Incremental correlation of multiple well logs following geologically optimal neighbors

Xinming Wu¹, Yunzhi Shi¹, Sergey Fomel¹, and Fangyu Li²

Abstract

Well-log correlation is a crucial step to construct cross sections in estimating structures between wells and building subsurface models. Manually correlating multiple logs can be highly subjective and labor intensive. We have developed a weighted incremental correlation method to efficiently correlate multiple well logs following a geologically optimal path. In this method, we first automatically compute an optimal path that starts with longer logs and follows geologically continuous structures. Then, we use the dynamic warping technique to sequentially correlate the logs following the path. To avoid potential error propagation with the path, we modify the dynamic warping algorithm to use all the previously correlated logs as references to correlate the current log in the path. During the sequential correlations, we compute the geologic distances between the current log and all of the reference logs. Such distances are proportional to Euclidean distances, but they increase dramatically across discontinuous structures such as faults and unconformities that separate the current log from the reference logs. We also compute correlation confidences to provide quantitative quality control of the correlation results. We use the geologic distances and correlation confidences to weight the references in correlating the current log. By using this weighted incremental correlation method, each log is optimally correlated with all the logs that are geologically closer and are ordered with higher priorities in the path. Hundreds of well logs from the Teapot Dome survey demonstrate the efficiency and robustness of the method.

Introduction

Well-log correlation aims to identify corresponding depth samples among well logs. Each set of such corresponding samples belongs to the same geologic layer with similar rock properties (Wheeler and Hale, 2014). Manually identifying multiple sets of corresponding samples among many well logs is highly labor intensive and subjective. Therefore, computer-based automatic methods have been proposed to correlate single pairs of logs and multiple logs.

The earliest automatic methods (e.g., Rudman and Lankston, 1973; Mann and Dowell, 1978) use the cross-correlation algorithm to compute shifts that correlate a single pair of logs. The cross-correlation algorithm, however, is accurate only to estimate slowly varying shifts. Therefore, some authors (e.g., Smith and Waterman, 1980; Waterman and Raymond, 1987) propose to use a dynamic waveform matching method to better estimate rapidly varying shifts. Such a dynamic warping method is further improved by incorporating human interpretations (e.g., Lineman et al., 1987; Wu and Nyland, 1987; Lalier et al., 2012, 2016) and seismic constraints (Julio et al., 2012) into well-log correlations.

The crosscorrelation and dynamic warping methods can be directly applied for correlating multiple well logs by choosing a reference and then independently aligning all the other logs to the reference (Le Nir et al., 1998). This independent pair-wise correlation method, however, is not able to compute consistent and globally optimal correlations of multiple logs. Fang et al. (1992) compute three pairwise correlations among a cycle of three logs and repeatedly perform such a three-pass cycle processing to obtain multiple-log correlations. Wheeler and Hale (2014) first compute shifts for all possible pair-wise correlations among multiple logs without choosing a correlation path, and then they compute the least-squares fit of the shifts to obtain consistent correlations of the logs. Lalier et al. (2016) sequentially compute pair-wise correlations following a path, and then they iteratively reduce inconsistent correlations until a stable and minimal cost correlation is achieved.

In this paper, we present a highly efficient method, as shown in Figure 1, to compute robust correlations of multiple well logs. First, we discuss how to compute a geologically optimal path that starts with longer logs, follows shorter distances, and avoids faults and uncon-
formities. In computing such an optimal path, we construct a pseudovelocity map by taking into account a seismic structural discontinuity map and a smoothed log-length map. From the pseudovelocity map, we solve for an eikonal equation to compute a travelt ime map, which is further used to sort the logs in a geologically reasonable order. Then, we discuss an incremental correlation method to use an increasing number of previously correlated logs as references to correlate a subsequent log in the correlation path. In correlating the current log to the multiple references, geologically closer references are weighted more than the further ones that are separated from the current log by faults and unconformities. We also discuss how to compute correlation confidences to quantitatively evaluate the correlation results. Such confidences are also used to weight the references (previously correlated logs) for correlating the current log. Finally, we apply a similar incremental correlation method to simultaneously compute consistent correlations of multiple types of logs.

