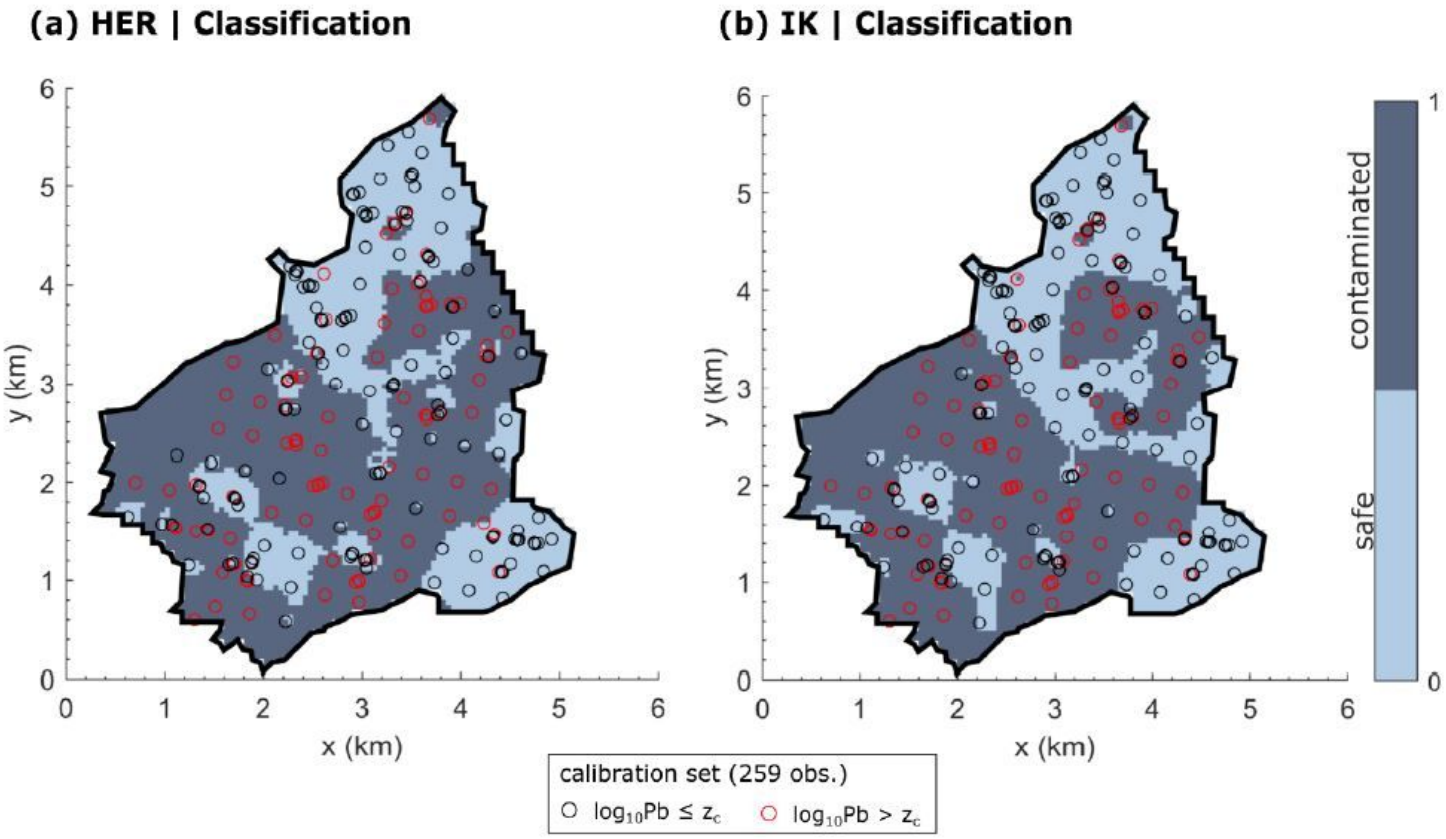

Figure 7

Classification of locations as contaminated by Pb on the basis that the probability of exceeding the critical threshold z_c is larger than the marginal probability of contamination (0.421). (a) HER method, and (b) IK method.

