

Bayesian Inversion of Well Log Data into Facies Units based on a Spatially Coupled Model

Maria Vigsnes

Master of Science in Physics and Mathematics
Submission date: June 2006
Supervisor: Karl Henning Omre, MATH

Problem Description

The purpose of the work is to define and evaluate a spatial model for facies classification based on real well log data. The well log data should be presented and analyzed, the spatially coupled model for facies classification should be defined, the classification algorithm should be constructed and implemented and the results evaluated.

Assignment given: 27. January 2006
Supervisor: Karl Henning Omre, MATH

Preface

This report is the result of the subject TMA4905 Statistics Thesis in the final semester of my Master of Science degree in Industrial Mathematics. The study is performed at the Department of Mathematical Sciences, Norwegian University of Science and Technology. The work corresponds to one semesters work load.

I would like to thank my supervisor Professor Henning Omre for all support and guidance through the entire semester. I would also like to thank Statoil Research Center at Rotvoll for all involvement, and a special thanks to Kjetil Nordahl and Erlend Gudding for the lessons about the Statfjord Formation and well-logging.

Abstract

Through a study of cored wells from the Statfjord Formation in the Tampen Area, we derive a spatially coupled classification model for facies units. We consider a Bayesian framework for the problem. A likelihood model is defined from the log-response of the formation, where response from neighbour observations are considered. A first order Markov chain prior model is estimated from the cores. From the posterior pdf, the marginal maximum posterior solution can be calculated and simulations can be generated. Since the posterior pdf can be factorized, it can be calculated by a recursive Forward-Backward algorithm for hidden Markov models. The classification model is complex, and if the model assumptions does not coincides with the underlying model, the classification might give poor results due to the large number of estimated model parameters. The results from the classification of a blind well were not as good as we expected, but gave good results for the small classes, compared to a classification model without spatial coupling.

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 2 | Description of the well log data | 2 |
| 2.1 | Core interpretation and classification | 6 |
| 2.2 | Well logs | 7 |
| 3 | Stochastic models and computational algorithm | 11 |
| 3.1 | Likelihood model | 11 |
| 3.2 | Prior model | 12 |
| 3.3 | Posterior model | 16 |
| 3.4 | Forward-Backward algorithm | 16 |
| 3.5 | Bayesian model adapted to the data from the Statfjord Formation | 19 |
| 3.6 | Location wise model | 23 |
| 4 | Parameter estimation | 23 |
| 4.1 | Likelihood parameters estimated based on visual inspection | 24 |
| 4.2 | Likelihood parameters estimated by minimizing the squared error | 31 |
| 4.3 | Prior model parameters | 35 |
| 4.4 | Location wise model parameters | 39 |
| 5 | Implementation and test criteria | 42 |
| 5.1 | Test criteria | 44 |
| 6 | Results and discussions | 46 |
| 6.1 | Facies associations | 47 |
| 6.2 | Lithofacies | 70 |
| 7 | Closing Remarks | 92 |
| | References | 92 |

1 Introduction

In this study we derive a spatially coupled model for classification of facies units in wells based on real well log data from the Statfjord Formation in the Tampen Area. Both well logs and core interpretation is used, in which the latter is considered to be the truth. We look at two training wells, A and B, and one test well, C. The data from the two training wells is used in the derivation of the classification model, and the model is further evaluated on the test well in addition to the training wells. The cores from the wells considered are classified into a hierarchical system of classes of facies units defined by the Statfjord Formation. In this study we derive classification models for both classification into four classes of facies associations, which represents continental to nearshore deposits, and six classes of lithofacies, which is a finer classification based on grain size.

We have chosen a Bayesian framework for the inverse problem. The log-response from the formation of the wells considered in the likelihood model, and the spatial coupling in the formation in the Markov chain prior model. The model parameters are estimated based on a statistical approach, and we consider two methods of estimating the likelihood parameters. The first method is based on a visual study of the logs, while the second is a least-squares minimization. From Bayes rule, the prior and likelihood model gives us the posterior pdf. Given the observed log data, the posterior pdf gives the probabilities of all possible combinations of facies units along the profile. In order to compute the posterior pdf, we consider the exact and efficient Forward-Backward algorithm for hidden Markov models given in Chib (1996) and Scott (2002), where we extend the algorithm to include neighbours in the likelihood model. From the calculated posterior pdf the marginal maximum posterior solution can be calculated and simulated profiles can be generated.

An extensive search for similar articles has been done, and we are surprised that no statistical study of well logs for use in an inversion problem was found.

In Avseth and Mukerji (2002) seismic lithofacies are classified from well logs using statistical rock physics. Three different classification methods are considered, Mahalanobis discriminant analysis, a pdf classification and a neural network classification. All classifications are based on a study of gamma-ray logs and seismic p-wave velocity of the seismic lithofacies. Prior to the classification, the data have been filtered in order to remove noise. Inversion is not considered, and therefore the spatial coupling of the facies and the spatial response from the formation is not utilized.

In Ali and Lall (1998) a methodology for simulation of core profiles using a continuous homogeneous semi-Markov model is presented. The strategy is developed by using a transition intensity matrix to determine the transitions between states, and bootstrapping to determine the layer thickness corresponding to the new state. No conditioning upon well logs is considered, and neither is an inversion problem. The semi-Markov model could have been used as a prior model for the inversion problem in this study, if we formulate the prior in the form of a one-step transition probability matrix. However, since the layers of the wells are thick, this demands extremely large state spaces and therefore computing resources. In our study, we will only consider a simple first order Markov chain as prior model, since the focus will be on deriving a likelihood model adapted to the well log data.

The core interpretation and the well logs are presented in section 2, and the derivation of the stochastic model and computational algorithm are given in section 3. Further, the parameter estimation is given in section 4, followed by some information about implementation together with the test criteria in section 5. In section 6 the results with discussions are presented, and at last, closing remarks are given in section 7.

2 Description of the well log data

The well log data considered in this study are from the Statfjord Formation in the Tampen area. The three wells, A, B and C, are chosen because they are located relatively close to each other, and therefore should have some of the same qualities. The wells are cored, which means that they are drilled with a hollow, cylindrical drill, in such way that the cores can be removed from the well. Further, the cores are analyzed and the facies units are classified.

After the cores are removed from the wells, logging tools are lowered down to the bottom of the wells fastened on wires. The tools are raised at a constant speed to the top of the wells, while different information of the formation is recorded at regular intervals. Log data is normally recorded at every half feet, but after an interpolation the data is represented at every fifth centimeters. The extra data is a disadvantage for us, since redundant data make our computations unnecessary complicated. This is solved by only using data at every third realization, which means at every fifteenth centimeters. The log data and the core interpretation from the three wells are given in Figure 1, 2 and 3. The core interpretation is missing in some locations, which are hatched in the figures.

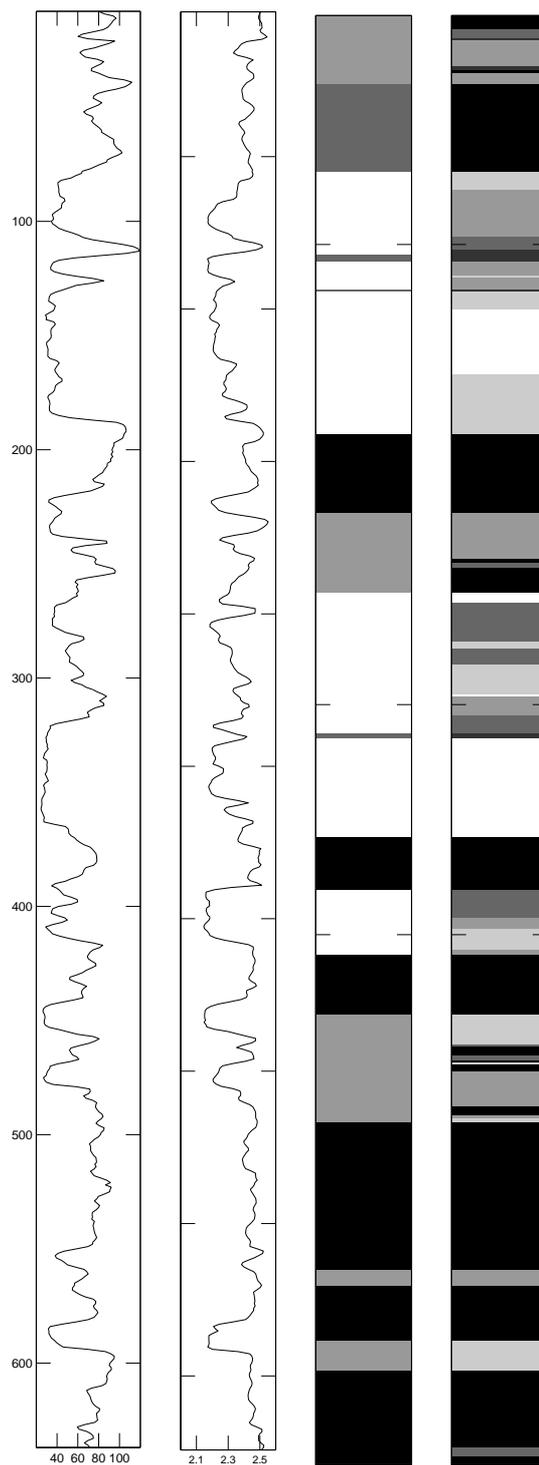


Figure 1: The first training well, A. From left is the gamma-ray log, the density log and core interpretations of facies associations and lithofacies respectively. The gray scale levels for the classes are given in Figure 4 and 6.

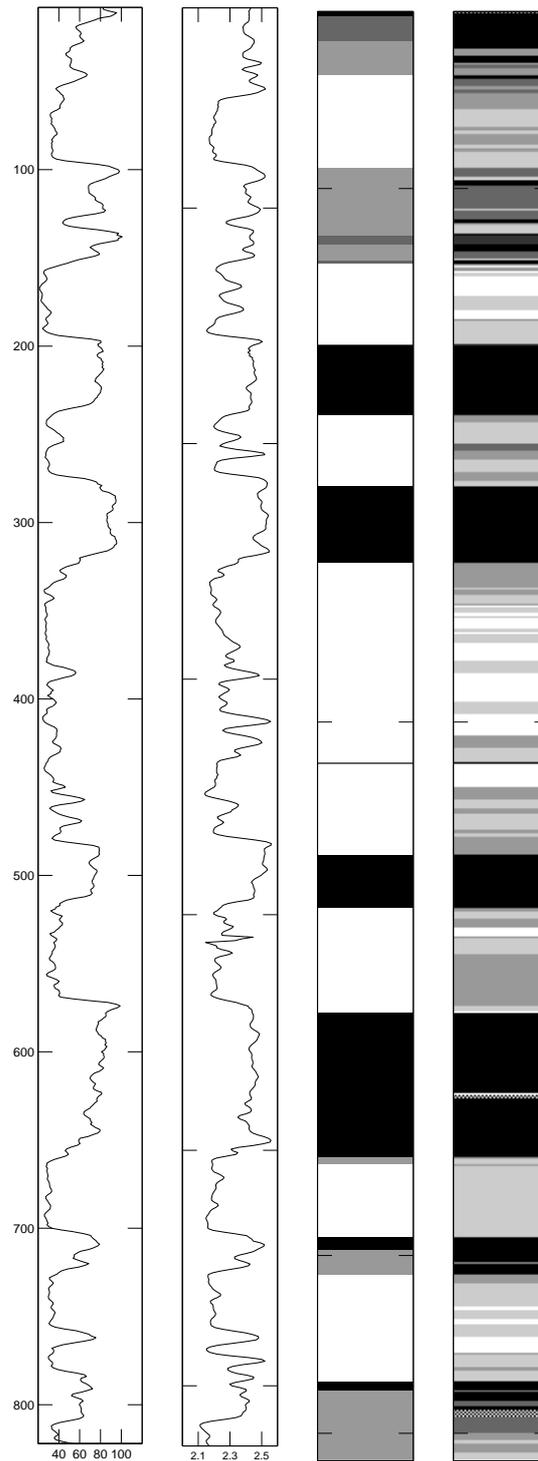


Figure 2: The second training well, B. From left is the gamma-ray log, the density log and core interpretations of facies associations and lithofacies respectively. The gray scale levels for the classes are given in Figure 4 and 6. In the locations that are hatched, the interpretations of lithofacies are missing.

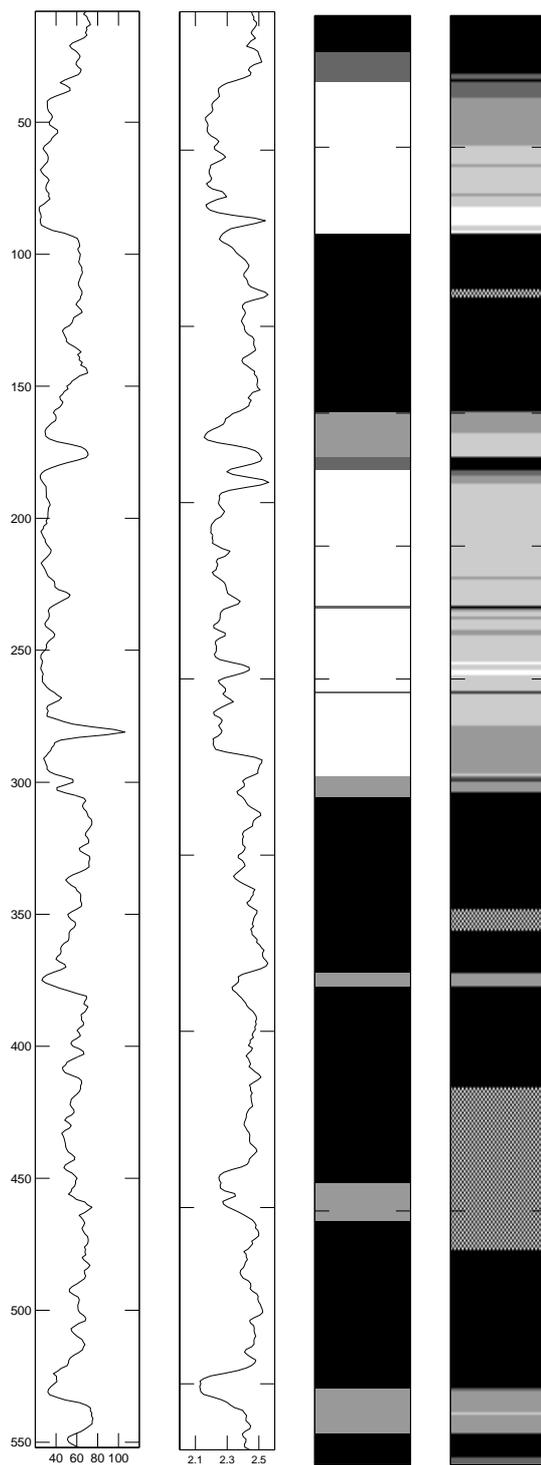


Figure 3: The test well, C. From left is the gamma-ray log, the density log and core interpretations of facies associations and lithofacies respectively. The gray scale levels for the classes are given in Figure 4 and 6. In the locations that are hatched, the interpretations of lithofacies are missing.

| Facies associations name | Class | Color |
|-------------------------------|--------------|------------|
| Multi-storey fluvial channels | π_1^{fa} | |
| Single-storey fluvial channel | | |
| Crevasse splay | π_2^{fa} | Light Gray |
| Crevasse channels | | |
| Floodplain fines undiff. | π_3^{fa} | Dark Gray |
| Channel abandonment | | |
| Lake/Pond | | |
| Bayfill | | |
| Mature, dry paleosol | π_4^{fa} | Black |
| Wet paleosol | | |

Figure 4: The descriptions, classes and gray scale levels of the facies associations, π^{fa} .

The locations for the data points are ordered from the top of the well and down. The three wells have different length and depth, but for notational convenience we denote the locations by $t = 1, 2, \dots, n$.

2.1 Core interpretation and classification

The formation of the wells can be described by the Statfjord Formation, which partition the formation in three hierarchic levels. First, the formation can be partitioned into 8 zones called picks, which are numbered from the bottom to the top of the formation. The picks can be separated into two super groups, pick 1-5 and pick 6-8, where the picks in each of the groups have some of the same qualities. We will consider pick 2-4 in this study.

Each of the picks can be partitioned into layers of facies associations at random order. There are 10 different facies associations, and not all of them occur in every pick. The facies associations describe the continental to nearshore deposits. The facies associations can be grouped into a two-, three- or four-class system, and in this study we will consider the four-class system. Figure 4 shows the descriptions of the facies associations, π^{fa} , where the gray scale levels correspond to the four classes denoted by $\pi_1^{fa}, \dots, \pi_4^{fa}$. Figure 5 shows the proportions of the classes in the three wells.

The facies associations can again be partitioned into layers of lithofacies. The 21 different types of lithofacies describes structure and contents of different geological aspects as sand, mud, clay, heterolith and others, and are the finest classifications of the well. The

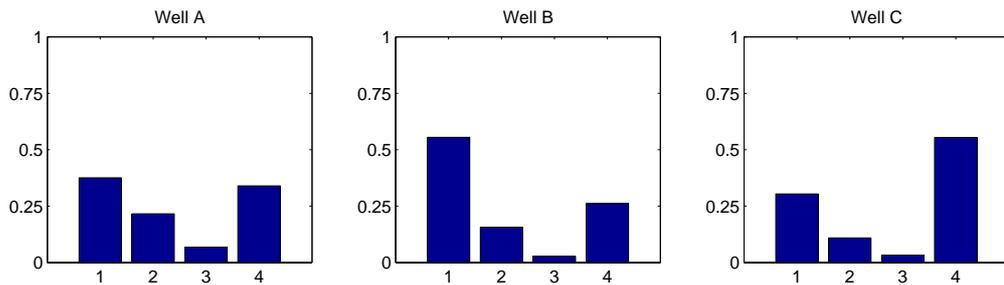


Figure 5: The proportion of the classes of facies associations, π^{fa} , in the three wells.

lithofacies can be grouped into either two or six groups, and in this study we will consider the six groups denoted by $\pi_1^{lf}, \dots, \pi_6^{lf}$, which partition the lithofacies according to their grain size. The descriptions and corresponding gray scale levels we can see in Figure 6, and the proportion of the classes in the three wells in Figure 7.

In this study we consider the classification into both facies associations and lithofacies. The cores, which we denote by $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$, take values from both π^{fa} and π^{lf} .

2.2 Well logs

We consider two different logs recorded in the three wells, which is the gammaray log, l^{gr} , and the density log, l^d .

The gammaray log, l^{gr} , is a passive log. The logging tool for the gammaray log consists of one receiver that measures the natural radioactivity of the formation. The response is a function of the concentration per weight of the radioactive mineral and the density of the rock.

The density log, l^d , on the other hand, is an active log. The logging tool for the density log consists of one source and two detectors. The source emits gamma rays and the density of the formation is measured by the attenuation of the gamma rays in the detectors. The gamma rays collide with electrons in the formation, which make them loose energy and change direction. The intensity that is measured in the detectors represents the electron density of the formation, which is correlated with the density of the formation.

In order to see how the response from the three wells behave, the density log versus the gammaray log is plotted in Figure 8. As we can see, the range of the response

| Lithofacies descriptions | Class | Color |
|--|--------------|-------|
| Granule-bearing or conglomeratic, poorly sorted SST, w/ unspecified parallel stratification | π_1^{lf} | |
| Granule-bearing or conglomeratic, poorly sorted SST, w/ cross bedding | | |
| Granule-bearing or conglomeratic, poorly sorted SST, apparently massive | π_2^{lf} | |
| Coarse- to very coarse-grained, +- clean SST, w/ unspecified parallel stratification | | |
| Coarse- to very coarse-grained, +- clean SST, displaying cross bedding | | |
| Coarse- to very coarse-grained, +- clean SST, apparently massive | π_3^{lf} | |
| Dominantly medium-grained SST, w/unspecified parallel stratification | | |
| Dominantly medium-grained SST, w/ ripples | | |
| Dominantly medium-grained SST, w/ cross bedding | | |
| Dominantly medium-grained SST, apparently massive | π_4^{lf} | |
| Dominantly fine- to very fine-grained SST, w/ unspecified parallel stratification | | |
| Dominantly fine- to very fine-grained SST, w/ ripples | | |
| Dominantly fine- to very fine-grained SST, w/ cross bedding | | |
| Dominantly fine- to very fine-grained SST, apparently massive | π_5^{lf} | |
| Sand-dominated heterolith, laminated and/or lenticular- to flaser-bedded | | |
| Mud-dominated heterolith, laminated and/or lenticular- to flaser bedded | | |
| massive-looking or very faintly silt/sand-stripped claystone or mudstone, may be sandy in places | π_6^{lf} | |
| sand-dominated paleosol, may display faint ripples or other primary lamination | | |
| silt-dominated paleosol, may display faint ripples or other primary lamination | | |
| clay-dominated paleosol, may display silt-stripes and/or other primary lamination | | |
| massive calcretes and silicified layers interbedded with clay- and mudstone | | |

Figure 6: The descriptions, classes and gray scale levels of the lithofacies, π^{lf} .

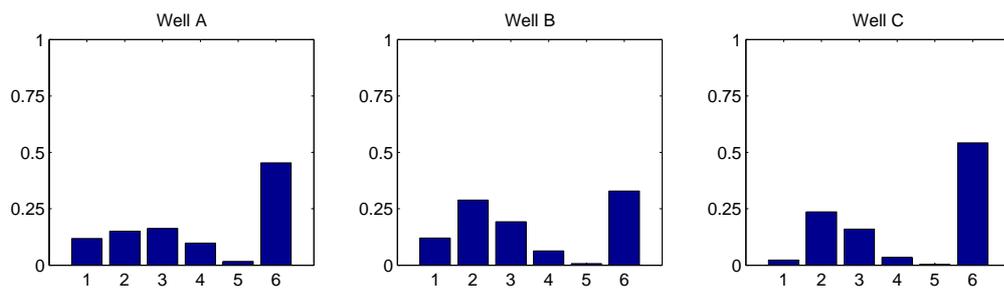


Figure 7: The proportion of the classes of lithofacies, π^{lf} , in the three wells.

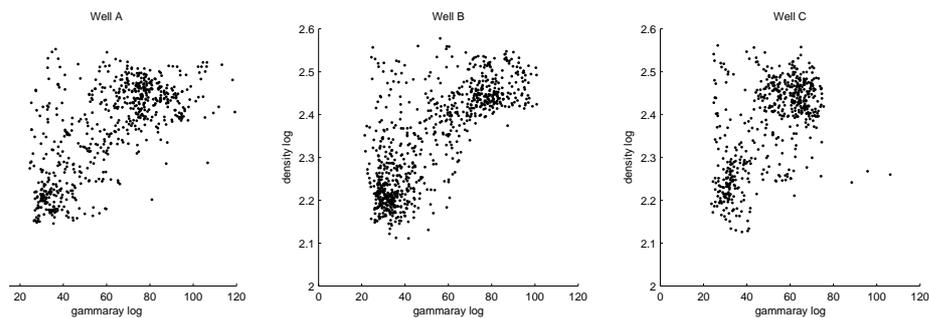


Figure 8: The gamma ray log, l^{gr} , versus the density log, l^d , for the three wells.

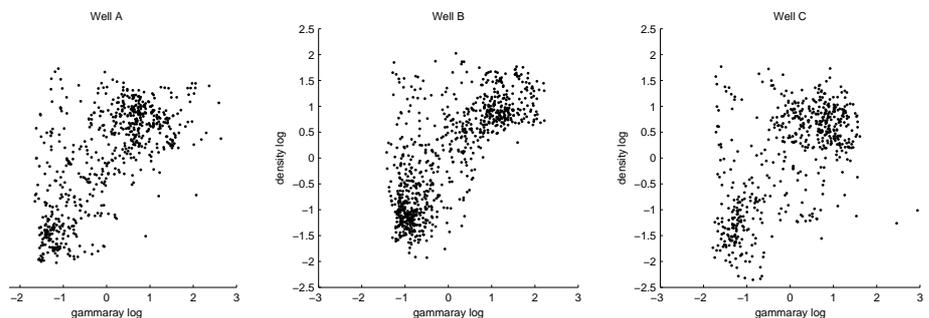


Figure 9: The gamma ray log, l^{gr} , versus the density log, l^d , for the three wells with standardized data.

varies between the wells, and especially for the gammaray log in the test well, C. If we disregard the few outliers at the right in the plot of well C, we notice that the response from the gammaray log lies in the interval $[20, 80]$, which is narrow compared to the interval $[20, 120]$ for well A and B.

If we want to define a classification model for well C, based on well A and B, the log data from all wells need to have similar levels of response. Since the range is different in the wells, the log data need to be standardized. We give the data mean 0 and variance 1. Figure 9 shows the standardized data, and as we can see, the range is now more similar in the three wells.

In order to see how the response from the different classes in the formation behave, the realizations in the plots are depicted by symbols, as we can see from Figure 10 and 11. In Figure 10 the symbols separate the realizations into facies associations, π^{fa} , and in Figure 11 into lithofacies, π^{lf} . If all realizations from each class are gathered into separated clusters, the classification would be easy. Unfortunately, the data are more mixed, because of variance and possible shoulder effects from neighbour observations.

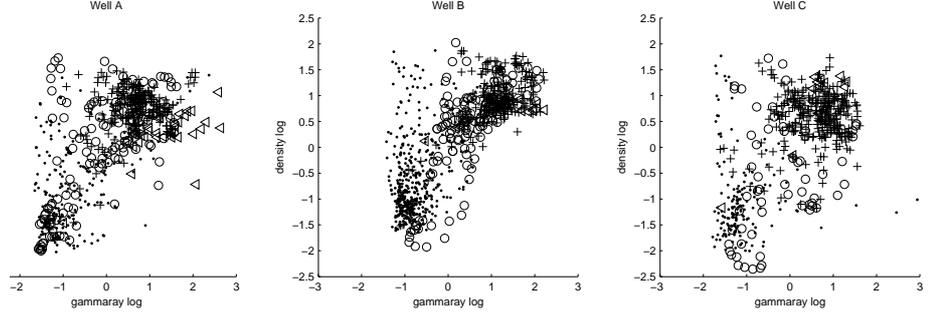


Figure 10: The gamma ray log, l^{gr} , versus the density log, l^d , for the three wells with standardized data classified into facies associations. The dots (\cdot) are class π_1^{fa} , the circles (\circ) are class π_2^{fa} , the plus signs ($+$) are class π_3^{fa} and the triangles (\triangle) are class π_4^{fa} .

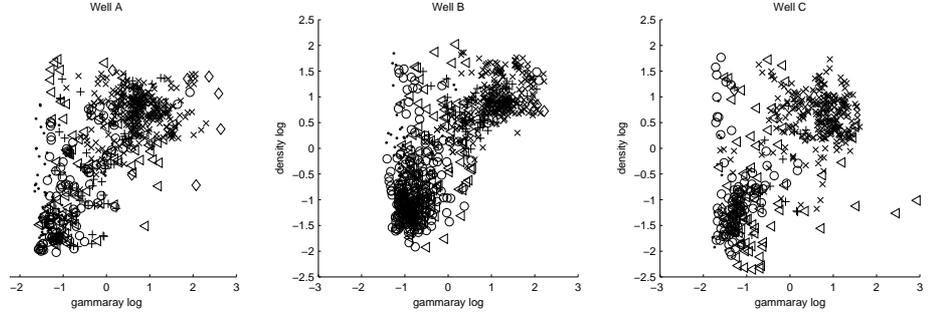


Figure 11: The gamma ray log, l^{gr} , versus the density log, l^d , for the three wells with standardized data, classified into lithofacies. The dots (\cdot) are class π_1^{lf} , the circles (\circ) are class π_2^{lf} , the plus signs ($+$) are class π_3^{lf} , the triangles (\triangle) are class π_4^{lf} , the diamonds (\diamond) are class π_5^{lf} and the crosses (\times) are class π_6^{lf} .

In Figure 10 we can see that the dots (\cdot) and the plus signs ($+$) are clustering separately, which indicates that the probability of misclassifying locations from class π_1^{fa} as class π_3^{fa} and vice versa is small. The same thing we can see for the circles (\circ) and crosses (\times), class π_2^{lf} and class π_6^{lf} , in Figure 11.

Apart from this, the classes of both facies associations and lithofacies are relatively mixed into each other, which means that we are not able to classify the wells from these plots alone, we will need a spatial classification model.

3 Stochastic models and computational algorithm

In this section we derive a general Bayesian model for an unknown core profile, $\mathbf{x} = (x_1, x_2, \dots, x_n)$, and corresponding log data, $\mathbf{d} = (d_1, d_2, \dots, d_n)$, followed by definition of an efficient algorithm for calculating the posterior pdf. Further, a Bayesian model adapted to the data from the Statfjord Formation at the Tampen Area is derived. Finally, as a comparison, a location wise model is defined.

Given the observed log data, $\mathbf{d} = (d_1, d_2, \dots, d_n)$, we want to classify the corresponding core profile. From a Bayesian setting, the unknown core profile, $\mathbf{x} = (x_1, x_2, \dots, x_n)$, has a prior pdf, $p(\mathbf{x})$, which is based on the assumptions we make for the model. The likelihood function, $p(\mathbf{d}|\mathbf{x})$, gives the probabilities of the log data, given all possible combinations of classes along the profile. The posterior pdf, $p(\mathbf{x}|\mathbf{d})$, gives the probability of all possible combinations of classes along the profile given \mathbf{d} , and from Bayes rule we have

$$p(\mathbf{x}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{d})},$$

where $p(\mathbf{d})$ is a constant.

3.1 Likelihood model

If the observed log data depends on the unknown profile as we can see from Figure 12, we can express the log data at location t as

$$d_t = g(x_t) + \epsilon_t \tag{1}$$

where $g(\cdot)$ is some known function and ϵ_t is the error at location t . As mentioned in the previous section, the likelihood function gives the probability of the log data given all possible combinations of classes along the profile. We assume that the error in expression (1) is independent, identical distributed as

$$\epsilon_t \sim \mathcal{N}(0, \sigma^2),$$

thus, only white noise is considered. This gives the distribution for the likelihood function

$$p(\mathbf{d}|\mathbf{x}) \sim \mathcal{N}_n(\mathbf{g}(\mathbf{x}), \sigma^2 \mathbf{I}),$$

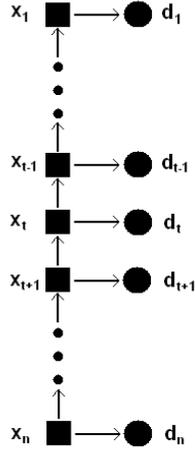


Figure 12: The graph illustrates the relation between the locations of the profile, x_1, \dots, x_n , and the observed log data, d_1, \dots, d_n .

where the mean is a vector, $\mathbf{g}(\mathbf{x}) = (g(x_1), g(x_2), \dots, g(x_n))^T$, and \mathbf{I} is the identity matrix of dimension $n \times n$. This again gives

$$p(\mathbf{d}|\mathbf{x}) = (2\pi)^{-\frac{n}{2}} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^n [d_t - g(x_t)]^2 \right\},$$

which can be written in the form

$$p(\mathbf{d}|\mathbf{x}) = \prod_{t=1}^n p(d_t|x_t), \quad (2)$$

where $p(d_t|x_t) \sim \mathcal{N}(g(x_t), \sigma^2)$.

3.2 Prior model

In this section we first define a Markov chain, and then introduce the Markov chain prior model.

Markov chains

We look at a sequence of random variables, $\mathbf{x} = (x_1, x_2, \dots, x_n)$, where each variable, x_t , take values from a finite set of classes, $\Psi = \{\psi_1, \psi_2, \dots, \psi_m\}$. This is a stochastic

process where the index t denotes the location at which x_t occurs. A good reference for stochastic processes is Ross (2003).

The stochastic process is said to be a Markov chain if it fulfils the following condition

$$p(x_t|x_1, x_2, \dots, x_{t-1}) = p(x_t|x_{t-1}) \quad \text{for} \quad t = 2, 3, \dots, n. \quad (3)$$

In words, the conditional distribution of x_t given x_1, x_2, \dots, x_{t-1} , is conditionally independent of x_1, x_2, \dots, x_{t-2} , given x_{t-1} . This is called the Markov property, and is illustrated in Figure 12.

We define the probabilities for making a transition from class ψ_i to class ψ_j for all $\psi_i, \psi_j \in \Psi$, that is

$$q(j|i) = p(x_t = \psi_j | x_{t-1} = \psi_i).$$

The probabilities $q(j|i)$ are called one-step transition probabilities, and since the number of classes is finite, they define the one-step transition probability matrix,

$$\mathbf{Q} = \begin{pmatrix} q(1|1) & q(2|1) & \cdots & q(m|1) \\ q(1|2) & q(2|2) & \cdots & q(m|2) \\ \vdots & \vdots & \ddots & \vdots \\ q(1|m) & q(2|m) & \cdots & q(m|m) \end{pmatrix}, \quad (4)$$

where the rows give the class at location $t - 1$, and the columns the class at location t . Since the process must make a transition into some class, we have that

$$\sum_{j=1}^m q(j|i) = 1.$$

The transition probabilities are independent of t , that is

$$p(x_t|x_{t-1}) = p(x_{t+u}|x_{t+u-1}) \quad \text{for} \quad t = 2, 3, \dots, n,$$

for arbitrary $u \in \mathcal{N}_+$. This means that the conditional probability of x_t given x_{t-1} is the same for every t . A Markov chain with this property is said to be homogeneous.

The initial distribution of a Markov chain is the probability at location $t = 1$, $p(x_1)$. The probability distribution for a homogeneous Markov chain is fully described by the initial distribution and the transition probability matrix. The probabilities at location

t can be calculated by

$$p(x_t) = p(x_1)\mathbf{Q}^{t-1},$$

for $t = 2, 3, \dots, n$. In words, the distribution at location t is the product of the initial distribution and the transition probability matrix powered to $t - 1$.

The joint distribution for the Markov chain is

$$\begin{aligned} p(\mathbf{x}) &= p(x_1, x_2, \dots, x_n) \\ &= p(x_n|x_1, \dots, x_{n-1})p(x_{n-1}|x_1, \dots, x_{n-2}) \cdots p(x_2|x_1)p(x_1) \\ &= p(x_n|x_{n-1})p(x_{n-1}|x_{n-2}) \cdots p(x_2|x_1)p(x_1) \\ &= \prod_{t=1}^n p(x_t|x_{t-1}), \end{aligned}$$

where $p(x_1|x_0) = p(x_1)$ is the initial distribution, and the transition probabilities $p(x_t|x_{t-1})$ for $t = 2, 3, \dots, n$ are given by the transition probability matrix, \mathbf{Q} , in expression (4).

We define the ν -step transition probabilities

$$q^\nu(j|i) = p(x_{t+\nu} = \psi_j | x_t = \psi_i),$$

which is the probability that a process at location $t + \nu$ is class ψ_j , when location t is class ψ_i , for all $\psi_i, \psi_j \in \Psi$.

Class ψ_j is said to be accessible from class ψ_i if $q^\nu(j|i) > 0$ for some $\nu \in \mathcal{N}_+$, which means that it is possible to reach class ψ_j from class ψ_i . Two classes ψ_i and ψ_j that are accessible to each other are said to communicate. A Markov chain, $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \Psi^n$, with transition probability matrix \mathbf{Q} , is irreducible if all the classes in Ψ communicate with each other.

Class ψ_i is said to have period d if $q^\nu(i|i) = 0$ whenever ν is not divisible by d , and d is the largest integer with this property. A class with period 1 is said to be aperiodic.

For an irreducible and aperiodic Markov chain, $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \Psi^n$, with transition probability matrix, \mathbf{Q} , it can be shown that there exists a unique stationary distribution, $p_s(x)$. The stationary distribution satisfies

$$p_s(x_t = \psi_j) = \sum_{i=1}^m p(x_t = \psi_j | x_{t-1} = \psi_i) p_s(x_{t-1} = \psi_i),$$

and

$$\sum_{j=1}^m p_s(x_t = \psi_j) = 1,$$

for $t = 2, 3, \dots, n$. If the Markov chain reaches equilibrium, the marginal distribution for all t will converge to the stationary distribution, regardless of what the initial distribution is. That is $p_s(x = \psi_j) = \lim_{\nu \rightarrow \infty} p(x_{t+\nu} = \psi_j | x_t = \psi_i)$ for all $\psi_i, \psi_j \in \Psi$.

