The Annals of Statistics
1991, Vol. 19, No. 3, 1512–1527

SPLINE FUNCTIONS AND STOCHASTIC FILTERING

BY CHRISTINE THOMAS-AGNAN

Université Paul Sabatier (Toulouse III) and Université des Sciences Sociales (Toulouse I)

Some relationships have been established between unbiased linear predictors of processes, in signal and noise models, minimizing the predictive mean square error and some smoothing spline functions. We construct a new family of multidimensional splines adapted to the prediction of locally homogeneous random fields, whose "m-spectral measure" (to be defined) is absolutely continuous with respect to Lebesgue measure and satisfies some minor assumptions. By considering partial splines, one may include an arbitrary drift in the signal. This type of correspondence underlines the potentialities of cross-fertilization between statistics and the numerical techniques in approximation theory.

1. Introduction. A frequent problem in many sciences is to reconstruct a function f: $\mathbb{R}^d \mapsto \mathbb{R}$ from possible noisy measurements y_i of $f(t_i)$ at a finite number of given irregularly spaced data points t_i, $i = 1, \ldots, n$, of $\mathbb{R}^d$ $d \in \mathbb{N}^*$. The smoothing spline functions are among the purely numerical techniques developed for this approximation problem. On the other hand, some prefer to explain the data with a stochastic model by considering that the observations y_i constitute one realization of a stochastic process $Y(t)$, observed at discrete points t_i, blurred by an additive noise ε_i. By making additional assumptions on this model and using a given prediction technique, one gets a predicted random process $\hat{Y}(t)$ which depends upon the random variables $Y(t_i)$; substituting the data y_i for $Y(t_i)$ then gives the desired approximation of f. But it turns out that some connections have been observed between these techniques in the sense that they yield the same solution in certain conditions: see Dolph and Woodbury (1952), Dubrule (1983), Heckman (1986), Kimeldorf and Wahba (1970a, b), Kohn and Ansley (1983), Matheron (1981), Salkauskas (1982), Thomas-Agnan (1987) and Watson (1983). This establishes a natural correspondence between types of concrete splines and the prediction by certain techniques of some types of processes. In some of these studies, the predicted random process is obtained by Bayes estimation, that is, $\hat{Y}(t) = E(Y(t)|Y_1, \ldots, Y_n)$, where Y_i is the random variable $Y(t_i) + \varepsilon_i$. In other works, the predicted random process is obtained by the best linear unbiased estimation method (BLUE hereafter). These two predictors coincide in the Gaussian case, and we will mainly focus on the second point of view here.

Received July 1989; revised May 1990.
AMS 1980 *subject classifications*. Primary 62M20; secondary 65D07, 62M15.
Key words and phrases. Spline functions, interpolation, smoothing, partial splines, inf-convolution splines, reproducing kernel, locally homogeneous random fields, variogram, kriging.

In Section 2, we construct a new family of concrete multidimensional splines called α-splines. We use the Fourier transform, in suitable spaces, to build a measure of smoothness for functions and we give explicit formulas for their computation. In Section 3, we present the stochastic filtering technique and recall some facts about locally homogeneous random fields. In Section 4, we prove the correspondence between partial α-splines and the prediction of a large class of processes connected to locally homogeneous random fields. We discuss a few examples [details can be found in Thomas-Agnan (1989)]. Finally the Appendix presents the technical proofs of some lemmas used in Section 2.

While this better understanding is already satisfactory from a theoretical point of view, it could also lead to a cross-fertilization of these techniques [see Kohn and Ansley (1987)].

2. Construction of α-splines.

2.1. *Abstract partial splines.* There are several types of multidimensional spline functions. We will restrict attention here to those which can be obtained as the solution of a quadratic optimization problem in a Hilbert space. Moreover, there are several types of partial splines. They are generally decomposed additively into a parametric and a nonparametric component, these two functions depending upon $t \in \mathbb{R}^d$ through the whole or only part of the variables $(t_1, \ldots, t_d)$ [see Heckman (1986), Laurent (1980), Wahba (1986) and Shiau, Wahba and Johnson (1986)]. We will use here the following definitions from Laurent (1988).

The function f to be approximated will belong to the space $\mathscr{E}$ of functions from $\mathbb{R}^d$ to $\mathbb{R}$, equipped with the pointwise convergence topology. The spline function we consider to approximate f will be the solution of one of the following quadratic optimization problems. Let H be a subspace of $\mathscr{E}$ equipped with a structure of Hilbert space (on $\mathbb{R}$ or $\mathbb{C}$), its scalar product being denoted by $(\cdot, \cdot)$. Let $L_1, \ldots, L_n$ be n linearly independent continuous functionals on $\mathscr{E}$, $Y = (y_1, \ldots, y_n)$ be a vector of $\mathbb{R}^n$, λ a positive real number and J be a seminorm on H. Let $w_1, \ldots, w_K$ be K, $K \in \mathbb{N}$, given linearly independent functions on $\mathbb{R}^d$. Then the interpolation problem is

$$\text{(1)} \qquad \begin{aligned} &\operatorname{Min} J(h), \\ &h \in H,\ \theta_k \in \mathbb{R},\ k = 1, \ldots, K,\ L_i\Big(h + \sum \theta_k w_k\Big) = y_i,\ i = 1, \ldots, n. \end{aligned}$$

The smoothing problem is

$$\text{(2)} \qquad \begin{aligned} &\operatorname{Min} \sum_{i=1}^{n} \Big(L_i\Big(h + \sum \theta_k w_k\Big) - y_i\Big)^2 + \lambda J(h), \\ &h \in H,\ \theta_k \in \mathbb{R},\ k = 1, \ldots, K. \end{aligned}$$

Under additional assumptions on $J, L_1, \ldots, L_n, w_1, \ldots, w_K$, Laurent (1988) proves the existence and uniqueness of the solution of these optimization problems; the corresponding solutions will be called interpolating partial splines for (1) and smoothing partial splines for (2) ("inf-convolution" splines in France). Similar definitions without the parametric part $\sum \theta_k v_k$ lead to ordinary (nonpartial) splines. Usually, the seminorm $J(\cdot)$ is a smoothness criterion and problem (2) expresses the balance between the smoothness of the spline and its fit to the data.