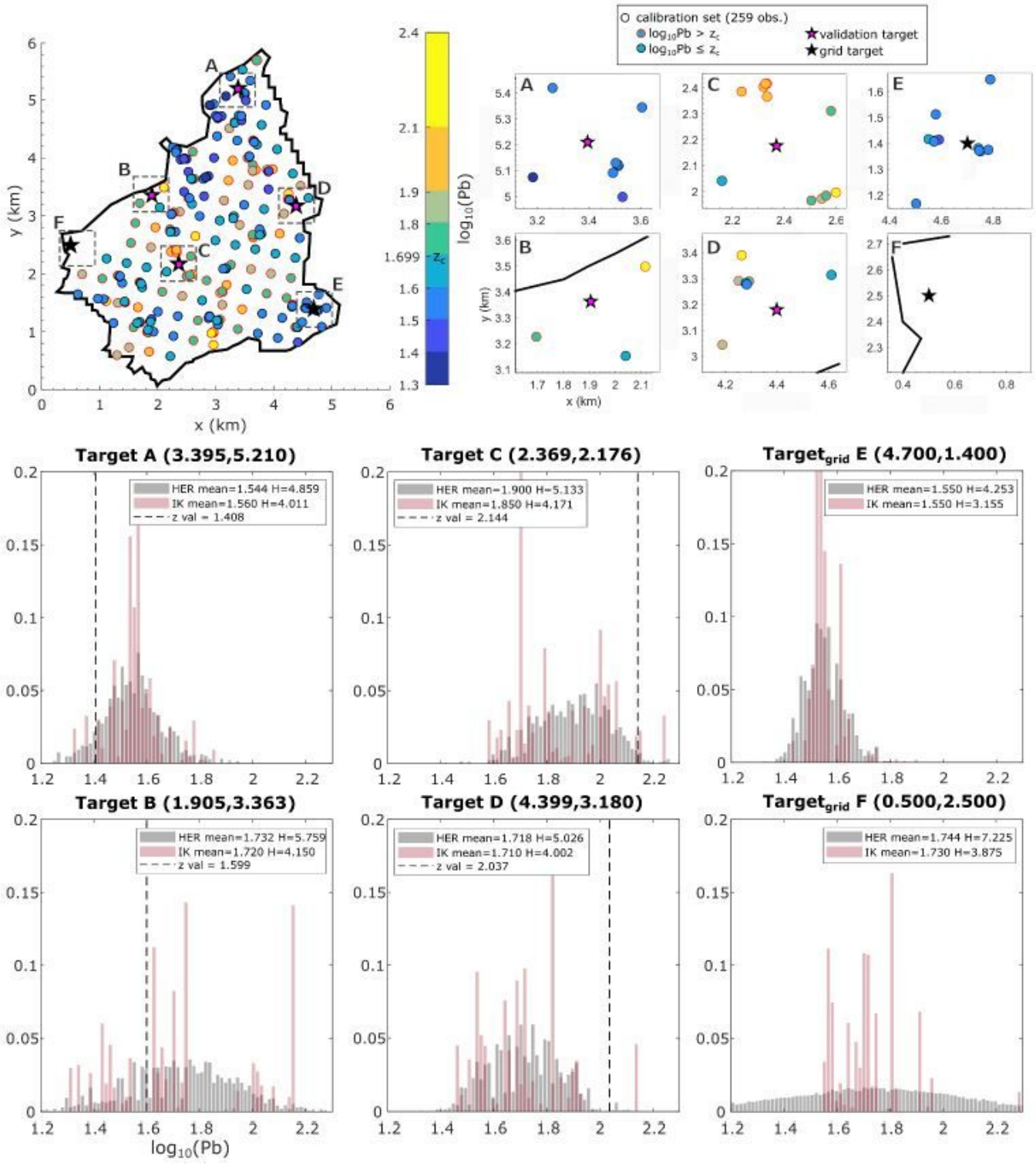

Figure 8

Local distribution of targets of the validation set (targets A to D) and grid (targets E and F) for HER (gray) and IK (red). Targets are identified by their coordinates (x,y). The location of each target is shown in a buffer of 600 m by 600 m.

