International Journal of
Earth & Environmental Sciences

Original Article **Open Access**

Special Issue: Sedimentary Environments and Facies

Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision

Valentina Baranova[1], Azer Mustaqeem[1*] and Aziz Mustaqeem[2]

[1]Petro-Explorers Inc. Calgary, AB, Canada
[2]Faculty of Mathematics, University of Waterloo, Canada

Abstract

Sequence stratigraphy is one of the methods in geoscientific interpretation that provides insight into stratigraphic processes and facies and assists in building meaningful geological earth model. Sequence stratigraphic methods depend heavily on human interpretation and level of expertise, are quite laborious, and mainly qualitative in nature. Although sequence stratigraphic interpretation of seismic data has been automated through various commercial software packages, well data-based sequence stratigraphy workflows remain largely as research projects and lack wide scale automation. Automatic identification of the major sequence stratigraphic surfaces (sequence boundaries and flooding surfaces) and their chronologic ranking remains the main obstacle. The above is attempted through automation of newly proposed Quantitative Sequence Stratigraphic technique based on earlier published research. Proposed automated workflow addresses three main areas: well log data preparation and conditioning, chronostratigraphic log interpretation, and application of quantitative sequence stratigraphy. Well log data preparation covers automation in data cleaning, normalization, geometry correction and format conversion. Chronostratigraphic log interpretation includes expert supervised automated picking and correlation of the formation tops and chronostratigraphic age assignments for every well in a large database. Quantitative sequence stratigraphic technique follows, to automatically identify major sequence stratigraphic surfaces (sequence boundaries and flooding surfaces) of different orders for all wells. Cross-sections, mapping, and log interpretation tools are developed to allow experts to quality control, correct and supervise the interpretation process at every stage of the workflow. The workflow and its application are demonstrated on two datasets: Blackfoot dataset from the Western Canadian Sedimentary Basin, Alberta, Canada and Teapot Dome dataset from Laramide Basin, Wyoming, USA.

Publication History:

Received: March 02, 2022
Accepted: March 26, 2022
Published: March 28, 2022

Keywords:

Automation sequence stratigraphic workflow, Quantitative analysis of geological well log data

Introduction

The pace of the decision making in oil and gas industry has been accelerated recently, following the introduction of cloud computing and the availability of large digital datasets. As a result, we are witnessing high demand for automation in various stages of exploration and production cycles. Many workflows from data processing and conditioning to the building of the final earth model are being subject to extensive automation. Recent advances in distributed computing allow us, not only to get the results faster and consistent, but also to process, filter, correct and understand big data with relative ease. The projected lack of investments in oil and gas industry, increase in energy demand, need for more efficient operations, and lower carbon business models are pressurizing companies to switch from the manual to automated applications.

Sequence stratigraphic interpretation methods have helped in identifying hydrocarbon plays and assisted in creating subsurface geological and basin models in nearly all geological settings. However, the sequence stratigraphic workflow continues to be arduous, time consuming and qualitative in nature. Human expertise and experience remained the central key, and thus was not extensively used to process large datasets under exploration time constraints. Traditionally, the methods of sequence stratigraphy include working with well logs to identify significant stratigraphic boundaries, assign geological ages, correlate formation tops, and build dip and strike sections and maps. Geologists would also use various additional datasets (outcrop descriptions, core analyses, and biostratigraphy) where available. In parallel, the seismic sequence stratigraphy method is often used to identify horizons and faults on seismic data. The resulting geological model is then built using well logs along with the seismic boundaries, faults and stratal terminations picked from 3D seismic. Normally, this would take weeks, if not months, for evaluation of an exploration or a development block.

There are numerous attempts and methods proposed, in the recent past, to automate some parts of the sequence stratigraphic workflow. Most of them addressed well tops correlation, petrophysics, and data preparation. We tried to provide a complete overview, of published work, for the automation of the different processes of the workflow where possible. However, there is not much research or attempts to automate the complete sequence stratigraphy workflow from well log data preparation and cleaning, major sequence stratigraphic surfaces identification, to creating geological cross-sections and maps.

This study is an attempt to achieve a sequence stratigraphic model by computerized automation of every step of the workflow. Quantifying the geological and sequence stratigraphic concepts is the key to the implementation of the workflow. The workflow thus acts like an assembly line, where experts provide quality control and critical corrections at various stages.

***Corresponding Author:** Mr. Azer Mustaqeem, Petro-Explorers Inc, P. Geo, Calgary, AB, Canada, Tel: +1(403)978-1753; E-mail: azer@petroexplorers.com

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

At the initial stage, the input digital well log data, is taken through various steps of data cleaning and header assignments, where some data will be rejected due to errors in the data or format. This step, data labelling, is also important to minimize errors in later stages. In parallel, well data geometry is also built to place the log data in their exact location in the sub-surface.

Reference wells are identified out of all wells available in the study. A reference well is a well that represents a geologically related cluster of wells in the area with a similar log motif. For any given area, there may be one or more reference wells. Then, all the stratigraphic tops are identified on those reference well(s). Geological age will be assigned to each stratigraphic marker using the established chrono-stratigraphy for any given area. Using a proprietary method of multi-window correlation and age sequencing, tops are automatically picked for all wells. This is followed by the identification of sequence boundaries and flooding surfaces on all wells using automation of the newly proposed “TSF Analysis” [1].

With the maturation of seismic based methods of automatically identifying discontinuities on 3D seismic data [2,3], it is possible to identify unconformities, relative geological time surfaces, and faults from seismic data. By automatically tying wells and updating the velocity model using reference horizons it is now possible to automatically build a geological model for any given area by using the RGT (Relative Geological Time) cube. Using the well based analyses described in this paper, full relative geological model is converted to an Absolute Geologic Time (AGT) Earth Model where each sample of the cube has an X, Y, Z, and an Age value.

Our research continues to address the data, algorithms, and geological complexities to achieve the above goals. The methodology presented in this publication is dealing with the automated sequence stratigraphy using well log data as the only input.

The automated workflow is applied to two datasets with different sets of lithologies (Blackfoot field, Alberta, Canada, and Teapot Dome field, Wyoming, USA).

Method

Major steps of the automated sequence stratigraphy workflow are shown in Figure 1. Input data include digital well logs, together with deviation files and well header information on the one hand and 2D/3D seismic datasets on the other. As mentioned, earlier, the present paper only deals with the digital well logs dataset along with associated well information. The results are hence limited vertically to well log data resolution. Additional project area information, prior stratigraphic studies, geological experience in the area, and computer data QC algorithms are used to initialize the processes and to fix the errors in input data.

Well log data is often sampled at various depth increments and mixed units (meters or feet); a resampling of the log data is the first step to create fixed increment logs with consistent measurement units. Log data files often do not have projected local coordinate information and could be missing information critical for well location and labeling. A separate automatic subroutine will search for Universal Well Identifier (UWI) and well names to compare and correct the log files using Well Header Information (header file) and log files thus reducing further errors.

Individual log extraction processes follow, where each log type, along with their aliases, are extracted into separate directories carrying the same header information as their parent log files. At present, we are only using Gamma Ray (GR) logs due to their prevalence in exploration well data. GR logs are normalized and converted to V-shale logs. The goal of the automation process is to eventually work with the various volumetric lithology logs such as V-shale, V-limestone, V-coal, and V-quartz to further improve the workflow. In the case of older well data where the GR data is not available Spontaneous Potential Log (SP) can be used to derive V-shale data.

For the given dataset, the reference well(s) should be identified. Formation tops are picked on reference well(s) and geological age is assigned to them. Using the automated top correlation subroutine, the

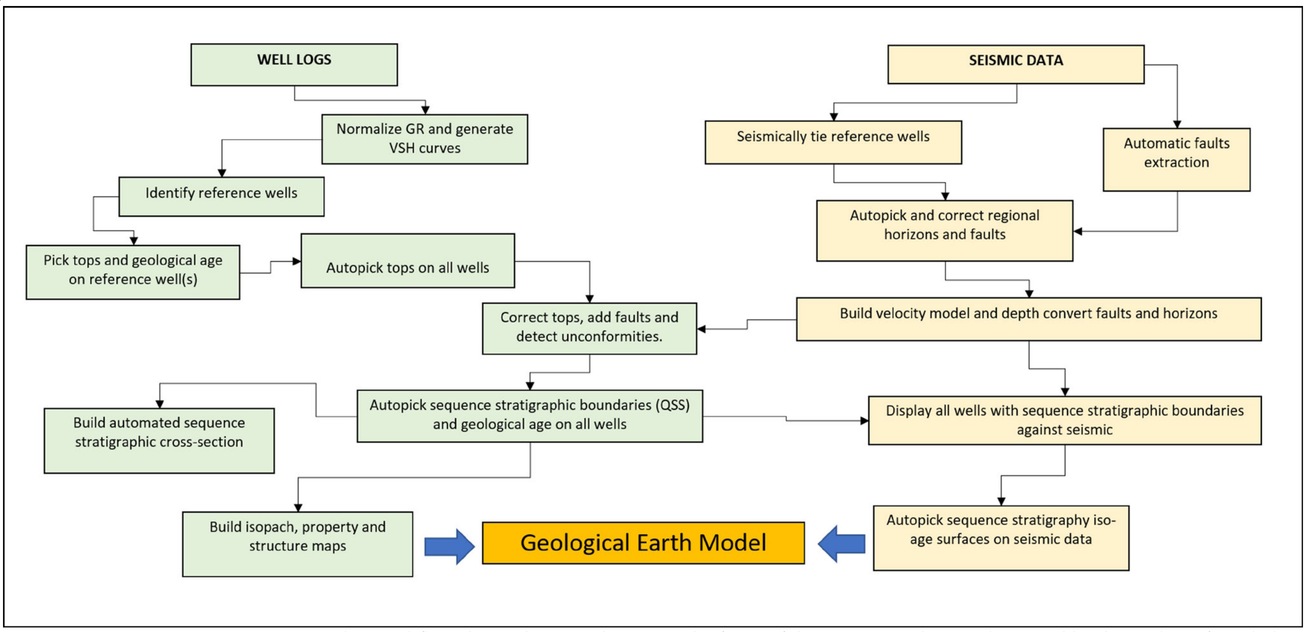

Figure 1: Automatic sequence stratigraphic workflow chart. The green boxes are the focus of the present study complemented by the analyses from light yellow boxes.

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

formation tops are to be picked on all wells followed by a correction of tops, adding faults, and detecting unconformities. This stage requires expert supervision. Then, major sequence stratigraphic boundaries are picked on each well automatically and needed cross-sections, isopach, property and structure maps are generated.

Below is detailed description of the workflow followed by the two case studies.

Normalization of GR and V-shale logs

GR or gamma ray logs are most used log for quantitative sequence stratigraphic analysis due to their close relation to grain size in siliciclastic depositional setting. GR logs require normalization through the area of investigation, such that the results are not skewed by variations in the baseline. Normalization is carried out using an average and standard deviation derived from a chosen set of 10-20 logs, that have deeper penetrations, spatially spread through the project area and the one covering most of the stratigraphic intervals. Normalization is important for large datasets as many of the well logs may have been recorded with different logging tools or using various tool calibrations over the years. To avoid effect of spikes in the data, a large window filter is used to calculate the standard deviation and to balance the logs. For the wells that do not have GR logs, Spontaneous Potential (SP) logs are used through the application of GR prediction and subsequent normalization.

In the area of investigation, there could be multiple runs of GR logs for the same well. We automate a log merging technique where computer algorithm detects the start and stop depth of the log, correlates with any other occurrence within the same depth range, than selects and discards based on depth window of coverage, correlation index and log run information, if available. Computer automation will also flag the well where there are contradicting logs within the same depth range.

Once the logs are normalized and merged, a simple equation is used to convert the GR logs to GR Index (GRI).

$$GRI = \frac{GR_{\log} - GR_{\min}}{GR_{\max} - GR_{\min}}$$

where GR_{max} and GR_{min} values are calculated using multiple well histogram after normalization. Volume of Shale (VSH) can be derived from GRI using various equations established by many authors [4-6]. In the simplest linear form VSH is equal to GRI.

Figure 2 shows an example of GR to VSH conversion using above methods. Once an optimum conversion equation is identified for an area, all the normalized GR logs are converted to VSH logs.

In most of the siliciclastic sedimentation there are no major difference between GR response and V-shale, although in carbonates and coal bearing strata one may have to go through a petrophysical processes to correct for volume of carbonate, volume of coal and low radioactive shales.

Although it is preferable to derive V-shale from GR due to its widespread availability, in some areas with older set of logs V-shale derived from Spontaneous Potential (SP) and resistivity (Res) logs can be integrated with the ones from GR. In such cases, a machine learning algorithm (Random Forest with XGBoost) is trained on SP, Resistivity and GRI using the wells that contain all three logs [7]. The trained algorithm is then applied to the wells with SP and Resistivity only to predict the missing GR logs.

