

Multi-scale Random Field Models

Marco A. R. Ferreira

Universidade Federal do Rio de Janeiro, Brazil

†

David Higdon

Los Alamos National Laboratories, USA

Herbert K. H. Lee

University of California at Santa Cruz, USA

Mike West

Duke University, USA

Summary. We introduce a class of multi-scale models for random fields. The novel framework couples standard Markov models for the random field stochastic process at different levels of resolution, and links them via error models to induce a new and rich class of structured linear models reconciling modelling and information at different levels of resolution. Jeffrey's rule of conditioning is used to revise the implied distributions and ensure that the probability distributions at different levels are strictly compatible. Bayesian estimation based on Markov Chain Monte Carlo methods is developed. To highlight the potential applications of our multi-scale framework, we provide two examples. In the first example, we illustrate with a simulated data set the procedures of multi-scale field simulation and parameter estimation. In the second example, we use our multi-scale model as a prior for permeability fields to solve a fluid flow inverse problem.

1. Introduction

Multi-scale modelling arises in a wide variety of applications. There are at least three classes of problems that can be modelled most effectively within a multi-scale framework. In the first type, data are observed at different spatial scales and the model is used to integrate the information from the different scales. In the second type, data are observed only at the finest scale and the model is used to induce a particular process at that scale. Finally, the observed data are related non-locally and nonlinearly to an underlying multi-scale process, and the model can be used as a prior for that process.

In this paper, we introduce a class of multi-scale random fields useful for modelling processes that live and possibly can be observed at different levels of resolution. Each level of resolution is connected with the immediately finer level through a linear function plus Gaussian noise. We start with a Markov random field process for each level, and then we use Jeffrey's rule of conditioning (Jeffrey 1988) to revise the implied distributions and ensure that the probability distributions of the different levels are compatible. Analogously to

†*Address for correspondence:* Marco A. R. Ferreira, Universidade Federal do Rio de Janeiro, Instituto de Matemática, Departamento de Métodos Estatísticos, CP: 68530, Rio de Janeiro - RJ, CEP: 21945-970, Brazil
E-mail: marco@im.ufrj.br

multi-scale time series analysis (see Ferreira et al. (2003) and references therein), this results in a novel class of models for random fields that exhibit a variety of spatial correlation functions based on a very parsimonious parameterisation, has the ability to combine information across levels of resolution and to emulate long range dependence spatial processes.

Multi-scale modelling has appeared mainly in the engineering literature, focused on the development of coarser representations of the phenomenon of interest in order to obtain fast computational algorithms. In most of that literature, the statistical structure is isolated from scale to scale, thus there is no consistent joint multi-resolution statistical model, as for example in the work of Saquib et al. (1996), Comer and Delp (1999) and Pizurica et al. (2002). An exception is the work developed by Allan S. Willsky and coauthors in the past decade, summarised in the review paper Willsky (2002). In that body of work, Willsky and coauthors developed multi-scale models in dyadic and quad trees, in which sites of a given level are conditionally independent given the immediate coarser level. This allows a state-space representation of those models, and thus a variant of the Kalman filter can be used for inference (for references on state-space models and the Kalman filter, see Harvey (1989) and West and Harrison (1997)). Fully Bayesian approaches generalising that body of work are reasonably new and have been applied to a variety of fields. For example, Kolaczyk (1999) introduces Bayesian multi-scale models based on recursive dyadic partitions for Poisson processes, Nowak and Kolaczyk (2000) use such framework to solve Poisson inverse problems, Nowak (1999) proposes a multi-scale hidden Markov model for Bayesian image analysis, Kolaczyk and Huang (2001) construct a multi-scale model for spatial aggregation, Huang et al. (2002) use multi-scale models to perform fast spatial prediction for global processes. Although those models lead to very efficient algorithms, they introduce artifacts in the analysis such as the blocky behaviour pointed out by Irving et al. (1997). A possible remedy to those artifacts is to define a multi-scale model on a more general graph as proposed by Huang and Cressie (2001), where each site at a given resolution level depends on more than one site at the immediate coarser level, and the knowledge of the immediate coarser level decorrelates the sites at that given resolution level. The model by Huang and Cressie (2001) leads to smoother processes than the tree-based models at each resolution level but, as the tree-based models, does not take into account the dynamics at each level of resolution. In contrast, our models assume the existence of dependence between sites of a given resolution level even conditional on the immediate coarser level, and as such they allow fairly smooth processes at the different levels of resolution and incorporate the dynamics at each level.

Here we construct multi-scale random field models with interconnected resolution levels and smoothness within each level. For example, Figure 1 represents a multi-scale random field with three levels of resolution. At the coarsest and intermediate levels of resolution, each site corresponds to four sites of the immediate finer level. In general, we consider a multi-scale random field with L levels of resolution, where x_l denotes the vectorised l -th level of resolution with n_l elements, $l = 0, \dots, L - 1$; x_{lj} denotes the value of the j -th site (or region) of the l -th level and (l, j) its corresponding index. We assume that site (l, j) has b_{lj} neighbour sites at the same level of resolution, one parent site at the immediate coarser level $l - 1$, and $m_{l+1,j}$ descendant sites at the immediate finer level $l + 1$. We denote by N_{lj} the set of sites of the l -th level that are neighbours of (l, j) and by D_{lj} the set of sites of the $(l + 1)$ -th level that are descendants of (l, j) . We shall impose some stochastic structure at each level, such as Markov Random Fields, and connect the levels through some coarsening operation, such as averaging or sampling, but in general this leads to incompatible probability distributions. We use Jeffrey's rule of conditioning to revise

the implied distributions and ensure that the probability distributions at different levels are strictly compatible.

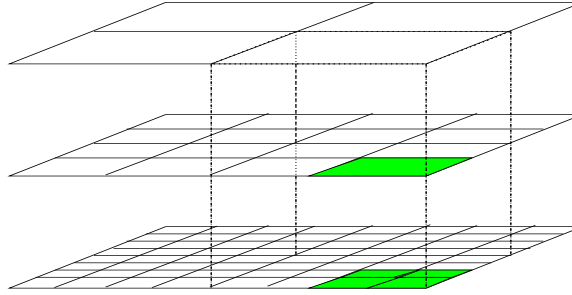


Fig. 1. Three level random field

The remainder of this paper is organized as follows. In Section 2 we present the construction of a two level model in order to introduce the basic ideas of the construction of multi-scale random fields. We generalise our class of models to an arbitrary number of levels in Section 3. In Section 4 we describe the algorithmic implementation of our model by means of a Markov chain Monte Carlo (MCMC) sampling scheme. In Section 5.2 we present two applications of our multi-scale random fields: In the first example, we illustrate with a simulated data set the procedures of multi-scale field simulation and parameter estimation. In the second example, we use our multi-scale model as a prior for 2-D permeability fields to solve a fluid flow inverse problem. We conclude with a brief discussion in Section 6.

