# DEVELOPMENT OF A FRAMEWORK TO ASSESS THE IMPACT OF SCALE DEPENDENT FACTORS ON THE CLASSIFICATION OF LANDCOVER MAPS

A.M.Lechner [a, *], S.D. Jones [a], S.A. Bekessy [b]

[a] School of Mathematical and Geospatial Sciences, RMIT University, Melbourne – a.lechner@student.rmit.edu.au, simon.jones@rmit.edu.au
[b] School of Global Studies, Social Science & Planning, RMIT University, Melbourne, sarah.bekessy@rmit.edu.au

**KEY WORDS:** Mapping, Classification, Scale, Spatial, Resolution, Landscape, Metric

**ABSTRACT:**

Remote sensing data type, classification technique and class description act together to produce, large differences in the classification of landcover. The resulting map will vary in the extent, patchiness and accuracy of classified areas. Differences in the classification of a landcover map are the result inter-relationships between a number of scale dependent factors such as pixel size, extent and smoothing filters. Many studies have investigated these factors individually using empirical data and have come to conclusions based on their unique case studies without isolating one factor from another. This study holistically investigates the different factors to better understand their interactions and relative importance.

The effect of scale dependent factors was tested on presence/absence tree cover maps; a common data layer used in landuse planning worldwide. Extent and pixel size were manipulated and a smoothing filter was applied to examine the differences in classification outcome. The aim of this project was to examine the relationship between scale dependent factors and landscape pattern as measured by total area and landscape metrics. It was found that changes in scale dependent factors affected the level of patchiness however total area remained constant. Furthermore the relationships between the factors generally appeared predictable. The study demonstrated that production of landcover maps can be subjective and that the final product can, to a large degree; result from the classification technique and sensors used.

## 1. INTRODUCTION

### 1.1 Scale Dependent factors

Scale dependent factors such as pixel size, study extent and smoothing filters affect the classification of landcover. These factors are dependent on the remote sensing data, classification techniques and class description used. Landcover maps will vary in their extent, patchiness and accuracy of classified areas based on the inter-relationship these factors. Many studies have investigated these factors using empirical data and have come to conclusions based on unique case studies investigating one factor in isolation (Hsieh et al., 2001). This study holistically investigated the different scale dependent factors to better understand their interaction and their relative importance.

In most studies data will be collected at the most appropriate scale however, for studies using remote sensing data, users are limited to specific scales available. The most appropriate scale for a study is a function of the environment (its spatial arrangement), the kind of information that is to be derived, and the classification technique used (Woodcock and Strahler, 1987). Numerous combinations of these factors are possible and their effects are usually interrelated and scale dependent.

At different spatial scales, landscape composition and configuration will change. Unfortunately, knowledge of how these spatial patterns change is limited (Wu et al., 2002). Variables such as area and spatial pattern will change when grain and extents is altered (Wiens, 1989). These variables can be used to quantify the degree of change in landscape composition and configuration resulting from changing the scale dependent factors.

The aim of this project is to investigate the effect of changing the spatially dependent factors in the context of vegetation extent mapping. It does not set out to solve the problem, rather to quantify its nature. The development of an integrated model is not new to remote sensing (e.g. Ju et al., 2005; Hsieh et al., 2001). Many studies have investigated scale dependent factors and have come to conclusions based on their scene and site specific evidence without considering their interactions (Hsieh et al., 2001). This paper aims to give a greater understanding of how they interact and to examine their relative importance.

The research objective for this project was to examine the relationship between the scale dependent factors and change in landscape pattern as measured by total area and landscape metrics. Users who base their analyses on a maps characterization of landscape pattern need to be aware that these patterns are scale dependent.

The interactions were investigated from the users' perspective through examining a number of landscape metrics. These metrics were chosen because they were simple and they summarised important patch characteristics. They have straightforward practical uses such as the measurement of total area and mean distance between patches rather than purely characterising fragmentation such as the fractal dimension index.

This study is unusual in that it uses real landscapes with a large study area and sample size. Other studies have used simulated landscapes (e.g. Li et al., 2005) and many studies which have used real landscapes tend to consider a small number of landscapes (e.g. De Clerq et al., 2006; Wu et al., 2002).

### 1.2 Landcover maps

This study utilizes the Tree25 presence / absence tree cover data set produced for the Department of Sustainability and Environment's (Victoria, Australia) Corporate Geospatial Data library (DSE, 2006) (Figure 1). This dataset is typical of woody / non-woody vegetation data layers used around the world in land use planning and habitat mapping.

