A General Approach to Spatio-Temporal Correlation Using The Efficacy of Counter-IED Intervention as an Example

Edward Tomme

CyberSpace Operations Consulting, Inc.

February, 2010

Abstract

This article is a direct descendant from Nair *et al.* [1] and Rivolo [2] with a similar end goal: assessment of the effectiveness of different types of interventions on the occurrence of successful and unsuccessful explosive events during a specific portion of a real-world counterinsurgency operation. Its primary purpose was initially to perform an independent check on the approach and results of these two references. However, it also considers effectiveness from several points of view including the same cumulative before-/after-event time period as previously used as well as a two new schemes used to investigate whether the cumulative nature of the previous method biased the conclusions at larger time periods. It is shown that the previously used cumulative scheme biases the results toward the low-time values.

A comparison of actual explosions to explosive devices found and cleared is performed to determine whether the interventions had a statistically-significant effect on the carelessness of the insurgents, as measured by the post-intervention variation on the found/cleared-to-exploded ratio.

A discussion of the potential effects of bad data points in the set is also performed, concluding that inadvertent inclusion of these points would have negligible effect on the results.

Intervention Type I is found to be counterproductive. While initially increasing enemy carelessness for a period of a few weeks, it ends up actually increasing enemy activity by 125% to 250% for almost six weeks following the intervention out to a range of about two kilometers. Interventions of this type should be discontinued.

Intervention Type II is found to be highly effective in reducing enemy activity over short (~1-2 km.) ranges for several months. It not only increases enemy carelessness but reduces the number of events by 25% to 50% for over 90 days.

1. Introduction

In this paper we investigate a data set consisting of the time and location of two kinds of *events*, 29,964 explosions and 20,223 found and cleared explosive devices occurring over the course of 547 days. We compare the timing and locations of these events to two different types of *interventions*, hereinafter referred to Intervention Types I and II. There were 107 occurrences of Intervention Type I and 201 occurrences of Intervention Type II. Collectively, events and interventions shall be referred to as *incidents*. The goal of the comparison was to determine whether the interventions had any statistically-significant effect on the occurrence of the events and, if so, the duration and spatial extent of the effect.

A time-series representation of the incidents is shown in Figure 1, while a geographic representation of the incidents is shown in Figure 2. The actual timings and locations of the events and interventions have been shifted for operational reasons; however the shifts were performed in such a manner as to not affect the results. Notice the significant extension of the event data both before and after the intervention data. Such extensions enabled the before/after comparison described below.

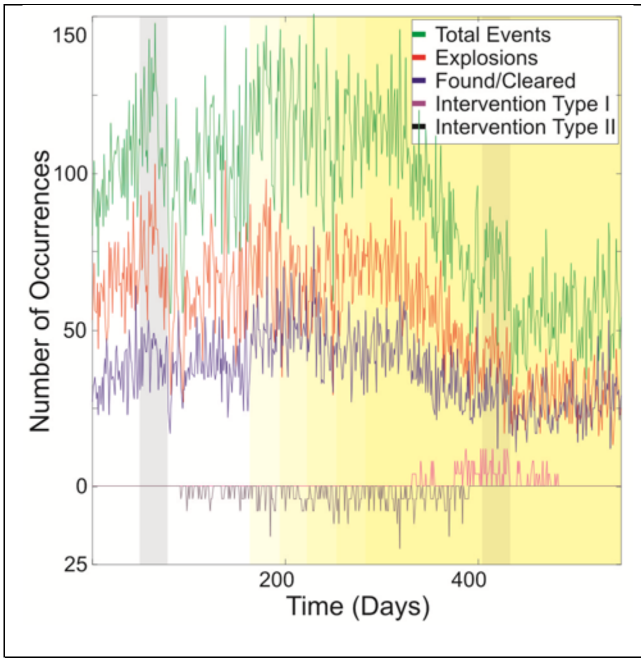


Figure 1. Temporal Distribution of Events. Intervention counts are multiplied by four to enhance visibility. Intervention Type II counts are inverted for clarity. Bin size is one day. Grey vertical bars indicate Ramadan. The yellow region indicates the American troop surge.

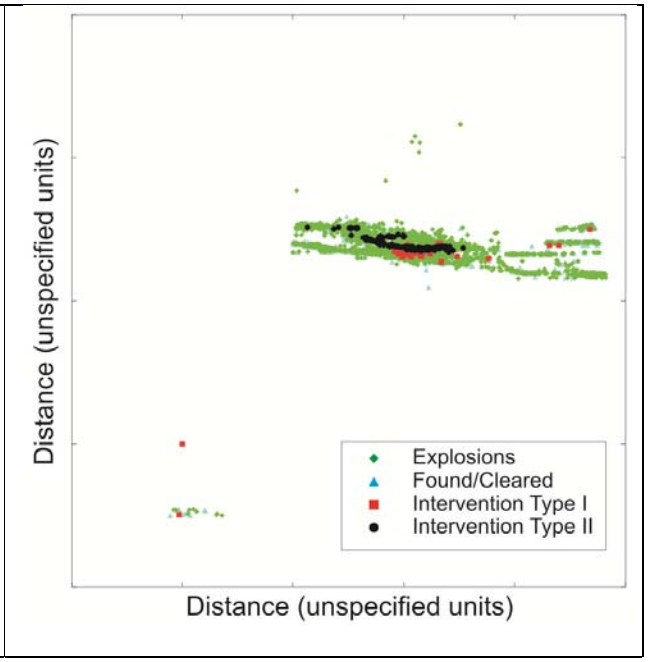


Figure 2. Spatial Distribution of Events. Extent of diagram is on the order of 10^6 meters on a side. Not all of the 50,000 plus points are distinguishable in this diagram.

Figure 3 and Figure 4 show the different events and interventions binned by the day of the week on which they occurred, with Figure 4 showing the total counts and Figure 5 showing each category as a fraction of the maximum number of counts for that category. It is worth noting that the Muslim weekend occurs on Thursdays and Fridays. Examination of Figure 4 does show some variation in the actual numbers of events, but the maximum and minimum total events do not vary from day to day in an extreme manner. More illustrative, however, is Figure 5. Notice that the total events on Friday, Saturday, and Sunday are the lowest, at 8-11% below maximum. . In contrast the remaining days' total events are all within 4% of the maximum level, which occurs on Thursday. Thursday is a popular day, having the maximum fraction of total events, explosions, and interventions of type II and the second largest fraction of found/cleared events. We shall discuss the variation in the number of explosions and found/cleared devices later in this paper.

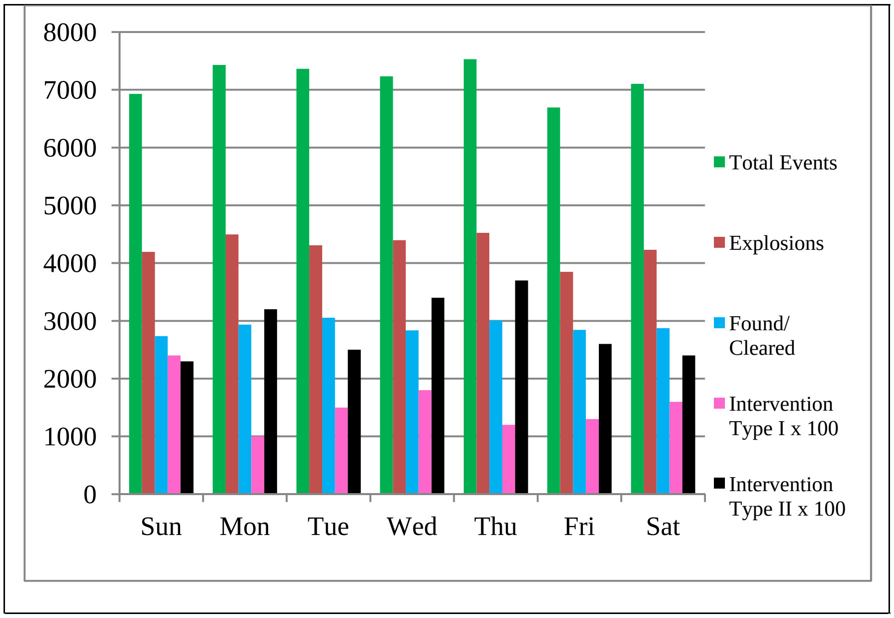


Figure 3. Incident Counts by Day of Week. Intervention counts multiplied by 100 for visibility.

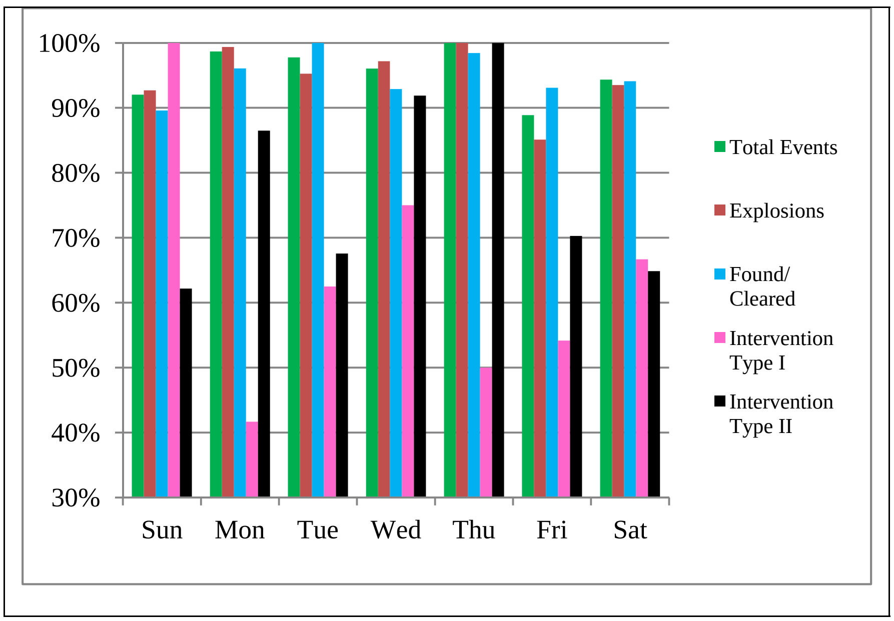


Figure 4. Incident Counts by Day of Week, Percent of Maximum Count.

Even more pronounced is the variation in which days the interventions were executed, especially type I interventions. While the relatively small numbers of interventions when compared with the number of events could add some statistical variation, there was still a sufficient number to get statistically meaningful results. Sundays were by far the most popular day to execute type I interventions, and a mere 40% of the maximum number were executed the following day. In fact, no more than 75% of the Sunday totals were executed on any other day. Such a result could indicate some level of predictability in the campaign. The type II intervention campaign was less predictable, with Mondays, Wednesdays, and Thursdays being about 10-20% more likely to have an intervention than the other days.

The variation in the number of events on each day, while not extremely pronounced, will be addressed when we discuss the binning of data in the next section.

2. Approach

A flow-chart overview of the process described below is presented in Appendix I.

To determine whether the interventions had any effect upon the events, a statistic similar to that used by Nair *et al.* in [1] was employed. The general approach of this study and those in [1] and [2] is to take each intervention of a particular type and note the timing and spatial proximity of the events occurring around it. First, the time and location of an individual intervention is noted. Next, each event was compared to the intervention to determine the differences in time and location. If both the time and location differences fall within certain predetermined limits, the event is added to an appropriate accumulation bin. The bins are arranged by range between the event and the intervention, and by number of days before or days after the intervention. If the event is outside either the time or space limits it is ignored. Once the event has been binned or ignored, the next event is similarly compared to the intervention. This process repeats until all events have been compared to the intervention. The next intervention is then selected and the event comparison process is repeated. The full comparison process is repeated until all interventions are compared with all events.

The algorithm used for this analysis also segregates data comparing the different types of interventions individually with the different types of events so that the maximum flexibility in analysis could be retained. Thus, for example, one may investigate the correlation of Intervention Type I with explosions, with found/cleared devices, and with a combination of the two event types.

Figure 5 gives an example of how the process works for a single intervention. The selected intervention is shown at the center of the spatial and temporal plots and labeled with the number 0. The selected range bins and time bins are shown as are the maximum range and time considered. Explosion 1 is not binned because it occurs outside the maximum considered range. Found/Cleared 3 is also not binned because, although it lies within the considered range, it occurs before the considered time period. Explosion 4 would be binned in range 3, day 2. The complete binning of all the events as related to the selected intervention is shown in the binning frame of the figure.

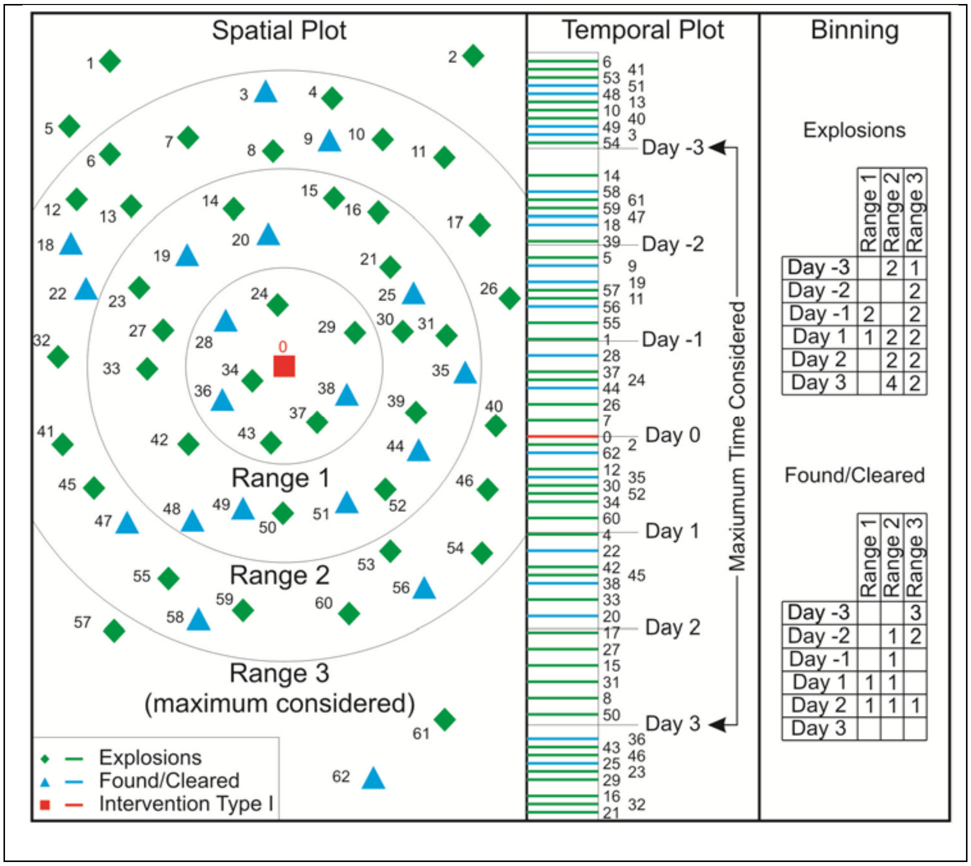


Figure 5. Illustration of the Algorithm. Notional spatial and temporal data is presented and the resulting binning is shown. Numbers next to events in the spatial plot correspond to numbers next to the temporal locations.

