

Topical Report

Predicting Porosity and Permeability for the Canyon Formation, SACROC Unit (Kelly-Snyder Field), Using the Geologic Analysis via Maximum Likelihood System

September, 2007

Performed by:

Reinaldo Gonzalez and
Scott R. Reeves

Advanced Resources International, Inc.
11490 Westheimer Rd., Suite 520
Houston, TX 77077

In conjunction with:

Eric Eslinger

Eric Geoscience, Inc.
10 Sussex Road
Glenmont, NY 12077

Prepared for:

U.S. Department of Energy
Contract No. DE-FC26-04NT15514



**Advanced Resources
International, Inc.**

Disclaimers

U.S. Department of Energy

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Advanced Resources International, Inc.

The material in this Report is intended for general information only. Any use of this material in relation to any specific application should be based on independent examination and verification of its unrestricted applicability for such use and on a determination of suitability for the application by professionally qualified personnel. No license under any Advanced Resources International, Inc., patents or other proprietary interest is implied by the publication of this Report. Those making use of or relying upon the material assume all risks and liability arising from such use or reliance.

Executive Summary

Accurate, high-resolution, three-dimensional (3D) reservoir characterization can provide substantial benefits for effective oilfield management. By doing so, the predictive reliability of reservoir flow models, which are routinely used as the basis for significant investment decisions designed to recover millions of barrels of oil, can be substantially improved. This is particularly true when Secondary Oil Recovery (SOR) or Enhanced Oil Recovery (EOR) operations are planned. If injectants such as water, hydrocarbon gasses, steam, CO₂, etc. are to be used; an understanding of fluid migration paths can mean the difference between economic success and failure.

SOR/EOR projects will increasingly take place in heterogeneous reservoirs where interwell complexity is high and difficult to understand. Although reasonable reservoir characterization information often exists at the wellbore, the only economical way to sample the interwell region is with seismic methods which makes today's standard practice for developing a 3D reservoir description to resort to the use of seismic inversion techniques. However, the application of these methods brings other technical drawbacks than can render them inefficient. The industry therefore needs improved reservoir characterization approaches that are quicker, more accurate, and less expensive than today's standard methods.

To achieve this objective, the Department of Energy (DOE) has been promoting some studies with the goal of evaluating whether robust relationships between data at vastly different scales of measurement could be established using advanced pattern recognition (soft computing) methods. Advanced Resources International (ARI) has performed two of these projects with encouraging results showing the feasibility of establishing critical relationships between data at different measurement scales to create high-resolution reservoir characterization. In this third study performed by ARI and also funded by the DOE, a model-based, probabilistic clustering analysis procedure is successfully applied to generate a high-resolution reservoir characterization outcome. The approach was applied in the Pennsylvanian-Permian reef carbonates (Cisco and Canyon Formations) of a subregion of the SACROC Unit, Horseshoe Atoll, Permian Basin, Texas, and acknowledged as a highly complex carbonate reservoir.

Due to the modest results achieved with the application of soft-computing methodologies to the available information (no crosswell data at hand), the original project target about creating a data-driven device relating surface seismic information, crosswell seismic attributes, geophysical logs and core parameters for the prediction of core-scale porosity and permeability profiles in locations where only 3D surface seismic data was available, had to be reformulated. It was shown that 3D seismic information was not capable of capturing the degree of vertical variability of SACROC. As a consequence, available seismic information was unincorporated from posterior reservoir characterization tasks, and a combination of data-driven procedures and geostatistical methods was utilized for reservoir characterization purposes.

A selected area within the SACROC Unit platform was used for this study. The suitable logs for the creation of an "intelligent" log-to-core device were not present for all wells. These logs were gamma ray (GR), neutron porosity (NPHI), bulk density (RHOB), and delta time (DT).

It was necessary to create a first “intelligent” tool, a log-to-log model to provide synthetic logs of RHOB and DT (or eventually of acoustic impedance derived from them) at well locations where only GR and NPHI were available (the most common situation in this reservoir). Once the “ideal” logs were completed, a second model, a log-to-core device, provides core scale estimates of porosity and permeability (P&P). The validity of these soft-computing devices was checked using “holdout” wells. In this way, “core” parameter profiles, with high vertical resolution, could be generated for many wells. This procedure permits to populate any well location with core-scale estimates of P&P and rock types facilitating the application of geostatistical characterization methods.

The first step procedure was to discriminate rock types of similar depositional environment and/or reservoir quality (RQ) using a specific clustering technique. The approach implemented utilized a model-based, probabilistic clustering analysis procedure called GAMLS^{1,2,3,4} (Geologic Analysis via Maximum Likelihood System) which is based on maximum likelihood principles. During clustering, samples (data at each digitized depth from each well) are probabilistically assigned to a previously specified number of clusters with a fractional probability that varies between zero and one. This permits individual samples to “belong” to more than one “rock type”, and so allows for gradational, or intermediate, rock types. In other words, a given sample might have characteristics of more than one rock type (or reservoir quality unit) bringing a strong component of geologic reality into the process that contrasts with other mathematical methodologies of data-driven applied for reservoir characterization. The developed “taxonomy” is used as a framework for ensuing calculation of reservoir parameter values.

The clusters were qualitatively related to RQ using data output tables, crossplots, and frequency plots. Also, cross sections were generated which permitted a visual and qualitative assessment of lateral “bed” continuity and vertical bed thickness and style. Clustering analyses indicated that the SACROC carbonate section can be divided into a suite of closely-related flow units that have a “good” RQ (average porosity ~ 11-13 %) and into a suite of closely-related flow units that have a “poor” RQ (average porosity generally < 5 %). As interpreted from clustering analysis output, the contacts between these good and poor suites is generally rather sharp, as opposed to the generally gradational contacts that exist among the several flow units that comprise the good and poor suites. The relatively “sharp” contacts are interpreted to represent 3rd to 4th order sequence boundaries and, practically, they would likely act as significant barriers to vertical fluid flow.

This integrated use of pattern-recognition and stratigraphic classification provided a meaningful technical advancement over conventional methodologies, specifically by giving the necessary vertical resolution for 3D reservoir characterization tasks, by reducing uncertainty, by facilitating results directly in the engineering terms needed for effective reservoir management, and by simplifying and streamlining the process, making it more time (and cost) efficient. Once generated P&P estimates for all wells with no core data, geostatistical methods can directly be used to build a 3D reservoir model.

Table of Contents

Disclaimers.....	ii
Executive Summary	iii
List of Tables	vii
List of Figures.....	viii
1.0 Introduction.....	1
2.0 Area of Study	4
3.0 Dataset.....	8
3.1 <i>Well Log Data</i>	8
3.2 <i>Core Data</i>	9
3.3 <i>Seismic Information</i>	10
3.4 <i>Crosswell Data</i>	12
4.0 Methods.....	15
4.1 <i>The GAMLS Clustering Engine</i>	15
4.2 <i>The GAMLS Workflow</i>	16
5.0 Discrimination and Interpretation of Facies Units.....	19
5.1 <i>Interpretation of Lithology of Facies Units</i>	19
5.2 <i>Correlation of Facies Units Among Wells</i>	19
5.3 <i>Estimation of Missing Data</i>	19
5.4 <i>Lithofacies Units versus Electrofacies Units versus Flow Units</i>	20
6.0 Reservoir Quality of Facies and Flow Units.....	21
6.1 <i>Background</i>	21
6.2 <i>Flow Unit Realizations and Reservoir Quality</i>	21
6.3 <i>Estimation of Porosity and Permeability in Non-Cored Wells</i>	22
6.4 <i>Porosity and Permeability via the “AI” Method</i>	23
7.0 Realizations and Their Results	30
7.1 <i>CIA - Clustering Using a 4-Curve Log Suite and Application of “P&P via AI” to this Realization</i>	30
7.1.1 <i>Cluster Setup Procedures</i>	30
7.1.2 <i>Interpretation of Modes (Assignment of Lithologies)</i>	31

7.1.3	<i>Properties (RQ) as Functions of Modes</i>	33
7.1.4	<i>Modes, Beds, and Flow Units</i>	37
7.1.5	<i>Porosity and Permeability of Non-Cored Wells Assigned to Modes from CIA</i>	46
7.1.6	<i>Variation of K0, K90 and Kv with Flow Units</i>	48
7.2	<i>C1B - Clustering with a 4-Curve Log Suite Realization Using 12 Wells with RHOB and DT Curves and 19 Wells without RHOB and DT</i>	54
7.3	<i>C1C - Realization from Clustering of 12 Wells Using 4 Log Curves and Core Porosity and Core Permeability from one Cored Well</i>	58
7.4	<i>C1D - Clustering Run Same as C1C Using 25 Modes</i>	70
7.5	<i>C1E - Clustering Run C1C Using Core Data Well 37-11 & 11-15</i>	71
7.6	<i>Analysis of Oriented Cross Sections</i>	74
8.0	Conclusions	82
9.0	References	84
10.0	List of Acronyms and Abbreviations	86
Appendix A: Exploratory Statistical Analysis		A-1
Appendix B: Seismic Attributes Computed from Well Log and Seismic Data, SACROC Oilfield, Texas		B-1
Appendix C: The Use of Seismic Information		C-1
Appendix D: GAMLS Methodology		D-1

List of Tables

Table 1: Clustering Procedures for the P&P Via AI Method for Predicting Core-Scale P&P Estimates	29
Table 2: Summary of Clustering Runs Utilized to Analyze Methodology Robustness	30
Table A-1: Statistical Parameters of the Core Permeability and Porosity Measurements	A-6
Table A-2: Statistical Parameters of the Core Permeability and Porosity Measurements (Well 11-15).....	A-6
Table A-3: Statistical Parameters of the Core Permeability and Porosity Measurements (Well 19-12).....	A-7
Table A-4: Statistical Parameters of the Core Permeability and Porosity Measurements (Well 37-11).....	A-7
Table A-5: $\log_{10}(K_0)$ vs. Porosity: Correlation Coefficients for Different Window-Averaged Values	A-12
Table A-6: Correlation Coefficients, Mean and Median	A-13
Table A-6: Statistical Parameters of the Well Logs for Wells 11-15, 19-12, and 37-11	A-15
Table A-7: Correlation Matrix of Well Log Parameters for Wells 11-15, 19-12, and 37-11	A-16
Table A-8: Variance Percentage by Factor for Wells 11-15, 19-12, and 37-11	A-17
Table A-9: Variable Contributions by Factor for Wells 11-15, 19-12, and 37-11	A-17
Table B-1: Well Locations & Nearest Seismic Trace in Local SACROC Coordinate Frame	B-7

List of Figures

Figure 1: Pathway to 3D High-Resolution Reservoir Description.....	2
Figure 2: Illustration of Different Scales of Measurement	2
Figure 3: Location of the SACROC Unit, Permian Basin (Modified from Raines, 2001)	4
Figure 4: Stratigraphic Column of the Permian Basin.....	5
Figure 5: General Lithologic Setting of SACROC Unit (Modified from Raines, 2001)	6
Figure 6: Three Areas of SACROC Unit (Modified from Raines, 2001).....	7
Figure 7: Well Locations at the Northern Platform of the SACROC Unit	9
Figure 8: Core Permeability (Log_{10}) vs. Core Porosity	11
Figure 9: Schematic of the Two-Step “Soft-Computing” Procedure.....	14
Figure 10: C1AA vs. C9, Left-Porosity, Right-Permeability (Log_{10}).....	23
Figure 11: Predicted Porosity and Permeability Tracks for Wells 33-15, 11-15, and 19-12	24
Figure 12: Porosity and Permeability (Log_{10}) Tracks for Wells 11-15 and 19-12.....	25
Figure 13: AI_log (1st track), Porosity (2nd track) and $\text{Log}_{10}(\text{K0})$ (3rd track) for Testing Hold-Out Well 11-15 Actual Parameter Values in Red; Predicted Values in Blue	28
Figure 14: Predicted vs. Actual Values for AI_Log, Porosity, and $\text{Log}_{10}(\text{K0})$ Ranges: AI_Log [20, 60]; Porosity [0.0, 0.3]; $\text{Log}_{10}(\text{K0})$ [-2.75, 2.75]	28
Figure 15: Screen View of the GAMLS Clustering Analysis “Setup” for Cluster C1A.....	31
Figure 16: Tracks for Wells 11-15 and 37-11.....	32
Figure 17: Cluster Run C1A – GAMLS “ModeAssign” Screen: Automatic Lithology Assignment Routine.....	33
Figure 18: Cluster Run C1A: RHOB vs. NPHI Crossplot for Well 37-11	34
Figure 19: 3D Plot (RHOB, NPHI, GR) of All Wells Included in C1A.....	34
Figure 20: Cluster Run C1A: Mean and Standard Deviation for Selected Variables	35
Figure 21: Cluster C1A - Well 37-11: Core Permeability (Log_{10}) vs. Core Porosity	36
Figure 22: Cluster C1A: Probabilistic Representation of Modes at Each Depth is Displayed	36
Figure 23: Cluster C1A: Probabilistic Representation of Modes at Each Depth	37
Figure 24: Tracks for the Non Cored Well 33-15.....	38
Figure 25: Cluster C1A - Original 10 Mode Results vs. 4 Mode Results.....	39
Figure 26: Cluster C1A- 12 Wells with 4 Modes after Dropping Application	40
Figure 27: Cluster C1A - 12 Wells with “Crisp” Modes at each Depth	41
Figure 28: Cluster C1A -12 Wells Using BTF with a Minimum Bed Thickness = 5 Ft.....	42
Figure 29: Cluster C1A -12 Wells Using BTF with a Minimum Bed Thickness = 10 Ft.....	43
Figure 30: Cluster C1A -12 Wells Using BTF with a Minimum Bed Thickness = 20 Ft.....	44
Figure 31: Cluster C1A - Bed Information for each of the 7 Limestone Beds	45
Figure 32: Cluster C1A - Bed Information, Minimum Bed Thickness = 10 Ft	46
Figure 33: Well 33-15: Mean & Standard Deviation of Predicted Porosity and Predicted Permeability (Log_{10}).....	47

Figure 34: Well 33-15 Predicted Permeability (Log10) vs. Predicted Porosity for each Mode	47
Figure 35: $\text{Log}_{10}(\text{K90})$ vs. $\text{Log}_{10}(\text{K0})$ for Well 37-11.....	48
Figure 36: $\text{Log}_{10}(\text{K90})$ vs. $\text{Log}_{10}(\text{K0})$ for Well 11-15.....	49
Figure 37: $\text{Log}_{10}(\text{K90})$ vs. $\text{Log}_{10}(\text{K0})$ for Well 19-12.....	49
Figure 38: $\text{Log}_{10}(\text{K0})$ vs. $\text{Log}_{10}(\text{K0/K90})$ for Well 11-15.....	50
Figure 39: Kv vs. K0 (Logarithm) for Well 37-11	51
Figure 40: Kv vs. K0 (logarithm) for Well 11-15.....	51
Figure 41: Kv vs. K0 (logarithm) for Well 19-12.....	52
Figure 42: Well 37-11: Frequency Distribution by Mode (Cluster C1A) of $\text{Log}_{10}(\text{Kv/Kh})$	53
Figure 43: $\text{Log}_{10}(\text{Kh})$ vs. $\text{Log}_{10}(\text{Kv/Kh})$ Cored Well 37-11.....	54
Figure 44: Cluster C1B: Probabilistic Representation of Modes Displayed at Each Depth	56
Figure 45: Comparison of Realizations from C1A (Top) and C1B (Bottom)	57
Figure 46: Cluster C1B, Predicted Permeability (Log_{10}) vs. Predicted Porosity for each Facies, Well 33-11	58
Figure 47: Cluster C1C: Porosity Values Distribution for 12 Wells Discriminated by Modes	59
Figure 48: Cluster C1C: Permeability (Log_{10}) Values Distribution for all 12 Wells Discriminated by Modes.....	60
Figure 49: Tracks of Clustering Results for the Non-Cored Well 33-15	61
Figure 50: C1C Predicted POR vs. C9 Predicted POR, Well 33-15.....	62
Figure 51: C1C Predicted PERM vs. C9 Predicted PERM (Log_{10}), Well 33-15	62
Figure 52: C1C Predicted POR vs. Actual Core Porosity, Well 11-15 as a Hold-Out Testing Well.....	63
Figure 53: C1C Predicted POR vs. Actual Core Porosity, Well 11-15 as a Hold-Out Testing Well.....	64
Figure 54: C1C Predicted POR vs. Actual Core Porosity, Well 19-12 as a Hold-Out Testing Well.....	64
Figure 55: C1C Predicted POR vs. Actual Core Porosity, Well 19-12 as a Hold-Out Testing Well.....	65
Figure 56: C1C Predicted PERM (Log_{10}) vs. Actual Core Permeability, Well 11-15 as a Hold-Out Testing Well....	66
Figure 57: C1C Predicted PERM (Log_{10}) vs. Actual Core Permeability, Well 19-12 as a Hold-Out Testing Well....	66
Figure 58: Frequency Distributions for Core Porosity and Predicted Porosity (Using Cluster C1C), Well 11-15.....	67
Figure 59: Frequency Distributions for Core Porosity and Predicted Porosity (Using Cluster C1C), Well 19-12.....	67
Figure 60: Frequency Distributions for Core Permeability and Predicted Permeability (Using Cluster C1C), Well 11-15.....	68
Figure 61: Frequency Distributions for Core Permeability and Predicted Permeability (Using Cluster C1C), Well 19-12.....	68
Figure 62: $\text{Log}_{10}(\text{K0})$ vs. $\text{Log}_{10}(\text{K0}) - \text{Log}_{10}(\text{K90})$, Well 11-15	69
Figure 63: $\text{Log}_{10}(\text{C1D_K0_EST})$ vs. $\text{Log}_{10}(\text{C1C_K0_EST})$, Well 19-12	70
Figure 64: $\text{Log}_{10}(\text{C1E_K0_EST})$ vs. $\text{Log}_{10}(\text{C1C_K0_EST})$, Well 19-12.....	71
Figure 65: $\text{Log}_{10}(\text{C1E_K0_EST})$ Versus $\text{Log}_{10}(\text{K0})$ (Actual Data): Well 19-12.....	72
Figure 66: Predicted $\text{Log}_{10}(\text{K0})$ vs Actual $\text{Log}_{10}(\text{K0})$, Well 19-12.....	73
Figure 67: Comparison of Realizations from Five Clustering Runs.....	74
Figure 68: Location Map Cross Sections A-A', B-B', and C-C'	75

Figure 69: Cross Section A-A'	76
Figure 70: Cross Section A-A' with Positions of Zones A, B, and C	76
Figure 71: Cross Section B-B' with Positions of Zones A, B, and C	77
Figure 72: Cross Section C-C' with Positions of Zones A, B, and C	78
Figure 73: C1A Wells with Order Based on an Apparent Structural Trend of a Prominent Limestone "Bed" Near the Bottom of the Section.....	79
Figure 74: C1A Wells Oriented According to Relative Structural Position Using a Prominent Low RQ "Bed" Near the Bottom of the Studied Interval	81
Figure A-1: Crossplots of $\text{Log}_{10}(\text{K}_0)$ vs. $\text{Log}_{10}(\text{K}_{90})$ - Wells 11-15, 19-12, and 37-11	A-2
Figure A-2: Permeability Histograms and Corresponding Probability Plots - Well 11-15.....	A-4
Figure A-3: Permeability Histograms and Corresponding Probability Plots - Well 19-12.....	A-4
Figure A-4: Permeability Histograms and Corresponding Probability Plots - Well 37-11.....	A-5
Figure A-5: Permeability Histograms and Corresponding Probability Plots - All Wells	A-5
Figure A-6: Crossplots $\text{Log}_{10}(\text{K}_0)$ vs. $\text{Log}_{10}(\text{K}_v)$ at All Cored Wells	A-9
Figure A-7: Crossplots $\text{Log}_{10}(\text{K}_{90})$ vs. $\text{Log}_{10}(\text{K}_v)$ at all Cored Wells	A-10
Figure A-8: Porosity Histograms at all Cored Wells	A-11
Figure A-9: Superimposed Crossplots of $\text{Log}_{10}(\text{K}_0)$ vs. Porosity	A-12
Figure A-10: $\text{Log}_{10}(\text{K}_0)$ vs. Porosity for Wells 11-15 (Red), 19-12 (Blue), and 37-11 (Green)	A-13
Figure A-11: $\text{Log}_{10}(\text{K}_0)$ vs. Porosity for Wells 11-15 (Red), 19-12 (Blue), and 37-11 (Green)	A-14
Window Median Filtered 10' (Left), 40' (Center), and 100' (Right)	A-14
Figure A-12: $\text{Log}_{10}(\text{K}_0)$ vs. Porosity for Wells 11-15 (Red), 19-12 (Blue), and 37-11 (Green)	A-14
Figure A-13: Bar Plot- Percentage of Variance Contribution of Each Factor	A-17
Figure B-1: Well 11-15, Log and Synthetic Trace Attributes Unfiltered	B-9
Figure B-2: Well 11-15, Log and Synthetic Trace Attributes Unfiltered in the Depth Domain with a 2-10 Hz Band Pass Filter	B-10
Figure B-3: Well 11-15, Log and Real Seismic Trace Attributes.....	B-11
Figure C-1: Core Porosity Depth-Plots on Wells 37-11, 11-15 and 19-12	C-2
Figure C-2: Seismic Acoustic Impedance (AI) Traces Overlain onto the Core Porosity Data for Cored Wells, AI = - 1000 left to 1000 right.....	C-3
Figure C-3: AI vs. Core Porosity for Cored Wells 37-11 and 19-12	C-4
Figure C-4: AI vs. Predicted Core Porosity for Nine Non-cored Wells	C-5
Figure C-5: Core Permeability vs. Core Porosity for the Three Cored Wells	C-6
Figure C-6: Core Permeability (Logarithm) vs. Core Porosity for the Three Cored Wells	C-7
Figure C-7: Rock Types, Some Seismic Attributes and AI_Log.....	C-8
Figure C-8: AI Profiles from Seismic (AI_Seis); Log Derived (AI_Log); Median Filtered 50' (AI_50'); Median Filtered 100' (AI_100'); and Median Filtered 150' (AI_150')	C-9

1.0 Introduction

Accurate, high-resolution, three-dimensional (3D) reservoir characterization can provide substantial benefits for effective oilfield management. By doing so, the predictive reliability of reservoir flow models, which are routinely used as the basis for significant investment decisions designed to recover millions of barrels of oil, can be substantially improved. Even a small improvement in incremental oil recovery for high-value assets can result in important contributions to bottom-line profitability.

This is particularly true when Secondary Oil Recovery (SOR) or Enhanced Oil Recovery (EOR) operations are planned. If injectants such as water, hydrocarbon gasses, steam, CO₂, etc. are to be used, an understanding of fluid migration paths can mean the difference between economic success or failure. In these types of projects, injectant costs can be a significant part of operating expenses, and hence their optimized utility is critical.

SOR/EOR projects will increasingly take place in heterogeneous reservoirs where interwell complexity is high and difficult to understand. Although reasonable reservoir characterization information often exists at the wellbore, the only economical way to sample the interwell region is with seismic methods. Surface reflection seismic has relatively low cost per unit volume of reservoir investigated, but the resolution of surface seismic data available today, particularly in the vertical dimension, is not sufficient to produce the kind of detailed reservoir description necessary for effective SOR/EOR optimization and planning.

Today's standard practice for developing a 3D reservoir description is to use seismic inversion techniques. These techniques make use of rock physics concepts to solve the inverse problem, i.e., to iteratively construct a likely geologic model and then upscale and compare its acoustic response to that actually observed in the field. This method suffers from the fact that rock physics relationships are not well understood, and the need to rely on porosity-permeability transforms to estimate permeability from porosity. Further, these methods require considerable resources to perform, and are applied to only a small percentage of oil and gas producing assets.

Since the majority of fields do not utilize these techniques (today), many fields are sub-optimally developed. The industry therefore needs an improved reservoir characterization approach that is quicker, more accurate, and less expensive than today's standard methods. This will permit more reservoirs to be better characterized, allowing recoveries to be optimized and significantly adding to recoverable reserves.

A new approach to achieve this objective was first examined in a Department of Energy (DOE) study performed by Advanced Resources International (ARI) in 2000/2001¹. The goal of that study was to evaluate whether robust relationships between data at vastly different scales of measurement could be established using virtual intelligence (VI) methods. The proposed workflow required that three specific relationships be established through use of data-driven modeling methods, in that case Artificial Neural Networks (ANN's): core-to-log, log-to-crosswell seismic, and crosswell-to-surface seismic as shown in Figure 1. A key attribute of this

approach, is the inclusion of borehole seismic (such as crosswell and/or vertical seismic profiling – VSP) in the data collection scheme. As shown in Figure 2, borehole seismic fills a critical gap in the resolution spectrum of reservoir measurements between the well log and surface seismic scales, thus establishing important constraints on characterization outcomes.

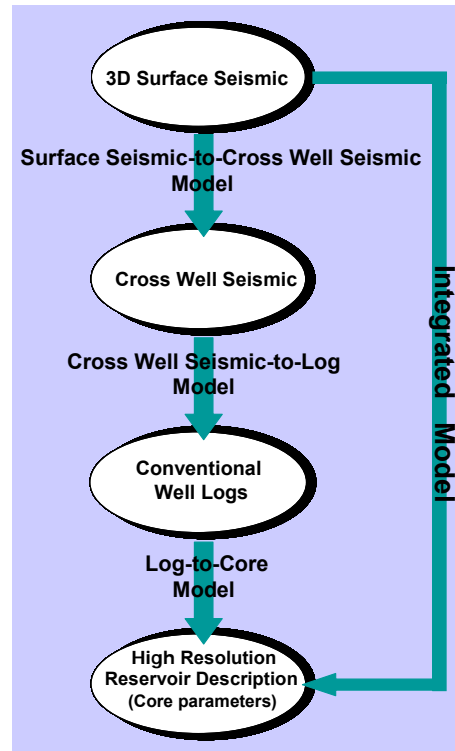


Figure 1: Pathway to 3D High-Resolution Reservoir Description

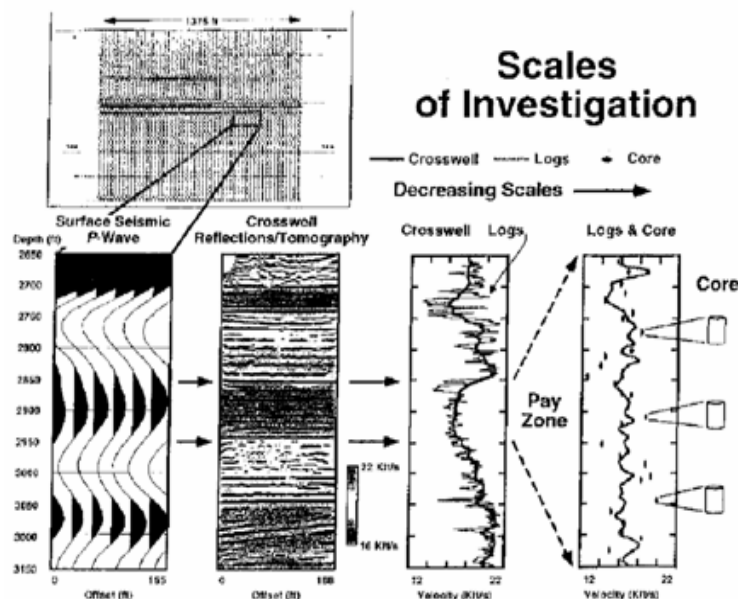


Figure 2: Illustration of Different Scales of Measurement

The results of the initial study showed that it is, in fact, feasible to establish the three critical relationships required, and that use of data at different scales of measurement to create high-resolution reservoir characterization is possible. Based on the results of the study, in September 2001, the DOE, again through ARI, launched a subsequent two-year government-industry R&D project to further develop and demonstrate the technology². The primary goals of the second project were to make improvements to the initial methodology by incorporating additional virtual intelligence (VI) technologies (such as clustering), using core measurements in place of magnetic resonance image (MRI) logs, and streamlining the workflow. The project was performed at the McElroy field on the Central Basin Platform of the Permian basin. The results of the study indicated that a reasonable reservoir characterization could be created using clustering methods. The model provided results that appeared consistent with known reservoir properties of the field, and identified potential areas of poor reservoir quality to avoid in future development.

The clustering approach was shown to have an advantage over ANN methods since the entire process could be performed with a single, integrated model as opposed to multiple, sequential models. However, experimentation with and without crosswell data suggested that, in that particular study, crosswell data actually deteriorated model performance. It is believed that the crosswell data was of poor quality, which may have introduced error into the process, creating the result.

The second study showed sufficient promise in the utility of soft-computing methods for reservoir characterization so that further refinement of the process was undertaken. In this third study, also performed by ARI and funded by the DOE, the clustering approach was again utilized to generate a high-resolution reservoir characterization outcome, and is the subject of this report. The project was performed at the SACROC Unit, operated by Kinder Morgan CO₂ Company (KMCO₂), in the Permian basin of West Texas.

2.0 Area of Study

The SACROC Unit includes most of the Kelly-Snyder field and some of the Diamond “M” field in Scurry County, Texas. It is a part of the Horseshoe Atoll located in the eastern half of the Midland basin which is the eastern sub-basin of the overall Permian basin of western Texas and southeastern New Mexico as shown in Figure 3. The 2.8 Billion Stock Tank Barrel (Original Oil in Place) Unit produces from the Pennsylvanian-aged Cisco and Canyon formations with productive carbonates actually belonging to both formations^{3,4} as depicted in Figure 4. The reservoir is typically called the “Canyon Reef” as a general description³. As is the case with many such Pennsylvanian reef complexes, the SACROC Unit exhibits a great deal of vertical relief and laterally complicated geometries.

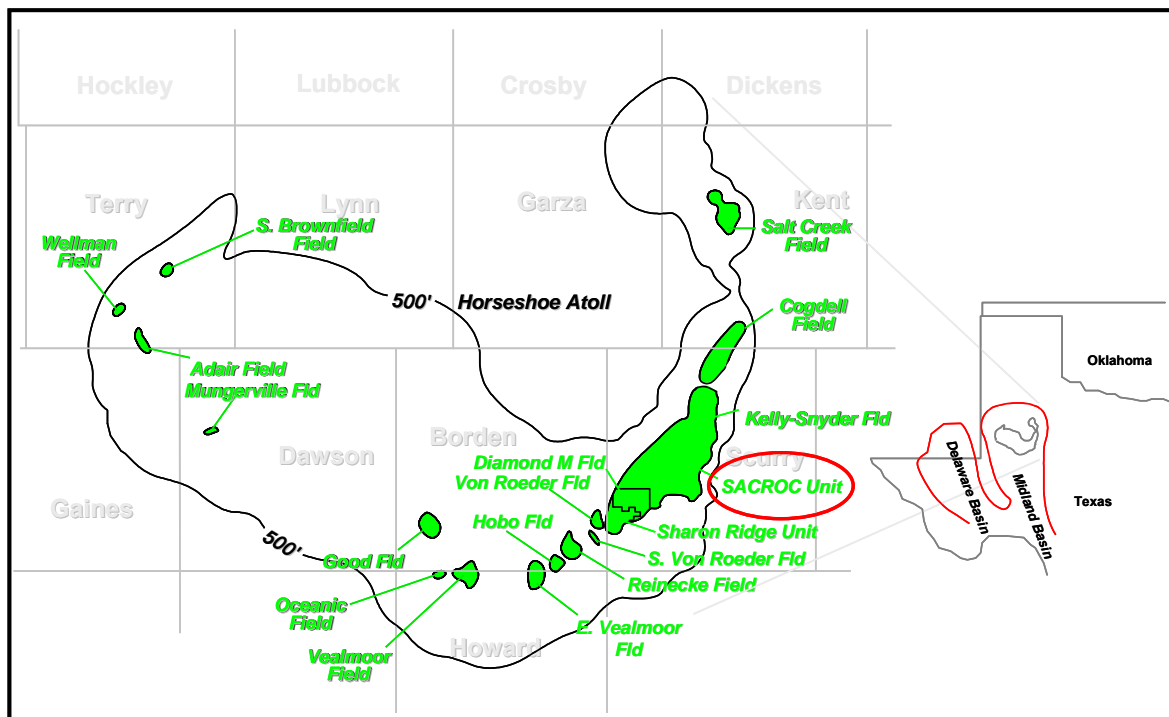


Figure 3: Location of the SACROC Unit, Permian Basin (Modified from Raines, 2001)

The productive interval of the Canyon Reef is mainly composed of limestone, although minor amounts of anhydrite, chert, sand and shale can be locally found. Dolostone is rare to nonexistent. The Wolfcamp shale is a natural seal above the carbonate and around the flanks of the SACROC Unit. Gross stratigraphy across the unit is shown in Figure 5. Toward the east and west boundaries, the Cisco-Canyon productive carbonate interval thins and drops below the regional oil-water contact. Carbonate accumulations present extremely complex geometries and steep sides, and seem to frequently commence on antecedent highs in one or more underlying zones.

Reservoir heterogeneity at SACROC has been attributed to tectonics, icehouse conditions, and post-burial geochemical processes^{3,5}. During Pennsylvanian time, tectonic activity strongly influenced growth initiation points and imposed stress conditions. Rapid sea level fluctuations which put carbonate organisms in a “preservation” mode, resulted in high vertical carbonate accumulations and a large variety of depositional environments and facies^{3,4,5}. This resulted in thick, steep-sided carbonate intervals with a good chance for erosion and diagenesis³. Subsequent geochemical processes modified the reservoir through karst overprints, dissolution, and precipitation.

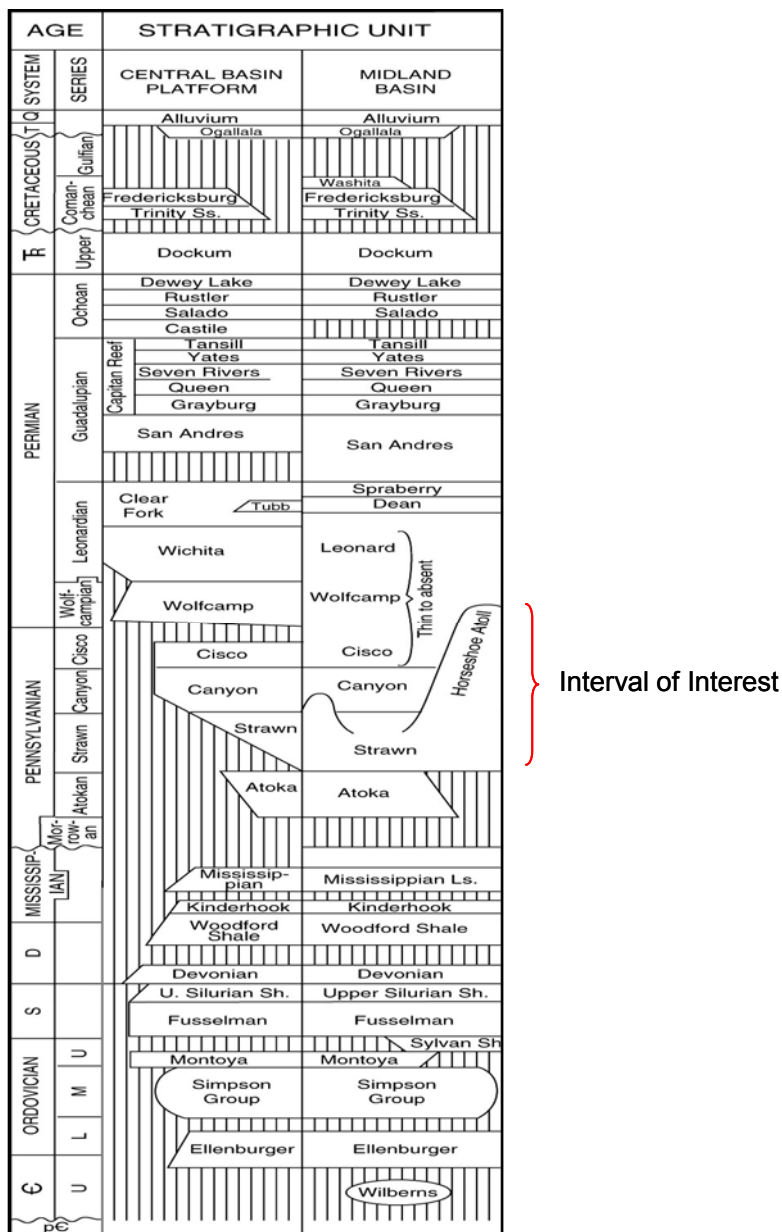


Figure 4: Stratigraphic Column of the Permian Basin.

These depositional complexities accompanied by later modifications produced a region where sudden and abrupt changes can be found within any zone (even in relatively flat areas), and where the relationship between flow units is hard to determine. Raines et al.³ shows this complexity by illustrating how two closely spaced wells exhibit substantial differences in the vertical distribution of porosity

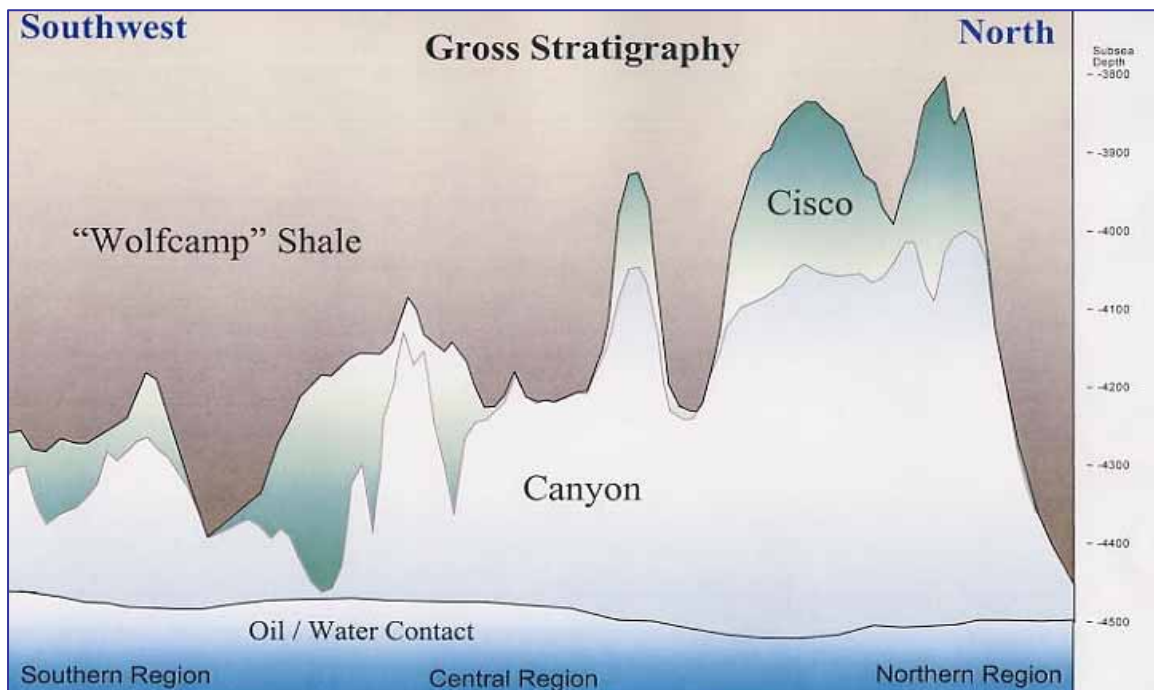


Figure 5: General Lithologic Setting of SACROC Unit (Modified from Raines, 2001)

Development operations commenced in the 1940's, thus production operations have impacted reservoir conditions for many years, further complicating the current reservoir setting. At present, SACROC is the focus of tertiary CO₂ recovery operations. Because of the delicate nature of tertiary recovery operations from the mechanical, physical, and economic viewpoint, it is vital to gain a better understanding of current reservoir fluid flow mechanics. This effort is currently under way in the area. While tertiary CO₂ flooding efforts are present in the central portion of the field, simultaneous reservoir characterization research is focused in the more complicated portions of the field.

SACROC has been divided into three broad geographic areas as seen in Figure 6. This study took place in a 0.5 mi² test site within the northern third of the unit which is frequently called the North Platform. According to Michael Raines⁵, the Canyon is the most productive formation and the Cisco is generally tight. However, a zone, called "The Green Zone", is the exception to the rule. "The Green Zone", which is about 1/2 Cisco and 1/2 Canyon, is the most productive interval/area in the SACROC Unit. Thickness of the Canyon is approximately 750' with several thick, laterally continuous zones, especially toward the oil water contact at the base of the Canyon interval.

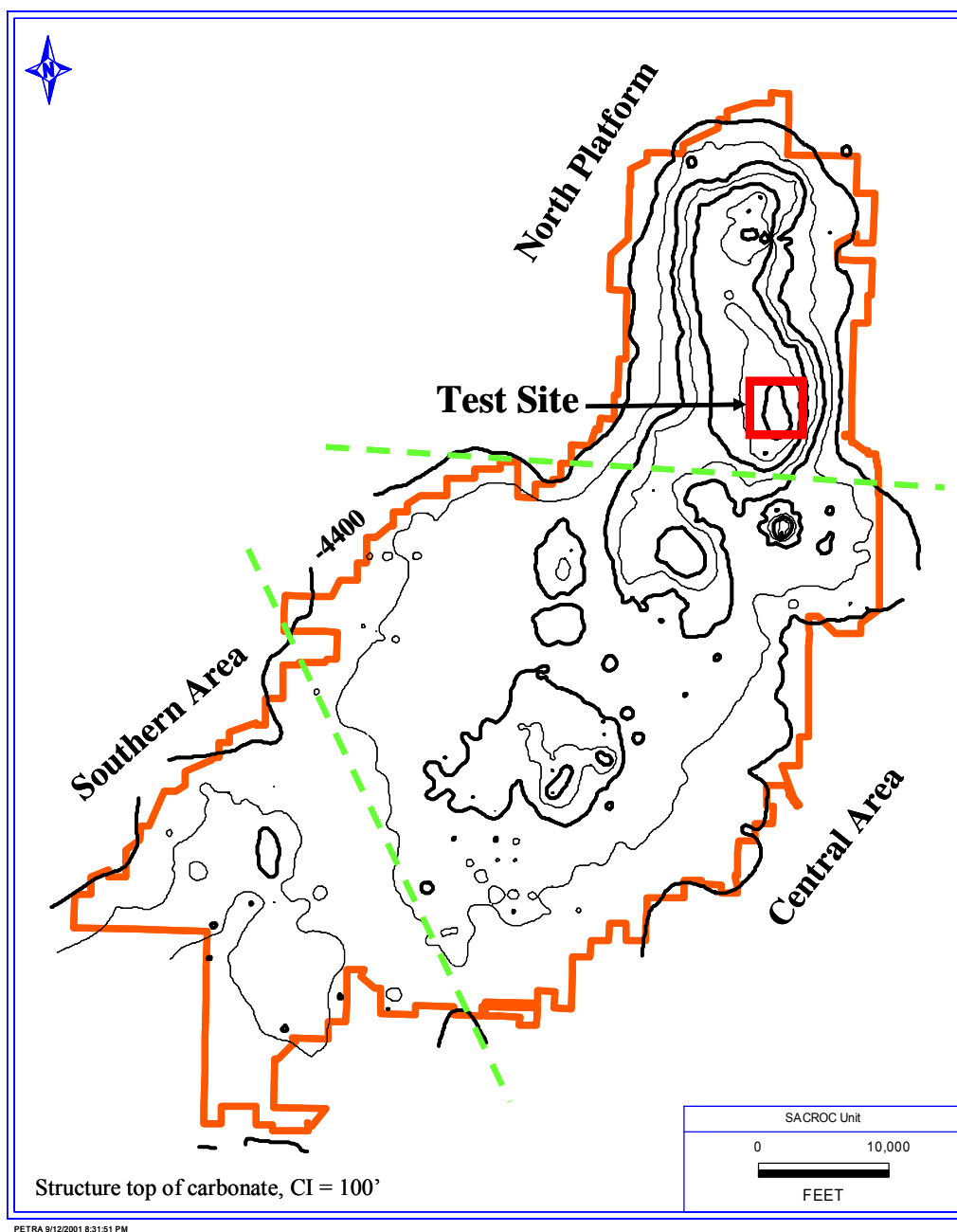


Figure 6: Three Areas of SACROC Unit (Modified from Raines, 2001)

3.0 Dataset

3.1 Well Log Data

More than four hundred (400) wells from the Northern Platform were available for study. A variety of geophysical well logs, common and uncommon, are present in these wells. Most of wells in SACROC Unit have gamma ray (GR), and neutron porosity (NPHI), however few have bulk density (RHOB) or delta time (DT) logs. Log response issues at SACROC are complex and deserve to be mentioned. Raines et al.³ reported that despite the Canyon Reef interval being composed primarily of limestone, false GR signatures (often in excess) are not uncommon, and are caused by the presence of uranium in the formation. With relative frequency, this response is seen in low porosity intervals, in the lowest portions of the Canyon series, and on the flanks of the carbonate buildup. Casing effects on the total or apparent GR as well as on the corrected GR, are also present and usually correctable. Modern NPHI log data is generally rather reliable, although may be influenced by the presence or lack of casing over a portion of the log, and NPHI logs do not always properly characterize the flow capacity of a well.

Nevertheless, it is believed that Wolfcamp shale (above the carbonate interval) may be used to normalize bad GR and NPHI log data, when digital data has been recorded over a sufficient portion of the shale interval. This feature is unique to the Horseshoe Atoll and has been helpful in maximizing the reliability of SACROC log data. More details about these log issues can be found in Raines et al.³.

Multivariable statistical methods were applied in order to select the most applicable log parameters for characterization tasks. These are discussed in detail in Appendix A. The availability of such logs at each well and their functionality for seismic calibration were important determining factors for the final selection. Taking into consideration statistical arguments, aspects linked to the capabilities of log tools, and data availability, the well log parameters judged suitable for characterization tasks were GR, NPHI, RHOB, and DT. Resistivity and photoelectric factor showed a good potentiality of being used for characterization tasks under the adopted procedures; however, their scarce presence at SACROC precluded them for use for the clustering analysis performance.

Because of log data complexity, only twelve (12) wells having an "ideal" log suite consisting of RHOB, NPHI, GR, and DT were selected for evaluation in the study area SACROC. The location of these wells is shown in Figure 7. In order to reduce the number of total variables in the clustering runs and in some specific stages of the entire characterization process, RHOB and DT were utilized to generate a well log Acoustic Impedance (hereafter denoted by AI_log) as a combined form of both logs. AI_log is computed by the formula

$$AI_log = 100 (RHOB/DT)$$

3.2 Core Data

While core data is available for 29 wells in the Northern Platform area, most are from older wells of which the quality of the data was in question. This study used core data from three wells cored between 2004 and 2005 by KMCO₂, which have porosity and permeability data from whole core covering nearly the entire ~800' Canyon Reef interval. The location of these wells is shown in Figure 7. Only one of these wells (37-11) was near the center of the study area (Figure 7).

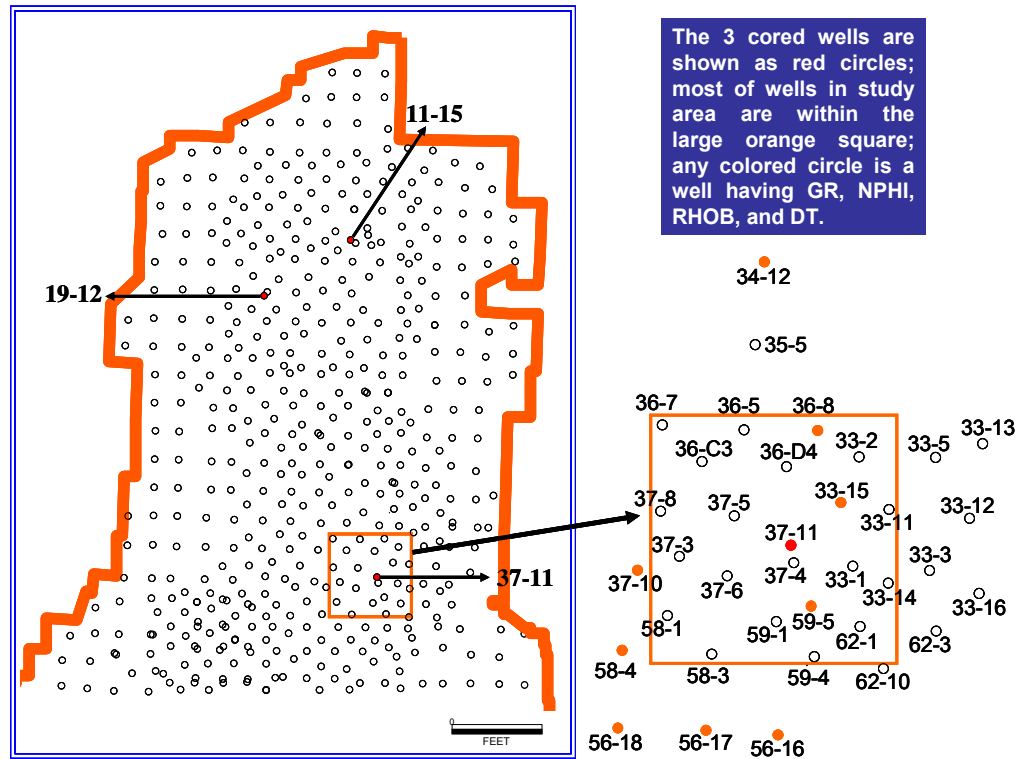


Figure 7: Well Locations at the Northern Platform of the SACROC Unit

Porosity and permeability data used here are measurements taken from whole core samples at one foot sampling intervals. Three values for permeability were measured on every sample: K0, K90, and Kv. The K0 whole core measurements were taken along an orientation line marked on the cores immediately after the cores were brought to the surface; the K90 whole core measurements were taken at 90 degrees from the K0 master line, and the Kv measurements were vertical whole core measurements. In our analyses, we used the K0 values based on the fact that no noticeable directional bias of this permeability data was identified.

Raines and Helms analyzed the effect of core size (scale) on permeability measurements values from these same three wells⁶. Permeability was measured on twelve-inch long core segments. These segments were then cut into 6-inch long segments and permeability measured, and then further cut into 2-inch long segments and permeability measured vertically and horizontally. They made several observations, perhaps the most important being that Kv is

sensitive to core size, and that permeability in samples less than 6" in height overestimate K_v at the 12" scale. A possible reason for this, they explained, is that longer samples are more likely to encounter depositional barriers to vertical flow. They concluded that "To model [vertical permeability in] flow units, sample heights must be greater than 6 inches". They also concluded there is little variation in horizontal porosity between short core plugs and longer core segments so that scaling is not an issue: "...porosity may be scaled up from plug scale or two inch scale with a reasonable expectation of achieving twelve inch equivalent data"⁶. Further discussion of core permeability/porosity data analysis is discussed in Appendix A.

For this study, core porosity was converted from percentage to fractional units, and the base 10 logarithm of permeability parameters was used. In general terms, core data indicates that porosity tends to increase from top to bottom in the upper half of Canyon Reef interval and then decreases through the lower portion of the interval. The core permeability (more specifically its logarithm value) follows a similar trend. Due to the thin-bedded nature of flow units, these changes are not smooth with depth and the well logs do not necessarily "see" these small-scale variations in porosity and permeability. Definitively, the Surface Seismic (SS) traces are not able to "see" these changes.

The relationship between permeability (K_0) and porosity for the three cored wells is shown in Figure 8. Despite the appearance of some linear trends (in terms of clouds alignment) which can be seen for each well data, overall there is not a convincing linear relationship between permeability (\log_{10}) and porosity values. This is not unexpected since the reservoir is made of numerous different depositional facies, each of which can be expected to have different porosity-permeability relationships (see next section).

3.3 *Seismic Information*

Surface Seismic 3D information covering the whole North Platform of the SACROC Unit was available from KMCO₂. The interval of interest occurred between a depth of 6,000 feet (MD) and 7,000 feet (MD) (essentially from just below the Wolfcamp shale to the oil-water contact), and was sampled at an interval of 1.0 foot. This seismic data was part of a modern 3D survey recorded within the SACROC field and was considered of good quality⁷.

