T-distributed Random Fields: A Parametric Model for Heavy-tailed Random Fields

Jo Røislien[2], Henning Omre[3]

[2]Department of Petroleum Engineering and Applied Geophysics, Norwegian University of Science and Technology, Norway; e-mail: jo.roislien@rikshospitalet.no

[3]Department of Mathematical Sciences, Norwegian University of Science and Technology, Norway; e-mail: omre@math.ntnu.no

ABSTRACT

Histograms of observations from spatial phenomena are often found to be more heavy-tailed than Gaussian distributions, which makes the Gaussian random field model unsuited. A T-distributed random field model with heavy-tailed marginal probability density functions is defined. The model is a generalization of the familiar Student-T distribution, and it may be given a Bayesian interpretation. The increased variability appears cross-realizations, contrary to in-realizations, since all realizations are Gaussian-like with varying variance between realizations. The T-distributed random field model is analytically tractable and the conditional model is developed, which provides algorithms for conditional simulation and prediction, so-called T-kriging. The model compares favourably with most previously defined random field models. The Gaussian random field model appears as a special, limiting case of the T-distributed random field model. The model is particularly useful whenever multiple, sparsely sampled realizations of the random field are available, and is clearly favourable to the Gaussian model in this case. The properties of the T-distributed random field model is demonstrated on well log observations from the Gullfaks field in the North Sea. The predictions correspond to traditional kriging predictions, while the associated prediction variances are more representative, as they are layer specific and include uncertainty caused by using variance estimates.

KEY WORDS: random fields, hierarchical models, sampling, kriging, parameter estimation.

INTRODUCTION

Histograms of observations from spatial phenomena, for example from well log data, often appear more heavy-tailed than the Gaussian distribution. This feature requires more general models than the ones usually defined in traditional geostatistics. The disadvantage is that these models are often complicated to handle mathematically.

In recent years, a series of papers have been presented on parametric stochastic models for continuous random fields (RF) with non-Gaussian marginal probability density functions (pdf). Gaussian RF (GRF) models are of course most widely used in practice (Chilès and Delfiner, 1999). The normal-score approach makes a univariate ϕ-transform of the random variables (RV) into a Gaussian marginal pdf, for then to assume a GRF model. The normal-score model, ϕ^{-1}-GRF model, is widely used as a model for RFs with non-Gaussian marginal pdfs (Chilès and Delfiner, 1999). The advantage of the ϕ^{-1}-GRF model is its simplicity and flexibility, but its disadvantage is the lack of analytical tractability. In the current paper RFs with symmetric, unimodal marginal pdfs, heavy-tailed or not, are considered. In a recent paper (Gunning, 2002) the use of Lévy-Stable RFs (LSRF) as models for such RFs is discussed. Gunning presents an interesting discussion on desirable features of stochastic models for continuous RFs, and the LSRF is evaluated with respect to these features. The current paper can be seen as an extension of Gunnings work. We elaborate on his list of features, but suggest another class of RF models. Our list of desirable features of parametric stochastic models for continuous RFs is as follows:

- *Fully specified probabilistic model*
  This is a requirement for simulation of the RF. RF models based on contrasts and increments often leave parameters in the trend unspecified, and hence they are unsuited for simulation.

- *Permutation invariance*
  The RF model must be exchangeable in its components. This is a requirement for RF models (Yaglom, 1962).

- *Probabilistic consistency*
  Finite dimensional marginal pdfs of different dimensions defined by the RF model must be such that they coincide when non-common dimensions are integrated out. This is another requirement for RF models (Yaglom, 1962).

- *Marginal invariance*
  Finite dimensional marginal pdfs defined by the RF model should all belong to the same parametric class of pdfs. This makes it easier to ensure probilistic consistency, and simplifies analytical work.

- *Closed under additivity*
  Often, observations are collected as averages over sampling volumes, so-called sample support. The objective is frequently to determine spatial averages over a given volume. If spatial averages define an integrated RF which belongs to the same parametric class as the initial RF, this will normally simplify the analytical work.

- *Analytical tractability*
  The objective is usually to determine the conditional RF model conditioned on a set of observations in arbitrary locations. In order to define this conditional RF model,

the non-observed dimensions of the RF model must be integrated out. This is very difficult unless the RF model is analytically tractable. Numerical integration in such high dimensions is virtually impossible. Analytical tractability will also make it easier to define efficient simulation algorithms for the RF.

- *General model*
  It is of course desirable that the parametric RF model is general in the sense that the admissible range of parameter values define a large variety of RF models. It would of course be convenient if marginal pdfs of all kinds could be modelled. Moreover, it would be fine if certain parameter combinations provide familiar RF models, the GRF model in particular.

- *Model parameter inference*
  The parametric RF model is by definition fully specified by a set of parameters. It is of course desirable that the number of parameters is low, and that the parameters are interpretable. Most important, however, is that reliable estimators can be defined based on a set of observations from the RF.

- *Diminishing spatial dependence*
  Random variables in two locations should approach independence as their interdistance increases. This property is related to the ergodic characteristics of the RF, and will make estimators for model parameters more reliable. Without diminishing spatial dependence one may not be able to define consistent estimators for the model parameters based on observations from one single realization of the RF.

GRF models are by far the most used in practice, and the GRF is normally defined as follows:

Definition 1 ***Gaussian Random Field (GRF)***
A RF $\{Z(x); x \in \mathcal{D} \subset \mathbb{R}^n\}$ is termed a GRF if

$$Z = [Z(x_1), \ldots, Z(x_m)]^T \quad \sim \quad N_m(\mu, \Phi)$$

with pdf

$$f(z) \quad = \quad \frac{1}{(2\pi)^{m/2}} |\Phi|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2}(z-\mu)^T \Phi^{-1} (z-\mu) \right\}$$

for all configurations $(x_1, \ldots, x_m) \in \mathcal{D} \times \ldots \times \mathcal{D}$ and all $m \in \mathcal{N}_+$ where $N_m(\mu, \Phi)$ represents the multivariate Gaussian distribution with parameters (μ, Φ) of proper dimensions. The GRF is parameterized $G_x(\mu_x, \phi_{xx})$ where

$$\begin{aligned} \mu_x &: \{\mu(x); x \in \mathcal{D}\} \\ \phi_{xx} &: \{\phi(x', x''); (x', x'') \in \mathcal{D} \times \mathcal{D}\} \end{aligned}$$

with μ_x the expectation function over $\mathcal{D}$, and ϕ_{xx} a positive definite covariance function over $\mathcal{D} \times \mathcal{D}$. ∎

The GRF model meets most of the desirable features listed above. Being a general model is the only feature really being violated. The GRF model constitutes an extreme case among RF models in the sense that it maximizes entropy, i.e., it is the smoothest model possible given the model parameters. This is also related to the central limit theorem – everything averages out to the GRF. The GRF model exhibits two unfavourable features in particular: The Gaussian marginal pdf has extremely light tails, and the prediction variance is only dependent on the location configuration of the observations – not on the values actually observed.

The objective of the current paper is to define a RF model with symmetric, unimodal marginal pdfs with large flexibility in tail behaviour. The T-distributed RF (TRF) model is introduced and its properties explored. The TRF model can also be given a Bayesian interpretation as an extension of the work in Kitanidis (1986), Omre (1987) and Hjort and Omre (1994).

The paper is organized as follows; the next section contains a motivation for the study, before the TRF model is defined and its properties discussed. Then T-kriging is introduced, and estimators for the model parameters of the TRF are defined. A case study on the real data explored in the introduction is included, and the two last sections contain a discussion of the characteristics of the TRF model and some concluding remarks.

MOTIVATION

Consider the density log from a well in the Gullfaks field in the North Sea (Fig. 1, left display). The well seems to penetrate several horisontal sedimentary layers with varying properties, and a segmentation into layers is made (Fig. 1, right display). The layered structure is assumed to be caused by abrupt changes in the depositional processes. The list of layer averages and empirical variances of the nine layers is presented in Table 1. If a classification of the layers based on the average values is made, layer 1, 2, 3, 5, 7 and 9 will be pooled in one class. One could of course have used empirical variances for further classification, but we have chosen not to do so here, because empirical variances are often very sensitive to individual samples. Histograms of the six pooled layers look fairly dense with light tails (Fig. 2). However, the pooled histogram from the six layers (Fig. 3, left display), appears as peaked with relatively heavy tails. In the right display of Figure 3, a kernel-smoothed pooled histogram is presented together with maximum likelihood adapted Gaussian and Student-T pdfs. Note that the Student-T pdf reproduces the histogram much more reliably than the Gaussian.

To summarize, by compiling relatively dense, light-tailed histograms from the individual layers, one obtains a composite peaked, heavy-tailed histogram. This change in shape appears since the layer-histograms have different variances. Recall that identical centering is ensured by the initial classification. Statisticians are aware that the composite of Gaussian RVs, identically centered with varying variances, may appear as Student-T distributed (Walpole and Myers, 1993). In the current study, the well logs in the individual layers are considered to be GRFs, equally centered but with different variances, which appear as outcomes from a unifying TRF.

The evaluation above is made in a well with numerous observations in each layer. If only a few observations in each layer are available, for example as core plug samples, one may wish to interpolate or simulate along the entire well trace and extend this into the three-dimensional reservoir. The layers may be classified with respect to averages, while further classification based on empirical variances may be considered unreliable. The composite histogram may appear Student-T like, hence a RF model with a Student-T marginal pdf would be preferred. Based on this RF model, interpolation and simulation in the individual layers conditional on the few available obervations in that particular layer can be made. This setting constitutes the challenge of the study.

T-DISTRIBUTED RANDOM FIELDS

In order to define a suitable RF model for the data presented in the previous section, a survey of multivariate T-distributed RVs is presented in the following subsection. In the subsequent subsections, TRFs, simple and conditional, are defined.

Multivariate T-distributed Random Variables

The multivariate T-distributed RV is a generalization of the Student-T distributed RV introduced in most introductory courses in statistics. The exposure here is mainly based on results in Cornish (1954), Johnson and Kotz (1972) and Welsh (1996).

