Assessment of predictive learning methods for the completion of gaps in well log data

Rui L. Lopes [a,*], Alípio M. Jorge [a,b]

[a] INESC TEC, Porto, Portugal
[b] DCC - FCUP, Universidade Do Porto, Portugal

A B S T R A C T

Well logs are records of petro-physical data acquired along a borehole, providing direct information about what is in the subsurface. The data collected by logging wells can have significant economic consequences in oil and gas exploration, not only because it has a direct impact on the following decisions, but also due to the subsequent costs inherent to drilling wells, and the potential return of oil deposits. These logs frequently present gaps of varied sizes in the sensor recordings, that happen for diverse reasons. These gaps result in less information used by the interpreter to build the stratigraphic models, and consequently larger uncertainty regarding what will be encountered when the next well is drilled.

The main goal of this work is to compare Gradient Tree Boosting, Random Forests, Artificial Neural Networks, and three algorithms of Linear Regression on the prediction of the gaps in well log data. Given the logs from a specific well, we use the intervals with complete information as the training data to learn a regression model of one of the sensors for that well. The algorithms are compared with each other using a few individual example wells with complete information, on which we build artificial gaps to cross validate the results. We show that the ensemble algorithms tend to perform significantly better, and that the results hold when addressing the different examples individually. Moreover, we performed a grid search over the ensembles parameters space, but did not find a statistically significant difference in any situation.

1. Introduction

The goal of the exploration activities is to identify commercially viable reserves of oil and gas (McNamara and McKenzie-Brown, 2013). For such reserves to have accumulated, a set of complex conditions is necessary that is largely dependent on past geological history, and present geological formations and structures. Particular combinations of potential source and reservoir rocks, together with migration pathways and trap structures, are needed for the hydrocarbon deposits to occur. Finding these reservoirs and estimating the likelihood of finding oil and gas is a technically complex process requiring the use of different techniques, including seismic surveys, drilling, coring, aero-magnetic/gravity surveys, and exploration and appraisal drilling.

Typically several logs are collected in each well, that complement each other's information. The data readings are usually acquired through sensing tools that are lowered into the hole by cable. The data acquired by each sensor is logged (recorded) at the surface as the tool is pulled up the hole. During the well drilling it is also possible to acquire data using instruments at the bottom of the driller/hole. The term "well log data" is then used to refer to data collected in or descriptive of the rock(s) and fluid(s) surrounding the hole.

These detailed and indirect measurements of rock and fluid properties in the subsurface typically include i) gamma ray intensity (related to the types of minerals present); ii) electrical resistance (related to the quantity and types of fluids); iii) density and iv) porosity (related to the pore-volume fraction); and v) sonic velocity, also referred to as transit time which relates to both rock and fluid properties. These are common examples found in the composite well logs of the North Sea, and merely illustrate the hundreds of well logs that may be collected.

The collected data frequently present gaps in the sensor recordings. Well log gaps result in less information on which to base the geoscientific models, and consequently larger uncertainty regarding what will be encountered when the next well is drilled. These gaps happen for different reasons, from tools that fail or malfunction, to operator errors (for instance, an operator may turn off recording equipment at the wrong time, or it may be discovered later that the wrong interval was logged). It

---

R.L. Lopes, A.M. Jorge

Table 1
Data set sample with five records. Each record includes the well Name, Latitude, and Longitude, as well as the normalised values for the Depth (m) of the readings, the bulk density (RHOB, g/c^3), the sonic log (DT, $\mu s/ft$), the gamma-ray (GR, $gAPI$), and the neutron porosity (NPHI, v/v).

Well	Depth	RHOB	DT	GR	NPHI	Latitude	Longitude
F04-02-A	0.57	0.39	0.19	0.09	0.19	0.804	0.512
F16-03	0.72	0.18	0.18	0.09	0.43	0.801	0.511
F12-01	0.33	0.36	0.32	0.12	0.42	0.803	0.513
F02-05	0.20	0.28	0.29	0.047	0.25	0.805	0.513
F09-03	0.63	0.44	0.19	0.067	0.24	0.804	0.513

is the main objective of this article to provide a methodology for completing the gaps in well log data, and compare the performance of different models on this prediction task.

Completing gaps in well logs can be interpreted either as an imputation or a prediction task. The former would be the case for logs where the gaps are frequent and small (low number of contiguous missing data points). The latter is found when the gaps consist of a large number of contiguous data points. In order to understand the characteristics of the problem in hands, the gaps present in the logs of the sensors of each of the 1026 wells in the North Sea are first identified. The frequency and size of the holes across the data are analysed, providing a descriptive and exploratory analysis of the gaps previously identified. The goal of this analysis is to enable the generation of artificial gaps that reflect the statistics collected before and thus constitute a relevant synthetic problem on which to train and test the predictors.

The typical approach to the forecast of log values uses data from a single well (or a few wells from the same block), typically surrounding a particular gap (see the next Section for particular examples). Here a similar strategy is adopted, by using all the data available from a single well (excluding the gap). Different techniques can be employed to predict the missing data. In this work Ordinary Least Squares (OLS), Bayesian Regression (BRR), RANdom SAmple Consensus (RANSAC), Artificial Neural Networks (ANN), Random Forests (RF), and Gradient Tree Boosting (GB) are compared on the prediction of missing gaps, on a well-by-well basis. In particular, a single complete well is used to generate realistic artificial gaps, and the models are trained to predict those values from the remaining sensor data. Therefore, the main contributions of this work are i) a methodology to build models for sensors based on a complete case analysis, using all the information from the well; ii) a cross-validated study of the predictors performance along a complete well; and iii) a comparison of the baseline linear models and ANNs (a popular non-linear tool in the area) with recent non-linear ensemble models (RF

and GB). To the best of the authors' knowledge the performance of the ensemble methods has not been assessed or compared before on a similar task of predicting gaps in well log data.

Our experimental work shows that the ensemble models outperform the remaining approaches with statistical significance, and that these results can be replicated across different individual wells. Further more we have not found a significant difference between the ensemble models, even after fine-tuning both methods using a grid search over the parameters space.

The next section provides the motivation for addressing this problem, and presents and summarises the recent related work.

2. Related work

Different techniques have been used to build models that predict and estimate various geophysical properties, by means of either regression or classification tasks including, but not limited to: deterministic petrophysical modelling, using shale, matrix, and fluid properties; stochastic modelling, where an approximate curve is used as input, and the reconstructed curve is the output; and different soft-computing algorithms (Aminzadeh and De Groot, 2006; Holdaway, 2014).

The typical approach to prediction tasks based on well logs uses data from a single well (or a few wells from the same block) as input, in order to extrapolate to a neighbouring well, or to a specific depth range in the same well (Holmes et al., 2003; Ayoub and Mohamed, 2014; Ahmadi et al., 2014). The prediction tasks may target for instance the estimation of the rock permeability and porosity (Aifa et al., 2014; Zerrouki et al., 2014) avoiding further core sampling, or the identification of the lithofacies types (Mohseni et al., 2015) to improve the stratigraphical analysis.

Table 2
Descriptive statistics of the gaps identified in all the well logs for each sensor, respectively the sonic log (DT), the gamma-ray (GR), the neutron porosity (NPHI), and the bulk density (RHOB). The values (except the count) represent depth (or gap size) in meters.

Sensor	DT	GR	NPHI	RHOB
count	1651.00	951.00	337.00	746.00
mean size (m)	59.14	26.86	144.76	97.31
std (m)	204.16	133.43	369.86	301.11
min size (m)	0.30	0.30	0.30	0.30
25% (m)	2.40	1.25	2.70	1.90
50% (m)	7.20	2.00	8.50	5.60
75% (m)	17.70	6.30	61.50	21.90
max size (m)	2228.60	1765.40	3361.50	3356.00
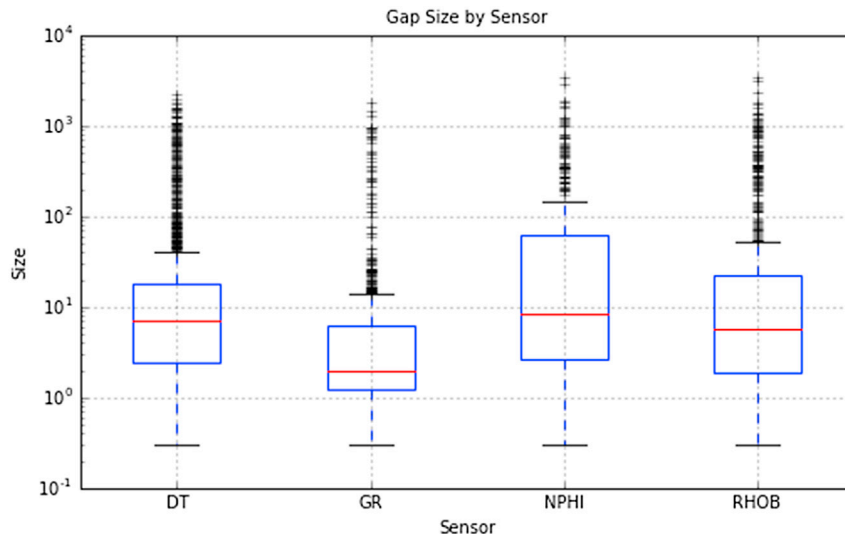
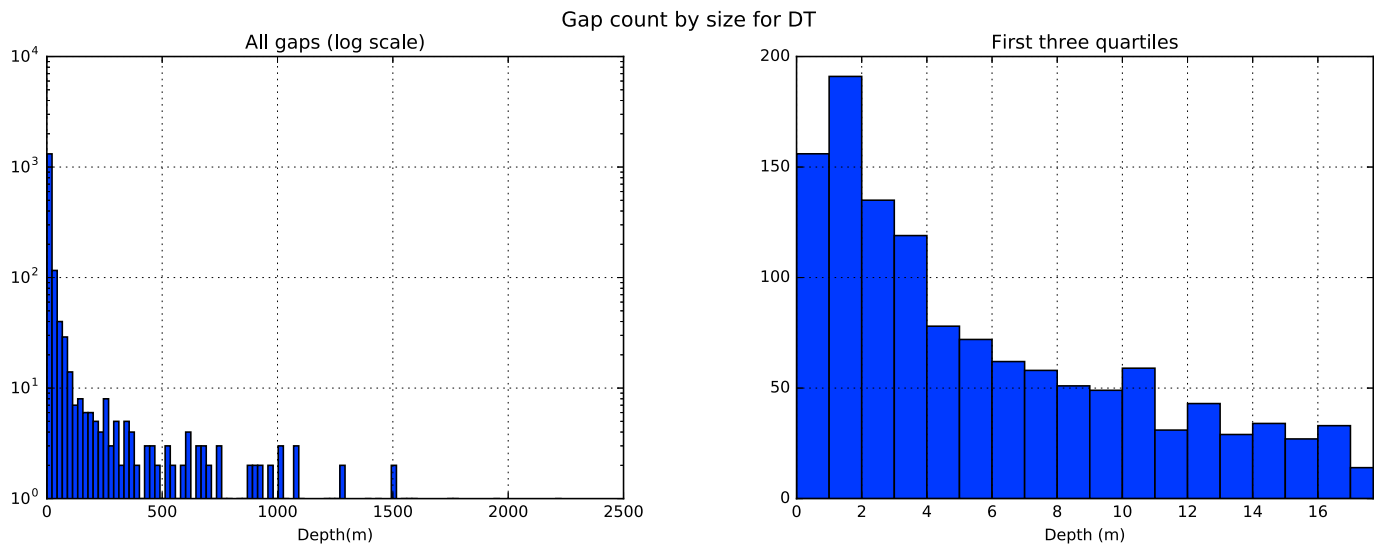**Fig. 1.** Distribution of the gap sizes by sensor.