The purpose of this dataset is varied, however, its initial purpose was to provide a comprehensive, consistent dataset for tree cover monitoring for the state of Victoria (Australia). Furthermore it is expected to provide an excellent source of data for applications which require the identification of remnant tree cover such as connectivity analysis and habitat modelling (DSE, 2006).

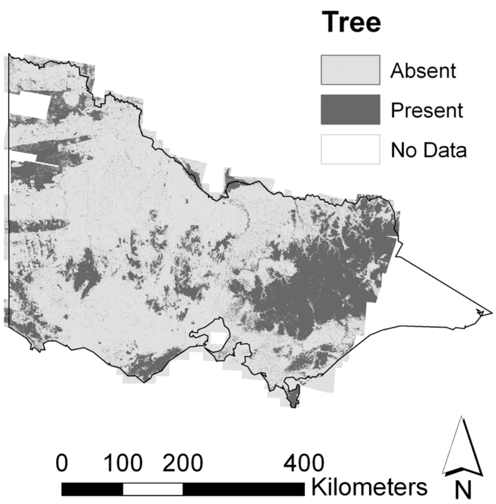


Figure 1. Map of the study area and Tree25, tree presence / absence data set overlaid.

### 1.3 Changing Scale Dependent factors

Pixel size (or spatial resolution), and extent were manipulated and a smoothing filter was used to examine the differences in classification outcome. All the variables were manipulated to simulate a range of conditions and determine how patchiness and patch area changes.

Pixel size is an important variable to investigate as using the default pixel size (i.e. sensor resolution) will result in a view of the world that relates to the sensor but may not necessarily reflect the needs of the question being asked (Fassnacht et al., 2006). Pixel size is one of the most important elements determining how the other scaling factors will change. Pixel size controls the limit of the smallest feature which can be extracted from an image. For areas where vegetation is highly fragmented such as urban areas and where patches appear as small as median strips and backyards, Jensen and Cowen (1999) concluded that at least 0.5 to 10m spatial resolution is required. Resolution was changed to simulate differing sensor resolutions by degrading the original classified image.

The second factor investigated was the use of a smoothing filter. Pixel based landscape classification can result in a salt and pepper effect because spatial autocorrelation is not incorporated in the classification technique (Ivits and Koch, 2002). A common practice used in remote sensing is smoothing the image by aggregating pixels to reduce classification error caused by this effect. The use of a smoothing filter will often result in the removal of edge complexity as well as increasing the minimum mappable unit (MMU). The MMU tends to be larger than the pixel size so that spatial and/or content information may be lost (Fassnacht et al., 2006). Larger MMUs may result in patches of interest being falsely combined within adjacent patches (Fassnacht et al., 2006). For this study the smoothing algorithm used was a majority filter. However, other filters can be used for the similar purposes such as mean or low pass filters.

The final variable investigated was extent, which is the total physical area covered by the data source. As the extent increases so does the probability of sampling rare classes (Wiens, 1989). Furthermore, if grain size is fixed, fragmentation increases with increasing extent (Riitters et al., 2000). The effect of extent was investigated by comparing between many landscape samples at different extents.

Landscape metrics were used to analyse the effects on landcover classification of varying pixel size, applying the smoothing filter and changing extents. These metrics were chosen because they describe simple patch characteristics that users of the Tree25 data layer in Victoria utilise. Users of landcover maps need a practical understanding of how scale dependent factors affect classification. For example, in the region of Victoria it is important to measure correctly the area of native vegetation, as a permit is required to remove, destroy or modify native vegetation from a landholding greater than 0.4 hectares (Cripps et al., 1999). Understanding the landscape metric, mean patch area is therefore critical when assessing the suitability of a particular landcover map for this purpose. Another example is to understand how the mean distance between patches changes by altering scale dependent factors. An understanding of distance between patches is useful for population modellers to calculate the probability of dispersal between populations based on this distance (e.g. RAMAS (Akcakaya, 2002)).

### 1.4 Data

The study area encompasses most of the state of Victoria which is approximately 227,416 km². The study area is dominated by broad acre cropping and crop pasture, vegetation and dryland pasture (Figure 2). There are a variety of abiotic and biotic processes occurring at multiple scales, resulting in a complex landscape composition and configuration.

