1996 MULD (Reprint)

GIS 94

GIS'94 Symposium

Vancouver
British Columbia
February 1994

Knowledge based ecosystem prediction: Field testing and validation

Jan A. Mulder
Alberta Research Council
6815 8th Street NE
Calgary, AB
T2E 7H7
E-mail: mulder@arc.ab.ca

Ian G.W. Corns
Canadian Forest Service
5320 122nd Street
Edmonton, AB
T6H 3S5
E-Mail: icorns@nofc.forestry.ca

Abstract

The Alberta Research Council's Naia program is concerned with the design and implementation of an ecologically- oriented spatial and knowledge based framework to support forest and land resource management. A decision support system with the capability of representing the knowledge used by a forest ecologist to infer a forest ecosystem from a variety of data sources is under continuing development. The system has been designed as a classification shell with the capability of representing uncertainty in a hierarchically structured knowledge base. The classification process implements a combination of symbolic and evidential reasoning and it predicts ecosystems from topography, forest cover, and soil maps. The shell operates in conjunction with a GIS. The system has been tested with different classification systems. It has a high prediction accuracy when tested against the ground truth data used to build the classification system (85 - 94%). Field tests of the system have thus far provided interesting and promising results. When applied to a particular forest management area the system requires a refinement of its knowledge base in terms of adjustment in the mass functions by means of which evidence for different ecosystems is updated. In addition, it is sometimes necessary to further refine the ecosystem classification system itself. The shell like design of the system makes it easy to accommodate such adaptations. After a knowledge base refinement the system's performance is very solid. A more surprising result is the system's ability to expose the lack of data integrity between different maps. The classification system is expected to reduce the cost/ha of ecosystem classification while at the same time improving the quality of the forest inventory.

Introduction

The construction and implementation of ecological site classification systems for the Canadian forestland is generally considered to be an important step towards sustainable management of our forestland. The process of building such classification systems has been an ongoing activity across Canada for several years. Ecological site classification systems are now in place for most of British Columbia under the biogeoclimatic ecosystem classification program (Mitchell and Green, 1981; Meidinger *et al.* 1988; Delong *et al.* 1990). In Alberta a field guide for the West Central Alberta area was produced by Corns and Annas (1986), and field guides for South Western Alberta and Northern Alberta are in preparation. Efforts to produce such

