MAPPING OF SALMON HABITAT PARAMETERS USING DIGITAL AIRBORNE IMAGERY

THOMAS PUESTOW

Mapping of Salmon Habitat Parameters Using Digital Airborne Imagery

by

Thomas Puestow

A thesis submitted to the School of Graduate
Studies in partial fulfillment of the
requirements for the degree of
Master of Science

Department of Geography
Memorial University of Newfoundland
1998

St. John's Newfoundland

Abstract

This study focuses on the application of airborne remote sensing and image classification to the mapping of bottom substrate, channel pattern and land cover as important freshwater habitat parameters for Atlantic salmon. A Compact Airborne Spectrometric Imager (CASI) was used to collect multispectral image data with approximately 20 nm wide bands centred at wavelengths of 510, 590, 660 and 730 nm. Image preprocessing included a first order atmospheric correction for path radiance and geometric registration to the UTM reference system. Numerical transforms on the imagery included principal component transformations on original and logarithmized spectral bands, as well as the derivation of a normalized difference vegetation index (NDVI). Ancillary information consisted of valley gradient and stream width. Valley gradient was derived from elevation data contained in a 1:50,000 digital map sheet. Stream width was extracted from the image data. The river course was divided in sections of approximately equal length (30 m), and the average width of each segment was calculated from its length and area. The importance of individual predictor variables for the extraction of the habitat parameters was established using the mean response for each predictor variable, standardized distance matrices and plots of group variability. Separate image classifications were carried out for substrate type, channel pattern and land cover using a hierarchical decision tree algorithm. The end nodes of the final classification trees were implemented as classification rules in a FORTRAN program. Classification accuracy was assessed using an independently collected test sample. The observed overall classification accuracies were 66.87 %, 38.11 % and 84.91 % for

substrate type, channel pattern and land cover, respectively. Overall accuracy was significantly improved for the habitat parameters substrate type and channel pattern by combining categories of these variables according to their significance in designating suitable spawning habitat. The revised overall accuracy values for these habitat parameters were 73.76 % and 64.47 %, respectively. Finally, substrate type and channel pattern were combined to create composite maps of spawning habitat suitability. The resulting stratification of salmon spawning habitats corresponds well with the findings of earlier investigations. Therefore, the value of the methodology developed in this study for the management and protection of freshwater salmon habitat was successfully demonstrated.

Acknowledgments

I would like to thank my supervisor Dr. Elizabeth L. Simms for her continued advice and support throughout this program. Many thanks also to my co-supervisors, Dr. Alvin Simms and Ms. Karyn Butler for valuable input and suggestions. Financial support for this research was obtained from Department of Fisheries and Oceans Canada. In addition, financial assistance was provided by the School of Graduate Studies and GEOIDAL, Memorial University of Newfoundland. All analysis was carried out using the facilities provided by GEOIDAL.

Table of Contents

List of Tables

List of Figures

Chapter 1.0: Introduction

1.1 Introduction

In recent years, worldwide concern has been voiced regarding the observed decline of Atlantic salmon (*Salmo salar*) throughout its habitat range. Among the most important factors responsible for this decline is the rapid decrease in suitable freshwater habitat (Gibson, 1993). Freshwater habitats fulfill important ecological functions pertinent to spawning and overwintering, usually involving a main river with tributaries and ponds. While young salmon are most sensitive to environmental disturbances during their first year after emergence from eggs, it is the availability and quality of spawning habitat that has the greatest impact on their production rate (Shearer, 1992). Any physical disturbance in the watershed such as increased erosion or artificial obstructions that limit access to upstream areas, may have a severe impact on the quality and availability of salmon habitat.

With increasing pressure from natural resource based industrial activities (e.g. forestry and mining), and the expansion of urban development into formerly pristine areas, effective management strategies have to be developed to ensure the protection of the endangered habitat. These strategies should be based on quantitative inventories and support repeated applications. This would permit the integration of relevant information in digital data bases and facilitate efficient resource management through research into the relationships between Atlantic salmon and its habitat parameters (Edgington *et al.*, 1987; Scruton *et al.*, 1992).

Traditionally, resource inventories of freshwater salmon habitat have been carried out with a substantial amount of ground based data collection. While offering the highest degree of accuracy, the cost and time effort required to conduct field surveys increases with the area covered and may not be viable for large and remote areas. Alternatively, air-photo interpretation has been successfully used fc cost and time efficient salmon habitat inventories and mapping (Dubois and Clavet. 1979; Amiro, 1983; Edgington *et al.*, 1987). In this case, the quality of the results generally depends on the experience and knowledge of the interpreter. Therefore, potentially severe limitations of this method exist for the objective reproduction of interpretation results.

Information about land cover and water bodies pertinent to habitat inventories can be derived in a quantitative, objective manner from remotely sensed, multispectral imagery (Lyzenga, 1978; Richards, 1987; Rimmer *et al.*, 1987; Dekker *et al.*, 1992; Bierwirth *et al.*, 1993). These images are digital representations of reflected or emitted radiation of the earth's surface. Typically, radiance is recorded at the sensor for several regions of the electromagnetic spectrum, or spectral bands. Presently, remotely sensed data are available with spatial resolution ranging from a few meters for airborne sensors to 20 or 30 m for satellite platforms. The number of optical spectral bands and spectral resolution available varies across different sensors. Generally, spectrometric airborne imagery offers a greater choice of spectral bands and higher spectral and spatial resolution.

Recently, researchers have discovered the benefits of using airborne remote sensing data in research questions related to the management of freshwater resources, such as the extraction of bottom substrate as an indicator of spawning habitat suitability (e.g. MacLeod *et al.*, 1992; Acornley *et al.*, 1995). Given previous research, it is the purpose of this study to contribute to a better understanding of the potential of multispectral remote sensing for the inventory and management of freshwater salmon habitats.

1.2 Research Objectives

The principal objective of this study is to explore the potential of multispectral remote sensing and digital ancillary data as a tool for the inventory and mapping of freshwater habitat parameters of Atlantic salmon (*Salmo salar*). Secondary objectives are as follows:

I. Identify important freshwater habitat parameters
II. Define an appropriate set of predictor variables for each parameter
III. Select an appropriate classification method and assess classification accuracy
IV. Combine individual parameters to model spawning habitat suitability

In order to achieve the principal objective it is necessary to review past and current research on relationships between Atlantic salmon and freshwater habitat components. The results of this review are used in the identification of habitat parameters to be predicted. While spectral information and its derivatives form the central data component, an effort is made to identify and incorporate relevant non-spectral information into the analysis. A set of potential predictor variables is selected for each habitat parameter, and an appropriate classification method identified. An assessment of classification accuracy is carried out for each habitat parameter. Individual habitat parameters are combined to yield composite maps of habitat suitability.

1.3 Thesis Organization

Basic concepts of the freshwater ecology of Atlantic salmon are introduced in Chapter 2.0. Important freshwater habitat parameters are identified, and the use of remote sensing data pertaining to the extraction of these parameters is examined. Special consideration is given to image classification techniques and the extraction of features submerged in water. Chapter 3.0 introduces the study area and contains an account of data collection procedures. Methods and procedures followed in this study are presented in Chapter 4.0. Chapter 5.0 describes processing results and statistical characteristics of all variables used in this analysis. The selection of predictor variables is explained and the classification of each habitat parameter is presented. Furthermore, this chapter contains an evaluation of the performance of the developed methodology. Finally, an example is given of how the results of this investigation can be used to model salmon spawning habitat suitability. A critical discussion of the results of this investigation is presented in Chapter 6.0. Error sources are identified for each habitat parameter and their impact on the classification accuracy is assessed. Chapter 7.0 contains a summary of results and recommendations for future research.

Chapter 2.0 Background

2.1 Introduction

As an anadromous species, Atlantic salmon spend part of their life cycle in freshwater and part in the ocean. In general, young salmon migrate into the ocean after a freshwater stage of one to five years. In most cases, they return to the location of their emergence after one winter at sea to spawn, although some fish spend two or more winters at sea. The majority of returned salmon do not survive spawning, but a certain proportion will migrate back into the ocean (Shearer, 1992).

Freshwater habitats of Atlantic salmon are determined by water velocity and depth, substrate size and the amount of cover, either in-stream (e.g. boulders and logs) or along the banks (vegetation) (Heller and Hohler, 1981; Hawkins *et al.*, 1993). The distribution of salmon of various age groups throughout these habitats varies considerably. Generally, fry (age < 1 year) occupy the spawning areas whereas parr (age > 1 year) migrate toward deeper, wider stream sections (Gibson, 1993). In the case of insular Newfoundland, habitat utilization is especially flexible due to the absence of many competing species. As a result, inter-species competition is comparatively low, and Atlantic salmon occupies ecological niches that are usually characterized by different species in mainland Atlantic Canada. Rivers and streams form the rearing habitat for Atlantic salmon during the freshwater stage. However, with respect to reproduction and, ultimately, survival of the population, their most important function is in providing suitable spawning grounds (Shearer, 1992; Gibson, 1993).

2.2 Variables Characterizing Freshwater Habitat of Atlantic Salmon

Barila *et al.* (1981) examined the relationship between biological parameters (i.e. number of species, number of individuals, diversity index) and selected geomorphological variables such as stream order, width, gradient and water depth. It was found that species diversity was highly correlated with stream order. Lanka *et al.* (1987) investigated the relationship between drainage basin geomorphology and trout standing stock (kg/100 m²) in 91 Rocky Mountain streams. The corresponding watersheds were either covered by high elevation forest (65 streams) or by low elevation rangelands (26 streams). Multiple regression analysis was used to predict trout standing stock from various geomorphological variables. In the case of forest streams, the predictive model consisted of the variables reach elevation, relief ratio, drainage density and average stream width (r^2 = 0.51). Trout standing stock in rangeland streams was predicted using basin elevation, basin perimeter, channel slope and basin relief (r^2 = 0.64). Frissell *et al.* (1986) present a hierarchical framework to characterize stream habitats according to different spatial resolutions and temporal periods. Table 2-1 gives an overview of the proposed levels and associated geomorphological processes. The most degrading impacts on riverine habitats occur from the level "Reach" on downwards to "Microhabitat". These impacts are either of natural cause or related to man-made activities. An efficient method of habitat surveillance and monitoring will focus on habitat parameters at these levels.

