Bayesian image analysis with coloured Voronoi tessellations and a view to applications in reservoir modelling

Jesper Møller[1] **and Øivind Skare**[2]

[1] Department of Mathematical Sciences, Aalborg University, Fredrik Bajers Vej 7G, DK-9220 Aalborg Ø, Denmark. E-mail: jm@math.auc.dk
[2] Norwegian Computing Center, P.O. Box 114 Blindern, N-0314 Oslo, Norway. E-mail: skare@nr.no

ABSTRACT: A new flexible prior for Bayesian image analysis and reservoir modelling is defined in terms of interacting coloured Voronoi cells described by a certain nearest-neighbour Markov point process. This prior can be defined in both 2 and 3 (as well as higher) dimensions, and simple MCMC algorithms can be used for drawing inference from the posterior distribution. Various 2D and 3D applications are considered.

Keywords: Markov chain Monte Carlo; Markov connected component fields; Markov random fields; Nearest-neighbour Markov point processes; Polygonal partitions; Three dimensional image analysis; Triangulation models.

1. Introduction

The purpose of this paper is to construct simple and flexible 2D and 3D priors in Bayesian image analysis and reservoir characterisation, where data are typically very sparse or obtained by indirect observations, and where the posterior is explored by means of Markov chain Monte Carlo (MCMC) methods. Careful specification of such priors is particularly important when Bayesian methods are used in oil reservoir applications, since fluid flow is described by large-scale properties, while reservoir-specific information about geology, e.g. from wells and seismic data, is limited (Damsleth, Tjølsen, Omre & Haldorsen 1990, Hjort & Omre 1994). General geological knowledge about the reservoir is usually available and may assist us in choosing realistic priors.

Various priors have been used in Bayesian image analysis and reservoir modelling. Pixel-based models such as Markov random fields often lead to slow MCMC algorithms due to the high dimensionality of the problem, particularly in 3D applications, while flexible tessellation models such as the 2D triangulation models in Nicholls (1998) are difficult to extend to the 3D case. Such and other priors are reviewed in Section 2. We consider next an alternative prior model based on a marked point process for a coloured Voronoi tessellation, which we model in Section 3 as a nearest-neighbour Markov point process (Baddeley & Møller 1989). This prior model can be easily defined in both 2 and 3 (as well as higher) dimensions. Compared to the pixel-based models and the triangulation models, our use of marked point processes for coloured Voronoi tessellations provides a parsimonious parameterisation, which is convenient when dealing with various posterior distributions based on sparsely observed data. Some other advantages compared to Nicholls triangulation models are that Voronoi cells are in one-to-one correspondence with their generating points (called nuclei in this paper), and we do not need to include border points, as we can truncate Voronoi cells that are partly outside the image area or reservoir volume. In particular the MCMC algorithm described in Section 4 become much simpler than those considered in Nicholls (1998). However, the advantages of our model are most clear in 3 dimensions, as Nicholls triangulation models are much more flexible in 2 dimensions.

Convergence properties of our MCMC algorithms are discussed in Section 4, while Section 5 concerns parameter specification for the prior and likelihood terms in the posterior distribution. Empirical results are discussed in Section 6 for a 2D image analysis experiment and for real well observations in a 3D north sea oil reservoir. Section 7 contains some concluding remarks.

2. Background

One approach to Bayesian image analysis and reservoir modelling is to use marked point processes for objects against a background (Baddeley & Van Lieshout 1993). This approach has e.g. been used when the objects are sand bodies in fluvial reservoirs (Holden, Hauge, Skare & Skorstad 1998) or fault patterns (Munthe, Holden, Mostad & Townsend 1994). A general marked point process model for reservoir modelling is given in Lia, Tjelmeland & Kjellesvik (1997). However, detailed prior information is requested if marked point processes for (possibly deformed) objects against a background are used in sparse data situations, see e.g. Holden et al. (1998). The objects considered in Lia et al. (1997) and Munthe et al. (1994) are modelled as simplified geometric objects. Rue & Hurn (1999) combine the marked point process approach with that of deformable templates (Grenander & Miller 1994) for purposes in Bayesian object recognition, and where objects e.g. correspond to different types of biological cells.

Another category of models is pixel-based random fields, particularly Markov random field (MRF) models which have been extensively used as priors in Bayesian image analysis (Geman & Geman 1984, Besag 1986). They are particularly useful when we model mosaic patterns where none of the components represent a background; or, more generally, when modelling complex and irregular geometries which are difficult to represent by the marked point process approach. Examples in reservoir modelling include reservoirs formed by a mixture of sedimentary processes and reservoirs with a high packing of sedimentary facies (rock types). Often simple pairwise interaction MRF priors with small neighbourhoods are used, but by allowing higher order interactions and larger neighbourhoods more realistic priors can be constructed, see e.g. Tjelmeland & Besag (1998). An alternative to MRF models is the Markov connected component field models of Møller & Waagepetersen (1998), where potentials are associated to each of the (maximal) connected components of pixels with the same 'colour' in the image (or reservoir) so that global information about the components can be more easily incorporated than for MRFs with bounded neighbourhoods.

A problem with these pixel-based models is the high dimensionality of the state space which can cause slowly converging simulation algorithms. This is particularly true in 3D applications.

An alternative to pixel-based random field models is a polygonal partitioning (tessellation) of the space (Arak, Clifford & Surgailis 1993, Nicholls 1997, Nicholls 1998). Nicholls promotes the tessellation models as an approach for intermediate pattern analysis, where the polygons serve as fundamental building blocks of structures in the image. An advantage of this approach compared to that of pixel-based models is that we can obtain the same resolution with a lower dimensionality of the state space, since in homogeneous areas of the image or reservoir a few large polygons could suffice. For instance, Voronoi cells (as defined in the next section) can more easily than static grid cells follow the contours of objects.

An appealing example of a 2D coloured triangulation model is found in Nicholls (1998). First a natural base measure is specified: in the interior and on the border of a simple polygonal set $S \subset \mathbf{R}^2$, independent point configurations x_I and x_B are drawn from homogeneous Poisson point processes, and a finite colouring c is then specified for any possible triangulation τ of S with vertex set $x_I \cup x_B \cup h$, where h is the set of boundary vertices of the polygon S. Secondly a general purpose prior is defined w.r.t. the base measure as an exponential family type model with a two-dimensional canonical sufficient statistic specified by the number of triangles and the total length of edges separating regions of different colours. It seems difficult to extend such models to 3D, for example by using three kinds of Poisson point processes defined in the interior, on the bounding surfaces and on the edges of the set S (which is now assumed to be a 3D polyhedron). Moreover, extending the updates for the MCMC algorithm in Nicholls (1998) would be rather intricate, and it is not obvious how irreducibility will be ensured.

