Bayesian kriging : a comparison with validation criteria to ordinary kriging for radiological characterisation

Martin Wieskotten, Marielle Crozet, Bertrand Iooss, Céline Lacaux, Amandine Marrel

▶ **To cite this version:**

Martin Wieskotten, Marielle Crozet, Bertrand Iooss, Céline Lacaux, Amandine Marrel. Bayesian kriging : a comparison with validation criteria to ordinary kriging for radiological characterisation. geoENV2022 - The 14th International Conference on Geostatistics for Environmental Applications, University of Parma; geoENVia, Jun 2022, Parme, Italy. pp.56-62. cea-04743399

HAL Id: cea-04743399
https://cea.hal.science/cea-04743399v1

Submitted on 18 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

A comparison between Bayesian and ordinary kriging based on validation criteria: application to radiological characterisation

Martin Wieskotten[1,3], Marielle Crozet[1], Bertrand Iooss[2,5], Céline Lacaux[3], Amandine Marrel[4]

1. CEA, DES, ISEC, DMRC, Univ. Montpellier, Marcoule, France, marielle.crozet@cea.fr, martin.wieskotten@cea.fr

2. EDF R&D, 6 quai Watier, 78400, Chatou, France, bertrand.iooss@edf.fr

3. LMA Université d'Avignon, EA 2151, 84029, Avignon, France, celine.lacaux@univ-avignon.fr

4. CEA, DES, IRESNE, DER, Cadarache, Saint-Paul-Lez-Durance, F-13108 France, amandine.marrel@cea.fr

5. Institut de Mathématiques de Toulouse, Toulouse, France

Abstract

Radiological characterisation is one of the main challenges of the decommissioning and dismantling projects of nuclear facilities. This is an important step in decommissioning projects as it aims to estimate the quantity and spatial distribution of different radionuclides. As the type of contamination (nature of radionuclides, radioactive intensity, etc.) will require specific dismantling protocols, this estimation will condition the dismantling plan. In order to carry out the estimation, measurements are performed on site to obtain preliminary information. The information being incomplete (since all spatial positions cannot be measured due to time and cost constraints), spatial interpolation, using a kriging tool for example, is often performed to predict the value of interest for the contamination (radionuclide concentration, radioactivity, etc.) at unobserved positions.

A strong assumption made when applying ordinary kriging is that most of the kriging model parameters (except for the mean) used in the model are known. However, this is rarely the case in practice. Furthermore, the estimation error made when these parameters are estimated from the data is never taken into account, although this can lead to biased kriging predictions and overoptimistic prediction variances. This problem is emphasised when only a few observations are available, since the variance of the parameters' estimators becomes larger. This case is quite common for decommissioning projects, especially in the decommissioning of nuclear infrastructures where the number of measurements can be constrained by environmental risks such as radioactivity.

To address this issue, we propose here to use Bayesian kriging where the model parameters are considered as random variables, which allows to take into account their uncertainties. The use of prior specifications in Bayesian kriging also allows for more robust parameter estimate when only a few observations are available. Despite its advantages, Bayesian kriging has the drawback of being more computationally heavy and needing a prior specification, which is often difficult to choose and can strongly influence the predictions.

As such, the present work focused on assessing the usefulness of Bayesian kriging whilst comparing its performance to that of ordinary kriging. First, in order to make a relevant comparison, a simulated data set with known parameters is initially studied and several cross-validation criteria such as the predictivity

coefficient (Q^2), the Predictive Variance Adequacy (PVA), and the α-CI plot Demay et al., (2021) are estimated for varying data sizes to quantify the performances of both kriging methods.

A new criterion, the Predictive Interval Adequacy (PIA), is also introduced and studied.

Secondly, the same comparison is applied on a real data set from the decommissioning project of the G3 reactor at the CEA's Marcoule site. Finally the sensitivity of Bayesian kriging to prior specification and parameter estimation is also studied.

