Using Data Science to Improve Air Safety

Distribution Statement A: Approved for Public Release per AMRDEC PAO

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

Presented by:

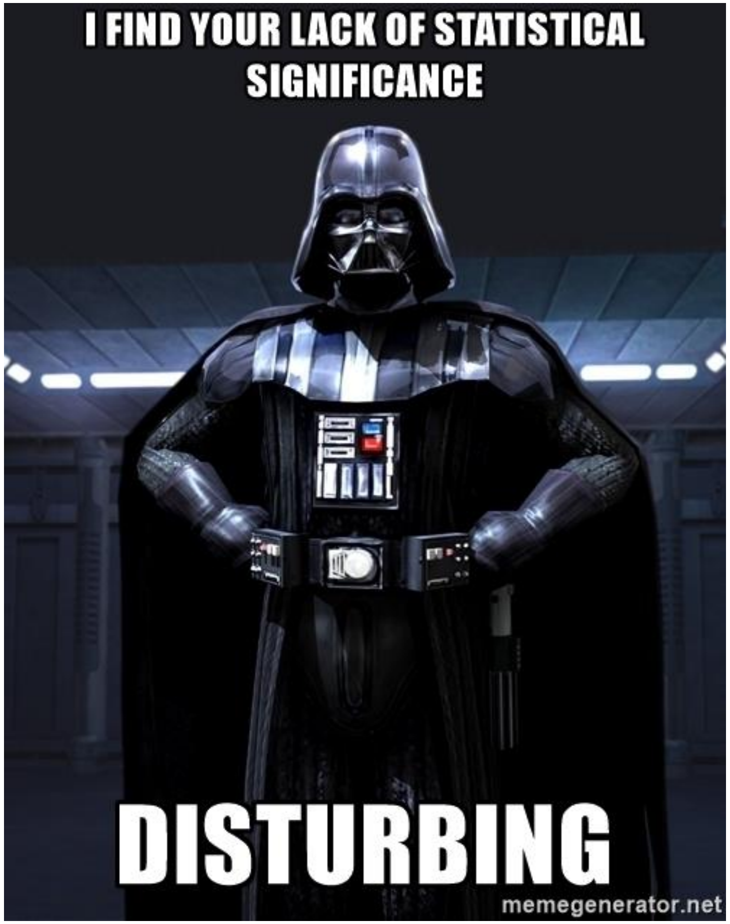
Daniel Wade

Team Lead Aerospace Engineer

U.S. Army Aviation and Missile Research, Development, and Engineering Center

Background

- **US Army Aviation Engineering Directorate**
  - Airworthiness Authority for the Army
  - TRL 7-9 Development and Qualification
- **Dynamics Branch**
  - Health and Usage Monitoring Systems and Aviation Data Science Team Lead
- **Bachelor and Master of Science in Mechanical Engineering**
  - Dynamics & Modal Analysis
  - I'm not a
    - Researcher
    - Statistician or
    - Data scientist



Who is AMRDEC?

U.S. Army Aviation and Missile Research, Development, and Engineering Center provides increased responsiveness to the nation's Warfighters through aviation and missile capabilities and life cycle engineering solutions.

- **Headquartered at Redstone Arsenal, AL**
- **5 Directorates**
- **9,000 scientists & engineers**
- **$2.45 billion in reimbursable funding, FY 16**
- **$339 million in Science & Technology funding, FY 16**

AMRDEC Priorities

Strategic Readiness – provide aviation and weapons technology and systems solutions to ensure victory on the battlefield

Future Force – develop and mature Science and Technology to provide technical capability to our Army's (and nation's) aviation and weapons systems

Soldiers & People – develop the engineering talent to support both Science and Technology and materiel enterprise

The Lexicon of Aviation Data Science

- Health and Usage Monitoring Systems (**HUMS**)
  - The child of FOQA (Flight Operations Quality Assurance)

- **True Positive**: *Sensitivity*; HUMS correctly identified a faulted state
  - False Negative: Missed Detection

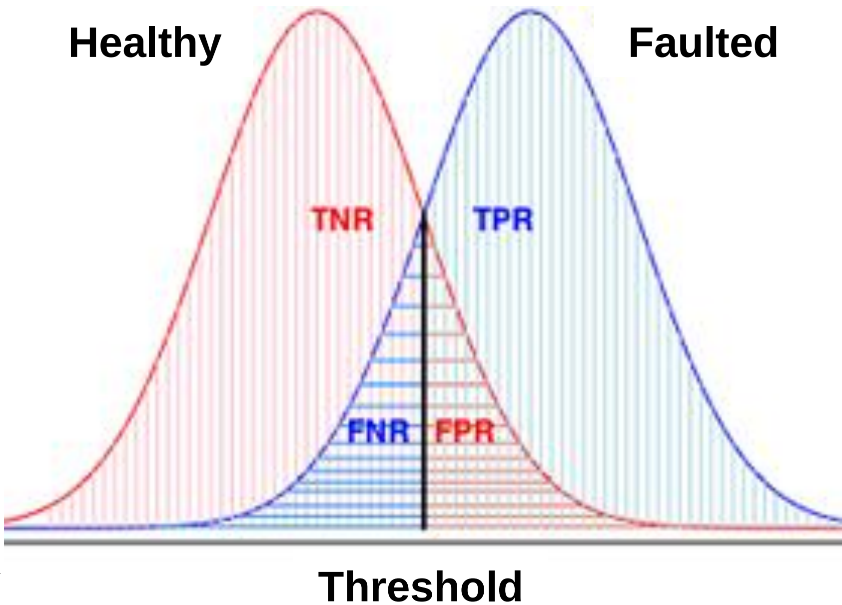
- True Negative: *Specificity*; HUMS correctly identified a healthy state
  - **False Positive**: False Alarm

- **Bookmakers Informedness** = TPR – FPR

- **Ground Truth**
  - Assets and Examples

- **ROC**: Receiver Operating Characteristic

- **Epicyclic Transmission**: Planetary Gearbox



What is HUMS?

Health and Usage Monitoring System

Flight Operations Data (Parametric Data)

e.g. altitude, pitch rate, engine torque

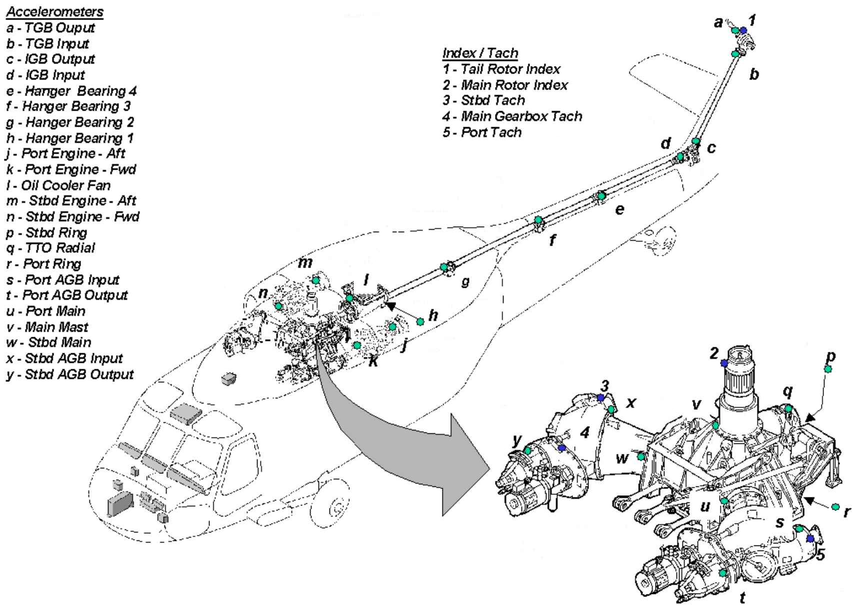
Sensor Data

Burst data (High Frequency)

e.g. accelerometers

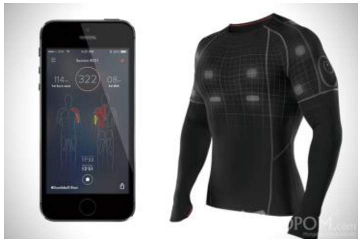
Continuous data (Low Frequency)

e.g. oil debris monitor



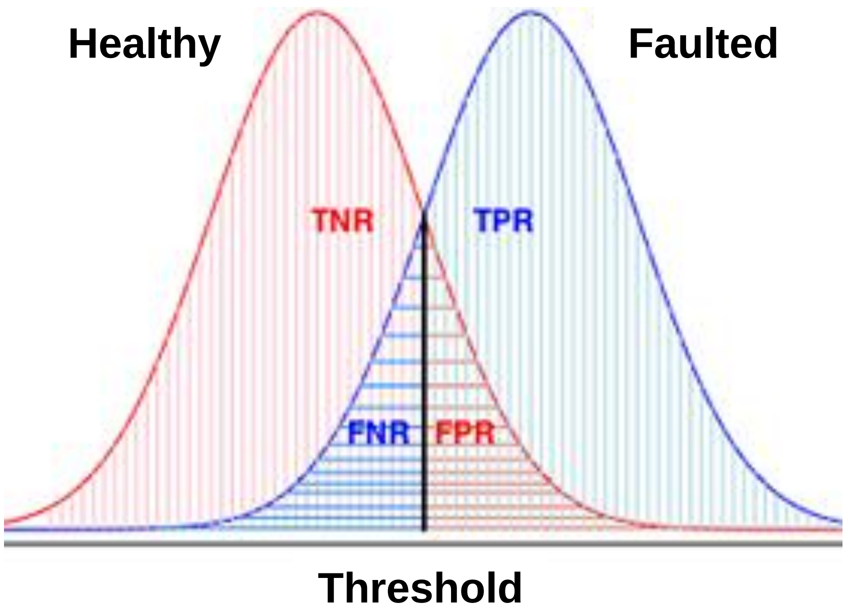
What do we use it for?

- Univariate exceedance monitoring during flight
  - Oil debris monitoring
- Health/Usage monitoring
  - Drive train vibration
  - Rotor vibration
  - Flight regime classification
- Accident Investigation
  - Cockpit voice
  - Flight data recording



Problems with HUMS?

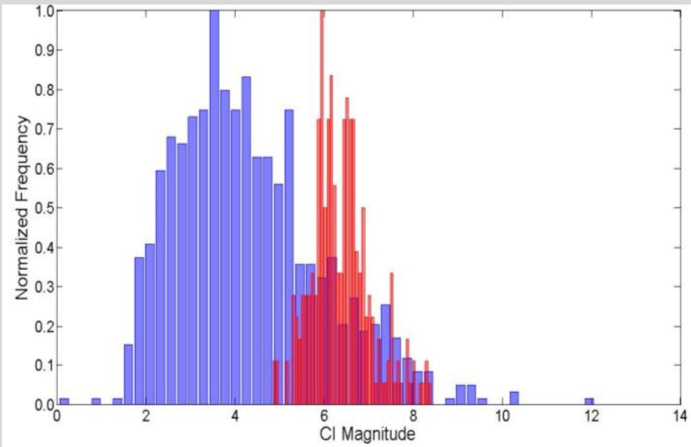
Exclusively uses univariate exceedance classification methods which are often prone to a False Positive/Negative problem.

