Evaluating POD in Real Situations and the ‘Delta’ Factor

Martin WALL *
* ESR Technology Ltd. National NDT Centre (NNDTC), 22 East Central, Olympic Way, Milton Park, Abingdon, Oxfordshire, OX14 4SA, UK
martin.wall@esrtechnology.com

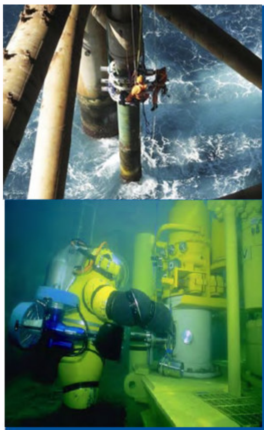
Abstract

The use of Probability of Detection (POD) to characterise the reliability of non-destructive testing is well established in safety critical and other industries including aerospace, nuclear and offshore. There are established methodologies for determining reliability from POD trials and increasingly by modelling. POD provides an important input to integrity and safety assessments.

For a variety of reasons POD in real situations may differ to that established in trials; for example due to human or environmental factors(HF) or application related (AP) factors such as defect tightness or characteristics, coatings, thresholds and sensitivities used, or component geometry or thickness. This difference between the POD in trials or modelling and real situations is referred to as the “delta-factor”. What would be the effect on POD if a crack was a facetted stress corrosion crack rather than fatigue, or if the crack was under compression? What would be the effect of combining NDT techniques or repeat inspection?

This paper shows a number of approaches used by NNDTC to evaluate the “delta” of reliability measurements to the field performance of NDT including: use of human reliability data, such as that arising from PISCIII; empirical measurements; the use of NDE simulators; derivation of the delta-factor using POD models; and use of POD trial data in real situations.

POD models developed by NNDTC are used to show how parametric differences can be assessed. An example is given of the use of “spot-the-ball” simulators intrinsic within these models and an NDE simulator developed for railway axle inspection can be used to evaluate operator effects. Methodologies for correcting for human factors using data from PISC III and comparison of automated and manual UT inspection are also described.



ESR Technology
A Hyder Consulting group company

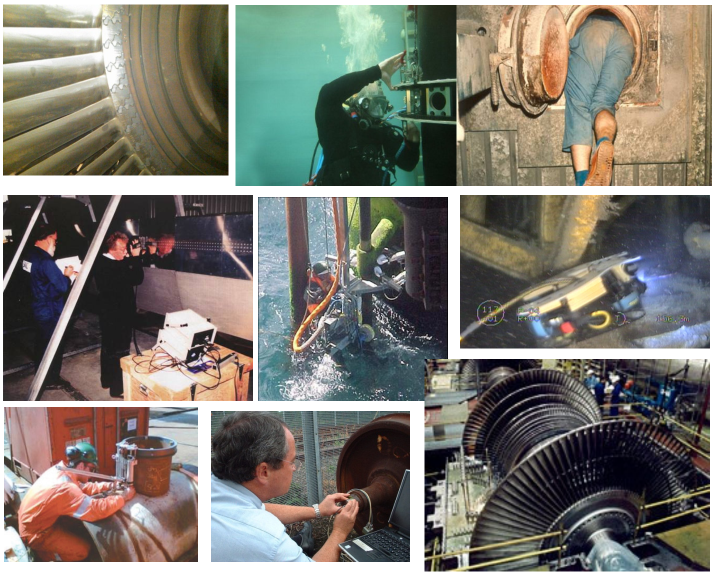
Evaluating POD in Real Situations and the 'Delta' Factor (Paper 33)

Dr Martin Wall
ESR Technology, National NDT Centre NNDTC

5th European-American Workshop on Reliability of NDE
October 7 – 10, 2013 in Berlin, Germany

'Field' v POD trial? May not have POD trial data for your situation

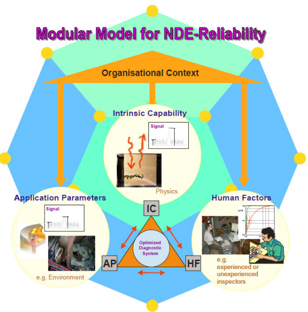
ESR Technology
A Hyder Consulting group company



- Environment
- Time pressures
- Access conditions
- Distractions
- Surface condition
- Component temperature
- Interpretative skills required (complexity)
- Operator skill level
- Hand – to – eye co-ordination
- Familiarity
- Diligence (supervision)
- Scrutiny
- Repetitiveness, including coverage of large areas

