WHEN TRUST MATTERS

Development and Validation of Automatic Data Quality Control using Probabilistic Bayesian Neural Network

DeepWind 2023

Anish Venu (28)

- Senior Data Science and Domain Expert
- Digitalization & Research
- 4 years with DNV

Agenda

- Introduction
- Project Data – Research at Alpha Ventus (RAVE)
- General Background
- Automatic Data Quality Control using BNN
- Case Study – First Results
- Future Works

Introduction – Standard data quality control of measurement data

High frequency Raw data signals → Calibrated data signals → Descriptive statistics → Manual/visual check of statistical data → Flagging abnormal data signals

Disadvantages

- Time consuming process
- High measurement operational cost
- Impossible to check the high frequency measurements
- Immediate detection of measurement errors are not possible
- Extended measurement campaign
- Added uncertainty

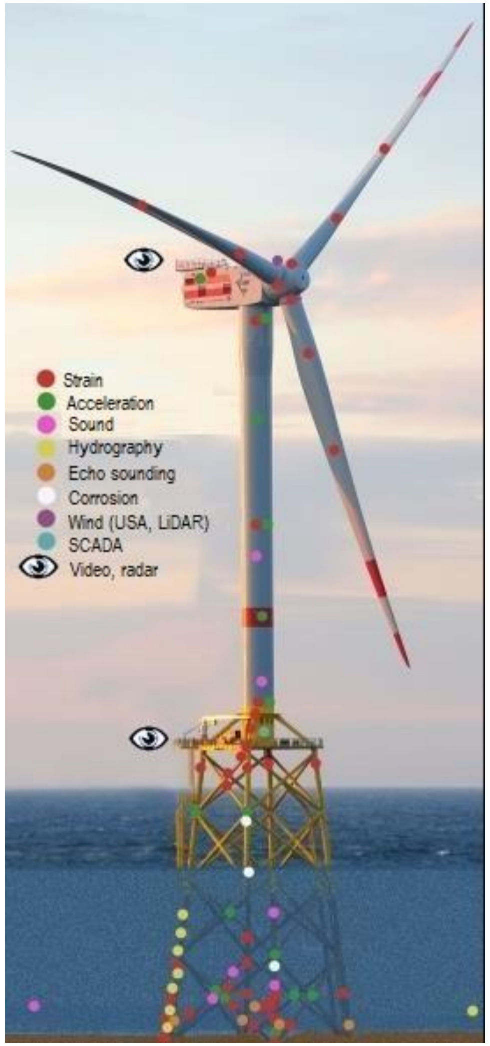
Project Data – Research at Alpha Ventus (RAVE)

- ➢ The research Initiative RAVE carries out research and development work on the offshore test field alpha Ventus.
- ➢ RAVE is funded by the Federal Ministry for Economic Affairs and Climatic Actions (BMWK) and coordinated by the Fraunhofer Institute for Wind Energy Systems (IWES).
- ➢ In more than 30 research projects, more than 60 partners from science and industry have been working on a wide range of research questions since 2008.
- ➢ The financial support from the BMWK so far amounted to more than 50 million euros.

Wind Farm Outlook

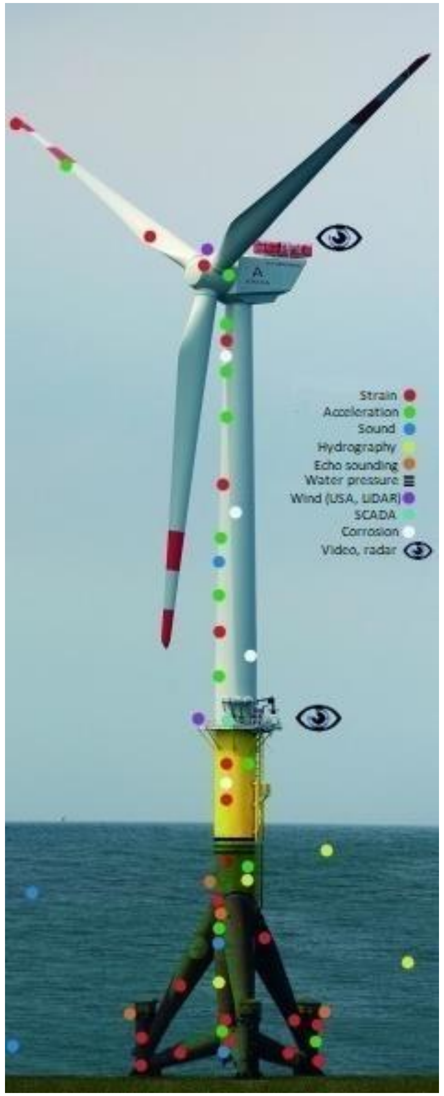
- ➢ 45 Km North von Borkum
- ➢ 30 m water depth
- ➢ 12 Wind turbines
  - 6 AREVA WIND M5000
  - 6 Senvion 5M
- ➢ CAPEX : 250 Million Euros
- ➢ More than 10 years of measurement data

