THE IDENTIFICATION OF DATA PRIMITIVES TO SEPARATE THE CONCEPTS AND SEMANTICS OF *LAND COVER* AND *LAND USE*: THE EXAMPLE OF 'FOREST'

Alexis Comber
Department of Geography, University of Leicester, Leicester, LE1 7RH, UK, Email ajc36@le.ac.uk

KEY WORDS: land use, land cover, primitives, semantics, text mining

ABSTRACT
This paper is concerned with data and classifications that confuse the concepts of land cover and land use. Conceptual confusion is problematic for data integration and modelling. This has resulted in calls for the separation of land use and land cover from the global land monitoring community. Text mining approaches are combined with social network analyses as a method for unravelling the different concepts embedded in land cover and land use semantics and applied to descriptions of forest cover and use. Whilst the results show the distinct biological dimension to land cover descriptions and the socio-economic character of land use, they reveal the deep degree of semantic confusion embedded in land cover and land use descriptions. The implications for this lack of internal semantic accuracy and consistency in land resource inventories are discussed and the case made for separating the concepts of land cover from land use.

1. INTRODUCTION

There are many ways of representing and describing land based features. Historically the overriding trend in land inventory, nationally and locally, has been to record information on *land use* (Fisher et al., 2005). Since the 1970's many land inventories have reported on *land cover* driven by the availability and machine processing of satellite imagery compared to the earlier demand and application driven surveys. In the process, land use and land cover have become inter-changeable concepts often because of the demands of different agents and actors involved in the commissioning process. Confused thinking in the reporting of land information hinders the translation of information from different surveys. In the past this may not have mattered as obtaining land data often involved extensive dialogue with the data producer and extensive metadata was included in the survey memoir. Now there is no memoir and data access is relatively easy and quick via various web-portals, eScience initiatives and spatial data infrastructures (e.g. the computing Grid and the EU's INSPIRE). These factors minimise the interaction between producers and users and therefore reduce the potential for them to clarify any inconsistencies in their shared understandings of, for instance, what they mean by the term 'forest'. The consequences of unavoidable inter-institutional negotiation over data specifications (see Comber et al., 2002; 2003) and the spatial and spectral limitations of satellite imagery compared to field survey are three-fold. First, each individual member or institution on the steering committee of any big mapping project is forced to accept some degree of compromise over the specification of the land features to be identified. Second, the agreed classification is a hybrid of land cover and necessarily *inferred* land use. Third, the data users always have to 're-work' or manipulate the data in some way in order to be to incorporate the data into *their* analyses and to answer *their* questions. These problems are in part caused by the process of accountable data commissioning but mainly by the lack of 'data primitives' in land information, especially in land data derived from remote sensing. Data primitives are here defined as the dimensions or measurements that describe at the most fundamental level the processes under investigation. This paper presents a rationale for the divorce of land cover from land use through the example of 'forest'.

2. BACKGROUND

2.1 Origins of the confusion of land cover and land use

The persistence and perpetuation of an inconsistent and counter-intuitive conceptual framework for measuring and monitoring land based resources can be seen in many national and international programmes (e.g. the IGBP Land Use and Cover Change programme, Nunes and Auge, 1999). The origins of this illogical paradigm lie in the 1970's when the availability of medium resolution satellite imagery coincided with the wish of governments to better manage their land resource for a range of objectives as exemplified by the most influential work in this area, Anderson et al (1976). Traditionally agencies concerned with tax, environmental management, planning, etc., had their own data specifications, data collection methodologies and classification schemes for recording land-based features. Anderson *et al*'s (1976) outline of the USGS Land Use and Land Cover Classification specified a hybrid land use and land cover classification. In developing a standard national remotely sensed land classification, the confusion of land cover and land use was driven by a number of factors:
- The need to accommodate the existing classifications of different agencies;
- The ability to machine process remotely sensed imagery (i.e. statistical discrimination of land features);
- The need for consistent information that could be compared across time, space and at different levels of aggregation;
- The need to accommodate differing agency interests;
- The need for a "resource oriented" classification to address the 95% of the national area not covered by previous "people-oriented" classification of the Standard Land Use Coding Manual (U.S. Urban Renewal Administration and the Bureau of Public Roads, 1965).

