SPE-201453-MS

A New Machine Learning Procedure to Generate Highly Accurate Synthetic Shear Sonic Logs in Unconventional Reservoirs

Ilia Alexandrovich Chaikine, Sproule; Ian Donald Gates, University of Calgary

This paper was prepared for presentation at the SPE Annual Technical Conference & Exhibition originally scheduled to be held in Denver, Colorado, USA, 5 – 7 October 2020. Due to COVID-19 the physical event was postponed until 26 – 29 October 2020 and was changed to a virtual event. The official proceedings were published online on 21 October 2020.

Abstract

Shear sonic travel time (DTS), along with compressional sonic travel time and bulk density are required in order to estimate rock mechanical properties which play an important role in fracture propagation and the success of hydraulic fracture treatments in horizontal wells. DTS logs are often missing from the log suite due to their costs and time to process. The following study presents a machine learning procedure capable of generating highly accurate synthetic DTS curves. A hybrid convolutional-recurrent neural network (c-RNN) was chosen in the development of this procedure as it can learn sequential data which a traditional neural network (ANN) cannot. The accuracy of the c-RNN was superior when compared to that of the ANN, simple baselines and empirical correlations. This procedure is a cost effective and fast alternative to running DTS logs and with further development, has the potential to be used for predicting production performance from unconventional reservoirs.

Introduction

Mechanical rock properties such as the Young's modulus and Poisson's ratio are important to understand the way fractures propagate within reservoir rocks during hydraulic fracturing. Furthermore, the mechanical properties in an unconventional reservoir are anisotropic and heterogenous, which makes the production performance of a horizontal well highly dependent upon its location and orientation within the reservoir. It remains unclear how fractures propagate within these tight reservoirs especially in the context of the state-of-stress but a full understanding of the mechanical properties can yield insights that enable the ability to ultimately predict the production performance of each stage and cumulatively, of the well.

In typical practice, measuring mechanical properties is done through direct and indirect methods (Maleki et al., 2014). Direct methods involve taking core samples from the reservoirs and performing laboratory tests, for example, triaxial stress tests, to understand the stress-strain behavior of the rock sample as well as the stress at which the rock fails. Due to the high costs and difficulty of rock sample extraction, typically only a few core samples are retrieved from the formation of interest. Also, laboratory tests are expensive and time consuming (Maleki et al., 2014). Even when obtained, core data only provides point-value (discrete) information on the formation of interest – the fraction of the reservoir sampled is typically under 1%.

Discrete data is not very useful when describing the entire formation since the mechanical properties of rock vary greatly with depth and the loading conditions of the surrounding formations which can vary areally due to variable thickness and densities of the formations and unconformities. A continuous mechanical property measurement along the thickness of the formation is desired.

Fortunately, mechanical rock properties can be derived indirectly from a combination of three sets of log measurements, specifically: bulk density (RHOB), compressional slowness (DTP), and shear slowness (DTS) (Fjar et al., 2008). The problem is that shear sonic data is often missing from a well log suite due to its high cost and length of time to acquire. Without log or core data, the only way to estimate the shear sonic value is either through empirical correlations or statistical methods (Hadi and Nygaard, 2018). Most empirical correlations attempt to express the relationship between the DTP and DTS measurements obtained from the wells with DTS logs. Empirical correlations were the main way to estimate DTS for many years with the more familiar works from Caroll et al. (1969), Castagna et al. (1985), and Han et al. (1986). The difficulty is that although DTS is highly correlated to DTP, multiple variables influence the DTS including pore pressure, fluid saturation, clay and shale content, stress profiles and other factors (Barre, 2009). Empirical correlations are also highly dependent upon the location and size of the study area, and result in poor estimates if taken outside that area.

Figure 1 plots DTS versus DTP data of the 14 wells used in this study (details on the formation is described below). It is easy to see the highly nonlinear, seemingly uncorrelated nature of the relationship which makes predicting DTS difficult. Since it is impossible to know all of the causal factors contributing to DTS, the better approach is to apply statistical methods. Machine learning algorithms have capabilities to integrate subtleties between data sets more than simple regression models. These algorithms work by learning patterns between input and output variables that are too complex and subtle for regular statistics to capture. There have been multiple studies applying various machine learning techniques to DTS prediction. Eskandari et al. (2004) used artificial neural networks (ANN) and showed that they performed better than empirical correlations. Rajabi et al. (2010) used genetic algorithms (GA) and fuzzy logic (FL) which were able to capture the general shape of the log but did not perform well on localized variations. Maleki et al. (2014) used both support vector regression (SVR) and ANN and found that although SVR was a better predictor its accuracy was limited and was not a good replacement to the true log. Al-Anazi and Gates (2015) used support vector regression (SVR) to predict Poisson's ratio and Young's modulus with a fuzzy-based ranking algorithm to select the most significant input variables and filter out dependency. There is no single algorithm that is superior to others and several types should be used in a particular study; however, ANNs have recently become the popular choice (Parapuram et al. 2018; Hadi and Nygaard, 2018). Recurrent neural networks (RNNs) are of particular interest as they are useful for sequence data such as logs as they retain a memory of previous patterns. RNNs have shown to overperform ANNs (Zhang et al. 2018). Studies to date have been able to capture the general trend of the relationship; however, they do not adequately capture smaller localized variations. These localized variations are important since they can play a major role in the mechanical properties which in turn can affect the way fractures propagate.

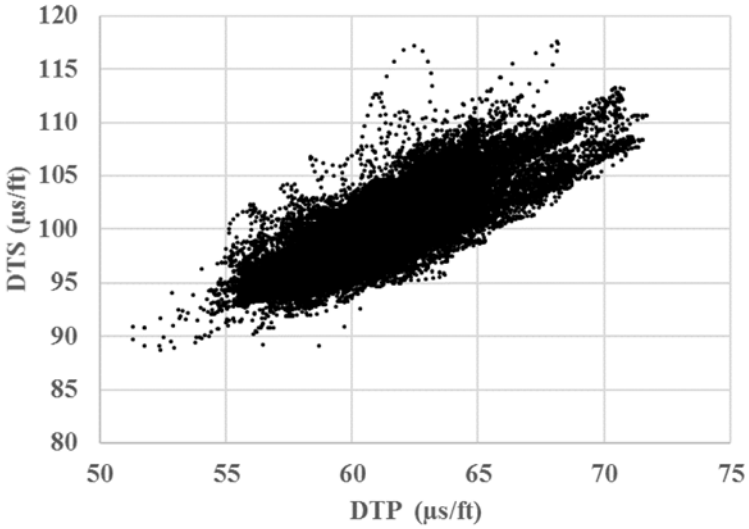

Figure 1—Plot of shear slowness (DTS) versus compressional slowness (DTP) measurements for 14 wells in the Montney formation.

Neural networks are a set of algorithms inspired by neural networks found in the brains of animals where the networks map input variables to output targets and find correlations by observing many samples of the data. They are made up of several layers composed of nodes. Typically, every node in one layer is connected to every node in the next via a weight factor, the network learns by updating these weights based on the data it sees. A network is said to be trained once it can achieve the highest accuracy on a validation set. Since there could be millions of weight values in a network, the only feasible way to update them is via backpropagation of a loss function (the difference between the true value and the one estimated by the network) initiated by an optimizer (Chollet, 2017).