If we let the initial distribution be the stationary distribution, we make sure that the Markov chain will always stay in the stationary distribution. The joint distribution for the Markov chain is fully described by the transition probability matrix, \mathbf{Q} , from which the stationary distribution, $p_s(x)$, can be determined.

Markov chain prior model

We want to classify the unknown profile, $\mathbf{x} = (x_1, x_2, \dots, x_n)$, where $x_t \in \Psi = \{\psi_1, \psi_2, \dots, \psi_m\}$ for every t . From Figure 12 we notice that x_t is dependent of $x_{t+1}, x_{t+2}, \dots, x_n$. We therefore need to define the Markov property in expression (3) in the opposite direction, and we have that

$$\begin{aligned} p(x_t | x_{t+1}, \dots, x_n) &= \frac{p(x_t, \dots, x_n)}{p(x_{t+1}, \dots, x_n)} \\ &= \frac{p(x_n | x_{n-1}) p(x_{n-1} | x_{n-2}) \cdots p(x_{t+1} | x_t) p(x_t)}{p(x_n | x_{n-1}) p(x_{n-1} | x_{n-2}) \cdots p(x_{t+2} | x_{t+1}) p(x_{t+1})} \\ &= \frac{p(x_{t+1} | x_t) p(x_t)}{p(x_{t+1})} \\ &= p(x_t | x_{t+1}). \end{aligned} \quad (5)$$

We now consider the unknown profile a stationary Markov chain defined upwards, given by

$$\begin{aligned} p(\mathbf{x}) &= p(x_1, x_2, \dots, x_n) \\ &= p(x_1 | x_2, \dots, x_n) p(x_2 | x_3, \dots, x_n) \cdots p(x_{n-1} | x_n) p(x_n) \\ &= p(x_1 | x_2) p(x_2 | x_3) \cdots p(x_{n-1} | x_n) p(x_n) \\ &= \prod_{t=1}^n p(x_t | x_{t+1}), \end{aligned} \quad (6)$$

where we have chosen the stationary distribution as initial distribution, $p(x_n | x_{n+1}) = p_s(x_n)$. The transition probabilities $p(x_t | x_{t+1})$ can be calculated from the transition

probability matrix \mathbf{Q} in expression (4), that is

$$p(x_t|x_{t+1}) = p(x_{t+1}|x_t) \frac{p_s(x_t)}{p_s(x_{t+1})},$$

where $p(x_{t+1}|x_t)$ are the transition probabilities from \mathbf{Q} , and $p_s(x_t)$ and $p_s(x_{t+1})$ are given by the stationary distribution.

3.3 Posterior model

The likelihood function in expression (2) and the prior pdf in expression (6) give the posterior pdf

$$\begin{aligned} p(\mathbf{x}|\mathbf{d}) &= \text{const} \times p(\mathbf{d}|\mathbf{x})p(\mathbf{x}) \\ &= \text{const} \times \prod_{t=1}^n p(d_t|x_t)p(x_t|x_{t+1}), \end{aligned} \quad (7)$$

where const is a normalizing constant and $p(x_n|x_{n+1}) = p_s(x_n)$.

The model in this study is a hidden Markov model (HMM). A HMM is a process where the behavior of the observed data, \mathbf{d} , can be expressed conditionally upon an unobserved, hidden Markov chain, $\mathbf{x} = (x_1, x_2, \dots, x_n)$. This phenomenon we capture in the likelihood function in expression (2). The Markov chain in the HMM has a stationary transition probability matrix and the initial distribution is often taken to be the stationary distribution for the chain, as for the prior pdf in expression (6).

The posterior pdf of the HMM given in expression (7) is a nonhomogeneous Markov chain, and give the probabilities of all possible outcomes of the unobserved profile, $\mathbf{x} = (x_1, \dots, x_n)$. The normalizing constant is therefore very difficult to determine since it requires a sum over all possible combinations of classes along the profile. Since the posterior pdf of the HMM is factorized, it can be calculated by the recursive Forward-Backward algorithm for HMM (Chib, 1996; Scott, 2002). Given the observed data, $\mathbf{d} = (d_1, d_2, \dots, d_n)$, the algorithm computes the posterior pdf, $p(\mathbf{x}|\mathbf{d})$, and its marginal posterior pdfs, $p(x_t|\mathbf{d})$, for all t .

3.4 Forward-Backward algorithm

The Forward-Backward recursion operates as follows. As the forward recursion moves upward through the profile it calculates the pdf at each step based on the data up to the actual location and the pdf in the most recently step. The backward recursion goes down

through the profile again, updating the pdf when data from above have been collected.

Forward recursion

The forward algorithm calculates the forward pdf recursively step by step, which means that in each step it calculates the joint pdf for x_{t+1} and x_t , conditioned upon the data in the locations n, \dots, t . The joint pdf can be partitioned like this

$$\begin{aligned} p(x_{t+1}, x_t | d_n, \dots, d_t) &= \text{const} \times p(d_t | x_{t+1}, x_t, d_n, \dots, d_{t+1}) p(x_{t+1}, x_t | d_n, \dots, d_{t+1}) \\ &= \text{const} \times p(d_t | x_t) p(x_{t+1}, x_t | d_n, \dots, d_{t+1}) \\ &= \text{const} \times p(d_t | x_t) p(x_t | x_{t+1}, d_n, \dots, d_{t+1}) p(x_{t+1} | d_n, \dots, d_{t+1}) \\ &= \text{const} \times p(d_t | x_t) p(x_t | x_{t+1}) p(x_{t+1} | d_n, \dots, d_{t+1}). \end{aligned}$$

The probability $p(d_t | x_t)$ is the likelihood of the data d_t given x_t , and $p(x_t | x_{t+1})$ is the transition probability from the prior Markov chain. Further, $p(x_t | d_n, \dots, d_t)$ is the marginal forward pdf at location t . The transition from line 1 to line 2 is valid because d_t is conditionally independent of x_{t+1} and d_n, \dots, d_{t+1} , given x_t . The transition from line 3 to 4 is also valid because x_t is conditionally independent of d_n, \dots, d_{t+1} , given x_{t+1} . The normalizing constant is given by

$$\text{const} = \left(\sum_{x_{t+1}} \sum_{x_t} p(x_{t+1}, x_t | d_n, \dots, d_t) \right)^{-1}.$$

The marginal forward probabilities are calculated from

$$p(x_t | d_n, \dots, d_t) = \sum_{x_{t+1}} p(x_{t+1}, x_t | d_n, \dots, d_t),$$

with the initial distribution

$$p(x_n | d_n) = \text{const} \times p_s(x_n) p(d_n | x_n),$$

where const is calculated from $\sum_{x_n} p(x_n | d_n) = 1$.

The forward pdf for the profile is a Markov chain because the distribution of x_t conditioned upon (x_n, \dots, x_{t+1}) and (d_n, \dots, d_t) is independent of (x_n, \dots, x_{t+2}) and (d_n, \dots, d_{t+1}) . The Markov chain is non-homogeneous because the transition probabilities are not con-

stant as t change.

Backward recursion

The backward recursion calculates the joint backward pdf for x_{t-1} and x_t given the entire set of observed data (d_1, d_2, \dots, d_n) . The joint pdf for x_{t-1} and x_t is given by

$$\begin{aligned}
 p(x_{t-1}, x_t | d_1, \dots, d_n) &= p(x_t | x_{t-1}, d_1, \dots, d_n) p(x_{t-1} | d_1, \dots, d_n) \\
 &= p(x_t | x_{t-1}, d_t, \dots, d_n) p(x_{t-1} | d_1, \dots, d_n) \\
 &= p(x_t | x_{t-1}, d_{t-1}, \dots, d_n) p(x_{t-1} | d_1, \dots, d_n) \\
 &= \frac{p(x_t, x_{t-1} | d_{t-1}, \dots, d_n)}{p(x_{t-1} | d_{t-1}, \dots, d_n)} p(x_{t-1} | d_1, \dots, d_n) \\
 &= \frac{p(x_t, x_{t-1} | d_n, \dots, d_{t-1})}{p(x_{t-1} | d_n, \dots, d_{t-1})} p(x_{t-1} | d_1, \dots, d_n),
 \end{aligned} \tag{8}$$

where $p(x_t, x_{t-1} | d_n, \dots, d_{t-1})$ and $p(x_{t-1} | d_n, \dots, d_{t-1})$ are the joint and marginal forward pdf respectively, and $p(x_{t-1} | d_1, \dots, d_n)$ is the marginal backward pdf. The transition between line 1 and 2 is valid because for every t , x_t is conditionally independent of (d_1, \dots, d_{t-1}) given x_{t-1} . Therefore the transition between line 2 and 3 also is valid. Since x_t is conditionally independent of d_{t-1} , it can be joined to the conditioning. Going from line 3 to line 4 we are using Bayes rule.

The marginal backward probabilities can be calculated as follows,

$$p(x_t | d_1, \dots, d_n) = \sum_{x_{t-1}} p(x_{t-1}, x_t | d_1, \dots, d_n),$$

and initially we have that the marginal backward pdf at location $t = 1$ equals the marginal forward pdf at the same location, $p(x_1 | d_1, \dots, d_n)$.

In order to generate simulations from the posterior pdf, we need to calculate the conditional backward pdf, which is

$$\begin{aligned}
 p(x_t | x_{t-1}, d_1, \dots, d_n) &= \frac{p(x_{t-1}, x_t | d_1, \dots, d_n)}{p(x_{t-1} | d_1, \dots, d_n)} \\
 &= \frac{p(x_t, x_{t-1} | d_n, \dots, d_{t-1})}{p(x_{t-1} | d_n, \dots, d_{t-1})}.
 \end{aligned}$$

The transition here is verified in expression (8). The conditional backward pdf is actually

expressed by the joint and marginal forward pdf.

Algorithm

In Algorithm 1 the Forward-Backward recursion is presented. The algorithm calculates the posterior pdf, $p(\mathbf{x}|\mathbf{d})$, from which it generates simulation profiles, $\mathbf{x}^S = (x_1^S, \dots, x_n^S)$. It also calculates the marginal pdf, $p(x_t|\mathbf{d})$, for all t , from which the maximum marginal posterior solution is calculated, $\mathbf{x}^M = \{\max_{x_t} p(x_t|\mathbf{d}), t = 1, 2, \dots, n\}$.

Algorithm 1 Forward-Backward algorithm

- 1: $p(x_n|d_n) = const \times p_s(x_n)p(d_n|x_n)$
 - 2: $const = [\sum_{x_n} p(x_n|d_n)]^{-1}$
 - 3: **for** $t = n - 1, n - 2, \dots, 1$ **do**
 - 4: $p(x_{t+1}, x_t|d_n, \dots, d_t) = const \times p(d_t|x_t)p(x_t|x_{t+1})p(x_{t+1}|d_n, \dots, d_{t+1})$
 - 5: $const = [\sum_{x_{t+1}} \sum_{x_t} p(x_{t+1}, x_t|d_n, \dots, d_t)]^{-1}$
 - 6: $p(x_t|d_n, \dots, d_t) = \sum_{x_{t+1}} p(x_{t+1}, x_t|d_n, \dots, d_t)$
 - 7: **end for**
 - 8: Generate $x_1^S \sim p(x_1|d_1, \dots, d_n)$
 - 9: $x_1^M = \max_{x_1} p(x_1|d_1, \dots, d_n)$
 - 10: **for** $t = 2, 3, \dots, n$ **do**
 - 11: $p(x_t|x_{t-1}, d_1, \dots, d_n) = \frac{p(x_t, x_{t-1}|d_n, \dots, d_{t-1})}{p(x_{t-1}|d_n, \dots, d_{t-1})}$
 - 12: $p(x_t|d_1, \dots, d_n) = \sum_{x_{t-1}} p(x_t|x_{t-1}, d_1, \dots, d_n)p(x_{t-1}|d_1, \dots, d_n)$
 - 13: Generate $x_t^S \sim p(x_t|x_{t-1}, d_1, \dots, d_n)$
 - 14: $x_t^M = \max_{x_t} p(x_t|d_1, \dots, d_n)$
 - 15: **end for**
-

3.5 Bayesian model adapted to the data from the Statfjord Formation

From Figure 1-3 we notice that in the locations where there are transitions between classes of facies associations or lithofacies, the transitions in the logs are smooth. We assume that this smoothness is not only a result of noise in the data, but also of the spatial response from the formation. We assume that every observation in the log responds to not only the corresponding location in the formation, but also to k neighbours at each side, where $k \ll n$, as we can see from Figure 13.

If the the posterior pdf adapted to the data from the Statfjord Formation can be factorized and formulated as in expression (7), we are able to calculate it by the Forward-

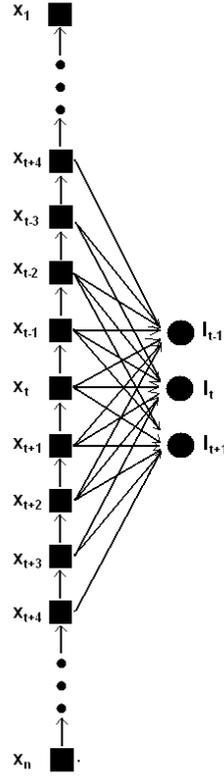


Figure 13: The graph illustrates the relation between the locations of the profile, x_1, \dots, x_n , and the observed log data, l_1, \dots, l_n .

Backward recursion in Algorithm 1. In this section we first derive a likelihood model based on the log-response from the formation, followed by a first order Markov chain prior model adapted to the likelihood model. There will obviously be irregularities at the borders, but a section will be ignored at the bottom and top of the wells due to this. Therefore, the boundary conditions will not be considered in the derivation of the model.

The methodology for the gammaray log, l^{gr} , and the density log, l^{gr} , is parallel, and therefore we only denote the logs by l . The same thing yields for the core classification into facies associations, π^{fa} , and lithofacies, π^{lf} , which we sometimes denote by π . The levels of response, the weights and the variance corresponding to either classification and log we denote by $r(\cdot)$, θ_i and σ^2 in the methodological discussion.

Likelihood model

As we can see from Figure 13, the observations of the log respond to more than one location in the core profile. We can express the log at location t as a weighted linear combination of the response in the corresponding location and k neighbours at each side, plus an error. The relation is given by

$$l_t = \sum_{i=-k}^k \theta_i r(x_{t+i}) + \epsilon_t, \quad (9)$$

where the weights are normalized, $\sum_{i=-k}^k \theta_i = 1$, and the response function, $r(\cdot)$, is the level of response for the class of the input location. We assume that also the error in expression (9) is independent, identical distributed as

$$\epsilon_t \sim \mathcal{N}(0, \sigma^2).$$

This gives the distribution of the likelihood function

$$p(\mathbf{l}|\mathbf{x}) = (2\pi)^{-\frac{n}{2}} \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^n [l_t - \sum_{i=-k}^k \theta_i r(x_{t+i})]^2 \right\},$$

which can be written in the form

$$p(\mathbf{l}|\mathbf{x}) = \prod_{t=1}^n p(l_t | x_{t-k}, \dots, x_{t+k}), \quad (10)$$

where $p(l_t | x_{t-k}, \dots, x_{t+k}) \sim \mathcal{N}(\sum_{i=-k}^k \theta_i r(x_{t+i}), \sigma^2)$.

In this study we consider two well logs, the gammaray log, \mathbf{l}^{gr} , and the density log, \mathbf{l}^d . The parameters in the likelihood function will be different for the logs, but the likelihood function for both logs are defined by expression (10). Conditioned on the formation of the well, the two logs are independent. Therefore the likelihood function for the entire model is the product of the likelihood function for the two logs, that is

$$p(\mathbf{l}|\mathbf{x}) = p(\mathbf{l}^{gr}|\mathbf{x})p(\mathbf{l}^d|\mathbf{x}).$$

Prior model

We have chosen to consider the prior model a first order Markov chain. In order to combine a first order Markov chain prior model with the likelihood function in expression (10), we need to consider the overlapping set of locations $(x_{t-k}, \dots, x_{t+k})$, as variables in the transition probabilities, that is

$$p(x_{t-k}, \dots, x_{t+k} | x_{t-k+1}, \dots, x_{t+k+1}).$$

The Markov chain prior model expressed by these transition probabilities is given by

$$p(\mathbf{x}) = \prod_{t=1}^n p(x_{t-k}, \dots, x_{t+k} | x_{t-k+1}, \dots, x_{t+k+1}). \quad (11)$$

Since the variables are overlapping, the transition probabilities $p(x_{t-k}, \dots, x_{t+k} | x_{t-k+1}, \dots, x_{t+k+1})$ are zero if the classes in the locations $x_{t-k+1}, \dots, x_{t+k}$ changes. If the classes in these locations are unchanged, we can express the transition probabilities by

$$\begin{aligned} p(x_{t-k}, \dots, x_{t+k} | x_{t-k+1}, \dots, x_{t+k+1}) &= p(x_{t-k} | x_{t-k+1}, \dots, x_{t+k+1}) \\ &= p(x_{t-k} | x_{t-k+1}). \end{aligned}$$

The transition in the first line is valid because reducing the probability space by removing locations that are also conditioned upon, will not change the probability. The transition between the first and the second line is valid because conditioned upon x_{t-k+1}, x_{t-k} is independent of $x_{t-k+2}, \dots, x_{t+k+1}$ because of the Markov property in expression (5). The probabilities $p(x_{t-k} | x_{t-k+1})$ are actually the transition probabilities in the first order Markov chain, $q(j|i)$, given by expression (4).

The transition probabilities in the prior model in expression (11) are now given by

$$p(x_{t-k}, \dots, x_{t+k} | x_{t-k+1}, \dots, x_{t+k+1}) = \begin{cases} p(x_{t-k} | x_{t-k+1}) & \text{for } x_{t-k+1}, \dots, x_{t+k} \text{ unchanged} \\ 0 & \text{otherwise} \end{cases}, \quad (12)$$

which define a transition probability matrix of size $m^{2k+1} \times m^{2k+1}$. The locations, $(x_{t-k}, \dots, x_{t+k})$, therefore take values from the state space defined by $\boldsymbol{\pi} \times \dots \times \boldsymbol{\pi}$, where $\boldsymbol{\pi} = (\pi_1, \dots, \pi_m)$.

Posterior model

The posterior pdf can now be expressed by the likelihood function for both logs and the prior pdf given in expression (10) and (11) respectively,

$$\begin{aligned} p(\mathbf{x}|\mathbf{l}^{gr}, \mathbf{l}^d) &= const \times p(\mathbf{l}^{gr}|\mathbf{x})p(\mathbf{l}^d|\mathbf{x})p(\mathbf{x}) \\ &= const \times \prod_{t=1}^n p(l_t^{gr}|x_{t-k}, \dots, x_{t+k})p(l_t^d|x_{t-k}, \dots, x_{t+k}) \\ &\quad \times p(x_{t-k}, \dots, x_{t+k}|x_{t-k+1}, \dots, x_{t+k+1}). \end{aligned} \quad (13)$$

The posterior pdf is factorized and can therefore be assessed by the Forward-Backward algorithm.

3.6 Location wise model

In order to illustrate the effect of taking advantage of the spatial coupling in the formation, and the neighbourhood effect in the logs in the stochastic model, we consider a simple model as a comparison. The simple model has a prior pdf without spatial coupling, that is

$$p(\mathbf{x}) = \prod_{t=1}^n p(x_t),$$

where $x_t \in \boldsymbol{\pi}$. A likelihood function without any neighbourhood effect is given in expression (2). The likelihood function for both logs together with the prior pdf above, define the location wise posterior pdf given by

$$p(\mathbf{x}|\mathbf{l}^{gr}, \mathbf{l}^d) = const \times \prod_{t=1}^n p(l_t^{gr}|x_t)p(l_t^d|x_t)p(x_t), \quad (14)$$

where *const* is a normalizing constant. The posterior pdf can be calculated location wise for every t , and from there the model has its name.

4 Parameter estimation

In this section we estimate the model parameters. The estimation is based on the training wells, A and B, because we want the test well, C, to be a blind well. At first, we estimate the likelihood parameters for the Bayesian model adapted to the data from the Staffjord Formation in expression (10), and we consider two different methods of estimating them.

The first method is based on visual inspection of the well log data, while in the second method we make use of the information we get from the first parameter estimation, and estimate the parameters by minimizing the squared error. Further, the prior transition probabilities in expression (11) are estimated by calculating the upward transitions in the cores. Finally, the model parameters for the location wise model in expression (14) are estimated, and the two different methods of estimating the likelihood parameters are considered.

In the following sections we denote the Bayesian model adapted to the data from the Statfjord Formation by the spatial model.

4.1 Likelihood parameters estimated based on visual inspection

At first, we estimate the weights, θ_i , followed by the response, $r(\cdot)$, and the variance, σ^2 .

Weights

As we can see from Figure 13, the observations in the logs respond to a set of locations in the formation. In order to know how many significant neighbours each of the logs respond to, we need to compare the logs with the core classification. From sections of the logs where we know from the core classification that there is a transition between two thick layers, we can calculate the number of significant neighbours, and how much influence each of them have on the logs, i.e. the weights.

Figure 14 shows five sections from well A and B where we have such transitions in both facies association and lithofacies. In some of the sections we can see that there is a shift in t between the log and the core, since the change of class is not corresponding to the center of the wavelet in the log. Shifts in t is a general problem in parameter estimation and facies classification. By looking at the logs isolated, we avoid this problem in the parameter estimation.

If we assume that the levels of response for each class are given, the exact weights for each of the sections in Figure 14 can be calculated directly from the log-wavelets. The weights are the gradients of the log-wavelets between the observations. From Figure 14 we can see that up to eight observations in the logs are involved in the transitions, which indicates that each observation in the logs might respond to up to three neighbours at each side of the corresponding location. We need to calculate the weights corresponding

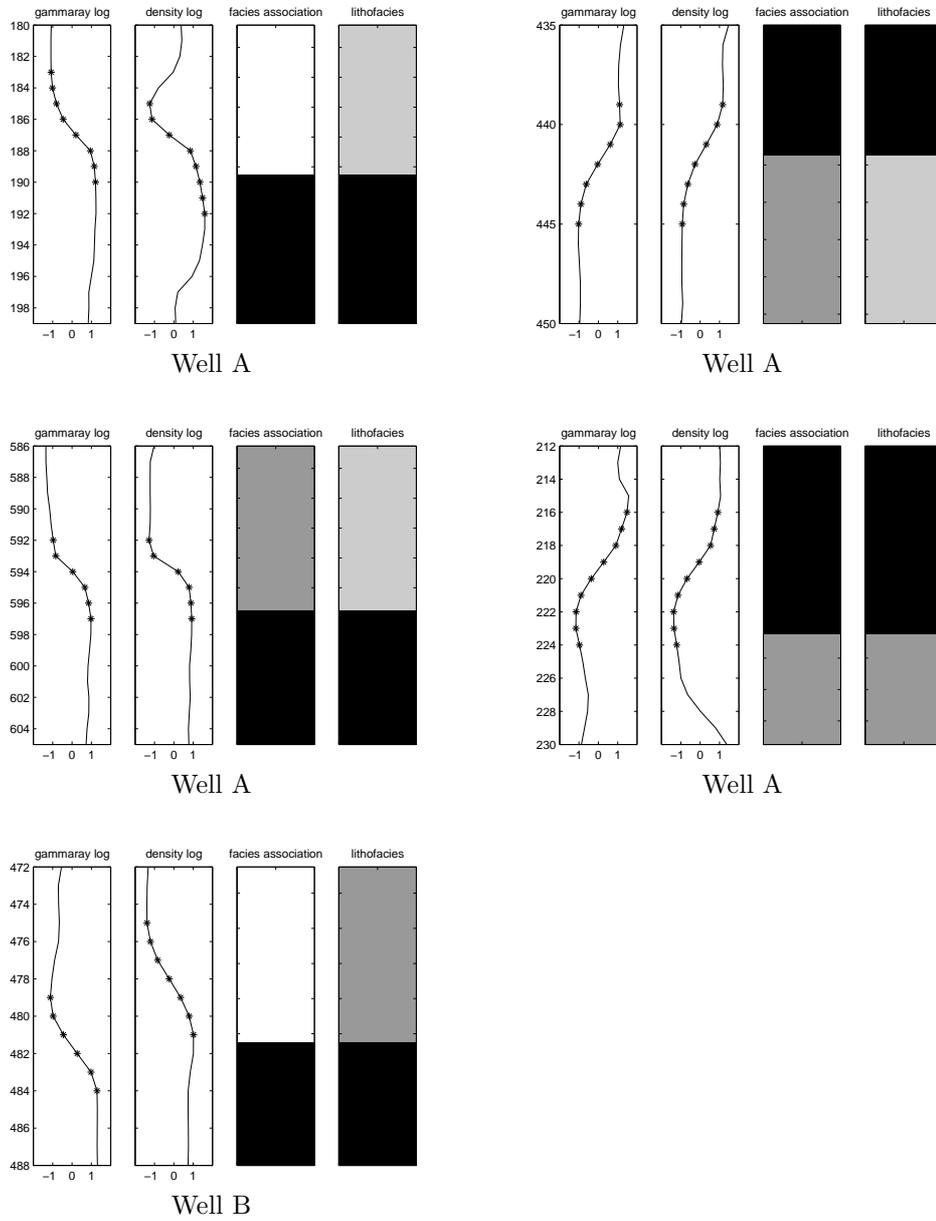


Figure 14: Five sections from the gammaray log, l^{gr} , the density log, l^{gr} , and the corresponding core classification into facies associations, π^{fa} , and lithofacies, π^{lf} from well A and B. The stars in the logs indicate the observations in the transition between the layers.

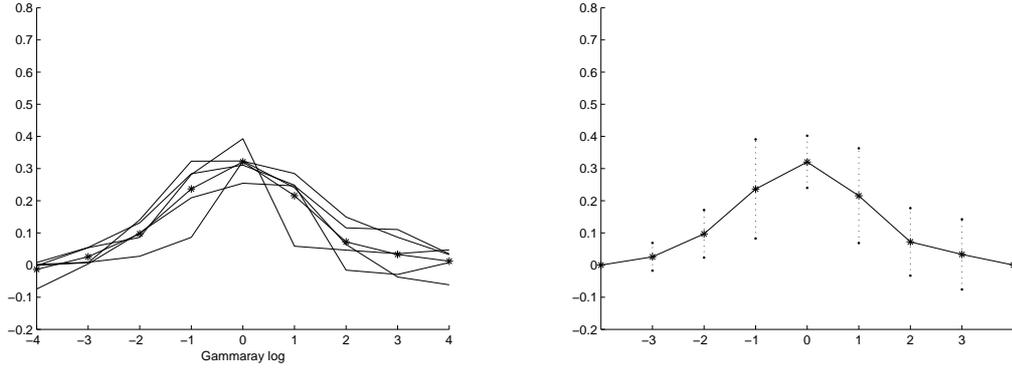


Figure 15: The calculated weights from the gammaray log, l^{gr} . Left are the weights calculated from the five sections in Figure 14, where the stars are the average of all the weights. Right is the average weights illustrated with 90% Gaussian confidence intervals.

to four neighbours at each side, in order to check if more than three neighbours are significant, which means θ_i for $i = -4, \dots, 4$.

The weights from the sections in Figure 14 are calculated and normalized, $\sum_{i=-4}^4 \theta_i = 1$, and illustrated in Figure 15 and 16. As we can see from Figure 16, one of the sections gives very small and negative θ_3 and θ_4 . We consider these as outliers, and adjust them to zero. Right in Figure 15 and 16 are plots of 90% Gaussian confidence intervals for all θ_i , and in the figures the stars are the average of the weights from all sections. In both the gammaray log and the density log the middle weight, θ_0 , is largest, descending at the sides. The weights are also approximately symmetric, which agrees with what we expect from the physics. The logs try to make their respective measurement in the actual location, but will at the same time get response from both sides of the location.

The average weights, 90% Gaussian confidence intervals and the estimates of the weights, $\hat{\theta}$, are presented in Table 1. From the table we can see that for both logs, θ_{-4} and θ_4 are not significantly different from zero, which means that there are not more than three significant neighbours.

Response

We know that the response in both logs is affected by neighbours and possible shifts in t in some sections, and therefore the average of the entire set of observations from each class is not a good estimate for the response. By removing some observations in the logs

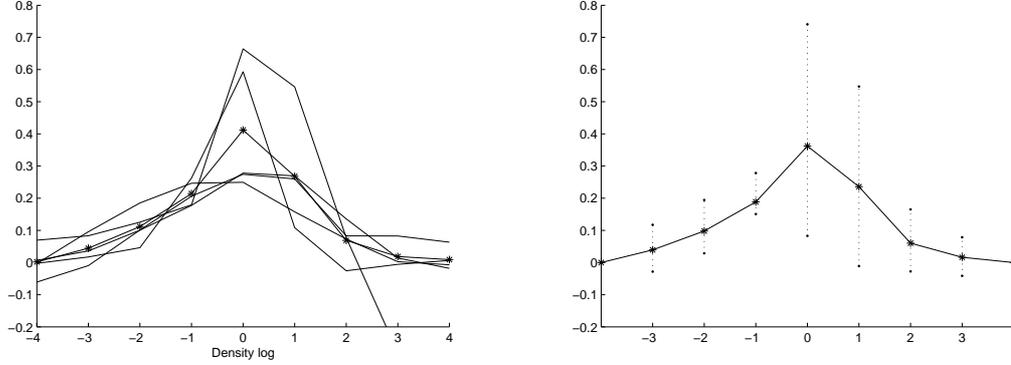


Figure 16: The calculated weights from the density log, \mathbf{l}^d . Left are the weights calculated from the five sections in Figure 14, where the stars are the average of all the weights. Right is the average weights illustrated with 90% Gaussian confidence intervals.

| | Gammaray log, \mathbf{l}^{gr} | | | Density log, \mathbf{l}^d | | |
|---------------|---------------------------------|-----------------------------------|----------------|-----------------------------|-----------------------------------|----------------|
| | Average | 90% Gaussian confidence intervals | $\hat{\theta}$ | Average | 90% Gaussian confidence intervals | $\hat{\theta}$ |
| θ_{-4} | -0.0136 | [-0.0699, 0.0427] | 0 | 0.0020 | [-0.0741, 0.0781] | 0 |
| θ_{-3} | 0.0257 | [-0.0175, 0.0688] | 0.0256 | 0.0445 | [-0.0284, 0.1174] | 0.0391 |
| θ_{-2} | 0.0972 | [0.0233, 0.1710] | 0.0970 | 0.1116 | [0.0289, 0.1944] | 0.0981 |
| θ_{-1} | 0.2368 | [0.0826, 0.3909] | 0.2364 | 0.2144 | [0.1507, 0.2782] | 0.1884 |
| θ_0 | 0.3209 | [0.2399, 0.4019] | 0.3204 | 0.4119 | [0.0829, 0.7408] | 0.3619 |
| θ_1 | 0.2158 | [0.0686, 0.3630] | 0.2154 | 0.2683 | [-0.0108, 0.5474] | 0.2357 |
| θ_2 | 0.0720 | [-0.0330, 0.1770] | 0.0719 | 0.0687 | [-0.0276, 0.1651] | 0.0604 |
| θ_3 | 0.0332 | [-0.0760, 0.1424] | 0.0331 | 0.0187 | [-0.0414, 0.0787] | 0.0164 |
| θ_4 | 0.0121 | [-0.0593, 0.0835] | 0 | 0.0089 | [-0.0433, 0.0611] | 0 |

Table 1: Estimates of the weights from the first parameter estimation. Average, 90% Gaussian confidence intervals and the best estimates, $\hat{\theta}$, for the gammaray log, \mathbf{l}^{gr} , and density, \mathbf{l}^d . The weights yields both classifications, i.e. facies associations, $\boldsymbol{\pi}^{fa}$, and lithofacies, $\boldsymbol{\pi}^{lf}$.

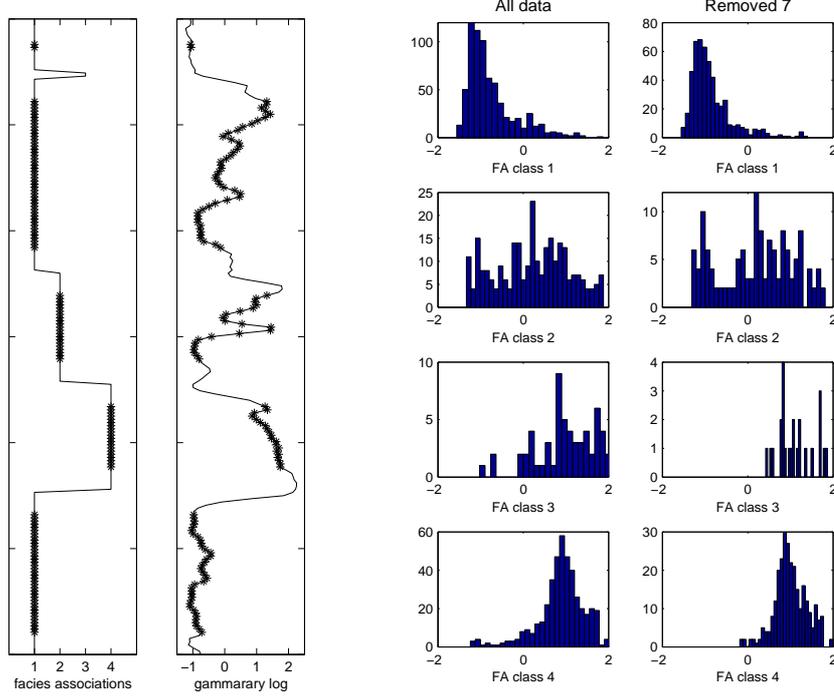


Figure 17: At the left is a section of core classification into facies associations, π^{fa} , and the same section of the gammaray log, l^{gr} , where the stars represent the observations that remain after seven observations at each side of every transition are removed. The left column of histograms are of all observations in the gammaray log, l^{gr} , from well A and B, for each of the classes of facies associations, π^{fa} . The right column of histograms are of the observations that remain after seven observations at each side of every transition are removed.

at both sides of every transition, we remove these effects.

At the left in Figure 17 and 18 we can see which observations in the logs that remain after removing seven observations at each side of the transitions between classes of facies associations, π^{fa} . We notice that the response in both logs is very unstable in the section of class π_2^{fa} . This may be due to misinterpretation of the core, but we decide to ignore this and accept the core classification to be the truth. The left column of histograms in Figure 17 and 18 is from the entire set of observations from the logs for every class, and the right column of histograms are after removing the observations in the transitions. From the histograms we can see that for class π_3^{fa} and π_4^{fa} the range of the response is considerably wider before removing the observations. The long tail that disappears in the histogram of class π_4^{fa} is a good illustration of the effects that we want to remove.

