

# Kriging approach for the experimental cross-section covariances estimation

S. Varet<sup>1,a</sup>, P. Dossantos-Uzarralde<sup>1</sup>, N. Vayatis<sup>2</sup>, A. Garlaud<sup>3</sup>, and E. Bauge<sup>1</sup>

<sup>1</sup> CEA-DAM-DIF, 91 ArpaJon Cedex

<sup>2</sup> ENS Cachan

<sup>3</sup> ENSIIE

**Abstract.** In the classical use of a generalized  $\chi^2$  to determine the evaluated cross section uncertainty, we need the covariance matrix of the experimental cross sections. The usual propagation error method to estimate the covariances is hardly usable and the lack of data prevents from using the direct empirical estimator. We propose in this paper to apply the kriging method which allows to estimate the covariances via the distances between the points and with some assumptions on the covariance matrix structure. All the results are illustrated with the  $^{25}_{55}\text{Mn}$  nucleus measurements.

## 1 Introduction

The use of a generalized  $\chi^2$  for the evaluated cross section uncertainty determination needs to take into account the correlations and the uncertainties of the experimental measurements. As the covariances are not provided, we need to estimate the covariance matrix  $\Sigma_F$  of the experimental cross sections with its inverse.

As in many cases only one measure per energy is available the direct empirical estimator is not available and the classical method of error propagation [1] is rarely usable due to the lack of information on the experimental parameters and due to the linearity assumption that is not fulfilled.

The kriging is a valuable tool for covariances evaluation under the stationnarity assumption (constant mean, constant variance and covariances unvariant by translation) ([2],[3] and [4]). Here this method is introduced in the field of experimental cross-section covariances determination. Most of the time and until now, we dispose of only one measurement per energy. In this case kriging is appropriate since it is a linear interpolation method.

The aim of this work is to estimate the experimental covariances through the kriging. The particularity of this method is that the coefficients of the linear decomposition are determined with what we call the variogram. The variogram is a continuous function which depends on the distance (on the x axis in dimension 1) of the data to interpolate. Under some assumptions, like the second order stationarity of the data, we can rely on the covariance function and the variogram. Thus, the resulting kriging model depends on the data covariance matrix and in particular it depends on the inverse covariance matrix. The covariance matrix is estimated through the estimation of the variogram. As the inverse covariance matrix is needed, the covariance matrix estimation must be invertible. In the litterature some functions for the variogram are known to lead to an invertible covariance matrix. Therefore in the kriging method the covariance matrix is estimated in adjusting a parametrical function to the empirical variogram.

---

<sup>a</sup> e-mail: [suzanne.varet@cea.fr](mailto:suzanne.varet@cea.fr)

Here, we illustrate this approach with real data measurement of  $^{25}_{55}\text{Mn}$  nucleus.

In the next section we introduce the needed notations. We present the kriging for the covariance evaluation in the third section. The entire approach is illustrated in the case of the  $^{25}_{55}\text{Mn}$  nucleus in the fourth section.

## 2 Notations

We assume that the response (in our case the experimental cross-section) is a stochastic process, written  $F(E)$ , which depends on the  $s$ -dimensional variable  $E$  (in our case  $s = 1$  and  $E$  is the energy) and that we observe only at a finite number  $N$  of points  $E_j$ ,  $j = 1, \dots, N$ . We write  $F = (F_1, \dots, F_N)^t$  the random vector of the measured cross section<sup>1</sup> at the  $N$  distinct energies. The covariance matrix of  $F$  is denoted by  $\Sigma_F$ . We assume that we can observe one realisation of the  $F$  vector that we write  $F_1^{mes}, \dots, F_N^{mes}$ .

## 3 The kriging

The kriging is a linear interpolation method. More precisely the kriging method supposes that we can write the response (in our case the experimental cross section) at a given point  $E \in \mathcal{E}$  (in our case at a fixed energy) as a linear combination of the observed responses  $F_1^{mes}, \dots, F_N^{mes}$ :

$$F(E) = \sum_{i=1}^N \lambda_i F_i^{mes}.$$

In order to find the coefficients  $\lambda_i$ ,  $i = 1, \dots, N$ , we suppose that we can write the experimental cross section