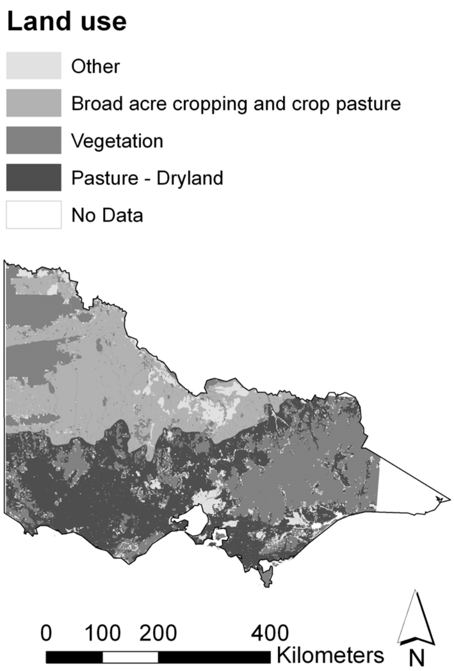


Figure 2. Map of land use in the study area.

Real landscapes were used instead of simulated landscapes such as those created by software such as Rule (Gardner, 1999) and SimMap (Saura and Mart´ýnez-Millán, 2000). Simulated landscapes are often used as replication at the landscape scale is often unfeasible. However Li et al. (2004) found that simulated landscape models had difficulties in capturing all the characteristics of real landscapes.

Comparison of the effects of scale between landscapes as well as within landscapes is important as the relationship between spatial patterns and scale may not be linear. Each landscape will vary in respect to the different processes operating at various scales (Wu et al., 2002). For example, disturbance can operate at many different scales from housing development to large fires to trees falls. Simulating landscapes at different scales which concurrently reflect reality is likely to be very difficult.

Numerous studies have been conducted on scaling effects but most of these studies have been confined to a few metrics or covered a narrow range of scales (Wu et al., 2002). This study is unusual in that the large study area allows for multiple replications at the landscape level of real landscapes. Studies which have a large sample size tend to use simulated landscapes (e.g. Li et al., 2005).

## 2. METHOD

### 2.1 Data

The original classified data were derived from SPOT panchromatic imagery with a 10 metre pixel size through a combination of digital classification and visual interpretation (DSE, 2006). No smoothing or filtering was applied at this layer creation stage. Tree cover is defined by the producers of the dataset as woody vegetation over 2 metres with crown cover greater than 10 percent.

### 2.2 Post - Processing

The original data were post-processesed to test the effect of resolution, extents and applying a smoothing filter on classification. All processing was performed using ArcGIS 9.1. The original image was first degraded to different pixel sizes. A filter was applied to the degraded images to smooth the image. Finally, each combination of filtered and degraded images were clipped to different extents.

#### 2.2.1 Pixel Size

Pixel size was changed by degrading the original image through interpolation techniques based on a nearest neighbour assignment using the centre pixel of the original image. This technique is particularly suitable for post processing of discrete data as it will not change the values of the cells (ESRI, 2007).

The original image was degraded from 10 metres to 100 metres at 10 metre increments. In this paper a decrease in resolution is analogous to an increase pixel size and vice versa.

#### 2.2.2 Smoothing filter

A majority filter was used to smooth the image. The majority filter replaces cells in a raster based on the majority of their contiguous neighbouring cells. The majority filter process has two criteria to fulfil before a replacement occurs. The number of neighbouring cells of a similar value must be in a majority and these cells must be contiguous around the centre of the filter kernel (ESRI, 2007). A 3 x 3 kernel was used for this process. A majority filter is useful for post processing as it works with discrete data.

#### 2.2.3 Extents

Subsets of this image were randomly clipped at 3000m, 10000m, and 20000m replicating landscapes of different extents (Figure 3). The extents represent the distance of a single side of a square. The image was clipped so that each replicant did not overlap. 20 samples were taken for each combination of smoothed image, extents and resolution with a total sample size of 600.

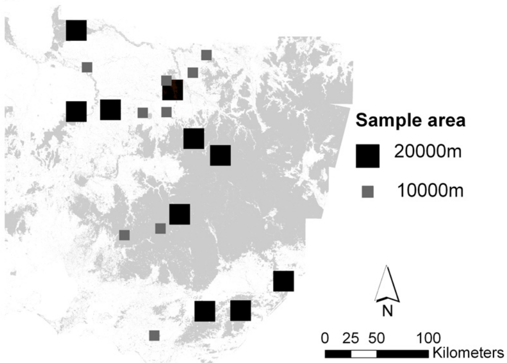


Figure 3. Clipped areas for western portion (50% of total area) of study area for extents 10000m and 20000m.

