

# ○ THE SPATIAL ANALYSIS OF SPECTRAL DATA

## EXTRACTING THE NEGLECTED DATA

*Brian Lees, School of Physical, Environmental and Mathematical Sciences, UNSW@ADFA, Canberra ACT 2600 Australia*

*Correspondence to Brian Lees: b.lees@adfa.edu.au*

Remotely sensed data are a key input to GIS-based spatial decision support systems for land cover and land use application areas. One of the major sources of error in the input of processed remotely sensed data to GIS is in the process of classification. Particularly important is the degradation of the data from the interval to nominal level of measurement. This is less significant in cultural landscapes where boundaries predominate, but it becomes an important source of error in natural, and disturbed natural, environments where gradients exist. Use of the  $G_i^*$  local statistic as an alternative approach to processing remotely sensed data proved very successful, replicating the level of discrimination achieved by conventional classification and field labelling in a much shorter time, whilst avoiding the errors associated with conversion of the data from the interval to nominal level of measurement.

## INTRODUCTION

Spatial decision making based on a combination of image-based and vector data is an important activity in many application domains. A strong theme of research within this has been to improve the quality of attribute information available from remotely sensed image data by combining the remotely sensed image data with ‘ancillary’ environmental and terrain data. Because of the data distribution information required for the diverse data sets, which is not always available, and the set-up time involved, parametric classifiers have become less favoured than the non-parametric alternatives in this sort of synthesis (Benediktsson et al, 1990).

Numerous studies of classification using decision trees, neural networks of various kinds, genetic algorithms and even cellular automata have been published to examine these alternatives (Lees, 1996a). In some disturbed natural environments the errors in prediction remain stubbornly high. Examination and tracking of the errors in prediction in these studies made it clear that the process of classification itself was the major source of error (Lees, 1996b). Classification achieves two things. Firstly, it simplifies image data into a form suitable for incorporation into entity-based cartography products but, in the process, degrades the data from the interval to the nominal level of measurement. In disturbed natural and natural environments there are both gradients and boundaries. Classification can represent the real boundaries faithfully but, because of the transformation of data values from the interval to the nominal level of measurement, inserts categorical boundaries where a gradient is present.

Gradients and boundaries are features in geographic space, and, with only few exceptions, most classification procedures are non-spatial analyses of spatial data. The spatial content of the image data usually remains un-investigated which tends to make it very difficult to replicate the boundary and gradient patterns present in the real world. This is unfortunate as image data, mostly digital remotely sensed data in this context, is a rich source of relatively high quality spatial information and only yields moderate to low quality attribute information. Very few studies have examined the explicitly spatial analysis of this sort of spatial data. Strategies such

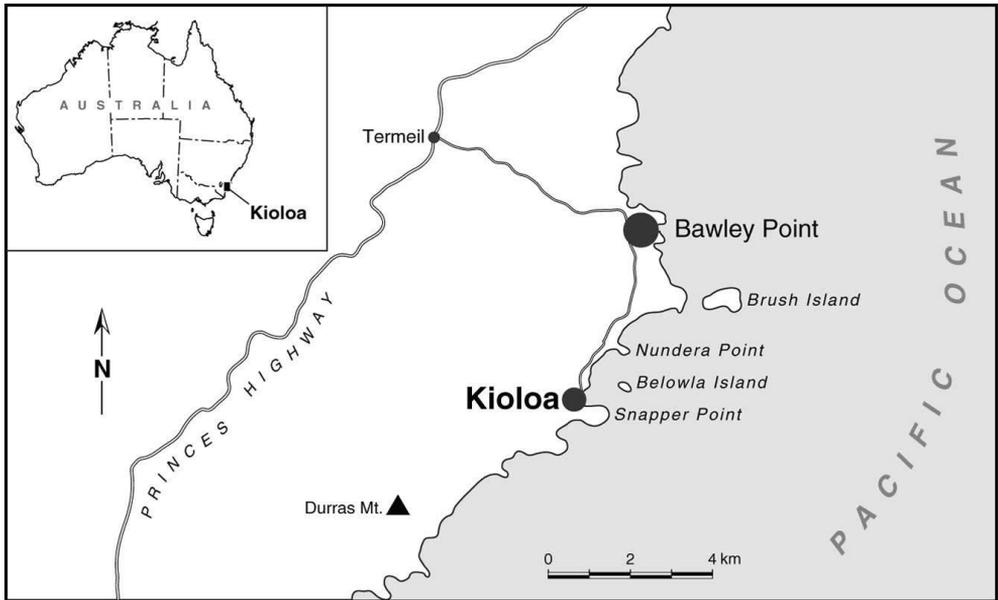
as spatial signal detection, local spatial autocorrelation and local regression are able to extract high quality spatial information from image data that can significantly enhance its use in spatial decision making.

