

Kriging-based subdivision schemes: Application to the reconstruction of non-regular environmental data

Jean Baccou, Jacques Liandrat

► **To cite this version:**

Jean Baccou, Jacques Liandrat. Kriging-based subdivision schemes: Application to the reconstruction of non-regular environmental data. *Mathematics and Computers in Simulation*, Elsevier, 2011, pp 2033-2050 10.1016/j.matcom.2010.12.009 . hal-01270181

HAL Id: hal-01270181

<https://hal.archives-ouvertes.fr/hal-01270181>

Submitted on 5 Feb 2016

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.

Kriging-based subdivision schemes: application to the reconstruction of non-regular environmental data

Jean Baccou^a, Jacques Liandrat^b

^a*IRSN, Centre de Cadarache, 13115 Saint Paul-Lez-Durance, France*

^b*Centrale Marseille and LATP, Technopôle de Château-Gombert, 13451 Marseille Cedex 20, France*

Abstract

This work is devoted to the construction of new kriging-based interpolating position-dependent subdivision schemes for data reconstruction. Their originality stands in the coupling of the underlying multi-scale framework associated to subdivision schemes with kriging theory. Thanks to an efficient stencil selection, they allow to cope the problem of non-regular data prediction while keeping the interesting properties of kriging operators for the quantification of prediction errors. The proposed subdivision schemes are fully analyzed and an application to the reconstruction of non-regular environmental data is given as well.

Key words: kriging, subdivision scheme, non-regular data prediction.

1. Introduction

Data modeling methods play a key role in the treatment of environmental problems. The information about the phenomenon under study is often discrete since it is provided by experimental measurements or numerical processing and therefore, modeling (or prediction) is required. Among data modeling methods, stochastic kriging-based approaches ([4]) are often used. Their main advantages compared to deterministic methods stand in the possibility to quantify the precision of the prediction thanks to an underlying probabilistic model. However, these methods usually assume that the phenomenon to predict is regular, which is not the case in practice such as in risk analysis where reliable reconstruction methods are crucial for the decision-making process. This paper is therefore devoted to the design of new stochastic modeling methods that improve the accuracy of the reconstruction of non-regular data. Contrary to classical approaches, their construction will be performed in two steps: a segmentation of data into different zones and a local kriging-based data prediction according to the information coming from the previous step. Taking apart the problem of

Email addresses: jean.baccou@irsn.fr (Jean Baccou),
jacques.liandrat@ec-marseille.fr (Jacques Liandrat)

data segmentation that refers to the wide literature of edge detector, we focus in the sequel on the construction of a new kriging-based subdivision scheme ([5]) for data prediction integrating local strategy according to the segmentation. Our work is organized as follows: Section 2 deals with the construction of our new subdivision schemes. Convergence results are provided in Section 3. Section 4 is devoted to applications to synthetic or real data.

2. Kriging-based position dependent prediction

In this section, we plug the kriging interpolation into the Harten's subdivision framework ([6]).

The general Harten's setting is a family of triplets $(V^j, D_j^{j-1}, P_{j-1}^j)$ where $j \in \mathbb{Z}$ is a scale parameter. For each value of j , V^j denotes a separable space of approximation associated to a resolution level 2^{-j} , D_j^{j-1} (resp. P_{j-1}^j) is a decimation (resp. prediction) operator connecting V^j to V^{j-1} (resp. V^{j-1} to V^j). Usually, these two operators are constructed from the pairs (D_j, R_j) of discretization and reconstruction operators as $P_{j-1}^j = D_j R_{j-1}$ and $D_j^{j-1} = D_{j-1} R_j$.

Since the prediction operator P_{j-1}^j maps any element of V^{j-1} in an element of V^j , it defines a subdivision scheme ([5]) that reads $\forall \{f_k^{j-1}\}_{k \in \mathbb{Z}}$, $\left(P_{j-1}^j f^{j-1}\right)_k = \sum_{m \in \mathbb{Z}} a_{k-2m}^{j,k} f_m^{j-1}$. Here, $\{a_{k-2m}^{j,k}\}_{m \in \mathbb{Z}}$, $(j, k) \in \mathbb{Z}^2$ is called the mask of the subdivision scheme. If it is independent of f^{j-1} , the subdivision is said to be linear. Moreover, one speaks about stationarity (resp. uniformity) when the mask does not depend on j (resp. on k).

The design of prediction operators is crucial for data reconstruction. As mentioned before, it depends on the discretization and reconstruction operators. From now on, we assume that D_j is the sampling operator defined by $(D_j f)_k = f(x_k^j)$ with $x_k^j = k2^{-j}$.

Here, we propose to use kriging theory to define the reconstruction operator. In kriging approaches ([4]), $\{f_k^j\}_{k \in \mathbb{Z}}$ are considered as realizations of a subset of random variables $\{\mathcal{F}(k2^{-j}), k \in \mathbb{Z}\}$ coming from a random process $\{\mathcal{F}(x), x \in \mathcal{D}\}$ (with \mathcal{D} a bounded domain of \mathbb{R}) that reads $\mathcal{F}(x) = m(x) + \delta(x)$, $x \in \mathcal{D}$, where $m(x)$ is the deterministic mean structure of $\mathcal{F}(x)$ and $\{\delta(x), x \in \mathcal{D}\}$ is a zero-mean random process.

Kriging theory assumes that the random variables $\delta(x)$, are spatially correlated. Under stationarity assumptions and constant deterministic mean structure of the random process, the spatial correlation structure of $\{\delta(x), x \in \mathcal{D}\}$ is identified to the spatial correlation of the data and is exhibited by computing the semi-variogram, $\gamma(h) = \frac{1}{2}E((\mathcal{F}(x+h) - \mathcal{F}(x))^2)$, where E denotes the mathematical expectation.