Fig. 2. Gap count for the transit time (DT). Left: Gap count by size for all the gaps, using the logarithmic scale; Right: Zoom over the first three quartiles, using the original scale.
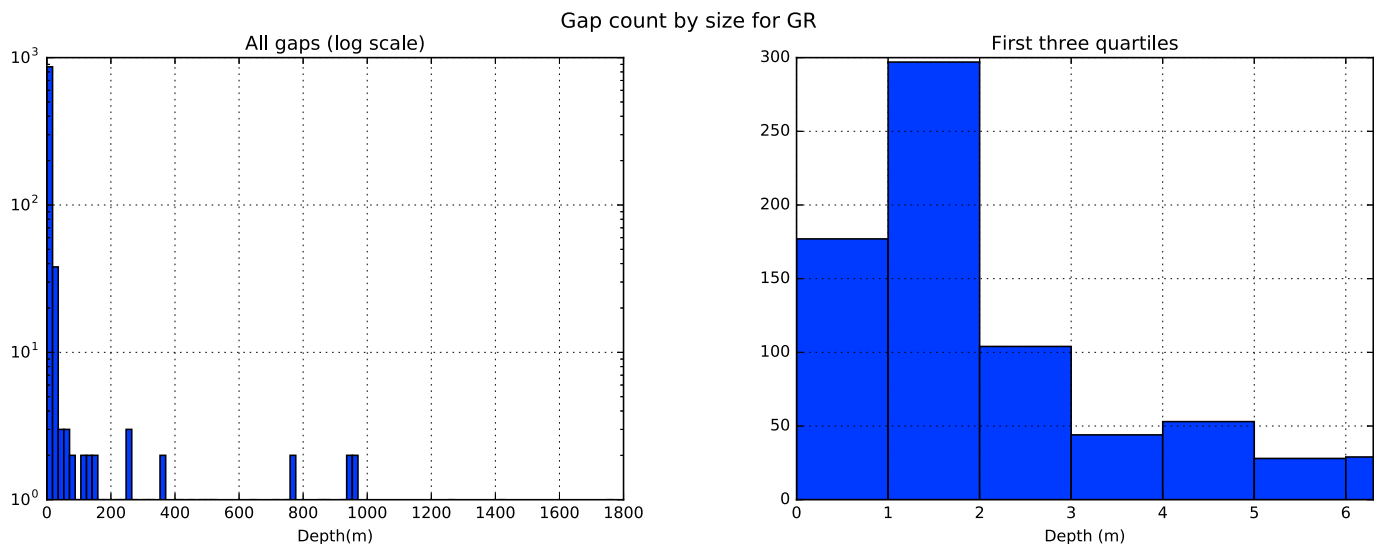

Fig. 3. Gap count for the gamma-ray (GR). Left: Gap count by size for all the gaps, using the logarithmic scale; Right: Zoom over the first three quartiles, using the original scale.

The existence of recent patents on how to predict some predetermined well log based on the information collected by the remaining sensors strongly motivates our approach (Yu et al., 2011). Nevertheless, the need for solutions to these types of tasks is well documented by other researchers targeting a variety of well logs and using different algorithms and models to capture the correlations between the logs. As an example, in the following paragraphs we present some recent works that target the comparison of different algorithms on the estimation of well log data and/or log-derived data.

Cranganu and colleagues (Cranganu and Bautu, 2010; Cranganu and Breaban, 2013) addressed the estimation of sonic log distributions in a field of the Anadarko basin, in Oklahoma. Their approach uses as input the gamma-ray and deep resistivity logs. Two different methods were applied to the same problem, first using Gene Expression Programming to evolve predictors, while in the most recent publication the authors used Support Vector Regression.

Singh et al. (2016) used artificial neural networks to predict the effective porosity in a gas field at Garden City, Kansas. The ANN was trained with the back-propagation algorithm, using the sonic, density, and resistivity log data as input. The authors argue that the ANN does not need prior knowledge of petrophysical properties (pore fluid type, matrix material type). However, the *density* used as input is actually the *porosity density* which indirectly includes information about the rock matrix and the pore fluid density.

Rafik and Kamel (2016) partitioned the well logs using an electrofacies classifier, based on principal component analysis, clustering, and discriminant analysis. Then, the authors used non-parametric regression techniques to predict permeability within each electrofacies. Three approaches were compared: generalised additive models, alternating conditional expectations, and neural networks. They show that the electrofacies characterization results in models that are more robust than the alternatives.

Other approaches (Soubotcheva and Stewart, 2004; Mojeddifar et al., 2014) employ and correlate seismic attributes with well log data in order to improve the prediction of the natural porosity. Amongst the most important attributes for log prediction the authors include the energy, envelope, spectral decomposition, similarity, AVO intercept, and seismic velocity. The methods compared include different clustering algorithms, neural networks, neuro-fuzzy systems, and pseudo-forward equations.

As one can see, typically the input data is gathered from a small number of sensors, which choice varies with each study-case. The cases where there are more attributes (features) being used as input are those
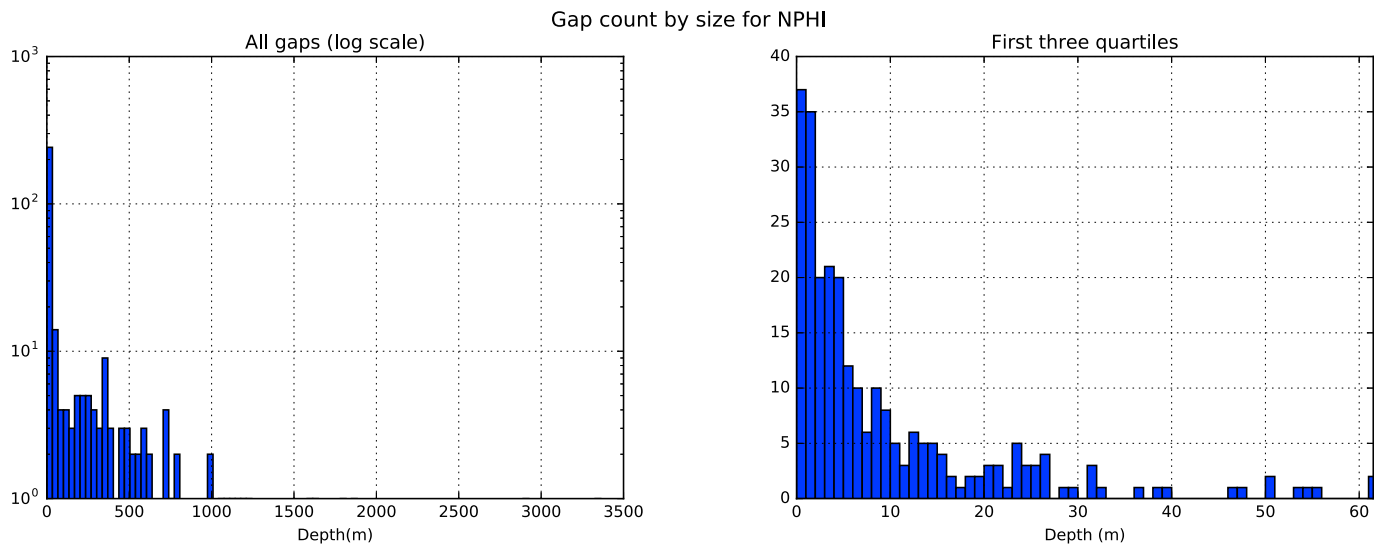
R.L. Lopes, A.M. Jorge

Gap count by size for NPHI

Fig. 4. Gap count for the neutron porosity (NPHI). Left: Gap count by size for all the gaps, using the logarithmic scale; Right: Zoom over the first three quartiles, using the original scale.
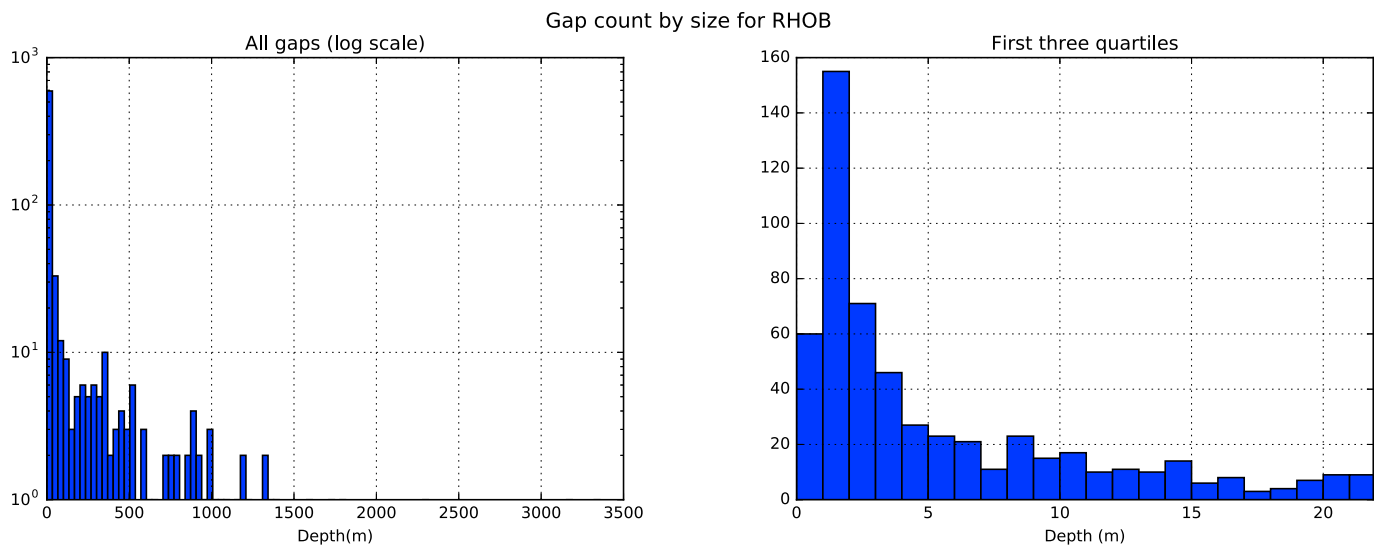
Gap count by size for RHOB

Fig. 5. Gap count for the bulk density (RHOB). Left: Gap count by size for all the gaps, using the logarithmic scale; Right: Zoom over the first three quartiles, using the original scale.

that integrate seismic information as well. Concerning the goals, it is more frequent to address the estimation of derived properties (natural porosity, permeability, etc) than the logs themselves. The former automates part of the interpretation work-flows, by correlating well log data with core samples, and human expert knowledge. The latter focus on estimating the values for the continuous gaps in the logs, providing better information for the following steps in the interpretative work-flows.