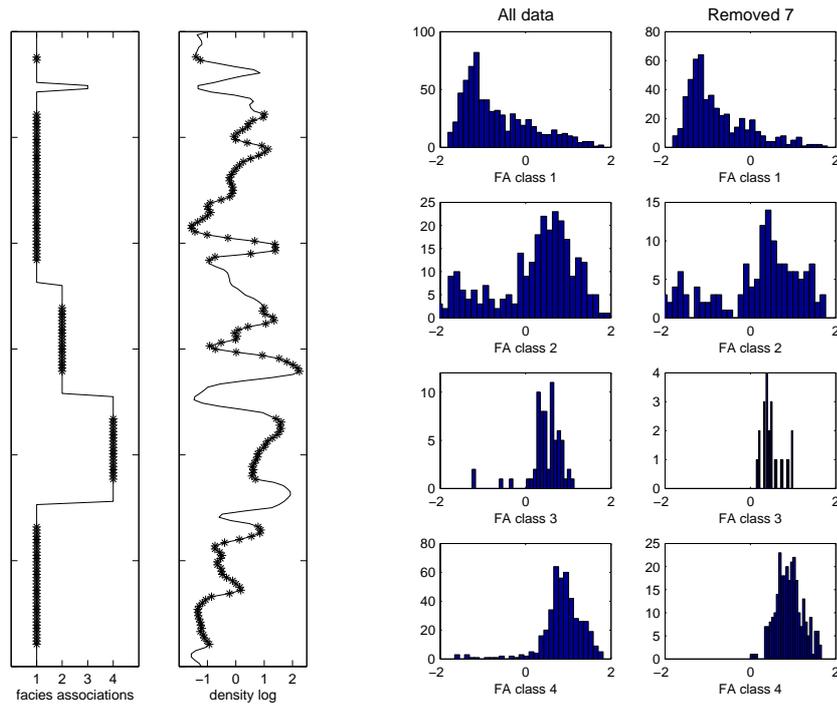


Figure 18: At the left is a section of core classification into facies associations, π^{fa} , and the same section of the density log, l^d , where the stars represent the observations that remain after seven observations at each side of every transition are removed. The left column of histograms are of all observations in the density log, l^d , from well A and B, for each of the classes of facies associations, π^{fa} . The right column of histograms are of the observations that remain after seven observations at each side of every transition are removed.

| | Gammaray log, \mathbf{l}^{gr} | | Density log, \mathbf{l}^d | |
|--------------|---------------------------------|------------------------------------|-----------------------------|------------------------------------|
| | $\hat{r}(\cdot)$ | 90% empirical confidence intervals | $\hat{r}(\cdot)$ | 90% empirical confidence intervals |
| π_1^{fa} | -0.9883 | [-1.3527, 0.2205] | -1.0010 | [-1.5575, 0.7966] |
| π_2^{fa} | 0.2510 | [-1.1836, 1.5271] | 0.4113 | [-1.7458, 1.4241] |
| π_3^{fa} | 1.0356 | [0.5334, 1.7876] | 0.4558 | [0.2039, 0.9866] |
| π_4^{fa} | 0.9685 | [0.3773, 1.6752] | 0.8968 | [0.4122, 1.5060] |

Table 2: Response estimates and 90% empirical confidence intervals for the response from the gammaray log, \mathbf{l}^{gr} , and density log, \mathbf{l}^d , for the classes of facies associations, π^{fa} , in the first parameter estimation.

| | Gammaray log, \mathbf{l}^{gr} | | Density log, \mathbf{l}^d | |
|--------------|---------------------------------|------------------------------------|-----------------------------|------------------------------------|
| | $\hat{r}(\cdot)$ | 90% empirical confidence intervals | $\hat{r}(\cdot)$ | 90% empirical confidence intervals |
| π_1^{lf} | -1.1677 | [-1.3887, -0.7977] | -1.1723 | [-1.4625, 0.1289] |
| π_2^{lf} | -1.1260 | [-1.2749, -0.8166] | -1.1313 | [-1.4442, 0.5377] |
| π_3^{lf} | -0.9846 | [-1.2301, -0.4995] | -1.1209 | [-1.3789, 1.7074] |
| π_4^{lf} | -0.8349 | [-0.8603, -0.8294] | 1.0336 | [0.4314, 1.0502] |
| π_5^{lf} | (1.2247) | NA | (0.6353) | NA |
| π_6^{lf} | 0.9918 | [0.3401, 1.7118] | 0.8236 | [0.3538, 1.4267] |

Table 3: Response estimates and 90% empirical confidence intervals for the response from the gammaray log, \mathbf{l}^{gr} , and density log, \mathbf{l}^d , for the classes of lithofacies, π^{lf} , in the first parameter estimation. The estimates of the response of class π_5^{lf} in parenthesis are made from all available locations for the class, since the layers of this class are too thin and no observations are left after removing observations in the transitions.

Estimates of the response for each of the classes of facies associations, for the gammaray log and the density log, $\mathbf{r}(x_t)$ for $x_t \in \pi^{fa}$, are the median of the remaining observations. These estimates and 90% empirical confidence intervals for the facies associations are presented in Table 2. Estimates of the response for the lithofacies, $\mathbf{r}(x_t)$ for $x_t \in \pi^{lf}$, are found similarly, and are presented in Table 3. The response for class π_5^{lf} was not possible to estimate this way, because the layers of this class are so thin that no observations were left after removing observations in the transitions. The estimates in parenthesis are made from all available observations of class π_5^{lf} .

| Core classification | Log | $\hat{\sigma}^2$ | 90% bootstrapping intervals |
|---------------------|----------|------------------|-----------------------------|
| π^{fa} | l^{gr} | 0.4297 | [0.2936, 1.1525] |
| π^{fa} | l^d | 0.5949 | [0.5493, 1.5814] |
| π^{lf} | l^{gr} | 0.4046 | [0.3200, 0.7294] |
| π^{lf} | l^d | 0.6004 | [0.4891, 1.2855] |

Table 4: Estimates of the variance, $\hat{\sigma}^2$, and 90% bootstrapping intervals for the four cases in the first parameter estimation.

Variance

Estimates of the variance we calculate from the expression,

$$\hat{\sigma}^2 = \widehat{Var}(\epsilon) = \frac{1}{n-1} \sum_{t=1}^n \epsilon_t^2 = \frac{1}{n-1} \sum_{t=1}^n \left\{ l_t - \sum_{i=-3}^3 \hat{\theta}_i \hat{r}(x_{t+i}) \right\}^2, \quad (15)$$

where the estimates of the response, $\hat{r}(\cdot)$, and the weights, $\hat{\theta}_i$ for $i = -3, \dots, 3$, are given in Table 1, 2 and 3. Histograms of the error, $\epsilon_t = l_t - \sum_{i=-3}^3 \hat{\theta}_i \hat{r}(x_{t+i})$ for $t = 1, \dots, n$, are presented in Figure 19, and we notice that the empirical distributions of the error are approximately Gaussian, though they are a little skew. The resulting estimates of the variance are given in Table 4. Since we have two logs and two ways of classifying the logs, there are four different estimates of the variance. The range of the variance is from 0.4 to 0.6, which is reasonable considering that the data alone have variance of 1 and we expect this variance to be less.

In order to make intervals of uncertainty for σ^2 , we need to use the method of bootstrapping, which is based on sampling from empirical distributions. A reference for bootstrapping is Efron and Tibshirani (2003). T times, one set of weights, θ_i^* for $i = -3, \dots, 3$, are drawn from the five possible in Figure 15 or 16, and levels of response, $r^*(\cdot)$, are drawn from their corresponding confidence intervals in Table 2 or 3, which we consider Gaussian. The error, σ^{*2} , is calculated by expression (15) at every instance, and 90% bootstrapping intervals are made from the T estimates of the error. The intervals are given in Table 4.

4.2 Likelihood parameters estimated by minimizing the squared error

In the previous section we noticed that the logs respond to the corresponding location in the formation and three neighbours at each side, and that the weights are approximately

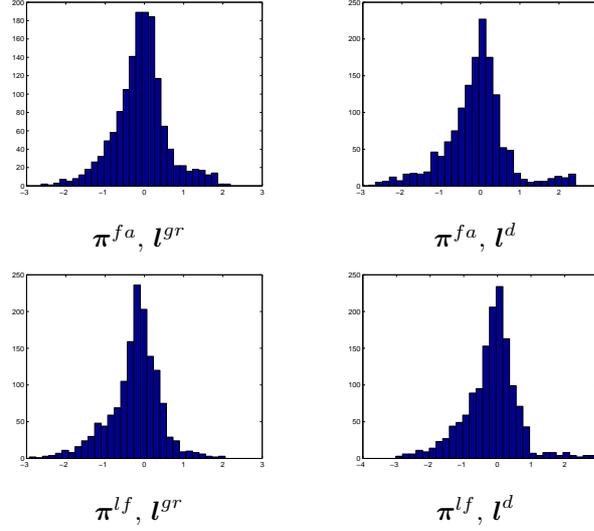


Figure 19: Histograms of the error, ϵ_t for $t = 1, \dots, n$, for the four cases of the spatial model in the first parameter estimation.

symmetric. If we decide that the weights are symmetric, $\theta_{-i} = \theta_i$, we have that $\theta_0 = 1 - 2\theta_1 - 2\theta_2 - 2\theta_3$, and we only need to estimate three weights, θ_1 , θ_2 and θ_3 . We have also observed that $\theta_0 \geq \theta_1 \geq \theta_2 \geq \theta_3$. These restrictions on the parameters we can express as intervals for the three parameters, that is

$$\begin{aligned} \theta_1 &\in [0, \frac{1}{3}(1 - 2\theta_2 - 2\theta_3)] \\ \theta_2 &\in [0, \theta_1] \\ \theta_3 &\in [0, \theta_2], \end{aligned} \quad (16)$$

where the weights are normalized, $\sum_{i=-3}^3 \theta_i = 1$. We want to estimate the parameters by minimizing the squared error, which gives us the expression

$$(\boldsymbol{\theta}, \mathbf{r}) = \arg \min_{(\boldsymbol{\theta}, \mathbf{r})} \sum_{t=1}^n \epsilon_t^2 = \arg \min_{(\boldsymbol{\theta}, \mathbf{r})} \sum_{t=1}^n \left\{ l_t - \sum_{i=-3}^3 \theta_i r(x_{t+i}) \right\}^2. \quad (17)$$

This is actually a linear regression problem if the weights are given, where the response for all classes are the regression coefficients. With the weights given, expression (17) can be equivalently expressed as

$$\mathbf{r} = (\mathbf{D}_\theta^T \mathbf{D}_\theta)^{-1} \mathbf{D}_\theta^T \mathbf{l}, \quad (18)$$

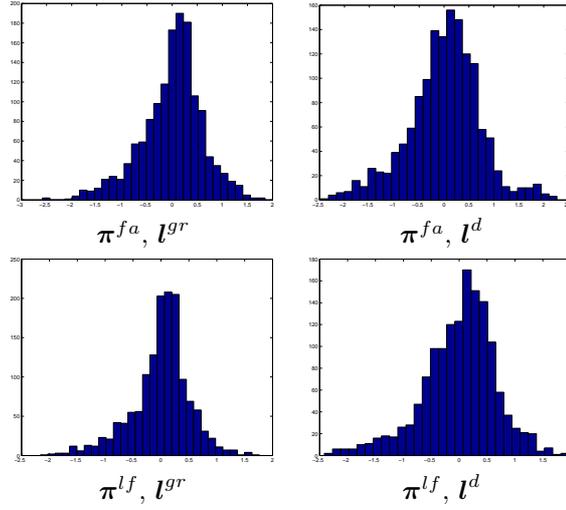


Figure 20: Histograms of the error, ϵ_t for $t = 1, \dots, n$, for the four cases of the spatial model in the second parameter estimation.

where \mathbf{r} is a vector of length m with response for the m classes, \mathbf{D}_θ is the design matrix for the regression problem where the weights corresponding to the core classification are represented, and \mathbf{l} is a vector of length n with the log data. For θ_i , $i = -3, \dots, 3$ given, the method of least squares in expression (18) calculates the minimizing response directly. By solving expression (18) for every combination of θ_1 , θ_2 and θ_3 on a regular grid given by the restrictions in expression (16), we can also identify the minimizing weights. We consider a grid with distance of 0.01 between the nodes, and further a zooming grid with distance of 0.001. We also check for possible shifts in t , which means that we solve the minimization problem for several shifts between the core and the log.

The resulting estimates for the weights, $\hat{\theta}_i$ for $i = -3, \dots, 3$, and the response, $\hat{r}(\cdot)$, are given in Table 5, 6, 7 and 8. Further, the variance, $\hat{\sigma}^2$ is calculated by

$$\hat{\sigma}^2 = \widehat{Var}(\epsilon) = \frac{1}{n-1} \sum_{t=1}^n \epsilon_t^2 = \frac{1}{n-1} \sum_{t=1}^n \left\{ l_t - \sum_{i=-3}^3 \hat{\theta}_i \hat{r}(x_{t+i}) \right\}^2.$$

Histograms of the error, $\epsilon_t = l_t - \sum_{i=-3}^3 \hat{\theta}_i \hat{r}(x_{t+i})$, are presented in Figure 20, and we notice that also in this parameter estimation we get approximately Gaussian distribution on the error. The estimates of the variance, $\hat{\sigma}^2$, are given in Table 9.

In order to get intervals for the uncertainty, we need to use the method of bootstrapping.

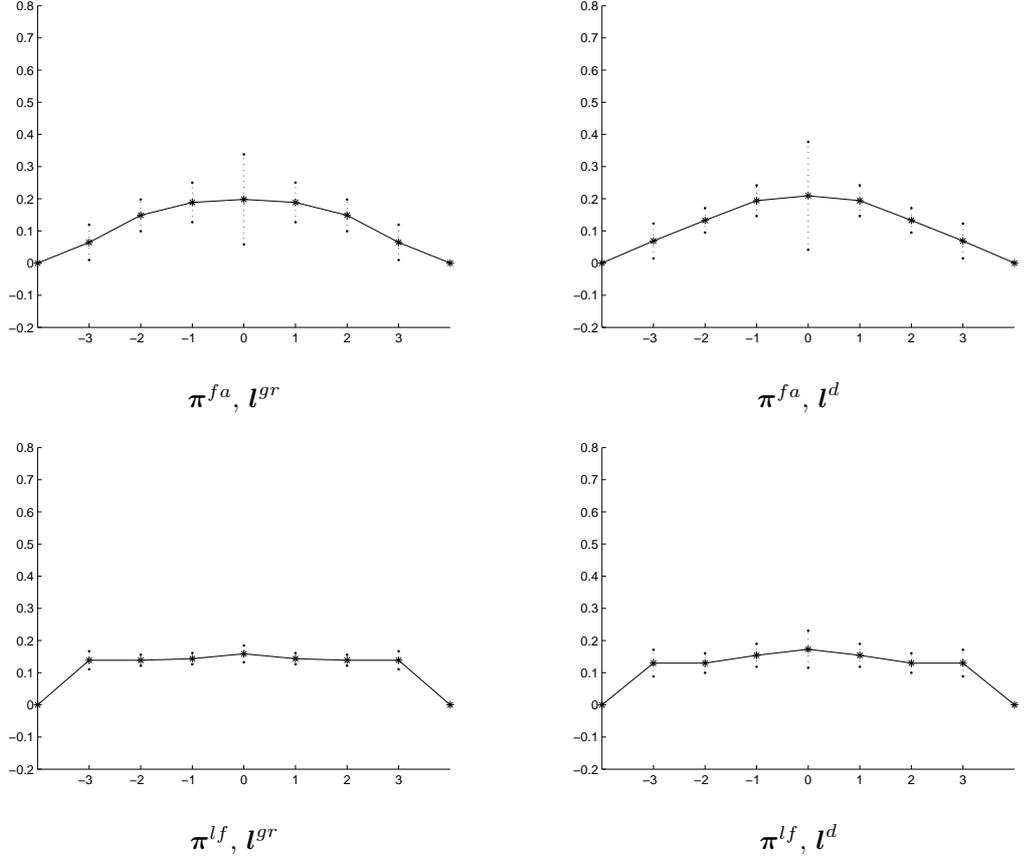


Figure 21: Estimated weights, $\hat{\theta}_i$ for $i = -3, \dots, 3$, with corresponding 90% bootstrapping intervals for the four cases in the second parameter estimation.

T times, a new set of log data, \mathbf{l}^* , is made from

$$l_t^* = \sum_{i=-3}^3 \hat{\theta}_i \hat{r}(x_{t+i}) + \epsilon^*,$$

where ϵ^* is drawn randomly from the error, $\epsilon_t = l_t - \sum_{i=-3}^3 \hat{\theta}_i \hat{r}(x_{t+i})$ for $t = 1, \dots, n$. From the log data, \mathbf{l}^* , the minimization problem in expression (17) is solved, and new estimates of the weights, θ_i^* for $i = -3, \dots, 3$, the response, $r^*(\cdot)$, and the variance, σ^{*2} , are made. From the T estimates of every parameter, 90% bootstrapping intervals are made. The estimates with the corresponding intervals are given in Table 5, 6, 7, 8 and 9.

The best estimates of the weights, $\hat{\theta}_i$ for $i = -3, \dots, 3$, together with the corresponding

| | Gammaray log, \mathbf{l}^{gr} | | Density log, \mathbf{l}^d | |
|---------------|---------------------------------|-----------------------------|-----------------------------|-----------------------------|
| | $\hat{\theta}$ | 90% bootstrapping intervals | $\hat{\theta}$ | 90% bootstrapping intervals |
| θ_{-3} | 0.0643 | [0.0093, 0.1192] | 0.0689 | [0.0147, 0.1230] |
| θ_{-2} | 0.1484 | [0.0989, 0.1978] | 0.1328 | [0.0946, 0.1709] |
| θ_{-1} | 0.1884 | [0.1270, 0.2497] | 0.1939 | [0.1463, 0.2415] |
| θ_0 | 0.1980 | [0.0578, 0.3382] | 0.2090 | [0.0413, 0.3767] |
| θ_1 | 0.1884 | [0.1270, 0.2497] | 0.1939 | [0.1463, 0.2415] |
| θ_2 | 0.1484 | [0.0989, 0.1978] | 0.1328 | [0.0946, 0.1709] |
| θ_3 | 0.0643 | [0.0093, 0.1192] | 0.0689 | [0.0147, 0.1230] |

Table 5: Estimates of the weights from the second parameter estimation for the facies associations, $\boldsymbol{\pi}^{fa}$. Best estimates, $\hat{\theta}$, and 90% bootstrapping intervals for the gammaray log, \mathbf{l}^{gr} , and density, \mathbf{l}^d .

bootstrapping intervals are illustrated in Figure 21. If we compare the weights in Figure 21 with the weights from the first parameter estimation in Figure 15 and 16, we notice that the weights from this parameter estimation are smoother than the others. This is due to the fact that we only check for global shifts in t when the weights are estimated this way. Different stretch in the logging tool wire at different depth result in different shifts in t along the well. The weights estimated when only global shifts in t are allowed, will be almost equalized due to this. In the first parameter estimation the weights are invariable for shifts in t , and therefore we do not have this problem.

The parameter estimates are relatively stable. The estimates of the response given in Table 7 and 8 have the same ordering as the estimates in the previous parameter estimation, given in Table 2 and 3, except for the facies associations for the density log. Due to estimating the parameters by minimizing the squared error, the estimates of the variance given in Table 9 are smaller than in the previous parameter estimation.

4.3 Prior model parameters

We need to estimate the transition probabilities in the prior Markov chain in expression (4), that is $q(j|i)$ for $i, j = 1, \dots, m$, for both facies associations, $\boldsymbol{\pi}^{fa}$, and lithofacies, $\boldsymbol{\pi}^{lf}$. In order to estimate these probabilities we count the upward transitions between all observations in well A and B, n_{ij} . From $\sum_{j=1}^m q(j|i) = 1$, we have that

$$q(j|i) = \frac{n_{ij}}{N_i}, \quad \text{where} \quad N_i = \sum_{j=1}^m n_{ij}.$$

| | Gammaray log, \mathbf{l}^{gr} | | Density log, \mathbf{l}^d | |
|---------------|---------------------------------|------------------|-----------------------------|------------------|
| | 90% bootstrapping intervals | | 90% bootstrapping intervals | |
| | $\hat{\theta}$ | | $\hat{\theta}$ | |
| θ_{-3} | 0.1386 | [0.1106, 0.1667] | 0.1299 | [0.0886, 0.1711] |
| θ_{-2} | 0.1386 | [0.1217, 0.1556] | 0.1299 | [0.0995, 0.1603] |
| θ_{-1} | 0.1435 | [0.1261, 0.1609] | 0.1539 | [0.1182, 0.1895] |
| θ_0 | 0.1585 | [0.1324, 0.1846] | 0.1727 | [0.1151, 0.2304] |
| θ_1 | 0.1435 | [0.1261, 0.1609] | 0.1539 | [0.1182, 0.1895] |
| θ_2 | 0.1386 | [0.1217, 0.1556] | 0.1299 | [0.0995, 0.1603] |
| θ_3 | 0.1386 | [0.1106, 0.1667] | 0.1299 | [0.0886, 0.1711] |

Table 6: Estimates of the weights from the second parameter estimation for the lithofacies, $\boldsymbol{\pi}^{lf}$. Best estimates, $\hat{\theta}$, and 90% bootstrapping intervals for the gammaray log, \mathbf{l}^{gr} , and density, \mathbf{l}^d .

| | Gammaray log, \mathbf{l}^{gr} | | Density log, \mathbf{l}^d | |
|--------------|---------------------------------|--------------------|-----------------------------|--------------------|
| | 90% bootstrapping intervals | | 90% bootstrapping intervals | |
| | $\hat{r}(\cdot)$ | | $\hat{r}(\cdot)$ | |
| π_1^{fa} | -0.7746 | [-0.9717, -0.6681] | -0.7077 | [-0.8653, -0.5710] |
| π_2^{fa} | 0.1385 | [-0.0384, 0.2946] | 0.1173 | [-0.0537, 0.2795] |
| π_3^{fa} | 1.2502 | [1.0521, 1.4860] | 0.6487 | [0.3972, 0.9621] |
| π_4^{fa} | 0.9594 | [0.7530, 1.1330] | 0.9151 | [0.7450, 1.0754] |

Table 7: Response estimates and 90% bootstrapping intervals for the response from the gammaray log, \mathbf{l}^{gr} , and density log, \mathbf{l}^d , for the classes of facies associations, $\boldsymbol{\pi}^{fa}$, in the second parameter estimation.

| | Gammaray log, \mathbf{l}^{gr} | | Density log, \mathbf{l}^d | |
|--------------|---------------------------------|--------------------|-----------------------------|--------------------|
| | 90% bootstrapping intervals | | 90% bootstrapping intervals | |
| | $\hat{r}(\cdot)$ | | $\hat{r}(\cdot)$ | |
| π_1^{lf} | -1.1494 | [-1.3793, -0.9961] | -0.6300 | [-0.8870, -0.3917] |
| π_2^{lf} | -0.8628 | [-1.0973, -0.6786] | -0.9555 | [-1.1361, -0.7586] |
| π_3^{lf} | -0.5977 | [-0.8025, -0.4550] | -0.5384 | [-0.7559, -0.4179] |
| π_4^{lf} | 0.2891 | [-0.0063, 0.3887] | 0.2800 | [-0.0314, 0.4512] |
| π_5^{lf} | 2.9801 | [2.6837, 3.6494] | 1.3542 | [0.8091, 2.2667] |
| π_6^{lf} | 1.0022 | [0.8366, 1.1469] | 0.9036 | [0.7070, 1.0796] |

Table 8: Response estimates and 90% bootstrapping intervals for the response from the gammaray log, \mathbf{l}^{gr} , and density log, \mathbf{l}^d , for the classes of lithofacies, $\boldsymbol{\pi}^{lf}$, in the second parameter estimation.

| Core classification | Log | $\hat{\sigma}^2$ | 90% bootstrapping intervals |
|---------------------|----------|------------------|-----------------------------|
| π^{fa} | l^{gr} | 0.3730 | [0.3806, 0.4464] |
| π^{fa} | l^d | 0.5281 | [0.4967, 0.5830] |
| π^{lf} | l^{gr} | 0.2696 | [0.2735, 0.3287] |
| π^{lf} | l^d | 0.4344 | [0.4125, 0.5620] |

Table 9: Estimates of the variance, $\hat{\sigma}^2$, and 90% bootstrapping intervals for the four cases in the second parameter estimation.

The number of upwards transitions, n_{ij} , in the facies associations, π^{fa} , are

$$\mathbf{n}^{fa} = \begin{bmatrix} 688 & 4 & 6 & 4 \\ 1 & 258 & 2 & 7 \\ 5 & 3 & 58 & 1 \\ 8 & 4 & 1 & 421 \end{bmatrix}. \quad (19)$$

From expression (19) we notice that the diagonal entries, n_{ii} , are much larger than the off-diagonal entries, which indicates that the layers of facies associations are thick. The number of transitions between layers, n_{ij} for $i \neq j$, differs and we can see that some transitions are very unlikely, for instance, $n_{21} = n_{34} = n_{43} = 1$. The transition probability matrix for the facies associations is given by

$$\mathbf{Q}^{fa} = \begin{bmatrix} 0.9801 & 0.0057 & 0.0085 & 0.0057 \\ 0.0037 & 0.9627 & 0.0075 & 0.0261 \\ 0.0746 & 0.0448 & 0.8657 & 0.0149 \\ 0.0184 & 0.0092 & 0.0023 & 0.9700 \end{bmatrix}, \quad (20)$$

which illustrates the spatial structure in the formation concerning the facies associations. Coming from class π_i^{fa} , the probability distribution for entering class π_j^{fa} for $j = 1, \dots, 4$ is multinomial with probability $q(j|i)$. From this we have that 90% confidence intervals for the probabilities are given by

$$\begin{bmatrix} [0.9797, 0.9804] & [0.0055, 0.0059] & [0.0083, 0.0088] & [0.0055, 0.0059] \\ [0.0034, 0.0041] & [0.9615, 0.9638] & [0.0069, 0.0080] & [0.0251, 0.0271] \\ [0.0682, 0.0811] & [0.0397, 0.0499] & [0.8573, 0.8740] & [0.0119, 0.0179] \\ [0.0179, 0.0189] & [0.0089, 0.0096] & [0.0021, 0.0025] & [0.9694, 0.9707] \end{bmatrix}.$$

The number of upwards transitions, n_{ij} , in the lithofacies, π^{lf} , in well A and B are,

$$\mathbf{n}^{lf} = \begin{bmatrix} 154 & 14 & 4 & 0 & 1 & 2 \\ 11 & 285 & 28 & 5 & 1 & 4 \\ 7 & 18 & 215 & 6 & 2 & 15 \\ 1 & 5 & 4 & 94 & 0 & 10 \\ 0 & 1 & 2 & 3 & 10 & 1 \\ 2 & 11 & 10 & 7 & 3 & 526 \end{bmatrix}. \quad (21)$$

The pattern in the matrix of lithofacies in expression (21) is more explanatory than for the facies associations. The four first classes, π_1^{lf} , π_2^{lf} , π_3^{lf} and π_4^{lf} all consists of sand of different grain size, where class π_1^{lf} is most coarse-grained, and class π_4^{lf} is the most fine-grained. From expression (21) we notice that starting in class π_1^{lf} , the most likely transition is to class π_2^{lf} , and from class π_2^{lf} , to class π_3^{lf} . From here the transition might be back to class π_2^{lf} again, or to class π_6^{lf} , while a transition into class π_4^{lf} is less likely. Except for the small probability of entering class π_4^{lf} from class π_3^{lf} , the pattern described above confirms the principle of fining upwards, which means that the fine-grained sand lies above the coarser-grained sand. In addition to this, some entries in the matrix in expression (21) are zero or very small, which justifies that there is a spatial structure in the formation. The transition probability matrix for the lithofacies is given by

$$\mathbf{Q}^{lf} = \begin{bmatrix} 0.8800 & 0.0800 & 0.0229 & 0 & 0.0057 & 0.0114 \\ 0.0329 & 0.8533 & 0.0838 & 0.0150 & 0.0030 & 0.0120 \\ 0.0266 & 0.0684 & 0.8175 & 0.0228 & 0.0076 & 0.0570 \\ 0.0088 & 0.0439 & 0.0351 & 0.8246 & 0 & 0.0877 \\ 0 & 0.0588 & 0.1176 & 0.1765 & 0.5882 & 0.0588 \\ 0.0036 & 0.0197 & 0.0179 & 0.0125 & 0.0054 & 0.9410 \end{bmatrix}, \quad (22)$$

with 90% confidence intervals for all probabilities given by

$$\begin{bmatrix} [0.8769, 0.8831] & [0.0774, 0.0826] & [0.0215, 0.0243] & [0, 0] & [0.0050, 0.0064] & [0.0104, 0.0124] \\ [0.0348, 0.0368] & [0.8384, 0.8424] & [0.0897, 0.0927] & [0.0156, 0.0170] & [0.0030, 0.0036] & [0.0124, 0.0136] \\ [0.0256, 0.0276] & [0.0669, 0.0700] & [0.8151, 0.8199] & [0.0219, 0.0237] & [0.0071, 0.0081] & [0.0556, 0.0585] \\ [0.0074, 0.0101] & [0.0409, 0.0468] & [0.0324, 0.0377] & [0.8191, 0.8300] & [0, 0] & [0.0836, 0.0918] \\ [0, 0] & [0.0361, 0.0816] & [0.0865, 0.1488] & [0.1396, 0.2134] & [0.5406, 0.6359] & [0.0361, 0.0816] \\ [0.0034, 0.0038] & [0.0193, 0.0201] & [0.0175, 0.0183] & [0.0122, 0.0128] & [0.0052, 0.0056] & [0.9403, 0.9417] \end{bmatrix}.$$

From expression (12) we have that

$$p(x_{t-k}, \dots, x_{t+k} | x_{t-k+1}, \dots, x_{t+k+1}) = \begin{cases} p(x_{t-k} | x_{t-k+1}) & \text{for } x_{t-k+1}, \dots, x_{t+k} \text{ unchanged} \\ 0 & \text{otherwise,} \end{cases}$$

where $(x_{t-3}, \dots, x_{t+3}) \in \boldsymbol{\pi} \times \dots \times \boldsymbol{\pi}$ since $x_t \in \boldsymbol{\pi}$. The state space will be of size 4^7 for the classes of facies associations, $\boldsymbol{\pi}^{fa}$, and of size 6^7 for the lithofacies, $\boldsymbol{\pi}^{lf}$. This gives transition probability matrices of size $4^7 \times 4^7$ and $6^7 \times 6^7$.

4.4 Location wise model parameters

For the location wise model in expression (14) the parameters that need to be estimated are the parameters in the likelihood function, which is the response, $r(\cdot)$, and variance, σ^2 , and the probabilities in the prior pdf, $p(x_t)$ for $x_t \in \boldsymbol{\pi}$. As for the spatial model, we consider two ways of estimating the likelihood parameters.

Likelihood parameters estimated based on visual inspection

Since the estimates of the response given in Table 2 and 3 are independent of the choice of model, we consider these parameters in the location wise model as well. If we assume that the log data respond to only one location in the formation, we get the relation

$$l_t = r(x_t) + \epsilon_t.$$

The estimates of the variance in this model are calculated from

$$\hat{\sigma}^2 = \widehat{Var}(\epsilon) = \frac{1}{n-1} \sum_{t=1}^n \epsilon_t^2 = \frac{1}{n-1} \sum_{t=1}^n [l_t - \hat{r}(x_t)]^2, \quad (23)$$

and the estimates are given in Table 10. Comparing these estimates of the variance with the estimates based on visual inspection for the spatial model given in Table 4, we can see that the variance in this model is larger in all cases. Since this model is actually too simple to express the structure of the underlying model, this agree with what we expect.

The intervals of uncertainty are made from T bootstrapping estimates, σ^{*2} , which are calculated from expression (23) with levels of response, $r^*(\cdot)$, drawn from their intervals in Table 2 and 3. The 90% bootstrapping intervals for σ^2 are given in Table 10.

| Core classification | Log | $\hat{\sigma}^2$ | 90% bootstrapping intervals |
|---------------------|----------|------------------|-----------------------------|
| π^{fa} | l^{gr} | 0.4484 | [0.4511, 0.9148] |
| π^{fa} | l^d | 0.6140 | [0.6004, 1.7387] |
| π^{lf} | l^{gr} | 0.5035 | [0.4441, 0.8530] |
| π^{lf} | l^d | 0.7047 | [0.6627, 1.6101] |

Table 10: Estimates of the variance, $\hat{\sigma}^2$, and 90% bootstrapping intervals for the four cases in the location wise model and the first parameter estimation.

Likelihood parameters estimated by minimizing the squared error

In this parameter estimation we estimate the response by minimizing the squared error, that is

$$\mathbf{r} = \arg \min_{\mathbf{r}} \sum_{t=1}^n \epsilon_t^2 = \arg \min_{\mathbf{r}} \sum_{t=1}^n [l_t - \hat{r}(x_t)]^2. \quad (24)$$

The variance is calculated from expression (23). The estimates of the response and variance are given in Table 11, 12 and 13. The estimates of the response have the same ordering and are not much different from the estimates in the spatial model given in Table 7 and 8. The estimates of the variance are larger than the estimates given in Table 9, as we expected.

In order to get intervals of uncertainty for these parameters, we bootstrap T sets of new log data, l^* , where

$$l_t^* = \hat{r}(x_t) + \epsilon^*,$$

and ϵ^* is drawn randomly from the error, $\epsilon_t = l_t - r(x_t)$, for $t = 1, \dots, n$. From the new sets of log data, we calculate the response, $r^*(\cdot)$, by expression (24), and further the variance, σ^{*2} , from expression (23). From the T estimates of the response and the variance, 90% bootstrapping intervals are calculated, which are given in Table 11, 12 and 13.

Prior model parameters

We want the probabilities for the different classes to reflect the proportion of locations in the formation that are classified to the respective classes. The proportion of each class of facies associations and lithofacies for the three wells are illustrated in Figure 5 and 7 respectively. The estimates of the probabilities in the prior model are the average of the proportions of well A and B. The estimates for the facies associations and lithofacies are

| | Gammaray log, \mathbf{l}^{gr} | | Density log, \mathbf{l}^d | |
|--------------|---------------------------------|-----------------------------|-----------------------------|-----------------------------|
| | $\hat{r}(\cdot)$ | 90% bootstrapping intervals | $\hat{r}(\cdot)$ | 90% bootstrapping intervals |
| π_1^{fa} | -0.7484 | [-0.7809,-0.7068] | -0.6541 | [-0.7037,-0.6133] |
| π_2^{fa} | 0.2186 | [0.1420, 0.3124] | 0.2073 | [0.1476, 0.2762] |
| π_3^{fa} | 1.1399 | [0.9959, 1.2551] | 0.4674 | [0.3178, 0.6177] |
| π_4^{fa} | 0.8971 | [0.8381, 0.9506] | 0.8553 | [0.8069, 0.9096] |

Table 11: Response estimates and 90% bootstrapping intervals for the response from the gammaray log, \mathbf{l}^{gr} , and density log, \mathbf{l}^d , for the classes of facies associations, π^{fa} , in the location wise model and the second parameter estimation.

| | Gammaray log, \mathbf{l}^{gr} | | Density log, \mathbf{l}^d | |
|--------------|---------------------------------|-----------------------------|-----------------------------|-----------------------------|
| | $\hat{r}(\cdot)$ | 90% bootstrapping intervals | $\hat{r}(\cdot)$ | 90% bootstrapping intervals |
| π_1^{lf} | -1.0255 | [-1.0837,-0.9641] | -0.5759 | [-0.6530,-0.4959] |
| π_2^{lf} | -0.7909 | [-0.8433,-0.7394] | -0.8019 | [-0.8426,-0.7284] |
| π_3^{lf} | -0.4079 | [-0.4690,-0.3559] | -0.4308 | [-0.4953,-0.3385] |
| π_4^{lf} | 0.2429 | [0.1252, 0.3120] | 0.1996 | [0.1081, 0.2955] |
| π_5^{lf} | 1.3076 | [1.0775, 1.0775] | 0.4835 | [0.1792, 0.7674] |
| π_6^{lf} | 0.8871 | [0.8483, 0.9286] | 0.7938 | [0.7486, 0.8458] |

Table 12: Response estimates and 90% bootstrapping intervals for the response from the gammaray log, \mathbf{l}^{gr} , and density log, \mathbf{l}^d , for the classes of lithofacies, π^{lf} , in the location wise model and the second parameter estimation.

| Core classification | Log | $\hat{\sigma}^2$ | 90% bootstrapping intervals |
|---------------------|-------------------|------------------|-----------------------------|
| π^{fa} | \mathbf{l}^{gr} | 0.4272 | [0.3913, 0.4579] |
| π^{fa} | \mathbf{l}^d | 0.5620 | [0.5292, 0.5900] |
| π^{lf} | \mathbf{l}^{gr} | 0.3785 | [0.3523, 0.4055] |
| π^{lf} | \mathbf{l}^d | 0.5351 | [0.5017, 0.5699] |

Table 13: Estimates of the variance, $\hat{\sigma}^2$, and 90% bootstrapping intervals for the four cases in the location wise model and the second parameter estimation.

| Facies associations | | Lithofacies | |
|---------------------|------------------|--------------|------------------|
| x_t | $\hat{p}(\cdot)$ | x_t | $\hat{p}(\cdot)$ |
| π_1^{fa} | 0.4766 | π_1^{lf} | 0.1194 |
| π_2^{fa} | 0.1826 | π_2^{lf} | 0.2278 |
| π_3^{fa} | 0.0455 | π_3^{lf} | 0.1794 |
| π_4^{fa} | 0.2953 | π_4^{lf} | 0.0784 |
| | | π_5^{lf} | 0.0116 |
| | | π_6^{lf} | 0.3834 |

Table 14: Prior model parameter estimates for the location wise model, for the facies associations and lithofacies.

given in Table 14. Since the prior probabilities in the location wise model are calculated from the same wells as the prior probabilities in the spatial model, the parameters are consistent with the parameters in expression (20) and (22).

