A Probabilistic Approach to Environmental Change Detection with Area-Class Map Data

Christopher B. Jones[1], J. Mark Ware[1], and David R. Miller[2]

[1] School of Computing, University of Glamorgan, Pontypridd, CF37 1DL, United Kingdom
{CBJones, JMWare}@glam.ac.uk
[2] Macaulay Land Use Research Institute, Aberdeen, United Kingdom
mi008@mluri.sari.ac.uk

Abstract. One of the primary methods of studying change in the natural and man-made environment is that of comparison of multi-date maps and images of the earth's surface. Such comparisons are subject to error from a variety of sources including uncertainty in surveyed location, registration of map overlays, classification of land cover, application of the classification system and variation in degree of generalisation. Existing geographical information systems may be criticised for a lack of adequate facilities for evaluating errors arising from automated change detection. This paper presents methods for change detection using polygon area-class maps in which the reliability of the result is assessed using Bayesian multivariate and univariate statistics. The method involves conflation of overlaid vector maps using a maximum likelihood approach to govern decisions on boundary matching, based on a variety of metrics of geometric and semantic similarity. The probabilities of change in the resulting map regions are then determined for each class of change based on training data and associated knowledge of prior probabilities of transitions between particular types of land cover.

1 Introduction

One of the major applications of geographically-referenced spatial data is that of monitoring change on the earth's surface. Change in natural and semi-natural environmental phenomena may reflect variation in climate, the dynamics of plant and animal communities and the influence of human activity in land and transport development, agriculture and the production of pollutants. Approaches to automated change detection based on existing spatial data sources differ according to the type of data that are used. A considerable volume of literature documents the application of satellite remote sensing, in which the evidence for change is based on comparison of either the radiometric pixel values and their combinations, such as in vegetation indices, or the classified pixels resulting from various methods of interpretation of the source data [1]. An alternative approach is based on the comparison of area-class polygon maps produced from interpretation of sources such as aerial photography and land surveys [2].

The results of change detection analyses based on comparisons of multi-date digital maps are subject to error from several sources. All methods may incur error in the surveyed location of spatial data elements and error in the locational registration of the respective datasets. In the case of classified data, errors arise due to variation in the definition of and the application of the classification systems and in the degree of generalisation. Comparison of radiometric data values, or their combinations or transformations, is also affected by factors such as variation in lighting conditions and in seasonal characteristics of vegetation and agriculture.

Studies of the reliability of change detection based on comparison of spatial datasets have been performed with regard to a range of methodologies, with the most extensive research being in the field of satellite remote sensing. Less effort has been expended on the development of automated change detection with area-class maps, whether derived from aerial photography, land survey, or indeed satellite image interpretation, yet they constitute a massive source of historical data on various aspects of the environment. Several studies have highlighted the problems of error in this context (e.g.[3], [4], [5], [6]). However there are many examples of automated change detection using either area-class maps or satellite imagery, in which reliability of the results is not reported (e.g. [7], [8], [9]), a situation which may be regarded as reflecting the lack of appropriate and readily implemented methods. Most geographical information systems do provide facilities that can be used to perform comparisons of polygon maps for purposes of change detection, but there is a shortage of effective tools for determining the nature of change. In particular it is difficult to evaluate the reliability of the results with regard to the sources of data on which they are based.

In this paper we address specifically the use of area-class polygon maps for automated change detection and present methods that are intended to assist in improving reliability while providing quantitative measures of the accuracy of the results. The approach adopted is based on the application of Bayesian statistical methods [10]. It is assumed that a variety of geometric and semantic measures associated with elements of the source data can be used, in combination with prior knowledge of the occurrence of transitions between different land-cover categories, to indicate that multi-date maps represent specific types of change. Training data, based on expert assertions of the presence and absence of change, are used to estimate the prior conditional density functions for each type of evidence. The training data used in this study relate to change in map boundaries and in land cover categories derived from aerial photographic surveys of Scotland. We present results of evaluating the method using expert-asserted ground truth data separate from that used in training.

In the remainder of the paper we start by reviewing existing methods of polygon map comparison and discuss the relevance of conflation methods in Section 2, before summarising the procedures employed here in Section 3. Section 4 describes the experimental results, and a conclusion is provided in Section 5.

