EZENKWU, C.P., GUNTORO, J., STARKEY, A., VAZIRI, V. and ADDARIO, M. 2023. Automated well-log pattern alignment and depth-matching techniques: an empirical review and recommendations. *Petrophysics* [online], 64(1), pages 115-129. Available from: https://doi.org/10.30632/PJV64N1-2023a9

Automated well-log pattern alignment and depth-matching techniques: an empirical review and recommendations.

EZENKWU, C.P., GUNTORO, J., STARKEY, A., VAZIRI, V. and ADDARIO, M.

2023

OpenAIR @RGU

This document was downloaded from https://openair.rgu.ac.uk

CC BY

PETROPHYSICS, VOL. 64, NO. 1 (FEBRUARY 2023); PAGES 115–129; 11 FIGURES, 5 TABLES. DOI:10.30632/PJV64N1-2023a9

Automated Well-Log Pattern Alignment and Depth-Matching Techniques: An Empirical Review and Recommendations

Chinedu Pascal Ezenkwu[1], John Guntoro[1], Andrew Starkey[1], Vahid Vaziri[1], and Maurillio Addario[2]

ABSTRACT

Well logging has been an integral part of decision making at different stages (drilling, completion, production, abandonment) of a well's history. However, the traditional human-reliant approach to well-log interpretation, which has been the most common practice in the industry, can be time consuming, subjective, and incapable of identifying fine details in log curves. Previous studies have recommended automated approaches as a candidate for addressing these challenges. Despite the progress made so far, what is not yet clear from the existing literature is the extent to which these automated approaches can dispense with human interventions in real-life scenarios. This paper presents an empirical review of different depth-matching techniques in real-life timelapse well logs, primarily focusing on gamma ray and the extent to which the outcomes of these techniques match the results from a human expert. Specifically, the performances of dynamic time warping (DTW), constrained DTW (CDTW), and correlation optimized warping (COW) are investigated. The experiments also consider the effects of filtering and normalization on the performance of each of the techniques. Concerning the correlations of each technique's outcome with the reference data and an expert-generated outcome, this research identifies and discusses its key challenges, as well as provides recommendations for future research directions. Although the COW technique has its limitations, as discussed in this paper, our experiments demonstrate that it shows more potential than DTW and its variants in the well-log pattern alignment task. The work entailed by this research is significant because identifying and discussing the limitations of these techniques is vital for solution-oriented future research in this area.

INTRODUCTION

Well logging is a fundamental decision-support practice in the oil and gas industry because of its crucial role in subsurface exploration and formation evaluation. Well logs provide insightful petrophysical and geomechanical information, which can be significant at different stages of a well's life cycle (Le et al., 2019; Torres Caceres et al., 2022). However, the acquisition of logs is often characterized by several uncertainties and limitations, making preprocessing a desideratum for the data analysis phase. While improvements in the quality of logging suites can help minimize noise and uncertainty due to random and systematic errors, depth misalignments of logs recorded at different passes or with disparate logging tools in the same well, and different resolutions, different noise levels between different logging passes have remained a challenging problem within the industry (Moore et al., 2011; Zimmermann et al., 2018; Bittar et al., 2021). Depth misalignments can be caused by several factors, depending on the logging technique (logging while drilling or wireline logging) used (Storey and Bolt, 2016; Torres Caceres et al., 2022). These factors include differences in weather conditions, varying sampling rates between tool types and logging passes, and friction or stick-slip between the wireline cable and the borehole wall, which can be higher if the borehole wall is rough or contains mudcake. The oil and gas industry, so far, relies hugely on the judgments of human log analysts to manually synchronize mismatched logs before providing interpretations for them. Nonetheless, this traditional human-dependent depth-alignment process is subjective, time consuming, and cannot match insightful minutiae within log curves. These limitations of the manual well-logging process have driven considerable attention toward automating the depth matching of logs as a potential solution to the problem. The automation of well-log preprocessing can speed up the well-

Manuscript received by the Editor May 12, 2022; revised manuscript received September 18, 2022; manuscript accepted October 25, 2022.
[1]School of Engineering, University of Aberdeen, UK, chineduezenkwu@abdn.ac.uk; r01jpsg@abdn.ac.uk; a.starkey@abdn.ac.uk; vahid.vaziri@abdn.ac.uk
[2]ANSA Data Analytics, maurillio.addario@ansa-data.com

integrity process, reduce costs, and make the process more practicable to nonspecialists.

Due to the availability of big historical data, high compute power, and the recent development of several powerful artificial intelligence (AI) algorithms, the oil and gas industry now leverages digital technology on virtually all fronts, including reservoir depth matching. In the literature, different machine-learning paradigms have found applications in the automation of well-log preprocessing and interpretations (Choubey and Karmakar, 2021; Koroteev and Tekic, 2021; Sircar et al., 2021). For example, Wang et al. (2020) and Hong and Kang (2020) both applied deep neural networks, a supervised learning technique, to the log-matching problem. Deep reinforcement learning has also been applied to the same task (Bittar et al., 2020, 2021). Although these techniques have been reported to have performed well at this task, they have several issues capable of limiting their real-life applications. Firstly, training these algorithms requires a large amount of data, which cannot always be guaranteed in geological problems (Downton and Hampson, 2019), and high-performance computing hardware, which can contribute substantially to climate change (Strubell et al., 2019). To illustrate, despite being described as the best in natural language processing (NLP) tasks (Edwards, 2021; Wang et al., 2021), the carbon dioxide emissions from training and applying a transformer (a type of deep neural network) are even more substantial than the lifetime emissions of an automobile (Strubell et al., 2019). With calls from different quarters to decarbonize the energy system, within which the oil and gas industry has a significant role to play, the industry has to reconsider operations that can further contribute to global warming (Nurdiawati and Urban, 2022; Smil, 2022; Zachmann et al., 2022). Moreover, due to the difficulty in implementing complex algorithms such as deep neural networks from scratch, the AI community depends, to a large extent, on third-party libraries such as TensorFlow (Abadi et al., 2016), Pytorch (Paszke et al., 2019), Theano (Al-Rfou et al., 2016), and Keras (Gulli and Pal, 2017). However, the combination of novel deep-learning approaches with industrial software is limited.

This paper drives towards an optimal but easy-to-implement and deploy lightweight solution to the depth-matching problem. The insights from this study are from collaborative research between academia and the industry on addressing practical challenges with automating the well-log alignment process. Due to the problems associated with the powerful but expensive AI techniques, as identified in the preceding paragraph, this paper focuses on dynamic time warping (DTW), constrained DTW (CDTW) (Sakurai et al., 2005; Müller, 2007), and correlation optimized warping (COW) (Nielsen et al., 1998). Our experiments with these algorithms on real-life well logs reveal that none of these algorithms can eliminate the need for a human specialist. Nonetheless, COW demonstrates to have more potential than DTW and CDTW. This paper discusses practical problems identified with these algorithms and recommends strategies to improve their performances.

THEORY

Mathematical Description of the Depth-Matching Problem

Garcia Manso (2020) presents the mathematical description of the depth-matching problem as follows—$s(t)$ and $r(t)$ are continuous functions representing the survey and the reference, respectively, as functions of a variable depth, t. A warping function $w(t)$ represents the misalignment between the survey and the reference curves. The most common misalignment is an offset. For example, with an offset of k meters between the curves, the warping function would be $w(t) = t - k$, and the relation between the curves is that $r(t)$ and $s(w(t)) = s(t - k)$ are similar patterns. Since $w(t)$ is unknown in reality, methods for estimating warping functions have attracted lots of research attention.

In practice, both the survey and the reference are discretized into a set of noisy samples. Vectors **s** and **r** denote the data points from the survey and the reference, respectively. The samples are taken at a certain depth, t_i. To avoid ambiguity, $s(t_{s,i})$ is the measurement recorded for the survey at $t_{s,i}$, and $r(t_{r,i})$ is the measurement for the reference at $t_{r,i}$. These are expressed in Eqs. 1 and 2.

$$\mathbf{r} = \begin{pmatrix} r(t_{r,1}) \\ \vdots \\ r(t_{r,N}) \end{pmatrix} \tag{1}$$

$$\mathbf{s} = \begin{pmatrix} s(t_{s,1}) \\ \vdots \\ s(t_{s,N}) \end{pmatrix} \tag{2}$$

Given the vectors **r** and **s**, the depth-matching problem seeks to estimate the warping function, $w(t)$. However, this problem, in practice, is more complicated than the description above. Factors such as noise, differences in sampling rates, and the logs' amplitudes, along with depth misalignment, can contribute to the complexity of the problem. The characteristics of the log curves determine the most suitable approach for the depth-matching process.