Assuming for the rest of our work that the semi-variogram is identified once and for all from the available data $\{f_k^0\}_{k \in \mathbb{Z}}$, it is approximated by a least square fit of

the discrete experimental semi-variogram, $\gamma_{exp}(h) = \frac{1}{2Card(N(h))} \sum_{(k,l) \in N(h)} (f_k^0 - f_l^0)^2$, with $N(h) = \{(k, l), h - \epsilon \leq |(k-l)2^{-j}| \leq h + \epsilon\}$ and for every h such that $Card(N(h))$, the cardinality of $N(h)$, is sufficiently large. For any integer D , the kriging-based reconstruction operator is finally defined as a piecewise interpolation operator such that for all $f^j \in V^j$,

$$R_j f^j(x) = \sum_{m=-l_{j+1,2k-1}}^{r_{j+1,2k-1}-1} \lambda_{j,m}^{l_{j+1,2k-1}, r_{j+1,2k-1}-1}(x) f_{k+m}^j, \text{ if } x \in [(k-1)2^{-j}, k2^{-j}]. \quad (1)$$

The parameter $l_{j+1,2k-1}$ (resp. $r_{j+1,2k-1}$) denotes the number of left (resp. right) points in the interpolating stencil associated to the interval $[x_{k-1}^j, x_k^j]$ such that $l_{j+1,2k-1} + r_{j+1,2k-1} = D+1$. The kriging weights, $\{\lambda_m^{l_{j+1,2k-1}, r_{j+1,2k-1}-1}(x)\}_m$ are solutions of the following problem:

$$\left\{ \begin{array}{l} \text{Minimize the estimation variance:} \\ E \left(\sum_{m=-l_{j+1,2k-1}}^{r_{j+1,2k-1}-1} \lambda_{j,m}^{l_{j+1,2k-1}, r_{j+1,2k-1}-1}(x) (\mathcal{F}((k+m)2^{-j}) - \mathcal{F}(x))^2 \right) \text{ for } x \in [(k-1)2^{-j}, k2^{-j}] \\ \text{under the unbiasedness constraint:} \\ E \left(\sum_{m=-l_{j+1,2k-1}}^{r_{j+1,2k-1}-1} \lambda_{j,m}^{l_{j+1,2k-1}, r_{j+1,2k-1}-1}(x) \mathcal{F}((k+m)2^{-j}) \right) = E(\mathcal{F}(x)). \end{array} \right. \quad (2)$$

The kriging-based prediction operator associated to the reconstruction (1) is then defined for all $k \in \mathbb{Z}$ as:

$$\left\{ \begin{array}{l} f'_{2k}{}^{j+1} = (P_j^{j+1} f'^j)_{2k} = f'_k{}^j, \\ f'_{2k-1}{}^{j+1} = (P_j^{j+1} f'^j)_{2k-1} = \sum_{m=-l_{j+1,2k-1}}^{r_{j+1,2k-1}-1} \lambda_{j,m}^{l_{j+1,2k-1}, r_{j+1,2k-1}-1} f'_{k+m}{}^j, \end{array} \right. \quad (3)$$

where $\forall k \in \mathbb{Z}$, $f'_k{}^0 = f_k^0$ and $\{\lambda_{j,m}^{l_{j+1,2k-1}, r_{j+1,2k-1}-1}\}_m$ are the kriging weights solution of Problem (2) when $x = -\frac{1}{2}$. They satisfy:

$$\begin{bmatrix} \gamma(0) & \dots & \gamma(D2^{-j}) & 1 \\ \gamma(2^{-j}) & \dots & \gamma((D-1)2^{-j}) & 1 \\ \dots & \dots & \dots & 1 \\ \gamma(D2^{-j}) & \dots & \gamma(0) & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_{j,-l_{j+1,2k-1}}^{l_{j+1,2k-1}, r_{j+1,2k-1}-1} \\ \lambda_{j,-l_{j+1,2k-1}+1}^{l_{j+1,2k-1}, r_{j+1,2k-1}-1} \\ \dots \\ \lambda_{j,r_{j+1,2k-1}-1}^{l_{j+1,2k-1}, r_{j+1,2k-1}-1} \\ \mu \end{bmatrix} = \begin{bmatrix} \gamma(l2^{-j} - 2^{-(j+1)}) \\ \gamma((l-1)2^{-j} - 2^{-(j+1)}) \\ \dots \\ \gamma((r-1)2^{-j} + 2^{-(j+1)}) \\ 1 \end{bmatrix}, \quad (4)$$

with μ the Lagrange multiplier enforcing the unbiasedness of the estimator.

System (3) appears as a subdivision scheme with the mask:

$$a^{j,2k} : \begin{cases} a_0^{j,2k} = 1, \\ a_m^{j,2k} = 0, \quad m \neq 0, \end{cases}, \quad a^{j,2k-1} : \begin{cases} a_{-2m-1}^{j,2k-1} = \lambda_{j-1,m}^{l_{j,2k-1}, r_{j,2k-1}}, \quad m = -l_{j,2k-1}, \dots, r_{j,2k-1} - 1, \\ a_m^{j,2k-1} = 0 \text{ otherwise.} \end{cases}$$

Given (D, l, r) with $l + r = D + 1$, there are many strategies to define the couple $(l_{j,2k-1}, r_{j,2k-1})$. Among them, one mentions the two following ones that will be considered in this paper. The first one is $l_{j,2k-1} = l, r_{j,2k-1} = r$ that leads to **translation-invariant stencils**. The corresponding subdivision scheme is then stationary and uniform. In the second strategy, $(l_{j,2k-1}, r_{j,2k-1})$ are functions of an a priori defined segmentation of the real line. This leads to **position-dependent stencils** and defines non-stationary and non-uniform subdivision scheme. In this case and since our goal is the approximation of piecewise smooth functions, we focus for the rest of the paper on the following position-dependent stencil selection, associated to a unique segmentation point $\{y_0\}$.

Definition 2.1.

