USING PRIOR INFORMATION IN DESIGNING
POINT IMPACT DETECTION NETWORKS

Peter Schumacher & James V. Zidek

Technical Report #120

SIAM INSTITUTE FOR MATHEMATICS AND SOCIETY

HEALTH SCIENCES PROGRAM
RAND CORPORATION

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY

DEPARTMENT OF STATISTICS
UNIVERSITY OF BRITISH COLUMBIA
Using Prior Information in Designing Point Impact Detection Networks
by
Peter Schumacher and
James V. Zidek
Technical Report No. 120

University of British Columbia
Department of Statistics

September 1988
Prepared in part under support to SIMS from the United States Environmental Protection Agency.
ABSTRACT

This paper investigates the effect of prior information on the design of monitoring networks for detecting the potential impact of an event which is to occur at a specified time. It is assumed that an $F$-test of space–event interaction is to be used to decide if an impact has occurred. Maximizing the power of this test, or rather the simpler, closely related goal of maximizing the noncentrality parameter is taken to be the designer’s objective. Some of the results obtained are qualitative. For example, for certain fairly realistic general models of how a subregional impact might distribute itself, it is shown that it is never optimal to place more than 50% of the monitoring sites in any one of the homogeneous subregions in which the impact might occur. Another qualitative, more intuitively obvious result is that it is essential to monitor subregions where the impact is not likely to occur (‘quasicontrols’); this would maximize the contrast created by the potential impact. A very general solution to the optimal design problem is given in a form which could be readily implemented in practice with the aid of a computer. Explicit solutions are also given for certain realistic impact models.

Keywords: Monitoring networks; optimal design; Bayesian designs; point impacts; environmental monitoring; assessing pollution.
USING PRIOR INFORMATION IN DESIGNING POINT IMPACT DETECTION NETWORKS

SUMMARY

Designing experiments invariably requires the use of prior information because the data the experiment is designed to produce is not yet in hand. For example, in the first phase of a major study of surface water in the United States, Linthurst and his co-investigators chose the population of lakes to be sampled as "lakes located within those regions expected to contain the most lakes in the U.S. characterized by alkalinity $< 400 \mu$ eq/L (i.e., those areas where acidic deposition would potentially have the most effect)". Because this prior information is usually introduced in an informal way, its influence on the selection of the ultimate design may well be obscured. This paper investigates this influence by introducing this prior information as an explicit datum through intuitively natural parameters in an objective function which expresses the goal of the design, taken here as the detection of a potential impact of an event occurring at a specified time. In this paper, it is supposed that the required design is to be balanced and the random measurement errors at the various sites before and after the potential impact are normally distributed and conditionally independent given the model parameters. The $F$-test of space-event interaction is a conventional method of looking for the potential effect and is adopted in this paper.

The methodology developed in this paper is potentially applicable in a variety of situations such as assessing the impact of opening a new power plant or shutting down an old one. Alternatively, the effects of a new abatement program might be the response of interest. The required before and after measurements could involve such things as air quality, water quality, or ecological attributes. An objective in
the situations envisaged in this paper, is the test of significance of the space-event interaction. It is clear that any region is susceptible to changes in ambient conditions between the times at which the before and after measurements are made. For this reason it is essential to have "pseudo-control" stations placed in subregions which are almost certainly not going to be affected by the potential impact in question.

In examples like those described above, the paucity of background data which may be at most baseline averages, and consequent inability of the designers to fit even spatial covariances, forces subjective choices to be made; these will involve educated "hunches" about the size and likelihood of impacts. In this paper, these hunches are explicitly expressed by an assumption that the impact field is random and has a joint probability distribution which summarizes the uncertainty in the prior opinions of the experimenters.

The power of the $F$-test at the root of this paper and developed for the general case where vectors rather than individual responses are being measured, depends upon the non-centrality parameter, which is a quadratic functional of the expected values of the differences between before and after measurements at the design stations. It contrasts the differences with the "norm"; that is, the average of such differences over all the design points. A marked departure of any difference from the norm indicates an impact at that station; the noncentrality parameter being zero means no impact occurred even though there may have been changes in the ambient level between the times at which the before and after measurements were taken. But even if there had been an impact, the noncentrality parameter might still be zero if the design points did not happen to include some of the stations affected by that impact. So it is
intuitively clear that to maximize the power of the test, the design points should be selected to maximize the (unknown) noncentrality parameter. Hence there is a need to resort to subjective strategies based on prior knowledge.

As noted earlier, in this paper uncertainty about the impact field is expressed by endowing it with a probability distribution. Thus the noncentrality parameter is a random variable dependent upon the choice of the design network. The power of the $F$-test is therefore random, so the test's actual power is the expected value of that of the $F$-test. The optimal design would maximize the expected power.

However, the combinatorial problem of finding the best subset of design points with an objective function as complex as that of the expected power is completely intractable. An intuitively similar and much simpler objective would be that of maximizing the expectation of the noncentrality parameter. Although the problem of maximizing this objective function would seem to be substantially simpler than that of maximizing the expected power itself, even this problem is well beyond the scope of modern methods except in very special cases.

So a further simplification is made to obtain a problem which is amenable to analysis. This simplification derives from the fact that in general, intuition about the potential area of impact will inevitably be somewhat vague. Clearly certain subregions will be seen to have a high risk of sustaining the impact, others a low risk. Overall, the region may be partitioned into a collection of homogeneous "zones". The boundaries of these zones will be somewhat vague and within the zones the stations will be more or less homogeneous. An individual's background knowledge will not be sufficiently fine as to enable the stations within each zone to be distinguished from their neighbours.
in that zone. A technical assumption based on this observation is imposed and this puts the problem in a tractable form. With this assumption there is a substantial simplification in the form of the noncentrality parameter which becomes a quadratic function in the sampling fractions of the numbers of the monitoring stations assigned to each zone. The problem of finding the optimal design now reduces to finding the optimal sampling fractions, which maximize the expected noncentrality parameter.

With the number of design points fixed, the reduction proposed in the last paragraph changes the original problem into a quadratic integer programming problem. While finding an exact solution to this problem remains difficult, our discretization of the original problem has an obvious approximate solution derived from treating the sampling fractions as continuous rather than discrete decision variables. The only constraints imposed in obtaining the approximate solution are that the sampling fractions be nonnegative and add up to one.

A particularly simple case may be treated in this approach, that in which the number of zones sustaining an impact is either zero or one. The possibility of placing design points outside of the homogeneous zones described above in a zone where no impact can occur, is permitted. A further simplification is that the impact will be uniform over the zone of impact. For this case, the approximately optimal sampling fractions may be found in an explicit form. Furthermore, in this case, the optimal integer programming solutions are derivable and the necessary algorithm is derived in this paper.

Unfortunately, the approximately optimal sampling fractions cannot be derived in the general case. But a number of properties of these sampling fractions are derived.
The results described above have shortcomings which are described in the paper and possible extensions are indicated. However, the results also show promise of being potentially valuable particularly in finding preliminary designs from which to work in a practical setting. And the results include some potentially interesting qualitative conclusions.

For example suppose there were only one potential impact zone which sustains a major impact with probability .95 and no impact with probability .05. How should the monitoring network stations be distributed? Because the overall change in ambient conditions is unknown it is clear that some stations must be placed outside the zone to provide baseline measurements against which those in the zone can be compared to discover whether there has been an impact. It may be surprising that, in fact, fifty percent of the stations should be placed outside.

Now consider a situation in which there are two zones (1 and 2) whose mutually exclusive impact probabilities are .2 and .7. Thus there is a 10% chance that neither zone will be hit but if 1 were to be hit, it would lead to a minor impact while for zone 2 it would be catastrophic. It is known that a third zone (0) will not sustain an impact. Again the problem is to decide how to distribute the monitoring stations. Obviously, there is no reason to place network stations in zone 0; since it is known that only one of zones 1 and 2 can be impacted, each serves as a pseudo-control for the other and hence baseline measurements from zone 0 are unnecessary. It does seem surprising however, that again, the optimal allocation is fifty percent of the stations to each of two monitored zones. In fact, it is always true for this sort of model that if there are exactly two potential impact zones that fifty percent of the monitoring
stations should be allocated to each of the two zones.

If, however, there are at least three zones, then for the sort of model just described, at least three zones must be incorporated into the approximately optimal sampling design. But in general this need not be true. It is possible under certain spatial correlation structures to have only one pair of zones in the approximately optimal design.

One of the more surprising results in this paper is that in a fairly general range of design situations, the approximately optimal design fractions can never exceed one-half for any zone. These sampling fractions may approach 1/2 for a given zone when the potential impact for that zone is large. This may seem surprising since large impacts are easy to detect and it might be thought that less sampling effort should be expended for such a zone. Less surprising is the result that 1/2 is approached as well if the probability of impact for a zone approaches 1.
1. **INTRODUCTION.**

1.1 **Origins of the Problem.**

Designing experiments invariably requires the use of prior information because the data the experiment is designed to produce is not yet in hand. For example, in the first phase of a major study of surface water in the United States, Linthurst and his co-investigators chose the population of lakes to be sampled as "lakes located within those regions expected to contain the most lakes in the U.S. characterized by alkalinity less than 400 μ eq/L (i.e., those areas where acidic deposition would potentially have the most effect)" (Linthurst *et al* 1986, p.4). Because this prior information is usually introduced in an informal way, its influence on the selection of the ultimate design may well be obscured. This paper investigates this influence by introducing this prior information as an explicit datum through intuitively natural parameters in an objective function which expresses the goal of the design, taken here as the detection of a potential impact of an event occurring at a specified time. It is supposed that (i) the required design is to be balanced, (ii) the random measurement errors at the various sites before and after the potential impact are normally distributed and conditionally independent given the model parameters. The *F*-test of space–event interaction (in the terminology of Millard and Lettenmeier, 1986) is a conventional method of looking for the potential effect (c.f. Green, 1979) and is adopted in this paper.

