Title
Alternative representations of in-stream habitat: classification using remote sensing, hydraulic modeling, and fuzzy logic

Journal
International Journal of Geographical Information Science, 19(1)

ISSN
1365-8816

Authors
Legleiter, Carl J.
Goodchild, M F

Publication Date
2005

Peer reviewed

International Journal of Geographical Information Science
Vol. 000, No. 000, Month 2005, 1–22

Alternative Representations of In-Stream Habitat: Classification using Remote Sensing, Hydraulic Modeling, and Fuzzy Logic

CARL J. LEGLEITER and MICHAEL F. GOODCHILD
Geography Department, University of California Santa Barbara, Ellison Hall 3611, Santa Barbara, CA, 93106; e-mail: carl@geog.ucsb.edu; good@geog.ucsb.edu

(*Received 24 January 2004; in final form 6 April 2004*)

Improved techniques are needed to characterize complex fluvial systems and monitor ecologically important, yet highly vulnerable riverine environments. This paper explores potential alternatives to traditional mapping of in-stream habitat and presents fuzzy set theory as a means of departing from the rigid, Boolean, object-based framework. We utilize hydrodynamic modeling, remotely sensed data, and fuzzy clustering to obtain classifications that allow for continuous partial membership and gradual transitions among habitat types. Methods of assessing cluster validity are available, but data quality is a crucial consideration. Crisp, vector-based representations can be derived from raster fuzzy classifications by applying a threshold to maximum membership values. This process results in conditional objects separated by ambiguous transition zones, and a compromise must be reached between the proportion of the channel assigned to polygons and the certainty with which this assignment can be made. Spatial patterns of classification uncertainty can also be used to identify areas of confusion, infer boundaries of variable width, and highlight areas of increased habitat diversity. Hydraulic modeling and remote sensing complement one another and, together with field work, could provide a more realistic representation of the fluvial environment.

Keywords: rivers; remote sensing; hydraulic modeling; fuzzy classification; indeterminate boundaries; uncertainty

1. Introduction

Fluvial systems vary continuously across a range of spatial and temporal scales as dynamic erosive and depositional processes establish the physical habitat template for aquatic biota (Ward 1989, Knighton 1998, Wohl 2000). Characterizing this complexity poses a considerable challenge to scientists and resource managers seeking concise, meaningful representations of channel morphology and in-stream habitat. This task takes on added importance due to the heightened biodiversity of riparian corridors and the sensitivity of riverine landscapes to disturbance (Malmqvist and Rundle 2002, Ward *et al*. 2002). This combination of ecological significance and vulnerability mandates consistent, quantitative, and spatially explicit characterization of fluvial environments (Newson and Newson 2000).

Conventional stream classification methods, however, suffer from several fundamental limitations (Poole *et al*. 1997, Goodwin 1999, Roper *et al*. 2002). Typical approaches include process-oriented categorizations (Whiting and Bradley 1993, Montgomery and Buffington 1997), visual qualitative schemes (Bisson *et al*.

International Journal of Geographical Information Science
ISSN 1365-8816 print/ISSN 1362-3087 online © 2005 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/13658810412331280220

1982, Hawkins *et al*. 1993), and Rosgen's (1994) popular descriptive framework. In practice, field mapping of in-stream habitats "retain(s) a strong flavour of qualitative ambiguity" (Burrough 1996, p. 18) due to imprecise formulation of the entities to be mapped. As a result, these techniques have been criticized for their inherent subjectivity, a basic flaw that "seriously compromises (their) repeatability, precision, and transferability" (Poole *et al*. 1997, p. 879) and casts doubt upon the mere notion of classifying streams (Goodwin 1999).

Having expressed their dissatisfaction with the traditional approach to habitat unit classification, Poole *et al*. (1997) recommend that stream monitoring programs "must instead focus on direct, repeatable, cost-efficient, and quantitative measures … spanning several scales of resolution" (p. 879). On a local level, for example, hydrodynamic models can be used to simulate flow patterns determining the suitability of in-stream habitat for various organisms (Ghanem *et al*. 1996, Lamouroux *et al*. 1998, Crowder and Diplas 2000). Remote sensing provides a more synoptic perspective, and previous research has demonstrated the feasibility of mapping in-stream habitats from digital image data (Wright *et al*. 2000, Legleiter *et al*. 2002, Marcus 2002, Whited *et al*. 2002, Marcus *et al*. 2003, Legleiter 2003). Advantages of remote sensing include expanded geographic coverage, allowing rivers to be examined on a watershed rather than a reach scale, and consistent, quantitative description of fluvial environments through image classification procedures (Marcus 2002).

Both ground-based and remote approaches to characterizing in-stream habitat have been limited by their rigid adherence to a Boolean, object-based data model (Burrough 1996, Fisher 1996, Fisher 1998). This conventional approach does not provide an adequate representation due to the internal heterogeneity of habitat units, the subjectivity of boundary delineation, and the lack of an accepted classification scheme. Fuzzy set theory provides an alternative descriptive framework that accommodates such uncertainty by allowing individual entities to exhibit continuous partial membership in more than one class, which, in a spatial setting, allows for zones of gradual transition (Burrough 1989, Burrough and Frank 1996). Burrough (1989) states that the use of fuzzy sets is "appropriate … whenever we have to deal with ambiguity, vagueness and ambivalence in … conceptual models of empirical phenomena," adding that "fuzziness is often a concomitant of complexity" (p. 479). Given the intricate process interactions controlling channel morphology and the inadequacies of conventional stream classification methods, fuzzy representations of in-stream habitat present an appealing alternative. The quantitative data obtained via hydraulic modeling and/or remote sensing also allows for more rigorous treatment of uncertainty.

In this paper, we apply fuzzy set theory and spatially explicit metrics of classification uncertainty to a new domain – the identification and delineation of aquatic habitats in shallow stream channels. From a GIScience perspective, these fluvial environments pose a number of unique challenges: 1) a narrow, elongate geometry; 2) a scale-dependent range of dynamic behavior in space and time; 3) a compelling need to address underlying geomorphic and ecologic processes; and 4) a paucity of suitable data sources and classification systems for characterizing habitat conditions. For the stream ecologist or resource manager, we propose that an innovative, fuzzy approach could circumvent the subjectivity of conventional habitat classification and provide a richer representation that more faithfully honors the complexity of the fluvial environment. In the following sections we use

100659
International Journal of Geographical Information Science gis51069.3d 17/11/04 11:15:01
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

hydrodynamic simulations, remotely sensed data, and fuzzy classification techniques to explore potential alternative representations of in-stream habitat and characterize spatial patterns of classification uncertainty.

2. Methods

2.1 *River2D Hydrodynamic Model*

The River2D model used in this study was developed at the University of Alberta and is freely distributed over the internet (Steffler and Blackburn 2002). Data requirements include: 1) channel bed topography; 2) an estimate of flow resistance, expressed in terms of the roughness height k_s; 3) parameters describing transverse eddy viscosity; and 4) boundary conditions, typically an inflow discharge and an estimated water surface elevation at the lower end of the reach. An important, often challenging aspect of two-dimensional hydraulic modeling is the generation of a finite element mesh to discretize the model domain. An iterative computational scheme is then invoked to obtain steady state solutions of the depth-averaged governing equations of fluid flow (Miller and Cluer 1998). Flow depth and downstream and transverse velocity components are estimated at each mesh node and other quantities can be derived from these basic results (Steffler and Blackburn 2002). This study utilized survey data from a 625 m reach of the Kananaskis River in Alberta, Canada, distributed with the River2D model; this regulated river is the site of ongoing research on in-stream flow requirements for habitat maintenance (Katopodis, 2003). Following the procedure outlined by Blackburn and Steffler (2002), a hydrodynamic simulation was performed for a typical base flow discharge of 2 m^3/s to obtain estimates of flow depth, velocity magnitude, Froude number, and shear velocity at each mesh node. These data were then extracted on a 1 m grid with a total of 12699 nodes.

2.2 *Remotely Sensed Data and Image Processing*

Airborne imagery of the Lamar River, a fifth-order stream in northeastern Yellowstone National Park, Wyoming, USA, was acquired by a pair of sensors: the Probe-1 hyperspectral and ADAR multispectral instruments. This stream has been the site of prior research into remote mapping of fluvial environments and detailed descriptions of the study area are provided in Marcus *et al.* (2003) and Legleiter (2003). Both data sets featured the fine spatial resolution required for small mountain streams, but provided different levels of spectral detail and radiometric precision (Table 1). Whereas the Probe-1 data were acquired on an experimental basis and are somewhat exceptional for their combination of fine spatial and spectral resolution, the ADAR instrument is operated commercially and is more typical of the data that resource managers might obtain in practice.

