Bulletin of Entomological Research (2005) **95**, 47–56 DOI: 10.1079/BER2004338

Analysis of spatial patterns at a geographical scale over north-western Europe from point-referenced aphid count data

N. Cocu[1], K. Conrad[2], R. Harrington[2] and M.D.A. Rounsevell[1]*

[1]Département de Géographie, Université catholique de Louvain, Place Louis Pasteur 3, 1348 Louvain-la-Neuve, Belgium: [2]Division of Plant and Invertebrate, Ecology, Rothamsted Research, Harpenden, Hertfordshire, AL5 2JQ, UK

Abstract

The spatial analysis by distance indices (SADIE) technique was developed to evaluate the spatial pattern of point-referenced count data as well as the spatial association between two sets of data sharing the same point locations. This paper presents an analysis of spatial patterns in aphid count data and the association of these data with climate across north-west Europe. The paper tests the applicability of the technique to large geographical areas. Aggregation and cluster indices were calculated for the total annual abundance of the peach–potato aphid *Myzus persicae* (Sulzer) and for the annual mean rainfall and temperature at aphid monitoring sites. Association indices demonstrated the stability in time of aphid spatial structures and the correlation between aphid density and climate patterns. Groups of relatively large numbers of aphids, termed patches, and groups of relatively small numbers of aphids, termed gaps, were located and their mean size estimated. The aphid patterns were quite stable in time and the spatial patterns of temperature and rainfall were weakly associated with *M. persicae* annual abundance. Similarities were observed between the results of SADIE and those from the more widely used technique of spatial autocorrelation (SAC). However, the SADIE association index has the advantage of quantifying the possible associations between aphid data and the factors that determine population distribution. Thus, high temperature and low rainfall were identified as environmental factors that were positively associated with aphid abundance across north-west Europe.

Introduction

Spatial patterns are concerned with the location of particular points relative to one another that give rise to random, clustered or regular distributions (Bailey & Gatrell, 1995). Determining whether the geographical distribution is random or not, and describing the pattern, can be used to infer the processes that are responsible for generating the patterns (Legendre & Legendre, 1998). Within this framework, a wide range of techniques can be applied, and these have been developed in various branches of science, e.g. ecology, geography and geology (Dale *et al.*, 2002). Their suitability depends primarily on the characteristics of the data, on the hypothesis to be tested, and on their potential for quantifying spatial structures at the global and local level (Perry *et al.*, 2002). The statistical analyses of spatial data used in ecology were the subject of a recent review that established some conceptual and mathematical relationships among the different methods (Dale *et al.*, 2002), but highlighted that most of these methods seek to answer slightly different questions.

---

*Author for correspondence
Fax: +32 (0)10 47 28 77
E-mail: rounsevell@geog.ucl.ac.be

In ecology, aggregated, regular or random spatial patterns were first differentiated by variance-mean methods (s^2/m, index of cluster size, index of cluster frequency, Lloyd's patchiness and crowding indices, Morisita's index Iδ, and the coefficient of Taylor's power law) (Upton & Fingleton, 1985), but these indices ignored the spatial location of samples, were based on the assumption that sample values were not correlated and sometimes failed to differentiate more complex spatial patterns (Liebhold *et al.*, 1991; Sharov *et al.*, 1996). These shortcomings have led to the development of other methods for specific ecological data types, such as block and quadrat variance methods, neighbour networks and spatial analysis by distance indices (SADIE; see Dale *et al.*, 2002 for a complete review). These latter methods are most often applied in the field of insect and plant ecology.

Statistics measuring spatial autocorrelation (SAC) such as Moran's *I* and Geary's *c*, which are commonly used in geography, are also capable of characterizing spatial patterns, of dealing simultaneously with the location and the values of ecological data, and are applicable to a wide range of data types: ordinal, nominal, or interval data and regularly or irregularly distributed points (Jumars *et al.*, 1977; Sokal, 1978; Sokal & Oden, 1978). This family of statistics is commonly used in transport geography (Black & Thomas, 1998) and in epidemiology (Burra *et al.*, 2002), but is increasingly used to characterize the spatial distribution of species (Judas *et al.*, 2002; Cocu *et al.*, in press, a) or fine-scale spatial structures in genetics and molecular ecology (Gömöry & Paule, 2002).

Methods originally developed in the field of geology, collectively known as geostatistics (variograms, kriging, etc.), have also been used relatively recently to quantify spatial patterns in ecological count data (Sokal *et al.*, 1987; Schotzko & O'Keefe, 1989; Liebhold *et al.*, 1991, 1993; Schotzko & Knudsen, 1992). These techniques are, however, employed less frequently when data are sparse and cover large geographical areas.

The focus of this paper is on point-counts of insects, with data in the form: *x*,*y*,*z*, where *x*,*y* references a spatial location and *z* is the count of insects at that point. The family of analytical techniques known as SADIE has been used to describe the spatial features of point-referenced individual count data and requires only the coordinates of each point (*x, y*) and its associated count (*z*) to be specified (Perry, 1995, 1998a,b; Perry *et al.*, 1996, 1999). Moreover, the technique places no restriction on the number and arrangement of the sample units and irregular spacing of units is perfectly acceptable (Perry *et al.*, 1996; Perry & Klukowski, 1997). The SADIE technique can be applied at two different spatial scales, at a global and a local level (i.e. for the entire study area and for each sample point), and allows hypothesis testing, e.g. determining whether the observed pattern is significantly different from random. The method detects and tests non-randomness and measures the overall spatial pattern by indices. Moreover, local contributions to the degree of global clustering can be identified and quantified for individual sample units. Other developments of the SADIE method allow determination of the spatial association and dissociation between two data sets that share the same spatial locations. This methodology has already proved useful in the study of within-field spatial distributions of insects (Winder *et al.*, 1999), for understanding predator–prey relationships (Holland *et al.*, 1999; Winder *et al.*, 2001) and for determining external factor influences on the spatial distributions of species (Holland *et al.*, 2000; Thomas *et al.*, 2001). However, there has not yet been a published study in which the technique has been used for data points that are distributed over large geographical areas such as north-west Europe. One purpose of this paper, therefore, is to test its applicability in such circumstances.

The methodology applied here is described and illustrated through the use of aphid annual counts for the pest peach–potato aphid *Myzus persicae* (Sulzer) (Hemiptera: Aphididae) collected in suction traps distributed across north-west Europe. *Myzus persicae* is a major pest of crops including potato and sugarbeet and, although it is present throughout north-west Europe, it is not uniformly distributed. The objectives are to assess whether SADIE can be used to reveal spatial structure in the aphid data over large areas, determine the average size of spatial clusters and measure how stable the spatial structure remains over time.