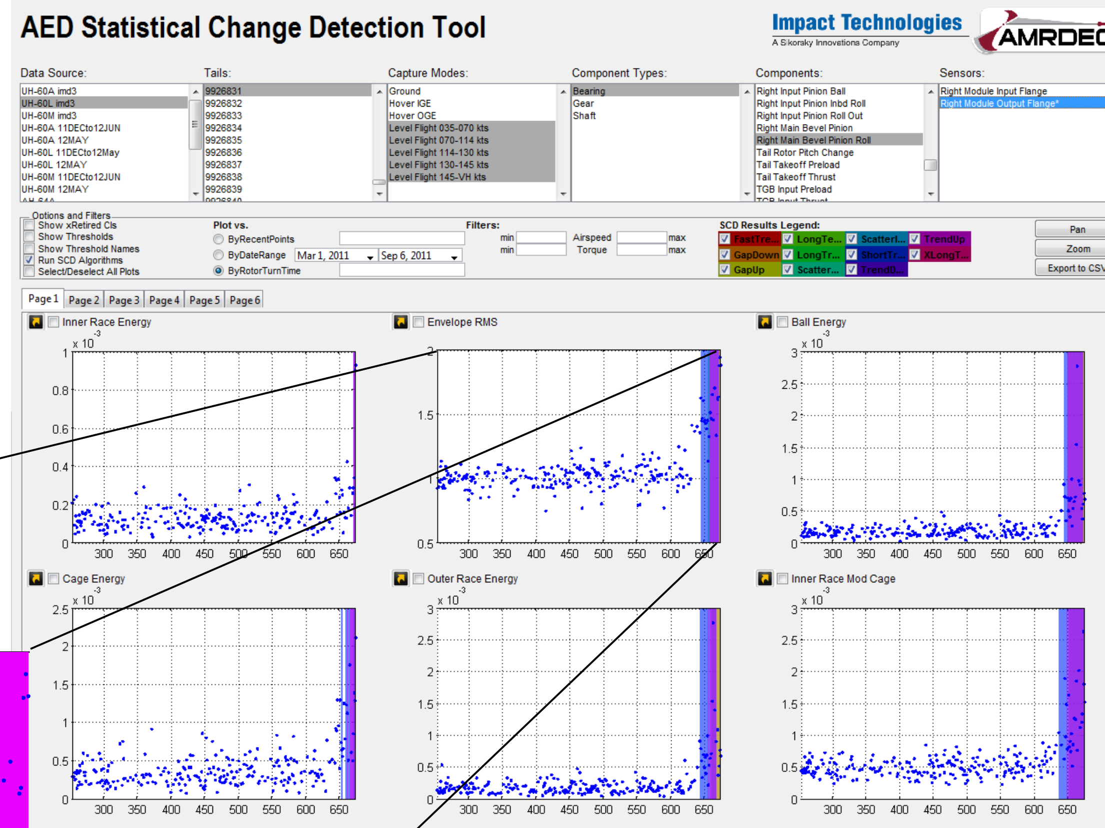


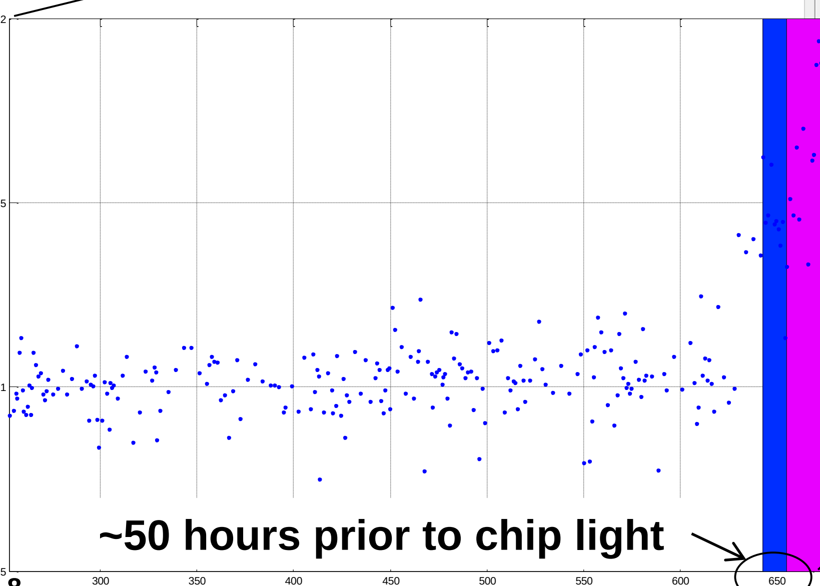
- The problem is temporal
- The variables are noisy
- Health is often relative
- Anomalous does not always mean broken or dangerous
- It does not account for other flight variables

An Example: Change Detection

The aircraft is not separated from the fleet

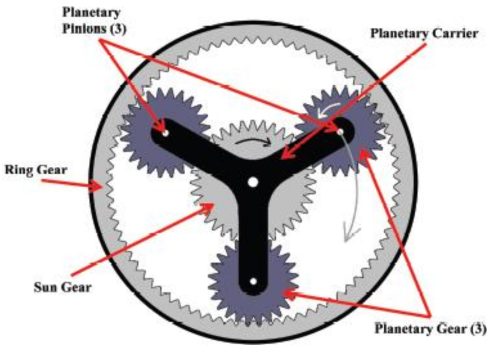


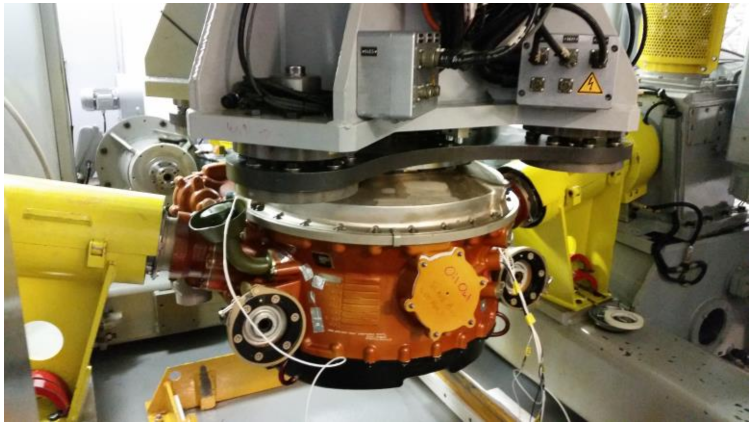




Case Study: Transmission Internal Failure

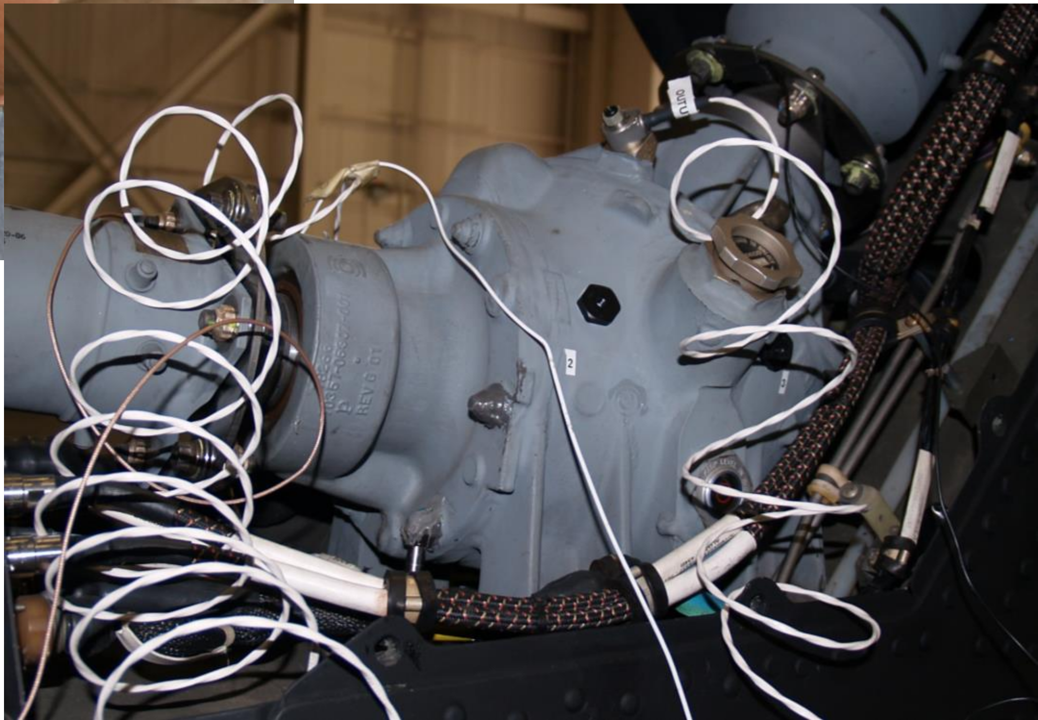
Epicyclic Transmission





Spiral Bevel Transmission

Can vibration transfer across an epicyclic transmission?



How well are we actually doing?

Epicyclic Transmission 1	HUMS Indicated Healthy	HUMS Indicated Fault	Assets
Actual Condition: Healthy	TN=100%	FP=0%	2
Actual Condition: Faulty	FN=100%	TP=0%	6
		Sum of Assets:	8

Epicyclic Transmission 2	HUMS Indicated Healthy	HUMS Indicated Fault	Assets
Actual Condition: Healthy	TN=100%	FP=0%	4
Actual Condition: Faulty	FN=100%	TP=0%	4
		Sum of Assets:	8

Epicyclic Transmission 3 Built In HUMS	HUMS Indicated Healthy	HUMS Indicated Fault	Assets
Actual Condition: Healthy	TN=0%	FP=100%	1
Actual Condition: Faulty	FN=100%	TP=0%	25
		Sum of Assets:	26

Epicyclic Transmission 4	HUMS Indicated Healthy	HUMS Indicated Fault	Assets
Actual Condition: Healthy	TN=91%	FP=9%	11
Actual Condition: Fault	FN=95%	TP=5%	21
		Sum of Assets:	32

Can we improve?

Epicyclic Transmission 3 Built In HUMS	HUMS Indicated Healthy	HUMS Indicated Fault	Assets
Actual Condition: Healthy	TN=0%	FP=100%	1
Actual Condition: Faulty	FN=100%	TP=0%	25
		Sum of Assets:	26

Epicyclic Transmission 3 Modified HUMS	HUMS Indicated Healthy	HUMS Indicated Fault	Assets
Actual Condition: Healthy	TN=100%	FP=0%	1
Actual Condition: Faulty	FN=56%	TP=44%	25
		Sum of Assets:	26

What about spiral bevel transmissions?

Tail Gearbox 1	HUMS Indicated Healthy	HUMS Indicated Fault	Assets
Actual Condition: Healthy	TN=100%	FP=0%	4
Actual Condition: Faulty	FN=0%	TP=100%	3
		Sum of Assets:	7

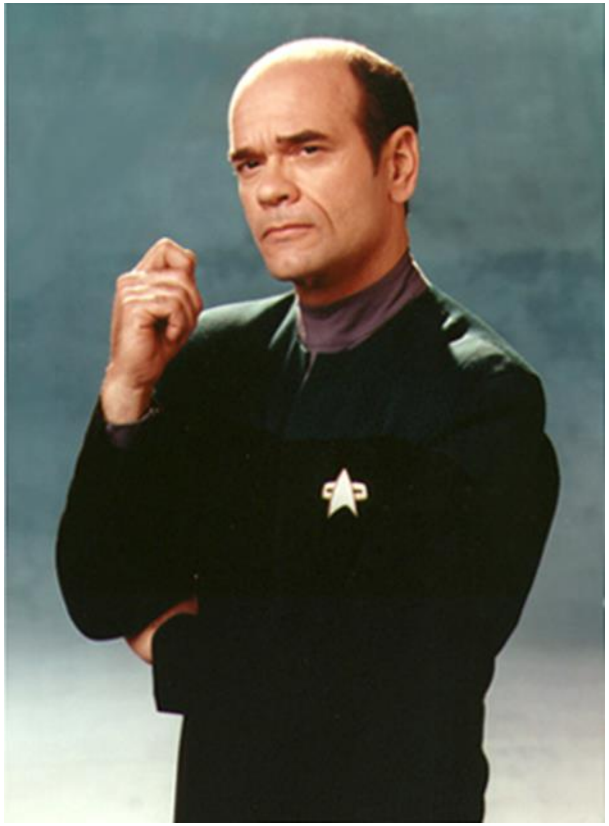
Tail Gearbox 2	HUMS Indicated Healthy	HUMS Indicated Fault	Assets
Actual Condition: Healthy	TN=71%	FP=29%	7
Actual Condition: Fault	FN=13%	TP=87%	15
		Sum of Assets:	22

What are we doing to fix the problem?

Remember the Emergency Medical Hologram?

What are we doing to fix the problem?

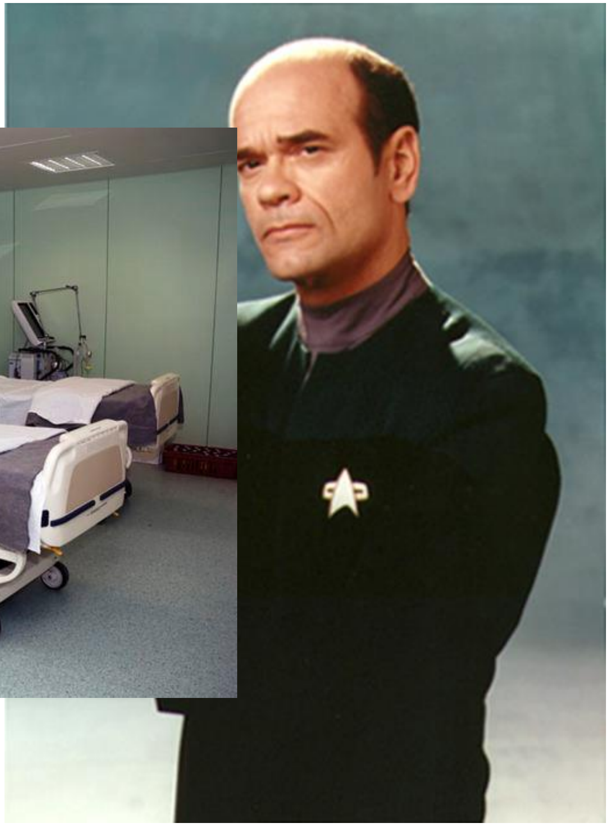
Remember the Emergency Medical Hologram?



