

Deep neural network application for 4D seismic inversion to changes in pressure and saturation: Optimizing the use of synthetic training datasets

Gustavo Corte^{1*}, Jesper Dramsch², Hamed Amini^{1,†} and Colin MacBeth¹

¹Heriot-Watt University, Edinburgh, EH14 4AS, UK, and ²Technical University of Denmark, Copenhagen, Denmark

Received February 2020, revision accepted May 2020

ABSTRACT

In this work, we tackle the challenge of quantitative estimation of reservoir dynamic property variations during a period of production, directly from four-dimensional seismic data in the amplitude domain. We employ a deep neural network to invert four-dimensional seismic amplitude maps to the simultaneous changes in pressure, water and gas saturations. The method is applied to a real field data case, where, as is common in such applications, the data measured at the wells are insufficient for properly training deep neural networks, thus, the network is trained on synthetic data. Training on synthetic data offers much freedom in designing a training dataset, therefore, it is important to understand the impact of the data distribution on the inversion results. To define the best way to construct a synthetic training dataset, we perform a study on four different approaches to populating the training set making remarks on data sizes, network generality and the impact of physics-based constraints. Using the results of a reservoir simulation model to populate our training datasets, we demonstrate the benefits of restricting training samples to fluid flow consistent combinations in the dynamic reservoir property domain. With this the network learns the physical correlations present in the training set, incorporating this information into the inference process, which allows it to make inferences on properties to which the seismic data are most uncertain. Additionally, we demonstrate the importance of applying regularization techniques such as adding noise to the synthetic data for training and show a possibility of estimating uncertainties in the inversion results by training multiple networks.

Key words: Reservoir geophysics, Time lapse seismic, Inversion, Machine learning.

INTRODUCTION

Estimating dynamic reservoir property change during a period of field production from four-dimensional (4D) seismic data has been a challenge and ambition for geoscientists in the oil and gas industry. These estimates are appealing for reservoir

monitoring and history matching purposes, because 4D seismic data offer information about reservoir property changes across the whole reservoir, at a specific production time. It complements well production information, which is spatially sparse but temporally dense. 4D seismic data provide information for the space between wells. But this information is encoded into the measured seismic amplitudes. So, we need to comprehend how the changes occurring inside the reservoir affect the seismic amplitudes we measure.

[†]Currently at Aker BP, Stavanger, Norway

*E-mail: gac2@hw.ac.uk

Along field production, the reservoir goes through constant change in properties such as fluid saturation, pore pressure, temperature, or even to changes in the reservoir rock architecture itself due to compaction and dissolution. The change in each of these properties has an independent impact on the seismic data, but they seldom act alone. Water injection for example leads to an increase in water saturation and an increase in pressure in the vicinities of the injector well. The observed 4D seismic amplitudes are a superposition of all the effects caused by the simultaneous variations in any dynamic property. The challenge is in quantitatively estimating the simultaneous contribution of each reservoir property to the final observed data. As is common in geophysical inversion, this is an underdetermined problem, prone to ambiguities and highly uncertain. Seismic information is limited and cannot provide enough independent measurements to characterize the whole reservoir state.

The information present in the variation of 4D amplitudes with offset (4D AVO) is crucial for quantifying multiple simultaneous reservoir property changes. We can highlight two major theoretical studies that use analytical solutions to show the possibility of quantifying the changes in both pressure and saturation directly from the 4D AVO data. In the first, Landrø (2001) follows a linearization of Smith and Gidlow's approximation to the reflection coefficient equation (Smith and Gidlow, 1987) to analytically derive a linear relation linking the 4D AVO gradient/intercept seismic attributes to the changes in two reservoir properties, pressure and water/oil saturation. The derived equations depend on rock physics-related parameters that can be estimated using laboratory measurements. In a different approach, Alvarez and Macbeth (2014) follow a linearization of the Aki and Richards' (1980) approximation to the reflection coefficient equation to derive an angle-dependent linear relation between the changes in pressure and oil/water saturations and the 4D seismic amplitudes. This relation also depends on reservoir petro-elastic parameters. Additionally, MacBeth *et al.* (2006) developed a data-based inversion method that assumes a linear link between any 4D seismic attribute and the changes in pressure and saturation. The parameters in the equations here are not directly related to any petro-elastic property, instead, they need to be previously calibrated using repeated well measurements, in a similar manner as the way neural networks are trained using direct observations. In this study, the authors run a principal component analysis to determine the best 4D seismic attributes for the simultaneous quantification of pressure and saturation changes (Floricich, 2006). The authors conclude that the 4D AVO related attributes contribute the most to the

inversion process and that the separation of effects could not be done without this AVO information.

Other methods take advantage of the 4D AVO information in many domains (impedance, gradient/intercept, seismic amplitudes) to invert for the changes in different reservoir properties such as pressure, compaction, and the saturations of water and gas (Trani *et al.*, 2011; Coleou *et al.*, 2013; Corzo *et al.*, 2013; Davolio *et al.*, 2013; Omofoma, 2017; Wong, 2017; Côte *et al.*, 2019). Most of these studies stress that there is a good deal of ambiguity and uncertainty in the solutions, thus any available external information should be used to constrain and/or regularize the inversion process (Blanchard and Thore, 2008; Blanchard, 2012). External information may come from the wells, as in MacBeth *et al.* (2006) and Coleou *et al.* (2013), where the authors use well-injected and produced volumes as global constraints to the saturation results. Reservoir simulation models can also be used to provide information to guide the inversion results. Davolio *et al.* (2013) and Omofoma (2017) use multiple realizations of a reservoir simulation model to define local hard bounds, constraining the possible inversion results. Côte *et al.* (2019) use the results of a history matched simulation model as local prior information in a Bayesian inversion approach to regularize the solution and provide soft constraints to the inversion results. Additionally, 4D seismic time-shift measurements have also been used as a data-based source of information in simultaneous inversion processes (Trani *et al.*, 2011; Thore and Hubans, 2012). In neural network solutions, we do not have the possibility of applying direct constraints to the inference process. Constraining the training dataset does not guarantee constrained results either, as the network can extrapolate beyond the training dataset. In this paper, we show a few techniques that can be used in the construction of the network architecture and training dataset that contribute to regularizing and constraining the inversion results.

The construction of the training dataset is a critical step that has great impact on the inversion results. Neural network applications, as opposed to most of the mentioned studies, do not rely on a physical model to establish the links between the seismic and reservoir domains. Instead, they rely on a training dataset composed of real input–output measurements, learning from it the non-linear relations that link inputs to outputs. The training dataset defines the ‘physical’ model that is used in the inversion, so it is important that it contains a good physical representation of the whole problem. In this case, a good training dataset should represent the whole reservoir, containing the global variability on reservoir quality and the possible dynamic property combinations.

Measured data to compose a training dataset can only come from repeated well measurements at seismic acquisition times. MacBeth *et al.* (2006) use well data in a model calibration approach analogous to neural network training. The authors pre-define a linear equational link between the seismic and reservoir domains and iteratively calibrate the equations' parameters to fit the data measured at a few well locations. The main difference to neural network training is that deep neural networks (DNNs) contain thousands more parameters to be calibrated, leading to much more complex non-linear relations. Consequently, deep neural networks need a much larger amount of data for satisfactory training. Nonetheless, Cao and Roy (2017) perform a synthetic study showing that a neural network can also be trained successfully using only information at well locations in a 4D reservoir property inversion application. In real reservoir cases though, the necessary repeated saturation well logs are not common and may be lacking as whole, as is true in our case study. This type of data is sparse, and it can be argued that it may be biased to good reservoir areas, where the wells are located, and thus, incapable of representing the entire reservoir. More often than not, just the well data are not sufficient to properly train a neural network.

The alternative is to use synthetic data to help in the construction of the training dataset. Ayzenberg and Liu (2014) present a real reservoir case of a neural network application to 4D pressure and saturation inversion where the authors populate their training dataset with reservoir simulation results and real seismic observations at a few well-understood areas. To extend their training dataset beyond the wells, they begin a shift to synthetic data, but only on the reservoir domain, keeping the real observed seismic data. Xue *et al.* (2019) use a fully synthetic dataset to train their neural network to quantify the changes in water saturation on a real reservoir case. The authors make use of a wedge model as a static frame and random sampling of the dynamic domain. More recently, Zhong *et al.* (2020) presented a solution using convolutional generative adversarial networks to invert impedance change images to reservoir property changes. Their convolutional approach analyses full images, incorporating a spatial correlation aspect into the inference results. For this reason, their synthetic training dataset is composed of full reservoir images, created by running 300 reservoir simulations with varying static models. Although convolutional networks are the state-of-the-art in image analysis, they require an immense amount of previous work to prepare synthetic training datasets.

This paper presents a DNN application to inverting 4D AVO seismic data into the simultaneous changes in three reservoir properties: pressure (ΔP), water saturation (ΔS_w) and gas saturation (ΔS_g). It provides a quick and practical alternative to more well-established inversion methodologies. As a good platform for comparison, we present a Bayesian model-based inversion approach applied to the same dataset in Côte *et al.* (2019), and a direct comparison of methods in Dramschet *et al.* (2019a).

The DNN is trained with synthetic data and applied to real 4D seismic data from a North Sea field. We use a reservoir simulator to seismic modelling approach (Sim2Seis) to construct four synthetic training datasets with the objective of assessing the impact of the distribution of data in the training dataset on the quality of the inversion results when applied to a real 4D seismic dataset. The training datasets presented differ essentially on how much external physical information is used to constrain and distribute the data. We show the value of using physics informed and fluid flow consistent realizations to create a realistic distribution of data in a synthetic training dataset. Furthermore, we show the importance of training the DNN on noisy synthetic data and the possibility of estimating uncertainties in the results by training multiple DNN models with varying signal to noise levels. With this we address the problems of constraining the results with external physical information and regularizing solutions to avoid overfitting of the training data and inverting noise.

FIELD AND DATASET

The field is composed of stacked turbidite channel and sheet-like sands ranging from 5 to 30 m in thickness and 25–30% in porosity. It is highly compartmentalized both laterally and vertically due to faults and intercalating shales. The sandstone reservoir is present in four adjacent fault blocks. Faults between blocks are sealing, creating four isolated segments with varying water oil contacts (Fig. 1b). This whole structure dips and thickens to the north-west (Fig. 1a). Inside each segment there are faults that may be sealing or not, leading to a few isolated compartments and a highly complex geological setup. Figure 1(c) shows the vertical sum of pore volume in the reservoir, where we see clearly the channel features. Detailed explanations of the depositional and stratigraphic evolution of the region can be found in Ebdon *et al.* (1995) and Lamers and Carmichael (1999).

