SPE-185077-MS

Multivariate Analysis Using Advanced Probabilistic Techniques for Completion Design Optimization

Bertrand Groulx, Verdazo Analytics; Jim Gouveia, Rose & Associates; Don Chenery, Verdazo Analytics

This paper was prepared for presentation at the SPE Unconventional Resources Conference held in Calgary, Alberta, Canada, 15-16 February 2017.

Abstract

Efforts to identify optimal completion technology and design parameters are complicated by the compounding impacts of broad statistical variability in operations, reservoir/fluid and completion/wellbore design. There are several analysis approaches available to identify and optimize key completion design parameters. Each approach offers limited insight on its own, but combining a set of approaches into a disciplined methodology can collectively present a unique understanding of optimal completion technology and design. Traditional parallel coordinates visualizations offer strong visual cues of correlations, but in datasets with broad statistical variability they often convey a lack of correlation and fail to distinguish statistical trends. Statistical methods are unique in their ability to provide insights into non-continuous correlations where upper and lower thresholds exist; however, they are not effective at providing a deterministic measure of an individual input's effect on an outcome. Modelling and regression analysis can provide a means to measure the effect of several input variables on an outcome, but lack transparency and are often perceived as a "black box" solution with outcomes that have limited supporting evidence, or supporting evidence that is difficult to understand.

We demonstrate a robust multivariate analysis methodology using a hybrid approach involving the principles of parallel coordinates, dimensional normalization and advanced probabilistic techniques. One of the benefits of this approach is that it can yield statistically significant insights on sample sets as small as 80 wells. The methodology involves six steps that offer transparency to the analysis and facilitate a narrative of understanding:

1. Selection of a performance measure set
2. Analogue well selection
3. Selection of numerical completion design input parameters
4. Parallel Coordinates Distributions: input parameter impact analysis
5. Evaluation of analogue fitness and subset selection
6. Input Optimization Distributions: input optimization process

We found that the use of consistent dimensional normalization on both inputs and outcomes better isolates the impact of an input parameter. The shape and position of parallel coordinates distributions can illustrate nuances of impact that are lost in other multivariate approaches.

In this paper we apply and test this methodology on three major resource plays in the Western Canadian Sedimentary Basin: a gas play, a liquids-rich gas play and an oil play.

Introduction and Background

With a decade of completion information now readily available to North American producers, the analysis of completion and production data has become the norm. Given the variability in the data and the many input parameters at play, it is difficult to determine the impact of a single input parameter on production performance. Correlation analysis proves difficult, since the correlations tend to be weak and vary over the full range of values. Regression tools can provide outcomes that have limited supporting evidence, or supporting evidence that is difficult to understand.

The authors looked outside the oil and gas industry for visual multivariate analysis techniques that could be combined with the strengths of probabilistic techniques and deliver accessible, explainable insights. The objective of our investigation was to provide a visual methodology that could be used to identify patterns of behavior, fuel meaningful discovery through targeted investigations and cultivate an understanding of a completion input parameter's impact on production performance.

Purpose of this Study

To develop a scalable and repeatable visual analysis approach that:

- uses statistical techniques that are readily accessible to a broad audience
- offers data-driven statistical insights with visual nuances that can inform completion modelling first principles and advanced regression analysis
- is suitable for very large datasets and datasets as small as 80 wells
- can be effective even when all inputs are not available for all wells.

Hybrid Visualization Approach

Parallel Coordinates

The approach presented in this paper leverages the principles of parallel coordinates combined with advanced probabilistic techniques. Parallel coordinates, for completion analysis, would involve taking the range of values for a particular production performance measure on the first axis and plotting lines to the corresponding input parameters on subsequent axes for each well. The approach typically involves using "brushing" to colour the performance measure (e.g. into quartiles) to visually determine if there is a discernable range for each input parameter that strongly contributes to optimal performance. In datasets with a small number of influencing input parameters (low dimensionality), this visual technique can yield clear visual patterns and insights. An example of an effective use of Parallel Coordinates is shown in Figure 1. The weakness of Parallel Coordinates is revealed when there are many input parameters (high dimensionality) and "over-plotting" occurs, obscuring any visual insights (Figure 2).

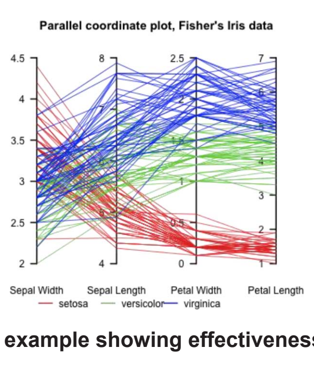


Figure 1—Parallel Coordinates example showing effectiveness of brushing (Wikipedia, 2016)

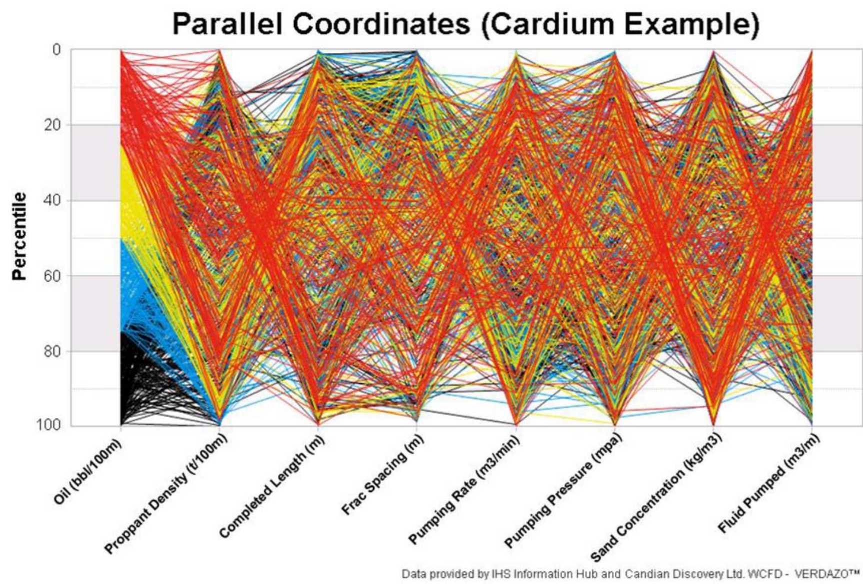


Figure 2—Parallel Coordinates Cardium completion parameters example showing over-plotting

Parallel Coordinates Distributions

Figure 3 illustrates how we applied the brushing technique of parallel coordinates to cumulative probability distributions in order to compare the distributions of an input parameter for each production performance measure quartile. This visual approach will be referred to as a Parallel Coordinates Distribution (PCD) chart.

