Transfer Function Model Assisted Probability of Detection for Lap Joint Multi Site Damage Detection

Quantitative Nondestructive Evaluation Conference 2011

July 18-22, 2011

Michel Bode, Justin Newcomer, Stephanie Fitchett

Sandia National Laboratories

Airworthiness Assurance NDI Validation Center (AANC)

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

MAPOD for WFD Detection
Outline

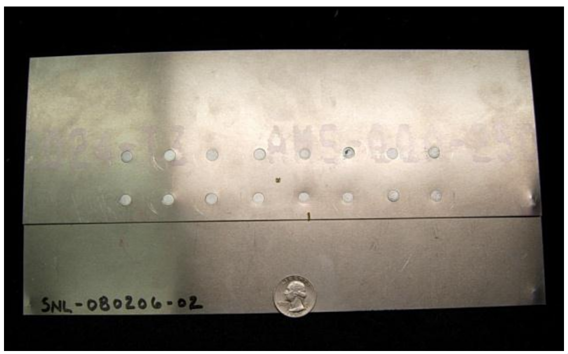
- Transfer function approach based on POD model parameters
- Unique validation because of having samples from Q4 – real defects characterized in retired real airplane structure
- Specimens and basic POD results from inspections
- Simple Transfer model
- First Approach – linear regression verified using bootstrapped parameter estimates
- Second Approach – multivariate regression from bootstrapped hit/miss data
- Results & Conclusions

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

Basic Specimen Setup

- 4 Inspectors (varying experience levels)
- 4 Sets of POD Specimen Panels

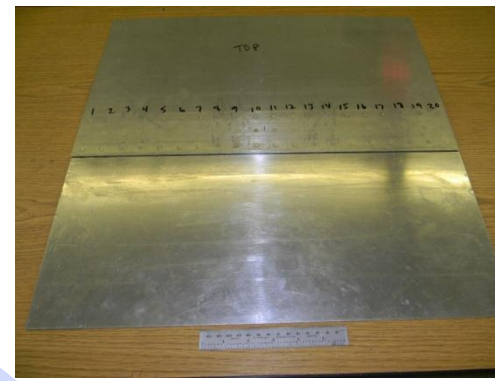


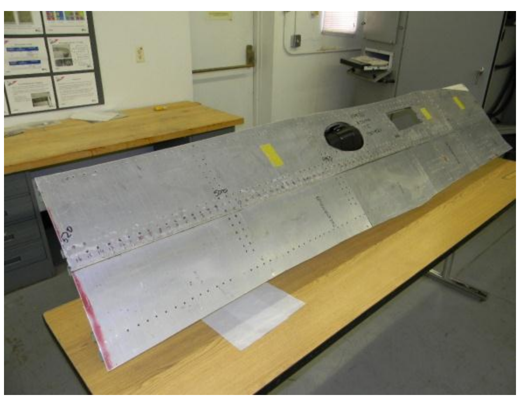
STRUCTURE

FLAW

Quad 1
.039-.141 flaw distribution
EDM Flaws, Simulated Structure

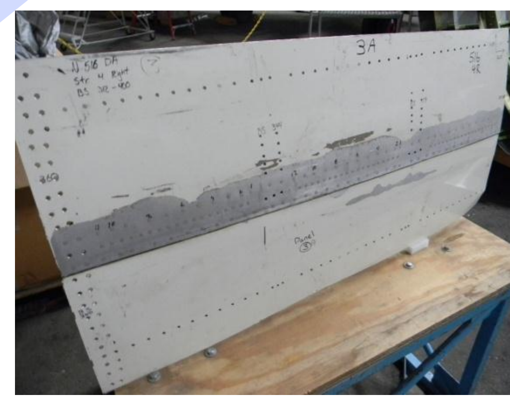
Quad 2
.035-.199 flaw distribution
Engineered Cracks, Simulated Structure





Quad 3
.010-.400 flaw distribution
EDM Flaws, Real Structure

Quad 4
.019-.274 flaw distribution
Real Cracks, Real Structure Paint

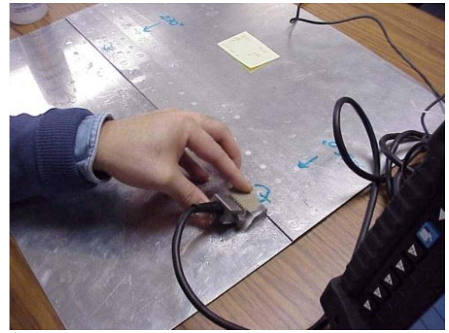


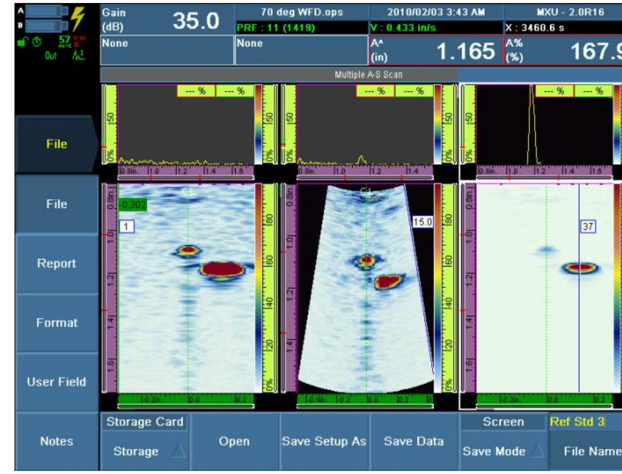
This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

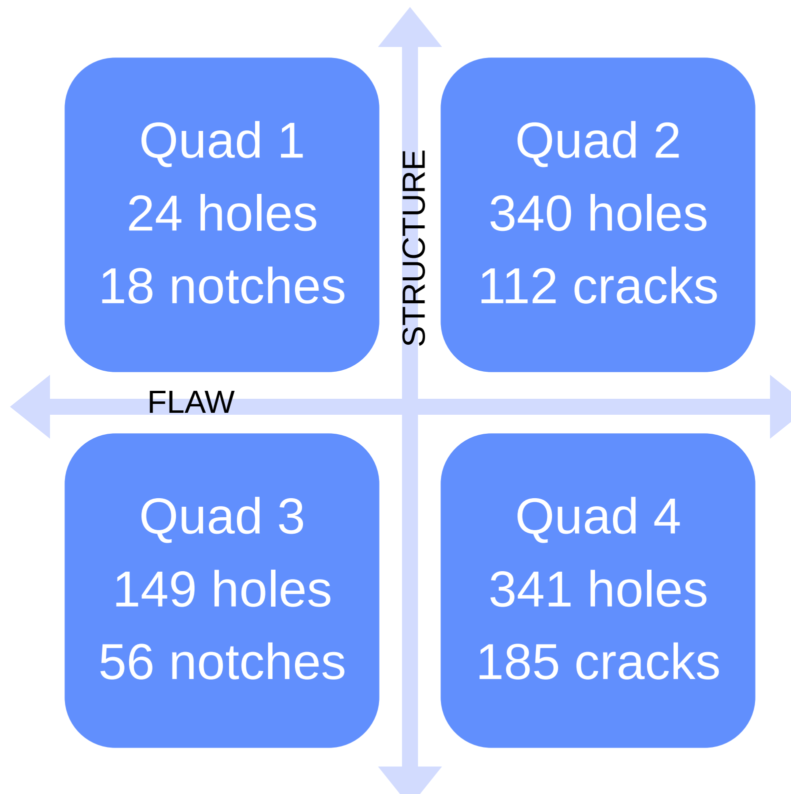
Inspection

- Method used: Ultrasonic Linear Array Shear Wave on Omniscan Instrument, 10L64 probe (10MHz, 64 element) with 70° wedge





- Each hole was inspected for flaws on the left and on the right sides, providing 2 detection opportunities per hole, each treated independently.
- Quad 4 panels included holes with multiple flaws on one or both sides, and a multi-flaw model was used
- All defect lengths are inches.



