

# A Random Field Model and its Application in Industrial Production. Supplementary materials 1.

Julie OGER<sup>\*†</sup>, Emmanuel LESIGNE<sup>\*</sup>, Philippe LEDUC<sup>†</sup>

<sup>\*</sup> Université François-Rabelais, CNRS, LMPT UMR7350, Tours, France

<sup>†</sup> STMicroelectronics, Tours, France

Let  $X$  be an abstract set.

We consider a prior random field  $Y_x = U + VW_x$ , where  $U$  is a real random variable following a uniform distribution on an interval  $[-m, m]$ , where  $V$  is a real and positive random variable following a uniform distribution on an interval  $[\epsilon, 1/\epsilon]$  and where  $(W_x)_{x \in X}$  is a centered normalized Gaussian field. Moreover, we suppose that  $U$ ,  $V$  and  $(W_x)_{x \in X}$  are independent.

The parameter characterizing the Gaussian field  $(W_x)_{x \in X}$  is the correlation function  $k$  (recall that the mean is zero and the variance is 1). We suppose here that  $n \geq 3$ . Let  $x_1, x_2, \dots, x_n \in X$  and  $\mathbf{y} := (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ . Denote  $\Sigma := (k(x_i, x_j))_{1 \leq i, j \leq n}$  the matrix of correlations and  $\mathbf{k}(\mathbf{x}) := (k(x, x_j))_{1 \leq j \leq n}$  the correlation vector. We suppose that we are in a generic position so that the matrix  $\Sigma$  is invertible.

**Theorem 3.** *The conditional distribution of the random field  $(Y_x)_{x \in X}$  knowing that  $(Y_{x_i} = y_i)_{1 \leq i \leq n}$  is given by explicit formulae of densities of finite dimensional marginals.*

*When the parameter  $m$  goes to infinity and  $\epsilon$  goes to zero, for  $n > 2$ , this conditional distribution becomes a multivariate Student distribution.*

*In particular, when  $m \rightarrow \infty$  and  $\epsilon \rightarrow 0$ , for  $n > 2$ , the univariate conditional distribution of the random variable  $Y_x$  becomes a Student distribution with  $n - 2$  degrees of freedom, with location parameter*

$$\mu + \mathbf{k}(\mathbf{x})\Sigma^{-1}(\mathbf{y} - \mu\mathbf{1})^T \quad \text{with} \quad \mu := \frac{\mathbf{y}\Sigma^{-1}\mathbf{1}^T}{\mathbf{1}\Sigma^{-1}\mathbf{1}^T}$$

*and scale parameter*

$$\sqrt{\frac{1}{n-2} ((\mathbf{y} - \mu\mathbf{1})\Sigma^{-1}\mathbf{y}^T) \left( 1 - \mathbf{k}(\mathbf{x})\Sigma^{-1}\mathbf{k}(\mathbf{x})^T + \frac{(1 - \mathbf{1}\Sigma^{-1}\mathbf{k}(\mathbf{x})^T)^2}{\mathbf{1}\Sigma^{-1}\mathbf{1}^T} \right)}$$

where  $\mathbf{1} = (1, 1, \dots, 1) \in \mathbb{R}^n$ .

**Proof. Part 1: Multivariate conditional distribution**

We look for the distribution of the random field  $(Y_x)_{x \in X}$  given  $(Y_{x_i} = y_i)_{1 \leq i \leq n}$ .

Let  $r$  be a positive integer and  $(t_1, \dots, t_r) \in X^r$ . We have:

$$(Y_{t_1}, \dots, Y_{t_r}, Y_{x_1}, \dots, Y_{x_n}) = (U, \dots, U) + V(W_{t_1}, \dots, W_{t_r}, W_{x_1}, \dots, W_{x_n})$$

where  $(W_{t_1}, \dots, W_{t_r}, W_{x_1}, \dots, W_{x_n})$  follows the distribution  $\mathcal{N}(0, \Delta)$  with:

$$\Delta := \begin{pmatrix} \Sigma_2 & \mathbf{k}(t) \\ \mathbf{k}(t)^T & \Sigma \end{pmatrix}$$

where  $\mathbf{k}(\mathbf{t}) := (k(t_i, x_j))_{1 \leq i \leq r, 1 \leq j \leq n}$  and  $\mathbf{\Sigma}_2 := (k(t_i, t_j))_{1 \leq i \leq r, 1 \leq j \leq r}$ . As before, we suppose that we are in a generic situation, so that the matrix  $\mathbf{\Delta}$  is positive-definite.

