Assessing regional scale weed distributions, with an Australian example using *Nassella trichotoma*

S W LAFFAN
School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW, Australia

Received 27 April 2005
Revised version accepted 4 October 2005

Summary

Knowledge of the spatial distribution of weed infestations over regional scales is essential for effective management of source populations and to assess future threats. To this end, the distributions of *Nassella trichotoma* across a study area in south-east New South Wales, Australia, were analysed using the geographically local Getis–Ord G_i* spatial hotspot clustering statistic. The clustering of *N. trichotoma* observations was analysed at three infestation levels: presence (at any density), patch level and the occasional plant level. The results indicate that there are *c.* 578 km^2 of cells containing *N. trichotoma* in strongly clustered infestations, 11.2 km^2 within weakly clustered infestations distinct from the main clusters, and 55 km^2 that are not clustered. There are 117 km^2 of strongly clustered patch level cells, 3 km^2 in distinct but weak clusters, and none outside of a cluster area. Of the occasional plant level cells, 329 km^2 are strongly clustered, 6.2 km^2 are in distinct but weak clusters, and 19 km^2 are not clustered. These results provide a mechanism by which control efforts can be prioritized. The analysis approach described in this paper provides a consistent, quantitative and repeatable approach to assess weed infestations across regional scales and can be applied to any weed species for which spatial distribution data are available.

Keywords: *Nassella trichotoma*, serrated tussock, spatial analysis, spatial clustering, weed management.

LAFFAN SW (2006) Assessing regional scale weed distributions, with an Australian example using *Nassella trichotoma*. *Weed Research* **46**, 194–206.

Introduction

The assessment of the spatial distribution of weed infestations over regional scales in a consistent, quantitative and repeatable way is essential for effective management of source populations and for future threat assessments. Spatial analyses of weed distributions provide a tool to assist managers to prioritize control options. This might be through identifying those infestations that are of a controllable extent and density now, those which will require longer-term control efforts, and perhaps those that are beyond any current control methods. Such analyses could also be used to provide benchmark data against which the performance of subsequent control measures can be assessed. Depending on what managers are trying to achieve, the analysis results could also be used in association with other spatial data such as site accessibility, soil and climate to aid in further planning, or as inputs to spatio-temporal modelling approaches to explore the potential spread of the infestations over time and conduct scenario analyses of the effect of possible control measures. However, the methods commonly used to assess weed distributions have potentially significant limitations when applied at the regional scale.

A key issue for regional weed assessment is the scale at which the problem is analysed. Most research into the spatial analysis of weed distributions has focused on the field scale rather than the regional scale (for example Zanin *et al.*, 1998; Dieleman & Mortensen, 1999; Colbach *et al.*, 2000; Rew *et al.*, 2001; Walter *et al.*, 2002; Ambrosio *et al.*, 2004), with study areas ranging from 190 m^2 to 1.73 km^2 (Rew & Cousens, 2001). While the analyses used by these researchers were appropriate for field scale data, the spatial complexity caused by environmental and spatial con-

Correspondence: S W Laffan, School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia. Tel: (+61) 2 9385 8093; Fax: (+61) 2 9385 1558; E-mail: shawn.laffan@unsw.edu.au

trols means that their analyses cannot be directly extended to the regional scale. A method is needed that can appropriately assess regional scale weed infestations in a robust manner and at an appropriate scale. It is this issue of the analysis method that is addressed in this paper.

There are many candidate weed species with regional distributions. The focus in this research is on *Nassella trichotoma* (serrated tussock) (Nees) Arech, for which there is a spatially extensive data set of aerial observations across the Snowy River, Cooma Monaro and Bombala Shires (local government areas) in New South Wales (NSW), Australia (Figs 1 and 2).

The objective of this study is therefore to identify the main hotspots of *N. trichotoma* infestations across the study region, although the approach is applicable to any weed species. Knowledge of these distributions will enable managers to identify and prioritize their control options, and serve as an indication of the possible source of future expansions of the infestations. The analyses are described following a consideration of the general principles of spatial analysis of weed distributions and the characteristics of *N. trichotoma*.

Spatial analysis of weed distributions

Maps of regional scale weed distributions (e.g. Fig. 2) enable broad patterns to be discerned. However, while providing very useful information about the scale and extent of weed infestations, it is difficult to visually compare such infestations across a regional study area in the consistent, repeatable and objective manner required to prioritize control efforts. Such an approach is possible using spatial statistics, for which some general considerations and approaches are now described.

The analysis of spatial data may be broadly categorized into the spatially implicit and the spatially explicit (Mackey & Laffan, 2002). Spatially implicit analyses take no account of the spatial relationships among the data, in ignorance of Tobler's (1970) First Law of Geography 'That everything is related to everything else; but that near things are more related than those far apart'. Spatially explicit analyses account for this, and quantifying the extent to which geographical phenomena fulfil Tobler's law is one objective of such analyses.

There are two main approaches for the spatially explicit analysis of spatial data. These are the spatially

