EXPLICIT CALIBRATION AND SIMULATION OF STOCHASTIC FIELDS BY LOW-ORDER ARMA PROCESSES

Steen Krenk
Department of Mechanical Engineering, Technical University of Denmark
Building 403, Nils Koppel’s Alle, DK-2800, Kgs. Lyngby, Denmark
e-mail: sk@mek.dtu.dk

Keywords: ARMA simulation, Stochastic fields, Wind field simulation.

Abstract. A simple framework for autoregressive simulation of stochastic fields is presented. The autoregressive format leads to a simple exponential correlation structure in the time-dimension. In the case of scalar processes a more detailed correlation structure can be obtained by adding memory to the process via an extension to autoregressive moving average (ARMA) processes. The ARMA format incorporates a more detailed correlation structure by including previous values of the simulated process. Alternatively, a more detailed correlation structure can be obtained by including additional ‘state-space’ variables in the simulation. For a scalar process this would imply an increase of the dimension of the process to be simulated. In the case of a stochastic field the correlation in the time-dimension is represented, although indirectly, in the simultaneous spatial correlation.

The model with the shortest memory – the single-step autoregressive model – is analyzed in detail, and an efficient multi-step calibration procedure is developed. The calibration makes direct use of conditional correlations and means, expressed explicitly in terms of the zero and k-step correlation matrices of the stochastic field. The correlation structure of an isotropic turbulent wind field is developed from the generalized von Kàrmàn spectrum in terms of the Airy function, and the simulation procedure is illustrated for turbulent wind with Airy or exponential function representation of the correlation structure. In spite of the basic simplicity of the simulation algorithm, the details of the transverse correlation, that should satisfy an integral condition for consistency, is represented well.
1 INTRODUCTION

Stochastic fields play an important role in several technical contexts, notably as models for loads from wind, waves and earthquakes, and for representation of spatial properties e.g. of materials. In the first class of applications there is a time dimension, as well as one or more spatial dimensions, and this makes it appealing to consider the field as a development of a spatial field in time. Stochastic fields are characterized by their correlation, and for stationary fields a Fourier transformation relates the correlation function to a spectral density representation. For physical processes such as waves and wind there is often a background theory, that is most easily expressed in terms of spectral properties, and this suggests simulation based on spectral densities via a Fourier representation of the field. However, a Fourier representation is typically based on a finite time interval, to be selected as basis for the representation before the simulation, and furthermore computational efficiency suggests the use of a number \(2^N\) of equal intervals in order to enable the use of the FFT formalism.

For stochastic fields with a time dimension there would be clear advantages in the use of sequential simulation techniques based on recurrence relations in time as expressed in the autoregressive moving average (ARMA) format. For scalar processes the details of the correlation structure is represented via a ‘memory’ that accounts for recent history. In principle the memory effect could be represented by including a number of auxiliary ‘state-space’ variables. For scalar processes this would imply an undesirable computational overhead, and the ARMA format is mostly retained. However, in the case of a stochastic field information of the correlation of the field may already be present in the correlation between simultaneous observations, and thus it may be possible to obtain a representation of the correlation structure, also in the time direction, via the additional information contained in a simultaneous observation of the field. This suggests the possibility of using ARMA models with rather short memory, corresponding to models with only a few coefficient matrices. The present paper deals with the shortest memory model – the single-step AR model – for stochastic fields, and develops a direct calibration procedure that overcomes the problem of lack of robustness associated with single-step calibration. The model is illustrated for simulation of a turbulent wind field, where the along-wind correlation is implicitly contained in the instantaneous spatial correlation.

2 LOW-ORDER ARMA MODELS

The family of autoregressive (AR) processes with or without moving average (MA) terms generates a series \(u_n\) of scalar or vector valued variables from a corresponding series of uncorrelated independent variables \(\xi_n\). In the context of stochastic fields the variables are vectors of dimension \(m\),

\[
u_n = [u_1, u_2, \cdots, u_m]^T, \quad \xi_n = [\xi_1, \xi_2, \cdots, \xi_m]^T.
\]

These variables are typically associated with individual points in space, and the vectors refer to a section through the field, while the sequence \(\cdots, n-1, n\) moves the section across the field as illustrated in Fig. [I].