3. Coloured Voronoi tessellation models

The idea of using Voronoi tessellations in image analysis is not new, see for example Ahuja & Schachter (1983) and Green (1995); for theoretical properties, applications and further references on Voronoi tessellations, see Okabe, Boots & Sugihara (1992) and Møller (1994). Below we follow Baddeley & Møller (1989) in defining a particular prior for coloured Voronoi tessellations as a nearest-neighbour Markov point process with interactions between neighbouring Voronoi cells of different colours. In contrast to the marked point process approach for objects against a background, we will demonstrate that neither a particular background colour nor a detailed prior information is requested. Instead of a Voronoi tessellation, one could use a Delaunay tessellation, which also can be defined in any dimension $d \geq 1$. This is the dual of a Voronoi tessellation with cells obtained by considering the convex hulls for neighbouring Voronoi nuclei; under regularity conditions, it constitutes a triangulation (see Møller 1994). However, using a Voronoi tessellation seems

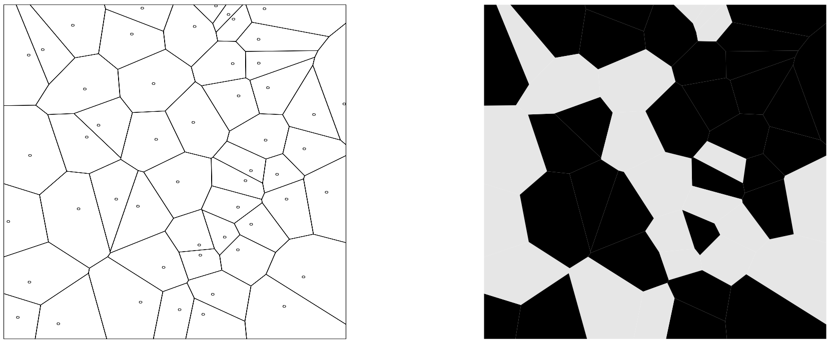

FIGURE 1. Left: A realisation of a Poisson-Voronoi tessellation of rate 50 on the unit square. Right: The same tessellation with a binary uniform colouring of the cells.

more attractive for at least two reasons: it is convenient that Voronoi cells are in one-to-one correspondence with their generating points, as a colouring of the Voronoi tessellation can be specified as a marked point process (see below); we do not need to include border points.

Let $S \subset \mathbf{R}^d$ be a bounded Borel set of volume $0 < |S| < \infty$. In our applications, S is a d-dimensional box corresponding to the image area or the reservoir volume, and $d \in \{2, 3\}$. For finite point configurations $x = \{x_1, \ldots, x_n\} \subset S$, we associate the Voronoi tessellation with cells

$$C(x_i|x) = \{s \in S : \|x_i - s\| \leq \|x_j - s\| \text{ for all } j \neq i\}, \quad i = 1, \ldots, n,$$

where $\|\cdot\|$ denotes Euclidean distance. Thus $C(x_i|x)$ is the set of points in S that has x_i as nearest 'nucleus' in x. For some applications it may be relevant to replace the Euclidean distance $\|\cdot\|$ by another metric, see Scheike (2000) and Section 6.2.

To each nucleus x_i we attach a mark $m_i \in M$, where $M = \{0, 1, \ldots, K-1\}$ is a finite set. We interpret m_i as the colour of Voronoi cell $C(x_i|x)$, and we let $y = \{(x_1, m_1), \ldots, (x_n, m_n)\}$ be the corresponding marked point pattern. Further, for $\beta > 0$, let μ_β denote the Poisson process on $S \times M$ under which the nuclei follow a homogeneous Poisson point process of rate β on S, and conditional on the nuclei, the colours of Voronoi cells are independent and uniformly distributed on M. Figure 1 shows a typical realisation under μ_β when $K = 2$, $S = [0, 1]^2$ is the unit square, and $\beta = 50$ is the mean number of nuclei.

Our prior has density proportional to $\exp(-\theta\, s(y))$ with respect to μ_β, where θ is a real parameter and

$$s(y) = \sum_{i<j:\, x_i \underset{x}{\sim} x_j} 1[m_i \neq m_j] \tag{1}$$

is the number of pairs of neighbouring Voronoi cells of different colours. Here $\underset{x}{\sim}$ denotes the neighbour relation on the points in x as defined by

$$x_i \underset{x}{\sim} x_j \quad \Leftrightarrow \quad C(x_i|x) \cap C(x_j|x) \neq \emptyset \tag{2}$$

for $x_i, x_j \in x$. Notice that the conditional distribution of the colouring of cells given the nuclei x is an Ising ($K = 2$) or Potts ($K > 2$) model with neighbour relation $\underset{x}{\sim}$ and coupling parameter θ. It is convenient to express the density of the prior with respect to the standard Poisson process $\mu = \mu_1$:

$$f(y|\beta, \theta) = \beta^{n(y)} \exp\left(-\theta s(y)\right)/c(\beta, \theta) \tag{3}$$

where $n(y)$ denotes the number of nuclei and $c(\beta, \theta)$ is a normalising constant. The density is well defined for all $\beta > 0$ and $\theta \in \mathbf{R}$, but in the sequel we assume that $\theta \geq 0$, reflecting our prior belief that neighbouring cells tend to be of the same colour. Clearly, if $\theta = 0$ we just obtain the Poisson process μ_β.

In the sense of Baddeley & Møller (1989), (3) defines a nearest-neighbour Markov point process with respect to $\underset{x}{\sim}$ (or more precisely the relation on y as induced by the Voronoi neighbour relation (2)). In fact, setting

$$N(x;\xi) = \{x_i \in x : x_i \underset{x \cup \{\xi\}}{\sim} \xi\} \quad \text{for } \xi \in S \setminus x,$$

the Papangelou conditional intensity defined by

$$\lambda_{\beta,\theta}(y; (\xi, m)) = f(y \cup \{(\xi, m)\}|\beta, \theta)/f(y|\beta, \theta) \quad \text{for } \xi \in S \setminus x,\ m \in M,$$

reduces to

$$\lambda_{\beta,\theta}(y; (\xi, m)) = \beta \exp\left(-\theta s(y; (\xi, m))\right) \tag{4}$$

where

$$s(y; (\xi, m)) = \sum_{x_i \in N(x;\xi)} 1[m_i \neq m] - \sum_{\{x_i, x_j\} \subseteq N(x;\xi):\, x_i \underset{x}{\sim} x_j,\, x_i \underset{x \cup \{\xi\}}{\not\sim} x_j} 1[m_i \neq m_j] \tag{5}$$

depends only on y through the coloured neighbouring nuclei to ξ. This is advantageous in the computations involved in the MCMC algorithm described in Section 4. Here, (5) follows from the fact that $x_i \underset{x \cup \{\xi\}}{\sim} x_j$ implies that $x_i \underset{x}{\sim} x_j$.

Figures 2, 3 and 4 show simulated realisations under the prior when $S = [0, 1]^2$ and different values of K, β and θ are used. While β specifies the 'resolution' in the image (see Figure 2), θ controls the degree of smoothing in the same way as the coupling parameter do for pixel-based Ising/Potts models (see Figures 3 and 4). The realisations show some flexibility in our prior model, though for larger values of θ realisations are typically dominated by one component of the same colour and many small components of different colours. Possibly a phase transition happens for some some value θ_{crit} of θ as S expands to infinity: Figures 3 and 4 may indicate that $\theta_{crit} \in [0.4, 0.5]$ when $K = 2$ and $\theta_{crit} \in [0.6, 0.7]$ when $K = 4$, but we have not investigated this in any further detail.