The data y_i in our problem being measurements of $f(t_i)$, we assume moreover that H is a reproducing kernel Hilbert space; we recall that it is a Hilbert space of functions on $\mathbb{R}^d$ in which, for all t of $\mathbb{R}^d$, there exists an element $R_t(\cdot)$ in H, called the representer of evaluation at t, such that for all f in H, $f(t) = (R_t(\cdot), f(\cdot))$ and that the kernel of H is the function R: $\mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}$ defined by

$$R(s,t) = (R_t, R_s) = R_t(s) = R_s(t). \tag{3}$$

The concrete splines are obtained by choosing specific Hilbert spaces and seminorms. In particular, we recall two important cases. The thin-plate splines (or multidimensional D^m-splines) are constructed in Beppo–Levi spaces with the seminorms:

$$\sum_{|\beta|=m} \binom{m}{\beta} \| D^\beta h \|^2_{L_2}, \tag{4}$$

where for a multiindex $\beta = (\beta_1, \ldots, \beta_d) \in \mathbb{N}^d$, we denote by $|\beta|$ the sum $\sum \beta_i$, by $\binom{m}{\beta}$ the multinomial coefficient $m!/\Pi\beta_i!$ and by D^β the differential operator $\partial^{|\beta|}/\partial x_1^{\beta_1} \ldots \partial x_d^{\beta_d}$ [see Meinguet (1979)]. If L is a linear differential operator, L-splines are constructed with the seminorms $\|Lh\|^2_{L_2}$ [see Kimeldorf and Wahba (1971)].

2.2. *Criterion of smoothness.* Some authors [see Duchon (1977) and Klonias (1984)] have used the idea of building a smoothness measure for a function in suitable spaces on the asymptotic behavior of its Fourier transform (hereafter denoted by $\mathcal{F}$). For periodic functions, Thomas-Agnan (1990b) introduces seminorms which are weighted l_2-norms of the sequence of Fourier coefficients, leading to a generalization of periodic D^m-splines. A similar generalization of Duchon's splines in dimension 1 is developed in Thomas-Agnan (1990a), where the smoothness of a function f is measured by a weighted L_2-norm of $\mathcal{F}f$. Our goal here is to construct a multidimensional version of these splines. This new family of concrete splines called α-splines will include as particular cases: the thin-plate splines (or multidimensional D^m-splines) and some L-splines.

The first challenge is to specify the Hilbert spaces adapted to work with this type of seminorms. Even though we want to deal with functions of $\mathcal{E}$ which are at least continuous, it appears more convenient for the construction of these spaces to consider that they belong to the space $\mathcal{S}'(\mathbb{R}^d)$ of Schwartz

tempered distributions (real-valued here) for which a theory of Fourier transform is known [see Schwartz (1966)].

2.3. *Construction of the spaces.* We will say that an element h of $\mathcal{S}'(\mathbb{R}^d)$ is represented by an element l of $\mathcal{E}$ (or alternatively that l is a representer of h) whenever the map which associates to an element φ of $\mathcal{S}(\mathbb{R}^d)$ (the space of infinitely differentiable functions with fast decay at ∞) the real number $\int \varphi(t)l(t)\,dt$, defines an element of $\mathcal{S}'(\mathbb{R}^d)$ which coincides with h. We denote by z^* the complex conjugate of a complex number z. Let m be an integer and α be a function from $\mathbb{R}^d$ to $\mathbb{C}$ verifying the following assumptions:

(A1) $\alpha(\omega) \neq 0$ for all $\omega \in \mathbb{R}^d$.
(A2) $\alpha(\omega) = \alpha(-\omega)^*$.
(A3) $|\alpha|^{-2}$ is locally integrable.
(A4) There exists a ball B of $\mathbb{R}^d$ centered at 0, such that $\omega \to |\alpha(\omega)|^{-2}\|\omega\|^{-2m}$ belongs to $L_1(\mathbb{R}^d \setminus B)$.

Then let $\mathcal{E}_{m\alpha}$ be the space of real-valued elements h of $\mathcal{S}'(\mathbb{R}^d)$ such that, for all $\beta \in \mathbb{N}^d$ satisfying $|\beta| = m$, $\mathcal{F}(D^\beta h)$ is a locally integrable function and $\alpha\mathcal{F}(D^\beta h)$ belongs to $L_2(\mathbb{R}^d)$.

LEMMA 2.3.1. *Under assumptions* (A1) *through* (A4), *any element of* $\mathcal{E}_{m\alpha}$ *is represented by a continuous function.*

PROOF. There exists an infinitely differentiable function ξ vanishing in a neighborhood of 0 and equal to 1 on $\mathbb{R}^d \setminus B$. By (A4), for h in $\mathcal{E}_{m\alpha}$ and any β such that $|\beta| = m$, we have $\xi(\omega)\omega^{-\beta}\mathcal{F}(D^\beta h)(\omega) \in L_1(\mathbb{R}^d)$. Using the following property of Fourier transforms [see Schwartz (1966)]:

$$\mathcal{F}(D^\beta h)(\omega) = (2\pi i\omega)^\beta \mathcal{F}h(\omega), \tag{5}$$

where ω^β denotes $\prod_{i=1}^d \omega_i^{\beta_i}$, one gets that the map $\omega \mapsto \xi(\omega)\mathcal{F}h(\omega) \in L_1(\mathbb{R}^d)$, and therefore $\mathcal{F}(\xi\mathcal{F}h)$ is represented by a continuous function. On the other hand, $(1 - \xi(\omega))\mathcal{F}h(\omega))$, being compactly supported, has a Fourier transform represented by a continuous function. Hence, by reciprocity of $\mathcal{F}$ in $\mathcal{S}'(\mathbb{R}^d)$, $h = \mathcal{F}(\xi\mathcal{F}h + (1 - \xi)\mathcal{F}h)$ is also represented by a continuous function. □

We will identify from now on any element of $\mathcal{E}_{m\alpha}$ with its continuous representer. By a similar argument, one can establish a connection between the growth rate of $\|\omega\|^{-m}|\alpha(\omega)|^{-1}$ at ∞ and the number of square integrable derivatives of elements of $\mathcal{E}_{m\alpha}$.