Keywords: geostatistics, bayesian kriging, ordinary kriging, validation criterion

Introduction

Radiological characterization is one of the main challenges encountered in the nuclear industry for the decommissioning and (D&D) of old infrastructures. Its main goal is to evaluate the quantity and spatial distribution of radionuclides. As such, measurements are made to constitute a data set and obtain preliminary information. While measurements are made, many problems can arise. The radioactivity present on site can be dangerous for operators and does not allow for many measurements. In some extreme cases, drones and robots have to be used (CEA DEN, (2017)), making measurements more expensive, therefore reducing data set's sizes. It is therefore quite common in nuclear decommissioning and dismantling characterisation to have little data available. A balance has to be found between information and costs, and statistical tools make it possible to optimise the information extracted, within a rigorous mathematical framework and with some confidence.

More precisely, spatial statistics or geostatistics are used to predict the variable of interest at unobserved position (prediction of the expected value), with an indication of the expected error in prediction (the prediction variance). The methodology is often based on two steps: first the construction of a statistical model with the estimation of its parameters, followed by prediction with a linear interpolator called kriging. Geostatistics also receive a common critic: its predictions do not take into account the uncertainty in the estimation of the model parameters. The variances of the predictions are too optimistic and the neglected model uncertainties can have a significant impact. This problem is made worse for smaller data sets, which can be common in D&D projects. The work of Desnoyers, (2010) is one of the first example of application to radiological characterisation. In this thesis a case study was analysed, but was based on many measurements, which is not realistic in the case of industrial nuclear D&D projects.

To overcome this, a Bayesian approach was first proposed by Kitanidis, (1986). Its main goal was to take into account uncertainties in the scale and mean parameters of the model. The work of Handcock and Stein, (1993) then completed the full Bayesian approach which considers all the parameters of the model as unknown. More recently, a slightly different approach was presented by Krivoruchko and Gribov, (2019) and is called empirical Bayesian kriging. While the equations are similar to the ones of regular Bayesian kriging, the prior choices are obtained through unconstrained simulations of the random field. This approach was adapted to allow for multi-fidelity applications, where Bayesian theory is used to update the initial data with new, more accurate data (classically used with cokriging if correlations between old and new data exist). Some examples can be found in meteorology with Gupta et al., (2017) or for oil extraction in Al-Mudhafar, (2019). Note that a more complete description of Bayesian kriging with an extension to generalised linear model is presented in Diggle and Ribeiro, (2007).

Material and Methods

Spatial Model and Predictions

The model considered is the following random field:

$$\{Z(x), x \in \mathbb{R}^2\}$$

which is constrained to $D \subset \mathbb{R}^2$. The random field $Z(.)$ is isotropic and stationary, meaning:

$$\forall\, x \in\ D, E[Z(x)] = \beta$$

$$\forall\, x, y\ \in\ D, Cov(Z(x), Z(y)) = \sigma^2\, C_\phi(|x - y|).$$

The term C_ϕ corresponds to a semidefinite positive function. Moreover, the random field is considered Gaussian. Thus for n observations at positions $\{x_1, \ldots, x_n\}$, we obtain the Gaussian random vector $\boldsymbol{Z} = (Z(x_1), \ldots, Z(x_n))'$:

$$\boldsymbol{Z}|\beta, \sigma^2, \phi \sim \mathcal{N}(\beta \mathbf{1}_n, \sigma^2 \boldsymbol{K}_\phi)$$

with $\mathbf{1}_n = (1, \ldots 1)'$, where β is the mean parameter, σ^2 the variance parameter, ϕ the range and $\sigma^2\, \boldsymbol{K}_\phi = (Cov(Z(x_i), Z(x_j)))_{1 \le i,j \le n}$ the covariance matrix. The observation of $\boldsymbol{Z}$ is written $\boldsymbol{z} = (z(x_1), \ldots, z(x_n))'$.

The model is therefore specified by 3 different parameters: the trend parameter $\beta \in\ D_\beta$, the scale parameter $\sigma^2 \in\ D_{\sigma^2}$ or variance and the range parameter $\phi \in\ D_\phi$. The first step of the geostatistical methodology is to estimate these parameters. Two main procedures exist: variographic analysis and maximum likelihood estimation. An extensive literature is available about parameter estimation with variographic analysis, such as Chilès and Delfiner, (2012), Webster and Oliver, (2007). We chose to use maximum likelihood estimation to avoid automatic fitting of variograms since the following applications will require the estimation of parameters for numerous data sets generated by simulations. Automatic fitting of variograms is strongly discouraged in most of the literature, Chilès and Delfiner, (2012), Webster and Oliver, (2007), so we avoid the methodology here.

