

Development of industrial process characterisation through data analysis

Aleksandar Josifovic
Design, Manufacture and
Engineering Management
University of Strathclyde

75 Montrose street, Glasgow, G1 1XJ
Email: aleksandar.josifovic@strath.ac.uk

Jonathan Corney
Design, Manufacture and
Engineering Management
University of Strathclyde

75 Montrose street, Glasgow, G1 1XJ
Email: jonathan.corney@strath.ac.uk

Abstract—Recently the world has seen an explosion of available data, gathered from sensors, experiments and direct measurements. The engineering world is shifting from model-centred simulations to the exploitation of large datasets for insights into machine behaviour. In the case of optimisation, where usually a simulation model is provided for the exploration of the parameter space, the new data-driven approach translates into the use of data for quantifying problems and delivering an optimal solution.

This paper presents a methodology developed for interpretation of condition monitoring data in the design process and site characterisation of hydraulic fracturing. Field data samples were processed to construct a nominal operating regime in a process pumping system.

After describing the generic process a case study implementation of field data from a North American hydraulic fracturing site is presented for illustration. The paper summarises sensor installation, data acquisition and cloud-based pre-processing. Consequent analysis demonstrates expected operational output in normal operating conditions for a specified geological terrain. The paper concludes with a base-case scenario for PD pump's process output.

I. INTRODUCTION

In recent years modern engineering systems have witnessed the start of a shift in the process of product development from in-depth physical modelling to increasing exploitation of real-time monitoring of on-site operation. Machine learning, developed from Artificial Intelligence, aims to provide solutions for problems based on developed algorithms and analysis of historical data.

The concept of machine learning has been previously used in many different industries and scientific studies such as agriculture [1], mineral extraction [2], optical recognition [3], robotics [4] and others. In industrial environment, two approaches prevail; data collection on site and processing in cloud or, both data collection and processing is done directly on site.

Hydraulic fracturing process was established in 1950s [5] and involves using large number of positive displacement (PD) pumps to pressurise and deliver slurry-water mix into a well [6]. This method for stimulation and increase of hydrocarbon recovery has had global economic impact in energy sector, as demonstrated in North America [7], [8], [9].

In this well stimulation techniques all the central processes are controlled manually with limited amount of feedback information related to equipment life and failure prediction. Data processing of site metrics is currently being recorded on site and evaluated at central locations. Initial assessment stage needs to consider present on-site practices to develop benchmark criteria for future exploration wells. Proposed initiative considers using characterized process metrics to develop performance envelopes for satisfactory equipment operation. With the expansion of gathered data potential for predictive maintenance will be feasible. This paper evaluates mechanisms for extending equipment's operational availability by introducing a methodology for a PD pump monitoring system.

Aim of this work is to establish methodology for development of data processing system for hydraulic fracturing.

Systematic approach for implementing machine learning system involves using previously recorded field metrics to characterise, evaluate amplitudes and assign operational envelopes with specific phases of the process. Specific milestones in this work include following:

- Sensor identification and installation on equipment at process location
- Data acquisition using specialised industrial hardware
- Upload to cloud-based system and access through bespoke data software
- Data evaluation using MatLab for identified process phases
- Establishing performance envelopes based on previously examined data

II. METHODOLOGY

In Figure 1 knowledge discovery diagram is presented.

Source data - For data processing and analysis of hydraulic fracturing processes a case study was required. Selected data was recorded in Texas, U.S. in 2015, the upcoming sections will present additional well properties.

Target data - Pumping stages in hydraulic fracturing typically last between 60 and 180 minutes. In addition to time constraint two other determinants that define individual location are required flow rates and treating pressures [10]. Case study identifies a single pumping stage.