These works are examples that motivate our effort to address the prediction of well log data, and illustrate the relevance of the problem for the practitioners. Moreover, they show the variety of techniques that have been applied to the problem, as well as the opportunity to apply recent learning algorithms such as the ensemble methods presented in Section 4.

3. Data description

In order to gain a better understanding of the problem of filling-in log gaps, an exploratory analysis of the gaps contained in the logs of dozens or hundreds of wells is necessary before the forecast of sensor values is conducted.

The data collected from the various sensors at some particular well is

typically subject to some corrections and merged into a composite log. This log includes the meta-data for the well, and the readings for each sensor aligned by depth (using hard-coded out-of-range values for the detected missing values and erroneous readings). The data is typically reported in a graphical format with the sensor readings side by side for easy visual inspection. In order to have a realistic descriptive analysis of the typical well log gap, we collected the composite logs from the off-shore Dutch wells in the North Sea, summing up to a total of 1033 wells.[1] These logs are furnished by the Dutch government and are freely available at the NLOG site, which is managed by TNO, Geological Survey of the Netherlands.

The most common logs found in these records were: i) *bulk density*, which relates to the seismic velocity of waves travelling through the medium; ii) *sonic logs*, which give a measure of the formations capacity to transmit seismic waves, i.e., the formations interval transit time; iii) *gamma ray*, which is based on the correlation between the radioactive isotope content and mineralogy; and iv) *neutron porosity*, which tracks the

---

[1] The data and meta-data for each well can be consulted at http://www.nlog.nl/en/listing-boreholes.
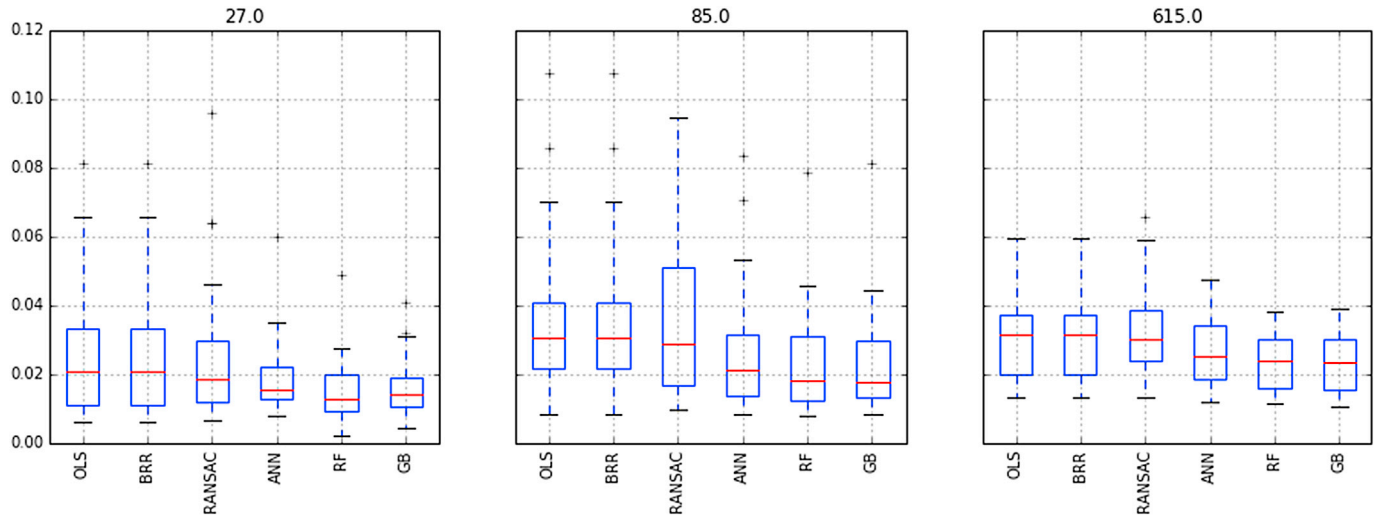
Fig. 6. Mean absolute error for each algorithm by gap size, over all folds. The values in the y-axis indicate percentages in the interval $[0, 1]$. The labels on the x-axis indicate the algorithm used, respectively Ordinary Least Squares, Bayesian Ridge Regression, Random Sample Consensus, Artificial Neural Network, Random Forest, Gradient Boosting.
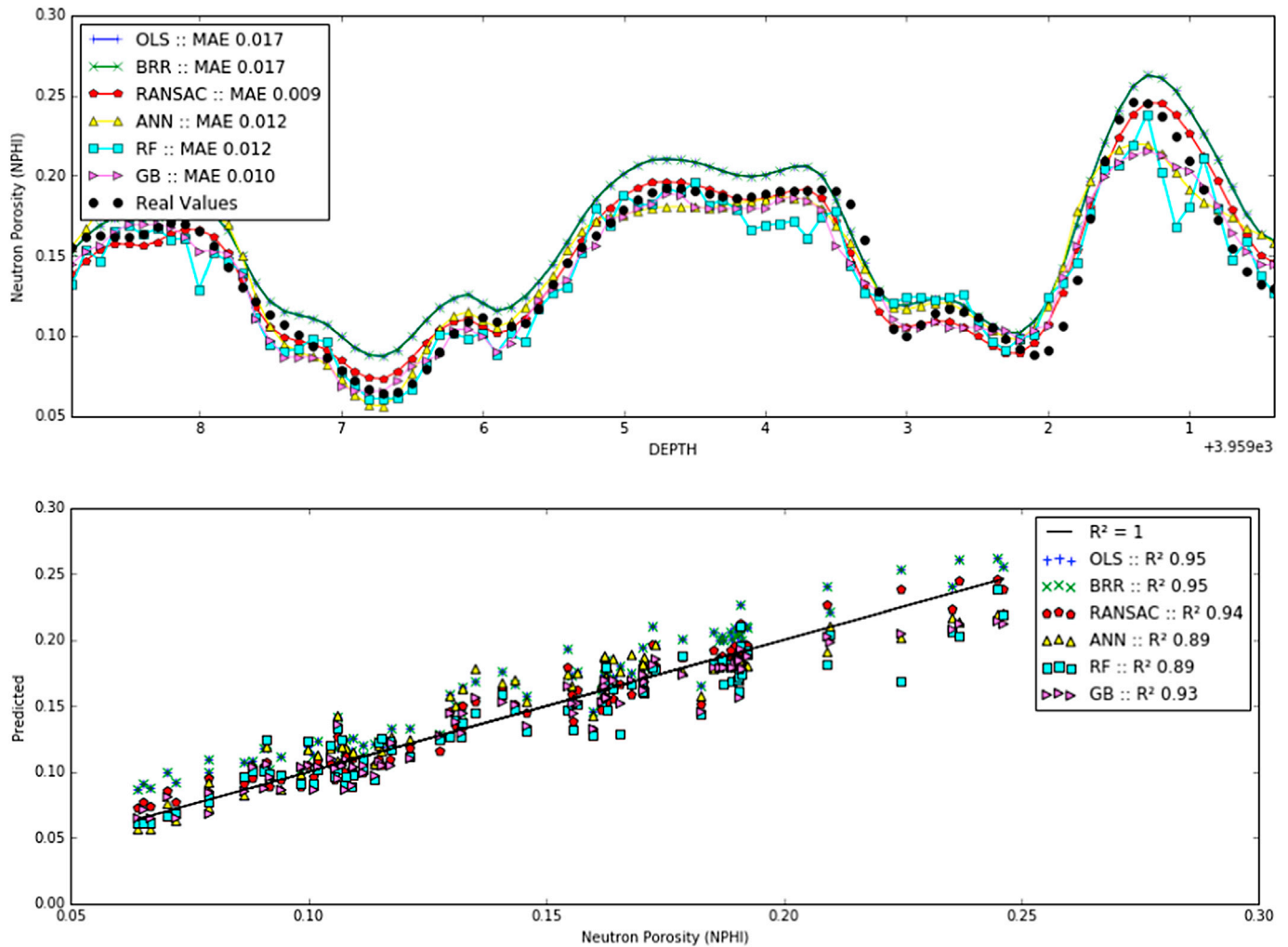
Fig. 7. Example for a fold of the second quartile where the models performed well. Top: Comparison of the predictions for each algorithm with the real values. Bottom: Scatter plot of the predictions for each algorithm with the real values (NPHI). The legend indicates the algorithm and the mean absolute error/R, respectively Ordinary Least Squares, Bayesian Ridge Regression, Random Sample Consensus, Artificial Neural Network, Random Forest, and Gradient Boosting.

average hydrogen density of the volume under investigation.

Each record has also the depth besides the sensor readings. For the training data set the records where any of these sensors are missing were not considered. Consequently, every well in the data set has at least one

Fig. 8. Example for a fold of the second quartile where the models had a median performance. Top: Comparison of the predictions for each algorithm with the real values. Bottom: Scatter plot of the predictions for each algorithm with the real values (NPHI). The legend indicates the algorithm and the mean absolute error/R, respectively Ordinary Least Squares, Bayesian Ridge Regression, Random Sample Consensus, Artificial Neural Network, Random Forest, and Gradient Boosting.
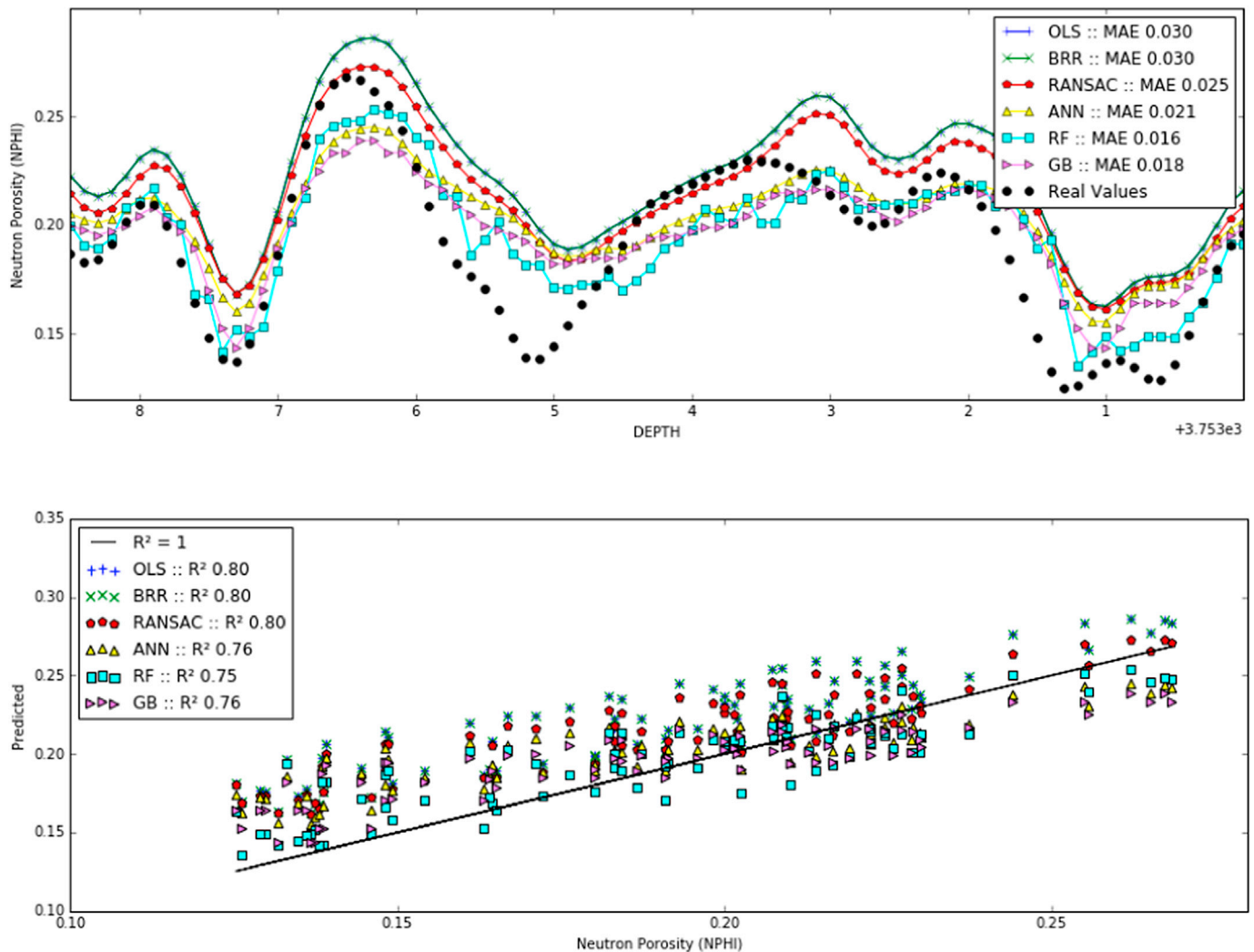
complete record with data from all the sensors acquired. All the variables were normalised to the range $[0, 1]$, with exception of the neutron porosity for which the valid records are originally provided inside the specified range. Table 1 provides a few sample records from different wells that illustrate this description.

