Geothermics 87 (2020) 101854

Contents lists available at ScienceDirect

Geothermics

journal homepage: www.elsevier.com/locate/geothermics

Lithofacies classification of a geothermal reservoir in Denmark and its facies-dependent porosity estimation from seismic inversion

Runhai Feng[a,*], Niels Balling[a], Dario Grana[b]

[a] Department of Geoscience, Aarhus University, Høegh-Guldbergs Gade 2, 8000 Aarhus C, Denmark
[b] Department of Geology and Geophysics, University of Wyoming, 1000 E. University Ave., Laramie, USA

ARTICLE INFO

Keywords:
Geothermal reservoir characterization
Markov priors
Artificial Neural Networks
Lithofacies classification
Porosity prediction
Seismic inversion results

ABSTRACT

Characterization of geothermal reservoirs is an important step for exploration and development of geothermal energy, which is reliable and sustainable for the future. Based on the inversion results of seismic reflection data, lithofacies and porosity are predicted beyond well locations on a potential geothermal reservoir in the north of Copenhagen, onshore Denmark. To classify the lithofacies, a new system of Artificial Neural Networks-Hidden Markov Models is proposed to consider the complex spatial distribution of rock properties and the intrinsic depositional rules. Artificial Neural Networks can overcome the common Gaussian assumption for the distribution of rock properties. At the same time, the transition matrix in Hidden Markov Models provides the conditional probability for the lithofacies transitions along the vertical direction. After classification, the resulting lithofacies are used to constrain the porosity prediction, in which the Artificial Neural Networks is trained and applied within each type of lithofacies, as a regression process. The novelty of this approach is in the integration of statistics and computer science algorithms that allows capturing hidden and complex relations in the data that cannot be explained by traditionally deterministic geophysical equations. This workflow could also improve the prediction accuracy and the uncertainty quantification of the porosity distribution given rock properties.

1. Introduction

As extracted from the earth's heat, the geothermal energy is a reliable and sustainable form of energy, which is an attractive substitute of fossil fuels to heat buildings with less CO_2 emissions (Eidesgaard et al., 2019). Opposed to other low-carbon options like solar or wind powers, this energy is a non-intermittent alternative, i.e. it is fully accessible at all weathers (Zwaan and Longa, 2019).

A geothermal fluid exploration relies on various geophysical methods, such as different inversion schemes applied to seismic, electromagnetic, gravity, and magnotelluric data (Ars et al., 2019), of which the magnetotellurics (MT) is generally more common. MT allows delineating the electrical conductivity structures of the geothermal area and measuring the difference in the resistivity between geothermal-rich rocks and surrounding formations (Maithya and Fujimitsu, 2019). Rosenkjaer et al. (2015) inverted 3D MT datasets from the Krafla and Hengill geothermal area, Iceland, for the characterization of subsurface structures. However, MT methods are applicable in geothermal places with high temperatures, whereas in sedimentary basins with average/low temperatures, seismic approaches are more promising.

Seismic data describe elastic contrasts at the layers' interfaces in the subsurface, and they provide high-resolution images of the deeper structures and are used for the exploration of geothermal reservoirs as well. Seismic amplitudes and travel-times could be interpreted to identify anomalies that might reveal geothermal potentials. Casini et al. (2010) analyzed a 3D seismic survey to identify fractured zones in a deep geothermal reservoir at the Travale test site (Italy). Lüschen et al. (2014) simulated the potential of a hydrothermal reservoir for sustainable usages at Unterhaching, Germany, using seismic data. A high-resolution 3D seismic data was analyzed by Krawczyk et al. (2019) to further develop a geothermal research platform in the Northeast German Basin.

In this paper, we propose to use seismic attributes derived from the elastic inversion of seismic data to characterize a potential geothermal reservoir in Denmark. Acoustic impedance (AI) and the ratio between P-velocity and S-velocity (Vp/Vs) obtained from seismic inversion are rock-physical parameters that are peculiar for different rock types and are highly correlated to porosity (Kumar et al., 2014). The degree of reservoir compartmentalization is reflected by different types of rocks, and the amount of energy resources is calculated based on the value of

* Corresponding author.
E-mail address: r.feng@geo.au.dk (R. Feng).

https://doi.org/10.1016/j.geothermics.2020.101854
Received 11 December 2019; Received in revised form 16 March 2020; Accepted 8 April 2020

reservoir porosity (Keelan, 1982; Bosch et al., 2002). In order to predict these reservoir properties, many methods have been presented. Mukerji et al. (2001) classified reservoir lithologies and pore fluids based on the statistical rock-physical models and seismic information. Under a Bayesian approach, Grana et al. (2017) and Grana (2018) developed a joint inversion of seismic data for the simultaneous estimation of reservoir rocks and fluid properties.

In this characterization study, to account for the vertical dependency of the lithofacies, a first-order stationary Markov chain in Hidden Markov Models (HMM) is introduced (Rabiner, 1989; Elfeki and Dekking, 2001). A Markov model is a statistical approach used to describe a sequence of unknown/hidden occurrences, such as the lithofacies sequence in the subsurface, based on the probability of the occurrence at a certain location conditioned to the occurrence at a previous location. Similar applications can be found in the hydrocarbon exploration (Grana et al., 2017; Feng et al., 2018a). However, instead of adopting a Gaussian assumption in HMM (Lindberg and Grana, 2015), which is usually inadequate to describe complex data distributions, the non-linear relationship between seismically inverted properties (AI and Vp/Vs) and reservoir lithofacies is modelled via the Artificial Neural Networks (ANN). ANN is a subclass of deep-learning methods capable of modelling the link between model variables and data responses as a system inspired by biological neuron functionalities. The algorithm learns the non-linear relations between the parameters of interest and the input data with the help of a number ('depth') of hidden layers.

In particular, an integrated system combining ANN and HMM (ANN-HMM) is proposed to classify reservoir lithofacies with seismic inversion results as inputs. Then, using a regression ANN, reservoir porosity is inverted under the constraint of previously classified lithofacies, to reduce the prediction uncertainty. The novelty of this proposed methodology is in the integration of advanced statistical methods, such as Markov chain models, and deep-learning algorithms, like neural networks, to simulate complex geological sequences where physical models cannot completely describe this information, and data-driven approaches are used to predict the model parameters based on higher-order relations hidden in the data. In traditional approaches, a relation between the model parameters and the measured data must be assumed and calibrated on the logging dataset. However, the geophysical models, including petrophysical relations, are only approximations of the physical responses measured in the subsurface and often cannot capture the behavior of the model properties in some intervals where the predictions deviate from the observed trends. Furthermore, traditional methods require the knowledge of several parameters that cannot be directly measured most of the time. In the proposed approach, the statistical model imposed on the prediction by the Markov chain approach allows preserving the geological realism of the sequential patterns, and the integration of the deep-learning algorithms enables extracting complex relations between the data and the model variables from a limited training dataset. Hence, the proposed approach is data-driven and does not require prior assumptions or the use of initial approximate models.

The content of this paper is organized as follows: first, geothermal resources in Denmark are outlined briefly; then, the new approaches are introduced and described; finally, the characterization of reservoir lithofacies and porosity, based on the inversion results from a seismic survey in the north of Copenhagen, Denmark is presented.

2. Geothermal resources in Denmark

The Danish subsurface is divided into five major structural parts: the North German Basin, the Ringkøbing–Fyn High, the Danish Basin, the Sorgenfrei–Tornquist Zone and the Skagerrak–Kattegat Platform (Fig. 1) (Mathiesen et al., 2010). These structural divisions are a decisive factor on the geothermal exploration in the Danish subsurface (Nielsen, 2003; Nielsen et al., 2004). The Mesozoic succession has been the main geothermal target and the most promising reservoirs occur within the Triassic–Lower Cretaceous succession in the Danish Basin and the North German Basin (Balling et al., 2002; Mathiesen et al., 2010). Four main stratigraphic units with a regional geothermal potential have been identified: the Lower–Upper Triassic Bunter Sandstone and Skagerrak Formation, the Upper Triassic–Lower Jurassic Gassum Formation, the Middle Jurassic Haldager Sand Formation and the Upper Jurassic–Lower Cretaceous Frederikshavn Formation (Nielsen et al., 2004; Vosgerau et al., 2016).

Due to the complex dependence on burial history, lithofacies, and mineralogical composition, the depth trends of porosity and permeability, and consequently the quantification of storage and flow abilities of geothermal waters, are often uncertain (Kristensen et al., 2016; Weibel et al., 2017), which may reduce the predictive strength of the current geological models for identifying the areas of interest. In order to provide essential information for reservoir indications and mitigate prediction uncertainties, advanced methods to estimate reservoir lithofacies and porosity based on seismic inversion results are necessary.