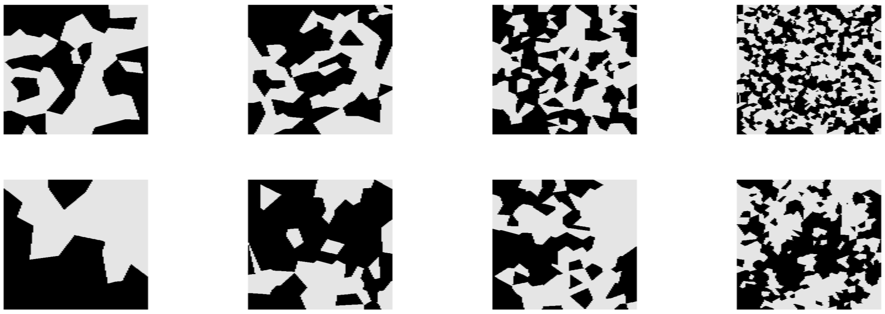

FIGURE 2. Realisations from the prior with $K = 2$, $\theta = 0$ (first row) and $\theta = 0.4$ (second row), and $\beta = 50, 100, 250, 1000$ (left to right).

Obviously our prior (3) can be modified in many ways. For example, the indicator function in (1) could be multiplied by the length of the common edge for $C(x_i|x)$ and $C(x_j|x)$; this would resemble the prior in Nicholls (1998) for triangulation models. One could also be inspired by the ideas in Møller & Waagepetersen (1998), attempting to model the size, shape and other geometrical properties of the connected components of cells with the same colour in the image. For instance, to avoid the many small components which occur in Figures 3 and 4, the prior might be modified to penalise the appearance of such small components, cf. the penalised Ising model studied in Møller & Waagepetersen (1998). Alternatively, following Tjelmeland & Besag (1998) we could consider larger neighbourhoods, though the irregular tessellation structure makes it difficult to have a similar detailed model specification. If the colours are not uniformly distributed, $\beta^{n(y)}$ in (3) could be replaced by $\prod_{m=0}^{K-1} \beta_m^{n_m(y)}$, where $n_m(y)$ is the number of cells with colour m and $\beta_m > 0$ is a parameter. We could also consider other 'reference measures' than μ_β. As an example, heterogeneities in the image or reservoir might be better represented by replacing the parameter β by a (possibly random) intensity function.

In reservoir modelling, available reservoir-specific data are often limited; hence, it may be difficult to validate the prior and to estimate parameters. However, additional data are usually available from other locations thought to have a similar geological structure. For example, vertical cross sections of outcrops may be used to choose a prior and to estimate the model parameters. In addition, geologists have knowledge about how geological processes develop.

4. Bayesian inference and simulation

Using a Bayesian setting, inference is based on the posterior density obtained by multiplying our prior density with a likelihood function for the data given the coloured Voronoi tessellation; specific examples are given in Section 6. The present section contains a more general discussion on Bayesian inference and simulation as related to the applications of our interest.

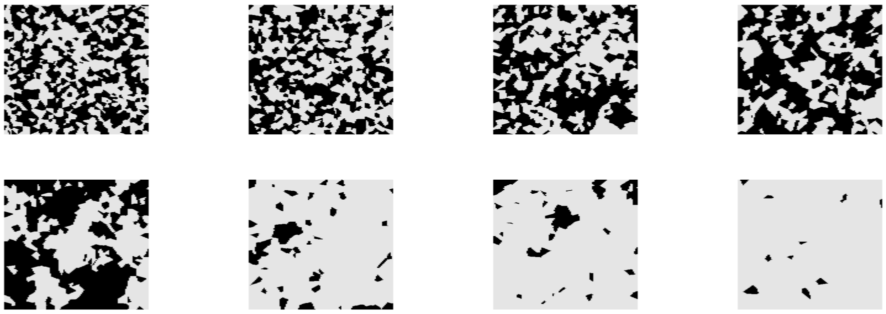

FIGURE 3. Realisations from the prior with $K = 2$, $\beta = 1000$ and $\theta = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7$ (left to right, top to bottom).

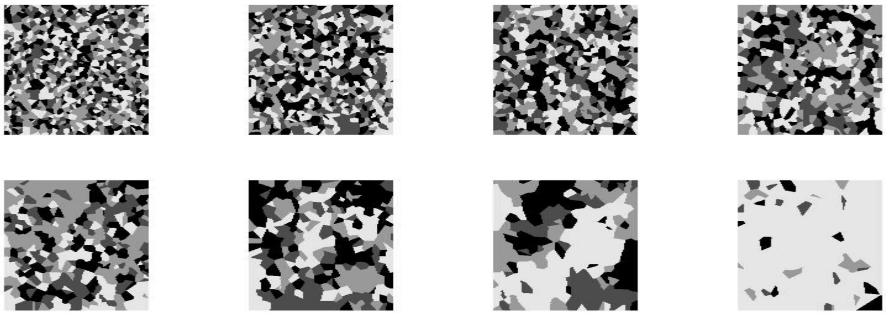

FIGURE 4. Realisations from the prior with $K = 4$, $\beta = 1000$ and $\theta = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7$ (left to right, top to bottom).

Specification of the likelihood term (the conditional density for the data given the 'true image') depends of course on the particular context of application. In Bayesian image analysis, data are most often available as a degraded image, corrupted by noise and (sometimes) smoothed; see, for example, Geman & Geman (1984), Besag (1986) and Ripley (1988). In oil, gas or ground water reservoirs, we have three main sources of reservoir-specific information: observations in wells, seismic data and production tests and history; see Bjørlykke (1989), Tjelmeland & Omre (1997) and Tjelmeland (1996). Well observations give accurate information about facies (rock types) at sparse locations. Seismic data are the result of reflected sound waves, where the seismic signals for each vertical trace can be represented as a convolution of the reflections at the boundaries between facies. The seismic data will have a high degree of spatial dependence. Production tests and history result from fluid flow between wells and can be represented by a fluid flow simulator.

As both the prior and the posterior are analytically intractable we shall use simulations in both cases based on the Metropolis-Hastings algorithm studied in Geyer & Møller (1994), Geyer (1999) and Møller (1999). At each update of the algorithm, either a birth or death or move is proposed, and the proposal is accepted with probability $\min\{1, H\}$, where H is the Hastings ratio as defined below. More precisely, for simplicity a birth and a death are each proposed with the same probability $q \le 0.5$, so a move is proposed with probability $1 - 2q$. Further, in case of a birth proposal $y \to y \cup \{(\xi, m)\}$ we have that ξ and m are independent and uniformly distributed in S and M, respectively, and $H = R(y; (\xi, m))$ where

$$R(y; (\xi, m)) = \frac{\pi(y \cup \{(\xi, m)\})|S|}{\pi(y)(n(y) + 1)}. \tag{6}$$

Here π denotes either the prior or posterior density. Furthermore, in case of a death proposal $y \to y \setminus \{(x_i, m_i)\}$, the marked point (x_i, m_i) is chosen at random from y (i.e. with probability $1/n(y)$), and $H = 1/R(y \setminus \{(x_i, m_i)\}; (x_i, m_i))$. Finally, in case of a move proposal $y \to (y \setminus \{(x_i, m_i)\}) \cup \{(x'_i, m'_i)\}$, we chose (x_i, m_i) at random from y, the d coordinates in $x'_i - x_i$ are independent and normally distributed with mean 0 and variance $\sigma^2_{\text{move}} > 0$. We select $m'_i \in M$ at random and $H = R(y \setminus \{(x_i, m_i)\}; (x'_i, m'_i))/R(y \setminus \{(x'_i, m'_i)\}; (x_i, m_i))$ (setting $H = 0$ if $x'_i \notin S$). As noticed in Section 3, for simulation of the prior and typically also for the posterior, the calculation of (6) and hence of the Hastings ratio involves only local computations.