Definition 2 ***T-distributed random variable***
A RV $Z \in \mathbb{R}^m$ is multivariate T-distributed

$$Z \sim T_m(\mu, \Omega, \nu)$$

if its pdf is

$$f(z) = \frac{\Gamma(\frac{\nu+m}{2})}{\Gamma(\frac{\nu}{2})(\nu\pi)^{m/2}} |\Omega|^{-\frac{1}{2}} \left[1 + \frac{1}{\nu}(z-\mu)^T \Omega^{-1}(z-\mu)\right]^{-\frac{\nu+m}{2}}$$

where $\Gamma(x)$ is the gamma function, and with $\mu \in \mathbb{R}^m$ a centering vector, $\Omega \in \mathbb{R}^m \times \mathbb{R}^m$ a positive definite scale/dependence matrix, and $\nu \in \mathbb{R}_+$ the degrees of freedom. ∎

This definition specifies a spherical-symmetric pdf centered at μ with Ω controlling scale and multivariate dependence, while ν controls the tail behaviour (Mardia, Kent and Bibby, 1979). In Figures 4 and 5 $T_2(0, \Omega, \nu)$ variables are displayed for varying values of Ω and ν. Note that all bivariate pdfs have spherical contour lines, but the general shape appears with varying peakedness and heavy-tailedness.

The properties of the multivariate T-distributed RV are summarized below:

Result 1 ***Properties of multivariate T-distributed random variables***
Let $Z \in \mathbb{R}^m$ have pdf as in Definition 2. Then the following properties can be demonstrated:

1. *Special cases:*

$$\begin{aligned} T_m(\mu, \Omega, \nu) &\overset{\nu\to\infty}{\longrightarrow} N_m(\mu, \Omega) \\ T_m(\mu, \Omega, 1) &= C_m(\mu, \Omega) \end{aligned}$$

   where $N_m(\mu, \Omega)$ and $C_m(\mu, \Omega)$ are the multivariate Gaussian and Cauchy distributions, respectively.

2. *Moments:*

$$\begin{aligned} \mathrm{E}\{Z\} &= \mu \quad ; \quad \nu \geq 2 \\ \mathrm{Cov}\{Z\} &= \tfrac{\nu}{\nu-2}\Omega \quad ; \quad \nu \geq 3 \end{aligned}$$

   while for ν less than the specified values the moments are infinite.

3. *Linear transform:*
   Let A be an arbitrary known $(k \times m)$-matrix, and
$$Z_A = AZ.$$
   Then
$$Z_A \sim T_k(A\mu, A\Omega A^T, \nu).$$
   Note that this entails that all marginal pdf s of Z are T-distributed as well, since Z_A is a marginal of Z for certain choices of binary A.

4. *Conditional distributions:*
   Let Z_A be defined as in Point 3. Then
$$[Z|Z_A = z_A] \sim T_m\left(\mu_{\cdot|z_A}, \Omega_{\cdot|z_A}, \nu + k\right)$$
   where
$$\begin{aligned} \mu_{\cdot|z_A} &= \mu + \Omega A^T(A\Omega A^T)^{-1}(z_A - A\mu) \\ \Omega_{\cdot|z_A} &= \xi(z_A)\left[\Omega - \Omega A^T(A\Omega A^T)^{-1}A\Omega\right] \end{aligned}$$
   with
$$\xi(z_A) = \frac{1}{1+\frac{k}{\nu}}\left[1 + \frac{1}{\nu}(z_A - A\mu)^T(A\Omega A^T)^{-1}(z_A - A\mu)\right].$$
   Note that this entails that all pdf s conditional on arbitrary linear combinations are T-distributed.

5. *Non-independence:*
   Let Z be arbitrarily decomposed as follows:
$$Z = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}, \ \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \ \Omega = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{bmatrix}$$
   with $Z_1 \in \mathbb{R}^{m_1}$ and $Z_2 \in \mathbb{R}^{m_2}$ and $m = m_1 + m_2$. Then
$$T_m(\mu, \Omega, \nu) \neq T_{m_1}(\mu_1, \Omega_{11}, \nu) \times T_{m_2}(\mu_2, \Omega_{22}, \nu)\ ;\ \nu < \infty$$
   even for the particular case Ω_{12} and Ω_{21} being matrices containing only zeros. This entails that for $\nu < \infty$, Z_1 and Z_2 will not be independent even if Ω is a block matrix. In the limit $\nu \to \infty$ independence will be obtained, as the multivariate Gaussian case is reached.

6. *Decomposition:*
   Let H and U be independent RVs. Then
$$Z = \mu + H^{-\frac{1}{2}}\Omega^{\frac{1}{2}}U$$
   with
$$\begin{aligned} \nu H &\sim \chi^2(\nu) \\ U &\sim N_m(0, I_m) \end{aligned}$$
   where $\chi^2(\nu)$ is a univariate chi-squared distribution with ν degrees of freedom, and I_m is a $(m \times m)$ unit-diagonal matrix.

7. *Hierarchical representation:*
Let Z *conditional on the random parameters* (β, ϕ^2) *be distributed as*

$$[Z|\beta, \phi^2] \sim N_m(G\beta, \phi^2 \Phi_0^Z)$$

where G *is an arbitrary known* $(m \times k)$ *matrix,* β *is a random* $(k \times 1)$ *vector,* ϕ^2 *is a univariate RV and* Φ_0^Z *is a known positive definite* $(m \times m)$ *correlation matrix, with*

$$\begin{aligned} [\beta|\phi^2] &\sim N_k(\mu_\beta, \phi^2 \Phi_0^\beta) \\ \phi^2 &\sim IG\left(\frac{\nu}{2}, \frac{\nu\omega^2}{2}\right) \end{aligned}$$

where μ_β *and* Φ_0^β *is the expectation vector and correlation matrix of appropriate dimensions, and* $IG(\kappa, \lambda)$ *represents the inverse gamma pdf*

$$f(\phi^2) = \frac{1}{\Gamma(\kappa)} \lambda^\kappa \left(\frac{1}{\phi^2}\right)^{\kappa+1} e^{-\frac{\lambda}{\phi^2}} \ ; \ \kappa, \lambda > 0, \phi^2 > 0.$$

Then

$$Z \sim T_m(\mu, \Omega, \nu)$$

where

$$\begin{aligned} \mu &= G^T \mu_\beta \\ \Omega &= \omega^2 \Omega_0 = \omega^2 [\Phi_0^Z + G^T \Phi_0^\beta G]. \end{aligned}$$

Note that this representation can be interpreted in a Bayesian setting with (β, ϕ^2) *being random hyperparameters.*

T-distributed Random Fields

The TRF is defined by the multivariate T-distributed RV along the lines of the GRF in Definition 1.

Definition 3 ***T-distributed Random Field (TRF)***
A RF $\{Z(x); x \in \mathcal{D} \subset \mathbb{R}^n\}$ *is termed a TRF if*

$$Z = [Z(x_1), \ldots, Z(x_m)]^T \sim T_m(\mu, \Omega, \nu)$$

for all configurations $(x_1, \ldots, x_m) \in \mathcal{D} \times \cdots \times \mathcal{D}$ *and all* $m \in \mathcal{N}_+$ *where* $T_m(\mu, \Omega, \nu)$ *represents the multivariate* T*-distribution with parameters* (μ, Ω, ν)*. The TRF is parameterized* $T_x(\mu_x, \omega_{xx}, \nu)$ *where*

$$\begin{aligned} \mu_x &: \{\mu(x); x \in \mathcal{D}\} \\ w_{xx} &; \{\omega(x', x''); (x', x'') \in \mathcal{D} \times \mathcal{D}\} \end{aligned}$$

with μ_x *the centering function over* $\mathcal{D}$ *and* ω_{xx} *a positive definite scale/dependence function over* $\mathcal{D} \times \mathcal{D}$*, and* ν *the degrees of freedom.* ∎

Note that from the definition of TRFs and Result 1.1 it is obvious that a TRF tends towards a GRF whenever $\nu \to \infty$. T-distributed RVs are in Result 1.3 shown to be closed under linear transformations. This entails that a differential TRF, if it exists, also will be a TRF. The same holds for an integrated TRF.

In Figures 6 through 9 realizations of $T_x(0, \omega_{xx}, \nu)$ with varying ν are displayed. The scale/dependence function ω_{xx} has form

$$\omega(x', x'') = \exp\left\{-\alpha(x'' - x')^2\right\}$$

with $\alpha = \frac{1}{50}$, and ν takes values 1, 3, 7, and ∞. Note that the variability between realizations. seems to decline as ν increases and the TRF approaches a GRF. The effect of Result 1.7 can be observed by the fact that each realization is Gaussian-like, with large cross-realization variability.

The hierarchical representation of T-distributed RVs in Result 1.7 can be used to define:

Result 2 ***Hierarchical representation of TRFs***
Let $\{Z(x); x \in \mathcal{D}\}$ *conditional on the random parameters* (β, ϕ^2) *be:*

$$\{[Z(x)|\beta, \phi^2]; x \in \mathcal{D}\} \sim G_x(g_x^T \beta, \phi^2 \phi_{0xx}^Z)$$

where

$$\begin{aligned} g_x &: \{g(x) = \left(g_1(x), \ldots, g_k(x)\right)^T; x \in \mathcal{D}\} \\ \phi_{0xx}^Z &: \{\phi_0^Z(x', x''); x', x'' \in \mathcal{D} \times \mathcal{D}\} \end{aligned}$$

with g_x *and* ϕ_{0xx}^Z *a known trend function and a spatial correlation function respectively, and* β *a random* $(k \times 1)$ *vector and* ϕ^2 *a univariate random variance. Further, let*

$$\begin{aligned} [\beta|\phi^2] &\sim N_k\left(\mu_\beta, \phi^2 \Phi_0^\beta\right) \\ \phi^2 &\sim IG\left(\frac{\nu}{2}, \frac{\nu\omega^2}{2}\right) \end{aligned}$$

where μ_β *and* Φ_0^β *are expectations and correlation matrix of appropriate dimensions. Then*

$$\{Z(x); x \in \mathcal{D}\} \sim T_x(\mu_x, \omega_{xx}, \nu)$$

with

$$\begin{aligned} \mu_x &: \{\mu(x) = g(x)^T \mu_\beta; x \in \mathcal{D}\} \\ \omega_{xx} &: \{\omega(x', x'') = \omega^2 \omega_0(x', x'') \\ &\qquad = \omega^2[\phi_0^Z(x', x'') + g(x')^T \Phi_0^\beta g(x'')]; x', x'' \in \mathcal{D} \times \mathcal{D}\}. \end{aligned}$$

∎

This hierarchical representation of a TRF can be given a Bayesian interpretation where the model parameters (β, ϕ^2) in a GRF is assigned appropriate prior models. Note that this corresponds to the Bayesian kriging models discussed in Kitanidis (1986), Omre (1987), Le and Zidek (1991), Handcock and Stein (1993) and Hjort and Omre (1994).

