PHM Solution Development in Wind: Problems & Solutions

Jamie Godwin

October 16th 2013
New Orleans, Louisiana, USA.

Durham University

Overview

- Motivation & Challenges
- Case study # 1
    - PHM for wind turbine pitch faults
    - Data-driven, using maintenance logs
- Case Study #2
    - PHM for wind turbine gearbox
    - Statistical, using normal operational behaviour
- Discussion
- Conclusions

Motivation

- Low correlation between SCADA data and maintenance records
  - (Less than 5% of alarms have an associated maintenance record)
- Maintenance on a wind turbine represents 20-25% of total asset cost
  - Up to 75% of this is unscheduled maintenance
  - Preventive maintenance can be (up to) 40 times cheaper!
- Typically, ~25% can be saved with the proper application of a n next generation maintenance philosophy
- Lower than expected penetration of CBM into industry

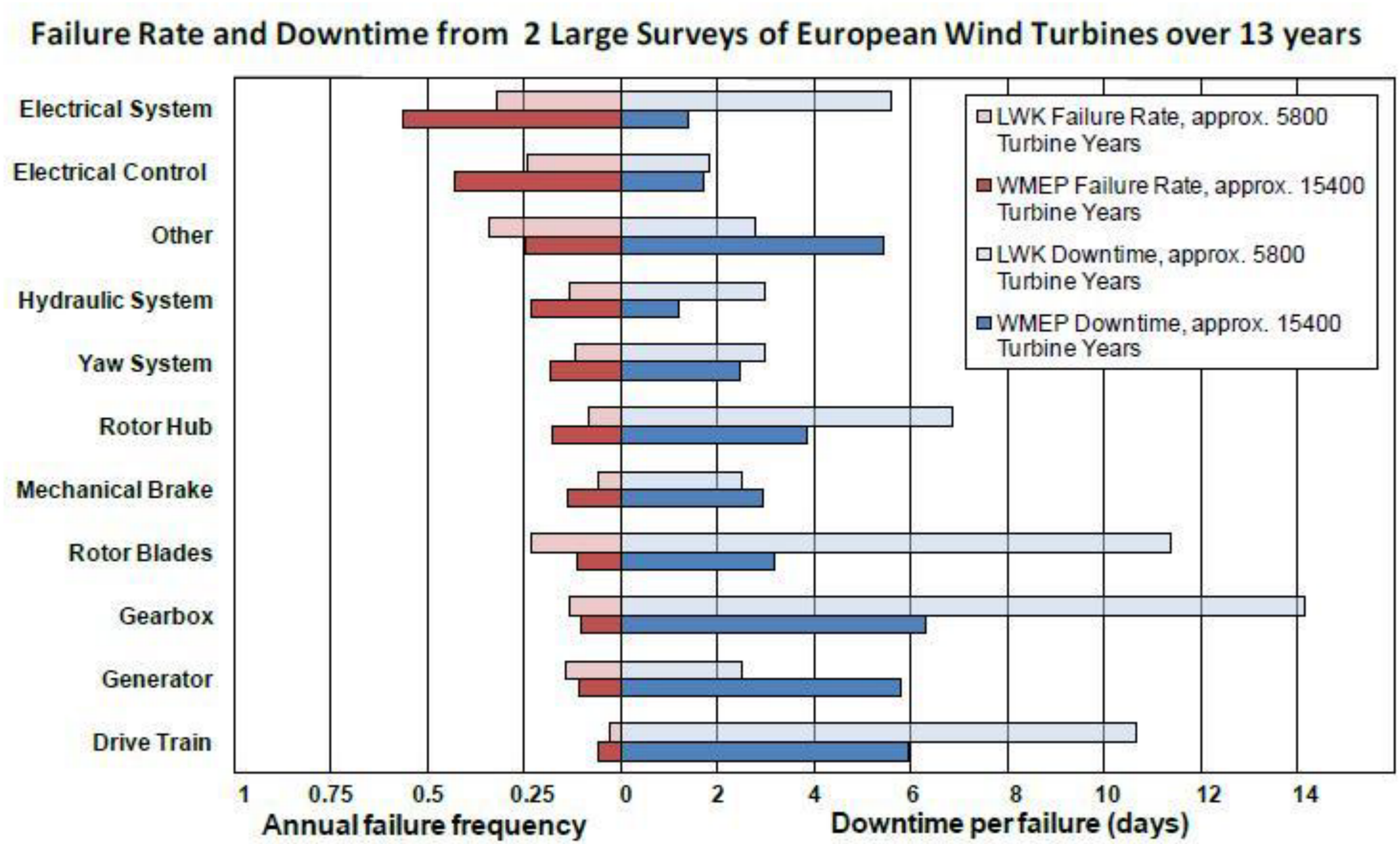
Challenges

- Data isn't labeled
  - We don't know how degraded components are
  - We don't know what attributes are relevant

