

A TWO-STAGE MODEL OF
RADIOLOGICAL INSPECTION:
SPENDING TIME^{*}

William S. Brown
Brookhaven National Laboratory
Upton, New York

^{*} This work was performed under the auspices of the U.S. Department of Energy

A TWO-STAGE MODEL OF RADIOLOGICAL INSPECTION: SPENDING TIME

William S. Brown
Brookhaven National Laboratory
Upton, New York 11973-5000

The paper describes a model that visually portrays radiological survey performance as basic parameters (surveyor efficiency and criteria, duration of pause, and probe speed) are varied; field and laboratory tests provided typical parameter values. The model is used to illustrate how practical constraints on the time allotted to the task can affect radiological inspection performance. Similar analyses are applicable to a variety of other tasks (airport baggage inspection, and certain types of non-destructive testing) with similar characteristics and constraints.

INTRODUCTION

In settings in which radioactive material is created or used (e.g., medical applications, power generation, and weapons production), radiological surveys are conducted to confirm that equipment and surfaces are free of contamination. Surveyors scanning for contamination must interpret the audible output or visual reading of a survey instrument to determine when the signal (clicks or visual readings) exceeds the background level by a margin sufficient to conclude that contamination is present. The task of detecting low levels of contamination is difficult because both the signal and the background are variable.

In recent years the human performance aspects of this task have been investigated (Brown & Abelquist, 1998), and approaches for taking human performance into account in designing radiological surveys have been proposed (Abelquist & Brown, 1999; Abelquist, Brown, Powers, & Huffert, 1998). The present paper extends this work by considering in detail how surveyors 'spend' time.

RADIOLOGICAL SURVEYING

Description of The Task

When inspecting for radioactive contamination, surveyors move a hand-held probe over surfaces while monitoring an instrument for evidence of a count rate in excess of background. While moving the probe, the surveyor must hold it very close to, but not allow it to touch, the surface. This requirement, coupled with the fact that the human auditory system is particularly well suited to picking up temporal changes, leads surveyors to perform the task primarily based on the audio output (i.e., clicks) of the survey instrument. The presence of radioactive sources is indicated by increases in the rate of clicks that the surveyor judges to be greater than those which normally occur owing to the random nature of radioactive decay.

In practice, surveyors do not make decisions based on a single indication. Rather, upon noting an increased number of counts, they pause briefly and then decide whether to move on or take further measurements. Thus the detection of radioactive sources by scanning consists of two stages: continuous monitoring and stationary sampling. In the first stage, characterized by slow, continuous movement of the probe, the sur-

veyor has only a brief "look" at potential sources. Hence sensitivity is relatively low. The surveyor's criterion (i.e., willingness to decide that a signal is present) at this stage is likely to be liberal, in that the surveyor should respond positively on scant evidence, since the only "cost" of a false positive is the time required for the pause. The second stage occurs only after a positive response has been made at the first stage. It is marked by the surveyor interrupting the scanning and holding the probe stationary for a period of time, while comparing the instrument output signal during that time to the background counting rate. Owing to the longer observation interval, sensitivity is relatively high. For this decision the criterion should be stricter, since the cost of a "yes" decision is to spend considerably more time and effort to take a static measurement. If the observation interval is sufficiently long, an acceptable rate of source detection can be maintained despite application of the more stringent criterion. Performance will depend on optimally allocating time to both stages of the decision.

A Model of Survey Performance

Owing to the statistics of radioactive decay, the performance theoretically achievable for a given source intensity and background can be predicted based on the length of the observation. Figure 1 shows theoretical operating characteristics for the detection of a source producing 120 counts per minute (cpm) in a background of 60 cpm, for observation intervals of 1 and 4 seconds. Thus sensitivity is a function of the observation interval, i.e., the time for which a source is under the moving probe, or the time for which a surveyor pauses over a suspected source.

As indicated earlier, performance depends on surveyors' criteria as well as sensitivity. For the 1-second observation interval (dotted line) shown in Figure 1, an observer adopting the operating point (open circle) near the middle of the characteristic would detect the source 80% of the time, and would also identify background activity as a source roughly 26% of the time. If the situation were such that missed signals should be strongly avoided, the observer might choose an operating point closer to the upper right corner. In this case, 95% of the sources would be detected, but the rate of false positives would increase to roughly 63%.