Project Data – Research at Alpha Ventus (RAVE)



Available Measurements

- Controller Signals
- Acceleration sensors on the tower and blades
- Multiple strain guages on the tower and blades
- Wind measurements
- Atmospheric measurements
- Sea-State Measurements
- Other critical structural measurements
- Other electrical signals



General Background – Automatic Data Quality Control

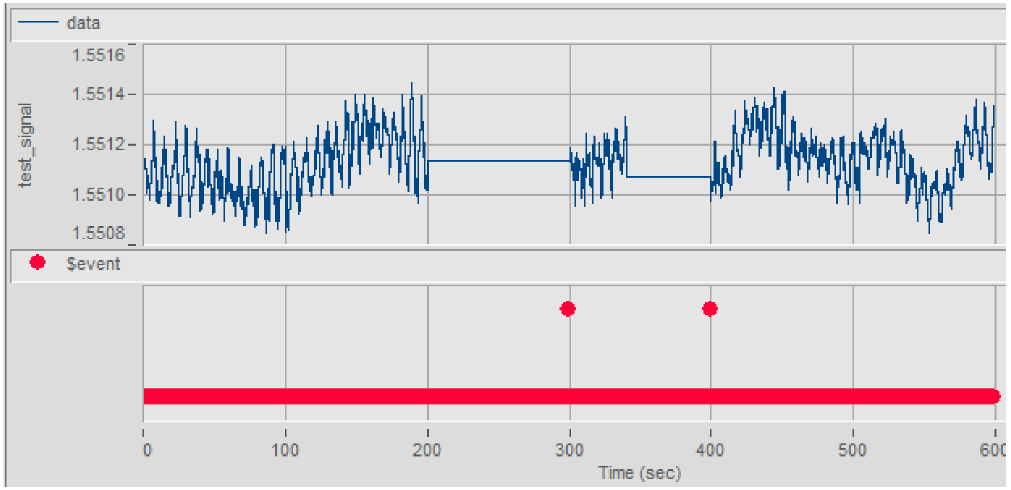
High frequency Raw data signals → Automatic data quality control → Calibrated Signals → Flagging abnormal data signals

Objective

- Control the data collected from RAVE wind farm
- Plausibility check on raw signals (0.2 to 50 HZ signals)
- Automating the control and flagging process
- Independent to sensor and measurement system
- Minimal input parameters (Robust model)
- Save time and operational cost
- High quality data for future applications

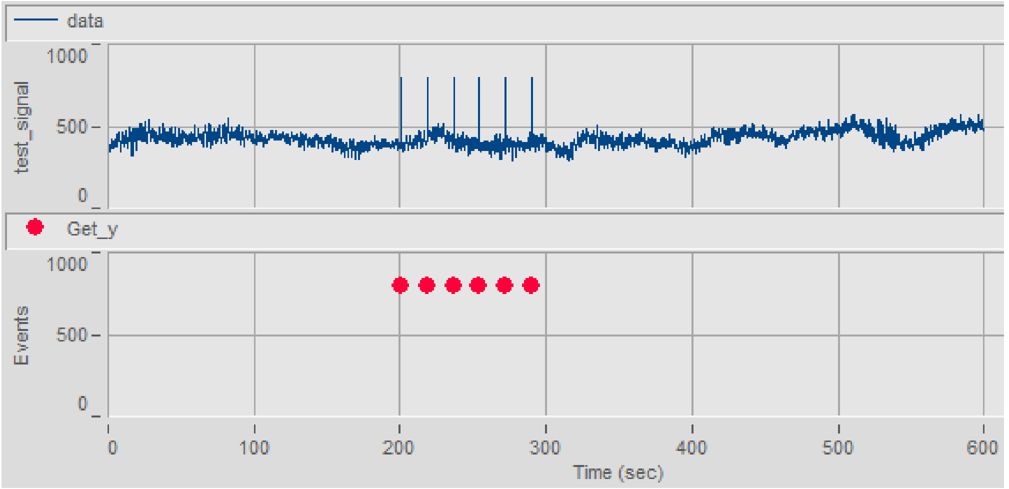
Position	Test Type	Meaning	Thresholds	Description
1	Length	Reduced data length	N_{crit}%	Data of length of some value N_{crit} deviating from N 100%
2	Flat Line	Constant Signal	N/A	All values the same (e.g. bad if sensor is strain gauge, Ok/Check if machine data)
3	Flat Line	Partially Constant	t_{crit}	Constant values for a period of > t_{crit} seconds (e.g. signal dropouts)
4	Pre-defined Limits	Measurement Range	$\sum (x_i > x_{crit}) > 0$	At least one value outside the measurement range (e.g. ±10 V)
5	Spike	Spike events exceeded	n_{crit}	Number of spikes found in signal exceeds critical value.
6	Spike	Low Correlation	r_{crit}	Despiked signal poorly correlated with uncorrected signal.
7	Visual/ Qualitative	Qualitative assessment	N/A	Data assessed manually (e.g. poor correlation with wind speed).
8–16	-	*- Spare -*	-	Further tests included here.