5 Implementation and test criteria

Since the set of overlapping locations, $(x_{t-3}, \dots, x_{t+3})$ for $t = 1, \dots, n$, constitute a part of the posterior pdf in expression (13), the Forward-Backward algorithm needs to be reformulated. The reformulated recursion is given in Algorithm 2, and we notice that calculating the constant in line 3 requires a sum over $4^{14} \approx 2.68 \times 10^8$ elements for the facies associations and $6^{14} \approx 7.84 \times 10^{10}$ elements for the lithofacies. This is too large for a normal computer to calculate. To be able to calculate the posterior pdf we need to take a look at the pattern in the transition probability matrices. Since the locations, $(x_{t-3}, \dots, x_{t+3})$ for $t = 1, \dots, n$, are overlapping, the majority of the elements in the transition probability matrices are zero.

The stars in Figure 22 illustrate the pattern of nonzero elements in the transition probability matrix for the facies associations. The nonzero elements are located four by four, and as we can see they form four declining rows. The nonzero elements of the transition probability matrix for the lithofacies will behave equivalently, in six declining rows. In order to calculate the posterior pdf in expression (13) by the Forward-Backward algorithm, we need to implement smartly. By removing all the zero elements, we are left with four and six rows of nonzero elements, and get transition probability matrices of size 4×4^7 and 6×6^7 instead. The sum in line 3 in the Forward-Backward algorithm is then over $4^8 = 65536$ and $6^8 = 1679616$ elements, which is manageable.

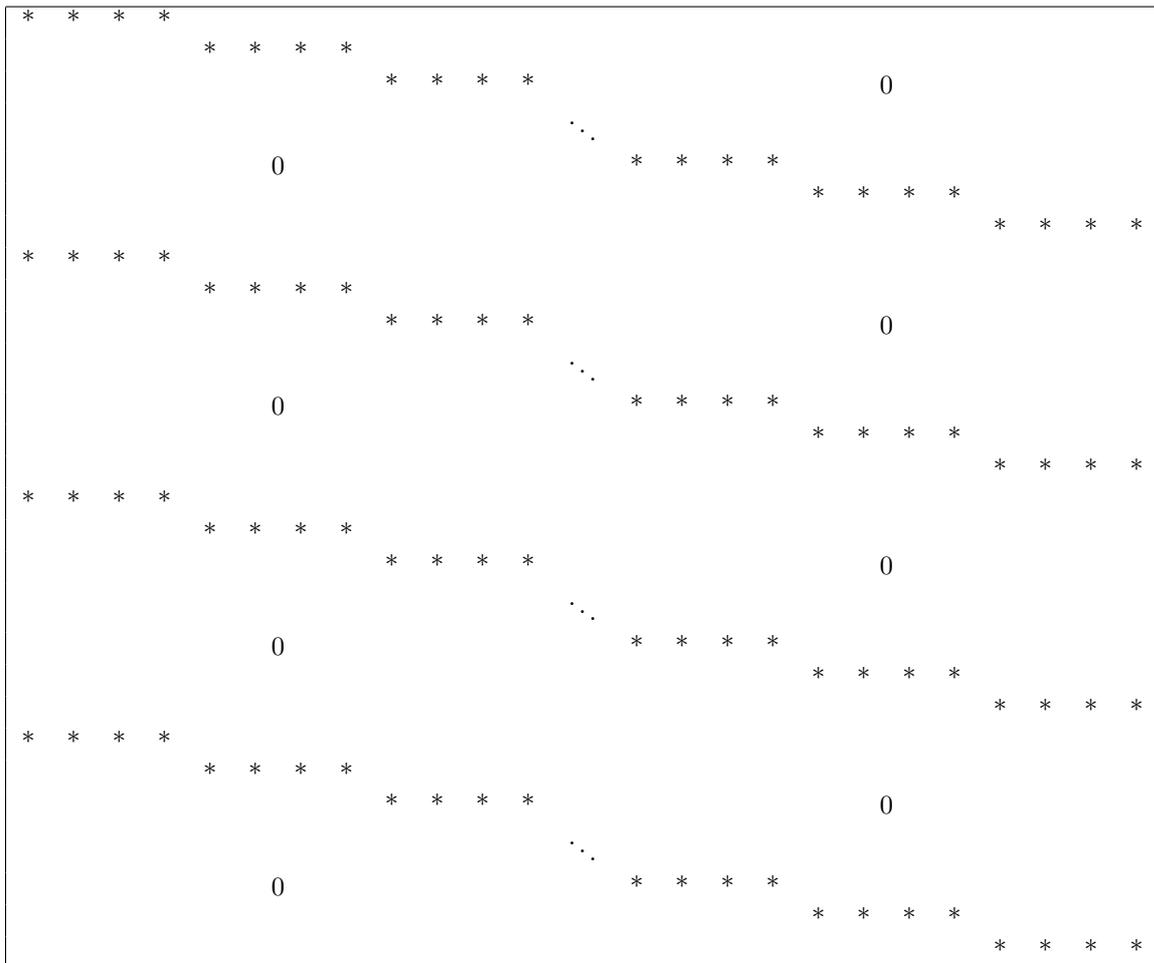


Figure 22: Illustration of the pattern in the transition probability matrix of size $4^7 \times 4^7$ for the facies associations. The stars represent the pattern of nonzero elements.

Algorithm 2 Forward-Backward algorithm for the data from the Statfjord Formation

```

1: for  $t = n - 1, n - 2, \dots, 1$  do
2:    $p(\{x_{t-2}, \dots, x_{t+4}\}, \{x_{t-3}, \dots, x_{t+3}\} | l_n, \dots, l_t) = \text{const} \times p(l_t | \{x_{t-3}, \dots, x_{t+3}\})$ 
      $\times p(\{x_{t-3}, \dots, x_{t+3}\} | \{x_{t-2}, \dots, x_{t+4}\}) p(\{x_{t-2}, \dots, x_{t+4}\} | l_n, \dots, l_{t+1})$ 
3:    $\text{const} = \left[ \sum_{\{x_{t-2}, \dots, x_{t+4}\}} \sum_{\{x_{t-3}, \dots, x_{t+3}\}} p(\{x_{t-2}, \dots, x_{t+4}\}, \{x_{t-3}, \dots, x_{t+3}\} | l_n, \dots, l_t) \right]^{-1}$ 
4:    $p(\{x_{t-3}, \dots, x_{t+3}\} | l_n, \dots, l_t)$ 
      $= \sum_{\{x_{t-2}, \dots, x_{t+4}\}} p(\{x_{t-2}, \dots, x_{t+4}\}, \{x_{t-3}, \dots, x_{t+3}\} | l_n, \dots, l_t)$ 
5: end for
6: for  $t = 2, 3, \dots, n$  do
7:    $p(\{x_{t-3}, \dots, x_{t+3}\} | \{x_{t-4}, \dots, x_{t+2}\}, l_1, \dots, l_n)$ 
      $= \frac{p(\{x_{t-3}, \dots, x_{t+3}\}, \{x_{t-4}, \dots, x_{t+2}\} | l_n, \dots, l_{t-1})}{p(\{x_{t-4}, \dots, x_{t+2}\} | l_n, \dots, l_{t-1})}$ 
8:    $p(\{x_{t-3}, \dots, x_{t+3}\} | l_1, \dots, l_n)$ 
      $= \sum_{\{x_{t-4}, \dots, x_{t+2}\}} p(\{x_{t-3}, \dots, x_{t+3}\} | \{x_{t-4}, \dots, x_{t+2}\}, l_1, \dots, l_n) p(\{x_{t-4}, \dots, x_{t+2}\} | l_1, \dots, l_n)$ 
9:   Generate  $\{x_{t-3}, \dots, x_{t+3}\}^S \sim p(\{x_{t-3}, \dots, x_{t+3}\} | \{x_{t-4}, \dots, x_{t+2}\}, l_1, \dots, l_n)$ 
10:   $p(x_t | l_1, \dots, l_n) = \sum_{x_{t-3}} \sum_{x_{t-2}} \sum_{x_{t-1}} \sum_{x_{t+1}} \sum_{x_{t+2}} \sum_{x_{t+3}} p(x_{t-3}, \dots, x_{t+3} | l_1, \dots, l_n)$ 
11:   $x_t^M = \text{max}_{x_t} p(x_t | l_1, \dots, l_n)$ 
12: end for

```

5.1 Test criteria

In this section we define the criteria we will use to evaluate the posterior pdf, $p(\mathbf{x} | \mathbf{l}^{gr}, \mathbf{l}^d)$, and the marginal posterior pdf, calculated by

$$p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d) = \sum_{x_{t-3}} \sum_{x_{t-2}} \sum_{x_{t-1}} \sum_{x_{t+1}} \sum_{x_{t+2}} \sum_{x_{t+3}} p(x_{t-3}, \dots, x_{t+3} | \mathbf{l}^{gr}, \mathbf{l}^d),$$

for $t = 1, \dots, n$.

Mmap solution versus true core profile

The predicted profile is the marginal maximum posterior solution (mmap), and is calculated from

$$x_t^M = \arg \max_{x_t} p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d),$$

where $\mathbf{x}^M = (x_1^M, \dots, x_n^M)$ and $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$ is the marginal posterior pdf for all t . The mmap solution is displayed next to the true core profile. Comparison can be done by visual inspection. The ideal mmap solution is identical to the true core profile. We expect the mmap solution to be smoother, because it is a prediction towards the mean.

Display of the marginal posterior pdfs

The probabilities for each class at every location in the posterior pdf is displayed separately, i.e. the marginal posterior pdf, $p(x_t|\mathbf{l}^{gr}, \mathbf{l}^d)$, for $x_t \in \boldsymbol{\pi}$. The presence of the classes in the true core profile is marked on the axis of the respective plots. In an ideal prediction the probability profiles would be square-shaped with probability of 1 when the class is present in the true core profile and 0 elsewhere.

Confusion matrix

The confusion matrix reflects the relation between the locations of the true core profile, x_t , and the mmap solution, x_t^M , for all t . The sum of the rows is the number of times each class appears in the true core profile, and the sum of the columns is the number of times they appear in the mmap solution. Each element of the matrix is the number of locations where the class given in the row appears in the true profile, and in the mmap solution is classified into the class given in the column. Thus, the sum of the diagonal elements of the confusion matrix is the number of locations where the mmap solution equals the true core profile. In an ideal prediction the off-diagonal elements of the confusion matrix would be zero.

Share of matches

The share of matches is the number of times the mmap solution equals the true core profile, divided by the number of locations in the profile. In an ideal prediction the share of matches would be one. If the proportion of one class in the true core profile is very large, a mmap solution with only this class will give a share of matches equally to this proportion. Therefore the share of matches alone is not a sufficient measure of the results from the classification.

Simulated profiles versus true core profile

We simulate several profiles from the posterior pdf, $\mathbf{x}^S \sim p(\mathbf{x}|\mathbf{l}^{gr}, \mathbf{l}^d)$. The simulated profiles show the variation in the posterior pdf. The simulated profiles we display next to the true profile, hence comparisons can be done by visual inspection.

Signal-to-noise ratio

As a measure of how much signal we get from the log data we calculate the ratio

$$\frac{signal}{noise} = \frac{\sigma_s^2}{\sigma_n^2}$$

where σ_s^2 is the variance in the signal and σ_n^2 is the variance in the noise. The variance in the signal is calculated by

$$\sigma_s^2 = \frac{1}{n-1} \sum_{t=1}^n \left\{ \hat{r}(x_t) - \frac{1}{n} \sum_{k=1}^n \hat{r}(x_k) \right\}^2.$$

The noise is the result we get by subtracting the signal from the log data, and the variance in the noise is calculated by

$$\sigma_n^2 = \frac{1}{n-1} \sum_{t=1}^n \left\{ [l_t - \hat{r}(x_t)] - \frac{1}{n} \sum_{k=1}^n [l_t - \hat{r}(x_k)] \right\}^2.$$

Both estimates of variance depend on the estimates of the response, $\hat{r}(\cdot)$, in the model. The larger the ratio, the better classification we expect. If the ratio is 1 we have just as much noise as signal in the log data, and if the ratio larger than 1 we have more signal than noise.

6 Results and discussions

In this section the results from the classifications with the Bayesian model adapted to the Statfjord data, named the spatial model, and the location wise model are presented. The posterior pdfs based on the well log data, \mathbf{l}^{gr} and \mathbf{l}^d , are calculated for both facies associations, $\boldsymbol{\pi}^{fa}$, and lithofacies, $\boldsymbol{\pi}^{lf}$, with likelihood parameters from the first and second parameter estimation. We present the results from the two training wells, A and B, and the test well, C. At first the results from the classification of facies associations are evaluated, followed by the results from the classification of lithofacies. At every incident, the results from the location wise model are displayed next to the results from the spatial model, thus the results can be compared. We expect the results from the spatial model to be better than the results from the location wise model because the spatial model is more adjusted to the underlying model than the location wise model.

6.1 Facies associations

Before we evaluate the results, some typical challenges in the classification of facies associations can be mentioned. Two examples of problems are that thin layers of class π_2^{fa} between thick layers of class π_4^{fa} are often assumed to be class π_1^{fa} , and thin layers of π_3^{fa} between layers of π_1^{fa} are assumed to be class π_4^{fa} . The first problem appears at approximately $t = 555$ and $t = 590$ in well A and at approximately $t = 450$ in well C. The second problem appears at approximately $t = 115$ and $t = 320$ in well A, at approximately $t = 430$ in well B and at approximately $t = 230$ and $t = 260$ in well C. In the following discussion of the results for the facies association, we will comment whether our model manages to make correct classifications in these locations.

Another challenge is to correctly classify the small classes. From Figure 5 we can see the proportions of all classes of facies associations in the three wells. Notice that class π_3^{fa} is a small class in all wells. The proportions are also different between the wells, and particularly we notice that the proportion of class π_4^{fa} is very high in the test well, C.

The results from the classification of facies associations are presented in Figure 23-40. First we consider the results from the first parameter estimation, followed by the results from the second.

Likelihood parameters estimated based on visual inspection

The results from the first parameter estimation for the facies associations are given in Figure 23-25 for the first training well, A, in Figure 26-28 for the second training well, B, and in Figure 29-31 for the test well, C.

By visually comparing the two mmap solutions for the first training well, A, in Figure 23, we find them equally good. Both logs seem to find class π_1^{fa} at most locations, which makes the pattern look similar to the truth. Notice that only the spatial model finds class π_3^{fa} at approximately $t = 115$ and class π_2^{fa} at approximately $t = 555$. From the display of the marginal posterior pdfs in Figure 24 we can see that the probability curves in the spatial model are smoother than for the location wise model, which indicates that the spatial model is less sensitive to noise than the location wise model. The probabilities for class π_3^{fa} in the location wise model are small at all locations, which is reflected in the confusion matrix where we can see that the number of locations in the mmap solution that are classified into class π_3^{fa} for this model is zero. The share of matches is of the

same reason smaller for the location wise model. The simulations in Figure 25 reflect the variation in the posterior pdfs, and especially the instability in the location wise model which results in many thin misclassified layers.

The mmap solutions for the second training well, B, in Figure 26 both classify class π_1^{fa} almost correctly, which makes the pattern of the mmap solutions good, except that they misclassify a section of class π_2^{fa} at around $t = 100$ into class π_4^{fa} . From the display of the marginal posterior pdfs in Figure 27 we can see that both models have small probabilities for class π_3^{fa} . This is reflected in the confusion matrices. Class π_3^{fa} is a small class in well B, and therefore it does not influence much on the share of matches, which are 0.78 and 0.76 for the spatial and location wise model respectively. The share of matches of 0.78 is the largest match we have when it comes to the facies associations for the spatial model, and there might be a connection with the large signal-to-noise ratio for the gammaray log, of 2.60. The simulations in Figure 28 for the spatial model are very similar to the mmap solution, and although we recognize the pattern from the true profile in the simulations for the location wise model, the instability of the posterior pdf is prominent.

We notice from the mmap solutions in Figure 29 that the spatial model seems to have problems finding class π_4^{fa} in the thick layers of the test well, C. As we can see from Figure 5 the proportion of class π_4^{fa} is very high in this well. The location wise model does not seem to have this problem, though the thick layers are partitioned by relatively thin layers of class π_2^{fa} . The display of the marginal posterior pdfs and the confusion matrices in Figure 30 tell the same story about class π_4^{fa} as the mmap solutions. Further, the spatial model over-classify class π_3^{fa} , while the location wise model does not find class π_3^{fa} at all. In addition both models over-classify class π_2^{fa} . This leads to share of matches of 0.5431 and 0.6936 for the spatial and location wise model respectively. We notice that the spatial model manages to find the thin layer of class π_3^{fa} at approximately $t = 230$ and some of the layer of class π_2^{fa} at approximately $t = 450$. The simulations in Figure 31 display the variation in the posterior pdfs, and we describe the results to be of only average quality.

Likelihood parameters estimated by minimizing the squared error

The results from the second parameter estimation are given in Figure 32-34 for the first training well, A, in Figure 35-37 for the second training well, B, and in Figure 38-40 for the test well, C. The least squares parameter estimates make the signal-to-noise ratios

smaller in all wells, which may lead to poorer results.

From the mmap solutions for the first training well, A, in Figure 32 we can see that the spatial model does not find class π_4^{fa} in the thick layer around $t = 200$, as apposed to the location wise model and both models with the first parameter estimates. Due to this layer, the share of matches given in Figure 33 is smaller with these likelihood parameters for the spatial model. Except from this, the results with the second parameter estimates does not differ much from the first for this well, and we can add that also here the spatial model finds class π_3^{fa} at approximately $t = 115$ and class π_2^{fa} at approximately $t = 555$. The location wise model has problems finding the small classes, π_2^{fa} and π_3^{fa} , but because of the good match of class π_4^{fa} , the share of matches is slightly better than for the spatial model. The simulations are displayed in Figure 34, and notice that also here the simulations reflect the degree of variation in the posterior pdf.

From Figure 35 we can see that with the second parameter estimates, in the mmap solution for the second training well, B, for the spatial model, many sections of class π_4^{fa} are misclassified into class π_3^{fa} . Class π_3^{fa} is a small class, which results in the location wise model not finding the class at all. The marginal posterior pdfs and the confusion matrices in Figure 36 confirm this, but still the share of matches are 0.7730 and 0.7718 for the spatial and location wise model respectively. The simulations from the spatial model in Figure 37 also contain many layers of class π_3^{fa} , while the simulations from the location wise model are unstable as we have seen before.

From the mmap solutions for the test well, C, in Figure 38 we can see that the spatial model finds class π_4^{fa} in more locations than with the first parameter estimates. But still, the method does not manage to correctly classify the entire layers of this class. This may be due to the high proportion of class π_4^{fa} in this well, compared to the training wells, A and B. As for the first parameter estimates, the location wise model finds more of this class than the spatial model. By comparing the display of the marginal pdfs and the confusion matrices for the models in Figure 39, we can see that the spatial model is a better classification model for the small classes, and we can see that the mmap solution for this model finds some of the layer of class π_2^{fa} at approximately $t = 450$. The location wise model does not find class π_3^{fa} and hardly class π_2^{fa} . With signal-to-noise ratios of 1.30 for the gammaray log and 1.09 for the density log, the share of matches of 0.6183 and 0.7266 are good, compared to the results from the first parameter estimation. In accordance with the mmap solution, the simulations from the spatial model in Figure 40 finds more of class π_4^{fa} than for the first parameter estimates.

Summary of the results for the facies associations

In a lot of the instances above the location wise model had a better share of matches than the spatial model. This can be explained by the good match the location wise model had in the large classes, even though the model had considerable problems finding the small classes. The spatial model found more of the small classes, and in addition, it managed to find some of the challenging layers that was introduced in the beginning. The location wise model did not find any of them. The location wise model was very affected by the noise in the data, which is reflected in the display of the marginal posterior pdfs and the simulations. The probability curves in the marginal posterior pdfs are not as smooth as for the spatial model, which leads to very unstable simulations.

The spatial structure in the prior transition probability matrix in expression (20), is kept in the simulations. This means that the transitions corresponding to the small entries of the transition probability matrix will occur less than the transitions corresponding to the large probabilities. The simulations from the location wise model does not have this structure. In order to show that this structure is present in the simulations from the spatial model only, transition probability matrices are calculated from all simulations from both models. The transition probability matrix calculated from the simulations from the spatial model is given by

$$\hat{Q}^{fa} = \begin{bmatrix} 0.9607 & 0.0131 & 0.0165 & 0.0097 \\ 0.0144 & 0.9172 & 0.0183 & 0.0501 \\ 0.0626 & 0.0376 & 0.8859 & 0.0140 \\ 0.0278 & 0.0113 & 0.0078 & 0.9530 \end{bmatrix},$$

and we notice that this matrix is similar to the transition probability matrix in expression (20). The transition probability matrix calculated from the simulations from the location wise model is given by

$$\hat{Q}_{lw}^{fa} = \begin{bmatrix} 0.8751 & 0.0884 & 0.0037 & 0.0327 \\ 0.2120 & 0.3104 & 0.0614 & 0.4162 \\ 0.0483 & 0.2310 & 0.1069 & 0.6138 \\ 0.0449 & 0.2229 & 0.0904 & 0.6418 \end{bmatrix},$$

which does not have the structure of expression (20) at all.

Since the observations from class π_1^{fa} and π_3^{fa} separates into different clusters in Figure

10, we expected that class π_1^{fa} and π_3^{fa} should be easy to separate from each other, which the results from the classifications confirm.

The results from the two different parameter estimates alternated in giving the best share of matches for the three wells. The second parameter estimates gave best match for the test well, C, which we did not expect since the signal-to-noise ratios are smaller with these parameters. The results from second parameter estimates had a tendency to over-classify class π_3^{fa} . Apart from this, it is difficult to point out one of the parameter estimates to be the best for the spatial model.

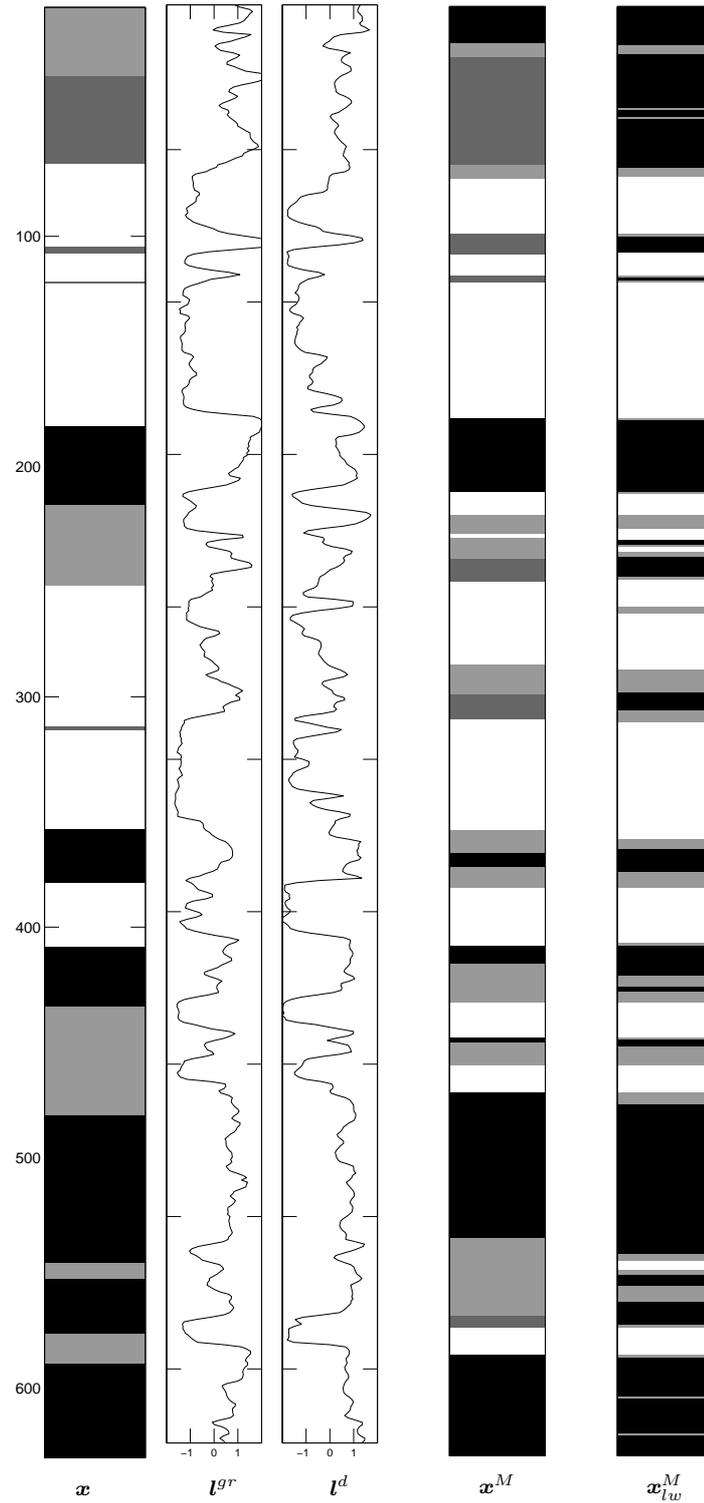


Figure 23: Results from the first parameter estimation (visual inspection). From left is the core classification of well A into facies associations, π^{fa} followed by the gammaray log, l^{gr} , density log, l^d , and mmap solutions, \mathbf{x}^M and \mathbf{x}_{lw}^M , for the spatial and location wise model respectively. The signal-to-noise ratio is 1.15 for the gammaray log and 1.03 for the density log. The gray scale levels of the respective classes are given in Figure 4, and we have that class π_1^{fa} is white and class π_4^{fa} is black.

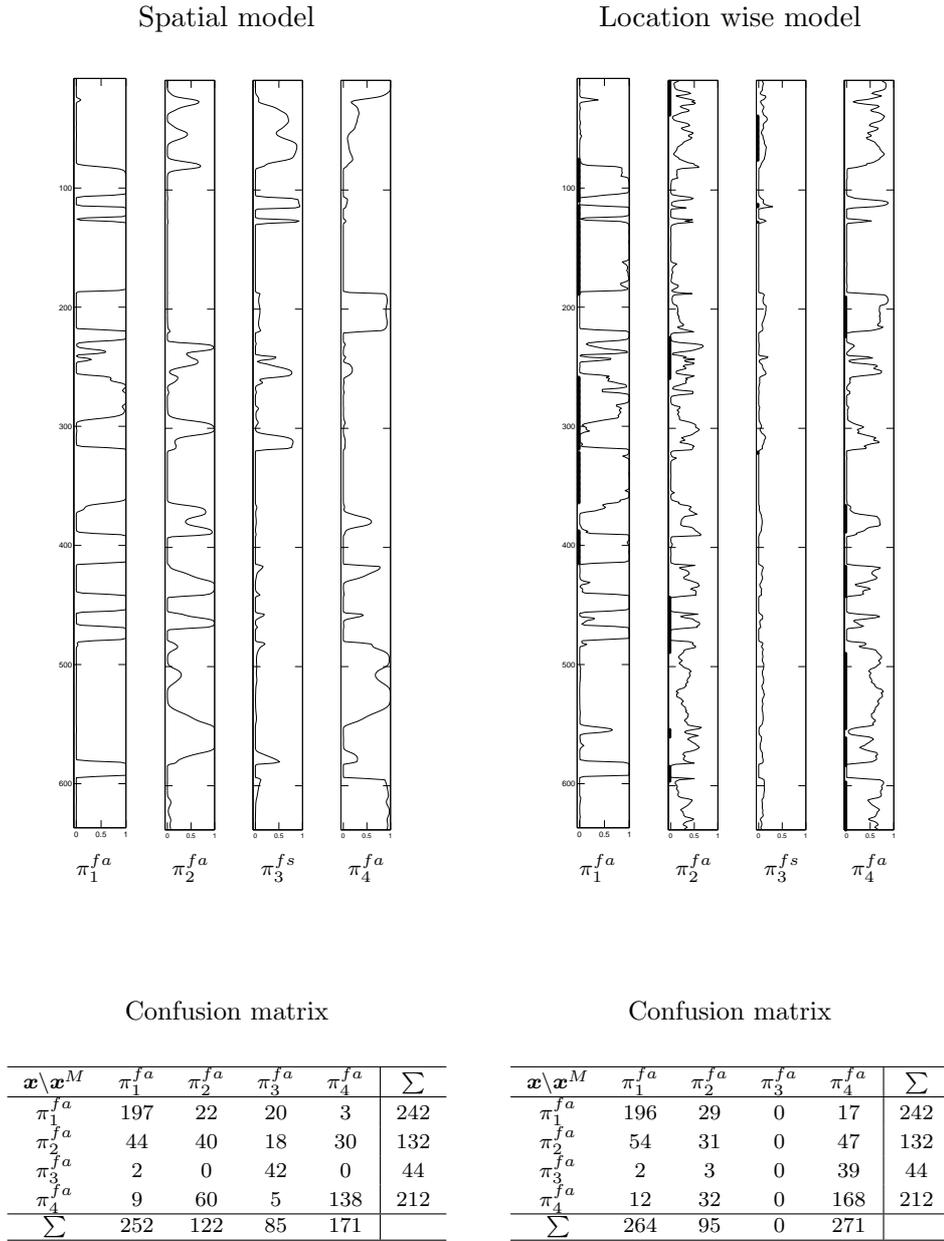


Figure 24: Results from the first parameter estimation (visual inspection). From the top the marginal posterior pdf for well A, $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$, is displayed for all t and all classes of facies associations, π^{fa} , for the spatial model (left) and the location wise model (right). The presence of the classes of facies associations at every location is marked on the axis of the respective plots. Further confusion matrices and share of matches for the two models.

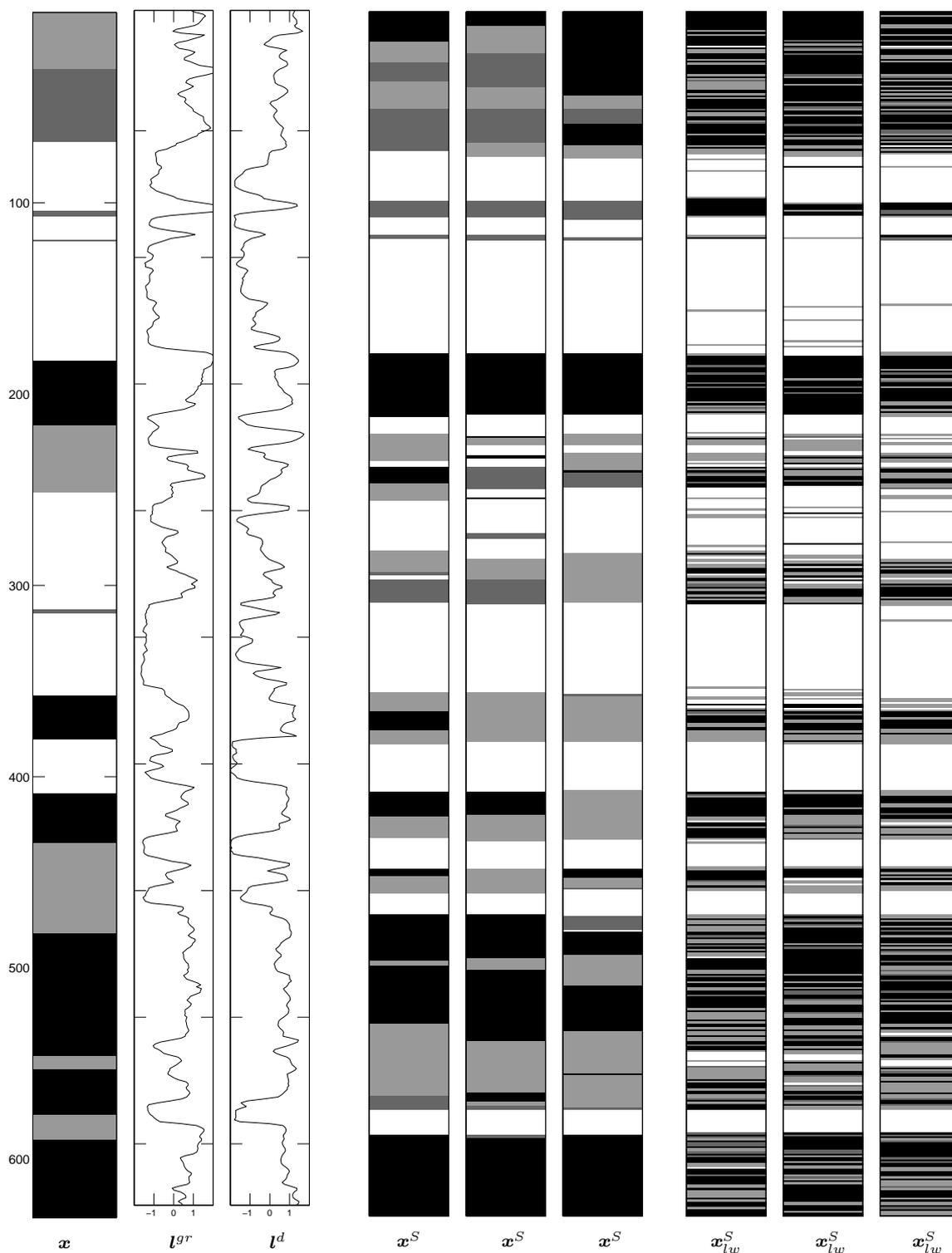


Figure 25: Results from the first parameter estimation (visual inspection). From left is the core classification of well A into facies associations, π^{fa} , followed by the gamma-ray log, l^{gr} , and density log, l^d . Further, three simulations from the posterior pdf for the spatial model, \mathbf{x}^S , and location wise model \mathbf{x}_{lw}^S . The signal-to-noise ratio is 1.15 for the gamma-ray log and 1.03 for the density log. The gray scale levels of the respective classes are given in Figure 4, and we have that class π_1^{fa} is white and class π_4^{fa} is black.

