Data Simulation of Matérn Type

JIA-YUE LI 1, XIN LU 2, MING LI 2, SHENGYONG CHEN 3

1 Key Lab of Geographical Inform. Sci., Ministry of Education; Sch. Resources and Environment Sci. East China Normal University
No. 3600, Zhongshan North Road, Shanghai 200062
PR. CHINA
jyl204@hotmail.com, jyl@singapore.com

2 School of Information Science & Technology
East China Normal University
No. 500, Dong-Chuan Road, Shanghai 200241
PR. CHINA
angleriaxin@yahoo.com.cn, ming_lihk@yahoo.com
http://www.ee.ecnu.edu.cn/teachers/mli/js_lm(Eng).htm

3 College of Computer Science
Zhejiang University of Technology
Hangzhou 310023
PR. CHINA
sy@ieee.org

Abstract: Recently, the correlation function of the Matérn’s receives increasing interests in geostatistics. This paper discusses our work in synthesizing the random data based on the Matérn’s correlation function. The analysis in this paper exhibits that the data of the Matérn type are in the domain of fractal time series. The present results suggest that the power spectrum (PSD) based method may be efficiently for synthesizing the random data of the Matérn type. We shall explain the reason to select the PSD based method and give the demonstrations of simulations. This paper may yet provide a pavement towards the generation of multidimensional random fields of the Matérn type.

Key-Words: Random data generation; Geostatistics; Fractal time series; Fractional Langevin equation; Fractional oscillator processes; The Matérn correlation function.

1 Introduction

Time series and stochastic processes gain applications to statistic issues in various fields, ranging from financial engineering to geosciences, as can be seen from Clements and Hendry [1], Box et al. [2], Fuller [3], Papoulis and Pillai [4], Bendat and Piersol [5], Elishakoff and Lyon [6], Preumont [7], Chatfield [8], Vege [9], Balescu [10], Stanislaw [11], Chakrabarti [12], Dorf and Bishop [13], Landahl and Mollo-Christersen [14], Li [15]. In the field of stochastic processes, simulation of random series is a topic since the simulated data can be used to repeatedly explore the statistical properties of real random data (Ross [16]). This paper is in the area of random data generation.

Nowadays, nonlinear time series is paid attention to, see e.g., Fan and Yao [17], Tong [18]. In this aspect, fractal time series and chaotic one are particularly attractive (Samorodnitsky and Taqqu [19], Mandelbrot [20], Beran [21], Boudec and Thiran [22], Liu [23], Li [24]). This paper is in the scope of the simulation of fractal time series.

Recently, geostatistics, a branch of statistics, is developing fast. The Matérn correlation function (MCF) plays a role in serving as a flexible correlation function in the field of geostatistics, see e.g., Ripley [25], Schabenberger and Gotway [26], Chiles [27], Webster and Oliver [28]. It was introduced by Matérn [29] and Whittle [30], independently. Gneiting [31] first called such a function by the term of the Whittle-Matérn correlation family. Late, Guttorp and Gneiting [32] detailed it using the term of the Matérn correlation family. We use the term as that in [32].

The MCF receives increasing attention in the academic society, see e.g., Minasny and McBratney [33], Handcock and Wallis [34], Dietrich [35], Lim et al. [36-38], Pardo-Igúzquiza et al. [39], Bivand et al. [40], Donner and Barbosa [41], just naming a
few. This paper aims at presenting a simulation method of random data based on the MCF.

Denote by \( r(\tau) \) the correlation function of a zero mean random function \( x(t) \) for \( t, \tau \in \mathbb{R} \), where \( \mathbb{R} \) is the set of real numbers. Then,

\[
    r(\tau) = E[x(t)x(t+\tau)],
\]

where \( E \) is the mean operator. A correlation function is said in the MCF family if

\[
    r(\tau) = \frac{c}{\Gamma(v)} |\tau|^{v} K_v(|\tau|),
\]

where \( c \) is a constant, \( \Gamma(v) \) is the Gamma function, and \( K_v(\cdot) \) is the modified Bessel function of the second kind of order \( v \).

