CONTRIBUTION OF SEISMIC AMPLITUDE ANOMALY

INFORMATION IN PROSPECT RISK ANALYSIS

A Thesis

Presented to

the Faculty of the Department of Earth and Atmospheric Sciences

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

By

Tuna Altay Sansal

May 2014

CONTRIBUTION OF SEISMIC AMPLITUDE ANOMALY

INFORMATION IN PROSPECT RISK ANALYSIS

____

Tuna Altay Sansal

APPROVED:

____

Dr. John Castagna, Chairman

____

Dr. Peter Copeland, Member

____

Dr. David DeMartini, Member

____

Dean, College of Natural Sciences and Mathematics

Acknowledgements

I would like to express my gratitude to my advisor, Dr. John Castagna, for his guidance, for sharing of his valuable knowledge, and his patience during this study. I would also like to thank my committee members Dr. Peter Copeland and Dr. David DeMartini for their support, wisdom and their valuable time.

I would like to thank Mike Forrest, Rocky Roden, and Roger Holeywell, who are consulting the DHI consortium where the data came from, for their help, kindness, and feedback.

Special thanks are given to Turkish Petroleum Corporation (TPAO) for their financial support during my graduate study.

Without the support of my friends, the fun and knowledge sharing experience we had during graduate school, it would be very hard to complete this work. So, I would like to thank all my friends for their presence.

Finally, I would like to thank my mom Belma Sansal, my dad Zeki Sansal, my sister Bengu Sansal, and my beloved Cassie DeBusk for giving their support and sticking with me through the process of completing my graduate studies.

CONTRIBUTION OF SEISMIC AMPLITUDE ANOMALY

INFORMATION IN PROSPECT RISK ANALYSIS

A Thesis

Presented to

the Faculty of the Department of Earth and Atmospheric Sciences

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

By

Tuna Altay Sansal

May 2014

Abstract

Stepwise linear regression of a database of 177 Class III hydrocarbon prospect outcomes and associated descriptions of Direct Hydrocarbon Indicator (DHI) observations indicate that the seismic characteristics can be used to predict well outcomes with a success rate better than 74% for out of sample tests. The most important seismic characteristics are presence of a phase change at the down dip edge of the anomaly, down dip conformance of the anomaly to structure (fit to closure), lack of unexplained anomalies in the same stratigraphic interval in the area, down-dip extent of the anomaly consistent with sealing capacity, and presence of prospect analogues. AVO analysis and results consistent with rock physics trends are also found to be significant factors in success/failure analysis. As seal capacity is an often neglected factor, its high ranking in the stepwise regression has significant practical implications. The mean-squared prediction error and residuals for all of the predictions are within acceptable limits. This shows that there is a relationship between the characteristics and quality of the interpreted DHI anomalies and the prospect outcome.

Table of Contents

List of Figures

Chapter 1 - Introduction

All hydrocarbon (HC) exploration projects have the common goal of finding reserves of oil and gas that are profitable. Companies involved in oil and gas exploration need to assess risk factors before drilling in potential exploration prospects using several factors. Eliminating the exploration risk is not possible. However, many companies greatly reduce their risk by implementing new principals of risk analysis and new technologies.

One of methods used in reducing risk of drillable prospects is understanding the mechanics and the impact of amplitude anomalies on prospects. The presence of Direct Hydrocarbon Indicators (DHI) on seismic data have a significant impact on uncertainty levels in risk analysis.

When evaluating DHIs, it is very important to understand the DHI anomalies correctly. To interpret DHIs in a correct manner, one must know the seismic data properties such as polarity and phase. Because the expected DHI anomalies vary depending on the rock properties, knowledge of the geologic setting of the area is a critical part of the DHI evaluation process.

1.1. DHI and AVO Concepts

Seismic DHIs are evidence of hydrocarbons directly seen on seismic data. In seismic exploration prior to amplitude preserving processing around 1960's, the true amplitudes were not carried on and automatic gain control (AGC) was applied; thus, interpretation of the amplitude anomalies was not possible. When the water in the pores is replaced with hydrocarbons, when the rock is more porous and when net-gross ratio increases, the acoustic impedance of reservoir rocks is reduced. And depending on the encasing lithology, the reservoirs produce amplitude anomalies in seismic sections (Brown, 2012).

The understanding of seismic polarity, phase and frequency content of the seismic data is a key when amplitude anomalies are observed. Without known polarity and phase, a lithological change can be easily interpreted as an amplitude anomaly caused by hydrocarbons. Conventional direct hydrocarbon indicators are bright spots, dim spots, polarity reversals (Figure 1.1), flat spots, and gas chimneys.

- **Bright spots:** If the acoustic impedance of a brine sand is less than the encasing lithology, it causes a soft (a *trough* in SEG Standard Polarity) reflection. When brine is replaced with hydrocarbons (dominantly gas), for Class IIn and III sands, the magnitude of the amplitude increases; thus creating a bright spot.
- **Dim Spots:** As opposed to bright spots, when the acoustic impedance of the brine sand is larger than the encasing lithology, it causes a hard (a *peak* in SEG Standard Polarity) reflection. As hydrocarbons are added to the rock frame; the impedance of the previous brine sand drops to a level where it does not go lower than the encasing lithology's impedance. This causes a drop in the magnitude of the amplitude at reservoir level, which is observed as a dim spot in the seismic section.
- **Polarity Reversals:** These occur when the impedance of brine sand is slightly more than the encasing lithology (still a hard reflection). However, when hydrocarbons are substituted for brine, the impedance of the fluid-filled sand drops below the encasing lithology. This causes the polarity of the seismic amplitude to change from a peak to trough or vice-versa in different polarity concepts.

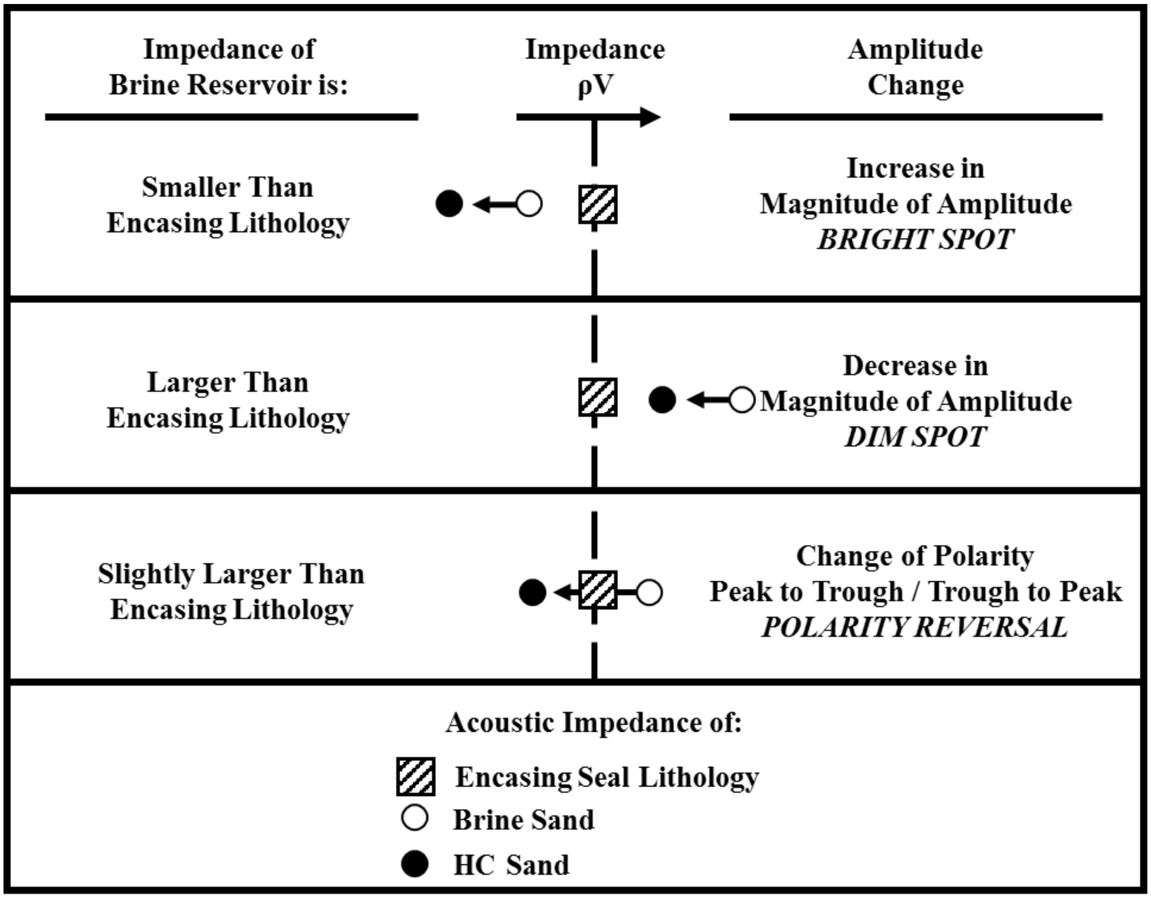


Figure 1.1: Variations in amplitude response due to changes in pore fluid. The impedance values are plotted according to SEG Standard Polarity. (Modified from AAPG Memoir 26)

- **Flat Spots:** In seismic sections, the hydrocarbon-brine contact produces a flat reflection, unconformable with the lithological reflections from the trap boundaries. If they are correctly mapped, the flat spot can give the interpreter a rough idea of the reservoir thickness (Backus and Chen, 1975). In practice, most of the time, observed flat spots have the largest impedance contrast compared to other reflections surrounding them, therefore it is easy to pinpoint. However, the reservoir must be thick enough to produce a flat spot. It is also possible that, in rare cases, the structure and the lithology can be observed as a fake flat spot.

- **Gas Chimneys:** Gas chimneys are seen in seismic sections when over-pressured gas breaches the seal and migrates towards the surface. The attenuation caused by gas chimneys can shadow most of the structure below and in it. They are indicators of gas presence in the basin and is a direct hydrocarbon indicator. Due to attenuation caused by the gas, the area where the chimney is, either loses higher frequencies or gets completely attenuated in P-wave sections. Incorporating S-wave data, which is less sensitive to fluid changes, to the interpretation will aid in a better determination of the structure. However, gas chimneys are also an indicator of seal breaches around the reservoir. Therefore they must be interpreted carefully to avoid drilling low gas saturated targets.

Figure 1.2, Figure 1.3, Figure 1.4, Figure 1.5, and Figure 1.6 show examples of described direct hydrocarbon indicators.

