CHAPTER 5: SPATIAL & TEMPORAL CORRELATION

5.1 INTRODUCTION

Monitoring data collected at monitoring sites that are too close may vary in a similar way over time. In effect, this means that the data are "correlated" and that if this "spatial correlation" is high, resources may be being wasted because similar information is being obtained from the two sites.

Similarly, data collected at regular intervals from a particular monitoring site may be correlated if the sequential data points vary in a similar way. This is "temporal correlation". For example, if daily sampling is being performed and the data from any particular day varies similarly to that of the previous day, again resources are being wasted because the current day's data contain similar information to that of the previous day.

Both these situations should ideally be avoided. The number of monitoring sites and the sampling frequency impact directly on the total costs of any monitoring programme. On the one hand monitoring sites should be chosen that are sufficiently far apart to be independent. On the other hand, the sampling frequency should be sufficiently low to avoid sequential samples being correlated. The nature of a monitoring variable, in particular the random variability in its measurement, can also place a constraint on sampling frequency (or *vice versa*). In this case it may impose a maximum frequency (*i.e.* a minimum time between samples.) The following sections present a framework that addresses some of the issues.

5.2 NATURE OF MONITORING VARIABLES

5.2.1 Introduction

Before a final choice of monitoring variables can be made, the frequency of sampling must be determined to ensure that associated sampling and analytical costs are not excessive.

The process outlined below allows the estimation of an appropriate maximum frequency considering, primarily, the likely random variation in experimental measurements. Sensitivity and bias are not issues that are considered here (however, see the Null Hypotheses Chapter for the ramifications of these).

The variability determines, for example, the confidence with which statistics can be compared for any given data set. Alternatively, if a specific confidence is required for such comparisons, then this can give guidance on how many samples might be needed (*i.e.* the sampling frequency). A high variability means that more samples are required in order to make such comparisons at a given level of confidence than would be the case if the variability was much lower. This has cost implications.

In essence therefore, we need to ensure that the choice of monitoring variables (some of which may have relatively, and unavoidably, high variabilities) does not place unacceptable demands on the NTMP's ability to report statistically meaningful results.

5.2.2 Maximum monitoring frequency

It is the Department that is ultimately accountable for either action or inaction in response to NTMP reports. The Department is therefore the single most important stakeholder in deciding on acceptable probabilities of both false negative and false positive results. In arriving at these estimates, careful consideration should be given to the implications of the errors for each null hypothesis (as discussed in the chapter on statistical considerations).

Specifically, three factors need to be given attention:

- The acceptable probability of reporting a false negative result.
- The acceptable probability of reporting a false positive result.
- The "effect size". This is defined as the difference between the means being compared divided by the standard deviation (assumed equal for the two samples whose means are being compared).

If each of these can be defined, the table below can be used to obtain the necessary number of data points. Since annual statistics are being assumed for the NTMP, this essentially provides the monitoring frequency. For example, if twelve samples are required, monthly sampling would suffice. It should be noted that a mean may not be the most appropriate statistic to use (a median may be better). However, should means be chosen, then the following analysis could be used.

It was noted above that there exists (a) natural variability and (b) variability inherent in the experimental technique. Natural variability can usually only be quantified by actually measuring changes over long periods (at least one hydrological year). In the absence of information on natural variability, one can, as a preliminary exercise, assume the natural variation is zero. In other words assume all variation arises out of the experimental method.

Experimental variations are easier to quantify and are reasonably well known for many potential monitoring variables. One can then go through the exercise of establishing appropriate monitoring frequencies. Since natural variability has not been considered, these frequencies will be the absolute maximum that is necessary.

If real data on actual natural variations are not available, sensible estimates of this variation should be obtained. The same exercise can then be carried out to obtain an appropriate frequency.

Table 5.1. Recommended minimum number of samples per year as a function of required probabilities of false negative and false positive errors and effect size [Faul and Erdfelder, 1992].

Probability of false positive (%)	Power (β)	Two-tailed tests: Probability of false negative (%)			One-tailed tests: Probability of false negative (%)		
		5%	10%	20%	5%	10%	20%
		α=0.05	α=0.1	α=0.2	α=0.05	α=0.1	α=0.2
		Small effects (size = 0.2)					
5%	0.95	1302	1084	858	1084	858	620
10%	0.9	1054	858	658	858	658	452
20%	0.8	788	620	452	620	452	284
		Medium effects (size = 0.5)					
5%	0.95	210	176	138	176	138	100
10%	0.9	172	140	106	140	106	74
20%	0.8	128	102	74	102	74	46
		Large effects (size = 0.8)					
5%	0.95	84	70	56	70	56	40
10%	0.9	68	56	42	56	42	30
20%	0.8	52	42	30	42	30	20
		Very large effects (size = 1.0)					
5%	0.95	54	46	36	46	36	26
10%	0.9	46	36	28	36	28	20
20%	0.8	34	28	20	28	20	12
		Extremely large effects (size = 2.0)					
5%	0.95	16	14	10	14	10	8
10%	0.9	14	12	8	12	8	4
20%	0.8	12	8	4	8	4	4

As an example, assume the following:

- Acceptable probabilities of false negative and false positive results are 10%.
- The standard deviation representing the variability (from whatever source) is 10 units.
- We want to be able to distinguish meaningfully between means that are 20 units apart.
- The null hypothesis requires us to determine whether one mean is significantly different from another mean.

The effect size is 20/10 = 2. A two-tailed test is required because a difference between means is being assessed (not whether one is greater or less than the other). A monthly monitoring frequency is therefore acceptable (*i.e.* 12 per year).