One of the goals in identifying spatial patterns in species distribution and abundance is to aid understanding and identification of the underlying mechanisms. The distribution and dynamics of *Myzus persicae*, like most aphids, are strongly influenced by environmental variables, especially temperature, which has a major effect on developmental and reproductive rates, as well as on life cycle strategy and winter survival (Harrington *et al.*, 1995). Few studies have been able to establish relationships between the spatial structure of environmental factors and the spatial structure of population distributions. An exploration of the spatial structure of climatic data may reveal explanatory relationships. Assuming that the spatial structures of the aphid and climate data are dependent, if the climate is distributed in a heterogeneous way then the aphid spatial structure will be rather similar. If the climate is more homogeneous, it can be supposed that *M. persicae* abundance will also be more homogeneous and, in the same way, a random climatic structure should give rise to a random distribution of abundance. The purpose of this paper, therefore, is to determine whether there are associations between climate and aphid spatial patterns. SADIE techniques are used to determine whether environmental variables are associated with the aphid distributions.

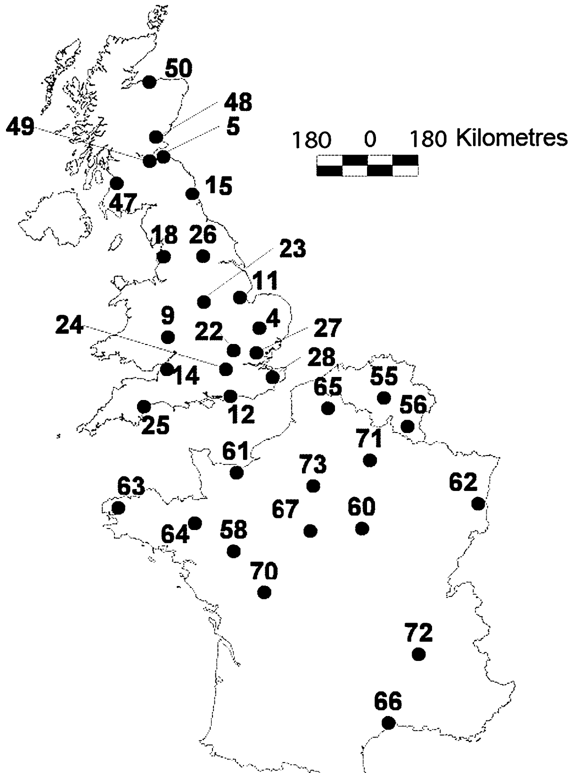
Cocu *et al.* (in press, a) have recently measured spatial autocorrelation (SAC) on the same dataset, using global and local Moran's I statistics. These analyses highlighted the presence of an aggregated structure in the aphid data except in 1989, 2000 and 2001. The technique also differentiated the location of homogeneous zones from heterogeneous zones, but did not give any information on the density in these zones, i.e. high or low aphid counts. Consequently, in this study, the potential contribution of each technique (SADIE and SAC) is discussed and the results are compared.

Materials and methods

Data

Aphid data

Aphid data were obtained from the Exploitation of Aphid Monitoring in Europe (EXAMINE) database which covers the entire European network of suction traps providing daily, standardized, quantitative information for many species (http://www.rothamsted.bbsrc.ac.uk/examine/; Harrington *et al.*, 2004). Data were selected for the peach–potato



4. Broom's Barn
5. East Fenton
9. Hereford
11. Kirton
12. Littlehampton
14. Long Ashton
15. Newcastle
18. Preston
22. Rothamsted
23. Shardlow
24. Silwood
25. Starcross
26. Tadcaster
27. Writtle
28. Wye
47. Ayr
48. Dundee
49. East Craigs
50. Elgin
55. Gembloux
56. Libramont
58. Angers
60. Auxerre
61. Caen
62. Colmar
63. Landerneau
64. Le Rheu
65. Loos
66. Montpellier
67. Orléans
70. Poitiers
71. Reims
72. Valence
73. Versailles

Fig. 1. Distribution map of the aphid data sampling points extracted from the European EXAMINE suction trap network.

aphid *M. persicae* because of its importance as a pest throughout north-west Europe. The data were aggregated to annual counts of individuals for 34 sites within Great Britain, France and Belgium (fig. 1). This small number of sample sites would limit the value of traditional methods of quantifying spatial patterns. However, the use of such sparse data is acceptable to SADIE because of the differences in its underlying philosophy compared to other techniques (Dale *et al*., 2002; Perry *et al*., 2002).

Climatic data

Climate data were obtained from the mean monthly data set generated by the European Union Advanced Terrestrial Ecosystem Analysis and Modelling (ATEAM) project (Mitchell *et al*., 2004). These data were interpolated from large numbers of weather stations across Europe to a 10-min longitude/latitude grid. The annual mean temperature and precipitation data were extracted for the 10-min grid cells to which each trap belongs for the years 1989 to 2000. These data were therefore considered also as point-referenced data with a climatic attribute that was converted into integer form, a requirement of the SADIE method.

Methods

General approach

For the data of *M. persicae* annual abundance, the work was based on four stages: (i) quantification of the global spatial pattern for each year; (ii) quantification of the global and local degree of clustering for each year; (iii) visualization of the results and quantification of the mean cluster size by generating smoothed surfaces from the cluster indices; and (iv) measurement of the spatial association indices for pairs of years for the aphid data. These tests allowed an investigation of the spatial, temporal and spatio-temporal changes in the distribution of *M. persicae*.

Analysis of climate data was based on four stages: (i) quantification of the global spatial pattern by the use of aggregation indices; (ii) quantification of the global and local degree of clustering; (iii) visualization of the results by generating smoothed surfaces from the cluster indices; and (iv) quantification of the spatial association between the spatial patterns of aphid and climate data.

Indices of aggregation

The SADIE technique consists of quantifying the spatial pattern in a sampled population by measuring the total distance that the individuals of the observed sample must move in order to reach a regular arrangement in space. Regularity corresponds to the situation where the sampled individuals are as dispersed as possible. Perry (1998a) defined I_a, the index of aggregation. A value of $I_a > 1$ indicates the presence of an aggregated sample, an index equal to 1 characterizes a spatially random sample and a value of $I_a < 1$ indicates the presence of a regular sample. The index has an associated probability (P_a) that the data are not distributed randomly. This probability is obtained by comparing the observed spatial pattern with the corresponding values obtained by random permutations of the observed counts among the sample units.