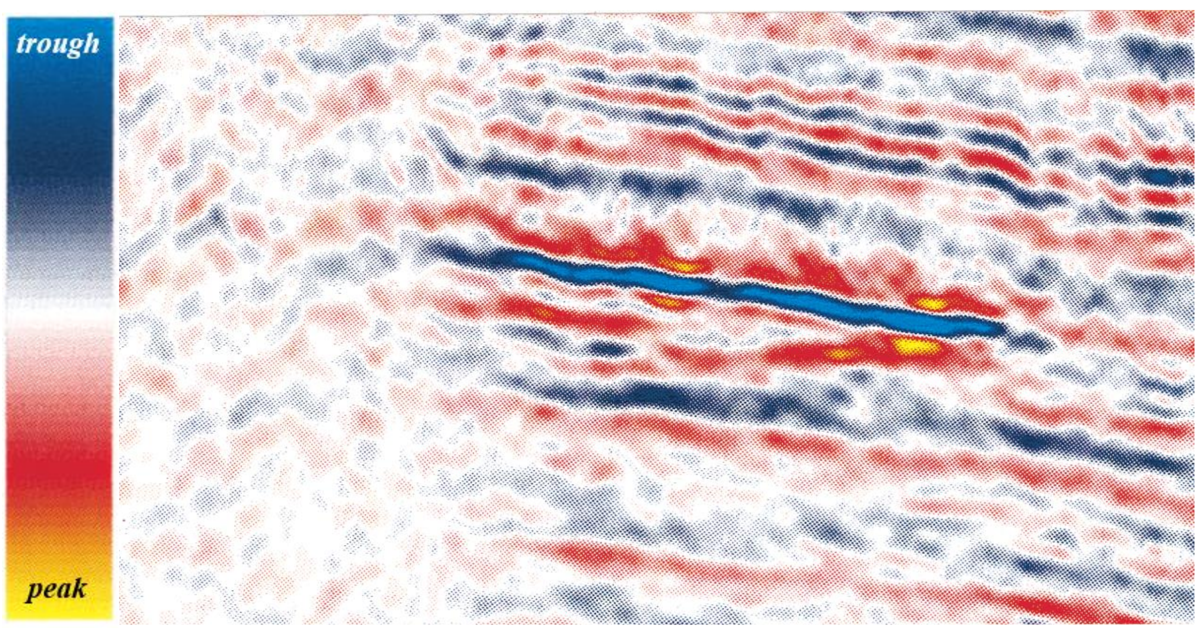


Figure 1.2: A vertical section from Vienna Basin, Austria. It shows a bright spot (blue), and a flat spot (red). The amplitudes are weak on the left side because of a gas chimney effect. (Brown, 2012)

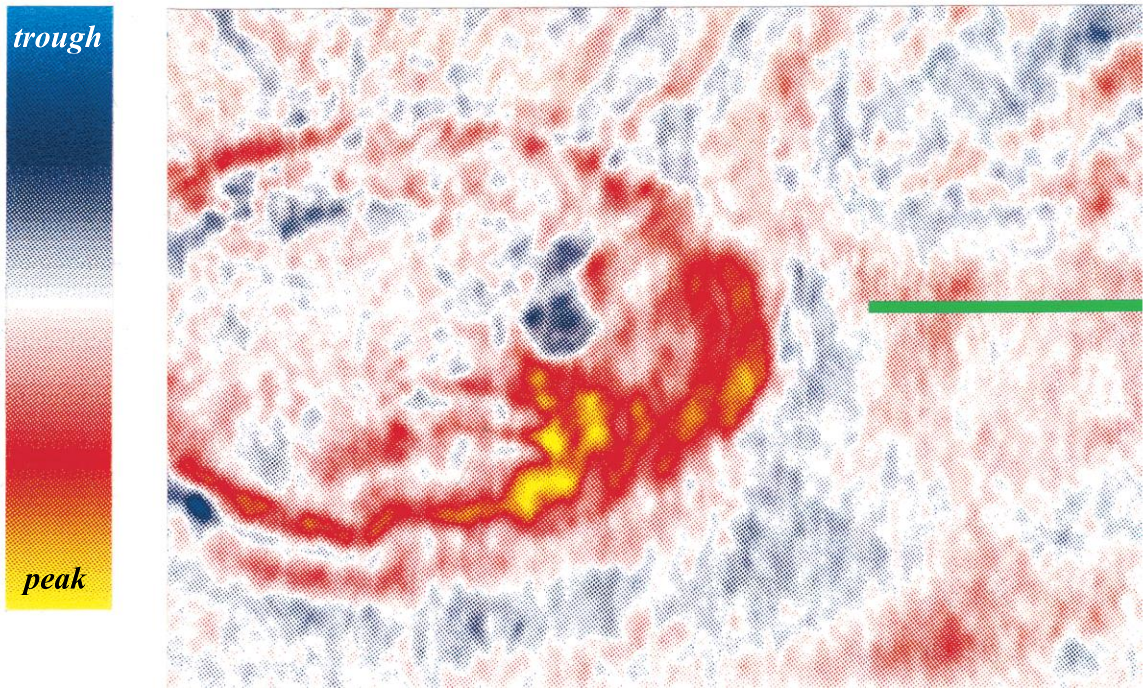


Figure 1.3: A time slice at the flat spot level for the vertical section in Figure 1.2. (Brown, 2012)

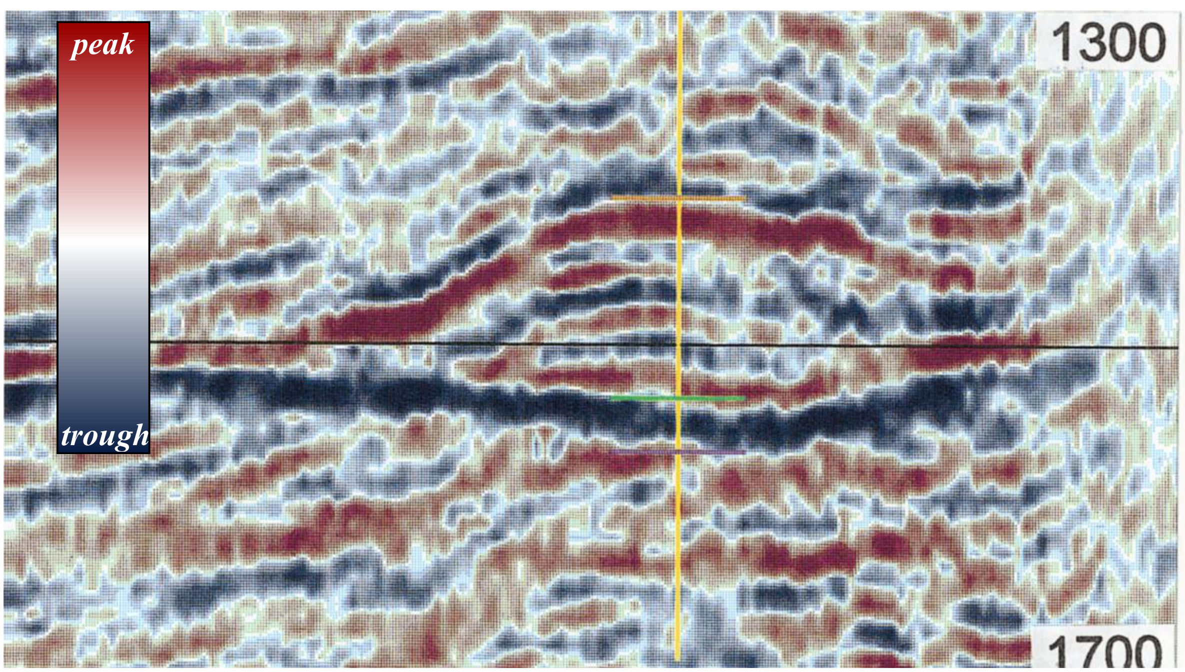


Figure 1.4: Another example of a bright spot (red), and a flat spot (blue). Data from Nile Delta, Egypt. Notice the velocity sag caused by the gas reservoir. (Brown, 2012)

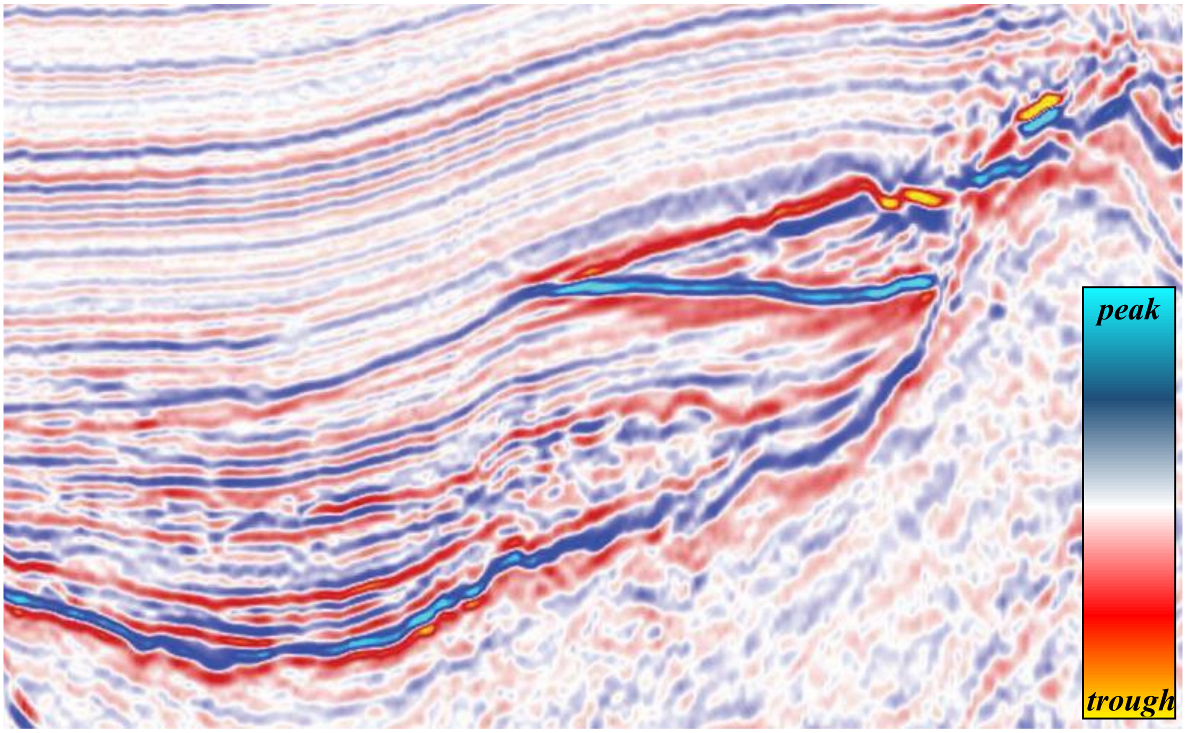


Figure 1.5: A polarity reversal and a gas water contact (blue) seen in stacked data from offshore Sabah, Malaysia. (Brown, 2012)

Despite the fact that the DHI anomalies help interpreters make smarter decisions on prospect evaluation, DHIs may also misguide exploration in several ways. It is well known that seismic amplitude anomalies can be caused by factors other than commercial hydrocarbons (Forrest et al., 2010):

- Low-saturation gas
- Clean blocky wet sand
- Low-velocity shale or marl
- Low-porosity gas sands can be interpreted as high-porosity oil sand

After determining a DHI anomaly in stacked data, an Amplitude Variation with Offset (AVO) study should be done to understand and classify the possible reservoir.

Zoeppritz's equations (1919) define the seismic amplitude variation with offset for reflected and transmitted planar waves between the boundaries of two elastic media. Due to the complication of the Zoeppritz equations, approximations were made. The most widely known ones are the Richards and Frazier (1976), and Aki Richards three-term (1980) approximation. Shuey (1985) also proposed an approximation to Aki-Richards to even more simplify the angle dependence.

Rutherford and Williams (1989) presented that gas-sand reservoirs can generally be classified as Class I, II, and III sands based on their AVO characteristics. Castagna et al. (1998) introduced class IV sands. Table 1 describes the character of these sands, and Figure 1.6 shows the typical amplitude variation with offset.