Slide 2

Outline



Christina MUELLER et al, 4th EAW,www.ndt.net/index.php?id=8314

Reliability = f(IC,AP,HF)

- Why defects are missed
- Real v Trial The 'delta' factor
- Human and Environmental Factors HF
- Approaches to evaluate delta (and de-rate POD)
- Examples
  - Use of Human Reliability data
  - Use of existing POD data
  - Expert Panel Approach
  - Monte-Carlo Analysis
  - Use of POD models
  - Bayesian updating
  - Use of Simulators
- Conclusions

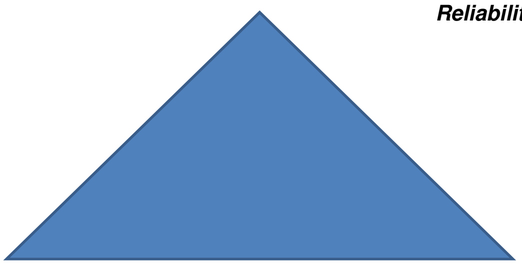
Slide 3

Why defects are missed?

Intrinsic Capability IC

Defect (Characteristics, type, tightness, orientation etc.) AP

Reliability = f(IC,AP,HF)



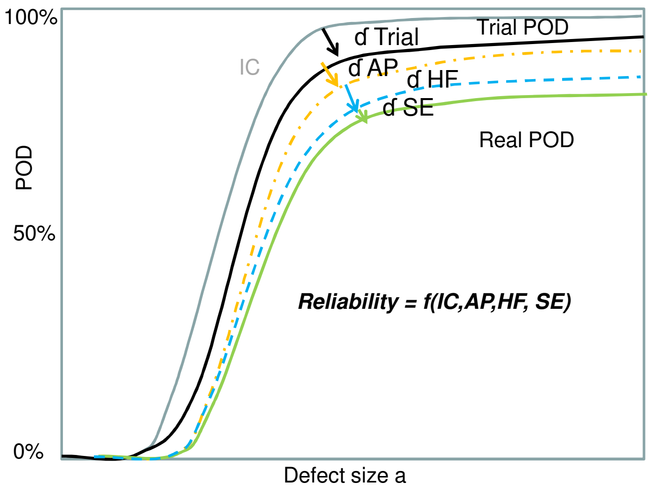
Component, coverage and Geometry AP

Human and Environmental Factors HF

Slide 4

The 'delta' factor

ESR Technology



Several ɗelta factors

Intrinsic Capability IC

ɗ Trial

ɗ Application AP

ɗ Human HF

ɗ Socio-Economic SE

European-American POD model

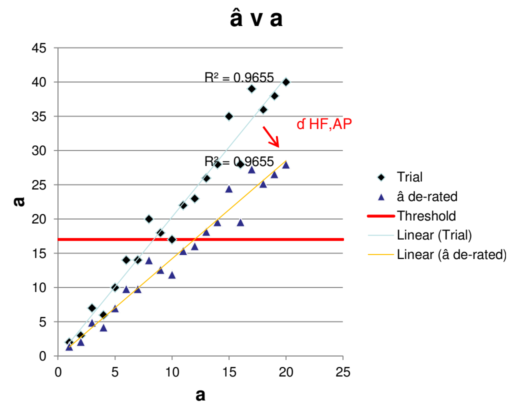
R = f(IC,AP,HF)

How significant 'ɗelta factors' are depends on closeness of trial data to real situation

Slide 5

'delta' factors- Transfer Function to raw data or POD curve?

ESR Technology



1. Apply transfer function to raw POD data: hit/miss, signal/noise distributions or â v a

or

2. Transfer function on POD curve

Depends on

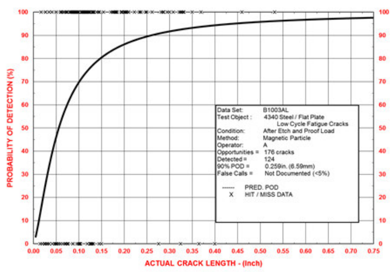
The application

The NDE method

The basis of the transfer function

The data type you have (*you may only have POD curves*)

