Kriging the Fields: a New Statistical Tool for Wave Propagation Analysis

Ph. De Doncker*, J-M Dricot*, R. Meys*, M. Hélier†, W. Tabbara†

Abstract **— A new statistical method for radiowave propagation analysis is presented based on the spatial statistics tools known as kriging and variographic analysis. In the proposed method, fields are considered as random variables of position. Using a few samples of those variables obtained by numerical or experimental means, unknown field values with confidence intervals are inferred. Validation of the new approach is carried out on measurements in indoor environment.**

1 INTRODUCTION

Radiowave propagation prediction is the key point of many applications for wireless network performance analysis or electromagnetic compatibility studies. It generally focus either on the local field behaviour by experimentation or numerical methods, or on the global field characterization at statistical level. However, when studying radiating systems, both approaches can be useless either because statistical characterization is too coarse to be meaningful at local scale, or because exact local behaviour prediction is too time-consuming or impossible to carry out numerically due to system modelization complexity. In those cases, a third way should be considered, where predictions would have to be drawn using few field values measured or computed. This third way can be stated as follows: to estimate at the local scale field values from data being too sparse or too coarse to be exploited by classical analysis. Spatial statistics provide the tools required to solve this problem.

Spatial statistics were historically developed in the framework of geostatistics to provide mineral content estimations for mine exploitation. Basically, the method proposed here is to apply the same technique to electromagnetic wave propagation analysis: the electric field values within a given area is considered as one realization of a random function (RF). Samples of this RF being known numerically or experimentally, spatial statistics is used to infer the RF values throughout the whole area and, more important, their estimation variance, i.e. their confidence interval.

*Université Libre de Bruxelles, CP 165, Av. Roosevelt, 50 1050 Brussels, Belgium, e-mail: `pdedonck@ulb.ac.be`, `jdricot@ulb.ac.be`, `rmeys@ulb.ac.be`, Fax: +32 2 6504713

†Supélec/DRE, Plateau de Moulon 91192, Gif-sur-Yvette, France, e-mail: `Marc.Helier@supelec.fr`, `tabbara@lss.supelec.fr`

Kriging has already been considered with success in electromagnetics [1], namely in transmission line study [2], but the whole spatial statistics framework has never been adapted to wave propagation. This approach participates to the authors' general effort to use statistics in order to reduce the cost of analysis of complex systems, and to estimate the accuracy of the results [3,4].

In the first part of this paper, the basic features of spatial statistics will be presented without all the mathematical details. Next, it will be shown how the field statistical properties can be theoretically developed in the general case of Rayleigh channels including the most familiar wireless network or electromagnetic compatibility applications such as communication channel estimation, EMC compliance tests or reverberation chambers analysis. Finally, the proposed method will be shortly validated on experimental data's.

2 SPATIAL STATISTICS

Spatial statistics basic features will be shortly introduced here. Further reading can be found e.g. in [5].

2.1 The regionalized variables

It is well known that in propagation phenomena through complex channels, field distributions cannot be described by any analytical expression due to their highly oscillatory behaviour. Generally, a statistical approach has rather to be considered. Channel definition in terms of log-normal, Rice or Rayleigh distributions for instance is usual in wireless network design [6], whereas the complex cavity statistics concept is widely used in the electromagnetic compatibility community [7]. Using the spatial statistics framework, the electromagnetic fields can be considered throughout the geographic region under study as *random functions* (RF) of position. These functions are moreover asked to be second order stationary, i.e. if $Z(x)$ represent the RF associated to the electric or magnetic field at x:

$$\begin{aligned} \mathcal{E}\{Z(x)\} &= m \\ \mathcal{E}\{[Z(x)-m][Z(x+h)-m]\} &= C(h) \end{aligned} \tag{1}$$

where $\mathcal{E}\{\}$ denotes the mathematical expectation and $C(h)$ the covariance function. This equation states that the mean and covariance of the RF $Z(x)$ are to be position independent. We shall restrict ourselves here to this class of RF, although it is evident that in propagation problems the mean electric field strength for instance is not always stationary when the distance to the transmitter greatly varies. In this case, either the basic spatial statistics tools are to be applied on subsets of the region under study, or generalized tools must be considered.

The RF $Z(x)$ itself can never be observed, and its realizations $z(x)$ are the only entities accessible to the experimenter. In spatial statistics, realizations are called *regionalized variables*. In fact, $z(x)$ as a whole is also impossible to obtain since it would ask an infinite number of measurements or computations, and all the studies have to be based on a finite set of values $\{z(x_\alpha)\}$ at the measurement or computation points $\{x_\alpha\}$. The denser the sampling grid $\{x_\alpha\}$, the more accurate predictions can be.