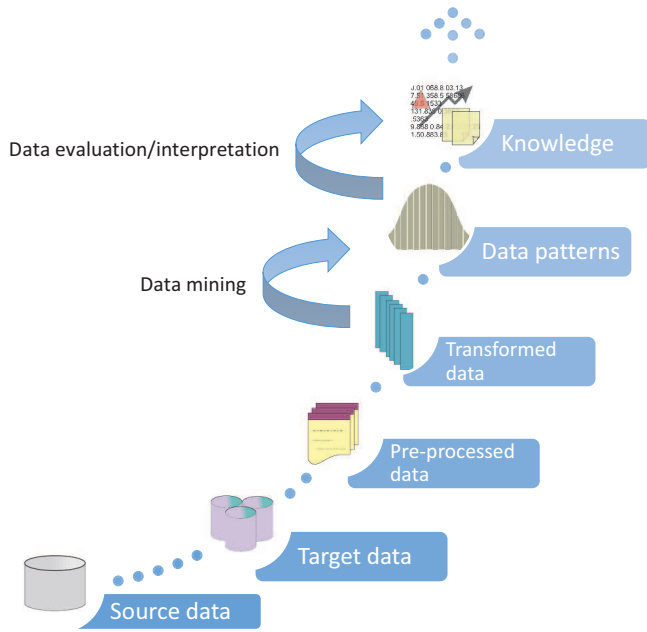


Figure 1. Data mining and knowledge discovery diagram

Pre-processed data - Processing site information requires selection of relevant operational metrics to be recorded for analysis. In this case, PD pump's inlet/outlet pressure, speed and vibration are measured.

Transformed data - Acquired pump metrics are transformed for display into SI units, pressure in *MPa*, speed in *RPM* and vibration in *g*.

Data patterns - Physical measurements are aligned and displayed in the same phases of pumping process. Data characterisation in form of mean, extreme and frequency values are presented for all sensor measurements.

Knowledge - Cumulative sensor output provides necessary knowledge for site characterisation in hydraulic fracturing by defining surface pump pressure and associated vibration for a given pump speed.

III. CASE STUDY DESCRIPTION

Hydraulic fracturing involves using multiple high pressure PD pumps to propel fluid downhole. A typical site consists of multiple frac-truck assemblies each of which carries a single PD pump.

Photograph of a hydraulic fracturing process site, with equipment annotation, is shown in Figure 2.

A. Sensor data

One of the pump units was instrumented with a slurry pressure sensors, a speed sensor, accelerometers on both power-end and fluid-end together with a number of pressure and temperature sensors for power-end lubrication, shown in Figure 3. In addition to PD pump instrumentation, engine and transmission operation metrics were recorded to monitor drivetrain performance.



1 - PD pump
2 - Frac-truck assembly
3 - Blender - slurry mixer
4 - Water container
5 - Sand container
6 - Manifold trailer
7 - Wellhead
8 - Data collection van

Figure 2. Equipment identification on site

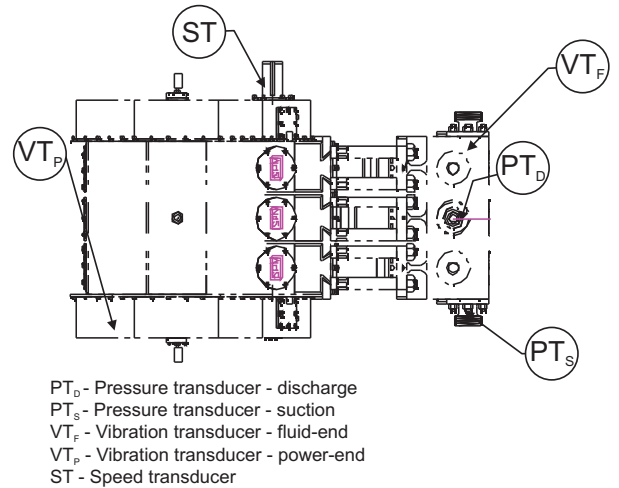


Figure 3. Sensor allocation on the field unit

All the reciprocating equipment on a frac-site can be related to the pumping frequency, therefore running the PD pump at higher speed will demand a higher acquisition rate from the sensors. For example, a higher acquisition rate is required for accelerometers to identify spectral energy at higher frequencies.

B. IMS gateway

A key component of the system described is the IMS gateway, shown in Figure 4. This unit is used to acquire sensor data, perform signal processing and algorithmic calculations on the data and utilise its communication interfaces to send collected information to a cloud server.

Data flow path, illustrated in Figure 5, progresses as follows:

- 1) Signal conditioning changes transducer output into a readable value, which is followed by noise attenuation. For example a 4-20 mA value, obtained from the sensor, is converted into 0-5 VDC.
- 2) The signal conditioned data is then sent to one of the two data acquisition (DAQ) cards where the analogue data is sampled and converted into a digital format



| Data acquisition module | | |
|-------------------------|---|----------------------------------|
| Electrical | Power Supply 20-28Vdc | 1.5A |
| DAQ | Inputs 40 analogue inputs 0-20mA 4-20mA +/- 10Vdc | Outputs Modbus Ethernet |
| Sample rate | 80kHz 3 Channels | Frequency 8kHz 10 Channels |
| Single Board Computer | Processor i7 | Software Windows 8.1 |
| Communication: | Ethernet/J1939/ WiFi/Celular/GPS | |

Figure 4. Data acquisition unit is installed in the field. Main characteristics are listed in the table.