Table 1: AVO behavior of class I, II, III, and IV sands at top of reservoir. (Castagna et al., 1998)

Class	**Relative Impedance**	**A (Intercept)**	**B (Gradient)**	**Remarks**
I	Higher than overlying unit	+	-	Reflection coefficient (and magnitude) decrease with increasing offset
II	About the same as the overlying unit	± (IIp/IIn)	-	Reflection magnitude may increase or decrease with offset, and may reverse polarity
III	Lower than overlying unit	-	-	Reflection magnitude increases with offset
IV	Lower than overlying unit	-	+	Reflection magnitude decreases with offset

The Shuey three-term approximation introduced the terms Intercept (A), Gradient (B), and Curvature (C) for AVO studies. However, the last term C is mostly neglected for AVO cross-plot studies.

$$R(\theta) \approx A + Bsin^2(\theta) + Csin^2(\theta)tan^2(\theta) \qquad \text{(Eq. 1.1)}$$

$$A = \frac{1}{2}\left(\frac{\Delta V_p}{V_p} + \frac{\Delta \rho}{\rho}\right), \qquad B \approx -2\frac{V_s^2}{V_p^2}\frac{\Delta \rho}{\rho} + \frac{1}{2}\frac{\Delta V_p}{V_p} - 4\frac{V_s^2}{V_p^2}\frac{\Delta V_s}{V_s},$$

$$C \approx \frac{1}{2}\frac{\Delta V_p}{V_p}$$

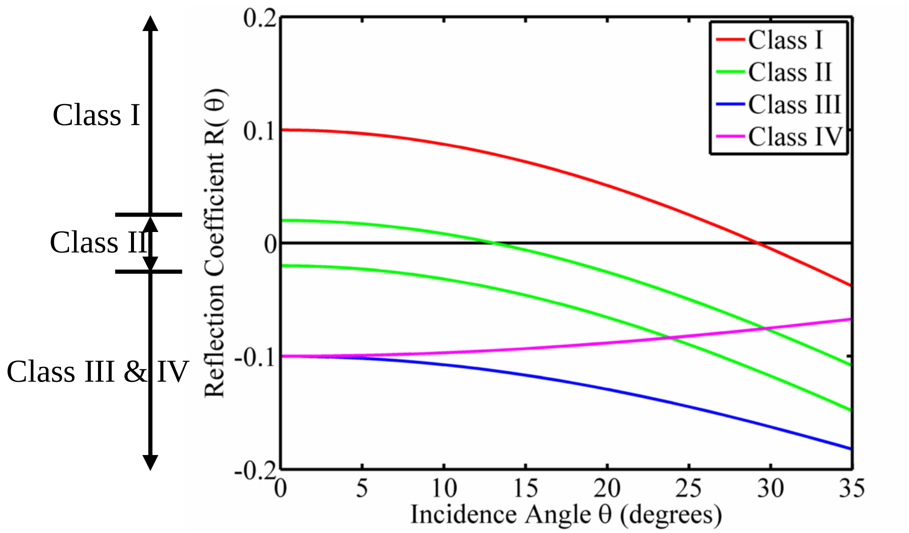


Figure 1.6: P-Wave reflection coefficients for a shale-gas sand interface (modified from Castagna et al., 1998)

Intercept and Gradient values are cross-plotted for further analysis of the amplitude variations in zones of interest. A background trend must be defined for wet formations and the outlying data points are grouped according to cross-plot quadrants to define the class.

1.2. Database and Characteristics

A database provided by the DHI Interpretation and Risk Analysis Consortium, which began in 2001 with the support of oil companies, was used. The software Seismic Amplitude Analysis Module (SAAM), developed by the DHI Consortium, is a powerful tool for DHI prospect evaluation. The database consists of 217 prospects all around the world with targeted reservoirs from the age Triassic to Pleistocene. Size of the prospects ranges from 100 to 10,000+ acres; and the depths range from 2000 to 20,000+ ft. Closures include structural, stratigraphic, and combination of both (Roden et al., 2012). Out of 217 prospects, 177 complete Class III prospects were elected for this study. Each AVO class has a specific list of questions about DHI characteristics, which were correlated to the prospect outcome. The DHI characteristics are graded on a scale of 1 to 5 according to the observed behavior, as seen on Table 2.

Table 2: List of characteristics and respective grade values (G). 1: Worst, 5: Best Observation

#	CHAR.	G	GRADE DESCRIPTION
1	Amplitude change (as viewed on stacked P wave seismic data)	1	Top of zone of interest shows a strong amplitude change that is a positive reflector.
		2	Barely perceptible change relative to chosen background
		3	Minor amplitude change relative to off closure event (water leg) OR isolated anomaly with no water leg event.
		4	Low to moderate (negative) amplitude change relative to off closure (strat or structure) event.
		5	Moderate to strong (negative) amplitude change relative to off closure (strat or structure) event. Preferably has an observable water leg.
2	Consistency within mapped target area (on stacked data)	1	Highly variable from background to max in no perceptible pattern within the mapped target area.
		2	Fairly variable, but predominantly higher within the mapped target area.
		3	Generally consistent within the mapped target area, but small areas show marked variations.
		4	Generally consistent within the mapped target area.
		5	No significant variation within the mapped target area.
3	Are unexplained anomalies seen on stacked data outside closure (within same stratigraphic sequence)?	1	Many similar, unexplained amplitude events seen outside likely closures.
		2	Some similar, unexplained amplitude events seen outside likely closures.
		3	Similar, unexplained amplitude events only occasionally seen outside probable closures.
		4	Similar, unexplained amplitude events rarely observed outside closures and then not as distinctive.
		5	This amplitude event is unique. No similar events can be seen outside the prospect area
4	Down-dip conformance (fit to closure) based on far-offset or stacked data	1	None, down-dip edge of amplitude anomaly cuts across structure.
		2	Down-dip edge cuts across structure, but small areas may be conformable. May include tilted paleo-contact.
		3	Neutral. Down-dip edge shows about equal mix of conformable and non-conformable characteristics. OR…no water leg, i.e. reservoir not present down-dip.
		4	Down-dip edge generally conformable, within limits of velocity model
		5	Almost perfectly conformable along entire down-dip extent-nearly perfect match to depth contours.
5	Lateral conformance based on far-offset or stacked data	1	Highly irregular anomaly pattern is difficult to explain by simple fault or channel model or contradicts accepted depositional model for the area.
		2	Anomaly forms a complex pattern which may fit a geological model but not established for this area.
		3	Anomaly largely fits a plausible depositional model, but one for which there is little direct evidence OR partially fits an established depositional model.
		4	Anomaly fits a simple, well-established depositional model that is known to work in the area.
		5	Structural trap. Lateral edges are structurally controlled by faults OR 4-way dip closure (no lateral conformance issues).

6	Flat Spots indicating fluid contacts	**1**	High confidence of having high quality, thick sand but no flat spot is present. Flat spots expected due to response seen in analog fields.
		2	None observed.
		3	Slight indication along only a portion of feature OR suspect origin. For instance, layering fails to extend beyond flat spot suggesting channel edge origin OR low net to gross sand ratio makes stratigraphic origin much more likely than HC column.
		4	Good indication, but might also be caused by stratigraphic changes. Signature typically downgraded due to significant lateral variation or because it is seen in only specific line orientations.
		5	Consistent signature along entire down-dip limit. If wide enough, amplitude variation realistically reflects change in overlying thickness.
7	Phase or character change at the down-dip edge of the anomaly ('tuning' or interference effect caused by a change in fluid content)	**1**	No phase or character change observed.
		2	Subtle phase or character change at down-dip edge of anomaly.
		3	Fair phase or character change at down-dip edge of anomaly.
		4	Seismic event broadens in up-dip direction. Approaching tuning thickness…but not thick enough for distinct flat spot.
		5	Good broadening of seismic event with strong peak at base (bottom of gas sand). May be associated with edge of flat spot.
8	Signature match vs expected (polarity and shape)	**1**	A well-defined response is the reverse of that predicted from modeling or pertinent analogs.
		2	Response appears significantly different from that predicted from modeling or pertinent analogs.
		3	Complex, indeterminate signature shows some similarities but differences from that predicted from modeling or pertinent analogs.
		4	Response appears generally similar to that predicted from modeling or pertinent analogs.
		5	A well-defined response closely agrees with that predicted from modeling or pertinent analogs.
13	Change in AVO compared to model (wet vs HC filled)	**1**	Clear mismatch with model.
		2	Poor match with model.
		3	Results are indeterminate.
		4	Fair match to model.
		5	Clear match to model.
14	Excluding possible stacked pays, is the AVO effect anomalous compared to reflectors above and below?	**1**	AVO response of the target is nearly identical to that of virtually all nearby reflectors. Good evidence for processing artifact or lithology effect, not HC's.
		2	AVO response of target is similar to essentially all other reflectors in the section.
		3	AVO response of target is shares some characteristics with other events.
		4	AVO response of the target is somewhat different from that of nearby reflectors.
		5	AVO response of the target is distinctly different from that of nearby reflectors outside likely closure.

15	Is the AVO effect anomalous compared to the same event outside closure?	**1**	AVO response of the target is nearly the same as the event shows outside closure. Strong evidence that this is a processing artifact or lithology effect, not HC's.
		2	AVO response of the target is generally similar to the same event outside closure. Suggests this may be a processing artifact or lithology effect, not HC's.
		3	AVO response of target shares some characteristics with the same event outside closure OR can't compare.
		4	AVO response of the target is somewhat different from that of the same event outside closure
		5	AVO response of the target is distinctly different from that of the same event outside closure.
19	Multiple stacked indicators on the same trap (same stratigraphic sequence)?	**1**	None
		2	Slight indication of second AA in closure at a different level.
		3	Moderate evidence for second AA, with some consistency to closure
		4	Second well-defined AA at a different level shows similar relation to closure
		5	Three or more levels of AA's with similar signatures and similar relationship to closure
20	Similar indicators on other parts of closure?	**1**	None
		2	Possible seismic indicator is seen in a related fault block on closure,
		3	Good evidence for a second related AA on another portion of same closure but not totally consistent.
		4	Good evidence for several related AA's on other portions of closure but not totally consistent.
		5	All separate fault blocks on this closure have similar seismic indicators.
21	Velocity pull-down	**1**	None
		2	Slight indication, but probably structural or stratigraphic effect
		3	Moderately strong velocity sag, likely caused by velocity, but might be due to other effects
		4	Strong indication of low velocity anomaly
		5	Strong, uniform velocity sag consistent enough to allow estimate of pay thickness
22	Amplitude and frequency shadow beneath anomaly (may be evident at different spectral decomposition frequencies)	**1**	No shadow-zone observed
		2	Data quality not sufficient to identify shadow-zone, if present.t
		3	Slight evidence for shadow-zone under strong anomaly.
		4	Fair evidence for shadow-zone under strong anomaly.
		5	Clear shadow-zone under strong anomaly.

23	Have indicators been proven nearby? (true positive)	**1**	No known positive tests
		2	At least one geological success from a similar seismic signature in this play.
		3	Several known geological successes in this play have similar seismic signatures.
		4	One or more known discoveries (commercial or non-commercial) in this play have similar seismic signatures.
		5	Several discoveries (commercial or non-commercial) nearby or at least one on adjacent structure at target formation have nearly identical seismic anomalies.
24	Have indicators been disproved nearby? (false positive)	**1**	Several dry holes nearby and at least one dry hole on adjacent structure have nearly identical seismic anomalies.
		2	Several known dry holes in this play have nearly identical seismic anomalies.
		3	At least one dry hole in this play has a nearly identical seismic anomaly.
		4	At least one dry hole from a similar seismic anomaly in this play.
		5	No known negative tests
25	Sealing capacity for interpreted Column Height (CH) of this anomaly	**1**	A well-controlled Effective Stress estimate of less than 700 psi indicates likelihood of no seal. Caution!! In deep-water settings this situation may occur even for small columns if prospect has low overburden pressure.
		2	An Effective Stress estimate or analog fields support maximum CH equal to OR less then apparent amplitude anomaly height
		3	No Effective Stress estimate or analog fields available - max CH unknown.
		4	An Effective Stress estimate or analog fields support maximum CH equal to OR greater than apparent amplitude anomaly height
		5	An Effective Stress estimate or analog fields support maximum CH much greater than apparent amplitude anomaly height.
26	At the anomaly level, how confident are you of preservation? (no late fault movement, breaching, tilting)	**1**	Good evidence of late fault leakage or large structure tilting.
		2	Fair evidence of late fault leakage or structure tilting.
		3	Some late fault movement or tilting, but not considered significant.
		4	No late faulting or tilting observed.
		5	Good paleo-structure with no late faulting or tilting.

