Advanced defect detection and classification in sandwich composite structures using statistical and AI-driven techniques

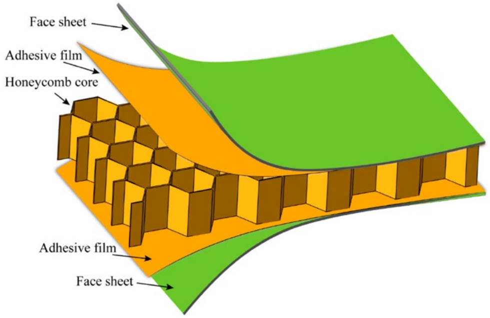
Dr. Joshua Aigbotsua

Prof. Bruce Drinkwater

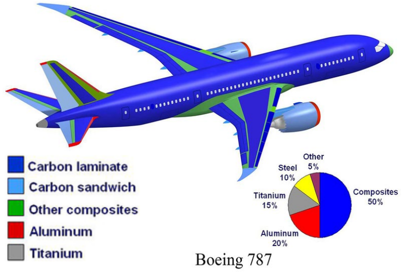
Prof. Robert Smith

Dr. Tom Marshall

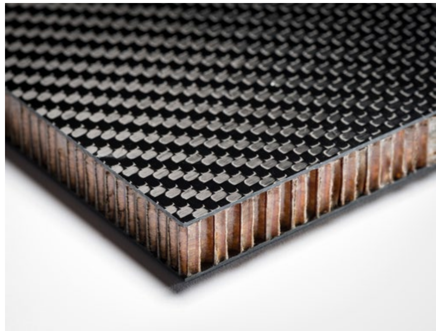
Bonded structures



- Sandwich panels and thick-section composites
  - Two face skins and a core (foam, balsa & honeycomb)
  - Uses: Aircraft interiors/exteriors, boats, wind turbines



- Defects (manufacturing or in-service)
  - Compromise structural integrity
  - Increases risk of failure
- Therefore, testing is required

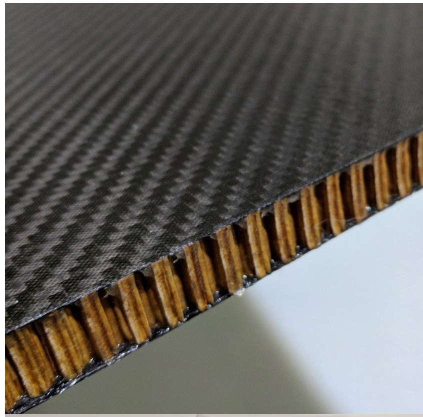


NDT of bonded structures

- These type of structures have posed a challenged for many years

- Conventional NDT methods struggle
  - **UT**: high attenuation & scattering, complex signals from internal geometry, couplant
  - **ECT**: can't be used on non-metallic parts, relatively shallow penetration
  - **Thermography and Radiography**: difficult to implement, expensive, have challenges and limitations

- Bond-testing NDT methods mitigates some of these issues
  - In most cases, no couplant is needed
  - Experiences far less attenuation
  - Signal are easy to interpret



The BondCheck

- Multi-mode bond testing instrument
  - Pitch-catch
  - Resonance
  - MIA (Mechanical Impedance)

- They work by exciting the structure and capturing the vibration response locally

"Two panes" view

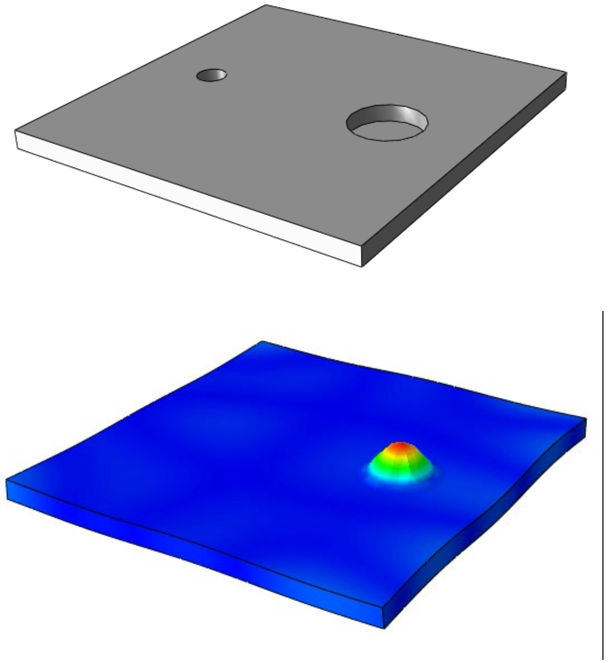
MIA probe

Resonance probe

Pitch-catch probe

Current methodology

- What is the current methodology?
- Bond testing, like all NDT methods, has inherent limitations
- Sensitivity decreases when calibration is based on a single defect
- Single-frequency, short-pulse excitation
    - limits depth sensitivity
    - limits capability to fully characterize defects
- Most sensitive to large, shallow defects
- Can these issues be mitigated?



8 kHz excitation

Motivation

- There is a need to improve sensitivity to defects of varying sizes and depth

- There is a lack of ability to characterise defects reliably (including depth information), using low-frequency inspection techniques

- There is a need for a relatively inexpensive and easy-to-implement method, not requiring trained inspectors (automation)

<u>These can be achieved by:</u>

- Implementing excitation techniques and advanced data analysis to improve sensitivity and detection capability

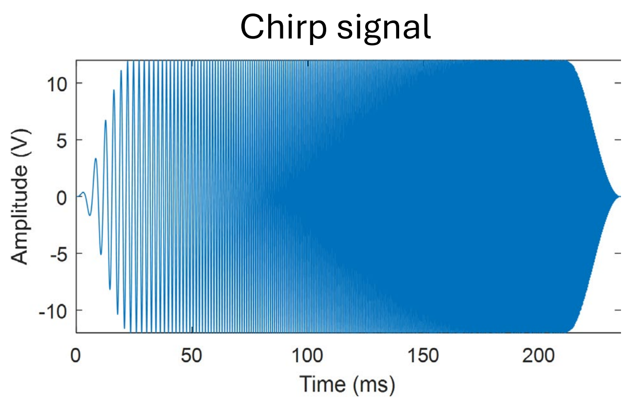
- The application of machine learning and statistical analysis to drive automation and improve predictive performance (classification).

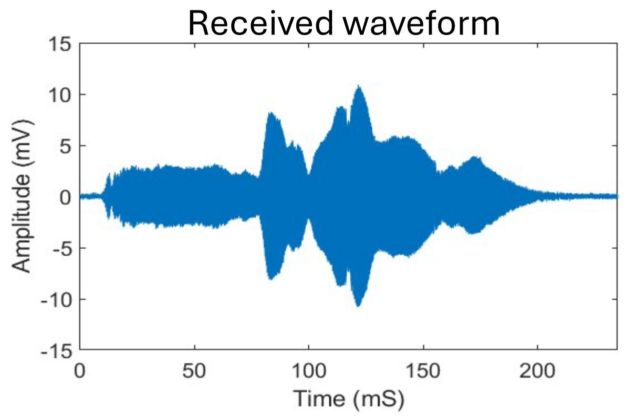
Methodology

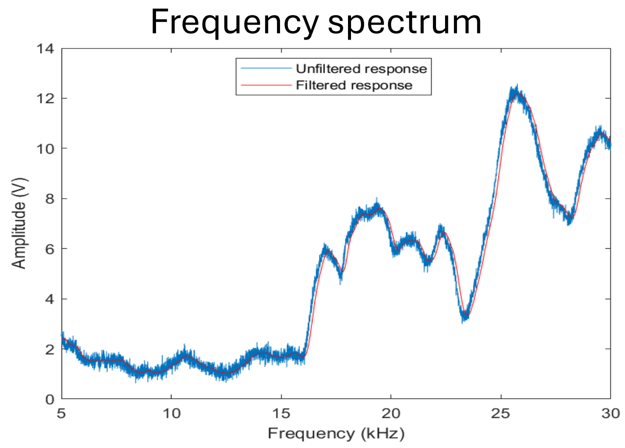
- Chirp excitation (5-30 kHz)
  - Improve sensitivity to a range of defects
  - Improve sensitivity to deep defects by exciting through-thickness resonance & harmonics