2 Map Overlay Methods for Change Detection

Given a pair of polygon maps relating to different dates, a simple approach to determining differences between them, and hence potential regions of change, is to overlay the maps and analyse the resulting regions of intersection. Differences in classifications derived from the parent maps for a particular polygon indicate that the region may be categorised as representing change. If the source maps contain independently generated boundaries that are approximately coincident then sliver polygons can be expected to occur and the problem arises of distinguishing them from polygons that may represent genuine change [3]. Conventional GIS functionality allows the slivers to be removed if they are smaller than some areal threshold. However, the problem then arises of selecting an appropriate threshold value and in any event the result may be an arbitrary increase in the area of the adjacent polygon with which it is merged.

Chrisman and Lester [3] pointed out that sliver area by itself is not a reliable measure and proposed distinguishing between change and non-change slivers on the basis of their shape. This still does not distinguish the case where, though a sliver may be relatively small, it may include a separation between its opposite sides that exceeds a locational error distance threshold. Cherrill and McClean [4] suggested confining measurements of change to the internal zones of polygons that exclude a boundary error buffer, and they presented results of applying this method using a range of buffer sizes. The results reduced the incidence of what they considered to be erroneous change, though the technique leads to the exclusion from change analysis of what could be a significant proportion of the map area. It should be remarked however that Cherrill and McClean observed in their study that boundary location error was a much less important factor than that of area class differences due to disagreement between interpretations rather than to ‘real world’ change.

An alternative approach to resolving the problem of slivers in conventional GIS is to increase the size of a ‘fuzzy tolerance’ value that is used to control the merger of line vertices that belong to adjacent line features. A technique for performing this merger has been described by [11], but the method appears to merge vertices in a manner which, though distance-constrained, may result in an arbitrary location for the merged lines within the limits of the distance constraint. For maps containing boundaries with errors of different magnitudes, this could result in the degradation of better quality boundaries that may be moved in the course of the merge procedure. As suggested by Chrisman [12], when two boundary segments are considered to be equivalent, it may be preferable simply to delete the poorer quality line. A method of map overlay in which differing locational errors of the map components are taken into account has been described in [13] and [14]. A shortcoming of this and related techniques, such as [15] and [16], is that they adopt a very localised, largely point-based approach which fails to take account of higher level geometric and semantic characteristics of the parent linear features. This issue is addressed in the feature matching procedures developed in the work described here.

2.1 Map Conflation

The idea of overlaying a pair of maps, such that the better quality versions of equivalent parts of the maps are retained, is integral to the concept of map conflation (e.g. [17], [18]). An important part of conflation is the matching of equivalent features. When a lower quality representation is merged with a better quality equivalent, it is assumed that it brings with it other topologically linked features for which there may be no match between the maps. The result is a map in which the best quality elements of the sources are preserved while non-common elements are also retained in a topologically consistent manner.

As originally conceived, map conflation was intended to be applied to multiple representations of the same phenomena relating to a particular point in time. The concept is extended here to combine maps relating to different points in time. Thus where linear features or boundaries cannot be distinguished from each other, the resulting map should include the best known representation of the feature which may be the result of a merge process. Recent work on conflation [19] addresses the problems of data uncertainty with a rule-based approach using geometric and attribute matching procedures. Our approach is related, but differs in being concerned with change and in the explicit use of training data to analyse the statistical characteristics of change.

3 Probabilistic Change Detection Procedure

In the present study there are two types of event, concerning a pair of maps relating to different times, that can be evaluated probabilistically with Bayesian statistics. These are:

- sections of area boundaries being equivalent,
- land cover changing from one category to another.

The major functional stages in the procedure for change detection are as follows:

- training to determine the prior conditional probability distributions of evidence associated with the presence of boundary section equivalence and land-cover category change,
- overlay the two maps under consideration and identify paired sections of boundary that are indistinguishable at some probability level,
- merge matched line features, if a conflation is required,
- determine the probability that land-cover change has occurred for map regions in which the source land cover attributes differ.