2. Construction of a two level model

In this section, we introduce the main concepts on multi-scale modelling through the construction of a two level multi-scale random field model. The definition of the model is somehow subtle – It is important to carefully consider on what information we are conditioning each equation, otherwise the model is inconsistent. To be more explicit, in order to impose smoothness within each resolution level, we assume that coarse and fine levels follow Markov random field processes. Moreover, we connect the levels through a linear link equation with Gaussian noise. Obviously, if these three probabilistic statements were conditioned on the same information, they would generally be contradictory. We reconcile these probabilistic statements through Jeffrey’s rule of conditioning, where the key is to consider the order of arrival of the different information and to revise *part* of the old information that is in conflict with the new information.

We start by assuming that the fine level $x_1 = (x_{11}, \dots, x_{1n_1})$ can be modelled by a proper Markov Random Field (PMRF) model as considered by Ferreira and de Oliveira (2004)

$$x_1 \sim N(\mu \vec{1}_{n_1}, \Sigma_1), \quad (1)$$

where $\mu \in \mathbb{R}$ is a location parameter, $\vec{1}_n$ is a n -dimensional vector of ones and $\Sigma_1^{-1} = \tau_1(\alpha_1 I_{n_1} + H_1)$ is the precision matrix, with $\tau_1 > 0$ a scale parameter, I_n the $n \times n$ identity

matrix and

$$(H_1)_{kj} = \begin{cases} h_{1k}, & j = k \\ -g_{1kj}, & (1, j) \in N_{1k} \\ 0, & \text{otherwise,} \end{cases} \quad (2)$$

where $g_{1kj} > 0$ is a ‘measure of similarity’ between sites $(1, k)$ and $(1, j)$ and $h_{1k} = \sum_{j \in N_{1k}} g_{1kj}$. The parameter $\alpha_1 > 0$ is a ‘spatial’ parameter that controls the strength of association between the components of x_1 and determines the main properties of the PMRF. Following Equations (1) and (2), the conditional distribution of x_{1k} conditional on all the other sites of the fine level field is $N(a_{1k}, C_{1k})$, where $C_{1k} = \{\tau_1(\alpha_1 + h_{1k})\}^{-1}$ and $a_{1k} = \tau_1 C_{1k}(\alpha_1 \mu + \sum_{j \in N_{1k}} g_{1kj} x_{1j})$. When $\alpha_1 \rightarrow 0$, the PMRF approaches the intrinsic autoregressive model (Besag *et al.* 1991; Besag and Koopberg 1995), which is an improper distribution that has been extensively used in spatial statistics as a prior for latent processes or random effects (Sun *et al.* 1999; Carlin and Banerjee 2003). In this paper, in order to guarantee that the matrix Σ_1^{-1} is positive definite we assume $\alpha_1 > 0$, since that implies Σ_1^{-1} diagonally dominant and so positive definite (Harville, 1997). The positive definiteness of Σ_1^{-1} and thus the existence of Σ_1 are fundamental for our multi-scale model construction.

We assume a coarsening operation through a linear link equation establishing that the value of each site at the coarse level is equal to a weighted average of the corresponding sites at the fine level plus an error term. More specifically, the link equation is defined as follows:

$$p(x_0|x_1) = N(x_0|A_1 x_1, \delta_1 I_{n_0}), \quad (3)$$

where the sum of the elements of each line of A_1 is equal to one. For example, the term $A_1 x_1$ can represent arithmetic block averages or sampling of the fine level x_1 . In that case, Equation (3) becomes

$$p(x_0|x_1) = \prod_{i=1}^{n_0} N(x_{0i}|m_{1i}^{-1} \sum_{j \in D_{0i}} x_{1j}, \delta_1) \quad (4)$$

The MRF model at the fine level x_1 and the link Equation (3) imply the particular model $p(x_0) = N(x_0|\mu 1_{n_0}, A \Sigma_1 A' + \delta_1 I_{n_0})$ for the coarse level. As discussed by Lakshmanan and Derin (1993), in general the Markovianity is lost in this coarsening operation. Thus, the resulting model at the coarse level is more complex than desired and does not lead to efficient algorithms to incorporate information at different resolutions. In order to deal with this problem, Lakshmanan and Derin (1993) approximate the process at the coarse level by a Markov random field, but then the joint probabilistic model of fine and coarse levels becomes inconsistent. We take here a different route, in which we revise part of the information contained in Equations (1) and (3) in order to have a simple process at the coarse level and to reconcile the models at the different levels of resolution.

Suppose that we receive additional information about the coarse level x_0 , information that supersedes the prior information on which $p(x_0)$ is based and directly revises $p(x_0)$ to an updated model $q(x_0)$. For example, suppose $q(x_0)$ is a PMRF:

$$q(x_0) = N(x_0|\mu \vec{1}_{n_0}, Q_0), \quad (5)$$

where $Q_0^{-1} = \tau_0[\alpha_0 I_{n_0} + H_0]$, and

$$\{H_0\}_{kl} = \begin{cases} h_{0k}, & k = l, \\ -g_{0kl}, & (0, k) \in N_{0l}, \\ 0, & \text{otherwise.} \end{cases}$$

The revision of $p(x_0)$ to $q(x_0)$ can be viewed as Bayesian updating with an implicit likelihood function proportional to $q(x_0)/p(x_0)$. Let us assume that in the revision of beliefs the coarse level x_0 is sufficient for the fine level x_1 , where this meaning of sufficiency is as defined by Diaconis and Zabell (1982), i.e., the updated conditional distribution of the fine level depends only on the values at the coarse level. Thus we can apply Jeffrey's rule of conditioning to revise the marginal distribution of x_1 (for details, see Jeffrey 1988 and Diaconis and Zabell 1982). The following theorem establishes the resulting multi-scale model with two levels of resolution.

THEOREM 2.1. *Consider the initial model (1) for the fine level x_1 , the link equation (3) and the revised model (5) for the coarse level x_0 . If in the revision of beliefs the coarse level x_0 is sufficient for the fine level x_1 , then*

i) The joint multi-scale model for coarse and fine levels is

$$q(x_0, x_1) = N(x_1 | \mu \vec{1}_{n_1} + B_1(x_0 - \mu \vec{1}_{n_0}), \Sigma_1 - B_1 W_1 B_1') N(x_0 | \mu \vec{1}_{n_0}, Q_0); \quad (6)$$

ii) The revised marginal model for the fine level is

$$q(x_1) = N(x_1 | \mu \vec{1}_{n_1}, \Sigma_1 - B_1(W_1 - Q_0)B_1'), \quad (7)$$

where $B_1 = \Sigma_1 A_1' W_1^{-1}$ and $W_1 = A_1 \Sigma_1 A_1' + \delta_1 I_{n_0}$.

Proof. See Appendix. □

3. Construction of a model with several levels

Here we generalise the construction of the multi-scale model to any arbitrary number of levels through the use of a result by Skyrms (1980), that is, the same effect of Jeffrey's rule can be obtained through the use of sufficiently richer sample space and by careful conditioning of each equation on different information. In addition, conditioning each equation on different information provides a different interpretation and enlightens interesting aspects of the multi-scale model.