Once all interventions have been correlated with all events and the results binned, the results need to be accumulated appropriately for analysis. The normal way of displaying such results is to show a statistic related to the effectiveness of the intervention on the vertical axis and the range between interventions and events on the horizontal axis, with a different plot shown for the various days following the intervention.

The method of accumulating the data for display on the various days can be done in a number of ways. Examples of three of these ways are shown in Figure 6. In [1] and [2], the statistic is applied cumulatively, so that results displayed for day one include only the day-one bins while, for example, results applied for day three include the day-one, day-two, and day-three bins. An advantage to this approach is that the additional counts accumulated in this way enable more accurate variances to be calculated during the bootstrapping process to be described later. However, that approach also tends to bias the results toward the earlier time bins since those data are used in all subsequent results.

An alternative method may also be applied that does not apply the same bias. Instead of a cumulative approach, a sliding window of dates located symmetrically about the intervention may be used. The width of this window can be adjusted to fit operational or statistical needs. Finally, instead of looking at symmetrical time periods before and after the event, it is possible to take the average results for some period immediately prior to the intervention and compare them with post-intervention time periods. An advantage of this method is that the pre-intervention

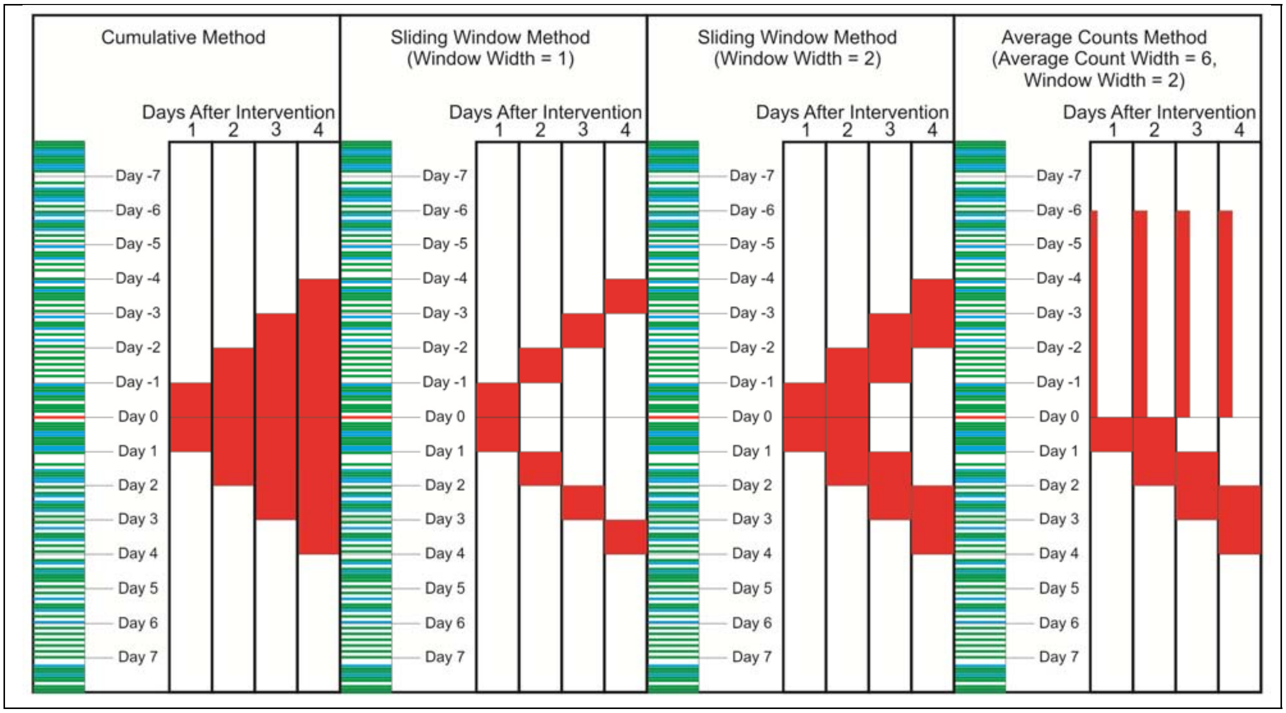


Figure 6. Temporal Windows Illustration. Frame 1: cumulative method; frame 2: sliding window method, window width = 1 day; frame 3: sliding window method, window width = 2 days; frame 4: average count method, average count width 6 days, window width 2 days (note the changing width of the averaged before-intervention data).

data is more relevant to warfighters since it provides a baseline of events immediately preceding their actions. It also could be less influenced by day-of-week considerations. For example, should it be discovered that a greater number of events occur on Saturdays, then the symmetrical spacing of time periods would inevitably end up comparing Saturdays with some other day of the week when a lesser number of events would occur. By using the average number over, say, a period of several weeks immediately prior to the intervention could lessen the effect of this biasing. Results using all of these methods of data accumulation are presented later in this paper.

Once the data had been appropriately correlated and binned, a method of determining the effectiveness of the intervention needed to be developed. A statistic similar to that used in [1] was employed in this analysis. A brief justification follows. Let *B* and *A* denote the number of events occurring within a certain time period before and after an intervention, respectively. Let *x* and *y* show the functional spatial relationships and *t* show the functional temporal relationship. The statistic used in this analysis was

$$\phi(x,y,t) = \begin{matrix} 2\left(\frac{A(x,y,t)}{A(x,y,t)+B(x,y,t)}\right)\ if\ \ A+B\ \neq 0 \\ 1\ if\ A+B=0 \end{matrix}. \tag{1}$$

In essence, this statistic compares the number of events occurring within the defined spatial limits but occurring within time *t* after an intervention to the total number of events occurring between time *t* before and time *t* after the intervention. If the number of events before and after the intervention within the specified spatial and temporal regions is roughly the same, then this statistic is on the order of unity. Should the intervention have the effect of reducing or increasing the number of events, the statistic will be less than or greater than unity, respectively. The factor of two is arbitrary and only serves to make the no-change case result in an easily discernable number. Rivolo gives a more detailed description of the development of such statistics in [2].

Once the statistic has been applied to the accumulated, binned results it remained to determine confidence intervals for the data. The well-known bootstrap method (e.g., Press *et al*. [3]) was used in this analysis to obtain these intervals. A set of interventions equal in size to the original intervention set drawn randomly with replacement from the original set of interventions was used as the bootstrap data set. The distribution of values at each time/range combination for all bootstrapped data sets was assumed to be normal, and the normal distribution was used to determine the confidence intervals.

The method of Rivolo (see [2]) was used to normalize the data. In essence, it was assumed that the effect of the intervention would eventually become zero at large distance. Thus, the statistic should approach unity as range increased. The method of normalization takes several points near the maximum considered range for each time step under consideration, computes their average, and determines the required shift for the time step. The results and confidence intervals for each time step are then shifted by the appropriate normalization constant.

While the statistic defined by Equation (1) is fairly straightforward to mathematicians, it may be fairly difficult to understand for the warfighter for whom the analysis is ultimately aimed. The reason for this difficulty is that an attribute of the statistic is that it is non-linear in the very quantity in which warfighters are likely to be interested: the fractional reduction of events following an intervention. A question a warfighter is anticipated to ask is, “By exactly what percentage were incidents reduced following our interventions?” Using the statistic defined by Equation (1), it would be difficult to answer this question directly.

To illustrate this problem, solving Equation (1) for A/B will give the fractional reduction or increase in the number of events following an intervention:

$$\alpha = \frac{\phi}{2-\phi}, \qquad (2)$$

where α is defined as A/B and the time and space functional dependencies are omitted for clarity. This function is plotted in Figure 7, and it is clearly non-linear. Thus, for the remainder of this paper we shall report α instead of ϕ.

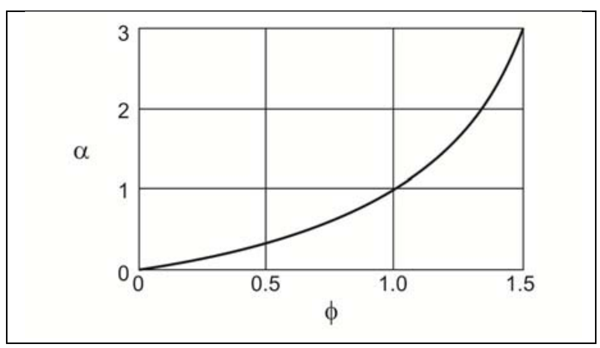


Figure 7. Plot of the fractional number events occurring after an intervention compared to the events occurring before the intervention as a function of the statistic developed in Equation (1).

3. Algorithm Validation

In order to confirm the validity of the model, several intervention/event data sets with predetermined characteristics were analyzed. Confidence intervals were not calculated for these tests as they were designed to demonstrate the validity of the comparison algorithm, not the bootstrap. Calculation of the confidence intervals was based on the same algorithm, so validation of that technique follows from validation of the base algorithm.

First, a data set involving a single intervention and 10^4 events randomly distributed in both time and space around the intervention time and location was used. Figure 8 shows the results of this test. As expected, the results showed statistical fluctuations above and below the no-change value of unity since the events and interventions in this case were completely uncorrelated. By adding more and more interventions, the statistical fluctuations decreased until, with over 10^6 events included, the statistic ϕ was almost indistinguishable from unity at all ranges in the displayed plot. The point of showing the data with 10^4 points is to illustrate the statistical fluctuations present in such data, fluctuations that will be even more pronounced in the results from the actual data, which typically will contain significantly fewer than 10^4 relevant, correlated points.

Next, the same intervention was correlated with a set of events that could be modified to increase in frequency at specific range intervals or time intervals from the intervention. For the first trial, events were all randomly distributed in space about the intervention location but were skewed in time so that three times as many events occurred after the intervention as before, or A = 3B.

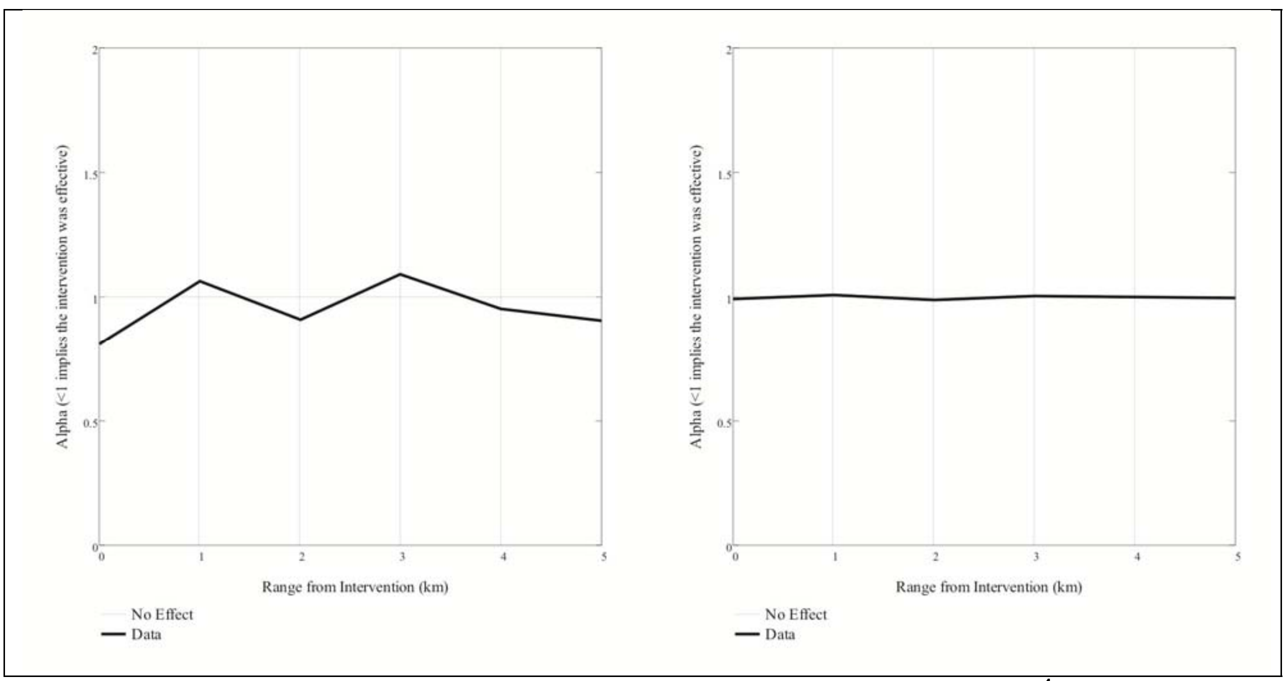


Figure 8. Algorithm Test Using Randomly Distributed Events. Left frame uses 10^4 events; right frame uses 10^6 events. As expected, larger statistical deviations are apparent in the smaller data set.

From equation (1), the statistic should approach 3 under these conditions. Figure 9 shows the results of that test for 10^4 and 10^6 events, respectively, showing the expected behavior.