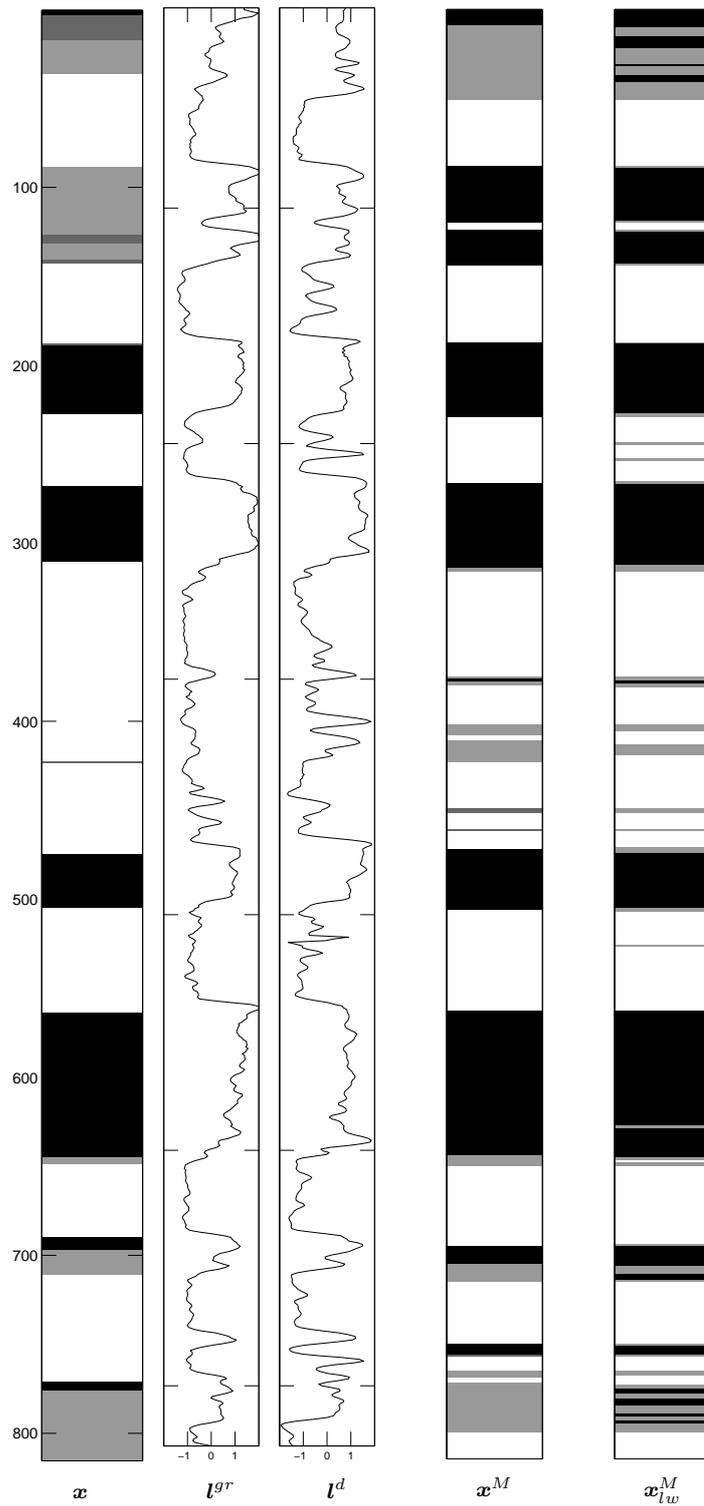


Figure 26: Results from the first parameter estimation (visual inspection). From left is the core classification of well B into facies associations, π^{fa} followed by the gamma-ray log, l^{gr} , density log, l^d , and mmap solutions, x^M and x_{lw}^M , for the spatial and location wise model respectively. The signal-to-noise ratio is 2.60 for the gamma-ray log and 1.29 for the density log. The gray scale levels of the respective classes are given in Figure 4, and we have that class π_1^{fa} is white and class π_4^{fa} is black.

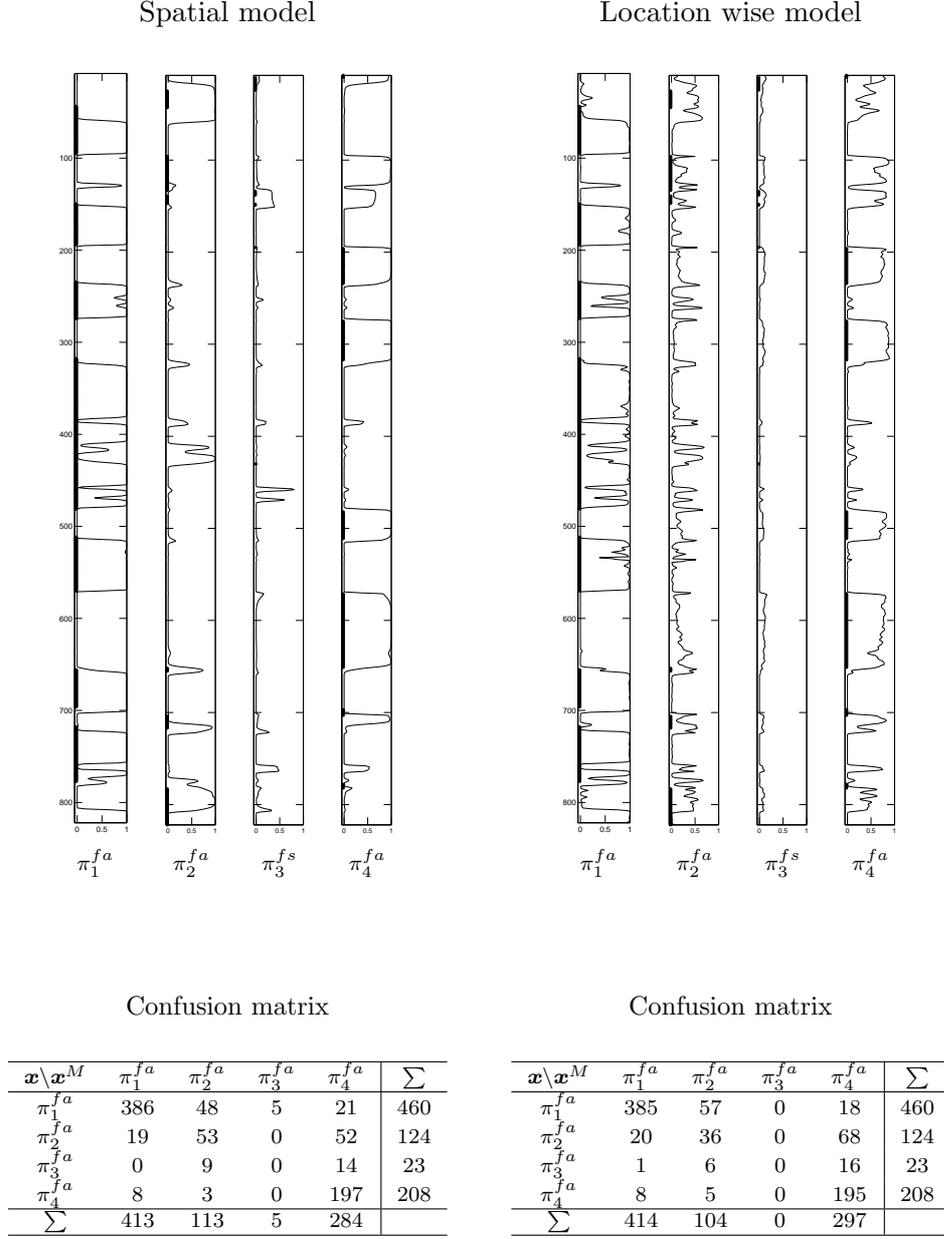


Figure 27: Results from the first parameter estimation (visual inspection). From the top the marginal posterior pdf for well B, $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$, is displayed for all t and all classes of facies associations, π^{fa} , for the spatial model (left) and the location wise model (right). The presence of the classes of facies associations at every location is marked on the axis of the respective plots. Further confusion matrices and share of matches for the two models.

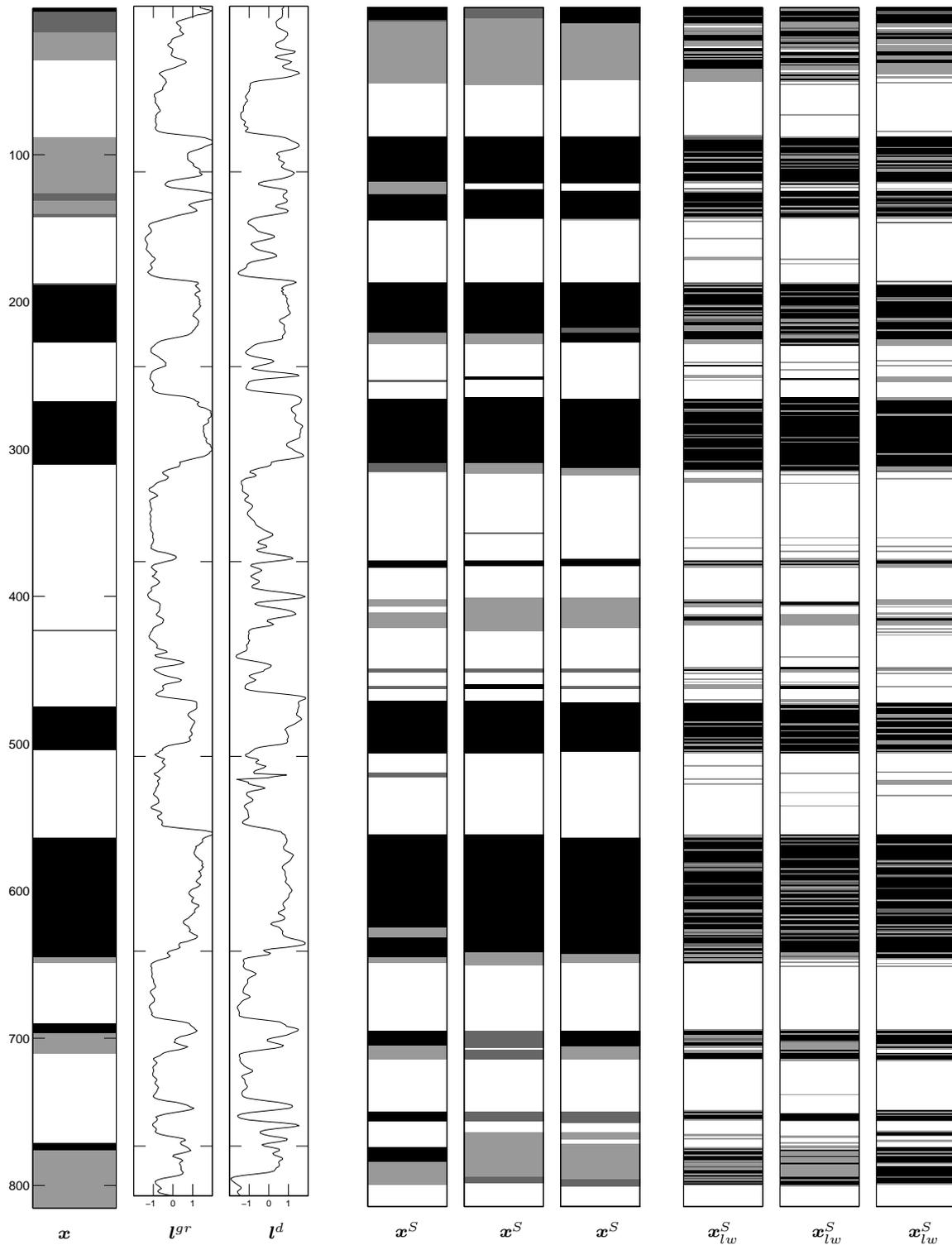


Figure 28: Results from the first parameter estimation (visual inspection). From left is the core classification of well B into facies associations, π^{fa} , followed by the gamma-ray log, l^{gr} , and density log, l^d . Further, three simulations from the posterior pdf for the spatial model, x^S , and location wise model x_{lw}^S . The signal-to-noise ratio is 2.60 for the gamma-ray log and 1.29 for the density log. The gray scale levels of the respective classes are given in Figure 4, and we have that class π_1^{fa} is white and class π_4^{fa} is black.

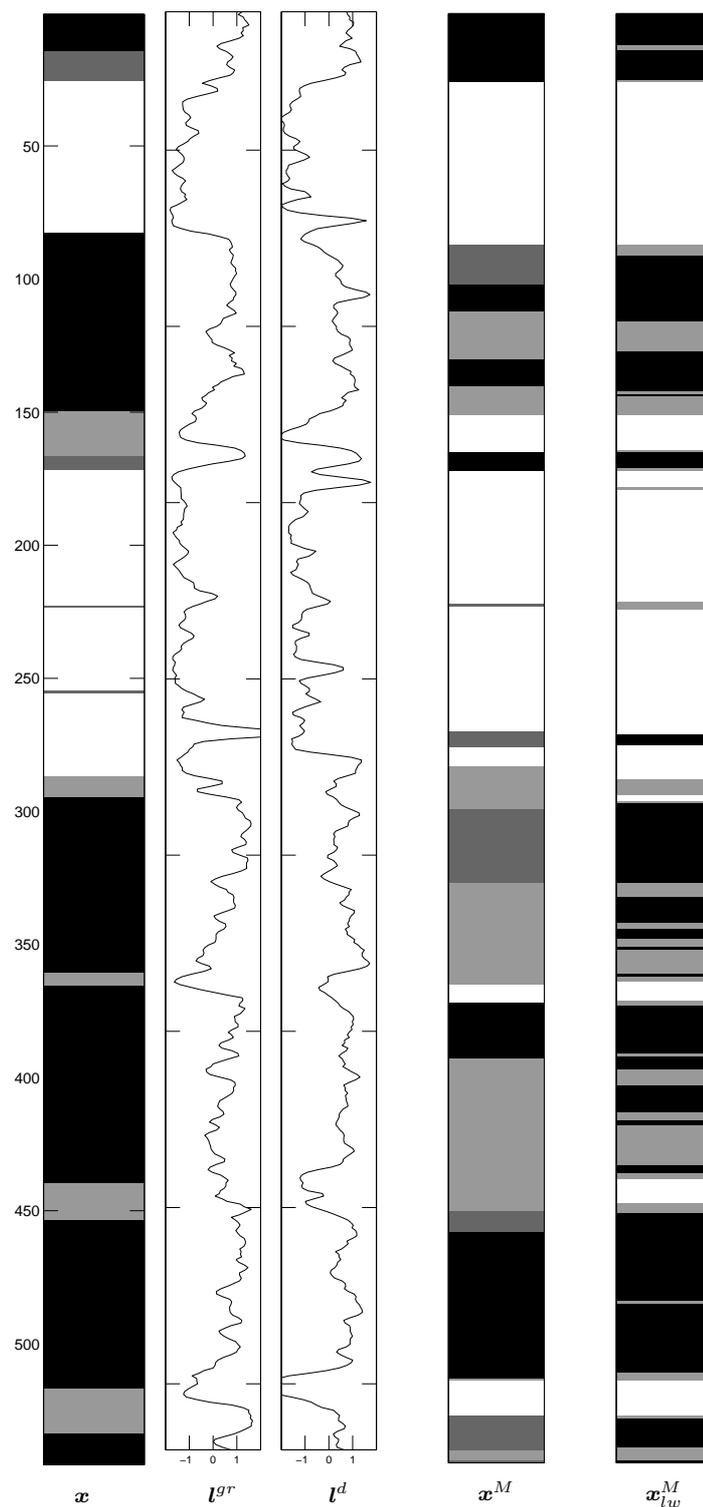


Figure 29: Results from the first parameter estimation (visual inspection). From left is the core classification of well C into facies associations, π^{fa} followed by the gamma-ray log, l^{gr} , density log, l^d , and mmap solutions, x^M and x_{lw}^M , for the spatial and location wise model respectively. The signal-to-noise ratio is 1.55 for the gamma-ray log and 1.28 for the density log. The gray scale levels of the respective classes are given in Figure 4, and we have that class π_1^{fa} is white and class π_4^{fa} is black.

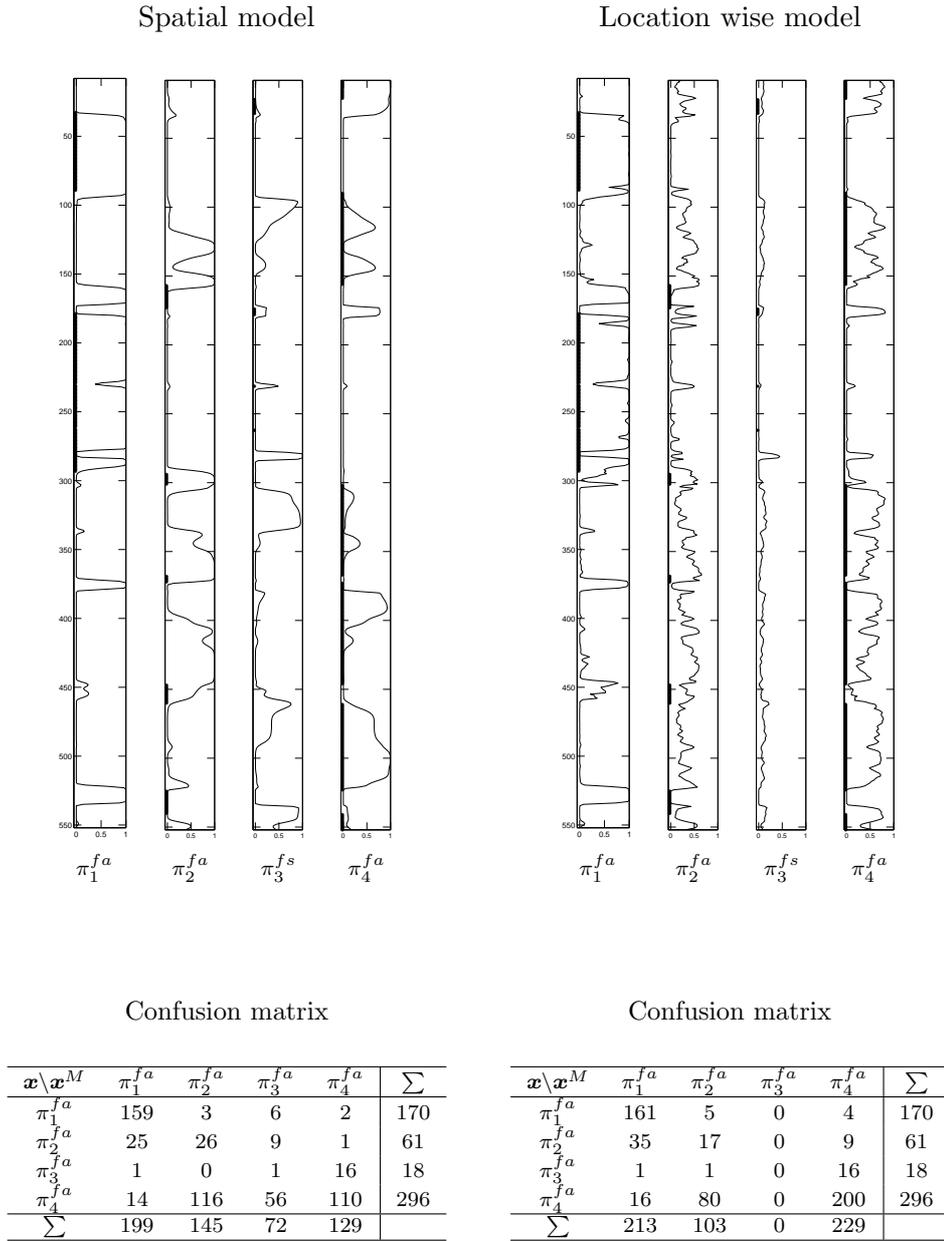


Figure 30: Results from the first parameter estimation (visual inspection). From the top the marginal posterior pdf for well C, $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$, is displayed for all t and all classes of facies associations, π^{fa} , for the spatial model (left) and the location wise model (right). The presence of the classes of facies associations at every location is marked on the axis of the respective plots. Further confusion matrices and share of matches for the two models.

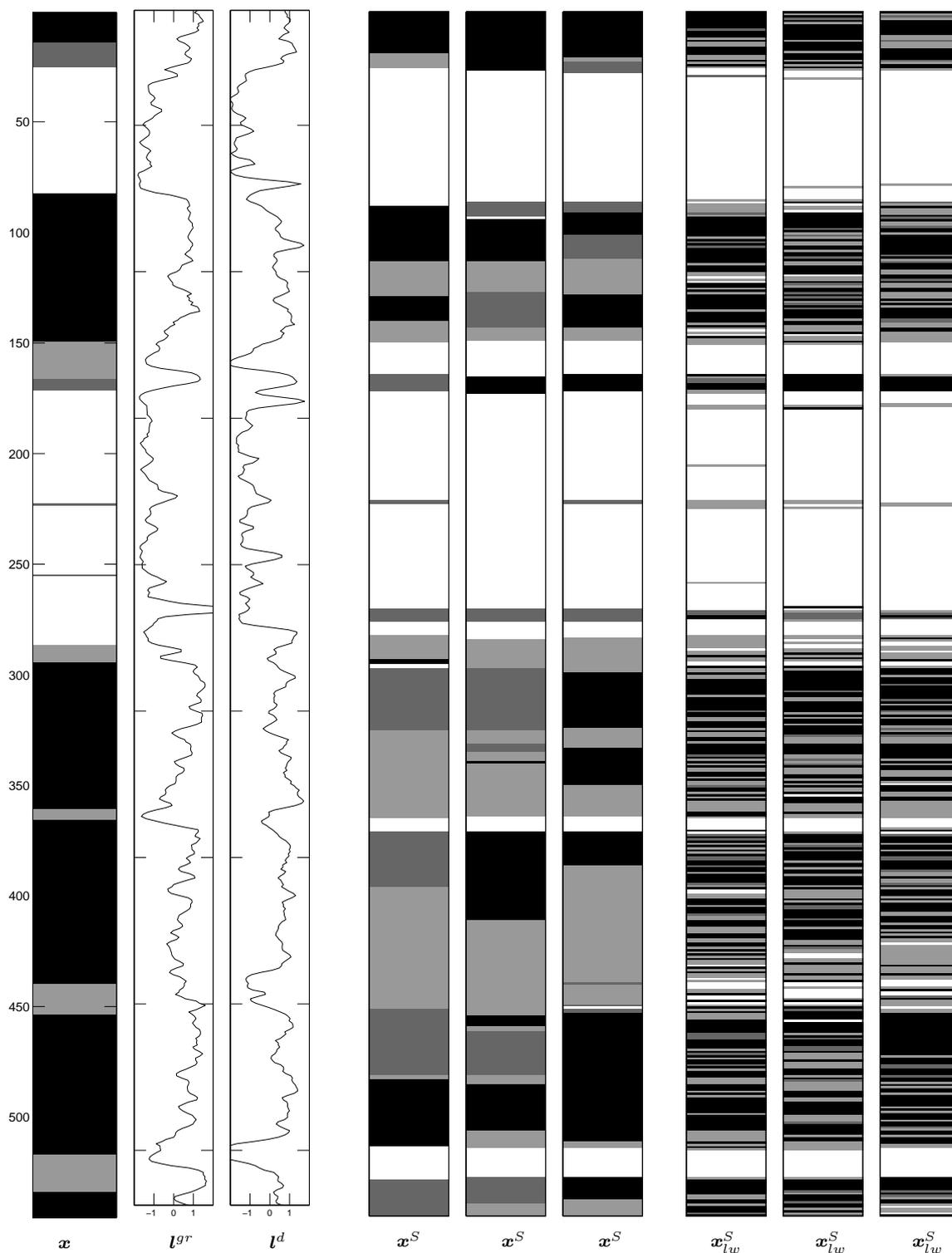


Figure 31: Results from the first parameter estimation (visual inspection). From left is the core classification of well C into facies associations, π^{fa} , followed by the gamma-ray log, l^{gr} , and density log, l^d . Further, three simulations from the posterior pdf for the spatial model, \mathbf{x}^S , and location wise model \mathbf{x}_{lw}^S . The signal-to-noise ratio is 1.55 for the gamma-ray log and 1.28 for the density log. The gray scale levels of the respective classes are given in Figure 4, and we have that class π_1^{fa} is white and class π_4^{fa} is black.

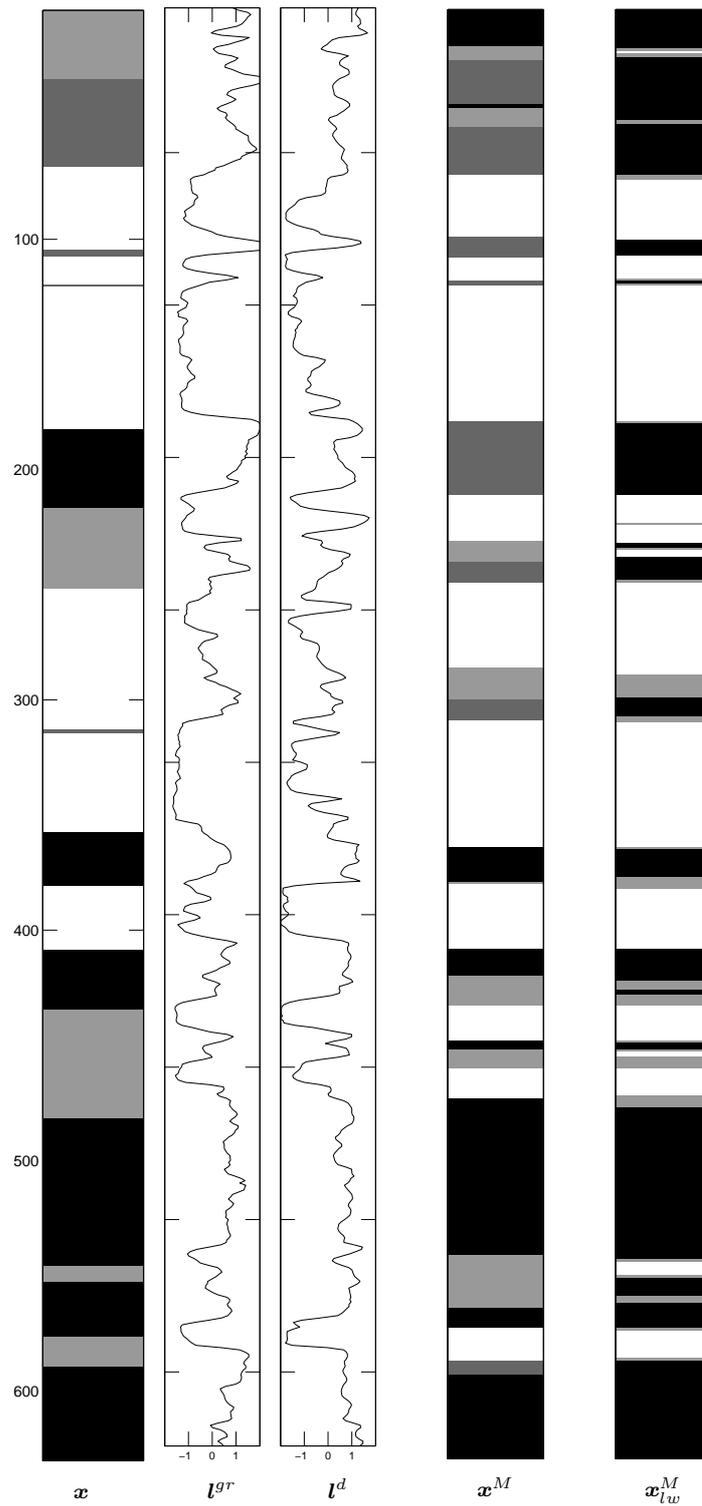


Figure 32: Results from the second parameter estimation (least squares). From left is the core classification of well A into facies associations, π^{fa} followed by the gammaray log, l^{gr} , density log, l^d , and mmap solutions, \mathbf{x}^M and \mathbf{x}_{lw}^M , for the spatial and location wise model respectively. The signal-to-noise ratio is 1.03 for the gammaray log and 0.80 for the density log. The gray scale levels of the respective classes are given in Figure 4, and we have that class π_1^{fa} is white and class π_4^{fa} is black.

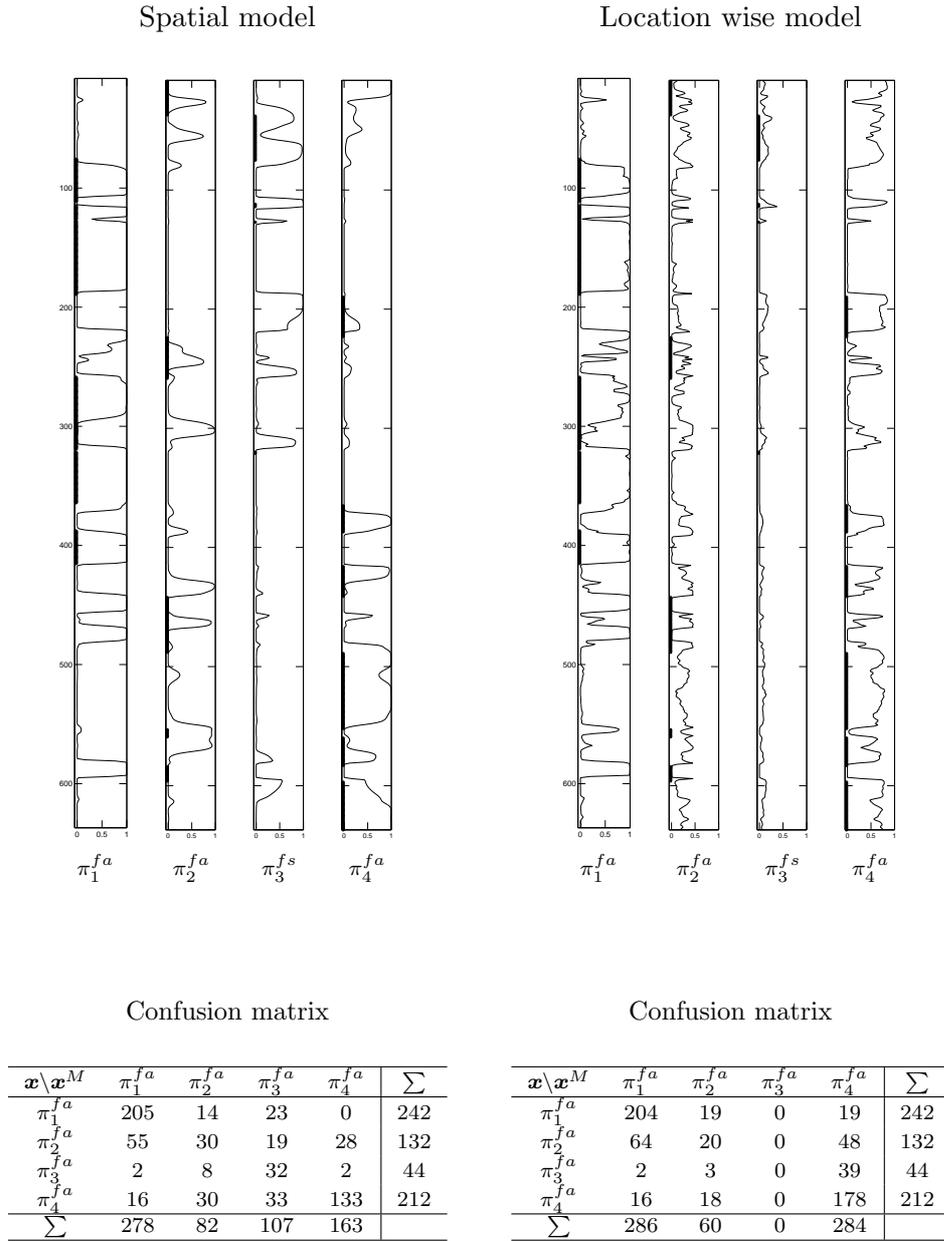


Figure 33: Results from the second parameter estimation (least squares). From the top the marginal posterior pdf for well A, $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$, is displayed for all t and all classes of facies associations, π^{fa} , for the spatial model (left) and the location wise model (right). The presence of the classes of facies associations at every location is marked on the axis of the respective plots. Further confusion matrices and share of matches for the two models.

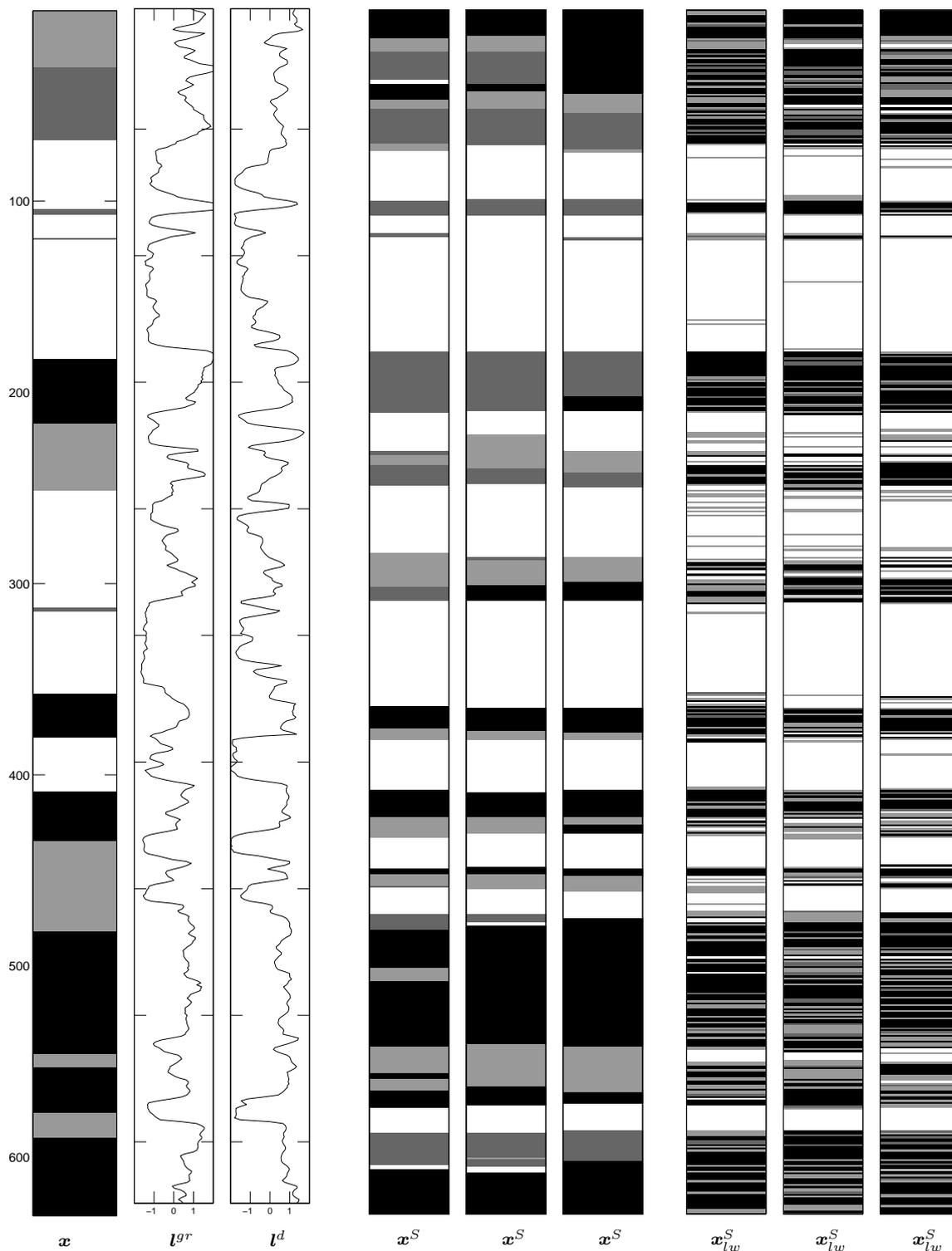


Figure 34: Results from the second parameter estimation (least squares). From left is the core classification of well A into facies associations, π^{fa} , followed by the gamma-ray log, l^{gr} , and density log, l^d . Further, three simulations from the posterior pdf for the spatial model, \mathbf{x}^S , and location wise model \mathbf{x}_{lw}^S . The signal-to-noise ratio is 1.03 for the gamma-ray log and 0.80 for the density log. The gray scale levels of the respective classes are given in Figure 4, and we have that class π_1^{fa} is white and class π_4^{fa} is black.