Overview of Automated Warping Methods

Despite being recorded in depth, well logs can also be construed as a time series due to their unidimensional domain. Throughout this paper, the time series is assumed to be a direct analog to the depth series of well logs. Several methods for the comparison of different time series exist in the literature. However, some of these approaches are linear and can only yield a good result if the signals align in the time dimension. For example, the Euclidean distance metric is useful in computing the similarity between different time series if similar attributes or patterns in the signals correspond in time. However, there are situations in practice that involve time series which do not align properly. These situations are better addressed using nonlinear distance similarity metrics such as DTW. Next, we review nonlinear similarity metrics that are suitable for signal alignment.

Dynamic Time Warping (DTW)

DTW is useful for measuring the similarity between two temporal sequences, which do not align in time, by warping these sequences in a nonlinear fashion to match each other. For example, Fig. 1a shows the linear element-wise distance between the two signals, which might not be representative for measuring their similarity. In contrast, Fig. 1b shows the nonlinear DTW measure between the two. Computing the similarity of the sequences in Fig. 1b requires a nonlinear metric such as DTW. This uses a recursive distance calculation method to generate a cost matrix that can be used to determine the optimal match between any two sequences.

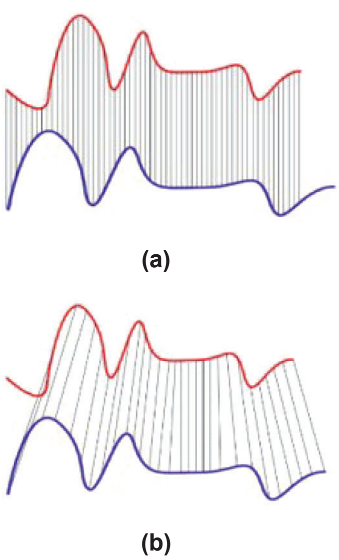


Fig. 1—(a) Linear matching and (b) nonlinear matching.

Given the data sets of Eqs. 1 and 2, DTW aims to transform **s** such that it will match **r** as closely as possible. The algorithm transforms each data point in **s** to match another point in **r** while satisfying three necessary conditions—boundary condition, monotonicity condition, and step size condition (Müller, 2007; Senin, 2008). DTW uses dynamic programming to generate a global cost matrix $\mathbf{D} \in R^{N \times M}$ from a local cost matrix $\mathbf{C} \in R^{N \times M}$, where each element $C(i,j) = c(s_i, r_j)$ is a local cost that measures the similarity between two points in the survey and the reference logs. $C(i,j)$ can be the Manhattan or Euclidean distance between the two points.

With $D(n,0)$, $\forall n \in [1: N]$ and $D(0,m)$, $\forall m \in [1: M]$ set to ∞ and $D(0,0)$ set to 0, each element of the matrix **D** is computed according to Eq. 3.

$$D(i,j) = \min\{D(i-1,j-1), D(i,j-1), D(i-1,j)\} + c(s_i, r_j) \quad (3)$$

Backtracking from the upper right corner to the bottom left corner through valleys or low costs on the global cost matrix **D**, DTW finds the alignment path. The corresponding elements between the two sequences on the alignment path define a match between them. Figure 2 shows an alignment path on a global cost matrix. The darker the shade, the lower the cost.

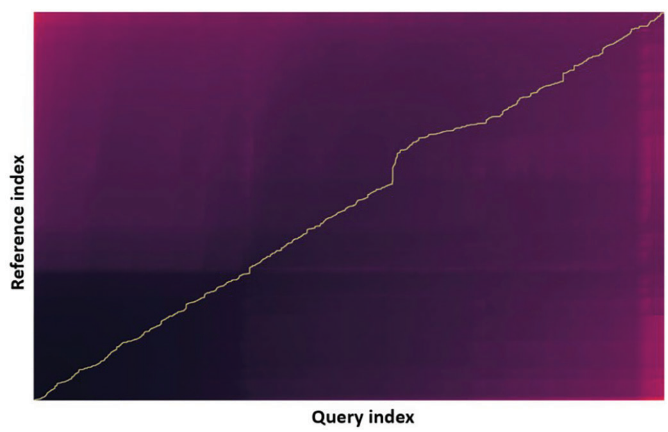


Fig. 2—Alignment path on a global cost matrix. Please note the darker the shade, the lower the cost.

DTW has a wide range of application areas, including speech recognition (Amin and Mahmood, 2008; Permanasari et al., 2019), handwriting and digital signature matching (Niels, 2004; Parziale et al., 2019), and sign language and gestures recognition (Jangyodsuk et al., 2014; Cheng et al., 2016). However, experiments in this paper reveal some of the challenges with the application of DTW in well-log depth matching. Some of these challenges are due to factors such as differences in amplitudes between the two sequences, noise, variations in the lengths of the logs, and the possibility of one-to-many or many-to-one mapping between the sequences. Our experiments empirically demonstrate different preprocesses to improve the effects of noise and amplitudes on DTW and other warping methods.

Constrained Dynamic Time Warping (CDTW)

In practice, there are situations where the result of standard DTW does not give the best solution to a log alignment problem; this could be either because of many-to-one matching or due to large amounts of shifts. The CDTW methods set a constraint to limit the extent of shifts within a given window as well as eliminating many-to-one or one-to-many matching of the indexes of the signals to prevent overstretching or overcompression. The two most popular members of CDTW are the Sakoe-Chiba band and the Itakura parallelogram methods (Itakura, 1975; Sakoe and Chiba, 1978; Geler et al., 2019). They derive a parameter to limit the warping extent around the diagonal of the cost matrix.

The Sakoe-Chiba band method is parameterized by a radius r (r is also called the warping window size), which is the number of off-diagonal elements to consider. The Itakura parallelogram method sets a maximum slope s for alignment paths, which leads to a parallelogram-shaped constraint. Figures 3a and 3b represent the Sakoe-Chiba band and the Itakura parallelogram methods. Choi et al. (2020) report that although the CDTW has a higher classification accuracy than standard DTW, fixing a correct window size can be problematic. A wrong choice for r or s can prevent relevant regions from participating in the alignment process (Zhang et al., 2017).

Correlation Optimized Warping (COW)

Although DTW and CDTW have been applied substantially in practice, they are not the best methods in shape-based pattern matching of signals because they make element-wise comparisons between time series.

Nielsen et al. (1998) proposed COW, which maintains the overall shape of a sequence by warping it into segments with limited allowable flexibility within these segments. COW first splits the sequence into equal segments, which, based on a parameter known as slack, uses different ranges of the segments to search for the optimal match between the sequence. The segments are scaled in time using linear interpolation.

The number of segment borders is determined using the ratio of the points in reference to the selected segment length. The cost matrix for COW consists of the normalized cross correlation between the different segments. Because the length of the segments is variable, the different configurations realized with different adjustments of slack must be considered; to obtain this, a secondary matrix is defined per segment. Similarly to DTW, the problem is finally solved using dynamic programming and backtracking the optimum path on the cost matrix.

COW has been popularly reported in the literature for its applications in chromatographic alignment problems (Nielsen et al., 1998; Tomasi et al., 2004; Bloemberg et al., 2013). One issue with COW is that it is very dependent on user-defined parameters such as the length of the segments and the slack (which determines the maximum stretch or squeeze of the warped segments). Slack and segment length contribute to the accuracy and speed of the method. For example, a larger segment length and a lower slack can speed up the computation time of COW. However, selecting a good segment size and an accurate slack can be challenging in the practical applications of COW.