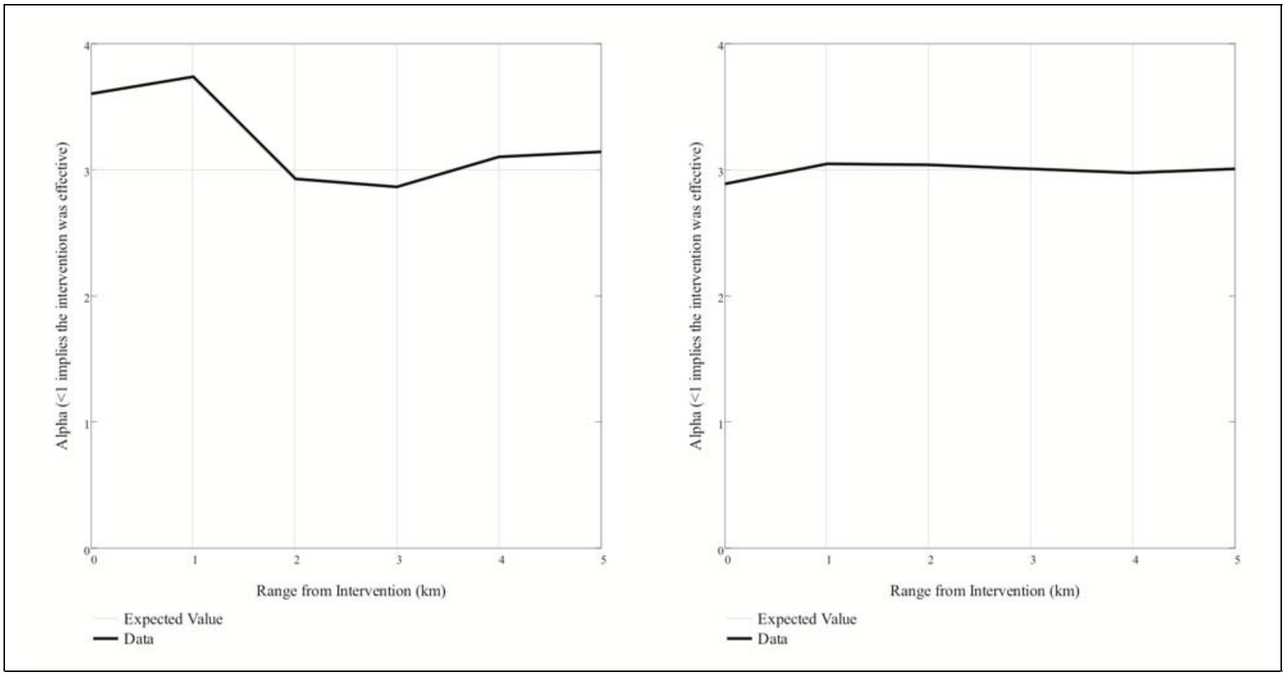


Figure 9. Algorithm Test Using Temporally Biased Data.   Randomly distributed spatial events and temporal events biased so that A = 3B. Left frame uses 10^4 events; right frame uses 10^6 events.

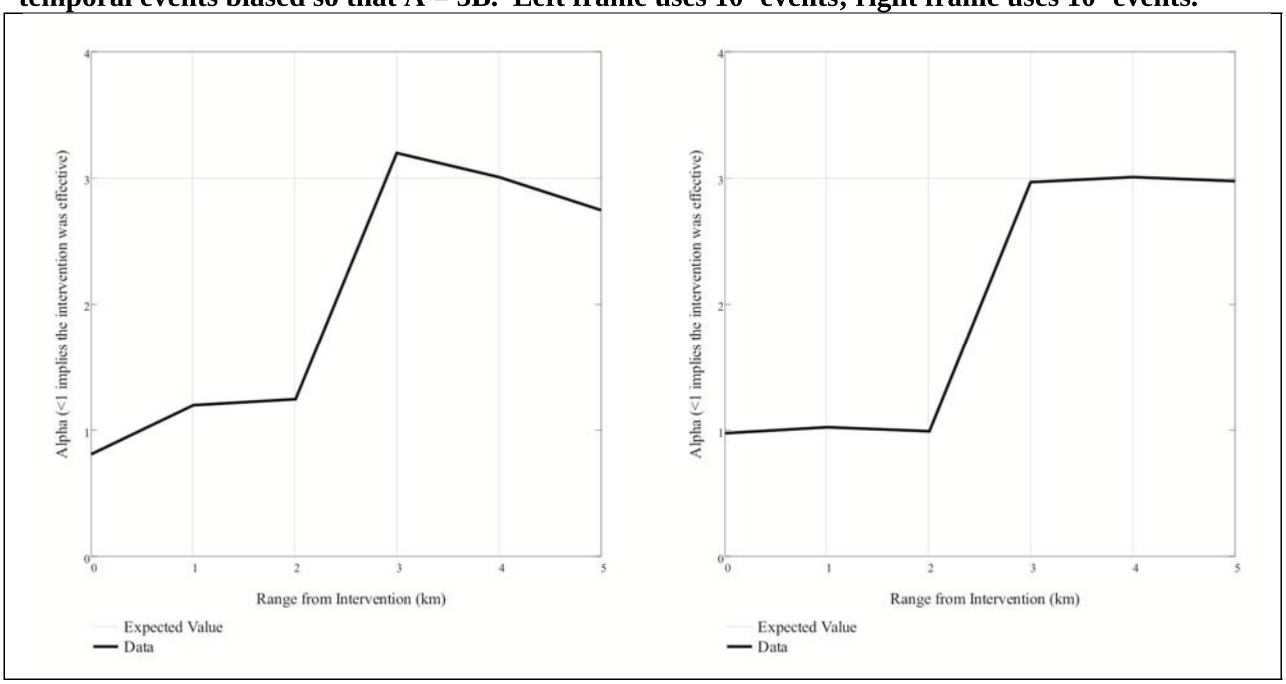


Figure 10. Algorithm Test Using Temporally and Spatially Biased Data. Spatial events biased so that events between 0 and 3 km. had A=B and events occurring greater than 3 km. from the intervention had A = 3B. Left frame uses 10^4 events; right frame uses 10^6 events.

Finally, a test data set that consisted of two distinct parts was tested with the algorithm. The first part was similar to that shown in Figure 8, while the second portion was similar to that shown in Figure 9, except that this portion was biased so that it only could occur at ranges greater than 3 kilometers from the interventions. Figure 10 shows the results of this test, and again these results match expectations of the model.

The model was also tested to ensure that the order in which the data points were read did not affect the results. Additionally, the model was checked to ensure that the maximum ranges and dates used did not affect the results. Other than slight vertical shifting that was expected when the maximum range was changed, shifting due to the slight variations about unity of the normalization constants, there were no changes in the model output.

Based on these tests of the algorithm we will assume the validity of the algorithm and now will apply it to the actual data.

4. Results

General Discussion of the Results

The algorithm was applied to the actual data, with results being shown in Figure 11 and Figure 12. These figures present a series of α vs. range plots as have been shown previously. Recall that an event is either an explosion or a found and cleared device. Also recall that α represents the fraction A/B, or the number of events occurring after the intervention divided by the number of events occurring before the intervention. It can also be interpreted as the fractional change in events following the intervention. In each case, the left frame of the pair is the result for Intervention Type I and the right frame is for Intervention Type II. Due to the binning of results, the range axis shows the *start* of each one-kilometer-wide bin; thus, the range labeled zero actually covers all events that occurred between zero and one kilometer from the intervention. The number of days following the intervention is shown on each figure pair. Confidence intervals representing 75%, 90%, and 95% levels are shown on the figures as well. These curves were based upon 100 iterations of the bootstrapping routine, which is exceedingly sufficient to get reasonable results [3]. Unless noted elsewhere, we shall only discuss the data curve and the 95th percentile curves for the remainder of this paper.

Note that the series of plots in Figure 11 and Figure 12 represents results using the <u>cumulative binning</u> scheme employed in [1] and [2] and illustrated in the left frame of Figure 6. From the figures it is apparent that as the number of days following the intervention increase, the width of the confidence intervals decreases as well. This result is consistent with the binning scheme because with each added day more points are added to the analysis, allowing for a more accurate calculation of mean, standard deviation, and consequently a tighter confidence interval. We shall return to discussions concerning the width of the confidence interval when we discuss the other binning schemes later in this paper.

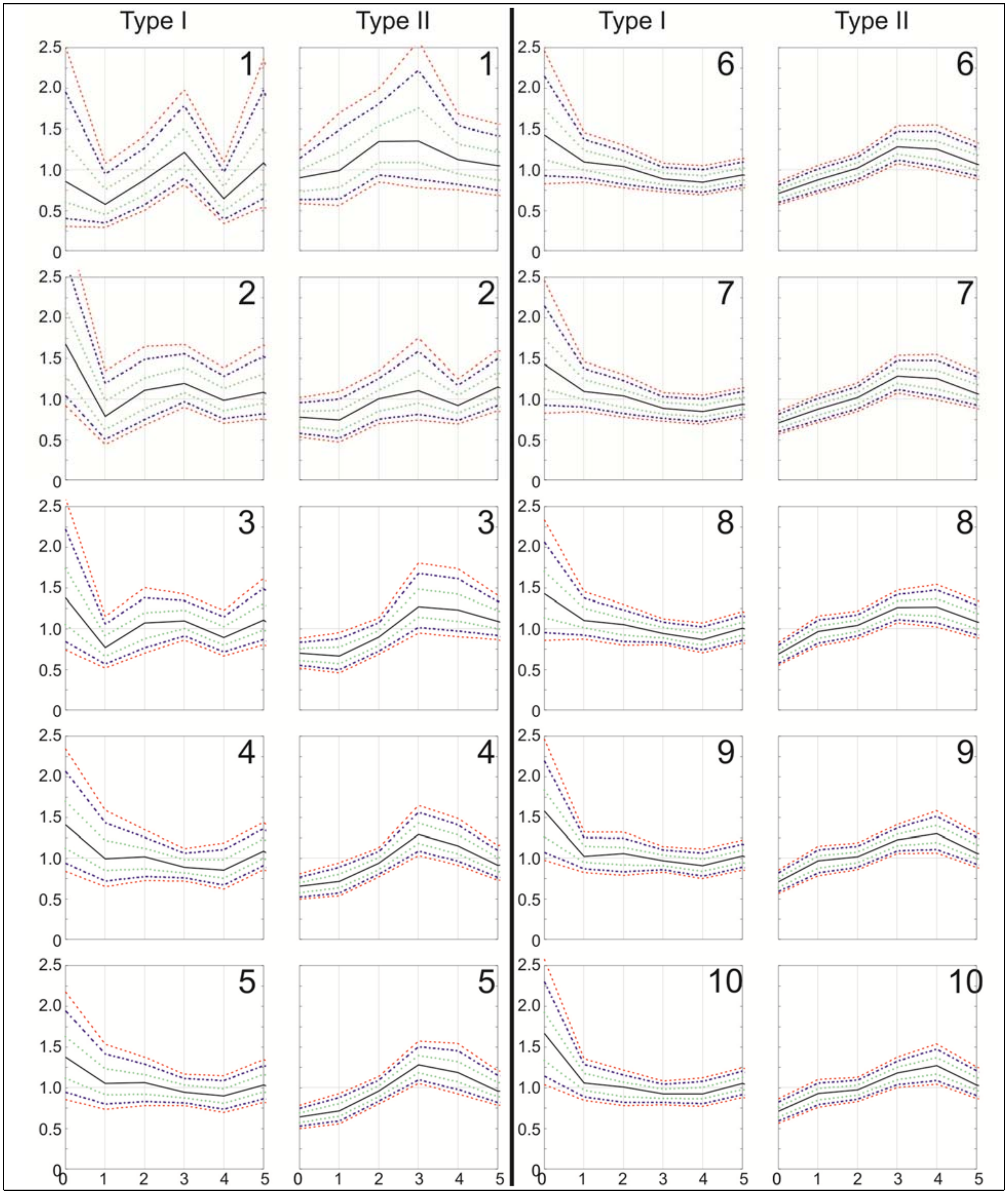


Figure 11. Results Using the <u>Cumulative Method</u>. Axes labels omitted on individual figures for clarity; vertical axis: percentage change in events following the intervention (dimensionless); horizontal axis: range from intervention (km). The number in upper right of each frame is the number of days following the intervention, binned one complete day at a time. For example, the number 3 indicates data from 2 to 3 days following the intervention. The black solid curve is the data. The green dotted, blue dash-dotted, and red dashed curves are the 75%, 90%, and 95% confidence intervals, respectively.

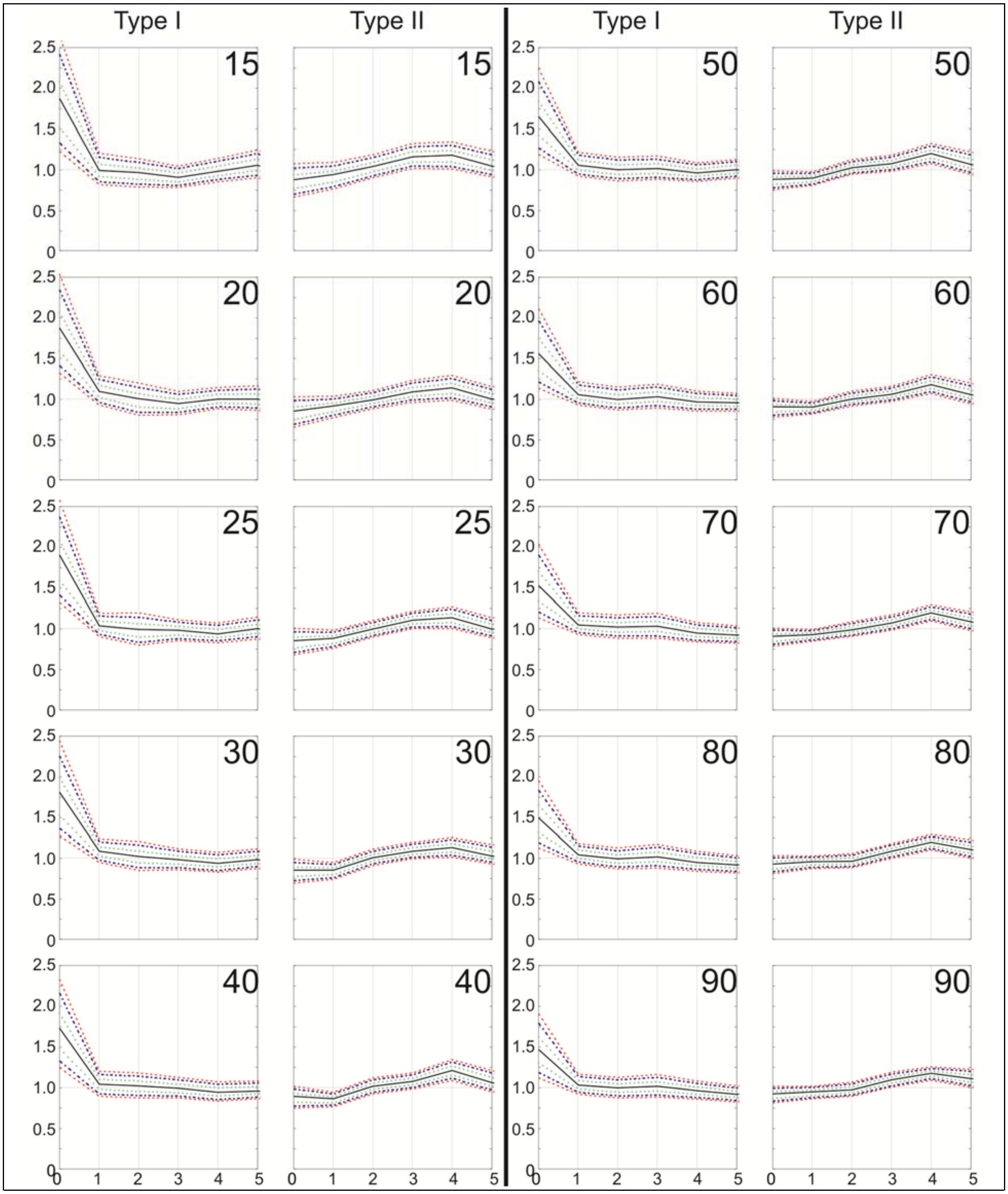


Figure 12. Results Using the <u>Cumulative Method</u> (continued). Axes and line descriptions as in the previous figure.