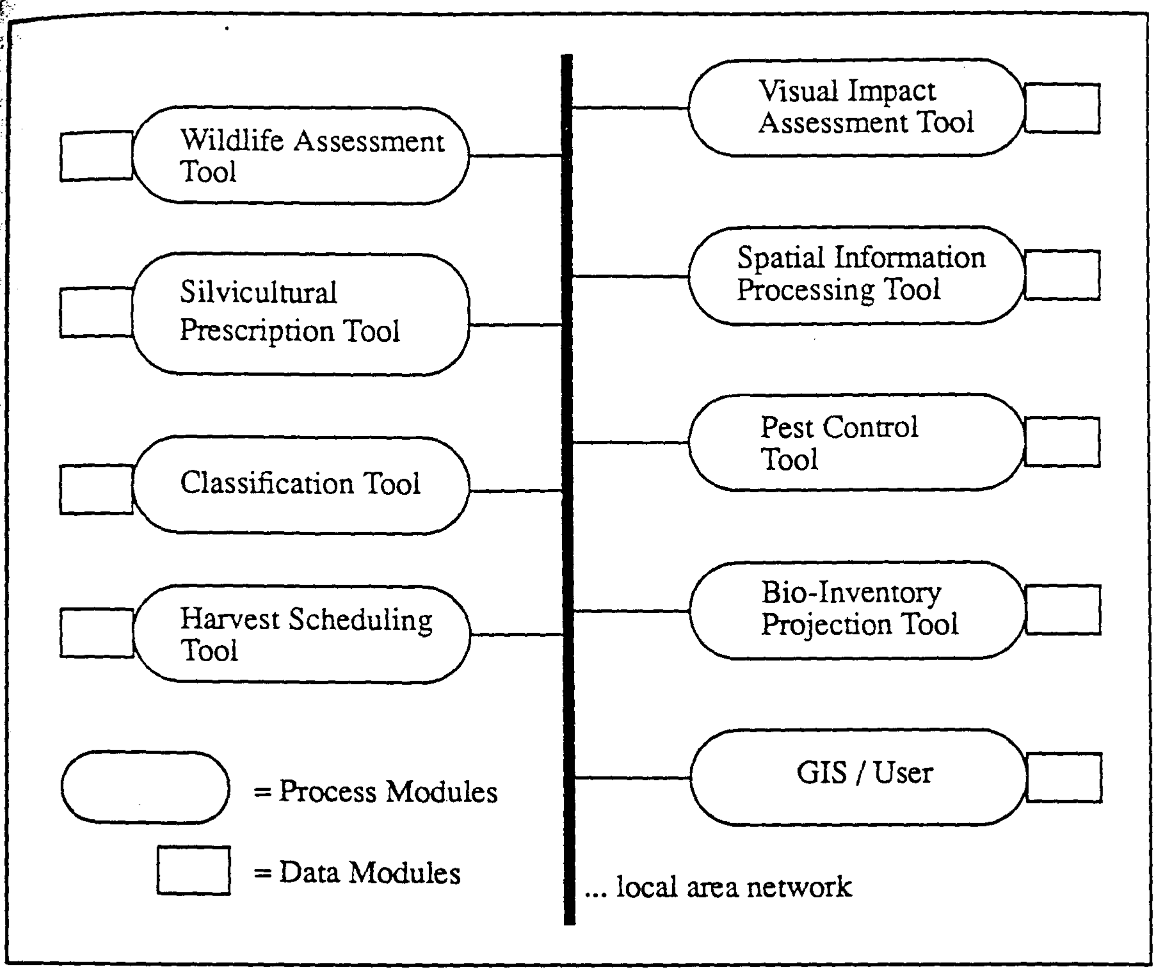


Figure 1: *Forest management decision support tool/network.*

field guides for Sasketchewan and Manitoba are in progress. In Manitoba, two pilot study areas were classified by Knapik *et al.* (1988, 1989). Similar site classification guides have been produced for parts of Ontario (Jones *et al.* 1983; Sims *et al.* 1989), New Brunswick (Zelazny *et al.* 1989) and Newfoundland (Meades and Moores, 1989).

Ecological classification field guides are an economical means of assessing the ecological characteristics and potential management opportunities and constraints for a tract of forest land. The largest disadvantage is that they do not delineate the location and areal extend of the unit classified like a map does. In order to satisfy the forester's need for mapped site information it is necessary to do a large amount of air photo interpretation and field survey transacts to sample the variability of the landscape. The cost of producing an ecological land classification map at the 1:15,000 to 1:20,000 scale desired by foresters is high. Costs of $5,50 / ha and more are to be expected. Methods for producing site inventory maps have been documented by Valentine (1986).

As part of the Naia program, the Alberta Research Council, with the cooperation of the Canadian Forest Service, Northern Forestry Centre, and Weldwood of Canada, Hinton Division, is developing a knowledge based predictive mapping system for ecosystem classification. The system, when fully operational, will act as a decision support system for forest ecosystem classification with the expectation that it will reduce the cost/ha while at the same time improving the quality of both the site classification process, the forest inventory, and ultimately forestland management decisions. The objective of the Naia program is to provide governments, forest industries, and other resource information users with practical, commercially supported, and ecologically-oriented decision support tools for forest planning. Designing the software and working closely with the forest industry is a multi-disciplinary team of specialists in natural resources and information technology. This project team is building a series of software tools using Geographic Information, Knowledge Based System, and computer modelling. From an implementation viewpoint the objective of the program is to create a series of special purpose software tools that interact with each other by means of a message passing protocol. The user encodes tasks and views results with a GIS. The different modules may all reside on one machine, or, alternatively, on different machines linked by a local or long distance network.