For the exploratory and descriptive analysis of the gaps, using the original (not normalised) depth we considered to exist a gap in a well if the distance between two consecutive records is larger than three times the depth step.[2] This excludes punctual small gaps which can be addressed using techniques other than prediction. Given these settings, approximately four and a half thousand gaps were identified across all sensors, in a total of 826 wells (the remaining wells do not exhibit gaps). Figure. 1 illustrates the descriptive statistics for the gaps in the data set, using the logarithmic scale. The values (except the count) represent the gap size (depth) in meters. For most of the wells each meter corresponds to ten data-points, since the typical depth step is $0.1m$. The detailed statistics are presented in Table 2.

The mean and standard deviation are biased by the rare, and very large gaps. Nevertheless, as can be observed by the quartiles data, most of the gaps are rather small for all sensors with exception of the neutron porosity (NPHI). Although the gap count is smaller in this case, the gap size is as

much as three times larger than the remaining sensors' gap sizes. The neutron porosity was then used as the target sensor (to be predicted), since it poses a challenge with increased difficulty. Figures 2-5 detail further the distribution of the gap count by size for each sensor, presenting the full histogram in the logarithmic scale, and zooming over the first three quartiles using the original scale.

The following section describes the learning methodologies and models used, and details the experimental settings used for the comparison of the predictors.

4. Problem and methodologies

The problem of forecasting missing well log data has been addressed using different techniques, from linear models to artificial neural networks (see Section 2). With the exception of a few works where "lengthy gaps" or complete logs are predicted, usually the size of the gaps that are found in the records is not explicitly reported and/or studied. Even though it is dependent on the total depth that was drilled, we have found that the typical sensor gap size goes from a few hundred to a few thousand points (also dependent on the sensor type, see the previous section), whereas a complete log has typically dozens of thousands of data points (the typical depth for a well goes from 3 km to more than 5 km, using 10 samples per meter).

In the next section we compare the results of using generalised linear

---

[2] For most of the wells the depth step is $0.1m$, but in a few cases it is $0.2m$.
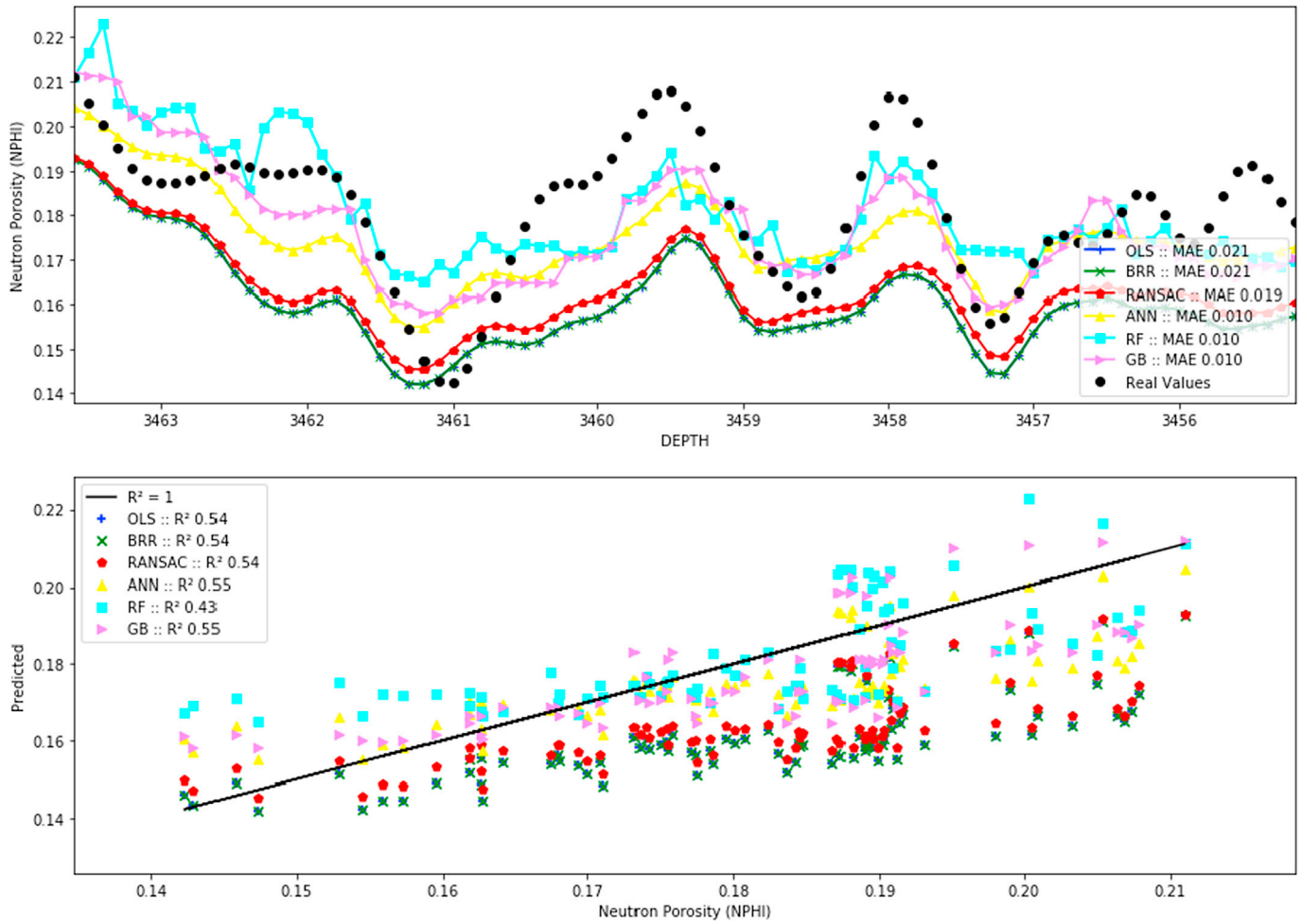
Fig. 9. Example for a fold of the second quartile where the models performed badly. Top: Comparison of the predictions for each algorithm with the real values. Bottom: Scatter plot of the predictions for each algorithm with the real values (NPHI). The legend indicates the algorithm and the mean absolute error/R, respectively Ordinary Least Squares, Bayesian Ridge Regression, Random Sample Consensus, Artificial Neural Network, Random Forest, and Gradient Boosting.

Fig. 10. Histograms of the mean absolute error (*binsize* = 1%) for each algorithm in the first quartile data (smallest gap size). The titles of the subplots indicate the algorithm used: Ordinary Least Squares (OLS), Bayesian Ridge Regression (BRR), Random Sample Consensus (RANSAC), Random Forest (RF), and Gradient Boosting (GB).

models - Ordinary Least Squares (OLS), Bayesian Ridge Regression (BRR), and RANdom SAmple Consensus (RANSAC) ([Rifkin and Lippert, 2007; Choi et al., 2009](#)) - and non-linear models - Artificial Neural Networks (ANN), Random Forests (RF), and Gradient Tree Boosting (GB)

([Breiman, 1999; Friedman, 2001](#)) -, on the prediction of missing gaps in the neutron porosity log of a single well.

The generalised linear models provide a consistent baseline for comparison with non-linear methods. The chosen models are well-

R.L. Lopes, A.M. Jorge

Fig. 11. Histograms of the mean absolute error (*binsize* = 1%) for each algorithm in the second quartile data (median gap size). The titles of the subplots indicate the algorithm used: Ordinary Least Squares (OLS), Bayesian Ridge Regression (BRR), Random Sample Consensus (RANSAC), Random Forest (RF), and Gradient Boosting (GB).
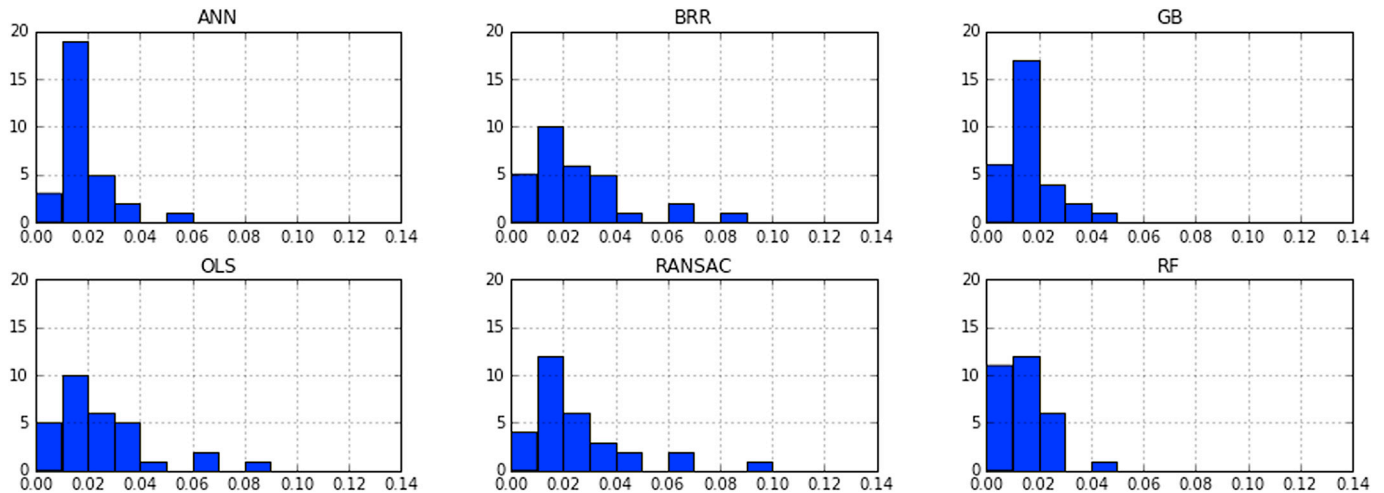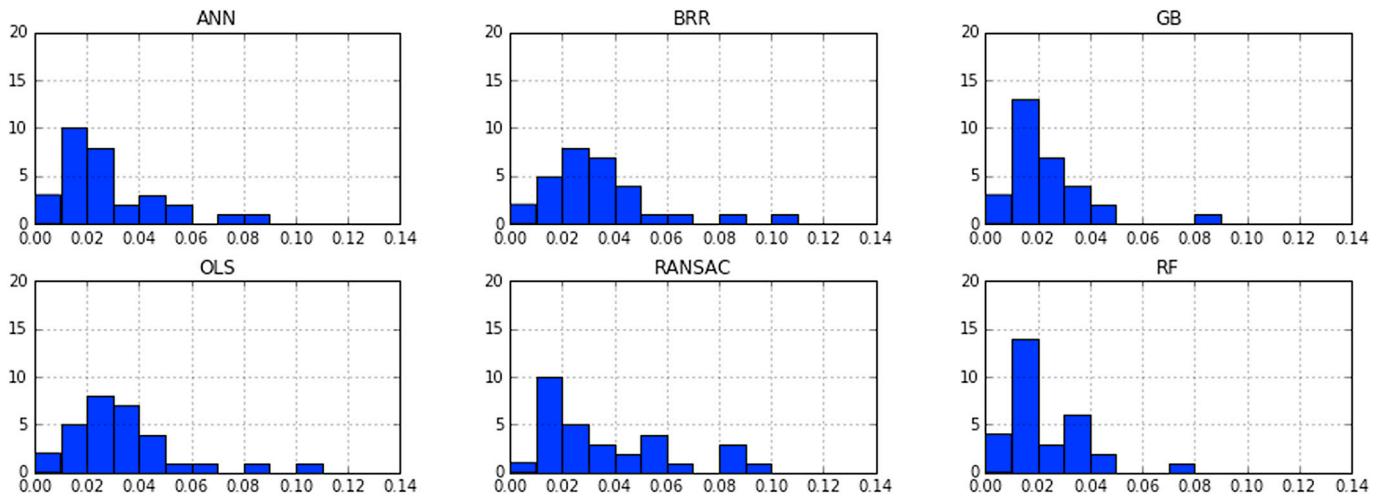

