Paper IV

Petro-Elastic Log-Facies Classification Using the Expectation-Maximization Method and Hidden Markov Models

David V. Lindberg & Dario Grana

Paper submitted

Petro-Elastic Log-Facies Classification Using the Expectation-Maximization Method and Hidden Markov Models

Abstract

Log-facies classification methods aim to estimate a profile of facies at the well location based on the values of rock properties measured or computed in well log analysis. Statistical methods generally provide the most likely classification of lithological facies along the borehole by maximizing a function that describes the likelihood of a set of rock samples belonging to a certain facies. However, most of the available methods classify each sample in the well log independently and do not account for the spatial distribution of the facies profile. In this work, a classification method based on hidden Markov models is proposed, a stochastic method that accounts for the probability of transitions from one facies to another one. Differently from other available methods where the model parameters are assessed using nearby fields or analogues, the unknown parameters are estimated using a statistical algorithm called the Expectation-Maximization algorithm. The method is applied to two different datasets: A clastic reservoir in the North Sea where four litho-fluid facies are identified and an unconventional reservoir where four lithological facies are defined. The application also includes a sensitivity analysis and a comparison to other statistical methods.

1 Introduction

Facies classification is one of the key modeling components in reservoir characterization. The distribution of rock properties in the static reservoir model depends on the facies classification, which is generally achieved from geophysical measurements. Reservoir facies are defined at the well location first. Traditional log-facies classification is based on depositional and sedimentological models. A detailed geological model of the facies strictly depends on the availability of core samples in the well. However, these models are difficult to extend to the entire reservoir model for the lack of accurate measurements far away from the well. In order to extend the facies classification to the entire reservoir model, log facies should be linked to well logs data. Generally, in conventional reservoirs, log-facies classification relies on petrophysical curves performed in formation evaluation analysis (such as, porosity and mineralogical volumes). However, facies classification, especially in unconventional reservoirs, should also include elastic and geomechanical properties (such as, P- and S- wave velocities, Poisson ratio, Young's modulus).

In many practical applications, facies are first defined based on core sample analysis combined with regional geological models. Then log-facies are re-classified at the well location using well logs. The accuracy of this reclassification is lower than the classification obtained from core samples and geological models, and some of the geological facies can be grouped into a broader lithological class. For example, in a conventional clastic reservoir, geological facies such as marine shale and flood plain are generally not distinguishable from petrophysical curves, because they both show low effective porosity and high clay content and they can be reclassified simply as shale. Similarly, distributary channels and crevasse plays show high quartz content and high effective porosity and could be reclassified as sand.

Different methods can be used to classify well logs data in terms of log-facies. Simple deterministic methods, such as cut-off methods, could be used to discriminate facies based on a limited number of well logs. For example one could use the gamma ray log and a cut-off value of 100 API, to discriminate sand and shale in a clastic reservoir, or use the effective porosity curve and a cut-off value of 0.2 to differentiate clean sand from shaley sand. However this method is not suitable in complex geological environments where multiple logs should be used to classify log-facies. If all the main acquired well logs should be used simultaneously (for instance neutron porosity, density, resistivity and gamma ray) and/or all the petrophysical curves related to the volumetric fractions, deterministic methods are not suitable. The first challenge is

the number of well logs, that does not allow to use cut-off values, since a certain facies could satisfy the cut-off value for one variable but not for another property. The second issue is related to the accuracy of the data: as a matter of fact well log measurements and the curves computed from them can contain errors which could affect the deterministic classification. In order to face these two issues, statistical methods should be introduced. Several statistical methods have been presented in literature. Most of these methods belong to the family of clustering algorithms. Examples of these algorithms are: linear and non-linear discriminant analysis, k-means clustering, support vector machine etc (Hastie et al. (2009)). Some of these algorithms require a statistically representative training dataset (supervised learning methods), others only require initial guess values (unsupervised learning methods). Examples of unsupervised learning algorithms are neural networks and self-organizing maps. Bayesian classification can also be used as long as a reliable likelihood function can be estimated from a facies dataset (for example from lab measurements).

All these methods estimate the most likely facies classification at each location in the well log, but fail to account for the vertical continuity in the facies profile. Each sample in the well log is classified independently from the adjacent samples. Therefore unrealistic facies sequences could be created in the classified profile, for example very thin layers if the well logs are very noisy. Furthermore some transitions between facies could be more likely to happen than others, and other transitions could be unfeasible. The simplest example can be described in a scenario with one facies, sand, and three fluids within it: gas, oil and water. Suppose focus is on classifying the three litho-fluid facies: gas sand, oil sand and water sand. If vertical continuity is not accounted for in the classification, one could obtain water sand on top of gas sand, which is not physically possible due to the gravity effect.

In this paper, a new approach to log facies classification is proposed where the vertical correlation of facies along the well profile is introduced through a spatial stochastic model, to ensure realistic vertical sequences along the well profile. The log facies classification we propose is based on three steps: 1) identification of lithological facies at the well log scale and geological interpretation; 2) selection of well logs data to perform the classification, including petrophysical, elastic and if available geomechanical properties; and 3) statistical methodology to classify log-facies based on Expectation-Maximization (EM) and hidden Markov models (HMM).

The combined use of petro-elastic properties was previously proposed in Grana et al. (2012) in a Monte Carlo classification workflow, and even though

the approach allows representing the posterior uncertainty in the classification, it does not account for vertical correlation in the profile. Therefore to improve the available classification methods, Markov models are here introduced. The use of Markov chains to model geological layering was first proposed by Krumbein and Dacey (1969), see also Elfeki and Dekking (2001) for an overview. A Markov model is a stochastic process where the conditional probability distribution of future states of the process depends only upon the present state, not on the sequence of events that preceded it. In this work, the facies at a given location is considered as a state of the process. Then, the use of a Markov model implies that the probability of having a given facies at a given location, depends only on the facies at the location above. This conditional probability guarantees vertical correlation in the facies profile, which is an advantage since it avoids artifacts in the classification, such as one-sample layers or unrealistic facies sequences. Non-physical transitions can also be avoided in Markov models by setting this transition probability equal to 0 in the prior model.

When the states following the Markov model, that is the facies profile, are unobserved and only the related outputs, that is the well logs, are observed, the process is called a hidden Markov model (HMM). Several applications of HMMs exist in computer science, speech recognition, signal theory and biology, see for example Dymarski (2011) for an overview. Eidsvik et al. (2004) proposed the use of HMMs for well log inversion into geological attributes in the same context as considered in this work, and estimated the parameters in a fully Bayesian framework requiring Markov chain Monte Carlo (McMC) calculations. We choose to estimate the parameters of the HMM using the Expectation-Maximization (EM) algorithm (Dempster and Rubin 1977). Compared to McMC, inference on the parameters using the EM algorithm is much less computer demanding. The particular application of the EM algorithm to HMMs is also known as the Baum-Welch algorithm (Baum et al. 1970) and has not been applied to well log inversion before. In the application it is assumed that the distribution of the log properties used in the classification is Gaussian within each facies, which results in a Gaussian mixture model where each component of the mixture corresponds to a lithofacies. Gaussian mixture have been previously used in a facies classification context in Grana and Della Rossa (2010) and Xu and Torres-Verdin (2014). Differently from these applications, the current approach aims to estimate the unknown parameters of the mixture, as well as the unknown parameters of the vertical sequence transitions through the EM method.

The mathematical methodology is first introduced, where the main features

of HMMs are summarized and the application of the EM method to estimate the HMM parameters is presented. Next, two datasets chosen for the application of the presented method are introduced. The first dataset includes well logs and computed curves from three different domains: petrophysics, geomechanics and elastic domain, acquired in an unconventional reservoir (Marcellus shale) in North America. The goal of this application is to classify the main lithological facies at the well location. A comparison of statistical classification methods to the application of the presented method to a representative sub-set of well logs is presented. The second dataset consists of a standard set of well logs data, sonic logs, petrophysical logs and curves performed in formation evaluation analysis, acquired in a clastic reservoir in the North Sea characterized by a complex sequence of thin sand and shale layers. The main goal of this application is to classify four litho-fluid facies: gas sand, oil sand, water sand and shale.

2 Methodology

In this section, the statistical methodology used for log facies classification is presented. First, Markov models are introduced, then HMMs, and finally the EM method for the estimation of the HMM parameters.

2.1 Introduction to Markov Models and Hidden Markov Models

The unknown facies profile can be represented as a sequence of unknown states of a process. A stochastic process has the (first-order) Markov property if the conditional probability distribution of future states of the process depends only upon the present state, not on the sequence of events that preceded it. A discrete-time stochastic process satisfying the Markov property is known as a Markov chain. A Markov chain is a sequence of random variables $\{X_t\}_{t=1,\dots,T}$ with the Markov property, that is for any $1 < t \leq T$

$$p(X_{t+1} = x | X_1 = x_1, \dots, X_t = x_t) = p(X_{t+1} = x_{t+1} | X_t = x_t)\,. \qquad (1)$$

In this work, categorical Markov processes with a discrete-valued state space only are considered, hence each state X_t belongs to one out of N classes, for example for N possible facies classes. The probability on the right hand side of Eq.1 is called the transition probability, which is assumed to be stationary throughout the sequence, that is independent of t. These transition probabilities can be collected in a $(N \times N)$ matrix P, called the transition matrix.

Define the probability of a transition from state i to state j as P_{ij}, then the transition matrix can be written as

$$P = \begin{pmatrix} P_{11} & \cdots & P_{1N} \\ \vdots & P_{ij} & \vdots \\ P_{N1} & \cdots & P_{NN} \end{pmatrix}. \tag{2}$$

Since there are only N possible states, the sum of transition probabilities from state i to the other states $j = 1,\ldots,N$ must be 1, that is each row of the matrix sums to 1, $\sum_{j=1}^{N} P_{ij} = 1$. Using the transition matrix it is possible to calculate, for example, the global proportions of the states, by taking the limit, $\lim_{k\to\infty} P^k$.