2.2 Variographic analysis

Any regionalized variable $z(x)$ presents a spatial structure due to the underlying physical process such as fast fading in communication channels or complex cavities. This spatial structure can be studied by looking at the associated RF $Z(x)$ *variogram* $\gamma(h)$ which evaluates the similarity between pairs of values of $Z(x)$ at x and $x+h$. It is defined by:

$$\gamma(h) \;=\; \frac{1}{2}\,\mathrm{Var}\left\{Z(x+h) - Z(x)\right\} \qquad (2)$$

where Var{} denotes the variance. Under the second order stationarity assumption, the variogram can be shown to be equivalent to the covariance function:

$$\gamma(h) \;=\; C(0) - C(h) \qquad (3)$$

However, the variogram is preferable because on the one hand it is insensible to any determination error of the mean value m, and on the other hand because it can be generalized to non-stationary RF, in contrary to the covariance. The variogram defined by (2) relies on the knowledge of the whole random process $Z(x)$. However, in practice, only a finite set of values $\{z(x_\alpha)\}$ is accessible and it is necessary to estimate $\gamma(h)$ from these values. An unbiased estimator of the variogram is given by the *experimental variogram*

$$\hat{\gamma}(h) \;=\; \frac{1}{2N_h} \sum_{|x_\alpha - x_\beta| = h} \left[\, z(x_\alpha) - z(x_\beta)\,\right]^2 \qquad (4)$$

where N_h is the count of pairs separated by the lag h.

2.3 Kriging

Let us suppose now that a set of values $\{z(x_\alpha)\}$ has been obtained by experimental or numerical means. The problem is how to estimate unknown values of the regionalized variable $z(x)$ using this set, or, at least, to obtain an approximation of unknown values with confidence intervals. Using classical methods, it is impossible to solve this problem because either the statistical properties are defined on a global scale, or the point value prediction cannot take advantage of the known results. In fact, to take into account what has already been measured or computed, the *local* spatial structure of the regionalized variable has to be considered. The variogram provides this possibility through the method known as *kriging*.

Let us suppose that the experimenter wants to estimate an unknown value $z(x_0)$. In the kriging process, the starting point is the estimation $\hat{Z}(x_0)$ of the corresponding random variable $Z(x_0)$. To take into account the already known values, this latter is written as a linear combination of the associated random variables:

$$\hat{Z}(x_0) \;=\; \sum_{\alpha=1}^{n} w_\alpha\, Z(x_\alpha) \qquad (5)$$

Imposing $\hat{Z}(x_0)$ to be the best linear unbiased estimator of $Z(x_0)$, it can be shown that the weights w_α are solution of the algebraic system:

$$\begin{cases} \displaystyle\sum_{\beta=1}^{n} w_\beta\, \gamma(|x_\alpha - x_\beta|) \;-\; \mu \;=\; \gamma(|x_\alpha - x_0|) & \\ & \alpha = 1..n \\ \displaystyle\sum_{\alpha} w_\alpha = 1 & \end{cases} \qquad (6)$$

It is then possible to compute $\hat{z}(x_0)$ for the particular realization considered according to

$$\hat{z}(x_0) \;=\; \sum_{\alpha=1}^{n} w_\alpha\, z(x_\alpha) \qquad (7)$$

Moreover, the great advantage inherent to kriging is coming from the fact that the estimation variance of the prediction can also be deduced from the weights w_α and the variogram $\gamma(h)$:

$$\sigma_e^2(x_0) \;=\; -\mu \;-\; \gamma(0) \;+\; 2\sum_{\alpha=1}^{n} w_\alpha\, \gamma(|x_\alpha - x_0|) \qquad (8)$$

so that to each estimated value, it is possible to link a confidence interval $[\hat{z}(x_0) - n_\sigma \sigma_e(x_0), \hat{z}(x_0) +$

$n_\sigma \sigma_e(x_0)]$ wherein the regionalized variable is present with a given certainty. To denser the sampling grid $\{x_\alpha\}$, the smaller this interval, but what is important to note is the fact that whatever the prior experimenter knowledge of the regionalized variable is, an interval can be given *locally* inside which the fields can be supposed to lie. In EMC problems for instance, this kind of information is generally sufficient because what matters is only the compliance to given regulatory rules, and not exact values which are most often beyond the experimenter knowledge due to system modeling complexity. For each kind of applications, the confidence interval must be first calibrated (n_σ must be chosen) according to the certainty level to be reached, but this work must only be done once.