If the standard deviation was 40 units (perhaps due to large variability) then the effect size would be 20/40 =0.5. So we would need 140 samples per year. This requires sampling almost every two days (2.6 to be more exact). Sampling every three days would not be sufficient.

If natural variability is very much higher than the variability due to experimental methods, then the above exercise will not necessarily restrict the choice of monitoring variables on the basis of their experimental variability (because natural variability will be the main factor to do so). In such a case the above exercise need not be undertaken (in respect of variability due to random errors in experimental method).

However, if experimental variability is equivalent to or higher than natural variability then limits placed on sampling frequency by available resources may restrict the choice of variables to those that are less variable.

5.3 TEMPORAL CORRELATION

Temporal correlation can only be addressed when real data are available. During the design phase of the project, data collected by the project focussing on endocrine disruptor compounds (funded by the Water Research Commission) may be used. Data at each monitoring site will need to be analysed separately to establish a minimum monitoring frequency (*i.e.* maximum time between sampling) required at each site to avoid temporal correlation. This information can be used to obtain better estimates of overall costs and frequencies to be used in the pilot studies.

The pilot studies that will follow this design phase should have as a major objective the collection of sufficient data to ensure sound datasets are obtained to get accurate estimates of minimum monitoring frequencies. Again, different frequencies may be calculated at the different monitoring sites. If the frequencies are similar in magnitude, it may be satisfactory to assume a single average frequency for the NTMP as a whole. However, if there is a large range of frequencies, it may be necessary to consider different frequencies for different monitoring variables (or possibly type of monitoring site). However, the logistical and managerial implications of such a decision would need to be carefully considered.

During the subsequent phased implementation phase it is also possible to examine the data that have been collected for temporal correlation. This should occur during the planned reviews of the monitoring programme as a whole. Again, the appropriateness of a single national average frequency can be weighed against the use of different frequencies under different circumstances.

5.4 SPATIAL CORRELATION

The existence of correlation between monitoring sites can only be established when sufficient data from the sites are available. It is conceivable that this will only be necessary once full-scale implementation has begun (after the pilot studies). Due to the inevitable phased implementation at increasing numbers of sites over the years, it is likely to be relatively easy to avoid spatial correlation initially (using common sense) by placing monitoring sites at significant distances from one another. The emphasis will initially be on obtaining data for a water management area (WMA) as a whole. Completely different river systems are likely to be chosen to achieve this. Therefore, spatial correlation is not highly likely (though should, nevertheless, be specifically borne in mind when choosing sites at this time). However, in subsequent years as more and more sites are added to each WMA to obtain better coverage, spatial correlation will naturally become increasingly likely. Therefore, an examination of spatial correlation should form part of the regular review of the NTMP.

5.5 COMPARISON OF APPROACHES

Statistical approaches are available that can provide information on optimum monitoring frequencies (*i.e.* that deal with temporal correlation) and minimising the number of sampling sites (*i.e.* that deal with spatial correlation). The following two approaches are compared briefly.

5.5.1 Entropy approach

This approach is based on a concept of "entropy" as being a measure of the uncertainty in random processes. Monitoring variables such as toxicant concentrations and degree of toxicity in natural waters can be subject to a wide variety of random processes that ultimately determine the value measured at the time of sampling. To this "natural variability" is then added the "experimental variability" as a result of random variations in the analytical method. Ozkul *et al.* (2000) note that if the primary objectives of the monitoring is to determine variability, then this method can be used to evaluate how informative the data are in time and space dimensions.

It might be noted here that it is not the primary objective of the NTMP to determine variability. It is the primary objective to determine status and trends. Status has been assumed to be represented by some kind of average (like an annual average). However, status could be re-interpreted to include some measure of variability (which could simply be a standard deviation). Nevertheless, such variability is important to the NTMP if only for the reason that the results reported to the water resources manager should reflect the degree of confidence. Typically, this depends on variability.

Ozkul *et al.* (2000) note the following advantages of the entropy method:

- It provides a quantitative measure of information content of any particular site and of an observed time series.
- It can assess the degree to which spatial and temporal correlation exists.
- It gives an indication of the usefulness of the data.
- It can simultaneously assess several features of a monitoring programme (*e.g.* sites, frequencies, variables, and duration).

Although these advantages seem impressive, they also note a major disadvantage of the method. It is sensitive to the choice of multivariate probability function that adequately represents the multivariate nature of the network. As a consequence, they recommend using different techniques in combination in order to investigate network features from different perspectives.

5.5.2 Principal components approach

Another general approach to determining the degree of spatial and temporal correlation involves determining the minimum number of "components" (in a mathematical sense) that adequately reproduces (or models) the total variability of the data. The components that do so are called the "principal components", implying that the other components are less important. This is equivalent to saying that these latter are correlated and therefore provide little or no independent information.

If components are identified as monitoring sites, then this allows redundant sites to be identified and discarded. The approach can also be used to establish temporal correlation at particular sites.

This method has a number of important advantages:

- It is well established and typically available in standard statistical software packages.
- Since it based on a so-called "covariance or correlation matrix", it does not depend on knowing the underlying statistical distributions of the data. (For example, normality need not be assumed.)

5.6 RECOMMENDATIONS

Accordingly, in summary, the following recommendations are made:

- For each monitoring variable chosen for the NTMP, the random experimental variability should be estimated, acceptable probabilities for false positives and false negatives chosen and typical effect sizes estimated. Based on these, the maximum monitoring frequency should be estimated from Table 5.1. It is recommended that the probability of false negatives be more stringent than for false positives.

- The presence of temporal and spatial correlation can only be assessed with real data. Until such time as these are available, common sense should prevail to ensure as far as possible that these are avoided.

- In future when such data become available, it is recommended that the principal components approach be used to assess the degree of correlation.