Table 2-1: Characteristics of Stream Habitats (after Frissell *et al.*, 1986)

Level	Spatial Resolution [m]	Time Period [years]	Boundaries	Processes
Stream	> 1000	100,000 to 1,000,000	Drainage Basin	Denudation
Segment	100 to 1000	1000 to 10,000	Junctions; Falls	Migration of Tributary Junctions
Reach	10 to 100	10 to 100	Slope Breaks	Bank Erosion
Pool-Riffle	1 to 10	1 to 10	Bed Profile	Bedform Changes
Microhabitat	0.1 to 1	0.1 to 1	Substrate Type; Depth	Microbial Activity

Benda *et al.* (1992) investigated the distribution of salmonid habitats over a watershed at three spatial resolutions, i.e. 0.1 to 1 km^2, 2 to 26 km^2 and 240 km^2. These numbers correspond to river sections, sub-basins, and the whole watershed, respectively. The most extensive areas of rearing and spawning habitats were found to be located along small reaches on a young fluvial terrace. On the sub-basin level, varying habitat quality was related to discharge and channel gradients. Overall, most habitats were located along extensive stretches of the main river valley

Scruton and Gibson (1993) defined habitat suitability indices for juvenile Atlantic salmon at 18 selected rivers in Newfoundland, Canada. The number of fish/100 m^2 was related to the following river characteristics: stream width, water depth, discharge, substrate type, in-stream cover and cover by overhanging vegetation. It was concluded that fry show preference for shallow (10 to 20 cm) and narrow (<3 m) stream sections with pebble and cobble substrate, whereas parr favoured wider (<7.5 m) and deeper (15 to 40 cm) sections with boulder-rich bottom material.

Accordingly, Gibson (1993) reports the vital importance of substrate grain size and heterogeneity for spawning habitat quality. In particular, habitats with coarse sediments (cobble and pebble substrate) are described as being favoured over locations with finer sediments. Sediment size has a profound impact on egg survival: the presence of large quantities of suspended fine sediment can hamper egg oxygen supply and fertilization and inhibit emergence. Preferred water depths for spawning lie within a range of 10 to 75 cm with discharge velocities ranging from 15 to 90 cm/s for normal flow conditions. Appropriate spawning areas are often located along upper reaches in the head water region of a river system, and full utilization is only achieved by free access to these areas without anthropogenic or natural obstacles.

Other significant features characterizing riverine salmon habitat are channel patterns such as pools, riffles, cascades, falls and rapids. Channel patterns reflect characteristic combinations of the type of current flow and water depth. Moreover, they often occur in relatively regular sequences, such as alternate successions of pools and riffles, and may constitute very distinct habitats. Land cover and land use pattern in the watershed are important sources of information about the type of riparian vegetation and the location and nature of obstructions in the river. Land cover information is necessary for the identification and quantification of potential sources of fine sediment and other pollutants, such as excessive urban development, road construction and quarrying (Heller and Hohler, 1981; Bisson *et al.*, 1981; Frissell *et al.*, 1986; Gibson, 1993; Hawkins *et al.*, 1993).

The geomorphologic-hydrological and anthropogenic characteristics of a river system are the predominant factors controlling the quality of freshwater salmon habitat. These characteristics can be described by the type of bottom substrate, channel pattern and land cover in the watershed.

2.3 Remote Sensing Approaches to Riverine Habitat Mapping

2.3.1 Air Photo Interpretation

Dubois and Clavet (1979) have recognized air photo interpretation as a valuable tool for the development of habitat inventories for salmon rivers in Quebec. Aerial photography at a scale of 1: 50,000 was used in the interpretation of channel pattern, land use along the rivers, and terrestrial and aquatic vegetation. This led to the successful identification of pools and spawning beds as important habitat types. The overall classification accuracy ranged from 62% for pools to 81% for spawning beds. Clavet (1980) proposed a salmon habitat inventory method to be carried out in six stages, including the identification of relevant parameters, their coding into a map legend, a field survey, preliminary and refined air photo interpretations, and a final cartographic representation of potential salmon habitat. Approximately 80% of the spawning beds and rearing habitats characterized by channel pattern and substrate type could be located correctly, provided that the photographs were acquired at water levels similar to those expected during spawning season. Therefore, photographs acquired at extremely high water levels are not appropriate since it is not possible to discern from them the actual state of submergence of gravel beds during spawning season. The presented method of sub-surface feature detection works best at water depths of less than 2 m. Difficulties encountered in photo interpretation include specular reflection of sun light on the water surface and the occurrence of different substrate types showing similar colour and brightness.

Interpretation of aerial photography was used by Côté *et al.* (1987) to identify substrate types at locations of shallow water. In the presence of deep water or specular reflection, the type of bottom substrate was inferred from riverbank topography, erosive processes and depositional microforms in the streambed. Rubin (1992) found color and color-infrared multitemporal aerial photography at a scale of 1:24,000 particularly useful for the identification of historic channels, the type of riparian vegetation and land use activities in the watershed. This information was subsequently used for the restoration of productive habitat for anadromous fish species in California.

The collection of habitat information by means of air photo interpretation is more cost and time efficient than conventional field based data acquisition due to the availability of aerial photographic data, short processing times and minimal costs (Dubois and Gosselin, 1994).

2.3.2 Multispectral Remote Sensing

Salmon habitat parameters such as substrate type and channel pattern are submerged or part of the water body. Water applications of satellite or airborne remote sensing have primarily focused on the identification of suspended sediment concentration, chlorophyll content and concentration of dissolved organic matter (Rimmer *et al.*, 1987; Lathorp and Lillesand, 1989; Dekker *et al.*, 1992; Goodin *et al.*, 1993; Hamilton *et al.*, 1993; Nichol, 1993; Jupp *et al.*, 1994). The applicability of multispectral imagery for the detection of bottom features has been demonstrated in conjunction with bathymetric mapping (Lyzenga, 1978, 1983; Lathorp and Lillesand, 1989; Philpot, 1989; Roberts *et al.*, 1992; Luczkovich *et al.*, 1993; Lyon and Hutchinson, 1995).

Attenuation of electromagnetic radiation through scattering and absorption in the water column has to be considered where the target features are either covered by water or water depth itself is the object of interest. Scattering is strongest at short wavelengths, whereas absorption affects radiation of longer wavelength. In addition, the absorption behaviour of water constituents such as dissolved organic matter or suspended sediment can substantially influence the water leaving radiance (Dekker *et al.*, 1992; Jupp *et al.*, 1994). Attenuation in the water column increases exponentially with water depth. Lyzenga (1978; 1981) accounts for this relationship by formulating a bottom type index. The following transformation is applied to the image data to obtain a variable that is linearly dependent on water depth:

$$X_i = \ln(L_i - L_{Si}) \qquad (2\text{-}1)$$

where

X_i = transformed radiance in band i
L_i = radiance in band i
L_{Si} = deep water radiance in band i

Training samples were collected over areas of uniform bottom reflectance and subsequently used in the calculation of a coordinate system rotation, yielding n-1 depth independent variables and one depth dependent variable from the transformed radiances. This method was applied by Lambert (1994) to the mapping of submerged kelp beds in Eastern Canada.

A similar approach was used by Bierwirth *et al.* (1993) for the mapping sea floor reflectance in shallow coastal waters. LANDSAT-TM imagery was converted to reflectance values to represent the spectral properties of the substrate. Water depth was calculated from deep-water reflectance and depth invariant bottom type reflectance. An estimate of the true depth could be obtained by assigning a value of zero for bottom type reflectance, thus assuming the mean substrate reflectance over all bands to be one. The slope of a regression line established from known bathymetry data and the natural logarithm of pixel reflectance yielded water attenuation coefficients for each spectral band.

Khan *et al.* (1992) applied principal component analysis (PCA) directly to LANDSAT-TM bands 1 and 2 without prior corrections. Single band thresholding of the second principal component was used to differentiate sand, rock, mud and seagrass cover in the Western Arabian Gulf.

Zacharias *et al.* (1992) could distinguish several types of intertidal seaweeds. The Compact Airborne Spectrographic Imager (CASI) was used to collect spectral data in 8 channels as listed in Table 2-2. Bands 1, 2, 3, 6 and 8 were subjected to PCA. Image classification was subsequently carried out applying the ISODATA[1] algorithm to the second, third and fourth principal components. At three different locations, at least two genera of seaweed could successfully be discriminated with overall accuracies ranging from 65 to 86 %.

Table 2-2: Spectral Band Configuration Used by Zacharias *et al.* (1992)

Band 1 [nm]	Band 2 [nm]	Band 3 [nm]	Band 4 [nm]	Band 5 [nm]	Band 6 [nm]	Band 7 [nm]	Band 8 [nm]
431 to 459.	480 to 590	545 to 559	602 to 614	646 to 660	666 to 678	746 to 750	871 to 879

MacLeod *et al.* (1992) used CASI imagery to map substrate types constituting aquatic habitat in Lake Ontario. A total of seven spectral bands were used in the analysis (Table 2-3). Image

[1] The ISODATA algorithm is described in detail in Section 2.4.

classification using the ISODATA algorithm was carried out with bands 1, 4 and 7 as well as with the band ratios 7/6, 4/1 and 6/3. The types of bottom substrate encountered in the study area consisted of vegetation, mud, limestone rubble with gravel and sand, and limestone rubble with boulders. Among these, the bottom types vegetation, mud, and limestone rubble could be differentiated. Similar results were obtained with both raw data and data converted to radiance units. This indicates that conversion to absolute radiance values is not essential when mapping bottom substrate.

