SCALING THE WALLS OF HISTORY: THE PERILS AND PITFALLS OF MULTI-SCALE LAND COVER COMPARISON

S.D. Jones *, J.G. Ferwerda, K.J. Reinke

School of Mathematical and Geospatial Sciences, RMIT University,
GPO Box 2476V, Melbourne 3001, VIC Australia - simon.jones@rmit.edu.au

Technical Commission II

KEY WORDS: Land Cover Mapping, Classification, Spatial Resolution, Remote Sensing

ABSTRACT:

The problem addressed is this: most environmental issues require context to solve them. Is the ocean getting warmer? Is the desert growing? Is the forest declining? Solution: measure the temperature / size / leaf area. But such measurements only have significance if there are other *comparable* historical measurements to compare them too. This paper is about that word comparable. Can we really compare landscape generalisations gathered at different times and at different spatial scales? Today we have the ability to produce land cover maps at a very high spatial resolution (grid cell sizes of 10, 5 or even 1 m). Historically, data has typically been collected at coarser spatial scales (grid cell sizes of 50, 100 or even 1000 m). To facilitate comparison, modern data is often re-scaled to match the historical data. To evaluate the validity of this process, a series of synthetic landscapes were created. These landscapes included the full range of possible dispersion from a random spatial distribution of scene elements to a highly clustered spatial pattern. Each simulated landscape was firstly classified and then degraded to four levels of generalisation (simulating a range of spatial resolutions). In parallel, the process was reversed and the simulated landscapes were degraded and then classified. The resultant classifications were then compared. In all cases the integrity of the data was best preserved when the image was more highly spatially autocorrelated. Changing the spatial scale (i.e. degrading) of classifications resulted in a rapid decline in information content, particularly in more random landscapes. The implications of these results are then discussed.

1. THE CHALLENGE

1.1 The need for long term databases

Land cover change mapping is an increasingly important activity. Many international treaties and protocols mandate a monitoring or repeat mapping process. For example, the key questions of: ‘how much?’ and ‘how fast?’ environmental degradations are proceeding is addressed in a number of United Nation (UN) conventions: The Convention on Biological Diversity (CBD), requires parties to: (i) regularly report how much landscape diversity and natural habitat is being lost[1)]; (ii) Report how much ecosystem diversity (quality) is being lost[2)]; (iii) "Rehabilitate and restore degraded ecosystems and promote the recovery of threatened species..."[3)] The UN Framework Convention on Climate Change (UNFCCC) emphasises the need for comprehensive policies and measures to address issues related to the sources, sinks, and reservoirs of greenhouse gases, taking into account different socioeconomic contexts[4)]. The Kyoto Protocol calls for national reporting systems for Carbon sinks and sources[5)].

Such monitoring or repeat mapping processes imply the creation and maintenance of long term spatial data sets. These can be created from interpolated observations, modelled predictions, and / or extrapolation. Remotely sensed imagery, from aircraft or satellite-based sensors, is increasingly used to aide in this process, or to independently to derive ‘total-sample’ thematic layers or classifications.

1.2 An Australian example

Australia’s National Carbon Accounting System provides information on land-based sources and sinks of greenhouse gases to fulfil international reporting obligations under the Kyoto Protocol, as well as providing annual estimates to Australia's National Greenhouse Gas Inventory. Fundamental to accounting for Carbon change is an understanding of the change in land cover. The impact of an event associated with land cover change may continue over many years and vary with time since the event took place. It is, therefore, necessary to monitor change in land cover over extended periods of time.

In Australia, ~370 Landsat Thematic Mapper (TM) scenes were used to create a continent-wide (690 million hectares) database of Forest / Non-Forest land cover for 12 time periods, spanning 1972 to 2002. Forests were defined as having a minimum of 20% tree crown cover and a minimum height of 2 metres at maturity. A minimum area of 0.2 hectares is also imposed (Furby, 2002, Jones *et al.*, 2004). However, the spatial resolution of this dataset is not constant over the thirty year archive period. The initial (1972) grid cell resolution is $50m^2$ whereas the more recent spatial data layers (e.g. 2002) have a grid cell size of $25m^2$. Comparison of these two epochs requires the rescaling of the modern dataset to that of the initial 1972 spatial resolution.