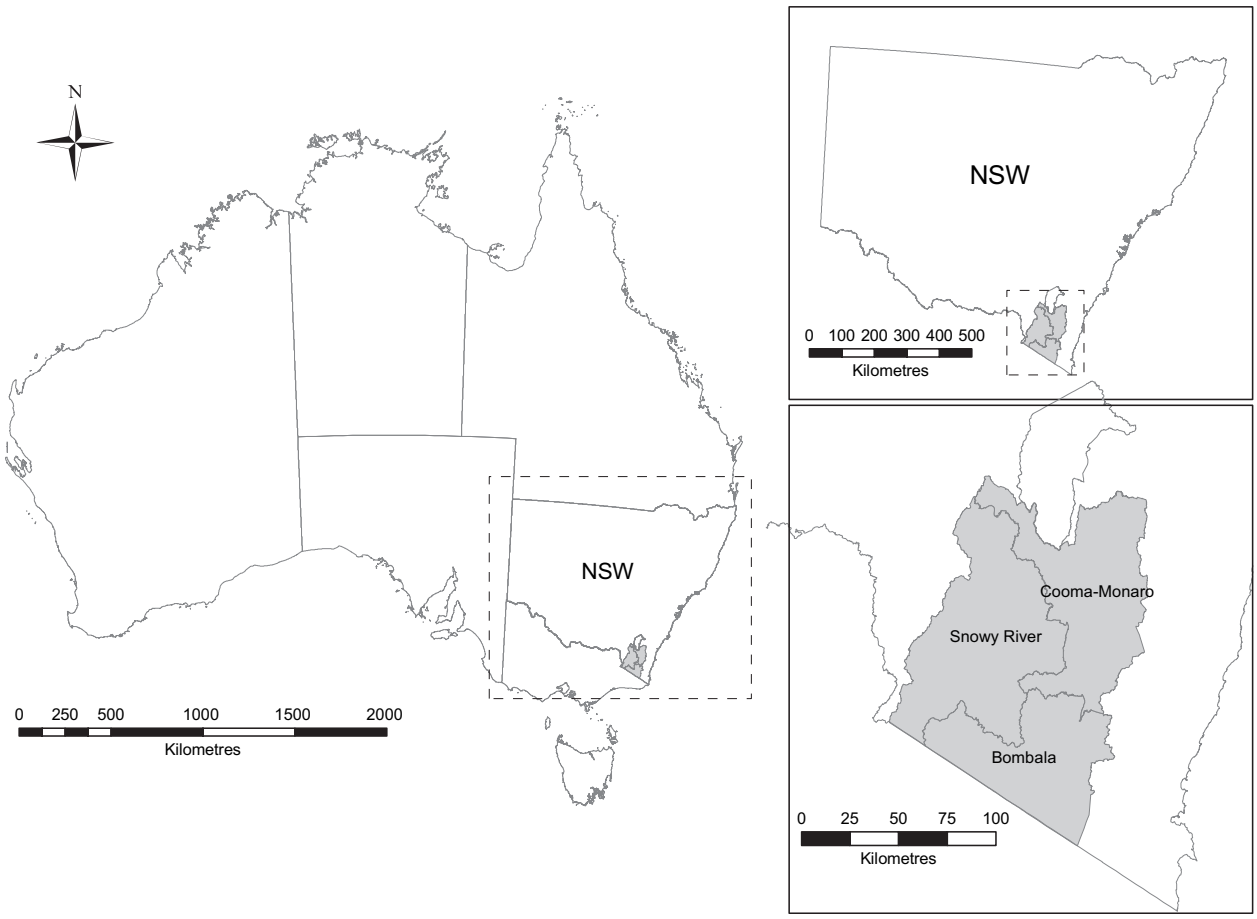

Fig. 1 The study area is located in south-east NSW, Australia.

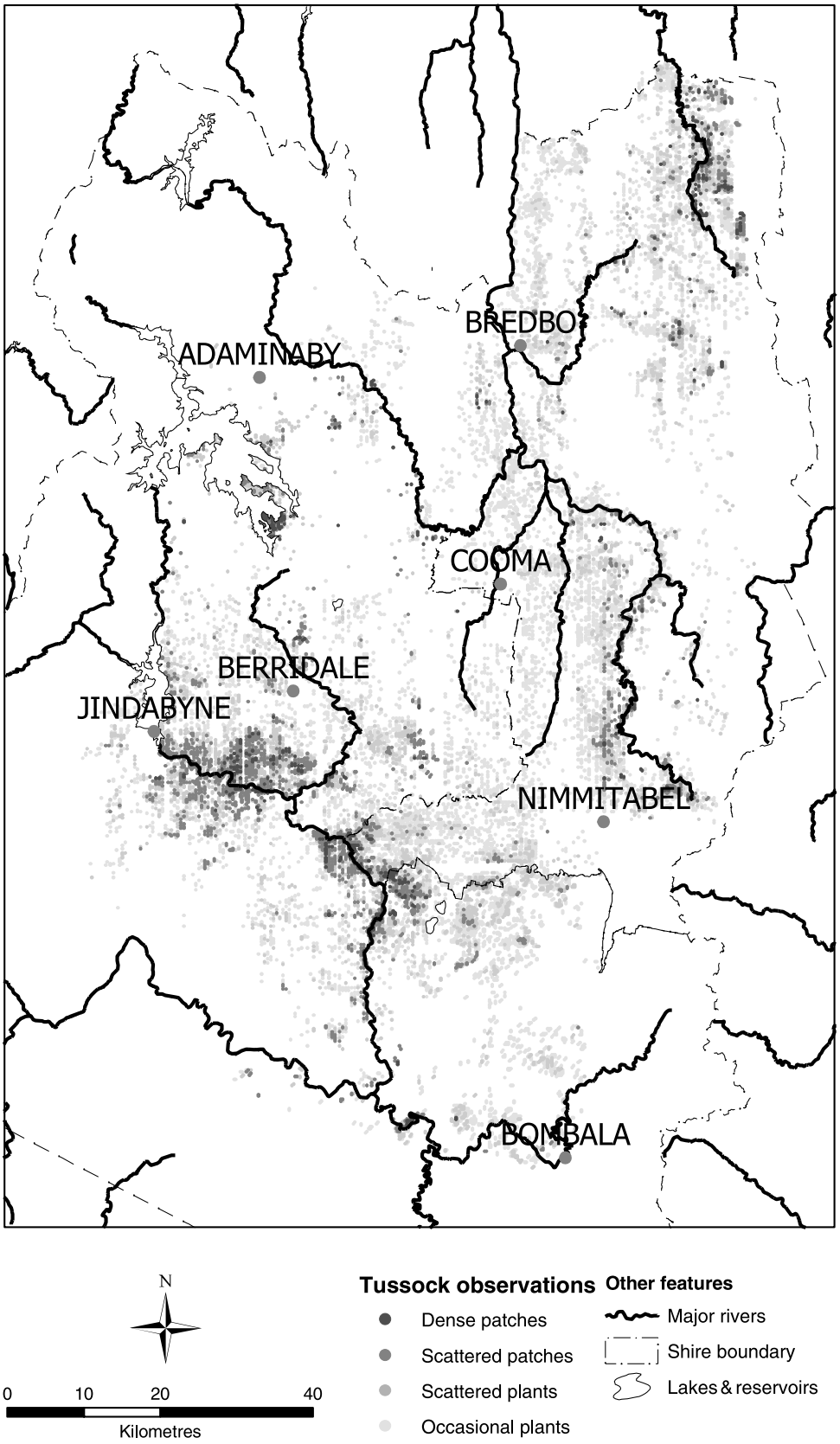

Fig. 2 The distribution of *N. trichotoma* infestations across the study area.

global and the spatially local. Both are normally implemented using a moving window to define the neighbourhood of sample locations used in the analysis, for which any window shape may be used, provided it is a reasonable representation of the process driving the distribution (Laffan, 2002). A circular shape is most often used because normally there is no certainty about which driving process applies where.

The difference between global and local approaches is in how the results are reported. Global approaches summarize the spatial relationships in the data as a single figure or graph. However, this represents the system as an average process and can conceal important local variations. It is also subject to the effects of spatial non-stationarity and inconsistent anisotropy (directional effects), particularly where distributions are not in spatial equilibrium. Geographically local analyses are often adaptations of global statistics, and result in a geographic surface of results. This means that local variations may be elucidated and that the analyses are less susceptible to the effects of spatial non-stationarity. The latter factor is particularly important when one considers that spatial non-stationarity is to be expected at regional scales. Local analyses require a more consistent density of data than do global statistics but, where such data are available, they usually produce better results than global analyses because they can cope with the effects of spatial non-stationarity (Fotheringham *et al*., 2002; Laffan & Lees, 2004; Laffan *et al*., 2005).

The semivariance is perhaps the most commonly used spatial analysis statistic (Eqn 1), and has previously been used to assess the spatial structure of weed infestations (Zanin *et al*., 1998; Rew *et al*., 2001; Walter *et al*., 2002). When applied as a global statistic, the semivariance is calculated as the average squared difference between samples separated by some distance. One can interpret the spatial structure of the phenomenon being studied when the semivariance is calculated for multiple distances and plotted as a semivariogram. An estimate of the proportion of variation that cannot be explained as spatial structure is obtained by taking the ratio of the nugget (the intercept of the variogram with the *Y*-axis) and the sill (the value at which the variogram plateaus). Of most interest for spatial distributions is the distance at which the sill begins (the range). This represents the distance to which the data are correlated and there is spatial structure. The range can also be interpreted as the average radius of spatial clusters in a data set. However, when applied as a global statistic representing an average process, the semivariance will not reveal the local spatial variation required for weed management when applied at the regional scale. It also does not provide an indication of whether the spatial relationship is of weed infestations, or if it is of the absence of weed infestations:

$$\gamma(d) = \frac{1}{2n_d}\sum_{i=1}^{n_d}\{Z(x_i) - Z(x_{i+d})\}^2 \qquad (1)$$

where d is the lag distance for which the statistic is to be calculated, n_d is the number of neighbours of location i at distance d. $Z(x_i)$ is the value at location i, $Z(x_{i+d})$ is the value at the neighbouring location at distance d.