The hierarchical representation in Result 2 can be used for simulation of TRFs, and a suitable algorithm is as follows:

Algorithm 1 ***Simulation of TRFs***
Simulate a TRF $\{Z(x); \in \mathcal{D}\} \sim T_x\Big(g_x^T\mu_\beta, \omega^2[\phi_{0xx}^Z + g_x^T\Phi_0^\beta g_x], \nu\Big)$ *by:*

- *generate* ϕ^2 *from* $IG(\frac{\nu}{2}, \frac{\nu\omega^2}{2})$
- *given* ϕ^2*, generate* β *from* $N_k(\mu_\beta, \phi^2\Phi_0^\beta)$
- *given* (β, ϕ^2)*, generate*

$$\{Z(x); x \in \mathcal{D}\} \sim G_x(g_x^T\beta, \phi^2\phi_{0xx}^Z).$$

This algorithm can of course be very efficient since β and ϕ^2 are low-dimensional, and any fast algorithm for simulation of GRFs can be used.

Conditional T-distributed Random Fields

Consider a TRF $\{Z(x); x \in \mathcal{D}\}$ and let $Z^d = [Z(x_1), \ldots, Z(x_n)]^T$ be a set of observations in arbitrary locations $(x_1, \ldots, x_n) \in \mathcal{D} \times \ldots \times \mathcal{D}$ with associated realization $z^d = (z_1, \ldots z_n)^T$. The conditional TRF is denoted $\{[Z(x)|z^d]; x \in \mathcal{D}\}$. From Result 1.4 the conditional TRF is given as
Result 3 ***Conditional TRF***
A TRF $\{Z(x); x \in \mathcal{D}\} \sim T_x(\mu_x, \omega_{xx}, \nu)$ *conditional on the* $(n \times 1)$*-dimensional vector of observations* $Z^d = z^d$ *is:*

$$\{[Z(x)|z^d]; x \in \mathcal{D}\} \sim T_x\left(\mu_{x|z^d}, \omega_{xx|z^d}, \nu + n\right)$$

where

$$\begin{aligned}
\mu_{x|z^d} &: \left\{[\mu(x)|z^d] = \mu(x) + \omega_{xd}^T\Omega_{dd}^{-1}(z^d - \mu_d); x \in \mathcal{D}\right\} \\
\omega_{xx|z^d} &: \left\{[\omega(x', x'')|z^d] = \xi(z^d)\left[\omega(x', x'') - \omega_{x'd}^T\Omega_{dd}^{-1}\omega_{x''d}\right]; (x', x'') \in \mathcal{D} \times \mathcal{D}\right\}
\end{aligned}$$

with

$$\begin{aligned}
\mu_d &= (\mu(x_1), \ldots, \mu(x_n))^T \\
\omega_{xd} &= (\omega(x, x_1), \ldots, \omega(x, x_n))^T \\
\Omega_{dd} &= \begin{bmatrix} \omega(x_1, x_1) & \cdots & \omega(x_1, x_n) \\ \vdots & \ddots & \vdots \\ \omega(x_n, x_1) & \cdots & \omega(x_n, x_n) \end{bmatrix} \\
\xi(z^d) &= \frac{1}{1 + \frac{n}{\nu}}\left[1 + \frac{1}{\nu}(z^d - \mu_d)^T\Omega_{dd}^{-1}(z^d - \mu_d)\right]
\end{aligned}$$

■

Note that this entails that all conditional TRFs are TRFs themselves. The conditional TRF also has a hierarchical representation that follows from Result 2:

Result 4 ***Hierarchical representation of conditional TRFs***
Let $\{[Z(x)|z^d]; x \in \mathcal{D}\}$ *conditional on the random parameters* (β, ϕ^2) *be:*

$$\left\{[Z(x)|z^d, \beta, \phi^2]; x \in \mathcal{D}\right\} \sim G_x\left(\mu_{x|z^d,\beta}\ ,\ \phi^2 \phi^Z_{0xx|z^d,\beta}\right)$$

where $\mu_{x|z^d,\beta}$, ϕ^2 *and* $\phi^Z_{0xx|z^d,\beta}$ *are the conditional expectation function, the variance and the conditional correlation function of a* $G_x(g_x^T\beta, \phi^2\phi^Z_{0xx})$. *Otherwise the notation is as in Result 2. Further, let*

$$\begin{aligned}
[\beta|z^d, \phi^2] &\sim N_k\left(\mu_{\beta|z^d}, \phi^2 \Phi^\beta_{0|z^d}\right)\\
[\phi^2|z^d] &\sim IG\left(\eta_{1|z^d}, \eta_{2|z^d}\right)
\end{aligned}$$

where $\mu_{\beta|z^d}$ *and* $\Phi^\beta_{0|z^d}$ *are the conditional expectation and correlation matrix of a* $N_k\left(\mu_\beta, \phi^2\Phi^\beta_0\right)$, *while* $\eta_{1|z^d}$ *and* $\eta_{2|z^d}$ *are the conditional parameters of a* $IG(\eta_1, \eta_2)$ *with* $\eta_1 = \frac{\nu}{2}$ *and* $\eta_2 = \frac{\nu\omega^2}{2}$. *Then*

$$\left\{[Z(x)|z^d]; x \in \mathcal{D}\right\} \sim T_x\left(\mu_{x|z^d}, \omega_{xx|z^d}, \nu + n\right)$$

with

$$\begin{aligned}
\mu_{x|z^d} &: \left\{[\mu(x)|z^d] = g(x)^T\mu_\beta + \omega^T_{xd}\Omega^{-1}_{dd}[z^d - G_d^T\mu_\beta]; x \in \mathcal{D}\right\}\\
\omega_{xx|z^d} &: \Big\{[\omega(x', x'')|z^d]\\
&\qquad = \xi(z^d)\left[\omega(x', x'') - \omega^T_{x'd}\Omega^{-1}_{dd}\omega_{x''d}\right]; (x', x'') \in \mathcal{D} \times \mathcal{D}\Big\}
\end{aligned}$$

where

$$\begin{aligned}
G_d &= \big(g(x_1), \ldots, g(x_n)\big)\\
\omega_{xx} &: \left\{\omega(x', x'') = \omega^2\left[\phi^Z_0(x', x'') + g(x')^T\Phi^\beta_0 g(x'')\right]; (x', x'') \in \mathcal{D} \times \mathcal{D}\right\}\\
\xi(z^d) &= \frac{1}{1 + \frac{n}{\nu}}\left[1 + \frac{1}{\nu}(z^d - G_d^T\mu_\beta)^T\Omega^{-1}_{dd}(z^d - G_d^T\mu_\beta)\right]
\end{aligned}$$

otherwise the notation is as in Result 2 and 3. The exact expressions are given in the Appendix. ∎

This hierarchical representation can of course be seen as the posterior model in a Bayesian setting. Moreover, this representation provides a suitable simulation algorithm for conditional TRFs:

Algorithm 2 ***Simulation of conditional TRFs***
Simulate a conditional TRF $\{[Z(x)|z^d]; x \in \mathcal{D}\} \sim T_x(\mu_{x|z^d}, \omega_{xx|z^d}, \nu + n)$ *with notation as in Result 4 by:*

- *given* z^d, *generate* ϕ^2 *from* $IG\left(\xi_{1|z^d}, \xi_{2|z^d}\right)$
- *given* z^d *and* ϕ^2, *generate* β *from* $N_k\left(\mu_{\beta|z^d}, \phi^2\Phi^\beta_{0|z^d}\right)$
- *given* z^d *and* (ϕ^2, β), *generate* $\{Z(x); x \in \mathcal{D}\} \sim G_x\left(\mu_{x|z^d,\beta}, \phi^2\phi^z_{0xx|z^d,\beta}\right)$

This algorithm can of course be very efficient, since β and ϕ^2 are low-dimensional, and any fast algorithm for simulation of conditional GRFs can be used.

PREDICTION IN T-DISTRIBUTED RANDOM FIELDS

A predictor for $Z(x_+)$ in an arbitrary location $x_+ \in \mathcal{D}$ based on the observations in $Z^d = [Z(x_1), \ldots, Z(x_n)]^T$ with $(x_1, \ldots, x_n) \in \mathcal{D} \times \ldots \mathcal{D}$ having realizations $z^d = (z_1, \ldots, z_n)^T$, can be obtained from the definition of TRFs in Definition 3 and Result 1.4. It is reasonable to term this predictor the T-kriging predictor.

Definition 4 ***T-kriging predictor***
A RV $Z(x_+)$ from a TRF $\{Z(x); x \in \mathcal{D}\} \sim T_x(\mu_x, \omega_{xx}, \nu)$ can be predicted from the $(n \times 1)$-dimensional vector of observations $Z^d = z^d$ by:

$$[Z(x_+)|z^d] \sim T_1 \left([\mu(x_+)|z^d], [\omega(x_+)|z^d], \nu + n \right)$$

where

$$\begin{aligned}
\left[\mu(x_+)|z^d\right] &= \mu(x_+) + \omega_{x_+d}^T \Omega_{dd}^{-1} (z^d - \mu_d) \\
\left[\omega(x_+)|z^d\right] &= \left[\omega(x_+, x_+)|z^d\right] \\
&= \xi(z^d) \left[\omega(x_+, x_+) - \omega_{x_+d}^T \Omega_{dd}^{-1} \omega_{x_+d}\right]
\end{aligned}$$

with notation as in Result 3. From Result 1.2 one has

$$\begin{aligned}
\mathrm{E}\{Z(x_+)|z^d\} &= \left[\mu(x_+)|z^d\right] & ; \quad \nu + n \geq 2 \\
\mathrm{Var}\{Z(x_+)|z^d\} &= \frac{\nu + n}{\nu + n - 2} \left[\omega(x_+)|z^d\right] & ; \quad \nu + n \geq 3
\end{aligned}$$

∎

Note that the conditional expectation predictor, which is optimal under squared error loss, when it exists, is linear in z^d, and coincides with the traditional kriging predictor. The prediction variance, however, is also dependent on z^d. This is contrary to traditional kriging where the prediction variance is dependent on the location configuration $(x_1, \ldots, x_n)$ only.