Denote by  $f$  the density of the random vector  $(Y_{t_1}, \dots, Y_{t_r}, Y_{x_1}, \dots, Y_{x_n})$ :

$$f(\zeta) = \frac{1}{2m} \frac{1}{\frac{1}{\epsilon} - \epsilon} \int_{-m}^m \int_{\epsilon}^{\frac{1}{\epsilon}} \frac{1}{(\sqrt{2\pi})^{n+r} v^{n+r} \sqrt{|\mathbf{\Delta}|}} \exp\left(-\frac{1}{2v^2} (\zeta - \mathbf{u}) \mathbf{\Delta}^{-1} (\zeta - \mathbf{u})^T\right) du dv$$

with  $\zeta := (y_{t_1}, y_{t_2}, \dots, y_{t_r}, y_{x_1}, y_{x_2}, \dots, y_{x_n})$  and  $\mathbf{u} := (u, u, \dots, u) \in \mathbb{R}^{n+r}$ .

The conditional density distribution of  $(Y_{t_1}, Y_{t_2}, \dots, Y_{t_r})$  given  $(Y_{x_i} = y_{x_i})_{1 \leq i \leq n}$  is:

$$\mathbf{z} := (y_{t_1}, y_{t_2}, \dots, y_{t_r}) \mapsto \frac{f(\mathbf{z}, \mathbf{y})}{\int_{\mathbb{R}^r} f(\mathbf{z}', \mathbf{y}) d\mathbf{z}'} =: g(\mathbf{z}).$$

After simplification, we get:

$$g(\mathbf{z}) = \frac{\int_{-m}^m \int_{\epsilon}^{\frac{1}{\epsilon}} \frac{1}{v^{n+r}} \exp\left(-\frac{1}{2v^2} ((\mathbf{z} \ \mathbf{y}) - \mathbf{u}) \mathbf{\Delta}^{-1} ((\mathbf{z} \ \mathbf{y}) - \mathbf{u})^T\right) du dv}{\int_{\mathbb{R}^r} \int_{-m}^m \int_{\epsilon}^{\frac{1}{\epsilon}} \frac{1}{v^{n+r}} \exp\left(-\frac{1}{2v^2} ((\mathbf{z}' \ \mathbf{y}) - \mathbf{u}) \mathbf{\Delta}^{-1} ((\mathbf{z}' \ \mathbf{y}) - \mathbf{u})^T\right) d\mathbf{z}' du dv}.$$

Denote:

$$\text{NUM}(\mathbf{z}) := \int_{-m}^m \int_{\epsilon}^{\frac{1}{\epsilon}} \frac{1}{v^{n+r}} \exp\left(-\frac{1}{2v^2} ((\mathbf{z} \ \mathbf{y}) - \mathbf{u}) \mathbf{\Delta}^{-1} ((\mathbf{z} \ \mathbf{y}) - \mathbf{u})^T\right) du dv.$$

Consider the  $u$ -variable polynomial:

$$((\mathbf{z} \ \mathbf{y}) - \mathbf{u}) \mathbf{\Delta}^{-1} ((\mathbf{z} \ \mathbf{y}) - \mathbf{u})^T = au^2 + 2bu + c.$$

We recall that, for  $a, b$  and  $c$  real numbers and  $a > 0$ :

$$\int_{-m}^m \exp\left(-\frac{1}{2v^2} (au^2 + 2bu + c)\right) du \xrightarrow{m \rightarrow +\infty} \frac{v\sqrt{2\pi}}{\sqrt{a}} \exp\left(-\frac{1}{2v^2} \left(c - \frac{b^2}{a}\right)\right)$$

which gives, using the monotone convergence theorem:

$$\text{NUM}(\mathbf{z}) \xrightarrow{m \rightarrow +\infty} \int_{\epsilon}^{\frac{1}{\epsilon}} \frac{\sqrt{2\pi}}{v^{n+r-1}\sqrt{a}} \exp\left(-\frac{1}{2v^2} \left(c - \frac{b^2}{a}\right)\right) dv.$$

