SPE-180710-MS

Three Approaches to Predicting ESP Pump Failures during SAGD Operations

Leon Fedenczuk Gambit Consulting, John Graham Suncor Energy, Trent Pehlke Suncor Energy

This paper was prepared for presentation at the SPE Canada Heavy Oil Technical Conference held in Calgary, Alberta, Canada, 7–9 June 2016.

Abstract

This paper presents a novel approach to a time scale discretization when predicting ESP pump failures at different scales. This study proves that models can be used to formalize failure predictions, prevention, and lead to optimizing the ESP's replacement and/or maintenance. The target parameters reflected two different time scale ranges. In the first approach 'Time to Failure' and its corresponding 'Active Time to Failure' were predicted. The second case excluded time periods when a well was off-line for other reasons than failure. These two targets (modeling parameters) represented low frequency events and were developed using geological or/and well geometry parameters. The Total Time to Failure model (Production Period model) based on a combined trajectory and geology data set showed acceptable and stable performance. A corresponding model with wellhead parameters summarized across each production period was introduced to complement the large scale analysis.

A second group of models of higher resolution was designed to detect failures in real time. In these cases estimations for the probability of a failure at a specific time using the most recent wellhead data while excluding well's non-active time periods related to workovers and other non-productive time periods.

These models used pre-processed wellhead data from a few selected wells and pads. Well data required pooling large amounts of data and developing a parameter summarization in time periods based on uninterrupted Motor Current Time Periods. These discrete time periods represented events with or without a failure depending on a reason for the current value to be zero. The probability of a pump failure was estimated using two approaches. In the first approach only the last two 'periods' that corresponded to non-failure and failure periods respectively were used. The second approach involved all non-failure periods leading to each corresponding failure period. The first approach overestimated the failures while the second approach overestimated the non-failure events.

Initial probability models predicted events with a relatively high success rate. However, more data and additional data transformations are required to verify the practicality of our approach. More refined sub-

period estimates in each Current Time Periods may help in developing improved models.

Introduction

Pump failures are expensive and predicting these events can reduce lost revenue and help in optimization of pump replacement, stopping and/or alerting operations to prevent catastrophic events.

Estimating mean failure rate values and fitting them into known distributions is a common practice in process improvements[1]. However, SAGD processes have been evolving very rapidly and it is difficult to get a good sample size of the same pump type, size, and working conditions to fit such distributions. Furthermore, making decisions to replace a pump when it seems to work well is a difficult decision that has to be supported by reliable models. Data driven analytics is emerging for monitoring and preventing problems with ESP pumps in the oil industry [2,3]. What has been missing given the number of SAGD operations is analyzing the contribution of other factors to pump failures. The goal of this study was to test how geology, well geometry, and well head parameters can help in predicing pump failures. Like any analytical project we had to acquire information about pump failures in a specific time period and all associated data: well head parameters, well geology description, and well completion information. All data sets were tested, cleaned, pre-processed, normalized, and analyzed in a recursive process, which is characteristic of data driven projects [4,5]. Major analytical tasks included:

- Developing metrics associated with ESP pump failures.
- Identifying TARGET parameters (to be predicted in order to estimate failures and help in detection).
- Identifying the most important geological and well trajectory parameters correlated to pump failures.
- Identifying the combination of wellhead measurements in time supporting predicting/avoidance of pump failures and pump replacement planning.

Failures by Well – Large Scale Analysis with Geology and Geometry

Several metrics associated with days between pump failures have been developed and tested. Eight pad data from Firebag between 2008 to 2013 was used. Failures for all wells were considered for modeling the expected days to failure based on geometry and geology associated with each well.
We considered several alterative target parameters or variations of the ‘days to failure’. Specifically, we estimated a weighted mean or median days and weighted harmonic mean or median days to failure (see sample of failure statistics in shown in Table 1 APPENIDX 2). The weighting was done by an associated pump size (Median catalog derived pump rate). The failure definition/flag was based on DIFA reports.