Geologically optimal path
Sequentially correlating multiple well logs is efficient, but it often requires an optimal path to obtain reliable results. An optimal path should start with relatively long logs and follow geologically continuous structures.

Figure 2 shows the method that we use to compute such an optimal path for the 165 density logs (the yellow circles in Figure 2a) of the Teapot Dome survey. In this method, we first count the number of samples in each log to compute a normalized log-length map (Figure 2b), which is further smoothed by applying a Gaussian smoothing filter with a smoothing half-width sigma $= 10$ (samples). In this smoothed and normalized log-length map $l(x)$ (Figure 2c), the logs with relatively longer length are denoted by higher values (white colors). The log length will be an important consideration in properly ordering the logs for a sequential correlation method. We expect the relatively longer logs (with geologically more complete rock properties) to be ordered at the beginning in a sequential correlation method because these beginning logs will serve as references for correlating the subsequent logs. However, log length is not the only consideration of choosing a proper correlation order of the logs. For example, in Figure 2b, the longest log is denoted by the red arrow, but we may not want to choose it as the reference at the very beginning because it is located at the edge of the survey and is far away from all the other logs. The Gaussian smoothing applied to the log length map helps to take into consideration the length of the nearby logs and avoid choosing the outlier longest log as the reference to start with. After the Gaussian smoothing, the smoothed longest logs are now located in the middle of the survey, as denoted by the red circle in Figure 2c.

A proper correlation path should also avoid passing across structural discontinuities. We use a corresponding 3D seismic image to compute the attributes, such as coherence (Marfurt et al., 1999), fault likelihood (Hale, 2013b; Wu and Hale, 2016), and unconformity likelihood (Wu and Hale, 2015), that can highlight structural discontinuities including faults and unconformities. We stack the 3D attribute volume vertically over time or depth to obtain a 2D spatial map of the structural discontinuities. As an example, we compute a 3D fault-likelihood image from the 3D Teapot Dome seismic image and we vertically stack the likelihood image to obtain a map $d(x)$ of fault zone show in Figure 2d. Such a discontinuity map is also normalized such that $0 \leq d(x) \leq 1$.

With these two maps, the log length ($l(x)$) and seismic structural discontinuities ($d(x)$), we further define a pseudovelocity map by $v(x) = l(x) \cdot [1 - d(x)]$ (Figure 2e), where relatively higher values denote areas with longer logs and more continuous structures. To compute a path following longer logs and more continuous structures, we place a wave source (the red circle in Figure 2e) at the position with the highest velocity and we propagate the wave in this velocity map to compute a travelt ime map by solving the following eikonal equation:

$$\nabla t(x) \cdot \nabla t(x) = \frac{1}{v^2(x)}. \tag{1}$$

We compute the travelt ime map $t(x)$ (Figure 2f) by solving a finite-difference approximation to the above eikonal equation using an iterative algorithm similar to that discussed by Jeong et al. (2007) and Hale (2009). In this travelt ime map (Figure 2f), the black curves

Figure 1. The workflow of the proposed methods to first order well logs in a geologically reasonable path and then sequentially correlate the logs using a weighted incremental correlation method.
denote traveltime contours or wavefronts, which generally follow the high velocity values in Figure 2e.