- Text maintenance logs
  - Often incomplete
  - Inconsistent vocabulary

- Faults may go unrecorded in archive data
  - Inclusion in PHM development will degrade performance

Motivation



Crabtree (2010)

Example Data

Time Stamp	Average Wind Speed	Max Wind Speed	Motor 1 torque maximum	Motor 2 torque maximum	Pitch torque average	Blade 1 angle	Blade 2 angle	alarm
01/01/2001 12:00	8.9	11.9	53.579998	53.029999	4.85	2.03	2.03	No
01/01/2001 12:10	10.2	14.6	61.369999	60.23	9.07	2.36	2.36	No
01/01/2001 12:20	1.1	1.7	0	0	0	85.919998	85.979996	No
01/01/2001 12:30	16.5	22.7	56.43	55.649998	51.610001	21.549999	21.549999	No
01/01/2001 12:40	9.6	12.4	44.119999	39.02	4.56	1.98	1.98	No
01/01/2001 12:50	14.8	20.5	49.369999	42.259998	19.959999	11.28	11.28	No
01/01/2001 13:00	7.1	9.1	17.83	17.779999	9.74	1.99	1.99	No
01/01/2001 13:10	9.2	14.3	39.93	33.469997	24.48	2.04	2.04	No
01/01/2001 13:20	12.8	19.2	68.019997	71.790001	47.289997	8.07	8.07	No
01/01/2001 13:30	1	2.6	0	0	0	0	-89.759995	No
01/01/2001 13:40	11.9	16.1	55.110001	59.989998	14.73	5.72	5.72	No
01/01/2001 13:50	4.4	6.2	13.88	12.21	4.17	1.98	1.98	No
01/01/2001 14:00	6.5	8.7	24.129999	23.619999	6.48	1.99	1.99	No
01/01/2001 14:10	4.4	5.4	61.189999	63.829998	0	70.659996	70.639999	No
01/01/2001 14:20	3.2	5	51.259998	52.02	7.15	2.01	2.01	No
01/01/2001 14:30	10.6	14.2	62.66	57.719997	7.86	2.41	2.41	No
01/01/2001 14:40	3	4	25.689999	29.619999	0	1.98	1.98	No
01/01/2001 14:50	12.6	17	63.93	63.02	41.200001	8.41	8.41	No
01/01/2001 15:00	6	8.2	0	0	0	86.169998	86.080002	yes
01/01/2001 15:10	7.4	8.9	18.369999	16.279999	6.85	1.98	1.98	No
01/01/2001 15:20	10.9	15.6	56.079998	56.84	13.23	4.15	4.15	No
01/01/2001 15:30	15.9	21.9	65.029999	65.639999	36.25	14.549999	14.549999	No
01/01/2001 15:40	7.2	10.5	16.6	18.76	6.62	1.98	1.98	No
01/01/2001 15:50	2.1	3.8	0	0	0	85.979996	85.939995	No
01/01/2001 16:00	2	2.7	0	0	0	85.959999	86.029999	No

144 records per day, across 190 channels (7 shown)
Up to 100 wind turbines on a farm

… SCADA systems aren't perfect!

Case Study #1

- Needed to reduce pitch fault alarms from the SCADA system

- Large drain on maintenance resources

- 3 sources of data available – SCADA data, SCADA alarms, maintenance records

- Hypothesis:
  - Can we automatically identify pitch faults, and we can determine if they are false positive?