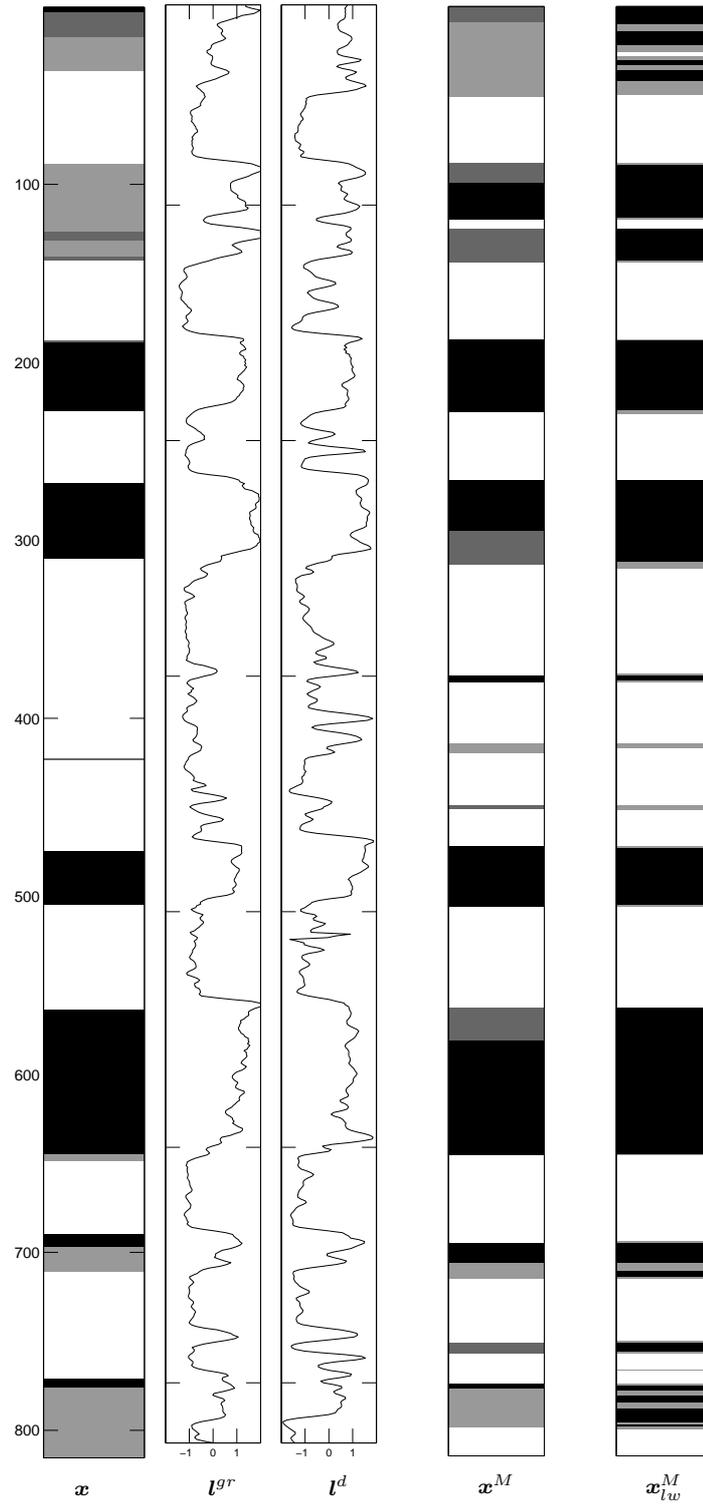


Figure 35: Results from the second parameter estimation (least squares). From left is the core classification of well B into facies associations, π^{fa} followed by the gammaray log, l^{gr} , density log, l^d , and mmap solutions, \mathbf{x}^M and \mathbf{x}_{lw}^M , for the spatial and location wise model respectively. The signal-to-noise ratio is 1.93 for the gammaray log and 0.92 for the density log. The gray scale levels of the respective classes are given in Figure 4, and we have that class π_1^{fa} is white and class π_4^{fa} is black.

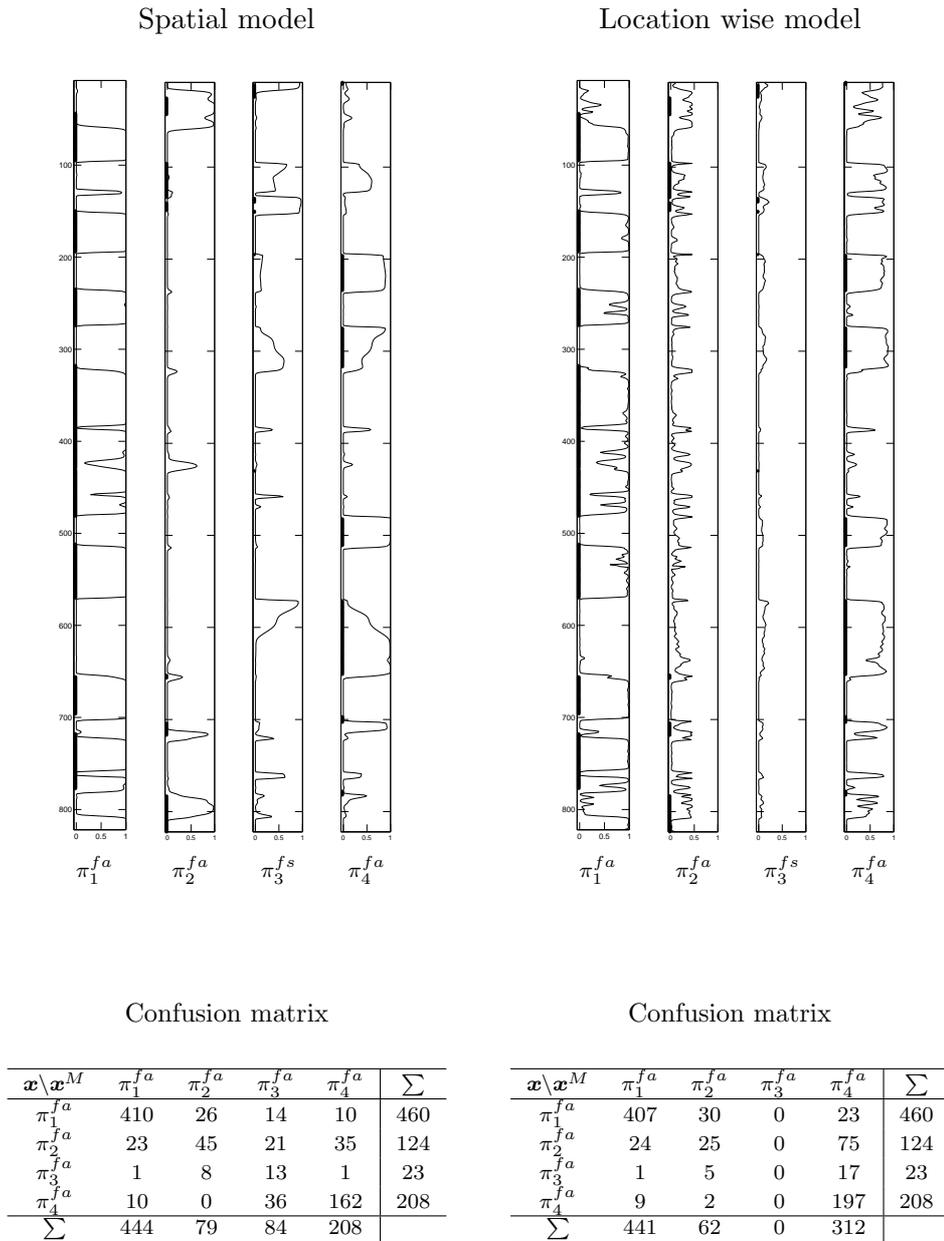


Figure 36: Results from the second parameter estimation (least squares). From the top the marginal posterior pdf for well B, $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$, is displayed for all t and all classes of facies associations, π^{fa} , for the spatial model (left) and the location wise model (right). The presence of the classes of facies associations at every location is marked on the axis of the respective plots. Further confusion matrices and share of matches for the two models.

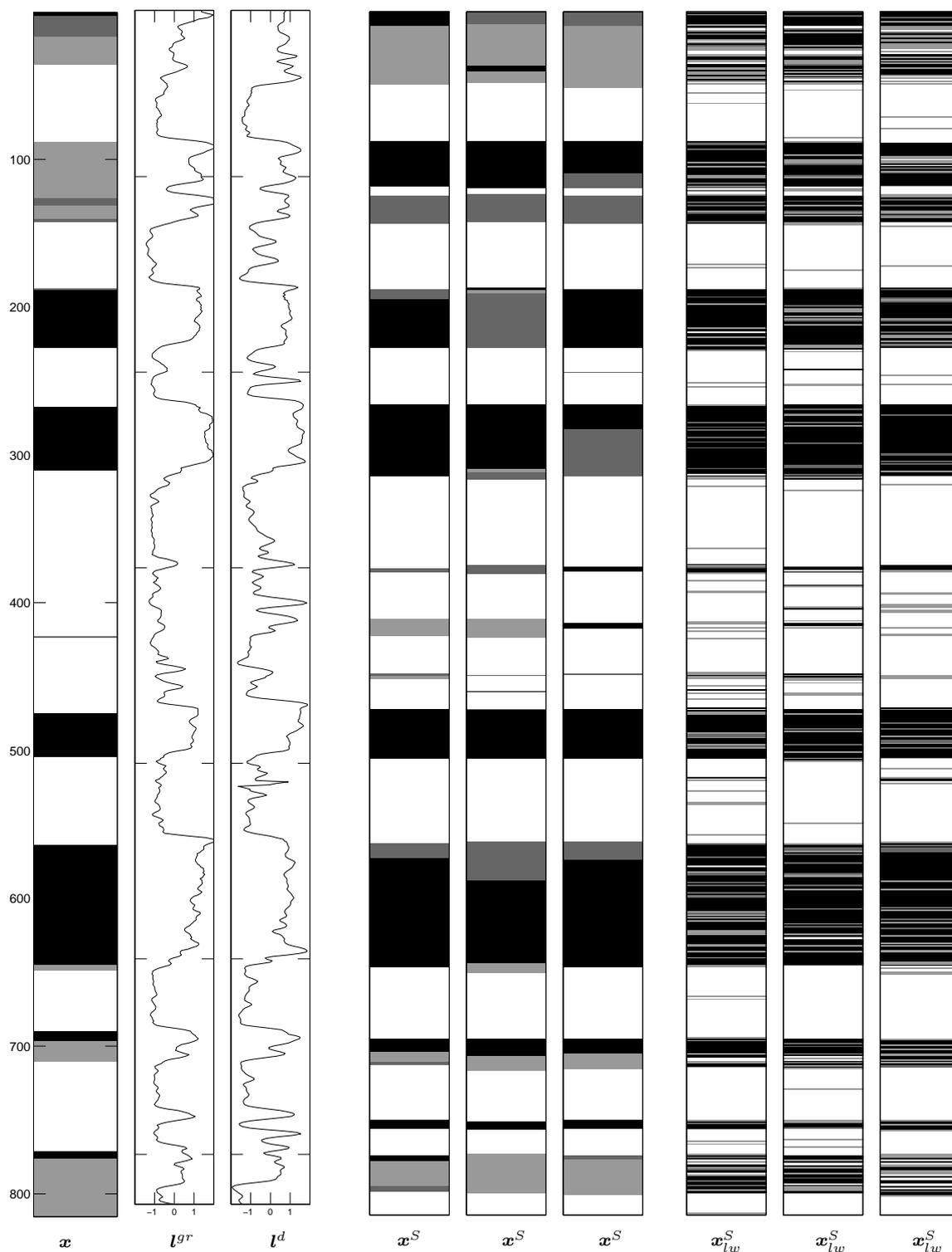


Figure 37: Results from the second parameter estimation (least squares). From left is the core classification of well B into facies associations, π^{fa} , followed by the gamma-ray log, l^{gr} , and density log, l^d . Further, three simulations from the posterior pdf for the spatial model, \mathbf{x}^S , and location wise model \mathbf{x}_{lw}^S . The signal-to-noise ratio is 1.93 for the gamma-ray log and 0.92 for the density log. The gray scale levels of the respective classes are given in Figure 4, and we have that class π_1^{fa} is white and class π_4^{fa} is black.

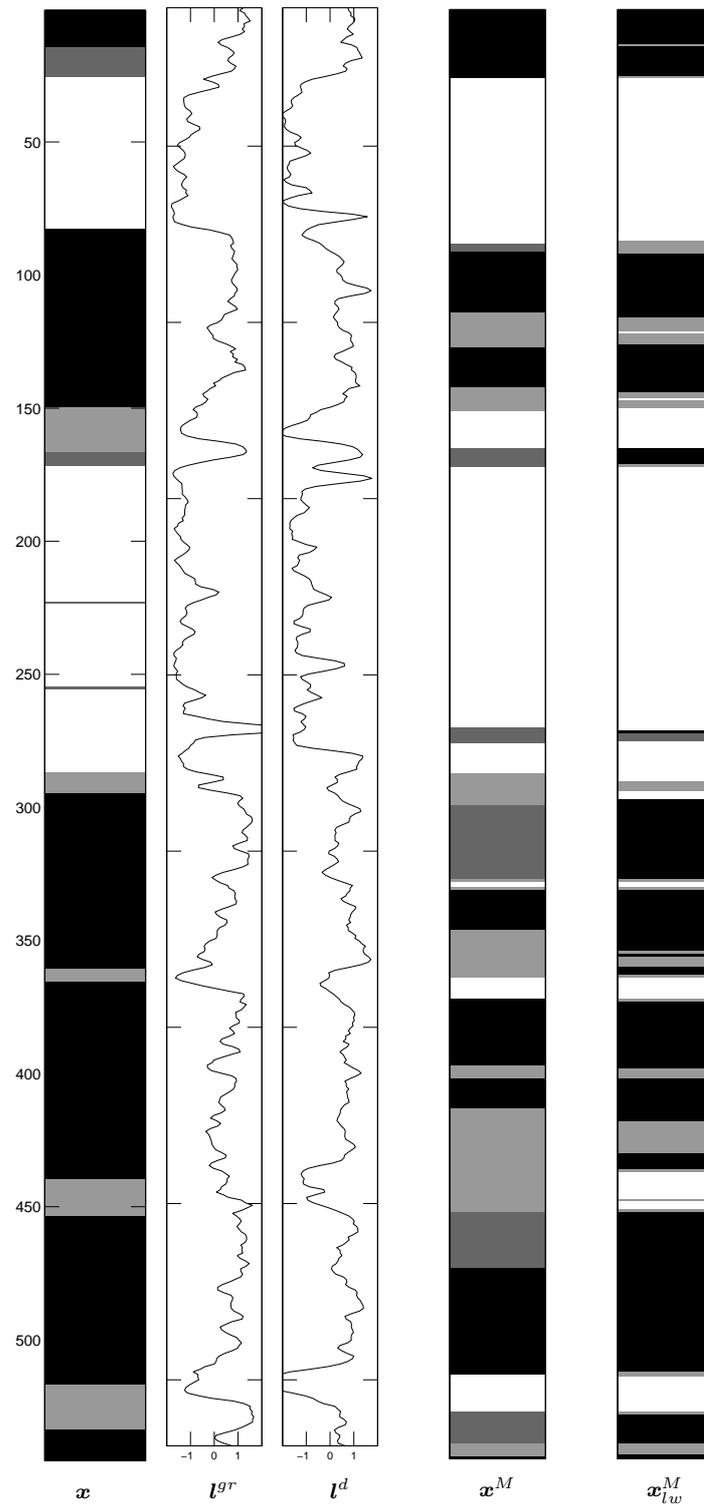


Figure 38: Results from the second parameter estimation (least squares). From left is the core classification of well C into facies associations, π^{fa} followed by the gamma-ray log, l^{gr} , density log, l^d , and mmap solutions, x^M and x_{lw}^M , for the spatial and location wise model respectively. The signal-to-noise ratio is 1.30 for the gamma-ray log and 1.09 for the density log. The gray scale levels of the respective classes are given in Figure 4, and we have that class π_1^{fa} is white and class π_4^{fa} is black.

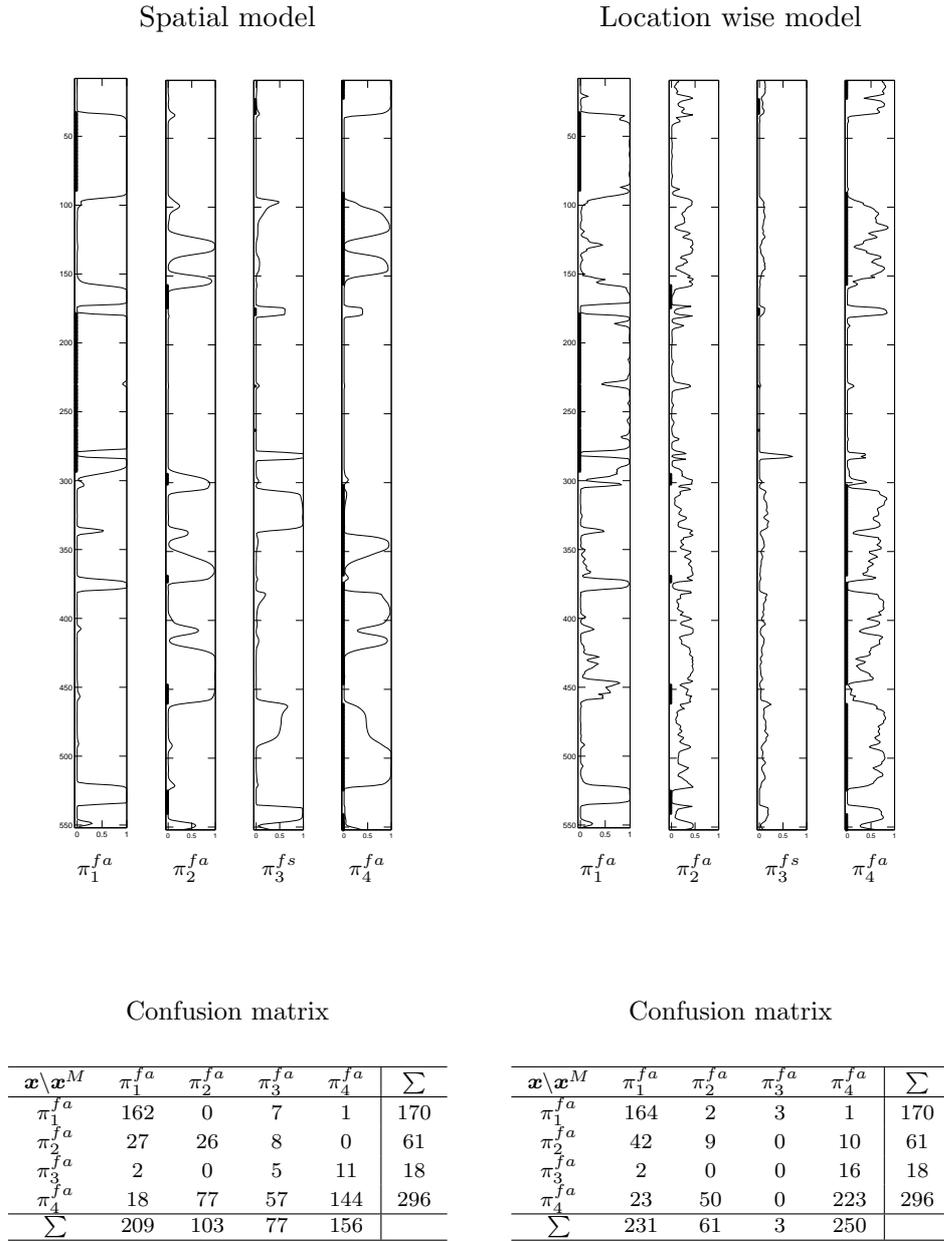


Figure 39: Results from the second parameter estimation (least squares). From the top the marginal posterior pdf for well C, $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$, is displayed for all t and all classes of facies associations, π^{fa} , for the spatial model (left) and the location wise model (right). The presence of the classes of facies associations at every location is marked on the axis of the respective plots. Further confusion matrices and share of matches for the two models.

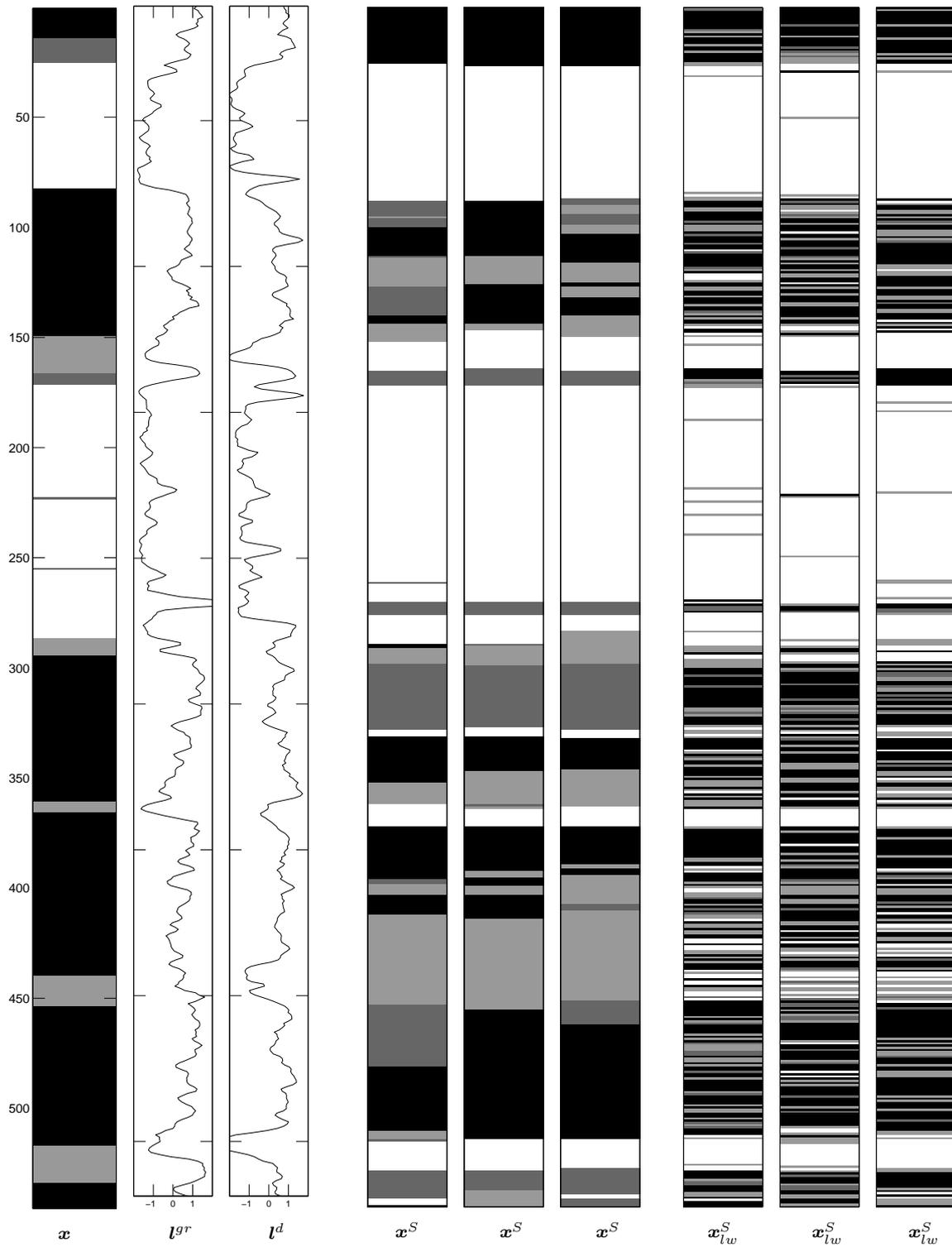


Figure 40: Results from the second parameter estimation (least squares). From left is the core classification of well C into facies associations, π^{fa} , followed by the gamma-ray log, l^{gr} , and density log, l^d . Further, three simulations from the posterior pdf for the spatial model, \mathbf{x}^S , and location wise model \mathbf{x}_{lw}^S . The signal-to-noise ratio is 1.30 for the gamma-ray log and 1.09 for the density log. The gray scale levels of the respective classes are given in Figure 4, and we have that class π_1^{fa} is white and class π_4^{fa} is black.

6.2 Lithofacies

The challenge in the classification of lithofacies is that there are a larger number of classes than for the facies associations, which makes the layers thinner and the proportion of some of the classes very small. Especially in the test well, C, we notice that the proportions of class π_1^{lf} , π_4^{lf} and π_5^{lf} are very small, as we can see from Figure 7. We expect the small classes to be hard to find. We also expect the models to have problems separating the four classes of sand of different grain size, $\pi_1^{lf} - \pi_4^{lf}$, from each other.

The results from the classification of lithofacies are presented in Figure 41-58. First the results from the first parameter estimation is considered, followed by the results from the second parameter estimation.

Likelihood parameters estimated based on visual inspection

The results from the first parameter estimation for the lithofacies are given in Figure 41-43 for the first training well, A, in Figure 44-46 for the second training well, B, and in Figure 47-49 for the test well, C.

The mmap solution for the spatial model for the first training well, A, in Figure 41 seems to have all the six classes represented, as apposed to the location wise model. From the display of the marginal posterior pdfs in Figure 42 we notice that for the location wise model the pattern of the probabilities for class π_1^{lf} , π_2^{lf} and π_3^{lf} are almost similar, with the probabilities of class π_2^{lf} a bit larger than the others. From the confusion matrix it seems that the location wise model is only capable of separating sand from not sand, since the mmap solution nearly almost consists of class π_2^{lf} and π_6^{lf} . The spatial model is better to separate the four classes of sand, $\pi_1^{lf} - \pi_4^{lf}$, from each other. We notice that both models have problems finding the correct locations of class π_4^{lf} , and also the smallest class in this well, class π_5^{lf} . In addition, the location wise model does not find class π_1^{lf} at all. From the simulations in Figure 43 we can see that the spatial model gives better simulations, since the noise is a big disturbance in the simulations from the location wise model.

None of the mmap solutions for the second training well, B, in Figure 44 seems to find class π_1^{lf} . From the display of the marginal posterior pdfs and the confusion matrices in Figure 45, we notice that none of the models find class π_5^{lf} either. Again we can see that the location wise model does not manage to separate the first three classes of sand, π_1^{lf} ,

π_2^{lf} and π_3^{lf} from each other, and we therefore get the same sand/not sand classification as in the first training well. The spatial model manages to find some of both class π_2^{lf} and π_3^{lf} , but still, the share of matches is smaller than for the location wise model. From the simulations in Figure 46 we notice that both models find it hard to separate the classes of sand, $\pi_1^{lf} - \pi_4^{lf}$, from each other. The the spatial model manages to find the correct class of sand in relatively many locations, but the location wise model on the other hand, is too sensitive to the noise in the log data.

From the mmap solutions for the test well, C, in Figure 47 we can see that the location wise model again has almost exclusively class π_2^{lf} and π_6^{lf} in the mmap solution. These are the largest classes in the test well, as we can see from Figure 7. The spatial model has all six classes represented in the mmap solution, but from the display of the marginal posterior pdfs and the confusion matrices in Figure 48 we notice that the model over-classifies both class π_1^{lf} and π_4^{lf} . Since the location wise model has such large proportions of the biggest classes, π_2^{lf} and π_6^{lf} , in the mmap solution, it gets the largest share of matches, of 0.7104, compared to a share of matches of 0.6258 for the spatial model. In the simulations in Figure 49 we can see that class π_1^{lf} is over-represented in both models. The simulations from the location wise model are also unstable due to noise, and the problem in this well also seems to be separating the classes of sand from each other.

Likelihood parameters estimated by minimizing the squared error

The results from the second parameter estimation are given in Figure 50-52 for the first training well, A, in Figure 53-55 for the second training well, B, and 56-58 for the test well, C. The signal-to-noise ratios are smaller with these parameters, which may lead to poorer results.

The mmap solution for the first training well, A, in Figure 50 for the location wise model seems to have more classes represented in this parameter estimation, but still we can see that the spatial model finds class π_1^{lf} better than the location wise model. The display of the marginal posterior pdfs and the confusion matrix for the location wise model in Figure 51 tell the same story as in the first parameter estimation, the model is only capable of separating sand from not sand. Apart from the spatial model finding class π_1^{lf} in more locations than the location wise model, the spatial model has problems finding class π_3^{lf} , π_4^{lf} and π_5^{lf} in the correct locations, which gives the small share of matches of 0.5032. The share of matches for the location wise model is larger, 0.5508, due to the large proportion of class π_6^{lf} . The simulations from the spatial model

in Figure 52, are more stable than for the location wise model, but still, the model has problems separating the classes of sand.

By visually comparing the mmap solutions for the second training well, B, in Figure 53 with the mmap solutions from the first parameter estimation in Figure 44, these seem better because they find class π_1^{lf} . From the display of the marginal posterior pdfs and the confusion matrices in Figure 54 we can see that as apposed to the spatial model, the location wise model hardly finds class π_1^{lf} , and never finds class π_4^{lf} and π_5^{lf} at all. The location wise model has a better share of matches than the spatial model because it has larger number of correct classifications of class π_2^{lf} and π_6^{lf} . The spatial model handles the separation of the classes of sand, $\pi_1^{lf} - \pi_4^{lf}$, better than the location wise model. The simulations in Figure 55 tell the same story as before, the spatial model is less sensitive to the noise in the log data.

The mmap solutions for the test well, C, in Figure 56 have even more layers misclassified into class π_1^{lf} than for the first parameter estimation, especially for the spatial model. From the display of the marginal posterior pdfs and the confusion matrices in Figure 57 we can see that the results for the location wise model are almost identical to the results from the first parameter estimation, while the results for the spatial model are poorer. The spatial model has a smaller match at class π_2^{lf} and π_6^{lf} than the location wise model, and both models have problems correctly classifying the rest of the classes. The simulations from the spatial model in Figure 58 seem poorer than before, due to the large representation of class π_1^{lf} . The simulations from the location wise model are as unstable as always.

Summary of the results for the lithofacies

As for the facies associations, the location wise model gave better share of matches than the spatial model in many of the instances, due to the good match in the large classes. Still, the posterior pdf from the spatial model managed to separate the four classes of sand, $\pi_1^{lf} - \pi_4^{lf}$, better than the location wise model. We noticed that for the location wise model, the probabilities in the marginal posterior pdfs are below 0.5 nearly everywhere for all classes except class π_6^{lf} , almost like it protected itself for making total misclassifications. The spatial model has larger probabilities for the class the model classifies in the respective locations, including for the misclassifications.

Notice that the simulations from the spatial model keep the same spatial structure that

was specified in the prior model in expression (22). This means that the transitions corresponding to the small entries of the prior transition probability matrix in expression (22), should appear less than the transitions corresponding to the large entries in the simulations. There are no such structure in the location wise model. Transition probability matrices are calculated from all simulations from both models, in order to confirm this structure. The transition probability matrix calculated from all simulations from the spatial model is given by

$$\hat{Q}^{lf} = \begin{bmatrix} 0.8868 & 0.0675 & 0.0210 & 0.0007 & 0.0082 & 0.0157 \\ 0.0293 & 0.8574 & 0.0651 & 0.0217 & 0.0096 & 0.0169 \\ 0.0234 & 0.0652 & 0.7791 & 0.0394 & 0.0172 & 0.0757 \\ 0.0097 & 0.0410 & 0.0299 & 0.8229 & 0.0042 & 0.0924 \\ 0 & 0.0752 & 0.1028 & 0.1153 & 0.6316 & 0.0752 \\ 0.0057 & 0.0151 & 0.0176 & 0.0178 & 0.0153 & 0.9284 \end{bmatrix},$$

and the transition probability matrix calculated from all simulations from the location wise model is given by

$$\hat{Q}_{lw}^{lf} = \begin{bmatrix} 0.2172 & 0.4410 & 0.2555 & 0.0407 & 0.0008 & 0.0448 \\ 0.2134 & 0.4354 & 0.2657 & 0.0371 & 0.0004 & 0.0480 \\ 0.1735 & 0.3637 & 0.2926 & 0.0544 & 0.0022 & 0.1137 \\ 0.0581 & 0.1185 & 0.1280 & 0.2382 & 0.0083 & 0.4491 \\ 0 & 0 & 0.0124 & 0.0248 & 0.0186 & 0.9441 \\ 0.0082 & 0.0235 & 0.0377 & 0.0712 & 0.0266 & 0.8328 \end{bmatrix}$$

As we can see, only the simulations from the spatial model have the structure of expression (22). The entries that are zero in expression (22) should have been zero in the matrix calculated from the simulations of the spatial model, but due to numerical instabilities, the entries are only almost zero.

Based on Figure 11 we expected the probabilities of misclassifying class π_2^{lf} as class π_6^{lf} and vice versa to be small, which the results from the classification confirm.

For the lithofacies also, the two parameter estimations alternated in giving the best share of matches, though the first parameter estimates gave the best results for the test well, C. The signal-to-noise ratio was larger for the first parameter estimates in this well, of 1.53 for the gammaray log and 1.37 for the density log, which may have been the reason.

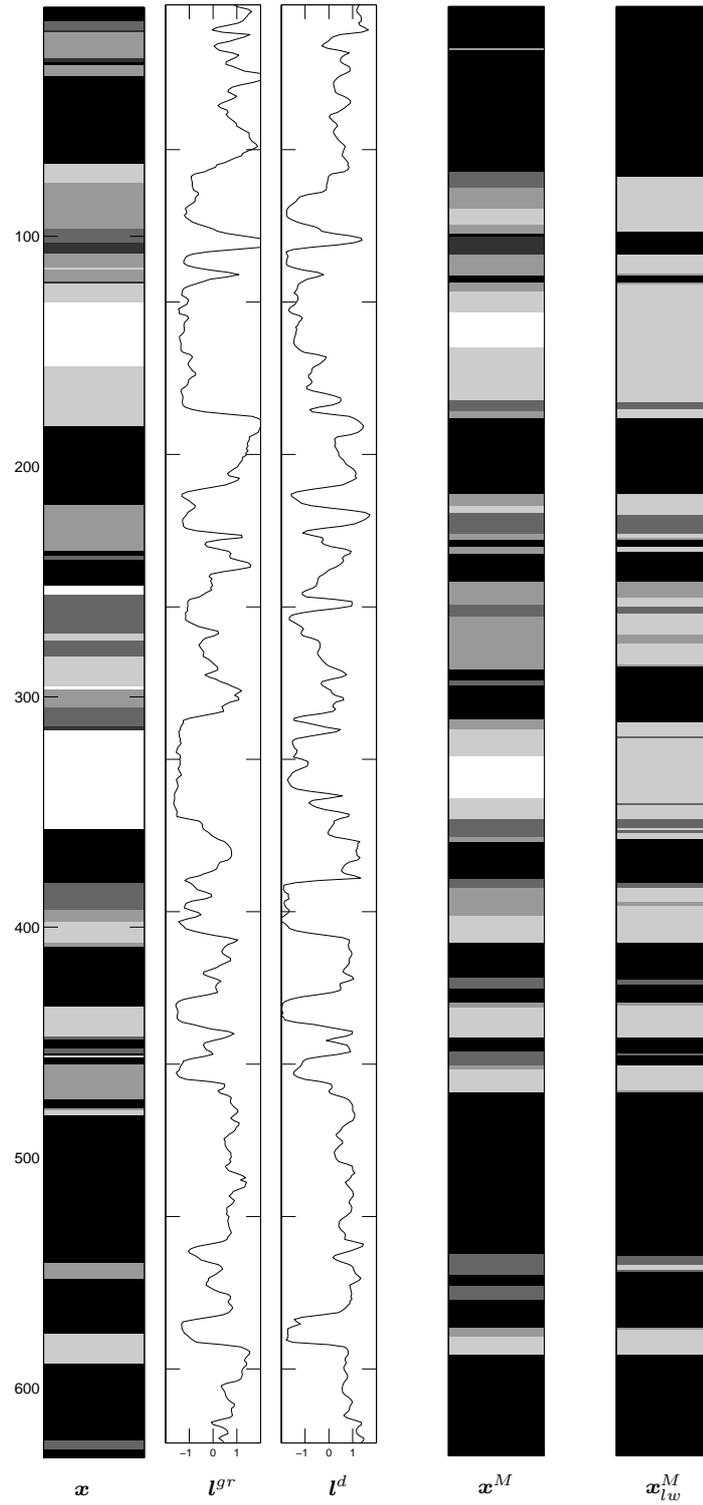


Figure 41: Results from the first parameter estimation (visual inspection). From left is the core classification of well A into lithofacies, π^{lf} followed by the gamma-ray log, l^{gr} , density log, l^d , and mmap solutions, x^M and x_{lw}^M , for the spatial and location wise model respectively. The signal-to-noise ratio is 1.64 for the gamma-ray log and 1.13 for the density log. The gray scale levels of the respective classes are given in Figure 6, and we have that class π_1^{lf} is white and class π_6^{lf} is black.