Let us define for all $j \in \mathbb{Z}$, the index k_{j-1} s.t. $y_0 \in [x_{k_{j-1}-1}^{j-1}, x_{k_{j-1}}^{j-1}]$. For all j and k such that $y_0 \in [x_{-l+k}^{j-1}, x_{r-1+k}^{j-1}]$, we set:

- If $y_0 \in [x_{2k_{j-1}-2}^j, x_{2k_{j-1}-1}^j[$, then

$$\begin{cases} \text{If } k < k_{j-1} \text{ then } r_{j,2k-1} = k_{j-1} - k \text{ and } l_{j,2k-1} = D + 1 - k_{j-1} + k, \\ \text{If } k \geq k_{j-1} \text{ then } r_{j,2k-1} = D + 1 + k_{j-1} - k \text{ and } l_{j,2k-1} = k - k_{j-1}. \end{cases}$$
- If $y_0 \in [x_{2k_{j-1}-1}^j, x_{2k_{j-1}}^j]$, then

$$\begin{cases} \text{If } k \leq k_{j-1} \text{ then } r_{j,2k-1} = k_{j-1} - k \text{ and } l_{j,2k-1} = D + 1 - k_{j-1} + k, \\ \text{If } k > k_{j-1} \text{ then } r_{j,2k-1} = D + 1 + k_{j-1} - k \text{ and } l_{j,2k-1} = k - k_{j-1}. \end{cases}$$

Figure 1 displays an example of this selection rule when $D = 3, l = 2$ and $r = 2$.

There are two main advantages of the proposed kriging-based position-dependent prediction. Since it uses kriging, local prediction error is available. This error depends on the local semi-variogram and on the distance between the points involved in the kriging system. Moreover, the data segmentation allows to specify the prediction in each zone of the domain and to take into account the position of the point inside its zone.

3. Convergence analysis of the kriging-based subdivision scheme

We first recall the classical definition of uniform convergence for stationary subdivision schemes ([5]).

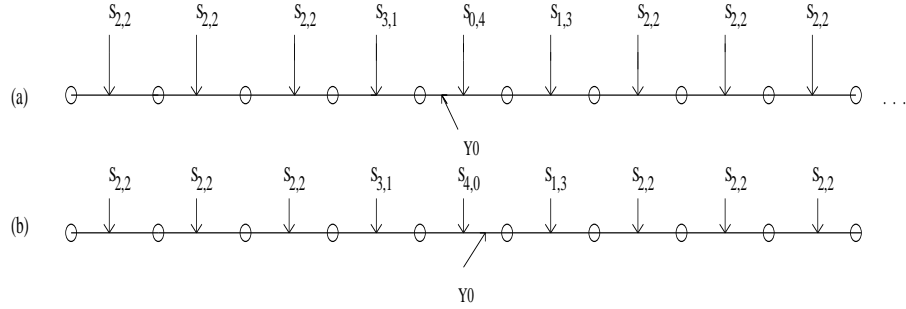


Figure 1: *Example of stencil selection associated to a unique segmentation point $y_0 \in [x_{k_{j-1}-1}^{j-1}, x_{k_{j-1}}^{j-1}]$, $S_{l,r}$ stands for the stencil with l left points and r right points. (a) $y_0 \in [x_{2k_{j-1}-2}^j, x_{2k_{j-1}-1}^j]$, (b) $y_0 \in [x_{2k_{j-1}-1}^j, x_{2k_{j-1}}^j]$.*

Definition 3.1.

The subdivision scheme \mathcal{S} is said to be L^∞ -convergent if for any real sequence $\{f_k^0\}_{k \in \mathbb{Z}} \in V^0$, there exists a continuous function f (called the limit function associated to f^0) such that: $\forall \epsilon, \exists J$ such that $\forall j \geq J, \|\mathcal{S}^j f^0 - f(\frac{\cdot}{2^j})\|_\infty \leq \epsilon$.

In the case of position-dependent subdivision scheme, the previous definition is slightly modified since the adaptation of the prediction around segmentation points can lead to discontinuous limit functions. We refer to [2] for further details.

Among the various tools available to establish the convergence of a subdivision scheme the so called matrix formalism ([7]) has been generalized in [2] to non-uniform schemes. In order to study the convergence of kriging-based subdivision schemes, we propose in what follows an extension to non-stationary prediction.

3.1. General convergence results

Extending [5] and [2], it turns out to be that one can associate to a non-stationary and non-uniform scheme, a set of refinement matrices for a translation-invariant strategy and a set of refinement and edge matrices (due to subdivision around segmentation points) when choosing the previously defined position-dependent scheme. More precisely, we have:

Definition 3.2.

- For a translation-invariant subdivision, let F_k^j be the minimal set of N points at level j that determines the values at dyadic points in the interval $[k2^{-j}, (k+1)2^{-j}]$ at level above j . For each j , the two $N \times N$ refinement matrices, $A_{0,j}$ and $A_{1,j}$ are defined by $F_{2k}^{j+1} = A_{0,j}F_k^j$ and $F_{2k+1}^{j+1} = A_{1,j}F_k^j$.

- Around segmentation points, let G_+^j (resp. G_-^j) be the set of M points at level j that determines the corresponding set of points G_+^{j+1} (resp. G_-^{j+1}) in the right (resp. left) vicinity of y_0 . For each j , the two $M \times M$ -edge matrices $A_{2,j}^+$ and $A_{2,j}^-$ are defined by $G_-^{j+1} = A_{2,j}^-G_-^j$ and $G_+^{j+1} = A_{2,j}^+G_+^j$.

According to [2], the subdivision scheme is then completely characterized in the translation-invariant strategy by the set of refinement matrices $\{A_{0,j}, A_{1,j}\}_{j \in \mathbb{Z}}$. In the position-dependent strategy, it is characterized by the two sets of refinement/edge matrices $\{A_{0,j}, A_{1,j}\}_{j \in \mathbb{Z}} / \{A_{2,j}^-, A_{2,j}^+\}_{j \in \mathbb{Z}}$ and, for any dyadic point $x = k_x 2^{-j_x}$ ($x \neq y_0$), by the integer T_x corresponding to the length of the range of scales where the two successive differences, $\delta f_{k 2^{j-j_x}}^j$ and $\delta f_{k 2^{j-j_x-1}}^j$, are computed mixing refinement and edge matrices (we note $\delta f_k = f_{k+1} - f_k$).