With the change of variable  $q = \frac{1}{2v^2} \left(c - \frac{b^2}{a}\right)$  and the definition of Euler function  $\Gamma$ , we get:

$$\text{NUM}(\mathbf{z}) \xrightarrow{m \rightarrow +\infty, \epsilon \rightarrow 0} \frac{\sqrt{2\pi}}{\sqrt{a}} 2^{\frac{n+r-4}{2}} \left(c - \frac{b^2}{a}\right)^{-\frac{n+r-2}{2}} \Gamma\left(\frac{n+r-2}{2}\right).$$

Note that  $au^2 + 2bu + c$  is always positive since we exclude degenerate situations. Thus, its discriminant is negative, that is  $(c - b^2/a) > 0$ . In addition, the expression  $(c - b^2/a)$  is a positive polynomial of degree 2 in the variable  $\mathbf{z}$ , and we write

$$c - \frac{b^2}{a} = Q(\mathbf{z}) + L(\mathbf{z}) + M$$

where  $Q$  is a positive-definite quadratic form,  $L$  is a linear form and  $M$  is constant.

By monotone convergence, the limit of the denominator in the expression of the density  $g$  is what we expect and we obtain:

$$g(\mathbf{z}) \xrightarrow{m \rightarrow +\infty, \epsilon \rightarrow 0} \frac{\frac{\sqrt{2\pi}}{\sqrt{a}} 2^{\frac{n+r-4}{2}} (Q(\mathbf{z}) + L(\mathbf{z}) + M)^{-\frac{n+r-2}{2}} \Gamma\left(\frac{n+r-2}{2}\right)}{\int_{\mathbb{R}^r} \frac{\sqrt{2\pi}}{\sqrt{a}} 2^{\frac{n+r-4}{2}} (Q(\mathbf{z}') + L(\mathbf{z}') + M)^{-\frac{n+r-2}{2}} \Gamma\left(\frac{n+r-2}{2}\right) d\mathbf{z}'},$$

that is:

$$g(\mathbf{z}) \xrightarrow{m \rightarrow +\infty, \epsilon \rightarrow 0} \frac{(Q(\mathbf{z}) + L(\mathbf{z}) + M)^{-\frac{n+r-2}{2}}}{\int_{\mathbb{R}^r} (Q(\mathbf{z}') + L(\mathbf{z}') + M)^{-\frac{n+r-2}{2}} d\mathbf{z}'}$$

As  $Q(\mathbf{z}) + L(\mathbf{z}) + M$  is a positive polynomial, we get with Lemma 1 proven hereafter, for  $n > 2$ :

$$\int_{\mathbb{R}^r} (Q(\mathbf{z}') + L(\mathbf{z}') + M)^{-\frac{n+r-2}{2}} d\mathbf{z}' < +\infty.$$

With Lemma 2, we conclude that  $g(\mathbf{z})$  is a multivariate Student distribution, when  $m$  goes to infinity and  $\epsilon$  goes to zero.

**Part 2: Univariate conditional distribution**

We can give more precise formulae for one dimensional marginals. If  $r = 1$ , we have

$$g(\mathbf{z}) \xrightarrow{m \rightarrow +\infty, \epsilon \rightarrow 0} \frac{(Q(\mathbf{z}) + L(\mathbf{z}) + M)^{-\frac{n-1}{2}}}{\int_{\mathbb{R}} (Q(\mathbf{z}') + L(\mathbf{z}') + M)^{-\frac{n-1}{2}} d\mathbf{z}'}$$