This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

2 Parameter Probit Model for POD Hit/Miss Data in each Quadrant

$$\text{POD} = \Phi(\text{Intercept} + \text{Slope}*\ln(\text{length}))$$
$$= \Phi(c + d*\ln(\text{length}))$$

where Φ is the cumulative normal distribution

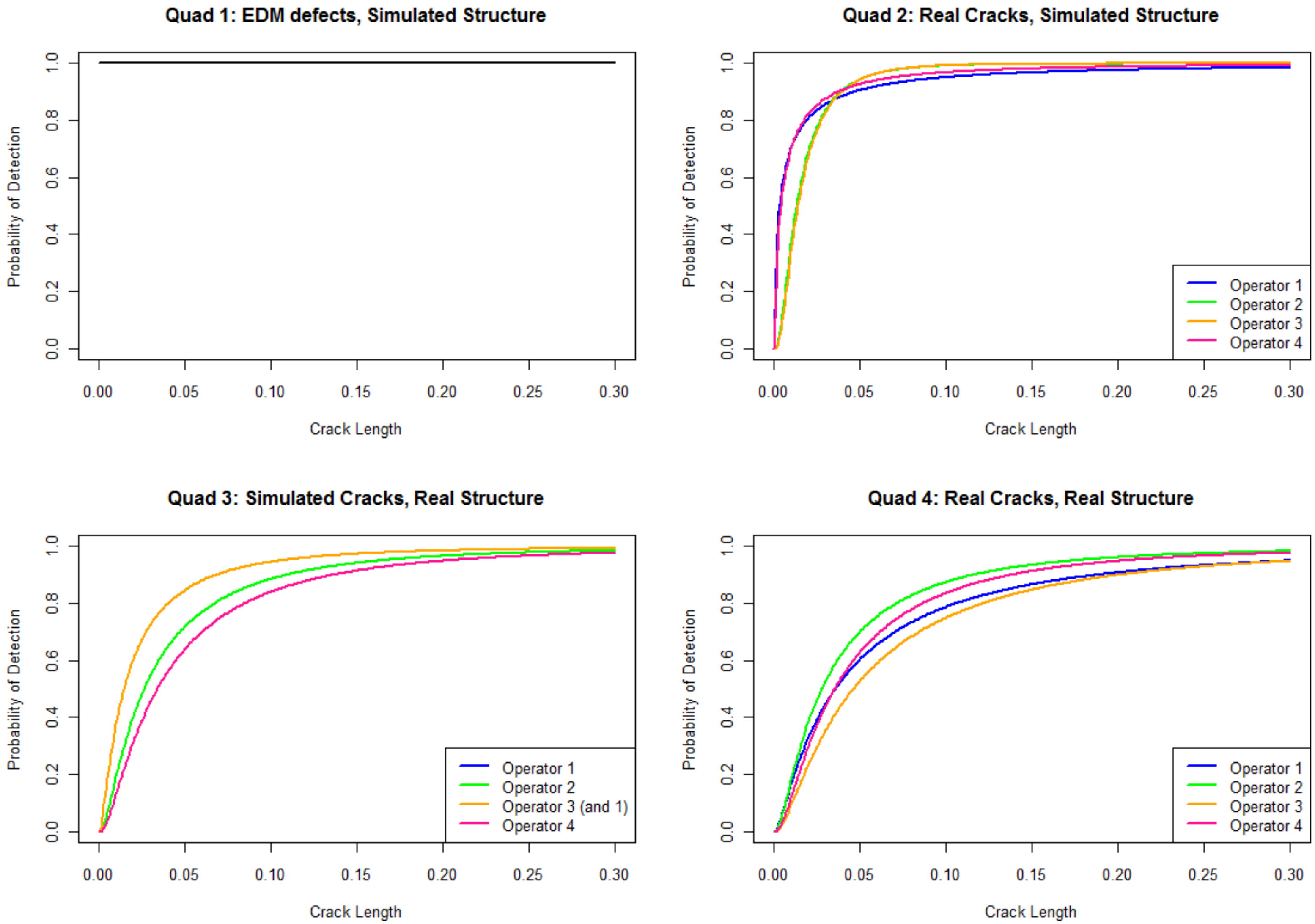
$$\text{POD}_{0.90} = \exp\{(\Phi^{-1}(0.90) - \text{Intercept})/\text{Slope}\}$$

= smallest crack length for which POD is at least 90%

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

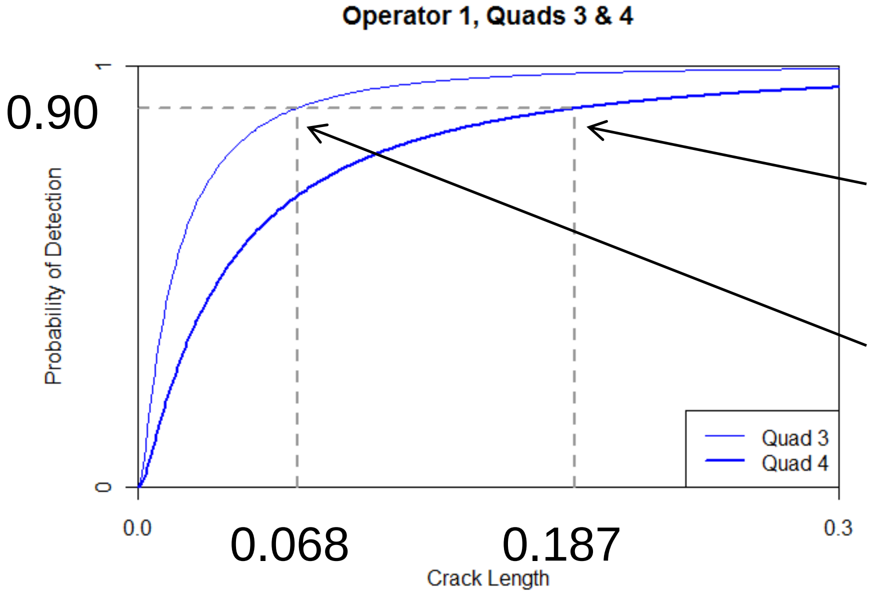
POD curves by Quadrant



Sandia National Laboratories

Parameters for POD curves

	Quad 2			Quad 3			Quad 4		
Operator	Int c_2	Slope d_2	$POD_{0.90}$	Int c_3	Slope d_3	$POD_{0.90}$	Int c_4	Slope d_4	$POD_{0.90}$
1	2.78	0.49	**0.046**	3.52	0.83	**0.068**	2.57	0.77	**0.187**
2	5.18	1.20	**0.039**	3.28	0.90	**0.109**	3.22	0.90	**0.116**
3	5.32	1.25	**0.039**	3.52	0.83	**0.068**	2.68	0.87	**0.200**
4	3.20	0.58	**0.037**	3.12	0.92	**0.136**	3.15	0.94	**0.138**



$$POD = \Phi(2.57 + 0.77 * \ln(\text{length}))$$
$$0.9 = \Phi(2.57 + 0.77 * \ln(0.187''))$$