* Corresponding author.
1) Convention on Biological Diversity, Articles 7, 8, & 9.
2) Convention on Biological Diversity, Articles 7, 8, & 9.
3) Convention on Biological Diversity, Articles 8(f).
4) United Nations Framework Convention on Climate Change, Article 4, para. 3.
5) Kyoto Protocol, Article 5.

1.3 Remote sensing data

As spatial scientists we are familiar with the concept of scale. We are used to ordering or classifying landscapes into convenient land cover / land use units at a variety of application specific spatial scales. Indeed, there are many mapping standards that suggest or impose a scale. For example, the US Federal Geographic Data Committee (FGDC) have Content Standards for Digital Geospatial Metadata (that relate to cadastral, soil or vegetation mapping etc.).

In practise, we use technologies to derive land cover maps. Remote sensing is a common method of deriving land cover datasets. Earth observation technologies have evolved rapidly in the past thirty years and continue to do so. There are now a multitude of satellite earth observing systems (see the CEOS – Committee for Earth Observation Satellites online database: http://alto-stratus.wmo.ch/sat/stations/SatSystem.html) for a comprehensive listing. Much of this data is in use for monitoring or repeat mapping processes. Such remotely sensed images are not uniform however, and vary markedly according to platform used, sensor (spatial, radiometric, spectral and temporal resolutions) and scene illumination / viewing geometry (Lillesand and Kiefer, 2004). Internationally *de facto* standards have emerged: the International Geosphere-Biosphere programme (IGBP) has created a standard global land cover classification (at 1km^2); whilst the ESA GLOBCOVER initiative will produce a 300m global land cover classification for 2005 based on the MERIS sensor.

Image analytical algorithms (such as classifications), and temporal analyses, are often carried out without any reference to the spatial scale of the study. Users assume, for example, that image classification will work equally well on 1.1km AVHRR (Advanced Very High Resolution Radiometer onboard the NOAA satellites) data as it does on 28.5m TM (Landsat Thematic Mapper 4 / 5) or 4m multi-spectral IKONOS data. But is this the case? The spatial composition of a pixel's various spectral 'components' and their mixing will vary on a per landscape basis but also as a function of spatial resolution. The spectral composition and mixing of a pixel is a function of not only the spectral resolution but also of the spatial resolution afforded by the sensor in addition to the distribution and organisation of the land cover units being mapped (Moody and Woodcock, 1994). Marcaeu and Hay (1999) and Turner et al. (1989) also note that as aggregation occurs information is lost at coarser scales and that measurements made at different scales may not be directly comparable.

The data used in monitoring or repeat mapping processes is very diverse. Integrations and comparisons between these diverse datasets are problematic but becoming increasingly frequent.

2. METHOD

2.1 Purpose

The aim of this experiment was to determine whether multi-scale land cover images can support accurate comparisons. This was explored by varying (i) the spatial resolution, (ii) the degree of landscape spatial autocorrelation and (iii) the order of processing (i.e. classification and resampling).

2.2 Simulating landscapes

A series of eight synthetic landscapes were generated using the Grid Cell Uncertainty Model (GCUM) (Hunter et al., 1994 for a detailed description). The GCUM generates synthetic landscapes that vary from a completely random distribution of landscape elements to a highly clustered or spatially auto-correlated pattern. Spatial autocorrelation describes the probability for the attributes of geographically neighbouring grid cells to be more similar than distant ones. The GCUM model creates a set of random landscape elements (r = 0 or random, where r is the degree of spatial autocorrelation). These can then be grouped or clustered into progressively more spatially autocorrelated landscapes (increasing r values) using the spatial autocorrelation index of Cliff and Ord (1981). The simulated landscapes can be conceptualised as being any biophysical variable (ocean temperature, desert or forest extent, biodiversity etc.).