## PREVIOUS WORK

Spatial signal detection enables ‘ancillary’ vector data to be used to remove known spatial patterns from image data prior to classification (Lees and Hafner, 2000). In the Liverpool Plains study a conventional classification of remotely sensed data of a region of intensive agriculture subject to minor salinisation gave a confusing result. In order to isolate that part of the signal which gave information about the spatial patterning and degree of salinisation, the other, known, patterns needed to be removed. To do this, a vector layer depicting the field pattern was created from the image data. Field polygons were then allocated the major class in each polygon, with the assumption that this was what the farmers intended. Because of the multiple crops being grown in this highly productive area, reflectance values across field boundaries were standardised to one crop type. As this was not a linear relationship, but rather a very fuzzy one, a neural network was used to optimise this. Once the cultural pattern of agriculture was removed the image could easily be interpreted as showing the distribution and intensity of low-level salinisation effects on the standardised crop. Like the other classification techniques, this was aimed at producing better attribute information. The explicit use of the spatial information contained in the image to facilitate the analysis of the image was very effective.

Using a similar technique, Burnett and Blaschke (2003) suggest that there is a scale hierarchy in landscapes which allows multi-scale segmentation, but considerable field examination of sites over which we have carried out experiments show that not all are that nicely organised. Our main test area is around Kioloa, on the N.S.W. South Coast, Australia (figure 1). Part of the problem is that the Kioloa region is a disturbed natural, rather than cultural, landscape and cannot be decomposed into objects or patches without creating boundaries where gradients exist. The problem is similar to that addressed by Pearson (2002) in northern Australia.

In environments where cultural patterning is less strong, or even non-existent, the patterning caused by changes in variables such as soil type, geology, topography, and soil wetness can be dealt with by image segmentation using multi-source data (Benediktsson et al, 1990; Lees and Ritman, 1991). Although in those papers this language is not used, effectively the procedure is similar to that used by Lees and Hafner (2000) in the Liverpool Plains. The difference between them is that the latter creates an intermediate vector layer, to deal with patches and fields as entities, and the former works directly at pixel level. A very similar procedure could be used to standardise the effects of variables such as soil type, geology, topography, and soil wetness leaving an image showing the distribution of, say, disturbance. Conversely, as Lees and Ritman (1991) intended when they added remotely sensed data to the analysis of environmental variables carried out by Moore et al. (1991), the analysis could use the remotely sensed data to provide disturbance information. It would be easy to see this approach as essentially non-spatial, but both the topographic derivatives and hydrological derivatives are the results of spatial modelling. Similarly, at broader scales, and for larger areas, climate variables used in similar analyses are derived by spatial modelling. The spatial information inherent in the data is being used implicitly.



**Figure 1** The location of the study area.

Local spatial statistics, on the other hand, are explicitly spatial and produce spatial information from the image data of a sort we have not yet become very used to in this application area. Zones of similarity and dissimilarity, patches and gradients are identified. In conjunction with the sorts of analysis we are used to, these additional sources of information greatly strengthen our ability to make better decisions.

There have been two strands of research in remote sensing which are relevant. In the first, much of the interest has been in the response of materials to incident radiation of various types. In the second, a much more macroscopic approach has been taken. In the forests that we have been dealing with it became very apparent, very early on, that the reflectance alone of the leaves of *Eucalyptus* sp., was a very poor way to discriminate between them (Lees and Ritman, 1991). However the second approach, outlined by Jupp and Walker (Jupp et al., 1986; Walker et al., 1986) showed that a great deal of structural information could be derived from the mixture of light and shade within a pixel.

This latter approach informed later work on the spatial context and scale relationships in raster data for thematic mapping (Fitzgerald and Lees, 1995). In an attempt to improve the classification of the forests around Kioloa we used a neural network to incorporate spatial context. The results showed that the incorporation of spatial context does indeed improve classification accuracy. However, the scale of the spatial context must be matched to the theme, or vegetation community. Particular levels of spatial context optimised different types of land cover. For the sea, the level of classification accuracy obtained with no spatial context (1x1 30m pixel) was as high as the accuracy obtained with spatial contexts up to a hundred metres radius (7x7 30m pixels). For the grasslands, classification accuracy was maximum with a spatial context of 3x3 30m pixels. The closed forests were classified best with a spatial context of 5x5 30m pixels, and

the more open forests with spatial context radius of 9x9 30m pixels. For the whole area, the best results were obtained with a spatial context radius of 3x3 30m pixels.

These results make absolute sense when one considers the crown-to-crown spacing between trees, and other vegetation, in each of these communities. It is important to bear in mind that, although the forest cover is discussed as though it formed discrete types, this forest is a continuum of change in structure, and in species. As a result, although the classification accuracy was increased significantly, the use of a classification procedure still degraded the data from the interval to the nominal level of measurement and meant that the errors induced by this process still persisted.

## **EXPLORATORY SPATIAL DATA ANALYSIS**

The results obtained by Fitzgerald and Lees (1995) were interesting, but our conclusion was that the time involved in setting up a spatially-aware neural network was probably not cost-effective. However, the development of local, rather than global, statistical techniques for detecting variations in spatial autocorrelation offer a new way to achieve the same objective with much less effort.

Exploratory spatial data analysis (ESDA), of which spatial autocorrelation measures are an important subset, examine an observed distribution in space and attempt to make inferences about the process that generated it. The main objective is to highlight patterns in the data that are considered by the investigator to be significant.