Let us come back to the definitions of  $a$ ,  $b$  and  $c$  and compute the expression  $(c - b^2/a)$ . After straightforward calculations, we get:

$$\Delta^{-1} = \frac{1}{\lambda} \begin{pmatrix} 1 & -\mathbf{k}(t)\Sigma^{-1} \\ -\Sigma^{-1}\mathbf{k}(t)^T & \lambda\Sigma^{-1} + \Sigma^{-1}\mathbf{k}(t)^T\mathbf{k}(t)\Sigma^{-1} \end{pmatrix}$$

with  $\lambda := 1 - \mathbf{k}(t)\Sigma^{-1}\mathbf{k}(t)^T$ . Then:

$$\begin{aligned} a &= \mathbf{1}\Delta^{-1}\mathbf{1}^T \\ b &= \frac{1}{\lambda} (z - \mathbf{k}(t)\Sigma^{-1}\mathbf{y}^T) (1 - \mathbf{k}(t)\Sigma^{-1}\mathbf{1}^T) + \mathbf{y}\Sigma^{-1}\mathbf{1}^T \\ c &= \frac{1}{\lambda} (z - \mathbf{k}(t)\Sigma^{-1}\mathbf{y}^T)^2 + \mathbf{y}\Sigma^{-1}\mathbf{y}^T \end{aligned}$$

with  $\mathbf{1} = (1, 1, \dots, 1) \in \mathbb{R}^n$ .

Set  $z' := z - \mathbf{k}(t)\Sigma^{-1}\mathbf{y}^T$ . This change of variable yields:

$$c - \frac{b^2}{a} = qz'^2 - 2lz' + m$$

with

$$\begin{aligned} q &= \frac{1}{\lambda} \left( 1 - \frac{1}{\lambda a} (1 - \mathbf{k}(t)\Sigma^{-1}\mathbf{1}^T)^2 \right) \\ l &= \frac{1}{\lambda a} (1 - \mathbf{k}(t)\Sigma^{-1}\mathbf{1}^T) (\mathbf{y}\Sigma^{-1}\mathbf{1}^T)^T \\ m &= \mathbf{y}\Sigma^{-1}\mathbf{y}^T - \frac{1}{a} (\mathbf{y}\Sigma^{-1}\mathbf{1}^T)^2. \end{aligned}$$

In the variable  $z'$ , we have a Student distribution with  $n - 2$  degrees of freedom, with location parameter

$$\frac{l}{q} = \mu (1 - \mathbf{k}(x)\Sigma^{-1}\mathbf{1}^T) \quad \text{with} \quad \mu := \frac{\mathbf{y}\Sigma^{-1}\mathbf{1}^T}{\mathbf{1}\Sigma^{-1}\mathbf{1}^T}$$

and scale parameter

$$\sqrt{\frac{1}{n-2} \left( \frac{m}{q} - \left( \frac{l}{q} \right)^2 \right)} = \sqrt{\frac{1}{n-2} ((\mathbf{y} - \mu\mathbf{1})\Sigma^{-1}\mathbf{y}^T) \left( 1 - \mathbf{k}(x)\Sigma^{-1}\mathbf{k}(x)^T + \frac{(1 - \mathbf{1}\Sigma^{-1}\mathbf{k}(x)^T)^2}{\mathbf{1}\Sigma^{-1}\mathbf{1}^T} \right)}.$$

Then, the conditional distribution of the random variables  $Y_x$  is a Student distribution with  $n - 2$  degrees of freedom, with location parameter:

$$\mu + \mathbf{k}(x)\Sigma^{-1}(\mathbf{y} - \mu\mathbf{1})^T \quad \text{with} \quad \mu := \frac{\mathbf{y}\Sigma^{-1}\mathbf{1}^T}{\mathbf{1}\Sigma^{-1}\mathbf{1}^T}$$

and scale parameter:

$$\sqrt{\frac{1}{n-2} ((\mathbf{y} - \mu\mathbf{1})\Sigma^{-1}\mathbf{y}^T) \left( 1 - \mathbf{k}(x)\Sigma^{-1}\mathbf{k}(x)^T + \frac{(1 - \mathbf{1}\Sigma^{-1}\mathbf{k}(x)^T)^2}{\mathbf{1}\Sigma^{-1}\mathbf{1}^T} \right)}$$

with  $\mathbf{1} = (1, 1, \dots, 1) \in \mathbb{R}^n$ . □

**Lemma 1.** *Let  $p(\mathbf{z})$  be a polynomial of degree 2 in  $r$  variables. Suppose that:*

- *for all  $\mathbf{z} \in \mathbb{R}^r$ ,  $p(\mathbf{z}) > 0$ ,*
- *the quadratic part of  $p$  is positive definite.*