Please state the nature of the medical emergency

What are we doing to fix the problem?

Remember the Emergency Medical Hologram?



*Please state the nature of the **engineering** emergency*

Machine Learning in a critical environment

We live in a common place with other industries when we talk about this topic:

- Medicine
- Nuclear Power
- Aviation

Development of multivariate machine learned diagnostics and prognostics requires

a process…

Our Machine Learning Process

Our Machine Learning Process

Our Machine Learning Axioms for Aviation

- Stirring the pile, is training
- Model evaluation, is training
- Model selection, is training
- Model validation, is training
- Looking under the hood, is training
- Stirring stops prior to testing
- Testing is done by the customer on a clean dataset

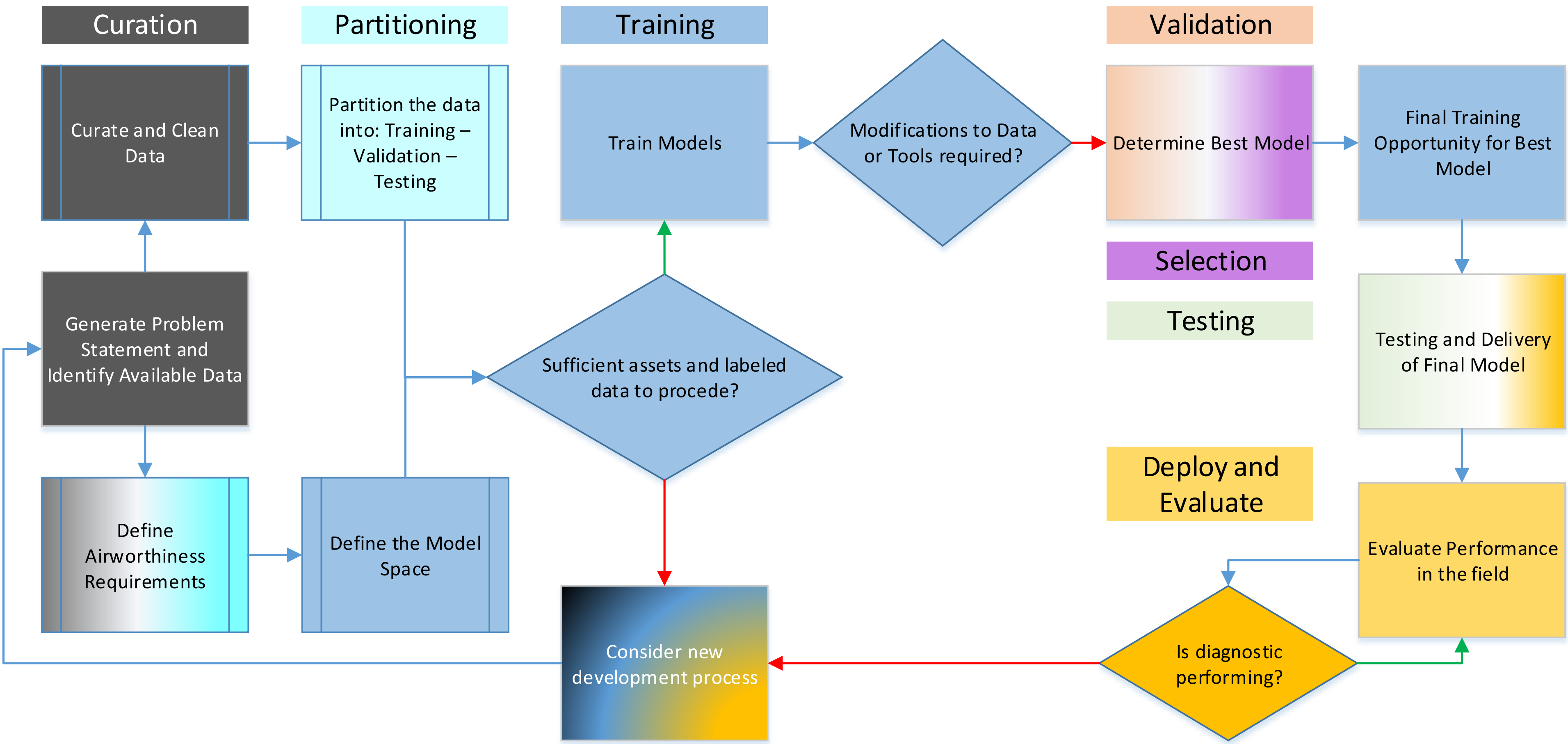
How did we implement our axioms on a real aviation problem?

- We put together a general path forward we expect to see when we take on a machine learning task.

- Demonstrated in our NGB internal failure classification work
  - Cleanse
  - Partition
  - Train
  - Validate
  - Select
  - Test
  - Deploy

- We built a flow chart!