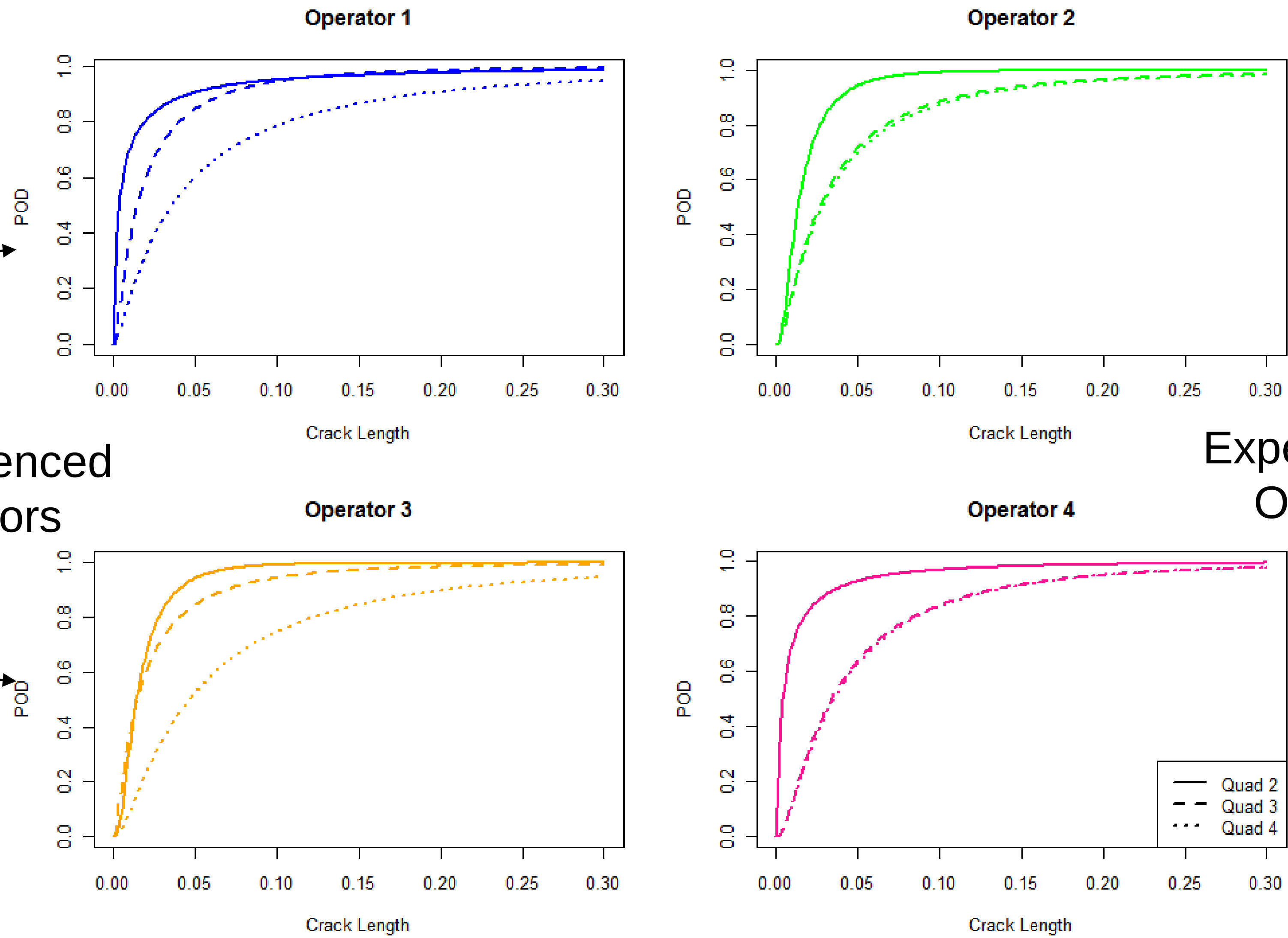
$$POD = \Phi(3.52 + 0.83 * \ln(\text{length}))$$
$$0.9 = \Phi(3.52 + 0.83 * \ln(0.068''))$$

Sandia National Laboratories

POD curves by Operator

Less Experienced Operators

More Experienced Operators



False Call Rates

	Quad 2			Quad 3			Quad 4		
Operator	False calls	Opps	Rate (%)	False calls	Opps	Rate (%)	False calls	Opps	Rate (%)
1	28	568	4.9	12	242	5.0	40	497	8.1
2	25		4.4	1		0.4	29		5.8
3	45		7.9	7		2.9	56		11.3
4	26		4.6	2		0.8	43		8.7

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

Model Assisted POD Generation

- Ideally, would like to estimate slope and intercept parameters for Q4 PODs from those in Q1, Q2, and Q3
  - Q1 provides minimal information for fitting
  - too few operators to estimate Q4 parameters jointly from Q2 and Q3 parameters, which would be best option
- First look at POD parameter estimates and their relationships
  - fairly strong linear relationships between c & d within a quadrant
  - some relationship between Q3 and Q4 parameters
  - negligible relationship between Q2 and Q4 parameters
- With data from 4 operators, we can try simple models
  - predict c_4 from c_2, c_3 & predict d_4 from d_2, d_3 using linear regression
    - $c4 = \beta0 + \beta1*c2 + \beta2*c3 + \varepsilon$
    - $d4 = \beta0 + \beta1*d2 + \beta2*d3 + \varepsilon$
  - other models are possible

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

Use Simulation to Demonstrate Potential MAPOD Methodologies

- Would like to improve on the simple strategies above, but need more data
- First Approach
  - Simulate many (c_2,c_3,d_2,d_3)'s assuming a joint normal distribution with observed covariance
  - Use independent regressions to estimate c_4 from c_2,c_3; and estimate d_4 from d_2,d_3 (best we can do with 4 operators)
  - Compare, for Q4,
    - Predicted PODs generated from regression estimates
    - PODs generated from original 4 operators
- Second Approach
  - Bootstrap hit/miss data for 36 additional operators
  - Generate POD curves and thus $(c_2,c_3,d_2,d_3,c_4,d_4)$ based on new hit/miss data
  - Use multivariate regression to jointly estimate (c_4,d_4) from (c_2,c_3,d_2,d_3) for 40 total operators
  - Compare, for Q4,
    - Predicted PODs generated from multivariate regression estimates for 4 operators
    - PODs generated from original 4 operators