The general format of the combined autoregressive moving average process is

\[
u_n = A_1 u_{n-1} + A_2 u_{n-2} + \cdots + A_j u_{n-j}
+ B_1 \xi_{n-1} + B_2 \xi_{n-2} + \cdots + B_k \xi_{n-k}, \quad n = 1, 2, \cdots
\]

where \(j\) denotes the number of regression terms, and \(k\) similarly the number of averaging terms. Typically, calibration of the process matrices makes use of direct matching to the covariance
properties of the field to be simulated via the Yule-Walker equations, or use is made of discrete spectral properties involving the a discrete Fourier transformation, [1].

2.1 First-Order Autoregressive Format

The present simulation algorithm will concentrate on the first order autoregressive format. In this format the current vector \( u_n \) is given as a linear combination of the previous vector \( u_{n-1} \) and a random vector \( \xi_{n-1} \),

\[
\begin{align*}
\quad u_n &= A u_{n-1} + B \xi_{n-1}, \\
& n = 1, 2, \ldots 
\end{align*}
\]

The random vectors \( \xi_n \) are here taken in the form of uncorrelated normalized normal components, whereby

\[
E[\xi_i \xi_j^T] = \delta_{ij} I.
\]

Hereby the role of the matrix \( B \) is to generate the correlation between the components of the input in the AR equation (4).

The matrices \( A \) and \( B \) determine the properties of the series, and are determined to reproduce desired properties of an underlying stochastic field. The AR format (3) corresponds to a field with zero expectation of all vectors \( u_n \). The properties of the stochastic field are given in terms of the covariance matrices

\[
C_k = E[u_n u_{n-k}^T].
\]

The stochastic field is assumed to homogeneous, and thus the covariance matrix is independent of the subscript \( n \), identifying the section. Furthermore the present algorithm relates to the first order single-step format \( \xi \). This corresponds to fields with one-step memory, the so-called Markov property. The covariance structure of this type of field can be generated from the single-step covariance properties, contained in the two matrices

\[
C_0 = E[u_n u_n^T], \quad C_1 = E[u_n \xi_{n-1}^T].
\]

These two covariance matrices determine the AR coefficient matrices \( A \) and \( B \).

2.2 Coefficient matrices and conditional field properties

The matrices \( A \) and \( B \) can be determined from basic operations involving pre-multiplication of the AR equation (4). However, it is illuminating to derive the expressions directly from the corresponding conditional expectation and covariance of the underlying field. When using the AR equation (3), the vector \( u_n \) is formed as the sum of a deterministic part \( Au_{n-1} \), determined by the previous step, and a stochastic zero-mean component \( B \xi_{n-1} \). Both coefficient matrices \( A \) and \( B \) are determined by rewriting the recurrence relation (3) in the form

\[
\begin{align*}
\quad u_n &= A u_{n-1} + B \xi_{n-1}. 
\end{align*}
\]
By definition the stochastic variable $\xi_{n-1}$ is independent of $u_{n-1}$ and has the expectation zero. The relation (7) then identifies the conditional expectation of $u_n$ as in the full sequence, when the previous vector $u_{n-1}$ is known, i.e.

$$E[u_n | u_{n-1}] = Au_{n-1}. \quad (8)$$

In the special case of zero-mean variables the conditional variance of $u_n$ for given $u_{n-1}$ is, see e.g. [2],

$$E[u_n | u_{n-1}] = C_1C_0^{-1}u_{n-1}. \quad (9)$$

Thus, the regression matrix $A$ is identified as the matrix in the conditional mean formula, whereby

$$A = C_1C_0^{-1}. \quad (10)$$

This equation gives the regression matrix $A$ explicitly in terms of the covariance matrices $C_0$ and $C_1$ of the stochastic field.