Deep networks (networks with many successive layers) were introduced in the early 1990's and due to their impressive performance have become one the most popular machine learning tools today. With each successive layer in the network, more complex representations are developed. Here, we use of three types of layers (Chollet, 2017):

1. Densely connected – this is the simplest type of layer, every node in one layer is fully connected to every node in the next.
2. Recurrent – this type of layer is used for sequence-based data such as time series or logs. It iterates through a sequence of a sample while maintaining a state of memory relative to what has already been seen.. The two most popular types of recurrent nodes are the Long Short-Term Memory (LSTM) and the more recent gated recurrent unit (GRU). Both units have similar performance and are related in that they gate information to prevent the vanishing gradient problem. In this study we use the GRU since it has a simpler structure and is more computationally efficient than the LSTM (Chung et al., 2014)
3. Convolutional – this is a layer that learns spatial hierarchies of patterns. When stacked together making a convolutional network, they are able to capture both local and global patterns. As a result, they are mostly used for image identification tasks; however, they have also found success in processing sequence data. Their performance is competitive with recurrent layers and is usually computationally less intensive.

Deep networks are able to combine many different types of layers, which opens nearly unlimited potential for experimenting with configurations to find an optimal structure for a certain type of problem. There is no optimal network structure since it depends largely on the problem at hand. Some rules of thumb exist that can be used as a starting guide. However, the only way to find the optimal configuration is by manually

tuning hyperparameters and running many experiments. Hyperparameters are parameters whose value is set before the training occurs. Important hyperparameters are the type and number of layers, number of nodes in each layer, dropout (which introduces randomness), batch size and optimizer.

In the following, we present a procedure that uses a hybrid variation of the RNN to generate synthetic DTS curves from DTP and RHOB logs and well deviation survey. The procedure is applied to a set of wells in the Montney Formation in Alberta, Canada.

Study Area and Data Processing

Montney Formation

The Montney Formation, shown in Figure 2, is a tight siltstone/shale oil and gas play that spans over 500 km from southeast Alberta to northwest British Columbia and covers ~130,000 km2. The quality of reservoir varies significantly, with the general trend of conventional sandstone in the southeast to pure shale in the northwest. The thickness and depth also vary greatly from <1 m thickness and 500 m depth in the southeast to over 350 m thickness and 4,000 m depth in the northwest. The in-place resource volumes for the play are estimated at over 2,000 Tcf of gas and over 150 billion barrels of oil and condensate (Reynolds et al. 2014; Wang and Chen, 2016). The Montney Formation has become a major producer, due to the enormous volumes of in-place hydrocarbons and the major developments in horizontal drilling and multi-stage hydraulic fracturing technologies. With thousands of multistage horizontal wells drilled to date, it accounts for ~38% of total Canadian natural gas production (Canada's Energy Future, 2018).

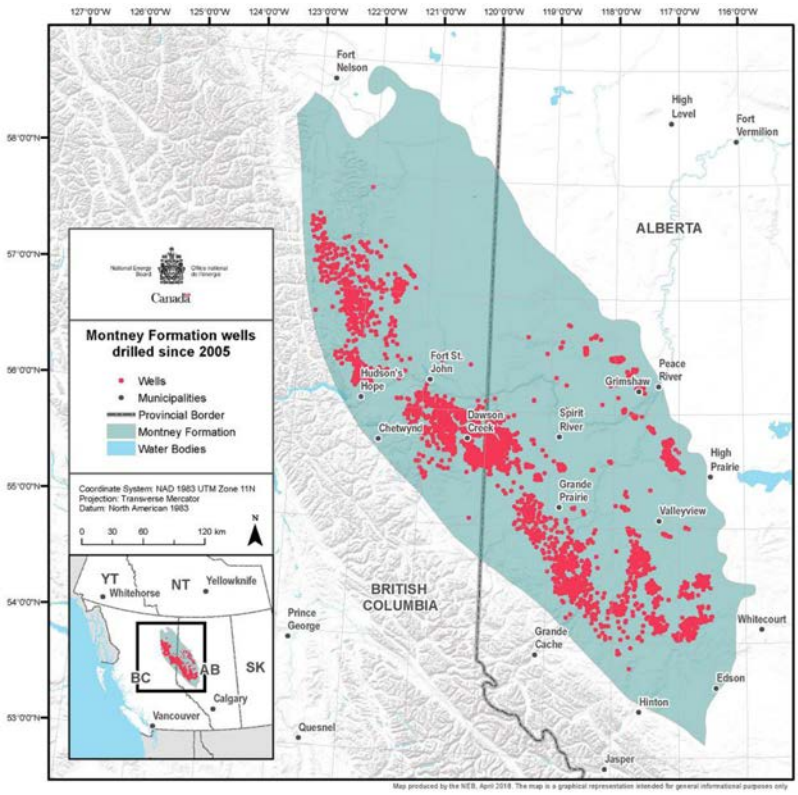

Figure 2—Areal extent of the Montney Formation (National Energy Board, 2018).

Study Area

The study is located near Dawson Creek in the province of British Columbia in a 30 km by 30 km area with average thickness 300 m. In this area, the Montney underlies the Doig Formation and overlies the Belloy Formation. The area contains 145 vertical wells, 35 deviated wells, and 255 horizontal wells that penetrate

the Montney Formation with only 60 of these wells penetrating all the way to the Belloy Formation. An areal view of all the wells in the study is shown in Figure 3. Most of the horizontal and vertical wells are located in the middle of the study area with some horizontal wells in the northwest. A three-dimensional (3D) structural model of the study was constructed to get a better visual understanding of the reservoir and to serve as a placeholder for data. Figure 4 shows the model along with the Montney and Belloy tops. All of the horizontal wells in the study target the Montney Formation with the majority drilled in the top 20% of its thickness.

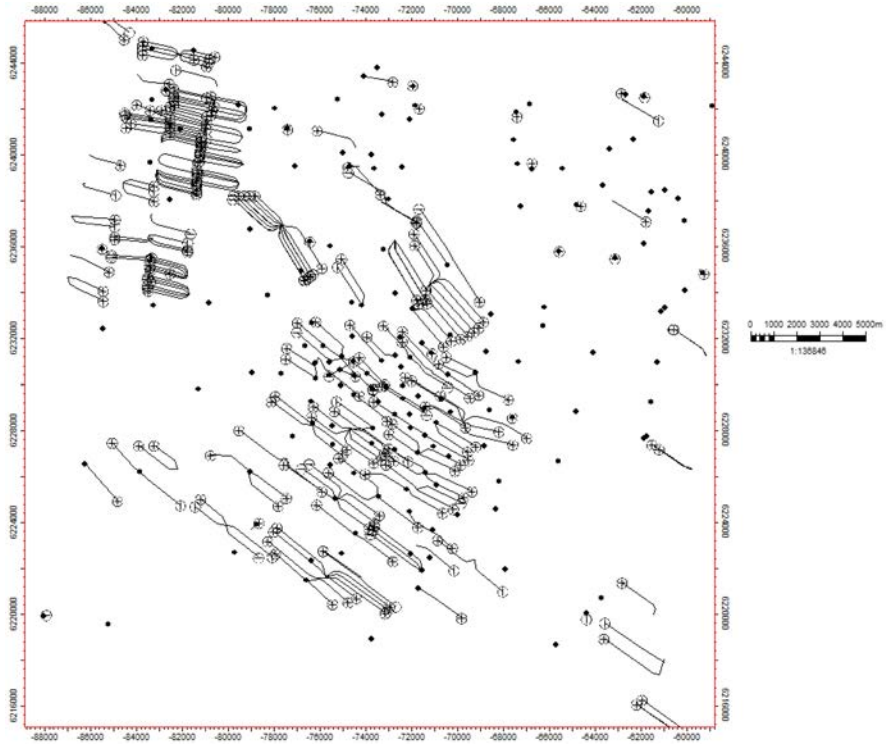

Figure 3—Aerial view of all the wells used in the study.