- Capture a full spectrum of information

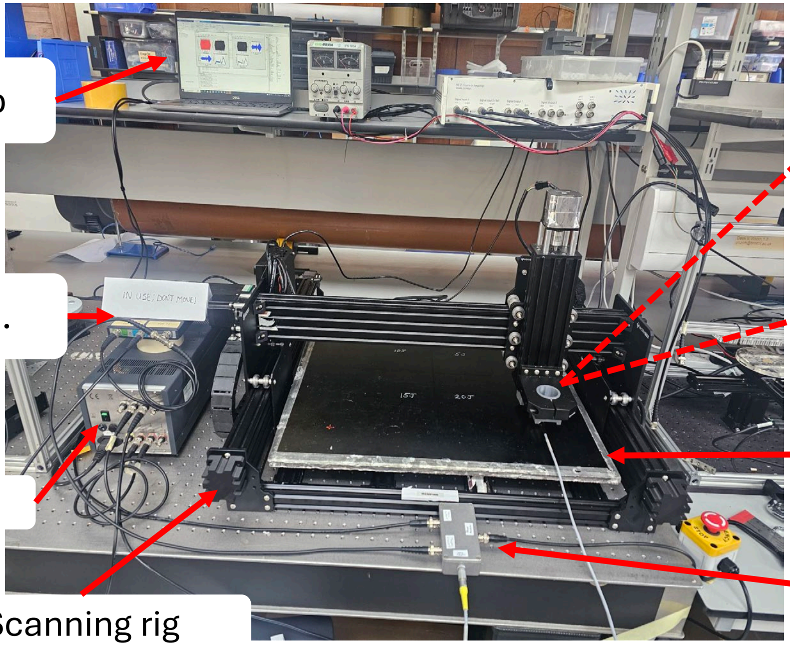
- Frequency-domain analysis on this data allows for:
  - Automated detection (for various defect sizes & depths)
  - Automated characterisation and classification of defects

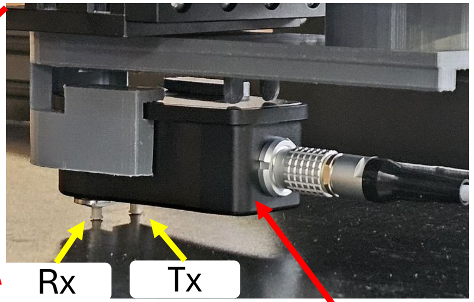


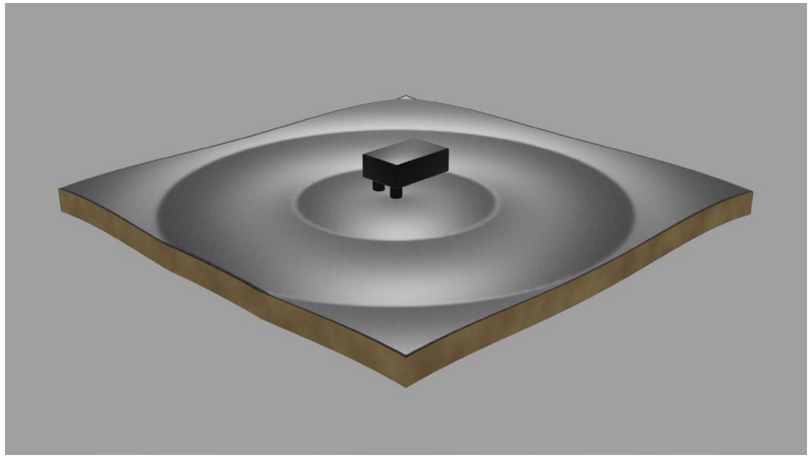


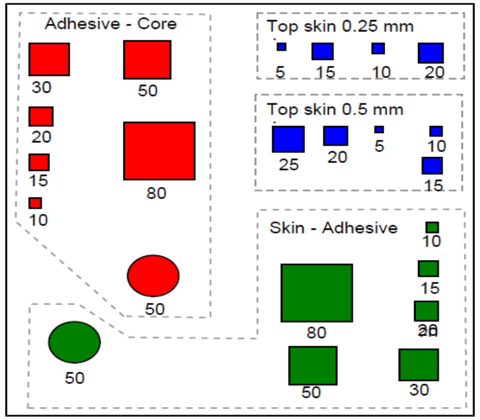


Demonstrator

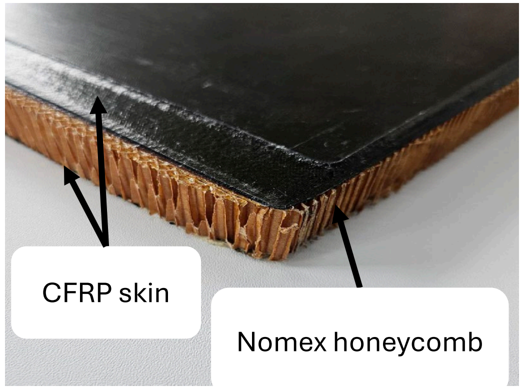








600 x 600 mm

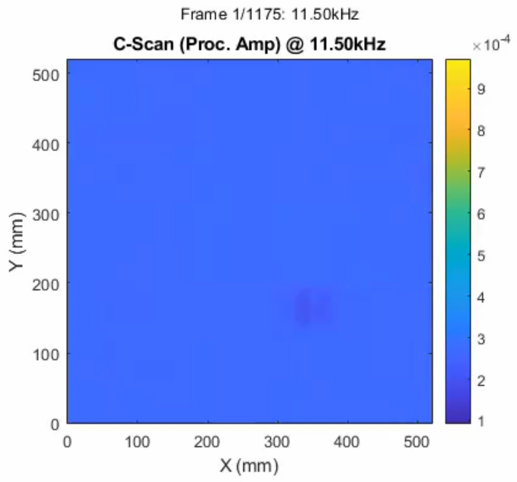


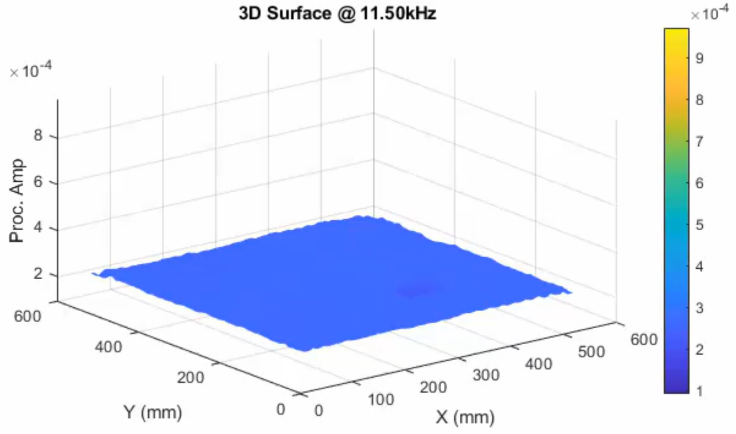
- Nomex honeycomb sandwich panel
- The panel has a range of artificially manufactured defects
- The panel is scanned in increments of 3mm
- The response is monitored and captured with a computer interface
- Scans are performed for the near-side & far-side for comparison

C-scan plots

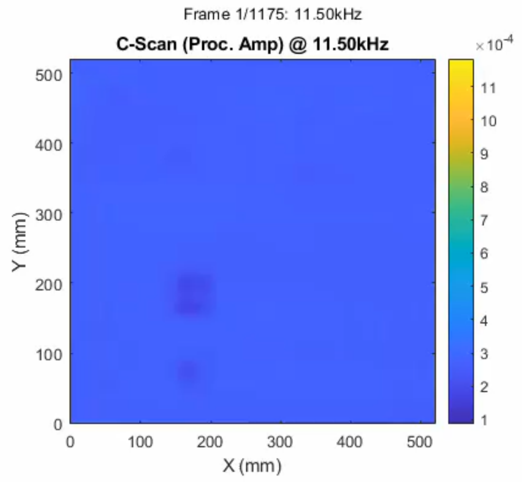
- Visualised the amplitude response as a function of the frequency
- Negligible drop in sensitivity to far-side defects when compared to near-side
- Detection remains dependent on the frequency using this method