Fig. 12. Histograms of the mean absolute error (*binsize* = 1%) for each algorithm in the third quartile data (largest gap size). The titles of the subplots indicate the algorithm used: Ordinary Least Squares (OLS), Bayesian Ridge Regression (BRR), Random Sample Consensus (RANSAC), Random Forest (RF), and Gradient Boosting (GB).
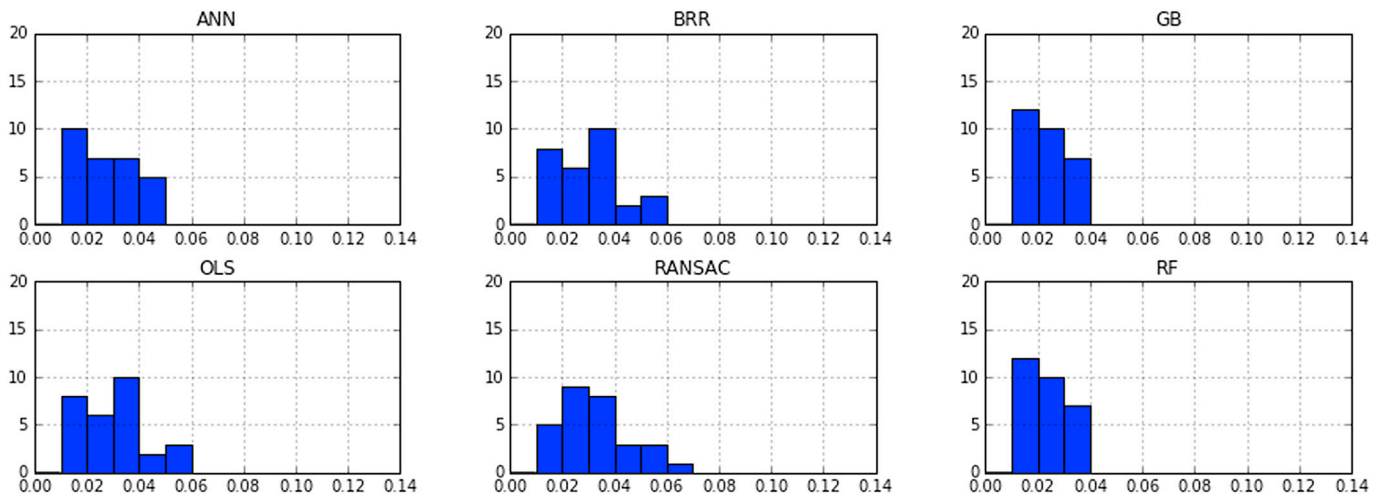
known: i) OLS infers a vector of coefficients that minimize the residual sum of squares between the observed responses in the dataset, and the responses predicted by the linear approximation; ii) BRR includes regularization parameters in the estimation procedure, i.e., the regularization parameter is not hard-coded but tuned to the data at hand; and iii) RANSAC is a non-deterministic algorithm that fits a model from random subsets of inliers from the complete data set, producing a reasonable result with a certain probability, which is dependent on the number of iterations (i.e., the probability increases with the number of iterations).

The non-linear models target the performance improvement over the typical generalised linear models. In this case ANNs, and two ensemble methods based on decision trees (RF and GB) were used. ANNs are computational models inspired by the biological neural networks that constitute animal brains. These models learn to do tasks by considering examples (supervised learning), generally without domain-specific programming.

In particular, a Multi-Layer Perceptron (MLP) was used in this work. It is composed of a collection of inter-connected artificial neurons (similarly to the neurons and synapses in biological brains). Each connection between neurons transmits a signal that varies with the strength of the connection. If the combined input signal is strong enough, the receiving neuron activates and propagates a signal to the downstream neurons connected to it. For the experimental work we used a single hidden layer composed of one hundred rectified linear units. The neurons are typically organized in fully-connected layers, with the signals travelling from the first (input), to the last (output), possibly with recurring connections. The weights associated to each connection are adaptable, with different approaches to the learning algorithm proposed over the years. The most common approach is the back-propagation of errors, typically combined with the stochastic gradient descent (SGD) algorithm.

The ensemble methods combine the predictions of several base estimators built with a given learning algorithm in order to improve generalization/robustness over a single estimator. In RF each tree in the ensemble is built from a sample drawn with replacement (i.e., a bootstrap sample) from the training set. In addition, when splitting a node during the construction of the tree, the split that is picked is the best split among a random subset of the features. Given this randomness, the bias of the forest usually slightly increases (with respect to the bias of a single tree) but its variance also decreases due to averaging. Typically this effect more than compensates for the increase in bias, hence yielding an overall better model. GB combines decision tress of fixed size (the weak learners/models) into a single strong learner, in a stage-wise fashion. Each step a new weak learner is added to the model and the residuals of the previous weak learners are used to improve the fit of the model. As in

R.L. Lopes, A.M. Jorge

Table 3
Results of the statistical significance tests for the pairwise comparisons. For each quartile fifteen results are presented ordered by the test statistic. The methods indicate the algorithm used: Ordinary Least Squares (OLS), Bayesian Ridge Regression (BRR), Random Sample Consensus (RANSAC), Artificial Neural Network (ANN), Random Forest (RF), and Gradient Boosting (GB).

	Quartile	Method1	Method2	statistic	p-value	adj. p-value
0	Q1	RF	OLS	4.39	0.0000	0.0002
1		RF	BRR	4.38	0.0000	0.0002
2		GB	OLS	4.05	0.0001	0.0008
3		GB	BRR	4.04	0.0001	0.0008
4		RF	RANSAC	3.63	0.0003	0.0042
5		GB	RANSAC	3.29	0.0010	0.0150
6		ANN	OLS	2.85	0.0044	0.0657
7		ANN	BRR	2.84	0.0045	0.0678
8		ANN	RANSAC	2.09	0.0365	0.5477
9		RF	ANN	1.54	0.1233	1.0000
10		GB	ANN	1.20	0.2305	1.0000
11		RANSAC	OLS	0.76	0.4484	1.0000
12		RANSAC	BRR	0.75	0.4543	1.0000
13		GB	RF	0.34	0.7324	1.0000
14		BRR	OLS	0.01	0.9921	1.0000
15	Q2	GB	BRR	5.39	0.0000	0.0000
16		GB	OLS	5.39	0.0000	0.0000
17		GB	RANSAC	5.36	0.0000	0.0000
18		RF	BRR	5.22	0.0000	0.0000
19		RF	OLS	5.22	0.0000	0.0000
20		RF	RANSAC	5.19	0.0000	0.0000
21		ANN	BRR	3.97	0.0001	0.0011
22		ANN	OLS	3.97	0.0001	0.0011
23		ANN	RANSAC	3.94	0.0001	0.0012
24		GB	ANN	1.42	0.1557	1.0000
25		RF	ANN	1.25	0.2118	1.0000
26		GB	RF	0.17	0.8643	1.0000
27		RANSAC	BRR	0.04	0.9704	1.0000
28		RANSAC	OLS	0.04	0.9704	1.0000
29		BRR	OLS	0.00	1.0000	1.0000
30	Q3	GB	RANSAC	5.30	0.0000	0.0000
31		GB	BRR	4.97	0.0000	0.0000
32		GB	OLS	4.95	0.0000	0.0000
33		RF	RANSAC	4.36	0.0000	0.0002
34		RF	BRR	4.03	0.0001	0.0008
35		RF	OLS	4.01	0.0001	0.0009
36		ANN	RANSAC	3.31	0.0009	0.0141
37		ANN	BRR	2.98	0.0028	0.0426
38		ANN	OLS	2.96	0.0031	0.0460
39		GB	ANN	1.99	0.0467	0.7007
40		RF	ANN	1.05	0.2947	1.0000
41		GB	RF	0.94	0.3467	1.0000
42		RANSAC	OLS	0.35	0.7288	1.0000
43		RANSAC	BRR	0.32	0.7465	1.0000
44		BRR	OLS	0.02	0.9813	1.0000

Table 4
Parameters used in the grid search for random forests and gradient boosting.

Algorithm	Random forests	Gradient boosting
Learning Rate	N.A.	$\{0.01, 0.05, 0.1, 0.2, 0.3\}$
Max. Tree Depth	$\{None, 10, 50, 100\}$	$\{2, 3, 5, 10, 50\}$
Number of estimators	$\{3, 5, 10, 100, 200\}$	$\{100, 200, 300, 400, 500\}$
Max. number of features	$\{1, 2, 3, 4\}$	
Min. samples per split	$\{5, 10, 50, 100, 200\}$	
Min. samples per leaf	$\{5, 10, 50, 100, 200\}$	

RFs the output of the model is a consensus between the weak learners, usually a weighted contribution from each sub-model.

5. Experiments and discussion

Three different sets of experiments were conducted to compare the algorithms' performances. First, a detailed study-case over a specific well is presented which compares the performance of the different algorithms using the default parametrisations. Second, using the same study-case a

Table 5
Parameters used in the best model found with the grid search for random forests and gradient boosting.