Following [5] and [2] we also introduce, for each j , $A_{0,j}^{(1)}$ and $A_{1,j}^{(1)}$ (resp. $A_{2,j}^{(1),-}$ and $A_{2,j}^{(1),+}$) the refinement (resp. edge) matrices of the subdivision scheme associated to the first difference.

The following convergence theorem then holds:

Theorem 3.1.

Translation-invariant strategy:

If there exists J and μ_0 such that $\forall j \geq J$,

$$\|\Pi_{m \leq j} \frac{1}{2} A_{\epsilon, m}^{(1)}\|_{\infty} \leq \mu_0 < 1, \text{ for all } \epsilon \in \{0, 1\}, \quad (5)$$

then the translation-invariant subdivision scheme is uniformly convergent.

Position-dependent strategy:

If there exists J, μ_0 and T such that $\forall j \geq J$,

$$\|\Pi_{m \leq j} \frac{1}{2} A_{\epsilon, m}^{(1)}\|_{\infty} \leq \mu_0 < 1, \text{ for all } \epsilon \in \{0, 1\}, \quad (6)$$

$$\|\Pi_{m \leq j} \frac{1}{2} A_{2, m}^{(1),-}\|_{\infty} \leq \mu_0 < 1 \quad , \quad \|\Pi_{m \leq j} \frac{1}{2} A_{2, m}^{(1),+}\|_{\infty} \leq \mu_0 < 1, \quad (7)$$

and $\forall x, T_x \leq T < \infty$, then the position-dependent subdivision scheme is uniformly convergent.

3.2. 2 examples of convergence analysis

This section is devoted to the study of the convergence of two kriging-based subdivision schemes with parameters (3, 2, 2). They correspond to two different identified semi-variograms that are used in the numerical tests of Section 4.

3.2.1. Exponential semi-variogram

We assume here that the semi-variogram identified from the data is of exponential type i.e $\gamma(h) = K(1 - e^{-ah})$ (K and $a > 0$). In order to prove the uniform convergence, we start by studying in the two next propositions the asymptotical subdivision scheme obtained to the limit of large values of j . Then, we derive the convergence result for the subdivision scheme itself.

Proposition 3.1.

The kriging weights involved in the translation-invariant and position-dependent subdivision schemes with an exponential semi-variogram satisfy:

$$(\lambda_{j,-2}^{2,2}, \lambda_{j,-1}^{2,2}, \lambda_{j,0}^{2,2}, \lambda_{j,1}^{2,2}) \rightarrow_{j \rightarrow +\infty} (0, \frac{1}{2}, \frac{1}{2}, 0), \quad (8)$$

$$(\lambda_{j,2}^{1,3}, \lambda_{j,1}^{1,3}, \lambda_{j,0}^{1,3}, \lambda_{j,-1}^{1,3}) = (\lambda_{j,-3}^{3,1}, \lambda_{j,-2}^{3,1}, \lambda_{j,-1}^{3,1}, \lambda_{j,0}^{3,1}) \rightarrow_{j \rightarrow +\infty} (0, 0, \frac{1}{2}, \frac{1}{2}), \quad (9)$$

$$(\lambda_{j,3}^{0,4}, \lambda_{j,2}^{0,4}, \lambda_{j,1}^{0,4}, \lambda_{j,0}^{0,4}) = (\lambda_{j,-4}^{4,0}, \lambda_{j,-3}^{4,0}, \lambda_{j,-2}^{4,0}, \lambda_{j,-1}^{4,0}) \rightarrow_{j \rightarrow +\infty} (0, 0, 0, 1). \quad (10)$$

Proof:

Here $D = 3$ and $(r, l) \in \{(2, 2), (3, 1), (4, 0), (1, 3), (0, 4)\}$. We focus on (8), the proofs for (9)-(10) being similar.

Since we are interested in the behavior of the kriging weights when $j \rightarrow \infty$, one writes $\gamma(h) = Kah + O(h^2)$. Then, the kriging system (4) becomes:

$$(\Gamma_{j,4}^0 + \Gamma_{j,4}^1) \Lambda_j = \gamma_{j,4}^0 + \gamma_{j,4}^1, \quad (11)$$

where the matrix $\Gamma_{j,4}^0$ (resp. $\Gamma_{j,4}^1$) stands for the first order (resp. second order) term with regards to h powers.

More precisely,

$$\Gamma_{j,4}^0 = \begin{bmatrix} 0 & a2^{-j} & 2a2^{-j} & 3a2^{-j} & 1 \\ a2^{-j} & 0 & a2^{-j} & 2a2^{-j} & 1 \\ 2a2^{-j} & a2^{-j} & 0 & a2^{-j} & 1 \\ 3a2^{-j} & 2a2^{-j} & a2^{-j} & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}, \quad \gamma_{j,4}^0 = \begin{bmatrix} \frac{3}{2}a2^{-j} \\ \frac{1}{2}a2^{-j} \\ \frac{1}{2}a2^{-j} \\ \frac{3}{2}a2^{-j} \\ 1 \end{bmatrix}.$$

It is elementary to show that:

$$\|\Gamma_{j,4}^1\|_\infty < K_{\Gamma^1} 2^{-2j}, \quad \text{and} \quad \|\gamma_{j,4}^1\|_\infty < K_{\gamma^1} 2^{-2j}. \quad (12)$$

The solution of (11) is then:

$$\Lambda_j = \left(I_4 + \Gamma_{j,4}^{0,-1} \Gamma_{j,4}^1 \right)^{-1} \Gamma_{j,4}^{0,-1} \gamma_{j,4}^0 + \left(I_4 + \Gamma_{j,4}^{0,-1} \Gamma_{j,4}^1 \right)^{-1} \Gamma_{j,4}^{0,-1} \gamma_{j,4}^1, \quad (13)$$

where I_4 is the 4×4 identity matrix. We start by expanding $\left(I_4 + \Gamma_{j,4}^{0,-1} \Gamma_{j,4}^1 \right)^{-1}$.