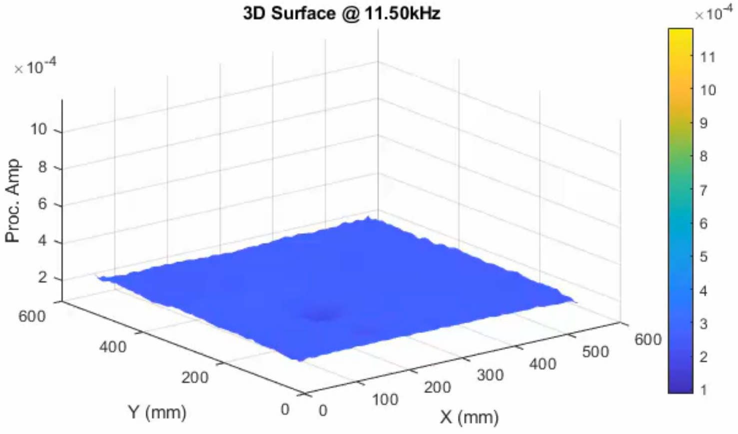
Near-side scan





Far-side scan





Detection methodologies

Contrast frequency method (baseline)

- Emulates current industry practice for bond testing
- Requires calibration for optimal performance
- Typically a single-frequency excitation
- Detection is frequency-dependent

RMSD algorithm

- Integrates information from multiple frequencies
- No prior calibration or knowledge of the structure is needed
- Detection is not frequency dependent
- Can be fully automated

Machine learning (anomaly detection)

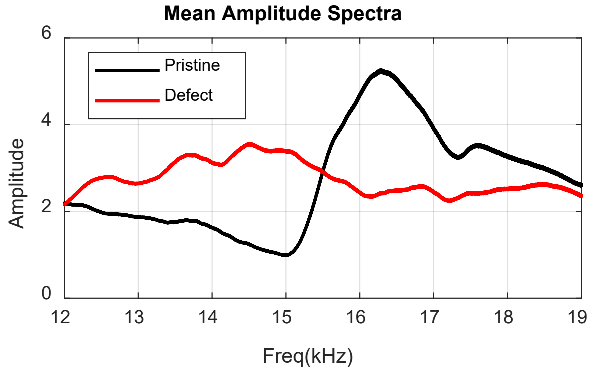
- Integrates information from multiple frequencies
- Can learn form a host of spectral features simultaneously (amplitude, phase, band power, frequency shifts, etc)
- Can be automated after initial training

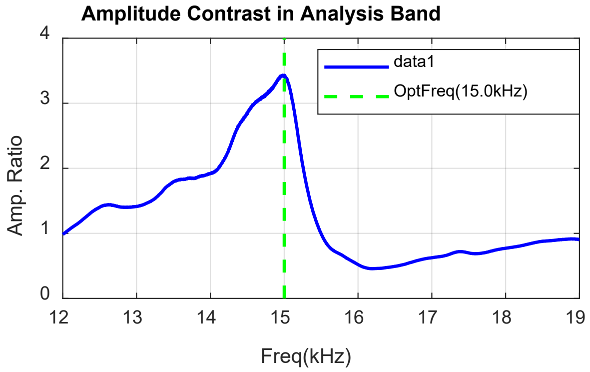
Method 1: Contrast frequency method (baseline)

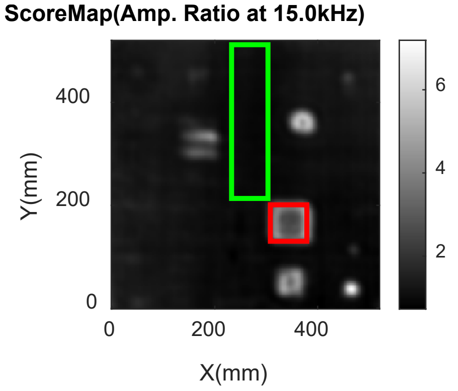
$$f_{opt} = \arg\max_{f} \left| \frac{\bar{A}_d(f)}{\bar{A}_p(f)} \right| \qquad \text{Amplitude}$$

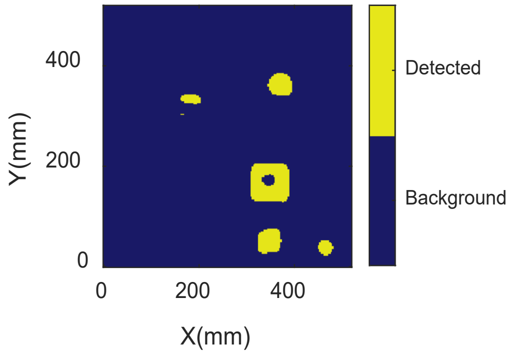
$$f_{opt} = \arg\max_{f} |\bar{\phi}_d(f) - \bar{\phi}_p(f)| \qquad \text{Phase}$$

- Contrast frequency method (baseline)
  - Emulates current industry practice
  - Inspection sensitivity is calibrated to a single defect
- First, an **optimal frequency** for inspection is chosen by the operator, using a calibration specimen
- Then, a score is calculated for each pixel. This is done by normalizing its response at the chosen frequency against the mean response of a user-selected pristine region.
- A **score map** is generated for the entire panel using these calculated scores.
- A threshold corresponding to 90% POD of the user-selected defect is used to produce a binary detection mask





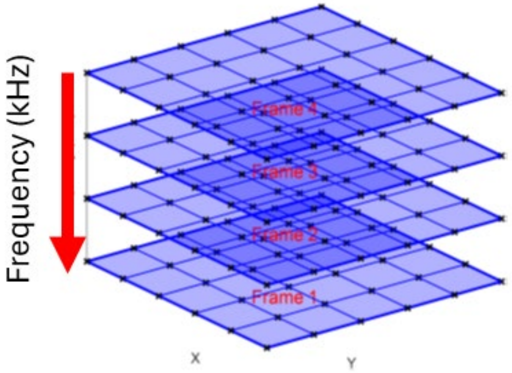


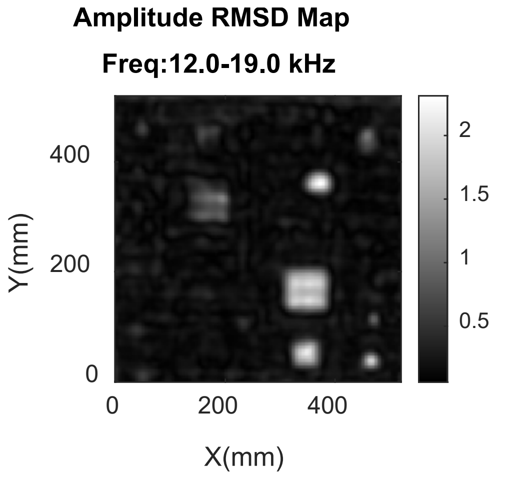


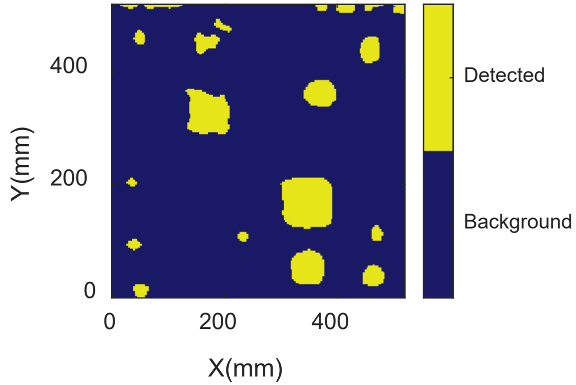
Method 2: RSMD algorithm

$$RSMD(x,y) = \sqrt{\frac{1}{N_f}\sum_{i=1}^{N_f}\left(s(x,y,f_i) - M(f_i)\right)^2}$$

- RMS deviation (RSMD) algorithm for automated detection
- Quantifies the deviation from a reference signal (in this case the median of the responses) across a spectrum of frequencies
- Selection of the reference signal is automated
- Plots the pixel-by-pixel output as a 2D spatial score map
- By applying a threshold of the 90th percentile value from all values in the score map, we can binarize the output into a mask





