On the Whittle-Matérn correlation family

Peter Guttorp Tilmann Gneiting

NRCSE

Technical Report Series

NRCSE-TRS No. 080

July 11, 2005

On the Whittle-Matérn correlation family

Peter Guttorp and Tilmann Gneiting
University of Washington

Abstract:
Handcock and Stein (1993) introduced the Matérn family of spatial correlations into statistics as a flexible parametric class with one parameter determining the smoothness of the paths of the underlying spatial field. In this note we describe the history of this family, and document its relationship to the Hankel transform. We argue that an appropriate name for this family is the Whittle-Matérn family.

Key words: Hankel transform, spatial covariance, Bessel function.

1. *Stationary processes in* $\mathbf{R}^d$

The idea of a wide sense stationary stochastic process was introduced by Khintchine (1934), who showed that the correlation function of a continuous such process could be written as a cosine transform

$$\varphi(s) = Corr(X(t), X(t+s)) = \int_{\mathbb{R}} \cos(xs)\, dF(x)$$

of a cdf F. Cramér (1940) generalized the result to a d-dimensional complex-valued process, assuming continuity at the origin of the correlation function, obtaining

$$\varphi(s) = E(\exp(is^T X)) \tag{1}$$

where X is a d-dimensional random variable. The special case of $d = 2$ for real stationary processes was discussed in detail by Matérn (1947), who demonstrated that the exponential and probability mixtures of exponential correlations are valid spatial correlation functions.

2. *Isotropic spatial correlations and the Hankel transform*

An important subclass of the stationary correlation functions are those that are isotropic, i.e., have spherical isocorrelation curves. Using the representation (1), we see that in the isotropic case, the right-hand side must depend only on $r = |s|$, so that

$$\varphi(r) = E(\exp(ir|X|\cos\theta),$$

where θ is the angle between the vectors s and X. Assuming for simplicity that X has a density $f_X(x)$ (which depends only on $|x| = u$) we can write (1) as

$$\varphi(r) = \int_0^\infty u^{d-1} f_X(u) \int_{S_d} \exp(iur\cos\theta)\, d\sigma_d\, du, \tag{2}$$

where the inner integral is over the surface of the unit sphere in d dimensions with respect to the uniform distribution over that surface. Following Schoenberg (1938) and Lord (1954), this inner integral becomes $(2\pi)^{d/2}(ur)^{-d/2+1} J_{d/2-1}(ur)$, where J_ν is a Bessel function of the first kind and order ν. Let $\Lambda_\nu(x) = \Gamma(\nu+1)(x/2)^{-\nu} J_\nu(x)$. Then (2) becomes

$$\varphi(r) = \int_0^\infty f_{|X|}(u)\Lambda_{d/2-1}(ru)\, du, \tag{3}$$

where $f_{|X|}$ is the density of $|X|$, i.e. the Hankel transform (Hankel 1875; Piessens, 1996) of the density $f_{|X|}$. Lord (1954, eq. 10) gives an inversion formula for the Hankel transform (see also Stein, 1999, p. 46). The inversion formula is essentially of the same form, so that any Hankel transform of a density function corresponds to an isotropic correlation function, and vice versa if the correlation function is integrable.

As conjectured by Schoenberg (1938) and proved by Crum (1956), the class of all measurable isotropic correlation functions in $\mathbf{R}^d$ for $d > 1$ consists of mixtures of isotropic correlation functions that are continuous at the origin, and the correlation function $1(r=0)$ (called the nugget effect in geostatistics).

3. *Turbulence*

von Karman (1948) developed an approximation to the correlation function for homogeneous turbulent fluctuations (e.g. in a wind tunnel) in the case of large Reynolds number and insignificant viscosity of the medium. Many authors had derived a log spectral decay of -5/3 for large frequencies, while others had argued that the spectral density behaves like a fourth power near the origin. von Karman suggested an interpolation between these values, yielding the correlation function

$$\rho(r) = \frac{2^{2/3}}{\Gamma(1/3)} r^{1/3} K_{1/3}(r)$$

which fits surprisingly well to some (but not other) wind tunnel data.