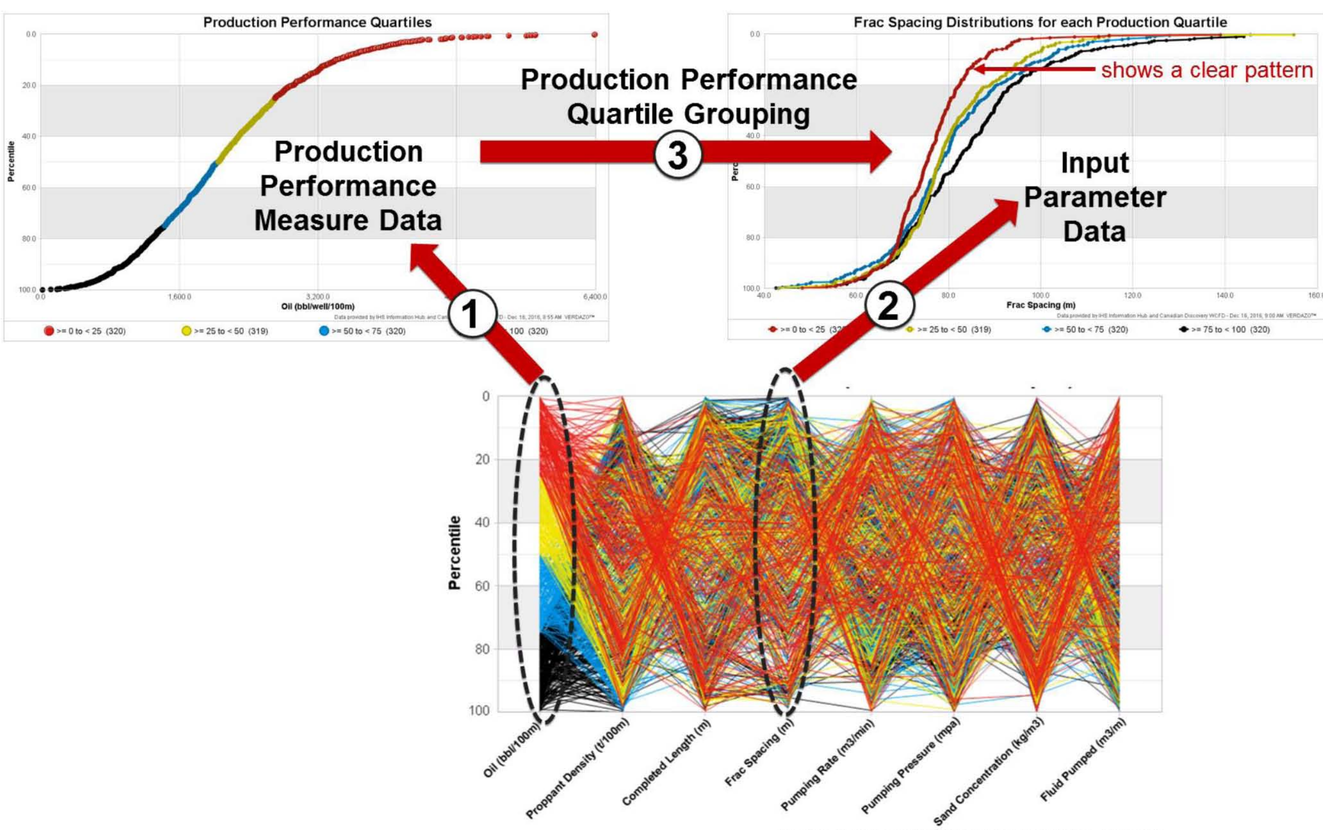


Figure 3—Parallel Coordinates Distribution chart methodology

The colouring convention used in this paper for all PCD charts is:

– the First Quartile (top quartile) as red (percentile range of 0 to 25)
– the Second Quartile as yellow (percentile range 25 to 50)
– the Third Quartile as blue (percentile range 50 to 75)
– the Fourth Quartile (bottom quartile) in black (percentile range 75 to 100)

Visual Nuances in the data

The relative placement of each quartile in PCD charts can illustrate visual nuances in one of, or a combination of, the patterns identified in Figure 4:

a. Clear Performance Pattern that shows a sequential progression from bottom quartile to top quartile (Figure 4a)
b. Thresholds where quartiles cross or coalesce (Figure 4b)
c. Correlation Windows which show a sequential performance progression between a lower threshold and an upper threshold (Figure 4c)
d. No Discernable Pattern, where the quartiles overlap and/or show no sequential progression (Figure 4d)

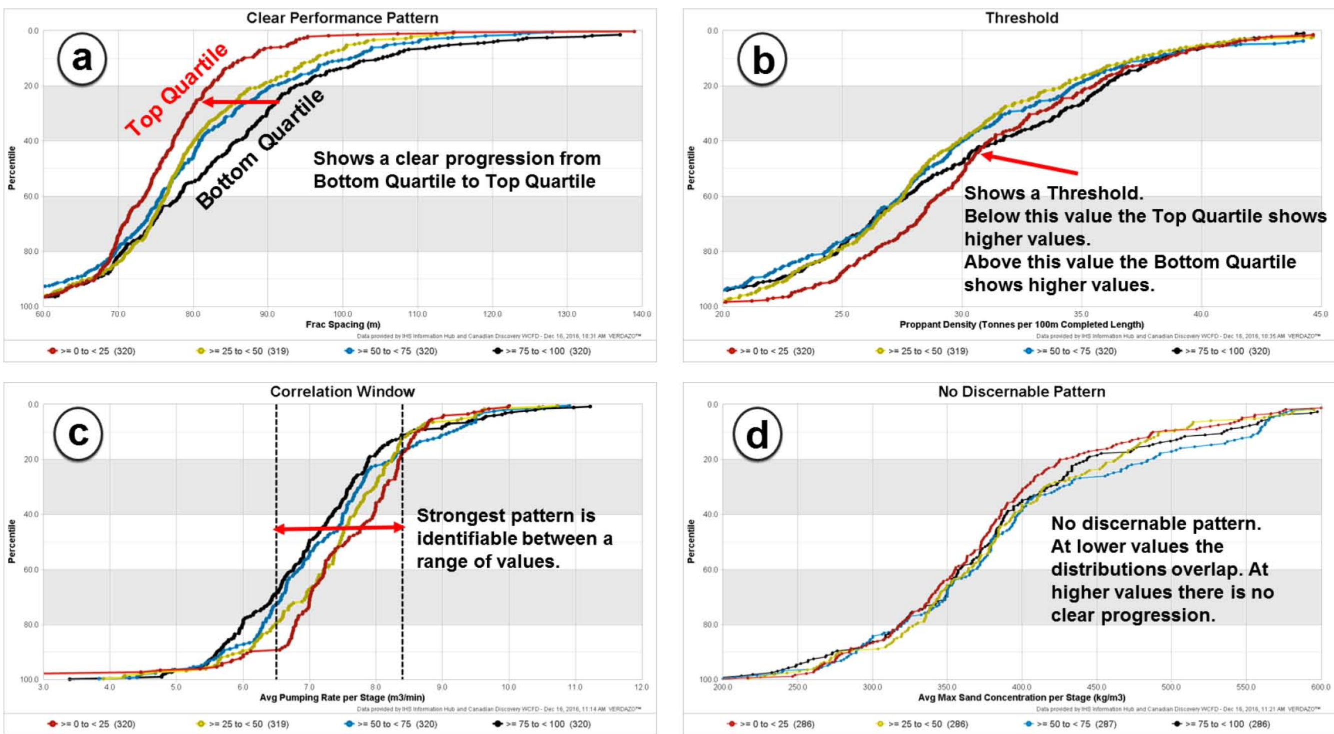


Figure 4—Parallel Coordinates Distribution Patterns

It is recommended that PCD charts be generated for multiple production performance measures and multiple input parameters. This is best presented in a Matrix of PCD charts (Figure 5) for easy identification of the input parameters that display the most discernable patterns. We found that dimensionally normalized variables (e.g. production/100 m) were consistently more effective at generating identifiable patterns, particularly when the input parameter was dimensionally normalized to the same, or a suitably comparable, measure. For example, if you are investigating Production/Stage, then suitable input parameters to use would be Proppant/Stage or Pump Rate/Stage.