A major impetus in the development and application of ESDA in the last few decades is linked to advances in computer processing power, the availability of large well structured spatial data sets, and increased dissatisfaction with the simple spatial manipulations of geographic data. In particular, a significant increase in spatial theory (Unwin, 1996) and the recasting of spatial statistical tools from global to local operation (Rogerson and Fotheringham, 1994; Fotheringham and Charlton, 1994; Bailey, 1994; Zhang and Griffith, 2000) has driven this.

The truth of Tobler's first law of geography (1970) which states that "everything is related to everything else, but near things are more related than distant things", can be estimated using spatial autocorrelation. Spatial autocorrelation is concerned with a comparison of similarity among the attributes and similarity in location of spatial objects or features. A positive spatial autocorrelation occurs when similar attribute values are adjacent, while negative spatial autocorrelation occurs when neighbouring values are more dissimilar than those that are further apart. Zero spatial autocorrelation occurs when attributes are spatially independent (Goodchild, 1986).

Remote sensing systems structure the landscape into a regular grid of pixels, independent of the land cover characteristics. There is thus often a high degree of spatial autocorrelation between pixels. While this may be seen as a source of error when using statistical methods that assume independence, it is a fundamental image characteristic that can be exploited. Wulder and Boots (1998) provide a useful summary of studies that use various aspects of image spatial autocorrelation to advantage.

## **LOCAL MEASURES OF SPATIAL AUTOCORRELATION**

Traditionally, spatial dependency has been explored using global measures such as joint-counts, Moran's I, and Geary's C (Hubert et al., 1981). Such techniques assume that the relationships

being estimated are stationary over space and that a single parameter measure is implicitly taken to depict the nature of such relationships for all points within the area of study (Fotheringham and Brunson, 1999). Techniques which generate only a sole average measure of spatial dependency tend to conceal, or suppress, any significant local variation of spatial nonstationarity in the nature and extent of spatial autocorrelation within a data set. Such a global estimate will be quite uninformative and may be very misleading. It is a little surprising, given their lack of utility, that they generated any interest other than the purely academic. The assumption of spatial homogeneity over large areas of the Earth's surface is extremely unrealistic. More subtly, the assumption of spatial homogeneity over large areas, at a constant scale, is also unrealistic. The results obtained by Fitzgerald and Lees (1995) demonstrate that, in practise, spatial autocorrelation is scale sensitive.

In response to this problem, scientists employing quantitative spatial data analysis have become increasingly involved in the development of local statistical techniques that focus primarily on highlighting and understanding differences across space, rather than similarities, as is the case with global measures (Fotheringham, 1997). More specifically, emphasis is on the dissection of global statistics into local constituents, concentration on local exceptions instead of global regularities and production of local mappable statistics, rather than 'whole-map' values.

Local indicators of spatial association (LISA) have been developed to overcome the above limitations (Anselin, 1995) and have the potential to reveal discreet spatial regimes that might be missed using existing global techniques. These measures focus on local variations within patterns of spatial dependence and evaluate the nature and extent of concentration in the values of the variable  $x$  in a local region, within an area of interest. LISA are defined as any statistic where each observation gives an indication of the extent of significant spatial clustering of similar values around and observation, and where the sum of all observations must be proportional to a global indicator of spatial association. This allows the global coefficient of spatial autocorrelation to be deconstructed, making it possible to identify the individual locations contributing to the global autocorrelation.

Moran's ( $I_i$ ) and Geary's ( $c_i$ ) coefficients are two LISA developed by Anselin (1995) and are versions of the established global univariate statistical techniques (Moran I and Geary C). Details of the statistical theory and formulae used to derive these are given in Anselin (1995), Sokal et al., (1998a, 1998b) and Tiefelsdorf and Boots (1997). Both compare values of a central location, minus the data set mean with neighbourhood locations, minus the data set mean. In this, they resemble the texture variable sometimes calculated in remote sensing analysis. Moran's ( $I_i$ ) measures joined covariation of neighbouring localities. A positive spatial autocorrelation is where these deviate strongly from the mean and have the same sign (+ve or -ve). A negative spatial autocorrelation is where the value of the central location deviates widely from the mean and has a sign opposite to its neighbours. Geary's ( $c_i$ ) measures squared differences between the values of neighbouring points and the central location. High values indicate a negative spatial autocorrelation. When values between the central location and its neighbours are all close to the mean, the data exhibits positive spatial autocorrelation. This would correspond to a weak, or a zero level, spatial autocorrelation using Moran's ( $I_i$ ) and indicates that these two statistics provide separate information about spatial structure. The Moran's ( $I_i$ ) has been successfully used by Anselin (1995) to analyse the spatial distribution of conflict in Africa, while Sokal et al., (1998a)

used it to identify variations in the spatial distribution and population structure of cactophilic fruit flies in eastern Australia.

## GETIS-ORD STATISTIC

Getis and Ord (1992) and Ord and Getis (1995) have formulated two local statistical measures,  $G_i$  and  $G_i^*$ , that evaluate the degree of spatial clustering in high (hot spot) and low (cold spot) values of the study variable,  $x$ , about a central location or cell,  $i$ . These are locations where the value clusters are consistently greater than, or less than, the mean.