As noted above, many local statistics are adaptations of global statistics, although this need not always be the case. Of the many possible approaches (e.g. Anselin, 1995; Whelan *et al*., 2001), the Getis–Ord G_i^* hotspot cluster statistic is of most use here (Getis & Ord, 1992, 1996; Ord & Getis, 1995). The G_i^* statistic measures the degree of spatial clustering of a local sample and how different it is from the expected value (Eqn 2). It is calculated as the sum of the differences between values in the local sample and the mean, and is standardized as a z-score with a mean of zero and a standard deviation of 1:

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j - W_i^*\bar{x}^*}{s^*\sqrt{\frac{(nS_{1i}^*) - W_i^{*2}}{n-1}}} \qquad (2)$$

where i is the centre of the local neighbourhood, d is the lag distance (bandwidth of the sample window), w_{ij} is the weight for neighbour j from location i, n is the number of samples in the data set, W_i^* is the sum of the weights, S_{1i}^* is the number of samples within d of the central location ($=W_i^*$ for a binary weights case), $\bar{x}^*$ is the mean of the whole data set, and s^* is the standard deviation of the whole data set.

The G_i^* statistic is two-tailed, so a score of -2 is as clustered as a score of $+2$. The difference is that positive values represent clusters that are, on average, greater than the mean (the expected value if there were no spatial clustering). Negative values represent clusters that are less than the mean. If values are coded such that high values represent weed infestations, then a positive G_i^* value represents a cluster of weed infestations. A negative value in this case represents a cluster of samples without weeds, or perhaps with only isolated weed samples surrounded by samples without weeds. The extent to which the G_i^* value is greater or less than the mean represents the strength of the spatial clustering in that local sample. Given it is a z-score, G_i^* values more extreme than *c*. ± 2 represent strong clustering, as 95% of the data under a normal distribution should be within 2 standard deviations of the mean. A value of 3 therefore means that the degree of clustering is more than 3 standard deviations more than what one would expect if there were no spatial clustering. Values between ± 2 may

be interpreted as weakly clustered, with values being less than 2 standard deviations away from what one would expect if there were no spatial clustering. These threshold values are, however, not exact because the correlation of spatial data violates the assumption of independence required for statistical significance. This correlation is due to Tobler's law and is, in any case, the object of interest in this study.

Nassella trichotoma

Nassella trichotoma is native to the pampas grasslands of South America. It is an important weed in Australia, New Zealand and South Africa, with smaller infestations in France, Italy and Scotland (Campbell & Vere, 1995). *Nassella trichotoma* currently occurs in the south-eastern parts of NSW and Victoria and parts of Tasmania (McLaren *et al*., 1998; Parsons & Cuthbertson, 2001), with *c*. 7000 km^2 of infested land in NSW (Parsons & Cuthbertson, 2001). In the study region it is classified as a class W2 noxious weed under the NSW Noxious Weeds Act 1993, which means it must be fully and continuously suppressed and destroyed. It is enough of a problem overall in Australia to be designated a Weed of National Significance (Agriculture & Resource Management Council of Australia & New Zealand & Australian & New Zealand Environment Conservation Council & Forestry Ministers, 2000).

Nassella trichotoma plants have a basal diameter of up to 15 cm, a height of up to 60 cm and a competition suppressing leaf spread of up to 50 cm (Healy, 1945; Campbell & Vere, 1995). *Nassella trichotoma* is tolerant to most soil types, and grows on both fertile and infertile soils (Healy, 1945). Each plant produces enormous quantities of seeds each year (more than 140 000 per plant), enough to infest a further hectare of land. Seed densities in the upper 2.5 cm of soil have been estimated at 444 000 000 ha^{-1} (Healy, 1945), and a dense infestation across 1 ha can produce 2 tonnes of seed (Parsons & Cuthbertson, 2001). *Nassella trichotoma* is a long-term problem because some portion of the seeds can remain viable in the soil seedbank for over 20 years in favourable conditions (Taylor, 1987 cited in Campbell, 1998). Further details on the physical characteristics of *N. trichotoma* are given in Healy (1945), Campbell and Vere (1995) and Parsons and Cuthbertson (2001).

The main problem posed by *N. trichotoma* is that it has such a low protein (4%) and very high fibre (86%) content (Campbell, 1982, 1998), making it unpalatable and difficult for stock to digest. Heavy infestations can reduce the carrying capacity of pastures by over 90% (Campbell, 1998; Jones & Vere, 1998), while moderate infestations cause a 40% reduction (Jones & Vere, 1998). The economic impacts of *N. trichotoma* are severe, and have been modelled in 1997 as costing the New South Wales lamb and wool industry $AU40 million per year (Jones & Vere, 1998).

The most common dispersal mechanism for *N. trichotoma* is through wind, although longer distances may be achieved through the movements of vehicles and livestock (Campbell, 1998). Some panicles have been observed blown 10 km from the source (Healy, 1945), while Jones and Vere (1998) report distances up to 20 km. While wind dispersal may normally be over short distances due to obstacles such as vegetation and fences, the large numbers of seeds produced and the cumulative dispersal over years can result in spatially extensive infestations. Water also plays a role where stream bank erosion and floods can redistribute soil containing *N. trichotoma* seeds.

Nassella trichotoma is a complex, regional scale problem for which there is no single management solution. The currently available control measures for *N. trichotoma* are preventive or reactive, depending on the location and level of infestation (Campbell, 1998). All control measures have their limitations, and the ease with which the weed can be spread by wind, the large number of seeds produced each year and the long residence times in seedbanks mean that control measures need to be conducted on a continuing basis. Preventive options include fencing and other barriers to interrupt the windborne spread of seed, or planting competitive pastures. Reactive options range from manual chipping of individual plants to spraying with herbicides such as glyphosate and flupropanate. Airborne spraying can be effective in combination with other control measures such as planting competitive pastures (Campbell, 1974). Jones *et al*. (2000) suggest that the most socially optimal control option for low rainfall and low fertility areas is to retire infested land and plant trees. Biological control options are also under investigation (Briese & Evans, 1998; Briese *et al*., 2001).

Materials and methods

The study region

The study region consists of the Snowy River, Cooma-Monaro and Bombala shires. It is *c*. 15 000 km^2 in area (Fig. 1), of which *N. trichotoma* occupies *c*. 710 km^2. A further 95 km^2 of tussock have been observed under management actions. The region has a cool, dry climate, similar to the South American pampas grasslands where *N. trichotoma* is native. Climate surfaces generated using ANUCLIM (Houlder *et al*., 2000) indicate that annual rainfall for the surveyed area varies between 495 and 1289 mm year^{-1}, with 5%, median and 95% values being 517, 607, and 872 mm year^{-1} respectively. The

minimum temperature of the coldest period varies from −4 to 0°C, with a median of −2°C. The maximum temperature of the warmest period varies from 20 to 27°C, with a median of 24°C. The main land use is grazing, although low rainfall and/or low fertility soils make it marginal in many cases. With the low economic returns on grazing comes a reduced ability for farmers to control *N. trichotoma* infestations. These conditions, along with the rugged topography in some areas of the catchment, are identified by Vere *et al.* (1993) as reducing the profitability of controlling the weed, creating a detrimental feedback relationship.

