

Advanced operational analytics with machine learning

Will Jowitt

Renewable Energy Analytics, DNV GL

Contents

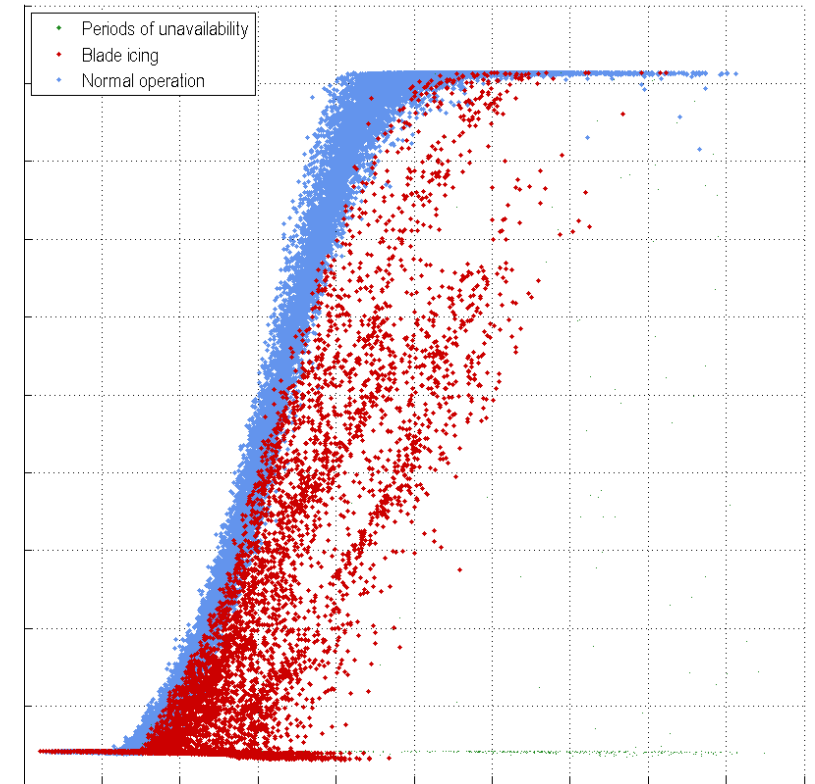
- Power curve performance analysis
 - Turbine blade icing loss
 - Machine learning labelling tool
- Data Mining
 - Turbine Degradation study
- Drivetrain Integrity Monitoring
 - New machine learning approach

Turbine Blade Icing Loss

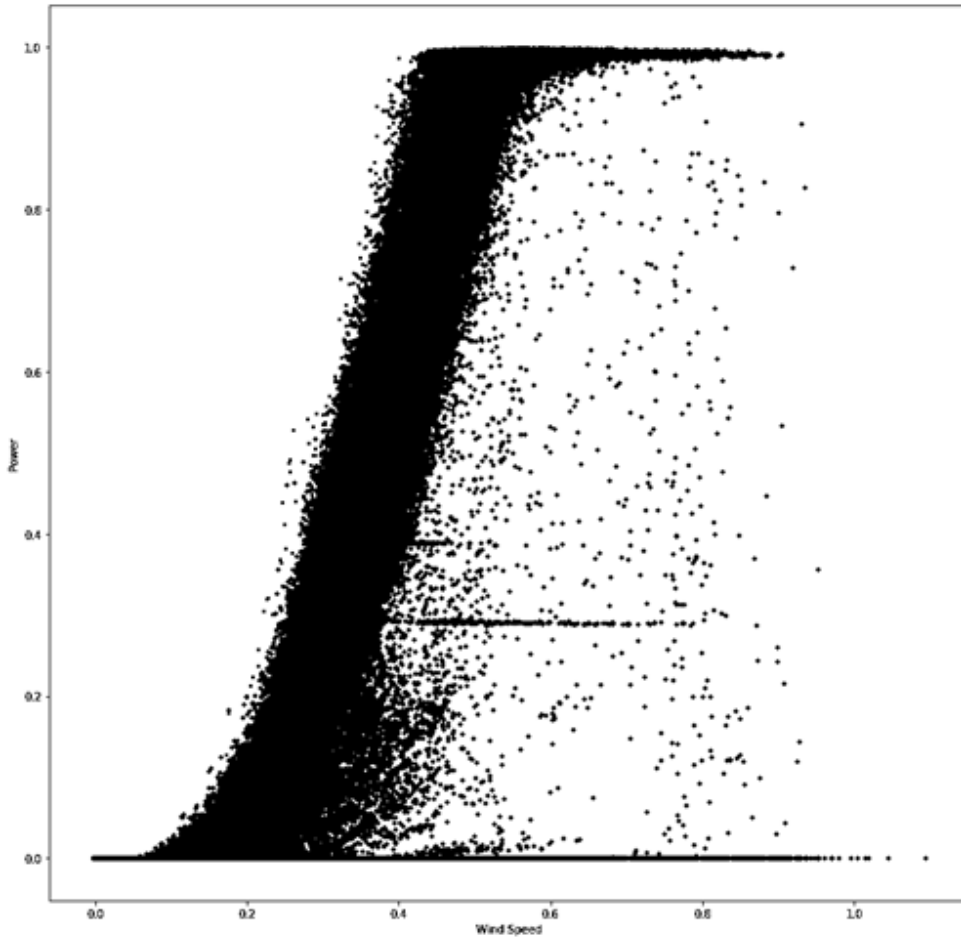


Photo: Kent Larsson, ABVEE

- **Blade icing** is the cause of large production losses through the winter months for many wind farms in the Nordics.
 - Identifying and consistently labelling ice affected production is very time consuming to do accurately
 - We have received a lot of interest surround **IPS warranty reviews**
 - SCADA data enables us to see how effective the systems are at maximizing performance during icing conditions.
-
- These analyses require a large amount of data from multiple wind farms under varied conditions in order to provide credible results.