LEMMA 2.3.2. *Under assumptions* (A1) *through* (A4),

$$J(h) = \sum_{|\beta|=m} \binom{m}{\beta} \|\alpha\mathcal{F}(D^\beta h)\|_{L_2} \tag{6}$$

defines a seminorm in $\mathscr{E}_{m\alpha}$ *whose kernel is the space* $\mathscr{P}_m$ *of polynomials on* $\mathbb{R}^d$ *whose total degree is less than or equal to* $(m-1)$.

PROOF. It clearly follows from (A1) and the fact that an element of $\mathscr{S}'(\mathbb{R}^d)$ satisfies $D^\beta h = 0$ for all β such that $|\beta| = m$ if and only if $h \in \mathscr{P}_m$ [see Schwartz (1966)]. Note that the dimension of $\mathscr{P}_m$ is $M = \binom{d+m-1}{d}$. □

THEOREM 1. *Under assumptions* (A1) *through* (A4), $\mathscr{E}_{m\alpha}$ *is a reproducing kernel Hilbert space.*

For the proof of Theorem 1, we need the following definition [see Meinguet (1979)].

DEFINITION. A set of $M = \binom{d+m-1}{d}$ points of $\mathbb{R}^d$: $x_1, \ldots, x_M$ such that for all $(\gamma_1, \ldots, \gamma_M)$ in $\mathbb{R}^M$, there exists a P in $\mathscr{P}_m$ such that $P(x_i) = \gamma_i$, $i = 1, \ldots, m$, is called a $\mathscr{P}_m$ unisolvent set.

To construct a scalar product in $\mathscr{E}_{m\alpha}$, we will use a $\mathscr{P}_m$ unisolvent set $x_1, \ldots, x_M$ (there always exists one). Its choice is arbitrary and the reproducing kernel of $\mathscr{E}_{m\alpha}$ is independent of this set, even though it occurs in the formula giving the kernel.

PROOF. The technical proofs of Lemmas 2.3.3 to 2.3.7 are to be found in the Appendix. Using Lemma 2.3.2 and the above definition, it is clear that, if u and v are elements of $\mathscr{E}_{m\alpha}$,

$$(7) \quad (u,v)_{m\alpha} = \sum_{i=1}^{M} u(x_i)v(x_i) + \sum_{|\beta|=m} \binom{m}{\beta} \left(\alpha\mathscr{F}(D^\beta u), \alpha\mathscr{F}(D^\beta v)\right)_{L_2}$$

defines a scalar product in $\mathscr{E}_{m\alpha}$. Let $\mathscr{E}^1_{m\alpha}$ be the set of elements of $\mathscr{E}_{m\alpha}$ such that $h(x_i) = 0$ for all $i = 1, \ldots, M$. Being the orthogonal of $\mathscr{P}_m$ in $\mathscr{E}_{m\alpha}$, for the above inner product, $\mathscr{E}^1_{m\alpha}$ is closed and it is clear that $\mathscr{E}_{m\alpha}$ is the direct sum of $\mathscr{P}_m$ and $\mathscr{E}^1_{m\alpha}$. Hence, $\mathscr{E}^1_{m\alpha}$ is isomorphic to $\mathscr{E}_{M\alpha}/\mathscr{P}_m$ which is complete, being a Beppo–Levi space [see Deny and Lions (1954)]. Therefore, $\mathscr{E}_{m\alpha}$ is complete and it remains to establish explicit formulas for the representers of evaluation in each subspace (R^0_t for $\mathscr{P}_m$ and R^1_t for $\mathscr{E}^1_{m\alpha}$). The representer R_t will then be the sum $R^0_t + R^1_t$. In the finite-dimensional subspace $\mathscr{P}_m$, by unisolvence of $x_1, \ldots, x_M$, there exist M polynomials P_j such that $P_j(x_i) = \delta_{ij}$, where δ_{ij} is Kronecker's symbol, for all i and $j = 1, \ldots, M$. It follows that

$$(8) \qquad R^0_t(s) = \sum_{i=1}^{M} P_i(t)P_i(s) = R^0_s(t) = R^0(s,t).$$

Let us recall some classical facts [see, e.g., Schwartz (1966)]. Let Δ^m denote the mth iterated Laplacian

$$\Delta^m = \sum_{|\beta|=m} \binom{m}{\beta} D^{2\beta}.$$

In $\mathscr{S}'(\mathbb{R}^d)$, the solutions of

$$\Delta^m u = 0, \qquad u \in \mathscr{S}'(\mathbb{R}^d) \tag{9}$$

are necessarily polynomials of total degree less than or equal to $2m$. They are called polyharmonic polynomials and in particular, elements of $\mathscr{P}_{2m}$ satisfy (9). There exists in $\mathscr{S}'(\mathbb{R}^d)$ solutions E of

$$\Delta^m E = \delta, \tag{10}$$

where δ is Dirac's distribution at 0. They are called fundamental solutions of the iterated Laplacian and are equivalently solutions in $\mathscr{S}'(\mathbb{R}^d)$ of

$$\|\omega\|^{2m} \mathscr{F}E(\omega) = 1 \quad \text{for all } \omega \in \mathbb{R}^d. \tag{11}$$

Any two solutions of (10) or (11) differ by a polyharmonic polynomial and a particular solution of (10) is given by

$$E_m(x) = \gamma_{m,d} \|x\|^{2m-d} \ln(\|x\|), \tag{12}$$

when $2m \geq d$ and d even, for some known proportionality constant $\gamma_{m,d}$, and by

$$E_m(x) = \mu_{m,d} \|x\|^{2m-d}, \tag{13}$$

otherwise, for some known proportionality constant $\mu_{m,d}$. From now on, let us denote by E_m any solution of (12).

LEMMA 2.3.3. *There exists in* $\mathscr{S}'(\mathbb{R}^d)$ *solutions of*

$$\|\omega\|^{2m} \mathscr{F}E(\omega) = |\alpha(\omega)|^{-2}, \tag{14}$$

and any two solutions of (14) *differ by a polyharmonic polynomial solution of* (9).