Given this study's focus on in-stream habitat, both images were masked to include only the active channel. The Probe-1 hyperspectral image was transformed using a common data reduction and noise removal technique - the minimum noise fraction (MNF; Boardman and Kruse 1994) – prior to image classification. The Probe-1 Lamar River scene was also subdivided into two subsets to contrast different stream segments: a 245 m-long lower reach and an upper reach approximately 365 m in length, corresponding to the ADAR image analyzed in this study.

In support of the Probe-1 overflight, a two-person team mapped in-stream habitat on the Lamar River using the Bisson *et al.* (1982) classification scheme popular

Table 1. Characteristics of the remote sensing instruments used in this study.

Sensor	Probe-1	ADAR
Type	Hyperspectral	Multispectral
Spatial resolution (IFOV)	1 m	0.75 m
Number of bands	60*	4
Spectral range	437.6–1301.7 nm*	450–860 nm
Spectral resolution	16–20 nm	60–110 nm
Radiometric resolution (range)	12 bit (0–4095)	8 bit (0–255)
Image acquisition date	August 3, 1999	October 7, 1999

*The Probe-1 instrument actually features 128 bands extending into the shortwave-infrared portion of the spectrum (2500 nm), but only the first 60 bands were considered in this study due to the strong absorption of infrared light by a water column (Bukata *et al.* 1995).

among resource management agencies. These field maps consisted of discrete channel-spanning polygons representing seven classes: eddy drop zones, high- and low-gradient riffles, glides, runs, rough-water runs, and pools; detailed descriptions of these units are available in Ladd *et al.* (1998) and Legleiter *et al.* (2002). Habitat units were mapped to hard copies of the imagery to facilitate co-registration and comparison with image-based classifications (Marcus 2002). Although buffering habitat polygons can improve classification accuracy by up to 15% (Legleiter *et al.* 2002, Legleiter 2003), in this study every pixel was assigned to a class to examine the uncertain transitional zones between adjacent units. Training sites selected from the field maps were used to produce maximum likelihood supervised classifications (Richards and Jia 1999) for both the original seven units mapped in the field and a simplified scheme consisting of eddy drop zones, riffles, a single run/glide category, and pools. The accuracy of these classifications was assessed using error matrices (Congalton and Green 1999) and mapped using uncertainty indices (Zhu 1997, 2001).

2.3 *Fuzzy Partitions and the Fuzzy c-Means Algorithm*

The basic premise of fuzzy set theory is that a space of objects (e.g., the pixels of an image) can be partitioned into a series of fuzzy sets by assigning to each individual entity a ‘‘grade of membership’’ (Kandel 1986, p. 3) in each of the sets. These fuzzy membership values range between 0 (no membership) and 1 (type specimen) and specify the degree to which each individual can be regarded as belonging to a specific fuzzy set. Defining the membership functions of fuzzy sets thus takes on critical importance, and, under the data-driven similarity relation model used in this study, the value of the membership function depends on the classification method used (Burrough, 1989).

The fuzzy *c*-means (FCM) algorithm described by Bezdek *et al.* (1984) has proven especially popular and has been used to produce land-cover maps from remotely sensed data (Zhang and Stuart 2001). The method is based upon the concept of a fuzzy *c*-partition U, which subdivides a data set comprised of n observations into a collection of $c \leqslant n$ fuzzy subsets, called clusters. The degree to which the k^{th} observation is a member of the i^{th} cluster is expressed in terms of a fuzzy membership value $\mu_{ik} \in [0,1]$. Each individual observation has a membership in each of the c clusters, with values close to 1 indicating a high degree of similarity between the observation and the center of a particular cluster. Beginning with a randomly generated c X n initial partition U_0, the FCM algorithm proceeds by minimizing a generalized least-squared error functional

$$J_m(U, v) = \sum_{k=1}^{n} \sum_{i=1}^{c} (\mu_{ik})^m d_{ik}^2, \tag{2}$$

where U is the fuzzy c-partition, v denotes a vector of cluster centers in p-dimensional space (p is the number of variables), m is a weighting exponent that controls the degree of fuzziness (typically between 1.5 and 3), and d_{ik} is the distance between the k^{th} observation and the i^{th} cluster center. Given values of c and m, the algorithm iteratively assigns observations to clusters and recalculates cluster centers until J_m achieves a local minimum (Bezdek 1981, Bezdek *et al.* 1984).

In this study, the FCM algorithm was implemented in MATLAB using a Euclidean distance metric. Three data sets, consisting of four variables each, were used to produce a series of fuzzy partitions: 1) flow depth, velocity magnitude, Froude number, and shear velocity from the River2D hydrodynamic model of the Kananaskis River; 2) brightness values in the blue, green, red, and shortwave-infrared bands of the ADAR imagery of the Lamar River; and 3) the first four MNF-transformed bands derived from the upper and lower subsets of the Lamar River Probe-1 scene. All observations (pixels or hydraulic model grid cells) were input to the FCM algorithm and unsupervised fuzzy classifications were generated for all combinations of the number of clusters ($2 \leqslant c \leqslant 10$, in increments of 1) and weighting exponent ($1.1 \leqslant m \leqslant 2.5$, in increments of 0.1), for a total of 135 FCM classifications of each data set.

2.4 *Indices of Cluster Validity*

Because the FCM technique will generate a fuzzy partition for any number of clusters c and any weighting exponent m, each (c, m) pair represents a unique algorithm. The fuzzy membership values obtained thus depend on the choices of c and m, but in many cases neither the appropriate number of clusters nor a suitable weighting exponent is known beforehand (Gath and Geva 1989). Even if there is some rationale for particular choices of c and m, a means of assessing the quality of the clusters generated by the FCM algorithm is necessary (Bezdek 1981).

Several decades of research in the fields of pattern recognition and numerical taxonomy have produced a variety of heuristic indices of cluster validity that can guide selection of optimal fuzzy c-partitions (Bezdek 1981, Bezdek and Pal 1992, Pal and Bezdek 1995, Zahid *et al.* 1999, Halkidi *et al.* 2001). Some of these criteria utilize only the final fuzzy c-partition U while other validity functions incorporate the geometric structure of the original data to identify partitions that consist of compact (internally cohesive), separate (mutually distinctive) clusters (Gath and Geva 1989, Zahid *et al.* 1999). "Thus, although the environment is fuzzy, the aim of the classification is generation of well-defined subgroups" (Gath and Geva 1989, p. 774), and the "most valid" c-partition is thus the "least fuzzy" (Bezdek 1981, p. 98). In the context of natural resource mapping, this basic heuristic is appropriate because the ambiguity and vagueness of fuzzy classifications are not readily accommodated by a cartographic legacy devoid of such inexactness (Fisher 1998).

In this study, a collection of eight cluster validity indices (Table 2) were utilized to evaluate the effects of the c and m parameters, infer an optimal number of in-stream habitat classes, and determine an appropriate degree of fuzziness. For detailed descriptions and derivations of these cluster validity criteria the interested reader is

Table 2. Cluster validity criteria used to evaluate fuzzy classification results obtained using different numbers of clusters c and various weighting exponents m.

Cluster Validity Criterion (reference)	Summary Description
Partition Coefficient PC (Bezdek 1981, Bezdek *et al.* 1984)	The value of PC is inversely proportional to the overall average overlap between pairs of fuzzy subsets. For c clusters, $1/c \leqslant PC(c) \leqslant 1$ and $PC(c)=1$ implies that the partition U is a hard (crisp or Boolean) partition. Conversely, a value of $PC(c)=1/c$ implies that $U=[1/c]$ and that membership is spread evenly over all classes (maximal fuzziness).
Partition Entropy PE (Bezdek 1981, Bezdek *et al.* 1984)	The value of PE is inversely proportional to the degree to which membership is spread across clusters. For c clusters, $0 \leqslant PE(c) \leqslant \log(c)$, with $PE(c)=0$ indicating a hard partition and $PE(c)=\log(c)$ implying that $U=[1/c]$ such that membership spread evenly over all classes.
Xie-Beni (1991) cluster validity function XB	Incorporates both the original data X and the fuzzy c-partition U. Smaller values of XB indicate a partition with clusters that are compact and separate from one another. The optimal c-partition minimizes XB.
Fukuyama-Sugeno cluster validity function FS (Zahid *et al.* 1999)	Combines the FCM functional J_m with a second term describing the separation between cluster centers and the overall mean of the data set. Smaller values of FS indicate a partition with compact, separate clusters, and the optimal c-partition minimizes FS.
Fuzzy hypervolume FHV (Gath and Geva 1989)	Calculates a fuzzy covariance for each cluster and computes the volume of p-dimensional space occupied by these clusters. Small values of FHV indicate compact clusters and thus better fuzzy c-partitions.
Average partition density D_{PA} (Gath and Geva 1989)	Computed using FHV above and the sum of central members, which takes into account only those data within a specified statistical distance of the cluster center. A fuzzy density is calculated for each cluster and then averaged over all clusters. Higher values of D_{PA} indicate better c-partitions, and a partition with both dense and loose clusters is considered good due to the dense substructures.
Partition density P_D (Gath and Geva 1989)	This criterion expresses the overall partition density according to the intuitive, physical definition of density. Higher values of P_D indicate compact, separate clusters and better c-partitions.
Separateness-compactness index SC (Zahid *et al.* 1999)	This criterion incorporates both the partition U and the data X using fuzzy union and intersection operations. Higher values of c indicate good cohesion within clusters and a small overlap between pairs of clusters; maximizing SC provides a well-defined fuzzy c-partition.

referred to the original publications; Zahid *et al.* (1999), Pal and Bezdek (1995), and Halkidi *et al.* (2001) provide useful summaries.