General Background – Automatic Data Quality Control



Flags : 001000/1

Master Flag

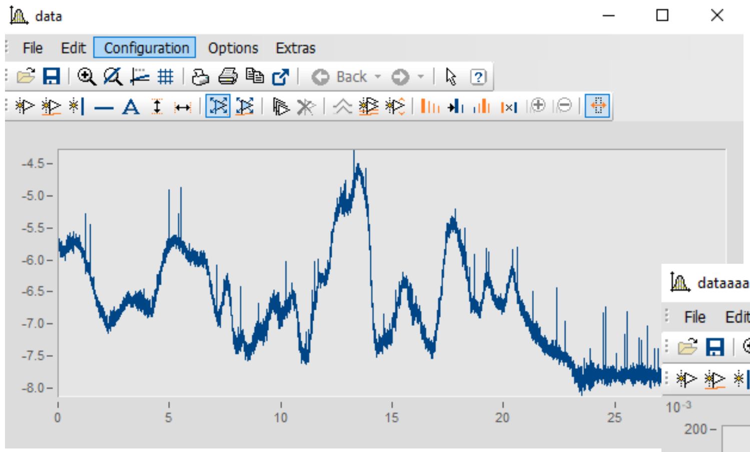


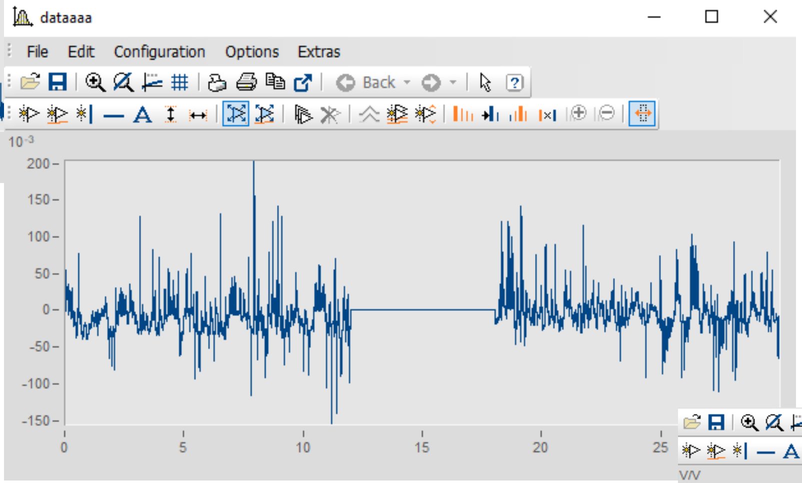
Flags : 000100/1

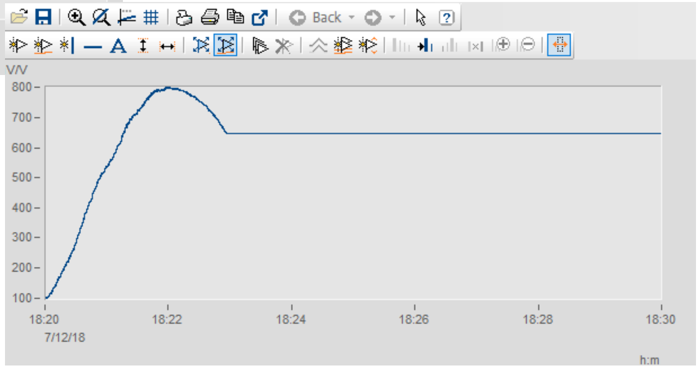
Master Flag

General Background – Automatic Data Quality Control

DQ	DR
Trot_hss_Nm	Master_check
000111	Fail
000111	Fail
000111	Fail
000111	Fail
000111	Fail
000111	Fail
000111	Fail
000111	Fail
000111	Fail
000111	Fail
000111	Fail
000111	Fail
000111	Fail
000111	Fail
010110	Fail
000111	Fail
000111	Fail
000111	Fail
000111	Fail
000111	Fail



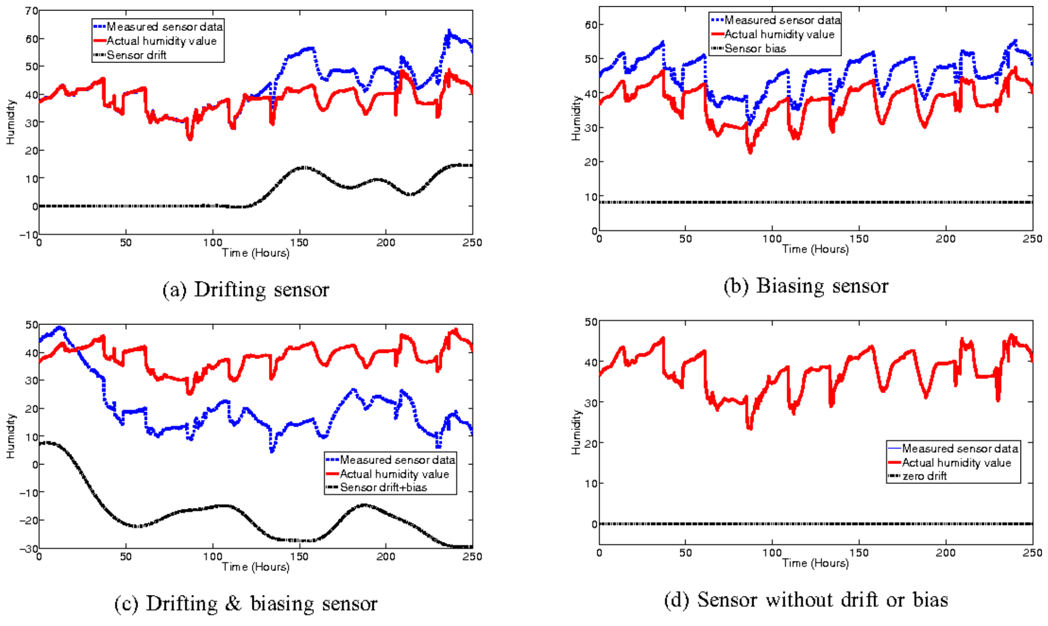




General Background – Automatic Data Quality Control

Limitations

- ➢ Detects only 70% of the commonly occurring events
- ➢ Time & environmental sensitive events are not detected
- ➢ Not using the historically available cleaned database
- ➢ No data filling/replacement method available
- ➢ No additional advantages