Table 2-3: Spectral Band Configuration Used by MacLeod *et al.* (1992)

Band 1 [nm]	Band 2 [nm]	Band 3 [nm]	Band 4 [nm]	Band 5 [nm]	Band 6 [nm]	Band 7 [nm]
470 to 500	515 to 536	540 to 561	575 to 597	625 to 647	670 to 692	740 to 760

Unsupervised cluster analysis was applied by Lyon *et al.* (1992) to map bottom sediment types with Daedalus 1260 data at St. Mary's River, Michigan. The four spectral bands listed in Table 2-4 were used to discriminate 50 initial clusters. These were subsequently grouped and identified as sand, silt/clay, silt/sand, sand/silt and sand-rock/silt.

Table 2-4: Spectral Band Configuration Used by Lyon *et al.* (1992)

Band 1 [nm]	Band 2 [nm]	Band 3 [nm]	Band 4 [nm]
400 to 450	500 to 550	550 to 600	600 to 650

Luczkovich *et al.* (1993) used Landsat TM spectral bands 1, 2 and 3 to map coral reefs, sand bottom and sea grass off the coast of the Dominican Republic. At water depths ranging from 0 to 5 m the variability in samples of the three bottom types was related to heterogeneity in the samples rather than to water depth. Borstad *et al.* (1992) demonstrated the potential of multispectral airborne remote sensing to detect fish schools. Using three spectral bands of a CASI sensor centered at 470 nm, 545 nm and 640 nm, several herring schools could be discriminated successfully against the background radiation of deep water and sea floor.

Acornley *et al.* (1995) demonstrated the potential of CASI imagery for the mapping of salmonid spawning habitat in the River Test, England. Reference data consisted of spectrometric and bathymetric measurements as well as positional measurements of redd locations. The selected CASI spectral band configuration is presented in Table 2-5. Lyzenga's (1978) method was applied to calculate a linear relationship between spectral response and water depth. Bathymetric measurements were correlated to all transformed spectral bands and showed the highest correlation with the ln- transformed Band 8 with $r = -0.82$. This relationship was used to derive a map of predicted water depths. Potential spawning habitats were mapped using the spectral bands 1 to 10 in a maximum likelihood classification. Qualitatively, the classification result was found to correspond well with the known location of spawning beds.

Table 2-5: Spectral Band Configuration Used by Acornley *et al.* (1995)

Spectral Band									
1	2	3	4	5	6	7	8	9	10
Central Wavelength [nm]									
510	555	590	620	645	660	670	701	740	800

Remote sensing has proven to be an efficient means to gather information about aquatic habitats. If applied to the mapping of rivers, remotely sensed data should have a sufficiently high spatial resolution for an appropriate coverage of narrow river sections. Moreover, requirements for the application of multispectral imagery include the selection of at least three spectral bands: two bands in the visible spectral region for the extraction of submerged features as well as one near-infrared band to separate water covered areas from dry land. The visible bands should be selected so as to minimize loss of information due to scattering and absorption in the water column. An overview of the spectral characteristics of sensors from selected applications in aquatic habitat mapping is presented in Table 2-6.

Table 2-6: Selected Remote Sensing Applications of Substrate Mapping

Author	Sensor	Spectral Bands [nm]	Method	Types of Substrate
Lyzenga, 1981	M-8	480 to 520 500 to 540 520 to 570 550 to 600 580 to 640 620 to 700	canonical transform; use of bathymery	hard, unvegetated bottom, white carbonate sand; seagrass beds
Borstad *et al.*, 1992	CASI	460 to 480 535 to 555 630 to 650	visual interpretation of color composites.	fish schools; deep water; bottom.
Khan *et al.*, 1992	Landsat TM	450 to 520 520 to 600	PCA applied to spectral bands with no prior transformation.	sand; beach rock, hard bottom; mud; seagrass.
Zacharias *et al.*, 1992	CASI	See Table 2-2	PCA applied to spectral bands with no prior transformation; ISODATA classification of principal components.	several types of intertidal seaweeds.
MacLeod *et al.*, 1992	CASI	See Table 2-3	ISODATA classification of raw bands and band ratios	vegetation, mud; limestone rubble.
Lyon *et al.*, 1992	Daedalus 1260	See Table 2-4	unsupervised classification of raw spectral bands.	sand, silt clay; silt sand, sand silt, sand-rock silt.
Bierwirth *et al.*, 1993	Landsat TM	450 to 520 520 to 600 630 to 690	bottom reflectance is obtained by applying radioative transfer model, use od bathymetric information	substrate reflectances
Lambert, 1994	Landsat TM and SPOT	TM2 TM3 XS1 XS2	canonical analysis; use of bathymetry	macrophytes, sand rock
Acomley *et al.*, 1995	CASI	see Table 2-5	maximum likelihood classification of raw spectral bands.	redd (nest) locations.

2.4 Automated Classification of Multispectral Imagery

Information is extracted from remote sensing data by means of image classification. This assumes that picture elements (pixels) showing similar spectral behaviour can be grouped into distinct spectral classes that correspond to features of interest, or informational classes. The relationships between informational and spectral classes are established by a chosen classification method. These methods are commonly divided into supervised and unsupervised techniques. The latter use only information that is inherent in an image without prior knowledge about the location of target features, and the classification is carried out by cluster analysis or histogram merging techniques. Supervised classification algorithms, on the other hand, use training data to derive spectral signatures for classes of interest. Unknown pixels in the image are then assigned to a class according to these signatures. If reliable training data are available, supervised classification methods generally outperform unsupervised approaches (Mather, 1987).

Relatively simple methods of supervised image classification assign pixels to classes based on the Euclidean distance from the class mean (minimum distance classifier) or the value range in that class (parallelepiped classifier). The ISODATA algorithm is a modification of the minimum distance method that combines the characteristics of both, supervised and unsupervised approaches (Duda and Hart, 1973). A set of training clusters is used to compute class centroids. New cases are subsequently classified according to their distance from these centroids. With each classification, the group centroids are recomputed, and the procedure is repeated until no further changes occur. In the case of overlapping spectral signatures, however, the methods described above can lead to large classification errors.

In the maximum likelihood procedure, statistical frequency distributions are used to classify pixels according to their likelihood of class membership. This method requires the estimation of probability density functions for all spectral classes from mean vectors and covariance matrices. In order to correctly estimate class membership probabilities, training pixels must be normally distributed with respect to all spectral bands used in the analysis. Furthermore, it is assumed that the covariance matrices in each class are equal, although techniques exist to account for inequality of covariance matrices as demonstrated by Kershaw and Fuller (1992). Pixels are assigned to classes according to the highest probability. In general, whenever spectral classes are not distinctly clustered around a mean or well separated by their value ranges, statistical image classification is the superior method (Gonzalez and Woods, 1992).

The maximum likelihood approach to image classification is an example of statistical discriminant function analysis (Lachenbruch, 1975). Serious implications can arise if assumptions regarding the distribution of the data and the equality of covariance matrices are violated (Basu and Odell, 1974; Mather, 1990). In addition, the presence of spatial autocorrelation, as is to be expected with most remotely sensed data, can result in underestimated variances (Campbell, 1981). Since group membership probabilities are estimated from these variances, the classification may be unreliable and no inferences may be made about the discriminating power of the underlying model (Cliff and Ord, 1981; Labovitz and Masuoka, 1984; Griffith, 1987; Odland, 1988; Chou *et al.*, 1990). Spatial autocorrelation effects in the analysis of remotely sensed data can be accounted for by choosing a random sampling scheme for training and test data to ensure the independence of sample pixels (Campbell, 1981).

Alternatively, decision tree analysis (DTA) has been applied to data that did not match the requirements for conventional statistical classification techniques (Kass, 1980; Hawkins and Kass, 1982; Breiman *et al.*, 1984; Quinlan *et al.*, 1987; Lees and Ritmann, 1991). Relationships between spectral and informational classes are detected by dividing a data set recursively into smaller portions according to a set of predictor variables and one response (dependent) variable. The final result is a division of the original data set into mutually exclusive and exhaustive sub-sets. No limiting assumptions about data distributions are necessary, and in the same data set categorical as well as continuous data can be handled simultaneously (Fabricius and Coetzee, 1992; Dymond and Luckman, 1994). With respect to the analysis of remotely sensed images, the incorporation of non-image data such as polygons digitized from maps or digital elevation models is facilitated (Walker and Moore, 1988; Lees and Ritmann, 1991). The property of generating rules from large, heterogeneous databases has recently led to the incorporation of decision tree algorithms in expert systems as tools for inductive knowledge generation. This is an efficient alternative to the expensive and slow knowledge acquisition by interviewing human experts (Hart, 1986; Wharton, 1987; Moller-Jensen, 1990; Kelly, 1991; Günther *et al.* 1993).

In DTA complex data sets are handled in a flexible manner through recursive partitioning. That is, a data set is progressively divided into smaller, more homogeneous sub-sets, and relationships between predictor and response variables are analyzed for each sub-set separately. Several approaches exist to partition the data set. Morgan and Sonquist (1963) developed an automatic interaction detection algorithm (AID). A data set consisting of a continuous dependent variable and categorical predictors is partitioned by collapsing categories of predictor variables. Categories are combined to maximize the between-group-sum-of-squares. The resulting split is always binary. Predictor variables are either monotonic (ordinal) or free (nominal). Categories of free predictors can be combined in any order, whereas categories of monotonic predictors can only be combined in an ordinal fashion. Significant improvements to the AID algorithm led to the development of the CHAID method (Kass, 1980). In this case the dependent variable is categorical, and the best predictor variable to define a split at a given node is selected according to statistical significance. That is, predictor variable categories are collapsed so as to maximize the χ^2-statistic, and the statistical significance of the resulting groupings of categories is calculated. Other improvements include the inclusion of a type of predictor variable that can handle missing data, and the possibility of k partitions, where $2 \leq k \leq c$ (c = number of categories in the predictor variable). At every node, a multiple search is conducted to find the most significant statistical relationship between predictor and response variable, resulting in a higher probability of detecting relationships by chance which, in reality, do not exist (increased Type I error rate of false acception). In order to counter problems associated with the detection of spurious relationships, statistical significance values for potential splits are divided by the Bonferroni factor. This factor is calculated based on the number of ways the original predictor variable categories can be combined to groups with the number of groups fixed to the final number of merged categories. Consequently, predictors with many categories are discriminated against in favour of variables with fewer categories. Biggs *et al.* (1991) developed exhaustive partitioning as a refinement of the CHAID method. As before, the selection of the best split is based on statistical significance. The Bonferroni adjustment, however, is calculated allowing for a variable number of groups so as to remove bias towards variables with few categories.