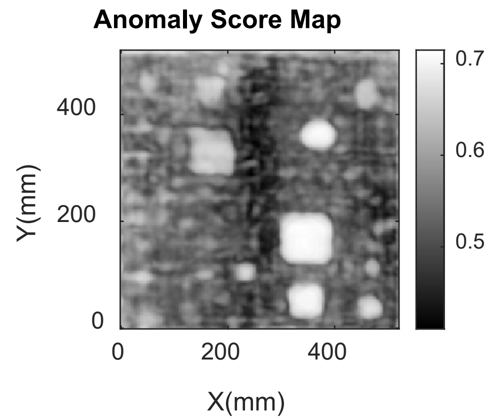


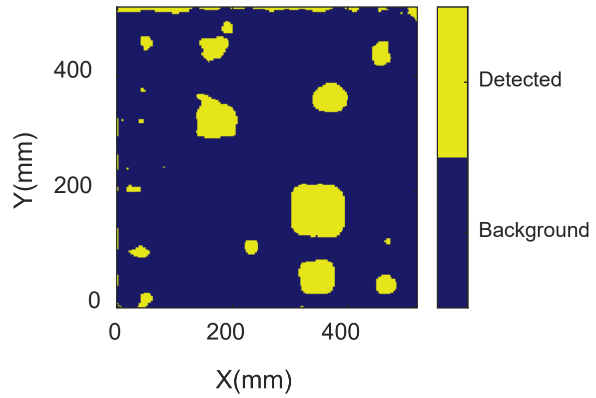
Method 3: Machine learning

- ML is used to implement automated anomaly detection
- Establish a reference median spectrum from pristine data $\tilde{A}_{ref}(f)$
- Quantify the deviation of each pixel's spectrum from this reference
- An **Isolation Forest model** is then trained on a collection of these residual spectra that are known to be pristine
- To generate the score map, an **anomaly score** is assigned to each spectrum, by the trained model
- A **high anomaly score** signifies an anomaly
- A 90th percentile threshold value of the score map is used to binarize the output in a mask

$$R_A(f) = \frac{A(f)}{\tilde{A}_{ref}(f) + e}$$ Amplitude

$$R_\phi(f) = \phi(f) - \tilde{\phi}_{ref}(f)$$ Phase

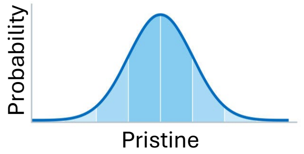




Anomaly detection workflow & operation

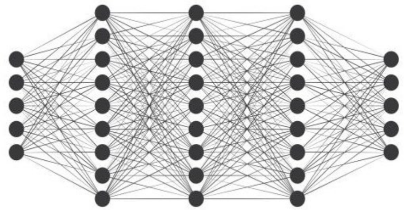
ML Workflow:

Defect-free training data → Chose suitable ML model (Decision trees or Neural networks) → Train the network or tree → Evaluate performance → Update model parameters and/or training data

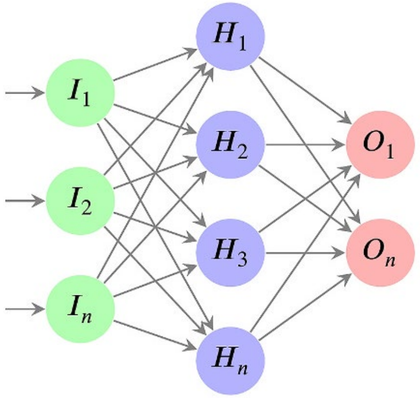


Isolation forest

Train the network or tree



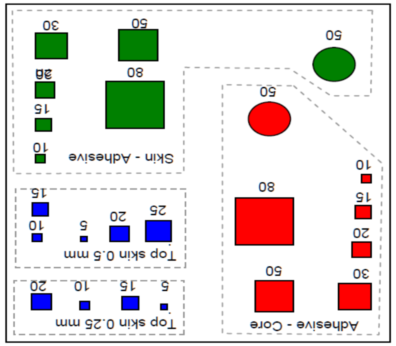
Operation:

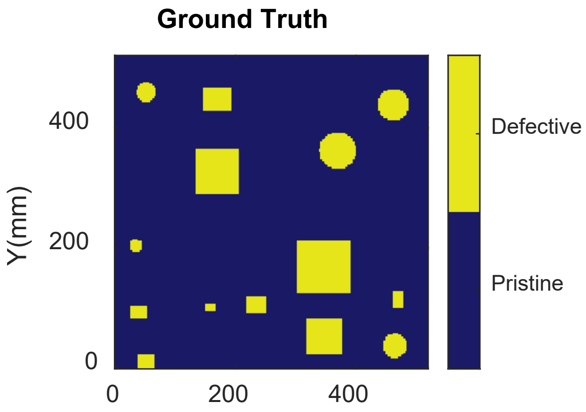
New measurement	Trained network or tree	Prediction
Amplitude Phase Engineered features		Normal (defect-free) Anomaly

Automated detection is based on a 90th percentile threshold

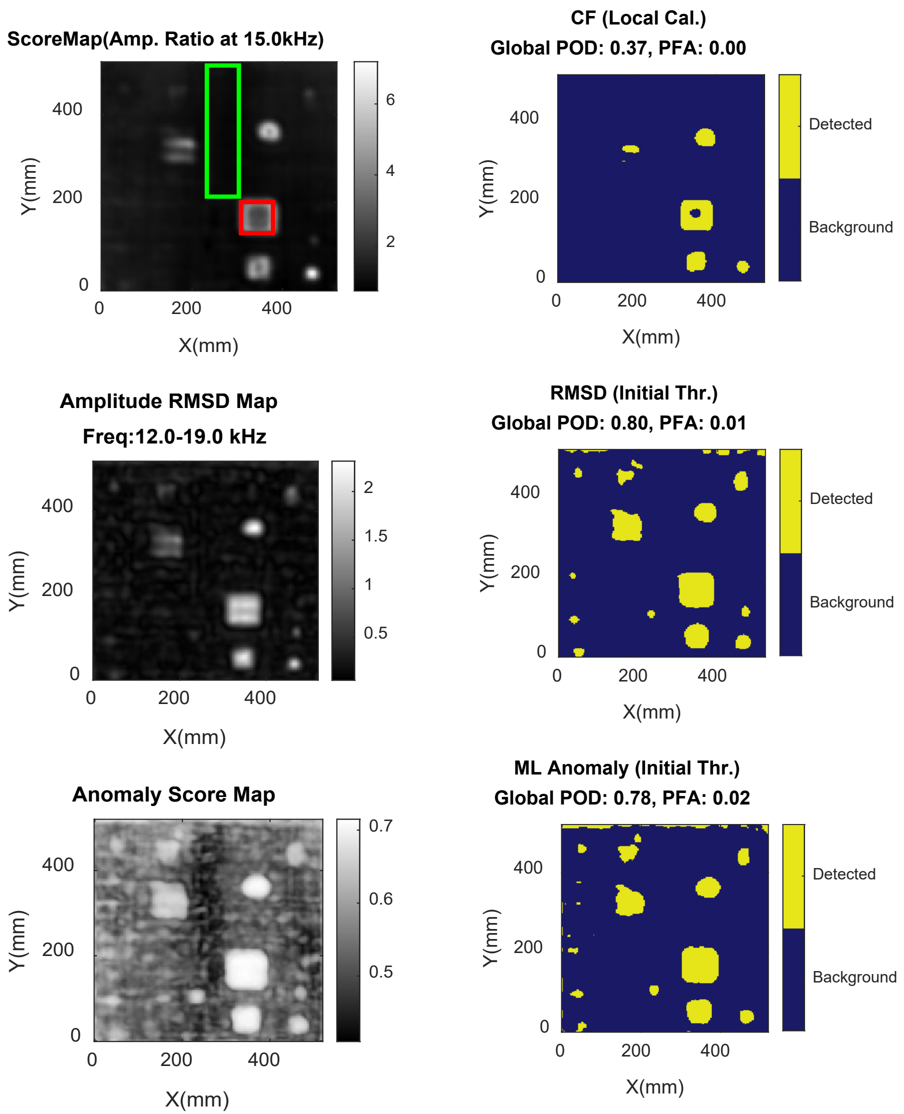
Establishing the ground truth

- The ground truth mask is a summary of what is detectible
- Detection is defined as any defect with a PFA of less than 40%, when it's excited at resonance
- Meaning it is distinguishable from the background noise
- Using this logic, 14 of 23 known defects are detectible using the amplitude information
- The ground truth mask is a reference, used to evaluate the performance of the proposed detection methods