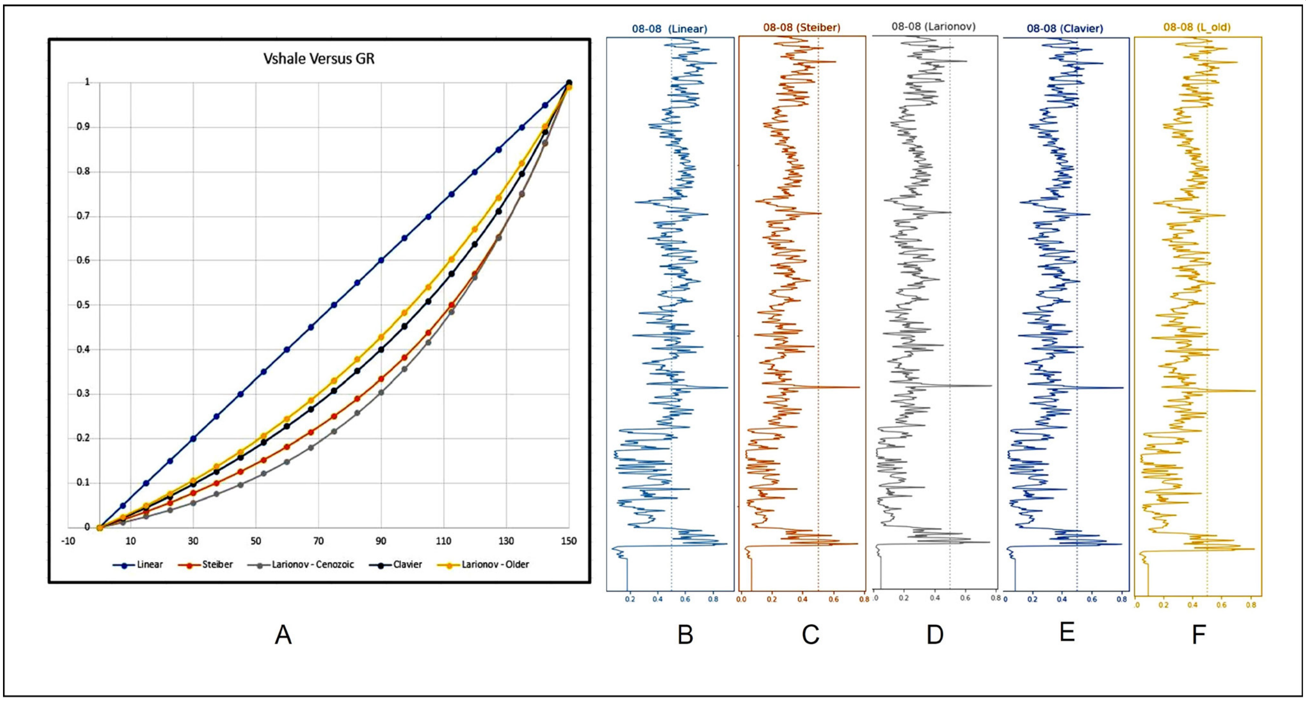

Figure 2: Graphical representation of various conversion equations from GR to Volume of Shale (VSH) and examples of VSH calculation applied to Blackfoot dataset. (A) – Conversion functions: (B)-(F) – Examples of GR to VSH conversions with different equations on well 08-08. Linear: VSH = GRI = $(GR_{log} - GR_{min}) / (GR_{max} - GR_{min})$; Larionov: VSH = $0.083\,(2^{3.7*GRI} - 1)$; Steiber: VSH = GRI / (3 – 2*GRI); Clavier: VSH = $1.7 - [(3.38 - (GRI + 0.7)^2]^{0.5}$; Larionov for older rock: VSH = $0.33\,(2^{2*GRI} - 1)$.

Reference well identification

The reference wells, representing large datasets, are the wells that are chosen and interpreted as type-wells. The emphasis in the choice of reference wells is on better log quality, core availability, bio-stratigraphic information, petrophysical analysis and completeness of log suite. A reference well represents all wells in its vicinity with similar log motif and geology (similar deposition environment). There can be one or numerous reference wells representing a dataset. All stratigraphic markers, with assigned geological age, are then picked manually on reference wells using locally established stratigraphic column and ages. Unconformities are also assigned their approximate age gaps.

Geological age assignment

After assigning absolute age to the tops, the tops of reference wells become chronostratigraphic surfaces. Based on the stratigraphic markers the Geological Time Scale is automatically adjusted and saved for each well. The geological time scale used in the process have the most updated values of epoch and stage boundaries [8]. It is important to note that in our sequence stratigraphic workflow we consider the tops as geological time markers rather than a lithologic interface and thus may contradict some of local established norms. Many oil and gas exploration companies that have dealt with large datasets, spanning various jurisdictions, have proposed geological time scale based tops interpretation than local formation top names.

In the case of unconformities and/or normal faults an age gap is either assigned through an age table or derived from the seismic data, if available. The missing tops will follow the iso-age lines thus overlapping each other with zero thickness.

Sea level curve assignment

To better identify and constrain sequence stratigraphic surfaces, the base sea level curves, global and local, have been added to the display as per Haq, [9] and Chevron Gulf of Mexico, respectively. These curves are also adjusted automatically and scaled by age values of the formation tops.

Well tops correlation

Picking formation tops is another labor-intensive part of any interpretation. There are several approaches to automate formation top picking, some of the published examples are described as following. Shi et al., [10], and Wu., [11], used coherence-weighted graphs, showing a workflow for sequential correlation of multiple well logs, following an optimal path that preserves maximum coherency between neighboring log traces. Gosses and Zhang [12], presented a supervised machine learning method, mimicking visual approaches of geologist. Zoltan Sylvester ([13]) used dynamic time warping along with the correlation within chronostratigraphic surfaces. Brazell et al. [14] applied a novel deep convolutional neural networks (CNN) to correlate tops for a big dataset. Karimi et al. [15], used well-to-well correlation approach based on principal component analysis to identify lithologic boundaries.

We take a supervised correlation approach where the reference well(s) is picked by an expert interpreter. Quite like Wu [11], we start with the first pass of sliding window cross-correlation between reference log and the target log to approximate the target window with highest cross-correlation coefficient. The process than follows with another small window to fine tune the results. The algorithm follows the logic of picking tops on nearby wells first and then extend it to other wells, radially moving away from the reference well(s). Correlation is an iterative process and uses correlation coefficients as well as stratigraphic position of a top in decision making. The weighting factor is also applied based on the distance from the reference well. Once a high correlation (e.g., correlation coefficient > 0.8) is achieved for a certain formation top, the new top becomes part of the seed pick as well as a reference in the location of the subsequent window location for the next iteration. Auto-picked tops are sequenced with geological age and the hierarchy will be preserved except in the cases of thrust faults where the tops may repeat due to thrusting.

Although the results obtained are quite reliable, they could still face problems due to missing logs, large distances and the presence of faults and unconformities. To solve the issue of faults and unconformities, we apply various processes to constrain the results. Firstly, the location and throw of the faults are taken from the depth converted RGT (relative geological time) seismic data and are considered as a gap in the stratigraphic age. Each fault plane in 3D is assigned a throw value defined as an age gap such that when the well crosses the fault plane a value is automatically extracted on the well log as an X, Y, Z and age gap (negative for thrust faults and positive for normal faults). To address unconformities, the top age table is assigned with the number of age gaps observed by the experts in the reference wells. Lastly, based on the depth of logs all the formation tops must be picked or estimated in the final iteration (phantom tops) as a geological age surface, even if they are missing in a fault zone or are eroded. The phantom tops will follow each other along a fault plane or unconformity with zero gap thus defining an age gap but depicting missing top and age.

Correlation versus geological picks

Figure 3 shows the test of iteration process between the reference well and the target well (Blackfoot dataset). To assist in identifying the minimum correlation threshold for reliable picks, the absolute error between auto-picks and that of the expert's picks are plotted. It is noted that except, for a few outliers caused by bad log data, most of the auto-picks are reliable when the correlation coefficient exceeds a value of 0.6. For each formation top the correlation coefficients are plotted on a map to see the effectiveness away from the reference well. During the development of the auto-picking algorithm a vast number of tests were carried out to improve the results. Figure 3 also shows the autopicked map of the Milk River formation (120 wells) with a maximum absolute difference between expert picks and computer picks as less than 2 meters.

Assigning major sequence stratigraphic boundaries to all wells

Using the corrected V-shale curves, major sequence stratigraphic boundaries are assigned to all wells using method of Quantitative Sequence Stratigraphy (QSS). QSS is the automation of so-called "TSF analysis"; a step-by-step explanation of the method follows:

TSF Analysis

In his paper Ainsworth [1] introduced the so-called "TSF Analysis" to quantitatively identify important sequence stratigraphic surfaces (flooding surfaces and sequence boundaries) of different orders by applying a ratio of rate of accommodation to the rate of sediment

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

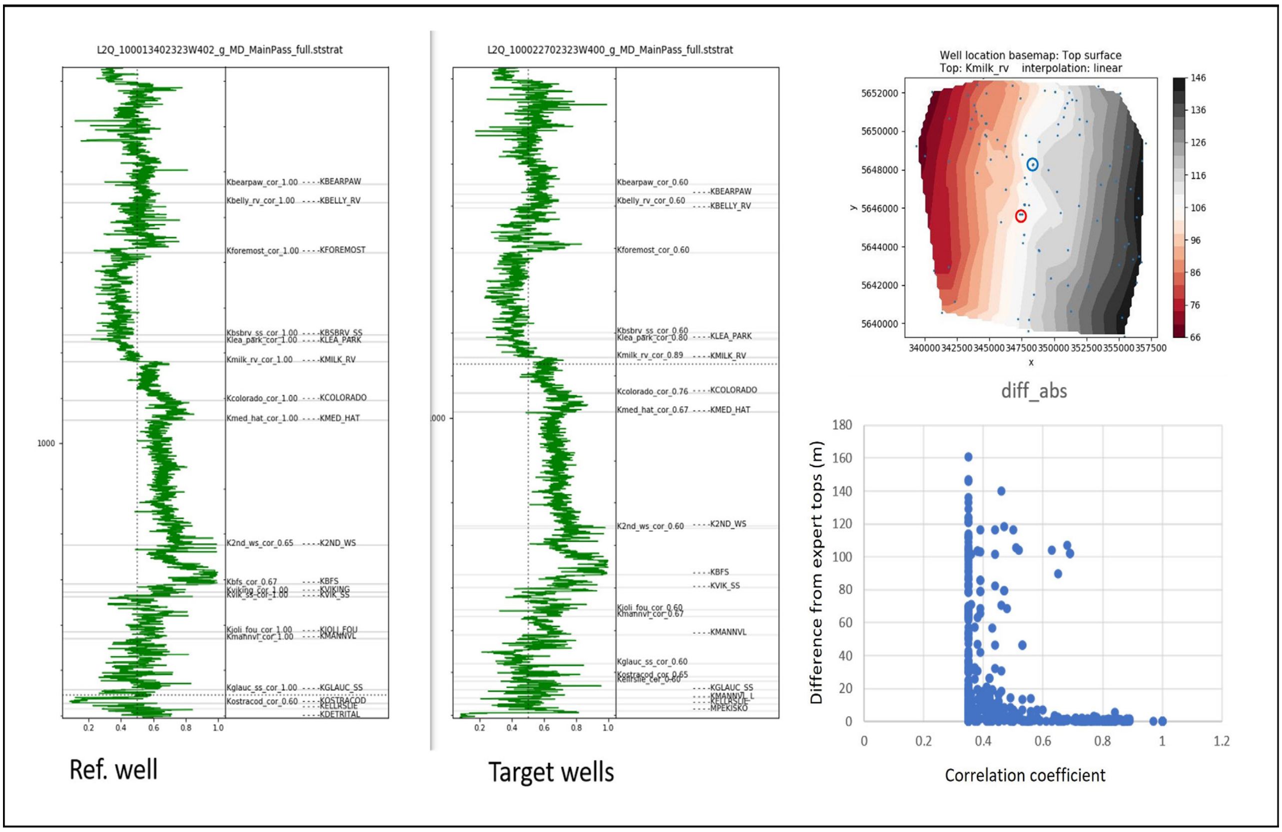

Figure 3: Automatic well tops correlation in the vicinity of reference wells, the cross-plot on bottom right shows relationship between auto-picks and expert picks. Auto-picked Milk River Fm. Map (sub-sea meters) is shown on the top right-hand side highlighting reference well (red) and target well (blue).