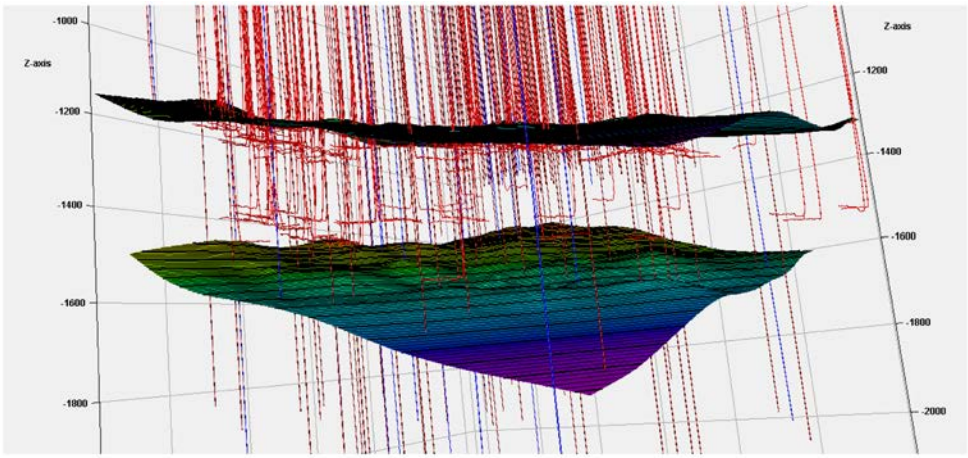

Figure 4—A 3D visualization of all wells in the study area. Red = horizontal, blue = deviated, and black = vertical.

The geological model was constructed in the Petrel geological modelling package (Schlumberger, 2018) by using well information, deviation surveys, formation tops as well as logs from public data sources. Public data requires quality checks which are performed on deviation surveys, surface coordinates, surface elevation and well tops. It was found that most publicly reported well top depths needed correction. This was

done by using the gamma ray (GR) and resistivity logs. The Montney formation top surface was generated by using ~400 well tops. Due to the high number and density of penetrations, the Montney surface is believed to have relatively low uncertainty. The Belloy Formation on the other hand only had 60 penetrations. The Montney has a much lower variation in thickness as compared to the variation in true vertical depth, so a thickness map should have more certainty than a Belloy surface generated using only 60 penetrations. The thickness map was generated using Montney thickness data from the 60 wells that penetrate the entire formation and the Belloy surface was generated by adding the thickness map to the Montney surface map. In this way, the uncertainty of the Belloy surface is low.

Having a 3D model not only helps with visualization of the study area but also provides a means to estimate reservoir thickness for the 85% of wells that do not penetrate the entire formation. The Montney Formation has a dip and varies in its true vertical depth, so a geological layer or feature at location x1, y1 and a given depth z would not necessarily be present at the same depth z at location x2, y2. The same layer is likely however to be present along the same thickness interval, so this type of measurement is a better baseline for comparison than true vertical depth.

Data Preparation

Estimating rock mechanical properties requires three sets of logs: RHOB, DTP, and DTS. The majority of the 180 vertical and deviated wells have RHOB and DTS logs but only 14 (8% of the wells) have all three sets of logs that covered >90% of the reservoir thickness. Figure 5 shows all the vertical and deviated wells without DTS marked by black points and the wells with DTS marked by a label. If accurate synthetic DTS logs could be generated for the remaining 166 vertical and deviated wells with no DTS, the certainty of a rock mechanical model could potentially increase significantly.

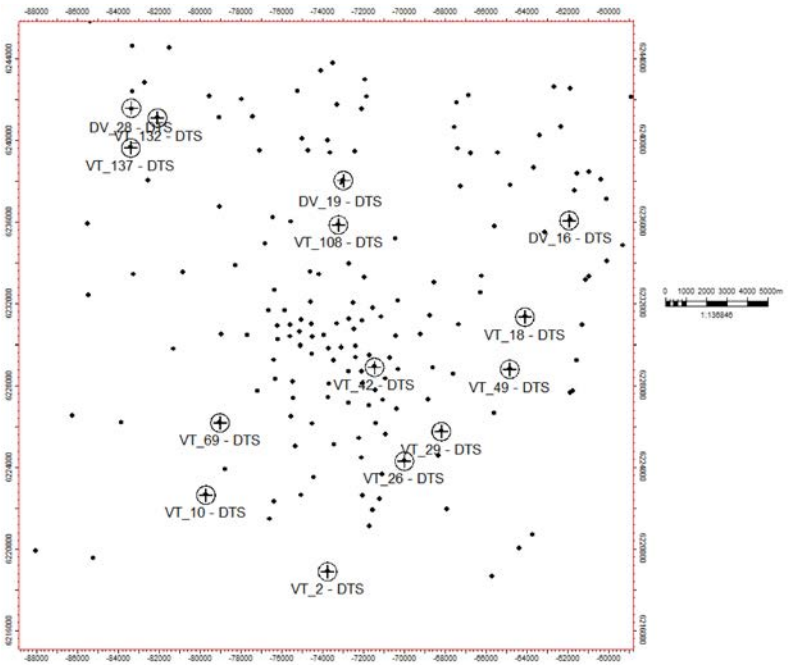

Figure 5—Areal view of the study area. Black points with no labels are wells with no DTS logs, labeled wells have the DTS logs.

The following data was extracted for the 14 wells in the Montney Formation: deviation survey, RHOB, DTP, and DTS logs. Initially, GR logs were also extracted, however through much preliminary testing it was found that GR did not improve model accuracy and, in some cases, even worsened the performance. Even after the GR log was converted to a volume of shale (Vsh) log, it was still found to be unhelpful since the GR reaction for similar layers of shale varied greatly from well to well.

Log measurements for each well have a specific resolution or measurement frequency that ranges from 1 to 6 inches (2.5-15 cm). This caused some wells to have 6 times more measurements than others for an interval of equal thickness. This is a problem since all wells need to carry the same weight to have an unbiased analysis. To combat this discrepancy all logs were converted to a resolution of 4,000 steps per 100% reservoir thickness, which is 0.025% per step. The 4,000 step scale was chosen since it represented the average number of measurements in the well logs and did not lower the resolution for the majority of the wells. Wells that did not penetrate the entire Montney Formation had an estimated formation thickness which was calculated using the top and bottom surfaces. As a result, every well had 4,000 measurements of RHOB, DTP, and DTS from top to bottom. Each step also contained x and y coordinates (constant for vertical wells). Deviated well coordinates for every step were calculated from the deviation survey. Finally, sections of the formation that had missing values of one log were removed. Although not frequent, anomalies in measurement (such as a sudden 40% spike in density) were also removed. This cleaning of data reduced the total number of points per well; however, the step change of 0.025% thickness remained the same leaving the experiment unbiased towards a certain well. The total number of data points for all 14 wells was 53,554.

Experimental Setup and Types of Neural Network Algorithms

Experimental Setup

The synthetic DTS curve was generated by using the following five inputs: x coordinate, y coordinate, depth measured in true vertical depth subsea (TVD subsea), RHOB and DTP for every available point along the Montney thickness. Since there are only 14 wells in the data set, the only suitable method to determine accuracy is the "leave-one-out" cross-validation method. In this method, 13 wells are used as training data with one well to validate (blind), the experiment is then run a total of 14 times (one for each blind well) and the results of all 14 experiments are examined. This "blind" well is not truly blind as its performance can be measured. However, it is important that an individual blind well does not change the way the tests are performed. This means the network structure and training procedure must be fixed throughout the evaluation. The network can be tuned after all 14 experiments are completed to see if an improvement is made. As the network makes 4,000-point sequence predictions, multiple metrics are required to understand the accuracy of the prediction. For this study we compare four types of accuracy metrics:MAPE – Mean Absolute Percentage Error, MaxPE – Maximum Percentage Error, amount of points with >3% error, and amount of points with >5% error.