2.3 Simulating historical / spatial resolution changes

To simulate the imaging of the landscapes at various spatial resolutions over an extended period of time, four levels of landscape aggregation were created (Figure 1). The *original* landscape (A1) contains all of the data and could be considered comparable to a 10m spatial resolution image. Level 2 (L2) has been resampled or degraded and can be considered comparable to a 20m spatial resolution image. Level 4 (L4) has been degraded from the original image to an equivalent pixel size of 40m. Level 8 (L8) is the coarsest level presented here and has an 80m spatial resolution. Each of the degraded landscapes was then classified using the Iterative Self-Organising Data Analysis Technique (ISODATA) (Tou and Gonzalez, 1977) to yield a 2-class binary result (Figures 1, B2, B3, B4) The process was then repeated but reversing the order of tasks so that the classification was performed first on the original un-degraded landscape and the resultant classified images then resampled to the aforementioned four levels of generalisation (Figures 1, C2, C3, C4). These images represent the landscapes as they would appear if classified first at a higher spatial resolution and then degraded to *match* a historical dataset.

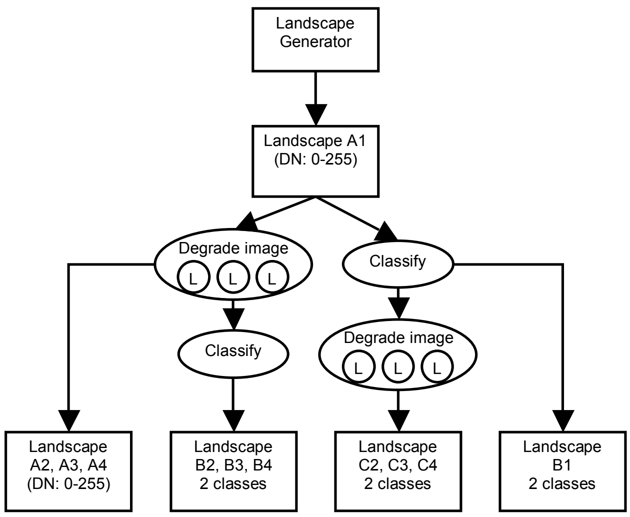
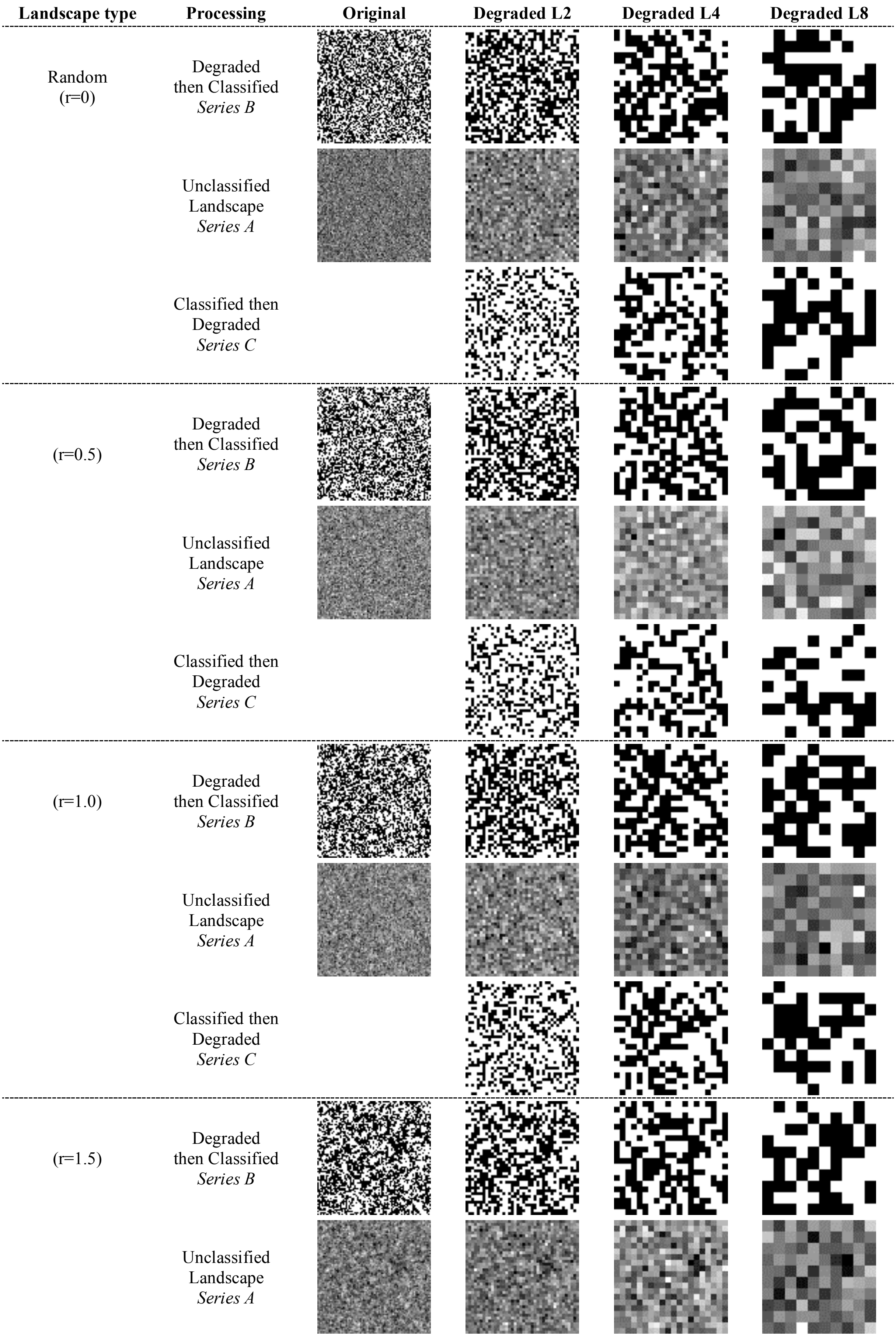