We built multivariate regression models that predicted the median number of days to pump failure using geology and/or trajectory based metrics. Multi-step model development analyzed the contributions and significance of variables and model strength for each set of final variables. This led to robust models with only the most significant parameters being utilized. A trajectory based model was stronger than geology based model. These models were significant but weak and characterized by R-square = 0.36 and 0.25 respectively.
A final model based on selected variables from both sets of parameters (geometry and geology) showed improved performance with R-square = 0.41 as shown in Table 2. This number was strong enough to

show that there were significant parameters associated with higher frequency of failures. The model presented here was based on one geometry variable and three geological parameters. All the explanatory estimates were significant and stable. A higher number of days to failure can be expected for:

- An injector with higher TVDSS variation
- A higher % of F2 channel
- A higher Continuous Gross Pore Volume (m3)
- A lower Continuous Gross Porosity (%)

The last two explanatory variables represent a combination of two measurements representing the same factor. This indicates that different formulas and combinations/ratios of explanatory parameters could improve future model performances. This and the following parts of the paper show the importance of the data processing steps.

Failures – Large Scale Analysis with Wellhead Data

Next a set of regression models for **Total Time to Failure** and **Hours-On to Failure** using only well-head parameters were built. The second target (hours-on) parameter excluded time periods from the well history when the pump was not operational for any reason. These models used heavily pre-processed wellhead (TAG) data. At this stage data was pooled for selected wells due to the large amount of wellhead data that had to be processed.

Specifically, time interval discretization and parameter integration in these time intervals was introduced. Thus, we compressed each time period during which a pump worked continuously into a set of estimates. The new data contained a sequential period number and summarized estimates of all parameters during each time period. We utilized two definitions. The first is a standard definition during SAGD operation and the second is our definition of the discretized time period.

- **Production Period is** a continuous time period between the pump installation date/time and its failure date/time.
- **Motor Current Period** is a continuous time period with all values of the motor current greater than zero. A new Motor Current Period starts after each pump motor restart.

Several new parameters were defined for wellhead measurements that described count, max, min, mean, median, and variation statistics during each discrete Motor Current Period. A sample of these parameters is presented in Table 3.

The first **Time to Failure** model used only three metrics: maximum value of Injector Casing Percent Open, maximum of Producer Pressure Casing at wellhead, and minimum of Producer Temperature Casing at wellhead.

A second version of the **Hours-On to Failure** regression model included more exotic metrics based on hours-on during motor current periods. Two metrics INJ_TBG_PCT_MAX (max of INJ_TBG_PCT during the whole Production Period) and PROD_T_TBG_MEAN_MAX (max value of the mean values of PROD_T_TBG that was estimated during each Motor Current Period) were used to build a final regression model. It had an R-Square = 0.75 and parameter estimates are presented in Table 4.

A corresponding decision tree model was characterized by R-Square = 0.82 and shown in Figure 1. The root split (top) was based on Inj_Tbg_Pct_Max, which was present in two-variable regression Hours-On model. The second level splits (in branches below) were based on Prod_Csg_Pct_Mean_Median and Stm_Tbg_StdDev_Mean. This model indicates that Prod_Csg_Pct_Min_Median results in higher Hours-On to failure. In the second branch with a three node split low and high variability in Stm_TBG result in lower Hours-On to Failure. Mid range variability may indicate a situation when the steam chamber is being better controlled, which would result in better pump life. The model diagnostics confirm that building predictive models for the pumps expected life is possible. However, we did not have enough data to perform enough model 'stress testing'. For example we did not find any difference based on a pump size.

Probability of Failures – Detailed Scale Analysis with Wellhead Data in the Last Two Current Periods

Two models were developed that predicted the actual probability of failure in the last motor current period when the actual event took place. As it was described before each production period (time from the pump installation to the corresponding failure) was divided into the motor current periods of continuous electric motor readings (Period of Motor Current >0). That represented predicting failures as they were happening.

Our initial model predicted a failure in the last motor current period and it used average well head parameters in the last two motor current periods. Thus, we looked at the wellhead parameter summarizes in the period when the failure happened to get the failure metrics. The non-failure metrics were obtained from the previous period when pump did not fail. The logistic model was developed to detect the impeding failure based on differences of these parameters during these two time periods.

Due to limited data pool the above target definition of the failure and the non-failure event had 42 motor-current period pairs with the failure and the non-failure events (Y/N). The corresponding logistic regression model was characterized by a low R-Square = 0.21 and Max-Rescaled R-Square = 0.28. At the same time nearly 10-12 parameters competed for the final model but none of the most important variable sets could support good models.

Table 5 presents a corresponding frequency statistics with the correct and incorrect model classification. Specifically, it shows a frequency table that shows observed levels (_FROM_) of the target parameter (Failed=N/Y) versus the predicted levels (Into_P) and corresponding metrics (Frequency, Percent, Row Pct, and Column Pct).