Synthetic seismic traces were computed using sonic and density logs from approximately 81 wells. Seismic attributes for the synthetic traces were computed alongside corresponding attributes from 3D, migrated seismic traces located near each of the wells. A total of 16 seismic attributes were computed, which comprise most of those commonly used in attribute analyses⁷ (see Appendix B).

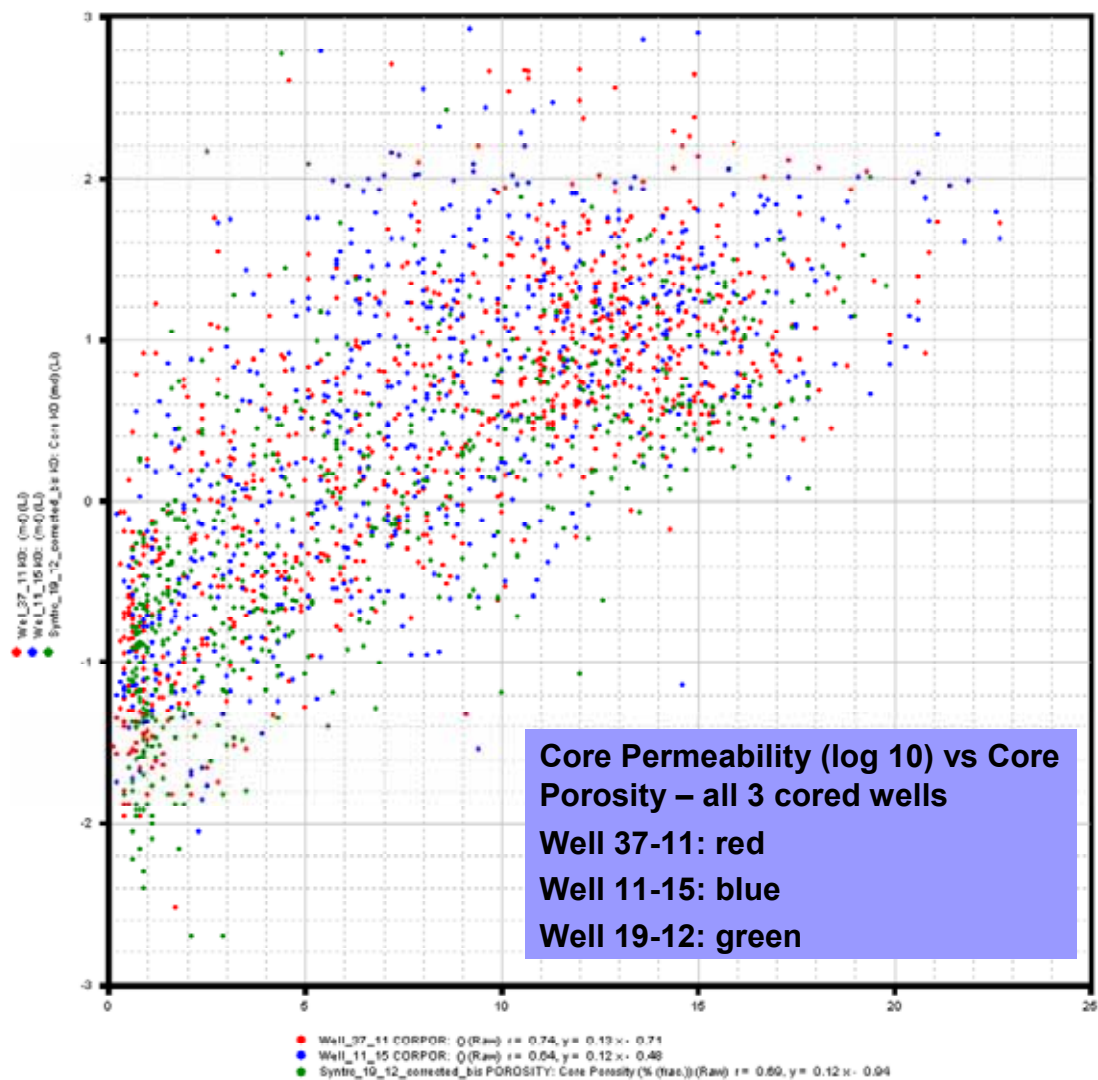


Figure 8: Core Permeability (Log₁₀) vs. Core Porosity

In order to align the seismic trace with the logs, a synthetic seismic trace was computed. This allowed for a more valid alignment, inasmuch as the synthetic and real traces theoretically respond to the same rock properties. Since a synthetic trace requires a sonic and density log, only wells with a sonic and density log could be used in this analysis. The seismic attributes were computed from only the nearest trace, and for the interval between the depths of 6,000 ft and the deepest point on the sonic or density log. For about one-half the wells, the shallowest log value was below 6,000 ft, in which case the trace attribute calculations began there. The real seismic traces were also computed and re-sampled to one foot and aligned with the synthetic trace by taking into account the KB elevation of the well and the 2300-foot datum elevation of the seismic data.

Three sets of output files were produced for each well. The first set contains the unfiltered seismic attributes computed from the synthetic trace. The second contains the same seismic attributes but filtered to more closely match the frequency bandwidth of the real data. The third set contains the attributes computed from the real seismic trace nearest the well.

The sequential order of traces in the real-seismic log files was:

1. Depth (in feet)
2. Density
3. Sonic
4. Reflection coefficient trace
5. Amplitude of synthetic seismic trace
6. Amplitude of real seismic trace
7. Absolute amplitude
8. Acoustic impedance
9. Trace derivative
10. Trace envelope
11. First derivative of trace envelope
12. Second derivative of trace envelope
13. Hilbert transform
14. Instantaneous phase
15. Response phase
16. Cosine of instantaneous phase
17. Instantaneous frequency
18. Response frequency
19. Perigram
20. Perigram multiplied by cosine of instantaneous phase.

The first five traces come directly from the well data; the remaining traces comprise the seismic trace and its computed attributes. The sequence of traces for the synthetic-trace files is the same except that the real seismic trace is omitted and the attributes apply to the synthetic trace. More details about the computation of seismic attributes can be found in Appendix B.

3.4 Crosswell Data

One of the primary goals of this project was to create three data-driven devices (or eventually an integrated one) capable of utilizing the raw data, as well as the clustering information, to relate

- Surface to crosswell seismic (specifically seismic attributes to crosswell traces).
- Crosswell attributes (computed from crosswell traces) to geophysical log responses.
- Geophysical logs to core permeability and porosity.

These data types should be explored and conditioned, individually and in combination, using clustering techniques which could identify patterns and commonalities in data. With these three “intelligent” devices, it is possible to provide insights into lithofacies and depositional environments, and consequently any surface seismic trace could be deconvolved from a low resolution elastic waveform to a high-resolution representation of permeability and porosity. These relationships would permit the development of a model capable of predicting core-scale porosity and permeability profiles even in locations where only 3D surface seismic data had been “shot”. We believe that crosswell measurements are a key factor to overcome the resolution gap problems between considered data sets.

However, as of July 2007, KMCO₂ had not been available to perform the planned crosswell survey. This delay constituted a serious drawback in reaching the original objectives of the project, and termination of the project was considered. Instead the characterization process was continued using the available information (i.e., core data, well log information and 3D surface seismic data), with the expectation that at any moment the specific crosswell information would be acquired, delivered, and incorporated into the characterization tasks.

After several attempts to generate an efficient model with the application of data-driven methodologies (consistent with the originally proposed objectives), and utilizing the available information (no crosswell data at hand), the goal of reconciling the gap between core data and 3D surface seismic information was only modestly achieved. The essential goal of multi-scale data integration for a high-resolution reservoir characterization uniquely using data-driven techniques was considered unsatisfactory. For an additional exposition of the failed utilization of seismic attributes in the attainment of an integrated data-driven device, please see Appendix C.

However, as a product of these efforts, a two-step “soft-computing” procedure was developed capable of efficiently generating core-scale porosity and permeability values (as well as rock types geologically consistent) at well locations where only GR and NPHI were available, which is the most common situation in this reservoir. Because the suitable logs for the creation of an “intelligent” Log-to-Core device are not present at all wells, it was necessary the creation of another intelligence tool, a Log-to-Log model, to provide such missing information. Figure 9 illustrates schematically this two-step procedure. The validity of these soft-computing devices was checked using “holdout” wells.

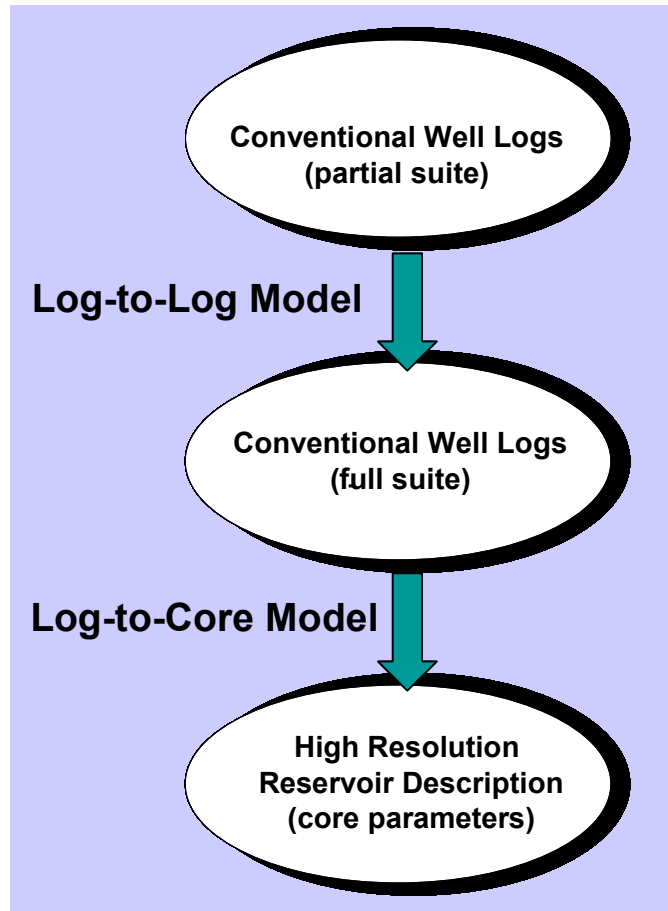


Figure 9: Schematic of the Two-Step “Soft-Computing” Procedure

4.0 Methods

4.1 The GAMLs Clustering Engine

*GAMLs*⁸ (Geologic Analysis via Maximum Likelihood System) derives its name from the maximum likelihood system by which it works. However, it could also be termed a model-based neural system (to separate it from neural nets, we abstain from the word “net”). *GAMLs* is NOT a simple back propagation neural net.

The mathematical framework of *GAMLs*, termed MLANS (Maximum Likelihood Adaptive Neural System), is described in Perlovsky and McManus⁹. MLANS has previously been used in pattern recognition problems such as missile tracking and discrimination (see Perlovsky^{10,11,12}).

MLANS is a neural system that combines *a priori* knowledge, adaptivity, and fuzzy logic. Also, it does not require “training data” (but it uses it if available), and it can process both numerical and symbolic information. Finally, MLANS performance has been demonstrated to approach the information-theoretic performance limits:

1. the Bayes error in classification and association accuracy (see Fukunaga¹³), and
2. the Cramer-Rao bound on learning efficiency or speed¹⁴

The structure of MLANS enabled its use by Nichols Research Corporation (the originator) to solve problems in the following areas (partial list):

- resource allocation and sensor fusion for missile defense
- missile attack warning for fighter aircraft using IR sensors
- submarine tracking and classification using active and passive sonar sensors
- ground target detection and classification and search & rescue of downed aircraft using synthetic aperture radar
- airborne target detection and classification for drug interdiction using over-the-horizon radars
- unmanned vehicle control using visible sensors
- data classification, archival retrieval, and compression
- medical image diagnosis
- extraction and interpretation of topside ionosonde data
- automatic target recognition algorithm development and evaluation
- fingerprint identification

GAMLs was the first application of MLANS in the geoscience field.

The expression “model-based” refers to the fact that a specific mathematical distribution, the model, is imposed upon the solution. Introduction of this requirement results in a significant operational advantage over non model-based systems, namely: 1 - the ability to perform with less

training data, by an order of magnitude or more, and therefore, 2 - speed. Even if the mathematical model imposed is not strictly accurate in all situations, the advantage of this “model-based” strategy almost always outweighs the disadvantages.

The model used by *GAMLS*, which is the model to which the solution converges, is that the frequency distribution curve for the values of each variable is approximated by the sum of n Gaussian curves where n equals the number of modes (clusters) used in the modeling. Thus, the solution provides, for each variable, a distribution (mean and standard deviation) for each mode.

A *GAMLS* solution minimizes the departure from the model, and the minimization is done using maximum likelihood techniques (see Press et al¹⁵ for a discussion of maximum likelihood techniques). The speed of *GAMLS* is attributed to the system “knowing where it is going” and it can arrive at its destination with little to no “training”. In addition, replicate computer runs will arrive at the same answer so long as the dataset and the initialization conditions remain constant.

GAMLS uses Gaussian (normal) or lognormal distributions for modeling each mode of each variable. These distributions are chosen because they provide a simple, concise representation that assumes the least prior knowledge in estimating an unknown probability distribution with a given mean and variance. *GAMLS* performs two major tasks: it clusters and it regresses. Taken together, clustering and regression with *GAMLS* permit many of the tasks routinely done by a geoscientist to be performed efficiently.

The clustering process is actually a classification process: rocks which have similar attributes are identified and then clustered (classified, grouped, etc.) together. Regression determines the relationships among the data and known parameters, and once this relationship is known, the relationship can be used to predict the values of parameters when they are not known. For more details about clustering, prediction and other *GAMLS* features, see Appendix D.

The regression is accomplished using a data-driven probabilistic superposition of linear relationships between each of the Gaussian clusters and the unknowns. This results in a generalized nonlinear relationship for a system having multiple clusters.

4.2 *The GAMLS Workflow*

From the reservoir management viewpoint, methods to create high-resolution 3D reservoir characterizations integrating well log data and seismic information (3D) remain expecting an answer accepted by everyone. In essence, the problem of establishing efficient relationships between data at different scales has not yet been solved.

The reservoir characterization procedure is based on the cluster analysis of each dataset (i.e., surface seismic, crosswell seismic, logs, and core data) to discriminate categories with similar geological properties (i.e., clusters of attributes, log signatures, core compositions, etc.). Each category corresponds to an identifiable stratigraphic unit or depositional environment. Each of the available data types is explored and conditioned, individually and in combination, using categorization techniques (forms of self organizing maps) which bring elements of geologic reality into the process.

Lithologies were analyzed and interpreted using the probabilistic clustering analysis procedure GAMLS. This soft-computing software permits well logs, seismic attributes and core reservoir parameters to be used as variables in a multi-dimensional clustering analysis that results in all samples (at their respective depths) being assigned probabilistically to a user-defined number of "modes". A "mode" is a mathematical term that has the same functional meaning as "cluster." The modes can be envisioned as electrofacies (or lithofacies or flow units) which have similar properties. Details of the GAMLS⁸ procedure are also given in Eslinger et al¹⁶.

In the multi-well clustering run for this study, ten (10) modes were chosen that were deemed appropriate for an analysis of this type in this complex reservoir. Nevertheless, the number of modes is somewhat arbitrary, and a greater or smaller number could be used, however ten modes usually provides sufficient detail..

A clustering run needs to be initialized under selected mathematical conditions. This initialization provides GAMLS⁸ with the initial means and covariances for each variable and for each mode of each well. Essentially, initialization gives a starting point for the clustering run in the space of solutions. It provides initial probability assignments to each sample, and can be done in a supervised or non-supervised manner (cf. Eslinger et al¹⁶). The type of initialization used for the clustering run was the Large Covariance (LC) method which is an unsupervised method used when little is known about the rocks, and when no single variable is believed to be most important in discriminating the samples into rock types. That is, no *a priori* assumptions are made about relationships among the variables or between the variables and rock type. This initialization method assigns equal "weight" to each of the variables.

Finally, two additional parameters must be set-up before initialization of any clustering run. These parameters establish the numerical criterion of convergence, and they are the "Number of Iterations" and the "Convergence Goal". The number of iterations is the number of runs executed by GAMLS⁸ to reach a maximum likelihood classification (the default value is 500). The "convergence goal" is a number which, when reached, terminates the clustering run. The convergence goal is the maximum difference in fractional probability assignments for any mode of any sample. Thus, the convergence goal is a low number (typically, 0.001) which, when reached, indicates that further iterations would not cause the mode assignments to change in a significant manner.

It should be mentioned that although detailed whole core descriptions have been made of the SACROC strata in the three selected cored wells, the descriptions were not available for use in this study to incorporate as an established sedimentary model. Access to this information would be useful in future research to improve results.

The upper-most depth used for each well was the top of the called "Canyon Reef" as picked by the dataset provided by KMCO₂, and ratified by the clustering analysis here developed.

A typical work flow utilizing GAMLS can be summarized as following:

1. Select well log curves to be utilized in clustering. Typically, RHOB, NPHI, and GR are considered. Other curves like DT, PEF, and resistivity profiles can also be very revealing and could be used as needed and as available.
2. Edit data as needed to remove erroneous or false data and to perform interwell curve normalization. Examination of interwell curve normalization necessity was evaluated but it was concluded that it was unnecessary.
3. Make appropriate core-to-log depth corrections to ensure that the core plug data is "on depth" with the well log curves. These corrections can be done using the core plug grain density data and also the core gamma ray scan. For this project, it was necessary to perform this activity because of depth variations in all three cored wells. Depths were corrected utilizing statistical tools.
4. Set up a clustering run using the following steps:
 - select wells and depth ranges (multi-well clustering is possible and usually desirable)
 - select the variables (well logs)
 - select the number of modes
 - initialize the "run" under conditions adapted to the problem.
 - iterate to convergence; convergence occurs when the sample probability assignments cease to change significantly for successive iterations; the model that drives towards a solution is that the frequency distribution for all modes of all variables will approach a normal distribution at the maximum likelihood solution; each clustering solution is considered to be a model "realization"; replicate runs give the same answer if the setup procedure is exactly replicated
5. Evaluate the results via examination of a series of a output tables and plots

In this particular project, the described procedure allowed the division of the geologic section of interest into rock types (facies units). This was done to interpret the lithology of the facies units, to examine the degree to which the facies units might be correlated among the wells, and to estimate values for missing data (particularly RHOB, DT, porosity, and permeability).

5.0 Discrimination and Interpretation of Facies Units

If the modes are considered to be facies units, then the clustering process automatically discriminates the clustered depth intervals into facies units. Since the sample assignments are probabilistic, each sample can be assigned probabilistically to more than one facies unit. This is referred as the “fuzzy” probability assignment in the concordance with the Fuzzy Logic theory. It is accepted that reality does not work in black and white, but in shades of grey. In essence, Fuzzy Logic theory expresses that any interpretation (rock type) is possible between two end points, but some are more probable than others^{17,18}. In a depth plot, the “fuzzy” probabilistic assignments are displayed as a stacked bar chart on a horizontal axis with axis ranging from zero (probability) at left to 1.0 (probability) at right. The sum of the mode probability assignments at each depth (each sample) is 1.0.

The probability (“fuzzy”) assignments can also be displayed as “crisp” assignments at each depth. The “crisp” mode assignment is the mode (facies unit) which has the highest “fuzzy” assignment so at each depth, a unique and definitive mode is declared. Contiguous “crisp” assignments in depth to the same mode define a “bed”.

5.1 Interpretation of Lithology of Facies Units

A “ModeAssign” routine within GAMLS⁸ automatically makes lithology assignments to each of the clustering modes (facies units). This routine is based largely on GR signal for clastic rocks and apparent grain density for carbonate rocks (see Eslinger¹⁹); however, the rules for these lithology assignments can be changed by the user in conformity with a preconceived understanding of the studied reservoir. The automatic assignments can also be overridden if available core data indicates the automatic assignments are not correct.

5.2 Correlation of Facies Units Among Wells

Depth plots displaying the probability assignments of all samples can be used to provide an easy method for visual examination and interpretation (see figures below). In complex reservoirs with rapid vertical changes in depositional environment or diagenesis, thin beds defined by clustering can be eliminated by a process called “bed thickness filter”. The filter, which operates on the “crisp” beds is a simple form of upscaling and with some stratigraphic intervals might be used to aid correlation of major (thick) facies units.

5.3 Estimation of Missing Data

A major feature of GAMLS is that values for missing (null) data are automatically generated during the clustering process. Data can be estimated for any curve used as a clustering

variable. It is important to note that the data which can be used as clustering variables is not restricted to well log data. Any type of data that can be digitized can be used as a variable. Furthermore, this data does not need to be continuous data, such as well log data, but can be discontinuous data, such as core plug porosity and permeability.

This means that missing log data can be generated during a clustering run. It also means that core plug data, which might be missing for a large portion of the interval, can be estimated during a clustering run.

In this study, the variables used in the majority of the clustering runs were RHOB, NPHI, GR, and DT. But, excluding the three cored wells, only nine wells inside and around the test area had all four of these curves. Most other wells had no RHOB or DT logs, so, RHOB and DT curves were generated during clustering and were incorporated in the characterization of the study area.

Additionally, since only three wells had core data, two of which were approximately 800 feet from the central part of the test area, the "estimation" capability was used to estimate core-scale porosity and permeability values for most of the non-cored wells. This procedure is discussed more fully below.

5.4 Lithofacies Units versus Electrofacies Units versus Flow Units

The term "facies" is used above as the name for a group of samples that have similar well log character as defined by a cluster analysis. The term "electrofacies" would be appropriate if all of the variables used in the clustering run were curves obtained from well log "tools". If the assumption is made that the well log curves are responding to lithology differences, then the term "lithofacies" would also be appropriate for the clustered groups. If one or more of the variables is a flow property (such as permeability) or is related to a flow property (which most well log curves are), then the term "flow unit" would also be appropriate. The choice of what term is best used to describe the "groups" that are generated by clustering is somewhat a semantic issue but also depends on the prejudice of the analyst. Here, the term "mode", "facies" and "flow unit" are used interchangeably.

6.0 Reservoir Quality of Facies and Flow Units

6.1 Background

A major goal of this study is to define flow units and to determine their reservoir quality (RQ). We here use RQ to mean collectively the porosity, permeability, bed thickness, and lateral bed continuity of the flow units.

Since facies units are considered to be equivalent to flow units (see discussion above), then flow units can be defined by clustering analyses using well log curves as variables or a combination of well log curves and core permeability and core porosity. Multiple clustering runs can be made where each run uses a different suite of variables. Each clustering run provides a different flow unit realization. Ideally, comparison of different clustering realizations would result in similar flow unit realizations.

There are several different work flows that can determine the RQ of flow unit realizations. It is not obvious, a priori, which work flow will produce the most realistic (closest to the "truth") flow unit realization and which work flow will produce the most credible definition of RQ for any given flow unit realization. It will be shown below that the GAMLS⁸ modeling procedure is quite robust in that the porosity and permeability profiles obtained are largely independent of the particular work flow used.

Because only three wells were fully cored and only one of those three wells is near the center of the study area, the database for building a model for estimating porosity and permeability for the non-cores wells is limited. On the positive side, the cored wells were plugged every foot over a depth range from approximately 600 to 900 feet, so there is a very good core control for the three cored wells.

6.2 Flow Unit Realizations and Reservoir Quality

For each clustering run, or series of clustering runs, the resulting realization can provide the following information for each flow unit:

1. the arithmetic mean and standard deviation of porosity,
2. the arithmetic mean and standard deviation of the logarithm (base 10) of permeability,
3. the total bed thickness, the number of beds, and the average bed thickness for the original beds (defined by clustering),
4. the number of beds after bed thickness filtering to eliminate beds thinner than, for example, 10 feet (any minimum bed thickness can be selected), and

5. a qualitative assessment of well-to-well lateral continuity of each flow unit; this is done by generating for each well either a depth plot of the “fuzzy” probabilistic assignment or a depth plot of the “crisp” (mode) assignment and correlating these representation by visual examination.

6.3 *Estimation of Porosity and Permeability in Non-Cored Wells*

The basic approach is to perform a clustering run using one or more cored wells with core porosity and/or core permeability as variables along with a selected log curve suite as variables. For any non-cored wells included in the clustering run, estimates of core porosity and/or core permeability are automatically made during the clustering process. Choices to be made for any given clustering run are:

1. how many cored wells to include,
2. the inclusion of one or both core porosity and core permeability for cored wells,
3. which non-cored wells to include, and
4. which log curve variables to include

Another choice is to define flow units using a clustering run that does or does not include core data as variables. This last approach would require the determination of porosity and permeability profiles utilizing a previous clustering run(s) that uses core porosity and/or permeability and then assignment of these profiles to a facies determined using a clustering run that used only log curves as variables. This is the method used for cluster runs C1A and C1B (see below).

An alternative would be to determine both the facies and the profiles for porosity and permeability in the same clustering run. This is the method used for cluster runs C1C, C1D, and C1E (see below).

For methodology purposes at the SACROC Unit test site, the petrophysical curves and core parameters were utilized to create a wells classification based on their availability:

- *Group G1* consists of wells with GR, NPHI, RHOB, DT and core data. The common core measurements for the cored wells were porosity, horizontal permeability (K0 and K90), vertical permeability (Kv), percentage of fluorescence, and grain density data. In the study area, only the well 37-11 satisfied this group conditions. However, the other two cored wells (11-15 and 19-12) were included in the procedure to enrich the learning process of the data-driven device conceived to model the log-to-core relationship.
- *Group G2* consisted of wells with GR, NPHI, RHOB and DT, but without core data. Three wells within the study area met this requirement (33-15, 36-8, and 59-5), and an additional six wells in the near vicinity met of the study area (34-12, 37-10, 56-16, 56-17, 56-18, and 58-4). The most distant of the six, the 34-12 was only 2,500 feet away.

- *Group G3* consisted of wells with only GR and NPHI. Twenty five wells were included in this group. Eighteen wells were within the study area and remaining wells are no farther than 1,400 feet from the study area boundaries.

Since most of the wells had no RHOB or DT logs, a choice as to "fill in" the "ideal" log suite by estimating both RHOB and DT profiles for these wells (using GAMLS methods), or to combine these two variables into only one variable, acoustic impedance (AI_log), needed to be made. Both approaches were used.

6.4 Porosity and Permeability via the "AI" Method

The calculation of log AI supplies an additional log which may possibly provide a better correlation with seismic acoustic impedance and which can be used for clustering analysis between well logs and seismic attributes. Additionally the use of an AI_log rather than a RHOB and DT, reduces the absence of two geophysical parameters to only one when clustering processes are later carried out on those wells having only GR and NPHI (group G3).

To elucidate the real effectiveness of using AI_log instead of RHOB and DT to predict porosity and permeability values, corresponding clustering runs were executed with and without these well logs. Predicted porosity and permeability values that were generated via a clustering run that utilized GR, NPHI and AI_log, (here called cluster C9) were equivalent to predicted values obtained via a clustering run using GR, NPHI, RHOB and DT (here named C1AA). Both clustering runs also included core porosity and core permeability (\log_{10}) as key parameters coming from wells 37-11 and 19-12 (cored well 11-15 was not included). Results are illustrated in Figure 10, Figure 11, and Figure 12.

Figure 10 shows two crossplots comparing predicted values of porosity (left) and permeability (right) using the clustering run C1AA, versus C9. Results for cored wells 19-12 (green), and 11-15 (red); and for the non-cored well 33-15 (blue) were superimposed in both plots. The corresponding aligned clouds (by colors) demonstrate a strong linear relationship between these two different types of predicted values. Correlations coefficients corroborate this fact.

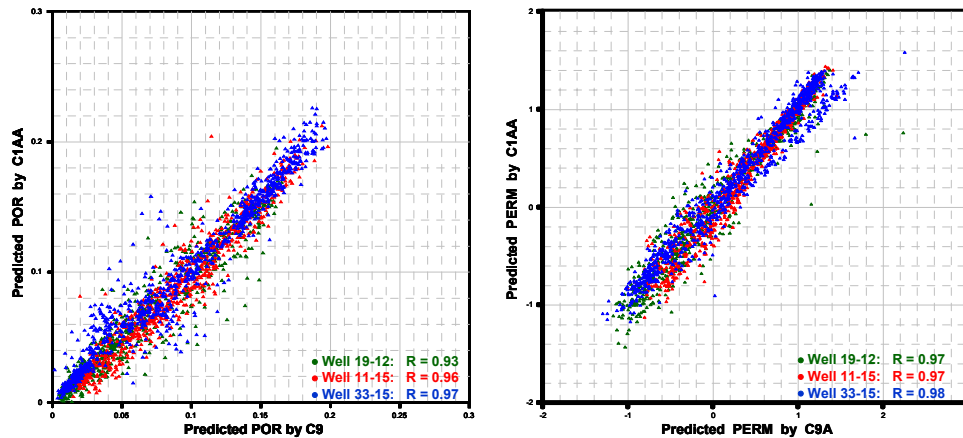


Figure 10: C1AA vs. C9, Left-Porosity, Right-Permeability (Log₁₀)

Figure 11 presents porosity and permeability tracks for wells 33-15, 11-15, and 19-12. Predicted porosity values by using cluster C1AA (dark blue) were superimposed to predicted porosity values by using cluster C9 (light blue). The range of variability in these tracks is from 0.0 to 0.30. Likewise, predicted permeability values (\log_{10}) using cluster C1AA (dark green) were superimposed to predicted permeability values (\log_{10}) by using cluster C9 (light green). The range of variability in these permeability tracks (\log_{10}) is from -3.0 to 3.0 . In both cases, it is hard to say that predicted values coming from different clustering runs are different.

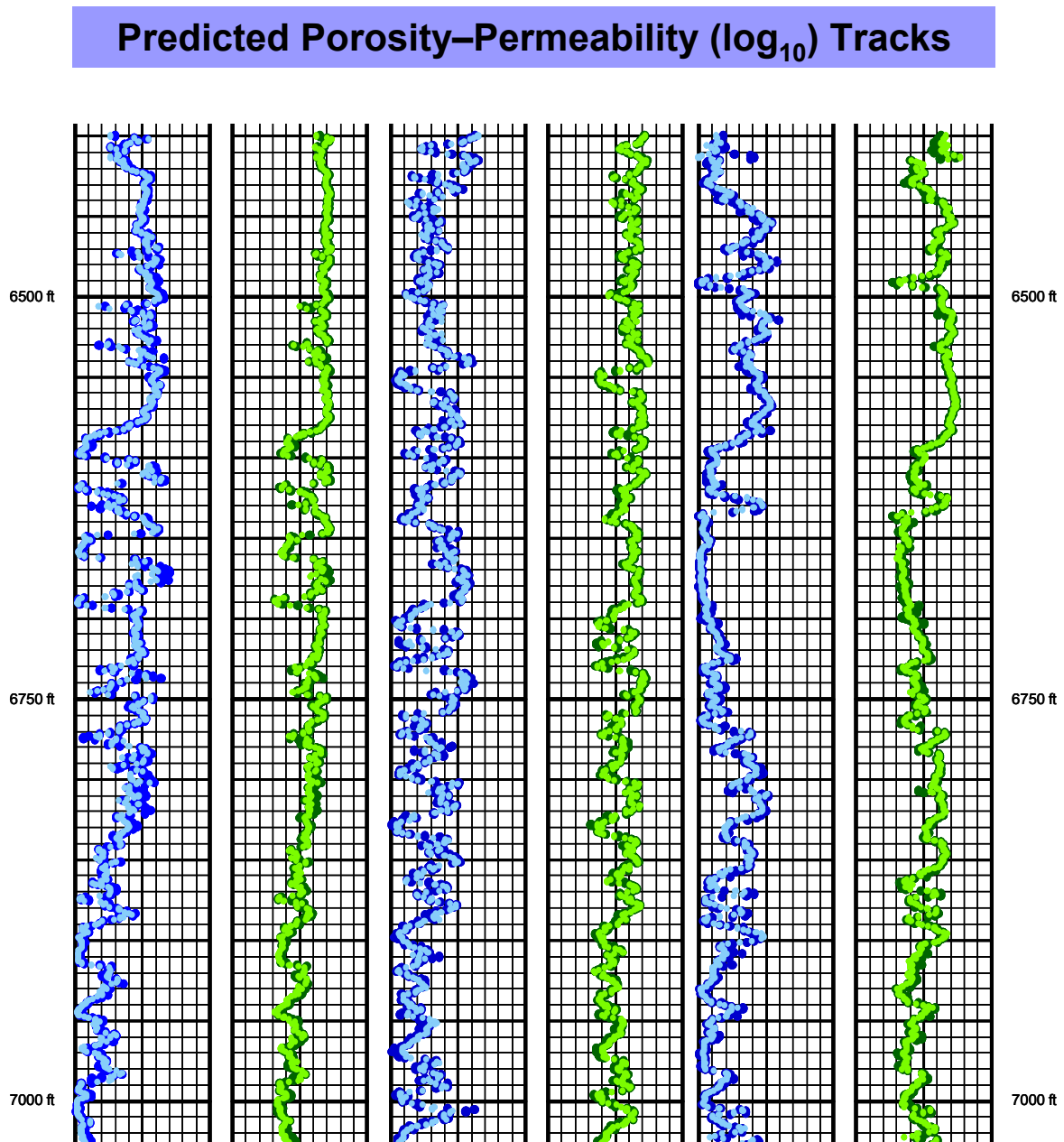


Figure 11: Predicted Porosity and Permeability Tracks for Wells 33-15, 11-15, and 19-12

Predicted porosity and the logarithm of permeability values are plotted jointly with actual core porosity and permeability data for cored wells 11-15 and 19-12 in Figure 12. Here, predicted porosity values using cluster C1AA are in dark blue, predicted porosity values using cluster C9 are in light blue, predicted permeability values using cluster C1AA are in dark green, predicted permeability values cluster C9 are in light green, and actual values are in red. The same variability range above exposed is been applied. It can be appreciated that both approaches show similar efficiency as predictors of core porosity and core permeability.

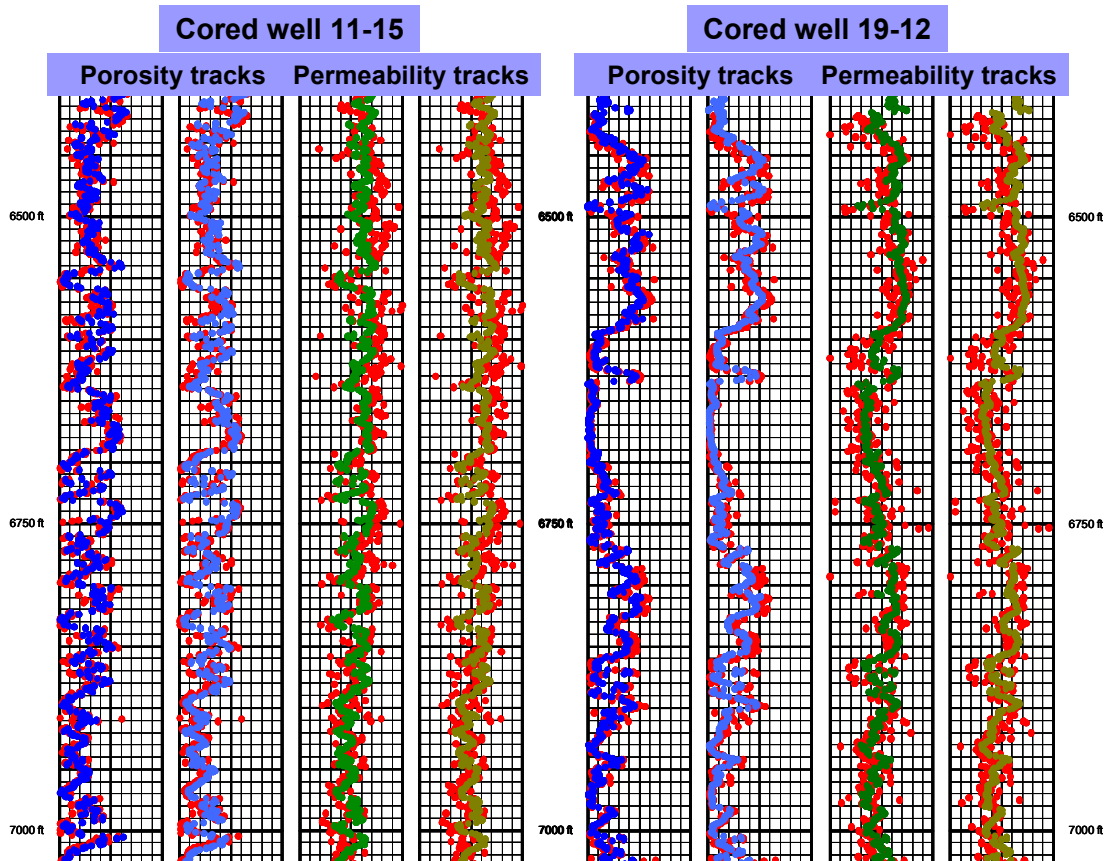


Figure 12: Porosity and Permeability (\log_{10}) Tracks for Wells 11-15 and 19-12

The procedure for combining RHOB and DT into AI_log and using AI_log as a clustering variable is herein termed the "P&P via AI" method. This method involves several clustering runs which included AI_log, selected additional log curves (NPHI and GR), plus one or both permeability and porosity. The end results are "predicted" profiles for porosity and permeability in non-cored wells. This "P&P via AI" method was done in the following step-wise manner:

1. Using wells in groups G1 and G2, a clustering run was performed which included only GR, NPHI, and AI_log as log parameters. Notice that all wells belonging to these two groups have real data for the variables. The target of such procedure was to

provide an “intelligence” device capable of predicting AI_log in wells without RHOB and DT (group G3). This first step also shows the feasibility of applying data-driven methods to simulate missing logs based on the presence of other.

2. Since an actual AI_log does not exist in wells of group G3 (RHOB and DT are absent), it was necessary to compute AI_log values in order to have a complete set of logs for predicting core parameters values. GAMLS permits the use of a previous clustering run to predict missing values of one of the variables utilized in such run at other locations. In consequence, the clustering results of the previous step were utilized to predict AI_log at wells in group G3. This provided wells of group G3 a complete set of parameters necessary for later prediction of porosity and permeability values at core-scale based on the presence of GR, NPHI and AI_log .
3. The next step was to generate core porosity and core permeability curves at wells belonging to groups G2 and G3. This task was executed individually for each group and under different conditions.
 - First, core porosity and core permeability values were simultaneously generated for wells in group G2. Utilizing wells in groups G1 and G2, a clustering analysis was run with actual data GR, NPHI, and AI_log of both groups, and core P&P measures for wells belonging only to group G1. This clustering analysis resulted in the generation of simulated core parameters values on wells of group G2. Thus conceived, this first step avoids the use of predicted AI_log curves of G3 wells as an input, and takes full advantage of the actual data of G1 and G2 wells.
 - Next porosity and permeability values were generated at wells belonging to the group G3. Using all well groups, a clustering run was executed with the following characteristics:
 - i. G1 wells - actual GR, NPHI, AI_log, and core P&P parameters
 - ii. G2 wells - actual GR, NPHI, and AI_log jointly with the pseudo core P&P generated previously
 - iii. G3 wells - actual GR and NPHI, and predicted AI_log

This clustering step resulted in the generation of pseudo core P&P estimated values at wells belonging to group G3. Table 1 summarizes the direct workflow followed in the “P&P via AI method”. Other clustering runs were omitted that were performed as part of the search for the optimum workflow. These omitted clusters were coded as C1, C3, C4, C5, C6, C7, C8, and C11. For instance, three of these clusters (C1, C7, and C8) didn’t include core well 11-15. Well 11-15 was used as a "hold out" well to test the accuracy of the predictions. The other clusters were used in various tests to help optimize the work flow, and some were designed to determine the most convenient way of predicting AI_log or core-scale P&P for wells without actual data. They were used to make decisions about whether or not estimate simultaneously porosity and permeability, or about how to use more efficiently the simulated values generated on those wells without any particular variable.

Figure 13 shows tracks of AI_log, core porosity, and core $\log_{10}(K_0)$ respectively for cored well 11-15, the hold-out well. Actual values are in red and predicted values are in blue. These results were obtained during the sequence of different clustering runs done to establish the “P&P via AI” method. For instance, cluster C1 was the first clustering run used to predict AI_log at wells. Cluster C1 had identical conditions to cluster C2 except that well 11-15 was excluded. The first two tracks indicate that agreement between predicted values and actual values are pretty good. In the track for $\log_{10}(K_0)$, we can appreciate that spatial tendencies are reproduced; however, larger numerical differences between actual and predicted values are evident.

Figure 14 shows crossplots of predicted vs. actual values. AI_log is shown at left; porosity is seen at center, and $\log_{10}(K_0)$ is on the right. All crossplots show a good correlation between predicted and actual values. The corresponding correlation coefficients (r) are 0.94 for AI_log, 0.74 for porosity, and 0.59 for $\log_{10}(K_0)$. As can be expected for permeability, the prediction is the least accurate in absolute terms. However, the vertical variability of permeability is reproduced quite well, the predictions simply do not fully capture the extreme values. Nevertheless, reproducing the vertical variability is arguably the most important objective when utilizing such results for reservoir flow simulation purposes. This subject is revisited in section 7.0, particularly in the discussion related to Figure 62.

Cluster run C8 was a precursor of cluster run C9 in providing estimates values of porosity and permeability. The set up for cluster run C8 was identical to cluster C9 except that well 11-15 was not included. Again (second track) porosity values were excellently reproduced. However, permeability (\log_{10}) predicted values were not as good as the porosity predicted values. This is discussed in more detail below.

In summary, the “P&P via AI” method allowed populating all wells included in the selected study area with core-scale P&P curves which facilitated the application of geostatistical methods for reservoir characterization.

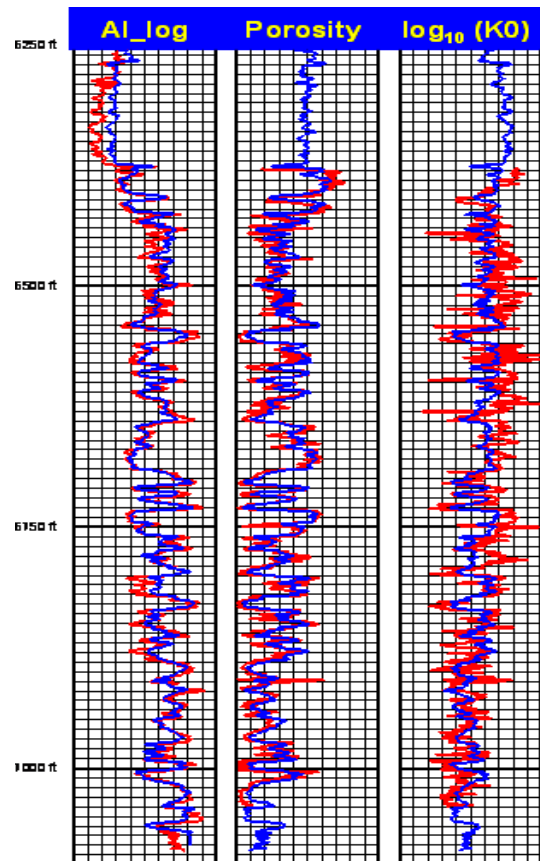


Figure 13: AI_log (1st track), Porosity (2nd track) and $\log_{10}(K0)$ (3rd track) for Testing Hold-Out Well 11-15 Actual Parameter Values in Red; Predicted Values in Blue

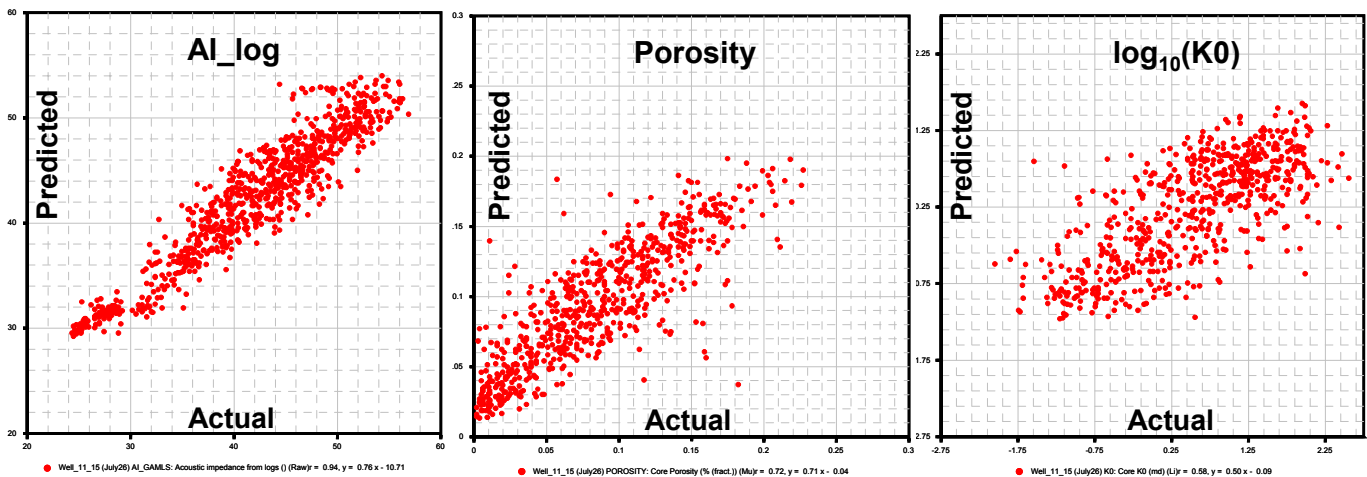


Figure 14: Predicted vs. Actual Values for AI_Log, Porosity, and $\log_{10}(K0)$ Ranges: AI_Log [20, 60]; Porosity [0.0, 0.3]; $\log_{10}(K0)$ [-2.75, 2.75]

Table 1: Clustering Procedures for the P&P Via AI Method for Predicting Core-Scale P&P Estimates

Cluster	Parameters	Wells	Observation
C2	GR, NPHI, AI	G1 + G2	Used for predicting AI on G3 wells (well 11-15 included)
C9	GR, NPHI, AI POR, K0	G1 + G2	Used for estimating POR and K0 on wells only of G2 (well 11-15 included)
C9_K90	GR, NPHI, AI POR, K90	G1 + G2	Clone of C9. Used for estimating K90 only on G2 wells (C9-POR used on G2 wells)
C9_Kv	GR, NPHI, AI POR, Kv	G1 + G2	Clone of C9. Used for estimating Kv only on G2 wells (C9-POR used on G2 wells)
C10	GR, NPHI, AI POR	G1 + G2 + G3	Used for estimating POR only on G3 wells (using C9-POR estimates on G2 wells)
C12	GR, NPHI, AI POR, K0	G1 + G2 + G3	Used for estimating K0 on only G3 wells (using C9-POR and C9-K0 estimates on G2 wells, and C10-POR estimates on G3 wells)
C12_K90	GR, NPHI, AI POR, K90	G1 + G2 + G3	Clone of C12. Used for estimating K90 on only G3 wells (using C9-POR and C9-K90 estimates on G2 wells, and C10-POR estimates on G3 wells)
C12_Kv	GR, NPHI, AI POR, Kv	G1 + G2 + G3	Clone of C12. Used for estimating Kv on only G3 wells (using C9-POR and C9-Kv estimates on G2 wells, and C10-POR estimates on G3 wells)

7.0 Realizations and Their Results

Five clustering runs are discussed below to analyze the robustness of the methodology for Rock Taxonomy and Reservoir Quality in the Canyon Formation of the SACROC Unit (Kelly-Snyder Field), from a Probabilistic Clustering of Well Logs Data and Core Information using Soft Computing Methods. The first three are discussed in greatest depth. The setup for these clustering runs is summarized in the Table 2.

Table 2: Summary of Clustering Runs Utilized to Analyze Methodology Robustness

Run	Wells	Variables	Modes	Initialization
C1A	12 (3 cored)	GR, NPHI, RHOB, DT	10	LC
C1B	as C1A + 19 more	as C1A but 19 had no RHOB or DT	10	prior results of C1A
C1C	as C1A	as C1A and P&P from well 37-11	10	LC
C1D	as C1A	as C1A and P&P from well 37-11	25	LC
C1E	as C1A	as C1A and P&P from wells 37-11 & 11-15	10	LC

7.1 *C1A - Clustering Using a 4-Curve Log Suite and Application of “P&P via AI” to this Realization*

7.1.1 *Cluster Setup Procedures*

Porosity and permeability (P&P) profiles were determined using the RQ via AI method (described above), and these profiles were then applied to the flow unit realization described below. Application of the P&P profiles is mechanically done by inserting the P&P data into the already imported “LAS” files. The mean and standard deviation of the imported data is then included in the “fuzzy” and “crisp” statistics tables for all clustering runs. This data can be readily examined and plotted.

Cluster 1A (C1A) used strictly log data from three cored wells and nine non-cored wells. Figure 15 shows the setup for this clustering run. The cored wells were 37-11 (in the center of the study area) and two wells, 19-12 and 11-15, located to the NNW and NW about 8000 feet from the center of the study area. The non-core wells used were 33-15, 23-12, 36-8, 37-10, 56-

16, 56-17, 56-18, 58-4, and 59-5. Except for well 34-12, these wells were within about 5000 feet of well 37-11.

There was no attempt to restrict the depth ranges clustered to the Cisco and Canyon sections. The tops and bottoms of the intervals selected for clustering extended up into the Wolfcamp Shale for most wells and also likely extended down into non-productive water zone of the Canyon or below the Canyon.

The clustering variables were RHOB, NPHI, GR, and DT. All of the wells had continuous logs for these curves, so there was no missing data. An initialization method ("large covariance") was used that puts no emphasis on any of the particular variables. Ten modes were selected for clustering.














GAMLS clustering analysis "setup" for C1A								
12 wells, 10 modes; variables used: RHOB, NPHI, GR, DT; no missing log data								
3 wells were cored: 37-11 (in the heart of the study area) and 19-12 and 11-15 (about 8000 ft to the NW & NNW of the central study area)								
Well	Top	Bottom	% To Use	Window	Curve #1	Curve #2	Curve #3	Curve #4
 C1A								
 Well_33_15	6094.00 ft	7054.00 ft	100.00 %	No	BULK_DENS (Raw)	DT_LOG (Raw)	GR (Raw)	NPHI (Raw)
 Well_34_12	6245.00 ft	7012.00 ft	100.00 %	No	BULK_DENS (Raw)	DT_LOG (Raw)	GR (Raw)	NPHI (Raw)
 Well_36_8	6000.00 ft	7053.00 ft	100.00 %	No	BULK_DENS (Raw)	DT_LOG (Raw)	GR (Raw)	NPHI (Raw)
 Well_37_10	6097.00 ft	7029.00 ft	100.00 %	No	BULK_DENS (Raw)	DT_LOG (Raw)	GR (Raw)	NPHI (Raw)
 Well_37_11	6087.50 ft	7076.50 ft	100.00 %	No	BULK_DENS (Raw)	DT_LOG (Raw)	GR (Raw)	NPHI (Raw)
 Well_56_16 (July26)	6147.50 ft	7028.50 ft	100.00 %	No	BULK_DENS (Raw)	DT_LOG (Raw)	GR (Raw)	NPHI (Raw)
 Well_56_17 (July26)	6089.00 ft	7021.00 ft	100.00 %	No	BULK_DENS (Raw)	DT_LOG (Raw)	GR (Raw)	NPHI (Raw)
 Well_56_10 (July26)	6247.00 ft	7055.00 ft	100.00 %	No	BULK_DENS (Raw)	DT_LOG (Raw)	GR (Raw)	NPHI (Raw)
 Well_58_4 (July26)	6340.00 ft	7039.00 ft	100.00 %	No	BULK_DENS (Raw)	DT_LOG (Raw)	GR (Raw)	NPHI (Raw)
 Well_59_5 (July26)	6000.50 ft	7047.50 ft	100.00 %	No	BULK_DENS (Raw)	DT_LOG (Raw)	GR (Raw)	NPHI (Raw)
 Well_11_15	6254.00 ft	7086.00 ft	100.00 %	No	BULK_DENS (Raw)	DT_LOG (Raw)	GR (Raw)	NPHI (Raw)
 Synthetic_19_12_corrected_bis	6000.00 ft	7080.00 ft	100.00 %	No	BULK_DENS (Raw)	DT_LOG (Raw)	GR (Raw)	NPHI (Raw)

Figure 15: Screen View of the GAMLS Clustering Analysis "Setup" for Cluster C1A

7.1.2 Interpretation of Modes (Assignment of Lithologies)

Clustering results and raw logs are shown for two of the cored wells in Figure 16. For each well, the first track shows GR and DT and the second track shows RHOB and NPHI. The third and fourth tracks are core porosity and core permeability, respectively. The fifth track is a cumulative mode probability (CMP) plot which displays the "fuzzy" probabilistic assignment at each depth (the horizontal axis is fractional probability). For each sample this plot shows the probability assignments for each mode. The sixth track is a "beds" plot which displays the "crisp" (mode) assignment at each depth (here, only the mode is shown that has the highest probability for each sample depth). The CMP plot is also termed a "fuzzy" plot and the "beds" plot is also termed a "crisp" plot.