MAPE is an indicator of the network performance on all points in the sequence, it does not reveal how much of the sequence has poor accuracy, or what the maximum error is. MaxPE shows the maximum error in the entire predicted sequence, but does not give insight into the average performance (Zhang et al. 2015). The amount of points greater than (3 or 5%) is a metric between MAPE and MaxPE and gives insight into what fraction of the predicted sequence is reliable. These metrics although helpful, are not very intuitive and the best way to compare algorithm performance is by plotting the entire synthetic log generated by the algorithm overtop the measured log. Visually comparing synthetic logs generated by different algorithms can more easily point to how well each algorithm predicts small variations and which log sections are predicted better than others.

Consistency is also important, if network performance differs too much for the same blind well experiment, there is a problem and the network cannot be trusted for true blind wells. To measure consistency, each of the 14 blind well experiments were reset and trained three times, without modifying any network parameters or training procedures. Metrics including the maximum minus minimum error and the standard deviation of errors of the three runs were measured. Running multiple experiments on each blind well is also a good way to average the results by removing really good or really bad results.

Comparisons

To test if an algorithm has statistical significance, it must be compared to a baseline such as: taking the average DTS/DTP ratio of training wells and applying it to the blind well, or taking the average DTS minus DTP value of training wells and applying it to the blind well.

The networks are also compared to the DTS estimates of two popular empirical correlations: Castagna et al. (1985) for shales: Vs(km/s) = 0.77 Vp − 0.8674 and Han et al. (1986) for shaly sandstones:Vs(km/s) = 0.85 Vp − 1.14

Networks used in the experiments

Two types of networks were chosen to run experiments: a simple feedforward network (ANN) and a convolutional-recurrent hybrid (c-RNN). The networks were programmed in Python using the keras library (Chollet, 2015). The ANN was chosen as it is the simplest type of network and a good baseline to compare to the more complicated c-RNN. The c-RNN hybrid network was chosen because it combines the speed and ability to process large amounts of data of a convolutional network with the sequence processing ability of a recurrent network. Traditional RNNs are only able to hold memory a few hundred steps and since each of the wells has 4,000 steps in the input sequence, a standalone RNN would not be able to predict with great accuracy. Convolutional networks are able to convert long sequences into shorter ones of higher-level features, which gives a c-RNN hybrid the ability to process sequences with thousands of steps (Chollet, 2017). Figure 6 is a visual representation of the input and output data format of the c-RNN.

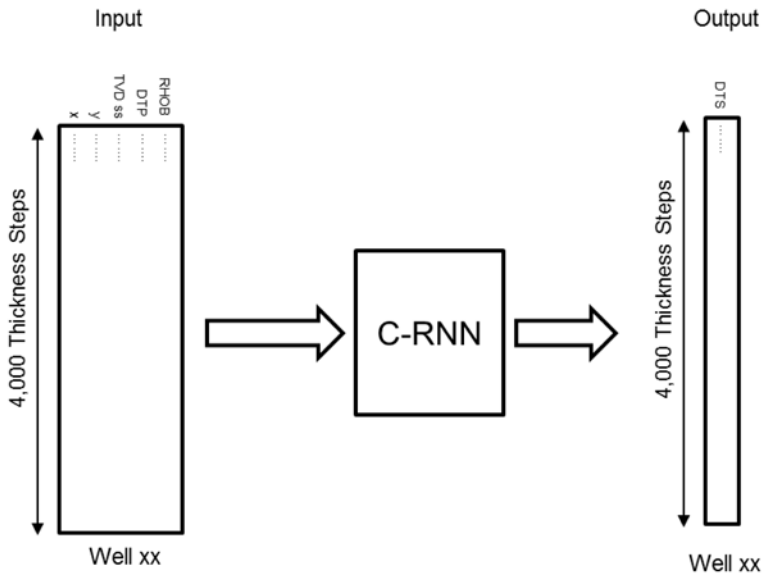

Figure 6—Format of input and output data of the c-RNN.

Overfitting

Deep neural networks are good at learning complicated relationships between inputs and outputs; however, if trained too long the networks start to confuse noise with signals and start to overfit to the training set which results in lower generalization capabilities (Srivatava et al. 2014). This is especially problematic when the amount of training data is limited as in the case in this study.

Since the aim of this study is to generate synthetic DTS for wells with no actual DTS there should be a way to determine the stopping point of the training without looking at blind well performance. Most machine learning experiments are run with three sets of data: training, validation and test. The model is trained on the training set, the validation set is used to monitor overfitting and the test set is then run after network has finished training. The typical approach is to stop the training at the point where validation error is at a minimum. This approach is good when there are at least a few hundred samples available; this study only

has 14 wells in total. Running multiple experiments showed that the point at which the validation error was at a minimum was inconsistent between different blind wells and in repeat trials of the same blind well; this implied that a way to limit overfitting was needed. To limit overfitting, we maximized the batch size and applied dropout.

Dropout is one of the most effective and commonly used methods to reduce overfitting (Chollet, 2017). The term "dropout" refers to the temporary removal (setting to zero) of a fraction of nodes from a network along with their incoming and outgoing connections during training. The units to drop are chosen at random with a probability of p which is known as the dropout rate which is usually set between 0.2 and 0.5 (Chollet, 2017). At test time, no units are dropped, instead layer output values are scaled down by p. At its core, dropout adds randomness which breaks up patterns that are not significant (Srivatava et al. 2014).

Batch size refers to the number of training examples that are seen before a network update is made. Batch size comes in three options (Brownlee, 2017): batch mode - batch size is equal to the total number of samples in the training set, mini-batch mode - batch size is smaller than that in batch mode but greater than one, and stochastic mode - batch size is equal to one. Figure 7 plots how the validation error evolves with different batch sizes (as a percent of the maximum batch size) for the VT_42 blind well test. The plot reveals that fluctuations in blind error become less aggressive and overfitting becomes less of an issue with larger batch sizes. For these reasons it was decided to run all experiments in full batch mode.

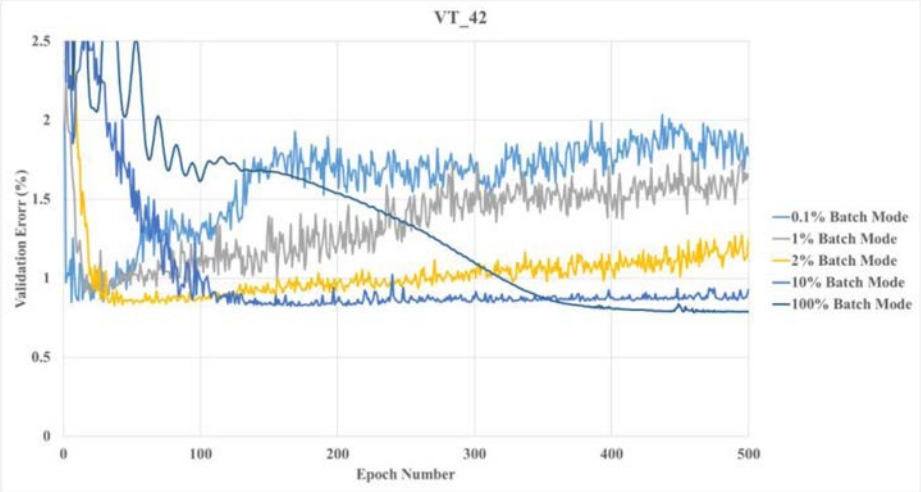

Figure 7—Effect of batch size on the evolution of the validation error for blind well VT_42.