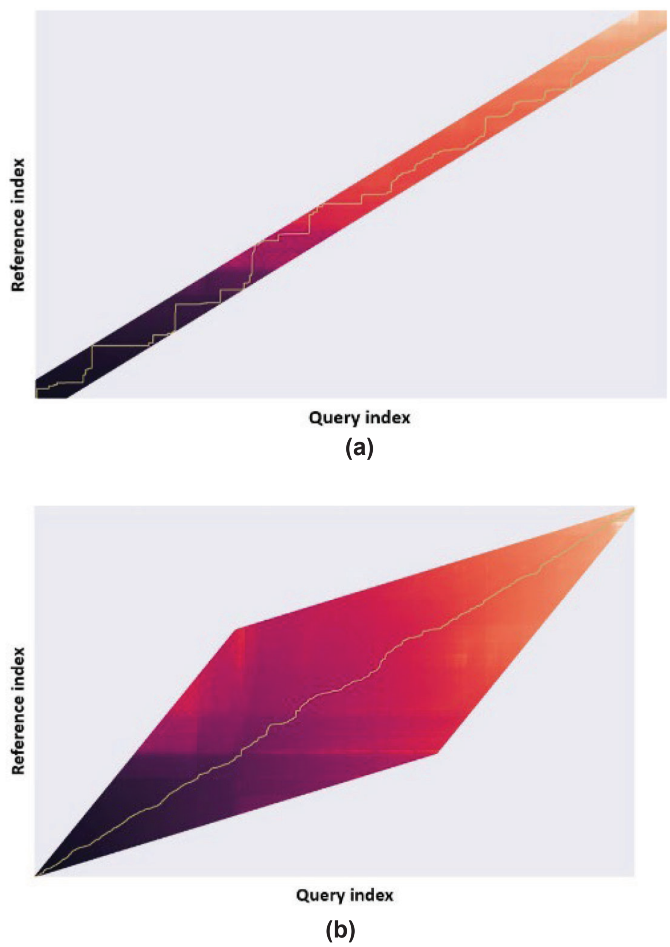


Fig. 3—Constrained DTW methods: (a) Sakoe-Chiba and (b) Itakura. Please note the darker the shade, the lower the cost.

EXPERIMENTS

Data Set

The experiments in this research are based on an anonymized real-world well. The data set, as shown in Fig. 4, consists of the reference log, the survey log, and a manually shifted log provided by a log analyst. The findings made in this paper are consistent with the results we obtained from applying these techniques to several wells from diverse localities. However, for the sake of this review, we focus on one representative well from our data set. Table 1 presents the summary statistics of the logs. Because manually shifting a log does not affect the statistics of the data set, the manually shifted log has the same statistics as the survey log. As can be seen in Fig. 4 and Table 1, the reference log has more than twice the amplitude of the survey log. Moreover, with a signal-to-noise ratio (SNR) (or a coefficient of variation (CV)) of 1.48 (or 67.57%), the reference log is noisier than the survey log with SNR (or CV) of 2.18 (or 45.87%), respectively.

Noise and differences in amplitudes can constitute a problem for some warping methods more than others. As is conventional in most well-log pattern alignment tasks, the survey log is to be shifted in depth such that its features align as closely as possible with similar features in the reference log. The manually shifted GR will be used as a benchmark throughout the experiments since the automation of the well-log pattern alignment process aims to achieve at least human-level accuracy in the task.

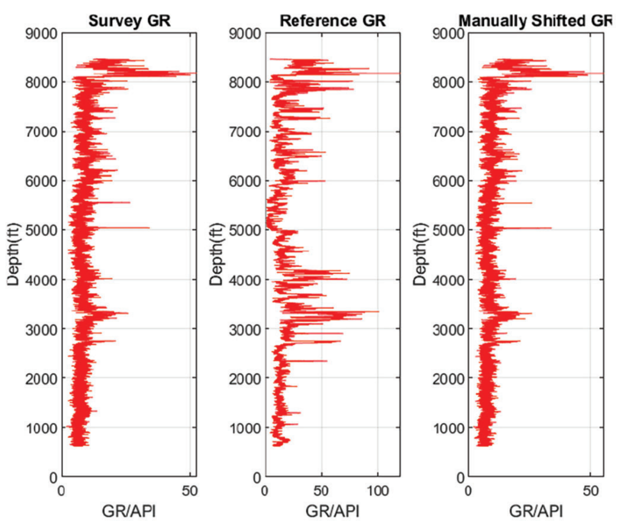


Fig. 4—Data set.

Table 1—Summary Statistics of the Logs

Statistics	Survey Log	Reference Log
size	15,678	15,678
mean	9.24	18.27
std	4.36	12.38
min	2.16	1.19
25%	6.86	11.07
50%	8.24	14.67
75%	10.16	20.80
max	52.53	140.11

Description of Experiments

Using the data set of Fig. 4, experiments were conducted to demonstrate how the performances of different automated well-log warping methods compare with a manually shifted log from a specialist. The experiments investigate the standard DTW technique, two CDTW (the Sakoe-Chiba band and the Itakura parallelogram) methods, and the COW. Each technique has been combined with one or more preprocessing steps to understand how each method can be improved using different preprocessing such as normalization (either z-score normalization or min-max normalization) and filtering (median filtering or wavelet denoising (WD)) (Chang et al., 2000). All possible combinations of the normalization and filtering techniques have been applied to the problem. The results obtained using these methods are compared against the manually shifted log. Figure 5 presents the experiment workflow. These experiments will help to identify methods that compare more closely to the performance of a human analyst. We also explore what works best for each method using different preprocessing techniques and, on this basis, recommend how to improve the automation of the well-log alignment process.

Implementation Details

The parameters for parametric methods, such as the Sakoe-Chiba band, the Itakura parallelogram, the COW technique, and the WD method, are empirically selected to ensure that they yield the best result they can for this task. For the sake of reproducibility, the parameters have been reported in Table 2. Moreover, the DTW and the CDTW have been implemented with the dtwalign Python package[1], while WD has been implemented using the Python scikit-image package.[2] For COW, we replicated exactly the MATLAB implementation by Tomasi et al. (2004) in Python.

[1]https://github.com/statefb/dtwalign
[2]https://scikit-image.org

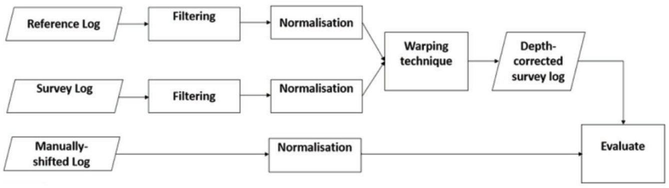
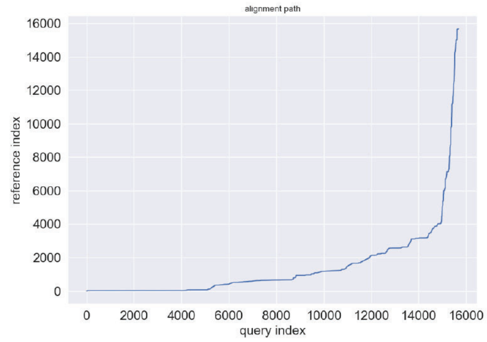


Fig. 5—Experiment workflow.

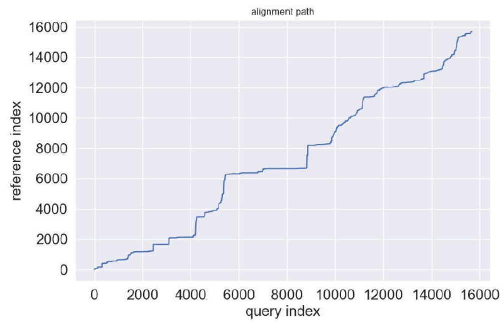
The algorithms are run on Python version 3.8 on a 64-bit Windows 10 computer with Intel® Core™ i5-4570 CPU @ 3.20 GHz and 16.0 GB (15.9 GB usable) RAM.

Table 2—Parameters

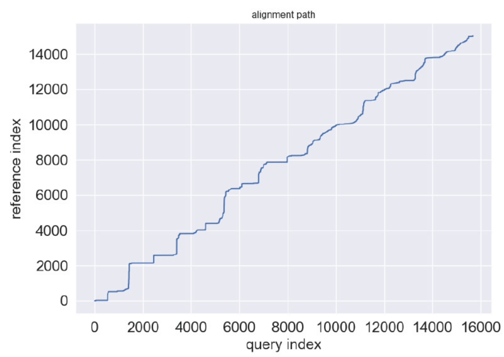
Technique	Parameter	Value	Description
Sakoe-Chiba	*r*	2	warping window size
Itakura	*s*	2	maximum slope
COW	slack seg	50 100	slack segment length
WD	method levels wavelet	BayesShrink 3 coif3	thresholding method wavelet levels kernel



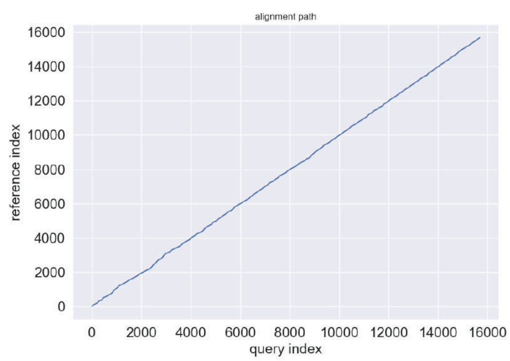
(a) DTW



(b) Itakura



(c) Sakoe-Chiba



(d) COW

Fig. 6—Alignment paths due to each of the warping techniques applied to the data sets without any preprocessing.