The initial pressure in the field was only around 3 MPa above bubble point pressure, making pressure maintenance to prevent gas exsolution the main production strategy. To maintain pressure, water injectors were drilled in the water leg, on the west flank of the reservoir and in other select zones

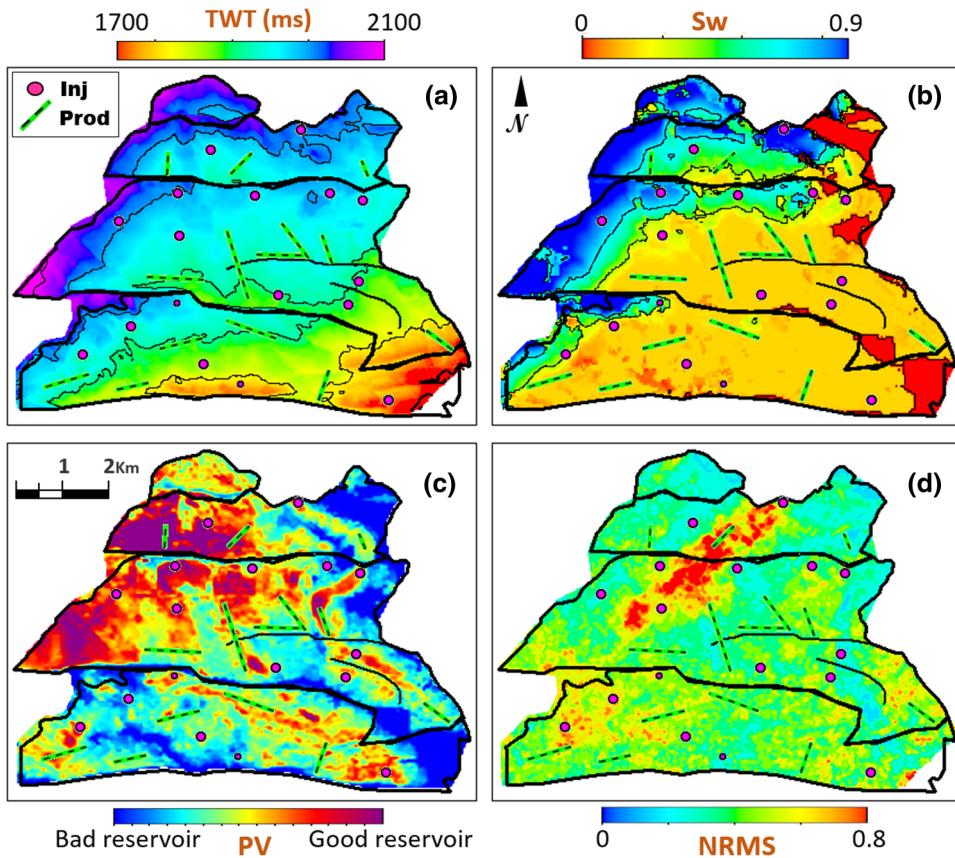


Figure 1 (a) Top reservoir horizon in time, (b) initial water saturation for reservoir sandstone, (c) pore volume for reservoir and (d) NRMS measure of non-repeatability.

around the reservoir. Even so, production in this complex compartmentalized structure led to areas with strong pressurization due to water injection into isolated compartments, while other areas lack the pressure support and experience gas release due to pressure depletion below bubble point pressure. This creates a complex dynamic setup on top of an already complex static framework. The challenge is to use four-dimensional (4D) seismic data to quantitatively estimate simultaneous changes in three dynamic properties: pressure (ΔP), water saturation (ΔS_w) and gas saturation (ΔS_g) across the reservoir. Both pressure and gas effects on seismic data are non-linear, so the inversion method should deal properly with the non-linearities due to changes in these two properties.

The production strategy for this reservoir included regular 4D seismic acquisitions to aid in monitoring reservoir production. In this paper, we present the results of the method applied to one of the many monitor seismic acquisitions acquired along the field life. The reservoir is thin to seismic standards, being identified in a seismic quadrature section as one single

trough (Fig. 2). For this reason, all of the analysis is done in map form. The seismic data used for inversion (Fig. 3) are the time-lapse difference in the sum of negative amplitudes map attribute (ΔSNA), extracted from quadrature seismic volumes along the reservoir time window. This map extraction consists of a vertical sum of the negative seismic amplitudes between the top and bottom reservoir horizons (shown in Fig. 2) for the baseline and monitor volumes, followed by a subtraction of these two maps (monitor – baseline). Calculated time-shifts are very small and show no correlation with the seismic amplitudes or production data, so unfortunately, they were not useful. We use the pre-production seismic acquisition as the baseline for generating the 4D seismic maps.

Figure 1(d) shows the normalized root mean squared (NRMS) map extracted from the monitor-baseline pair used for inversion. This was calculated in a 400 ms time window, 100 ms above the reservoir, so that it is away from any production-related effects, but deep enough not to capture acquisition footprints. NRMS is a measure of comparison between two seismic traces. When extracted in the overburden

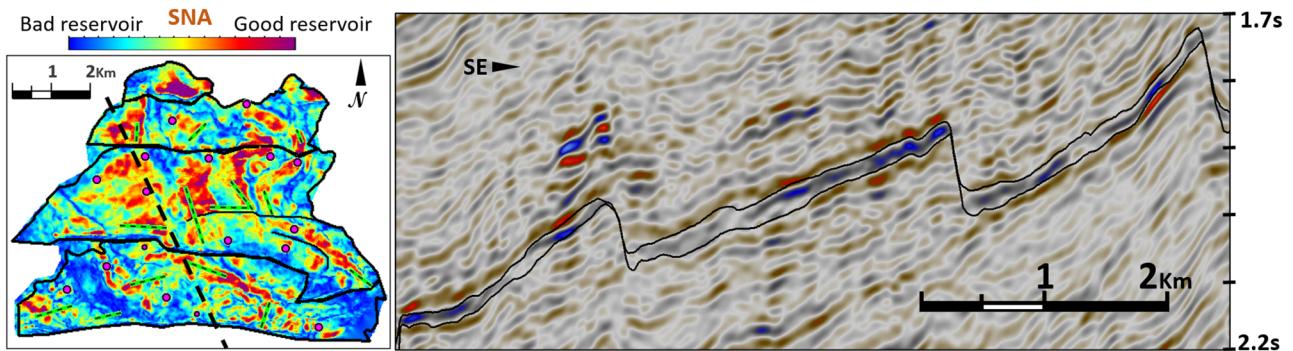


Figure 2 Sum of the negative amplitude (SNA) seismic map attribute for the baseline acquisition (left), extracted from the quadrature full stack seismic 1996 baseline acquisition (right). Seismic section on the right shows the top and bottom reservoir horizons between which seismic attributes are extracted.

region, NRMS maps are interpreted as a measure of non-repeatability between two seismic acquisitions, providing a relative estimation to the 4D seismic data quality across the reservoir. NRMS values range from 0 to 2. Values of 0 represent perfectly correlated traces, 1 is observed for uncorrelated random noise traces while 2 is observed for anti-correlated traces. It is commonly presented as percentage values, with a multiplication of the calculated NRMS values by 100. Calculated NRMS values range from 20% to 40%, except for the undershoot zone, where repeatability is poor due to the presence of a production platform. The poor repeatability in this area has a strong impact on the quality of the signal and, as we will see, in the inversion results for this area.

To aid in interpretations, Fig. 4 shows vertical average maps of the reservoir simulation results. In all maps in Figures 3 and 4, the colour scales are adjusted to represent softening effects as yellow–red and hardening effects in blue–green. Softening effects are defined as those that are related to a

decrease in the bulk reservoir rock seismic impedance, while hardening effects relate to increases in reservoir impedance. For example, increase in water saturation lead to an increase in reservoir impedance because water impedance is higher than the oil impedance. Increases in gas saturation on the other hand lead to decreases in the reservoir impedance because gas' seismic impedance is lower than that of the oil. For this reason, their colour scales are inverted. This scheme simplifies the comparison of reservoir property maps and seismic maps.

In Figs 3 and 4 some areas of interest are circled, showing important features that we will use as guides for a qualitative assessment of the inversion results. This initial interpretation is inherited from previous four-dimensional amplitudes with offset (4D AVO) studies done for this dataset (Côte *et al.*, 2019). Circled areas represent the reservoir property that dominates each seismic anomaly. This depends on the seismic sensitivity to changes in each property and on what other property changes are occurring simultaneously. Seismic

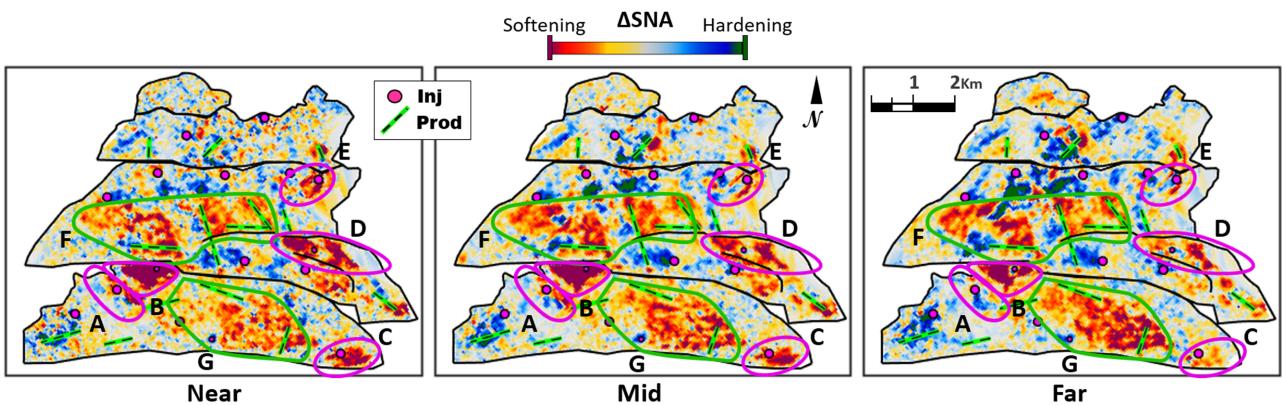


Figure 3 Seismic data used for inversion: time-lapse change in the sum of negative amplitudes attributes extracted from the near, mid and far stack.

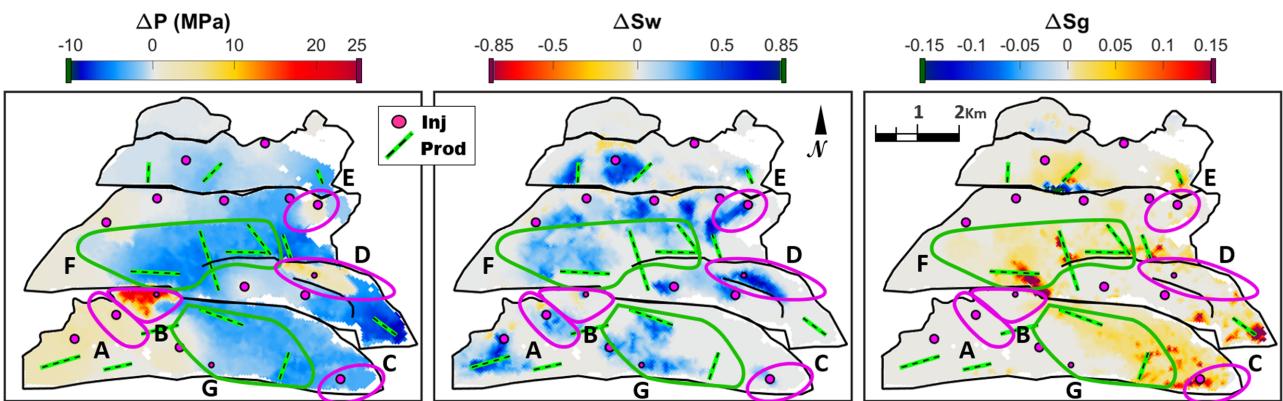


Figure 4 Vertical average maps of the results of the reservoir flow simulation for the changes in pressure (ΔP), water saturation (ΔS_w) and gas saturation (ΔS_g).

amplitudes may be more sensitive to changes in some property over other. If changes in both these properties are superposed, the seismic effects related to one property will overcome and dominate over the other. Uncertainties tend to be larger in the properties that have their effects dominated by co-occurring changes in other properties to which the seismic data are more sensitive.

