# Data-Driven Forecasting for Anomaly Detection in a Compressor Unit

Aycan Sapmaz, a\* Aslı Yasmal, a Gizem Kusoglu Kaya,a Yasin Utar,b Barış Akgünc

a Turkish Petroleum Refinery, Körfez, Kocaeli, 41780, Türkiye

b Turkish Petroleum Refinery, Hacılar, Kırıkkale, 7180, Türkiye

c Dept. of Computer Engineering, Koç University, Sarıyer, Istanbul, 34450, Türkiye

aycan.sapmaz@tupras.com.tr

Abstract

Equipment reliability is crucial for refineries and timely anomaly (equipment failures, sensor faults, wear and tear, unexpected inputs etc.) detection is essential to keep equipment running safely, improve performance, and have an effective maintenance strategy. Modern refineries generate large amounts of data. Combined with machine learning, models that can monitor the operation of complex processes and equipment in real-time can be learned. These models can guide operators, and engineers in identifying faults. This study proposes a data-driven approach to detect anomalies of a reciprocating compressor in a petrochemical refinery. The idea is to capture the regular operating behavior of the compressor with a learned model and compare its predictions with measurements. As such, a model that forecasts future sensor outputs given past measurements is trained from real-world historical data. Deep neural networks with recurrent layers are utilized. After training, the forecasted measurements are compared with the observed measurements and any large deviations are flagged as potential anomalies. The approach is evaluated both on historical and real-time data. The results demonstrate that the approach can be used as an anomaly detection decision-aid for operators and engineers. The approach has the potential to facilitate rapid actions, to help avoid major faults, and for reducing operator fatigue and cognitive load, letting them focus on higher level tasks such as monitoring entire processes versus single equipment.

**Keywords**: Reciprocating Compressor, Predictive Maintenance, Anomaly Detection, Fault Detection, Condition Monitoring.

* 1. Introduction

The petroleum industry has strict monitoring requirements, with operational security and safety serving as guiding values. Damaged equipment compromises safety leads to energy shortages and causes financial losses. Anomalous equipment or process behavior often signals such issues beforehand which motivates the development and deployment anomaly detection systems. We highlight that many chemical processes have normal/acceptable within a range and that these behaviors can be learned from historical data. This learned model can be deployed online and its predictions can be compared with observations, raising a warning when the two do not agree.

Due to the availability data, prevalence of machine learning and importance of anomaly detection, multiple data-driven approaches have been created. Recurrent Neural Network (RNN) models are popular in the literature due to their ability to model sequences (Rivas et al. 2020). For example, Canizo et al. (2019) used RNNs in industrial elevator systems for predictive maintenance. Koprinkova-Hristova et al. (2011) used RNNs for predictive maintenance on power plant mill-fans and to perform online monitoring. These methods use the deviations between observations and predictions to detect anomalies.

As a use case, we focus on reciprocating compressors which are an important part of any refinery. They are prone to performance degradation during extended operation due to the dynamic nature of oil refining. It is crucial to promptly identify any anomalies in their operation and warn operators to avoid hazardous situations and costly shutdowns. We use Long-Short Term Memory (LSTM) and Gated recurrent units (GRUs), both types of RNNs, to capture the typical behavior of a compressor and use it for anomaly detection.

* 1. Methodology
     1. Process Description

The process where the reciprocating compressor located is producing high octane reformate, a main component of gasoline, obtained from low octane heavy naphtha. High purity hydrogen gas (Net gas) is also produced with this process, making it a main hydrogen producer on top of its octanizing feature. Unit feed enters reactors to produce high octane aromatic compounds from paraffins and naphthene. After this reaction, liquid and gas phase products are separated. The gas phase is further separated to two, for the recycle gas compressor and the net gas compressor. After the compression section, liquid and gas phases are again separated and the produced hydrogen is distributed to whole process unit from the effluent of these net gas compressors and the separator. The remaining liquid phase goes to stabilizer section to split reformate from low hydrocarbon product which is the result of the cracking side reactions in the reactors. Within this process, high octane reformate and high purity hydrogen gas are obtained.

Net gas compressors are reciprocating type compressors and use pistons (Mobley, R. K. 2001). They consist of 3 stages and the first stage consists of 2 cylinders. They have high maintenance costs comparted to the centrifugal compressors (Brown, R. N.1997). Since they supply hydrogen, they are crucial for other units as well.

* + 1. Data Description

We posit that the compressor discharge temperature and cylinder vibration can be used to detect anomalies. Based on our expertise and discussions with process engineers, we selected 18 sensor measurements involving stage suction temperatures (oC), cylinder and crank case vibration values (rpm), gas flow rate (m3/h), stage suction pressure (kg/cm2) and stage outlet pressure (kg/cm2). Our idea is to predict the near-future (15 minutes) discharge temperature based on a temporally local history of these measurements.