The matrix $B$ is also determined directly from the relation (7). Both sides represent a vector variable with zero mean. The vector $u_{n-1}$ is known, and thus the covariance matrix of the right hand side must correspond to the conditional covariance of $u_n$,

$$\text{Cov}[u_n u_{n}^T | u_{n-1}] = B E[\xi_{n-1}\xi_{n-1}^T] B^T. \quad (11)$$

The conditional covariance matrix of two sets of variables of which one is known is given by, see e.g. [2],

$$\text{Cov}[u_n u_{n}^T | u_{n-1}] = C_{0|1} = C_0 - C_1 C_0^{-1} C_1^T. \quad (12)$$

The vector $\xi_{n-1}$ has independent normalized components according to (4), and thus the expectation on the right hand side of (11) gives the unit matrix. Comparison of the two expressions for the conditional covariance $C_{0|1}$ then gives the following equation of the matrix $B$,

$$B B^T = C_{0|1} = C_0 - C_1 C_0^{-1} C_1^T. \quad (13)$$

This equation determines the product $B B^T$.

The role of the matrix $B$ in the autoregressive relation (7) is to generate correlated input from the uncorrelated components of the random vectors $\xi_n$. This leaves a certain indeterminacy of the matrix $B$ as only the product $B B^T$ contributes to the mutual correlation correlation of the components in the equation. A simple and direct solution consists in assuming that the coefficient matrix $B$ is in the form of a lower triangular matrix. The equation (13) then immediately identifies $B$ via the Cholesky factorization of the conditional covariance matrix $C_{0|1}$ of the stochastic field. Alternatively, the matrix $B$ may be expressed in symmetric form in terms of the eigenvalues $\Lambda$ and eigenvectors $U$ of the conditional covariance matrix $C_{0|1}$, defined by

$$C_{0|1}U = U\Lambda. \quad (14)$$

It then follows from the orthogonality relations of the eigenvectors $U$ that the coefficient matrix $B$ can be expressed as

$$B = U\Lambda^{1/2}U^T. \quad (15)$$

This completes the single-step calibration procedure.
3 MULTI-STEP MODEL CALIBRATION

It is observed that the two parameter matrices of the AR filter are determined by the conditional expectation and the conditional covariance with the previous variable $u_{n-1}$ known. Direct use of these relations is justified, when the correlation properties of the stochastic field are captured well by the relation between two neighboring vectors $u_{n-1}$ and $u_n$. However, the use of closely spaced vectors may introduce undesirable sensitivity in the calibration, and a more robust procedure can be obtained by calibrating the model via the properties of vectors with larger separation. A simple direct procedure using the vectors $u_{n-k}$ and $u_n$, separated by $k$ steps, is described in the following.

3.1 Recurrence matrix $A$

The single-step recurrence relation may be extended by substituting $u_n$ from the previous relation and so forth. This leads to the $k$-step relation

$$u_n = A^k u_{n-k} + \left( A^{k-1} B \xi_{n-k} + \cdots + A B \xi_{n-2} + B \xi_{n-1} \right).$$

(16)

In the previous section the matrix $A$ was determined via the conditional expectation of $u_n$ for known $u_{n-1}$, i.e. for one-step separation. This formula is now generalized to $k$-step separation. First the term containing $u_{n-k}$ is moved to the left side of the equation,

$$u_n - A^k u_{n-k} = \left( A^{k-1} B \xi_{n-k} + \cdots + A B \xi_{n-2} + B \xi_{n-1} \right).$$

(17)

The terms on the right are statistically independent with mean zero, and thus the conditional expectation of $u_n$ for known $u_{n-k}$ is

$$E[u_n|u_{n-k}] = A^k u_{n-k}.$$  

(18)

The conditional expectation also follows directly from the unconditional covariance between $u_n$ and $u_{n-k}$ as

$$E[u_n|u_{n-k}] = C_k C_0^{-1} u_{n-k}.$$  

(19)

Comparison of these two expressions for the $k$-step conditional expectation gives the following expression for the $k$-power of the recurrence matrix,$$
A^k = C_k C_0^{-1}.  

(20)

This relation identifies the recurrence matrix $A$ from the covariance properties at a separation of $k$ steps.