The characteristics in Table 2 are extracted from SAAM. The characteristic numbering was kept the same even though some of them were discarded for this study to avoid confusion. The criteria in discarding characteristics were either not having enough data points or being not well related to the amplitude anomaly itself. Figure 1.7 shows a matrix of values plotted to see the distribution of data points on determining variables to use. The variables without a significant distribution of values were removed from the model. Even though the stepwise regression approach would eliminate these statistics, it is good practice in any regression analysis to exclude these variables to avoid any errors or biased prediction.

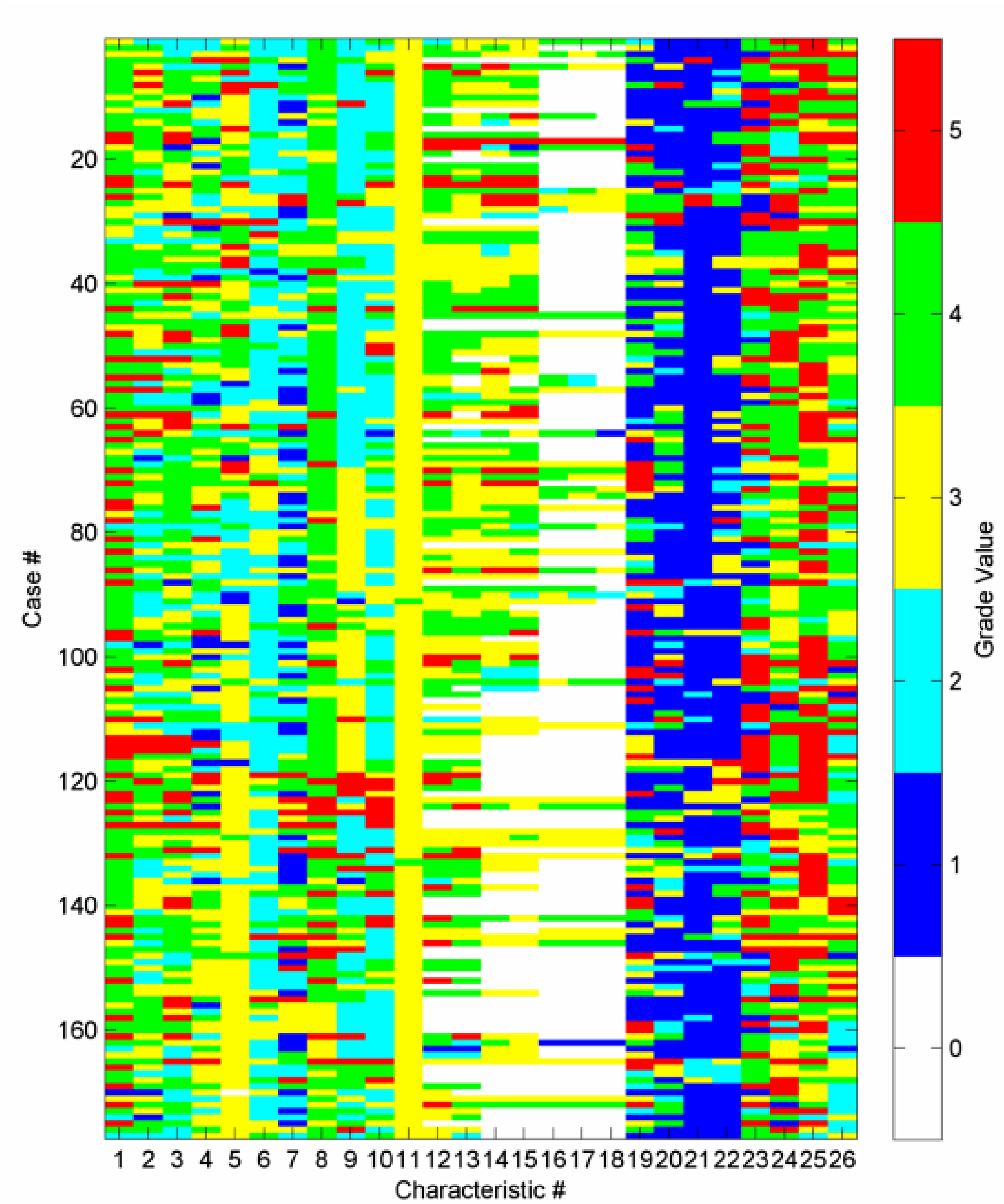


Figure 1.7: A matrix including all characteristics and values colored according to the grades.

The distribution of grade values of a variable within itself would also bias the prediction. Figure 1.8 shows the number of answers for grade of each characteristic. The characteristics without a close-to-even distribution will have negligible significance to the outcome. A mostly repeating grade (e.g. characteristic #11, and #21), especially zero values, are to be expected to be excluded with the stepwise regression routine.

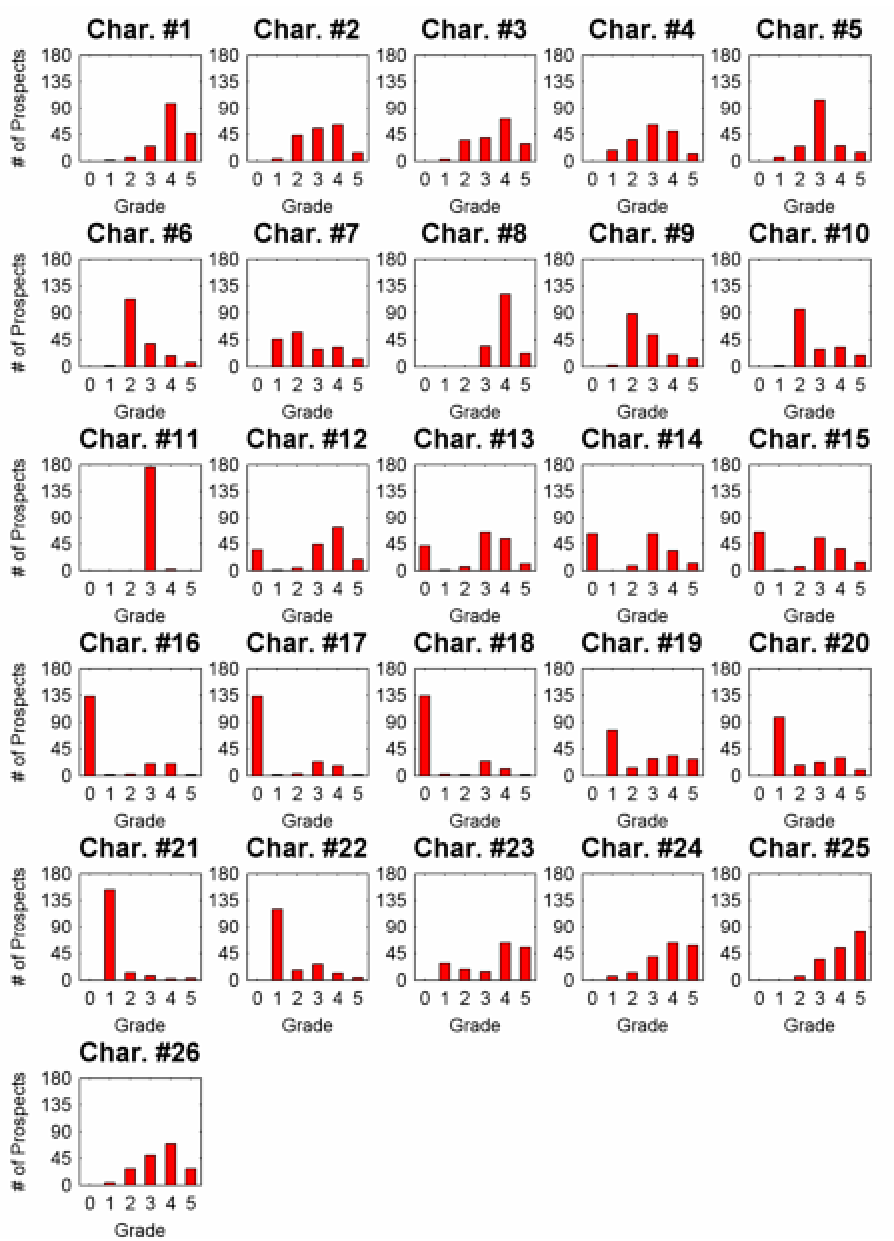


Figure 1.8: Distribution of grade values for every DHI characteristic.

The chosen input characteristics were:

[1] Amplitude change (as viewed on stacked P wave seismic data)

[2] Consistency within mapped target area (on stacked data)

[3] Are unexplained anomalies seen on stacked data outside closure (within same stratigraphic sequence)?

[4] Down-dip conformance (fit to closure) based on far-offset or stacked data

[5] Lateral conformance based on far-offset or stacked data

[6] Flat Spots indicating fluid contacts

[7] Phase or character change at the down-dip edge of the anomaly ('tuning' or interference effect caused by a change in fluid content)

[8] Signature match vs expected (polarity and shape)

[13] Change in AVO compared to model (wet vs HC filled)

[14] Excluding possible stacked pays, is the AVO effect anomalous compared to reflectors above and below?

[15] Is the AVO effect anomalous compared to the same event outside closure?

[19] Multiple stacked indicators on the same trap (same stratigraphic sequence)?

[20] Similar indicators on other parts of closure?