3.1 Training

The training stage requires that an expert asserts the presence of each of the two types of event within training data that are representative of the context of change detection. Associated metrics of potentially relevant evidence are then analysed to estimate their conditional probability distribution. We now summarise metrics that have been considered in this study. It is important to stress however that the approach is generic and there are many potential items of evidence that might prove to be effective.

Evidence for Boundary Equivalence (Hypothesis H_1). Given an overlay of a pair of time-differing maps of the same area, an expert must identify nearby sections of boundary that represent the same (unchanged) boundary phenomenon on the ground. These pairs of asserted equivalent line portions are then analysed to evaluate the following metrics:

E1: Hausdorff distance between line sections a and b. This is a measure of the maximum lateral separation of *a* and *b*. If d_{ai} is the shortest distance from edge *i* of line *a* to its nearest neighbouring edge in *b* (taking account of all locations on each of the edges), and d_{bj} is the distance from edge *j* of line *b* to its nearest neighbour in *a*, then

$$E1 = \max(\max(d_{ai}), \max(d_{bj})) .$$

E2: approximate average distance between line sections consisting of m *and* n *vertices, given by*

$$E2 = \frac{1}{m+n} \Sigma d_{ai} + \Sigma d_{bj} .$$

E3: difference in line length.

E4: difference in alternation (i.e. intersections with 'anchor line' connecting start and end).

E5: difference in trend orientation as determined by the 'anchor line'.

E6: difference between the pair of left-right attribute codes of each section of boundary. The difference between a pair of attribute codes can be measured as the weighted path length between the two codes within the hierarchical classification system to which they belong. If P is the set of edges e in the shortest path and each edge has a weight determined by a function **W** of the depth in the hierarchy of the node at the lower (most distant from the root) end of the edge, then combining the weighted left P_l and right P_r path distances

$$E6 = \sum_{e \in P_l} W(e) + \sum_{e \in P_r} W(e) .$$

Evidence for Change in Land Cover Type (Hypothesis H_2). The following items of evidence were considered in an effort to support the hypothesis of change in a single polygon for one type of land cover X to another type Y:

E7: (0 or 1) the attribute from map A *is type* X*, while the attribute from map* B *is type* Y.

E8: (0 or 1) there is a polygon of type Y *adjacent to the polygon of type* X *in map* A.

Conditional Probability Density Functions. For each of these items of evidence, the mean is determined and stored in the vector M and a covariance matrix C is constructed using the items of evidence (such as E1 to E6). Provided that the individual items of evidence are approximately normally distributed, or can be transformed to normality, the conditional probability density function [10] for a set of items of evidence E is represented by:

$$p(E|H) = \frac{1}{(2\pi)^{n/2}\,|C|^{1/2}} \exp[-\tfrac{1}{2}(E - M)^T C^{-1}(E - M)] .$$

In the event of only a single item of evidence being used, then the conditional probability may be estimated directly from the training data.

3.2 Posterior Probabilities and Prior Probabilities

Given a set of items of evidence E it will then be possible to estimate the posterior probability of an event using the conditional probability density function with Bayes' rule:

$$p(H|E) = \frac{p(E|H)\,p(H)}{p(E)}$$

The prior probability p(E) can be derived in this case from the summation of the product $(p(E|H_i).p(H_i))$ for the hypothesis and its negation. The prior probability of a hypothesis p(H) should be based on expert knowledge that is relevant to the context. In the absence of such knowledge it may be approximated on the basis of the training data [20]. Thus $p(H_1)$ is estimated by the frequency with which an expert has asserted that a neighbouring line within a predetermined maximum tolerance distance of the source line is in fact equivalent to the source line. $p(H_2)$ may be found from analysis of land-cover transition matrices based on expert assertions of change between one land use type and another. An example of such a study that is relevant to the research described here is that of Hester *et al* [21]. In the case of types of transition for which no such prior knowledge is available it can be approximated from the training data by the proportion of the area of type X polygons that become type Y polygons, as verified by the expert assertions.