- Expansion of $\left(I_4 + \Gamma_{j,4}^{0,-1} \Gamma_{j,4}^1 \right)^{-1}$

From a classical result of matrix calculus,

$$\Gamma_{j,4}^{0,-1} = \begin{bmatrix} B^{-1} (I_3 + \mathbb{1}' S^{-1} \mathbb{1} B^{-1}) & -B^{-1} \mathbb{1}' S^{-1} \\ -S^{-1} \mathbb{1} B^{-1} & S^{-1} \end{bmatrix}.$$

with

$$B = 2^{-j} \begin{bmatrix} 0 & a & 2a & 3a \\ a & 0 & a & 2a \\ 2a & a & 0 & a \\ 3a & 2a & a & 0 \end{bmatrix}, \quad \mathbb{1} = (1, 1, 1, 1) \text{ and } S = -\mathbb{1}B^{-1}\mathbb{1}'.$$

From these definitions, it is then straightforward that

$$B^{-1}(I_3 + \mathbb{1}'S^{-1}\mathbb{1}B^{-1}) = 2^j \mathcal{M}_1, \quad B^{-1}\mathbb{1}'S^{-1} = \mathcal{V}_2, \quad (S^{-1}\mathbb{1}B^{-1})' = \mathcal{V}_3, \quad S^{-1} = K_4 2^{-j},$$

where $\mathcal{M}_1, \mathcal{V}_2, \mathcal{V}_3$ and K_4 are independent of j .
Therefore, $\|\Gamma_{j,4}^{0,-1}\|_\infty \leq K_{\Gamma^0} 2^j$. Combined with (12) it gives

$$\|\Gamma_{j,4}^{0,-1}\Gamma_{j,4}^1\|_\infty \leq K_{\Gamma^0} K_{\Gamma^1} 2^{-j}, \quad (14)$$

and finally, for sufficiently large j $(I_4 + \Gamma_{j,4}^{0,-1}\Gamma_{j,4}^1)^{-1} = \sum_{i=0}^{\infty} (-1)^i (\Gamma_{j,4}^{0,-1}\Gamma_{j,4}^1)^i$.

Expression (13) becomes:

$$\Lambda_j = \Gamma_{j,4}^{0,-1}\gamma_{j,4}^0 + \left(\sum_{i=1}^{\infty} (-1)^i (\Gamma_{j,4}^{0,-1}\Gamma_{j,4}^1)^i \Gamma_{j,4}^{0,-1}\gamma_{j,4}^0 + \sum_{i=0}^{\infty} (-1)^i (\Gamma_{j,4}^{0,-1}\Gamma_{j,4}^1)^i \Gamma_{j,4}^{0,-1}\gamma_{j,4}^1 \right). \quad (15)$$

Note that this step also provides the following inequalities that are useful in the sequel:

$$\|\Gamma_{j,4}^{0,-1}\gamma_{j,4}^0\|_\infty \leq \mathcal{K}_0, \quad \|\Gamma_{j,4}^{0,-1}\gamma_{j,4}^1\|_\infty \leq \mathcal{K}_1 2^{-j}. \quad (16)$$

Coming back to (15), in order to conclude the proof, we now show that the norm of second second term tends to 0 when $j \rightarrow +\infty$, that the first term does not depend on j and that its four first components are equal to $(0, \frac{1}{2}, \frac{1}{2}, 0)$.

Let us first focus on the second term of (15). From (14) and (16), it is straightforward that:

$$\begin{aligned} \left\| \sum_{i=1}^{\infty} (-1)^i (\Gamma_{j,4}^{0,-1}\Gamma_{j,4}^1)^i \Gamma_{j,4}^{0,-1}\gamma_{j,4}^0 \right\|_\infty &\leq \mathcal{K}_0 \sum_{i=1}^{\infty} (K_{\Gamma^0} K_{\Gamma^1})^i 2^{-ij}, \\ \left\| \sum_{i=0}^{\infty} (-1)^i (\Gamma_{j,4}^{0,-1}\Gamma_{j,4}^1)^i \Gamma_{j,4}^{0,-1}\gamma_{j,4}^1 \right\|_\infty &\leq \mathcal{K}_1 2^{-j} \sum_{i=0}^{\infty} (K_{\Gamma^0} K_{\Gamma^1})^i 2^{-ij}, \end{aligned}$$

which becomes for sufficiently large j ,

$$\begin{aligned} \left\| \sum_{i=1}^{\infty} (-1)^i \left(\Gamma_{j,4}^{0,-1} \Gamma_{j,4}^1 \right)^i \Gamma_{j,4}^{0,-1} \gamma_{j,4}^0 \right\|_{\infty} &\leq \mathcal{K}_0 K_{\Gamma^0} K_{\Gamma^1} \frac{2^{-j}}{1 - K_{\Gamma^0} K_{\Gamma^1} 2^{-j}}, \\ \left\| \sum_{i=0}^{\infty} (-1)^i \left(\Gamma_{j,4}^{0,-1} \Gamma_{j,4}^1 \right)^i \Gamma_{j,4}^{0,-1} \gamma_{j,4}^1 \right\|_{\infty} &\leq \mathcal{K}_1 \frac{2^{-j}}{1 - K_{\Gamma^0} K_{\Gamma^1} 2^{-j}}. \end{aligned}$$

Therefore

$$\lim_{j \rightarrow \infty} \left\| \sum_{i=1}^{\infty} (-1)^i \left(\Gamma_{j,4}^{0,-1} \Gamma_{j,4}^1 \right)^i \Gamma_{j,4}^{0,-1} \gamma_{j,4}^0 + \sum_{i=0}^{\infty} (-1)^i \left(\Gamma_{j,4}^{0,-1} \Gamma_{j,4}^1 \right)^i \Gamma_{j,4}^{0,-1} \gamma_{j,4}^1 \right\|_{\infty} = 0.$$

For the first term of (15), since by construction $\Gamma_{j,4}^0$ is invertible, the vector U is solution of $\Gamma_{j,4}^0 U = \gamma_{j,4}^0$. Direct calculation gives $U = (0, \frac{1}{2}, \frac{1}{2}, 0, 0)$, that concludes the proof. \blacksquare

Using (8)-(10) one gets the following result:

Proposition 3.2.