Note further that when the number of conditioning observations increases, i.e., $n \to \infty$, the T-kriging predictor tends towards the traditional kriging predictor, since the degrees of freedom in the predictor pdf tends towards infinity, which entails Gaussianity. Broadly spoken, the T-kriging predictor accounts for the fact that the variance of the RF is unknown and must be estimated. In traditional kriging the estimation uncertainty of the variance is ignored. These considerations have a parallel in the use of Student-T distributions in traditional statistical inference.

PARAMETER ESTIMATION IN T-DISTRIBUTED RANDOM FIELDS

In this section a TRF $\{Z(x); x \in \mathcal{D}\} \sim T_x\left(g_x^T \mu_\beta, \omega^2[\phi_{0xx}^Z + g_x^T \Phi_0^\beta g_x], \nu\right)$ with notation as in Result 2 will be considered. Focus is on estimation of the model parameters. Assume that several independent outcomes of this TRF exist: $\{Z_i(x); x \in \mathcal{D}\}$; $i = 1, \ldots, m$. In each of these outcomes a number of observations are made $Z_i^d = [Z_i(x_{i1}), \ldots, Z_i(x_{in_i})]^T$; $i = 1, \ldots, m$ with realizations $z_i^d = (z_{i1}, \ldots, z_{in_i})^T$. The objective is to estimate the model parameters $(\mu_\beta, \omega^2, \phi^Z_{0xx}, \Phi^\beta_0, \nu)$ from these observations. The spatial correlation function ϕ^Z_{0xx} is particularly difficult to determine, and it is assumed known in the exposition below. Procedures for estimating ϕ^Z_{0xx} may be developed based on estimation of spatial correlation functions in GRF models (Chilès and Delfiner, 1999).

It is of course possible to specify the full likelihood function based on the definition of the TRF and perform maximum likelihood estimation. It turns out that the parameter ν, which defines the tail behaviour of the marginal pdf of the TRF, is particularly hard to determine by this approach. This is unfortunate, since this tail behaviour is of primary interest in the study. An alternative estimation procedure, termed hierarchical maximum likelihood, is recommended here, since it appears to provide more stable estimates of ν in particular.

The hierarchical maximum likelihood estimator is based on the hierarchical representation of TRFs in Result 2. It draws on the property that given (β, ϕ^2) the individual outcomes of the TRF will be GRFs. Hence, based on the m outcomes, (β_i, ϕ^2_i); $i = 1, \ldots, m$ can be estimated by maximum likelihood estimators. The parameters of interest, $(\mu_\beta, \omega^2, \Phi^\beta_0, \nu)$, can thereafter be estimated by maximum likelihood estimators based on $(\beta_i|\phi^2_i)$ and $\phi^2_i; i = 1, \ldots, m$ and their pdfs as specified in Result 2.

The likelihood function for (β_i, ϕ^2_i); $i = 1, \ldots, m$ is

$$L(\beta_1, \ldots, \beta_m, \phi^2_1, \ldots, \phi^2_m) = \prod_{i=1}^m \frac{1}{(2\pi\phi^2_i)^{\frac{n_i}{2}}} \left|\Phi^Z_{0d_id_i}\right|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2\phi^2_i} M^d_i\right\}$$

with

$$M^d_i = (z^d_i - G^T_{d_i}\beta_i)^T (\Phi^Z_{0d_id_i})^{-1} (z^d_i - G^T_{d_i}\beta_i)$$

where $G_{d_i} = \big(g(x_{i1}), \ldots, g(x_{in_i})\big)$ is a known $(k \times n_i)$ matrix and $\Phi^Z_{0d_id_i}$ is a $(n_i \times n_i)$-correlation matrix defined by ϕ^Z_{0xx}. The corresponding maximum likelihood estimates are

$$\begin{aligned} \hat{\beta}_i &= [G_{d_i}(\Phi^Z_{0d_id_i})^{-1} G^T_{d_i}]^{-1} G_{d_i} (\Phi^Z_{0d_id_i})^{-1} z^d_i \; ; \; i = 1, \ldots, m \\ \hat{\phi}^2_i &= \frac{1}{n_i} \hat{M}^d_i \; ; \; i = 1, \ldots, m \end{aligned}$$

with $\hat{M}^d_i$ being M^d_i where $\hat{\beta}_i$ is substituted for β_i. The likelihood function for $(\mu_\beta, \Phi^\beta_0)$ is

$$\begin{aligned} L(\mu_\beta, \Phi^\beta_0) &= \prod_{i=1}^m \frac{1}{(2\pi\hat{\phi}^2_i)^{\frac{n_i}{2}}} |\Phi^\beta_0|^{-\frac{1}{2}} \\ &\quad \cdot \exp\left\{-\frac{1}{2\hat{\phi}^2}(\hat{\beta}_i - \mu_\beta)^T (\Phi^\beta_0)^{-1} (\hat{\beta}_i - \mu_\beta)\right\}. \end{aligned}$$

By maximizing this likelihood function, one obtains the following estimators:

$$\begin{aligned} \hat{\mu}_\beta &= \frac{1}{\sum_{i=1}^m \frac{1}{\hat{\phi}^2_i}} \cdot \sum_{i=1}^m \frac{\hat{\beta}_i}{\hat{\phi}^2_i} \\ \hat{\Phi}^\beta_0 &= \frac{1}{m} \sum_{i=1}^m \frac{1}{\hat{\phi}^2_i} (\hat{\beta}_i - \hat{\mu}_\beta)(\hat{\beta}_i - \hat{\mu}_\beta)^T. \end{aligned}$$

Finally, the likelihood function for (ω^2, ν) is

$$L(\omega^2, \nu) = \prod_{i=1}^{m} \frac{1}{\Gamma(\frac{\nu}{2})} \left(\frac{\nu\omega^2}{2}\right)^{\frac{\nu}{2}} \left(\frac{1}{\hat{\phi}_i^2}\right)^{\frac{\nu+2}{2}} \exp\left\{-\frac{\nu\omega^2}{2\hat{\phi}_i^2}\right\},$$

and maximizing provides the following expressions, defining the maximum likelihood estimators;

$$\begin{aligned} \hat{\omega}^2 &= \left[\frac{1}{m}\sum_{i=1}^{m}\frac{1}{\hat{\phi}_i^2}\right]^{-1} \\ \frac{\Gamma'(\frac{\nu}{2})}{\Gamma(\frac{\nu}{2})} - \ln\frac{\nu}{2} &= \ln\left[\prod_{i=1}^{m}\frac{1}{\hat{\phi}_i^2}\right]^{1/m} + \ln\hat{\omega}^2 \end{aligned}$$

where $\Gamma(x)$ and $\Gamma'(x)$ is the gamma function and its derivate, respectively. The last equation must be solved numerically in order to define an estimate for ν.

Properties of maximum likelihood estimators in RFs in general are discussed in Stein (1999). For this particular hierarchical maximum likelihood approach one has; $\hat{\beta}_i$ and $\hat{\phi}_i^2$ are consistent estimators in the sense that $n_i \to \infty$ by expanding $\mathcal{D} \to \mathbb{R}^n$, then $\hat{\mu}_\beta$, $\hat{\Phi}_0^\beta$, $\hat{\omega}^2$ and $\hat{\nu}$ are consistent estimators in the sense that $m \to \infty$. Note in particular that both $n_i \to \infty$, by expanding $\mathcal{D}$, and $m \to \infty$ are required to ensure consistent estimators for the model parameters of interest.

An interesting case appears if $\hat{\phi}_i^2; i = 1, \ldots, m$ all turn out identical to, for example, $\hat{\phi}^2$. Then it is easy to show that $\hat{\omega}^2 = \hat{\phi}^2$ and $\hat{\nu} = \infty$. This entails that the inferred TRF coincides with a GRF. Hence the degrees of freedom, ν, is related to the dispersion in $\hat{\phi}_i^2; i = 1, \ldots, m$.

Another interesting case appears when only one realization is available, i.e., $m = 1$. Hence $z_1^d = (z_{11}, \ldots, z_{1n_1})^T$ is the only available data vector. With $n_1 \geq 2$ estimates of β_1 and ϕ_1^2 can be obtained, and the estimators are consistent in the sense specified above. The estimate of μ_β is obtainable while Φ_0^β is left unspecified. Assume for the moment that β is a constant vector μ_β which makes Φ_0^β irrelevant. The interesting feature is that ω^2 and ν are estimated to $\hat{\phi}_1^2$ and ∞. This entails that the inferred TRF collapses into a GRF, since they coincide for $\nu = \infty$. Consequently, one cannot make inference about TRFs with observations from only one realization of the RF. This is related to the fact that the TRF model does not exhibit diminishing spatial dependence, and the heavy-tailedness only appears cross-realizations and not in-realizations. Recall that all realizations appear GRF-like with parameter values varying between realizations.

CASE STUDY

Consider the density log from the well in the Gullfaks field (Fig. 1, left display) . Assume that the log has only been sampled every third meter, while the layer geometry has been determined from elsewhere (Fig. 10). The layers are classified with respect to the average value in each layer, which results in layer 1, 2, 3, 5, 7 and 9 being pooled, as in the "Motivation" section. The variance estimates, however, are highly unreliable when based on so few samples in each layer, and are therefore not used for classification. The composite histogram of all samples in the pooled layers is displayed in the left display in Figure 11.

Two alternative modelling approaches can be seen. In the traditional GRF modelling approach, the layers are considered to be realizations from one common GRF model. The variances in all layers are considered to be identical, and the differences in empirical variances are tributed to the sampling variance. The compiled set of observations can of course be used for inference about the parameters of the GRF model. Alternatively, a TRF modelling approach can be used, where the layeres are considered to be realizations from a unifying TRF model. This approach gives room for different variances in the individual layers. Also in this case, the compiled set of observations can be used for model parameter inference. Consequently, the difference in the two approaches is the interpretation of the varying empirical variances in the layers. Note further that the TRF model contains the GRF model as a limiting case, and hence the former can be seen as a generalization of the latter.