We make one final point prior to discussing the specific results: as could be logically expected and as was postulated in the earlier normalization scheme discussion, the effect drops off as range from the intervention increases, eventually becoming indistinguishable from the pre-intervention level. In each of these figures only ranges up to five kilometers from the incident are shown, although calculations were performed out to a range of 40 kilometers. In all cases, the 95% confidence interval was centered on and very close to a value of $\alpha = 1$, validating our assumption and our normalization scheme.

Recall that the normalization scheme not only assumed that α would asymptotically approach unity, but used the actual asymptotic level to determine a normalization constant that would force a value of unity. The normalization constants used to shift the results vertically so that they approach $\alpha = 1$ at great range are shown in Figure 13. For these results, the asymptote was determined by taking the average of the values of α at one-kilometer intervals at ranges from 30 to 40 kilometers from the intervention.

Referring back to Figure 1, in general the number of events is a decreasing function of time. Thus, when comparing before and after results one would expect to see a natural reduction irrespective of any effects specifically due to interventions, especially when looking at large numbers of days following the intervention. For example, when displaying a result using the cumulative binning scheme for a time period of 30 days, recall that the comparison is between data taken 30 days before the intervention with data taken 30 days after the intervention, for a total time lapse of 60 days. As is evident from Figure 1, the average slope of the number of events over such a period can be significant. The slope of the normalization constant plot is steeper for Intervention Type I because the overall decrease in events during the period of that campaign was greater.

As discussed previously, the method of Rivolo [2] accounts for this trend by forcing the large-

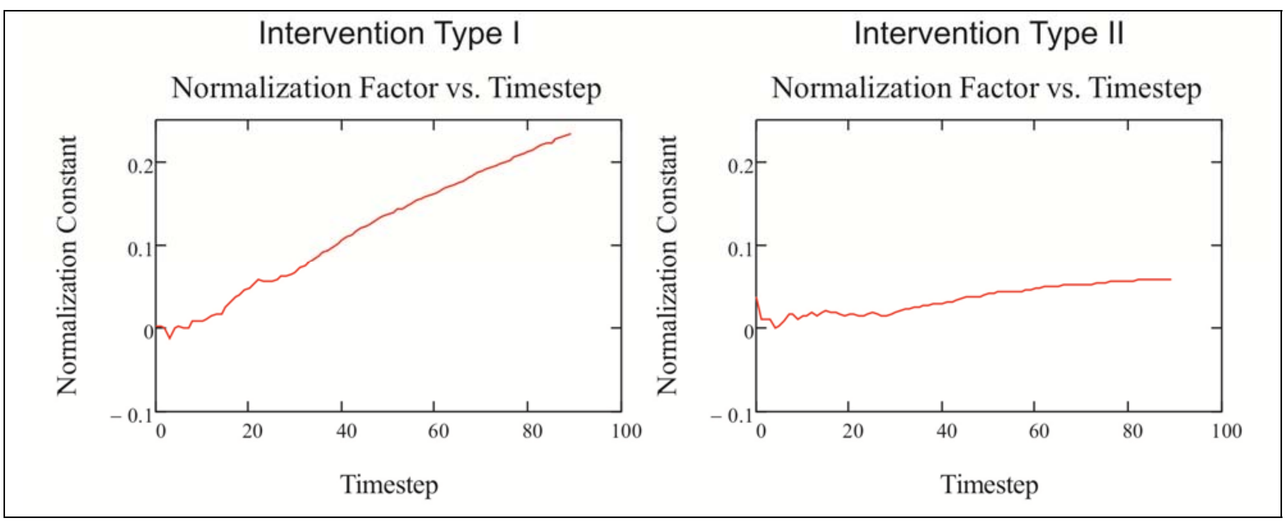


Figure 13. Normalization Constants for Both Types of Interventions, Cumulative Binning Method.

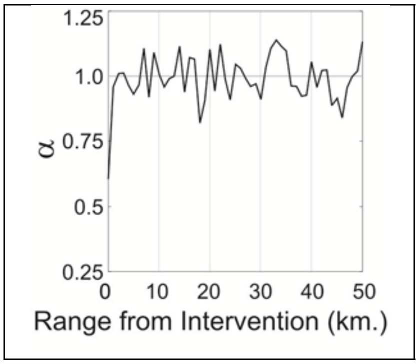


Figure 14. Example of the asymptotic approach of the results to unity at larger range. Statistical fluctuations about α = 1 are apparent.

range results to show no effect by adding a constant to all results for that data set. The generally

upward trend of the normalization constants reflects the fact that the daily rate of events was decreasing over the course of the time period considered in this data set; as things naturally get better overall, the asymptote would be further and further below unity as things got better and better. To compensate, progressively larger normalization constants would have to be added to shift the results back to a no-effect asymptote. Additionally, small time deltas would be expected to have less difference than larger time deltas. The normalization constants shown in Figure 13 show exactly these trends.

Cumulative Binning Scheme

With the generalities now taken care of, it is time to examine the results in detail. The frames showing the results one day post-intervention are essentially statistical variations about the no-change ($\alpha = 1$) line. We postulate that this lack of results shortly following an intervention is due to the finite time delay inherent in the bomb-planting system, where events already in the system continue to play out regardless of the intervention incident.

From days two to 90, however, the different intervention types have markedly different results. Intervention Type II appears to have the desired effect of reducing the level of events following the intervention. On day two, the reduction in events between zero and one kilometer is indicated to be about 25%, but could be anywhere between about 0% and 45%, based on the confidence interval. This effect remains roughly the same out to three kilometers, where it becomes statistically indistinguishable from $\alpha = 1$.

As can be seen in the figures, the reduction of events between zero and one kilometer from the incident stays relatively constant at about 25% to 30% for the first ten days, with the only real change being the previously mentioned tightening of the confidence intervals. By day six, the spatial extent of the effect, previously extending to about three kilometers, has reduced to only one kilometer from the intervention. By 15 days even the zero-to-one kilometer effect has diminished to about a 10% reduction, with the confidence interval indicating that it is quite possible that the effect has completely disappeared. For the remainder of the investigated time period, the effect of Intervention Type II is essentially negligible at all ranges.

Conversely, Intervention Type I appeared to have an effect contrary to that desired. Instead of inducing a reduction in the number of events, events actually increased following this type of intervention. Again, the data from the first day following the intervention is statistically indistinguishable from no-change. On day two, however, the intervention appears to have stirred up a hornet's nest, increasing the number of events somewhere between 0% and 300%. As the confidence intervals tighten, it can be seen that from days two onward, the number of events occurring between zero and one kilometer are either very slightly reduced or increased by factors of two to three until day nine. Subsequent days' results do not get any better, either, with the confidence interval clearly showing between a 25% and a 250% increase in events for two months following the intervention, an increase that only slightly diminishes by the end of the three-month time span over which the data was analyzed. Clearly, such a result is not what was desired.

Sliding Window Binning Scheme

As was discussed previously, the cumulative binning scheme has a flaw of biasing the results of larger time intervals toward those of earlier time intervals. To investigate the magnitude of this bias we shall now repeat the analysis using the sliding-window binning scheme, illustrated in frames 2 and 3 of Figure 6. Results obtained from the algorithm using the sliding window method are displayed in Figure 15. We have previously alluded to our suspicion that different days of the week could have higher or lower propensities for event occurrences. In this figure, a seven-day sliding window was used to guarantee that the same number of each day of the week was captured in both the before- and after-intervention bins. Seven days is the minimum number of days for which this criterion can be met. Seven days also tends to provide sufficient data to establish reasonably tight confidence intervals, as can be seen in the progression of day ranges in Figure 11, where the very wide confidence intervals in the first few frames tighten significantly by the seven-day frame pair.

Figure 15 is arranged exactly like Figure 12 for ease of comparison of the results using the two different methods. By examining Figure 6 it can be seen that for a seven-day window the cumulative and sliding window methods will produce exactly the same results for the first seven days of analysis. That was the case in this analysis. Days eight through ten, though not shown here, differed slightly from the cumulative method. Both of these examples provide further validation of the technique used to generate the results.

As can be seen when comparing the figures, the cumulative method displayed previously and used in [1] and [2] did tend to bias the results toward those obtained on lower days-after-intervention numbers. Using the cumulative method, for example, there is a measureable increase in the number of events following Type I Interventions all the way out to the end of the 90-day period of analysis, using the 95% confidence interval to determine the effect. Using the sliding window method, the increase in events disappears after about 50 days; the lower 95% confidence band drops below unity at that point and never goes back above that level. For Intervention Type II, we noted previously that using the cumulative method approximately a 10% reduction (and perhaps no reduction at the 95% level) was indicated at 15 days. Although the 80- and 90-day frames in the figure imply a reappearing harmful effect, examination of the

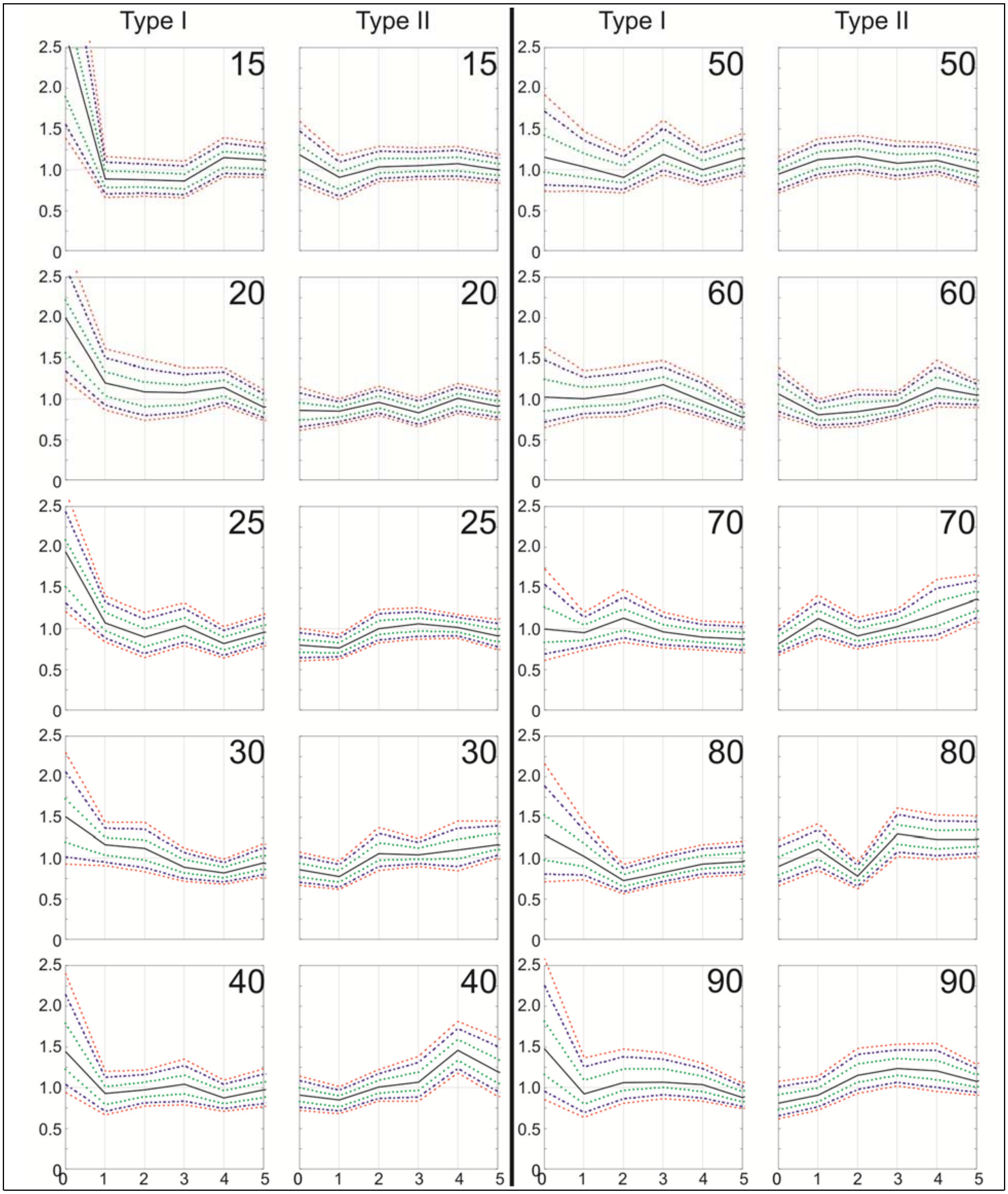


Figure 15. Results Using the <u>Sliding Window Method, 7-Day Window Width</u>. Axes and line descriptions as in previous figures.

frames from about day 50 on show this to be merely statistical fluctuation and not a real trend; other frames near those shown actually show decreases in activity. Using the sliding window method to remove bias, the 15-day result shows no effect and, although not shown, the effect essentially disappears after approximately 11 days.

In order to get a better feel for where the effects actually peter out, we present Figure 16. Notice that this figure is arranged differently than the other 20-pane figures in this paper; the left side relates to Intervention Type I and the right side relates to Intervention Type II. It uses a 4-day sliding window, the smallest width that would give reasonably-sized confidence intervals, as determined from visual inspection of Figure 11. The fact that either the before or the after bin would very likely have more weekend days could contribute to additional statistical fluctuations.

Selected days near the termination of the effects for both Intervention Type I and Intervention Type II are shown. In the left half of the figure, the behavior of the statistic for Type I Interventions shows, ignoring the random jumps in the curve due to statistical fluctuations, that while the case can be made that the increased level of activity due to the intervention remains up to about 42 days, by 44 days the effect has essentially vanished. Similarly, the right half of the figure shows the diminishment of activity extends to a range of about two kilometers through about six days, to one kilometer through about 10 days, and essentially vanishes after that time.

In summary of this binning method, it is clear that even with the wider confidence intervals and increased statistical fluctuations that result from the smaller number of available points for analysis, the results of both types of intervention showed markedly more persistent effects using the cumulative method than were present in the sliding window method. We believe these results clearly demonstrate the systematic bias in the cumulative method.

Average Counts Binning Scheme

While the sliding window binning scheme could offer a less biased analysis of the data, it still suffers from the fact that it compares data from widely separated time periods. For example, in both the previous methods time periods symmetrically spaced about the intervention are compared. When, as we have done, we specify results for 90 days following an intervention, the time period in question actually compares data taken almost six months apart. The relevance of the before-intervention data to the post-intervention data seem tenuous at best for these widely separated time periods.

