IMPROVED WELL COST ESTIMATION THROUGH UNCERTAINTY QUANTIFICATION

A Thesis

by

KYLE ONEIL PARADIS

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
In partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Samual F. Noynaert
Committee Members,	Jerome J. Schubert
	Charles P. Aubeny
	Duane A. McVay
Head of Department,	A. Daniel Hill

May 2016

Major Subject: Petroleum Engineering

Copyright 2016 Kyle Oneil Paradis

ABSTRACT

The study was initiated to develop more accurate well cost estimations for drilling and completion Authorization for Expenditures (AFE). Specifically, a privately funded company (Company A) is interested in analyzing historical drilling and completion AFEs to determine a more accurate and reliable estimation process to efficiently develop their core assets in the Montney play of northeast British Columbia.

In 2014, Company A and working interest affiliates underestimated well cost AFEs by 23% reducing the amount of available capital which could be used for further development. The uncertainty in these estimates was quantified at a sub-cost level which determined the locations of focus for strategic intervention and a workflow was created which utilized previous years' estimates, actual costs, and a probabilistic cost model to convert deterministic well cost estimates into probabilistic to better estimate well costs.

The uncertainty can be represented through a series of tornado charts showing the primary areas requiring improvement, based on an evaluation of 97 cost codes related to drilling and completion operations. The areas selected, are those with the largest ranges of uncertainty in terms of dollar value. For example, 11 cost codes were found to exceed ranges of $500k per well and these require immediate attention. The developed workflow examined each cost code through a probabilistic analysis determining the optimal procedure for adjusting future AFE estimates'. As a result, the workflow reduced well cost underestimation from 23% to only 5%. The improved well cost estimations provide the opportunity for a more accurate allocation of funds.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Samuel Noynaert, and my committee members, Dr. Duane McVay, Dr. Jerome Schubert, and Dr. Charles Aubeny for their guidance and support throughout the course of this research. Thanks also go to my friends, colleagues, and the Harold Vance Department of Petroleum Engineering faculty and staff for making my time at Texas A&M University a great experience.

I would also like to thank Ariel Brown, as well as my parents Paul and Dianne Paradis for their continued support throughout this process.

NOMENCLATURE

ΔA	Delta Actual
AFE	Authorization for Expenditure
AFE Bust	Exceeding AFE Estimate
CAPEX	Capital Expenditures
CDF	Cumulative Distribution Function
C_T	Total AFE Cost
C_i	Deterministically Estimated AFE Sub-Cost i
DB	Directional Bias
Det.	Deterministic
ED	Expected Disappointments
i	Index of AFE Sub-Cost
m	Slope
n	Number of Elements in Data Set
NPT	Non-Productive Time
P10	10^{th} Percentile
P30	30^{th} Percentile
P50	50^{th} Percentile
P70	70^{th} Percentile
P90	90^{th} Percentile
Prob.	Probabilistic

VBA	Visual Basic for Applications
WI	Working Interest
X_i	Historically Calculated Correction Factor

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

In 2014, an oil and gas operation company (Company A-small privately funded partner with multiple companies in both operational and non-operational roles) and its working interest affiliates underestimated well cost AFEs by 23%. As part of an effort to better allocate funds, Company A is interested in utilizing a probabilistic cost model which analyzes historical drilling and completion AFEs and determines more accurate and reliable estimations for efficient development of their core assets in the Montney play shale formation of northeast British Columbia (Fig. 1).

Fig. 1—Montney Shale Play (Nieto *et al*., 2013)

Companies and engineers tend to think that deterministic cost estimates are sufficient when estimating drilling and completion costs. This perception is even more common in regards to unconventional reservoirs due to the perceived similarity in the large quantity of wells required for development. Evidence however, proves this perception to be false, often due to overconfidence in their ability to estimate, hindering the economic efficiency of hydrocarbon development through less than optimal allocation of funds.

The economic success of oil and gas developments depends largely on the economic appraisal of multiple projects and, by association, the accuracy of cost estimates. These early cost estimates are often used to determine and compare rate of returns for multiple projects. As project funding is often limited, the most profitable projects must be determined. Failure to accurately estimate these costs can lead to the selection of less profitable or even worse unprofitable projects for development. In order to make more accurate estimates, one must sufficiently understand the uncertainty incorporated within these estimates, ultimately allowing for better decision making. In order to make more informed decisions, one must sufficiently understand the uncertainty incorporated within these estimates. Understanding and finally incorporating this uncertainty into the cost estimates is crucial to optimal field development.

A method was created to incorporate the uncertainty in well cost estimations in the work of Valdes *et al.* (2013), of which this study is a continuation. Valdes *et al*. created a five-step procedure for generating probabilistic well cost estimations from deterministic estimates provided by engineers. The procedure involved converting a deterministic cost model (Eq. 1), into a probabilistic cost model (Eq. 2), by using external historical correction factors (Eq. 3).

$$\mathrm{C_T} = \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{X_i} = \mathrm{C_1} + C_2 + C_3 + \cdots + C_\mathrm{n} \quad (1)$$

$$\mathrm{C_T} = \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{X_i C_i} = \mathrm{X_1 C_1} + \mathrm{X_2 C_2} + \mathrm{X_3 C_3} + \cdots + \mathrm{X_n C_n} \quad (2)$$

$$\mathrm{X_i} = \frac{\text{Actual well cost}}{\text{Estimated well cost}} \quad (3)$$

where C_T is the total AFE cost, C_i is a deterministically estimated AFE Sub-cost, and X_i is the historically calculated correction factor. This probabilistic cost model was developed for a single data set and often required cumbersome manual intervention in the form of distribution definition, truncation boundary settings, and data quantity selection and serves as a proof of concept. In order to have the capability of handling multiple data sets, methods for automated distribution fitting, truncation boundaries, and data quantity adaptations were presented (described later in further detail).

Valdes, *et al* built on prior work within the oil and gas industry which studied uncertainty and its effect on cost estimation. To better explore the concept of uncertainty, Capen (1976) journeyed across Society of Petroleum Engineering events in the United States, where he was able to test 1200+ people on their knowledge about uncertainty. The experiment required each person to put ranges around their answers to 10 random trivia questions. Table 1 shows that the average number of missed questions in comparison to the expected number of misses for the given requested range. He then concluded, among other results, that when people are uncertain about their answers, they most often do not know the degree of uncertainty in their answer.

Table 1—Summary of 95-percent ranges shows the average number of missed questions in comparison to the expected number of misses from Capen (1976) study.

SPE-AIME Section	Number of Usable Responses	Requested Range (percent)	Expected Number of Misses	Actual Number Average Misses	95-Percent Confidence Interval
Hobbs Petroleum	34	98	0.2	6.26	5.45 to 7.07
Oklahoma City	111	98	0.2	7.00	6.64 to 7.36
Los Angeles Basin (1)	28	90	1	5.96	5.24 to 6.68
San Francisco	61	90	1	6.41	5.89 to 6.93
Oxnard	26	90	1	7.38	6.64 to 8.12
Long Beach (1)	28	90	1	6.04	5.20 to 6.88
New York	29	90	1	6.52	5.76 to 7.28
Bridgeport/Charleston (1)	16	90	1	7.63	6.89 to 8.37
Anchorage	63	90	1	6.54	6.00 to 7.08
Bartlesville	44	90	1	6.30	5.61 to 6.99
Lafayette	79	90	1	6.51	6.03 to 6.99
Shreveport	41	90	1	6.83	6.18 to 7.48
Vernal	13	80	2	7.23	6.30 to 8.16
Denver	129	80	2	6.46	6.12 to 6.80
Cody	42	80	2	7.31	6.74 to 7.88
Columbus	27	50	5	6.96	6.47 to 7.45
Lansing	30	50	5	6.83	6.16 to 7.50
Chicago	41	50	5	6.54	5.97 to 7.11
Tulsa	53	50	5	6.79	6.33 to 7.25
Los Angeles Basin (2)	27	30	7	7.00	6.26 to 7.74
Long Beach (2)	28	30	7	7.39	6.80 to 7.98
Bridgeport/Charleston (2)	15	30	7	7.82	6.97 to 8.67

Since the upstream oil and gas industry is a capital intensive business, good business practices dictate the use of accurate estimations of the costs of projects. When uncertainty is evaluated, the industry is often asked to express their estimates in ranges of probability (P10, P50, and P90). However as Capen notes, “Having no good quantitative idea of uncertainty, there is an almost universal tendency for people to understate it. Thus, they overestimate the precision of their own knowledge and contribute to decisions that later become subject to unwelcome surprises.” As petroleum industry profit margins continue to shrink (due to lower crude oil prices and greater difficulty of extracting resources), the importance of the precision in these ranges to eliminate "unwelcome surprises," continues to increase. These "unwelcome surprises" are often the difference between profits and losses.

Brashear *et al*. (2001) noted that the average return on investment amongst oil and gas companies in the 1990s was approximately 7%. This is particularly interesting because the average hurdle rate for oil and gas companies during this time period was at least 15%. Building on this observation, McVay and Dossary (2014) identified chronic overconfidence and optimism as the problems in project evaluation. In this study, optimism is defined as, "the tendency to ignore or not consider possible negative outcomes, or the tendency to give greater weight to possible positive outcomes than possible negative outcomes," while chronic overconfidence is defined as the, "underestimation of uncertainty." For their study, the authors found the fundamental issue to be overconfidence and its subsequent negative effects on project evaluation outcomes.

McVay and Dossary (2014) explained how overconfidence and optimism can work together. As we are optimistic and fail to see the downsides of the project development, we become overconfident. Our natural optimism makes it necessary to constantly examine our estimates in comparison to the actual values to access our uncertainty and ultimately reduce chronic overconfidence. When engineers do this, we can combat our biases and reduce the number of overestimations and expected disappointments (ED). Ridding projects of ED will vastly improve industry performance allowing for the more frequent identification of truly superior projects.

As industry projects continue to expand in both scope and cost, estimates based on probabilistic methods are increasingly preferred over the previous standard deterministic estimates due to their more reliable nature. With this in mind, Williamson *et al*. (2006), provided a five-step Monte Carlo technique work flow as well as some of the commonly encountered pitfalls in its application to the estimation of time on single well costs. The steps are as follows:

1. Defining the model
2. Data gathering
3. Defining input distributions
4. Sampling input distributions
5. Interpreting and using the results

Using a probabilistic model allows an engineer to rely on the normally more accurate mean estimate of a distribution instead of the typically used "Best Estimate," which is most often an underestimation of the actual cost. This is particularly helpful as AFE actual costs are most often skewed to the right, presenting a log-normal distribution due to factors previously discussed (Fig. 2).

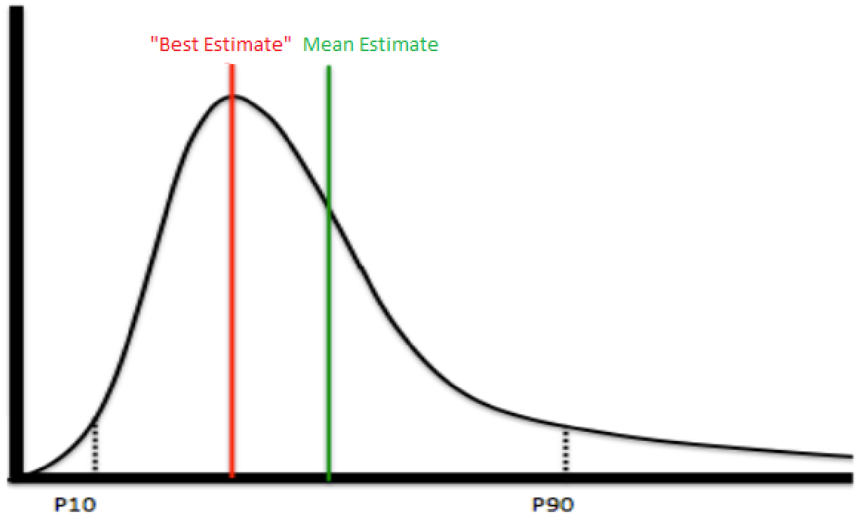

Fig. 2—"Best Estimate" vs. Mean Estimate

Amani and Rostami (2007) supported the probabilistic estimation through Monte Carlo simulation methodology by showing the capability of utilizing different probabilities of occurrence depending on location specific uncertainty. For example, they claimed the procedure could be standardized by using P50 (median) values as the base case for decision making. However, if a problem was to be expected (either due to a new area or new contractor) they recommend using a P75 value to account for the more likely encountered problems. This however, is not the optimal approach as using a consistent P75 for expected problems is almost equivalent to just using an additional contingency value in an AFE.

Not only is it evident that probabilistic models utilizing Monte Carlo simulation, when used appropriately, can increase the accuracy of well cost estimations, but, Hariharan *et al.* (2006) showed they can also be applied in respect to testing the economic impact of new technologies. Hariharan states that when a probabilistic framework is utilized, one can more accurately determine the range of potential economic impacts of different cost-saving technologies and procedures, allowing for a better understanding of the associated risks and rewards. The ability to easily test the economic impact on a probabilistic basis by showing the range of possible scenarios (i.e., does the technology work as expected, better than expected or worse than expected?) will greatly aid in the technological advances of a change-resisting drilling industry. Hariharan *et al*. also presented a survey to the SPE Drilling Technical Interest Group in which 57 members participated. The survey results indicate that 91% of the respondents believe that there is value in doing probabilistic analysis (Fig. 3); however only 54% of the respondents say that they are utilizing a probabilistic method now (Fig. 4).

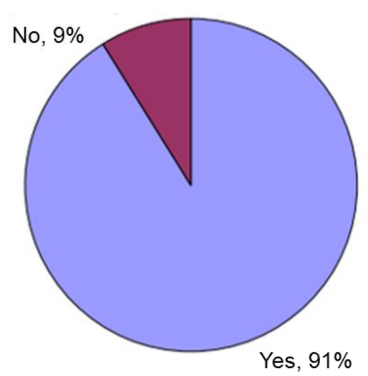

Fig. 3—Is there value in using probabilistics?

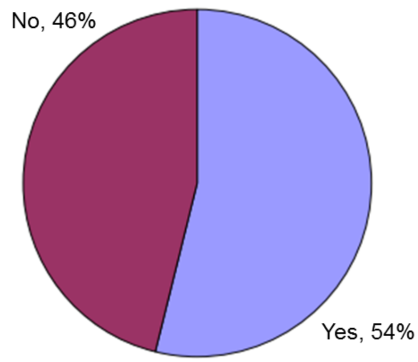

Fig. 4—Do you use probabilistics now?

Survey respondent comments indicating why probabilistic analysis was not able to be implemented focused on management. A few examples are shown below:

- "Results not understood by management, partners"
- "Low acceptance by engineers and managers"
- "Managements old habits"

These survey results demonstrate that engineers are struggling to explain their results in terms that allow management to better see the benefits.

Noticing the increasing occurrence of Monte Carlo-generated probabilistic well estimates, Adams *et al.* (2015) decided to revisit and improve the estimation models of Kitchel *et al.* (1997), Peterson *et al.* (2005), and Williamson *et al.* (2006). The primary improvements included advocating the importance of not blindly asserting a probability distribution function as well as the necessity of validating the probabilistic model against historical data for calibration.

Adams *et al.* accomplished this by plotting the cumulative density function data onto probability scales (Fig. 5). This type of plot compares the actual data to the modeled best fit distribution providing visual representation of distribution calibration. This is more accurate than using the chi-squared fitting program as it becomes easier to check for trends and outliers.

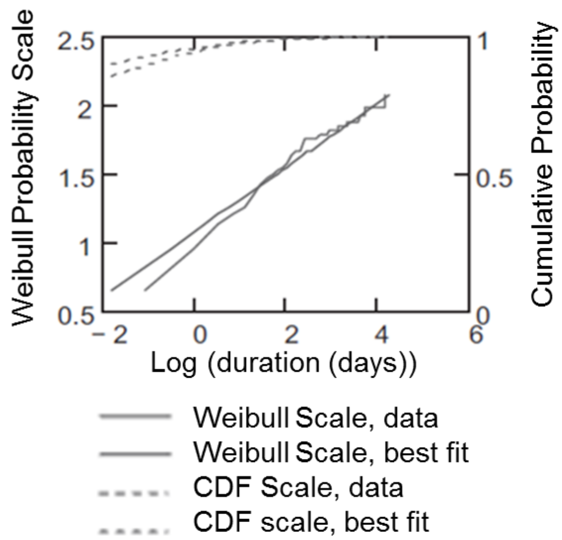

Fig. 5—Cumulative density function plotted on probability scale.

Gathering, examining, and reporting AFE data can be incredibly time consuming, sometimes requiring weeks of valuable time from engineers and support staff. Knowing this, Shilling and Lowe (1990) began the process of automating drilling AFE cost estimating and tracking. By tracking actual field costs with corporate accounting systems, companies can confirm that estimated costs match actual costs from the field allowing for a more accurate representation of spending. On top of realizing the importance of the time required to monitor AFE progress, the potential of obtaining AFE knowledge on both a total well and detail level basis in a real-time scenario was also recognized. This knowledge could potentially allow the operator to predict expenditures exceeding AFE estimates before they happen. These two points must be considered when improving a well cost estimation process as the more complex the model becomes, the more time and training is required to monitor it. This means procedures and models should be automated as much as possible to increase overhead efficiency.