Figure 1 also shows that, for the same source and background intensities, a 4-second observation will allow the

source to be detected 90% of the time with a false positive rate of 5%.

The earlier discussion of the task implies a combination of a short and a long observation; this is depicted in Figure 2. The dotted line represents ideal performance for a 1-second observation. The open circle represents the observer's decision at the first stage. The combined performance of the first and second stages is represented by the solid line, which represents a 4-second observation. The line begins at the initial decision point (open circle), not at the upper right corner of the plot; this is a consequence of the assumption that the stages are sequential, i.e., a 'miss' at the first stage removes an item from further consideration, setting an upper limit on the final true-positive rate. The observer's final decision is represented by the filled circle. In this example, the observer's goal was to achieve better than 90% true positives while limiting false positives to 10% or less.

The implication of this view of the scanning task is that to perform within specified limits on missed sources and false calls, surveyors adopt criteria at each of stage of the scanning process to keep the joint performance within limits. It also follows from the preceding discussion that, since the time available to perform a survey is necessarily limited, surveyors' performance will depend on how they allocate time; i.e., surveyors balance the deliberateness with which they move the probe and the frequency with which they pause against the time they spend on each. Pausing more often and moving more slowly allows more sources to be caught in the first phase; longer pauses allow higher sensitivity in the final decision. Since the time available is fixed, the time spent owing to the *frequency* of pausing would be expected to trade off against the time actually spent paused. Put simply, if time is limited, the greater the number of pauses, the faster the probe must be moved between pauses, or the shorter each pause must be. Performance of surveyors in field trials was examined in detail for evidence of this relationship.

MODELING FIELD PERFORMANCE

Controlled field tests were conducted to examine the detection of sources by actual surveyors. The tests employed actual radioactive sources and scanning instrumentation. The surveyors' behavior was recorded in a way that allowed both true positive and false positive rates to be estimated.

The behavior of surveyors in the field tests conformed to the expectations described above. That is, surveyors typically adopted liberal criteria during scanning (i.e., they paused often), and more stringent criteria for their final decisions (i.e., they made relatively few false calls error). However, surveyors varied as to how well they performed the tasks. The performance of three surveyors who completed the survey in about the same amount of time with differing levels of accuracy were fit to the model of survey behavior; the model's depiction of their performance is shown in Figures 3-5.

The model used a normal distribution to approximate, for assumed background and source intensities, the numbers of counts expected to occur in specified intervals. For this, operating characteristics and operating points were generated to match surveyors' field performance. This process is described below.

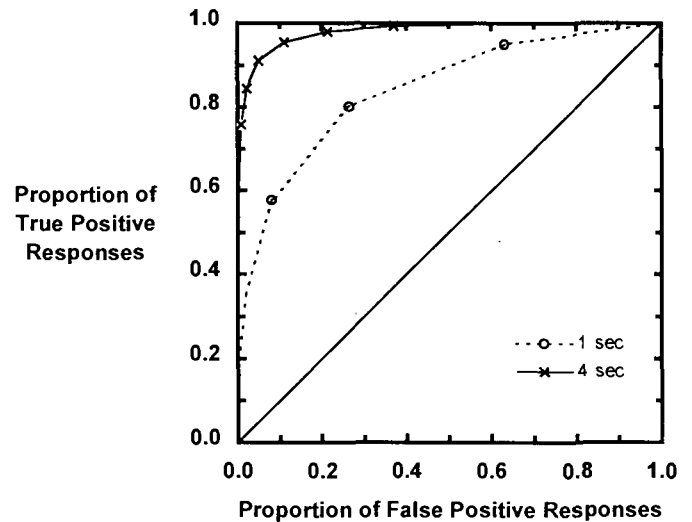


Figure 1. Theoretical operating characteristics for observation intervals of 1 second and 4 seconds.