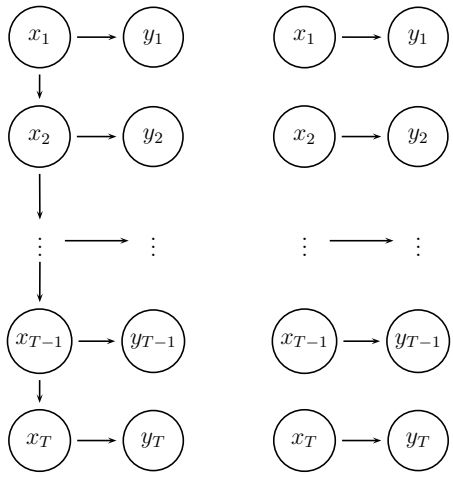
Consider for example a set of three facies: sand, silt and shale. An example of transition probabilities can be

$$P = \begin{array}{c} \begin{matrix} sand & silt & shale \end{matrix} \\ \begin{pmatrix} 0.95 & 0.025 & 0.025 \\ 0.04 & 0.95 & 0.01 \\ 0 & 0.05 & 0.95 \end{pmatrix} \end{array} \begin{matrix} sand \\ silt \\ shale \end{matrix} \tag{3}$$

Rows correspond to sand, silt, and shale at generic depth z, and columns refer to sand, silt, and shale at depth $z+1$ (downward transition). The matrix is read (by row) as follows: if one knows that at depth z the actual facies is sand, then at depth $z+1$ the probability of finding sand is 0.95, the probability of finding silt is 0.025, and the probability of finding shale is 0.025. Similarly for the other rows. In the example the transition from shale to sand is impossible (probability equal 0). Impossible transitions can be often found in transition probabilities of litho-fluid classes, to honor the gravity effects. The terms on the diagonal of the transition matrix are related to the thickness of the layers: the higher the numbers on the diagonal, the higher the probability to observe no transition (that is, high probability that a facies has a transition to itself).

A HMM is a model in which the process being modeled is assumed to follow a Markov process with unobserved (hidden) states. Generally the Markov process itself cannot be observed (in other words the states cannot be measured) but indirect observations, related to the states, are available. For example,

facies cannot be measured in the subsurface, but rock properties that depend on the facies type can be measured, such as porosity for example. Let the random variable Y_t be the observation at time t, which relation to the corresponding hidden state at time t, X_t, is given by the probability density function (pdf) of Y_t conditioned on X_t, $p(Y_t|X_t)$, denoted the output probabilities. The relation is displayed in Fig.1a, notice the Markov property in the hidden level and the elementwise relations between the observations and the hidden states represented by arrows. In this work it is assumed that the observations are continuous, typically from a Gaussian distribution. The parameters of a HMM are then: the transition matrix defining the transition probabilities $p(X_t|X_{t-1})$ and the parameters of the conditional pdf defining the output probabilities $p(Y_t|X_t)$.



(a) HMM **(b)** Model without vertical spatial dependency

Figure 1: Directed acyclic graph of (a) a standard HMM and (b) a model without spatial dependency, for example for LDA, QDA and NBM. Here, $\mathbf{y} = (y_1, \ldots, y_T)$ are observed data (well logs) and $\mathbf{x} = (x_1, \ldots, x_T)$ unobserved variables (facies). The directed arrows represent dependencies between the variables.

If the observed variable is a M-dimensional vector distributed according to a multivariate Gaussian distribution for each state, then for each class there are M parameters for the vector of means of the observed variables and $M(M+1)/2$ parameters for the covariance matrix, which results in $N\frac{M(M+3)}{2}$ parameters in total. In the application, if there are $N = 3$ facies and $M = 4$

rock properties (for example, well logs of porosity, clay volume, P-wave velocity, and S-wave velocity), then the hidden Markov model requires the estimation of 42 parameters. The assumption that the observed variable is a M-dimensional vector distributed according to a multivariate Gaussian distribution for each state means that the conditional distribution of the observed variables given the state is a Gaussian Mixture model (GMM).

The complete set of HMM parameters for a given model by is described by $\lambda = \{A, B, \pi\}$ where A represents the transition matrix, B represents the parameters of the output probabilities, and π is the vector of global proportions of the states. These parameters can be easily estimated once the facies classification is known (supervised learning), for example for the transition probabilities by counting the number of transitions throughout the facies classification and normalizing. However, one can still assume a Markov model even if the facies classification is not known (unsupervised learning). To obtain a facies classification at the well location, the problem constitutes finding the most probable sequence of states (facies) and the corresponding parameters that maximize their probability. In order to accomplish this task, the EM method is applied, which for the particular case of HMMs is also called the Baum-Welch algorithm (Baum et al. 1970).

2.2 The Expectation-Maximization Method for Hidden Markov Models

As shown in the previous section, HMMs can require the estimation of a large number of parameters. Several methods have been proposed. In the log-facies classification methodology presented, the EM method (Dempster and Rubin 1977) is adopted. The EM method finds the maximum likelihood estimates of parameters in probabilistic models in the presence of missing data, that is in this context when the facies classification is unknown. This method is not completely new in geophysics. Grana and Della Rossa (2010) introduced this method to estimate the parameters of a Gaussian mixture model but ignored the spatial correlation of the Markov model in the estimation step. The application of the EM-method to hidden Markov models for log-facies classification is new.

Denote the complete data by $Z = (X, Y)$ for which X are the unobserved (hidden) variables (the facies classification) and $Y = y$ the observations (the well logs), and denote the parameters to estimate by λ. The EM method is an iterative procedure that consists in two main steps. The expectation (E) step computes the expectation of the complete data log-likelihood function

with respect to the current estimate of the parameters λ^*

$$\begin{aligned} Q(\lambda,\lambda^*) &= E_{X|Y=y,\lambda^*}\left[\log\{p(X,Y=y|\lambda)\right] \\ &= \sum_X \log\{p(X,Y=y|\lambda)\} \times p(X|Y=y,\lambda^*)\,. \end{aligned} \tag{4}$$

The maximization (M) step maximizes the expected likelihood computed in the expectation part

$$\lambda = \arg\max_\lambda \left\{Q(\lambda,\lambda^*)\right\}\,. \tag{5}$$

When certain criteria are fulfilled (Wu 1983), the algorithm converges to the maximum likelihood solution

$$\widehat{\lambda} = \arg\max_\lambda \{p(y|\lambda)\}\,, \tag{6}$$

that is the set of parameters that maximize the probability of the observation under the current model.

If the rock properties are assumed to take a Gaussian mixture distribution, and the components of the mixture with the facies are identified, then the EM algorithm can be used as a clustering method to classify the facies. Indeed, the EM algorithm can be applied to the set of rock properties to estimate the parameters of the mixture and for each sample one can assign the classified facies by taking the argument of the maximum of the likelihood of the components. However, this application does not account for the spatial correlation and the transition probabilities. In this work, the EM algorithm is therefor combined with HMMs, with the Gaussian mixture assumption. The parameters to estimate are: 1) the transition matrix $P = \{P_{ij}\}_{i,j=1,\ldots,N}$, where N is the number of possible states of the hidden random variable X, and 2) the means μ_i and the covariance matrices Σ_i of the Gaussian probability distributions $b_i(y)$ of the observation variables Y, for each state $i = 1,\ldots,N$

$$b_i(y) := p(y|X=i) = N(y;\mu_y^{(i)},\Sigma_y^{(i)}) \tag{7}$$

or, in terms of Gaussian mixture models

$$b(y) = \sum_{i=1}^{N} \pi_i N(y;\mu_y^{(i)},\Sigma_y^{(i)})\,, \tag{8}$$

where π_i are the weights of the components for the mixture which corre-

spond to the probability of the states. The initial state distribution (for $t = 1$) is given by $\pi_i = p(X(1) = i)$ for $i = 1, \ldots, N$. Thus a HMM can be described by $\lambda = \{A, B, \pi\}$ where $A = \{P\}$ and $B = \{\mu_i, \Sigma_i\}_{i=1\ldots,N}$. Baum et al. (1970) proved that the parameters maximizing the Q-function in the M-step in Eq.5 for a HMM are analytical tractable. In fact the E and M steps are performed simultaneously after first computing a set of marginal posterior probabilities for the hidden states (the facies) through a series of forward and backward recursions (Baum et al. 1970). The full algorithm is known as the Baum-Welch algorithm, and is shown in detail in Appendix A.

3 First Application

In this section the proposed facies classification methodology is applied to a dataset from an unconventional reservoir in North America. Data are acquired in a well in the Marcellus shale in two different depth intervals: between $2,072m$ and $2,195m$ and between $2,398m$ and $2,473m$. Three different lithologies are present: shale, limestone and sandstone; shale is predominant (more than 60% in the two intervals), sand is characterized by very low porosity (less than 5%). The dataset includes a set of measured well logs: sonic (P- and S- wave velocity V_p and V_s) and petrophysical (neutron porosity, density, gamma ray and resistivity). A standard formation evaluation analysis has been performed in order to compute the main petrophysical curves: total porosity ϕ_t, and volumetric fractions of lithological components of the solid phase (volumes of clay, quartz and calcite: Vol_{clay}, Vol_{qua} and Vol_{cal}). Finally a set of elastic and geomechanics attributes such as Young's modulus E, Poisson ratio P_r, and photoelectric index pe, is also available. The well log data are displayed in Fig.2. Two different intervals of the well log profile are considered, an upper part $(2,072 - 2,195m)$ and a lower part $(2,398 - 2,473m)$, see Fig.2.