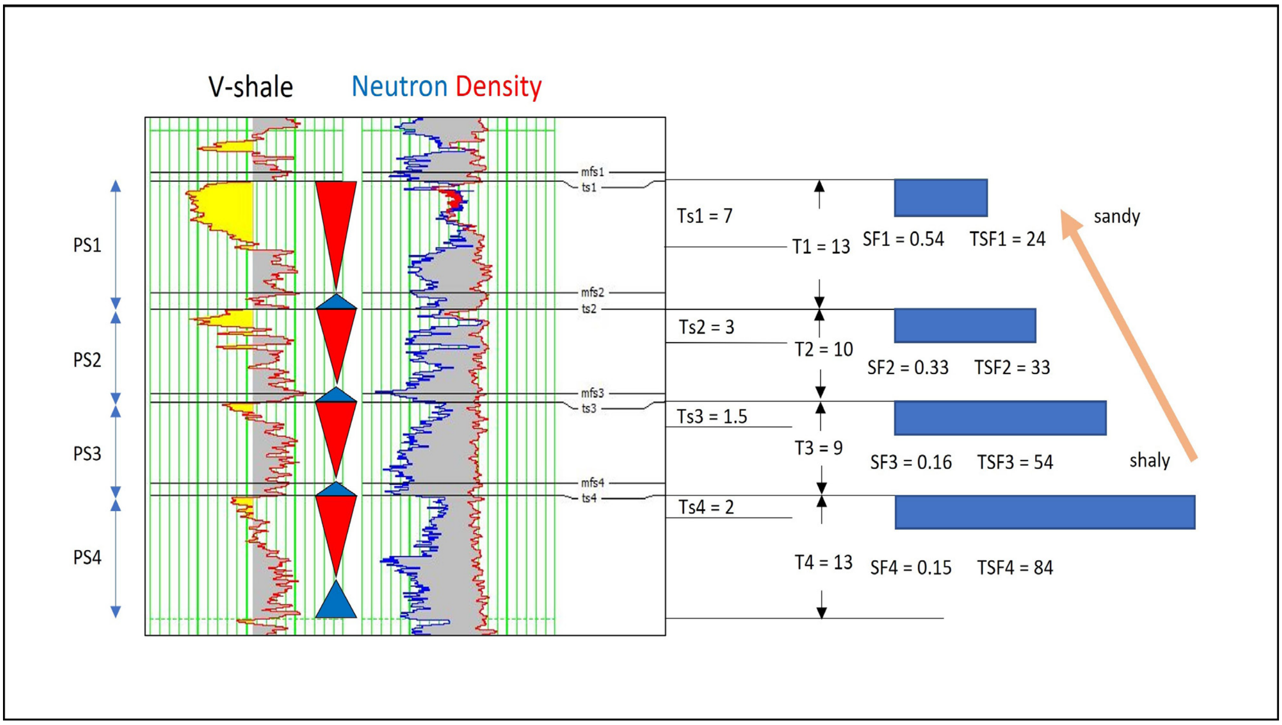

Figure 4: TSF analysis after Ainsworth [1], where TSF = T/SF where T = thickness of parasequence, Ts = thickness of sand within a parasequence and SF (sand fraction) = Ts/T. By re-arrangement of the equation TSF = T^2/Ts.

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

supply within individual parasequences. Application of the TSF analysis to a sample well is shown in Figure 4 where the sand ratio is increasing from older to younger parasequences.

In our present work a fully automated TSF analysis, referred to as the Quantitative Sequence Stratigraphy (QSS) analyses, is applied to discriminate sequence boundaries through application of the modified cascaded smoothing algorithms on a V-shale log. Traditional method of identifying sequence boundaries through picking T-R cycles (blue triangle to show fining upward sequence and red colored triangle for coarsening upward sequences), as published by Al-Husseini [16], shown in Figure 5. In our studies, the traditional T-R cycles are shown to the left of the QSS curve (Figure 4) to confirm that the sequences boundaries picked through the QSS method adhere to the sequence stratigraphic principles. The QSS curve represents an automatic calculation of TSF analysis, where the minimum value corresponds to the sequence boundary and the maximum flooding surface of different orders. It can be shown that QSS method is consistent and can be applied to large datasets to reach a solution in a fraction of time.

Application of smoothing algorithm

The process of calculating the QSS curve is shown in Figure 6. Ainsworth [1] has utilized a single baseline to differentiate sand and shale. Using a straight line could miss T-R cycles that are depicted by small variation in lithology especially within a silty/shaley sequence. In our case, we have modified the concept by applying a variable smooth background line derived from the log itself. We use a cascaded Stavitzky-Golay filter (SGF) first introduced by Savitzky and Golay [17] to achieve a baseline. For consistency in achieving similar parasequences output, we use first pass of SGF with 1001 samples rolling window followed by another pass of 101 samples. At large window length the SGF provides a centered moving average of V-shale log while removing all associated sand and shale lobes.

Next step is a short window smoothing of the V-shale curve. Most of the smoothing methods may obliterate the necessary interfaces required to be preserved in any V-shale log, except the Weiner filter [18]. Two passes of short window Weiner filter (WF) are applied to protect large interfaces while smoothing local variations. In the next

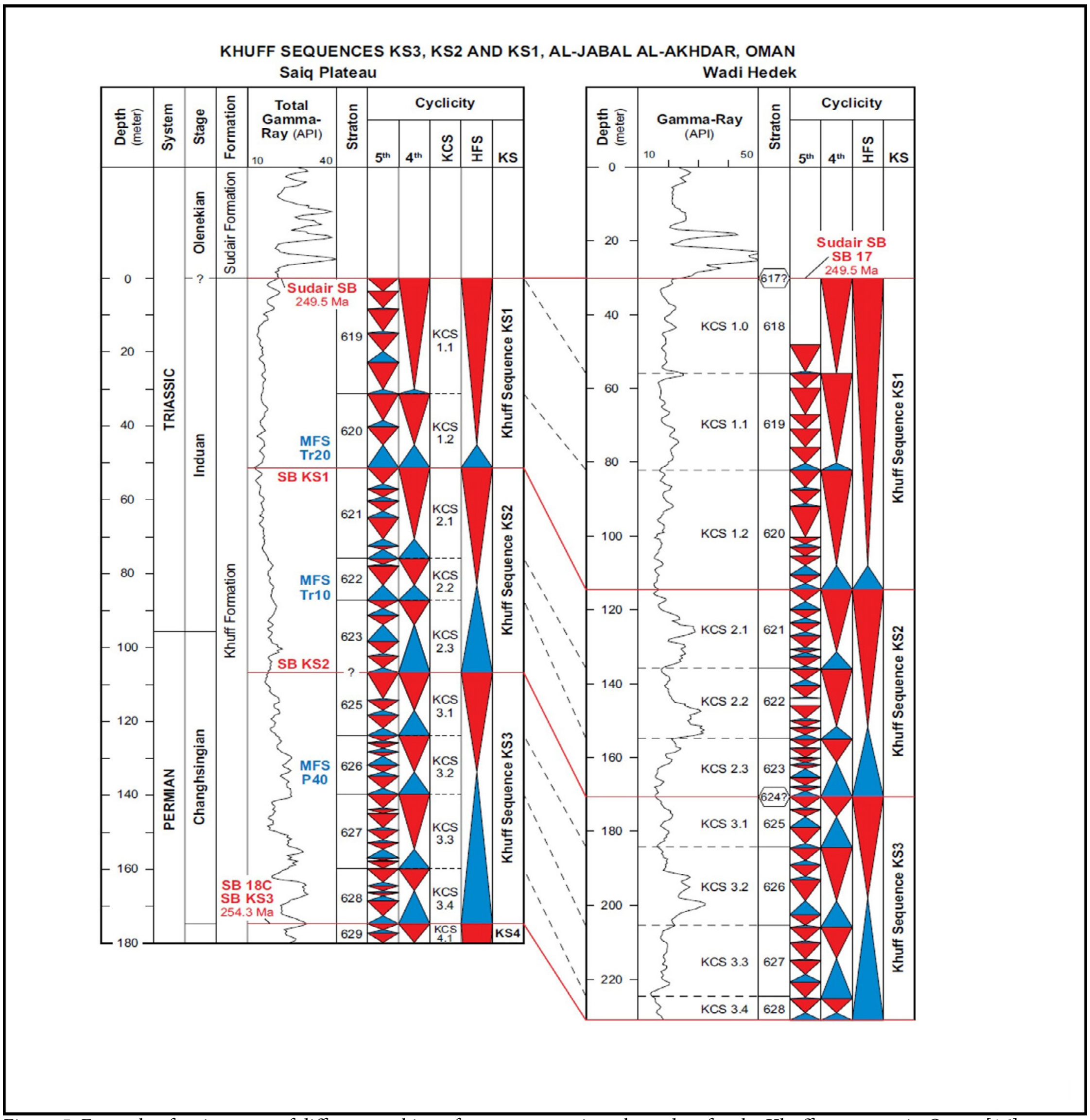

Figure 5: Example of assignment of different ranking of sequence stratigraphy orders for the Khuff sequences in Oman [16].

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

step the intersection values are found between WF log and SGF background log. In essence, when the WF log crosses to the left of the SGF log, it is marked as sequence boundary (red line) or the top of a parasequence and when it crosses to the right it is marked as base of the sand interval (blue line) (Figure 6). The total thickness of a parasequence, T, will be the distance between two consecutive red lines and sand thickness, Ts, will be between red line to blue line. The derived values are thus automatically carried over for calculation of the QSS log. To assist the expert, T-R- sequences (blue and red triangles) are also automatically generated.

Recognition of orders (ranking) of sequence stratigraphic surfaces

The QSS method not only allows identification of sequence stratigraphic surfaces of higher frequencies (5th order parasequences), but it can also automatically assign lower orders (4th order, 3rd order and even 2nd order in some cases) (from 10^4 -10^5 yr for the 5th order surfaces to 10^7-10^8 yr for the 3rd order sequences) as per Ainsworth [1] and Schlager [19].

The third and fourth order sequences [20] record base-level oscillations, tectonically induced changes in topography, sediment supply, and climatic variations. Traditionally, it is quite a difficult task, which is very time consuming and, in some cases, cannot even be achieved.

There are not many attempts made to address automatic identification of sequence stratigraphic surfaces and the assignment of hierarchy issue. Ye et al. [21] has proposed an interesting approach in automatically identifying stratigraphic boundaries and their hierarchy, where a well-log signal is transformed into a two-dimensional wavelet-scale and log depth representation using a continuous wavelet transform. The resulting multi-scale phase image of a well-log exhibits oval-shaped patterns, the hierarchy of which then assigned by the so-called a significance-of-cone method.

Figure 7 shows, schematically, how the ranking is calculated for one well. From the QSS curve derived from the automatic TSF analysis, one can see the depositional patterns of increasing or decreasing TSF values. If a pattern of increasing or decreasing TSF values exists, all the 5th order parasequences within the pattern are amalgamated to create a 4th order sequence. A new TSF value is derived by calculating total sand thicknesses and total thickness of the new sequence (parasequence stack). The 5th order parasequences are thus converted to 4th order sequences through the length of the log data. The calculation is again repeated to achieve the 3rd order sequence sets. Once applied to all the wells, geological stacking patterns and regional sequence boundaries can be observed.

As a result, we achieve a full suite of logs over the large area with the major sequence boundaries and marked tops identified with

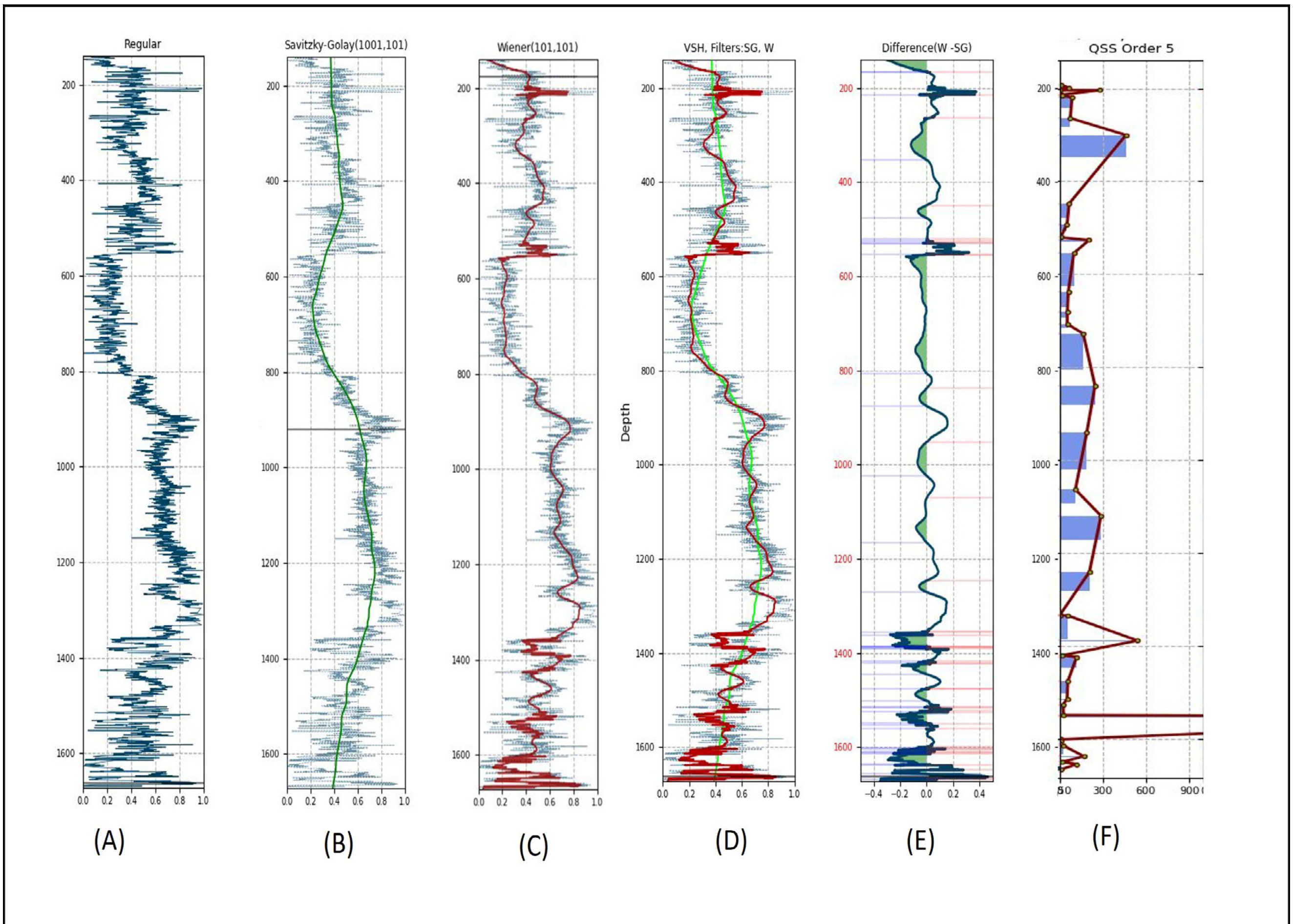

Figure 6: Smoothing algorithm applied to create Quantitative Sequence Stratigraphy (QSS) on Blackfoot data, well 01-08. (A) regular Vsh log of the reference well; (B) Savitsky-Golay filter, double pass; (C) Wiener filter, double pass; (D) Wiener (red), Savitsky-Golay (green) filters and VSH log (blue); (E) Difference = Wiener (double pass) - Savitsky-Golay (double pass) filter. (F) TSF results with QSS envelope.