In  $G_i$  the value of  $x$  at  $i$  is excluded from the local sum. In  $G_i^*$  it is included and is considered to be more appropriate for this problem since it allows a window of user-defined dimensions to be employed in the statistical calculation. Calculation of the  $G_i$  by excluding the central cell of  $G_i^*$  results in an analysis that is analogous to the neighbourhood component of Morans I (Laffan, 2002). There has been little to no guidance on which of these statistics to use in any given situation but, as we wish to examine the effect of different window sizes on spatial autocorrelation, the  $G_i^*$  statistic looks to be the most useful. The Getis-Ord statistics are not strictly local measures of spatial autocorrelation, but more specifically of spatial association. Sokal et al., (1998b) argue that, as neither satisfy the second requirement of Anselin's (1995) definition of LISA, where the sum of all observations must be proportional to a global indicator of spatial association, these are not properly LISA. Nevertheless, the Getis-Ord statistics have some attractive features, which are discussed below.

The implementation of  $G_i^*$  is as a  $z$ -score, calculated as the difference between the sum of the local sample and the weighted global mean, divided by the weighted global standard deviation.  $G_i^*$  is positive when a cluster is greater than the mean and negative when a cluster is less than the mean. More extreme values denote greater clustering of similar values and values of zero are not clustered (Laffan, 2002). It expresses the sum of the weighted variate values within a specified distance of a particular observation  $I$ , as the proportion of the sum of the variate values for the entire study area (Wulder and Boots, 1998). This can be compared with the statistic's expected value under a hypothesis of no local spatial autocorrelation to indicate if the degree of clustering of  $x$  values in the vicinity of  $i$  is greater than or less than chance would dictate (Getis, 1994).

In formal terms, the Getis-Ord statistic  $G_i^*(d)$  for some distance  $d$  is defined as:

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j}{\sum_j x_j} \quad (1)$$

Where:  $\{ w_{ij}(d) \}$  is the spatial weights matrix consisting of symmetric binary weights with 1s assigned to allocations within distance  $d$  of observation  $i$ , including  $i$  itself (i.e.  $w_{ii} = 1$ ), and 0 otherwise. In a remote sensing context, a window can be defined around an observation  $i$  by specifying an appropriate value of  $d$ .

Steps to derive a standard version of  $G_i^*$  have been provided by Ord and Getis (1995). Wulder and Boots (1998) summarise these in the context of application to remotely sensed imagery as follows:

Calculating the mean ( $\bar{x}$ ) and variance ( $s^2$ ) of  $x$  for the entire image:

$$\bar{x} = \frac{\sum_j x_j}{n} \quad \text{and} \quad s^2 = \frac{\sum_j x_j^2}{n} - \bar{x}^2 \quad (2)$$

The expected value of  $G_i^*$  is:

$$E(G_i^*) = W_i^* / n \quad (3)$$

With a variance of:

$$\text{Var}(G_i^*) = \frac{W_i^*(n - W_i^*)}{n^2(n - 1)} \left[ \frac{s}{x} \right]^2 \quad (4)$$

This results in the following  $z$ -score standardised form:

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j - W_i^* x}{s[W_i^*(n - W_i^*)/(n - 1)]^{1/2}} \quad (5)$$

Where:  $W_i^* = \sum_j w_{ij}(d)$

In terms of remotely sensed imagery,  $G_i^*$  values measure the extent to which a pixel is surrounded by a cluster of high or low reflectance values. Large positive  $G_i^*$  values denote a cluster of bright pixels, whilst large negative  $G_i^*$  values denote a cluster of darker pixels. This additional information makes the  $G_i^*$  useful when dealing with remotely sensed data as it combines both the spatial association information with some spectral information. This makes its information content greater than the texture variable which is only an index of local variance.

Where local statistical measures have been used in environmental studies, most have used the Moran or Geary versions (e.g. Pearson, 2002). Relatively few have used the Getis-Ord method. Wulder and Boots (1998) summarise the use of the Getis-Ord method with remotely sensed data. More recently, Holden et al., (2000) used this method to evaluate coral ecosystem health by identifying spatial autocorrelation patterns in multi-temporal SPOT images. They postulated that zones of spatial homogeneity characterised dead algae-dominated reef, while more heterogeneous areas were more likely to reflect healthy conditions. Derksen et al., (1998) used spatial dependence information, generated using the Getis-Ord method, to link patterns of prairie snow cover and atmospheric circulation phenomena.

## STUDY AREA

A 250 km square area around Kioloa, on the New South Wales south coast (figure 1), was used to compare the results obtained using the Getis-Ord method with other techniques. The forests around Kioloa present a difficult classification problem. The terrain runs northwest-southeast, the worst possible alignment for illumination using satellite remote sensing. The forests have been burnt in patches, logged in patches, and are extremely diverse.