Combining statistical models and deep-learning methods, this study is focused on a geothermal exploration project at the Hillerød area, north of Copenhagen, which is close to the Margretheholm plant (Fig. 2). The Hillerød area is located in the eastern part of the Danish Basin, with the Gassum Formation as the primary reservoir target (Nielsen, 2003), and the reservoir rocks are predominantly moderately sorted, fine- to medium-grained sandstone. This formation is already being exploited in the geothermal district heating plants at Thisted and Sønderborg (Røgen et al., 2015) (Figs. 1 and 2). Overlain on the Gassum Formation, the sandstone-dominated Lower Jurassic Reservoir Unit (LJRU) is the second reservoir target (Bredesen et al., 2020; Feng, 2020). As shown by samples and logs from local wells, the LJRU contains relatively homogeneous reservoir rocks, whereas the Gassum Formation is more heterogeneous with interbedded thin shales. The thermal waters are characterized by low-to-average temperatures (40–90 °C) within this area (Fuchs et al., 2020). Therefore, seismic data are more suitable than MT data, of which the latter one is applicable to reservoirs of high-temperature thermal waters (> 100 °C).

3. Methodology

The Methodology includes two main parts: the lithofacies classification and the porosity prediction.

3.1. ANN-HMM for lithofacies classification

Artificial Neural Networks (ANN) is one of the most popular deep-learning tools for classification problems. ANN is represented by hidden layers and neurons, with activation functions embedded (Saggaf et al., 2003), and is used to describe the highly complex relationship between input features and output targets. The relationship is honored independently at each spatial location, which means that the algorithm does not account for the spatial correlations within the data. On the other hand, Hidden Markov Models (HMM) is a geostatistical method commonly applied to mimic the distribution of spatial patterns. HMM can decode the natural ordering of spatially distributed random variables based on indirect observations (Rabiner, 1989). Typically, in HMM, a Gaussian assumption is made to describe the relationship between the categorical latent variables and the observations, which may be insufficient.

In order to address these problems in ANN and HMM, here, we propose a new system: Artificial Neural Networks-Hidden Markov Models (ANN-HMM) for the classification of categorical hidden states (lithofacies, in our application) based on observations represented by continuous variables (elastic attributes, in our case), as shown in Fig. 3.

There are three parameters in ANN-HMM:

$$\boldsymbol{\lambda} = \{A, B, C\} \tag{1}$$

The transition matrix A contains the conditional probabilities

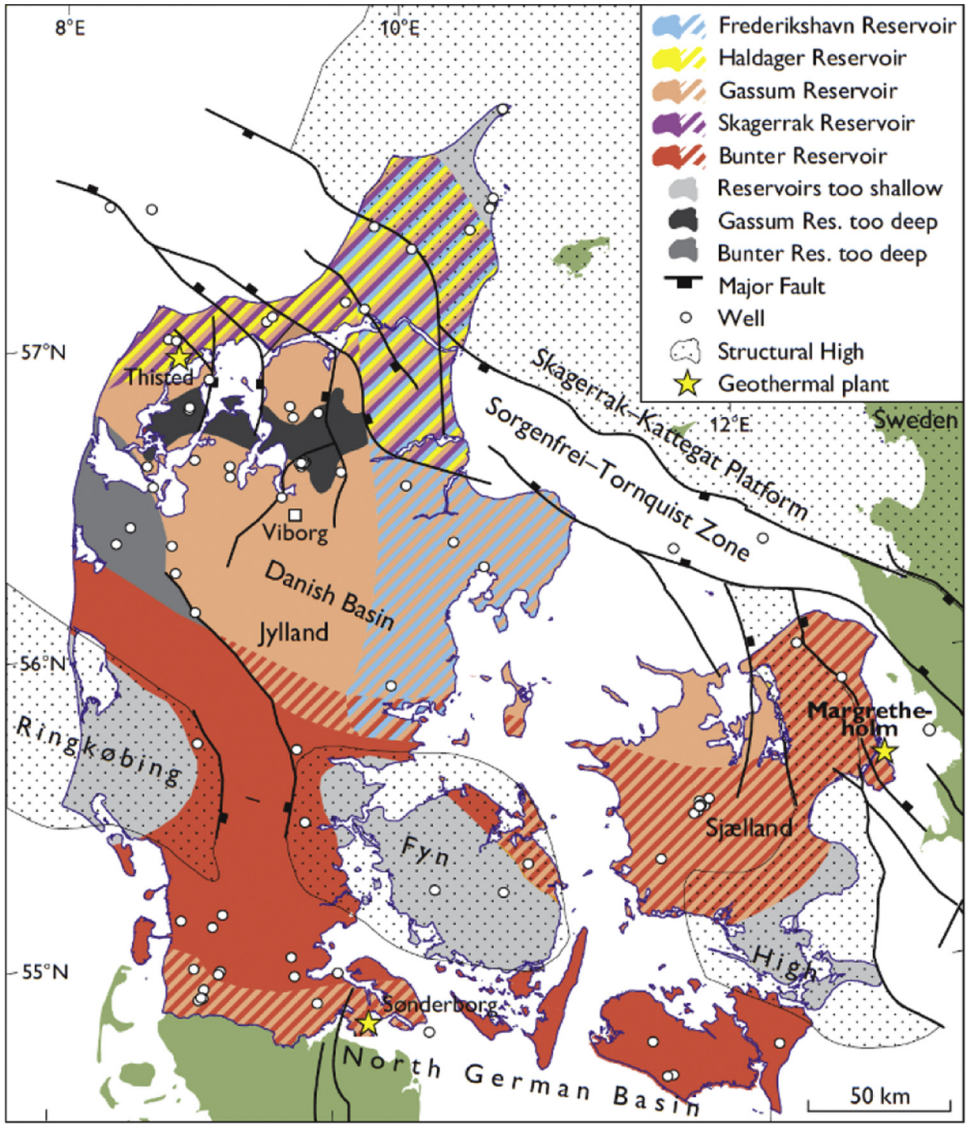

Fig. 1. Distribution of potential geothermal reservoirs in Denmark (modified from Mathiesen et al., 2010).

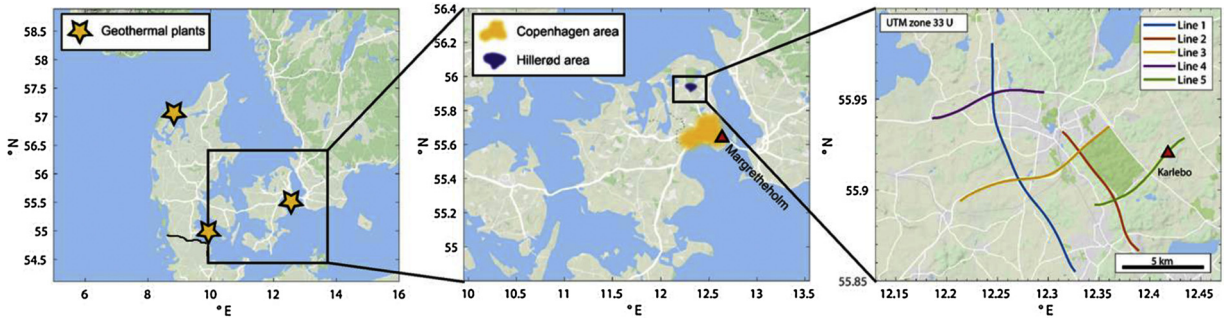

Fig. 2. Map of geothermal plants and location of the study area for the geothermal reservoir characterization (Bredesen et al., 2020).

between different states or lithofacies (**S**) at adjacent spatial locations along the vertical direction and is generally obtained from a scanning of known lithofacies at well locations:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & a_{ij} & \vdots \\ a_{m1} & \cdots & a_{mm} \end{bmatrix} \tag{2}$$

where a_{ij} specifies the transitional probability from state S_i to state S_j ($\mathbf{S} = \{S_1, S_2, \cdots, S_m\}$, m is the total number of categorical lithofacies).

The stationary probability (i.e. the global proportion) of the lithofacies — B is derived by taking the limit of transition steps (f) in matrix A (Eq. (2)) to an infinite number (Elfeki and Dekking, 2001):

$$\lim_{f\to\infty} a_{ij}^{(f)} = b_j,\ j = 1,\ \cdots, m \tag{3}$$

in which, i disappears because the stationary probability does not depend on the starting states, and b_j is one element in B ($B = \{b_j\}$).