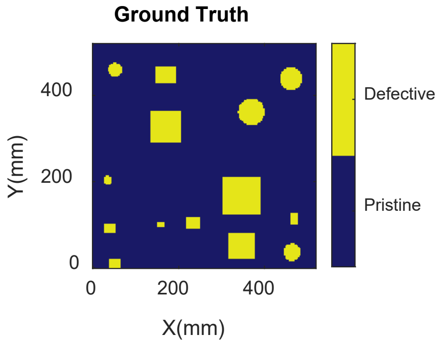


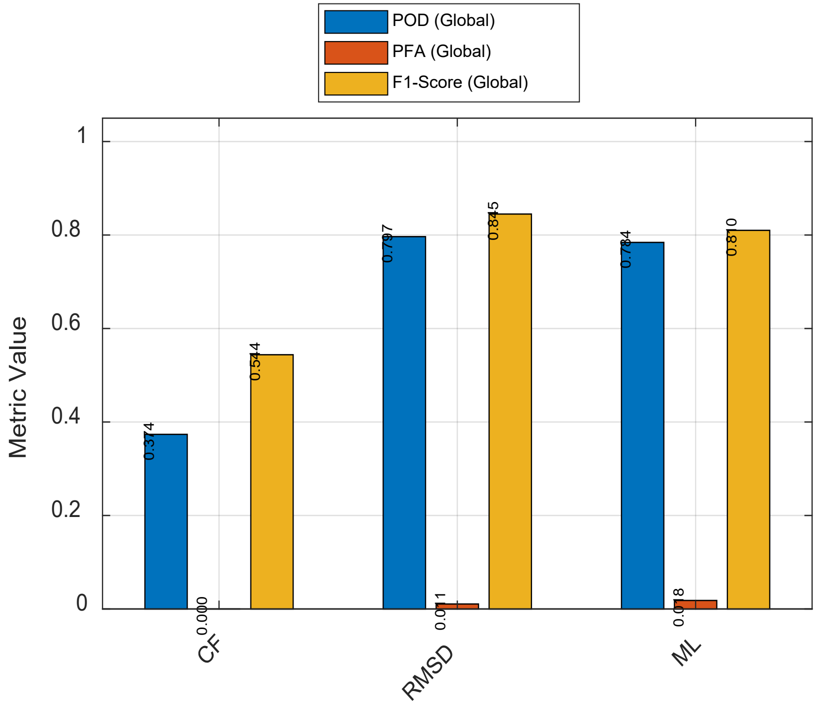


Comparison at 90% POD & percentile

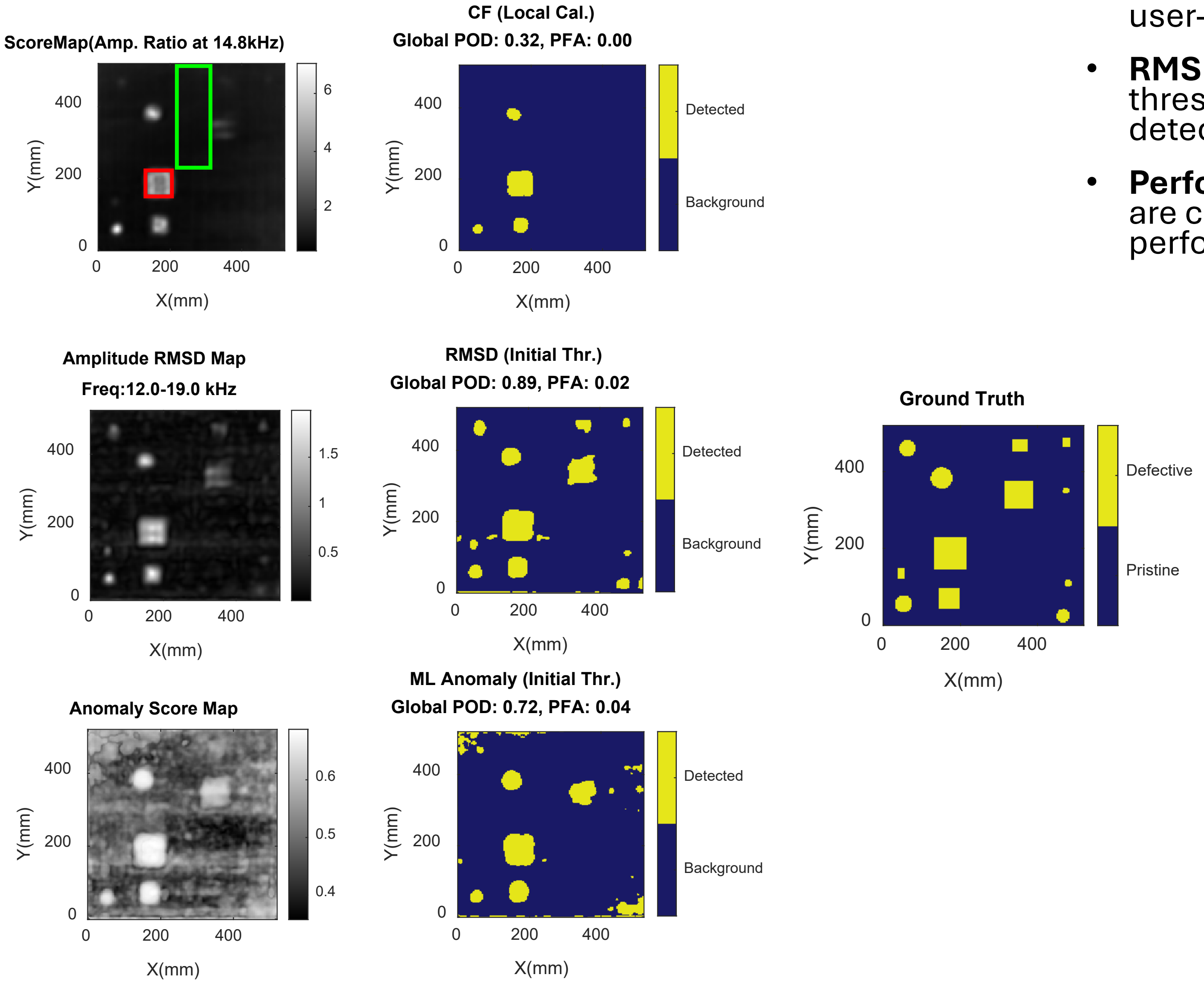


- **CF Method:** A threshold is established based on a 90% Probability of Detection (POD) for defective pixels from a user-selected defect, creating a binary detection mask
- **RMSD & ML Anomaly Detection:** The score maps are thresholded at the 90th percentile to create a binary detection mask
- **Performance Evaluation:** The resulting sensitivity maps are compared to ground truth data to measure performance.

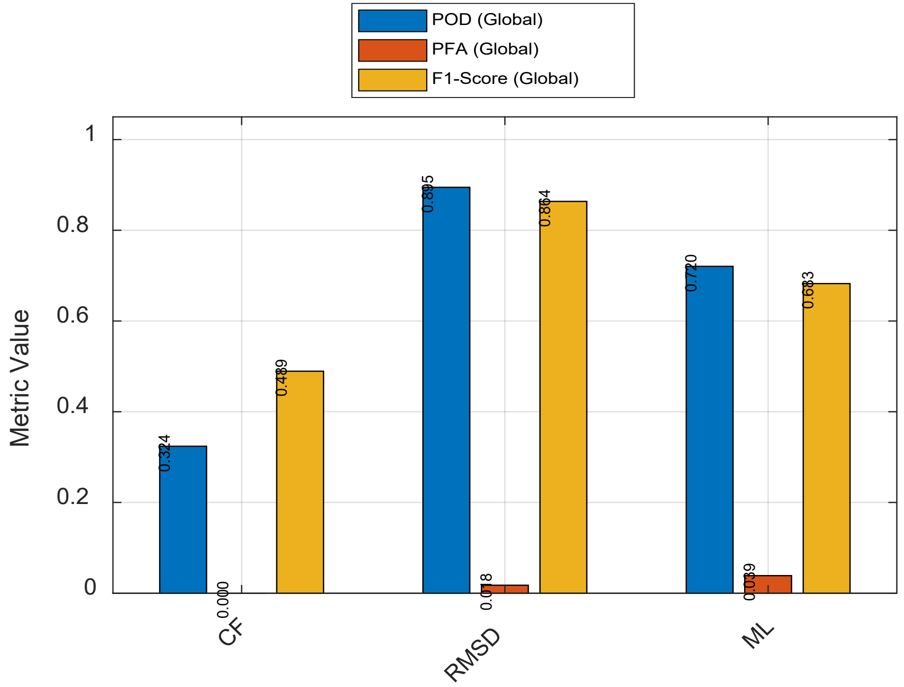




Far-side comparison

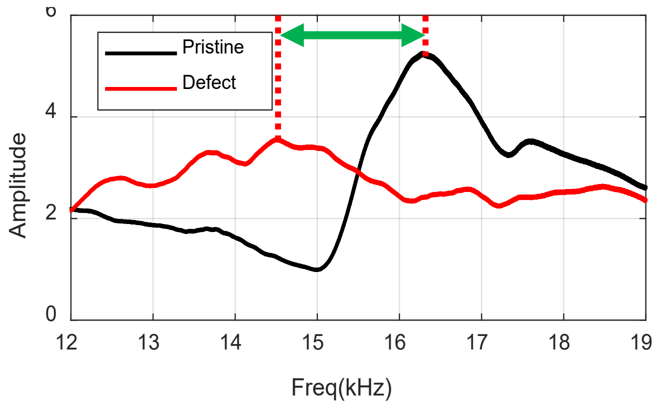


- **CF Method:** A threshold is established based on a 90% Probability of Detection (POD) for defective pixels from a user-selected defect, creating a binary detection mask
- **RMSD & ML Anomaly Detection:** The score maps are thresholded at the 90th percentile to create a binary detection mask
- **Performance Evaluation:** The resulting sensitivity maps are compared to ground truth data to measure performance.