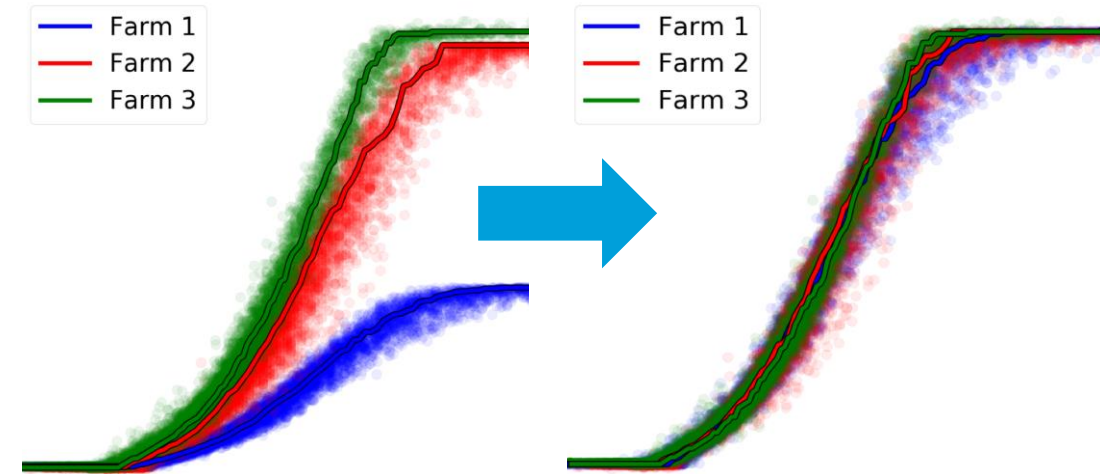
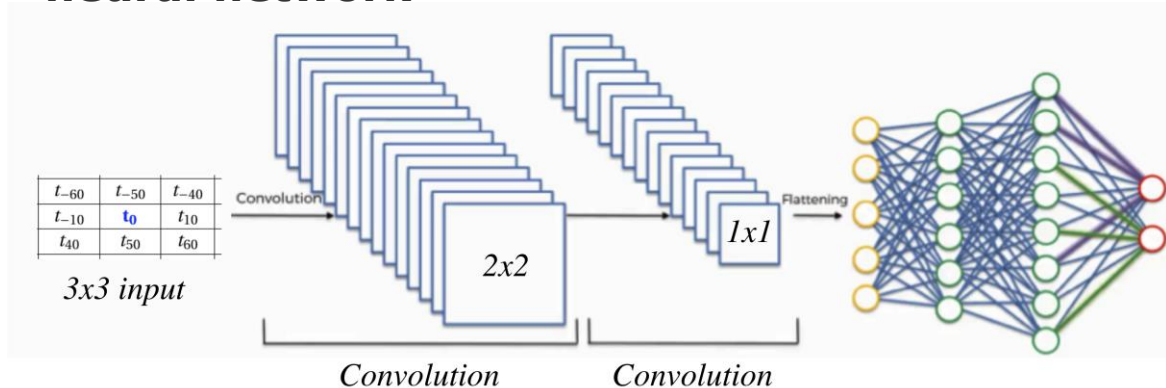
Power Curve Performance Analysis



- SCADA analysis is essential for good analysis of turbine performance.
- Labelling performance issues accurately is time consuming and susceptible to human error
- Much of the effort in performance assessments are driven by SCADA analysis section of the work
- Labelling the data using simple logical filters is prone to errors as they miss the nuances in the data and legitimate differences between turbines
- Old automatic labelling models also fail to identify new performance issues