Zones circled in magenta (A, B, C, D and E) are well-understood softening signals related to pressurization around water injectors. In this case, the hardening signals related to the increase in water saturation are overcome and dominated by the stronger softening signals related to the pressure increase. Zones circled in green (F and G) are well-understood softening signals related to gas saturation increase. In these zones, the gas saturation softening response dominates, but it is always in competition with hardening signals due to water saturation increase. In zone F, the water comes mainly from the aquifer located to the north and west, aided by injectors in the water leg. This zone is particularly complicated because, aside from the water-gas competition, it lies under a platform in an area of low seismic repeatability (Fig. 1d), thus the seismic data here are very noisy and uncertain. The AVO gradient is especially affected, crippling the data capability of differentiating between pressurization-related softening effects and gas saturation related ones. For this reason, inversion results in zone F may show leakage between gas and pressure effects. In zone G, the water comes from the two injectors placed on its southwest edges. We see from the simulation results (Fig. 4) that a considerable amount of water has been injected in this zone, but no hardening signal can be seen in Fig. 3. As seismic data are much more sensitive to increases in gas than in water saturations, the gas-related response dominates.

We see some hardening signals related to water saturation increase in Fig. 3, but no evident hardening signal related to pressure depletion. The non-linear nature of pressure effects on 4D seismic data means that even though pressure increase leads to strong softening signals, pressure depletion results in very low hardening anomalies. To complicate further, in the present case, pressure depletion is always accompanied by gas coming out of solution, so the gas softening effects always dominate over the weak hardening pressure depletion signals. This makes it particularly difficult to quantify pressure depletion values from the 4D seismic data.

The main challenges for this inversion are to quantify the pressure increase values in the pressurized compartments, to differentiate between pressure related and gas related softening signals, to determine areas of pressure depletion and to locate water fronts in areas where water saturation related hardening signals are dominated by other competing effects.

DEEP NEURAL NETWORK ARCHITECTURE AND TRAINING

We employ a deep neural network (DNN) with the encoder-decoder architecture to translate the mapped sum of negative amplitudes (ΔSNA) seismic attributes (Fig. 3) into the corresponding changes in three reservoir properties: ΔP , ΔS_w and ΔS_g . Although we mention maps, this network is not working on full images, instead it makes pixel-wise inferences, analysing each pixel individually and independently of others. In the training phase, it uses all the pixels (or samples) provided in the training dataset to define one general non-linear function that best links the values given in the input domain to the output domain. In the inference phase, it applies this non-linear function individually to each pixel in the map

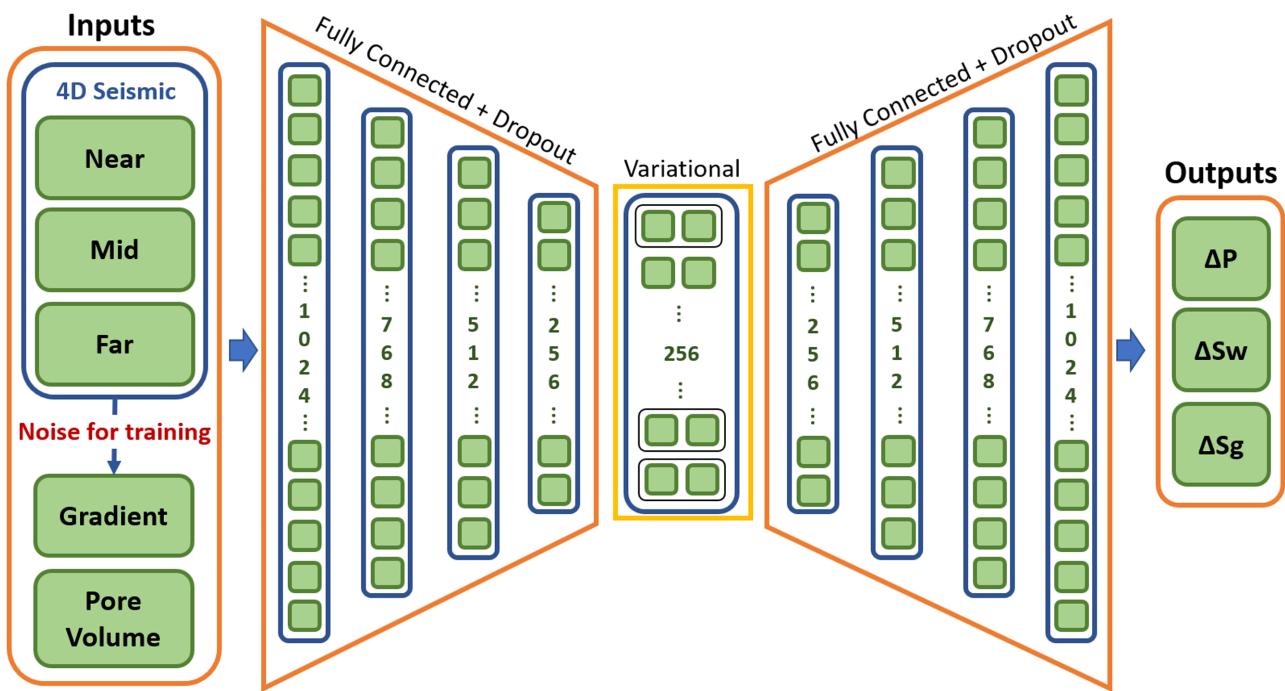


Figure 5 Deep neural network architecture.

provided. Being so, there is no lateral correlation constraint or smoothing technique to ensure lateral correlation and control noise content. For this reason, we employ three important techniques that help to avoid overfitting and inverting noise: dropout regularization (Srivastava *et al.*, 2014), variational Bayes encoding layer (Kingma and Welling, 2014) and training with noisy samples (Bishop, 1995).

The full DNN architecture consists of four encoding layers that compress the input information from 1024 neurons in the first layer to 256 neurons in the last, a central layer where variational encoding is implemented and a mirror decoder structure that decompresses the information back into 1024 neurons (Fig. 5). Compressing the information serves as forcing function for the network to learn the most meaningful features in the data in regard to the optimization objective. In each of the encoder-decoder layers we use a dropout regularization technique in the training phase, this randomly excludes 20% of the neuron connections in each training iteration. Dropout regularization is commonly used as a technique to prevent overfitting of the training data, leading to a more general model and helping in dealing with noisy datasets (Srivastava *et al.*, 2014).

The central encoding layer is arranged so that each neuron defines two outputs which are used to define Gaussian distributions (mean and standard deviation). In the follow-

ing layer, each neuron draws one random value from each Gaussian distribution. Consequently, a neuron from the central layer feeds slightly different values to each neuron in the following layer, as opposed to feeding the same value to all neurons as it is done in all other connections in the network. Using variational encoding in the central layer, instead of a fixed link from encoding to decoding, provides a flexibility to the network, making it more general and robust to noise (Kingma and Welling, 2014). To be able to train the network with back propagation of the gradients, we use the ‘reparameterization trick’ as in Kingma *et al.* (2015). To construct this architecture, we used a Tree of Parzen scheme for estimation of the hyperparameters (Bergstra *et al.*, 2013). This is an optimization scheme that uses a subset of the training data to find the best hyperparameters for the problem at hand. Adjusted hyperparameters were the number of layers and neurons per layer.

The input layer contains the time-lapse difference in the ΔSNA seismic attribute, extracted for the near, mid and far angle-stacks. From these three 4D seismic attributes, the network calculates the four-dimensional amplitudes with offset gradient to be used as an input as well. In the training phase, before calculating the gradient, we add white Gaussian noise to the synthetic data. This step is crucial for achieving meaningful results when making inferences on the noisy seismic

data. Training with noisy synthetics is equivalent to Tikhonov regularization of inversion processes (Bishop, 1995), so it controls overfitting and prevents the DNN from treating the noise as signal. The magnitude of the added noise is controlled by one single parameter, the signal to noise ratio, that defines the standard deviation of the Gaussian distribution from which random noise values are drawn. In a later section of this paper, we present an analysis of the impact of this noise parameter on the inversion results and elaborate on how to assess the performance of the trained models and select an optimal value for this parameter.

In addition to the time-lapse seismic attributes, we also include in the inputs the reservoir pore volume, calculated from the reservoir simulation model. This static parameter is relevant as reservoir pore volume acts as a scalar on the 4D seismic amplitudes, leading to stronger responses in areas with higher reservoir pore volume. We observed that the addition of this static parameter within the network architecture is essential in achieving a more accurate regression result. The pore volume as an input parameter abstracts the information the neural network has to learn from the seismic input maps, alleviating the learning process for the network (Dramsch *et al.*, 2019b).

Supervised training of neural networks relies on an ensemble of samples of known input-output pairs which define the training dataset. We use the Adam optimization method (Kingma and Ba, 2015) for training. It is a stochastic gradient descent optimization approach with Nesterov momentum and an adaptive learning rate. The algorithm iteratively updates the neuron weights that define the network state, in each iteration outputs are calculated from the inputs provided in the training dataset. The algorithm then compares the calculated outputs to the outputs in the training dataset, using a mean squared error objective function. More details on the architecture and training strategies used can be seen on Dramsch *et al.* (2019b).

The training data are presented to the network in the form of $N \times 1$ vectors, with N being the number of samples in the training dataset. Each of the parameters in the input ($\Delta S_{\text{A}_{\text{near}}}$, $\Delta S_{\text{A}_{\text{mid}}}$, $\Delta S_{\text{A}_{\text{far}}}$ and Pore Volume) and output (ΔP , ΔS_w and ΔS_g) are presented as a separate $N \times 1$ vector.

CONSTRUCTION OF THE SYNTHETIC TRAINING DATASETS

Defining the training dataset is one of the most important steps in a neural network workflow. In the present application, the network represents (or replaces) the physics that links the seismic domain to the reservoir domain, but in fact it has no

knowledge or any information about the physics it represents. It learns this implicitly from the training dataset, so the quality of the training dataset will define how well the network will mimic the physics of the problem and its capability of inferring meaningful results from unseen data.

To construct the training dataset, we need to form an ensemble of input-output realizations. These are independent single pixel realizations of ΔP , ΔS_w , ΔS_g and Pore Volume and their resulting sum of negative amplitudes values for near, mid and far stacks. In the lack of a good and sufficient measured dataset, the alternative is to use synthetic data, generated based on a physical model that represents as best as possible the problem at hand.

In this application, we employ a reservoir simulator to seismic modelling (Sim2Seis) technique (Amini, 2014) to create synthetic seismic data from a reservoir flow simulation model. We use one fixed reservoir model that has been previously history matched to production data. The model grid spans the whole reservoir, so it contains a good representation of the variability of the static reservoir properties. We use this static geological model as the frame for creating four different training datasets that differ in the distribution of the sampled realizations in ΔP , ΔS_w , ΔS_g and Pore Volume. For each realization, we extract a pseudo-log from a certain location in the static reservoir model. The pore volume value is calculated from this pseudo-log, as the vertical sum of the pore volumes for active cells in the reservoir zone. Next, we define a sample realization for the ΔP , ΔS_w and ΔS_g values. These values are distributed vertically in all reservoir cells in the extracted pseudo-log, always respecting initial and residual saturation values in each cell. As this is a map-based approach, we do not model different vertical distributions of fluids or pressure in the reservoir. With this approach, we maintain the vertical resolution of the reservoir model on the static properties, but the dynamic properties represent a vertical average. Angle-stacked seismic traces are then calculated for baseline and monitor states using Sim2Seis and from them the sum of negative amplitudes attributes are extracted along the reservoir time window.