[21] Velocity pull-down

[22] Amplitude and frequency shadow beneath anomaly (may be evident at different spectral decomposition frequencies)

[23] Have indicators been proven nearby? (true positive)

[24] Have indicators been disproved nearby? (false positive)

[25] Sealing capacity for interpreted Column Height (CH) of this anomaly

[26] At the anomaly level, how confident are you of preservation? (no late fault movement, breaching, tilting)

Chapter 2 - Method

2.1. Problem and Analysis

This study was designed to investigate the relationships between several DHI characteristic observations and prospect outcomes. To analyze these relationships, a simple multiple linear regression was applied to all the characteristics. After determining which single characteristic has the most influence out of all available options, a step-wise approach to update the starting model was applied. The criterion in picking the next characteristic was to look at the total mean-squared-error of the prediction (MSPE) while it is being added. Figure 2.1 shows the algorithm used to pick characteristics that uses the following steps:

- Start with empty model
- Test for each character by itself
- Pick the one with the lowest MSPE
- Add the picked variable to the model
- Keeping the previously selected variable(s), test adding the remaining variables
- Pick the next variable with the lowest MSPE
- Add the picked variable to the previous model
- REPEAT until a pre-determined number of iterations is reached

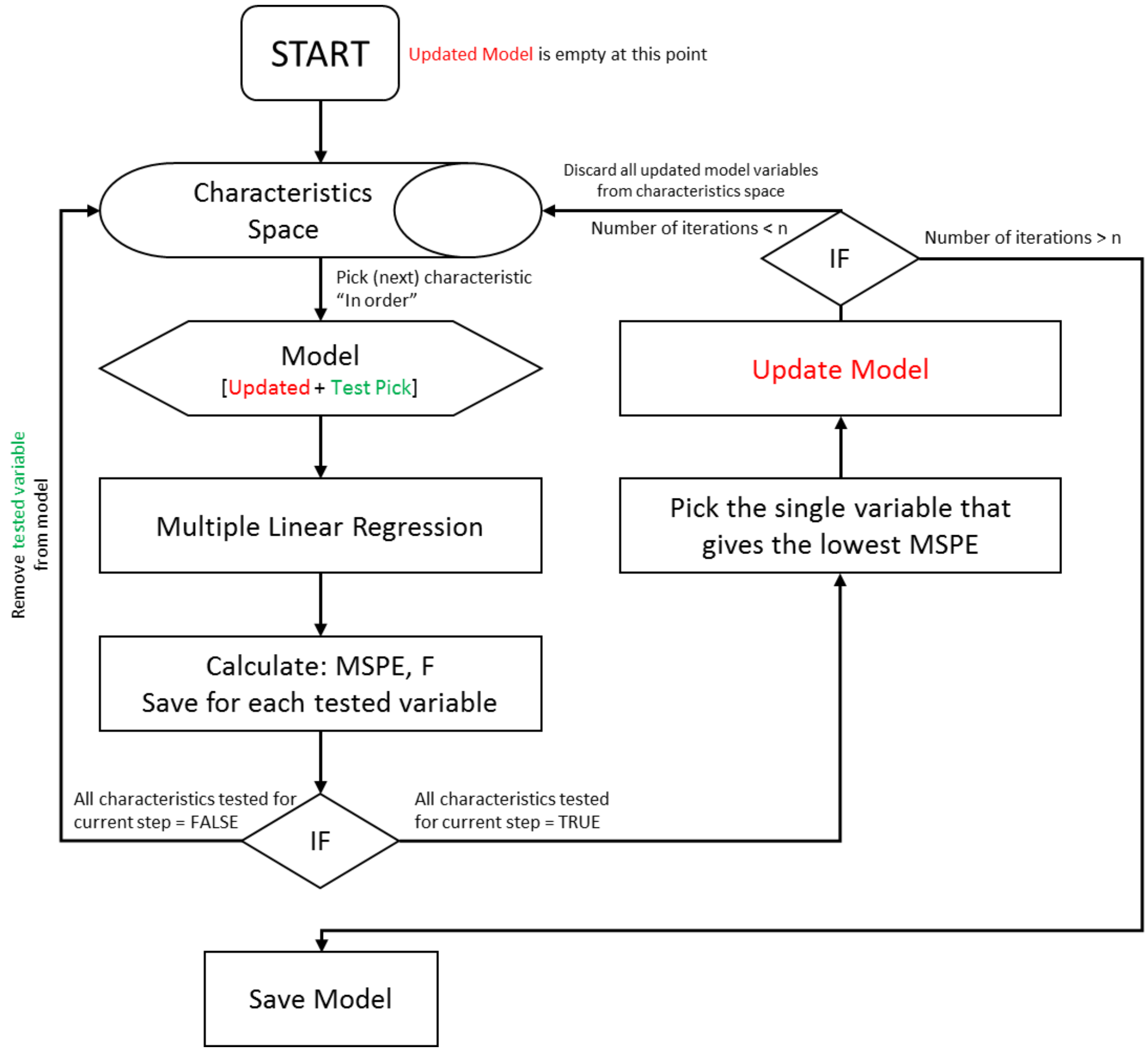


Figure 2.1: Flow Chart for the algorithm used

2.2. Multiple Linear Regression

A linear model of multiple variables was used to approximate the drilling outcomes. Many of the regression problems involve using multiple variables. Regression is a supervised learning technique. The goal of multiple linear regression is to predict the value of one or more target dependent variables given the value of multiple independent variables (characteristics in this study). The form of regression worked in this study was to use a linear model function to fit to the data. A possible multiple regression model is:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k \qquad \text{(Eq. 2.1)}$$

where Y is the binary outcome, β_k are the regression coefficients for the independent variables including the β_0 intercept term, and x_k (k = 1, 2, …, 26) are the independent variables. Least-squares minimization could be used to solve for the regression coefficients for this model. Assuming there are more observations n ($n = 177$) than number of variables, a regression table can be written as Table 3.

Table 3: Regression table (x).

$\boldsymbol{Y}$	$\boldsymbol{x_1}$	$\boldsymbol{x_2}$	…	$\boldsymbol{x_k}$
$\boldsymbol{Y_1}$	x_{11}	x_{12}	…	x_{1k}
$\boldsymbol{Y_2}$	x_{21}	x_{22}	…	x_{2k}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
$\boldsymbol{Y_n}$	x_{n1}	x_{n2}	…	x_{nk}

The least-squares function would be given by

$$L = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{ij} \right)^2 \qquad \text{(Eq. 2.2)}$$

The estimates must satisfy equations 2.3 and 2.4.

$$\left.\frac{\partial L}{\partial \beta_0}\right|_{\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k} = -2 \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \right) = 0 \qquad \text{(Eq. 2.3)}$$

$$\left.\frac{\partial L}{\partial \beta_j}\right|_{\hat{\beta}_0, \hat{\beta}_1, \ldots, \hat{\beta}_k} = -2 \sum_{i=1}^{n} \left(y_i - \hat{\beta}_0 - \sum_{j=1}^{k} \hat{\beta}_j x_{ij} \right) x_{ij} = 0, \qquad j = 1, 2, \ldots, k \qquad \text{(Eq. 2.4)}$$

Solving these equations, the least-square normal equations can be derived.

$$
\begin{aligned}
n\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{ik} &= \sum_{i=1}^{n} Y_i \\
\hat{\beta}_0 \sum_{i=1}^{n} x_{i1} + \hat{\beta}_1 \sum_{i=1}^{n} x_{i1}^2 + \hat{\beta}_2 \sum_{i=1}^{n} x_{i1}x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{i1}x_{ik} &= \sum_{i=1}^{n} x_{i1}Y_i \\
&\vdots \\
\hat{\beta}_0 \sum_{i=1}^{n} x_{ik} + \hat{\beta}_1 \sum_{i=1}^{n} x_{ik}x_{i1} + \hat{\beta}_2 \sum_{i=1}^{n} x_{ik}x_{i2} + \cdots + \hat{\beta}_k \sum_{i=1}^{n} x_{ik}^2 &= \sum_{i=1}^{n} x_{ik}Y_i
\end{aligned}
\qquad \text{(Eq. 2.4)}
$$

The solution of the normal equations (2.4) lead to the least-square estimations of the regression coefficients β_k.

After generating the model the mean-squared-error for prediction (2.5) and F value (2.6) is calculated for the analysis of the results.

$$
MSPE = \left(\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i \right)^2 \right) / n \qquad \text{(Eq. 2.5)}
$$

$$
F = \frac{R^2/k}{[(1 - R^2)/(n - k - 1)]} \qquad \text{(Eq. 2.6)}
$$

Where Y_i is the observed outcome, $\hat{Y}_i$ is the predicted outcome, and R^2 is the ratio of explained variation to total variation, also known as the coefficient of determination.

The MSPE and F Value were used in the regression routine to figure out which characteristics to add and when to stop adding more to the model.

Chapter 3 - Results

3.1. In-Sample Calibration

To begin, the first 89 samples of the Class III database were chosen to generate the model weights. Figure 3.1 shows the MSPE while adding characteristics. Figure 3.2 shows the F values for the same process.

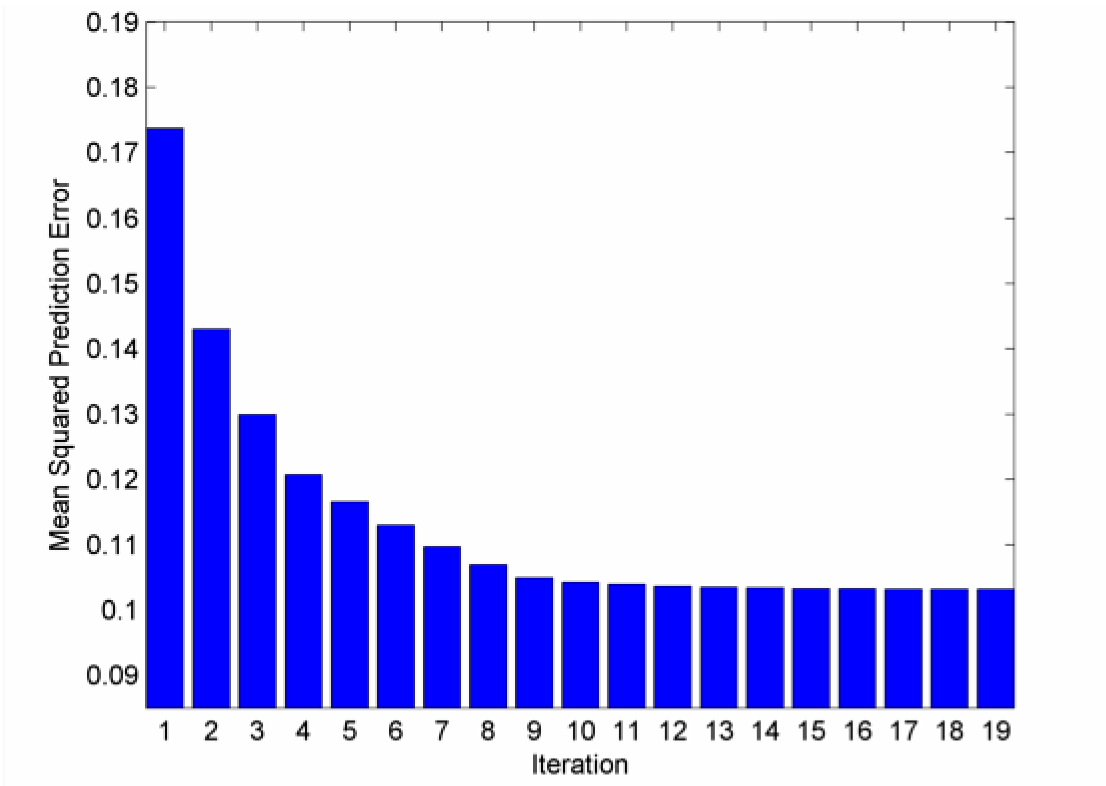


Figure 3.1: Mean-squared-error for model prediction on every iteration.

On the MSPE graph, it can be clearly seen that there is a big drop when adding the significant characteristics in the first iterations. As the number of iterations is increased, the decrease in MSPE becomes smaller, because the less significant characteristics start to influence the model. The F value also keeps dropping as the degrees of freedom are increased when adding new characteristics.

By looking at these results, it is a safe assumption to say that the characteristics added after the 10th iteration can be excluded from calibration. The characteristics that were chosen to be excluded from the models in this way still keep to drop the F value while keeping the MSPE almost the same, which means they are not contributing any positive information to the prediction.