Note that the literature of synthesizing random data is rich, see, e.g., the early work by Cox and Muller [42] in 1958, Press et al. [43], Saucier [44], Smozuka [45]. Recently, Li [46] discussed the reasons that the simulation of fractal time series is an issue worth studying. As will be mentioned in the next section, random function following the Matérn correlation family, which we call the random function of the Matérn type, is a type of fractal time series rather than conventional ones.

In methodology, the generation of random data can be classified into two categories. One is to synthesize data according to a given probability density function (PDF), see, e.g., Press et al. [43], Saucier [44]. The other is based on a given power spectrum density function (PSD), see, e.g., Smozuka [45], Li and et al. [47-50]. For the generation of fractal time series with long-range dependence (LRD), autocorrelation function (ACF) based generation is considered, see, e.g., Lim and Teo [36], Lim and Eab [29], Lim and Teo [36]. However, the random function of the Matérn type is of short-range dependence (SRD). Thus, we shall address that we prefer the method to synthesize the random function of the Matérn type to be PSD-based.

The rest of paper is organized as follows. In Section 2, we shall give the problem statement and the preliminaries about this problem. In Section 3, the reason to select PSD-based method to synthesize the random function of the Matérn type will be explained. Demonstrations are given in Section 4. Discussions are in Section 5. Finally, we conclude the paper in Section 6.

2 Problem Statement

Recall that the process of the Matérn type is Gaussian (Lim and Teo [36]). A Gaussian process is uniquely determined by its ACF, equivalently its PSD (Papoulis and Pillai [4]).

Conventionally, a Gaussian time series, say \( y(t) \), can be taken as the output (or response) of a system of integer order under the excitation of white noise \( w(t) \). The system expressed by a stochastically differential equation of integer order is given by

\[
    \sum_{i=0}^{q} a_i \frac{d^{\nu-i}}{dt^{\nu-i}} y(t) = \sum_{i=0}^{q} b_i \frac{d^{\nu-i}}{dt^{\nu-i}} w(t).
\]

Denote by \( g(t) \) the impulse function of the linear system (3). Denote the Fourier transforms of \( y(t) \), \( g(t) \), \( w(t) \) by \( Y(\omega) \), \( G(\omega) \), and \( W(\omega) \), respectively, where \( j = \sqrt{-1} \) and \( \omega \) is angular frequency. Then, according to the theorem of convolution, one has

\[
    Y(\omega) = G(j\omega)W(\omega).
\]

Denote the PSDs of \( y(t) \) and \( w(t) \) by \( S_y(\omega) \) and \( S_w(\omega) \), respectively. Then, \( S_y(\omega) = 1 \) if \( w(t) \) is the normalized white noise. Therefore, one has

\[
    S_y(\omega) = |G(j\omega)|^2.
\]

Thus, the solution to (3) in time is given by

\[
    r(\tau) = F^{-1}[S_y(\omega)],
\]

where \( F^{-1} \) stands for the inverse of the Fourier transform. We note that the conventionally stochastically differential equation (3) may not produce the time series of the Matérn type as can seen from the following.

Let \( f(t) \) be a piecewise continuous on \((0, \infty)\). It is integrable on any finite subinterval of \([0, \infty)\). For \( t > 0 \), denote by \( \alpha_D^{\nu} f(u) \) the Riemann-Liouville integral operator of order \( \nu > 0 \). It is given by

\[
    \alpha_D^{\nu} f(u) = \frac{1}{\Gamma(\nu)} \int_0^t (t-u)^{\nu-1} f(u)du,
\]

see Lubich [53], Miller and Ross [54], Podlubny [55], Hilfer [56], Lakshmikantham et al. [57].