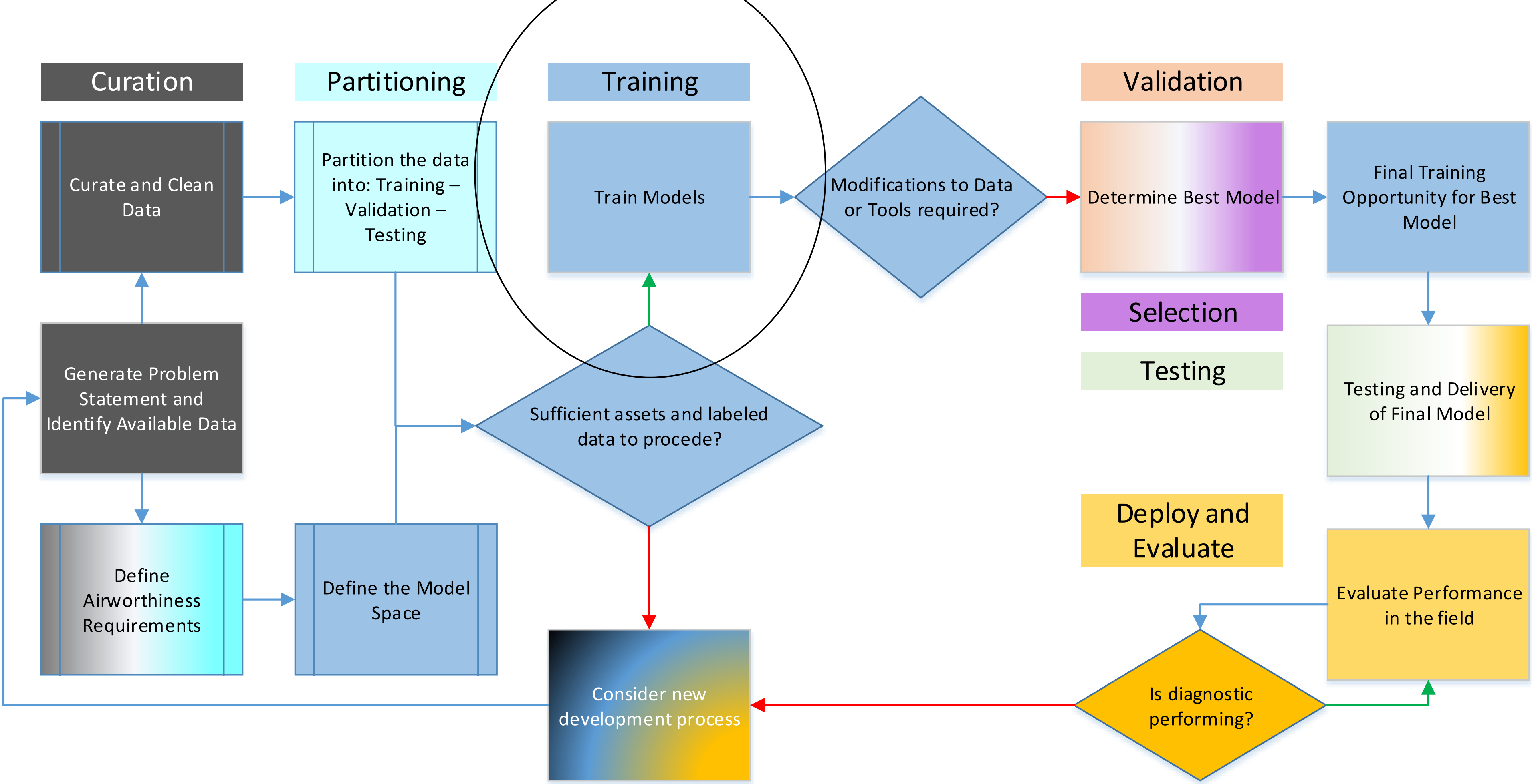
Aviation Machine Learning Process



TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

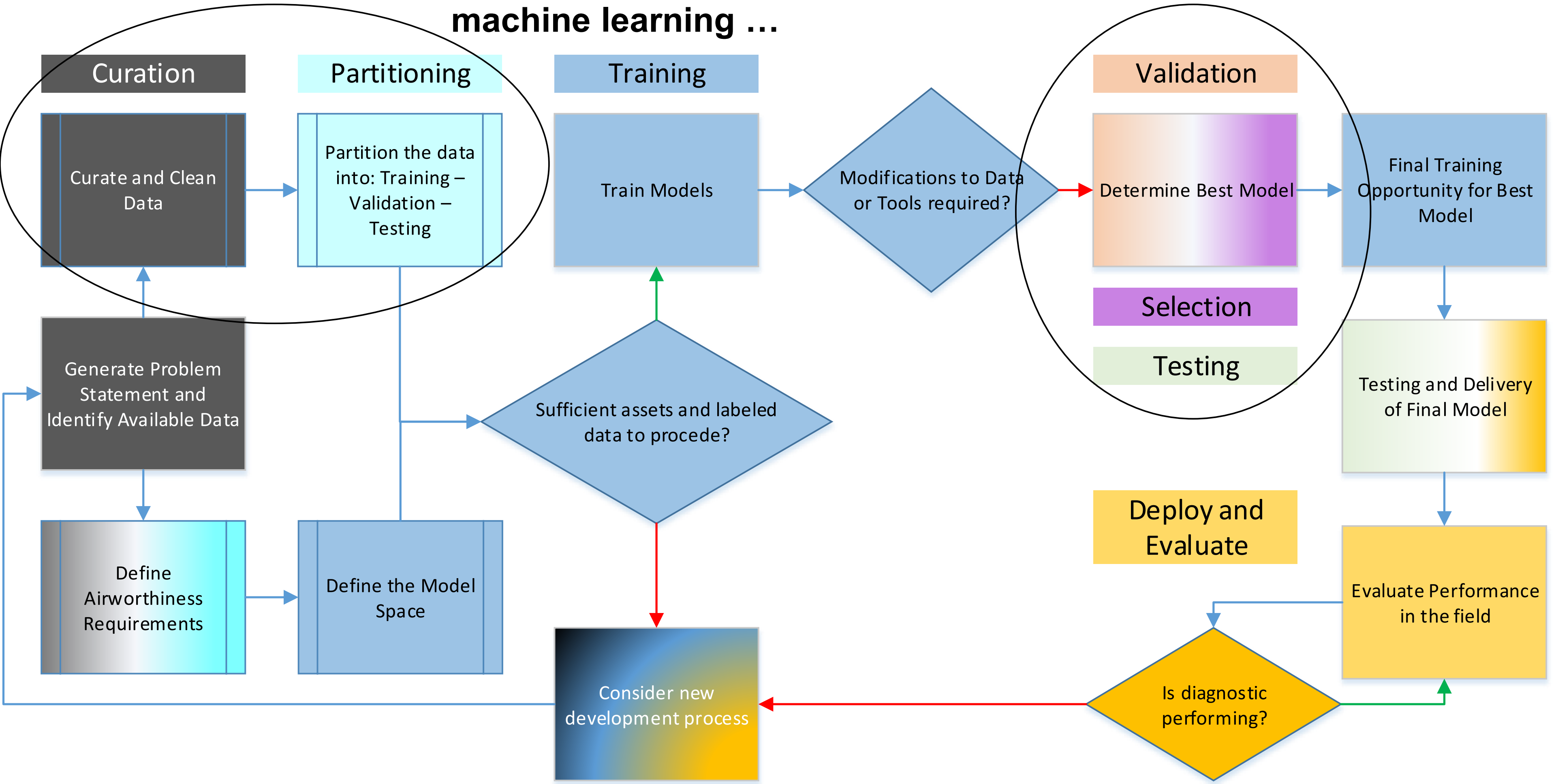
Aviation Machine Learning Process

What people think when I say machine learning

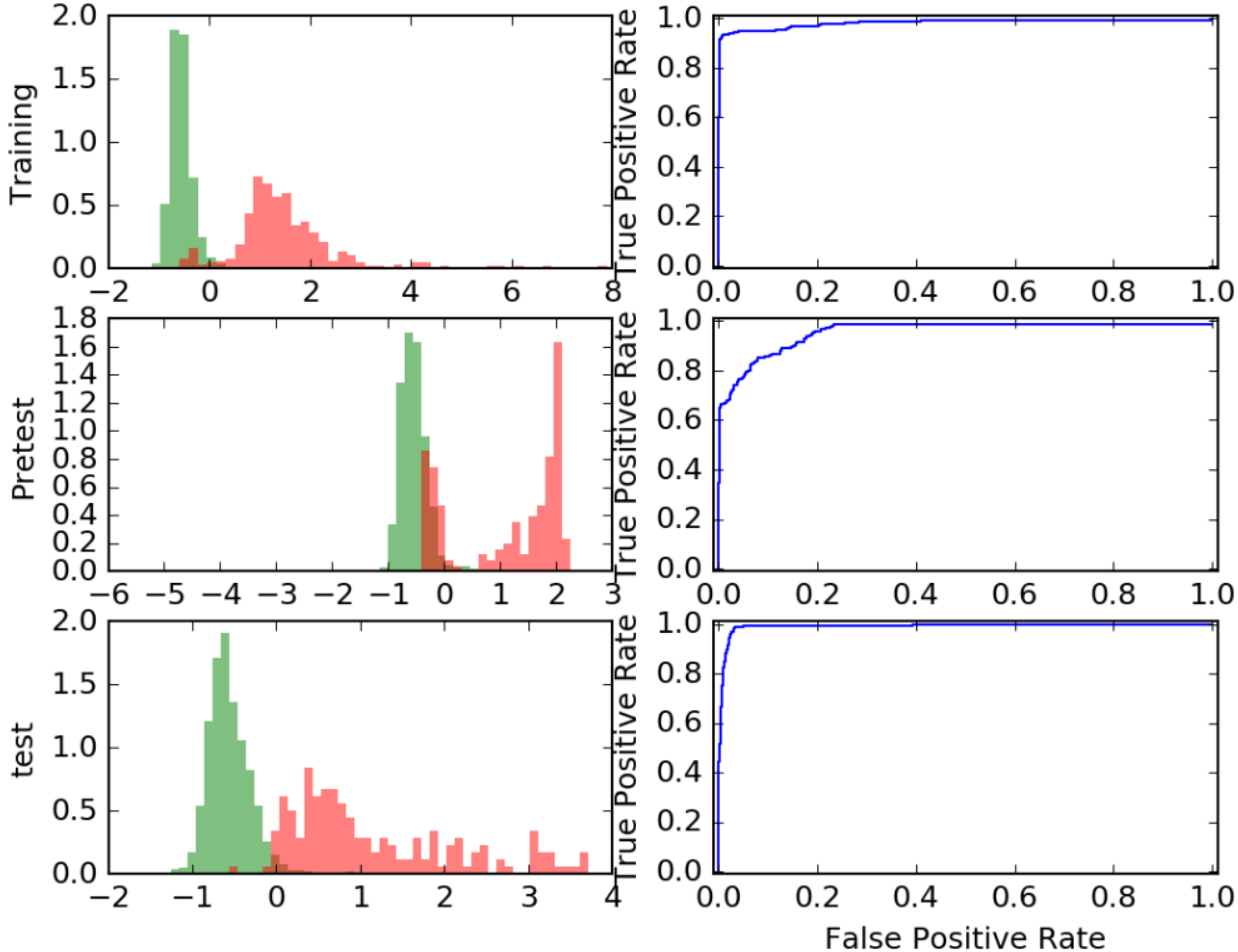


Aviation Machine Learning Process

What I've realized is the important part of machine learning …



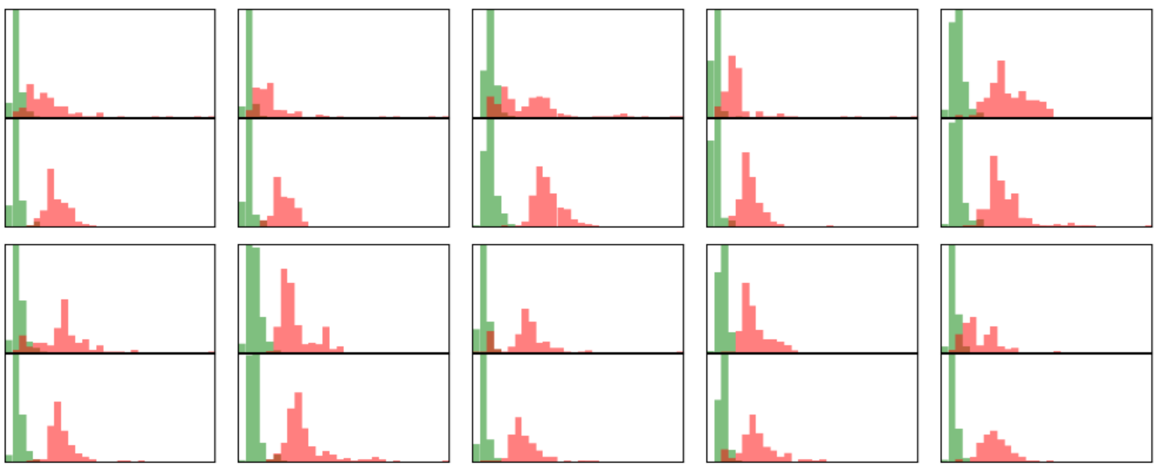
ROC curves

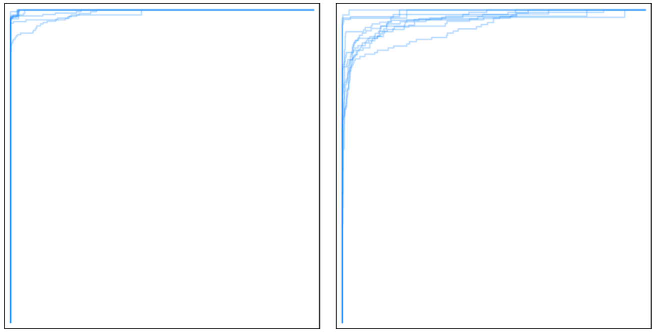

Training
Pretest
test
True Positive Rate
False Positive Rate

How did it perform in cross validation?

CROSS-VALIDATION w/ 10 FOLDS

Fold #	True Positive Rate	False Positive Rate	True Positive Accuracy	True Negative Accuracy	Informedness
Fold #0	1.00	0.08	0.22	1.00	0.92
Fold #1	0.84	0.06	0.26	1.00	0.78
Fold #2	0.90	0.02	0.50	1.00	0.87
Fold #3	1.00	0.06	0.34	1.00	0.94
Fold #4	1.00	0.07	0.23	1.00	0.93
Fold #5	0.98	0.07	0.24	1.00	0.92
Fold #6	1.00	0.06	0.29	1.00	0.94
Fold #7	0.83	0.02	0.55	1.00	0.82
Fold #8	0.99	0.03	0.49	1.00	0.96
Fold #9	0.82	0.02	0.52	0.99	0.79





Dr. Andrew Wilson

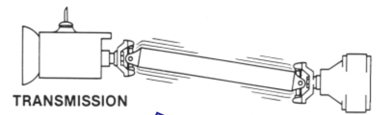
SPECTRUM SURROGATE MODELING FOR VIBRATION PROBLEMS

Overview

- Army Rotorcraft Vibration Problems
  - Mechanical Diagnostics
  - Aeroelastic Dynamics
- Three applications of surrogate modeling
  - Sandia/Army Collaboration
    - MD: Sensor Redundancy
    - Aeroelastic: Surrogate DNS
  - Army Sustainment Innovation
    - MD: Spectrum Reconstruction

Background

IMAGENET



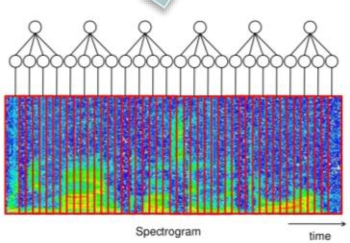
Aviation Engineering

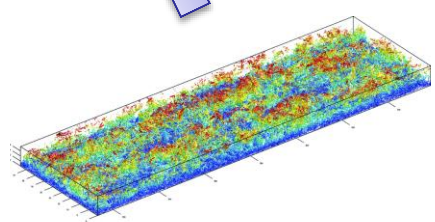
theano

Sandia/AED Collaboration: Applying CNNs to frequency-response modeling

Microsoft CNTK

KDD





Sandia National Laboratories

Test Problem Background

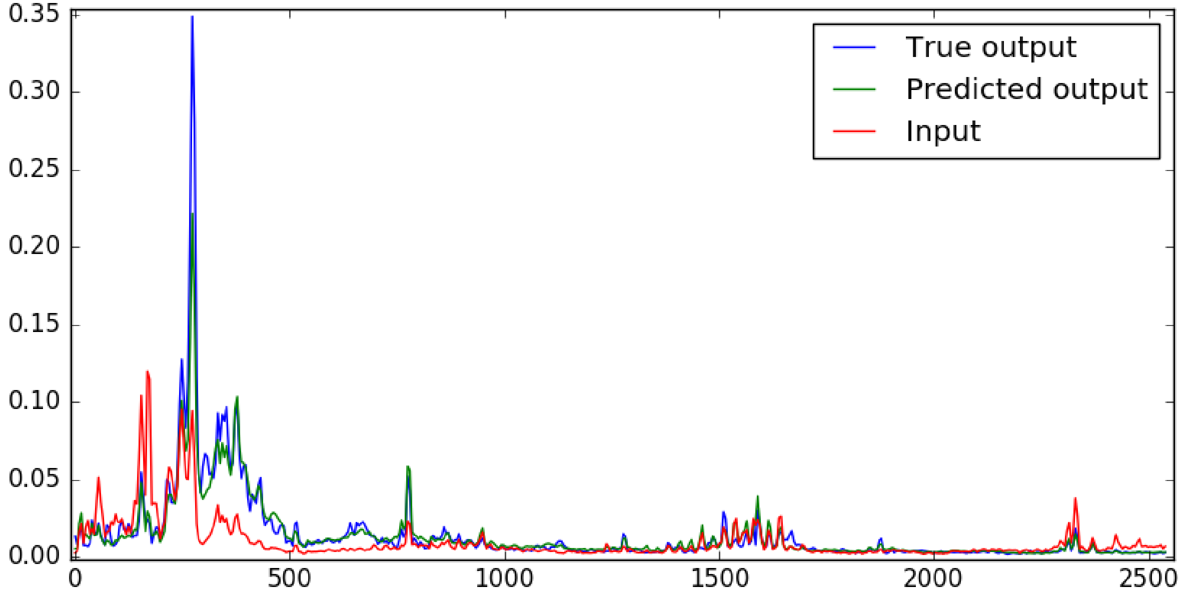
Sandia National Laboratories

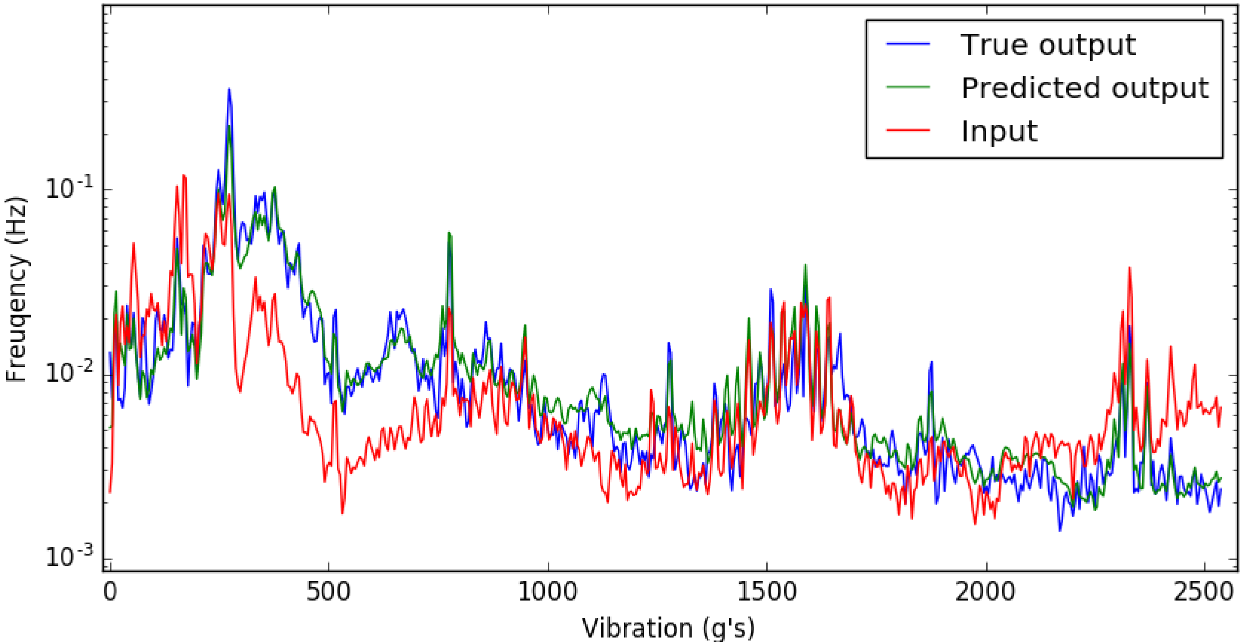
Aviation ENGINEERING

- Turbulent Boundary Layer Wall PSD
  - LES is relatively cheap but misses near-wall dynamics
  - DNS is very expensive but high fidelity
  - Can CNNs use free-stream PSDs to predict wall PSDs?
- Axial/Vertical Sensor Redundancy
  - Two accelerometers fielded to all aircraft in perpendicular axes
  - Sensors + wiring costly (lbs on aircraft)
  - Years of collected operational spectra
  - Can CNNs use axial axis spectrum to predict vertical axis spectrum?