Many subsequent inventories and initiatives have copied the land classification confusion of Anderson *et al* (1976), developing hybrid classifications that confuse land use and land cover. Indeed the 'land cover/land use' couplet has become the *modus operandi* for many initiatives and most surveys where the differences between land cover and land use are frequently noted, but rarely accommodated. Land use and land cover are often used interchangeably in many studies, surveys, programmes of research and reports. Whilst land use dynamics are the major determinant of land cover changes, they are in essence very different things. The fundamental difference between land cover and land use is that the former describes the physical characteristics of the earth's surface and the latter describes the activities upon it. Their differences are described in the reports of many mapping projects that incorporate a hybrid classification. Despite widespread acknowledgement of the differences the two concepts continue to be intertwined.

2.2 Land Cover

Land cover is the physical material at the surface of the earth. Land covers include grass, asphalt, trees, bare ground, water, etc. There are two primary methods for capturing information on land cover: field survey and through analysis of remotely sensed imagery. Field survey involves the detailed recording of land cover features. Typically, surveyors record attributes of floristic and landscape features by annotating base maps (traditionally paper but increasing digital). The emphasis of field based land cover surveys is usually ecological (e.g. the Countryside Surveys in the UK – Barr et al., 1993; Haines-Young et al., 2000) capturing information on the distribution of plant species, vegetation communities and phyto-sociological associations. Field surveys are time consuming and labour intensive but do capture *data primitives* – that is, data on the fundamental component features (e.g. plant species) that constitute the basic units of land cover. The reported land cover classes are created by aggregating data primitives and reported in terms of the data primitive composition or aggregation, as determined by the imposition of such things as minimum mapping units. This paper is primarily concerned with the differentiation of land cover and land use features as recorded in remotely sensed imagery but cite the case of the nature of the information captured by field survey here as a contrast to that captured from remotely sensed data.

Land cover in remote sensing terms is the material which we see and which directly interacts with electromagnetic radiation and causes the level of reflected energy which determines the tone or the digital number at a location in an aerial photograph or satellite image. Tone or digital number alone may not be enough to distinguish between the different cover types, but remote sensing of land features supported by empirical investigation, that with measurement of tone in discreet wave bands different land covers are increasingly separable, although context, pattern and texture may also be used (Lillesand and Kiefer, 2000). The land cover classes that are discerned are clusters of pixels in the *N* wave bands that are within some defined statistical tolerance or distance in that feature space. Because of the nature of the information that remotely sensed land cover records and the way that the information is reported, the data primitives are not explicit. Yet land cover classes are described in not in terms of their primitives but by their ecological or use characters. The reasons for this are many but in part due to the fact that the spectral characteristics of many features of interest (e.g. woodlands, urban areas) are not consistent across different scenes, sensors, landscape contexts and spatial scales (Comber et al., 2004). Land cover is essential for environmental models (e.g. climatic and hydrologic), but is not directly useful for most policy and planning purposes (planning of the human or the natural environment), where land use is the relevant phenomenon.

2.3 Land Use

Land use is a description of how people *utilize* the land. Urban and agricultural land uses are two of the most commonly recognised high-level classes of use. Residential land, sports grounds, commercial areas etc. are also all land uses: land use describes socio-economic activity. Nunes and Auge (1999: p. 37) describe land use as involving "considerations of human behaviour, with particularly crucial roles played by decision makers, institutions, initial conditions of land cover".

The recording of land use and land use classifications have a number of characteristics that result in the concepts and measurements of land use being more contested than for land cover. First, the relationship between land use and land cover is complex and cannot be directly inferred, although it frequently is, as indicated by the quote from Nunes and Auge (1999) above. Fisher et al. (2005) noted that land cover and land use have complex *many-to-many* relationships and cited the example of the cover "Grass" which can occur in a number of different land uses: sports grounds, urban parks, residential land, pasture, etc. Also very few areas of homogenous land use have a single land cover. Second, that land use classifications do not fulfil the criteria of allocating features on the land surface into uniquely to one class: a single point in space may quite legitimately have a number of different land uses at any given moment. Much land has multiple states of use which may be simultaneous or alternate: the field with cows may be the village football pitch at weekends; the reservoir may provide flood control but also angling; and plantation forestry may also be used for several forms of recreation, including hunting and hiking, and even for grazing. The specification of any particular land use at any specific point in space is more problematic and contested because of these issues compared to land cover. For example, Hoeschele (2000) revealed serious differences in how land is used and regarded by indigenous commercial and subsistence farmers, on the one hand, and by forestry technocrats, on the other in the Attappadi district of India.