The emission or likelihood probability C ($\Pr(\mathbf{Y}|\mathbf{S})$) is the connection

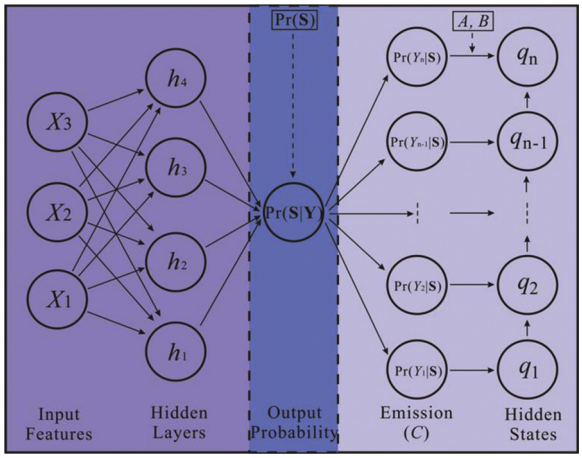

Fig. 3. Proposed method for lithofacies classification: Artificial Neural Networks-Hidden Markov Models (ANN-HMM). X_i is the i^{th} input feature such as AI or Vp/Vs; h_i is the i^{th} neuron in hidden layers; Pr(**S**|**Y**) tells the output probability of lithofacies or states (**S**) given observations (**Y**) (**Y** = $\{X_i\}$, a combination of the input features), without any Markov prior information; q_d represents the depth point (d) at which its state needs to be decoded; Pr(**S**) is the relative frequencies of states (**S**) in the training dataset; A and B are the priors; C stands for the emission (likelihood) function which describes the distribution of **Y** given **S**.

between observations **Y** (elastic attributes) and states **S** (lithofacies), and can be estimated from logging data. In HMM, means and covariances in Gaussian functions are used to provide this value, which are to be replaced by weights and biases of hidden neurons from ANN (the left part in Fig. 3). In particular, the multilayer perceptron (MLP), a specific configuration of ANN, is used to describe this relationship.

The system must be trained using direct measurements of the properties of interest. Input features (X_i) such as AI and Vp/Vs, are fed into the neural architecture (left part in Fig. 3). Based on the target variables (lithofacies), weights and biases in the hidden neurons (h_i) of ANN are updated to minimize a pre-defined error function.

In the classification process, the so-calibrated algorithm can then calculate the output probability Pr(**S**|**Y**) of ANN using a *softmax* function (Goodfellow et al., 2016), with inputs of logging data or seismic inversion. Without prior information (A and B in Eq. (1)), the classified result from Pr(**S**|**Y**) by selecting lithofacies of the highest probability value is a degenerated case of ANN-HMM. Then, according to the Bayes' theory (Scales and Snieder, 1997; Ulrych et al., 2001; Calvo-Zaragoza et al., 2019), the emission probability C (Pr(**Y**|**S**)) can be obtained:

$$C{=}\Pr(\mathbf{Y}|\mathbf{S}) = \frac{\Pr(\mathbf{S}|\mathbf{Y})\Pr(\mathbf{Y})}{\Pr(\mathbf{S})} \propto \frac{\Pr(\mathbf{S}|\mathbf{Y})}{\Pr(\mathbf{S})} \tag{4}$$

where Pr(**S**|**Y**) is the output of ANN, as calculated by the *softmax* equation; Pr(**Y**) is a constant from independent observations; Pr(**S**) is assumed to be equal by the relative frequencies of lithofacies in the training dataset, which is similar to the global proportion of lithofacies (B) (Eq. (3)). However, in order to keep things analytically tractable, a distinction is made for them.

After a complete definition of the model **λ**, the Viterbi algorithm (Viterbi, 1967) is applied for the decoding task of hidden lithofacies. Making using of a dynamic programming trellis along the vertical path, the most probable sequence of states $Q = q_1 q_2 \cdots q_n$ ($q_d = S_i$, $1 \le d \le n$, $1 \le i \le m$, n is the total number of data points and m is total number of lithofacies), is able to be found from a backtracing step. For implementation details of the Viterbi method, please refer to Rabiner (1989).

3.2. Lithofacies-dependent ANN for porosity prediction

In the second step of the proposed characterization process, we predict the special distribution of porosity conditioned by elastic attributes and lithofacies, with the goal to quantify the volume of geothermal waters and the link with rock permeability (Fang and Yang, 2015). The relationship between porosity and elastic properties is generally facies-dependent and non-linear, which makes the solution of the inverse problem non-unique. In general, the uncertainty in the porosity prediction is relatively large when only elastic properties are considered. However, due to the intrinsic petrophysical characteristics of lithofacies, a less uncertain porosity estimation can be achieved by introducing the facies-dependent relations.

In this present approach, we propose to predict reservoir porosity based on the inverted seismic attributes, with a constraint imposed by the classified lithofacies from ANN-HMM. The porosity inversion is obtained by applying ANN with the *sigmoid* activation function adopted to extract the complex relationship between input rock properties (AI and Vp/Vs) and target porosities in a regression process, within each lithofacies.

4. Real case study

4.1. Lithofacies classification

The proposed method is applied to a real dataset from an onshore potential geothermal reservoir, north of Copenhagen, Denmark (Hillerød area in Fig. 2). Five 2D seismic lines have been acquired across the area and a borehole, namely the Karlebo well, is drilled in the vicinity of line No. 5 (Fig. 2). The Margretheholm well (part of the Margretheholm geothermal plant) is approximately 40 km away from the Karlebo well and presents similar depositional environments and reservoir features (Vosgerau et al., 2016, 2017). A preliminary seismic inversion was performed to predict elastic attributes of line No. 5, and two rock properties (AI and Vp/Vs) are estimated.

At the Margretheholm and Karlebo well locations, rock properties are calculated from sonic and density logs, and lithofacies profiles are available (Fig. 4a and b). Shale, Shaly Sand, and Sand have been identified based on an interpretation of the Gamma Ray log, sedimentological and depositional rules. The lithofacies-dependent bivariate joint distribution of elastic properties is displayed in Fig. 4c. Sand and Shaly Sand show skewed marginal distributions, whereas Shale has a bimodal behavior probably due to different mineral compositions at the top and bottom of the interval (Fig. 4a and b). Sand and Shaly Sand also have a large overlap because of similar rock properties.

The proposed ANN-HMM is trained using AI, Vp/Vs and reference lithofacies from the two wells. Both datasets are used simultaneously in the workflow to represent the variability of the data and increase the amount of information available to extract the non-linear relationship between model variables and measured data.

Then, to validate the method, AI and Vp/Vs log data of the Margretheholm well are used for the classification of lithofacies at the well location. The classified lithofacies by the Viterbi decoding in ANN-HMM are shown in Fig. 5. Results obtained by applying ANN and HMM independently are also displayed. In ANN, lithofacies of the highest value in the *softmax* function are assigned along the vertical profile. The same Viterbi method is used in HMM as well, where C (the likelihood) is calculated by the Gaussian function that is replaced by trained weights and biases in ANN-HMM. Note that the classification in ANN-HMM and HMM is performed from the bottom to the top, as illustrated by vertical arrows in q_i (Fig. 3), in order to fully honor a depositional process. The Matthews correlation coefficient (*MCC*) is adopted to quantify the performance of different classifiers (Matthews, 1975). It is computed based on a multiclass confusion matrix and the value is between -1 and 1, where -1 represents an inverse outcome between truth and prediction, and 1 means a perfect classification. The *MCC* value for