For this application, four lithological facies are defined: 1 sandstone cemented with calcite, 2 sandstone cemented with quartz, 3 limestone and 4 shale, represented by the colors white, light-grey, dark-grey and black respectively. These facies have been characterized using core samples from the lower part of the well log profile. A stratigraphic profile for the lower interval derived from depositional and sedimentological models is also available and identifies three main layers: a thick shaley layer on top, a limestone layer in the middle, and a tight cemented sandstone layer on the bottom split into a quartz-filled layer and a calcite-filled layer, see Fig.2. However the stratigraphic profile does not show local heterogeneities and thin layers that are

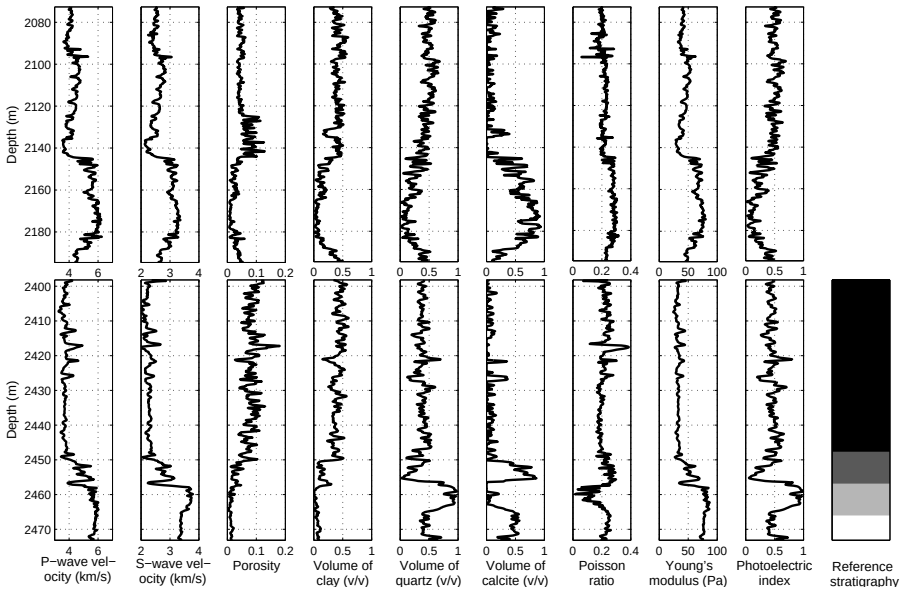


Figure 2: Set of well logs for the upper (top) and lower (bottom) interval for the first application. For the lower interval, a reference stratigraphic facies profile is also displayed.

visible in petrophysical curves in well logs. The facies classification method is therefor applied to obtain a more detailed log-facies classification.

The application of the method is divided into three parts: 1) Classification of facies with given proportions, transition probabilities and output probabilities, that is supervised learning; 2) Classification of facies with given output probabilities and estimation of proportions and transition probabilities, that is partly unsupervised learning; and 3) Classification of facies and estimation of proportions, transition probabilities and output probabilities, that is fully unsupervised learning. In the first part, all the parameters of the HMM are thus assumed to be known and the classification method is applied using these parameters. The facies classifications are given by the maximum a posteriori (MAP) prediction which is the most probable classification under the given models. In the second part, the output probabilities only are assumed to be known and estimation of the most likely classification together with proportions and transition probabilities is performed simultaneously. In the last part, the most likely classification and all the HMM parameters are estimated simultaneously. For all three examples, the estimation and classification is first performed in the lower part, where the re-

sults can be compared to the stratigraphic profile. Furthermore, the same method is applied to the upper part, applying the parameter estimates from the lower part, to obtain a consistent log-facies classification in both depth intervals. The classification is also compared applying a) the petro elastic well logs $\{V_p, V_s, \phi_t, Vol_{clay}, Vol_{qua}, Vol_{cal}\}$ only, b) the geomechanical well logs $\{P_r, E, pe\}$ only and c) all well logs.

In the last part of this work, a sensitivity analysis on the number of well logs that should be used in the classification is performed, in order to reduce the dataset to a set of relevant well logs. Finally, the results of the application are compared to traditional statistical classification techniques, such as naive linear and non-linear (quadratic) discriminant analysis and the Gaussian mixture model without spatial correlation.

3.1 Classification of Facies with Given Proportions, Transition Probabilities and Output Probabilities

For the first part of the application, the stratigraphic profile is used to empirically estimate all the parameters of the HMM, hence a supervised learning case. The empirical proportions of the facies are estimated from the global proportions in the stratigraphic profile $\hat{\pi}_s = (0.10.0.12, 0.12, 0.66)$, which is chosen as initial distribution, that is $\hat{\pi} = \hat{\pi}_s$. The prior model transition probability parameters are estimated by an empirical counting process throughout the stratigraphic profile

$$\hat{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0.02 & 0.98 & 0 & 0 \\ 0 & 0.02 & 0.98 & 0 \\ 0 & 0 & 0.01 & 0.99 \end{pmatrix}. \tag{9}$$

Notice the high occurrence of zero-probabilities in the estimated transition matrix due to the lack of transitions in the reference facies. For example, calcite-cemented sandstone (white) act as an absorbing state, that is if a transition into it occurs the process can never have transitions out of it again. However in order to achieve a higher resolution classification and to achieve a more flexible model by avoiding absorbing states, all the entries are manually set to be non-zero because geologically none of these transitions are impossible. In the following cases, with unknown stratigraphic profile, a minimum probability of 0.01 is similarly set in each entry before running the EM

algorithm. The adjusted transition matrix estimate is

$$\hat{P}^{adj} = \begin{pmatrix} 0.97 & 0.01 & 0.01 & 0.01 \\ 0.02 & 0.96 & 0.01 & 0.01 \\ 0.01 & 0.02 & 0.96 & 0.01 \\ 0.01 & 0.01 & 0.02 & 0.96 \end{pmatrix} . \tag{10}$$

Furthermore, due to the large thickness of the layers in the stratigraphic log, the probabilities on the diagonal might still be too large. Finally, the Gaussian likelihood mean and covariance matrix parameters are estimated by standard ML methods for each facies. In other words, the well log data $\mathbf{y}$ is seperated by stratigraphic facies into $N = 4$ subsets $\{\mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(N)}\}$ with respective dimensions $\{n_1, \ldots, n_N\}$. The maximum likelihood parameter estimates for each class $l \in \{1, \ldots, N\}$ are then given by

$$\hat{\boldsymbol{\mu}}_l = \frac{1}{n_l} \sum_{i=1}^{n_l} y_i^{(l)} \quad , \quad \hat{\boldsymbol{\Sigma}}_l = \frac{1}{n_l - 1} \sum_{i=1}^{n_l} \left(y_i^{(l)} - \hat{\boldsymbol{\mu}}_l \right) \left(y_i^{(l)} - \hat{\boldsymbol{\mu}}_l \right)^T . \tag{11}$$

The resulting Gaussian likelihood model for the output probabilities is displayed in Fig.3 by pairwise 90% confidence regions separated by facies. In the pairwise plot, diagonality of the regions indicate correlation between variables. In addition, if there is almost one-to-one correspondence between variables it is represented by very narrow region width, which indicate that the variables are essentially the same (up to some linear scale). Notice that this is the case for the pairs a) V_p and E and b) Vol_{qua} and pe, hence one could choose only one log out of each of these pairs for the classification. The likelihood models displaying the most significant separation between the facies, with significant horizontal separation between the region centers, are for the well logs V_p, ϕ_t, Vol_{clay} and E. Notice that the Gaussian pdfs in some plots cover unrealistic negative regions, for example negative volumetric fractions, hence the assumption of a Gaussian likelihood models might not be best suited for these well logs. The Gaussianity is here chosen to be kept however, due to its superior analytical tractability. Gaussian pdfs should be then truncated to avoid non-realistic values.

The HMM facies MAP classifications with plug-in parameter estimates are given in Fig.4 for both the upper and lower part. The MAP predictions have coarse layers as expected because of the Markov chain prior model. For the lower part, with the reference stratigraphic profile, all three classifications are

