PREAMBLE

Inspiration and genius — one and the same.

Victor Hugo

This paper, in an ingenious way, mingles the themes of information content of data and the latent spatial dependence of geo–referenced data. To date most of the spatial statistics literature has been devoted to model specification issues, in order to handle or accommodate spatial dependence. But rather than adjusting computations in order for standard statistical tables to serve as proper references in a conventional way, Richardson exploits the redundant information aspect of geo–referenced data so that standard statistical tables can be used in a novel way. Accordingly, the purpose of this paper is to present a modified test of association based upon the correlation coefficient, and then evaluate its performance. One question Sen asks is whether or not these two approaches are equivalent? Richardson's presentation should inspire much subsequent research on this equivalence topic, as well as various extensions of the modification to other classical statistics. The notion of degrees of freedom as an index of information contained in data is employed in the implementation of Richardson's modification, which essentially filters out the redundant information seemingly quiescent in geo–referenced data by determining the correct degrees of freedom index; radical changes in this index can occur, as is indicated by tabulated results. But, except for extremely small degrees of freedom, the magnitude of the t–statistic does not change very much for a given probability level across its degrees of freedom range. This fact suggests that spatial dependence may make little difference to the drawing of inferences, other than in cases that are close to a selected critical value; fortunately other statistics, such as chi square, may better demonstrate Richardson's point. Another prominent drawback Sen notes is that this approach fails to acknowledge the impact spatial dependence has on important statistical properties of estimators, such as efficiency. Nevertheless, the elegance of Richardson's approach should open new areas of research in spatial statistics.

The Editor

Some Remarks on the Testing of Association Between Spatial Processes

Sylvia Richardson *

INSERM U.170 - 16 Avenue Paul Vaillant Couturier, 94807 Villejuif Cedex, France

Overview: This paper is concerned with the problem of testing for association between spatially defined variables exhibiting spatial autocorrelation. This problem occurs commonly in geography with variables typically defined as area averages of dichotomous or continuous variables recorded at the individual level. Classical statistical methods, such as correlation or regression analysis, are not directly applicable to the situation of spatially autocorrelated variables. A modified test of association based on the correlation coefficient is reviewed and results of its performance (Type I error, power, robustness to departure from normality) as well as that of a nonparametric measure of association, Tjøstheim's index, are given. This comparison shows the weak performance of the nonparametric index of association. For the case of several variables, the modified test can be extended to assess the significance of the partial correlation coefficient. Alternatively, regression analysis with the spatially parametrised error variance–covariance matrix can be performed. The two approaches are discussed and compared on some examples of geographical epidemiology concerning the relationship between lung cancer mortality and industrial factors.

1. Introduction

The problem of testing association between spatially defined variables occurs commonly in science and its applications. Standard examples are found in the fields of geography and regional science, for instance when relating consumption of agricultural output to road accessibility (Cliff and Ord, 1981). Upton and Fingleton (1985) discuss cases in ecology concerned with the relationships between, for example, plant species or island flora and locational or environmental characteristics. Examples in sociology and political science are described by Doreian (1981), with reference to voting behaviour and socio–economic or political factors. In epidemiology, etiological clues to environmental risk–factors are sometimes sought through their joint analysis with disease incidence or mortality maps (Doll, 1984; Armstrong and Doll, 1975).

As seen in these studies, the data consist of variables observed at different locations that can be considered as observations of stochastic processes exhibiting typically some spatial dependence (*i. e.*, dependence between variables indexed by nearby points). This dependence or *spatial autocorrelation* can arise in different ways. It might be an intrinsic characteristic of the process itself, such as the existence of interactions between the sites of a diffusion or of a contagious phenomenon. Spatial autocorrelation might also result indirectly from the influence on the variable considered of other related factors varying continuously

* Chantal Guihenneuc and Virginie Lasserre of the Laboratoire de Statistiques Médicales de l'Université de Paris V provided invaluable computing assistance. Evelyne Przybilski is warmly thanked for her secretarial assistance. This work was supported by a EURATOM contract # BI6-126F.

through space. Such is the case for variables defined as area averages of dichotomous variables recorded at the individual level. Even if there is *a priori* no direct inter–individual influence on a particular dichotomous variable, spatial autocorrelation between aggregated values will often be observed since spatially close individuals will often share some common influencing characteristics. The epidemiological examples that will be discussed in the final section of this paper belong to this category.

Classical statistical methods based on correlation or regression analysis are not directly applicable to the situation of spatially autocorrelated variables. The consequences of neglecting existing autocorrelation in regression analysis have been pointed out by Johnston (1972) for time series and Cliff and Ord (1981) for spatial series.

The problem of dealing with spatial autocorrelation in testing for association may be tackled in various ways. Time series methods proposed by Haugh (1976) such as the prewhitening of each series before analysis of the cross correlations could be extended to the spatial domain by considering appropriate filtering. Using this approach, Pierce (1977) found only weak evidence of a relationship between pairs of economic time series that were traditionally considered as related. In a subsequent paper, Geweke (1981) compares several tests of independence between stationary time series, in particular Haugh's test and an F test on the regression parameters of a mixed regressive–autoregressive model between the two series. He concludes that in many cases the proportion of Type II errors of Haugh's test is larger than that of F tests of the regression coefficients. For spatial series, no comparable study has been done to date although the prewhitening of series has been discussed by Griffith (1980). It is likely that a similar conclusion to Geweke would hold and that tests of independence or estimates of regression coefficients based on the original series will be more efficient than those based on their residuals after prewhitening.

Alternatively standard measures of association such as the correlation coefficient can be adapted to take account of autocorrelation. This approach has been developed by Clifford, Richardson and Hémon (1989) (CRH 89) and will be reviewed in Section 2.

New indices of association can also be proposed. A non–parametric index of association was developed by Tjøstheim (1978). A comparative study (Type I error and power) of this index and of the aforementioned modified test of the correlation coefficient will be presented in Section 3. Section 4 will complement this discussion by studying the robustness of the modified test of the correlation coefficient for departures from normality. Finally, classical regression analysis can be extended specifically to take into account the spatial structure of the data by modelling the variance–covariance error matrix. In Section 5 this approach is reviewed and a modified test of partial correlation, which is an extension to the case of several variables of the modified test of simple correlation, is presented. In addition the results given by different choices of models for the variance–covariance error matrix are compared with those given by the modified test of partial correlation on some examples of geographical epidemiology.

2. Modified test of association

We are interested in data sets that consist of a set $\mathbf{A}$ of $\mathbf{N}$ locations numbered from $\mathbf{1}$ to $\mathbf{N}$ and a set of pairs of observations $\{(\mathbf{X}_\alpha, \mathbf{Y}_\alpha), \alpha \varepsilon \mathbf{A}\}$, where each pair is indexed by its location. We shall use the following notation:

$$\bar{\mathbf{X}} = N^{-1}(\sum X_\alpha)$$
$$\mathbf{s_{XY}} = N^{-1} \sum (X_\alpha - \bar{X})(Y_\alpha - \bar{Y})$$
$$\mathbf{s_{X^2}} = N^{-1} \sum (X_\alpha - \bar{X})^2$$

similar expression can be written for $\bar{\mathbf{Y}}$ and $\mathbf{s_{Y^2}}$.

Modified tests of association based either on $\mathbf{s_{XY}}$, the empirical covariance between pairs of observations $\{(\mathbf{X}_\alpha, \mathbf{Y}_\alpha), \alpha \varepsilon \mathbf{A}\}$, or based on $\mathbf{r_{XY}}$, the corresponding empirical correlation coefficient, have been proposed by Clifford, Richardson and Hémon (1989). These tests rely on an estimation of the variance of $\mathbf{s_{XY}}$ and $\mathbf{r_{XY}}$ that takes the internal autocorrelations into account.

In the classical case where the elements of $\mathbf{Y}$ are normal i.i.d. random variables conditional on $\mathbf{X}$ or vice-versa, then $\mathbf{r_{XY}}$ has the standard null distribution with p.d.f.

$$\mathbf{f_N(r)} = (\mathbf{1} - \mathbf{r^2})^{1/2(\mathbf{N}-4)}/\mathbf{B}(\mathbf{1/2}, \mathbf{1/2}(\mathbf{N}-\mathbf{2})), \qquad |\mathbf{r}| \leq \mathbf{1}$$

where $\mathbf{B}$ is the beta function.

The expectation of $\mathbf{r_{XY}}$ is zero and its variance is equal to $(\mathbf{N}-\mathbf{1})^{-1}$. Critical values of $\mathbf{r_{XY}}$ are usually obtained from t-tables since $(\mathbf{N}-\mathbf{2})^{1/2}\mathbf{r}/(\mathbf{1}-\mathbf{r^2})^{1/2}$ has a t-distribution with $\mathbf{N}-\mathbf{2}$ degrees of freedom under these assumptions. We shall refer to this statistic as $t_{\mathbf{N}-\mathbf{2}}$. This is also the t-statistic that is calculated in testing the significance of the linear regression either of $\mathbf{Y}$ on $\mathbf{X}$, or of $\mathbf{X}$ on $\mathbf{Y}$.

2.1. The variance of $\mathbf{r_{XY}}$

Suppose now that $\mathbf{X}$ and $\mathbf{Y}$ are independent but that both $\mathbf{X}$ and $\mathbf{Y}$ are multivariate normal vectors with constant means and variance–covariance matrices $\mathbf{\Sigma_X}$ and $\mathbf{\Sigma_Y}$ respectively. The variance of $\mathbf{r_{XY}}$ is inflated by positive autocorrelation. This was shown asymptotically in the time-series context by Bartlett (1935) and in the spatial context by Richardson and Hémon (1981). This needs to be taken into account in the testing method.

It can be shown that, to the first order, the variance of $\mathbf{r_{XY}}$, $\sigma_\mathbf{r}^2$ is:

$$\sigma_\mathbf{r}^2 = \frac{\mathbf{var(s_{XY})}}{\mathbf{E(s_{X^2})E(s_{Y^2})}}, \tag{1}$$

and that this approximation is exact in some special cases. The variance of $\mathbf{s_{XY}}$ can be evaluated if some hypotheses on the spatial structure of $\mathbf{\Sigma_X}$ and $\mathbf{\Sigma_Y}$ are imposed. We suppose that pairs in $\mathbf{A} \times \mathbf{A}$ can be divided into strata $\mathbf{S_0}, \mathbf{S_1}, \ldots \mathbf{S_K}$ such that the covariances within strata remain constant, *i. e.*,

$$\mathbf{cov(X_\alpha, X_\beta) = C_X(k)} \quad \text{if} \quad (\alpha, \beta) \varepsilon \mathbf{S_k},$$

and

$$\mathbf{cov}(\mathbf{Y}_{\alpha}, \mathbf{Y}_{\beta}) = \mathbf{C_Y}(\mathbf{k}) \quad \text{if} \quad (\alpha,\beta)\,\varepsilon\,\mathbf{S_k},$$

with $S_0 = \{(\alpha,\alpha),\, \alpha\,\varepsilon\, A\}$. This formulation is flexible enough to permit non-isotropy or other aspects of inhomogeneity to be taken into account.

An estimate of the variance of $\mathbf{s_{XY}}$ is then derived:

$$\mathbf{N}^{-2} \sum_{\mathbf{k}} \mathbf{N_k} \hat{\mathbf{C}}_{\mathbf{X}}(\mathbf{k}) \hat{\mathbf{C}}_{\mathbf{Y}}(\mathbf{k}), \tag{2}$$

$\mathbf{N_k}$ is the number of pairs in strata $\mathbf{S_k}$ and $\hat{\mathbf{C}}_{\mathbf{X}}(\mathbf{k})$ [respectively $\hat{\mathbf{C}}_{\mathbf{Y}}(\mathbf{k})$] is the estimated autocovariance:

$$\hat{\mathbf{C}}_{\mathbf{X}}(\mathbf{k}) = \sum_{\mathbf{S_k}} (\mathbf{X}_{\alpha} - \bar{\mathbf{X}})(\mathbf{X}_{\beta} - \bar{\mathbf{X}})/\mathbf{N_k}$$

Thus the estimate (2) takes into account the autocorrelation of both $\mathbf{X}$ and $\mathbf{Y}$, and leads to an estimate of the variance of $\mathbf{r_{XY}}$.

$$\hat{\sigma}_{\mathbf{r}}^{2} = \sum [\mathbf{N_k} \hat{\mathbf{C}}_{\mathbf{X}}(\mathbf{k}) \hat{\mathbf{C}}_{\mathbf{Y}}(\mathbf{k})] / [\mathbf{N}^2 \mathbf{s_{X^2} s_{Y^2}}]. \tag{3}$$

In a simulation study reported in CRH 89, mutually independent autoregressive processes $\mathbf{X}$ and $\mathbf{Y}$ were generated on a 12-by-12 lattice. This allowed a comparison to be made between the asymptotic variance of $\mathbf{r_{XY}}$, the first order approximation (1), the average estimated $\hat{\sigma}_{\mathbf{r}}^{2}$ given by (3) and the empirical variance of $\mathbf{r_{XY}}$ over 4000 trials. The results are illustrated in Figure 1 and clearly show that the asymptotic value is far too large for highly autocorrelated processes. Note that in time series analysis a first assessment of the cross-correlogram is traditionally done *via* a similar asymptotic expression. For moderate autocorrelation there is little difference between the empirical and the average estimated variance given by (3). For high autocorrelation in each processes the estimated variance is consistently too low.