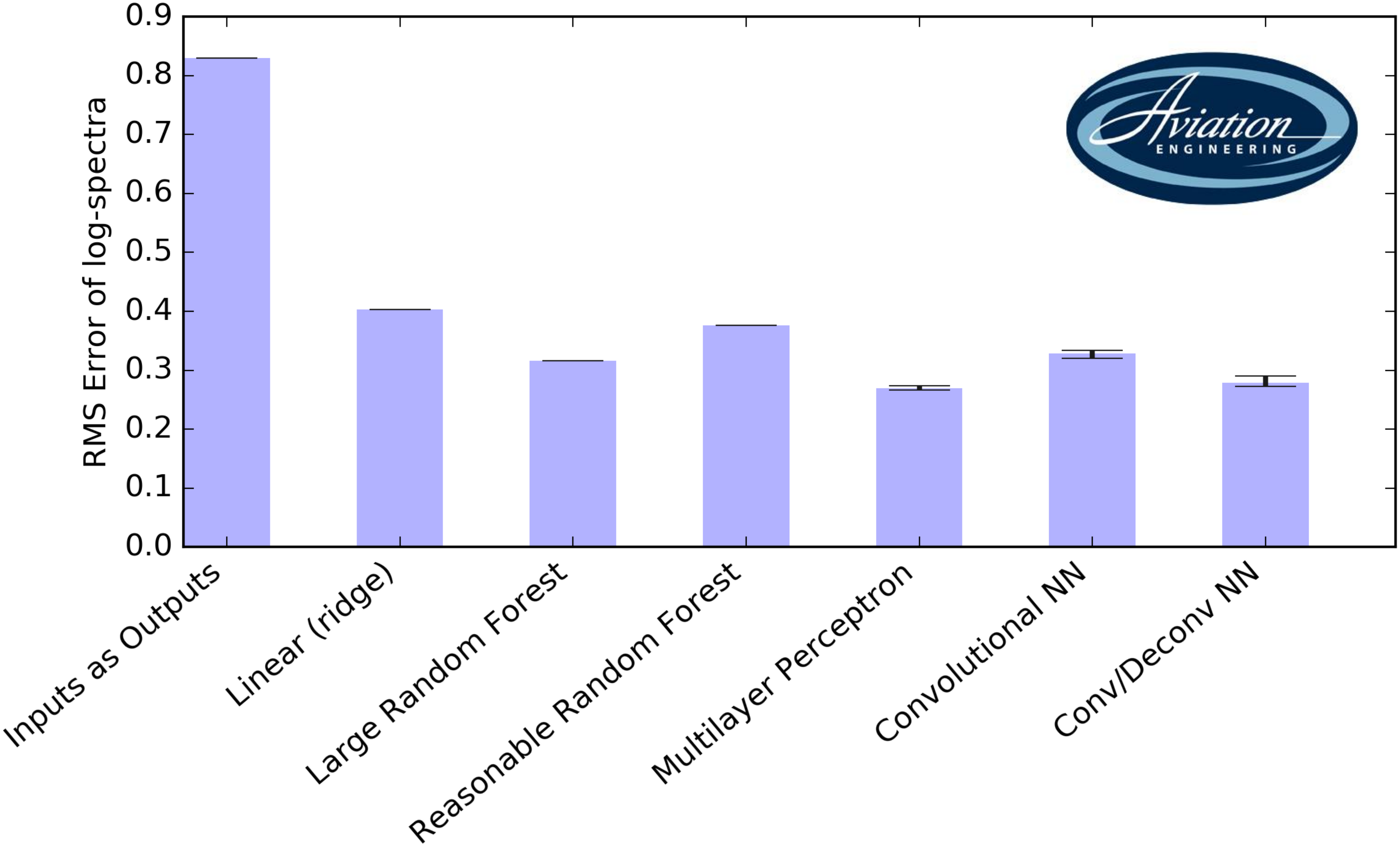
AED Results: Surrogate Model for Sensor

AED Perpendicular Axis: Results

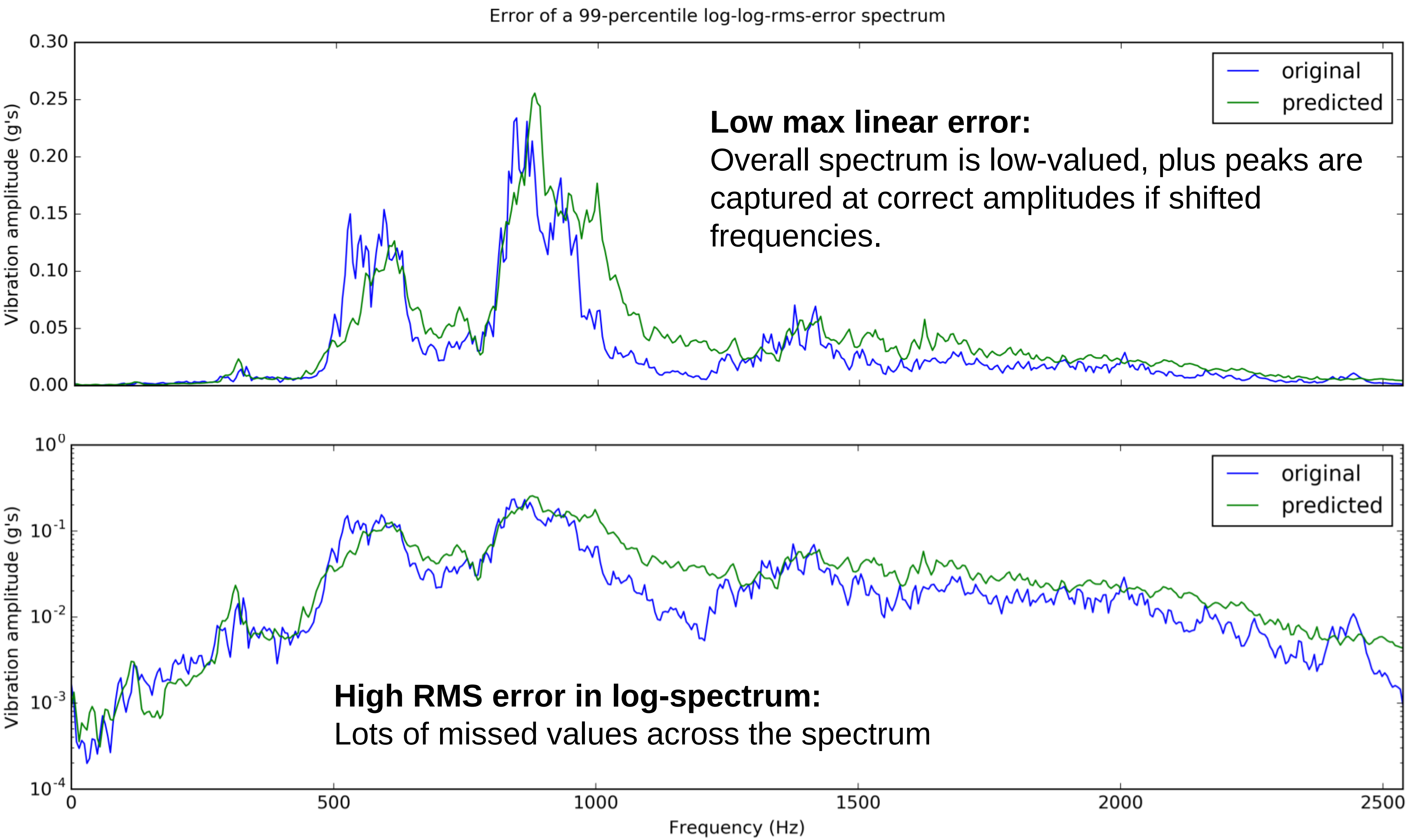




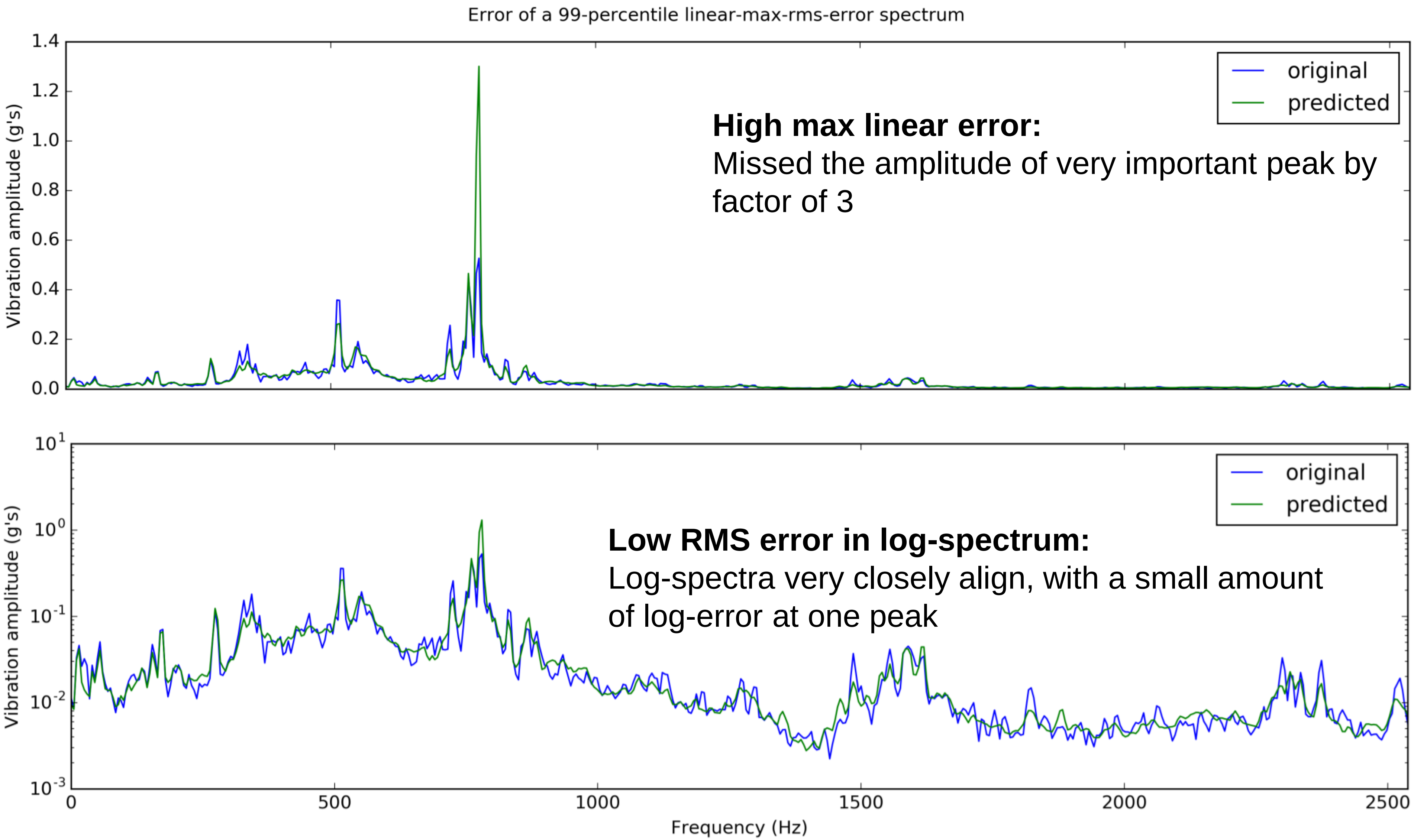
AED Perpendicular Axis: Results



RMS vs Max
Linear vs Log



RMS vs Max
Linear vs Log



High max linear error:
Missed the amplitude of very important peak by factor of 3