Cluster indices and graphical displays

Two standardized dimensionless cluster indices: ν_i and ν_j (Perry *et al*., 1999) have been used in order to provide a local measure of the degree of small-scale clustering for each individual sample point. These indices quantify the degree to which the sampled count contributes toward clustering either as part of a patch (a cluster of relatively large numbers of aphids) or a gap (a cluster of relatively low numbers of aphids). Winder *et al*. (1999, 2001) defined clusters as areas enclosed by contour levels of $+1.5$ or -1.5. When $\nu_i > 1.5$, the index indicates patchiness; when $\nu_J < -1.5$, it conveys the idea that the sampled point has membership of a gap, and when the cluster index equals one, it indicates a random placement of that point value in relation to others. Consequently, a map of the degree of clustering, the cluster plot (red-blue plot), can help to visualize the areas where population counts are relatively large or small. A formal test of the overall degree of clustering into patches and gaps in the

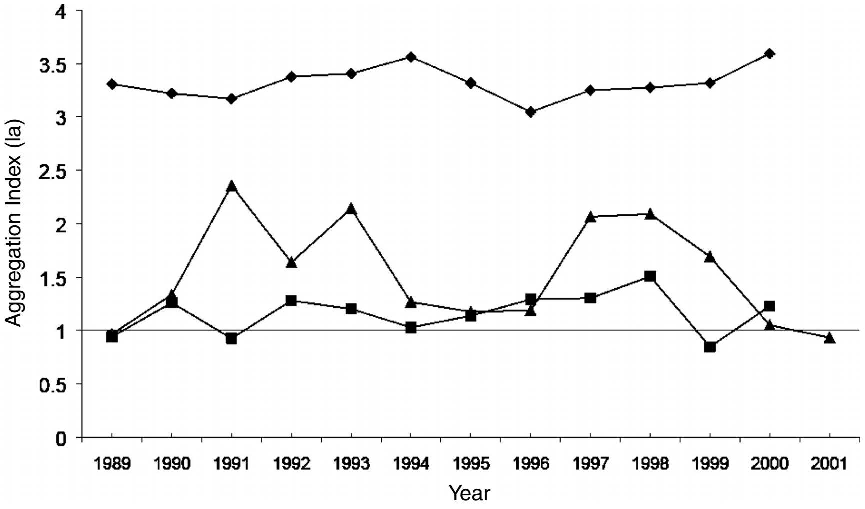


Fig. 2. Global spatial pattern analysis of the annual abundance of *Myzus persicae* (log(n + 1)) and climate in north-western Europe. Plots of the spatial aggregation index I_a versus year for *M. persicae* annual abundance (▲); mean annual temperature (◆) and mean annual rainfall (■).

entire dataset is provided by their overall average values $\bar{v}_i$ and $\bar{v}_j$ that are compared with their corresponding values from randomization (Perry *et al.*, 1999).

Overall association indices

Two sets of data may be spatially positively associated, negatively associated or occur at random with respect to one another (Perry, 1998b). An overall spatial association index, *X* (Perry & Dixon, 2002) may be used to quantify the relationship of the spatial structure between two sets of counts sharing the same location. *X* is equivalent to the correlation coefficient between the clustering indices for the two sets of counts. *X* also corresponds to the mean of the local spatial association values χ_k and its significance can be tested by Dutilleul-adjusted randomizations (Dutilleul, 1993; Winder *et al.*, 2001; Perry & Dixon, 2002). At the local scale, positive association suggests coincidence of patch clusters or of gap clusters for the two sets, whereas negative association is indicated by a patch coinciding with a gap (Perry & Dixon, 2002).

Results

Aphid spatial pattern

The distribution of *M. persicae* in 1989, 2000 and 2001 was different from the other years (fig. 2). The index of aggregation I_a was minimal, nearly equal to 1 and was not significant, which indicates that for these years there were no major clusters at a global scale in the data set. Conversely, a value of I_a higher than 1 and a low value of P_a suggest that there was a strong and significant aggregation, e.g. for 1991 to 1993 and for 1997 to 1999. For the other years, the value of I_a indicates the presence of a moderately strong spatial pattern, but one which is not significant.

Compared to the other years, which all show some spatial structuring, the average clustering indices of the SADIE technique shown in fig. 3 also confirm that 1989, 2000 and 2001 were different since no overall spatial clusters either in patches or in gaps can be identified: $|\bar{v}_i|$ and $|\bar{v}_j|$ were minimal, nearly equal to 1 and were not significant. These indices emphasize the observation that there was strong clustering into patches and gaps in 1991, 1993, 1999 ($|\bar{v}_i|$ and $|\bar{v}_j| > 1.5$); into patches in 1992 and 1994 ($|\bar{v}_i| > 1.5$), and into gaps in 1997 and 1998 ($\bar{v}_j < -2$), but the overall clustering indices were not significant in 1990, 1994, 1995 and 1996. Clusters were located by the position of their centroid and there appears to be a trend in their location over the period considered (only the years 1989, 2000, 2001 and 1991 to 1993 are shown in fig. 4): the south-east of Great Britain and the north-east and south-east of France were usually clustered as patches (the maximum counts are recorded at Valence (site 72) (1553 *M. persicae* in 1991 and 990 *M. persicae in* 1993) and at Poitiers (site 70) (4976 *M. persicae* in 1997)), while the Scottish traps and the western part of Great Britain were mainly grouped as gaps. It is likely that these patterns reflect a climatic or a land use effect because individuals of *M. persicae* were abundant in warmer areas that were rich in host plants, and less abundant in wetter zones characterized by less arable land.

Measurement of the association indices for pairs of aphid data (table 1) showed that the spatial pattern was quite stable in time at the beginning of the considered time-series. For the first five years (1989 to 1993), there were positive associations at lag one (just above the diagonal), i.e. the spatial pattern was fairly stable from year to year, but then the system switched to more-random changes between years. The association indices suggest that the spatial patterns observed in 1994 and 1995 were still similar to the structure observed in the previous years but, after this, dissimilarity occurred, especially in 1996, 1997, 1999 and 2000. Moreover, even if both 1989 and 2001 were characterized by the absence of an overall spatial structure, the negatively significant value of the association index suggests that the patterns were dissimilar. The plot of the mean overall association index, *X*, versus lag period (analogous to a correlogram of Moran's I)