RESULT AND DISCUSSIONS

Quality of Log Alignment

While the well-log pattern alignment task aims to match the survey log to the reference log, the aim is not to realize an output log that overfits the reference log to the extent of distorting the survey log unnecessarily. While we intend to match peaks in the survey log with those of the reference log, the aim is to achieve it similarly to the manner a log analyst solves the task. We define the quality of log alignment as a measure of how well a shift table solves the pattern alignment task without unnecessarily distorting the survey log. A warping technique can yield an excessively shifted log due to overfitting or many-to-one mapping of log features. For example, Figs. 6a through 6d show alignment paths for the different warping techniques applied on the survey log before preprocessing. In this case, COW produces the most similar alignment path to the manually shifted depths in Fig. 7a. The alignment path in Fig. 7a does not indicate any noticeable shifts because the maximum depth shift from the manual picks is only about 6 ft, as illustrated in Fig. 7b. From Fig. 6a, DTW resulted in an extreme distortion of the survey log due to its many-to-one mapping of depths. CDTW methods do not result in as much distortion of the log as in DTW; this is because of the constraints the CDTW methods impose on the alignment path. However, the outputs from the CDTW methods on the unprocessed logs are worse than COWs, using the manually shifted alignment path as a benchmark.

Table 3 shows that the outputs of DTW and CDTW (Itakura and Sakoe-Chiba) are poorly correlated with the original survey GR except when they are normalized (although, unlike in other cases, normalization has not improved DTW when combined with WD). The COW's outputs give the highest correlation with the original survey log throughout Table 3, except for the z-score normalized data sets in which COW gives a similar correlation as other methods. From the table, after filtering, the COW's correlation with the survey log is 0.74, with or without normalization. This value is the same as the correlation of the manually shifted log with the original survey log. This result implies that COW provides a similar depth alignment as the manually shifted log after filtering, with either median or WD filtering.

Unlike DTW and CDTW, normalization does not show any noticeable effect on the COW's outputs; this outcome can be explained by the fact that COW is shape based and does not realize depth alignments using pointwise distance metrics similar to DTW and CDTW. Table 3 also reveals that the outputs of DTW and CDTW (Itakura and Sakoe-Chiba) give a relatively low correlation with the original survey GR except when they are normalized (although, unlike in other cases, normalization has not improved DTW when combined with WD). Figure 8 shows preprocessing techniques that can lead to the best alignment path for each algorithm. While a combination of median filtering and z-score normalization can improve the alignment paths of DTW and CDTW, COW does not need normalization. With only median filtering, the COW's outputs compare more closely to the manually shifted log than the other warping techniques, as has been previously demonstrated.

Correlation With the Reference Log

Table 4 shows that the output logs due to DTW and CDTW give correlations of between 0.88 and 0.93 with the reference log, while the COW's logs consistently give a correlation of 0.63 with the reference log. From Table 4 alone and assuming a higher correlation to mean improvement, DTW and CDTW can be adjudged better than COW and even the manually shifted log, which gives a correlation of 0.58 with the reference log. These results have been illustrated in Fig. 9. This section seeks to understand whether a high correlation with the reference log implies a better performance of a warping technique.

In machine-learning parlance, reference logs are the training labels in well-log pattern alignment tasks. Warping techniques should be used to correctly reposition the survey log in depth while comparing the features of the two logs. However, while reference logs provide an example for the task, a high correlation between the output log and the reference log does not necessarily guarantee a good result. A possible factor as to why the outputs of DTW and CDTW result in a high correlation with the reference log is overfitting. DTW and CDTW cost function is based on pointwise distance metrics such as Euclidean distance or Manhattan distance. These distance measures do not pay attention to the shape of the logs to be aligned. Using pointwise distance comparisons, DTW and CDTW try to fit the reference log as much as possible, which leads to the overstretching of the survey log (and, in some cases, the reference log as well) due to many-to-one or one-to-many mapping between the two logs. Although CDTW seeks to prevent overstretching using constraints around the alignment paths, the problem persists with the current CDTW approaches.

Figure 10 illustrates overfitting in the well-log pattern alignment task using DTW as an example. From the graphs, DTW overstretched the survey and the reference logs to enable them to match more closely with each other. In this case, DTW extrapolated the logs from the original maximum length of ~15,000 to ~30,000 ft. Not only that this overstretching resulted in an incorrect solution to the task, but the magnitude of the overstretching was also physically impossible. We argue that the reason why DTW and CDTW give a high correlation, as reported in Table 4, is that the two algorithms optimize the overfitting of the reference and the survey logs.

Table 4 shows that the correlations of the COW's outputs with the reference log are closest to that of the manually shifted log. COW is a shape-preserving segment-based warping technique. COW uses segments and slacks to prevent overstretching of signals while optimizing their shape alignment. COW relates more closely to what a human analyst does in practice. A log analyst seeks to identify related features between the logs using their shapes while ignoring the numerical value and magnitude of each point on the curves. Figure 11 presents the windowed correlations of the survey log, the manually shifted log, and a COW output log against the reference log. From left to right of each of Figs. 11a, 11b, and 11c, the window size is decreased in order to reveal small differences between each log and the reference log. Figures 11a, 11b, and 11c show that the COW output correlates better with the reference log than the survey log but in a similar manner to the manually shifted log. Figure 11d provides a zoomed-in section of the logs for easier visual inspection of the extent to which the COW-automated warping of the survey log compares with the manually shifted log.

Direct Correlation With the Manually Shifted Log

The numerous research in automating the well-log alignment process intends to realize a technique that can guarantee the same quality of service obtainable with log analysts. Hence, a suitable comparison for this task is the manually shifted log. In previous sections, the performances of the different techniques have been indirectly compared to the manually shifted log. For completeness, we discuss the direct correlation between the outputs from the automated well-log alignment techniques and the manually shifted log.

Table 5 demonstrates that COW shows the overall best correlation with the manually shifted log when compared with DTW and CDTW. The table further confirms that while normalization (especially the z-score normalization) improves the performances of DTW and CDTW, only filtering improves the COW's performance.

While the evaluations of the automated well-log alignment techniques have been on the basis of their direct or indirect comparisons with the output from a human log analyst, it is not clear at this stage how best to determine which, between COW and the manually shifted log, is the most efficient. This warrants future research.

Table 3—Correlations Between Shifted GR due to the Different Methods and the Original Survey GR*

	Preprocessing (Filtering + Normalization) Mean (standard error)								
Warping Technique	None	Median	WD	Z-Score	Min-Max	Median + Z-Score	Median + Min-Max	WD + Z-Score	WD + Min-Max
DTW	0.35(0.00)	0.35(0.00)	0.24(0.00)	0.72(0.00)	0.71(0.00)	0.73(0.00)	0.71(0.00)	0.27(0.00)	0.27(0.00)
Itakura	0.55(0.00)	0.55(0.00)	0.56(0.00)	0.72(0.00)	0.71(0.00)	0.73(0.00)	0.71(0.00)	0.71(0.00)	0.70(0.00)
Sakoe-Chiba	0.33(0.00)	0.45(0.00)	0.37(0.00)	0.72(0.00)	0.71(0.00)	0.72(0.00)	0.72(0.00)	0.70(0.00)	0.72(0.00)
COW	0.72(0.00)	0.74(0.00)	0.74(0.00)	0.72(0.00)	0.72(0.00)	0.74(0.00)	0.74(0.00)	0.74(0.00)	0.74(0.00)

*The correlation between the manually shifted GR and the survey GR is 0.74, while that of reference with the survey is 0.48. These are the results of 10 trials for each experiment.