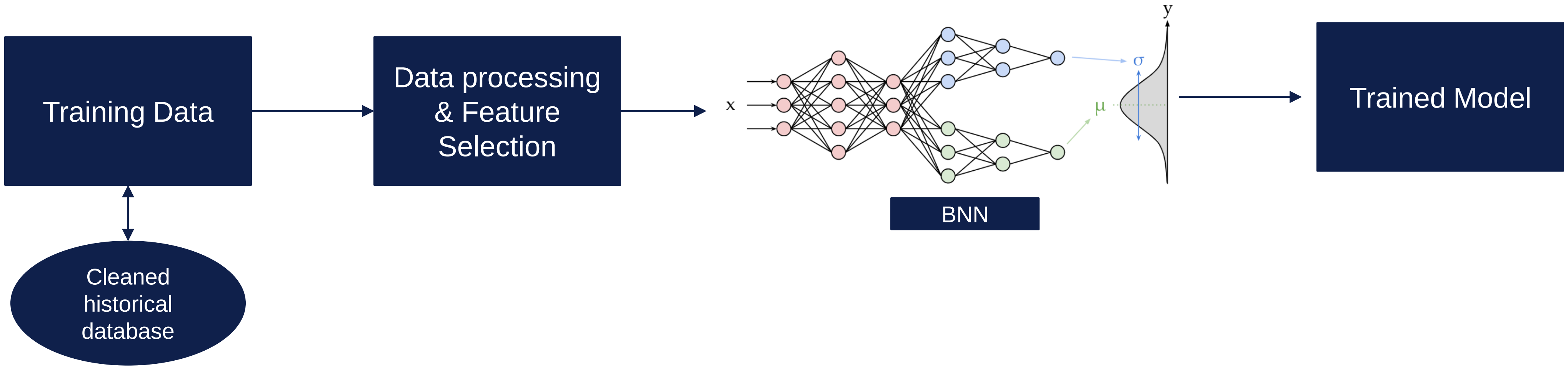


(a) Drifting sensor (b) Biasing sensor (c) Drifting & biasing sensor (d) Sensor without drift or bias

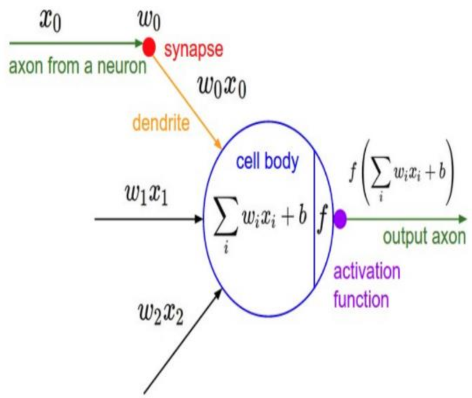
ADQC Output – 000000/0

No Events Found

Automatic Data Quality Control Using Bayesian Neural Network



Neutral network state of Art



Conditional Probability : Bayes Theorem

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Standard Neural Net / **Bayesian Neural Net**

Standard Neural Net

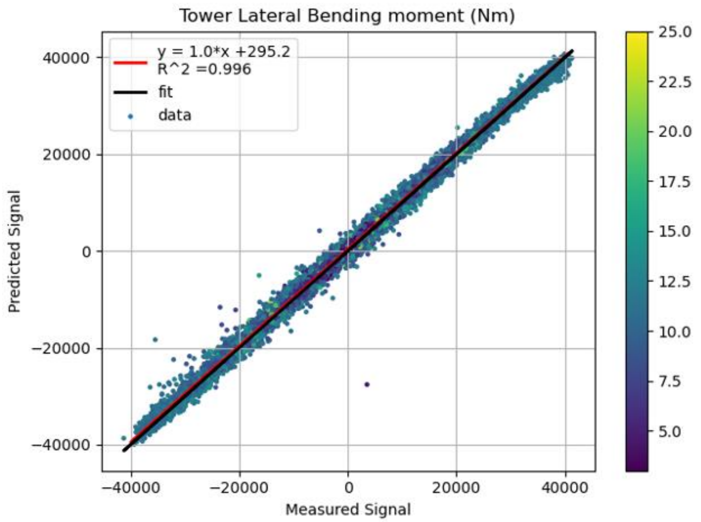
- Parameters represented by *single, fixed values (point estimates)*
- Conventional approaches to training NNs can be interpreted as *approximations* to the full Bayesian method (equivalent to MLE or MAP estimation)