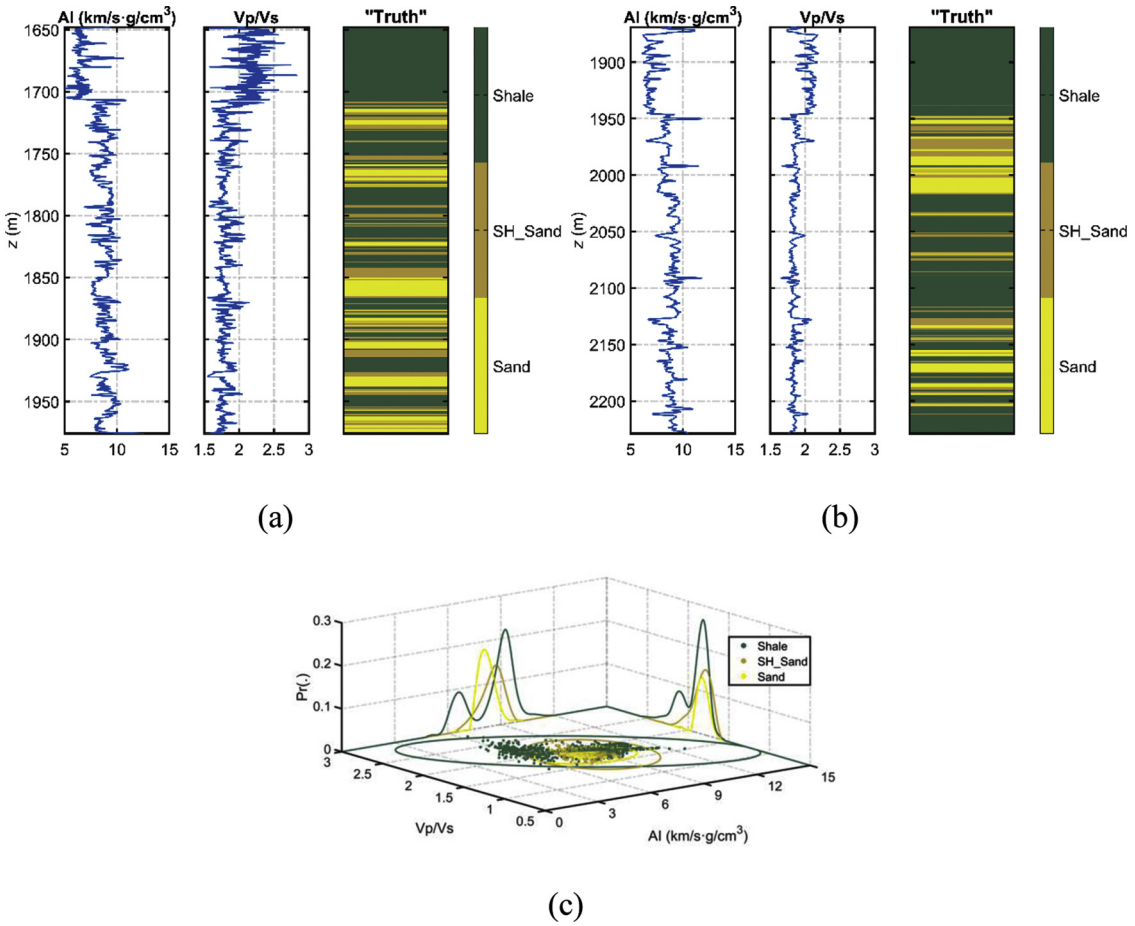

Fig. 4. Elastic properties (AI and Vp/Vs) and reference lithofacies from geological interpretations ("Truth") at the (a) Margretheholm and (b) Karlebo well. (c) Lithofacies-dependent bivariate joint distribution of elastic properties (Shale in green, Shaly Sand (SH_Sand) in brown, and Sand in yellow). On the horizontal plane, contour plots of the approximate Gaussian distributions are shown. On the vertical planes, the smoothed univariate marginal distributions are displayed per each lithofacies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ANN is 0.5954; HMM: 0.5399; ANN-HMM: 0.6261. Therefore, the result obtained by ANN-HMM is more accurate than the ones by the other two traditional approaches. As a probability curve, ROC (Receiver Operating Characteristics) could also be utilized to measure the classifiers' performances, at various thresholds. However, the Viterbi path in ANN-HMM and HMM does not estimate a posterior probability, which makes the ROC analysis inapplicable here.

We then apply the method for lithofacies classification to the Karlebo location and make a comparison with other traditional approaches as well (Fig. 6). In HMM, wrong classifications of Shale as Shaly Sand can be observed from 2020 m to 2110 m. Similarly, between 1950 m and 2000 m the algorithm over-predicts the Sand proportion. This is probably due to the non-stationarity of the reference data that cannot be accounted for by a first-order Markov chain model, as well as

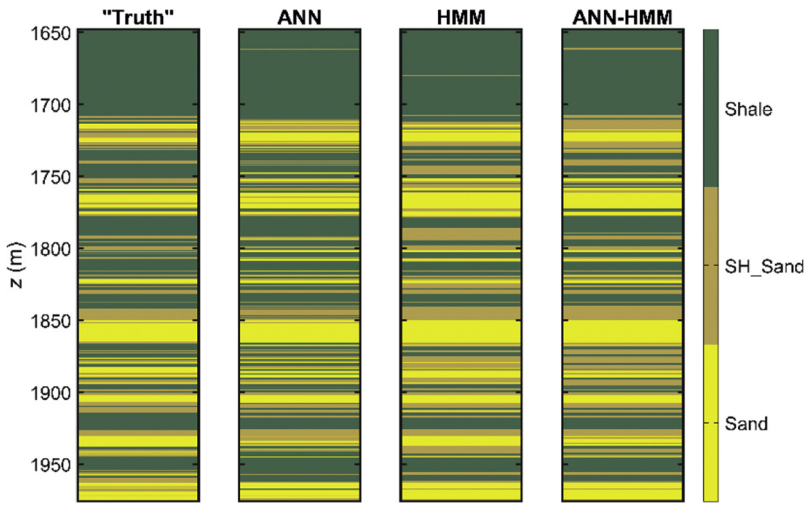

Fig. 5. Classification results of ANN, HMM and ANN-HMM applied to elastic properties at the Margretheholm well in Fig. 4a.

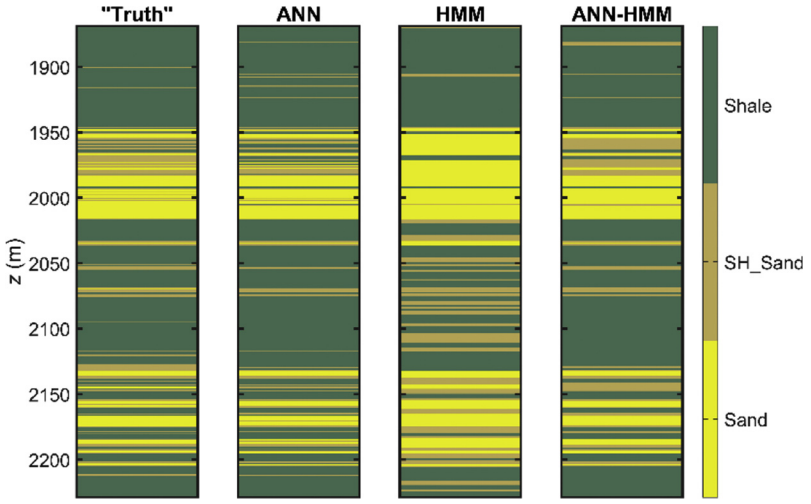

Fig. 6. Classification results of ANN, HMM and ANN-HMM applied to elastic properties at the Karlebo well in Fig. 4b.

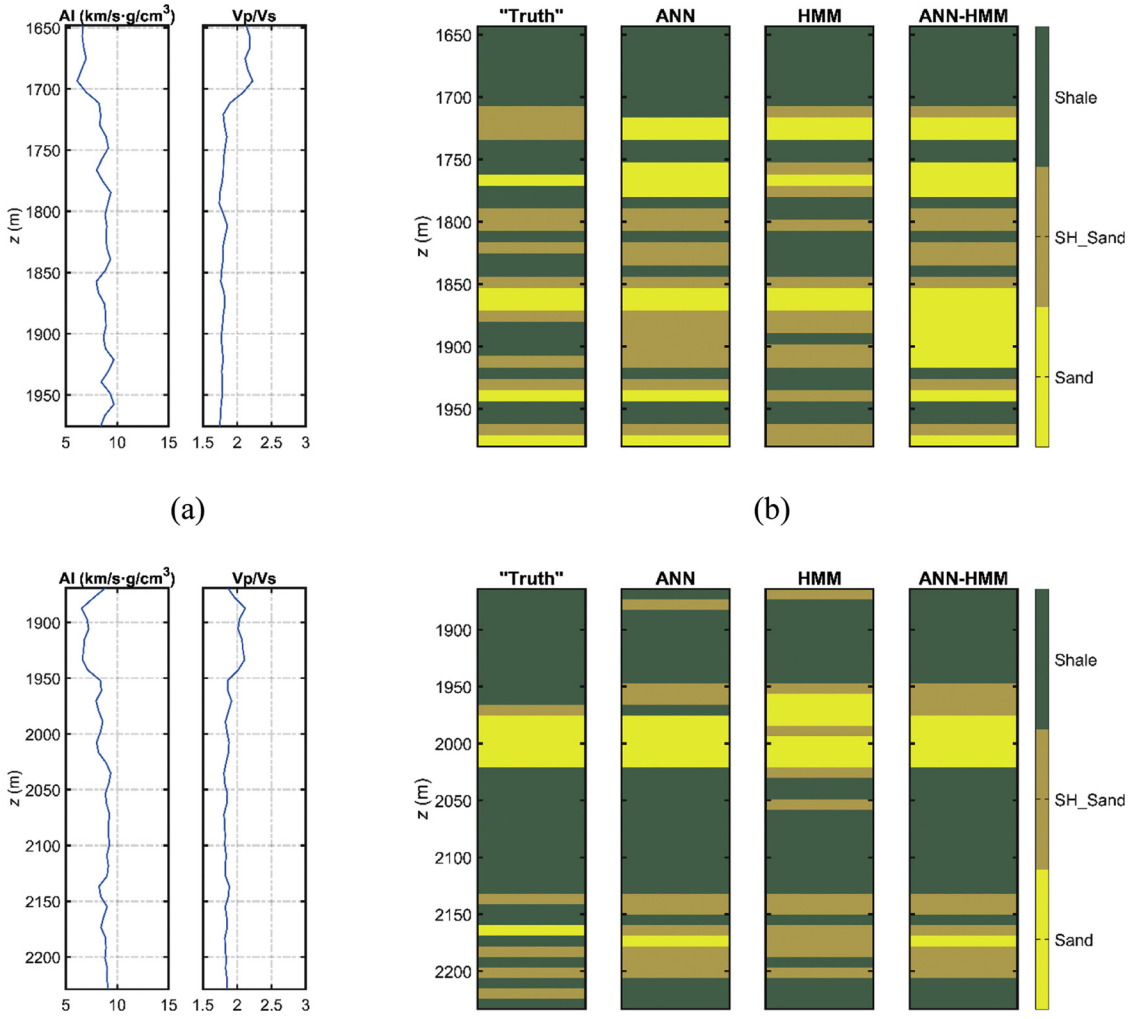

Fig. 7. (a) Upscaled logs and (b) predicted lithofacies at the Margretheholm well. (c) Upscaled logs and (d) predicted lithofacies at the Karlebo well.

the insufficient Gaussian assumption. In ANN, the result is overall more accurate than that by HMM, despite some misclassifications of thin interbedded layers in the two Sand reservoir units. The combined ANN-HMM further improves the classification result. The *MCC* value for ANN is 0.7798; HMM: 0.4227; and ANN-HMM: 0.7954.