2.2. Modified tests

A modified t-test ($\mathbf{t}_{\widehat{\mathbf{M}}-\mathbf{2}}$) was proposed based on an estimated *effective sample size* $\widehat{\mathbf{M}}$, $\widehat{\mathbf{M}} = \mathbf{1} + \hat{\sigma}_{\mathbf{r}}^{-\mathbf{2}}$, that rejects the null hypothesis of no association when:

$$|(\widehat{\mathbf{M}} - \mathbf{2})^{\mathbf{1/2}} \mathbf{r} (\mathbf{1} - \mathbf{r}^{\mathbf{2}})^{-\mathbf{1/2}}| > \mathbf{t}^{\alpha}_{\widehat{\mathbf{M}}-\mathbf{2}} \tag{4}$$

where $\mathbf{t}^{\alpha}_{\widehat{\mathbf{M}}-\mathbf{2}}$ is the critical value of the t-statistic with $\widehat{\mathbf{M}} - \mathbf{2}$ d.f. The quantity $\widehat{\mathbf{M}}$ takes into account the spatial autocorrelation in the variables $\mathbf{X}$ and $\mathbf{Y}$ and is typically less than $\mathbf{N}$ for positively autocorrelated processes.

Equivalently a standardised covariance can be used:

$$\mathbf{W} = \mathbf{N s_{XY}} [\sum \mathbf{N_k} \hat{\mathbf{C}}_{\mathbf{X}}(\mathbf{k}) \hat{\mathbf{C}}_{\mathbf{Y}}(\mathbf{k})]^{-\mathbf{1/2}}$$

and tested as a standard normal relying upon central limit theorems for spatially dependent variables (Bolthausen, 1982; Guyon and Richardson, 1984).

Figure 1.

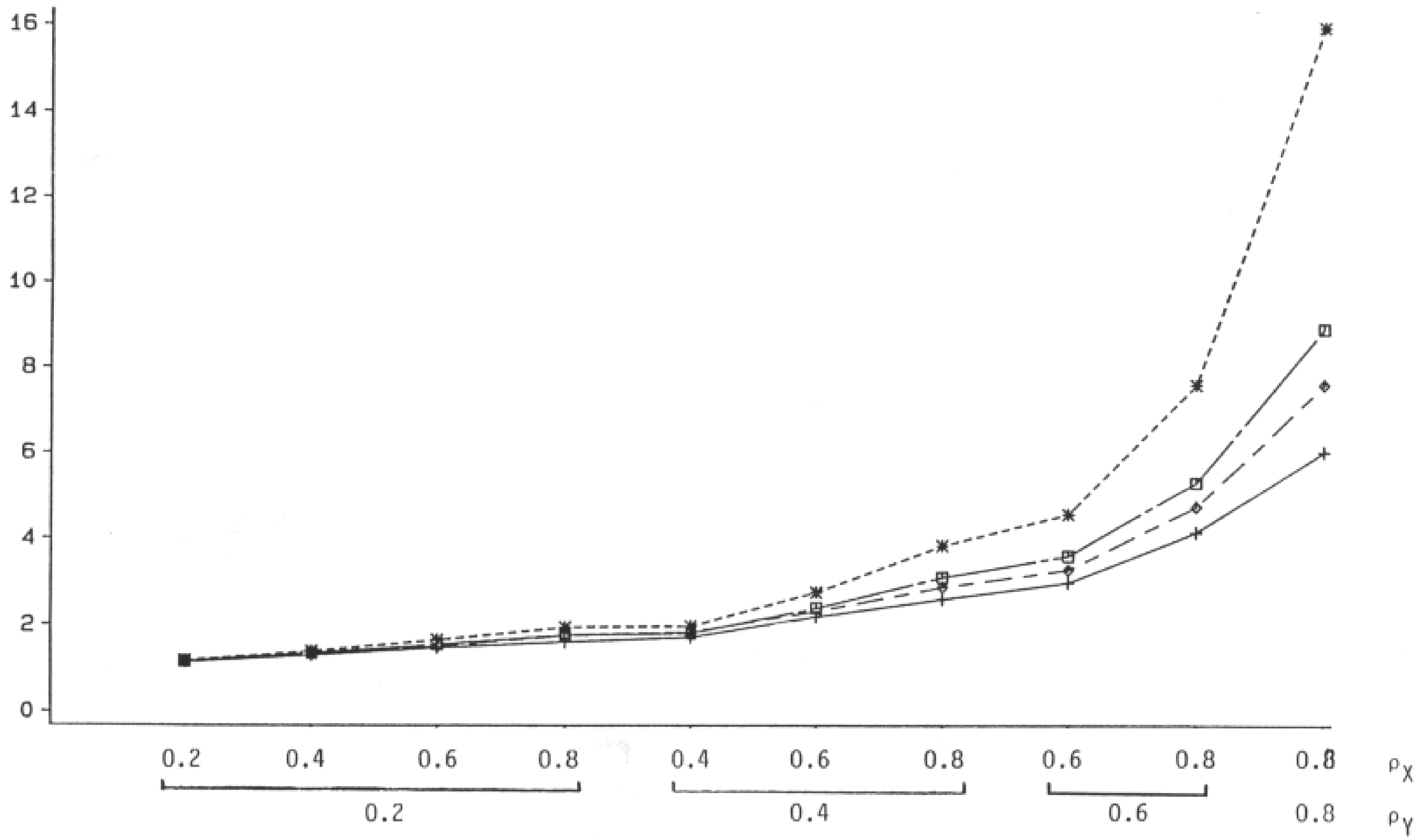

Figure 1 : The variance of r for two mutually independant simultaneous autoregressive processes on a 12x12 lattice : comparison of the asymptotic value (*-----------*), the first order approximation (1) (□- - - - □), the empirical variance of r over 4000 simulations (◊_____ ◊) the average v_r of $\hat{\sigma}_r^2$ given by (3) (+ _____+).

The values on the abscissa are $\rho_X(1)$ and $\rho_Y(1)$, the nearest neighbour autocorrelations for X and Y.

The Type I errors and the power of the modified tests were investigated in several simulation studies and were shown to be satisfactory (CRH 89, Richardson and Clifford, 1988). For small spatial domains, the Type I errors of the $t_{\widehat{M}-2}$ test are closer to their nominal value than those of the $\mathbf{W}$ test. Otherwise the two tests give equivalent results. Their performance will again be illustrated in the following section, which is concerned with a study of the performance of non–parametric tests in cases of spatial autocorrelation.

3. Comparison of non–parametric tests of association and the modified tests

A classical non–parametric measure of association between two variables is Spearman's rank correlation $\mathbf{r_S}$ which evaluates a correlation coefficient between the ranks of $\mathbf{X}_\alpha$ and $\mathbf{Y}_\alpha$. It can be calculated as

$$\mathbf{r_S} = 1 - \frac{6}{n(n^2-1)} \sum_{\alpha=1}^{N} d_\alpha^2,$$

where $\mathbf{d}_\alpha$ is the difference between the ranks of $\mathbf{X}_\alpha$ and $\mathbf{Y}_\alpha$.

A non–parametric spatial index of association was proposed by Tjøstheim (1978) and generalised by Hubert and Golledge (1982). It is based on the sum of the "distances" between locations of similar ranks for the variables $\mathbf{X}$ and $\mathbf{Y}$, with a suitably chosen distance function. The values $\{\mathbf{X}_\alpha, \alpha = 1, \ldots, N\}$ and $\{\mathbf{Y}_\alpha, \alpha = 1, \ldots, N\}$ are first ranked. Let $\mathbf{X}(i)$ and $\mathbf{Y}(i)$ denote respectively their i^{th} rank and let $\{k_X(i), l_X(i)\}$ and $\{k_Y(i), l_Y(i)\}$ denote the respective coordinates of $\mathbf{X}(i)$ and $\mathbf{Y}(i)$. Supposing that the coordinate system has been centered, Tjøstheim's index $\mathbf{A}$ can be written as:

$$\mathbf{A} = \frac{\sum_{i=1}^{N} \{k_X(i)k_Y(i) + l_X(i)l_Y(i)\}}{\sum_{i=1}^{N} \{k_X(i)^2 + l_X(i)^2\}}.$$

Note that the denominator of $\mathbf{A}$ is a constant of the coordinate system and does not depend on $\mathbf{X}$ (or $\mathbf{Y}$).

To evaluate the moments of $\mathbf{r_S}$ or $\mathbf{A}$ under the null distribution of no association between $\mathbf{X}$ and $\mathbf{Y}$, it is assumed in both cases that the $N!$ permutations of the values of one of the variables, the other staying fixed, are equally likely. For autocorrelated variables $\mathbf{X}$ and $\mathbf{Y}$, this is no longer true and the existing autocorrelations are perturbed by the permutations. As in the case of the Pearson correlation coefficient $\mathbf{r_{XY}}$, the variance of $\mathbf{r_S}$ and $\mathbf{A}$ are increased by positive autocorrelation whereas the variances classically used to test under the null hypothesis are assumed to be fixed. This leads to over–significant tests in the case of positive autocorrelation. A Monte Carlo simulation was carried out in order to evaluate this and to compare the power of the different tests.

Simulation model

In order to stay close to the examples that will be analysed in the last section, a simulation was carried out on an irregular grid of points defined by the administrative centres of the French *département* ($N = \mathbf{82}$). Spatial dependence was introduced directly on the variance–covariance matrix of the multivariate normal distributions by considering a disc model where the covariance between 2 points is defined as being proportional to the intersection area of 2 discs centered on those points (see §5.3 for a precise definition). The shape of the covariance function of this model shows a fairly linear decrease with distance. This shape is similar to the observed variograms; *i.e.*, the plot of

$$\mathbf{N}_{\mathbf{k}}^{-1} \sum_{(\alpha,\gamma)\,\epsilon\,\mathbf{S}_{\mathbf{k}}} (\mathbf{X}_{\alpha} - \mathbf{X}_{\gamma})^2$$

against the average distance between locations in $\mathbf{S}_{\mathbf{k}}$, of a number of variables that will be considered in our examples. The parameter of the disc model is chosen so that the autocorrelation for a distance of 40 km between points is equal to 0.2, 0.4, 0.6 or 0.8. We denote this autocorrelation by $\rho(\mathbf{1})$ indexed by the name of the variable.

For each chosen value of $\rho_{\mathbf{X}}(\mathbf{1})$ an $\mathbf{N}$-by-$\mathbf{N}$ matrix $\mathbf{\Sigma_X}$ is generated, with diagonal elements equal to 1, following the disc model. $\mathbf{\Sigma_X}$ is then triangularised, $\mathbf{\Sigma_X} = \mathbf{LL^t}$, and a realisation of $\mathbf{X}$ distributed as $\mathbf{N}(\mathbf{0}, \mathbf{\Sigma_X})$ is obtained by first generating a vector of $\mathbf{N}$ i.i.d. $\mathbf{N}(\mathbf{0}, \mathbf{1})$ and then pre-multiplying this vector by L.

The distances between the centres of *départements* were partitioned into 15 classes of 50 km intervals each. This gives 15 strata $\mathbf{S}_1, \ldots, \mathbf{S}_{15}$; the stratum $\mathbf{S}_0 = \{(\alpha, \alpha), \alpha\,\epsilon\,\mathbf{A}\}$. These strata are used in the calculation of $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ as defined in §2 formula (4).

3.1. Type I error for $\mathbf{r_S}$ and A

Exhibit 1 gives the observed percentages of Type I errors together with their 95% confidence limits for 4 statistics: $\mathbf{t}_{\widehat{\mathbf{M}}-2}$, $\mathbf{t}_{\mathbf{N}-2}$, $\mathbf{r_S}$ and $\mathbf{A}$, testing at a 5% nominal level the association between $\mathbf{X}$ and $\mathbf{Y}$ under the null hypothesis of independence between $\mathbf{X}$ and $\mathbf{Y}$. Five hundred simulations were done for each combination of $\rho_{\mathbf{X}}(\mathbf{1})$ and $\rho_{\mathbf{Y}}(\mathbf{1})$. The performance of the modified $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ statistic is satisfactory as the value 5% belongs to all the confidence intervals and there is no systematic variation with increasing autocorrelation. On the other hand, the Type I errors of the non-adjusted test $\mathbf{t}_{\mathbf{N}-2}$ and $\mathbf{r_S}$ are clearly increasing with increasing autocorrelation, reaching values around 30% instead of the nominal 5% in the most highly autocorrelated case. This result had already been observed (CRH 89) and here we see also that, not unexpectedly, the behaviours of $\mathbf{t}_{\mathbf{N}-2}$ and $\mathbf{r_S}$ are quite similar since they both use $\mathbf{N}-\mathbf{2}$ degrees of freedom.

For Tjøstheim's index $\mathbf{A}$, the influence of the autocorrelation on the significance level is less important. The difference with the nominal 5% level is only clearly apparent in the most highly autocorrelated case (0.8×0.8) with Type I errors nearly tripled. We checked the empirical variance of $\mathbf{A}$ in this case and found it to be around 1.6 times its theoretical value. Hence as for $\mathbf{r_S}$, Tjøstheim's index $\mathbf{A}$ leads to over-significant tests but only in the presence of high autocorrelation in both the variables.