The reservoir model cell dimensions are 50×50 m in the lateral dimensions and 3 m vertically. The regular seismic grid separation is 25 m in both inline and crossline directions. For all synthetic seismic calculations, we use one fixed petro-elastic model that was previously calibrated to the well logs and the observed seismic data (Amini and MacBeth, 2015; Amini, 2018). The petro-elastic model is based on a mixture of sand and shale grains following volume fractions given by reservoir net to gross. Rock frame elastic moduli are calculated

Table 1 Training dataset sizes and training times

Training Datasets	Number of Pseudo-Logs	$\Delta P \times \Delta S_w \times \Delta S_g$ Realizations per Pseudo-Log	Total Number of Samples	Average Training Time (minutes)
1	300	1130	339,000	~17
2	300	475	142,500	~7
3	12,944	7	30,608	~4
4	12,944	100	1,294,400	~60

following Nur's critical porosity model (Nur *et al.*, 1998). Pressure dependence follows the compliance model by MacBeth (2004), using for model parameters the values measured for a sandstone reservoir in the west of Shetland islands, analogous to Shiehallion, which are provided in his paper. Effective fluid elastic moduli are calculated using a homogeneous saturation model and Gassmann's fluid substitution equations (Gassmann, 1951) bring all the pieces together to calculate the saturated rock elastic moduli. Seismic traces are calculated by a convolution of a source wavelet with the reflectivity series calculated with the petro-elastic model. We use separate wavelets for each angle-stack, which are extracted individually from the each of the observed angle-stacked seismic volumes. Wavelets are all zero phase with central frequencies of 29 Hz (near), 25 Hz (mid) and 20 Hz (far).

We tested four training datasets with the objective of assessing the impact of data size and the use of physics informed realizations to populate the training dataset. The amount of physics information used to construct the training datasets increases from set 1 to 2 and to 3. For Training dataset 4, data augmentation techniques are used to assess the impact of dataset size on the inversion results, while keeping the similar levels of physics information as in Training dataset 3.

Table 1 shows a comparison of the amount of data, how it is distributed and the resulting training runtimes for all four training datasets presented. Figure 6 shows the global distribution of data in each training dataset.

Regularly sampled realizations (Training datasets 1 and 2)

In neural network applications, it is often desirable to have general trained networks that can be applied satisfactorily to many different cases. In the current application, this would mean training one general network that can be applied to many reservoirs with differing static and dynamic situations. The most general training dataset should contain realizations representing all possible situations, both on the static (Pore Volume) and on the dynamic ($\Delta P \times \Delta S_w \times \Delta S_g$) domains.

Training dataset 1 is constructed in a way as to be the most general. For this, we generate synthetic realizations for every point in a regular four-dimensional sample grid (Pore Volume $\times \Delta P \times \Delta S_w \times \Delta S_g$). The sample grid covers the whole range of possible situations on all four properties. To keep the saturation values realistic, we apply a unity constraint on the sum of the saturations, and always respect residual oil and water saturations, so that as an example, if $\Delta S_w = 0.5$, then ΔS_g is constrained to values between 0 and 0.15. For each Pore Volume value defined in the sample grid, we extract one single pseudo-log from the simulation model that best represents this Pore Volume value. We then proceed to calculating time-lapse seismic traces for all realizations in the dynamic domain. The reservoir simulation model is used here only to define the static frame for calculating synthetic data, but the reservoir simulation results are not used.

The same sampling strategy is used for Training datasets 1 and 2. The only difference between these datasets is one simple constraint used in Training dataset 2. This constraint comes from our external understanding of the physical processes governing the problem at hand. As has been mentioned, in this reservoir, the increase of gas saturation is a response to pressure depletion. As pressure goes below bubble point pressure, gas comes out of solution from the oil phase. Wherever pressure increases from initial pressure, which is above bubble point pressure, we expect no gas saturation change to occur. Thus, training sample realizations containing simultaneous increases in pressure and gas saturation are not representative of the reality analysed and could be interfering negatively in the solutions. For Training dataset 2, we delete all samples with simultaneous increase in pressure and gas saturation, consequently making it less general, more specific to the reservoir situation. The comparison of Training datasets 1 and 2 pinpoints the impact of one simple constraint, showing the benefits that can be achieved by adding one bit of physical information to constrain the realizations in the synthetic training dataset.

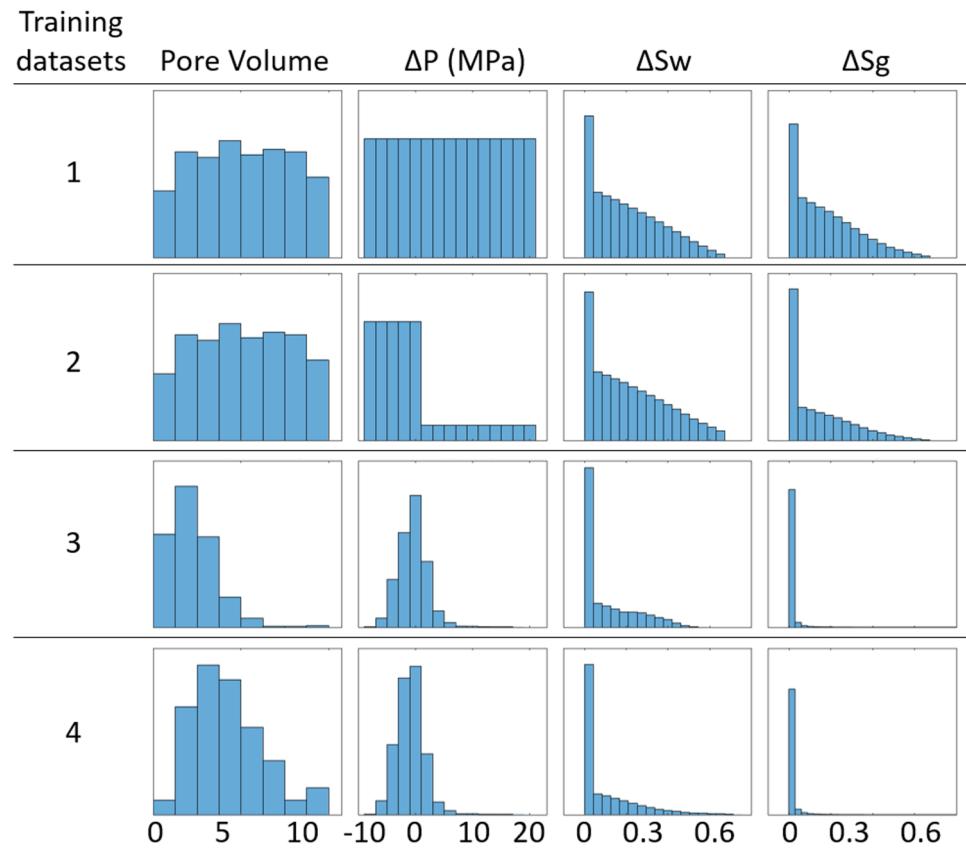


Figure 6 Histograms showing the global distribution of data in Training datasets 1, 2, 3 and 4.

Fluid flow consistent realizations (Training datasets 3 and 4)

A full 3D reservoir flow simulation offers $\Delta P \times \Delta S_w \times \Delta S_g$ realizations that respect a full range of physical processes, including the causal relationship between pressure depletion and gas saturation increase, but also processes related to wettability, capillary forces, relative permeability, etc. Restricting our training samples to the combinations offered by the results of a reservoir simulation model means that the training dataset will respect all of these physical processes and thus be even more realistic and similar to what an unbiased measured dataset would offer. The resulting training datasets are even less general and more specific to the reservoir situation. Reservoir simulation results are used in Training datasets 3 and 4.

For Training dataset 3, we use the reservoir simulation results as the only $\Delta P \times \Delta S_w \times \Delta S_g$ combinations in the training dataset. To maintain consistency with the previous approaches we do not run simulator to seismic modelling directly on the reservoir simulation results, first we extract vertical average maps for ΔP , ΔS_w and ΔS_g at a given time-step,

then we apply the same forward modelling process as previously described (distributing the averaged maps vertically in the extracted pseudo-logs). We extracted simulation results for eight time-steps corresponding to the seismic acquisition dates. This choice on time-steps was done simply because it is common to use a reservoir simulator to seismic modelling workflow at seismic acquisition times to compare synthetic with observed seismic data, so in this case we did not need to extract additional simulation results, all the data were already available and had been generated and used for other previous purposes, this would usually also be true in a regular industry setup. There is no real need to restrict the time-steps to seismic acquisition dates though.

There is a great difference on the distribution of data using this approach (Fig. 6). In this case, we run synthetic traces for every trace location in the simulation model, with this we have 12,944 static pseudo-logs instead of the 300 of the previous models, on the other hand, we only have seven $\Delta P \times \Delta S_w \times \Delta S_g$ realizations per pseudo-log. Thus, pore volume is much more finely sampled, and the global distribution of pore volume is no longer uniform, here it resembles the global

Table 2 Performances and training SNRs of the best models in the synthetic and well validations

Training Datasets	Synthetic Validation					Well Validation	
	Total NMSE	ΔP NMSE	ΔS_w NMSE	ΔS_g NMSE	Training SNR	ΔP NMSE	Training SNR
1	3.08	2.68	2.19	4.37	21.5	6.60	39.5
2	1.56	0.71	1.66	2.31	42	2.50	19
3	0.66	0.62	0.70	0.64	17	0.46	14
4	0.56	0.50	0.73	0.45	11	0.40	12

distribution that could be found in reality. Regarding the dynamic domain sampling, though the simulation results cannot be taken as the real reservoir fluid flow state, the global distributions should resemble reality, as the model has been history matched to well production and injection volumes and pressure measurements. The resulting training dataset is smaller in total number of samples but is much more representative of the reality of the reservoir.

It remains true that DNNs benefit from larger amounts of data. For this reason, in Training dataset 4 we make an effort to augment the previous training dataset, while maintaining all the physical relationships present in the reservoir simulation results. We do this by grouping all the dynamic domain samples ($\Delta P \times \Delta S_w \times \Delta S_g$) in the previous training dataset to create an ensemble of possible samples that respect the reservoir flow physics. For each static pseudo-log extracted from the simulation model, we draw 100 random samples from this global ensemble of dynamic realizations to run time-lapse synthetic seismic traces. In practice, we take the reservoir simulation results found at one trace location and apply it to a different trace location, always respecting end member saturation limits. This approach maintains global distributions in pore volume, ΔP , ΔS_w and ΔS_g that are similar to Training dataset 3 (Fig. 6), while augmenting the data size by 100 times.

MODEL PERFORMANCE QUANTIFICATION

To assess the performance of each trained model, we use two validation approaches. The first one is based on synthetic data, where ground truth is exact, but the seismic data are more well behaved and may not represent the real observed data properly. The second is applied to the observed data itself, using as ground truth for validation the well measurements at seismic acquisition times. In the second approach, the seismic data are a good representation of the real data, but the ground truth is not exact and carries itself some uncertainty. A summary of the performance quantification results for all four training datasets can be found in Table 2.

As has been mentioned, we add random noise to the synthetic training samples in the training phase, to take advantage of the regularization property this technique carries (Bishop, 1995). A deep neural network that is trained on noisy synthetics is more generalized and capable of interpreting the noisy character of the observed data and avoid overfitting this noise. This capability varies with the amplitude of the noise added. If we add too much noise, this corrupts the amplitudes with offset information in the training phase and inferences are compromised. To assess the impact of the training noise parameter on the inference performance, we train 100 models for each training dataset, with varying training noise levels, ranging from 0 to 50% in noise to signal ratio (SNR). Intuitively, the best performance should be achieved by the model that is trained using the same SNR levels as are present in the inference data. As the final objective is to apply the inversion to the real observed 4D seismic, it is worth estimating the SNR level present in this dataset. For this, we consider the noise power as the root mean square (RMS) of the observed seismic amplitudes in the northernmost segment, where no production has occurred and no 4D seismic signal is expected, and the signal power as the RMS across the whole reservoir. Observed 4D seismic data noise level was estimated to be 14%.