To test the validity of the method at the seismic scale, we also apply the classification to the upscaled well logs of Margretheholm and Karlebo at the seismic resolution that is 1/4 of wavelength. A sequential Backus averaging is used to upscale the log data and the most frequent lithofacies in a running window are chosen along the vertical direction (Lindsay and Koughnet, 2001; Avseth et al., 2005). Fig. 7 shows the upscaled logs (AI and Vp/Vs), the upscaled reference profiles ("Truth")

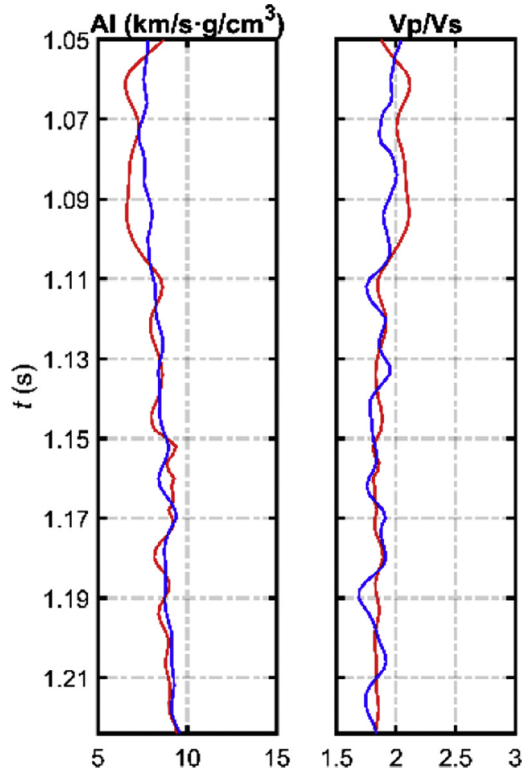

Fig. 8. Inverted rock properties at the Karlebo well. Blue lines represent the inverted results, and red lines are the upscaled logs. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and the classified lithofacies obtained by ANN, HMM and ANN-HMM, from which it can be seen that ANN-HMM achieves a better performance at the seismic scale as well, compared with ANN and HMM. At the Margretheholm well, the *MCC* value for ANN is 0.5388; HMM: 0.4868; ANN-HMM: 0.5436, and it is 0.5608 for ANN; 0.5033 for HMM; 0.6642 for ANN-HMM at the Karlebo well.

Finally, the lithofacies classification is applied to the inverted elastic properties (AI and Vp/Vs) from seismic data, which are independent from the training datasets before. An AVO (amplitude versus offset) inversion scheme is employed on offset gathers (Mallick and Adhikari, 2015) and it is based on the Aki-Richards linear approximation of Zoeppritz equations (Aki and Richards, 2002) where the reflection coefficients are expressed as a function of incident angles. AI is related to the intercept at the near-zero offset in the AVO curve, and Vp/Vs is a function of the gradient which requires intermediate-far offsets. The inverted elastic properties of AI and Vp/Vs at the Karlebo well are displayed in Fig. 8.

Based on the inverted elastic properties in Fig. 8, the proposed ANN-HMM is applied and the classified lithofacies are shown in Fig. 9, together with the results by ANN and HMM as well. It can be observed that ANN-HMM performs better than ANN, and among the three, the result by HMM is the worst (ANN: 0.3610; HMM: 0.1871; ANN-HMM: 0.4399, for *MCC* values). It should be pointed out that the *MCC* value represents a quality index for the classification performance in the entire interval. Due to the mismatch between the true and inverted rock properties (Fig. 8), these values are relatively low. In Fig. 9, the proportions of Sand and Shaly Sand are slightly over-estimated by ANN-HMM, compared with the ones by ANN; however, its *MCC* value is still higher than the corresponding index for ANN.

Confusion matrices showing the success and failure rates of classifications in each lithofacies are presented in Fig. 10. Along the diagonal space, the correctly classified data samples and their corresponding percentages are shown, in which a perfect classification should have a score of 100 %. Classified samples are considered as suboptimal, if they are close to the diagonal, such as Shaly Sand predicted as Sand. The off-diagonal elements represent the worst misclassifications and are expected to have a value of 0 %. ANN and ANN-HMM have a similar performance for the classification of Sand. Compared with ANN, more Shale have been predicted as Shaly Sand in ANN-HMM, which are mediocre. For Shaly Sand, ANN-HMM can achieve a higher success rate than ANN. The classification ratio by HMM is the worst because of the distributed data samples across the truth and the prediction categories.

From the AVO inversion of cross-section seismic data (Fig. 11a), inverted AI and Vp/Vs (Fig. 11b) are obtained and then used as inputs for the classification of lithofacies (Fig. 11c). The result of ANN-HMM is recovering the major characteristics of the subsurface, such as the boundary of Top Gassum between Sand and Shaly Sand, and the complex geological structures are honored.

4.2. Porosity prediction

The porosity distribution based on rock properties suffers the problem of non-uniqueness, and there is a large overlap between the high-porosity reservoir unit and the low-porosity non-reservoir unit (Fig. 12a). In this proposed method for porosity prediction, the above classified lithofacies are included as a constraint to improve the prediction accuracy and reduce the uncertainty (Fig. 12b).

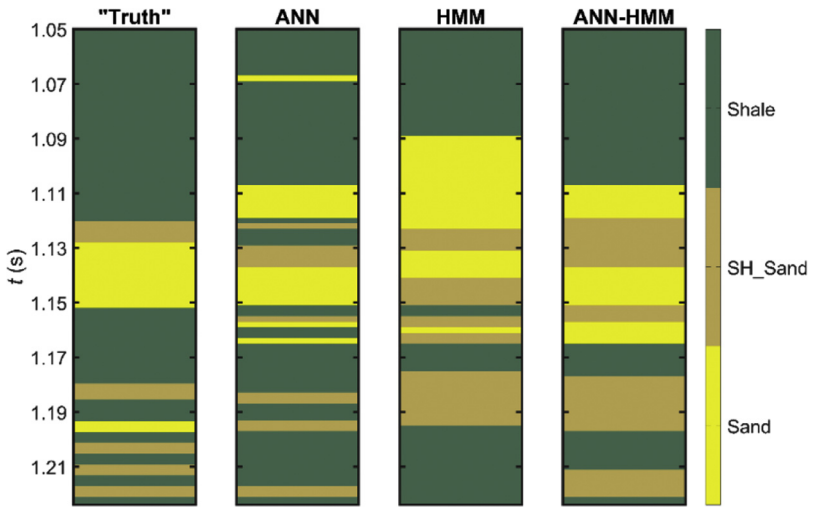

Fig. 9. Classification results by ANN, HMM and ANN-HMM applied to seismically inverted elastic properties (Fig. 8) at the Karlebo well.

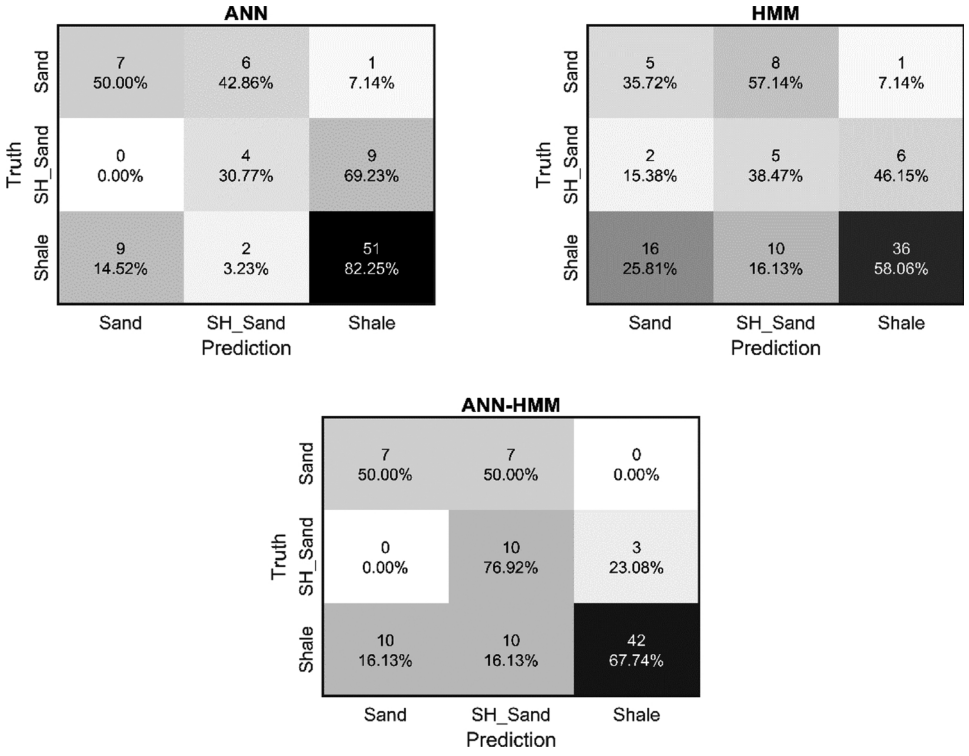

Fig. 10. Confusion matrices of ANN, HMM and ANN-HMM. Black background color represents a higher classification score.