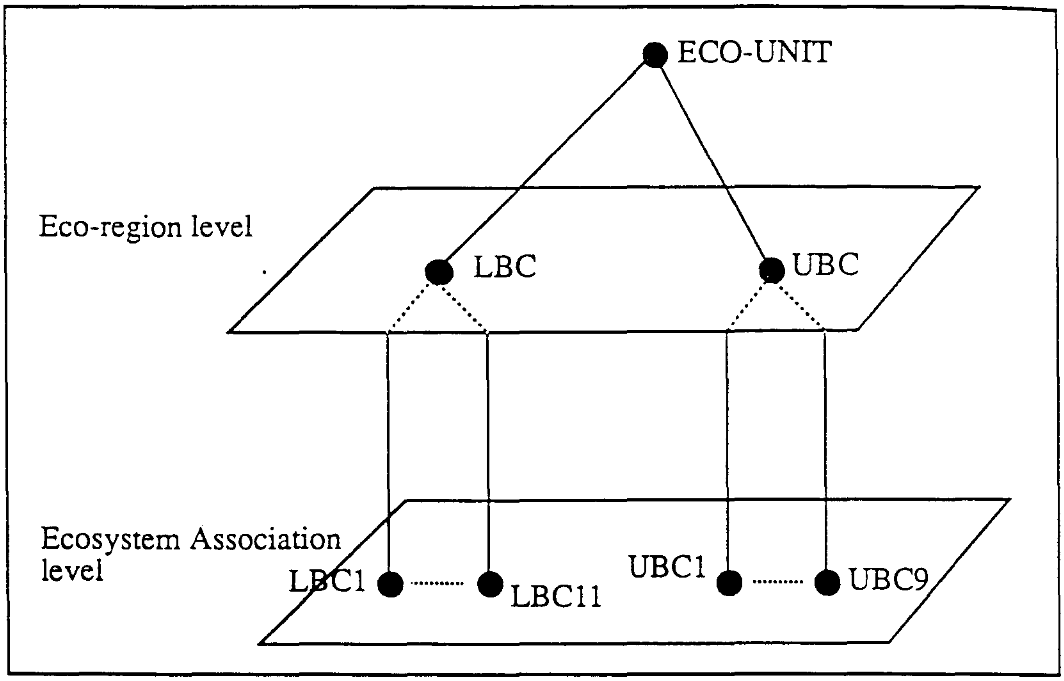


Figure 2: *Part of specialization hierarchy for West Central Alberta.*

Fig. 1 illustrates the vision of such a network. Most of the work thus far has been done on the classification tool. Some preliminary work has taken place on the wildlife and on the projection tool (Leishman and Mulder, 1993). The purpose of this paper, however, is to report on the current state of development of the classification tool. Section 2 describes the design and implementation of this tool. In sections 3 and 4 we discuss the results of recent field tests.

Design and implementation of classification tool

Naia's classification tool is a generic tool for performing classification tasks. The tool has been designed as a shell and it consists of two components: a knowledge base and a process component. This section only summarizes these processes. The reader is referred to other Naia publications for more detail. The project is described in detail in a series of papers published in the GIS-93 symposium (Jones, 1993; Mulder and Corns, 1993a; Skye, 1993; Crain *et al.* 1993). A more detailed description of the evidential reasoning component of the classification tool is provided in Mulder and Corns, 1993b, and Mulder, 1994.

The knowledge base

The knowledge base is object oriented and it recognizes three types of objects: *primitives, models,* and *features.* Primitives and features are both derived from a data domain, whereas the models constitute an interpretation domain. Primitives are the unit of interpretation. They are interpreted by models. Both primitives and models can be characterized by means of a property list which takes the form of an attribute - value list.

In the Naia classification tool the primitives are *polygons*. Models are *ecosystems*. Ecosystems are organized as a hierarchy. Fig. 2 illustrates part of such a hierarchy for West Central Alberta. The hierarchy consists of two levels: an ecoregion level and an ecosystem association level. West Central Alberta is represented by several ecoregions including Lower Boreal Cordilleran (LBC) and Upper Boreal Cordilleran (UBC). All ecosystem associations are named after their parent ecoregion (e.g. UBC1). The shell allows the use of any number of levels of specialization.

Primitives and models are linked by *features*. A feature is an entity that is derivable from input data. A feature can be discrete or continuous and it can assume a range of possible values. A feature links a primitive with a model by means of a *mass function*. Such a function is discrete or continuous and can assume any value between -1 and +1. Mass functions are obtained through knowledge elicitation sessions with domain experts. They express the belief of the expert that a particular model will occur given a certain value for a particular feature. Mass functions enable the system to compute a belief value for each possible ecosystem.