Wind turbine pitch fault

- A deviation of the wind turbine blade angle from a pre-defined optimum
  - Modern wind turbines feather the blades to regulate power generation

- Faults can be due to pitch motor degradation
- Or they an be due to electrical system failure/malfunction

- Each blade angle should be identical (but this may not always be the case)

Motivation

- Wind turbine pitch fault represents the most common SCADA alarm on the wind turbine
  - Up to 45% (!) of SCADA alarms are pitch system related
  - Alarms can switch off turbine to prevent damage

- In some cases, the alarms are active for over 100 days.
  - ~1,700 alarms per year (> 4 per day!)
  - Large drain on maintenance resources to analyse all alarms

- Difficult to determine pitch fault through SCADA analysis
  - SCADA systems have many imperfections

- Remote reset can be a cost effective strategy! (If it’s safe to do)

Data description

- 8 Wind turbines used in analysis
  - ~ 1 million SCADA records (10 min intervals, 28 months)
  - 243 recorded pitch faults in the maintenance log
  - Over 20,000 SCADA records with pitch fault alarms

- All wind turbines are from the same wind farm
- All wind turbines are the same model

- Attributes determined by entropy & expert guidance (after labeling)

Data labelling

- 3 classifications were derived:
  - No Pitch fault present
    - All data not in the other categories
  - Potential pitch fault
    - Data associated with a SCADA pitch alarm
  - Established pitch fault
    - SCADA records directly associated to a maintenance action in the maintenance log (within 48 hours)

- These labels allow traditional data mining to be undertaken

Model selection

- 4 wind turbines used for training RIPPER (of the 8 available)
  - Any classifier could have been used
- Classes were balanced to remove majority bias
- 70 models developed (8 choose 4)
  - Ensure methodology is not sensitive to training data
  - Rule accuracy: 69.99% - 87.41% (*M*=82.70%, *SD*=4.26%)
  - Rule base: 6 – 38 (*M*=16.5, *SD*=7.65).
- Weak correlation (r=.056) between number of rules & accuracy
  - Beneficial to choose a smaller rule base which is easier for domain experts to understand.

- 21 Models were dominant for their rule base size
- Chosen model had 14 rules with 85.50% accuracy

Classifier Post-processing

- In order to filter SCADA data & remove noise, post-processing was performed
- Needed to ensure persistence
- A 90 minute threshold was set
    - Partly due to past experience
    - Partly due to analysis of the SCADA data
    - Partly due to expert knowledge

- If the threshold was breached, an alarm was raised

Results

- Post-processing provided filtering of SCADA alarms
  - Number of alarms was reduced by 35.80% - 52.26% (*M*=44.69%, *SD*=6.62%)
  - Average alarm length was reduced by 28.06% - 49.90% (*M*=35.68%, *SD*= 8.60%).

- 74 of 85 Maintenance actions were identified by the expert system (>87% accuracy) – 11 maintenance actions were missed
  - 7 of these 11 were due to missing data from the SCADA system
  - The remaining 4 are currently under investigation

Case Study #1 Conclusions

- Able to significantly reduce the number of SCADA alarms
- Able to significantly reduce the length of SCADA alarms

- Strong model classification accuracy (>85%)

- High model diagnosis accuracy (>87%)

- Needed significant quantities of data
- Needed a physical model
- Used maintenance logs to guide system

Case Study #2

- Need to identify gearbox degradation to enable efficient maintenance strategies

- Very limited failure data available
  - 1 failure from >14,000 hours of data
  - ~$5million if gearbox fails – excessive maintenance performed.

- New paradigm is required to identify degradation

Wind turbine gearbox

- So much data is collected (>14,000 records/day)
  - Why not use it?!

- Traditional data mining tries to encapsulate failure conditions
  - We have more "normal" data ... can we get more information out of this?
  - (We can!)

Developing a condition index

- If we can determine "normal" behaviour; we can measure deviations from this.