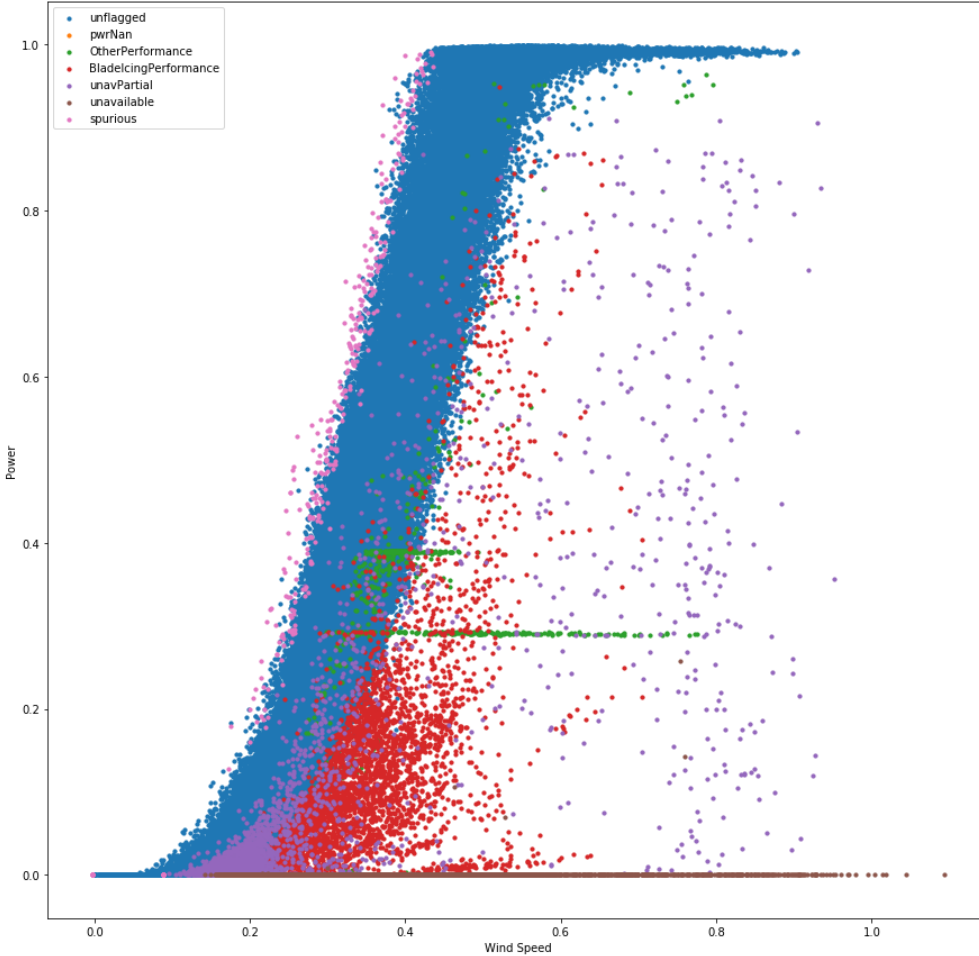
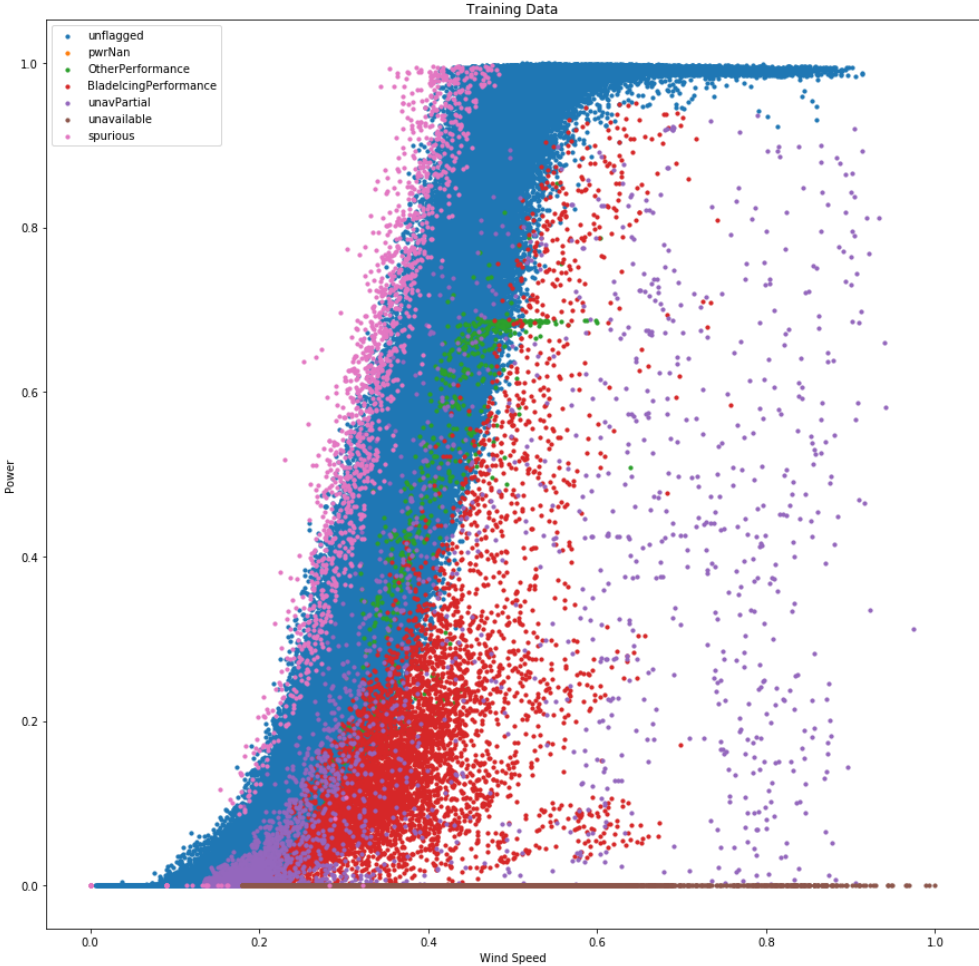
Machine Learning Labelling Tool

- A training dataset is prepared for each wind farm analysed by manually labelling a subset of the SCADA data from the wind farm
- Normalisation of power curves is necessary to create a consistent training dataset – this greatly improves the predictive performance across the wind farm
- The normalised data is used to train a **convolutional neural network**