Stopping procedure

ANN. The ANN is meant to serve as a baseline to compare the c-RNN too. If the c-RNN performs better than the best possible ANN results than the complexity and longer run time of the c-RNN are justifiable. To reach the best possible result of the ANN a cheat run was performed for every blind well. In the cheat run, the blind well error was observed during training and the network was forced to stop when the blind error value was at its lowest. The point at which the blind error is minimum was unique for every blind well and the only way to find it was to observe the blind error rate at every epoch. Each blind well test was reset and run multiple times and the run with the lowest blind error was used to generate a synthetic DTS log. This type of procedure is only meant to test the maximum potential of the network and is not applicable when generating synthetic DTS for wells with no DTS since the error on these wells cannot be observed.

c-RNN. Figure 8a shows how the training and validation error evolves over 1,000 training epochs for a blind well experiments. "Tremors" in both the training set and blind well error start to occur at fairly regular intervals after ~400 epochs and the error rate seems to stabilize at a very low constant error between them.

These tremors were found to be universal across all wells. A long run was performed on one of the wells to see if this trend continued. Figure 8b shows the train and test error for 20,000 iterations of data. Overfitting seemed to have stopped and the blind well error is more or less constant between tremors. After about 2,000 iterations the tremors in test data get very large but still converge back to a low and stable error rate. The tremors in the blind well point to a possibility for improvement in accuracy if a way to stop the algorithm at the low point of the blind tremor is found.

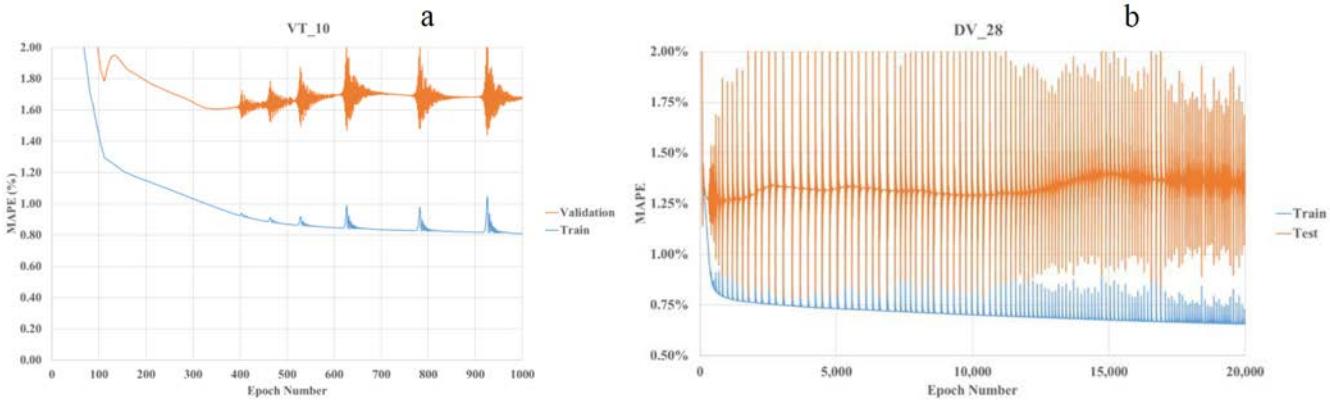

Figure 8—a (left) – 1,00 iterations of the training and validation error for well VT_10, Figure 8b (right) = 20,000 iterations for well DV_28

Since the validation error does not exist when generating synthetic DTS curves for wells with no DTS, one can only use the training error to determine when to stop the network. From the figures it is seen that the tremors in both the training set and blind well occur nearly at the same time (validation tremor starts a bit earlier and ends a bit later than the training tremor). Since the training and validation tremors where found to coincide with one another, a simple empirical procedure was developed to determine the stopping criteria for the c-RNN: run network for 1,000 epochs (when network has stopped improving results); if training error is stable and far enough away from the previous tremor, stop training; if training error is experiencing a tremor or very close to previous tremor, continue running network for 50-100 more epochs (but before next tremor occurs) until you reach a state in step 2, stop training; if another tremor occurs, repeat step 3 until you reach a state at step 2.

Finally, to see how close this procedure comes to the maximum potential of this network, a cheat run was performed for every blind well. Like the ANN cheat run, the blind well error was observed during training and the network was stopped at the lowest blind value.

Summary of Procedure

Figure 9 shows the procedure required to generate a trained version of the synthetic DTS tool presented in this study. Figure 10 shows how to use the trained tool to generate synthetic DTS curves. The training and generating procedures are not limited to the Montney and can be applied to any formation of interest.

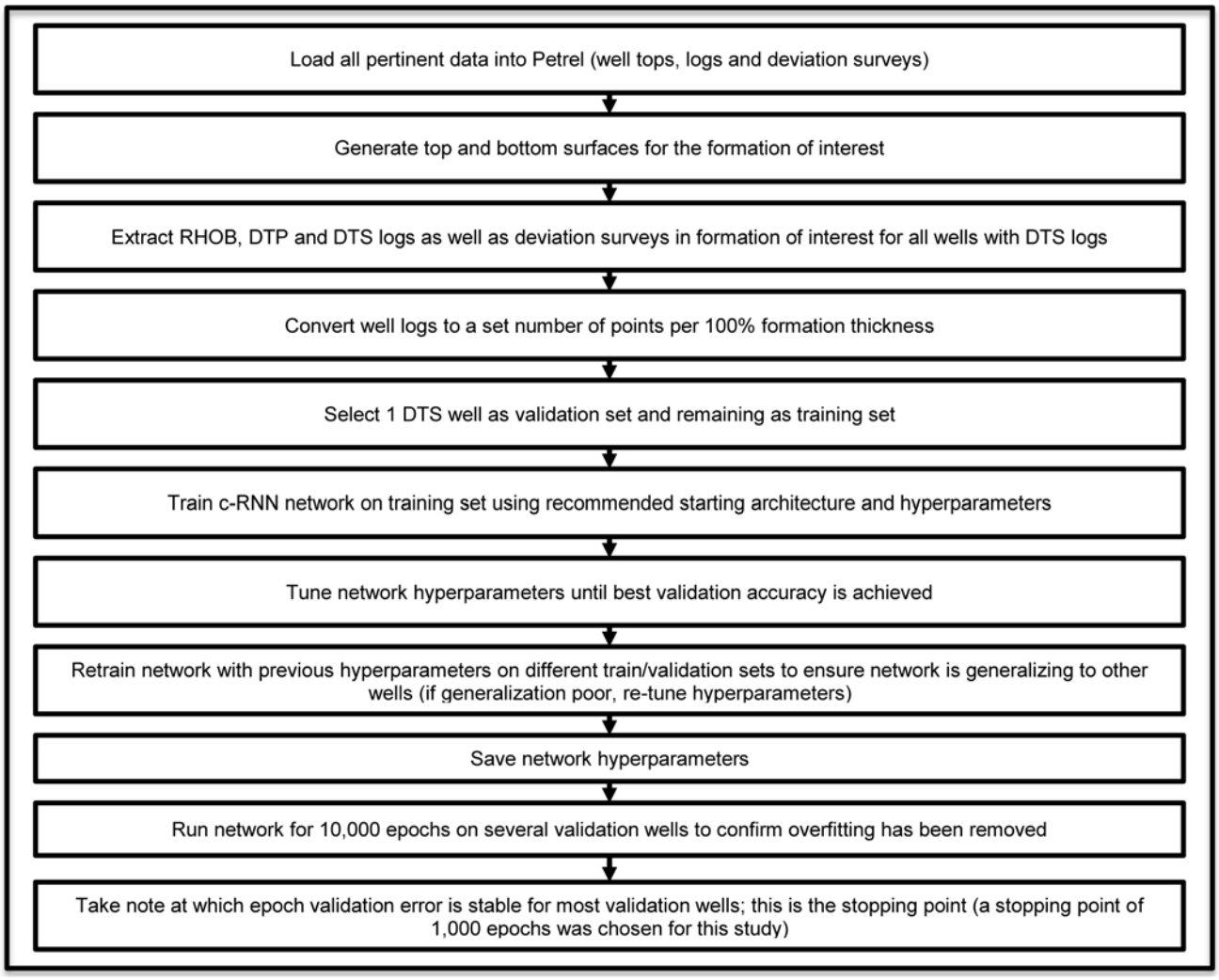

Figure 9—General procedure for training the synthetic DTS tool that can be applied to any formation.