The matrices involved in the asymptotical subdivision process are:

1) *Refinement matrices:*

$$A_0 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

with $Sp(A_0) = Sp(A_1) = \{1, \frac{1}{2}, 0\}$.

2) *Edge matrices: their expression depend on the dyadic decomposition of y_0 ,*

$$A_2^- \in \left\{ \left[\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{bmatrix} \right\}, \quad A_2^+ \in \left\{ \left[\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \right\}.$$

In any case, $Sp(A_2^+) \subset \{1, \frac{1}{2}, 0\}$ and $Sp(A_2^-) \subset \{1, \frac{1}{2}, 0\}$.

Proposition 3.2 finally leads to the convergence result.

Proposition 3.3.

The translation-invariant and position-dependent subdivision schemes of parameter (3, 2, 2) associated to an exponential semi-variogram are uniformly convergent.

Proof:

Following [5], the eigenvalues of $\frac{1}{2}A_0^{(1)}$ and $\frac{1}{2}A_1^{(1)}$ (resp. $\frac{1}{2}A_2^{(1),-}$ and $\frac{1}{2}A_2^{(1),+}$) are the eigenvalues of A_0 and A_1 (resp. A_2^- and A_2^+) except $\lambda_0 = 1$. Thus, from Proposition 3.2, these eigenvalues are strictly bounded by 1. This is enough to ensure Inequalities (5), (6) and (7) for the asymptotical subdivision scheme and therefore for the subdivision scheme. Theorem 3.1 finally leads to the uniform convergence of the translation-invariant scheme but also to the position-dependent one since the extra assumption related to the uniform bound for the transition zone is satisfied following Proposition 2.3 of [2]. ■

3.2.2. Gaussian semi-variogram

We assume here that the semi-variogram identified from the data is $\gamma(h) = K \left(1 - e^{-(ah)^2}\right)$ (K and $a > 0$). As previously, the uniform convergence is established by first studying the asymptotical behavior of the scheme.

Proposition 3.4.

The kriging weights involved in the translation-invariant and position-dependent subdivision schemes with a Gaussian semi-variogram satisfy:

$$\begin{aligned} (\lambda_{j,-2}^{2,2}, \lambda_{j,-1}^{2,2}, \lambda_{j,0}^{2,2}, \lambda_{j,1}^{2,2}) &\rightarrow_{j \rightarrow +\infty} \left(-\frac{1}{16}, \frac{9}{16}, \frac{9}{16}, -\frac{1}{16}\right), \\ (\lambda_{j,2}^{1,3}, \lambda_{j,1}^{1,3}, \lambda_{j,0}^{1,3}, \lambda_{j,-1}^{1,3}) &= (\lambda_{j,-3}^{3,1}, \lambda_{j,-2}^{3,1}, \lambda_{j,-1}^{3,1}, \lambda_{j,0}^{3,1}) \rightarrow_{j \rightarrow +\infty} \left(\frac{1}{16}, -\frac{5}{16}, \frac{15}{16}, \frac{5}{16}\right), \\ (\lambda_{j,3}^{0,4}, \lambda_{j,2}^{0,4}, \lambda_{j,1}^{0,4}, \lambda_{j,0}^{0,4}) &= (\lambda_{j,-4}^{4,0}, \lambda_{j,-3}^{4,0}, \lambda_{j,-2}^{4,0}, \lambda_{j,-1}^{4,0}) \rightarrow_{j \rightarrow +\infty} \left(-\frac{5}{16}, \frac{21}{16}, -\frac{35}{16}, \frac{35}{16}\right). \end{aligned}$$

Proof:

The proof is the same as in the exponential case taking into account that $\gamma(h) = K(ah)^2 - \frac{(ah)^4}{2} + O(h^6)$. Note that we consider here the Taylor series up to degree 4, otherwise the matrix involved in the first order kriging system is not invertible. ■

From the previous proposition, the asymptotical kriging weights correspond to the coefficients of the mask of degree 3 Lagrange interpolating subdivision scheme ([2]) with the same translation-invariant and position-dependent strategies. This scheme has been proved to uniformly converge in [2]. As previously, this is enough to ensure the uniform convergence for our scheme.

Convergence result for exponential and Gaussian-type subdivision schemes of parameter $(3, 2, 2)$ can be generalized. In a forthcoming paper, we show that the convergence is still ensured in the case of a scheme of parameter (D, r, r) for any $(D, r) \in \mathbb{N}^2$ with $r \leq D$.

4. Numerical tests

This section is devoted to three numerical tests. Here, the question of the accuracy of the semi-variogram identification, that refers to cross-validation techniques ([4]), is not addressed even though it remains a crucial point for data representation. We start by illustrating the capability of position-dependent kriging-based subdivision schemes on a synthetic example. Then, we show applications to real cases.

4.1. Synthetic example

Figure 2, top, left, displays a synthetic set of data coming from a regular sampling ($\Delta x = 2^{-10}$) of the test function:

$$\begin{cases} 15000 * \frac{\sin(-5+20x)}{-5+20x}, & \text{if } x \in \text{Zone 1} = [0; 0.5] \\ 2500 * \frac{\sin(1.5 * (-5+10(x-1/2)))}{-5+20(x-1/2)} + 300, & \text{if } x \in \text{Zone 2} =]0.5; 1] \end{cases}$$

Assuming that some data (cross signs) are missing, our goal is to evaluate the capabilities of our models to accurately predict this missing information.

We provide a comparison of the results obtained with the translation invariant and the position-dependent kriging-based subdivision schemes with parameters $D = 3$, $l = 2$, $r = 2$ and $y_0 = 0.5$.

The experimental and theoretical semi-variograms for each zone (position-dependent strategy) and for the whole data are plotted on Figure 2. They have been obtained using the statistical software SUNSET developed at IRSN ([3]). Note that in each case the semi-variogram exhibits a Gaussian-type structure but with different parameters.

Figure 3 shows the predicted data while Table 1 exhibits the l^2 -error and the average estimation variances.

Method	l^2 -error	Average estimation variance
TI	38.1	5.1
PD	0.08	Zone 1: 7, Zone 2: 0.2

Table 1: l^2 -error and average estimation variance for a Translation-Invariant (TI) and a Position-Dependent (PD) one.