Figure 5. Methodology for data acquisition from the PD pump unit on site.

that can be read by the Single board computer (SBC). Accelerometers are recorded at 4kHz in DAQ card B and the remaining sensors are recorded at 1 kHz in DAQ card A.

- 3) The DAQ data is then sent to the SBC via USB interface. All the sensor data is stored in a database with time stamps and associated sensor information (e.g. sensor type and units of measure)
- 4) Data is then saved to the solid state drive (SSD). All the data firstly goes through a buffer and all the data can be approached almost instantaneously (no need to be pre-programmed). It can also be automated to acquire specific data at any available frequency.
- 5) Data is finally uploaded to the cloud server or physically transferred by one of the field engineers. The main challenge with current DAQ system is the volume of data storage space required to record high speed time series information. This can quickly run to terabytes of data and consequently the system is programmed always to erase the oldest data.

C. Rate of acquisition and data storage

Stage data is stored using three different sampling rates. Low speed was used to capture the whole stage. Due to the length of the stage, which could be up to three hours in length, it was sufficient to have 1 Hz sampling frequency to capture accurately well-head metrics (e.g. sampling frequency of 1 kHz is used for all pressure sensors, temperature sensors and rotational encoders installed on the field unit). In contrast, the accelerometers on the PD pump's power-end and fluid-end

acquire data at 4 kHz owing to the speed of the reciprocating drives and the induced equipment vibration on site.

D. Data transfer method

Physical data was downloaded using a specialised software, 'SiWorkspace', from the cloud to a local workstation where data was processed using MatLab. Although SiWorkspace is a configurable tool it currently only has limited data processing capabilities and is in principle only used to download the data from the cloud.

IV. DATA ANALYSIS

The system aims to allow all significant events, seen in the operation, to be detected and recorded. By evaluating their magnitude and impact the opportunity arises to predict and control events or foresee their development. The system would also be able to send information to the operator and signal any operational irregularity. For example, excessive vibration is common on site and to date it's been difficult to get data to investigate the cause. In addition, the handlers on site needs to walk between units during operation and report any leaks or extreme vibration. There are potential safety benefits which a control system can provide.

In addition to the speed encoder (measuring PD pump's rpm) and discharge pressure sensor, location of which can be seen in Figure 3, pump is also instrumented with two accelerometers. In Figure 6 precise orientation and placement on two main subsections of the pump, power-end and the fluid-end, can be seen.

To illustrate the system data PD pump pressure and speed are displayed in Figure 7. It is clear that changes occur in 44th, 53rd, 94th and 126th minute. All of the changes can be correlated to variations seen in hydraulic fracturing stage, thereby, highlighted areas will be analysed further by using high frequency data. However, there hasn't been enough information to justify in-depth discussion of each of these states as they show great similarities between one another.

Selected data shows operational measurements from two accelerometers, suction (inlet), discharge (outlet) pressure and pump speed. Recorded metrics from two accelerometers are