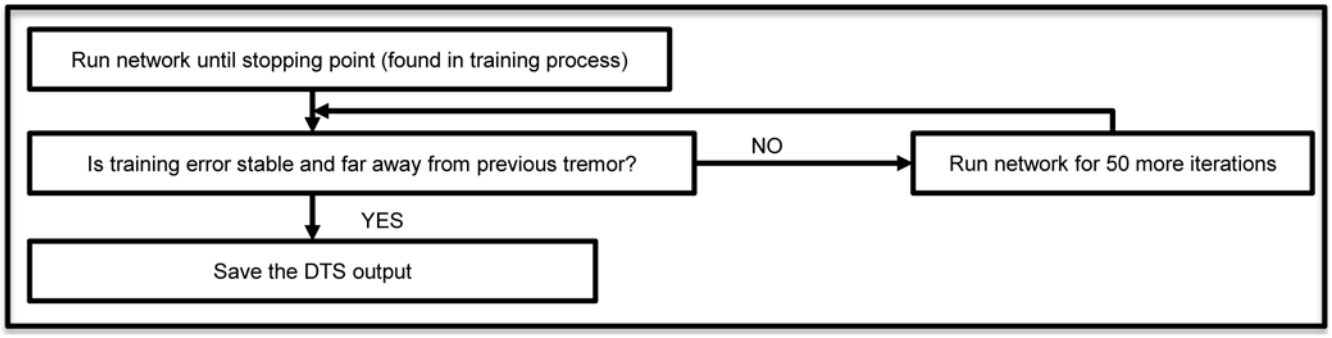

Figure 10—Process for using the training tool to generate synthetic DTS curves.

Results and Discussion

Tables 1a to 1e summarize the results of the leave-one-out experiments. Tables 1a to 1d compare the four different metrics for the two empirical correlations, the two simple baselines, the best possible results achieved with the ANN cheat run, the average of the three c-RNN runs along with the best possible result achieved with the c-RNN cheat run. Table 1e show the the difference between maximum and minimum error values of of the three c-RNN runs for each blind well. The MAPE and MaxPE (Tables 1a and 1b) are not the best metrics as they do not describe the amount of local errors, but they do provide more information

about the results. The more revealing metric is the fraction of points >3% error (Table 1c) and the visual comparison.

Table 1—Results of the study comparing the performance of the various methods of generating synthetic DTS curves.

Table 1a - MAPE

Well	Castagna et al.	Han et al.	DTS/DTP	DTS-DTP	ANN Cheat	Ave of 3 c-RNN Runs	c-RNN Chea
VT_02	2.80%	2.19%	1.24%	1.03%	0.91%	1.00%	0.83%
VT_10	2.31%	3.17%	1.61%	1.53%	1.88%	1.76%	1.27%
VT_18	1.79%	3.47%	1.35%	0.93%	0.62%	0.66%	0.50%
VT_26	3.39%	2.55%	1.80%	2.29%	1.12%	1.02%	0.83%
VT_29	3.80%	1.89%	1.87%	2.15%	1.16%	1.27%	1.09%
VT_42	7.05%	3.27%	4.07%	3.11%	1.17%	1.02%	0.85%
VT_49	1.53%	3.95%	1.66%	1.01%	1.81%	1.40%	0.66%
VT_69	6.41%	2.84%	3.29%	2.11%	1.28%	1.20%	0.91%
VT_108	2.05%	3.01%	1.43%	1.31%	0.92%	0.99%	0.82%
VT_132	2.73%	2.73%	1.36%	1.37%	1.01%	0.95%	0.75%
VT_137	4.59%	2.93%	2.17%	1.37%	1.26%	1.99%	1.31%
DV_16	1.60%	4.11%	2.17%	1.75%	0.74%	1.01%	0.75%
DV_19	2.57%	4.92%	2.95%	2.32%	1.59%	1.27%	1.18%
DV_28	2.67%	3.12%	1.78%	2.15%	1.62%	1.52%	0.85%
Total	**3.21%**	**3.14%**	**2.03%**	**1.73%**	**1.23%**	**1.22%**	**0.89%**

Table 1b - MaxPE

Well	Castagna et al.	Han et al.	DTS/DTP	DTS-DTP	ANN Cheat	Ave of 3 c-RNN Runs	c-RNN Cheat
VT_02	8.37%	9.50%	4.96%	4.06%	4.36%	4.18%	3.52%
VT_10	10.44%	14.00%	12.90%	13.32%	14.63%	14.00%	12.98%
VT_18	6.85%	10.93%	5.87%	5.71%	4.72%	4.44%	3.82%
VT_26	14.76%	11.14%	8.45%	6.15%	5.09%	5.32%	5.52%
VT_29	11.42%	7.70%	7.03%	6.59%	5.11%	4.87%	4.12%
VT_42	16.84%	13.59%	9.34%	7.61%	5.66%	6.07%	5.18%
VT_49	6.02%	10.54%	6.31%	5.50%	4.79%	7.18%	4.32%
VT_69	15.84%	12.57%	8.54%	6.31%	3.87%	4.43%	3.73%
VT_108	6.79%	11.06%	7.10%	5.93%	3.97%	4.03%	3.77%
VT_132	12.15%	11.89%	6.65%	8.88%	6.07%	7.08%	6.75%
VT_137	14.48%	11.36%	7.13%	7.47%	6.15%	5.98%	6.74%
DV_16	8.73%	13.20%	6.73%	5.88%	5.71%	5.05%	5.51%
DV_19	10.59%	14.82%	10.53%	7.53%	7.32%	6.32%	5.67%
DV_28	11.49%	14.23%	8.64%	9.65%	10.07%	10.36%	11.32%
Total	**16.84%**	**14.82%**	**12.90%**	**13.32%**	**14.63%**	**14.00%**	**12.98%**

Table 1c – Percentage of points above 3% error							
well Well	**Castagna et al.**	**Han et al.**	**DTS/DTP**	**DTS-DTP**	**ANN Cheat**	**Ave of 3 c-RNN Runs**	**c-RNN Cheat**
VT_02	41.53%	27.17%	5.39%	2.12%	1.97%	1.92%	0.43%
VT_10	32.30%	50.65%	10.33%	9.37%	19.00%	15.50%	8.58%
VT_18	14.60%	53.84%	4.25%	1.18%	0.15%	0.18%	0.13%
VT_26	47.01%	34.42%	18.33%	25.63%	3.19%	2.63%	1.15%
VT_29	61.24%	19.55%	19.08%	25.81%	4.80%	7.87%	4.20%
VT_42	94.29%	42.67%	69.03%	56.65%	3.55%	4.01%	1.99%
VT_49	13.59%	65.92%	11.69%	1.71%	10.88%	2.96%	0.68%
VT_69	92.09%	32.35%	49.67%	18.99%	3.07%	3.34%	0.60%
VT_108	23.95%	44.54%	8.97%	5.84%	2.38%	1.71%	0.54%
VT_132	35.08%	43.78%	6.70%	7.08%	0.78%	0.97%	1.20%
VT_137	62.19%	38.55%	28.08%	5.22%	3.57%	24.84%	5.25%
DV_16	13.36%	68.09%	21.18%	16.37%	0.89%	3.13%	1.01%
DV_19	35.29%	67.53%	42.41%	28.32%	13.14%	5.13%	5.46%
DV_28	31.48%	48.66%	13.05%	17.75%	8.35%	5.50%	2.10%
Total	**42.30%**	**45.42%**	**21.45%**	**15.37%**	**5.26%**	**5.66%**	**2.31%**