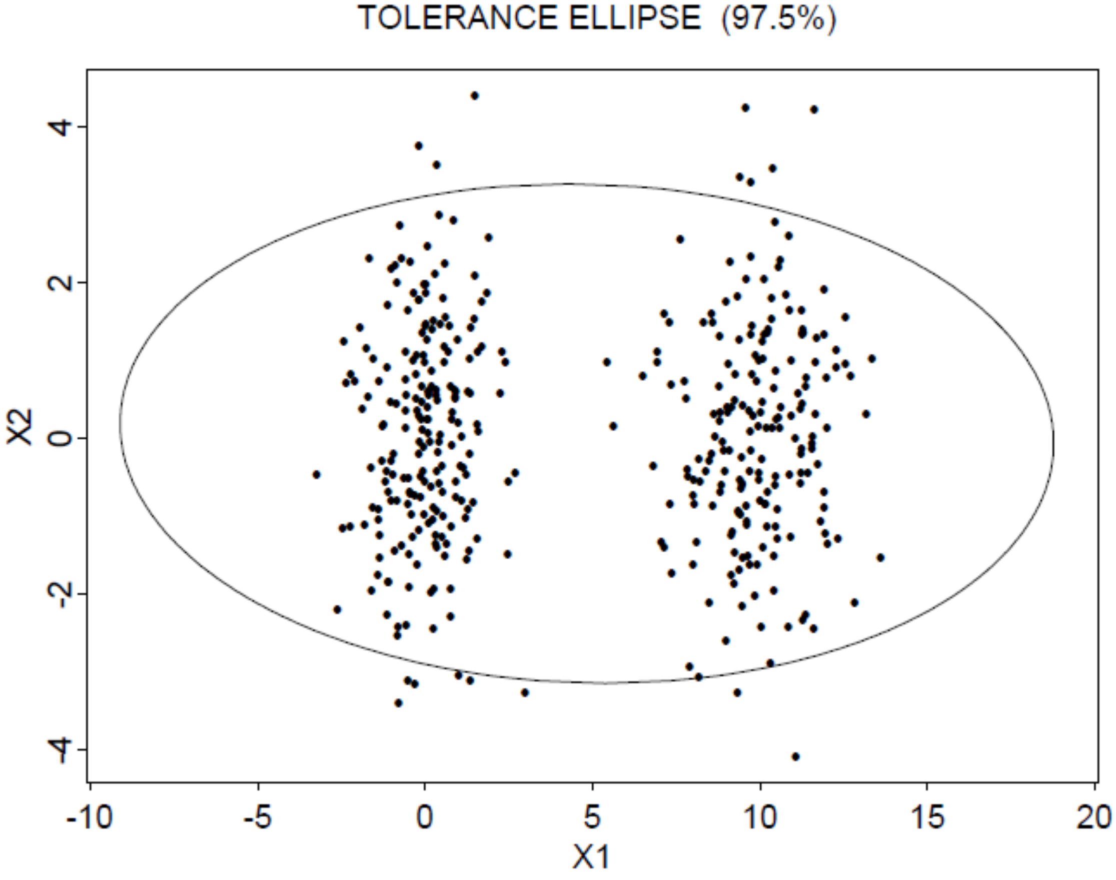
- More "normal" data, means a stronger understanding of this behaviour.
  - It's a win-win situation!

- Can use multivariate distance metrics, such as the Mahalanobis distance (or robust derivatives):

$$RMD_i = \sqrt{(x_i - \hat{\mu}')^T MCD^{-1}(x_i - \hat{\mu}')}$$

Why the MCD?