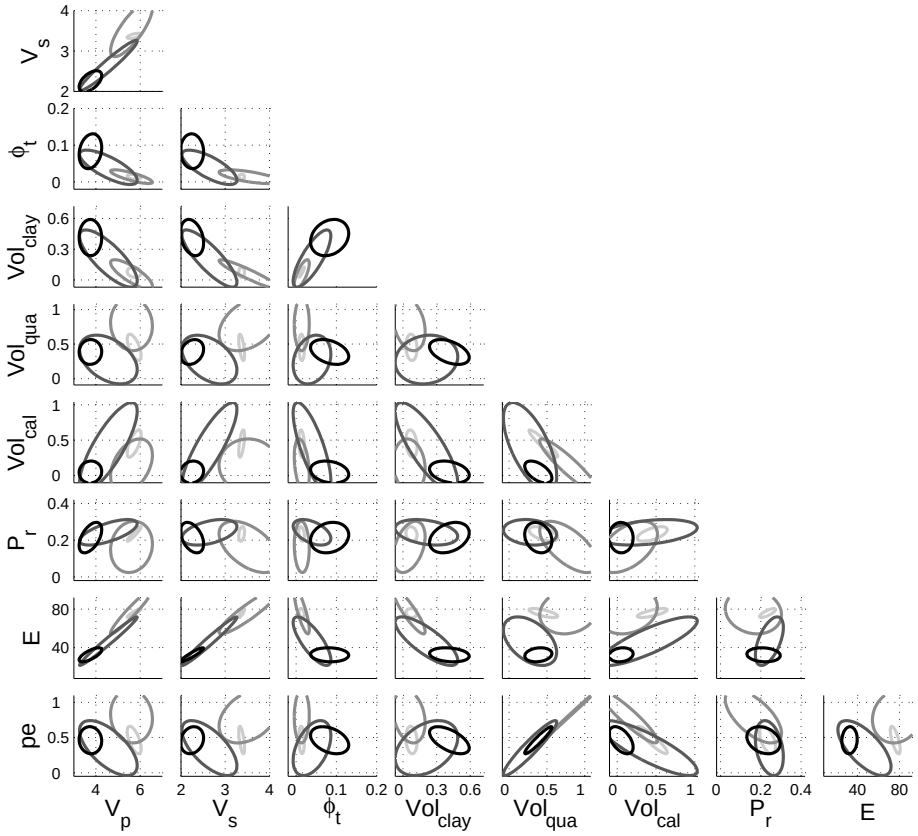
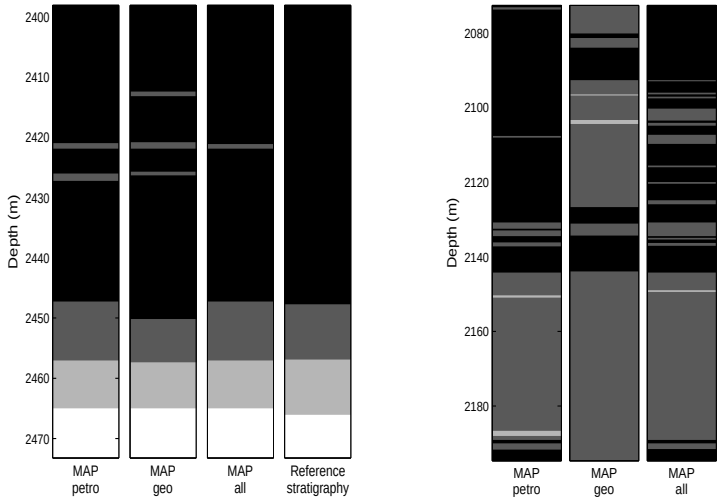


Figure 3: 2D plots of 90% confidence regions of the Gaussian likelihood model for the output probabilities for each facies class estimated empirically.

similar to the reference stratigraphy. In the upper part, in which there is no reference stratigraphic profile, notice that the MAP prediction applying the geomechanical well logs only has significantly coarser layers than the other two, and also that the classification at depth about 2,072m-2,145m differs significantly.

3.2 Classification of Facies with Given Output Probabilities and Estimation of Proportions and Transition Probabilities

For the second part of the application, the Gaussian likelihood parameters estimated in the first part is used, while all the prior parameters of the Markov model are reestimated assuming the facies to be unknown, which is denoted a partly unsupervised learning case. The parameters are estimated by the



(a) Lower part **(b)** Upper part

Figure 4: Comparison of the MAP predictions for case 1.

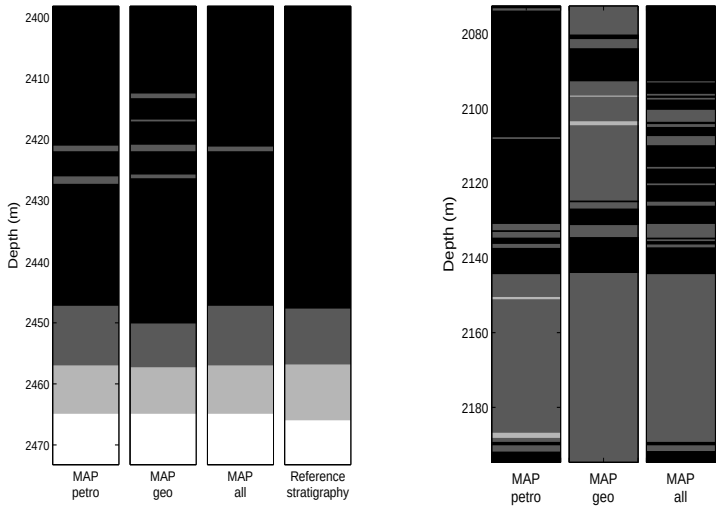
Baum-Welch algorithm in Appendix A, in which the likelihood parameters are not updated but are set. The EM-algorithm is run with 100 iterations to ensure convergence (in practice all runs converged in less than 10 iterations). The estimation is performed for the three well log sets mentioned, that is for the petro elastic well logs, the geomechanical well logs and all well logs with respective superscripts p, g, a. The estimated transition matrices are

$$\hat{P}^p = \begin{pmatrix} 0.97 & 0.01 & 0.01 & 0.01 \\ 0.02 & 0.96 & 0.01 & 0.01 \\ 0.01 & 0.01 & 0.95 & 0.03 \\ 0.01 & 0.01 & 0.01 & 0.97 \end{pmatrix}, \quad \hat{P}^g = \begin{pmatrix} 0.97 & 0.01 & 0.01 & 0.01 \\ 0.02 & 0.96 & 0.01 & 0.01 \\ 0.01 & 0.02 & 0.92 & 0.05 \\ 0.01 & 0.01 & 0.02 & 0.96 \end{pmatrix}$$

$$\hat{P}^a = \begin{pmatrix} 0.97 & 0.01 & 0.01 & 0.01 \\ 0.02 & 0.96 & 0.01 & 0.01 \\ 0.01 & 0.02 & 0.95 & 0.02 \\ 0.01 & 0.01 & 0.01 & 0.97 \end{pmatrix}. \tag{12}$$

The estimated transition matrices resemble the reference matrix in Eq.(9) with large probabilities along the diagonal. The high values on the diagonal of the matrix combined with the low values outside the diagonal show the presence of thick layers (high probability of transitions from one facies to itself) and few transitions between different facies.

The corresponding HMM facies MAP classifications are given in Fig.5. For the lower part, all three classifications are again similar to the stratigraphic profile. This is consistent with the information from depositional models and core photos where there was observed a higher heterogeneity in the upper part.



(a) Lower part **(b)** Upper part

Figure 5: MAP predictions for case 2.

3.3 Classification of Facies and Estimation of Proportions, Transition Probabilities and Output Probabilities

For the third part of the application, all parameters are reestimated by the Baum-Welch algorithm in Appendix A, assuming the facies stratigraphy to be unknown, hence a fully unsupervised learning case. The estimated transition

matrices are

$$\hat{P}^p = \begin{pmatrix} 0.93 & 0.02 & 0.04 & 0.01 \\ 0.02 & 0.96 & 0.01 & 0.01 \\ 0.03 & 0.01 & 0.90 & 0.06 \\ 0.01 & 0.01 & 0.02 & 0.96 \end{pmatrix}, \hat{P}^g = \begin{pmatrix} 0.97 & 0.01 & 0.01 & 0.01 \\ 0.02 & 0.95 & 0.01 & 0.02 \\ 0.01 & 0.03 & 0.90 & 0.06 \\ 0.01 & 0.01 & 0.02 & 0.96 \end{pmatrix}$$

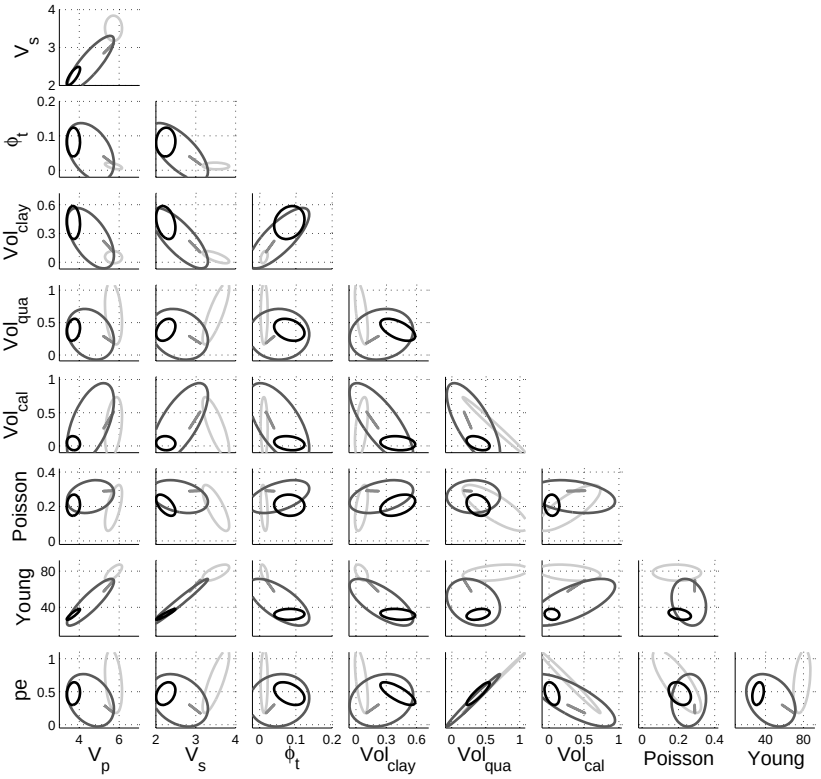
$$\hat{P}^a = \begin{pmatrix} 0.97 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.49 & 0.49 & 0.01 \\ 0.03 & 0.01 & 0.91 & 0.05 \\ 0.01 & 0.01 & 0.02 & 0.96 \end{pmatrix}. \quad (13)$$

Notice again the large diagonal probabilities, however for this unsupervised case there is more uncertainty present, hence some of the off-diagonal elements are a bit larger than for the reference transition probability matrix in Eq.(10). The Gaussian likelihood model for the output probabilities estimated by the EM algorithm is presented in Fig.6 by pairwise 90% confidence regions for the respective well log sets. The regions resemble the empirical estimates in Fig.3 quite reliably.