Figure 5—Parallel Coordinates Distribution Matrix

The PCD methodology was tested on datasets ranging from more than 1000 wells to datasets as small as 80 wells, yielding 4 distributions of 20 wells per quartile. Figure 6 demonstrates how the 80-well pattern is similar to the pattern show in Figure 6b, which represents 1279 wells. The subset of 80 wells is denoted in the map in Figure 6a. Less than 80 wells would result in quartile distributions of less than 20 wells and may not adequately represent the range of values required for meaningful insights. The minimum number of wells required is a function of the variance of the dataset being analyzed. For lognormal distributions (for example, rates and reserves) the P10:P90 ratio is a proxy for variance. For a 20-well dataset, a P10:P90 ratio of 3 would have approximately a +/− 10% uncertainty around the mean outcome, and a P10:P90 ratio of 4 would be +/− 15%. As discussed in SPE 175527 (McLane & Gouveia) the uncertainty around the mean, and hence our confidence in the representativeness of the datasets decreases as we reduce the sample count.

When working with small datasets the user must ensure that the observed deltas between the datasets' mean values exceed the innate uncertainty of the sample size, as discussed above.

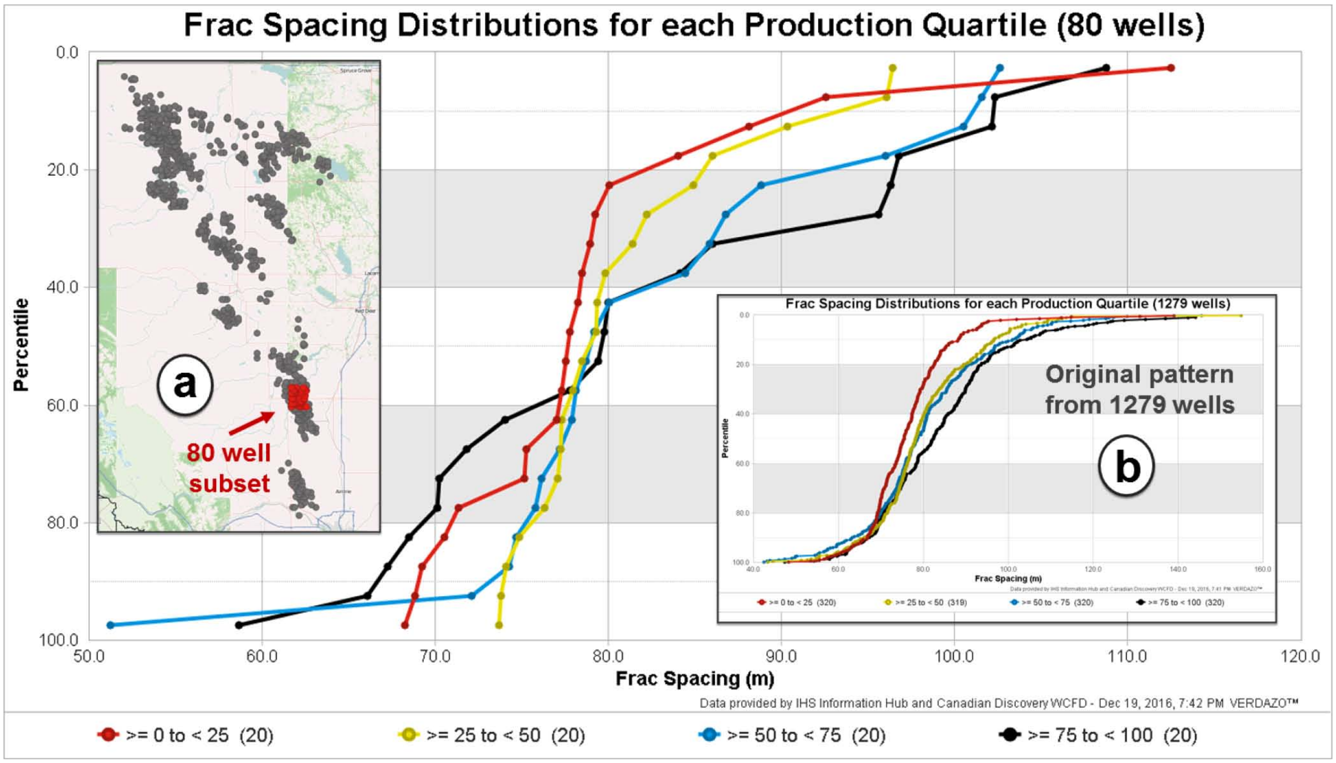
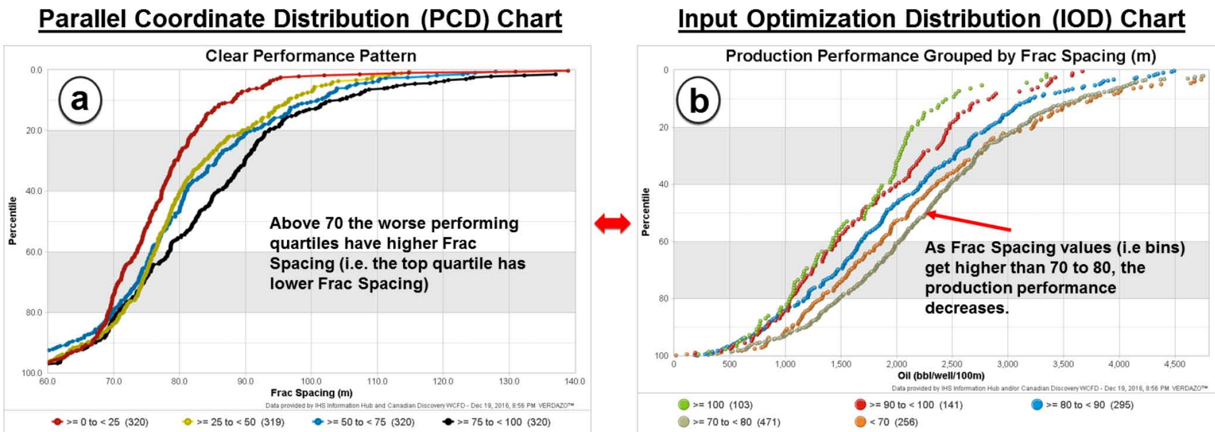