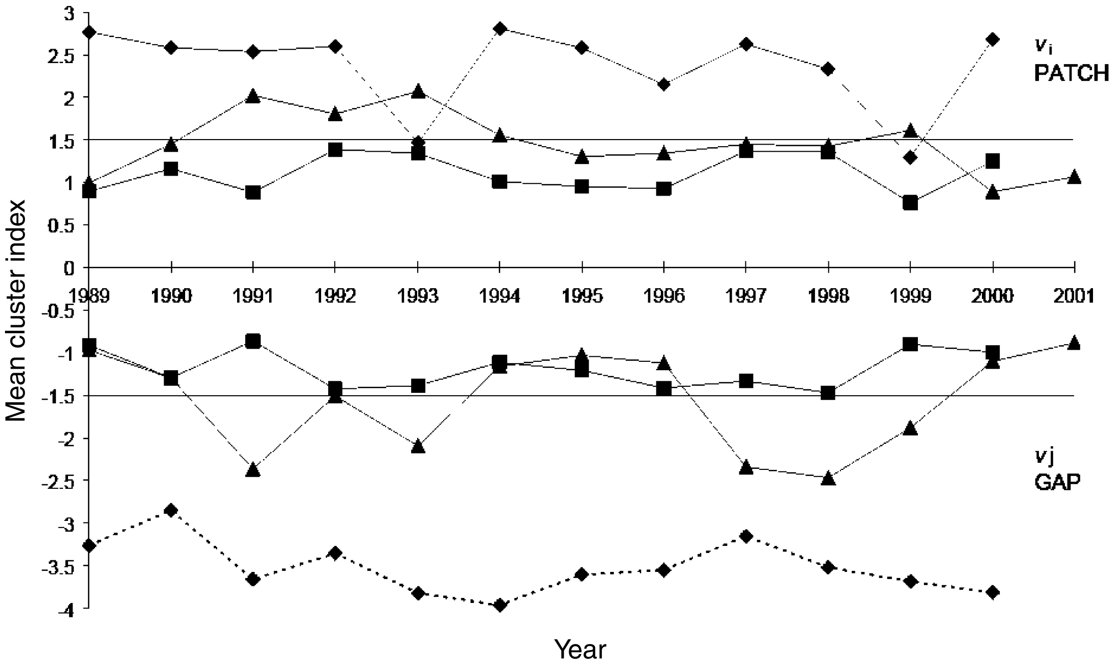


Fig. 3. Local spatial pattern analysis of the annual abundance of *Myzus persicae* (log(n + 1)) and climate (around each aphid monitoring site). Plots of mean clusters indices ($|\bar{v}_i|$and $|\bar{v}_j|$) versus year for *M. persicae* annual abundance (▲); mean annual temperature (◆) and rainfall (■).

confirms these observations and indicates that similarity in aphid catches between years decreased during the first seven years, reaching zero at lag 7 (fig. 5).

Climatic spatial pattern

The aggregation index (fig. 2) was not significant for precipitation. There was no particularly strong spatial structure associated with the mean annual precipitation data across north-west Europe (I_a nearly equal to 1) except in 1998, where I_a was maximum and significantly exceeded the value of 1, implying the presence of an aggregated pattern. Thus, the precipitation data were not characterized by a particular spatial structure at this scale, and this finding is broadly valid for the whole of the considered period. Conversely, the index calculated for temperature demonstrates the presence of a strong and highly significant aggregated structure. The index of aggregation I_a represented in fig. 2 is higher than 3 for each year.

The mean cluster indices for rainfall (fig. 3) confirm the absence of spatial clusters at the global level as well as in the form of patches or gaps. However, at the local level, the location of patches and gaps was rather stable; the south-east of the UK as well as the north-east of France form zones of low precipitation, i.e. gaps, and the west of the UK form zones of high precipitation, i.e. patches (e.g. years 1991 to 1993 in fig. 6). For temperature, the indices (fig. 3) confirm the presence of a strong spatial structure defined by the presence of both patches and gaps. The cluster plots allow visualization of these structures at a local level (e.g. years 1991 to 1993 in fig. 6). The location of the patches and the gaps was stable over time: the south-west of France was a patch area since the climate there was milder, whilst the UK and the north-east of France form homogeneous zones where the annual mean temperature was cooler.

Spatial association between the spatial patterns of aphids and climate

The results of the cluster analysis on climate and the spatial distribution of *M. persicae* abundance can be compared visually in figs 4 and 6. The location of *M. persicae* patches (zones of high abundance) appear to correspond with the location of rainfall gaps (zones of low levels of precipitation) and vice versa and, in the same way, the *M. persicae* patches seem linked to the temperature patches (zones of high temperature), although this could correspond with more favourable zones in terms of land use.

The overall association indices obtained between *M. persicae* and rainfall, and between *M. persicae* and temperature were generally weak (fig. 7) and do not indicate either a significant association or disassociation. However, for some specific years, the results suggest a slight negative and significant relationship with rainfall (in 1992, 1994, 1995 and 1998) and a slight positive and significant association with temperature (in 1991, 1997 and 1998).

Discussion

In spite of the description of a random pattern for rainfall and of an aggregated distribution of temperature over the period 1989–2000, the results do not suggest that particular climatic conditions explain the random spatial structure of *M. persicae* abundance observed in 1989, 2000 and 2001. There are, however, indications of a possible influence of climate, both temperature and precipitation, on the aphid