3 APPLICATION TO PROPAGATION

3.1 Analytical variogram derivation

In an indoor or outdoor communication channel, or in an enclosure or reverberation chamber, the fields are made up many waves arriving from the various scatterers. The electric field can thus be well modeled according to the integral representation

$$\vec{E} \;=\; \int_\Omega \vec{F}(\Omega)\, e^{-j\vec{k}\cdot\vec{r}}\, d\Omega \tag{9}$$

where the integration is performed over all real angles $\Omega = (\theta, \phi)$ and where, in a spherical coordinate system, the plane wave spectrum $\vec{F}(\Omega)$ can be written as

$$\vec{F}(\Omega) \;=\; F_\theta(\Omega)\, \vec{1}_\theta \;+\; F_\phi(\Omega)\, \vec{1}_\phi \tag{10}$$

with both F_θ and F_ϕ complex functions.

Assuming that the electric field is made up many uncorrelated plane waves having equal energy and random phases, the statistical properties of the plane wave spectrum can be deduced as shown in [11]. The statistical properties of each field component can then be studied. Let us for instance consider the complex vertical component $E_z(\vec{r})$. The correlation function between the two values $\vec{E}_z(\vec{r}_1)$ and $\vec{E}_z(\vec{r}_2)$ seen as random variables is given by [8]

$$\rho(\vec{r}_1, \vec{r}_2) \;=\; \frac{1}{E_0^2/3}\, \mathcal{E}\{E_z(\vec{r}_1)\, \vec{E}_z^*(\vec{r}_2)\} \tag{11}$$

where $E_0^2 = \mathcal{E}\{|\vec{E}|^2\}$ is the electric field mean power and * denotes complex conjugate. In the xy plane, this expected value can be computed in closed form [11]:

$$\rho(d) \;=\; \frac{3}{2}\, \frac{\sin kd}{kd} \left(1 - \frac{1}{(kd)^2}\right) \;+\; \frac{3}{2}\, \frac{\cos kd}{(kd)^2} \tag{12}$$

where $d = |\vec{r}_1 - \vec{r}_2|$. Next, assuming that both the real and imaginary parts of each field component satisfy a gaussian distribution with zero mean and equal variances, as usual in communication channels or reverberation chambers, the correlation function for $|E_z|$ can be shown [7] to be

$$\rho_m(d) = \frac{\frac{2}{\pi}\, \rho(d) \mathrm{Arcsin}(\rho(d)) + \frac{2}{\pi}\, \sqrt{1 - \rho^2(d)} - \frac{\pi}{4}}{1 - \pi/4} \tag{13}$$

The analytical variogram for $|E_z|$ can then be deduced from (4). This expression will depend on E_0 and it has thus to be adjusted according to the situation.

3.2 Kriging algorithm

Let us suppose that a set of $|E_z|$ values are known. In our notations, this set is written as $\{z(x_\alpha)\}$. To define an estimator $\hat{z}(x_0)$ of an unknown value $z(x_0)$, the following algorithm has to be followed:

1. Calculate the experimental variogram $\hat{\gamma}(h)$ by (4) using $\{z(x_\alpha)\}$.
2. Adjust E_0 in the analytical variogram to fit $\hat{\gamma}(h)$.
3. Choose n in (7) to compute $\hat{z}(x_0)$. Usually, the two nearest neighbouring known values are sufficient, so that $n = 2$.
4. Solve (6) using the analytical variogram.
5. Compute $\hat{z}(x_0)$ and $\sigma_e^2(x_0)$ by (7) and (8).
6. Choose n_σ to define the confidence interval. This step must be carried out thanks to a calibration procedure which will not be described here. See [9] for further details. In case of fast fading, $n_\sigma = 1$ gives good results.

4 RESULTS

To validate the approach, measurements at 2 GHz were undertaken in indoor environment at the University of Brussels. A single result will be shown here, but further application possibilities will be explained during the presentation.

On Figure 1 an experimental variogram and the corresponding adjusted analytical one are drawn. The agreement is quite good, showing that the plane wave representation of the electric field is justified. After a few centimeters corresponding to $\lambda/2$, the variogram reaches a sill indicating the maximal spatial lag before uncorrelation. If the known experimental values are separated by a larger distance, classical methods cannot predict field values in between. Using kriging and the analytical variogram, the field spatial structure can be taken into account for any distance, and inference can be done.