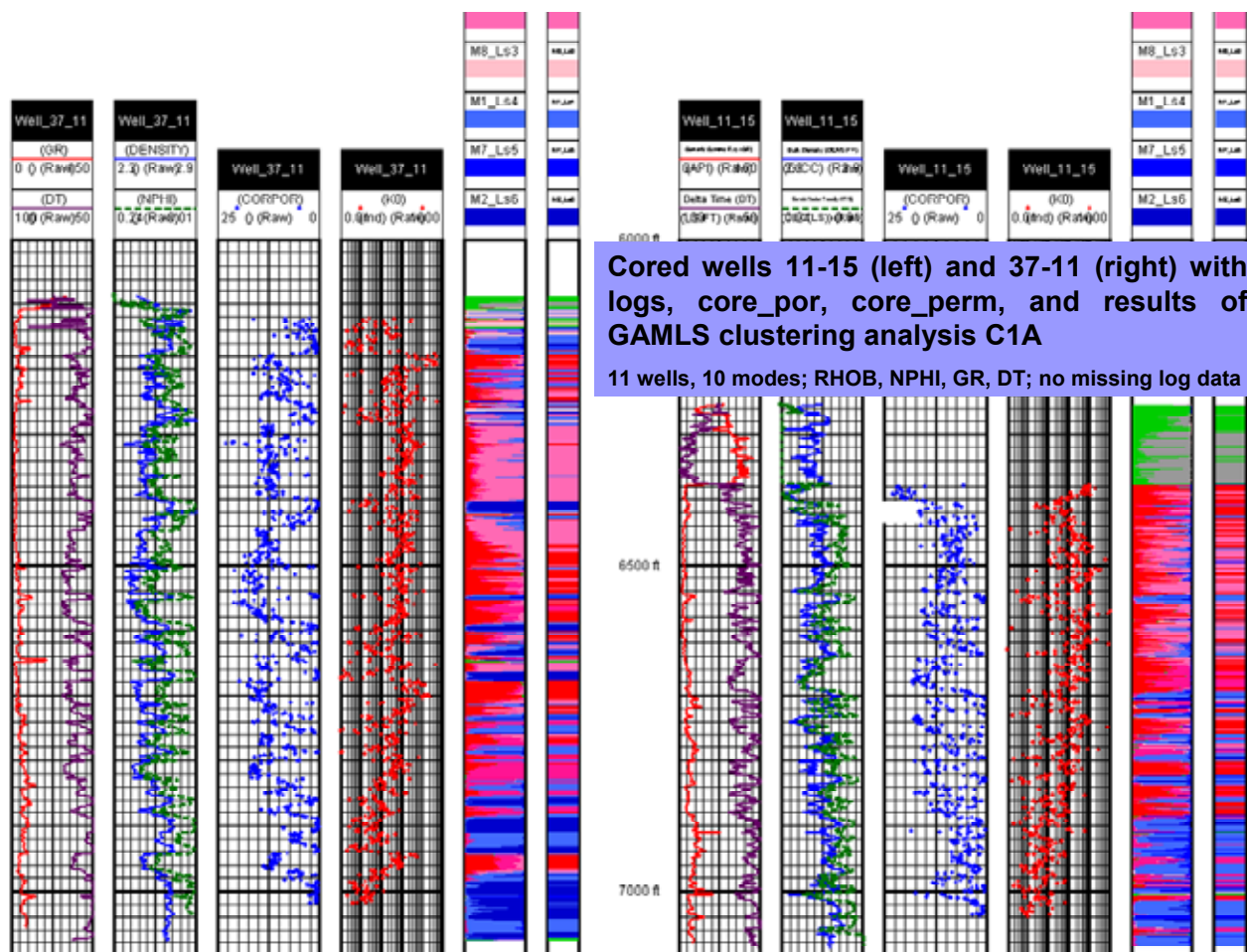


Figure 16: Tracks for Wells 11-15 and 37-11

We used a combination of the “ModeAssign” routine of GAMLS (which gives default lithologies for each mode), geological expertise, and published information to interpret all of the modes that comprise the Cisco and Canyon interval. All lithology assignments are subject to user-input logic criteria and can be overruled given core data or other information. In this cluster run (C1A), all modes in the Cisco and Canyon sections were interpreted to be limestone. Mode 10 (M10) was interpreted to have the best reservoir quality (RQ) because it had the highest mean NPHI value (0.13) which in turn suggested that it had the highest porosity. Clustering results which tabulate mean core porosity and permeability for each mode show that M10 generally had the best reservoir quality in the cored wells and is depicted in Figure 17. This mode, the apparent "best" limestone, was named (M10_LsBest). For plotting purposes, this mode was assigned a bright red color. In Figure 17, and based on apparent porosity from the mean NPHI values, the limestone modes were ordered from top to bottom in order of decreasing reservoir quality.

Of the other five limestone modes, three had relatively high apparent porosities based on the mean NPHI values (Figure 17): mode M3_LS1 (mode 3 – limestone 1) with mean NPHI = 0.12 and M5_LS2 and M8_LS3 with mean NPHI = 0.11. The other three limestone modes had

relatively low apparent porosities: M1, M7, and M2 with mean NPHI values of 0.04, 0.04, and 0.01, respectively. To discriminate among the "high porosity" modes and the "low porosity" modes, M3, M5, and M8 were colored deep pink, hot pink, and pink respectively, and the other limestone modes were shades of blue (the deeper the blue shade, the lower the porosity).

The modes assigned as siltstones (M9_Slt, in green) and shales (M6_Sh1, M4_Sh2, in grays) were mostly (but not entirely) above the carbonate sections and presumably mostly in the overlying Wolfcamp shale. The upsection change from carbonates to shale indicated by the clustering results is generally sharp and provides a good means to pick the contact between the Cisco and the Wolfcamp.

Mode	Name	Lith	Plot	Color	%	Hold	Mode	GD	BULK...	DT_L...	GR	NPHI
10	M10_LsBest	Ls	Yes	red	17.67	<input checked="" type="checkbox"/>	10	2.575	2.487	63.1	14.0	0.130
3	M3_Ls1	Ls	Yes	deep pink	12.63	<input checked="" type="checkbox"/>	3	2.699	2.513	61.7	20.9	0.119
5	M5_Ls2	Ls	Yes	hot pink	12.68	<input checked="" type="checkbox"/>	5	2.697	2.521	60.7	9.3	0.105
8	M8_Ls3	Ls	Yes	pink	1.33	<input checked="" type="checkbox"/>	8	2.667	2.626	55.1	23.8	0.105
1	M1_Ls4	Ls	Yes	royal blue	16.57	<input type="checkbox"/>	1	2.728	2.633	53.8	21.8	0.040
7	M7_Ls5	Ls	Yes	blue1	2.84	<input type="checkbox"/>	7	2.701	2.663	55.8	41.8	0.040
2	M2_Ls6	Ls	Yes	blue3	15.81	<input type="checkbox"/>	2	2.707	2.692	49.7	22.6	0.007
9	M9_Slt	Slt	Yes	green3	5.05	<input type="checkbox"/>	9	2.521	2.552	74.9	77.0	0.215
6	M6_Sh1	Sh	Yes	gray60	5.80	<input type="checkbox"/>	6	2.660	2.555	82.5	106.9	0.247
4	M4_Sh2	Sh	Yes	gray40	9.60	<input type="checkbox"/>	4	2.570	2.595	79.0	113.6	0.253

Figure 17: Cluster Run C1A – GAMLs “ModeAssign” Screen: Automatic Lithology Assignment Routine

7.1.3 Properties (RQ) as Functions of Modes

Figures 18 and 19 are representative crossplots and 3-dimensional plots respectively that show samples color coded to the “crisp” assignment. The ellipses are 4D ellipsoids projected onto the RHOB-NPHI plane, and are drawn around two standard deviations from the centre (mean values) of each ellipsoid. Samples are color coded according to their “crisp” (most likely) mode assignment. Figure 20 is a portion of a "fuzzy statistics" table generated as output from the clustering run. This table presents the mean and standard deviation of all curves in the well log data file for each clustering mode. The well log data file is initially the imported LAS file, but it can be supplemented by any derived data.

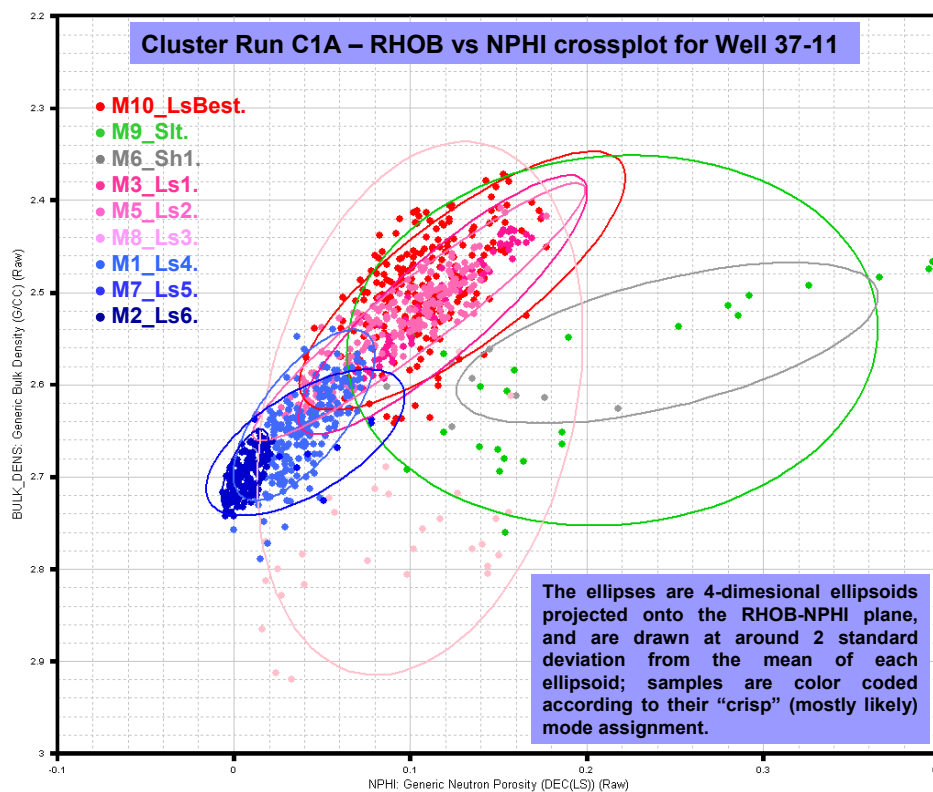


Figure 18: Cluster Run C1A: RHOB vs. NPHI Crossplot for Well 37-11

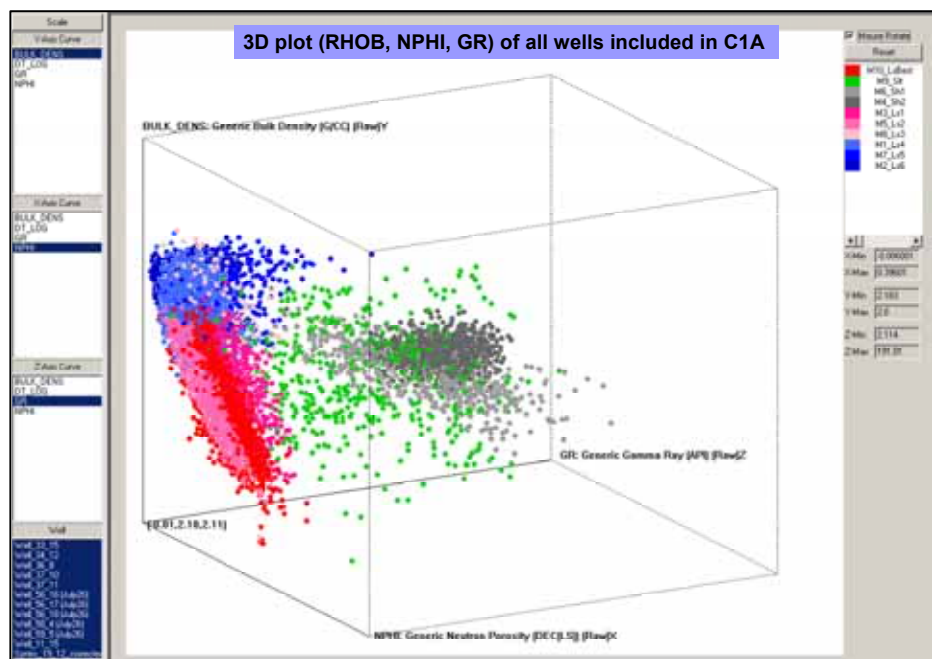


Figure 19: 3D Plot (RHOB, NPHI, GR) of All Wells Included in C1A

Cluster Run C1A – mean and standard deviation for selected variables (well log & core data) for each mode (rock type) in Well 37-11											
5. Well_37_11	M10_LsBest	M9_Slt	M6_Sh1	M4_Sh2	M3_Ls1	M5_Ls2	M8_Ls3	M1_Ls4	M7_Ls5	M2_Ls6	
* BULK_DENS (Raw) *	2.5030± 0.0595	2.5872± 0.0849	2.6131± 0.0389		2.5196± 0.0623	2.5332± 0.0537	2.7755± 0.0851	2.6485± 0.0452	2.6568± 0.0385	2.6997± 0.0228	
CALI (Raw)	8.9255± 0.1484	10.1070± 1.3352	9.5395± 0.7225		8.8905± 0.1449	8.9906± 0.1365	9.1655± 0.2254	8.9318± 0.2594	8.9824± 0.2907	8.8393± 0.2381	
DEEP_LAT (Li)	1.4749± 0.3176	2.2803± 0.6185	1.5810± 0.3645		1.4341± 0.3051	1.5546± 0.3143	2.5018± 0.6478	2.0042± 0.4215	1.9443± 0.5232	2.6050± 0.3027	
* DT_LOG (Raw) *	63.4054± 3.5051	72.1205± 9.1531	70.8988± 6.3237		62.0009± 4.0167	60.9194± 3.7950	54.5901± 3.6175	53.8908± 2.1889	57.7415± 3.0568	49.7427± 1.0683	
* GR (Raw) *	12.3696± 3.6071	52.5278± 31.7508	66.9891± 11.7120		17.8723± 6.3499	8.2781± 2.3817	11.4906± 8.1320	18.6463± 7.6785	25.7615± 13.0468	19.5859± 8.1553	
* NPHI (Raw) *	0.1048± 0.0281	0.1977± 0.1004	0.1588± 0.0393		0.1076± 0.0339	0.0932± 0.0331	0.0851± 0.0495	0.0406± 0.0171	0.0621± 0.0258	0.0073± 0.0079	
SHALL_LAT (Raw)	28.6904± 29.1576	59.3097± 63.2434	22.0235± 28.0379		28.1903± 34.1594	38.2356± 55.6812	128.7428± 91.2911	100.4734± 100.201	95.3825± 83.8905	385.0851± 271.927	
AI_GAMLS (Raw)	39.6463± 3.1336	36.4719± 4.9281	37.1830± 3.7138		40.8762± 3.7135	41.8065± 3.6025	51.0995± 4.1702	49.2544± 2.7129	46.1587± 2.8442	54.3057± 1.5016	
K0 (Li)	1.0389± 0.6878	-0.0148± 0.8122	0.5583± 0.5059		0.6703± 0.6106	0.8940± 0.6777	0.1766± 0.8327	-0.0974± 0.8117	0.1084± 1.1788	-0.7066± 0.6618	
K90 (Li)	1.0420± 0.6404	0.3146± 0.9160	0.4296± 0.5035		0.6827± 0.5955	0.9229± 0.6159	0.4664± 0.7780	-0.0042± 0.8490	0.2001± 1.3392	-0.7004± 0.7187	
KV (Li)	0.7005± 0.8737	-0.1379± 1.0625	-0.3468± 0.8144		0.2338± 0.7766	0.4789± 0.9201	-0.2563± 1.0407	-0.7655± 0.7996	-0.4704± 0.9876	-1.5996± 0.5366	
CORPOR (Mu)	0.1267± 0.0358	0.0558± 0.0372	0.0516± 0.0111		0.1185± 0.0367	0.1079± 0.0379	0.0563± 0.0325	0.0503± 0.0287	0.0563± 0.0334	0.0135± 0.0134	

Cluster Run C1A – enlargement (from above) of the 3 modes with the “best” reservoir quality; NOTE: the perm values listed are logarithm (base 10) of core perm											
5. Well_37_11	M10_LsBest	M3_Ls1	M5_Ls2								
%	2.5030± 0.0595	2.5196± 0.0623	2.5332± 0.0537								
DT (Raw)	8.9255± 0.1484	8.8905± 0.1449	8.9906± 0.1365								
* BULK_DENS (Raw) *	1.4749± 0.3176	1.4341± 0.3051	1.5546± 0.3143								
CALI (Raw)	63.4054± 3.5051	62.0009± 4.0167	60.9194± 3.7950								
DEEP_LAT (Li)	12.3696± 3.6071	17.8723± 6.3499	8.2781± 2.3817								
* GR (Raw) *	0.1048± 0.0281	0.1076± 0.0339	0.0932± 0.0331								
* NPHI (Raw) *	28.6904± 29.1576	28.1903± 34.1594	38.2356± 55.6812								
AI_GAMLS (Raw)	39.6463± 3.1336	40.8762± 3.7135	41.8065± 3.6025								
K0 (Li)	1.0389± 0.6878	0.6703± 0.6106	0.8940± 0.6777								
K90 (Li)	1.0420± 0.6404	0.6827± 0.5955	0.9229± 0.6159								
KV (Li)	0.7005± 0.8737	0.2338± 0.7766	0.4789± 0.9201								
CORPOR (Mu)	0.1267± 0.0358	0.1185± 0.0367	0.1079± 0.0379								

Figure 20: Cluster Run C1A: Mean and Standard Deviation for Selected Variables

Figure 21 shows a crossplot of (logarithm of) core permeability (K0) versus core porosity at cored well 37-11. The samples are color coded according to “crisp” mode assignment. This is a graphical representation of how the defined modes are related to reservoir quality. The graph can be supplemented with the “fuzzy statistics” table (Figure 20) which gives, for instance, the mean porosity and permeability for the mode with the best reservoir quality (M10_LsBest) to be 12.7% and 1.04 (log₁₀), respectively.

Figures 22 and 23 are depth plots that have probability assignments on the horizontal axis. Except for wells 11-15 and 19-12, which are to the N and NW out of the central study area, the wells were arranged left-to-right by structurally “correlating” wells on a prominent limestone bed near the bottom of the section (Figure 22) and on a prominent limestone bed near the top of the section (Figure 23).

Regardless of how the wells are arranged in cross sections, it is clear that the upper half of the carbonate section is generally better reservoir quality than the lower half, and within the upper half there is a zone of “best” reservoir quality approximately between 6,400 and 6,600 feet (MD).

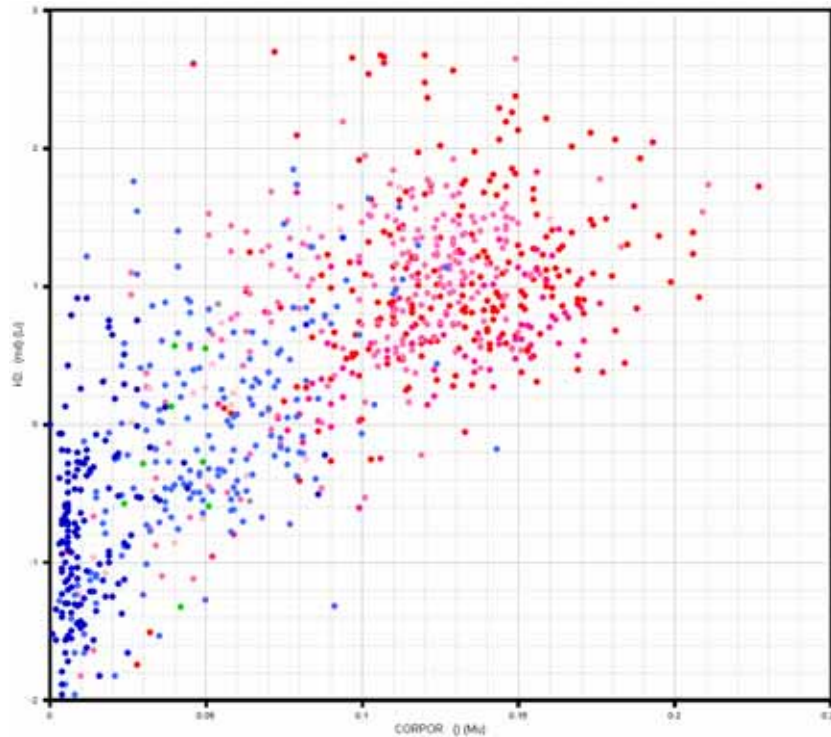


Figure 21: Cluster C1A - Well 37-11: Core Permeability (\log_{10}) vs. Core Porosity

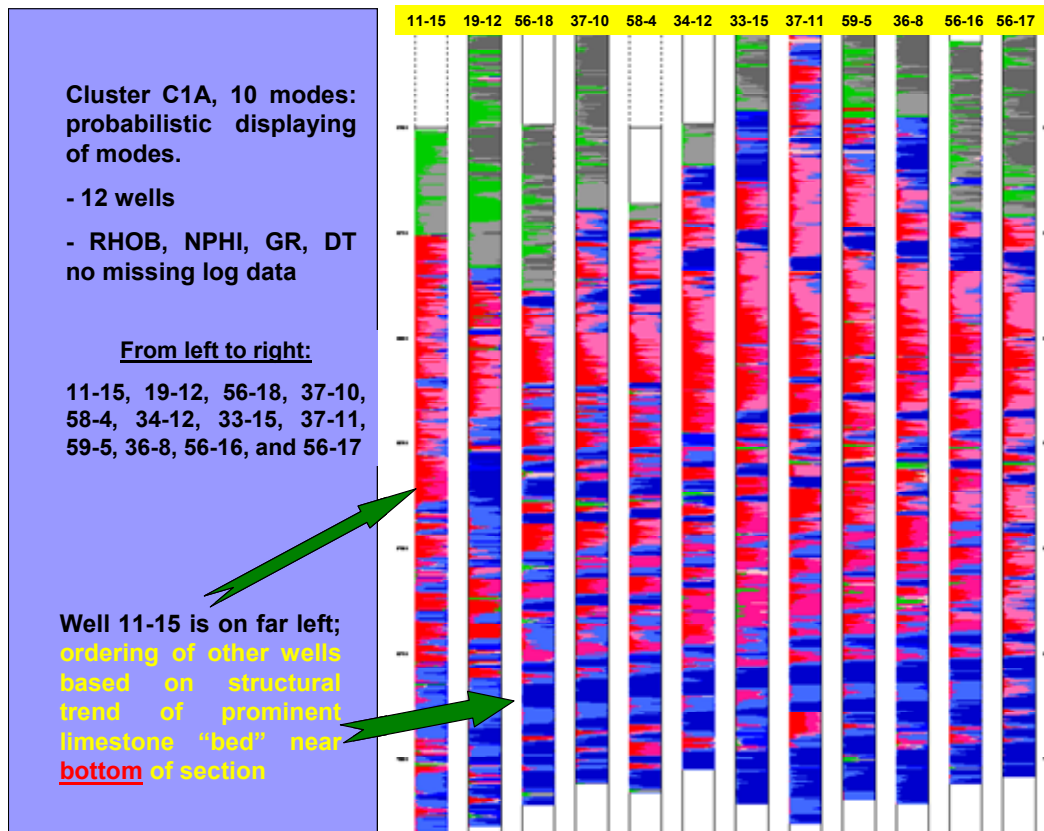


Figure 22: Cluster C1A: Probabilistic Representation of Modes at Each Depth is Displayed

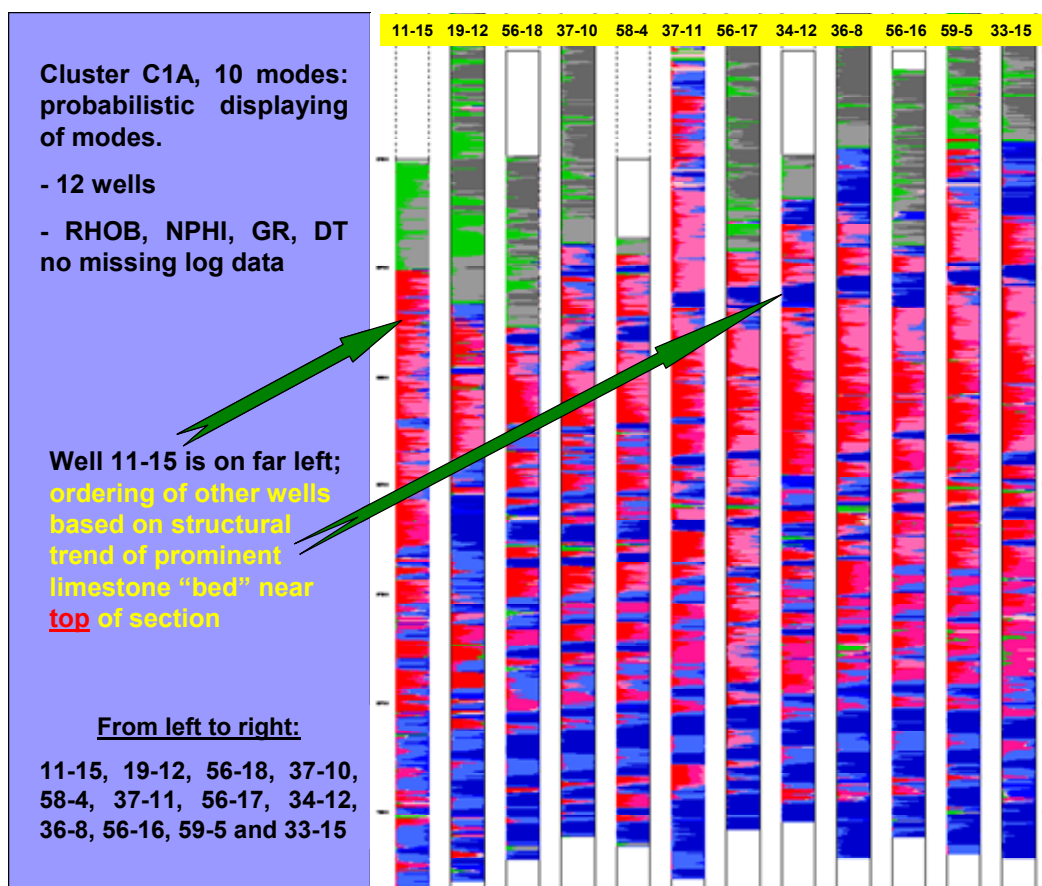


Figure 23: Cluster C1A: Probabilistic Representation of Modes at Each Depth

7.1.4 Modes, Beds, and Flow Units

Figure 24 shows C1A clustering results (tracks 1 and 2) and raw well logs (track 5) for non-cored well 33-15. The permeability and porosity values plotted in tracks 3 and 4, respectively, are predicted values using the "P&P via AI" method. This figure is an example of deriving flow units from one clustering run, deriving predicted porosity and permeability from another (series) of clustering runs, and then relating the flow units to these two RQ parameters. This was done for all of the non-cored wells in C1A.

The highest porosities and permeabilities are in the zones indicated by the bright red Mode 10 color. It is also apparent that the four modes with highest RQ (red and pink colors) have transitional contacts with one another but have fairly sharp contacts with the modes with lower RQ (blue colors). That is, the reds and pinks could be considered as one series of closely related flow units and the blues as a second series of closely related flow units.

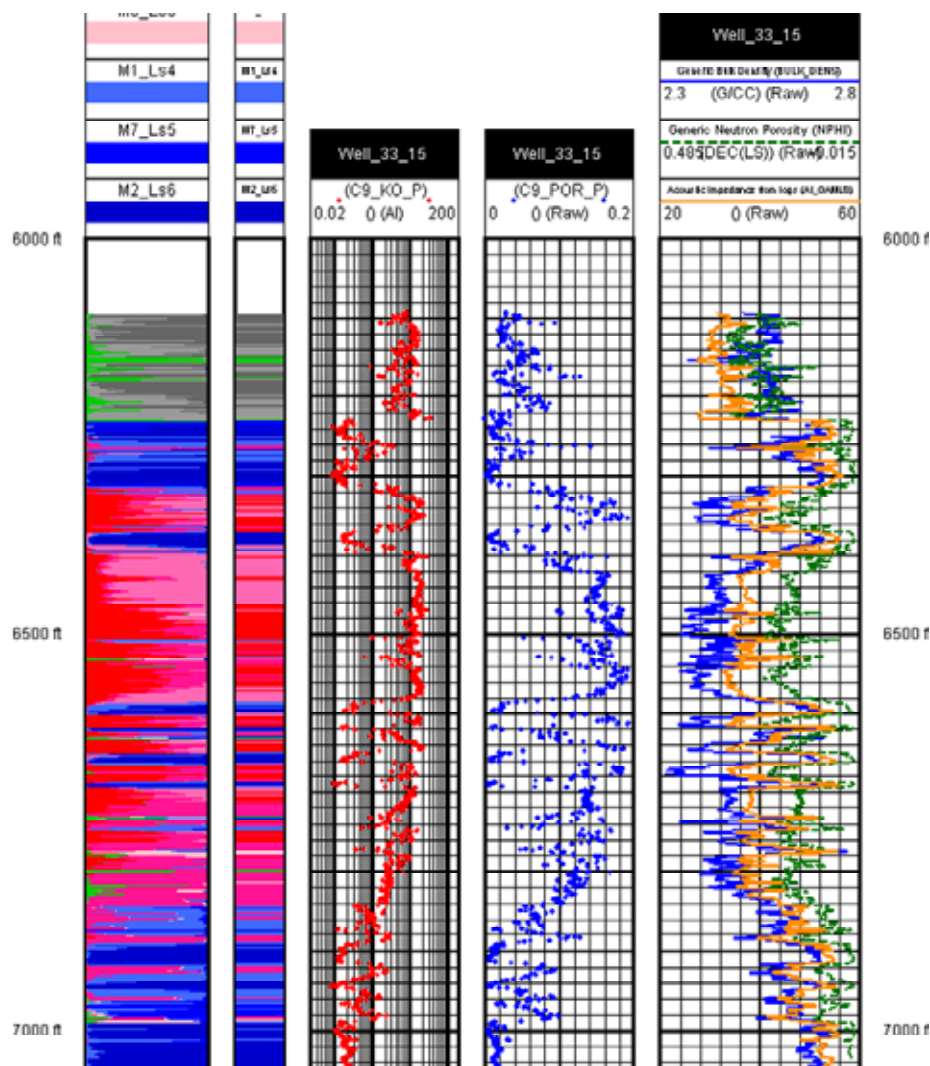


Figure 24: Tracks for the Non Cored Well 33-15

The flow unit realization provided by C1A into seven flow units (modes) can be simplified by eliminating all but one of the red and pink modes and all but one of the blue modes, so that there is only one red mode and only one blue mode. This is done by using the "drop modes" feature in a clustering run that has already converged and then reiterating to convergence. This was performed for C1A and the results are shown in Figures 25 and 26. Six modes were dropped retaining only the most abundant good RQ limestone (dropping 3 modes), the most abundant poor quality RQ (dropping 2 more modes) and keeping the siltstone and 2 shale modes (which are mostly in the overlying shale). Figure 25 compares the 10-mode CMP plots with the 4-mode CMP plots for the 3 cored wells. Figure 26 shows the CMP plots for the 4-mode realization for all 12 wells. Despite the wells alignment does not follow any directional orientation; similar patterns in the spatial distribution of the reformulated modes can be appreciated in most of the wells.

Another way to simplify the flow unit scenario shown in Figure 24 is to eliminate thin beds by using a "bed thickness filter" (BTF). Figures 27-30 are depth plots that show "crisp" assignments. Figure 27 shows beds from Cluster C1A without any BTF. In other words, the original "crisp" modes have been displayed. Wells were arranged left-to-right in same order as in Figure 20. In particular, 201 beds were discriminated in Well 37-11.

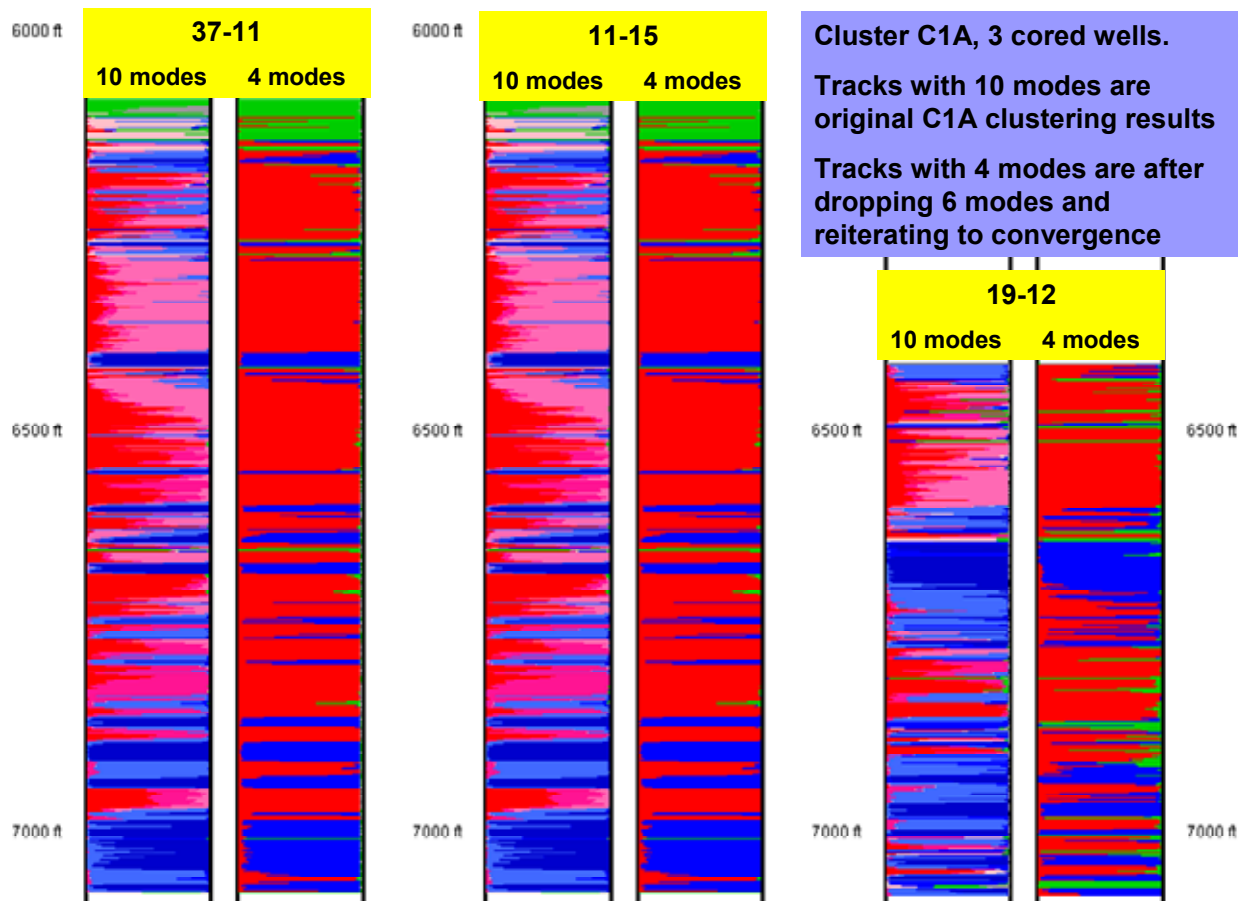


Figure 25: Cluster C1A - Original 10 Mode Results vs. 4 Mode Results

Figures 28, 29, and 30 show beds after applying BTF with minimum bed thicknesses of 5, 10, and 20 feet, respectively. Filtering can be used to aid correlation of zones of high reservoir quality and also as a type of "curve blocking" or "upscaling" for reservoir modeling (Eslinger, 2007).

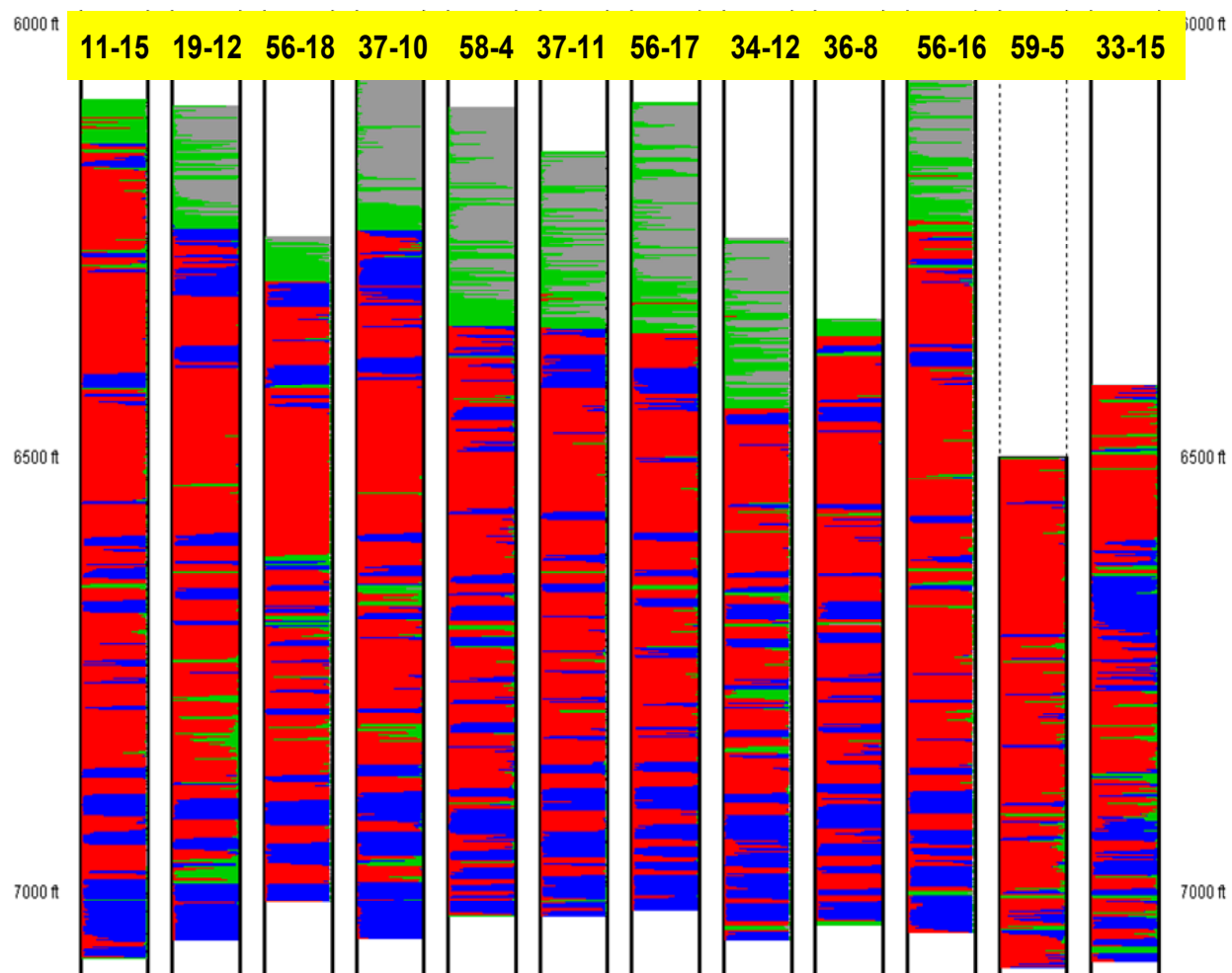


Figure 26: Cluster C1A- 12 Wells with 4 Modes after Dropping Application

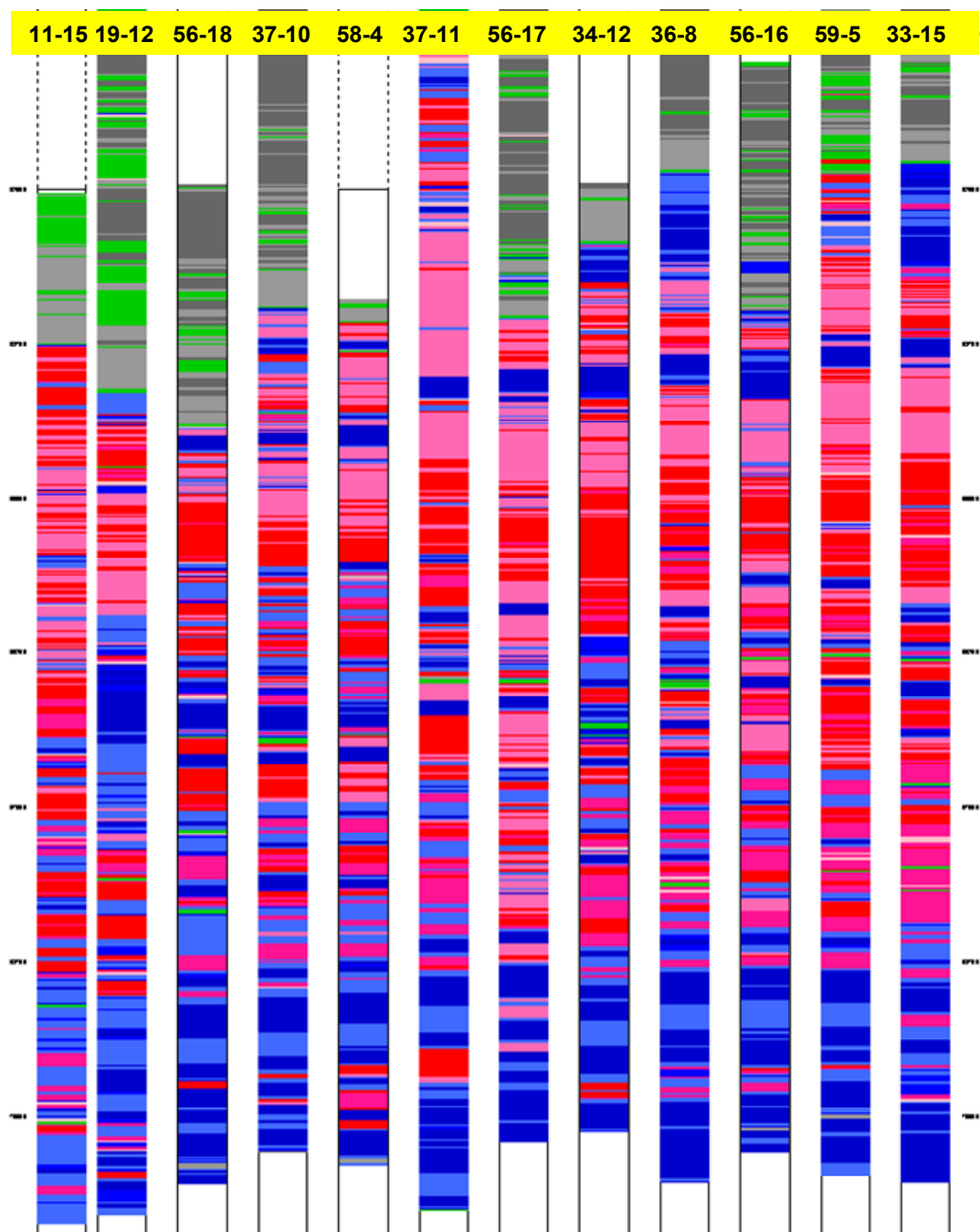


Figure 27: Cluster C1A - 12 Wells with “Crisp” Modes at each Depth

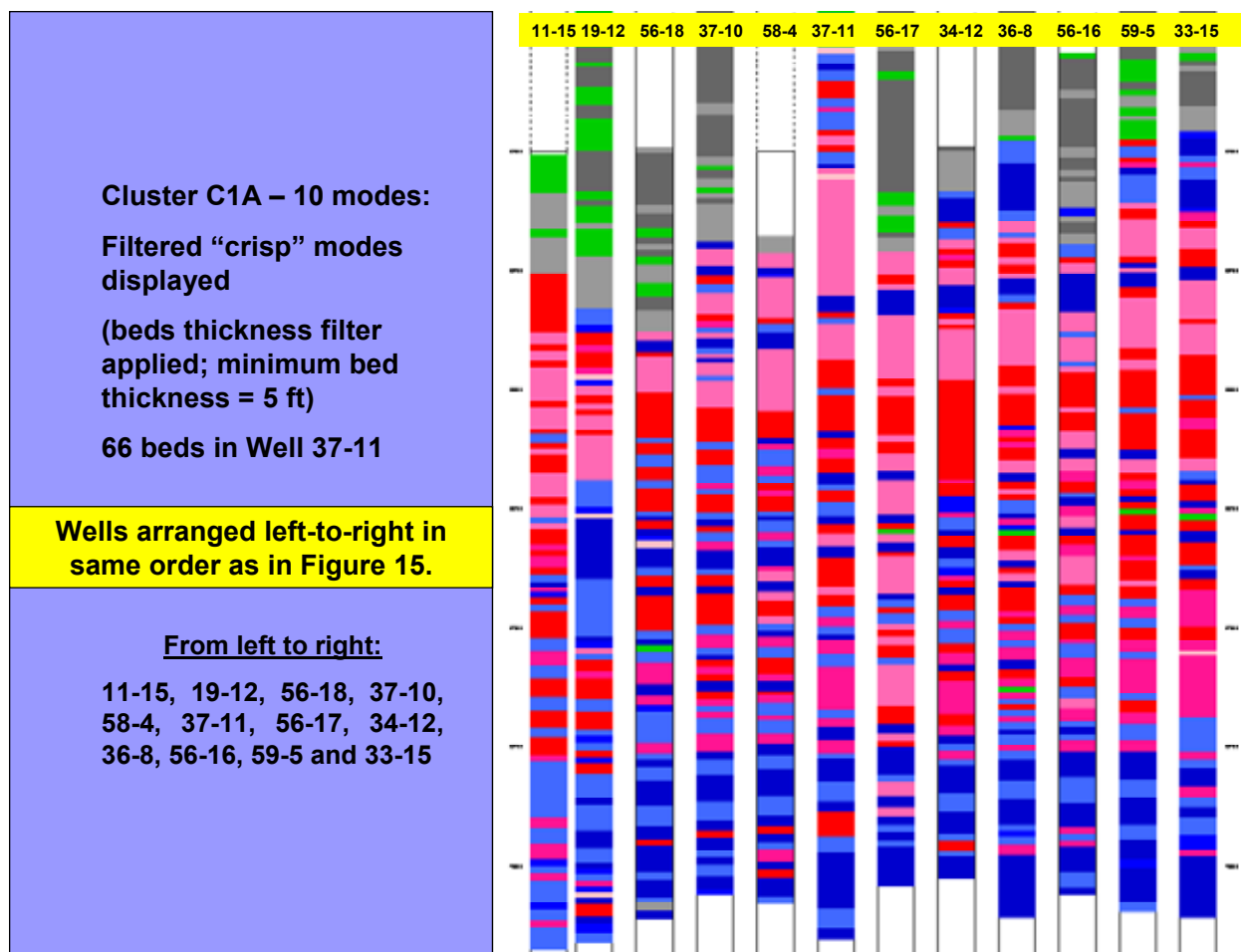


Figure 28: Cluster C1A -12 Wells Using BTF with a Minimum Bed Thickness = 5 Ft

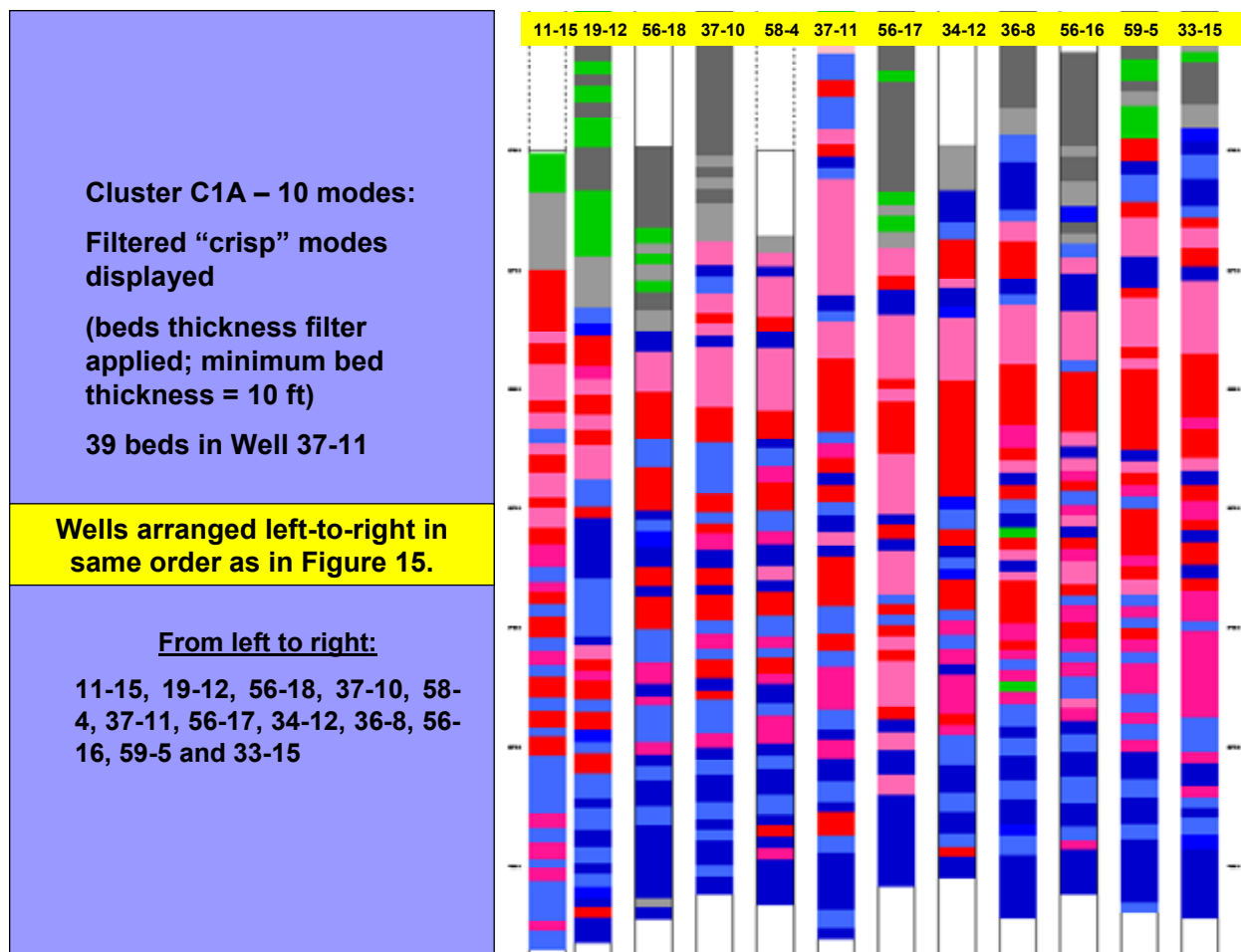


Figure 29: Cluster C1A -12 Wells Using BTF with a Minimum Bed Thickness = 10 Ft

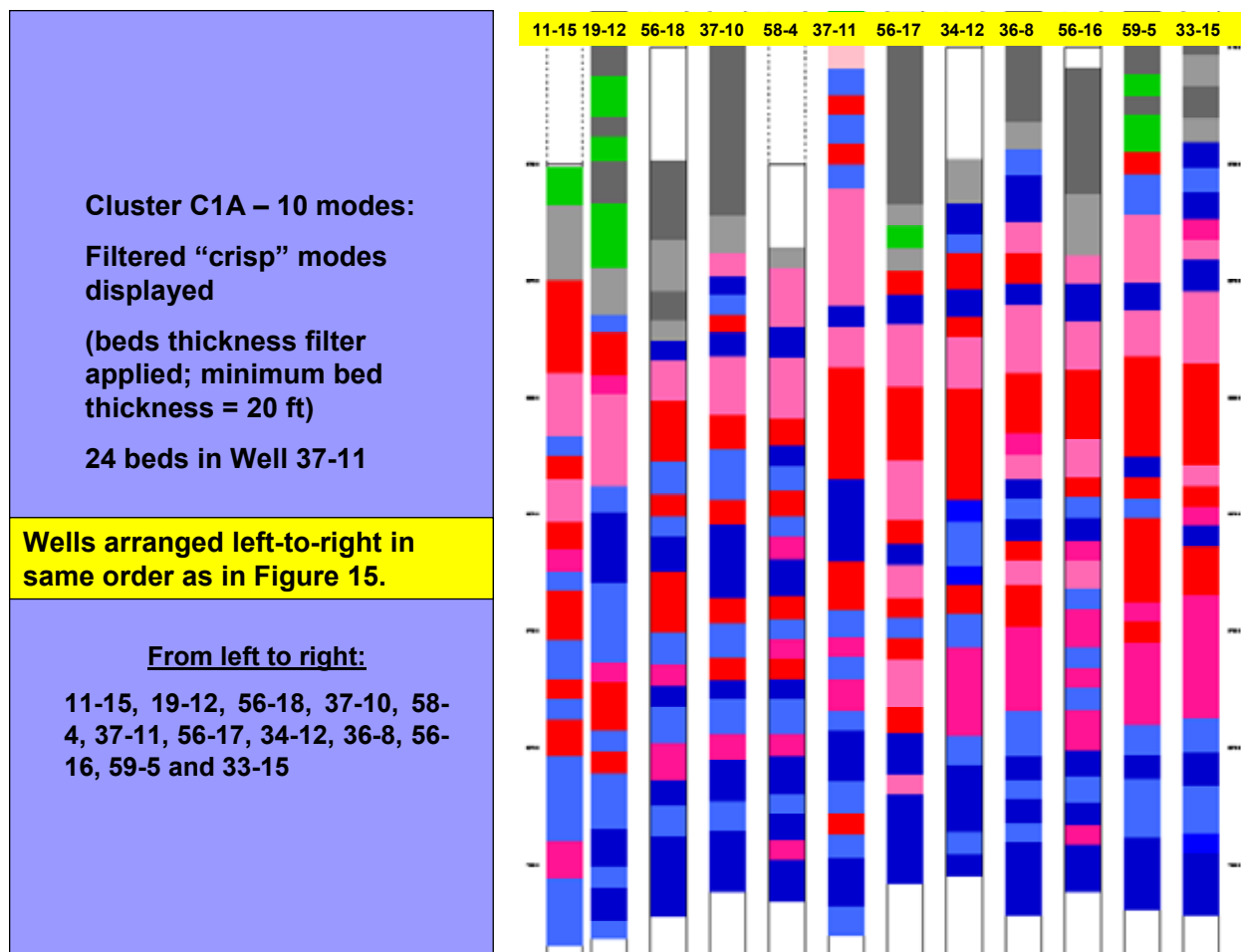


Figure 30: Cluster C1A -12 Wells Using BTF with a Minimum Bed Thickness = 20 Ft

With any clustering analysis, two tables (“fuzzy” statistics and “crisp” statistics) for each well in the clustering run output the mean and standard deviation of all curves in the well log “curves” file for each mode, the percentage of samples assigned to each mode, the total footage assigned to each mode, the number of beds defined by each mode, the mean bed thickness for each mode, and the mean probability assignment of the “crisp” mode for each mode. This information is useful for evaluating the overall RQ of each mode. Figures 31-33 show examples of the data that is generated in these tables. The tables shown give data for all 12 C1A wells taken together and also for non-cored well 33-15.