We denote by $E_{m\alpha}$ any solution of (14). Let us define a function θ_t on $\mathbb{R}^d$ by

$$\theta_t(\omega) = \exp(2\pi i(\omega, t)) - \sum_{i=1}^{M} P_i(t) \exp(2\pi i(\omega, x_i)), \tag{15}$$

where $(\cdot, \cdot)$ denotes the standard scalar product in $\mathbb{R}^d$. Let $K^t_{m\alpha}$ be the element of $\mathscr{S}'(\mathbb{R}^d)$ defined by

$$\mathscr{F}K^t_{m\alpha} = \theta_t \mathscr{F}E_{m\alpha}. \tag{16}$$

LEMMA 2.3.4. *For all* $t \in \mathbb{R}^d$, $K^t_{m\alpha} \in \mathscr{E}_{m\alpha}$.

Then for any element $u \in \mathscr{E}^1_{m\alpha}$, we can define an element v of $\mathscr{E}$ by

$$v(t) = \left(u, \pi^1 K^t_{m\alpha}\right)_{m\alpha}, \tag{17}$$

where π^1 is the orthogonal projection onto the space $\mathscr{E}^1_{m\alpha}$.

LEMMA 2.3.5. $v \in \mathscr{S}'(\mathbb{R}^d)$.

LEMMA 2.3.6. $u - v$ *is the solution of* (9).

LEMMA 2.3.7. $v \in \mathscr{E}_{m\alpha}$.

By Lemmas 2.3.5 and 2.3.6, we conclude that $u - v$ is a polyharmonic polynomial of degree less than or equal to $2m$. But Lemma 2.3.7 shows that its degree is necessarily less than or equal to $(m - 1)$. Finally since $\theta_{x_j} \equiv 0$ for all $j = 1, \ldots, M$, by definition of the P_j, we have $v(x_j) = 0$ and therefore v and $u - v$ belong to $\mathscr{E}^1_{m\alpha} \cap \mathscr{P}_m$. Thus, for all $u \in \mathscr{E}^1_{m\alpha}$, $u(t) = (u, \pi^1 K^t_{m\alpha})_{m\alpha}$ and $\pi^1 K^t_{m\alpha} = R^1_t$. □

We now give an alternate formula for the reproducing kernel in $\mathscr{E}^1_{m\alpha}$, which will be useful for the computations of α-splines [it appears in Wahba and Wendelberger (1980) for the thin-plate spline case].

COROLLARY.

$$\begin{aligned} R^1(s,t) = {} & E_{m\alpha}(t - s) - \sum_{i=1}^{M} P_i(t) E_{m\alpha}(x_i - s) \\ & - \sum_{i=1}^{M} P_i(s) E_{m\alpha}(t - x_i) + \sum_{i;j=1}^{M} P_j(t) P_i(s) E_{m\alpha}(x_j - x_i). \end{aligned} \tag{18}$$

PROOF. $\psi^t(s) = E_{m\alpha}(t - s) - \sum P_i(s) E_{m\alpha}(t - x_i)$ defines a function of $\mathscr{S}'(\mathbb{R}^d)$, which clearly satisfies $\mathscr{F}\psi^t = \theta_t \mathscr{F} E_{m\alpha} = \mathscr{F} K^t_{m\alpha}$; hence, $\psi^t = K^t_{m\alpha}$ in $\mathscr{S}'(\mathbb{R}^d)$. On the other hand,

$$\begin{aligned} K^t_{m\alpha}(s) &= \left(\pi^0 K^t_{m\alpha}, R^0_t\right)_{m\alpha} + \left(\pi^1 K^t_{m\alpha}, R^1_t\right)_{m\alpha} \\ &= \sum_{i=1}^{M} P_i(t) K^t_{m\alpha}(x_i) + R^1(s,t), \end{aligned}$$

which is equivalent to (18). Note that one may also write

$$R^1(s,t) = \int_{\mathbb{R}^d} \frac{\theta_t(\omega)\theta^*_s(\omega)}{|\alpha(\omega)|^2 \|\omega\|^{2m}}\, d\omega. \tag{19}$$

□

2.4. *The optimization problem and its solution.* It is clear from Section 2.3 that one can now define concrete partial splines as in Section 2.1 by letting H be $\mathcal{E}_{m\alpha}$, the seminorm being defined by (6). Denote by W the $n \times K$ matrix with elements $w_{ij} = w_j(t_i)$, $i = 1, \ldots, n$; $j = 1, \ldots, K$, and by P the $n \times M$ matrix with elements $p_{ij} = P_j(t_i)$. We make the following additional assumption:

(A5) The $n \times (K + M)$ matrix (WP) has rank $(K + M)$.

Note that (A5) implies $n \geq (K + M)$ and $\text{rank}(P) = M$. In dimension $d = 1$, for ordinary (nonpartial) splines, (A5) reduces to $n \geq m$. Note also that $\text{rank}(P) = m$ if and only if the set of points $\{t_1, \ldots, t_n\}$ contains a $\mathcal{P}_m$ unisolvent subset.

We can apply classical results from partial spline theory [see, e.g., Laurent (1988) or Bates, Lindstrom, Wahba and Yandell (1987)] to show that, under assumptions (A1) through (A5), problems (1) and (2) have a unique solution in $\mathcal{E}_{m\alpha}$, and to write explicit formulas for their solution. For both problems, the solution is of the form

$$h = \sum_{k=1}^{K} \theta_k w_k + \sum_{j=1}^{M} \zeta_j P_j + \sum_{i=1}^{n} \mu_i R(t_i, \cdot) \tag{20}$$

with $\theta = (\theta_1, \ldots, \theta_K)'$, $\zeta = (\zeta_1, \ldots, \zeta_M)'$ and $\mu = (\mu_1, \ldots, \mu_n)'$ satisfying the linear system

$$\begin{aligned} K\mu + W\theta + P\zeta &= Y, \\ W'\mu &= 0, \\ P'\mu &= 0, \end{aligned} \tag{21}$$

where the matrix K with elements k_{ij} is defined by

$$k_{ij} = R(t_i, t_j) \tag{22}$$

for the interpolating problem, and by

$$k_{ij} = R(t_i, t_j) + \lambda \delta_{ij} \tag{23}$$

for the smoothing problem.

For computational purposes, one can use the following lemma which generalizes the corresponding well-known property in the thin-plate case.