Figure 1. Processing method


Landscape type
Processing
Original
Degraded L2
Degraded L4
Degraded L8
Random
(r=0)
Degraded
then Classified
Series B
Unclassified
Landscape
Series A
Classified then
Degraded
Series C
(r=0.5)
Degraded
then Classified
Series B
Unclassified
Landscape
Series A
Classified then
Degraded
Series C
(r=1.0)
Degraded
then Classified
Series B
Unclassified
Landscape
Series A
Classified then
Degraded
Series C
(r=1.5)
Degraded
then Classified
Series B
Unclassified
Landscape
Series A

Classified then Degraded *Series C*

(r=2.0) Degraded then Classified *Series B*

Unclassified Landscape *Series A*

Classified then Degraded *Series C*

(r=2.3) Degraded then Classified *Series B*

Unclassified Landscape *Series A*

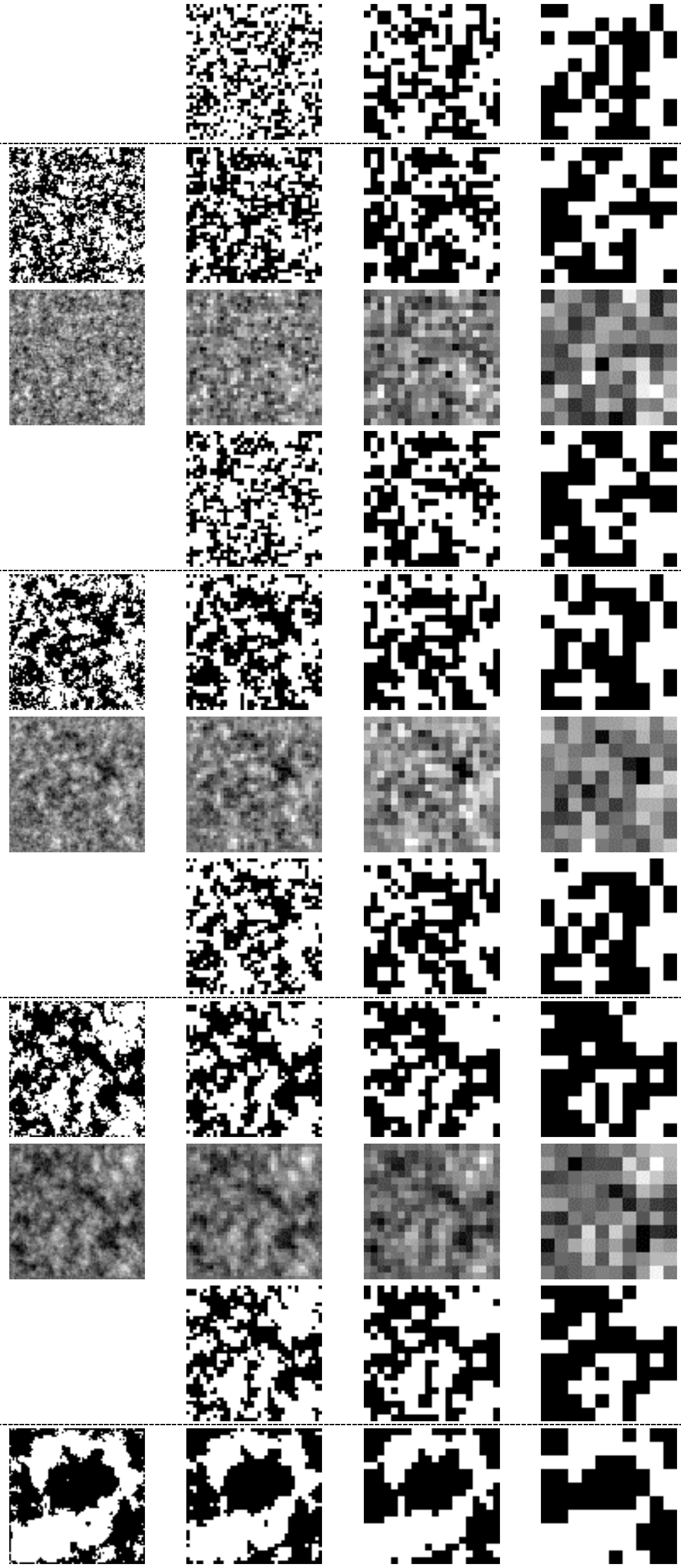
Classified then Degraded *Series C*

(r=2.4) Degraded then Classified *Series B*

Unclassified Landscape *Series A*

Classified then Degraded *Series C*

Highly spatially autocorrelated (r=2.49) Degraded then Classified *Series B*



Unclassified
Landscape
Series A

Classified then
Degraded
Series C

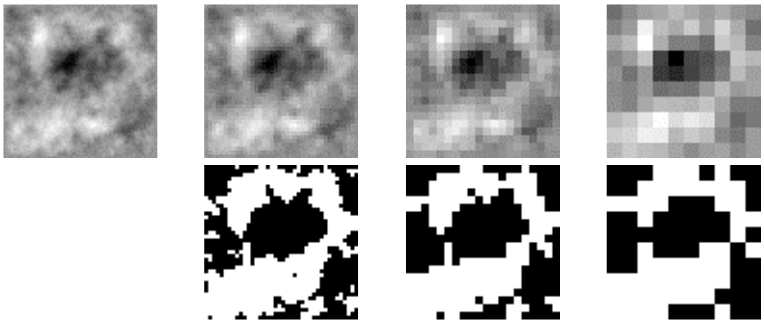


Figure 2. Synthetic landscapes created (at eight levels of spatial autocorrelation). Each was: degraded to four levels of aggregation and classified; and, classified and then degraded