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 208 doi: https://doi.org/10.15344/2456-351X/2022/208

assigned absolute age for every well. Figure 8 shows the process of automatic identification of sequence stratigraphic surfaces of the 5^{th} order for one of the well from the Western Canadian sedimentary Basin (WCSB) with assigned absolute geological age.

Automatic cross-section

Automatic cross-section building provides a mean to observe large amount of data for the QC of the workflow, regional understanding, and sequence stratigraphic interpretation. The cross-section builder can read data of, automatically picked tops, well logs and sequence stratigraphic interpretation of individual wells and automates the correlation process through efficient QC processes. The output of top correlation subroutine will automatically feed into the map program.

Automatic seismic interpretation and extraction of horizons/ faults

Automatic extraction of horizons and faults has been a challenge for many developers. Many methods have been tested and some provide quite impressive results for simple datasets, but generally fail in more complex regions such as salt domes, multiple fault families and unconformities. However, by implementing modern research, researchers have been approaching a completely automated seismic interpretation algorithm for an overly complex 3D seismic dataset [22]. The seismic based process is beyond the scope of this publication, but the resulting RGT cube is used to improve well log correlations.

Relative geological time cube (Wheeler Domain)

The concept of relative geological time (RGT cube) comes from the interpretation of seismic data where the seismic horizons are presented as age equivalent time markers. The RGT cube is derived from the seismic data after applying algorithms and methods described by Wu and Hale [2]. The relative geological time is converted to true geological time based on the age column of each well.

Case Studies

Blackfoot dataset (Western Canada Sedimentary Basin)

The Blackfoot field is in Township 23, Range 23, West of the 4^{th} Meridian, near Strathmore, Alberta (Figure 9). The Alberta Basin, a part of the Western Canadian Sedimentary Basin, extends from the Rocky Mountain Thrust Belt in the west to the Precambrian Canadian Shield in the northeast, and to the Sweetgrass Arch in the south. Blackfoot field, southeast of Calgary, has produced oil and gas mainly from a Glauconitic sandstone, a part of Cretaceous compound incised-valley system [23].

Geological history of the Blackfoot field is governed by its location on a wide coastal plain generated in response to the relative sea level fluctuation of different hierarchy (local, regional, and global). Sea level cycles and coastal on-laps can be identified based on occurrence of sand, shale, and coal.

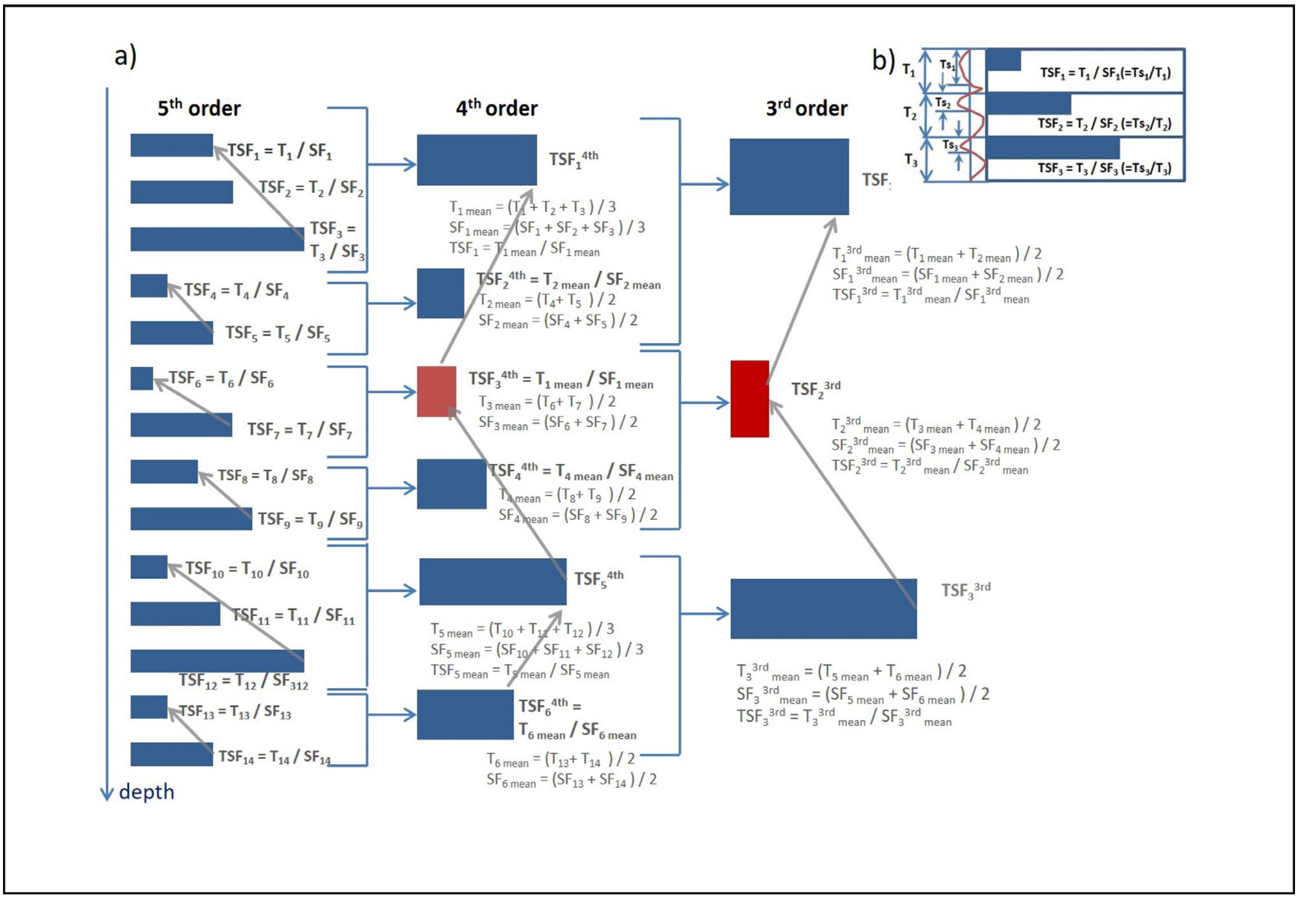

Figure 7: Automatic identification of different ranking orders of sequence stratigraphic surfaces: a. Calculation of lower rank sequences from higher to lower rank (from the 5^{th} to the 3^{rd} order); b. Estimation of parasequence thickness Ti and sand fraction SFi from VSH well log: TSFi = Ti / SFi.

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 208 doi: https://doi.org/10.15344/2456-351X/2022/208

Figure 9 also shows an isopach map of the Lower Manville depicting incised valley fill. 150 wells from the Blackfoot area including 10 wells targeting the incised valley are used for the Blackfoot case study, many of the wells are slightly deviated, especially those that fall on the 3D seismic.

All the wells in the Blackfoot dataset are drilled close to each other within uniform tectono-geological area and thus we have chosen only one reference well 01-08 for this study. Tops were manually picked on the reference well and then automatically extended for the rest of the wells through the correlation algorithm. Geological age is assigned based on the Absolute Age Chart (Table 1) for all wells in the study area. The Absolute Age Chart defines the depositional time for each formation and thus provides an additional control to the accuracy of the automated interpretation of the formation tops. The Absolute Age Chart for Blackfoot dataset is built on the bases of the Table of Formation of Alberta by AGAT Laboratories [24] and the Stratigraphic Correlation Chart [25].

Figure 10 shows the application of QSS to the well 01-08. QSS results under various filters separate para-sequences and coastal on-laps. Minimum QSS values correspond to the sequence boundaries of different orders while maximum values correspond to the flooding surfaces.

The flooding surfaces (FS) coincide well with the major geological events, from older to younger, flooding after the filling of the detrital sandstones above Mississippian unconformity, the end of Glauconite

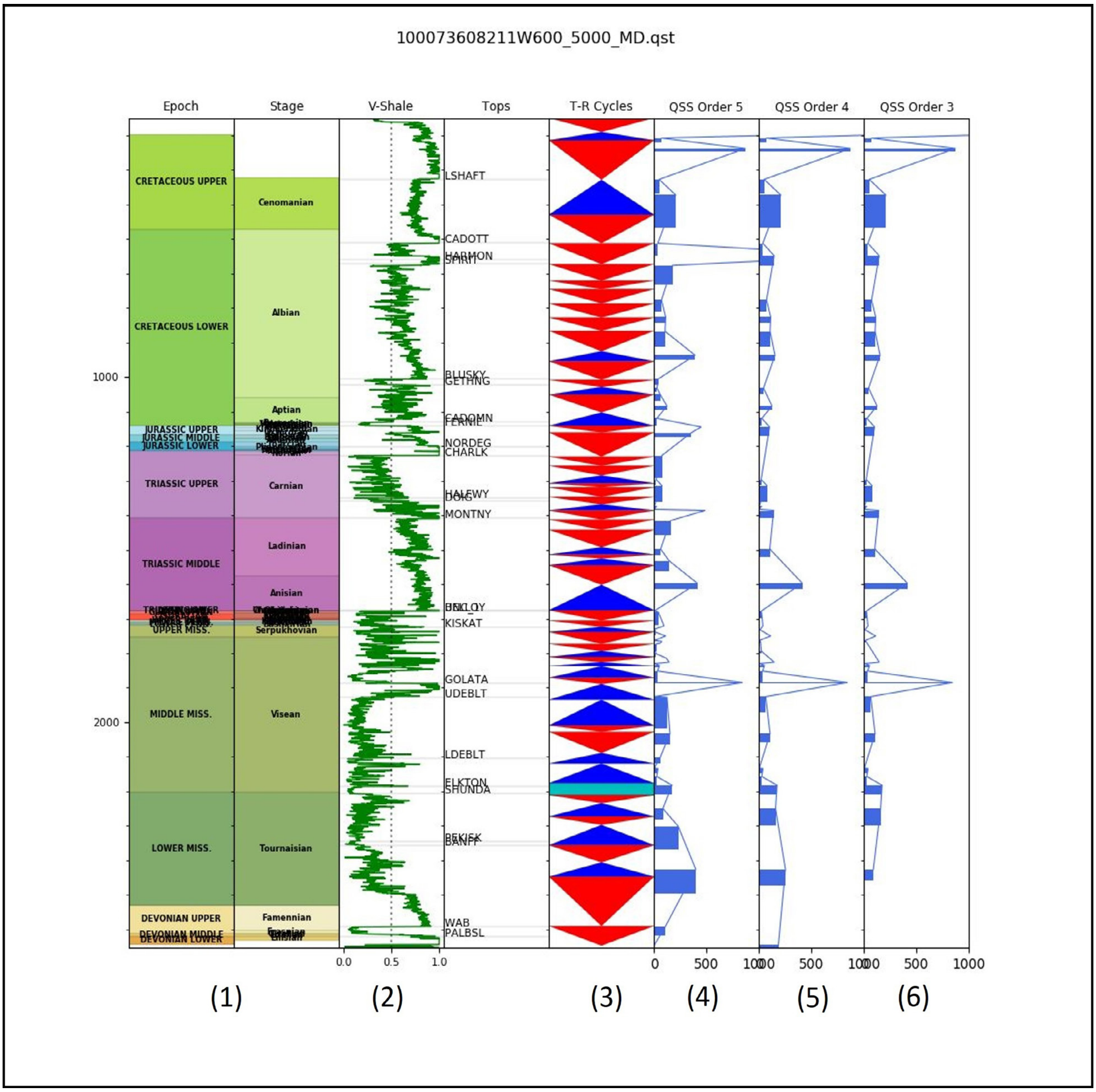

Figure 8: Automatic identification of sequence stratigraphic surfaces of the 5th order in the Montney well (WCSB). On the left (columns from left to right): 1 – Geological Age 2 -Vsh log with picked tops; 3 – traditional identification of sequence boundaries using T-R sequences; 4 – QSS curve Order 5; 5 – QSS Curve Order 4; 6 – QSS Curve Order 3.