LEMMA 2.4.1. *The solution defined by* (20), (21) *and* (22) *or* (23) *is unchanged if one replaces in* (20) $R(t_i, \cdot)$ *by* $E_{m\alpha}(t_i - \cdot)$ *and in* (22) *or* (23) $R(t_i, t_j)$ *by* $E_{m\alpha}(t_i - t_j)$.

We note that α-splines provide an interpolation method which is translation equivariant, and moreover rotation equivariant when α is isotropic. It is clear that thin-plate splines are obtained for $\alpha \equiv 1$ and L-splines for α equal to a polynomial with no real root.

3. Stochastic filtering.

3.1. *Model and prediction method.* The stochastic model we consider to explain the data is the following.

Let $Y(t)$, $t \in \mathbb{R}^d$, be a stochastic process with second moments (also called a random field when $d > 1$). Let $\varepsilon_1, \ldots, \varepsilon_n$ be n random variables independent from the process $Y(t)$, satisfying the usual assumptions for a white noise:

$$E\varepsilon_i = 0; \qquad E\varepsilon_i\varepsilon_j = 0, \qquad i \neq j; \qquad E\varepsilon_i^2 = \sigma^2, \qquad i, j = 1, \ldots, n,$$

where E denotes mathematical expectation. The observations y_i are explained as one realization of the random variable $Z_i = Y(t_i) + \varepsilon_i$. Moreover, the process $Y(t)$ is often decomposed into the sum of a drift and a fluctuation:

$$Y(t) = m(t) + X(t), \tag{24}$$

where $m(t)$ is a deterministic component linear combination of a finite number L of known functions u_i on $\mathbb{R}^d$ with unknown parameters, and where the process $X(t)$ has mean 0 and given covariance structure $R(s, t) = E(X(t)X(s))$.

We furthermore specialize the stochastic model by assuming that the process $X(t)$ belongs to certain class of stochastic processes convenient for statistical inference.

A theory analogous to the well-known theory of stationary random processes has been developed by Yaglom (1957) and Guelfand and Vilenkin (1967), in the context of generalized random processes. It has been exploited later by Matheron (1973) in its theory of intrinsic random functions, restricting to the case where the generalized processes with stationary increments of order m are mere random functions (ordinary processes) with stationary increments of order m. The interest of these processes is that one can develop a spectral analysis of their second-order structure by Bochner-type theorems. The role of the function R naturally related to the covariance structure of an ordinary stationary process by $R(s, t) = R(s - t)$ is taken here by the variogram or structure function. It is defined, for an ordinary process with stationary increments of order m, as any function V such that

$$E\left(\sum_{i=1}^{n} c_i X(t_i)\right)^2 = \sum_{i,j=1}^{n} c_i c_j V(t_i - t_j) \tag{25}$$

for any $(t_1, \ldots, t_n) \in (\mathbb{R}^d)^n$ and any set of coefficients $(c_1, \ldots, c_n)$ satisfying $\sum c_i P(t_i) = 0$, for all P in $\mathscr{P}_m$. In the same way as the function R is even and of positive type, the function V is even and conditionally of positive type of order m. Guelfand and Vilenkin (1967) proved that the class of tempered distributions U conditionally of positive type of order m is characterized by the fact that their Fourier transform, in $\mathscr{S}'(\mathbb{R}^d)$, satisfies the following property: For any polynomial P with complex coefficients and total degree less than or equal to $(m - 1)$, $|P|^2\mathscr{F}U$ is a positive slowly increasing measure. Hence, for an ordinary process with stationary increments of order m, and variogram V, one can define a positive slowly increasing measure by

$\|\omega\|^{2m}\mathcal{F}V(\omega)$; since it plays a role parallel to the spectral measure of a stationary process, we will refer to it as to the m-spectral measure.

Now we can state the additional assumptions for model (24):

(B1) The span of $u_1, \ldots, u_L$ contains $\mathcal{P}_m$.
(B2) $X(t)$ is a mean zero second-order ordinary process with stationary increments of order m and variogram V.
(B3) The m-spectral measure $d\mu_x$ is absolutely continuous with respect to Lebesgue measure.
(B4) $(1 + \|t\|^2)^{-m}$ is integrable with respect to the measure $d\mu_x$.

The rationale behind (B3) and (B4) is that we want the m-spectral measure to have a density of the form $|\alpha|^{-2}$ with a function α satisfying (A1) through (A4). Note that (B2) does not entirely define the process $X(t)$ since there exists nontrivial such processes with variogram 0 (polynomials of degree less than or equal to m and mean zero random coefficients); but they are the only characteristics of X that play a role in the solution of the subsequent problem. One could as well use Matheron's "$(m - 1)$ IRF" without drift.

The predicted random process $\hat{Y}(t)$ obtained by BLUE is the linear combination of the random variables Z_i, which satisfies $E(\hat{Y}(t)) = E(Y(t))$, for any value of the parameters in $m(t)$, and minimizes the predictive mean square error $E(\hat{Y}(t) - Y(t))^2$.

3.2. *Computation of the prediction.* To underline the equivalence in view of the following paragraph, we use a similar notation.

Let $\hat{Y}(t) = \sum_{i=1}^{n}\gamma_i(t)Y(t_i)$ and $\Gamma = (\gamma_1(t), \ldots, \gamma_n(t))'$. Let $w_1, \ldots, w_K$, $k = 1, \ldots, K = L - M$, be functions that span a subspace of the span of $\{u_1, \ldots, u_n\}$ in direct sum with $\mathcal{P}_m$. Let $W(t) = (w_1(t), \ldots, w_K(t))'$ and $P(t) = (P_1(t), \ldots, P_M(t))'$, where $P_1, \ldots, P_M$ span $\mathcal{P}_m$ as previously. Then the unbiasedness condition gives

$$P'\Gamma = P(t) \quad \text{and} \quad W'\Gamma = W(t). \tag{26}$$

Condition (26) is equivalent to saying that the set of $(n + 1)$ coefficients $(\gamma_1(t), \ldots, \gamma_n(t), -1)$ satisfies

$$\sum_{i=1}^{n}\gamma_i(t)Q(t_i) - Q(t) = 0,$$

for any $Q \in \mathcal{P}_m$, and hence the predictive mean square error is given by

$$\begin{aligned} E\big(\hat{Y}(t) - Y(t)\big)^2 &= E\left(\sum_{i=1}^{n}\gamma_i(t)X(t_i) - X(t) + \sum_{i=1}^{n}\gamma_i(t)\varepsilon_i\right)^2 \\ &= \sum_{i;j=1}^{n}\gamma_i(t)\gamma_j(t)V(t_i - t_j) \\ &\quad - \sum_{i=1}^{n}\gamma_i(t)V(t_i - t) + \sigma^2\sum_{i=1}^{n}\gamma_i^2(t). \end{aligned}$$