Tatarskii (1961, pp. 7 and 18; originally 1959) derived the isotropic correlation function

$$\rho(r) = \frac{2^{1-\nu}}{\Gamma(\nu)} (\alpha r)^{\nu} K_{\nu}(\alpha r) \tag{4}$$

by Fourier inversion, and Shkarofsky (1968) generalized (4) by allowing two complementary scale parameters. He argued that a correlation function for turbulence needs to have no cusp, zero derivative at the origin, and a second derivative that is negative and finite. His generalized version, satisfying these requirements, is

$$\rho(r) = \frac{(\alpha\sqrt{r^2+\delta^2})^{(\nu-1)/2} K_{(\nu-1)/2}(\alpha\sqrt{r^2+\delta^2})}{(\alpha\delta)^{(\nu-1)/2} K_{(\nu-1)/2}(\alpha\delta)} \tag{5}$$

and if we let $\delta \to 0$ we obtain the Tatarskii (or Whittle-Matérn) correlation function with scale parameter $1/\alpha$. Both Tatarskii and Shkarofsky restrict attention to dimensions $d = 1$ and $d = 3$.

4. *Whittle's generalization of the exponential correlation*

The exponential correlation function is a natural correlation in one dimension, since it corresponds to a Markov process. In two dimensions this is no longer so, although the exponential is a common correlation function in geostatistical work. Whittle (1954) determined the correlation corresponding to a stochastic differential equation of Laplace type

$$\left[\left(\frac{\partial}{\partial t_1}\right)^2 + \left(\frac{\partial}{\partial t_2}\right)^2 - \alpha^2\right] X(t_1,t_2) = \varepsilon(t_1,t_2)$$

where ε is white noise. The corresponding discrete lattice process is a second order autoregression. The correlation function is

$$\rho(r) = \alpha r K_1(\alpha r) \tag{6}$$

which is derived from (2) using the spherically symmetric density

$$f_X(x) \propto (x_1^2 + x_2^2 + \alpha^2)^{-2}.$$

In fact, Whittle calculated a more general form using the Fourier transform of

$$(x_1^2 + x_2^2 + \alpha^2)^{-(\nu+1)} \tag{7}$$

where $\nu > 0$, which results in the covariance function

$$c(r) = \left(\frac{r}{2\alpha}\right)^{\nu} \frac{K_\nu(\alpha r)}{\Gamma(\nu + 1)}. \tag{8}$$

However, he only applied this to the case d=2, and presented (6) as the natural spatial covariance, in the sense that the exponential is natural for d=1.

5. *The Whittle-Matérn family*

Matérn (1960) used the representations (2) and (3) to derive some families of correlation functions in any dimension d. For example, the squared exponential correlation $\exp(-\alpha^2 v^2)$ comes from (3) with a radial density of the form

$$f_{|X|}(u) \propto u^{d-1} \exp(-u^2 / 4\alpha^2).$$

The Whittle-Matérn family results from the representation (2) with the spherically symmetric density

$$f_X(x) \propto (\alpha^2 + x_1^2 + \cdots + x_d^2)^{-(d/2+\nu)}$$

yielding the following version of (8) (or (4))

$$\rho(r) = \frac{2^{1-\nu}}{\Gamma(\nu)} (\alpha r)^{\nu} K_\nu(\alpha r) \tag{9}$$

whenever $\nu > 0$. Actually, Matérn arrived at this correlation function in a rather circuitous route, letting the parameter in the squared exponential correlation function follow a gamma distribution which yields (9) as a spherically symmetric density, and then noting that its Fourier transform actually is a density as well, as long as $\nu > 0$. A different derivation of (9) is obtained by Matérn (1960, Chapter 3) from a shot noise process in $\mathbf{R}^d$.

There has been some confusion in the literature about the valid parameter values for (9). Matérn himself, having argued in his derivation of the family that $\nu > 0$, then claims that the model is valid also for $\nu = 0$. This is, however, incorrect. The error has been repeated by several authors, some even allowing negative values of ν.

6. *Some uses of the Whittle-Matérn family*

Goff and Jordan (1988) is, to our knowledge, the first place where this class of correlation functions is proposed to parameterize the smoothness of the realizations of the corresponding random field. Their emphasis is on the Hausdorff (fractal) dimension.