One question warfighters are anticipated to ask is akin to "What have you done for me lately?" They may want to know what the reduction in events is compared to a fixed, recent time period. To answer this question we shall now repeat the analysis using the average counts binning scheme, as illustrated in the fourth frame of Figure 6. Using a 14-day comparison period, we will be able to answer the question "What was the percentage change in events at a certain number of days following the intervention, as compared to the previous two weeks?"

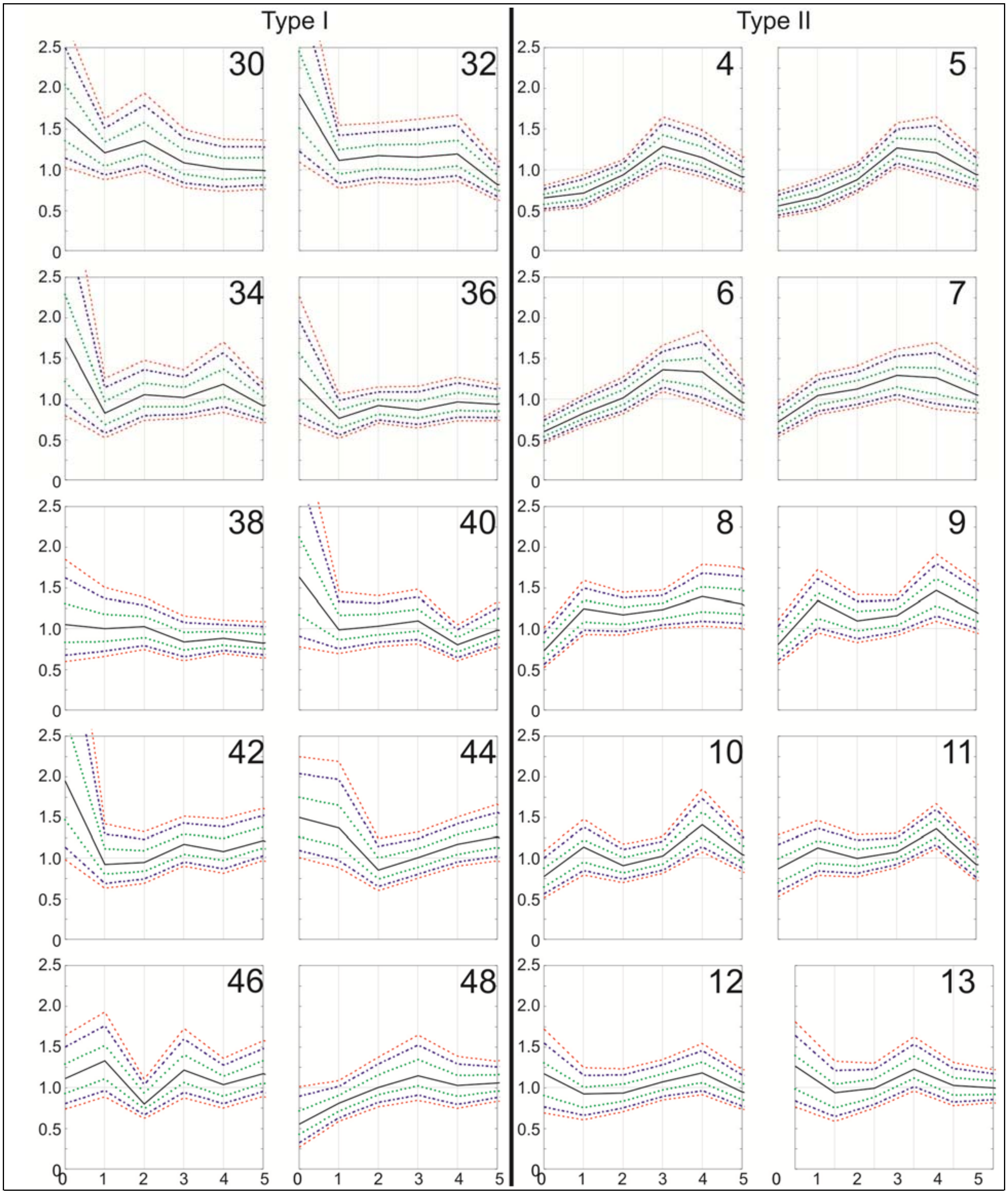


Figure 16. Results Using the <u>Sliding Window Method, 4-Day Window Width</u>. Axes and line descriptions as in previous figures.

In order to make the statistic deliver comparable numbers, we will have to weight the before and after numbers appropriately. Referring to Equation 1, it is seen that for a no-change result to yield a value of unity, A and B must be equal. When finding the average number of events in the past two weeks, for example, we will come up with the number of events per day. Should the window used for the accumulation of results after the analysis be three days, for example, then if there were no change our average, essentially using one day's worth of before-intervention data would be compared to our result, using three days' worth of after-intervention data, yielding a result of 1.5 instead of 1. To avoid this problem, we must multiply the before-intervention average by the width of the after-intervention data accumulation window. By using a sliding window of fixed width, we also avoid the biasing problems inherent in the cumulative method.

A final advantage to this method is that the denominator of the statistic calculated in Equation 1 is much more stable, with B being a constant instead of being subject to the variations encountered in both the sliding window and cumulative binning methods. Refer again to Figure 6 for help in visualizing how these variations could arise.

The results of the correlations using the average counts method are similar to the other two methods, but the persistences of the effects are different. As can be seen in Figure 17 and Figure 18, Intervention Type I still produced a significant undesirable increase in events but the effect essentially ends at 38 days to the 95% confidence level and 47 days to the 75% confidence level. Intervention Type II produced the desired effect, decreasing the number of total events by roughly 25% to a distance of 1-2 kilometers. This effect remains relatively constant for the entire 90-day period of the analysis.

We believe the biasing and irrelevant data problems discussed above make the cumulative and the sliding window methods less attractive than the average counts method, and shall use the average counts method as our standard for the remainder of this paper.

The Carelessness Factor

While the analysis technique developed in [1] and [2] and furthered in this paper is robust, there is one other way to tease information about the efficacy of the interventions from the data set. Events in the data set are clearly delineated as explosions and found/cleared. It may be reasonably supposed that in addition to suppressing events following an intervention, another result could be that the bomb making/planting system becomes careless due to fear of subsequent interventions. One indication of such carelessness could be that the ratio of found/cleared events to explosion events increases. Such an indication could be the result of bomb-makers producing more duds or bomb-planters doing a poor job of camouflage in their increased haste to avoid detection and reprisal. We shall define this ratio as

$$\beta = \frac{\text{Number of Explosion Events}}{\text{Number of Found/Cleared Events}} \tag{2}$$

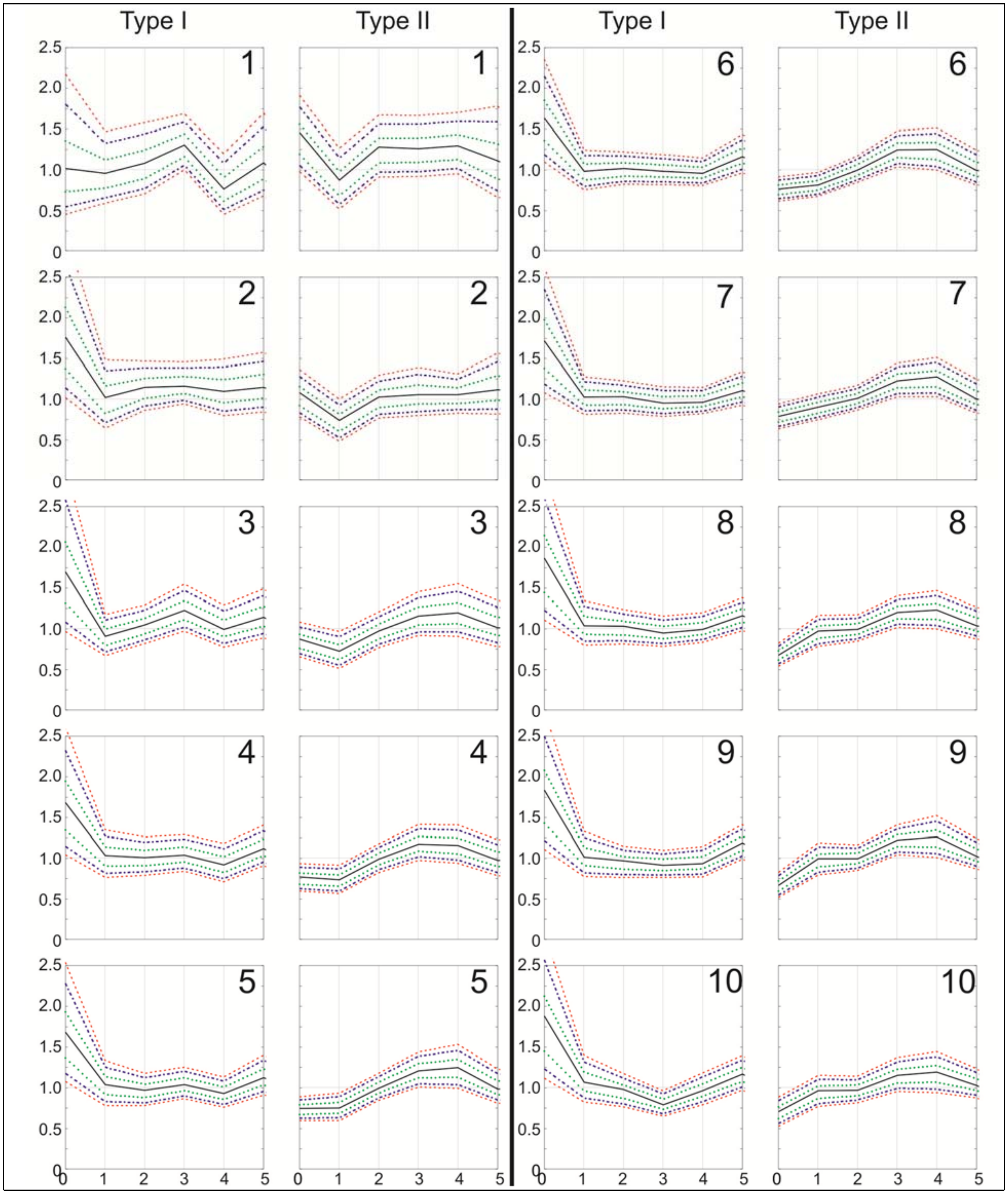


Figure 17. Results Using the Average Counts Method. Window width: 7 days; comparison: average counts over the two weeks prior to the intervention. Axes and line descriptions as in previous figures.

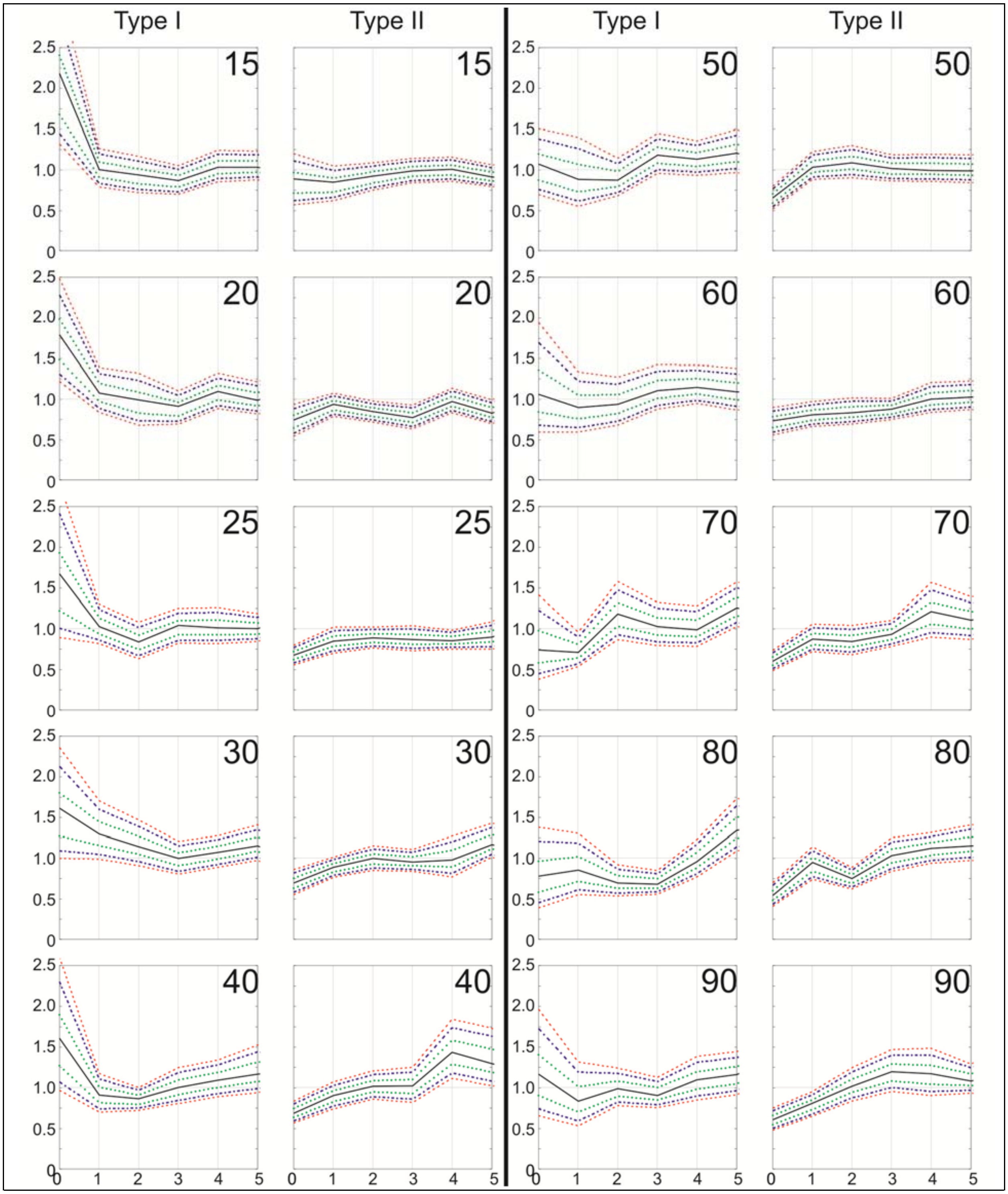


Figure 18. Results Using the <u>Average Counts Method (continued).</u> Window width: 7 days; comparison: average counts over the two weeks prior to the intervention. Axes and line descriptions as in previous figures.