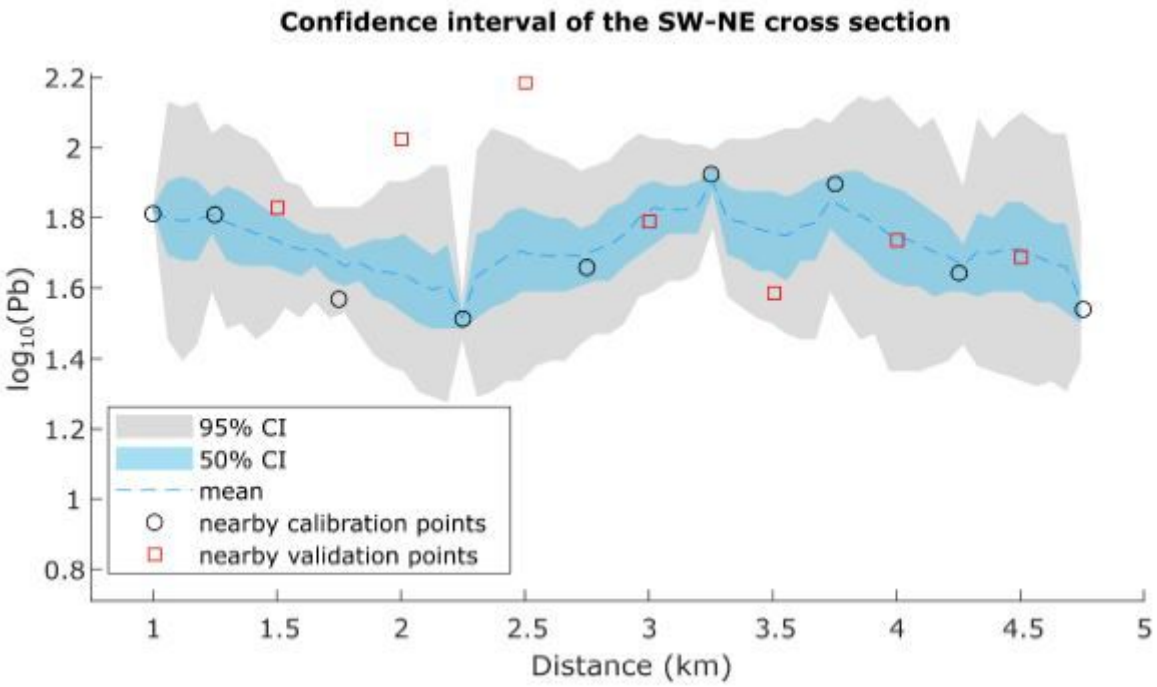

Figure 9

HER confidence interval (CI) of the SW-NE cross section (identified in Fig. 3a).