Figure 6—PCD presents a recognizable pattern with just 80 wells

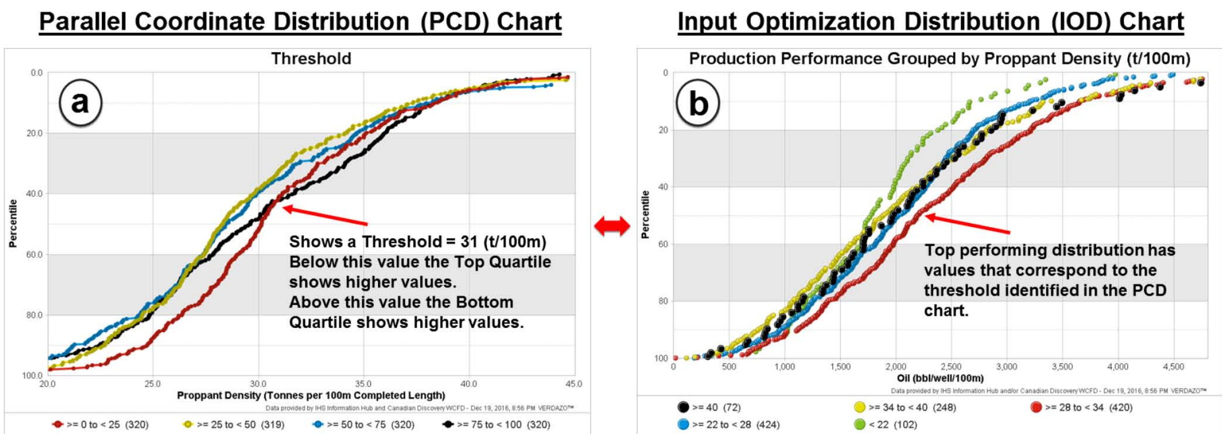
Input Optimization Distributions

Figure 8a is a PCD chart for Proppant Density (t/100 m), showing the production performance measure quartile colouring of 12-month Cumulative Oil (bbl/100 m completed length) for the Cardium Oil study area. It illustrates a threshold phenomenon where a Proppant Density higher than 31 t/100 m shows a crossover between the top quartile and the bottom quartile. This suggests that above the identifiable threshold, the worst performing wells have a higher proppant density and increased proppant begins to have a detrimental effect on production performance. This is supported by Figure 8b, an Input Optimization Distribution (IOD) chart that shows distributions of the production performance measure grouped into bins of Proppant Density (the input that is trying to be optimized). Bin sizes were selected to ensure that well count was statistically representative for each bin, that a reasonable number of bins were created to identify an optimal input parameter range and that the identified threshold value was represented by a specific bin. Figure 8b shows that distributions of performance improved as Proppant Density increased up to the 28-to-34 t/100 m bin. As proppant density increased above this threshold, the distributions resulted in lower production performance. Figure 8b provides supporting evidence of the threshold value that was identified in Figure 8a.



Note: colours on each chart represent different groupings

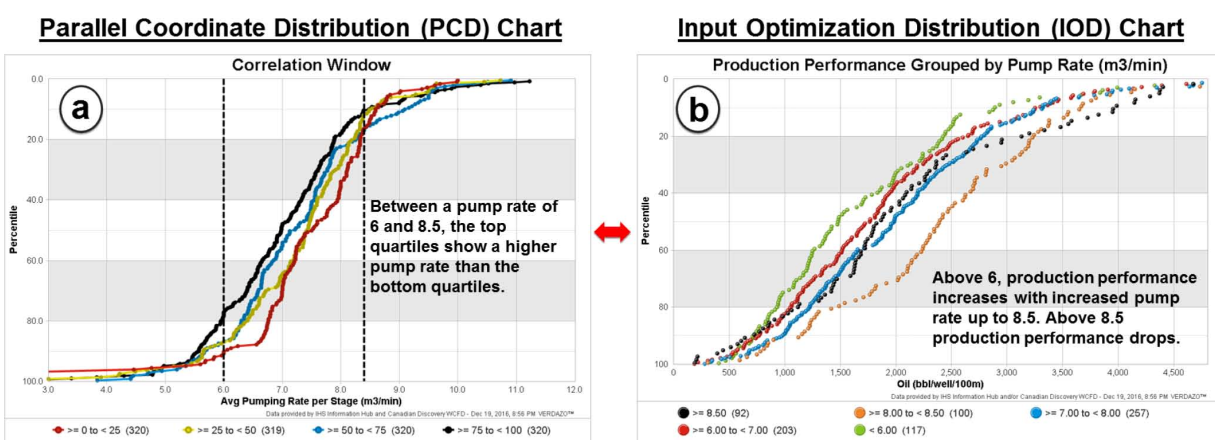
Figure 7—Clear Pattern: PCD provides bin value indicators that are supported by IOD



Note: colours on each chart represent different groupings

Figure 8—Threshold Pattern: PCD provides bin value indicators that are supported by IOD

The approach of using a PCD chart combined with the IOD chart yields reinforced data-driven visual insights on an input parameter's impact on production performance. The PCD patterns can indicate suitable binning values for IOD charts where behavioral shifts are expected to occur (i.e. thresholds). Figures 7 through 10 provide illustrations of how the IOD approach can be effectively applied to each of the PCD patterns identified in Figure 4.



Note: colours on each chart represent different groupings

Figure 9—Correlation Window: PCD provides bin value indicators that are supported by IOD

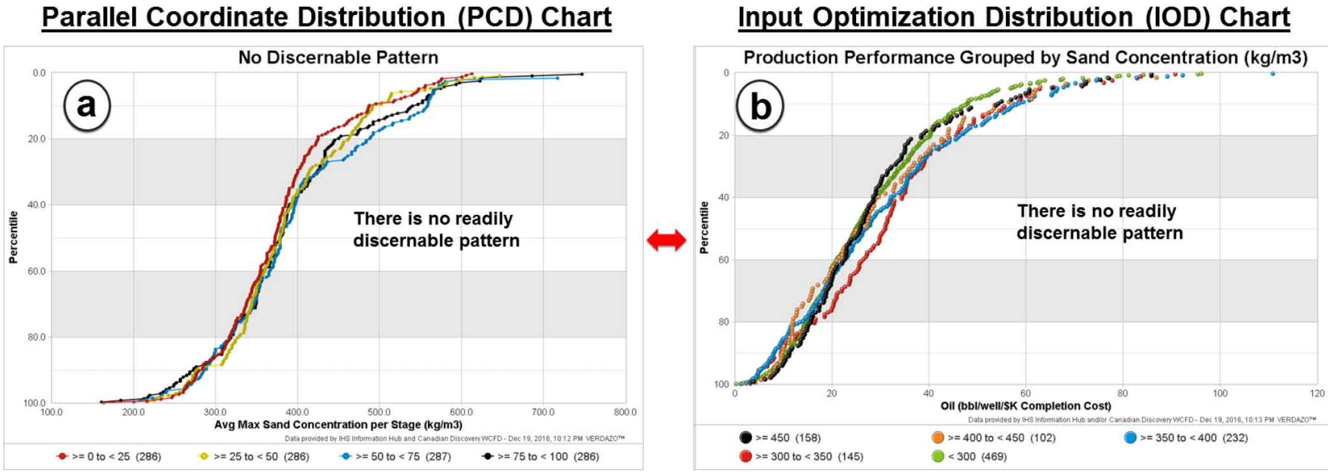


Figure 10—No Discernable Pattern: IOD supports PCD indications of no discernable pattern

Recent industry trends have seen dramatic increases in proppant loading with associated unprecedented production capability. The industry norm has evolved to rate restrict these wells, in some cases, for years. Pressure data can be used to IPR-correct the production values. Where pressure data is not available it is our recommendation to use a cumulative production performance measure. The longer the duration of the cumulative production period, the more the impact of rate restricting is mitigated in the results. It is for these reasons that peak and IP90 performance measures are not recommended for plays where rate restriction occurs.

Workflow Overview

1) Selection of Production Performance Measures

Common production performance measures include production rates (e.g. IP90, peak rate, first N-month average rate), cumulative production volumes at N months, and Estimated Ultimate Recoverable. These same performance measures can be dimensionally normalized to correct the production values for a particular input parameter (e.g. completed length). To develop and test the visual analysis concepts of this paper, we used:

a. 12-month Cumulative Production, referred to as Oil (bbl/well) or Gas (Mcf/well)
b. 12-month Cumulative Production/100 m completed length, referred to as Oil (bbl/well/100 m) or Gas (Mcf/well/100 m) or unit production performance measure

2) Analogue Well Selection

There are many criteria available for the selection of an analogue well dataset. While similarity is desirable in the dataset, it is important to have enough variability in the data to be able to visually identify patterns in production performance behavior. It was discovered that when variability in an input parameter had a P10:P90 ratio of 1.6 or lower, there was not enough variability to identify optimal values. The main criteria used in this paper for Analogue Well Selection include:

c. Producing Formation
d. Open/Cased
e. Fluid Type
f. Vintage (On Production Year)

3) Selection of Numerical Completion Design Input Parameters
The Input Parameters that were used to test the concepts presented in this paper included:

a. Completed Length (m)
b. Frac Spacing (m): (note that the number of stages was excluded as it is a function of completed length and frac spacing)
c. Proppant Density (t/100 m)
d. Fluid Pumped (m3/m completed length)
e. Avg Max Sand Concentration per Stage (kg/m3)
f. Avg Pumping Pressure per Stage (Mpa)
g. Avg Pumping Rate per Stage (m3/min)

Other parameters and analogue selection criteria that the user may consider include:

- Reservoir & Fluid Considerations
  - Pore pressure - the impact of varying overpressure (psi/ft variation)
  - Effective porosity
  - Impact of regions with known natural fracturing
  - Regional stress variations
- Completion & Wellbore Considerations
  - Casing size
  - Open vs cased hole
  - Proppant type
  - Frac fluid selection
  - Energized vs non-energized fluids
  - Well spacing
  - Lateral landing (interval selected)
  - Geo-steering vs non-geo-steering into a preferred landing zone

4) Parallel Coordinates Distributions: input parameter impact analysis
By assembling the PCD charts into a Matrix (Figure 5) the impact of each of the inputs can be quickly and easily compared, identifying specific patterns that warrant further analysis using IOD charts.