In the Naia classification tool the features are derived from available maps. Examples are: elevation, aspect, slope, forest cover, and soil. Each feature is linked to each ecosystem by means of a mass function provided by a forest ecologist. Elevation is mostly linked to ecoregions, whereas the other features constrain ecosystems at the ecosystem association level only. The ecosystem hierarchy also acts as an inheritance hierarchy. For example, a mass function constraining LBC (Fig. 2) is inherited by each of LBC's successors at the ecosystem association level.

Process component

The classification process is based on two principles: the principle of *least commitment,* and the principle of *graceful degradation.* The first principle requires that input data constrain the ecosystems at a level of specialization that is appropriate for those data. As mentioned before, elevation mostly constrains ecosystems at the ecoregion level, because this feature cannot distinguish very accurately between ecosystem associations. The second principle requires that the classification process reflects the reliability and availability of input data. For this purpose, the classification process maintains a *confidence* factor for each polygon. This factor is represented by a number between 0 and 1, 0 meaning no confidence, 1 meaning maximum confidence. The principle of graceful degradation, among other things, enables the system to deal with missing data.

The classification process takes a set of polygons as input. These polygons are created by a map overlay process in which each polygon inherits the features of its "parent" maps. Classification is performed polygon by polygon. At first each polygon is interpreted as an ecounit (see Fig. 2) with a belief value 1. As no input data have thus far been considered the confidence at this point will be 0. Input data are represented as features. The features are introduced one by one. The mass functions of each feature constrain the ecoregions and ecosystem associations possible. As classification progresses the ecounit is replaced by one or more of its successor ecosystems in the specialization hierarchy (Fig. 2). For each polygon the classification process creates: a listing of possible ecosystems each with an associated belief value, a confidence value, and an explanation trace.

The classification tool is loosely connected with a GIS. Map overlay is done with the GIS. Classification is performed by a Knowledge Based System. Visualization of results, in turn, is done again with the GIS. routines were developed with ARCINFO™ . The Knowledge Based System was implemented in Common Lisp Object System (CLOS).

Testing and validation

The classification tool has already been implemented and tested in several regions of Alberta and Manitoba. Different classification systems were used for Alberta and Manitoba. Based on the shell concept the implementation of the system for different areas of the country has thus far proven to be easy. The original validation of the system was done against ground truth points collected as part of the creation of the local ecosystem classifications, and independent mapping efforts. The prediction accuracy of the system in the first set of trials was in the 63% range (Mulder and Corns, 1993a), but was subsequently improved to 85 - 94% (Mulder and Corns, 1993b).

The first field test took place in the summer of 1993 in the forest management area of Weldwood in Hinton. Fig. 3 illustrates one of the two areas for which ecosystem prediction was made by the classification tool. The classification used in this experiment was the ecosystem classification for West Central Alberta (Corns and Annas, 1986). The results thus far have been very interesting and promising.

In essence, every kind of prediction made by the system must be verified in the field. Fig. 3 shows for each polygon the ecosystem association with the highest belief. In the GIS, however, the forest ecologist can view the complete list of possible classifications and their associated beliefs. The field work is very much driven by these listings. For example, for polygons in which the ecosystem association with the highest belief far exceeds the belief of its closest competitor a few sample points for that area may suffice. If field observations prove, that the system's prediction is correct, then one may trust the overall prediction for that polygon. On the other hand, in polygons for which the beliefs are nearly equal, a much more concentrated field validation effort is necessary. One form of potential cost savings therefore comes from the fact that the predictions of the classification tool can serve as a focus for the field work.

The West Central Alberta classification system has been constructed with ground truth points collected in several areas in West Central Alberta. Such points, however, were limited in the Hinton area. When this particular knowledge base setting was tried for the Hinton area the results were at first disappointing. A very large number of polygons were incorrectly classified. The good news, on the other hand, was that most errors were systematic. These problems were easily solved by making some minor corrections to the mass functions.

Figure 3 also shows a number of ambiguous classifications. For example, the legend shows areas with UBC3/4 and UBC5/6. This means that the system could not distinguish between UBC3 and 4 in the one case and between UBC5 and 6 in the other. In the case of UBC3 and 4 both ecosystem associations have lodgepole pine as the dominant tree species. The associations differ, however, in the secondary species (black spruce for UBC3 and white spruce for UBC4). Although the presence of a secondary species was apparent in the field, the forest cover map data sometimes did not indicate a secondary species, which resulted in the ambiguity in classification.