Now, we turn to the Langevin equation. The standard Langevin equation is given by

\[
    \left( \frac{d}{dt} + \lambda \right) X_i(t) = w(t),
\]

where \( \lambda > 0 \) (Papoulis and Pillai [4]). The solution to (8) in frequency domain or in time is given by

\[
    S_{X_i}(\omega) = \frac{1}{\lambda^2 + \omega^2},
\]

\[
    r_{X_i}(\tau) = e^{-|\tau|\lambda}.
\]

Different from the derivations of either the Matérn’s [29] or the Whittle’s [30], Lim and Eab [58] derived the correction function of the Matérn family from the following fractional Langevin equation with a single parameter \( \beta > 0 \)

\[
    \left( \frac{d}{dt} + \lambda \right)^\beta X_i(t) = w(t).
\]

We address the MCF following [58-60].
Denote $g_{x_1}(t)$ the impulse response function of the above equation. Then, it is the solution to (12)
\[
\left( \frac{d}{dt} + \lambda \right)^\beta g_{x_1}(t) = \delta(t),
\]
(12)
where $\delta(t)$ is the Dirac-$\delta$ function. Denote by $G_{x_1}(\omega)$ the Fourier transform of $g_{x_1}(t)$. Then, doing the Fourier transforms on the above yields
\[
G_{x_1}(\omega) = \frac{1}{(\lambda - j\omega)^\beta},
\]
(13)
Then,
\[
S_{x_1}(\omega) = \frac{1}{(\lambda^2 + \omega^2)^\beta}.
\]
(14)
The above is the solution to (11). The inverse Fourier transform of (14) yields the solution to (11) in time domain. It is given by
\[
C_{x_1}(\tau) = F^{-1}\left[ S_{x_1}(\omega) \right] = \frac{\lambda^{-2\nu}}{2\sqrt{\pi}\Gamma(\nu + 1/2)} \left| \lambda \tau \right|^\nu K_\nu \left( \left| \lambda \tau \right| \right),
\]
(15)
where $\nu = \beta - \frac{1}{2}$.

Note that, for $\nu = H \in (0, 1)$, (14) becomes
\[
S_{x_1}(\omega) = \frac{1}{(\lambda^2 + \omega^2)^{H+1/2}}.
\]
(16)
Thus, a random function of the Matérn type is SRD. We note that an SRD series does not mean that it is smooth. Denote by $D$ the fractal dimension of the random function of the Matérn type. It is a measure of the local irregularity and is expressed by
\[
D = 2.5 - \beta.
\]
(17)
The larger the value of $D$ the stronger the local irregularity. One thing essential is that the processes of the Matérn type are in the domain of fractal time series. Therefore, the random data of the Matérn type is not in the conventional sense. Its simulation is worth studying. Fig. 1 indicates the plots of the PSD of the Matérn type for $\beta = 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, \text{ and } 0.95$. 

(a)  1
0  4  8  12  16  20  f (Hz)
0  0.5  1  PSD
(b)  1
0  4  8  12  16  20  f (Hz)
0  0.5  1  PSD
(c)  1
0  4  8  12  16  20  f (Hz)
0  0.5  1  PSD
(d)  1
0  4  8  12  16  20  f (Hz)
0  0.5  1  PSD
(e)
Fig. 1. Illustrations of the PSD of the random data of the Matérn type. (a). $\beta = 0.15$. (b). $\beta = 0.25$. (c). $\beta = 0.35$. (d). $\beta = 0.45$. (e). $\beta = 0.55$. (f). $\beta = 0.65$. (g). $\beta = 0.75$. (h). $\beta = 0.85$. (i). $\beta = 0.95$.

3 Simulation Method

Let $U$ be a uniformly distributed random number between 0 and 1. Let $p(x)$ and $F(x)$ be the PDF and cumulative distribution function of $x$, respectively. Then,

$$ F(x) = \int_{-\infty}^{x} p(t)dt. \quad (18) $$

From the above, one sees that if the PDF of $x$ is given, $F(x)$ is known. Therefore, the PDF-based method, i.e., the inverse transform method, says that if

$$ x = F^{-1}(U), \quad (19) $$

then $x$ has the given PDF $p(x)$, referring Press et al. [43], Saucier [44] for this method.

Though the random function of the Matérn type is Gaussian, the PDF-based method may be difficult to use since that method can only assure the generated random data to follow the given PDF instead of the given form of either ACF or PSD.