Table 4—Correlations Between Shifted GR due to the Different Methods and the Reference GR*

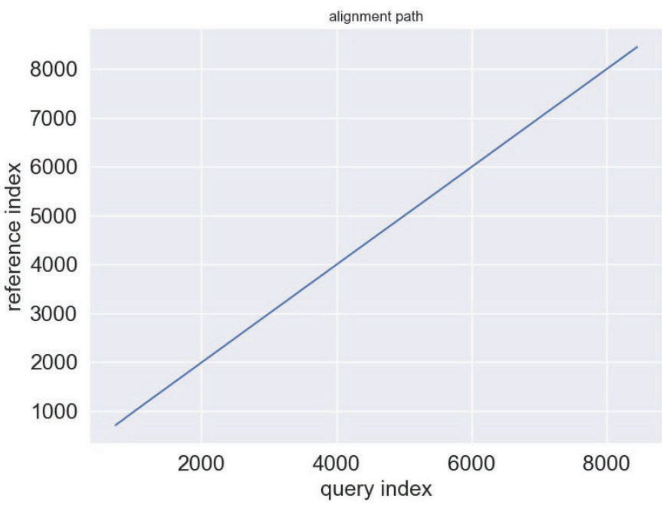
	Preprocessing (Filtering + Normalization) Mean (standard error)								
Warping Technique	None	Median	WD	Z-Score	Min-Max	Median + Z-Score	Median + Min-Max	WD + Z-Score	WD + Min-Max
DTW	0.88(0.00)	0.88(0.00)	0.92(0.00)	0.93(0.00)	0.93(0.00)	0.93(0.00)	0.92(0.00)	0.91(0.00)	0.91(0.00)
Itakura	0.90(0.00)	0.88(0.00)	0.90(0.00)	0.93(0.00)	0.93(0.00)	0.92(0.00)	0.92(0.00)	0.93(0.00)	0.93(0.00)
Sakoe-Chiba	0.91(0.00)	0.89(0.00)	0.90(0.00)	0.93(0.00)	0.93(0.00)	0.93(0.00)	0.92(0.00)	0.93(0.00)	0.93(0.00)
COW	0.63(0.00)	0.63(0.00)	0.63(0.00)	0.63(0.00)	0.63(0.00)	0.63(0.00)	0.63(0.00)	0.63(0.00)	0.63(0.00)

*The correlation between the manually shifted GR and the reference GR is 0.58. These are the results of 10 trials for each experiment.

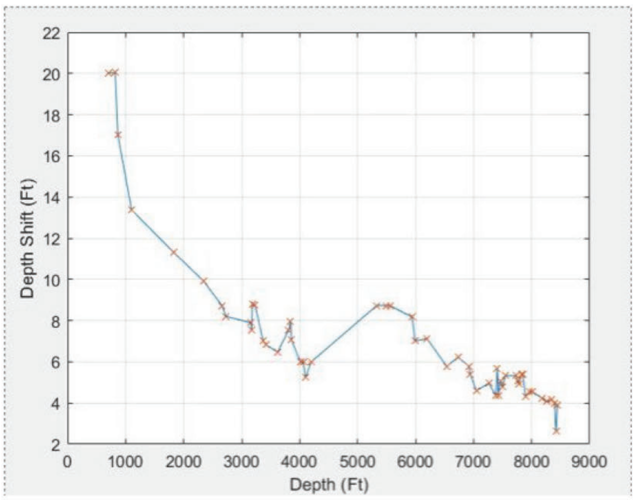
Table 5—Correlations Between Shifted GR due to the Different Methods and the Manually Shifted GR*

	Preprocessing (Filtering + Normalization) Mean (standard error)								
Warping Technique	None	Median	WD	Z-Score	Min-Max	Median + Z-Score	Median + Min-Max	WD + Z-Score	WD + Min-Max
DTW	0.34(0.00)	0.34(0.00)	0.46(0.00)	0.79(0.00)	0.78(0.00)	0.79(0.00)	0.77(0.00)	0.39(0.00)	0.39(0.00)
Itakura	0.59(0.00)	0.58(0.00)	0.57(0.00)	0.79(0.00)	0.78(0.00)	0.79(0.00)	0.77(0.00)	0.79(0.00)	0.76(0.00)
Sakoe-Chiba	0.46(0.00)	0.44(0.00)	0.43(0.00)	0.79(0.00)	0.77(0.00)	0.79(0.00)	0.76(0.00)	0.79(0.00)	0.75(0.00)
COW	0.79(0.00)	0.81(0.00)	0.81(0.00)	0.79(0.00)	0.79(0.00)	0.81(0.00)	0.81(0.00)	0.81(0.00)	0.81(0.00)

*These are the results of 10 trials for each experiment.

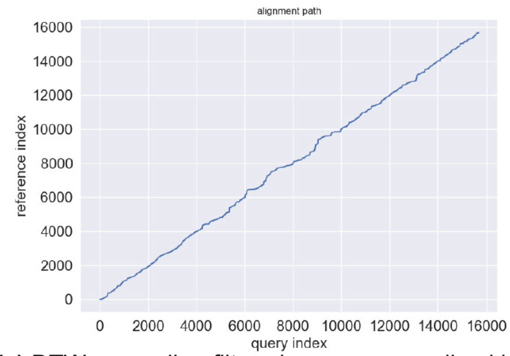


(a)

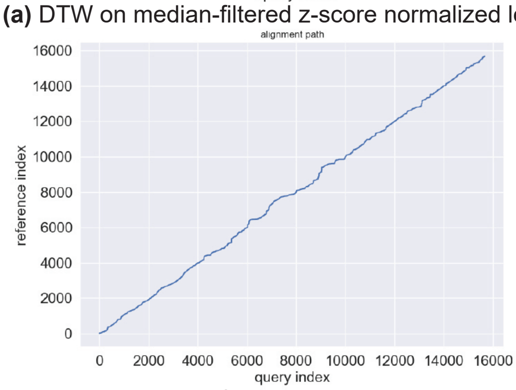


(b)

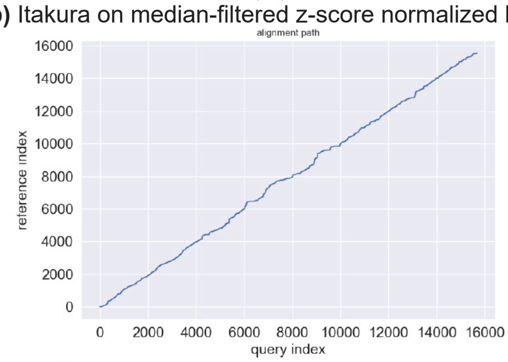
Fig. 7—(a) Alignment paths for the manually shifted depths. (b) Shift profile for the manually shifted depths.



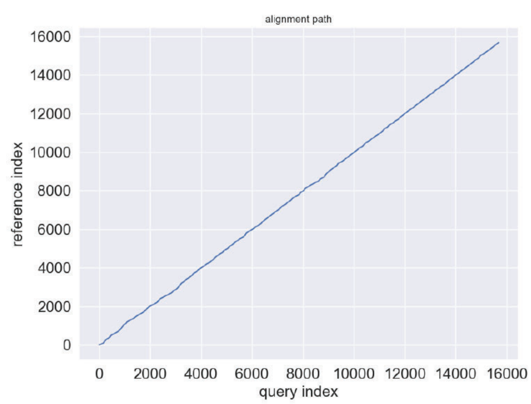
(a) DTW on median-filtered z-score normalized log



(b) Itakura on median-filtered z-score normalized logs



(c) Sakoe-Chiba on median-filtered z-score normalized logs



(d) COW on median-filtered logs

Fig. 8—Preprocessing techniques that lead to the best alignment for each warping technique.

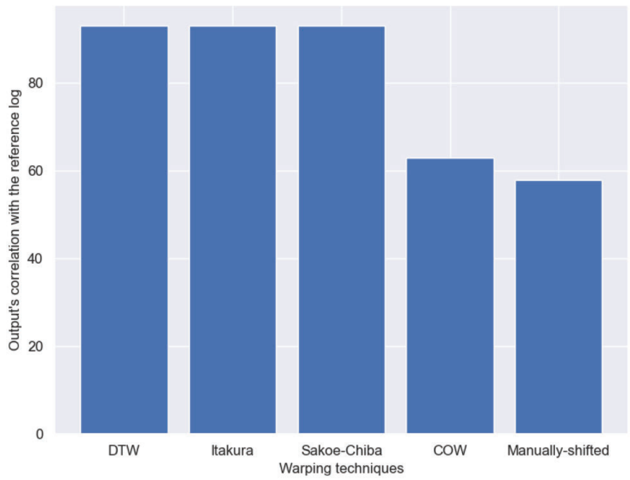


Fig. 9—Comparing the highest correlations each method's outputs give with the reference log against that of the manually shifted log.