Low RMS error in log-spectrum:
Log-spectra very closely align, with a small amount of log-error at one peak

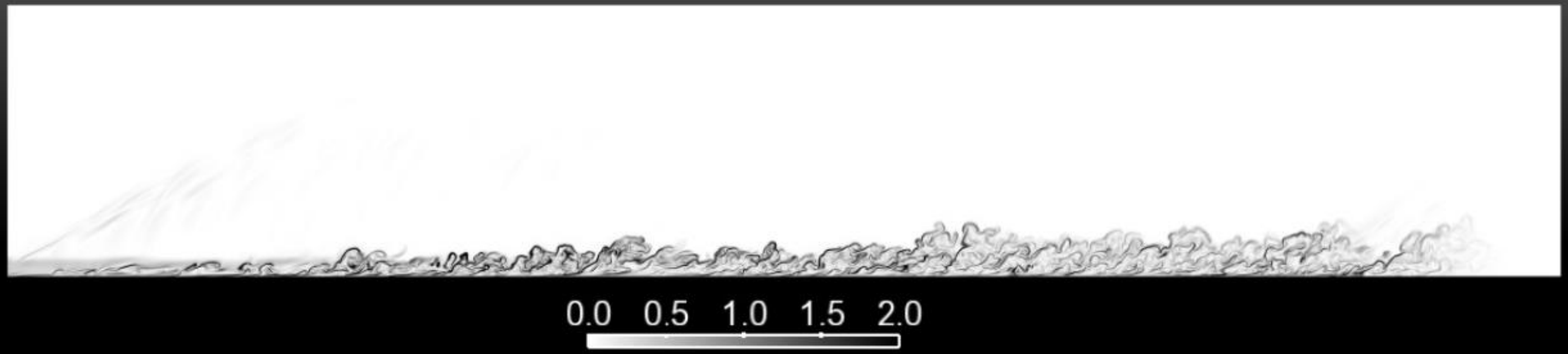
Sandia Results: Surrogate Model for DNS

Sandia National Laboratories

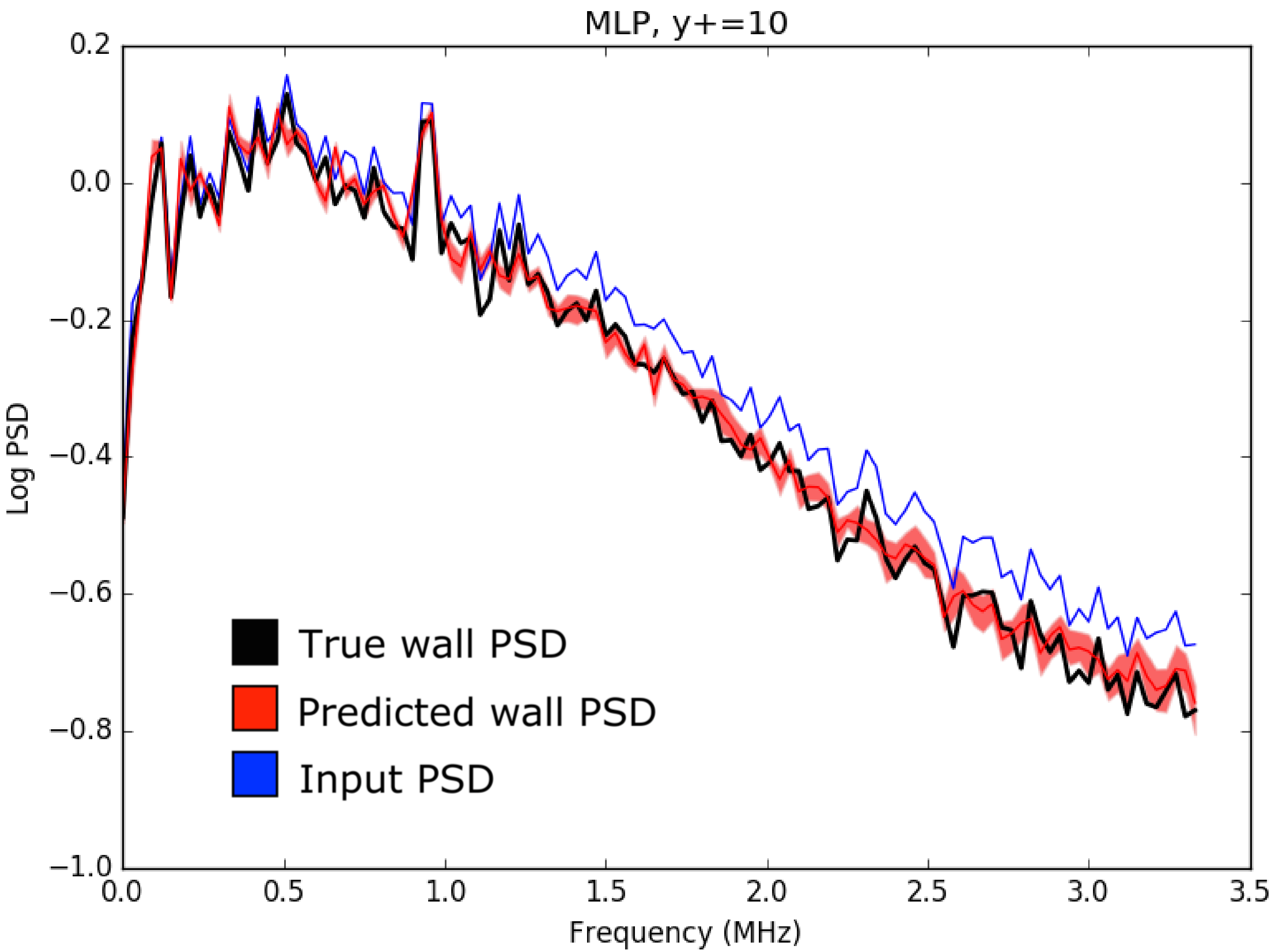
Sandia DNS Data Set

1. Mach 2.0 compressible flat plate turbulent boundary layer
2. Low-dissipation 5th order upwind biased flux-reconstruction scheme
3. Fourth order explicit Runge Kutta time integration
4. 100.7 M mesh cells
5. Near wall resolution: $\Delta x+ < 5$, $\Delta y+ < 0.2$, $\Delta z+ < 4$
6. $1075 < Re\Theta < 1310$
7. Run for > 1200τ (where $\tau = \delta 0/ U\infty$)

Sandia National Laboratories



Sandia Wall PSD Results



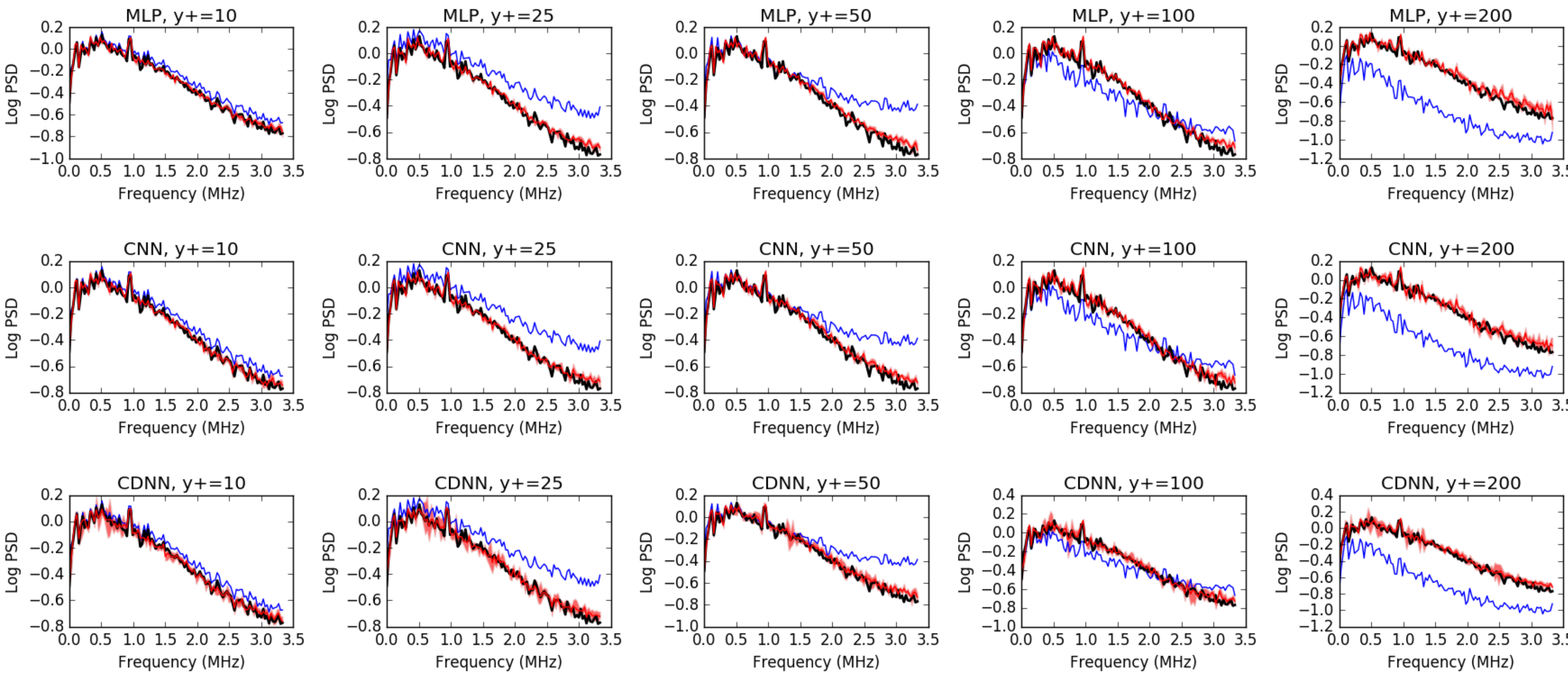
Sandia Wall PSD Results

True wall PSD

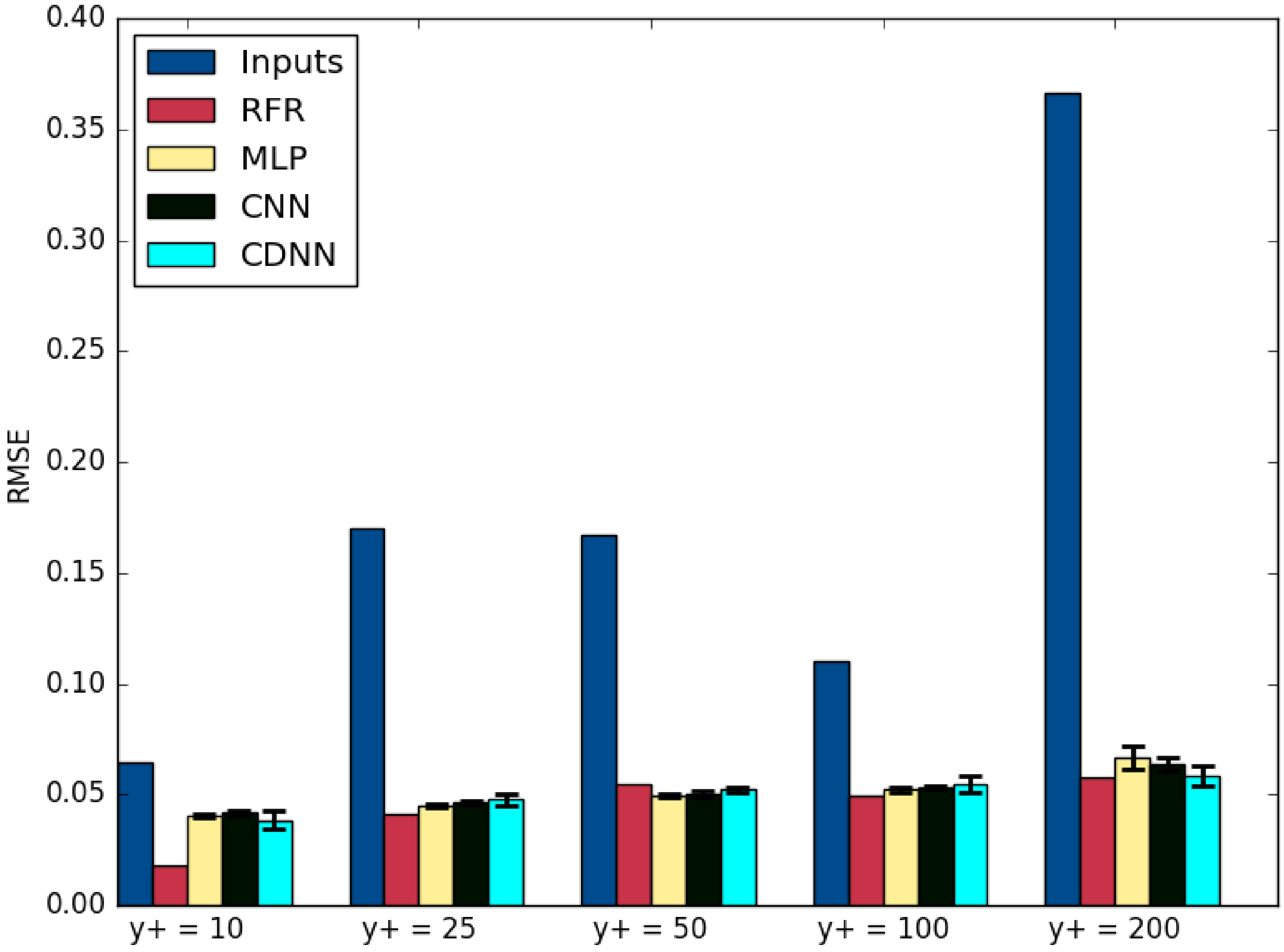
Predicted wall PSD

Input PSD

Sandia National Laboratories



Sandia Wall PSD Results



Sandia National Laboratories

Conclusions and Future Work

Sandia National Laboratories

Aviation Engineering

- Performance of all methods similar
- Can predict PSD at wall, even out to y+ = 200
- High frequency predictions require further work
- Data partitioning methodology