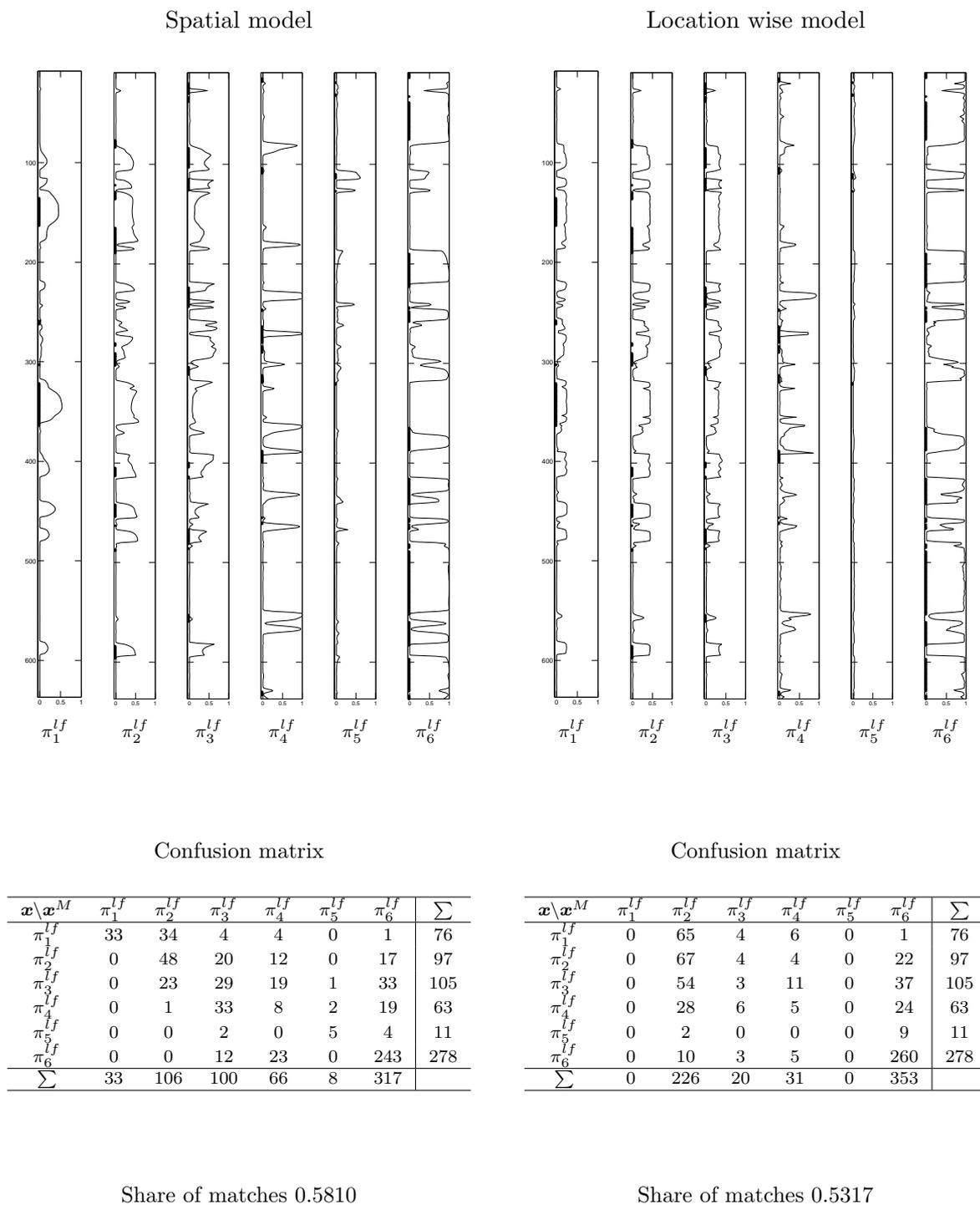


Figure 42: Results from the first parameter estimation (visual inspection). From the top the marginal posterior pdf for well A, $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$, is displayed for all t and all classes of lithofacies, π^{lf} , for the spatial model (left) and the location wise model (right). The presence of the classes of lithofacies at every location is marked on the axis of the respective plots. Further confusion matrices and share of matches for the two models.

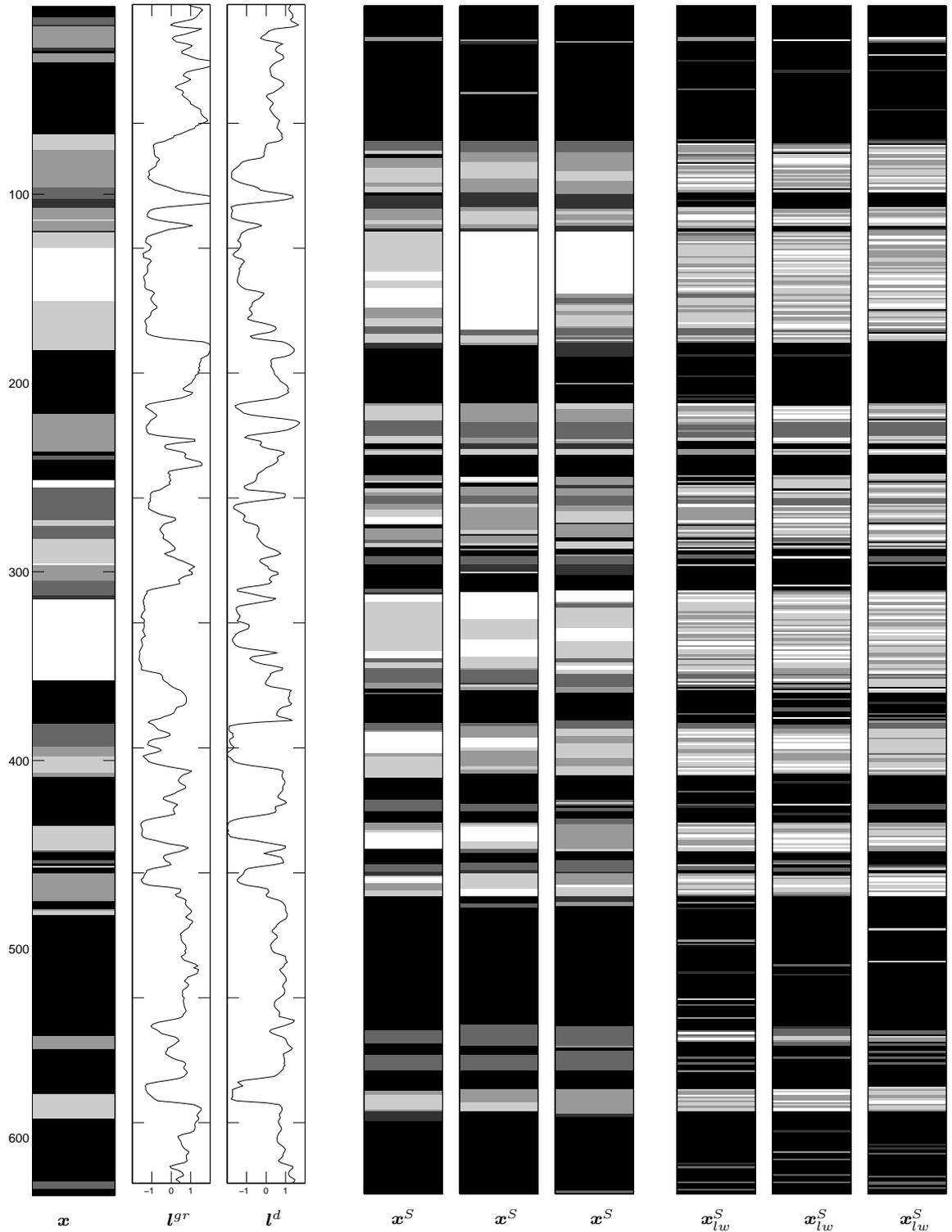


Figure 43: Results from the first parameter estimation (visual inspection). From left is the core classification of well A into lithofacies, π^{lf} , followed by the gamma-ray log, l^{gr} , and density log, l^d . Further, three simulations from the posterior pdf for the spatial model, \mathbf{x}^S , and location wise model \mathbf{x}_{lw}^S . The signal-to-noise ratio is 1.64 for the gamma-ray log and 1.13 for the density log. The gray scale levels of the respective classes are given in Figure 6, and we have that class π_1^{lf} is white and class π_6^{lf} is black.

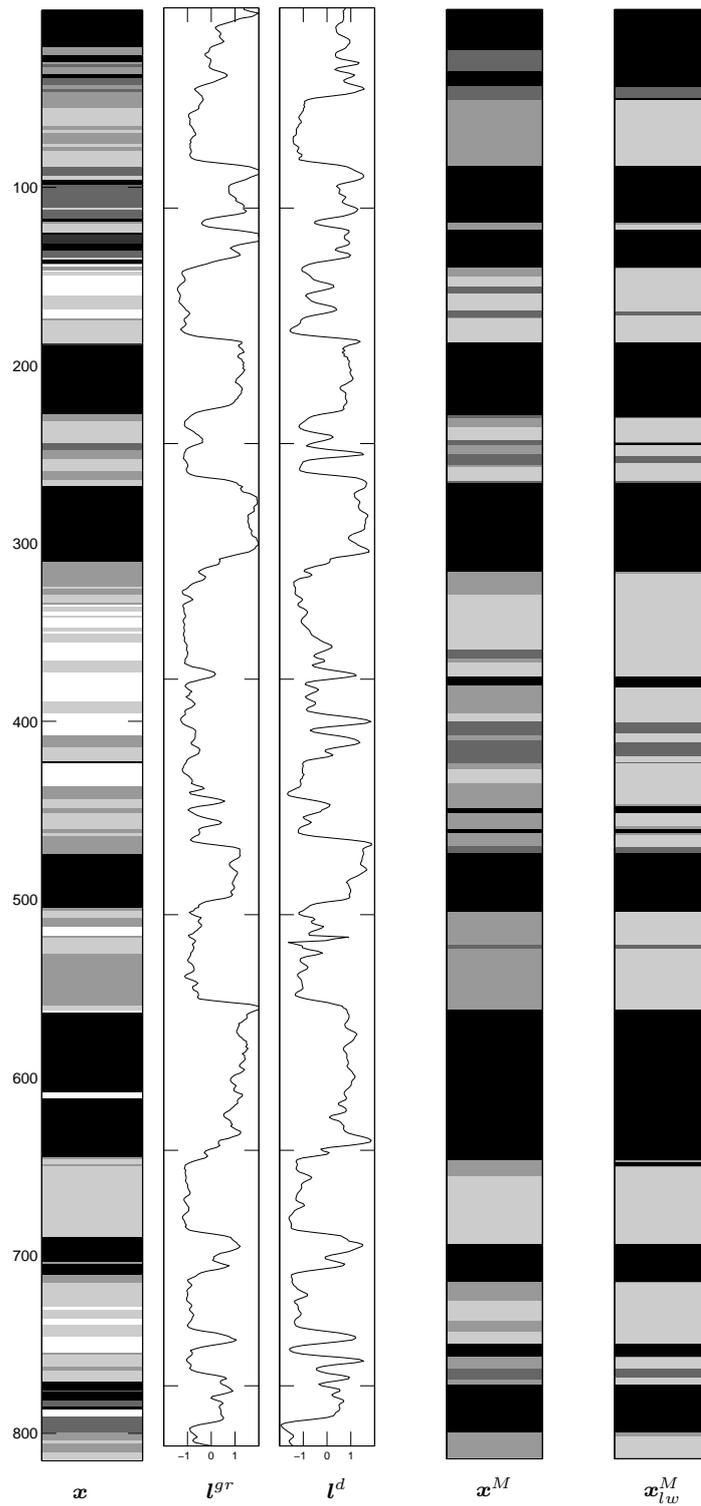


Figure 44: Results from the first parameter estimation (visual inspection). From left is the core classification of well B into lithofacies, π^{lf} followed by the gamma-ray log, l^{gr} , density log, l^d , and mmap solutions, x^M and x_{lw}^M , for the spatial and location wise model respectively. The signal-to-noise ratio is 2.32 for the gamma-ray log and 1.61 for the density log. The gray scale levels of the respective classes are given in Figure 6, and we have that class π_1^{lf} is white and class π_6^{lf} is black. In the locations that are hatched the core classification is missing.

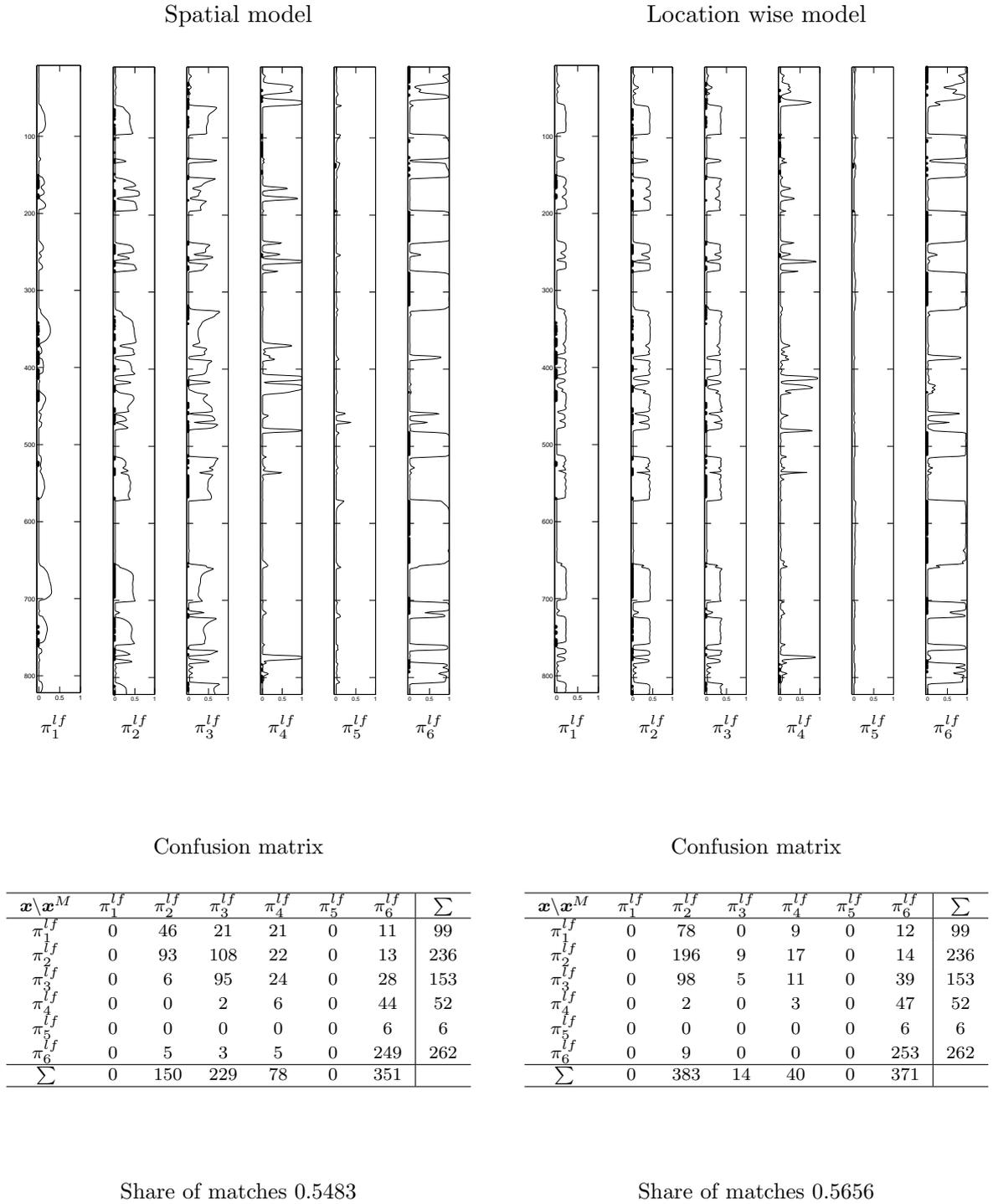


Figure 45: Results from the first parameter estimation (visual inspection). From the top the marginal posterior pdf for well B, $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$, is displayed for all t and all classes of lithofacies, π^{lf} , for the spatial model (left) and the location wise model (right). The presence of the classes of lithofacies at every location is marked on the axis of the respective plots. Further confusion matrices and share of matches for the two models.

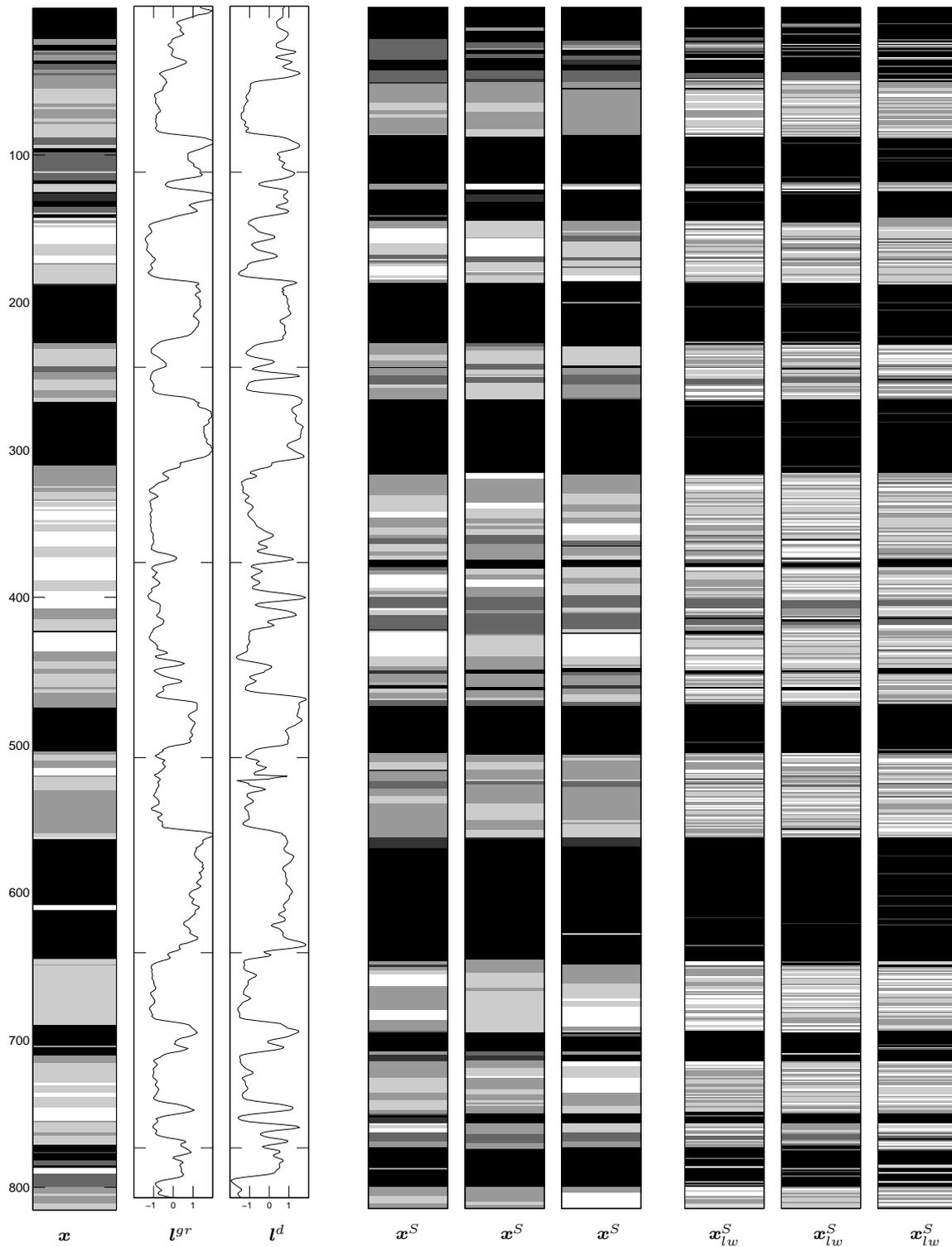


Figure 46: Results from the first parameter estimation (visual inspection). From left is the core classification of well B into lithofacies, π^{lf} , followed by the gamma-ray log, l^{gr} , and density log, l^d . Further, three simulations from the posterior pdf for the spatial model, \mathbf{x}^S , and location wise model \mathbf{x}_{lw}^S . The signal-to-noise ratio is 2.32 for the gamma-ray log and 1.61 for the density log. The gray scale levels of the respective classes are given in Figure 6, and we have that class π_1^{lf} is white and class π_6^{lf} is black. In the locations that are hatched the core classification is missing.

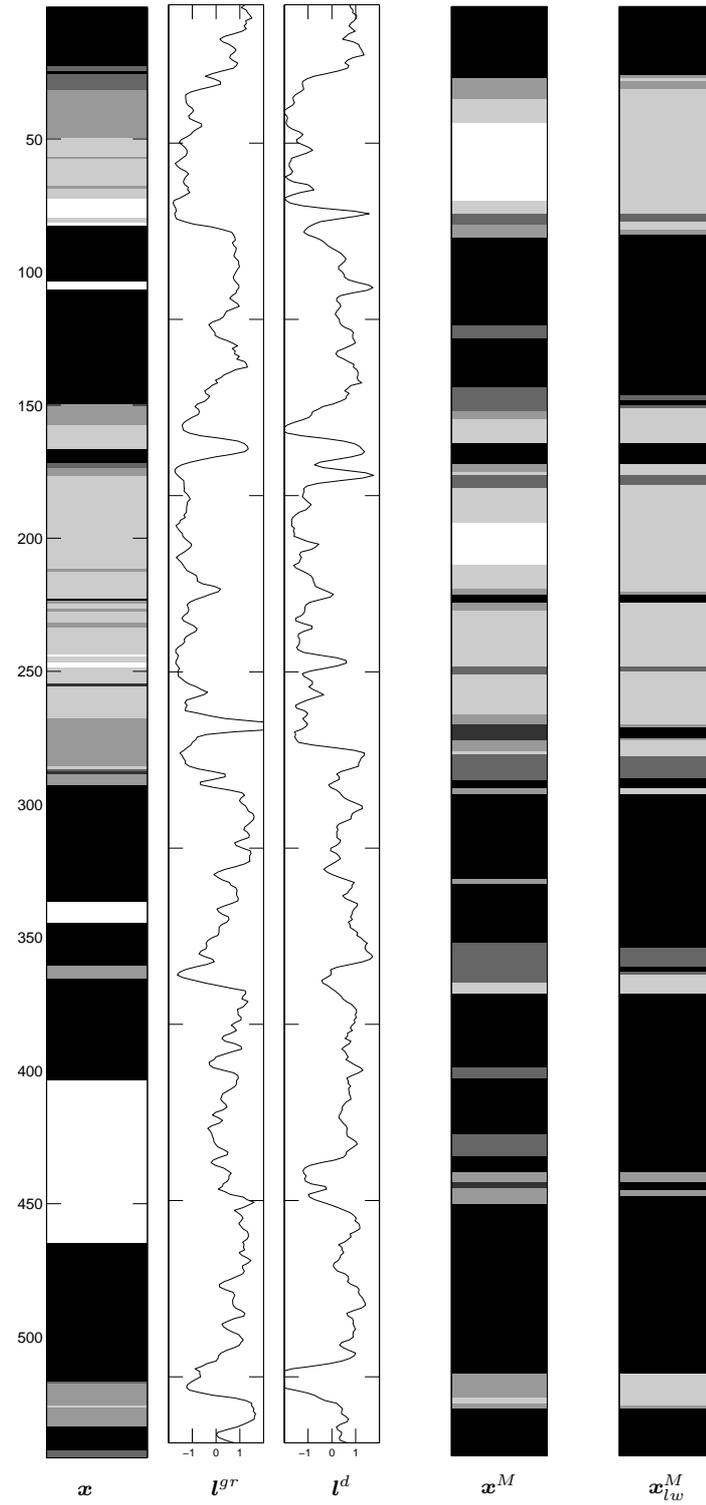


Figure 47: Results from the first parameter estimation (visual inspection). From left is the core classification of well C into lithofacies, π^{lf} followed by the gamma-ray log, l^{gr} , density log, l^d , and mmap solutions, x^M and x_{lw}^M , for the spatial and location wise model respectively. The signal-to-noise ratio is 1.53 for the gamma-ray log and 1.37 for the density log. The gray scale levels of the respective classes are given in Figure 6, and we have that class π_1^{lf} is white and class π_6^{lf} is black. In the locations that are hatched the core classification is missing.

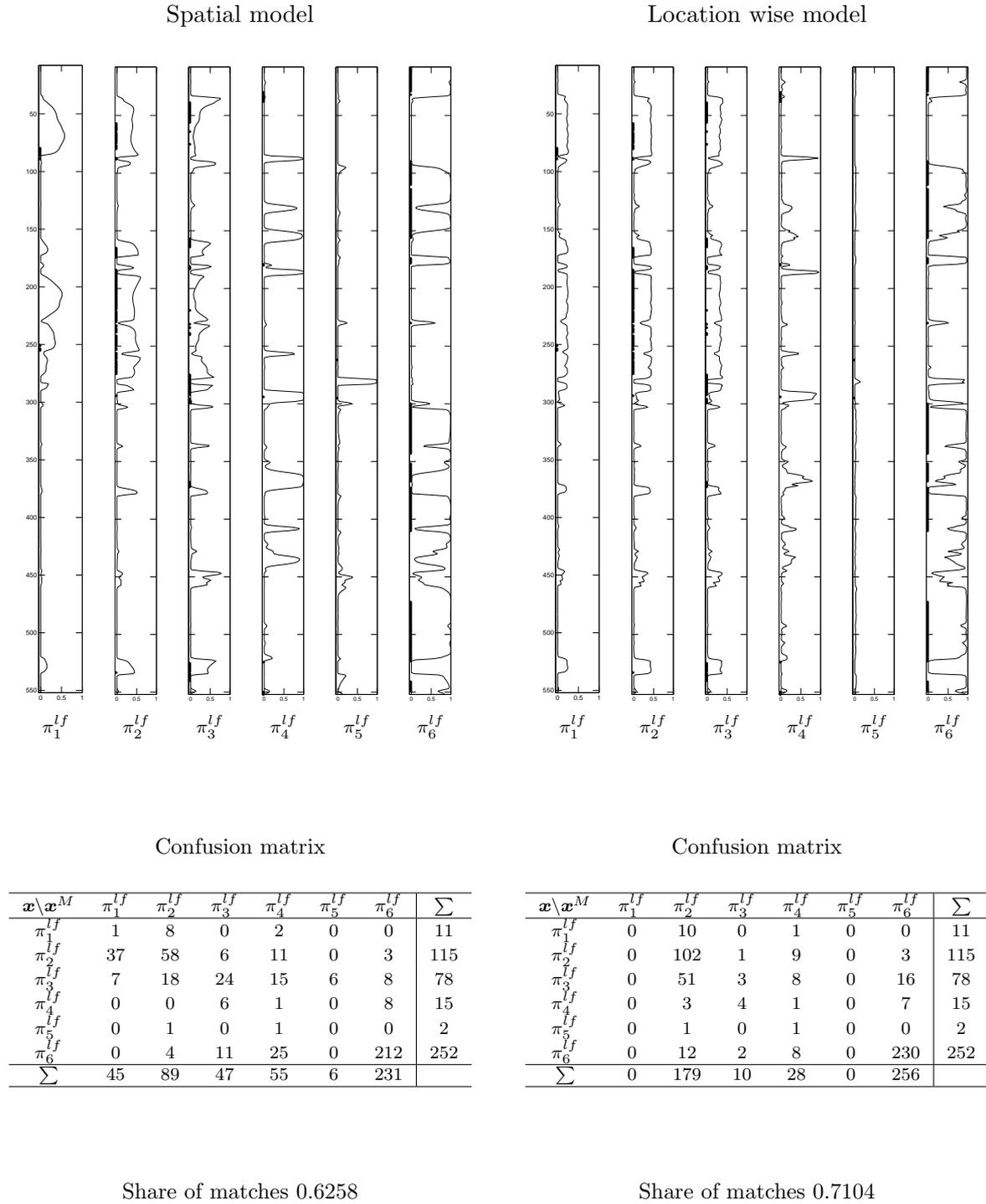


Figure 48: Results from the first parameter estimation (visual inspection). From the top the marginal posterior pdf for well C, $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$, is displayed for all t and all classes of lithofacies, π^{lf} , for the spatial model (left) and the location wise model (right). The presence of the classes of lithofacies at every location is marked on the axis of the respective plots. Further confusion matrices and share of matches for the two models.

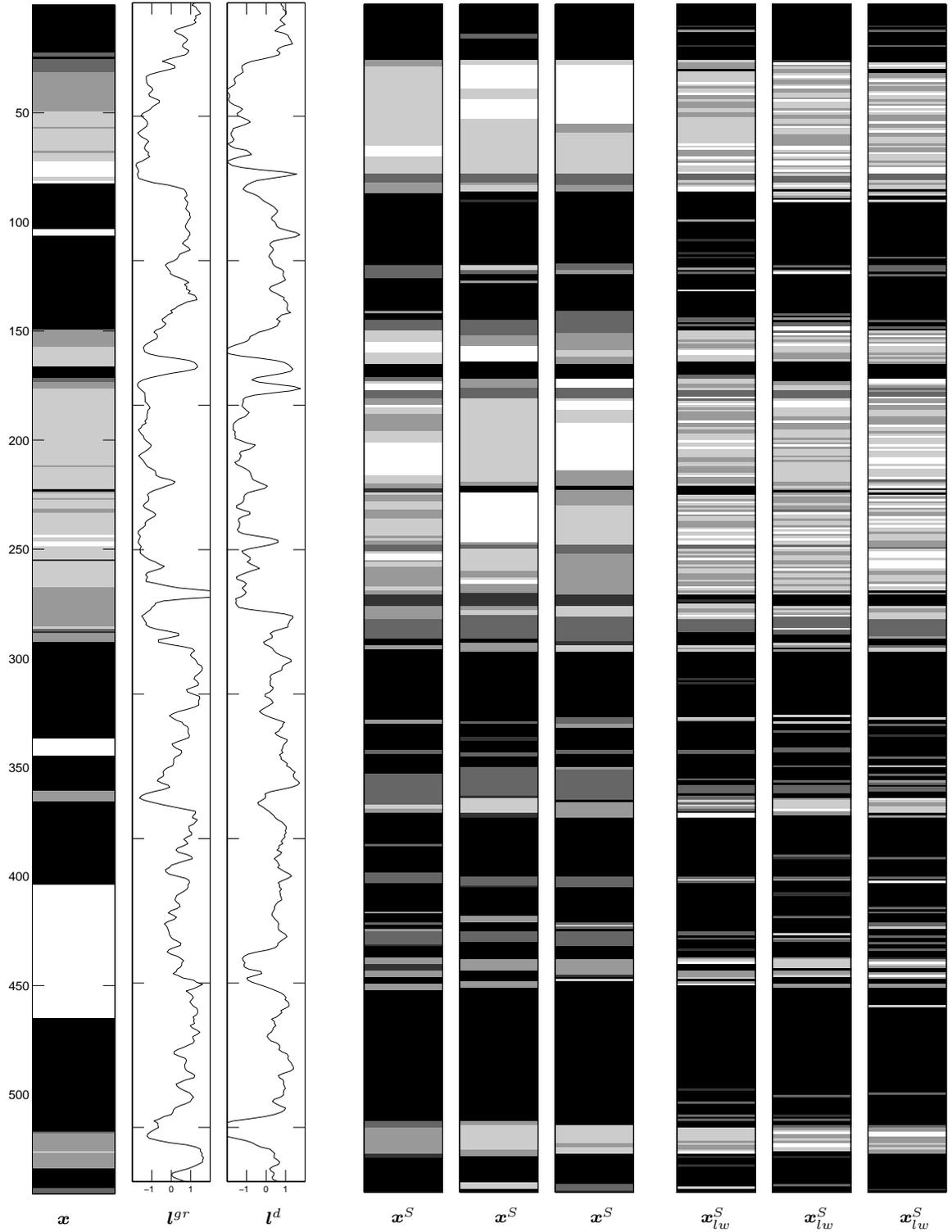


Figure 49: Results from the first parameter estimation (visual inspection). From left is the core classification of well C into lithofacies, π^{lf} , followed by the gamma-ray log, l^{gr} , and density log, l^d . Further, three simulations from the posterior pdf for the spatial model, \mathbf{x}^S , and location wise model \mathbf{x}_{lw}^S . The signal-to-noise ratio is 1.53 for the gamma-ray log and 1.37 for the density log. The gray scale levels of the respective classes are given in Figure 6, and we have that class π_1^{lf} is white and class π_6^{lf} is black. In the locations that are hatched the core classification is missing.

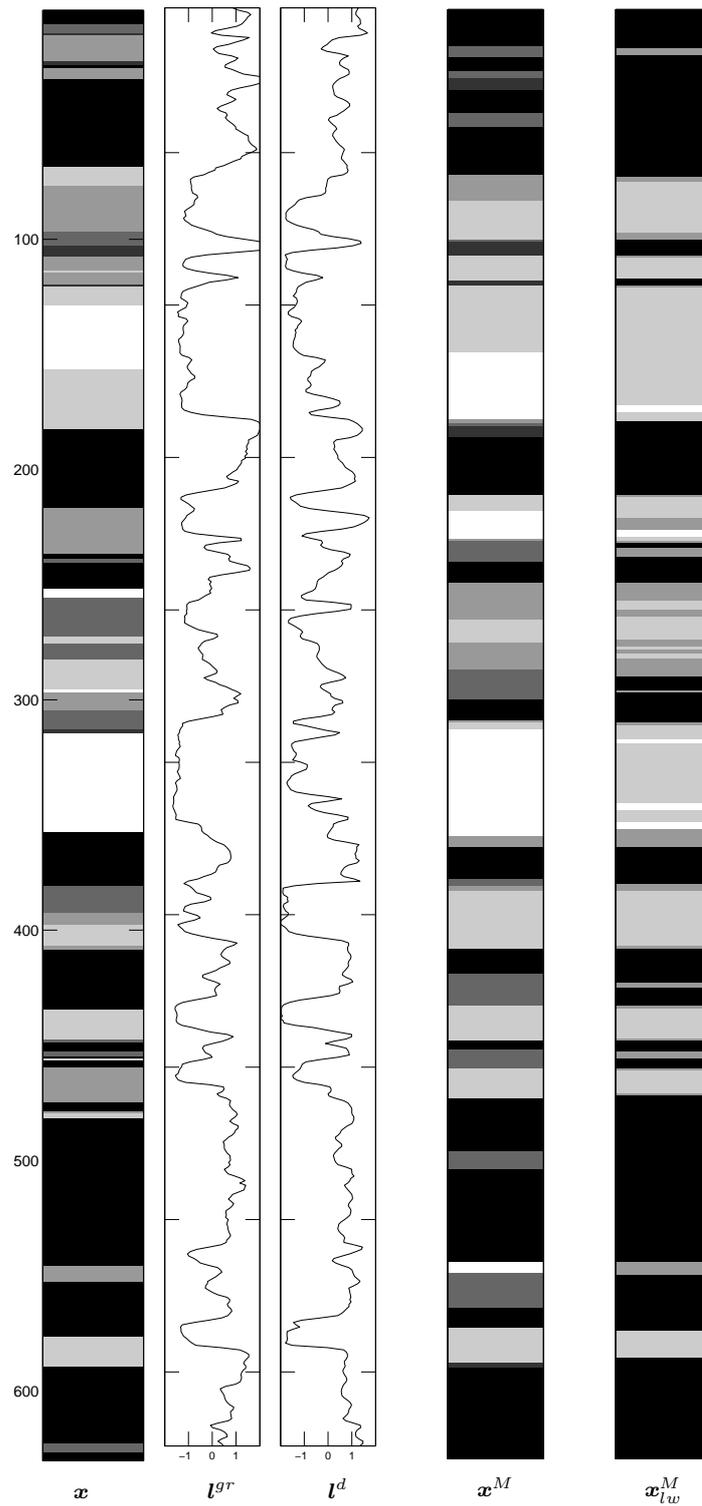


Figure 50: Results from the second parameter estimation (least squares). From left is the core classification of well A into lithofacies, π^{lf} followed by the gamma-ray log, l^{gr} , density log, l^d , and mmap solutions, x^M and x_{lw}^M , for the spatial and location wise model respectively. The signal-to-noise ratio is 1.57 for the gamma-ray log and 0.95 for the density log. The gray scale levels of the respective classes are given in Figure 6, and we have that class π_1^{lf} is white and class π_6^{lf} is black.