3. METHOD

In this work a text mining approach was applied to the various different types of forest use and cover descriptions. Generating information from text using automated computer techniques ("mining") is a complex process and despite extensive research including the development of sophisticated software (e.g. General Architecture for Text Engineering[1]) Natural Language Processing (NLP) remains very complex. The complexity arises because a word or term can have many meanings. Simple text mining is used by many internet search engines, which rank the documents found by relevance derived from the similarity with phrase entered by the user. Comparing multiple descriptions of forest land use and forest land cover is an extension to information retrieval and has been used to explore the semantic relations between different land cover datasets. Wadsworth et al. (2005) analysed the conceptual overlaps between different global land cover data. Wadsworth et al. (2006) applied computer characterisation to the textual descriptions of two UK land cover maps in order to be able to integrate them and found the integrative approach based on text mining to be more effective than human experts.

3.1 Data

The website "Definitions of forest, deforestation, afforestation, and reforestation" (Lund 2006) contains hundreds of different descriptions of forest activity and forest cover. These descriptions are organised into different definitional groupings of forest which were "based upon literal interpretations of the definitions" (Lund, 2006). The descriptions in the 'As a land use type' and 'As a land cover type' categories were extracted from the General, National and International groups for analysis (the state and provincial data were not analysed). These were cleaned to get rid of the references to source of the data descriptions.

[1] http://gate.ac.uk/

3.2 Initial processing

Each description was converted into a word list. Some words were gathered into terms or phrases, e.g. "25 m", "per cent". A matrix was constructed of descriptions (or classes) against words (or terms) used in the different forest cover and forest use descriptions, where the cells in the matrix contained the number of times each term appears in each class. The terms in the matrix were weighted using the "tf.idf" (total frequency x inverse document frequency) scheme (Robertson & Jones 1976):

$$W_{ij} = \frac{n_i}{\sum n_i} \ln \frac{D}{n_j} \quad (1)$$

Where W_{ij} is the weight of the i^{th} word in the j^{th} class
n_i is the number of times the word appears in the j^{th} class
Σn_i is the total length of the j^{th} class description.
D is the total number of classes
n_j is the number of classes containing the i^{th} word
The weighting has the effect that a word that appears in all class descriptions has a zero weight, but a word appearing frequently in a few short classes has a high weight.

3.3 Analysis of terms and identification of primitives

The significance of the terms in the weighted matrices was evaluated using a standard principal components analysis (PCA) technique based on a correlation matrix. For each of the 6 matrices relating to the different types of forest use and forest cover descriptions (general, national and international) the PCA identified:

- The number of components that explained the variation in class descriptions;
- The amount of variation explained by each component;
- The terms with the greatest loading for each component

For each of the 6 matrices, the PCA identified the components with eigenvalues greater than unity. Within each component the terms associated with highest component loadings were identified as those within 10% of the highest loading value. The weight of the component loading indicated relative strength of correlation to each principal component.

The six sets of descriptions were also grouped into two sets of use and cover. These were then independently processed in the same way.

4. RESULTS

The analyses presented in this section describe the terms in each component (for each type of use and cover description) with the highest loading i.e. they are within 10% of the maximum loading. The results describe the analysis of these *significant* terms. The numbers of principal components with eigenvalues greater than 1, terms with high loadings and the amount of variation in the weighted matrices explained by them are shown in Table 1 for each type of forest description.

4.1 Differences between forest *cover* and forest *use*

The significant terms identified during the PCA from the weighted matrices described in section 3.2 were placed into 3 groups related to the general nature of land use and land cover descriptions: activity and surface respectively. The object was to start to draw out the primitives associated with the concepts of forest use and cover. Each term was characterised as being

- 'Biological': those relating to vegetation, the environment and plants. This was expected to be more clearly associated with forest cover;
- 'Socio-economic': those relating to commercial activities, maintenance, and management. This was expected to be more clearly associated with forest use;
- 'Spatial / Structural': those relating to measurements specifications such as height, spatial extent and area as well as structural aspects such as crown closure. This was expected to be an important aspect within both sets of descriptions.

Other terms, such as prepositions, the verbs to be and to have etc., that could not be placed into the three categories were ignored. The distributions for the different categories of forest description are shown in Table 2. From the results in Table 3 the following statements can be made:

- Cover descriptions have a higher proportion of *Biological* terms than Use ones;
- *Socio-economic* are more frequent in Use descriptions;
- Use has a lower proportion of *Spatial / Structural* terms than cover does;
- National descriptions have generally fewer *Biological* terms.