3.2. Comparison of the powers of the modified $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ test and of the index A

The discussion of the comparative performances of Spearman's $\mathbf{r_S}$ and Tjøstheim's index $\mathbf{A}$ in terms of a general cross-product statistic between a measure of spatial proximity and a measure of non-spatial proximity was initiated by Glick (1982), and taken up by Hubert and Golledge (1982) and Upton and Fingleton (1985). This formulation is helpful in highlighting the contrast between the index $\mathbf{A}$ and $\mathbf{r_S}$. The index $\mathbf{A}$ uses a relatively sophisticated measure of spatial proximity, the distances between locations of similar ranks, but a $0-1$ classification of non-spatial proximity in terms of identical ranks. On the other hand, $\mathbf{r_S}$

Exhibit 1.

Proportion of type I errors and 95 % confidence limits for the modified $t_{\hat{M}-2}$ test, the standard t_{N-2} test, Spearman's r_S and Tjøstheim's index A, for a nominal 5 % test level in the case of two mutually independent processes generated by a disc model on a network of 82 points.

$\rho_Y(1)$	$\rho_X(1)$	0.0	0.2	0.4	0.6	0.8
0.0	$t_{\hat{M}-2}$	3.4 ± 1.6				
	t_{N-2}	3.2 ± 1.5				
	r_S	4.4 ± 1.8				
	A	4.0 ± 1.7				
0.2	$t_{\hat{M}-2}$	6.0 ± 2.1	6.4 ± 2.1			
	t_{N-2}	6.2 ± 2.1	7.4 ± 2.3			
	r_S	6.2 ± 2.1	5.8 ± 2.1			
	A	4.6 ± 1.8	4.8 ± 1.9			
0.4	$t_{\hat{M}-2}$	4.6 ± 1.8	5.8 ± 2.1	4.4 ± 1.8		
	t_{N-2}	5.6 ± 2.0	7.0 ± 2.2	8.0 ± 2.4		
	r_S	6.0 ± 2.1	7.0 ± 2.2	8.2 ± 2.4		
	A	5.4 ± 2.0	6.2 ± 2.1	5.4 ± 2.0		
0.6	$t_{\hat{M}-2}$	5.6 ± 2.0	4.4 ± 1.8	5.0 ± 1.9	4.8 ± 1.9	
	t_{N-2}	5.8 ± 2.1	5.8 ± 2.1	9.8 ± 2.6	13.6 ± 3.0	
	r_S	5.0 ± 1.9	6.2 ± 2.1	9.0 ± 2.5	16.8 ± 3.3	
	A	4.6 ± 1.8	5.6 ± 2.0	4.2 ± 1.8	6.4 ± 2.1	
0.8	$t_{\hat{M}-2}$	4.4 ± 1.8	5.0 ± 1.9	4.2 ± 1.8	5.2 ± 1.9	3.0 ± 1.5
	t_{N-2}	5.2 ± 1.9	7.8 ± 2.4	11.4 ± 2.8	20.2 ± 3.5	35.6 ± 4.2
	r_S	4.6 ± 1.8	9.4 ± 2.6	12.0 ± 2.8	18.8 ± 3.4	34.4 ± 4.2
	A	5.0 ± 1.9	4.2 ± 1.8	8.8 ± 2.5	6.4 ± 2.1	14.0 ± 3.0

The parameters $\rho_X(1)$ and of the disc models for X and Y respectively are equal to the autocorrelations at 40 km. 500 simulations were carried out in each cell.

only considers a $0-1$ measure of spatial proximity (identical locations or not) and a less crude measure of non–spatial proximity, the square of the differences between the ranks. Extension of the index $\mathbf{A}$ to a statistic that equally involves both types of proximity has been proposed by Hubert and Golledge.

We now discuss the results of a Monte Carlo study aimed at comparing the power of the modified $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ test and of the index $\mathbf{A}$ under two sets of alternative hypotheses:

$\mathbf{H_1}$: $\mathbf{Y}$ and $\mathbf{X}$ are related by a linear relationship with autocorrelated error term,

$\mathbf{H_2}$: $\mathbf{Y}$ and $\mathbf{X}$ are related by a local permutation with a random disturbance,

for different levels of autocorrelation in $\mathbf{X}$. The comparison is limited to these two statistics because their significance levels can be controlled and fixed at around 5% (except for $\mathbf{A}$ in one case of very high autocorrelation). Comparison with $\mathbf{r_S}$ is consequently not possible except in unautocorrelated cases, which are of little interest here. Two contrasting alternative hypotheses were chosen: $\mathbf{H_1}$ in the classical framework of linear regression particularly suited to the modified $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ test of the correlation coefficient and $\mathbf{H_2}$ as a local rearrangement particularly adapted to the index $\mathbf{A}$.

3.2.1. Observed power of $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ *and* $\mathbf{A}$ *under* $\mathbf{H_1}$

The alternative hypothesis of linear regression between $\mathbf{Y}$ and $\mathbf{X}$ was defined as:

$$\mathbf{H_1}: \mathbf{Y} = \mathbf{aX} + \mathbf{W}, \quad \mathbf{X} \sim \mathbf{N}(\mu_\mathbf{X}, \Sigma_\mathbf{X}), \quad \mathbf{W} \sim \mathbf{N}(\mu_\mathbf{W}, \Sigma_\mathbf{W})$$

and $\mathbf{X}$ and $\mathbf{W}$ independent. It is difficult to calculate theoretically the power of the modified test $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ or that of $\mathbf{A}$ because their distribution under $\mathbf{H_1}$ is not precisely known. Their power can be assessed by simulations.

Two independent spatially autocorrelated processes $\mathbf{X}$ and $\mathbf{W}$ were generated on the grid of the administrative centres of French *départements* as Gaussian variables with a disc model for their autocovariance. Without loss of generality $\sigma^2_\mathbf{X} = \sigma^2_\mathbf{W}$ was chosen and hence the correlation $\rho_{\mathbf{XY}}$ between $\mathbf{X}$ and $\mathbf{Y}$ was only dependent on the parameter $\mathbf{A}$. Five hundred trials were carried out for several levels of autocorrelation in $\mathbf{X}$ and $\mathbf{W}$ and for the values $\rho_{\mathbf{XY}} = 0.2$ and 0.4. The grid contained $\mathbf{N} = \mathbf{82}$ points. Results for higher values of $\rho_{\mathbf{XY}}$ are not reported because the power of the $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ was very close to 1. In these simulations a 5% nominal level was chosen. The results are presented in Exhibit 2.

With respect to $\mathbf{t}_{\widehat{\mathbf{M}}-2}$, one can see that the observed power decreases when there is strong autocorrelation in the $\mathbf{X}$ or the $\mathbf{W}$ variable. This is to be expected since the variance of $\mathbf{r_{XY}}$ increases and the effective degrees of freedom diminish with increasing autocorrelation. In Richardson and Clifford (1988), the power of $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ was shown to be quite close to a reference value given by the power of the classical t–test based on a number of observations compatible with the observed variance of $\mathbf{r_{XY}}$. The observed power of the $\mathbf{W}$ statistic is not shown because it is nearly identical to that of $\mathbf{t}_{\widehat{\mathbf{M}}-2}$. A theoretical approximation of the power of $\mathbf{W}$ was also given in this earlier paper.

On the other hand, the power of Tjøstheim's index $\mathbf{A}$ under $\mathbf{H_1}$ is very low. Even in the case $\rho_{\mathbf{XY}} = 0.4$ where the power of $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ is in most cases over 80%, that of the index $\mathbf{A}$ does not exceed its significance level. Indeed it is around 5% in most cases and only reaches

Exhibit 2.

(continued)

Observed power and 95 % confidence intervals for the modified $t_{\hat{M}-2}$ test and Tjostheim's index A under an alternative hypothesis of linear model : Y = aX + W where both X and W follow a disc model, X and W independent and of equal variance and a is chosen so that the correlation ρ_{XY} between X and Y takes the value 0.2 or 0.4.

		$\rho_x(1) = 0$		$\rho_x(1) = 0.2$		$\rho_x(1) = 0.4$		$\rho_x(1) = 0.6$		$\rho_x(1) = 0.8$	
$\rho_W(1)$		$\rho_{XY} = 0.2$	$\rho_{XY} = 0.4$	$\rho_{XY} = 0.2$	$\rho_{XY} = 0.4$	$\rho_{XY} = 0.2$	$\rho_{XY} = 0.4$	$\rho_{XY} = 0.2$	$\rho_{XY} = 0.4$	$\rho_{XY} = 0.2$	$\rho_{XY} = 0.4$
0.4	$t_{\hat{M}-2}$	**44.6** [40.2-49.0]	**95.4** [93.6 - 97.2]	**36.8** [32.6-41.0]	**93.0** [90.8-95.2]	**37.6** [33.4-41.8]	**91.0** [88.5-93.5]	**34.4** [30.2-38.6]	**88.6** [85.8-91.4]	**21.4** [17.8-25.0]	**79.2** [75.6-82.8]
	A	**5.4** [3.4-7.4]	**6.0** [3.9 - 8.1]	**6.4** [4.3-8.5]	**5.6** [3.6- 7.6]	**5.6** [3.6 - 7.6]	**7.4** [5.1 - 9.7]	**5.2** [3.3 - 7.1]	**5.4** [3.4 - 7.4]	**7.6** [5.3 - 9.9]	**11.0** [8.3 - 13.7]
0.6	$t_{\hat{M}-2}$	**48.4** [44.0-52.8]	**96.4** [94.8-98.0]	**36.4** [32.2-40.6]	**94.8** [92.9-96.7]	**36.2** [32.0-40.4]	**93.0** [90.8-95.2]	**28.0** [24.1-31.9]	**81.0** [77.6-84.4]	**20.0** [16.5-23.5]	**56.4** [52.1-60.7]
	A	**3.8** [2.1-5.5]	**3.8** [2.1 - 5.5]	**4.4** [2.6-6.2]	**4.4** [2.6- 6.2]	**4.6** [2.8 - 6.4]	**5.0** [3.1 - 6.9]	**4.4** [2.6- 6.2]	**6.0** [3.9 - 8.1]	**9.2** [6.7 - 11.7]	**11.0** [8.3 - 13.7]
0.8	$t_{\hat{M}-2}$	**52.4** [48.0-56.8]	**96.8** [95.3 - 98.3]	**48.4** [44.0-52.8]	**93.4** [91.2-95.6]	**42.4** [38.1-46.7]	**89.4** [86.7-92.1]	**28.6** [24.6-32.6]	**76** [72.3-79.7]	**12.0** [9.2 - 14.8]	**41.8** [37.5-46.1]
	A	**5.0** [3.1-6.9]	**5.4** [3.4 - 7.4]	**6.4** [4.3-8.5]	**4.2** [2.4- 6.0]	**7.2** [4.9- 9.5]	**4.8** [2.9- 6.7]	**4.6** [2.8 - 6.4]	**7.4** [5.1 - 9.7]	**13.0** [10.1-15.9]	**15.4** [12.2-18.6]

The parameters $\rho_X(1)$ and $\rho_W(1)$ of the disc models for X and W are equal to the autocorrelation at 40 km. 500 simulations were carried out in each cell.

Exhibit 2.

Observed power and 95 % confidence intervals for the modified $t_{\hat{M}-2}$ test and Tjøstheim's index A under an alternative hypothesis of linear model : Y = aX + W where both X and W follow a disc model, X and W independent and of equal variance and a is chosen so that the correlation ρ_{XY} between X and Y takes the value 0.2 or 0.4.

		$\rho_x(1) = 0$		$\rho_x(1) = 0.2$		$\rho_x(1) = 0.4$		$\rho_x(1) = 0.6$		$\rho_x(1) = 0.8$	
$\rho_W(1)$		$\rho_{XY} = 0.2$	$\rho_{XY} = 0.4$	$\rho_{XY} = 0.2$	$\rho_{XY} = 0.4$	$\rho_{XY} = 0.2$	$\rho_{XY} = 0.4$	$\rho_{XY} = 0.2$	$\rho_{XY} = 0.4$	$\rho_{XY} = 0.2$	$\rho_{XY} = 0.4$
0.0	$t_{\hat{M}-2}$	**42.4** [38.1-46.7]	**96.6** [95 - 98.2]	**44.2** [39.8-48.6]	**96.0** [94.3-97.7]	**43.2** [38.9-47.5]	**96.6** [95.0-98.2]	**39.2** [34.9-43.5]	**92.2** [89.8-94.6]	**35.8** [31.6 - 40.0]	**87.2** [84.3-90.1]
	A	**5.8** [3.8-7.8]	**5.6** [3.6 - 7.6]	**7.4** [5.1-9.7]	**6.0** [3.9- 8.1]	**5.8** [3.8 - 7.8]	**5.0** [3.1 - 6.9]	**4.8** [2.9-6.7]	**6.8** [4.6 - 9.0]	**7.6** [5.3 - 9.9]	**10.0** [7.4 - 12.6]
0.2	$t_{\hat{M}-2}$	**42.4** [38.1-46.7]	**97.4** [96.0 - 98.8]	**42.2** [37.9-46.5]	**95.6** [93.8-97.4]	**39.0** [34.7-43.3]	**94.0** [91.9-96.1]	**34.6** [30.4-38.8]	**9.2** [89.6-94.4]	**30.4** [26.4 - 34.4]	**83.6** [80.4-86.8]
	A	**4.8** [2.9-6.7]	**5.4** [3.4 - 7.4]	**6.2** [4.1-8.3]	**6.8** [4.6- 9.0]	**3.8** [2.1 - 5.5]	**5.6** [3.6 - 7.6]	**4.2** [2.4 - 6.0]	**5.4** [3.4 - 7.4]	**6.2** [4.1 - 8.3]	**9.4** [6.8 - 12.0]

The parameters $\rho_X(1)$ and $\rho_W(1)$ of the disc models for X and W are equal to the autocorrelation at 40 km. 500 simulations were carried out in each cell.