3.3 Locational Matching (Via Triangulation)

In order to identify candidate sections of pairs of lines that may prove to be regarded as indistinguishable, a search is performed to find for each edge of each line segment all edges of other line segments that are within a given tolerance distance, based on the locational accuracy of the relevant lines. The results of these searches are used to construct pairs of line sections that are mutually within the tolerance distance of each other, *i.e.* there is no part of one line that is further (when measured by shortest distance) than the tolerance distance from the other line section. Each section of a line segment is a sequence of component edges. However, because one part of an edge could be within tolerance and the other part out of tolerance, the source edges are split where such ambiguities arise. The spatial search for neighbouring edges is performed using a main-memory constrained Delaunay triangulation of the vertices of the linear features. Details of this type of procedure can be found in [22] and in [23].

For each line section there may then be several other line sections with which it might be regarded as equivalent. The posterior probabilities of equivalence with all candidates are estimated as summarised above. If a probability exceeds a specified threshold then the pair of line sections with the highest probability (maximum likelihood) is selected as equivalent.

3.4 Merging of Line Sections Deemed Equivalent for a Given Probability Level

Merging of equivalent line sections is carried out in a way that will give preference to one of the line sections if it is known to be of higher accuracy than the other line section. This could be the case if the source maps had been produced at different nominal scales or if the procedure for producing one of the maps was known to be more reliable than that for the other.

Assuming that two line sections a and b are to be merged, then the merged line section c is found by a weighted averaging of a and b where the weighting is governed by the relative accuracy of a and b. There are three steps in the merging process. First, vertices are added to a and b such that: (i) a and b have the same number of vertices; and (ii) vertices appear at proportionally equivalent distances along a and b. Each vertex in a now has a corresponding vertex in b. Next, c is found by calculating the weighted average of corresponding vertices in a and b (*i.e.* $v_c = Wv_a + [1-W]v_b$, where $0<=W<=1$). If one line section is of a much higher accuracy than the other, the averaging procedure may be replaced by a procedure which simply selects the more accurate. As a result of the weighted averaging there will be sharp breaks between the merged and the adjacent unmerged line sections. The final step involves applying a local progressive weighting to provide a smooth transition (see Fig. 4 and Fig. 5). More details can be found in [24].