*If  $n > 0$  then*

$$\int_{\mathbb{R}^r} p(\mathbf{z})^{-\frac{n+r}{2}} d\mathbf{z} < \infty.$$

*Proof of Lemma 1.* With an appropriate change of orthonormal basis, we can write:

$$p(\mathbf{z}) = \sum_{i=1}^r d_i e_i^2 + \sum_{i=1}^r s_i e_i + M$$

where  $(e_1, e_2, \dots, e_r) := \mathbf{e}$  are the coordinates of  $\mathbf{z}$  in the new basis, and where, for all  $1 \leq i \leq r$ ,  $d_i > 0$ .

We consider a ball  $B$  in  $\mathbb{R}^r$ , containing zero, such that for all  $\mathbf{z} \notin B$ , we have:

$$p(\mathbf{z}) \geq \frac{1}{2} \sum_{i=1}^r d_i e_i^2 \geq \frac{C}{2} \sum_{i=1}^r e_i^2$$

where  $C = \min(d_1, d_2, \dots, d_r)$ . Denote  $B^c$  the complementary of  $B$  in the space  $\mathbb{R}^r$ . As  $p > 0$ , the integral over  $B$  of the function  $p^{-\frac{n+r}{2}}$  is finite. We have:

$$\int_{B^c} p(\mathbf{z})^{-\frac{n+r}{2}} d\mathbf{z} \leq C' \int_{B^c} \left( \sum_{i=1}^r e_i^2 \right)^{-\frac{n+r}{2}} de_1 de_2 \dots de_r$$

with  $C' := \left(\frac{C}{2}\right)^{-\frac{n+r}{2}}$ .

That means:

$$\int_{B^c} p(\mathbf{z})^{-\frac{n+r}{2}} d\mathbf{z} \leq C' \int_{B^c} \|\mathbf{e}\|^{-(n+r)} de_1 de_2 \dots de_r.$$

But,

$$\int_{B^c} \|\mathbf{e}\|^{-(n+r)} de_1 de_2 \dots de_r < \infty \quad \text{if and only if} \quad n > 0$$

so we conclude:

$$\int_{\mathbb{R}^r} p(\mathbf{z})^{-\frac{n+r}{2}} d\mathbf{z} < \infty \quad \text{if} \quad n > 0.$$

□

**Lemma 2.** Let  $n$  and  $r$  be two positive integers with  $n \geq 3$ . Let  $Q$  be a positive-definite quadratic form in  $r$  variables,  $L$  be a linear form in  $r$  variables and  $M$  be a positive constant. Suppose that for all  $\mathbf{z} \in \mathbb{R}^r$ ,  $Q(\mathbf{z}) + L(\mathbf{z}) + M > 0$ . Define  $C := \left( \int_{\mathbb{R}^r} (Q(\mathbf{z}) + L(\mathbf{z}) + M)^{-\frac{n+r-2}{2}} d\mathbf{z} \right)^{-1}$ . The density of probability  $f$  on  $\mathbb{R}^r$  of the form:

$$f : \mathbf{z} \mapsto C (Q(\mathbf{z}) + L(\mathbf{z}) + M)^{-\frac{n+r-2}{2}}$$

is a multivariate Student density with  $n - 2$  degree of freedom.

*Proof of Lemma 2.* Denote  $\mathbf{z} := (z_1, z_2, \dots, z_r)$ ,  $Q(\mathbf{z}) := \sum_{i=1}^r q_{ii} z_i^2 + 2 \sum_{1 \leq i < j \leq r} q_{ij} z_i z_j$  and  $L(\mathbf{z}) := \sum_{i=1}^r \ell_i z_i$ . Denote also  $\mathbf{L} = (\ell_1, \ell_2, \dots, \ell_r)$  and  $\mathbf{Q} = (q_{ij})_{1 \leq i, j \leq r}$ . Then:

$$L(\mathbf{z}) = \mathbf{L}\mathbf{z}^T \quad \text{and} \quad Q(\mathbf{z}) = \mathbf{z}\mathbf{Q}\mathbf{z}^T.$$