The data set

A regional scale data set of *N. trichotoma* infestations has been collected by aerial surveys over a 4-year period from 2001 to 2004 by local government and landcare agencies (Snowy River Shire, Cooma-Monaro Shire, Bredbo Landcare, Nimmitabel Landcare, Berridale Rocky Plains Landcare and Bungarby Landcare) (Fig. 2). For most of the year such surveying is difficult, as *N. trichotoma* resembles native tussock grasses (Campbell, 1982). To allow for this, aerial surveys were conducted at the end of winter when plants turn a distinctive golden yellow colour and during flowering periods when abundant purple seed heads are evident (Parsons & Cuthbertson, 2001).

Surveys were conducted by helicopter, with flight lines aligned approximately north–south and spaced *c.* 400 m apart. The helicopter altitude was *c.* 20–25 m above the ground surface (J. Clarke, pers. comm.). Forward speed was *c.* 30 m s^{-1}, consistent with Hyde-Wyatt (1979). An observer in the helicopter called out observations of *N. trichotoma* for areas on the ground nominally 200–250 m on a side. Calls were registered by a second operator using a computer with mapping software interfaced with a differential global positioning system (GPS). Calls used a four-class ordinal scale of densities similar to that described in Jones and Vere (1998) (Table 1), using representative aerial photographs as references for each class. The mapping software stored GPS way-points at regular intervals (this varied between surveys), and so an absence of *N. trichotoma* (null data) is implied by a GPS co-ordinate with no call made. In some cases the null data were available as vectors of the helicopter flight lines. Other weeds, evidence of *N. trichotoma* control measures, and rabbit burrow densities were also recorded during the survey, but were coded as null values for these analyses. All data were stored and analysed using the Map Grid of Australia, Zone 55, co-ordinate system.

There are three main sources of spatial error in the surveys. First, there is a time delay between *N. trichotoma* being observed and the infestation code being entered into the computer. This results in a spatial offset along the helicopter flight line and varies between observers and surveys. Second, the GPS and the mapping software were sometimes out of synchronization, usually resulting in groups of two to five observations (and up to 20) being assigned to the one GPS co-ordinate. In these cases the first observation code assigned is used. Third, an observer can sometimes continue to call 'runs' of infestation codes after the density of plants has changed, for example where scattered patches (C2) are observed for long periods and it takes a short time to realize that the infestation density has changed to dense patches (C1). This is most common where infestation densities are near to the class boundaries.

The spatial errors in the data set preclude its use for detailed local planning. However, the highly dispersive nature of *N. trichotoma* means that management is normally conducted at the property scale, for which very high accuracy is not needed. On-ground observations of *N. trichotoma* infestations indicate the data are accurate to this level of detail (T. Fletcher, pers. comm.). In addition, the use of the G_i* index in a spatial window means it acts as a spatial smoother, reducing the effect of short range spatial errors. Following from this, any *N. trichotoma* observations that do not fall within a positive G_i* cluster might be errors in the observations that can be verified using field surveys.

Despite its errors, the aerial survey method is very cost effective in comparison with on-ground sampling across regional scales (Rew & Cousens, 2001; Rew *et al.*,

Table 1 Classification system used for aerial observations of serrated tussock densities

Density class	Percentage ground cover	Description	Area of observations (km^2)	Frequency of observations (cells)
C1	80–100	Dense patches	25	626
C2	50–80	Scattered patches	98	2440
C3	10–50	Scattered plants	179	4482
C4	<10	Occasional plants	410	10 253
Null		None observed		

Ground cover percentages are approximate and areas are based on 200 m by 200 m cells reported to the nearest km^2.

2001). The sampling approach is also advantageous in that human observers are able to identify isolated *N. trichotoma* plants. This is not possible using multispectral remote sensing systems because of spatial and spectral resolutions and associated mixed pixels (Lamb & Weedon, 1998; Lamb *et al.*, 1999; McGowen *et al.*, 2001). Such limitations might be overcome using hyperspectral remote sensing data, and any weed surfaces derived from hyperspectral images could be used as the input data for the approach described in this paper. However, costs are currently high for hyperspectral data with sufficiently fine spatial resolution across regional scales (see Lass *et al.*, 2005).

Data pre-processing

The density classes used for the aerial surveys are arbitrary and have limited ecological meaning. An analysis that pays strict attention to these classes will propagate observer errors that occur when infestation densities are near to the class boundaries. To alleviate this effect the data were recoded into a series of indicator classes, where locations that exceed a threshold are given a value of 1 and all other locations are assigned zero. This does not remove all of the classification uncertainties, but does limit them to only one class boundary in each analysis.

The indicator classes used here are: (1) the presence of *N. trichotoma* at any density (where C1, C2, C3 and C4 are assigned a value of 1, Null is assigned zero); (2) patch level infestations (C1 and C2 are assigned a value of 1, C3, C4 and Null are assigned zero); (3) infestations consisting of scattered plants, where C4 infestations are assigned a value of 1 and all other values a zero (C1, C2, C3 and Null). The three indicator classifications allow an analysis of (1) the presence of *N. trichotoma*, (2) patch level (possibly core) infestations, and (3) occasional plants which may develop into patches but are currently most easily controlled.

The continuous timed recording of GPS co-ordinates by the mapping software means that there is an overwhelmingly greater number of null data observations than there are *N. trichotoma* observations. This is not realistic, as these null values do not represent observations of an absence of *N. trichotoma*. They also commonly occur within the on-ground observation area for *N. trichotoma* records, and would distort the results if retained as is. These null values must therefore be resampled so they are spatially comparable with the tussock observations.

To make the null records spatially comparable with the *N. trichotoma* observations, the data set was aggregated to a lattice with mesh points spaced 200 m apart to represent the smaller range of the nominal on-ground observation area. These were then converted to a raster data set such that any cell containing an observation of *N. trichotoma* was assigned that code. Cells containing more than one *N. trichotoma* observation were assigned the highest infestation code in the cell to allow for a worst-case scenario. Cells containing only null records were retained as a single null value.

The null data are essential for the G_i^* statistic, as it will otherwise return a biased representation of the degree of spatial clustering. There are several areas of the Cooma-Monaro surveys where no null data are available along the flight lines, and so artificial null data were generated. These were based on random values using the Normal() function in ArcInfo GRID to generate pseudo-random values using a Gaussian distribution. The Normal() function uses a pseudo-random number generator with known flaws (Van Neil & Laffan, 2003), but is adequate for this purpose because the random values were thresholded to be included or excluded. Different threshold values were assessed until the distribution approximated that of the null data from the remainder of the data set, where the distributions were for the number of observations or thresholded values within a five cell radius of observed *N. trichotoma* samples. This is not an ideal way to generate the null data, but is more reliable than visually estimating flight lines from the *N. trichotoma* observations.

Spatial windows

The use of a local moving window gives local spatial statistics a marked advantage over global statistics. They remain, however, sensitive to the weights in the sample window. A binary sample window, where all sample observations are given equal weighting in the calculations, results in circular artefacts in the G_i^* surfaces. The relative rarity of *N. trichotoma* observations in comparison with the null observations makes the arithmetic mean of the whole data set close to zero, so even a single observation will have an effect on the G_i^* statistic, normally resulting in a weak positive cluster (see Eqn 2). These artefacts are normally caused by an *N. trichotoma* observation occurring near the edge of the sample window. They are most pronounced when the observations are spatially isolated, but can also artificially extend the boundaries of a cluster. While these artefacts are normally of low intensity, they are a distraction in interpretation.