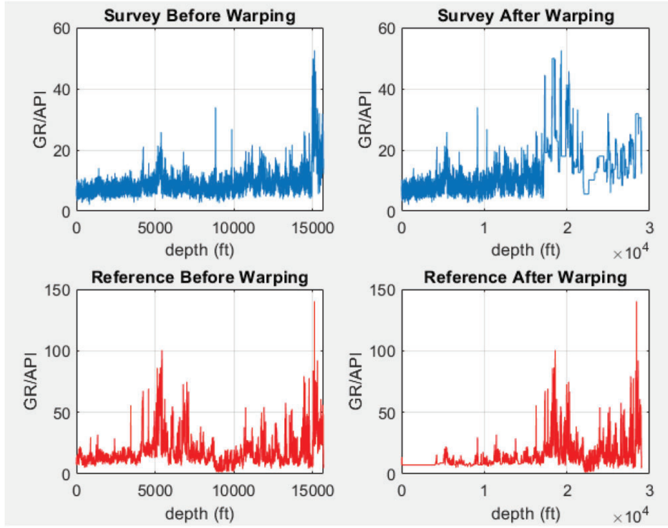


Fig. 10—Illustration of overfitting between the survey and the reference logs using DTW.

RECOMMENDATIONS AND FUTURE RESEARCH

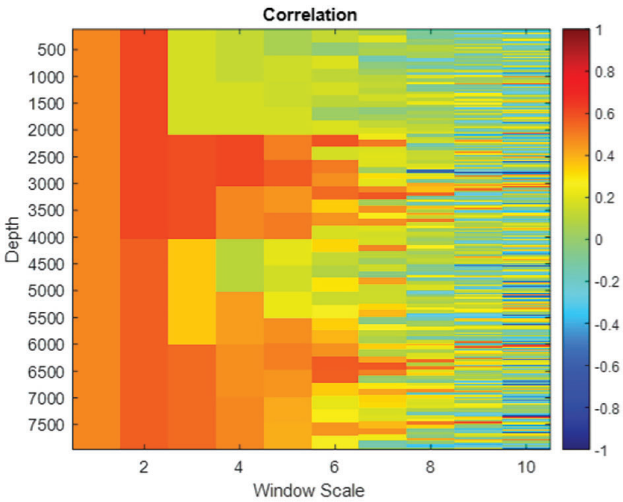
Given the enormous research interest in the automation of the well-log pattern alignment process in the oil and gas sector, the findings made from this research are significant both for academia and industry. Firstly, this research has demonstrated that a high correlation between an output log from a warping technique and the reference log does not guarantee good performance. Key attributes for suitable warping techniques for well-log pattern alignment include good quality of alignment, shape preservation during alignment, not overfitting, and indirectly learning to generate outputs similar to a manually shifted log. These attributes can be evaluated using the analyses provided in this paper.

DTW and CDTW perform poorly in the well-log pattern alignment task because they distort the logs to optimize their overfitting. Their warping process results in shifted logs which do not represent a good solution to the task. A fundamental property of these techniques that contributes to this outcome is their dependency on pointwise distance metrics. One way to improve DTW could be by re-engineering its cost function to use a shape-based similarity metric. However, if, for any reason, DTW or CDTW is a requirement for a well-log pattern alignment task, then the normalization of the logs before warping is recommended.

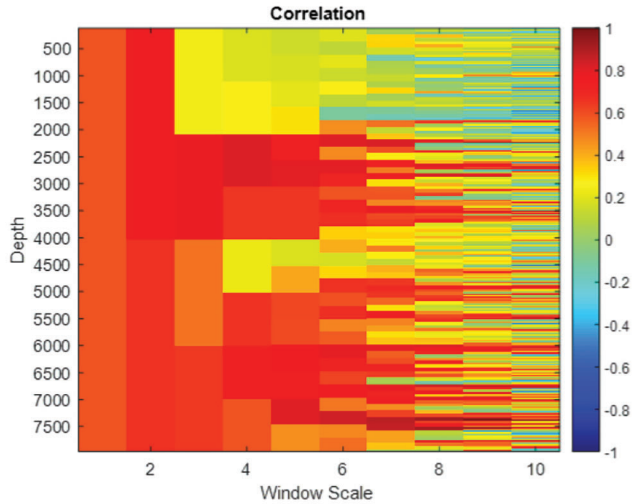
CDTW techniques are designed to overcome the many-to-one challenge of DTW. But because these methods are parametric, choosing suitable parameters for them remains a challenging task. Hence, integrating an adaptive parameter selection strategy for CDTW techniques warrants future research.

Among the different techniques, COW compares more favorably to the manually shifted log than others. However, there are several issues with the current implementation of COW. Firstly, COW works well when the survey and the reference logs are of equal length. The reason for this is that COW assumes that the ends of the logs are already matched and fixed. This situation will not always be true, especially when the logs are of different lengths. Solving this limitation of COW will bring an overarching improvement to COW and the automated well-log alignment process.

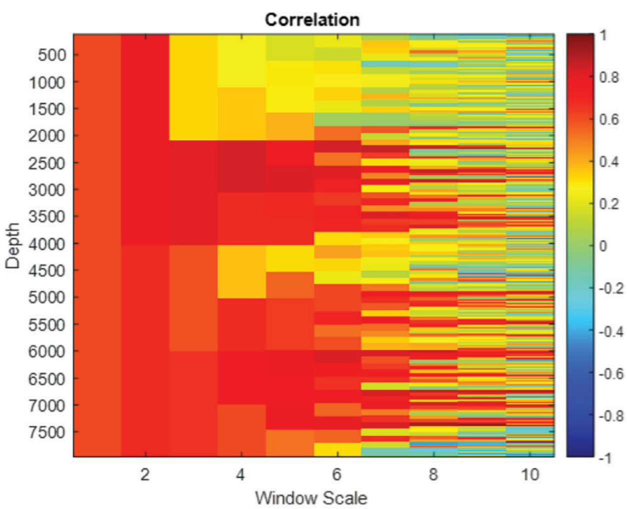
Moreover, COW relies on two essential parameters—segment length and slack. As with any parametric algorithm, selecting correct parameters can be challenging; therefore, an automated technique for choosing the right segment length and slack variable will hugely improve the COW's performance. Furthermore, the current implementation of COW uses the same slack and equal segment length except for the last segment, which can be of a different length. Since logs are not equally dense or sparse throughout, varying segment lengths and slacks could be considered for COW. Our future research will consider heuristics for choosing optimal parameters for the different regions of the log.



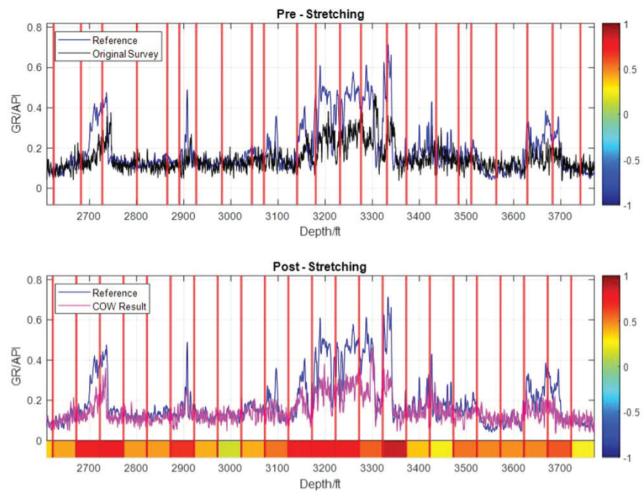
(a) Windowed correlation between the reference and the survey logs. Global correlation = 0.4826.



(b) Windowed correlation between the reference and the manually shifted log. Global correlation = 0.5828.



(c) Windowed correlation between the reference log and a COW output log. Global correlation = 0.6307.



(d) Zoomed-in section of the logs.

Fig. 11—Windowed correlation profile between the reference and the other logs, including a COW output log. Please note the window scale allows correlations to be computed in segments. For example, with a window scale of 1, the entire data set is used at once; with a window scale = 2, the correlations are computed in two segments and so on.

Finally, the COW's outputs in this research have closely compared to a manually shifted log better than other techniques. However, we have not been able to confirm the extent to which COW performs in comparison to the manually shifted logs. Although the COW's outputs yielded higher correlations with the reference log than the manually shifted log, it cannot be guaranteed that it outperforms the manually shifted log. The preceding statement is the case because Section \ref{overfitting} reveals that DTW and CDTW show good correlations to the reference log due to overfitting. Future research will seek to quantify how well COW performs in comparison to an analyst using a blind evaluation of COW outputs and manually shifted logs.

CONCLUSIONS

This paper provides an empirical review of automated well-log pattern alignment techniques. The paper focuses primarily on DTW, CDTW, and COW because of their ease of implementation, sample efficiency, and ease of deployment in everyday petrophysical software. The methods have been experimented on gamma ray—a reference log, a survey log, and a manually shifted log as the ground truth. For each method, a combination of filtering and normalization techniques are applied to the logs. The results show that COW compares more closely to the manually shifted logs than the other techniques. Furthermore, COW performance largely improves with filtering, while DTW and CDTW perform better with normalization. DTW and CDTW have not shown a good result because they use pointwise distance metrics, which overstretches both the survey log and the reference log. Although COW performs relatively better than the other techniques, it still requires several improvements to enable it to cope with logs of different lengths as well as choose the right parameters.