2.5 *Defuzzification and α-cuts*

Although fuzzy clustering provides valuable data on the uncertainty associated with the assignment of an observation to a class, traditional categorical maps are often required in practice (Zhang and Stuart 2001). Fuzzy membership values quantify the similarity of the in-stream habitat at each location to each of the clusters, which can be visualized as a series of fuzzy surfaces. A crisp classification can be derived by assigning each observation to the cluster in which it exhibits the highest grade of membership, a process which has been referred to as "a kind of defuzzification" (Zhang and Stuart 2001, p. 182). This simple maximization operation, itself a form of maximum likelihood classification, was used to obtain crisp classifications of the remotely sensed data and hydraulic model output. The maximum fuzzy membership values were also retained as a measure of the association of each pixel with the class to which it was assigned.

Vector representations were derived following the fuzzy-fuzzy object model developed by Cheng *et al.* (2001) for features that have a vague thematic description, an uncertain spatial extent, and, potentially, some degree of spatial overlap with one another – an appropriate description of the in-stream habitats mapped in this study. Boundaries can only be ascribed to fuzzy surfaces to produce discrete objects under a specific set of conditions, specified as thresholds called α-cuts (Kandel 1986), that the maximum fuzzy membership value of a pixel must meet or exceed (Cheng *et al.* 2001). This process has been described as a "transformation between attribute uncertainty in the raster domain and boundary uncertainty in the vector domain" (Zhang and Stuart 2001, p. 182) and can be used to translate fuzzy surfaces into categorical maps consisting of objects meeting a predetermined α-level of certainty separated by gaps representing zones of uncertainty, called epsilon bands (Zhang and Stuart 2001). In this study, conditional in-stream habitat polygons were derived from fuzzy classifications by applying a series of α-cuts. The corresponding epsilon bands provided an indication of the extent, location, and overall abundance of boundary zones between habitat units.

2.6 *Spatial Variability of Classification Uncertainty*

For both supervised, crisp classifications and unsupervised, fuzzy clustering of in-stream habitat, the confidence with which pixels could be assigned to classes was expected to vary spatially. Indices of classification uncertainty developed in the context of soil classification (Zhu 1997, 2001) were used to examine these spatial patterns and identify areas of the channel that were not typical of any single class. Zhu (2001) draws a distinction between two types of classification uncertainty. Exaggeration uncertainty describes the error incurred by assuming that an observation has full membership in a class and quantifies the dissimilarity between the entity and the class to which it is assigned. Ignorance uncertainty summarizes the membership values of an entity in the classes other than that to which it has been assigned and is proportional to the fuzziness of the entity. Classification uncertainty indices are based upon similarity vectors S_s^c that describe the membership of the entity at each location s in each of the c classes. For the FCM classifications, fuzzy membership values provided the similarity vectors, and Zhu's (2001, p. 341) procedure was used to calculate similarity vectors for supervised classifications of the Lamar River Probe-1 scene.

Having obtained similarity vectors, each pixel is assigned to the class to which it is most similar; this is equivalent to the defuzzification procedure described above and has also been called "hardening" (Zhu 2001, p. 335). An entropy measure H_s is then used to estimate the ignorance uncertainty associated with this hardening:

$$H_s = -\frac{1}{\log_e c}\sum_{i=1}^{c}\left[\left(S_s' i\right)\log_e\left(S_s' i\right)\right], \tag{3}$$

where i indexes the c classes and $S_s' i$ denotes the i^{th} element of a similarity vector that is normalized to sum to one. The entropy H_s ranges between 0 and 1, with a value of 0 indicating that the entity at s has full membership in the class to which it was assigned (zero ignorance uncertainty) and a value of 1 implying that the entity exhibits an equal degree of membership in all c of the classes (complete ignorance uncertainty). The exaggeration uncertainty associated with the hardening of a classification is defined as

$$E_s = \left(1 - S_s^g\right), \tag{4}$$

where S_s^g is the similarity between the entity at location s and the class g to which it is assigned. These two indices were computed for each in-channel location and used to produce maps displaying the spatial variation of classification uncertainty.

Identifying those areas of the channel where classifications were least certain could also be used to infer boundaries of variable width (Burrough 1996). To explore this possibility, the maximum and second highest fuzzy membership values at each location were compared using difference and ratio forms of Burrough's confusion index:

$$CI_{diff} = 1 - (maxfuz - maxfuz2) \text{ or } CI_{ratio} = maxfuz2/maxfuz.$$

3. Results

3.1 *Crisp, Supervised Classification of In-stream Habitat: Lamar River Probe-1 Scene*

Traditional classifications of in-stream habitat were produced by applying a maximum likelihood procedure to the first ten MNF-transformed bands of the masked Probe-1 imagery using training data derived from field maps of the Lamar River. Both 4- and 7-unit classifications were generated, and an accuracy assessment including all pixels not used as training sites resulted in low overall accuracies of 64.7% and 47.5%, respectively (Table 3). Although Marcus *et al.* (2003) obtained 85.5% accuracy by applying a 2 m buffer to field map polygons, the reduction in accuracy incurred when transitional boundary zones between habitat units are not excluded suggests that conventional image classification procedures might be inadequate for characterizing the full continuum of fluvial form. Supervised classifications are only as valid as the training sites from which they are derived, and sophisticated technology can not resolve the ambiguity among similar in-stream habitats. Low overall accuracies could thus be the inevitable consequence of applying a subjective field classification scheme to polythetic natural phenomena characterized by indeterminate boundaries (Burrough 1996). These complications render the Boolean, object-based data model inappropriate, and poor accuracies could result from the inadequacy of the representation rather than a deficiency in the remotely sensed data or classification algorithm.

100659 International Journal of Geographical Information Science gis51069.3d 17/11/04 11:15:03
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

Table 3. Confusion matrices and accuracy assessment statistics for 4- and 7-unit supervised maximum likelihood classifications of the Lamar River Probe-1 scene (first ten MNF bands). Habitat unit abbreviations: EDZ=eddy drop zone; HGR=high-gradient riffle; LGR=low-gradient riffle; RW run=rough-water run.

	Field-Mapped In-Stream Habitat				
Image Classification	EDZ	Riffle	Run/Glide	Pool	Total
EDZ	590	710	1798	8	3106
Riffle	57	6940	3710	40	10747
Run/Glide	133	3070	20936	48	24187
Pool	22	2057	4164	496	6739
Total	802	12777	30608	592	44779

Overall Accuracy=64.7% (28962/44779); Kappa Coefficient=0.3682

	Field-Mapped In-Stream Habitat							
Image Classification	EDZ	HGR	LGR	Glide	Run	RW run	Pool	Total
EDZ	586	93	527	794	298	127	9	2434
HGR	14	1057	761	144	393	100	12	2481
LGR	43	823	4387	1929	596	320	26	8124
Glide	93	289	594	7713	1230	416	35	10370
Run	15	273	562	2498	6055	638	46	10087
Rwrun	44	356	1673	1981	1487	1029	24	6594
Pool	7	265	1117	1826	736	298	440	4689
Total	802	3156	9621	16885	10795	2928	592	44779

Overall Accuracy=47.5% (21267/44779); Kappa Coefficient=0.3465

3.2 *Fuzzy c-means Classifications*

Given these concerns with the crisp field maps and image classifications, a series of fuzzy classifications were generated using a fuzzy *c*-means algorithm implemented in MATLAB. Figure 1 presents an example from the River2D hydraulic model of the Kananaskis River, illustrating the spatial distribution of fuzzy clusters. In this raster representation of in-stream habitat, each 1 m^2 area of the channel was assigned a membership value in each of the four clusters, corresponding to the images in Figure 1. This allowance for partial membership in multiple classes provided a marked improvement over conventional object-based habitat maps. Rather than forcing the river continuum into a finite set of rigid classes separated by abrupt boundaries, these fuzzy partitions enabled a more faithful depiction of gradual transition zones. The fuzzy clusters illustrated in Figure 1 were spatially continuous, compact, and hydraulically reasonable. Phase-space plots of modeled flow depth and velocity (Figure 2) could be used to interpret cluster centers in ecologically meaningful terms. Similar fuzzy classifications were also produced from the ADAR multispectral and Probe-1 hyperspectral imagery of the Lamar River.

