SAND2007-4346C

Characterization of Intrusion Detection System Performance using Design of Experiments

Kim W. Mitchiner, Mark Snell, Frank Griffin
Sandia National Laboratories
Albuquerque, NM

INMM Conference
Tucson, AZ
July 8 – 12, 2007

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Why Alternative Test and Evaluation Processes?

- Generic needs for Test and Evaluation
  - Deliver data in a cost-effective, timely manner
  - Provide data that are predictive and reusable
  - Ensure traceability of experimental variables and data
  - Document system's performance under simultaneously varying conditions
  - Store data in good data-management systems
  - Provide the basis for alternative, formal testing protocols
- Design of Experiments
  - Addresses each of these bullets

Conventional Test and Evaluation Methods

- 30 detections / 30 tests
  - Uses binomial tables to determine probability of sensing at a given confidence level
  - Problems:
    - Pass / Fail test
    - Manpower intensive
    - Does not evaluate interactions among variables
    - Environmental factors and weather conditions are not included as test variables

- One-variable-at-a-time experiment
  - Used to determine the effects of specific variables on the measured response while holding all other variables constant.
  - Problems:
    - Does not consider the interactions between the variables
    - Impractical to perform enough tests to cover the entire set of operating conditions
      - For example, 108 experiments with no repetitions are required to evaluate 4 variables with 3 values each

Overview of Design of Experiments

- Statistical testing methodology
  - Uses all of the factors considered to influence the measured performance response
  - Generates randomized test matrix of unique experiments as a function of the factors
  - Measures response as a function of the interaction of all of the factors
  - Varies multiple factors in the matrix simultaneously for each unique experiment
- Statistical, model-based evaluation methodology
  - Uses statistical methods, such as "regression analysis," to
    - Generate an equation for the performance response as a function of the experimental test factors
    - Identify the significant factors influencing system performance
    - Create a model for predicting responses under different operating conditions

Test and Evaluation of a Radar Intrusion Detection System

- Problem:
  - Determine the viability of an extended-detection radar technology for intrusion detection
  - Assess the impact of degradation factors, such as terrain and environment
- Approach:
  - Conduct 30/30 tests based on target and radar factors only
  - Conduct DOEx experiments based on target, radar, and degradation-factor interactions
- Consideration:
  - Key to any successful DOEx test is the selection of an all-inclusive set of the right factors to test
- Measured Response:
  - Sensing Distance from radar at first detection in meters

Table of Factors and Values

Factor	Low Value	High Value
Vegetation Height (in)	12	48
Number of Targets	1	2
Starting distance from radar (m)	1200	1400
Speed of Target(s) (m/sec)	1	3
Installation Height of Radar above ground (ft)	4	15
Antenna Tilt (deg)	0	30
Scan Angle (deg)	60	180

Test and Evaluation of the Radar Intrusion Detection System (30/30 tests)

Results

- 5 probability of sensing values based on 150 Pass/Fail tests
- Tests – pros:
  - Provided data set that the customer expected
  - Repeatable under similar conditions
  - Required minimal setup time
- Tests – cons:
  - Manpower intensive
  - Did not take into account vegetation height, terrain, etc.
    - Speculation is required to determine performance under a different set of conditions
  - Test variables were limited to one condition at a time

Starting Distance	Average Sense Range from System	Average Distance Traveled	Standard Deviation	Alarms/ Tests	P_s @ 95% Confidence
Hands-and-Knees Radial Crawler (Speed = 0.3 m/s)					
498 m	487 m	11 m	6.3 m	30/30	0.90
Radial Walker (Speed = 1.0 m/s)					
1008 m	996 m	12 m	2.2 m	30/30	0.90
Tangential Walker (Speed = 1.0 m/s)					
1000 m	1000 m	47.4 m	23.4 m	28/30	0.80
Radial Pickup Truck (Speed = 15 mph)					
1405 m	1243 m	162 m	32.8 m	27/30	0.76
Radial Small Vehicle (Speed = 15 mph)					
1405 m	1361 m	44 m	9.7 m	30/30	0.90