The kriging predictor is a linear interpolator which expressions are derived from supplementary conditions, such as minimizing the prediction variance. For a detailed description of kriging and its construction, the reader can refer to the reference books of Chilès and Delfiner, (2012), Cressie, (1993) for geostatistics, but also Rasmussen and Williams, (2006) for computer code surrogate models. Let x_0 be an unobserved position at which we wish to predict the expected value and the variance of $Z(x_0)|\sigma^2, \phi, \boldsymbol{Z} = \boldsymbol{z}$ (the mean is considered unknown), the ordinary kriging equations are:

$$E[Z(x_0)|\sigma^2, \phi, \boldsymbol{Z}] = \left(\boldsymbol{k} + \mathbf{1}_n \frac{1 - \mathbf{1}_n' \boldsymbol{K}_\phi^{-1} \boldsymbol{k}}{\mathbf{1}_n' \boldsymbol{K}_\phi^{-1} \mathbf{1}_n}\right)' \boldsymbol{K}_\phi^{-1} \boldsymbol{Z}$$

$$Var[Z(x_0)|\sigma^2, \phi, \boldsymbol{Z}] = \sigma^2 \left(\mathbf{1} - \boldsymbol{k}' \boldsymbol{K}_\phi^{-1} \boldsymbol{k} + \frac{\left(1 - \mathbf{1}_n' \boldsymbol{K}_\phi^{-1} \boldsymbol{k}\right)^2}{\mathbf{1}_n' \boldsymbol{K}_\phi^{-1} \mathbf{1}_n}\right)$$

with $\sigma^2 \boldsymbol{k} = (Cov(Z(x_0), Z(x_j)))_{1 \le j \le n}$.

A major concern for applications of these equations is that they are conditional on the knowledge of the variance and correlation length parameters, which is mostly unrealistic since they are estimated. This assumption yields overoptimistic prediction variances and narrower confidence intervals. This problem is made worse for small data set where parameter estimation is sensible to each observation. To address this issue, more robust methods exist such as cross-validation estimation presented by Bachoc, (2013). Another solution is to consider the parameters as random variables. Bayesian approach seems natural in this case and leads to Bayesian kriging.

Indeed, Bayesian kriging deals simultaneously with estimation and predictions by considering the parameters as random variables that must be predicted conditionally to the observed data. We will use here the approach

described by Diggle and Ribeiro, (2002). For ease of notation, densities will be noted as $p(.)$ and the conditioning to parameters will be simplified from $Z|\beta = \tilde{\beta}\ to\ Z|\beta$.

Bayesian kriging predictions are derived from the predictive distribution as follows:

$$\begin{aligned} p(Z(x_0)|\boldsymbol{Z} = z) &= \int_{D_\beta \times D_{\sigma^2} \times D_\phi} p(Z(x_0), \beta, \sigma^2, \phi|\boldsymbol{Z} = \boldsymbol{z}) d\beta d\sigma^2 d\phi \\ &= \int_{D_\beta \times D_{\sigma^2} \times D_\phi} p(Z(x_0), \beta, \sigma^2|\phi, \boldsymbol{Z} = \boldsymbol{z}) p(\phi|\boldsymbol{Z} = \boldsymbol{z}) d\beta d\sigma^2 d\phi \\ &= \int_{D_\phi} p(Z(x_0)|\phi, \boldsymbol{Z} = \boldsymbol{z}) p(\phi|\boldsymbol{Z} = \boldsymbol{z}) d\phi \,. \end{aligned}$$

As per usual in Bayesian framework, we choose a joint prior distribution for β, σ^2:

$$\pi(\beta, \sigma^2) \propto \frac{1}{\sigma^2}.$$

For the correlation length, the prior is reduced to a uniform law between the minimum and the maximum distance allowed by the data set:

$$\phi \sim U(D_{min}, D_{max}).$$

Validation Criteria

The validation criteria are given by Demay et al., (2021) as means to compare and choose different covariance models for geostatistical predictions. These criteria aim to assess the quality of both the predictions of the model and the associated prediction variances. Their expressions are given here in their leave-one-out cross-validation formulation, but can be extended to K-fold cross-validation or to validation sets.