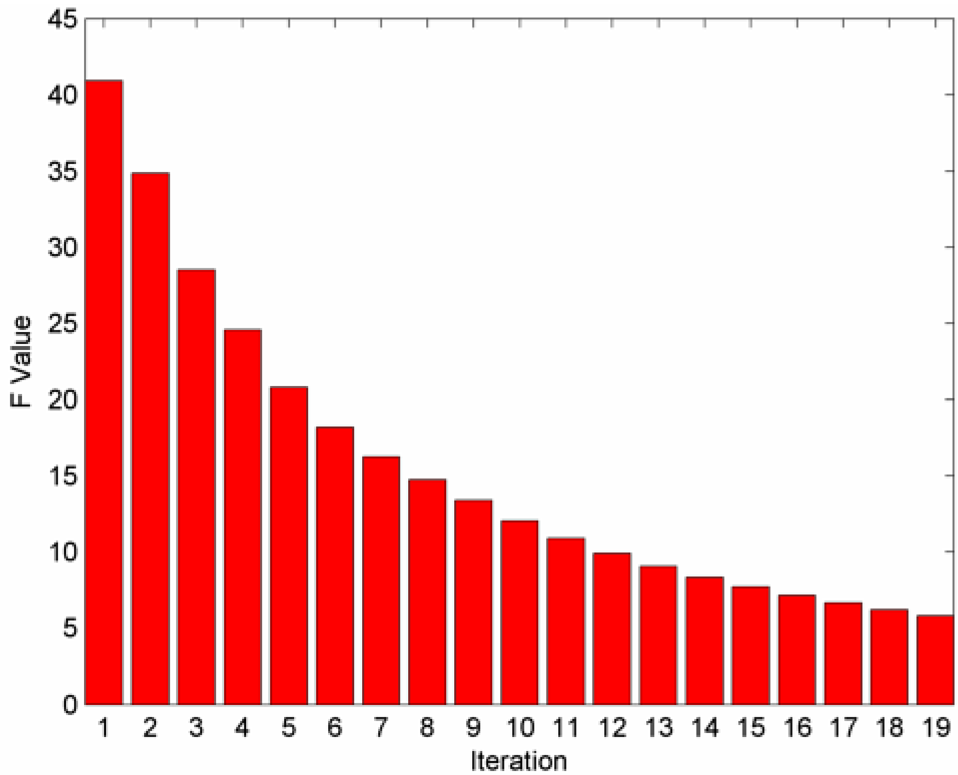


Figure 3.2: F Statistic for model prediction on every iteration.

The predictions for the in-sample prospects are calculated at each step and plotted with the corresponding observed drilling outcomes. Figure 3.3 shows the predictions for a model generated by 10 iterations. From left to right, it is seen that the prediction results get better. To look at the deviation from the observed outcomes, the residuals must be plotted. Figure 3.4 shows the residuals distribution for the calculated model. The results show a distribution which has a minimal number of large errors (flanks) and a big number of very close predictions.

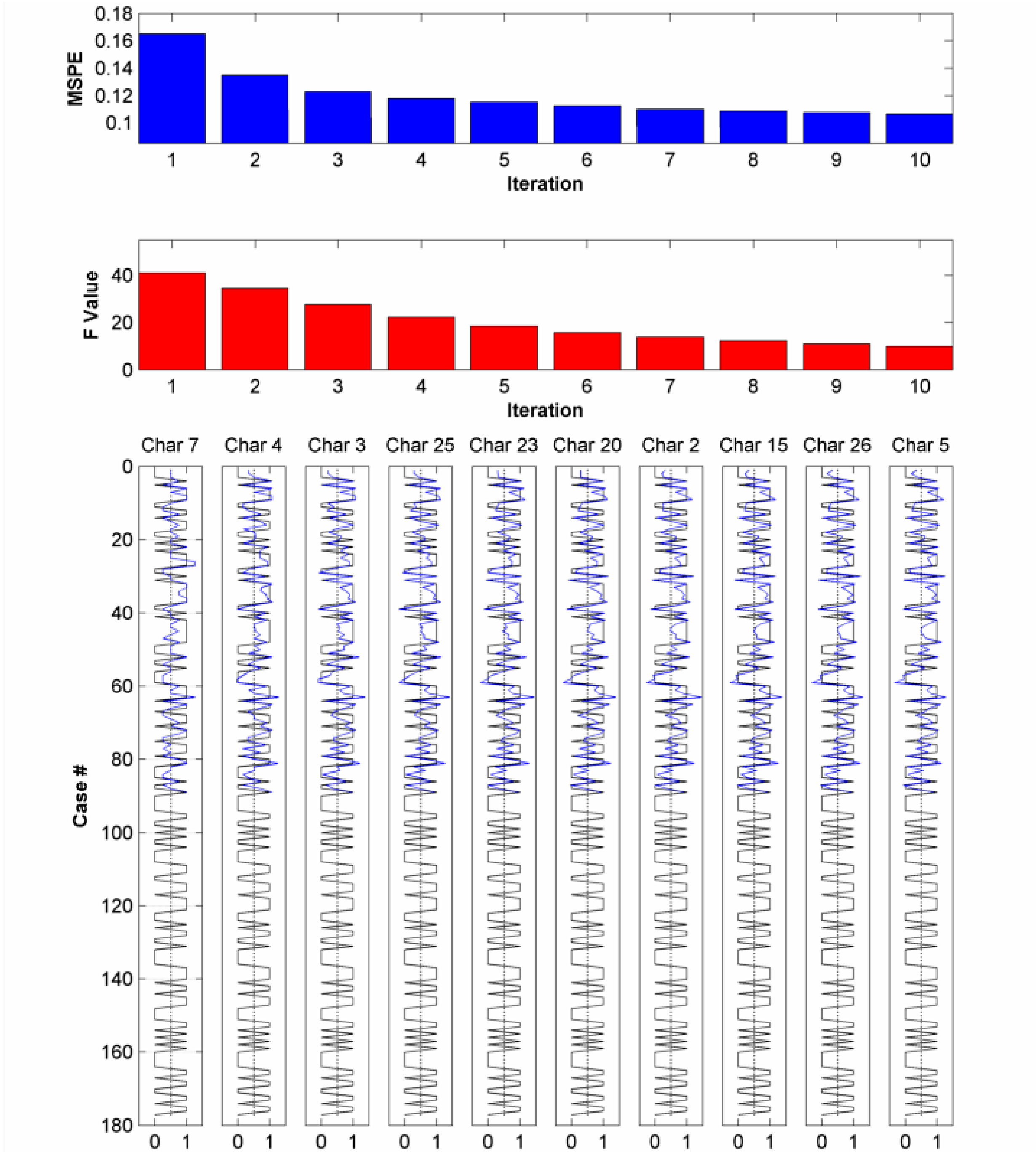


Figure 3.3: In-sample predictions vs. observations for 10 iterations.

According to the stepwise regression routine, the order of importance of added characteristics is: **7, 4, 3, 25, 23, 20, 2, 15, 26, and 5**, which will be discussed in the conclusions chapter.

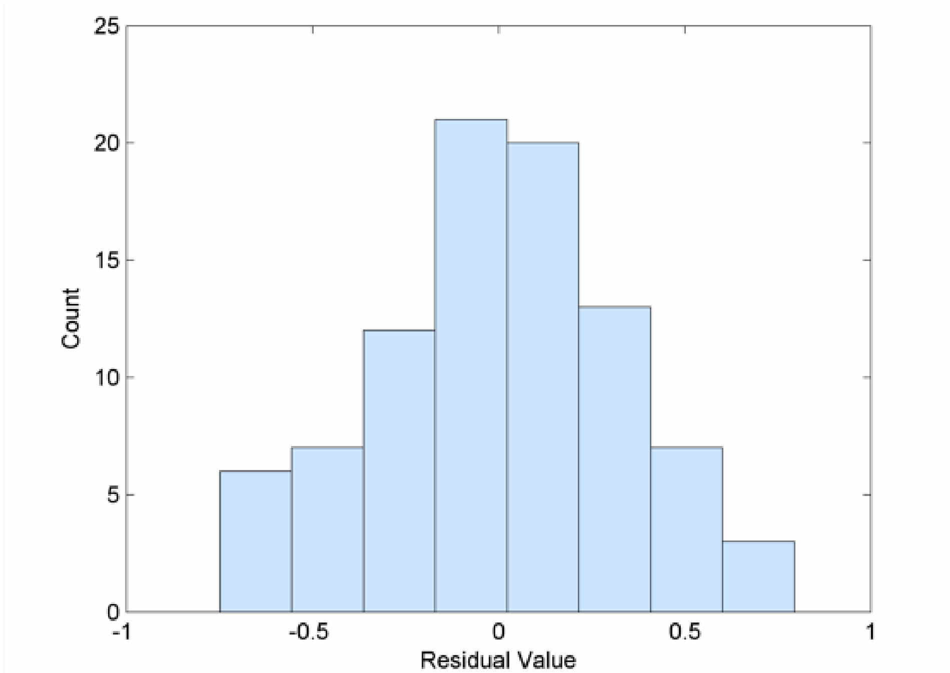


Figure 3.4: Residuals for the in-sample calibration.

The predicted outcomes are rounded to the nearest valid answer (0 for values smaller than 0.5, and 1 for values bigger than 0.5). After this calculation, the percentage of correct predictions was calculated. The accuracy of the prediction in the calibration dataset is 88.5% for successful wells, and 83.8% for failure wells.

3.2. Out-Sample Predictions

For the remaining 88 cases, the results were predicted using the in-sample calibrated model. This test is intended to ensure that the prediction is reliable. The residuals and the amount of error are expected to be higher than the in-sample data, but should still be in acceptable limits. The accuracy of out-sample tests are calculated to be 84% for successful wells, and 74% for failure wells. Figure 3.5 shows the results displayed in Figure 3.3 with the addition of out-sample predictions. By only using the first picked characteristic, the prediction is biased towards “success”. Addition of extra variables centers the prediction and generates more accurate results.