ACKNOWLEDGMENTS

This work was supported by the Scottish Funding Council, Advanced Innovation Voucher, and ANSA Data Analytics.

NOMENCLATURE

Abbreviations

AI = artificial intelligence
COW = correlation optimized warping
CDTW = constrained dynamic time warping
CV = coefficient of variation
DTW = dynamic time warping
GR = gamma ray
SNR = signal-to-noise ratio
WD = wavelet denoising

Symbols

$\mathbf{C}$ = local cost matrix
$\mathbf{D}$ = global cost matrix
k = offset
$\mathbf{r}$ = reference vector
r = warping window in the Sakoe-Chiba band
$\mathbf{s}$ = sample vector
s = maximum slope for Itakura parallelogram
seg = segment length
t = depth
$w(t)$ = warping function

REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D.G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., and Zheng, X., 2016, TensorFlow: A System for Large-Scale Machine Learning, *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 265–283. ISBN: 978-1-931971-33-1.

Al-Rfou, R., Alain, G., Almahairi, A., Angermueller, C., Bahdanau, D., Ballas, N., Bastien, F., Bayer, J., Belikov, A., Belopolsky, A., Bengio, Y., Bergeron, A., Bergstra, J., Bisson, V., Snyder, J.B., Bouchard, N., Bounlanger-Lewandowski, N., Bouthillier, X., de Brebisson, A., Breuleux, O., Carrier, P.-L., Cho, K., Chorowski, J., Christiano, P., Cooijmans, T., Cote, M.-A., Cote, M., Courville, A., Dauphin, Y.N., Delalleau, O., Demouth, J., Desjardins, G., Dieleman, S., Dinh, L., Ducoffe, M., Dumoulin, V., Ebrahimi Kahou, S., Erhan, D., Fan, Z., Firat, O., Germain, M., Glorot, X., Goodfellow, I., Graham, M., Gulcehre, C., Hamel, P., Harlouchet, I., Heng, J.-P., Hidasi, B., Honari, S., Jain, A., Jean, S., Jia, K., Korobov, M., Kulkarni, V., Lamb, A., Lamblin, P., Larsen, E., Laurent, C., Lee, S., Lefrancois, S., Lemieux, S., Leonard, N., Lin, Z., Livezey, J.A., Lorenz, C., Lowin, J., Ma, Q., Manzagol, P.-A., Mastropietro, O., McGibbon, R.T., Memisevic, R., van Merrienboer, B., Michalski, V., Mirza, M., Orlandi, A., Pal, C., Pascanu, R., Pezeshki, M., Raffel, C., Renshaw, D., Rocklin, M., Romero, A., Roth, M., Sadowski, P., Salvatier, J., Savard, F., Schluter, J., Schulman, J., Schwartz, G., Serban, I.V., Serdyuk, D., Shabanian, S., Simon, E., Spieckermann, S., Subramanyam, S.R., Sygnowski, J., Tanguay, J., van Tulder, G., Turian, J., Urban, S., Vincent, P., Visin, F., de Vries, H., Warde-Farley, D., Webb, D.J., 47 Willson, M., Xu, K., Xue, L., Yao, L., Zhang, S., and Zhang, Y., 2016, Theano: A Python Framework for Fast Computation of Mathematical Expressions, *arXiv e-prints*. URL: https://arxiv.org/pdf/1605.02688.pdf. Accessed November 22, 2022.

Amin, T.B., and Mahmood, I., 2008, Speech Recognition Using Dynamic Time Warping, *2008 2nd International Conference on Advances in Space Technologies*, 74–79. DOI: 10.1109/ICAST.2008.4747690.

Bittar, M., Wang, S., Chen, J., and Wu, X., 2020, Reinforced Learning Technique for Multi-Well Logs Depth Matching Yield Better Reservoir Delineation, *SEG Technical Program Expanded Abstracts 2020*, 3852–3856. DOI: 10.1190/segam2020-w13-02.1.

Bittar, M., Wang, S., Wu, X., and Chen, J., 2021, Multiple Well-Log Depth Matching Using Deep Q-Learning, *Petrophysics*, **62**(4), 353–361. DOI:10.30632/PJV62N4-2021a1.

Bloemberg, T.G., Gerretzen, J., Lunshof, A., Wehrens, R., and Buydens, L.M.C., 2013, Warping Methods for Spectroscopic and Chromatographic Signal Alignment: A Tutorial, *Analytica Chimica Acta*, **781**, 14–32. DOI: 10.1016/j.aca.2013.03.048.

Chang, S.G., Yu, B., and Vetterli, M., 2000, Adaptive Wavelet Thresholding for Image Denoising and Compression, *IEEE Transactions on Image Processing*, **9**(9), 1532–1546. DOI: 10.1109/83.862633.

Cheng, H., Dai, Z., Liu, Z., and Zhao, Y., 2016, An Image-to-Class Dynamic Time Warping Approach for Both 3D Static and Trajectory Hand Gesture Recognition, *Pattern Recognition*, **55**, 137–147. DOI: 10.1016/j.patcog.2016.01.011.

Choi, W., Cho, J., Lee, S., and Jung, Y., 2020, Fast Constrained Dynamic Time Warping for Similarity Measure of Time Series Data, *IEEE Access*, **8**, 222841–222858. DOI:10.1109/ACCESS.2020.3043839.

Choubey, S., and Karmakar, G.P., 2021, Artificial Intelligence Techniques and Their Application in Oil and Gas Industry, *Artificial Intelligence Review*, **54**, 3665–3683. DOI: 10.1007/s10462-020-09935-1.

Downton, J.E., and Hampson, D.P., 2019, Use of Theory-Guided Neural Networks to Perform Seismic Inversion, Oral presentation at the GeoConvention 2019, Calgary, Canada, 13–17 May. URL: https://geoconvention.com/wp-content/uploads/abstracts/2019/GC2019_205_Use_of_theory-guided_neural_networks_to_perform_seismic_inversion.pdf. Accessed November 22, 2022.

Edwards, C., 2021, The Best of NLP, *Communications of the ACM*, **64**, 9–11. DOI: 10.1145/3449049.

Garcia Manso, A., 2020, Automatic Depth Matching for Petrophysical Borehole Logs, URL: http://resolver.tudelft.nl/uuid:c6d43d2d-cb33-450f-b59a-636ec07bc34a. Accessed November 22, 2022.

Geler, Z., Kurbalija, V., Ivanović, M., Radovanović, M., and Dai, W., 2019, Dynamic Time Warping: Itakura vs. Sakoe-Chiba, *2019 IEEE International Symposium on Innovations in Intelligent SysTems and Applications (INISTA)*, 1–6. DOI: 10.1109/INISTA.2019.8778300.

Gulli, A., and Pal, S., 2017, Deep Learning With Keras: Implementing Deep Learning Models and Neural Networks With the Power of Python, Packt Publishing Ltd. ISBN: 9781787128422.

Hong, Y., and Kang, C., 2020, Automatic Well Correlation by Aligning Multiple Well Using Deep Neural Networks, Paper SPE-202908 presented at the Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, UAE, 9–12 November. DOI: 10.2118/202908-MS.

Itakura, F., 1975, Minimum Prediction Residual Principle Applied to Speech Recognition, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **23**(1), 67–72. DOI: 10.1109/TASSP.1975.1162641.

Jangyodsuk, P., Conly, C., and Athitsos, V., 2014, Sign Language Recognition Using Dynamic Time Warping and Hand Shape Distance Based on Histogram of Oriented Gradient Features, *Petra '14: Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments*, **50**, 1–6. DOI: 10.1145/2674396.2674421.

Koroteev, D., and Tekic, Z., 2021, Artificial Intelligence in Oil and Gas Upstream: Trends, Challenges, and Scenarios for the Future, *Energy and AI*, **3**, 100041. DOI: 10.1016/j.egyai.2020.100041.

Le, T., Liang, L., Zimmermann, T., Zeroug, S., and Heliot, D., 2019, A Machine-Learning Framework for Automating Well-Log Depth Matching, *Petrophysics*, **60**(5), 585–595. DOI: 10.30632/PJV60N5-2019a3.