The methodology developed in this paper is potentially applicable in a variety of situations such as assessing the impact of opening a new power plant or shutting down an old one. Alternatively, the effects of a new abatement program might be the
response of interest. The required before and after measurements could involve such things as air quality, water quality, or ecological attributes.

A recent example is the permanent closure in 1985 of a large copper smelter in Tacoma, Washington. Before and after measurements of the chemical composition of rain water were made to determine the impact of this large pollution source on downwind rainwater composition. The second author’s interest in this problem arose from a request to assist in developing a before–and–after mud sampling program when exploratory drilling commenced in Harrison Bay on the north slopes of Alaska. A preliminary version of the results discussed in this paper were used in developing the design recommended in the contractor’s report.

An objective in the situations envisaged in this paper, is the test of significance of the space–event interaction. It is clear that any region is susceptible to changes in ambient conditions between the times at which the before and after measurements are made. For this reason it is essential to have “pseudo–control” stations placed in subregions which are almost certainly not going to be affected by the potential impact in question. This fact is not always recognized and there are examples where all the monitoring is carried out in risky regions. The space–event interaction test fails because there is nothing against which to compare the changes in potentially affected subregions.

In examples like those described above, the paucity of background data which may be at most baseline averages, and consequent inability of the designers to fit even spatial covariances, forces subjective choices to be made; these will involve educated “hunches” about the size and likelihood of impacts. In this paper, these hunches are
explicitly expressed by an assumption that the impact field is random and has a joint probability distribution which summarizes the uncertainty in the prior opinions of the experimenters.

1.2 Optimality Criteria

The power of the \( F \)-test at the heart of this paper and developed in Section 3 for the general case where vectors rather than individual responses are being measured, depends upon the noncentrality parameter, \( \delta^2 \). It is easily seen that \( \delta^2 \) is a quadratic functional of the expected values of the differences between before and after measurements at the design stations. It contrasts the differences with the “norm”; that is, the average of such differences over all the design points. A marked departure of any difference from the norm indicates an impact at that station; \( \delta^2 = 0 \) means no impact occurred even though there may have been a change in the ambient level between the times at which the before and after measurements were taken. But even if there had been an impact, \( \delta^2 \) might still be zero if the design points did not happen to include some of the stations affected by that impact. So it is intuitively clear that to maximize the power of the test, the design points should be selected to maximize \( \delta^2 \). But this is unknown since the expected before and after levels of the measurements of interest are, in general, unknown and hence there is a need, in practice, to resort to subjective strategies based on prior knowledge.

As noted earlier, in this paper uncertainty about the impact field is expressed by endowing it with a probability distribution. Thus \( \delta^2 \) is a random variable dependent upon the choice of the design network. The power of the \( F \)-test is therefore random so the actual power of the test is the expected value of that of the \( F \)-test. The optimal
design would maximize the expected power.

1.2 Simplifying the Problem

However, the combinatorial problem of finding the best subset of design points with an objective function as complex as that of the expected power is completely intractable. An intuitively similar and much simpler objective would be that of maximizing the expectation of $\delta^2$, denoted by $E(\delta^2)$. If 1 and 0 are used respectively to represent stations which are included and excluded from the design network, a design can be specified by the vector of all these 1’s and 0’s. Then $E(\delta^2)$ is just a quadratic function of this design vector. Although the problem of maximizing this objective function would seem to be substantially simpler than that of maximizing the expected power itself, even this quadratic binary programming problem is well beyond the scope of modern programming methods except in very special cases like that of Section 2 of this paper (Professor M. Queyranne, personal communication).

So in this paper, a further simplification is made to obtain a problem which is amenable to mathematical programming methods. Our simplification derives from the fact that in general, intuition about the potential area of impact will inevitably be somewhat vague. Clearly certain subregions will be seen to have a high risk of sustaining the impact, others a low risk. Overall, the region may be partitioned into a collection of homogeneous "zones" labelled $1, \ldots, K$. The boundaries of these zones will be somewhat vague and within the zones the stations will be more or less homogeneous. If the zones are reasonably compact, an individual's background knowledge will not be sufficiently fine so as to enable the stations within each zone to be distinguished from their neighbours in that zone. Denote the expected differences
between before and after measurements as $Z_{ij}$, where $i$ denotes the zone and $j = 1, \ldots, m_i$ denotes the station within the zone. Our assumption of homogeneity of opinion within zones will be formally embodied in a property we will refer to as

**SECOND ORDER ZONEWISE EXCHANGEABILITY (SOZE):** $E(Z_{ij}) = \mu_i$, $E(Z_{ij}^2) = \beta_i$, $E(Z_{ij}Z_{ij'}) = \gamma_i$, for certain positive constants $\mu_i, \beta_i, \gamma_i$ and all $i, j$ not equal to $j'$ and at the same time $E(Z_{ij}Z_{kJ}) = \beta_{ik}$ for all values of these subscripts, with $i$ not equal to $k$. With this assumption there is a substantial simplication in the form of $E(\beta_i^2)$ which becomes a quadratic function in $f = (f_1, \ldots, f_K)$ where $f_i$ denotes the fraction of the total number, $d$, of design stations which are in zone $i$, for $i = 1, \ldots, K$. The problem of finding the optimal design now reduces to finding the optimal sampling fractions, $f$, which maximize the expected non-centrality parameter. It is assumed that the stations within zones will be distributed at random.

### 1.4 Approximately Optimal Designs

With the number of design points fixed, the reduction proposed in the last paragraph changes the original quadratic binary programming problem into a quadratic integer programming problem for which an algorithm is developed in the special case of Section 2. While finding an exact solution to this integer programming problem remains difficult in general, our discretization of the original problem has an obvious approximate solution derived from treating the sampling fractions as continuous rather than discrete decision variables. The only constraints imposed in obtaining the approximate solution are $f_i \geq 0$ and $\sum f_i = 1$. In general, it may be necessary to impose the additional constraints that $f_i \leq F_i$, and this is permitted in the algorithm presented in Section 3. However, in the environmental situation the number of poten-
tial sites available in any zone will usually be large relative to the number of stations to be included in the design, so \( F_i = 1 \) is a realistic bound and the one used for most of the mathematical analysis in Sections 2 and 3. The resulting design will be called the **APPROXIMATELY OPTIMAL DESIGN (AOD)**.

### 1.5 Summary

In Section 2 we consider a particularly simple case in which the number of zones sustaining an impact is either zero or one. The possibility of placing design points outside of the \( K \) zones in a zone labelled 0 where no impact can occur, is allowed. The sample fraction allocated to that zone is \( f_0 \). A further simplification is that if the impact is sustained in zone \( i \), the impact will be uniform over the zone. For this case, the approximately optimal sampling fractions, \( f_{oi} \), are found in an explicit form. Furthermore, in this case, the optimal integer programming solutions, \( f_{oi} \), are derivable and the necessary algorithm is derived in Section 2.

Unfortunately, the AOD sampling fractions cannot generally be found in an explicit form. However, in Section 3, they are characterized in certain special cases including an obvious extension of that of Section 2. As well a number of properties of these sampling fractions are presented Section 3. In general, finding these fractions in specific cases entails the use of standard quadratic programming algorithms, one of which is illustrated in examples given in that section.

The results derived in Sections 2 and 3 are discussed in Section 4. Some of the shortcomings of the present approach are noted and possible extensions are indicated.
1.6 Qualitative Implications

To conclude this section we will present situations which show some of the qualitative implications of our results. First, suppose there is only one potential impact zone which sustains a major impact with probability .95 and no impact with probability .05. How should the monitoring network stations be distributed? Because the overall change in ambient conditions is unknown it is clear that some stations must be placed outside the zone to provide baseline measurements against which those in the zone can be compared to discover whether there has been an impact. It may be surprising that, in fact, fifty percent of the stations should be placed outside.

Now consider a situation in which there are two zones (1 and 2) whose mutually exclusive impact probabilities are .2 and .7. Thus there is a 10% chance that neither zone will be hit but if 1 were to be hit, it would lead to a minor impact while for zone 2 it would be catastrophic. It is known that a third zone (0) will not sustain an impact. Again the problem is to decide how to distribute the monitoring stations. Obviously, there is no point to placing any network stations in zone 0; since it is known that only one of zones 1 and 2 can be impacted, each serves as a pseudo-control for the other and hence baseline measurements from zone 0 are unnecessary. It does seem surprising however, that again, the optimal allocation is fifty percent of the stations to each of two monitored zones, 1 and 2. In fact, it is always true for the model of Section 2 and its substantial generalization of Section 3, that if there are exactly two potential impact zones \((K = 2)\) then fifty percent of the monitoring stations should be allocated to each of the two zones.

If, however, \(K \geq 3\), then for the model of Section 2 and its generalization in Section
3, at least three zones must be incorporated into the approximately optimal sampling design. But in the general case considered in Section 3, this is no longer true. It is possible under certain spatial correlation structures to have only one "distinguished pair" of zones in the AOD.

For the model of Section 2, it is easy to see that the AOD sampling fractions never exceed one-half for any zone. This is shown to be true even for the exact integer programming solution to the problem. Perhaps one of the most surprising results of this paper is that the same result holds true for the AOD sampling fractions for the generalization in Section 3 of the model of Section 2. The AOD sampling fractions in Section 2 approach \( \frac{1}{2} \) for a given zone when the potential impact for that zone is large. This may seem surprising since large impacts are easy to detect and it might be thought that less sampling effort should be expended for such a zone. Less surprising is the result that \( \frac{1}{2} \) is approached as well if the probability of impact for a zone approaches 1.