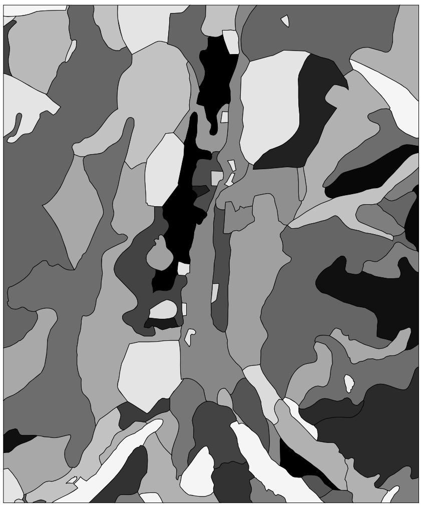

Fig. 1. 1946 test data

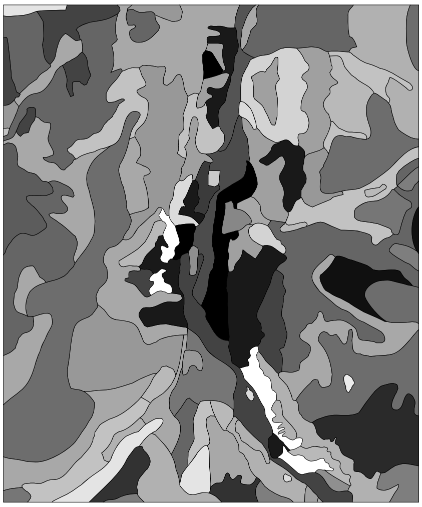

Fig. 2. 1988 test data

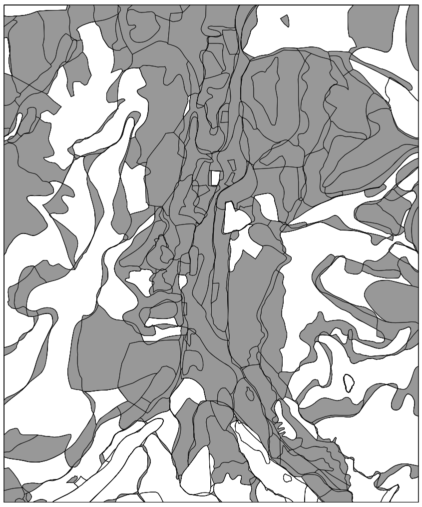

Fig. 3. 1946/1988 change map based only on feature attribute difference

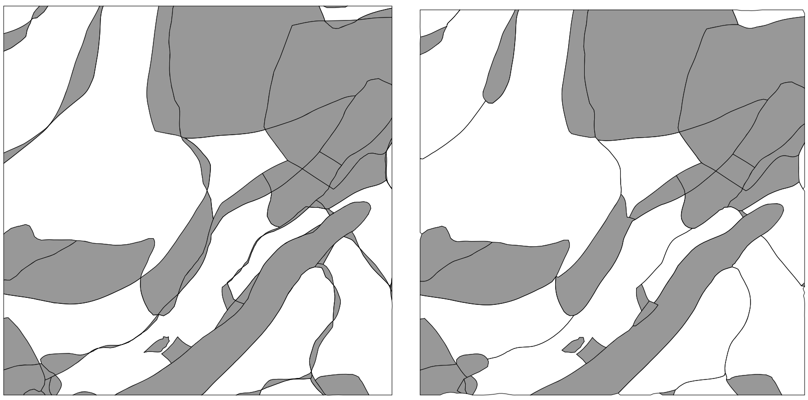

Fig. 4. Part of the 1946/1988 change map before (left) and after (right) boundary merging

3.5 Determine Change Probabilities for Each Region of the Map

The final stage of the change detection process is that of attaching posterior probabilities of land cover change to polygons resulting from the combination of the source maps. This stage may be performed on the conflated map or on the 'raw' intersection map.

Taking Account of Differences in Generalisation of Land Cover Classification. When comparing land cover maps in which the classification systems differ, or in which the nature of application of a classification system differs, care must be taken to identify differences in land cover category that are attributable only to differences in the degree of generalisation of the categories. Thus if one category is a generalisation of the other, then they cannot be treated as evidence of change. The semantic closeness metric for a single 'path' used in E6 may be appropriate to a consistently applied individual classification, but it provides a finite difference between a class and its immediate superclass. An alternative closeness measure E9 (based on that given in [25]), determines the number of non-common superclasses of a pair of categories, so that a parent-child relation gives a distance of zero. A third metric E10, which is potentially applicable to diverse classification schemes, determines the semantic distance based on the difference between the set of terms used to define a category [26]. Metrics E9 and E10 have been implemented as part of the current research project and enable candidate change polygons to be screened out of the result of the probabilistic change detection procedure.

Visualisation of the Results. Several methods have been used to present the results of the change detection procedures. At the simplest level the source maps and their intersected overlay can be displayed (Figs. 1-4). On a change map, those polygons that represent change with a confidence exceeding a specified probability may be highlighted. Alternatively the variation in confidence between different change polygons can be represented by a variable range of colour lightness and saturation (not illustrated here). A further refinement of the latter is to display the variable confidence, via lightness and saturation, of polygons that represent selected types of change. The degree of confidence with which area boundaries are merged can be illustrated by modifying the thickness of the corresponding boundary line symbol, as in Fig. 5.

4 Experimental Results

The functionality described in Section 3 has been implemented in the C programming language in combination with the Avenue scripting language of the ArcView GIS. The change detection functions can all be called from a modified version of the conventional ArcView user interface.

In order to evaluate the method, land cover maps and associated training data for 1946, and 1988 (Fig. 1 and Fig. 2) were obtained for the same 30km^2 region (at UK National Grid origin 282000E 790000N) in the Cairngorms from the Macaulay Land

Use Research Institute (MLURI). A requirement was that the maps be generated independently which was not initially the case. This resulted in a reinterpretation (by MLURI staff) of the selected 'squares' of the earlier map using the original aerial photography.

The newly generated maps were then overlaid in pairs and, for each boundary section on each map, a set of candidate equivalent boundaries was generated automatically. The expert mapper was asked to indicate which pairs of boundary sections were regarded as logically equivalent in the sense that they could not be distinguished from each other on the basis of location (though there could be an associated attribute change).