The lower bounds of the sampling size was set at 3 kilometres as suggested by Forman and Godron (1986) although it is recognized that in principle landscape size is related to the scale at which an organism perceives their environment. The upper limit was based on the approximate area size of a small catchment at around 20 kilometres. Furthermore, as the extents were increased beyond this amount, computer processing time increased markedly.

### 2.3 Calculating Landscape Metrics

Area was calculated based on pixels classified as either tree present or absent as identified by ArcGIS. Landscape metrics were then calculated using the fragstats package (McGarigal et al., 2002). Five landscape metrics were used: patch number, mean patch area, mean patch density, mean nearest neighbour distance, and mean perimeter to area ratio.

## 3. RESULTS

The study found that while the total area classified remained relatively constant when the image resolution changed there were large differences in the patchiness resulting from changing resolution and using a smoothing filter. As image spatial resolution decreased (i.e. pixel size increased) or a smoothing filter was applied the subtle levels of patchiness disappeared. Small patches either aggregated into larger patches or disappeared (Figure 4). Some measures of patchiness appeared to be non-random in relation to the spatial dependent factors, however this was not always the case. For most metrics used it was impossible to test the effects of changing the extents due to the low sample size and high variability.

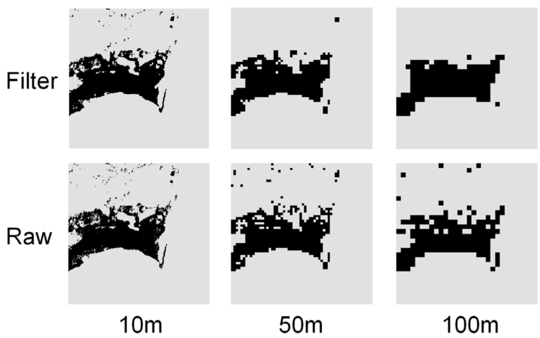


Figure 4. Example of processing. The original (raw) image at 10 metre spatial resolution was degraded up to 100 metres. For each degraded image a majority filter was used to smooth the image.

It was found that the greater the extent the greater the mean number of patches, and the lower the spatial resolution the lesser number of patches (Figure 5). Additionally using the smoothing filter also resulted in a lesser number of patches.

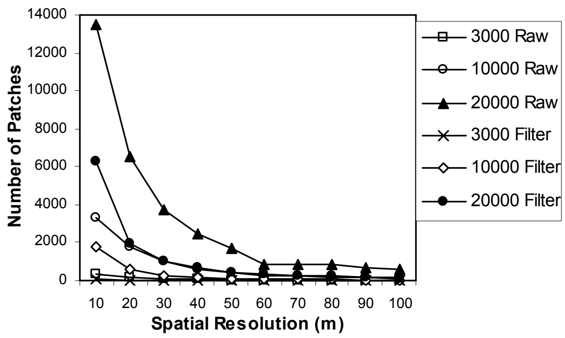


Figure 5. Comparison of the effect of changing extents, spatial resolution and applying a smoothing filter on the number of patches.

The relationship between mean patch area and the spatially dependent factors was the opposite to mean number of patches. Mean patch area increased as resolution decreased (Figure 6a). Furthermore, using a smoothing filter and decreasing the spatial resolution resulted in an increase in mean patch area. The mean number of patches changed, as a result of changing the spatial resolution, however the total area classified as tree or non-tree remained constant (Figure 6b). Due to the high standard error resulting from the small sample size (n = 20) a comparison between extents could not be conducted. The differences between the value of proportion classified as present or absent for different extents is the result of high variability in the landscape. However the filtered data tended to have a significantly ($P < 0.05$) lower proportion of cells classified as present for both 3000m and 20000m extents.

The relationship between patch area and resolution was not perfectly linear. Whilst the overall trend was to increase patch area with increasing resolution this was not always the case. Applying the majority filter caused a greater change at lower spatial resolutions. At 10m resolution there was a drop in the increase in mean patch area of 5% compared to 115% at 100m resolution for 3000m extents and at 10m resolution there was a drop in the mean patch area of 93% compared to 505% at 100m resolution for 30000m extents.