2. SINGLE SUBREGION IMPACTS.

Here at most a single random subregion, \( I \), may be affected by the impact. Let \( p_i = P(I = i) \) denote the conditional probability it is zone \( i \), given that an impact has occurred, \( i = 1, \ldots, K \). Assume the impact is uniform over the zone where it occurs and that for the generic station representing zone \( i \), \( Z_i \) is \( \Delta_I + \Delta \) or \( \Delta \) according as \( i = I \) or \( i \neq I \). We adopt the convention of letting \( \Delta_I = 0 \) when no impact occurs. It is convenient to designate the complement of zones 1, \ldots, \( K \) as zone 0, a zone where it is certain that no impact can occur, so that \( Z_0 = \Delta \) with probability 1.

Here \( \Delta \) represents the overall change in the ambient level, ignoring the potential
impact, and it is unspecified (unknown). Although the $\Delta_i$'s need not be completely specified, we will require that $\Delta_i = a_i \Delta_i$ where the $a_i$'s at least, are specified by the designer as a basic input based, like the specification of the $p_i$'s, on prior experience. It will become clear that without loss of generality we may take $\Delta_i = 1$ even when it is unknown and thus, to avoid additional clutter, we will suppose the $\Delta_i$'s themselves are specified.

Let $D$ consist of the co-ordinates $(i, j)$ for zones $i$ and stations $j$ specifying the network design. Then the average change in level over $D$ is

$$\overline{Z}_D = \frac{1}{d} \sum_{(i,j) \in D} Z_{ij},$$

where $d$ is the total number of stations to be included in the network, including those in zone 0, if any. It is readily seen that $\overline{Z}_D = \Delta + f_I \Delta_I$. Thus $Z_{ij} - \overline{Z}_D = (1 - f_I) \Delta_I$ or $-f_I \Delta_I$ according as $i = I$ or $i \neq I$. Thus the noncentrality parameter, $\delta^2$, described in the Introduction is

$$\delta^2 = n \sum_{(i,j) \in D} (Z_{ij} - \overline{Z}_D)^2$$

$$= nd \{ f_I[(1 - f_I) \Delta_I]^2 + (1 - f_I)[-f_I \Delta_I]^2 \}$$

$$= nd \Delta_I^2 f_I (1 - f_I), \quad (1)$$

where $n$ is the number of replicate measurements made at each monitoring station before and after the time of the potential impact.

With the convention that $\Delta_I = 0$ in the absence of an impact, the expected value of the noncentrality parameter in equation (1) is

$$E(\delta^2) = nd P \sum_{i=1}^{K} p_i \Delta_i^2 f_i (1 - f_i), \quad (2)$$
where \( P \) is the probability an impact occurs, and the remaining factor on the right of (2) represents the conditional expectation of \( \delta^2 \) given that an impact has occurred.

2.1 An Approximately Optimal Design.

The approximately optimal design (AOD) fractions, \( \{f_{ai}\} \), will be found by maximizing the expression displayed in (2) subject only to the conditions \( f_{ai} \geq 0 \) for all \( i \) and \( \sum f_{ai} = 1 \). The first result is trivial to prove, and intuitively natural.

Theorem 1. If \( K = 1 \) the AOD fractions are \( f_{a0} = f_{a1} = \frac{1}{2} \).

Let

\[
H(t) = \sum_{i=1}^{K} p_i \Delta_i^2 f_i (1 - f_i).
\]

Then \( H \) is a strictly concave function over the convex set

\[
\mathcal{F} \triangleq \{ f : f_i \geq 0, \ i = 0, \ldots, K, \ \sum f_i = 1 \}
\]

as long as \( p_i \Delta_i^2 > 0, \ i = 1, \ldots, K \) as we are now assuming conditionally. Thus \( H \) has a unique global maximum, \( f_a \), over the range of feasible \( f \)'s and no local maxima other than \( f_a \).

Let \( \mathcal{J} \) denote the set of all labels, \( i \), for which \( f_i \geq 0 \) and \( f_i \leq 1 \) are "nonbinding" constraints. By this we mean that \( H(f_a) > H(f_a + h) \) when \( h = (h_0, \ldots, h_K) \), \( h_i = 0, \ , i \notin \mathcal{J}, \ \sum h_i = 0 \) and all the \( \{h_i\} \) are sufficiently small. In other words, \( \mathcal{J} \) represents those \( \{f_{ai}\} \) for which small increases or decreases will result in a decrease in \( H(f_a) \). The complement of \( \mathcal{J} \) in \( \{0, \ldots, K\} \) will represent those \( \{f_{ai}\} \) for which \( f_i \geq 0 \) or \( f_i \leq 1 \) are "binding". But since we can ignore the possibility that \( f_i = 1 \) for any \( i \), \( f_i \geq 0 \) is the only constraint in the problem which could potentially be binding. Note that \( f_{ai} = 0 \)
is possible even though \( f_i \geq 0 \) is not binding. Example 9 in Section 3 illustrates this point.

If \(|h|\) is small and terms of second order in the \( \{h_i\} \) are ignored, \( H(f, h) \approx H(f) + G(f, h) \), where \( G(f, h) = \sum p_i \Delta_i^2 h_i(1 - 2f_{ai}) \). It follows that for any such \( h \), \( 0 > G(f, h) \). Now choose \( h \) so that \(|h|\) is sufficiently small, \( \sum h_i = 0 \), and \( h_i = 0 \) for \( i \in \mathcal{J} \). Then \( 0 > G(f, h) \). But at the same time \( 0 > G(f, -h) = -G(f, h) \). Hence \( 0 = \sum h_i p_i \Delta_i^2 (1 - 2f_{ai}) \) for every such \( h \). It follows that \( p_i \Delta_i^2 (1 - 2f_{ai}) = \lambda \) for \( i \in \mathcal{J} \).

Now for any admissible \( h \) with \(|h|\) sufficiently small, \( G(f, h) = \lambda \sum h_i + \sum h_i p_i \Delta_i^2 = \sum h_i (p_i \Delta_i^2 - \lambda) - \lambda h_0 \) where \( \mathcal{J}_0 = \mathcal{J} \cap \{1, \ldots, K\} \). Thus since \( h_i \geq 0 \) for all \( i \in \mathcal{J} \), \( p_i \Delta_i^2 < \lambda \) for all \( i \in \mathcal{J} \). Now if \( f_{a0} > 0 \) we may choose \( h_i = 0, \ i \in \mathcal{J} \), and \( h_0 < 0 \) or \( h_0 > 0 \), providing \( h_0 \) is sufficiently small, while satisfying \( \sum h_i = 0 \). It follows that \( \lambda = 0 \), \( f_{ai} = \frac{1}{2} \) for \( i \in \mathcal{J} \), and \( p_i \Delta_i^2 < 0 \), \( i \in \mathcal{J} \). Clearly this is possible if and only if \( K = 1 \) and \( f_{a1} = f_{a0} \), the case covered in Theorem 1.

In summary, if \( K > 1 \), \( f \) is a local maximum of \( H \) only if \( f_{a0} = 0 \) and for some constant \( \lambda > 0 \), \( f_{ai} = \frac{1}{2}[1 - \lambda(p_i \Delta_i^2)^{-1}] \) whenever \( f_{ai} > 0 \), while \( p_i \Delta_i^2 < \lambda \) whenever \( f_{ai} = 0 \) and this is a binding constraint. But conversely, as indicated above, this condition is also sufficient to assure that \( f \) is a global maximum of \( H \). The existence of a maximum for \( H \) on \( \mathcal{F} \) assures therefore the existence of \( \lambda \) and the set \( \mathcal{J} \) defined above. We now derive \( f \) in explicit form for \( K > 1 \).

Arrange \( \{p_i \Delta_i^2\} \) in nondecreasing order of size. It follows that \( \mathcal{J} = \{0, 1, \ldots, m\} \) for some \( m \geq 0 \). Furthermore, \( 1 = \sum_{i=m+1}^{K} \frac{1}{2}[1 - \lambda(p_i \Delta_i^2)^{-1}] \) so \( \lambda = [K - m - 2][(\sum_{i=m+1}^{K} (p_i \Delta_i^2)^{-1})^{-1}]. \)

That such an integer \( m \) exists follows from considering

\[
L(m) \triangleq (K - m - 1)\tau(m) - \tau(m) - \cdots - \tau(K), \quad m = 1, \ldots, K,
\]
where $\tau(i) = (p_i \Delta_i^2)^{-1}$, for all $i$. Observe that $L(K) = -2\tau(K) < 0$. Suppose $L(m) \geq 0$.

Then

$$L(m - 1) - L(m) = [K - (m - 1) - 1]\tau(m - 1) - \tau(m - 1) - \cdots - \tau(K)$$

$$- [K - m - 1] \tau(m) + \tau(m) + \cdots + \tau(K)$$

$$= [K - m - 1][\tau(m - 1) - \tau(m)]$$

$$\geq 0.$$  

Thus $L$ has at most one sign change.

But

$$\lambda - p_{m+1} \Delta_{m+1}^2 \propto \lambda \tau(m + 1) - 1$$

$$\propto [K - (m + 1) - 1][\tau(m + 1) - \tau(m + 1) - \cdots - \tau(K)]$$

$$= L(m + 1),$$

for $m = 0, 1, \ldots, K - 1$. Thus either $\lambda < p_i \Delta_i^2$ for all $i$ and hence $m = 0$, or there is a unique $m$ for which $p_m \Delta_m^2 < \lambda \leq p_{m+1} \Delta_{m+1}^2$ when $\lambda$ (a function of $m$) is as defined above. These results may be summarized in a Theorem.

**Theorem 2.** For $K > 1$, the sampling fractions which maximize the expected noncentrality parameter are:

$$f_{ai} = \begin{cases} 0, & i = 0, 1, \ldots, m \\ \frac{1}{2}[1 - \lambda(p_i \Delta_i^2)^{-1}], & i = m + 1, \ldots, K \end{cases}$$

where $\lambda = \lambda(m) = [K - m - 2][\sum_{i=m+1}^{K} (p_i \Delta_i^2)^{-1}]^{-1}$, and $m = 0$ unless a unique positive $m < K - 2$ can be found which satisfies $p_m \Delta_m^2 < \lambda(m) \leq p_{m+1} \Delta_{m+1}^2$.

