Individual Research Project: Preliminary Report

Deirdree A. Polak

June 2019

1 Project Introduction

The proposed project title is

"Automated Slugging Detection using Machine Learning"

Company Supervisors
Dr Meindert Dillen & Peter Kronberger
Wintershall Dea GmbH
meindert.dillen@wintershalldea.com & peter.kronberger@wintershalldea.com

Academic Supervisors
Prof Olivier Dubrule & Lukas Mosser
Imperial College London
olivier.dubrule@imperial.ac.uk & lukas.mosser15@imperial.ac.uk

This project is undertaken as part of an internship at Wintershall Dea GmbH in Kassel, Germany and Bergen, Norway.

1.1 Project Motivation

As part of Wintershall Dea's digitalization strategy to automate and improve company processes, they are developing a digital twin of the BRAGE oilfield in the North Sea. One of the initiative projects is the development of an automated slug detection process.

A slug flow in a well is characterized by a mass of liquid flowing to the surface, followed by large gas pockets. The uneven flow content causes instabilities in the production facility, due to the fluid separator not being able to handle outbursts of gas. The large gas outflow in the shared oilfield separator and compressors reduces the system's capacity.

An automated slug detector should be able to predict the occurrence of slugs allowing operators to take immediate action, therefore reducing production costs to the company. The goal would be to be able to proactively predict the well's pressure behaviour as a slug comes to the surface. This would allow production engineers and operators to make decisions in real time to minimise costs and optimise production.

Two wells within the BRAGE will be considered in the scope of the project, A-19 and A-40, both of which have frequently been affected by slugging. Well A-19 presents rare slugging characteristics, whilst the A-40 has strong intermittent slugging. For this reason, the A-40 has been closed for approximately a year at great expense to Wintershall.

1.2 Project Objectives

The objectives of this project are as follows:

1. Collect relevant well data for slugging detection
2. Develop slug indicator label based on the historical data available.
3. Research and Compare different machine learning models, primarily based on supervised learning using the slug indicator.
4. Predict slugging events using machine learning based on historical data of two wells: A40 (frequent slugging) and A19 (lighter slugging). Train model and validate with testing data.
5. Predict pressure and other parameters trends using machine learning. This will be based on historical data and modelled data of the A40 and A19 wells. Train and validate model.

Furthermore to increase the value of the automated slugging detection for Wintershall Dea, another objective should be to predict the pressure and other parameters if slugging control is applied. Slug control is performed when the operators try to control the slug through closing the production choke.

1.3 Project Data

The data available for the project dates 10 years back, with one data point every minute. This totals to 5.3 million data points (10 years x 365 days x 1440 minutes). The data features available from the oilfields are:

- Pressure and temperature at the well head in Barg and Celsius Degrees
- Pressure and temperature at the bottom hole of the well Barg and Celsius Degrees
- Production Choke as a percentage
- Gas-Lift Choke(%), rate(Sm3/HR) and injection pressure (BarG)

2 Literature Review

A through understanding of platform operations and time series is required prior to starting modelling. In this section, a summary of the relevant practices in both fields is outlined. Moreover, an overview of the work previously performed by Wintershall Dea GmbH on the slugging issue is summarized.

2.1 Slugging

A slug flow is a multiphase-fluid flow characterized by large amounts of liquids followed by pockets of gas. The flow is shown in the second diagram in the Figure 1.

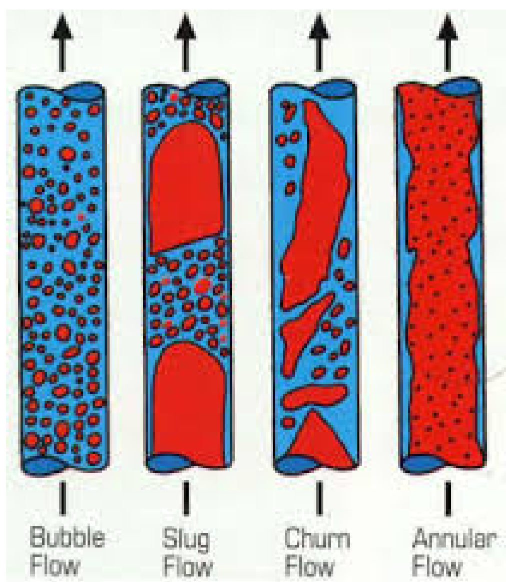

Figure 1: Different Flow Regimes in Vertical Pipes, Red is Gas, Blue is Liquid

2.1.1 Slug Types

There are three main types of slugging that can occur in a well or production [12]:

- Terrain Slugs occur at the base of a sharp incline in the well where the liquids accumulate and the pressure is not sufficient to lift it
- Turn-up Slugs occur when there is an increase in fluid rate through the system
- Hydrodynamic Slugs are caused by the shear stresses between the liquids and gas

In the context of this project, only terrain slugs will be considered. As outlined in the Previous Work section, the geometry of the wells were found to be the main cause of slugs within the oilfield.