		National	International	General
Use	Total terms	1261	484	511
	Descriptions	117	26	51
	Variation	91%	100%	100%
	Significant terms	152	64	124
Cover	Total terms	1179	390	810
	Descriptions	136	32	152
	Variation	85%	99%	77%
	Significant terms	106	75	78

Table 1. The original data, the number of terms and components with eigenvalues greater than 1 and significant terms for each type of forest description.

		Biological	Non-specific	Socio-economic	Spatial / structural
Use	National	8.6%	48.7%	25.7%	17.1%
	International	12.5%	28.1%	34.4%	25.0%
	General	12.9%	41.9%	25.8%	19.4%
	All	9.4%	46.3%	28.9%	15.4%
Cover	National	17.0%	42.5%	9.4%	31.1%
	International	25.3%	34.7%	6.7%	33.3%
	General	17.9%	42.3%	9.0%	30.8%
	All	20.0%	45.3%	6.5%	28.2%

Table 2. Characterisation of the significant terms for forest use and cover, major differences between use and cover in bold.

4.2 Primitives for forest *cover* and *use* data at different scales

The significant terms at between forest use and cover and at different scales (national, international and general) were ex-

plored. The aim was to reveal the nature of the concepts and terms that were unique to cover and use specific to different scales, as well as those that were shared. For each of the three characterisation groups (biological, socio-economic, spatial) 4 analyses were performed.
1) Terms at different scales (International, National and General), extracted using PCA, were compared at for land cover;
2) Terms at different scales were compared at for land use;
3) Land use and land cover were compared by looking at the overlap between the terms extracted at different scales;
4) Land use and land cover were compared by looking at the overlapping terms when all use and all cover descriptions were compared as two groups.
For the first two analyses the overlapping terms were visualised using social network software NetDraw[2] to create network diagrams. The results of the second two analyses, having only 2 elements, are displayed in tabular form.

4.2.1 Biological Terms

The only 'biological' terms that are shared by all the different scales of forest cover descriptions are 'ecosystem' and 'natural' (Figure1). There are no terms that are shared amongst all the different scales of forest use descriptions (see Figure 2).

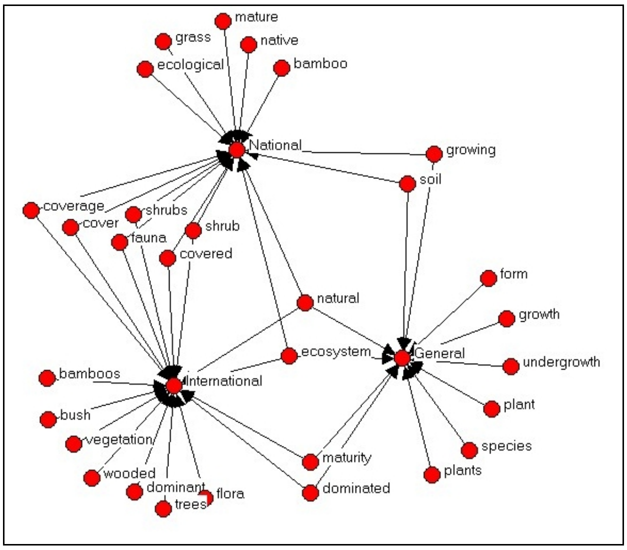

Figure 1. Forest Cover – key biological terms at different scales.

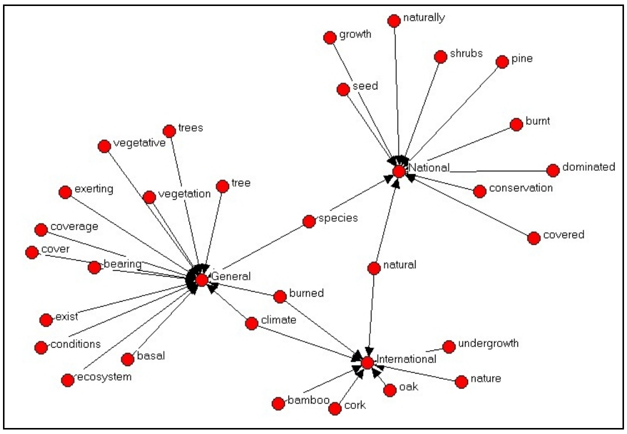

Figure 2. Forest Use – key biological terms at different scales.

Table 3 shows that when the descriptions are grouped into just two sets, then cover has many more unique significant biological terms.