The next metric analysed was patch density which is calculated as the number of patches in the landscape divided by the total landscape area. As the spatial resolution decreased (i.e. pixel size increased) the mean patch density decreased for all extents (Figure 7). This decrease was quite dramatic at 10m resolution there was a drop in the mean patch density from 18 to .44 at 100m resolution for 3000m extents and from 33.6 to 1.4 for 20000 extents. The results of applying a filter had similar affect as decreasing resolution, that is, decreasing patch density. However applying the filter caused a greater change at lower resolutions. At 10m resolution there was a drop in the mean patch density from of 53% compared to 71% at 100m resolution for 3000 extents and at 10m resolution there was a drop in the mean patch density from of 53% compared to 78% at 100m resolution for 30000m extents. Figure 8 shows the relationship between patch density and resolution for single samples compared to figure 7 which shows the mean of all the samples. Figure 7 shows that as spatial resolution decreases patch density will predictably decrease. The relationship appears to fit an inverse exponential function.

The next metric investigated isolation and proximity. This was done by calculating the nearest neighbouring value based on the shortest edge-to-edge distance for a patch of the same type. As spatial resolution increased the nearest neighbour distance generally increased (Figure 9). However, of all the measures of patchiness this appeared to be least predictable. The variability appeared to be inconsistent and unrelated to resolution. It was impossible to compare extents because of the high standard error. Figure 10 shows that there was no relationship between resolution and using a smoothing filter.

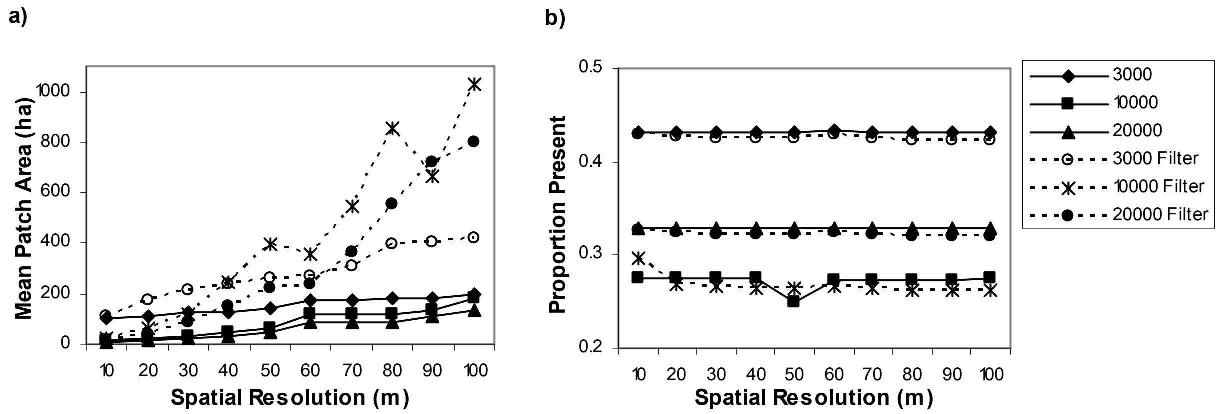


Figure 6. Mean patch area in hectares. a) Raw data. b) Data smoothed with a majority filter.

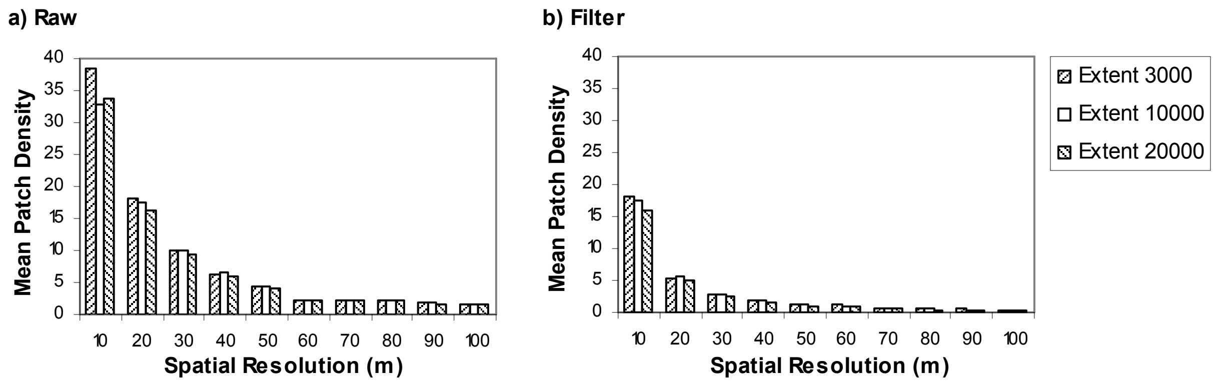


Figure 7. Mean patch density: number of patches in the landscape, divided by total landscape area. a) Raw data. b) Data smoothed with a majority filter.