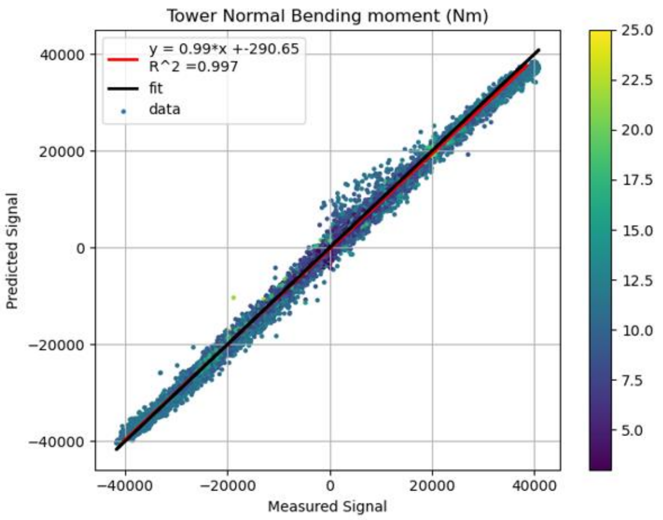
Bayesian Neural Net

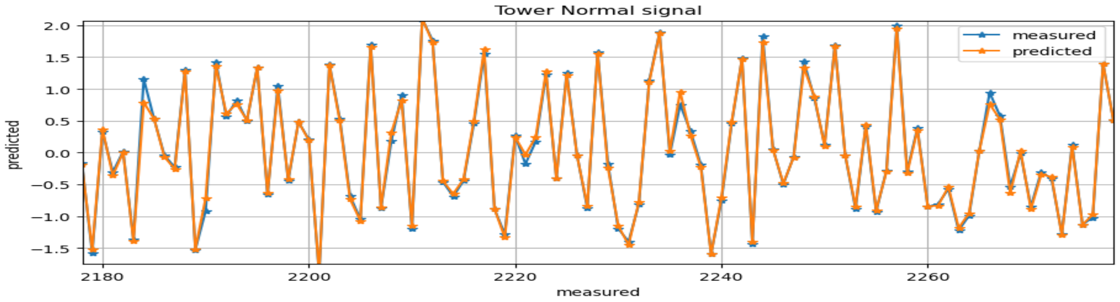
- Parameters represented by *distributions*
- Introduce a prior distribution on the weights $P(\mathbf{w})$ and obtain the posterior $P(\mathbf{w} \mid \mathcal{D})$ through Bayesian learning
- Regularization arises naturally through the prior $P(\mathbf{w})$
- Enables principled model comparison