Slug flow is to be avoided during oil production as the process facilities can't handle the large chunks of gas. It saturates the separator, and conveys the excess gas to the flare. This has poor environmental and economical consequences.

2.1.2 Slugs Indicator

Slugging is typically indicated in the data through the pressure and temperature at the bottom-hole of the well and at the well-head. There is to date no actual variable in the data that indicates a slug, so far operators and engineers have been able to identify them through visual inspection only. This highlights the need to create a slug label for supervised machine learning applications. Having this standardised prediction will be more objective than the current engineer's classification.

The graph in Figure 2 shows the well head pressure and bottom hole pressure in well A-40 over two years. The sharp increase in pressure at the well head are characteristics of slugging. The bottom hole pressure is more stable, but responds to slugs by spiking.

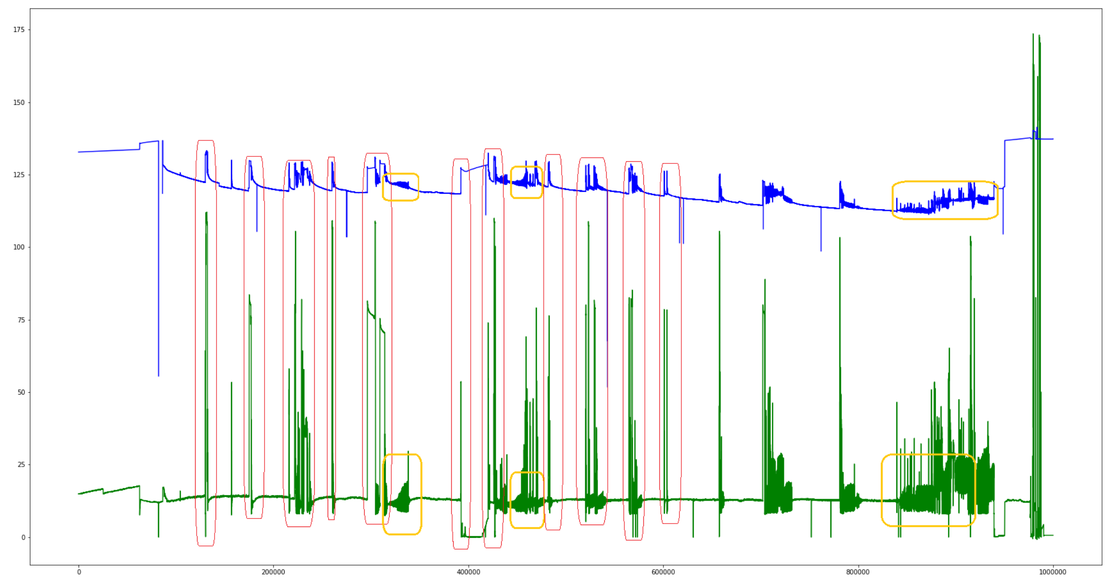

Figure 2: Pressure Fluctuations at Slug, Green is well head pressure, Blue is bottom hole pressure. The slugs are indicated in red. The small yellow boxes indicate smaller scale slugs that will be ignored in the scope of this project. The pressure is in BarG on the y-axis and the time is in minutes on the x-axis.

The liquid (oil and water) and gas flow rate would be the best indicator, however the values are only estimated at the well head. A flow-meter that detects both liquids and gas is too costly to be fitted at each well. Therefore there are no accurate values for the slug flow. The only accurate recorded data available for the flow of one well is when the well's production is rerouted to the test separator.

A separator in oil production is the first stage of the processing manifold in offshore production. The separator takes in the fluids of all the wells in the production oil field. It's a pressurized vessel that segregates oil, gas and water of the incoming flow. At this stage, it is possible to calculate accurately the full flow of all the connected wells. It gives accurate outflow values for the whole oilfield [5].