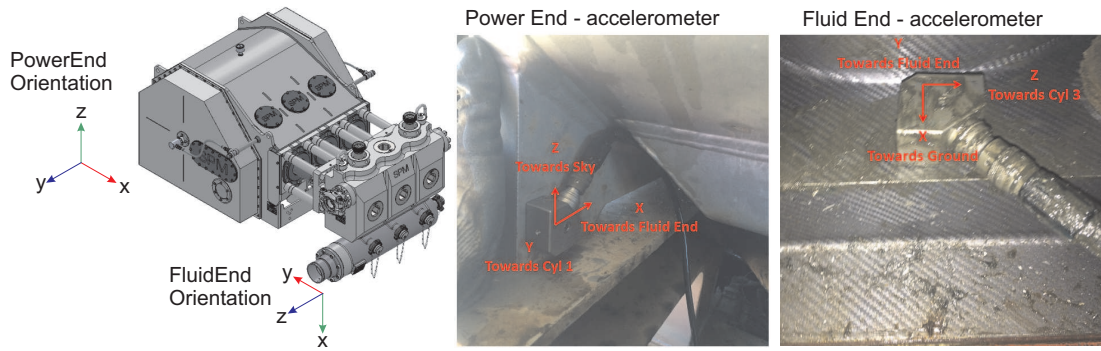


Figure 6. Orientation of accelerometers on the PD pump

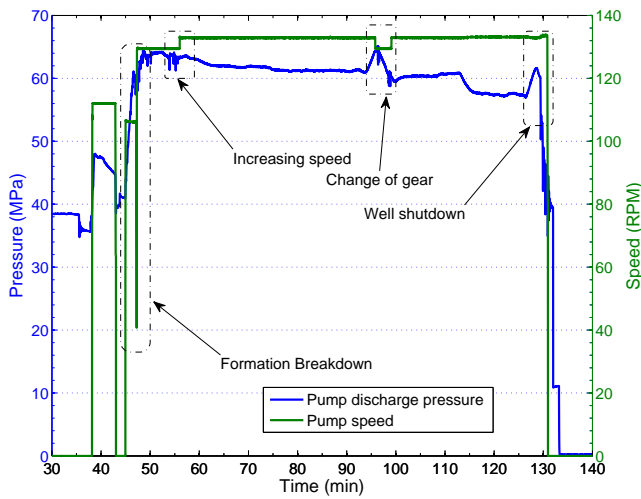


Figure 7. Identifying transitional parts of the stage

Table I
DATA FROM TRI-AXIAL ACCELEROMETER LOCATED ON THE PD PUMP'S
POWER END

| Sensor | Peak (g) | | Mean (g) | RMS (g) |
|-----------------|----------|----------|----------|---------|
| | Positive | Negative | | |
| x-accelerometer | 5.09 | -5.00 | -0.03 | 0.83 |
| y-accelerometer | 0.39 | -0.36 | -0.02 | 0.08 |
| z-accelerometer | 0.62 | -0.69 | -0.03 | 0.11 |

shown in three coordinate axes (X,Y and Z) at acquisition frequency of 4 kHz. The remaining sensors are all acquiring at 1 kHz.

1) *Power-End (PE) Tri-axial accelerometer*: Time domain analysis of the crankshaft accelerometer in state one is shown in Table I.

Analysis in time domain demonstrates significantly higher energy in x-axis of the crankshaft assembly. Calculating RMS values shows approximately 10 times more energy in x-axis compared to y-axis and 7 times more energy than in the z-axis.

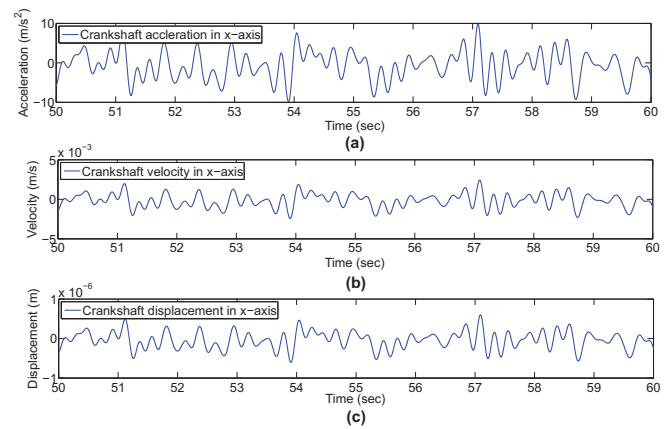


Figure 8. Crankshaft accelerometer and data metrics displaying velocity and displacement

Peak values are mostly symmetrical in positive and negative magnitude. Mean values are relatively low in all three orientations.

To date the evidence suggests there is only limited interaction between the vibration in the three axes, so each of the axis is considered as an independent function. Because of the comparatively high energy in one direction further PE accelerometer analysis and discussion focuses solely on the x-axis.

Time series data obtained from the PE accelerometer can be further developed by double integration to evaluate the crankshaft displacement. Pump's PE displacement along the x-axis is quantified in Figure 8 (c). The noisy signal was processed with 50 Hz filter.

Analysis in the frequency domain of accelerometer in x-axis is displayed in Figure 9.