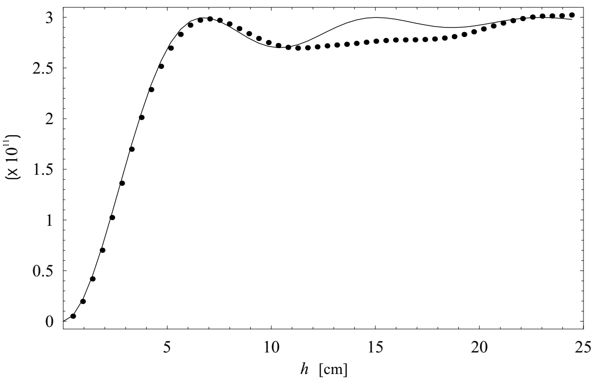

Figure 1: Theoretical (solid line) and experimental (dotted line) variograms

Figure 2 shows the kriging result obtained with field values measured each 40cm (more than 2λ). The thick line is the field estimator, the thin solid lines limit the confidence interval, and the dashed line is the full experimental result drawn to validate the kriging estimator. The supposed known values lie at the points where the confidence interval collapses (kriging is an exact interpolator: at the known values, the estimator is exact and the confidence interval is thus zero). It is possible to see that the estimator does not follow the exact experimental values. The sampling is indeed too coarse to allow fast variation prediction. However, the confidence interval defines variation bounds wherein 92% of the experimental values lie. Using kriging in combination with a rough sampling scheme allows to predict field envelope variation only, which is often sufficient. To obtain more accurate estimator, a denser sampling grid is necessary, as will be shown during the presentation.

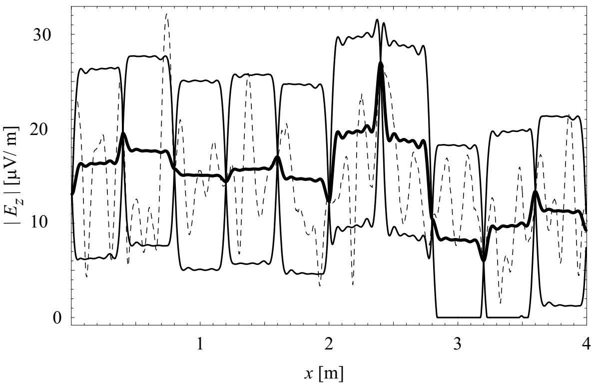

Figure 2: Kriging estimator (thick line), confidence interval (solid lines) and experimental values (dashed line)

5 CONCLUSIONS

A new statistical method for radiowave propagation analysis has been proposed. The method considers fields as random variables of position, and inference is carried out taking into account their spatial structure via the variographic analysis. Validation has been shown on experimental data's obtained in indoor environment.

References

[1] J. Lefebvre, H. Roussel, E. Walter, D. Lecointe, W. Tabbara, "Prediction from wrong models: the kriging approach",IEEE Ant.Prop.Mag.,38,pp35-45, Aug.1996.

[2] V. Rannou, F. Brouaye, M. Hélier, W. Tabbara, "Kriging the quantile: application to a simple transmission line model",Inv.Prob.,18, pp.37-48,2002.

[3] W. Tabbara, Ph. De Doncker, M. Hélier, D. Lecointe, "Electromagnetic waves propagation and coupling: the statistical approach to information retrieval",Proc. Europ. Test and Telemetry Conf., Toulouse, France, June 2003.

[4] Ph. De Doncker, M. Hélier, W. Tabbara, "Statistical analysis of broadband personal area networks disturbances on a transmission line in indoor environment",Proc. Int. Symp. on Electromag. Comp. EMC Europe, pp.83-86, Sorrento, Italy, Sept. 2002.

[5] H. Wackernagel, "Multivariate geostatistics, an introduction with applications", Springer, New York, 1995.

[6] S.R. Saunders, "Antennas and propagation for wireless communication systems", Wiley, New York, 1999

[7] T.H. Lehman, "A statistical theory of electromagnetic fields in complex cavities", Interaction Notes, Note 494, May 1993

[8] Ph. De Doncker, "Spatial correlation functions for fields in three-dimensional Rayleigh channels", PIERS, 40, pp55-69, 2003

[9] Ph. De Doncker, J-M. Dricot, R. Meys, M. Hélier, W. Tabbara, "Application of spatial statistics to electromagnetic wave propagation", to be published.