Using the techniques in Geyer & Møller (1994) and Geyer (1999) we can under fairly weak condition establish convergence of the Markov chain generated by this Metropolis-Hastings algorithm. In fact, in all specific examples considered in this paper, we have geometric ergodicity as the target distribution can be easily shown to be locally stable (see Geyer (1999) and Kendall & Møller (2000) for a discussion of the role of the local stability condition). However, it is difficult exactly to quantify the convergence rates of the chain, cf. Appendix B in Møller (1999).

Figure 5 shows time series for the statistics $n(y)$ and $s(y)$ under the prior (3) with $S = [0, 1]^2$, $\beta = 1000$, $(K, \theta) = (2, 0.4)$ or $(K, \theta) = (4, 0.6)$, and when the chain is started in one of three rather different states. Note that $(n(y), s(y))$ is a minimal sufficient statistic for the parameter (β, θ), and the values of θ were chosen just below θ_{crit}, see Section 3. For all initial values of the chain, the time series seem to be stable after about $25,000$ updates of the Metropolis-Hastings algorithm. We estimated the integrated auto-covariance time (IACT) from the samples of $s(y)$ using methods as described in Geyer (1992). If the chain $y_1, y_2, \ldots$ is in equilibrium, then IACT= $\lim_{N\to\infty} N\mathbf{Var}(\bar{s}_N)/\mathbf{Var}(s(y_1))$ is the ratio between the asymptotic variance of the Monte Carlo estimate $\bar{s}_N = \sum_1^N s(y_i)/N$ of the mean $\mathbf{E}s(y_1)$ and the variance $\mathbf{Var}(s(y_1))/N$ as obtained for the IID case. The estimates were IACT= 6300 for $(K, \theta) = (2, 0.4)$ and IACT= 6900 for $(K, \theta) = (4, 0.6)$. Empirical results when simulating posterior distributions are given in Section 6.

In the algorithm we construct the Delaunay tessellation and thereby obtain the dual Voronoi tessellation. The main computational part of the algorithm is to update the Delaunay cells for each point which is to be added, moved or removed. Two different algorithms are used in the 2D and 3D cases. In the 2D case we modified the Hull software by Ken Clarkson (see http://cm.bell-labs.com/netlib/voronoi/hull.html). This algorithm uses an incremental method for building a convex hull; the Delaunay triangulation may then be found as the projection of this convex hull. In the 3D case we used an incremental method for building a Delaunay tetrahedrisation, partly based on algorithm described in Mücke (1993). If a new point is added, we make first an initial tetrahedrisation by searching for the tetrahedron, t, that contains the new point, and then adding four new tetrahedras inside t while removing t. Second, a number of edge and face flip operations are performed until the Delaunay property is restored: the circumsphere of each tetrahedron does not contain any other points than those belonging to the tetrahedron. The remove operations also involves edge and face flip operations, going in the reverse direction of the add operation.

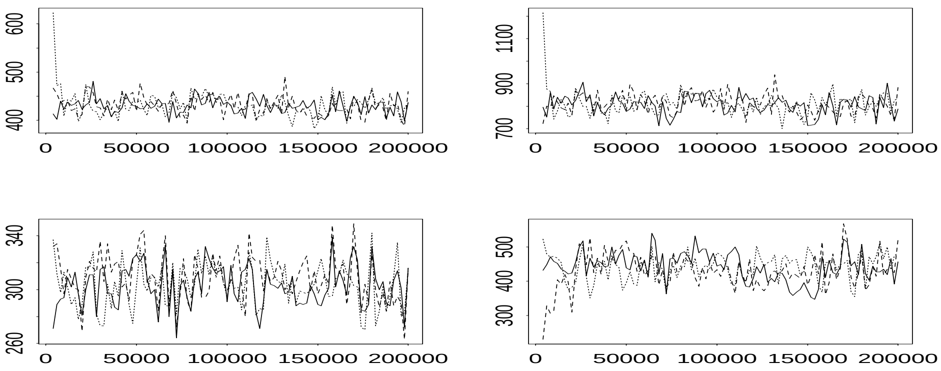

FIGURE 5. Plots of $n(y_i)$ (first column) and $s(y_i)$ (second column) versus 200,000 basic updates when $\beta = 1000$, $(K, \theta) = (2, 0.4)$ (first row) or $(K, \theta) = (4, 0.6)$ (second row), and different initial states are used: $n(y_1) = 0$ (solid line); $n(y_1) = 2\,\mathbf{E}n(y_1)$ (dotted line); $n(y_1) = \mathbf{E}n(y_1)$ and all colours in y_1 are 0 (dashed line).

5. Parameter specification

We now address the question of how to choose or estimate the parameter (β, θ) of the prior (3) and an unknown parameter ϕ of the likelihood function $f(w|y, \phi)$, where w denotes the data; see also Section 6 for specific examples. We assume for simplicity that (β, θ) and ϕ vary independently of each other.

In reservoir modelling applications, parameter estimation may be a difficult task when data are limited. A simple ad-hoc strategy for choosing the parameters in the prior is as follows. We simulate from the prior for different values of (β, θ) and then choose a value that produces realisations that correspond to expert experience of what the image or reservoir should look like. A variant of this approach is adopted by Nicholls (1998). He interactively varies the parameters as the run proceeds and then observes the effect on sample appearance.

On the other hand there may often be training data available that may be used to tune the model parameters. One method uses a MCMC run to estimate unknown normalising constants in the prior or likelihood function, whereby an approximate maximum likelihood estimate (MCMCMLE) may be obtained (Geyer & Thompson 1992, Geyer 1999). This method is used by Tjelmeland (1996) and Tjelmeland & Besag (1998) for higher order Markov random field models, by Møller & Waagepetersen (1998) for Markov connected component field models, and by Syversveen & Omre (1997) in a simple 2D reservoir model where the only available data are scarce well observations.

A much less computationally intensive method is maximum pseudo-likelihood estimation (Besag 1975, Besag 1977, Jensen & Møller 1991, Baddeley & Turner 2000); for models with stronger interactions, this can be less efficient than the MCMCMLE. In the present setting of a marked point process model for the prior parameter (β, θ), the pseudo-likelihood function is given by

$$\mathrm{PL}(\beta, \theta; y) = \exp\left(-\int_S \sum_{m\in M} \lambda_{\beta,\theta}(y; (\xi, m))\, \mathrm{d}\xi \,/K\right) \prod_i \lambda_{\beta,\theta}(y; y_i)$$

where $\lambda_{\beta,\theta}(y; y_i) := \lambda_{\beta,\theta}(y \setminus y_i; y_i)$ for $y_i \in y$, and the maximum pseudo-likelihood estimate is the value of (β, θ) which maximises this function.

Another strategy is to put independent hyper-priors $\pi(\beta, \theta)$ and $\pi(\phi)$ on the parameters and use a fully Bayesian approach. Adopting the approach in Besag (1986) and Heikkinen & Högmander (1994), we can alternately update y, ϕ and (θ, β) using Metropolis-Hastings updates. The update of (β, θ) involves the unknown normalising constant $c(\beta, \theta)$ of our prior (3), so instead this prior term is replaced by the pseudo-likelihood approximation when updating (β, θ); or use the same approach as in MCMC MLE (Higdon 1994). To the best of our knowledge the properties of such an approach for parameter estimation in a point process setting has not yet been investigated.