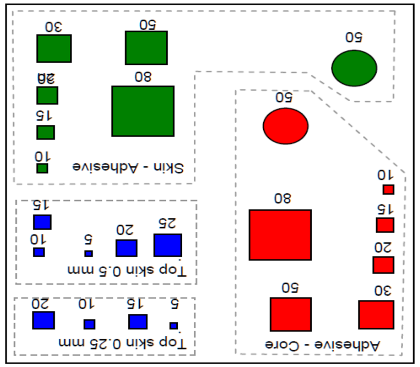


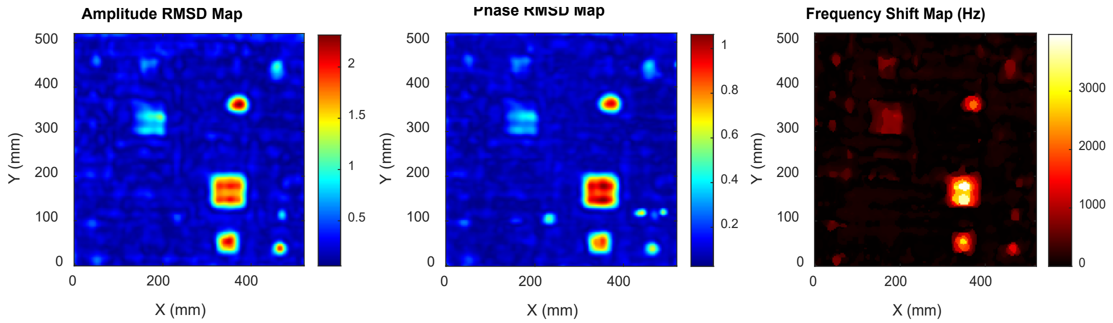
Multi-mode detection

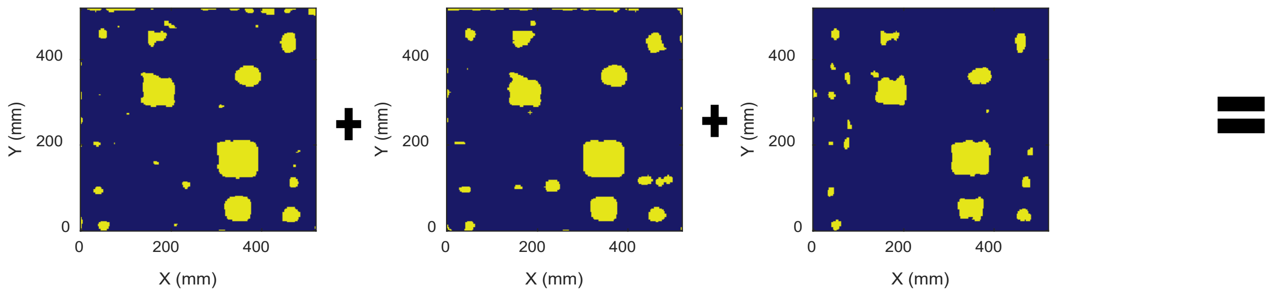
- Panel can be interrogated using various available information when a chirp excitation is used
- **Amplitude**, **phase** and **frequency shifts** (can detect subtle changes in local mass and stiffness)
- Fusion of these maps creates a comprehensive map of the defects, enhancing overall detection
- Fusion can lead to a slight increase in speckled noise but within the acceptable threshold



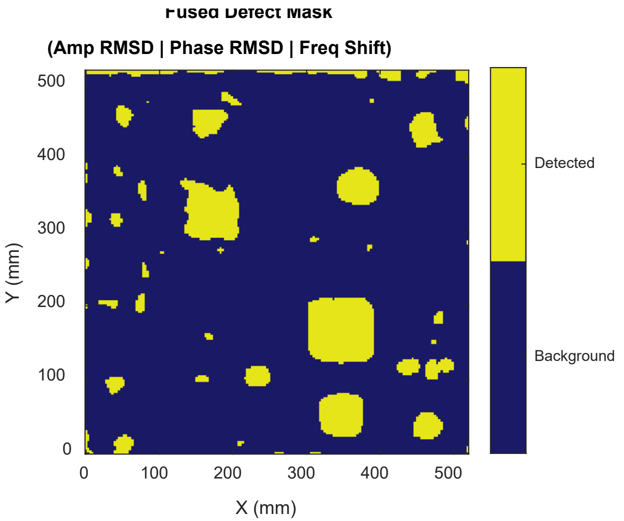
Resonance frequency shift







90th percentile threshold detection maps



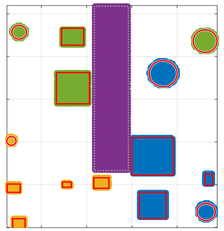
Conclusions on detection

- RSMD and ML methods show consistent improvement over the currently implemented methodology (CF)
- Careful selection of the pristine training data, and fine-tuning of the model parameters is needed for accuracy using ML
- RSMD method appears to be the most reliable, however, the accuracy of this method is dependent on the quantity of ‘normal’ response when calculating the median reference spectrum (it relies the ‘normal’ response being the most common)
- A chirp signal allows for panel interrogation using various available information (amplitude, phase and frequency shift)
- Fusion of these detection modes can lead to improved detection capability compared to using a single mode for detection

Defect classification workflow & operation

ML Workflow:

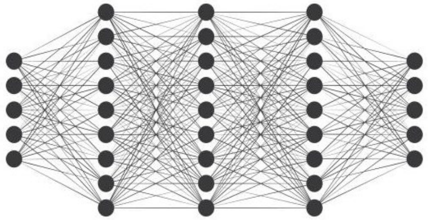
Labelled training data → Define data split (train/Val) → Chose suitable ML model (Decision trees or Neural networks) → Train the network or tree → Evaluate performance on validation data → Update model parameters and/or training data → Train the network or tree



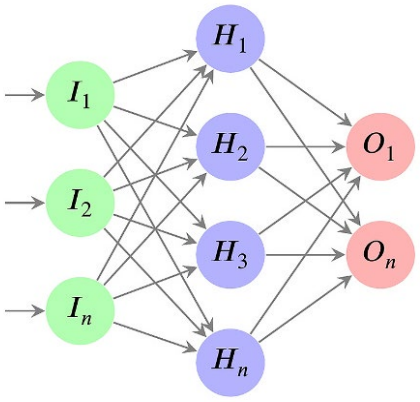
80/20

Random forest

Train the network or tree



Operation:

New measurement	Trained network or tree	Class prediction
Amplitude Phase Engineered features		Response type: Pristine or a defect class