Predictivity coefficient (Q^2)

The main goal of this coefficient is to evaluate the predictive accuracy of the model by normalising the errors, allowing a direct interpretation in terms of explained variance. Its definition is the following:

$$Q^2 = 1 - \frac{\sum_{i=1}^{n} (z(x_i) - \hat{z}_{-i})^2}{\sum_{i=1}^{n} (z(x_i) - \hat{\mu})^2}$$

where $\hat{z}_{-i}$ is the value predicted by the model built without the i-th observation and $\hat{\mu}$ is the empirical mean of the data set.

This coefficient measures the quality of the predictions and how near they are to the observed values. It is similar to the coefficient of determination used for regression (with independent observations), but estimated here by cross-validation. The closer its value is to 1, the better the predictions are (relative to the observations). As a rule of thumb, if the Q^2 is smaller than 0.5 (i.e. less than 50% of output variance explained), the model is not considered valid.

Predictive variance adequacy (PVA)

This second criterion aims to quantify the quality of the prediction variances given by the model and kriging. Initially introduced by Bachoc, (2013), it is defined by the following equation:

$$PVA = \left| log\left(\frac{1}{n} \sum_{i=1}^{n} \frac{(z(x_i) - \hat{z}_{-i})^2}{\hat{s}_{-i}^2} \right) \right|$$

where $\hat{s}_{-i}^2$ is the prediction variance of the model built without the i-th observation.

It estimates the average ratio between the squared observed prediction error and the prediction variance. It therefore gives an indication of how much a prediction variance is bigger or smaller than the one expected. The closer the PVA is to 0, the better the prediction variances are. For example, a $PVA \approx 0.7$ indicates prediction variances that are on average two times bigger or smaller than the squared errors.

Predictive interval adequacy (PIA)

The PVA is criterion for variance adequacy but does not consider possible asymmetry in the predictive distribution. In the Gaussian case, mean and variance characterise completely the distribution, but in the case of Bayesian kriging where the predictive distribution is not Gaussian, the Q^2 and PVA are not enough to evaluate the quality of the model and its prediction. As such, we propose a geometrical criterion called the predictive interval adequacy (PIA) defined as follows:

$$PIA = \left| log\left(\frac{1}{n} \sum_{i=1}^{n} \frac{(z(x_i) - \hat{z}_{-i})^2}{\left(\hat{q}_{0.31,-i} - \hat{q}_{0.69,-i}\right)^2} \right) \right|$$

where $\hat{q}_{0.31,-i}$ (respectively $\hat{q}_{0.69,-i}$) is the estimation of the quantile of order 0.31 (respectively 0.69) of the predictive distribution without the i-th observation. It has been defined to be identical to the PVA for a Gaussian distribution, but rather than comparing squared errors to the predictive variance, it compares the width of prediction intervals with the squared errors. Another main difference is that the intervals defined by the PIA are centred on the median whereas the PVA is centred around the mean. Finally, an estimation of the predictive distribution is therefore necessary to compute this criterion, whereas the PVA only needs mean and variance to be computed.

α-CI plot and Mean Squared Error α ($MSE\alpha$)

The Gaussian process model allows to estimate credible intervals for predicted values, for $0 \leq \alpha \leq 1$:

$$CI_\alpha(z(x_i)) = [\hat{z}_{-i} - \hat{s}_{-i} q_{1-\alpha}; \hat{z}_{-i} + \hat{s}_{-i} q_{1-\alpha}]$$

This expression is only valid if all parameters are known. For example if the scale parameter is incorrectly estimated, the width of the predicted confidence intervals will not reflect what we might observe. But how can we validate a confidence interval without prior knowledge of the model parameters? The idea behind this criterion is to evaluate empirically the number of observations falling into the predicted confidence intervals and to compare this empirical estimation to the theoretical ones expected:

$$\Delta_\alpha = \frac{1}{n} \sum_{i=1}^{n} \delta_i \text{ where } \delta_i = \begin{cases} 1 \; if \; z(x_i) \in CI_\alpha(z(x_i) \\ 0 \; else \end{cases}.$$

This value can be computed for varying α, and can then be visualised against the theoretical values, yielding what Demay et al., (2021) called the α-CI plot.