Automatic Data Quality Control Using Bayesian Neural Network

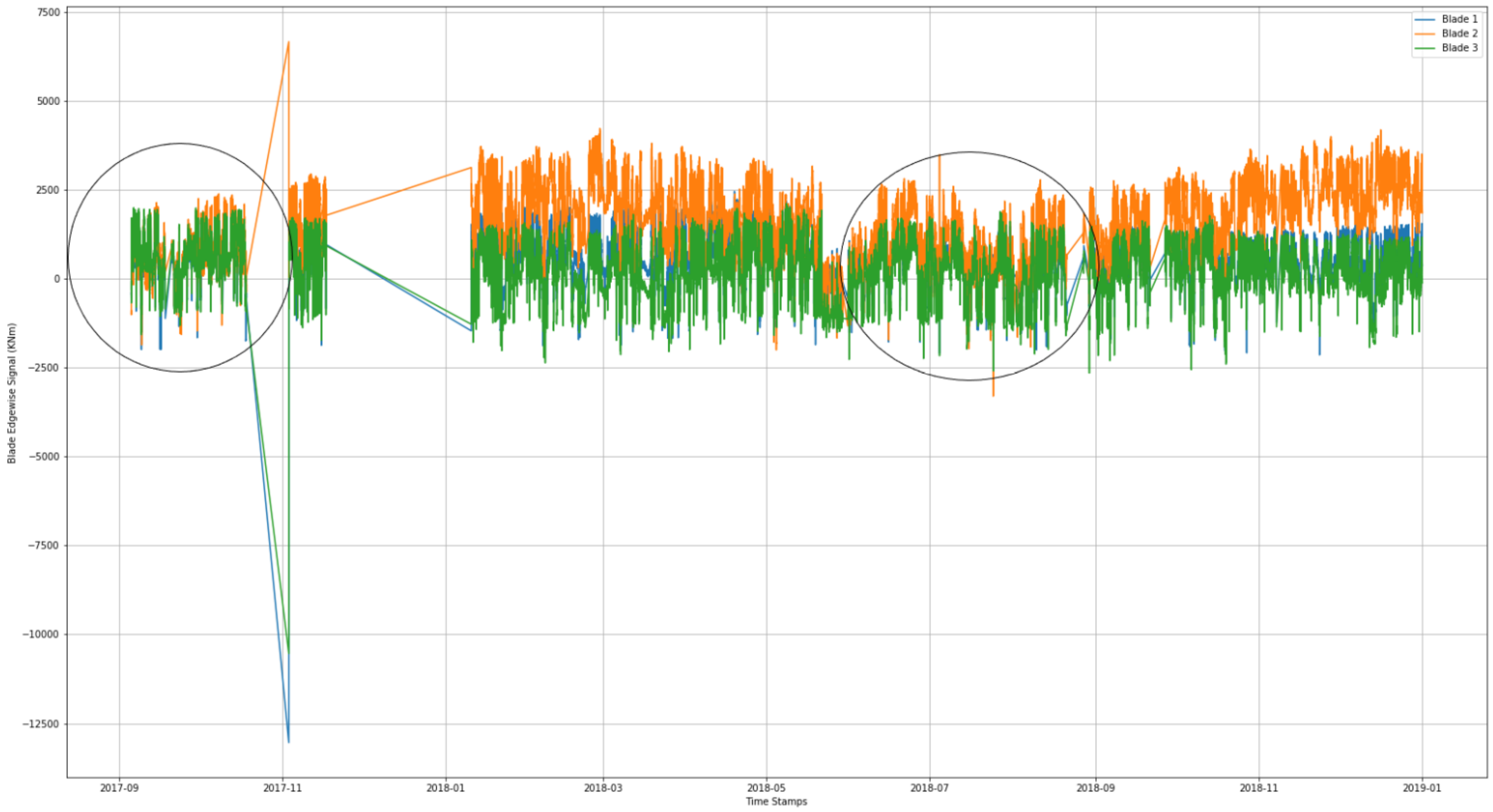
Example : Tower Signals







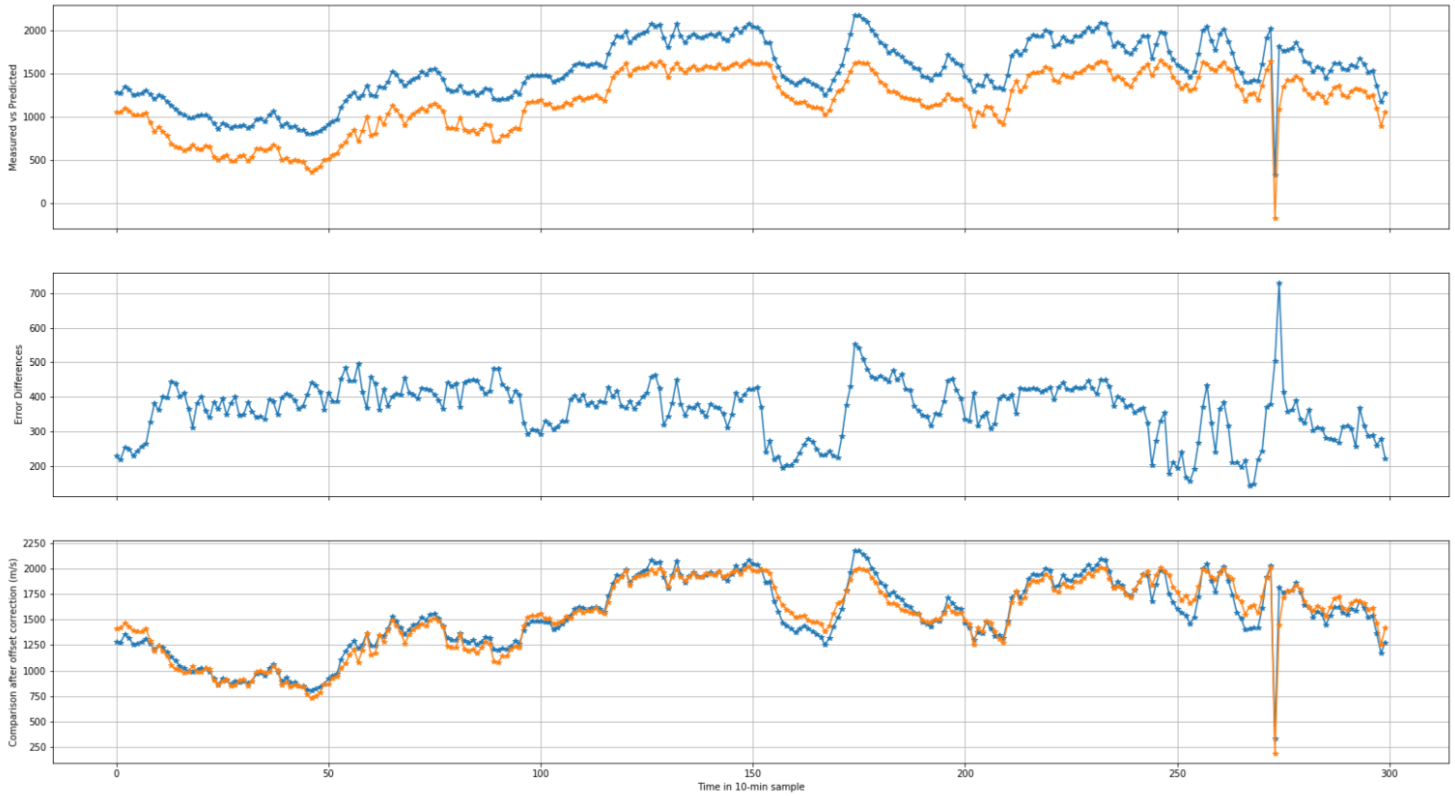
Case Study – First Results



- ➢ **Sensor installed and calibrated in Autumn (Black circles)**
- ➢ **Drifting problem in the other seasons**