Not every possible combination of feature values leads to an ecosystem classification. This is quite apparent in Fig. 3 where each white area implies that the system could not come up with a classification. Such a failure has one of three causes: a data problem, a model problem, or a combination of both. Errors in a single data

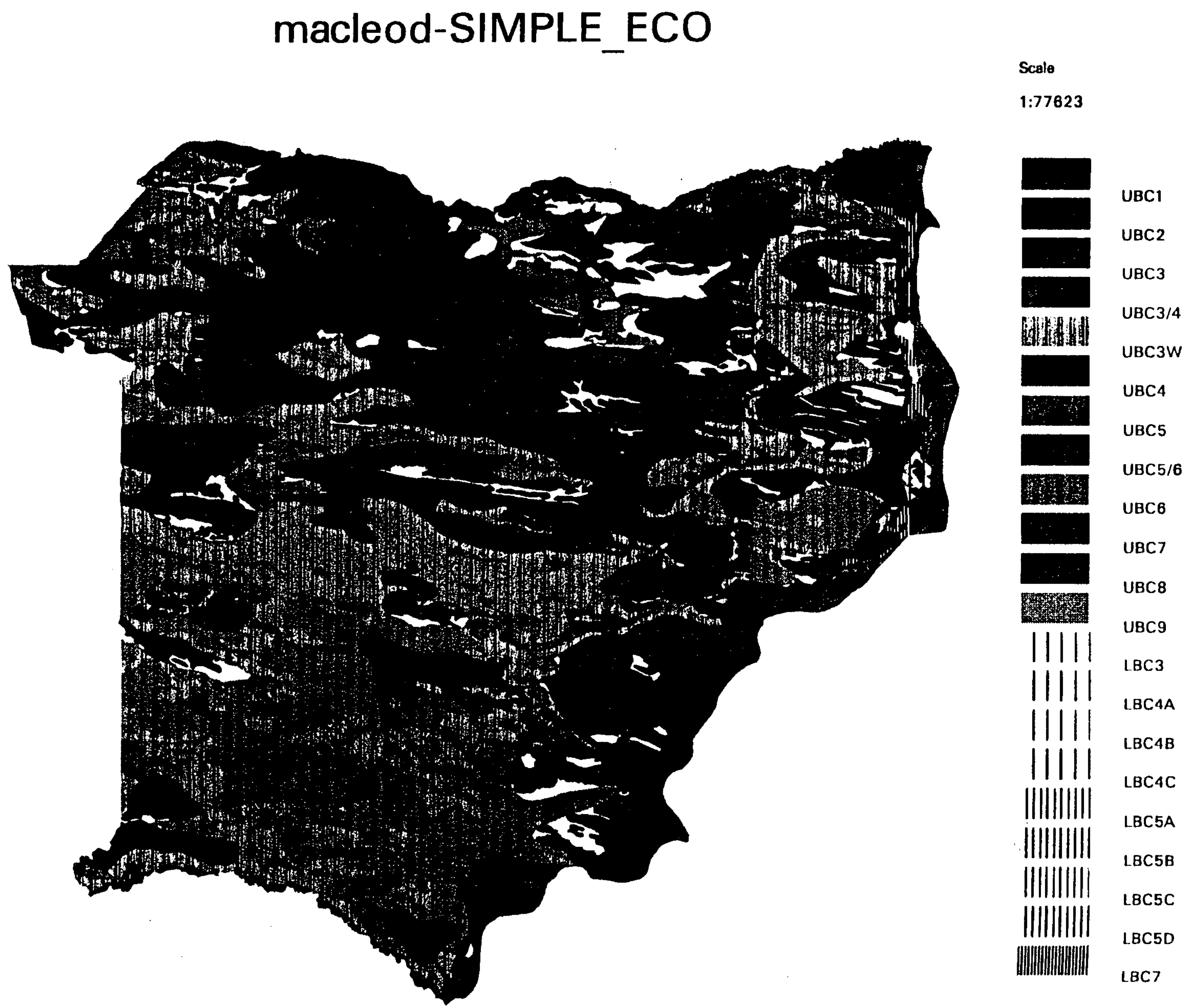


Figure 3: *An ecosystem map generated by the Naia system.*

source turned out to be a rarity. In some cases, however, data were incomplete, as was exemplified above with forest cover. Many of the unclassified polygons are slivers. These are mostly caused by a misalignment of forest cover and soil maps. The cause of this misalignment is a difference in scale. The soil maps for the area have a much smaller scale than the forest cover. Field checks indicated, as expected, that the larger scale forest cover data are more reliable than the soil. Efforts are underway to remove slivers caused by a forest cover/soil inconsistency by using knowledge about the compatibility between tree species and soil series.

As mentioned before, the ground truth data used to build the classification system did not include data from the Hinton area. Some unusual data combinations were found that were not covered by the classification system. More specifically, areas with lodgepole pine and black spruce are generally found on well drained to imperfectly drained soils, but not on poorly drained soils. The latter condition was regularly found in the test area. A new ecosystem association had to be introduced to cover these data. The classification UBC3W does therefore not occur in the classification field guide (Corns and Annas, 1986). The modular shell like design of the system makes it easy to add new ecosystems to the classification.

Discussion

The lessons learned from the system thus far have been interesting and promising. A first series of field tests have indicated that Naia can be made to work well after some adjustments in the knowledge base to reflect local situations. Adjustments include minor modifications in mass functions and, possibly, the addition of new associations to the classification system. The modular design of the system makes such adjustments easy to implement.

The more surprising result of the field tests, however, was that Naia does much more than provide ecosystem classifications. As part of the classification process the system points at weaknesses and ambiguities in existing data sets. As well, it will put a finger on conflicts between different data sets. We discussed this issue in section 3. Finally, the system will point at gaps in the classification system used. All of these capabilities will assist with improving the quality of the existing data sets. Higher quality data will allow for more informed decisions which, in turn, will lead to increased efficiency in forest operations.