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

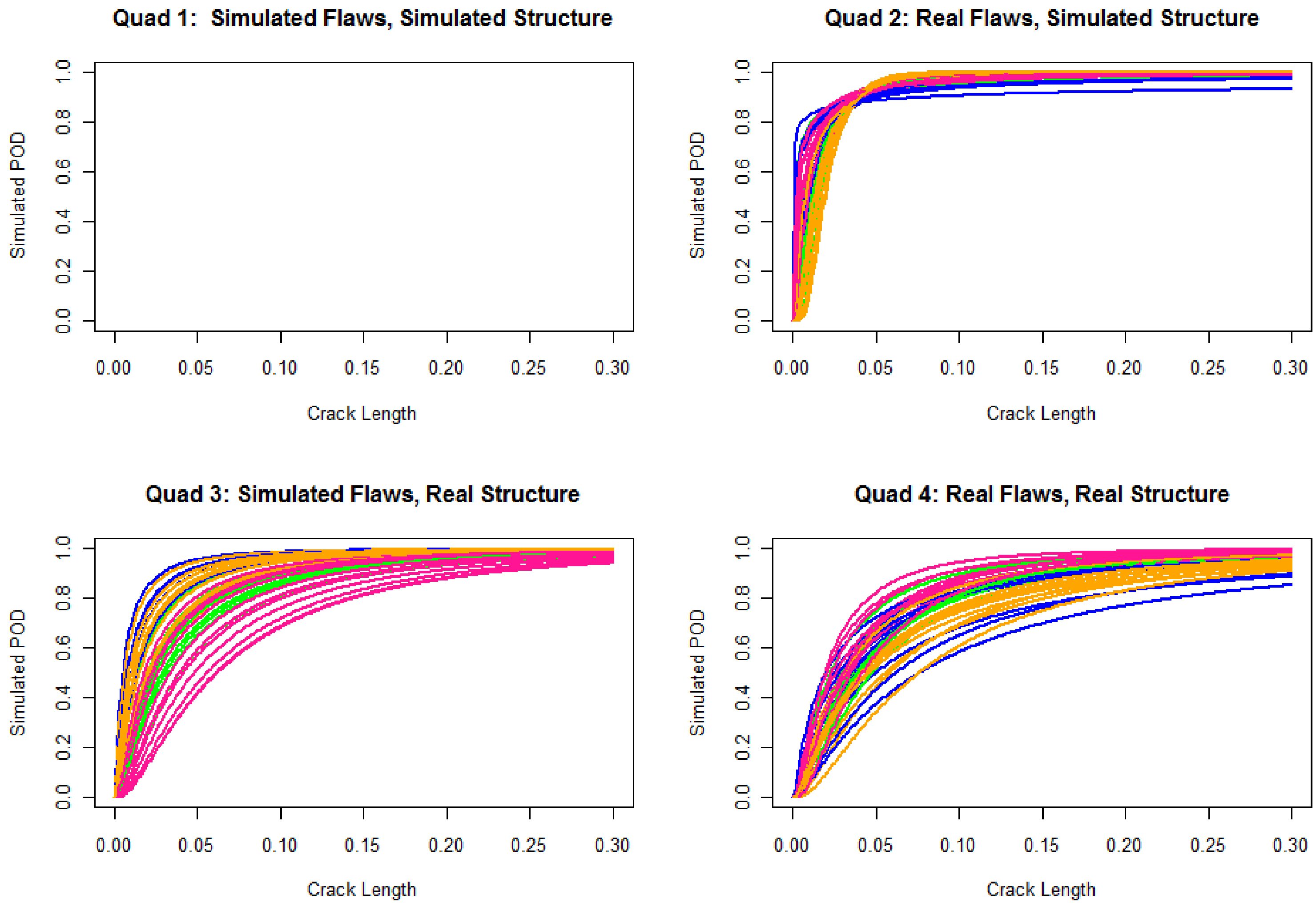
First Simulation Approach:
Generate parameter estimates and predict Q4 c's & d's using independent regressions.

- For each operator and observed (c_2,d_2,c_3 d_3) quadruple, generate 2500 more quadruples (total of 10,000 quadruples) assuming
  - joint normal distribution
  - mean equal to the observed quadruple value
  - covariance equal to the covariance matrix from the four observed operators
- Use the least squares line from the original operators to predict 10,000 values for c_4 based on c_2 and c_3 and 10,000 values for d_4 based on d_2 and d_3
- Look at resulting 10,000 POD curves for Q4 (sample of 40 curves on next slide)

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

First Simulation Approach: A Selection from 10,000 Simulations



This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

$POD_{0.90}$ Estimates
First Simulation Approach

	Quad 2	Quad 3	Quad 4
Median from Observed Operators	0.039	0.088	0.164
Median from Sims (upper 95% CL)	0.039 (0.057)	0.090 (0.184)	0.156 (0.279)

- In Quad 2, the intercept and slope parameter estimates were highly variable for the 4 observed operators (despite the corresponding POD curves being quite similar), causing some simulated Quad 2 POD curves to level off early, or to have negative parameter values

- **Conclusion**: The basic linear regressions used to predict c_4 and d_4 lead to adequate predictions of $POD_{0.90}$ in Quad 4… but perhaps we can do better!

