Sujet de thèse au laboratoire Heudiasyc UMR 7253 :

Title	Uncertainties and imprecision in Spatial Interpolations for Urban Risk Mapping
Supervising team	UTC : Sébastien Destercke, Benjamin Quost ; IRIT : Romain Guillaume, Hélène Fargier ; BRGM : Jérémy Rohmer, S. Belbeze, D. Guyonnet
Descriptif du sujet	Spatial data interpolation is a common operation in the field of geosciences. While the number of applications is numerous, an important observation is the low proportion of studies that estimate uncertainties (<5%). Uncertainties can be multiple, of different nature, and more specifically, the uncertainty due to the imperfection of knowledge (epistemic) can be significant in high-stakes applications in urban environments. A promising approach for an exhaustive and transparent consideration of uncertainties is the theory of imprecise probabilities (including the possibility theory, Dubois and Prade, 1988 and Dempster Shafer theory, Shafer 1976, Dempster 1967). This framework has foundations in classical probability theory and can be seen as a generalization of the Bayesian framework by providing an additional degree of flexibility to express different types of uncertainties. In the context of this PhD thesis (in partnership between UTC, IRIT and BRGM), we aim to explore two approaches. The first approach involves extending the classical framework and Bayesian approaches while working under the same assumptions. This approach may rely on the introduction of multiple priors (in the form of sets, Fig. 2b,c) instead of just one. This may either rely on the generalization of Gaussian processes as proposed by Mangili (2016) or on a post-processing approach like "conformal prediction" (Mao et al., 2022) by applying an order statistics based method to the kriging problem. This option is particularly promising in situations with sparse and/or clustered data. The second approach, which is particularly interesting for problems with imprecision and noise in measurements (or even the presence of outliers), aims to depart from classical assumptions by adopting a completely different viewpoint. This can be done either (i) by solving a fuzzy optimization problem with constraints (Dubois et al., 2014), or (ii) by adopting the information fusion perspective (Shinde et al., 2021), which has shown interesting properties in cases where the Bayesian model is misspecified. These developments will be applied to the case of estimating urban pedogeochemical backgrounds, particularly for Toulouse Métropole (Belbèze 2019), whose sparse, imprecise, clustered data/knowledge context is representative of a range of practical situations.

	The candidate will be located at UTC with expected travel to partner institutes (Orleans and Toulouse).
References	Belbeze, S., Djemil, M., Béranger, S., Stochetti, A. (2019). Détermination de FPGA – Fonds Pédo-Géochimiques Anthropisés urbains. Agglométation pilote : Toulouse métropole. Rapport finale BRGM/RP-69502-FR. Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. The Annals of Mathematical Statistics. 38 (2): 325–339. doi:10.1214/aoms/1177698950 Dubois D, Prade H (1988) Possibility Theory: An Approach to Computerized Processing of Uncertainty, Plenum, New York. Dubois, D., Fargier, H., Ababou, M., & Guyonnet, D. (2014). A fuzzy constraint-based approach to data reconciliation in material flow analysis. International Journal of General Systems, 43(8), 787-809. Mangili, F. (2016). A prior near-ignorance Gaussian process model for nonparametric regression. International Journal of Approximate Reasoning, 78, 153-171. Mao, H., Martin, R., & Reich, B. J. (2022). Valid model-free spatial prediction. Journal of the American Statistical Association, (just-accepted), 1-28. Shafer, G. (1976). A Mathematical Theory of Evidence, Princeton University Press Shinde, K., Feissel, P., & Destercke, S. (2021). Dealing with inconsistent measurements in inverse problems : set-based approach. International Journal for Uncertainty Quantification, 11(3).
Sought profile	- Training in applied mathematics : probability/statistics, machine learning, data science, optimisation… - Provable high skills in a coding language commonly used for scientific computing, such as Python, R, Octave, … - A taste for applied as well as theoretical studies - Sought skills : autonomy, open mindedness, teamwork, …
Contacts	UTC : Sébastien Destercke sebastien.destercke@hds.utc.fr IRIT : Romain Guillaume Romain.Guillaume@irit.fr BRGM : Jérémy Rohmer j.rohmer@brgm.fr