- Estimation of covariance is sensitive to noise

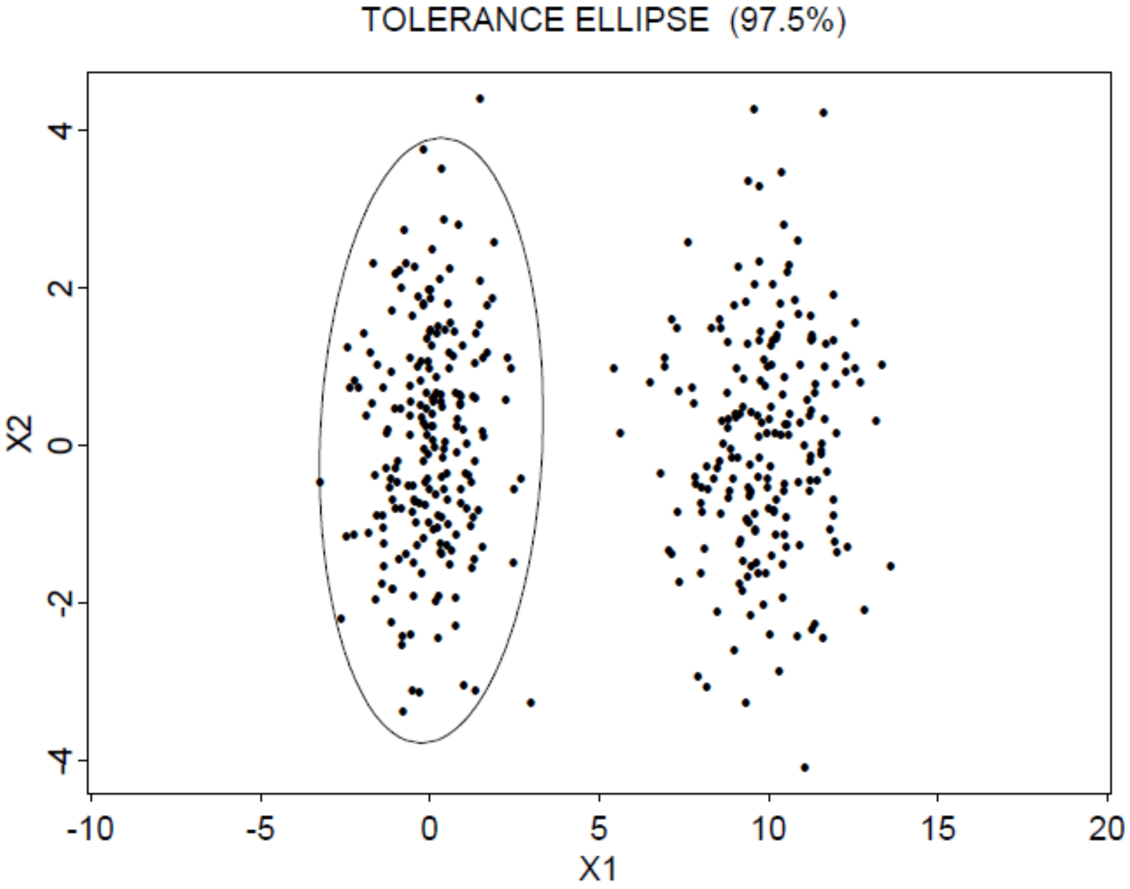


Taken from Rousseeuw & Van Driessen (1999)

(bivariate)

Why the MCD?

- Estimation of covariance is sensitive to noise



Taken from Rousseeuw & Van Driessen (1999)

(bivariate)