DOEx Test Matrix and Results for Upright Walker Target

Variables								Distance from Radar			Results	
Run	Vegetation Ht (ft)	No. of Targets	Starting Distance (m)	Speed (m/sec)	Install Ht (ft)	Antenna Tilt	Scan Angle (deg)	Y1	Y2	Y3	Y avg	S avg
1	12	1	1200	3	15	3	60	1166.3	1169.6	1081.2	1139	50
2	12	1	1400	3	4	0	180	1358.1	1344.2	1356.2	1353	8
3	12	1	1200	1	4	0	60	1190.8	1192.3	1155.9	1180	21
4	48	1	1200	1	4	3	180	437.6	407.7	392.7	413	23
5	48	2	1200	3	4	0	60	434.1	472.9	464.7	457	20
6	48	2	1400	3	15	3	180	831.8	771.1	848.4	817	41
7	12	1	1400	1	15	3	180	1389.6	1383.6	1383.6	1386	3
8	48	1	1200	3	15	0	180	1158.9	1166.6	1168.2	1165	5
9	12	2	1400	3	15	0	60	1363.7	1380.3	1380.4	1375	10
10	48	1	1400	3	4	3	60	478.5	463.4	479.2	474	9
11	48	2	1200	1	15	3	60	1183.6	1184.9	1183.4	1184	1
12	12	2	1400	1	4	3	60	1386.6	1386.6	1383.5	1386	2
13	48	2	1400	1	4	0	180	478.4	486.8	477.3	481	5
14	12	2	1200	3	4	3	180	1151.2	1166.4	1192.1	1170	21
15	48	1	1400	1	15	0	60	1378.1	1381.1	1381.3	1380	2
16	12	2	1200	1	15	0	180	1187.9	1198.5	1193.9	1193	5

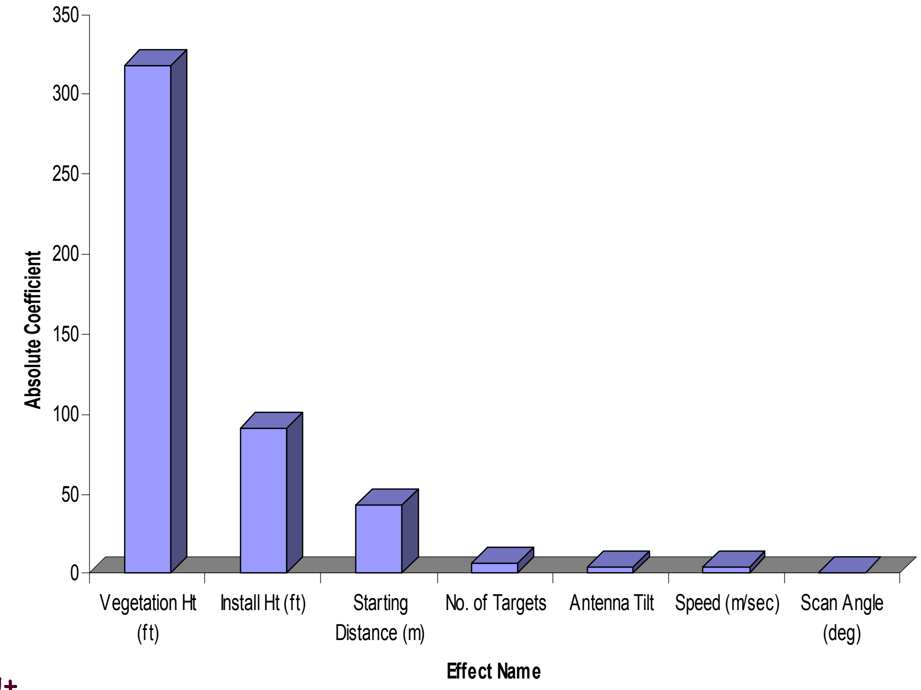
- 16 unique experiments were generated for the 7 factors identified
  - DOEx screening experiment (fractional factorial)
  - Only main effects are considered
- Each experiment was repeated three times
- The averages were used as input into the statistical regression analysis