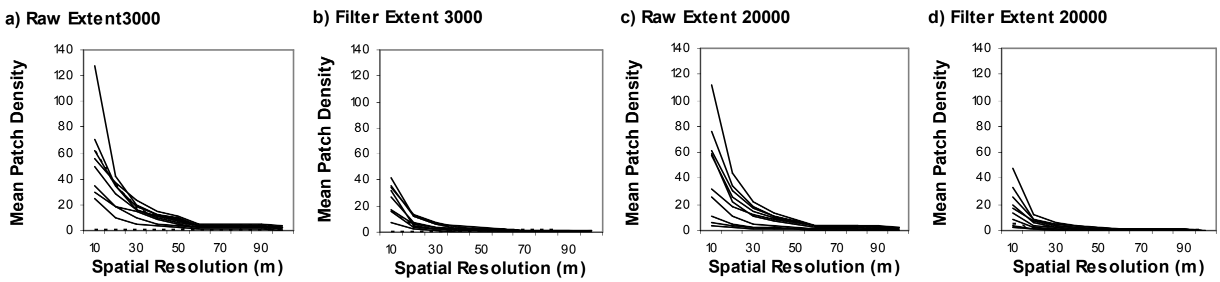


Figure 8. Mean patch density (number of patches in the landscape, divided by total landscape) for 10 samples at extents 3000 and 20000 for data before and after smoothed with majority filter. This demonstrates a predictable decline in patch density as spatial resolution is degraded from 10 to 100 meters for each sample.

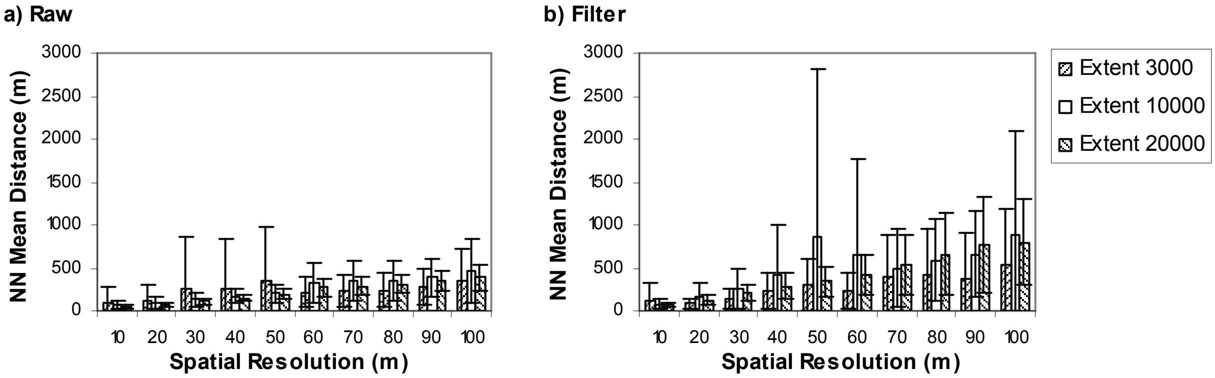


Figure 9. Mean Euclidian Nearest Neighbour Distance in metres, Error bars indicate standard deviation. a) Raw data b) Data smoothed with a majority filter.

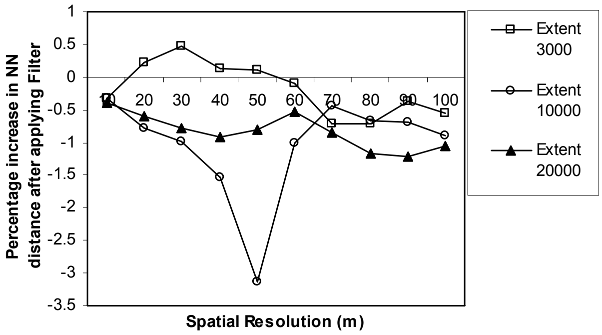


Figure 10. Percentage change in mean nearest neighbour distance between patches after applying the majority filter to images from 10 m to 100m spatial resolutions.