3.3 *Assessment of Cluster Validity*

Anticipating the sensitivity of the clustering algorithm to the number of clusters c and weighting exponent m, FCM classifications of each data set were generated using various (c, m) combinations and eight cluster validity functions were used to evaluate the effects of these two parameters (Table 4). Typically, the minimum or

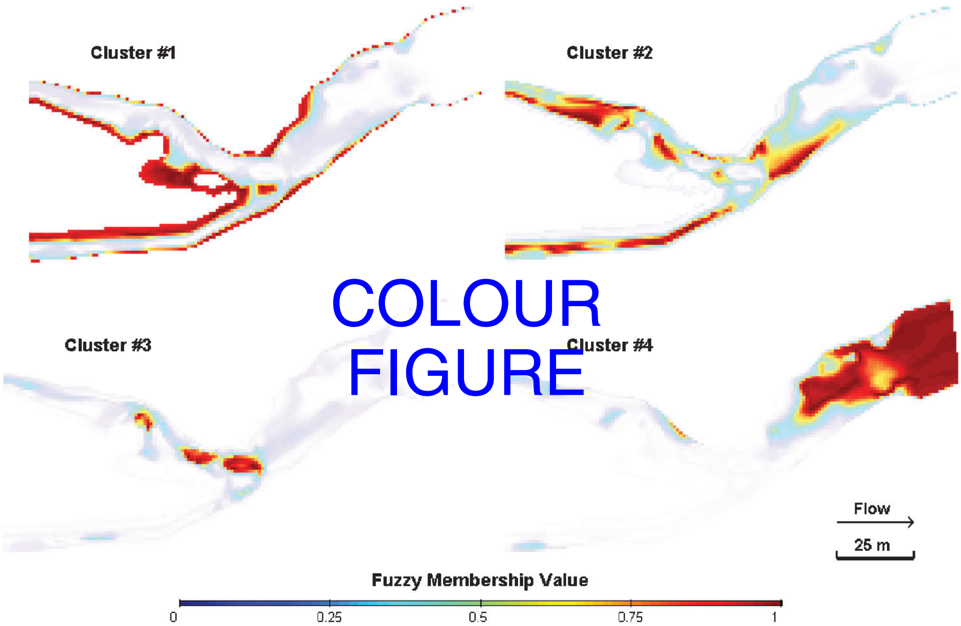

Figure 1. Fuzzy *c*-means classification of the Kananaskis River derived from the River2D hydrodynamic model. Transparency of images is scaled to the fuzzy membership value. Weighting exponent $m=2$.

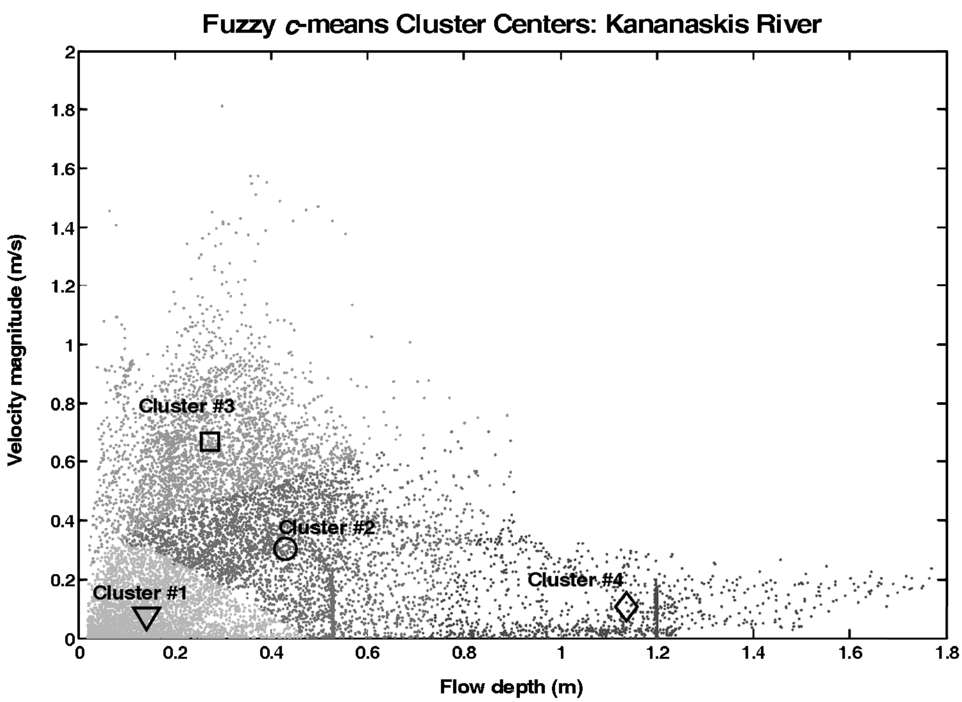

Figure 2. Phase-space plot of cluster centers from a fuzzy *c*-means classification of the River2D hydrodynamic model results (weighting exponent $m=2$). Cluster numbers correspond to the images in Figure 1.

International Journal of Geographical Information Science gis51069.3d 17/11/04 11:15:07
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)
100659

Table 4. Cluster validity indices at their minima/maxima, along with the number of classes c and weighting exponent m at these extrema.

	River2D Hydraulic Model: Kananaskis River, Alberta			Probe-1 Hyperspectral: Upper Lamar River, WY			Probe-1 Hyperspectral: Lower Lamar River, WY			ADAR Multispectral: Lamar River, WY		
	c	m	Index Value	c	m	Index Value	c	m	Index Value	c	m	Index Value
Partition Coefficient	2	1.2	0.983	2	1.1	0.976	2	1.1	0.976	2	1.1	0.993
Partition Entropy	2	1.1	0.030	2	1.1	0.041	2	1.1	0.040	2	1.1	0.012
Xie-Beni Function	10	2.5	0.000	10	2.5	32.183	9	2.5	116.948	10	2.5	21.930
Fukuyama-Sugeno	10	1.1	−2111.093	10	1.1	−1.663E+06	10	1.1	−1.52E+06	10	1.1	−3.07E+06
Fuzzy Hypervolume	2	1.2	0.000	2	1.3	6607.874	3	1.1	20092.121	2	2.4	1233.958
Avg. Partition Density	6	1.2	1.661E+09	7	1.3	0.416	8	1.2	0.108	7	1.2	20.402
Partition Density	2	1.2	1.278E+08	3	1.3	0.278	3	1.1	0.062	2	2.4	2.846
Separate-Compact	5	2.5	2.567	2	1.3	0.073	3	1.4	0.050	6	2.5	3.774

maximum of a cluster validity function is assumed to correspond to a "good" partition (Table 2 and references therein), but many of these criteria exhibited monotonic trends with increasing numbers of clusters (Halkidi *et al.* 2001); their asymptotic behavior as $m \rightarrow 1$ (a hard partition) or as $m \rightarrow \infty$ (infinitely fuzzy) also limited the utility of some indices (Pal and Bezdek 1995).

Nevertheless, cluster validity functions can be used to guide selection of c and m parameters under appropriate circumstances. For example, the separateness-compactness index *SC* (Zahid *et al.* 1999) can be computed for various (c, m) pairs to produce a three-dimensional surface (left panel, Figure 3). Transects across this surface can then be used to infer an appropriate level of fuzziness m for a specified number of clusters or, for a given weighting exponent, to determine an optimal number of clusters c. In the case of the ADAR imagery, for a four-unit classification the *SC* index attains a maximum for the fuzziest possible partition (m=2.5), suggesting that most pixels will have membership values relatively evenly distributed across the four clusters. For a fixed m value of 2 (the default in many FCM implementations), *SC* is maximized for three clusters. Of course these results are also a function of the input data set, and could be due to insufficient spectral detail for discriminating among in-stream features (Wright *et al.* 2000, Legleiter *et al.* 2002). In practice, cluster validity criteria like the *SC* index are best used as a guide to supplement field experience and prior knowledge of the stream under study.