Only 24 cases of NO Failure were properly classified, which is 57.14% of all cases without failure. The other 18 cases were classified as failures. We had more luck with the failure rate (2^{nd} row of the table). 30 cases of failure were properly classified with 71.43% success rate.
A high number of misclassifications and especially false positive outcomes (non-failure misclassified as failure) show that the model was too sensitive. This was due to unknown pump age and equal number of No-Fault (N) and Fault (Y) events (heavily weighting Y events).

Probability of Failures – Detailed Scale Analysis with Wellhead Data and All Current Periods

Our final model used estimated well-head parameters in all current periods during the production period leading to failure. It was built with all Non-Failure events in a series of the current periods leading to a specific failure. Thus, a 'failed' parameter/indicator was created and it was set to 'Y' only during the last period (29 cases) and 'N' in all previous motor current periods without failures (891 cases). These two counts changed with iterations and depended on specific data pre-processing and number of variables in the logistic regression model. During this modeling stage we included the motor age at the beginning of each motor current period.

A preliminary logistic model had good R-Square = 0.74. Some of the most important parameters included: Age, Annular Gas Flow, Frequency, Motor Current, casing and tubing parameters. In the modeling process we discovered that too many parameters with significant statistics (Pr > ChiSq <0.05) showed up on the most important list of parameters. The same phenomenon characterized the previously presented model based only on the last two current periods. Too many variables typically result in model instabilities due to multivariate correlations. Furthermore, in this case we are modeling a rare case scenario because the modeled target level occurs approximately 3% of time.

Thus, the model performance (in Table 6) was probably too optimistic and biased because the data used to fit the model was also used in the assessment of the model. We tried two ways of dealing with this problem. In the first we split the available data into training and validation data sets. We trained the model using the training data set and then assessed by applying the model to the validation data set.
However, our available data set had very few 'Failed' cases (29 out of 920) and splitting data into two sets resulted in an unacceptably small failure count in both sets. In the second approach we applied cross-validation process, which could provide an unbiased assessment of the model without reducing the training data set. In both cases we did not get much better results.

Both modeling cases showed that well-head parameters in time periods leading to failures support early failure detection. More data should help in developing more robust models and more work had to be done with the data transformation/normalization, and variables elimination. In data driven modeling it is essential to create as many physics based parameters in the initial stages of the project. Some of these variables could be based on combinations of monitored variables as it was shown when modeling the intake temperature during SAGD [6]. Advanced metrics could be based on a combination of motor current periods and fitting 'known' shapes (shapelets) into these series [7]. In the next iteration we are planning to include parameters associated with quality of power being supplied to a pump.

Potential Implementation

The above study indicates that one model could not satisfy our needs. Different target parameter definitions resulted in developing models with different detection sensitivity and at different time resolution scale.
Thus, the most likely solution should be based on a model predicting:

- Expected Hours to Failure
- Oversensitive model for Probability of Failure based on the last few time periods associated with the motor current
- Probability of Failure based on all motor current periods from the past history (less sensitive model).

Their weighted contributions could be used to estimate a final pump wellness indicator. The first oversensitive model prone to the false positive detection would be used for flagging potential problems, while the second model would identify the most likely failure candidates. The above models can be implemented as pop-up solution in the currently available field applications to display expected 'Time to Failure' and a status for the Failure Probability.

Summary

This study was intended as a proof of concept for linking the ESP pump failure to geology and well geometry. Furthermore, we proved that wellhead parameters could be used to predict pump failures with enough time to set a warning or stopping the pump prior to the event.
In order to identify parameters that could help in pump failure detection we introduced parameter discretization based on the motor current.

Implemented data pre-processing and definitions proved that more than one model type was needed to address all aspects of predicting pump failures.

A large scale model could be build for predicting Time to Failure. A more detailed scale could be represented by Probability of Failure models. Different target definitions and advanced data pre-processing lead to models of variable sensitivity resulting in too many false detections or too few true detections. In our opinion, we would have to implement these different model types and use their predictions in a combined fashion.