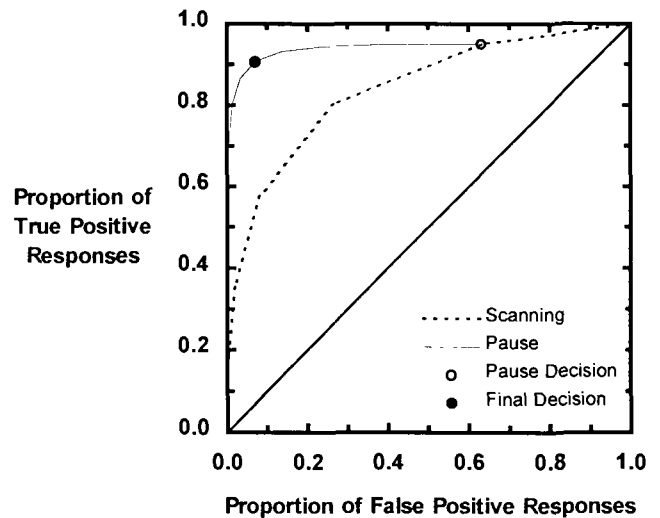


Figure 2. Example of theoretical performance of a two-phase detection process.

First, values representing the observation interval in the first phase (the inverse of probe speed) were selected so that the operating characteristic produced by the model (assuming reasonable values the average intensity and size of the sources and for the efficiency of the surveyors) passed through a point in ROC space corresponding to the surveyor's performance for that phase (i.e., the decision to pause); this is shown by the dotted lines in the figures. The assumed criterion value was then adjusted so that the operating point (open circle) corresponded to the observed performance. Then the observation interval for the second phase (the length of the pause) was chosen so that the operating characteristic (the solid line) passed through the point representing the surveyor's final overall performance (the filled circle).

Surveyors #1 and #2 (Figures 3 and 4) both paused over all of the sources during scanning; this is indicated by the open circles near the top of each plot. Though neither missed any

sources at this initial phase, Surveyor #1 paused less frequently than Surveyor #2; the initial operating point in Figure 1 is further to the left than that in Figure 2. This put #2 on a 'lower' operating characteristic, as indicated by the dotted curves passing through the points. In the model, this indicates that #2 was scanning more quickly than #1; i.e., the initial observation interval assumed was shorter.

Surveyor #1 did not perform as well as #2; the surveyors correctly flagged 12 and 13 of the sources, respectively. The fit of the model for the final decision point (the filled circle near the 'crosshairs' on in each figure) was produced by varying the assumed time for which the surveyor paused. The trajectory of the operating characteristic is lower in the ROC space for a shorter pause (Figure 3), higher for a longer pause (Figure 4).

Surveyor #3 paused over all but one of the sources during scanning; the operating point (open circle) and therefore the operating characteristic (dotted line) from the model are lower in the plot than for either of the other surveyors (neither of whom missed any sources initially). As before, the fit assumes that other things are equal, and that performance depends on the observation interval – in this case the surveyor is modeled as moving more quickly than the others.

While performing less well than the other surveyors in the initial phase (missing one of the sources during scanning), Surveyor #3 correctly flagged all of the sources paused over, with virtually no false calls (as indicated by the filled operating point falling very close to the left side of the plot). The overall performance of #3, therefore, was much better (in terms of a d' measure of sensitivity) than that of the other two surveyors.

The surveyors' performance, as depicted in the model, can be summarized as follows. Surveyor #1 moved slowly, pausing as needed - but not for very long (therefore possibly failing to flag sources paused over and making some false calls). Surveyor #2 moved more quickly but pauses often (thereby not missing sources in the initial stage), and for a longer time (correctly flagging all sources paused over). Surveyor #3 moved most quickly (missing a source at the initial stage), but with long pauses (flagging all sources paused over and making very few false calls). Thus, other things being equal, performance can be seen to depend on the surveyors' strategies for allocating time to the different components of the surveying task.

ACCOUNTING FOR TIME

In view of the apparent trading relationship among the deliberateness of scanning, the frequency of pausing and the duration of pauses, an additive expression was devised to describe the time spent in the survey task. It was based on the performance of Surveyors #1 and #2, who, as mentioned earlier, completed the survey with comparable accuracy in the same amount of time (roughly 45 minutes). The aim was to use the observation interval estimates that came from fitting the surveyors' performance and the number of times they were actually observed to pause during the field test to account for the time taken to complete the task.