We consider a multi-scale random field model with L levels of resolution. The l -th level is denoted by x_l , $l = 0, 2, \dots, L - 1$ where x_0 is the coarsest level and increases in l mean progressively finer levels. The "movement" from coarser to finer levels is interpreted as an increase in the resolution analogously to the wavelet framework, but in our multi-scale framework each level has a meaningful practical interpretation. For example, our multi-scale model can potentially be used to combine spatial information at the zip code area, county and state levels.

Each equation is conditional on different information, as the equations would be incompatible if conditioned on the same information. Here, we denote by I_l the initial knowledge about the behaviour on the l -th level, and by G_l the accumulated knowledge from the coarsest level up to the l -th level x_l . More specifically, $G_0 = I_0$ denotes the knowledge about the generator mechanism of the coarsest level x_0 and the accumulated knowledge up to level x_0 as well. In addition $G_1 = I_0 \cap I_1$. In general, we recursively define $G_l = G_{l-1} \cap I_l$.

3.1. Model

We define the general multi-scale random field model in a hierarchical way; conditional on all the information used to construct the model, the true dependence direction is from

coarser to finer levels. To be concrete, we develop the ideas assuming that the level x_l given I_l , the initial information about level l , follows a PMRF process. Then, we have the following equation conditional on $I_l, l = 0, \dots, L - 1$:

$$x_l | I_l \sim N(\mu \vec{1}_{n_l}, \Sigma_l) \quad (8)$$

where $\Sigma_l^{-1} = \tau_l(\alpha_l I_{n_l} + H_l)$ and the role and definitions of τ_l, α_l and H_l are analogous to the role and definitions of τ_1, α_1 and H_1 in Section 2. In order to develop estimation and simulation methodologies, it is useful to rewrite Equation 8 as (Ferreira and De Oliveira 2004)

$$\begin{aligned} p(x_l | I_l) &\propto \tau_l^{0.5n} \prod_{k=1}^n (\lambda_{lk} + \alpha_l)^{0.5} \\ &\times \exp \left\{ -0.5\tau_l \left(\mu^2 n_l \alpha_l - 2\mu\alpha_l \sum_{k=1}^n x_{lk} + \alpha_l \sum_{k=1}^n x_{lk}^2 + x_l' H_l x_l \right) \right\}. \end{aligned} \quad (9)$$

where $\lambda_{l1} \geq \lambda_{l2} \geq \dots \geq \lambda_{l,n-1} > \lambda_{ln} = 0$ are the eigenvalues of H_l .

In addition, we assume that conditional on I_l there exists a known linear transformation mapping the level x_l to the immediate coarser level x_{l-1} , plus noise ($l = 1, \dots, L - 1$):

$$p(x_{l-1} | I_l, x_l) = N(x_{l-1} | A_l x_l, U_l), \quad (10)$$

where $U_l = \delta_l I_{n_{l-1}}$ and the sum of the elements of each line of A_l is equal to one. This representation is useful when the relation between coarser and finer levels occurs in a non-regular grid and the number of descendants or the weights are not constant.

In the simplest case we have the following equation in which each element of x_{l-1} is written as an arithmetic average of its descendants at level l plus noise:

$$x_{l-1,s} = m_{ls}^{-1} \sum_{w \in D_{l-1,s}} x_{lw} + u_{l-1,s} \quad (11)$$

where $u_{l-1,s}$, ($l = 1, \dots, L$), ($s = 1, \dots, n_l$), are mutually uncorrelated zero-mean, normally distributed noise terms with $u_{l-1,s} \sim N(0, \delta_l)$ for some *between levels* variance δ_l .

Integrating out level x_l to obtain the distribution of level x_{l-1} given I_l yields $p(x_{l-1} | I_l) = N(x_{l-1} | \mu \vec{1}_{n_{l-1}}, A_l \Sigma_l A_l' + U_l)$.

Now, we assume that we receive updated information G_{l-1} about the generator mechanism of the coarser level, information that partially supersedes I_l . Conditional on this new information, the coarser level x_{l-1} follows a Gaussian process with mean $\mu \vec{1}_{n_{l-1}}$ and covariance matrix Q_{l-1} :

$$p(x_{l-1} | G_{l-1}) = N(x_{l-1} | \mu \vec{1}_{n_{l-1}}, Q_{l-1}). \quad (12)$$

As the new information goes down the hierarchy, we want to revise the information on the process at the l -th resolution level. Analogously to the construction of the two level model, the basic principle is that the accumulated knowledge G_l supersedes the initial knowledge I_l on which the initial distribution for x_l was based, and is then sent down to finer levels to update the joint distribution throughout the hierarchy.

In order to obtain the joint multi-scale model for (x_0, \dots, x_{L-1}) given all the knowledge G_{L-1} , we assume the following two hypothesis:

- (a) $p(x_0, \dots, x_{l-1} | G_l) = p(x_0, \dots, x_{l-1} | G_{l-1}, I_l) = p(x_0, \dots, x_{l-1} | G_{l-1})$
 (b) $p(x_l | G_l, x_0, \dots, x_{l-1}) = p(x_l | G_{l-1}, I_l, x_{l-1}) = p(x_l | I_l, x_{l-1})$

Hypothesis (a) states that the coarser levels x_0, \dots, x_{l-1} are independent of the generator mechanism of the finer level x_l given their own generator mechanism; this hypothesis is equivalent to the subjective update of the joint distribution of x_0, \dots, x_{l-1} . Hypothesis (b) states that given the initial information I_l on the generator mechanism of level l and the value of x_{l-1} , level x_l is independent of the revised information G_{l-1} on the generator mechanism of level $l-1$; this hypothesis is equivalent to the Jeffrey's rule sufficiency hypothesis. The following theorem establishes the resulting multi-scale model with L levels of resolution.

THEOREM 3.1. *The model defined by Equations (8) and (10) and Hypotheses (a) and (b) has the following properties:*

i) *The joint multi-scale model for (x_0, \dots, x_{L-1}) given all the knowledge G_{L-1} is*

$$p(x_0, \dots, x_{L-1} | G_{L-1}) = N(x_0 | \mu \vec{1}_{n_0}, Q_0) \prod_{l=1}^{L-1} N(x_l | \mu \vec{1}_{n_l} + B_l(x_{l-1} - \mu \vec{1}_{n_{l-1}}), \Sigma_l - B_l W_l B_l');$$

ii) *The marginal model for $x_l, l = 1, \dots, L-1$ given all the knowledge G_{L-1} is*

$$p(x_l | G_{L-1}) = N(x_l | \mu \vec{1}_{n_l}, Q_l);$$

iii) *The conditional model of x_l given the knowledge from the coarsest up to the $(L-1)$ -th resolution level and the realized process from the coarsest up to the $(l-1)$ -th level is*

$$p(x_l | G_{L-1}, x_0, \dots, x_{l-1}) = N(x_l | \mu \vec{1}_{n_l} + B_l(x_{l-1} - \mu \vec{1}_{n_{l-1}}), \Sigma_l - B_l W_l B_l'),$$

where $B_l = \Sigma_l A_l' W_l^{-1}$, $W_l = A_l \Sigma_l A_l' + \delta_l I_{n_{l-1}}$ and $Q_l = \Sigma_l - B_l (W_l - Q_{l-1}) B_l'$.