References

1. Anwar Khalil Sheikh, Muhammad Younas, Dahham Matar Al-Anazi, Weibull Analysis of Time Between Failures of Pumps Used in an Oil Refinery, The 7th Saudi Engineering Conference, KFUPM, Dharan, December 2002.
2. Gupriya Gupta, Michael Nikolaou, Luigi Saputelli3 and Shyam Panjwani, Applying Predictive Analytics to Detect and Diagnose Impending Problems in Electric Submersible Pumps Used for Lifting Oil from Wellbores, 15 AIChE Annual Meeting, Nov. 8-13, 2015
3. Joe Vandevier, ESP-Conclusion: Multiple factors affect electrical submersible pump run life, Oil&Gas Journal, 11/01/2010.
4. John C. Davis, Statistics and Data Analysis in Geology, John Wiley & Sons 1986.
5. L. Fedenczuk, K. Hoffman, and T. Fedenczuk, Predicting Waterflood Responses with Decision Trees, Journal of Canadian Petroleum Technology, No.5, Vol.46, 2007.
6. Leon Fedenczuk, John Graham, Trent Pehlke , and Nicholas Cherniwchan, "Predicting SAGD ESP Intake Temperature" SPE-180710-MS, SPE Canada Heavy Oil Conference to be held 7-9 June, 2016 in Calgary, Alberta, Canada.
7. Josif Grabocka, Nicolas Schilling, Martin Wistuba, Lars Schmidt-Thieme, Learning Time-Series Shapelets, SPE-170680-MS.

Appendix 1 – Figures

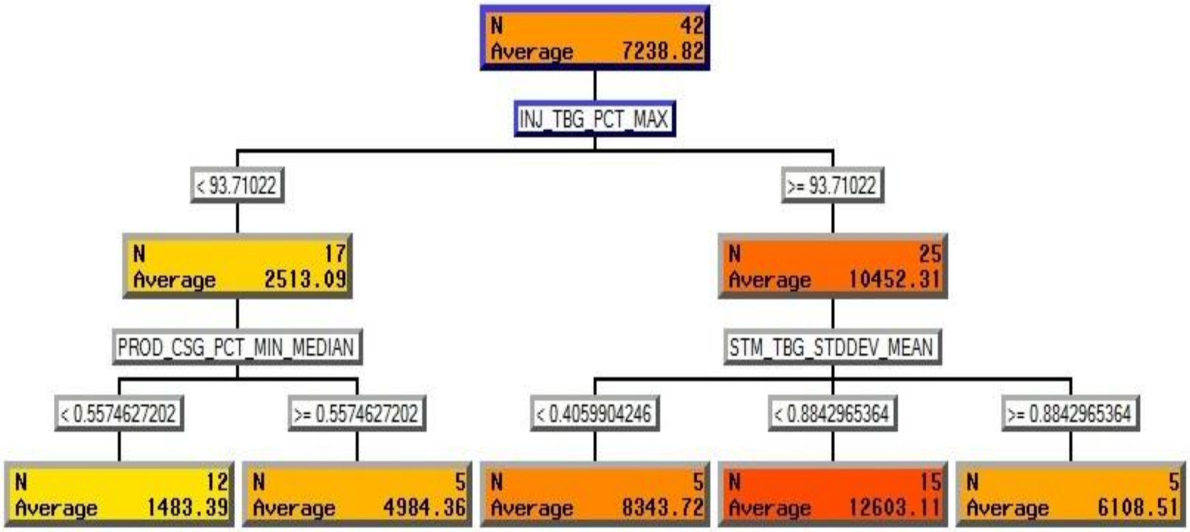

Figure 1: Decision Tree for Hours-On Time to Failure (R-Square = 0.82)

Appendix 2 – Tables

1	Well_Name II	Date_Well_Started	Date_Well_Last_Event	Event_Count	Mean_Duration_Days	Min_Duration_Days	Max_Duration_Days	Pump Weighted Mean Duration Days	Pump Weighted Har-Mean Duration Days	Median Duration Days	Event_Annual_Frequency	Well_Age
2	P2P4	20-Oct-05	4-Dec-13	4	295	21	547	272	63	305	0.49	8.13
3	P1P9	9-Nov-05	17-May-11	3	663	467	856	663	624	666	0.54	5.52
4	P2P2	14-Mar-07	25-Aug-11	3	399	130	561	399	262	505	0.67	4.45
5	P2P5	16-Oct-05	21-Oct-12	5	367	67	570	409	287	403	0.71	7.02
6	P2P10	25-Apr-09	2-Feb-12	2	484	474	493	482	482	484	0.72	2.78
7	P1P7	11-Mar-07	13-Nov-13	5	375	181	949	375	267	256	0.75	6.68
8	P1P4	13-Apr-09	20-Dec-12	3	240	114	410	240	184	197	0.81	3.69
9	P1P3	3-Mar-09	25-Aug-13	4	403	9	1,013	477	38	295	0.89	4.48
10	P1P1	9-Dec-05	9-Aug-13	7	391	137	748	405	303	344	0.91	7.67
11	P1P10	24-Nov-05	2-Mar-12	7	312	1	668	327	9	235	1.12	6.27
12	P2P8	29-Jun-06	6-Jun-13	8	292	27	960	325	88	127	1.15	6.94
13	P2P1	21-Jun-05	19-Aug-10	6	180	32	656	133	69	106	1.16	5.16
14	P2P7	18-Oct-05	16-Oct-11	7	303	87	703	282	195	224	1.17	6.00
15	P1P8	11-Mar-08	27-Jan-13	6	292	29	724	292	82	260	1.23	4.88
16	P1P5	29-Mar-09	3-Jan-14	6	279	12	591	278	43	290	1.26	4.77
17	P1P2	15-Nov-05	18-Sep-13	10	281	9	601	286	44	301	1.27	7.85
18	P1P6	17-Jul-06	18-Jul-13	9	258	9	539	269	75	258	1.28	7.01
19	P2P9	21-Mar-09	28-Mar-12	4	224	4	644	171	11	124	1.32	3.02
20	P2P6	13-Oct-05	23-Mar-10	7	190	29	612	222	70	95	1.58	4.44
21	P2P3	6-Jul-06	25-May-13	12	163	3	821	173	25	45	1.74	6.89