Defect classification with ML

- Residual spectrum (amplitude & phase) for each pixel is calculated relative to the median spectrum of a carefully selected pristine dataset

$$R_A(f) = \log\left(\frac{A(f)}{\tilde{A}_{ref}(f) + e}\right)$$ Amplitude

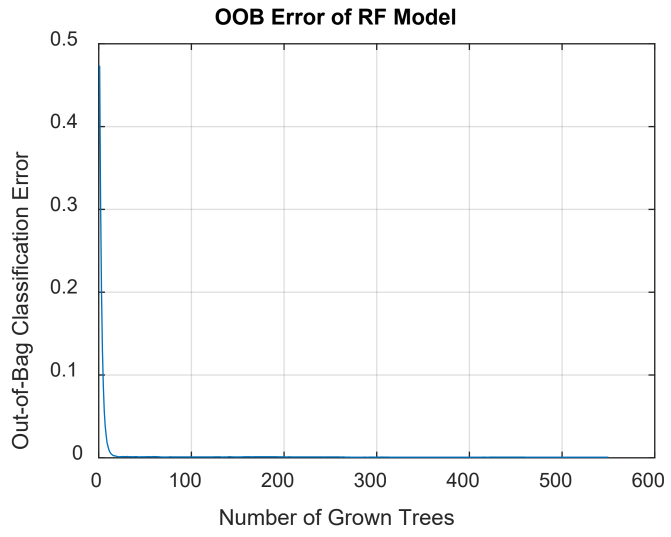
$$R_\phi(f) = \phi(f) - \tilde{\phi}_{ref}(f)$$ Phase

- Then class-specific template were generated from the training data, using a **prototype spectrum** (relating to the class central tendency – median)

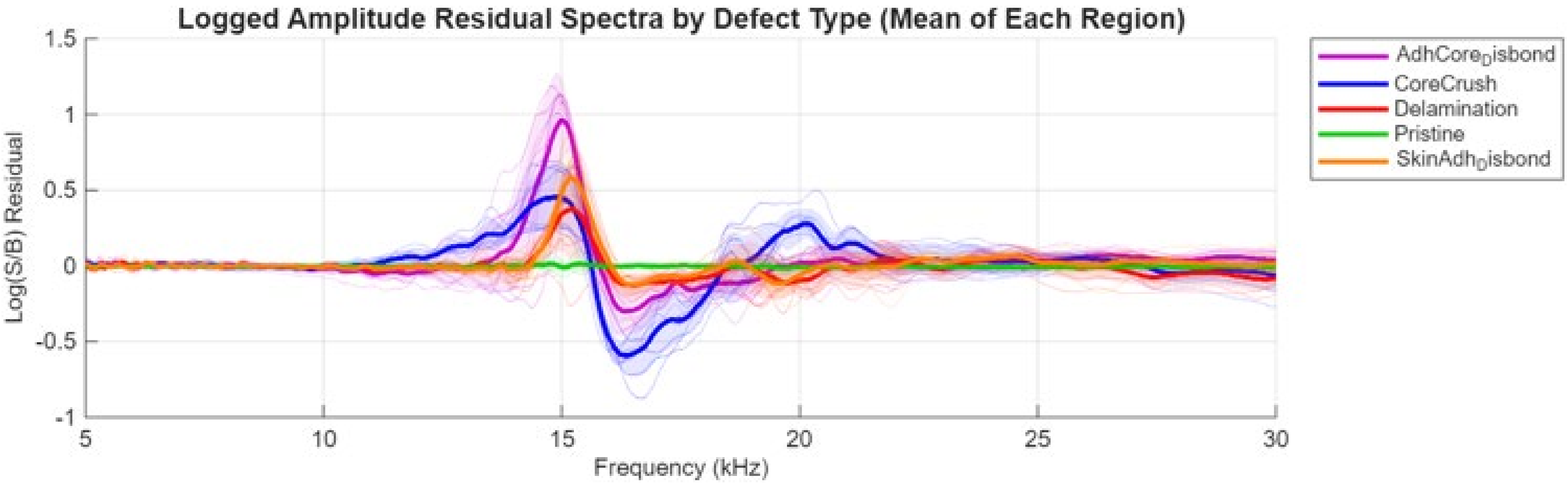
- For each pixel within a class training set, the **Euclidean distance**, **Cosine similarity** and **Mahalanobis distance** are calculated using the **prototype spectrum** and fed to the ML model as engineered features

- Synthetic Minority Over-sampling Technique (SMOTE), and costing which penalises wrong classification of lesser classes, was used to improve the class imbalance of the training data

- Out-of-Bag errored is monitor for model convergence



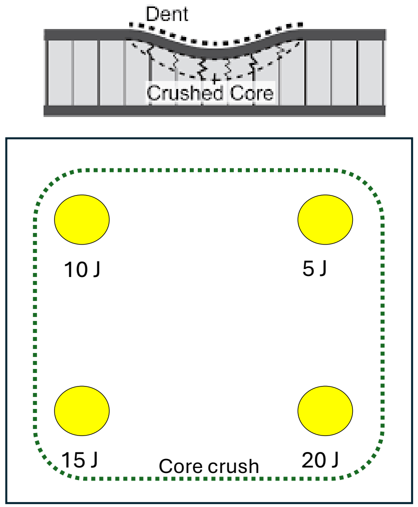
The defect fingerprint



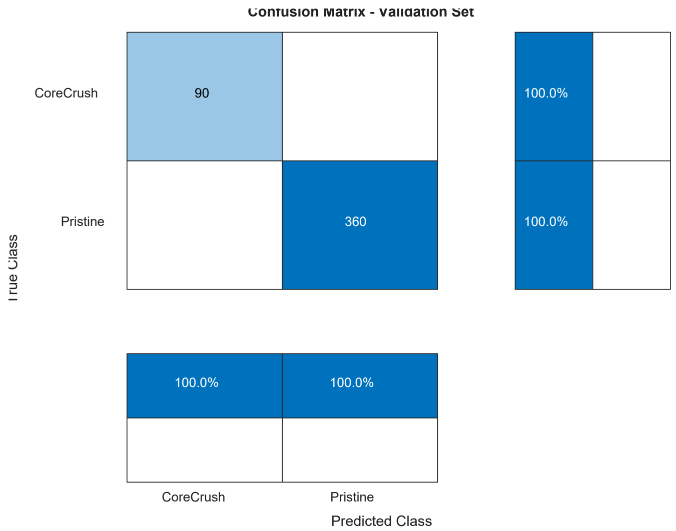
- Spectral information for training the ML models
- Statistical engineered features were derived from this information to increase the accuracy of the models