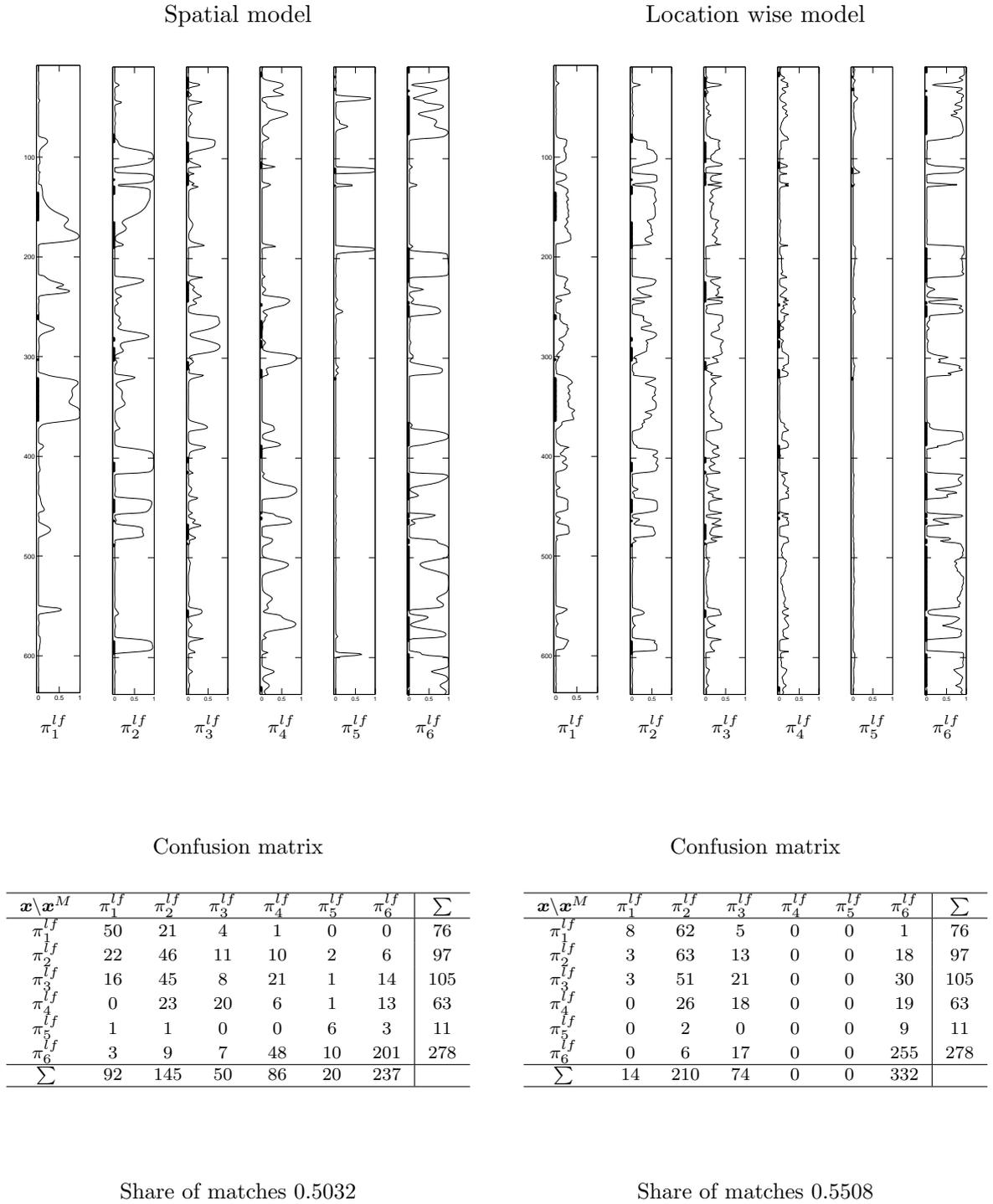


Figure 51: Results from the second parameter estimation (least squares). From the top the marginal posterior pdf for well A, $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$, is displayed for all t and all classes of lithofacies, π^{lf} , for the spatial model (left) and the location wise model (right). The presence of the classes of lithofacies at every location is marked on the axis of the respective plots. Further confusion matrices and share of matches for the two models.

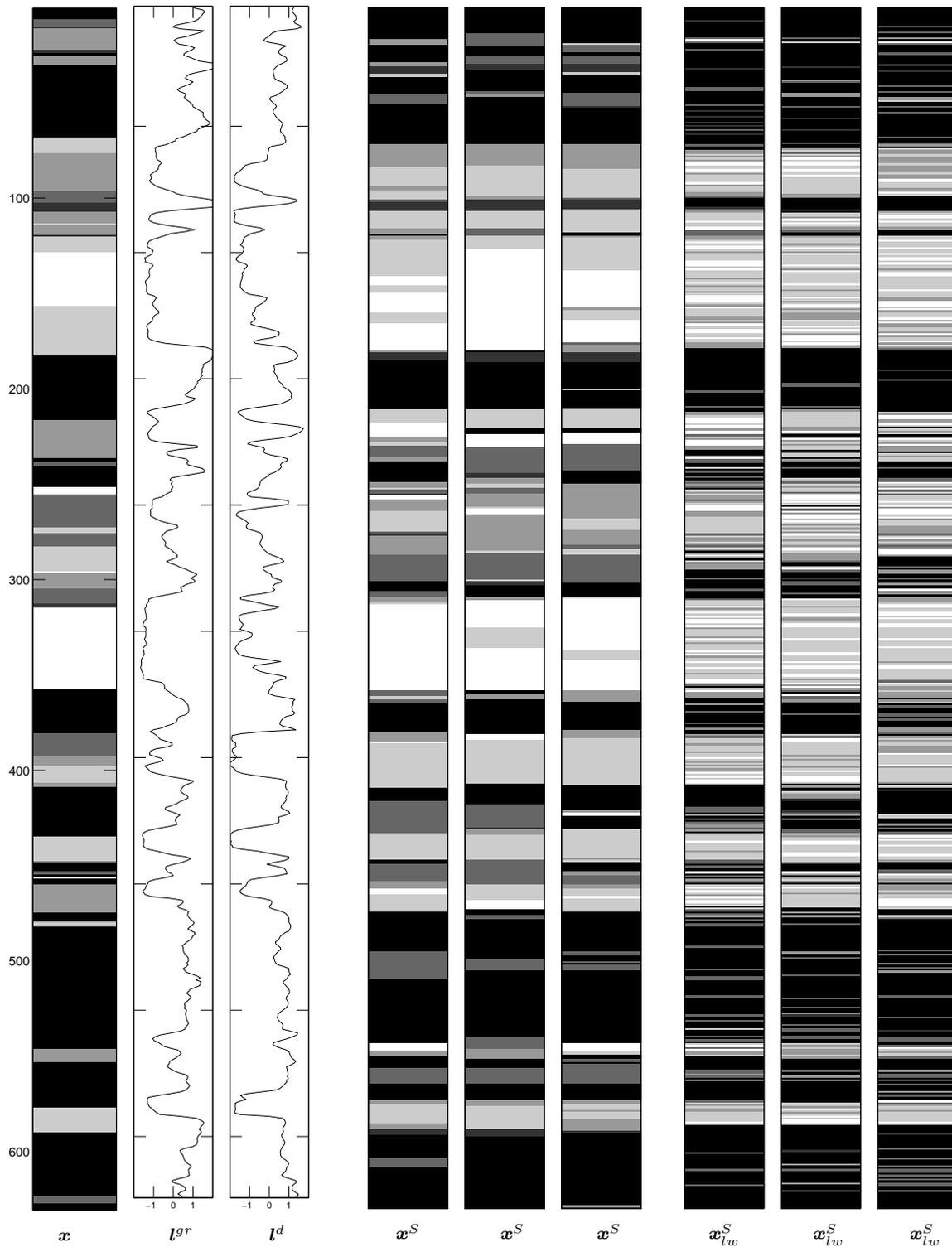


Figure 52: Results from the second parameter estimation (least squares). From left is the core classification of well A into lithofacies, π^{lf} , followed by the gamma-ray log, l^{gr} , and density log, l^d . Further, three simulations from the posterior pdf for the spatial model, \mathbf{x}^S , and location wise model \mathbf{x}_{lw}^S . The signal-to-noise ratio is 1.57 for the gamma-ray log and 0.95 for the density log. The gray scale levels of the respective classes are given in Figure 6, and we have that class π_1^{lf} is white and class π_6^{lf} is black.

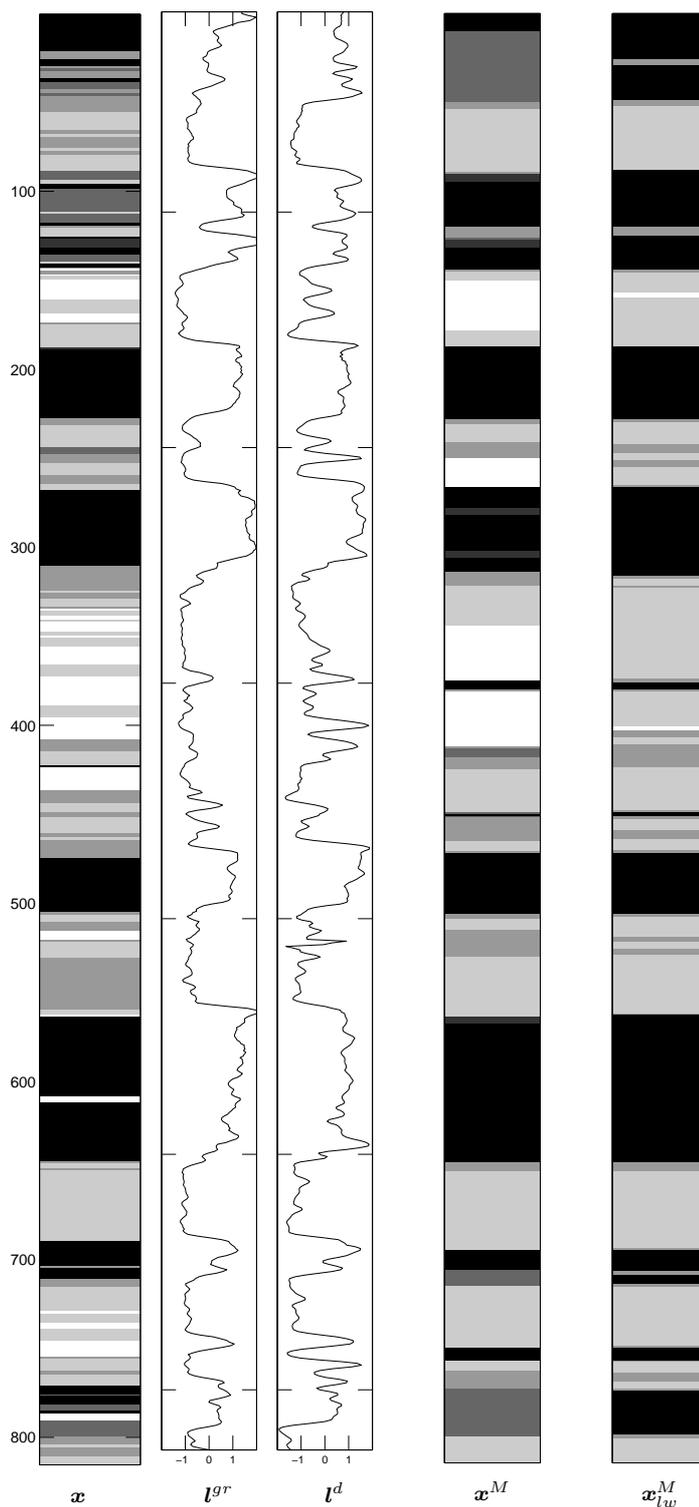


Figure 53: Results from the second parameter estimation (least squares). From left is the core classification of well B into lithofacies, π^{lf} followed by the gamma-ray log, l^{gr} , density log, l^d , and mmap solutions, x^M and x_{lw}^M , for the spatial and location wise model respectively. The signal-to-noise ratio is 2.57 for the gamma-ray log and 1.27 for the density log. The gray scale levels of the respective classes are given in Figure 6, and we have that class π_1^{lf} is white and class π_6^{lf} is black. In the locations that are hatched the core classification is missing.

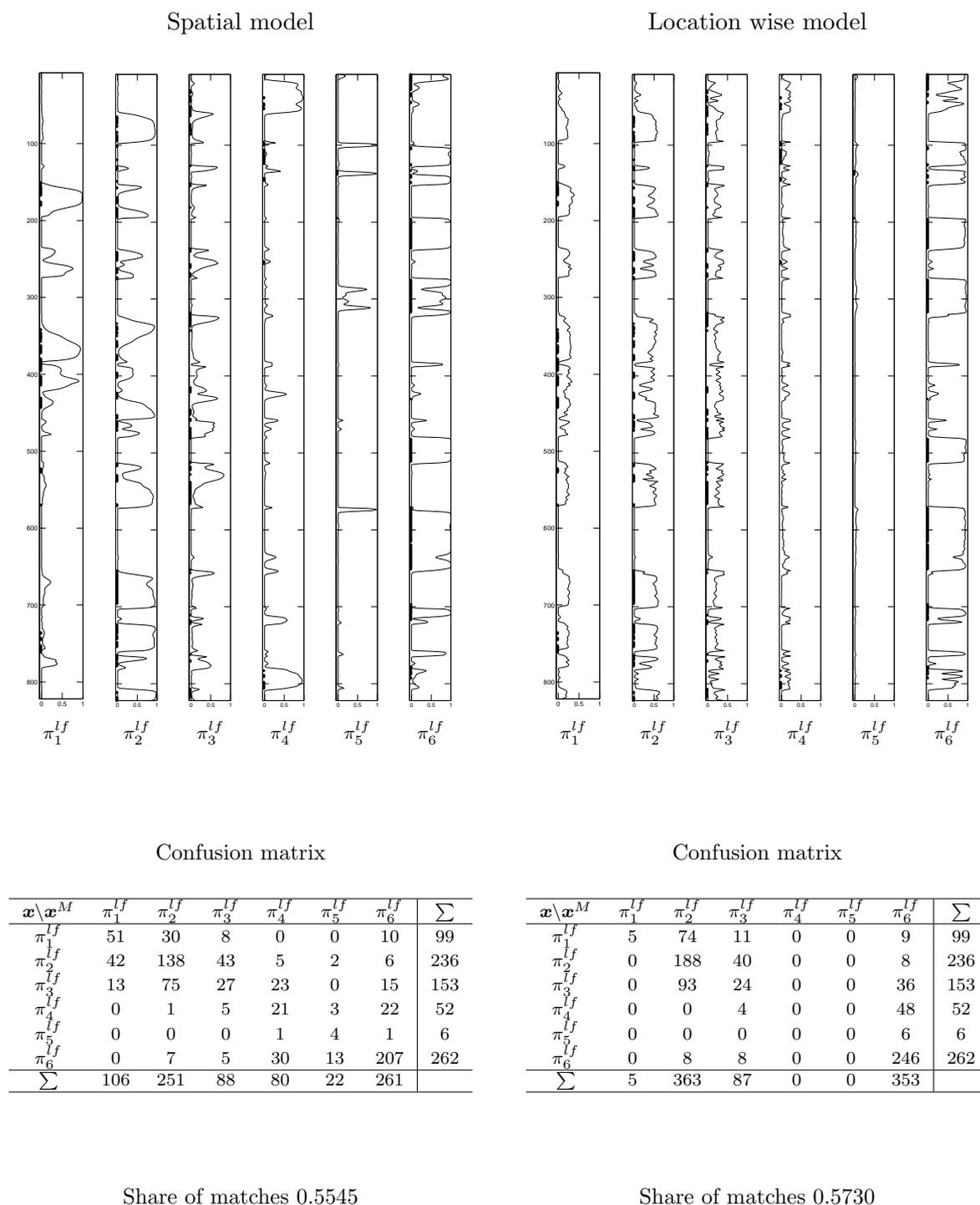


Figure 54: Results from the second parameter estimation (least squares). From the top the marginal posterior pdf for well B, $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$, is displayed for all t and all classes of lithofacies, π^{lf} , for the spatial model (left) and the location wise model (right). The presence of the classes of lithofacies at every location is marked on the axis of the respective plots. Further confusion matrices and share of matches for the two models.

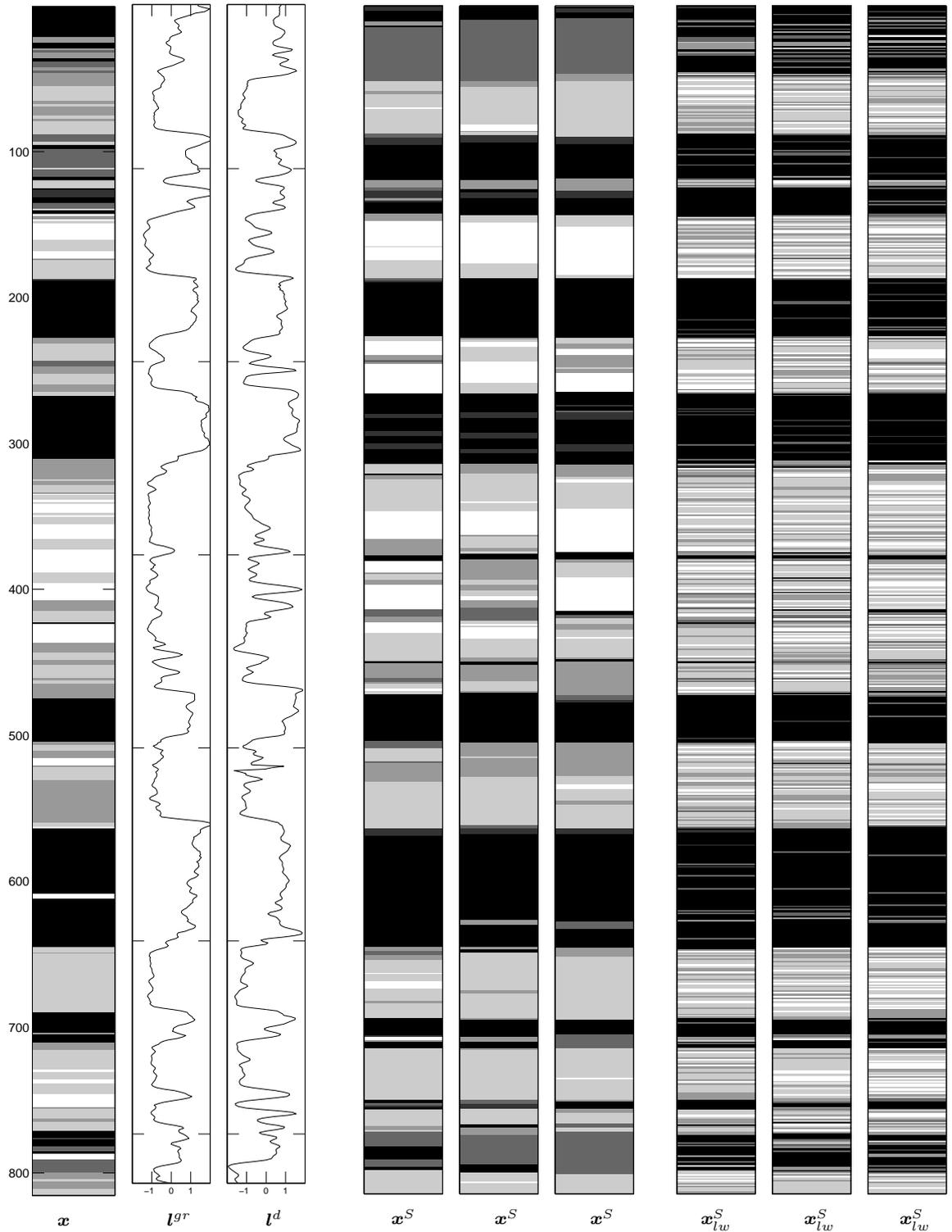


Figure 55: Results from the second parameter estimation (least squares). From left is the core classification of well B into lithofacies, π^{lf} , followed by the gamma-ray log, l^{gr} , and density log, l^d . Further, three simulations from the posterior pdf for the spatial model, \mathbf{x}^S , and location wise model \mathbf{x}_{lw}^S . The signal-to-noise ratio is 2.57 for the gamma-ray log and 1.27 for the density log. The gray scale levels of the respective classes are given in Figure 6, and we have that class π_1^{lf} is white and class π_6^{lf} is black. In the locations that are hatched the core classification is missing.

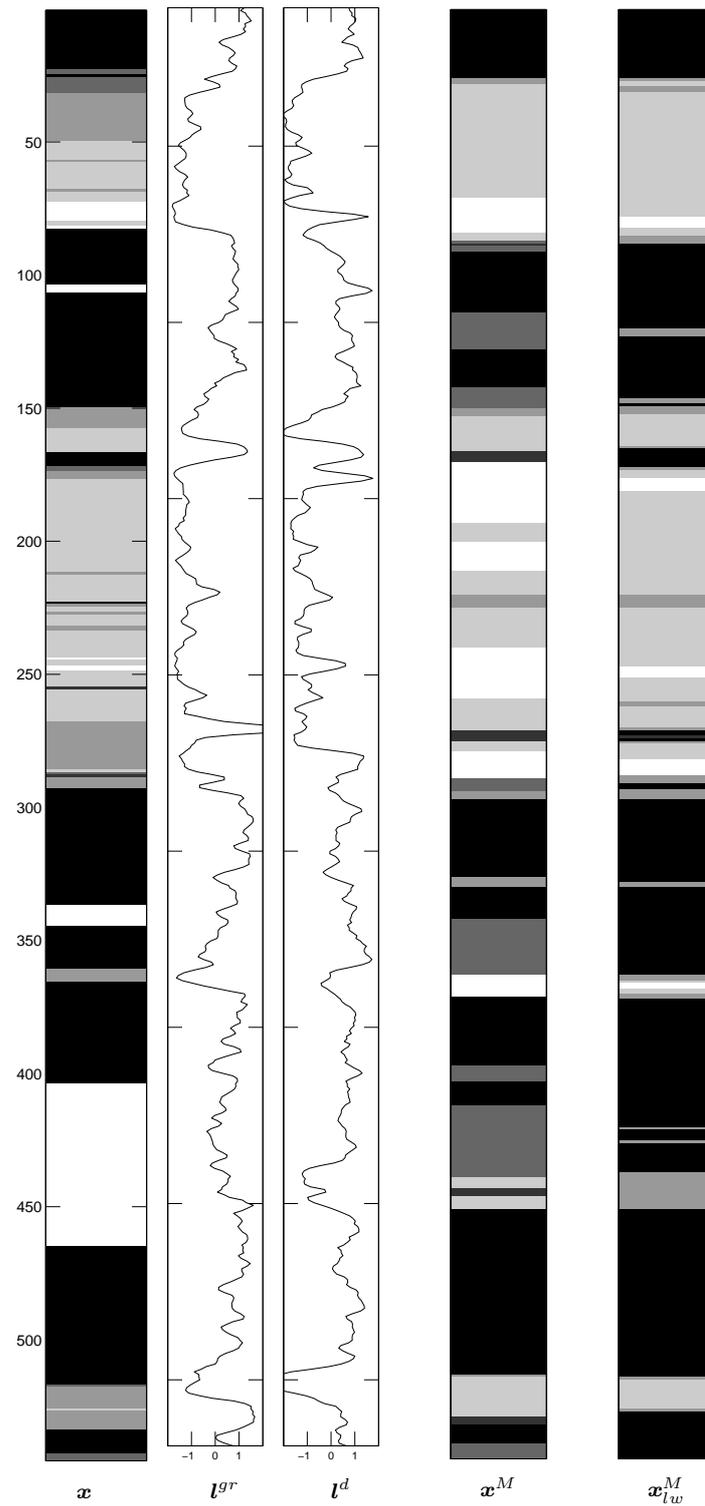


Figure 56: Results from the second parameter estimation (least squares). From left is the core classification of well C into lithofacies, π^{lf} followed by the gamma-ray log, l^{gr} , density log, l^d , and mmap solutions, x^M and x_{lw}^M , for the spatial and location wise model respectively. The signal-to-noise ratio is 1.21 for the gamma-ray log and 1.07 for the density log. The gray scale levels of the respective classes are given in Figure 6, and we have that class π_1^{lf} is white and class π_6^{lf} is black. In the locations that are hatched the core classification is missing.

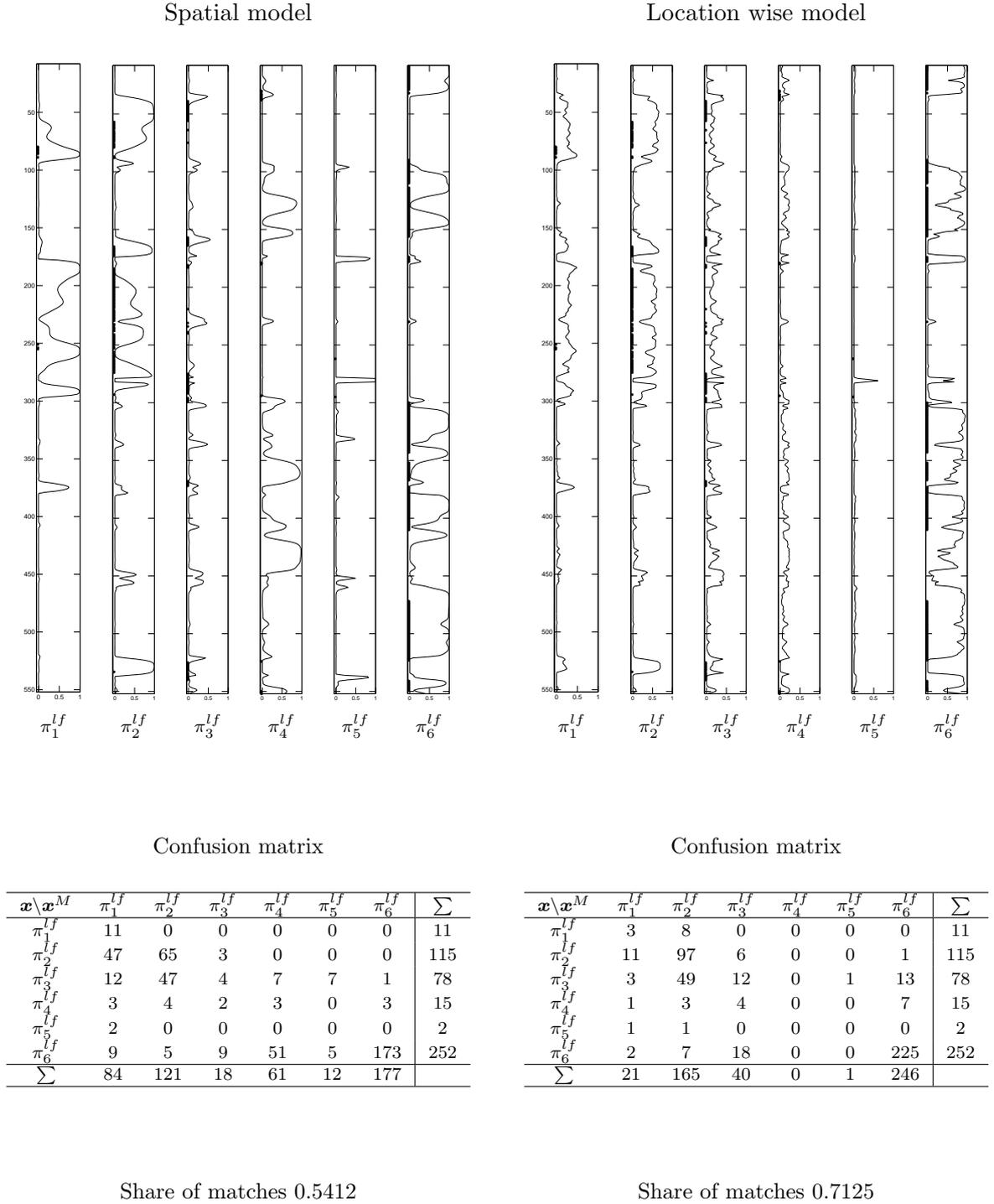


Figure 57: Results from the second parameter estimation (least squares). From the top the marginal posterior pdf for well C, $p(x_t | \mathbf{l}^{gr}, \mathbf{l}^d)$, is displayed for all t and all classes of lithofacies, π^{lf} , for the spatial model (left) and the location wise model (right). The presence of the classes of lithofacies at every location is marked on the axis of the respective plots. Further confusion matrices and share of matches for the two models.

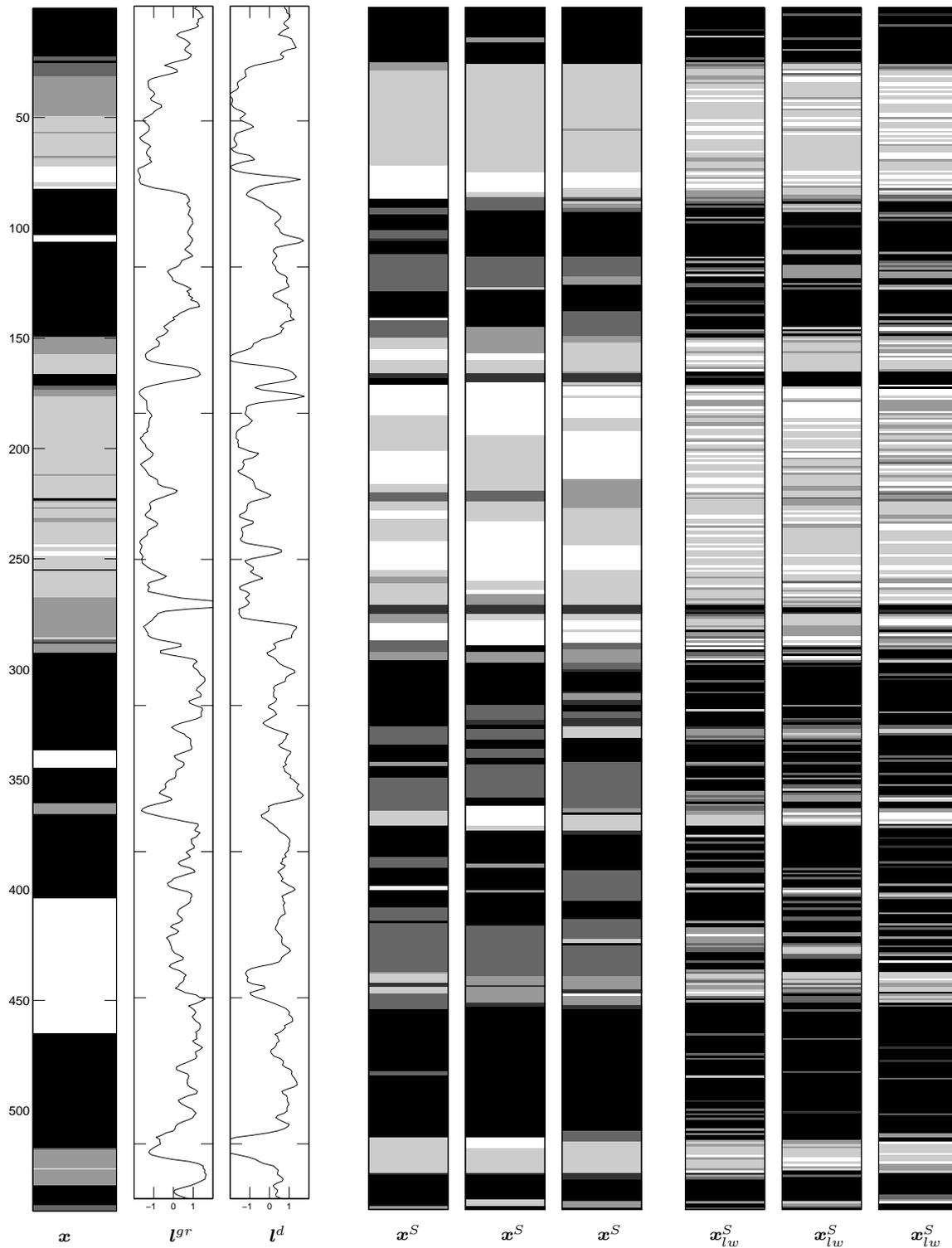


Figure 58: Results from the second parameter estimation (least squares). From left is the core classification of well C into lithofacies, π^{lf} , followed by the gamma-ray log, l^{gr} , and density log, l^d . Further, three simulations from the posterior pdf for the spatial model, x^S , and location wise model x_{lw}^S . The signal-to-noise ratio is 1.21 for the gamma-ray log and 1.07 for the density log. The gray scale levels of the respective classes are given in Figure 6, and we have that class π_1^{lf} is white and class π_6^{lf} is black. In the locations that are hatched the core classification is missing.

7 Closing Remarks

In this study we have shown that a spatial classification model for wells from the Statfjord Formation at the Tempen Area can be derived and evaluated. Through a study of the well logs and the formation of the wells, we have adapted a Bayesian model to the well log data considered.

From the study of the well logs, we have seen that the levels of response from the formation varies between the wells. The need of a standardization method for the log data is therefore present. In this study we considered a simple standardization of the log data, but a class-dependent standardization may be considered in future research.

We found a considerable spatial structure in the classes of facies associations and lithofacies, which is reflected in the prior model. The proportions of the classes in the wells are different, which makes it hard to derive a good prior model for classification of blind wells. In further studies a prior model may be estimated from several representative training wells. A semi-Markov or a second order Markov prior model can also be considered, though the problem will be extremely expanded. The well logs also contains spatial structure. The log-response is considerably affected by neighbour observations in the formation, and we have tried to recreate the underlying structure in the likelihood model. The parameter estimates in the likelihood model were relatively stable. In further work a class-dependent error may be considered.

The classification model in this study is complex, and there are many model parameters to be estimated. If the model assumptions and the underlying model does not coincides completely, a complex model may give poorer results than a simpler model. The results from the classifications of the blind well were not as good as we had hoped, which may be due to incorrect model assumptions, but also to the differences between the wells considered. Still, our classification model managed some of the classical challenges in facies classification. Further, the spatial model managed to classify the small classes of the wells relatively good, compared to the location wise model.

There is obviously a spatial structure in the well log data and the formation. The expenses of the spatial classification model are the large state spaces, and therefore the need of large computing resources. The extra effort with the spatial model compared to a location wise model, is yet not proven to give significantly improved results. If a spatial model should be considered in further studies, more representative training data must be used.

References

- Ali, A. I. and Lall, U. “A Continuous Parameter Homogeneous Semi-Markov Model for Stratigraphic Analyses from Borehole Data.” *The log analyst: a journal of formation evaluation and reservoir*, 29(2):26–37 (1998).
- Avseth, P. and Mukerji, T. “Seismic Lithofacies Classification from Well Logs using Statistical Rock Physics.” *Petrophysics: the SPWLA journal of formation evaluation and rese*, 43(21):70–81 (2002).
- Chib, S. “Calculating posterior distributions and modal estimates in Markov mixture models.” *Journal of Econometrics*, 75:79–97 (1996).
- Efron, B. and Tibshirani, R. *An Introduction to the Bootstrap*. Chapman & Hall (2003).
- Ross, S. *Introduction to Probability Models*. Academic Press, eight, international edition (2003).
- Scott, S. “Bayesian Methods for Hidden Markov Models: Recursive Computing in the 21st Century.” *Journal of the American Statistical Association*, 97(457):337–351 (2002).