Cover		Shared	Use	
bamboos	growing	bamboo	exerting	basal
maturity	mature	cover	exist	bearing
dominant	bush	coverage	naturally	burned
ecological	native	covered	nature	burnt
fauna	plant	dominated	oak	climate
flora	plants	ecosystem	pine	conditions
form	shrub	growth	seed	conservation
grass	soil	natural	tree	cork
	wooded	shrubs	vegetative	
		species		
		trees		
		undergrowth		
		vegetation		

Table 3. The significant biological terms unique to and shared between forest cover and forest use descriptions.

4.2.2 Socio-economic terms

There are no 'socio-economic' shared by all the different scales of forest use descriptions (i.e. International, National and General) – see Figure 3. 'Woodland' is the only term shared amongst all the different scales of forest cover descriptions (see Figure 4) and there are many fewer significant socio-economic terms associated with cover.

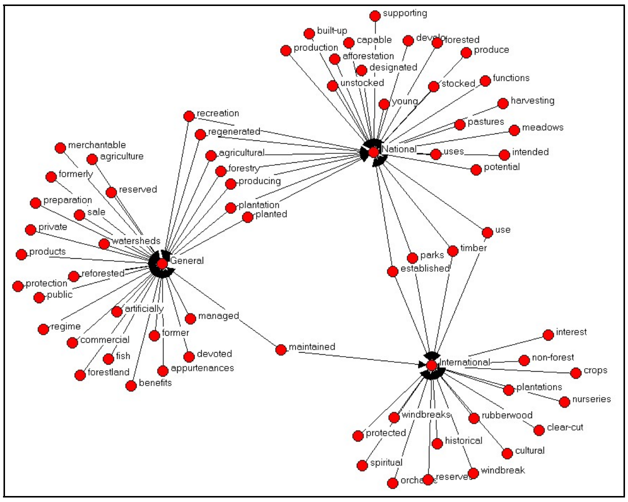

Figure 3. Forest Use: key socio-economic terms

[2] http://www.analytictech.com/

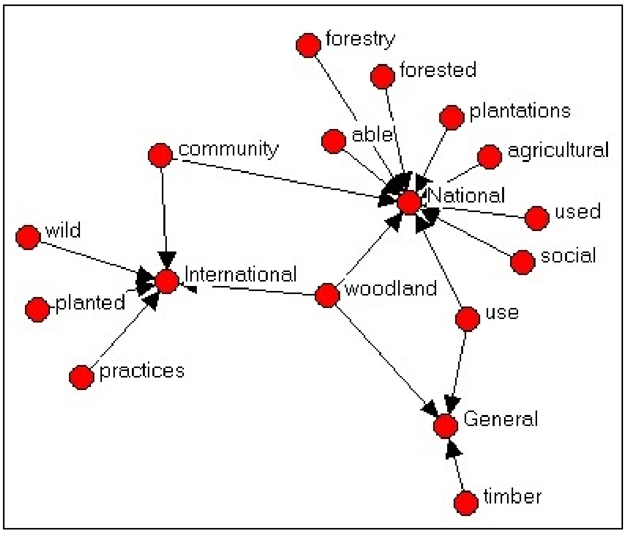

Figure 4. Forest Cover: key socio-economic terms.

Analysing all forest use and cover descriptions reveals the same pattern although with fewer significant terms (Table 4).

Cover	Use			
able	reserves	fish	pastures	developed
community	stocked	former	plantation	rubberwood
practices	young	formerly	potential	designated
social	benefits	functions	preparation	agriculture
used	built-up	uses	private	supporting
wild	capable	historical	produce	unstocked
woodland	clear-cut	intended	producing	watersheds
Shared	public	interest	production	windbreak
agricultural	crops	sale	products	windbreaks
forested	cultural	managed	protected	appurtenances
forestry	spiritual	meadows	protection	merchantable
plantations	reserved	regime	harvesting	maintained
planted		nurseries	recreation	afforestation
timber		orchards	reforested	commercial
use		parks	regenerated	established
		devoted	non-forest	artificially

Table 4. The significant socio-economic terms unique to and shared between forest cover and forest use descriptions

4.2.3 Spatial / Structural terms

The only significant 'spatial / structural' term that is shared at all the different scales of forest cover is 'Dense' (Figure 5). For forest use descriptions arte different scales, the only significant shared term is 'Primarily' (Figure 6).