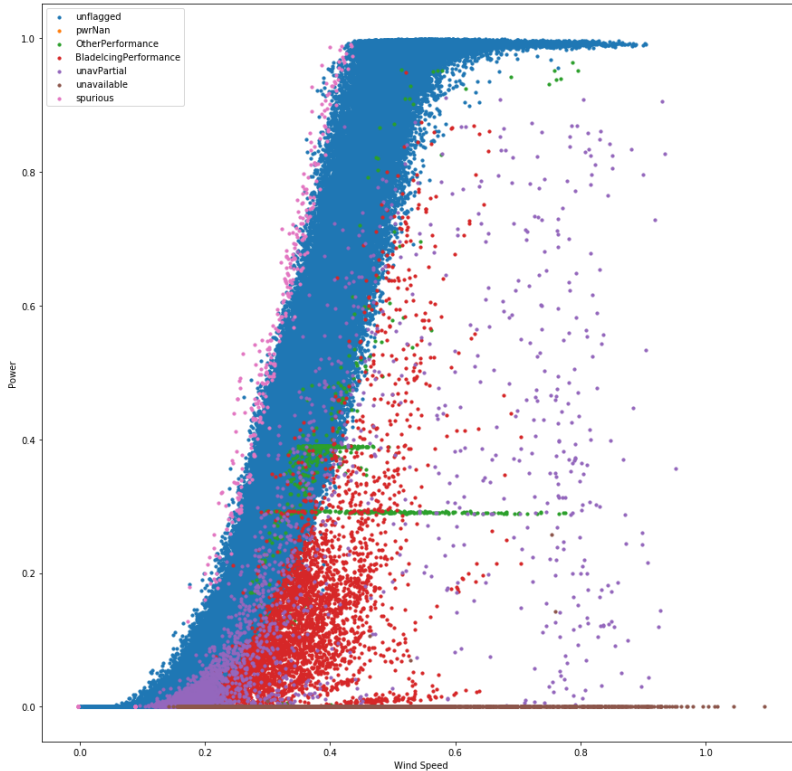
- Time is considered when training the model in order to capture downtime/performance events
- The model is then used to label the remaining data from the wind farm – provided the training data was high quality **the model produces consistent results efficiently**

Example: Turbine Blade Icing Loss



Example: Turbine Blade Icing Loss

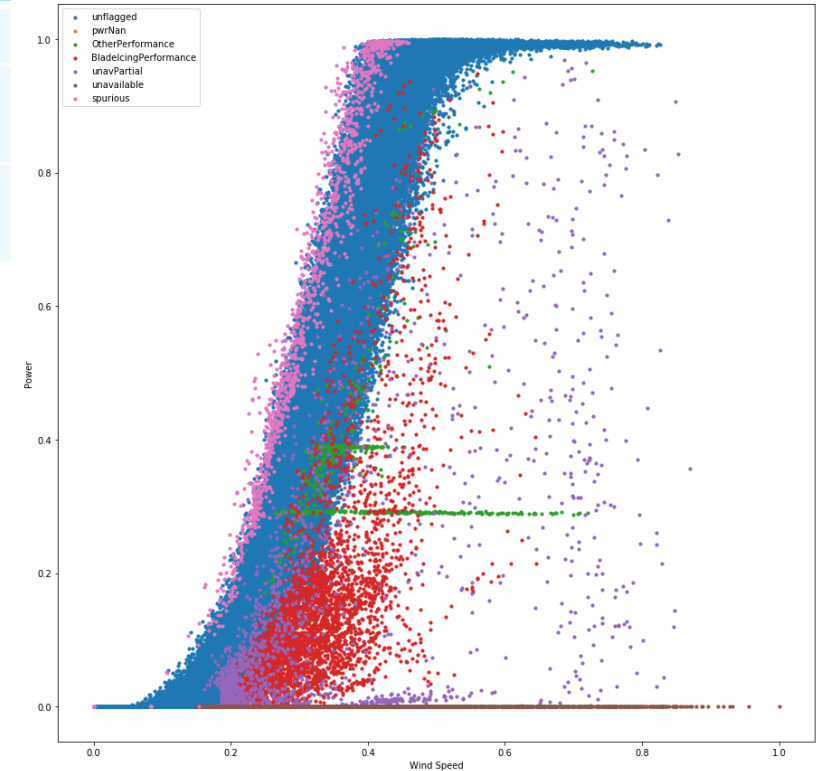
Machine learning labels



		Predicted	
		Blade icing	No icing
Actual	Blade icing	86.0%	14.0%
	No icing	0.3%	99.7%

- The model applies the learned flags with **high precision**
- Minimal manual flagging is required which can be carried out when checking the results