The effect of observations near the edge of the sample window can be reduced using a weighting scheme. For these analyses, the weights were assigned using a bisquare function so that the contribution of sample locations to the statistic decreases (decays) with distance from the centre of the sample window (Eqn 3; Fig. 3).

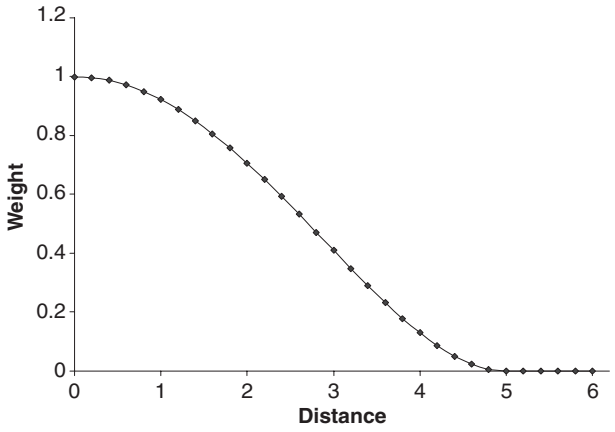

Fig. 3 The bisquare weighting function decreases from a weight of 1 at distance zero to a weight of zero at the selected bandwidth (in this case 5).

In simple terms, values closer to the centre of the analysis window have a greater contribution to the result than do values from further away. The bisquare function has a Gaussian form and decays to zero at the edge of the sample window (the bandwidth). Isolated locations will still be identified, but their G_i* value will be greatly reduced in intensity when they are near the edge of the sample window:

$$w_{ij} = \begin{cases} (1 - (d/b)^2)^2; & d < b \\ 0; & d \geq b \end{cases} \quad (3)$$

where w_{ij} is the weight, d is the lag distance between central location i and sample j, and b is the bandwidth.

Implementation

The G_i* statistic was applied to the three indicator classifications using the bisquare weighting function within a circular window. To gain an understanding of how the clustering changes with changing spatial scale, the bandwidth of the sample window was varied from 1 to 6 km (5–30 cells) using a 1 km increment. These results were then aggregated on a cell-by-cell basis to generate surfaces representing the maximum positive value across all analysis scales, and the scale at which the maximum value occurred. Negative clusters, which indicate clusters of no infestation or isolated observations, were assigned a value of −1 for display and subsequent analysis purposes.

The fact that the seeds are easily dispersed, produced in large numbers and can have long residence times in seedbanks means that the most manageable *N. trichotoma* cases are the occasional plant infestations. Dense core areas will require large amounts of effort over many years before the weed is brought under control using currently available control measures. Consequently, the G_i* results were used to identify spatial distributions as possible core infestations, possible less established infestations, and occasional plants that may be manageable.

To identify the three distribution types, any infested cell with a maximum positive value exceeding 2 (G_i* > 2) was considered as part of a strong cluster of infestations. Fringe areas with positive values where $0 < G_i* < 2$ were not considered as part of these strong clusters. Any infested cell that had a negative G_i* response was considered as spatially isolated. These should be more easily managed than those areas in strong positive clusters. As noted above, these isolated cases may also be errors in the observations that can be verified using on-ground surveys. Any observation that occurred within a weakly positive but spatially distinct cluster ($0 < G_i* < 2$ and not part of the fringe or halo of a strong positive cluster) may be less established and therefore manageable, albeit not as easily as an observation with a negative G_i* response. These may also be a primary source of any future spatial expansion of the weed distribution (Moody & Mack, 1988).

Semivariograms were also calculated to provide a spatially global comparison for the G_i* results. These used a lag size of 200 m and were calculated to a distance of 40 km. This distance was used to assess if any global structures were present beyond the sample radius used for the G_i* analyses, and also because the smallest axis of the sample data is 80 km across. Samples beyond 40 km will be biased because the data pairs from which the semivariance values are calculated will be largely derived from the extremes of the data set.

Results

The semivariogram results (Fig. 4) are difficult to interpret in exact terms. There are no distinct sills, and there are clear fluctuations at *c*. 25 km for the presence and isolated plants analyses. The patch level variogram indicates a slight decline after *c*. 18 km. Some very general interpretations are that the spatial distance to which the data are related is *c*. 15 km for presence of *N. trichotoma*, *c*. 10–15 km for patch scale infestations, and *c*. 5 km for isolated cases. Comparison with the G_i* results indicates that the clusters vary widely around these values (Fig. 5). The ratio of the nugget to the sill is also greater than 50% for each of the infestation levels analysed (presence = 62%, patch level = 52%, isolated = 81%), indicating that a large proportion of the variance cannot be explained as global spatial structure.

The G_i* results indicate that large areas of the study region fall within *N. trichotoma* clusters with a G_i* value of greater than 2, and that much of this area is represented by a small number of large clusters (Fig. 5,

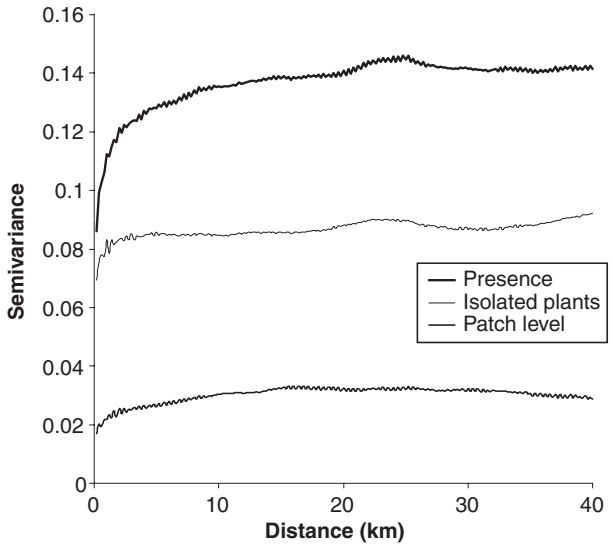

Fig. 4 Semivariograms for the three analysis classifications. The legend order is the order in which they plot.

Table 2). Approximately 14 448 cells (578 km^2) of the total 17 801 cells (710 km^2) with *N. trichotoma* observations occur in strong positive presence clusters where $G_i^* > 2$. Similarly, 2918 cells (117 km^2) of the 3066 patch level cells (123 km^2), and 8215 cells (329 km^2) of the 10 253 occasional plant level cells (410 km^2), occur in strong positive clusters.

There are 1366 cells (55 km^2) of *N. trichotoma* observations with negative G_i^* responses in the presence analysis. There are also 468 cells (19 km^2) of *N. trichotoma* observations in the occasional plant analysis that have a negative G_i^* response, albeit 363 of these (14.5 km^2) also occur within the presence clusters where the G_i^* score exceeds 2. The remainder are scattered plant infestations (code C3), as none of the patch level observations have a negative G_i^* response for the patch level analysis (Fig. 6).

There are also many spatially distinct, weakly positive clusters ($G_i^* < 2$), with 155 presence, 79 patch level, and 140 occasional plant clusters. These tend to be small, with the median size 0.28 km^2 for the presence clusters, 1 km^2 for the patch level clusters, and 0.48 km^2 for the occasional plant clusters. These patches contain, respectively, 280 cells (11.2 km^2) of observations used in the presence analysis, 74 cells (3 km^2) of observations for the patch level analysis, and 155 cells (6.2 km^2) of observations for the occasional plant analysis.