Proof. See Appendix. □

In general, direct computation of the matrices B_l , W_l and Q_l , $l = 1, \dots, L-1$, will not be necessary. Instead, the following result will be used:

$$p(x_l | G_l, x_0, \dots, x_{l-1}) = \frac{N(x_{l-1} | A_l x_l, U_l) N(x_l | \mu \vec{1}_{n_l}, \Sigma_l)}{N(x_{l-1} | \mu \vec{1}_{n_{l-1}}, A_l \Sigma_l A_l' + U_l)}. \quad (13)$$

This factorisation allows efficient simulation of multi-scale random fields and their parameters.

4. Posterior simulation

The posterior distribution is explored by using a Markov chain Monte Carlo procedure (Gamerman 1997), which generates a sample from the joint posterior density of the unknown quantities of the model. From this sample, posterior summaries such as means, medians, variances and credible intervals can be computed through Monte Carlo integration. In Subsection 4.1 we describe the simulation of multi-scale random fields, while in Subsection 4.2 we describe the simulation of the related parameters.

4.1. Simulation of multi-scale random fields

Simulation of multi-scale random fields is necessary for the inference process when part of the field is observed and part is not or when these models are used as priors for hidden processes. In the latter case, if the observed process is linearly related to the hidden process, fast direct simulation can be accomplished using linear algebra results presented by Rue (2001) for PMRFs. Conversely, when the relationship between the hidden process and the observations is nonlinear and non local, as in the inverse problem application presented in Section 5.2, simulation of the multi-scale random field is performed using MCMC-based iterative methods. Here we focus on iterative simulation methods.

As the coarsest level x_0 follows a priori a Markov Random Field, its simulation is straightforward. In order to explore the joint distribution of the random field sites, we use single site Metropolis steps. More specifically, the proposal for x_{0k} is simulated from $x_{0k}^{(prop)} \sim N(x_{0k}^{(old)}, C_{0k}/\phi_{0k})$, where $C_{0k} = \{\tau_0[\alpha_0 + h_{0k}]\}^{-1}$ and ϕ_{0k} is adjusted to yield good acceptance rates. The proposal is accepted with Metropolis acceptance probability (Gamerman 1997).

The simulation of the t -th level ($l = 1, \dots, L-1$) is performed by blocks using a checkerboard pattern, with each block corresponding to one site of the immediate coarser level. Let $B_{l-1,j}$ and $R_{l-1,j}$ be the black and red sites of level l corresponding to site $(l-1, j)$ respectively. As simulation of red and black sites is analogous, here we develop the latter.

The prior full conditional density for black sites at level l corresponding to site $(l-1, k)$ is:

$$\begin{aligned}
p(x_{B_{l-1,k}} | G_{L-1}, x_{\sim B_{l-1,k}}, x_{l-1}) &\propto p(x_l | I_l, x_{l-1}) = \frac{p(x_{l-1} | I_l, x_l) p(x_l | I_l)}{p(x_{l-1} | I_l)} \\
&\propto p(x_{l-1,k} | I_l, x_{D_{l-1,k}}) \prod_{j \in B_{l-1,k}} p(x_{lj} | I_l, x_{N_{lj}}) \\
&\propto \exp \left\{ -\frac{1}{2\delta_l} \left(x_{l-1,k} - m_l^{-1} \sum_{j \in D_{l-1,k}} x_{lj} \right)^2 \right\} \\
&\quad \prod_{j \in B_{l-1,k}} \exp \left\{ -\frac{l}{2C_{lj}} (x_{lj} - a_{lj})^2 \right\}, \quad (14)
\end{aligned}$$

where $C_{lj} = \tau_l(\alpha_l + h_{lj})$ and $a_{lj} = \tau_l C_{lj}(\alpha_l \mu + \sum_{i \in N_{lj}} g_{lij} x_{li})$.

The joint proposal for $x_{B_{l-1,k}}$ is simulated as follows. The proposal for each site in $B_{l-1,k}$ is simulated independently from $x_{lk}^{(prop)} \sim N(x_{lk}^{(old)}, C_{lk}/\phi_{lk})$, where ϕ_{lk} is a tuning parameter adjusted to yield good acceptance rates. The joint proposal for $x_{B_{l-1,k}}$ is accepted with Metropolis acceptance probability.

4.2. Parameters update

Here we address the simulation of the parameters and their prior specification.

We assume that the parameters are independent a priori with joint prior density

$$p(\mu) \prod_{l=0}^{L-1} p(\alpha_l) p(\tau_l) \prod_{l=1}^{L-1} p(\delta_l)$$

More specifically, we assume $p(\mu) = N(\mu|m_\mu, s_\mu^2)$, $p(\delta_l) = Ga(\delta_l|0.5n_{\delta_l}, 0.5n_{\delta_l}s_{\delta_l}^2)$, $p(\alpha_l) \propto \left[\sum_{k=1}^{n_l-1} (\lambda_{lk} + \alpha_l)^{-2} - (n_l - 1)^{-1} \left\{ \sum_{k=1}^{n_l-1} (\lambda_{lk} + \alpha_l)^{-1} \right\}^2 \right]^{0.5}$, the reference prior derived by Ferreira and de Oliveira (2004) for the spatial parameter of a PMRF, and $p(\tau_l) = Ga(\tau_l|0.5n_{\tau_l}, 0.5n_{\tau_l}s_{\tau_l}^2)$.

The mean level μ is simulated with a Metropolis step. The acceptance probability is computed with the help of Equation (13).

The simulation of δ_l , $l = 1, \dots, L - 1$ is performed with a Metropolis-Hastings step $\delta_l^{(prop)} \sim U(\delta_l^{(old)}/\phi_\delta, \delta_l^{(old)}\phi_\delta)$, where $\phi_\delta > 1$ is a tuning parameter adjusted to yield good acceptance.

The parameters α_l and τ_l , $l = 0, \dots, L - 1$, are strongly correlated a posteriori. Thus we jointly generate α_l and τ_l using individual Metropolis-Hastings steps $\alpha_l^{(prop)} \sim U(\alpha_l^{(old)}/\phi_\alpha, \alpha_l^{(old)}\phi_\alpha)$ and $\tau_l^{(prop)} \sim U(\tau_l^{(old)}/\phi_\tau, \tau_l^{(old)}\phi_\tau)$. The joint proposal $(\alpha_l^{(prop)}, \tau_l^{(prop)})$ is accepted with the appropriate Metropolis-Hastings probability.