Synthetic validation

In order to make a fair comparison between the performances achieved with the four different training datasets, we created an additional set of synthetic data, which is not used for training in any of the models. We want the validation dataset to represent the real data as well as possible, so that the performances calculated can be interpreted as the capability of a certain model to make correct inferences on the real data. For this reason, the validation dataset is constructed by extracting the reservoir simulation results for one separate time step, which is not used in Training datasets 3 and 4. With this the validation dataset contains 12,944 samples. Then, we add random noise with 14% noise to signal ratio (SNR)

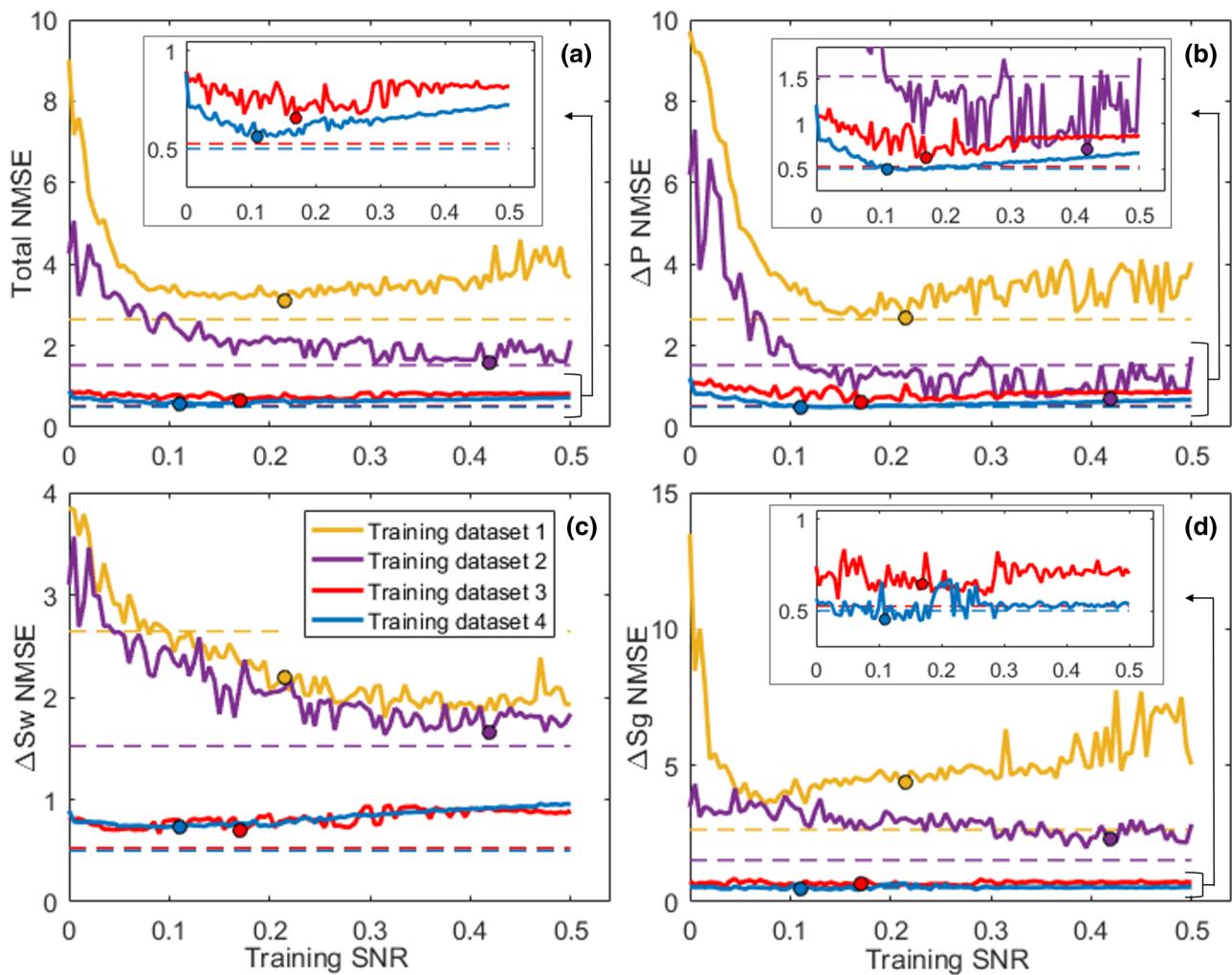


Figure 7 NMSE results for the synthetic validation. (a) Total, (b) ΔP , (c) ΔS_w and (d) ΔS_g .

levels, so that the synthetic validation dataset contains noise levels comparable to the observed 4D seismic. We apply the inference step to the same noisy synthetic validation dataset for all trained models and assess the performance by comparing the inference results to the reservoir property values in the validation dataset. Comparisons are made individually for each reservoir property using the normalized mean squared error (NMSE) metric, these are then individually normalized using the standard deviation of each target reservoir property and averaged to achieve a global performance metric.

Figure 7 shows plots for the performances achieved as a function of the training noise to signal ratio (SNR) for all four training datasets. Plots (b)–(d) show the performances on each property individually and plot (a) shows the global performances for all the trained models. To provide a reference

we also show the performances achieved in the ideal noiseless case (dashed lines), where the models are trained without noise addition and applied to a noiseless version of the synthetic validation dataset. Models that achieved the best total performance for each training dataset are highlighted in Fig. 7 and their results are plotted in map form in Fig. 8.

We observe a global increase in performance (Fig. 7a) with the addition of training noise for all training datasets. Best performances do not achieve the performance values for the ideal noiseless synthetic case (dashed lines) but are all considerably better than when no training noise is added. The addition of training noise is responsible for a decrease of 30–60% in the total normalized mean squared error (NMSE). For Training datasets 1, 3 and 4, the total NMSE reaches a region of lower values with training SNRs around the noise level applied to the validation dataset (14%). Training

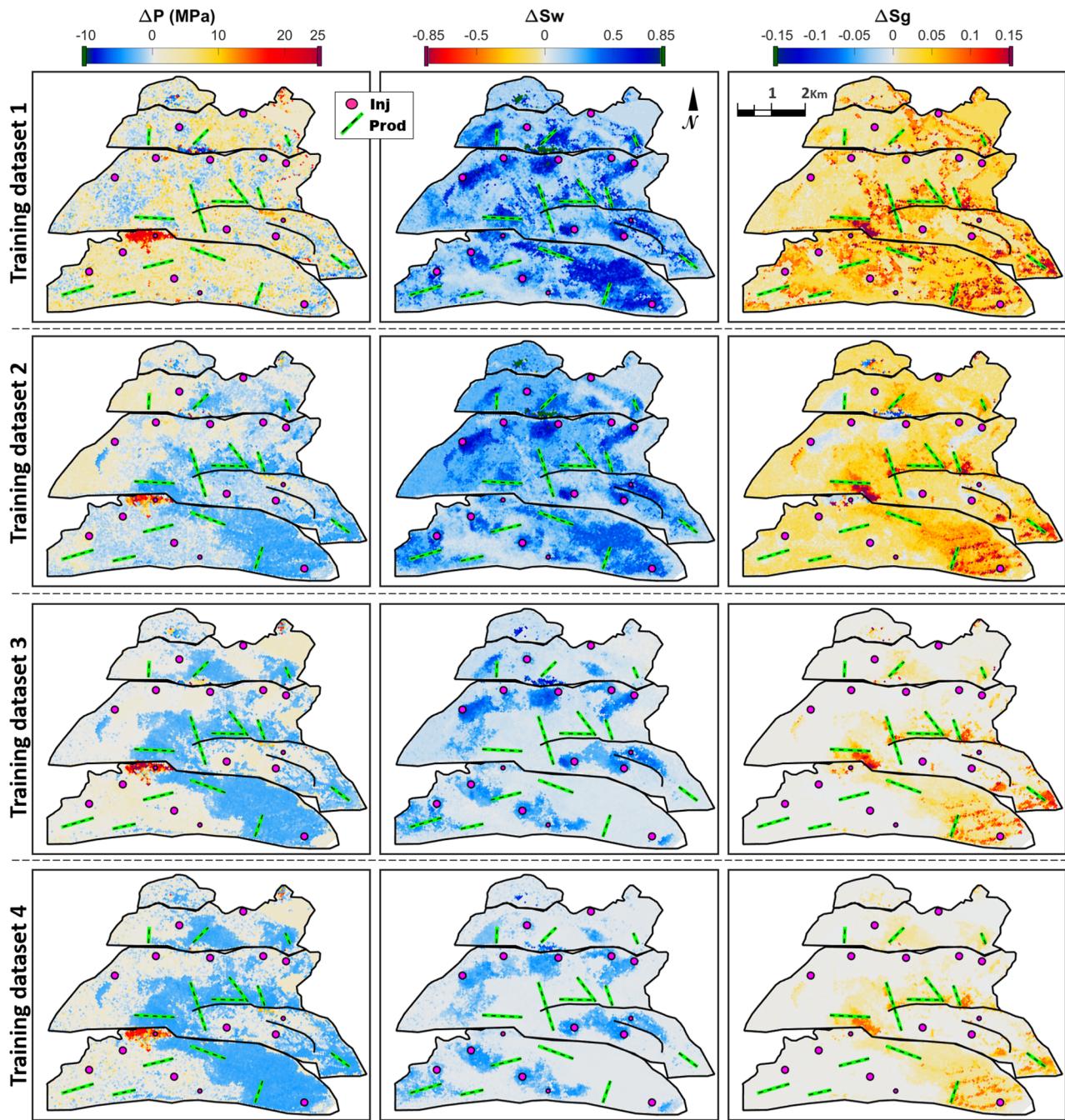


Figure 8 Map results from the best models in the synthetic validation.

dataset 2, on the other hand, shows an increase in performance with training SNRs up to 49%. Intriguingly, for this training dataset, it has been beneficial to train the model with noise levels three times higher than the noise content of the data it was applied to.

In a comparison between the training datasets, we see a global increase in the performances from Training dataset 1 to 2(a). This is mostly due to better estimations of the pressure (b) and the gas saturation (d). Performances for the water saturation estimation are similar for Training datasets 1 and 2.

This shows us that the use of a simple physics-based constraint on the pressure and gas saturation has a positive impact on the estimations of these two properties, while it has little impact on the water saturation estimations. Training datasets 3 and 4 show considerably better performances than the two previous ones, showing how much of a benefit can be achieved by using training datasets populated with fluid flow consistent samples. Training dataset 4 achieves slightly better total performances (a) than Training dataset 3, indicating that the data augmentation has been beneficial to the quality of inferences. Again this is due to a difference in performances for estimating pressure and gas saturation, as water saturation performances are similar. For these two training datasets, the addition of noise has no impact on the performance of the gas saturation estimations.

Well data validation

Although we have made an effort to create a validation dataset that mimics the real data, the performances achieved when applying a model to synthetic data may not represent the truth when the model is applied to the observed four-dimensional seismic data. As has been mentioned, labelled data for the real case are only present as measurements made at wells during seismic acquisition times. For Schiehallion, we only have bottom hole pressure measurements, which is the main reason why we need synthetic data for training in the first place. There is no real data to assess the inference results on the saturation values, but we can nonetheless assess the pressure estimations only. The synthetic validation indicates that total performances are mostly driven by the performance on estimating pressures. This provides some confidence in evaluating the models using well bottom hole pressure measurements only, as the performance on the pressure estimates may be a reasonable representation of the global performances.