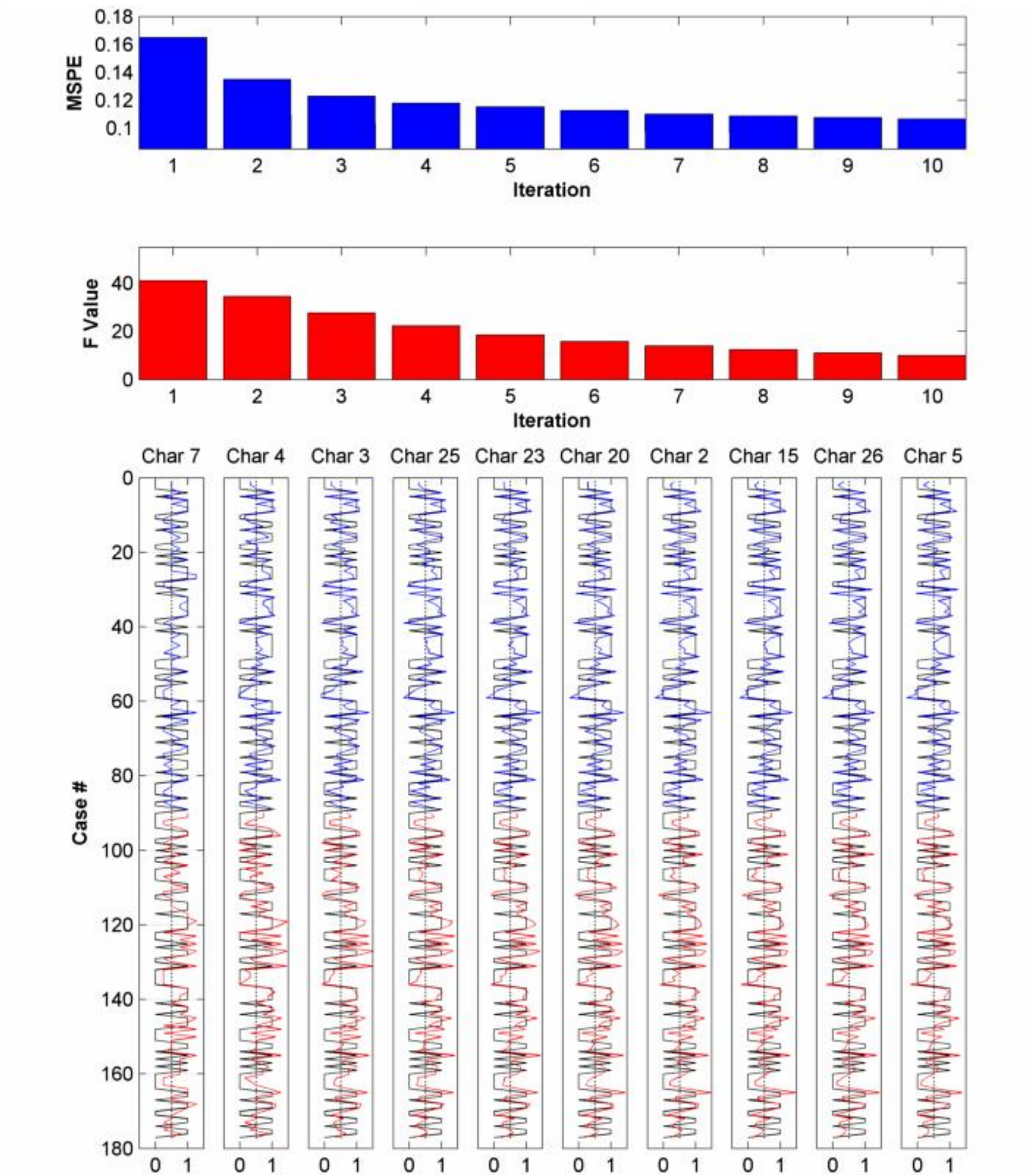


Figure 3.5: Out sample (red) and in sample (blue) predictions vs. observations for 10 iterations.

The residual values for the out-sample tests shown on Figure 3.6 shows a good distribution with possible outliers. The last samples from 170 to 177 seem to have poor prediction compared to other cases. This could be due to data quality, lack of data, or ambiguous DHI characteristics.

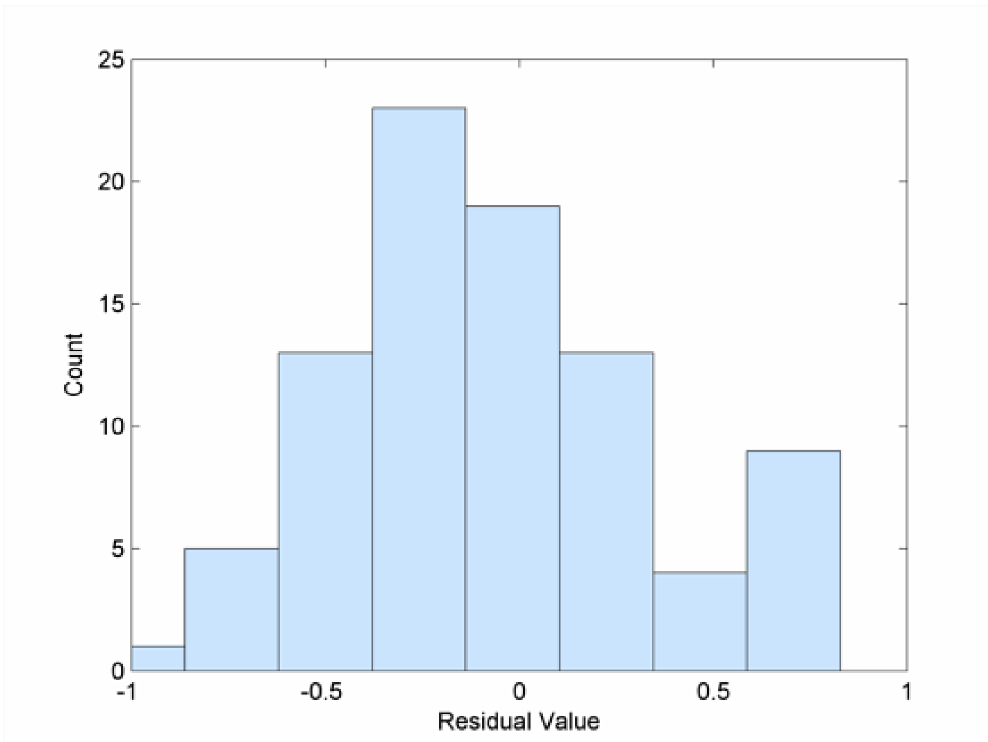


Figure 3.6: Residuals for the out sample predictions.

Chapter 4 - Conclusions

4.1. Limitations

The results of multiple linear regression in a complex scenario discussed in this study mostly rely on initial estimates of the interpreters, which can be subjective or biased. It seems likely that constraining or calibrating the regression models increases the chance of getting the best model. Furthermore, the possibility of having inter-correlations of DHI characteristics will also affect the final result.

Not all prospects carry the same kind of DHIs. Some of the characteristics used to define prospects may or may not be evident for one, whereas another prospect may show highly descriptive DHIs. However, grouping of prospects with similar DHIs would reduce the number of data points to regress and may lead to incorrect or biased results.

Theoretically, using as many as possible data points play a key role in achieving the best models. With the increasing number of prospects in SAAM database, it will be possible to separate the data into smaller and more detailed groups to analyze the characteristics in depth.

4.2. Resulting Characteristics

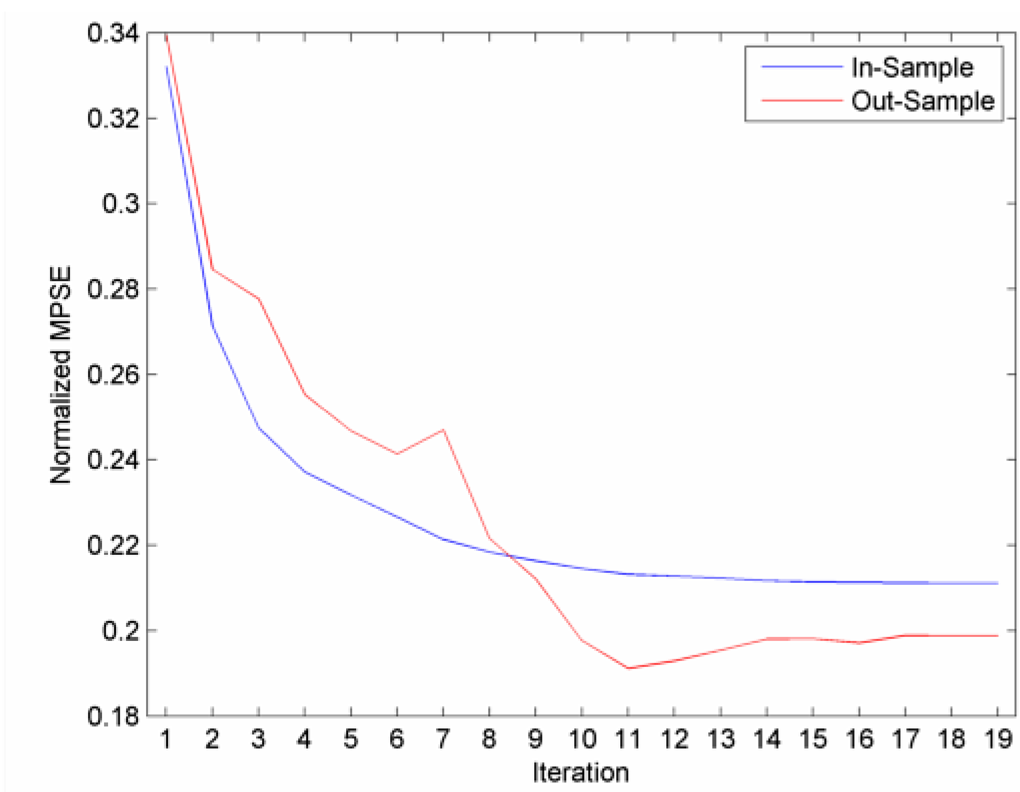


Figure 4.1: MPSE comparison for all used characteristics by iteration.

Calculation of predicted values for in-sample and out sample tests should allow one to generate MPSE values for both with each iteration. Figure 4.1 shows that there is a mutual decrease of MPSE in both tests until the 11th iteration. The out-sample tests' MPSE starts to increase after iteration 11. This supports the previous decision of discarding parameters after iteration 10.

The order of characteristics selected by the algorithm gives an insight to the importance of the characteristics. The order and the description of the results is shown in Table 4.

Table 4: Order of selected characteristics using multiple linear regression, and descriptions.

Order	Char #	Description
1	7	Phase or character change at the down-dip edge of the anomaly ('tuning' or interference effect caused by a change in fluid content)
2	4	Down-dip conformance (fit to closure) based on far-offset or stacked data
3	3	Are unexplained anomalies seen on stacked data outside closure (within same stratigraphic sequence) ?
4	25	Sealing capacity for interpreted Column Height (CH) of this anomaly
5	23	Have indicators been proven nearby? (true positive)
6	20	Similar indicators on other parts of closure?
7	2	Consistency within mapped target area (on stacked data)
8	15	Is the AVO effect anomalous compared to the same event outside closure?
9	26	At the anomaly level, how confident are you of preservation? (no late fault movement, breaching, tilting)
10	5	Lateral conformance based on far-offset or stacked data
11	13	Change in AVO compared to model (wet vs HC filled)
12	24	Have indicators been disproved nearby? (false positive)
13	21	Velocity pull-down
14	22	Amplitude and frequency shadow beneath anomaly (may be evident at different spectral decomposition frequencies)

15	8	Signature match vs expected (polarity and shape)
16	1	Amplitude change (as viewed on stacked P wave seismic data)
17	6	Flat Spots indicating fluid contacts
18	19	Multiple stacked indicators on the same trap (same stratigraphic sequence)?
19	24	Have indicators been disproved nearby? (false positive)

The results clearly show that the presence and correct interpretation of DHIs are correlated to the DHI prospect success/failure. This study shows that the importance of structurally conformant and continuous amplitude anomalies, and its seismic phase behavior, as well as, presence of proven-to-be-successful DHI analogs have strong correlation with the outcome.

4.3. Discussion and Recommendations

It is important to note that these results are isolated to the samples used in this study. When using different inputs as the calibration (i.e. picking different or random in-sample cases), the order of the output characteristics may change due to the sensitivity of the method to the kinds of observed DHIs. However, the first 10 picked characteristics seem not to be significantly affected for the majority of the tests, with very little number of exceptions.

By looking at the results, the seal capacity is in the Top 5 characteristics. This is an unexpected result. However, as seal capacity is an often neglected factor, its high ranking in the stepwise regression may have significant practical implications.

One important discarded characteristic, that is not included in the Top 10, is the presence of flat spots. It is well known that presence of flat spots is a highly trusted DHI in real world exploration. The reason it is not included in the significant characteristics list is the very low statistical variation of the input. Characteristic #6 on Figure 1.8 shows that most of the prospects do not have a flat spot present. The number of remaining prospects with observed flat spots are not statistically significant for regression analysis.

In this study it is chosen to linearly represent the relationship. It is challenging to represent DHI properties with fixed values, due to the complexities and non-uniqueness of real world conditions. The very likely nature of the problem is to be non-linear and must be studied in more detail. By analyzing all characteristics individually and then all-together to come up with a basis function to be used in the regression routine is highly recommended and might improve the results.