5. Applications

5.1. A simulated example

Here we present a simulated multi-scale random field with two levels of resolution: a 16 by 16 coarse level and a 32 by 32 fine level. Figure 2 presents the multi-scale field simulated with parameters $\beta_0 = 15.0$, $\mu = 8.5$, $\alpha_0 = 1.0$, $\beta_1 = 100.0$, $\alpha_1 = 0.1$ and $\tau_1 = 0.0001$. Note that the fine level has a local behaviour similar to that of Markov random fields, but yet retains the coarse level induced global behaviour. This global behaviour is analogous to the long-memory type of behaviour observed in multi-scale time series models (Ferreira *et al.* 2003).

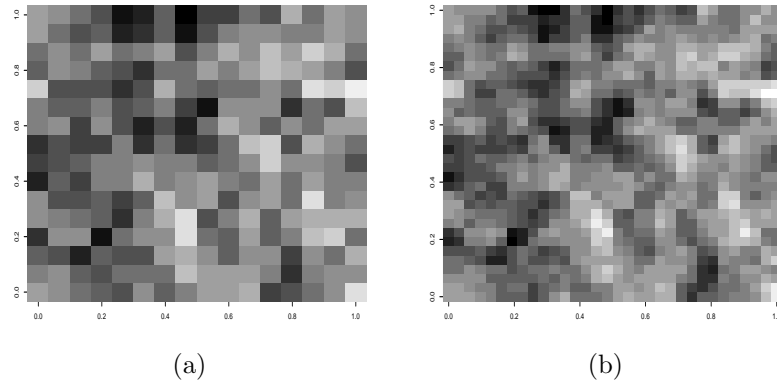


Fig. 2. Simulated multi-scale random field with parameters $\mu = 15.0$, $\tau_0 = 1.0$, $\alpha_0 = 0.5$, $\tau_1 = 8.0$, $\alpha_1 = 1.0$ and $\delta_1 = 0.0016$. (a) Coarse level, (b) Fine level.

The parameters were estimated using the MCMC-based procedure presented in Section 4. Vague priors were used with $m_\mu = 0$, $s_\mu^2 = 10^4$, $n_{\delta_1} = n_{\delta_1}s_{\delta_1}^2 = n_{\tau_0} = n_{\tau_0}s_{\tau_0}^2 = n_{\tau_1} =$

Table 1. Simulated multi-scale random field - Posterior summaries

	True value	Mean	Standard deviation
μ	15.0	15.024	0.053
τ_0	1.0	1.066	0.123
α_0	0.5	0.308	0.197
τ_1	8.0	7.970	0.446
α_1	1.0	0.955	0.123
δ_1	0.0016	0.00161	0.00009

$n_{\tau_1} s_{\tau_1}^2 = 10^{-3}$. Table 1 presents some posterior summaries, showing that the estimation procedure works very well.

5.2. Modelling of 2-D permeability fields

Here we present an application of the multi-scale random field models to the estimation of 2-D permeability fields. This generalises the work of Ferreira et al. (2003) on the multi-scale estimation of 1-D permeability fields. The increase in dimensionality poses new problems for the estimation of permeabilities because with this increase the fluid has many more possible paths to flow through. This increase in the number of possible paths means for example that small areas of very low permeability have little effect on the fluid flow, as the fluid can go around those areas. As a consequence, the problem of permeability field estimation is severely ill-posed in the sense that for given observed dynamic data and given agreement level, there is an infinite number of possible permeability fields. Many of those fields are too rough to be plausible, and proposed approaches generally impose some stochastic regularity constraints on the permeability field. For example, Cunha et al. (1996) and Oliver et al. (1997) assume that the permeability field follows a Gaussian process model, while Lee et al. (2002) assume that the permeabilities follow a Markov random field process. While the Gaussian process model often implies permeability fields that are too smooth, the Markov random field assumption can lead to an undesirably fast decay of the spatial correlation. As an alternative, multi-scale random field models have the local behaviour of Markov random fields but globally emulate long range dependence processes, being well suited to capture global features of the permeability field. Thus, we present here the use of multi-scale random field models as priors for permeability fields.

The setup of the examples in this section is the same as in Lee et al. (2002). In order to estimate the permeability field of an aquifer, the following experiment is performed: water is pumped into the aquifer through one or more injection wells at a fixed rate. The water is extracted by producer wells while keeping constant the pressure at the bottom of each well. After equilibrium is reached, a tracer is injected at the injection well(s) and the concentration of the tracer is measured over time at each of the producer wells. Lee et al. (2002) found that the first time of arrival of the tracer at each well, called the breakthrough time, is practically a sufficient statistic for a unimodal concentration curve. In the petroleum engineering literature, Vasco et al. (1998) and Yoon et al. (1999) have also used the breakthrough times as summaries of the concentration curves. Hence we use the breakthrough times as our data and we denote the observed vector of breakthrough times by $t = t_1, \dots, t_n$. We assume that all other important physical quantities such as initial pressure and porosity are known. Moreover, we assume an incompressible medium and fluid, and an ideal tracer.

The permeability field affects the fluid flow and therefore the dynamic data in a highly non-linear manner. When the physical characteristics of the porous media such as permeability field and porosity are known, physical models such as Darcy's Law and the law of conservation of fluid mass can be used to solve the forward problem, that is, to predict the fluid flow and the dynamic data. These physical models are described by a system of differential equations that can be numerically solved by computer codes called Fluid Flow Simulators (FFS). For references on FFS, see King and Datta-Gupta (1998) and Vasco and Datta-Gupta (1999). Here we use the S3D stream-tube code of King and Datta-Gupta (1998), which is fast enough to make Markov chain Monte Carlo solutions practical. As Ferreira et al. (2003), we assume that a good approximation for the likelihood function at the l -th level of resolution, $l = 1, \dots, L - 1$, is

$$p(t_1, \dots, t_n | x_l, \sigma_l^2) = (2\pi\sigma_l^2)^{-0.5} \exp \left[-0.5\sigma_l^{-2} \sum_{i=1}^n \{ \log(t_i) - \log(f_i^l(x_l)) \}^2 \right] \quad (15)$$

where x_l is the log-permeability field at resolution level l , and $f_i^l(x_l)$ is the i -th output breakthrough time from the FFS running at resolution level l with input permeability field x_l . Assuming (15), we take advantage of running the FFS at different levels of resolution in order to accelerate the procedure for estimation of the permeability field. More specifically, we first run our algorithm at the coarse level and then we run the algorithm at the fine level conditional on the coarse level results. As a consequence, our algorithm converges faster at the fine level when compared with algorithms based on traditional Markov random field models, analogously to results on multi-grid methods (for reference on multi-grid methods, see Briggs *et al.* 2000). Moreover, as the information goes only in the coarse to fine direction, we can parallelise the code by, along the generation of the coarse level realizations, sending the generation of each fine level realization correspondent to an already obtained coarse level realization to a different parallel node. Because there is little communication between the nodes, this parallelisation has excellent granularity and allows very efficient use of parallel machines.