Hence, if we denote by K the matrix with elements k_{ij} defined by

$$k_{ij} = V(t_i - t_j) + \sigma^2 \delta_{ij}, \tag{27}$$

and by $Z(t)$ the vector $(V(t_1 - t), \ldots, V(t_n - t))'$, the vector Γ is obtained by minimizing $\Gamma' K \Gamma - Z'(t)\Gamma$ under the constraints $P'\Gamma = P(t)$ and $W'\Gamma = W(t)$. We are reduced to a quadratic optimization problem with linear constraints. Now, if we denote by Ω the $n \times n$ matrix $K^{-1}(I - P(P'K^{-1}P)^{-1}P'K^{-1})$, easy linear algebra shows that Γ satisfies

$$\begin{aligned} \Gamma' Y = &\left[\Omega Y - \Omega W(W'\Omega W)^{-1} W'\Omega Y\right]' Z(t) \\ &+ \left[(P'K^{-1}P)^{-1} P'K^{-1}\left\{Y - W(W'\Omega W)^{-1} W'\Omega Y\right\}\right]' P(t) \\ &+ \left[(W'\Omega W)^{-1} W'\Omega Y\right]' W(t). \end{aligned} \tag{28}$$

Note that $\Gamma' Y$ is the result of substituting y_i for $Y(t_i)$ in $\hat{Y}(t)$.

4. The equivalence. Now we can state and easily check the equivalence result corresponding to our particular conditions.

THEOREM 2. *Under assumptions* (B1) *through* (B4), *the BLUE of* $Y(t)$ *in model* (24), *when one substitutes the data* y_i *for the random variable* $Y(t_i)$, *is a partial* α*-spline of order* m, *where the functions* w_k *of* (2) *span a subspace of the span of* $\{u_1, \ldots, u_L\}$ *in direct sum with* $\mathcal{P}_m$, *where the smoothing parameter* λ *of* (2) *is the variance of the noise* σ^2, *and where the function* α *is related to the variogram* V *by*

$$\|\omega\|^{2m} \mathcal{F}V(\omega) = |\alpha(\omega)|^{-2} \quad \text{for all } \omega \in \mathbb{R}^d. \tag{29}$$

Different methods have been used for similar problems. We choose here to simply check that the solutions of both optimization problems, which we have characterized in Sections 2 and 3, are identical. An algebraic point of view could be taken as in Matheron (1981) and Kimeldorf and Wahba (1970a). In a particular case, one can take an even more elegant approach which could possibly be generalized: When $m(t) = 0$ in (24), it is not difficult to see that these optimization problems are dual problems in the sense of convex optimization duality [see Ciarlet (1982)]. Note that the presentation of the filtering problem as an optimization problem in a Hilbert space dates back to Parzen (1961).

PROOF. We already remarked that the density of the m-spectral measure $d\mu_X$ of X with respect to Lebesgue measure can be written $|\alpha|^{-2}$, with a function α satisfying (A1) to (A4). Hence, V satisfies (29) which implies that V is one of the fundamental solutions of (14). Therefore, (20) and (28) are

similarly linear combinations of the same functions and it is enough to check that the coefficients given in (28) satisfy the system (21), which is easy algebra. □

In the theory of stationary processes, particular interest has been given to those with a rational spectral density (the so-called ARMA process for discrete time). In continuous and one-dimensional time, the general rational spectral density can be written [see Yaglom (1962)]:

$$f_x(\omega) = \frac{|Q(2\pi i\omega)|^2}{|P(2\pi i\omega)|^2}, \tag{30}$$

where P and Q are real coefficient polynomials of degree p and q respectively, and where all the 0's of P have positive real part and the 0's of Q have nonnegative real part. By analogy in dimension d, one can call $X(t)$ an isotropic ARIMA(p, m, q) process any second-order process with stationary increments of order m, and m-spectral density of the form (30), where we substitute $\|\omega\|$ for ω in the right-hand side, and satisfying $m + p - q > d/2$. The corresponding α function is then given by

$$\alpha(\omega) = \frac{|P(2\pi i\omega)|}{|Q(2\pi i\omega)|}. \tag{31}$$

A classical example is the correspondence between thin-plate splines in odd dimension and the prediction of processes whose variogram is proportional to $\|t\|^{2k+1}$, which can be classified as ARIMA$(0, k + (d + 1)/2, 0)$. Furthermore, (30) and (31) show that the family of splines obtained for (31) with constant P is associated to the prediction of processes with a "polynomial generalized covariance" type of variogram [see Matheron (1973)], which can be classified as ARIMA$(0, k + (d + 1)/2, k - j)$ [see Thomas-Agnan (1989) for details].

APPENDIX

PROOF OF LEMMA 2.3.3. Note that $|\alpha|^{-2} \in \mathcal{S}'(\mathbb{R}^d)$ and that it is enough to find a solution F in $\mathcal{S}'(\mathbb{R}^d)$ to the division problem: $\|\omega\|^{2m}F(\omega) = |\alpha(\omega)|^{-2}$. Let $T_{2m-1}(\varphi)$ be the Taylor expansion of order $(2m - 1)$ at 0 of an element φ of $\mathcal{S}(\mathbb{R}^d)$, and let F be the following linear functional on $\mathcal{S}(\mathbb{R}^d)$:

$$\begin{aligned} F(\varphi) = &\int_{\mathbb{R}^d} \left(\frac{1 - \xi(\omega)}{|\alpha(\omega)|^2} \right) \left(\frac{\varphi(\omega) - T_{2m-1}(\varphi)(\omega)}{\|\omega\|^{2m}} \right) d\omega \\ &+ \int_{\mathbb{R}^d} \left(\frac{\xi(\omega)}{|\alpha(\omega)|^2} \right) \left(\frac{\varphi(\omega)}{\|\omega\|^{2m}} \right) d\omega. \end{aligned} \tag{32}$$