Accelerometer reading in the x-axis of the PE are shown in Figure 9. Given that the pump is operating at 106 rpm (i.e. 1.76 Hz) it can be seen that the dominant amplitude spikes are at 455 Hz, 15 Hz and 135 Hz which are listed according to their power amplitude.

2) *Fluid-End (FE) Tri-axial accelerometer*: Time domain analysis of the accelerometer placed on the fluid-end chamber

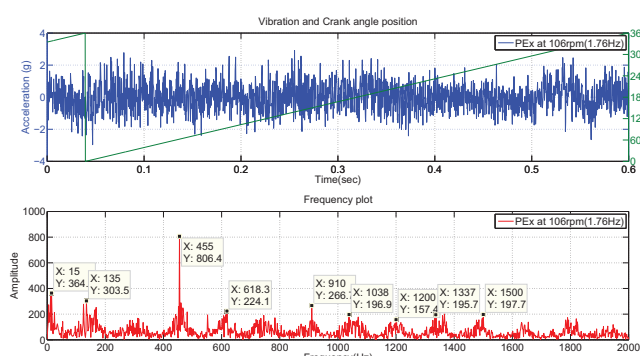


Figure 9. Frequency plot of accelerometer placed on the PD pump's Power End in the direction of the x-axis

Table II
DATA FROM TRI-AXIAL ACCELEROMETER LOCATED ON THE PD PUMP'S FLUID END

| Sensor | Peak (g) | | Mean (g) | RMS (g) |
|---------------|----------|----------|----------|---------|
| | Positive | Negative | | |
| x-accelerator | 0.27 | -0.28 | -0.03 | 0.04 |
| y-accelerator | 4.65 | -5.49 | -0.03 | 0.17 |
| z-accelerator | 0.28 | -0.29 | -0.02 | 0.06 |

in state one is shown in Table II.

Once again energy in the direction of one of the axes is visibly higher compared to the other two. Again, displacement can be evaluated using double integration of the recorded acceleration. Figure 10 shows fluid-end displacement over the course of ten seconds.

Because of the nature of the signal it was decided that no filtering was needed as it will affect the final output. Because of this it is clear that pulses are equally spaced. This means that pump running at a constant speed will produce synchronized pulses that correspond well with plunger movement inside the pump. This effect is naturally more prominent on the fluid-end side of the pump because of the closeness to the reciprocating source (plunger) and the operating medium (slurry).

Analysis of the fluid-end accelerometer in frequency domain can be seen in Figure 11. Presented data is extracted from the same time segment as in the analysis of the crankshaft metrics.

From the data displayed in Figure 11 it can be seen that there are certain similarities to the data presented in the crankshaft section, which is to be expected. Primarily, the occurrence of the peak at 455 Hz matches earlier results. The remaining two peaks at 150 Hz and 10 Hz are close to the ones exhibited in the PE section. Power amplitudes are lower in the FE section, vibration peak in time series is lower by 70% compared to the PE.

3) *Suction pressure analysis*: PD pump's inlet pressure is a good indicator of the overall performance and can be used very efficiently for predicting cavitation [11]. Data is processed and presented in form of time series and frequency domain.

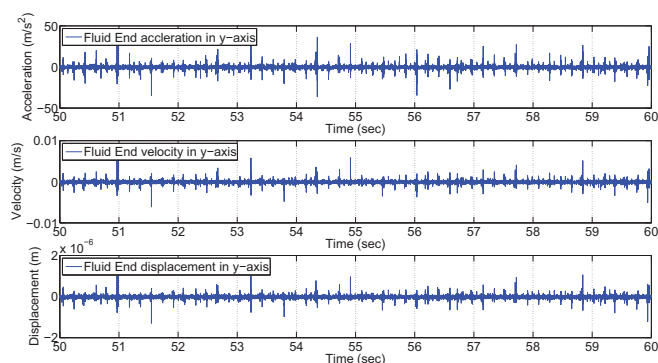


Figure 10. Acceleration, velocity and displacement plot from accelerometer located on the Fluid End in the direction of the y-axis

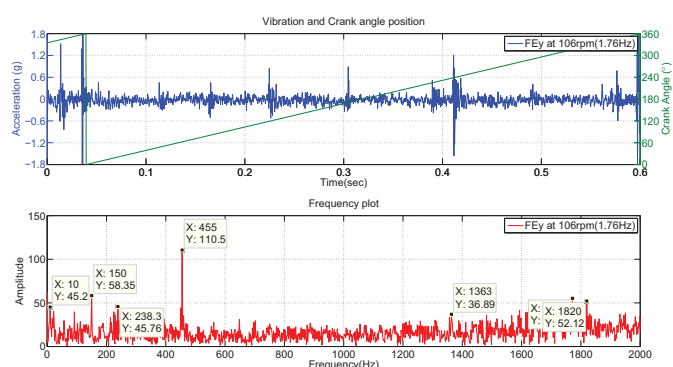


Figure 11. Frequency analysis plot from accelerometer located on the Fluid End of a PD pump in the direction of the y-axis

Table III shows peak pressure and mean values during one second.