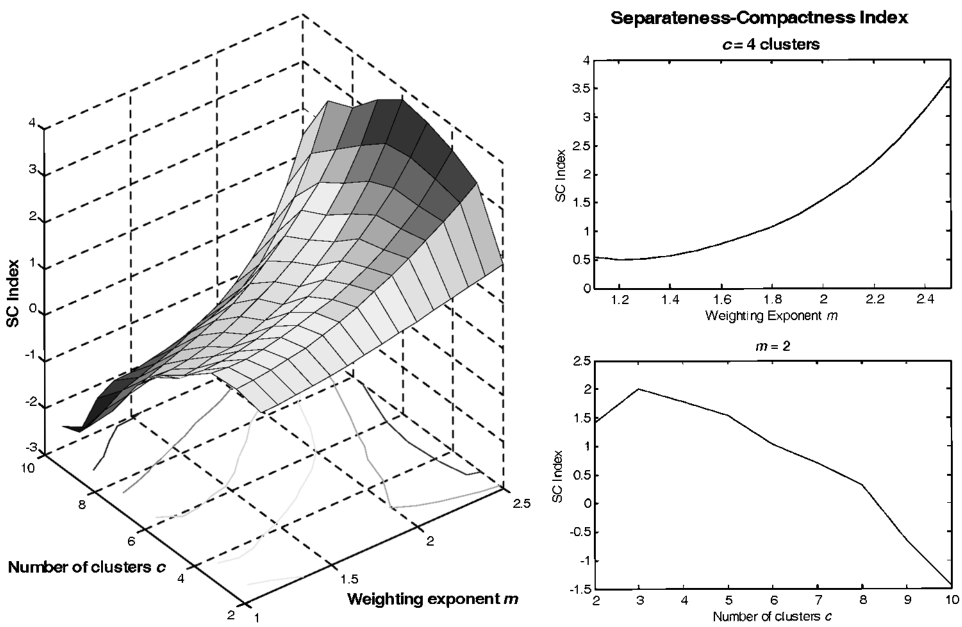

Figure 3. Separateness-compactness cluster validity index (Zahid *et al.* 1999) for fuzzy c-means classifications of the Lamar River derived from ADAR multispectral imagery, plotted as a function of the number of clusters c and the weighting exponent m in the FCM algorithm. Higher values of the SC index tend to indicate better fuzzy partitions. This criterion can be used to select an appropriate degree of fuzziness for a specified number of classes (upper right) or an optimal number of classes for a fixed weighting exponent (lower right).

3.4 *α-Cuts and Classification Hardening*

Although fuzzy classifications appeared to provide a more realistic, more uncertain representation of the fluvial environment, a traditional crisp classification is often desirable for purposes of display and summary. Fuzzy clustering results were hardened by assigning each pixel to the cluster in which its membership value was highest and vector maps preserving the spatial pattern of classification uncertainty were obtained by applying α-cut thresholds. Only those pixels with maximum fuzzy membership values exceeding the specified α were retained in these crisp classifications, and those that did not meet this criterion became gaps (i.e., epsilon bands) in the hardened image. Figure 4 provides an example from the upper Lamar River Probe-1 scene, with the width of the epsilon bands clearly increasing for higher α-levels. Higher α thresholds mandate a greater degree of classification certainty and areas of the channel that cannot be allocated to any single class with that level of confidence remain unclassified. A tradeoff must therefore be made between the certainty with which pixels can be assigned to classes and the proportion of the channel to be retained in a crisp classification.

This compromise will also depend upon the input data set and the c and m parameters used in the FCM algorithm. Figure 5 plots the proportion of the channel that would be excluded from a crisp classification created by applying various α-cut thresholds to each of the four stream reaches examined in this study. These results suggest that, for a given α threshold, data acquired by the 4-band ADAR sensor could be more readily assigned to crisp classes than the Probe-1 hyperspectral imagery, with the River2D hydraulic model intermediate (Table 5). The contrast between the upper and lower subsets of the Probe-1 Lamar River scene indicates

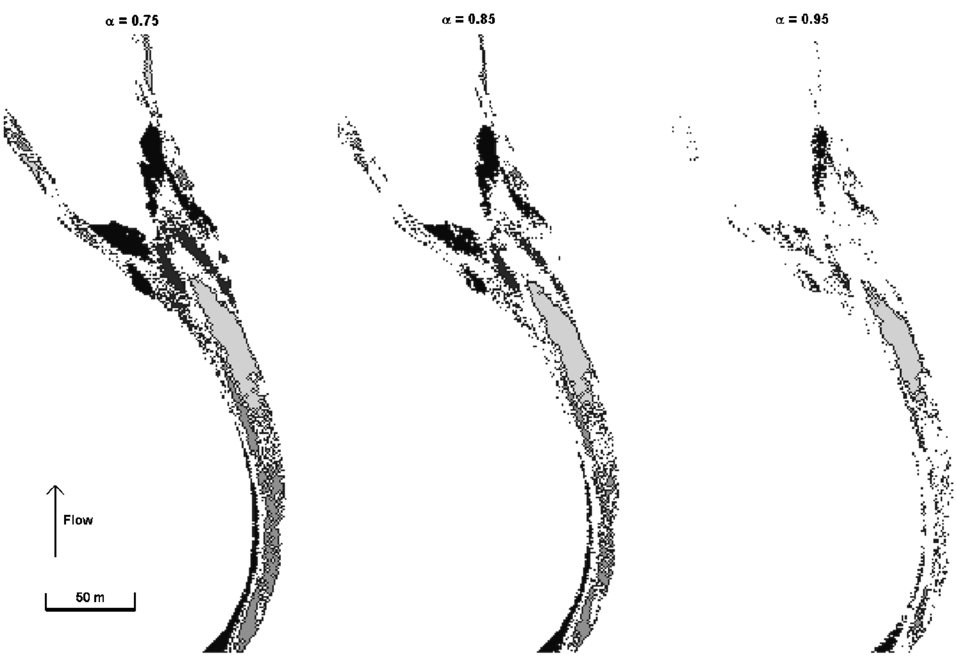

Figure 4. α-cuts and resulting crisp classification of the Lamar River derived from Probe-1 hyperspectral imagery. Portions of the channel with maximum fuzzy membership values less than α indicate zones of uncertainty. Weighting exponent m=1.5.

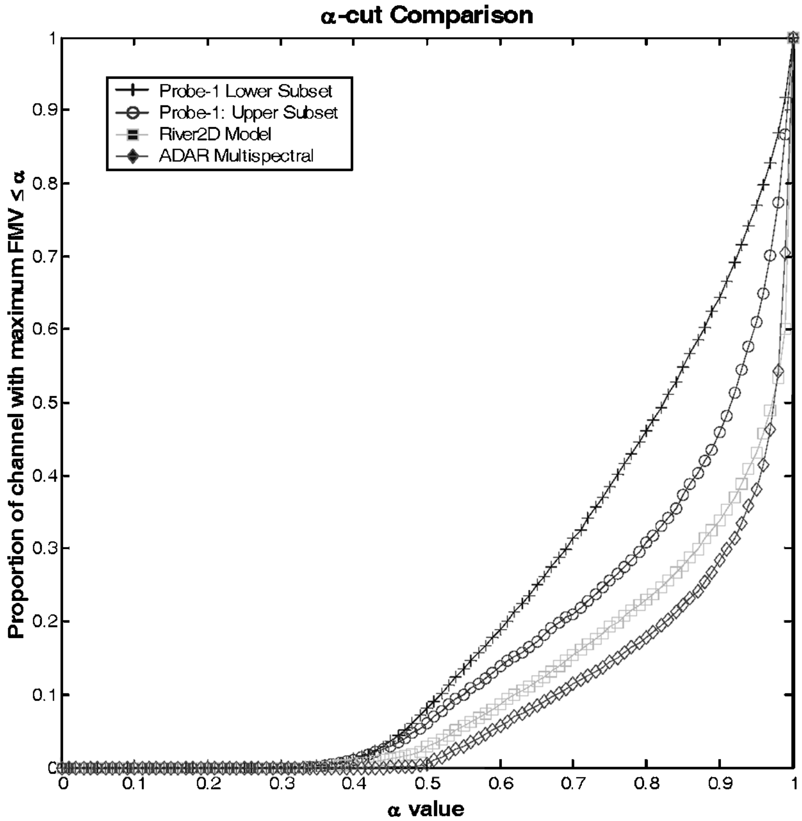

Figure 5. α-cut comparisons of the four stream segments examined in this study, using fuzzy c-means classification with 4 clusters and a weighting exponent m of 1.5. These curves illustrate the proportion of the channel that would be excluded from a crisp classification created by taking an α-cut at the corresponding maximum fuzzy membership value threshold.

that fuzzy c-partitions and derived α-cut classifications are highly site- (and data-) specific. This observation raises the important point that these numerical procedures are entirely data-driven and thus highly sensitive to data quality. In the ADAR scene, for example, the concentrated variability of membership values along shallow channel margins might have been due to portions of the bank included in the stream mask. The linear geometry of rivers dictates that a sizable fraction of pixels will be mixed, particularly for narrow channels or low-flow conditions. Practical

Table 5. Results of defuzzification at different α-cut thresholds for each data set (c=4 clusters, weighting exponent m=1.5). Table lists the proportion of channel with maximum fuzzy membership values exceeding the specified α. Those areas excluded from the resulting crisp classification represent ambiguous, complex, or gradational portions of the channel.