These results point out two main advantages of position-dependent kriging-based approaches:

- The position-dependent scheme outperforms the translation-invariant one in term of prediction. According to Figure 3, it allows to remove oscillations (or Gibbs phenomenon) around segmentation point.

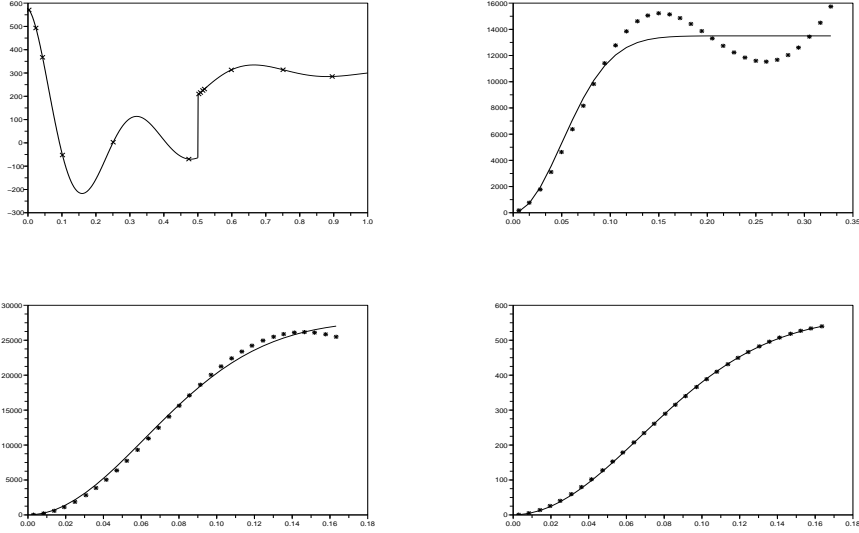


Figure 2: Synthetic example. Top, left, synthetic data with missing information (cross signs). From top, right to bottom right, experimental (star) and theoretical (continuous line) semi-variograms for the whole data, $\gamma(h) = 2.7(1 - e^{-\frac{1}{0.7^2}h^2})$, for Zone 1, $\gamma(h) = 5.58(1 - e^{-\frac{1}{0.08^2}h^2})$ and for Zone 2, $\gamma(h) = 0.1(1 - e^{-\frac{1}{0.09^2}h^2})$.

- As it can be seen in Table 1, the zone-dependent strategy drastically affects the computation of the average estimation variance. Keeping in mind that the estimation variance is related to an a priori prediction error, a refined estimation of this quantity is a real improvement of the method.

4.2. Applications to environmental data

4.2.1. Rainfall measurements

Figure 4, top, left, displays the yearly rainfall measurements of the Sahelian zone in Gouré, Niger, Africa between 1935 and 1999 (digitalized from [8]). As in the previous example, we assume that some measurements (cross signs) are missing. The data of Figure 4 exhibit two main behaviors which are confirmed by analysts: a wet period before 1967 (Zone 1) and a dry period after (Zone 2)

Following the previous approach, we compare a global kriging prediction to a two zone position-dependent one. Figure 4 displays a zoom of the predicted data with the corresponding a priori prediction error (computed for each predicted point f_k^j as $I = [f_k^j - 2\sigma_k^j; f_k^j + 2\sigma_k^j]$ with σ_k^j the square root of the estimation variance).

It appears that the position-dependent approach always leads to confidence

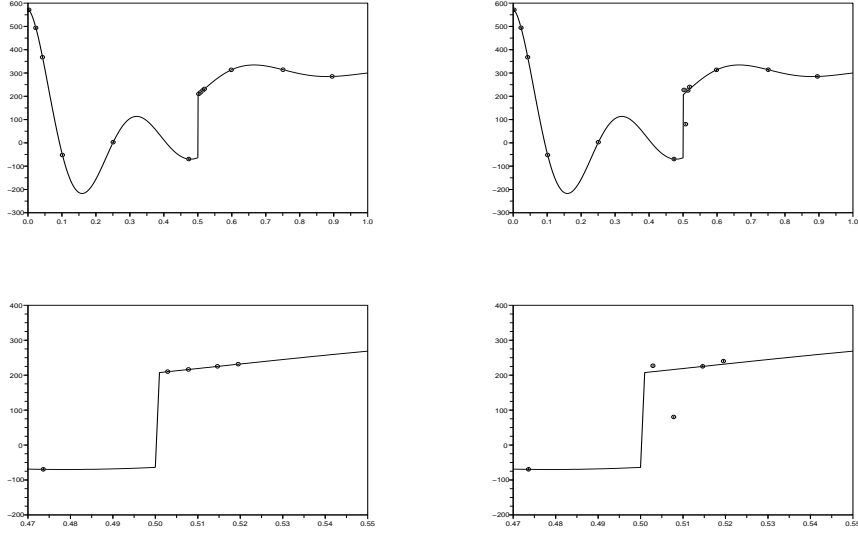


Figure 3: Predicted data. Top, Predicted data (circle) and exact measurements (continuous line) with a position-dependent (left) and translation-invariant (right) strategy. Bottom: zoom around the segmentation point.

intervals encompassing the true value of the measurement, in opposition to what happens with the translation-invariant strategy.

4.2.2. Simplified example of radioactive contamination mapping

We consider here a simplified example of a radio-nuclide concentration map in a nuclear test area. The available data are given on a fixed dyadic grid of grid step 2^{-4} .

According to the measurements (Figure 5, top, left), high concentrations of radio-nuclide are localized around ground zero area (i.e. where tests have been conducted, Zone 1) whereas there is a drastic decrease in the concentration for further measurement points (Zone 2).

Using an alternate direction strategy, a bivariate position dependent kriging based prediction has been defined and used. After identifying two semi-variograms, we provide on Figure 5 the concentration map provided by our prediction scheme up to scale $J = 6$. In order to compare with translation-invariant strategy and since using a unique semi-variogram is not satisfactory due to strong non-regularity, we compare our scheme to the translation-invariant Lagrange interpolating scheme classically used in subdivision process ([5], [1]).