Instead of using typical rock-physical models (Avseth et al., 2010), ANN is applied with the *sigmoid* function embedded to derive the unknown yet complex relationship between elastic properties (AI and Vp/Vs) and reservoir porosity, in a fully data-driven way. Three different networks for Shale, Shaly Sand and Sand are designed and independently trained using lithofacies-dependent elastic properties from the wells. Then, these trained ANNs are used for the prediction of reservoir porosity with previously classified lithofacies as additional constraints, based on the elastic inversion results. In Fig. 13, the prediction quality of porosity is improved after including the constraint of lithofacies (plots b–d), compared with the porosity prediction independent from the lithofacies classification (plot a). The improved predictions are also shown by an increase of correlation coefficient between the prediction and the reference porosity, of which the predicted porosity constrained with ANN-HMM classified lithofacies is the best (Fig. 13d).

The predicted porosity along the seismic cross section is shown in Fig. 14. With constraints of the lithofacies by ANN-HMM (Fig. 11c), the result has more geologically realistic continuity inside the Lower Jurassic Reservoir Unit (LJRU) Formation, which makes that overall the LJRU Formation has a better reservoir quality than the underlied Gassum Formation. For Fjerritslev and Vinding Formations, they are associated with Shale and Shaly Sand, which are the confining rock units (Fig. 11c).

5. Discussion

In this study, the reservoir lithofacies is classified using a new system combining ANN and HMM. Subsequently, the porosity is predicted by a totally data-driven ANN, with a constraint of previously classified lithofacies.

Belonging to the deep-learning methods, ANN is a non-parametric approach that can overcome the Gaussian assumption of many statistical algorithms, leading to a more accurate description of the complex distribution of reservoir properties. To impose the spatial continuity model observed in the reference data, ANN is integrated with HMM using the Viterbi algorithm to classify spatially correlated lithofacies. Then, with a constraint of lithofacies, uncertainty in the prediction of porosity could be largely reduced, and complicated data distributions are modeled by ANN. Hence, ANN has been applied for the classification (discrete lithofacies) and regression (continuous porosity) problems in a two-step procedure, in which different activation functions have been selected, such as *softmax* and *sigmoid* (Goodfellow et al., 2016). Other types of neural networks such as Convolutional Neural Networks or Recurrent Neural Networks could also be used (Grana et al., 2020).

The new ANN-HMM integrates the information from the data with the knowledge from the transition matrix, which has the function to preserve the geological realism of the classification by imposing the estimated transition probability. In complex geological environments, with alternating sequences of thin and thick layers, some of the thin layers (below the data resolution) might be miss-predicted due to constraints imposed by the transition matrix of the Markov chain in the facies classification. A potential solution would be to divide the intervals of interest into multiple layers and apply the method with different transition matrices. When associated emission probabilities of the lithofacies in thin layers are high enough, these thin layers can be correctly predicted, such as the ones in Fig. 6 (ANN-HMM, around 1900 m).

To train the neural part in ANN-HMM, well log data have been randomly split into training (80 %), development (dev) (10 %) and test (10 %) subsets. ANN shares the same parameters with ANN-HMM in terms of neural weights and biases. HMM accounts for prior information (A and B), while a Gaussian assumption is made to estimate means and covariances for C (Eq. (1)), instead of hidden neurons assigned in ANN-HMM. For ANN used in the porosity prediction, the same ratio between training, dev and test (80-10-10 %) datasets is adopted for rock properties within each lithofacies (three neural networks corresponding to three lithofacies (Sand, Shaly Sand and Shale), in this case). Then, these subsets are combined individually to train the fourth ANN without any discrimination of lithofacies types, as applied in Fig. 13a.

In the proposed workflow, an accurate lithofacies classification is

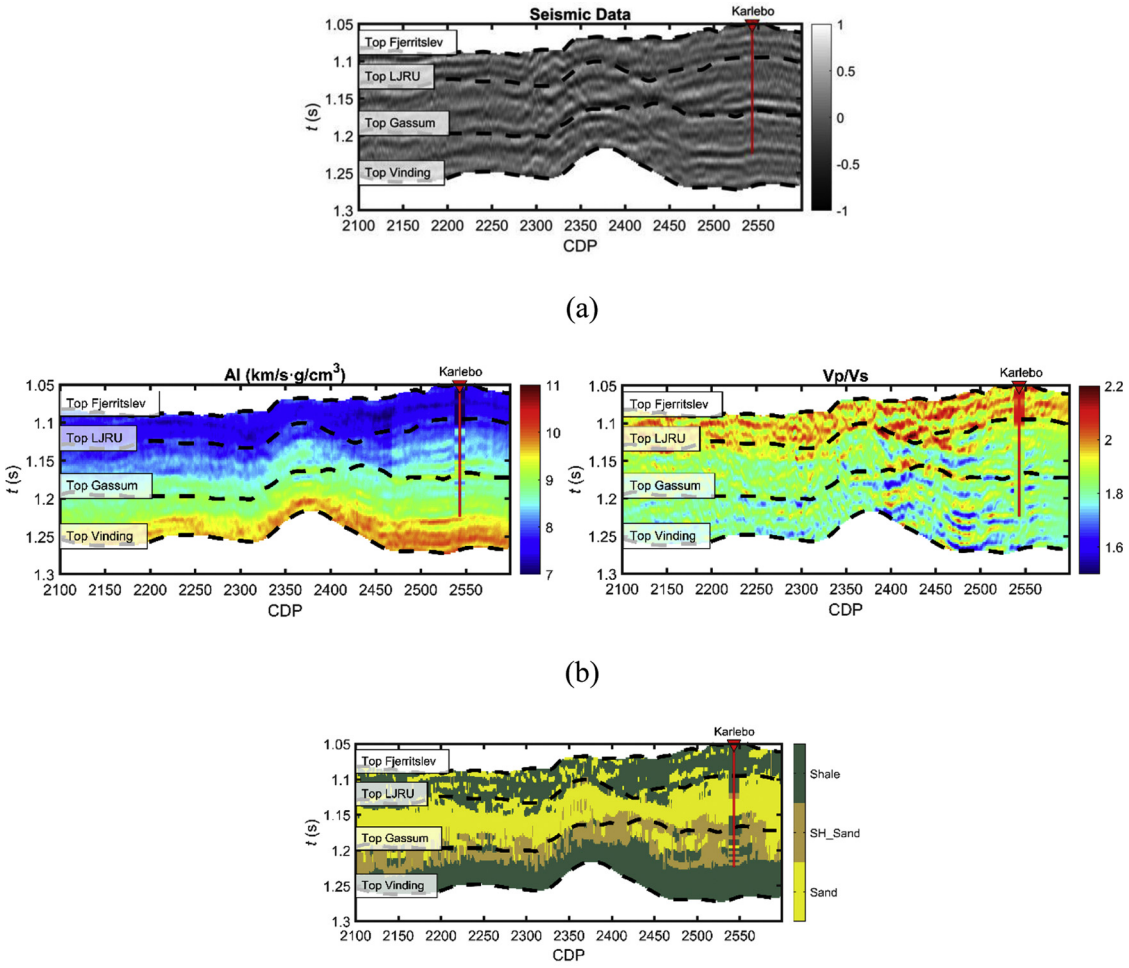

Fig. 11. (a) Cross section of seismic data along survey line No. 5 (Fig. 2); (b) Elastic inversion results in terms of AI and Vp/Vs; (c) Classified lithofacies by applying ANN-HMM trace-by-trace. CDP means common-depth point. True rock properties and reference lithofacies at the Karlebo site are superposed on the inverted (b) and classified (c) results (red lines). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