An entire pumping cycle is displayed in Figure 12 where pressure is shown in blue and plunger displacement, traveling from 0 to 360°, is shown in green on the right axis. Unsurprisingly, the three cylinder pump has three distinct pressure spikes at 0°, 120° and 240° of crank rotation. Pressure fluctuations, shown in Figure 12, are influenced by opening and closing of the inlet valves which gives rise to the transient behaviour inside the inlet manifold of the PD pump. Instrumented PD pump was operating in the fleet of PD pump units thereby the possibility of suction pressure fluctuation due to system interaction must not be ignored.

Peak to peak variation is illustrated in Table III. Positive

Table III
SUCTION PRESSURE OPERATIONAL METRICS

| Sensor | Peak (MPa) | | Mean (MPa) | RMS (MPa) |
|------------------|------------|----------|------------|-----------|
| | Positive | Negative | | |
| Suction Pressure | 0.53 | 0.22 | 0.36 | 0.36 |

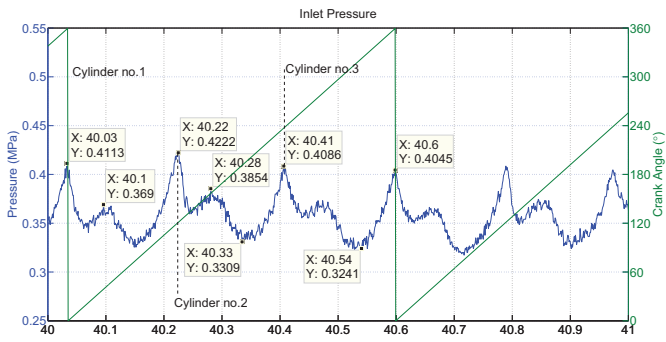


Figure 12. Suction pressure on the inlet side of the pump during one pumping stroke

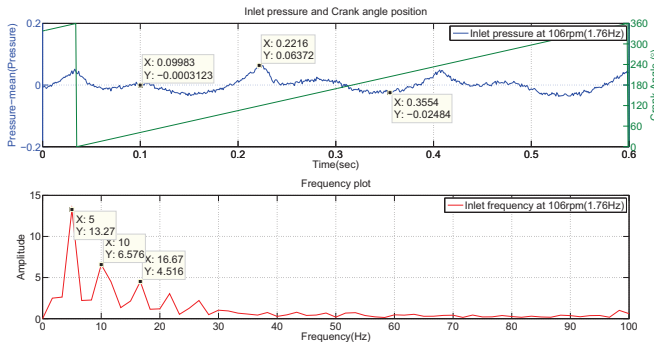


Figure 13. Frequency spectrum of the suction pressure on the inlet side of the pump during one pumping stroke

pressure peak is approximately 46% higher than the mean value. Pressure variation in the negative pressure peak is 60%.

It can be seen that peaks accurately align with the crank angle position and the annotations in Figure 12 indicate precise timing of the individual cylinders during one pumping cycle. In Figure 13 a single pumping cycle is analysed in frequency domain. The regular pressure pattern is visible in the frequency plot. Results show regular energy spikes at 5 Hz, 10 Hz and 16.7 Hz.

4) *Discharge pressure analysis:* Time domain analysis of the discharge pressure has significance due of its influence on pump performance. Pressure was recorded at the discharge plenum located inside the PD pump. In the field the possibility that other PD pumps in the system are influencing operation of the test unit must also be considered.

Mean pressure during the recording time of two minutes, along with details on peak values, are shown in Table IV.

Figure 14 displays operation during one pumping cycle. The slow modulation of the outlet pressure can be the result of the load and the flow rate. Overall complexity of the signal is shown in the Figure 14.