α Cut Threshold	River2D Model	Probe-1: Upper	Probe-1: Lower	ADAR
0.75	0.8081	0.7438	0.6156	0.8563
0.8	0.7708	0.6925	0.5394	0.8213
0.85	0.723	0.6272	0.4522	0.7775
0.9	0.6613	0.5413	0.3567	0.7167
0.95	0.5693	0.3907	0.2298	0.6187

application of fuzzy logic in the context of stream classification thus requires familiarity with the limitations of the available data.

For a given data set and FCM algorithm (*c* and *m* parameters), graphs like Figure 5 can be used to determine the appropriate α threshold for assigning a user-specified proportion of the channel to crisp classes. Conversely, given a proportion, the plot could be used to estimate the level of uncertainty that would be introduced by allocating that fraction of the channel to discrete classes. Graphs like Figure 5 thus indicate how much of the channel *cannot* be allocated to any single class at the specified α level, potentially an important piece of information. These unclassified areas might contain several habitat types, features unlike those found elsewhere in the channel, or variability at a scale finer than the spatial resolution of the digital imagery or finite element mesh. In any case, the heterogeneity, uniqueness, and/or complexity of these zones of ambiguity make them, in a sense, the most interesting portion of the stream. From an ecological perspective, these areas might also be the most valuable, perhaps representing stream segments with a greater diversity of flow conditions and substrate characteristics. The ability to identify and monitor these critical habitats thus represents a significant potential advantage of fuzzy approaches to stream classification.

3.5 *Spatial Patterns of Classification Uncertainty*

Transitional zones of increased habitat diversity can also be inferred from maps of classification uncertainty. Crisp, maximum likelihood classifications of the Lamar River Probe-1 scene, based on the field-mapped in-stream habitats, were used to calculate Zhu's (1997, 2001) classification entropy and exaggeration uncertainty indices (Figure 6). Although the assignment of individual 1 m^2 pixels to classes was a marked improvement over the channel-spanning polygons of the field map, these results, together with low overall accuracies (Table 3), suggested that a rigid Boolean model might be inappropriate. Uniformly high exaggeration uncertainty implied that most pixels were only weakly similar to the class to which they were

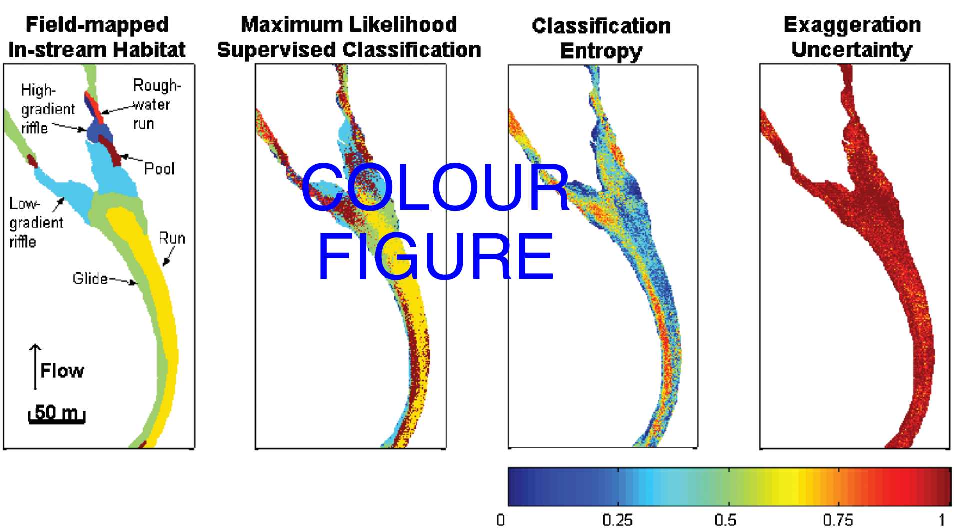

Figure 6. Maximum likelihood supervised classification of the upper reach of the Lamar River Probe-1 scene and associated spatial patterns of classification uncertainty.

assigned, and high classification entropy indicated that much of the channel could be considered a partial member of several habitat types. These results suggested that constraining the remotely sensed data to a discrete collection of arbitrarily defined classes and assigning each pixel to only one of these categories failed to fully capitalize on the information content of the imagery and/or to characterize the full spectrum of habitat conditions.

Fuzzy classifications allowing for continuous variation of membership among clusters and across space thus appeared to provide a more realistic representation of the complex fluvial environment. Burrough's (1996) observation that boundaries "concentrate confusion in the smallest zone possible" (p. 27) also implied that maps of classification uncertainty could be used to identify the most "confused" areas of the channel and, presumably, boundaries among habitat units. An example is given in Figure 7, where uncertainty indices calculated from a fuzzy classification of the lower reach of the Lamar River are displayed as three-dimensional surfaces. Those portions of the channel exhibiting the greatest degree of classification uncertainty can thus be visualized as topographic highs, whereas areas that could be relatively confidently assigned to individual clusters appear as valleys. Extending this analogy, the distinct ridge visible at the lower right of the classification entropy and exaggeration uncertainty images could be interpreted as a relatively sharp boundary separating distinctive habitat features. At the upper end of the reach (far upper left on the images), however, the presence of a high plateau suggests that this zone of flow convergence could contain a variety of habitat conditions. The degree to which

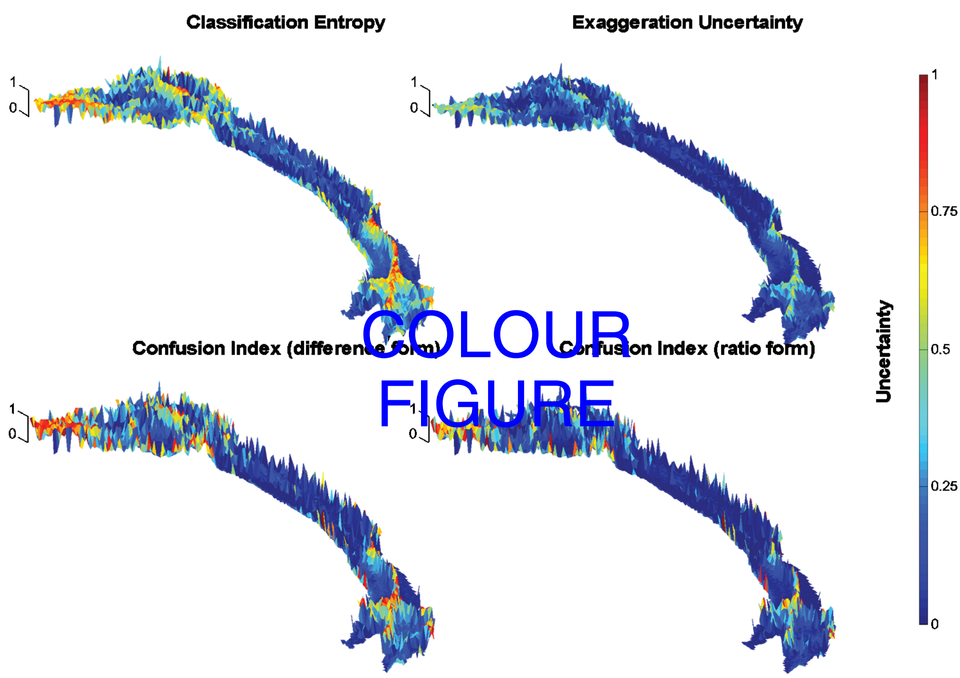

Figure 7. Spatial variation of classification uncertainty on the lower reach of the Lamar River, with areas exhibiting the greatest uncertainty represented by topographic highs. Fuzzy c-means classification derived from Probe-1 hyperspectral imagery, with c=4 clusters and weighting exponent m=1.5. Flow direction is toward the lower right of each image, and the reach is approximately 245 m long.

the stream can be bounded thus appeared to be highly spatially variable in its own right.