Estimating the accuracy of a decision tree classification requires elimination of redundant branches (pruning) to find the optimal sized tree for a given application. The pruning process is necessary since redundancy of rules will decrease the accuracy of the decision tree through over-fitting. Pruning is generally realized either by cross-validation procedures using the training sample, or by using a test sample that was collected independently from the training data set Quinlan *et al.*, 1987; Safavian and Landgrebe, 1991). According to Breiman *et al.* (1984), using an independent test sample is the preferred and statistically more robust method. In this case, the decision tree is reproduced on the test data set. The branches of the tree are successively removed while observing the overall accuracy at every step. The process is stopped when the accuracy decreases with the removal of a branch.

Reddy and Bonham-Carter (1991) have demonstrated the relevance of decision trees for spatial analysis. Exhaustive partitioning was used to analyze geological, geophysical and remotely sensed data to predict mineral occurrence. The resulting hierarchical tree structure is shown in Figure 2-1. Decision rules consisting of "IF...THEN" statements were established by following down the branches of the tree to the end nodes. The rules were implemented using a geographic information system (GIS), and mineral occurrences could successfully be predicted.

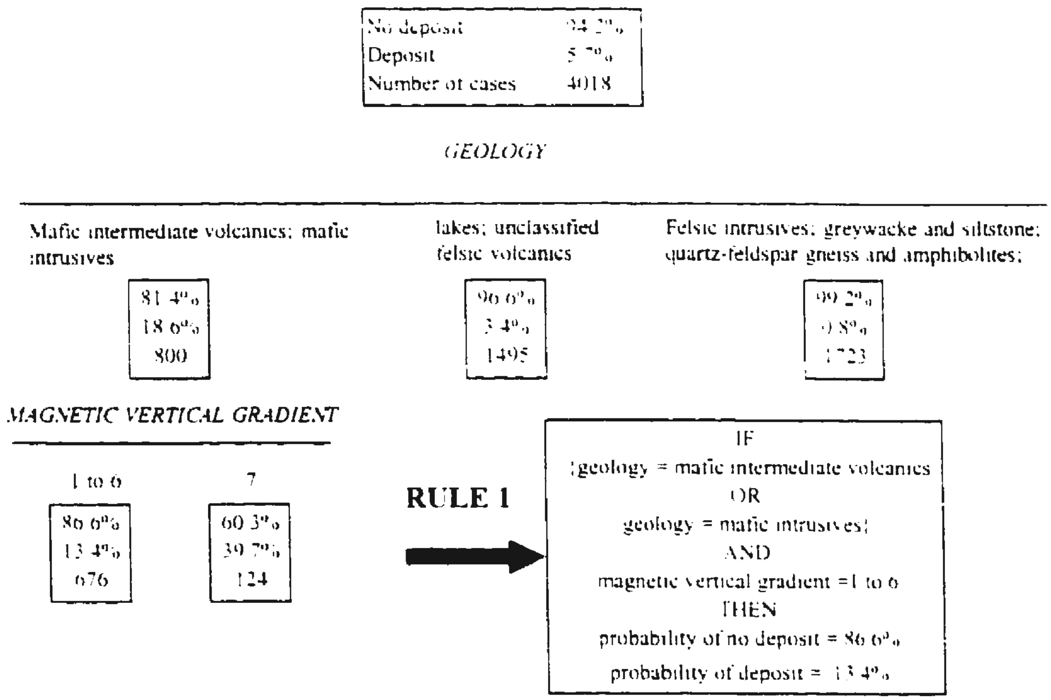

Figure 2-1: Decision Tree for Mineral Mapping (after Reddy and Bonham-Carter, 1991)

Lees and Ritmann (1991) used DTA to integrate remotely sensed and digital map data for vegetation mapping. Landsat TM spectral bands 1, 2, 3, 4, 5, and 7 were used together with relief and geological information in a binary decision tree algorithm to predict distributions of eight vegetation classes in a hilly environment. Only the categories "dry sclerophyll vegetation" and "cleared forest" were classified with acceptable accuracy levels of 70 and 88 %, respectively. For the remaining classes the proportion of correctly classified cases ranged from 19 to 49 %. Nevertheless, this result was found to be superior to using either imagery or thematic map data alone under similar conditions. Belward and de Hoyos (1987) applied a supervised binary decision tree to the classification of agricultural crops from LANDSAT-MSS imagery. Eight types of crop were distinguished successfully at per class accuracies ranging from 48 to 99 %. This result was found to be comparable to the result of a conventional maximum likelihood procedure. However, the decision tree was found to be computationally more efficient and required less time for training area generation. Recently, Hess *et al.* (1995) used decision tree classification to estimate inundated area and vegetation in the Amazon floodplain from multi-frequency, polarimetric synthetic aperture radar (SAR) imagery. The land cover categories of water, clearing, macrophytes, non-flooded forest and flooded forest were identified with per class accuracies above 90%.

Bottom substrate, channel pattern, and type of land cover have proved to be key factors in the characterization of freshwater habitat for Atlantic salmon. While conventional air photo interpretation is in some instances applied routinely to the mapping of freshwater salmon habitat, the full potential of digital imagery has yet to be explored. Studies concerned with bottom type mapping have mostly concentrated on either coastal or lacustrine environments. The results from these investigations, particularly in conjunction with the recent findings of Acornley *et al.* (1995), strongly suggest that an extension into the mapping of habitat in rivers and streams is feasible. Further advances could be made by including non-spectral information about habitat parameters in the analysis. For example, bottom substrate and channel pattern can also be described by geomorphologic-hydrological measures such as stream width and gradient. Decision tree analysis has proven to be a reliable method for the efficient and robust statistical analysis of large data sets with varying data types. This characteristic makes DTA ideally suited for the integration of both spectral and non-spectral data in an analysis of freshwater salmon habitat.

Chapter 3.0 Study Area and Data

3.1 Study Area

The Come By Chance River study area is characterized by abundant freshwater salmon habitat, ample road access. A limited occurrence of shaded river sections provides the basis for the use of remotely sensed data. The river is located on the isthmus of the Avalon Peninsula in eastern Newfoundland (Figure 3-1) and has an axial length of about 17 km, draining a watershed of approximately 64 km^2. Stream width varies from 33 m at the mouth to 4.5 m close to the headwaters (Fisheries and Oceans Canada, 1994) (Figure 3-2). Ponds and tributaries were excluded from the analysis. The overall flow conditions have been described as relatively stable, with little change in bedforms and substrate during high water conditions (Harmon, 1966). In the western part of the watershed, the topography is characterised by comparatively steep hills rising from 170 m to 300 m above sea level. In the eastern part, the slopes are gentler with elevations varying from 100 m to 170 m. The bedrock material is composed of Palaeozoic volcanic rocks (Agriculture Canada, 1991). Glacial and glacio-fluvial sediments dominate the surficial geology, with occasional bedrock outcrops on steep slopes. Stony, humic-ferric podzols are the principal soil types. The vegetation cover in the watershed includes dwarf shrub heath dominated by *Kalmia angustifolia*, bogs with *Sphagnum sp.* mosses, fens composed of grasses and sedges and forest with *Abies balsamea* and *Picea mariana* (Damman 1983). Coniferous forest is predominantly found in the sheltered valley and along the lower slopes. Most of the higher elevations and hilltops are covered with dwarf shrub heaths.