Writing AFEs is a common task among engineers in the petroleum industry. As described by Peterson *et al*. (1993), this task typically "consists of artfully incorporating offset well data, engineering calculations, projections regarding operational improvements, and judgements about suitable contingencies." In an effort to optimize allocations of funds, AFE estimates of finished projects are generally compared to their actual value. This comparison takes place partly to help the estimating process become more accurate, but mostly to ensure the proper allocation of funds. It is in the best interest of every company to have the accurate amount of money set aside for each investment. Not only is it in the company's best interest internally, but it also supports compliance with the Sarbanes-Oxley Act Section 103. This act requires all publically traded companies to "provide reasonable assurance that transactions are recorded as necessary to permit preparation of financial statements in accordance with generally accepted accounting principles, and that receipts and expenditures of the issuer are being made only in accordance with authorizations of management and directors of the issuer."

Accurate estimates have large repercussions on company earnings. Too little cash apportionment could potentially cause cash flow problems, and too much means the loss of an opportunity to put the money to work elsewhere. In order to mitigate this issue

Peterson *et al.* (1993) proposed an AFE-generating model incorporating risk analysis and Monte Carlo simulation. The idea behind this model was to use historical drilling time data to generate probabilistic drilling AFEs to be compared to deterministic estimates and most importantly *actual costs* for two complex offshore well AFEs. The case study determined that the AFE estimates created using Monte-Carlo simulation more accurately predicted the number of well days than the deterministically-generated AFEs. In addition to being more accurate, the Monte Carlo simulation-generated AFEs also provide information on the contribution of problem versus problem-free days to the total days required.

1.1. Study Objectives

The objectives of study are as follows:

- Use probabilistic models to generate more accurate and reliable well cost AFE estimates for better allocation of funds. This method is presented as AFE correction factors and will be validated using previous years' AFE estimated and actual costs.
- Identify the accuracy of historical estimates submitted by Company A and working interest affiliates.
- Quantify the uncertainty in the estimation process to a sub-cost level to determine the locations of focus for strategic intervention. The locations of focus will be determined as the sub-costs with the most associated uncertainty and highest percentage of AFE cost. This allows Company A to obtain the optimal reward for increased accuracy if the ability to focus on all sub-costs does not exist.

1.2. Data

The AFE data used for this analysis was provided directly from Company A and was compiled through their AFE software. The data contained estimated and actual costs by sub-cost for 174 drilling, completion, and facilities AFEs from the years 2008-2014 for Company A and the following working interest affiliates:

- Company A: Privately funded organization focused on E&P of North American unconventional resources.
- Company B: No company information was provided.
- Company C: Public natural gas producer operating in Northeast British Columbia.
- Company D: Subsidiary of national oil company.
- Company E: Subsidiary of national oil company.
- Company F: Public international oil and gas company.

Of the 174 original AFEs, only 109 were used in the analysis due to incomplete data. The incomplete data was in the form of AFEs that were put together for planned projects that were never started.

The operator generated the AFE's using a deterministic workflow. Engineers estimated the cost of each AFE by estimating the costs of each subcategory and summing them up. A contingency cost was deterministically estimated in an attempt to account for uncertainty. As the projects were being completed, invoices were turned in for each cost and entered in for each subcategory. This workflow presents two challenges associated with data quality.

The first challenge occurs within the AFE reporting process. Invoices are often received by someone who did not make the AFE estimate, which sometimes lead to miscoding of the invoices. Common discrepancies associated with the miscoding issue are invoices being coded to the wrong cost code. This results in estimated costs with no actual costs and vice versa. For example, cost codes 415 and 615 both have the description Mud/Chemicals. Invoices which fall into this category are often miscoded into the wrong Mud/Chemicals cost code, resulting in an estimate with no actual cost for 415 (Mud/Chemicals) and an actual cost with no estimate for 615 (Mud/Chemicals). To mitigate this discrepancy, cost codes with similar descriptions were combined. Contingency cost codes were not included in the model as the probabilistic cost model is intended to replace the use of contingency.

The second challenge was that the data was presented as one Excel file consisting of a separate sheet for each of the 174 AFEs. It was necessary to write a Microsoft® Visual Basic® for Applications (VBA) code to retrieve necessary information from each separate sheet and compile them into a master database. This process should be adapted to the accounting system to present a fully automated process.

1.3. Company Base Line Statistics

It is necessary to first identify the accuracy of the AFE estimates by Company A and working interest affiliates for completed projects to determine base line statistics. This is accomplished by first looking at the total correction factor (Eq. 4) for each company through all years of data (Fig. 6). Summations of costs and estimates were utilized in place of averages to account for instances of miscoding. This was also utilized to mitigate the effect of project termination represented by estimates with no actual costs or vice versa. Total correction factors greater than 1 represent cost under estimations, while total correction factors less than 1 represent over estimations.

$$X_T = \frac{\text{Sum Actual Costs}}{\text{Sum Estimates}} \qquad (4)$$

Analysis of the data showed that all companies, with the exception of Company C, required total AFE correction factors greater than 1 meaning they underestimated their well cost AFEs (Fig. 6). Although overestimating can seem like an ideal way to prepare for cost overruns, Company C's conservative estimates effectively took away $1.75 million dollars of which could be allocated to other projects (Fig. 7). This is because for most companies, once an AFE has been approved, funds are allocated to that project until all invoices are in and accounted for. This means that while additional funds may eventually be returned to the available capital pool, the time these additional funds spend unavailable represents lost opportunities to appropriately utilize this capital.

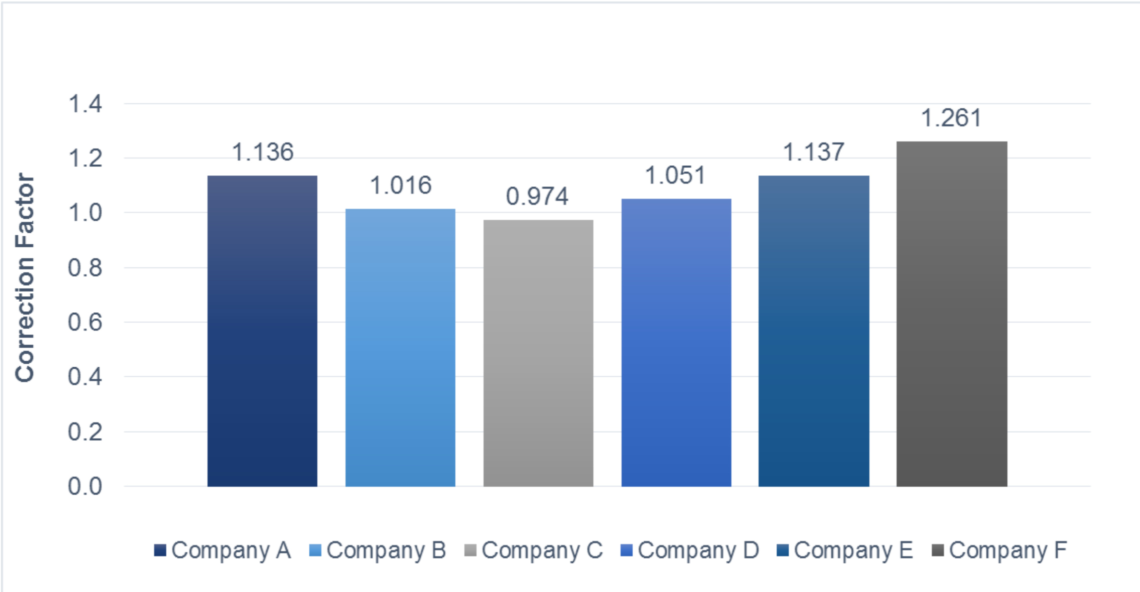

Fig. 6—Total correction factor for each company.

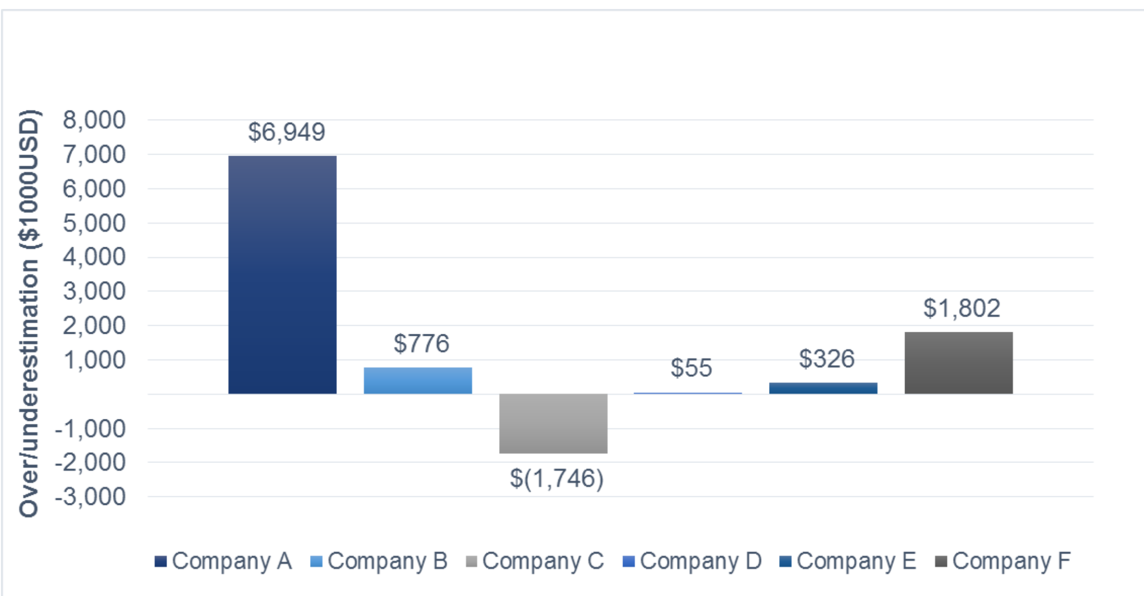

Fig. 7—Magnitude of over or underestimating for each company in dataset.

After the companies are analyzed over all the years, each company's ability to estimate costs must be investigated for each year separately (Figs. 8-12) to look for trends of improvement or retrogression. Company A's estimation data shows an overcorrection in AFE estimating (switch from overestimating costs to underestimating) between 2011 and 2012. Company B represents an improving trend. Company C-F represent no noticeable trends. While some were better than others, no company successfully considered uncertainty through the use of deterministic contingencies and all companies studied are capable of improvement with the inclusion of probabilistic cost estimating practices.

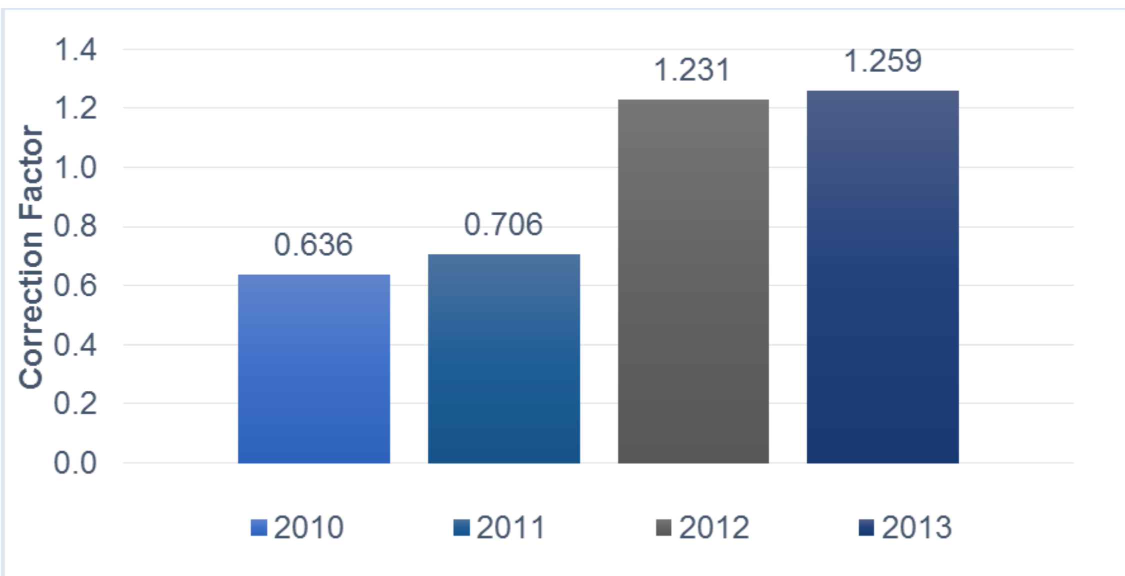

Fig. 8—Total correction factor for Company A.

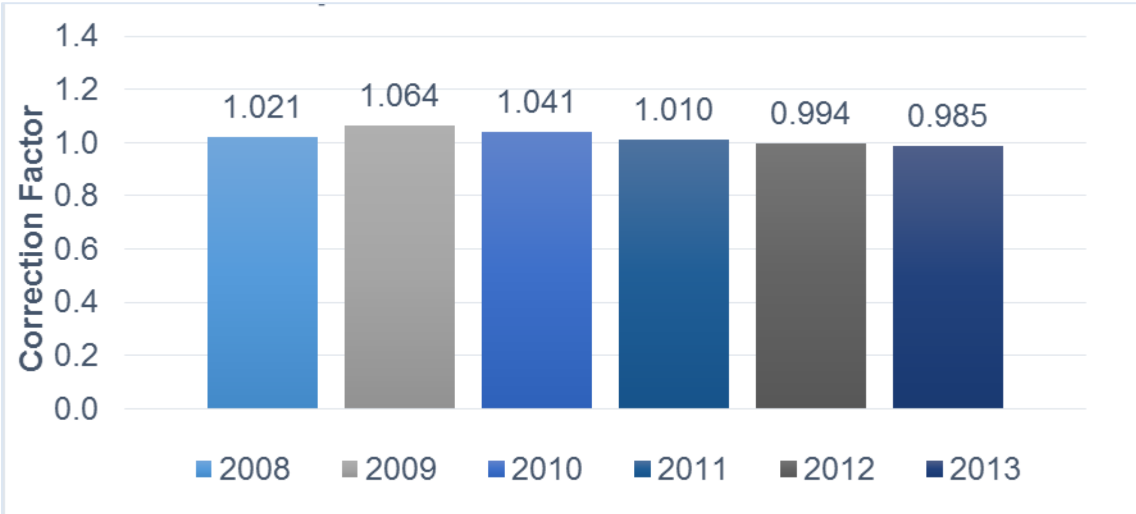

Fig. 9—Total correction factor for Company B

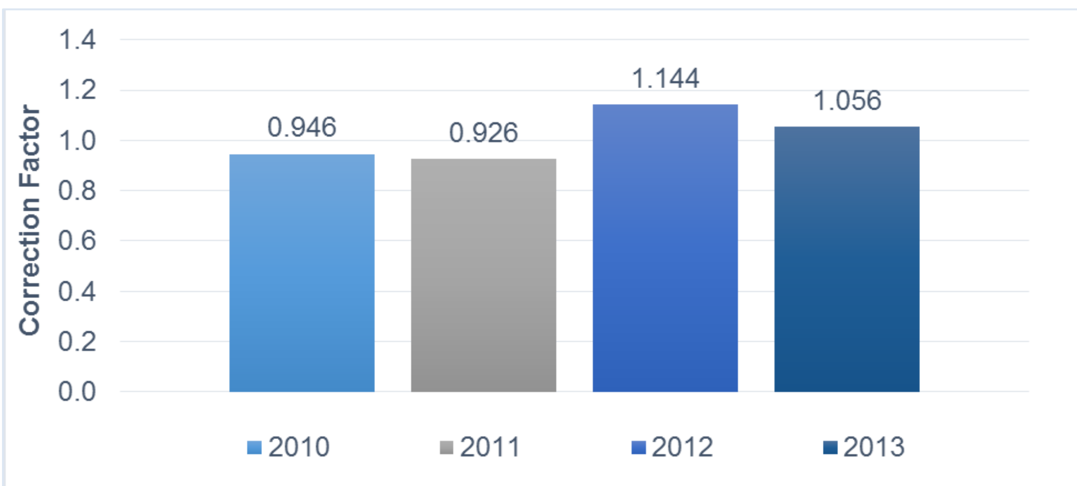

Fig. 10—Total correction factor for Company C.

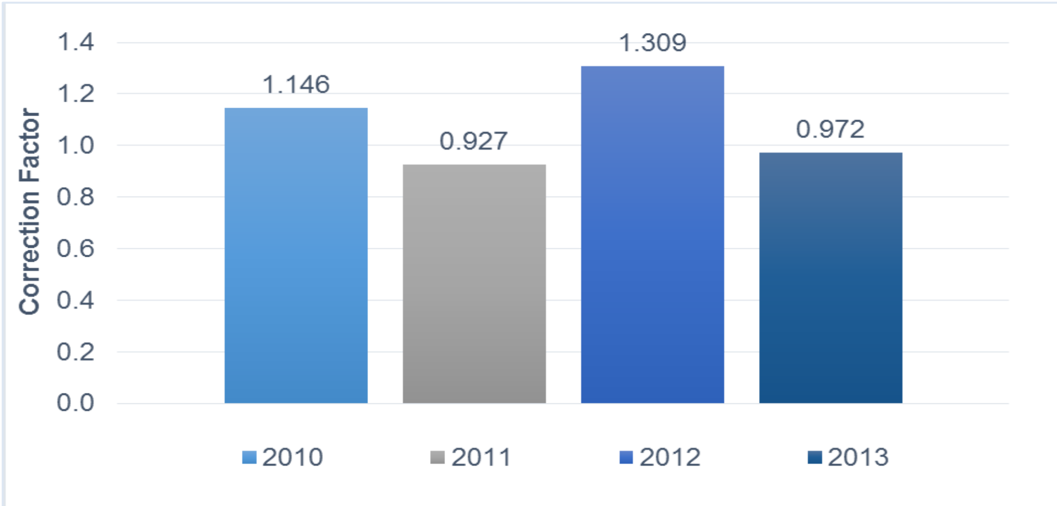

Fig. 11—Total correction factor for Company E.