The use of the POD

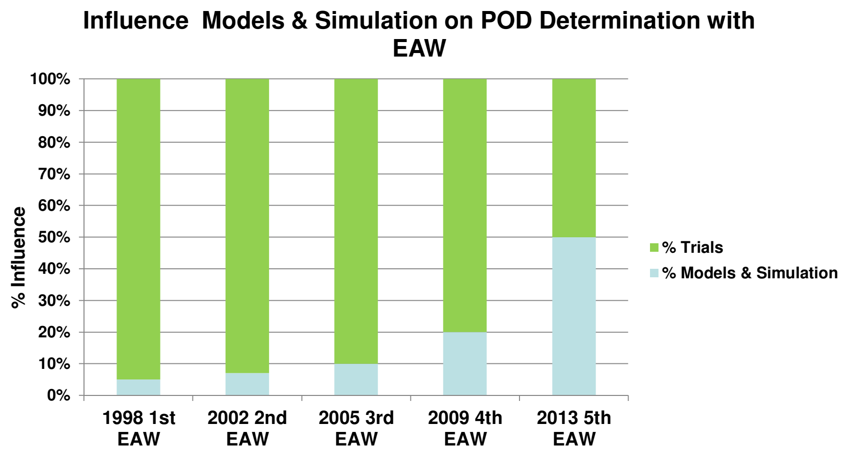


Most accurate to apply transfer function to raw experimental data: e.g. hit/miss, signal/noise distributions or â v a ; or to raw predictions if model assisted POD.

Slide 6

Influence of POD Models & Simulation with EAW

ESR Technology



Slide 7

Development of POD Models and Simulations with EAW

ESR Technology

Year	% Models & Simulation	% Trials	Label
1998 1st EAW	5%	95%	AEA NNDTC (UK), IOWA CNDE (US)
2002 2nd EAW	7%	93%	AEA NNDTC (UK), BAM (DE), IOWA CNDE (US)
2005 3rd EAW	10%	90%	Iowa (US), AEA NNDTC (UK), CEA (FR), BAM (DE)
2009 4th EAW	20%	80%	BAM (DE), ESR NNDTC (UK), HOIS (UK), CEA (FR), TNO (NE), Others
2013 5th EAW	50%	50%	BAM (DE), CEA (FR),ESR NNDTC (UK), HOIS (UK), IZFP(DE), TNO (NE), Others

- NNDTC Models Aerospace -> Nuclear -> Manufacturing ->Space (ESA) -> Oil &Gas Modular Experimental or Model Input
- UT -> Radiography -> ET -> Technique Independent- UT pit detection

Slide 8

Allowing for the 'delta' factor in POD estimates

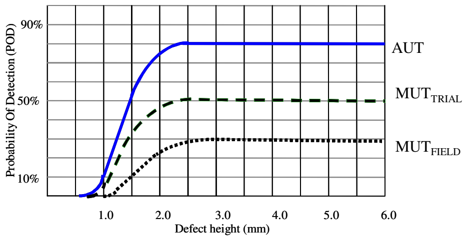
Models can examine application parameters -→ '*Transfer functions*'. Human factors difficult to model. Use *empirical data* or data from simulations or models:

- Human reliability trial data (PISC III, PANI 1-3 etc.)
- Neural Network (Simulate inspector -Automated defect Detection, Windsor, Wall 1991)
- Expert panel approach
- POD models
- Bayesian updating
- Simulators
- Monte-Carlo methods
- Trial data

Similar approaches may be applied to model or trial data (trial data may already include some human factors (HF)

Slide 9

Manual NDT inspection of plate welds NIL data



Manual ultrasonics POD values typically much lower than would be expected (circa ~50%)

Usually involves *hand-eye coordination*

Automated methods typically found to give much better reliability (NIL Trials)

Worse reliability can be found in using manual methods in field situations ('the field factor')

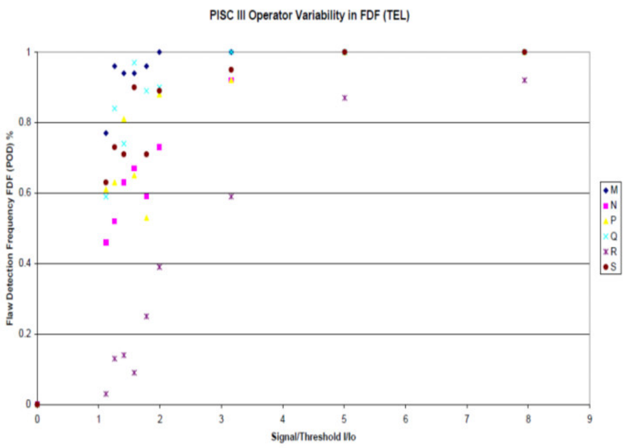
Detection threshold is unaffected

More factors are at play in the field compared to laboratory trials, therefore *plateau must be affected* (*Lilley et al ECNDT ndt.net)*