A comparison test is underway for a 5,000 ha area in the Weldwood forest management area in which the traditional field survey methodology is being compared with the Naia approach with respect to prediction accuracy and cost. Based on field tests completed thus far, the prediction is that Naia will save costs both directly and indirectly. The classification tool has a focussing effect on the field work. Field work will concentrate on areas with high classification ambiguity as predicted by Naia. This is expected to lead to a reduction in the overall field work that is required with the traditional approach. Direct cost savings can be expected as a result of this focussing effect. A further direct savings is expected from the fact that the Naia approach requires no digitization of data, as all data already exist in digital form. Indirect cost savings can be expected from an increase in classification consistency. When classifying a large area the traditional method will require the use of different field teams. This will potentially lead to discrepancy in interpretations by different teams. With the field effort focus of the Naia approach a single team can cover a much larger area. The most important indirect cost savings, however, will come from an improvement in the quality of the forest inventory, a side effect of the classification process. With all these capabilities Naia is expected to provide ecosystem classification for much less than $5.50/ha which is the estimated cost of classification with the traditional methodology.

Acknowledgments

The research and field testing work described in this paper has been supported by a grant from the Natural Sciences and Engineering Research council (NSERC) of Canada (OG0121247) to the first author. In-kind support for this project has been provided by the Alberta Research Council, the Canadian Forest Service, Northern Forestry Centre, and Weldwood of Canada, Hinton division. The authors are grateful to Don Kuzenko for assistance with the software development of the classification tool. Thanks also to Dan Jaliff for comments on a draft version of this paper. Sean Curry and Dave Pressley of Weldwood are gratefully acknowledged for providing materials and assistance with the field validation. ARCINFO is a geographic information system marketed by ESRI.

References

Crain I.K., Gong P., Chapman M.A., Lam S., Alai J., Hoogstraat M., Implementation Considerations for Uncertainty Management in an Ecologically-oriented GIS, Proceedings of GIS 93, Vancouver B.C., 1993, pp. 167-172.

Corns I.G.W., Annas R.M., Field Guide to Forest Ecosystems of West Central Alberta, Northern Forestry Centre, Canadian Forest Service, Edmonton, 1986.