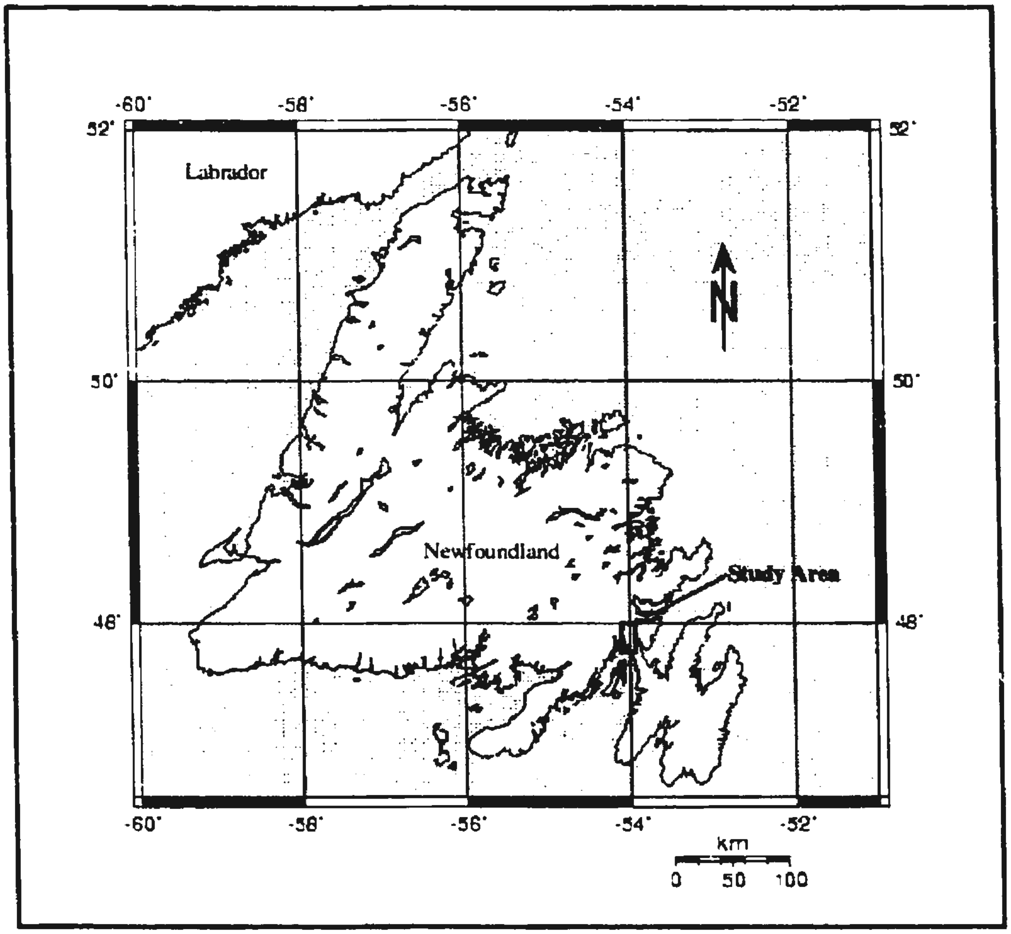

Figure 3-1: Study Area Location

At locations exposed to strong winds, *Empetrum eamesii* replaces *Kalmia angustifolia* as the dominant species. The riparian vegetation is composed of forest, shrubs, grass, sedges and occasionally more extensive swamps with stands of alder. Urban development is concentrated in the communities of Come By Chance and Goobies (Figure 3-2). Industrial activity is limited to an oil refinery located about 5 km south of Come By Chance. The main transportation routes are the Trans-Canada Highway, the Burin Peninsula highway and the former Canadian National railbed, now used as a gravel road. In addition, numerous trails are present throughout the watershed.

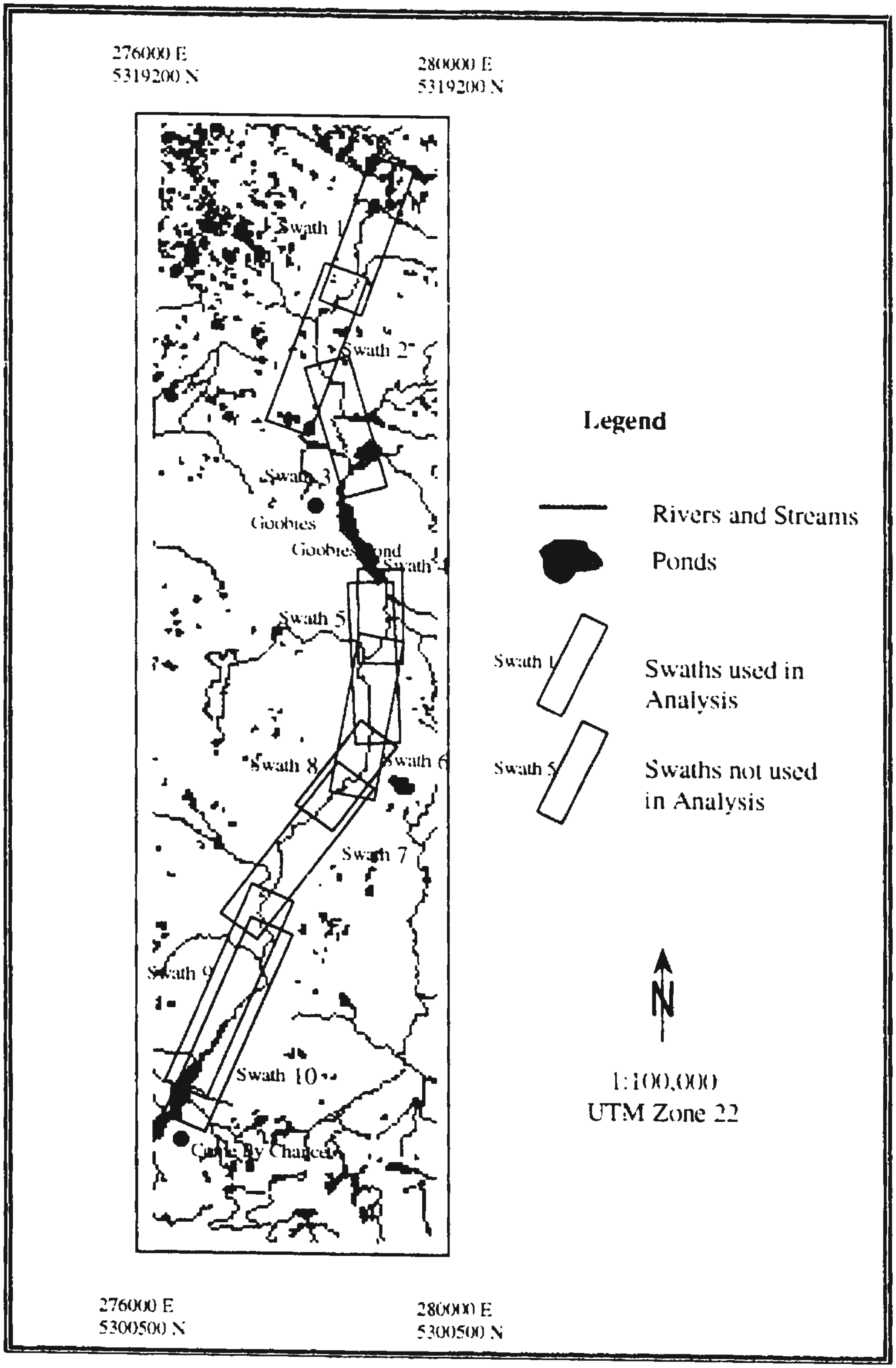

Figure 3-2: Come By Chance River Study Area

3.2 Data

In Chapter 2.0, important factors were identified that determine the quality of freshwater habitat for Atlantic salmon. Accordingly, this study is concerned with the extraction of substrate type, channel pattern and land cover as major habitat parameters from remotely sensed data. In this section all the data sets used in the analysis will be described in detail, including field data, aerial photography, multispectral imagery and ancillary map data.

3.2.1 Field Data

A stream habitat survey conducted in October 1993 served as the primary source of reference data for the classification of channel pattern and substrate type. For this purpose the river course was divided into segments that served as basic sampling units and varied in length from 50 m to 200 m. Each segment is bordered by continuously numbered transects across the river course. The segments were identified by their upstream transect and marked in the field (Figure 3-3). Variables measured over segments and along transects are presented in Table 3-1 (Fisheries and Oceans Canada, 1994).

Table 3-1: Field Data

Parameter	Measurement Unit	Mode of Measurement
Bank Height; Bank Gradient	meter; degree	at transect
Channel Pattern	category	per segment
Grain Size	percent	per segment
Habitat Unit	square meter	per segment
Segment Length	meter	per segment
Stream Width	meter	at transect
Water Depth	centimeter	3 to 6 measurements across transect

All river transects were numbered according to their distance from Transect 0 in the estuary. Water depth was determined at each transect at three to six equally spaced locations across the width of the river. Each segment was assigned one type of channel pattern, such as run, riffle, steady, flat and rapid (Table 3-2). Substrate was recorded for each segment as proportions of seven basic grain size classes, including fines, gravel, pebble, cobble, rubble, boulder and bedrock (Table 3-3). Habitat units were calculated as the area of each segment.

Table 3-2: Definition of Channel Pattern (after Scruton *et al.*, 1992)

Channel Pattern	Definition
Run	Swift, turbulent flow with broken surface; mean depth > 25 cm; stream width is less than average; boulder and rubble substrate;
Riffle	Average to rapid flow with broken surface; mean depth < 25 cm; gravel through boulder substrate;
Steady	slow flow with smooth surface; stream width and depth greater than average; extend over several segments
Flat	Slow flow with smooth surface; stream width and depth greater than average; occurs within a segment;
Rapid	Areas of steep gradient; rapid to turbulent flow; white water; rubble, boulder and bedrock substrate;

Table 3-3: Definition of Grain Size Classes

Grain Size Class	Diameter
Fines	< 0.2 cm
Gravel	0.2 to 3 cm
Pebble	3 to 5 cm
Cobble	6 to 13 cm
Rubble	14 to 25 cm
Boulder	> 25 cm
Bedrock	Bedrock

3.2.2 Aerial Photography

In order for the survey data to be used in the delineation of training and test areas, it was necessary to identify the surveyed river segments on the imagery. Since the field data were not geo-referenced *in situ*, a second survey was undertaken in September 1994 with the objective to identify river segments as marked in the field on panchromatic aerial photographs. The characteristics of the photographic data are given in Table 3-4.

Table 3-4: Characteristics of Aerial Photography

Altitude	Focal Length	Approximate Scale	Date
1341 m	152.7 mm	1:8800	May 1992
Roll: 92208		**Frames:** 142, 144, 146, 152, 154, 156, 158, 166, 168, 170	

According to the scope of this study, aerial photography served a dual purpose: (1) the principal source of reference with regards to the location of river segments, and (2) in the identification of training and test sites for the land cover classification. The scale of the aerial photographs was large enough to identify features such as individual trees, shrubs and boulders. Digital imagery and aerial photographic data were recorded within 16 months of each other. No major changes in land cover have occurred during this time period.

Land cover categories are defined in Table 3-5. An overview of all habitat parameter categories is presented in Table 3-6. The last column in Table 3-6 indicates the tables in which the respective variable categories are explained in detail.