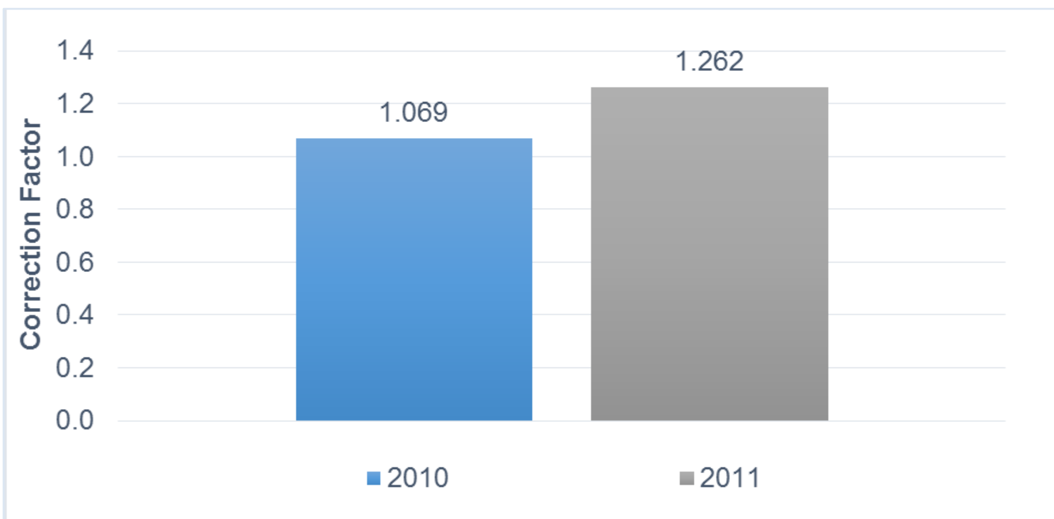

Fig. 12—Total correction factor for Company F.

2. UNCERTAINTY QUANTIFICATION

2.1. Methodology Overview

Since it has been determined by Valdes *et al.* (2013) that AFEs could be improved with the addition of probabilistic, instead of deterministic, cost estimation, uncertainty quantification can be implemented to determine primary areas of focus for optimized improvement. The uncertainty quantification is a three-step procedure:

1. Determine distributions of correction factors by cost code.
2. Use the distributions to calculate P10 and P90 estimates.
3. Create tornado plots for visual representation of the uncertainty.

The areas of focus chosen are those with the most potential to obtain the greatest value for the company improvement. This potential is found by finding the areas which have the largest range in dollar value uncertainty.

2.2. Step #1: Determine Distributions of Correction Factors by Cost Code

Using the Palisade Corporation's @RISK® commercial software, distributions of the correction factors were calculated for each cost code with the distribution fit function. For example, the distribution for cost code 406, drilling day work, can be seen in Fig. 13 as a right skewed distribution with a mode, median, and mean equal to 1.09, 1.18, and 1.31, respectively.

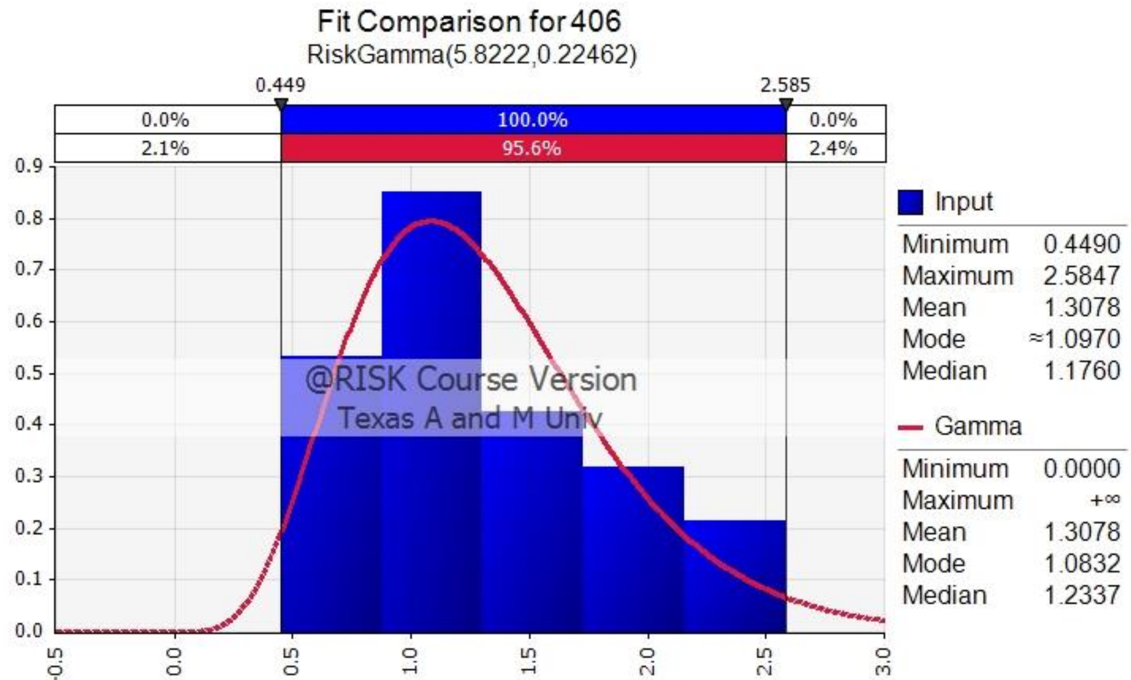

Fig. 13—Example of distribution of correction factors for a particular cost code (code 406 -drilling day work).

2.3. Step #2: Determine P10 and P90 Estimations

With the distributions of the correction factors determined, the deterministic estimates were converted into distributions by multiplying the deterministic estimates by the correction factor distributions. P10 and P90 estimations were then determined utilizing the *Riskpercentile* function, which returns the requested percentile of a simulated distribution, creating a range of estimates that represents the uncertainty.

2.4. Step #3: Create Tornado Plots and Determine Areas of Focus

Once the ranges of estimates representing uncertainty have been determined, these results can be modeled through the use of tornado plots. The values of the tornado plots were calculated in Eq. 5 and 6, respectively, and labeled Delta Actual (ΔA).

$$\Delta A_{P10} = \mathrm{P}_{10}\text{Estimate} - \text{Average Estimate} \quad (5)$$

$$\Delta A_{P90} = \mathrm{P}_{90}\text{Estimate} - \text{Average Estimate} \quad (6)$$

The difference between ΔA_{P90} and ΔA_{P10} is the associated uncertainty within each cost code estimation. Six tornado plots were created representing the highest amount of capital expenditure (CAPEX) uncertainty to lowest amount of CAPEX uncertainty. The intervals are as follows.

1. Uncertainty Quantification CAPEX>$500K (Fig. 14)
2. Uncertainty Quantification $500K>CAPEX>$250K (Fig. 15)
3. Uncertainty Quantification $250K>CAPEX>$1200K (Fig. 16)
4. Uncertainty Quantification $120K>CAPEX>$60K (Fig. 17)
5. Uncertainty Quantification $60K>CAPEX>$30K (Fig. 18)
6. Uncertainty Quantification $30K>CAPEX (Fig. 19)

2.5. Discussion

Sub-costs of importance were determined based on highest range of uncertainty. These areas are clearly highlighted in Fig. 14 and include the cost code descriptions: Casing-Surf/Int Accessories, Line pipe/Flowlines, and Drilling Fluids.

After further inspection into Casing-Surf/Int Accessories (426), it was determined that the uncertainty is a result of miscoding. Often casing string actual costs (Casing-Surface 423) were coded as accessories, resulting in estimates for Casing-Surface (423) with no coded actual costs and actual costs for Casing-Surf/Int Accessories (426) being unnaturally large. For example, Casing-Surf/Int Accessories (426) has been underestimated by 66% from 2008-2014 while Casing-Surface (423) was overestimated by 21% during the same period. This uncertainty can best be mitigated through deletion or combination of multiple cost codes to reduce the field reporting confusion. If this is done, care must be taken to do so in accordance with accounting practices, especially in regards to determining tangible and intangible costs. After combining Casing-Surf/Int Accessories and Casing-Surface (426 and 423, respectively), the cost codes only represented a 50% underestimation. The miscoding of Casing-Surf/Int Accessories resulted in approximately $2.2 million worth of uncertainty in AFE estimations. Once input errors or miscoding has been eliminated, the uncertainty of other associated cost codes can be quantified allowing for better well cost estimations.

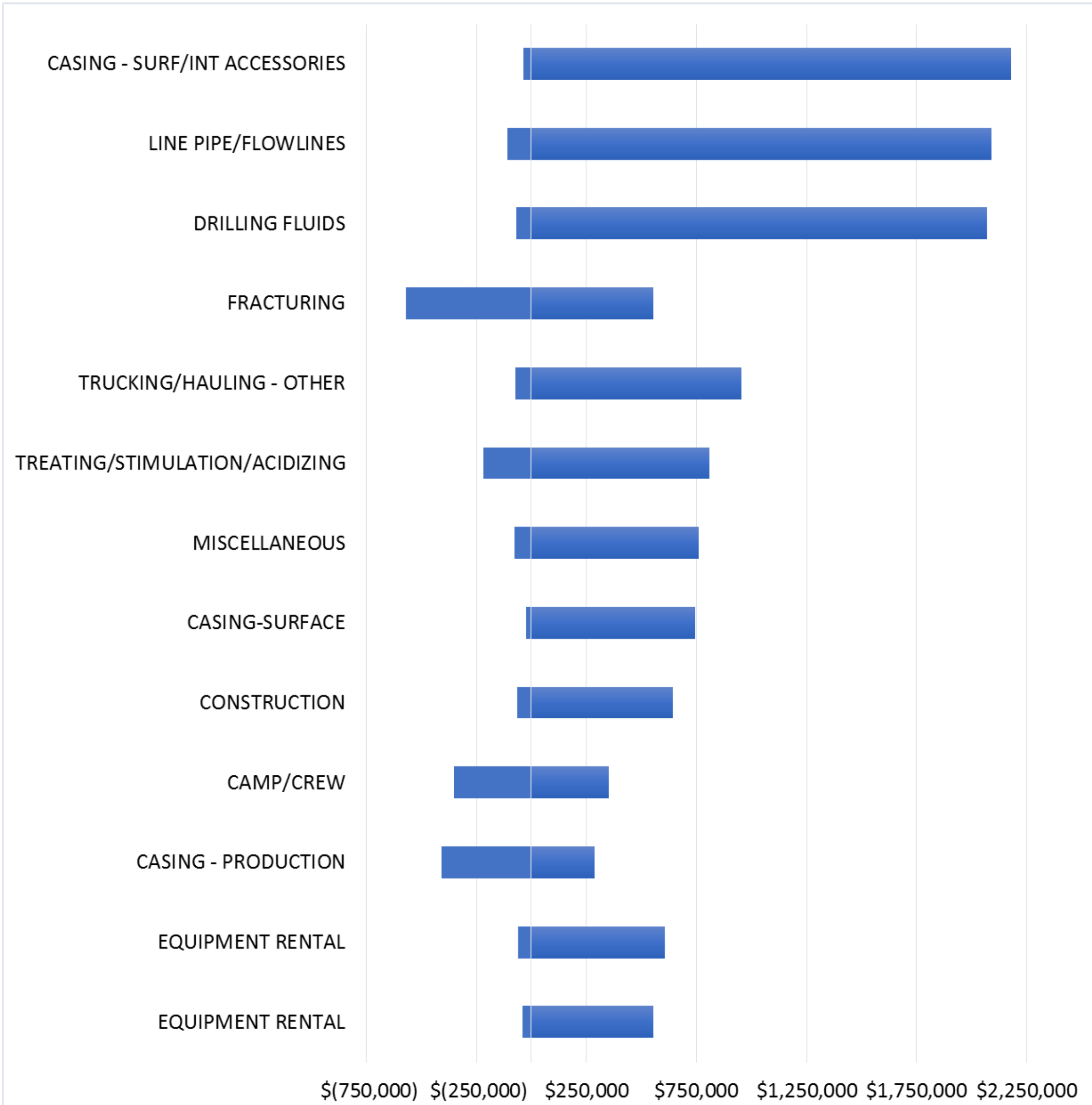

Fig. 14—Uncertainty quantification CAPEX>$500K shows the AFE sub-costs with the highest associated uncertainty. For example, "Casing-Surf/Int Accessories" has the highest magnitude of estimation uncertainty ($2.2 million).

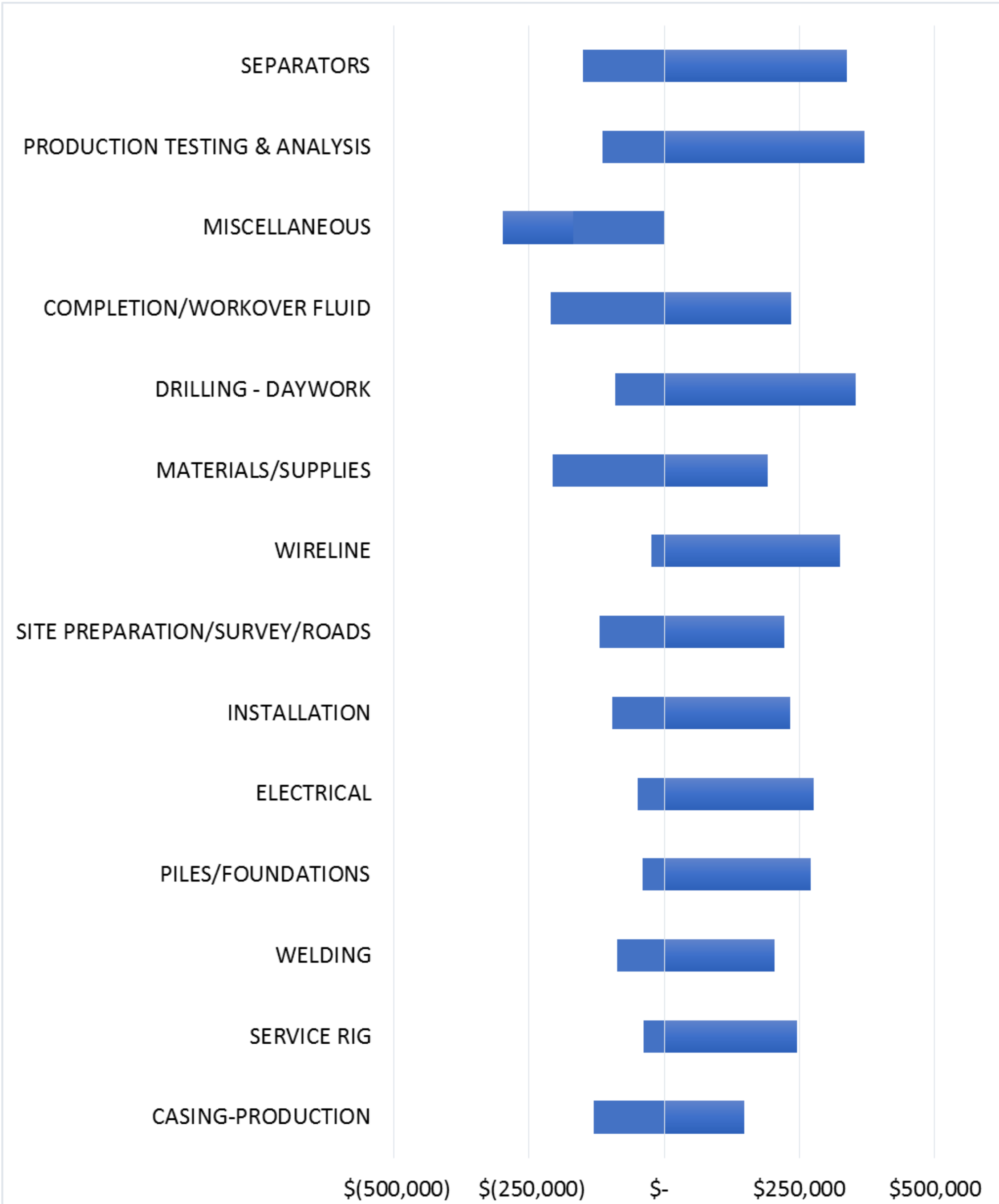

Fig. 15—Uncertainty quantification $500K>CAPEX>$250K shows the estimate uncertainty for each sub-cost.