First, since some time is spent on activity other than surveying (e.g., moving from one location to the next), the total time to be accounted for was reduced by a reasonable amount.

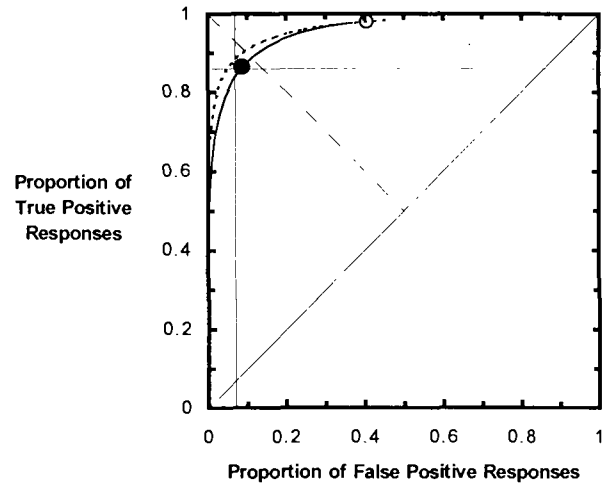


Figure 3. Performance of Surveyor #1, who paused over all 14 sources and correctly flagged 12.

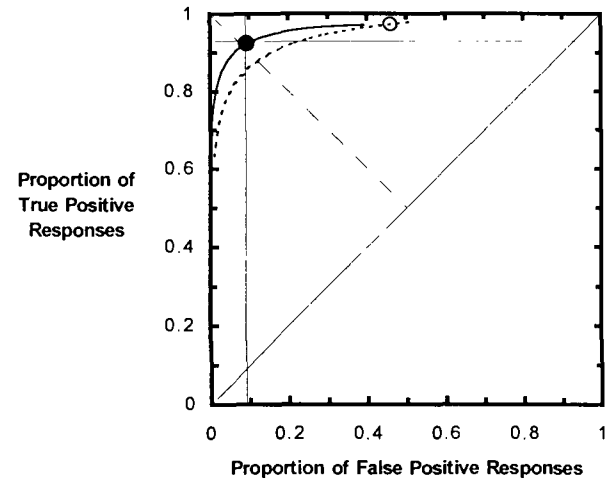


Figure 4. Performance of Surveyor #2, who paused over all 14 sources and correctly flagged 13.

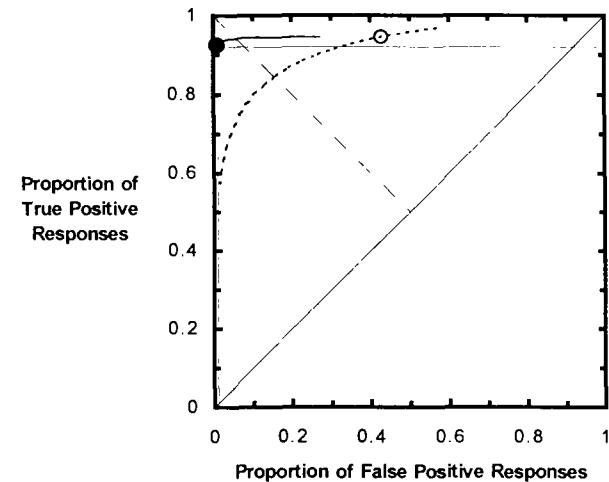


Figure 5. Performance of Surveyor #3, who paused over 13 sources and correctly flagged 13.

Next, the actual number of times each surveyor was observed to pause during the field trials was multiplied by the length of that surveyor's pauses as estimated by fitting the model; this represented the amount of time for which the probe was stationary. Finally, the remaining time (for which the probe was assumed to be moving) was applied to the area that was surveyed in proportion to the initial observation interval estimated from the fit for each surveyor. That is, the area to be covered and the time remaining were used to calculate roughly how much time it took each surveyor to cover a given distance.

The result was an expression for the time spent surveying that was based on the scanning speed and pause duration values inferred from the surveyors' performance using the model described earlier. While it was encouraging that the estimates of the rate at which the probe was moved that were obtained by this method were roughly comparable to what would be expected based on a surveyor's rule of thumb, a further test was also done. The performance of a surveyor who completed the field test in the shortest time was fitted as described earlier, and the obtained values were used to estimate the time spent on the task. The result was roughly 30 minutes, which matched the time recorded for that surveyor in the field test.