10 to 15% in cases of high autocorrelation in $\mathbf{X}$, where results reported in Exhibit 1 show that its significance level also increases. This is the reason for the apparent increase of power of the index $\mathbf{A}$ with autocorrelation that has also been noted by Glick (1982), although he did not relate it to Type I errors.

Hence, as expected from the construction of the two statistics $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ and $\mathbf{A}$, $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ has good power in the regression framework whilst the index $\mathbf{A}$ is clearly unable to recognize that type of association if there is any random error.

3.2.2. Observed power of $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ *and* $\mathbf{A}$ *under* $\mathbf{H_2}$

The second alternative hypothesis was constructed as follows. The 82 *départements* were separated into 16 groups of contiguous *départements* (composed of from 4 to 7 *départements* to balance their total surface area). The process $\mathbf{X}$ was generated with unit variance following a disc model, and then its values were randomly permuted within each of the 16 groups. This created a variable $\widetilde{\mathbf{X}}$ deduced from $\mathbf{X}$ by a local permutation. The variable $\mathbf{Y}$ was then defined as: $\mathbf{Y} = \widetilde{\mathbf{X}} + \mathbf{dW}$, where $\mathbf{W}$ are i.i.d. $\mathrm{N}(\mathbf{0}, \mathbf{1})$ variables and $\mathbf{0} \leq \mathbf{d} \leq \mathbf{1}$. Five hundred simulations were carried out for several values of $\rho_{\mathbf{X}}(\mathbf{1})$ and $\mathbf{d}$. The results are presented in Exhibit 3. Several observations can be made:

(a) When there is no autocorrelation, the variable $\widetilde{\mathbf{X}}$ is considered as independent of $\mathbf{X}$ by the statistic $\mathbf{t}_{\widehat{\mathbf{M}}-2}$, which only looks at association at the same site. The index $\mathbf{A}$ has maximum power when $\mathbf{d} = \mathbf{0}$, but its power decreases dramatically as soon as there is a random disturbance, even of small variance. For instance, it reaches 1/4 of its value when the variance of $\mathbf{dW}$ is $(\mathbf{1/4})^{\mathbf{2}}$ of the variance of $\widetilde{\mathbf{X}}$.

(b) When $\mathbf{X}$ is autocorrelated, then the power of $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ is substantially increased, being not too far from that of $\mathbf{A}$ when $\mathbf{d} = \mathbf{0}$ and $\rho_{\mathbf{X}}(\mathbf{1}) = 0.8$ and always higher than $\mathbf{A}$ even when $\mathbf{d}$ is small.

In conclusion, it is apparent that even under an alternative hypothesis well adapted to the index $\mathbf{A}$, the power of $\mathbf{A}$ is weak as soon as a more realistic situation including random disturbances is analysed. This considerably limits the interest of using the index $\mathbf{A}$ to detect a spatial shift. On the other hand, when there was autocorrelation in the $\mathbf{X}$ variable, the statistic $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ had a reasonable power to detect a local spatial rearrangement. One could conjecture that an extension of the $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ to *spatially lagged cross–correlations* (*i. e.*, between $\mathbf{X}_\alpha$ and $\mathbf{Y}_\beta$, β in a neighbourhood of α) would prove more powerful.

4. Robustness of the modified tests of association

The construction of the modified tests as well as the study of their performance was done under Gaussian hypotheses. In this section we investigate some aspects of the robustness of these tests to departure from normality. Three types of departure from normality will be considered, namely

(a) truncated Gaussian variables

(b) lognormal variables

(c) mixtures of Gaussian variables.

For each of these cases the Type I errors of $\mathbf{t}_{\widehat{\mathbf{M}}-2}$, $\mathbf{t}_{\mathbf{N}-2}$, and $\mathbf{A}$ are evaluated by Monte Carlo simulations with 500 trials and at a 5% nominal level of significance. In (a) mutually

Exhibit 3.

Observed power and 95 % confidence limits for the modified $t_{\hat{M}-2}$ test and Tjøstheim's index A under an alternative hypothesis of linear model : $Y = \tilde{X} + dW$ (see §3.2.2) where X is generated by a disc model and W are i.i.d N(0,1).

	$\rho_X(1)$	0.0	0.2	0.4	0.6	0.8
	$\hat{C}_Y(k)$	- 0.02	0.01	0.05	0.13	0.33
d=0.0	$t_{\hat{M}-2}$	5.6 ± 2.0	6.6 ± 2.2	10.6 ± 2.7	33.2 ± 4.1	79.2 ± 3.6
	A	100	100	100	100	100
	$\hat{C}_Y(k)$	- 0.02	0.01	0.04	0.12	0.32
d=0.125	$t_{\hat{M}-2}$	5.2 ± 1.9	6.0 ± 2.1	10.6 ± 2.7	32.4 ± 4.1	80.2 ± 3.5
	A	52.2 ± 4.4	54.6 ± 4.4	54.8 ± 4.4	61.0 ± 4.3	73.8 ± 3.9
	$\hat{C}_Y(k)$	- 0.02	0.01	0.04	0.10	0.30
d=0.25	$t_{\hat{M}-2}$	5.8 ± 2.0	5.6 ± 2.0	10.8 ± 2.7	32.8 ± 4.1	78.4 ± 3.6
	A	25.4 ± 3.8	22.6 ± 3.7	26.8 ± 3.9	26.6 ± 3.9	48.8 ± 4.4
	$\hat{C}_Y(k)$	- 0.02	0.0	0.03	0.09	0.24
d=0.5	$t_{\hat{M}-2}$	5.0 ± 1.9	6.2 ± 2.1	9.8 ± 2.6	28.2 ± 3.9	72.6 ± 3.9
	A	10.0 ± 2.6	12.2 ± 2.9	10.4 ± 2.7	16.8 ± 3.3	38.4 ± 4.3
	$\hat{C}_Y(k)$	- 0.02	- 0.01	0.01	0.05	0.14
d =1	$t_{\hat{M}-2}$	5.8 ± 2.0	5.6 ± 2.0	8.6 ± 2.5	18.2 ± 3.4	51.2 ± 4.4
	A	4.6 ± 1.8	7.2 ± 2.3	7.8 ± 2.4	9.4 ± 2.6	22.0 ± 3.6

The parameters $\rho_X(1)$ and $\rho_W(1)$ of the disc models for X and W are equal to the autocorrelation at 40 km and $\hat{C}_Y(1)$ is the overage observed autocorrelation of Y in the first strata S_1 defined in §3. 500 simulations were carried out in each cell.

independent variables $\mathbf{X}$ and $\mathbf{Y}$ were generated following a disc model, an then their values were set equal to the chosen truncation limits if they exceeded these limits. In (b) mutually independent autocorrelated variables $\mathbf{U}$ and $\mathbf{V}$ were first generated by a disc model based upon i.i.d. $\mathbf{N(0, \ln(2))}$ variables [rather than $\mathbf{N(0,1)}$], and then $\mathbf{X}$ and $\mathbf{Y}$ were defined as follows:

$$\mathbf{X} = [\ \mathbf{exp(U)}/\sqrt{\mathbf{2}}] - \mathbf{1},$$

and

$$\mathbf{Y} = [\ \mathbf{exp(V)}/\sqrt{\mathbf{2}}] - \mathbf{1}.$$

This definition implies that $\mathbf{X}$ and $\mathbf{Y}$ follow centered lognormal distributions, with unit variance. The correlation between $\mathbf{X}_\alpha$ and $\mathbf{X}_\beta$ is equal to $2^{\rho_U(\alpha,\beta)} - 1$, with

$$\rho_{\mathbf{U}}(\alpha, \beta) = \mathbf{E}(\mathbf{U}_\alpha \mathbf{U}_\beta).$$

In (c) mutually independent autocorrelated variables $\mathbf{U}$ and $\mathbf{V}$ were first generated by a disc model, and then the variable $\mathbf{X}$ was defined as equaling $\mathbf{U}$ with probability γ and $\mathbf{bU}$ with probability $\mathbf{1} - \gamma$, $\mathbf{0} \leq \gamma \leq \mathbf{1}$; $\mathbf{Y}$ was defined similarly with respect to $\mathbf{V}$. The variance of $\mathbf{X}$ thus was equal to $[\gamma + \mathbf{b^2(1} - \gamma)]$, and the correlation between $\mathbf{X}_\alpha$ and $\mathbf{X}_\beta$ was equal to:

$$\frac{[\gamma + \mathbf{b(1} - \gamma)]^{\mathbf{2}}}{[\gamma + \mathbf{b^2(1} - \gamma)]} \cdot \rho_{\mathbf{U}}(\alpha, \beta).$$

Results are presented in Exhibit 4 for several symmetric truncation levels, in Exhibit 5 for the lognormal case, and in Exhibit 6 for several mixture coefficients γ and $b = 3$. The observed Type I errors are, on the whole, not much different from those appearing in Exhibit 1. For $\mathbf{t}_{\widehat{\mathbf{M}}-\mathbf{2}}$ the significance levels are not altered in the lognormal case (b), but there is a slight tendency toward over–conservative levels in cases of higher autocorrelations with some Type I errors being below 5% in cases (a) or (c). The results are nearly identical for W. Non–symmetric truncations also were tried and the results were found to be similar.

In case (c), mixtures with $b = 6$, 9 or 12 also were simulated for highly autocorrelated $\mathbf{U}$ and $\mathbf{V}$ processes (see Exhibit 7). The tendency of $\mathbf{t}_{\widehat{\mathbf{M}}-\mathbf{2}}$ to be over–conservative again was apparent, with Type I errors around 2% in nearly all instances. Note that when the coefficient $\mathbf{b}$ is increased, the internal autocorrelation is decreased, with this balancing effect rendering comparisons awkward; other combinations of $\mathbf{b}$ and γ appear to be less informative. For $\mathbf{t}_{\mathbf{N}-\mathbf{2}}$ and the index $\mathbf{A}$, the results are comparable to those presented in Exhibit 1, with high autocorrelation resulting in inflated significance levels.

In conclusion, the $\mathbf{t}_{\widehat{\mathbf{M}}-\mathbf{2}}$ test is shown to be quite robust to small departures from normality in terms of its significance level, with a tendency to give over–conservative test results in situations of high autocorrelation. It would be interesting to develop a family of permutation tests based upon a restricted set of permutations that would preserve some aspect of the spatial structure and to compare the performance of those tests to that of the modified $\mathbf{W}$ and $\mathbf{t}_{\widehat{\mathbf{M}}-\mathbf{2}}$ tests.

Exhibit 4.

Proportion of type I errors and 95 % confidence limits for the modified $t_{\hat{M}-2}$ test and Tjøstheim's index A, for a nominal 5 % test level in the case of two mutually independent processes generated by a disc model on a network of 82 points and truncated at chosen limits.

Truncations limits		$\rho_Y(1) = \rho_X(1)$		
		0.0	0.4	0.8
[- 1.64 ; 1.64]	$t_{\hat{M}-2}$ t_{N-2} A	4.8 ± 1.9 4.2 ± 1.8 5.4 ± 2.0	5.8 ± 2.0 8.2 ± 2.4 6.4 ± 2.1	3.2 ± 1.5 36.4 ± 4.2 8.8 ± 2.5
[- 1.28 ; 1.28]	$t_{\hat{M}-2}$ t_{N-2} A	4.4 ±1.8 4.6 ± 1.8 8.4 ± 2.4	5.0 ± 1.9 8.2 ± 2.4 8.6 ± 2.5	3.0 ± 1.5 35.6 ± 4.2 10.8 ± 2.7
[- 1.04 ; 1.04]	$t_{\hat{M}-2}$ t_{N-2} A	4.6 ±1.8 4.2 ± 1.8 7.2 ± 2.3	5.4 ± 2.0 8.0 ± 2.4 8.8 ± 2.5	3.4 ± 1.6 34.0 ± 4.2 7.8 ± 2.4
[- 0.84 ; 0.84]	$t_{\hat{M}-2}$ t_{N-2} A	4.6 ±1.8 4.6 ± 1.8 10.2 ± 2.7	4.6 ± 1.8 7.8 ± 2.4 7.8 ± 2.4	3.6 ± 1.6 31.6 ± 4.1 10.0 ± 2.6

The parameters $\rho_X(1)$ and $\rho_Y(1)$ of the disc models for X and Y are equal to the autocorrelations at 40 km. 500 simulations were carried out in each cell.

Exhibit 5.

Proportion of type I errors and 95 % confidence intervals for the modified $t_{\hat{M}-2}$ test, the standard t_{N-2} test and Tjøstheim's index A, for a nominal 5 % test level in the case of two mutually independent lognormal processes X and Y generated from processes U and V following a disc model on a network of 82 points.