There are a number of notable consequences of this theorem. These are given in the following corollaries whose proofs are immediate. The approximately optimal design (AOD), $D_a$, is that having the sampling fractions stated in Theorem 2.

**Corollary 3.** If $K = 2$, $f_{a1} = f_{a2} = \frac{1}{2}$. ☑
Corollary 4. For any $K \geq 3$, the number of zones in $D_a$ is $K - m \geq 3$.

Proof. The result is obvious since otherwise $\lambda(m) \leq 0$ for all $m$ and $p_m \Delta_m^2 < \lambda(m)$ cannot then be satisfied for any $m$, a contradiction. 

Corollary 5. If $K = 4$, it is approximately optimal to monitor just $K - m = 3$ zones if and only if

$$(p_1 \Delta_1^2)^{-1} > (p_2 \Delta_2^2)^{-1} + (p_3 \Delta_3^2)^{-1} + (p_4 \Delta_4^2)^{-1},$$

when zone 1 is excluded from $D_a$.

2.2 A Discrete Allocation Approach.

Thus far we have considered the design problem from the point of view of choosing continuous sampling fractions. Now let us consider the discrete analogue: allocating a fixed number of monitoring stations among the $K$ zones. We start with expression (2) of the preceding section, for the expected value of the noncentrality parameter:

$$E(\delta^2) = ndP \sum_{i=1}^{K} p_i \Delta_i^2 f_i (1 - f_i).$$

On suppressing the constants $n$ (number of replicates at each station) and $P$ (probability an impact occurs at all), and denoting the number of stations to place in zone $i$ by $n_i$, so that $f_i$ is replaced by $n_i/d$, this becomes

$$E(\delta^2) \propto \sum_{i=1}^{K} p_i \Delta_i^2 \frac{n_i (d - n_i)}{d}$$

or

$$E(\delta^2) \propto \Phi(n_1, \ldots, n_K) \equiv \sum_{i=1}^{K} \phi_i(n_i),$$

(3)

with $\phi_i(n_i) = \frac{p_i \Delta_i^2 n_i (d - n_i)}{d}$. The optimization problem is to maximize $\Phi$ over non-negative integers $n_i$ satisfying $\sum_{i=1}^{K} n_i = d$, where $d$ is specified in advance.
Before considering a general solution to this problem, let us first look at some trivial yet illustrative special cases. Dispensing with these now serves the additional purpose of allowing us to present the general case more compactly.

Suppose one had only a single station to allocate. It is intuitively obvious that there is no point in monitoring at all, since one is anticipating a test for interaction, and data on only one zone is useless. This is consistent with the result of applying the criterion in (3): \( E(\delta^2) = 0 \) regardless of which zone receives the station. Thus \( \delta^2 \) is zero, and consequently the power of the test cannot rise above its level, say \( \alpha \).

Next suppose that, besides zone zero, there are exactly two zones with \( p_i \Delta_i^2 \neq 0 \). (This subsumes the case \( K = 2 \).) Consistent with Corollary 3 of the last section, the optimal allocation here attempts to split the \( d \) stations as evenly as possible between the two zones, without regard to the relative magnitudes of the \( p_i \Delta_i^2 \). Specifically, if \( d \) is even, both zones receive \( \frac{d}{2} \) stations. Otherwise, \( \frac{d}{2} - \frac{1}{2} \) stations are first placed in each zone, and then the final station is allocated at random to either of the two. Interestingly enough, although the incremental contribution of this last station to the value of \( \Phi \) is zero, it so happens that \( \Phi \) is still strictly greater than it would be with an even allocation of \( d - 1 \) stations. (Recall that the total number of stations to be allocated is specified in advance; this number appears in the objective function \( \Phi \), as well as in the constraint.)

From here on, it is implicitly assumed that \( K \geq 3 \) and \( d \geq 2 \). For this general case, we make use of an algorithm presented in *FOX* (1966), to which we were introduced by Professor Maurice Queyranne (personal communication). In the present notation, it is:
1. set all \( n_i \) to zero,

2. find the zone \( i \) with the largest value of \( \phi_i(n_i + 1) - \phi_i(n_i) = p_i \Delta_i^2 (d - 2n_i - 1)/d \), and increment \( n_i \) by one station,

3. repeat step 2 until all stations have been allocated \( \sum_{i=1}^{K} n_i = d \).

This is a "greedy" algorithm, in that at each stage one allocates a station to the zone whose potential incremental contribution to the objective function is greatest. Intuitively, there are two features of this problem which argue for a greedy algorithm as a method of solution. Firstly, the improvement in the objective function if one increments the \( i \)th component, \( n_i \), depends only on \( n_i \) and not on the current allocations in other zones (\( \Phi \) is linear in its arguments). Hence, at any given stage, one does not 'forego' or alter the potential benefit of adding a station to some zone, by choosing another; it is still available later. Secondly, the (cumulative) contribution of a zone to the objective function, is concave in the number of stations placed there (the \( \phi_i \) are parabolas opening downward). This is a kind of foreknowledge: given \( n_i \) at some stage, the incremental improvement in \( \Phi \) if one added a station to zone \( i \) now, is an upper bound on the incremental improvements available later through zone \( i \).

FOX (1966) shows that the algorithm terminates at an optimum, but for a class of problems which unhappily does not include the one at hand, with respect to either the objective function or the constraint. Nonetheless, it turns out that the algorithm works in this case too. Let \( \mathbf{n} \) denote the (partial) allocation\(^\dagger\) \((n_1, \ldots, n_K)\), and call \( \mathbf{n} \) "\( p \)-undominated" when it satisfies (i) \( \sum_{i=1}^{K} n_i = p \), and (ii) \( \Phi(\mathbf{n}) \geq \Phi(\mathbf{m}) \), for all \( \mathbf{m} \) with

\(^\dagger\) The term 'allocation' is used throughout to mean either partial or complete allocations unless otherwise indicated.
$\sum_{i=1}^{K} m_i = p$. (Note that a $p$-undominated allocation is feasible if and only if $p = d$.)

Observe that an optimum is just a $d$-undominated allocation. Hence the algorithm terminates at an optimum if the following is true.

**Lemma 6.** If $x = (x_1, \ldots, x_K)$ is $p$-undominated, then $y = (x_1, \ldots, x_h + 1, \ldots, x_K)$ is $(p + 1)$-undominated, where $h$ is an index maximizing

$$\phi_i(x_i + 1) - \phi_i(x_i).$$

**Proof.** When $p = 0$ the result is obvious. For $p \geq 1$, choose any allocation $v$ with $\sum v_i = p + 1$. We show in successive cases that $\Phi(v) \leq \Phi(y)$. Choose any index $\ell$ such that $v_{\ell} > 0$. Decrement $v_{\ell}$ by one to get an allocation $u$. Now $\sum u_i = p$, and $\Phi(x) \geq \Phi(u)$, since $x$ is $p$-undominated. Suppose $u_{\ell} \geq x_{\ell}$. Then

$$\Phi(v) - \Phi(u) = \phi_{\ell}(u_{\ell} + 1) - \phi_{\ell}(u_{\ell})$$

$$\leq \phi_{\ell}(x_{\ell} + 1) - \phi_{\ell}(x_{\ell}), \quad \text{by concavity}$$

$$\leq \phi_h(x_h + 1) - \phi_h(x_h), \quad \text{by assumption}$$

$$= \Phi(y) - \Phi(x).$$

Together with $\Phi(u) \leq \Phi(x)$, this implies $\Phi(v) \leq \Phi(y)$.

Now suppose $u_{\ell} < x_{\ell}$. Since $\sum u_i = \sum x_i$, there is an index $j$ with $u_j > x_j$. Thus $u_j \geq 1$. Note that $u_j$ is also the $j^{th}$ component of $v$, since $j \neq \ell$. Consider an allocation $w$ having $w_j = u_j - 1 \geq 0$, and $w_i = u_i$ for all $i \neq j$. Since $\sum w_i = p$, it follows that $\Phi(w) \leq \Phi(x)$. Now

$$\Phi(v) - \Phi(w) = \phi_j(u_j) - \phi_j(u_j - 1)$$

$$\leq \phi_j(x_j + 1) - \phi_j(x_j), \quad \text{by concavity, and } u_j > x_j$$

$$\leq \phi_h(x_h + 1) - \phi_h(x_h), \quad \text{by assumption}$$

$$= \Phi(y) - \Phi(x).$$

16
Together with $\Phi(w) \leq \Phi(x)$, this implies $\Phi(v) \leq \Phi(y)$.

There are some notable properties of the optimal allocation scheme which we now summarize.

**Lemma 7.** For $K \geq 3$, the optimal allocation places no more than half the available stations in any one zone. Specifically, for $d$ even, the optimal allocation $n^*$ has $n^*_i \leq \frac{d}{2}$ for all $i$, while for odd $d$, $n^*_i \leq \frac{d}{2} - \frac{1}{2}$.

This is straightforward; the reader may convince himself of its validity by considering the $\phi_i$ as parabolas in $n_i$ which rise for $n_i < \frac{d}{2}$ and fall beyond that.

**Lemma 8.** For $K \geq 3$, the value of $\Phi$ at the optimum allocation is strictly monotonically increasing in the number of stations $d$.