In fitting the performance, all of the surveyors were assumed to be operating with equal (though not perfect) efficiency. In fact, there is evidence that surveyors differ (Brown & Abelquist, 1998), probably owing to such factors as training, experience and technique. It was noted that another participant in the field test, whose results are shown in Figure 6, took 45 minutes to complete the task, but performed no better than the surveyor who took 30 minutes. In this case, the efficiency parameter could be estimated from the model, based on the time; a test of the approach would require an independent estimate of efficiency, which was not available.

CONCLUSIONS

The values assumed for the durations of pauses in the calculations were plausible based on 'part-task' laboratory simulations of scanning, informal observations of surveyors, and the opinions of survey personnel. However, the actual duration of pauses was not measured during the field tests. Neither was the actual speed at which surveyors moved the probe; values used in the calculation were inferred based on the area that was scanned (roughly converted to linear distance based on the size of the probe) and the time available.

It is also important to note that the sources in the field tests (which were designed to approximate actual field contamination configurations) varied as to intensity and area. The values used in the calculation, therefore, are only representative.

Although it was necessary to infer some key parameters (rather than having of actual field measurements), the results described here nevertheless demonstrate how one might begin to take into account the explicit or implicit time constraints

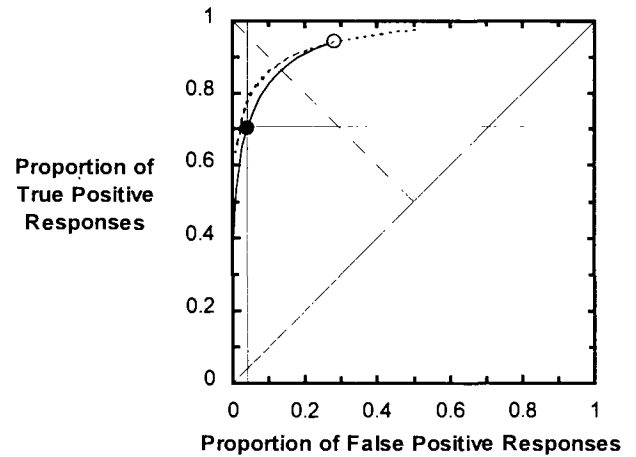


Figure 6. Performance of Surveyor #4, who paused over 13 sources and correctly flagged 10.

that exist in radiological surveying and other types of inspection. It is quite feasible to perform field tests so as to provide actual data for quantities that had to be estimated in the analyses reported in this paper.

Such an experiment could then be combined with one employing psychophysical techniques (see Brown, 1995; Brown & Abelquist, 1998) in order to determine whether individual differences in scanning performance are a result of differences in sensitivity or are in fact a reflection of how a surveyor 'spends time.'

The analysis described here illustrates how practical constraints on the time allotted to the task might affect radiological inspection performance. Similar analyses are applicable to other tasks (e.g., airport baggage inspection, or certain types of non-destructive testing) with similar characteristics and constraints. This approach should allow the effects of time constraints on inspection performance to be anticipated and perhaps contribute to the development of strategies that optimize the inspectors' use of time.

REFERENCES

- Abelquist, E.W., & Brown, W.S. (1999). Estimating minimum detectable concentrations achievable while scanning building surfaces and land areas. *Health Physics*, 76, 3-10.
- Abelquist, E.W., Brown, W.S., Powers, G.E., & Huffert, A.M. (1998). *Minimum detectable concentrations with typical radiation survey instruments for various contaminants and field conditions* (NUREG-1507). Washington, DC: U.S. Nuclear Regulatory Commission.
- Brown, W.S. (1995). *Signal detection theory analysis of performance in a simulated radiological survey task*. Paper presented at the Human Factors Society 39th Annual Meeting, San Diego, CA.
- Brown, W.S., & Abelquist, E.W. (1998). *Human performance in radiological survey scanning* (NUREG/CR-6364). Washington, DC: U.S. Nuclear Regulatory Commission.