$\rho_V(1) = \rho_U(1)$	0.0	0.2	0.4	0.6	0.8
$2\rho(1) - 1$	0	0.15	0.32	0.52	0.74
$t_{\hat{M}-2}$	**5.2** [3.3 - 7.1]	**5.6** [3.6 - 7.6]	**6.2** [4.1 - 8.3]	**5.8** [3.8 - 7.8]	**5.2** [3.3 - 7.1]
t_{N-2}	**5.0** [3.1 - 6.9]	**6.0** [3.9 - 8.1]	**7.2** [4.9 - 9.5]	**11.6** [8.8 - 14.4]	**29.4** [25.4 - 33.4]
A	**4.4** [2.6 - 6.2]	**6.2** [4.1 - 8.3]	**5.8** [3.8 - 7.8]	**6.0** [3.9 - 8.1]	**11.2** [8.4 - 14]

The parameters $\rho_U(1)$ and $\rho_V(1)$ of the disc models for U and V respectively are equal to the autocorrelations at 40 km. The resulting autocorrelation $\rho_X(1)$ the process in X is equal to $2\rho(1) - 1$ and similarly to Y. 500 simulations were carried out in each cell.

Exhibit 6.

Proportion of type I errors and 95 % confidence intervals for the modified $t_{\hat{M}-2}$ test, the standard t_{N-2} test and Tjøstheim's index A, for a nominal 5 % test level in the case of two mutually independent processes X and Y, each generated from a mixture of disc processes U with probability γ and 3U with probability $1-\gamma$.

	$\rho_U(1)=\rho_V(1)$	0.0	0.2	0.4	0.6	0.8
	$\rho_X(1)=\rho_Y(1)$	0.0	0.16	0.32	0.48	0.64
	$t_{\hat{M}-2}$	**3.2** [1.7 - 4.5]	**3.4** [1.8 - 5.0]	**2.2** [0.9 - 3.5]	**3.0** [1.5 - 4.5]	**2.2** [0.9 - 3.5]
$\gamma=0.9$	t_{N-2}	**3.6** [2.0 - 5.2]	**4.0** [2.3 - 5.7]	**6.0** [3.9 - 8.1]	**10.2** [7.5 - 12.9]	**34.2** [30.0 - 38.4]
	A	**5.8** [3.8 - 7.8]	**5.4** [3.4 - 7.4]	**3.4** [1.8 - 5.0]	**5.0** [3.1 - 6.9]	**10.8** [8.1 - 13.5]
	$\rho_X(1)=\rho_Y(1)$	0.0	0.15	0.30	0.45	0.60
	$t_{\hat{M}-2}$	**4.8** [2.9 - 6.7]	**3.6** [2.0 - 5.2]	**4.2** [2.4 - 6.0]	**2.8** [1.4 - 4.2]	**3.4** [1.8 - 5.0]
$\gamma=0.8$	t_{N-2}	**5.0** [3.1 - 6.9]	**4.8** [2.9 - 6.7]	**6.8** [4.6 - 9.0]	**13.4** [10.4 - 16.4]	**36.8** [32.6 - 41.0]
	A	**6.4** [4.3 - 8.5]	**4.2** [2.4 - 6.0]	**4.4** [2.6 - 6.2]	**5.4** [3.4 - 7.4]	**11.4** [8.6 - 14.2]
	$\rho_X(1)=\rho_Y(1)$	0.0	0.15	0.30	0.45	0.60
	$t_{\hat{M}-2}$	**4.6** [2.8 - 6.4]	**3.8** [2.1 - 5.5]	**4.2** [2.4 - 6.0]	**5.0** [3.1 - 6.9]	**3.4** [1.8 - 5.0]
$\gamma=0.7$	t_{N-2}	**5.0** [3.1 - 6.9]	**4.4** [2.6 - 6.2]	**6.8** [4.6 - 9.0]	**14.6** [11.5 - 17.7]	**35.6** [31.4 - 39.8]
	A	**7.4** [5.1 - 9.7]	**2.8** [1.4 - 4.2]	**5.0** [3.1 - 6.9]	**5.4** [3.4 - 7.4]	**11.8** [9.0 - 14.6]

The parameters $\rho_U(1)$ and $\rho_V(1)$ of the disc models for U and V respectively are equal to the autocorrelation at 40 km.

The resulting autocorrelation $\rho_X(1)$ in the processes X is equal to $(9-8\gamma)^{-1}(3-2\gamma)^2 \rho_U(1)$ and similarly for $\rho_Y(1)$. 500 simulations were carried out in each cell.

Exhibit 7.

Proportion of type I errors and 95 % confidence intervals for the modified $t_{\hat{M}-2}$ test, the standard t_{N-2} test and Tjøstheim's index A, for a nominal 5 % test level in the case of two mutually independent processes X and Y each generated from a mixture of disc processes U with probability γ and bU with probability 1-γ.

		$\rho_U(1) = \rho_V(1) = 0.8$			
	b	3	6	9	12
	$\rho_X(1) = \rho_Y(1)$	0.77	0.65	0.52	0.41
γ=0.99	$t_{\hat{M}-2}$	**1.8** [0.6 - 3.0]	**1.8** [0.6 - 3.0]	**1.8** [0.6 - 3.0]	**1.8** [0.6 - 3.0]
	t_{N-2}	**32.8** [28.7 - 36.9]	**32.8** [28.7 - 36.9]	**32.8** [28.7 - 36.9]	**32.6** [28.5 - 36.7]
	A	**11.0** [8.3 - 13.7]	**11.2** [8.4 - 14]	**10.8** [8.1 - 13.5]	**10.6** [7.9 - 13.3]
	$\rho_X(1) = \rho_Y(1)$	0.69	0.45	0.31	0.24
γ=0.95	$t_{\hat{M}-2}$	**2.0** [0.8 - 3.2]	**1.8** [0.6 - 3.0]	**1.8** [0.6 - 3.0]	**2.2** [0.9 - 3.5]
	t_{N-2}	**35.2** [31.0 - 39.4]	**33.6** [29.5 - 37.7]	**33.4** [29.3 - 37.5]	**34.0** [29.8 - 38.2]
	A	**9.0** [6.5 - 11.5]	**14.2** [11.1 - 17.3]	**9.6** [7.0 - 12.2]	**11.0** [8.3 - 13.7]
	$\rho_X(1) = \rho_Y(1)$	0.64	0.4	0.29	0.23
γ=0.9	$t_{\hat{M}-2}$	**2.2** [0.9 - 3.5]	**2.8** [1.4 - 4.2]	**3.2** [1.7 - 4.7]	**3.8** [2.1 - 5.5]
	t_{N-2}	**34.2** [30.0 - 38.4]	**36.2** [32.0 - 40.4]	**34.2** [30.0 - 38.4]	**33.0** [28.9 - 37.1]
	A	**10.8** [8.1 - 13.5]	**11.0** [8.3 - 13.7]	**9.8** [7.2 - 12.4]	**9.6** [7.0 - 12.2]

The parameters $\rho_U(1)$ and $\rho_V(1)$ of the disc models for U and V respectively are equal to the autocorrelation at 40 km.

The resulting autocorrelation $\rho_X(1)$ in the processes X is equal to :

$[\gamma + b^2(1-\gamma)]^{-1} [\gamma + b(1-\gamma)]^2 \cdot \rho_U(1)$ and similarly for $\rho_Y(1)$. 500 simulations were carried out in each cell.

5. Partial correlations and multiple regressions

In the Gaussian framework the modified tests of association can be extended to test partial correlations. Alternatively, regression models can be specified and estimated, in which case spatial autocorrelation should be included in the variance–covariance matrix of the residuals, if necessary. We have thought it interesting to compare the results of these two different approaches on some examples. For this purpose we shall first describe how the modified $\mathbf{t}_{\widehat{M}-2}$ test is extended and indicate some results on its performance (§5.1); then we shall discuss the regression approach (§5.2); finally, we shall give some comparative results based upon examples concerning the relationship between male lung cancer mortality, smoking and industrial exposure analysed at the geographical level of the French *départements* (§5.3).

5.1. Extension of the modified $\mathbf{t}_{\widehat{M}-2}$ statistic to the testing of partial correlations.

This extension was detailed in Richardson (1989) and we shall now briefly describe its main features. For the sake of clarity the method is going to be described for testing the association between two variables $(\mathbf{Y}_\alpha, \mathbf{Z}_\alpha)$ adjusted for a third one, $\mathbf{X}_\alpha, \alpha \varepsilon \mathbf{A}$. Its generalisation to any number of adjustment variables is straightforward.

We suppose that the $\mathbf{3N}$ vector $\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{pmatrix}$ follows a multivariate normal distribution. Then the joint distribution of $(\mathbf{Y}, \mathbf{Z})$ conditional on $\mathbf{X}$ is also multivariate normal. Therefore, we can test the following hypothesis: $\rho_{\mathbf{YZ}\,.\,\mathbf{X}} = \mathbf{0}$, where $\rho_{\mathbf{YZ}\,.\,\mathbf{X}}$ is the partial correlation between $\mathbf{Y}$ and $\mathbf{Z}$ conditional on $\mathbf{X}$, by testing that the correlation between the residuals of the regressions of $\mathbf{Y}$ on $\mathbf{X}$ and of $\mathbf{Z}$ on $\mathbf{X}$ is zero. Hence, the method outlined in §2.1 and 2.2 equations (1) through (4) can be extended to test a partial correlation coefficient.

In practice this leads to using the modified $\mathbf{t}_{\widehat{M}-2}$ statistic on the residuals of the linear regression of $\mathbf{Y}$ on $\mathbf{X}$ and of $\mathbf{Z}$ on $\mathbf{X}$, respectively. These residuals are estimated by ordinary least squares (OLS), since the OLS regression estimates are unbiased.

In summary, the following steps are followed:

Step 1: regress $\mathbf{Y}$ on $\mathbf{X}$ by OLS, giving estimated residuals $\widehat{\mathbf{U}}$,

Step 2: regress $\mathbf{Z}$ on $\mathbf{X}$ by OLS, giving estimated residuals $\widehat{\mathbf{V}}$,

Step 3: test the correlation coefficient between $\widehat{\mathbf{U}}$ and $\widehat{\mathbf{V}}$ using the modified test statistic $\mathbf{t}_{\widehat{M}-2}$ given in formula (4), §2.2 with

$$\widehat{\mathrm{M}} = [\widehat{\mathbf{var}(\mathbf{r}_{\widehat{U},\widehat{V}})}]^{-1} + \mathbf{1}.$$

Thus, the degrees of freedom are adjusted with respect to the spatial autocorrelation in the conditional distributions.

Simulation results reported in Richardson (1989) indicate that the performance of this extended $\mathbf{t}_{\widehat{M}-2}$ test is satisfactory with regard both to significance levels and to power. Indeed, its observed power was close to that of a classical test of partial correlation based upon a number of observations compatible with the observed variance of $\mathbf{r}_{\mathbf{YZ}\,.\,\mathbf{X}}$, the empirical partial correlation coefficient.

5.2. Regression with a spatially parametrised variance–covariance error matrix.

In the regression models involving spatially distributed variables, the spatial structure can be taken into account by allowing spatial autocorrelation in the error variable; in other words, considering models of the form (where the subscript indicates the dimension of the matrix):

$$\mathbf{Y}_{n\times 1} = \mathbf{X}_{n\times p}\beta_{p\times 1} + \mathbf{U}_{n\times 1}, \tag{5}$$

with $\mathbf{U}$ following a multinormal distribution $\mathbf{N}(\mathbf{0}, \mathbf{\Sigma_U})$.

In the classical framework of independent errors distributed as $\mathbf{N}(\mathbf{0}, \sigma^2\mathbf{I})$, tests of a regression coefficient β_i, $1 \leq i \leq p$, and tests of the partial correlation between $\mathbf{Y}$ and $\mathbf{X}_i$ conditional on $\{\mathbf{X}_j, j \neq i\}$ are equivalent. In the more general framework of (5), tests of the coefficients β_i are done *conditionally* on an estimated structure for $\mathbf{\Sigma_U}$, since the matrix $\mathbf{\Sigma_U}$ is only known for theoretical cases.

Different approaches to the modelling of $\mathbf{\Sigma_U}$ have been suggested. They basically follow three lines:

(i) assuming a specific parametric model for $\mathbf{U}$,

(ii) assuming a known functional form for $\mathbf{\Sigma_U}$, and

(iii) direct estimation of $\mathbf{\Sigma_U}$.

Estimation in the framework of (i) was first discussed by Ord (1975), who considered a specific autoregressive model for $\mathbf{U}$ of the form:

$$\mathbf{U} = \mathbf{cWU} + \varepsilon, \qquad \varepsilon \sim \mathbf{N}(\mathbf{0}, \sigma^2\mathbf{I}), \tag{6}$$

where $\mathbf{W}$ is a known matrix of weights representing contiguity between the spatial locations. Examples of the fitting of models (5) and (6) are given in Bodson and Peeters (1975), Cliff and Ord (1981), Doreian (1981), Bivand (1984), and Haining (1987). The regression model defined by equations (5) and (6) is analogous in the spatial context to the commonly used time series method of ARMA modelling of residual errors in regression, which has been extensively used (see Glasbey, 1988, for a recent discussion of this topic).