Similarly to the PIA, the α-CI plot must be adapted to the Bayesian kriging since the posterior distribution is not Gaussian. We therefore introduce a slightly different criterion based on the quantiles of the predictive distribution. More precisely, the α-CI plot relies now on credible intervals defined as:

$$CI_\alpha(z(x_i)) = \left[\hat{q}_{\frac{1-\alpha}{2}}; \hat{q}_{\frac{1+\alpha}{2}}\right]$$

where $\hat{q}_{\frac{1-\alpha}{2}}$ (respectively $\hat{q}_{\frac{1+\alpha}{2}}$) is the estimation of the quantile of order $\frac{1-\alpha}{2}$ (respectively $\frac{1+\alpha}{2}$) of the predictive distribution without the i-th observation.

Once again we obtain a criterion that is identical for both methods when the predictive distribution is Gaussian. Illustration of α-CI plot for ordinary and Bayesian kriging (for a same learning sample) is given by Figure 1.

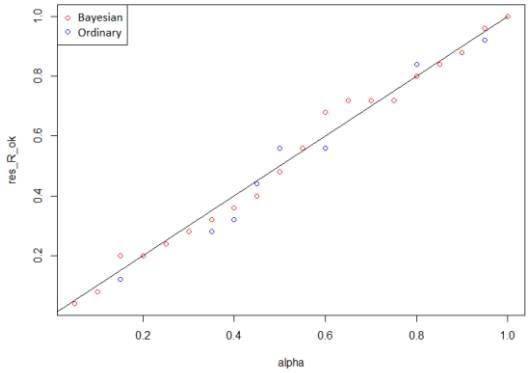

Figure1 Example of an α-CI plot with both Bayesian and ordinary kriging

To summarise the α-CI, we also introduce a new criterion called the Mean Squared Error α defined as follows:

$$MSE\alpha = \sum_{i=1}^{n_\alpha} (\Delta_{\alpha_i} - \alpha_i)^2$$

where n_α is the number of width considered for confidence intervals and α_i the width of the i-th confidence interval. The closer this criterion is to 0, the better the confidence/credible intervals are.

The different criteria aforementioned provide additional information to evaluate the prediction quality of the kriging model, either in terms of mean, variance or confidence/credible intervals. They will be used in the following to compare the performance of ordinary and Bayesian kriging.

Results

Protocol

Our goal is to compare Bayesian and ordinary kriging (which is the more commonly used kriging method) with our validation criteria and to identify if Bayesian kriging gives better results, and if so for which data set sizes.

We start with simulated data sets. We simulate 100 data sets for varying data set sizes, i.e. $16, 25, 36, 49, 64$ and 81 observations. The data sets are simulated in the space $[0,10]^2$ and the parameters chosen for the Gaussian field are:

$$\beta = 0.5, \sigma^2 = 0.1, \phi = 4.5$$

We also consider an exponential covariance model for the covariance function. For each of these data sets a cross-validation is applied to estimate each aforementioned validation criterion. Then the boxplot of each estimated criterion is represented against the data set sizes to compare both kriging's performances.

We wish to apply a similar protocol to a real data set for the G3 reactor in CEA Marcoule. This data set is made of 70 observations of radioactivity in the space $[0,10] \times [0,7]$. To generate multiple data sets, we sampled various data set sizes i.e. $20, 30, 40, 50, 60$ and 70 observations, with the last one being the real size of the original data set. Once again, for each data set a cross-validation is applied to estimate the validation criteria.

The boxplot of each estimated criteria is represented against the data set sizes.