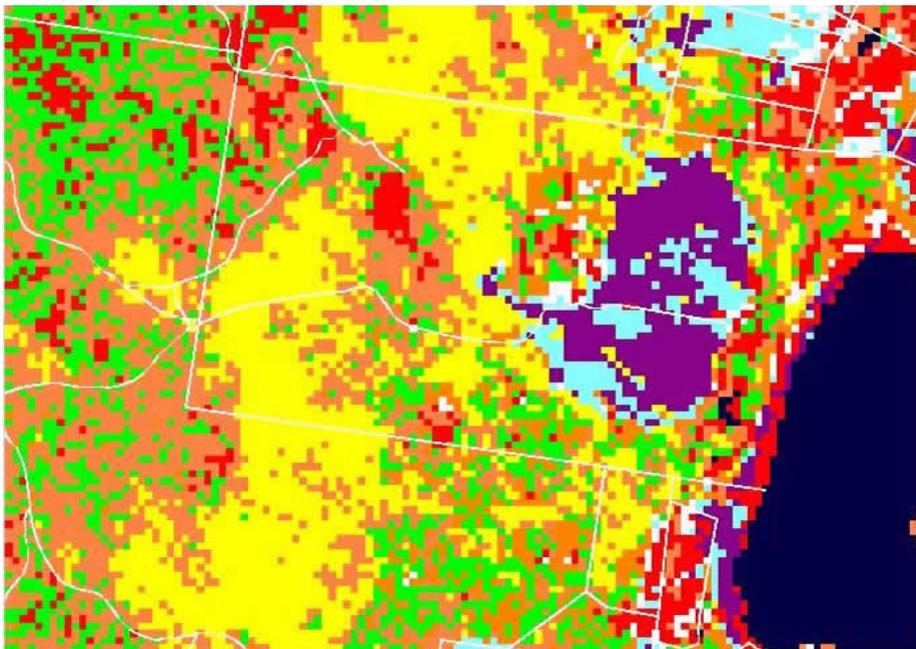
The complex interrelationships between geology, soils, landforms, local climatic factors and flora recorded in the area are reflected in the highly variable nature and distribution of vegetation types (Lees and Ritman, 1991). A range of regional ecosystems are represented including rainforest, both open and closed wet and dry sclerophyll forest, woodlands, heath, swamp and wetland communities and as well as cleared grasslands. Unlike cultural landscapes where disturbance is usually contained within cultural boundaries, here disturbance crosses boundaries in haphazard way and the successive patterns of recovery from disturbance are often themselves disturbed.

This makes approaches like those described by Lees and Hafner (2000), where cultural patterns were removed to facilitate analysis of regional soil geochemistry, difficult to implement successfully.

This study evaluates one particular type of ESDA tool, the Getis-Ord statistic. The focus of the task is the mapping of the forested terrains around Kioloa, where a great deal of previous work of this type has been done. A number of previous studies in this area have tested decision trees, spatially and temporally-aware neural networks, Dempster Shafer, and hybrid systems in the integration of remotely sensed and environmental data sets. The aim of all of these studies was to determine the most efficient and effective way of producing land cover data as input to spatial decision support systems.

## METHODOLOGY

The sole data set used in this study was a Landsat TM image. No ancillary data was used. In order to provide a comparison, an unsupervised classification using bands 2,4, and 7 was generated. The 56 spectral classes produced in this way were linked into 11 land cover types on the basis of field checking (figure 2). This was a very conventional procedure and, with the field work, took about two and half days.



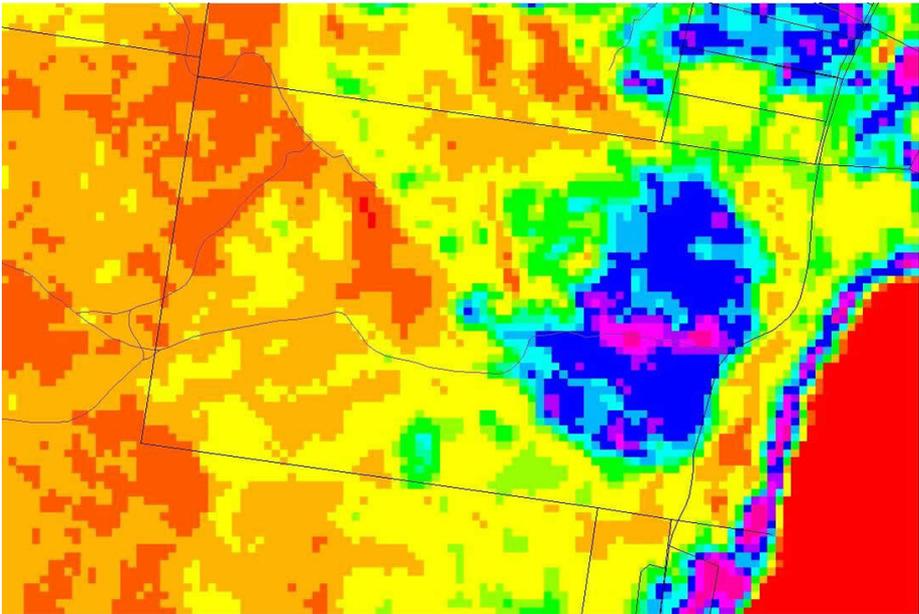
**Figure 2** Classified image using Landsat TM bands 2, 4, and 7. The classifier was ER-Mappers ISOCCLASS algorithm and 56 original classes were linked into 8 cover classes following field checking. The data is categorical and the within forest boundaries are mostly artefacts of the classification process.