$$F(E) = \mu(E) + \delta(E)$$

where  $\mu(E)$  is a deterministic component which represents the trend function and  $\delta(E)$  is a centered stochastic process which represents the residuals. In the simple kriging  $\mu(E)$  is assumed to be a known constant, in the ordinary kriging  $\mu(E)$  is assumed to be an unknown constant and in the universal kriging  $\mu(E)$  is assumed to be a linear combination of functions which depend on the position  $E$  that must be estimated. As a first approximation we suppose in this paper that we are in the ordinary kriging case. The coefficients  $\lambda_i$ ,  $i = 1, \dots, N$  are estimated in minimising the variance of the prediction with a non bias constraint. To deal with the unicity of the observation at a given point  $E_i$ , the kriging imposes a stationarity assumption on  $F(E)$  (or equivalently in the ordinary kriging case on  $\delta(E)$ ). In kriging, there is two types of stationarity assumptions: the second order stationarity (constant mean and covariance invariant by translation) and the intrinsic stationarity (the increment process has a constant mean and a variance invariant by translation). If  $F(E)$  is assumed to be a second order stationary process then it is a intrinsic stationary process but the inverse is not true in general. The variogram  $\gamma$  is define as the variance of the increment process:

$$2\gamma(h) = \text{Var}(F(E+h) - F(E)). \quad (1)$$

We assume now that  $F(E)$  is a second order stationary process. That is

$$\begin{cases} \mathbb{E}(F(E)) = m \quad \forall E \in \mathcal{E} \\ \text{Cov}(F(E+h), F(E)) = \mathcal{C}(h) \quad \forall E, E+h \in \mathcal{E}. \end{cases} \quad (2)$$

The  $\mathcal{C}$  function is the covariance function also called the covariogram. Under the second order stationarity assumption, the covariogram can be relied on the variogram by the following equation

$$2\gamma(h) = 2(\mathcal{C}(0) - \mathcal{C}(h)). \quad (3)$$

<sup>1</sup>  $()^t$  denotes the transpose of a vector.

This relation is a convenient way for the covariances estimation. Thus, for the experimental cross section covariances estimation, we assume that  $F(E)$  is a second order stationary process. Now we must estimate the variogram  $\gamma$ . The empirical estimation  $\hat{\gamma}$  of  $\gamma$  is

$$\hat{\gamma}(h) = \frac{1}{2|\mathcal{D}(h)|} \sum_{(i,j) \in \mathcal{D}(h)} (F_i^{mes} - F_j^{mes})^2 \quad (4)$$

where  $\mathcal{D}(h) = \{(i, j) \in \mathbb{N}^* \text{ such that } |E_i - E_j| = h\}$  and  $|\mathcal{D}(h)|$  is the cardinal of  $\mathcal{D}(h)$ . The value  $\mathcal{C}(0)$  is the variance of  $F(E)$  (assumed constant by the second order stationarity hypothesis) estimated by

$$\hat{\sigma} = \frac{1}{N-1} \sum_{i=1}^N (F_i^{mes} - \bar{F})^2 \quad (5)$$

where  $\bar{F} = \frac{1}{N} \sum_{i=1}^N F_i^{mes}$  is the empirical mean. Then we can estimate the covariances by

$\hat{\mathcal{C}}(h) = \hat{\sigma} - \hat{\gamma}(h)$ . However the obtained matrix can be non invertible. Moreover we need a continuous covariance function in order to interpolate at energies where no measurement have been done. In the litterature some continuous functions for the covariogram are known to lead to an invertible covariance matrix. Therefore we must find the most adapted covariogram function to the empirical variogram. In this paper we have used the exponential, gaussian and spherical functions:

exponential:  $\mathcal{C}_\theta(h) = \sigma^2 \exp\left(-\frac{|h|}{\theta}\right)$

gaussian:  $\mathcal{C}_\theta(h) = \sigma^2 \exp\left(-\frac{h^2}{2\theta^2}\right)$

spherical:  $\mathcal{C}_\theta(h) = 0$  if  $h > \theta$  and  $\mathcal{C}_\theta(h) = \sigma^2 \left(1 - \frac{3h}{2\theta} + \frac{1}{2}\left(\frac{h}{\theta}\right)^3\right)$  otherwise.

The covariance matrix estimation  $\widehat{\Sigma}_F$  is then given by:

$$\widehat{\Sigma}_F = \begin{pmatrix} \mathcal{C}_\theta(0) & \mathcal{C}_\theta(h_{12}) & \dots & \mathcal{C}_\theta(h_{1N}) \\ \vdots & \vdots & \dots & \vdots \\ \mathcal{C}_\theta(h_{N1}) & \mathcal{C}_\theta(h_{N2}) & \dots & \mathcal{C}_\theta(0) \end{pmatrix} \quad (6)$$

The kriging model  $F$  is then given, in minimising the mean square error, by:

$$F(E) = m + (\mathcal{C}_\theta(|E - E_1|), \dots, \mathcal{C}_\theta(|E - E_N|)) \widehat{\Sigma}_F^{-1} \begin{pmatrix} F_1^{mes} - m \\ \vdots \\ F_N^{mes} - m \end{pmatrix}.$$