- Definite difference in performance
- Depth of NN important?
- Max-errors unacceptable in linear domain
- RMS errors very good

- Pursue max-error loss functions for NN training
- Need to further explore validation/evaluation criteria
- Powerful and promising methods

AED Project: Surrogate Model For Spectrum Reconstruction

TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.

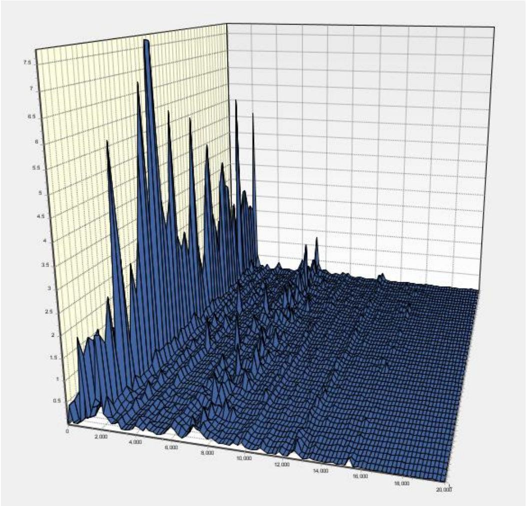
The "Raw Data" Problem

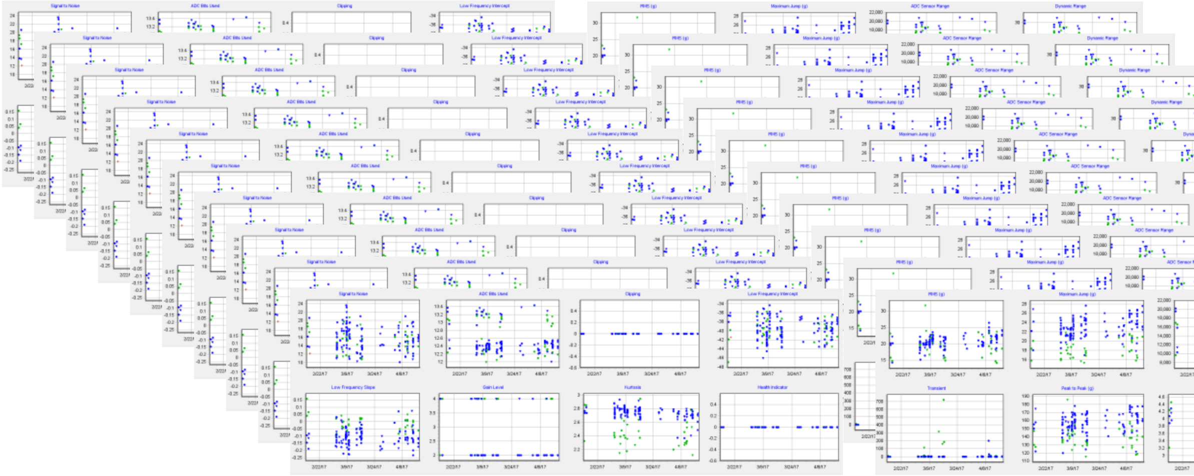
"Raw" vibration data

- Large data storage (?)
- Sensor data
- 30% of capture events
- Widely understood
- 8,193 point spectra

Condition indicators (CIs)

- Reduced storage (?)
- Features
- 100% of capture events
- Highly specialized
- 1,500 CIs





Problem/Opportunity

- IVHMS computes 1000's of CIs
  - Main Mod: >4,000 CIs generated from 5 sensors.
  - One Main Mod sensor:
    - >1500 CIs total
    - >580 CIs available for essentially all cases

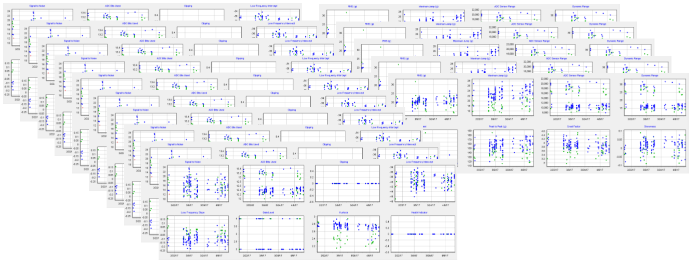
- **Question: Is there enough information in these CIs to reconstruct a reasonable approximation to the original spectrum?**

Machine Learning

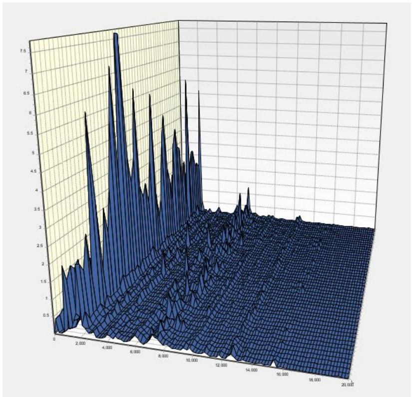
- One “semester” (half-year) of data
  - 90,000 acquisitions with raw data (spectra)
  - 580+ CIs (including sensor health CIs)
  - ~20 additional useful variables (Torque, Nr, …)

- Regression model:
  - Inputs are the CIs
  - Outputs are the spectrum bin values

- ***Ideal machine learning problem***
  - Surrogate model for *math*, not *physics*: CI computation
  - Non-independence of frequency features

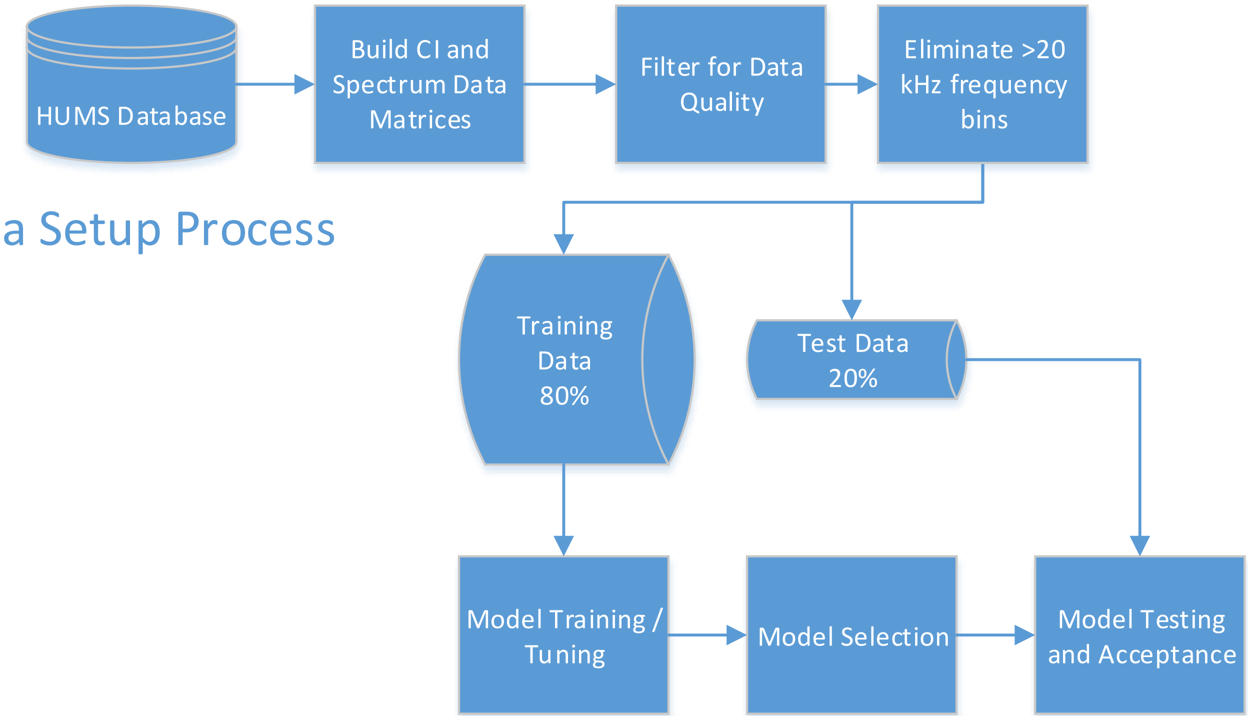


Machine learning,
a.k.a.,
“black magic”