5) Evaluation of analogue fitness and subset selection
While reviewing the PCD chart matrix there are four important considerations that should be used to determine if the existing analogue selection is appropriate and if subsets should be considered:

a. Extreme ranges of input parameter values
b. Inexplicable or conflicting PCD Patterns (see Figure B-3 for an example)
c. Including additional analogue selection criteria (that were not used to establish the original dataset)
d. Verify results geospatially (with a map) to see if identifiable geological or operational subsets should be analyzed separately

6) Input Optimization Distributions: input optimization process

Create IOD charts for the PCD charts that demonstrate patterns of interest (see Figures 4a, 4b and 4c). Use the insights from the pattern(s) to identify:

a. bottom threshold bin value
b. bin size (to highlight a particular threshold value)
c. top threshold value
d. adjusting all of the above to get reasonable well counts in each bin and a visually manageable number of bins

Isolate specific bins of interest in an analogue subset, and use smaller bin values to draw more detailed conclusions.

Studies

In this paper, we test this methodology on three major Resource Plays (Figure 11) in the Western Canadian Sedimentary Basin:

a. Cardium Oil Play (Fig A-1)
b. Montney Dry Gas Play (Fig B-1)
c. Montney Liquids-Rich Gas Play (Fig C-1)

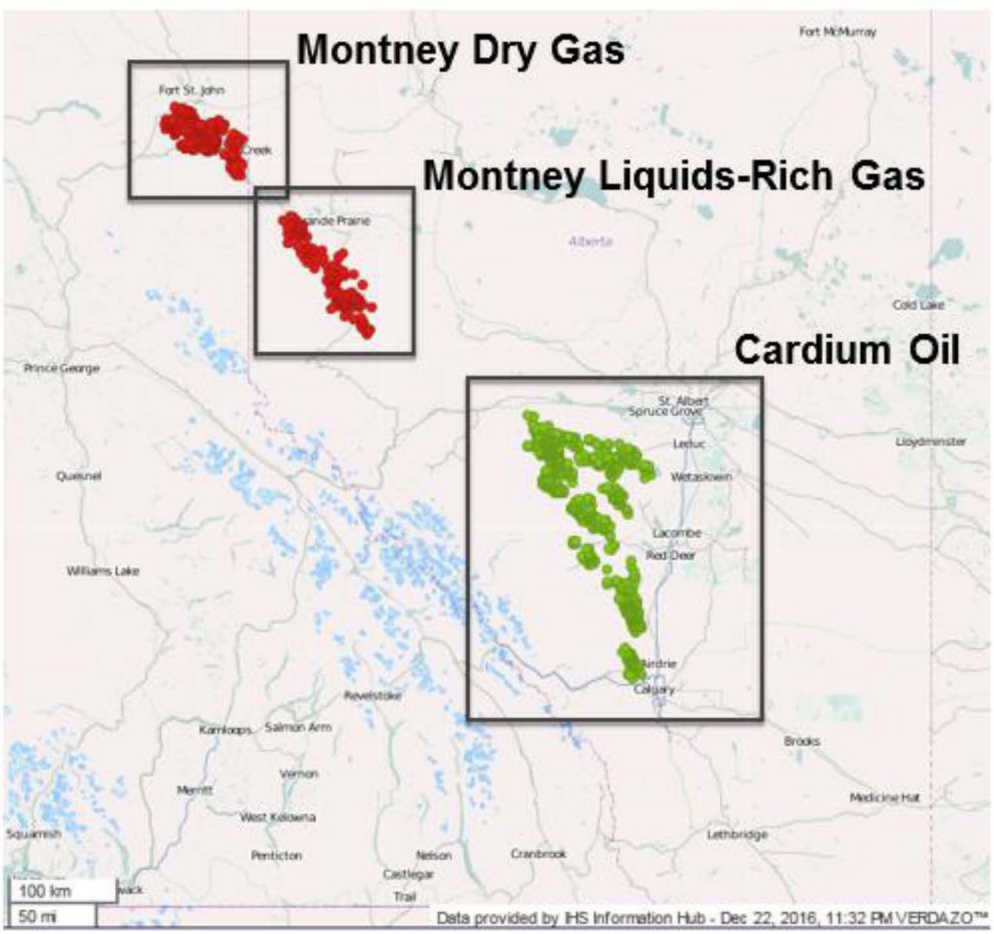


Figure 11—Three Study Areas

The purpose of these studies was to illustrate the concepts presented in this paper. A more rigorous and detailed analysis incorporating proprietary geological knowledge, analogue investigation and subset selection would be required to draw discernable conclusions for optimal completion parameter design.

In all study areas, Completed Length was by far the dominant driver of overall production (Figures A-2, B-2 and C-2). As such, any analysis of impact of other input parameters could be highly misleading. This necessitated correcting for Completed Length in any subsequent analyses by using the dimensionally normalized production performance measure of Production/100 m completed length, also referred to as unit production performance.

Using Production/100 m completed length, a PCD was created for each Resource Play to assess the impact that increasing completed length has on production performance (Figures A-3, B-3 and C-3).

Cardium Oil Play

The Cardium Oil Play dataset was used to create all the patterns shown in Figure 4 with the most notable impacts on Production/100 m being driven by Frac Spacing (Figure 7) and Pumping Rate (Figure 9). Thresholds were readily identified in Proppant Density (Figure 8) and Completed Length (Figure A-3), where values above 1400 m started to demonstrate a reduction in unit production performance (i.e. per 100 m).

Montney Dry Gas Play

In this dataset, increasing Completed Length demonstrated a reduction on production performance when Completed Length was below 2000 m. Above 2000 m, there was an increase in unit production performance as Completed Length increased. The unexpected or contradictory behavior at a particular threshold suggested that the analogue be scrutinized to see if there was a need to establish analogue subsets. It was determined that most of the longer wells also experienced much higher proppant density (t/100 m). By excluding the extremely high proppant density wells, the pattern of the chart becomes clearer (Figure B-3), and suggests that the unit production experiences a minor decrease as length increases, with the P50 of the top quartile being 8% below all other quartiles.

The most dramatic impact on unit production performance was Proppant Density (Figure B-4a). Unlike the Cardium where an upper threshold for optimal production performance was recognized, Figure B-4b suggests that the Montney Dry Gas Play has not yet reached an upper threshold.

Montney Liquids-Rich Gas Play

In this dataset, increasing Completed Length demonstrated a clear reduction in production performance (Figure C-3). Proppant Density demonstrated an upper threshold of approximately 105 (t/100 m) where a crossover between top and bottom quartiles occurs in Figure C-4a. This is further supported by Figure C-4b where the distribution of the greater-than-120 (t/100 m) bin demonstrates a reduction in unit production performance. Like the Cardium, the Montney Liquids-Rich Gas Play showed unit production performance improvements with increased Pumping Rate (Figure C5-a), with average pumping rate per stage greater than 9 (m3/min) demonstrating the best unit production performance.