**Proof.** Let $n^*$ denote an optimal allocation of $d$ stations. For emphasis write $\Phi(n^*, d) = \sum_{i=1}^K \phi_i(n^*_i, d)$. Now consider the problem of allocating $d + 1$ stations. Start with the (incomplete) allocation $n^*$. Some algebra shows that $\Phi(n^*, d + 1) - \Phi(n^*, d) = \sum_{i=1}^K p_i \Delta^2_i(n^*_i)^2/d(d+1) > 0$. Moreover, the new station can be allocated in such a way that $\Phi$ is increased further. To see this, note that Lemma 7 together with $K \geq 3$ implies that there must be a zone $j$ with $n^*_j \leq \frac{d}{2} - 1$ for $d$ even; for $d$ odd, let $j$ be any zone. In either case, putting the new station in zone $j$ constitutes an increase in the value of $\phi_j(\cdot; d+1)$, since the function is increasing to the left of $(d+1)/2$.

Thus we have found a (possibly sub-optimal) allocation of $d + 1$ stations which is strictly superior to the optimum allocation of $d$ stations. This proves the claim.

Lemma 8 says that one can always increase the expected value of the noncentrality parameter by increasing the number of stations available for allocation. This implies that the optimal design displays a kind of efficiency: for a given number of stations $d$
and an optimal allocation $n^*$ of these, one cannot achieve the corresponding value of $E(\delta^2)$ with less than $d$ stations.

Professor Maurice Queyranne has conjectured (personal communication) that the optimal discrete design may be found by first finding the optimal (continuous) sampling fractions, $f_{ai}$, $i = 1, \ldots, K$. Then the products $df_{ai}$ are each rounded down when they are nonintegral. The resulting partial allocation of, say, $p$ stations, is conjectured to be $p$-undominated. Thus if the greedy algorithm is applied to allocate the remaining $d-p$ stations, the result will be an optimal solution to the discrete problem. We have not yet shown this conjecture to be true, but it is well-supported by our empirical studies.

Example 9. In this example, there are $K = 7$ zones where an impact can occur as well as zone 0 where it is known that an impact is impossible. The values of $p_i$ for zones 1, $\ldots$, $K$ are respectively, .30, .15, .20, .12, .05, .04, and .14, and of the $\Delta^2$'s, 33.333, 133.333, 150.000, 583.333, 1600.000, 2250.000, and 714.286 to three decimal places. It follows that to the nearest integer, the $\tau$'s defined just above Theorem 2, are 10, 20, 30, 70, 80, 90, and 100, respectively. The value of $L(m)$ above Theorem 2 must be negative for $m = 5, 6, and 7$ according to theory. $L$ is easily computed and found to be $-0.0194$ and $+0.0188$ for $m = 4$ and 3, respectively. It follows from Theorem 2, that the AOD sampling fractions in this example are, $f_{ai} = 0$ for $i = 0, 1, 2, and 3$ and $f_{ai} = 0.203, 0.238, 0.268,$ and $0.291$ for $i = 4, 5, 6,$ and 7 respectively. If the number of monitoring stations is to be $d = 10$, say, then these AOD fractions would prescribe, after rounding off the products of the $f_{ai}$ with $d$ to the nearest integer, 2 stations in each of zones 4 and 5, and 3 stations in each of zones 6 and 7.
The true optimal design may be found by the greedy algorithm developed above. It terminates after the tenth iteration and yields exactly the same result.

3. **Subregionally Homogeneous Impacts.**

Suppose that the data are to consist of equal numbers, \( n \), of vectors of measurements which are to be made at each site before and after the time of the potential impact. The objective, as before, is that of detecting any change in the level of the measurements which may occur in the region other than that due to an ambient change affecting all sites. As before, homoscedasticity, the multinormality of data vector distributions, and approximate independence of observations across space and time is assumed.

In the situation described in the last paragraph, the noncentrality parameter, \( \delta^2 \), of the appropriate multivariate test for space-time interaction based on a design, \( D = \{(i, j)\} \), consisting of \( d \) stations, is proportional to

\[
\sum_{(i,j) \in D} (Z_{ij} - \bar{Z}_D)^T \sigma^{-2} (Z_{ij} - \bar{Z}_D)
\]

where \( \bar{Z}_D = \frac{1}{d} \sum_{(i,j) \in D} Z_{ij} \), and \( \sigma^2 \) denotes the covariance matrix of the distribution possessed in common by the data vectors.

Ordinarily, \( \sigma^{-2} \) will be unknown. However, since only the expected value of \( \delta^2 \) is required in our analysis, \( \sigma^{-2} \) may be replaced by its prior expectation provided that, a priori, \( \sigma^{-2} \) and the \( Z \)'s are independent. The authors believe this condition will be plausible in many applications and its validity will be assumed here. If this condition does not apply, a different approach to solving the problem addressed here would be needed.
Actually the prior expectation of $\sigma^{-2}$ needs to be specified only up to a multiplicative constant for the purposes of our analysis, since the optimal design is free of this constant. In any event, with $\langle \cdot, \cdot \rangle$ defined by $\langle x, y \rangle = nx^T E(\sigma^{-2})y$, and $\|x\|^2$ by $(x, x)$, for arbitrary vectors $x$ and $y$, $E(\sigma^2) = \sum_{(i, j) \in D} E\|Z_{ij} - Z_D\|^2$.

Let us now impose the further conditions of second order zonewise exchangeability (SOZE) described for the univariate case in Section 1. For any fixed zone $i = 1, \ldots, K$, the $\{E\|Z_{ij}\|^2, j = 1, \ldots, n_i\}$ and $\{E(Z_{ij}, Z_{ij'})$, $j, j' = 1, \ldots, n_i$, $j \neq j'\}$ have common values, $\beta_i$ and $\gamma_i$ respectively. As well, $\{E(Z_{ij}, Z_{ij'}), i, i' = 1, \ldots, K$, $i \neq i'\}$ have a common value, $\beta_{ii'}$, and we assume the matrix $(\beta_{ii'})$ has full rank.

Under the assumptions of SOZE,

\[
E(\sigma^2) = \sum_{(i, j) \in D} E\|Z_{ij}\|^2 - dE(Z_D, Z_D)
\]

\[
= \sum_{(i, j) \in D} n_i \beta_i - d^{-1} \sum_{(i, j) \in D} \sum_{(i', j') \in D} E(Z_{ij}, Z_{i'j'})
\]

\[
= \sum_{(i, j) \in D} n_i \beta_i - d^{-1} \sum_{i} n_i \beta_i - d^{-1} \sum_{i} n_i (n_i - 1) \gamma_i - d^{-1} \sum_{i} \sum_{i' \neq i} n_i n_{i'} \beta_{ii'}
\]

\[
= \sum_{i} f_i ((d - 1) \beta_i + \gamma_i) - d \sum_{i} f_i^2 \gamma_i - d \sum_{i} \sum_{i' \neq i} f_i f_{i'} \beta_{ii'}
\]

\[
\overset{\Delta}{=} H(f),
\]

say.

As in Section 2, $H$ is a strictly concave function of $f$ and hence has a unique global maximum, $f_\ast$, on the convex set $\mathcal{F} = \{f : 0 \leq L \leq f \leq U \leq 1, \sum f_i = 1\}$, where 1 denotes the vector all of whose elements are 1, and L, U are any feasible bounds on $f$ imposed by the problem. As before, $f_\ast$ will be referred to as the AOD sampling fractions and the proof of Theorem 2 can be adapted to this more general case. To state the result, Theorem 7, let $J_L$ and $J_U$, respectively, denote the subscripts $i$ for which $f_{si} = L_i$, $f_{si} = U_i$, and the corresponding constraints are binding. Let $\mathcal{F}$ denote the complement
of $J_L \cup J_U$ in $\{1, \ldots, K\}$. For simplicity we assume zone 0 of Section 2, where no impact can occur, has been relabelled and incorporated among the designated $K$ zones so that $K \geq 2$. The proof is omitted.

**Theorem 10.** For $K \geq 2$, the sampling fractions which maximize the expected noncentrality parameter are, for a uniquely determined partition $\{1, \ldots, K\} = J_L \cup \overline{J} \cup J_U$, the solution of the equations

$$\lambda = M_i(f_a), \; i \in \overline{J}$$

subject to $f_{ai} = L_i, \; i \in J_L$, and $f_{ai} = U_i, \; i \in J_U$, where $M_i(f_a) = \beta_i - \frac{1}{2}(\beta_i - \gamma_i) - 2\gamma f_{ai} - 2 \sum_{j \neq i} f_{aj} \beta_{ij}, \; M_i(f_a) < \lambda$ or $> \lambda$ according as $i \in J_L$ or $i \in J_U$, and $\lambda$ is a uniquely determined constant. \[ \blacksquare \]

In general Theorem 10 does not admit an explicit solution. A numerical solution is, however, readily found in specific cases using an adaptation of the simplex algorithm from the theory of linear programming. It is used purely as a device for systematically searching the set of feasible solutions for the one which uniquely satisfies all the conditions in Theorem 10. However, this point will not be elaborated on here because standard quadratic programming computer packages can be used to solve the quadratic programming problem directly in specific cases. But the characterization of the optimal solution given in Theorem 10 is required in subsequent discussion.