Slide 10

Use of Human reliability trial data (PISC III Action & PANI 1-3)

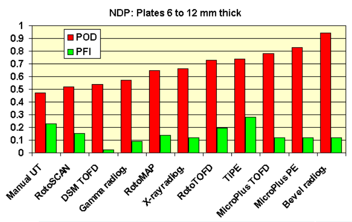
- Errors incurred whilst plotting and reporting data.
- Operators tired during a shift and became de-motivated.
- Tiredness due to length of shift and repetitiveness of the work
- Feelings of isolation, temperature and humidity also affected operators.
- Proficiency in mental arithmetic an issue.

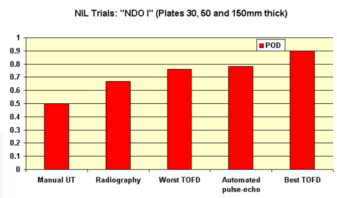


Slide 11

NIL Trial data Automated v Manual Ultrasonics (UT)

ESR Technology





Commercial equipment Automated methods typically found to give much better reliability (~80% in NIL Trials) compared to manual methods (typically~50%)
Used to distinguish application parameters from human factors (*Lilley, ECNDT*)

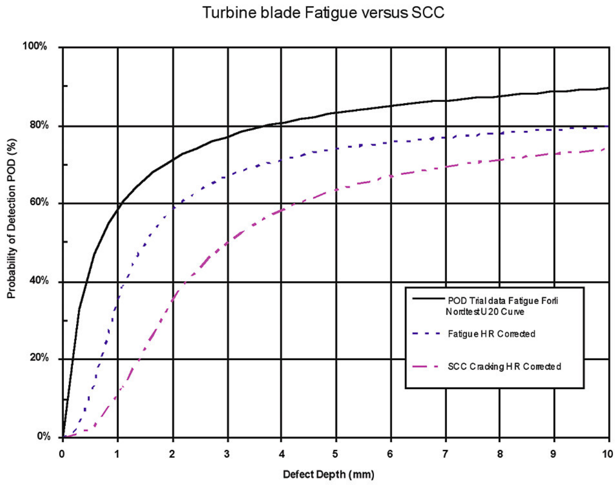
Technique	6+8mm	8+10mm	10+12mm	15mm	Average
µ+ TOFD	80	79	75	96	82%
µ+ PE	82	84	82	86	84%
Manual PE	46	46	48	69	52%
X-RT	69	63	66	67	66%
γ-RT	63	53	54	71	60%
Bevel RT	94	94	96	95	95%

Slide 12

UT inspection of Steam Turbine Blade Root Operator variation and defect type SCC v Fatigue

PISC III Operator data used for Human reliability correction. Empirical rules.

Experimental UT measurements on Steam Turbine blade SCC and fatigue give 'transfer function' for defect type.



Human factors in POD modelling and use of trial data M Wall, S Burch and J Lilley Insight Vol 51 No 10 October 2009

Slide 13

Expert Panel approach

- Common approach in use of POD data in technical justifications or inspection planning
- Benefits from a structured approach
- Identify relevant POD trial data
- Expert panel review differences between site situation and trial
- How will these affect the slope and plateau of the POD data (Spreadsheet assessment)
- Can be done qualitatively or made more quantitative based on experimental data.
- Benefits from sensitivity studies
- *Lilley et all, ECNDT*

Slide 14

Expert Panel Approach

Steam turbine blade ultrasonic UT inspection

ESR Technology
A Hyder Consulting group company

- Reflectivity (SCC v fatigue)
- Surface geometry
- Environment
- Pressure
- Access
- Interpretation
- Operator skill level
- Hand to eye coordination
- Familiarity
- Diligence
- Repetition
- Third party observations
- Supervision
- Permanent record of data
- Scrutinised

Detectability	Description	Original POD	Turbine blade	POD Slope	POD Level
UT-Reflectivity	Are the defects as detectable as those used to generate the POD curves? e.g. defect orientation, roughness.	Clear cut defects giving specular or corner trap effects.	UT – sensitive to large cracks, but not small	-2	+2
Surface condition	Are the surface conditions as comparable?	Good cond', ground.	OK	0	0
Surface geometry	Does the component geometry have an influence on the detectability? e.g. curvature, shadow effects.	No geometry effects.	Very strong detrimental impact.	0	-3
Interference	Is masking possible? e.g. delayed geometric echoes, satellite defects	None present.	Unlikely	0	0
Equipment	How does this compare to the suitability of the equipment used during the trials?	Top quality equipment.	Assume OK	0	0
Reliability					
Environment	How does the working environment compare?	Laboratory, except TEL	Hot, noisy site conditions	-1	-1
Pressure	Are the work pressures different?	Minimal.	Present, but not to XS	-1	-1
Access	Do access condition vary from those which apply to the study environment?	Good access.	Awkward – large impact.	-2	-3
Interpretation	Will interpretative skills differ from those used during the POD trials?	Simplistic configurations with clear cut defects.	Very difficult – significant impact	-3	-3
Operator Skill Level	How do operator skills differ from those participating in the trial? Will they always be available?	Top level operators used.	Same.	0	0
Hand to eye co-ordination	How does this differ from the trial conditions?	Easy access.	Very difficult – major impact.	-3	-3
Familiarity	How familiar are the operators with the application and the procedures compared to the trials?	Operators generally familiar with procedures.	V good	+2	+2
Diligence	How could this differ from that of the operators participating in the trials?	The operators were under scrutiny.	Operators worked in isolation	-1	-1
Repetition	What effect will this have compared to the trials?	Not specifically a problem	Highly repetitive	-2	-2

Slide 15

Models and Simulators

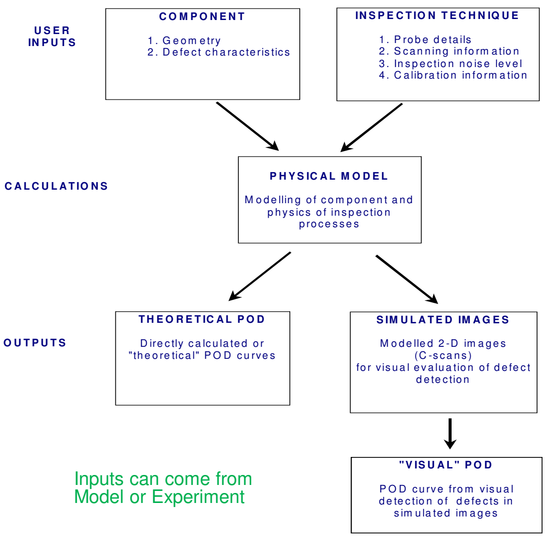
ESR Technology
A Hyder Consulting group company

- Prior history of using simulators to understand human and environmental factors (PISC III)
- Models provide insight on application parameters (AP)
- Simulators provide insight on human errors (HF) in interpretation of data
- Form a basis for deriving transfer functions
- Bayesian update methods can be used tolink model to experiment
- Particularly useful for image based NDE methods
- Simulated POD trials 'Spot-the-Ball'.

Slide 16

NNDTC Modular POD models

ESR Technology



Developed since early 1990s
Nuclear, manufacturing, space (ESA), oil & gas

Inputs:

- Details of component, defects and inspection technique

Physical modelling of inspection process

- Defects in random locations in component

Outputs:

- Directly calculated "theoretical" POD curves
- Simulated NDT data

User evaluation of simulated images ("Spot the defect")

- Alternative measures of POD, PFI

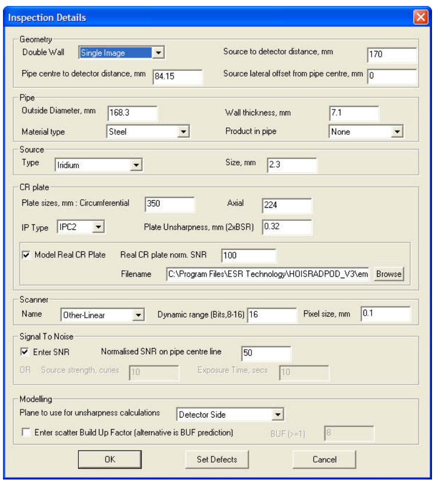
Modular approach

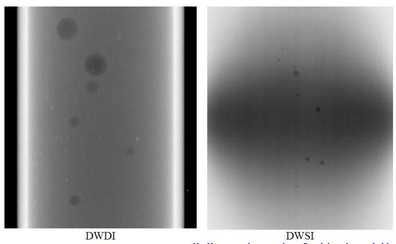
- Can be extended to other techniques

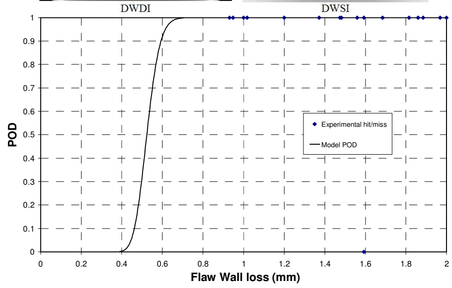
Slide 17

Model Example
HOIS RADPOD Model

ESR Technology





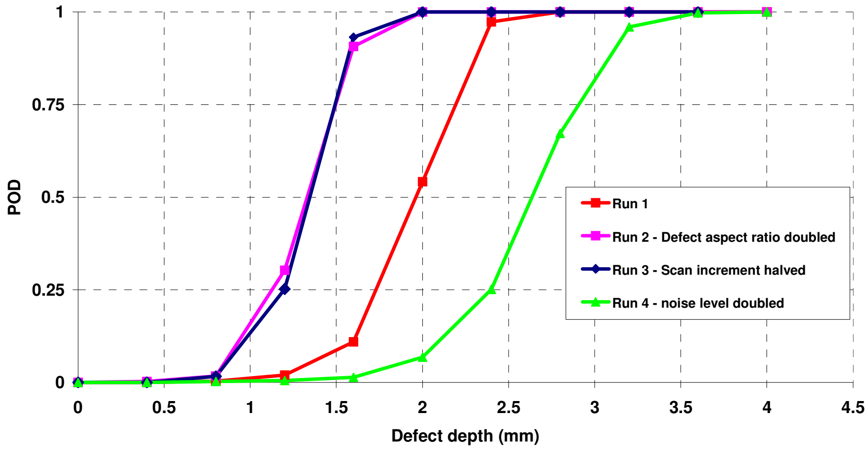


Slide 18

Results of POD modelling for four cases

ESR Technology

Variations in all parameters modelled have a significant effect on POD
Ultrasonic C-scan - Corrosion pit detection



Slide 19

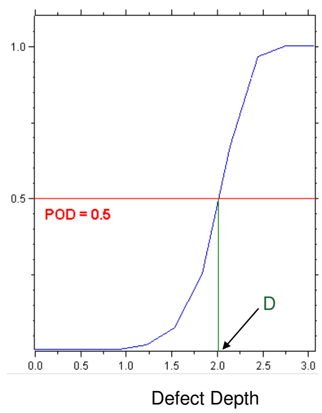
More extensive sensitivity study

ESR Technology

Many runs of POD model to assess effects on POD of variations in:

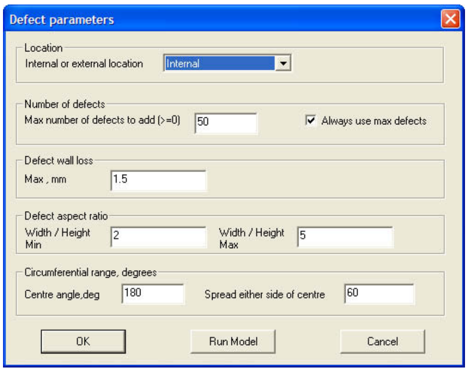
- Noise level on corrosion maps (in mm)
- Scan increment (mm)
- Defect aspect ratio (dimensionless)
- Defect depth (mm)
- Probe beamwidth at backwall (mm)

Analysis of results obtained in terms of the defect depth, D, corresponding to a particular POD value (e.g. 0.5)



Slide 20

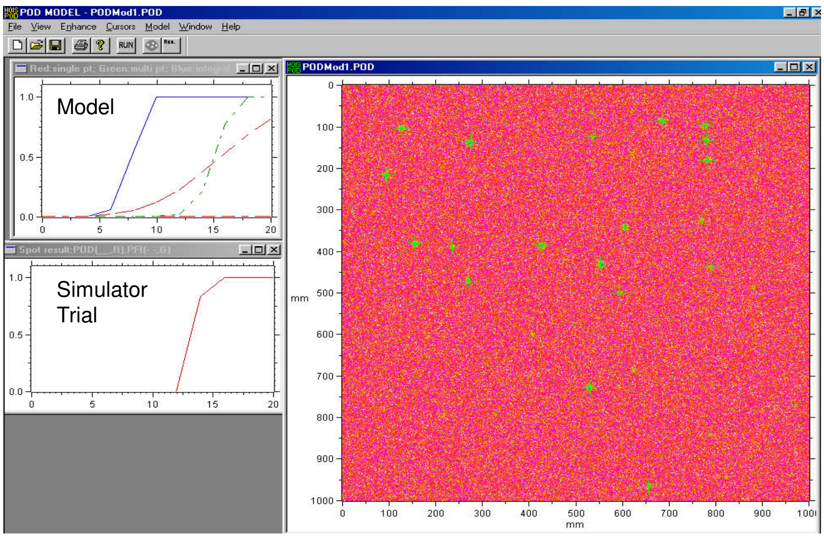
'Spot-the-Defect' Simulators



- Model produces simulated data (examples ESA POD models, HOIS POD Generators UT, CR radiography))
- Can include random numbers, size and types of defects in the image
- Images presented to the inspector in real time as NDT data
- Inspector(s) assess data and mark where they can see a defect
- Software analyses the results to give POD and false calls (POF)
- Useful training aid, complementary to trials, helps understand human errors in interpretation of data
  - *Burch S F and Stow B A BINDT 2005 and Insight*
  - *Burch S F and Stow B A Insight Vol 50 No 9 Sep 2008 p485-489*

Slide 21

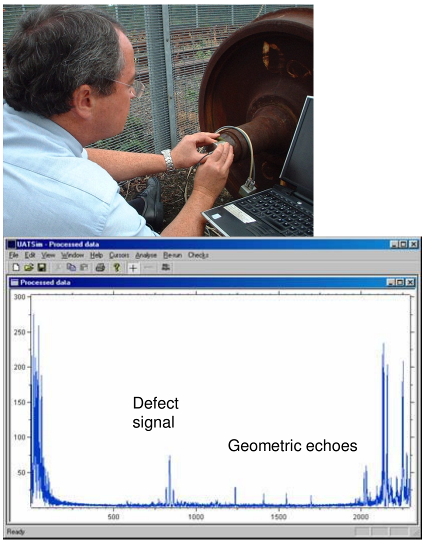
HOIS POD generator: example (fine scan increment) 'Spot the ball'



Slide 22

Rail Axle UT Simulator
Effect of geometric echoes on detection

ESR Technology



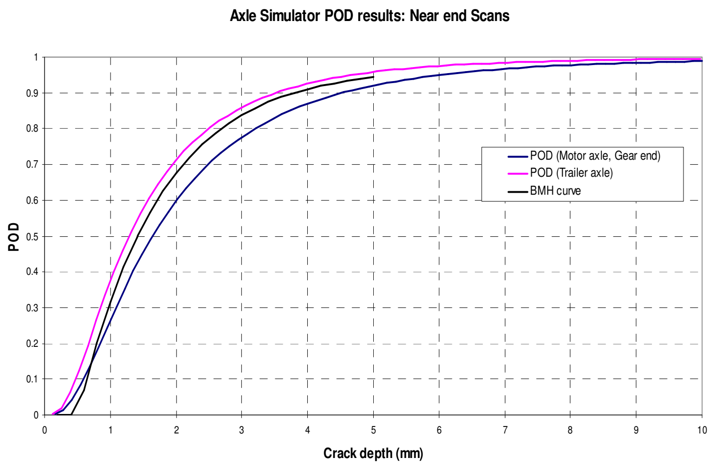
- Record Digital A-Scans from axle
  - Simulator uses *actual A-scan data* (Signal and noise)
- Record signals from reference defects
- Correction for crack signals
- Model generates series of A-Scans with defects in random locations in axle.
- Run the UAT simulator many times, to produce results for at least 200 simulated defects for each case.
- Analyse by '*Spot-the Ball*' hit-miss method . Use MIL Standard Software to produce POD curves for specific axle type.
- Full windows based software

To understand the effect of geometric echoes and crack detectability for complex geometry componets

Slide 23

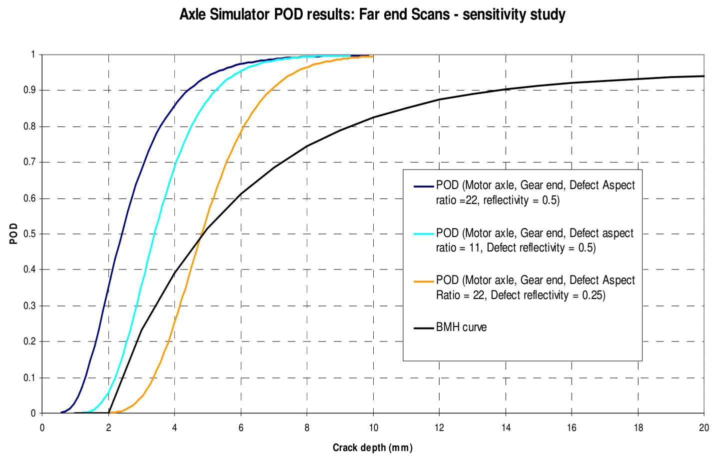
POD Simulator Results
Near-End Scan

ESR Technology



Slide 24

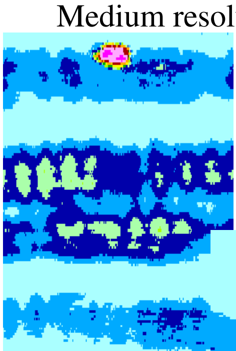
POD Simulator – Far-End Sensitivity Studies



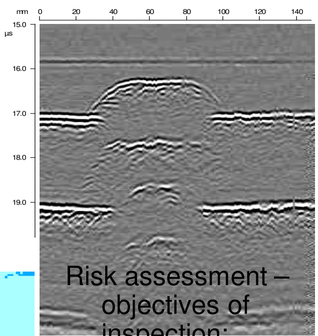
Slide 25

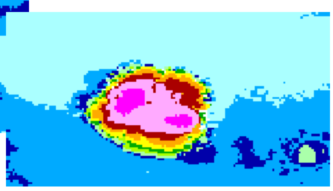
Improvement – Automation & Technical

Medium resolution



TOFD





High resolution

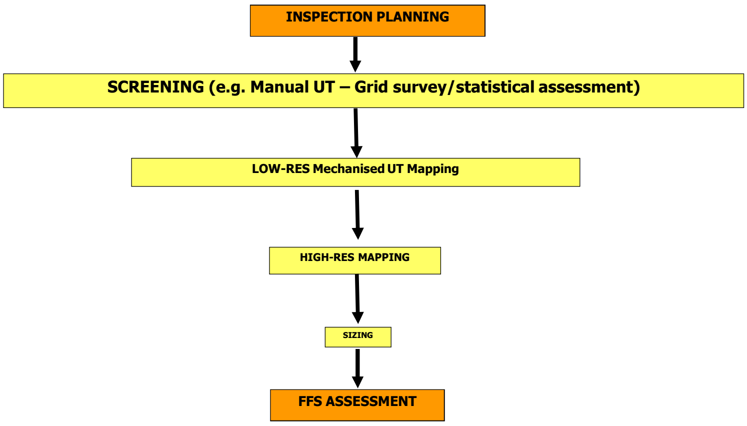
Risk assessment – objectives of inspection:

- Technical – supplementary techniques?
- Human factors – controls, automation?

Slide 26

Improvement Reduce δ–Factors
Scope of inspection

ESR Technology



Slide 27

Conclusions

ESR Technology

- Methods are evolving to take understand and take account of human and application factors in estimating POD
- These involve the use of human reliability data (PISC III, PANI), simulation models and expert judgement
- It is not currently possible to model human reliability. Therefore it is necessary to use empirical data or data from simulations or models.
- This understanding is important to improve the reliability of inspection, since POD is often adversely affected by *field issues*
- Improving NDT in the context of risk assessment requires a particular approach:
  - *Objectives of inspection* – What do we need to find?
  - *Scope* – Where do we need to look, and how?
  - *Sensitivity* – What techniques?
  - *Reliability* – What level of control, automation and verification is required?

Slide 28

Questions

?

Slide 29