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

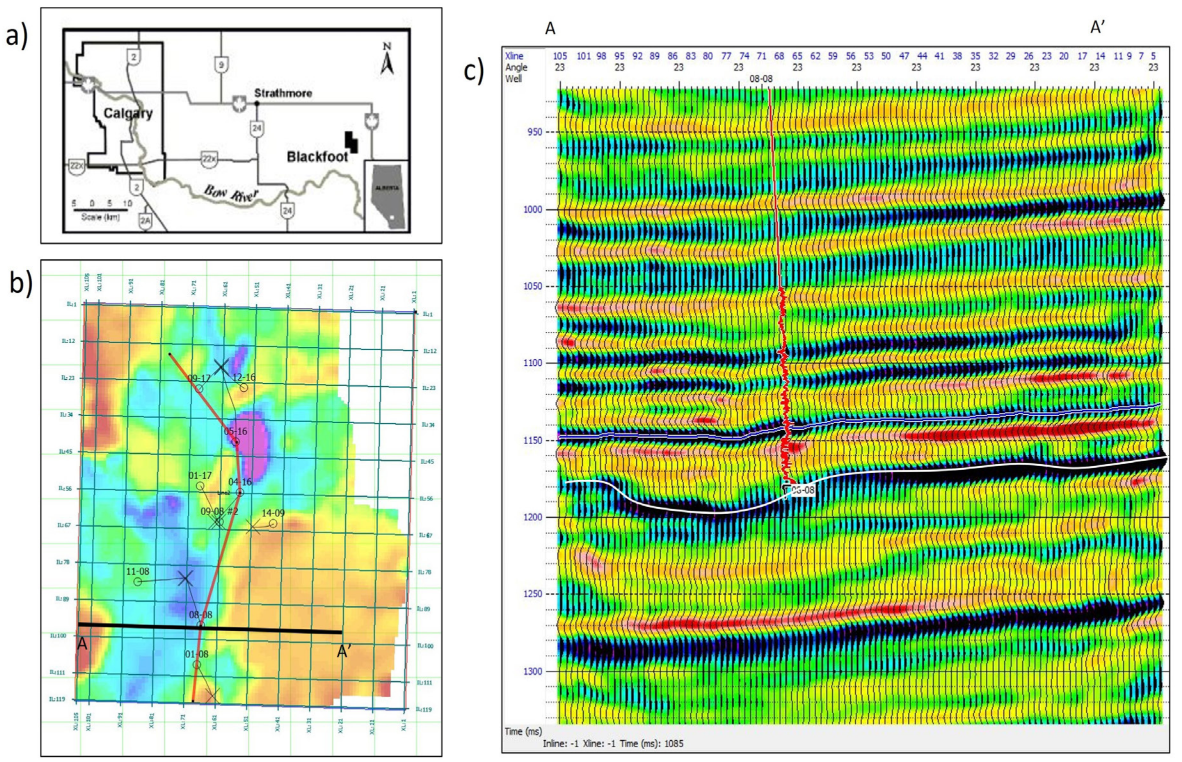

Figure 9: a) Location Map of Blackfoot Field Area. b) An isopach map showing the incised valley filled with Glauconitic sandstone. c) An E-W inline through 3D seismic showing the Well 08-08 penetrating the Glauconitic sandstone.

Top Full Name	TOP NAME	Age(Top)	Age(Bottom)
Bearspaw	Kbearpaw	74000	75000
Belly River	Kbelly_rv	75000	78000
Basal Belly River Sandstone	Kbsbrv_ss	78000	79000
Lea Park	Klea_park	80000	83500
Milk River	Kmilk_rv	83500	84000
Wapiabi	Kwapiabi	84000	84200
Colorado	Kcolorado	84200	84500
Medicine Hat	Kmed_hat	84500	84600
Second White Spec	K2nd_ws	93000	101000
Base Fish Scales	Kbfs	101000	103000
Viking Sandstone	Kvik_ss	103000	104000
Joli Fou	Kjoli_fu	104000	104500
Unconformity 1	Unc_1	104500	105500
Mannville	Kmannvl	105500	111000
Glauconitic Sandstone	Kglauc_ss	111000	111500
Ostracod zone	Kostracod	111500	124000
Detrital	Kdetrital	124000	125500
Unconformity 2	Unc_2	125500	350000
Pekisko	Mpekisko	350000	353000

Table 1: Absolute Age Chart used for Blackfoot dataset used as an input for geological age assignment.

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

regression (Lower Manville), the opening of Western Interior Seaway (Joli Fou Fm.) and the development of local flooding (Oldman Fm.). As for the marking of sequence boundaries, changes from coarsening upward to fining upward motifs, there is one right at the onset of the Upper Manville Group (a local lowstand event observed well on the QSS graph), another local sequence boundary corresponds to a lowstand event of Viking formation often recognized as a strandplain. Another major sequence boundary on the QSS curve is observed at Dinosaur Park Fm which marks changes from coarsening upward to fining upward events.

To discriminate the ranking of sequence stratigraphic boundaries the workflow described above is applied. Figure 11 shows result of automated identification of the sequence stratigraphic surfaces of lower ranking. Four 3rd order sequence boundaries are identified, SB1 corresponds with the top of Glauconitic sand, SB2 with Viking Fm, SB3 locates above 2nd White Spec Fm. (deeper equivalent of Cardium regression) and SB4 corresponds to the Dinosaur Park Fm. There are 5 major flooding surfaces automatically highlighted by QSS: FS1 is located just above sequence boundary of Glauconitic sands; FS2 corresponds to Joli Fou Fm.; FS3 is the event of the deepening of WIS,

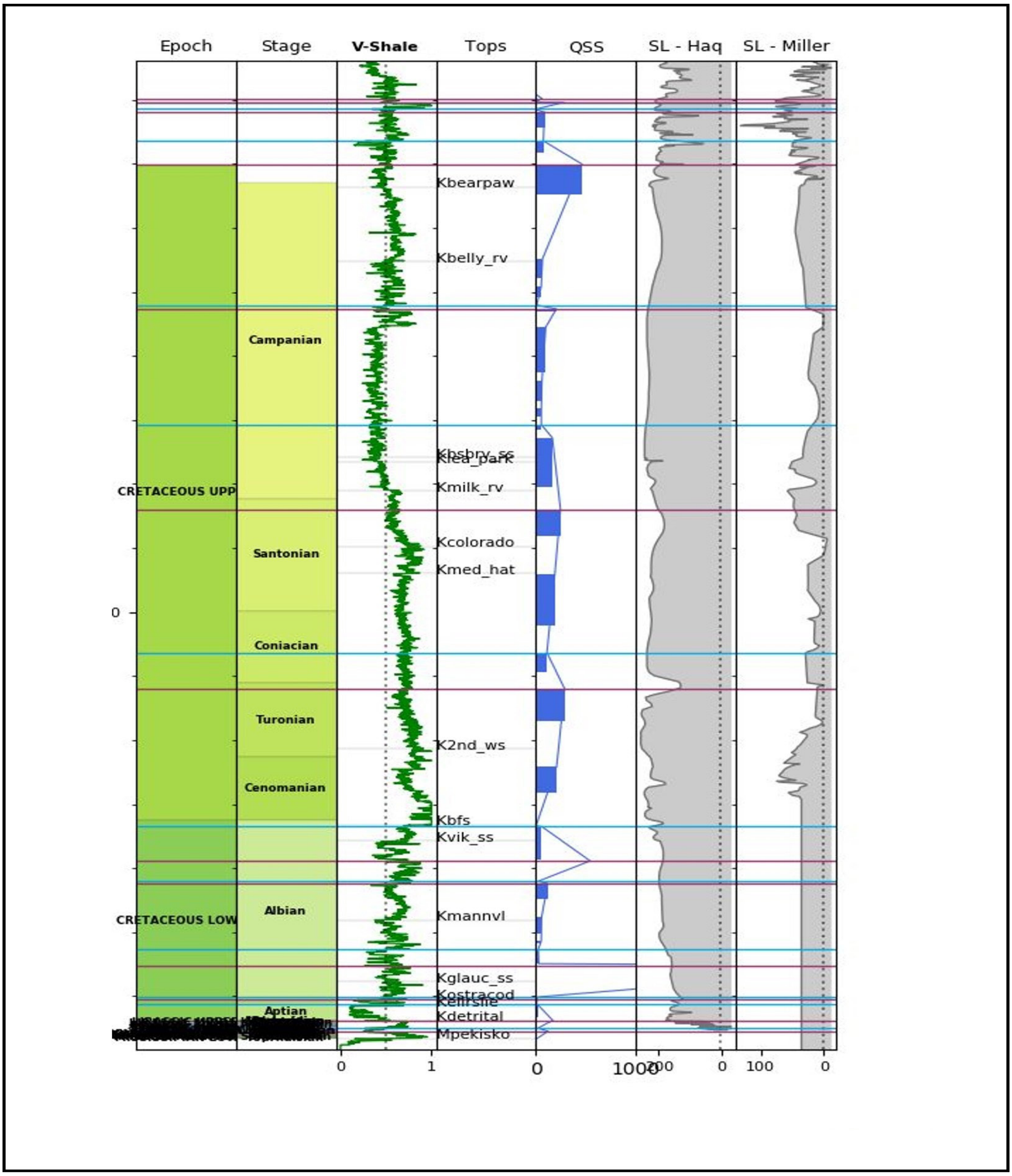

Figure 10: Example of automatically picked tops, geological age, sea-level curves and QSS applied to the well 01-08 with sequence boundaries (in red) and flooding surfaces (in blue).

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

FS4 is the last flooding event before separation of GOM and WCSB and FS5 is the local flooding event above Dinosaur Park Fm.

The cross-section built through the wells is shown on Figure 12. Each well in the cross-section displays VSH log, followed by QSS results and geological age. Sequence boundaries and flooding surfaces of different stratigraphic orders are picked automatically for each well and added to the cross-section. Then, based on the ranking of the sequence stratigraphic surfaces, orders of the surfaces were identified and added automatically to the cross-section. Sequence boundaries of the 3rd order are surfaces around Glauconitic sandstones and Dinosaur Park Fm, while flooding surfaces of the 3rd order are surfaces around Joli Fou and above Dinosaur Park formation. Those 3rd order surfaces could be traced throughout the whole area of Blackfoot field as regional surfaces.

Teapot Dome dataset (Wyoming, USA)

The Teapot Dome oilfield, located in central Wyoming on the southwest flank of the Powder River Basin, contains three shale formations, the Steele, Niobrara, and Mowry, which are known to produce as unconventional reservoirs in several nearby basins (Figure 13) [26]. As in the case of Blackfoot field, the geological history of the Teapot Dome is governed by its location on a coastal plain in response to the relative sea level fluctuation of different hierarchies (local, regional, and global).

There are more than 1300 wells in the Teapot Dome dataset with tops picked on 964 wells ranging from shallow up to 500 meters depth to deeper up to 1800 meters (~5500 ft) (317 deep well logs and 747 shallow well logs). Figure 13 shows the wells used for this study displayed over Tensleep time surface.

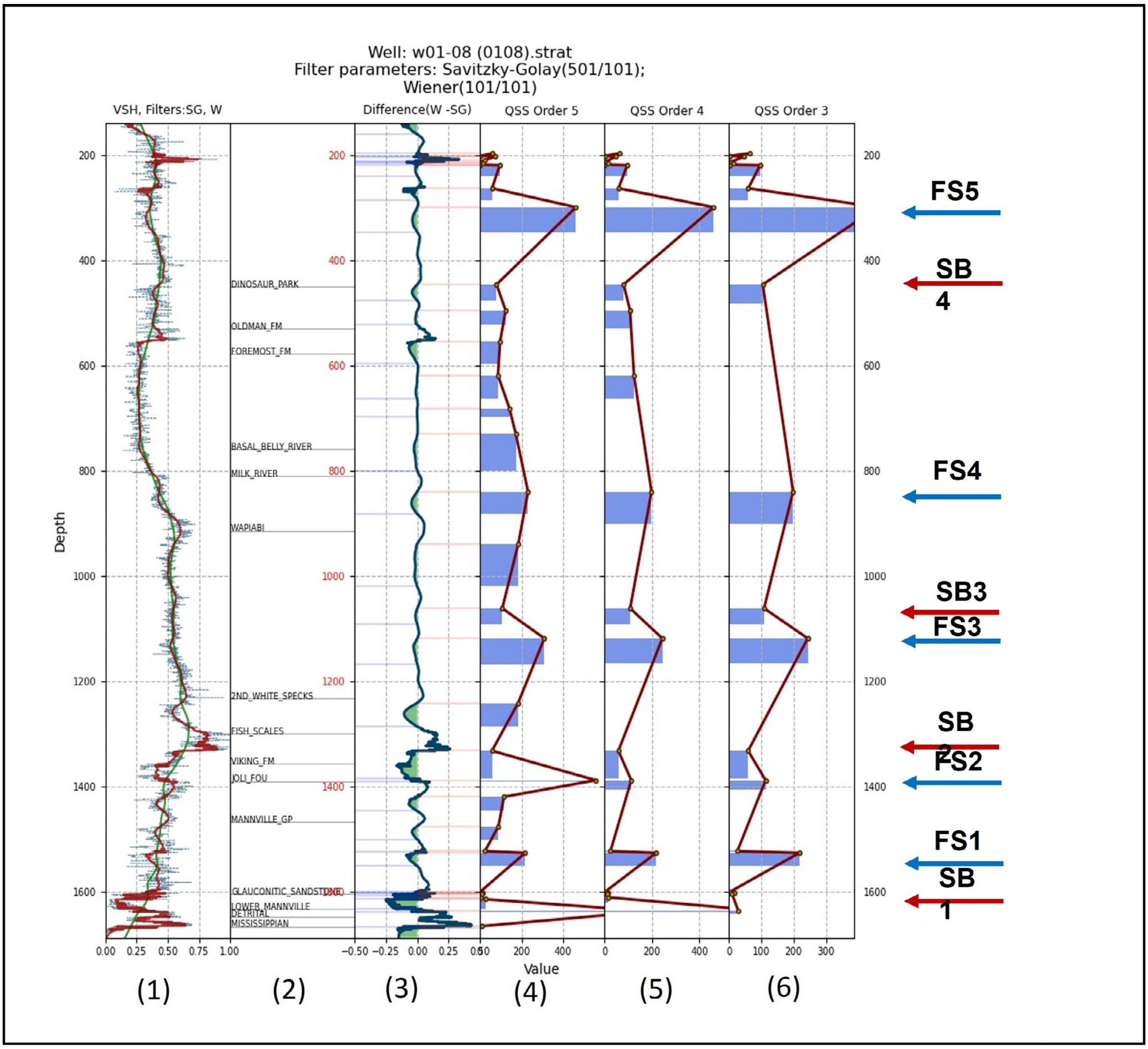

Figure 11: Automatic identification of sequence stratigraphic surfaces of (from left to right): 1 - application of Savitsky-Golay filter; 2 – automatically picked tops; 3 - application of Wiener filter and automatically identified sequence boundaries (in red) and flooding surfaces (blue); 4 - 5th order events; 5 - 4th order events; 6 - 3rd order events, red arrows - sequence boundaries (SB) and blue arrows flooding surfaces (FS) for reference well 01-08 identified for the 3rd order event.