In order to extract the matrix $A$ from (20) an eigenvalue decomposition is used. To this end the notation

$$A_k = C_k C_0^{-1}$$

(21)

is introduced. This matrix is non-symmetric and the eigenvalue decomposition takes the form

$$A_k P = P \Gamma^k, \quad A_k^T Q = Q \Gamma^k,$$

(22)

where $P$ and $Q$ are the right and the left eigenvector matrices of $A_k$, respectively. The eigenvalues are contained in the diagonal matrix $\Gamma^k = [\gamma_1^k, \gamma_2^k, \cdots, \gamma_m^k]$. The eigenvector matrices are normalized to satisfy the bi-orthogonality relations

$$Q^T P = P^T Q = I.$$  

(23)
Pre-multiplication of the eigenvalue equation (22a) with $Q^T$ and use of orthogonality gives the representation

$$Q^T A_k P = \Gamma^k.$$  

(24)

Now, introducing $A^k = A_k$ and using the orthogonality relations (23) this relation takes the form

$$(Q^T A P)^k = \Gamma^k.$$  

(25)

From this relation and the orthogonality relations (23) it follows that

$$A = P \Gamma Q^T.$$  

(26)

This formula determines the recurrence matrix $A$ from its $k$-step properties and generalizes the single-step procedure from Section 2.2. Clearly, the single-step result is recovered for $k = 1$, and it is seen that the eigenvalue decomposition becomes superfluous in that case.

3.2 Input matrix $B$

The input matrix $B$ is determined from the conditional variance by generalizing the single-step procedure from Section 2.2. In the relation (17) both sides represent a random vector with zero mean. When $u_{n-k}$ is known the covariance of the left side is seen to constitute the conditional covariance of $u_n$ for given $u_{n-k}$, i.e.

$$\text{Cov}[u_n u_n^T | u_{n-k}] = E[(A^{k-1}B\xi_{n-k} + \cdots + B\xi_{n-1})(A^{k-1}B\xi_{n-k} + \cdots + B\xi_{n-1})^T].$$  

(27)

In this case the conditional variance of the stochastic field is given by

$$\text{Cov}[u_n u_n^T | u_{n-k}] = C_{0|k} = C_{0} - C_k C_{0|k-1} C^T_k.$$  

(28)

The expression in terms of the matrices $A$ and $B$ of the model follow from evaluating the expectation on the right side of (27),

$$\text{Cov}[u_n u_n^T | u_{n-k}] = A^{k-1}B B^T (A^{k-1})^T + \cdots + A B B^T A^T + B B^T.$$  

(29)

The two expressions (28) and (29) for the conditional variance establish an equation between the conditional variance $C_{0|k}$ of the stochastic field, and a series expansion in terms of powers of the matrix $A$.

It follows from the $l$-factor product of the eigenvalue representation (26) that the $l$’th power of $A$ can be expressed as

$$A^l = P \Gamma Q^T \cdots P \Gamma Q^T = P \Gamma^l Q^T,$$  

(30)

where the orthogonality relation (23) has been used to cancel the inner matrix products. Substitution of this representation into the expression (29) for the conditional covariance matrix leads to the equation

$$C_{0|k} = P \Gamma^{k-1} Q^T B B^T Q \Gamma^{k-1} P^T + \cdots + P \Gamma Q^T B B^T Q \Gamma P^T + B B^T.$$  

(31)

The factors $P$ and $P^T$ are now eliminated by pre-multiplication with $Q^T$ and post-multiplication with $Q$. This gives the relation

$$Q^T C_{0|k} Q = \sum_{l=0}^{k-1} \Gamma^l [Q^T B B^T Q] \Gamma^l.$$  

(32)
It is seen that the right side of the equation is a summation in which the components of the symmetric matrix
\[ D = Q^T B B^T Q, \quad (33) \]
are multiplied by powers of the eigenvalues \( \gamma_1, \gamma_2, \ldots, \gamma_m \). The summation is conveniently carried out by using the sum of a \( k \)-term series with geometric progression. In component form
\[
\left[ Q^T C_{0|k} Q \right]_{ij} = \sum_{l=0}^{k-1} \left[ (\gamma_i \gamma_j)^l D_{ij} \right] = \frac{1}{1 - (\gamma_i \gamma_j)^k} D_{ij}, \quad (34)
\]
This equation gives the components of the matrix \( D \) as
\[
D_{ij} = \frac{1}{1 - (\gamma_i \gamma_j)^k} \left[ Q^T C_{0|k} Q \right]_{ij}. \quad (35)
\]
When the components of \( D \) have been computed from this relation, the input matrix \( B \) is determined from (33). Pre-multiplication with \( P \) and post-multiplication with \( P^T \) gives the equation
\[
B B^T = P D P^T. \quad (36)
\]
This is an equation of the same form as \((13)\) for the single-step procedure.