For the Getis-Ord analysis, TM band 5 data was selected because of its high-level of vegetation discrimination. The Getis-Ord computation was carried out using the G.\* function in ARC INFO. The remote sensing data was not resampled to either a finer or coarser resolution. The raw 30m

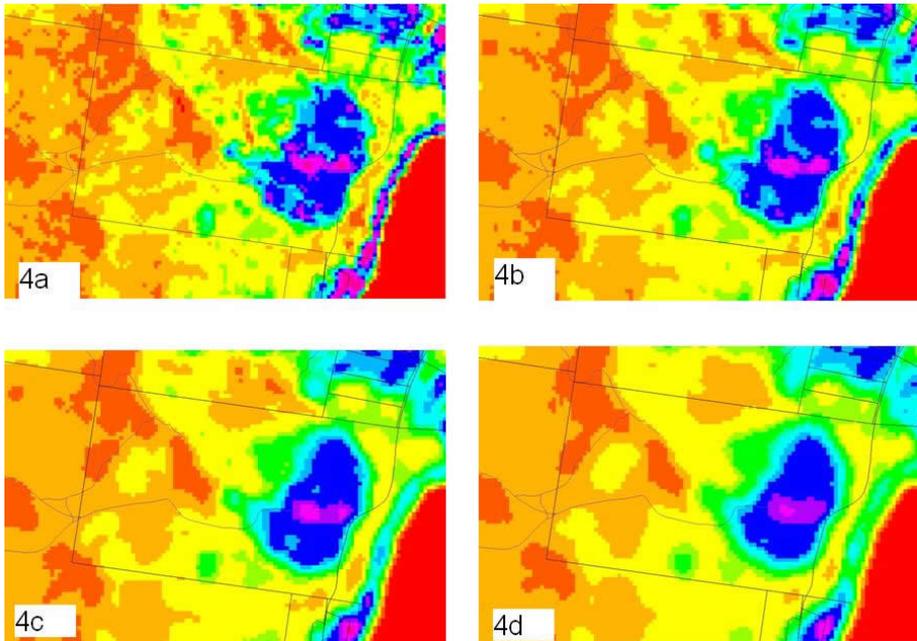
pixel size was left unchanged. The analysis was run with a number of spatial radii and the results compared to the final land cover mapping produced by conventional classification of remotely sensed data, and the results of previous exercises using spatially-aware neural networks.

## RESULTS

Changing the spatial scale of the analysis had an important impact on the results. Like Pearson's (2002) findings where, although the most effective cell size was small (0.2m), the most effective radius was a relatively large 30m, it was found that a similar radius (30m) bore the closest relationship to the observed vegetation patterns. A  $G_i^*$  with a 30m radius (approximating the 3x3 Fitzgerald and Lees (1995) window) gave the best results in terms of level of detail and correspondence with the preferred level of discrimination. A subset of this is shown in figure 3. The effect of applying different  $G_i^*$  cell radii values to image data is summarised in figure 4. It can be seen that significant smoothing of the data and merging of clusters occurs with increasing cell radius and indicates that the smallest radius produces the best results in terms of spatial detail.



**Figure 3**  $G_i^*$  output using a radius of 45m on Landsat TM band 5 data. The data is interval and the apparent boundaries within the forest are the result of density slicing for display purposes only. The overall appearance is similar to that of figure 2, but the utility of the data for input to GIS is very much greater. Where boundaries do exist, such as along the coast, these have been represented as such by the analysis.



**Figure 4** The smoothing effect of increasing the window size is readily apparent in these images.  
**Figure 4a:**  $G_i^*$  output using a radius of 45m on Landsat TM band 5 data  
**Figure 4b:**  $G_i^*$  output using a radius of 75m on Landsat TM band 5 data  
**Figure 4c:**  $G_i^*$  output using a radius of 110m on Landsat TM band 5 data  
**Figure 4d:**  $G_i^*$  output using a radius of 130m on Landsat TM band 5 data

A comparison between the  $G_i^*$  output (figure 3) and the thematic map produced by classification and field labelling (figure 2) indicates that the local statistical method effectively highlights areas of higher reflectance values corresponding to cleared land, buildings, areas of heath-related native grassland and beach sand deposits. More interestingly, relatively high spatial autocorrelation within these areas points to a considerable degree of spatial heterogeneity, which is clearly not represented in the vegetation classification or in the image band data. The analysis has picked up differences in quality of the grassland which clearly exist, but which are suppressed in the transition from spectral classes to land cover types. Within the forest, the levels of heterogeneity identified by Fitzgerald and Lees (1995) have been discriminated very effectively. The very disturbed areas of forest show essentially no spatial autocorrelation. The more open forests show a small amount of spatial autocorrelation and the closed forests show a fairly high degree of spatial autocorrelation. Similarly, the better quality grasslands show a high degree of spatial autocorrelation, as does the beach.