Simulations

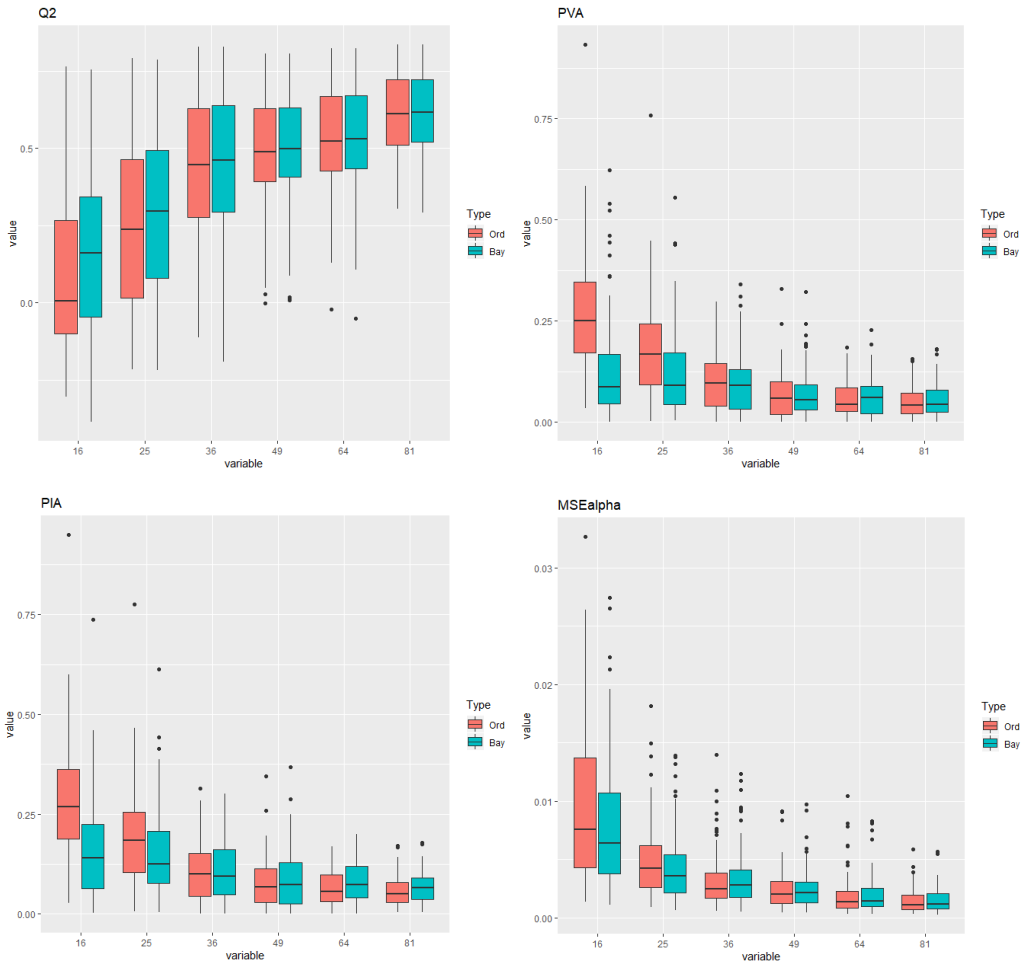

Figure 2 - Validation criterion against data size for simulated data for Q^2 criterion, PVA criterion, PIA criterion and MSEα criterion.

The results shown in Figure 2 indicate better prediction and prediction variances for smaller data sets. We can see on the different graphs that for data sets below 40 observations, Bayesian kriging seems to outperform ordinary kriging. This result is especially visible for the *PVA* and *PIA* for small data sets and shows that the main difference between both kriging methods still lies in the prediction variance estimation. It is mainly

explained by the fact that Bayesian kriging takes more of the parameter's estimation uncertainty into account than ordinary kriging. It therefore yields larger and more accurate confidence intervals, both represented by the PVA, PIA, and $MSE\alpha$ criteria.