When comparing well bottom hole pressures to seismic inverted pressures, it is important to keep in mind that the pressure data measured at wells are not exact ground truth, it carries uncertainties due to mainly two reasons:

- Spatial: In deviated wells, there is uncertainty in locating where along the well perforations the bottom hole pressure measurements should be related to. This is not an uncertainty intrinsic to the well measurements, it becomes relevant only when comparing inverted pressures along the well perforations to the measured values.
- Temporal: A seismic acquisition may take weeks or even months to be finished. Along this time, reservoir pressures

are not constant and bottom hole pressure measurements may vary substantially. There is uncertainty in selecting the ideal time to extract bottom hole pressure measurements for the validation (Omofoma *et al.*, 2019).

Figure 9 shows the normalized root mean squared (NMSE) values achieved by all the trained models. In general, the results corroborate the analysis made in the synthetic validation study. All training datasets present an increase in performance with the addition of training noise and a region of higher performances around 14%, the estimated signal to noise ratio (SNR) for the data they are applied to. In addition to these, Training datasets 1 and 2 show also other models with comparably good performances with higher training SNR values. Training dataset 4 performs only slightly better than Training dataset 3, considering the orders of magnitude higher than computational cost.

Although each trained model offers a deterministic solution, training multiple models with varying SNR values offers the possibility of entering a statistical mind-set. As we have multiple models that present comparable performances, we can create a more general solution by averaging the results of a few best performance models. Additionally, we can use the standard deviation of these solutions as an estimative to the uncertainties in each estimation. Inversion results presented in the next section are all averages of the 10 best solutions, assessed using this well NMSE, out of the 100 trained models. The models used for average and standard deviation quantifications are marked in Figure 9.

INVERSION RESULTS

To provide a visual representation of the benefits of training our deep neural network (DNN) with noise, in Figure 10 we present the inversion results when the DNNs are trained without noise (Training SNR = 0). Figure 11 shows the final inversion results, corresponding to the average between the solutions of the 10 best models selected in the well data validation study. Zero training signal to noise ratio (SNR) results are extremely noisy for Training datasets 1 and 2. The noise content is lower for Training datasets 3 and 4, results for water and gas saturations are similar to the final inversion results but noisier. It is the pressure results that improve most by adding training noise.

In this section, we interpret the final inversion results (Fig. 11) for each training dataset individually. Interpretations are mostly qualitative, using the circled zones as references. Zones circled in magenta are pressure increase areas and zones circled in green are gas accumulation areas.

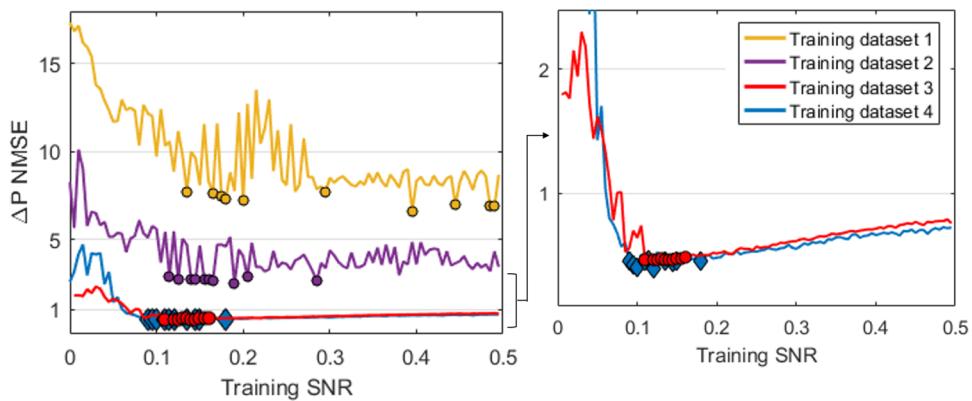


Figure 9 NMSE results for the well data validation.

Training dataset 1

The inversion results for this approach are quite noisy, especially in the gas and the water saturations. This model is incapable of defining well the areas that are impacted by gas saturation, showing very high increases in gas saturation all across the reservoir. Areas dominated by water saturation-related hardening signals are marked reasonably well, but we see a biased background resulting in increases in water saturation across the whole reservoir as well. Furthermore, water saturation values are above the physical limits in many areas, all water saturation values shown in dark green (limit of the colour scale) are above the 85% physical limit. The maximum values for water saturation are of around 1.7, which would mean an increase of 170% of the total porosity in water saturation. Although we can apply constraints to the training data, this does not guarantee that the results will also be constrained, deep neural networks are capable of extrapolating the training data and there is no other way to apply constraints to the inference results. This may lead to unrealistic solutions if the training dataset is inappropriate.

In the pressure results we see well marked pressure increase in all the zones circled in magenta, which is what is expected, but quantitatively the values are far from what is measured at the wells. These areas are where the pressure effects dominate the seismic signal; everywhere else the pressure results are noisy and present an apparent bias towards pressure increase. We also see some leakage from gas saturation-related softening signals into pressure results in zone F. As has been mentioned, in this case, the seismic data are nearly insensitive to pressure depletion. Furthermore, pressure depletion is accompanied by gas breakout, and the seismic data are very sensitive to gas saturation increases. When these two effects are superposed, the much stronger gas effects dominate so that

pressure depletion effects are comparable to noise. The seismic data do not offer any considerable information on pressure depletion (Côte *et al.*, 2019), so it is comprehensible that we see no pressure depletion in the pressure estimations, aside from noise.

Training dataset 2

The use of the constraint in the training dataset has a positive impact on the estimations of changes in pressure and gas saturation (Fig. 11). Pressure results are less biased and leakage in zone F is better controlled. Gas saturation results now mark reasonably well the areas of gas saturation dominance (zones G and F), but we still see a biased background showing gas saturation increases of around 7% across the whole reservoir. Water saturation results are nearly not affected by the constraint, estimations remain biased and values are above the physical limits in many areas.

Interestingly, we now see pressure depletion values in correlation with gas saturation increase in zone G. It is unlikely that these results are based on information from the seismic data so it could be regarded as noise, but as we will see clearly in the results of Training dataset 3 and 4, this is an indication that the network is learning an additional level of information from the training dataset.

Training dataset 3

We can see that having a better representation of reality in the training dataset samples pays off in the quality of the results (Fig. 11). For Training dataset 3 we have much clearer results, noise is contained for all three dynamic properties and we no longer see the general bias present in the water and gas saturation estimations for the previous models, instead, we have

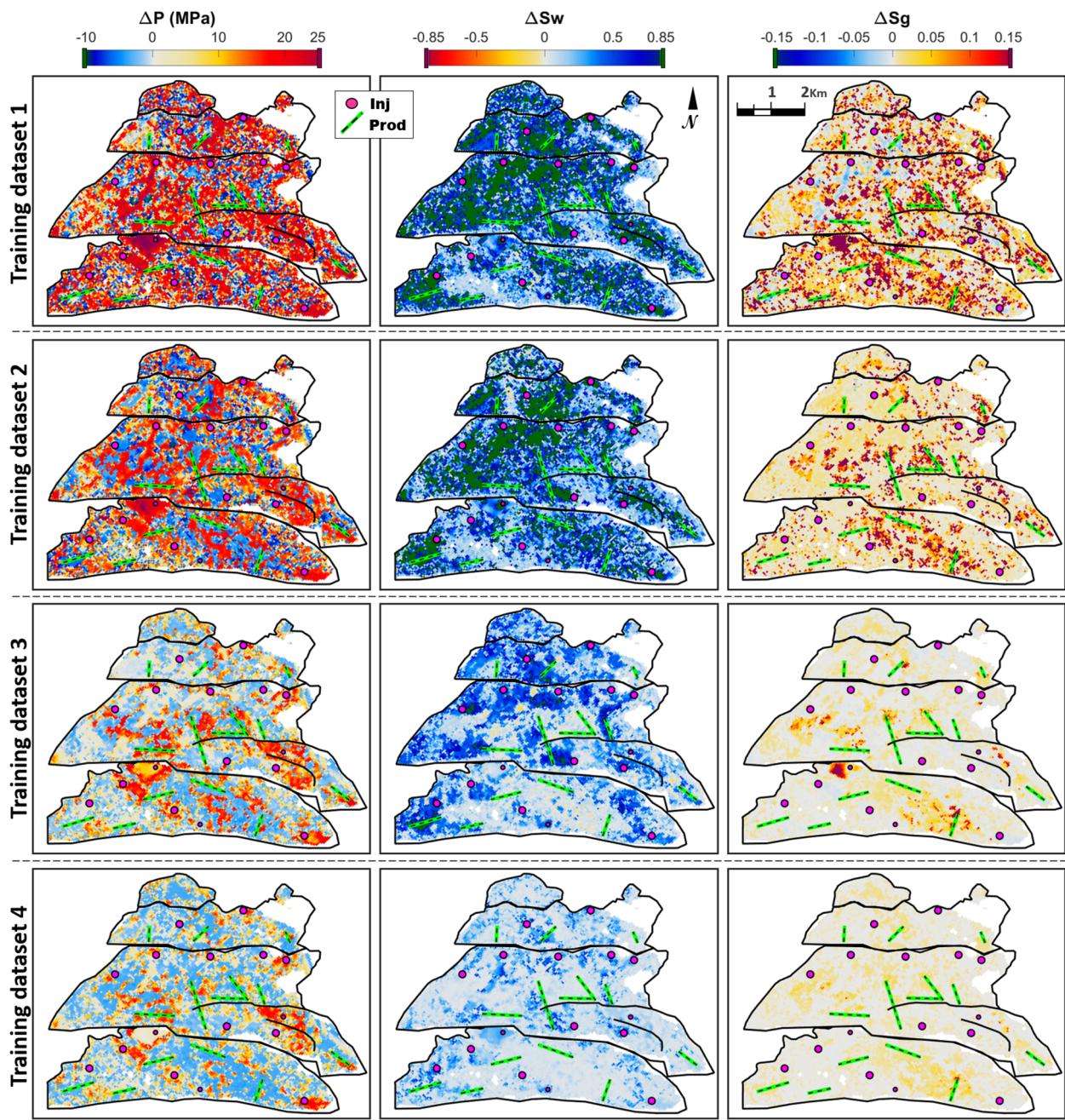


Figure 10 DNN inversion results using training SNR = 0, for training datasets 1, 2, 3 and 4.

a general background of zero values in the areas where there is no seismic evidence of change in each dynamic property. This leads to more contained water estimations forming well-defined bodies that can be connected to the injection wells. Quantitatively, water and gas saturation values are more realistic, falling below the physically possible threshold, except for the water anomalies in the undershoot zone, where the seismic

data are corrupted by low repeatability issues (Fig. 1d). Water front estimations are compromised in some areas where water effects on seismic are dominated by pressure (zones C and E) or gas saturation (zones F and G) effects. On the southwest edge of zone G, there are two injectors that inject a considerable amount of water (as can be seen in the reservoir simulation results in Fig. 4), but we have no seismic evidence of where