DOEx Screening Analysis Results for the Radar System

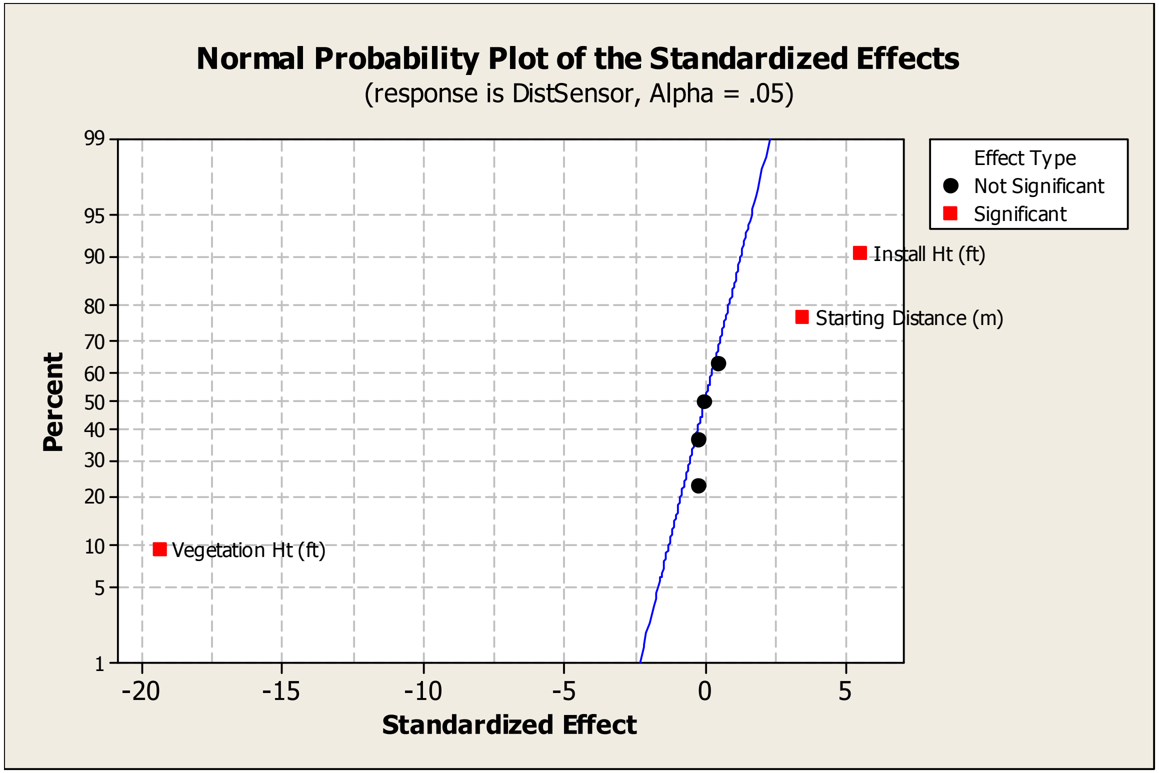
- DOEx Screening Analysis provided data on the main significant effects and some limited data on interactions
  - Indication that Vegetation Ht X Install Ht could be significant
- Significant Main Factors:
  - Vegetation Height (in)
  - Installation Height (ft)
  - Starting Distance (m)
- Regression Equation:

Detection Distance (m) = 588.73 – 17.67 X Vegetation Ht + 16.46 X Installation Ht + 0.56 X Starting Distance