### 3.1. Particular Covariance Structures.

A specialization of the model leading to Theorem 10, for which the AOD sampling fractions can be found explicitly, is analyzed in this Section. The particular covariance structure which gives this specialization is expected to be fairly broadly applicable.
Let:

$$\beta_{ij} = B_0 + B_1 \phi_i \phi_j, \text{ for all } i \neq j, \quad (5)$$

$$\gamma_i = \beta_i - C_0 \phi_i^r = B_0 + B_1 \phi_i^2 - C_1 \phi_i^r, \text{ for all } i,$$

where \( r = 0 \) or \( 2, \ 0 \leq \phi_1 \leq \cdots \leq \phi_K, \ B_1 > 0, \ C_0 \geq 0, \ C_1 < 0 \) and \( B_0 \) are all specified constants and \( f \) is restricted only by \( 0 \leq f \leq 1 \) so that \( 0 = L, \ U = 1. \) If \( r = 2, \ \phi_1 > 0 \) is assumed to avoid degeneracy. Example 12 below shows how such a covariance structure can arise. With this added structure (4) becomes

$$\lambda = B_0 + B_1 \phi_i^2 + d^{-1} C_0 \phi_i^r + 2C_1 \phi_i^r f_{st} - 2 \sum_{s \neq j} f_{st} (B_0 + B_1 \phi_i \phi_j), \ i \in \mathcal{I}. \quad (6)$$

Consider first the case where \( r = 0. \) Here equation (6) becomes

$$\lambda_i = \phi_i^2 + 2C_2 f_{st} - 2\phi_i \mu, \ i \in \mathcal{I} \quad (7)$$

where \( \lambda_i = (\lambda + B_0 - d^{-1} C_0) / B_1, \ C_2 = C_1 / B_1 \) and \( \mu = \sum f_{st} \phi_j. \) Multiply both sides of equation (7) by \( \phi_i \) and sum the result over all \( i \in \mathcal{I}. \) The result:

$$\lambda_i S_i = S + 2C_2 \mu - 2S_2 \mu, \quad (8)$$

where \( S_i = \sum_{i \in I} \phi_i, \ -\infty < i < \infty. \) Now sum both sides of (7) over \( i \in \mathcal{I} \) to get

$$\lambda S = S_2 + 2C_2 - 2S_1 \mu, \quad (9)$$

Equations (7), (8) and (9) are readily solved for \( f_{st}, \lambda_i \) and \( \mu. \) The results are:

$$f_{st} = \frac{1}{2} C_2^{-1} [\lambda_i - \phi_i^2 + 2\phi_i \mu], \ i \in \mathcal{I}$$

$$\mu = \frac{1}{2} [c(\phi^2, \phi) - 2C_2 S_1 S_0^{-2}) [c(\phi, \phi) - C_2 S_0^{-1}]^{-1}, \quad (10)$$

$$\lambda_i = (S_2 + 2C_2 - 2S_1 \mu) / (S_0),$$

22
where, in general, \( c(x, y) = \frac{1}{n_0} \sum_{i \in \mathcal{F}} (x_i - \bar{x})(y_i - \bar{y}) \).

But \( f_{ai} \geq 0 \) entails \( \lambda_1 \leq \phi_i^2 - 2\phi_i \mu \) for \( i \in \mathcal{F} \) while Theorem 10 requires \( \lambda_1 > \phi_i^2 - 2\phi_i \mu \) for \( i \in \mathcal{F} \equiv \mathcal{F}_L \). Since \( x \mapsto x^2 - 2x \mu \) is a quadratic function which attains its minimum at \( x = \mu \), it follows that \( \mathcal{F} = \{ i : \ell \leq i \leq u \} \) for integers \( \ell, u \in \{ 1, \ldots, K \} \) unless \( \mathcal{F} = \{ 1, \ldots, K \} \) when \( \mathcal{F} = \emptyset \).

AOD sampling fractions are now readily found by trying successively smaller \( \mathcal{F}'s \) until a \( \lambda_1 \) and \( \mu \) are found for which the conditions of Theorem 10 are satisfied. That a unique \( \lambda_1 \) and \( \mu \) exist follows from Theorem 10 as well. The procedure is illustrated below in Example 12.

The case where \( r = 2 \) and \( \phi_i > 0 \) for all \( i \) in equations (5) is handled in a similar fashion. Instead of equation (7) we obtain

\[
\lambda_2 = \phi_i^2 + 2C_3\phi_i^2f_{ai} - 2B_3\phi_i\mu, \quad i \in \mathcal{F}
\]

where \( \lambda_2 = (\lambda + B_0)/B_2, \quad B_2 = (B_1 + C_0/d), \quad C_3 = C_1/B_2, \quad B_3 = B_1/B_2, \) and \( \mu \) is as defined above. If \( 0 < \phi_i \) for all \( i \), equation (11) may be solved in the same manner as (7) to yield

\[
f_{ai} = \frac{1}{2}C_3^{-1}\phi_i^{-2}[\lambda_2 - \phi_i^2 + 2B_3\phi_i\mu], \quad i \in \mathcal{F}
\frac{\mu}{2} = \frac{1}{2}[c(\phi, \phi^{-2}) + 2C_3S_{-1}]/[C_3S_{-2} - B_3c(\phi^{-1}, \phi^{-1})],
\]

\[
\lambda_2 = [S_0 + 2C_3 - 2B_3\mu S_{-1}]/S_{-2},
\]

with \( c(-, -) \) as defined just below equation (10). As before we may deduce that \( \mathcal{F} = \emptyset \) or \( \mathcal{F} = \{ i : \ell \leq i \leq u \} \) so that \( \mathcal{F} \) can be found without resorting to a combinatorial search. The next theorem summarizes the results of this section.
Theorem 11. With the assumptions in equations (5) added to those of Theorem 10, the AOD sampling fractions are given in equations (10) or (12) according as \( r = 0 \) or 2 in equations (5). In any case, zones selected for monitoring by the AOD are the complement of \( \mathcal{J} = \emptyset \) or \( \mathcal{J} = \{ i : \ell \leq i \leq u \} \) for integers \( \ell, u \in \{1, \ldots, K\} \).

The following example involves a special case of some interest in its own right.

**Example 12.** Suppose the data vectors are one–dimensional, \( \beta_i = \sigma^2 + \mu_i^2 \), \( \gamma_i = \sigma^2 \rho_w + \mu_i^2 \) for all \( i \), and \( \beta_{ij} = \sigma^2 \rho_b + \mu_i \mu_j \) for all \( i \neq j \) with \( \rho_b < \rho_w \). Then assumptions (5) obtain with \( B_0 = \sigma^2 \rho_b \), \( B_1 = 1 \), \( \phi_i = \mu_i \) for all \( i \), \( C_0 = \sigma^2 (1 - \rho_w) \), \( r = 0 \) and \( C_1 = \sigma^2 (\rho_b - \rho_w) < 0 \).

A natural alternative to the last covariance model would have \( \text{Var}(Z_i) \propto \mu_i^2 \) provided \( \mu_i > 0 \) for all \( i \). In this case equations (5) obtain with \( r = 2 \).

To illustrate how to get the optimal design corresponding to the first of the two models in this example, let \( K = 5 \), \( \rho_b = .1 \), \( \rho_w = .8 \), \( \sigma^2 = 16 \) and the \( \{ \mu_i \} \) be those given below:

\[
i = 1 \ 2 \ 3 \ 4 \ 5 \\
\mu_i = 1 \ 2 \ 4 \ 5 \ 10
\]

Both \( \mu \) and \( \lambda_1 \) are computed successively for the various (but not all) possible choices of \( \mathcal{J} \), each a subinterval of \( \{1, \ldots, 5\} \) to illustrate the process of finding the optimal subset. The results, shown in Table I, reveal that \( \mathcal{J} = \{1,2,5\} \) is the optimal selection of zones since \( \lambda_1 > \mu_i^2 - 2\mu_i \mu \) for \( i = 3,4 \) while \( \lambda_1 \leq \mu_i^2 - 2\mu_i \mu \) for \( i = 1,2,5 \). Then using equations (10) it follows that \( f_{a_1} = f_{a_5} = \frac{1}{2} \) while \( f_{a_2} = 0 \). Thus, equivalently, \( \mathcal{J} = \{1,5\} \) is optimal. However, \( f_2 \geq 0 \) is a nonbinding constraint; \( f_{a_2} = 0 \) cannot be improved on by choosing \( f_{a_2} < 0 \) in violation of this constraint.
PLACE TABLE 1 ABOUT HERE

3.2. Multi-Region, Uniform Impacts.

To complete this section and achieve additional insight into the nature of AOD's, a substantial generalization of the model of Section 2 will be considered in this subsection. Theorem 10 will apply but the results of Section 3.1 will not. The assumptions added to those hypothesized in Theorem 3 are: \( Z_{ij} = Z_i \) for all \( j \) in zone \( i \); \( L = 0 \), and \( U = 1 \). Consequently \( \beta_i = \gamma_i \) and \( \beta_{ii} = \beta_i \). Assume \( E\|Z_i - Z_j\|^2 > 0 \) when \( i \neq j \).

Theorem 10 simplifies substantially with these added assumptions and the result is the next theorem.

**Theorem 13.** The AOD sampling fractions, \( \{f_{ai}\} \), are uniquely characterized by the conditions

\[
E\|Z_i - \overline{Z}\|^2 = \lambda, \quad i \in \mathcal{I},
\]

\[
E\|Z_i - \overline{Z}\|^2 < \lambda, \quad i \in \mathcal{I}
\]

(13)

for some constant \( \lambda \), where \( \overline{Z} = \sum_k f_{ak}Z_k \).

It is easily shown that the constant \( \lambda \) of Theorem 13 is \( \lambda = E\sum f_{ai}\|Z_i - \overline{Z}\|^2 \).

The proof of the following result is straightforward and omitted.

**Corollary 14.** The AOD places sampling sites in at least two subregions.

The next result is not intuitively obvious.

**Corollary 15.** The following statements are equivalent:

(i) The AOD places all the monitoring stations in zones \( i \) and \( j \).

(ii) \( f_{ai} = f_{aj} = \frac{1}{2} \).