From this computed traveltime map, we can extract traveltimes at all the well-log positions; smaller times should correspond to longer log lengths, more continuous structures, and shorter distances from the source position. Therefore, we can simply sort all the logs (Figure 3a) by their corresponding traveltimes to order the logs in a geologically optimal path that starts with longer logs, follows shorter distances, and avoids discontinuous structures, such as faults and unconformities. With this method, the randomly ordered logs in Figure 3a are ordered in Figure 3b, where the longer ones have been rearranged with higher priorities on left and will be correlated earlier. We can also observe that the density measurements are laterally more continuous in the optimally ordered logs (Figure 3b) compared with those in the randomly ordered ones (Figure 3a). Figure 4a shows the same density logs but are poorly ordered, where the short logs are ordered at the beginning, whereas the relatively long logs are ordered at the end. In the next section, we will apply the same correlation method to the optimally (Figure 3b) and poorly (Figure 4a) ordered logs for comparison.

Although the logs are ordered in a geologically optimal path, it can still be challenging to obtain accurate correlations of the logs using a sequential correlation method. The missing data and measurement errors within the logs may generate correlation errors, which can propagate and accumulate with the path. In addition, the logs ordered close to each other in this geologically optimal path can be spatially far away from each other. In the next section, we propose a weighted incremental correlation method to deal with these potential problems.

**Weighted incremental correlation**

To avoid potential error propagation with the correlation path, we propose to use all the previously correlated logs as references to correlate the current log, so that each log is optimally correlated to all the logs with higher priorities in the path. Because the number of reference logs increases with the correlation

---

**Figure 2.** (a) In total, 165 Teapot Dome logs are displayed on a time slice of the corresponding 3D seismic volume. A map of the normalized length of all the logs (b) before and (c) after Gaussian smoothing. (d) A map of a fault zone is computed from the seismic fault likelihood. Maps of the (c) smoothed well-log length and (d) a fault zone are used to define (e) a pseudovelocity map, (f) where a wave source is placed at the position with the highest velocity to compute a traveltime map. A geologically optimal correlation path of the logs is found by sorting traveltimes at all the log positions.
path, we call this method incremental correlation. In addition, we compute geologic distances between the current log and all the previously correlated logs and we use the distances to weight the references in correlating the current log. Such distances are proportional to the Euclidean distances between the current log and the reference logs, but they increase dramatically across faults or unconformities.

**Weighted incremental dynamic warping**

In this method of computing correlation shifts \( u_k[i] \) for each \( k \)th log \( f_k[i] \), we use all the previously corre-
lated logs as references and solve the following constrained minimization:

$$\text{arg min}_{u_k[i]} \sum_{i=1}^{N} \sum_{j=1}^{k-1} w_{j,k}[f_j[i] + u_j[i] - f_k[i] + u_k[i]]^p,$$

subject to $u_{\min} \leq u_k[i] \leq u_{\max}$ and $|u_k[i] - u_k[i-1]| \leq \varepsilon,$

(2)

where $i = 1, 2, \ldots, N$ represents the sample index in a well log; $f_j[i] + u_j[i], j = 1, \ldots, k - 1$ represent all the previously correlated logs; and $u_j[i]$ are the previously computed correlation shifts for the $j$th log. The positive exponential number $p$ is set to be one in all examples in this paper. The shift ranges $u_{\min}$ and $u_{\max}$ are defined to be $u_{\min} = -250$ (samples) and $u_{\max} = 250$ (samples), which is large enough to correlate the logs in the Teapot Dome survey. The value $\varepsilon$ is the shift strain, which defines squeezing ($u_k[i] - u_k[i-1] \leq 0$) and stretching ($u_k[i] - u_k[i-1] \geq 0$) (Hale, 2013a) of a well log in correlation. In the Teapot Dome example, there are no significant angular unconformities (missing layers) apparent in the measured well logs; therefore, we choose a small strain $\varepsilon = 0.1$ ($-0.1 \leq u_k[i] - u_k[i-1] \leq 0.1$) to avoid unreasonably large stretching and squeezing and to preserve the length of the logs during the correlation.