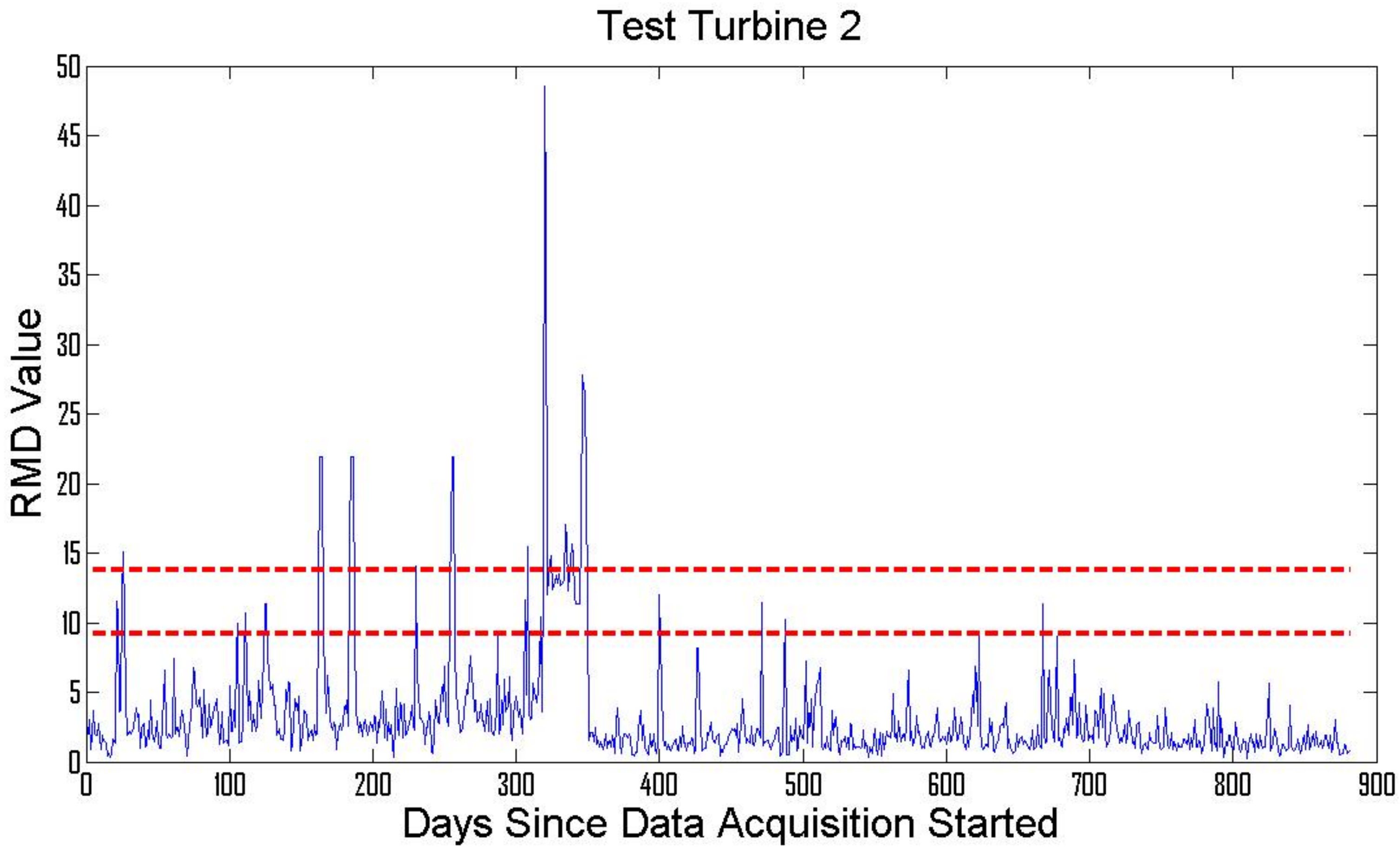
Model Attributes

- A (primitive) physics of failure model is utilised to determine attribute condition

- As the gearbox degrades, inefficiencies are created
- These inefficiencies cause increased friction
- This friction manifests as heat
- This heat is read by the sensors on the SCADA system

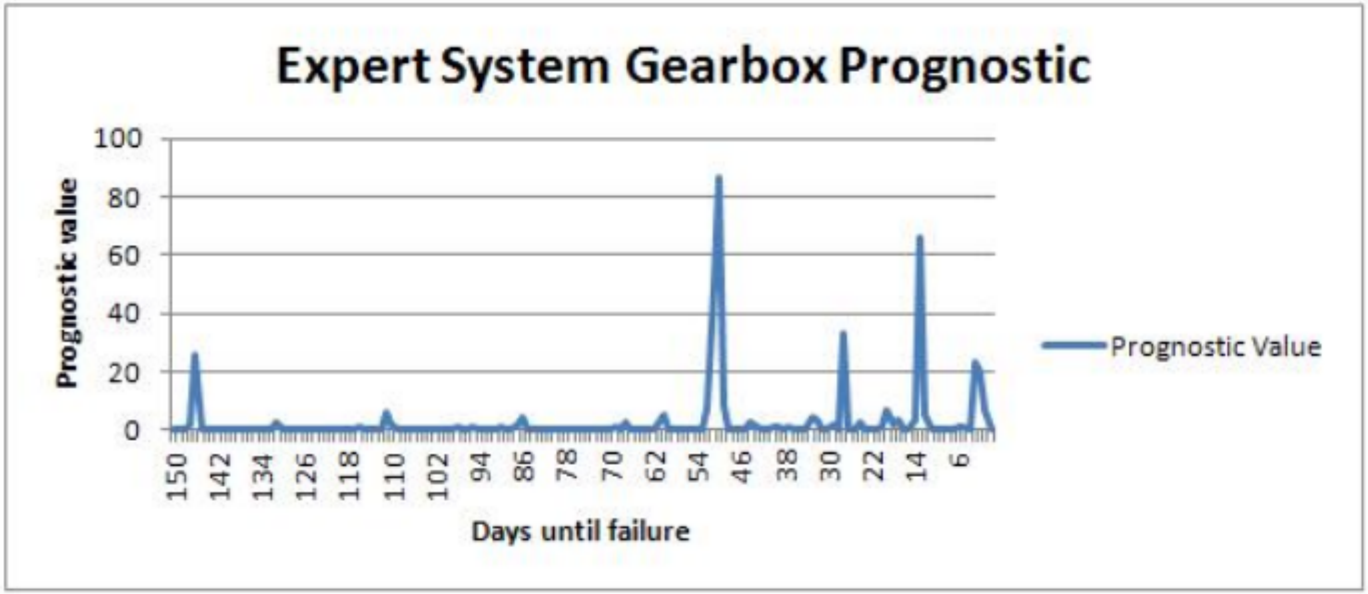
- Data is normalised for ambient temperature and loading