In this case study both the GRF and the TRF approach will be used, and the results based on each of them compared. The two models are $G_x(\mu, \phi^2 \phi_{0xx})$ and $T_x(\mu, \omega^2 \omega_{0xx}, \nu)$ respectively, with μ a constant level. The spatial correlation function ϕ_{0xx} and the spatial dependence function ω_{0xx} will, with constant μ, have similar interpretation, see Result 2, and they are assumed known to be

$$\phi_{0xx} = \omega_{0xx} = \varrho(x', x'') = \exp\left\{-\alpha(x'' - x')\right\}$$

with $\alpha = 1/0.65$, which is in good agreement with the data. The maximum likelihood estimators based on the GRF model provides $\hat{\mu} = 2018$ and $\hat{\phi}^2 = 6047$. Alternatively, the TRF model provides hierarchical maximum likelihood estimates $\hat{\mu} = 2018$, $\hat{\omega}^2 = 1383$ and $\hat{\nu} = 1.71$. The kernel smoothed histograms of the compiled observations and the inferred marginal pdfs from the two models are presented in the right display of Figure 11. The TRF model seems to reproduce the observations better than the GRF model.

Layers 2 and 3 are studied in more detail under the two competing models, since they constitute the extremes with respect to variability. Each of the layers are considered as independent realizations of the respective models, and they are evaluated under the two models inferred above. In Figure 12 conditional realizations in the two layers under the inferred GRF and TRF model respectively, are presented. Under the GRF model the inferred variance is considered to be the true one, and hence not adapted to the observations in the layer under study. In layer 2 this entails that the variablity in the conditioning observations completely dominates the variability of the model. The opposite effect is observed in layer 3, since the variability in the observations is less. Under the TRF model simulation is made according to Algorithm 2. The variability of the model is adaptive to the value of the available observations and the realizations look much more trustworthy. To summarize, the adaptivity of the variance in the conditional TRF model makes the realizations more realistic than under the GRF model with non-adaptive variance.

In Figure 10 two non-observed locations are marked, one in each layer. In Figure 13, the predictive conditional pdfs in these two locations under the inferred GRF and TRF model respectively, are displayed. Under the GRF model the simple kriging predictor with inferred model parameters is used. The conditional pdf is Gaussian with expectation dependent on the value of the observations and variance dependent on the location of the observations only. Hence the prediciton variances in the two layers appear very similar. Under the TRF model the prediction is made according to the results in the section on prediction. The conditional pdf is T-distributed with expectations coinciding with the expectations under the GRF model. The prediction variances, however, are adaptive to the actual values of the observations in the layer under study. Hence the prediction variance in layer 2 is much larger than in layer 3. Moreover, the degrees of freedom in the conditional T-distribution increases with the number of observations, since the estimation of the variance is accounted for. Hence the conditional pdf in layer 3 is more Gaussian-like than the one in layer 2. To summarize, the adaptive prediction variance in the TRF model makes the predictive pdfs more realistic than under the GRF model, where the variability is averaged over the layers. Moreover, the prediction is T-distributed under the TRF model, and accounts for the fact that the variance is unknown and estimated. Under the GRF model, the estimated variance is considered to be the true one, and the esimation effect is ignored.

One may argue that the empirical variance obtained from the observations in layers 2 and 3 are so different that the layers should not be pooled when making inference of the GRF and TRF models. It may be so, but in order to evaluate some of the really thin layers, some pooling must be made in order to infer a RF model. Then the effect demonstrated above will appear, although not as dramatic as seen here. If each layer is considered individually, and inference of a GRF model is done, the results will be overly optimistic because the uncertainty in estimating the variance is not accounted for.

CHARACTERISTICS OF T-DISTRIBUTED RANDOM FIELDS

The example in the "Motivation" section demonstrates that a RF with heavy-tailed marginal pdfs may result from a composite of RFs with light-tailed marginal pdfs with similar expectations but varying variances. The individual RFs may be GRFs with identical expectations, but varying variances, and this may be modelled by a unifying TRF.

The TRF model has most of the favourable properties for parametric models listed in the introduction. The model is fully specified; it exhibits permutation invariance and probabilistic consistency; it has marginal invariance; it is closed under additivity, and it is analytically tractable. The TRF model is a fairly general RF model, since it can reproduce a large class of RFs with symmetric, unimodal marginal pdfs, heavy-tailed or not. Moreover, the GRF model is a limiting case of the TRF model, while the Cauchy RF model appears as a special case. The generality of the TRF model has clear limitations, however, since RFs with neither skewed nor multimodal marginal pdfs can be reproduced. Model parameter inference is somewhat complicated, since consistent estimators can only be defined from sets of observations from several realizations of the TRF. If observations from only one single realization are available, the inferred TRF model collapses into a GRF model. The TRF model does not exhibit diminishing spatial dependence, and this is the reason for lack of consistency of the estimators based on observations from one realization only.

The TRF model compares favourably with the GRF model. The fact that the GRF is a special, limiting case of the TRF model emphasizes this. The TRF model can be seen as a composite of GRF models with identical expectations but varying variances. The heavy tails of the marginal pdfs of TRF models are caused by cross-realization variability and not in-realization variability, since all realizations are GRF-like with varying variance parameters from one realization to the other. For GRF models the variance parameter is a fixed constant. If sets of observations from several realizations are available, assuming the more general TRF model will increase flexibility. Inference under the TRF and GRF models coincide if observations from only one realization is available. Conditional TRF models will have both expectations and variances which depend on the values of the observations, while conditional GRF models will have only expectations depending on these values; the corresponding variance is only dependent on the location of these observations. The TRF model accounts for the fact that the model variance is largely unknown and must be estimated, while the GRF model considers the estimated variance to be the true value. Consequently, prediction variances will generally be larger under the TRF model than under the GRF model. To summarize, by consequently using a TRF model one will obtain the corresponding GRF model when it is appropriate, and otherwise benefit from the additional flexibility of the TRF model. Larger mathematical complexity constitutes the down-side of this approach. Moreover, by assuming a GRF model, extensions are sometimes easier to make, for example when adding measurement error to the observations.

The TRF model compares favourably with ϕ^{-1}-GRF models for RFs (Chilès and Delfiner, 1999) with symmetric, unimodal marginal pdfs. Note, however, that ϕ^{-1}-GRFs can be used for a larger class of RFs, although the reliability of the results is hard to judge. The ϕ^{-1}-GRF model is generally not closed under additivity, and it lacks analytical tractability, hence most results must be obtained through simulation or approximations. Lastly, strong assumptions about spatial stationarity of the RF are required in order to make inference about the ϕ-transformation. Consequently, the TRF model is considered favourable to the ϕ^{-1}-GRF model when they both apply.

The TRF model compares favourably with LSRF models as defined in Gunning (2002). Both models apply for RFs with symmetric, unimodal marginal pdfs. The LSRF models lack analytical tractability and most results must be obtained through simulation. The fact that the LSRF model is only defined through its characteristic function, with no closed form expression for the corresponding pdf, demonstrates this lack of analytical tractability. Both the TRF and the LSRF models lack diminishing spatial dependence. This seems to have less consequences for the former than the latter, since the analytical tractability of the TRF model makes it possible to define efficient simulation algorithms even for conditional TRF models.

In practice, the most favourable case for TRF models is in evaluation of RFs where multiple, sparsely sampled realizations are available. Inference of the TRF model is made on the compiled set of observations, and evaluation is done on each realization individually, based on the inferred TRF model. In reservoir evalutaion, several similar geological layers may be penetrated by a well and a limited number of core samples may be collected in each layer. Note further that the TRF model is defined for a reference space of arbitrary dimensions, hence extending the model into three dimensions representing the entire geological layer is simple.

CONCLUSIONS

Histograms of observations from spatial phenomena are often found to be more heavy-tailed than Gaussian distributions. Well log data in petroleum applications constitute one example. The heavy-tailedness of well log histograms may be caused by the well penetrating several layers with similar well log averages, but varying variances. A parametric T-distributed random field (RF) model able to capture this heavy-tailedness in the marginal pdf is defined. The model appears as a generalization of the familiar Student-T distribution, and it may be given a Bayesian interpretation. The large variability appears as cross-realization variability, contrary to in-realization variability, since all realizations are Gaussian-like with varying variances between realizations. This T-distributed RF model exhibits many favourable features: It is a fully defined, valid random field model; it defines a class of RF models closed for linear operators; it is analytically tractable; for RFs with symmetric, unimodal marginal pdfs it is fairly general, and contains the Gaussian RF model as a limiting case; reliable parameter estimators are defined based on observations from multiple realizations.

An alternative RF model with heavy-tailed marginal pdfs could be imagined. A Gaussian RF model with the variance being a RF would also exhibit heavy-tailed marginal pdfs. Moreover, if the variance RF is ergodic, the resulting heavy-tailed RF is expected to appear with diminishing spatial dependence. However, we have so far not been able to define a parametric RF model of this type which is analytically tractable.

The T-distributed RF model compares favourably with the Gaussian, ϕ-transformed and Lévy-Stable RF models. The model appears as a generalization of Gaussian RF models, and captures heavy-tailed marginal pdfs. It is analytically tractable, which is not the case for neither the ϕ-transform nor the Lévy-Stable RF models. Hence, for RFs with symmetric, unimodal marginal pdfs, the T-distributed RF model provides a recommendable alternative.

REFERENCES

Chilès, J-P. and Delfiner, P., 1999, Geostatistics: Modelling Spatial Uncertainty: Wiley, New York, 695 p.

Cornish, E.A., 1954, The multivariate t-distribution associated with a set of normal sample eluminates: Australian Journal of Physics, v.7, p.531-542.

Gunning, J., 2002, On the Use of Multivariate Lévy-Stable Random Field Models for Geological Heterogeneity: Mathematical Geology, v.34, no.1, p.34-62.

Handcock, M.S. and Stein M.L., 1993, A Bayesian Analysis of Kriging: Technometrics, v.35, no.4, p.403-410.

Hjort, N.L. and Omre, H., 1994, Topics in spatial statistics, Scandinavian Journal of Statistics, v.21, no.4, p.289-358.

Johnson, N.L. and Kotz, S., 1972, Distributions in Statistics: Continuous Multivariate Distributions: Wiley, New York, 333 p .

Kitanidis, P.K., 1986, Parameter uncertainty in estimation of spatial functions: Bayesian analysis: Water Resources Research, v.22, p.499-507.