Manual labels

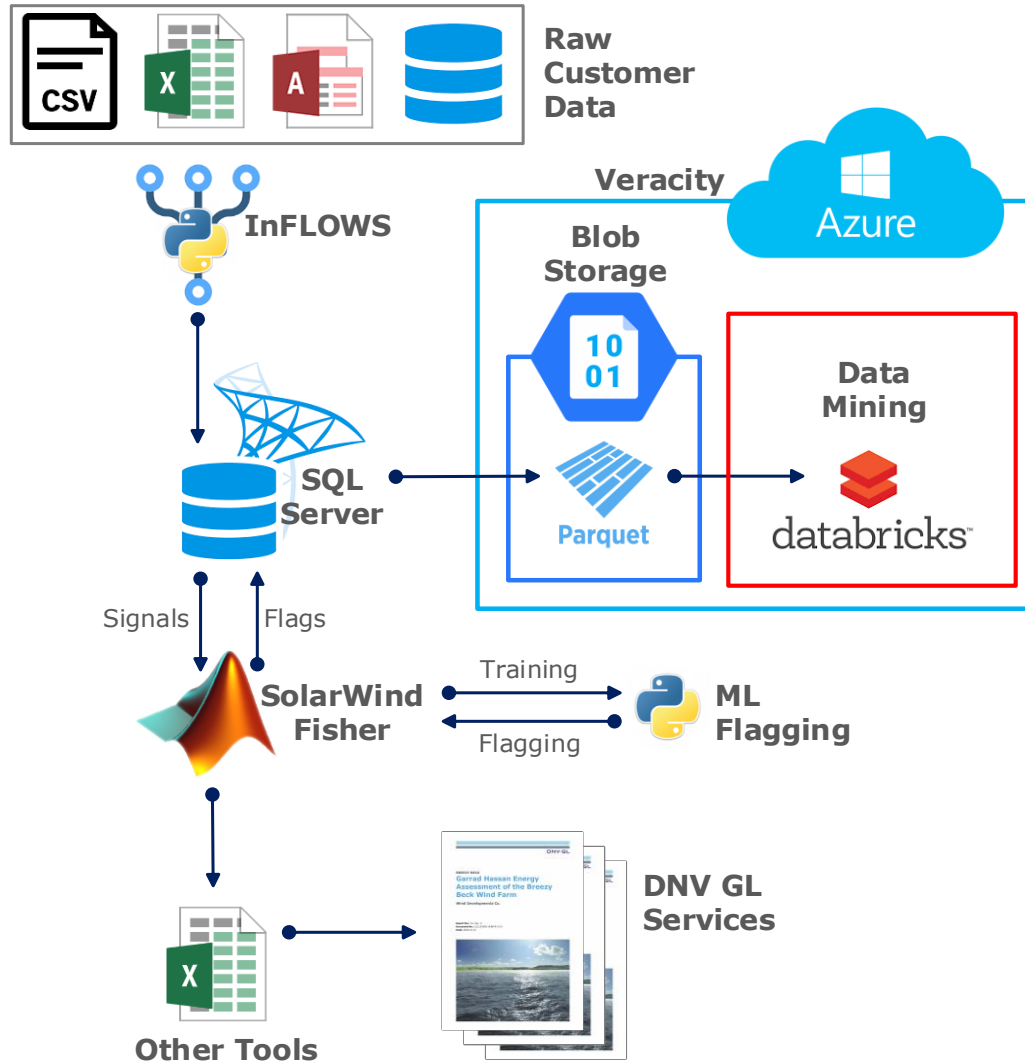


Data Mining Overview

600+ Windfarms,
6,000+ turbines and
20,000+ turbine years

- **50 GW** of operating wind and solar plant data
- A suite of data mining tools can be used for analysis of the entire portfolio.
- 'Big questions' can now be tackled efficiently using analytical methods
- **What questions would you ask 50 GW-worth of data?**