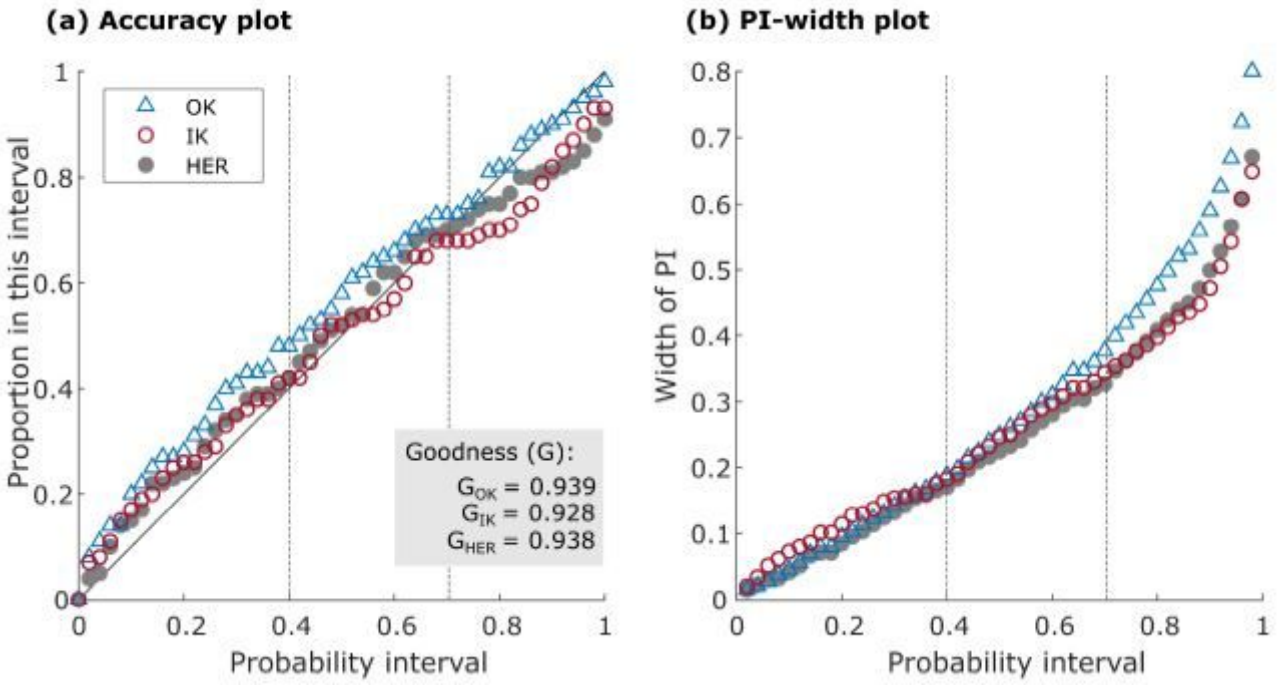

Figure 10

OK, IK, and HER performance. (a) Proportion of the true Pb values falling within the probability intervals ($\boldsymbol{p}$-PI) of increasing sizes, and (b) width of these intervals versus $\boldsymbol{p}$-PI. The goodness statistic ($\boldsymbol{G}$) quantify the similarity between the expected and observed proportions in the accuracy plots.

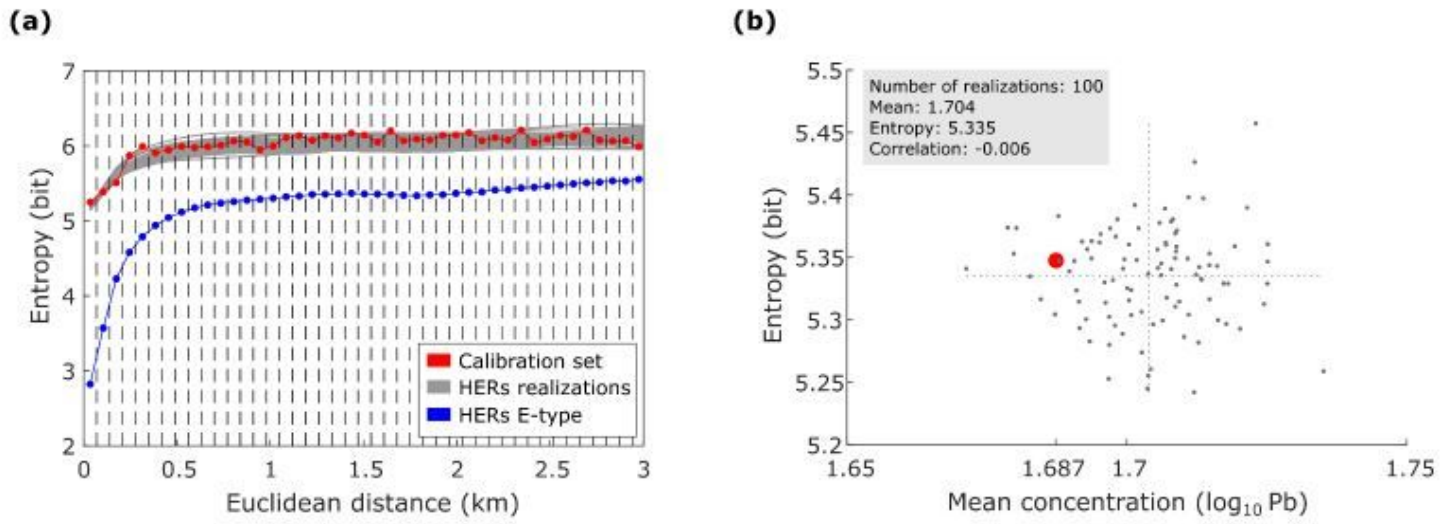

Figure 11

Ergodic fluctuations of 100 realizations generated with HERs. (a) Infogram and (b) scatterplot of the mean and entropy values.