A test separator is the same, except it only takes the inflow of one single well. This is done to get an accurate measure of the outflow of each well. This therefore gives accurate view of when slugs happen in one specific well's production. For the A40 well for example, the outflow was rerouted once in 2011. This flow information could then tell of the liquid and gas inflow and therefore if labelled for slugging validate whether the slug indicator developed from the pressure and temperature data is accurate.

The slug indicator can either

- be developed manually, through rolling mean of the well head pressure, and then used for supervised learning of slug prediction. This is the preferred option, as the data is noisy and it gives more control
- be developed through unsupervised learning. This will be discussed further in the anomaly detection section of this report.

2.2 Previous work

Within Wintershall Dea GmbH, previous work has been commanded to understand the root cause of the slugging flow in well A40. Part of the task was to investigate the effect of slug control in the well. Three separate consultancies reported on it and here are their findings.

2.2.1 Consultancy A

In November 2017, the first consultancy delivered a report that outlined ways of predicting slugs and offered ways to control them using the production choke. Using an XGBoost model with 400 random trees they modelled the output flow with no time dependencies. They created synthetic variables based on the cycles of the slug occurrences in well A-40. These variables are outlined in Figure 3. The model is only dependant on these parameters and is not treated as a time series, In this model, the event parameter to predict was set as the rolling mean of the absolute value of the well head pressure differential shifted 20 minutes into the future.

Whilst this is an interesting approach that yielded promising results to predict slugs, it is too dependant on the periodicity of the A-40 well slugs. This model would not be applicable to other wells and to slug control.

Synthetic Variable	Description
time_snc_valley2	Time since oil export was lower than 50 Sm3/h
time_snc_valley	Time since oil export was lower than 90 Sm3/h
time_snc_peak2	Time since oil export was higher tan 125 Sm3/h
time_snc_peak	Time since oil export was higher tan 90 Sm3/h
valley_duration2	Total time during which oil export is lower than 50 Sm3/h
valley_duration	Total time during which oil export is lower than 90 Sm3/h
peak2_duration	Total time during which oil export is greater than 125 Sm3/h
peak_duration	Total time during which oil export is greater than 90 Sm3/h
mnth	Month
BRA_PZT_13_404_sd	Past Rolling mean of the standard deviation of the wellhead pressure in a40

Figure 3: Synthetic Variables created to train Slugging Model by Consultancy A

2.2.2 Consultancy B

Consultancy B delivered a thorough report on the root cause of slugs in well A40. The report was produced in March 2019. The main take-away from this report relevant to the current project is that the slugging is caused by the geometry of the well.

The report also highlights that the pressure fluctuations due to slugging in A40 are reputed through the production system, at compressor stages 1 to 4. Stage 4 is where the fluctuations are at the highest. This highlights the harmful impact of slugging in one well of the oilfield, as it not only impacts the well, but also the production line along the way.

The report also researched ways to mitigate the slugs using gas lift and choking. Under simulations, the well could not be stabilised fully, and slugs still occured in the model of the A-40 . The stabilization was optimised using different choking settings. This shows the importance of further improving the model and integrate slug control through choke setting.

Moreover, the report concentrates on the periodicity of well A40's slugging. It finds that the slugging occurs in 6-minute long period, which is also observed in the separator. This information can be useful for modelling, as a different strategy could be used due to regular cycles. On the other hand, for well A19 slugging detection will most likely take the form of anomaly detection.

2.2.3 Consultancy C

Consultancy C is concurrently working on a similar project. Their approach to predicting slug using machine learning is to use time series modelling such as Echo State Network and Long Short Term memory networks. They are using the same slug indicator as Consultancy A. This already paves the road to using Recurrent Neural Networks for modelling the pressure.

2.3 Machine Learning

Whilst the project is still at an early stage, there are already a number of machine learning practices that can be considered. A short literature review of the current flow detection through machine learning is performed, however the two main areas of interest for this project will be Time Series Modelling and Anomaly Detection.