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

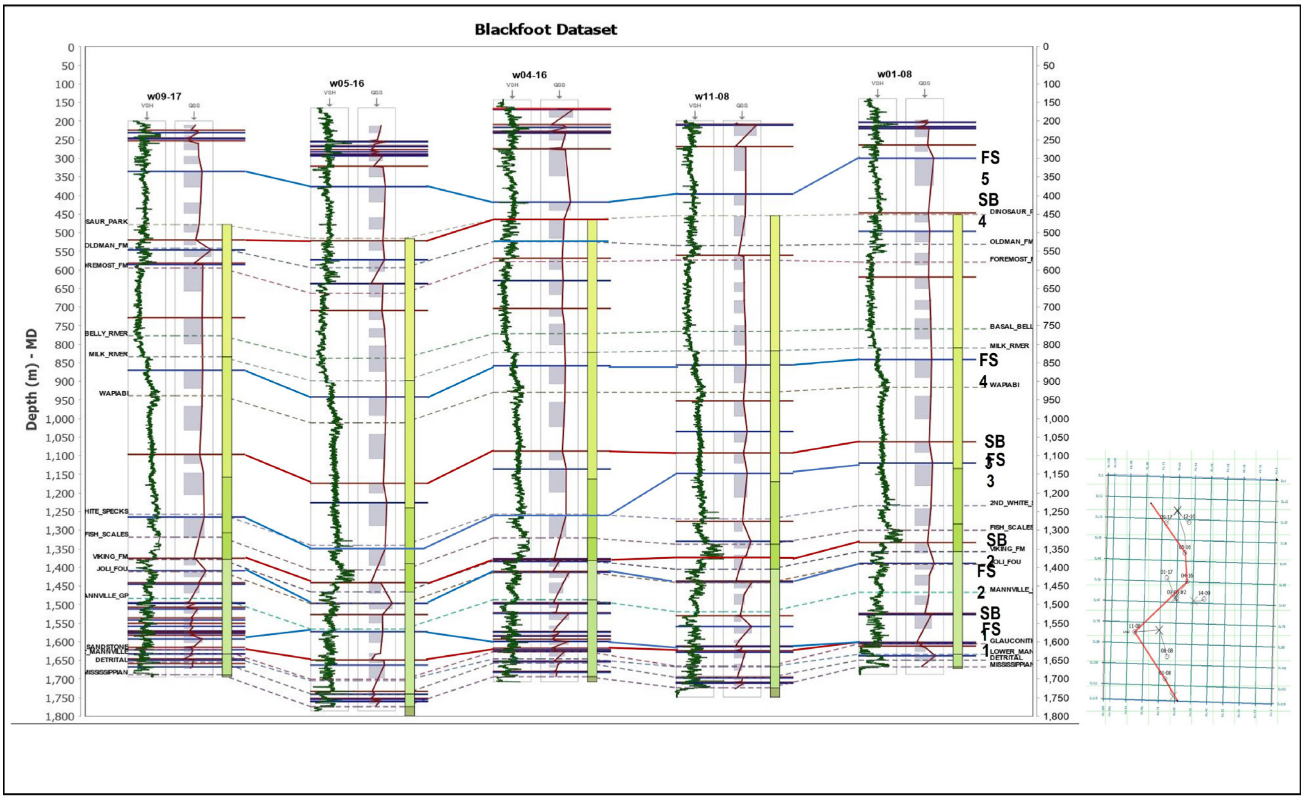

Figure 12: Automatic cross-section built for Blackfoot Dataset. Each well displays Vsh curve, QSS curve, Geological Age and sequence stratigraphic surfaces picked automatically. Regional 3rd order sequences are connected.

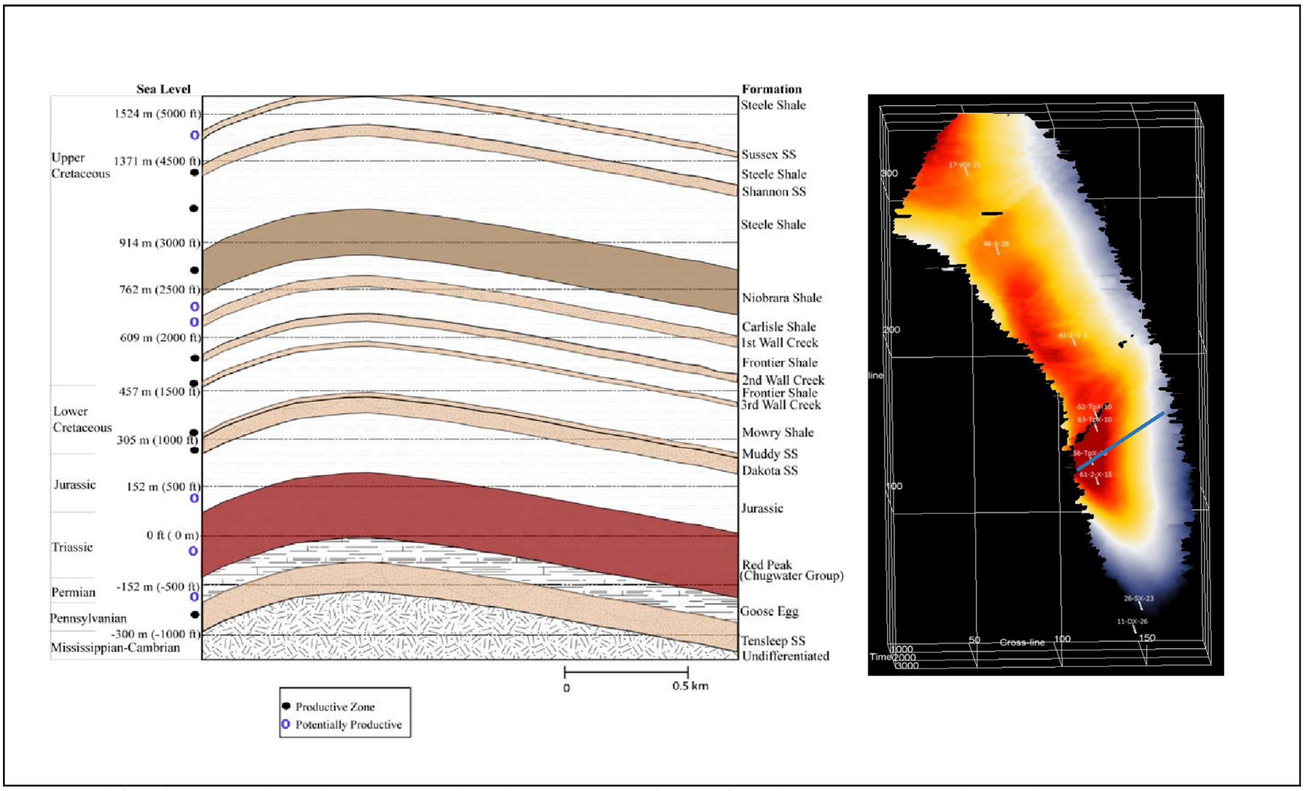

Figure 13: (left) Schematic geological cross-section through Teapot Dome dataset after Okojie-Ayoro, [27] (right) Tensleep time structure map Teapot Dome 3D dataset.

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

For the Teapot Dome dataset, as all the wells drilled in proximity to each other, only one reference well (Department X #2-3, 490252305400) is selected to run top autocorrelation analysis. Figure 14 shows an example of automatic top picking for the 45 wells with the reference well displayed in red.

Table 2 shows Absolute Age Charts for Teapot Dome data set used to assign geological age for every well in the dataset. Note that there were no tops between Jurassic and Lower Cretaceous, and Triassic and Permian, thus these tops were interpolated to find the boundaries between Age Units. In this case, Upper Cretaceous split into Steele, Niobrara Shale, Carlisle Shale and Frontier formations, and the main advantage of the automation is that this could be done at once for all wells after Absolute Age Chart is created or updated.

Figure 15 shows the QSS method applied to the reference well 25-1-x-14 as well as tops and geological age picked automatically on the left and on the right the figure shows the same plus identification of sequence boundaries (red lines) and flooding surfaces (blue lines) of the 5th order picked automatically as well for this well. There are many sequence stratigraphic surfaces identified throughout the well from bottom to top stretching from Late Carboniferous (Gzhelian) until Late Cretaceous (Campanian). The absolute age chart (Table 2) is created based on Wyoming Stratigraphic Nomenclature Chart [28].

Most of the flooding events of the 5th order corresponds well with the main shaly formations such as Niobrara, Mowry, Dakota etc. while the sand deposit (Tensleep, CRMT, MRSN and F2WC) relate to the sequence boundaries, reflecting local lowstand events, coinciding well with the coastal onlap data.

Figure 16 displays identification of the ranking of sequence stratigraphic events (sequence boundaries and flooding surfaces) from 5th order sequences through 3rd order sequences. There are 5 sequence boundaries and 4 flooding surfaces identified as 3rd order events. From bottom to top, SB 1 event corresponds well to MNKT Fm, FS1 marks significant change in depositional environment, SB2 corresponds with the SNDCu FM; FC2 with F3WC, and SB3 with F2Wc Fm, FS3 with CRLL Fm.

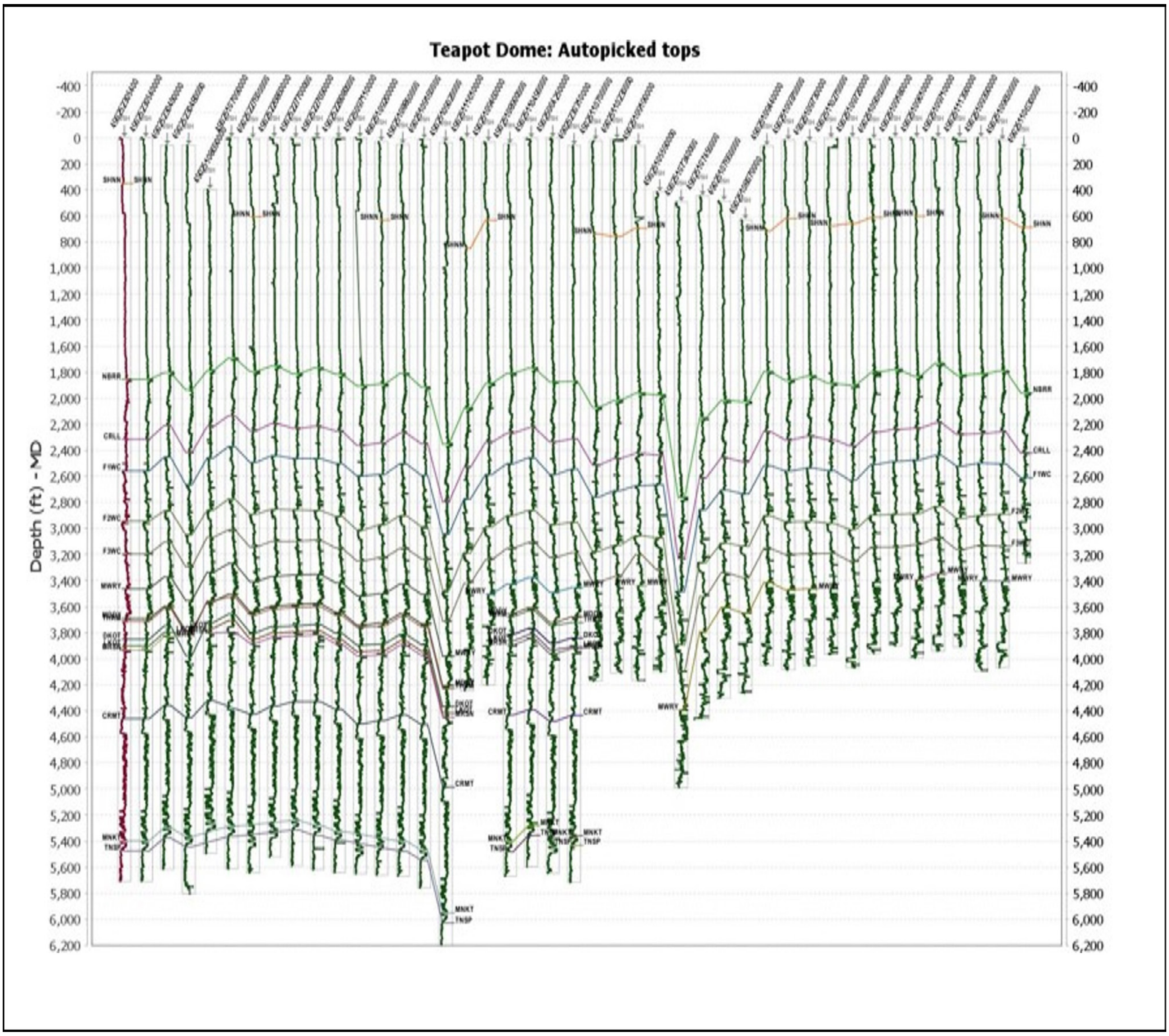

Figure 14: Example of well tops auto-picking performed on 45 randomly selected wells. Reference well Department X #2-3 displayed on the left in red.