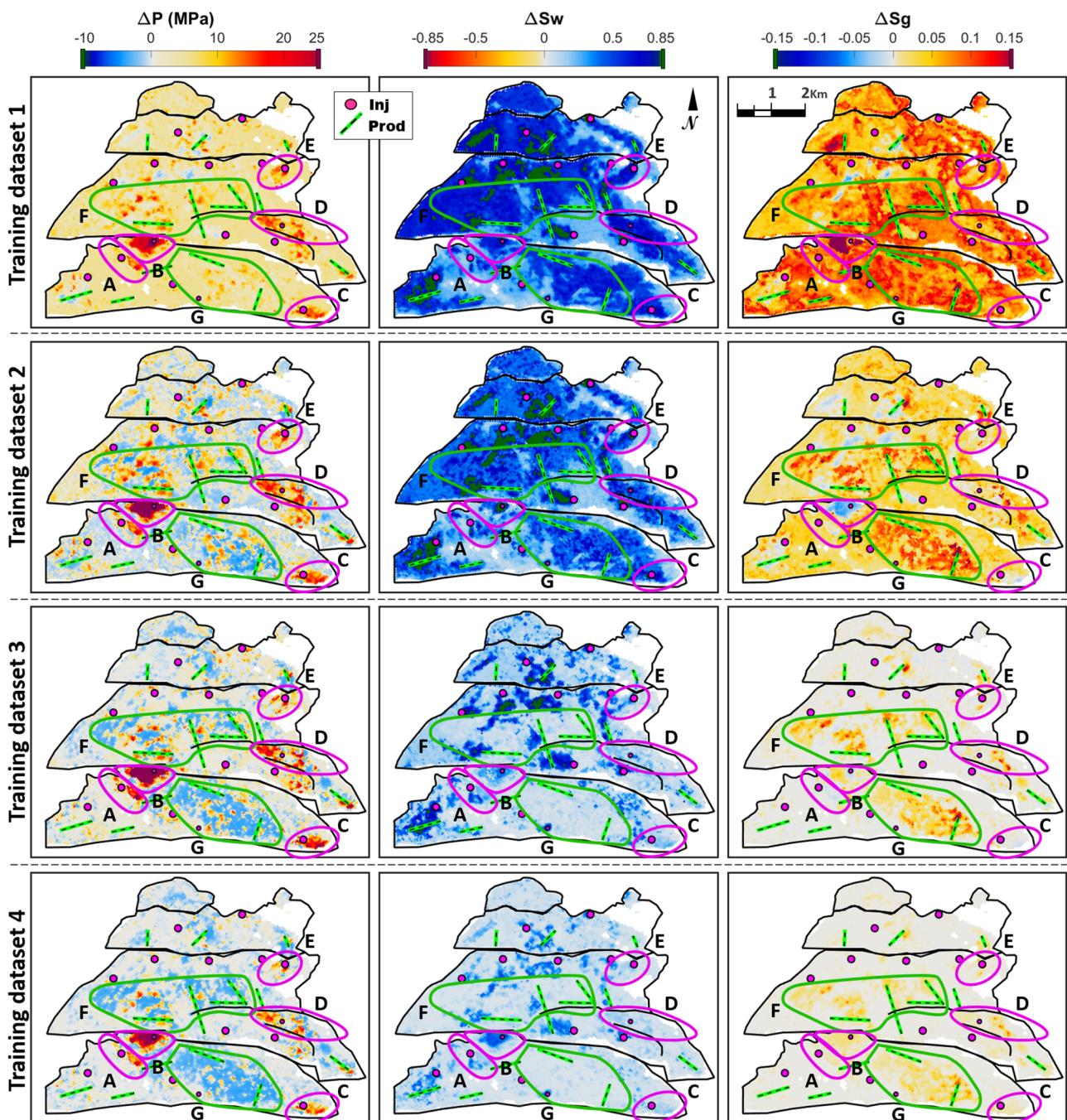


Figure 11 Final DNN inversion results for training datasets 1, 2, 3 and 4. Maps represent the average between the 10 best solutions. Zones circled in magenta are pressure increase areas and zones circled in green are gas accumulation areas.

this water has gone in the inversion results, because the water effects are completely obscured by the gas accumulation here. In zone E, pressure increase effects dominate the seismic response, here the water saturation results are also inconsistent with the amount of water that has been injected into this zone.

Pressure increase is seen in all areas where it dominates the seismic response (magenta circles), furthermore, the quantitative values are more in line with the pressures measured at the wells. Outside these zones we have a global trend of zero values and, interestingly, we see pressure depletion values in

correlation with gas saturation increase (zones F and G). It is clear that this information is not coming from the seismic data. If any seismic information related to pressure depletion could be detected, it would be away from those gas accumulation zones, where the pressure effect is not so overwhelmingly dominated. Instead, the inversion results show pressure depletion values only in direct correlation with the gas saturation increase. This pressure depletion – gas saturation increase correlation is present in the training dataset, as gas exsolutions is one of the physical processes modelled by the reservoir simulation. This is a clear indication that the deep neural network (DNN) is learning the correlations present in the training dataset and using this information to make inferences on the dominated properties, where the seismic data cannot provide useful information. We also see a curious gas saturation decrease accompanying the pressure increase in zone B. As there is no initial gas in the reservoir, gas saturation decrease is not present in the training datasets. It does represent a possible reality between monitors, but as we use a pre-production baseline, negative changes in gas saturations are not realistic. We see this result as an indication that the DNN is not only learning the correlation between pressure depletion and gas saturation increase, but also extrapolating this correlation to result in gas saturation decrease in pressure increase zones. This extrapolation is undesirable, as it leads to non-physical results, but it cannot be prevented.

Knowing that the DNN learns not only relations that link input to output but also the correlations between variables in the training dataset emphasizes the argument that the dynamic domain sampling used for creating the training dataset should be retained to physically realistic combinations, because the DNN will embed this physical knowledge into the non-linear transformations it learns. It is also important to note that the data size of this training dataset is much smaller than in the previous models, which means training runtimes are also much faster (Table 1). In this case, it is not necessarily the size of the dataset that matters most, instead the ability of this dataset to represent specifically the global reality of the problem is more important.

Training dataset 4

Augmenting the data size in this approach affected the results only slightly (Fig. 11), but nonetheless the results are generally more consistent. In the pressure results, we see less noise and leakage in the areas dominated by gas (zones F and G). Water saturation results are now below the physical limit everywhere in the map, though the shape of all the water bodies

is very similar to the previous case and the definition of water-fronts does not improve in areas of overwhelming dominance by other properties (zones B, C, D, F and G). Gas saturation results are similar but generally smaller than with Training dataset 3 and zone B no longer shows unrealistic gas saturation decrease values. Given the uncertainty in the estimations, the uncertainty in the reservoir simulation results in this property and the lack of measured saturation logs, it is impossible to define which gas saturation solution is more precise.

Globally the results using Training dataset 4 are slightly better than in model 3, but this comes at a high computational cost. In Table 1, we see that Training dataset 3 contains the smallest data size and quickest runtimes and nonetheless it performs much better than Training datasets 1 and 2, and nearly as well as Training dataset 4, which takes 15 times longer to train. It is also relevant to consider the time it takes to compute the synthetic seismic data to build these training datasets, which is around 100 times longer in Training dataset 4 than in 3.

From this analysis, we see the critical importance in constraining the realizations of our synthetic training dataset to realistic physics informed and fluid flow consistent combinations that represent the specific problem at hand. This constraining will make the model less general, so it should not be applied to a different case that may not contain the same constraint assumptions (e.g. gas injection, gas caps and reservoir compaction), instead it will be more specialized to provide the best results to one specific case.

UNCERTAINTIES

The presented deep neural network workflow is essentially a deterministic solution, but as discussed, we could produce multiple slightly different but equally viable solutions varying the signal to noise ratio (SNR) parameter for training. We can use these multiple solutions to create a simple estimation to uncertainties in the results. Uncertainties here represent the instability in the solutions with varying SNR values. Figure 12 shows maps of the standard deviations of the selected 10 best solutions for all four training datasets.

For Training dataset 1, uncertainty results offer little useful information. From Training dataset 1 to 4 gradually the uncertainties decrease globally, but some patches stand out with high uncertainty values. For the water saturation, uncertainties tend to be higher in the pressure increase zones, and around the high normalized root mean squared (NRMS) zone, which is represented well in the uncertainty results. On the other hand, uncertainty results do not represent well the

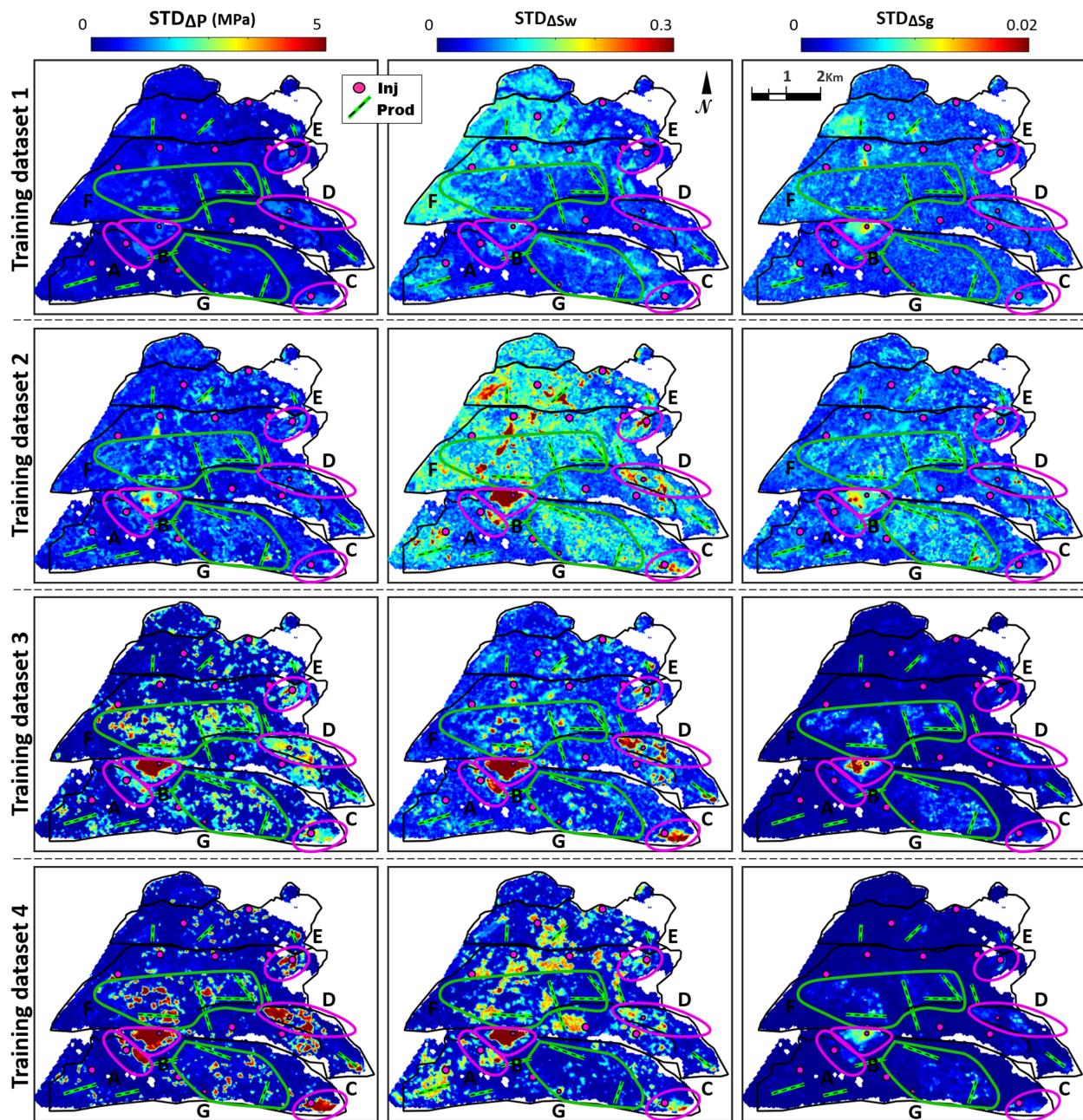


Figure 12 Uncertainty estimations for the four training datasets.

uncertainties expected in areas where water saturation effects are dominated by gas saturation effects (Zones D and E). Uncertainty results for pressure seem to show a direct correlation with the pressure values themselves. This is the opposite of what is expected, as seismic data become more sensitive to pressure as pressure increases. It is unclear what value can be brought from these types of uncertainty estimations, as they do not represent seismic or modelling uncertainties. One thing

does stand out though. Uncertainty values for all three properties highlight zone B as a high uncertainty area.