Figure 31 shows the bed information for each of the seven limestone beds in C1A with no BTF. Data for well 33-15 is at the top and data for all twelve wells taken together is at the bottom. Figure 32 shows the same information but after a bed thickness filter was applied to eliminate all beds less than 10 feet thick.

C1A, well 33-15				
bed and bed thickness data for limestone flow units (no bed thickness filtering)				
1. Well_33_15	M10_LsBest	M3_Ls1	M5_Ls2	
Beds	39	35	21	
Thickness (ft)	182.00	167.00	113.00	
Avg Thick (ft)	4.67	4.77	5.38	
%	19.0%	17.4%	11.8%	
Avg Prob	0.6544	0.7760	0.6737	
1. Well_33_15	M8_Ls3	M1_Ls4	M7_Ls5	M2_Ls6
Beds	4	40	11	24
Thickness (ft)	11.00	126.00	32.00	185.00
Avg Thick (ft)	2.75	3.15	2.91	7.71
%	1.1%	13.1%	3.3%	19.3%
Avg Prob	0.6171	0.7709	0.7871	0.9132
C1A, all 12 wells				
bed and bed thickness data for limestone flow units (no bed thickness filtering)				
All Wells	M10_LsBest	M3_Ls1	M5_Ls2	
Beds	483	277	348	
Thickness (ft)	1993.99	1028.01	1646.00	
Avg Thick (ft)	4.13	3.71	4.73	
%	18.2%	9.4%	15.0%	
Avg Prob	0.6870	0.7200	0.6764	
All Wells	M8_Ls3	M1_Ls4	M7_Ls5	M2_Ls6
Beds	63	518	119	250
Thickness (ft)	115.00	1914.00	258.00	1812.00
Avg Thick (ft)	1.83	3.69	2.17	7.25
%	1.0%	17.4%	2.3%	16.5%
Avg Prob	0.7523	0.7664	0.7866	0.9063

Figure 31: Cluster C1A - Bed Information for each of the 7 Limestone Beds

C1A, well 33-15				
bed and bed thickness data for limestone flow units (bed thickness filter applied – minimum thickness = 10 feet)				
5. Well_37_11	M10_LsBest	M3_Ls1	M5_Ls2	
Beds	8	3	4	
Thickness (ft)	236.00	81.00	189.00	
Avg Thick (ft)	29.50	27.00	47.25	
%	23.9%	8.2%	19.1%	
Avg Prob	0.5824	0.5927	0.6946	
5. Well_37_11	M8_Ls3	M1_Ls4	M7_Ls5	M2_Ls6
Beds	1	12		10
Thickness (ft)	41.00	236.00		186.00
Avg Thick (ft)	41.00	19.67		18.60
%	4.1%	23.9%	0.0%	18.8%
Avg Prob	0.4974	0.5669		0.7590
C1A, all 12 wells				
bed and bed thickness data for limestone flow units (bed thickness filter applied – minimum thickness = 10 feet)				
All Wells	M10_LsBest	M3_Ls1	M5_Ls2	
Beds	89	54	54	
Thickness (ft)	2172.00	947.01	1596.00	
Avg Thick (ft)	24.40	17.54	29.56	
%	19.8%	8.6%	14.5%	
Avg Prob	0.5645	0.5898	0.5917	

Figure 32: Cluster C1A - Bed Information, Minimum Bed Thickness = 10 Ft

7.1.5 Porosity and Permeability of Non-Cored Wells Assigned to Modes from C1A

Figure 33 shows the mean and standard deviation for porosity and \log_{10} of permeability (K0) for each mode for non-cored well 33-15. Similar data was produced for all of the non-cored wells, and are from the predictions made using the "P&P via AI" method.

Figure 34 shows predicted permeability (\log_{10}) versus predicted porosity for non-cored well 33-15. Again, the predicted permeability and porosity values are from the "P&P_AI" method. The samples are color coded by mode from C1A. Correlations by mode are given in the legend. The shale and siltstone samples (grays and green) are almost all from the overlying shale. The main point here is that RQ in terms of porosity and permeability varies by mode and the modeled (predicted) porosity and permeability have fairly good correlations.

C1A, well 33-15

mean & standard deviation of predicted porosity and predicted permeability (log 10) for each of the 7 limestone facies

the predicted porosity and permeability values were generated in cluster run C9 – the flow units were defined in cluster run C1A

1. Well_33_15	M10_LsBest		M3_Ls1		M5_Ls2	
C9_KO_P (Raw)	1.0105±	0.2344	0.6020±	0.3641	1.0341±	0.3532
C9_POR_P (Raw)	0.1493±	0.0285	0.1229±	0.0329	0.1305±	0.0358

1. Well_33_15	M8_Ls3		M1_Ls4	
C9_KO_P (Raw)	0.1712±	0.2526	-0.1559±	0.3045
C9_POR_P (Raw)	0.0860±	0.0163	0.0511±	0.0204

1. Well_33_15	M7_Ls5		M2_Ls6	
C9_KO_P (Raw)	-0.4245±	0.3292	-0.7421±	0.1584
C9_POR_P (Raw)	0.0463±	0.0255	0.0138±	0.0057

Figure 33: Well 33-15: Mean & Standard Deviation of Predicted Porosity and Predicted Permeability (Log₁₀)

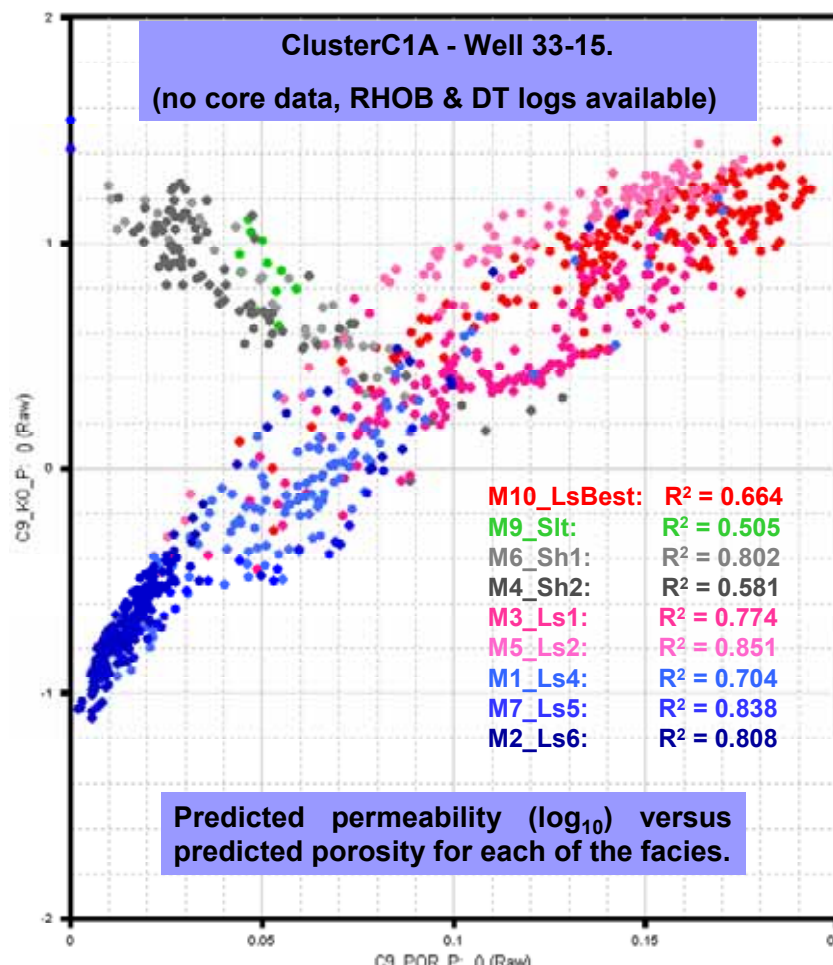


Figure 34: Well 33-15 Predicted Permeability (Log₁₀) vs. Predicted Porosity for each Mode

7.1.6 Variation of K0, K90 and Kv with Flow Units

According Raines and Helmes⁶, the cored well 37-11 “...targeted the typical Canyon section and the higher portions of the Cisco formation, including the highest area in the field.” Well 11-15 “...targeted a region of extreme porosity variation in the Canyon interval ...”, and well 19-12 “...targeted the thickest portion of the Cisco formation, where debris and bioherm deposits dominate the section...”.

The variation of actual values of K0, K90, and Kv were examined as a function of the C1A flow units. Figures 35-37 show K90 versus K0 for cored wells 37-11, 11-15, and 19-12, respectively (the base 10 logarithm of these parameters). Rather than reporting the traditional horizontal permeabilities K_{\max} and K_{\min} , K0 and K90 were measured relative to the master orientation line in order to identify any directional bias in the horizontal permeability. These figures color coded by modes of cluster C1A, show a lack of significant variation between K0 and K90 (\log_{10}) which suggests that there is not a noticeable directional bias of the data (at least in these cored well locations).

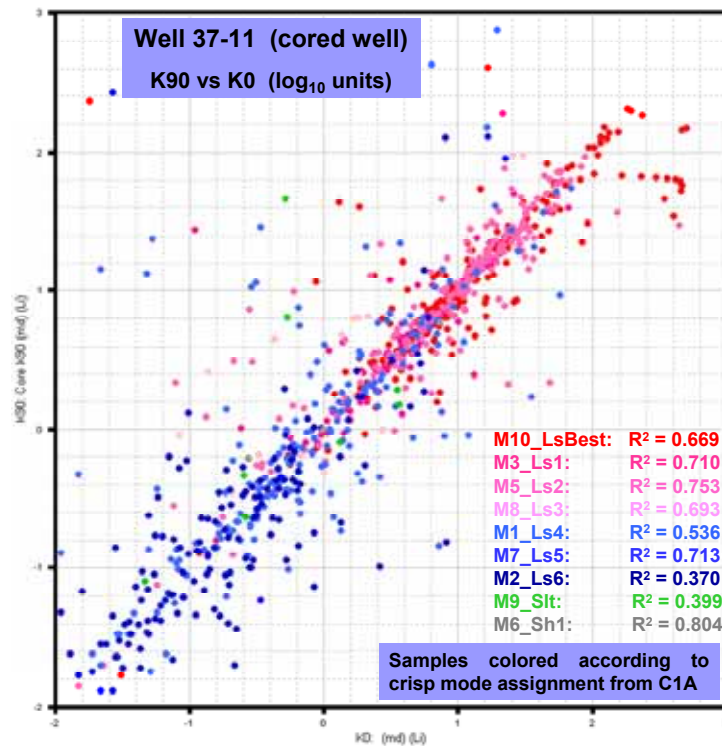


Figure 35: $\log_{10}(K90)$ vs. $\log_{10}(K0)$ for Well 37-11

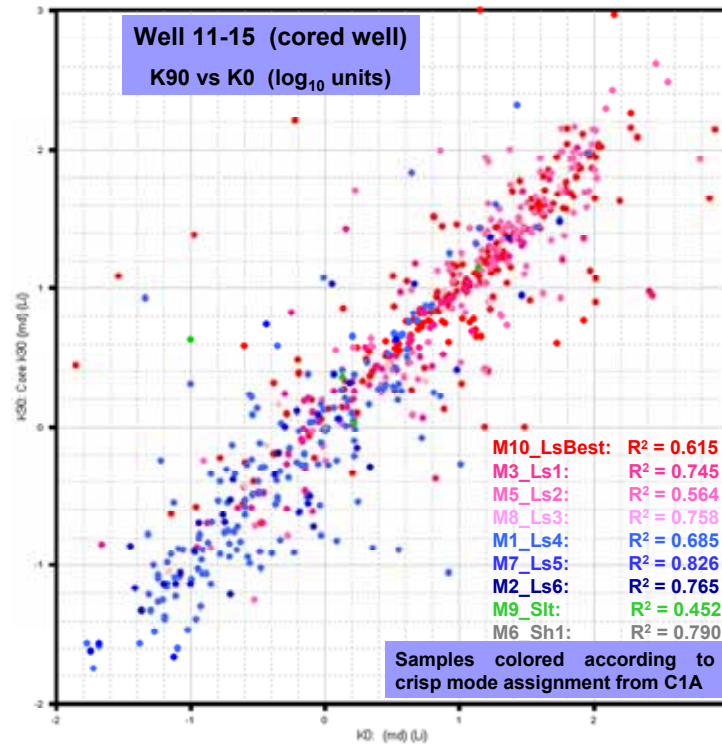


Figure 36: Log₁₀(K90) vs. Log₁₀(K0) for Well 11-15

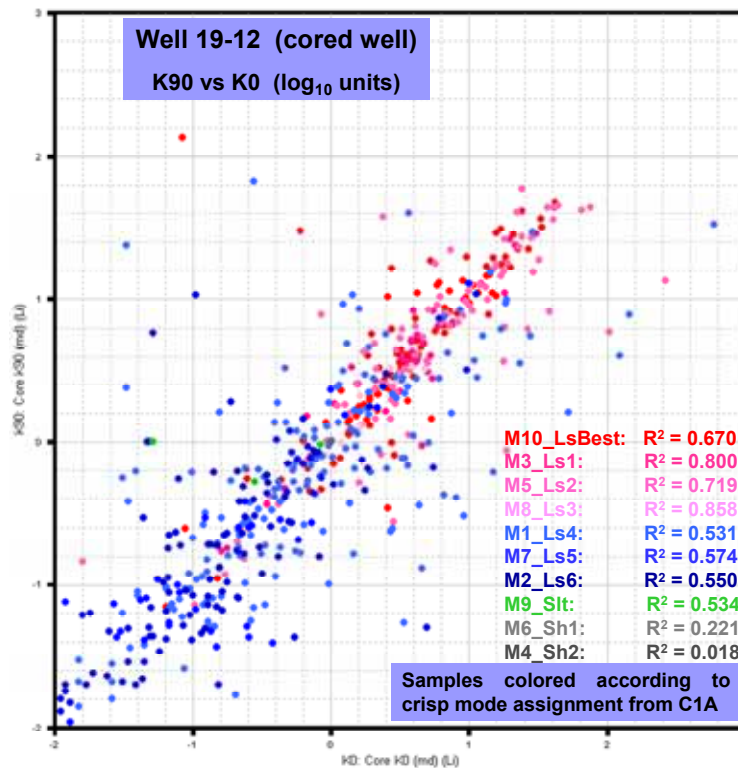


Figure 37: Log₁₀(K90) vs. Log₁₀(K0) for Well 19-12

Figure 38 shows the crossplot for well 11-15 of $\log_{10}(K0)$ versus the algebraic difference between $\log_{10}(K0)$ values and $\log_{10}(K90)$ values, i.e., $\log_{10}(K0) - \log_{10}(K90)$ (which is equivalent to $\log_{10}(K0/K90)$). This difference between $K0$ and $K90$ is slightly smaller for the flow units with higher reservoir quality; that is, the better RQ units are slightly more homogeneous. Figures 39-41 show the logarithm on base 10 of Kv versus the logarithm on base 10 of $K0$ for cored wells 37-11, 11-15, and 19-12, respectively. All these figures present in common that most of the ordered pairs $(\log_{10}(K0), \log_{10}(Kv))$ are concentrated below the concordance straight line ($y = x$) which illustrates the fact that most of Kv measurements are less than their corresponding $K0$ measurements independently of the particular mode. Certainly, in the better limestone modes (red and pinks), $K0$ and Kv tend to be better aligned and more concentrated around the concordance line, which is indicative of more similarity between $\log_{10}(K0)$ and $\log_{10}(Kv)$ values, and less reservoir anisotropy in these better RQ units. Similar arguments can be developed for a comparison between $K90$ and Kv (logarithms).

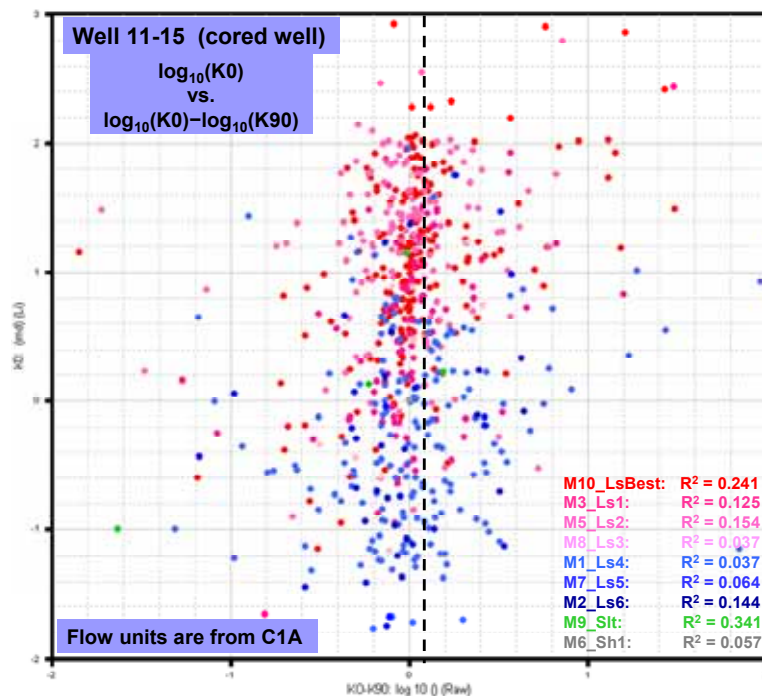


Figure 38: $\log_{10}(K0)$ vs. $\log_{10}(K0/K90)$ for Well 11-15

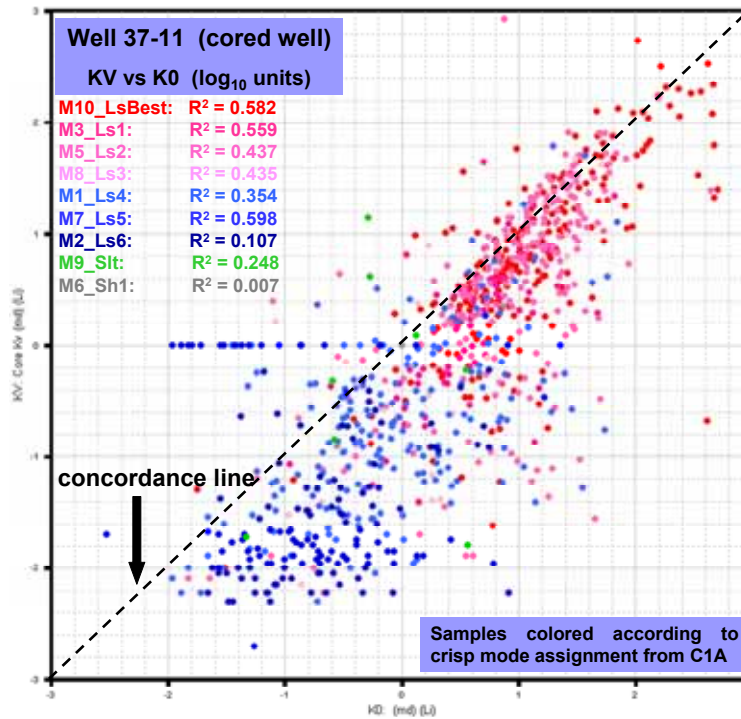


Figure 39: Kv vs. K0 (Logarithm) for Well 37-11

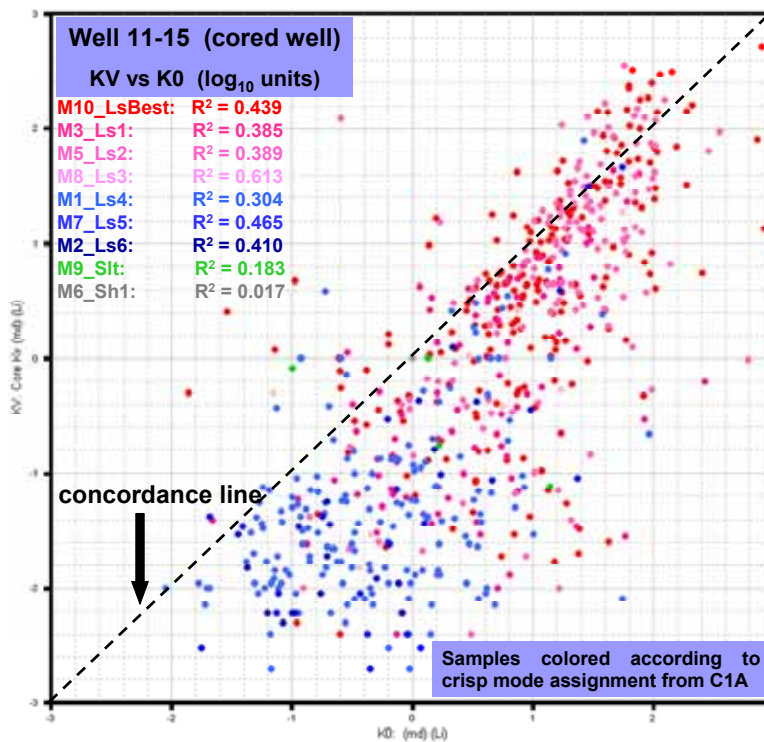


Figure 40: Kv vs. K0 (logarithm) for Well 11-15

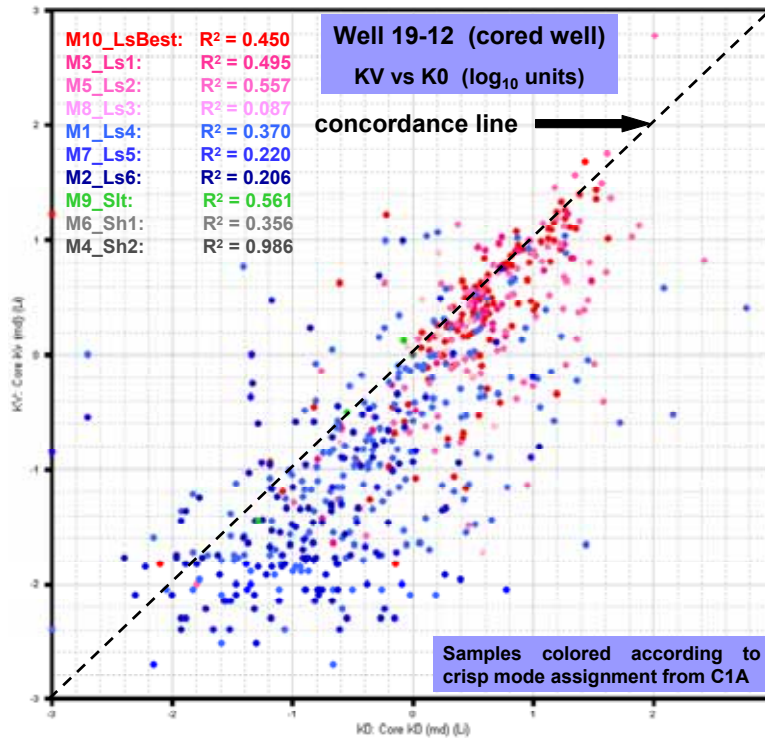


Figure 41: Kv vs. K0 (logarithm) for Well 19-12

Figure 42 shows the frequency distribution of $\log_{10}(K_v/K_h)$ for cored well 37-11, where K_h is the arithmetic mean of K_0 and K_{90} . Isotropic permeability exists when $K_v/K_h = 1.0$ (or when the logarithm of this ratio is 0.0). For most samples, this ratio between logarithms (base 10) of K_v and K_h is negative, indicating that most samples have K_v less than K_h . This is not unusual, and is typically interpreted to mean that barriers to vertical flow exist. Increasingly lower K_v/K_h values indicate greater number of flow barriers and increasing permeability anisotropy. Note that the modes with poorer RQ (blue colors) generally have lower K_v/K_h ratios than the modes with higher RQ (red and pink modes).

Figure 43 is another way of viewing the same relationship and shows $\log_{10}(K_h)$ versus $\log_{10}(K_v/K_h)$ for well 37-11. There are many more samples of low RQ than high RQ, or in other words, the poorer RQ units are more anisotropic than the better RQ units in the comparison of vertical permeability against horizontal permeability.

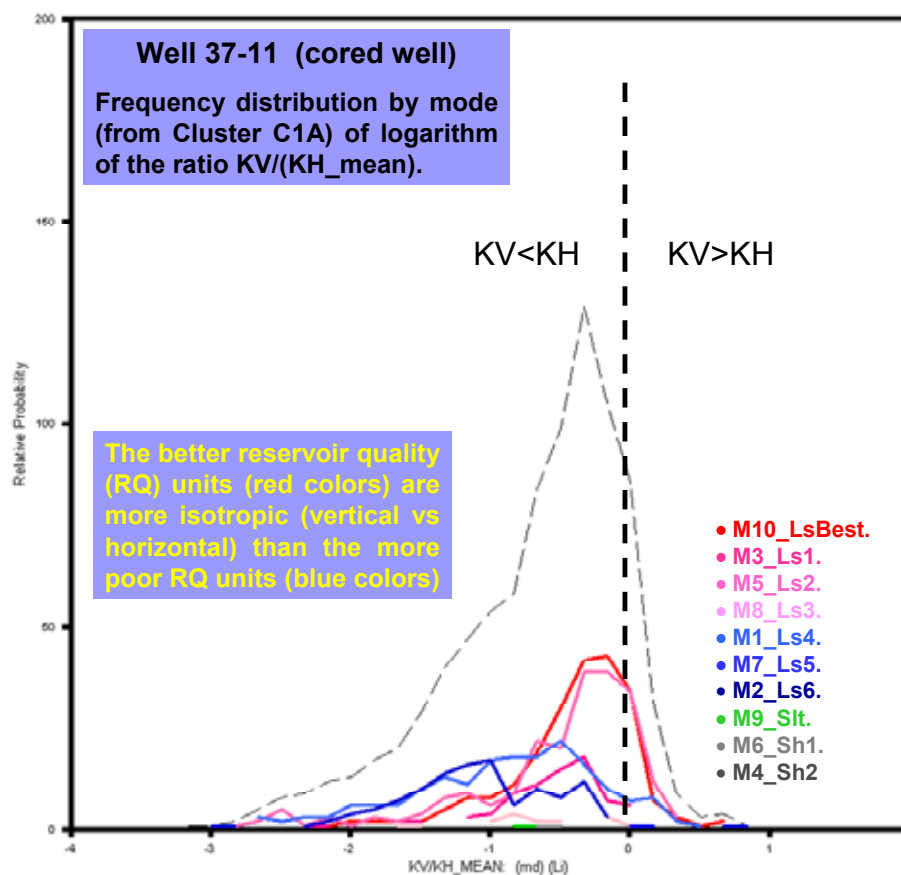


Figure 42: Well 37-11: Frequency Distribution by Mode (Cluster C1A) of $\text{Log}_{10}(K_v/K_h)$

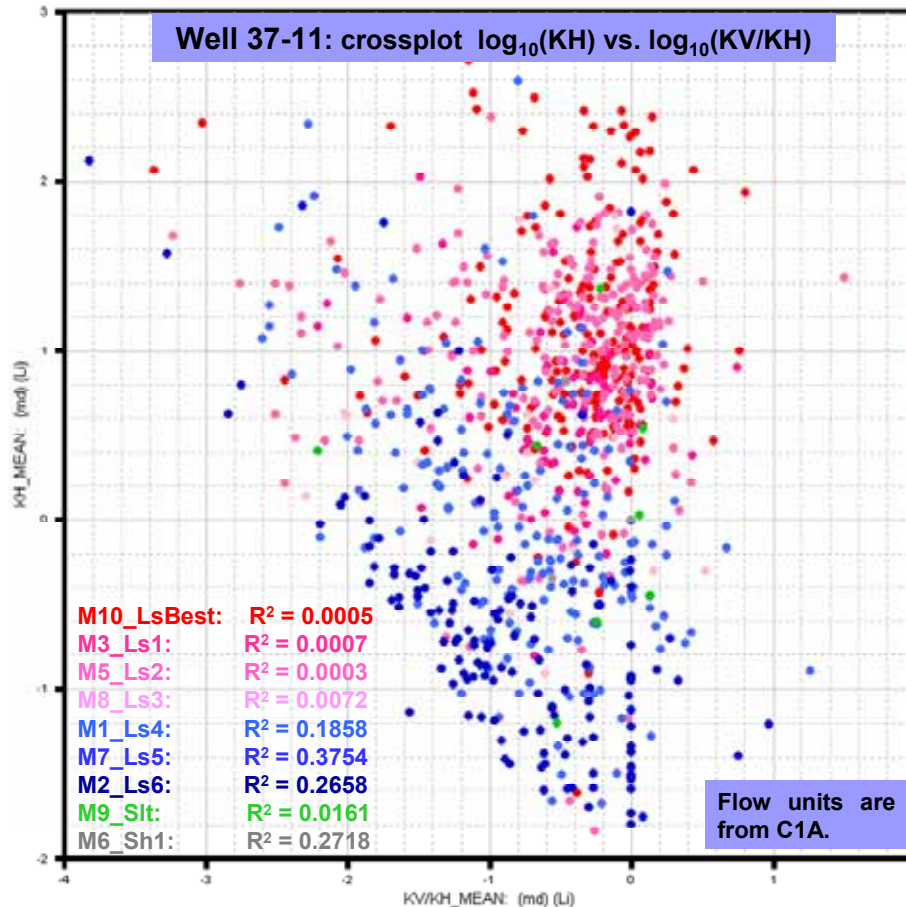


Figure 43: $\log_{10}(Kh)$ vs. $\log_{10}(Kv/Kh)$ Cored Well 37-11

7.2 C1B - Clustering with a 4-Curve Log Suite Realization Using 12 Wells with RHOB and DT Curves and 19 Wells without RHOB and DT

Many of the available wells did not have a RHOB or DT log. Nineteen of these wells were selected and a cluster run (C1B) was made to estimate RHOB and DT profiles. The cluster run also included the 12 wells used in C1A, and the same four variables were used as in C1A (RHOB, NPFI, GR, and DT). Curves containing “null” or missing values were generated as RHOB and DT tracks for the 19 wells. A “null” value is essentially a blank value, not a zero value, and is used for any log curve samples where data is missing or obviously incorrect. During clustering these null values are replaced by estimated values. Thus, this is a “filling in” or “prediction” process. These curves were inserted into the run setup and then during the clustering process, estimated values were computed and these values then replaced the null values. The porosity and permeability (P&P) profiles determined using the RQ via AI were then applied to this flow unit realization in the same manner as described above.

Because C1B included more wells than C1A, it is a different realization than C1A. Therefore, the means and covariances for each mode and the probabilistic mode assignments for a given depth for a given well (of the 12 wells used in both C1A and C1B) are not exactly the same in the two runs (from a mathematical point of view, problem conditions were altered). The "ModeAssign" logic for run C1B resulted in three "good" limestone modes with apparent NPHI porosities between 10 and 14 porosity units (PU), two "poor" limestone modes, three shale modes, and two dolostone modes. The automatic assignment of two modes to dolostone was made on the basis of apparent grain density. However, because of the known presence of anhydrite, and the lack of reported dolomite in these rocks, the mode assignments for these two modes (M8 and M1) were overridden and manually assigned to be anhydritic limestone (AnhLs). These two modes were assigned brown colors. Figure 44 shows a depth plot for C1B for the same 12 wells used in C1A. The wells are arranged in the same order left-to-right as they are in Figure 22 for C1A.

Figure 45 compares the "fuzzy" (cumulative mode probability) depth plot for C1B with C1A. Although the details for a given well are not exactly the same, it is clear that both realizations show the same reservoir quality trends.

Finally, Figure 46 shows permeability (\log_{10}) versus porosity from the "P&P via AI" method color coded according to the C1B facies for well 33-11, a well with neither core data nor RHOB or DT data. The predictions for porosity and permeability are a "blind prediction" as opposed to a "hold out" prediction test because there is no core data to verify the credibility of the prediction for this well. Again, there are good correlations between permeability and porosity for given flow unit.

C1B, 10 modes; probabilistic modes displayed

31 wells (but, only the 12 wells also used in C1A are displayed here)

RHOB, NPHI, GR, DT

The 19 wells for which RHOB & DT were missing are not displayed here

From left to right:

11-15, 19-12, 56-18, 37-10, 58-4, 34-12, 33-15, 37-11, 59-5, 36-8, 56-16, and 56-17

Well 11-15 is on far left; ordering of other wells based on structural trend of prominent limestone "bed" near bottom of section

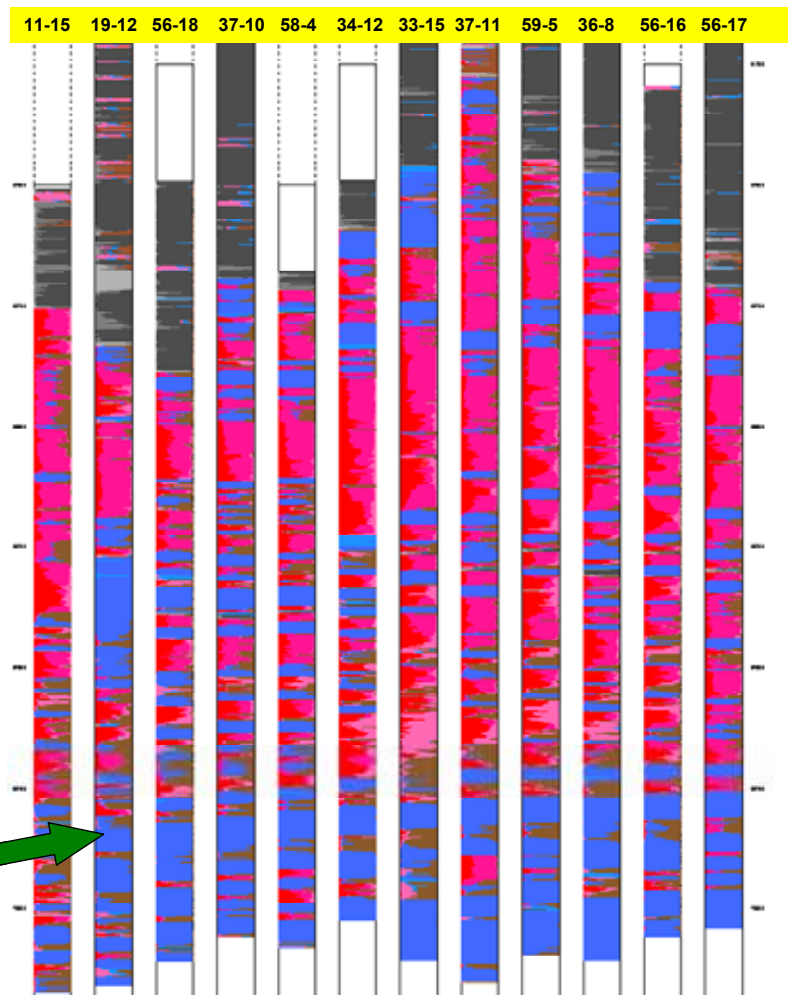


Figure 44: Cluster C1B: Probabilistic Representation of Modes Displayed at Each Depth

Comparison of realizations from C1A (above) and C1B (below).

Both clustering runs using 10 modes and 4 variables (RHOB, NPFI, GR, DT)

C1A: 12 wells with no missing data

C1B: 31 wells (the above 12 wells plus 19 wells with missing RHOB and DT)

Only the 12 wells (with no missing data) common to the two clustering runs are displayed here

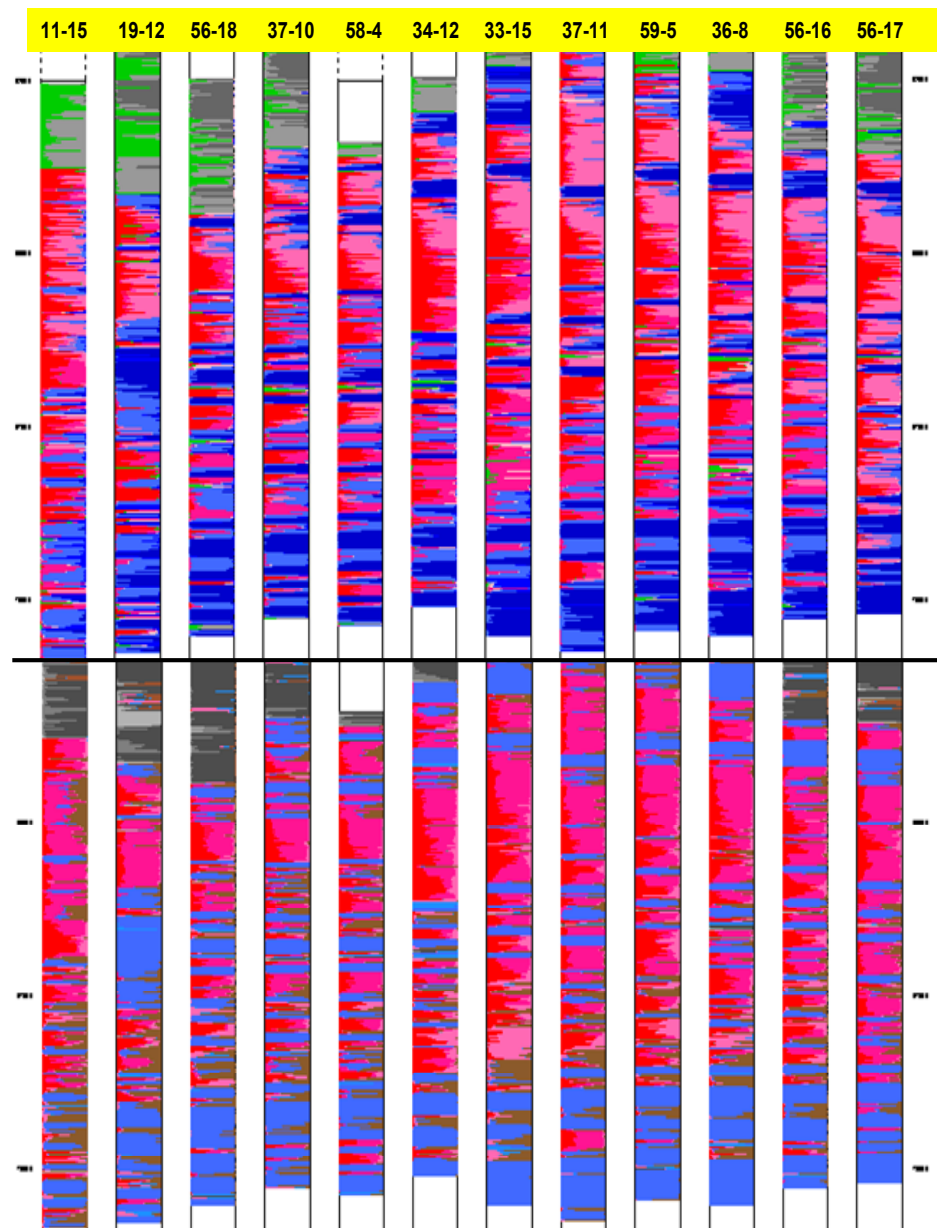


Figure 45: Comparison of Realizations from C1A (Top) and C1B (Bottom)

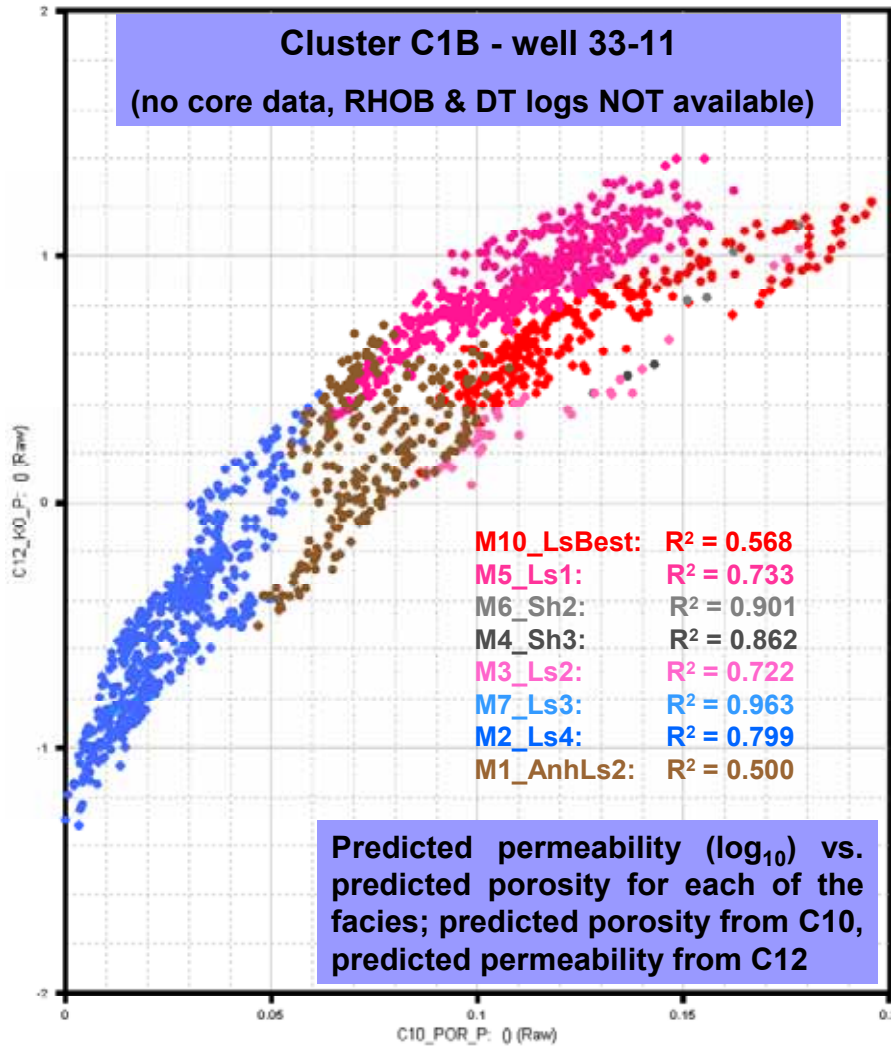


Figure 46: Cluster C1B, Predicted Permeability (\log_{10}) vs. Predicted Porosity for each Facies, Well 33-11

7.3 C1C - Realization from Clustering of 12 Wells Using 4 Log Curves and Core Porosity and Core Permeability from one Cored Well

Cluster run C1A (12 wells and 4 log curves as variables) was duplicated, and core porosity and core permeability from cored well 37-11 were added as variables. The core porosity and permeability from cored wells 11-15 and 19-12 were held out (operationally annulled), and place-holder curves with all null values were inserted for porosity and permeability for the nine non-cored wells. This clustering run is called C1C. For C1C, porosity and permeability were predicted for all nine non-cored wells and for the cored wells 11-15 and 19-12. Cored well 37-11 was the only "calibration" well.

The frequency distribution of porosity and permeability values for each mode is shown in Figures 47 and 48 for all twelve wells. The data for these plots is real core data for well 37-11

and predicted data for the other 11 wells in the clustering run. In both figures, the unit identified as the best limestone (red) presents a clear bias to high values of the corresponding parameter. In contrast, the poorer RQ units (blue) show a bias to lower values of both reservoir parameters. Other modes (units) tend to be spread on the whole parameter range of variability.

Figure 49 shows the same information for non-cored well 33-15 as Figure 24 except track one is a new track depicting the modes probabilistic representation from C1C (the cluster run that included core porosity and core permeability from 37-11). Track 2 depicts modes probabilistic representation modes bed representation from C1A are in track 3, and in tracks 4 and 5, respectively, are the predicted permeability and predicted porosity values using cluster run C1C superimposed to the predicted values coming from cluster C9. Note that the predictions from C1C (pink and light blue) agree closely with the predicted permeability and porosity from C9 (red and dark blue). C9 was the clustering run used to predict porosity and permeability using the "P&P via AI" method. Thus, these two methods of predicting porosity and permeability in a non-cored well (a "blind test") produced nearly identical results.