6. Examples

6.1. **Image analysis experiment.** Consider the binary image A1 in Figure 6 showing a horizontal cross-section from a 3D realisation of the marked point process of Lia et al. (1997). It consists of 50×100 rectangular pixels of size 0.02×0.01 when the image is identified with a unit square. The noisy image A2 is obtained from A1 by adding IID Gaussian white noise with mean 0 and variance σ; for this simple experiment we consider $\sigma = 1$ as a known parameter. Let $w = (w_1, \ldots, w_{5000})$ denote the corresponding values. Given a marked point pattern $y = \{(x_1, m_1), \ldots, (x_n, m_n))\}$, we let $c_j(y) \in \{0, 1\}$ denote the colour of the Voronoi cell containing the midpoint u_j of pixel j, i.e. $c_j(y) = m_i$ if $u_j \in C(x_i|x)$. In fact, c_j is almost surely well defined under the prior (3). The likelihood then takes the form

$$f(w|y) = (2\pi)^{-2500} \exp\left(-\sum_j (w_j - c_j(y))^2/2\right). \tag{7}$$

The image A3 is the corresponding MLE reconstruction, which has an error rate of 30.6% misclassified pixels.

Plot	β	θ	$\mathbf{E}(n(y)\|w)$	$\mathbf{E}(s(y)\|w)$	error	accept
B1	250	0.0	85 (74)	105 (98)	6.0	13
B2	250	0.8	34 (29)	38 (29)	7.8	5
C1	1000	0.0	804 (795)	913 (879)	8.9	47
C2	1000	0.8	397 (399)	210 (226)	6.1	34

TABLE 1. Posterior means of $n(y)$ and $s(y)$, with the actual values of $n(y)$ and $s(y)$ for the realisations in Figure 8 in parenthesis. The column "error" shows the percentage of mis-classified pixels when the MMP estimate is used. The column "accept" shows estimated acceptance probabilities (in percentage) of Metropolis-Hastings updates.

The MCMC algorithm described in Section 4 with $q = 1/3$ and $\sigma_{\text{move}} = 0.05$ was used to obtain samples from the posterior of length 20,000,000 for each value of (β, θ). Table 1 shows simulated results for various values of the parameter (β, θ). Note that the posterior means of the number of cells are considerably smaller than the corresponding prior means. As expected the acceptance probabilities for the Metropolis-Hastings updates increases as the 'resolution-parameter' β increases or as the interaction parameter θ decreases. Figure 7 shows the marginal posterior probabilities of the 5000 pixels. These grey level images are noisy for β large and θ small. At the other extreme, for β small and θ large, we see instead a stronger smoothing of the original image in Figure 6. Maximum marginal posterior (MMP) estimates can be obtained from Figure 7; the corresponding misclassification rates are clearly smaller than the 30.6% for the MLE reconstruction, cf. Table 1.

Realisations from the posterior are shown in Figure 8, which may be compared with Figure 6; see also the comparison of the actual values of the statistics $n(y)$ and $s(y)$ for these realisations and their posterior means in Table 1. Figure 9 shows the underlying structure of the Voronoi tessellations. Some of these tessellations have a variation in cell size that partly reflect the heterogeneity of the data image (which perhaps is more visible in the original image). In any case, we note that with few Voronoi cells, we could represent the large-scale structures and obtain realisations that were close to the original image. For example, the image $B2$ in Figure 8 contains only 29 Voronoi cells. A fixed grid of the same size would give a coarser resolution, in particular when estimating the marginal posterior probabilities.

We have considered various diagnostics for convergence of the algorithm. Three natural statistics to consider are $n(y)$, $s(y)$ and the residual sum of squares $u(y) = \sum_j (w_j - c_j(y))^2$ of the likelihood (7). The statistic $u(y)$ converged more slowly than the others. The time series for $u(y)$ (not shown here) are most slowly mixing for combinations of β small and θ large corresponding to the small acceptance probabilities in Table 1.

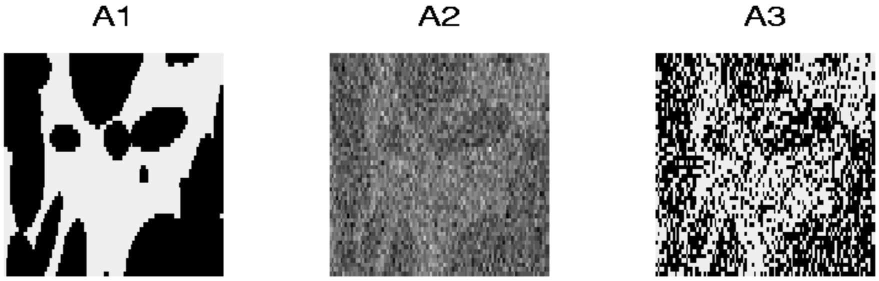

FIGURE 6. A1: the true scene. A2: the true scene corrupted with Gaussian white noise. A3: MLE reconstruction of the true scene.

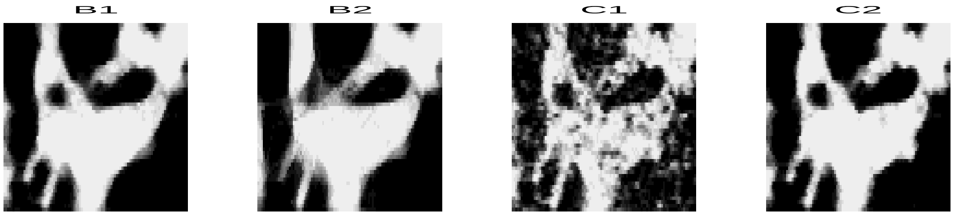

FIGURE 7. Marginal posterior probabilities with values of β and θ as given in Table 1.

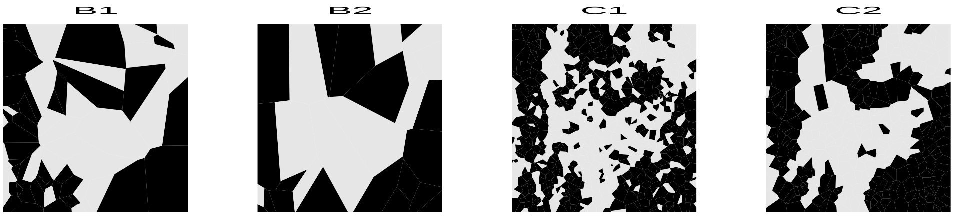

FIGURE 8. Realisations from the posterior with values of β and θ as given in Table 1.

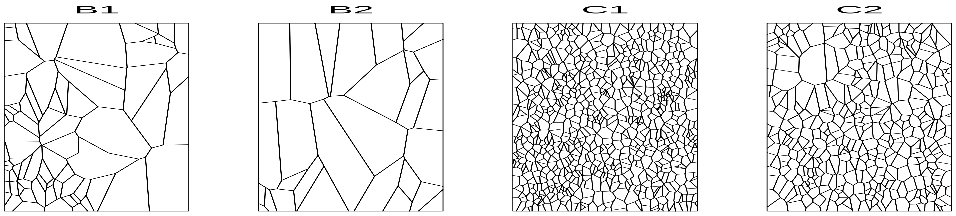

FIGURE 9. Voronoi cell structure of the realisations in Figure 8.