Zone B is an isolated compartment that was pressurized due to water injection. The injector well was online for around three years and measured pressure increases as high as 20 MPa in 2004, when the seismic acquisition was shot. A previous feasibility study in the area indicated that in order to reach the observed seismic amplitudes, the pressures would need to

be above the estimated rock fracture limit. There are many assumptions in the current petro-elastic model that make it unfit for representing the elastic behaviour of a fractured rock. If the rock has indeed been fractured due to high injection pressures, this means the petro-elastic model used in the creation of the training dataset cannot represent well the rock physics of this zone. Thus, in zone B we have an example of the use of an inadequate training dataset, constructed with an improper rock physics model. This results in solutions that are more unstable with small changes in the noise parameter, which reflects as high uncertainty values. So the uncertainty estimations here were useful for identifying a zone where the synthetic training data are unfit and thus inversion results are not trustworthy.

CONCLUSIONS

The present study has shown that deep neural networks (DNNs) trained exclusively using synthetic data can provide good solutions to the problem of inverting time-lapse seismic data to the simultaneous changes in pressure, water saturation and gas saturation. We show clear evidence of the benefit of adding noise to the synthetic data in the training phase to achieve less noisy and more accurate estimations. Training multiple models with varying training signal to noise ratio (SNR) values and assessing model performance using well measured data offers a possibility of selecting multiple equally probable solutions to create a more general average result. This may also lead to an estimation of the related uncertainties, by calculating the standard deviation of the solutions provided by the selected models. Uncertainty estimations here represent instability in the solutions with respect to the noise parameter and may indicate areas where the training dataset is inadequate.

We show the critical importance of using physics informed sampling of the dynamic domain in creating the training dataset, illustrated by the gradual increase in inversion quality from Training datasets 1 to 3. From datasets 1 to 2, we see that adding a simple external physical knowledge to constrain the samples has a positive impact on the solutions. Using the results of a three-dimensional reservoir flow simulation as the only samples in the dynamic domain (Training datasets 3 and 4) guarantees that the training dataset respects all physical processes modelled by the reservoir simulation. Training datasets then contain physical boundary constraints, physical correlations between dynamic domain parameters and global distributions that resemble the reality of the reservoir. This removes bias in the results and prevents the model from extrapolating beyond the training dataset and leading to extreme un-

realistic results. Solutions also become more stable, less noisy and more precise.

Additionally, we observe that when trained using fluid flow consistent data, the DNN learns not only the relations that link input to output, but also the correlations present in the training dataset, making use of those correlations to make inferences on dominated properties, where the seismic data lack information. This is observed in this case in the correlations between pressure depletion and gas saturation increase. Learning these correlations allows the DNN to resolve some ambiguity present in the seismic data, resulting in better solutions both in the pressure and in the gas saturation results.

We show that, in the present application, an ideal training dataset is one that resembles the most what an unbiased measured dataset would be, both in the sense of containing all the physical correlations in the dynamic domain and also maintaining a realistic global distribution on all related properties. This makes the DNN model less generalized, more specific to the problem at hand, so it should not be applied to describe situations that do not respect the constraints used in the training dataset.

Augmenting the data size while maintaining physical constraints achieved slightly better solutions, improving noise content and leakage in the results. This comes at a high computational cost though. In this application, it seems more important to constrain the sample realizations to physically informed and fluid flow consistent combinations than to chase large data sizes.

ACKNOWLEDGEMENTS

We thank the sponsors of the Edinburgh Time-Lapse Project, Phase VII (AkerBP, BP, CGG, Chevron, ConocoPhillips, ENI, ExxonMobil, Hess, Landmark, Maersk, Nexen, Norsar, OMV, PGS, Petrobras, Shell, Equinor, Woodside and Taqa) for supporting this research and Schlumberger for providing Eclipse software. This work was financed in part by the Brazilian National Research Council, CNPq (200014/2016-1) and the Danish Hydrocarbon Research and Technology Centre under the Advanced Water Flooding program. This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF). We thank Linda Hodgson and Ross Walder for informative discussions. The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Gustavo Côrte  <https://orcid.org/0000-0002-5063-2036>

Jesper Dramsch  <https://orcid.org/0000-0001-8273-905X>
 Hamed Amini  <https://orcid.org/0000-0001-9588-6374>
 Colin MacBeth  <https://orcid.org/0000-0001-8593-3456>

REFERENCES

- Aki, K. and Richards, P. (1980) *Quantitative Seismology: Theory and Methods V1*. New York, NY: W.H. Freeman & Co.
- Alvarez, E. and Macbeth, C. (2014) An insightful parametrization for the flatlander's interpretation of time-lapsed seismic data. *Geophysical Prospecting*, 62, 75–96.
- Amini, H. (2014) *A pragmatic approach to simulator-to-seismic modelling for 4D seismic interpretation*. PhD Thesis, Heriot-Watt University.
- Amini, H. (2018) Comparison of Xu-White, simplified Xu-White (Keys and Xu) and Nur's critical porosity in shaly sands. 80th EAGE Annual meeting, Copenhagen, Denmark, Extended Abstracts, We A11 08.
- Amini, H. and MacBeth, C. (2015) Calibration of rock stress-sensitivity using 4D seismic data. 77th EAGE Annual meeting, Madrid, Spain, Extended Abstracts, We A11 08.
- Ayzenberg, M. and Liu, S. (2014) Saturation and pressure inversion - from 4D seismic to reservoir model updating. 76th EAGE Annual meeting, Amsterdam, Netherlands, Extended Abstracts, We G102 16.
- Bergstra, J., Yamins, D. and Cox D.D. (2013) Making a science of model search: hyperparameter optimization in hundreds of dimensions for vision architectures. *30th International Conference on Machine Learning, ICML, Atlanta, Georgia, USA*, 115–123.
- Bishop, C.M. (1995) Training with noise is equivalent to Tikhonov regularization. *Neural Computation*, 7, 108–116.
- Blanchard, T.D. (2012) Inherent uncertainties in 4D AVO and the implications on pressure and saturation inversion. 74th EAGE Annual meeting, Copenhagen, Denmark, Extended Abstracts, Y2020.
- Blanchard, T.D. and Thore, P. (2008) Introducing prior information to pressure and saturation inversion: the key for success? 83rd SEG Annual meeting, Houston, Texas, USA, Extended Abstracts, 4971–4975.
- Cao, J. and Roy, B. (2017) Time-lapse reservoir property change estimation from seismic using machine learning. *The Leading Edge*, 36, 234–238.
- Coleou, T., Roustiau, A., Machecler, I., Ayzenberg, M., Fayemendy, C., Skjei, N., et al. (2013) 4D Petrophysical seismic inversion - case studies. 75th EAGE Annual meeting, London, UK, Extended Abstracts, We 17 03.
- Côte, G., MacBeth, C. and Amini, H. (2019) North Sea field application of 4D Bayesian inversion to pressure and saturation changes. 81st EAGE Annual meeting, London, UK, Extended Abstracts, Tu_P04_09.
- Corzo, M., Macbeth, C. and Barkved, O. (2013) Estimation of pore-pressure change in a compacting reservoir from time-lapse seismic data. *Geophysical Prospecting*, 61, 1022–1034.
- Davolio, A., Maschio, C. and Schiozer, D. (2013) A methodology to constrain pressure and saturation estimation from 4D seismic using multiple simulation models and observed data. *Journal of Petroleum Science and Engineering*, 105, 51–61.
- Dramsch, J., Corte, G., Amini, H., Lüthje, M. and MacBeth, C. (2019a) Deep learning application for 4D pressure saturation inversion compared to Bayesian inversion on North Sea Data. Second EAGE Workshop Practical Reservoir Monitoring, We PRM 11.
- Dramsch, J., Corte, G., Amini, H., MacBeth, C. and Lüthje, M. (2019b) Including physics in deep learning - an example from 4D seismic pressure saturation inversion. 81st EAGE Annual meeting, London, UK, Extended Abstracts, WS10_05.
- Ebdon, C.C., Granger, P.J., Johnson, H.D. and Evans, A.M. (1995) Early Tertiary evolution and sequence stratigraphy of the Faeroe-Shetland basin: implications for hydrocarbon prospectivity. *Geological Society Special Publication*, 90, 51–69.
- Floricich, M. (2006) *An engineering-consistent approach for pressure and saturation estimation from time-lapse seismic data*. PhD Thesis, Heriot-Watt University.
- Gassmann, F. (1951) Über die Elastizität poröser Medien. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, 96, 1–23.
- Kingma, D.P. and Ba, J. (2015) Adam: a method for stochastic optimization. *International Conference for Learning Representations, San Diego, California, USA*.
- Kingma, D.P. and Welling, M. (2014). Auto-encoding variational Bayes. *Proceedings of International Conference for Learning Representations, Banff, Canada*.
- Kingma, D.P., Salimans, T. and Welling, M. (2015) Variational dropout and the local reparameterization trick. *Advances in Neural Information Processing Systems*, 28, 2575–2583.
- Lamers, E. and Carmichael, S.M.M. (1999) The Paleocene deepwater sandstone play West of Shetland. *Petroleum Geology Conference series*, 5, 645–659.
- Landrø, M. (2001) Discrimination between pressure and fluid saturation changes from time lapse seismic data. *Geophysics*, 66, 836–844.
- MacBeth, C. (2004) A classification for the pressure-sensitivity properties of a sandstone rock frame. *Geophysics*, 69, 497–510.
- MacBeth, C., Floricich, M. and Soldo, J. (2006) Going quantitative with 4D seismic analysis. *Geophysical Prospecting*, 54, 303–317.
- Nur, A., Mavko, G., Dvorkin, J. and Galimudi, D. (1998) Critical porosity: a key to relating physical properties to porosity in rocks. *The Leading Edge*, 17, 357–362.
- Omomofoma, V. (2017) *The quantification of pressure and saturation changes in clastic reservoirs using 4D seismic data*. PhD Thesis, Heriot-Watt University.
- Omomofoma, V., MacBeth, C. and Amini, H. (2019) Intra-survey reservoir fluctuations – implications for quantitative 4D seismic analysis. *Geophysical Prospecting*, 67, 282–297.
- Smith, G.C. and Gidlow, P.M. (1987) Weighted stacking for rock property estimation and detection of gas. *Geophysical Prospecting*, 35, 993–1014.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R. (2014) Dropout: a simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15, 1929–1958.

- Thore, P. and Hubans, C. (2012) 4D seismic-to-well tying, a key step towards 4D inversion. *Geophysics*, 77(6), R227–R238.
- Trani, M., Arts, R., Leeuwenburgh, O. and Brouwer, J. (2011) Estimation of changes in saturation and pressure from 4D seismic AVO and time-shift analysis. *Geophysics*, 76(2), C1–C17.
- Wong, M.Y. (2017). *Pressure and saturation estimation from prm time-lapse seismic data for a compacting reservoir*. PhD Thesis, Heriot-Watt University.
- Xue, Y., Araujo, M., Wang, K., Lopez, J., Kumar, G. and Brew, G. (2019) Machine Learning to reduce cycle time for time-lapse seismic data assimilation into reservoir management. 81st EAGE Annual meeting, London, UK, Extended Abstracts, We_R07_08.
- Zhong, Z., Sun, A.Y. and Wu, X. (2020) Inversion of time-lapse seismic reservoir monitoring data using cycleGAN: a deep learning-based approach for estimating dynamic reservoir property changes. *Journal of Geophysical Research: Solid Earth*, 125, e2019JB018408.