Handcock and Stein (1993), in introducing the name "Matérn family" into the statistical literature, also emphasized the relationship between the parameter ν and the smoothness of paths of the realization: a random field with this covariance will have $\lceil \nu \rceil - 1$ times mean square differentiable paths. In the Gaussian case the derivatives will be continuous almost surely (a.s). Their reference (perhaps only for the case of Gaussian random fields) is to Cramér and Leadbetter (1967). Kent (1989) questioned the direct generalization of the Cramér and Leadbetter results to higher dimensions, and developed his own criterion for a.s. continuity of sample realizations, using the Whittle-Matérn correlation function to show that his condition is best possible. Stein (1999) derived conditions for mean square continuity and differentiability for stationary processes; see also Banerjee and Gelfand (2003). Handcock and Stein further argued that the Matérn family is a flexible family, allowing varying degrees of smoothness of the underlying field, and having the exponential correlation function as a special and the squared exponential correlation function as a limiting case.

Lord (1954) looked at spherically symmetric distributions in $\mathbf{R}^d$. As we saw in section 2, characteristic functions for such distributions are equivalent to continuous isotropic correlation functions. His multivariate gamma distribution has a density proportional to (9). The class of generalized hyperbolic distributions in $\mathbf{R}^d$ (Barndorff-Nielsen et al. 1982) includes spherically symmetric members with density functions or characteristic functions that are proportional to Shkarofsky's generalization (5).

Mejia and Rodriguez-Iturbe (1974) discussed the generation of synthetic hydrological spatial processes. In particular, they derived a generalized random phase shift model with the Whittle correlation (6). Rodriguez-Iturbe and his students have had an immense influence on the use of stochastic process models in hydrology.

The field of geodesy also has a long tradition of the use of stochastic processes to analyze gravitational fields and to diagnose geodetic networks (see Meier, 1981). The Whittle correlation function (6) is here called a first-order Markov model (Shaw et al., 1969), while the corresponding spectral function (7), which also can be used as a correlation function, is called a (generalized) Hirvonen model.

Handcock and Wallis (1994) applied the Whittle-Matérn family (9) to analysis of historical temperature data in northern United States. They concluded that the temperature field is not particularly smooth, although it is difficult to assess accurately the smoothness parameter ν. Another interesting environmental application is that of Fuentes (2002) to the statistical assessment of deterministic air quality models.

7. *Discussion*

There are many names for the family (9). It has been called the Basset family (Wackernagel, 1995, p. 219), the Bessel model (Matheron, 1965, p. 43; Chilès and Delfiner, 1999, pp. 86-87), the generalized Markov model (Meier and Keller, 1990, p. 190), the Matérn class (quite the most common name, starting with Handcock and Stein, 1993, p. 406), the Whittle-Matérn class (Gneiting, 1998, p. 381; Wackernagel, 1998, p. 245), the Whittle model (Dietrich, 1995, p. 150), Tatarski's model (Shkarofsky, 1968, p. 2136; he also calls the Whittle model Bessel or Norton, and the exponential model the Booker correlation) and the von Karman class (Goff et al., 1994, p. 493).

von Karman (1948) derived the correlation (9) with _ = 1/3 by interpolation in the spectral domain. Whittle (1954) also used the fact that correlation functions are Fourier transforms. He referred to Matérn (1947), where this representation was developed for the spatial case. He calculated a general Fourier transform, yielding (essentially) Matérn's (1960) equation 2.4.7. However, he only used it as a correlation function for $d = 2$, applying it to Mercer & Hall's uniformity trial data, where it fits better than an exponential correlation function.

Tatarskii (1959) derived the general form (9), but only for $d = 1$ (his equation 1.12, p. 10) or $d = 3$ (equation 1.33, p. 21). Matérn set down the general form (9), but apparently never applied it. His applications tend to use mixtures of exponentials. While he seems to be the first to write it down as a full parametric family valid in any dimension, he referred both to Lord (1954) and to Whittle (1954) in this context. Since Whittle actually gave the general expression for the family, and since Lord only viewed it as a density function, we argue that the family should be called the Whittle-Matérn family.

Acknowledgements:
This work was done while the first author was the Environmental Research Professor of the Swedish Organization of Graduate Engineers, whose support is gratefully acknowledged. The second author was partially supported by the National Science Foundation Career Award DMS-0134264.