Algorithm	Random forests	Gradient boosting
Learning Rate	N.A.	0.1
Max. Tree Depth	*None*	2
Number of estimators	5	400
Max. number of features	2	3
Min. samples per split	100	10
Min. samples per leaf	2	200

grid-search is performed on the parameter space for the ensemble algorithms. Finally, three different wells are addressed individually in order to confirm the trends identified in the detailed study-case.

5.1. Performance comparison

For the performance comparison of the different models, a single well with complete logs (without gaps) was selected, in particular the well identified by "L01-06". These well logs are composed of eighteen thousand samples with valid values for every sensor. This data set was split into k = number_samples//gap_size folds with non-overlapping test sets composed of contiguous points, where gap_size = $\{27, 85, 615\}$ samples. In the case of the largest gap size the split results in a 29-fold cross-validation. In the remaining gap sizes the number of folds is prohibitive (667 folds for the smaller gap size), hence thirty folds were randomly selected without replacement from the complete fold set. These will be used to average the models' performance, that depends not only on the size of the gap but also on the values of the gap itself. The comparison of the different models for each gap size was performed using the library defaults for every algorithm. The most promising algorithms identified in this analysis were then selected in order to be fine-tuned through a grid search over the model parameters.

The comparison results are summarised in Fig. 6, where the mean absolute error (MAE) is presented for each algorithm in each quartile of the gap size. Even if there are very few outliers for each algorithms' performance measurements, it is clear that it depends on the gap size and on the values of the gap itself since there is a large variance in the folds' results.

Given the high overlap between the interquartile ranges of the scores of the generalised linear models (OLS, BRR, and RANSAC), it is clear that over the complete set of folds the difference between them is not statistically significant. A similar overlap is observed when comparing the non-linear methods (RF and GB) with each other. Nevertheless, it is interesting to note that RF performs slightly better for the smallest gap size, and gradient boosting for the largest size, while in the case of the median gap the performance is identical. The ANN always performs worse than the ensemble methods. While the median error increases with the gap size for every algorithm, the difference between the linear and the non-linear models is smaller as the gap size grows. The question remains if the difference is significant for any of the gap size quartiles.

Figs. 7–9 are examples for folds of the second quartile where the models had from good to bad performance, respectively. The top chart shows the predictions of each algorithm and the real values along the well depth, while the bottom scatter plot shows the predicted values against the real values.

For the fold presented in Fig. 7, even though the mean absolute error (MAE) is lower for the non-linear methods, the linear models show a slightly better fit. The scatter plot shows more dispersion for the non-linear methods, and one can see the differences in the predictions specially in the right part of the gap. In the median case presented in Fig. 8 the MAE is much lower for the non-linear methods, particularly the ensemble methods (RF and GB). However, as observable by the larger dispersion on the scatter plot for every model, in this case the fit is not as good. This plot also shows a tendency of the non-linear methods to overestimate low porosities and underestimate higher porosities. Finally,
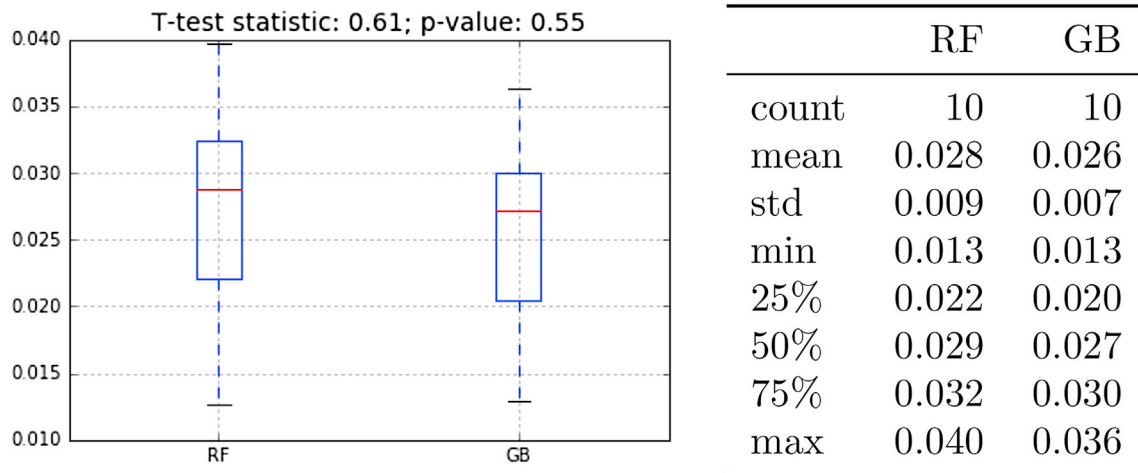
	RF	GB
count	10	10
mean	0.028	0.026
std	0.009	0.007
min	0.013	0.013
25%	0.022	0.020
50%	0.029	0.027
75%	0.032	0.030
max	0.040	0.036

Fig. 13. Descriptive statistics for the testing scores of the best random forest (RF) and gradient boosting (GB) model, over the 10 folds of cross-validation. The *t-test* statistic and p-value are also included.

Fig. 9 shows the case where the fit is poor, as shown by the cloud of points in the scatter plot, even though the MAE is not far from the median.

In order to assess the statistical significance of the differing performances the score distributions were subject to non-parametric tests. These are more appropriate when the distributions are not normal (see Figs. 10–12). Even though the number of samples (30) is enough to ignore the normality of the distributions, it is preferable to be conservative in the analysis.

An analysis of variance by ranks was performed using the Friedman Aligned Ranks (Hodges et al., 1962) (appropriate for a small number of groups), followed by the Nemenyi post-hoc tests for multiple pairwise comparisons (Dunn, 1961). The results are detailed in Table 3, where the fifteen comparisons are presented for each quartile, ordered by the test statistic. Three trends are visible: i) the differences amongst the generalised linear models are never statistically significant ($adj. p - value = 1.0$); ii) the differences between the non-linear methods are also not statistically significant in any quartile; and iii) the differences between the ensemble methods and the linear models are statistically significant for every quartile ($adj. p - value \ll 0.05$). In the case of the ANN it can only be confidently considered better than the linear models in the second quartile, since the differences are not statistically significant in the first quartile ($adj. p - value > 0.05$) and are very close to the confidence limit in the third quartile ($adj. p - value \simeq 0.05$).

5.2. Fine-tuning the ensemble models

It is not clear which of the two ensemble methods is more appropriate for the task. Since these were the best models, this deserves further exploration by fine-tuning each algorithm through a grid search on the parameter space. The values used for each parameter were adjusted by trial and error (detailed in Table 4), in order to find value ranges that do not restrict each other. Note that when *None* is used as the maximum tree depth for the random forests it means that the trees are fully expanded until no further splits are possible. A 10-fold cross validation strategy was applied to evaluate the methods using the mean absolute error (non-overlapping folds of contiguous points).

Despite some tradeoffs amongst parameters, typically both methods should perform better the more estimators are used (with diminishing returns). Moreover, considering the amount of variance in the data, it should be expected better results with shallower decision trees. As one can see in Table 5 these expectations were met in the case of the gradient boosting ensemble. The best model was obtained using the maximum number of estimators provided and very shallow trees (2). However, for

the random forests the best strategy was to use a large maximum depth, and a relatively small number of trees (5). Accordingly, GB used a large number of samples per leaf (200) with a small number of samples per split (10), while for RF the best used a large number of samples per split (100) with a small number of samples per leaf (2). The details of the performance of the best estimator over the cross-validation folds are presented in Fig. 13. Overall, both the mean and median absolute errors of the best configuration are lower for the gradient boosting algorithm (by 0.2%). However, the difference is not statistically significant as indicated by the results of the *t*-test, included on the figure's title.

5.3. Other study-cases

A pertinent question arising from the previous experiments is whether these results hold for wells of different exploration blocks. While a definitive answer cannot be provided, the analysis above was repeated for three more wells, each from a different block. The wells "F12-03", "K05-02", and "M04-03" were chosen from the subset of wells without gaps. After cleaning these have approximately fifteen thousand samples each, thus the resulting number of folds for the largest gap size is not much lower compared to the original well analysed in this article. The algorithms' performance shows the same trend identified before. The ensemble methods (RF and GB) perform better than the remaining, as shown by the distribution of the scores over the sets of folds for each gap size in Fig. 14. The statistical analysis of the results confirms these differences, as displayed in Table 7 at the end of this article. The ensemble methods are always better than the linear regressors ($adj. p - value \ll 0.05$), and in some cases better than the ANN. Similarly to well "L01-06" the differences between the ANN and the linear methods are typically not statistically significant ($adj. p - value = 1.0$).