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

The concept of introducing the strata and age gaps related to the faults is explained through a very good synthetic well tie for the well 62-TpX-10 to the seismic data. The well crosses through a fault at approximately 5050 ft TVD. A thin Tensleep formation is encountered at the hanging wall of the fault crossing to the footwall with a throw of 100 ft. The missing 100 ft is approximately equal to 3000 years in geological time.

Then, the cross-section through the Teapot Dome field is created automatically using the selection of wells along strike. As in the case of cross-section for the Blackfoot field, the one depicted on Figure 18 displays for each well VSH, QSS result, and geological age. All previously identified sequence stratigraphic surfaces (SB and FS) of the 3rd and the 5th orders added for each well. The 3rd order stratigraphic surfaces of sequence boundaries are (from bottom to top) around Tensleep Fm and MDDY Fm, and as for the flooding surfaces there is only one identified as 3rd order sequence - surface in between MNKT and CRMT and considered as the regional sequence stratigraphic surface that can be traced throughout the whole area.

Formation Name	Top Name	Age(Top)	Age(Bottom)
STEELE SHALE 1	StSH1_76600	76600	80700
SUSSEX SST	SSXSS	80700	80900
STEELE SHALE 2	StSH2	80900	81000
SHANNON SST	SHNNu_0800	81000	81250
STEELE SHALE 3	StSH3	81250	82500
NIOBRARA SHALE	NBRRws_0850	82500	89250
CARLISLE SHALE	CRLL_0880	89250	90000
FRONTIER 1ST WELL CREEK	F1WC_0910	90000	90600
FRONTIER SHALE 1	F1SH	90600	92000
FRONTIER 2ND WALL CREEK	F2WC_0930	92000	93000
FRONTIER SHALE 2	F2SH	93000	93900
FRONTIER 3RD WALL CREEK	F3WC_0950	93900	94300
BELLE FOURCHE	F3SH	95250	97250
MOWRY SHALE	MWRY_0980	97250	100500
MUDDY FORMATION	MDDY_1005	100500	103500
THERMOPOLIS SHALE	THRM_1020	103500	108000
DAKOTA SANDSTONE	DKOT_1080	108000	112000
LAKOTA SANDSTONE	LKOT_1160	112000	145500
MORRISON FORMATION	MRSN_1450	145000	155700
SUNDANCE FORMATION	SNDCu_1573	155700	165000
CHUGWATER	CHWR_220000	220000	227000
CROW MOUNTAIN	CRMT_2270	227000	250000
GOOSE EGG FORMATION	GSEG_250000	250000	252000
MINNEKAHTA LIMESTONE	MNKT_2688	268800	279200
OPECHE SHALE	OPCH_2710	279200	298900
TENSLEEP	TNSP_298900	298900	308000
AMSDEN FORMATION	AMSD_308000	308000	332000
MADISON LIMESTONE	MDSN_332000	332000	348000
Devonian and Older	BSMT_358000	358000	500000

Table 2: Absolute Age Chart used for Teapot Dome dataset.

Conclusions

An automated sequence stratigraphic workflow is proposed where raw data is intelligently investigated using the geological principles and expert supervision. There are three significant points addressed in our research. Firstly, it is the automation of sequence stratigraphic workflow on its own, from receiving raw well log data through the building of the geological cross-sections and maps allowing significant reduced time of interpretation through large datasets. Secondly, it is auto-correlation of formation tops over large datasets. Lastly, it is the automation of the identification of parasequences and major stratigraphic surfaces (sequence boundaries and flooding surfaces) that provides sequence stratigrapher a base canvas and insight to creativity.

During the study we have managed to overcome many complexities of automation, though there are few issues and limitations that remain to be solved and they are our goals for continued research. Use of single log type (V-shale) does not allow the automation to achieve the best correlation and sequence interpretation. Petrophysics automation in the future would permit creation of logs representing volumes of carbonates, quartz, clay, coal and porosity and thus the top correlation process could be further refined. Use of faults and unconformities rely on their good time-to-depth conversion, where quite often the seismic data does not have good enough well-ties to accurately build a velocity model. The future work will also expand into and include dipmeter, geochemical, petrography, outcrops and core analyses, as well as biostratigraphy information to automatically assign age.

The technique described above has been continuously tested and applied on various additional datasets. One of the examples is the application of QSS analyses to identify sequence stratigraphic surfaces for the Montney dataset in the Western Canada Sedimentary Basin by the authors where three stages of deposition within Montney formation were identified automatically [29]. The results were then compared with the published data. The authors also applied workflow to Western Canada Sedimentary Basin well log dataset which contains more than 50000 well logs, spread over 4000 sq km, achieving automated QSS curves within 3 days of computer time. The tops correlation was not possible for such large area without expert interpretation of large number of reference wells.

The proposed workflow can find various practical applications, in addition to its use in building geological model. It can be used for preparation, creation and maintaining of well log database for the companies to fast identification of sequence stratigraphic surfaces for divestiture and investments in hydrocarbon exploration. The workflow can also be used to screen formations that could function as Carbon Sequestration and Storage (CSS) targets.

Competing Interests

The authors declare that they have no competing interests.

Author's Contributions

The general development of QSS analysis concept and automated sequence stratigraphy workflow are done together by Valentina Baranova and Azer Mustaqeem while the coding and application of proper computer algorithms are performed by Aziz Mustaqeem who had an intern position with the company. Azer Mustaqeem and Aziz Mustaqeem developed a mathematical approach behind the quantitative sequence stratigraphy algorithm. Valentina Baranova was mainly involved in the drafting and preparation of the manuscript for the publication. Aziz Mustaqeem contributed a lot with the final English proof-reading of a manuscript.

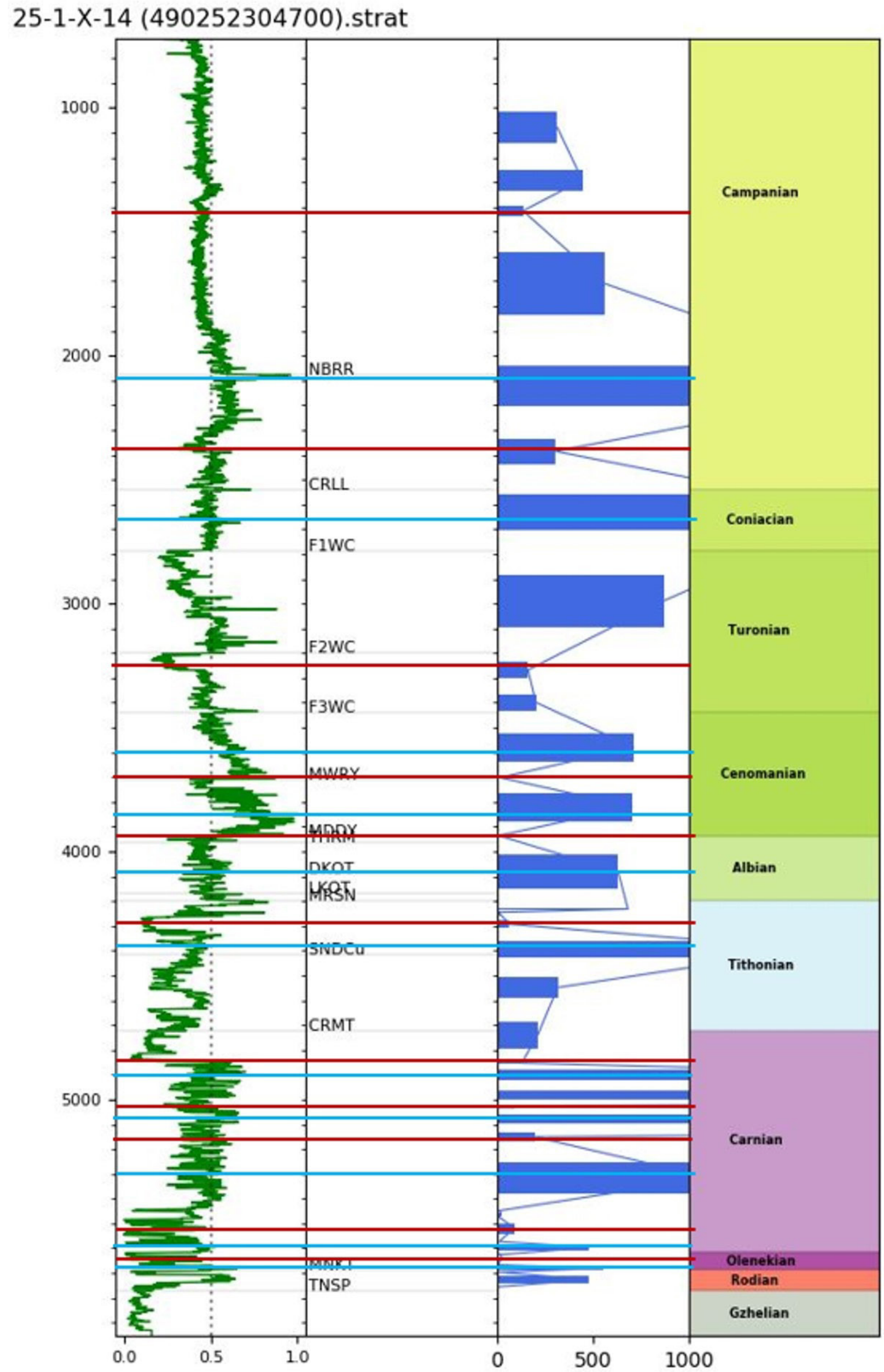

Figure 15: Example of geological age, tops assignment and QSS method applied to the well 25-1-x-14. Left - tops, geological age, QSS; Right - tops, geological age, QSS, sequence boundaries (in red), flooding surfaces (in blue) of different orders generated automatically.