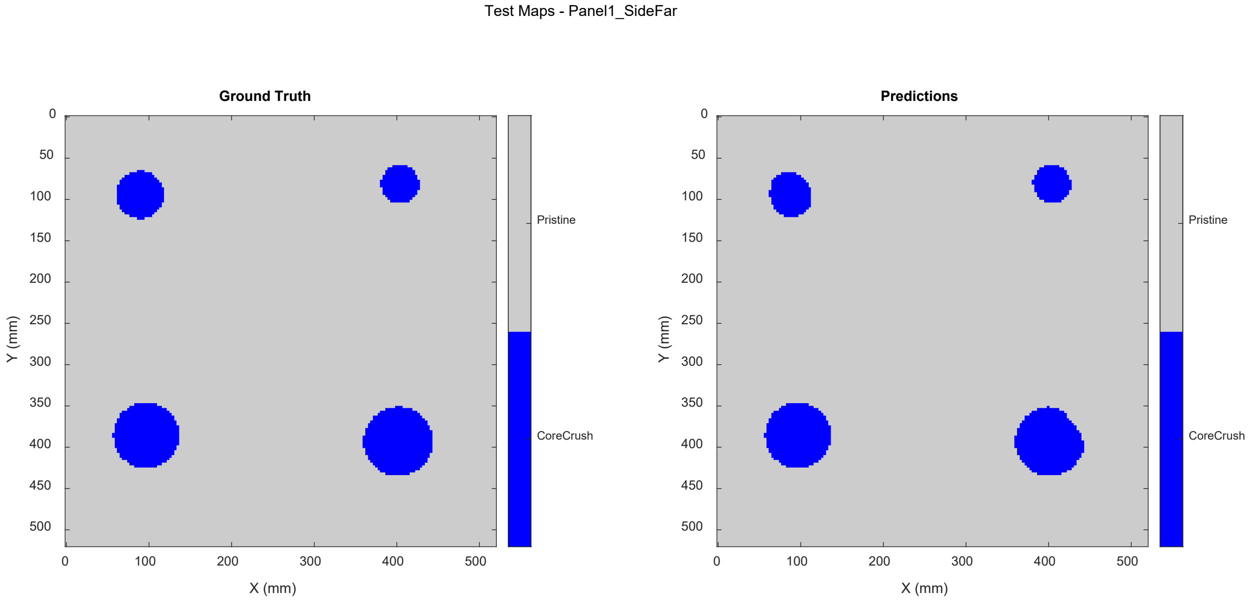
Scenario 1: 2 response classes

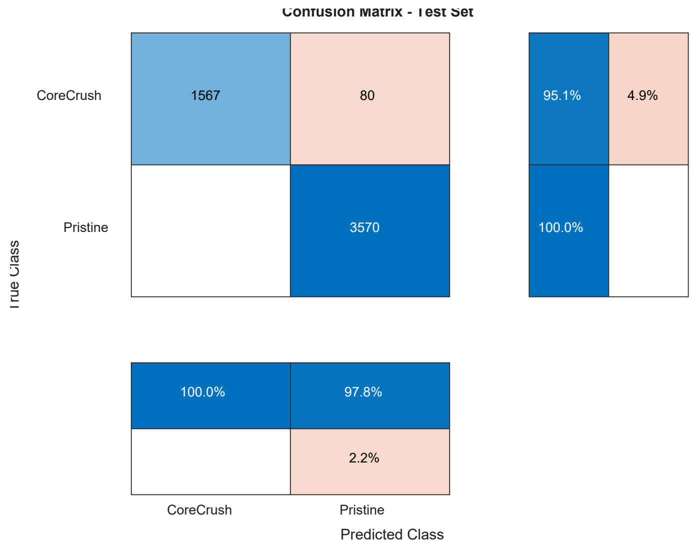
- Skin 1mm, Core 10mm
- Core-crush created with blunt impact of varying energies
- Train/Val on near side data, with an 80/20 split
- Tested on unseen data from the far-side scans



600 x 600 mm

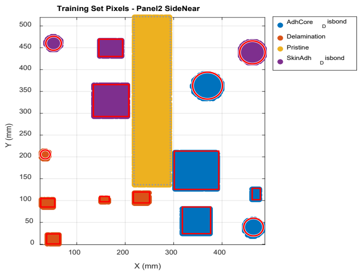






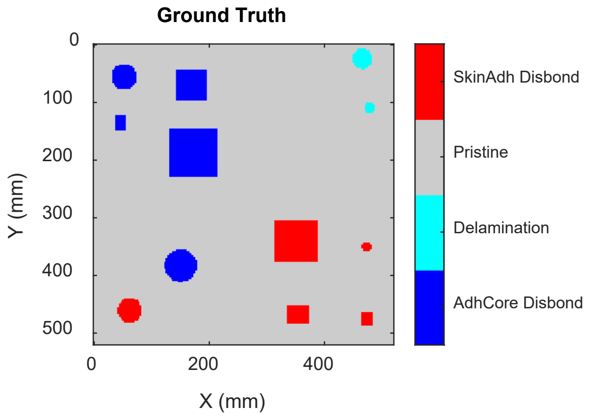
Scenario 2: multiple responses

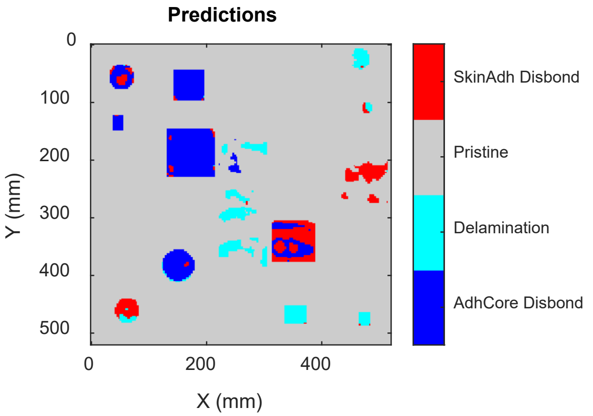
- Skin 2mm, Core 20mm
- Artificially manufactured defects
- Train/Val on near side data, with an 80/20 split
- Tested on unseen data from the far-side scans

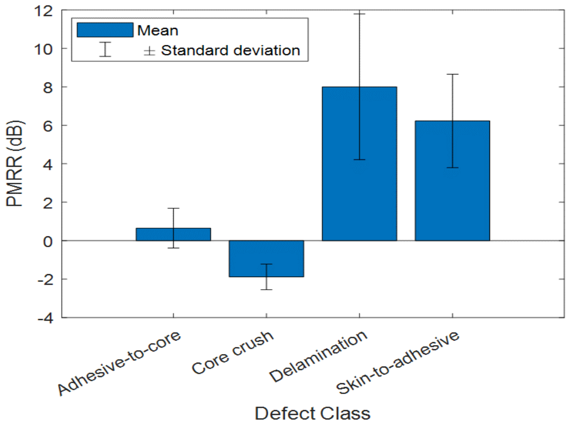


Confusion Matrix - Validation Set

True Class \ Predicted Class	AdhCore_Disbond	Delamination	Pristine	SkinAdh_Disbond		
AdhCore_Disbond	177		1		99.4%	0.6%
Delamination		31			100.0%	
Pristine			325		100.0%	
SkinAdh_Disbond				98	100.0%	
	100.0%	100.0%	99.7%	100.0%		
			0.3%			







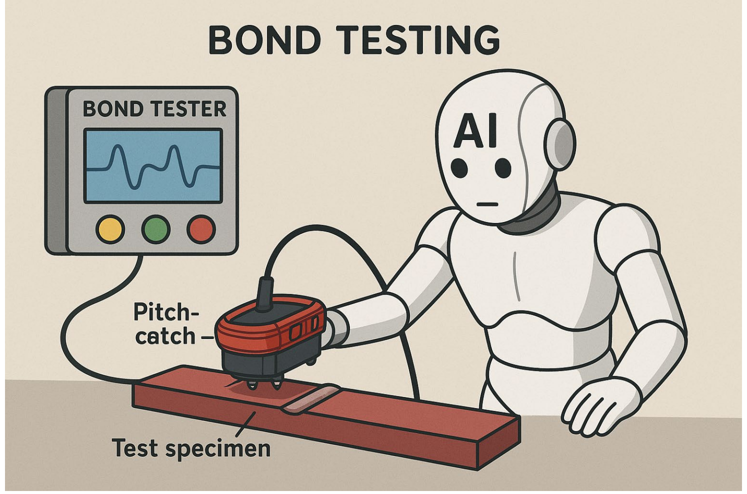
Confusion Matrix - Test Set

True Class \ Predicted Class	AdhCore_Disbond	Delamination	Pristine	SkinAdh_Disbond		
AdhCore_Disbond	1493	21	2	91	92.9%	7.1%
Delamination		109	13	15	79.6%	20.4%
Pristine	44	387	4525	235	87.2%	12.8%
SkinAdh_Disbond	203	227	32	520	53.0%	47.0%
	85.8%	14.7%	99.0%	60.4%		
	14.2%	85.3%	1.0%	39.6%		

Limitations & implications of ML

Limitations:

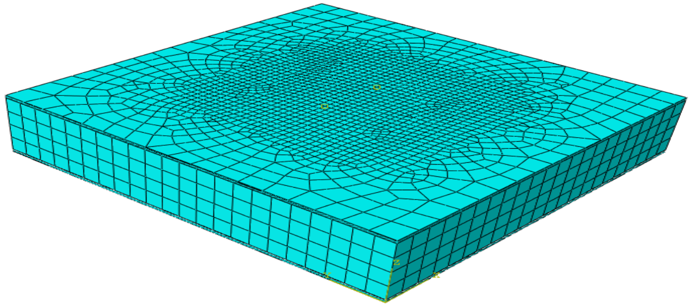
- Dependent on quality of training data
- The model requires careful parameter tuning
- It is sensitive to geometry (defect response is learnt in context of the panel geometry)



Implications:

- Enables interpretation of new measurements with high confidence
- Enable real-time automatic characterisation and response prediction, including depth information

Future work



Finite element!!

- Detection and characterisation of kissing-bond defects
- Transition from **data-driven** → **physics-based** prediction
- Data-driven: Uses historical data to make predictions for new measurements
- Physics-based: A digital twin of the panel can generate large amounts of synthetic training data, improving class prediction and depth estimation
- Can also be achieved via model-inversion algorithms (optimization-based or database-driven approaches)
- The aim is to reduce reliance on "**known patterns**" and move towards capturing the **underlying physics** of the vibration system
- **Explainability**!!

Acknowledgements

- Supervision team: Prof. Bruce Drinkwater, Prof. Robert Smith, Dr Tom Marshall
- QinetiQ: Test specimens
- University of Bristol: Laboratory & Software

Thank you for listening!
Questions?