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

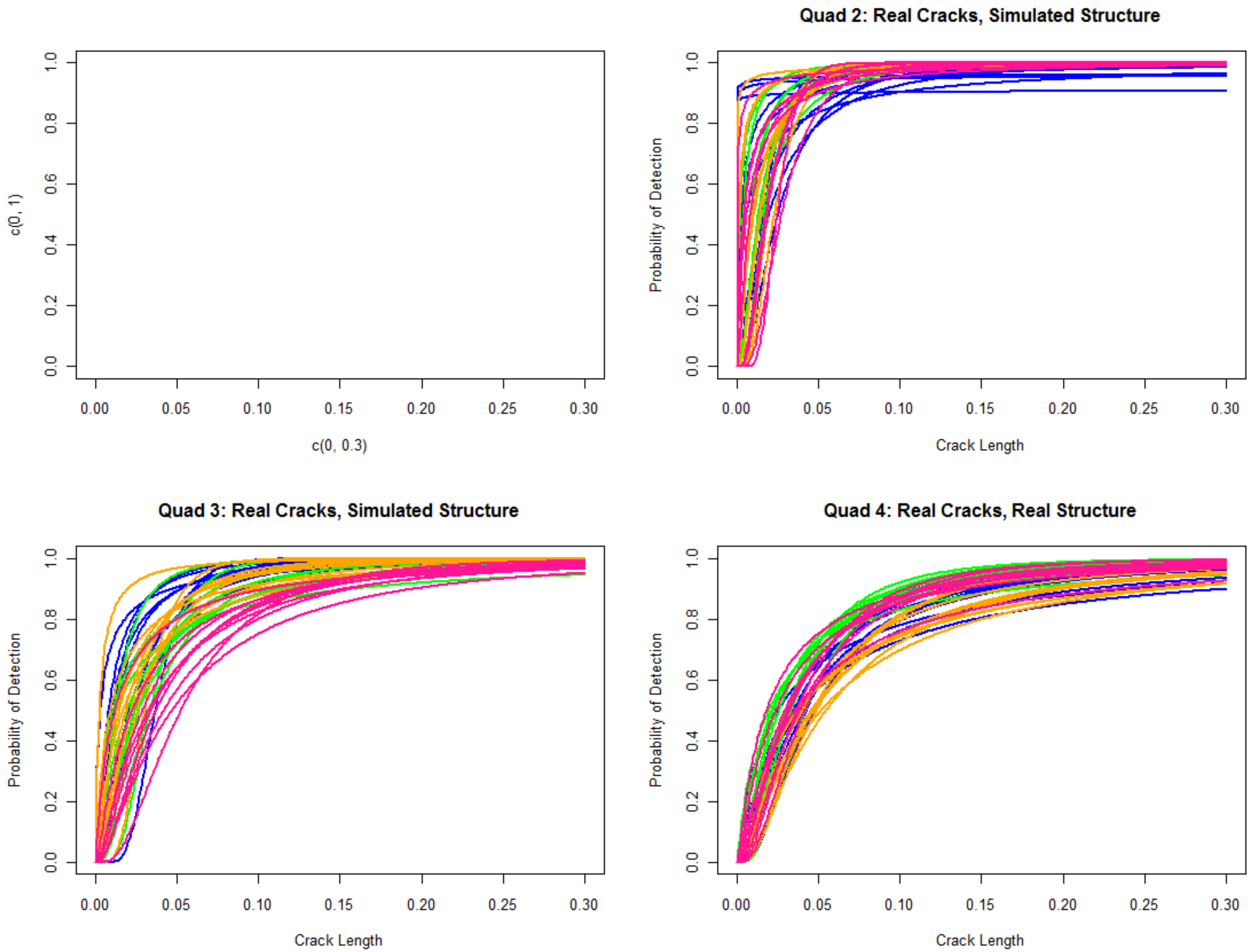
Second Simulation Approach: Bootstrap datasets for more operators and generate each POD.

- Generate POD curves for each operator in Q2, Q3, Q4
- Use the POD curves and random Bernoulli trials to bootstrap hit/miss data for 9 simulated operators based on each of the four observed operators
- Fit POD curves in all quadrants for 40 operators (4 original and 36 bootstrapped), generating $(c_2,d_2,c_3,d_3,c_4,d_4)$ for each
- Use multivariate regression to jointly estimate (c_4,d_4) from (c_2,d_2,c_3,d_3)
- Resulting 40 PODs (spaghetti plots) in Q2, Q3, and Q4, on next slide

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

Bootstrapped POD curves



This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

$POD_{0.90}$ Estimates
Second Simulation Approach

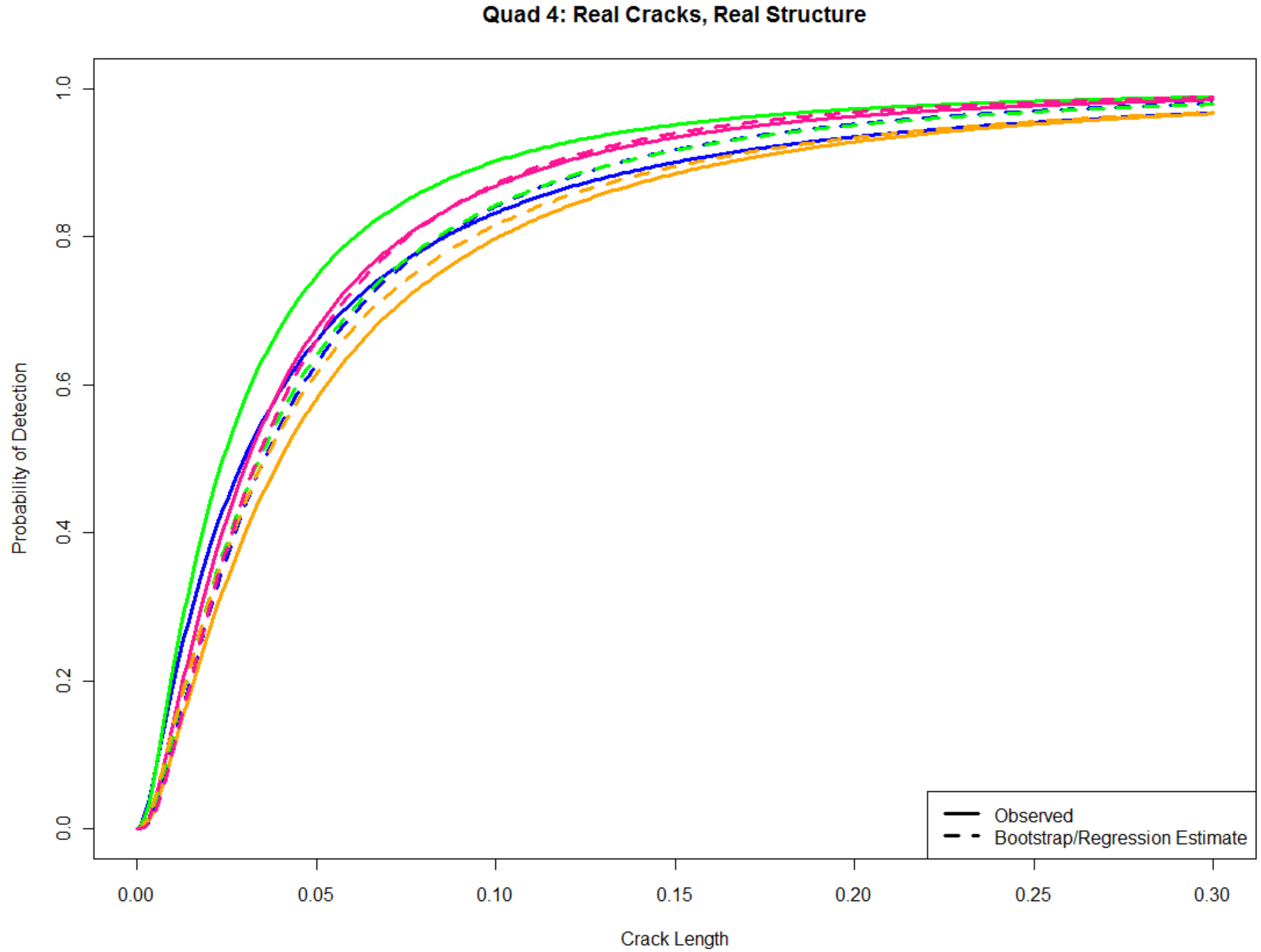
	Quad 2	Quad 3	Quad 4
Median from Observed Operators	0.039	0.088	0.162
Median from Bootstrap Operators (Upper 95% CL)	0.039 (0.071)	0.088 (0.156)	0.155 (0.258)