Data Mining Methodology

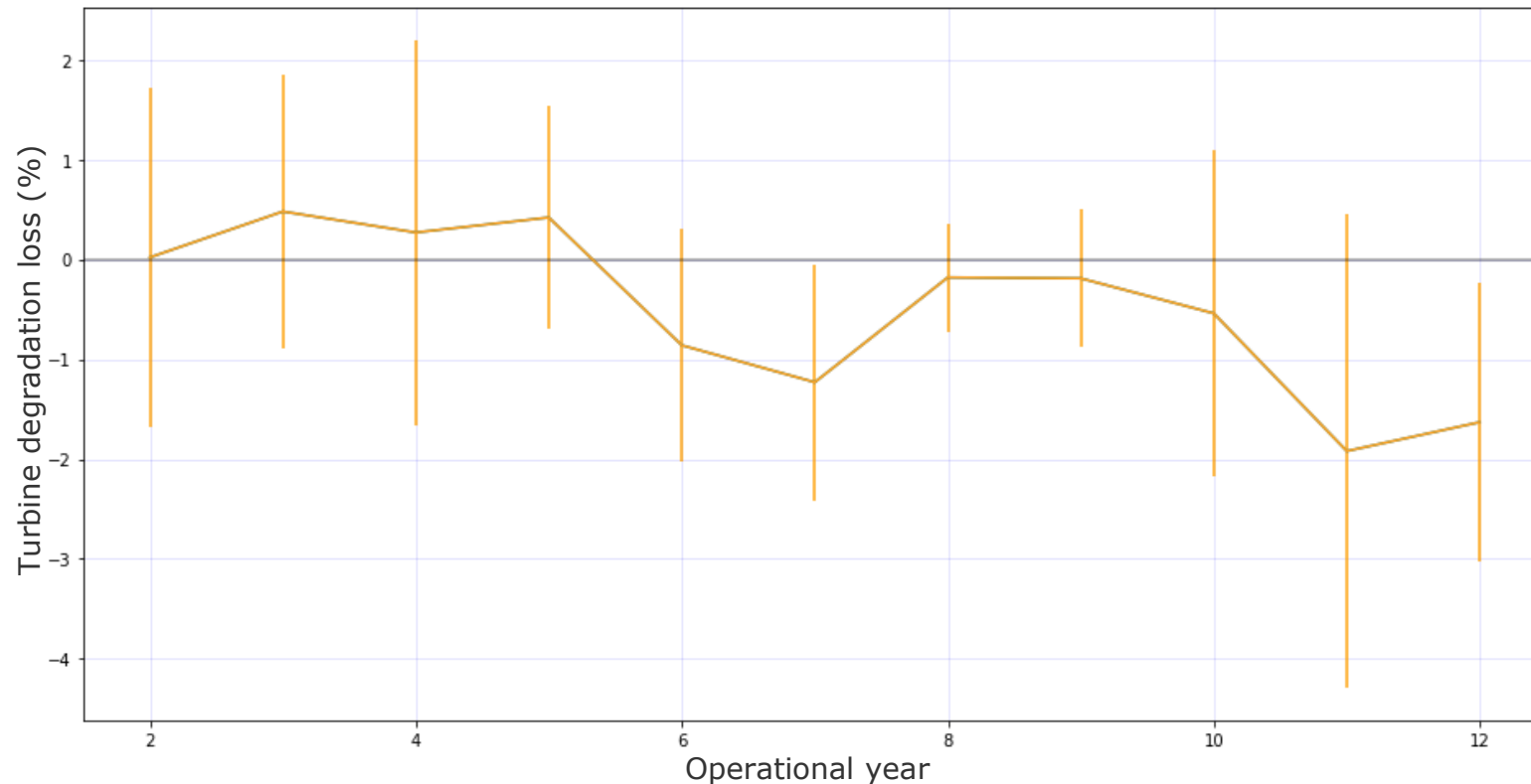


- 'InFLOWS' universal SCADA data ingest tool improving efficiency and standardising data.
- ML Flagging tool to speed up data labelling.
- Storage of standardised, cleaned and labelled datasets in Veracity data container.
- Data mining of Veracity data using Azure Databricks.

Example: Turbine Degradation Study

- Assumed turbine degradation rate is an important input to an energy assessment of a wind farm – with a **loss ranging from 1.0% to 1.3% over 20 years.**

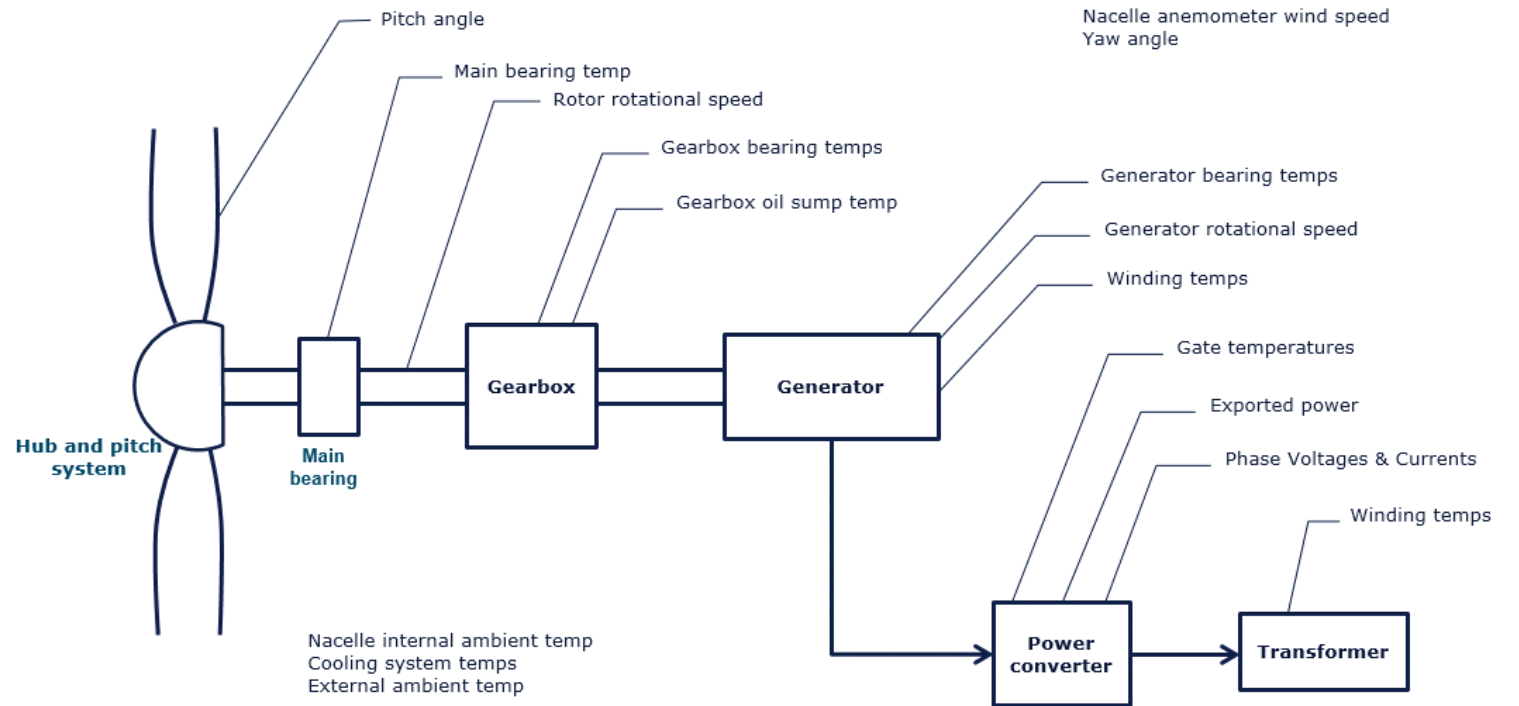
- This study confirmed our assumptions**
- More work is being done to investigate how projects located in the Nordic region align with our turbine degradation assumptions