(iii) \( E(Z_{it} - Z_i, Z_{jt} - Z_j) < 0 \) for all \( t \neq i, j \).
Proof. Suppose $f_{at} = 0$ except when $t = i$ or $j$. With $\alpha = f_{ai}$ and $(1 - \alpha) = f_{aj}$, Theorem 10 implies

$$0 = E||Z_i - \alpha Z_i - (1 - \alpha) Z_j||^2 - E||Z_j - \alpha Z_i - (1 - \alpha) Z_j||^2$$

$$= (1 - 2\alpha) E||Z_i - Z_j||^2.$$ 

Thus $\alpha = \frac{1}{2}$ and this establishes the equivalence of (i) and (ii). Now suppose (iii) holds: $0 > E(Z_{\ell} - Z_i, Z_{\ell} - Z_j)$ for all $\ell \neq i, j$. Then $-E(Z_i, Z_j) > E(Z_{\ell}, Z_{\ell}) - 2E(Z_{\ell}, \bar{Z})$, where $\bar{Z} = \frac{1}{2}Z_i + \frac{1}{2}Z_j$. Thus $E(Z_{\ell} - \bar{Z}, Z_{\ell} - \bar{Z}) < E(Z_i, Z_j) - E(Z_{\ell}, Z_{\ell}) = \frac{1}{2}E||Z_i - Z_j||^2 = \frac{1}{2}E||Z_i - \bar{Z}||^2 + \frac{1}{2}E||Z_j - \bar{Z}||^2 = \lambda = E||Z_i - \bar{Z}||^2 = E||Z_j - \bar{Z}||^2$, which proves that (iii) implies (ii) by Theorem 13.

Conversely, suppose (ii) holds. Theorem 10 implies $E||Z_{\ell} - \bar{Z}||^2 < \lambda$ for every $\ell \neq i, j$ where $\lambda = \frac{1}{2}E||Z_i - Z_j||^2$, and $\bar{Z} = \frac{1}{2}(Z_i + Z_j)$. A simple calculation shows that this is equivalent to the assertion that $0 > E(Z_{\ell} - Z_i, Z_{\ell} - Z_j)$ for all $\ell \neq i, j$, so the proof of the converse and hence the theorem is complete. \[\square\]

It is not obvious that if stations $i, j$ satisfying (iii) exist, they are unique. A simple direct proof is the following. For any two pairs of zone labels,

$$0 < E||Z_i + Z_j - Z_m - Z_n||^2$$

$$= E(Z_i - Z_m, Z_i - Z_n) + E(Z_j - Z_m, Z_j - Z_n)$$

$$+ E(Z_m - Z_i, Z_m - Z_j) + E(Z_n - Z_i, Z_n - Z_j).$$

So it is impossible that $(i, j)$ and $(m, n)$ both satisfy (iii), proving the assertion of uniqueness.

**Corollary 16.** $f_{ai} \leq \frac{1}{2}$ for all $i$.

**Proof.** Suppose to the contrary $f_{ai} > \frac{1}{2}$ for some $i$, say $i = 1$ for simplicity of
exposition. By Theorem 10, \(0 = E\|Z_1 - \bar{Z}\|^2 - E\|Z_j - \bar{Z}\|^2\) for every \(j \in \mathcal{J}\). But

\[
E\|Z_1 - \bar{Z}\|^2 - E\|Z_j - \bar{Z}\|^2 = E\|Z_1\|^2 \left(1 - 2f_{a_1}\right) - 2f_{a_1} E\langle Z_1, Z_j \rangle - \sum_{k \neq 1, j} 2f_{a_k} E\langle Z_1, Z_k \rangle - E\|Z_j\|^2 \left(1 - 2f_{a_j}\right) + 2f_{a_1} E\langle Z_1, Z_j \rangle + \sum_{k \neq 1, j} 2f_{a_k} E\langle Z_j, Z_k \rangle \\
= \left(1 - 2f_{a_1}\right) (E\|Z_1\|^2 - E\langle Z_1, Z_j \rangle) - \left(1 - 2f_{a_j}\right) (E\|Z_j\|^2 - E\langle Z_1, Z_j \rangle) \\
- \sum_{k \neq 1, j} 2f_{a_k} (E\langle Z_1, Z_k \rangle - E\langle Z_j, Z_k \rangle).
\]

But \(-\left(1 - 2f_{a_j}\right) = (1 - 2f_{a_1}) - \sum_{k \neq 1, j} 2f_{a_k}\). Thus \(0 = (1 - 2f_{a_1}) E\|Z_1 - Z_j\|^2 \\
- \sum_{k \neq 1} 2f_{a_k} E\langle Z_j - Z_1, Z_j - Z_k \rangle\). So \(f_{a_1} > \frac{1}{2}\) implies \(0 > \sum_{k \neq 1} f_{a_k} E\langle Z_j - Z_1, Z_j - Z_k \rangle\) for every \(j \neq 1, j \in \mathcal{J}\). Thus

\[
0 > \sum_{j \neq 1} \sum_{k \neq 1} f_{a_j} f_{a_k} E\langle Z_j - Z_1, Z_j - Z_k \rangle
\]

or

\[
0 > \sum_{k \neq 1} \sum_{j \neq 1} f_{a_k} f_{a_j} E\langle Z_k - Z_1, Z_k - Z_j \rangle,
\]

where the last inequality is obtained from its predecessor simply by interchanging the indices of summation, \(j\) and \(k\). But the sum of the right hand sides of these last two inequalities is \(\sum_{j \neq 1} \sum_{k \neq 1} f_{a_j} f_{a_k} E\|Z_j - Z_k\|^2 > 0\), and this is a contradiction. Thus \(f_{a_1} > \frac{1}{2}\) is impossible and the conclusion of the Corollary obtains. □

Theorem 13 does not in general admit an explicit solution. However, explicit solutions of interesting particular cases of this theorem are given by equations (5) when \(C_0 = 0\) (since here \(\gamma_i = \beta_i\) for all \(i\)) and Theorem 11. For brevity the details will not be presented here.

4. DISCUSSION.

This paper has investigated design problems involving subjective choices based upon prior information, explicitly admitted and used in a fairly conventional setup.
where a network is to be designed to detect the potential spatial impact of an event occurring at a known time. Very general optimal designs are derived and some interesting qualitative conclusions are reached. Our results will be discussed below.

The main products of this paper are (i) the general numerical algorithm in Theorem 7 for deriving the approximately optimal sampling fractions for a very general class of models, (ii) the identification of potentially useful special cases and the derivation of their explicit solutions, and (iii) the qualitative insights derived from the investigation of these special cases like those reported in the Introduction.

It may well be objected that it is inconsistent to “break the Bayesian eggs” by expressing the prior information through probabilities and then to not enjoy the “Bayesian omelette” opting instead to use the classical approach of hypothesis testing. This is a valid criticism. However, our goal is not to find alternatives to the $F$-test which as noted above is a conventional method in this context, but rather to discover how to incorporate prior information into design if the $F$-test is to be used and to discover something about the effects prior information has on the character of the design. Our results may well suggest considerations in other design contexts as well.

A deficiency in our approach lies in our concession to simplicity in choosing the expected noncentrality parameter of the $F$ statistic rather than the power of the test itself for the objective function. Although this simpler course has some intuitive appeal, the quality of the resulting design with respect to the power function needs further study. A major impediment to such a study is the great practical difficulty of finding the optimum design in realistic examples.

This paper does investigate a fairly general optimization paradigm from which
some insight into the effect of prior information can be gleaned. Although our models are fairly general, the approximately optimal designs (AOD's) implied by these models are explicit enough as to make the effect of the prior information quite visible. Subdividing the area of interest into small, homogeneous subregions as a first step in defining these designs does not seem excessively strong given the inevitable limitations of the designer's prior experience. It is this step in conjunction with the assumption of second order zonewise exchangeability that permits us to replace the insoluble binary (0–1) programming problem, with a relatively simple quadratic programming problem involving continuous sampling fractions for the subregions involved. The optimal solutions are then derived by conventional quadratic programming techniques. This approach is presented by Zidek (1984) where a special case of that treated in Section 2 of this paper is considered.

4.1 Practical Considerations

Of course we recognize that in practice, "good" rather than "optimal" designs are needed and optimal designs like those in this paper must be considered as tentative proposals susceptible to modification depending on the circumstances prevailing in the context of their implementation. Practical aspects of the design problem are discussed by Gilbert (1987), Millard and Lettenmeier (1986), and Green (1979). We believe "optimal" designs are valuable starting points since they can be explicated in terms of their axiomatic underpinnings and proposed changes to these optimal designs can be described in terms of modifications of the axioms. This can provide a degree of confidence and clarity in the typically complex situation confronting an environmetrician.
Political pressure exemplifies the sort of nonstatistical consideration which distorts a proposed network's design. A municipality may argue that a station should be closer to or even inside its boundaries to ensure that if there is an impact within those boundaries it will be discovered. The designer will contend that it should be well outside the boundaries to avoid spurious contamination from purely local sources like automobile exhaust, say. In such situations the site may well be chosen by a process of political compromise.

Logistical aspects of monitoring may also play a major role in siting the network's stations. Accessibility and proximity to an analytical laboratory are two examples of the sort of practical considerations which may ultimately confront the designer. A potential advantage of our approach with respect to logistical and political problems is that only the sampling fractions for entire subregions are prescribed and not the sitings within subregions. While randomized allocation of stations within subregions is deemed to be important, even when external considerations do dictate a relaxation of this requirement the optimal sampling fractions themselves may nevertheless be preserved.

4.2 Discussion of Assumptions

A more statistical category of problems derives from concerns about the model assumptions. Calling for a balanced design as we have done seems reasonable. But the supposition underlying the assumptions of normality and homoscedasticity, that the (conditional) variability in the data vectors derives from replication error alone is simplistic. In fact this variation may be contaminated by such things as interlaboratory differences, spatial variation, and temporal variation. The latter may obtain
even when the replicates are measurements of fractions of a completely homogeneous mixture when those measurements are made at differing times for logistical or other reasons.

The effect of the contamination may be serious, invalidate the key assumptions, lead to a loss of conditional independence between the data vectors obtained at different sites and times, and ultimately call into question the value of the network itself. In this paper we are assuming that due regard has been paid to the very important practical aspect of experimental design, that of ensuring this potential contamination is insignificant.