For the alternative approach (ii), varied functional forms for $\mathbf{\Sigma_U}$ have been considered. Ripley (1981) discusses classes of spatial covariance functions that ensure $\mathbf{\Sigma_U}$ always is non-negative definite. These include the family of functions first introduced by Whittle (1954) in which the covariance at $\mathbf{r}$ is proportional to $\mathbf{r}^{\mathbf{v}}\mathbf{K}_{\mathbf{v}}(\mathbf{ar})$, $\mathbf{v} > \mathbf{0}$, where $\mathbf{r}$ denotes the distance between points in $\mathbf{R}^2$, and $\mathbf{K}_{\mathbf{v}}$ are the modified Bessel functions of the second kind. Setting $\mathbf{v} = \mathbf{1/2}$ gives an exponential correlation function depending upon only one parameter. Cook and Pocock (1983), in their study of the association between water hardness and cardiovascular deaths, used a more general form of the exponential decline function that depended upon two coefficients.

Including the possibility of anisotropy, Vecchia (1988) considered a general form for a rational spectral density function of a two-dimensional process, whose covariance can be expressed in terms of derivatives of the Bessel function $\mathbf{K}_0(\mathbf{ar})$. This family of covariances overlaps partially with those introduced by Whittle.

There is another interesting family of functions for $\mathbf{\Sigma_U}$ in which the covariance between two points is defined as being proportional to the intersection area of two discs of common radius centered on those points. This family, often referred to as the disc model, was selected for simulating non–lattice autocorrelated processes in our work on the modified tests.

Other functional forms for $\mathbf{\Sigma_U}$ that do not necessarily ensure that $\mathbf{\Sigma_U}$ is non–negative definite also have been tried. This is the case for the quadratic distance function proposed by Agterberg (1984) and used by Haining (1987) for modelling the autocorrelation in trend surface analysis.

The choice of parameterisation of $\mathbf{\Sigma_U}$ is sometimes made by plotting the variogram of the OLS residuals. Care has to be taken when interpreting the variogram since it is sensitive to the number of pairs of data points used to estimate the empirical covariance at a particular distance. The number of pairs will vary with the distance, typically increasing at first. Ripley (1981) advises the use of cross–validation by successive deletion of data points to assess the fit of the model chosen for the covariance function.

Direct estimation of $\mathbf{\Sigma_U}$ in approach (iii) was advocated by Arora and Brown (1977) in a case where spatial data at different time intervals were available. In defining $\mathbf{\Sigma_U}$, Haining also has considered using direct estimates of the residual autocorrelation or a finite number of spatial lags (defined through a contiguity matrix), and zero elsewhere. Here, again, the resulting covariance matrix is not necessarily non–negative definite.

Once the model for $\mathbf{\Sigma_U}$ has been specified, estimation can be carried out in a Gaussian framework by maximum likelihood (ML) techniques, in cases (i) and (ii), when $\mathbf{\Sigma_U}$ is non–negative definite, or through some iterative method based upon generalised least squares, similar to that first proposed by Cochrane and Orcutt for time series (1949). Mardia and Marshall (1984) have studied the asymptotic properties of the maximum likelihood estimators (MLE) for model (5) in the cases of (i) and (ii). Assuming that $\mathbf{Y}$ is a Gaussian process, they give conditions that ensure the consistency and the asymptotic normality of $(\widehat{\beta}, \widehat{\theta})$, the MLEs of (β, θ), θ being the vector of parameters of $\mathbf{\Sigma_U}$. Conditional upon $\widehat{\theta}$, the variance–covariance matrix for $\widehat{\beta}$ can be calculated, and hence tests of the regression coefficients can be performed. Because of the conditionality involved, the standard errors of $\widehat{\beta}$ might be underestimated, which leads some authors to consider a Bayesian approach. Hepple (1979) carried out a Bayesian analysis of the model defined by equations (5) and (6), assuming a simple linear regression with a constant β_1, a slope β_2, and uniform diffuse priors for $\beta = (\beta_1, \beta_2)$, $\log(\sigma)$, and ρ. He shows how the conditional posterior distribution of β_2 is sensitive to values of ρ, and he derives the bivariate posterior distribution for β_2 and ρ, with mode corresponding to MLEs of β_2 and ρ.

When the data set contains a large number of points, MLE becomes computationally very heavy, and is fraught with difficulties. Mardia and Marshall (1984) used the Fisher scoring technique. A study by Warnes and Ripley (1987) showed that this method usually converges only to the nearest local maximum, and in some cases did not converge at all. Furthermore, they found cases of multimodal profile likelihoods that renders hazardous the search for a maximum. Moreover, Ripley (1988) reports some simulation results where the global maximum found is well away from the true value. In a recent paper, Vecchia (1988) proposed carrying out estimation, within the class of covariance functions defined earlier in this paper, with the help of successive *approximate likelihood functions* that are much easier to handle. He applied his method successfully to simulated data sets and to water level data where he wanted to estimate a trend.

Trend fitting, or regression when $\mathbf{X}$ is a matrix of polynomial powers and cross–products of geographic coordinates, is indeed an area where the estimation of model (5) has been much used. As pointed out by Haining (1987), residual autocorrelation can arise either from false specification of the order of the trend, or from local scale effects arising from spatial processes that operate at an intermediate scale between the regional trend and the local residuals. Haining analysed aerial survey data of marine pollution using three approaches: model (6) fitted via ML techniques, Agterberg's quadratic covariance function, and a direct estimation of $\mathbf{\Sigma_U}$ based upon six spatial lags. He encountered some problems of convergence in the last two approaches, due possibly to the matrix $\mathbf{\Sigma_U}$ not satisfying the non–negative definite property. When comparison was possible, he found some moderate differences between the resulting trend estimates given by the three methods. He also commented that the direct estimation of $\mathbf{\Sigma_U}$ is the least satisfactory.

Haining's results seem to indicate that regression estimates can be quite sensitive to the choice of model for $\mathbf{\Sigma_U}$. This is an interesting problem, and we have chosen to investigate it in a different manner by restricting the comparisons to models for $\mathbf{U}$ and $\mathbf{\Sigma_U}$ that satisfy the non–negative definiteness property, and by using the same estimation technique (namely ML) in all cases. Finally, the same examples will be analysed by means of the modified $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ test for partial correlations.

5.3. Comparison of regression results obtained by different modelling of the variance–covariance error matrix and by modified $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ tests

The examples analysed in this section concern the relationship between male lung cancer, some industrial factors, and smoking. Among the different cancer sites, male lung cancer has been frequently associated with industrial exposure (Pastorino *et al.*, 1984; Benhamou *et al.*, 1988), and it has been estimated that 15% of all lung cancer arising in men in the U. S. A. could be due to occupational risk factors (Doll and Peto, 1981). This figure has been the subject of recent debates (Simonato et al, 1988) and several authors have tried to estimate it using data from different case–control studies (Vineis *et al.*, 1988; Ronco *et al.*, 1988). Hence it is particularly interesting to test the link between male lung cancer and industrial exposure at a geographical level. If the results that emerge are consistent with those of case–controls or cohort studies, it becomes possible to calculate an estimation of attributable risk due to occupational factors based on comprehensive geographical census data rather than data from individual epidemiological studies. A study of the biases associated with estimation from ecological (grouped) data was done in Richardson, Stücker and Hémon (1987). In view of the strong link between smoking and lung cancer, some adjustment based on a measure of smoking consumption is also needed. We present some results on four branches of industry: metal, general engineering, mining and textile works.

5.3.1. The data

Male lung cancer mortality rate has been standardised over the age 35-74 and over a 2 year period, 1968-1969. The data were provided by the French National Institute for Health and Medical Research (INSERM) at the level of the French *départements*. Cigarette sales data were compiled by the French Nationalised Tobacco Company (SEITA). To take into account the time lag between smoking and the onset of a lung pathology, cigarette sales per inhabitant were recorded in 1953. Demographic data on the percentage of employed males in the metal industry, in general engineering, in the textile industry and in mining were taken from the 1962 census (INSEE). After the grouping of the *départements* around Paris into one area and the exclusion of four others owing to the poor quality of the data, $\mathbf{N} = \mathbf{82}$ locations were retained indexed by the coordinates of the administrative centres of the *départements*.

5.3.2. Spatial models for the variance–covariance error matrix

Four models for the variance–covariance error matrix Σ_U in the regression model:

$$\mathbf{Y} = \mathbf{Xb} + \mathbf{U}$$

[cf. (5) §3.2] were chosen. The first two involve only one shape parameter for the autocorrelation.

(a) $\quad \mathbf{U} = \mathbf{cWU} + \varepsilon, \qquad \varepsilon \sim \mathbf{N}(\mathbf{0}, \sigma^2\mathbf{I}),$

$\mathbf{W}$ is defined as an (82×82) $0-1$ contiguity matrix on the French *départements* normalised such that each row sum is equal to unity, and with non–zero $(\mathbf{i},\mathbf{j})$ element if *départements* $\mathbf{i}$ and $\mathbf{j}$ have a common border length. In this case the covariance matrix Σ_U is equal to $\sigma^2(\mathbf{I} - \mathbf{cW})^{-1}(\mathbf{I} - \mathbf{cW}^t)^{-1}$.

(b) $\quad \Sigma_U$ follows a disc model.

Letting $\mathbf{d_{ij}}$ be the distance between locations of *départements* $\mathbf{i}$ and $\mathbf{j}$, the $(\mathbf{i},\mathbf{j})$ element of Σ_U is then given by $\sigma^2\mathbf{f_a}(\mathbf{d_{ij}})$ with:

$$\mathbf{f_a}(\mathbf{r}) = \frac{2}{\pi}[\cos^{-1}(\frac{\mathbf{r}}{2\mathbf{a}}) - \frac{\mathbf{r}}{2\mathbf{a}}(1 - \frac{\mathbf{r}^2}{4\mathbf{a}^2})^{1/2}] \quad r \le 2a,$$
$$\mathbf{f_a}(\mathbf{r}) = 0 \quad r > 2a.$$

This covariance function exhibits a fairly linear decrease with increasing distance (cf. Ripley, 1981 p. 56) with the value for $\mathbf{f}'_\mathbf{a}(\mathbf{0})$, the slope of the tangent at zero, equal to $-\mathbf{2}(\mathbf{a}\pi)^{-1}$. To illustrate its behaviour using a linear approximation for small $\mathbf{r}$, we can say that if the autocorrelation is equal to 0.8 for a distance $\mathbf{r} = 40$ km [implying $\mathbf{f}'_\mathbf{a}(\mathbf{0}) = 0.005$] then the autocorrelation will become zero after a distance of $\mathbf{2a} = \mathbf{2} \times \mathbf{2}(\pi \times \mathbf{0.005})^{-1} = \mathbf{255}$ km.

Next we consider classes of covariances with two shape parameters for the autocorrelation.

(c) $\quad \Sigma_U$ follows an exponential model.

The $(\mathbf{i},\mathbf{j})$ element of Σ_U is given by $\sigma^2\gamma\mathbf{e}^{-\lambda\mathbf{d_{ij}}}$. This model was used by Cook and Pocock (1983) after inspection of the variogram of the OLS. residuals of their regression model.

(d) $\quad \Sigma_U$ follows a Bessel model.

The $(\mathbf{i},\mathbf{j})$ element of Σ_U is given by $\sigma^2\mathbf{g_{v,a}}(\mathbf{r})$ with

$$\mathbf{g_{v,a}}(\mathbf{r}) = \frac{1}{2^{\mathbf{v}-1}\Gamma(\mathbf{v})}(\mathbf{ar})^{\mathbf{v}}\mathbf{K_v}(\mathbf{ar}), \quad \mathbf{v} > \mathbf{0}, \quad \mathbf{a} > \mathbf{0}.$$

All models were fitted by ML. The numerical maximisation was performed with a safeguarded quadratic interpolation method for models (a) and (b), simplified by the use of eigenvalues in case (a) following the remark made by Ord (1975). For models (c) and (d) a quasi–Newton method with finite difference gradient was used. Starting points were provided after visual comparison between the shapes of the theoretical correlograms and those calculated from the OLS residuals. New starting points were tried when the program indicated convergence problems. Clearly the possibility of having encountered only local maxima cannot be discarded. The only lengthy maximisation was the one for model (d). The parameter v was restricted to lie in the range [0.1 – 2.9] after a visual inspection. All maximisations were performed on a Compaq 386 microcomputer using IMSL library routines (1987) UVMIF (1–parameter) or BCONF (2–parameter).

5.3.3. Relative efficiency of OLS with regard to generalised least squares (GLS)

In a recent paper, Krämer and Donninger (1987) have given some results on the relative efficiency of OLS with regard to GLS for the autoregressive model (a). They define the relative efficiency $\mathbf{e}$ by the quotient of the traces of the covariance matrices for GLS and OLS respectively:

$$\mathbf{e} = (\mathbf{X}^t\mathbf{\Sigma}_U^{-1}\mathbf{X})^{-1}/[(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{\Sigma}_U\mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}].$$

They show that for $\mathbf{\Sigma}_U = (\mathbf{I} - \mathbf{c}\mathbf{W})^{-1}(\mathbf{I} - \mathbf{c}\mathbf{W}^t)^{-1}$, the limit of $\mathbf{e}$ is unity as $\mathbf{c}$ tends to its maximal value of $\mathbf{1}$ provided the regressors $\mathbf{X}$ include a constant term. This result is of limited practical significance since, as they point out, for intermediate values of $\mathbf{c}$, the loss of efficiency can be substantial. We have computed this ratio for all the fitted regression models in order to pursue this problem further.