For applications where vector-based representations are desired, boundaries could thus be obtained by isolating high values of these uncertainty measures. An important advantage of this approach is the allowance for boundaries of variable width, wider where the environment is more heterogeneous or complex and narrower where adjacent habitat types are relatively distinct from one another. Although the mere notion of imposing boundaries on complex natural phenomena like the Lamar and Kananaskis Rivers introduces a certain degree of subjectivity, raster-based, fuzzy representations of these phenomena allow these boundaries to be produced in a defined, internally consistent manner while retaining quantitative information on the associated uncertainty.

100659

International Journal of Geographical Information Science gis51069.3d 17/11/04 11:15:17
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

4. Discussion

4.1 *Limitations of the Technique*

Although a fuzzy representation of in-stream habitat appeared to possess several important advantages, this approach introduces a unique suite of problems as well. Foremost among these is that of data quality. The fuzzy *c*-means algorithm, like any unsupervised clustering procedure, will be highly sensitive to flaws in the data. For the ADAR multispectral image examined in this study, the definition of the mask for the active channel of the Lamar River appeared to be problematic, and the algorithm, seeking numerical structure rather than physical meaning, highlighted mixed pixels along the bank rather than habitats within the stream. Any other sources of error, such as instrument noise, will be transferred to the clustering results and diminish their interpretability. In this sense, a potential advantage of more advanced hyperspectral instruments lies in the redundancy of the information they provide, which allows for noise removal techniques like the MNF transform. In any case, users must be not just conscientious but paranoid about the data they deliver to fuzzy clustering algorithms.

Granting that premise, remotely sensed data are perhaps not the ideal tool for implementing fuzzy classification of in-stream habitats. As Poole *et al.* (1997) have emphatically noted, efforts to inventory and monitor streams should be based on "direct, repeatable, … quantitative measures" (p. 879), and remote sensing is, by definition, not direct. Although previous research has demonstrated the feasibility of mapping fluvial environments from image data, the underlying physical processes governing the remotely sensed signal remain poorly understood. The upwelling spectral radiance from a stream channel is a complex function of depth, surface turbulence, substrate characteristics, suspended sediment, and perhaps other factors. Establishing theoretically sound linkages between this integrated spectral response and ecologically relevant properties of the channel thus becomes a priority for future research.

Given this inherent limitation of remotely sensed data, the River2D hydrodynamic model was perhaps the most useful data set. Though closely linked to the habitat characteristics of interest, hydraulic modeling is a complicated endeavor in its own right. Intensive field survey data are required to adequately characterize channel topography, and the selection of appropriate flow resistance and eddy viscosity parameters is typically an exercise in trial and error (Miller and Cluer

1998). Perhaps more importantly if long-term or large-scale management is a primary objective, modeled reaches are typically less than 100 m in length (Crowder and Diplas 2000). Ultimately, a combination of fieldwork, modeling, and image processing will likely prove most powerful.

100659

International Journal of Geographical Information Science gis51069.3d 17/11/04 11:15:17
The Charlesworth Group, Wakefield +44(0)1924 369598 - Rev 7.51n/W (Jan 20 2003)

4.2 *Recommendations and Future Research Directions*

While remote sensing, hydraulic modeling, and fuzzy logic all have the potential to advance our understanding of fluvial processes and the aquatic habitat they create, field work must remain a central component of any stream research program. Fuzzy classifications of remotely sensed data will achieve their greatest utility when used to complement ground-based measurement campaigns. Such an effort would be truly synergistic. Field personnel could provide biologically meaningful interpretations of the algorithmically generated clusters, with the remotely sensed data being used to direct efficient sampling strategies. Specifically, maps of classification uncertainty could be used to identify heterogeneous or complex areas of the channel, which ground crews could then examine in greater detail. In this manner, intensive measurements performed at certain critical locations could then be extrapolated to longer reaches and even entire watersheds on the basis of spectrally-driven classifications.

The ideas presented here remain in a developmental stage, and additional research will be required before these techniques can be implemented in practice. If remote sensing is to become a useful tool for characterizing aquatic habitat, physically-based models describing the interaction of light with an irregular water surface, an optically complex water column, and a heterogeneous substrate, must be developed. Incorporating spatial information is a related goal, and geostatistical methods could be used to produce continuous maps of depth and substrate characteristics from ground-based point measurements and remotely sensed data. If accurate estimates of water depth can be obtained from imagery of sufficient spatial resolution, remotely sensed data could conceivably act as an input to hydraulic modeling.

5. Conclusion

Dynamic fluvial systems cannot be adequately represented by conventional, object-based data models and this study presented fuzzy set theory as an alternative to this rigid framework. We utilized hydrodynamic modeling and remotely sensed data to explore different methods of characterizing aquatic habitat in the Lamar and Kananaskis Rivers. These environments are problematic for GIScience due to their linear geometric configuration, scale-dependent dynamic behavior, closely coupled physical and biological processes, and by the lack of a widely accepted ontology. Our comprehensive approach included ground-based mapping, supervised image classification, fuzzy clustering, and quantitative assessment of both cluster validity and spatial patterns of classification uncertainty. Introducing these techniques to the fluvial environment constitutes an innovative application of GIScience and an advance in stream studies.

Partitions generated by the fuzzy c-means algorithm allowed for partial membership in multiple classes and gradual transitions between adjacent habitat types, an important advantage relative to crisp field map polygons and supervised image classifications. Cluster validity indices can guide selection of an appropriate number of clusters and/or degree of fuzziness, but data quality must be a crucial consideration in any numerical clustering procedure. Crisp, vector representations can be derived from raster fuzzy classification results by establishing a threshold maximum membership value. Higher thresholds will result in smaller, more homogenous conditional objects separated by more extensive zones of ambiguity, and a compromise must be reached between the proportion of the stream retained in a crisp classification and the uncertainty involved in assigning pixels to a single class. Spatial patterns of classification uncertainty highlight ambiguous transitional areas that might contain a greater diversity of habitat conditions and can be used to infer natural boundaries of variable width.

At present, because hydrodynamic modeling directly quantifies the flow conditions determining habitat suitability for various organisms, it is perhaps more valuable than remotely sensed data for reach-scale studies. The synoptic perspective afforded by remote sensing constitutes an important advantage, however, and we suggest that a synergistic combination of field work, modeling, and image processing will ultimately be most useful. Although the application of these techniques to in-stream habitat mapping remains in a developmental stage, fuzzy classification and spatially explicit, quantitative assessment of uncertainty can be used to provide a more realistic alternative representation of the fluvial environment.

Acknowledgements

The Probe-1 data collected by Earth Search Science, Inc., and ADAR imagery acquired by Positive Systems, Inc., were purchased through a NASA EOCAP grant (Stennis Space Flight Center) administered through Yellowstone Ecosystem Studies. Special thanks are due to W. Andrew Marcus and Robert L. Crabtree for the opportunity to pursue this project and to Jim Rasmussen and Rob Ahl for their help in the field. Peter Steffler made the Kananaskis River data set available, Jingxiong Zhang provided valuable advice, Phaedon Kyriakidis developed sections of code adapted for use in this study, and three thoughtful reviewers helped improve the original manuscript. The American Society for Engineering Education, National Science Foundation, and California Space Institute provided financial support.

References

Bezdek, J.C., 1981, *Pattern recognition with fuzzy objective function algorithms* (New York: Plenum Press).

Bezdek, J.C., Ehrlich, R. and Full, W., 1984, FCM - the Fuzzy C-Means Clustering-Algorithm. *Computers & Geosciences*, **10**, pp. 191–203.

Bezdek, J.C. and Pal, S.K., 1992, *Fuzzy models for pattern recognition: methods that search for structures in data* (New York: The Institute of Electrical and Electronics Engineers, Inc.).

Bisson, P.A., Nielsen, J.L., Palmason, R.A. and Grove, L.E., 1982, A system of naming habitat types in small streams, with examples of habitat utilization by salmonids during low streamflow. In *Acquisition and utilization of aquatic habitat inventory information*, by N.A. Armatrout, editors (Bethesda, Maryland: American Fisheries Society), pp. 62–73.

Blackburn, J. and Steffler, P., 2002, River2D tutorial - the basics, http://www.river2d.ualberta.ca/

BOARDMAN, J.W. and KRUSE, F.A., 1994, *Automated spectral analysis: a geological example using AVIRIS data, north Grapevine Mountains, Nevada*, In ERIM Tenth Thematic Conference on Geologic Remote Sensing (San Antonio, TX: Environmental Research Institute of Michigan), pp. I-407–I-418.

BUKATA, R.P., JEROME, J.H., KONDRATYEV, K.Y. and POZDNYAKOV, D.V., 1995, *Optical Properties and Remote Sensing of Inland and Coastal Waters* (Boca Raton, FL: CRC Press).