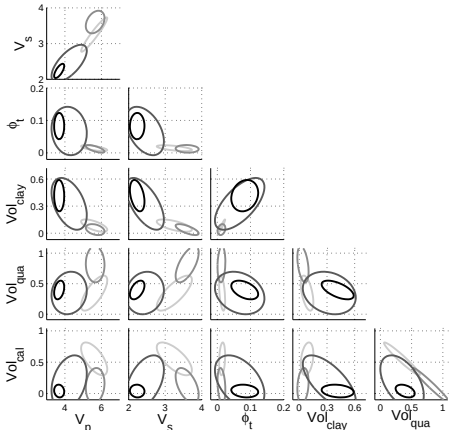
The HMM facies MAP classifications are given in Fig.7 displaying more rapid transition than for the previous cases due to being a fully unsupervised learning case. For the lower part, the classification from the petro elastic logs and all logs are poorer, the MAP prediction from all well logs in particular does not recognize the quartz-filled sandstone layer at all. The classification from the geomechanical logs is however reliable, being the smallest subset of well logs. In the upper part, where it is speculated that the absence of sandstone is due to the lack of visible evidence from core photos, all classifications provide a satisfactory results.

3.4 Sensitivity Study

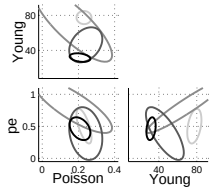
For this part, a sensitivity analysis is performed in the lower well part on the number of well logs that should be used in the classification. Classification of facies and estimation of all parameters, according to the fully unsupervised case of the previous part, is performed for all (511) subsets of the nine well



(a) All well logs

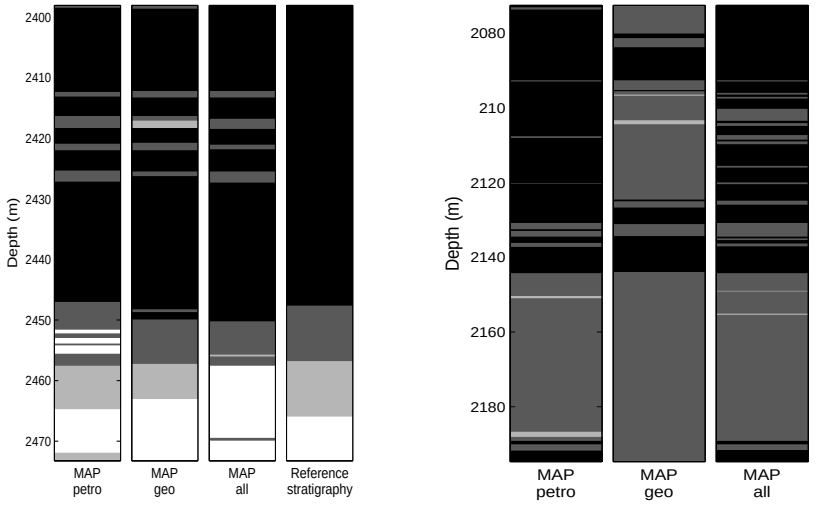


(b) Petro elastic



(c) Geomechanical

Figure 6: 2D plots of 90% confidence regions of the Gaussian likelihood model for the output probabilities for each facies class estimated by EM.



(a) Lower part **(b)** Upper part

Figure 7: MAP predictions for case 3.

logs. As a measure of fit, the predictive performance is evaluated by the test statistics: location-wise mismatch ratio c_1, mismatch in total number of transitions between different facies layers c_2 and difference in estimated transition matrices c_3. These are defined for $c_i \in [0,1]$, $i = 1,2,3$ as

$$c_1 = \frac{1}{T}\sum_{t=1}^{T} I(\hat{x}_t = x_t) \tag{14}$$

$$c_2 = 1 - \frac{1}{100}\min\{100, |\rho(\hat{\mathbf{x}}) - \rho(\mathbf{x})|\} \tag{15}$$

$$c_3 = 1 - \frac{1}{N/2}\min\left\{N/2, \sum_{i=1}^{N} |\hat{P}_{ii} - \hat{P}_{ii}^{adj}|\right\} \tag{16}$$

where $\mathbf{x} = (x_1, \ldots, x_t)$ is the reference stratigraphy, $\hat{\mathbf{x}} = (\hat{x}_1, \ldots, \hat{x}_t)$ the MAP prediction, $\rho(\cdot)$ the number of layer transitions in its input, $\hat{P}$ the transition matrix estimated by the EM algorithm and $\hat{P}^{adj}$ the reference transition matrix in Eq.(10). As there are only three transitions between different facies layers in the reference stratigraphy, the statistic c_2 will favor predictions with few such class transitions. Observe that c_3 is defined from the estimated transition matrices diagonal elements only, and can hence be regarded as a measure of how well the thickness of the facies layers are predicted. For both c_2 and c_3 upper limits are set, that is a maximum on mismatch in number

of total transitions and on transition matrix diagonal. All three statistics become one for a perfect match, when the MAP prediction is identical to the reference facies stratigraphy. The three statistics are given weights 0.5, 0.25 and 0.25 respectively, hence most emphasis is on the facies mismatch. The set of the well logs that maximizes the corresponding weighted sum of the test statistics

$$c_{sum} = 0.50c_1 + 0.25c_2 + 0.25c_3 \tag{17}$$

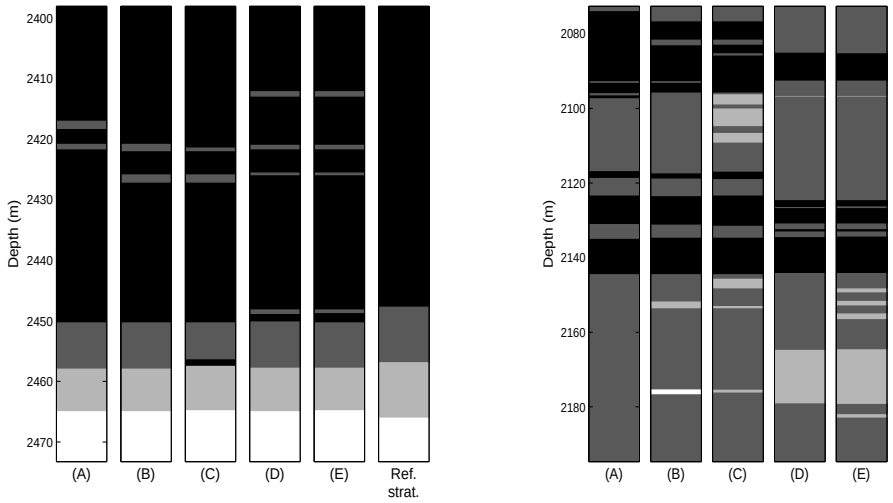
will be defined as the best subset of logs.

The ten best subsets of well logs in this study are given in Table 1, for which the best subset was found as $\{V_p, V_s, Vol_{clay}\}$. Notice also that the small subset of $\{V_s, Vol_{cal}\}$ was actually third best, reducing the needed number of well logs from nine to two. The sonic S-wave well log V_s has most occurrences, appearing in all but two of the ten best subsets, while the sonic P-wave well log V_p appears only once, however in the best subset. It therefore seems that only one of the two sonic logs is needed to make a reliable prediction, as they are highly correlated. It should also be pointed out that the resolution of sonic logs is generally lower than the resolution of petrophysical logs and petrophysical computed curves. The five best subset MAP predictions are displayed in Fig.8. Notice that the classifications in the upper part of the well have thicker layers than the previous classifications in cases 1 through 3. For the upper part, some quartz-filled sandstone layers are predicted for subsets (C) through (E).

It might seem counterintuitive that smaller subsets, even as small as two well logs only, have better predictive performance than the full set of well logs, as taking advantage of more information should provide more reliable results. Reasons for the opposite can be: the model is not perfect, that is there are model errors, for example the Gaussian assumption is probably not best suited for all well logs, also there is model parameter uncertainty as the estimates are used as plug-in values only. Too much information thus contains much noise, that is for the current model some of the well logs might act more as noise than being productive. Besides, different logs can be affected by different errors, resulting in contradictory information. This might also be the reason to why the smallest subset in case 3, with the three geomechanical well logs only, has a more reliable MAP prediction than for the others, see Fig.7. A more sophisticated hierarchical Bayesian model could account for model parameter uncertainty by assigning prior models to the parameters, also output probability functions other than the Gaussian could be chosen, however this is not considered further in this study.

Well logs	c_{sum}	
V_s, ϕ_t, Vol_{clay}, Vol_{qua}	0.9302	
V_s, ϕ_t, Vol_{clay}, $Poisson$, pe	0.9307	
V_s, Vol_{qua}, Vol_{cal}, $Poisson$	0.9310	
V_s, ϕ_t, Vol_{qua}, Vol_{cal}, $Poisson$	0.9313	
V_s, Vol_{clay}, $Poisson$, pe	0.9344	
ϕ_t, $Poisson$, $Young$	0.9352	(E)
V_s, ϕ_t, $Poisson$	0.9384	(D)
V_s, Vol_{cal}	0.9389	(C)
Vol_{clay}, Vol_{cal}, $Young$	0.9396	(B)
V_p, V_s, Vol_{clay}	0.9415	(A)

Table 1: Sensitivity study results: The ten best subsets.



(a) Lower part **(b)** Upper part

Figure 8: MAP predictions from the sensitivity study for the five best subsets (A) through (E) of Table 1.