Peak to peak oscillation varies in the sampled period. Section from the later part of the recording, shown in Figure 15, exhibit higher amplitudes.

Analysis in the frequency domain for the discharge pressure was done on the same time range of one second. Figure 16

Table IV
DISCHARGE PRESSURE

| Sensor | Peak (MPa) | | Mean (MPa) | RMS (MPa) |
|--------------------|------------|----------|------------|-----------|
| | Positive | Negative | | |
| Discharge Pressure | 43.80 | 70.18 | 57.18 | 57.34 |

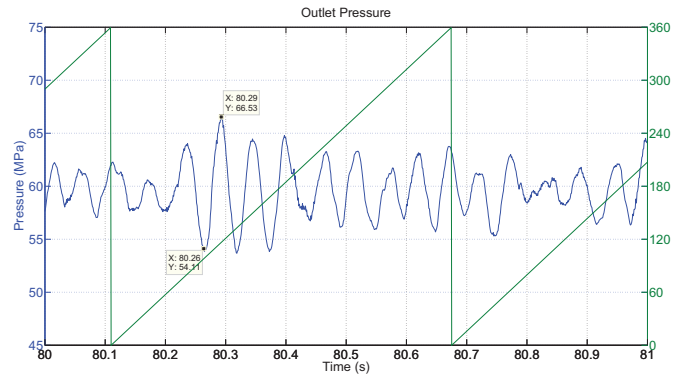


Figure 14. Discharge pressure inside outlet plenum of the pump during one pumping stroke

displays results from this analysis.

Frequency spectrum, shown in Figure 16, depicts impulsive nature in the low frequency part of the pressure signal. It is clear that there is a strong frequency concentration at 18 Hz forming 10:1 ratio with plunger frequency of 1.76 Hz.

V. RESULTS

Synchronization in the time domain and alignment against crankshaft position provides useful insight towards understanding pump operation in the field environment. Clear correlations between different sensor measurement can be seen and some of their relationship have been presented in this paper. In Figure 17 these observations are highlighted.

For example the accelerometer on the fluid-end, in the direction of the y-axis, shows equally spaced pulses that

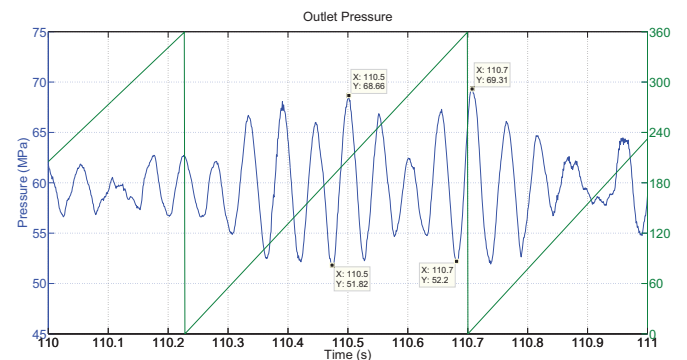


Figure 15. High amplitude variation of the discharge pressure inside the outlet plenum of the PD pump during one pumping stroke

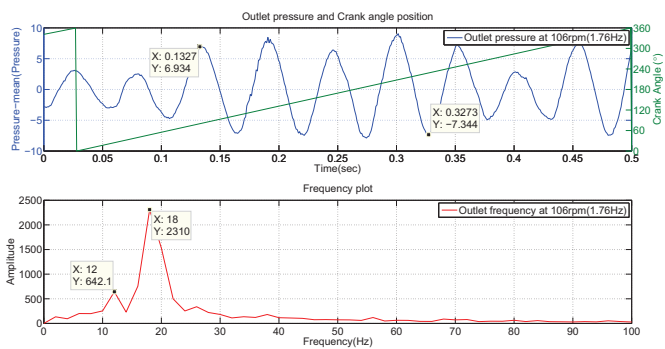


Figure 16. Frequency spectrum of the discharge pressure on the outlet side of the PD pump during one pumping stroke