6.2. **A reservoir modelling example.** In this section we model a part of the Heidrun/Tilje formation, which is situated in the north sea outside the coast of Norway. This reservoir is represented as a regular box S of size $2000 \times 5800 \times 15$ m^3 in which the original sedimentation has taken place. The reservoir facies result from tidal and fluvial sedimentary processes which have created a complex facies geometry. Four different facies are modelled: shoal, tidal channel sand, cement and background. The shoal facies consist of large lateral continuous sand objects. The shoal facies and the tidal channel sand are permeable and porous, and the oil and gas will mainly flow through these facies. The calcite cemented facies act as barriers of flow. Furthermore, the background consist of heterolithic facies with low permeability. The data consist of seven (almost vertically) drilled wells with horizontal locations and facies observations as shown in Figures 10 and 11, respectively. The vertical positions of the facies observations are transformed to the interval $[0, 15]$, where 0 and 15 corresponds to top and bottom of S, respectively.

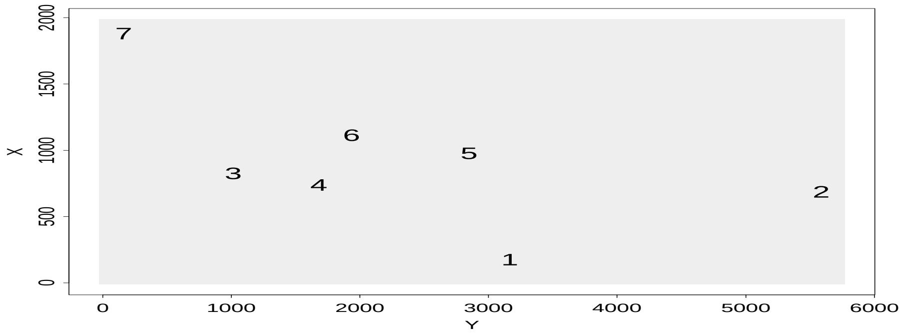

FIGURE 10. The reservoir with the seven drilled wells.

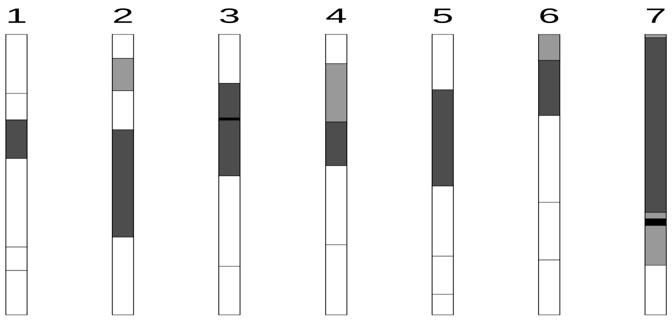

FIGURE 11. The facies observations for each of the seven drilled wells: shoal (dark shade), tidal channel sand (light shade), cemented facies (black) and background facies (white).

We use the following notation. Each well $k \in \{1, \dots, 7\}$ is represented as a (piecewise linear) line segment in 3D that is partitioned into intervals $w^k = \{[w_i^k, w_{i+1}^k[: i = 1, \dots, n_k\}$ corresponding to the different types of facies $p_i^k,\ i = 1, \dots, n_k$ (so $p_i^k \neq p_i^{k+1}$), cf. Figure 11. We assume here an ordering from top to bottom of S, such that w_1^k and $w_{n_k+1}^k$ are at the top and bottom of S, respectively. The observation vector for well k is then $z^k = \{([w_i^k, w_{i+1}^k[, p_i^k) : i = 1, \dots, n_k\}$. The coloured Voronoi tessellation defined by the marked point pattern $y = \{(x_1, m_1), \dots, (x_n, m_n)\}$ will partition well k into intervals $t^k = \{[t_i^k, t_{i+1}^k[: i = 1, \dots, n_k^*\}$, and we let $m_i^k \in \{0, 1, 2, 3\}$ denote the "true" facies type for the interval $[t_i^k, t_{i+1}^k[$ (so $m_i^k \neq m_i^{k+1}$). Let $s^k = \{[s_i^k, s_{i+1}^k[: i = 1, \dots, n_k^{**}\}$ be the coarsest partition so that each interval of t^k and w^k is a union of intervals from s^k. Finally, for each interval $[s_i^k, s_{i+1}^k[$, define $d_i^k = s_{i+1}^k - s_i^k$; the indicator D_i^k which is 1 if and only if observed and true facies agree in this interval; the indicator J_i^k which is 1 if and only if $q_i^k \neq q_{i+1}^k$ where q_i^k is the observed facies type in $[s_i^k, s_{i+1}^k[$. As we have observations of the facies below S, the indicator J_i^k is defined also for $i = n_k^{**}$. For five of the wells we have $J_{n_k^{**}}^k = 0$, while for the two remaining wells we have $J_{n_k^{**}}^k = 1$.

Our likelihood for the data $z = \{z^1, \dots, z^7\}$ given the coloured Voronoi tessellation y is specified as follows. Conditionally on y, the z^k are mutually independent. Letting μ_0 and μ_1 denote two given positive parameters, the conditional distribution of z^k given y is specified by considering each interval $[s_i^k, s_{i+1}^k[$: First, let $1 < i < n_k^{**}$. If we also condition on the observation for the previous intervals and on q_i^k, suppose that τ_i^k is exponentially distributed with parameter $\mu_{D_i^k}$, so that we observe $\tau_i^k = d_i$ if $J_i^k = 1$, while we only know that $\tau_i^k > d_i$ if $J_i^k = 0$ (conceptually one may think of the latter case as "censoring"). As we want to penalise discrepancies between observed and true facies values in wells, we may require that $\mu_0 \gg \mu_1$. If we furthermore condition on the event $J_i^k = 1$, then q_{i+1}^k is uniformly distributed on $\{0, 1, 2, 3\} \setminus \{q_i^k\}$. Next, specifying the conditional distribution in a similar way for the cases where $i = 1$ or $i = n_k^{**}$, we obtain the likelihood

$$L(z|y) = \prod_{k=1}^{7} \prod_{i=1}^{n_k^{**}} (\mu_{D_i^k}/3)^{J_i^k} \exp\{-\mu_{D_i^k} d_i^k\} = (\mu_0/3)^{N_0} (\mu_1/3)^{N_1} \exp\{-L_0\mu_0 - L_1\mu_1\}$$

where

$$N_0 = \#\{(k,i) : D_i^k = 0, J_i^k = 1\}, \qquad N_1 = \#\{(k,i) : D_i^k = 1, J_i^k = 1\}$$
$$L_0 = \sum_{(k,i):D_i^k=0} d_i^k, \qquad L_1 = \sum_{(k,i):D_i^k=1} d_i^k.$$

Since $N_0 + N_1$ is the total number of observed facies changes and $L_0 + L_1$ is the total length of wells,

$$L(z|y) \propto \left(\frac{\mu_0}{\mu_1}\right)^{N_0} \exp\{-L_0(\mu_0 - \mu_1)\}.$$

This likelihood has the following appealing property. We can redefine the meaning of true facies types by considering a finer partioning t^k (and then using this when defining s^k) by allowing m_i^k and m_i^{k+1} to be equal, since this will not influence the likelihood. We found it computational convenient to use the finer partioning as given by the intersection of wells and Voronoi cells.

Note that under the prior (3), we have a uniform prior for the facies proportions, so the posterior proportions will only differ as a result of spatial correlation with observed facies in wells; however, this can easily be modified and other prior information about e.g. size and shape of facies may be modelled as discussed in Section 3. Actually, as the sedimentary facies have a larger vertical than horizontal variation, we modify the Euclidean distance by scaling the z-distance by a factor of 400. We could also have included an anisotropy in the x-y direction to take into account the assumed direction of the sedimentation due to the position of the coastline at the time of deposition.