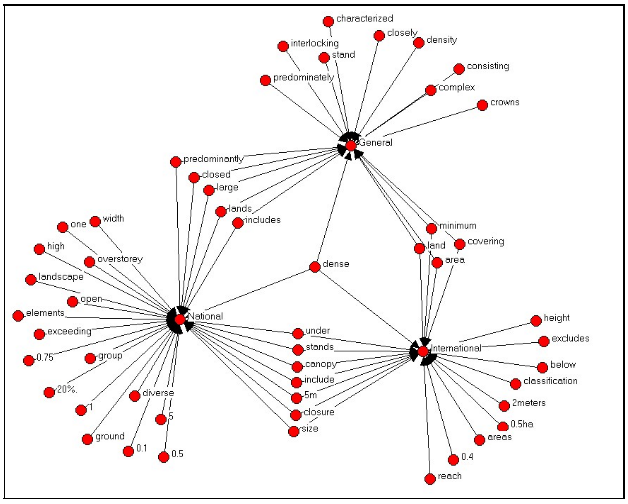

Figure 5. Forest Cover – key spatial / structural terms.

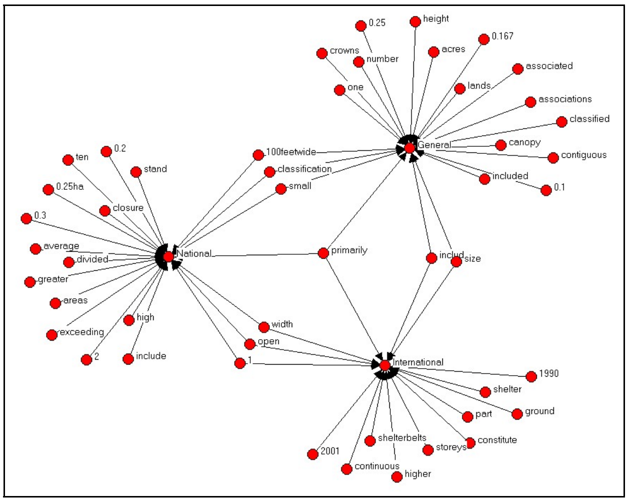

Figure 6. Forest Use - key spatial / structural terms.

Combining all the descriptions into two groups (use or cover) shows that cover descriptions have many more significant spatial / structural terms associated with its descriptions than use and that very few are shared (Table 5).

Cover		Shared	Use	
Closed	consisting	canopy		associated
land	covering	classification	0.3	associations
dense	large	includes	2	average
20%.	minimum	open	storeys	classified
2meters	characterized	size	acres	constitute
5m	closely	closure	small	continuous
area	complex	height	ten	divided
stands	density	include	0.167	greater
under	diverse	lands	0.2	higher
0.4	elements	width	0.25	included
0.5	excludes	exceeding	1990	number
0.75	group	ground	2001	contiguous
5	interlocking	high	0.25ha	100feetwide
0.5ha	landscape	stand	part	shelterbelts
below	predominantly	1	shelter	primarily

overstorey	one
predominately	0.1
	areas
	crowns

Table 5. The key spatial / structural terms unique to and shared between forest use and forest cover.

5. DISCUSSION

5.1 Data Primitives

The notion of data primitives is to identify the fundamental building blocks or foundations that underpin the concepts of the phenomenon under investigation, such as land use and land cover. Identifying data primitives – the underlying data concepts, what the data mean and represent – allows data to be better integrated into analyses alternative to the original purpose of the data. It facilitates better data re-aggregation, data re-use, sharing, and enables the uncertainties of data integration for specific analysis to be quantified.

In order to be able to effectively integrate datasets, data need to be consistent in terms of what they are reporting. Fisher et al. (2005) have described the internal data inconsistencies that may exist if concepts of land use and land cover are combined in Boolean classifications. Land use and land cover do not have a one to one relationship. Different covers may be subject to the same use and vice versa. Importantly, land uses may not be temporally consistent – alternative uses are possible for the same piece of land. The classification land use is described is much more open to contention (e.g. Heocshele, 2000).

Integration activities incorporating land data that confuse and combine the concepts of cover and use have to overcome the internal dataset inconsistency. This is problematic for models that incorporate land cover or land use data (e.g. evaluation of the impact of climate change, of the interaction between terrestrial and atmospheric environments, etc). For these reasons the IGBP have called for the explicit separation of the concepts of land use and land cover. For example, the Global Land Project science plan associates use with socio-economic systems and cover with biophysical systems (GLP, 2005).