As the dynamic data provide very little information about the smoothness of the field, the parameters of the model have to be specified a priori based on knowledge of specialists. As we model the log-breakthrough times, the variances $\sigma_{\epsilon_0}^2$ and $\sigma_{\epsilon_1}^2$ are related to the amount of allowed relative differences between the observed and the adjusted breakthrough times. Thus these variances have clear interpretation and informative priors can be easily assigned for them. In general, their simulated values are large in the beginning of the MCMC chain and become smaller as the multi-scale random field realizations converge to a draw from the posterior distribution. This behaviour of the traces of $\sigma_{\epsilon_0}^2$ and $\sigma_{\epsilon_1}^2$ leads to an effect similar to simulated annealing for the simulation of the different levels of the field, with higher temperatures in the beginning and cooler temperatures in the end of the chain.

The following subsections present two examples of the application of our multi-scale framework for the estimation of permeability fields. The first example is based on dynamic data simulated from a synthetic Gaussian log-permeability field. The second example refers to data obtained from a real experiment in the Hill Air Force Base in Utah.

5.2.1. Example: Gaussian field

The setup in this example is the inverted 9-spot pattern, that is, we have a square field with a single injection well in the centre of the field and eight producer wells, one at each corner

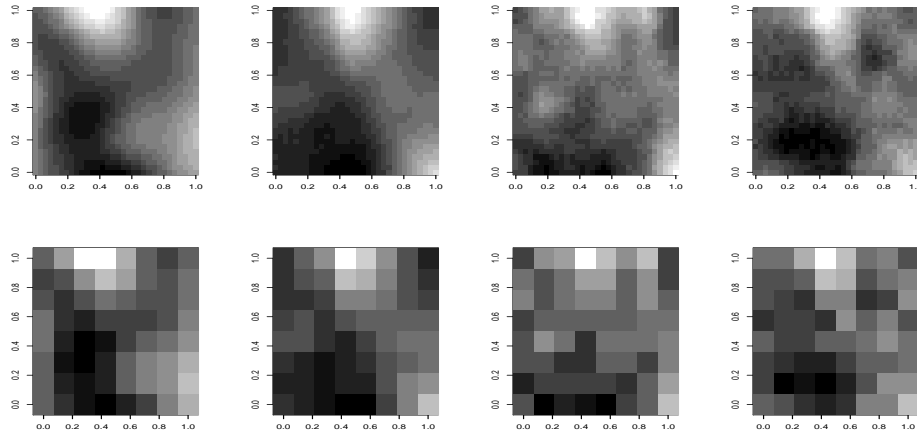


Fig. 3. Gaussian log-permeability field: fine level (first line), coarse level (second line), original field (first column), posterior mean (second column), realizations (third and fourth columns). Darker colours represent higher permeabilities.

and one on the middle of each edge. Thus we have eight breakthrough times to estimate the whole permeability field.

Figure 3 presents a log-permeability field as well as some results of the multi-scale analysis. The log-permeability field was simulated from a Gaussian process with a Gaussian covariogram. In the plots, darker colours indicate higher permeability values.

The original 32 by 32 log-permeability field is shown in the upper left of Figure 3. The original field was used as input for the fluid flow simulator, generating 8 breakthrough times. Using our multi-scale framework, these 8 data points were used to estimate the permeability field. In the lower left of Figure 3, we show the 8×8 coarse version of the original field. In the second column, we show the posterior means of the fine and coarse levels of the log-permeability field.

The multi-scale framework recovers the original field very well. The posterior mean has most of the major features of the original field, such as the regions of lower permeability in the southeast corner and in the centre-north, and the region of higher permeability in the central region and in the centre-south.

The last two columns show some fine and correspondent coarse level realizations of the log-permeability field. The fine level realizations are smoother than the correspondent coarse level realizations. There is high variability between the fine level realizations, nevertheless they show some of the main features of the original permeability field. Moreover, very small regions of low permeability appear in each of the realizations because the fluid can easily go around those regions.

5.2.2. Example: Hill Base example

In this example we use our multi-scale framework to estimate the permeability field of a test site contaminated with several contaminants at Hill Air Force Base in Utah. Multiple

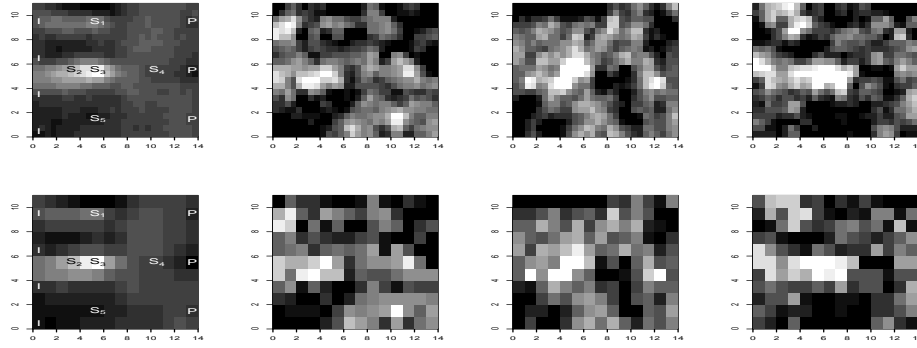


Fig. 4. Hill Base example. Log-permeability field: fine level (first line), coarse level (second line), posterior mean (first column), realizations (second, third and fourth columns). Darker colours represent higher permeabilities.

tracer experiments were run to estimate the physical characteristics in order to support the clean-up of the aquifer (Annable *et al.* 1998, Yoon 2000). We focus on the data from an experiment with a conservative tracer whose interaction with the contaminants is negligible. As we can assume that there is no interaction, the physical process can be well modelled by the S3D stream-tube code of King and Datta-Gupta (1998).

We model the 14 feet by 11 feet test site using two levels of resolution, a coarse resolution on a 14 by 11 grid and a fine resolution on a 28 by 22 grid. The left plots of Figure 4 contain the locations of the wells, where I, P and S stand for injection, production and sampling wells respectively. There are four injection wells along a short edge and three production wells along the opposite edge. The concentration of the tracer is measured only at five sampling wells in the middle of the site at locations S_1 , S_2 , S_3 , S_4 , S_5 with observed breakthrough times 0.42, 0.34, 0.93, 0.52 and 0.17 respectively.

Figure 4 shows the posterior means and realizations of the permeability field at both the coarse and fine level resolutions. The plots in the left column are the posterior means and the plots in the other columns are realizations. The upper and bottom lines correspond respectively to the fine and coarse levels of resolution. As the lower sampling well has the earliest breakthrough time, the lower left region has the highest posterior mean permeability. In addition, as the central sampling well has the latest breakthrough time and a sampling well slightly to its left has the second earliest breakthrough time, there is a sudden reduction in the posterior mean permeability close to the central well. Moreover, the top left region has reasonably high and the right region has intermediate level posterior mean permeability. Note that even though the posterior means at both resolution levels are reasonably smooth the posterior realizations are quite noisy indicating that many different pathways are consistent with the observed breakthrough times. Nevertheless, the posterior realizations clearly indicate that the region close to the central well has the lowest permeability.

6. Discussion

We have introduced a class of multi-scale random field models that can be used to integrate information from data observed at different scales, to induce long memory spatial processes at the finest scale and as a prior for an underlying multi-scale process. We have developed Bayesian inference for this class of models. Finally, we have presented the use of this multi-scale framework to the estimation of permeability fields from production data using fluid flow simulators running at different levels of resolution.