In addition the expert was asked to label those intersection polygons which corresponded to genuine (‘real world’) change in land cover, as well as those for which change was assumed to be unlikely. A distinction was made at this stage between changes involving human activity (such as plantation and felling) and those regarded as natural or semi-natural. The map data were then partitioned into training data and evaluative data and the parameters of the conditional probability density functions were estimated. The analyses referred to subsequently relate to comparison of the 1946 and 1988 maps.

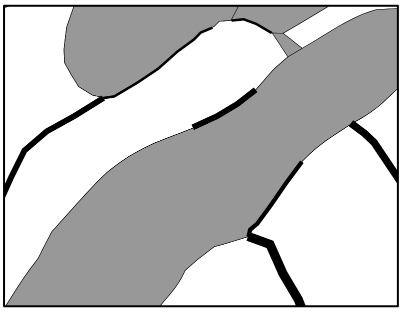

Fig. 5. Part of the 1946/1988 conflated change map. Line thickness relates to confidence in conflated lines (thicker lines represent lower confidence)

4.1 Boundary Merging

In a comparison of the 1946 and 1988 datasets, a buffering procedure was used to create a set of 294 candidate matching pairs of boundary segments. This set was divided into a training set of 150 with the remainder to be used as an evaluation dataset. In the training set, Kolmogorov-Smirnov tests for normality on the relevant items of evidence revealed that E3 was normally distributed. E1 and E2 required transformation to normality using the Box-Cox power transformation. E4 and E5 were found to be indistinguishable from random. However E5 was used at one level to filter out pairs of lines that exceeded 45 degree difference, as no such pairs were found to be equivalent in training. Preliminary observations indicate that E6 was not normally distributed and not a significant discriminator in this experiment. It may be envisaged however that it or a similar metric is likely to be of value in preventing merger of conceptually different types of boundary.

In the evaluation dataset, 64 out of 126 candidate matches were known to be matches on the basis of expert assertion. Of these, 55 were selected correctly as matches at the 50% confidence level, with 21 matches being identified incorrectly, giving 72% correct identifications. The remaining 9 expert asserted matches were recognised with a confidence level of less than 50%. At the 90% confidence level, 88% of identified matches were correct based on 30 correct matches (4 incorrect). The size of the samples at intermediate levels is too small to warrant reporting. However, the 90% confidence level results may be regarded as reasonable estimates of the reliability of the boundary matching results. These results were obtained when using the items of evidence E1, E2 and E3. It was found that lower success rates obtained when only using either E1, or E1 and E2.

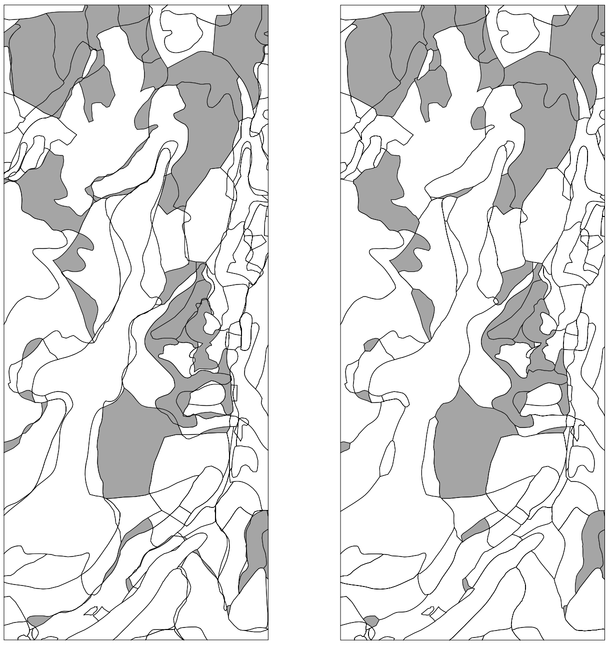

Fig. 6. Unconflated (left) and conflated (right) maps showing change between heather types at the 56% confidence level. Boundaries in the right hand map have been merged at the 70% confidence level