It can also be noted that for larger data sets, Bayesian and ordinary kriging yield similar results. This observation is to be expected, since Bayesian and inferential methodology coincide for larger data sets. It can be therefore argued that Bayesian kriging loses some of its advantages for data set larger than 40 observations, since its computational cost is heavier than that of ordinary kriging. The Q^2values are also extremely low for 49 observations or less, but it is to be expected for very small data sets such as sets of 16 observations.

G3's data set

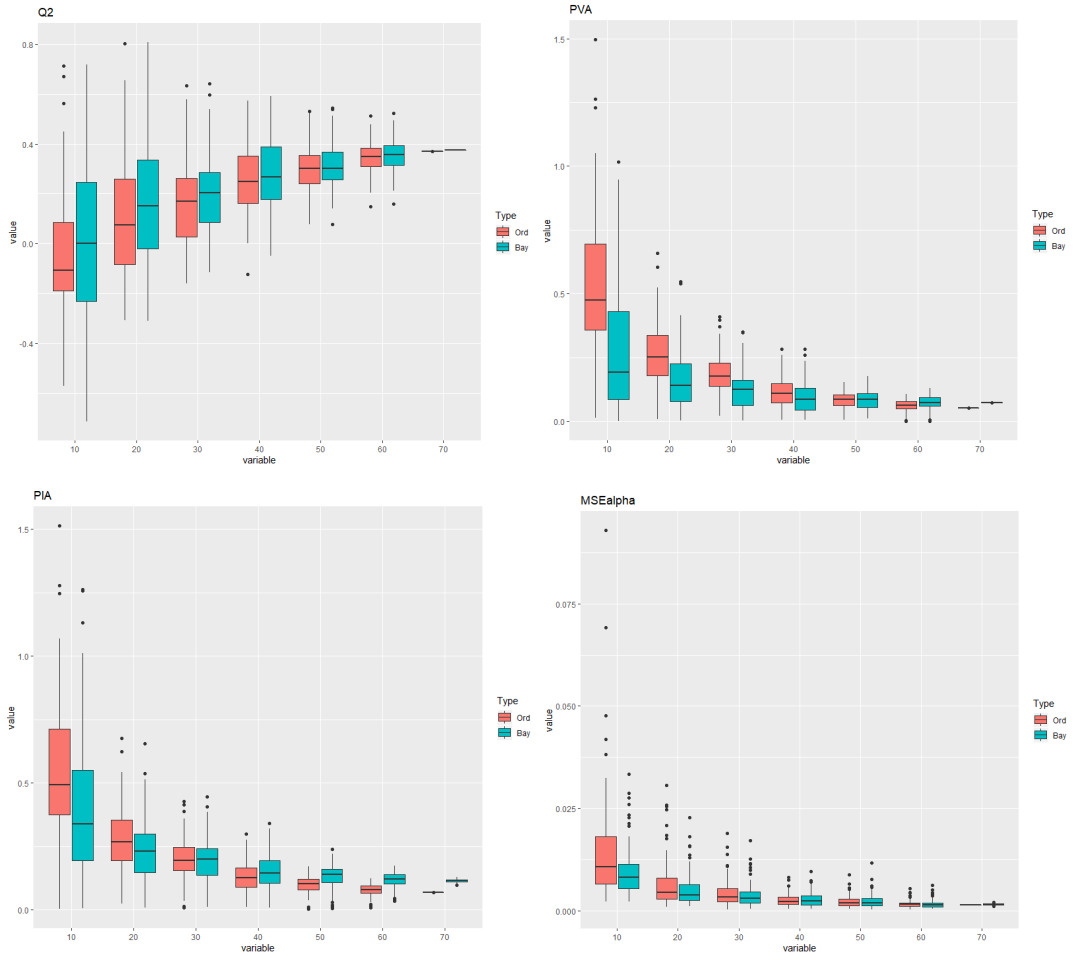

Figure 3 - Validation criterion against data size for the G3 data set for Q^2 criterion, PVA criterion, PIA criterion and $MSE\alpha$ criterion.