- Bootstrapped datasets did a good job of replicating operator behavior and variability

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

Comparison of Observed and Predicted Q4 PODs for 4 Operators (Second Approach)



This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

$POD_{0.90}$ Estimates from Original POD and Regression Estimated POD

	Original POD	Regression Estimated POD
Operator 1	0.187	0.134
Operator 2	0.116	0.135
Operator 3	0.200	0.155
Operator 4	0.138	0.116
Average	**0.160**	**0.135**

- Multivariate regression estimates for Q4 Bootstrap Operator PODs tended towards anticonservative estimates of $POD_{0.90}$
  - More work is needed to investigate the results of the multivariate regression and other possible correlation structures
  - *â* vs *a* POD curves may provide additional information which would allow for a better multivariate regression fit

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

Summary

- First simulation strategy offers the potential to generate both $POD_{0.90}$ and an 95% upper confidence bound on $POD_{0.90}$ from limited data
  - Still requires *some* data in Q4
  - Somewhat to our surprise, these independent linear regression models from the first example do quite well

- First step of second strategy, bootstrapping to simulate additional operators, appears to have good potential
  - Also requires *some* data in Q4
  - More work (investigation of correlation structure) is needed to determine if the strategy is viable for developing better transfer functions for Q4

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Conclusions & Future Work

- We did not have complete signal strength data readily available, but may pursue *â* vs *a* MAPOD when they are available
- Q3 POD appears to be strong predictor of Q4 POD for some operators in our particular setting
- Q3 carries information to predict Q4, while Q2 appears to carry very little
  - This is good news since Q3 panels (simulated flaws, real structure) are less expensive to produce than Q2 panels (real flaws, simulated structure)
  - Simulated flaws in real structure (Q3) may be more representative of real flaws in real structure (Q4) than real flaws in simulated structure (Q2) because of the influence of the structure
- We know paint attenuates the UT signal, which we suspect is at least part of the reason for the degeneration of POD curves from Q3 to Q4
  - The degree of attenuation is measurable so quantifying that change, and its influence on the predicted POD, is on the agenda
- Ultimate goal is to extend MAPOD methodology to large area composite airplane inspections

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Extra Slides
Details of Model Fits, Correlation Structure

This work was supported by the FAA, William J. Hughes Technical Center,
Atlantic City International Airport, New Jersey.

Sandia National Laboratories

Simple Model for c4 (as function of c2, c3)

```
Call:
lm(formula = c4 ~ c2 + c3, data = q)

Residuals:
0.04040  0.09013 -0.07623 -0.05430

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.8619     1.3504   5.822    0.108
c2            0.0897     0.0603   1.488    0.377
c3           -1.5834     0.4051  -3.909    0.159

Residual standard error: 0.1361 on 1 degrees of freedom
Multiple R-squared: 0.9416,
Adjusted R-squared: 0.8249

F-statistic: 8.065 on 2 and 1 DF,  p-value: 0.2416
```

back

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Simple Model for d4 (as function of d2, d3)

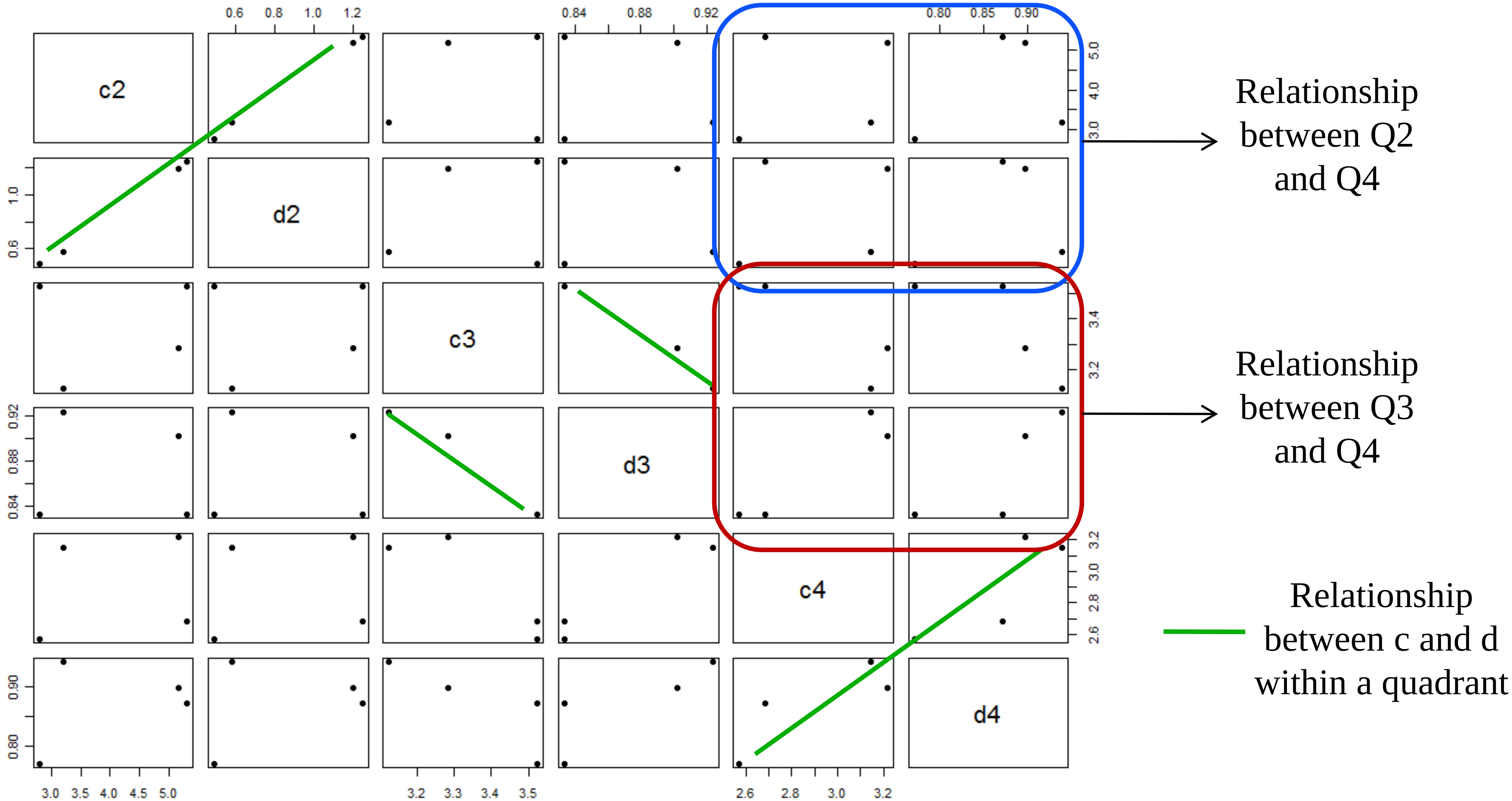
```
Call:
lm(formula = d4 ~ d2 + d3, data = q)

Residuals:
-0.02005 -0.03301  0.02784  0.02522

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.34447    0.59658  -0.577    0.667
d2           0.07047    0.07822   0.901    0.533
d3           1.32049    0.67127   1.967    0.299

Residual standard error: 0.05387 on 1 degrees of freedom
Multiple R-squared: 0.815,
Adjusted R-squared: 0.4451

F-statistic: 2.203 on 2 and 1 DF,  p-value: 0.4301
```