Moore, W.R., Ma, Y.Z., Urdea, J., and Bratton, T., 2011, Uncertainty Analysis in Well-Log and Petrophysical Interpretations, Chapter *in* Ma, Y.Z., and La Pointe, P.R., editors, *AAPG Memoir: Uncertainty Analysis and Reservoir Modeling: Developing and Managing Assets in an Uncertain World*, **2011**, 17–28. DOI: 10.1306/13301405M963478.

Müller, M., 2007, Dynamic Time Warping, Chapter 4, *Information Retrieval for Music and Motion*, 69–84, Springer, Berlin, Heidelberg. DOI: 10.1007/978-3-540-74048-3_4.

Niels, R., 2004, Dynamic Time Warping: An Intuitive Way of Handwriting Recognition?, Master's thesis, Radboud University Nijmegen, Nijmegen, The Netherlands. URL: https://ralphniels.eu/pubs/niels-dtw.pdf. Accessed November 22, 2022.

Nielsen, N.P.V., Carstensen, J.M., and Smedsgaard, J., 1998, Aligning of Single and Multiple Wavelength Chromatographic Profiles for Chemometric Data Analysis Using Correlation Optimised Warping, *Journal of Chromatography A*, **805**(1–2), 17–35. DOI: 10.1016/S0021-9673(98)00021-1.

Nurdiawati, A., and Urban, F., 2022, Decarbonising the Refinery Sector: A Socio-Technical Analysis of Advanced Biofuels, Green Hydrogen and Carbon Capture and Storage Developments in Sweden, *Energy Research & Social Science*, **84**, 102358. DOI: 10.1016/j.erss.2021.102358.

Parziale, A., Diaz, M., Ferrer, M.A., and Marcelli, A., 2019, SM-DTW: Stability Modulated Dynamic Time Warping for Signature Verification, *Pattern Recognition Letters*, **121**(1), 113–122. DOI: 10.1016/j.patrec.2018.07.029.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Alykhan, T., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S., 2019, PyTorch: An Imperative Style, High-Performance Deep Learning Library, *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, **721**, 8026–8037. DOI: 10.5555/3454287.3455008.

Permanasari, Y., Harahap, E.H., and Ali, E.P., 2019, Speech Recognition Using Dynamic Time Warping (DTW), *Journal of Physics Conference Series*, IOP Publishing, **1366**(1), 012091. DOI: 10.1088/1742-6596/1366/1/012091.

Sakoe, H., and Chiba, S., 1978, Dynamic Programming Algorithm Optimization for Spoken Word Recognition, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **26**(1), 43–49. DOI: 10.1109/TASSP.1978.1163055.

Sakurai, Y., Yoshikawa, M., and Faloutsos, C., 2005, FTW: Fast Similarity Search Under the Time Warping Distance, *Proceedings of the Twenty-Fourth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, 326–337. DOI: 10.1145/1065167.1065210.

Senin, P., 2008, Dynamic Time Warping Algorithm Review, tech report, University of Hawaii at Manoa Honolulu, USA, **855**, 40. URL: https://csdl.ics.hawaii.edu/techreports/2008/08-04/08-04.pdf.

Sircar, A., Yadav, K., Rayavarapu, K., Bist, N., and Oza, H., 2021, Application of Machine Learning and Artificial Intelligence in Oil and Gas Industry, *Petroleum Research*, **6**(4), 379–391 DOI: 10.1016/j.ptlrs.2021.05.009.

Smil, V., 2022, Decarbonization Algebra: The Cop26 Calls for Impossibly Steep Cuts in Carbon Emissions: Numbers Don't Lie, *IEEE Spectrum*, **59**(2), 20–21. DOI: 10.1109/MSPEC.2022.9706399.

Storey, M., and Bolt, H., 2016, Trends in Well Depth Errors, *EAGE Workshop on Velocities: Reducing Uncertainties in Depth*, **2016**, 1–3. DOI: 10.3997/2214-4609.201600083.

Strubell, E., Ganesh, A., and McCallum, A., 2019, Energy and Policy Considerations for Deep Learning in NLP, *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3645–3650, Florence, Italy, 28 July–2 August. DOI: 10.18653/v1/P19-1355.

Tomasi, G., van den Berg, F., and Andersson, C., 2004, Correlation Optimized Warping and Dynamic Time Warping as Preprocessing Methods for Chromatographic Data, *Journal of Chemometrics*, **18**(5), 231–241. DOI: 10.1002/cem.859.

Torres Caceres, V.A., Duffaut, K., Yazidi, A., Westad, F.O., and Johansen, Y.B., 2022, Automated Well-Log Depth Matching–1D Convolutional Neural Networks vs. Classic Cross Correlation, *Petrophysics*, **63**(1), 12–34. DOI:10.30632/PJV63N1-2022a2.

Wang, L., Niu, D., Zhao, X., Wang, X., Hao, M., and Che, H., 2021, A Comparative Analysis of Novel Deep Learning and Ensemble Learning Models to Predict the Allergenicity of Food Proteins, *Foods*, **10**(4), 809. DOI: 10.3390/foods10040809.

Wang, S., Shen, Q., Wu, X., and Chen, J., 2020, Automated Gamma-Ray Log Pattern Alignment and Depth Matching by Machine Learning, *Interpretation*, **8**(3), SL25–SL34. DOI: 10.1190/int-2019-0193.1.

Ye, Y., Niu, C., Jiang, J., Ge, B., and Yang, K., 2017, A Shape Based Similarity Measure for Time Series Classification With Weighted Dynamic Time Warping Algorithm, *2017 4th International Conference on Information Science and Control Engineering (ICISCE)*, 104–109, Changsha, China, 21–23 July. DOI: 10.1109/ICISCE.2017.32.

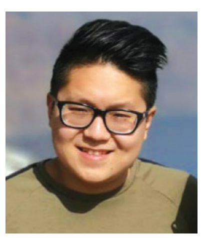
Zachmann, G., Holz, F., Kemfert, C., McWilliams, B., Meissner, F., Roth, A., and Sogalla, R., 2022, Decarbonisation of the Energy System, *Bruegel Policy Contribution Issue*, **1**. URL: http://aei.pitt.edu/103775/. Accessed November 22, 2022.

Zhang, Z., Tavenard, R., Bailly, A., Tang, X., Tang, P., and Corpetti, T., 2017, Dynamic Time Warping Under Limited Warping Path Length, *Information Sciences*, **393**, 91–107. DOI: 10.1016/j.ins.2017.02.018.

Zimmermann, T., Liang, L., and Zeroug, S., 2018, Machine-Learning-Based Automatic Well-Log Depth Matching, *Petrophysics*, **59**(6), 863–872. DOI: 10.30632/PJV59N6-2018a10.

ABOUT THE AUTHORS

Chinedu P. Ezenkwu obtained a PhD degree in autonomous artificial intelligence from the University of Aberdeen, where he is currently a research fellow. He has experience with several industrial and interdisciplinary research projects involving the applications of responsible AI in different domains. Pascal has been recognized as a global talent in AI by the Royal Academy of Engineering, UK. He is currently developing an AI toolkit for health and safety monitoring and management in work environments.



John Guntoro began his studies at the University of Southampton, from which he holds an MSci degree in geophysics. Upon graduation in 2014, he worked in Indonesia as a geophysicist at the centre of collaboration and research of oil and gas with Trisakti University. He is currently a PhD candidate and research assistant at the University of Aberdeen, working in collaboration with ANSA, whose research is focused on the application of machine-learning algorithms on seismic and well-log data.

Andrew Starkey is a senior lecturer at the University of Aberdeen and has research interests in the area of explainable AI, green AI, automated AI, and autonomous learning methods. He has been awarded an Enterprise Fellowship from the Royal Society of Edinburgh and Scottish Enterprise. He is also responsible for Blueflow Ltd, a spin-out company from the University of Aberdeen that proposed a solution for a wide range of data analysis areas, such as financial, textual, and web data, analysis of blogs and discussion threads, and condition monitoring.

Vahid Vaziri, with degrees in control engineering, complex systems, and nonlinear engineering systems, has more than 10 years of research experience and published more than 40 peer-reviewed papers in engineering systems. Vahid is currently a senior lecturer at the University of Aberdeen and PI of many R&D projects, including one that supported the current paper. He has research interests in the application of AI/ML to engineering problems.

Maurillio Addario is the data analytics and software engineering director at ANSA Data Analytics. He oversees the development and application of state-of-the-art artificial intelligence and machine-learning techniques, as well as the creation of novel methods to unlock the true value of customers' data sets. He is based in Aberdeen and has more than 20 years of data analytics experience in the industry, graduating in his native Brazil with a BSc degree in computer systems engineering and an MSc degree in information engineering from RGU.