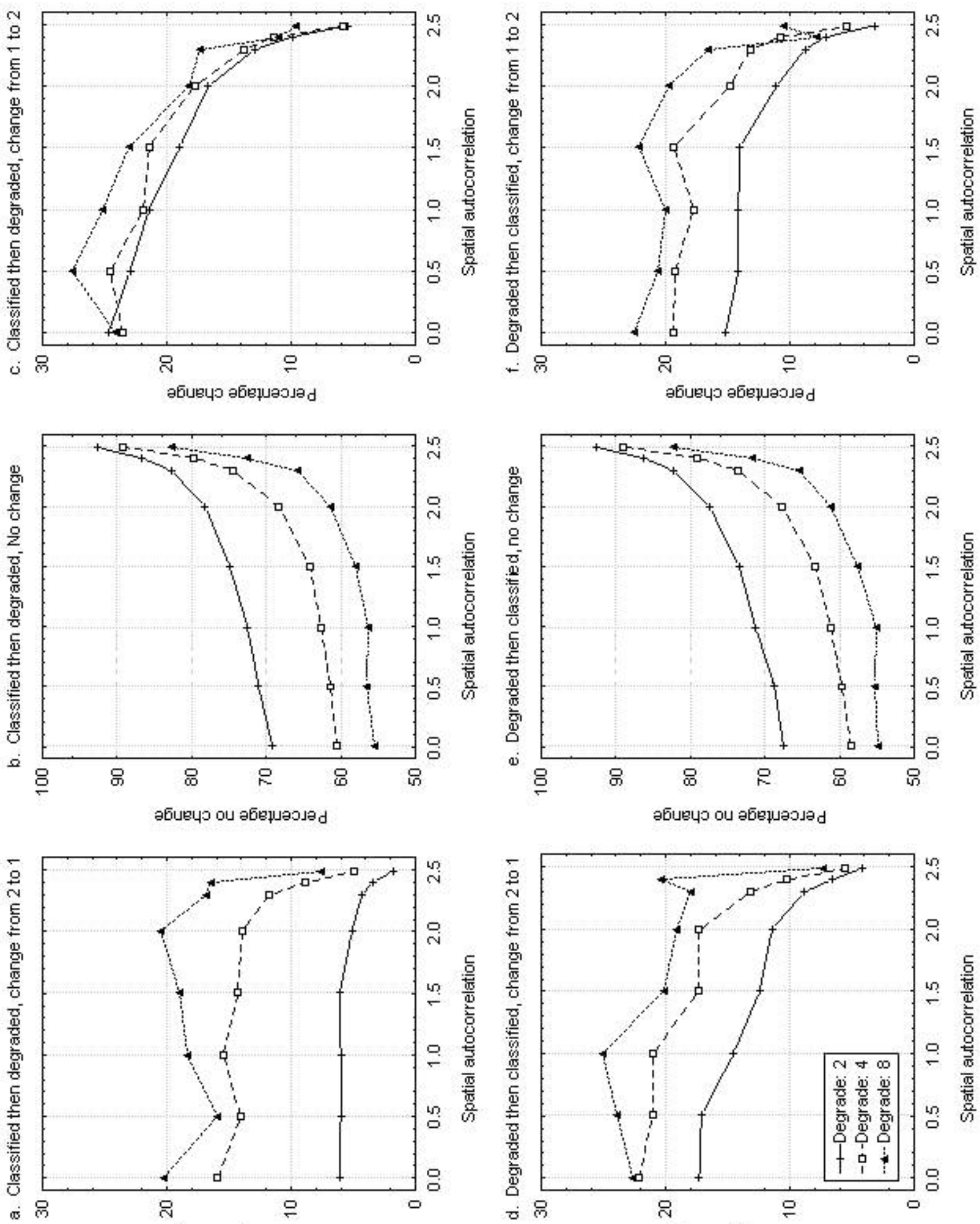


Figure 3. Percentage change between categories in the synthetic landscapes

3. RESULTS AND DISCUSSION

3.1 Differences and changes between the evaluations

The first point to note here is that the two 'creation processes' (*degraded and then classified;* and, *classified and then degraded* – Figure 1) do not result in the same landscape representation. This is particularly the case for random landscapes (r = 0) and even moderately clustered ones (r > 2). Visual inspection of Figure 2 confirms this with gross differences evident.

To investigate these resolution induced landscape changes further, the changes in class memberships between the two process outputs were calculated for each landscape (Figure 3). The top row (Figure 3a-c) illustrates the percentage change when landscapes were first classified and then degraded. The bottom row (Figure 3d-e) the percentage change when landscapes were first degraded and then classified. Random (r = 0) landscapes showed the most change, with 30-45% of pixels changing class. For all levels of aggregation a steady decline in the number of pixels changing class occurred as the spatial autocorrelation increased (and landscapes become more clustered). At a spatial autocorrelation of ~r = 0.21 all landscapes became much less susceptible to class changes showing marked increases in no percentage change (Figure 3b and 3e). Changes from class 1 to class 2 (and vice-versa) were also plotted (Figure 3c & g and 3a & d). Changes from class 1 to class 2 were marginally greater in landscapes that were classified and then degraded; whilst changes from class 2 to class 1 occurred more frequently in degraded and then classified landscapes. Again, landscapes with a spatial autocorrelation of r = 0.21 and above (i.e. highly clustered scene elements) seemed to exhibit far fewer changes.