#### 4 Application to the ${}^{55}_{25}Mn$ nucleus

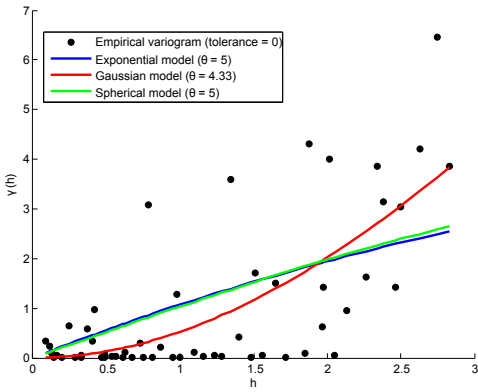
We present in this section the results of the variogram study for the  ${}^{55}_{25}Mn$  nucleus measurements. The measurements of the total and the  $(n,\gamma)$  cross sections are extracted from the EXFOR database ([5]) and the  $(n,2n)$  cross sections measurements are extracted from [6]. Please note that in order to prevent from the effect of high variations in the energy range and in the cross section range, we center and reduce the data. Then we calibrate the variogram model to the empirical variogram of the reduced data. Note also that when the number of measurements is high enough, in order to find a compromise between the  $\mathcal{D}(h)$  cardinal and the  $\hat{\gamma}$  sample size, we introduce a tolerance  $\varepsilon$  (a positive real) on the  $h$  distance in the  $\mathcal{D}(h)$  definition. That is  $\mathcal{D}_\varepsilon(h) = \{(i, j) \in \mathbb{N}^* \text{ such that } h - \varepsilon \leq |E_i - E_j| \leq h + \varepsilon\}$ . The variogram

model parameters are calibrated via a cross validation with the estimation of the mean square error as criterion:

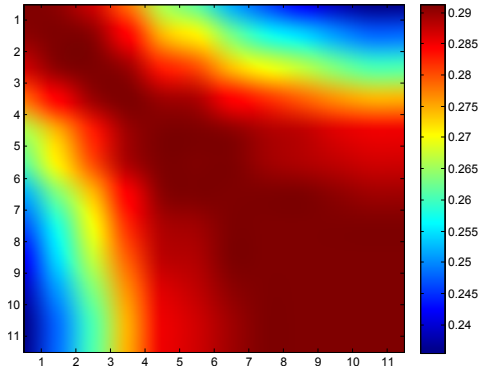
$$\text{mse}(\hat{\gamma}, \gamma) = \frac{1}{H} \sum_{k=1}^H (\hat{\gamma}(h_k) - \gamma(h_k))^2$$

where  $H$  is the  $\hat{\gamma}$  sample size.

For the  $^{55}_{25}\text{Mn}$  (n,2n), total and (n, $\gamma$ ) cross sections, the empirical variogram and the three variogram models (exponential, gaussian and spherical) are illustrated on figures 1(a) 2(a) and 3(a) respectively. For each of these cross sections, the covariance matrix associated to the lowest mean square error model is represented (1(b) 2(b) and 3(b) respectively). For the  $^{55}_{25}\text{Mn}$  (n,2n) cross section, the covariance matrix obtained with the variogram calibration seems to indicate that the correlations are higher at high energies. In the contrary, the  $^{55}_{25}\text{Mn}$  (n, $\gamma$ ) covariance matrix seems to indicate that the correlations are higher at low energies. Finally, in the case of the total cross section, the correlations seems similar for all energies.

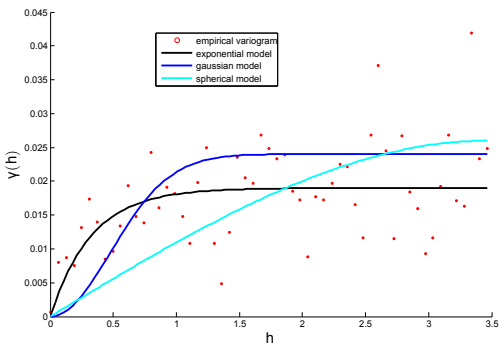


(a) Empirical variogram for the  $^{55}_{25}\text{Mn}$  (n,2n) cross section.