Le, N.D. and Zidek, J.V., 1991, Interpolation With Uncertain Spatial Covariances: A Bayesian Alternative To Kriging: Journal of Multivariate Analysis, v.43, p.351-374.

Mardia, K.V., Kent, J.T. and Bibby, J.M., 1979, Multivariate Analysis: Academic Press, London, 518 p.

Omre, H., 1987, Bayesian Kriging — Merging Observations and Qualified Guesses in Kriging: Mathematical Geology, v.19, no.1, p.25-39.

Stein, M.L., 1999, Interpolation of Spatial Data: Some Theory for Kriging: Springer-Verlag, New York, 247 p.

Walpole, R.E and Myers, R.H., 1993, Probability and Statistics for Engineers and Scientists: Prentice-Hall, Upper Saddle River, NJ, 739 p.

Welsh, A.H., 1996, Aspects of Statistical Inference: Wiley, New York, 464 p.

Yaglom, A.M., 1962, An Introduction to the Theory of Stationary Random Functions: Dover Publications, New York, 256 p.

APPENDIX: PARAMETERS IN HIERARCHICAL CONDITIONAL TRFS

The exact relations defined in Result 4 are defined here:

$$
\begin{aligned}
\mu_{x|z^d,\beta} &: \left\{[\mu(x)|z^d,\beta] = g(x)^T\beta + (\phi^Z_{0xd})^T(\Phi^Z_{0dd})^{-1}(z^d - G_d^T\beta); x \in \mathcal{D}\right\} \\
\phi^Z_{0xx|z^d,\beta} &: \left\{[\phi^Z_0(x',x'')|z^d,\beta] = \phi^Z_0(x',x'') - (\phi^Z_{0x'd})^T(\Phi^Z_{0dd})^{-1}\phi^Z_{0x''d}; x',x'' \in \mathcal{D}\times\mathcal{D}\right\}
\end{aligned}
$$

with

$$
\begin{aligned}
\phi^Z_{0xd} &= \big(\phi^Z_0(x,x_1),\ldots,\phi^Z_0(x,x_n)\big)^T \\
\Phi^Z_{0dd} &= \begin{bmatrix} \phi^Z_0(x_1,x_1) & \cdots & \phi^Z_0(x_1,x_n) \\ \vdots & \ddots & \vdots \\ \phi^Z_0(x_n,x_1) & \cdots & \phi^Z_0(x_n,x_n) \end{bmatrix} \\
G_d &= \big(g(x_1),\ldots g(x_n)\big) \\
\mu_{\beta|z^d} &= \mu_\beta + \Phi^\beta_0 G_d\left[\Phi^Z_{0dd} + G_d^T\Phi^\beta_0 G_d\right]^{-1}(z^d - G^T\mu_\beta) \\
\Phi^\beta_{0|z^d} &= \Phi^\beta_0 - \Phi^\beta_0 G_d\left[\Phi^Z_{0dd} + G_d^T\Phi^\beta_0 G_d\right]^{-1} G_d^T\Phi^\beta_0 \\
\eta_{1|z^d} &= \eta_1 + \frac{n}{2} \\
\eta_{2|z^d} &= \xi(z^d)\left[\eta_2 + \frac{\omega^2 n}{2}\right]
\end{aligned}
$$

List of Figures

Table 1: Averages and empirical variances in the layers of Figure 1.

layer	average	empirical variance
1	1962	351
2	1986	56402
3	2038	744
4	2331	54659
5	2014	147
6	2361	67417
7	1997	448
8	2554	41636
9	2053	5969

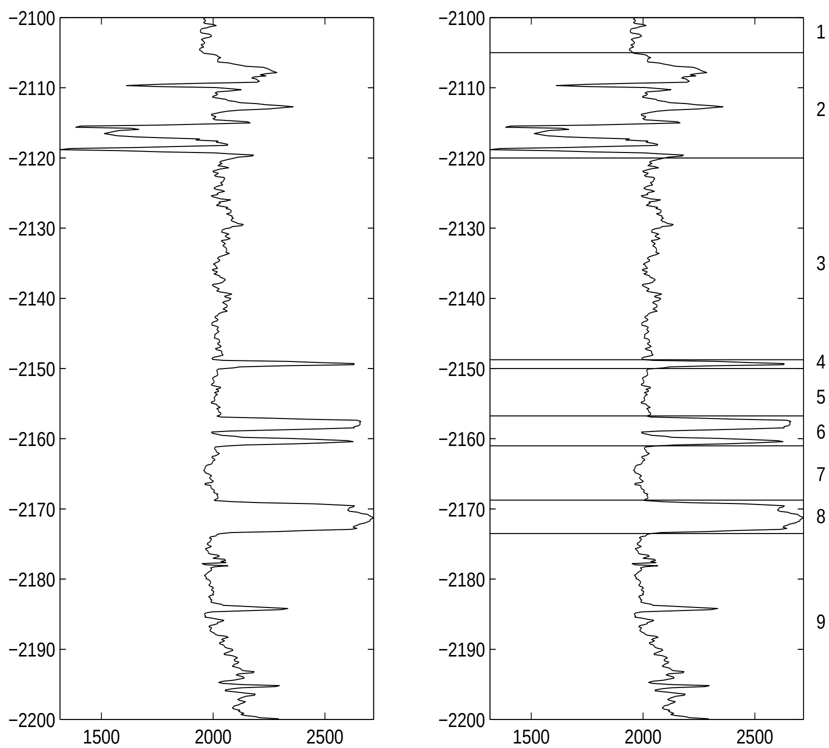

Figure 1: *Left: Bulk density log from well C33 of the Gullfaks field in the North sea. Right: Layer interpretation of the log.*

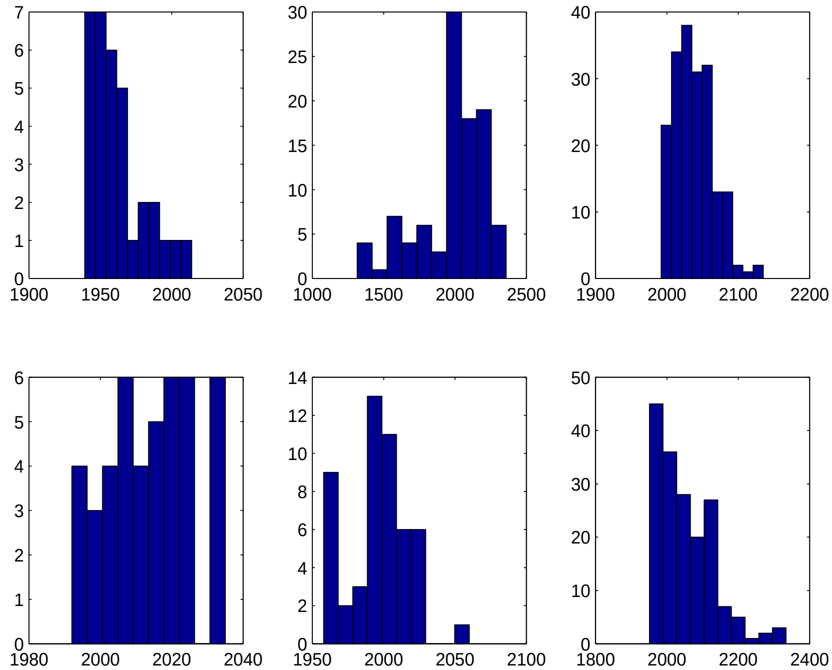

Figure 2: *Histogram of density observations in each of the pooled layers. Upper: Layer 1,2,3. Lower: Layer 5,7,9.*

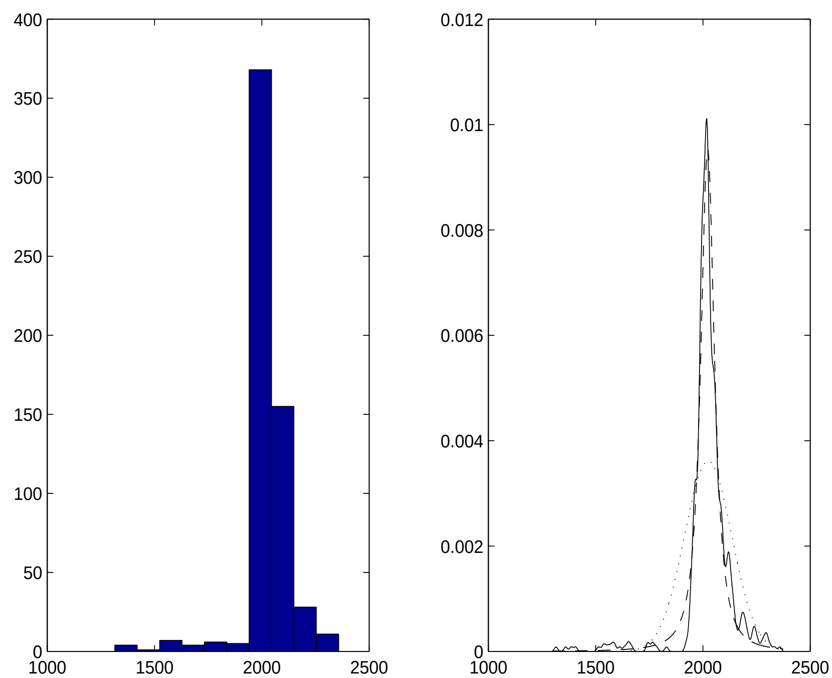

Figure 3: *Left: Histogram of density observations in pooled layers. Right: Kernel estimate of pdf (—); Gaussian estimate of pdf (⋯); and Student-T estimate of pdf (− − −).*