Almost without exception, higher spatial resolution classifications, regardless of their creation process, produced lower landscape change scores. That is to say, Level 2 aggregations produced a lower percentage change than Level 4 and Level 8 and Level 4 aggregations fewer than Level 8.

3.2 Implications for comparing landscape generalisations gathered at different spatial scales

Earlier we posed the question: Can we really compare landscape generalisations gathered at different temporal and spatial scales? In a way the question is mute, since we will be forced to make these comparisons regardless of how dubious they may be. Perhaps a more pertinent question is: Can we facilitate and enhance land cover change mapping by ensuring the measurements really are comparable?

Results obtained using the synthesised landscapes suggest that it is best practice to acquire a landscape (regardless of its spatial autocorrelation) at the spatial resolution it is to be mapped at. In all cases the integrity of the data was best preserved when the landscape was more highly spatially autocorrelated. Changing the spatial scale (degrading) of classifications resulted in a rapid decline in information content, particularly in more random landscapes. If degradation (resampling) is required, for example for comparison with historical datasets, imagery should be resampled to the appropriate resolution and then classified. Unfortunately this involves considerable effort, since the original data must be reprocessed, rather than just results (or outputs) being compared. Comparisons of classifications of different lineages (instrument / sensors) and degradations of existing classifications should only be used as a last resort and appropriate metadata information provided to the data user.

4. CONCLUSIONS

Today we can map our ocean, desert or forest attribute from space, with IKONOS or Quickbird imagery, with a high degree of accuracy, to within ~1-5m. If we wish to understand and compare it with historical data, as mandated in a number of UN treaties and conventions, we would be often forced into using episodic and discontinuous data with a much poorer spatial resolution. The results from this paper suggest making such comparisons between datasets derived at different times using different sensing systems is perilous. Even simple binary classifications in highly clustered landscapes (the best case scenario) can result in some errors. In landscapes with low spatial autocorrelation (i.e. disaggregation in the attribute measured), and / or where the differences in spatial resolution between the systems is large, the errors may impede coherent analysis. There is however, a lack of research into the theory behind these integrations and spatial scale changes (Jensen et al., 1998).

Further work: This paper has examined class changes in only the most simple of classifications (a two class image or binary matrix). Subsequent investigations will examine the effects of spatial resolution and classification accuracy on multi-class classification schemes.

REFERENCES

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

Furby, S., (2002) Land cover Change: Specification for Remote Sensing Analysis. National Carbon Accounting System, Technical Report No. 9, Australian Greenhouse Office, Canberra, ACT ISSN: 1442 6838.

Hunter, G.J., Goodchild, M.F. & Robey, M., (1994) A Toolbox for Assessing Uncertainty in Spatial Databases. Proceedings of the AURISA '94 Conference, Sydney, Australia, pp. 367-79.

Jensen, J.R., Saalfeld, A., Broome, F., Cowen, D., Price, K., Ramsey, D., and Lapine, L. (1998) Spatial Data Acquisition and Integration. Washington, D.C., University Consortium for Geographic Science Research White Paper (available at http://www.ucgis.org/research_white/data.html.)

Jones, S.D., Richards, G., Lowell, K., Woodgate, P., Buxton, L., (2004) Towards an Understanding of Uncertainty in Greenhouse Forest Assessments, ISPRS Congress, Turkey, 12-23 July 2004

Lillesand, T.M., Kiefer, R.W., (2004) Remote sensing and Image Interpretation, 5th edition, John Wiley and Sons.

Marcaeu, D.J. and Hay, G.J. (1999) Remote Sensing Contributions to the Scale Issue, Canadian Journal of Remote Sensing, 25 (4): 357–366.

Moody, A., Woodcock, C.E., (1994) Scale-dependent errors in the estimation of land-cover proportions – Implications for global land-cover datasets, Photogrammetric Engineering & Remote Sensing. Vol. 60, no. 5, pp. 585-594.

Tou, J.T., and R.C. Gonzalez. (1977) Pattern recognition principles. Addison-Wesley, Reading, Massachusetts. 377 pp.

Turner, M.G., O'Neill, R.V., Gardner, R.H., Milne, B.T., (1989) Effects of changing spatial scale on the analysis of landscape pattern, Landscape Ecology, 3 (3-4): 153-162