Although we have limited our examples here to the case of regular grids rather than irregular areal data, we note that the underlying framework is sufficiently general to handle either case. Areas warranting investigation include the extension of our multi-scale framework to the case of non-Gaussian errors, as well as to the case of correlated link equation errors. The latter extension would allow the accommodation of systematic bias in the relationship between fine and coarse level processes.

Our class of models can potentially be applied to several other real world problems. For example, our class of models can be used to validate computer models. A computer model is a computational implementation of a solver of a very complicated and specific mathematical problem that possibly involves a set of differential equations. These models are used to simulate the behaviour of a physical system for a given set of conditions, and provide information about a process of interest at a coarse scale. In general, there is additional information on the same process at a fine scale, but this information is sparse. The computer model can be validated by comparison of the fine scale observations with fine scale predictions provided by our multi-scale framework. Another possible application is to model a process observed at a very fine level, but such that there is information from another model for a much cruder level. For example, in meteorology typically there is information from a few stations on the amount of rain at those particular sites, and there is information from satellites and from a numerical model on the amount of rain for very large partitions of an entire region. Our multi-scale model can be used in this context in order to combine the information from the stations with the information from the satellites. The final product would be a very fine map of the amount of rain through the region of interest.

Acknowledgements

This work was partially supported by National Science Foundation grant DMS 9873275. Marco A. R. Ferreira was partially supported by CNPq grant 402010/2003-5 and FUJB grant 10.723-9. Herbert K. H. Lee was partially supported by National Science Foundation grant DMS 0233710.

A. Appendix

A.1. Proof of Theorem 2.1

The assumption that in the revision of beliefs the coarse level x_0 is sufficient for the fine level x_1 means that conditional on x_0 , x_1 is independent of the new information that led to the revision of beliefs about x_0 . That is, $q(x_1|x_0) = p(x_1|x_0)$. This last p.d.f. can be obtained by Bayes Theorem:

$$\begin{aligned} p(x_1|x_0) &\propto p(x_1)p(x_0|x_1) \\ &= N(x_1|\mu\vec{1}_{n_1}, \Sigma_1) N(x_0|A_1x_1, \delta_1I_{n_0}). \end{aligned}$$

Therefore:

$$p(x_1|x_0) = N(x_1|\mu\vec{1}_{n_1} + B_1(x_0 - \mu\vec{1}_{n_0}), \Sigma_1 - B_1W_1B_1')$$

where $B_1 = \Sigma_1A_1'W_1^{-1}$ and $W_1 = A_1\Sigma_1A_1' + \delta_1I_{n_0}$. Thus, the revised joint multi-scale model for coarse and fine levels is

$$q(x_0, x_1) \propto q(x_1|x_0)q(x_0) \\ N(x_1|\mu\vec{1}_{n_1} + B_1(x_0 - \mu\vec{1}_{n_0}), \Sigma_1 - B_1W_1B_1')N(x_0|\mu\vec{1}_{n_0}, Q_0).$$

This concludes part i).

Moreover, the revised marginal model for the fine level can be obtained by the integration of the above joint density with respect to the coarse level, and this is equivalent to the application of Jeffrey's rule of conditioning. Thus:

$$q(x_1) = \int q(x_0, x_1)dx_0 \\ = N(x_1|\mu\vec{1}_{n_1}, Q_1),$$

where $Q_1 = \Sigma_1 - B_1(W_1 - Q_0)B_1'$. This concludes part ii).

A.2. Proof of Theorem 3.1

Let us start proving part iii).

Hypothesis (a) implies that $p(x_0, \dots, x_{l-1}|G_{L-1}) = p(x_0, \dots, x_l|G_l)$, thus

$$p(x_l|G_{L-1}, x_0, \dots, x_{l-1}) = p(x_l|G_l, x_0, \dots, x_{l-1}) \\ = p(x_l|I_l, x_{l-1})$$

by Hypothesis (b). Using Bayes theorem and Equations (8) and (10),

$$p(x_l|I_l, x_{l-1}) \propto p(x_l|I_l)p(x_{l-1}|x_l, I_l) \\ = N(x_l|\mu\vec{1}_{n_l}, \Sigma_l)N(x_{l-1}|A_lx_l, U_l)$$

Thus, through Bayesian linear regression,

$$p(x_l|G_{L-1}, x_0, \dots, x_{l-1}) = N(x_l|\mu\vec{1}_{n_l} + B_l(x_{l-1} - \mu\vec{1}_{n_{l-1}}), \Sigma_l - B_lW_lB_l'),$$

where $B_l = \Sigma_lA_l'W_l^{-1}$ and $W_l = A_l\Sigma_lA_l' + \delta_lI_{n_{l-1}}$. This concludes part iii).

Therefore:

$$p(x_0, \dots, x_{L-1}|G_{L-1}) = p(x_0|G_{L-1}) \prod_{l=1}^{L-1} p(x_l|G_{L-1}, x_0, \dots, x_{l-1}) \\ = p(x_0|I_0) \prod_{l=1}^{L-1} p(x_l|I_l, x_{l-1}) \\ = N(x_0|\mu\vec{1}_{n_0}, Q_0) \\ \prod_{l=1}^{L-1} N(x_l|\mu\vec{1}_{n_l} + B_l(x_{l-1} - \mu\vec{1}_{n_{l-1}}), \Sigma_l - B_lW_lB_l').$$

This concludes part i).

Let us now prove part ii) by induction. By Hypothesis (a) we have that $p(x_0|G_{L-1}) = p(x_0|I_0) = N(x_0|\mu\vec{1}_{n_0}, Q_0)$. Let us assume that $p(x_{l-1}|G_{L-1}) = N(x_{l-1}|\mu\vec{1}_{n_{l-1}}, Q_{l-1})$. Then:

$$\begin{aligned} p(x_l|G_{L-1}) &= \int p(x_l|G_{L-1}, x_{l-1})p(x_{l-1}|G_{L-1})dx_{l-1} \\ &= \int p(x_l|I_l, x_{l-1})p(x_{l-1}|G_{L-1})dx_{l-1} \\ &= \int N(x_l|\mu\vec{1}_{n_l} + B_l(x_{l-1} - \mu\vec{1}_{n_{l-1}}), \Sigma_l - B_lW_lB_l')N(x_{l-1}|\mu\vec{1}_{n_{l-1}}, Q_{l-1})dx_{l-1} \\ &= N(x_l|\mu\vec{1}_{n_l}, Q_l), \end{aligned}$$

where $Q_l = \Sigma_l - B_l(W_l - Q_{l-1})B_l'$. This concludes part ii).