Table 1d – Percentage of points above 5% error							
Well	**Castagna et al.**	**Han et al.**	**DTS/DTP**	**DTS-DTP**	**ANN Cheat**	**Ave of 3 c-RNN Runs**	**c-RNN Cheat**
VT_02	14.16%	5.49%	0.00%	0.00%	0.00%	0.00%	0.00%
VT_10	8.98%	16.23%	3.81%	3.13%	5.51%	4.64%	3.11%
VT_18	2.13%	21.71%	0.15%	0.10%	0.00%	0.01%	0.00%
VT_26	20.03%	9.26%	5.65%	1.71%	0.03%	0.09%	0.08%
VT_29	25.81%	3.85%	4.95%	4.93%	0.08%	0.00%	0.00%
VT_42	75.28%	21.73%	26.90%	5.40%	0.17%	0.49%	0.06%
VT_49	0.85%	25.53%	0.53%	0.13%	0.00%	0.05%	0.00%
VT_69	65.80%	15.14%	16.33%	0.52%	0.00%	0.00%	0.00%
VT_108	2.87%	19.19%	0.38%	0.41%	0.00%	0.00%	0.00%
VT_132	14.03%	10.55%	0.25%	0.68%	0.20%	0.30%	0.60%
VT_137	39.38%	17.82%	6.06%	0.23%	0.10%	0.74%	0.15%
DV_16	1.77%	32.82%	0.63%	0.51%	0.25%	0.03%	0.10%
DV_19	12.51%	42.26%	16.96%	4.51%	1.99%	0.34%	0.33%
DV_28	16.35%	16.33%	0.90%	3.20%	1.15%	0.52%	0.53%
Total	**21.09%**	**18.22%**	**5.67%**	**1.76%**	**0.63%**	**0.49%**	**0.34%**

Table 1e – Maximum minus minimum error of the three c-RNN runs				
Well	**MAPE**	**MaxPE**	**Percentage of points > 3% error**	**Percentage of points > 5% error**
VT_02	0.29%	1.24%	2.58%	0.00%
VT_10	0.78%	1.36%	11.94%	0.62%
VT_18	0.18%	1.07%	0.08%	0.03%
VT_26	0.14%	0.70%	0.39%	0.11%
VT_29	0.07%	0.32%	0.95%	0.00%
VT_42	0.42%	2.14%	4.89%	1.25%
VT_49	0.57%	4.67%	3.06%	0.05%
VT_69	0.50%	1.28%	5.28%	0.00%
VT_108	0.03%	0.55%	0.56%	0.00%
VT_132	0.52%	0.89%	0.08%	0.03%
VT_137	0.29%	0.21%	8.64%	0.30%
DV_16	0.32%	0.76%	1.49%	0.05%
DV_19	0.06%	0.84%	0.50%	0.06%
DV_28	0.16%	0.49%	1.20%	0.03%
Average	**0.31%**	**1.18%**	**2.97%**	**0.18%**

The results show that the empirical correlations did not work for this study. The likely reason for this is because the empirical correlations were developed from a very broad (global) set of reservoirs and do not translate well to specific (local) reservoirs well. The simple baselines performed about twice as well as the empirical formulas, this is no surprise as they were derived from the average of 13 training wells, making them very local and only applicable to this particular study area. Both neural networks show statistical significance as they outperformed the simple baselines and the empirical correlations. The metrics show that the c-RNN performance is similar that of the ANN cheat run; however, a visual comparison reveals that the c-RNN has much better predictive capabilities. Figure 11 compares the ANN cheat and average c-RNN synthetic logs for the central well VT_42. The visual comparison shows how the average c-RNN run can predict the small variations much better than the best result achieved by the cheat ANN run.

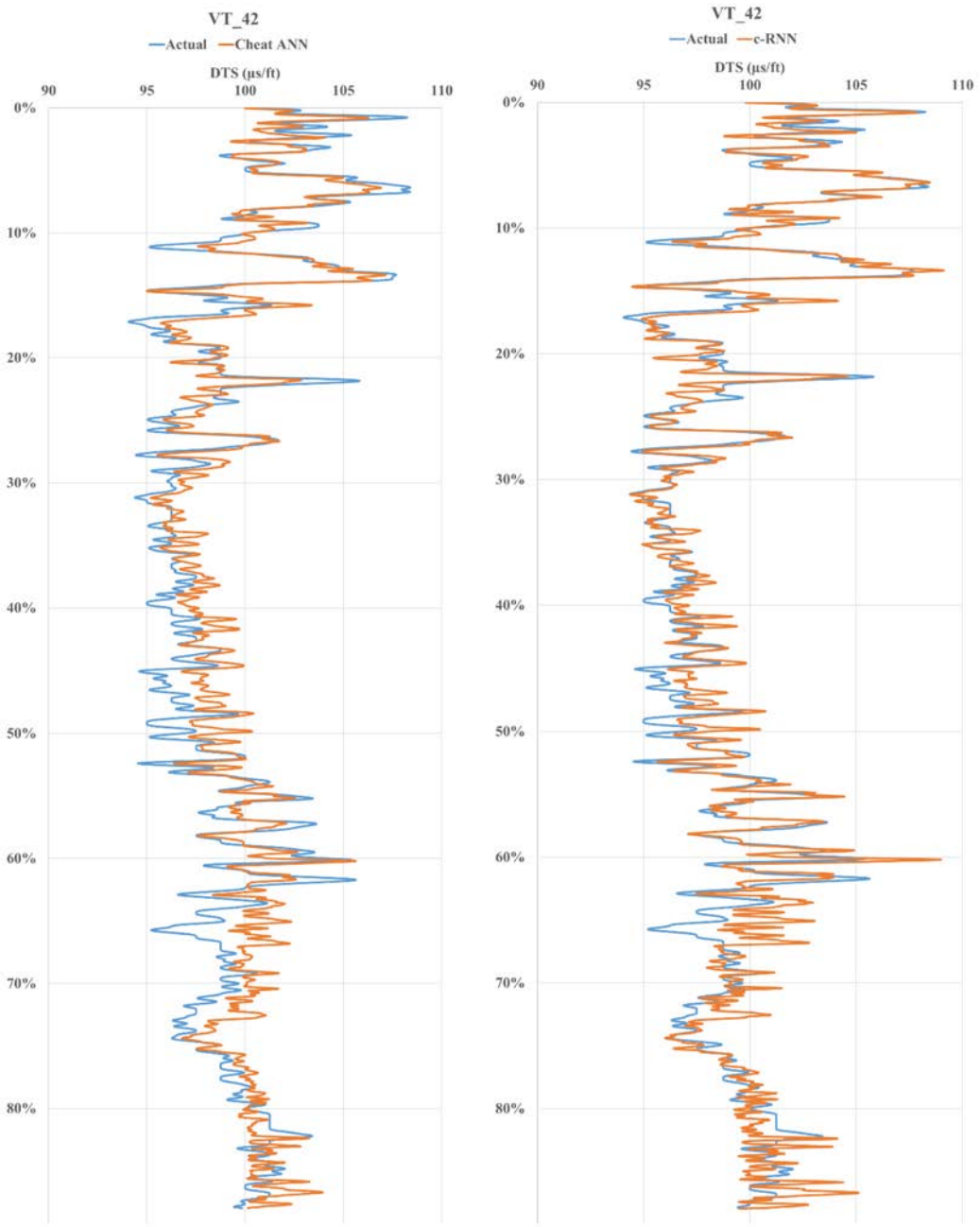

Figure 11—Synthetic logs generated by the cheat ANN (left) and the average of three c-RNN (right) runs compared to the true logs for the most important well - VT_42.