Delong C., McLeod A., MacKinnon A., Jang L, A Field Guide for Identification and Interpretation of Ecosystems of the Northeast Portion of the Prince George Forest Region, B.C. Min. For., Land Management Handbook No 22, Victoria B.C., 1990.

Jones R.K., Next Generation Forest Site Classification: Ecologically - Oriented Predictive Mapping Technology, Proceedings of GIS 93, Vancouver B.C., 1993, pp. 143-152.

Jones R.K., Pierpoint G., Wickware G.M., Jeglum J.K., Arnup R.W., Bowles J.M., Field Guide to Forest Ecosystem Classification for the Clay Belt, Site Region 3E, Ont. Min. Nat. Res., Queens Printer, Toronto, 1983.

Knapik L.J., Russell W.B., Riddell K.M., Stevens N., Forest Ecosystem Classification and Land System Mapping Project, Duck Mountain, Manitoba, Northern Forestry Centre, Canadian Forestry Service, Edmonton, Alberta, July 1988.

Knapik L.J., Ellis R.A., Phillips B., Forest Ecosystem Classification and Land System Mapping Pilot Project, Sandilands, Manitoba, Northern Forestry Centre, Canadian Forestry Service, Edmonton, Alberta, March 1989.

Leishman D., Mulder J.A., Preliminary Functional Design of a Dynamic Spatially Oriented Biophysical Inventory Projection System, study report, Canadian Forest Service, Northern Forest Centre, 1993.

Meades W.G., Moores L., Forest Site Classification Manual: A Field Guide to the Damman Forest Types of Newfoundland, Newfoundland Forestry Centre, Forestry Canada, St. Johns, and Newfoundland Dept. of For. and Agriculture, 1989.

Meidinger D., McLeod A., MacKinnon A., DeLong C., and Hope J., A Field Guide for Identification and Interpretation of Ecosystems of the Rocky Mountain Trench, Prince George Forest Region, B.C. Min. For., Land Management Rep. No 22, Victoria B.C., 1988.

Mitchell W.R., Green R.N., Identification and Interpretation of Ecosystems of the Western Kamloops Forest Region., First approx., B.C. Min. For., Land Management Rep. No 2, Victoria B.C. 1981.

Mulder J.A., Corns I.G.W., A Decision Support System for Predicting and Consolidating Ecosystems from Existing Map Data and Classification Systems, Proc. of GIS93 Symposium, Vancouver B.C., 1993a, pp. 153 - 160.

Mulder J.A., Corns I.G.W., Naia: A Decision Support System for Predicting Ecosystems from Existing Land Resource Data, Int. Joint Conf. on Artificial Intelligence, Workshop on Applications of AI in Agriculture, Natural Resources and Environment, Chambery, France, September 1993b, pp. 137 - 150.