Conclusions

1. The PCD and IOD visual analysis methodology is suitable for testing any input's impact on production performance with careful consideration of dimensionally normalizing both performance measures and inputs. Multiple performance measures should be used to maximize the number of patterns identified.
2. Given that demonstrable patterns were identifiable in all study areas, this methodology appears to be suitable for all types of plays where inputs have enough statistical variability to see measurable results.

3. This approach is effective at communicating nuances in the data, such as thresholds and correlation windows that are not easily identified using other techniques. Further, specific threshold values and correlation window ranges can be valuable inputs to other regression techniques or modelling efforts.
4. The approach is scalable, accommodating datasets containing greater than 1000 wells down to as few as 80 wells.
5. Inexplicable patterns are an effective way of identifying the need for analogue review and subset selection. It is essential that analogue subsets be explored to refine insights before drawing discernable conclusions.

Acknowledgements

The authors wish to acknowledge IHS Markit's Canadian Information Hub for production data used in the studies and Canadian Discovery Ltd.'s Well Completion and Frac Database for completion data used in the studies. The authors also wish to acknowledge the support of Verdazo Analytics Inc. and Rose and Associates.

Nomenclature

IPR = Inflow Performance Relationship

IP90 = the average daily production rate calculated using the first 2,160 hours on-stream

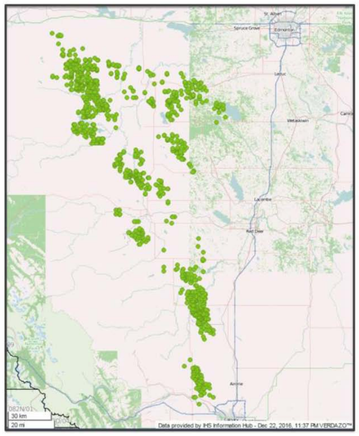
References

McLane, Mark, Gouveia, Jim, 2015, Validating Analog Production Type Curves for Resource Plays, Paper SPE 175527, presented at the SPE Liquids-Rich Basins Conference - North America, Midland, Texas, USA, 2–3 September

Wikipedia. 2016. Parallel Coordinates (24 October 2016 revision), https://en.wikipedia.org/wiki/Parallel_coordinates#/media/File:ParCorFisherIris.png (accessed Dec. 2016)

Appendix A – Cardium Study

Cardium Oil



1279 wells within the study area:

- Formation = Cardium Oil
- Primary Product = Oil
- Open Hole
- Base Fluid = Slickwater
- Horizontal
- Production Year >2010
- Frac information available
- > 12 months of production

Figure A-1—Cardium Oil Study Area

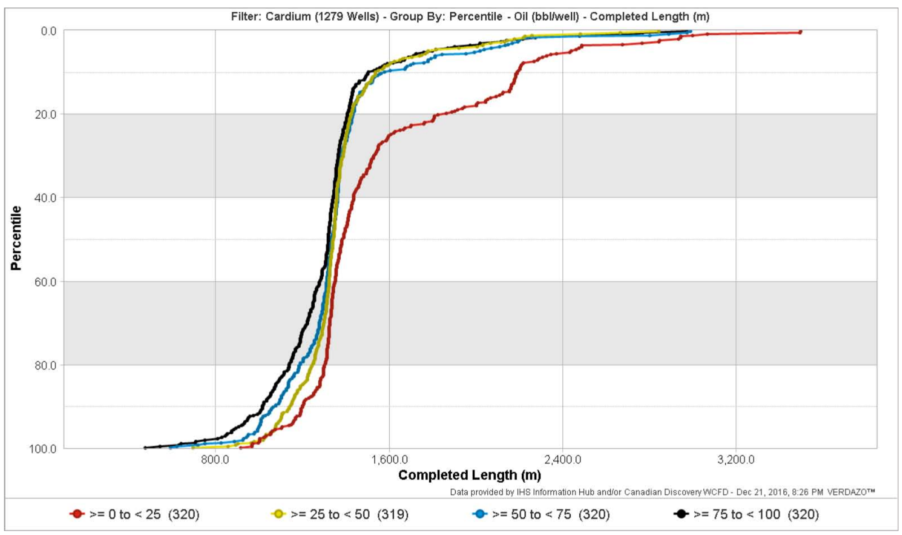


Figure A-2—PCD Chart of Completion Length Impact on Well Production

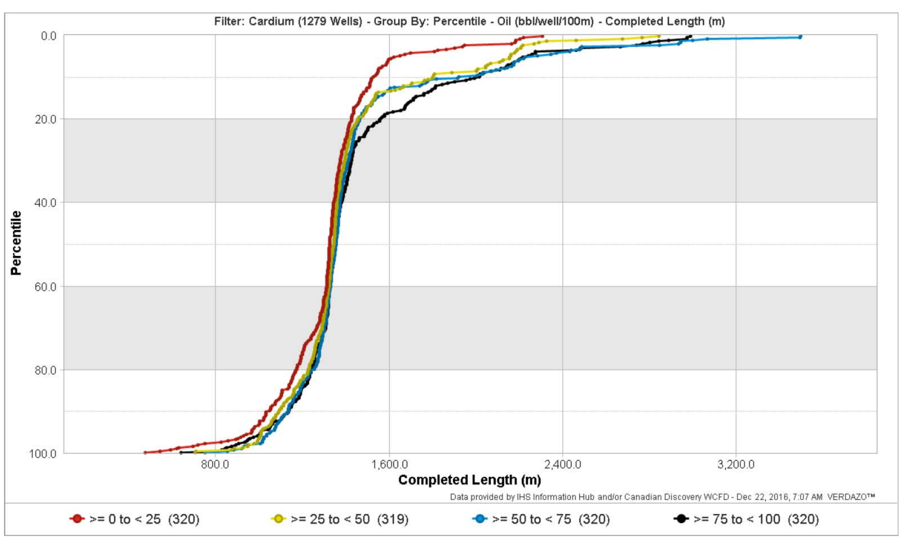
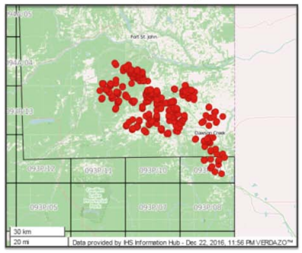


Figure A-3—PCD Chart of Completion Length Impact on Unit Production Performance

Appendix B – Montney Dry Gas Study

Montney Dry Gas



452 wells within the study area:
- Formation = Montney
- Primary Product = Gas
- Cased
- Base Fluid = Slickwater
- Horizontal
- Production Year >2009
- Frac information available
- > 12 months of production