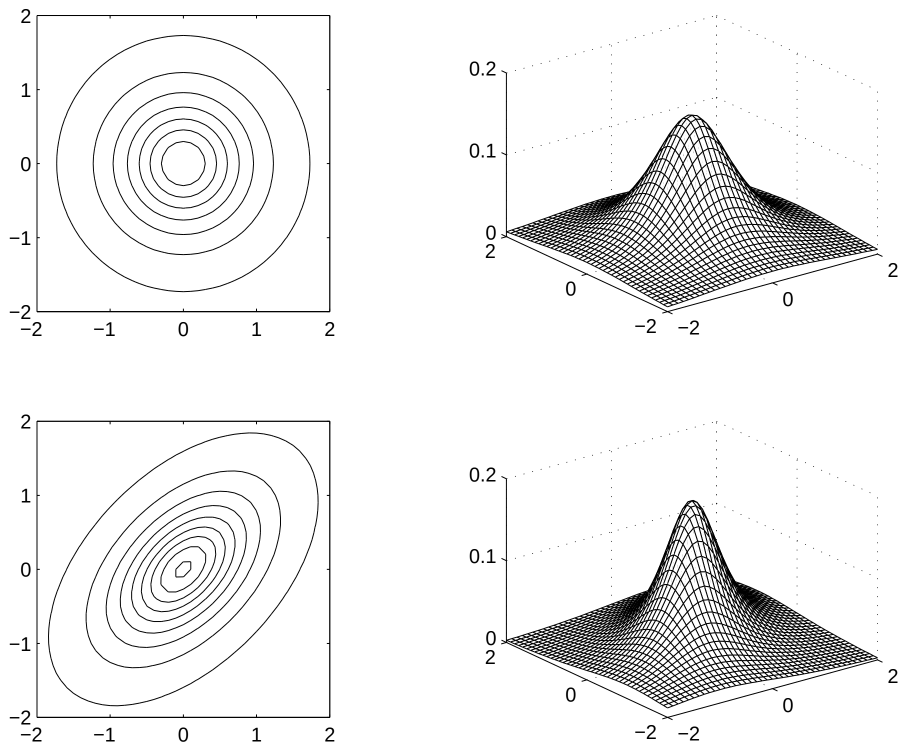

Figure 4: *Upper:* $T_2(0, \Omega, 1)$*-distribution with unit variance and no correlation. Lower:* $T_2(0, \Omega, 1)$*-distribution with unit variance and correlation 0.5.*

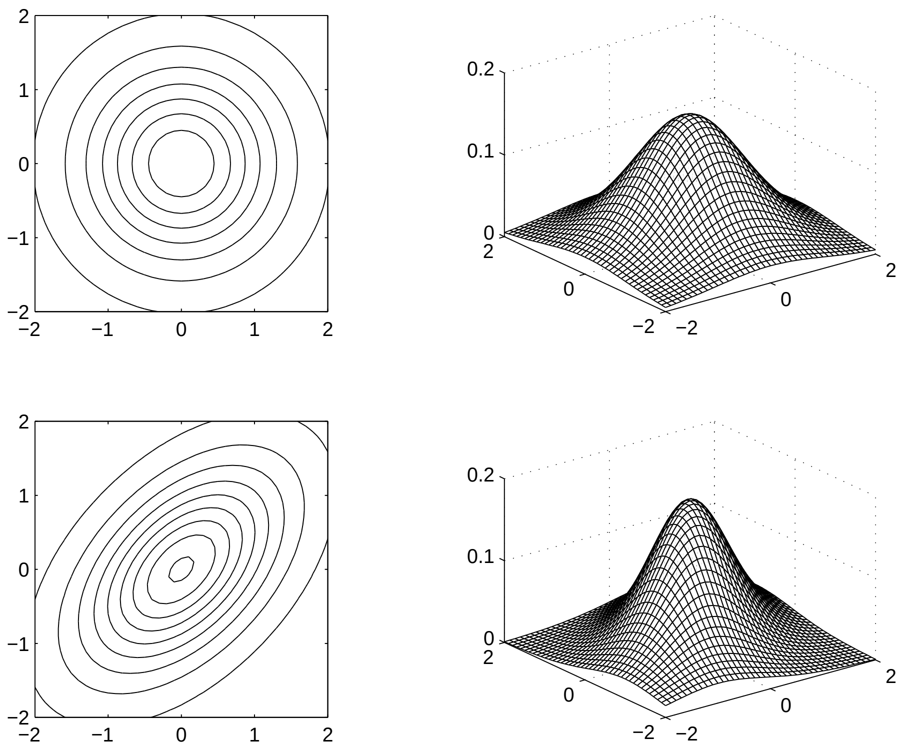

Figure 5: *Upper:* $T_2(0, \Omega, 7)$*-distribution with unit variance and no correlation. Lower:* $T_2(0, \Omega, 7)$*-distribution with unit variance and correlation 0.5.*