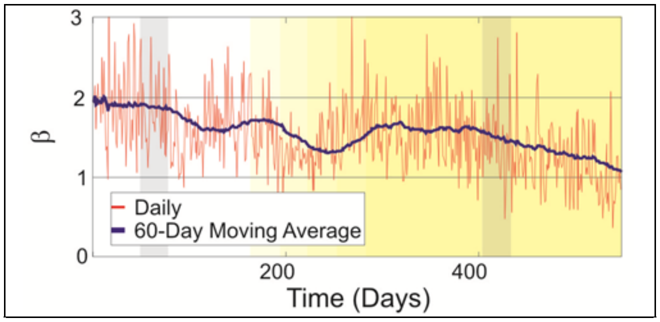


Figure 19. β During the Period of Study. The moving average is calculated using the greater of 60 days or the number of days from the start of the period of study. Grey vertical bars indicate Ramadan. The yellow region indicates the American troop surge.

During the period of study, β showed a significant decrease, as displayed in Figure 19. At the start of the period there were approximately twice as many explosive events as found/cleared events, a ratio that decreased almost to parity at the end of the period. There are several reasons that could cause this decrease in β, including more effective search procedures and increased carelessness on the part of the enemy.

Examination of the distribution of the β by day of week, displayed in Figure 20 is also illustrative. Notice that the ratio is roughly constant for all but two days of the week. In fact, the standard deviation of the ratio excluding Tuesdays and Fridays is only one half the value of the entire data set, indicating that the values for these two days are quite extreme. We propose one theory for these excursions. Some people avoid buying cars that were manufactured on Friday due to the perceived higher incidence of defects due to increasing worker carelessness as they approached the weekend. Recalling that the Muslim weekend begins on Thursday, and that a bomb planted on Thursday night would likely be completed on Wednesday, the pre-weekend carelessness factor could be an explanation for the higher incidence of failed bombing attempts on Friday. Also, referring back to Figure 4 and remembering that the Intervention Type II campaign was implemented on Mondays almost twice as frequently as the average of the remaining days of the week, bomb planters practicing their trade on Mondays after hearing of one of these interventions may have been rushed, trying to avoid another intervention, with their haste causing carelessness. Of course, these theories are purely speculative. There are other ways to analyze the carelessness factor with much more rigor.

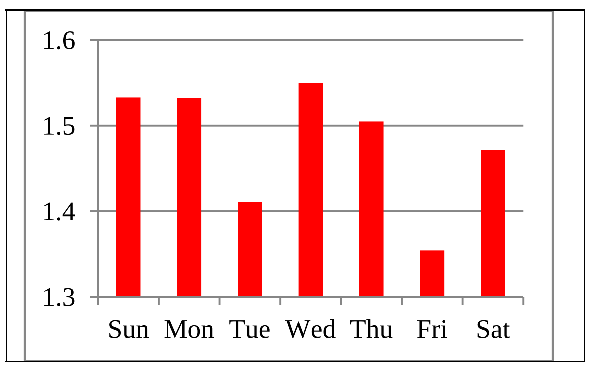


Figure 20. Found/Cleared to Explosions Ratio by Day of Week.

Of course, looking at the aggregate behavior of β says nothing about the effectiveness of the specific intervention campaigns on any sort of carelessness factor. To get this information we need to apply our algorithm separately to the explosion and found/cleared event types and determine what sort of information can be teased from those results. In Figure 21 and Figure 22 we present results using the average counts method for found/cleared events and explosions separately. As the number of found/cleared events was approximately half the number of

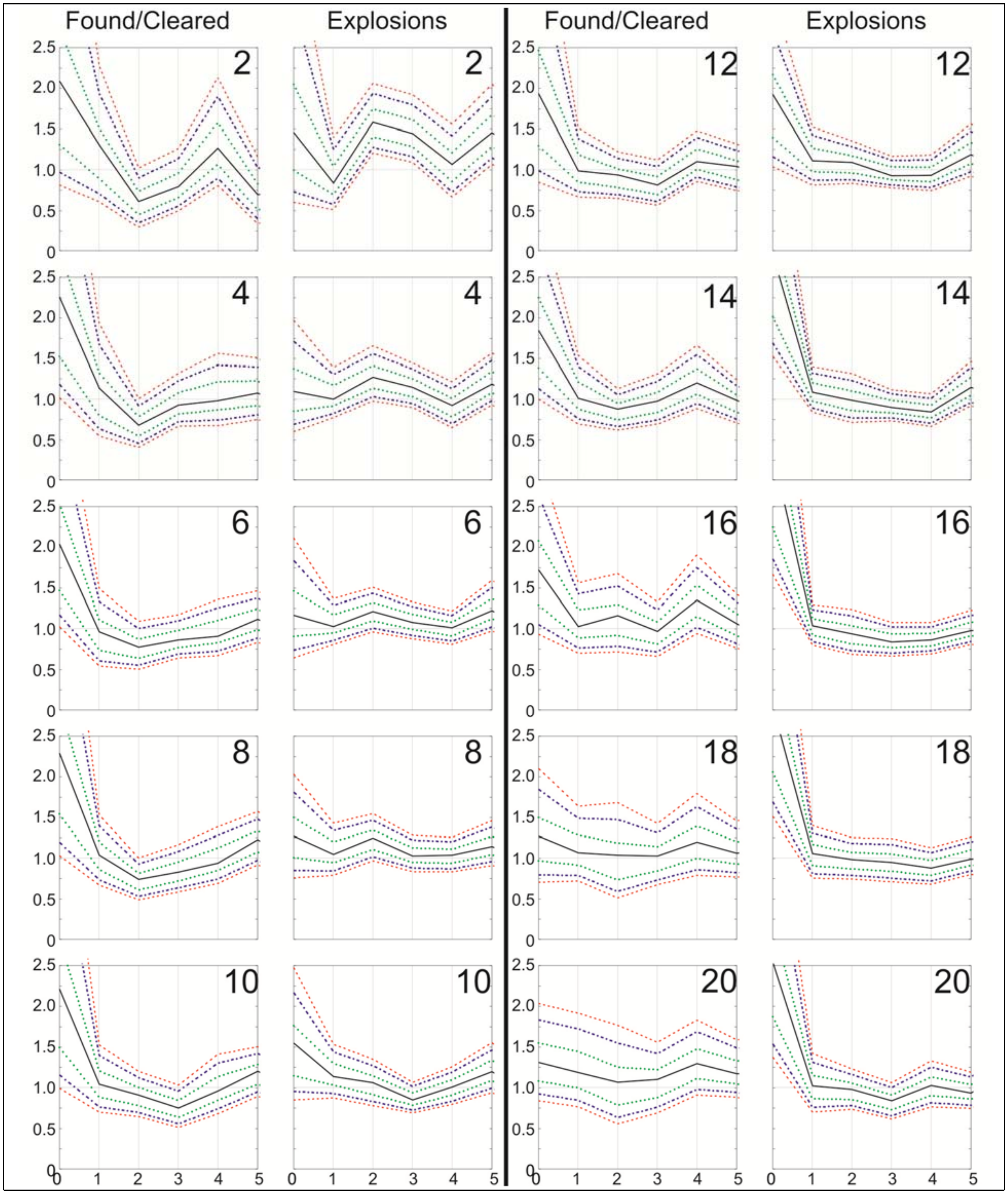


Figure 21. Comparison of Found/Cleared Events to Explosions Following <u>Intervention Type I</u>. Average Counts Method. Window width: 7 days; comparison: average counts over the two weeks prior to the intervention. Axes and line descriptions as in previous figures.

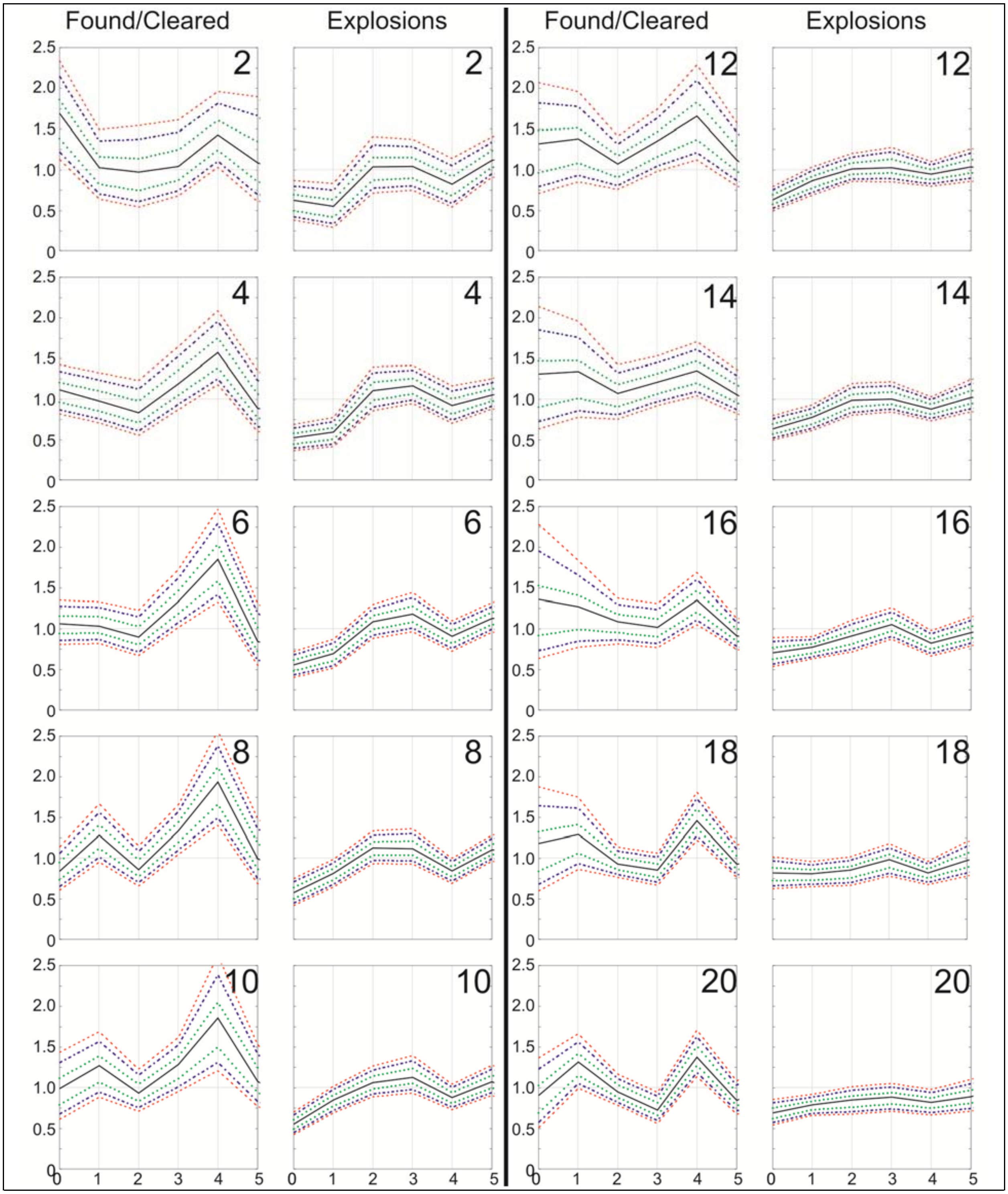


Figure 22. Comparison of Found/Cleared Events to Explosions Following <u>Intervention Type II</u>. Average Counts Method. Window width: 7 days; comparison: average counts over the two weeks prior to the intervention. Axes and line descriptions as in previous figures.

explosive events, the confidence intervals are significantly wider for the found/cleared results. The first figure shows the effects following Type I interventions and the second figure for Type II interventions. Recall from Figure 17 and Figure 18 that the effect of Type I interventions was to increase the number of total events within a few kilometers almost immediately. The breakdown of total events into found/cleared and explosions categories sheds additional light onto what is actually happening. Notice that it is actually the found/cleared events that cause the initial rise in total events. For the first ten or so days, the increase in the number of explosions is statistically indistinguishable from zero. It is only after that point that explosions dominate the increase in event occurrences, with the number of found/cleared events becoming less and less significant.

A completely different result is obtained from Type II Interventions. Here, the number of found/cleared events increases slightly for no more than two days following the intervention, and then returns to essentially no effect. The number of explosions, however, does drop down to about the 50% of pre-intervention levels, persisting with a drop of between 25% and 50% for the entire period shown in the figure.

One explanation of these results is increased carelessness. The Type I intervention causes an increase in the enemy's activity, but the activities are either so poorly planned or poorly executed that their devices are cleared before they can explode. After some time, though, the enemy becomes less shaken and their attacks become more effective, as measured by the decreasing ratio of found/cleared events to explosive events.

While it would be possible to apply a simple ratio method similar to our development of β, say

$$\gamma = \frac{\alpha_{\text{found cleared}}}{\alpha_{\text{explosions}}}, \tag{3}$$

to the results shown in Figure 21 and Figure 22, these results would not necessarily display a definitive carelessness factor, a number that would intuitively tell us whether the enemy had become more careless. A quick glance at these two figures will demonstrate why. If we use the frames showing the results from days four and 20 from Figure 21 as our examples, we see that in the former γ for the leftmost point would be about 2.25/1 = 2.25 while for the latter it would be about 1.25/2.5 = 0.5. It seems apparent that using this methodology a value for the factor of greater than unity would indicate a more careless enemy. At the four-day point the number of found/cleared devices has increased significantly over what was experienced during the two-week period prior to the intervention while the number of explosions has remained roughly constant, at least very near the intervention location. The greater-than-unity result might be seen as an indication of increased carelessness. Similarly at the 20-day point the ratio is less than unity, perhaps indicating less carelessness.

This method falls apart, however, when the value of α for the explosions is less than unity, as is the case in the frame for the results for four days post-intervention in Figure 22. Here, γ would be about 1.1/0.5 = 2.2, a result that is quite similar to the first one calculated from the very different situation at the four-day point in Figure 21. In fact, the value of γ for almost the entire 90-day period of analysis for Intervention Type II is greater than unity, even though it is readily apparent that instead of being careless enemy activity is actually being suppressed at short ranges. We need to develop a better measure of carelessness.

To determine this measure of carelessness, we begin by defining a careless situation as belonging to one of three cases: 1) the total number of events increases but the number of explosions do not; 2) the number of found/cleared events increases but the number of explosions do not; and 3) the fractional increase in found/cleared events is greater than the fractional increase in explosion events, where the fractional increase is defined as (events after intervention)/(events before intervention). We realize that there may be significant overlaps between these three cases. To determine whether there are patterns in the α values for found/cleared events and explosion events, we varied α for both of these event types between zero and two (representative of most of the results shown in Figure 21 and Figure 22 and varied β between one and two (representative of the values shown in Figure 19). We then computed the number of found/cleared, explosions, and total events that would occur under each of the combinations of α and β. If we define

$$\delta = \alpha_{\text{found cleared}} - \alpha_{\text{explosions}}, \qquad (4)$$

it becomes apparent that carelessness as defined above occurs every time that $\delta > 0$, a result that was essentially independent of β.