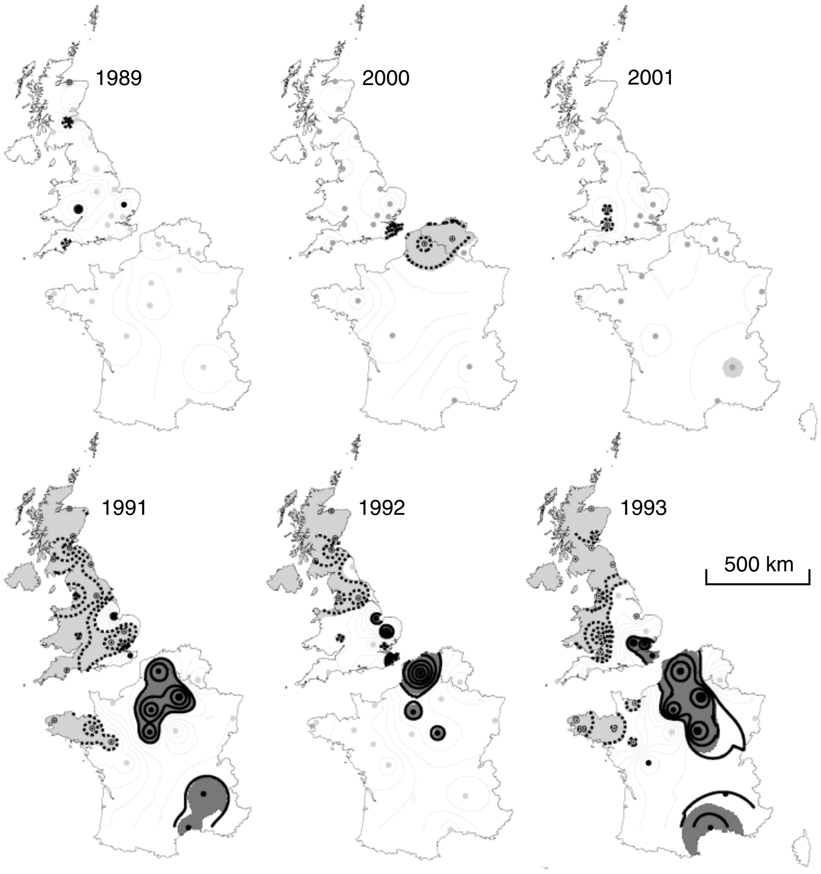


Fig. 4. Local spatial pattern analysis of the annual abundance of *Myzus persicae* ($\log(n+1)$). The years 1989, 2000 and 2001 represent the case where no structure has been identified at the global scale and the years 1991 to 1993 are presented as examples of the years with a global aggregated spatial pattern. Contour levels have been created on the basis of cluster indices interpolated in Arc View 3.2 by using the inverse distance weights method and a nearest neighbour approach. Light grey shading and dashed lines indicate strong gaps with index values < -1.5 ; dark grey shading and continuous lines indicate strong patchiness with index value > 1.5.

distribution, which suggests an association between these different patterns. The pattern of the results is more interesting than the significance level. Indeed, the results suggest that temperature and rainfall had opposing effects: precipitation was almost always negatively associated and temperature almost always positively associated with the aphid spatial pattern. Thus, there may be a slight influence of temperature (positive) and of rainfall (negative) on the aphid spatial structure, but these relationships were not strong at the global level. It is possible that climate influences the spatial structure of *M. persicae* abundance by reflecting the geographical distribution of suitable habitat. Thus, the climate may affect aphid abundance either directly or indirectly through its effect on host plant distribution.

Table 1. Results of the spatial association analysis of the annual abundance of *Myzus persicae* across north-west Europe.

Years	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
1989	0.58*	0.16	0.42	0.23	0.62*	0.33	−0.11	0.22	0.20	0.17	0.23	−0.48*
1990		0.49*	0.55*	0.54*	0.35	0.52*	−0.14	−0.12	0.30	−0.12	−0.20	−0.01
1991			0.68*	0.45*	0.05	0.31	0.25	−0.23	0.47*	0.00	−0.11	0.46*
1992				0.62*	0.22	0.51*	0.02	−0.10	0.41	−0.07	−0.05	0.31
1993					0.04	0.39*	−0.14	−0.21	0.34	−0.16	−0.13	0.35
1994						0.51*	−0.30	−0.21	0.36	−0.08	0.04	−0.17
1995							−0.22	−0.46*	0.13	−0.04	−0.23	−0.12
1996								0.32	0.00	0.14	−0.09	0.27
1997									0.09	0.13	0.56*	−0.06
1998										0.23	0.48	0.44
1999											0.39	0.01
2000												0.02

Values of the overall spatial association index X between pairs of years; $*P < 0.05$.

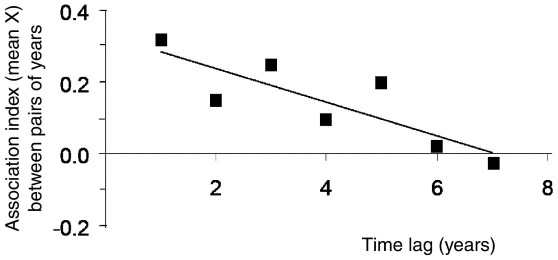


Fig. 5. Plot of the overall association index, *X*, between pairs of years, averaged for each lag period. Only the first seven years have been plotted as calculations for longer lags involve less than one half of the association measures. A downward linear trend suggests that similarity in *Myzus persicae* catches decreases to 0 at lag 7 ($R^2 = 0.70$).

However, disaggregating these two effects has not been possible here due to a lack of data on host plants. Analysing the spatial pattern of host plant distribution, i.e. of land use, would be relevant to aid understanding of the mechanisms that drive *M. persicae* abundance. Indeed, Cocu *et al*. (in press, b) highlighted that specific land use types within agricultural crops played a key role in determining total aphid numbers in north-west Europe.

Comparison of methodologies

This paper illustrates the use of the SADIE technique to study spatial patterns in *M. persicae* counts from suction traps. Aggregation at a global scale was quantified and local clusters, either patches or gaps, were identified and localized. SADIE is not the only methodology that could have been applied at a global and local scale to aphid count