We used the MCMC algorithm described in Section 4 with $q = 0.35$ and $\sigma_{\text{move}} = 100$ to obtain samples from the posterior. A simulation of $10,000,000$ iterations was run with a burnin of $2,000,000$ iterations. The likelihood parameters μ_0 and μ_1 were fixed to 10.0 and 0.5, respectively. Table 2 shows various results for four models I–IV with different values of the parameter (β, θ). As expected, since the data are scarce, the posterior means of $n(y)$ and $s(y)$ do not differ much from that of the prior. Considering the IACT, we see that the chains for models II and IV mix slowly. This may partly be due to large value of θ in models II and IV. This is also seen in time series plots for the statistic $s(y)$ (not shown here). The mean value of $\Delta L_0 = L_0/(L_0 + L_1)$, the discrepancy ratio, is moderately low, between 3% and 6%. This ratio is influenced by the likelihood parameters μ_0 and μ_1: increasing μ_0 (or decreasing μ_1) will reduce ΔL_0 and would cause slower mixing of the Markov chain. We would also expect that more (well) data would result in slower mixing. Then another MCMC algorithm, e.g. based on simulated tempering (Marinari & Parisi 1992, Geyer & Thompson 1995, Mase, Møller, Stoyan, Waagepetersen & Döge 1999), may be more appropriate.

We have also studied various plots of different kinds of sectional realisations for the models I–IV. Though such plots may be of geological interest, they are omitted here: as expected, they show somewhat similar behaviour at the wells, and smoother realisations are observed under model II and IV, as the colours of the Voronoi tessellation become more correlated with the data when θ increases. Similar features are seen in Figure 12, where the averages of marginal probabilities at horisontal locations for occurence of the tidal channel sand facies are shown.

Table 3 gives the posterior means of the facies proportions and the observed facies proportions in the wells. For the models I and III, the posterior means are close to the prior mean of 0.25, while for the models II and IV with high interaction the posterior means agree better with the observed facies proportions in wells.

model	$\beta\|S\|$	θ	$\mathbf{E}(n(y))\|\boldsymbol{z})$	$\mathbf{E}(s(y)\|\boldsymbol{z})$	$\mathbf{E}(\Delta L_0\|\boldsymbol{z})$	IACT
I	1000	0.0	952 (1000)	5069 (5420)	0.036	9,900
II	1000	0.5	430 (337)	1172 (573)	0.028	77,600
III	4000	0.0	3931 (4000)	21974 (22463)	0.059	28,000
IV	4000	0.5	1425 (1161)	3094 (1389)	0.037	263,500

TABLE 2. Posterior means of $n(y)$, $s(y)$ and $\Delta L_0 = L_0/(L_0 + L_1)$. The column "IACT" shows the computed integrated auto-covariance time for the $s(y)$ statistic. The corresponding results for the prior are given in parenthesis.

model	shoal	tidal	cement	background
I	0.2534	0.2297	0.2200	0.2969
II	0.2029	0.0615	0.0395	0.6961
III	0.2512	0.2423	0.2381	0.2684
IV	0.1526	0.0429	0.0320	0.7725
Well	0.3087	0.0848	0.0046	0.6019

TABLE 3. Posterior means of facies proportions and observed facies proportions in the wells.

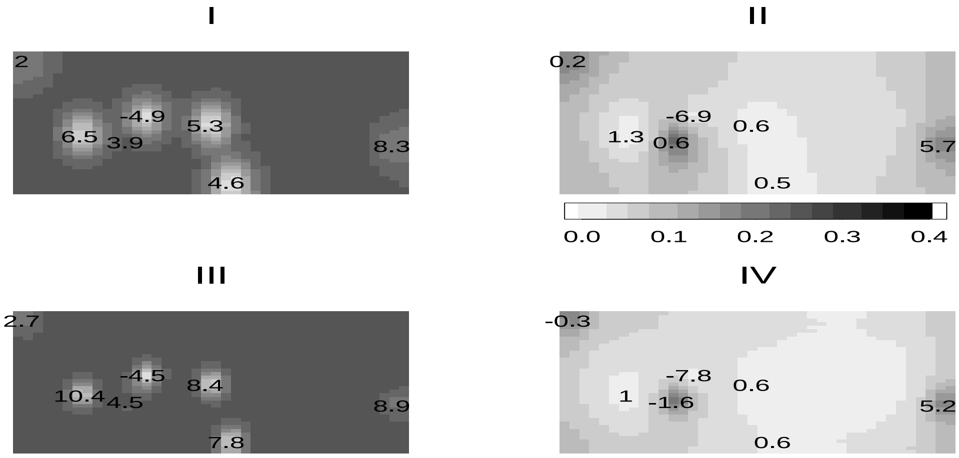

FIGURE 12. Averages of marginal probabilities at horisontal locations for occurence of the tidal channel sand facies, with values of β and θ as given in Table 2. The figures at the seven well locations show the difference, multiplied by 100, between these averages and the observed averages.

7. Closing remarks

Our coloured Voronoi tessellation model serves as a general purpose prior model for applications in Bayesian image analysis and reservoir modelling. It may be particularly useful when modelling complex and irregular geometries, which may be difficult to represent by marked point processes for objects against a background. The advantage to pixel-based random field models is that we may capture large-scale characteristics by a lower dimensionality of the state space, particularly in the 3D case of applications. The model may easily be elaborated, e.g. by using higher order interactions (Tjelmeland & Besag 1998) or by introducing penalising terms (Møller & Waagepetersen 1998). The examples discussed in Section 6 show clearly the sensitivity and importance of the prior and the parameter specification part, particulaly in sparse data situations, and it remains to elaborate further on this, possibly in combination with one or several of the methods mentioned in Section 5. For reservoir characterisation with sparse data, it may be even more relevant to use the geological knowledge when making an appropriate specification.

The Metropolis-Hastings algorithm used in this paper can easily be refined to ensure faster convergence towards the posterior, e.g. by generating points near locations where data indicate shift in colour, by letting the colours of a Voronoi cell be drawn conditioned on the neighbour colours, and by using simulated tempering as mentioned in Section 6.

Acknowledgements: JM was supported by MaPhySto (Centre for *Ma*thematical *Phy*sics and *Sto*chastics, funded by a grant from the Danish National Research Foundation). ØS was supported by NORFA and the Research Council of Norway. The collaboration producing this paper was supported by the European Union's research network "Statistical and Computational Methods for the Analysis of Spatial Data. ERB-FMRX-CT96-0095". We thank Geir Aamodt, Ragnar Hauge, Juha Heikkinen, Jacob Kooter Laading, Stein Overvoll, Håvard Rue, Bård Storvik and Håkon Tjelmeland for helpful comments. We thank Statoil for providing us with the 3D well data, and thank in particular Oddvar Lia and Arve Næss for help and support.

References

Ahuja, N. & Schachter, B. J. (1983), *Pattern Models*, John Wiley & Sons, New York.

Arak, T., Clifford, P. & Surgailis, D. (1993), 'Point based polygonal models for random graphs', *Adv. Appl. Probab.* **25**, 348–372.

Baddeley, A. J. & Van Lieshout, M. N. M. (1993), Stochastic geometry models in high-level vision, *in* K. V. Marida & G. K. Kanjii, eds, 'Statistics and image', Special issue Journal of Applied Statistics, Carfax, Abingdon, pp. 231–256.