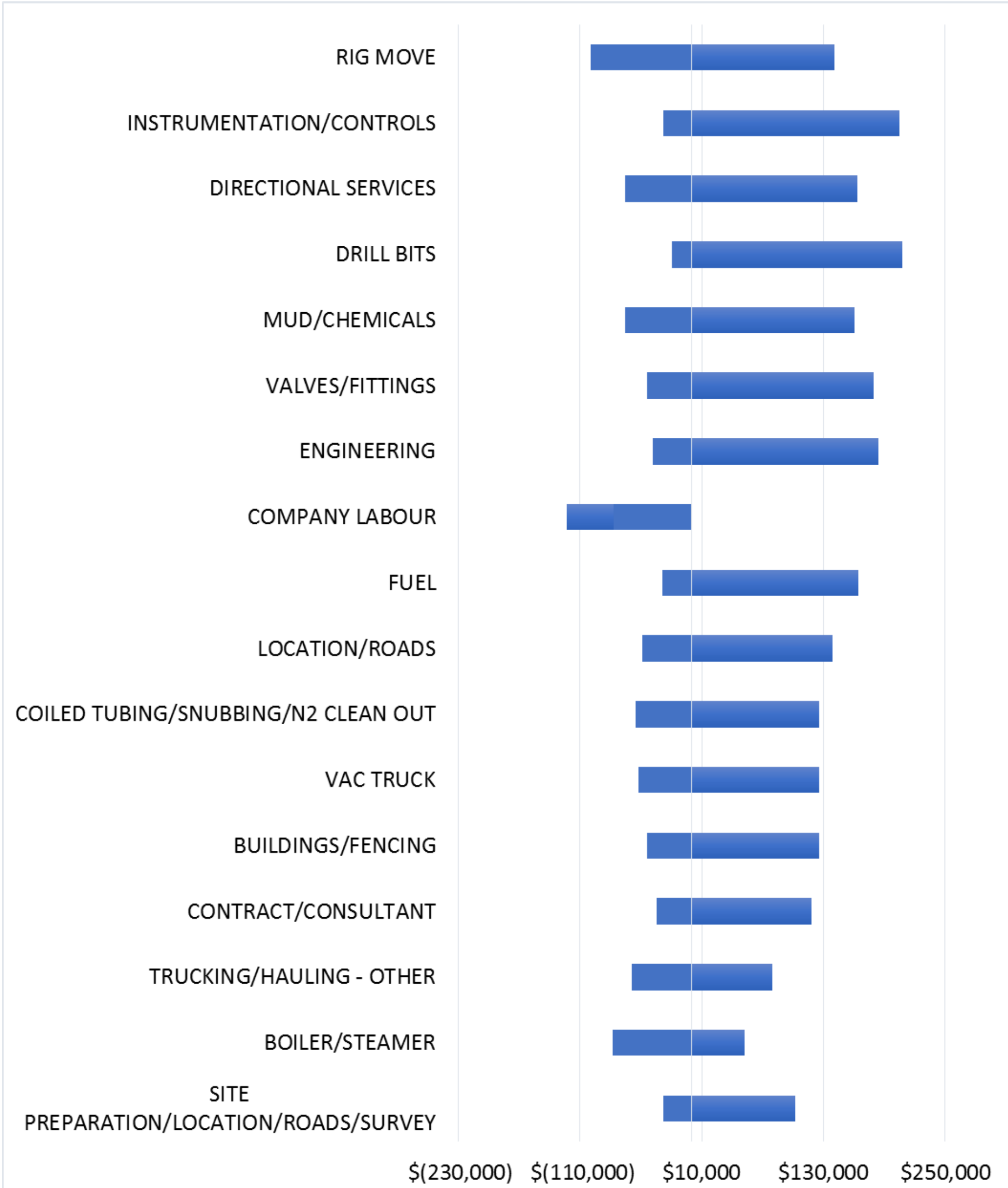

Fig. 16—Uncertainty quantification $250K>CAPEX<$120K shows the estimate uncertainty for each sub-cost.

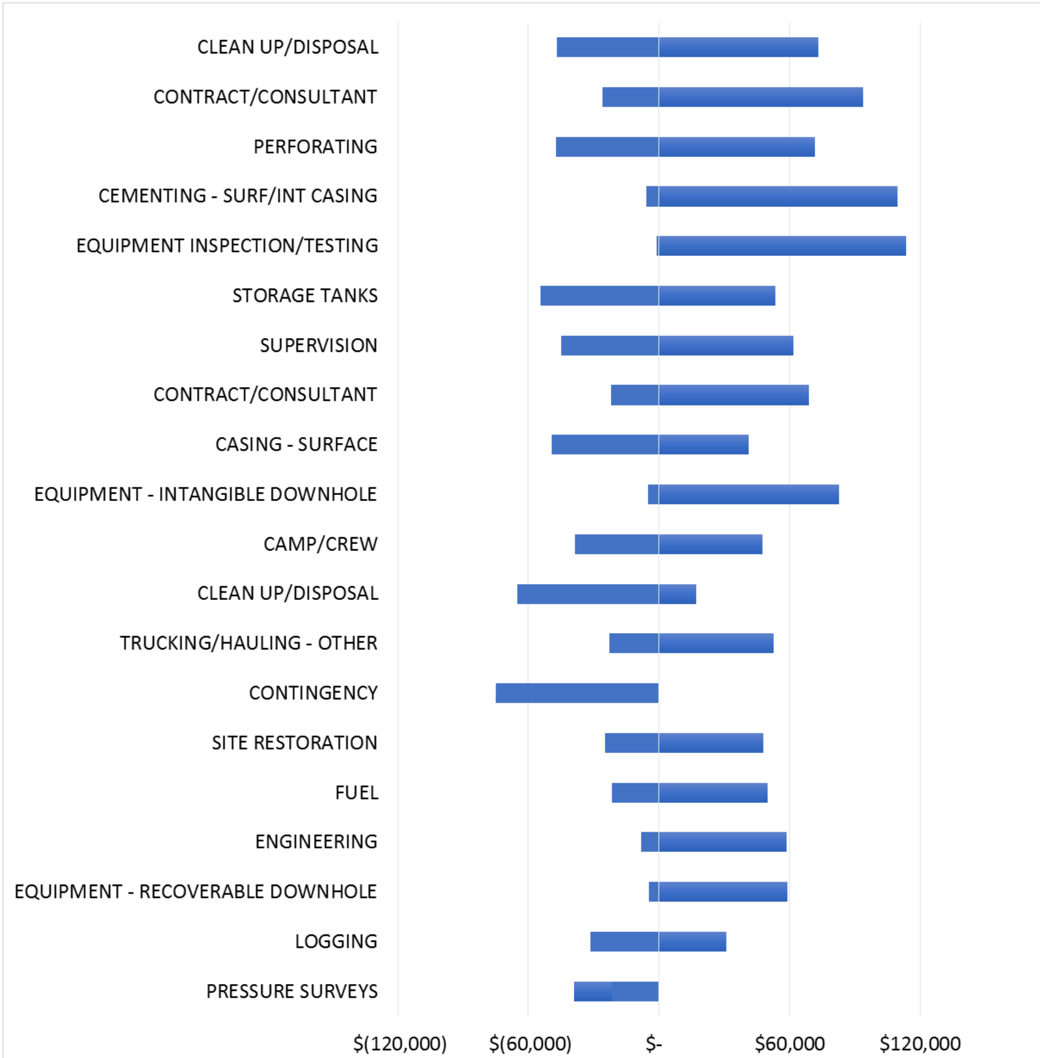

Fig. 17—Uncertainty quantification $120K>CAPEX>$60K shows the estimate uncertainty for each sub-cost.

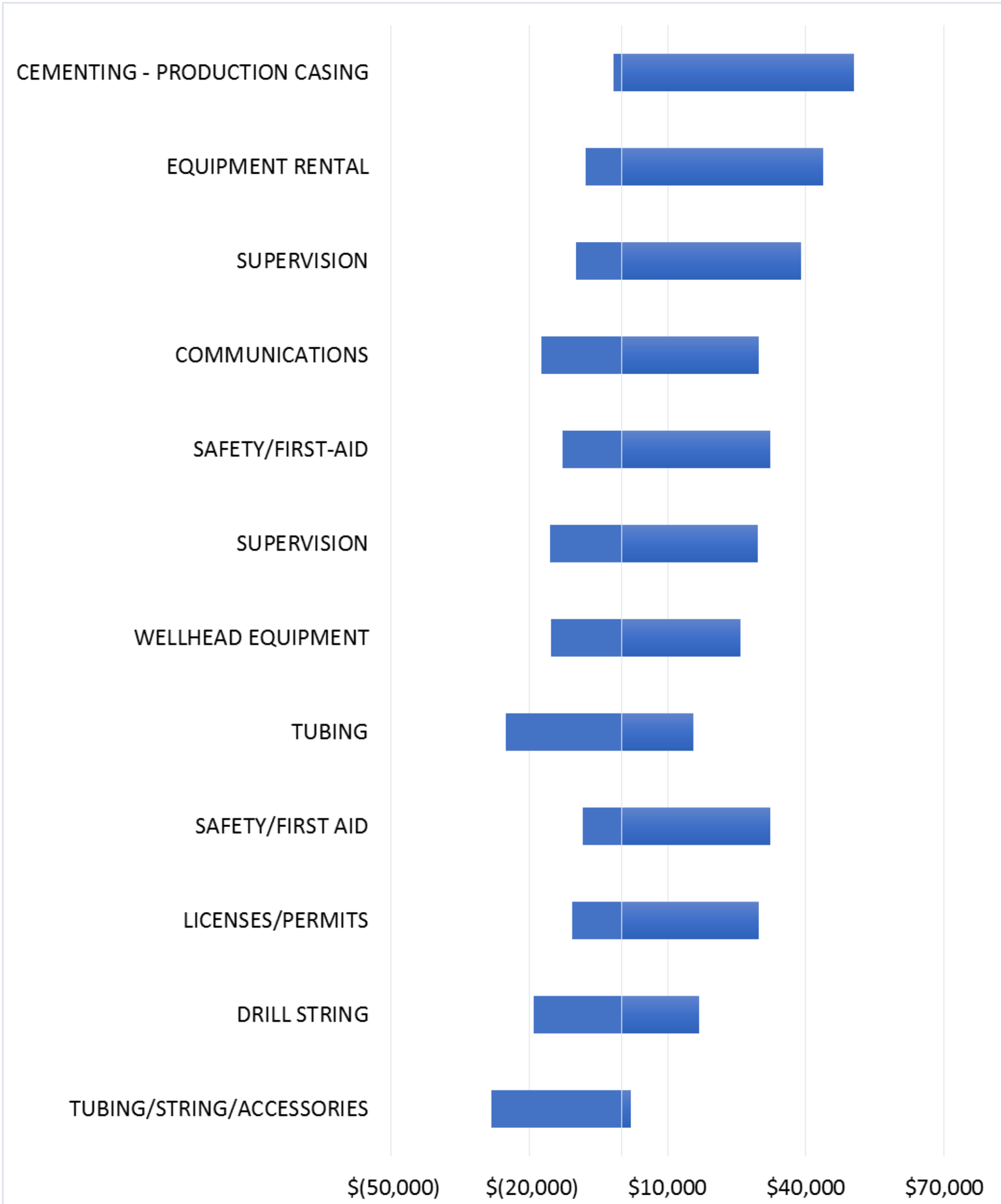

Fig 18—Uncertainty quantification $60K>CAPEX>$30K shows the estimate uncertainty for each sub-cost.

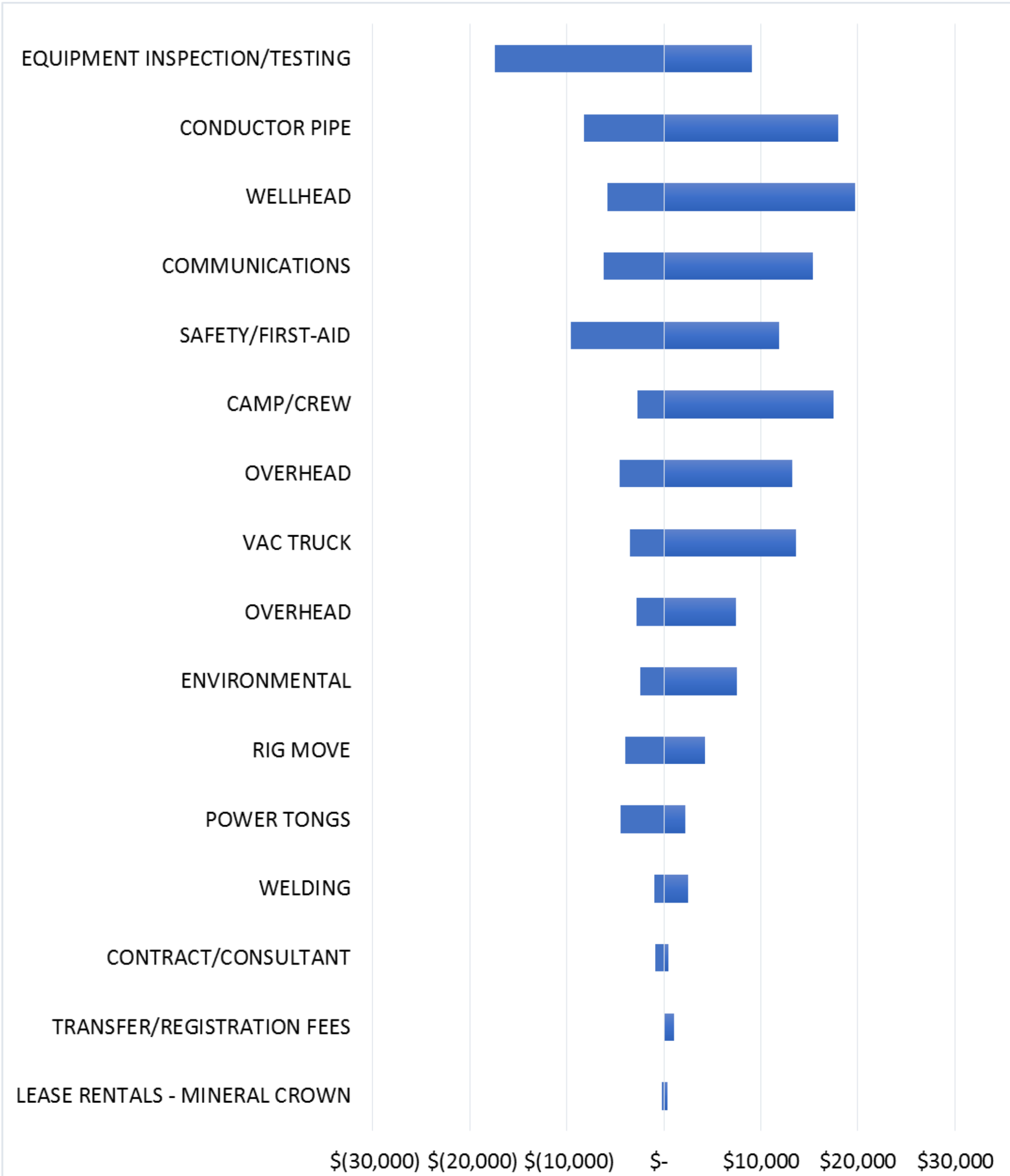

Fig. 19—Uncertainty quantification $30K>CAPEX shows the estimate uncertainty for each sub-cost.

3. PROBABILISTIC COST ESTIMATION

3.1. Methodology Overview

The probabilistic cost model was initially constructed in accordance with Valdes *et al.* (2013), as described in Section 1. The five-step procedure is as follows:

1. Determine historical correction factors for each cost code
2. Identify correlations between cost codes to incorporate dependencies
3. Truncate distributions for improved accuracy
4. Perform Monte-Carlo simulation
5. Verify results with known actual data.

The cost model of Valdes *et al.* (2013) was subsequently expanded to allow for multiple data sets to convert a proof of concept to an operational workflow which can be used as a decision making tool. The analysis was performed for various time periods of historical AFE data (2008-2013) to better predict 2014 estimates.

3.2. Step #1: Determine Distributions of Correction Factors by Cost Code

Similar to Section 2.2, distributions of the correction factors were calculated for each cost code with the distribution fit function of commercial software @RISK® by Palisades. The distribution fit function compares multiple distributions and selects the one that best fits the supplied data. This is done for each of the 97 drilling and completion sub-costs (i.e., cost codes ranging from 400-598). Cost codes Mud/Chemicals (615), Casing-Surface (626), and Casing-Production (627) were also included and combined with cost codes Mud/Chemicals (415), Casing-Surface (423), and Casing-Production (424), respectively, as they have the same cost code descriptions and could be included in drilling and completion AFEs. A full list of cost code descriptions is provided in Appendix A.

Many cost codes are infrequently used, and do not have an adequate number of data points for @RISK to complete a distribution fit (minimum five data points required). In the event that a cost code does not have at least five data points, unitary distributions are imposed to keep the automated functionally of the spreadsheet while honoring the deterministic estimate. Unitary distributions are defined as normal distributions with mean = 1 and standard deviation = 0.1 for this study.

3.3. Step #2: Determine Distribution Correlations

To better improve the accuracy and reliability of the study, the distribution correlations were then determined in the form of a 97x97 correlation matrix. A section of the correlation matrix can be seen in Fig. 20.

@RISK Correlations	401	402	403	405	406	407	409	412	413	414	416	417	418	419	420	421	422	425
401	1																	
402	0	1																
403	0.153948	0	1															
405	-0.03321	0	0.07663	1														
406	0.299786	0	0.142017	0.256342	1													
407	-0.14641	0	0.044137	0.131082	-0.32913	1												
409	0.015959	0	0.073249	0.187178	0.691106	-0.1767	1											
412	0.164369	0	0.079955	0.238024	0.700749	0.090538	0.819336	1										
413	0.095539	0	0.857913	0.022411	-0.36873	-0.16973	-0.46273	-0.67821	1									
414	0.068297	0	-0.17749	0.057839	0.432402	0.358932	0.102703	0.303335	-0.67579	1								
416	0	0	0	0	0	0	0	0	0	0	1							
417	0	0	0	0	0	0	0	0	0	0	0	1						
418	0	0	0	0	0	0	0	0	0	0	0	0	1					
419	0	0	0	0	0	0	0	0	0	0	0	0	0	1				
420	-0.23602	0	-0.48168	0.257777	-0.18962	-5E-05	-0.60037	-0.23847	-0.68931	-0.00902	0	0	0	0	1			
421	-0.55822	0	-0.6752	-0.03004	0.437985	-0.67358	0.35864	-0.40053	-0.94979	0.816353	0	0	0	0	0.185361	1		
422	0.287404	0	0.073517	0.248369	0.888004	-0.09217	0.849805	0.843129	-0.51522	0.388423	0	0	0	0	-0.48279	0.57434	1	
425	0.352741	0	0.027217	0.089055	0.774992	0.347801	0.259407	0.30949	-0.81406	0.7674	0	0	0	0	-0.1679	-0.45331	0.723584	1

Fig. 20—Correlation matrix for cost codes 401-425.