## DISCUSSION AND CONCLUSION

Given that management maps of land cover are more closely related to structural communities, rather than species, this simple analysis has proven to be very powerful. The examination of the results suggests that it can be used “as is”, although it could also be used as an ancillary data source in a more elaborate secondary analysis. In this context, it is important to recognise that

the output data is interval and the import of such a data set into a subsequent analysis avoids the errors resulting from the degradation of the data from the interval to the nominal level of measurement.

This simple local spatial statistic has achieved comparable results to some of the more complex classifications of remotely sensed data we have attempted in the past. Importantly, the result takes the form of field, rather than entity, data. Where real boundaries exist, they have been captured but where gradients exist, they have been retained. Although the best results were obtained with a radius of only 30 m, this technique has discriminated some forest types which the spatially-aware neural network and analysis suggested needed 120 m radius. Whilst it is clear that the analysis has not produced information unobtainable through other means, the information on the local spatial characteristics of the forest were obtained rapidly and at low-cost. The best comparison is with the results following combined unsupervised classification and field checking. The field checking, which links spectral classes into land cover types, effectively produces a map of structural communities. This takes several days of work. Running  $G_i^*$  over a single band takes minutes to achieve an almost comparable result.

One problem of this approach is that we do not yet understand how to field-check this sort of analysis. The comparison with other forms of analysis is informative in a qualitative sense, but a method for quantitative assessment still escapes us. This is an area for further research. It is clear that the local spatial statistic has not discriminated the rainforest class from other closed forest types. The rainforest is composed of distinctly different species to the other closed forests in this region and one of the constant problems thrown up here has been to discriminate these one from another. To date, the most effective strategies have been the integration of remotely sensed and environmental data in decision tree and neural network classifiers (Lees and Ritman, 1991; Fitzgerald and Lees, 1992). Neither data type on its own has allowed a useful level of discrimination, no matter which modelling techniques have been used. Environmental data on its own overpredicts the distribution of rainforest in these highly disturbed environments. Remotely sensed data, on the other hand, cannot discriminate it effectively because of the high spectral variance of the rainforest canopy, confusion with water, and confusion with shadow.

As in the Liverpool Plains study, spatial analysis of remotely sensed data has generated useful information, very simply. The Getis-Ord statistic, with its added information about brightness values, has shown itself to be particularly useful in dealing with remotely sensed data. The output is readily interpretable and easily incorporated into more complex analysis as a useful spatial variable. Importantly, it has not been necessary to invoke a very complex, multi-scale segmentation/object relationship modelling methodology of the sort described by Burnett and Blaschke (2003) to analyse the complexity of this landscape at a useful level. The final advantage has been the avoidance of errors associated with the transfer from the interval to nominal level of measurement.

## ACKNOWLEDGEMENTS

Shawn Laffan wrote the  $G_i^*$  AML used in this study.

## REFERENCES

- Anselin, L. 1995. 'Local Indicators of spatial association – LISA'. *Geographical Analysis* 27 (2): 93–115.
- Bailey, T.C. 1994. 'A review of statistical spatial analysis in geographical information systems'. In *Spatial Analysis and GIS*, edited by Fotheringham, A.S.; Rogerson, P. London: Taylor and Francis. 13–44.
- Benediktsson, J. A.; Swain, P. H.; Ersoy, O. K. 1990. 'Neural Network Approaches Versus Statistical Methods in Classification of Multisource Remote Sensing Data'. *IEEE Transactions on Geoscience and Remote Sensing* 28 (4): 540–551.
- Burnett, C; Blaschke, T. 2003. 'A multi-scale segmentation/object relationship modelling methodology for landscape analysis'. *Ecological Modelling* 168: 233–249.
- Derksen, C; Wulder, M; LeDrew, E; Goodison, B. 1998. 'Associations between spatially autocorrelated patterns of SSM/I-derived prairie snow cover and atmospheric circulation'. *Hydrological Processes* 12: 2307–2316.
- Fitzgerald, R.W.; Lees, B.G. 1992. 'The application of Neural Networks to the floristic classification of remote sensing and GIS data in complex terrain (II)'. *Proceedings of the 17th ISPRS*, Washington DC, USA. American Society for Photogrammetry and Remote Sensing. Bethesda. Vol XXIX, Part B7; 570–573.
- Fitzgerald R.W.; Lees, B.G. 1995. 'Spatial context and scale relationships in raster data for thematic mapping in natural systems'. In *Advances In GIS Research*, edited by Waugh, T. C.; Healey, R.G. Southampton: Taylor and Francis. 462–475.
- Fotheringham, A.S. 1997. 'Trends in quantitative methods I: stressing the local'. *Progress in Human Geography* 21 (1): 88–96.
- Fotheringham, A.S; Brunson, C. 1999. 'Local forms of spatial analysis'. *Geographical Analysis* 31 (4): 340–358.
- Fotheringham, A.S; Charlton, M.E. 1994. 'GIS and exploratory spatial data analysis: an overview of some research issues'. *Geographical Systems* 1: 315–327.
- Getis, A. 1994. 'Spatial dependence and heterogeneity and proximal databases'. In *Spatial analysis and GIS*, edited by Fotheringham, A. S.; Rogerson, P. London: Taylor and Francis. 105–120.
- Getis, A; Ord, J.K. 1992. 'The analysis of spatial association by use of distance statistics'. *Geographical Analysis* 24 (3): 188–205.
- Goodchild, M.F. 1986. *Spatial Autocorrelation*. CATMOG 47. Norwich: Geobooks.
- Holden, H.; Derksen, C.; LeDrew, E. 2000. 'Coral reef ecosystem change detection based on spatial autocorrelation of multispectral satellite data'. *GIS Development*. Accessed 26 May 2003. Available from: <http://www.gisdevelopment.net/aars/acrs/2000/ts3/cost006pf.htm>
- Hubert, L.J; Gollidge, R.G; Constanzo, C.M. 1981. 'Generalised procedures for evaluating spatial autocorrelation'. *Geographical Analysis* 13 (3): 224–233.
- Jupp, D.L.B; Walker, J; Penridge, L.K. 1986. 'Interpretation of vegetation structure in Landsat MSS imagery: A case study in disturbed semi-arid Eucalypt woodlands; Part 2, Model-based analysis'. *Journal of Environmental Management* 23: 35–57.
- Laffan, S.W. 2002. 'Using process models to improve spatial analysis'. *International Journal of Geographical Information Science* 16 (3): 245–257.
- Lees, B.G. 1996a. 'Inductive modelling in the spatial domain'. In *Neural Network Applications in the Geosciences* edited by Lees, B.G. Special Volume, *Computers and Geosciences* 22: 9, 955–957.
- Lees, B.G. 1996b. 'Improving the spatial extension of point data by changing the data model.' In *Proceedings of the Third International Conference on Integrating GIS and Environmental Modeling*, edited by Goodchild, M. et al. Santa Fe, New Mexico: National Centre for Geographic Information and Analysis, WWW; CD.
- Lees, B.G; Ritman, K. 1991. 'A decision tree and rule induction approach to the integration of remotely sensed and GIS data in the mapping of vegetation in disturbed or hilly environments'. *Environmental Management* 15: 823–831.