It is noted that in the case \( k = 1 \) the first factor in (35) is unity, and eigenvalue decomposition becomes superfluous, as the matrix on the left side of (36) is simply equal to the conditional covariance matrix \( C_{0|1} \) in that case. In the multi-step case the input over \( k \) steps has components occurring at different times, and the memory effect is accounted for by the dependence on the eigenvalues \( \Gamma \) as illustrated in (35).

4 WIND FIELD SIMULATION

An application area of considerable current interest is that of a turbulent wind field. Specific applications range from structures like towers, high-rise buildings and long bridges to wind turbines. Traditionally, the level of ambition regarding the representation of the turbulent wind load in these applications has been very different, with structural applications often making use of simple analytical approximations for the resulting load, while wind turbine design typically makes use of rather detailed wind field simulation, often based on FFT techniques, see e.g. [3].

4.1 Isotropic wind field correlation

The turbulent wind field used in the present paper is represented in the form of isotropic incompressible turbulence as described by Batchelor [4]. It follows from invariance to coordinate transformations that the general form of the covariance between the turbulent velocity components at two points separated by the spatial vector \( r \) as shown in Fig. 2 is of the form
\[
R(r) = E[\mathbf{v}(r_0 + r) \mathbf{v}(r_0)^T] = \sigma_v^2 \left[ \left( f(r) - g(r) \right) \frac{r r^T}{r^T r} + g(r) \mathbf{I} \right], \quad (37)
\]
where \( r = |r| \) is the distance between the two points, and \( \sigma_v^2 \) is the variance of a single component at a point. The functions \( f(r) \) and \( g(r) \) describe the lengthwise and transverse correlation, respectively.
At the typical wind speeds in the natural wind the flow can be assumed to be incompressible. It then follows from the incompressibility condition $\nabla^T \mathbf{v}$ that the correlation functions are related by

$$g(r) = f(r) + \frac{d}{2} f(r). \tag{38}$$

Thus, the isotropic incompressible stochastic field is described entirely in terms of a single scalar correlation function, e.g. the lengthwise correlation function $f(r)$. The lengthwise correlation function $f(r)$ is often represented in terms of its spectral density function $F(k)$, where $k$ is the wavenumber,

$$\sigma^2_k f(r) = \int_{-\infty}^{\infty} F(k) e^{ikr} dr. \tag{39}$$

Analytically tractable results are obtained, when using the generalized form of the von Kármán spectral density, [5, 6],

$$\sigma^2_k F(k) = \frac{1}{\sqrt{\pi}} \frac{\Gamma(\gamma)}{\Gamma(\gamma - \frac{1}{2})} \frac{\sigma^2_u \ell}{1 + (k\ell)^2}^{\gamma}. \tag{40}$$

Here $\ell$ is a length-scale of the turbulence, and the Kolmogorov cascade theory implies that the exponent is $\gamma = 5/6$. The analytical details of this case has been investigated by Kristensen and Jensen [7], and it was demonstrated that coherence and correlation functions can be expressed in terms of modified Bessel functions of the second kind with fractional index. In the present context it is convenient to use an alternative representation of the correlation functions in terms of Airy functions [8]. This formulation makes use of a non-dimensional transformed variable $z$ to represent the distance,

$$z = \left(\frac{3r}{2\ell}\right)^{2/3}, \tag{41}$$

whereby

$$f(r) = \frac{\text{Ai}(z)}{\text{Ai}(0)}, \quad g(r) = f(r) + z \frac{\text{Ai}'(z)}{3 \text{Ai}(0)}. \tag{42}$$

Here $\text{Ai}(z)$ is the Airy function and $\text{Ai}'(z)$ its derivative.

A simpler approximate formulation can be obtained by using the exponent $\gamma = 1$, whereby

$$f(r) = e^{-r/\lambda}, \quad g(r) = \left(1 - \frac{1}{2} \frac{r}{\lambda}\right)e^{-r/\lambda}. \tag{43}$$

Here the parameter $\lambda$ is the integral spatial length-scale of the turbulence, defined by

$$\lambda = \int_0^\infty f(r) \, dr. \tag{44}$$

When substituting the representation (42) of the correlation functions in terms of the Airy function in the definition of the integral length-scale, it is found that $\ell = 1.339\lambda$. 