5.2 Results and Method

The results show the association of land use with socio-economic aspects of land management and of land cover with biological. Interestingly there are few differences across different scales of use and cover descriptions, implying that the significant terms are not associated with international, national or general descriptions (*NB* ‘significant’ in this context means that the terms had loadings that were in the top 10% of the correlations between the terms and the principal components). Rather descriptions of forest are homogenous within different scales of use and cover description.

The results show the degree of confusion between land cover and land use descriptions and the difficulty of separating the concepts of forest use from forest cover as exemplified by analysis of the distribution of significant biological, socio-economic and spatial / structural terms. The biological terms that are shared are difficult to characterise as they are difficult to differentiate from use and cover which have overlapping management, context and structural descriptions. Anecdotally, the use terms could be said to relate more to biological functions, concepts and actions (verbs and adverbs) and cover terms to biological objects (nouns). Few of the socio-economic terms are significant for both use and cover and uniquely to use, although the shared terms do relate very strongly to aspects of management.

The results also show the current lack of primitives in land data. Whilst land cover or land use derived from remotely sensed imagery have spectral primitives - their position in spectral feature space – this is not how they are described. Their semantics describe their supposed characteristics on the ground – not the way that they were actually defined in the data. This is in comparison to land cover information generated from field survey whose primitives indicate the number and types of different plant species found in plant communities. Land use data derived from reflectance is an anachronism unless that use is consistent in terms of cover (which it is not – see Fisher et al., 2006). Instead land use of being related to unique positions in spectral feature space, land use can only be inferred from land cover due to the many to many relationships between use and cover.

Text mining with frequency and document size weighting has proved a useful tool in extracting the terms that contribute to the variation in class descriptions. However, no clear discernable pattern of primitives has been revealed by text mining of class descriptions – the significant terms are just words. This approach has been shown to be effective in separating differences in data semantics in many other applications (e.g. changes in soil classifications, land cover and in vegetation communities (Wadsworth et al., submitted). However, this simple approach, one required few assumptions to be made about the problem in hand, was unable to separate forest use and forest cover. One reason may have been the lack of text that was used in processing – Wadsworth et al. (2006) recommend at least 100 words for each class description. Another may be due to the way that the class descriptors were ordered in Lund (2007) and a further reason may be the genuine confusion over the concepts and descriptions of forest use and forest cover: they are so intertwined as to be inseparable. Of interest is that different sets of words are significant at different scales and very few words are co-associated as all scales.

6. CONCLUSIONS

Lund (2007) quite admirably has catalogued many classifications of forest into national, international and general use and cover groups and included the descriptions of forest as encapsulated by each dataset (with a reference). The hanging question from this analysis is whether these terms are truly data primitives or are they simply words? It is difficult to state that this work has identified *the* unique terms associated with cover and use, although subtle differences in flavour may (or may not) be discernable. The major finding of this work is the extent to which the confusion between land use and cover is embedded in so many land datasets, perpetuating the confusion between the two concepts perpetuate via their descriptions: the concepts of land use and land cover are mis-used everywhere.

It should be noted that there is no disagreement amongst practitioners that such a separation is desirable: land use ought to describe the activities on the earth’s surface and cover the surface itself. However, in practice these concepts are frequently or usually confused, not only within the same classification or database but also in the way that individual use and cover classes are described. This is in part due to the legacy of Anderson et al. (1976) which admitted its confusion of use and cover in order to satisfy and reach consensus amongst multiple agencies, and in part due to the nature of classifying remotely sensed imagery. Classification of remotely sensed imagery identifies areas that have similar statistical characteristics, as determined by their values in spectral space. This identifies areas of homogenous land cover. However, historically policy makers have been interested in activities and land use. The confusion between use and cover can therefore can also be seen to be data driven: the cheap, frequent, extensive and easy availability of satellite imagery has resulted in headlong rush for applications and the fudging of internal data consistency. Separation of the concepts of land use and land cover is needed to foster a culture of consistency in data recording in order to facilitate data integration and interoperability.

7. REFERENCES

Anderson, J.R., Hardy, E.E., Roach, J.T. and Witmer, R.E. 1976. *A Land Use and Land Cover Classification System for Use with Remote Sensor Data*. U.S. Geological Survey, Professional Paper 964, p 28, Reston, VA.