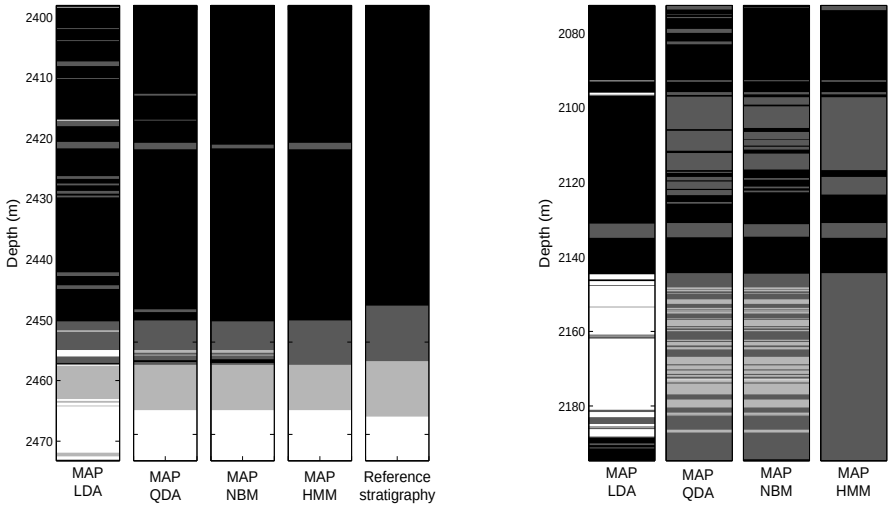
3.5 Comparison to Other Statistical Classification Methods

In the last part, the facies predictions from the HMM are compared to that of three other statistical classification techniques (Hastie et al. (2009)), namely linear discriminant analysis (LDA), quadratic discriminant analysis (QDA) and a naive Bayesian model (NBM) (Li and Anderson-Sprecher 2006). All three comparison methods neglect spatial correlation, see the graph in Fig.1b. LDA and QDA are regression techniques and correspond to classification from the Gaussian output probabilities, as in Eq.(7), only, also for LDA equal covariance between the Gaussian class functions is assumed. NBM is a Bayesian model without spatial correlation, where the prior model on the facies for each depth is set to the global proportions. The HMM predictions are thus expected to be coarser than those of the three other methods, as spatial correlation and thicker layers are enforced through the Markov chain assumption, see Fig.1a.

In the predictions it is applied the best subset $\{V_p, V_s, Vol_{clay}\}$ as found in the previous part, assuming that all parameters are known (corresponding to the supervised case 1). The MAP predictions from the four classification methods are given in Fig.9. All classifications have quite reliable location-wise match, however for NBM, QDA and especially LDA some false thinner layers are predicted. In the lower part, although some of the interbedded layers within the main shale layer could be realistic, there are several misclassification of thin layers, especially in the limestone. Hence, incorporating spatial dependency in the facies through the Markov property seems the approach providing the best result. In the upper part, LDA clearly misclassify the main limestone layer at the bottom of the interval, whereas QDA and NBM underestimates the spatial continuity of the layer.

4 Second Application

In this section the proposed facies classification methodology is applied to a conventional reservoir dataset in the North Sea in the susurface depth interval between $1,797m$ and $1,933m$. The geological environment is a conventional clastic reservoir with three fluids: gas, oil and water. The well log dataset contains standard wireline logs including sonic and petrophysical data, and computed curves from formation evaluation analysis such as porosity, mineralogical fractions and saturation curves (water, oil and gas saturation s_w, s_o, s_g), presented in Fig.10. From well measurements it is known that the contact between gas and oil occurs at depth about $1,870m$, see Fig.10.



(a) Lower part **(b)** Upper part

Figure 9: MAP predictions from the comparison study.

For this application, four litho-fluid facies are defined: gas-filled sandstone, oil-filled sandstone, water-filled sandstone and shale represented by the colors white, light-grey, dark-grey and black respectively. The gas cap at the top of the reservoir is about $20m$ thick and is located at the top of an anticline structure characterized by high-porosity sandstone. The lower part of the interval includes thin layers of oil sand and water sand alternated to shale layers. In this application, it is assumed that all parameters are unknown, as there is no reference stratigraphy, hence it is a fully unsupervised case to which the Baum-Welch algorithm is applied.

In this case there are gravity effects present, with impossible transitions downwards, that is gas-filled sandstone is never below oil- and water-filled sandstone and oil-filled sandstone is never below water-filled sandstone. These restrictions are enforced in the Markov model transition probability matrix P by setting the current probabilities to 0. A nice property of the Baum-Welch algorithm is that 0-probabilities in P are kept when updating, hence one only needs to set these probabilities to 0 in the initial guess on P. When running the EM-algorithm, the minimum non-zero transition probability is set to 0.001.

Clearly the fluid discrimination can be achieved by using saturation properties only, however saturation curves can be noisy and uncertain, because they are computed from resistivity measurements. Another problem with the use

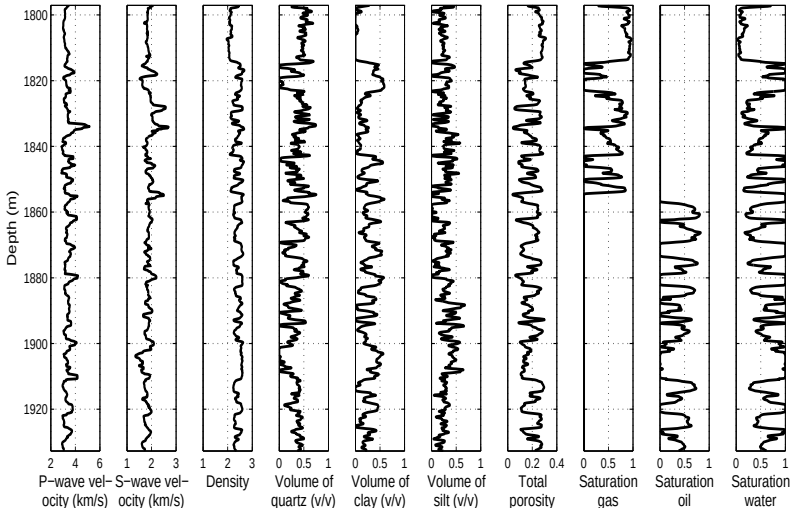


Figure 10: Well log dataset for the second application example.

of saturation curves in the classification is related to the applicability of the classification to the entire reservoir where saturation information is not available. A first sensitivity analysis on the use of saturation logs is performed to determine the impact on the facies model. Then, the four litho-fluid facies are classified based on the full set of parameters, then classification from the elastic properties only. The goal is to assess the value of information of sonic logs in the classification. As a matter of fact, after the calibration at the well location, litho-fluid classification is generally applied to seismic data where only elastic properties, such as velocities, are available. Therefore the facies discrimination based on elastic properties is crucial in seismic reservoir characterization, even though the best classification at the well location is generally achieved through petrophysical properties. For each of the two cases under investigation, full set of well logs (case 1) and petrophysical properties only (case 2), the proposed methodology is applied to estimate the HMM parameters and the most probable facies profile (cases 1a and 2a); then these parameters are used to estimate the most probable facies profile from elastic properties only (cases 1b and 2b). Although seismic data and inverted seismic attributes are not available for this case study, the facies classification from elastic properties only is a reliable test to assess the feasibility of the application.

4.1 Litho-Fluid Classification with Sensitivity to Saturation Logs

The first classification is performed by applying the Baum-Welch algorithm to the set of petroelastic curves. In Fig.11 the results obtained under three different settings are shown: petroelastic logs assuming the gas-oil contact known, petroelastic logs assuming the gas-oil contact unknown and entire set of well logs including saturation curves assuming the gas-oil contact unknown. For the first setting, the well log interval is divided in two subintervals: above and below the contact and the proposed methodology is applied by assuming three facies {gas-filled sandstone, water-filled sandstone, shale} above and {oil-filled sandstone, water-filled sandstone, shale} below. By comparing the results of the first two settings it is concluded that the contact can be automatically detected (Fig.11). This result is speculated to occur due to the clear petro-elastic signature of gas sand: sand filled by gas are characterized by low V_p/V_s ratio and low density; furthermore sand at the top of the reservoir also show higher porosity and lower clay content compared to the lower part. However the discrimination between water sand and oil sand and between water sand and shale is less satisfactory (Fig.12) due to the similar response of these facies. As previously stated, if the saturation logs are included in the classification, the discrimination between fluids clearly improve (Fig.11, last plot), however the final classification almost exclusively depends on the data whereas the spatial model becomes secondary.

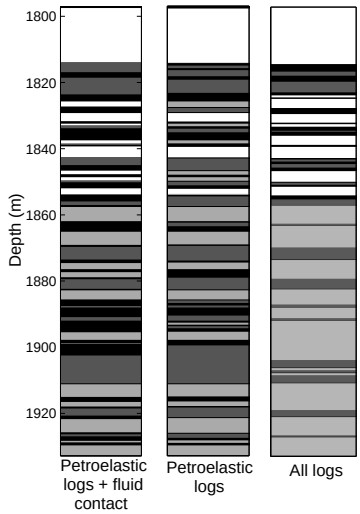


Figure 11: MAP predictions for three subsets of logs: petroelastic properties (assuming the gas-oil contact known); petroelastic properties (assuming the gas-oil contact unknown); all logs (assuming the gas-oil contact unknown).

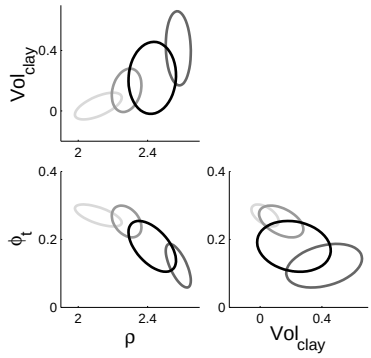


Figure 12: 2D plots of 90% confidence regions of the Gaussian output probabilities for each facies class estimated by EM for the second application (petrophysical dataset).