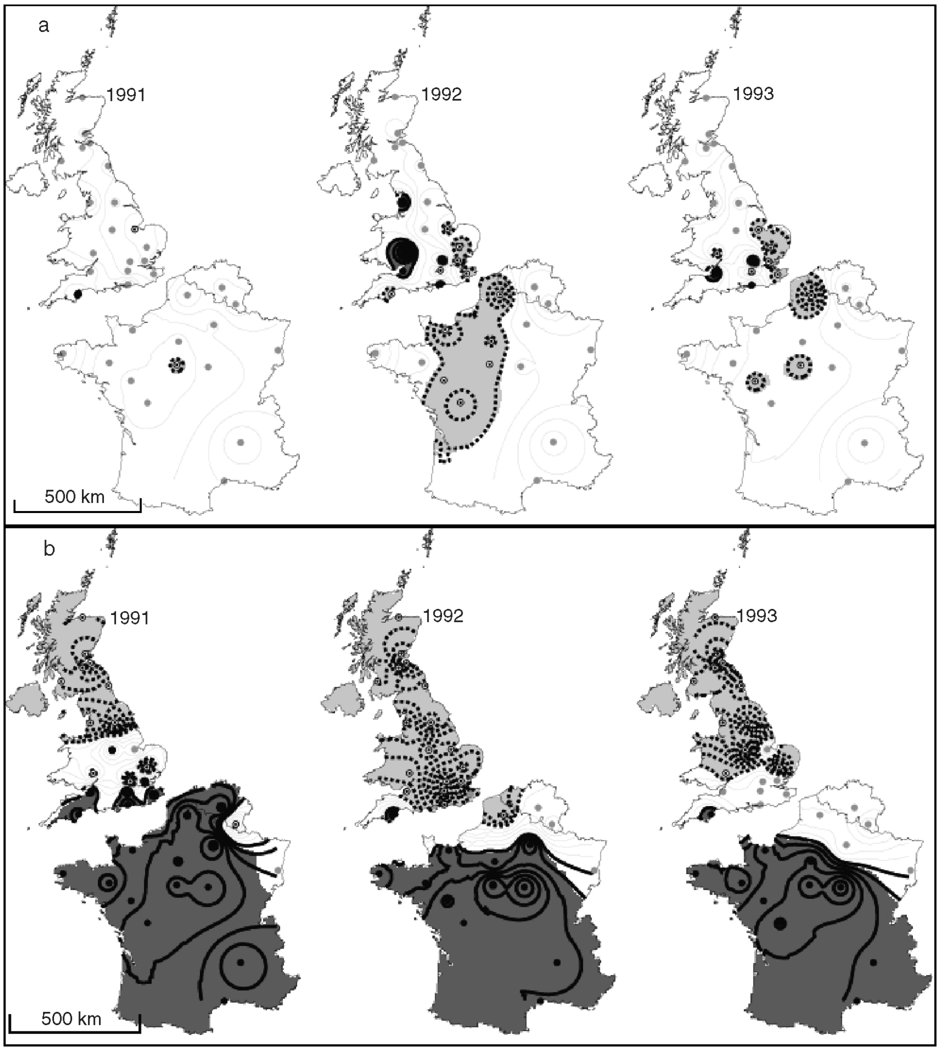


Fig. 6. Local spatial pattern analysis of climate data (annual mean): Plots and overlaid contour of SADIE clustering indices for precipitation (a) and temperature (b) for the years 1991 to 1993. Contour levels have been created on the basis of cluster indices interpolated in Arc View 3.2 by using the inverse distance weights method and a nearest neighbour approach. Light grey shading and dashed lines indicate strong gaps with index values < -1.5 ; dark grey shading and continuous lines indicate strong patchiness with index value > 1.5.

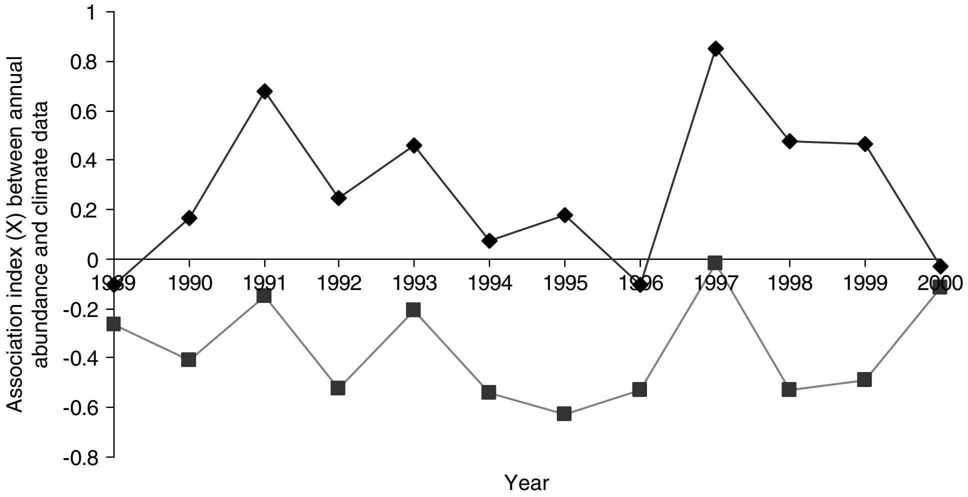


Fig. 7. Association in the spatial structure of *Myzus persicae* abundance and climate data. Plot of the overall spatial association index, *X* for annual abundance (log($n+1$)) of *M. persicae* and mean annual temperature (dashed line) and for annual abundance (log($n+1$)) of *M. persicae* and mean annual rainfall (continuous line) over 1989–2001.

data that would fulfil the same prerequisites (coping with irregularly spaced data, hypothesis testing, providing information on multiple scales). A technology much more familiar to geographers, spatial autocorrelation analysis (SAC), is also useful (table 1 in Perry *et al.*, 2002). In a previous study, Cocu *et al.* (in press, a) used the Moran index on the same data set as reported here to estimate the presence or absence of spatial autocorrelation at all sites across the study area (global level), to identify the possible contributions of each site to the global index (local Moran index of spatial autocorrelation – LISAs) and to try to quantify the area that is represented by a suction trap. The two techniques address slightly different questions: spatial autocorrelation analysis focuses on the identification and differentiation of homogeneous zones (either of low or high counts) and heterogeneous zones, whereas SADIE is not concerned with heterogeneous groups, but allows differentiation between homogeneous areas with relatively large counts and areas with small counts.

Despite these differences, similarities exist between the results obtained with the SADIE technique and the Moran index. Both techniques were unable to distinguish a spatial pattern at the global scale for the years 1989, 2000 and 2001. However, although the presence of local spatial autocorrelation or local clusters was identified by both techniques, the location of these areas of similarity at a local scale was slightly different. These differences may result from the fact that, for Moran's I, the weightings are not conditional on the abundance values, but rather on the distance between points. Conversely, the SADIE technique measures and maps spatial pattern, giving little weight to isolated, large, or small values (Perry *et al.*, 2002).

Finally, the SADIE cluster plots localize patches and gaps and quantify their mean size while the spatial autocorrelation analysis can identify local pockets of spatial autocorrelation (by the use of LISAs) and quantify the range of distances over which the pattern exists. The distance over which the Moran index is positive can be interpreted as an estimate of the dominant cluster size (Perry *et al.*, 2002). Therefore, the spatial autocorrelation analysis of *M. persicae* annual abundance indicated similarity between traps over a large range; *c.* 500 km (Cocu *et al.*, in press, a). The clusters can be relatively large and their size can be quantified by the area of the region enclosed by the appropriate contour level. It is, however, necessary to pay attention to the edge effect for the clusters located at the margins of the sampled areas. If the size of the closed clusters entirely within the study area is averaged over the full period, the distance of influence of a trap (radius or width) can be estimated coarsely at between 110 km and 180 km if the form of the cluster is approximated to a circle or to a square, respectively. Assuming that the clusters are closed by the borders of the study area, the radius or width change can be estimated coarsely to be between 160 km and 290 km. These values are much lower than for the Moran results, but the values are probably underestimations because of the approximation of the cluster shapes and by the constraint of closing clusters by the border of the study area. Although this observation agrees with values quoted in the literature, e.g. a circle of about 100 km (Hullé & Gamon, 1990), there appear reasonable grounds to believe that the similarity between traps can occur over larger distances (Cocu *et al.*, in press, a).