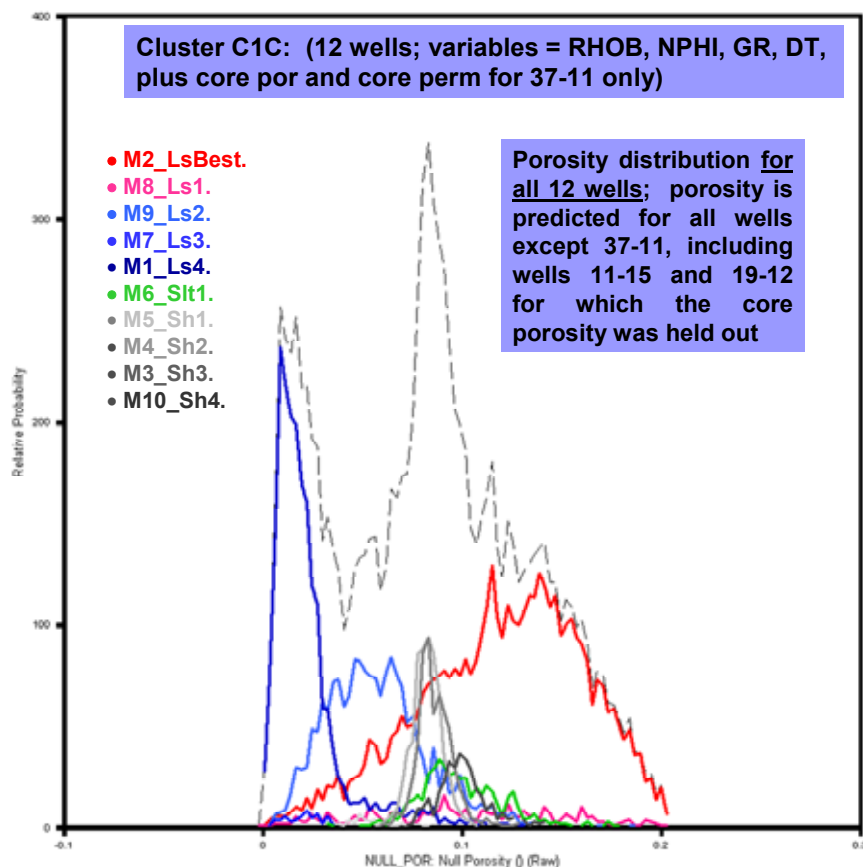


Figure 47: Cluster C1C: Porosity Values Distribution for 12 Wells Discriminated by Modes

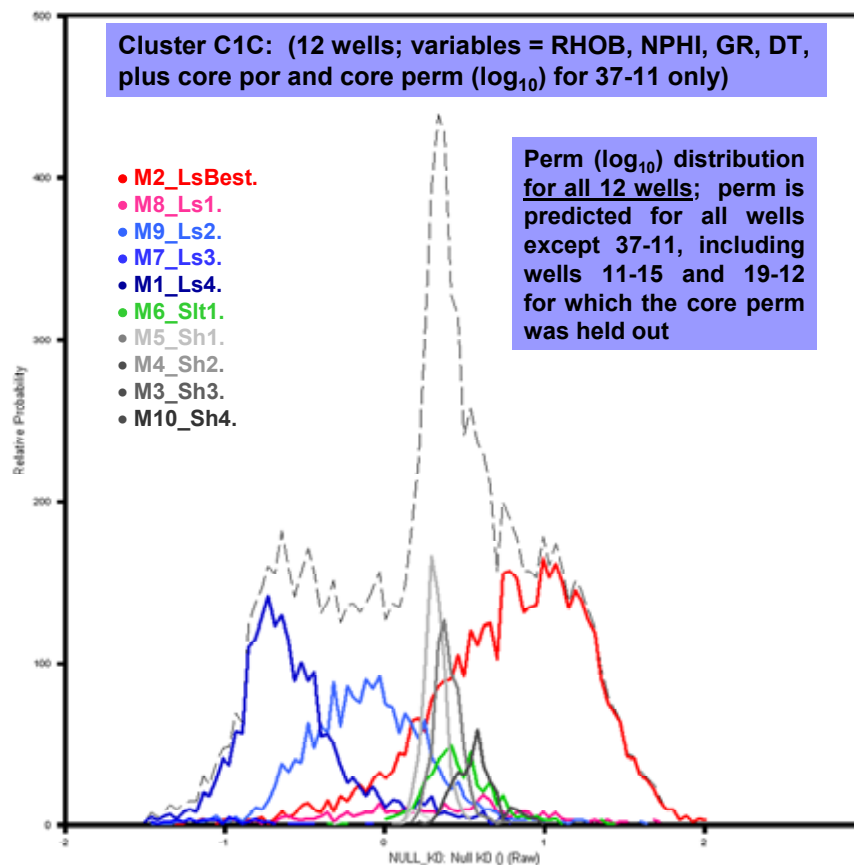


Figure 48: Cluster C1C: Permeability (\log_{10}) Values Distribution for all 12 Wells Discriminated by Modes

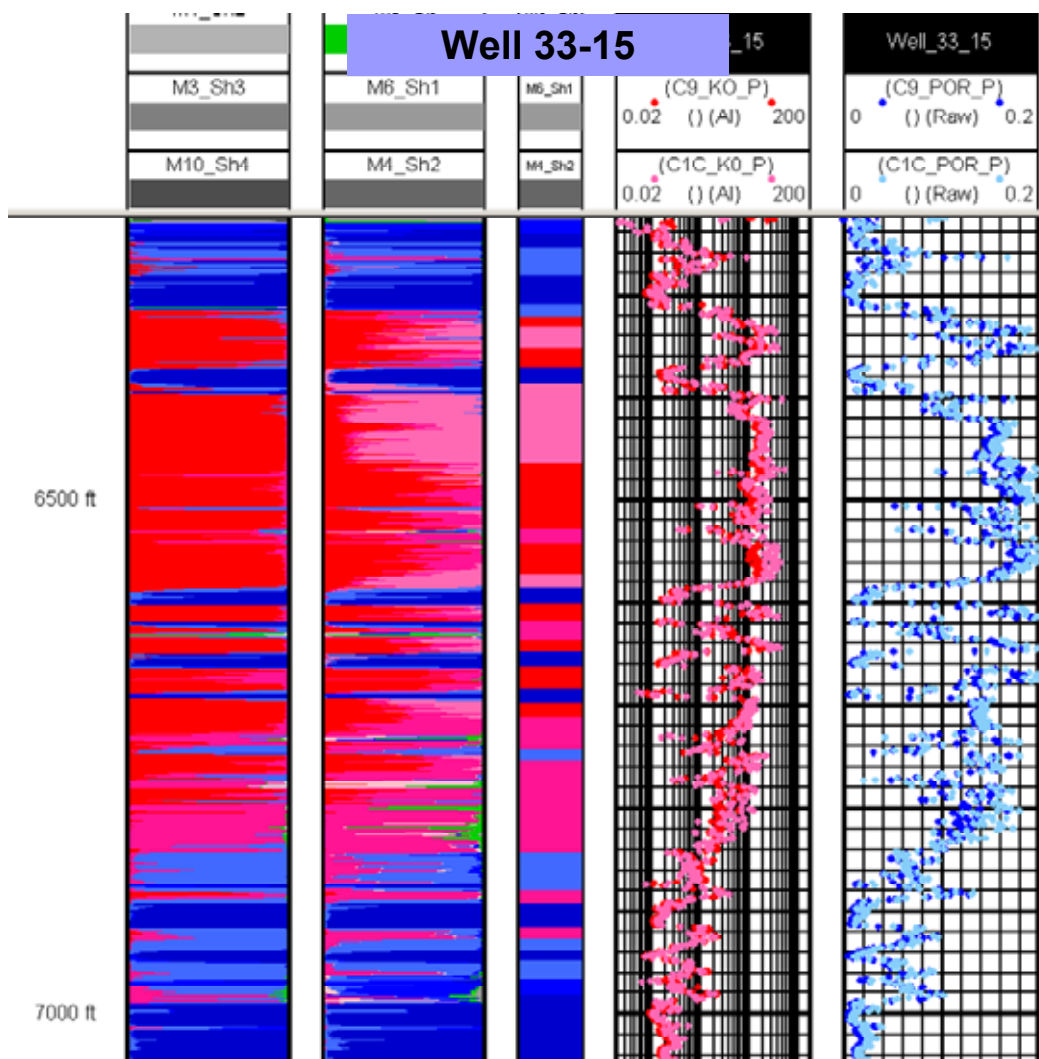


Figure 49: Tracks of Clustering Results for the Non-Cored Well 33-15

Figures 50 and 51 show the correlations, by mode, for the two methods for predicting porosity and permeability for well 33-15. Figure 50 shows predicted porosity using cluster C1C (a “straight clustering estimation” method) versus predicted porosity using cluster C9 (the “RQ via AI” method). Likewise, Figure 51 presents the predicted permeability using cluster C1C (a “straight clustering estimation” method) versus the predicted permeability using cluster C9 (the “RQ via AI” method). Most of the clouds representing modes show a marked alignment which is indicative of high similarity between compared values. Correlation coefficients support this fact and there are obviously better correlations at better RQ modes. In both figures the predicted parameter values using cluster C1C tend to be slightly greater than their analogues using cluster C9 (in particular the best RQ unit). This is due to the incorporation of the cored well 19-12 in the configuration of cluster C9. Porosity values as well as permeability measurements of this well are generally lower than their corresponding pairs of wells 37-11 and 11-15.

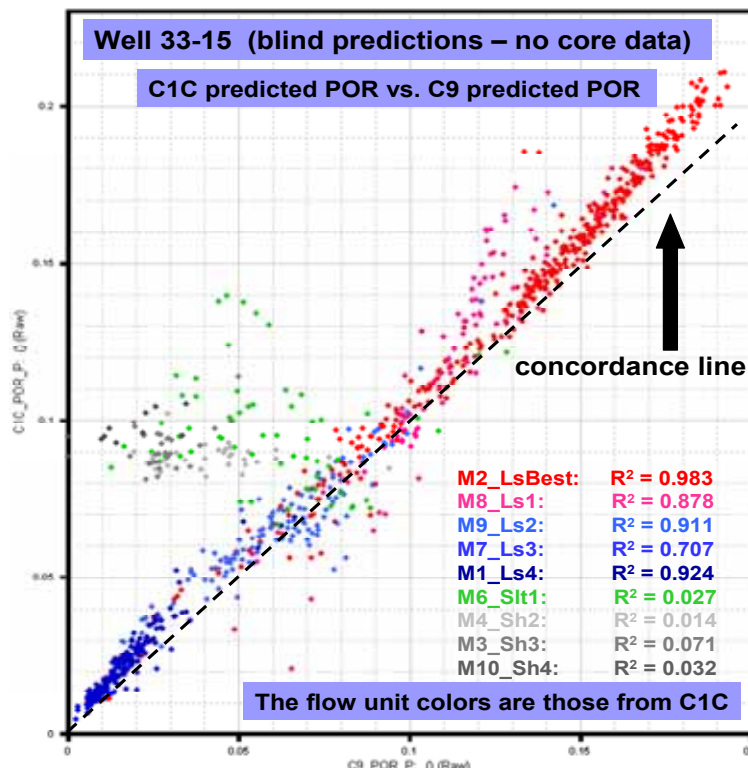


Figure 50: C1C Predicted POR vs. C9 Predicted POR, Well 33-15

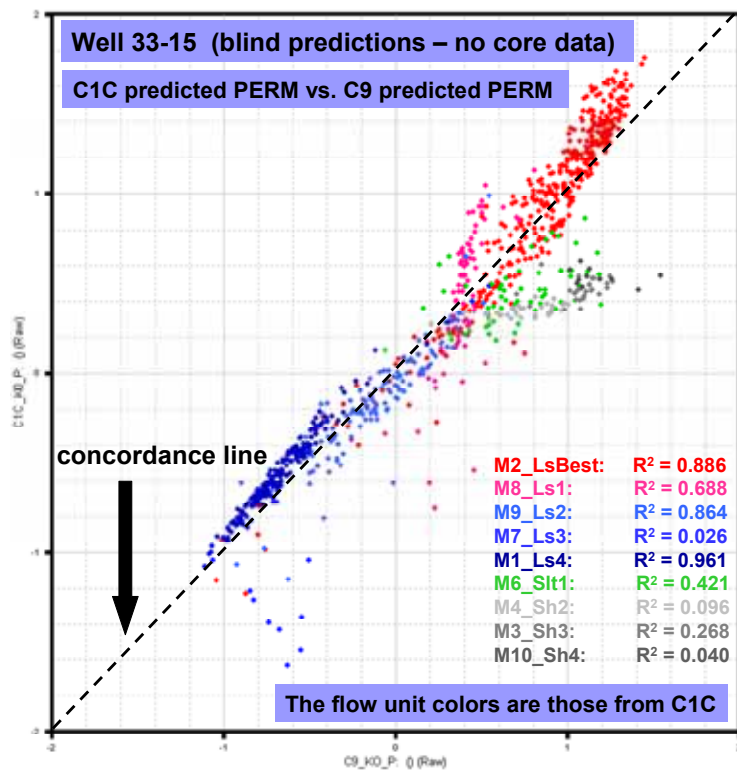


Figure 51: C1C Predicted PERM vs. C9 Predicted PERM (Log₁₀), Well 33-15

Figure 52 presents the predicted porosity using cluster C1C (a “straight clustering estimation” method) versus actual core porosity at well 11-15. In this graphic, the flow unit colors are those from cluster C1C. On the other hand, Figure 53 also shows predicted porosity using cluster C1C versus actual core porosity, but now the flow unit colors are guided by cluster C1A. Notice that the correlations for three of the four limestone flow units (Figure 52) are better using the flow unit definitions from cluster C1A (fed with information of three cored wells) than with using the low unit definitions from C1C (Figure 53). These holdout tests provide a measure of estimating how valid the predictions for non-cored wells might be.

Similar results were obtained for the cored well 19-12 which also was a hold-out testing well. Figures 54 and 55 both display crossplots of the predicted porosity using cluster C1C (a “straight clustering estimation” method) versus actual core porosity at well 19-12. Results on Figure 54 are guided by modes of cluster run C1C while results on Figure 55 are those from modes of cluster C1A. Note, on Figure 55, that the correlations for most of the limestone flow units are better using the flow unit definitions from cluster C1A (fed with three cored wells data) than with using the low unit definitions from C1C (Figure 54).

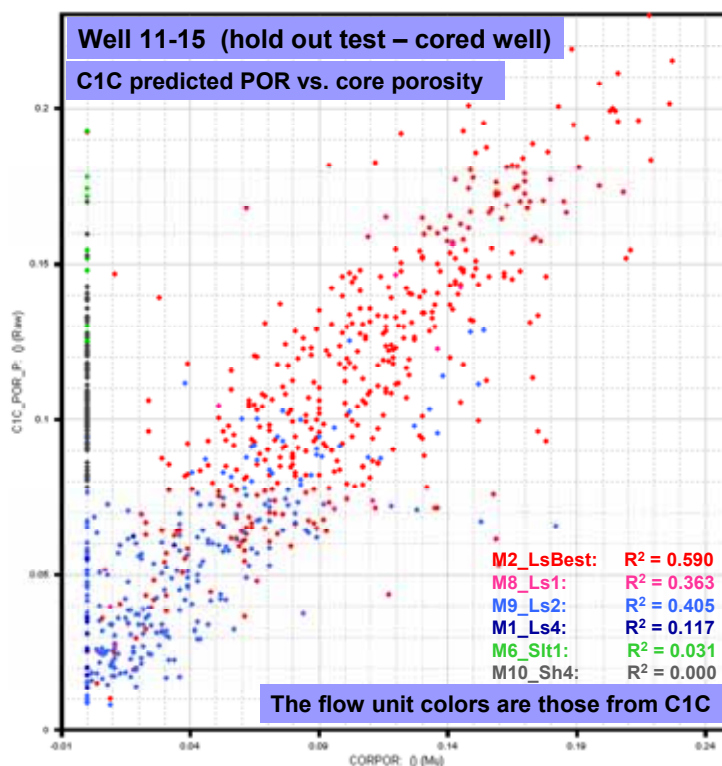


Figure 52: C1C Predicted POR vs. Actual Core Porosity, Well 11-15 as a Hold-Out Testing Well

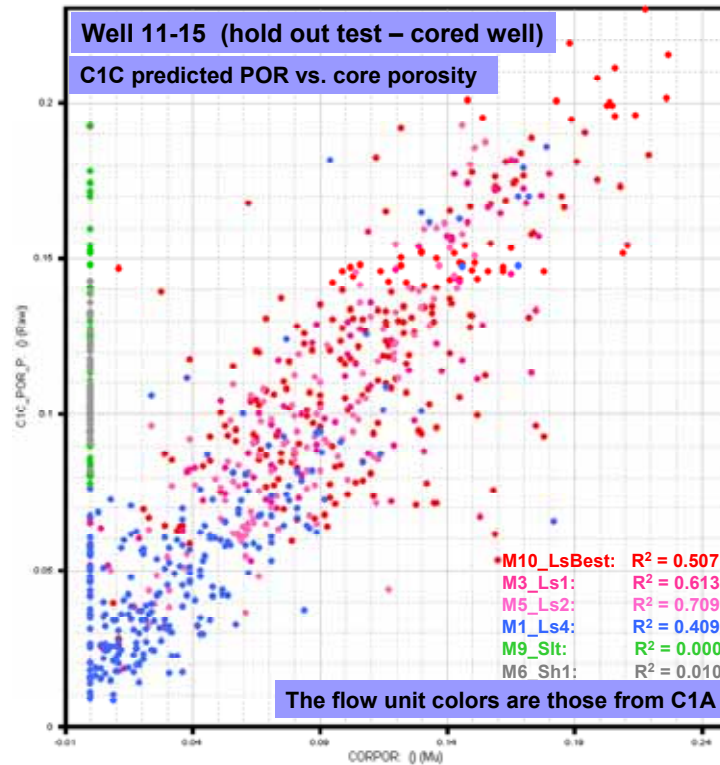


Figure 53: C1C Predicted POR vs. Actual Core Porosity, Well 11-15 as a Hold-Out Testing Well

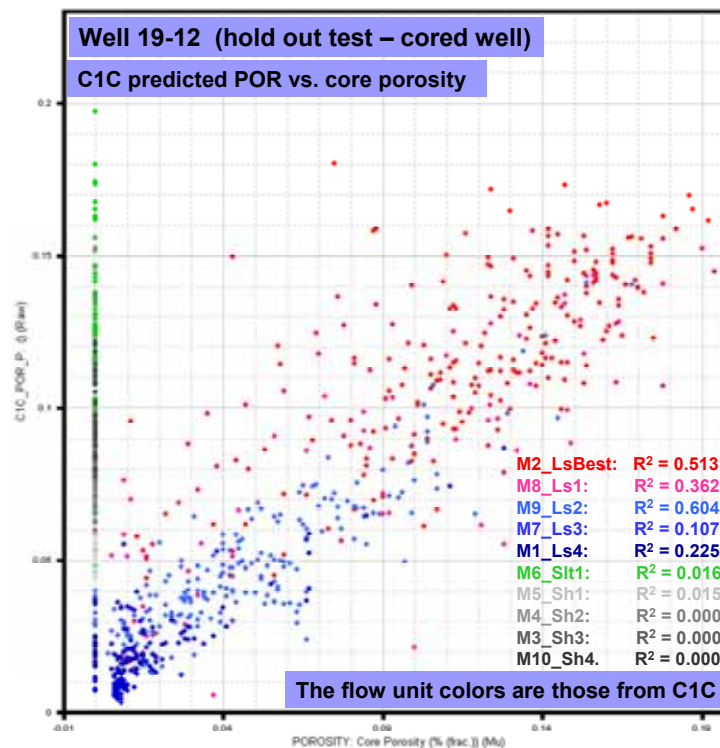


Figure 54: C1C Predicted POR vs. Actual Core Porosity, Well 19-12 as a Hold-Out Testing Well

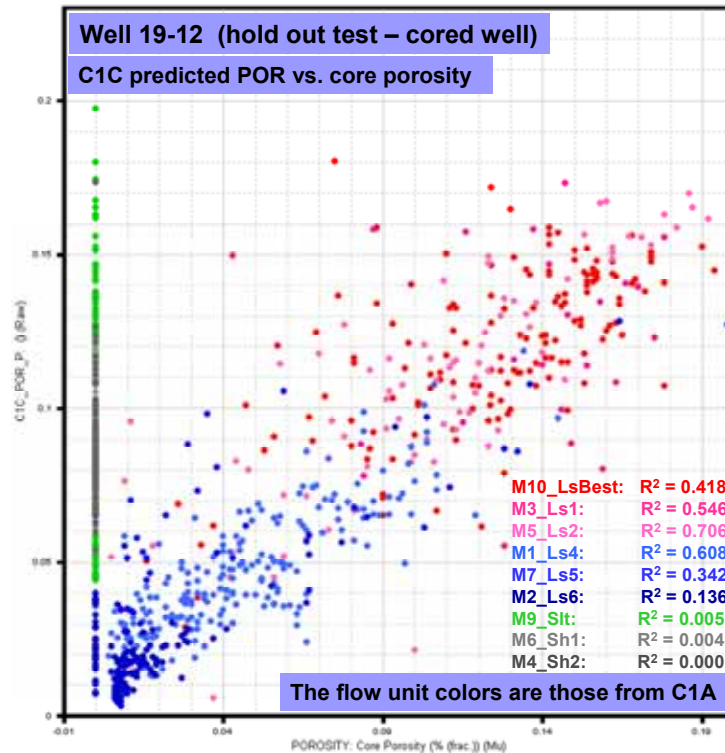


Figure 55: C1C Predicted POR vs. Actual Core Porosity, Well 19-12 as a Hold-Out Testing Well

Predicted permeability versus actual core permeability (logarithm) for cored wells 11-15 and 19-12 are shown respectively in Figures 56 and 57 using C1C flow units. For well 11-15, more “under-predictions” were made than “over-predictions” for the better RQ flow units. For well 19-12, more “over-predictions” were made than “under-predictions” for the better RQ flow units.

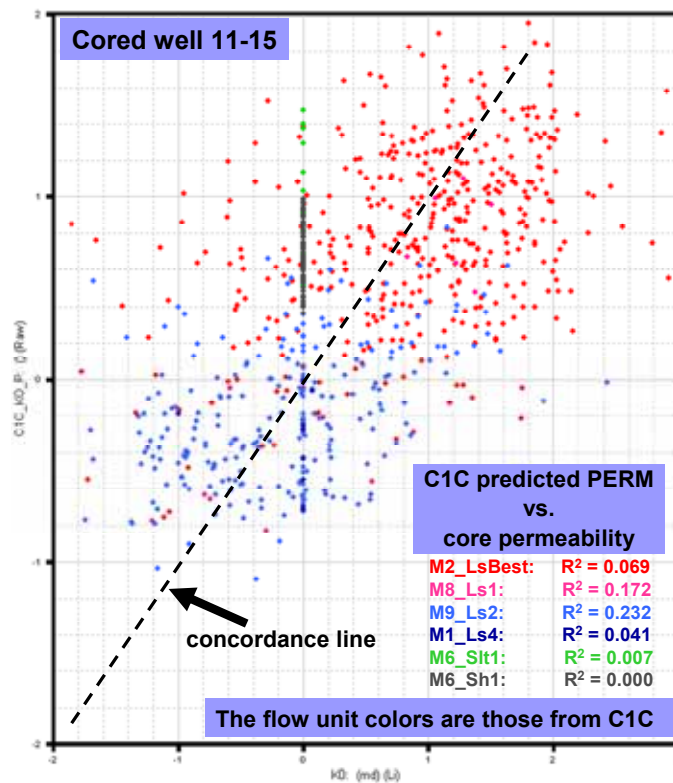


Figure 56: C1C Predicted PERM (Log₁₀) vs. Actual Core Permeability, Well 11-15 as a Hold-Out Testing Well

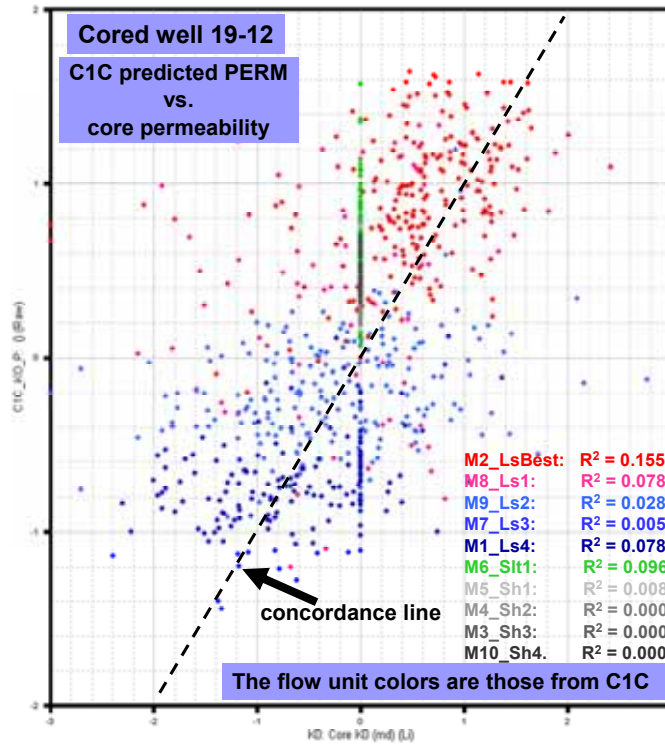


Figure 57: C1C Predicted PERM (Log₁₀) vs. Actual Core Permeability, Well 19-12 as a Hold-Out Testing Well

The frequency distributions for actual core porosity and predicted porosity are shown in Figures 58 and 59 for wells 11-15 and 19-12, respectively. In both graphics, the predicted porosity frequency has a marked tendency of behaving symmetrically (which is a consequence of the Gaussian models adopted by GAMLS). In both wells, the range of variability of predicted values is the same as the range of variability of actual porosity values. This is indicative that predicted values are neither overestimated nor underestimated.

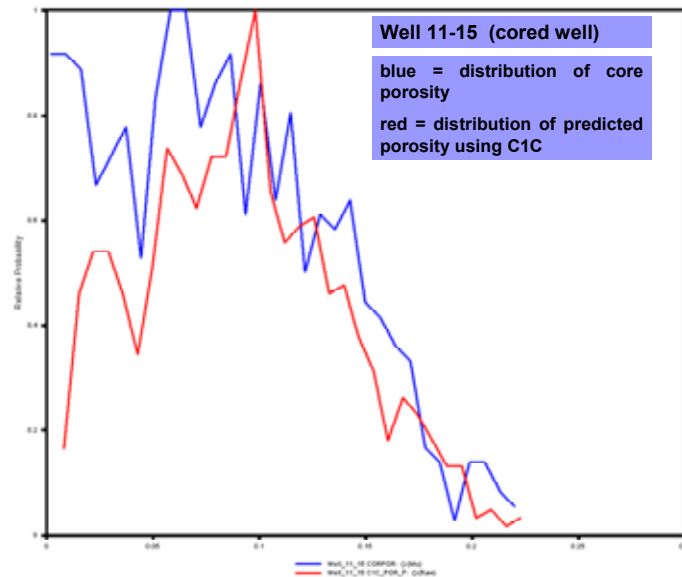


Figure 58: Frequency Distributions for Core Porosity and Predicted Porosity (Using Cluster C1C), Well 11-15

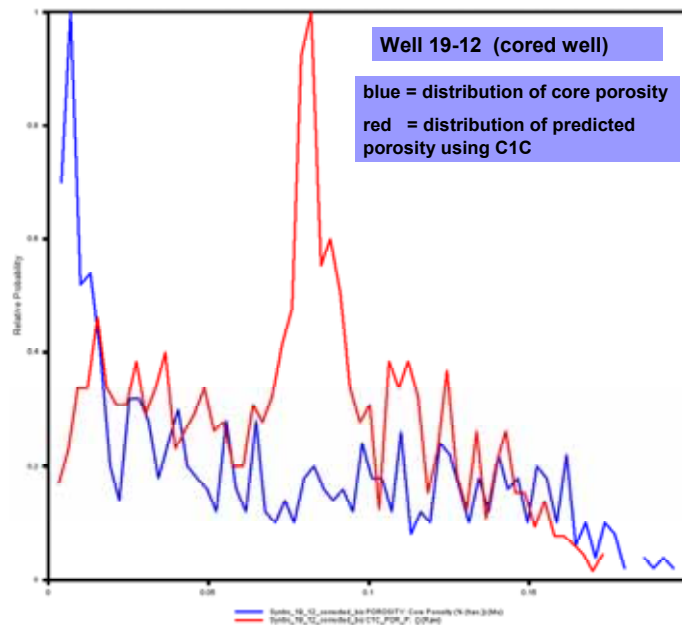


Figure 59: Frequency Distributions for Core Porosity and Predicted Porosity (Using Cluster C1C), Well 19-12

The frequency distributions for actual core permeability and predicted permeability are shown in Figures 60 and 61 for wells 11-15 and 19-12, respectively. Examination of these plots shows that, in contrast to porosity, permeability can be over-predicted or under-predicted by up to an order of magnitude. Figure 62 depicts actual data of $\log_{10}(K_0)$ versus the \log_{10} of K_0 minus the \log_{10} of predicted K_0 permeability for well 11-15 using C1A modes, and indicates that there is little to no bias in over-prediction versus under-prediction for permeability.

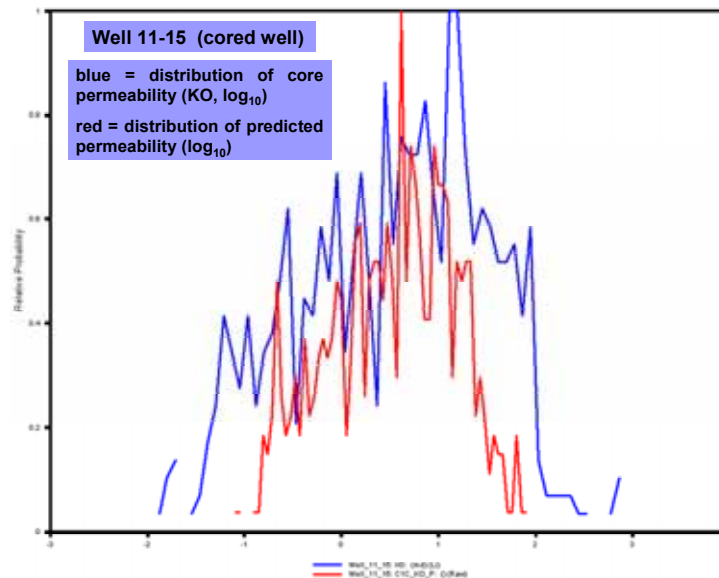


Figure 60: Frequency Distributions for Core Permeability and Predicted Permeability (Using Cluster C1C), Well 11-15

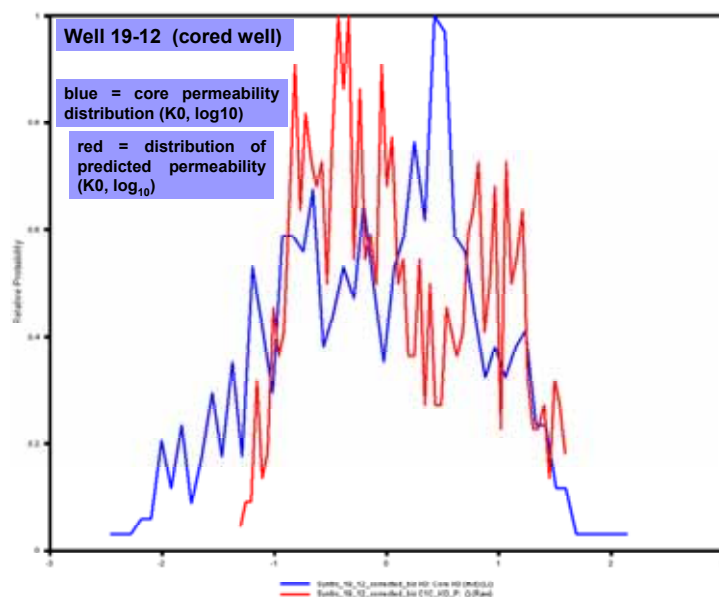


Figure 61: Frequency Distributions for Core Permeability and Predicted Permeability (Using Cluster C1C), Well 19-12

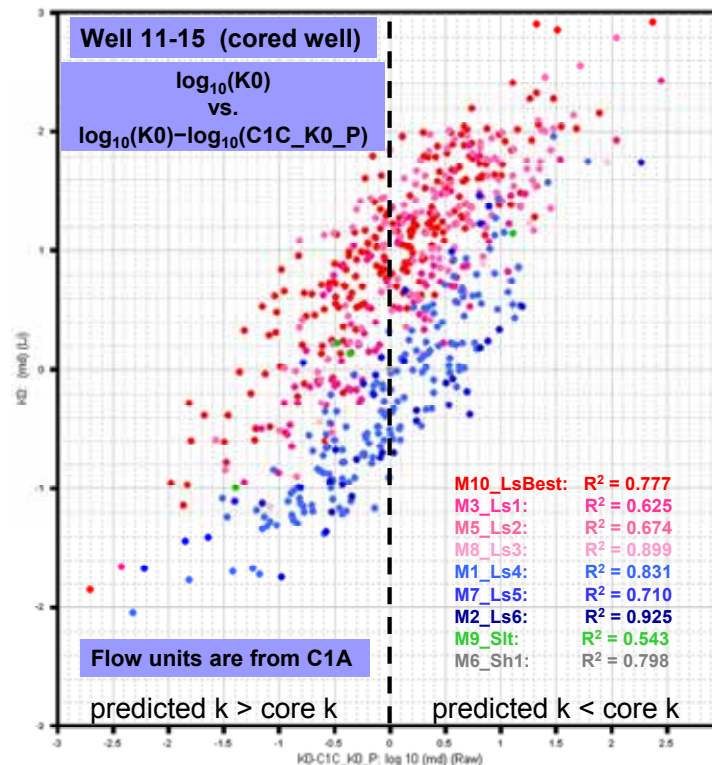


Figure 62: $\log_{10}(K0)$ vs. $\log_{10}(K0) - \log_{10}(K90)$, Well 11-15

The "accuracy" of porosity prediction is much better than that for permeability prediction. This is not uncommon in carbonates (in our experience). Low permeability samples tend to be over-predicted and high permeability samples tend to be under-predicted. This is due to a combination of factors:

1. the poor sensitivity of log tools to properties that control permeability,
2. the poor vertical log tool resolution which tends to average rock properties over 2 to 3 feet within which depth range permeability of one-inch diameter samples might vary by 2 to 3 orders of magnitude,
3. the inherent vertical and horizontal heterogeneity of the reservoir making applicability of the generated model (using core data of well 37-11) modestly transferred to wells hundreds of feet away, and
4. the nature of the GAMLS modeling process which, based on Gaussian distributions, tends to emphasis the central bulk of the distributions as opposed to the tails.

Based on these hold-out tests, core permeability values tend to be under-predicted and over-predicted by up to an order of magnitude in some locations. However, that is not necessarily negative, since whole core measurements are not representative of permeability data over larger scales, even to 12 inches (see discussion of Raines and Helms, 2007) above. That is, estimations of permeability made by using GAMLS appear to have a built-in inherent upscaling process that results in permeability trends that might be more representative for flow simulation modeling than the more erratic trends depicted by the actual core measurements. This inherent

upscaling is due to the fact that the GAMLS process is correlating core data with well log curves, the latter derived from a signal that averages the properties that control permeability over vertical intervals of 2 to 4 feet.

7.4 C1D - Clustering Run Same as C1C Using 25 Modes

Figure 63 shows the crossplot of the $\log_{10}(K_0)$ estimated by C1D ($\log_{10}(\text{C1D_K0_EST})$) versus the $\log_{10}(K_0)$ estimated by C1C ($\log_{10}(\text{C1C_K0_EST})$) at cored well 19-12, and where samples have been colored according to C1D “crisp” mode assignment. Coefficient of correlation values (R^2) are greater than 0.93 for all of the interpreted limestone lithologic units implying pretty similar estimates. Therefore, increasing the number of clustering modes from 10 to 25 can provide a more detailed "flow unit" realization but the estimates of K_0 (logarithm) not necessarily are improved.

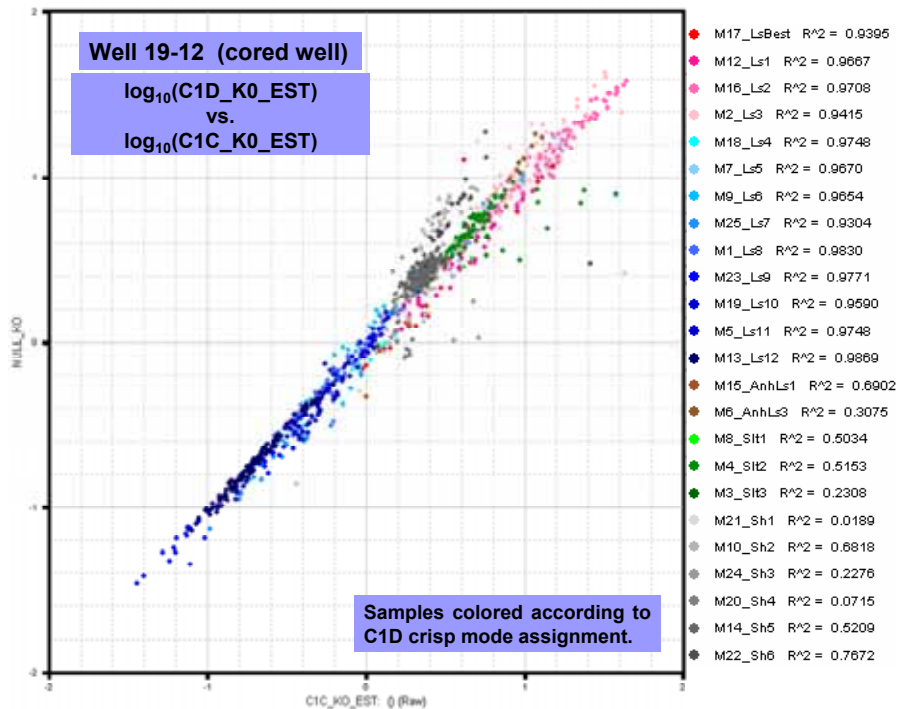


Figure 63: $\log_{10}(\text{C1D_K0_EST})$ vs. $\log_{10}(\text{C1C_K0_EST})$, Well 19-12

7.5 C1E - Clustering Run C1C Using Core Data Well 37-11 & 11-15

The crossplot of $\log_{10}(K0)$ estimated by C1E ($\log_{10}(C1E_K0_EST)$) versus the $\log_{10}(K0)$ estimated by C1C ($\log_{10}(C1C_K0_EST)$) at cored well 19-12 is shown in Figure 64. Coefficient of correlation values (R^2) are greater than 0.89 for four of the five interpreted limestone lithologic units. Both permeability predictions are equivalent. An a priori conclusion is that incorporation of core data from an additional well does not improve the quality of permeability estimates. However, in this case, the huge quantity of core measurements of well 37-11 covering the whole SACROC unit, and geological similarities at locations of these cored wells can sufficiently explain this fact.

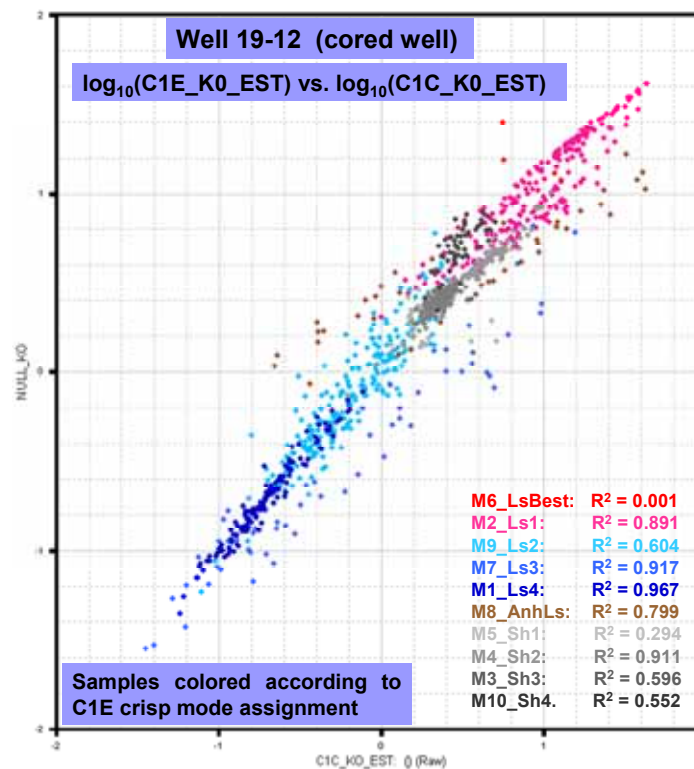


Figure 64: $\log_{10}(C1E_K0_EST)$ vs. $\log_{10}(C1C_K0_EST)$, Well 19-12

Figure 65 shows the $\log_{10}(K0)$ estimated by C1E ($\log_{10}(C1E_K0_EST)$) versus $\log_{10}(K0)$ (logarithm of actual core $K0$ permeability). This figure is analogous to Figure 57, which shows estimated permeability $\log_{10}(C1C_K0_EST)$ versus $\log_{10}(K0)$. Figures 65 and 57 are placed side-by-side in Figure 66 for direct comparison. The average coefficients of correlation for the five limestone lithologies are 0.069 for C1C and 0.073 for C1E. This indicates a negligible gain in prediction "accuracy" by addition of a second calibration well.

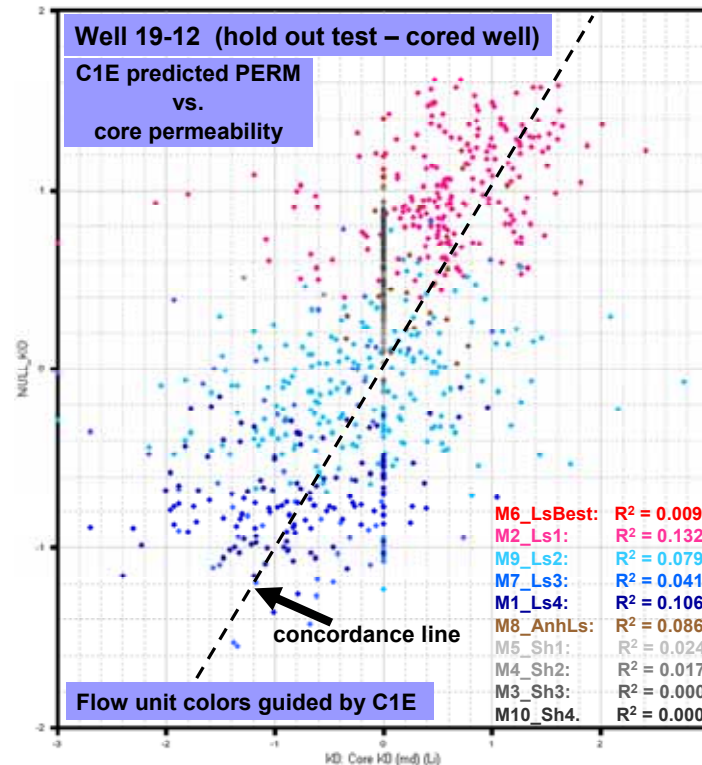


Figure 65: $\text{Log}_{10}(\text{C1E_K0_EST})$ Versus $\text{Log}_{10}(\text{K0})$ (Actual Data): Well 19-12

Figure 67 shows CMP plots for Well 37-11 for five clustering runs. Variations in clustering run setups included wells considered, clustering variables, and number of modes (Table 2). In each CMP plot, the three to four red-colored modes represent the best reservoir quality limestones, with porosities ranging from 10 to 14 percent, and the three to four blue-colored modes represent poor reservoir quality limestones with porosities generally less than 6 percent.

The permeabilities have a similar division into a high range and a low range. Most of the red-colored modes have permeabilities greater than 2.5 md and most of the blue-colored modes have permeabilities less than 2.5 md.

In the simplest lithologic division, the limestone modes could be divided into only two: good (red) and poor (blue). An example of this was shown in Figure 26. Regardless of the realization, the good RQ limestones are clearly discriminated from the poor RQ limestones (Figure 67). This indicates a fundamental difference in depositional or diagenetic features that has affected reservoir quality. The poor RQ limestones represent barriers to vertical flow. If the good (or poor) RQ units pinch out laterally, then this presents a significant problem to be addressed via geostatistical modeling in the development of geological and reservoir flow models.

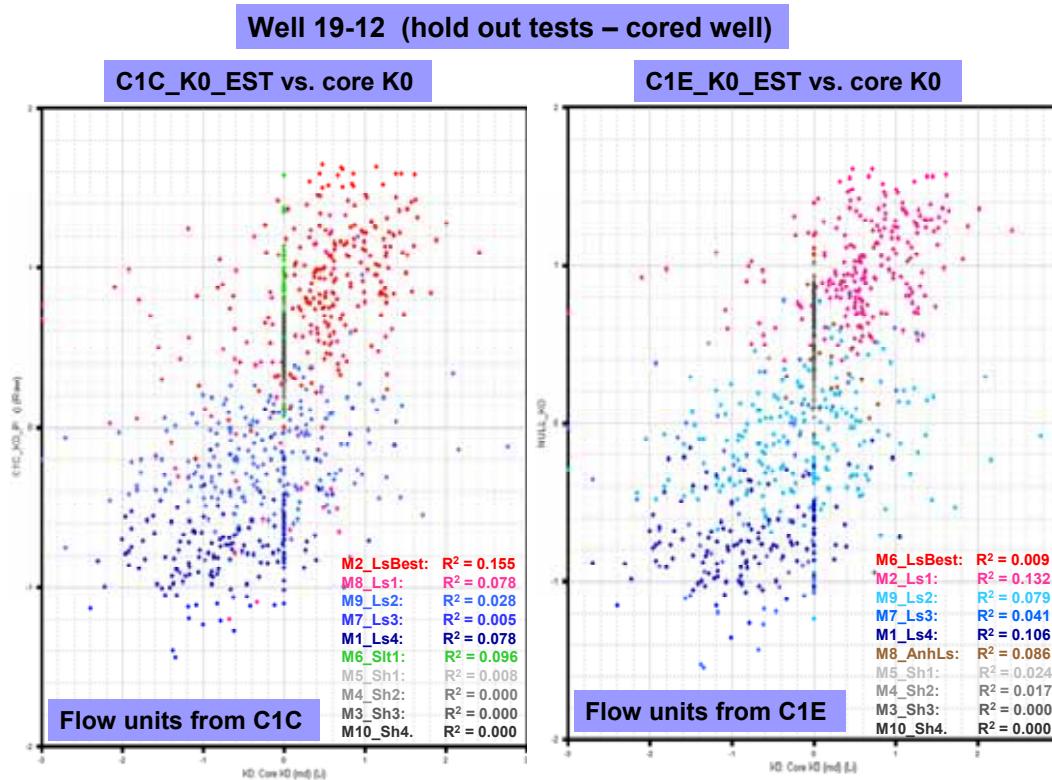
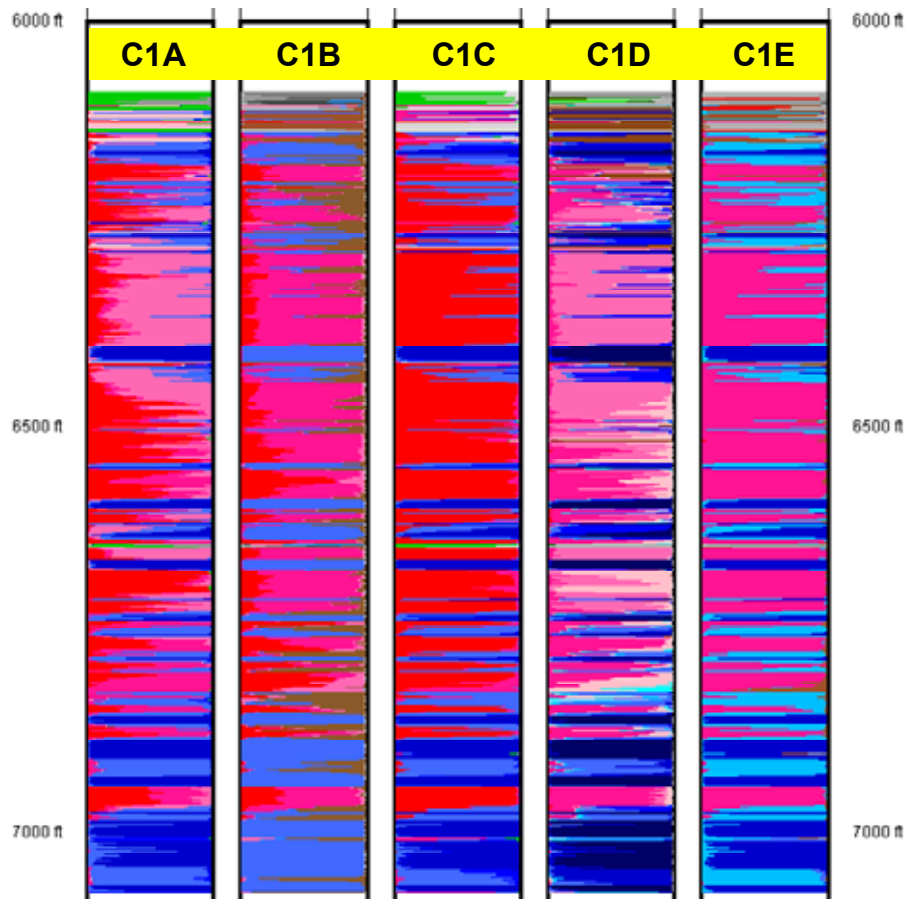


Figure 66: Predicted $\text{Log}_{10}(\text{K0})$ vs Actual $\text{Log}_{10}(\text{K0})$, Well 19-12
Left: Predicted Values from Clustering Run C1C. Right: Predicted Values from Clustering Run C1E

Well 37-11 (cored well): Comparison of Realizations from Five Clustering Runs



**Figure 67: Comparison of Realizations from Five Clustering Runs
Modes Probabilistic Representation Displayed According to Each Respective Cluster**

7.6 Analysis of Oriented Cross Sections

To attempt to gain a better understanding of the lateral and vertical variability of the flow units in the study area, two E-W cross sections and one N-S cross section were generated. Figure 68 depicts the locations of these cross sections named A-A', B-B', and C-C'. In particular, the cross section C-C' was studied and correlated by M. Raines⁶ before the development of this work.

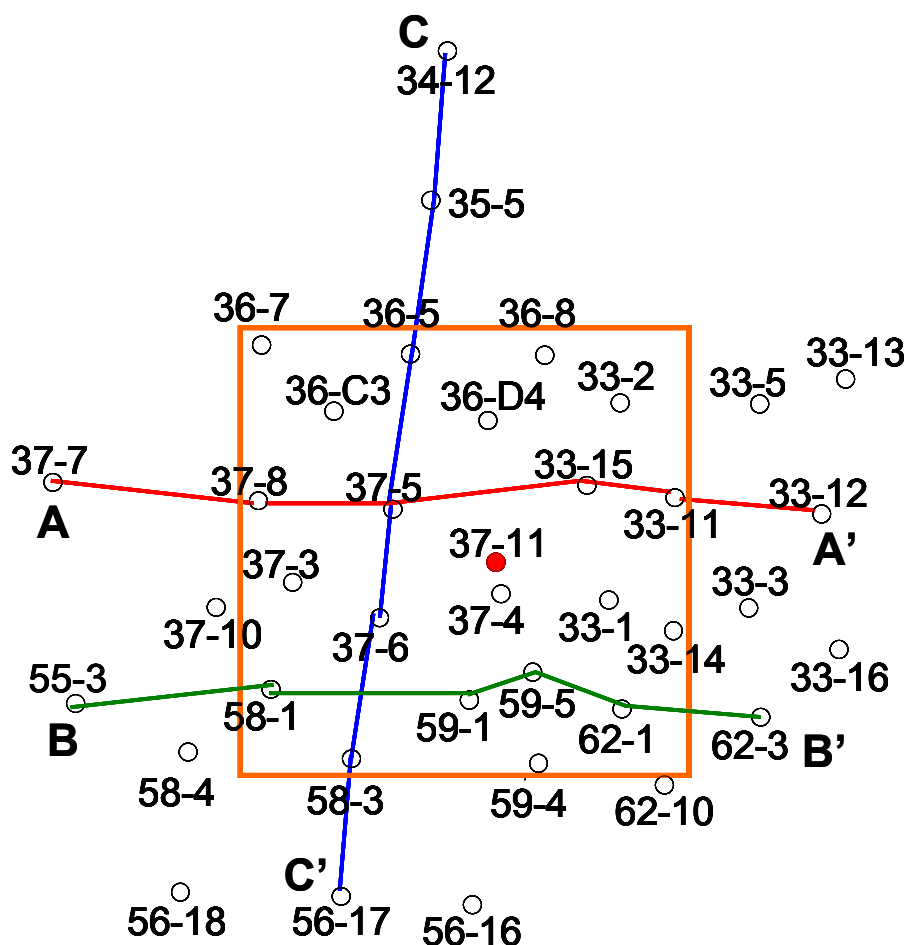


Figure 68: Location Map Cross Sections A-A', B-B', and C-C'

Figure 69 depicts cross section A-A' with interpreted "facies" using GAMLS clustering analysis, in particular, guided by cluster C1A. In this northern cross section, modes probabilistic representation of wells 37-7, 37-8, 37-5, 33-15, 33-11, and 33-12 were included. The bright red layers represent best reservoir quality (RQ) and dark blue represents the poorest RQ. Some limestone unit "tops" have been visually correlated (thin red lines). As it can be seen, thickest and best RQ units are in the upper half of section (but with a thin section of generally poor RQ above that).

In Figure 70, cross section A-A' is divided into three flow unit groups and labeled A, B, and C respectively from top to bottom at well 37-5. Zone B appears to have the best overall RQ and is separated from overlying zone A and underlying zone C by apparently laterally continuous flow-barrier units of low RQ (dark blue units). In addition, within zones A, B and C some individual flow units exist that can possibly be correlated.