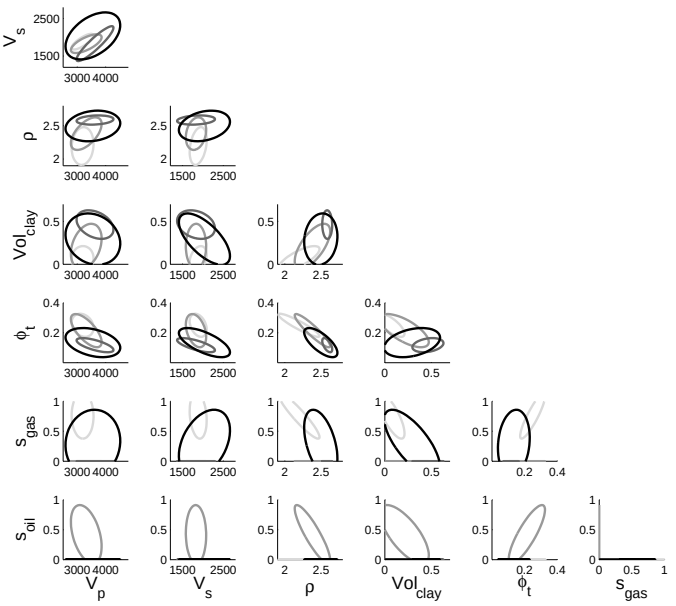
4.2 Litho-Fluid Classification From the Full Set of Well Logs

The inversion and estimation is here performed from the full set of well logs chosen as $\{V_p, V_s, \rho, \phi_t, Vol_{clay}, s_g, s_o\}$, denoted Case 1a). The estimated transition matrix is:

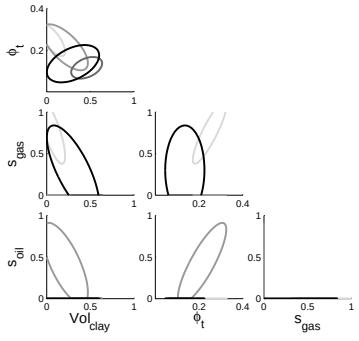
$$\hat{P} = \begin{array}{c} \begin{matrix} gas & oil & water & shale \end{matrix} \\ \begin{pmatrix} 0.961 & 0 & 0 & 0.0389 \\ 0 & 0.973 & 0.026 & 0.001 \\ 0 & 0 & 0.896 & 0.104 \\ 0.072 & 0.098 & 0.062 & 0.768 \end{pmatrix} \end{array} \begin{matrix} gas \\ oil \\ water \\ shale \end{matrix} \tag{18}$$

Notice the estimated 0-probabilities in P_{12} and P_{13}, hence gas-filled sandstone will only have transitions into gas-filled sandstone or shale. Oil-filled sandstone will have almost exclusively transitions into oil- or water-filled sandstone. The Gaussian likelihood model for the estimated outcome probabilities is presented in Fig.13a. Notice that the well logs ρ and Vol_{clay} has the most significant separation between the litho-fluid facies. Notice also the estimated negative correlation between ϕ_t and ρ, with increasing density leading to decreasing porosity as expected. Finally it is pointed out that the distributions of fluid saturations are represented by degenerate Gaussian pdfs due to the non-Gaussian behavior of these properties, characterized by distribution peaks at the extreme values of the range (that is saturation close

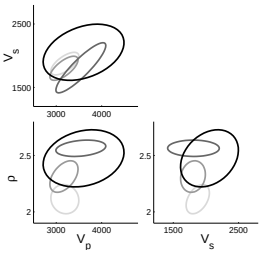
to 0 and 1). A normal score transformation could be introduced to better describe this behavior. The EM algorithm at certain locations mix the two classes shale and water sand. The main reason may be the similar elastic properties, in terms of density and velocity, of this clay type compared to higher porosity sandstone filled by water.



(a) Case 1a)



(b) Case 2a)



(c) Case 2b)

Figure 13: 2D plots of 90% confidence regions of the Gaussian likelihood model for the output probabilities for each facies class estimated by EM for the second application.

The HMM facies MAP classification for Case 1a) is given in Fig.14. The match between the predicted gas- and oil-filled sandstone layers and the gas and oil saturation well logs is reliable, that is gas is only present in the upper section while oil is only present in the lower section. Notice that the enforced gravity effects are honored, due to the 0-probabilities in the transition matrix, thin shale layers are enforced between some layers of hydrocarbon- and water-filled sandstone. Because of the possible mix between water-filled sandstone and shale some of these may be misspecified by each other, for example shale is highly under-represented in the lower section compared to the base case MAP. The hydrocarbon classes of most interest seem to be reliably predicted, however the predicted oil-filled sandstone layers are thicker than in the base case MAP.

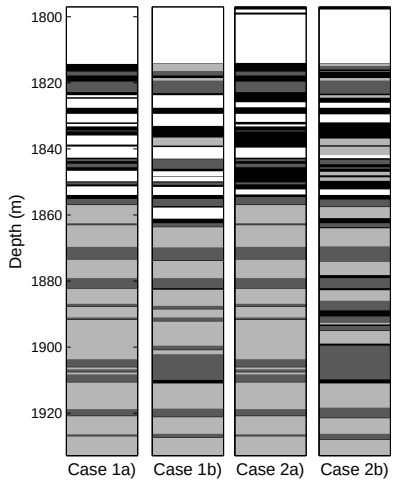
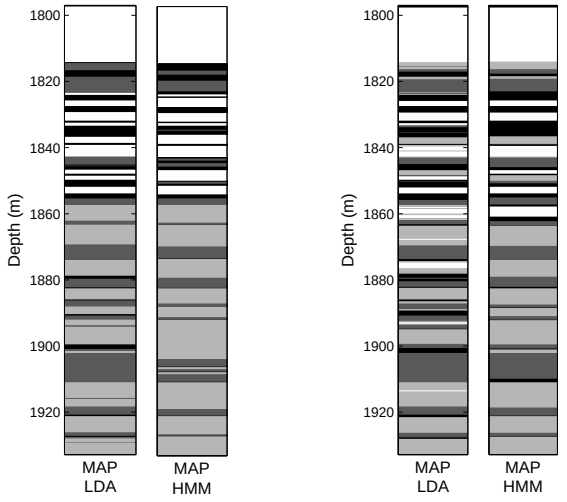


Figure 14: The MAP classifications for cases 1a-b) and 2a-b) for the second application.

For Case 1b), the inversion is performed based on the elastic well logs $\{V_p, V_s, \rho\}$ only, with all initial parameters set to the estimates from Case 1a). The corresponding facies MAP prediction is presented in Fig.14, and is quite similar to the MAP prediction from the full set of well logs. For this case, there are a few thin oil-filled sandstone layers present also above the contact, but no gas-filled sandstone layers are present below the contact. Hence inversion from the elastic well logs only gives a reliable facies prediction, as is the case of seismic data.

For Case 1, a comparison to linear discriminant analysis (LDA) is also per-

formed, which is the simplest model out of the three comparison models considered in the first application. The MAP prediction comparisons for Case 1a) are displayed in Fig.15a and are quite similar between the classification methods. For this application, the data (through the likelihood model) therefore seem to dominate the prior model. For the HMM however, thin shale layers are enforced between water-filled sandstone and oil-filled sandstone downwards as mentioned, to avoid impossible transitions due to gravitational effects. These shale layers should be thicker, parts of the water-filled sandstone layers next to it might be shale as the two classes are mixed. The MAP prediction comparisons for Case 1b) are displayed in Fig.15b. This MAP prediction for LDA has more rapid layer transitions than for the HMM, hence the spatial Markov prior model has more weight. When removing the saturation logs in particular, the data do not seem to dominate the prior model as much. Also, the LDA classification has some false gas-filled layers also below the contact.



(a) Case 1a) **(b)** Case 1b)

Figure 15: Comparison of the MAP predictions between LDA and HMM in the second application computed computed from (a) the full set of well logs and (b) the elastic well logs only.

4.3 Litho-Fluid Classification From Petrophysical Properties

The inversion and estimation is here performed from the petrophysical set of well logs $\{\phi_t, Vol_{clay}, s_g, s_o\}$ only, denoted Case 2a). The estimated transition

matrix is:

$$\hat{P} = \begin{array}{c} \begin{array}{cccc} gas & oil & water & shale \end{array} \\ \left(\begin{array}{cccc} 0.962 & 0 & 0.001 & 0.037 \\ 0 & 0.973 & 0.026 & 0.001 \\ 0 & 0 & 0.899 & 0.101 \\ 0.076 & 0.105 & 0.067 & 0.752 \end{array} \right) \end{array} \begin{array}{c} gas \\ oil \\ water \\ shale \end{array} \tag{19}$$

The estimated transition matrix is very similar to the estimate of Eq.(18) from the full set of well logs. The Gaussian likelihood model for the estimated outcome probabilities is presented in Fig.13b. Notice again the degenerate Gaussian pdfs of the fluid saturation.

In Case 2b), the inversion and estimation is performed from the elastic well logs $\{V_p, V_s, \rho\}$ only, similarly to Case 1b), but keep the estimated transition matrix in Eq.(19), hence estimating the likelihood parameters only. The Gaussian likelihood model is presented in Fig.13c. Again, ρ has the most significant litho-facies separation.

The HMM facies MAP predictions for Case 2a) and b) are given in Fig.14. For both classifications, gas-filled sandstone is only present above the contact as desired. In the classification for the elastic well logs of Case 2b), gas-filled sandstone is slightly under-represented, and again some oil-filled sandstone layers are present above the contact. Both MAP predictions however resemble the MAP prediction from all well logs in Case 1a) reliably.