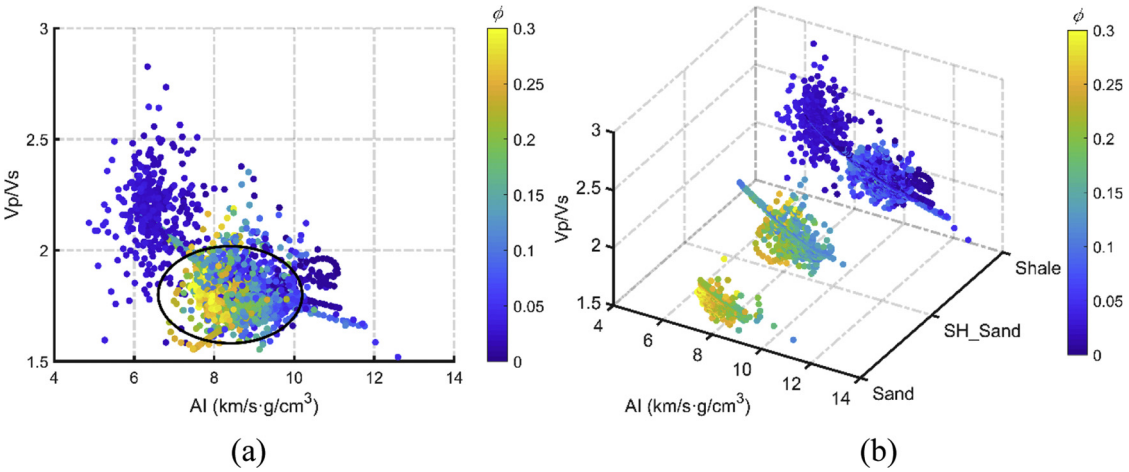

Fig. 12. Crossplots between rock properties (AI and Vp/Vs) and the color is coded by porosity. (a) The back region shows the high-porosity reservoir unit that overlaps with the low-porosity non-reservoir unit; (b) Lithofacies-dependent crossplots to increase the porosity discrimination.

essential for a successful prediction of the reservoir porosity, since lithofacies are used as a constraint in the porosity prediction. Large overlaps between Sand and Shaly Sand are observed (Fig. 4c), leading to a difficult discrimination in the presented application. Thus, in complex geological scenarios with large facies overlaps, additional geophysical measurements such as large-offset seismic data or controlled-source electromagnetic data might be necessary such that other rock properties, like density or resistivity can be included in the

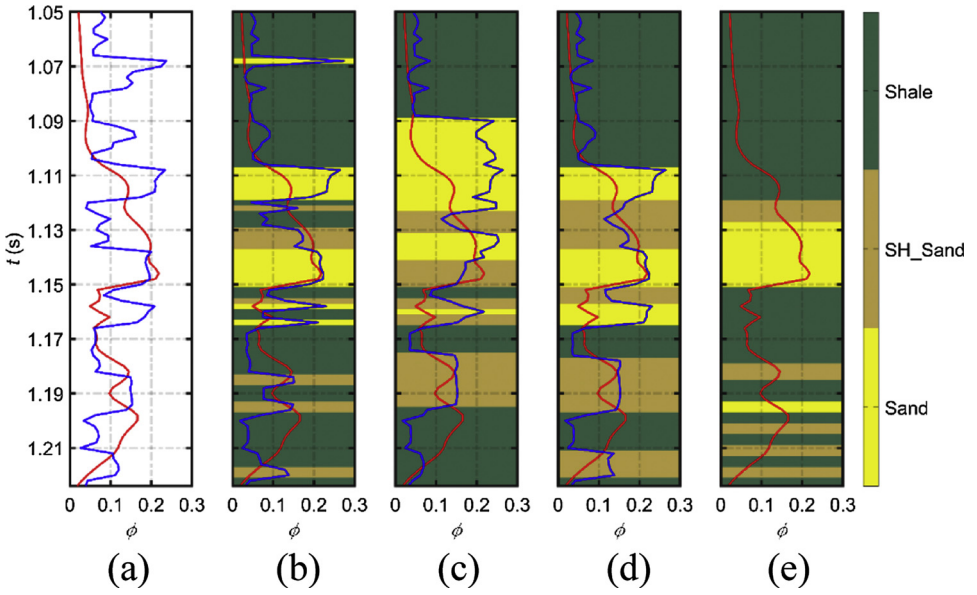

Fig. 13. Predicted porosities based on elastic inversion results at the Karlebo well (Fig. 8). The first plot (a) shows the prediction result without the use of lithofacies. The prediction results with the constraint imposed by classified lithofacies from ANN, HMM, ANN-HMM (Fig. 9) are displayed in plots from (b) to (d). The upscaled reference lithofacies and porosity are shown in (e). Red line represents the upscaled porosity (based on the density log) and blue line is the prediction. The correlation coefficient between predicted and upscaled porosities is: 0.2234, 0.5095, 0.4320 and 0.6180 in plots from (a) to (d). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

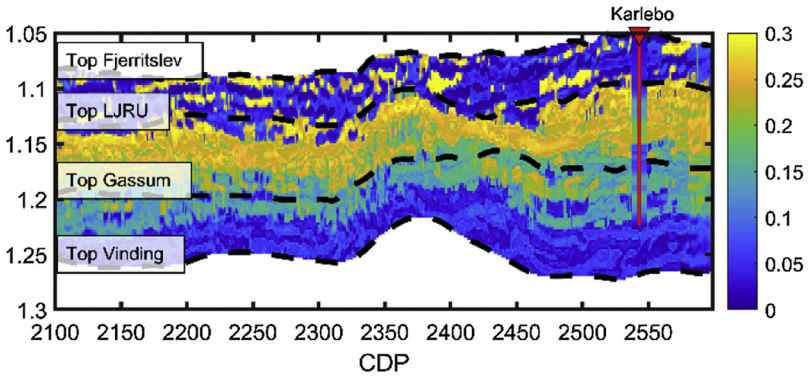

Fig. 14. Predicted porosity along the seismic line No. 5 (Fig. 2), with lithofacies constraints from ANN-HMM. Porosity at the Karlebo well is superposed on the cross section, as represented by the red line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

inversion algorithm.

To make results comparable to the seismic data (Fig. 11a), classified lithofacies (Figs. 9 and 11c) and predicted porosity (Figs. 13 and 14) based on inversion results (Figs. 8 and 11b) are in the time domain. The depth-to-time conversion was performed using a smooth velocity model. Well log and seismic data were also pre-processed to remove the systematic noises. Angle-dependent wavelets were statistically extracted based on seismic data for the spectral amplitude, whereas the phase was estimated from the well log reflectivities (Edgar and van der Baan, 2011). All these processing steps include approximations and might be uncertain.

Future research could include the horizontal correlations between lithofacies as shown in Middleton (1973) and Feng et al. (2018b). Furthermore, the estimation of permeability would be incorporated as well. As another important reservoir parameter, the permeability indicates pore connections in the subsurface. A significant amount of core data analysis on the relationship between the rock porosity and the hydraulic permeability is available for most of the Danish potential geothermal reservoirs (Kristensen et al., 2016; Weibel et al., 2017). Despite complex relationships, the porosity could be considered as a good hint of the permeability, and together with the reservoir thickness, information on the geothermal production potential is able to be inferred. For the present case study, our results clearly indicate the existence of sandstone units with a sufficient porosity (about 20 % or more) and the significant lateral extent for high-quality geothermal reservoir conditions.

6. Conclusion

A new procedure is proposed to classify the lithofacies based on seismic inversion results for a geothermal reservoir study in Denmark. Combining Artificial Neural Networks and Hidden Markov Models (ANN-HMM), the method overcomes the common assumption of Gaussian distributions in elastic properties and imposes a spatial correlation model on the vertical sequence of lithofacies using Markov transition probabilities. Next, the reservoir porosity is predicted using Artificial Neural Networks based on seismically inverted elastic attributes, with an additional constraint provided by classified lithofacies from ANN-HMM. This sequentially integrated approach reduces the uncertainty in the prediction of reservoir properties and improves the accuracy of the reservoir characterization. Using inversion results from seismic data, the characterization process could be performed away from well locations. As a result, potential geothermal reservoir units are identified in 2D sections.

CRediT authorship contribution statement

Runhai Feng: Conceptualization, Methodology, Visualization, Formal analysis, Writing - review & editing. **Niels Balling:** Writing - review & editing, Supervision, Project administration. **Dario Grana:** Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was performed within the framework of project GEOTHERM – Geothermal energy from sedimentary reservoirs – Removing obstacles for large scale utilization (project #6154-00011B) funded by Innovation Fund Denmark. The authors would like to thank Qeye and Esben Dalgaard for providing the inversion results of seismic data. Processing of well data was done by Kenneth Bredesen from GEUS (Geological Survey of Denmark and Greenland).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.geothermics.2020.101854.