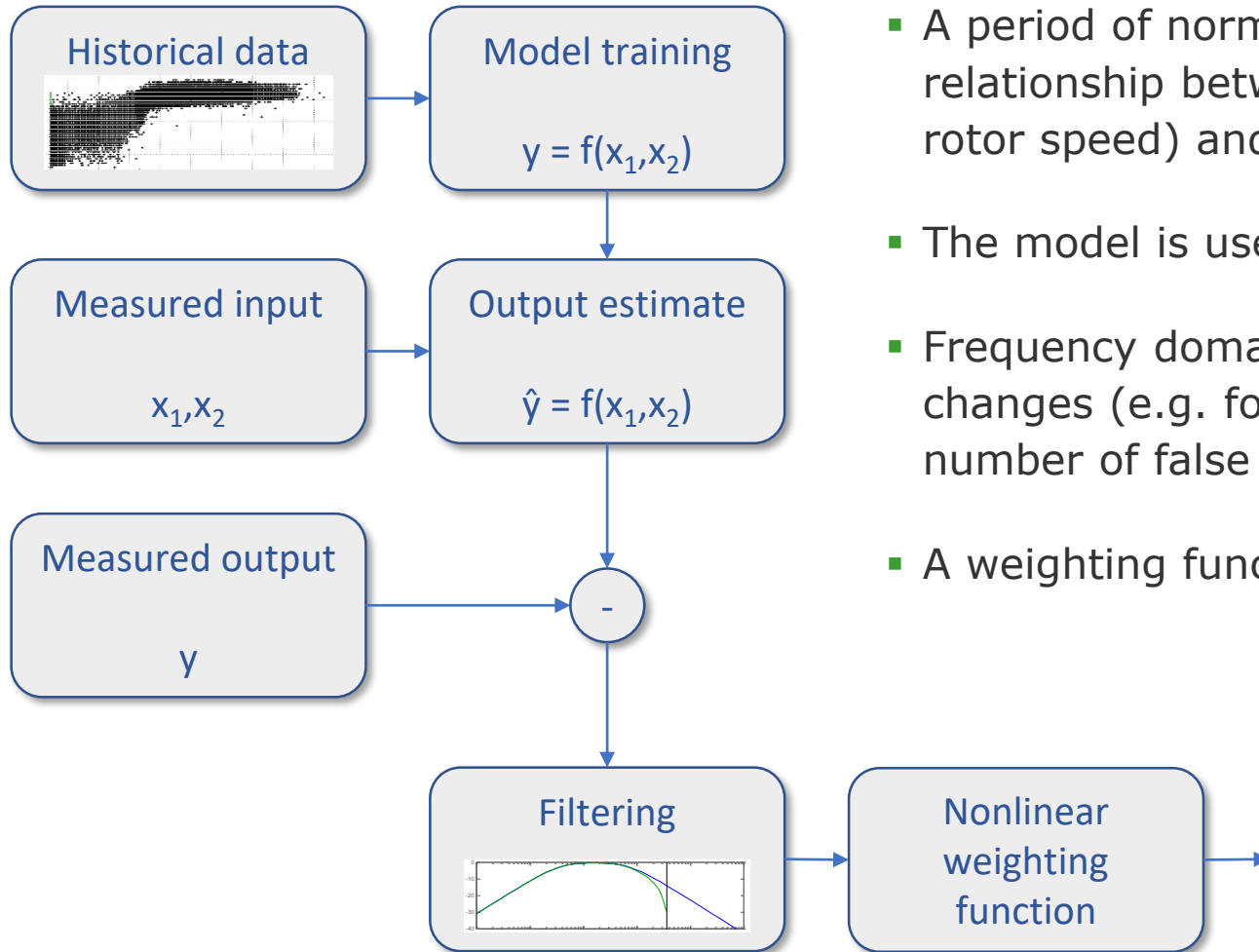
Year	Loss	WF_count	Turb_count	Std
2	0.02	241	6516	1.7
3	0.48	167	5362	1.37
4	0.27	98	3422	1.93
5	0.42	68	2996	1.12
6	-0.86	51	2484	1.17
7	-1.23	35	2333	1.18
8	-0.18	12	262	0.54
9	-0.19	7	74	0.69
10	-0.54	4	26	1.63
11	-1.92	3	17	2.37
12	-1.63	2	6	1.4

Drivetrain Integrity Monitor

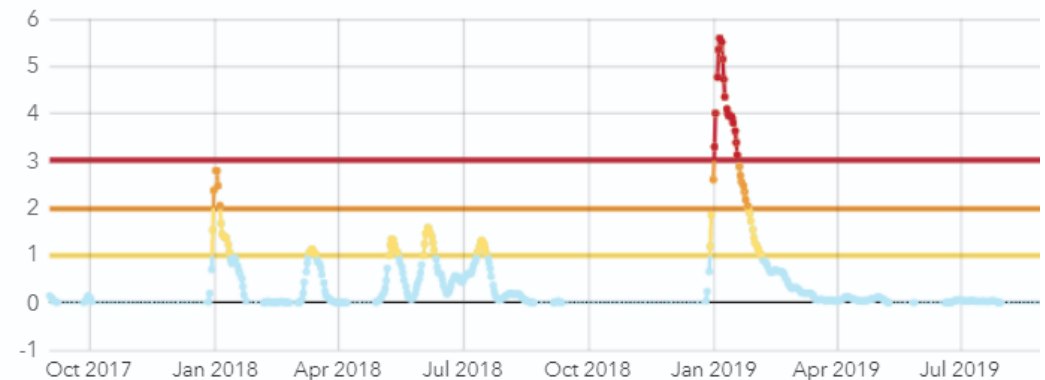
- One of the six modules contained within WindGEMINI
- Failure detection algorithm that uses existing 10-minute SCADA data
- Based on trending of the temperature signals from the wind turbine drivetrain
- The assumption is that an increase in temperature is indicative of dissipation caused by an anomaly
- No additional sensors are required