As far as the ACF-based method was concerned, for a given ACF $C_X(\tau)$, we have the synthesized random data expressed by

$$ y = w * F^{-1}\{[F(C_X)]^{0.5}\}, \quad (20) $$

where $*$ represents the convolution and $F$ the Fourier transform, and $w$ the white noise [46]. Note that both $C_X(\tau)$ and $S_X(\omega)$ are ordinary functions. Thus, the ACF-based method expressed by (20) works in principle. However, it may be time consuming compared to the PSD-based method interpreted below.

Let $x(t)$ be a random function with the frequency bandwidth $(0, f_{\text{max}})$, where $f_{\text{max}}$ is the maximum effective frequency of $x(t)$. Let $S_x(f)$ be its PSD. The discrete $S_x(f)$ is given by $S_x(n\Delta f)$ ($n = 1, 2, ..., N$), which implies that its PSD is divided into $N$ equal increments of width $\Delta f$ between 0 and $f_{\text{max}}$. According to the analysis in [11,12,45], one has

$$ x(t) = \sum_{n=1}^{N} a_n \cos[2\pi n\Delta f t + \vartheta(n\Delta f)], \quad (21) $$

where $0 \leq t \leq T_r$ and $\vartheta(n\Delta f)$ is a uniformly distributed random phase between $-\pi$ and $\pi$. The coefficients $a_n$ are given by

$$ a_n = a(n\Delta f) = \sqrt{2S_x(n\Delta f)\Delta f}. \quad (22) $$

Eq. (21) produces the random data whose PSD $S_x(f)$ is predetermined. The data generated with (21) contain periodicity. Let $T_r$ be the period of the generated random data. Then,

$$ T_r = \frac{1}{\Delta f}. \quad (23) $$

In the case of the finite Fourier transform,

$$ N\Delta f = f_{\text{max}}, \quad (24) $$
\[ \frac{N}{T_r} = f_{\text{max}}. \]  

(25)

Within \([0, T_r]\), the generated random data are not repeated.

4 Demonstrations

For \(T_r = 10\), we plot the generations for \(\beta = 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, \) and \(0.95\) in Fig. 2, which exhibits the periodicity of 10.

We now increase \(T_r\) up to 100. The simulations for \(\beta (= 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, \) and \(0.95\)) for \(T_r = 100\) are shown in Fig. 3, which imply the periodicity of 100. If we further increase \(T_r\) up to 1000, the simulations are not repeated within \([1, 1000]\), see Fig. 4.
Fig. 2. Simulations for $T_r = 10$. (a) Simulation for $\beta = 0.15$. (b) Simulation when $\beta = 0.25$. (c) $\beta = 0.35$. (d) Simulation for $\beta = 0.45$. (e) Simulation for $\beta = 0.55$. (f) Simulation for $\beta = 0.65$. (g) Simulation for $\beta = 0.75$. (h) Simulation for $\beta = 0.85$. (i) Simulation for $\beta = 0.95$. 
Fig. 3. Simulations for $T_r = 100$. (a). Simulation when $\beta = 0.15$. (b). Simulation when $\beta = 0.25$. (c). Simulation for $\beta = 0.35$. (d). Simulation for $\beta = 0.45$. (e). Simulation for $\beta = 0.55$. (f). Simulation for $\beta = 0.65$. (g). Simulation for $\beta = 0.75$. (h). Simulation for $\beta = 0.85$. (i). Simulation for $\beta = 0.95$. 

(a)
5 Discussions

Note that the fractal dimension represents the local irregularity of a sample path [61-72]. Therefore, smaller value of $\beta$ implies higher local irregularity. Such an effect is implied in Figs. 2-4. In Fig. 5, where $T_r = 200$, we purposely plot two simulations for $\beta = 1.5$ (i.e., $D = 1$) and $\beta = 0.6$ (i.e., $D = 1.9$) to demonstrate the effect of the local irregularity.

There are two things we shall work on in future. We shall study the simulation efficiency of the PSD-based method. In addition, we shall study the method to synthesize random field of the Matérn type.

6 Conclusion

We have discussed the famous correlation function of the Matérn’s. We have explained the PSD-based method of simulating the random series of the Matérn type. The demonstrations of simulations have been indicated.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (NSFC) under the project grant numbers 60573125, 60873264, 60870002, and Science and Technology Department of Zhejiang Province (2009C21008).

References:


