ELEMENTS OF DETECTION THEORY APPLIED TO NDE

Derek Sturges
GE Aircraft Engines
Engine Support Operation
Cincinnati, Ohio

Tutorial Workshop on Probability of Detection
Spring 1993 ASNT Conference
OUTLINE

0 DEFINING "DETECTION"
   - Practical aspects of how we think about detection
   - Differences between detection for specific NDE techniques
   - Effects of automation on detection
   - Effects of imaging on detection

0 DETECTION IN A SIGNAL-TO-NOISE CONTEXT
   - A basis for considering detection for all NDE processes?
   - Selected topics in Detection Theory
   - Applying Detection Theory to NDE
DEFINING "DETECTION"

0 NDE "DETECTION" IS RARELY DISCUSSED EXPLICITLY
   - We probably understand it - but we don't define it

0 DICTIONARY DEFINITIONS (Webster's New World Dictionary)
   - Detect: - to catch or discover, as in a misdeed
     - to discover or manage to perceive
       (something hidden or not easily noticed)
   - Detection: - a finding out (said especially of what tends to elude notice)

0 A RADAR DEFINITION (D.K. Barton, Radar System Analysis)
   - Target Detection: the process by which the presence of a sought-after object, or target, is sensed in the presence of competing indications which arise from background radiation, undesired echoes, or noise generated in the receiver
DEFINING "DETECTION"

WE NEED TO DEFINE "DETECTION" BEFORE WE MEASURE "POD"
- Standard NDE handbooks offer little help
- The Radar definition is close to what we need

LET'S REVIEW SOME NDE TECHNIQUES
- Clarify our implicit definition of detection
- Assess whether the Radar definition is suitable for NDE use
INSPECT, TEST, OR EVALUATE?

DISTINGUISH BETWEEN NDT, NDI, AND NDE?
- NDT, NDI, and NDE are often used interchangeably
- Nondestructive Evaluation can be used as a more comprehensive term
  - NDE comprises all NDT and NDI activities
  - NDE comprises detection, location, and characterization

DETECTION AND CHARACTERIZATION ARE DISTINCT PROCESSES
- In what follows, POD will refer to detection only
- Caution: this distinction is not always made
  - Example: "POD" is sometimes used to express a measurement of performance in detecting and classifying (e.g. sizing) defects
  - Better termed "Joint Probability of True Positive" (Joint P\{TP\})
    or "Probability Of Detection and Correct Interpretation" (PODCI)
DEFECTION DETECTION OR DEFECT-FREE MATERIAL?

0 SEPARATE BUT RELATED ISSUES
  - High POD alone does not guarantee defect-free material

0 DEFECTS SUCCESSFULLY DETECTED
  - High POD suffices for selecting NDE techniques
  - High POD is a necessary but insufficient condition for high product life

0 DEFECTS ESCAPING DETECTION
  - High product life requires a low probability that inspected material still contains any defects (compare Avioli's definition of Reliability)
  - Requires high POD and low probability of occurrence of defects
  - High POD may be less important if there is a very low probability of there being a defect in the uninspected material
MAGNETIC PARTICLE AND PENETRANT INSPECTIONS

- SENSOR (the eye)
  - Responds to indication color contrast (dye) or brightness (fluorescent)
  - Responds to indication shape and size (principally length)

- DISCRIMINATOR (the brain)
  - Records high-contrast and/or large indications - the "signal"
  - Classifies these as "flaws" or "defects" (i.e. above threshold)
  - Ignores low-contrast and/or small indications - the "noise"
  - Classifies these as "irrelevant" or "background" (i.e. below threshold)

- THRESHOLD (set by Specifications, Drawing Notes, etc.)
  - Typically defines maximum acceptable indication length
RADIOGRAPHIC INSPECTION

- SENSOR (the eye)
  - Responds to film contrast and density
  - Responds to indication shape and size

- DISCRIMINATOR (the brain)
  - Records high-contrast and/or large indications - the "signal"
    - Classifies these as "flaws" or "defects" (i.e. above threshold)
  - Ignores low-contrast and/or small indications - the "noise"
    - Classifies these as "irrelevant" or "background" (i.e. below threshold)

- THRESHOLD (set by Specifications, Drawing Notes, etc.)
  - Typically defines maximum acceptable indication diameter (or length)
EDDY-CURRENT OR ULTRASONIC A-SCAN INSPECTION

- SENSOR (electronic instrumentation)
  - Responds to indication amplitude
  - Responds to indication depth

- DISCRIMINATOR (an electronic threshold)
  - Records large indications - the "signal"
    - Classifies these as "flaws" or "defects" (i.e. above threshold)
  - Ignores small indications - the "noise"
    - Classifies these as "irrelevant" or "background" (i.e. below threshold)

- THRESHOLD (set by Specifications, Drawing Notes, etc.)
  - Typically defines a maximum acceptable indication amplitude
    - Threshold may be related to indication phase (ET) or depth (UT)
EFFECTS OF AUTOMATION

- **SENSORS**
  - The eye is replaced by electronic or electro-optical devices
    - Example: machine vision for automated penetrant systems

- **DISCRIMINATORS**
  - The brain is replaced by an automatic electronic device
    - Example: ultrasonic indications signalled by electronic alarm
  - The discrimination may be more quantitative, and computer-based
    - Example: penetrant indications may involve measured brightness and size

- **THRESHOLDING**
  - Thresholds may be defined more precisely, and may be more complex
    - Example: "detection" of an indication may be determined by the value of
      a function of several measured variables
EFFECTS OF IMAGING

0 SPATIAL CORRELATION

- Imaging presents data from successive interrogations more efficiently
  - Takes advantage of spatial correlation of information

- Facilitates use of automatic pattern-recognition and decision-making
  - Example: a computer can decide if X-ray CT wall-thickness is acceptable

0 IMAGING DOES NOT CHANGE THE BASIC DETECTION PROCESS

- POD is not necessarily improved by imaging
  - Example: an A-scan strip-chart contains the same data as a C-scan image

- POD may be improved as a result of use a lower discrimination threshold
  - In effect, the threshold may be lowered on a local basis
DETECTION AS A SIGNAL-TO-NOISE PROCESS

- **NDE TECHNIQUES ARE THRESHOLDED S:N PROCESSES**
  - Easiest to see this for "electronic" techniques (e.g. ET, UT) and for automated inspection techniques (e.g. computed tomography)
  - Basically true for other techniques, too (e.g. MT, PT, RT)

- **"THRESHOLD"**
  - Represents a level of discrimination between "signal" and "noise"

- **"SIGNAL"**
  - Represents the response from the defects - the sought-after target

- **"NOISE"**
  - Competing signals from sources other than defects
DETECTION AS A SIGNAL-TO-NOISE PROCESS

Elements of a receiving system

Signals with additive noise

SONAR DETECTION SYSTEM

- Many similarities to ultrasonic NDE
- Fundamental concepts are common to most NDE processes
SOURCES OF NOISE

NOISE FROM THE MATERIAL UNDERGOING INSPECTION

- Generic examples:
  - Surface roughness
  - Edge effects

- Process-specific examples:
  - Local changes in conductivity (MT, ET)
  - Indications from surface pores in castings (PT)
  - Grain-boundary reflections (UT)
  - Scattering (RT)

NOISE FROM THE INSPECTION PROCESS ITSELF

- Examples:
  - Electronic noise (ET, UT)
  - Grain structure in film (RT)
VARIATION IN THE DETECTION PROCESS

o DEFECTS
  - Each defect has slightly different characteristics
  - Differences in size, shape, orientation, location, nature

o MATERIAL
  - Each material sample has slightly different characteristics
  - Differences in conductivity, permeability, absorbtivity, etc.
  - Differences in shape, surface texture, etc.

o INSPECTION SYSTEM
  - Each inspection process has slightly different characteristics
  - Differences in penetrant concentration, etc.
  - Differences in sensitivity, linearity, frequency response, etc.
  - Each inspector has slightly different capabilities
EFFECTS OF VARIATION

- NUMEROUS PARAMETERS CAN INFLUENCE DETECTION
  - Some have predictable (deterministic) effects
  - Many effects are not quantitatively predictable
  - Consequently, detection is a probabilistic phenomenon

- IMPERFECTLY CONTROLLED INSPECTION PARAMETERS
  - Perfection is a worthy goal - but always out of reach!

- SCATTER IN MEASURED DATA
  - Both signal and noise data incorporate the effects of variation
  - Measurements will yield distributions of values
EFFECTS OF SCATTER ON POD

**Probability density function of crack detection probabilities for cracks of a specific length**

**Probability of Detection data with lower one sided confidence bound representing the effects of sample size and scatter in the data**
DISTRIBUTIONS OF SIGNAL AND NOISE

A signal-to-noise view of detection

- AREAS ENCLOSED RELATE TO THE ALTERNATIVE OUTCOMES
  - Detection of a defect that is present (POD)
  - Non-detection of a defect that is present (1 - POD)
  - Apparent detection of a defect that is not present (PFA)
  - Non-detection of a defect that is not present (1 - PFA)
DISTRIBUTIONS OF SIGNAL AND NOISE

- POD AND PFA ARE CO-VARIABLES
  - POD is not a function of PFA
  - Both vary with the threshold
  - A lower threshold increases POD and PFA

- CHOICE OF THRESHOLD
  - Difficult to keep design and manufacturing engineers happy!
  - Reducing Type I errors represents improved product life
  - Increasing Type II errors represents increased manufacturing cost
DEPENDENCE ON DEFECT SIZE

- Larger defects generally give larger signals
- For a given POD, the PFA will tend to decrease with increasing defect size
- For a given PFA, the POD will tend to increase with increasing defect size
- For a given defect "size", POD also depends on other parameters
  - Shape, orientation, location, nature
DISTRIBUTIONS OF SIGNAL AND NOISE

- An alternative way of looking at signal and noise
  - Probability distributions of envelope of signal-plus-noise
  - Plotted as a function of signal-to-noise ratio (S/N)
DISTRIBUTIONS OF SIGNAL AND NOISE

A GATEWAY TO EXISTING SIGNAL DETECTION THEORY
- Fifty years of theoretical development and practical application
- NDE applications have been largely ignored

STATISTICAL THEORIES OF SIGNAL DETECTION
- Original applications were to radio communications and radar
  - Predominantly devoted to analyzing reception of pulsed signals
- Based on statistical criteria for testing hypotheses and making decisions
  - Allow derivation of POD and PFA from noise and signal-plus-noise
  - Allow optimization of receiver/detector design
  - Provide rationale for choosing a specific detection threshold
DETECTION THEORY

- RESPONSE TO A BINARY STIMULUS
  - Two possible states of the stimulus
    - Noise alone (N)
    - Signal plus noise (SN)
  - Two possible detection decisions (responses)
    - Noise alone present
    - Signal plus noise present
  - Four possible outcomes
    - N true; choose N .......... true negative ..... probability (1 - PFA)
    - N true; choose SN .......... false positive ...... probability PFA
    - SN true; choose SN .......... true positive ...... probability POD
    - SN true; choose N .......... false negative .... probability (1 - POD)

- THRESHOLDED DETECTION
  - All observed values greater than the threshold are classed as SN
DETECTION THEORY

o OBSERVED PROBABILITIES
  - Probabilities of decision outcomes are estimated from observed frequencies
  - These are conditional probabilities (conditional upon the stimuli)
    - PFA is the probability of an N response given an SN stimulus
    - POD is the probability of an SN response given an SN stimulus

o DECISION CRITERIA
  - Rules to help make a choice between the two responses
    - Based on attaching relative importance to the four outcomes
  - Of many such criteria, two are of most interest
    - Bayes
    - Neyman-Pearson
DETECTION THEORY

- **BAYES CRITERION**
  - Assumptions: Observer has information about the source prior to the test
    - Prior (or "a priori") probabilities ($P_1, P_0$) are known
    - A "cost" ($C_{10}, C_{00}, C_{01}, C_{11}$) can be associated with each outcome
      - (These may be truly known or may be educated guesses)
  - Goal: design the decision criterion to minimize average cost
  - Problem: it is often difficult to assign prior probabilities and costs

- **NEYMAN-PEARSON CRITERION**
  - Assumptions: only the measured conditional probabilities are available
  - Goal: design a test to minimize PFA and maximize POD
  - Problem: these are usually conflicting objectives
  - Solution: set an upper limit to PFA, and maximize POD
DETECTION THEORY

- **LIKELIHOOD RATIO TEST**
  - Bayes and Neyman-Pearson criteria both lead to likelihood ratio tests
  - Likelihood Ratio = \( \frac{\text{Probability of choosing SN if SN is true}}{\text{Probability of choosing N if N is true}} \)
  - The decision depends on whether the likelihood ratio exceeds a threshold
  - The threshold used depends on the test
    - Bayes test: threshold involves the prior probabilities and costs
      \[ E_t = \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} \]
    - Neyman-Pearson test: threshold chosen so that PFA = constrained value

- **AN OPTIMUM RECEIVER**
  - One that presents at its output the likelihood ratio for each receiver input
  - \( LR = \frac{\text{Probability that given amplitude represents SN if SN is true}}{\text{Probability that given amplitude represents N if N is true}} \)
DETECTION THEORY

0 ALTERNATIVE DECISION CRITERIA
  - "Optimal" performance can be defined in many ways
    - Maximize True Positive while restricting False Positive
    - Maximize expected value (minimize "costs") of decision
    - Maximize percentage of correct decisions
  - Individual inspectors may apply their own criteria
    - Give priority to avoiding false alarms, or to avoiding misses, etc.

0 HUMAN FACTORS IN THE DECISION PROCESS
  - Inadequately trained inspectors use widely varying decision criteria
    - Human criteria tend to change with experience and/or training
      - e.g. both POD and PFA tend to decrease with time if SN is small

0 AUTOMATED INSPECTION
  - Decision criteria may vary, but are applied consistently
DETECTION THEORY

Binary communication signals

Signals of random amplitude and phase

Rayleigh fading signals of unknown frequency

EXAMPLES OF OPTIMUM RECEIVERS

- Signals with known or random parameters
- Signals corrupted by additive white Gaussian noise
DETECTION THEORY

- MEASUREMENT OF CONDITIONAL PROBABILITIES
  - Produce the stimulus and measure the response
  - Base estimates of probabilities of outcomes on observed frequencies

- COMMUNICATIONS, RADAR, SONAR: relatively easy to accomplish
  - Switch on the signal source(s), or tell the pilot where to fly

- NONDESTRUCTIVE EVALUATION: much more of a challenge
  - Stimuli are much more varied
  - True stimuli are natural, not manufactured
  - A "referee" technique is needed to identify the SN stimulus
    - For surface defects this might be optical microscopy
    - For subsurface defects no such technique exists
DETECTION THEORY

0 USE OF NATURAL DEFECTS

- Most desirable way of sampling the real defect population
- May be difficult to obtain a truly representative sample
- Good material quality means that defects are rare

- Adequate characterization may present difficulties
- Destructive examination is costly - and destroys the sample!

0 USE OF ARTIFICIAL DEFECTS

- Allows control of at least some of the defect characteristics
  - Size, shape, orientation, location, nature

- Difficult to truly represent the real defect population
  - What we measure is the POD for the artificial defect type
  - Still need to validate these measurements against natural defects
DETECTION THEORY

0 SELECTION OF MEASUREMENT CONDITIONS

- Goal of measurement must be clearly defined
  - Determines the inspection parameters

- Measure "Capability"
  - One inspection of a sample of defects of "identical" or differing size

- Measure "Repeatability"
  - Multiple measurements of a defect sample by the same inspector/system

- Measure "Reproducibility"
  - Multiple measurements of a defect sample using "identical" conditions

- Measure "Variability"
  - Multiple measurements of a defect sample by different inspectors/systems

- Measure "Reliability"
  - Multiple measurements of a defect sample using all these factors
DETECTION THEORY

- **MEASUREMENT OF POD ALONE**
  - Acceptable: POD is not dependent on PFA
  - Knowledge of POD is adequate for determining capability and reliability
  - Knowledge of POD is adequate for predicting product life

- **MEASUREMENT OF POD AND PFA**
  - Recommended: a more complete description of the detection process
  - Knowledge of PFA helps select an optimum threshold
  - Knowledge of PFA helps identify manufacturing costs

- **MEASUREMENT OF ALL FOUR OUTCOMES**
  - Not recommended: POD and PFA fully describe the detection process
  - Probability of a false negative (a miss) is the complement of POD
  - Probability of a true negative is the complement of PFA
SOURCES USED

- D.K. Barton, "Radar Systems Analysis" (Prentice-Hall, 1964)


- S.B. Richmond, "Statistical Analysis" (Ronald, 1964)