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

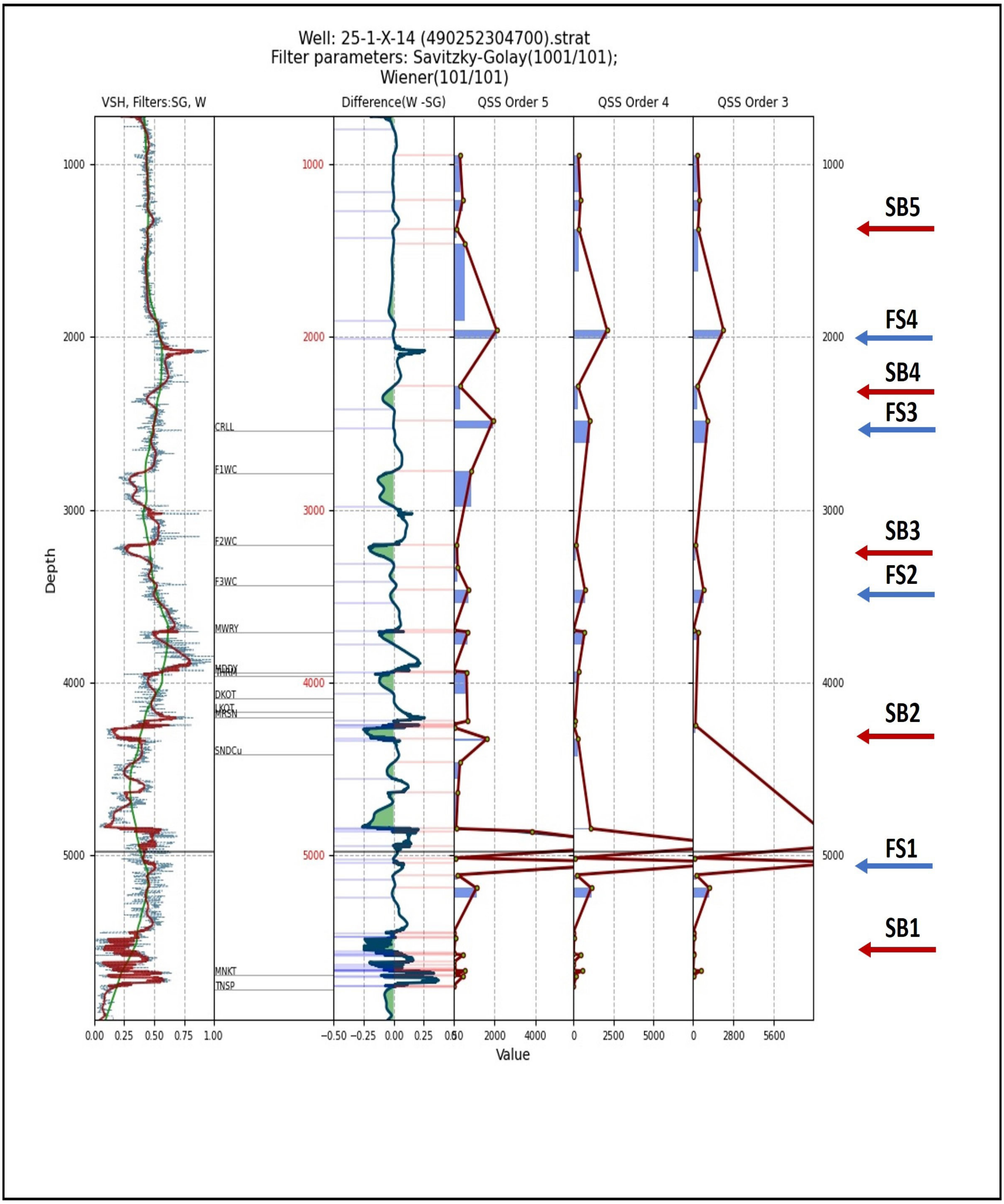

Figure 16: Automatic identification of lower ranking for the well 25-1-x-14 Teapot Dome dataset. Automatic identification of sequence stratigraphic surfaces of (from left to right): 1 - application of Savitsky-Golay filter; 2 – automatically picked tops; 3 - application of Wiener filter and automatically identified sequence boundaries (in red) and flooding surfaces (blue); 4 – QSS curve for the 5th order events; 5 – QSS curve for the 4th order events; 6 – QSS curve for the 3rd order events, red arrows - sequence boundaries (SB) and blue arrows for flooding surfaces (FS) for the well 25-1-x-14 identified for the 3rd order event.

Citation: Baranova V, Mustaqeem A, Mustaqeem A (2022) Automated Workflow for Sequence Stratigraphy Interpretation of Big Data with Expert Supervision. Int J Earth Environ Sci 7: 196 doi: https://doi.org/10.15344/2456-351X/2022/196

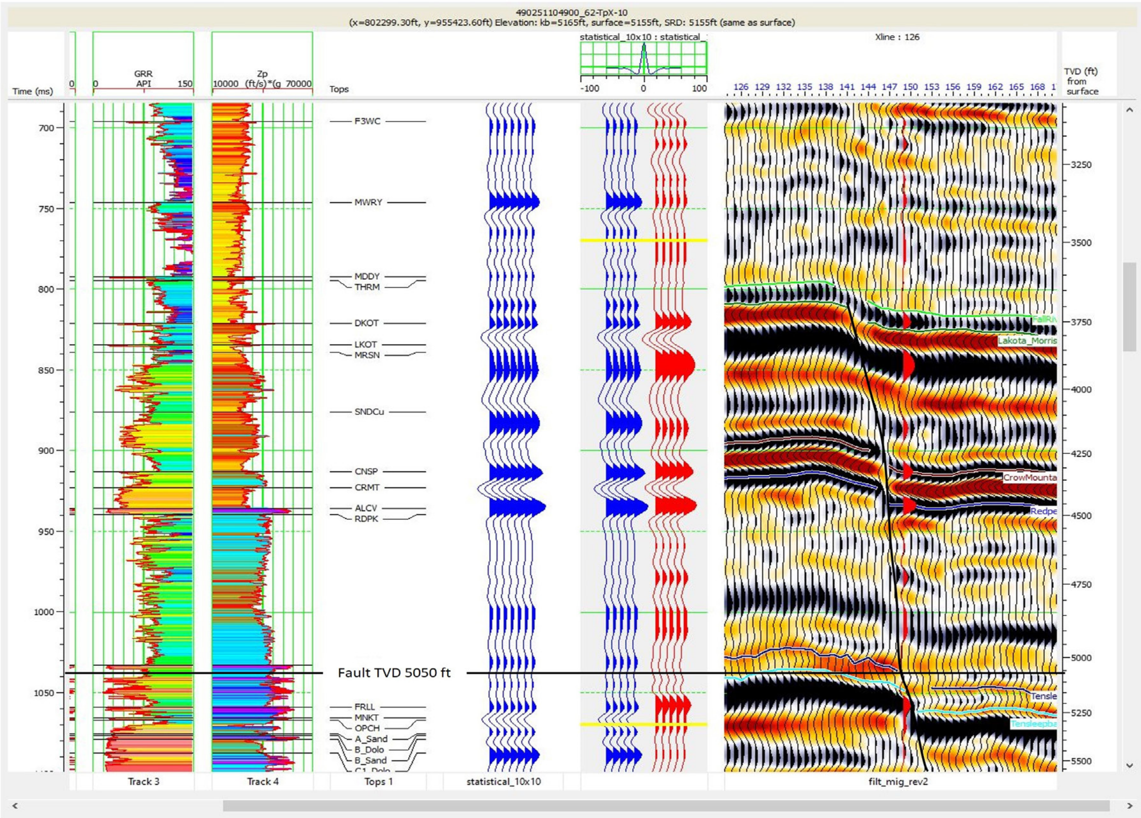

Figure 17: Synthetic Tie between Well 62-TpX-10 and crossline 126 through the Teapot Dome 3D. The well passes through a fault and a 50 ft fault gap is observed at ~5050 ft TVD.

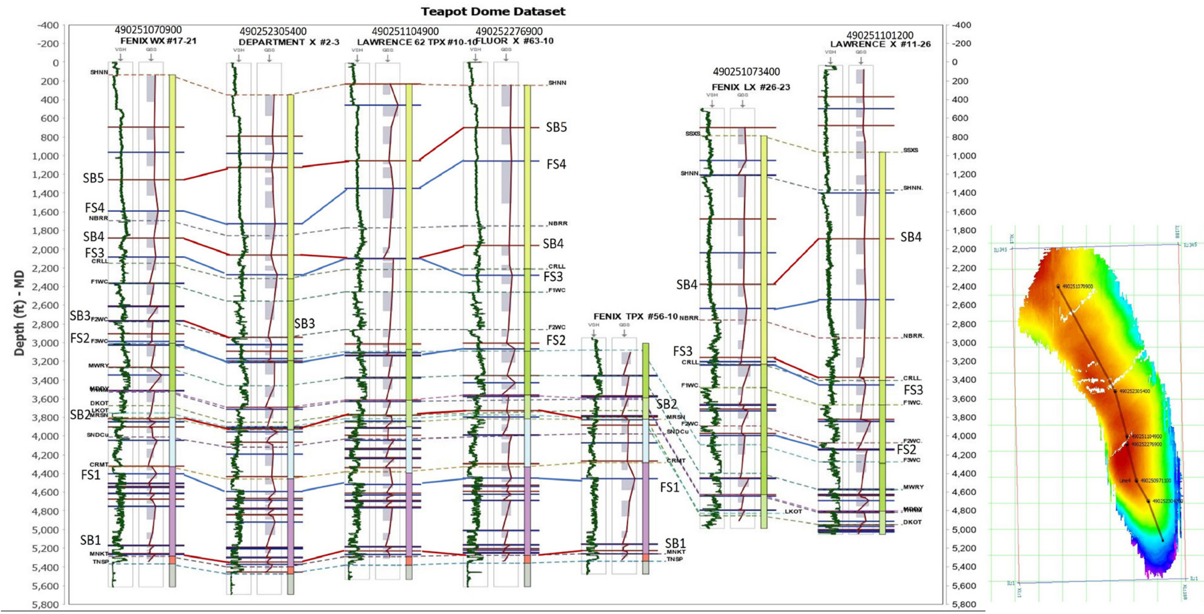

Figure 18: Automatic cross-section built for Teapot Dome dataset.

Acknowledgements

We would like to acknowledge dGB Earth Sciences for providing Sequence Stratigraphy plug-in to use in this paper, CREWES project of the University of Calgary for sharing Blackfoot dataset and Statcom Inc. for reprocessing it and RMOTC (US Department of Energy) for sharing Teapot Dome dataset.

References

1. Ainsworth RB, McArthur JB, Lang SC, Vonk AJ (2018) Quantitative sequence stratigraphy. AAPG Bulletin V 102: 1913-1939.
2. Wu X, Hale D (2016) 3D seismic image processing for faults. Geophysics 81: IM1-IM11.
3. Wu X (2017) Structure-, stratigraphy- and fault-guided regularization in geophysical inversion. Geophysical Journal International.
4. Stieber RG (1973) Optimization of shale volumes in open hole logs. Journal of Petroleum Technology 31: 147-162.
5. Larionov VV (1969) Radiometry of boreholes (in Russian): NEDRA, Moscow.
6. Clavier C, Hoyle W, Meunier D (1971) Quantitative interpretation of thermal neutron decay time logs: part I. Fundamentals and techniques. J Petrol Technol 23: 743-755.
7. Jaglan H, Kocsis G, Lakhliffi A de Groot P (2021) Experiences with machine learning and deep learning algorithms for seismic, wells and seismic-to-well applications. EAGE 2021 Annual Conference, Oral Presentation in Digitalization and AI: Reservoir and Wells.
8. Cohen KM, Harper DAT, Gibbard PL (2020) ICS International Chronostratigraphic Chart 2020/03: International Commission on Stratigraphy, IUGS.
9. Haq BU, Shutter SR (2008) A chronology of Paleozoic sea-level changes. Science 322: 64-68.
10. Shi Y, Wu X, Fomel S (2017) Finding an optimal well-log correlation sequence using coherence-weighted graphs. SEG International Exposition and 87th Annual Meeting.
11. Wu X, Shi Yu, Fomel S, Li F (2018) Incremental Correlation of Multiple Well Logs Following Geologically Optimal Neighbors. Interpretation.
12. Gosses J, Zhang LA (2019) Supervised Machine-Learning Approach to Stratigraphic Surface Picking in Well Logs from the Mannville Group of Alberta, Canada. Search and Discovery Article.
13. Friedman B (2020) A New Approach to Automated Well-Log Correlation in Three Dimensions. AAPG Explorer.
14. Brazell S, Bayeh A, Ashby M, Burton D (2019) A Machine-Learning Based Approach to Assistive Well-Log Correlation. Petrophysics 60: 469-479.
15. Karimi AM, Sadeghnejad S, Mansoor R (2021) Well-to-well Correlation and Identifying Lithological boundaries by Principal Component Analysis of Well-logs. Computers and Geosciences 157: 104942.
16. Al-Husseini M, and Koehrer B (2013) Chrono-and sequence stratigraphy of the Mid-Permian to Early Triassic Khuff sequences of the Arabian Plate. Geoarabia Manama 18: 103-130.
17. Savitzky A, Golay MJE (1964) Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry 36: 1627-1639.
18. Norbert W (1949) Extrapolation, Interpolation, and Smoothing of Stationary Time, Series. New York: Wiley.
19. Wolfgang S (2010) Ordered hierarchy versus scale invariance in sequence stratigraphy. Int J Earth Sci.
20. Catuneanu O, Bhattacharya JP, Blum MD, Dalrymple RW, Eriksson PG, et al. (2010) Sequence stratig-raphy: common ground after three decades of develop-ment. First Break 28, 21–34.
21. Ye S, Wellner RW, Dunn PA (2017) Rapid and Consistent Identification of Stratigraphic Boundaries and Stacking Patterns in Well-Logs - An Automated Process Utilizing Wavelet Transforms and Beta Distributions. SPE-187264-MS.
22. Wu Y, Schuster M, Chen Z, Le QV, Norouzi M, et al. (2016) Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. ArXiv abs/1609.08144.
23. Margrave GF, Lawton DC, Stewart RR, Miller S, Yang G, et al. (1997) The Blackfoot 3C-3D seismic survey: A case study. CREWES Research Report.
24. Table of Formations of Alberta (2019) AGAT Laboratories.
25. Stratigraphic Correlation Chart (2019) Core Laboratories, Calgary.
26. Adams LM (2017) New plays in an old field: depositional history and source rock characterization at Teapot Dome, Wyoming. MS thesis at Louisiana State University at Lafayette.
27. Okojie-Ayoro AO (2007) An Approach to Mapping of Shallow Petroleum Reservoirs Using Integrated Conventional 3D and Shallow P- and SH-Wave Seismic Reflection Methods at Teapot Dome Field in Casper, Wyoming. Theses and dissertations.
28. George L, Cardinal L, Winter G (2017) 2014 Wyoming Stratigraphic Nomenclature Chart. Search and Discovery Article.
29. Baranova VV, Mustaqeem A, Teitel A, Kim JJ (2021) Data analytics and application of automated sequence stratigraphic approach to Montney deposition in Fort St. John graben area. GeoConvention.

This article was originally published in a special issue:

Sedimentary Environments and Facies

Handled by Editor(s):

Prof. Damir Bucković
Department of Geology
University of Zagreb
Croatia