References

- Annable, M. D., P. S. C. Rao, K. Hatfield, W. D. Graham, A. L. Wood, and C. G. Enfield (1998). Partitioning tracers for measuring residual NAPL: Field-scale test results. *Journal of Environmental Engineering* 124, 498–503.
- Besag, J. and C. Kooperberg (1995). On conditional and intrinsic autoregressions. *Biometrika* 82, 733–746.
- Besag, J., J. York, and A. Mollié (1991). Bayesian image restoration, with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics* 43, 1–59.
- Briggs, W. L., V. E. Henson, and S. F. McCormick (2000). *A Multigrid Tutorial, 2nd edition*. Society for Industrial and Applied Mathematics.
- Carlin, B. P. and S. Banerjee (2003). Hierarchical multivariate CAR models for spatio-temporally correlated survival data. In J. Bernardo, M.J.Bayarri, J.O.Berger, A.P.Dawid, D.Heckerman, A.F.M.Smith, and M.West (Eds.), *Bayesian Statistics 7*, pp. 45–63. Oxford University Press.
- Comer, M. and E. Delp (1999). Segmentation of textured images using a multiresolution Gaussian autoregressive model. *IEEE Transactions on Image Processing* 8, 408–420.
- Cunha, L. B., D. S. Oliver, R. A. Redner, and A. C. Reynolds (1996). A hybrid Markov chain Monte Carlo method for generating permeability fields conditioned to multiwell pressure data and prior information. In *SPE Annual Technical Conference and Exhibition*, Denver. SPE.
- Diaconis, P. and S. L. Zabell (1982). Updating subjective probability. *Journal of the American Statistical Association* 77(380), 822–830.
- Ferreira, M. A. R., Z. Bi, M. West, H. Lee, and D. Higdon (2003). Multi-scale modelling of 1-D permeability fields. In J. Bernardo, M.J.Bayarri, J.O.Berger, A.P.Dawid, D.Heckerman, A.F.M.Smith, and M.West (Eds.), *Bayesian Statistics 7*, pp. 519–527. Oxford University Press.

- Ferreira, M. A. R. and V. de Oliveira (2004). Bayesian analysis for a class of Gaussian Markov random fields. Technical Report 177, DME - UFRJ, Downloadable from www.dme.im.ufrj.br/~marco/propermrf.pdf. Submitted.
- Ferreira, M. A. R., M. West, H. Lee, and D. Higdon (2003). Multi-scale and hidden resolution time series models. Technical Report 159, UFRJ - DME, Downloadable from www.dme.im.ufrj.br/~marco/mshrtsm.pdf. Submitted.
- Gamerman, D. (1997). *Markov Chain Monte Carlo*. Chapman Hall.
- Harvey, A. C. (1989). *Forecasting. Structural Time Series Models and the Kalman Filter*. Cambridge University Press.
- Huang, H.-C. and N. Cressie (2001). Multiscale graphical modeling in space: applications to command and control. In M. Moore (Ed.), *Spatial Statistics: Methodological Aspects and Applications*, pp. 83–113. Springer.
- Huang, H.-C., N. Cressie, and J. Gabrosek (2002). Fast, resolution-consistent spatial prediction of global processes from satellite data. *Journal of Computation and Graphical Statistics* 11(1), 63–88.
- Irving, W. W., P. W. Fieguth, and A. S. Willsky (1997). An overlapping tree approach to multiscale stochastic modeling and estimation. *IEEE Transactions on Image Processing* 6, 1517–1529.
- Jeffrey, R. C. (1988). Conditioning, kinematics, and exchangeability. In B. Skyrms and W. L. Harper (Eds.), *Causation, Chance, and Credence*, Volume 1, pp. 221–255. Dordrecht: Kluwer.
- King, M. J. and A. Datta-Gupta (1998). Streamline simulation: a current perspective. *In Situ* 22(1), 91–140.
- Kolaczyk, E. D. (1999). Bayesian multiscale models for Poisson processes. *Journal of the American Statistical Association* 94, 920–933.
- Kolaczyk, E. D. and H. Huang (2001). Multiscale statistical models for hierarchical spatial aggregation. *Geographical Analysis* 33(2), 95–118.
- Lakshmanan, S. and H. Derin (1993). Gaussian Markov random fields at multiple resolutions. In R. Chellapa and A. Jain (Eds.), *Markov Random Fields: Theory and Applications*, pp. 131–157. New York: Academic Press.
- Lee, H., D. Higdon, Z. Bi, M. A. R. Ferreira, and M. West (2002). Markov random field models for high-dimensional parameters in simulations of fluid flow in porous media. *Technometrics* 44, 230–241.
- Nowak, R. D. (1999). Multiscale hidden Markov models for Bayesian image analysis. In P. Müller and B. Vidakovic (Eds.), *Bayesian Inference in Wavelet Based Models*, pp. 243–265. New York: Springer-Verlag.
- Nowak, R. D. and E. D. Kolaczyk (2000, August). A statistical multiscale framework for Poisson inverse problems. *IEEE Transactions on Information Theory* 46(5), 1811–1825.

- Oliver, D. S., L. B. Cunha, and A. C. Reynolds (1997). Markov chain Monte Carlo methods for conditioning a permeability field to pressure data. *Mathematical Geology* 29, 61–91.
- Pizurica, A., W. Philips, I. Lemahieu, and M. Acheroy (2002, May). A joint inter and intrascale statistical model for Bayesian wavelet based image denoising. *IEEE Transactions on Image Processing* 11(5), 545–557.
- Rue, H. (2001). Fast sampling of Gaussian Markov random fields. *Journal of the Royal Statistical Society ser.B* 65, 325–338.
- Saquiib, S. S., C. A. Bouman, and K. Sauer (1996, September). A non-homogeneous MRF model for multiresolution Bayesian estimation. In *IEEE International Conference on Image Processing*, Volume 2, Lausanne, Switzerland, pp. 445–448.
- Skyrms, B. (1980). *Causal Necessity*. New Haven: Yale University Press. Appendix 2.
- Sun, D., R. K. Tsutakawa, and P. L. Speckman (1999). Posterior distribution of hierarchical models using CAR(1) distributions. *Biometrika* 86, 341–350.
- Vasco, D. W. and A. Datta-Gupta (1999). Asymptotic solutions for solute transport: A formalism for tracer tomography. *Water Resources Research* 35(1), 1–16.
- Vasco, D. W., S. Yoon, and A. Datta-Gupta (1998). Integrating dynamic data into high-resolution reservoir models using streamline-based analytic sensitivity coefficients. Society of Petroleum Engineers 1998 Annual Technical Conference, SPE 49002.
- West, M. and J. Harrison (1997). *Bayesian Forecasting and Dynamic Models (2nd Edition)*. Springer-Verlag: New York.
- Willisky, A. S. (2002, August). Multiresolution Markov models for signal and image processing. *Proceedings of the IEEE* 90(8), 1396–1458.
- Yoon, S. (2000). *Dynamic data integration into high resolution reservoir models using streamline-based inversion*. Ph. D. thesis, Texas A&M, Department of Petroleum Engineering.
- Yoon, S., A. H. Malallah, A. Datta-Gupta, D. W. Vasco, and R. A. Behrens (1999). A multiscale approach to production data integration using streamline models. Technical Report 56653, Society of Petroleum Engineers.