Table 1: Example of failure statistics for two pads associated with Y/N true ESP failure flag.

Parameter Estimates						
Variable	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	902.6043	374.429	2.41	0.019	0	0
Std_TVDSS_I	152.0854	59.1681	2.57	0.0127	0.2892	1.2973
Channel_F2_P	1611.452	633.094	2.55	0.0135	0.2550	1.0287
Cont_Gross2	0.000152	4.8E-05	3.18	0.0023	0.3660	1.353
Cont_Gross_P	-2952.15	1248.17	-2.37	0.0213	-0.269	1.330

Table 2: Model statistics for Days to Failure derived from well's geometry and geology (R-Square = 0.41, Adjusted R-Square = 0.38)

Inj_Csg_Pct_Max	Maximum value of Injector Casing Percent Open Choke (Inj_Csg_Pct)
Prod_P_csg_Max	Maximum of Producer Pressure Casing at wellhead (Prod_P_csg)
Prod_T_Csg_Min	Minimum of Producer Temperature casing at wellhead (Prod_T_Csg)
Inj_Tbg_Pct_Max	Maximum value of Injector Tubing Percent Open Choke (Inj_Tbg_Pct)
Prod_P_Tbg_Max	Maximum of Producer Pressure Tubing at wellhead (Prod_P_Tbg)
Prod_T_Tbg_Min	Minimum of Producer Temperature tubing at wellhead (Prod_T_Tbg)
Prod_T_Tbg_Mean_Max	Max of Mean T_Tbg by Current Periods during the specific Production Period
Prod_Csg_Pct_Mean_Median	Median of Mean Producer Casing Percent Open Choke by Motor Current Period
Stm_Tbg_StdDev_Mean	Mean of Standard Deviation for Tubing Steam by Motor Current Period

Table 3: Important derived metrics (statistics) used in the modeling process included.

Parameter Estimates							
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	1	-66553	9568.43	-6.96	<.0001	.	0
Inj_Tbg_Pct_Max	1	130.46	23.9061	5.46	<.0001	0.8273	1.2087
Prod_T_tbg_Mean_Max	1	336.94	55.34	6.09	<.0001	0.8273	1.2087

Table 4: Model parameters for Hours-On to Failure derived from wellhead data (R-Square = 0.75, Adjusted R-Square = 0.73)

Frequency Percent Row Pct Col Pct	Table of _FROM_ by Into_P			
	From Observed Response)	Into_P		
		N	Y	Total
	N	24	18	42
		28.57	21.43	50
		57.14	42.86	
		66.67	37.5	
	Y	12	30	42
		14.29	35.71	50
		28.57	71.43	
		33.33	62.5	
	Total	36	48	84
		42.86	57.14	100

Table 5: Classification table for Failure weighted model.

Frequency Percent Row Pct Col Pct	Table of _FROM_ by Into_P			
	From Observed Response)	Into_P		
		N	Y	Total
	N	890	1	891
		96.74	0.11	96.85
		99.89	0.11	
		99.55	3.85	
	Y	4	25	29
		0.43	2.72	3.15
		13.79	86.21	
		0.45	96.15	
	Total	894	26	920
		97.17	2.83	100

Table 6: Classification table for Non-Failure weighted model**.**