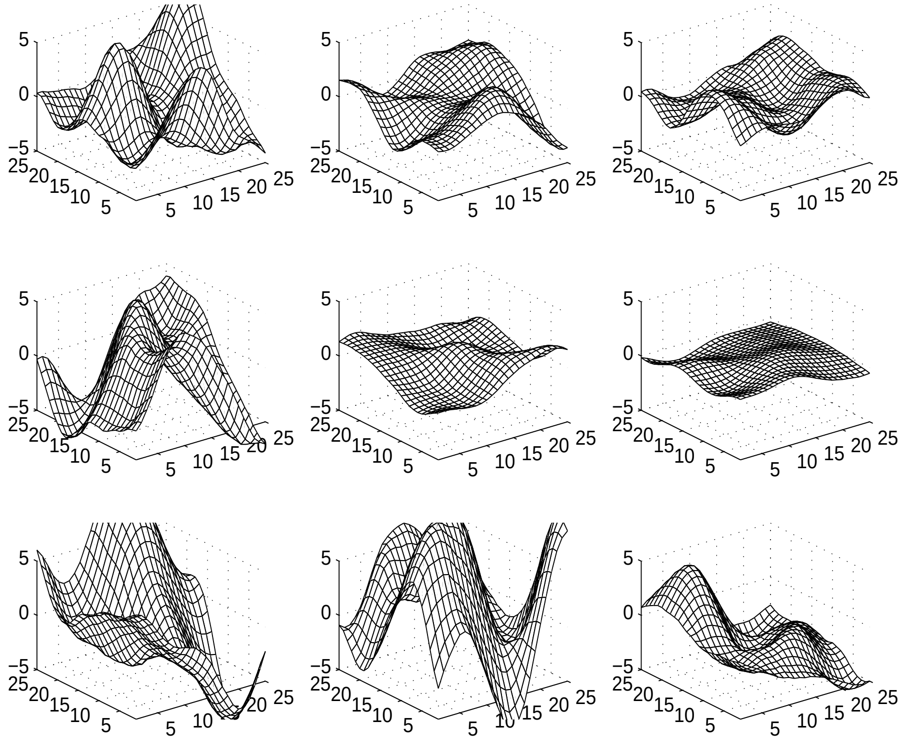

Figure 6: *Realizations of T-distributed RFs with $\nu = 1$.*

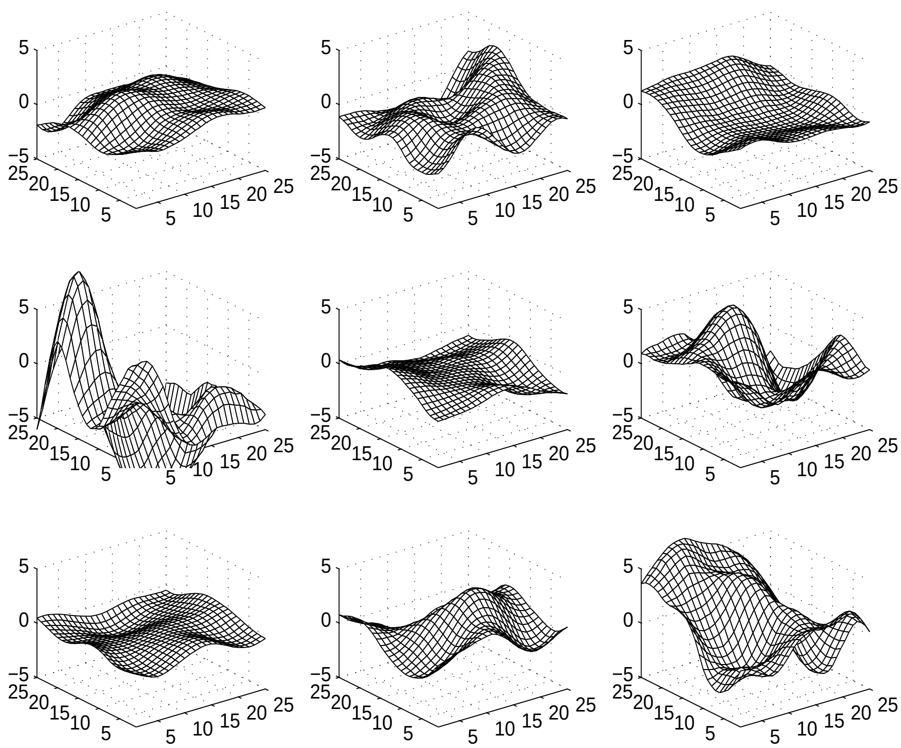

Figure 7: *Realizations of T-distributed RFs with $\nu = 3$.*

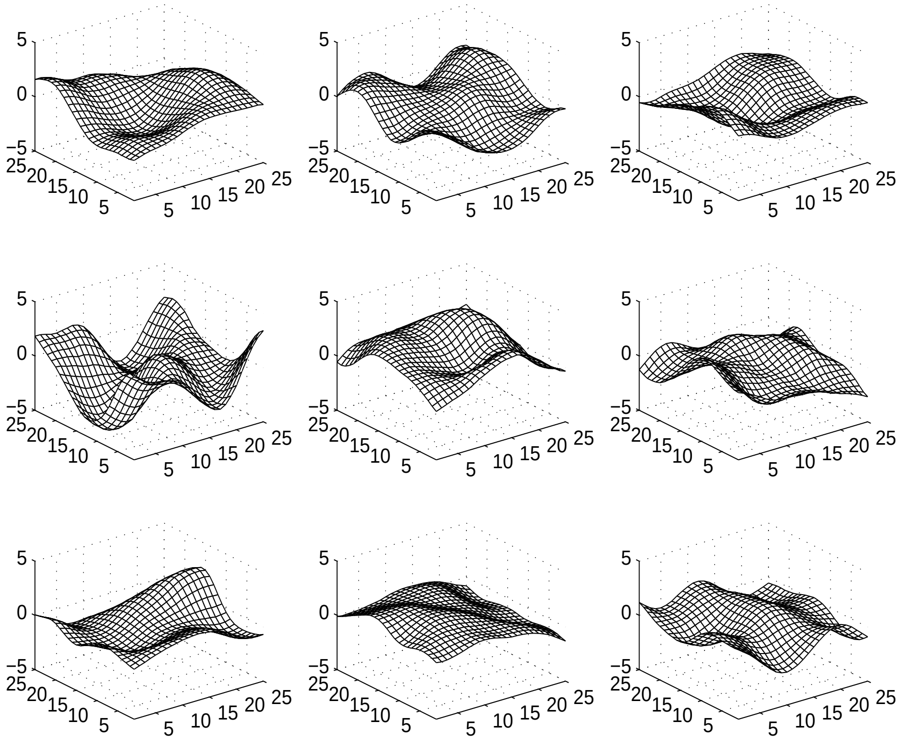

Figure 8: *Realizations of T-distributed RFs with $\nu = 5$.*

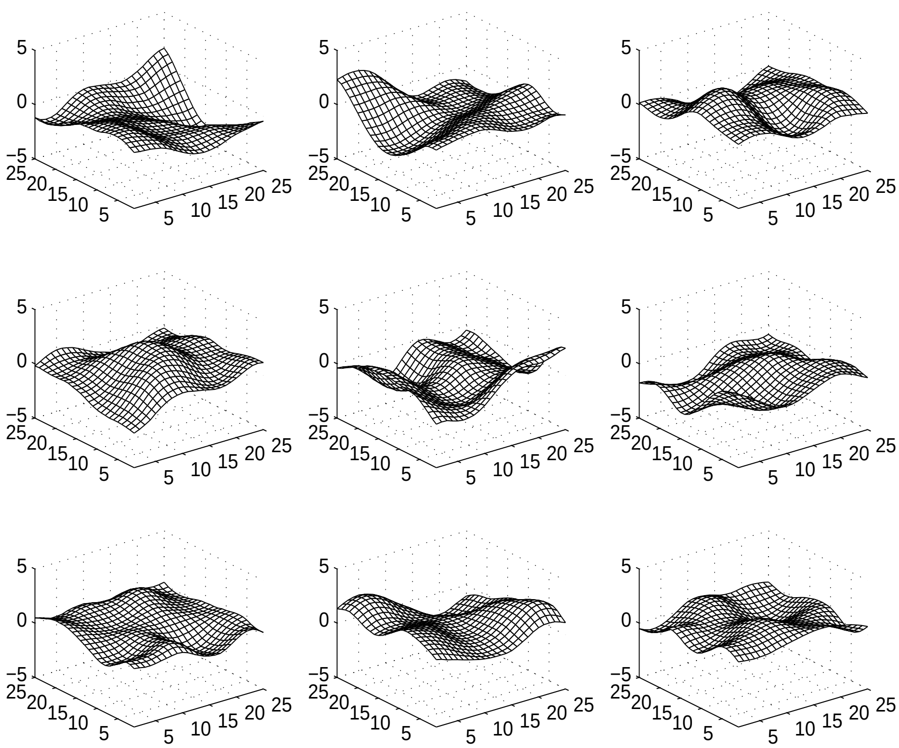

Figure 9: *Realizations of T-distributed RFs with $\nu = \infty$, i.e. Gaussian RFs.*

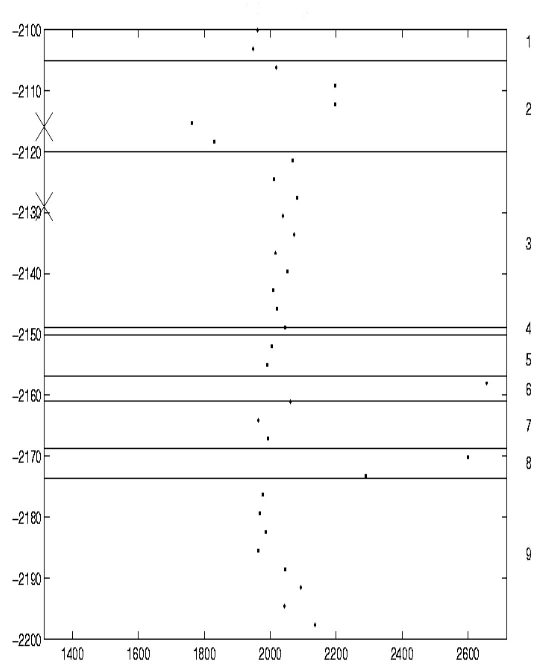

Figure 10: *Bulk density log observations from well C33 from the Gullfaks field in the North sea with layer interpretation. The two locations to be predicted are marked (×).*

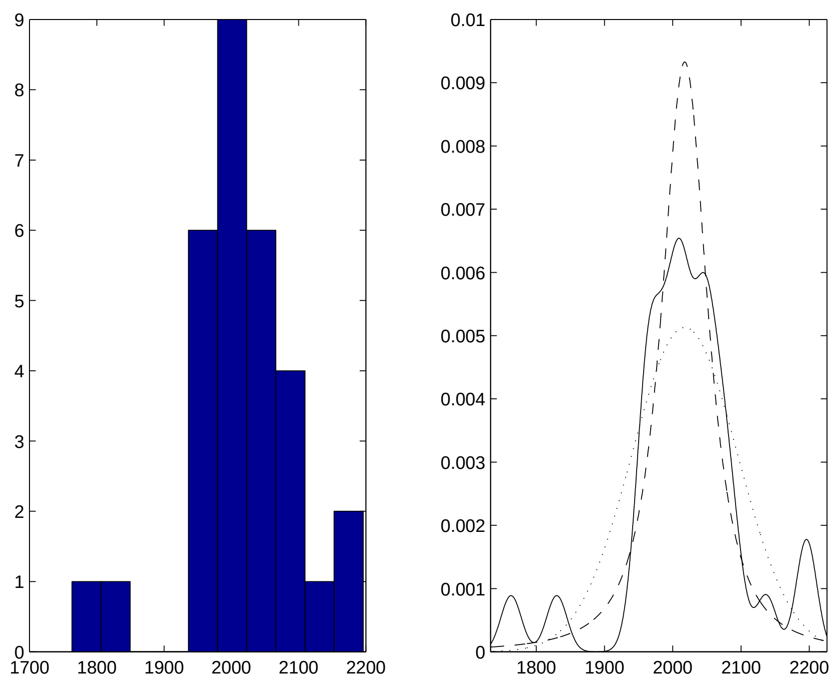

Figure 11: *Left: Histogram of density observations in pooled layers. Right: Kernel estimate of pdf (—); Gaussian estimate of pdf (· · ·); and Student-T estimate of pdf (− − −).*

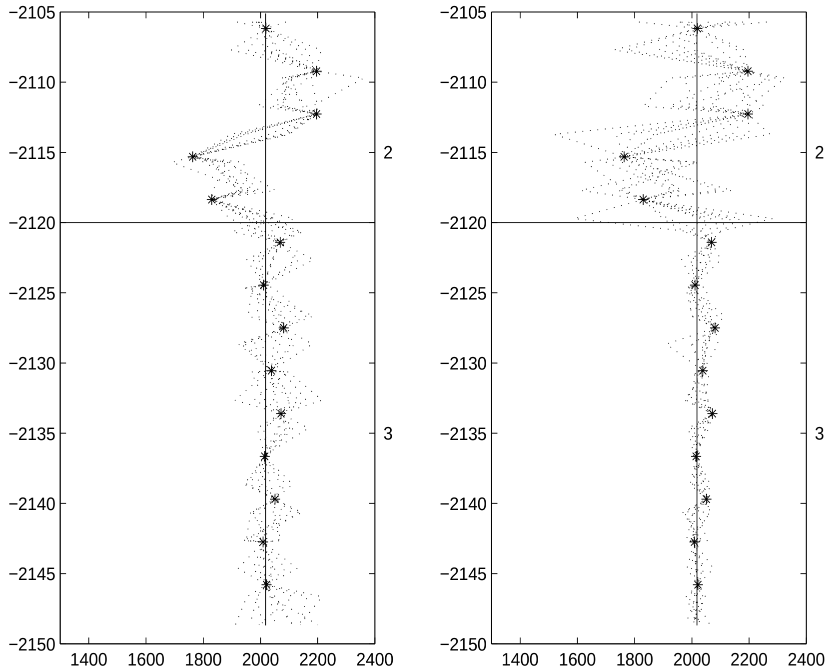

Figure 12: *Left: Ten conditional realizations from the inferred Gaussian RF model. Right: Ten conditional realizations from the inferred T-distributed RF model.*

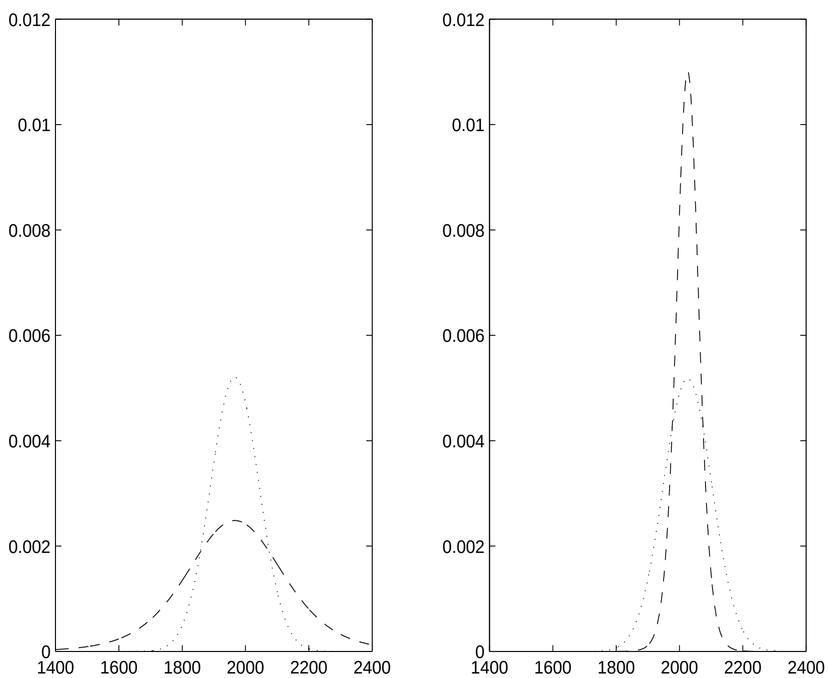

Figure 13: *Prediction pdf based on the inferred Gaussian RF model ($\cdots$) and the inferred T-distributed RF model ($- - -$) in layer 2 (left) and layer 3 (right).*