The term $w_{j,k}$ denotes a weighting map that depends on the $j$th and $k$th logs. We define such a map by combining three factors: (1) geologic distance between the two logs, (2) missing data in both logs, and (3) correlation confidence (equation 3) of the $j$th log. We set $w_{j,k} = 0$ when the correlation confidence of a reference log is smaller than some threshold or either the log sample $f_j[i] + u_j[i]$ or $f_k[i] + u_k[i]$ is missing. This is helpful to avoid using poorly correlated reference logs and missing data for the correlation.

For other reference logs with high correlation confidences, we define the weights $w_{j,k}$ with geologic distances that are defined as traveltimes $t_{j,k}$ from the current log ($f_k[i]$) position to the locations of all the previously correlated logs ($f_j[i] + u_j[i], j = 1, \ldots, k - 1$). Then, we compute a new traveltime map by placing a wave source at the $k$th log position and solving the same eikonal equation (equation 1) with the same pseudove-

**Figure 4.** (a) Density logs are poorly ordered with short logs at the beginning, whereas the relatively longer ones at the end. (b) The poorly ordered density logs are correlated by using the weighted incremental correlation method.
locity map (Figure 2e). From the new traveltime map, we can simply extract the traveltimes $t_{j,k}$ at the reference-log positions and define such traveltimes as geologic distances between the current log $f_k[i]$ and the reference logs $f_j[i + u_j[i]], j = 1, \ldots, k-1$. The extracted traveltimes or geologic distances will be proportional to the Euclidean distances and will dramatically increase across faults and unconformities where the pseudo velocities are set to be significantly low (nearly zeros). In correlating the current log to the reference logs, we set larger weights $w_{j,k}$ corresponding to smaller times $t_{j,k}$ so that the geologically closer references are weighted more than those geologically further ones that may be spatially far away from the current log or geologically separated from the current log by faults and unconformities.

To solve the above weighted and constrained minimization, we use a dynamic programming algorithm with three steps of computing alignment errors, forward accumulation, and backward tracking, as discussed by Hale (2013a). In the first step, the alignment error map $e[i, u_k]$ is a 2D array with respect to the sample index ($i = 0, 1, \ldots, N$) and shift lag ($u_{\text{min}} \leq u_k \leq u_{\text{max}}$). Each error in this map is computed as the summation of $p$ norms of the differences between the current log and all the previously correlated reference logs as defined in equation 2. To find the globally optimal correlation shifts that minimize the objective function in equation 2 is equivalent to finding an optimal path that passes though globally minimum alignment errors in the map $e[i, u_k]$. However, it is often difficult to extract such an optimal path directly from the alignment error map. The dynamic programming method performs a second step of forward accumulating the alignment errors to obtain an accumulated error map for the minimum path picking. In this second step,
the strain constraint \( |u_k[i] - u_k[i-1]| \leq \epsilon \) is incorporated into the dynamic programming method by controlling the accumulation of errors. In the forward accumulated error map, the globally minimum path is often more obvious than in the original error map and can be picked by back tracking the minimum accumulated errors. The same strain constraint is also incorporated in this back-tracking process as in the second step of forward accumulation.

By solving such minimizations sequentially for all logs in the path, we obtain all the correlated logs in Figure 3c. We can observe that all the logs are visually accurately correlated (the same layers in different logs are all horizontally aligned), although a lot of missing samples and noisy or singular values (e.g., the blue colors in Figures 3) are apparent in many of the logs. We do not try to detect these noisy values, and we eliminate them during the correlation. Using weights and multiple references in the incremental correlation method is helpful to reduce the influence from these noisy values.

In Figure 4b, we apply the same correlation method to correlate the same density logs, which, however, are not properly ordered as in Figure 4a. We observe that this correlation method still works well to reasonably correlate most of the density logs, especially these shallow ones. The misalignments of the logs, denoted by arrows in Figure 4b, are caused by the poorly chosen correlation path, where the reference logs on the left are short ones containing only the shallow rock properties. The misaligned logs, however, contain only the deep properties but miss the shallow properties and are therefore not able to be accurately aligned to the reference logs. This problem was solved by more reasonably order the correlation logs as shown in Figure 3b.