Barr CJ, Bunce RGH, Clarke RT, Fuller RM, Furze MT, Gillespie MK, Groom GB, Hallam CJ, Hornung M, Howard DC, Ness MJ 1993. *Countryside Survey 1990: main report. Countryside 1990 Series: Volume 2*. Department of the Environment, London.

Comber A.J.. Fisher P.F., Harvey, F., Gahegan., M and Wadsworth R.A, 2006. Using metadata to link uncertainty and data quality assessments. pp 279 – 292 in *Progress in Sptial Data Handling, Proceedings of SDH 2006*, eds. Andreas Riedl, Wolfgang Kainz, Gregory Elmes, Springer Berlin.

Comber, A., Fisher, P., Wadsworth, R., 2003 Actor Network Theory: a suitable framework to understand how land cover mapping projects develop? *Land Use Policy*, 20: 299–309.

Comber, A.J., Fisher, P.F. and Wadsworth, R.A., 2002. Creating Spatial Information: Commissioning the UK Land Cover Map 2000. pp. 351-362 in *Advances in Spatial Data*, eds. Dianne Richardson and Peter van Oosterom, Springer-Verlag, Berlin.

Comber, A.J., Fisher, P.F., Wadsworth, R.A., 2005a. What is land cover? *Environment and Planning B: Planning and Design*, 32:199-209.

Comber, A.J., Fisher, P.F., Wadsworth, R.A., 2005b. You know what land cover is but does anyone else?…an investigation into semantic and ontological confusion. *International Journal of Remote Sensing*, 26 1: 223-228

Comber, A.J., Law, A.N.R., Lishman, J.R., 2004. Application of knowledge for automated land cover change monitoring. *International Journal of Remote Sensing*, 2516: 3177-3192.

Fisher, P.F. 2003. Multimedia Reporting of the Results of Natural Resource Surveys, *Transactions in GIS*, 7 309-324.

Fisher, P.F., Comber, A.J., Wadsworth, R.A., 2005. Land use and Land cover: Contradiction or Complement. Pp. 85-98 in *Re-Presenting GIS*, eds. Peter Fisher, David Unwin, Wiley, Chichester.

GLP 2005. *Science Plan and Implementation Strategy*. IGBP Report No. 53/IHDP Report No. 19. IGBP Secretariat, Stockholm. 64pp.

Haines-Young RH, Barr CJ, Black HIJ, Briggs DJ, Bunce RGH, Clarke RT, Cooper A, Dawson FH, Firbank LG, Fuller RM, Furse MT, Gillespie MK, Hill R, Hornung M, Howard DC, McCann T, Morecroft MD, Petit S, Sier ARJ, Smart SM, Smith GM, Stott AP, Stuart RC, Watkins JW 2000 *Accounting for nature: assessing habitats in the UK countryside*, DETR, London

Hoeschele, W. 2000 Geographic Information Engineering and Social Ground Truth in Attappadi, Kerala State, India. *Annals of the Association of American Geographers,* 90 2, 293-321.

Lillesand, T.M., and Kiefer, R.W., 2000. *Remote Sensing and Image Interpretation*. 4th Edition. Wiley and Sons, Chichester.

Lund, H. Gyde coord. 2006. *Definitions of Forest, Deforestation, Afforestation, and Reforestation.* [Online] Gainesville, VA: Forest Information Services. Available from the World Wide Web: http://home.comcast.net/~gyde/DEFpaper.htm. Misc. pagination.

Nunes, C. & Auge, J.I. 1999 *International geosphere-biosphere programme: a study of global change of the International Council of Scientific Unions*, IGBP, Stockholm.

Robertson, S.E., Spärck Jones, K., 1976, Relevance weighting of search terms, *Journal of the American Society for Information Science*, 273, 129-46.

U.S. Urban Renewal Administration and the Bureau of Public Roads, 1965

Wadsworth R.A, Comber A.J., & Fisher P.F., 2006. Expert knowledge and embedded knowledge: or why long rambling class descriptions are useful. pp 197 – 213 in *Progress in Sptial Data Handling, Proceedings of SDH 2006*, eds. Andreas Riedl, Wolfgang Kainz, Gregory Elmes, Springer Berlin.

Wadsworth R.A., Fisher P.F., Comber A., George C., Gerard F. & Baltzer H. 2005. Use of Quantified Conceptual Overlaps to Reconcile Inconsistent Data Sets. Session 13 Conceptual and cognitive representation. *Proceedings of GIS Planet 2005*, Estoril Portugal 30th May - 2nd June 2005. ISBN 972-97367-5-8. 13pp