The correlations were determined utilizing Microsoft Excel® *correl* function and incorporated into the sampling of the cost code distributions with Palisades' @RISK® *RiskCorrmat* function. Correlation values equal to 1 represent perfect correlations while values equal to -1 represent perfectly inverse correlations. Perfect correlations are of course seen along the center diagonal as they are the distributions being correlated with themselves. For example, cost code 409 correlates perfectly with cost code 409. Excluding these known perfect correlations, all correlation values equal to 1 or -1 were due to small number of data points. These values were filtered out of the correlation matrix and replaced with no correlation which is shown with a zero. As stated in Section 3.2 above, many cost codes do not have the required amount of data points for reliable distributions and therefore represent unrealistic correlations, which if imposed would lead to unreliable estimates.

3.4. Step #3: Truncate Distributions

Once the distributions were determined and the distribution correlations were incorporated, it was necessary to truncate the distributions to filter unrealistic results. In many cases the correction factor distribution is log-normal in shape. In some instances, the upper end of the distribution (99^{th} percentile) can extend an abnormal length, resulting in unrealistically high Monte Carlo simulation results. These values can be in the range of billions of dollars for a single well estimate, which is not realistic. The distributions were truncated with Palisades' @RISK® *RiskTruncateP* function, which eliminates the top and bottom distribution percentiles from the simulation. The process of truncation narrows the sampling window of the distributions by placing upper and lower bounds. It is important to note that data itself is not cut (the full range of data is used to determine the distribution), only the sampling window of the distribution in subsequent steps is narrowed. Fig. 21 shows the mode, median, and mean of the truncated 406 drilling day work distribution are still equal to 1.09, 1.18, and 1.31, respectively.

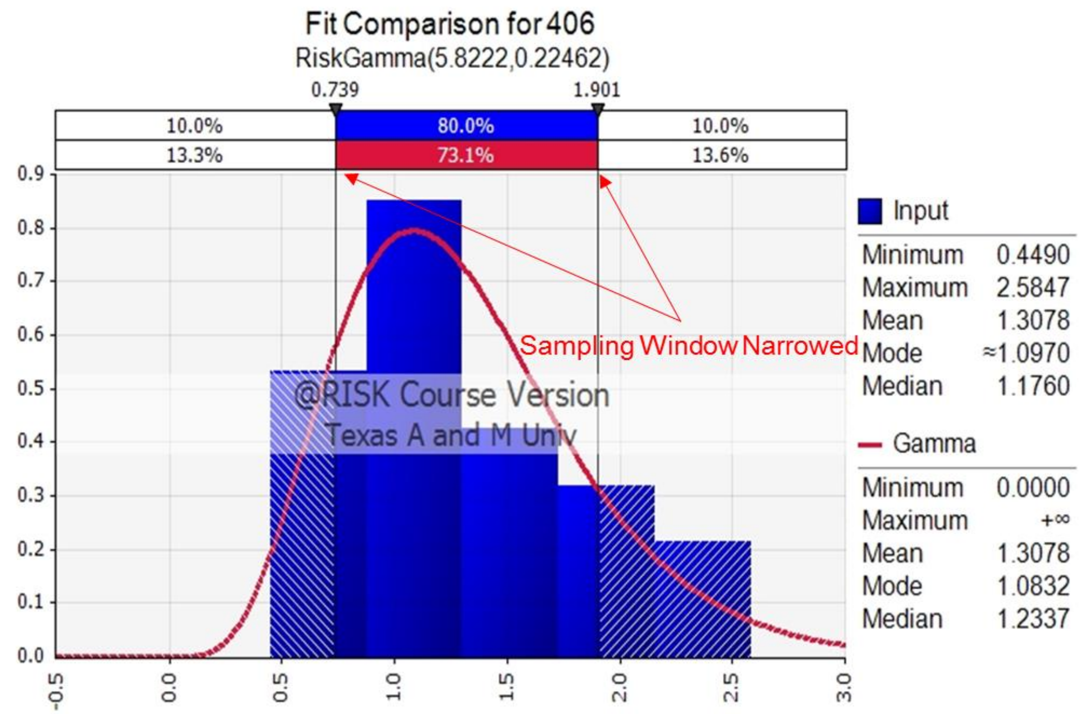

Fig. 21—406 drilling day work 10% truncation shows that truncation is not eliminating data. The truncation only narrows the upper and lower bounds for sampling. 406 drilling day work with no truncation is seen previously in Fig. 13.

The truncated distributions were analyzed for accuracy by utilizing the entire data set of deterministic estimates (2008-2014) to predict known actual costs for three truncation profiles: no truncation, 5% truncation, and 10% truncation (Fig. 22). The deterministic estimates were underestimated by approximately 16%. The probabilistic cost estimates with no truncations results were unreasonably over estimated by values ranging as high as several billions of dollars; however, the probabilistic cost estimates with a 5% truncation were over estimated by 11%. Continuing the trend of greater accuracy, the 10% truncation estimates were overestimated by 4%. Due to the increased accuracy, the 10% truncation profile was utilized in this study. This can be clearly seen in Fig. 22, as the sum of the probabilistic estimates with 10% truncation is located nearest to the perfect estimate line, which represents where the estimated cost is equal to actual cost.

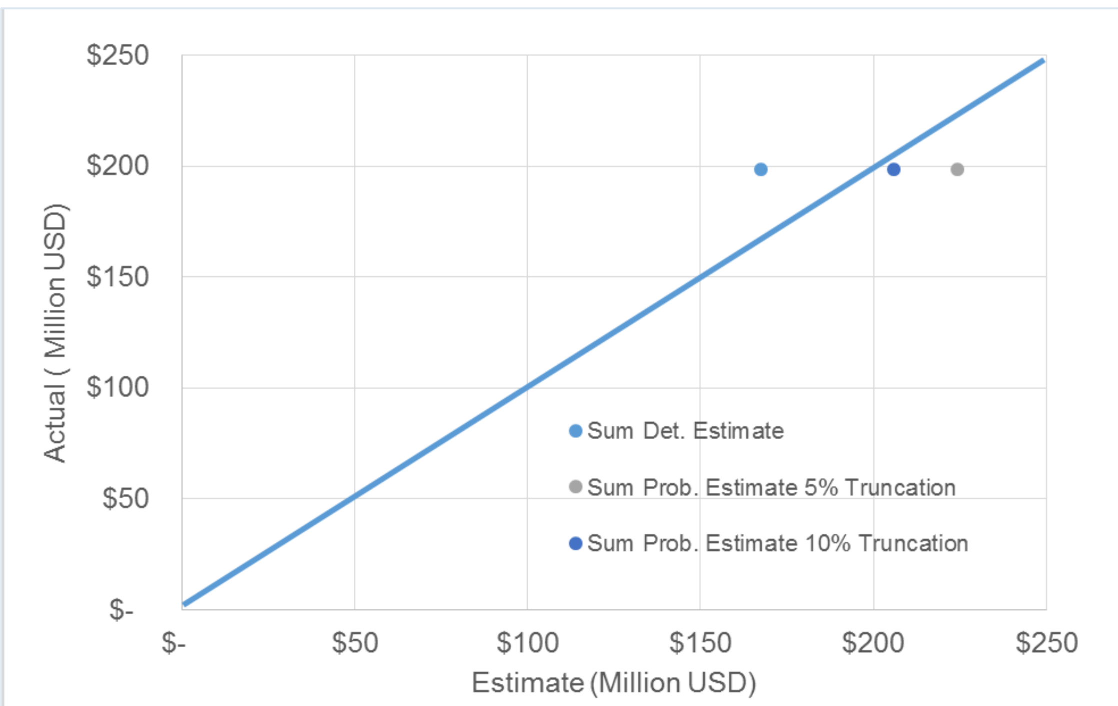

Fig. 22—Truncation profile accuracy shows the improving trend of increasing truncation percentile.

3.5. Step #4: Determine Probabilistic Estimates Using Monte-Carlo Simulation

With all the distributions correlated and truncated for maximum accuracy, the Monte-Carlo simulation was performed. This was completed by utilizing the probabilistic cost model, Eq. 2 and Eq. 3.

$$C_T = \sum_{i=1}^{n} X_i C_i = X_1 C_1 + X_2 C_2 + X_3 C_3 + \cdots + X_n C_n \quad (2)$$

$$X_i = \frac{\text{Actual well cost}}{\text{Estimated well cost}} \quad (3)$$

The deterministic estimate for each sub-cost was first multiplied by the correlated and truncated distribution to create a probabilistic sub-cost estimation distribution. A simulation was run with a 1,000 iterations to determine the mean values for each sub-cost estimate. The sub-cost mean values were summed taken to find the total mean probabilistic estimate for each AFE.

3.6. Step #5: Validate Results with Known Actual Data

This procedure was completed to calculate mean probabilistic estimates for each of the 2014 well cost AFEs using historical data ranges from 2008-2013, 2010-2013, 2011-2013, 2012-2013, and 2013 to determine the optimum amount of historical data to be included. The data range 2009-2013 was not examined as the year 2009 had too few AFEs to significantly change results. The results of each simulation were then compared to the deterministic 2014 AFE estimates (Figs. 23, 26, 29, 32 and 35). Following each simulation figure are percentile charts (Figs. 24, 27, 30, 33 and 36). These charts are the first step in creating a decision making tool by associating P10, P30, P50, P70 and P90 percentiles with the mean probabilistic estimate. The percentile charts show the effect of uncertainty ranges on whether or not each AFE is dependent or independent of each other. As it is most likely that the AFEs will be neither completely dependent nor completely independent of each other, the range of uncertainty of the 2014 mean probabilistic estimates was found within the bounds of the lines plotted by the completely dependent and independent points.. After percentiles were determined, it was important to then calibrate the percentiles to determine the optimism or pessimism and the confidence level in the cost model (Figs. 25, 28, 31, 34 and 37) in the form of calibration charts.

Fig. 23 shows the sum of the 2014 mean deterministic estimates was approximately $63 million dollars while the sum of the actual costs of the 2014 AFE's was approximately $82 million dollars. This is a $19 million dollar (23%) underestimation. The sum of the 2014 mean probabilistic cost estimates using 2008-2013 data was approximately $76 million dollars, which is only a $6 million dollar (7%) underestimation. The probabilistic cost model increased the accuracy of the deterministic estimate by $13 million dollars (16%).

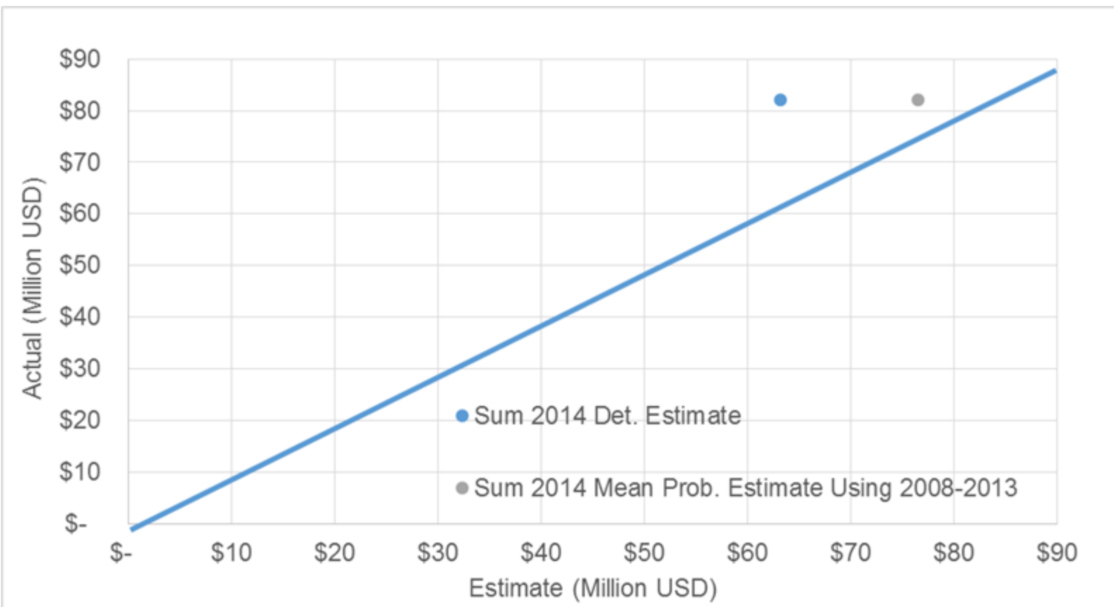

Fig. 23—2014 probabilistic estimate using 2008-2013 data shows the mean probabilistic estimate (7% underestimation) and deterministic estimate (23% underestimation) in comparison to the blue perfect estimate line.

Fig. 24 shows the effect of uncertainty ranges on if each AFE is dependent or independent of each other. The associated uncertainty in assuming all AFEs are completely dependent on each other is $39 million dollars. The associated uncertainty in assuming all AFEs are completely independent of each other is $24 million dollars. As previously stated, since both opposite assumptions are both likely to be wrong, the actual uncertainty is between $24 and $39 million dollars.

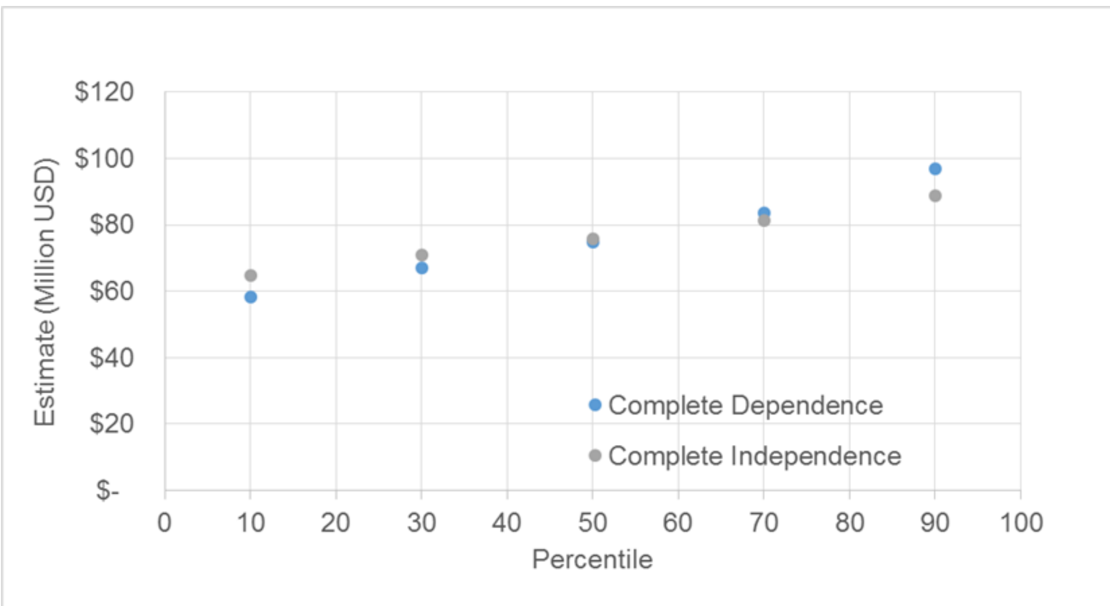

Fig. 24—2014 AFE dependence/independence relationship for 2008-2013 data shows the uncertainty in the sum of 2014 AFEs in the cases of complete AFE dependence or independence with each other.

The calibration chart (Fig. 25) shows overconfidence (slope less than 1), meaning the distributions are too narrow, and pessimistic Directional Bias (Db) which means overly conservative estimates.

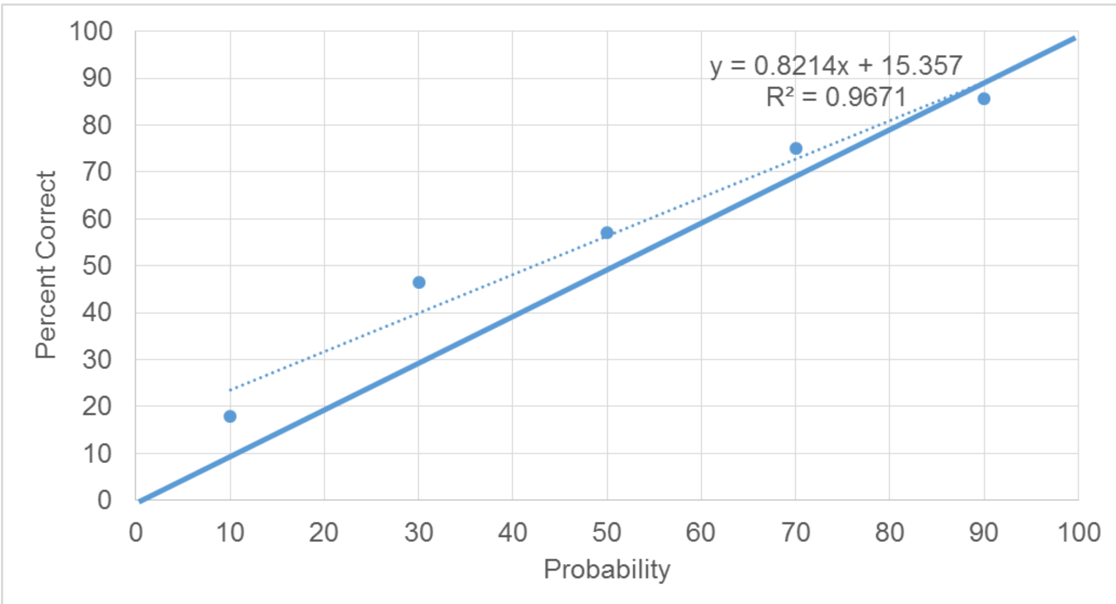

Fig. 25—2014 probabilistic estimate calibration chart for 2008-2013 data shows calculated percentiles in comparison to blue perfect calibration line.