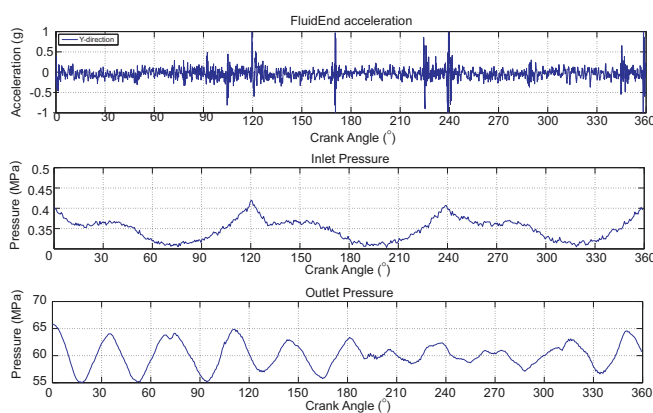


Figure 17. Important pump metric and their relationship during one full pumping cycle

correlate well with angle of the crankshaft. Similarly, inlet and outlet pressure show strong association with the type of a PD pump (i.e. triplex) in the output signal. Pressure variation is also presented and $\pm 10\%$ fluctuation is seen around the mean value.

VI. CONCLUSION

In conclusion, the information provided in this paper illustrates the complexity of data being collected from industrial operation to monitor site behaviour. This analysis highlights the field operation and the challenges that have to be overcome before the goal of a comprehensive mechatronic site control system can be realized.

The analysis of the accelerometers has shown that the vibration along the length of the pump is dominant. This presents valuable information that needs to be further developed to support structural analysis. For example, similar field data analysis has previously been reported for wind turbines where the fatigue cycle was measured and implemented in computational model to assess possible improvements and design modification [12], [13].

In pressure pulsation studies, scenarios showing steady state and transitional operation were displayed and discussed which have possibility to be used as system training data sets.

Future implementation of this approach in industrial systems could provide live monitoring and feedback to equipment operator. Lastly, the condition monitoring system could provide information about any outstanding trends or predict failure based on this method for developing training data sets.

REFERENCES

- [1] R. J. McQueen, S. R. Garner, C. G. Nevill-Manning, and I. H. Witten, "Applying machine learning to agricultural data," *Computers and electronics in agriculture*, vol. 12, no. 4, pp. 275–293, 1995, [URL](#).
- [2] J. Tessier, C. Duchesne, and G. Bartolacci, "A machine vision approach to on-line estimation of run-of-mine ore composition on conveyor belts," *Minerals Engineering*, vol. 20, no. 12, pp. 1129–1144, 2007, [URL](#).
- [3] A. Fathi and G. Mori, "Action recognition by learning mid-level motion features," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8, [URL](#).
- [4] J. R. Peters, "Machine learning of motor skills for robotics," Ph.D. dissertation, University of Southern California, 2007, [URL](#).
- [5] G. E. King and Apache, "Hydraulic fracturing 101: What every representative environmentalist regulator reporter investor university researcher neighbor and engineer should know about estimating frac risk and improving frac performance in unconventional gas and oil wells," in *SPE Hydraulic Fracturing Technology Conference*. Society of Petroleum Engineers, 2012, [URL](#).
- [6] Daneshy Consultants International, "Hydraulic fracturing to improve production," Tech101, 2010, [URL](#).
- [7] Rohan North, "Hydraulic fracturing market (online)," Accessed: 16 April 2016, 2016, [URL](#).
- [8] API, "Hydraulic fracturing unlocking americas natural gas resources," Accessed: 16 April 2016, 2014, [URL](#).
- [9] Oil and Gas - Financial journal, "Hydraulic fracturing market projected to continue growth," Accessed: 16 April 2016, 2015, [URL](#).
- [10] M. J. Economides, K. G. Nolte, U. Ahmed, and D. Schlumberger, *Reservoir stimulation*. Wiley Chichester, 2000, vol. 18, ISBN: 978-0471491927.
- [11] P. J. Singh and S. D. Able, "Determination of npshr for reciprocating positive displacement pumps—a new approach," in *PROCEEDINGS OF THE INTERNATIONAL PUMP USERS SYMPOSIUM*. TEXAS A&M UNIVERSITY SYSTEM, 1996, pp. 131–142, [URL](#).
- [12] B. Jouan, S. Bergholz, and J. Rudolph, "Fatigue monitoring approaches for power plants," in *ASME 2014 Pressure Vessels and Piping Conference*. American Society of Mechanical Engineers, 2014, pp. V001T01A035–V001T01A035, [URL](#).
- [13] S. Bergholz, J. Rudolph, and A. Willuweit, "Famos 4 wind a new method for the fatigue monitoring of wind energy plants," *Procedia Engineering*, vol. 133, pp. 84–89, 2015, [URL](#).