Keywords: SCADA, Anomaly Detection, Machine Learning, Gearbox

Comparison of Anomaly Detection Techniques for Wind Turbine Gearbox SCADA Data

C. McKinnon[a], J. Carroll[a], A. McDonald[a], S. Koukoura[a] and Conaill Soraghan[b]

Anomaly detection is a technique being investigated for use in the Wind Energy industry. This has been introduced due to the increased interest in condition based maintenance[1] as a method to reduce the operations and maintenance (O&M) costs of the assets. O&M costs are dependant on monitoring the condition of the turbines, and the analysis of the data collected.

This analysis looks at the use of anomaly detection to assess the condition of wind turbine gearboxes based on data from a number of operational turbines. A comparison is made between various methods of anomaly detection, these being one class support vector machine[2] (OCSVM), random forests[3], and nonlinear autoregressive neural networks with exogenous inputs[4] (NARX). These were selected based on a review of literature.

To complete the comparison of the three anomaly detection algorithms, operational data had to be obtained and pre-processed. Operational data was obtained from an industrial partner and prepared for use. The data preparation involved using a method of clustering the power curve produced based on the SCADA data and then removing outlier data based on the Euclidean distance of the points in each cluster to the centre of the clusters, this is referred to as cleaning. This was done to find a normal behaviour with no anomalies to compare new data to. Features were also selected, from the SCADA Temperature and Pressure data provided, using Univariate Statistics that found the feature with minimal variance from the target.

This analysis assesses the different algorithms to find the most suitable of the three based on the results from the SCADA data. This analysis looks at how the performance is influenced by the number of features used in the training, and the influence of cleaning both the training and testing data. An example of the results is shown in Figure 1, which is for the cleaned training and testing data using two temperature features. Training data being taken from a month long dataset one year before gearbox failure, and the testing data being from one month before failure. This example is of the OCSVM technique, showing on the left the training and testing data plotted on the same graph. On the right is the anomalies detected plotted on top of the Power time series for one turbine. The best performing of the three previously mentioned algorithms will then be used on new data from the Offshore Renewable Energy Catapult's Demonstration Wind Turbine to detect anomalies.

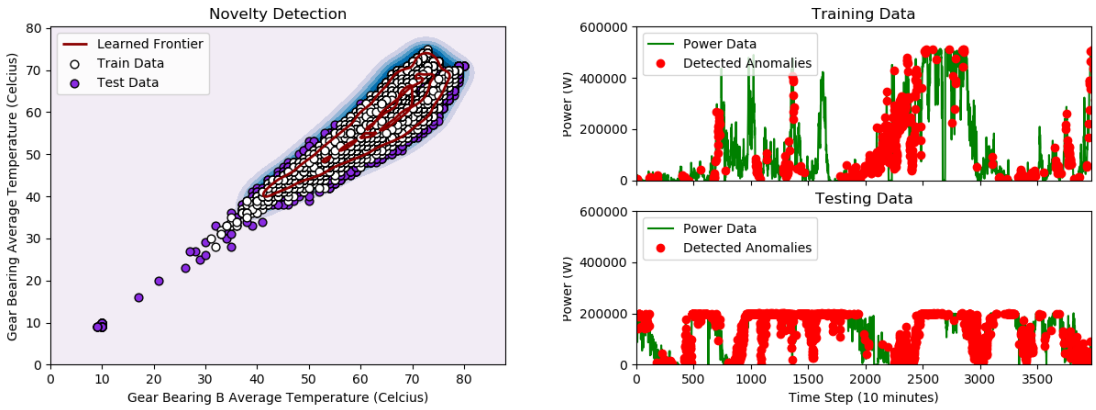

Figure 1: Left: Anomaly Detection results using OCSVM with cleaned training and testing data. Right: Detected anomalies plotted against the Average Power Output.

---

[a] CDT Wind and Marine Energy Systems, University of Strathclyde, Glasgow, UK [b] Offshore Renewable Energy Catapult, Inovo, 121 George St., Glasgow, UK

[1] Nielsen and Sørensen, *RESS,* **96**, 218 (2011) [2] Qu, Qin, et al. *LNICST,* **227**, 437 (2018) [3] McArthur, Shahbazi, et al., *Advanced Engineering Informatics*, **38**, 101 (2018) [4] Cui, Member, et al., *2018 Power Systems Computation Conference* (2018)

Accepted abstract of the following output: Mckinnon, C. (2019). Comparison of anomaly detection techniques for wind turbine gearbox SCADA data. Abstract from Wind Energy Science Conference 2019, Cork, Ireland.