The sum of the 2014 mean probabilistic cost estimates using 2010-2013 data (Fig. 26) was approximately $77 million dollars which is only a $5 million dollar (6%) underestimation. When compared to the deterministic $63 million dollar sum, $19 million dollar underestimation (23%), the probabilistic cost model increased the accuracy of the deterministic estimate by $14 million dollars (17%).

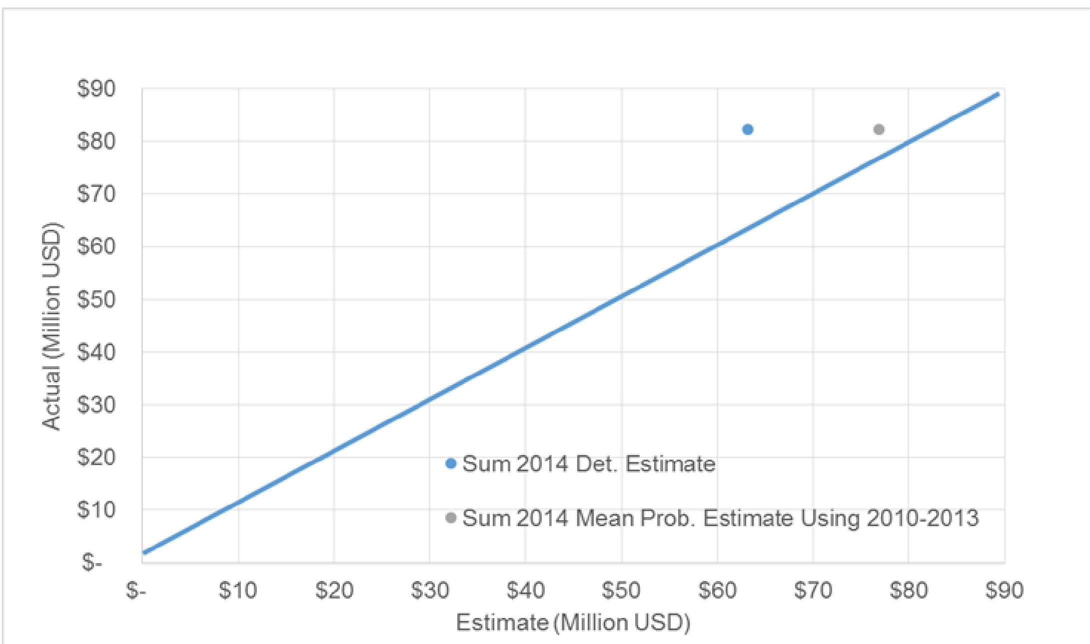

Fig. 26—2014 probabilistic estimate using 2010-2013 data shows the mean probabilistic estimate (6% underestimation) and deterministic estimate (23% underestimation) in comparison to the blue perfect estimate line.

Fig. 27 shows the effect of uncertainty ranges based on AFE dependence on each other. The associated uncertainty in assuming all AFEs are completely dependent on each other is $37 million dollars. The associated uncertainty in assuming all AFEs are completely independent of each other is $23 million dollars. As previously stated, since both assumptions are likely to be wrong, and thus it can be assumed the actual uncertainty is between $23 and $37 million dollars.

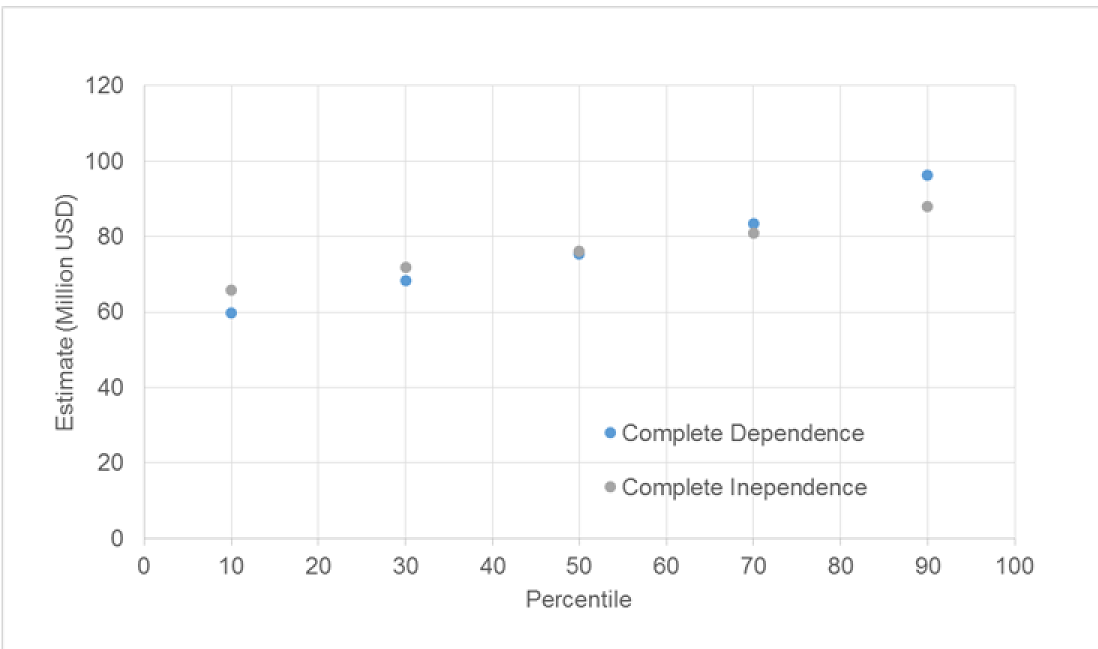

Fig. 27—2014 AFE dependence/independence relationship for 2010-2013 data shows the uncertainty in the sum of 2014 AFEs in the cases of complete AFE dependence or independence with each other.

The calibration chart (Fig. 28) shows an increase in both overconfidence (slope less than 1 meaning the distributions are too narrow) and pessimistic DB (overly conservative estimates) with decreased data quantity.

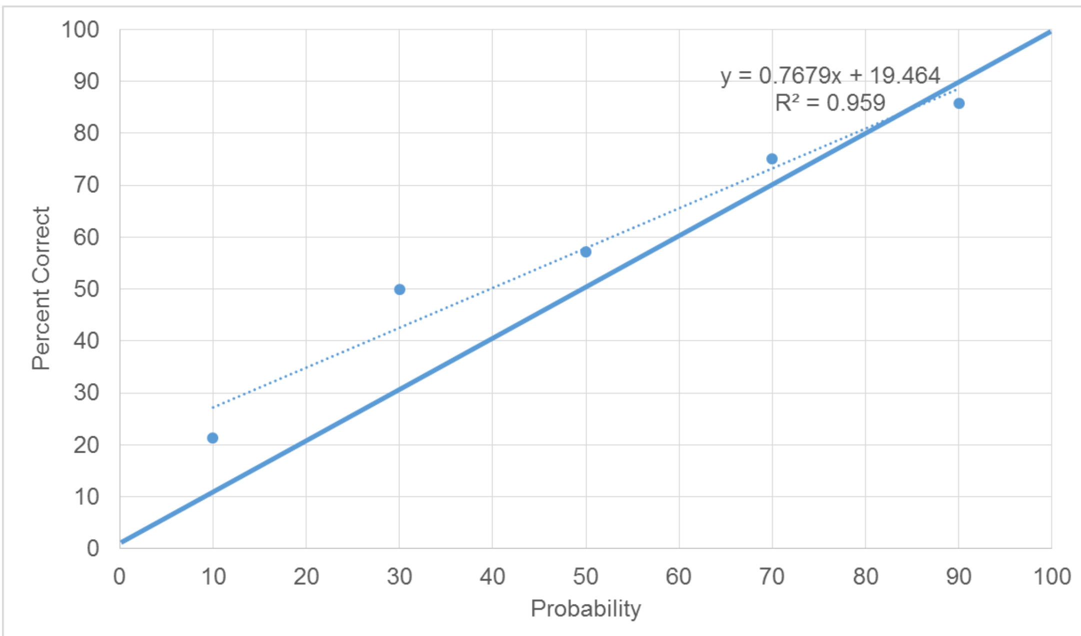

Fig. 28—2014 probabilistic estimate calibration chart using 2010-2013 data shows calculated percentiles in comparison to blue perfect calibration line.

The sum of the 2014 mean probabilistic cost estimates using 2011-2013 data (Fig. 29) was approximately $78 million dollars which is only a $4 million dollar (5%) underestimation. When compared to the deterministic $63 million dollar sum, $19 million dollar underestimation (23%), the probabilistic cost model increased the accuracy of the deterministic estimate by $15 million dollars (18%).

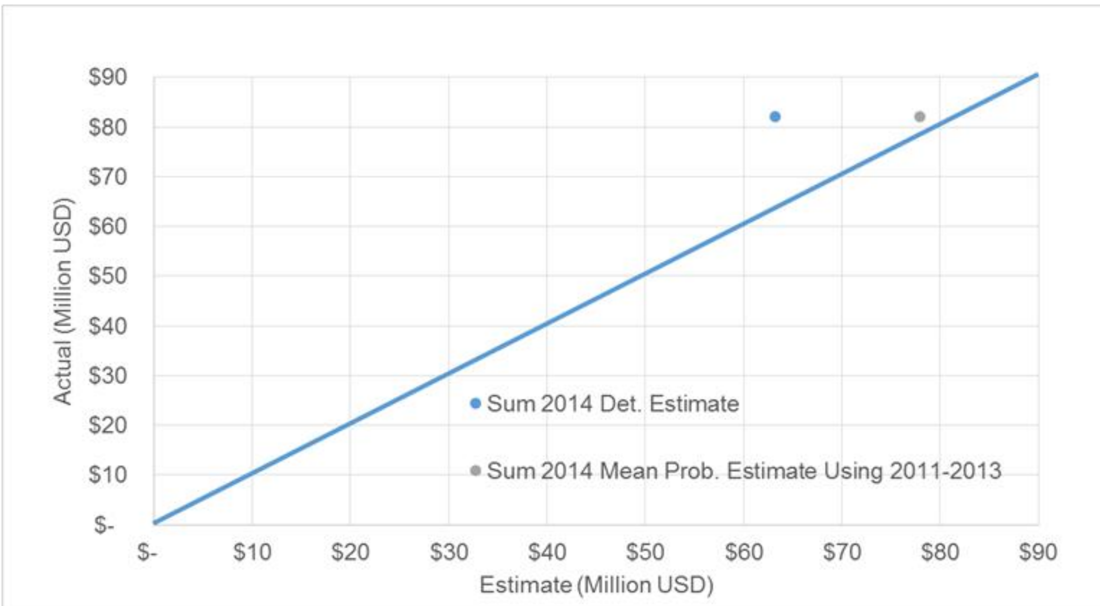

Fig. 29—2014 probabilistic estimate using 2011-2013 data shows the mean probabilistic estimate (5% underestimation) and deterministic estimate (23% underestimation) in comparison to the blue perfect estimate line.

Fig. 30 shows the effect of uncertainty ranges based on AFE dependence on each other. The associated uncertainty in assuming all AFEs are completely dependent on each other is $35 million dollars. The associated uncertainty in assuming all AFEs are completely independent of each other is $22 million dollars. As previously stated, since both assumptions are likely to be wrong, the actual uncertainty is between $22 and $35 million dollars.

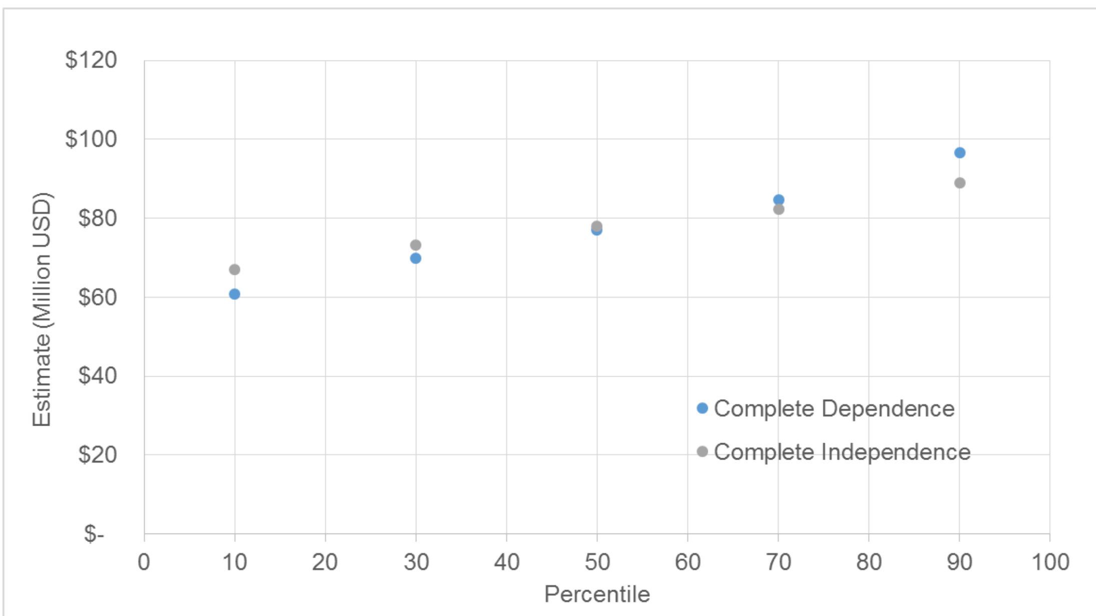

Fig. 30—2014 AFE dependence/independence relationship for 2011-2013 data shows the uncertainty in the sum of 2014 AFEs in the cases of complete AFE dependence or independence with each other.

The calibration chart (Fig. 31) shows an increase in both overconfidence (slope less than 1 meaning the distributions are too narrow) and pessimistic DB (overly conservative estimates) with decreased data quantity.

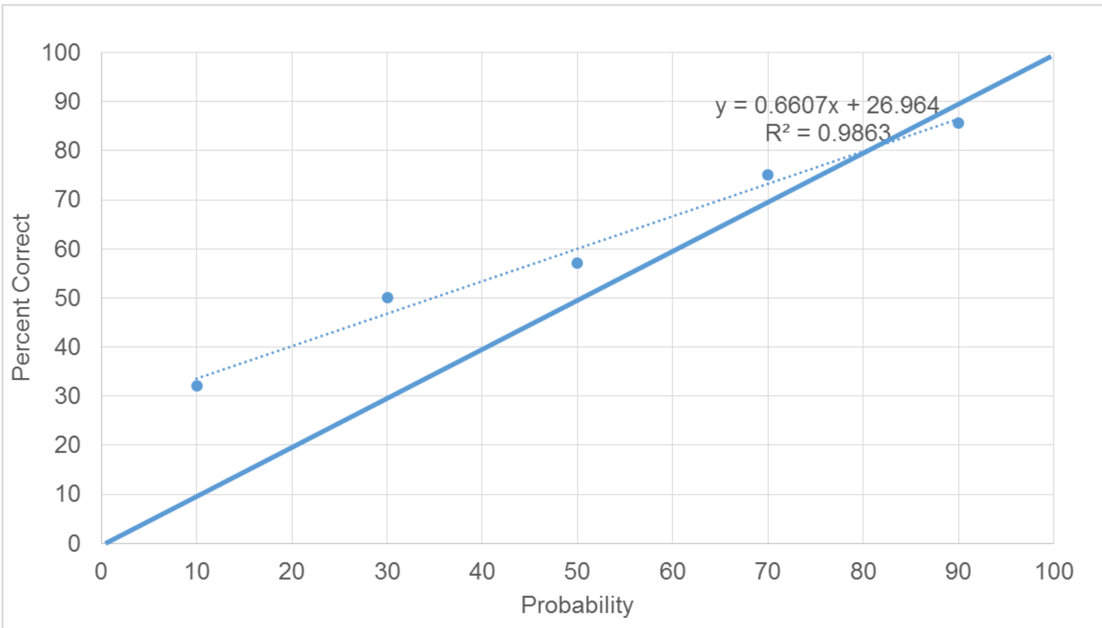

Fig. 31—2014 probabilistic estimate calibration chart using 2011-2013 data shows calculated percentiles in comparison to blue perfect calibration line.

The sum of the 2014 mean probabilistic cost estimates using 2012-2013 data (Fig. 32) was approximately $91 million dollars which is a $9 million dollar (11%) overestimation. When compared to the deterministic $63 million dollar sum, $19 million dollar underestimation (23%), the probabilistic cost model increased the accuracy of the deterministic estimate by $14 million dollars (17%).