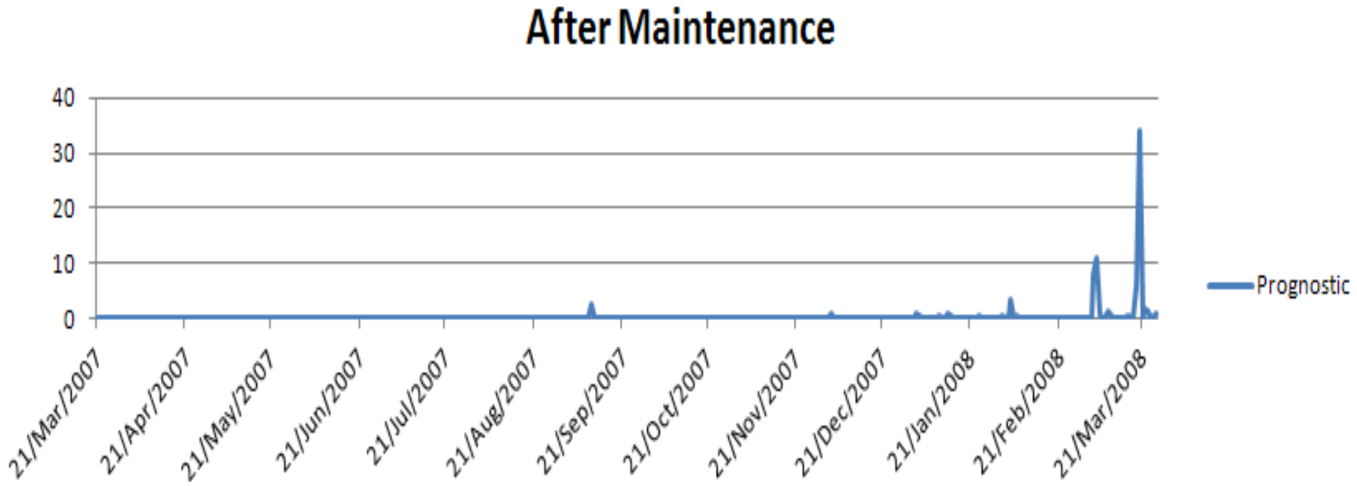
Condition Index



Thresholds determined based upon statistical properties of the RMD.

Condition Index – Smoothing





Condition Index

- Normal operational behaviour: ~94% of the time
- After maintenance, gearbox remained normal for 362 consecutive days
- Running in of the new gearbox is present in the data (not shown)
- 14 opportunities to inspect gearbox – first 6 months before failure

Condition Index

- Can accurately identify maintenance events in the data
- Can quantify the effectiveness of the maintenance performed
- Computationally tractable (can be performed on-line)
- Can be used for fault identification, RUL prediction, prognosis
    - Using ANN/SVM/RVM/regression etc.

Novel extras – rule extraction

•Can use statistical levels to provide class labels to enable data mining.

	Test Turbine 1	Test Turbine 2	Test Turbine3	Average
Normal Operation	860 (97.62%)	821 (93.12%)	829 (94.10%)	846 (96.03%)
Inspection Suggested	12 (1.36%)	32 (3.86%)	32 (3.63%)	21 (2.39%)
Potential Damage	9 (1.02%)	26 (2.95%)	20 (2.27%)	14 (1.59%)

Novel extras – rule extraction

- **Example Rules (from model using RIPPER):**
- *Planetary Gear Temperature >= 51.20 Degrees*

  Then Potential Damage
- *Rotor Speed >= 17 RPM*

  Then Potential Damage.
- *Energy Generated <= 0.33 MW and Rotor Speed >= 11 RPM*

  Then Potential Damage

Thank You

- Any Questions?

- E-Mail: j.l.godwin@durham.ac.uk

Durham University