References

Aki, K., Richards, P.G., 2002. Quantitative Seismology. University Science Books.
Ars, J.M., Tarits, R., Hautot, S., Bellanger, M., Coutant, O., Maia, M., 2019. Joint inversion of gravity and surface wave data constrained by magnetotelluric: application to deep geothermal exploration of crustal fault zone in felsic basement. Geothermics 80, 56–68.
Avseth, P., Mukerji, T., Mavko, G., 2005. Quantitative Seismic Interpretation: Applying Rock Physics Tools to Reduce Interpretation Risk. Cambridge University Press.
Avseth, P., Mukerji, T., Mavko, G., Dvorkin, J., 2010. Rock-physics diagnostics of depositional texture, diagenetic alterations, and reservoir heterogeneity in high-porosity siliciclastic sediments and rocks — a review of selected models and suggested work flows. Geophysics 75 (5), A31–A47.
Balling, N., Hvid, J.M., Mahler, A., Møller, J.J., Mathiesen, A., Bidstrup, T., Nielsen, L.H., 2002. Denmark. In: Hurter, S., Haenel, R. (Eds.), Atlas of Geothermal Resources in Europe. Publication No. EUR 17811. The European Commission, pp. 15–16 27–28 and plates.
Bosch, M., Zamora, M., Utama, W., 2002. Lithology discrimination from physical rock properties. Geophysics 67 (2), P573–P581.
Bredesen, K., Dalgaard, E., Mathiesen, A., Rasmussen, R., Balling, N., 2020. Seismic characterization of geothermal sedimentary reservoirs: a field example from the Copenhagen area, Denmark. Interpretation 8 (2), T275–T291. https://doi.org/10.1190/int-2019-0184.1.
Calvo-Zaragoza, J., Toselli, A.H., Vidal, E., 2019. Hybrid hidden Markov models and artificial neural networks for handwritten music recognition in mensural notation. Pattern Anal. Appl. 22, 1573–1584.
Casini, M., Ciuffi, S., Fiordelisi, A., Mazzotti, A., Stucchi, E., 2010. Results of a 3D seismic survey at the Travale (Italy) test site. Geothermics 39 (1), 4–12.
Edgar, J.A., van der Baan, M., 2011. How reliable is statistical wavelet estimation? Geophysics 76 (4), V59–V68.
Eidesgaard, Ó.R., Schovsbo, N.H., Boldreel, L.O., Ólavsdóttir, J., 2019. Shallow geothermal energy system in fractured basalt: a case study from Kollafjørður, Faroe Islands, NE-Atlantic Ocean. Geothermics 82, 296–314.
Elfeki, A., Dekking, M., 2001. A markov chain model for subsurface characterization: theory and applications. Math. Geol. 33 (5), 569–589.
Fang, Z., Yang, D., 2015. Inversion of reservoir porosity, saturation, and permeability based on a robust hybrid genetic algorithm. Geophysics 80 (5), R265–R280.
Feng, R., 2020. A Bayesian approach in machine learning for lithofacies classification and its uncertainty analysis. IEEE Geosci. Remote Sens. Lett. https://doi.org/10.1109/LGRS.2020.2968356.
Feng, R.H., Luthi, S.M., Gisolf, A., Angerer, E., 2018a. Reservoir lithology classification based on seismic inversion results by Hidden Markov Models: applying prior geological information. Mar. Pet. Geol. 93, 218–229.
Feng, R.H., Luthi, S.M., Gisolf, A., Angerer, E., 2018b. Reservoir lithology determination by hidden Markov random fields based on a Gaussian mixture model. IEEE Trans. Geosci. Remot. Sens. 56 (11), 6663–6673.
Fuchs, S., Balling, N., Mathiesen, A., 2020. Deep basin temperature and heat-flow field in Denmark – new insights from borehole analysis and 3D geothermal modelling. Geothermics 83, 1–18.
Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.
Grana, D., 2018. Joint facies and reservoir properties inversion. Geophysics 83 (3), M15–M24.
Grana, D., Fjeldstad, T., Omre, H., 2017. Bayesian Gaussian mixture linear inversion for geophysical inverse problems. Math. Geosci. 49 (4), 493–515.
Grana, D., Azevedo, L., Liu, M., 2020. A comparison of deep machine learning and Monte Carlo methods for facies classification from seismic data. Geophysics. https://doi.org/10.1190/geo2019-0405.1.
Keelan, D.K., 1982. Core analysis for aid in reservoir description. J. Pet. Technol. 34 (11), 2483–2491.
Krawczyk, C.M., Stiller, M., Bauer, K., Norden, B., Henninges, J., Ivanova, A., Huenges, E., 2019. 3-D seismic exploration across the deep geothermal research platform Groß Schönebeck north of Berlin/Germany. Geotherm. Energy 7 (15), 1–18.
Kristensen, L., Hjuler, M.L., Frykman, P., Olivarius, M., Weibel, R., Nielsen, L.H., Mathiesen, A., 2016. Pre-drilling assessments of average porosity and permeability in the geothermal reservoirs of the Danish area. Geotherm. Energy 4 (6), 2–27.
Kumar, D., Sugianto, H., Li, S., Patel, H., Land, S., 2014. Using relative seismic impedance to predict porosity in the Eagle Ford shale. SEG Annual Meeting.
Lindberg, D.V., Grana, D., 2015. Petro-elastic log-facies classification using the expectation-maximization algorithm and hidden Markov models. Math. Geosci. 47 (6), 719–752.
Lindsay, R., Koughnet, R.V., 2001. Sequential Backus Averaging: upscaling well logs to seismic wavelengths. Lead. Edge 20 (2), 188–191.
Lüschen, E., Wolfgramm, M., Fritzer, T., Dussel, M., Thomas, R., Schulz, R., 2014. 3D seismic survey explores geothermal targets for reservoir characterization at Unterhaching, Munich, Germany. Geothermics 50, 167–179.
Maithya, J., Fujimitsu, Y., 2019. Analysis and interpretation of magnetotelluric data in characterization ofgeothermal resource in Eburru geothermalfield, Kenya. Geothermics 81, 12–31.
Mallick, S., Adhikari, S., 2015. Amplitude-variation-with-offset and prestackwaveform inversion: a direct comparison using a real data example from the Rock Springs Uplift, Wyoming, USA. Geophysics 80 (2), 45–59.
Mathiesen, A., Nielsen, L.H., Bidstrup, T., 2010. Identifying potential geothermal reservoirs in Denmark. Geol. Survey Denmark Greenl. Bull. 20, 19–22.
Matthews, B.W., 1975. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochim. Biophys. Acta (BBA) - Protein Struct. 405 (2), 442–451.
Middleton, G.V., 1973. Johannes Walther's law of the correlation of facies. GSA Bull. 84 (3), 979–988.
Mukerji, T., Jørstad, A., Avseth, P., Mavko, G., Granli, J.R., 2001. Mapping lithofacies and pore-fluid probabilities in a North Sea reservoir: seismic inversions and statistical rock physics. Geophysics 66 (4), 988–1001.
Nielsen, L.H., 2003. Late Triassic – Jurassic development of the Danish basin and the Fennoscandian border zone, southern Scandinavia. Geol. Survey Denmark Greenl. Bull. 1, 459–526.
Nielsen, L.H., Mathiesen, A., Bidstrup, T., 2004. Geothermal energy in Denmark. Geol. Survey Denmark Greenl. Bull. 4, 17–20.
Rabiner, L.R., 1989. A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE 77 (2), 257–286.
Røgen, B., Ditlefsen, C., Vangkilde-Pedersen, T., Nielsen, L.H., Mahler, A., 2015. Geothermal energy use, 2015 country update for Denmark. In: Proceedings, World Geothermal Congress 2015. Melbourne, Australia.
Rosenkjaer, G.K., Gasperikova, E.G., Newman, G.A., Arnason, K., Lindsey, N.J., 2015. Comparison of 3D MT inversions for geothermal exploration: casestudies for Krafla and Hengill geothermal systems in Iceland. Geothermics 57, 258–274.
Saggaf, M.M., Toksöz, M.N., Mustafa, H.M., 2003. Estimation of reservoir properties from seismic data by smooth neural networks. Geophysics 68 (6), 1969–1983.
Scales, J.A., Snieder, R., 1997. To Bayes or not Bayes? Geophysics 62 (4), 1045–1046.
Ulrych, T.J., Sacchi, M.D., Woodbury, A., 2001. A Bayes tour of inversion: a tutorial. Geophysics 66 (1), 55–69.
Viterbi, A.J., 1967. Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. IEEE Trans. Inf. Theory 13 (2), 260–269.
Vosgerau, H., Mathiesen, A., Andersen, M.S., Boldreel, L.O., Leth Hjuler, M., Kamla, E., Kristensen, L., Brogaard Pedersen, C., Pjetursson, B., Nielsen, L.H., 2016. A WebGIS portal for exploration of deep geothermal energy based on geological and geophysical data. Geol. Survey Denmark Greenl. Bull. 35, 23–26.
Vosgerau, H., Gregersen, L., Kristensen, S., et al., 2017. Towards a geothermal exploration well in the Gassum Formation in Copenhagen. Geol. Survey Denmark Greenl. Bull. 38, 29–32.
Weibel, R., Olivarius, M., Kristensen, L., Friis, H., Hjuler, M.L., Kjøller, C., Mathiesen, A., Nielsen, L.H., 2017. Predicting permeability of low enthalpy geothermal reservoirs: a case study from the Upper Triassic−Lower Jurassic Gassum Formation, Norwegian–Danish Basin. Geothermics 65, 135–157.
Zwaan, B.V.D., Longa, F.D., 2019. Integrated assessment projections for global geothermal energy use. Geothermics 82, 203–211.