2.3.1 Multi-phase Flow Pattern Detection

Previous research has been performed to categorize flow patterns in horizontal and vertical pipes. In recent research, Guillen-Rondon et al [6] achieved 97% correct classification of three different types of flow: Intermittent Flow (like slugs), Dispersed flows (for example with bubbles) and Segregated flows (such as stratified or annular flows). The method used to get such a high accuracy was supervised classification using Support Vector Machine (SVM). SVM in this case ignores the time-series aspect of the data: it does not predict a slug, but recognizes a slugging flow as it occurs. The data used for this research was larger than that available for the project and included velocity of the gas and liquids, viscosity, density, surface tension, inclination angle and pipe diameters.

Further research outlies Artificial Neural Network (ANN) as tools to determine flow patterns. In particular, E.S. Rosa et al's [9] research recommends using Multi-layer Perceptron networks and Probabilistic Neural networks to in categorize the flows. Similarly, M. Al-Naser et al [2] uses a range of Feed Forward neural network.

Overall, the data available for the research described above is more precise and analytic than what is available on offshore wells. The data used is not time dependant, rather it's instant data taken instantly from different types of flow. It is therefore not directly applicable to this project, in terms of the machine learning techniques used.

2.3.2 Times Series Classification

In this section, different machine learning models are discussed. The different models will be compared on the data at a later stage. These methods can both classify the event (slug or no slug), and also predict performance in the future. A combination of the models described below will probably be used.

Regression

Support Vector Machines (SVM) ignore the temporal element the data set put provide a strong classification system. It's a binary classification system which looks for the hyperplane with the largest margin in the feature space. Data points on either side of this hyperplane are attributed to different classes. This model can be non linear by adding a kernel function.

Logistics regression works by attributing a data point a probability of belonging in either of the classes. This could be an beneficial trait in predicting slugs, as it would give engineers the likelihood of the slug, and therefore an indication of the urgency.

Decision Trees and Random Forests

Decision Trees and Random Forests models are common black box models to classify time series for supervised learning. The model takes in the set of data points, and builds a random decision tree to extract the relevant features. The more data points, the more robust the decision tree will be. With the random forest approach ensures there is no over-fitting.

Recurrent Neural Network (RNN)

RNNs are the state-of-the art approach for time series modelling. They differ from feed forward neural networks by having a recurring hidden layer that loops over itself at each iteration to learn from the previous step. This can be viewed conceptually in Figure 4. Simply, a RNN has two inputs: the current input, and the previous time step input. This assists the network in understanding temporally related data points or sequences [7], [**?**].

The activation function at the hidden layer can be simplified as:

$$h_t = a(\sigma_0 + \sigma_1 x_t + \sigma_2 x_{t-1})$$

Where h_t is the hidden layer, σ are the different weights and x_t and x_{t-1} is the data. Note that the activation function a is usually a sigmoid or tanh function.

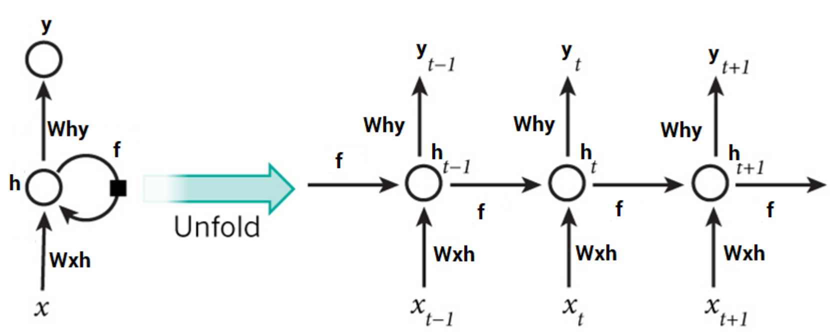

Figure 4: RNN diagram. The hidden layer is self-connected. Image from [7]

In the context of this project, where the slug are short outburst in time, RNNs are likely to perform well. However, they tend to become unreliable when dealing with long time sequences, and the gradients have a tendency to explode or vanish.

Long Short-Term Memory (LSTM)

LSTM is a variation of a RNN which was proposed in 1997 by Sepp Hochreiter and Jürgen Schmidhuber. They are stable and do not have exploding gradient characteristics.

The network structure is slightly different to that of the RNNS, in that instead of having a single hidden layer repeat itself, the hidden layer contains four layers with structures called gates. To explain this simply, there are three types of gates within the layer: an input gate, an output gate and a forget gate. Each has their own activation function ensuring the most valuable information goes through to the next time step [3].

LSTM is an accepted type of neural networks and examples of different networks are available on Pytorch documentation.