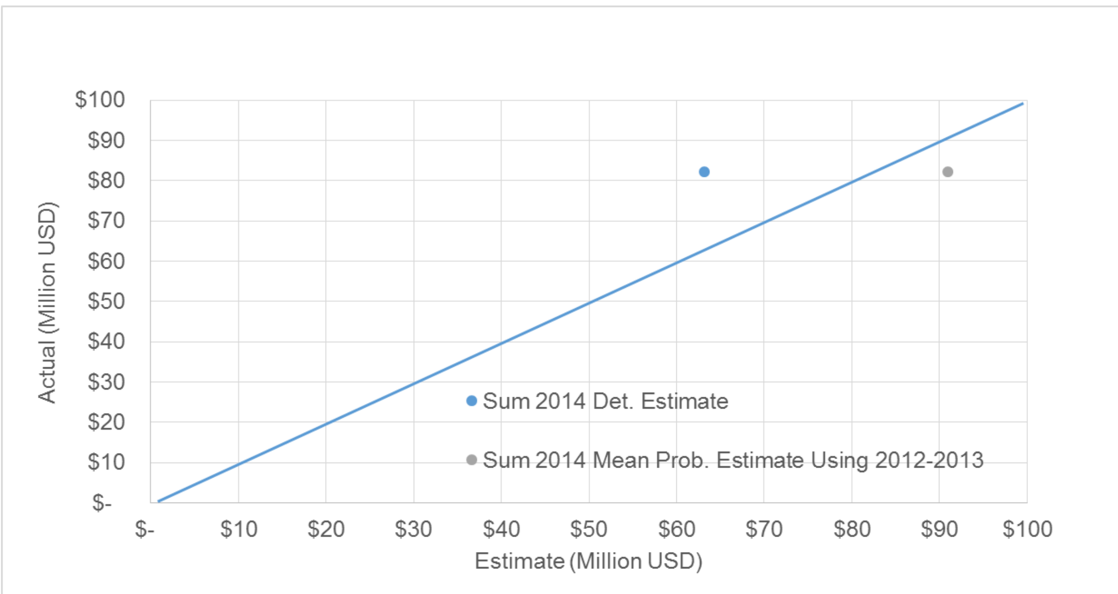

Fig. 32—2014 probabilistic estimate using 2012-2013 data shows the mean probabilistic estimate (11% over-estimation) and deterministic estimate (23% underestimation) in comparison to the blue perfect estimate line.

Fig. 33 shows the effect of uncertainty ranges based on AFE dependence on each other. The associated uncertainty in assuming all AFEs are completely dependent on each other is $45 million dollars. The associated uncertainty in assuming all AFEs are completely independent of each other is $28 million dollars. As previously stated, since both opposite assumptions are both likely to be wrong, the actual uncertainty is between $28 and $45 million dollars.

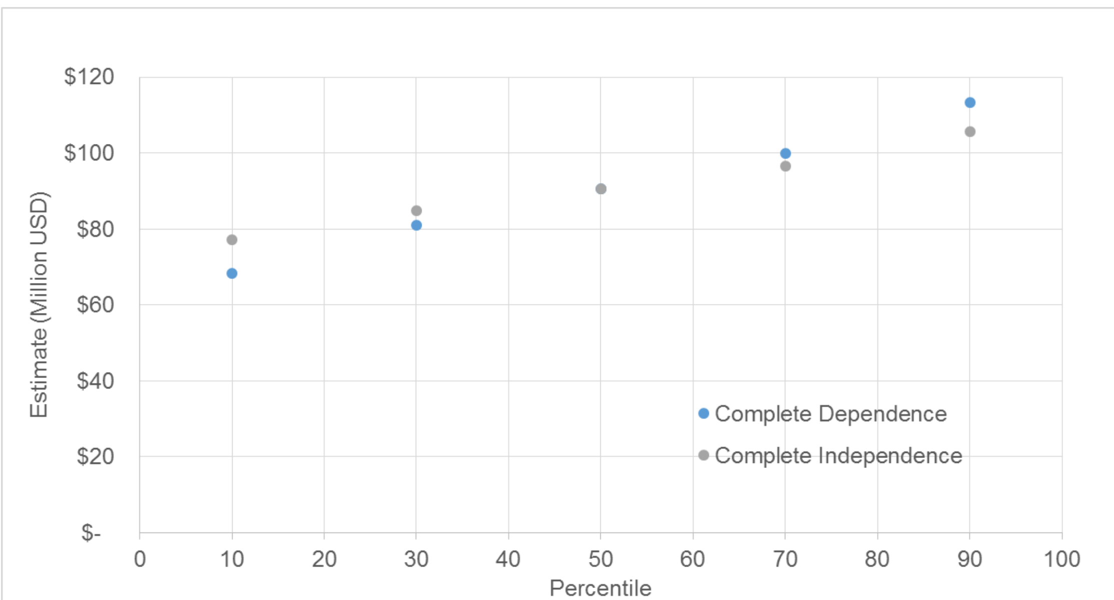

Fig. 33—2014 AFE dependence/independence relationship for 2012-2013 data shows the uncertainty in the sum of 2014 AFEs in the cases of complete AFE dependence or independence with each other.

The calibration chart (Fig. 34) shows an increase in both overconfidence (slope less than 1 meaning the distributions are too narrow) and pessimistic DB (overly conservative estimates) with decreased data quantity.

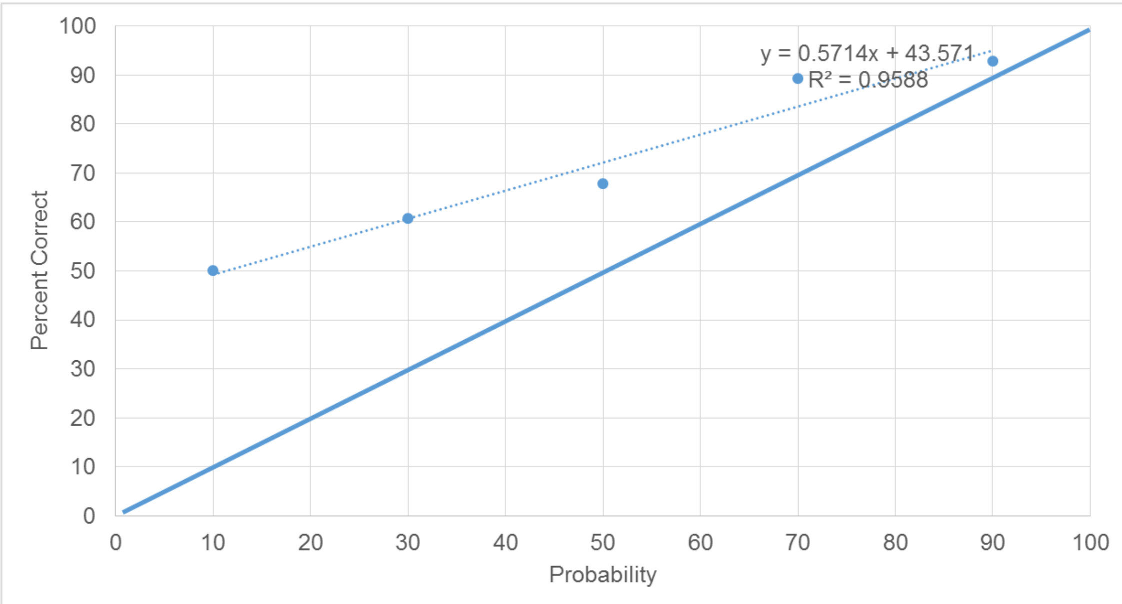

Fig. 34—2014 probabilistic estimate calibration chart using 2012-2013 data shows calculated percentiles in comparison to blue perfect calibration line.

The sum of the 2014 mean probabilistic cost estimates using only 2013 data (Fig. 35) was approximately $70 million dollars which is a $12 million dollar (15%) underestimation. When compared to the deterministic $63 million dollar sum, $19 million dollar underestimation (23%), the probabilistic cost model increased the accuracy of the deterministic estimate by $4 million dollars (8%).

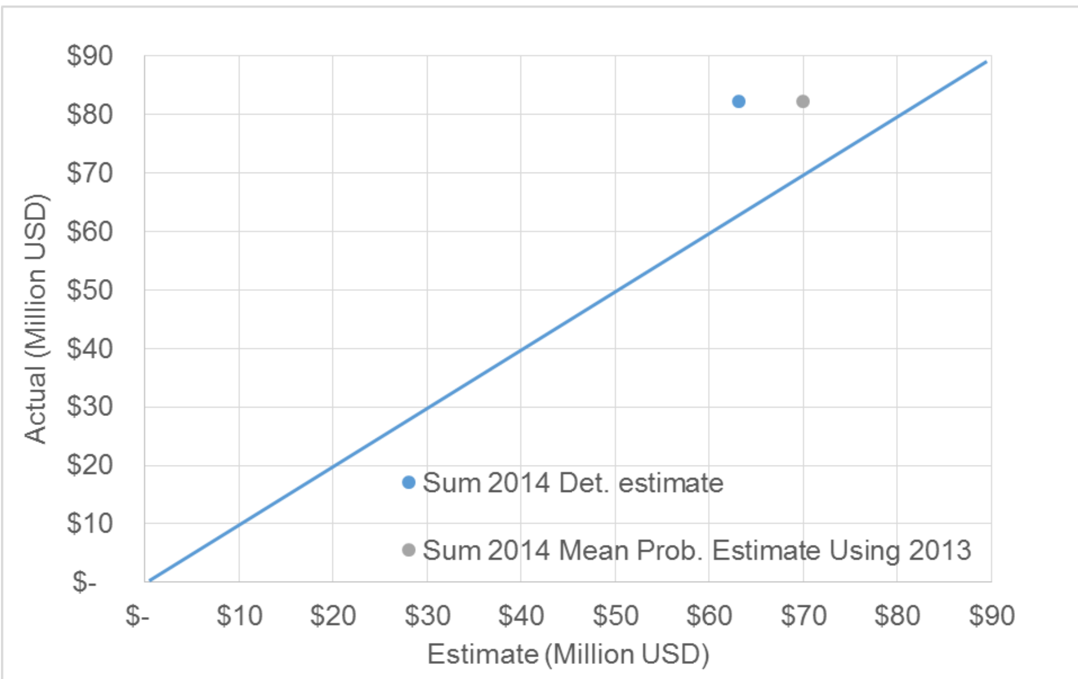

Fig. 35—2014 probabilistic estimate using 2013 data shows the mean probabilistic estimate (15% underestimation) and deterministic estimate (23% underestimation) in comparison to the blue perfect estimate line.

Fig. 36 shows the effect of uncertainty ranges based on AFE dependence on each other. The associated uncertainty in assuming all AFEs are completely dependent on each other is $27 million dollars. The associated uncertainty in assuming all AFEs are completely independent of each other is $17 million dollars. As previously stated, since both opposite assumptions are likely to be wrong, the actual uncertainty is between $17 and $27 million dollars.

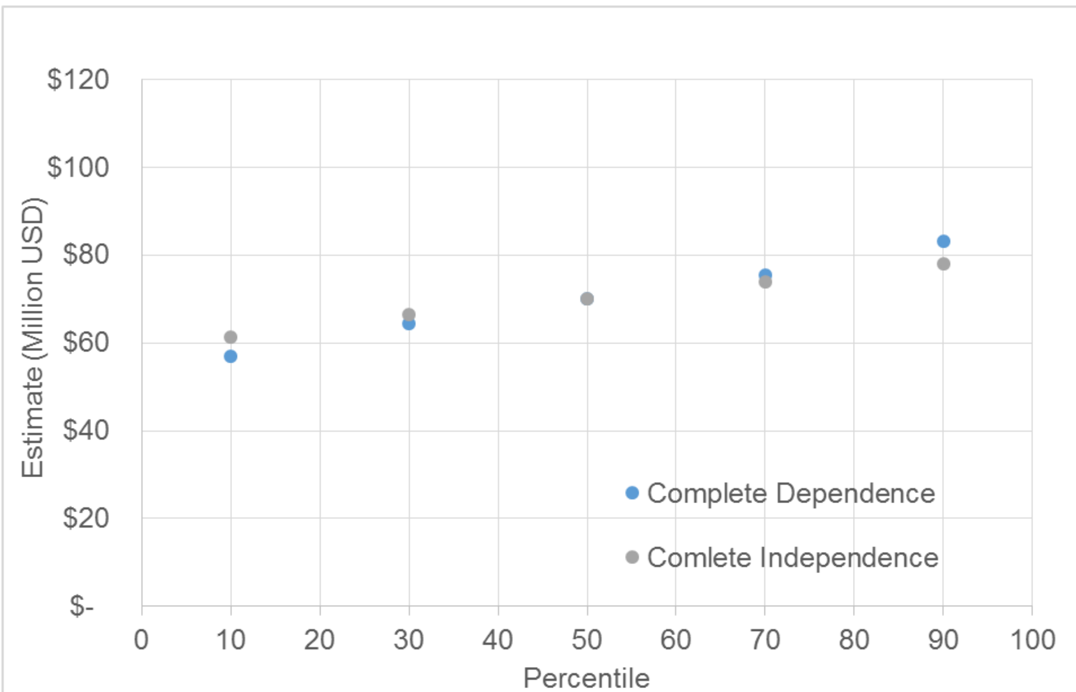

Fig. 36—2014 AFE dependence/independence relationship for 2013 data shows the uncertainty in the sum of 2014 AFEs in the cases of complete AFE dependence or independence with each other.

The calibration chart (Fig. 37) shows an increase overconfidence (slope less than 1 meaning distributions are too narrow) but a decrease in pessimistic DB as data becomes most current.

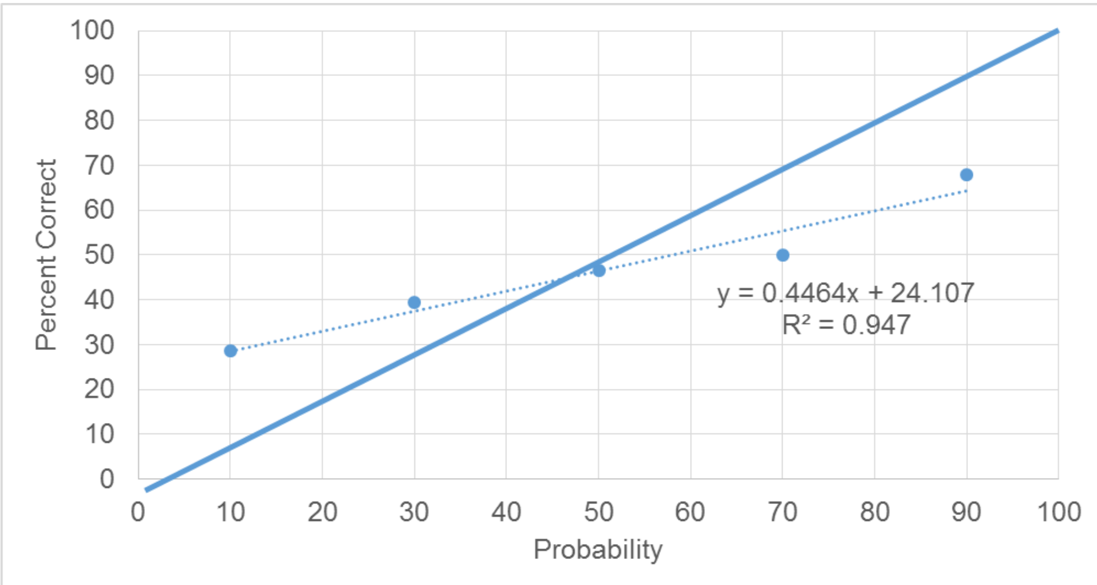

Fig. 37—2014 probabilistic estimate calibration chart using 2013 data shows calculated percentiles in comparison to blue perfect calibration line.

3.7. Discussion

Fig. 38 shows the summary of the comparisons of the 2014 deterministic estimates to the 2014 probabilistic estimates for each of the historical data set simulations. It is important to note that all probabilistic estimates are more accurate than the deterministic estimates. The simulations using 2008-2013, 2010-2013, and 2011-2013 represent an increasing accuracy trend while simulations using 2012-2013 and only- 2013 data begin to decrease in accuracy (Table 2). For all data sets, excluding 2012-2013, optimistic estimates (underestimations) are observed.

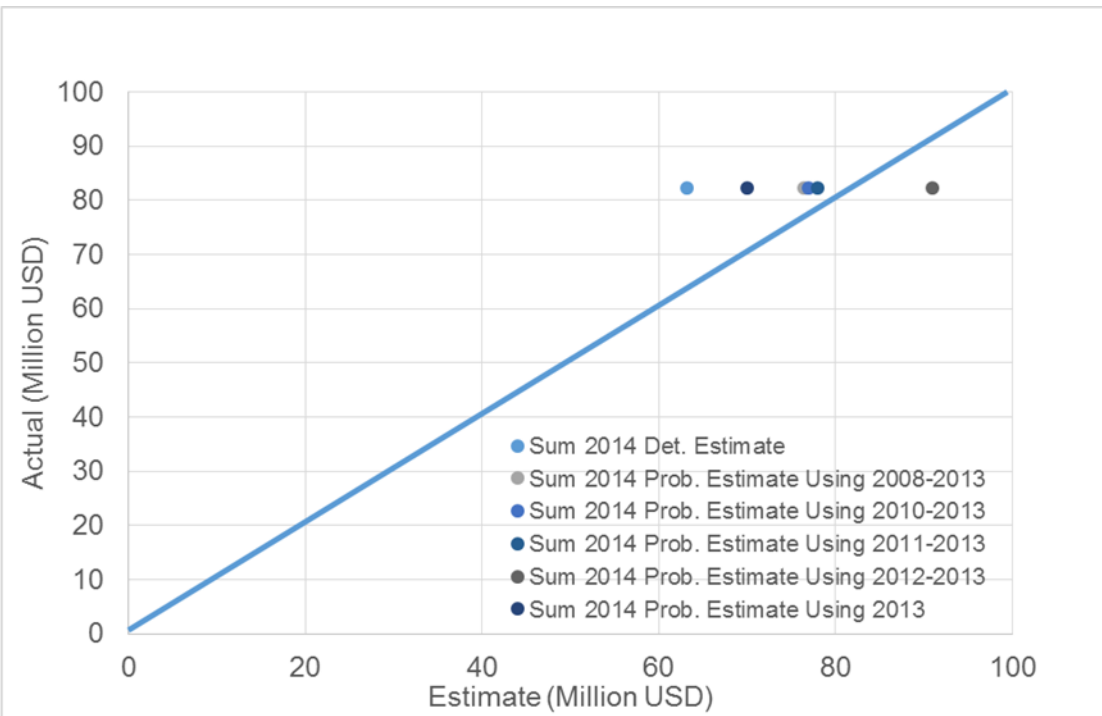

Fig. 38—Comparison of 2014 probabilistic estimates for each historical data set simulation (2008-2013; 2010-2013; 2011-2013; 2012-2013; 2013).