Table 3-5: Land Cover Characteristics (after Scruton *et al.*, 1992)

Land Cover Category	Description
Coniferous	Mature coniferous trees
Alder	Large, deciduous shrubs up to 2 m height
Shrub	Softwood shrubs
Wetland	Bog, fen, and herbaceous vegetation
Noveg	Bare soil pavement and buildings
Water	Water

Table 3-6: Habitat Parameter Categories

Habitat Parameter	Category	Compare Table
Substrate Type	*Gravel* *Rubble* *Boulder* *Bedrock*	Table 3-3
Channel pattern	*Run* *Riffle* *Steady* *Flat* *Rapid*	Table 3-2
Land cover	*Coniferous* *Alder* *Shrub* *Wetland* *Noveg* *Water*	Table 3-5

3.2.3 Multispectral Image Data

Image data were collected with a Compact Airborne Spectrometric Imager (CASI) on October 23rd, 1993. The sensor allows for a flexible setting of spectral bands from 428 to 946 nm and operates in either spectral or spatial mode. In spectral mode, the sensor records reflected radiation in up to 288 channels at a low spatial resolution, whereas spatial mode data are collected at spatial resolutions of few centimeters or meters. The peak Spectral Radiance Unit (SRU) expressed in [$\mu W \cdot cm^{2} \cdot sr^{-1} \cdot nm^{-1}$] is selected during the flight in order to determine the optimal saturation point for scaling radiance values into a 12 bit range (Borstad, 1992). In this study, the CASI sensor was operated in spatial mode. A nominal spatial resolution of 1.5 x 1.5 m was chosen to allow for a proper coverage of the widest and narrowest river sections. At the given spatial resolution, reflected radiation could be registered in a total of four spectral channels. Details of the spectral band configuration are listed in Table 3-7. The images were recorded as 12 bit data and subsequently re-scaled into an 8 bit range as required for input in the image processing software.

Table 3-7: Spectral Characteristics of the CASI Sensor

	Band 1	Band 2	Band 3	Band 4
Bandwidth [nm]	499.5 to 521.1	579.0 to 600.7	648.3 to 671.9	718.1 to 741.8
Band Center [nm]	510.30	589.85	660.10	729.95

A peak SRU of 2.5 was chosen for all spectral bands to highlight features submerged in water while preserving sufficient spectral variability over land areas. Band 1 was positioned at 510 nm to maximize water penetration. No channel was selected in the blue spectral region due to strong scattering in atmosphere and water column. Bands 2 and 3 were selected to collect information at locations of shallow water while minimizing scattering in the water column. Band 4 was used to separate land and water. All four channels were used in deriving land cover features.

Initially, ten flight swaths were defined to account for the sinuosity of the river (Figure 3-2). Swath 3 had to be eliminated due to extreme geometric distortions. As a consequence, a river section of approximately 300 m length north of Goobies Pond was excluded from the analysis. In the case of overlapping swaths the image data with the least distortions were used. This reduced the initial image data set to 6 individual scenes. Swath 8 was used only in the assessment of radiometric normalization of the imagery since the area covered by this flight line is fully accounted for by Swaths 6 and 7 (Figure 3-2).

An example of multispectral imagery, aerial photography and recorded field data is given in Figure 3-3. In Figure 3-3(a), several types of land cover are distinguished on the false colour composite image. Examples of field survey data are overlaid onto the aerial photograph in Figure 3-3(b). Of particular interest is the fact that the river segment identified by Transect 6175 (channel pattern = steady; substrate = bedrock) shows homogeneous tone and texture. On the other hand, tone and texture clearly vary in the segment corresponding to Transect 6300 (channel pattern = riffle; substrate = bedrock). This apparent heterogeneity in field data sampling units has implications for the classification of substrate type and channel pattern which are discussed in Chapter 6.0.

3.2.4 Ancillary Data

The National Topographic Service (NTS) digital map sheet 1N/13 served as an ancillary data source. Elevation data contained as contour lines in this map sheet were used in the calculation of valley gradient. The derivation of valley gradient as a potential predictor variable for bottom substrate and type of channel pattern is described in detail in Chapter 4.0.

Figure 3-3(a): False Colour Composite (FCC) Image
Band 3 (Blue) - Band 2 (Green) - Band 4 (Red)
white/blue – no vegetation
purple – wetland/shrub
red – trees
black - water

Figure 3-3(b): Aerial Photograph
Transect 6175:
Channel Pattern = Steady
Substrate = Bedrock
Transect 6300:
Channel Pattern = Riffle
Substrate = Bedrock

0 100 m

Figure 3-3: False Colour Composite Image and Aerial Photograph of Sub-Area

Chapter 4.0: Methodology

4.1 Introduction

In this chapter, the procedures for data pre-processing, training data collection and selection of predictor variables are presented. Classification of habitat parameters based on decision tree analysis and methods for the assessment of classification accuracy are explained. An example of habitat suitability mapping using individual habitat parameters is given at the end of this chapter. The methodology followed in this study is summarized in Figure 4-1.

4.2 Pre-Processing

The data sets used in this study were subjected to various forms of pre-processing prior to their use in subsequent analyses. Pre-processing of the field survey data involved the categorization of continuous substrate data by means of cluster analysis. An ancillary data set was created by calculating stream width and valley gradient. All images were subjected to atmospheric, radiometric and geometric corrections as well as to the derivation of image transforms.

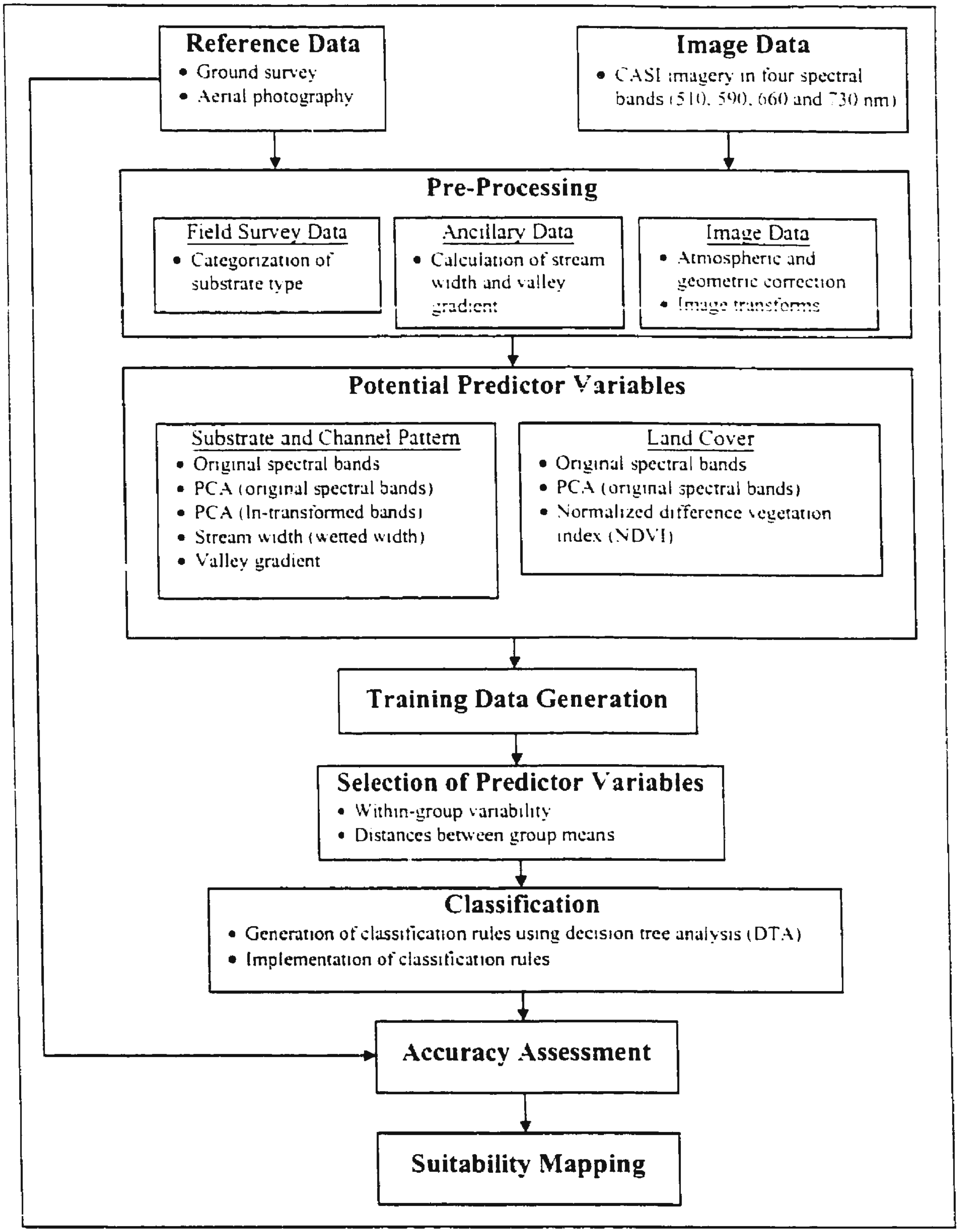

Figure 4-1: Methodology

4.2.1 Field Survey Data

Data from the field survey were used in the classification of the habitat parameters *Channel Pattern* and *Substrate Type*. The type of channel pattern was recorded as categorical data and subsequent pre-processing was not required. Bottom substrate composition, however, was initially described as proportions of seven basic grain size classes. In order to use bottom substrate as dependent variable, the continuous grain size data was reduced to only one variable consisting of discrete substrate categories using cluster analysis.

Cluster analysis is a method of grouping observations together that are similar with respect to a set of discriminating variables (Mather 1976; Davis, 1986). Observations are grouped together based on measures of similarity (correlation measures) or dissimilarity (distance measures). Clusters are formed so as to minimize differences within groups while maximizing differences between groups (Griffith and Amrhein, 1997). Approaches to cluster analysis include hierarchical and non-hierarchical techniques. Hierarchical clustering is further divided into agglomerative and divisive methods. Agglomerative methods start with all initial observations and form classes by grouping the most similar cases together. In divisive clustering, an initial cluster encompassing all observations is recursively split into smaller, homogeneous sub-sets. Non-hierarchical techniques, such as the k-means procedure, require the definition of a set of initial clusters. Class membership for these groups is computed for all cases. As new cases are added, the initial clusters change. The procedure is repeated until the observed changes are below a pre-defined threshold.