2.3.3 Outlier Detection

The slug detection problem could also be approached as an anomaly/outlier detection problem. This could work particularly for light-slugging wells such as the A19. It would make the assumption that there is a normal range for the data to fall in, and that the slugs are anomalies in the data. This would work particularly well for automated slug detection. This is an unsupervised classification approach.

There are two main approaches that have been considered in this literature review: One-Class Categorization with Support Vector Machines and Self Organizing Neural Networks, both applied for time series.

One-Class SVMs are most commonly used for outlier detection [8], [11], [4], [10]. In recent years, SVMs have been adapted to take into account time series. There are different approaches to converting the times series into SVM. J. Ma et al [8] uses use feature space. Scholkpof et al [10], look at using a kernel function to identify 'normal' regions and 'anomaly' regions.

On the other hand, Self-organizing Neural networks [1] work by using non-recurrent neural network with kernel functions. This approach is however deemed too complex for the scope of this project.

2.3.4 Discussion

There are two approaches that are preferred at this stage: using RNNs such as LSTM or Anomaly detection using One-Class SVMs. There are two very different approaches that need to be tested before affirming one is better than the other for the project applications. However, at this stage, times series modelling is preferred, as it is more straight forward, and can be more forgiving regarding the noisy data.

It should also be mentioned that the research performed here is non-exhaustive the approach might change along the remaining of the project.

3 Proposed approach

3.1 Project Plan

The following milestones have been identified for this project:

05/07	- Performed Research into Slugging - Visit to Bergen, Norway to meet team and understand project objectives - Researched and understood previous work for the project - Preliminary Research into possible supervised machine learning models - Raw Data Visualized
05/07	- Preliminary Report Due
17/07	- Data Processed and Cleaned - (Initial) Slug Indicator Defined - Select machine learning models
29/07	- Train Models to detect slugging events - Validate Models using a validation set - Select best performing model
02/08	- Research further into best models to predict parameters behaviour - Adapt Slug Detector Model
16/08	- Train Models to predict parameters behaviour - Validate Models using a validation set - Select best performing model
23/08	- Combine Models if required - Share Results with Production team for testing
30/08	- Software Due - Final Report Due
~15/09	- Presentation

References

[1] Leonardo Aguayo and Guilherme A. Barreto. Novelty detection in time series using self-organizing neural networks: A comprehensive evaluation. *Neural Processing Letters*, 2017.

[2] Mustafa Al-Naser, Moustafa Elshafei, and Abdelsalam Al-Sarkhi. Artificial neural network application for multiphase flow patterns detection: A new approach. *Journal of Petroleum Science and Engineering*, 145:548–564, 2016.

[3] Colah. Understanding lstm networks, Aug 2015.

[4] Santanu Das, Aditi Chattopadhyay, and Ashok N. Srivastava. Classifying induced damage in composite plates using one-class support vector machines. *AIAA Journal*, 48(4):705–718, 2010.

[5] Alphonsus Fagan. An introduction to the petrleum industry. *Government of Newfoundland and Labrador*, Nov 1991.

[6] Pablo Guillen-Rondon, Melvin D Robinson, Carlos Torres, and Eduardo Pereya. Support vector machine application for multiphase flow pattern prediction. Jun 2012.

[7] Dishashree Gupta. Fundamentals of deep learning – introduction to recurrent neural networks. *Analytics Vidhya*, Dec 2017.

[8] J. Ma and S. Perkins. Time-series novelty detection using one-class support vector machines. *Proceedings of the International Joint Conference on Neural Networks, 2003.*, 2003.

[9] E.s. Rosa, R.m. Salgado, T. Ohishi, and N. Mastelari. Performance comparison of artificial neural networks and expert systems applied to flow pattern identification in vertical ascendant gas–liquid flows. *International Journal of Multiphase Flow*, 36(9):738–754, 2010.

[10] Bernhard Scholkopf, Robert Williamson, Alex Smola, John Shawe-Taylor, and John Platt. Support vector method for novelty detection. *NIPS'99 Proceedings of the 12th International Conference on Neural Information Processing System*, 1999.

[11] John Shawe-Taylor and Blaž Žličar. Novelty detection with one-class support vector machines. *Studies in Classification, Data Analysis, and Knowledge Organization Advances in Statistical Models for Data Analysis*, page 231–257, 2015.

[12] Schlumberger Software. Slugging. Technical report, 2016.