Table 2—2014 probabilistic estimation summary comparison (negative percent errors represent underestimates, positive percent errors represent overestimates)

2014 Estimation	Sum Det. Estimate (Million USD)	Sum Prob. Estimate (Million USD)	Sum Actual (MMUSD)	Percent Error Det. -Actual	Percent Error Prob. -Actual
2008-2013	$ 63	$ 76	$ 82	-23%	-7%
2010-2013	$ 63	$ 77	$ 82	-23%	-6%
2011-2013	$ 63	$ 78	$ 82	-23%	-5%
2012-2013	$ 63	$ 91	$ 82	-23%	11%
2013	$ 63	$ 70	$ 82	-23%	-15%

The reasons for the decrease in accuracy for the 2012-2013 and 2013 data sets is seen in Table 3, which shows the number of unitary distributions (described in Section 3.2) in relation to the total amount of cost code distributions. Within this table, we notice two trends. First, the probabilistic estimate becomes more accurate as the years of the historical data set become more current. Second, as the unitary distributions become more than half of the total distributions (as seen in historical data set 2012-2013), the probabilistic estimate loses accuracy as it begins to convert into a more traditional deterministic estimate. This allows us to conclude that using the most current historical data set, while maintaining enough data points to stay probabilistic is the most optimal estimating practice. The most current historical data set will represent the most accurate industry related trends (service costs, commodity pricing, etc.).

Table 3--2014 probabilistic estimation distribution count

2014 Estimation	Percent Error Det. -Actual	Percent Error Prob. -Actual	Unitary Dist.	Total Dist.
2008-2013	-23%	-7%	37	94
2010-2013	-23%	-6%	38	94
2011-2013	-23%	-5%	40	94
2012-2013	-23%	11%	59	94
2013	-23%	-15%	76	94

Table 4--Number of AFEs supplied and analyzed in each year. The difference is due to data quality issues as previously discussed.

Year	Supplied	Used
2008	11	5
2009	8	5
2010	30	19
2011	35	22
2012	19	13
2013	28	17
2014	43	28
Total	174	109

Table 5 shows a comparison of the 2014 mean probabilistic estimates uncertainty ranges, confidence intervals, and directional biases for each data set. The range of uncertainty decreases for each data set until the unitary distributions become more than half of the total distributions (historical data set 2012-2013). At this point, as previously mentioned, the probabilistic estimate loses accuracy as it begins to convert into a more traditional deterministic estimate. All of the distributions for the 2014 mean probabilistic estimates are overconfident (too narrow) and pessimistic DB, as the slope is less than 1 and the intercept is greater than zero, respectively, of the calibration chart for each data set (Table 5). The level of overconfidence is increasing for each data set (slope approaching 0) as the amount of unitary distributions is increasing. The pessimistic DB is also increasing for each data set until the data set of only- 2013 data. This supports the claim that the most optimal practice is using the most current historical data set, while maintaining enough data points to stay probabilistic

Table 5— Comparison of 2014 mean probabilistic estimates uncertainty ranges, confidence intervals, and directional bias.

2014 Estimation	Uncertainty (Million USD)		Confidence		Directional Bias		Unitary Dist.
	Independent	Dependent	Confidence	Slope	DB	Intercept	
2008-2013	24	39	Overconfident	0.82	Pessimistic	15.35	37
2010-2013	23	37	Overconfident	0.77	Pessimistic	19.46	38
2011-2013	22	35	Overconfident	0.66	Pessimistic	26.96	40
2012-2013	28	45	Overconfident	0.57	Pessimistic	43.57	59
2013	17	27	Overconfident	0.45	Pessimistic	24.11	76

For all data sets, excluding 2012-2013, optimistic estimates (underestimation of 2014 mean probabilistic estimates) are observed, however, calibration charts show pessimistic distributions. This is not a contradiction as the 2014 mean probabilistic estimate is less than the 2014 mean actual cost (optimistic estimate) for all data sets excluding 2012-2013 and the 2014 median probabilistic estimate is greater than the 2014 median actual cost (pessimistic estimate). Fig. 39 shows overlaid distributions of all of the 2014 mean probabilistic AFE estimates and all of the 2014 AFE actual costs for the 2011-2013 data set. This illustrates how mean estimates can be optimistic (probabilistic mean estimate less than mean actual cost) while calibration charts suggest pessimistic distributions (probabilistic median estimate greater than median actual cost). Fig. 39 shows one actual cost data point (approximately $13.5 million dollars) pulls the mean actual cost above the mean estimate. Removal of this data point would drop the mean actual cost below the mean estimate switching the mean optimistic estimate to pessimistic.

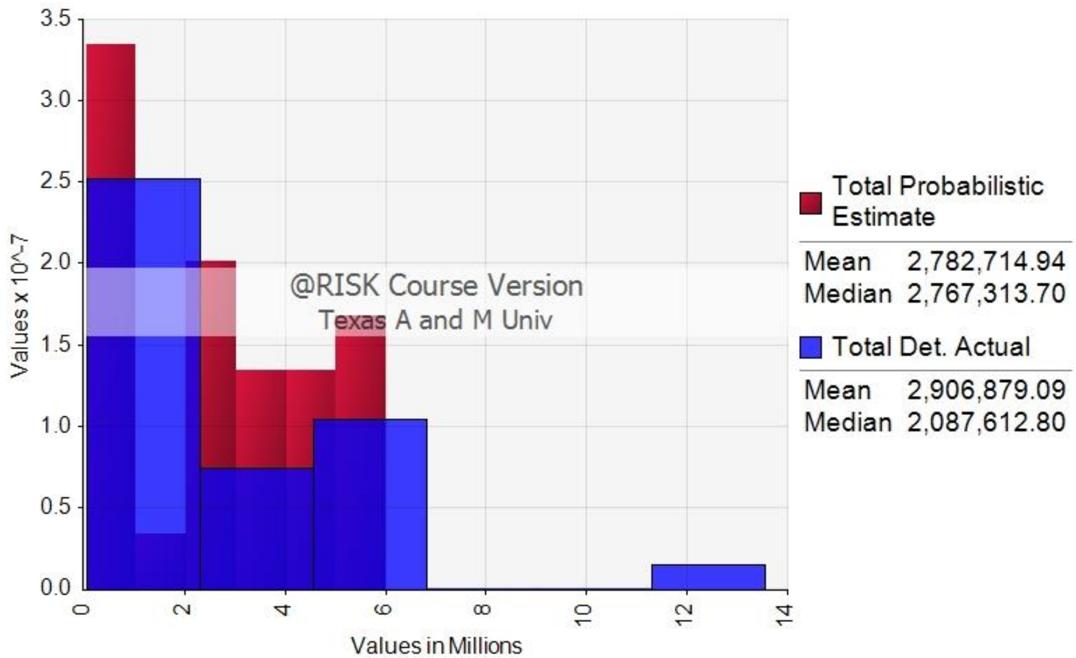

Fig. 39—2014 comparison of 2014 mean probabilistic estimates to 2014 actual costs for 2011-20013 historical data set shows how probabilistic estimates can be optimistic (Red mean<Blue Mean) and calibration charts suggest pessimistic distributions (Red median>Blue Median).

As all probabilistic estimate simulations for this study are more accurate than the deterministic, there is no minimum required amount of data points to implement this procedure. At the very minimum, cost codes with the highest amount of estimation uncertainty should be estimated probabilistically to obtain the most benefit for the least effort.   For ideal results, at least half of the cost codes should have enough data points to fit with distributions, making the cost model more probabilistic than deterministic. Following our conclusion of using the most current historical data set, while maintaining enough data points to stay probabilistic, the 2011-2013 data set would have been selected for use in the probabilistic cost model to more reliably estimate 2014 AFEs which would have been the most optimal data set selection.

4. CONCLUSION

The following conclusions have been drawn based on the results of the uncertainty quantification and conversion of deterministic AFE estimates to probabilistic estimates using the proposed probabilistic cost model workflow..

- The 2014 mean probabilistic estimates calculated through the proposed workflow, for all historical data sets, are more accurate than the deterministic estimates.
- The mean probabilistic estimates become more accurate as the years of the historical data set become more current, however the amount of unitary distributions increases. As the unitary distributions become more than half of the total distributions (as seen in historical data set 2012-2013), the probabilistic estimate loses accuracy as it begins to convert into a more traditional deterministic estimate.
- Using the most current historical data set, while maintaining enough data points to stay probabilistic is the most optimal estimating practice.
- Mean estimates can be optimistic (probabilistic mean estimate less than mean actual cost) while calibration charts suggest pessimistic distributions (probabilistic median estimate greater than median actual cost).
- Sub-costs of importance were determined based on highest range of uncertainty. These areas are clearly highlighted in Fig. 14 and include the cost code descriptions: Casing-Surf/Int Accessories, Line pipe/Flowlines, and Drilling Fluids.

- The miscoding of Casing-Surf/Int Accessories resulted in approximately $2.2 million worth of uncertainty in AFE estimations. Once input errors or miscoding has been eliminated, the uncertainty of other associated cost codes can be quantified allowing for better well cost estimations.
- While some were better than others, no company successfully considered uncertainty through the use of deterministic contingencies and all companies studied are capable of improvement with the inclusion of probabilistic cost estimating practices.
- The workflow further automates the correlation and truncation processes allowing for imperfect data sets and begins the calibration process.
- Reducing the number of cost codes and simplifying the definitions of each cost code should be done for better field estimating accuracy.
- Truncation sensitivity analysis is required to determine optimum required truncation and was determined to be 10% for this study.
- Probabilistic cost models can eliminate the necessity of deterministic contingency cost codes.

4.1. Recommendations for Future Work

This probabilistic cost model should be updated with the completed 2015 AFE historical data and should be applied to the 2016 deterministic estimates in order to more accurately estimate future AFEs. A completely automated method should be developed with a data cycling function that adds current data while discarding oldest data. This will keep the required amount of data points to keep estimates probabilistic while allowing the probabilistic cost model to remain current.

As the dependence of each AFE on each other is likely between the bounds of completely dependent and completely independent, another correlation matrix should be made on total AFEs to calculate and incorporate this dependence to create a better decision making tool.

The presented calibration charts show overconfidence and pessimism in the probabilistic cost model. The probabilistic cost model should be continuously calibrated to eliminate this bias in the estimation process. This will ensure that P10's are actually P10's, P50's are actually P50's, and P90's are actually P90's.

The effects of service costs associated with oil price could also be included in the model to incorporate fast increases or declines in prices, which will ultimately affect the model. Faster AFE reporting time will mitigate the impact of service cost changes when combined with the cycling function. This can be done through incorporating commercial AFE software compatibility for a completely automated process.

REFERENCES

Adams, A.J., Grundy, K.C., and Kelly, C.M. 2015. Probabilistic Well Time Estimation Revisited: Five Years On. *SPE Drilling & Completion* **25** (4): 472-499.SPE-173028-MS http://dx.doi.org/10.2118/119287-PA.

Amani, M. and Rostami, R. 2007. New Methodology for Afe Estimate and Risk Assessment: Reducing Drilling Risk in an Iranian Onshore Field. Paper presented at the Digital Energy Conference and Exhibition, Houston, Texas. SPE-107546-MS http://dx.doi.org/10.2118/107546-MS.

Brashear, J.P., Becker, A.B., and Faulder, D.D. 2001. Where Have All the Profits Gone? *Journal of Petroleum Technology* **53** (6): 20-73.SPE-73141-JPT http://dx.doi.org/10.2118/73141-JPT.

Capen, E.C. 1976. The Difficulty of Assessing Uncertainty (Includes Associated Papers 6422 and 6423 and 6424 and 6425). *Journal of Petroleum Technology* **28** (8): 843-850.SPE-5579-PA http://dx.doi.org/10.2118/5579-PA.

Hariharan, P.R., Judge, R.A., and Nguyen, D.M. 2006. The Use of Probabilistic Analysis for Estimation of Drilling Time and Costs When Evaluating Economic Benefits of New Technologies. Paper presented at the IADC/SPE Drilling Conference, Miami, Florida. SPE-98695-MS http://dx.doi.org/10.2118/98695-MS.

Kitchel, B.G., Moore, S.O., Banks, W.H. et al. 1997. Probabilistic Drilling Cost Estimating. *SPE Computer Applications* **9** (4): 121-125.SPE-35990-PA http://dx.doi.org/10.2118/35990-PA.

McVay, D.A. and Dossary, M.N. 2014. The Value of Assessing Uncertainty. *SPE Economics and Management* **6** (2): 100-110.SPE-160189-PA http://dx.doi.org/10.2118/160189-PA.

Nieto, J., Batlai, B., and Delbecq, F. 2013. Seismic Lithology Predicition: A Monteny Shale Gas Case Study. *CSEG Recorder*: 34-42

Peterson, S.K., de Wardt, J.P., and Murtha, J.A. 2005. Risk and Uncertainty Management-Best Practices and Misapplications for Cost and Schedule Estimates. Paper presented at the SPE Annual Technical Conference and Exhibition, Dallas, Texas. SPE-97269-MS http://dx.doi.org/10.2118/97269-MS.

Peterson, S.K., Murtha, J.A., and Schneider, F.F. 1993. Risk Analysis and Monte Carlo Simulation Applied to the Generation of Drilling Afe Estimates. Paper presented at the SPE annual Technical Conference and Exhibition, Houston, Texas. SPE-26339-MS http://dx.doi.org/10.2118/26339-MS.

Shilling, R.B. and Lowe, D.E. 1990. Systems for Automated Drilling Afe Cost Estimating and Tracking. Paper presented at the Petroleum Computer Conference, Denver, Colorado. SPE-20331-MS http://dx.doi.org/10.2118/20331-MS.

Valdes, A., McVay, D.A., and Noynaert, S.F. 2013. Uncertainty Quantification Improves Well Construction Cost Estimation in Unconventional Reservoirs. Paper presented at the SPE Unconventional Resources Conference, Alberta, Canada. SPE-167139-MS http://dx.doi.org/10.2118/167139-MS.

Williamson, H.S., Sawaryn, S.J., and Morrison, J.W. 2006. Monte Carlo Techniques Applied to Well Forecasting: Some Pitfalls. *SPE Drilling & Completion* **21** (3): 216-227.SPE-89984-PA http://dx.doi.org/10.2118/89984-PA.

APPENDIX A

Intangible Drilling Costs	400	Non-op & accruals
	401	Site preparation/survey/roads
	402	Company labor
	403	Rig move
	405	Camp/crew
	406	Drilling - daywork
	407	Contract/consultant
	409	Drill bits
	411	Utilities & power
	412	Fuel
	413	Boiler/steamer
	414	Directional services
	415	Mud/chemicals
	416	Mud logging
	417	Centrifuge
	418	Water access/hauling
	419	Materials/supplies
	420	Vac truck
	421	Power tongs
	422	Equipment rental
	423	Casing-surface
	424	Casing-production
	425	Equipment inspection/testing
	426	Casing - surf/int accessories
	427	Cementing - surf/int casing
	428	Safety/first aid
	429	Fishing
	431	Drill string
	432	Welding
	435	Drill stem test
	437	Air drilling
	439	Coring/core analysis
	440	Trucking/hauling - other
	441	Logging
	443	Abandment/plugback
	446	Casing - production accessories
	447	Cementing - production casing

	450	Wellhead equipment
	465	Drilling fluids
	468	Clean up/disposal
	469	Site restoration
	470	Property tax
	472	Licenses/permits
	474	Communications
	483	Engineering
	484	Supervision
	485	Environmental
	489	Insurance
	494	Miscellaneous
	495	Contingency
	498	Overhead
Intangible Completion Costs	500	Non-op & accruals
	501	Location/roads
	502	Company labour
	503	Rig move
	505	Camp/crew
	506	Drilling rig
	507	Contract/consultant
	509	Drill bits
	511	Utilities & power
	512	Fuel
	513	Boiler/steamer
	515	Mud/chemicals
	518	Water access/hauling
	519	Materials/supplies
	520	Vac truck
	522	Equipment rental
	523	Equipment - intangible downhole
	525	Equipment inspection/testing
	528	Safety/first-aid
	529	Fishing
	531	Drill/work string
	532	Welding
	534	Pressure surveys
	537	Air drilling

	540	Trucking/hauling - other
	541	Logging
	543	Abandonment/plugback
	545	Tubing/string/accessories
	546	Casing - production accessories
	548	Cementing
	549	Perforating
	550	Wellhead equipment
	551	Fracturing
	555	Wireline
	561	Service rig
	562	Coiled tubing/snubbing/n2 clean out
	563	Production testing & analysis
	565	Completion/workover fluid
	566	Treating/stimulation/acidizing
	567	Swabbing
	568	Clean up/disposal
	569	Site restoration
	570	Property tax
	572	Licenses/permits
	574	Communications
	583	Engineering
	584	Supervision
	585	Environmental
	586	Load oil
	589	Insurance
	594	Miscellaneous
	595	Contingency
	598	Overhead
Plant & Battery Equipment	615	Mud/chemicals
	626	Casing - surface
	627	Casing - production