It comes out that the removal of the Gibbs phenomenon, when performing a position-dependent strategy avoids the unrealistic negative concentrations obtained when following a translation-invariant strategy. One notes however that

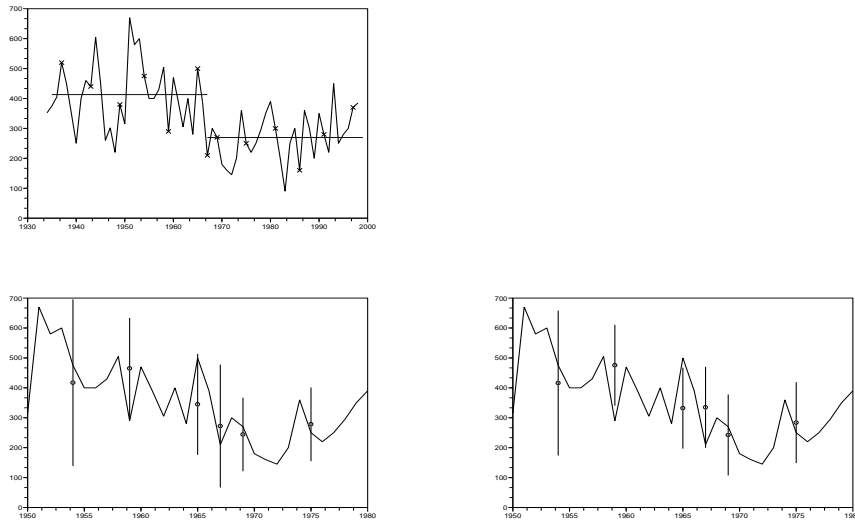


Figure 4: Rainfall measurements. Top, measurements, average rainfall for each zone and missing data (cross signs). Bottom, zoom of the predicted rainfall measurements. Predicted data (circle) and exact measurements (continuous line) with prediction error (vertical line): left, position-dependent strategy, right, translation-invariant strategy.

the model provides a bigger standard deviation along the segmentation curve, due to the extrapolation performed in the vicinity of this curve.

5. Conclusion

A new kriging-based subdivision scheme adapted to data segmentation has been introduced and fully analyzed. It is efficient for the prediction of non-regular data since it integrates a position-dependent mask taking into account the information given by the segmentation. It also provides an a priori prediction error which is relevant in risk analysis. It has been successfully applied to two environmental problems. Further developments concern the extension of this approach to non-dyadic and non-regular data grid point and its integration in a complete risk analysis study.

References

- [1] F. Arandiga, R. Donat, and A. Harten. Multiresolution based on weighted averages of the hat function I: linear reconstruction techniques. *SIAM J. Sci. Comput.*, 36(1):160–203, 1999.
- [2] J. Baccou and J. Liandrat. Position-dependent lagrange interpolating multiresolutions. *Int. J. of Wavelets, Multiresol. and Inf.*, 5(4):513–539, 2005.

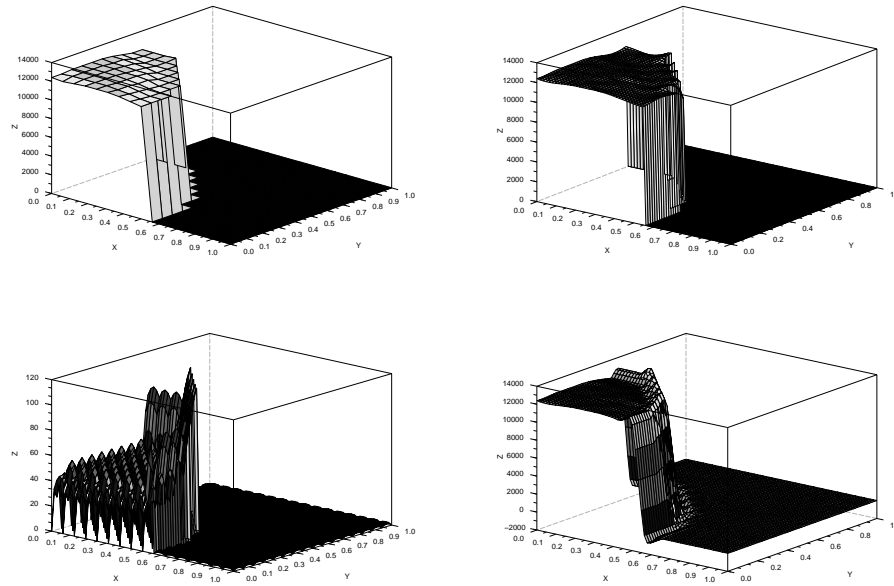


Figure 5: Mapping of radio-nuclide concentration. Top left, initial mapping (available data) and (Top right) predicted concentration map by position-dependent strategy. Bottom left, estimation standard deviation associated to the position-dependent prediction and (Bottom right) concentration map provided by the classical interpolating Lagrange-based prediction.

- [3] E. Chojnacki and A. Ounsy. Description of the ipsn method for the uncertainty and sensitivity analysis and the associated software: Sunset. *Proc. of ASME/JSME ICONE, Louisiana*, 3:545–550, 1996.
- [4] N.A. Cressie. *Statistics for spatial data*. Wiley Series in Probability and Mathematical Statistics, 1993.
- [5] N. Dyn. Subdivision schemes in computer-aided geometric design. In W.A Light, editor, *Advances in Numerical analysis II, Wavelets, Subdivision algorithms and Radial Basis functions*. Clarendon Press, Oxford, 1992.
- [6] A. Harten. Multiresolution representation of data: a general framework. *SIAM J. Numer. Anal.*, 33(3):1205–1256, 1996.
- [7] C.A. Miccheli and H. Prautzsch. Uniform refinement of curves. *Linear Algebra and Applications*, 114/115:841–870, 1989.
- [8] P. Ozer, C. Bodart, and B. Tychon. Climatic analysis of the gouré area, eastern niger: recent changes and environmental impacts. *Cybergeo, Environnement, Nature, Paysage*, 308, 2007.