back

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

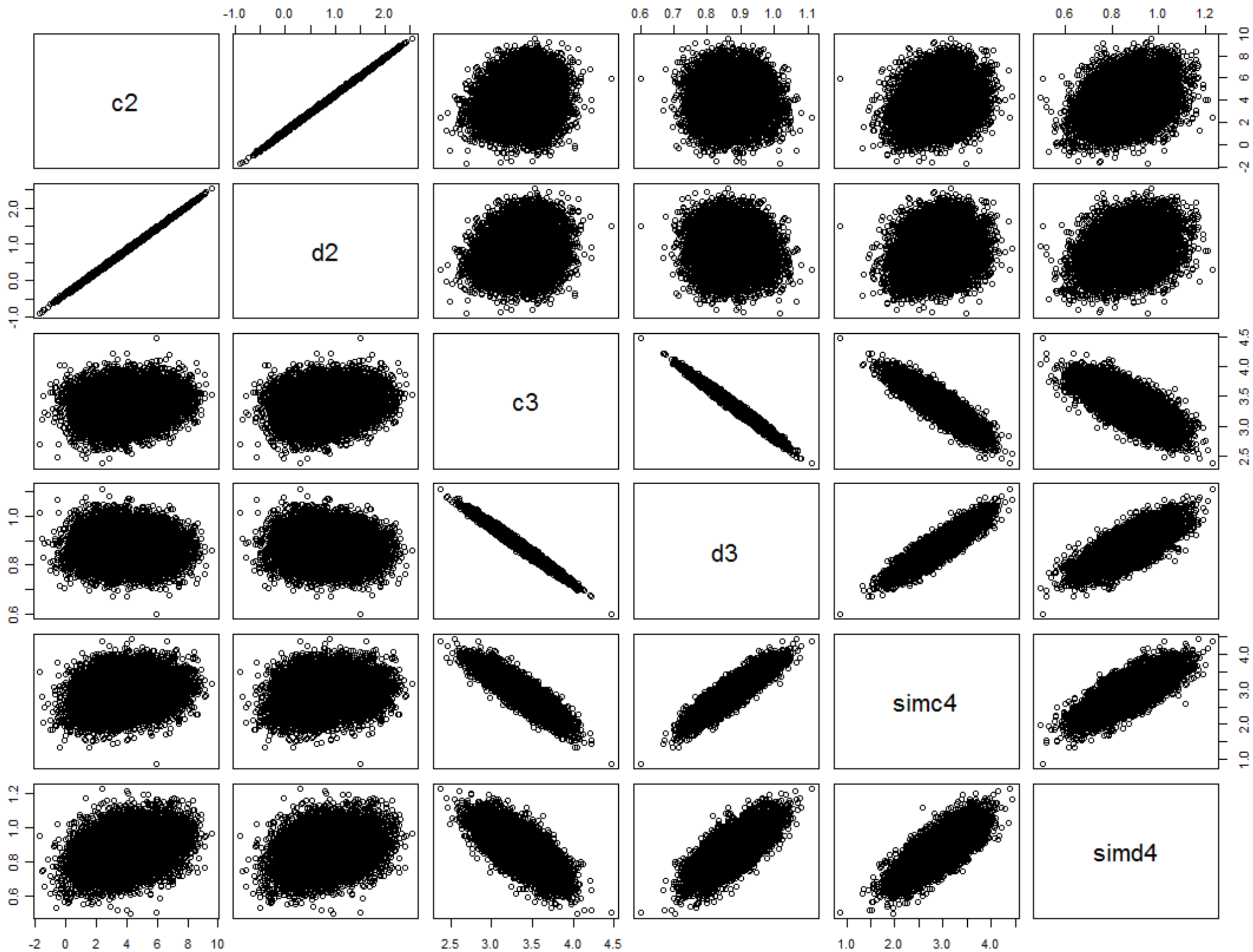
Correlation Structure Between Parameter Estimates



This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

Correlation Structure in First Simulation



This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Model for c4 as function of c2, d2, c3, d3

```
Call:
lm(formula = c4 ~ c2 + d2 + c3 + d3, data = fits)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.01166 -0.46895 -0.07504  0.28905  1.49891 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   3.6953     0.6501   5.684 2.02e-06 ***
c2            0.1758     0.3269   0.538   0.5941    
d2           -0.9903     1.0051  -0.985   0.3313    
c3           -0.4872     0.2456  -1.984   0.0552 .  
d3            1.4738     0.7490   1.968   0.0571 .  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 0.6164 on 35 degrees of freedom
Multiple R-squared: 0.2295,     Adjusted R-squared: 0.1414 
F-statistic: 2.606 on 4 and 35 DF,  p-value: 0.05239
```

back

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Sandia National Laboratories