Pareto Chart of Degree of Factor Significance

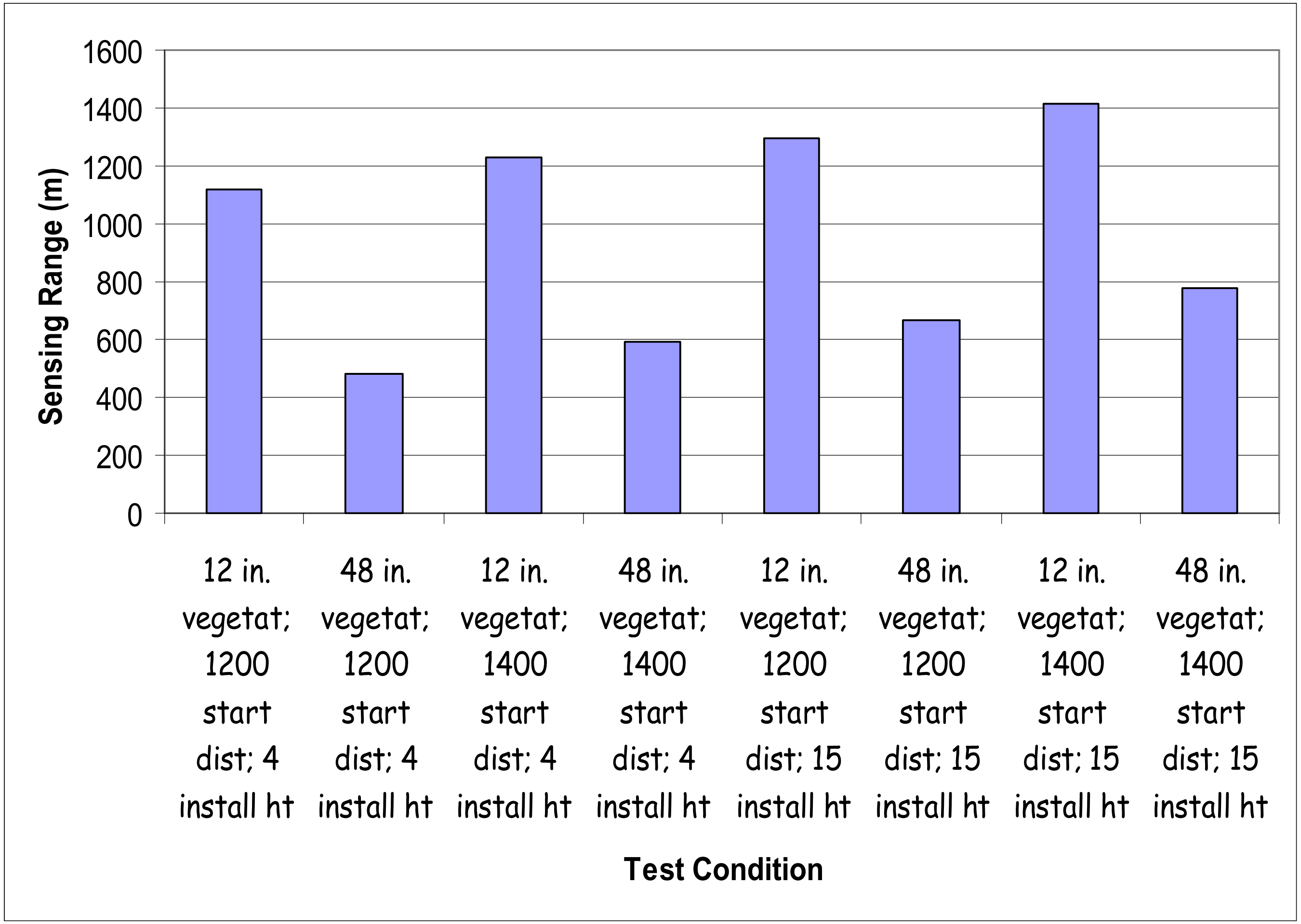


Comparison of Factors based on Normal Probability



The greater the distance from the line, the larger the influence

Sensing / Detection Distance (m) for Upright Walker Target



- Vegetation Height was the predominant discriminator for sensing distance
- Starting distance and installation height had relatively smaller effect

National Nuclear Security Administration

Sandia National Laboratories

Summary of DOEx and 30/30 Results for the Upright Walker Target

- 30/30 Tests
  - For each set of tests, only one condition was tested
    - Crawler (1 m/sec)
    - Walker (either radial or tangential) (3 m/sec)
    - Vehicle (large or small) (15 mph)
  - 150 tests were required to generate 5 probability of sensing values
  - Possible to extrapolate, if all of the variables (terrain, target, speed, radar) are held constant:
    - Approximate time required for the intruder to reach the sensor
    - Approximate point of first detection

- DOEx Tests
  - Evaluated 7 factors simultaneously in 16 unique experiments (repeated three times (48) for the walker/crawler only)
  - Provided a "predictive equation" for sensing distance at first detection
    - Enables the experimenter to evaluate conditions and locations not tested
    - Generates data with a given statistical accuracy
  - Probabilities of sensing can be calculated for a wide range of conditions
  - Provides data on
    - Sensitivity of sensing to variations in conditions/factors
    - Approximate points of first detection for several conditions
    - Approximate times for the intruder to reach the sensor

Test and Evaluation of the Radar Intrusion Detection System -Summary-

- Pros for DOEx
  - Fewer overall tests could be performed than for 30/30 tests, and an unlimited number of 0.90 P_S range values could potentially be provided under varying "tested" conditions using predictive model
  - Predictive model could be used to generate a theoretical detection envelope
  - Although test results were expected (most vs. least significant variables), the predictive model would help the SME or analyst to determine expected performance values of a system under untested, site-specific conditions

- Cons for DOEx
  - Very hard to find many of the test variable conditions at a site; just because they exist does not necessarily mean they exist together
  - System configuration changes may not save man hours in long run vs. 30/30 tests
  - Variable combinations need to be thoroughly evaluated and thought out
    - For example, performing a high-speed belly crawl at night in tall vegetation with undulating terrain does not make sense

Conclusion

"Customers are beginning to require more rigorous T&E methodologies"

Design of Experiments:

- Provides early input into the characterization of systems
- Provides insight into the performance of systems
- Optimizes and minimizes the number of tests required
- Can include defeat and degradation conditions as varying factors
- Maximizes information per observation
- Can be cost effective over the long term (as we get smarter!)