Figure 2: Two-point correlation of isotropic wind field.
The lengthwise and the transverse correlation functions are illustrated in Figs. 3a and 3b, respectively. The exponential representation \((43)\) is shown in full line, while the Airy function representation is shown by a dotted curve. It is seen that the curves from the two representations cross at \(r \approx \lambda\). The AR simulation format format is closely related to exponential attenuation, and the curves suggest the the distance \(r \approx \lambda\) may be appropriate for calibration of the AR model.

### 4.2 Simulated wind field

The AR field simulation is based on selecting a number of points in a plane, e.g. the \(xz\)-plane. The coordinates of these points are arranged in a global vector

\[
x_n^T = [r_1^T, r_2^T, \ldots, r_m^T]_n.
\]

and the corresponding velocities are similarly arranged in the global vector

\[
u_n^T = [v_1^T, v_2^T, \ldots, v_m^T]_n.
\]

The global covariance matrix is then given by

\[
C_k = E[u_n u_{n-k}^T] = \begin{bmatrix} \vdots & \vdots & \vdots \\
E[v_{i,n} v_{j,n-k}^T] & \vdots & \vdots \\
\vdots & \vdots & \vdots 
\end{bmatrix} = \begin{bmatrix} \vdots & \vdots & \vdots \\
R(r_{i,n} - r_{j,n-k}) & \vdots & \vdots \\
\vdots & \vdots & \vdots 
\end{bmatrix}.
\]

where the final form consists of the \(3 \times 3\) block matrices representing the covariance between the point \(r_i\) in layer \(n\) and the point \(r_j\) in layer \(n-k\). The simulation procedure consists of a \(k\)-step calibration of the matrices \(A\) and \(B\), based on the covariance matrices \(C_0\) and \(C_k\) as described in Section [3] followed by sequential use of the single-step recurrence \((7)\).

Simulation of a turbulent wind field by the present procedure was used in [9] for wind turbine response analysis. The field is represented by 24 radial lines of length 43 m with 8 points in each. In the example the integral length-scale is \(\lambda = 120\) m. The simulation is based on sections separated by \(\Delta y = 0.5\) m, and the calibration distance is taken to be the integral length-scale \(\lambda\), corresponding to \(k = 240\). The model is calibrated by use of the exponential representation of the wind field correlation. Lengthwise and transverse correlation for points with axial separation \(y\) are shown in Fig. [3]. It is seen that the lengthwise correlation represented by \(f(y)\) retains much of the exponential form, while the simulated transverse correlation represented by \(g(y)\) closely follows the theoretical Airy-function representation from the von Kármán spectrum.
5 CONCLUSIONS

A simple framework for autoregressive simulation of stochastic fields has been presented. The autoregressive format leads to a simple exponential correlation structure in the time-dimension. In the case of scalar processes a more elaborate correlation structure can be obtained by adding memory to the process via an extension to autoregressive moving average (ARMA) processes. The ARMA format incorporates a more detailed correlation structure by including previous values of the simulated process. Alternatively, a more detailed correlation structure can be obtained by including additional ‘state-space’ variables in the simulation. For a scalar process this would imply an increase of the dimension of the process to be simulated. In the case of a stochastic field the correlation in the time-dimension is represented indirectly in the simultaneous spatial correlation.

The model with the shortest memory – the single-step autoregressive model – has been analyzed, and an efficient multi-step calibration procedure has been developed. The calibration makes direct use of conditional correlations and means, expressed explicitly in terms of the zero and \( k \)-step correlation matrices of the stochastic field. The correlation structure of an isotropic turbulent wind field is developed from the generalized von Kármán spectrum in terms of the Airy function, and the simulation procedure is illustrated for turbulent wind with Airy or exponential function representation of the correlation structure. In spite of the basic simplicity of the simulation algorithm, the details of the transverse correlation, that should satisfy an integral condition for consistency, is represented well.

REFERENCES