The theory of Section 3 applies even when the unknown ambient change over the entire region is not approximately constant. However, in that case, these ambient changes will be confounded with those due to the impact and this will cast doubt on the meaning of a conclusion of a significant interaction. The impact effect becomes an unidentifiable component of the space-event interaction. It is the coefficient of variation of the ambient space-time series that matters since the mean value, known or unknown, drops out of the expression for the noncentrality parameter and cannot itself lead to a rejection of the null hypothesis. If it is believed that this coefficient of variation is small, the results of this paper will be valid and this is the assumption under which we are proceeding.

If the coefficient of variation in the spatial variation of the ambient level is large, it may be possible in some cases to design the sampling experiment to incorporate this variation into the replicated measurements while not losing the vital conditions of independence and homoscedasticity. This will inflate the residual variation and
make it harder to detect the impact if any. But unless an impact were significant when measured against ambient variation it would presumably not be considered important. In any case this gives another approach to overcoming the problem arising from shifts in ambients levels which are not approximately constant over space.

The plausibility of the assumption of an approximately constant ambient change is inversely related to the size of the overall region so this should be constrained to be as small as possible subject to the requirements that the potential impact, if any, occurs somewhere in the region and there is a subregion to serve as a pseudo-control where the impact is almost certain not to occur.

If it were supposed somewhat unrealistically that the ambient change could be specified reasonably accurately, this change could be removed from the data at the outset. Any observed differences could then be ascribed to the impact itself. It is easily seen in this case that all the monitoring stations should be placed in the riskiest zone. This strategy would not seem very wise even though it is optimal. The problem is a problem of the $F$–test itself which takes no account of the economic consequences of missing the impact altogether (a type two error). A more realistic approach to design than that taken in this paper would incorporate such costs. However, it would also complicate the analysis, add extraneous factors which would obscure the role of the prior information which we have tried to expose, and finally, it would introduce a host of tangential issues about how such costs are to be quantified. We have elected to treat what seems to us to be the most fundamental paradigm here and leave the consideration of practical refinements to future work.

Another situation for which the optimal design puts all the stations into the
riskiest zone is that in which the before and after measurements are well described by a fixed effects model which may incorporate different levels from a variety of factors such as "season". However, such a model seems unlikely to be realistic for many environmental processes. Moreover, its implementation requires that a series of before and after measurements be made in order to estimate the model and this will not always be feasible.

There is empirical evidence that the assumption of conditional independence between data vectors at different sites and times is not as unrealistic as might be supposed on hypothetical grounds. The studies of Hirsch and Gilroy (1985), Le and Petkau (1988), and Egbert and Lettenmaier (1986) suggest that if sampling points are reasonably well separated then even the unconditional correlations may be expected to be modest. This may well be due to enormous complexity and noise in environmental processes which dwarfs any underlying structural features.

4.3 Our Results and Their Implications

Section 2 demonstrates a valuable method for optimization and may therefore be of some independent interest. The problem treated there is somewhat specialized so the solution obtained will not be so broadly applicable as those of Section 3. But it is explicit and so suggests ideas which are not apparent from the results of Section 3. It is Theorem 1 which gives the result quoted in the Introduction that if there are only two subregions or "zones", and it is known that the potential impact cannot occur in the first but may occur in the second, then put 50% of the sampling effort in each of them. If a third zone were added then according to Corollary 3, zone 1 can be dropped. What is somewhat surprising however, is that the allocation of sampling
sites is once again 50% to each of the two monitored zones (2 and 3) regardless of the conditional probability, \( p_i \), that the impact will be in zone \( i, i = 2, 3 \), or the sizes of the potential impacts in these zones. These parameters are important in general, and Theorem 2 says that the fraction of sites in zone \( i \) should be proportional to \( 1 - (\lambda / \tau_i) \), where

\[
\tau_i = p_i \Delta_i^2.
\]

Here \( \lambda \) is a positive constant, and this fraction is non-negative. Thus zones with small potential impacts and likelihoods of impacts will be played down in the monitoring program, except when there are just two potential impact zones as noted above. This result for two zones obtains as well in the general context of Section 3.2 where both zones may potentially be hit, so it seems qualitative in nature.

Another conclusion which the authors find somewhat surprising and cannot rationalize is that of Corollary 4 which says that if there are three zones which can potentially be impacted, monitoring stations must be placed in all three, never two, regardless of the underlying parameters. This does not carry over to the generalization of the model in Section 2 presented in Section 3.2, where monitoring just two zones may sometimes be optimal. The conditions for this latter result are described in detail below.

When there are four or more potential impact zones, no conclusions of a general nature can be made about optimum number of zones to be monitored. It can be as few as three or as great as the total number of zones available, \( K \). If \( K = 4 \) one of these four zones, say \( i \), would be dropped from the optimal design if and only if \( \tau_i \) were significantly less than the others; more precisely, less than \( \frac{1}{3} \) of the harmonic
mean of the remaining \( \tau \)'s.

Theorem 10 is most general result given in the paper. It characterizes the optimal design for a very general impact model and covers the case where the sampling fractions are constrained by upper and lower bounds. The latter may arise in practice because the potential number of monitoring sites available in a small zone limits the fraction of the total sample which may be placed in that zone. This constraint is unlikely to be binding in a structural sense because of budgetary constraints which limit the sample's size and the infinite divisibility of space. However, the need to separate sites to gain independence, for example, might impose practical limitations on the fraction of sites in a zone. The lower bound may derive for logistical or political considerations of the type indicated above.

Theorem 10 does not yield an explicit result. It can be used in conjunction with a modified simplex algorithm to obtain an optimal design in specific applications once the required parameters have been specified. Since this method of implementing the Theorem is standard in reference works on the theory of quadratic programming no details are given in this paper. However, the Theorem is used in Section 3 to obtain explicit solutions in interesting special cases.

Section 3.1 gives a broad class of problems defined in terms of their special covariance structure for which explicit formulas for the approximately optimal sampling fractions can be found. The covariances are taken to be specified in terms of certain positive constants \( \phi_1 < \phi_2 < \ldots < \phi_K \), where \( E(Z_i^2) \) is an increasing linear function of \( \phi_i \). Theorem 11 says the approximately optimal design will incorporate those zones for which the \( \phi \)'s are small and large and drop those with \( \phi \)'s in the midrange, if
any are dropped. This is intuitively natural; to maximize the expected noncentrality parameter is to maximize the anticipated contrasts among the zones selected and the Theorem simply expresses that goal since the expected sizes of the $Z^2$'s indicate the relative sizes of the $Z$'s themselves.

For the special case of Theorem 11 specified by $r = 0$, the squared length of the interval of $i$'s corresponding to excluded zones is $4(\sigma^2_{\phi} + (\mu - \phi_{avg})^2) + \text{constant}$, where $\mu$ denotes the average of the $\phi$'s with respect to the approximately optimal sampling fractions while $\phi_{avg}$ and $\sigma^2_{\phi}$, respectively, denote the sample mean and variance of the $\phi$'s corresponding to the zones actually incorporated into the approximately optimal design. So if the $\phi$'s are quite variable in size, a large value of $\sigma^2_{\phi}$ is achievable and this interval of exclusion will be large; that is, a relatively large number of zones will be excluded as might be expected on intuitive grounds.

The scenario embodied in Section 3.2 is a substantial extension of that in Section 2. Explicit formulas can be obtained in an interesting special case with the help of the results of Section 3.1. However, most of the results presented there are qualitative in nature. Corollary 15 says that in a certain very special case, only two zones, say 1 and $K$, need to be monitored. This special case is when the variances in, $\sigma^2$, and cross–correlations between, $\rho$, respective zones are both constants over zones and $(\mu_{\ell} - \mu_1)(\mu_K - \mu_1) > \sigma^2(1 - \rho)$ for all $\ell$, $1 < \ell < K$, after the $\mu$'s have been ordered from smallest ($\mu_1$) to largest ($\mu_K$). This condition says that the anticipated values of the impacts in the two zones, 1 and $K$, are well separated from the rest, an intuitively appealing result. This result also says it is harder to come down to monitoring just two zones when stations are uncorrelated while if they are nearly perfectly correlated,
only two zones are needed, another appealing conclusion.

The most striking result in the paper is perhaps that in Corollary 16 which says that in the fairly general situation addressed in Section 3.2, there are no circumstances when more than 50% of the network's sites should be placed in any one zone. We find this result unintuitive and have not been able to rationalize it. Further study remains to be done to discover how much more generally this sort of result obtains.

ACKNOWLEDGEMENTS.

The authors benefitted from stimulating discussions with Dr. Jonathan Berkowitz and Dr. Harrry Joe about the problems treated here. Dr. Maurice Queyranne provided invaluable advice. Helpful comments were provided by the SIMS research group at the University of British Columbia.

REFERENCES.


**TABLE 1** Computations for Selecting the Optimal Monitoring Zones in Example 12 Using Equations (10)

<table>
<thead>
<tr>
<th>Network Zones</th>
<th>$\mu$</th>
<th>$\lambda_1$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>-17</td>
<td>-1</td>
<td>0</td>
<td>8</td>
<td>15</td>
<td>80</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>-116</td>
<td>-19</td>
<td>-36</td>
<td>-64</td>
<td>-75</td>
<td>-100</td>
</tr>
<tr>
<td>1, 5</td>
<td>5.5</td>
<td>-18</td>
<td>-10</td>
<td>-18</td>
<td>-28</td>
<td>-30</td>
<td>-10</td>
</tr>
<tr>
<td>1, 2, 5</td>
<td>5.5</td>
<td>-18</td>
<td>-10</td>
<td>-18</td>
<td>-28</td>
<td>-30</td>
<td>-10</td>
</tr>
<tr>
<td>1, 3, 4, 5</td>
<td>5.6</td>
<td>-24.5</td>
<td>-10.2</td>
<td>-18.4</td>
<td>-28.8</td>
<td>-31</td>
<td>-12</td>
</tr>
</tbody>
</table>