Given these results it is relevant to inspect if there is a trend on the configuration of the ensemble algorithms as well, or if the best parametrisation depends on the logs of the well that is being addressed. For that purpose a grid search over the ensembles' parameters was run for each of the three wells, using the same ranges as described in Table 4. The results for the three wells are summarised in Table 6, where one can see that the gradient boosting performed slightly better for two out of the three wells. The trend on the number of estimators is also visible for most wells, with a low number of estimators preferred for the random forests, and high for gradient boosting. However, there is no consistency over all the wells for any of the parameters. Similarly to the results presented before the differences between the mean absolute errors of the two algorithms is minimal (always inferior to 0.2%) and not statistically
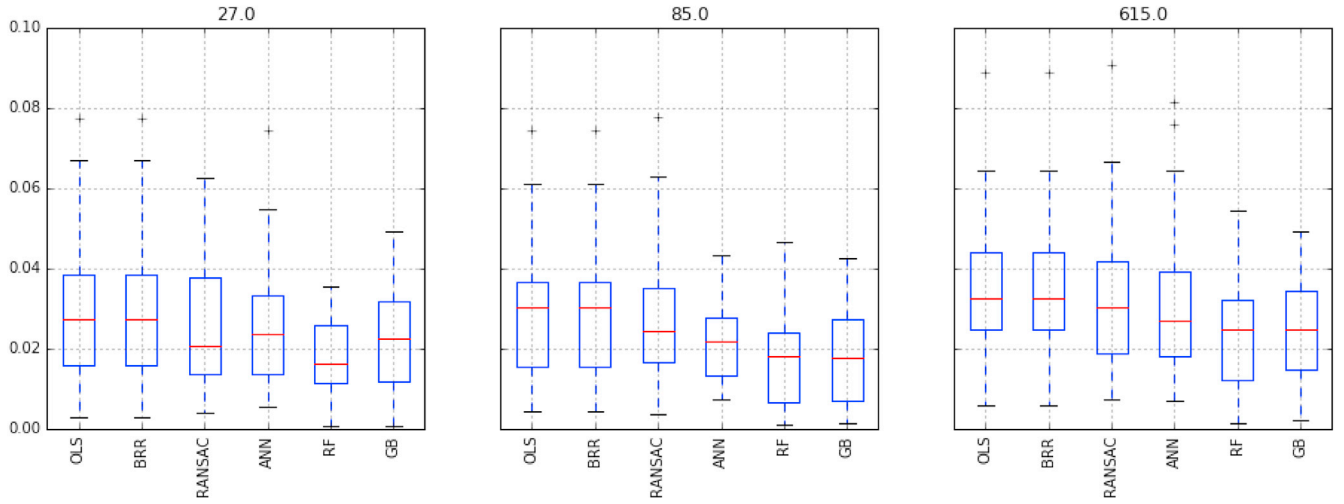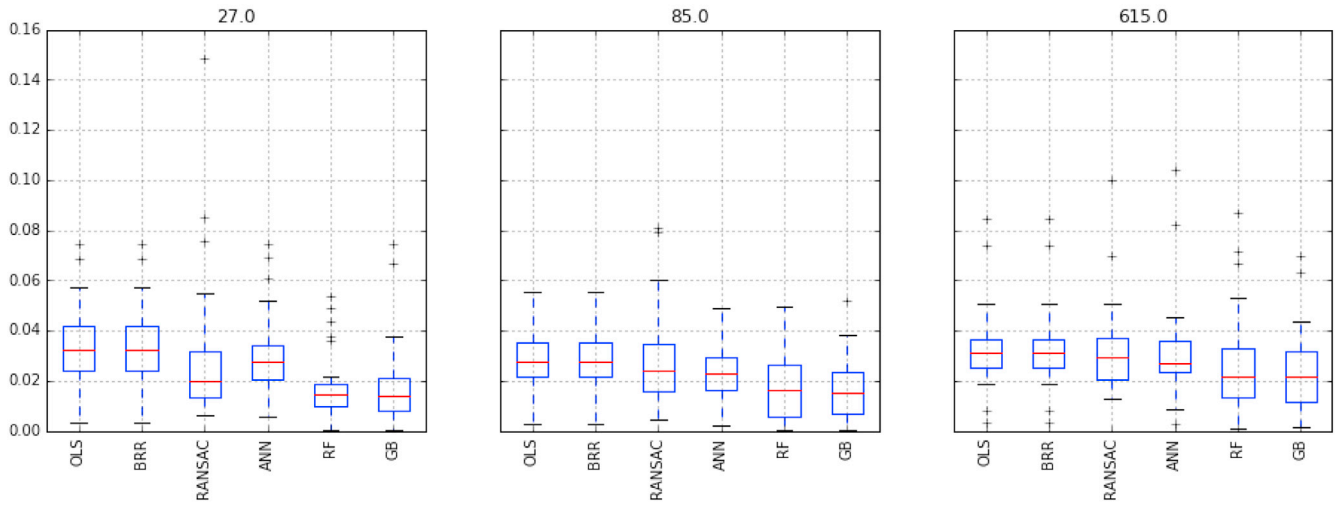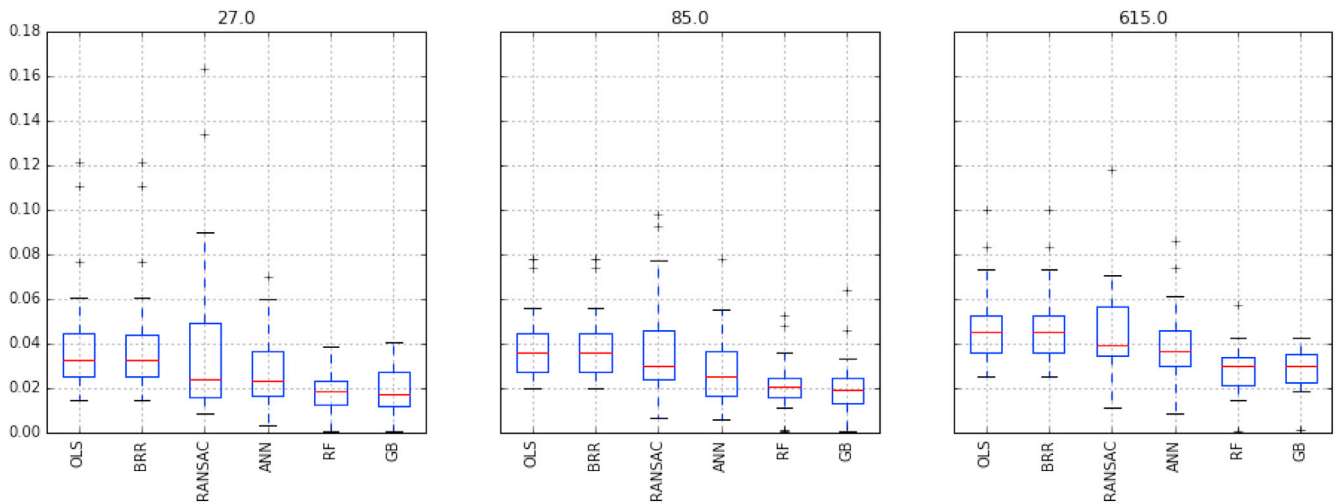
Fig. 14. Mean absolute error for each algorithm by gap size, over all folds in each well. The values in the y-axis indicate percentages in the interval [0, 1]. The labels on the x-axis indicate the algorithm used, respectively Ordinary Least Squares, Bayesian Ridge Regression, Random Sample Consensus, Artificial Neural Network, Random Forest, and Gradient Boosting.

Table 6
Parameters used in the best model for each well, found with the grid search for random forests and gradient boosting.

Algorithm	Random forests			Gradient boosting		
Well	F12-03	K05-02	M04-03	F12-03	K05-02	M04-03
Learning rate	–	–	–	0.1	0.2	0.3
Max. tree depth	*None*	5	*None*	10	5	5
Num. estimators	10	10	200	100	200	300
Max. features	2	4	2	4	3	2
Samples p/split	50	10	10	50	50	200
Samples p/leaf	3	5	3	200	5	100
MAE	0.0233	0.0235	0.0287	0.0247	0.0225	0.0285

significant.

The experiments described above were implemented using Pythons scientific environment (Oliphant, 2007), using the models available in the Scikit-Learn (Pedregosa et al., 2011) machine learning library. For the statistical tests the Stac library was used (Rodrguez-Fdez et al., 2015).

6. Conclusion and future work

The main goal of this work was to evaluate different methods for the regression of missing well log data. The first step consisted of a descriptive and exploratory analysis in order to gain a better understanding of the gaps' size and frequency across a large set of wells. Then, different data-driven models were trained for filling-in gaps in sensor data from a specific well log, using the remaining sensors as features, at any depth of that well. Finally, the study was extended to three wells on different exploration blocks to check whether the results hold for different data sets.

The analysis of the gap sizes has shown that in most cases the gaps are rather small when compared to the complete depth of the wells. Nevertheless, for every sensor the typical gap size goes up until a few hundreds of contiguous points. The neutron porosity stands out from the remaining sensors, the gap count is larger and the typical gap size is around three times larger than the remaining gaps. Since this is apparently the most unreliable sensor it was chosen as the prediction target.

The results of the experiments show that the larger the gap the higher the average error. This could be expected for different reasons, in particular because with larger gaps less data remains for the training, and there is typically more variance in the target values. Moreover, the results for each algorithm present a large variation amongst the folds, showing that the performance depends on the gap itself, not only on the gap size.

The ensemble methods (random forests and gradient boosting) performed better than the linear approaches for every gap size quartile, with statistical significance. One can see that as the gap size increases the difference between the two types of methods decreases. Nevertheless, those differences remain statistically significant. In the case of the artificial neural network its performance was not significantly different of either the linear or the ensemble methods. Given the similarity between the results of both the ensemble methods, a grid search was performed over the parameter space. The grid results indicate that the strategy for each algorithm could not be more opposite. In the case of the random forests the preference was for a small number of estimators and high maximum tree depth, while for gradient boosting the best result used a large number of estimators and shallow decision trees. Despite these differences and gradient boosting presenting lower mean absolute error, the overall performance difference is not statistically significant over the complete set of folds.

Applying the same methodology to different data sets confirmed the results concerning the superior performance of the ensemble algorithms. It also confirmed that the differences between the ensemble methods are not statistically significant even after fine-tuning. However, it has shown that the appropriate parametrisation is dependant on the data set (well) being addressed and cannot be prescribed based on the choice of

Table 7
Results of the statistical significance tests for the pairwise comparisons for wells "F12-03", "K05-02", and "M04-03" respectively. For each quartile fifteen results are presented ordered by the test statistic, numbered independently for each well. The methods indicate the algorithm used: Ordinary Least Squares (OLS), Bayesian Ridge Regression (BRR), Random Sample Consensus (RANSAC), Artificial Neural Network (ANN), Random Forest (RF), and Gradient Boosting (GB).

Well	Quartile	Method1	Method2	statistic	p-value	adj. p-value
F12-03	Q1	RF	BRR	5.15	0.0000	0.0000
		RF	OLS	5.14	0.0000	0.0000
		GB	BRR	3.59	0.0003	0.0049
		GB	OLS	3.58	0.0003	0.0052
		RF	ANN	3.33	0.0009	0.0129
		RF	RANSAC	2.94	0.0033	0.0499
		RANSAC	BRR	2.22	0.0266	0.3988
		RANSAC	OLS	2.20	0.0276	0.4143
		ANN	BRR	1.82	0.0686	1.0000
		ANN	OLS	1.81	0.0709	1.0000
		GB	ANN	1.77	0.0765	1.0000
		GB	RF	1.56	0.1185	1.0000
		GB	RANSAC	1.38	0.1691	1.0000
		ANN	RANSAC	0.40	0.6918	1.0000
		BRR	OLS	0.01	0.9881	1.0000
	Q2	RF	BRR	5.96	0.0000	0.0000
		RF	OLS	5.95	0.0000	0.0000
		GB	BRR	5.26	0.0000	0.0000
		GB	OLS	5.25	0.0000	0.0000
		RF	RANSAC	4.48	0.0000	0.0001
		GB	RANSAC	3.78	0.0002	0.0023
		ANN	BRR	3.48	0.0005	0.0075
		ANN	OLS	3.47	0.0005	0.0078
		RF	ANN	2.48	0.0133	0.1998
		ANN	RANSAC	2.01	0.0445	0.6675
		GB	ANN	1.77	0.0761	1.0000
		RANSAC	BRR	1.47	0.1411	1.0000
		RANSAC	OLS	1.46	0.1438	1.0000
		GB	RF	0.70	0.4832	1.0000
		BRR	OLS	0.01	0.9921	1.0000
		RF	BRR	6.30	0.0000	0.0000
		RF	OLS	6.29	0.0000	0.0000
		GB	BRR	6.17	0.0000	0.0000
	Q3	GB	OLS	6.16	0.0000	0.0000
		RF	RANSAC	4.69	0.0000	0.0000
		GB	RANSAC	4.56	0.0000	0.0001
		RF	ANN	3.55	0.0004	0.0057
		GB	ANN	3.42	0.0006	0.0094
		ANN	BRR	2.75	0.0060	0.0893
		ANN	OLS	2.74	0.0062	0.0927
		RANSAC	BRR	1.61	0.1064	1.0000
		RANSAC	OLS	1.60	0.1091	1.0000
		ANN	RANSAC	1.14	0.2561	1.0000
		GB	RF	0.14	0.8926	1.0000
		BRR	OLS	0.01	0.9902	1.0000
K05-02	Q1	RF	BRR	5.83	0.0000	0.0000
		RF	OLS	5.80	0.0000	0.0000
		GB	BRR	5.66	0.0000	0.0000
		GB	OLS	5.62	0.0000	0.0000
		RF	ANN	4.42	0.0000	0.0002
		GB	ANN	4.24	0.0000	0.0003
		RF	RANSAC	3.28	0.0011	0.0158
		GB	RANSAC	3.10	0.0020	0.0293
		RANSAC	BRR	2.56	0.0105	0.1573
		RANSAC	OLS	2.52	0.0116	0.1737
		ANN	BRR	1.42	0.1557	1.0000
		ANN	OLS	1.39	0.1661	1.0000
		ANN	RANSAC	1.14	0.2544	1.0000
		GB	RF	0.18	0.8584	1.0000
		BRR	OLS	0.03	0.9723	1.0000
	Q2	GB	RANSAC	5.85	0.0000	0.0000
		GB	BRR	5.77	0.0000	0.0000
		GB	OLS	5.74	0.0000	0.0000
		RF	RANSAC	5.50	0.0000	0.0000
		RF	BRR	5.43	0.0000	0.0000
		RF	OLS	5.40	0.0000	0.0000