BURROUGH, P.A., 1989, Fuzzy mathematical methods for soil survey and land evaluation. *Journal of Soil Science*, **40**, pp. 477–492.

BURROUGH, P.A., 1996, Natural Objects with Indeterminate Boundaries. In *Geographic Objects with Indeterminate Boundaries*, by P.A. Burrough and A.U. Frank, editors (Bristol, PA: Taylor and Francis), pp. 3–28.

BURROUGH, P.A. and FRANK, A.U., 1996, *Geographic Objects with Indeterminate Boundaries* (Bristol, PA: Taylor and Francis).

CHENG, T., MOLENAAR, M. and LIN, H., 2001, Formalizing fuzzy objects from uncertain classification results. *International Journal of Geographical Information Science*, **15**, pp. 27–42.

CONGALTON, R.G. and GREEN, K., 1999, *Assessing the acuracy of remotely sensed data: principles and practices* (Boca Raton, Florida: CRC Press).

CROWDER, D.W. and DIPLAS, P., 2000, Using two-dimensional hydrodynamic models at scales of ecological importance. *Journal of Hydrology*, **230**, pp. 172–191.

FISHER, P., 1996, Boolean and Fuzzy Regions. In *Geographic Objects with Indeterminate Boundaries*, by P.A. Burrough and A.U. Frank, editors (Bristol, PA: Taylor and Francis), pp. 87–94.

FISHER, P.F., 1998, Is GIS hidebound by the legacy of cartography? *Cartographic Journal*, **35**, pp. 5–9.

FRISSELL, C.A., LISS, W.J., WARREN, C.E. and HURLEY, M.D., 1986, A Hierarchical Framework for Stream Habitat Classification - Viewing Streams in a Watershed Context. *Environmental Management*, **10**, pp. 199–214.

GATH, I. and GEVA, A.B., 1989, Unsupervised optimal fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **PAMI-11**, pp. 773–781.

GHANEM, A., STEFFLER, P., HICKS, F. and KATOPODIS, C., 1996, Two-dimensional hydraulic simulation of physical habitat conditions in flowing streams. *Regulated Rivers-Research & Management*, **12**, pp. 185–200.

GOODWIN, C.N., 1999, Fluvial classifcitation: Neanderthal necessity or needless normalcy. In *Wildland Hydrology*, by D.S. Olson and J.P. Potyondy, editors (Middleburg, Virginia: American Water Resources Association), pp. 229–236.

HALKIDI, M., BATISTAKIS, Y. and VAZIRGIANNIS, M., 2001, On clustering validation techniques. *Journal of Intelligent Information Systems*, **17**, pp. 107–145.

HAWKINS, C.P., KERSHNER, J.L., BISSON, P.A., BRYANT, M.D., DECKER, L.M., GREGORY S.V., MCCULLOUGH D.A., OVERTON, C.K., REEVES, G.H., STEEDMAN R.J. and YOUNG, M.K., 1993, A hierarchical approach to classifying stream habitat features. *Fisheries*, **18**, pp. 3–12.

KANDEL, A., 1986, *Fuzzy mathematical techniques with applications* (Reading, Massachusetts: Addison-Wesley Publishing Company).

KATOPODIS, C., 2003, 2-D ice-free and ice-covered hydrodynamic modeling of riverine habitat in the Canadian Prairies. In *Proceedings of International IFIM User's Conference Fort Collins, CO, June 1–5, 2003*, http://www.mesc.usgs.gov/conferences/ifimconf/IFIMGraphic.asp?PresID=13, accessed January 5, 2004.

KNIGHTON, A.D., 1998, *Fluvial forms and processes: a new perspective* (New York: Oxford University Press).

LADD, S.C., MARCUS, W.A. and CHERRY, S., 1998, Differences in trace metal concentrations among fluvial morphologic units and implications for sampling. *Environmental Geology*, **36**, pp. 259–270.

Lamouroux, N., Capra, H. and Pouilly, M., 1998, Predicting habitat suitability for lotic fish: Linking statistical hydraulic models with multivariate habitat use models. *Regulated Rivers-Research & Management*, **14**, pp. 1–11.

Legleiter, C.J., Marcus, W.A. and Lawrence, R., 2002, Effects of sensor resolution on mapping in-stream habitats. *Photogrammetric Engineering and Remote Sensing*, **68**, pp. 801–807.

Legleiter, C.J., 2003, Spectrally driven classification of high spatial resolution, hyperspectral imagery: a tool for mapping in-stream habitat. *Environmental Management*, **32**, pp. 399–411.

Malmqvist, B. and Rundle, S., 2002, Threats to the running water ecosystems of the world. *Environmental Conservation*, **29**, pp. 134–153.

Marcus, W.A., 2002, Mapping of stream microhabitats with high spatial resolution hyperspectral imagery. *Journal of Geographical Systems*, **4**, pp. 113–126.

Marcus, W.A., Legleiter, C.J., Aspinall, R., Boardman, J.W. and Crabtree, R., 2003, High spatial resolution hyperspectral mapping of in-stream habitats, depths, and woody debris in mountain streams. *Geomorphology*, **55**, pp. 363–380.

Miller, A.J. and Cluer, B.L., 1998, Modeling considerations for simulation of flow in bedrock channels. In *Rivers Over Rock*, by K.J. Tinkler and E.E. Wohl, editors (Washington, D.C.: American Geophysical Union), pp. 61–104.

Montgomery, D.R. and Buffington, J.M., 1997, Channel-reach morphology in mountain drainage basins. *Geological Society of America Bulletin*, **109**, pp. 596–611.

Newson, M.D. and Newson, C.L., 2000, Geomorphology, ecology and river channel habitat: mesoscale approaches to basin-scale challenges. *Progress in Physical Geography*, **24**, pp. 195–217.

Pal, N.R. and Bezdek, J.C., 1995, On Cluster Validity for the Fuzzy C-Means Model. *Ieee Transactions on Fuzzy Systems*, **3**, pp. 370–379.

Poole, G.C., Frissell, C.A. and Ralph, S.C., 1997, In-stream habitat unit classification: Inadequacies for monitoring and some consequences for management. *Journal of the American Water Resources Association*, **33**, pp. 879–896.

Richards, J.A. and Jia, X., 1999, *Remote sensing digital image analysis: an introduction* (New York: Springer).

Roper, B.B., Kershner, J.L., Archer, E., Henderson, R. and Bouwes, N., 2002, An evaluation of physical stream habitat attributes used to monitor streams. *Journal of the American Water Resources Association*, **38**, pp. 1637–1646.

Rosgen, D.L., 1994, A classification of natural rivers. *Catena*, **22**, pp. 169–199.

Steffler, P. and Blackburn, J., 2002, River2D two-dimensional depth averaged model of river hydrodynamics and fish habitat, http://www.river2d.ualberta.ca/

Ward, J.V., 1989, The 4-Dimensional Nature of Lotic Ecosystems. *Journal of the North American Benthological Society*, **8**, pp. 2–8.

Whited, D., Stanford, J.A. and Kimball, J.S., 2002, Application of airborne multispectral digital imagery to quantify riverine habitats at different base flows. *River Research and Applications*, **18**, pp. 583–594.

Whiting, P.J. and Bradley, J.B., 1993, A process-based classification system for headwater streams. *Earth Surface Processes and Landforms*, **18**, pp. 603–612.

Wohl, E., 2000, *Mountain Rivers* (Washington, D.C.: American Geophysical Union).

Wright, A., Marcus, W.A. and Aspinall, R., 2000, Evaluation of multispectral, fine scale digital imagery as a tool for mapping stream morphology. *Geomorphology*, **33**, pp. 107–120.

Xie, X.L.L. and Beni, G., 1991, A Validity Measure for Fuzzy Clustering. *Ieee Transactions on Pattern Analysis and Machine Intelligence*, **13**, pp. 841–847.

Zahid, N., Limouri, N. and Essaid, A., 1999, A new cluster-validity for fuzzy clustering. *Pattern Recognition*, **32**, pp. 1089–1097.

ZHANG, J.X. and STUART, N., 2001, Fuzzy methods for categorical mapping with image-based land cover data. *International Journal of Geographical Information Science*, **15**, pp. 175–195.

ZHU, A.X., 1997, Measuring uncertainty in class assignment for natural resource maps under fuzzy logic. *Photogrammetric Engineering and Remote Sensing*, **63**, pp. 1195–1202.

ZHU, A.-.X., 2001, Modeling spatial variation of classification accuracy under fuzzy logic. In *Spatial uncertainty in ecology*, by C.T. Hunsacker, M.F. Goodchild, M.A. Friedl and T.J. Case, editors (New York: Springer-Verlag), pp. 330–350.