The Figure 3 shows similar results to the ones obtained for simulated data sets. One thing to note is that the variance of each validation criterion is reduced as the data sets size grows. This is both explained by the larger data sets, but also by our protocol, which chooses randomly points in the original 70 observations, so that when the data set sizes grows, the number of difference between randomly sampled observations is reduced. It can be noted that ordinary kriging seems to be slightly better than Bayesian kriging for larger data sets,

reinforcing our precedent argument that Bayesian kriging should be reserved for smaller data sets where uncertainty in parameter estimation is high.

Discussion and Conclusions

In conclusion the usage of Bayesian kriging for decommissioning and dismantling projects shows promising results, as its usefulness is made evident for small data sets. Even though Bayesian kriging gives better results, usual kriging methods such as ordinary kriging retains some advantages as their computational cost is lower and give similar results (sometimes even better results) for larger data sets. Therefore use of Bayesian kriging should be restricted to smaller data sets or cases in which prior information on parameters is well known. Our future works aim at a better modelling of measurement uncertainty, which could be achieved with the implementation of the work of Ng and Yin, (2012) for heteroscedastic models.

References

Al-Mudhafar, W.J., 2019. Bayesian kriging for reproducing reservoir heterogeneity in a tidal depositional environment of a sandstone formation. Journal of Applied Geophysics 160, 84–102. https://doi.org/10.1016/j.jappgeo.2018.11.007

Bachoc, François, 2013. Cross Validation and Maximum Likelihood estimations of hyper-parameters of Gaussian processes with model misspecification. Computational Statistics & Data Analysis 66, 55–69. https://doi.org/10.1016/j.csda.2013.03.016

Bachoc, F., 2013. Estimation paramétrique de la fonction de covariance dans le modèle de krigeage par processus gaussiens: application à la quantification des incertitudes en simulation numérique (PhD Thesis). Paris 7.

CEA DEN, 2017. L'assainissement-démantèlement des installations nucléaires, Le Moniteur. ed, Monographie CEA.

Chilès, J.-P., Delfiner, P., 2012. Geostatistics : Modeling Spatial Uncertainty, Second Edition. ed, Wiley Series In Probability and Statistics. Wiley.

Cressie, N., 1993. Statistics for spatial data. John Wiley & Sons.

Demay, C., Iooss, B., Le Gratiet, L., Marrel, A., 2021. Model selection based on validation criteria for Gaussian process regression: An application with highlights on the predictive variance. Quality and Reliability Engineering International. https://doi.org/10.1002/qre.2973

Desnoyers, Y., 2010. Approche méthodologique pour la caractérisation géostatistique des contaminations radiologiques dans les installations nucléaires (Thèse de doctorat). Paris, ENSMP.

Diggle, P.J., Ribeiro, P.J., 2007. Model-based Geostatistics, Springer Series in Statistics. Springer.

Diggle, P.J., Ribeiro, P.J., 2002. Bayesian Inference in Gaussian Model-based Geostatistics. Geographical and Environmental Modelling 6, 129–146. https://doi.org/10.1080/1361593022000029467

Gupta, A., Kamble, T., Machiwal, D., 2017. Comparison of ordinary and Bayesian kriging techniques in depicting rainfall variability in arid and semi-arid regions of north-west India. Environ Earth Sci 76, 512. https://doi.org/10.1007/s12665-017-6814-3

Kitanidis, P.K., 1986. Parameter Uncertainty in Estimation of Spatial Functions: Bayesian Analysis. Water Resources Research 22, 499–507. https://doi.org/10.1029/WR022i004p00499

Krivoruchko, K., Gribov, A., 2019. Evaluation of empirical Bayesian kriging. Spatial Statistics 32, 100368. https://doi.org/10.1016/j.spasta.2019.100368

Ng, S.H., Yin, J., 2012. Bayesian Kriging Analysis and Design for Stochastic Simulations. ACM Trans. Model. Comput. Simul. 22, 17:1-17:26. https://doi.org/10.1145/2331140.2331145

Rasmussen, C.E., Williams, C.K.I., 2006. Gaussian Processes for Machine Learning. MIT Press.

Webster, R., Oliver, M.A., 2007. Geostatistics for environmental scientists. John Wiley & Sons.