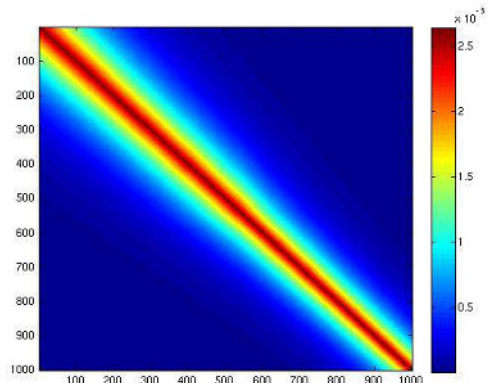


(b) Covariance matrix for the  $^{55}_{25}\text{Mn}$  (n,2n) cross section with gaussian model.

**Fig. 1.** Variogram and covariance matrix for the  $^{55}_{25}\text{Mn}$  (n,2n) cross section.

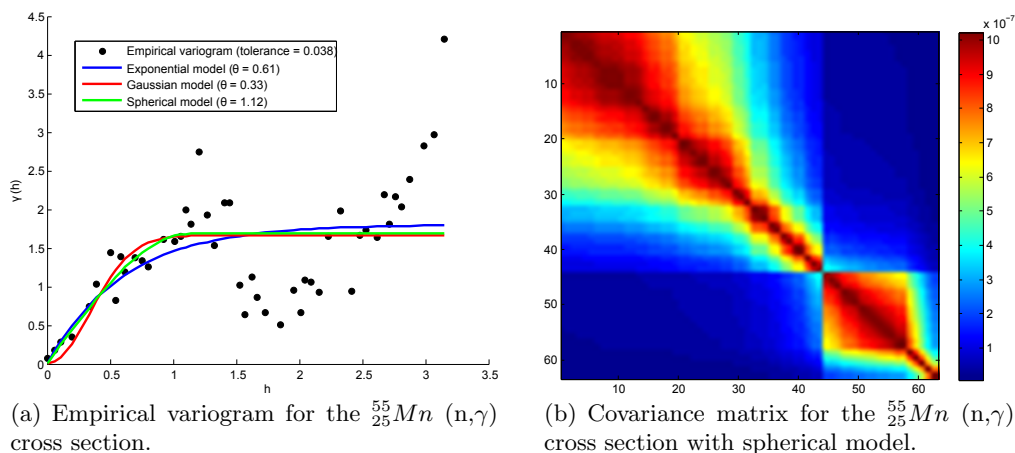


(a) Empirical variogram for the  $^{55}_{25}\text{Mn}$  total cross section.



(b) Covariance matrix for the  $^{55}_{25}\text{Mn}$  total cross section with exponential model.

**Fig. 2.** Variogram and covariance matrix for the  $^{55}_{25}\text{Mn}$  total cross section.



**Fig. 3.** Variogram and covariance matrix for the  $^{55}_{25}\text{Mn}$  ( $n,\gamma$ ) cross section.

## 5 Conclusion

In this paper an ordinary kriging method for experimental cross section covariance matrix estimation has been presented. The aim of this work is to apply this method to observed experimental cross section data. We focus on the situation where the lack of measurements prevents from using the direct empirical estimator and the informations on the experimental protocol are not available. To deal with the problem of one measurement per energy, the kriging approach makes strong assumptions on the covariance matrix structure. In this work we have assumed the variance stationarity and the decreasing correlations as the distance between measurements increases. In this context the covariance matrix can be easily approximated via the variogram study. The results could be compared with those obtained with another optimisation method [7]. Finally the consideration of less restrictive assumptions, like a non stationary mean, lead to a universal kriging approach which still remains a point to explore.

## References

1. M. Ionescu-Bujor and D. G. Cacuci, *A comparative review of sensitivity and uncertainty analysis of large-scale systems – I: Deterministic methods*, Nuclear Science and Engineering **147** (2004), 189–203
2. O. Roustant, D. Ginsbourger and Y. Deville, “DiceKriging, DiceOptim: Two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization”, (re-sent after minor revision in June 2012), Journal of Statistical Software, 2012.
3. M. Stein, “Interpolation of spatial data, some theory for kriging”, 1999, Springer.
4. E. Vasquez and E. Walter, “Choix d’une covariance pour la prédiction par krigeage de séries chronologiques échantillonnées irrégulièrement” I-Revue, GRETSI, Groupe d’Etudes du Traitement du Signal et des Images, 2005.
5. <http://www.oecd-nea.org/dbdata/x4/>
6. A. Milocco, A. Trkov, and R. Capote Noy *Nuclear data evaluation of  $^{55}\text{Mn}$  by the EMPIRE code with emphasis on the capture cross-section*, Nuclear Engineering and Design **241**(2011), 1071-1077.
7. S. Varet, P. Dossantos-Uzarralde, N. Vayatis and E. Bauge, “Pseudo-measurement simulations and bootstrap for the experimental cross-section covariances estimation with quality quantification”, WONDER 2012.