Every clustering technique will result in the forming of clusters. It is therefore necessary to ensure that the clusters represent an actual grouping structure present in the data and are not artifacts of a particular algorithm. Bailey and Gatrell (1995) verified cluster analysis results by using two different clustering procedures. A general agreement of both cluster solutions is an indication that a natural grouping was correctly captured. On the other hand, widely different results suggest an artificial grouping due to the chosen clustering method.

Cluster analysis of substrate data was realized in two stages. First, initial hierarchical, agglomerative clustering was applied using the median method to merge two groups based on the distance between their centroids for all discriminating variables. At every step, equal weight is given to both groups to be combined. Squared Euclidean distances were used as dissimilarity measure. It was therefore possible to retain the influence of individual river segments on the overall cluster characterization. The number of clusters was plotted against the distance measure to reveal the number of natural groups inherent in the data (Griffith and Amrhein, 1997). The second stage in reducing and categorizing substrate information involved the use of non-hierarchical k-means clustering. Both results were compared to verify the general correspondence of the obtained cluster solutions.

4.2.2 Ancillary Data

Knighton (1984) identified stream width and valley gradient as pertinent to the characterization of bottom substrate and channel pattern. Consequently, these variables are used as potential predictors in the classification of *Substrate Type* and *Channel Pattern*. In order to derive the valley gradient, a digital elevation model (DEM) of the study area was created. A DEM is a digital, discrete, three-dimensional (x, y, z) representation of a continuous surface. The generation of a DEM is divided in 2 steps: the definition of a regularly spaced grid (x, y) covering the area of interest, and the choice of an appropriate interpolation algorithm to calculate elevation (z) for all grid cells from a number of locations with known elevations. Elevation data for this study were digitized contour lines contained in the 1:50,000 digital map sheet NTS 1N.13. The corresponding spatial resolution was accepted to be 25 m. A grid cell size of 30 m was chosen for the DEM. Since the locations of known elevation followed digitized contour lines rather than being randomly distributed over the study area, the INTERCON algorithm was used for interpolation (Eastman, 1997). In this procedure, elevation at unknown locations is determined by linear interpolation between contour lines.

The next step in the extraction of valley gradient was the overlay of the river course from the image data onto the DEM. At the intersection of river course and DEM, elevation for each cell and distance between cell centers were recorded. Valley gradient was calculated as the gradient between two cells provided that the down-stream cell showed a lower elevation than the corresponding up-stream cell. If this was not the case, the next down-stream cell with a lower elevation was used. This procedure is based on the assumption that elevation along the course of the river is continuously decreasing in the down-stream direction. As a result the river was divided into discrete sections with one gradient value each.

Stream width was calculated as the average width of discrete river sections. These sections were created using the center locations of all DEM cells that intersected with the course of the river. Voronoi polygons were created around these center locations. The river course was divided into sections of an approximately equal length of 30 m by overlaying water mask and Voronoi polygons. Given length and area of each section, average stream width was calculated according to:

$$W = A / L \qquad (4\text{-}1)$$

where

W = average width
A = area
L = length

4.2.3 Image Data

A first order correction for atmospheric path radiance was applied to the imagery using the method proposed by Chavez (1998). This type of correction is mandatory in the case of conversion of digital counts to radiance or reflectance units, for the comparison of data collected at different dates or prior to combining spectral bands through mathematical operations (Mather, 1987). Following Chavez's procedure, an initial value for path radiance was determined for one spectral band using the lowest value in the image histogram. This assumes that in any scene there are dark areas such as deep water or shadows where the expected reflected radiation is next to zero. In the presence of path radiance, however, the minimum value in a histogram will be greater than zero. The starting DN value corresponding to atmospheric path radiance was selected using *Band 1* since it is the spectral band most affected by atmospheric scattering. Once this value was found, an appropriate relative atmospheric scattering model was selected. The scattering models have the following form:

$$p = c\lambda^{x} \tag{4-2}$$

where

p = path radiance
λ = wavelength
x = parameter with values ranging from −4 to 0
c = constant

The DN equivalent to path radiance was calculated using Equation 4-2 given the initial path radiance value as extracted from the histogram. Table 4-1 lists different atmospheric conditions and the corresponding values of x.

Table 4-1: Atmospheric Conditions (after Chavez, 1988)

Atmospheric Conditions	**Value of x**
very clear	-4
clear	-2
moderate	-1
hazy	-0.7
very hazy	-0.5

Given the atmospheric conditions at the time of the image acquisition, the Rayleigh scattering model was chosen ($x = -4$) to correct for atmospheric path radiance. Atmospheric absorption was not corrected for due to the lack of appropriate data. This is not viewed as a problem because the images were recorded in the visible and near-infrared spectral regions where absorption effects are negligible (Van Stokkam *et al.*, 1993; Cracknell and Hayes, 1991).

The time required to record all images was 37 minutes. Therefore, the effect of changes in solar elevation was evaluated. While constant atmospheric and radiometric conditions are assumed within each scene, the same is not necessarily true between scenes. Spectral radiance, the physical quantity measured by the CASI sensor, varies with solar elevation. Slater (1980) proposed the following relationship between surface reflectance, radiance registered at a remote sensor, and solar elevation:

$$R_i = \{d^2 \cdot \pi \cdot (L_i - L_{pi})\} / E_i \cdot \cos(\varphi) \qquad (4\text{-}3)$$

where

R_i = surface reflectance
d = earth-sun distance
L_i = radiance at sensor in band i
L_{pi} = atmospheric path radiance in band i
E_i = spectral irradiance at top of atmosphere
φ = solar zenith angle

Equation 4-3 contains two important points. First, similar surfaces in two scenes might show different values of spectral radiance if there is a large discrepancy in solar elevation. Second, areas of high surface reflectance are more severely affected than features showing low reflectance. In order to verify if radiometric normalization of the flight swaths is necessary, areas of relatively high and low reflectance values were identified on images of the flight swaths 1 and 8. High reflectance targets were represented by sections of the Trans-Canada Highway (TCH). Segments of the Come By Chance River were selected as low reflectance test areas. In both cases, constant surface conditions throughout the study area were assumed.

All images were geometrically corrected and registered to UTM coordinates (Zone 22, NAD83). Ground control points (GCP's) with known coordinates were identified on reference maps (scale ranging from 1:5,000 to 1:12,500) and on the imagery. A highly accurate approach to the geometric correction of airborne imagery is the thin plate spline method. However, this requires a large number of GCP's as well as the recording of reference coordinates using a global positioning system (GPS), which was not available at the time of this investigation. Therefore, first order polynomial regression analysis and nearest-neighbour resampling were applied to transform image coordinates. Registration accuracy was assessed using the root-mean-square (RMS) error, which is the standard deviation of the residuals of both Easting and Northing coordinates. The resolution of the corrected images was set to be 2 m to allow for appropriate coverage of the narrowest river sections.

For the prediction of *Substrate Type* and *Channel Pattern* a binary mask was created to separate water covered areas from land. The river course was digitized on-screen to exclude extensive areas of shadow over land. Visual inspection of *Band 4*, colour composite images, air photos and the histogram of *Band 4* were used to select a threshold value for the separation of water from land. The threshold value for pixels representing water was set at DN < 90 in *Band 4*.

Extracting information on bottom substrate or channel pattern required that the impact of the water column on the reflected radiation be minimized. Due to the absence of reliable depth measurements, the method proposed by Lyzenga (1978) was modified. First, pixels of submerged areas over bedrock substrate were transformed using the natural log. The linear variability observed in the log-transformed DN values reflects variations in water depth since the type of bottom substrate is kept constant. Next, principal component analysis (PCA) was applied to these pixels. The first principal component is aligned along the direction of maximum sample variance, thus representing water depth variability. Consequently, the remaining components contain information unrelated to water depth but related to bottom type variation. In addition, the approach of Khan *et al.* (1992) was followed by applying PCA directly to pixels of varying bottom types and without prior log-transformation.

In the land cover classification, the normalized difference vegetation index (NDVI) was used besides the original spectral bands. This index is sensitive to spectral differences between vegetated and non-vegetated areas and is calculated as:

$$NDVI = (L_{ir} - L_r) / (L_{ir} + L_r) \tag{4-4}$$

where

L_{ir} = radiance in the near-infrared spectral band
L_r = radiance in the red spectral band

The NDVI has also been used to differentiate between types of vegetation (Curran, 1983). In addition, PCA was applied to the land cover training data to enhance discrimination between land cover classes. The resulting principal components were included in the pool of potential predictor variables for the type of land cover.

4.3 Training and Test Area Generation

For each of the habitat parameters *Channel Pattern* and *Substrate Type*, 50 river segments were selected at random. Where applicable, both features were recorded for the same segment. The resulting database was randomly split in half to yield one training data set and one verification, or test, data set. A buffer of 20 m was created around each transect to account for errors in transect positioning. No pixel was collected within this buffer. Approximately 50 pixels were collected at random from each segment to assure correct representation of actual class occurrences in training and test data. Training and test areas for the classification of *Land Cover* were identified with the aid of aerial photographs. A total of 403 locations were selected at random over all images. Polygons containing an average of 60 pixels were delineated around these locations. Boundary pixels were not included. One half of the polygons were used to extract training data, while the other half was used to collect test data. The large number of training and test sites was required to account for any radiometric differences between the scenes. In both training and verification data sets the number of pixels was further reduced by randomly selecting one third of all cases for subsequent analysis.

4.4 Selection of Predictor Variables

The pool of potential predictor variables for each habitat parameter is presented in Table 4-2. For each variable, plots of mean response and standard deviations of all habitat parameter categories were examined. The coefficient of variation (CV) was used as a measure of group variability independent from the magnitude of the mean. Since the CV is only defined for positive values, variables containing negative values (i.e. principal components, NDVI) were adjusted by adding the variable minimum to the mean in each category before calculating the coefficient of variation (Neter *et al.*, 1993). Bivariate correlations between independent variables and standardized distances, or differences, between group means were analyzed to select the final predictor variables for each habitat parameter.