Machine Learning Approach

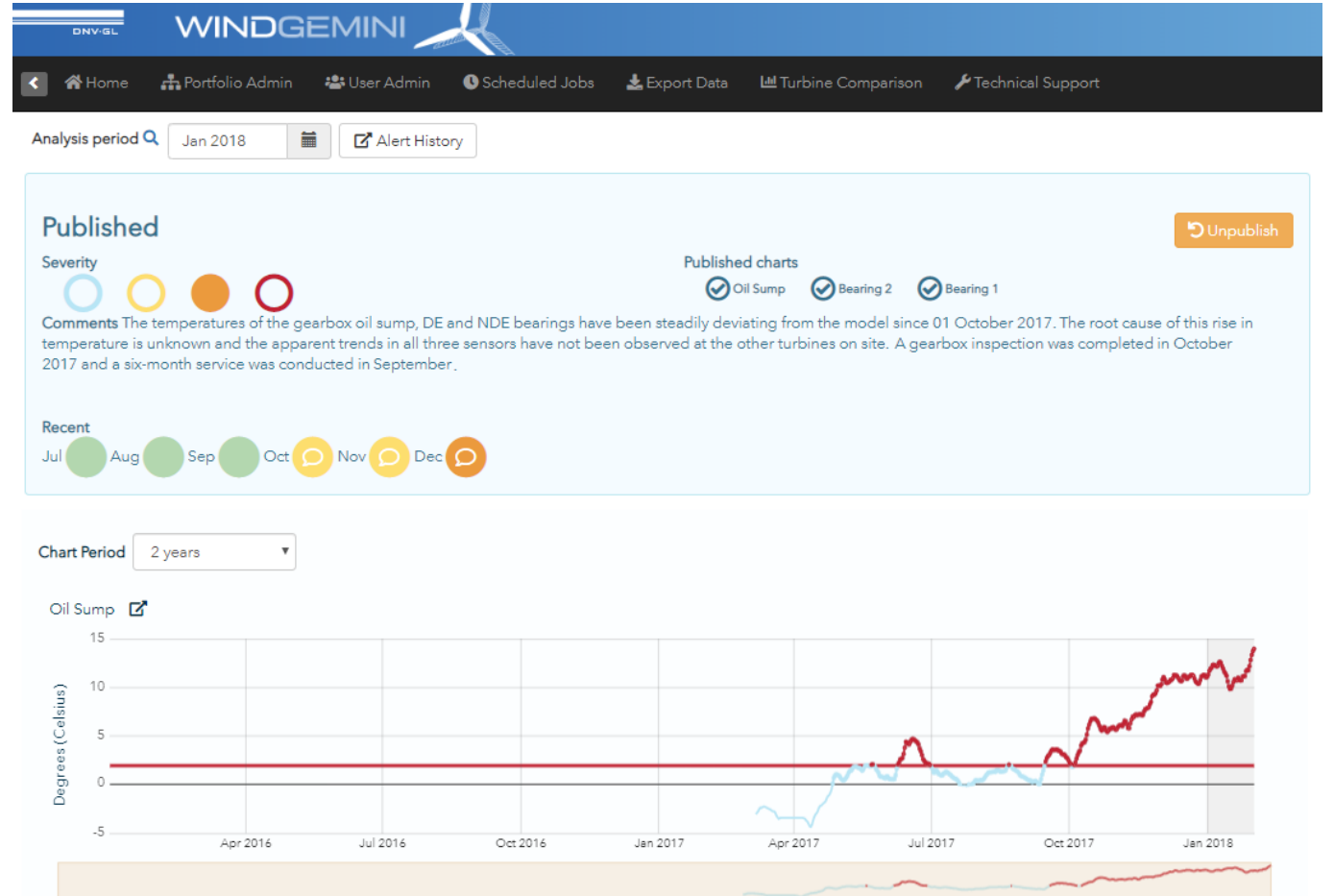


- A period of normal operation is used to establish an expected relationship between input signals (power, ambient temperature, rotor speed) and the drivetrain temperatures being monitored
- The model is used for ongoing monitoring of the real temperatures
- Frequency domain filtering reduces noise and the impact of step changes (e.g. following sensor replacements), reducing the number of false positives
- A weighting function automatically assigns a gravity to the alerts



Case study – October 2017

- In March 2017 a 2MW turbine experiences a gearbox failure
- WindGEMINI identifies an anomaly following the return to service
- Investigation showed one of the replaced components was defective
- Further analysis showed this was also having an effect on power curve efficiency (~94%)
- Issue was raised with the OEM and rectified, with an increase in performance and avoiding the risk of another failure and claim



Thank you

Will Jowitt

will.jowitt@dnvgl.com

+4420 3816 6816

www.dnvgl.com

SAFER, SMARTER, GREENER

The trademarks DNV GL®, DNV®, the Horizon Graphic and Det Norske Veritas® are the properties of companies in the Det Norske Veritas group. All rights reserved.