We then applied this equation to the results displayed in Figure 21 and Figure 22, calculating the confidence intervals by bootstrapping the results of equation (4). The carelessness factors for Intervention Type I and Type II are shown in Figure 23.

It is possible to interpret the results for Intervention Type I as showing a slight increase in carelessness out to a range of about one kilometer through about day 10. After that time the carelessness factor essentially drops to zero at all ranges until about day 16, where the carelessness factor drops below zero indicating an increased effectiveness of attacks. Although not shown, this increased enemy effectiveness effect disappears after about day 20.

The results for Intervention Type II clearly show an increase in carelessness to one to two kilometers through about 16 days after the intervention. After that point the carelessness factor drops to essentially zero where it remains for the rest of the period of analysis.

5. Potential Errors Due to Bad Data Points

Initial examination of the data set showed that a significant number of incidents were lacking geographic coordinate locations. The data set included many events and a single intervention for which the spatial coordinates were unknown. We shall refer to these data points for which the

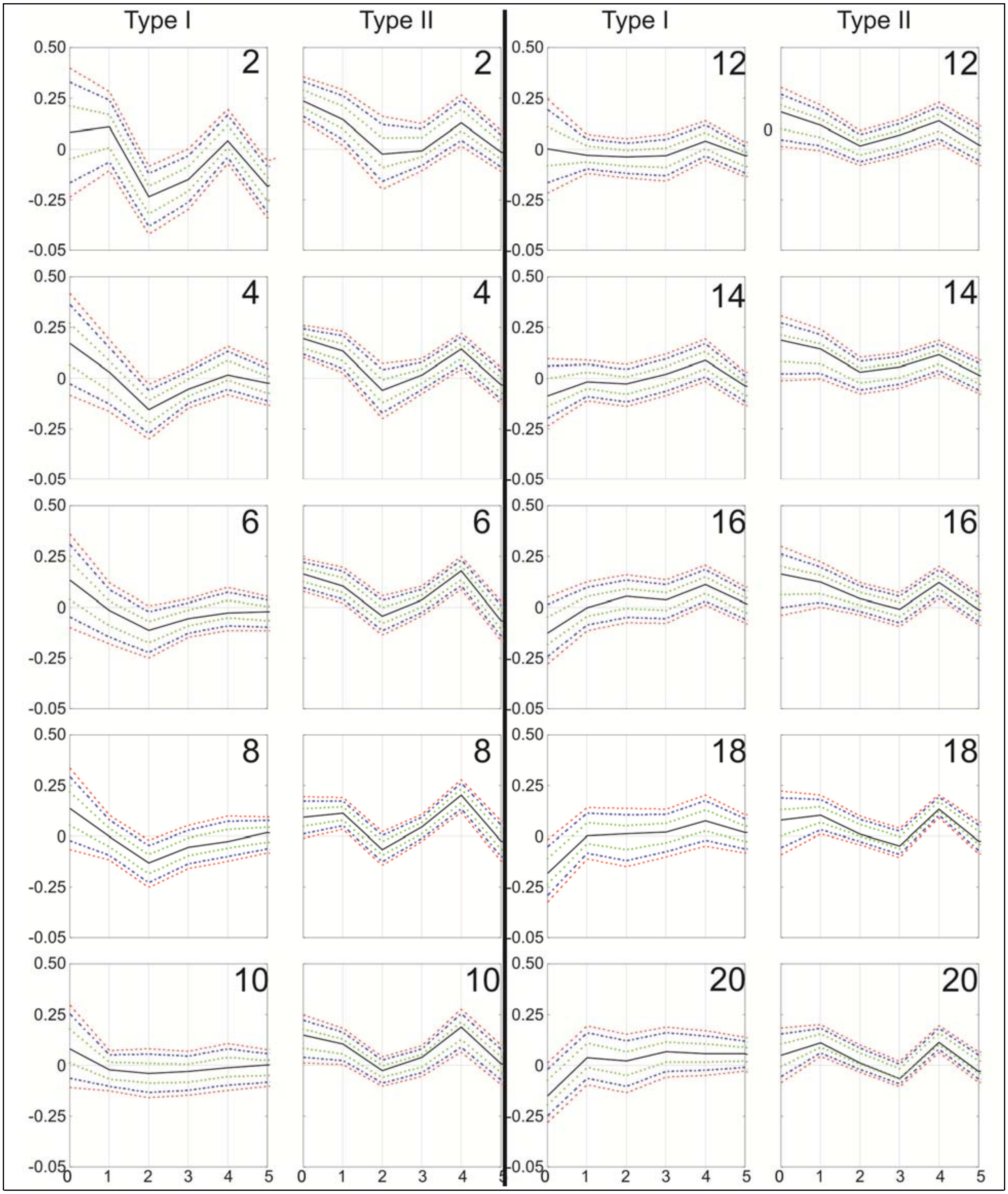


Figure 23. Carelessness Factor. Average Counts Method. Window width: 7 days; comparison: average counts over the two weeks prior to the intervention. Axes and line descriptions as in previous figures.

time was known but the location unknown as *zero-coordinate incidents*. Table 1 shows the number of zero-coordinate occurrences for each type of incident. In many data handling programs, blank fields are converted numerically to become zeros. Thus, these zero-coordinate incidents, were they accidentally included in the analysis, could potentially be a source of erroneous results due to their automatic passing of the maximum range test. They would only affect the result of the first range bin, the bin that contributes the most to the observed effect, as is clear from all of the figures showing results. We examined the effects of these zero-coordinate incidents to determine the magnitude of these potential errors.

Type Incident	Total in Data Set	Number of Zero-Coordinate Incidents
Intervention Type I	108	1
Intervention Type II	201	0
Explosions	29,999	343
Found/Cleared	20,284	61

Table 1. Summary of Zero-Coordinate Incidents.

Only the single zero-coordinate Intervention Type I and the approximately 400 zero-coordinate events actually interacted in this study, as the greatest range considered when computing the statistic was on the order of 50 kilometers while the nearest data point not in this set was on the order of 1000 kilometers. distant.

Days from Intervention	Number Before Intervention (B)	Number After Intervention (A)	Difference (A-B)
0-10	13	14	1
10-20	15	23	8
20-30	31	37	6
30-40	45	58	13
40-50	54	67	13
50-60	60	75	15
60-70	69	85	16
70-80	87	90	3
80-90	100	100	0

Table 2. Counts of Before- and After-Intervention Zero-Coordinate Events.

Figure 24 shows the temporal proximity of these zero-coordinate incidents, with the left panel showing all events and the right panel showing only those events within 90 days of the intervention. Table 2 illustrates the preponderance of events occurring after the incident, a fact that would tend to bias the statistic at zero distance for Intervention Type I to be greater than one.

When the algorithm was applied to the entire data set including the zero-coordinate incidents, the results of Type II interventions were unchanged. The effect of erroneously including the zero-coordinate incidents is shown in Figure 25. The results of Type I at ranges other than the first bin were also unchanged, as expected. The primary effect of adding these extra data points was to decrease the width of the confidence interval for the zero-to-one-kilometer range bin, especially for those plots where the days following intervention were small. Some slight shifting of the first data point was also seen, again early in the analysis. The effect is not as pronounced for larger days-after-intervention due to the relatively small number of zero-coordinate events being overwhelmed by the ever-increasing number of events with actual geo-coordinates. Thus, it may be seen that the effect of erroneously including the zero-coordinate incidents is negligible.

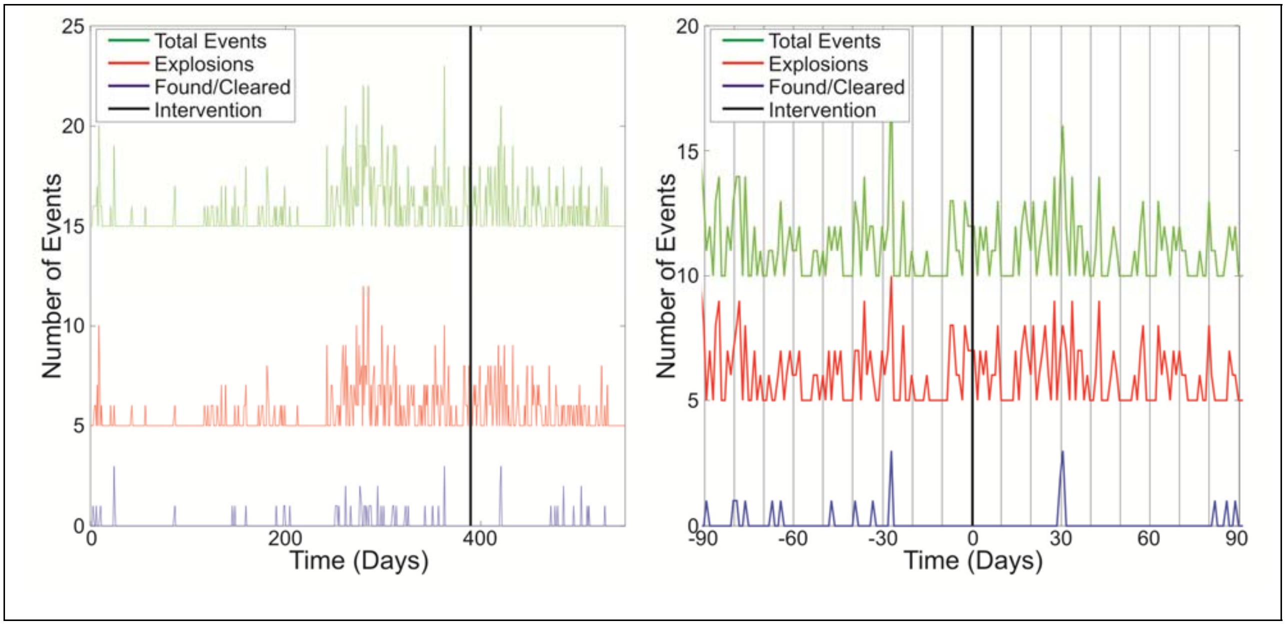


Figure 24. Time Phasing of Zero-Coordinate Events. Data sets shifted by multiples of five for clarity.

6. Summary

In this paper we investigated several methods of binning data: 1) the cumulative method (used in [1] and [2]), 2) the sliding window method and 3) average counts methods. We determined that the comparison of data from widely separated time periods in methods 1 and 2 as well as the biasing toward low-time data present in method 1 indicated that method 3 would give results more relevant to the warfighters who will need to use them.

We showed that in aggregate, Intervention Type II had the desired effect of reducing the number of events, particularly explosive events, by 25% to 50% up to two kilometers from the intervention for over 90 days. We also showed that in aggregate, Intervention Type I had the opposite of the desired effect. It actually increased enemy activity for a significant period following the intervention. The initial effect of this type of intervention was to increase found/cleared events. However, from about ten to 40 days following the intervention the number of explosive events increased somewhere between 125% and 250% for a distance of up to 2 kilometers from the intervention. Neither type of intervention appeared to have any long-range effect.

We also showed that a significant carelessness factor could be in effect. Careless activity, activity in which devices were planted but were planted or constructed so poorly that they were found and cleared before they were able to explode, increased for a short time following both types of intervention, being much more pronounced and in evidence over about twice the range for Intervention Type II than Intervention Type I. The carelessness effect lasted for about two

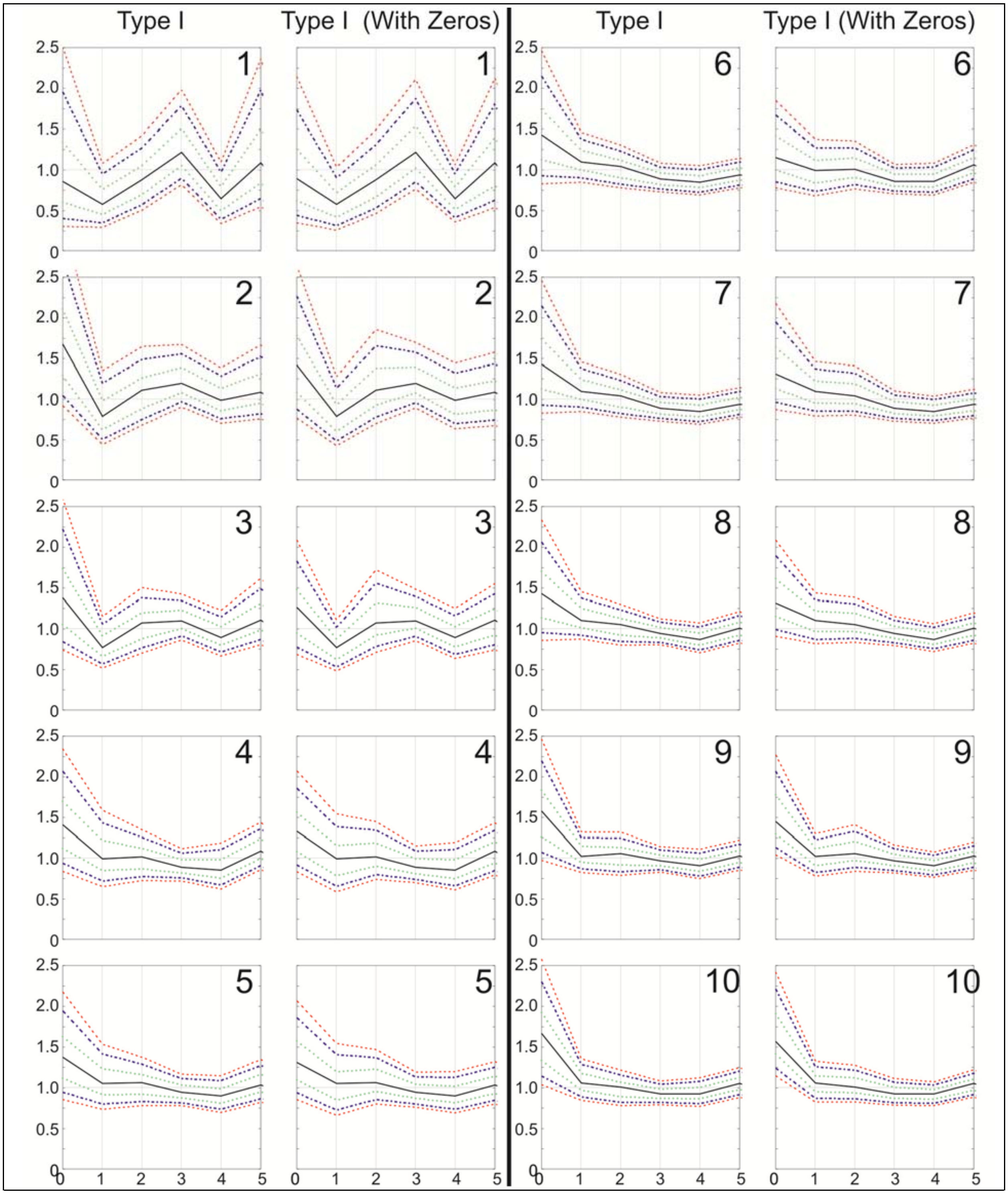


Figure 25. <u>Type I Intervention</u> Results Using the <u>Cumulative Method</u> and the <u>Zero-coordinate incidents</u>. Axes labels omitted on individual figures for clarity; vertical axis: percentage change in events following the intervention (dimensionless); horizontal axis: range from intervention (km). The number in upper right of each frame is the number of days following the intervention, binned one complete day at a time. For example, the number 3 indicates data from 2 to 3 days following the intervention. The black solid curve is the data. The green dotted, blue dash-dotted, and red dashed curves are the 75%, 90%, and 95% confidence intervals, respectively.

weeks in both cases.