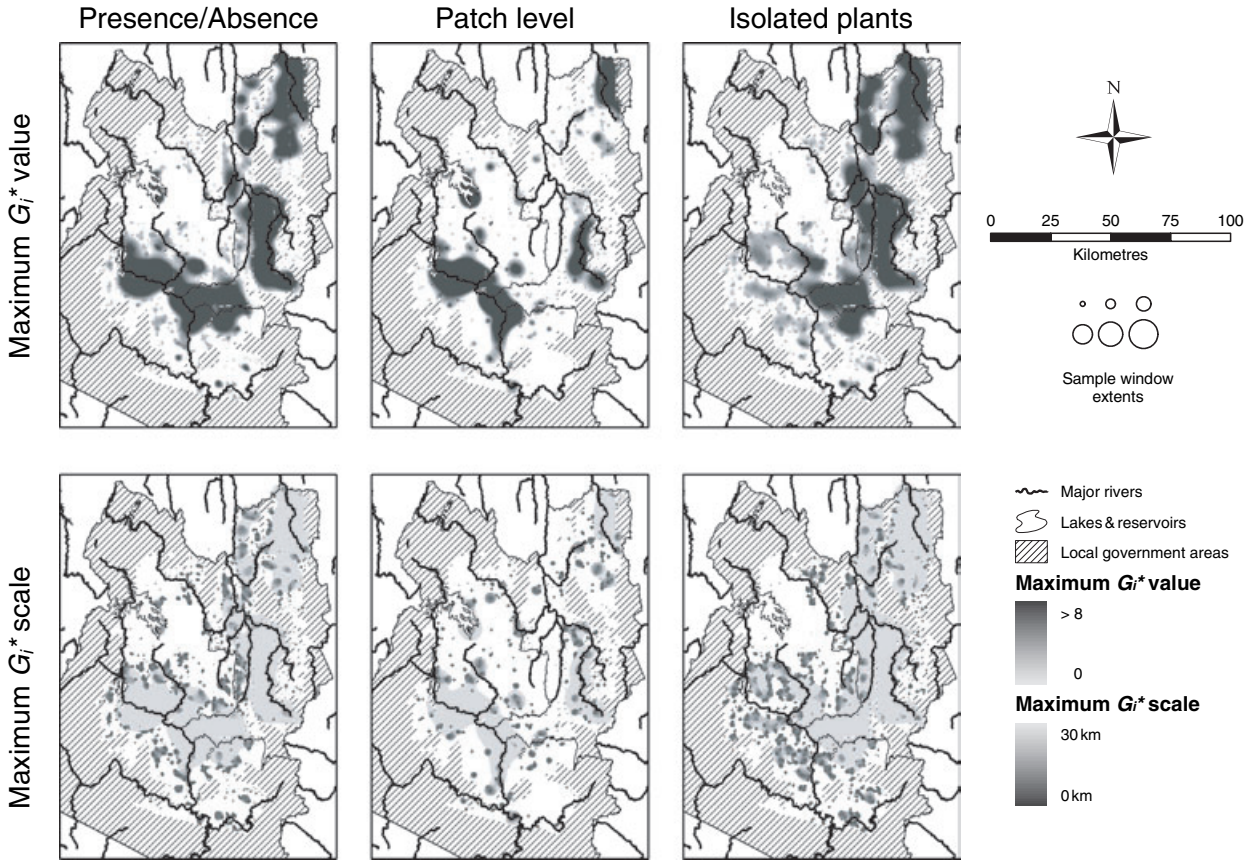

Fig. 5 The spatial distribution of positive clusters of *N. trichotoma* presence, patches and occasional plants. Cells which are always negative clusters are displayed as white. The images show the value of the maximum positive clustering and the distance at which that value was obtained. Sample window extents are shown on the right.

Table 2 Results aggregated across all bandwidths, including the spatial extent and frequency of clusters with a G_i^* score > 2 (excluding fringe areas where $0 < G_i^* < 2$), the numbers and areas of observations in three cluster types: strong positive, negative and spatially distinct weak positive (not a fringe component of a strong positive cluster), and the number and median size of spatially distinct, weakly positive clusters

Infestation level	Total area of clusters >2 km^2	Number of strong positive clusters	Number of strong positive clusters >5 km^2	Number of strong positive clusters >10 km^2	Observations in strong positive clusters ($G_i^* > 2$)	Observations where $G_i^* \leq 0$	Observations in spatially distinct weakly positive clusters ($0 < G_i^* < 2$)	Number of spatially distinct, weakly positive clusters (median size in brackets)
Presence (C1, C2, C3 & C4)	2942	123	17	8	578 km^2 (14 448 cells)	55 km^2 (1366 cells)	11.2 km^2 (280 cells)	155 (0.28 km^2)
Patch level (C1 & C2)	1624	74	18	11	117 km^2 (2918 cells)	0 km^2 (0 cells)	3 km^2 (74 cells)	79 (1 km^2)
Occasional plants (C4)	2941	145	14	10	329 km^2 (8215 cells)	19 km^2 (468 cells)	6.2 km^2 (155 cells)	140 (0.48 km^2)

Areas are based on a 200 m by 200 m cell size. Note that both the patch level and the occasional plant hotspots extensively overlap with the presence hotspots (see Fig. 5).

Fig. 6 The distribution of cells with each infestation category within each cluster type and for each analysis level. (A) Strongly positive clusters ($G_i^* > 2$), (B) negative clusters ($G_i^* < 0$) and (C) weak positive clusters ($0 < G_i^* < 2$ and spatially distinct).

The breakdown of infestation density classes within the cluster types (Fig. 6) reflects the overall distributions, with frequencies generally decreasing as the infestation level increases. The exceptions to note are the absence of any patch level infestation cells in the negative patch level clusters, and the higher number of scattered plants in the negative occasional plant clusters.

Discussion

A comparison of the G_i^* results with those of the semivariograms indicates how much information is suppressed by global spatial statistics when they are extended to the regional scale. There is a wide spread of the size of the local clusters around the variogram

ranges, with a spatially variable anisotropy where the clusters are aligned with the prevailing winds as they flow along the valleys. It should, however, be noted that previous research on the spatial structure of weeds have been applied across small spatial extents, comparable with the smaller analysis window sizes used in this research (e.g. Zanin *et al*., 1998; Dieleman & Mortensen, 1999; Colbach *et al*., 2000; Rew *et al*., 2001; Walter *et al*., 2002; Ambrosio *et al*., 2004). The issue at hand is the extension of such global statistics to regional scales, where one expects spatial non-stationarity, anisotropy and the absence of an equilibrium state between current weed distributions and the landscape. Spatial statistics applied at the local scale across regions, such as G_i^*, provide a much better understanding of the spatial distribution of the weeds.

Spatially, the *N. trichotoma* presence clusters occur in three main groups located in the Snowy River and Cooma-Monaro shires, one of which crosses into the north-west corner of Bombala (Fig. 5). The infestations in the surveyed part of Bombala shire are smaller and generally of lower intensity than in the Cooma-Monaro or Snowy River shires. This possibly reflects the distance from source areas, but is expected to change if the upwind infestations are not managed or if seed is brought into the area.