The fact that F belongs to $\mathcal{S}'(\mathbb{R}^d)$ follows from (A3) and (A4). Moreover, since

$T_{2m-1}(\|\cdot\|^{2m}\varphi) = 0$, we have

$$F(\|\cdot\|^{2m}\varphi) = \int_{\mathbb{R}^d} \frac{\varphi(\omega)}{|\alpha(\omega)|^2}\, d\omega,$$

which shows that F solves the division problem. □

PROOF OF LEMMA 2.3.4. For $\beta \in \mathbb{N}^d$ such that $|\beta| < m$, $P(t) = t^\beta \in \mathscr{P}_m$ and hence by (8), $P(t) = (R_t^0, P)_{m\alpha} = \sum_{i=1}^M P(x_i)P_i(t)$ which easily implies that

$$D^\beta \theta_t(0) = 0. \tag{33}$$

Therefore, $\mathscr{F}(D^\beta K_{m\alpha}^t)$, which by (16) is equal to the map: $\omega \mapsto (2\pi i\omega)^\beta \theta_t(\omega)\mathscr{F}E_{m\alpha}(\omega)$, can be represented by the locally integrable function: $\omega \mapsto (2\pi i\omega)^\beta \theta_t(\omega)|\alpha(\omega)|^{-2}\|\omega\|^{-2m}$. Furthermore, $\alpha\mathscr{F}(D^\beta K_{m\alpha}^t)$ is then represented by the function: $\omega \mapsto (2\pi i\omega)^\beta \theta_t(\omega)\alpha^*(\omega)^{-1}\|\omega\|^{-2m}$ which lies in $L_2(\mathbb{R}^d)$. □

PROOF OF LEMMA 2.3.5. Using (33) and the Taylor formula, there exists a neighborhood V_0 of 0 such that, for all $\omega \in V_0$:

$$|\theta_t(\omega)| \le \frac{\|\omega\|^m}{m!} \sup_{|\beta|=m} (D^\beta \theta_t)(0).$$

Let Q be defined by

$$Q(t) = \sup_{|\beta|=m} (D^\beta \theta_t)(0) = \sup_{|\beta|=m} \left| t^\beta - \sum_{i=1}^M x_i^\beta P_i(t) \right|.$$

On the other hand, for $\omega \in \mathbb{R}^d \setminus V_0$, we have

$$|\theta_t(\omega)| \le 1 + \sum_{i=1}^M |P_i(t)| = S(t).$$

Therefore, we have

$$|v(t)| \le \sum_{|\beta|=m} \binom{m}{\beta} \left[Q(t) \int_{V_0} |\alpha(\omega)\mathscr{F}(D^\beta u(\omega))|\, |\alpha^*(\omega)|^{-1}\, d\omega \right.$$
$$\left. + S(t) \int_{\mathbb{R}^d \setminus V_0} |\alpha(\omega)\mathscr{F}(D^\beta u(\omega))|\, |\alpha^*(\omega)|^{-1} \|\omega\|^{-m}\, d\omega \right].$$

This shows that v is a locally integrable function, and the growth rate at ∞ of Q and S being a polynomial rate completes the argument. □

PROOF OF LEMMA 2.3.6. For $\gamma \in \mathbb{N}^d$ such that $|\gamma| \ge m$, since the total degree of P_i is less than or equal to $(m-1)$, we have, for $\varphi \in \mathscr{S}(\mathbb{R}^d)$,

$$\int_{\mathbb{R}^d} \theta_t(\omega) D^\gamma \varphi(t)\, dt = (2\pi i\omega)^\gamma \mathscr{F}\varphi(\omega).$$

Therefore, using the proof of Lemma 2.3.4, we get

$$\int_{\mathbb{R}^d} D^\gamma v(t)\varphi(t)\,dt = \sum_{|\beta|=m} \binom{m}{\beta} \int_{\mathbb{R}^d} \left[\alpha(\omega)(2\pi i\omega)^\beta \mathcal{F}u(\omega)\right] \times \left[\frac{(2\pi i\omega)^{\gamma+\beta}\mathcal{F}\varphi^*(\omega)}{\alpha(\omega)\|\omega\|^{2m}}\right] d\omega. \tag{34}$$

Hence, since $\mathcal{F}\Delta^m v = \sum_{|\gamma|=m}\binom{m}{\gamma}\omega^{2\gamma}\mathcal{F}v$,

$$\int_{\mathbb{R}^d} \Delta^m v(t)\varphi(t)\,dt = \sum_{|\beta|=m} \binom{m}{\beta} \int_{\mathbb{R}^d} \left[\alpha(\omega)(2\pi i\omega)^\beta \mathcal{F}u(\omega)\right] \times \left[\frac{(2\pi i\omega)^{\beta}\mathcal{F}\varphi^*(\omega)}{\alpha(\omega)}\right] d\omega$$

$$= \int_{\mathbb{R}^d} \Delta^m u(t)\varphi(t)\,dt. \qquad \square$$

PROOF OF LEMMA 2.3.7. By (34), if $|\gamma| = m$, $\mathcal{F}D^\gamma v$ can be represented by the function

$$\omega \mapsto \sum_{|\beta|=m} \binom{m}{\beta} \left[\alpha(\omega)\mathcal{F}(D^\beta u)(\omega)\right] \left[\frac{(2\pi i\omega)^{\beta+\gamma}}{\alpha(\omega)\|\omega\|^{2m}}\right] = \mathcal{F}(D^\gamma u).$$

Therefore $v \in \mathcal{E}_{m\alpha}$ follows from $u \in \mathcal{E}_{m\alpha}$. $\square$

PROOF OF LEMMA 2.4.1. If $P'\mu = 0$, we have $\sum \mu_i R(t_i, t) = \sum \mu_i R^1(t_i, t)$ and using (18):

$$\sum_{i=1}^{v} \mu_i R(t_i, t) = \sum_{i=1}^{n} \mu_i \left[E_{m\alpha}(t_i - t) - \sum_{j=1}^{M} P_j(t) E_{m\alpha}(t_i - x_j)\right].$$

Let $\zeta_j^1 = \zeta_j - \sum \mu_i E_{m\alpha}(t_i - x_j)$. Then we get

$$\sum_{j=1}^{M} \zeta_j P_j + \sum_{i=1}^{n} \mu_i R(t_i, \cdot) = \sum_{j=1}^{M} \zeta_j^1 P_j + \sum_{i=1}^{n} \mu_i E_{m\alpha}(t_i - \cdot)$$

and evaluating this at $t = t_j$ for $j = 1, \ldots, n$ gives

$$P\zeta + K\mu = P\zeta^1 + K^1\mu,$$

where K^1 is the matrix with elements $k_{ij}^1 = E_{m\alpha}(t_i - t_j)$. $\square$

REFERENCES

BATES, D., LINDSTROM, M., WAHBA, G. and YANDELL, B. (1987). GCVPACK—routines for generalized cross-validation. *Comm. Statist. Simulation Comput.* **16** 263–297.