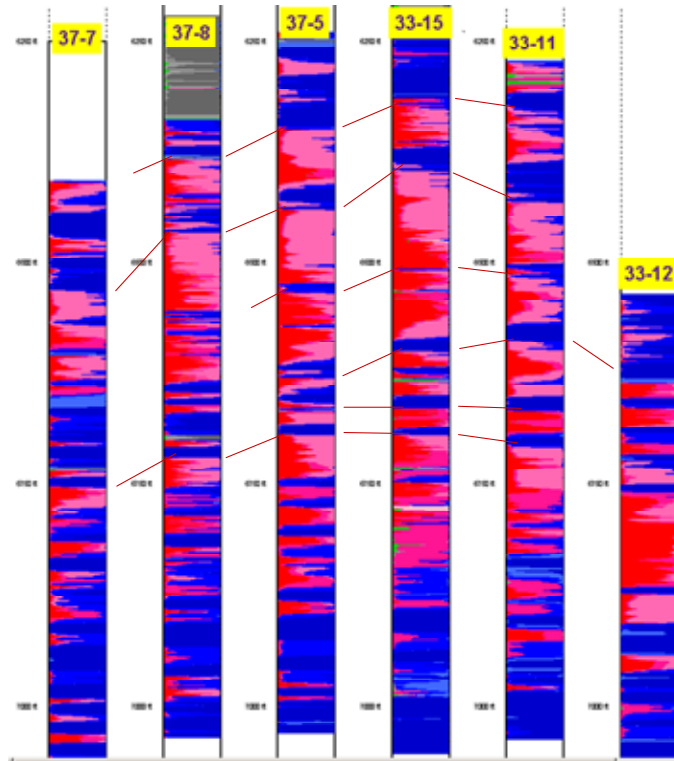


Figure 69: Cross Section A-A'

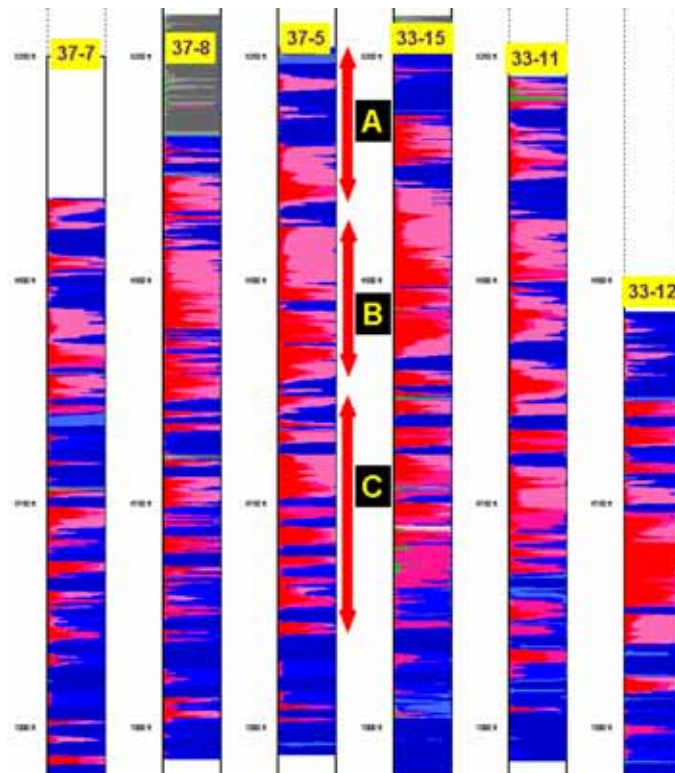


Figure 70: Cross Section A-A' with Positions of Zones A, B, and C

Figure 71 depicts cross section B-B' with interpreted “facies” using clustering analysis results of C1A. Modes probabilistic representation of wells 55-3, 58-1, 59-1, 59-5, 62-1, and 62-3 are here presented. Here, approximate positions of the zones A, B, and C are also shown at well 59-1. The zones and the individual flow units within the zones are less well-defined in this southern part of the study area. Low RQ zones that are laterally continuous would act as vertical flow barriers and tend to compartmentalize flow units causing any injected fluids to be “channeled” between the flow barriers. In consequence, local azimuthal anisotropy should be considered when planning injection well locations, perforation intervals, and sweep directions.

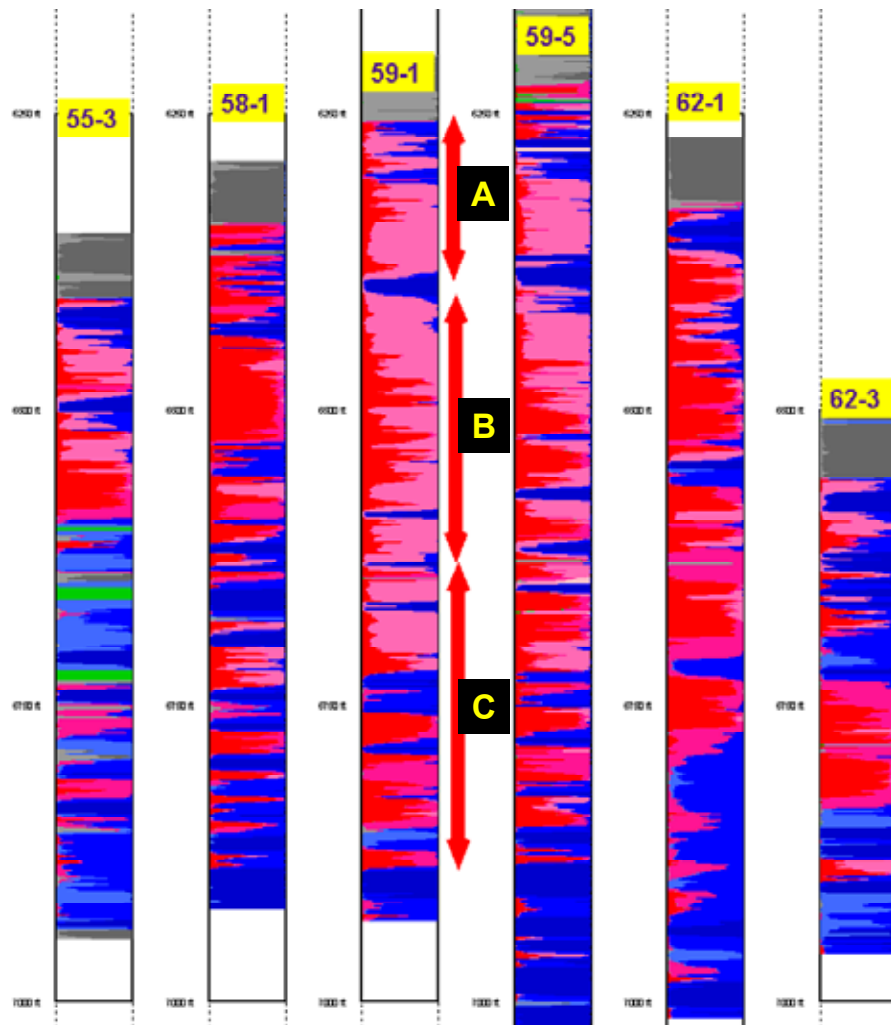


Figure 71: Cross Section B-B' with Positions of Zones A, B, and C

Cross section C-C' is presented in Figure 72, and includes CMP depth plots of wells 34-12, 35-5, 36-5, 37-5, 37-6, 58-3, and 56-17. Zones A, B, and C respectively from top to bottom are labeled at well 36-5. In this north-south section, zone B has the best overall RQ and it is separated from overlying zone A and underlying zone C by apparently laterally continuous flow-

barrier units of low RQ (dark blue units). In addition, within zones A, B and C some individual flow units exist that can be correlated.

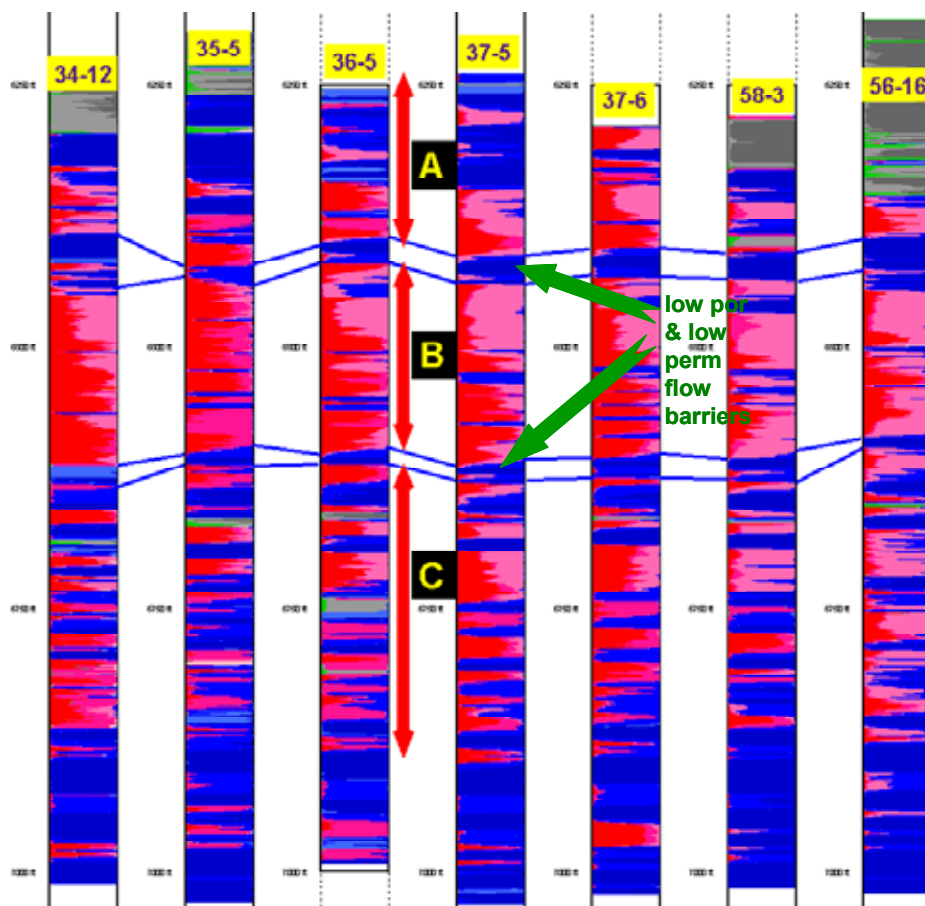


Figure 72: Cross Section C-C' with Positions of Zones A, B, and C

Figures 70, 71 and 72 can be used to analyze the lateral continuity of the mentioned zones A, B, and C. Cross section C-C' presents more similarities zone by zone when the CMP depth plots of all wells are compared. This is particularly evident in zone B where better RQ units are predominant. This fact can be related with the pattern of deposition of SACROC sediments with a dominant N-S trend of its structural axis. The same depositional setting could explain why in E-W cross sections A-A' and B-B' it is more difficult to find stratigraphic “empathy” between same zones of different wells or among the individual flow units within the zones.

Figure 73 is the lower portion of Figure 22 in which study wells were presented without a specific orientation. In Figure 73, the wells used in clustering run C1A were shown left-to-right in an order based on an apparent structural trend of a prominent limestone “bed” near the bottom of the section. This structural trend is “apparent” because it assumes a monoclin dip; in fact,

cross sections A-A', B-B', and C-C', based on defined geographic orientations, show the true structure to be slightly domal due to the reef buildup near the center of the core study area.

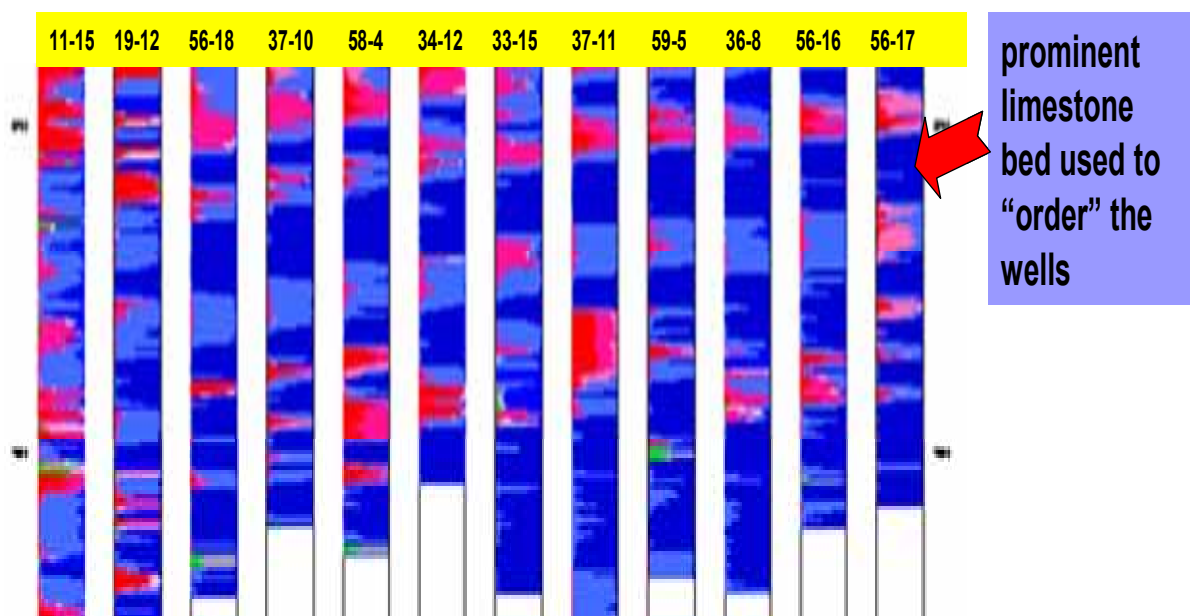


Figure 73: C1A Wells with Order Based on an Apparent Structural Trend of a Prominent Limestone “Bed” Near the Bottom of the Section (modified from Figure 22)

The compromise of ordering the wells according to apparent structure is done so that cross sections can be made that include wells which are not aligned along a particular geographic direction but that do show and emphasize the degree to which particular beds can be correlated even though they might not be nearest-neighbor wells.

To help define flow unit trends in a larger context, we summarize key points from Waite²⁰ (1993). These summary comments refer to the Horseshoe Reef in general and all features mentioned might not occur locally in the Kelly-Snyder Field.

Waite²⁰ identified four 3rd order sequences in the reef. The sequence boundaries are marked by prominent seismic reflectors. Each of these boundaries apparently marks the end of an overall shallowing (drop in sea level) and the onset of a major flooding episode (rise in sea level). Shallowing episodes might have been sufficient to generate an interval of sub-aerial erosion during maximum sea level drop during which time karsting and leaching (possibly accompanied by enhanced porosity development) occurred.

Chronology of the sequences is via fusulinid foraminiferal biostratigraphy which divides the Pennsylvanian section into 14 zones (7 in the earlier Pennsylvanian Strawn and 7 in the later Pennsylvanian Canyon and Cisco (Waite²⁰, Figure 4). Each of the foraminiferal zones is

interpreted to represent about 1 million years of deposition. The four 3rd-order sequences that have been recognized via seismic are marked by sequence boundaries (from oldest to youngest) at the top of the Strawn (top of foraminiferal zone DS7), at the top of the Canyon A (top of zone MC1), at the top of the Canyon B (top of zone MC4), and at the top of the Canyon B/Cisco (top of zone MC7). Onlapping relatively deep water Permian Wolfcampian shales overlie the uppermost carbonate reef deposits of foraminiferal zone MC4.

From well log analysis, 4th and 5th order sequences have been interpreted within some of the 3rd order sequences and the tops of some of these higher-order sequences are interpreted to be exposure surfaces (Waite²⁰, Figure 9). Although the Canyon A (50-100 feet) is interpreted to be a single parasequence set, in at least one area of the Horseshoe Atoll, the Canyon B (0-400 feet) consists of at least seven parasequence sets.

Each 3rd order sequence represents an evolution through time of carbonate deposits changing from platform to bank to reef mound. During each sequence, water became shallower and there was a decrease in the areal extent of the deposits often culminating in a "pinnacle" structure. Facies geometry of the four 3rd order sequences as interpreted from seismic indicate that the Strawn is a "mounded discontinuous" facies, the Canyon A is a "distinct clinoform" facies, the Canyon B is a "mounded coherent" facies, and the Canyon C/Cisco is a "lensoid" and/or "chaotic" facies. The rock types are mostly wackestones and packstones with occasional grainstones.

Using the study by Waite²⁰ as a model, we have attempted to correlate the data-driven defined flow units with his 3rd order sequences (Waite²⁰, Figure 19). In general, our database did not permit us to identify the top of the Strawn using well logs (many of our logs stopped before penetrating the Strawn). Since that footage was not included in the clustering runs, we could not search for a sequence boundary there.

In Figure 74, the 12 wells (G1 and G2 wells) used in C1A are positioned left-to-right according to relative structural position using a prominent low RQ "bed" near the bottom of the studied interval. This "bed" is interpreted to be the base of sequence boundary "Top C1" identified by Waite (1993) representing a flooding surface.

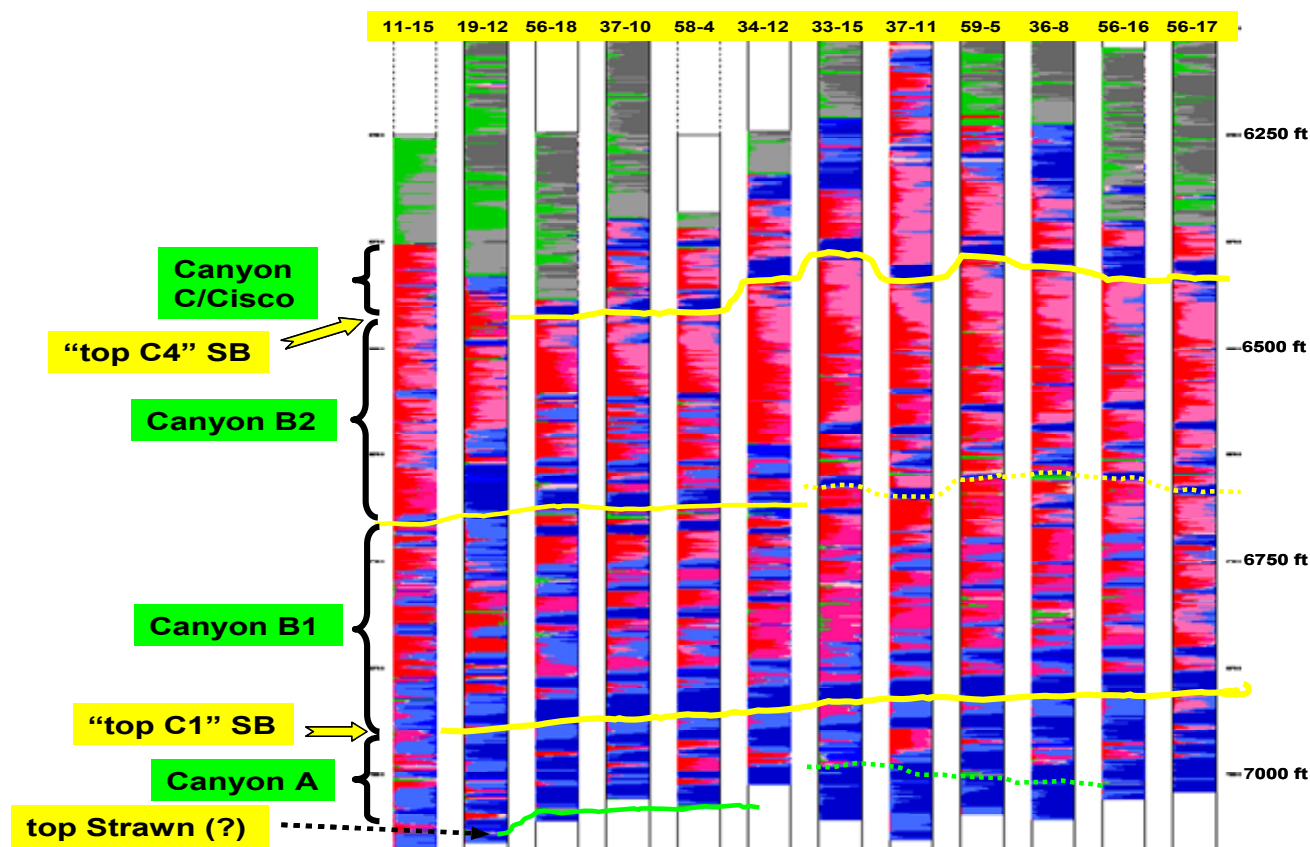


Figure 74: C1A Wells Oriented According to Relative Structural Position Using a Prominent Low RQ “Bed” Near the Bottom of the Studied Interval

There are two prominent horizons with relatively low average porosities and these were shown in Figure 72 as flow barriers with low RQ. These we suggest are the "top C1" (separating the Canyon A from the Canyon B) and the "top C4" (separating the Canyon B from the Canyon C/Cisco) sequence boundaries of Waite, and are illustrated in Figure 74 by the upper and lower yellow lines.

Within the Canyon B, there are numerous cycles of low porosity units overlain by high porosity units. In cored well 37-11, for instance, there are approximately 12 to 14 of these cycles. We interpret these to be 4th order cycles. One of these cycles is more prominent than the others and we use it to subdivide the Canyon B into a lower Canyon B1 and an upper Canyon B2. This subdivision is shown as the middle yellow line in Figure 74.

This provides a four-part division of strata between the Strawn and the Wolfcamp. Not surprisingly, the upper two of these divisions are coincident with the A and B and intervals depicted in Figures 70, 71, and 72, and the next division down section would be coincident with the C interval if the C were not arbitrarily stopped at its bottom but extended down to the low porosity horizon interpreted to be the "top C1" sequence boundary.

8.0 Conclusions

1. An approach for reservoir characterization based upon advanced pattern recognition techniques has been tested successfully in the Pennsylvanian-Permian reef carbonates (Cisco and Canyon Formations) of a subregion of the SACROC Unit, Horseshoe Atoll, Permian Basin, Texas.
2. A two-step pattern recognition procedure (soft-computing) was developed capable of efficiently generating core-scale P&P profiles at well locations where no core data existed. This procedure permitted to populate any well location with core-scale estimates of P&P and rock types facilitating direct application of geostatistical methods to build 3D reservoir models.
3. Using a database of 24 wells, 3 with whole core measurements of porosity and permeability through most of the section of interest, but with no core sedimentology description, no mineralogy or petrography, and no seismic, a probabilistic clustering procedure (using available well log and core data) was used to:
 - estimate profiles for RHOB and DT in wells that had only GR and NPHI logs;
 - estimate profiles for porosity and permeability in non-cored wells;
 - identify zones (electrofacies ~ flow units) with varying RQ based on variations in porosity and permeability;
 - evaluate (qualitatively) the degree to which the flow units can be correlated among wells;
 - identify a vertical cyclicity that is semi-pervasive throughout the area of study and relate this cyclicity to published information concerning the seismic sequence stratigraphy of the area.
4. Clustering analyses indicated that the SACROC carbonate section can be divided into a suite of closely-related flow units that have a "good" RQ (average porosity ~ 11-13 %) and into a suite of closely-related flow units that have a "poor" RQ (average porosity generally < 5 %). As interpreted from clustering analysis output, the contacts between these good and poor suites is generally rather sharp, as opposed to the generally gradational contacts that exist among the several flow units that comprise the good and poor suites. The relatively "sharp" contacts are interpreted to represent 3rd to 4th order sequence boundaries and, practically, they would likely act as significant barriers to vertical fluid flow. This has important implications for enhanced recovery performance.
5. Modest results were achieved with the application of soft-computing methodologies to the available information (no crosswell data at hand) in order to generate high-resolution reservoir characterization integrating multiple scale types of data. It was shown that 3D seismic information was not capable of capturing the degree of vertical variability of SACROC. As a consequence, available seismic information was unincorporated from posterior reservoir characterization tasks

6. The absence of crosswell seismic data in the available reservoir information constituted a serious drawback in reaching the goal of reconciling the gap between core data and 3D surface seismic information and in establishing a soft-computing methodology for high-resolution reservoir characterization integrating four different scale types of data.
7. Methodology here applied was successfully utilized to establish relationships between data from different origins and different scales, and to discriminate units with similar properties (i.e., clusters of log signatures, and core compositions), each corresponding to an identifiable stratigraphic unit or depositional environment. These classifications brought a strong component of geologic reality into the process that contrasts with other mathematical methodologies of data-driven applied in reservoir characterization tasks.
8. This integrated use of pattern-recognition and stratigraphic classification provided a meaningful technical advancement over conventional methodologies, specifically by providing the necessary vertical resolution for 3D reservoir characterization tasks, by reducing uncertainty, by providing results directly in the engineering terms needed for effective reservoir management, and by simplifying and streamlining the process, making it more time (and cost) efficient.
9. Advanced pattern recognition techniques can be either combined efficiently with other technical procedures (as it was shown in this work) or can be used in exclusivity on reservoir characterization studies depending on the quality and quantity of available information. This adaptability certainly represents an additional advantage to address reservoir characterization tasks.

9.0 References

1. Reeves, S.R., Mohaghegh, S.D., Fairborn, J.W., and Luca, G.; "Feasibility Assessment of a New Approach for Integrating Multiscale Data for High-Resolution Reservoir Characterization", paper SPE 77795 presented at the SPE Annual Technical Conference and Exhibition, San Antonio, Texas, Sept.29 to Oct.2, 2002.
2. Steen, J.W., and Reeves, S.R.; "Development, Testing and Validation of a Neural Model to Predict Porosity and Permeability from Well Logs, Grayburg Formation, McElroy Field, West Texas", Topical Report, U.S. Department of Energy, DE-FC26-01BC15357, July 2003.
3. Raines, M., Dobitz, J.K., and Wehner, S.C.; "A Review of the Pennsylvanian SACROC Unit", paper presented at West Texas Geological Society Fall Symposium, Midland, TX, Oct. 24-25, 2001.
4. Vest, E. L. Jr.; "Oil Fields of Pennsylvanian-Permian Horseshoe Atoll, West Texas" in Halbouty, Michael T. (ed.) *Geology of Giant Petroleum Fields*, AAPG Memoir Number 14. American Association of Petroleum Geologists, Tulsa Oklahoma, pp. 185-203, 1970.
5. Raines, M.: Personal Communication, 2007.
6. Raines, M., and Helms, W.: "Sample Size Impact on Permeability Values and Implications for Reservoir Models", West Texas Geological Society's Bulletin, Vol. 46, Number 4, March 2007.
7. Fairborn, J.; "Seismic Attribute Calculations. Final Report", WellSeismic Computing services, June 2006.
8. Eslinger, E.; "GAMLS (Geologic Analysis via Maximum Likelihood System). User's Manual", ericgeoscience.com, Eric Geoscience, Inc., 2000.
9. Perlovsky, L.I. and McManus, M.M.; "Maximum Likelihood Neural Networks for Sensor Fusion and Adaptive Classification", *Neural Networks*, Vol. 4, pp.89-102, 1991.
10. Perlovsky, L.I.; "Maximum likelihood adaptive neural system application to nontraditional signal classification", 3rd Annual Office of the Naval Technology Full Spectrum Review, Proceedings, 1, p. 121-146, NUWC, New London, CT., 1992.
11. Perlovsky, L.I.; "Computational concepts in classification: Neural Networks, Statistical Pattern Recognition, and Model Based Vision", *Journal of Mathematical Imaging and Vision*, Vol. 4 (1), p. 81-110, 1994.
12. Perlovsky, L.I.; "A model based neural network for transient signal processing", *Neural Networks*, Vol. 7 (3), p. 565-572, 1994.
13. Fukunaga, K.; "Introduction to Statistical Pattern Recognition", second edition, Boston, Academic Press, Inc., p. 591, 1990.
14. Cramer, H.; "Mathematical Methods of Statistics", Princeton, Princeton University Press, 1963.
15. Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T.; "Modeling of Data, in *Numerical Recipes. The Art of Scientific Computing*", Chapter 14, p. 498, New York, Cambridge University Press, 1990.

16. Eslinger, E., Burdick, B., and Cooper, J.; "Reservoir Characterization Using a Probabilistic Clustering Procedure", (in progress), 2007.
17. Hambalek , N., and González, R.J.; "Fuzzy Logic Applied to Lithofacies and Permeability Forecasting Case Study: Sandstone of Naricual Formation, El Furrial Field, Eastern Venezuela Basin", SPE 81078, prepared for presentation at the SPE Latin American and Caribbean Petroleum Engineering Conference held in Port-of-Spain, Trinidad, 27–30 April 2003.
18. Cuddy, S. J.; "Litho-Facies and Permeability Prediction from electrical logs using fuzzy logic", SPE Reservoir Evaluation & Engineering. Vol.3, Number 4, p. 319-324, 2000.
19. Eslinger, E.; "Procedures for Lithology Characterization and Probabilistic Upscaling (Curve "Blocking") Using Petrophysical and Core Data", AAPG National Meeting (poster session), Long Beach, CA, April 3 2007.
20. Waite, L.E.; "Upper Pennsylvanian Seismic Sequences and Facies of the Eastern and Southern Horseshoe Atoll, Midland Basin, West Texas, in Carbonate Sequence Stratigraphy", AAPG Memoir 57, Robert G. Loucks and J. Frederick Sarg (eds.), pp. 213-240, The American Association of Petroleum Geologists, Tulsa, OK, 1993.
21. Montgomery, D.C., and Runger, G.C.; "Applied Statistics and Probability for Engineers", second edition, John Wiley & Sons, Inc., New York, 1999.
22. Jensen, J.L., Lake, L.W., Corbett, W.M., and Goggin, D.J.; "Statistics for Petroleum Engineers and Geoscientists", first edition, Prentice Hall, Inc., New Jersey, 1997.
23. Johnson, D.E.; "Applied Multivariate Methods for Data Analysts", first edition, Brooks Cole Publishing Company, 1998.

10.0 List of Acronyms and Abbreviations

AI_log	=	Acoustic Impedance from RHOB and DT
bott(x,y)	=	bottom depth at location (x,y)
CMP	=	Cumulative Mode Probability
CorPerm	=	core permeability
CorPor	=	core porosity
DT	=	delta time
GAMLS	=	Geologic System via Maximum Likelihood System
GR	=	gamma ray
K0	=	permeability along master orientation line
K90	=	permeability at 90 degrees from K0 master line
Kv	=	vertical permeability
LC	=	Large Covariance
NPHI	=	neutron porosity
P&P	=	porosity and permeability
RHOB	=	bulk density
RQ	=	reservoir quality

Appendix A: Exploratory Statistical Analysis

Introduction

The data used for this study included core and log data from three (3) cored wells which sampled the whole SACROC unit; log data from twenty one (21) non cored wells inside the study area, and log data from other thirteen (13) non cored wells nearly surrounding this area. Two of the three SACROC cores covered the entire reservoir, from the deepest part of the “Wolfcamp” shale until reaching some feet below the original water-free producing oil-water contact. The other cored well covers all except the top twenty-six feet of the reservoir.

The three recently cored wells were well 11-15, well 19-12, and well 37-11. These wells are all located in the North Platform and were cored between 2004 and 2005, but only one of these wells (37-11) was near the center of the study area. Porosity and permeability data at one-foot intervals were available from these wells over the entire Canyon Reef interval (~ 800 ft). There were 26 older wells with some whole core data, but because of their age, the quality of the porosity and permeability analyses from these wells was not believed to be adequate for use in this study.

Continuous cores from single locations can provide a consistent, continuous, high quality one-dimensional datasets. In fact, the permeability measurements from a single continuous core provided an excellent dataset for this study because all zones present at the core location are seen in their natural relative thickness, abundances, and position. Additionally, the nature of bed and sequence boundaries (either permeable or sealing) and the permeability contrast between zones can be directly compared without problems caused by variations in measurement technique or equipment.

The measured core porosity was converted from percentage to fractional units, and the base 10 logarithm of all permeability parameters were directly used. In general terms, core data indicates that porosity tends to increase down-section in the upper half of the Canyon Reef, and then decreases through the lower part of the interval. Core permeability (more specifically its logarithm value) follows a similar trend. Due to the thin-bedded nature of flow units, these changes are not smooth with depth, and often well logs often do not have sufficient resolution to detect the rapid vertical changes in permeability.

Porosity and permeability data used here are measurements from whole core at one foot sampling intervals. One vertical and two horizontal whole core measurements of permeability were taken at each sampling depth-point. Rather than using the traditional horizontal permeabilities K_{max} and K_{min} , the horizontal measurements were recorded in two perpendicular directions and were denoted K_0 and K_{90} . A master orientation line was marked on cores where the core fit together with no rubble zones or spinners. This master line constituted the directional reference for whole core measurements of the K_0 permeability, and K_{90} was measured perpendicular to K_0 master line.

Different log suites were present for each well. These consisted of GR, NPHI, RHOB, DT, PEF, resistivity measurements, caliper, etc. Likewise, seismic attributes were computed⁷ and initially considered for application purposes on the data-driven methods. Well log data and seismic attributes were statistically explored; however, this appendix is focused on those results utilized for the achievement of the project objectives.

Core Data: Permeability Distributions

Statistical analyses of the two horizontal permeability measurements (K0 and K90), the average horizontal permeability (Kh), and vertical permeability (Kv) can help to determine the pattern of permeability distribution. The preliminary analyses were based on histograms, calculation of statistical indicators, and the study of correlations between permeability measurements.

In these analyses, we emphasized on the K0 values because no noticeable directional bias of this permeability data was identified. Both horizontal permeabilities have similar ranges, with a similar density profile. In deed, a direct comparison of logarithm of K0 and logarithm of K90 via crossplots shows that the values are well aligned around the concordance line ($y = x$). Figure A-1 shows corresponding crossplots of these parameters for well 11-15, well 19-12, well 37-11, and all cored wells together discriminated by colors (see next section for correlation coefficients).

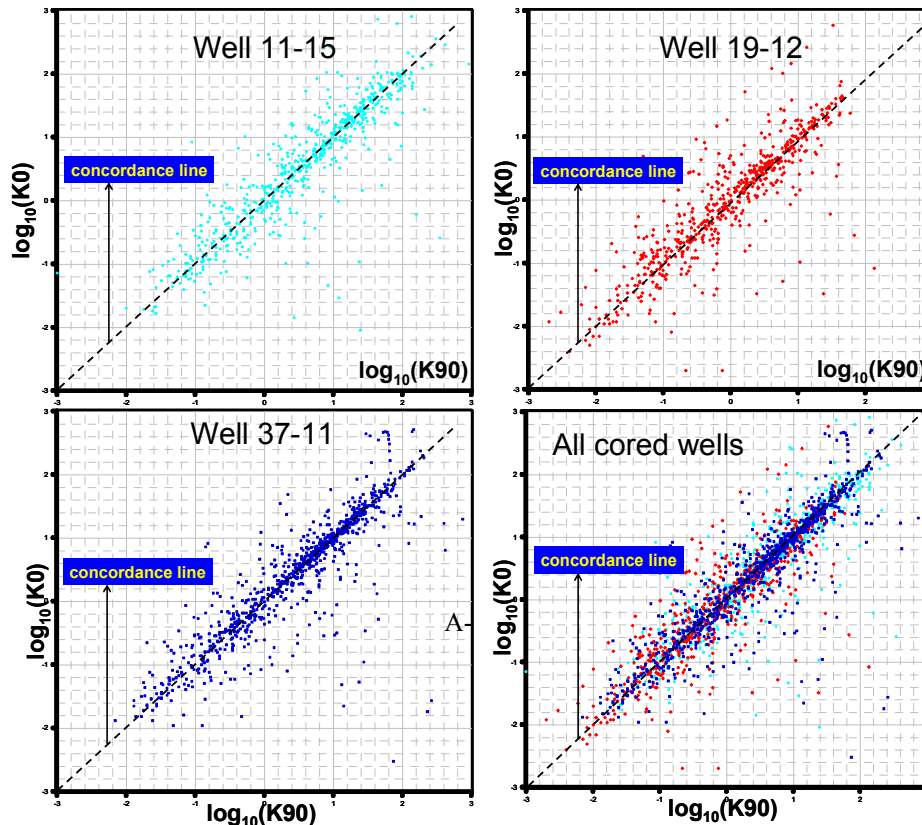


Figure A-1: Crossplots of $\text{Log}_{10}(K0)$ vs. $\text{Log}_{10}(K90)$ - Wells 11-15, 19-12, and 37-11

The absence of significant variation in the logarithm of permeability suggests that there is not a strong directional bias at these locations. Consequently, we considered K_0 to be representative of the horizontal permeability. We decided to assume this result for the whole North Platform because all three cores drilled have sampled all or almost all of the Cisco Formation. The coring program was designed to cover the entire reservoir at different locations thus permitting rock to be recovered from all zones. As an alternative, the arithmetic average of K_0 and K_{90} , denoted by K_h and conceived as a representative horizontal permeability, was eventually used for comparison purposes with other datasets and visual simplification of plots.

Histograms of horizontal permeabilities (K_0 , K_{90} , and K_h), and vertical permeability K_v are illustrated in Figures A-2, A-3, A-4, and A-5 for wells 11-15, 19-12, 37-11, and all cored wells respectively. The histograms exhibit distributions positively skewed with few outlier values. In addition, most permeability values are below 50 md. This can be better observed from the first vertical bar of the K_{90} histogram of all cored wells (Figure A-5). This bar indicates 94% of K_{90} values between 0 and 51 md. Similar analyses for other types of permeability in Figure A-5 lead to similar conclusions: a very high concentration of low values of permeability. The corresponding probability plots displaying the actual data values versus expected lognormal values are also included in Figures A-2, A-3, A-4, and A-5. The excellent alignments of all probability plots to the theoretical straight line (where the standard lognormal distribution appears) reveal the lognormal nature of these permeability distributions.

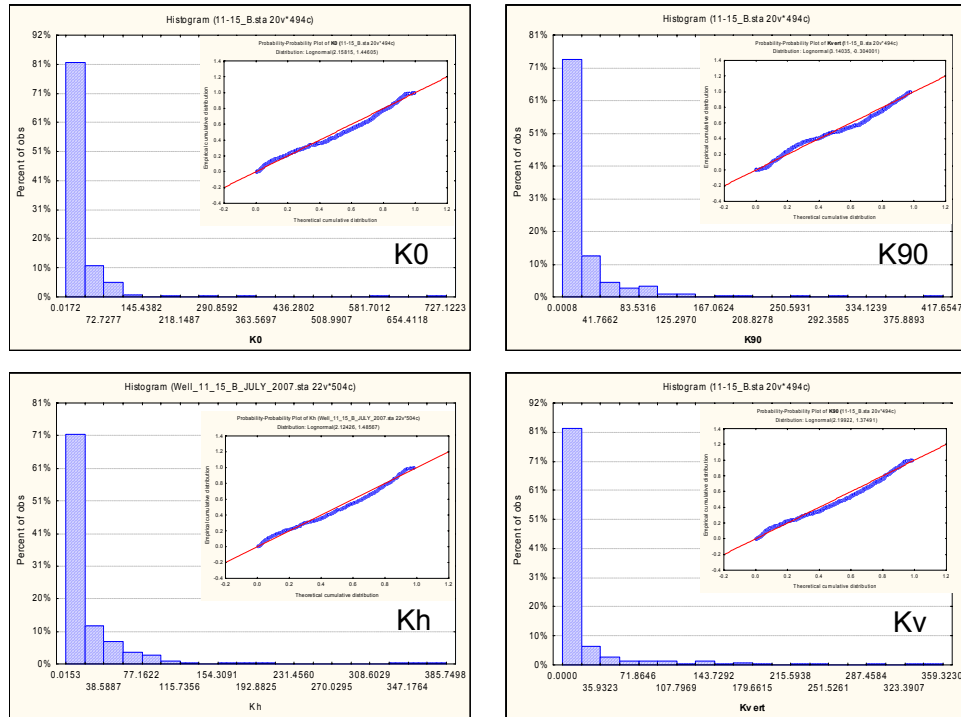


Figure A-2: Permeability Histograms and Corresponding Probability Plots - Well 11-15

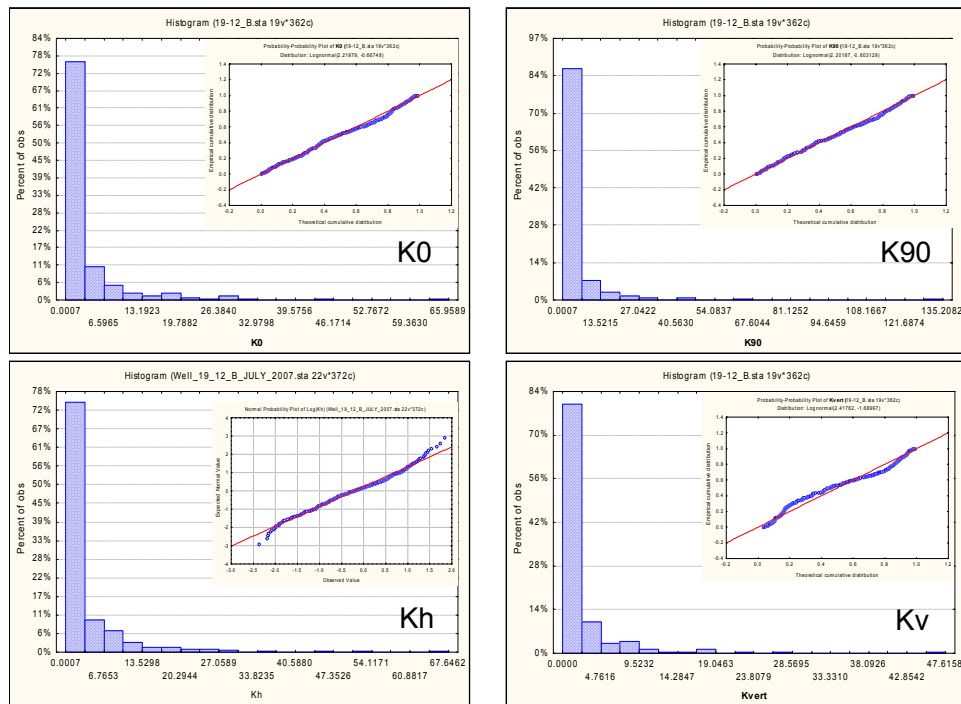


Figure A-3: Permeability Histograms and Corresponding Probability Plots - Well 19-12

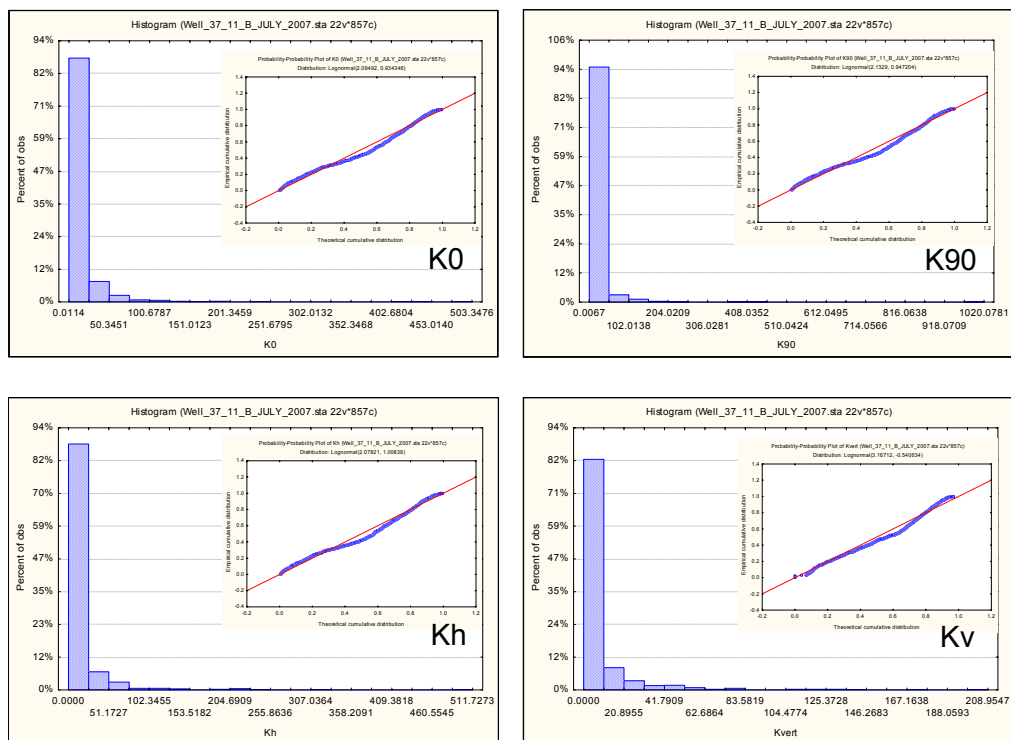


Figure A-4: Permeability Histograms and Corresponding Probability Plots - Well 37-11

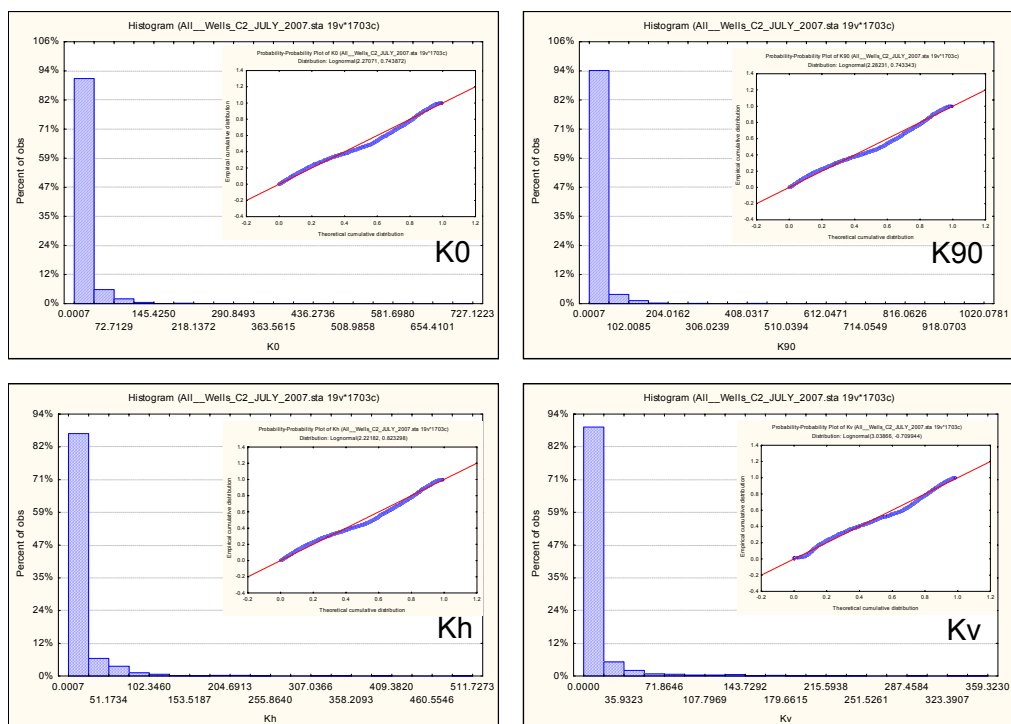


Figure A-5: Permeability Histograms and Corresponding Probability Plots - All Wells

A summary of the main statistical indicators,^{14,21,22} representing each set of measurements are presented in Table A-1 for the data set constituted by all three cored wells. This includes values for the arithmetic, geometric and harmonic mean, the standard deviation, the coefficient of variation together with skewness, percentiles 10 and 90. Results indicate that all different means of vertical permeability are smaller than the corresponding means of horizontal permeability measurements. All values of percentile 90 show that 90 % of sampled values of any kind of permeability are less than 34 md.

**Table A-1: Statistical Parameters of the Core Permeability and Porosity Measurements
(Wells 11-15, 19-12, and 37-11)**

Statistical Parameters	K0	K90	Kh	Kv	Porosity
Mean	13.28	13.38	13.33	8.56	0.0815
Geometric Mean	2.10	2.10	2.28	0.49	0.0551
Harmonic Mean	0.13	0.14	0.19	0.00	0.0272
Median	3.16	3.05	3.29	0.58	0.0802
Minimum	0.00	0.00	0.00	0.00	0.0009
Maximum	727.12	1020.08	511.73	359.32	0.2274
Lower Quartile	0.37	0.40	0.43	0.05	0.0293
Upper Quartile	11.94	11.94	12.51	4.95	0.1271
Percentile 10	0.09	0.08	0.10	0.01	0.0093
Percentile 90	32.13	31.22	33.63	19.32	0.1545
Standard Deviation	38.30	39.03	32.86	26.62	0.0548
Coefficient of Variation	2.88	2.92	2.47	3.11	0.6725
Skewness	10.22	13.27	7.02	6.63	0.1967

**Table A-2: Statistical Parameters of the Core Permeability and Porosity Measurements
(Well 11-15)**

Statistical Parameters	K0	K90	Kh	Kv	Porosity
Mean	21.92	19.68	20.80	15.72	0.0871
Geometric Mean	4.25	3.95	4.43	0.74	0.0641
Harmonic Mean	0.48	0.21	0.51	0.00	0.0344
Median	5.53	5.53	5.89	0.98	0.0820
Minimum	0.02	0.00	0.02	0.00	0.0018
Maximum	727.12	417.65	385.75	359.32	0.2274
Lower Quartile	0.84	0.93	1.03	0.06	0.0420
Upper Quartile	21.34	22.62	23.25	9.30	0.1266
Percentile 10	0.17	0.14	0.18	0.01	0.0163
Percentile 90	59.26	54.29	56.41	39.12	0.1632
Standard Deviation	53.98	37.84	41.41	42.11	0.0542
Coefficient of Variation	2.46	1.92	1.99	2.68	0.6223
Skewness	8.36	4.88	5.13	4.56	0.3341

When similar tables are developed for each individual well, we find that data coming from well 19-12 is somewhat different when it is compared to the other two cored wells. These differences are mainly due to local geological differences at this location associated possibly to different processes of erosion and diagenesis, and later geochemical processes (karst overprints, dissolution, and precipitation)^{3,5}. Tables A-2, A-3, and A-4 show the same group of statistical indexes for wells 11-15, 19-12, and 37-11 respectively.

**Table A-3: Statistical Parameters of the Core Permeability and Porosity Measurements
(Well 19-12)**

Statistical Parameters	K0	K90	Kh	Kv	Porosity
Mean	3.10	3.62	3.36	1.78	0.0649
Geometric Mean	0.51	0.55	0.59	0.18	0.0382
Harmonic Mean	0.04	0.06	0.07	0.03	0.0209
Median	0.49	0.58	0.60	0.14	0.0431
Minimum	0.00	0.00	0.00	0.00	0.0045
Maximum	65.96	135.21	67.65	47.62	0.1979
Lower Quartile	0.12	0.10	0.13	0.02	0.0123
Upper Quartile	3.17	3.15	3.41	1.76	0.1116
Percentile 10	0.03	0.03	0.04	0.01	0.0079
Percentile 90	9.25	10.24	9.56	5.57	0.1537
Standard Deviation	6.59	9.84	7.18	4.25	0.0562
Coefficient of Variation	2.12	2.72	2.14	2.39	0.8662
Skewness	4.55	8.22	4.58	5.54	0.5793

**Table A-4: Statistical Parameters of the Core Permeability and Porosity Measurements
(Well 37-11)**

Statistical Parameters	K0	K90	Kh	Kv	Porosity
Mean	12.58	13.86	13.22	7.30	0.0855
Geometric Mean	2.55	2.58	2.74	0.58	0.0589
Harmonic Mean	0.30	0.28	0.32	0.00	0.0274
Median	3.82	3.91	4.10	0.99	0.0894
Minimum	0.01	0.01	0.01	0.00	0.0009
Maximum	503.35	1020.08	511.73	208.95	0.2266
Lower Quartile	0.54	0.55	0.58	0.08	0.0359
Upper Quartile	12.11	12.91	12.87	5.95	0.1297
Percentile 10	0.12	0.12	0.13	0.01	0.0089
Percentile 90	29.00	29.02	29.87	19.40	0.1509
Standard Deviation	33.89	46.02	32.78	18.33	0.0534
Coefficient of Variation	2.69	3.32	2.48	2.51	0.6244
Skewness	9.05	14.44	7.59	5.45	-0.0065

As can be noticed, statistical parameters (extremes, means, quartiles, and percentiles) of well 19-12 are lower than the corresponding numbers for wells 11-15 and 37-11. Statistical parameters of wells 11-15 and 37-11 tend to be more similar to each other; however, better values belong to well 11-15.

A statistical parameter commonly utilized to define various classes of heterogeneity is the coefficient of variation (CV)²². This coefficient is defined as the ratio of the standard deviation to the mean, and is used in describing the variability of any population of data values. Calculating CV for permeability data, Jensen et al²² established ranges of variation for CV in terms of the scale and depositional environment. In particular, if CV is greater than 1.0, permeability distributions are considered very heterogeneous. Carbonates are at the top of Jensen²² scale with CV values that can pass the 3.0 threshold. This is the case for permeability distributions of the cored wells of SACROC where all tables show CV values greater than 1.0.