The final metric considered was the perimeter to area ratio which describes the relationship between shape and area. As spatial resolution increased the ratio decreased (Figure 11). The mean perimeter to area ratio and spatial resolution is the inverse of patch area. By default, the mean perimeter to area ratio is strongly related to patch area. For example, if shape is held constant and patch size increased there will be a decrease in the ratio. Applying the smoothing filter resulted in a predictable decrease in the mean perimeter to area ratio.

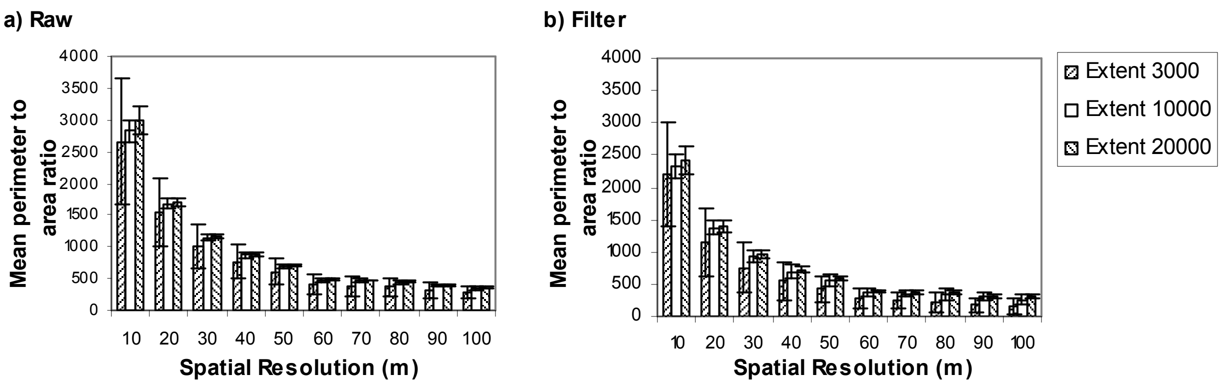


Figure 11. Mean perimeter to area ratio. Error bars indicate standard deviation for a) Raw data b) Data smoothed with a majority filter.

## 4. DISCUSSION

It can be seen that changes in scale dependent factors affect the patchiness and total area classified. Sometimes the relationships between factors are predictable, however, that is not always the case and not all metrics varied in the same way.

The relationship between mean patch area and resolution is not constant and is likely to be the result of landscape patterns, whilst the relationship between resolution and applying the smoothing filter and mean patch density appeared predictable. Applying the smoothing filter at lower spatial resolutions had a greater effect on mean patch density and mean patch area lower resolutions. Furthermore, applying the smoothing filter resulted in a significant difference in the total area classified.

Due to the small sample size and large variability for most metrics it was impossible to compare the effects of changing the study area extents. We would expect greater variability in smaller extents than larger ones and that a larger extent will have a greater probability of containing all the variability within a landscape. Also, if the sample size was increased the mean of these samples should reflect the mean of the landscape. However, increasing the sample size or the sample area could be problematic as the area of real landscapes is finite.

## 5. CONCLUSION

The measurement of landscape pattern from landcover maps has become a common practice in various fields such as landscape ecology. However many people are unaware of the scale dependency of this phenomena. This study demonstrates that characterization of landscape patterns by landcover maps are the product of the inter-relationship of a number of scale dependent factors such as the spatial resolution of the imagery, applying a smoothing filter and study extents used. Landcover maps will vary in the extent and patchiness of classified areas based on this inter-relationship.

Landscape pattern will change as result of the interaction of the scale dependent factors. For example, the effect of using a majority filter at low spatial resolutions will not be the same when used at high resolutions. Techniques that are used at one resolution are not necessarily transferable to different resolutions and may result in a very different classification. This has wide ranging consequences for users transferring techniques used on medium resolution imagery from sensors such as Landsat to high resolution imagery from sensors such as IKONOS and Quickbird.

This study was the first step in the development of a framework to quantify the magnitude of the effect of different spatial dependent factors on the landcover classification. It demonstrated that there is considerable interaction between the scale dependent factors, indicating that the investigation of spatial dependent factors need to be done simultaneously.

Future research is needed to assess the effect of these spatially dependent factors on accuracy as well as patchiness and area. Furthermore as the landscape patterns found in the study area may be site specific it is difficult to generalise to other areas. Thus, there is a need to perform the same spatial analysis for a wide range or resolutions using different smoothing filters and extents in multiple real landscapes settings to create a significant volume of data. This will allow for wide ranging generalisations to be made which will be the basis for the development of guidelines for map users.