Table 7 (*continued*)

Well	Quartile	Method1	Method2	statistic	p-value	adj. p-value
		GB	ANN	3.33	0.0009	0.0131
		RF	ANN	2.98	0.0029	0.0428
		ANN	RANSAC	2.52	0.0117	0.1761
		ANN	BRR	2.45	0.0145	0.2170
		ANN	OLS	2.42	0.0157	0.2356
		GB	RF	0.34	0.7306	1.0000
		RANSAC	OLS	0.10	0.9171	1.0000
		RANSAC	BRR	0.07	0.9407	1.0000
		BRR	OLS	0.03	0.9763	1.0000
	Q3	GB	BRR	4.71	0.0000	0.0000
		GB	OLS	4.66	0.0000	0.0000
		GB	RANSAC	4.44	0.0000	0.0001
		GB	ANN	3.67	0.0002	0.0036
		RF	BRR	3.53	0.0004	0.0063
		RF	OLS	3.48	0.0005	0.0075
		RF	RANSAC	3.26	0.0011	0.0165
		RF	ANN	2.49	0.0127	0.1900
		GB	RF	1.18	0.2380	1.0000
		ANN	BRR	1.03	0.3018	1.0000
		ANN	OLS	0.99	0.3230	1.0000
		ANN	RANSAC	0.77	0.4409	1.0000
		RANSAC	BRR	0.26	0.7935	1.0000
		RANSAC	OLS	0.22	0.8278	1.0000
		BRR	OLS	0.04	0.9647	1.0000
M04-03	Q1	RF	OLS	6.12	0.0000	0.0000
		RF	BRR	6.10	0.0000	0.0000
		GB	OLS	6.01	0.0000	0.0000
		GB	BRR	6.00	0.0000	0.0000
		RF	RANSAC	4.58	0.0000	0.0001
		GB	RANSAC	4.48	0.0000	0.0001
		ANN	OLS	3.21	0.0013	0.0200
		ANN	BRR	3.19	0.0014	0.0211
		RF	ANN	2.91	0.0036	0.0544
		GB	ANN	2.80	0.0050	0.0755
		ANN	RANSAC	1.67	0.0940	1.0000
		RANSAC	OLS	1.53	0.1251	1.0000
		RANSAC	BRR	1.52	0.1288	1.0000
		GB	RF	0.10	0.9171	1.0000
		BRR	OLS	0.01	0.9881	1.0000
	Q2	RF	OLS	7.11	0.0000	0.0000
		RF	BRR	7.11	0.0000	0.0000
		GB	OLS	7.09	0.0000	0.0000
		GB	BRR	7.08	0.0000	0.0000
		RF	RANSAC	5.89	0.0000	0.0000
		GB	RANSAC	5.87	0.0000	0.0000
		ANN	OLS	4.10	0.0000	0.0006
		ANN	BRR	4.09	0.0000	0.0006
		RF	ANN	3.01	0.0026	0.0388
		GB	ANN	2.99	0.0028	0.0418
		ANN	RANSAC	2.88	0.0040	0.0594
		RANSAC	OLS	1.22	0.2238	1.0000
		RANSAC	BRR	1.21	0.2257	1.0000
		GB	RF	0.02	0.9822	1.0000
		BRR	OLS	0.00	0.9960	1.0000
	Q3	RF	BRR	5.90	0.0000	0.0000
		RF	OLS	5.89	0.0000	0.0000
		GB	BRR	5.87	0.0000	0.0000
		GB	OLS	5.86	0.0000	0.0000
		RF	RANSAC	5.21	0.0000	0.0000
		GB	RANSAC	5.18	0.0000	0.0000
		RF	ANN	2.95	0.0032	0.0473
		ANN	BRR	2.95	0.0032	0.0483
		ANN	OLS	2.94	0.0033	0.0493
		GB	ANN	2.92	0.0035	0.0525
		ANN	RANSAC	2.26	0.0237	0.3552
		RANSAC	BRR	0.68	0.4942	1.0000
		RANSAC	OLS	0.68	0.4984	1.0000
		GB	RF	0.03	0.9740	1.0000
		BRR	OLS	0.01	0.9948	1.0000

the algorithm.

Future work may address the improvement of the results through the integration with seismic data at the target well location, if available.

Another possible approach is to cluster samples by log similarity and develop models for the different log characteristics. A common trait amongst this and the related work is the small amount of features used as input. Feature engineering (extracting new attributes from the records, combining attributes, and selecting the most discriminative) are also appropriate avenues of research in order to improve upon these results. Finally, this work could be extended to a multi-well scenario, where the model is inferred from several wells in a given region, and used for prediction in a different well of the same region.

References

Ahmadi, M.-A., Ahmadi, M.R., Hosseini, S.M., Ebadi, M., 2014. Connectionist model predicts the porosity and permeability of petroleum reservoirs by means of petrophysical logs: application of artificial intelligence. J. Petrol Sci. Eng. 123, 183–200.

Aifa, T., Baouche, R., Baddari, K., 2014. Neuro-fuzzy system to predict permeability and porosity from well log data: a case study of Hassi R Mel gas field, Algeria. J. Petrol Sci. Eng. 123, 217–229.

Aminzadeh, F., De Groot, P., 2006. Neural Networks and Other Soft Computing Techniques with Applications in the Oil Industry. EAGE Publications. URL. https://books.google.pt/books?id=SzFQAQAAIAAJ.

Ayoub, M.A., Mohamed, A.A., 2014. Estimating the lengthy missing log interval using group method of data handling (GMDH) technique. Appl. Mech. Mater. 695, 850. URL. http://akademiabaru.com/wvcarmea/docu/157.pdf.

Breiman, L., 1999. Random forests. UC Berkeley TR567. URL. http://machinelearning202.pbworks.com/w/file/fetch/60606349/breiman_randomforests.pdf.

Choi, S., Kim, T., Yu, W., 2009. Performance Evaluation of RANSAC Family. British Machine Vision Association, 81.1–81.12. URL. http://www.bmva.org/bmvc/2009/Papers/Paper355/Paper355.html.

Cranganu, C., Bautu, E., 2010. Using gene expression programming to estimate sonic log distributions based on the natural gamma ray and deep resistivity logs: a case study from the Anadarko Basin, Oklahoma. J. Petrol Sci. Eng. 70 (3), 243–255.

Cranganu, C., Breaban, M., Mar. 2013. Using support vector regression to estimate sonic log distributions: a case study from the Anadarko Basin, Oklahoma. J. Petrol Sci. Eng. 103, 1–13.

Dunn, O.J., Mar. 1961. Multiple comparisons among means. J. Am. Stat. Assoc. 56 (293), 52. URL. http://www.jstor.org/stable/2282330?origin=crossref.

Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Statist. 29 (5), 1189–1232 bibtex: friedman_greedy_2001. URL. http://www.jstor.org/stable/2699986.

Hodges, J.L., Lehmann, E.L., others, 1962. Rank methods for combination of independent experiments in analysis of variance. Ann. Math. Statist. 33 (2), 482–497. URL. http://projecteuclid.org/euclid.aoms/1177704575.

Holdaway, K., 2014. Harness Oil and Gas Big Data with Analytics: Optimize Exploration and Production with Data Driven Models, first ed. Wiley Publishing. bibtex: Holdaway:2014.

Holmes, M., Holmes, D., Holmes, A., 2003. Generating missing logs b techniques and pitfalls. In: AAPG Annual Meeting, Salt Lake City, Utah. URL. http://www.searchanddiscovery.com/documents/2003/holmes/images/holmes.pdf.

McNamara, A., McKenzie-Brown, P., Sep. 2013. Petroleum Exploration and Production. URL. http://www.thecanadianencyclopedia.ca/en/article/petroleum-exploration-and-production/.

Mohseni, H., Esfandyari, M., Habibi Asl, E., 2015. Application of artificial neural networks for the prediction of carbonate lithofacies, based on well log data, Sarvak Formation, Marun oil field, SW Iran. Geopersia 5 (2), 111–123. URL. http://journals.ut.ac.ir/article_56089_1dfe9f3ecdbc56e9b1876321eae020d7.pdf.

Mojeddifar, S., Kamali, G., Ranjbar, H., Salehipour Bavarsad, B., 2014. A comparative study between a pseudo-forward equation (PFE) and intelligence methods for the characterization of the north sea reservoir. Int. J. Min. Geo Eng. 48 (2), 173–190. URL. http://journals.ut.ac.ir/article_53105_48.html.

Oliphant, T.E., May 2007. Python for scientific computing. Comput. Sci. Eng. 9 (3), 10–20 bibtex: oliphant_python_2007.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: machine learning in Python. J. Mach. Learn. Res. 12, 2825–2830 bibtex: scikit-learn.

Rafik, B., Kamel, B., Nov. 2016. Prediction of permeability and porosity from well log data using the nonparametric regression with multivariate analysis and neural network, Hassi RMel Field, Algeria. Egypt. J. Petrol. URL. http://linkinghub.elsevier.com/retrieve/pii/S1110062116300290.

Rifkin, R.M., Lippert, R.A., 2007. Notes on regularized least squares. URL. https://dspace.mit.edu/handle/1721.1/37318.

Rodrguez-Fdez, I., Canosa, A., Mucientes, M., Bugarn, A., 2015. STAC: a web platform for the comparison of algorithms using statistical tests. In: Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on. IEEE, pp. 1–8. URL. http://ieeexplore.ieee.org/abstract/document/7337889/.

Singh, S., Kanli, A.I., Sevgen, S., Jan. 2016. A general approach for porosity estimation using artificial neural network method: a case study from Kansas gas field. Studia

Geophys. Geod. 60 (1), 130–140. URL. http://link.springer.com/10.1007/s11200-015-0820-2.

Soubotcheva, N., Stewart, R.R., 2004. Predicting porosity logs from seismic attributes using geostatistics. CREWES Res. Rep. 16. URL. http://crewes.org/ForOurSponsors/ResearchReports/2004/2004-03.pdf.

Yu, Y., Seyler, D., McCormack, M., Nov. 2011. Method for estimating missing well log data. Google Patents, US Patent 8,065,086. URL. https://www.google.com/patents/US8065086.

Zerrouki, A.A., Aïfa, T., Baddari, K., 2014. Prediction of natural fracture porosity from well log data by means of fuzzy ranking and an artificial neural network in Hassi Messaoud oil field, Algeria. J. Petrol Sci. Eng. 115, 78–89.