The patch level observations occur almost exclusively within strong positive clusters (97.5%, see Fig. 6), indicating that development to patch level corresponds with the development of core areas of *N. trichotoma* infestation. The patch level hotspots most clearly follow the major valleys along which the dominant westerly winds are funnelled. This is particularly evident along the Snowy River valley below Jindabyne (the south-western clusters in Fig. 5). The largest patch level clusters are in the Snowy River shire, followed by Cooma-Monaro and then Bombala. These clusters are often in nature reserves along difficult to access valleys,

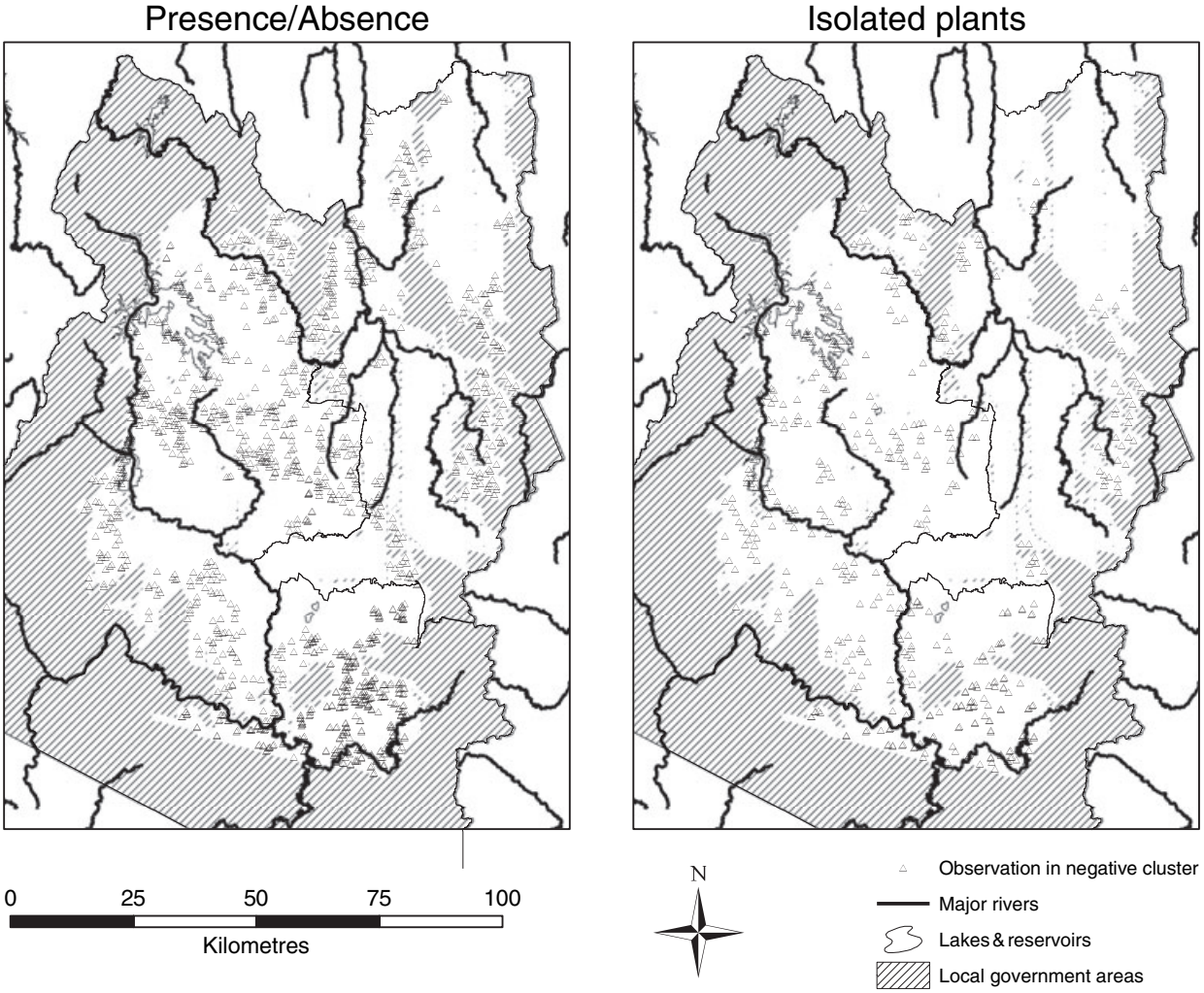

Fig. 7 The spatial distribution of *N. trichotoma* observations that occur within negative clusters. The total analysis extent is superimposed over the local government areas in white; 78% of the occasional plant observations fall within the presence clusters (363 of 468).

compounding the effects of topographically controlled wind dispersal.

The distribution of the occasional plant clusters across the entire study area also indicates that large areas could readily become patch level infestations if left unmanaged. The breakdown of infestation type for the occasional plant analysis (Fig. 6) indicates that occasional plants are very often isolated from the main infestations, which is where the denser infestation classes occur. One should also note that the infestations in Cooma-Monaro are largely occasional plants (C4 infestations), while those in the Snowy River shire are of the patch level. This is possibly because of observer bias between surveys, but this is unlikely to occur across such a large area because of the clear visual difference between scattered and occasional plants and groups of plants within a patch. It is more likely due to the less rugged terrain in the Cooma-Monaro shire allowing wind dispersal across a larger area from a large number of isolated infestations, with the resulting infestations yet to reach patch level (J. Clarke, pers. comm.).

In terms of management options, if one considers the strong positive *N. trichotoma* clusters as core areas, then these locations would be almost impossible to eradicate with current control measures, and efforts focused on containment may be the only practical option. Any eradication efforts can then be focused on the *N. trichotoma* observations that fall outside of the strong positive clusters, possibly beginning with those in the less rugged terrain. This remains a complex task, as there are large numbers of observations that do not occur within strong positive clusters. However, the results indicate that these are most commonly occasional or scattered plants, requiring significantly less control effort per case than the patch level infestations. These are also the infestations that should be focused on, as future expansion of weed populations can be greater from such cases (Moody & Mack, 1988). In addition, 363 cells (14.5 km^2) of the occasional plant (C4) observations with a negative G_i* response occur within presence clusters (Fig. 7), leaving 105 cells (4.2 km^2) that do not occur within a positive cluster. As noted earlier, some of these could actually be erroneous observations that do not require any control action.

Conclusions

Regional scale weed distributions are readily assessed using local scale, explicitly spatial analysis techniques. Local spatial statistics allows management approaches to be prioritized in a consistent, objective and repeatable way. The main clusters of *N. trichotoma* infestations across a regional scale study area in Australia have been identified, as have potential source areas for future patch level infestations. The identification of these clusters allows an understanding of where resources might be applied to most effectively control weed infestations and also where further investigations may be focused. This was extremely difficult to achieve using visual assessment of the original survey data in its raw form, and also using a spatially global analysis method. In addition, the effectiveness of any control actions could be assessed by comparing the local analysis results for this and subsequent surveys, as well as monitoring the spatio-temporal dynamics of the infestations. Only one weed species has been considered in this research, but the analysis approach is readily extended to other weed species at regional scales where reliable data are available.

Acknowledgements

Survey data and advice were kindly provided by Jo Clarke, John Porter, Tim Fletcher, Charlie Litchfield and Phil Zylstra. Vector data sets used in the maps are copyright Geoscience Australia.