Model for d4 as function of c2, d2, c3, d3

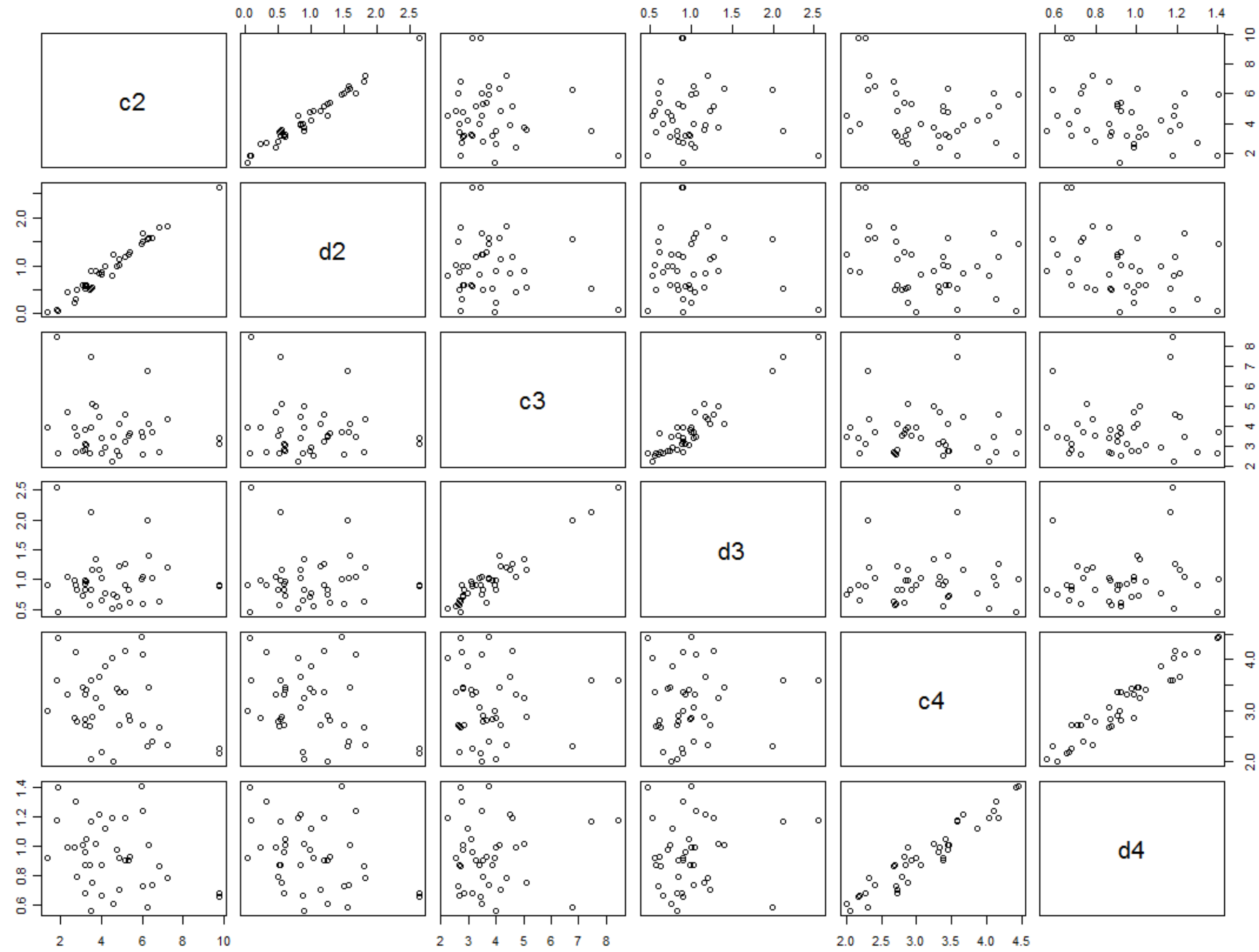
```
Call:
lm(formula = d4 ~ c2 + d2 + c3 + d3, data = fits)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.35653 -0.13859 -0.02136  0.12273  0.52678 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  1.00471    0.22184   4.529  6.6e-05 ***
c2           0.09111    0.11153   0.817    0.420    
d2          -0.42320    0.34298  -1.234    0.225    
c3          -0.11051    0.08380  -1.319    0.196    
d3           0.36019    0.25558   1.409    0.168    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 0.2103 on 35 degrees of freedom
Multiple R-squared: 0.2004,     Adjusted R-squared: 0.109 
F-statistic: 2.193 on 4 and 35 DF,  p-value: 0.09
```

back

This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

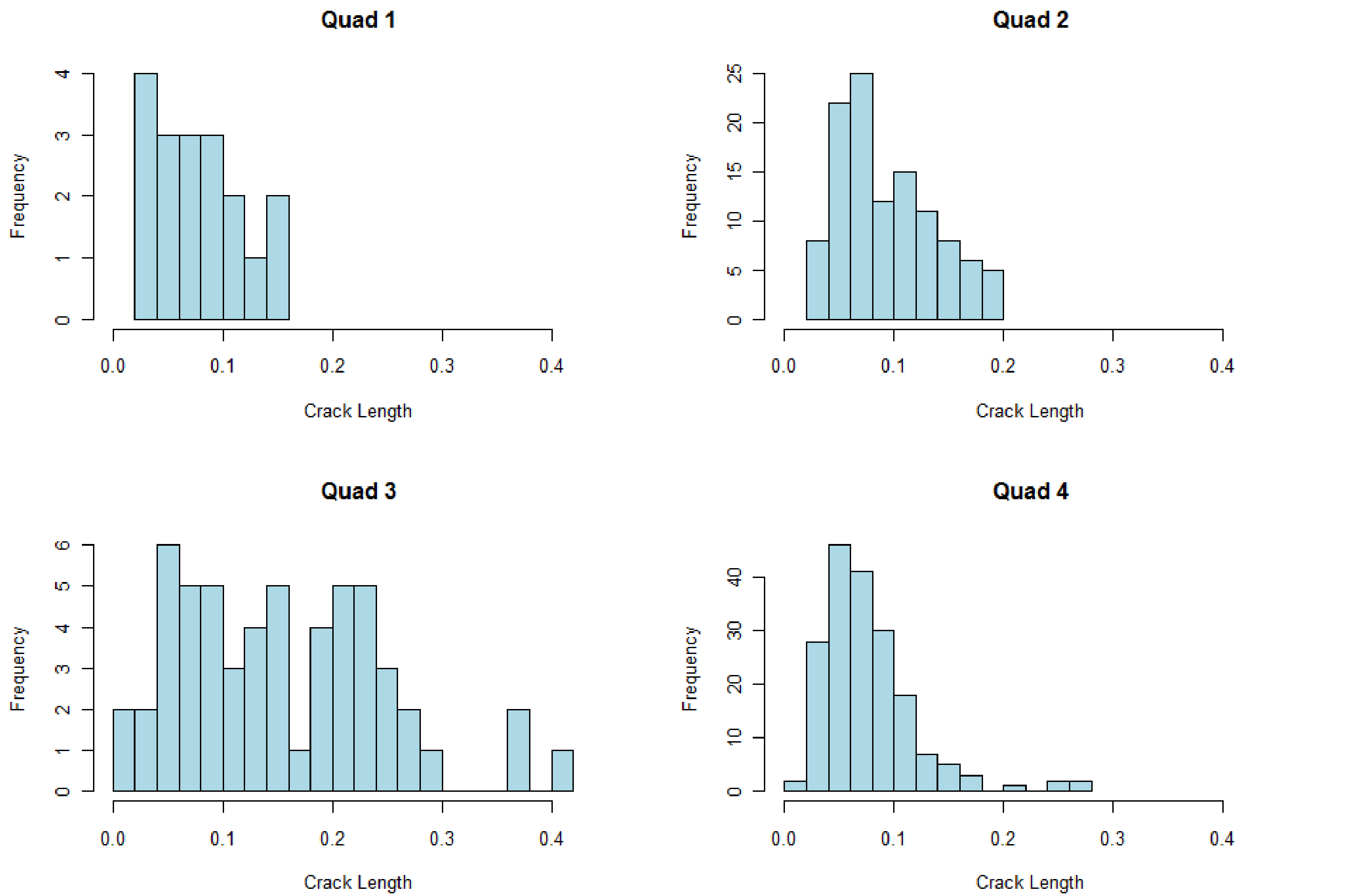
Sandia National Laboratories

Correlation Structure in Second Simulation



This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.

Distributions of Crack Lengths



This work was supported by the FAA, William J. Hughes Technical Center, Atlantic City International Airport, New Jersey.