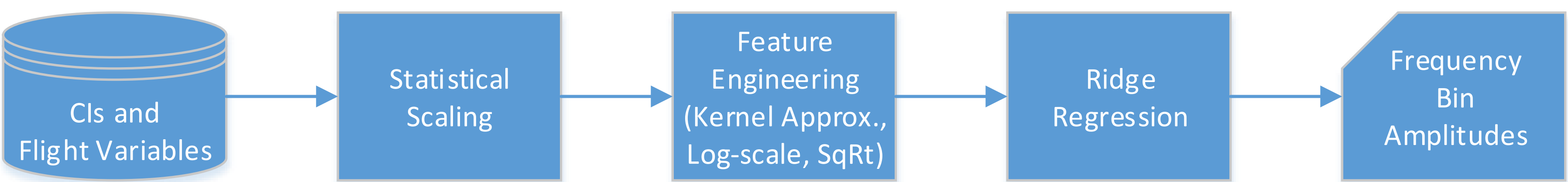


Model Architecture and Training

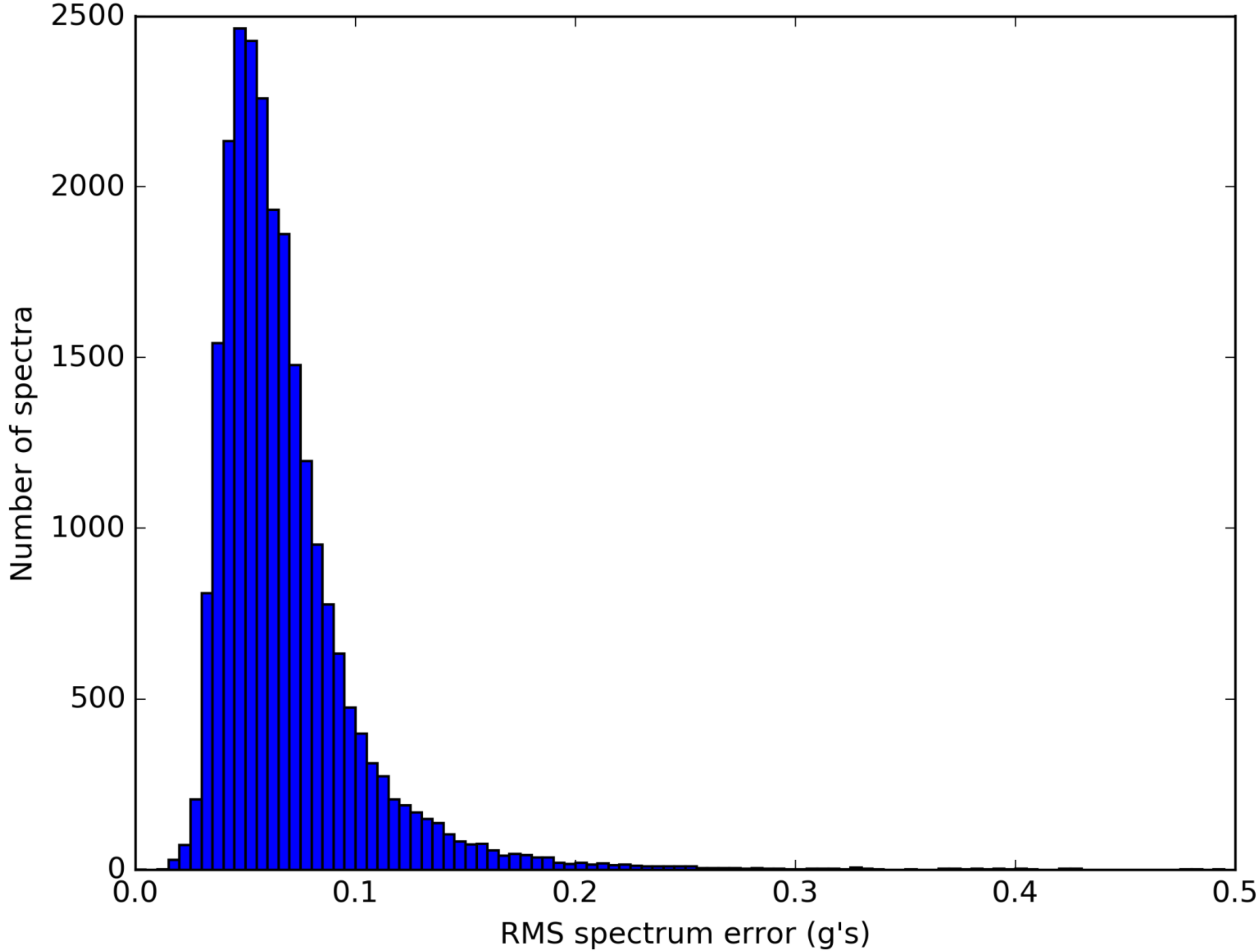
Data Setup Process



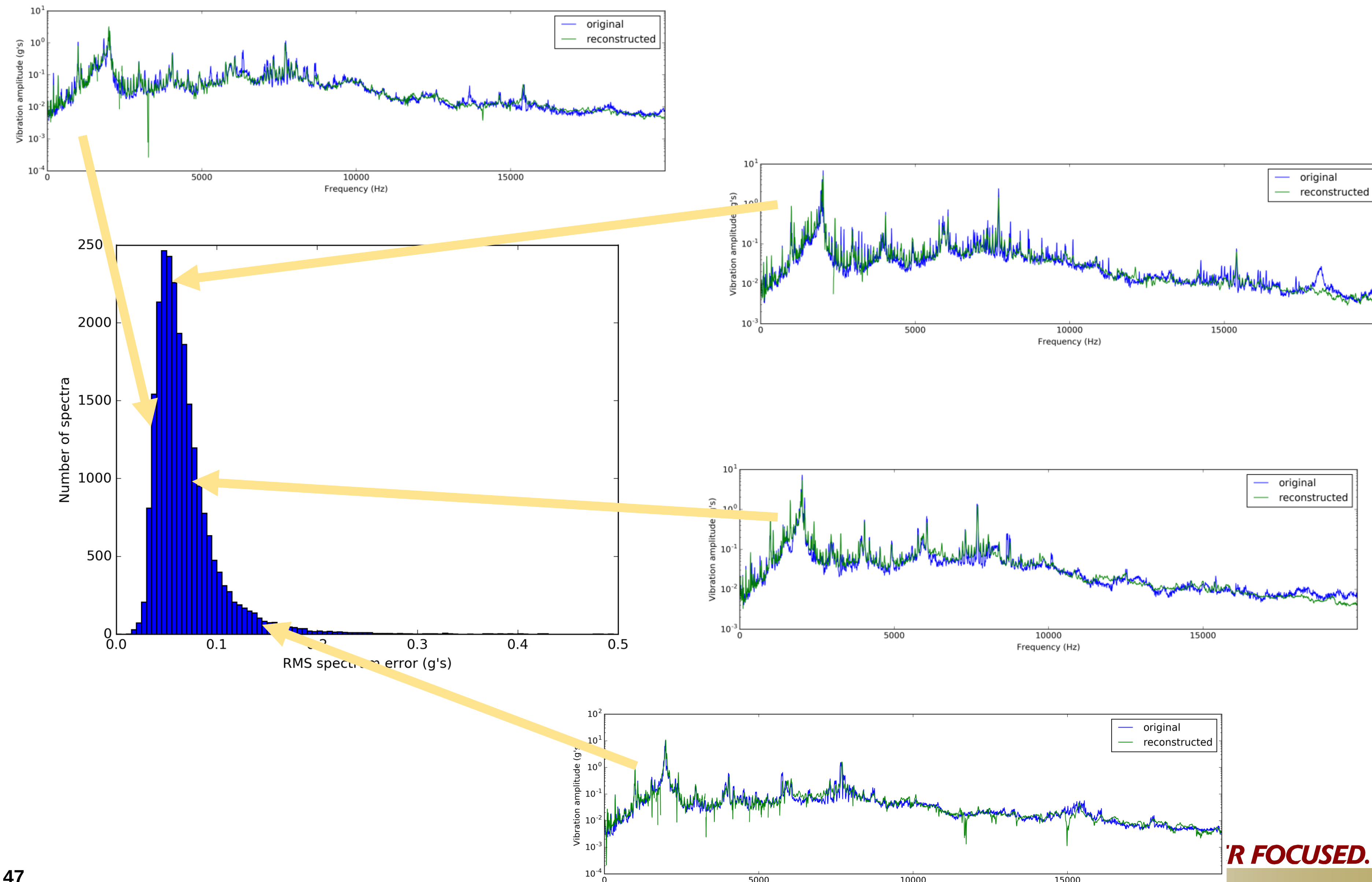
Model Architecture



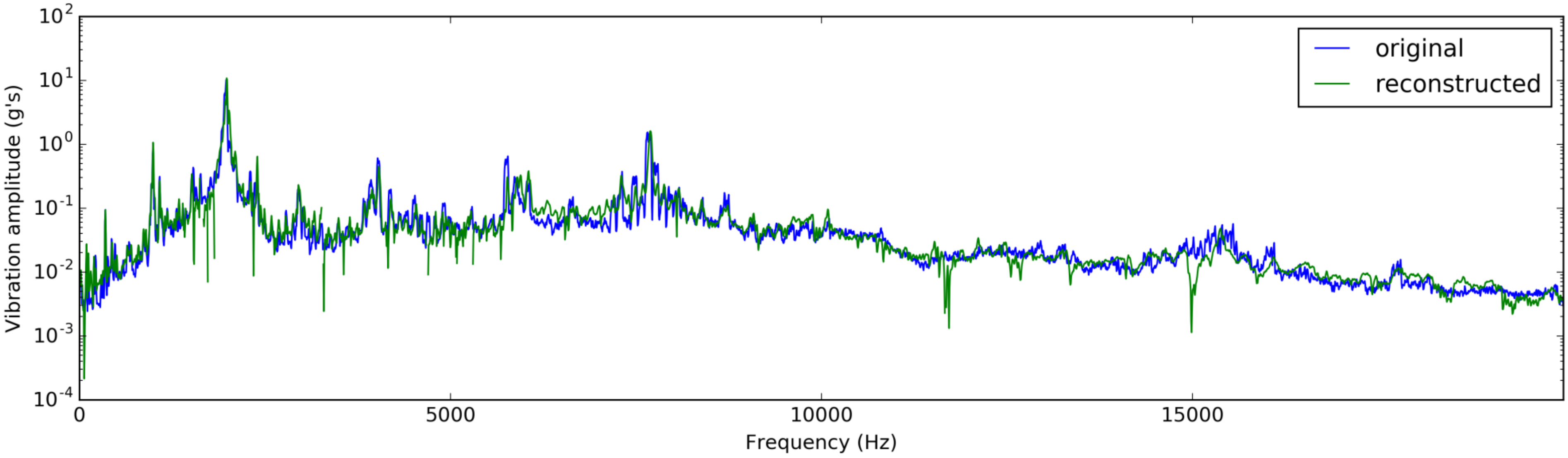
Results

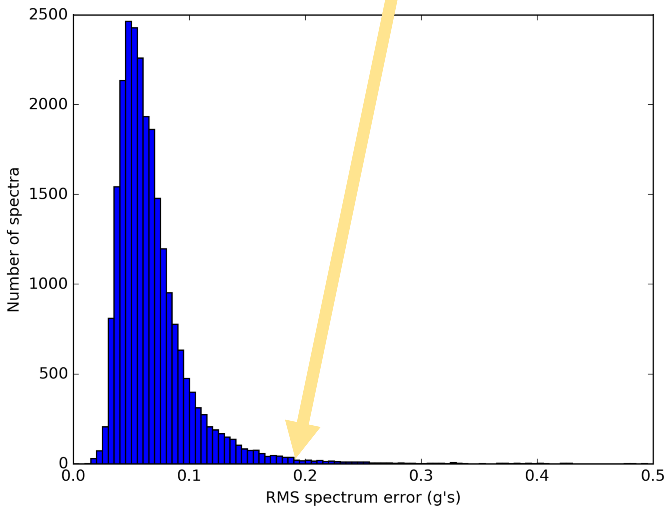

2500
2000
1500
1000
500
0
Number of spectra
0.0
0.1
0.2
0.3
0.4
0.5
RMS spectrum error (g's)

Results


original
reconstructed
Vibration amplitude (g's)
Frequency (Hz)
0
5000
10000
15000
Number of spectra
RMS spectrum error (g's)
0.0
0.1
0.3
0.4
0.5
500
1000
1500
2000
250

Results





99th percentile error

Implications

- Perceived value of data history / tear-down analyses
  - Value of a tear-down driven by data availability; three cases:
    - No data
    - No raw data
    - Full raw data
  - Many tear downs were not performed or (if performed for other reasons) were judged to be uninformative due to the lack of raw data
  - Present work brings value of *no raw data* to almost the same value as *raw data*

- Perceived value of “data reduction” (fewer variables stored)
  - Data reduction is not needed (full raw vibe data < 10% of total data)
  - Data reduction not all that pronounced
    - Too many variables for engineers/maintainers to consider
    - Not significant reduction of data size vs full spectrum

Conclusions

- Data-driven surrogate modeling can be highly effective for problems that are driven by vibration spectra.

- Both simple linear models and incredibly complex deep neural network models can be used very effectively.

Our Relevant Publications on this Topic

1. Wilson, A., Wade, D., Albarado, K., Partain, J., and Statham, M., “A Classifier Development Process for Mechanical Health Diagnostics on US Army Rotorcraft”, Proceedings of the ML and PHM Workshop, SIGKDD 2016, San Francisco, CA, August 2016.
2. Wilson, A., and Wade, D., “Reconstructing Spectra from IVHMS Condition Indicators,” Proceedings of the 73rd American Helicopter Society Annual Forum, Fort Worth, TX, May 2017.
3. Wilson, A., Wade, D., Ling, J., Chowdhary, K., Davis, W., Barone, M., and Fike, J., “Convolutional Neural Networks for Frequency Response Predictions,” Proceedings of the Verification and Validation Symposium, Las Vegas, NV, May 2017.
4. Wade, D., and Wilson, A., “Applying Machine Learning-Based Diagnostic Functions to Rotorcraft Safety”, Proceedings of the Tenth Australian Defence Science and Technology Group International Conference on Health and Usage Monitoring Systems, Melbourne, VIC, Australia, February 2017.
5. Wade, D. et al, “Measurement of Vibration Transfer Functions to Inform Machine Learning Based HUMS Diagnostics,” Proceedings of the 72nd Annual Forum of the American Helicopter Society, May 2016.

Good References

- Cal Tech: “Learning From Data”
  - **FREE** on YouTube
  - https://work.caltech.edu/telecourse
- NASA work in Flight Operations Data and the Future ATC System
  - https://www.nasa.gov/content/air-traffic-operations-lab-answering-big-questions-about-the-future-of-air-travel
- Journal of Aerospace Information Systems
  - https://arc.aiaa.org/loi/jais
- SIGKDD (Association for Computing Machinery: Special Interest Group on Knowledge Discovery and Data Mining)
  - http://www.kdd.org/
- ASME V&V Symposium
  - https://www.asme.org/events/vandv

Thank you for your time and attention

AMRDEC Web Site
www.amrdec.army.mil

Facebook
www.facebook.com/rdecom.amrdec

YouTube
www.youtube.com/user/AMRDEC

Twitter
@usarmyamrdec

Public Affairs
AMRDEC-PAO@amrdec.army.mil