CIARLET, P. G. (1982). *Introduction à l'Analyse Numérique matricielle et à l'optimisation*. Masson, Paris.

DENY, J. and LIONS, J. L. (1954). Les espaces du type de Beppo-Levi. *Ann. Inst. Fourier (Grenoble)* **5** 305–370.

DOLPH, C. L. and WOODBURY, M. A. (1952). On the relation between Green's functions and covariances of certain stochastic processes and its application to unbiased linear prediction. *Trans. Amer. Math. Soc.* **72** 519–550.

DUBRULE, O. (1983). Two methods with different objectives: Splines and kriging. *J. Internat. Assoc. Math. Geol.* **15** 245–257.

DUCHON, J. (1977). Splines minimizing rotation invariant semi-norms in Sobolev spaces. In *Constructive Theory of Functions of Several Variables, Oberwolfach 1976* (W. Schempp and K. Zeller, eds.) 85–100. Springer, Berlin.

GUELFAND, I. M. and VILENKIN, N. Y. (1967). *Les distributions, applications de l'analyse harmonique* (Dunod, ed.) **4**.

HECKMAN, N. (1986). Spline smoothing in a partly linear model. *J. Roy. Statist. Soc. Ser. B* **48** 244–248.

KIMELDORF, G. S. and WAHBA, G. (1970a). Spline functions and stochastic processes. *Sankhyā Ser. A* **32** 173–180.

KIMELDORF, G. S. and WAHBA, G. (1970b). A correspondence between Bayesian estimation on stochastic processes and smoothing by splines. *Ann. Math. Statist.* **41** 495–502.

KIMELDORF, G. and WAHBA, G. (1971). Some results on Tchebicheffian spline functions. *J. Math. Anal. Appl.* **33** 82–95.

KLONIAS, V. K. (1984). On a class of non-parametric density and regression estimators. *Ann. Statist.* **12** 1263–1284.

KOHN, R. and ANSLEY, C. F. (1983). On the smoothness properties of the best linear unbiased estimate of a stochastic process observed with noise. *Ann. Statist.* **11** 1011–1017.

KOHN, R. and ANSLEY, C. F. (1987). A new algorithm for spline smoothing based on smoothing a stochastic process. *SIAM J. Sci. Statist. Comput.* **8** 33–48.

LAURENT, P. J. (1980). Spline functions using inf-convolution. *International Workshop on Approximation Theory*. Technion, Haïfa, Israel.

LAURENT, P. J. (1988). Inf-convolution splines. Technical Report RR 749-M-IMAG, Grenoble.

MATHERON, G. (1973). The intrinsic random functions and their applications. *Adv. in Appl. Probab.* **5** 439–468.

MATHERON, G. (1981). Splines and kriging: Their formal equivalence. In *Computer Applications in the Earth Sciences: An Update of the 70s* (D. F. Merriam, ed.). Plenum Press, New York.

MEINGUET, J. (1979). Multivariate interpolation at arbitrary points made simple. *J. Appl. Math. Phys.* **30** 292–304.

PARZEN, E. (1961). An approach to time series analysis. *Ann. Math. Statist.* **32** 951.

SALKAUSKAS, K. (1982). Some relationships between surface splines and kriging. In *Multivariate Approximation Theory. II. Internat. Ser. Numer. Math.* **61** 313–325.

SCHWARTZ, L. (1966). *Théorie des distributions*. Hermann, Paris.

SHIAU, J. J., WAHBA, G. and JOHNSON, D. R. (1986). Partial spline models for the inclusion of tropopause and frontal boundary information in otherwise smooth two and three-dimensional objective analysis. *J. Atmospheric Oceanic Tech.* **3** 714–725.

THOMAS-AGNAN, C. (1987). Statistical curve fitting by Fourier techniques. Thesis, Univ. California, Los Angeles.

THOMAS-AGNAN, C. (1989). Smoothing noisy data by two equivalent techniques. In *Data Analysis, Learning Symbolic and Numeric Knowledge* (E. Diday, ed.). Nova Science Publishers, New York.

THOMAS-AGNAN, C. (1990a). A family of splines for nonparametric regression and their relationships with kriging. *Statistics* **21** 533–548.

THOMAS-AGNAN, C. (1990b). Smoothing periodic curves by a method of regularization. *SIAM J. Sci. Statist. Comput.* **11** 482–502.

WAHBA, G. (1986). Partial and interaction spline models for the semiparametric estimation of functions of several variables. In *Computer Science and Statistics: Proceedings of the 18th Symposium on the Interface*. (T. J. Boardman, ed.) Amer. Statist. Assoc., Washington, D.C.

WAHBA, G. and WENDELBERGER, J. (1980). Some new mathematical methods for variational objective analysis using splines and cross-validation. *Monthly Weather Rev.* **108** 1122–1145.

WATSON, G. S. (1983). Smoothing and interpolation by kriging and with splines. Technical Report 241, series 2, Dept. Statistics, Princeton Univ.

YAGLOM, A. M. (1957). Some classes of random fields in n-dimensional space, related to stationary random processes. *Theory Probab. Appl.* **2** 273–320.

YAGLOM, A. M. (1962). *Theory of Stationary Random Functions*. Prentice-Hall, Englewood Cliffs, N.J.

UNIVERSITÉ PAUL SABATIER
118 ROUTE DE NARBONNE
31400 TOULOUSE, FRANCE