An important advantage of the SADIE technique is the graphic and cartographic facilities that allow presentation of visual information to support conclusions drawn from the indices. The cluster plots allow identification of global trends in the spatial pattern of data such as, in this particular case, a south-east towards north-west gradient of *M. persicae* abundance. Moreover, SADIE is a useful mapping tool that is lacking in spatial autocorrelation analysis, since the location of homogeneous high or low aphid counts allows the construction of hypotheses regarding the factors that explain aphid spatial structure. Contrary to spatial autocorrelation analysis, SADIE is not able to quantify spatial data structures in relation to distance lags, but spatial association can be quantified in relation to time lags. The spatial correlogram of spatial autocorrelation is important as it gives clues to the factors which can explain the underlying spatial structure and the associated structure function can have implications for interpolation between unsampled counts. However, as pointed out by Legendre & Legendre (1998) the information provided by a correlogram is not specific and a blind interpretation may often be misleading if not supported by maps. It seems reasonable to conclude, therefore, that different methods should be used in combination to cover all aspects of the spatial characteristics of a data set (Perry *et al*., 2002). The results indicate that the use of SADIE together with spatial autocorrelation analysis allows characterization of most spatial aspects of the aphid data.

The SADIE technique is also interesting because of its use of association measures, which allow an evaluation of the stability of the aphid spatial patterns over time. The results of the analysis presented here showed that the spatial structure of *M. persicae* was rather stable, suggesting a relationship between the species and surrounding environmental factors. For example, the stability in the aphid spatial patterns could reflect stability in underlying environmental factors such as the mean annual temperature. The results of the analysis undertaken here suggest that climate plays a role in structuring the annual abundance of *M. persicae* across north-west Europe. However, neither the pattern of the precipitation variable nor the spatial structure of the temperature was able to explain anomalous years. Thus, other variables may be acting on the aphid data structure such as land-use variables that determine food resources, the presence of suitable host plants and natural enemies. Further analyses, based on multiple linear regression, residual maximum likelihood and artificial neural networks, are thus planned in order to improve understanding and knowledge of the environmental variables which contribute to these processes.

Conclusions

The analysis presented here investigated the spatial, temporal and spatio-temporal changes in the distribution of *M. persicae*. Over the period 1989–2001, the spatial pattern of *M. persicae* was shown to be quite stable in time and generally aggregated either into groups of relatively large annual counts (patches) and small annual counts (gaps). SADIE is an effective technique for analysing biogeographical spatial patterns even over large geographical areas. The absence of satisfactory tools to describe structures at small spatial scales, i.e. over wide geographic areas, and for relatively restricted data sets, is the principal reason why SADIE provides an interesting way to progress. Such techniques should, however, be combined with other types of analysis. For example, SADIE can be supplemented by correlograms which add to the quantification of the global and local spatial patterns, providing information concerning the evolution of this structure at several distances. It is essential to consider the whole range of existing spatial analysis techniques, independently of their applicability, because each was developed with its own characteristics, in particular concerning the level of prerequisites (i.e. organization of the data) and objectives. Moreover, the SADIE association tool provides many interesting possibilities for geographical analyses. In this analysis, the spatial patterns of the mean annual temperature and rainfall were shown to be weakly associated with the structure of *M. persicae* annual abundance. The use of these association indices has already been effective in other ecological fields, for example, linking patterns of virus infection with patterns of insect vectors (Korie *et al*., 2000) and understanding predator–prey interactions (Winder *et al*., 2001). Furthermore, recent developments that permit the calculation and mapping of local spatial association measures could allow further exploration of the relationships between a species and its habitat.

Acknowledgements

This work was supported by the European Commission funded project EXAMINE (contract no. EVK2-CT1999-20001) and by a grant from the Université catholique de Louvain. The authors are very grateful to: Maurice Hullé (Institut National de la Recherche Agronomique (INRA), Le Rheu), Jon Pickup (Scottish Agricultural Science Agency, Edinburgh), and Jean-Louis Rolot (Centre de Recherches Agronomique de Gembloux) for providing their aphid data; Colin Denholm (Rothamsted Research), Paul Verrier (Rothamsted Research) and Damien Maurice (INRA, Le Rheu) for managing the aphid database; the ATEAM project for providing the climatic data and Joe Perry (Rothamsted Research) for commenting on a previous version of the manuscript. Rothamsted Research receives grant aided support from Biotechnology and Biological Sciences Research Council of the United Kingdom.

References

Bailey, T.C. & Gatrell, A.C. (1995) *Interactive spatial data analysis*. Longman Scientific & Technical.

Black, W.R. & Thomas, I. (1998) Accidents on Belgium's motorways: a network autocorrelation analysis. *Journal of Transport Geography* **6**, 23–31.

Burra, T., Jerrett, M., Burnett, R.T. & Anderson, M. (2002) Conceptual and practical issues in the detection of local disease clusters: a study of mortality in Hamilton, Ontario. *Canadian Geographer* **46**, 160–171.

Cocu, N., Harrington, R., Hullé, M. & Rounsevell, M.D.A. (in press, a) Spatial autocorrelation as a tool for identifying the geographic patterns of aphid annual abundance. *Agricultural and Forest Entomology*.

Cocu, N., Harrington, R., Rounsevell, M.D.A., Worner, S.P. & Hullé, M. (in press, b). Geographic location, climate and land use influences on the phenology and numbers of the aphid, *Myzus persicae* in Europe. *Journal of Biogeography*.

Dale, M.R.T., Dixon, P., Fortin, M.-J., Legendre, P., Myers, D.E. & Rosenberg, M.S. (2002) Conceptual and mathematical relationships among methods for spatial analysis. *Ecography* **25**, 558–577.

Dutilleul, P. (1993) Modifying the t-test for assessing the correlation between two spatial processes. *Biometrics* **49**, 305–314.

Gömöry, D. & Paule, L. (2002) Spatial and microgeographical genetic differentiation of black alder (*Alnus glutinosa* Gaertn.) populations. *Forest Ecology and Management* **160**, 3–9.

Harrington, R., Bale, J.S. & Tatchell, G.M. (1995) Aphids in a changing climate. pp. 125–155 *in* Harrington, R. & Stork, N.E. (*Eds*) *Insects in a changing environment*. London, Academic Press.

Harrington, R., Verrier, P., Denholm, C., Hullé, M., Maurice, D., Bell, N., Knight, J., Rounsevell, M., Cocu, N., Barbagallo, S., Basky, Z., Coceano, P.-G., Derron, J., Katis, N., Lukášová, H., Marrkula, I., Mohar, J., Pickup, J., Rolot, J.-L., Ruszkowska, M., Schliephake, E., Seco-Fernandez, M.-V., Sigvald, R., Tsitsipis, J. & Ulber, B. (2004) 'EXAMINE' (EXploitation of Aphid Monitoring in

Europe): a EU thematic network for the study of global change impacts on aphids. pp. 45–49 *in* Simon, J.C., Dedryver, C.A., Risp, C. & Hullé, M. (*Eds*) *Aphids in a new millennium*. Versailles, INRA Editions.

