Post-doctoral position 10 month

Handing epistemic uncertainty in spatial interpolation problem

Post-doc duration: 10 months – expected starting date: September 2025

Location: IRIT, Cr Rose Dieng-Kuntz, 31400 Toulouse

Main supervisor: Romain Guillaume

This post-doctorate is part of the ANR HOUSES project, and the post-doc will work in collaboration with the other project partners, particularly BRGM.

Interpolation of spatially distributed data is a key element in geo-environmental sciences [1]. Despite the importance of these methods and their widespread use, an analysis of the literature [2] reveals major limitations when data are “SIC”, i.e. Sparse (i.e. small number of observations, e.g., ~100 over several 10km²), Imprecise (subject to measurement errors), and Clustered (heterogeneously distributed), which are commonly encountered in Earth sciences (see examples Figure 1 for geophysics, geology, soil pollution and natural hazards). While there are innumerable examples of its application, one important observation is the low proportion of studies proposing the estimation of uncertainties (<5%).

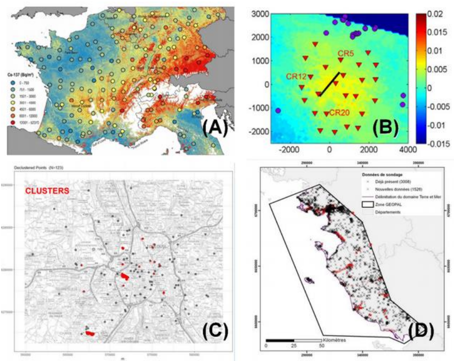

Figure 1. Illustration of SIC situations: (A) Sparse data of Cesium concentration [3]; (B) Sparse and clustered measurement locations for ground motion monitoring [4]; (C) Sparse and clustered data for geochemical background mapping in Toulouse city [5]; (D) Large number but clustered data for loose sediment thickness mapping in Pays de Loire [6].

Yet uncertainties can be multiple and of different natures, and more specifically, the uncertainty linked to the imperfection of knowledge (epistemic) can be significant in applications with high societal stakes in urban environments. One promising avenue for taking full and transparent account of uncertainties is that of imprecise probability theory (including, in particular, possibility theory [7] and Dempster Shafer theory [8],[9]). This framework has its foundations in classical probability theory and can be seen as a generalization of the Bayesian framework, bringing an additional degree of flexibility to express different types of uncertainty.

In machine learning, the total uncertainty of a prediction sums from two uncertainty types, arising from different sources: aleatoric and epistemic [10]. The former reflects the irreducible noise and ambiguity in the data due to class overlap, while the latter is related to the lack of knowledge about model parameters, and can be reduced by expanding the training dataset. These approaches are of growing interest due to the need for confidence in the prediction model [11], [12]. recently this approach has been extended in context of imprecise probability [13].

The aim of this post doc is to study and propose methods for decomposing and evaluating uncertainties in the context of spatial interpolation such as kriging [14] [15], IDW [16], based on the above theories of uncertainty.

[1] Li, J., & Heap, A. D. (2014). Spatial interpolation methods applied in the environmental sciences: A review. Environmental Modelling & Software, 53, 173-189.
[2] Belbeze, S., 2020. http://infoterre.brgm.fr/rapports//RP-70282-FR.pdf
[3] Meusburger, K., Evrard, O., Alewell, C., Borrelli, P., Cinelli, G., Ketterer, M., ... & Ballabio, C. (2020). Plutonium aided reconstruction of caesium atmospheric fallout in European topsoils. Scientific reports, 10(1), 11858.
[4] Rohmer, J., Loschetter, A., Raucoules, D., de Michele, M., Raffard, D., & Le Gallo, Y. (2015). Revealing the surface deformation induced by deep CO2 injection in vegetated/agricultural areas: The combination of corner-reflectors, reservoir simulations and spatio-temporal statistics. Engineering Geology, 197, 188-197.
[5] Belbèze, S., 2019. http://infoterre.brgm.fr/rapports//RP-69502-FR.pdf
[6] Thinon et al. (2013). https://infoterre.brgm.fr/rapports/RP-62383-FR.pdf
[7] Didier Dubois and Henri Prade. Possibility theory: an approach to computerized processing of uncertainty. Plenum Press, New York, 1988.
[8] A. P. Dempster. Upper and Lower Probabilities Induced by a Multivalued Mapping. The Annals of Mathematical Statistics, 38(2):325 – 339, 1967.
[9] G Shafer. A Mathematical Theory of Evidence. Princeton University Press, 1976.
[10] Armen Der Kiureghian and Ove Ditlevsen. Aleatory or Epistemic? Does It Matter? Structural Safety, 31(2):105–112, 2009.
[11] Liu, J., Lin, Z., Padhy, S., Tran, D., Bedrax Weiss, T., & Lakshminarayanan, B. (2020). Simple and principled uncertainty estimation with deterministic deep learning via distance awareness. Advances in neural information processing systems, 33, 7498-7512.
[12] Kotelevskii, N., Artemenkov, A., Fedyanin, K., Noskov, F., Fishkov, A., Shelmanov, A., ... & Panov, M. (2022). Nonparametric uncertainty quantification for single deterministic neural network. Advances in Neural Information Processing Systems, 35, 36308-36323.
[13] Hofman, P., Sale, Y., & Hüllermeier, E. (2024). Quantifying Aleatoric and Epistemic Uncertainty with Proper Scoring Rules. arXiv preprint arXiv:2404.12215.
[14] Chilès, J. P., & Desassis, N. (2018). Fifty years of kriging. Handbook of mathematical geosciences: Fifty years of IAMG, 589-612.
[15] Kevin Loquin and Didier Dubois. Kriging and epistemic uncertainty: a critical discussion. Methods for handling imperfect spatial information, pages 269–305, 2010.
[16] Shepard, D. (1968, January). A two-dimensional interpolation function for irregularly-spaced data. In Proceedings of the 1968 23rd ACM national conference (pp. 517-524).