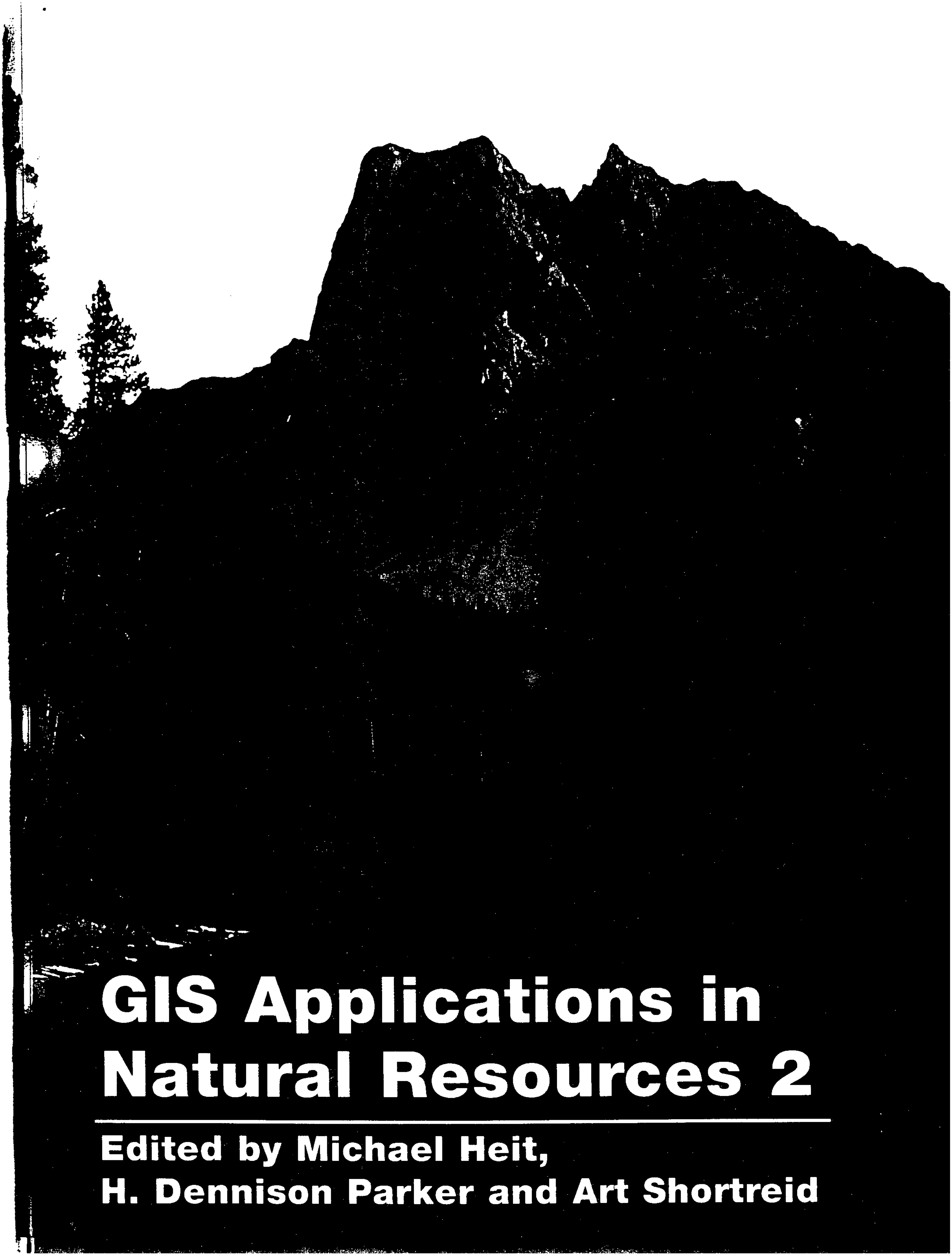
Mulder J.A., Two Views of Uncertainty in the Dempster-Shafer Theory: Hypothesis Uncertainty versus Belief Uncertainty, in preparation.

Sims R.A., Towill W.D., Baldwin K.A., and Wickware G.M., Field Guide to the Forest Ecosystem Classification for NorthWestern Ontario, Great Lakes For. Centre, Forestry Canada, Sault Ste Marie, Ont., and Ontario Min. Nat. Res., 1989

Skye D., Naia: Ecologically-oriented Spatial and Knowledge- based Framework to Support Forest and Land resource Management, Proceedings of GIS-93, Vancouver B.C., 1993, pp. 161-166.

Valentine K.W.G., Soil Resource Surveys for Forestry: Soil Terrain and Site Mapping in Boreal and Temperate Forests, Clarendon Press, Oxford. 1986.

Zelazny V.F., Ng T.T.M., Hayter M.G., Bowling C.L., Bewick O.A., Field Guide to the Forest Site Classification in New Brunswick, Canada - New Brunswick For. Subsidy Agreement, Publ. N.B. Dept. Nat. Res. and Energy, Federicton N.B., 1989.



GIS Applications in Natural Resources 2

Edited by Michael Heit, H. Dennison Parker and Art Shortreid

GIS Applications in Natural Resources 2

Edited by
Michael Heit, H. Dennison Parker, and Art Shortreid

GIS Applications in Natural Resources 2 is a selected compilation of unedited papers from the GIS '92–'95 Symposia held in Vancouver, British Columbia.

Library of Congress Cataloging-in-Publication Number: 92-191861

ISBN 1-882610-17-2

For more information contact:
E-mail: *books@gisworld.com*
WWW: *htttp//www.gisworld.com*
GIS World Books
155 E. Boardwalk Drive, Suite 250
Fort Collins, Colorado 80525
USA

Watershed Management

Managing Sustainable Development