References

Agriculture & Resource Management Council of Australia & New Zealand & Australian & New Zealand Environment Conservation Council & Forestry Ministers (2000) *Weeds of National Significance: Serrated Tussock (Nassella trichotoma)* Strategic Plan. National Weeds Strategy Executive Committee, Launceston, http://www.weeds.org.au.

Ambrosio L, Iglesias L, Marin C & del Monte JP (2004) Evaluation of sampling methods and assessment of the sample size to estimate the weed seedbank in soil, taking into account spatial variability. *Weed Research* **44,** 224–236.

Anselin L (1995) Local indicators of spatial association – LISA. *Geographical Analysis* **27,** 93–115.

Briese DT & Evans HC (1998) Biological control of serrated tussock (*Nassella trichotoma*): is it worth pursuing? *Plant Protection Quarterly* **13,** 94–97.

Briese DT, Pettit W & Anderson F (2001) *Biological Control of Serrated Tussock and Chilean Needle Grass*. Report for the Rural Industries Research and Development Corporation, Australia.

Campbell MH (1974) Efficiency of aerial spraying for long term control of serrated tussock. *Australian Journal of Experimental Agriculture and Animal Husbandry* **14,** 405–411.

Campbell MH (1982) The biology of Australian weeds 9. *Nassella trichotoma* (Nees) Arech. *Journal of the Australian Institute of Agricultural Sciences* **48,** 76–84.

Campbell MH (1998) Biological and ecological impact of serrated tussock (*Nassella trichotoma* (Nees) Arech.) on pastures in Australia. *Plant Protection Quarterly* **13,** 80–86.

Campbell MH & Vere DT (1995) *Nassella trichotoma* (Nees) Arech. In: *The Biology of Australian Weeds* (eds RH Groves, RCH Shepherd & RG Richardson), Vol. 1, 189–202. Richardson, Frankston, Victoria.

Colbach N, Dessaint F & Forcella F (2000) Evaluating field-scale sampling methods for the estimation of mean plant densities of weeds. *Weed Research* **40,** 411–430.

Dieleman JA & Mortensen DA (1999) Characterizing the spatial pattern of *Abutilon theophrasti* seedling patches. *Weed Research* **39,** 455–467.

Fotheringham AS, Brunsdon C & Charlton M (2002) *Geographically Weighted Regression, the Analysis of Spatially Varying Relationships*. Wiley, New York.

Getis A & Ord JK (1992) The analysis of spatial association by use of distance statistics. *Geographical Analysis* **24,** 189–206.

Getis A & Ord JK (1996) Local spatial statistics: an overview. In: *Spatial Analysis: Modelling in a GIS Environment* (eds P Longley & M Batty), 261–277. GeoInformation International, Cambridge.

Healy AJ (1945) Nassella tussock. Field studies and their agricultural significance. *DSIR New Zealand Bulletin* **91**.

Houlder D, Hutchinson M, Nix H & McMahon J (2000) anuclim User's Guide. Centre for Resource and Environmental Studies, The Australian National University, Canberra.

Hyde-Wyatt BH (1979) *Senecio jacobaea* location from the air. In: *Proceedings of the 7th Asian Pacific Weed Science Society Conference*, Tokyo, 45–46.

Jones RE & Vere DT (1998) The economics of serrated tussock in New South Wales. *Plant Protection Quarterly* **13,** 70–75.

Jones RE, Vere DT & Campbell MH (2000) The external costs of pasture weed spread: an economic assessment of serrated tussock control. *Agricultural Economics* **22,** 91–103.

Laffan SW (2002) Using process models to improve spatial analysis. *International Journal of Geographical Information Science* **16,** 245–257.

Laffan SW & Lees BG (2004) Predicting regolith properties using environmental correlation: a comparison of spatially global and spatially local approaches. *Geoderma* **120,** 241–258.

Laffan SW, Nielsen O, Silcock H & Hegland M (2005) Sparse grids: a new predictive modelling tool for the analysis of geographic data. *International Journal of Geographical Information Science* **19,** 267–292.

Lamb DW & Weedon MM (1998) Evaluating the accuracy of mapping weeds in fallow fields using airborne digital imaging: *Panicum effusum* in oilseed rape stubble. *Weed Research* **38,** 443–451.

Lamb DW, Weedon MM & Rew LJ (1999) Evaluating the accuracy of mapping weeds in seedling crops using airborne digital imaging: *Avena* spp. in seedling triticale. *Weed Research* **39,** 481–492.

Lass LW, Prather TS, Glenn NF, Weber KT, Mundt JT & Pettingill J (2005) A review of remote sensing of invasive weeds and example of the early detection of spotted knapweed (*Centaurea maculosa*) and babysbreath (*Gypsophila paniculata*) with a hyperspectral sensor. *Weed Science* **53,** 242–251.

Mackey BG & Laffan SW (2002) Case studies in GIS and environmental modeling. In: *Geographic Information Systems and Environmental Modeling* (eds KC Clarke, BE Parks & MP Crane), 223–251. Prentice Hall, New Jersey.

McGowen IJ, Frazier P & Orchard P (2001) Mapping serrated tussock and Scotch thistle in temperate pastures using Landsat TM imagery. In: *Proceedings of the IEEE 2001 International Geoscience and Remote Sensing Symposium*, University of NSW, Sydney, Australia.

McLaren DA, Stajsic V & Gardener MR (1998) The distribution and impact of South/North American stipoid grasses (Poaceae: Stipeae) in Australia. *Plant Protection Quarterly* **13,** 62–69.

Moody ME & Mack RN (1988) Controlling the spread of plant invasions: the importance of nascent foci. *Journal of Applied Ecology* **25,** 1009–1021.

Ord JK & Getis A (1995) Local spatial autocorrelation statistics: distributional issues and an application. *Geographical Analysis* **27,** 286–306.

Parsons WT & Cuthbertson EG (2001) *Noxious Weeds of Australia*, 2nd edn. CSIRO Publishing, Collingwood, Australia.

Rew LJ & Cousens RD (2001) Spatial distribution of weeds in arable crops: are current sampling and analytical methods appropriate? *Weed Research* **41,** 1–18.

Rew LJ, Whelan B & McBratney AB (2001) Does kriging predict weed distributions accurately enough for site-specific weed control? *Weed Research* **41,** 245–263.

Taylor NJ (1987) *Ecological Aspects of Nassella tussock*. Botany Division, DSIR, Lincoln, New Zealand.

Tobler WR (1970) A computer movie simulating urban growth in the Detroit region. *Economic Geography* **46**(Suppl.)**,** 234–240.

Van Neil K & Laffan SW (2003) Gambling with randomness: the use of pseudo-random number generators in GIS. *International Journal of Geographical Information Science* **17,** 49–68.

Vere DT, Auld BA & Campbell MH (1993) Economic assessments of serrated tussock (*Nassella trichotoma*) as a pasture weed. *Weed Technology* **7,** 776–782.

Walter AM, Christensen S & Simmelsgaard SE (2002) Spatial correlation between weed species densities and soil properties. *Weed Research* **42,** 26–38.

Whelan BM, McBratney AB & Minasny B (2001) Vesper – spatial prediction software for precision agriculture. In: *ECPA 2001, Proceedings of the 3rd European Conference on Precision Agriculture* (eds G Grenier & S Blackmore), 139–144, Montpellier, France.

Zanin G, Berti A & Riello L (1998) Incorporation of weed spatial variability into the weed control decision-making process. *Weed Research* **38,** 107–118.