* + 1. Computational Details and Proposed Methodology

As mentioned in Section 1, we utilize Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) layers to build deep neural networks and to model the normal behavior of our compressor. These layers are selected as they are good at modelling sequences. To avoid overfitting, a drop-out layer is added after the recurrent layer. These layers drop random unit outputs based on a given probability which enhances robustness and forces generalization. Then a feed-forward layer is added with Rectified Linear Unit (ReLU) activations, followed by the prediction layer.

The resulting models take a window of past observations, including the most recent one, as input to make a prediction. This window slides over the sequence one step at a time. This is called the sliding window technique and is usually employed with a fixed window size denoted as 'W' which we set to 10.

We first train this model with substantial amount of data from a timeframe that we know to not include anomalies. Then we deploy it and compare its predictions with new observations. However, due to the dynamic nature of refineries, the measurement time-series are non-stationary, i.e., even their regular behavior changes over time. As such, a learned model will lose its predictive power and result in erroneous detection. Re-training the models with the recently observed data is a way to alleviate this issue. We use a retraining frequency of 100 and use the last 1000 data points to fine-tune the current model. This is called moving window re-training and helps the model to adapt to changing patterns in the data over time. Such an approach has been shown to be effective for tasks such as stock price prediction, weather forecasting, and many others (Hota et al., 2017).

* + 1. Performance Analysis

We use the Mean Absolute Error (MAE) and Root mean squared error (RMSE) metrics for model selection. In addition to performance metrics, model results are also examined visually. Exponentially Weighted Moving Average (EWMA) control charts are used for this. The chart consists of upper and lower control lines notated as UCL and LCL, respectively. Calculations of UCL and LCL are shown below where *σ* is the standard deviation of data, is the target value, is weighting factor and *L* is the number of standard deviations.

(1)

(2)

* 1. Experiments, Results and Discussion
     1. Modelling Regular Behavior

We first make use of a dataset that we know to be devoid of significant anomalies, since the measurements come after a maintenance period. The dataset is sampled with 15 minutes intervals and contains a total of 4375 data points. We divide this dataset into training set – used to learn the model, test set – used to evaluate the model. The idea is to make sure the natural behaviour of the process is captured. We trained multiple models with GRU and LSTM layers, composed of 64 and 128 hidden units with this data. The dropout ratio was fixed at 0.5 for each model. The MAE and RMSE values of this first data set is given in Table 1. Similarly, the residuals are provided in Figure 1. We can see that the GRU-based model with 128 hidden units has the best performance, with the same sized LSTM-based model close behind. Based on this, the GRU-128 model was selected to be the basis of the anomaly detection approach.

* + 1. Anomaly Detection Dataset

For evaluating anomaly detection, a larger dataset spanning 2 years of operation has been selected. This dataset gas 64,875 data points, sampled with 15 mins frequency. It includes anomalies, shutdowns, interventions, times where the compressor works alone and other operational changes. From the perspective of the model, these are outside the regular behaviour and should be flagged as anomalies. A total of 7 anomalies were observed in 2-years dataset. Anomalies 1 and 3 are operational changes, anomalies 2 and 6 involve vibration problems, anomalies 4 and 5 belong to date when compressor works alone and anomaly 7 is due to a change of vibration sensor. Not all of these are actual anomalies (1,3,4,5 and 7).

Table 1. Model performance metrics for the first dataset

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Train** | | | | **Test** | | | |
|  | GRU(64) | GRU(128) | LSTM(64) | LSTM(128) | GRU(64) | GRU(128) | LSTM(64) | LSTM(128) |
| MAE | 0.49 | 0.25 | 0.41 | 0.28 | 1.03 | 0.52 | 0.78 | 0.63 |
| RMSE | 0.66 | 0.38 | 0.54 | 0.37 | 1.39 | 0.85 | 1.09 | 0.90 |

A graph of a graph

Description automatically generated with medium confidence

Figure 1. Residual compressor outlet temperatures for the first dataset with models: (a) GRU(64), (b) GRU(128), (c) LSTM(64), (d) LSTM(128)

* + 1. Anomaly Detection without Retraining

We first evaluate the model learned with the first dataset on the second dataset. We use control charts for visual inspection, Figure 2 for this case. The y-axis shows the difference between the actual and model-predicted values, and the x-axis the time. Horizontal lines represent the LCL and UCL which are set to ±2. This decision was based on Equations. 1 and 2, and discussions with process engineers. The part between the limits (control zone) is taken as normal operation and the outside grey points are labelled as anomalies.