The CV, as a measure of relative permeability variation, can also be used for comparisons of the variability of different intervals, facies and formations. As shown in Table A-1 (all cored wells), the CV for vertical permeability is greater in all cases than corresponding values for horizontal permeabilities, implying, that for this reservoir, a greater variability in the vertical permeability distribution exists. This is also corroborated when CV values of Kh and Kv are compared in all tables. However, each location presents its own particularities; for instance, CV values of K90 are greater than other permeability CV values in wells 19-12 and 37-11 due to the presence of high values of this particular parameter in relation with their respective mean values.

Skewness is a statistical indicator of the deviation of a distribution from symmetry. As shown in Tables A-1, A-2, A-3, and A-4, all permeability distributions show positive values for this parameter, which is indicative of the positively skewed nature of the distributions. As a reference, the bias of the perfectly symmetrical normal distribution is zero.

Permeability Correlations

We calculated Pearson's linear correlation coefficients (R) between pairs of $\log_{10}(K_0)$ and $\log_{10}(K_{90})$ measurements to determine the correlation between these parameters (and others). For wells 11-15, 19-12, and 37-11 the crossplot for the pair ($\log_{10}(K_{90})$, $\log_{10}(K_0)$) are shown in Figure A-1. The aligned distribution of points around the concordance line and the very high positive correlation between the two horizontal measurements $\log_{10}(K_0)$ and $\log_{10}(K_{90})$ are indicative of a close similarity among these two types of horizontal permeability values. The correlation coefficients (R) between $\log_{10}(K_0)$ and $\log_{10}(K_{90})$ measurements for wells 11-15, 19-12, and 37-11 are respectively equal to 0.88, 0.85 and 0.87.

On the other hand, the crossplots for the pair ($\log_{10}(K_v)$, $\log_{10}(K_0)$) and ($\log_{10}(K_v)$, $\log_{10}(K_{90})$) for wells 11-15, 19-12, and 37-11 are now shown in Figures A-6 and A-7 respectively. As mentioned before, similar correlation patterns were registered on all these wells. We can see that correlations between vertical and horizontal measurements are more modest although they can still be considered good. The correlation coefficients (R) between $\log_{10}(K_0)$ and $\log_{10}(K_v)$ measurements for wells 11-15, 19-12, and 37-11 are respectively 0.75, 0.74 and 0.78. Likewise, for wells 11-15, 19-12, and 37-11, the corresponding correlation coefficients

between $\log_{10}(K_{90})$ and $\log_{10}(K_v)$ were 0.76, 0.75, and 0.76. Scrutiny of data in these figures indicate that all clusters of points are mostly concentrated above the concordance line which reveals that the magnitude of vertical measurements of permeability tend to be definitively less than the corresponding horizontal values of permeability.

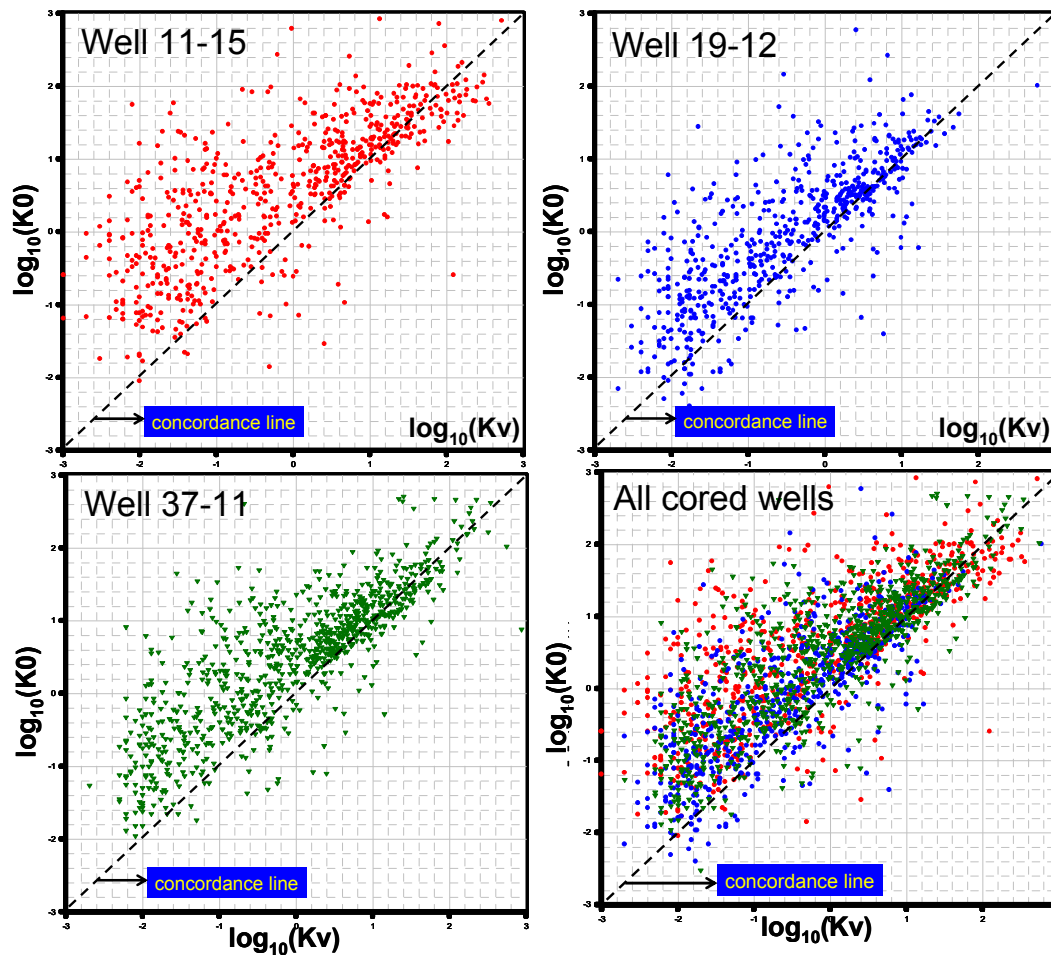
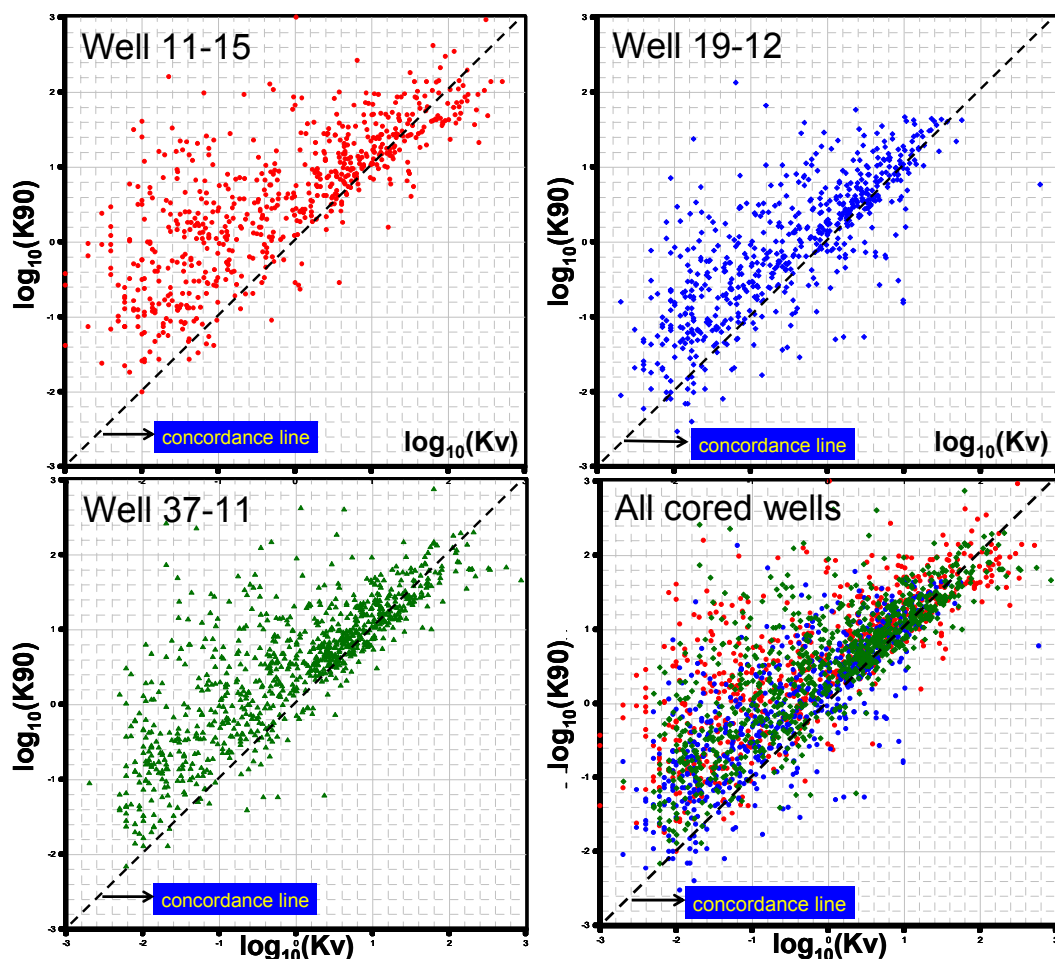


Figure A-6: Crossplots $\log_{10}(K_0)$ vs. $\log_{10}(K_v)$ at All Cored Wells
Vertical and Horizontal Scales Range from -3.0 to 3.0



**Figure A-7: Crossplots $\text{Log}_{10}(\text{K}_{90})$ vs. $\text{Log}_{10}(\text{K}_v)$ at all Cored Wells
Vertical and Horizontal Scales Range from -3.0 to 3.0**

Different geological factors control the horizontal and vertical permeability in a reservoir, and this can explain the more modest correlation between the horizontal permeabilities and K_v (logarithm). For instance, bedding planes and laminations are dominant controls favoring horizontal permeability, but these same features could eventually work as barriers to the vertical flow in the reservoir²². Nevertheless, the presence of vertical and/or inclined fractures or bioturbation processes may change this general horizontal behavior enhancing the flow in the vertical direction.

About Porosity

Summaries of statistical indicators for porosity are presented in Tables A-1, A-2, A-3, and A-4. The same data are also summarized in the form of histograms in Figure A-8. Porosity distribution of well 19-12 is noticeably different when it is compared with the other histograms. As mentioned before, local geological differences^{3,5} at this location can explain these disparities. A strong component of low values of porosity seen in this histogram are possibly associated to a larger presence of shales and modes here described as “poor” limestones (see section 7.6 for a

more detailed discussion). Histogram of well 37-11 suggests the co-existence of two populations, one representing “poor” limestones (and shales), and “good” limestones. The “cut-off” porosity value separating these two populations appears to be approximately 10 %. This bimodal nature of the porosity distribution can also be detected (although less pronounced) in the histogram for porosity values of all cored wells.

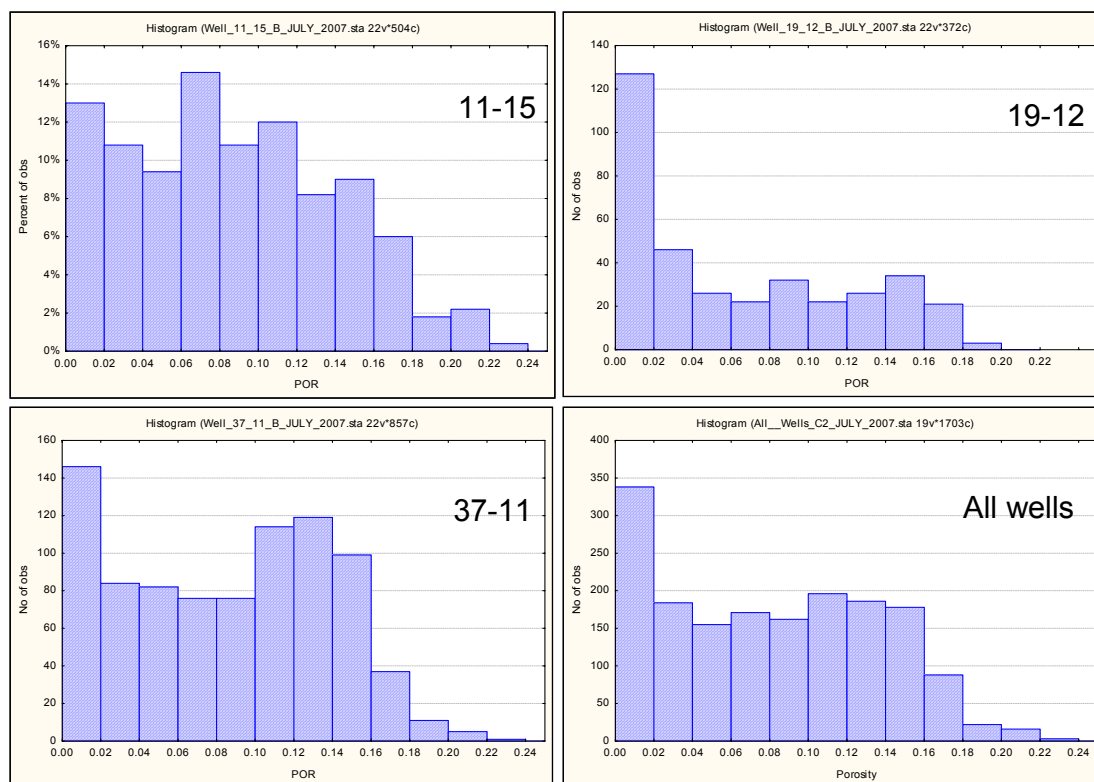


Figure A-8: Porosity Histograms at all Cored Wells
Porosity Values Range from 0.0 to 0.24

Permeability and Porosity Correlations

The relationship between whole core measurements of permeability (K0) and porosity for the three cored wells is shown in Figure A-9. Despite linear-like "cloud" trends, there is not a convincing linear relationship between the logarithm of permeability (K0) and porosity for any of the wells. This is not unexpected because the reservoir is made of different depositional facies, each of which might be expected to have different porosity-permeability relationships.

Mathematical analyses were performed to investigate the levels of correlation between $\log_{10}(K0)$ and porosity. We calculated two different central measures for porosity and logarithm of permeability (K0) for all samples. These central measures were arithmetic mean values and median values corresponding to adopted size-windows around the sample depth. For instance, if a 10-point arithmetic mean is used, a moving window consisting of 5 points above and 5 points below a sampled depth is used to compute the arithmetic mean value of the 10 points at the

center of the window. Likewise, if a 10-point median filter⁸ is selected, a moving window is made for 10 contiguous points, and the mid-point is replaced by the median of all 10 points.

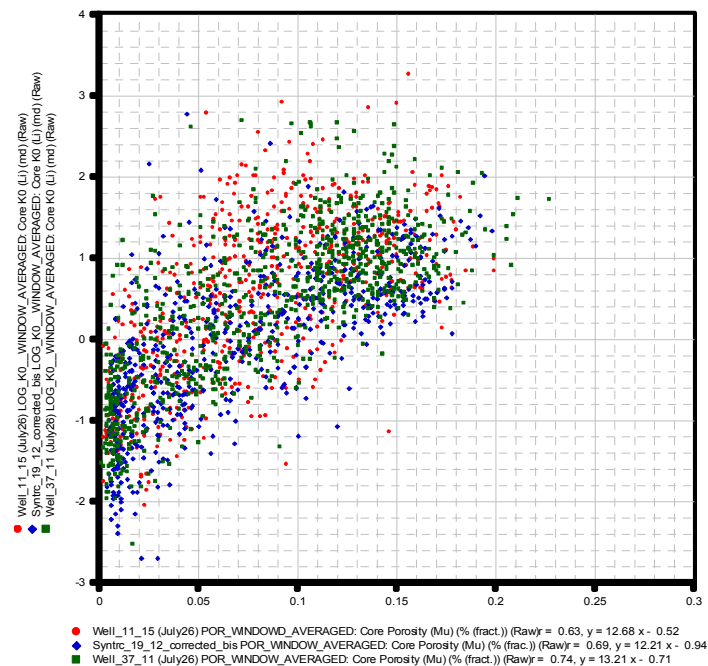


Figure A-9: Superimposed Crossplots of $\text{Log}_{10}(\text{K0})$ vs. Porosity Wells 11-15 (Red), 19-12 (Blue), and 37-11 (Green)

Table A-5 shows correlation coefficients for these experiments for the three cored wells. The middle three rows in the upper section show results using mean values within windows of various sizes, with the average values present in bold. The middle three rows of the lower section show results using median values within windows of various sizes, with the average values in bold the bottom row.

Table A-5: $\text{Log}_{10}(\text{K0})$ vs. Porosity: Correlation Coefficients for Different Window-Averaged Values

MEAN	Well	Raw	Mean-10'	Mean-20'	Mean-40'	Mean-100'
	11-15	0.640	0.680	0.660	0.680	0.860
	19-12	0.690	0.770	0.800	0.840	0.930
	37-11	0.740	0.800	0.800	0.810	0.850
	Avg.	0.687	0.750	0.753	0.777	0.880
MEDIAN	Well	Raw	Median-10'	Median-20'	Median-40'	Median-100'
	11-15	0.640	0.680	0.630	0.640	0.810
	19-12	0.690	0.870	0.890	0.910	0.950
	37-11	0.740	0.810	0.820	0.820	0.810
	Avg.	0.687	0.787	0.780	0.790	0.857

We conclude from inspecting the average values in rows 5 and 10, that these parameters experience a slight increase in correlation as the window sizes are incremented. Additionally,

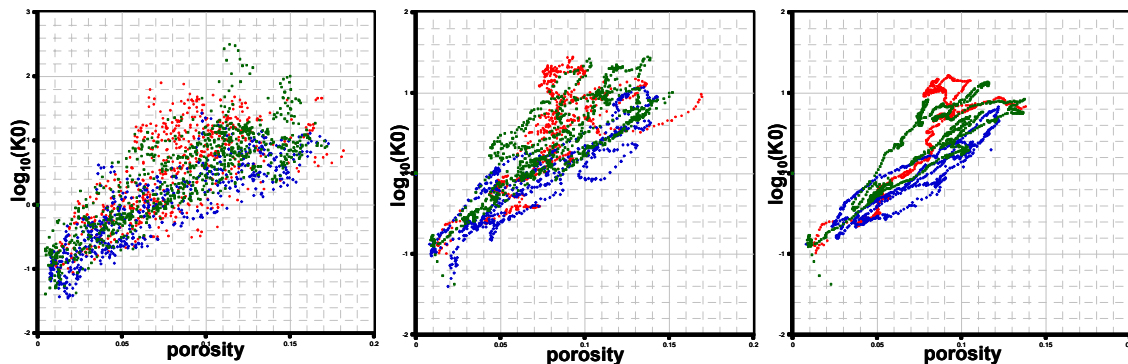
correlation coefficients were calculated for the median procedure when the window sizes were 15 ft (vertical resolution of the simulation grid block), and 112 ft (horizontal resolutions of the simulation grid block). The results, given in Table A-6, are consistent with results of Table A-5.

Table A-6: Correlation Coefficients, Mean and Median

WELL	Median-15 ft	Median-112 ft
11-5	0.6300	0.8500
19-12	0.8800	0.9500
37-11	0.8100	0.8100
Avg.	0.7733	0.8700

Figure A-10 shows three different crossplots of $\log_{10}(K_0)$ vs. porosity when the original values were averaged using the arithmetic mean procedure above described. The different colors are associated with the different cored wells: red for well 11-15, blue for well 19-12, and green for well 37-11. At left, the utilized window size was 10 ft; at center, the window size was of 40 ft; finally at right, results belong to a window size of 100 ft. All crossplots show a good correlation between these two parameters at increasing scales. Similar results were derived analyzing crossplots of $\log_{10}(K_0)$ vs. porosity when the original values were “averaged” utilizing the median filter⁸ procedure. Figure A-11 depicts these results under the same characteristics of the Figure A-10. Finally, Figure A-12 shows results associated to the median filter procedure when window sizes of 15 ft (left), and 112 ft (right) were applied. These sizes coincide with the vertical and horizontal dimensions of the grid block utilized in the grid that represents the selected SACROC study area for characterization and simulation purposes.

The problem of averaging permeability at different scales is a controversial subject, and awaits a universally accepted solution. Different aspects of this problem are related to the type of reservoir and type of fluids present in the reservoir. We have utilized results from Table A-5 as guidelines for developing a coherent methodology combining data-driven techniques with geostatistical methods to characterizing a reservoir. The topic of averaging permeability, with its consequent effect over the possible porosity-permeability correlation at different volume, was beyond the scope of this project. Currently, there are several institutions worldwide (academic and industry related), giving attention to this important subject.



**Figure A-10: Log₁₀(K₀) vs. Porosity for Wells 11-15 (Red), 19-12 (Blue), and 37-11 (Green)
Window Averaged 10' (Left), 40' (Center), and 100' (Right)**

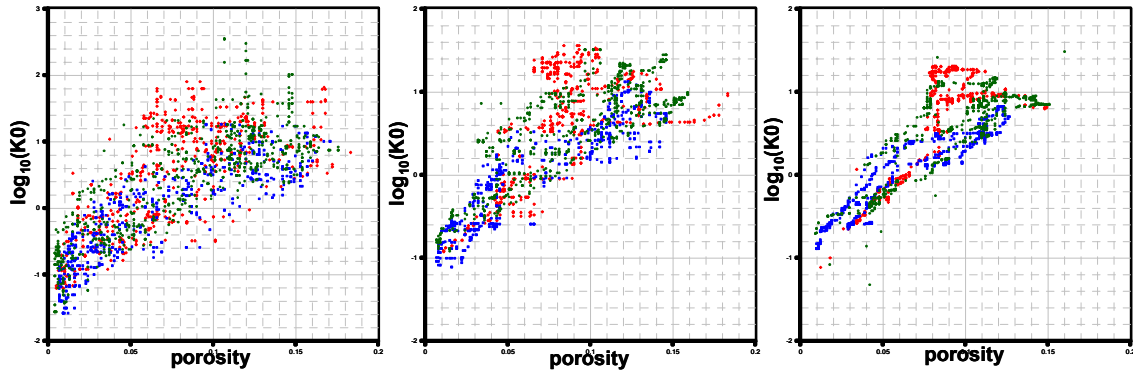


Figure A-11: $\text{Log}_{10}(\text{K0})$ vs. Porosity for Wells 11-15 (Red), 19-12 (Blue), and 37-11 (Green) Window Median Filtered 10' (Left), 40' (Center), and 100' (Right)

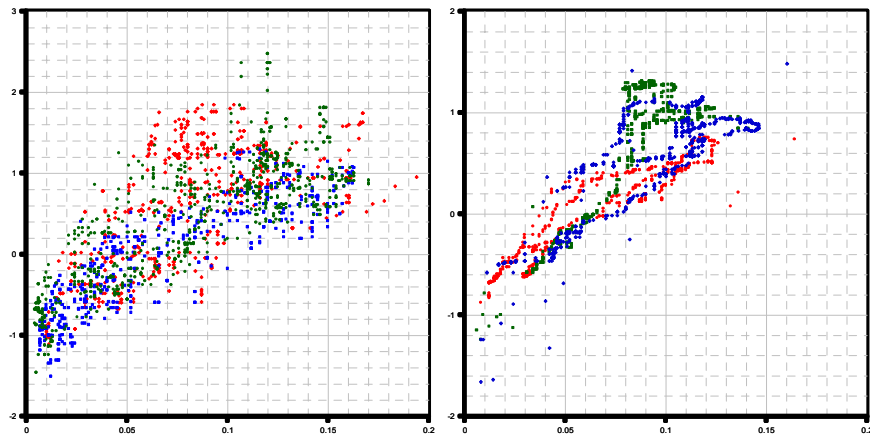


Figure A-12: $\text{Log}_{10}(\text{K0})$ vs. Porosity for Wells 11-15 (Red), 19-12 (Blue), and 37-11 (Green) Median Filtered 15' at Left; Median Filtered 112' at Right

Principal Component Analysis

Principal Component Analysis (PCA) is a linear dimensionality reduction technique with which orthogonal directions of maximum variance in the original data are identified. Data is projected into a new space of lower-dimensionality and constituted of a subset of the highest-variance components²³. The collected data in a reservoir is expressed as different variables that can be correlated, and also can be very large in number. In consequence, the interpretation of the data and the detection of possible structures or patterns are often hard to distinguish. PCA can make these two tasks easier by transforming the original variables to a smaller number of uncorrelated variables.

Utilizing PCA, it is possible to summarize the data effectively and to reduce the dimensionality of the data without a significant loss of information. PCA is used as an alternative form of displaying the data, thus permitting a better understanding of its possible structures

without impacting the information essences. Additionally, due to the total variance in a data set can be expressed as the sum of the individual variances associated with each principal component, only the first few principal components that explain most of the original variables variability is often useful to reveal the structure and patterns in the data. This action reduces the dimensionality of the problem and complexity for the clustering analysis purposes.

Considering the data quality, the field-wide availability, and their functionality for seismic calibration a suite of well logs was selected for the analysis. In this field, we considered the following nine (9) well logs: bulk density (RHOB), Caliper (Cal), Deep Laterolog Resistivity (LLD), Density Porosity (DPHI), Delta Time (DT), Gamma Ray (GR), Photo Electric Factor (PEF), Neutron Porosity (NPHI), and Shallow Laterolog Resistivity (LLS).

With the objective of creating a Log-to-Core model to predict core-scale P&P values at wells with only electrical logs, we focused our analyses on data belonging to the three cored wells 11-15, 19-12 and 37-11. Individual analyses for each cored well data set were carried out, however results are here omitted. Because all these data sets were included for clustering analysis purposes and the creation of the Log-to-Core model, we decided to analyze an integrated data set of these three cored wells.

Table A-6 presents a summary of selected statistical parameters of the well logs based on data sets of wells 11-15, 19-12, and 37-11.

Table A-6: Statistical Parameters of the Well Logs for Wells 11-15, 19-12, and 37-11

Descriptive Statistics	RHOB	Cal	LLD	DPHI	DT	GR	PEF	NPHI	LLS
Mean	2.58	8.29	139.25	0.0803	58.51	17.37	6.68	0.0751	111.19
Median	2.58	8.78	43.43	0.0800	58.50	14.99	6.56	0.0700	34.45
Minimum	2.32	6.04	6.02	-0.2120	47.96	4.57	4.03	-0.0060	4.25
Maximum	3.05	9.63	2281.77	0.2400	90.06	74.84	20.65	0.2400	1722.19
Lower Quartile	2.50	8.64	21.27	0.0270	53.15	9.72	6.13	0.0340	17.22
Upper Quartile	2.67	8.94	119.50	0.1300	63.38	22.43	7.03	0.1120	95.52
Percentile 10	2.46	6.10	14.70	0.0000	50.01	7.20	5.75	0.0090	11.81
Percentile 90	2.71	9.12	388.96	0.1600	66.99	29.83	7.62	0.1440	299.75
Standard. Deviation	0.10	1.15	246.48	0.0609	6.27	10.08	1.03	0.0504	201.26
Skewness	0.01	-1.34	3.45	-0.0014	0.22	1.55	3.48	0.3645	3.60

Likewise, Table A-7 presents the correlation matrix of all well log parameters here considered. These values are a measure of the possible linear relation between two variables, and can range from -1.0 to +1.0. In particular, this correlation coefficient is the well known Pearson correlation coefficient (R). The value of -1.0 represents a perfect linear negative correlation while a value of +1.0 will represent a perfect positive linear correlation. A null value is interpreted as a lack of linear correlation.

Table A-7: Correlation Matrix of Well Log Parameters for Wells 11-15, 19-12, and 37-11

	RHOB	Cal	LLD	DPHI	DT	GR	PEF	NPHI	LLS
RHOB	1.0	-0.1685	0.5158	-0.9981	-0.8803	0.3487	0.0289	-0.8738	0.5249
Cal	-0.1685	1.0	-0.1075	0.1680	0.1810	-0.3197	0.1942	0.1951	-0.1980
LLD	0.5158	-0.1075	1.0	-0.5136	-0.5463	0.1644	-0.0659	-0.4966	0.9122
DPHI	-0.9981	0.1680	-0.5136	1.0	0.8779	-0.3487	-0.0267	0.8717	-0.5226
DT	-0.8803	0.1810	-0.5463	0.8779	1.0	-0.3344	0.0401	0.8798	-0.5902
GR	0.3487	-0.3197	0.1644	-0.3487	-0.3344	1.0	-0.2668	-0.2974	0.1985
PEF	0.0289	0.1942	-0.0659	-0.0267	0.0401	-0.2668	1.0	0.0490	-0.1411
NPHI	-0.8738	0.1951	-0.4966	0.8717	0.8798	-0.2974	0.0490	1.0	-0.5508
LLS	0.5249	-0.1980	0.9122	-0.5226	-0.5902	0.1985	-0.1411	-0.5508	1.0

Table A-7 shows the resistivities LLD and LLS are highly correlated with a correlation coefficient of 0.9122. This allows the discarding of one of these parameters when multivariate statistical methods or other less conventional data-driven methods are applied for clustering and regression purposes²³. Other good correlations (as expected) exist between RHOB, DPHI, and NPHI.

PCA is applied in order to obtain the principal components (or factors) from the well log data set. In Table A-8 it can be observed how much (in percentage) a factor “explains” the total variability in the original well log variables. For instance, we can attribute the 52.80% of the total variability to the first factor. Analyzing the cumulative percentage column of this table, we conclude that almost 96% of the total variability can be “explained” by the first five factors allowing the selection of those well log variables better correlated with these five initial factors. Figure A-13 illustrates the behavior of variance percentage associated to each individual factor.

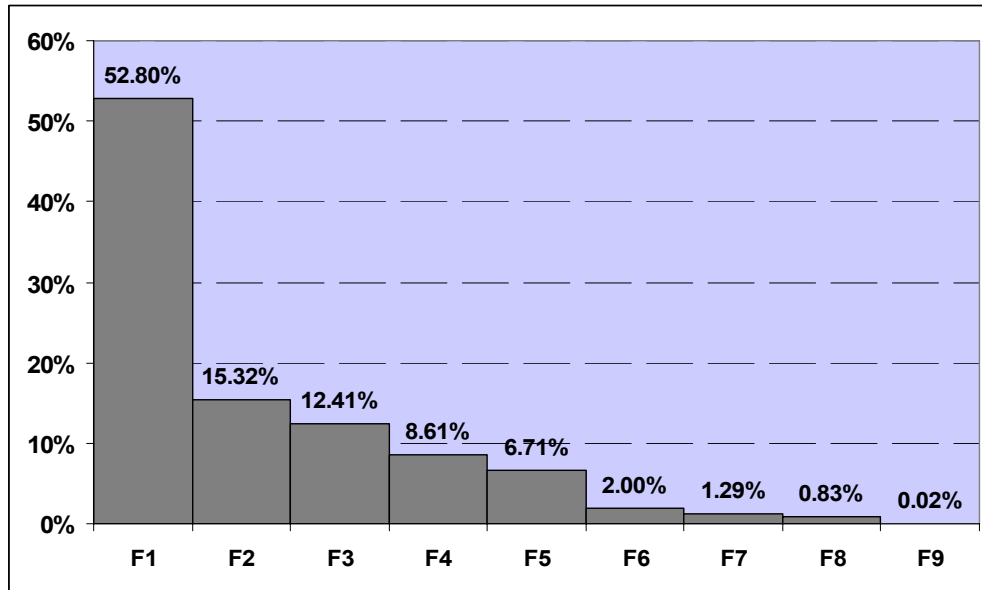


Figure A-13: Bar Plot- Percentage of Variance Contribution of Each Factor

Table A-8: Variance Percentage by Factor for Wells 11-15, 19-12, and 37-11

Factor	Variance Percentage	Cumulative Percentage
F1	52.80%	52.80%
F2	15.32%	68.13%
F3	12.41%	80.54%
F4	8.61%	89.15%
F5	6.71%	95.86%
F6	2.00%	97.86%
F7	1.29%	99.15%
F8	0.83%	99.98%
F9	0.02%	100.00%

Table A-9: Variable Contributions by Factor for Wells 11-15, 19-12, and 37-11

Factor	RHOB	Cal	LLD	DPHI	DT	GR	PEF	NPHI	LLS
F1	0.18200	0.01610	0.10922	0.18140	0.17938	0.03860	0.00185	0.17204	0.11940
F2	0.01744	0.29442	0.00313	0.01714	0.00737	0.25406	0.39825	0.00763	0.00056
F3	0.05796	0.01141	0.36286	0.05855	0.01942	0.08040	0.04234	0.03179	0.33527
F4	0.00358	0.57338	0.00928	0.00386	0.00660	0.01075	0.37383	0.00641	0.01231
F5	0.00462	0.09891	0.04165	0.00476	0.01354	0.60684	0.17013	0.04858	0.01098
F6	0.21631	0.00226	0.03403	0.22887	0.20004	0.00717	0.01235	0.27907	0.01991
F7	0.00045	0.00062	0.00027	0.00056	0.56383	0.00219	0.00033	0.42950	0.00225
F8	0.01070	0.00290	0.43956	0.01185	0.00979	0.00000	0.00091	0.02497	0.49931
F9	0.50694	0.00000	0.00000	0.49300	0.00003	0.00000	0.00001	0.00002	0.00000

Table A-9 depicts the results of the PCA applied to these well log variables. Each row (each factor) shows a fractional value associated to each well log variable. This fraction expresses the corresponding contribution (in variability) of this specific variable in the factor considered. We can see that RHOB, DPHI, DT, and NPHI are the most influential variables. Because many multivariate statistical techniques and data-driven methods are sensitive to the high co-linearity between predictor variables²³, we decide to exclude DPHI of the final set of well log variables due to its high linear correlation with RHOB (-0.9981). Due to its importance to seismic calibration purposes, RHOB and DT were included in the final selection of utilized well log parameters (despite of their modest presence at wells of SACROC). NPHI was the “ideal” parameter: it was a driver variable and had presence in all wells. Resistivity data could provide input for the first factor, and we initially considered the LLD as a final variable. However, because so few wells at SACROC have this log available, its use was precluded.

Table A-9 shows in the second row, that Cal, GR, and PF have the most impact in the variability of the second factor. Due to the significance of the second factor, and the fact that GR was present in all wells, its inclusion in the final set of well log variables was justified. PEF presented the same conditions of LLD, but it is rarely available consequently, it was not included. The caliper had a modest presence at SACROC wells, however anomalies within the data that was not possible to explain, lead to its exclusion. Thus the final suite of well log parameters we considered was the GR, NPHI, RHOB, and DT.

Appendix B: Seismic Attributes Computed from Well Log and Seismic Data, SACROC Oilfield, Texas

John Fairborn
WellSeismic Computing Services

Introduction:

This report summarizes the seismic attribute calculations of real and synthetic seismic traces from the SACROC oilfield in West Texas. The synthetic traces were computed from the sonic and density logs of about 81 wells. Seismic attributes for the synthetic traces were computed alongside corresponding attributes from 3-d, migrated seismic traces located near each of the wells. The attributes were written to las-formatted ascii files for multi-variate, neural-network-based, characterization of the SACROC reservoir.

A total of 16 seismic attributes were computed, which comprise most of those commonly used in attribute analyses. An excellent summary of these attributes including their potential geologic significance has been written by Turhan Taner and can be found on the website: www.rocksolidimages.com/pdf/attrib_revisited.htm#_494264900.

Three sets of output files were produced for each well. The first set contains the unfiltered seismic attributes computed from the synthetic trace. The second contains the same seismic attributes but filtered to more closely match the frequency bandwidth of the real data. The third set contains the attributes computed from the real seismic trace nearest the well. The seismic data is part of a modern 3-d survey recorded within the SACROC field and is of excellent quality.

Synthetic Traces

The synthetic traces were computed using a one-dimensional algorithm derived by Goupillaud and originally published in the 1966 volume of Geophysics. It computes the primary reflections, transmission loss, and all possible multiple reflections for a P-wave normally incident on a series of horizontal layers of equal travel-time thickness. The reflection coefficient at each hypothetical interface was computed from the sonic and density logs after resampling the logs to the corresponding equal time intervals. The output of the synthetic trace is therefore given in time; it is then converted to depth using the sonic log and resampled to one foot intervals.

To account for transmission loss of the synthetic traces and to also account for whatever amplitude recovery programs were applied to the real traces, the amplitudes of both the real and synthetic traces were balance by the reflection coefficient trace. That is, the rms amplitude of the synthetic and real traces were set equal to the rms amplitude of the reflection coefficient trace. The reflection coefficients were computed directly from the sonic and density logs. The maximum value of the reflection coefficients was arbitrarily set to 1000 in order to avoid the very small numbers associated with typical reflections coefficients.

The real seismic traces were also resampled to one foot and aligned with the synthetic trace by taking into account the KB of the well and the 2300-foot datum elevation of the seismic data. The alignment was fine-tuned by cross-correlation.

Log and attribute traces

The sequential order of traces in the real-seismic log files is:

1. depth in feet
2. density
3. sonic
4. reflection coefficient trace
5. amplitude of synthetic seismic trace
6. amplitude of real seismic trace
7. absolute amplitude
8. acoustic impedance
9. trace derivative
10. trace envelope
11. first derivative of trace envelope
12. second derivative of trace envelope
13. Hilbert transform
14. instantaneous phase
15. response phase
16. cosine of instantaneous phase
17. instantaneous frequency
18. response frequency
19. perigram
20. perigram multiplied by cosine of instantaneous phase.

The first five traces come from the well data; the remaining traces comprise the seismic trace and its computed attributes. The sequence of traces for the synthetic-trace files is the same except that the real seismic trace is omitted and the attributes apply to the synthetic trace.

Attribute Descriptions

1. Amplitude of the seismic trace.

This is copied directly from the migrated data, resampled to one foot, and aligned with the synthetic trace. As mentioned earlier the amplitude was adjusted to have the same r.m.s. value as that of the reflection coefficients.

2. Absolute amplitude.

All values of the trace amplitude were made positive.

3. Acoustic Impedance

The acoustic impedance trace is the running sum of the amplitude trace and is based on the assumption that the amplitude represents a band-limited version of the reflection coefficients. The running sum is meant to approximate trace integration, which when applied to a reflection coefficient trace gives the natural log of acoustic impedance. Consequently the trace is proportional to the log of the acoustic impedance, rather than the acoustic impedance itself.

Because of the limited frequencies in the seismic data, this attribute is more commonly referred to as relative acoustic impedance.

4. Trace derivative

This is the depth-derivative of the amplitude and is computed in the frequency domain. The maximum value of the derivative trace is set equal to 1000.

5. Envelope

The envelope of the trace is the square root of the sum of the trace squared and its Hilbert Transform, also squared; that is,

$$E = \text{SQRT} (F^{**2} + G^{**2})$$

where

E = trace envelope

F = original seismic trace

G = Hilbert Transform of the trace.

The trace envelope is a measure of the instantaneous trace energy and is related to the magnitude of the reflection coefficients. The Hilbert Transform is also a trace attribute and will be defined later.

6. First derivative of the envelope

The depth derivative of the envelope is computed in the same way as the depth derivative of the trace. This attribute is a measure of the sharpness, or rise time, of a reflection, which in turn depends on such things as the absorption properties of the rock and the transmission losses of the reflection waveform as it propagates through thin layers. As with the amplitude derivative (attribute 4) the maximum value of both the first and second envelope derivatives are set equal to 1000.

7. Second derivative of the envelope

The second derivative of the envelope is a measure of the sharpness of the envelope peaks and is influenced by a complex interaction of interfering reflections and energy absorption.

8. Hilbert Transform of the trace

The Hilbert Transform is a useful mathematical tool for transforming real functions (in this case the seismic trace) into complex analytic functions. Looking at the complex extension of an otherwise real signal allows for a number of computational advantages and insights into the nature of the signal. In the seismic case those insight come from variations in instantaneous phase, frequency, and envelope. The latter two attributes will be discussed later. The Hilbert Transform of the trace is the imaginary part of this analytic signal while the actual trace makes up the real part.

9. Instantaneous phase

The instantaneous phase refers to that of the complex analytic trace and is defined as:

$$IP = \text{arctangent} (G/F)$$

Where IP = instantaneous phase, and as before F and G represent the seismic trace and its Hilbert Transform. The instantaneous phase defines specific points on the reflection waveform and therefore is useful for visualizing reflector continuity and shape.

10. Response phase

Each seismic attribute discussed up to this point has been computed at every sample point. Consequently they can change rapidly and are influenced by noise and the normal oscillations of the reflection waveform. The response phase (and the response frequency discussed later) is supposed to represent a single value for a given waveform.

The length of the waveform is defined as the distance between local minimum on the envelope trace, and the response phase is the instantaneous phase computed at the maximum point on the envelope between the local minimum. The envelope peak determines the sample point at which the phase is computed, but the phase itself is computed from the amplitude trace and its Hilbert Transform. The envelope peak is where the signal energy is highest relative to noise, and so the instantaneous phase at that point would most represent something statistically related to the true waveform phase.

11. Cosine of instantaneous phase

The instantaneous phase varies between plus and minus π , so there is a discontinuity in the trace as the phase wraps from plus to minus π . In order to avoid the discontinuity, the cosine of the instantaneous phase can be computed, which results in a trace smoothly varying between plus and minus one.

12. Instantaneous frequency

The instantaneous frequency is the depth derivative of the instantaneous phase. It can be shown that the instantaneous frequency represents the average frequency of the power spectrum of the seismic wavelet. High frequencies can be associated with thin beds and sharp interfaces.

13. Response frequency

As in the case of response phase, the response frequency is the instantaneous frequency computed at the sample point corresponding to each peak of the trace envelope. It therefore represents a single frequency value associated with a given wavelet, and since it is computed at the point of maximum signal energy, it would also be the most reliable calculation of this attribute.

14. Perigram

The perigram is the trace envelope with its average value removed. Consequently the perigram looks more like a regular trace, oscillating between positive and negative values, but at the same time it remains directly related to reflection strength.

15. Perigram multiplied by the cosine of instantaneous phase

This combination of trace attributes provides amplitude information to the instantaneous phase. Otherwise it is impossible to tell if interesting features related to instantaneous phase were due to valid reflections or simply low-amplitude coherent noise.

Seismic trace and well correlations

In order to align the seismic trace with the logs a synthetic seismic trace had to be computed. This allows for a more valid alignment, inasmuch as the synthetic and real traces theoretically respond to the same rock properties. Since a synthetic trace requires a sonic and density log, only wells with a sonic and density log could be used in this analysis. Table B-1 lists such wells along with the ffile and cdp (field file id and common depth point) of the seismic trace nearest the well. The X and Y coordinates of the wells and traces are in the local SACROC oilfield frame of reference.

Attributes were computed from just the nearest trace. Since the data quality was excellent no trace summation for the purpose of improving signal quality was necessary. The seismic attributes were computed between the depths of 6000 ft and the deepest point on the sonic or density log. For about one-half the wells, the shallowest log value was below 6000 ft, in which case the trace attribute calculations began there.

Examples

Figure B-1, B-2, and B-3 are from the three LAS files associated with well 11-15. The first two figures show the unfiltered and filtered seismic attributes computed from the synthetic trace. Figure B-3 shows the attributes computed from the real seismic trace. For plotting purposes the amplitude of each trace was scaled independently.

Note on the synthetic trace that the reflection response from some very thin beds disappears upon filtering. The large reflection coefficient at a depth of about 6550 is a good example. Since the filtered synthetic trace more nearly represents the frequency bandwidth and therefore the resolving power of the seismic trace, we see here a case where the limited bandwidth of surface seismic data can miss a thin layer.

Also note the level of agreement between the synthetic trace and the real trace. It is reasonably good, but there are sections where the two traces are different. This is typical for all the wells. Generally there are one or two large reflections within the short window over which the attributes were computed. The synthetic and real traces line up well at the large reflections, but the fit is not as good across the low-amplitude, fine-structure of the two traces. This is common, and there are many reasons for it. The synthetic trace is a zero-offset simulation under strict geologic assumptions. The real trace, on the other hand, is the summation of many traces (perhaps hundreds) recorded over a range of shot-receiver offsets and azimuths. The real traces also contain noise (including Shear waves) and a distribution of multiple reflection noise that is different from the synthetic trace.

Table B-1: Well Locations & Nearest Seismic Trace in Local SACROC Coordinate Frame
All Units are in Feet - Note, 2 Wells are Outside the Bounds of the Seismic Survey

ffid = field file i.d. of seismic trace
 cdp = cdp number of seismic trace
 R = distance from well to nearest seismic trace
 Well name followed by the letters ST = sidetrack

Wellname	X-well	Y-well	ffid	cdp	X-seis	Y-seis	R
9-15	46497.8	77482.4	1253	87	46533.5	77491.5	36.8
17A-11A	52910.2	73841.8	1223	147	52932.5	73817.0	33.3
37-11	53273.3	59788.3	1096	157	53226.5	59806.9	50.3
56-16	53173.6	56819.5	1069	158	53165.3	56835.4	17.9
56-17	52040.5	56853.0	1069	148	52067.1	56898.8	53.0
27-18	51201.7	69610.3	1184	133	51148.1	69622.8	55.1
27-19	48854.7	69716.8	1184	112	48841.9	69755.8	41.0
10-16	53986.6	78990.6	1271	154	54005.2	79043.9	56.5
18-14	49834.0	71196.2	1198	120	49809.1	71242.5	52.6
11-11A	49998.2	76475.6	1246	119	50003.3	76520.1	44.8
20-2	46664.2	71735.3	1201	91	46643.4	71755.6	29.1
20-13	46236.8	70040.0	1185	88	46212.6	70017.5	33.0
20-14	46288.3	71197.6	1196	88	46282.3	71225.5	28.5
20-15	46262.2	72492.7	1207	87	46242.1	72439.0	56.5
26-14	46201.7	68530.9	1172	89	46240.1	68583.6	65.2
27-16	51494.0	68417.1	1173	137	51517.7	68389.4	36.5
27-17	51212.7	67263.8	1163	135	51234.8	67303.9	45.8
28-15	52703.1	65841.4	1150	149	52689.9	65787.7	55.3
28-17	51113.4	65882.7	1150	135	51152.4	65876.3	39.5
29-9	54863.8	67035.7	1162	168	54852.4	66985.1	51.8
33-15	54035.5	60476.6	1103	164	54039.5	60531.2	54.8
34-12	52708.3	64190.8	1136	150	52711.0	64243.9	53.2
36-8	53633.1	61592.7	1112	160	53657.2	61544.9	53.5
37-10	50882.3	59313.8	1090	136	50882.3	59280.9	32.9
46-13	42834.7	55397.3	1051	65	42838.3	55447.6	50.4
46-14	41722.7	55293.2	1049	55	41727.5	55291.3	5.1
49-13	45458.7	58319.1	1078	87	45425.3	58273.4	56.5
49-15	44250.7	57042.9	1066	77	44251.1	57018.9	24.0
49-16	45413.6	59036.4	1085	86	45359.8	59048.	55.1
49-17	44206.1	59078.4	1085	76	44261.6	59111.8	64.7
50-6	44007.4	54527.6	1043	76	43995.6	54499.4	30.5
53-4	46524.6	55750.7	1056	98	46493.9	55787.8	48.2
55-8	49110.4	57158.7	1070	121	49108.4	57179.5	20.9
56-18	50655.8	56843.5	1068	135	50633.1	56871.3	35.9
56-19	50647.9	55610.6	1057	136	50673.3	55656.9	52.8
56-20	48050.2	55734.1	1056	112	48031.4	55699.1	39.8
56-21	49356.6	55964.5	1059	124	49368.2	55952.6	16.6
56-22	48050.0	54516.8	1045	113	48071.5	54484.8	38.5
56-23	50570.3	54505.7	1046	136	50603.6	54448.9	65.8
56-3A	48800.4	55343.1	1053	119	48781.1	55325.3	26.3
58-4	50682.7	58055.7	1079	135	50702.8	58079.3	31.0
59-5	53625.3	58847.8	1088	161	53615.1	58903.0	56.1

10-7	49650.1	79818.9	1276	114	49644.1	79846.3	28.1
11-12	50038.9	75323.2	1235	120	50043.4	75305.8	17.9
11-15	52039.6	75582.6	1239	138	52045.5	75631.1	48.9
19-12	47951.9	72969.8	1213	102	47927.4	73003.8	41.9
46-2B	41617.2	54805.0	1045	54	41592.3	54858.4	58.9
46-15	40195.2	56617.3	1060	40	40149.8	56594.3	50.9
46-13ST	42027.6	55785.2	1054	57	41978.8	55827.7	64.7
48-1A	42145.8	57692.2	1071	58	42196.2	57688.3	50.5
48-5	42219.8	58537.9	1079	58	42246.9	58566.9	39.7
49-3A	44095.4	57992.1	1075	75	44088.5	58019.9	28.6
49-4A	45476.7	58962.2	1084	87	45463.3	58932.3	32.7
49-5A	45236.9	57838.1	1074	86	45290.1	57840.4	53.2
49-2A	46586.2	57644.5	1073	98	46601.6	57654.6	18.4
49-6A	44089.7	58942.3	1083	75	44139.1	58898.5	66.0
50-5A	43274.0	56104.5	1057	69	43315.6	56081.2	47.7
50-2A	43857.7	55441.0	1051	74	43826.6	55390.6	59.2
50-2ST	43845.7	55168.4	1049	74	43813.9	55171.0	31.9
51-2ST	43737.5	52533.1	1025	75	43771.8	52529.0	34.6
51-5	43024.5	53382.8	1032	68	43047.5	53342.1	46.8
51-6	43064.9	52369.6	1023	69	43100.3	52347.4	41.8
51-7	44056.5	53352.5	1033	77	44042.1	53394.9	44.8
51-8	42379.7	54153.8	1039	62	42432.8	54148.8	53.4
51-1A	43682.8	54140.1	1040	73	43647.1	54189.0	60.6
52-4	44507.0	54804.5	1046	80	44453.9	54803.6	53.1
52-2A	45189.9	55284.4	1051	86	45144.5	55314.7	54.6
53-2A	46366.2	55249.4	1051	97	46352.5	55245.0	14.4
53-3A	45831.8	57137.9	1068	91	45801.2	57149.9	32.9
54-1AST2	44736.3	54203.7	1041	83	44751.7	54235.5	35.3
54-5	45269.9	53319.9	1033	88	45250.1	53325.3	20.5
54-6	45667.2	54464.3	1044	91	45649.2	54514.3	53.1
54-7	47156.6	53385.3	1035	105	47129.6	53437.3	58.6
56-22A	47212.1	54701.2	1047	105	47205.6	54755.1	54.3
56-3A	48021.8	55300.7	1052	112	48006.0	55259.8	43.8
56-20A	47254.5	55900.9	1057	105	47268.9	55853.2	49.8
57-8	48375.4	54364.6	1044	116	48394.6	54355.9	21.1
57-9	47113.6	52072.7	1023	106	47163.5	52113.1	64.2
60-6ST	50731.7	52147.2	1025	138	50690.3	52130.1	44.8
60-9	51384.0	53105.0	1034	144	51406.2	53080.5	33.1
71-9	50088.9	51669.6	1021	133	50115.9	51722.5	59.4
71-8	50582.1	51049.2	1015	138	50627.0	51031.9	48.1
74-9	46196.8	50181.0	1010	98	46202.6	50736.1	555.1
72-2ST	47373.3	51460.0	1017	108	47345.1	51441.5	33.8
74-5	45661.0	50813.7	1010	93	45653.5	50767.8	46.5
74-12	44386.5	52260.6	1023	81	44418.1	52271.4	33.4
74-6	45358.5	51806.4	1019	90	45381.1	51775.2	38.5
74-10	44776.9	51137.6	1013	85	44794.0	51147.9	20.0
74-11	44740.4	50025.0	1010	85	44775.0	50818.4	794.2

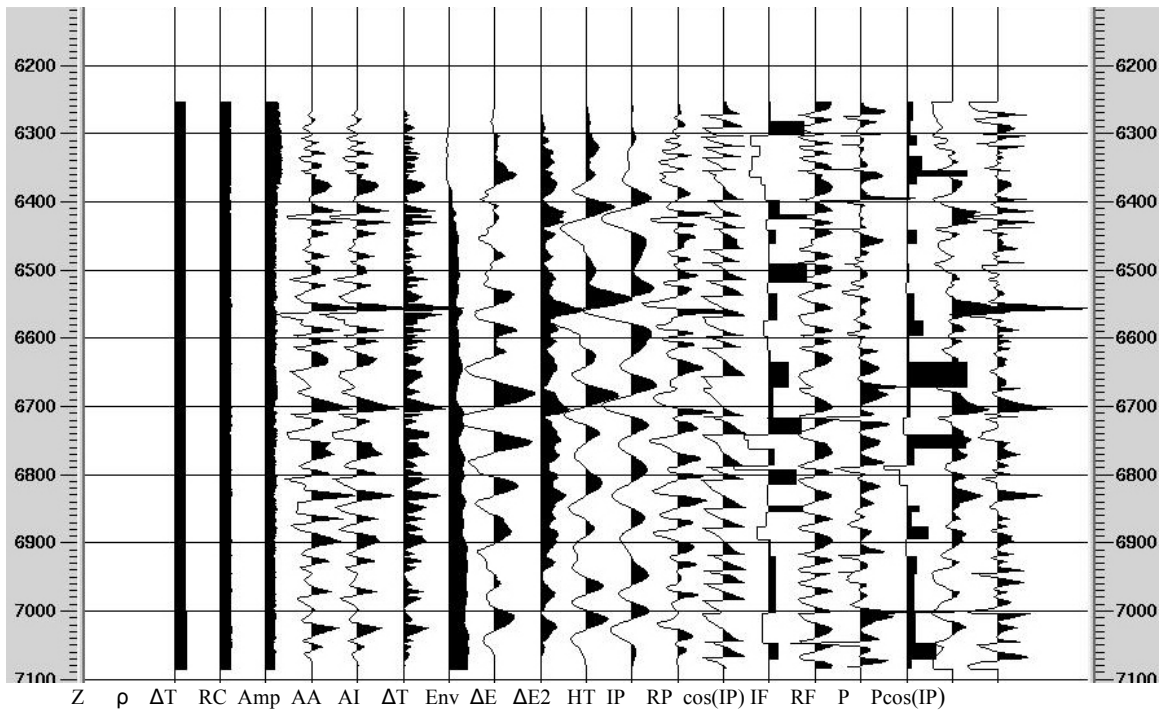


Figure B-1: Well 11-15, Log and Synthetic Trace Attributes Unfiltered

Z=depth, ρ =density, ΔT =sonic, RC=reflection coefficients, Amp=synthetic trace amplitude, AA=absolute value, AI=acoustic impedance, ΔT =trace derivative, Env=envelope, ΔE =first derivative of env., $\Delta E2$ =second derivative of env., HT=Hilbert Transform, IP=instantaneous phase, RP=response phase, $\cos(IP)$ =cosine inst. phase, IF=instantaneous frequency, RF=response frequency, P=perigram, $P\cos(IP)$ =perigram times cosine of inst. phase.