Figure B-1—Montney Dry Gas Study Area

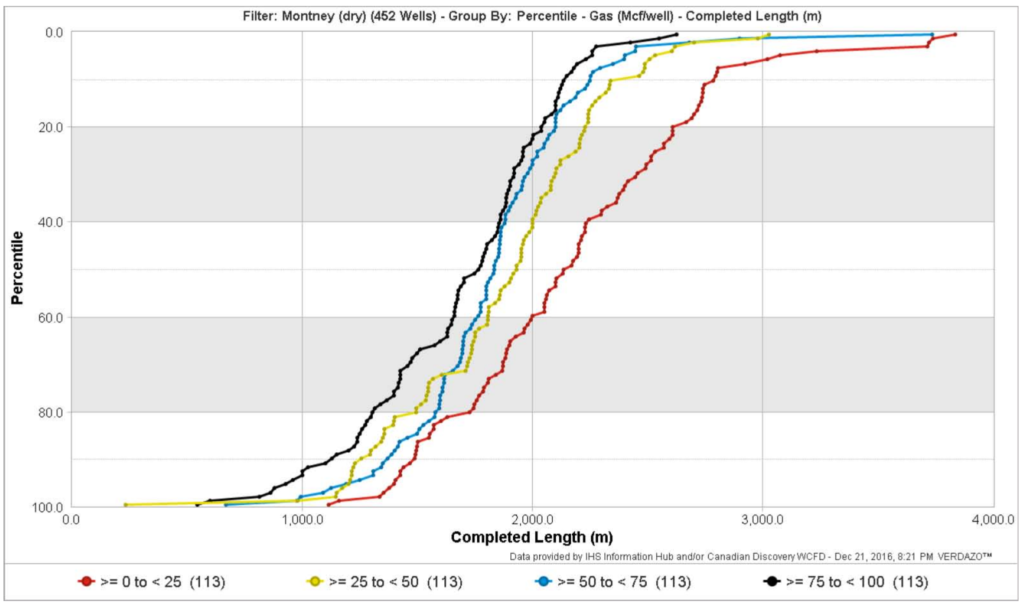
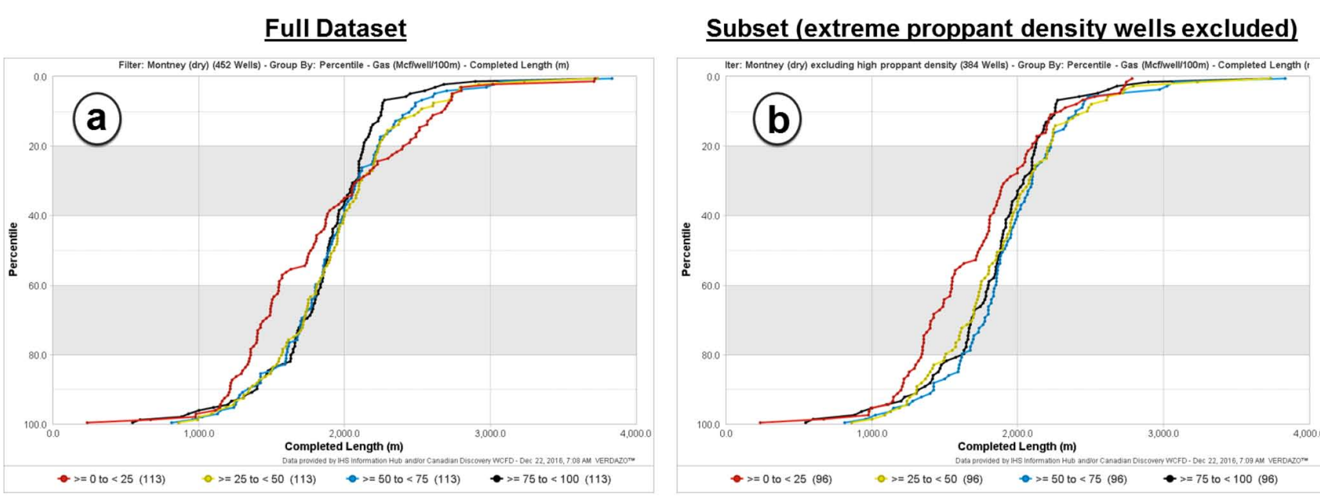
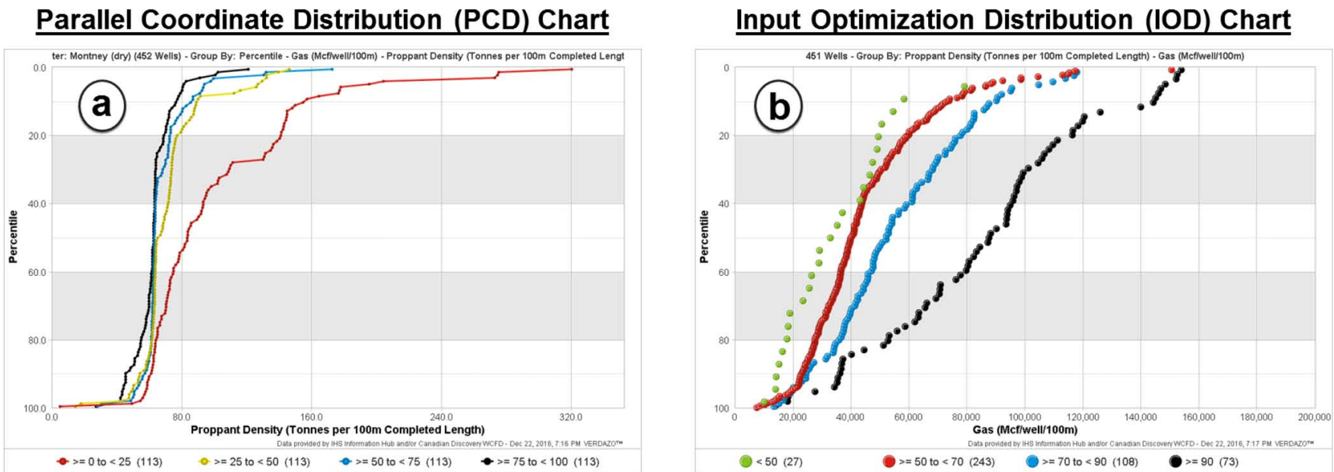