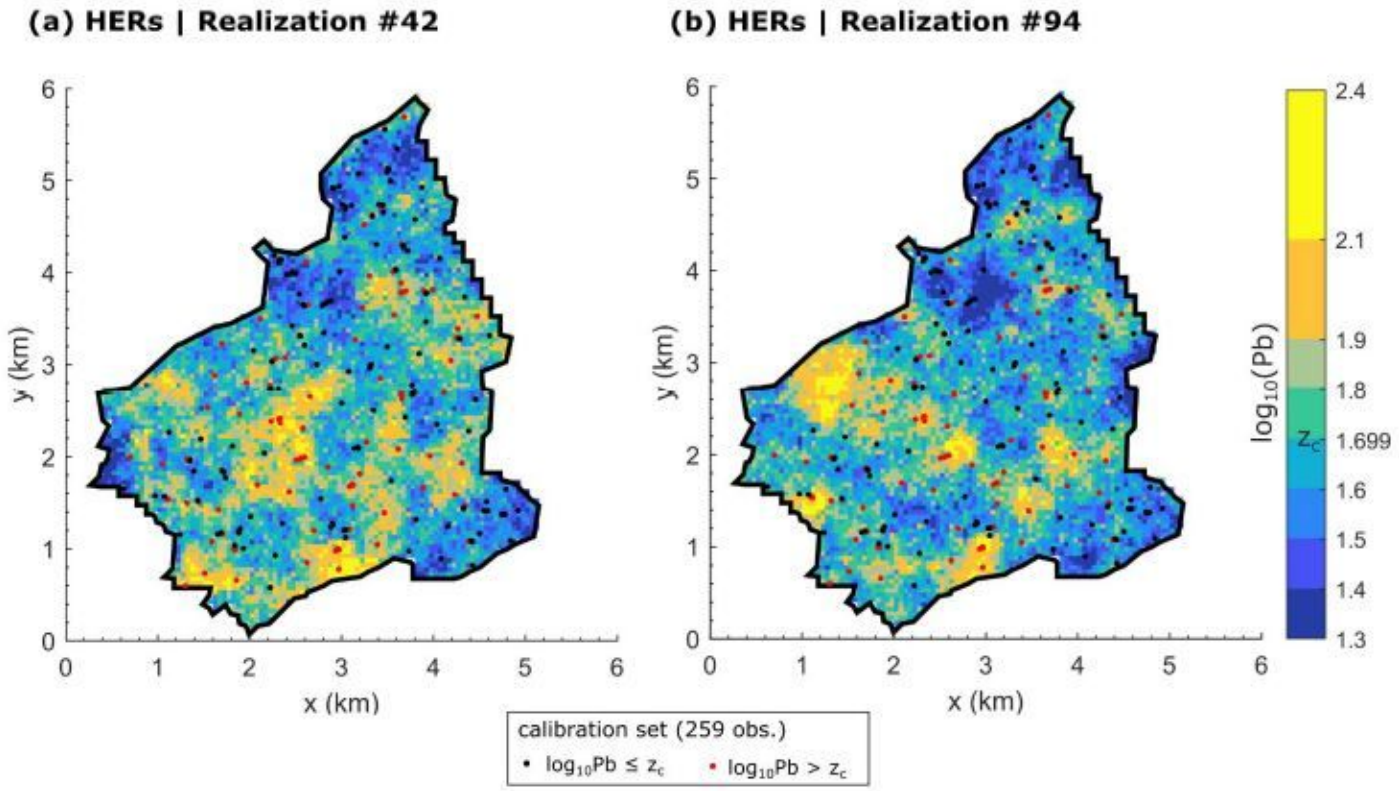

Figure 12

Realizations generated using HERs. (a) Realization #42 and (b) realization #94. Simulation grid size of 0.05 km x 0.05 km.