5 Discussion

Several methods, deterministic and statistical algorithms, can be used in facies classification. The previous applications show that different methods can provide different results in terms of facies discrimination. Furthermore, the same method can provide different results when different assumptions are introduced. For example, the same method applied to two different set of well logs from the same well, can provide a different classification. The only reliable approach to assess the validity of the facies classification is through an extensive analysis of core samples and integrated study combining quantitative log interpretation, rock physics, sedimentology and stratigraphy and laboratory measurements. The goal of this paper is to show the necessity of

introducing a spatial statistics model to represent the vertical continuity of the facies profile and quantify the transitions between different facies.

Once the spatial model is known and the transition probabilities are determined, the most probable classification can be estimated. In most of the real case applications, the spatial model is not known. The Baum-Welch algorithm can be used to estimate the most probable classification and the underlying parameters of the spatial model and the rock physics likelihood simultaneously. However it is pointed out that this inverse problem of facies classification is very challenging to solve due to the high number of unknown parameters and the uncertainty and noise in data measurements. Therefore the solution is not necessarily unique and multiple solution can equally honor the same measured dataset. The Baum-Welch algorithm provides the most likely facies profile by maximizing a likelihood function. This method relies then on the quality of the measured dataset.

In case studies where an extensive dataset in terms of well logs, laboratory measurements, core samples and prior geological information is available, the spatial continuity model as well as the other parameters of the HMM can be assumed set from the available prior information and the Baum-Welch algorithm can be used to predict the most likely facies classification with the given prior parameters (as shown in the first application on the first dataset). The advantage of this approach is that prior geological information can be included in the methodology. For example, when several wells are available in the same field, the overall facies proportions could be assumed from the reservoir model and the transition probabilities could be determined from the knowledge of the layer thicknesses and stratigraphic sequences of the reservoir model. The proposed method can then be applied for the log-facies classification at the well classification with the previously assessed parameters. These parameters are always the results of assumptions and models, therefore a sensitivity analysis of the effect of these assumptions should be performed after the classification.

Finally, the selection of the dataset to use in the facies classification is another key aspect of the method. In the first application, a statistical method to determine the optimal dataset to be used in the classification was shown. From the statistical point of view, the optimal dataset is the one that allows obtaining the best discrimination between the assigned facies. However, in reservoir studies, the dataset to be used in the classification should also account for data availability, quality of the data and goal of the classification. Some of the well logs have lower resolution or are noisier than others; other logs can be available only at limited well locations. The second application

also shows that some of the unknown parameters can be estimated from the full dataset, but the classification should then be applied only to a subset of logs, in order to be extended to the entire reservoir model where only elastic properties are available.

6 Conclusion

In this paper, a new methodology to classify log-facies, lithological or litho-fluid types, from well log data is presented. The proposed methodology aims to estimate log-facies based on common rock, fluid and elastic properties but also to mimic realistic vertical distributions. The method is based on two statistical tools: hidden Markov chain models and the Expectation-Maximization algorithm. The introduction of hidden Markov models allows us to introduce constraints on facies transitions. The transition probabilities are then combined with the likelihood function of rock samples to belong to a certain facies given its rock and fluid properties. Since hidden Markov models require a set of parameters for the transition probabilities as well as for the parametric likelihood functions, the Expectation-Maximization algorithm is applied to automatically estimate those parameters from well logs, the particular algorithm known as the Baum-Welch algorithm. The main advantage of this method is the ability of describing realistic geological sequences as shown in the presented applications. Furthermore, the method is flexible and can be combined with prior geological information such as overall proportion or likelihood functions estimated from petrophsyical and rock physics models. In other applications, some of the parameters could be assumed based on regional geological information or analogues from nearby fields. The so-obtained classification can then be extended to the entire reservoir field by using traditional Markov chain methods based on the likelihood functions estimated at the well locations.

7 Acknowledgments

Authors acknowledge Enerplus Resources (USA) Corp. for data assistance and University of Wyoming for the support. The work is partially funded by the Uncertainty in reservoir Evaluation (URE) activity - consortium at Department of Mathematical Sciences, NTNU, Trondheim, Norway.

References

Baum, L. E., Petrie, T., Soules, G., and Weiss, N. (1970). "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains." *The Annals of Mathematical Statistics*, 41, 3, 164–171.

Dempster, A. P., L. N. M. and Rubin, D. B. (1977). "Maximum likelihood from incomplete data via the EM algorithm." *Journal of the Royal Statistician Society Ser. B*, 39, 1–38.

Dymarski, P., ed. (2011). *Hidden Markov Models, Theory and Applications*. InTech.

Eidsvik, J., Mukerji, T., and Switzer, P. (2004). "Estimation of geological attributes from a well log: An application of hidden Markov chains." *Mathematical Geology*, 36, 3.

Elfeki, A. and Dekking, M. (2001). "Markov chain model for subsurface characterization: Theory and applications." *Mathematical Geology*, 33, 5.

Grana, D. and Della Rossa, E. (2010). "Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion." *Geophysics*, 75, 3, O21–037.

Grana, D., Pirrone, M., and Mukerji, T. (2012). "Quantitative log interpretation and uncertainty propagation of petrophysical properties and facies classification from rock-physics modeling and formation evaluation analysis." *Geophysics*, 77, 3, WA45–WA63.

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. 2nd ed. Springer.

Krumbein, W. C. and Dacey, M. F. (1969). "Markov chains and embedded Markov chains in geology." *Mathematical Geology*, 1, 1, 79–96.

Li, Y. and Anderson-Sprecher, R. (2006). "Facies identification from well logs: A comparison of discriminant analysis and naive Bayes classifier." *Journal of Petroleum Science and Engineering*, 53, 3-4, 149–157.

Viterbi, A. J. (1967). "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm." *IEEE Transactions on Information Theory*, 13, 2, 260–269.

Wu, C. F. J. (1983). "On the convergence properties of the EM algorithm." *The Annals of Statistics*, 11, 1, 95–103.

Xu, C. and Torres-Verdin, C. (2014). "Petrophysical rock classification in the Cotton Valley tight-gas sandstone reservoir with a clustering pore-system orthogonality matrix." *Interpretation*, 2, 1, T13–T23.

Appendix A: The Baum-Welch Algorithm

Let forward probabilities α indicate conditioning on data up to the current step, that is $\alpha_i(t) = p(X_t = i|y_1,\ldots,y_t)$ and $\alpha_{ij}(t) = p(X_{t-1} = i, X_t = j|y_1,\ldots,y_t)$ while backwards probabilities γ indicate conditioning on all the data, that is $\gamma_i(t) = p(X_t = i|y_1,\ldots,y_T)$ and $\gamma_{ij}(t) = p(X_{t-1} = i, X_t = j|y_1,\ldots,y_T)$. The Baum-Welch algorithm according to Baum et al. (1970) follows, with multivariate Gaussian likelihood models $\phi(y;\mu,\Sigma)$ for each state.

ALGORITHM: BAUM-WELCH ALGORITHM

Set initial parameters $\lambda^* = \{A^*, B^*, \pi^*\}$ where $A^* = \{P^*\}$ and $B^* = \{\mu_i^*, \Sigma_i^*\}_{i=1\ldots,N}$.

Iteratively do:

Forward recursions:

- Initiate:
$$C_1 = \left[\sum_{i=1}^N \phi(y_1;\mu_i^*,\Sigma_i^*) \times \pi_i^*\right]^{-1}$$
$$\alpha_i(1) = C_1 \times \phi(y_1;\mu_i^*,\Sigma_i^*) \times \pi_i^*\,, \quad i = 1,\ldots,N$$
- Iterate for $t = 2,\ldots,T$:
$$C_t = \left[\sum_{i=1}^N \sum_{j=1}^N \phi(y_t;\mu_j^*,\Sigma_j^*) \times P_{ij}^* \times \alpha_i(t-1)\right]^{-1}$$
$$\alpha_{ij}(t) = C_t \times \phi(y_t;\mu_j^*,\Sigma_j^*) \times P_{ij}^* \times \alpha_i(t-1)\,, \quad i,j = 1,\ldots,N$$
$$\alpha_j(t) = \sum_{i=1}^N \alpha_{ij}(t)\,, \quad j = 1,\ldots,N$$

Backward recursions:

- Initiate:
$$\gamma_j(T) = \alpha_j(T)\,, \quad j = 1,\ldots,N$$
- Iterate for $t = T,\ldots,2$:

$$\gamma_{ij}(t) = \frac{\alpha_{ij}(t)}{\alpha_j(t)} \times \gamma_j(t)\,, \quad i,j = 1,\ldots,N$$
$$\gamma_i(t-1) = \sum_{j=1}^{N} \gamma_{ij}(t)\,, \quad i = 1,\ldots,N$$

Update the parameters:

- Update the parameters λ by

$$\pi_i = \gamma_i(1)\,, \quad i = 1,\ldots,N$$
$$P_{ij} = \frac{\sum_{t=2}^{T} \gamma_{ij}(t)}{\sum_{t=1}^{T} \gamma_i(t)}\,, \quad i,j = 1,\ldots,N$$
$$\mu_i = \frac{\sum_{t=1}^{T} y_t \times \gamma_i(t)}{\sum_{t=1}^{T} \gamma_i(t)}\,, \quad i = 1,\ldots,N$$
$$\Sigma_i = \frac{\sum_{t=1}^{T} (y_t - \mu_i)(y_t - \mu_i)' \times \gamma_i(t)}{\sum_{t=1}^{T} \gamma_i(t)}\,, \quad i = 1,\ldots,N$$

- Set $\lambda^* = \lambda$

After convergence is reached, the final parameter set λ is the maximum likelihood estimate according to Eq.6. Moreover, the maximum a posteriori prediction of the categorical sequence $\mathbf{x}$ is obtained from the backward probabilities $\{\gamma_i(1), \gamma_{ij}(2), \ldots, \gamma_{ij}(T)\}_{i,j=1,\ldots,N}$ through the Viterbi algorithm (Viterbi 1967).