Case Study – First Results

Detection of sensor drift in the blade signals due to temperature change



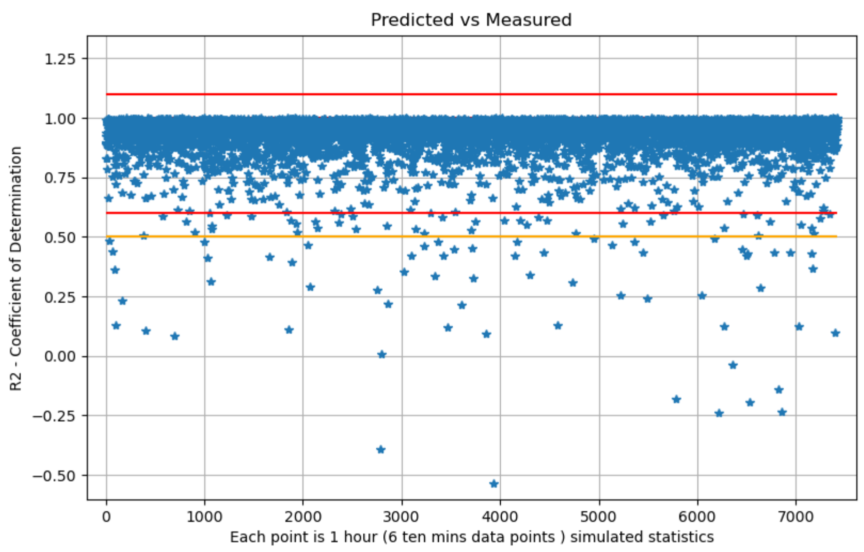
Measured vs Estimated

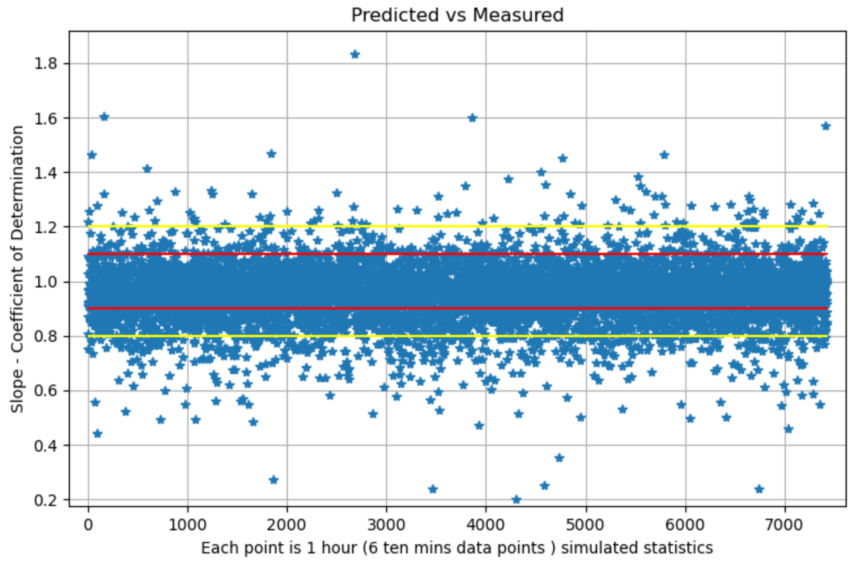
Error difference

Correcting the measured signal
based on prediction

Future Work

- Past performance based data flagging strategy





- Calculating the metrics (Predicted vs Measured) for 1 day window
- Comparing the calculated metrics with the historical performance
- Flagging the data falling outside the threshold range

Thank You

Anish Venu
E-Mail: anish.venu@dnv.com

Hans-Peter Link
E-Mail: Hans-Peter.Link@dnv.com

BUNDESAMT FÜR SEESCHIFFFAHRT UND HYDROGRAPHIE

Projektträger Jülich
Forschungszentrum Jülich

RAVE
RESEARCH AT ALPHA VENTUS

Bundesministerium für Wirtschaft und Energie