Figure 2 shows that the residuals of the model fluctuate wildly and have large magnitudes most of the time. This leads to labelling most measurements as anomalies. Furthermore, such large magnitudes signal learning issues and non-stationary behaviour.

A graph showing the temperature of a month

Description automatically generated

Figure 2. Compressor outlet temperature residuals for the second data without re-training.

A graph showing the temperature of a person

Description automatically generated with medium confidence

Figure 3. Compressor outlet temperature residuals for the second data with re-training.

* + 1. Anomaly Detection with Moving Window Re-Training

For this part of the study, we start with the model learned with the first dataset but fine-tune it every 100-time steps with the last 1000 data points. We only perform 2 epochs so that the model does not overfit. Figure 3 depicts the results. We can see that anomalies 1,3 and 7 were detected. There were signals for anomalies 3 and 5 as well but not as strong. These correspond to compressor working alone and perhaps do not need to be labelled as anomalies. A clean set of measurements by the second 8th month were wrongly labelled as anomalies. Quantitatively, our method was able to detect 71% of the anomaly points. Lastly, anomalies 2 and 6 went unnoticed which were due to vibrations issues.

* + 1. Anomaly Detection with Vibration Target

Based on our last observation, we decided to utilize vibration information for anomaly detection instead of temperature with the same model architecture and re-training strategy. The results are depicted in Figure 4.

A graph showing a number of different colored lines

Description automatically generated with medium confidence

Figure 4. Residual of compressor vibration for 2 years (with mw)

With the vibration information, we were able to detect anomalies 1,4,5 and 6. Higher loads when working alone probably led to higher vibrations so that anomaly 5 was caught. Using vibration led to catching anomaly 6 as well but we still missed 2. Quantitatively, model labelled 57% of the anomalous points.

* 1. Conclusion

In this paper, we presented an anomaly detection approach for reciprocating compressors which is a significant challenge faced by most oil and gas refineries. We used clean historical data to build a regular operation model of the compressor and compared its output with new observations to detect anomalies. We used moving window re-training to adapt our model in the face of non-stationary. Lastly, we showed that multiple sensors may be required to detect different types of anomalies.

Our next steps include using further features, training multi-output models and devising an anomaly detection approach that can work with these multiple outputs. We also plan to expand our work on different equipment’s.

References

A. Bao, E. Gildin, J. Huang, E. J. Coutinho, (2020, July). Data-driven end-to-end production prediction of oil reservoirs by enkf-enhanced recurrent neural networks. In SPE latin american and caribbean petroleum engineering conference. OnePetro.

A. Rivas, J. M Fraile, P. Chamoso, A. González-Briones, I. Sittón, J. M. Corchado, (2020). A predictive maintenance model using recurrent neural networks. In 14th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2019) Seville, Spain, May 13–15, 2019, Proceedings 14 (pp. 261-270). Springer International Publishing.

G. Qi, W. T. Tsai, Y. Hong, W. Wang, G. Hou, Z. Zhu, (2016, March). Fault-diagnosis for reciprocating compressors using big data. In 2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService) (pp. 72-81). IEEE.

H. S. Hota, R. Handa, A. K. Shrivas, (2017). Time series data prediction using sliding window based RBF neural network. International Journal of Computational Intelligence Research, 13(5), 1145-1156.

H. Purohit, K. Phogat, P. S. V. Nataraj, Performance Monitoring of Centrifugal Compressor System using LSTM based Deep RNN.

M. Canizo, I. Triguero, A. Conde, E. Onieva, (2019). Multi-head CNN–RNN for multi-time series anomaly detection: An industrial case study. *Neurocomputing*, *363*, 246-260.

P. Koprinkova-Hristova, M. Hadjiski, L. Doukovska, S. Beloreshki, (2011). Recurrent neural networks for predictive maintenance of mill fan systems. International Journal of Electronics and Telecommunications.

P. Kamat, R. Sugandhi, (2020). Anomaly detection for predictive maintenance in industry 4.0-A survey. In E3S web of conferences (Vol. 170, p. 02007). EDP Sciences.

R. N. Brown, (1997). Compressors: Selection and sizing. Gulf Professional Publishing.

R. K. Mobley, (2001). Compressors. In Plant Engineer's Handbook (pp. 601-614). Butterworth-Heinemann.

S. Bisgaard, M. Kulahci, (2011). Time series analysis and forecasting by example. John Wiley & Sons.

W. Yu, T. Dillon, F. Mostafa, W. Rahayu, Y. Liu, (2019). A global manufacturing big data ecosystem for fault detection in predictive maintenance. IEEE Transactions on Industrial Informatics, 16(1), 183-192.