Figure B-2—PCD Chart of Completion Length Impact on Well Production



Note: colours on each chart represent different groupings

Figure B-3—PCD Unit Production Performance Pattern Change with Analogue Subset



Note: colours on each chart represent different groupings

Figure B-4—PCD and IOD Charts Showing No Upper Threshold on Proppant Density

Appendix C – Montney Liquids-Rich Gas Study

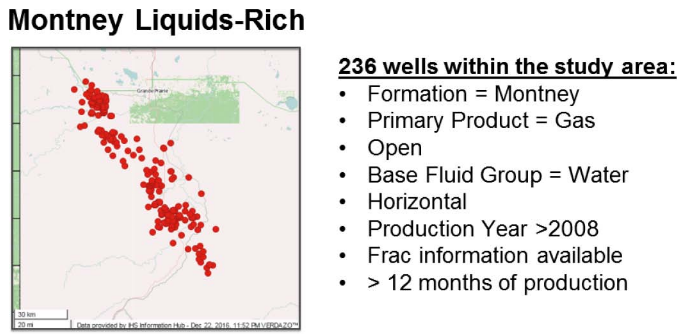


Figure C-1—Montney Liquid-Rich Gas Study Area

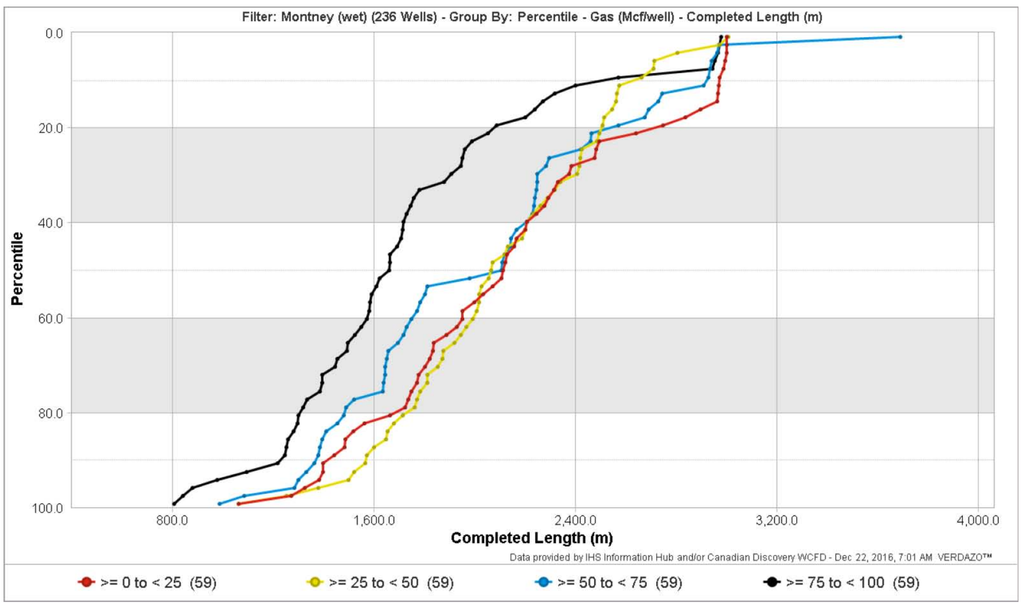


Figure C-2—PCD Chart of Completion Length Impact on Well Production

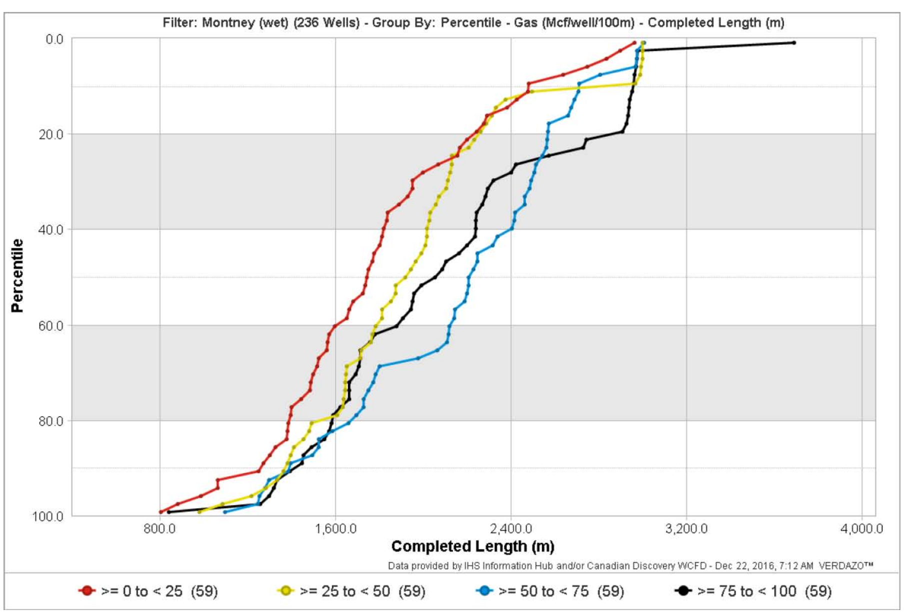
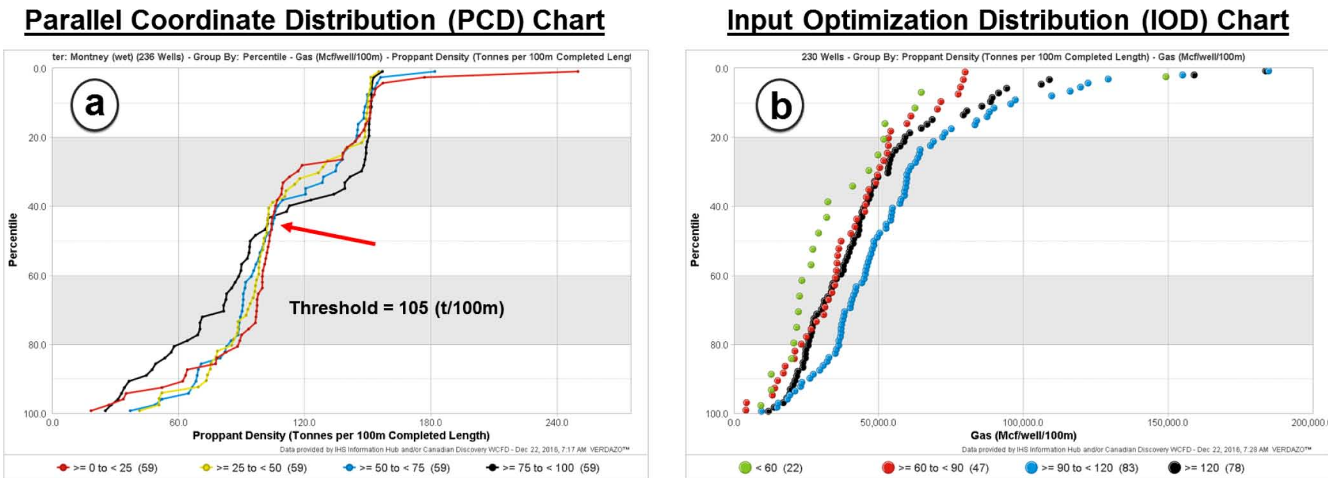
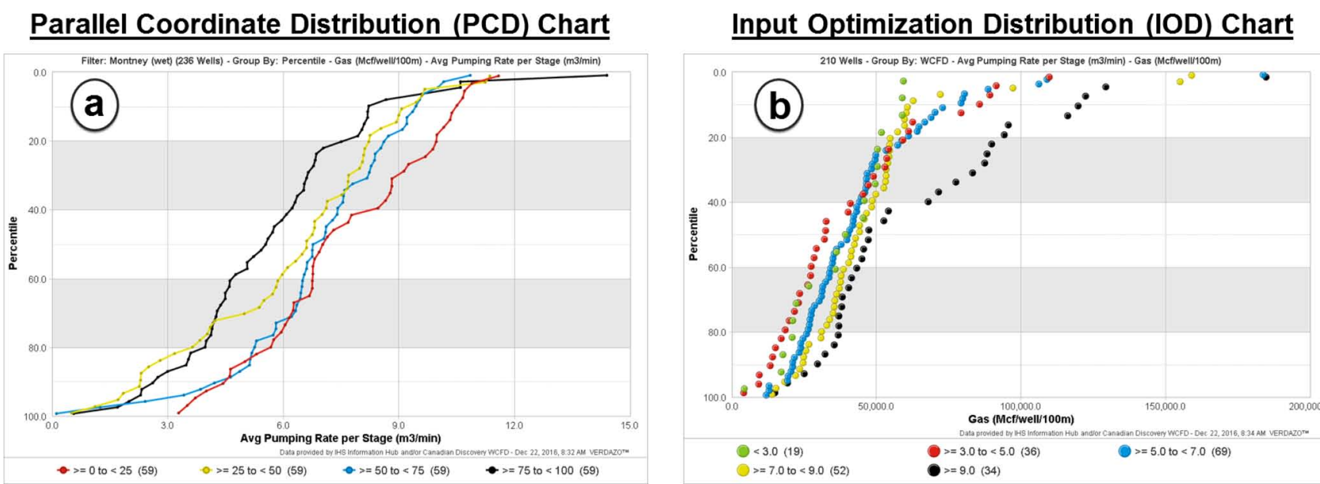


Figure C-3—PCD Chart of Completion Length Impact on Unit Production Performance



Note: colours on each chart represent different groupings

Figure C-4—PCD and IOD Charts Showing Upper Threshold on Proppant Density



Note: colours on each chart represent different groupings

Figure C-5—PCD and IOD Charts Showing No Upper Threshold on Pumping Rate