5.3.4. Regression results

Metal industry and general engineering workers

There is broad agreement amongst the four models in finding a statistically significant link between lung cancer rates and these two industries, both with and without adjustment for cigarette sales (Exhibits 8 and 9). Nevertheless, we observe some amount of variation among the models concerning the regression slopes or their significance levels, with closer similarity within the 1–parameter or the 2–parameter models. The inclusion of cigarette sales in the regression clearly improves the fit in all models with higher ML values. Even after this adjustment, there is still a substantial amount of autocorrelation among the residuals as shown in the estimated autocorrelation at 40 km. For both industries the disc model gives lower estimates of this residual autocorrelation and consequently higher t values. The fit of the exponential and the Bessel model are very close, with slightly higher likelihoods for the exponential model. From visual inspection of the observed and estimated correlograms, the Bessel model seems to give the closer fit and the 1–parameter disc model is clearly not flexible enough.

Mining industry

A non–significant association between male lung cancer rates and the percent of workers employed by the mining industry is found in all four models (Exhibit 10). The slopes are decreased after adjustment for cigarette sales and there is close agreement on residual autocorrelation and on the values of the t–statistics between all four models.

Exhibit 8.

Linear regression between lung cancer mortality rates for men and metal industry workers with and without adjustment on cigarette sales : standard, modified $t_{\hat{M}-2}$ tests and results for different parametrisations of the error variance-covariance matrix.

Metal industry workers	no adjustment		Adjustment on cigarette sales	
Standard test : β	29.1×10^{-4}		17.3×10^{-4}	
t (p value)	7.1	$(< 10^{-9})$	5.27	(10^{-6})
$t_{\hat{M}-2}$	3.00	(9.5×10^{-3})	3.48	(1.4×10^{-3})
	$\hat{M} = 16$		$\hat{M} = 37$	

Autoregressive model (a)

Metal industry workers	no adjustment		Adjustment on cigarette sales	
β	18.9×10^{-4}		12.6×10^{-4}	
t (p value)	4.34	(4×10^{-5})	3.70	(3×10^{-4})
log M.L.	1443.5		1490.0	
model parameter : $\hat{c}$	0.574		0.379	
e+	0.54		0.82	

Disc model (b)

Metal industry workers	no adjustment		Adjustment on cigarette sales	
β	18.7×10^{-4}		15.5×10^{-4}	
t (p value)	4.22	(6×10^{-5})	4.69	(10^{-5})
log M.L.	1431.1		1487.4	
model parameter : $2\hat{a}$	100.4		65.8	
autocorrelation ρ (1)*	0.493		0.227	
e+	0.66		0.834	

Exponential model (c)

Metal industry workers	no adjustment		Adjustment on cigarette sales	
β	15.7×10^{-4}		10.1×10^{-4}	
t (p value)	3.51	(7×10^{-4})	2.90	(4.8×10^{-3})
log M.L.	1445.4		1491.6	
model parameters : $\hat{\gamma}, \hat{\lambda}$	0.725 , 0.35×10^{-2}		0.569 , 0.79×10^{-2}	
autocorrelation $\rho(1)$*	0.629		0.414	
e+	0.28		0.58	

Bessel model (d)

Metal industry workers	no adjustment		Adjustment on cigarette sales	
β	17.1×10^{-4}		10.9×10^{-4}	
t (p value)	3.8	(2.8×10^{-4})	3.17	(2.2×10^{-3})
log M.L.	1442.3		1491.3	
model parameters : $\hat{v}, \hat{a}$	0.304 , 0.98×10^{-2}		0.211 , 0.01	
autocorrelation $\rho(1)$*	0.499		0.378	
e+	0.50		0.69	

* autocorrelation calculated using estimated model parameters at a distance of 40 km.
\+ relative efficiency as defined in §5.3.3.

Exhibit 9.

Linear regression between lung cancer mortality rates for men and general engineering workers with and without adjustment on cigarette sales : standard, modified $t_{\hat{M}-2}$ tests and results for different parametrisations of the error variance-covariance matrix.

General engineering workers	no adjustment		Adjustment on cigarette sales	
Standard test : β	32.4×10^{-4}		18.9×10^{-4}	
t (p value)	4.23	(6×10^{-5})	3.54	(7×10^{-4})
$t_{\hat{M}-2}$ (p value)	2.72	(1.03×10^{-2})	2.88	(5.7×10^{-3})
	$\hat{M} = 35$		$\hat{M} = 55$	

Autoregressive model (a)

General engineering workers	no adjustment		Adjustment on cigarette sales	
β	25.7×10^{-4}		16.7×10^{-4}	
t (p value)	4.77	(0.8×10^{-5})	3.67	(4×10^{-4})
log M.L.	1449.0		1492.5	
model parameter : $\hat{c}$	0.736		0.586	
e+	0.31		0.33	

Disc model (b)

General engineering workers	no adjustment		Adjustment on cigarette sales	
β	21.6×10^{-4}		18.4×10^{-4}	
t (p value)	4.28	(5×10^{-5})	3.96	(2×10^{-4})
log M.L.	1443.2		1485.2	
model parameter : $2\hat{a}$	190.1		85.4	
autocorrelation ρ (1)*	0.732		0.403	
e+	0.21		0.73	

Exponential model (c)

General engineering workers	no adjustment		Adjustment on cigarette sales	
β	21.8×10^{-4}		14.23×10^{-4}	
t (p value)	4.0	(0.4×10^{-3})	3.24	(1.7×10^{-3})
log M.L.	1449.9		1495.5	
model parameters : $\hat{\gamma}, \hat{\lambda}$	0.884 , 0.34×10^{-2}		0.787 , 0.49×10^{-2}	
autocorrelation ρ(1)*	0.772		0.647	
e+	0.22		0.37	

Bessel model (d)

General engineering workers	no adjustment		Adjustment on cigarette sales	
β	22.3×10^{-4}		15.1×10^{-4}	
t (p value)	4.05	(1.2×10^{-4})	3.36	(1.2×10^{-3})
log M.L.	1448.7		1493.6	
model parameters : $\hat{v}, \hat{a}$	0.381 , 0.61×10^{-2}		0.417 , 0.013	
autocorrelation ρ(1)*	0.695		0.537	
e+	0.28		0.52	

* autocorrelation calculated using estimated model parameters at a distance of 40 km.
+ relative efficiency as defined in §5.3.3.

Exhibit 10.

Linear regression between lung cancer mortality rates for men and mining industry workers with and without adjustment on cigarette sales : standard, modified $t_{\hat{M}-2}$ tests and results for different parametrisations of the error variance-covariance matrix.

Mining industry workers		no adjustment		Adjustment on cigarette sales	
Standard test :	β	19.9×10^{-4}		9.8×10^{-4}	
t	(p value)	3.16	(2.2×10^{-3})	2.14	(3.5×10^{-2})
$t_{\hat{M}-2}$	(p value)	2.37	(2.2×10^{-2})	2.23	(2.8×10^{-2})
		$\hat{M} = 47$		$\hat{M} = 89$	

Autoregressive model (a)

Mining industry workers	no adjustment		Adjustment on cigarette sales	
β	5.2×10^{-4}		3.0×10^{-4}	
t (p value)	1.03	(0.31)	0.78	(0.43)
log M.L.	1429.6		1480.3	
model parameter : $\hat{C}$	0.687		0.539	
e+	0.42		0.66	

Disc model (b)

Mining industry workers	no adjustment		Adjustment on cigarette sales	
β	4.2×10^{-4}		1.8×10^{-4}	
t (p value)	0.95	(0.34)	0.55	(0.58)
log M.L.	1425.7		1478.5	
model parameter : $2\hat{a}$	240.1		184.1	
autocorrelation $\rho(1)$*	0.788		0.723	
e+	0.21		0.31	

Exponential model (c)

Mining industry workers	no adjustment		Adjustment on cigarette sales	
β	4.5×10^{-4}		2.4×10^{-4}	
t (p value)	0.94	(0.35)	0.66	(0.51)
log M.L.	1435.9		1485.7	
model parameters : $\hat{\gamma}, \hat{\lambda}$	0.852 , 0.32×10^{-2}		0.743 , 0.50×10^{-2}	
autocorrelation $\rho(1)$*	0.749		0.609	
e+	0.35		0.60	

Bessel model (d)

Mining industry workers	no adjustment		Adjustment on cigarette sales	
β	4.6×10^{-4}		2.9×10^{-4}	
t (p value)	0.97	(0.33)	0.80	(0.43)
log M.L.	1434.2		1484.4	
model parameters : $\hat{v}, \hat{a}$	0.358 , 0.63×10^{-2}		0.311 , 0.92×10^{-2}	
autocorrelation $\rho(1)$*	0.663		0.524	
e+	0.411		0.65	

* autocorrelation calculated using estimated model parameters at a distance of 40 km.
\+ relative efficiency as defined in §5.3.3.

Textile industry

There is a discrepancy between the results given by the disc model that finds a borderline association after adjustment for cigarette sales, and other models that find a non-significant association with the textile industry (Exhibit 11). Note that for all models, the inclusion of cigarette sales has led to a higher t-value for the regression coefficient of mining. In one case (exponential model without adjustment for cigarettes), the maximisation encountered numerical problems and the maximum found might be only local.

Amongst the investigated associations with industrial exposure we thus found overall general agreement between the results given by different parametrisations of the variance-covariance error matrix. Nevertheless, significance levels did vary by a non-negligible ratio and in one case this variation could lead to a different interpretation of the results. Hence, it is important, when using this approach, to check the coherence of the results for at least two different models of $\mathbf{\Sigma_U}$. A further note of caution is also warranted when one compares the results given by the autoregressive model (a) for two choices of $\mathbf{W}$: a non-standardised $(0-1)$ matrix (results not shown) and the same matrix with row sums equal to 1. Some discrepancies arise even though the matrices use the same definition of "contiguity". With the non-standardised $\mathbf{W}$ matrix we found, for instance, a significant association with mining ($t = 2.02$, $p = 0.046$) before adjustment for cigarette sales and a borderline association ($t = 1.69$, $p = 0.095$) after adjustment. Overall the choice of standardising $\mathbf{W}$ led to higher ML values and to closer results with the other models than when using the non-standardised version. Finally one needs to recall that contrary to the time series case, it is difficult to ensure that a global maximum has always be found.

Relative efficiency

All the relative efficiencies calculated were far from unity even though a constant term was always included in the regression. This reinforces the necessity of modelling the spatial structure of the residuals. The relative efficiency of the exponential model was always lower than that of the Bessel model, pointing again to the exponential model as having a slightly better performance than the other models in our examples.

5.3.5. Comparison between the regression results and the modified $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ *statistic results*

For the mining industry and general engineering workers, the $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ test agrees overall with the results of the regressions by finding a statistically significant link. The agreement is closer once adjustment for cigarette sales has been performed, with levels of significance of the same order of magnitude as those given by the exponential or Bessel models for both industries. The agreement is poorer before adjustment when there is a higher residual autocorrelation with values of $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ lower than those given by the models. Recall that the adjustment of the d.f. carried out by $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ uses directly the autocorrelations $\hat{\mathbf{C}}_{\mathbf{X}}(\mathbf{k})$ and $\hat{\mathbf{C}}_{\mathbf{Y}}(\mathbf{k})$ for the $\mathbf{k}$ strata and consequently is adapted to the whole correlograms, whereas the model approach centres its estimations mainly on the first few classes of higher autocorrelation. In cases of lower autocorrelation levels, the first strata give most of the contribution to the modified d.f., and hence agreement might be expected to be closer between the two approaches.

For the mining industry there is a divergence of results since the modified $\mathbf{t}_{\widehat{\mathbf{M}}-2}$ statistic finds a significant link at the 5% level whereas none of the four models do. Recall that a

Exhibit 11.

Linear regression between lung cancer mortality rates for men and textile industry workers with and without adjustment on cigarette sales : standard, modified $t_{\hat{M}-2}$ tests and results for different parametrisations of the error variance-covariance matrix.

Textile industry workers	no adjustment		Adjustment on cigarette sales	
Standard test : β	18.3×10^{-4}		11.2×10^{-4}	
t (p value)	2.57	(1.2×10^{-2})	2.28	(2.5×10^{-2})
$t_{\hat{M}-2}$ (p value)	1.52	(0.14)	1.89	(6.4×10^{-2})
	$\hat{M} = 30$		$\hat{M} = 57$	

Autoregressive model (a)

Textile industry workers	no adjustment		Adjustment on cigarette sales	
β	0.6×10^{-4}		4.2×10^{-4}	
t (p value)	0.1	(0.9)	0.96	(0.37)
log M.L.	1428.6		1480.6	
model parameter : $\hat{c}$	0.709		0.543	
e+	0.36		0.66	

Disc model (b)

Textile industry workers	no adjustment		Adjustment on cigarette sales	
β	1.8×10^{-4}		9.1×10^{-4}	
t (p value)	0.33	(0.74)	2.06	(0.042)
log M.L.	1426.4		1474.6	
model parameter : $2\hat{a}$	189.8		83.4	
autocorrelation ρ (1)*	0.732		0.39	
e+	0.24		0.72	

Exponential model (c)

Textile industry workers	no adjustment		Adjustment on cigarette sales	
β	-0.5×10^{-4}		3.28×10^{-4}	
t (p value)	- 0.093	(0.92)	0.80	(0.43)
log M.L.	1435.0		1485.9	
model parameters : $\hat{\gamma}, \hat{\lambda}$	0.871 , 0.31×10^{-2}		0.755 , 0.52×10^{-2}	
autocorrelation ρ(1)*	0.771		0.613	
e+	0.21		0.42	

Bessel model (d)

Textile industry workers	no adjustment		Adjustment on cigarette sales	
β	1.8×10^{-4}		4.6×10^{-4}	
t (p value)	0.32	(0.75)	1.09	(0.28)
log M.L.	1431.4		1484.4	
model parameters : $\hat{v}, \hat{a}$	0.471 , 0.01		0.353 , 0.011	
autocorrelation ρ(1)*	0.649		0.518	
e+	0.35		0.59	

* autocorrelation calculated using estimated model parameters at a distance of 40 km.
\+ relative efficiency as defined in §5.3.3.

comparable significant link was also obtained when a non–standardised $\mathbf{W}$ matrix was used in model (a). This example was also investigated after removal of a linear trend component (see Richardson, 1989), and after this further adjustment the $\mathbf{t}_{\widehat{M}-2}$ statistic becomes clearly non–significant. This points to the existence of a simple gradient–like structure for the residuals of mining, cigarette sales and lung cancer that is taken into account by the regression models, but not by the modified $\mathbf{t}_{\widehat{M}-2}$ test (without trend removal).