References:
Banerjee, S and Gelfand, A. E. (2003) On smoothness properties of spatial processes. *J. Multivar. Anal.* **84:** 85–100.
Barndorff-Nielsen, O., Kent, J. and Sørensen, M. (1982) Normal variance-mean mixtures and z distributions. *Int. Statist. Rev.* **50:** 145–159.
Chilès, J.-P. and Delfiner, P. (1999) *Geostatistics: Modeling spatial uncertainty*. New York: Wiley.
Cramér, H. (1940) On the theory of stationary stochastic processes. *Ann. Math.* **41:** 215–230.
Cramér, H. and Leadbetter, M. R. (1967) *Stationary and related random processes.* New York: Wiley.
Crum, M. M. (1956) On positive-definite functions. *Proc. London Math. Soc. (3)* **6:** 548–560.
Dietrich, C. R. (1995) A simple and efficient space domain implementation of the turning bands method. *Water Resourc. Res.* **31:** 147–156.
Fuentes, M. (2002) Spectral methods for nonstationary processes. *Biometrika* **89:** 197–210.
Gneiting,T. (1998) Closed form solutions of the two-dimensional turning bands method. *Math. Geol.* **30:** 379–390.
Goff, J. A. and Jordan, T. H. (1988) Stochastic modeling of seafloor morphology: Inversion of sea beam data for second-order statistics. *J. Geophys. Res.* **93:** 13589–13608.

Goff, J. A., Holliger, K. and Levander, A. (1994) Model fields: A new method for characterization of random seismic velocity heterogeneity. *Geophys. Res. Lett.* **21:** 493–496.
Handcock, M. S. and Stein, M. L. (1993) A Bayesian analysis of kriging. *Technometrics* **35:** 403–410.
Handcock, M. S. and Wallis, J. R. (1994) An approach to statistical spatio-temporal modeling of meteorological fields. *J. Amer. Statist. Assoc.* **89:** 368–378.
Hankel, H. (1875) Bestimmte Integrale mit Cylinderfunctionen. *Math. Ann.* **8:** 453–479.
Kent, J. T. (1989) Continuity properties for random fields. *Ann. Probab.* **17:** 1432–1440.
Khintchine, A. (1934) Korrelationstheorie der stationären stochastischen Prozesse. *Math. Ann.* **109:** 604–615.
Lord, R. D. (1954) The use of the Hankel transform in statistics. I. General theory and examples. *Biometrika* **41:** 44–55.
Matérn, B. (1947) Metoder att uppskatta noggrannheten vid linje- och provytetaxering. *Medd. Statens Skogsforskningsinstitut* **36** no. 1. (In Swedish with substantial English summary)
Matérn, B. (1960) Spatial variation–Stochastic models and their application to some problems in forest surveys and other sampling investigations. *Medd. Statens Skogsforskningsinstitut* **49** no. 5.
Matheron, G. (1965) *Les variables regionalisées et leur estimation.* Paris: Masson.
Meier, S. (1981) Planar geodetic covariance functions. *Rev. Geophys. Space Phys.* **19:** 673–686.
Meier, S. and Keller, W. (1990) *Geostatistik*. Wien: Springer-Verlag.
Piessens, R. (1996) The Hankel Transform. In A. D. Poularias (ed.) *The Transforms and Applications Handbook*, Chapter 9. The Electrical Engineering Handbook Series. Boca Raton: CRC Press.
Schoenberg, I. J. (1938) Metric spaces and completely monotone functions. *Ann. Math.* **39:** 811–841.
Shaw, L., Paul, L., and Henrikson, P. (1969) Statistical models for the vertical deflection from gravity anomaly models. *J. Geophys. Res.* **74:** 4259–4265.
Shkarofsky, I. P. (1968) Generalized turbulence space-correlation and wave-number spectrum-function pairs. *Can. J. Phys.* **46:** 2133–2153.
Stein, M. L. (1999) *Interpolation of spatial data. Some theory for kriging.* New York: Springer-Verlag.
Tatarskii, V.I. (1959) *Teoriia fluktuatsionnykh iavlenii pri rasprostranenii voln v turbulentnoi atmosfere*. Moscow: Akademia Nauk.
Tatarskii, V. I. (1961) *Wave propagation in a turbulent medium.* New York: McGraw-Hill.
von Karman, T. (1948) Progress in the statistical theory of turbulence. *J. Marine Res.* **7:** 252–264.
Wackernagel, H. (1995) *Multivariate geostatistics*. Berlin: Springer-Verlag.
Wackernagel, H. (1998) *Multivariate geostatistics*. 2^{nd} ed. Berlin: Springer-Verlag.
Whittle, P. (1954) On stationary processes in the plane. *Biometrika* **41:** 434–449.