4.2 Probability of Land Cover Change

A problem arose with determining conditional probability density functions for categorical change in that the large number of MLURI attribute codes (including mosaic codes) resulted in only a few specific transitions occurring sufficiently frequently to generate significant statistics. In order to generate more robust statistics, a subset consisting of the more frequent occurrences of the 49 MLURI codes was selected for experimental purposes. Theoretical considerations indicate that metric E8 would need to be used selectively for particular types of change. This item of evidence is that prior to change the change unit had a neighbour of the same type that it changed to. This is expected to be of most relevance to evaluation of change in natural and semi-natural vegetation. Preliminary results have not demonstrated that it is a very sensitive indicator and hence here we confine the probability analysis to the use of evidence E7 alone (i.e. that the feature attribute codes indicate a specific type of change). As is to be expected the results vary considerably according to the type of change.

In the case of change from heather to coniferous woodland, training indicated that following boundary matching all feature codes indicating change to coniferous woodland corresponded to all expert assertions of such change. As a consequence the posterior probability (confidence level) was found to be 100% and the evaluation data confirmed this level.

In contrast, for example, training indicated that feature codes showing change from one type of heather to another (Fig. 6) are prone to a high level of uncertainty, giving a posterior probability of 56%. In the corresponding evaluation data only 47% of polygons exhibiting change between heather types represented actual change. This discrepancy may reflect the relatively small sample sizes that were used. Thus of an initial 157 polygons indicating change between heather types, only 68 remained following boundary merging. These were split into a training set and an evaluation set, each of size 34.

5. Conclusions

This project has addressed the problem of evaluating and improving the reliability of automated environmental change detection using area-class polygon maps. Bayesian statistics have been applied to estimate the probability of equivalence of boundary segments and the probability of land-cover change within polygons resulting from overlaying the maps. The approach enables the possibility of merging boundary sections to create a conflated map representing the best quality geometric elements of the source maps. A variety of geometric and semantic evidence for change may be considered, and the methods lead to the possibility of improving the reliability of the results by increasing the amount of training data used to estimate conditional probability density functions. Preliminary results using training data from land cover maps of Scotland found a reasonably close match between the generated probability levels and those of expert assertions. It is anticipated that larger training datasets and refinement of the choice of metrics could improve the robustness of this approach.

There is clearly scope for further research to build on the results of this project. This includes evaluating alternative geometric and semantic difference metrics and investigating automatic adoption of appropriate metrics and constraints depending upon the source data types. There is also potential for developing more versatile, hybrid methods to combine conventional quantitative multivariate statistics with qualitative sources of evidence. Thus certain types of land cover change may be associated with particular spatial relationships with neighbouring features, or with physical characteristics of terrain and climate (introducing an analogy with contextual methods in remote sensing such as those of [27] and [28].

Acknowledgement

This research was funded by Natural Environment Research Council grant GR3/10569.