**Correlation confidence**

To provide a quantitative quality control of each correlated log \( f_k[i + u_k[i]] \), we compute a correlation confidence \( c_k \) by averaging the zero-lag crosscorrelations of this log and all the previously correlated logs \( \langle f_j[i + u_j[i]], j = 1, \ldots, k-1 \rangle \):

\[
c_k = \frac{1}{k-1} \sum_{j=1}^{k-1} \frac{\langle f_j[i], f_k[i] \rangle^2}{\langle f_j[i], f_j[i] \rangle \langle f_k[i], f_k[i] \rangle},
\]

where \( \langle x[i], y[i] \rangle \) denotes the dot product between two sequences of \( x[i] \) and \( y[i] \), \( f_j = f_j[i + u_j[i]] \) represents the previously correlated logs, and \( f_k = f_k[i + u_k[i]] \) represents the currently correlated log whose confidence is to be computed. The black curve at the bottom of Figure 3c shows the computed correlation confidences for all correlated logs. The confidences of most logs are higher than 0.8, which indicates these logs are well-aligned as shown in the image above the confidence curve. The two relatively low confidence values (the dashed vertical lines in Figure 3b) highlight the corresponding two logs, which are not well-aligned.

Using the same incremental correlation methods, we also automatically correlate the gamma-ray (Figure 5) and velocity (Figure 6) logs measured in the Teapot Dome survey. We observe that most of these logs are also well-aligned (Figures 5b and 6b) and their corresponding correlation confidences are larger than 0.8.

![Figure 6](https://example.com/image6.png)

**Figure 6.** Geologically optimally ordered velocity logs (a) before and (b) after correlation. (b) Horizontal lines in the correlated logs correspond to (c) geologic horizons in the original logs.
The low-confidence logs are denoted by the vertical dashed lines in Figure 5b. The misalignments and low correlation confidences of these logs are caused by their large measurement errors (17th, 76th, and 102th logs) and missing data (all the other logs highlighted by vertical dashed lines in Figure 5b). Although some of these logs are correlated relatively earlier, these logs are not used as references for the subsequent correlations. In these examples, we set the weights \( w_{j,k} = 0 \) in equation 2 when the correlation confidences are smaller than 0.8 to filter out these unreliably correlated logs.

After the well-log correlation, all of the logs are mapped to the relative geologic time (RGT) domain, in which the geologic layers in all the logs are horizontally aligned. It is convenient to pick horizons across well logs in this RGT domain and map them back to the original depth domain to construct lithostratigraphic cross sections. As shown in Figures 3c, 5b, and 6b, the vertical axis of the correlated logs represents RGT, which means that horizontally aligned samples of different logs correspond to a same geologic time and belong to the same geologic layer. We can simply draw horizontal lines, as in Figure 6b, to pick geologic horizons across multiple logs in this RGT domain. Then, we can simply add the computed correlation shifts to the RGT of these horizontal lines to obtain the corresponding horizons in the original depth domain and build a cross section as shown in Figure 6c.

**Correlating multiple types of logs**

In practice, multiple types of logs, corresponding to different rock properties, are often measured in a same well. It is worthwhile to simultaneously correlate multiple types of logs to obtain consistent correlation shifts. In addition, simultaneous correlation is helpful to com-

![Figure 7](https://library.seg.org/Interpretation/August2018/720.png)

*Figure 7.* Four types of well logs, including (a) velocity, (b) gamma ray, (c) density, and (d) porosity are simultaneously correlated as shown in (e-h), respectively.
pute more reliable correlations by incorporating more constraints from different types of logs.