The bottom line of this study was to highlight the fact that Intervention Type I is, at best, counterproductive and should be discontinued. Intervention Type II, however, is very effective in reducing enemy activity over short (~1-2 km.) ranges for several months.

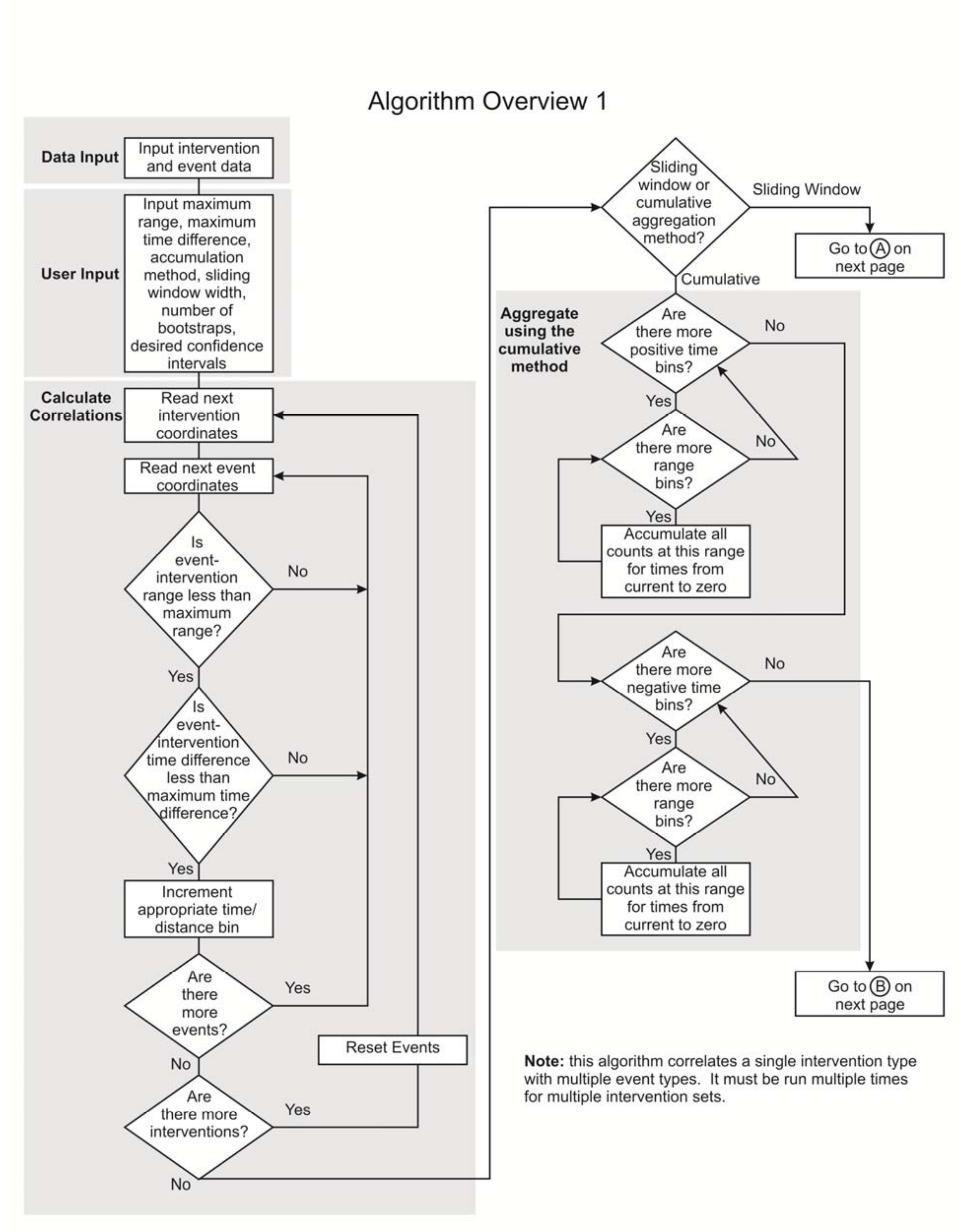
7. References

[1] Vijay Nair, A. Rex Rivolo, and Thomas Allen, "Spatio-Temporal Analysis of the Effectiveness of Interventions," Preprint, November 2008.

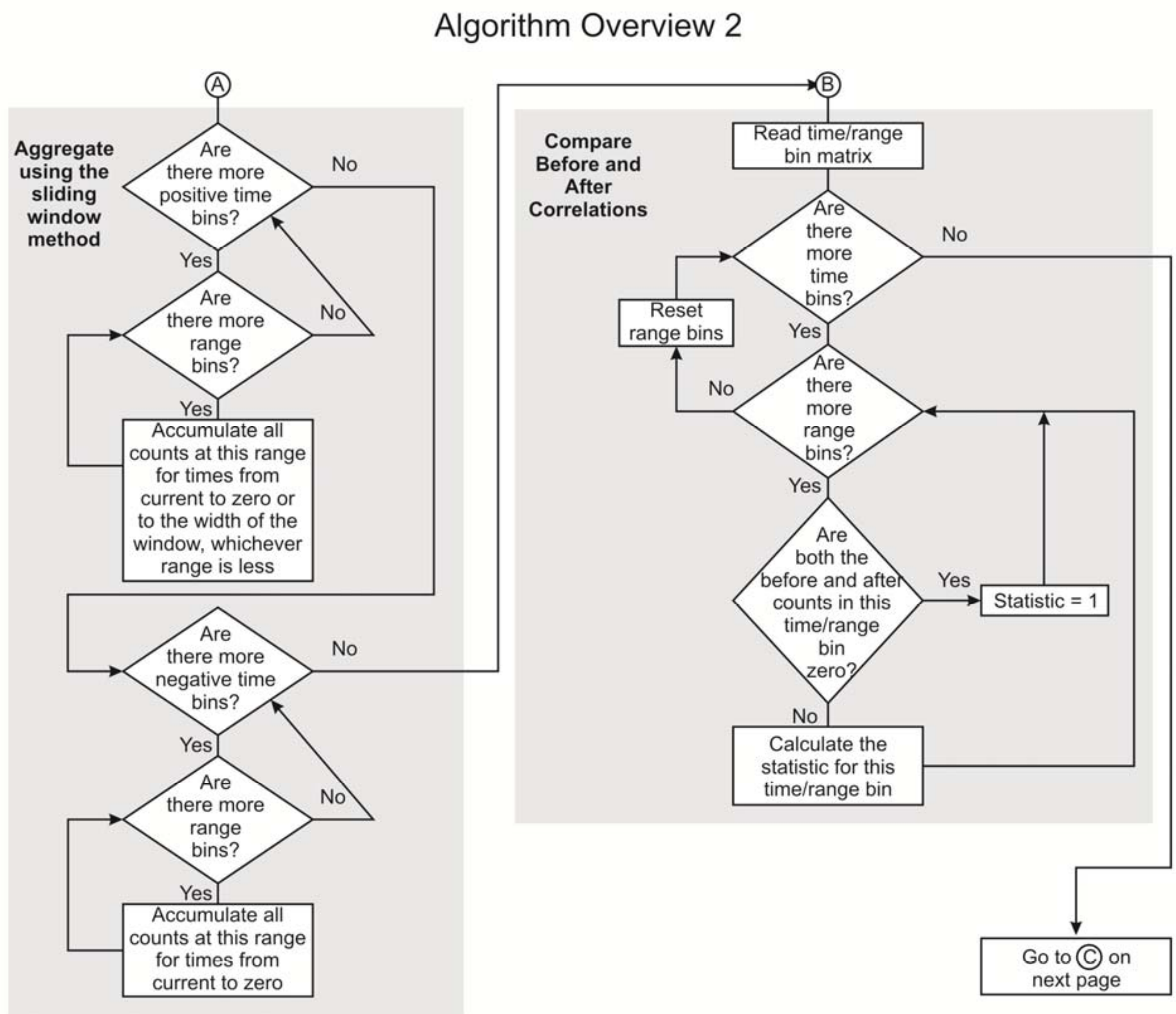
[2] A.R. Rivolo, "Assessing Operational Effectiveness Using Two-Point Space-Time Correlation Functions," Preprint, June 2008.

[3] W.H. Press, S.A. Teuklosky, W.T. Veterling, and B.P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing.* 2nd Edition. Cambridge: Cambridge University Press, 1992, pp. 691-703.

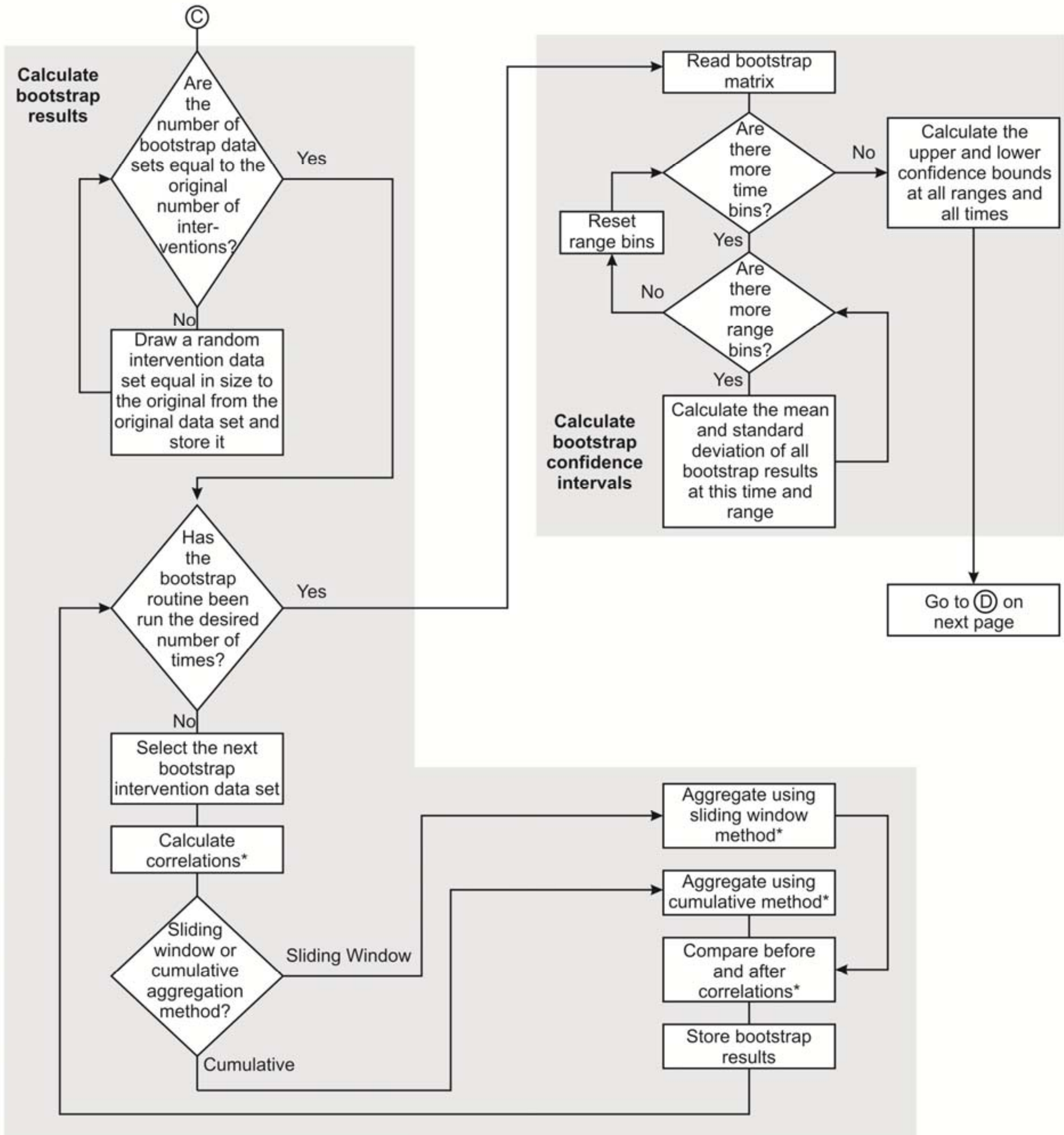
Appendix I.



Algorithm Overview 1
Data Input
Input intervention and event data
User Input
Input maximum range, maximum time difference, accumulation method, sliding window width, number of bootstraps, desired confidence intervals
Calculate Correlations
Read next intervention coordinates
Read next event coordinates
Is event-intervention range less than maximum range?
No
Yes
Is event-intervention time difference less than maximum time difference?
No
Yes
Increment appropriate time/distance bin
Are there more events?
Yes
No
Reset Events
Are there more interventions?
Yes
No
Sliding window or cumulative aggregation method?
Sliding Window
Go to Ⓐ on next page
Cumulative
Aggregate using the cumulative method
Are there more positive time bins?
No
Yes
Are there more range bins?
No
Yes
Accumulate all counts at this range for times from current to zero
Are there more negative time bins?
No
Yes
Are there more range bins?
No
Yes
Accumulate all counts at this range for times from current to zero
Go to Ⓑ on next page
Note: this algorithm correlates a single intervention type with multiple event types. It must be run multiple times for multiple intervention sets.


Algorithm Overview 2
A
Aggregate using the sliding window method
Are there more positive time bins?
No
Yes
Are there more range bins?
No
Yes
Accumulate all counts at this range for times from current to zero or to the width of the window, whichever range is less
Are there more negative time bins?
No
Yes
Are there more range bins?
No
Yes
Accumulate all counts at this range for times from current to zero
B
Compare Before and After Correlations
Read time/range bin matrix
Are there more time bins?
No
Reset range bins
Yes
Are there more range bins?
No
Yes
Are both the before and after counts in this time/range bin zero?
Yes
Statistic = 1
No
Calculate the statistic for this time/range bin
Go to C on next page

Algorithm Overview 3



* Refers to a previously named routine in this algorithm

Algorithm Overview 4