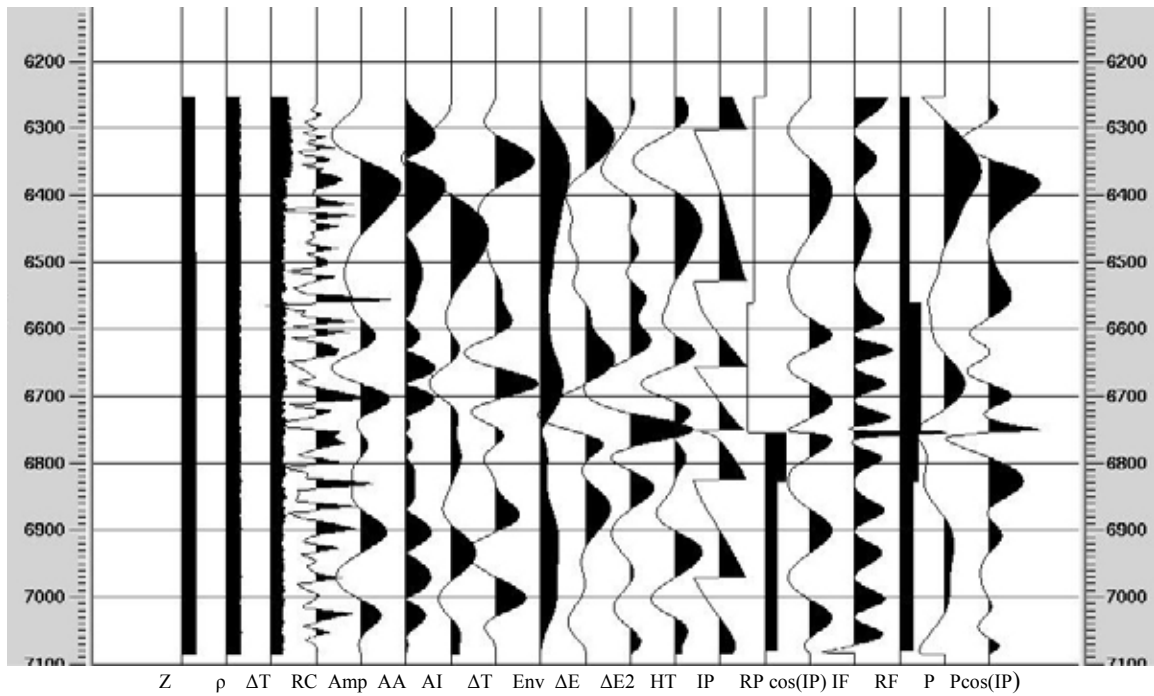


Figure B-2: Well 11-15, Log and Synthetic Trace Attributes Unfiltered in the Depth Domain with a 2-10 Hz Band Pass Filter

Z=depth, ρ =density, Δt =sonic, RC=reflection coefficients, Amp=synthetic trace amplitude, AA=absolute value, AI=acoustic impedance, ΔT =trace derivative, Env=envelope, ΔE =first derivative of env., $\Delta E2$ =second derivative of env., HT=Hilbert Transform, IP=instantaneous phase, RP=response phase, $\cos(IP)$ =cosine inst. phase, IF=instantaneous frequency, RF=response frequency, P=perigram, $P\cos(IP)$ =perigram times cosine of inst. phase.

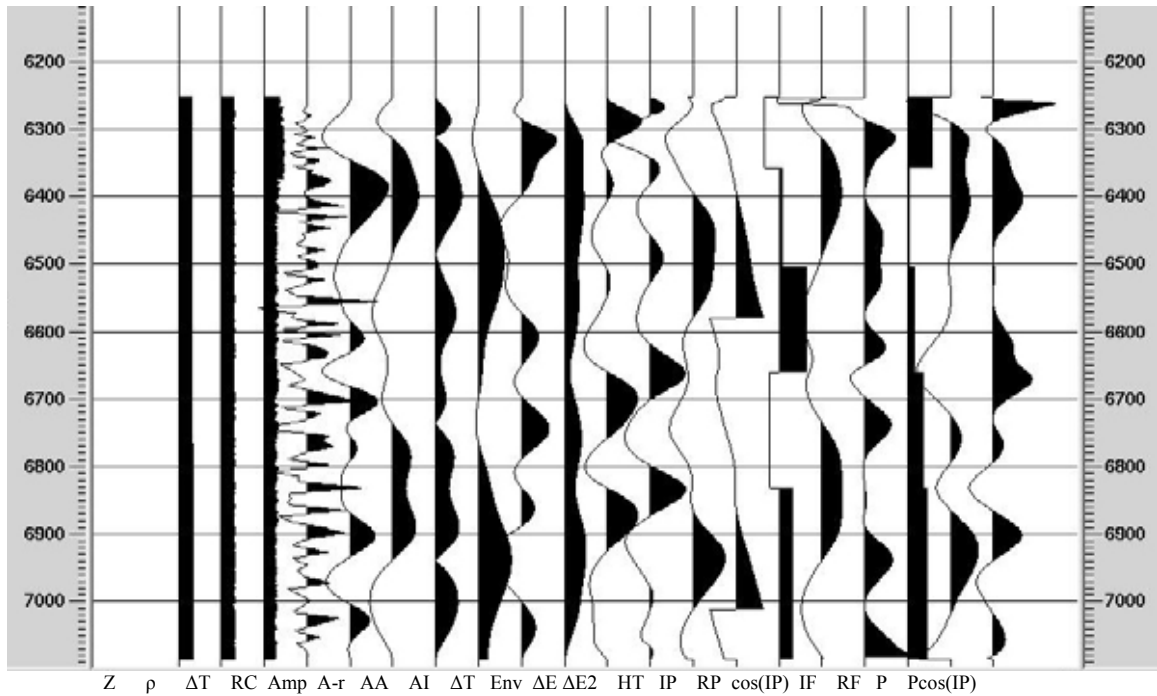


Figure B-3: Well 11-15, Log and Real Seismic Trace Attributes

Z=depth, ρ =density, Δt =sonic, RC=reflection coefficients, Amp=synthetic trace amplitude, A-r= real trace amplitude, AA=absolute value, AI=acoustic impedance, ΔT =trace derivative, Env=envelope, ΔE =first derivative of env., $\Delta E2$ =second derivative of env., HT=Hilbert Transform, IP=instantaneous phase, RP=response phase, $\cos(IP)$ =cosine inst. phase, IF=instantaneous frequency, RF=response frequency, P=perigram, $P\cos(IP)$ =perigram times cosine of inst. phase.

Appendix C: The Use of Seismic Information

One of the major goals of the SACROC project was to establish relationships between well information (wireline well logs and core data) and seismic information (surface and crosswell data) that would permit the development of a model for the prediction of porosity and permeability profiles along any surface seismic trace in the SACROC field study area.

Our original approach to reach this goal was the utilization of only advanced pattern recognition techniques to establish unconventional relationships between data at different scales, and by doing so, creating the pathway or “transformation” to derive core-scale reservoir properties from 3D surface seismic data. An essential part of the original plan, we considered the utilization of intermediate frequency data (specifically crosswell seismic) to bridge the resolution gap between 3D surface seismic and geophysical well logs, so that the resulting “transformation” could be more constrained and hence have a much narrower band of uncertainty.

The procedure thus conceived could incorporate and honor basic geological and rock physics concepts by using cluster analysis of each dataset (i.e., 3D surface seismic, crosswell seismic, logs, and core data) and identify units with similar properties (i.e., clusters of attributes, log signatures, core compositions, etc.), each corresponding to identifiable depositional environments. With these classifications, one could extend elements of geologic reality into the process.

In order to reach these objectives, we adopted the software GAMLS⁸ which is a data-driven tool strongly oriented to geosciences applications. The basic process performed by GAMLS⁸ is a clustering analysis, but predictions can be also made using the relationships developed within a clustering analysis.

Due to consecutive postponements in the execution of the planned crosswell survey, the characterization process was started with the available information, i.e., core data, well logs information and 3D surface seismic data. However, after several attempts to generate an efficient model with the application of data-driven methodologies (in congruity with the originally proposed objectives), and utilizing only the available information (no crosswell data at hand), the goal of reconciling the gap between core data and 3D surface seismic information was only modestly achieved. Therefore the essential target of a multi-scale data integration for a high-resolution reservoir characterization uniquely using data-driven techniques was considered unsatisfactory. The following paragraphs discuss the causes of this situation.

1. Three new wells with core porosity and permeability were used in developing a model. Despite a coring program designed to cover the entire reservoir at three different locations, in a reservoir as complex as SACROC there is no guarantee of capturing completely its geologic variability. This could influence the efficient feeding of the pattern recognition methods when they were used including seismic information.

2. The porosity/depth trends in the three cored wells generally are not well defined because the rocks are vertically heterogeneous with porosity varying over 20 porosity units within vertical ranges of 20 to 25 feet which can be seen in Figure C-1.

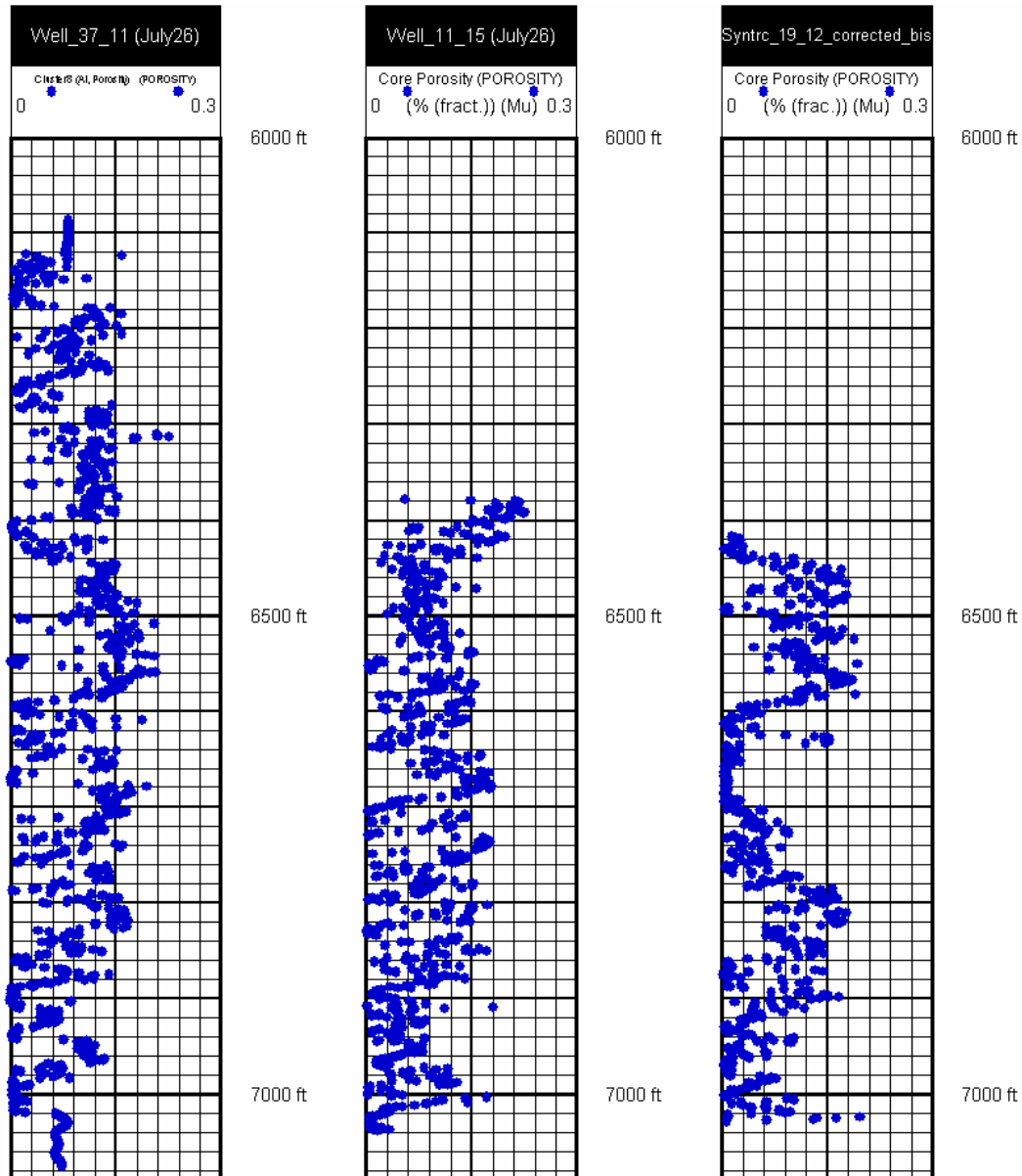


Figure C-1: Core Porosity Depth-Plots on Wells 37-11, 11-15 and 19-12

3. The seismic attribute Acoustic Impedance (AI) has low frequency with cycles on the order of 100 to 200 feet which is much larger (lower frequency) than the poorly defined cycles in core porosity (Figure C-2). Therefore, only porosity trends on the order of approximately 100 feet or greater would likely be detected if a successful prediction model were developed.

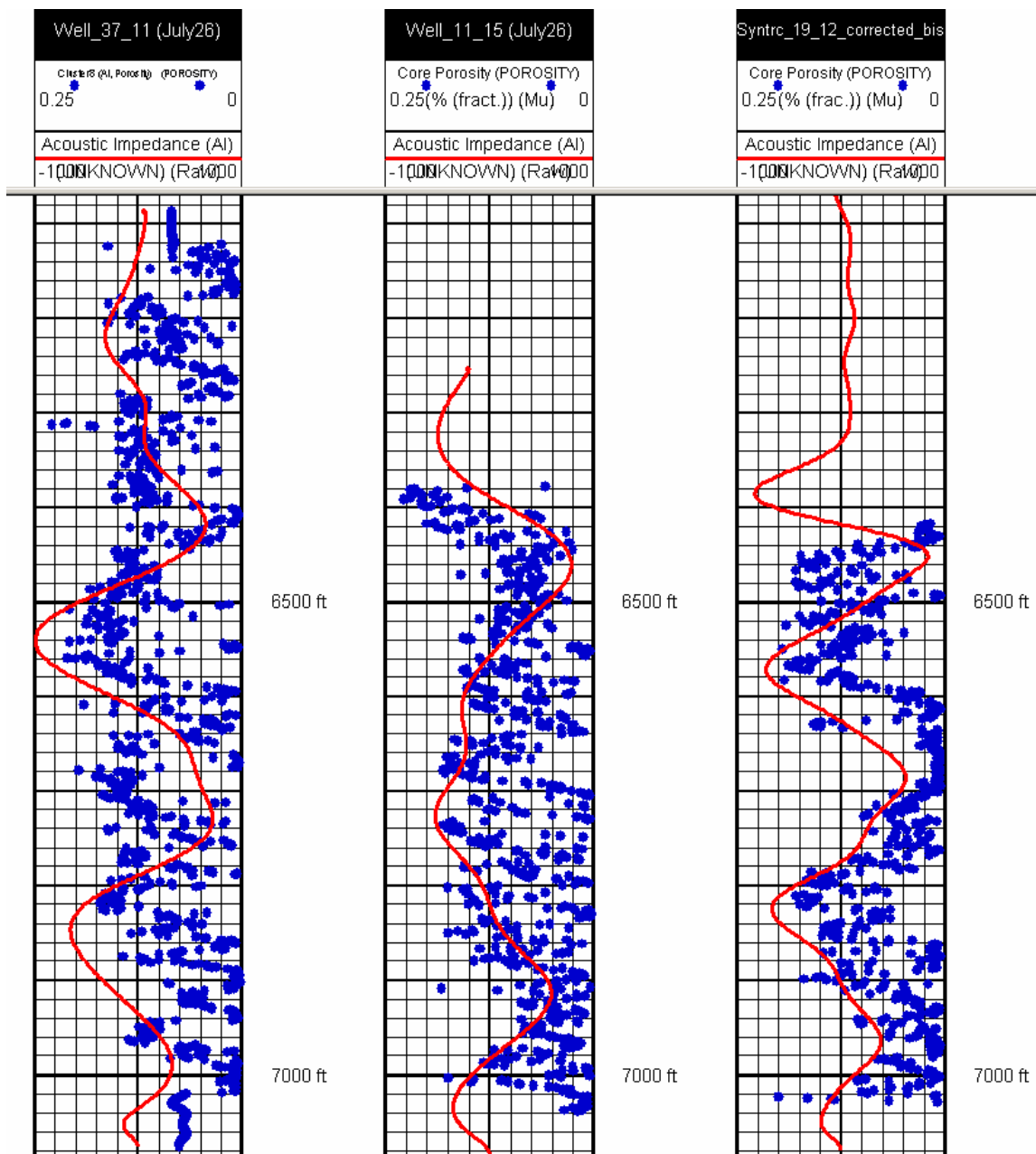


Figure C-2: Seismic Acoustic Impedance (AI) Traces Overlain onto the Core Porosity Data for Cored Wells, AI = -1000 left to 1000 right

4. AI is the attribute that is most likely to correlate (inversely) with porosity. If AI does not correlate with porosity, it is unlikely that other attributes will correlate with porosity. For the three cored wells, only one well (19-12) shows a relatively acceptable inverse correlation with AI in Figure C-2.
5. Crossplots of AI vs. actual core porosity for wells 37-11 and 19-12 show a poor trend with AI inside the different modes or clusters derived from one of the clustering run including seismic information as seen in Figure C-3. Also notice the wide range for porosity values.

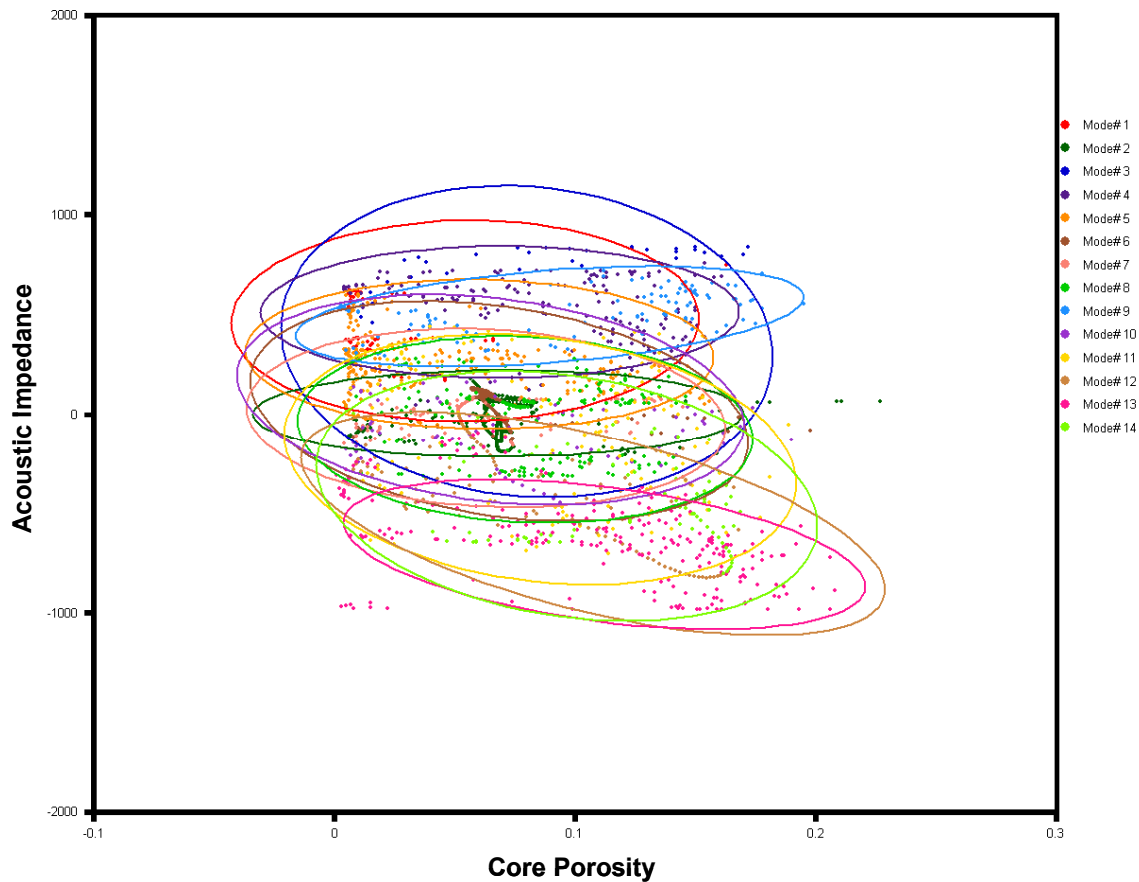


Figure C-3: AI vs. Core Porosity for Cored Wells 37-11 and 19-12
All Clusters (ovals) Show Poor Correlation between AI Measurements and Actual Core Porosity Values

6. Plots of AI vs. predicted porosity made from clustering runs show a relatively low porosity range for predicted values and a poorly defined trend with AI (compare Figures C-3 and C-4). A relatively small porosity range is predicted, compared with the original

core porosity range (see Figure C-3). All different clusters (ovals) show no correlation between AI measurements and predicted core porosity values.

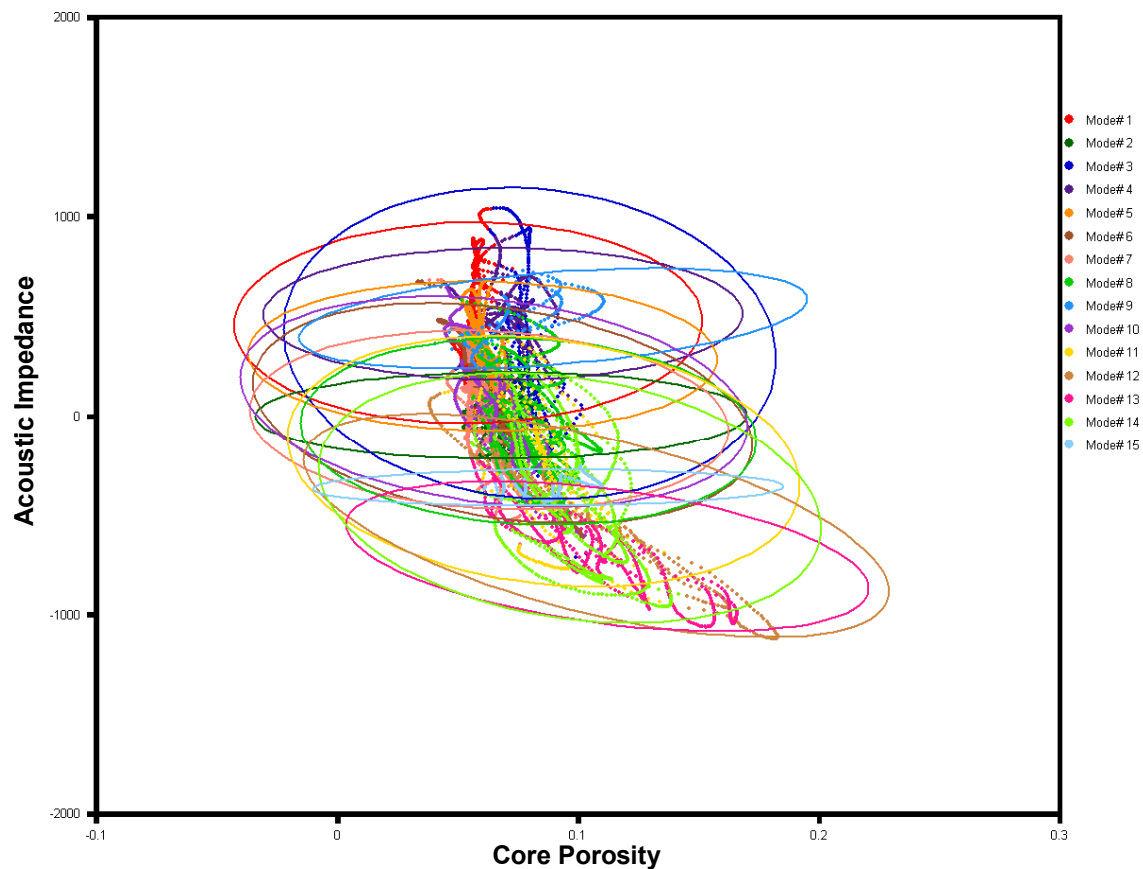


Figure C-4: AI vs. Predicted Core Porosity for Nine Non-cored Wells

7. According to geophysicist J. Fairborn⁷, “ties” of seismic depths to well log depths might vary as much as 15 to 20 feet. This means that since the vertical heterogeneity of the core porosity is such that porosity often changes 20 PU over this same depth range, then even if the seismic curve frequency were on the same order (~ 20 feet instead of 100 to 200 feet), it might still be difficult to directly correlate 3D surface seismic with well log and core data.
8. The correlation of core permeability with core porosity is poor as seen in Figure C-5, so even if porosity could be predicted with a confidence, predictions of permeability could be questionable. However core porosity and the logarithm of core permeability present pretty acceptable correlations shown in Figure C-6. This relationship motivated the utilization of the transformation on consequent tasks of prediction.
9. The one good stratigraphic marker that exists (based on GAMLS clustering and consistent with well perforation histories) is the reef-top of the SACROC reservoir.

Although one might surmise that surface seismic traces should correlate with this marker, the cored well 37-11 (located in the middle of the test area) is anomalous in terms of the seismic signal across this marker. This cast doubt about the correct calibration or quality of seismic attributes. A double cross-check was requested⁷ and it was confirmed that the originally delivered attributes were adequate as seen in Figure C-7.

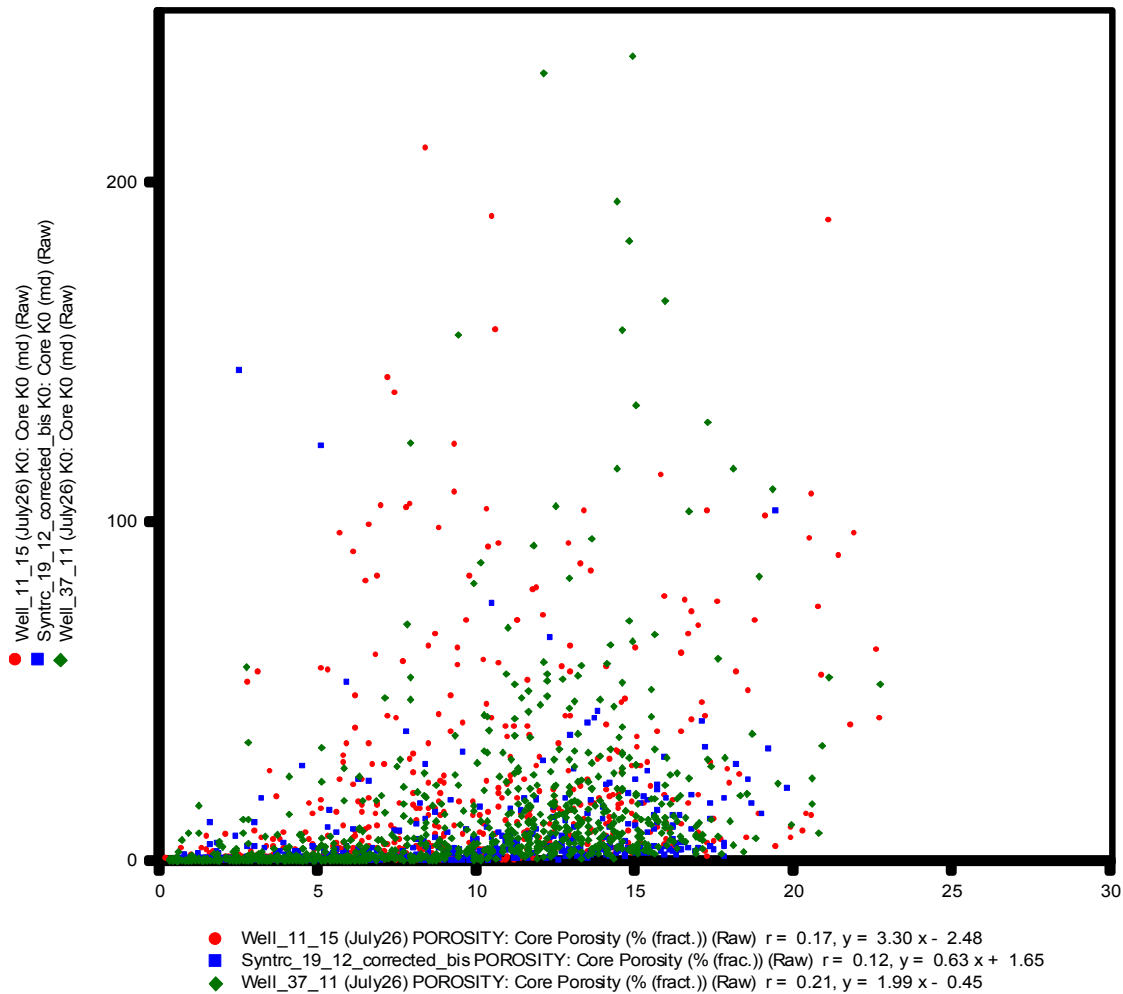


Figure C-5: Core Permeability vs. Core Porosity for the Three Cored Wells

10. A synthetic well log (AI_log) was generated from RHOB and DT well log curves. This new log, jointly with its smoothed versions (median filter transform⁸ at different window sizes), were compared with the seismic AI. For example, if a 10-point median filter was selected, a moving window was made for 10 contiguous points, and the mid-point is replaced by the median value of all 10 points. It was expected that the progressive upscaling of this synthetic log towards larger window sizes would provide progressive better correlations with the original seismic attribute AI. However, results associated to window sizes from 25' to 150' didn't show such results. These comparisons reinforced

the suspicion that the seismic information was not capable of capturing the degree of vertical variability of SACROC (Figure C-8).

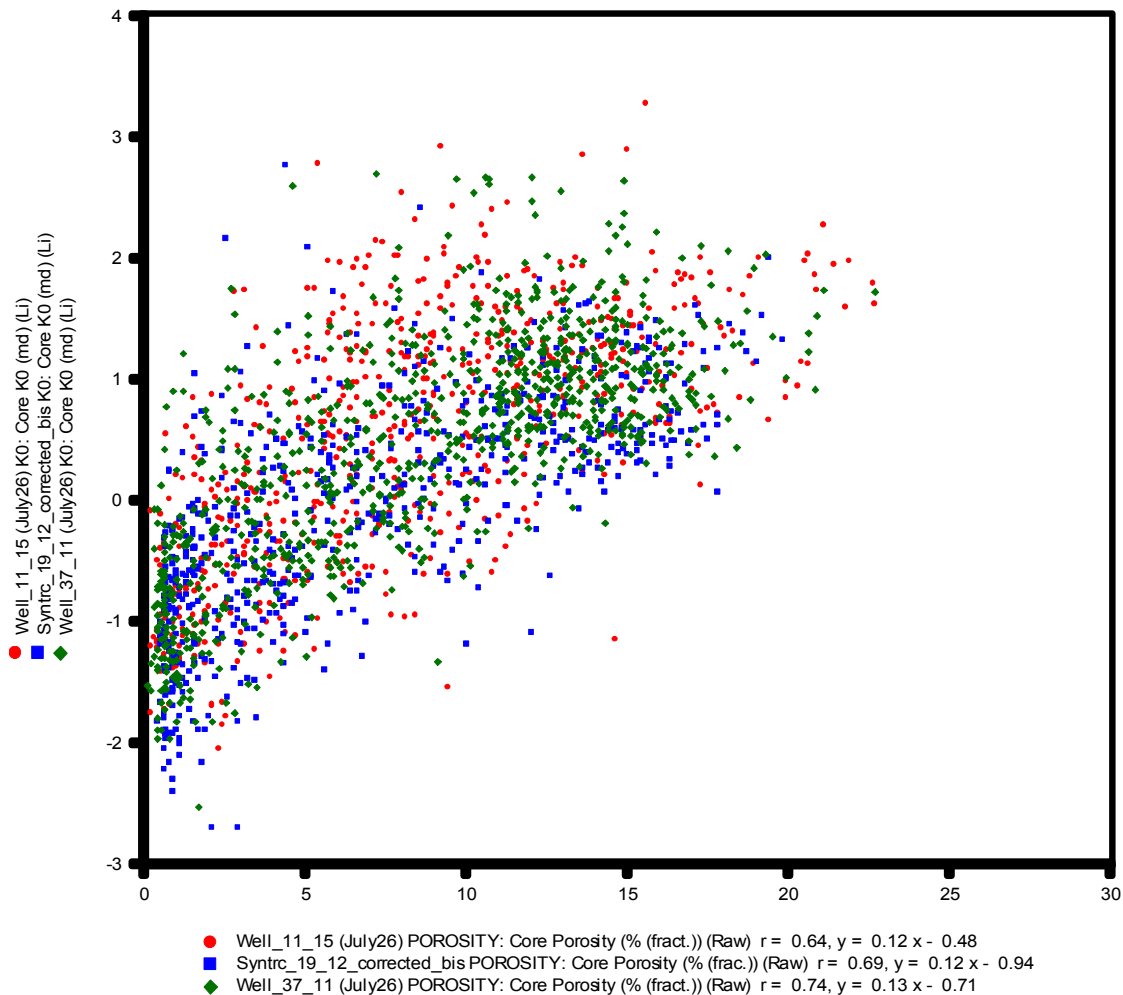


Figure C-6: Core Permeability (Logarithm) vs. Core Porosity for the Three Cored Wells

11. As of July 31, 2007, the planned crosswell seismic had not been acquired. In consequence, the first goal of establishing relationships between well logs and crosswell seismic data (and then between crosswell seismic and 3D surface seismic data) was not carried out. These relationships probably would have permitted the development of a model capable of predicting core-scale porosity and permeability profiles even in locations where only 3D surface seismic data had been “shot”. We believe that crosswell measurements could have helped to overcome some of the exposed problems.

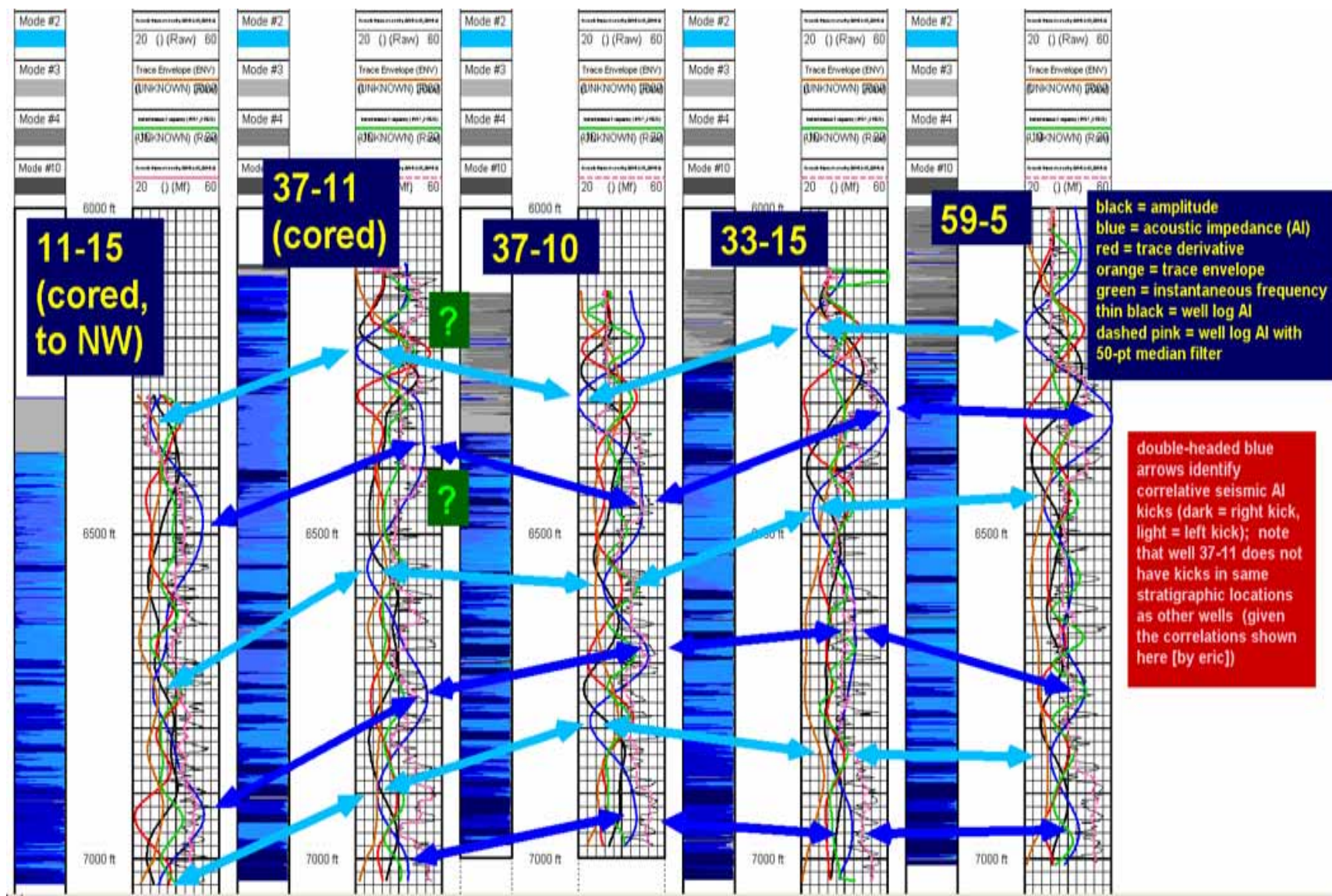


Figure C-7: Rock Types, Some Seismic Attributes and AI_Log

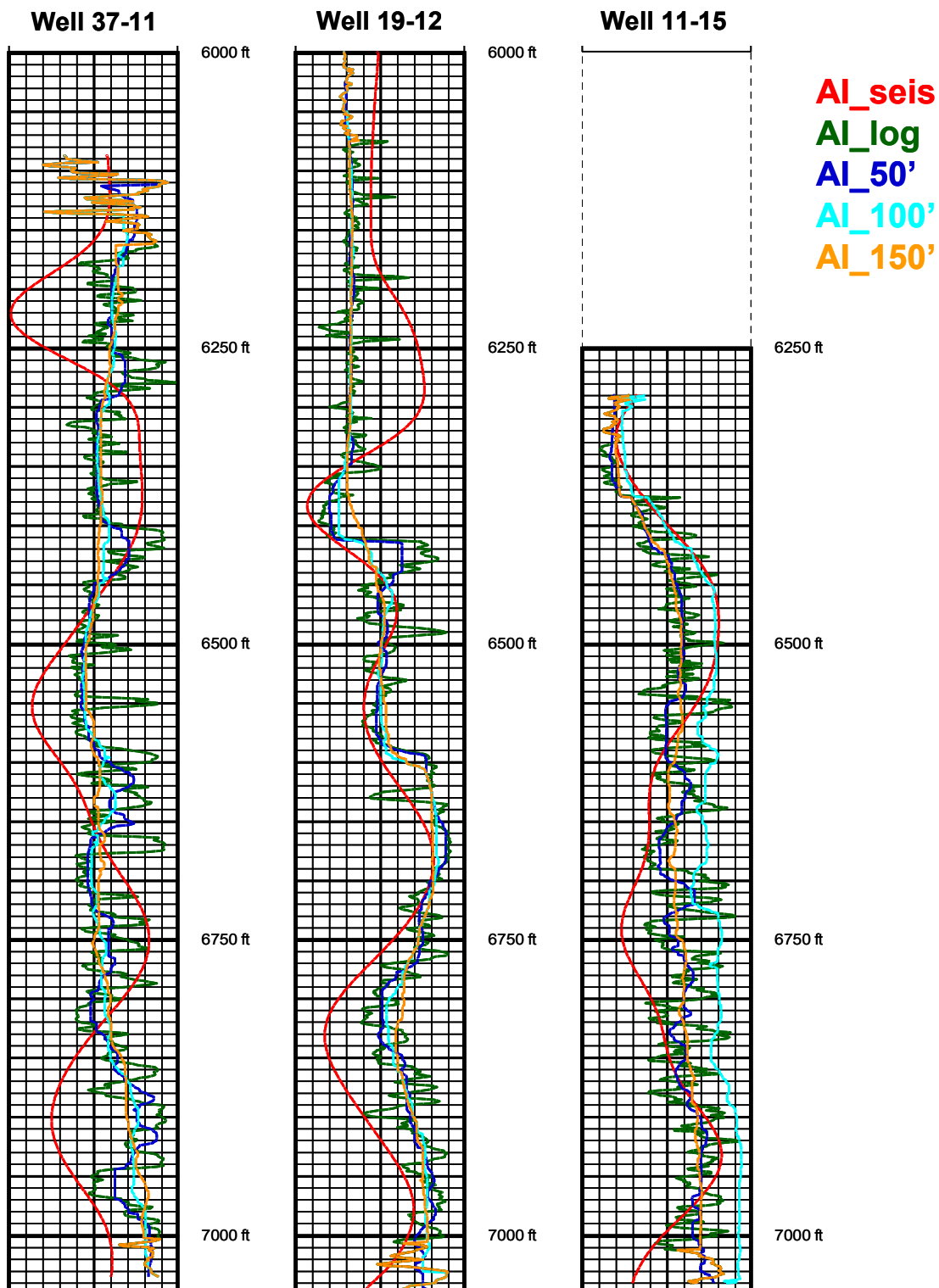


Figure C-8: AI Profiles from Seismic (AI_Seis); Log Derived (AI_Log); Median Filtered 50' (AI_50'); Median Filtered 100' (AI_100'); and Median Filtered 150' (AI_150')

Appendix D: GAMLS Methodology

Introduction to GAMLS

GAMLS⁸ develops a generalized multi-dimensional solution that results in a different n-dimensional relationship for each mode. This data-driven software is a multipurpose computer program strongly oriented toward geosciences applications.

The essential process performed by GAMLS⁸ is a clustering analysis but predictions can be also made using the relationships developed within a clustering analysis. GAMLS⁸ performs a multi-dimensional, nonlinear, probabilistic clustering of samples. Any kind of data can be used as variables in the procedure, and missing data are tolerated. Estimates of values for missing data are generated during the clustering process, and in fact this is the basis for the generation of porosity and permeability profiles at the areal positions of well logs and/or surface seismic traces.

A clustering analysis assigns all samples to a "cluster" or group within which the samples have similar characteristics (i.e., similar behavior of parameter values). These "clusters" or groups could also be termed rock types, facies, geoclasses, or flow units. The term "mode" can be also used. The results of the clustering analysis depend on the variables that are included in the analysis. Typical variables are well log "tools": gamma ray (GR), bulk density (RHOB), etc. Variables can also be core data: plug porosity, vertical permeability (Kv), grain size, etc. In this study, seismic attributes (surface and/or crosswell) are also considered as variables.

Each clustering analysis produces different results. The results can vary based on the variables, the number of modes requested, and the initialization procedures. Reference 8 discusses results of clustering analyses (runs) using different types and combinations of variables. GAMLS⁸ analyses are based on clustering and prediction methods that incorporate the following features:

1. maximum likelihood principles,
2. "fuzzy" probability,
3. a particular n-dimensional model utilized to achieve a solution,
4. the ability to cluster without training (unsupervised clustering),
5. the ability to initialize a clustering run using a training set (supervised clustering),
6. the ability to cluster using a large number of variables (>15),
7. the ability to cluster using a large number of modes (>15),
8. the ability to cluster using datasets with much missing data,
9. the ability to cluster many wells within the same run (multi-well clustering),
10. the ability to predict missing parameters in "Test" wells based on the clustering statistics of "Design" well(s).

Clustering

GAMLS⁸ performs clustering analyses as follows:

1. Variables (log tools, core data, etc.) are clustered into modes (clusters, flow units, facies, etc.) using maximum likelihood principles. A particular model is imposed on the solution (described below).
2. The user designates the variables to use, the number of modes to use, and the initialization procedure.
3. Each digitized depth is assigned to each of the modes with a fractional probability that varies from 0.0 to 1.0.
4. The imposed model is such that the frequency distribution of each mode for each variable is a Gaussian distribution. The sum of all the mode Gaussian distributions of each variable is modeled to approximate the true frequency distribution of that variable. Convergence to solution (an iterative process) is done in n-space where n = the number of variables.
5. Because of the imposed model, GAMLS⁸ can attain a solution without training. (Training is an essential initial step in neural net programs.).
6. Initialization is important because it defines the initial means and covariances of each mode of each variable. Convergence of the data to the “maximum likelihood” solution for the imposed model is affected by the initialization method used. One would like to initialize as close to the “answer” as possible.
7. Several wells can be clustered at the same time (in the same “run”).
8. Clustering can be done on datasets for which much missing data occurs for one or more variables. The values of the missing cells are estimated during iteration to convergence. These “estimated” values are different from “predicted” values (see below).

Prediction

GAMLS⁸ performs predictions as follows:

1. Predictions are based on the statistics of a previous clustering analysis. The statistics of the model well (well originally clustered) and the test well (well for which predictions are made) are assumed to be the same. Thus, the most valid predictions will be those wherein:
 - a. the lithologies of the model well and the test well are the same
 - b. the log tools used for clustering and prediction are the same, or equivalent
 - c. no interwell log tool normalization is necessary or, if it is necessary, normalization has been done properly
 - d. if core data is used in the clustering or prediction, appropriate core-to-log depth corrections have been made
2. Any variable that was included in a previous clustering analysis can be predicted.

3. More than one variable can be predicted in the same prediction analysis.

Initialization of a Clustering Run

Initialization refers to the method of computing the initial mean and covariance of each mode prior to iteration to convergence. The initialization procedure selected can affect the clustering results, particularly when using a dataset with a small number of samples. After initialization, the run is begun. Each iteration consists of a recomputation of the means and covariances plus a recomputation of the *fractional mode probability* (fmp) assignments at each sample depth. Convergence is defined to be when the largest change in fmp assignment for any mode of any depth, for the last two iterations, is <0.01 .

Types of initialization procedures commonly used are:

1. *Ordered-by-Variable (V)*: In this method, one of the variables used is selected for initial ordering. The data values for this variable are serially sorted according to magnitude (with the values of the other variables being co-sorted), and then $1/m$ of the samples is (initially) assigned to each of m groups (clusters) where m = the number of modes to be used during clustering. The $1/m$ samples with the smallest range of values are assigned to Mode 1, the next smallest to Mode 2, etc. The means and covariances of each mode are then computed for each mode using all of the variable data assigned to each mode, and this serves as the initial distribution for each of the modes. During iteration, the distribution of the samples is continually reassigned to satisfy the imposed model (Gaussian distribution of each mode) used. The departure from this model is minimized in v -space (v = number of variables used) using maximum likelihood principles.
2. *Large Covariance (LC)*: In this method, the initial mean and covariance of each mode is assumed to be large and close to one another. The initial perturbation of the means and covariances of one mode from another is given by a "Scale" value that varies from 0.05 to 2.0, with 2.0 being the largest perturbation. The LC method is used when each of the variables is believed to contribute to the solution in a roughly equal manner. The Ordered-by-Variable method is used when one of the variables is believed to be the most important for clustering.
3. *Prior Statistics*: The statistics of a previous clustering run may be used to initialize a run. This might be done if there is a particular "model" well, previously clustered, that is used as a basis for clustering additional wells.
4. *Core Data*: Lithologies identified from a core description can be assigned a numeric code and inserted into the LAS file as a variable. This variable can then be used to initialize a clustering run. Then, the code examines the well log values at each depth of each lithology and thereby computes the mean and covariance of each lithology. From this initialization, GAMLS then iterates to satisfy the imposed model. The original core description lithology assignments will be honored at solution to the extent that the core description happens to honor the imposed model.

Transformations

Pre-clustering data transformations can be made on any variable prior to clustering. These include:

1. logarithm (base 10 conversion); it is recommended that this be done for any variables that have or approach log-normal distributions
2. addition: a constant value is added; this can be done for interwell tool normalization
3. low/high cutoffs: ignore values less than or greater than a specified value
4. autoscale: converts range to mean of zero and standard deviation of plus and minus one.

Procedure

Well log and core data (depth corrected) are imported into GAMLS⁸ as LAS files. Data are examined with the goal of selecting the variables to use and the transformations to make (if any). Transformations are then made.

During the “SetUp” for a *clustering* run, the user inputs the following:

1. the well or wells to be clustered
2. the depth interval to be used for each run
3. the variables (tools, core data, etc.) to be used
4. the number of modes (clusters = ideal rock types) to be used
5. the initialization procedure

During the “SetUp” for a *prediction* run, the user inputs the following:

1. the well or wells for which predictions are to be made (test wells)
2. the clustering run whose statistics are to be used for the prediction (determined from clustering a “model” well)
3. the variables from the test wells that match the variables from the model well; if the test well(s) do not contain data for one or more of the variables, null values are inserted

Each clustering analysis produces a Statistics and a Mode Probabilities file. In addition, crossplots, frequency plots, and depth plots can be generated using the data contained in these two files. Depth plots that can be made include Cumulative Mode Probability (CMP) plots and BEDS plots. CMP plots show the fractional mode probability assignment made for each mode at each digitized depth. BEDS plots are a simplification of the CMP plots, and where each digitized depth is assigned to the mode with the highest probability.