Another recommended approach would be to partition the prospects with similar DHIs for each class. Different DHIs are expected to have different properties that may not be applicable to others. The partitioning process would eliminate these errors. Unfortunately, with the number of data points in the current database, partitioning would greatly reduce the number of variables for each segment. This may cause the regression to be rank-deficient (not enough prospects for the number of characteristics). With increasing number of prospects in the database, it will be possible to do this in the future.

References

Aki, K., and Richards, P. G., 1980, Quantitative seismology, Theory and methods: W.H. Freeman & Co.

Backus, M. M., and R. L. Chen, 1975, Flat spot exploration: Geophysical Prospecting, **23**, 533-577

Bishop, C. M., 2009, Pattern Recognition and Machine Learning, Springer, 8th Edition

Brown, A. R., 2011, Interpretation of three-dimensional seismic data, AAPG Memoir, **42**, 7th Edition

Castagna, J., Swan H., and Foster D., 1998, Framework for AVO gradient and intercept interpretation: Geophysics, **63**, no. 3, 948–956

Forrest, M., Roden, R., and Holeywell R., 2010, Risking seismic amplitude anomaly prospects based on database trends: The Leading Edge, **29**, 936–940

Ostrander, W. J., 1984, Plane-wave reflection coefficients for gas sands at non-normal angles of incidence: Geophysics, **49**, 1637–1648

Roden, R., Forrest M., and Holeywell R., 2012, Relating seismic interpretation to reserve/resource calculations: Insights from a DHI consortium: The Leading Edge, **31**, 1066-1074

Roden, R., Forrest M., and Holeywell R., 2005, The impact of seismic amplitudes on prospect risk analysis: The Leading Edge, **24**, 706–711

Rutherford, S. E. and Williams R. H., 1989, Amplitude versus offset variation in gas sands: Geophysics, **54**, 680–688

Shuey, R.T., 1985. A simplification of the Zoeppritz equations: Geophysics, **50**, p. 609-614

Zoeppritz, K., 1919, Erdbebenwellen VIIIB, On the reflection and propagation of seismic waves: Göttinger Nachrichten, I, 66-84

Appendix A: DHI Consortium & SAAM Information

The DHI Interpretation and Risk Analysis Consortium began in 2001 with the support of Rose & Associates and contributions of 41 oil companies. The software SAAM developed under this consortium provides interpreters with a powerful tool for resource evaluations. The conventional approach to determine the chances of finding flowable hydrocarbons, which is Pg, is to risk five main independent risk factors source, timing/migration, reservoir, closure, and containment (Table A.1) accordingly and multiplying the percentages. If there is a seismic amplitude anomaly related with the prospect, the analysis of such anomaly can decrease the uncertainty of analysis (Forrest et al., 2010; Roden et al., 2005).

The goals of the DHI Consortium have been and continue to be (Roden et al., 2012):

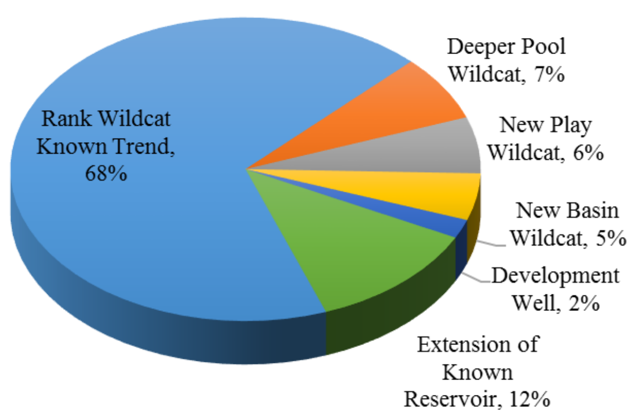
1) Gain a better understanding of how DHI anomalies impact predrill chance of geological success.
2) Characterize DHI anomalies observed using previous prospect reviews and discussions about risk analysis.
3) Create and archive a database of drilled prospect results.
4) Use the archived prospect database for improving the prediction ability of Pg and reducing the uncertainty.
5) To be a checklist and an educational tool for DHI interpreters in the analysis process to help risk seismic amplitude anomalies.
6) To discuss technologies for amplitude interpretation.

Table A.1: Conventional geologic chance factors for determining Initial Pg independent of seismic amplitude anomaly (Modified from Forrest et al., 2010; Roden et al., 2005)

Risk Element	Confidence of…
Source Rock	• Area and thickness • Richness • Thermal maturity • Kerogen type
Timing/Migration	• Closure timing (before/during migration) • Migration distance and pathways
Reservoir Rock	• Facies and Extent • Minimal Thickness • Reservoir Quality
Closure	• Depth/Shape of closure • Structural or stratigraphic • Confidence in mapping
Containment	• Sealing capacity • Preservation

The SAAM software includes 217 Class I, II, III, and IV prospects from all around the world with an approximately equal number of successful wells and dry. Closure types include structural, stratigraphic, and combinations of the both. Figure A.2 shows that the database is dominated by exploration wells with only 14% coming from development or extension wells (Roden et al., 2012).

Figure A.2: Prospect types of well in the database (Modified from Roden et al., 2012)



The DHI prospects in the database are categorized and risked as AVO classes I – IV (Figure 1.6). 76% of the prospects are class 3, 22% are class II sands. Only a small number of class I and class IV prospects are present.

SAAM software uses the following input information:

- Geologic context of the prospect
- Initial Pg, which is the product of geologic chance factors independent of the anomaly
- Seismic data quality
- Rock physics and data quality
- DHI characteristics for each relevant class

By this information, the DHI characteristics are graded on a scale of 1 – 5 by the estimator(s) based upon a series of objective criteria, with 1 being the worst and 5 being the best case scenario. Then, each grade is converted into a "Grade Value" ranging from -10 to 10, which describes how favorable is the characteristic to the prospect. Then a weight is applied to each characteristic and the values are summed to come up with an overall score for that prospect. This score is then normalized to a range from 0 to 100 out of all possible outcomes. Then a similar process is applied to seismic data quality information and this gives a data quality score ranging from 0 (little or no data reliability) to 100 (highest possible reliability). The prospect score is then multiplied by this value to achieve the final DHI Index. For example, in a perfect world scenario, if all characteristics are graded as 5, the prospect score would be 100. If the data quality operator is 50%, the DHI Index will be the half of the DHI Index if data quality was neglected. Finally, the Initial Pg value is modified by the DHI index to get the Revised Pg.

After calculating prospect specific Revised Pg values, using the prospect library, the outputs of SAAM can be calibrated (Figure A.3) and recalculated for all the prospects to have a better correlation with the drilling outcome. Figure A.4 shows a simplified version of the SAAM workflow.

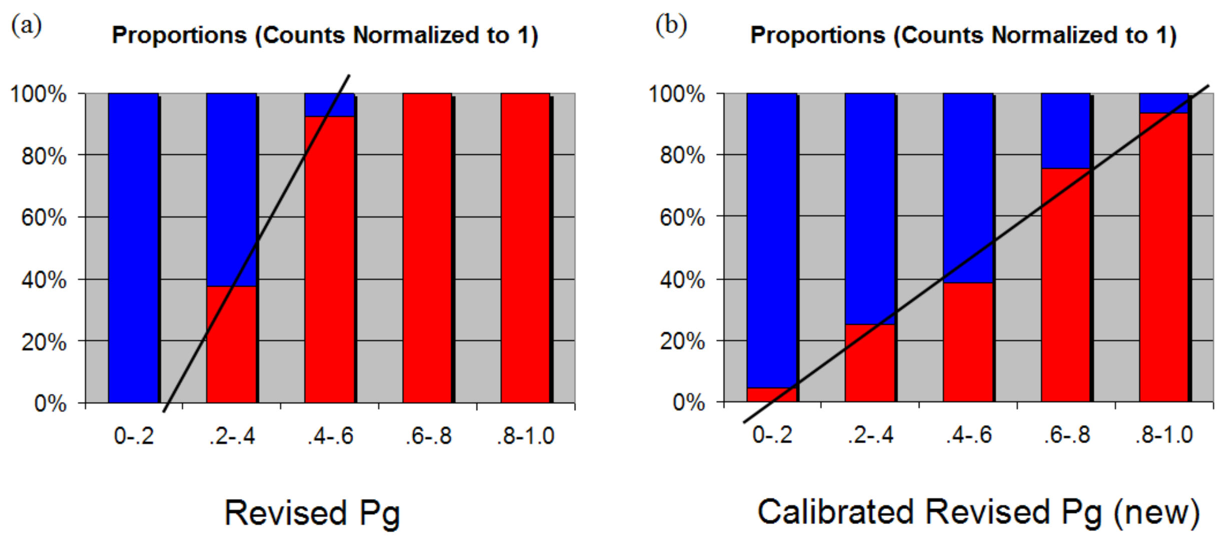
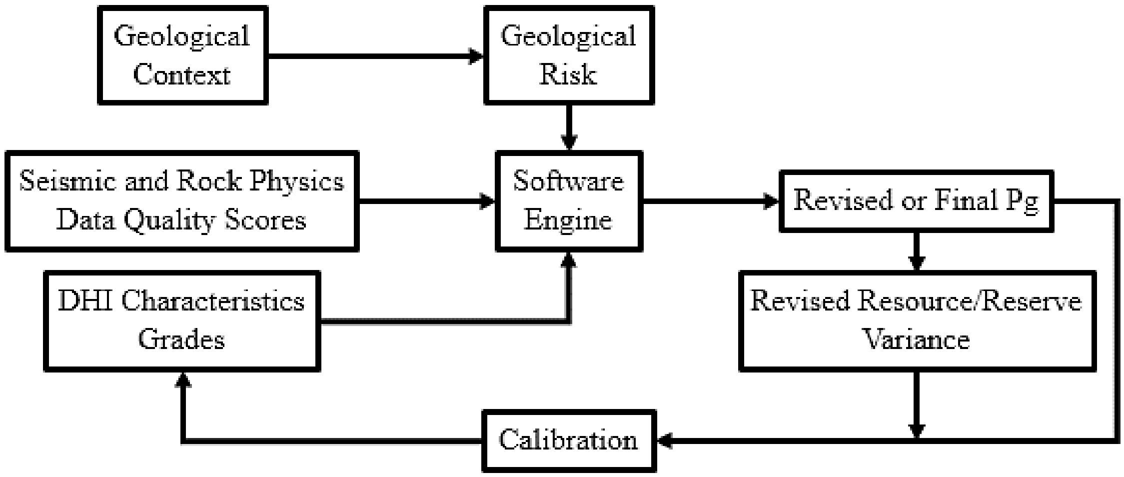


Figure A.3: On the left: uncalibrated Revised Pg values (a) which overestimate risk on the high end and underestimate risk on the low end; on the right: calibrated Pg estimate (b) which takes advantage of the prospect library. A well calibrated value should be a 45 degrees line on this chart (SAAM User's Guide)

Figure A.4: DHI Interpretation and workflow concept (Modified from Roden et al., 2012)



Currently SAAM offers three different calibration methods. The first method is *"Calibration from Revised Pg (Initial Pg + DHI Quality Index"* and it starts with Initial Pg + DHI Index and applies a calibration from known drilling outcomes from the database. However, the disadvantage of this method is that, it includes the Initial Pg in the calculation. The estimator must be very careful in making an Initial Pg estimate. The second method is *"Calibration from DHI Index"* which is the same method as the previous, but excludes the Initial Pg and only starts with the DHI Index. It only takes into account the strength of the DHI characteristics and the data quality which some may consider highly subjective. The third calibration option is *"Binary Bayesian Conditioning"*, which is an experimental approach. Further research on this method may produce additional enhancements.