## ACKNOWLEDGEMENTS

This work was supported by ARC Discovery Grant DP0450889 and by the Landscape Logic research hub. The authors would also like to acknowledge the help of Michael Conroy and John White at the Department of Sustainability and Environment and Bill Langford and Ascelin Gordon from Re-imagining the Australian Suburbs.

## REFERENCES

Akcakaya, H.R., 2002. *RAMAS GIS: linking spatial data with population viability analysis*, Applied Biomathematics, New York.

Cripps, E., Binning, C., Young, M., 1999. *Opportunity Denied: Review of the legislative ability of local government to conserve native vegetation.*, Environment Australia, Canberra.

De Clercq, E.M., Vandemoortele, F., De Wulf, R.R., 2006. A method for the selection of relevant pattern indices for monitoring of spatial forest cover pattern at a regional scale, *International Journal of Applied Earth Observation and Geoinformation*, 8(2), pp. 113-25.

DSE 2006. *Corporate Geospatial Data Library*, Melbourne, Australia. http://www.dse.vic.gov.au/dse/. (7 Aug. 2006)

ESRI 2007. *Arc GIS Desktop Help*. http://webhelp.esri.com/arcgisdesktop/9.1/index.cfm?TopicName=welcome. (7 Aug.2006)

Fassnacht, K.S., Cohen, W.B., Spies, T.A., 2006. Key issues in making and using satellite-based maps in ecology: A primer, *Forest Ecology and Management*, 222(1-3), pp. 167-81.

Forman, RTT & Godron, M 1986. *Landscape ecology,* Wiley, New York.

Gardner, R.H., 1999. RULE: map generation and a spatial analysis program, In: *Landscape Ecological Analysis: Issues and Applications*, Springer, New York, p. 280–303.

Hsieh, P.F., Lee, L.C., Chen, N.Y., 2001. Effect of spatial resolution on classification errors of pure and mixed pixels in remote sensing, *Ieee Transactions on Geoscience and Remote Sensing*, 39(12), pp. 2657-63.

Ivits, E., Koch, B., 2002. Landscape connectivity studies on segmentation based classification and manual interpretation of remote sensing data, *eCognition User Meeting, October 2002, München*.

Jensen, J.R., Cowen, D.C., 1999. Remote sensing of urban/suburb an infrastructure and socio-economic attributes, *Photogrammetric Engineering and Remote Sensing*, 65(5), pp. 611-22.

Ju, J., Gopal, S., Kolaczyk, E.D., 2005. On the choice of spatial and categorical scale in remote sensing land cover classification, *Remote Sensing of Environment*, 96(1), pp. 62-77.

Li, X., He, H.S., Wang, X., Bu, R., Hu, Y., Chang, Y., 2004. Evaluating the effectiveness of neutral landscape models to represent a real landscape, *Landscape and Urban Planning*, 69(1), pp. 137-48.

Li, X., He, H.S., Bu, R., Wen, Q., Chang, Y., Hu, Y., Li, Y., 2005. The adequacy of different landscape metrics for various landscape patterns, *Pattern recognition*, 38(12), pp. 2626-38.

McGarigal, K., Cushman, S.A., Neel, M.C., Ene, E., 2002 *FRAGSTATS: Spatial Pattern Analysis Program for Categorical Maps. Computer software program produced by the authors at the University of Massachusetts, Amherst*, http://www.umass.edu/landeco/research/fragstats/fragstats.htm. (30 Jan. 2007).

Riitters, K., Wickham, J., O'Neill, R., Jones, B., Smith, E. 2000. Global-Scale Patterns of Forest Fragmentation, *Conservation Ecology*, 4(2). http://www.consecol.org/vol4/iss2/art3/. (14 Jan. 2007)

Saura, S., Mart´ýnez-Millán, J., 2000. Landscape patterns simulation with a modified random clusters method, *Landscape Ecology*, 15(7), p. 661–78.

Wiens, J.A., 1989. Spatial scaling in ecology, *Functional ecology*, 3(4), pp. 385-97.

Woodcock, C.E., Strahler, A.H., 1987. The Factor of Scale in Remote-Sensing, *Remote Sensing of Environment*, 21(3), pp. 311-32.

Wu, J.G., Shen, W.J., Sun, W.Z., Tueller, P.T., 2002. Empirical patterns of the effects of changing scale on landscape metrics, *Landscape Ecology*, 17(8), pp. 761-82.