Holland, J.M., Winder, L. & Perry, J.N. (1999) Arthropod prey of farmland birds: their spatial distribution within a sprayed field with and without buffer zones. *Aspects of Applied Biology* **54**, 53–60.

Holland, J.M., Winder, L. & Perry, J.N. (2000) The impact of dimethoate on the spatial distribution of beneficial arthropods in winter wheat. *Annals of Applied Biology* **136**, 93–105.

Hullé, M. & Gamon, A. (1990) Relations entre les captures des pièges à succion du reseau AGRAPHID. pp. 165–177 *in* Cavalloro, R. (*Ed.*) *Euraphid network: trapping and aphid prognosis*. Luxembourg, Commission of the European Communities.

Judas, M., Dornieden, K. & Strothmann, U. (2002) Distribution patterns of carabid beetle species at the landscape-level. *Journal of Biogeography* **29**, 491–508.

Jumars, P.A., Thistle, D. & Jones, M.L. (1977) Detecting two-dimensional spatial structure in biological data. *Oecologia* **28**, 109–123.

Korie, S., Perry, J.N., Mugglestone, M.A., Clark, S.J., Thomas, C.F.G. & Mohamad Roff, M.N. (2000) Spatiotemporal associations in beetle and virus count data. *Journal of Agricultural, Biological and Environmental Statistics* **5**, 214–239.

Legendre, P. & Legendre, L. (1998) *Numerical ecology*, 2nd edn. Amsterdam, Elsevier Science B.V.

Liebhold, A.M., Zhang, X., Hohn, M.E., Elkinton, J.S., Ticehurst, M., Benzon, G.L. & Campbell, R.W. (1991) Geostatistical analysis of gypsy moth (Lepidoptera: Lymantriidae) egg mass populations. *Environmental Entomology* **20**, 1407–1417.

Liebhold, A.M., Rossi, R.E. & Kemp, W.P. (1993) Geostatistics and geographic information systems in applied insect ecology. *Annual Review of Entomology* **38**, 303–327.

Mitchell, T.D., Carter, T.R., Jones, P.D., Hulme, M. & New, M. (2004) *A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901–2000) and 16 scenarios (2001–2100)*. Working Paper, Tyndall Centre for Climate Change Research, University of East Anglia, Norwich (in press).

Perry, J.N. (1995) Spatial analysis by distance indices. *Journal of Animal Ecology* **64**, 303–314.

Perry, J.N. (1998a) Measures of spatial pattern for counts. *Ecology* **79**, 1008–1017.

Perry, J.N. (1998b) Measures of spatial pattern and spatial association for counts of insects. pp. 21–33 *in* Baumgartner, J., Brandmayr, P. & Manly, B.F.J. (*Eds*) *Population and community ecology for insect management and conservation. Proceedings of the Ecology and Population Dynamics Section of the 20th International Congress of Entomology, Florence, Italy, 25–31 August 1996*. Rotterdam, Balkema.

Perry, J.N. & Dixon, P. (2002) A new method for measuring spatial association in ecological count data. *Ecoscience* **9**, 133–141.

Perry, J.N., Bell, E.D., Smith, R.H. & Woiwod, I.P (1996) SADIE: software to measure and model spatial pattern. *Aspects of Applied Biology* **46**, 95–102.

Perry, J.N. & Klukowski, Z. (1997) Spatial distributions of counts at the edges of sample areas. pp. 103–108 *in Proceedings of the VI Conference of the Biometric Society (SpanishRegion), Córdoba, 21–24. September 1997*.

Perry, J.N., Winder, L., Holland, J.M. & Alston, R.D. (1999) Red-blue plots for detecting clusters in count data. *Ecology Letters* **2**, 1106–1113.

Perry, J.N., Liebhold, A.M., Rosenberg, M.S., Dungan, J., Miriti, M., Jakomulska, A. & Citron-Pousty, S. (2002) Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data. *Ecography* **25**, 578–600.

Schotzko, D.J. & Knudsen, G.R. (1992) Use of geostatistics to evaluate a spatial simulation of Russian wheat aphid (Homoptera: Aphididae) movement behavior on preferred and nonpreferred hosts. *Environmental Entomology* **21**, 1271–1282.

Schotzko, D.J. & O'Keefe, L.E. (1989) Geostatistical description of the spatial distribution of *Lygus hesperus* (Heteroptera: Miridae) in lentils. *Environmental Entomology* **82**, 1277–1288.

Sharov, A.A., Liebhold, A.M. & Roberts, E.A. (1996) Spatial variation among counts of gypsy moths (Lepidoptera: Lymantriidae) in pheromone-baited traps at expanding population fronts. *Environmental Entomology* **25**, 1312–1320.

Sokal, R.R. (1978) Spatial autocorrelation in biology. 2. Some biological implications and four applications of evolutionary and ecological interest. *Biological Journal of the Linnean Society* **10**, 229–249.

Sokal, R.R. & Oden, N.L. (1978) Spatial autocorrelation in biology. 1. Methodology. *Biological Journal of the Linnean Society* **10**, 199–228.

Sokal, R.R., Oden, N.L. & Barker, J.S.F. (1987) Spatial structure in *Drosophila buzzatii* populations: simple and directional spatial autocorrelation. *American Naturalist* **129**, 122–142.

Thomas, C.F.G., Parkinson, L., Griffiths, G.J.K., Fernandez Garcia, A. & Marshall, E.J.P. (2001) Aggregation and temporal stability of carabid beetle distributions in field and hedgerow habitats. *Journal of Applied Ecology* **38**, 100–116.

Upton, G. & Fingleton, B. (1985) *Spatial data analysis by example. Point pattern and quantitative data*. Volume 1. New York, John Wiley & Sons.

Winder, L., Perry, J.N. & Holland, J.M. (1999) The spatial and temporal distribution of the grain aphid *Sitobion avenae* in winter wheat. *Entomologia Experimentalis et Applicata* **93**, 277–290.

Winder, L., Alexander, C., Holland, J.M., Woolley, C. & Perry, J.N. (2001) Modelling the dynamic spatio-temporal response of predators to transient prey patches in the field. *Ecology Letters* **4**, 568–576.

(Accepted 23 September 2004)