To be applied to the simultaneous correlations of multiple types of logs, the weighted incremental correlation method requires only minor modifications. In this case, we use the dynamic programming method to compute the same shifts \( u_k[i] \) for all \( M \) types of logs at each \( k \)th well by solving the following minimization:

\[
\arg \min_{u_k[i]} \sum_{i}^{M} \bar{e}_m[i, u_k],
\]

subject to \( u_{\min} \leq u_k[i] \leq u_{\max} \) and \( |u_k[i] - u_k[i-1]| \leq \varepsilon. \)

(4)

where \( \bar{e}_m[i, u_k] \) represents normalized alignment errors \( e_m[i, u_k] \) computed for each \( m \)th type of logs:

\[
e_m[i, u_k] = \sum_{j=1}^{k-1} w_{m,j,k} |f_{m,j}[i + u_j[i]] - f_{m,k}[i + u_k[i]]|^p.
\]

(5)

In this equation, \( f_{m,j} \) denotes previously correlated \( j \)th (\( j = 1, 2, \ldots, k - 1 \)) well at the \( m \)th log type; \( u_j[i] \) represents the shifts that are already computed for the \( j \)th well; \( f_{m,k} \) is the \( k \)th well (at the \( m \)th type) to be correlated; and \( u_k[i] \) represents the shifts to be computed for this \( k \)th well. Notice that the shifts \( u_k[i] \) are independent on the log type \( m \) because we expect the shifts to be the same for all different types of logs at the same well. The weights \( w_{m,j,k} \) are defined similarly to those in equation 2, but they also depend on the log type \( m \). Note that the alignment errors (equation 5) are normalized before summing them together in equation 4 because the magnitude orders of different log types are much different from each other as shown in Figure 7.

Using the correlation shifts computed with this method, we simultaneously correlate the four types of velocity, gamma ray, density, and porosity logs, as shown in Figure 7e–7h, respectively. We can observe that all the well logs are accurately and consistently aligned. These 21 wells are still sequentially correlated from well index \( k = 1 \) to \( k = 21 \). However, the processing order (from index \( k = 1 \) to \( k = 21 \)) is not chosen to be geologically optimal as in the previous examples of correlating a single type of logs.

The processing order shown in Figure 7 is actually randomly chosen for testing the method. We can observe that the all the velocity logs are missing at the first eight wells except the first well. In addition, the four types of logs at the ninth well are all short but are ordered and correlated before many other logs. Therefore, this randomly chosen processing order in Figure 7 is obviously not an optimal order for the sequential correlation. However, the weighted incremental correlation method still provides reliable correlations as shown in Figure 7e–7h, which indicates that this proposed method does not highly depend on the path.

Conclusion

We have discussed methods to compute a geologically optimal path by solving an eikonal equation defined with log length and seismic structural discontinuities, to sequentially correlate multiple logs by using an increasing number of references, and to quantitatively evaluate the correlation results by computing correlation confidences.

The geologically optimal path starts with relatively longer logs, follows shorter Euclidean distances, and avoids geologically discontinuous structures such as faults and unconformities that are recognized from a corresponding 3D seismic image. To avoid correlation errors propagating with the path, we use all the previously correlated logs as references to correlate a subsequent log. We also compute geologic distances and correlation confidences of the reference logs and use the distances and confidences to weight these references in correlating the current log. By using weighted and increasing number of references, each log is optimally correlated to all the logs that are geologically closer and are ordered with higher priorities in the path. This weighted incremental correlation method is highly efficient and takes less than 1 min to correlate the four types of logs of 21 wells (Figure 7) with an eight-core computer.

Acknowledgments

This research is supported by the sponsors of the Texas Consortium for Computation Seismology. All the well-log data are provided by the Rocky Mountain Oilfield Test Center.

References


Julio, C., F. Lallier, and G. Caumon, 2012, Accounting for seismic trends in stochastic well correlation, in P. Abrahamsen, R. Hauge, and O. Kolbjørnsen, eds., Geostatis-


Biographies and photographs of the authors are not available.