Baddeley, A. & Møller, J. (1989), 'Nearest-neighbour Markov point processes and random sets', *Internat. Statist. Rev.* **57**, 89–121.

Baddeley, A. & Turner, R. (2000), 'Practical maximum pseudolikelihood for spatial point patterns', *Australian and New Zealand Journal of Statistics* **42**, 283–322.

Besag, J. (1975), 'Statistical analysis of non-lattice data', *The Statistician* **24**, 179–195.

Besag, J. (1977), 'Some methods of statistical analysis for spatial data', *Bulletin of the International Statistical Institute* **47**, 77–92.

Besag, J. (1986), 'On the statistical analysis of dirty pictures', *J. R. Statist. Soc.* B **48**, 259–302.

Bjørlykke, K. (1989), *Sedimentology and Petroleum Geology*, Springer-Verlag Inc., Berlin.

Damsleth, E., Tjølsen, C. B., Omre, H. & Haldorsen, H. H. (1990), A two-stage stochastic model applied to a north sea reservoir, *in* '65th Annual Technical Conference and Exhibition', Soc. of Petroleum Engineers, New Orleans. SPE 20605.

Geman, S. & Geman, D. (1984), 'Stochastic relaxation, Gibbs distribution, and Bayesian restoration of images', *IEEE Trans. Pattern Analysis and Machine Intelligence* **6**, 721–741.

Geyer, C. J. (1992), 'Practical Markov chain Monte Carlo', *Statistical Science* **7**, 473–483.

Geyer, C. J. (1999), Likelihood inference for spatial point processes, *in* O. Barndorff-Nielsen, W. Kendall & M. van Lieshout, eds, 'Stochastic Geometry: Likelihood and Computation', Chapman and Hall/CRC, Boca Raton, pp. 79–140.

Geyer, C. J. & Møller, J. (1994), 'Simulation procedures and likelihood inferences for spatial point processes', *Scand. J. Statist.* **21**(4), 369–373.

Geyer, C. J. & Thompson, E. A. (1992), 'Constrained Monte Carlo maximum likelihood for dependent data', *J. R. Statist. Soc.* B **54**, 657–699.

Geyer, C. J. & Thompson, E. T. (1995), 'Annealing Markov chain Monte Carlo with applications to ancestral inference', *J. Amer. Statist. Assoc.* **90**, 909–920.

Green, P. J. (1995), 'Reversible jump Markov chain Monte Carlo computation and Bayesian model determination', *Biometrika* **82**, 711–732.

Grenander, U. & Miller, M. (1994), 'Representation of knowledge in complex systems', *J. R. Statist. Soc.* B **56**, 549–603.

Heikkinen, J. & Högmander, H. (1994), 'Fully Bayesian approach to image restoration with an application in biogeography', *Appl. Statist.* **43**(4), 569–582.

Higdon, D. (1994), 'Spatial applications of Markov chain Monte Carlo for Bayesian inference', Ph.D. thesis, Department of Statistics, University of Washigton.

Hjort, N. L. & Omre, H. (1994), 'Topics in spatial statistics', *Scand. J. Statist.* **21**(4), 289–357.

Holden, L., Hauge, R., Skare, Ø. & Skorstad, A. (1998), 'Modelling of fluvial reservoirs with object models', *Math. Geol.* **30**(5), 473–496.

Jensen, J. & Møller, J. (1991), 'Pseudolikelihood for exponential family models of spatial point processes', *Ann. Appl. Prob.* **3**, 445–461.

Kendall, W. & Møller, J. (2000), 'Perfect simulation using dominating processes on ordered spaces, with application to locally stable point processes', *Adv. Appl. Prob.* **32**, 844–865.

Lia, O., Tjelmeland, H. & Kjellesvik, L. (1997), Modeling of facies architecture by marked point models, *in* E. Y. Baafi & N. A. Schofield, eds, 'Geostatistics Wollongong '96', proc. '5th Inter. Geostat. Congr.', Wollongong Australia, 1996, Kluwer Academic Publ., Dordrecht, pp. 386–397.

Marinari, E. & Parisi, G. (1992), 'Simulated tempering: a new Monte Carlo scheme', *Europhys. Lett.* **19**, 451–458.

Mase, S., Møller, J., Stoyan, D., Waagepetersen, R. & Döge, G. (1999), 'Packing densities and simulated tempering for hard core Gibbs point processes', Research Report R-99-2002, Department of Mathematical Sciences, Aalborg University. To appear in *Ann. Inst. Statist. Math.*

Møller, J. (1994), *Lectures on Random Voronoi Tessellations*, number 87 *in* 'Lecture Notes in Statistics', Springer-Verlag Inc., New York.

Møller, J. (1999), Markov chain Monte Carlo and spatial point processes, *in* O. Barndorff-Nielsen, W. Kendall & M. "van Lieshout", eds, 'Stochastic Geometry: Likelihood and Computation', Chapman and Hall/CRC, Boca Raton, pp. 79–140.

Møller, J. & Waagepetersen, R. P. (1998), 'Markov connected component fields', *Adv. Appl. Prob.* **30**, 1–35.

Mücke, E. (1993), Shapes and Implementations in Three-Dimensional Geometry, Phd thesis, University of Illinois.

Munthe, K., Holden, L., Mostad, P. & Townsend, C. (1994), Modelling sub-seismic fault patterns using a marked point process, *in* 'ECMOR IV, 4th European Conference on the Mathematics of Oil Recovery', Røros, Norway.

Nicholls, G. K. (1997), Coloured continuum triangulation models in the Bayesian analysis of two dimensional change point problems, *in* C. A. G. K. V. Mardia & R. G. Aykroyd, eds, 'The art and science of Bayesian image analysis', Proceedings of the Leeds Annual Statistics Research Workshop 1997, Leeds University Press.

Nicholls, G. K. (1998), 'Bayesian image analysis with Markov chain Monte Carlo and coloured continuum triangulation models', *J. R. Statist. Soc.* B **60**, 643–659.

Okabe, A., Boots, B. & Sugihara, K. (1992), *Spatial Tessellations - Concepts and Applications of Voronoi Diagrams*, John Wiley & Sons, Chichester.

Ripley, B. D. (1988), *Statistical Inference for Spatial Processes*, Cambridge University Press, Cambridge.

Rue, H. & Hurn, M. (1999), 'Bayesian object identification', *Biometrika* **86**(3), 649–660.

Scheike, T. H. (2000), 'Anisotropic growth of voronoi cells', *Adv. Appl. Prob.* **26**, 43–53.

Syversveen, A. & Omre, H. (1997), 'Conditioning of marked point processes within a Bayesian framework', *Scand. J. Statist.* **24**(3), 341–352.

Tjelmeland, H. (1996), Stochastic Models in Reservoir Characterization and Markov Random Fields for Compact Objects, PhD thesis, Department of Mathematical Sciences, Norwegian University of Science and Technology.

Tjelmeland, H. & Besag, J. E. (1998), 'Markov random fields with higher-order interactions', *Scand. J. Statist.* **25**(3), 415–433.

Tjelmeland, H. & Omre, H. (1997), A complex sand-shale facies model conditioned on observations from wells, seismics and production, *in* E. Y. Baafi & N. A. Schofield, eds, 'Geostatistics Wollongong '96, Proceedings of the Fifth International Geostatistical Congress', Kluwer Academic Publishers, Wollongong.