References

1. Singh A. (1989) 'Digital change detection techniques using remotely sensed data'. *International Journal of Remote Sensing* 10: 989-1003.
2. Kirby R.P. (1992) The 1987-1989 'Scottish national aerial photographic initiatives', *Photogrammetric Record* 14(80): 187-200.
3. Chrisman N.R. and M. Lester (1991) 'A diagnostic test for error in categorical maps'. *Proc Auto-Carto 10*, Baltimore, MD, ACSM/ASPRS, pp 330-348.
4. Cherrill A. and C. McClean (1995) 'An investigation of uncertainty in field habitat mapping and the implication for detecting land cover change'. *Landscape Ecology* 10(1): 5-21.
5. Aspinall R.J. and D.M. Pearson (1995) 'Desribing and managing uncertainty of categorical maps in GIS'. In Fisher P. (ed) *Innovations in GIS* 2, Taylor and Francis, pp 71-83.
6. Edwards G. and K.E. Lowell (1996) ‘Modeling uncertainty in photointerpreted boundaries’. *Photogrammetric Engineering and Remote Sensing* 62(4): 377-391.
7. Schmidt M., H.E. Schreier and P.B. Shah (1995) ‘A GIS evaluation of land use dynamics and forest soil fertility in a watershed in Nepal’. *International Journal of Geographical Information Systems* 9(3): 317-327.
8. Peccol E., A.C. Bird and T.R. Brewer (1996) ‘GIS as a tool for assessing the influence of countryside designations and planning policies on landscape change’. *Journal of Environmental Management* 47: 355-367.
9. Green K., D. Kempka and L. Lackey (1994) ‘Using remote sensing to detect and monitor land-cover and land-use change’. *Photogrammetric Engineering and Remote Sensing* 60(10): 1243-1251.
10. Johnson A.J. and D.W. Wichern (1998) *Applied Multivariate Statistical Analysis*. Prentice Hall.
11. Chrisman N.R., J.A Dougenik. and D. White (1992) ‘Lessons for the design of polygon overlay processing from the Odyssey Whirlpool algorithm’, *Proc. 5th Int. Symp. on Spatial Data Handling*, 2, International Geographical Union, pp 401-410.
12. Chrisman N.R. (1987) 'The accuracy of map overlays: a reassessment', *Landscape & Urban Planning* 14: 427-9.
13. Harvey F. 1994, ‘Defining unmovable nodes/segments as part of vector overlay: The alignment problem’, *Proceedings 6th Int. Symp. on Spatial Data Handling*, 1, International Geographical Union, pp 159-176.
14. Harvey F. and F. Vauglin (1996) 'Geometric match processing: applying multiple tolerances', *Proceedings 7th Int. Symp. on Spatial Data Handling*, 1, pp 4A.13-29.
15. Pullar D. (1991) ‘Spatial overlay with inexact numerical data’, *Proceedings AUTOCARTO 10*, ACSM-ASPRS, Falls Church, Va., pp 313-329.
16. Zhang G. and J. Tulip (1990) ‘An algorithm for the avoidance of sliver polygons and clusters of points in spatial overlay, *Proceedings 4th Int. Symp. on Spatial Data Handling*, *1*, International Geographical Union, pp 141-150.
17. Lupien A.E. and W.H. Moreland (1987) 'A general approach to map compilation', *AUTOCARTO 8*, ACSM-ASPRS, Falls Church, Va., pp 630-639.

18. Saalfield A. (1988) 'Conflation: automated map compilation'. *International Journal of Geographical Information Systems* 2(3): 217-228.
19. Cobb M.A., M.J. Chung H. Foley, F.E. Petry and K.B. Shaw (1998) 'A rule-based approach for the conflation of attributed vector data'. *GeoInformatica* 2(1): 7-35.
20. Sonka M., V.Hlavac and R. Boyle (1993*) Image Processing, Analysis and Machine Vision.* International Thomson Computer Press, London
21. Hester A.J., D.R. Miller and W. Towers (1996) ‘Landscape-scale vegetation change in the Cairngorms, Scotland, 1946-1988: implications for land management’. *Biological Conservation* 77: 41-51.
22. Jones C.B. and J.M. Ware (1998) 'Proximity search with a triangulated spatial model'. *Computer Journal* 41(2): 71-83.
23. Jones C.B., J.M. Ware and C.D. Eynon (1999) 'Triangulated spatial models and neighbourhood search: an experimental comparison with quadtrees', *The Visual Computer* 15(5): 235-248.
24. Ware J.M. and C.B. Jones (1998), 'Matching and aligning features in overlayed coverages', *Proceedings of 6th International Symposium on Advances in Geographical Information Systems (ACM-GIS’98)*, pp 28-33.
25. Sintichakis M. and P. Constantopoulos (1997) 'A method for monolingual thesauri merging'. *Proceedings 20th Annual ACM SIGIR Conference on Research and Development in Information Retrieval*, Philadephia, USA, pp 129-138.
26. Tversky A. (1977) ‘Features of similarity’. *Psychological Review* 84(4): 327-352.
27. Barr S. and M. Barnsley (1997) 'A region-based, graph-theoretic data model for the inference of second-order thematic information from remotely-sensed images'. *International Journal of Geographical Information Science* 11(6): 555-576.
28. Kontoes C., G.G. Wilkinson *et al* (1993). 'An experimental system for the integration of GIS data in knowledge-based image analysis for remote sensing of agriculture. *International Journal of Geographical Information Systems* 7(3): 247-262.