- Lees, B.G.; Hafner, S. 2000. 'The separation of natural from cultural spatial patterns in an area of intensive agriculture using GIS and artificial intelligence'. In *GIS and Geocomputation*, edited by Atkinson, P.; Martin, D. London: Taylor & Francis. 188–205.
- Moore, D.M.; Lees, B.G.; Davey, S. 1991. 'A new method for predicting vegetation distributions using Decision Tree Analyses in a Geographic Information System'. *Environmental Management* 15: 59–71.
- Ord, J.K.; Getis, A. 1995. 'Local spatial autocorrelation statistics: distributional issues and an application'. *Geographical Analysis* 27 (4): 286–305.
- Pearson, D.M. 2002. 'The application of local measures of spatial autocorrelation for describing pattern in north Australian landscapes'. *Journal of Environmental Management* 64: 85–95.
- Rogerson, P.A.; Fotheringham, A.S. 1994. 'GIS and spatial analysis: Introduction and overview'. In *Spatial Analysis and GIS*, edited by Fotheringham, A.S.; Rogerson, P.A. London: Taylor and Francis. 1–10.
- Sokal, R.R.; Oden, N.L.; Thompson, B. 1998a. 'Local spatial autocorrelation in a biological model'. *Geographical Analysis* 30 (4): 331–353.
- Sokal, R.R.; Oden, N.L.; Thompson, B. 1998b. 'Local spatial autocorrelation in biological variables'. *Biological Journal of the Linnean Society* 65: 41–62.
- Tiefelsdorf, M.; Boots, B. 1997. 'A note on the extremities of local Moran's I<sub>j</sub>s and their impact on global Moran's I\*'. *Geographical Analysis* 29 (3): 248–257.
- Tobler, W. 1970. 'A computer movie simulating urban growth in the Detroit region'. *Economic Geography* 46: 234–240.
- Unwin, D.J. 1996. 'GIS spatial analysis and spatial statistics'. *Progress in Human Geography* 20 (4): 540–551.
- Walker, J.; Jupp, D.L.B.; Penridge, L.K.; Tian, G. 1986. 'Interpretation of vegetation structure in Landsat MSS imagery: A case study in disturbed semi-arid Eucalypt woodlands; Part 1, Field data analysis'. *Journal of Environmental Management* 23: 19–33.
- Wulder, M.; Boots, B. 1998. 'Local spatial autocorrelation characteristics of remotely sensed imagery assessed with the Getis statistic'. *International Journal of Remote Sensing* 19 (11): 2223–2231.
- Zhang, Z.; Griffith, D.A. 2000. 'Integrating GIS components and spatial statistical analysis in DBMSs'. *International Journal of Geographical Information Science* 14 (6): 543–566.

Cite this article as: Lees, Brian. 2006. 'The spatial analysis of spectral data: Extracting the neglected data'. *Applied GIS* 2 (2): pp. 14.1–14.13. DOI: 10.2104/ag060014.