Conclusions

A procedure to generate highly accurate synthetic DTS curves has been developed. This procedure can act as a cost effective and fast alternative to running actual DTS logs. This does come with the condition that the area of interest is small enough and that there are enough wells with true DTS measurements available. Generation of a synthetic DTS measurement at a given point along the reservoir thickness only required five inputs: x and y coordinates, depth subsea, RHOB and DTP. The GR log was tested and was found to not improve the performance of the network. As shown by the cheat run, the procedure does still have room for improvement. Some potential reasons for the error in the synthetic DTS curves are:

- 13 training wells are not enough to learn every possible signal.
  MAPE=0.62%

MAPE=0.66%

- Some signals are caused by variables not captured in the study (ie. fractures, stress profiles, pore fluids, high porosity, shale, thin beds, gas presents, high total organic content (TOC)).
- DTS measurement error is caused by the tool itself or poor borehole conditions.

Although the network did perform poorly on two of the blind well tests (VT_137 and VT_10) it is more important to look at the performance of wells located closer to the majority of vertical and horizontal wells. As there is no real way to measure synthetic DTS curve accuracy for wells with no DTS one can only rely on the results of the network for nearby wells. Most wells are located in the center of the study area and most of the horizontal wells reside in the top 25% thickness and the network has shown to be able to generate fairly accurate curves for that region of the study. The c-RNN has also shown to have fairly consistent results between the runs, but it is still suggested to take the average of three runs when generating synthetic curves for wells with no DTS. This procedure is not limited to the Montney Formation, as long as the area is small enough and contains enough wells with DTS logs it could be applied to any other formation. Even in this small window of the Montney, the 250 existing horizontal wells only cover a small fraction of the thick reservoir and there is remaining opportunity for further development. Having a predictive procedure like this could potentially give the user a way to optimize where to drill and how to complete a horizontal well before it is actually drilled. If the number of wells per area is large enough the procedure could be applied to any other formation.

Acknowledgements

The authors acknowledge support from Sproule Associates Limited as well as the Natural Science and Engineering Research Council of Canada, the University of Calgary Canada First Research Excellence Fund program titled the Global Research Initiative in Sustainable Low Carbon Unconventional Resources.

References

1. Al-Anazi, A. F. and Gates, I. D. On Support Vector Regression to Predict Poisson's Ratio and Young's Modulus of Reservoir Rock, Chapter 5 in Cranganu et al. (eds.), Artificial Intelligent Approaches in Petroleum Geosciences. 2015, Springer International Publishing, DOI 10.1007/978-3-319-16531-8_5.
2. Barre, D. (2008). Mechanical Properties Log Processing and Calibration (accessed 29 May 2019)
3. Bhande, A., 2018, What is underfitting and overfitting in machine learning and how to deal with it, https://medium.com/greyatom/what-is-underfitting-and-overfitting-in-machine-learning-and-how-to-deal-with-it-6803a989c76,(Accessed June 1, 2019)
4. Brownlee, J., 2017, A Gentle Introduction to Mini-Batch Gradient Descent and How to Configure Batch Size, https://machinelearningmastery.com/gentle-introduction-mini-batch-gradient-descent-configure-batch-size/, (Accessed June 1, 2019)
5. Castagna, J. P., 1985, Shear-wave time-average equation for sandstones: Presented at the 55th Ann. Internat. Mtg., Soc. Expl. Geophys.
6. Chollet, F. (2015) keras, GitHub. https://github.com/fchollet/keras
7. Chollet, F. Deep Learning with Python, Manning Publications Co., Greenwich, CT, 2017
8. Eskandari H, Rezaee MR, Mohammadnia M (2004) Application of multiple regression and artificial neural network techniques to predict shear wave velocity from well log data for a carbonate reservoir, south-west Iran. In: *CSEG RECORDER*, pp 42 – 48
9. Fjar, E.; Holt, R. M.; Raaen, A. M.; Risnes, R. Petroleum Related Rock Mechanics, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 2008; pp. 309 – 339.

10. Hadi, F., and Nygaard, R. 2018. Shear Wave Prediction in Carbonate Reservoirs: Can Artificial Neural Network Outperform Regression Analysis? ARMA Paper 18-905 presented at the 52nd US Rock Mechanics / Geomechanics Symposium. Seattle, Washington, US, 17-20 June 2018.
11. Han, De-hua, Nur, A., and Morgan, D., 1986, Effects of porosity and clay content on wave velocities in sandstones: Geophysics, **51**, 2093–2107.
12. Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. arXiv:1412.3555 [cs], December 2014. URL http://arxiv.org/abs/1412.3555
13. Maleki, S., A. Moradzadeh, R. Ghavami, and R. Gholami. 2014. Prediction of shear wave velocity using empirical correlations and artificial intelligence methods. *NRIAG Journal of Astronomy and Geophysics*, **3**(1), 70–81. https://doi.org/10.1016/j.nrjag.2014.05.001
14. National Energy Board, 2018, Market Snapshot: Evolving technology is a key driver of performance in modern gas wells: a look at the Montney Formation, *one of North America's biggest gas resources*, http://www.neb-one.gc.ca/nrg/ntgrtd/mrkt/snpsht/2018/04-04-1vlvngtchnlg-eng.html (Accessed June 1, 2019)
15. National Energy Board, 2019, Canada's Energy Future 2018 Supplement: Conventional, Tight, and Shale Oil Production, http://www.neb-one.gc.ca/nrg/ntgrtd/ftr/2018cnvntnll/index-eng.html, (Accessed June 1, 2019)
16. Parapuram, G.; Mokhtari, M.; Ben Hmida, J. An Artificially Intelligent Technique to Generate Synthetic Geomechanical Well Logs for the Bakken Formation. Energies 2018, **11**, 680.
17. Rajabi M, Bohloli B, Ahangar EG (2010) Intelligent approaches for prediction of compressional, shear and Stoneley wave velocities from conventional well log data: a case study from the Sarvak carbonate reservoir in the Abadan Plain (Southwestern Iran). *ComputGeosci* **36** (5): 647 – 664
18. Reynolds, M., Bachman, R., Peters, W., 2014: "A Comparison of the Effectiveness of Various Fracture Fluid Systems Used in Multi-Stage Fractured Horizontal Wells: Montney Formation, Unconventional Gas"; SPE 168632 presented at the Hydraulic Fracturing Technology Conference, The Woodlands TX, Feb. 4 – 6.
19. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: a simple way to prevent neural networks from overfitting. J. Machine Learning Res. **15**, 1929 – 1958 (2014)
20. Schlumberger 2018, Petrel E&P Software Platform.
21. Wang, S., and Chen, S. 2016. Evaluation and Prediction of Hydraulic Fractured Well Performance in Montney Formations Using a Data-Driven Approach. Presented at the SPE Western Regional Meeting, Anchorage, Alaska, 23–26 May. SPE-180416-MS. http://dx.doi.org/10.2118/180416-MS.
22. Zhang, D., Yuntian, C., Jin, M., 2018. Synthetic well logs generation via Recurrent Neural Networks. *Pet. Explor. Dev.* **45**, 629 – 639.
23. Zhang, J., Florita, A., Hodge, B. M., Lu, S., Hamann, H. F., Banunarayanan, V., & Brockway, A. M. (2015). A suite of metrics for assessing the performance of solar power forecasting. *Solar Energy*, **111**, 157–175. http://dx.doi.org/10.1016/j.solener.2014.10.016