For the textile industry, the $\mathbf{t}_{\widehat{M}-2}$ finds a non–significant link, agreeing qualitatively with the results of three of the models. Nevertheless, the significance levels are lower than those given by these three models and not far from the borderline significance given by the disc model. Hence the results given by the $\mathbf{t}_{\widehat{M}-2}$ statistic are in this case intermediate between the disc model and the other three models.

In conclusion, except for mining, the results given by the modified $\mathbf{t}_{\widehat{M}-2}$ statistic would lead to the same conclusions as those given by the regression models. We note however that the values of the $\mathbf{t}_{\widehat{M}-2}$ statistic tend to be more moderate in their adjustment for spatial autocorrelation than those obtained by the regressions with parametrised variance–covariance error matrix, which find the link either still strongly significant or not at all. This is probably due to the more flexible nature of the adjustment for spatial structure carried out by $\mathbf{t}_{\widehat{M}-2}$. It is also clearly apparent that the use of standard regressions with i.i.d. errors would lead to quite erroneous conclusions.

6. Concluding remarks

In this paper we have compared different approaches to the testing of association between spatially autocorrelated variables.

One approach consists of adjusting classical tests on simple or partial correlation coefficients to take into account spatial autocorrelation. The proposed modified tests that we developed were shown to have satisfactory Type I error and power. These methods do not require the identification of a particular parametric model for the type of spatial autocorrelation and they only involve straightforward calculations that can be done on small computers. As these tests are derived in the Gaussian framework, a study of the robustness of their performance to departure from normality was also carried out. There was no measurable effect on Type I errors except in highly autocorrelated cases where the tests became too conservative.

Alternatively new measures of association could be employed such as Tjøstheim's non–parametric index $\mathbf{A}$. In a Monte Carlo study we found an expected increase in Type I errors of the index $\mathbf{A}$ in cases of high autocorrelation. Furthermore its observed power was lower than that of the modified test in most cases, even under an alternative hypothesis of local spatial permutation provided that a random error term is included.

Finally, the results of the modified tests were compared to those obtained by generalised regression with different spatial models for the variance–covariance error matrix on four examples of geographical epidemiology. There was general agreement between the two approaches except in one example. The correspondence between the modified test and general regressions was greater in the two cases where there was also close agreement between the results of the different spatial parametrisations. The modified tests can thus be seen as a first step in the testing of association, to be carried out before any parametric modelling that probably involves some degree of arbitrariness and some computational difficulties.

7. References

Agterberg, F. (1984) Trend surface analysis, in *Spatial Statistics and Models*, edited by G. Gaile and C. Willmot, pp. 147-171. Dordrecht: Reidel.

Armstrong, B., and R. Doll. (1975) Environmental factors and cancer incidence and mortality in different countries, with special reference to dietary practices. *International Journal of Cancer*, **15**, 617-631.

Arora, S., and M. Brown. (1977) Alternative approaches to spatial autocorrelation: an improvement over current practice. *International Regional Science Review*, **2**, (1), 67-78.

Bartlett, M. (1935) Some aspects of the time–correlation problem in regard to tests of significance. *Journal of the Royal Statistical Society*, **98**, 536-543.

Benhamou, S., E. Benhamou, and R. Flamant. (1988) Occupational risk factors of lung cancer in a French case–control study. *British Journal of Industrial Medicine*, **45**, 231-233.

Bivand, R. (1984) Regression modeling with spatial dependence: an application of some class selection and estimation methods. *Geographical Analysis*, **16**, (1), 25-37.

Bodson, P., and D. Peeters. (1975) Estimation of the coefficients of a linear regression in the presence of spatial autocorrelation. An application to a Belgian labour–demand function. *Environment and Planning A*, **7**, 456-472.

Bolthausen, E. (1982) On the central limit theorem for stationary random fields. *Annals of Probability*, **10**, 1047-1050.

Cliff, A., and J. Ord. (1981) *Spatial Processes: Models and Applications*. Pion: London.

Clifford, P., S. Richardson, and D. Hémon. (1989) Assessing the significance of the correlation between two spatial processes. *Biometrics*, **45**, (1), 123-134.

Cochrane, D., and G. Orcutt. (1949) Applications of least–square regressions to relationships containing autocorrelated error terms. *Journal of the American Statistical Association*, **44**, 32- 61.

Cook, D., and S. Pocock. (1983) Multiple regression in geographic mortality studies with allowance for spatially correlated errors. *Biometrics*, **39**, 361-371.

Doll, R. (ed.). (1984) The geography of disease, *British Medical Bulletin*. Published for the British Council by Churchill Livingstone.

Doll, R., and P. Peto. (1981) The causes of cancer. *Journal of the National Cancer Institute*, **66**, 1191-1308.

Doreian, P. (1981) Estimating linear models with spatially distributed data, in *Sociological Methodology*, edited by S. Leinhardt, pp. 359-388. San Francisco: Jossey–Bass.

Geweke, J. (1981) A comparison of tests of the independence of two covariance–stationary time series. *Journal of the American Statistical Association*, **76**, (374), 363-373.

Glasbey, C. (1988) Examples of regression with serially correlated errors. *The Statistician*, **37**, 277-291.

Glick, B. (1982) A spatial rank order correlation measure. *Geographical Analysis*, **14**, (2), 177-181.

Griffith, D. (1980) Towards a theory of spatial statistics. *Geographical Analysis*, **12**, (4), 325-339.

Guyon, X, and S. Richardson. (1984) Vitesse de convergence du théorème de la limite centrale pour des champs faiblement dépendants. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, **66**, 297-317.

Haining, P. (1987) Trend surface models with regional and local scales of variation with an application to aerial survey data. *Technometrics*, **29**, 461-469.

Haugh, L. (1976) Checking the independence of two covariance– stationary time series: a univariate residual cross–correlation approach. *Journal of the American Statistical Association*, **71**, 378-385.

Hepple, L. (1979) Bayesian analysis of the linear model with spatial dependence, in *Exploratory and Explanatory Statistical Analysis of Spatial Data*, edited by C. Bartels and R. Ketellapper, pp. 179-199 . Boston: Martinus Nijhoff.

Hubert, L., and R. Golledge. (1982) Measuring association between spatially defined variables: Tjøstheim's index and some extensions. *Geographical Analysis*, **14**, (3), 273-278.

Johnston, J. (1972) *Econometric Methods*, 2nd edition. New York: McGraw-Hill.

Krämer, W., and C. Donninger. (1987) Spatial autocorrelation among errors and the relative efficiency of OLS in the linear regression model. *Journal of the American Statistical Association*, **82**, (398), 577-579.

Mardia, K., and R. Marshall. (1984) Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika*, **71**, 135-146.

Ord, K. (1975) Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, **70**, 120-126.

Pastorino, U., F. Berrino, A. Gervasio, V. Pesenti, E. Riboli, and P. Crosignomi. (1984) Proportion of lung cancers due to occupational exposure. *Journal of the National Cancer Institute*, **33**, 231-237.

Pierce, D. (1977) Relationships and the lack thereof between economic time series, with special reference to money and interest rates. *Journal of the American Statistical Association*, **72**, (357), 11-26.

Richardson, S., and D. Hémon. (1982) On the variance of the sample correlation between two independent lattice processes. *Journal of Applied Probability*, **18**, 943-948.

Richardson, S., and P. Clifford. Testing association between spatial processes. Presented at the A.M.S. conference on *Spatial Statistics and Imaging*, Bowdoin College (1988).

Richardson, S. (1990) A method for testing the significance of geographical correlations with application to industrial lung cancer in France. *Statistics in Medicine*, **9**, 515-528.

Richardson, S., I. Stücker, and D. Hémon. (1987) Comparison of relative risks obtained in ecological and individual studies: some methodological considerations. *International Journal of Epidemiology*, **16**, 111-120.

Ripley, B. (1981) *Spatial Statistics*. New York: Wiley.

Ripley, B. (1988) *Statistical Inference for Spatial Processes*. Cambridge: Cambridge University Press.

Ronco, G., G. Ciccone, D. Mirabelli, B. Troia, and P. Vineis. (1988) Occupation and lung cancer in two industrialized areas of Northern Italy. *International Journal of Cancer*, **41**, 354-358.

Simonato, L., P. Vineis, and A. Fletcher. (1988) Estimates of the proportion of lung cancer attributable to occupational exposure. *Carcinogenesis*, **9**, 1159-1165.

Tjøstheim, D. (1978) A measure of association for spatial variables. *Biometrika*, **65**, (1), 109-114.

Upton, G., and B. Fingleton. (1983) *Spatial Data Analysis by Example.* Chichester: Wiley.

Vecchia, A. (1988) Estimation and model identification for continuous spatial processes. *Journal of the Royal Statistical Society*, B **50**, (2), 297-312.

Vineis, P., T. Thomas, R. Hayes, W. Blot, T. Mason, L. Williams Pickle, P. Correa, E. Fontham, and J. Schoenberg. (1988) Proportion of lung cancers in males, due to occupation, in different areas of the USA. *International Journal of Cancer*, **42**, (6), 851-856.

Warnes, J., and B. Ripley. (1987) Problems with likelihood estimation of covariance functions of spatial Gaussian processes. *Biometrika*, **74**, (3), 640-642.

Whittle, P. (1954) On stationary processes in the plane. *Biometrika*, **41**, 434-449.

DISCUSSION

"Some remarks on the testing of association between spatial processes"

by Sylvia Richardson

The problem of deciding if spatially distributed variables are associated is one the analyst of spatial data frequently encounters. Richardson has considered three commonly used (and one somewhat less common used) measures in the study of associations between pairs of variables. When one or all of the variables are spatially correlated, the distributions of the measures can be substantially affected. Richardson's paper proposes a method for correcting for these effects and shows by Monte Carlo means that the corrected statistics behave rather well.

Since the theoretical discussion in this paper deals mainly with the conditional distribution of these statistics given all but one of the variables, I, too, will focus on this situation. This is the familiar regression model where one considers the independent variables as given. Under spatial correlation, regression estimates are affected in at least two ways. One is the loss of efficiency, which affects parameter estimates as well as measures of association. The other, perhaps more serious effect is that the distribution of any statistic under the hypothesis of independence is often quite far from that under spatial correlation.

There have been some efforts to address both problems simultaneously using generalized least squares. In such an approach the covariance matrix $\mathbf{\Sigma}$ needs to be estimated, and given that usually there are far more elements of such a matrix than observations, some restrictions need to be placed on this matrix. Richardson's assumption that elements of $\mathbf{Cov}(\mathbf{U})$ be given by a step function on $\mathbf{A} \times \mathbf{A}$ is such an assumption, and, if the steps are large enough, can form the basis on which to estimate $\mathbf{\Sigma}$. Such an estimate would fall under the general rubric of EGLS estimation (Judge, *et al.*, 1985). Several methods exist for obtaining estimates of $\mathbf{\Sigma}$ and thereby the parameter estimates and measures of association (including Rao's MINQUE theory — see Rao, 1973). The difficulty of this approach in many applications is that the estimate of $\mathbf{\Sigma}$ is not positive definite. In that case it is not always clear what needs to be done, although a sometimes satisfactory way out is to delete some observations.

Richardson has taken a different approach, addressing only the second problem. She uses the traditional measures, but corrects them for spatial correlation. The method, though ingenious, is not entirely new. Approximating the distribution of beta–distributed random variables by a beta distribution with parameters chosen on the basis of the first few parameters has been used before (*e. g.*, Theil and Nagar, 1961). The formulæ are easy to compute and the Monte Carlo comparisons show the measures to be functioning well.

If this commentator has any significant questions, they are with the Monte Carlo comparisons. Richardson has compared modified statistics largely with statistics that totally ignore spatial correlation. I would like to see comparisons with EGLS versions of the statistics. If her approach performs as well as or better than these, then her contribution would have been most valuable.

References

Judge, G., W. Griffiths, R. Hill, H. Lutkepohl, and T-C. Lee. (1985) *The Theory and Practice of Econometrics*, 2nd ed. New York: Wiley.

Rao, C. (1973) *Linear Statistical Inference and its Applications*. New York: Wiley.

Theil, H., and A. Nagar. (1961) Testing the independence of regression disturbances. *Journal of the American Statistical Association*, **56**, 793-806.

Ashish Sen, University of Illinois/Chicago