Diagonalize the positive-definite matrix  $\mathbf{Q}$ ; there exists an orthogonal matrix  $\mathbf{P}$ , of type  $(r, r)$  such that:

$$\mathbf{Q} = \mathbf{P}\mathbf{A}\mathbf{P}^{-1} = \mathbf{P}\mathbf{A}\mathbf{P}^T$$

where  $\mathbf{A}$  is a diagonal matrix with diagonal coefficients  $a_1, a_2, \dots, a_r > 0$ .

Denote  $\mathbf{s} = \mathbf{z}\mathbf{P}$  and  $\mathbf{s} = (s_1, s_2, \dots, s_r)$ . We can write:

$$f(\mathbf{z}) = C (M + \mathbf{L}\mathbf{P}\mathbf{s}^T + \mathbf{s}\mathbf{A}\mathbf{s}^T)^{-\frac{n+r-2}{2}}.$$

With  $\mathbf{L}\mathbf{P} = (b_1, b_2, \dots, b_r)$  we deduce:

$$\mathbf{L}\mathbf{P}\mathbf{s}^T + \mathbf{s}\mathbf{A}\mathbf{s}^T = \sum_{i=1}^r b_i s_i + a_i s_i^2 = \sum_{i=1}^r a_i \left( s_i + \frac{b_i}{2a_i} \right)^2 - \frac{b_i^2}{4a_i}.$$

Denote  $\mathbf{B}$  the vector  $\left( \frac{b_i}{2a_i} \right)_{1 \leq i \leq r}$  and  $\mathbf{t} := \mathbf{s} + \mathbf{B}$ . Denote also  $d = M - \sum_{i=1}^r \frac{b_i^2}{4a_i}$ .

We have:

$$f(\mathbf{z}) = C (d + \mathbf{t}\mathbf{A}\mathbf{t}^T)^{-\frac{n+r-2}{2}}.$$

Remark that  $d > 0$  since  $d + \mathbf{t}\mathbf{A}\mathbf{t}^T = Q(\mathbf{z}) + L(\mathbf{z}) + M$  which is positive for all  $\mathbf{z}$ , thus for all  $\mathbf{t}$ . It is therefore possible to factor our  $d$ :

$$\begin{aligned} f(\mathbf{z}) &= C d^{-\frac{n+r-2}{2}} \left( 1 + \frac{1}{d} (\mathbf{z}\mathbf{P} + \mathbf{B}) \mathbf{A} (\mathbf{z}\mathbf{P} + \mathbf{B})^T \right)^{-\frac{n+r-2}{2}} \\ &= C d^{-\frac{n+r-2}{2}} \left( 1 + \left( \mathbf{z} + \mathbf{B}\mathbf{P}^T \right) \left( \frac{1}{d} \mathbf{Q} \right) \left( \mathbf{z} + \mathbf{B}\mathbf{P}^T \right)^T \right)^{-\frac{n+r-2}{2}}. \end{aligned}$$

With  $C' := C d^{-\frac{n+r-2}{2}}$ ,  $\boldsymbol{\mu} := -\mathbf{B}\mathbf{P}^T$  and  $\boldsymbol{\Pi} := \frac{d}{n-2} \mathbf{Q}^{-1}$ , finally, the following formula is obtained for the density of probability:

$$f(\mathbf{z}) = C' \left( 1 + \frac{1}{n-2} (\mathbf{z} - \boldsymbol{\mu}) \boldsymbol{\Pi}^{-1} (\mathbf{z} - \boldsymbol{\mu})^T \right)^{-\frac{n+r-2}{2}}.$$

According to the book of Kotz and Nadarajah (2004), page 1, we recognize the multivariate t-distribution in  $r$  dimension, with  $n - 2$  degrees of freedom, with scale matrix  $\boldsymbol{\Pi}$  and location vector  $\boldsymbol{\mu}$ .  $\square$

## References

Kotz, S. and Nadarajah, S. (2004). *Multivariate t-distributions and their applications*. Cambridge University Press.