Table 4-2: Potential Predictor Variables

Potential Predictor Variables	**Habitat Parameters**		
	Substrate Type	*Channel Pattern*	*Land Cover*
Original spectral bands	*Band 1* *Band 2* *Band 3* *Band 4*	*Band 1* *Band 2* *Band 3* *Band 4*	*Band 1* *Band 2* *Band 3* *Band 4*
Principal component transform applied to ln-transformed spectral bands	*PC1_ln* *PC2_ln* *PC3_ln* *PC4_ln*	*PC1_ln* *PC2_ln* *PC3_ln* *PC4_ln*	- - - -
Principal component transform applied to original spectral bands;	*PC1* *PC2* *PC3* *PC4*	*PC1* *PC2* *PC3* *PC4*	*PC1* *PC2* *PC3* *PC4*
Stream width	*Width*	*Width*	-
Valley gradient	*Gradient*	*Gradient*	-
Normalized difference vegetation index;	-	-	*NDVI*

Several predictor variables were generated as linear combinations of the original spectral bands. Therefore, it was necessary to assess the degree of multicollinearity present among the independent variables. Multicollinearity exists if two or more predictors are highly correlated. Correlation coefficients exceeding a value 0.70 can lead to inflated significance levels and logical problems due to duplication of information (Tabachnick, 1996). Extreme correlations of greater than 0.90, for example, are likely to prevent matrix inversion calculations in multiple regression and discriminant function analysis (Mather, 1976). Correlation among independent variables was analyzed using the non-parametric Spearman rank correlation coefficient (Walford, 1995).

Standardized distances between category means were calculated in two stages. First, all predictor variables were standardized according to:

$$z = (x - m) / s \qquad (4\text{-}5)$$

where

z = standardized score
x = raw variable score
m = variable mean
s = variable standard deviation

Secondly, the distance between two categories was calculated in each standardized predictor variable as following:

$$d_{ij} = m_{zi} - m_{zj} \qquad (4\text{-}6)$$

where

d_{ij} = distance between category means in standardized variable z
m_{zi} = mean of category i in standardized variable z
m_{zj} = mean of category j in standardized variable z

As a result, variables measured in different units were made directly comparable. Moreover, the calculated differences between category means were expressed in units of standard deviation and therefore provided a better indication of group separability than the original units (Davis, 1986). To minimize the risk of falsely accepting spurious relationships, significance levels and confidence intervals were calculated for each distance using the Scheffé procedure. This method is extremely conservative and allows for the simultaneous comparison of all group differences (Tabachnick, 1996). Group differences were accepted to be statistically significant at minimum significance level of 0.05.

A simple strategy was adopted in selecting the final predictor variables for each habitat parameter. First, independent variable showing the largest standardized distance was identified. If this value was significantly different from the corresponding difference in any other variable at a significance level of 0.05, the variable was selected. If two or more variables showed group differences not significantly different from the largest distance observed, they were included as predictors, provided that none of these variables were highly correlated with each other. However, if high correlations were observed, these variables were dropped from the analysis. High correlations between predictors were accepted when the largest mean difference observed was significantly different from any other variable in more than one instance. In this case, the presence of multicollinearity was accounted for by adjusting significance levels during the classification process. This procedure is discussed in detail in the following section.

4.5 Classification of Habitat Parameters

Classification of habitat parameters was carried out using decision tree analysis (DTA). The exhaustive partitioning procedure developed by Biggs *et al.* (1991) establishes relationships between predictor and response variable based on statistical significance. It is a statistically robustness technique and permits the integration of continuous and categorical data. For these reasons, exhaustive partitioning was selected as classification algorithm. A separate decision tree was grown for each of the habitat parameters *Substrate Type*, *Channel Pattern* and *Land Cover*. Only splits showing a significance level of 0.05 or higher were accepted so as to ensure statistically strong relationships between predictor and dependent variables. If several predictor variables passed the significance threshold at a given node, the variable with the highest significance was selected to partition the data. All observed significance levels were corrected using the Bonferroni adjustment factor to minimize the risk of falsely accepting relationships that were not significant. The Bonferroni adjustment factor was also used to account for the presence of highly correlated predictor variables. The effect of this adjustment was a further reduction of the initially obtained significance levels so as to suppress relationships which are caused by the duplication of information in the predictor variables.

Before applying classification rules derived from DTA to the classification of unknown cases, redundant branches of a decision tree must be removed to avoid overfitting. This process is called pruning and generally involves the use of an independent test sample that was not used in the creation of the decision tree (Breiman *et al.*, 1984). If DTA is applied to remotely sensed data, three data sets are required: one for training, one for pruning, and one to assess classification accuracy (Friedl and Brodley, 1997). However, for practical reasons it was not possible to allocate sampling areas for three separate data sets in the present study. Instead, the risk of overfitting was minimized by specifying a threshold sample size, or stop size, for the creation of new nodes. That is, if a given node contained fewer observations than the specified stop size it was not further partitioned.

The selected threshold sample size for each habitat parameter was related to the average sample size in an individual sampling unit. In the case of *Substrate Type* and *Channel Pattern*, individual sampling units are defined by the average length and width of river sections used in the derivation of stream width and valley gradient. Given an average width of 20 m, an average length of 30 m and a pixel resolution of 2 m, the corresponding stop size was set to be 150. Individual sampling units for the habitat parameter *Land Cover* are defined by the size of individual training areas. Accordingly, a threshold sample size of 60 was selected corresponding to the average sample size per training area. The decision trees were interpreted as statistically significant classification rules in "IF...THEN" format. These rules were subsequently implemented as a FORTRAN program to classify all of the imagery.

Integration of continuous predictors in decision tree analysis required the selection of discrete categories for each variable. Class intervals were derived with the aid of exploratory data analysis as described by Vellemann and Hoaglin (1981). With this approach, a data set is described and partitioned around the median. Furthermore, the data can be divided into equal parts to contain an eighth, a fourth, half, etc. of all observations. All image variables were divided into eight intervals of equal size. Stream width was divided into four discrete categories, while three equal sized intervals were selected for the variable valley gradient.

4.6 Accuracy Assessment

Classification was evaluated using a verification sample that was not previously used in the derivation of decision trees for the habitat parameters *Substrate Type*, *Channel Pattern* and *Land Cover*. The result of this analysis was summarized using contingency tables (Congalton and Green, 1993; Green *et al.*, 1993; Fitzgerald and Lees, 1994; Janssen and Van der Wel, 1994; Lark, 1995; Stehman, 1997).

Table 4-3 contains a schematic contingency table to illustrate the various measures of accuracy that were extracted. The columns of this table refer to the reference data, whereas the rows represent the classification result. Accuracy measures used in the assessment of classification performance are defined in Table 4-4. The misclassification rate of pixels in the test data set belonging to a particular substrate type is given by the respective errors of omission, designated as "OE" Table 4-4. Conversely, errors of commission refer to pixels wrongly assigned to groups to which they do not belong. Commission errors for each category are obtained from Table 4-4 as "CE". Accuracy measures include overall classification accuracy "OA", user's accuracy "UA" and producer's accuracy "PA". The overall classification accuracy represents the proportion of correctly classified observations across all categories. The user's accuracy value denotes the probability that a classified pixel actually belongs in the category it was classified. This measure is related to the error of

Table 4-3: Schematic Contingency Table

Classified Data	Reference Data: $i=1$	$i=2$	...	$i=q$	Row Total
$j=1$	c_{11}	c_{12}	...	c_{1q}	$\sum_{i=1}^{q} c_{1i}$
$j=2$	c_{21}	c_{22}	...	c_{2q}	$\sum_{i=1}^{q} c_{2i}$
...	...	...	...	...	...
$j=q$	c_{q1}	c_{q2}	...	c_{qq}	$\sum_{i=1}^{q} c_{qi}$
Column Total	$\sum_{j=1}^{q} c_{j1}$	$\sum_{j=1}^{q} c_{j2}$	...	$\sum_{j=1}^{q} c_{jq}$	

i = category in reference data
j = category in classified data
q = number of categories
c_{ij} = elements of the confusion matrix

Table 4-4: Measures of Accuracy

Accuracy Measure	Definition	Calculation
Overall Accuracy OA	the proportion of overall correctly classified samples	$OA = \frac{1}{n} \cdot \sum_{i,j=1}^{q} c_{ij}$
Overall Classification Error OE	the proportion of overall incorrectly classified samples	$OE = 1 - OA$
User's Accuracy UA	the conditional probability $p(j=i \mid j=k)$ that a sample is correctly allocated given that it was classified as k	$UA = \frac{c_{kk}}{\sum_{j=1}^{q} c_{kj}}$
Commission Error CE	for two classes (k,l) the conditional probability $p(i=l \mid j=k)$ that a sample classified as k actually belongs to category l	$CE = 1 - UA$
Producer's Accuracy PA	the conditional probability $p(j=i \mid i=k)$ that a sample which actually belongs to class k is correctly allocated	$PA = \frac{c_{kk}}{\sum_{i=1}^{q} c_{ik}}$
Omission Error OE	for two classes (k,l) the conditional probability $p(j=k \mid i=l)$ that a sample which actually belongs to category l is classified as k	$OE = 1 - PA$
Kappa Index of Agreement κ	proportion of overall correctly classified samples, adjusted for the possibility of random assignment	$\kappa = \frac{OA - CA}{1 - CA}$

commission. Likewise, producer's accuracy is related to errors of omission. It designates the probability that a pixel belonging to a given category in the verification data set is correctly classified. The kappa index of agreement is a measure of overall classification performance which takes into account correct assignment of pixels by chance. It is interpreted as the proportion by which a given overall accuracy exceeds the accuracy expected for a random classification. Calculation of the chance agreement term (CA) requires the derivation of a new matrix consisting of the products of row and column totals of the original contingency table. CA is then calculated by summing the diagonal and dividing by the grand total.

Jensen (1986) suggests the derivation of confidence intervals around estimated classification accuracy values using the normal approximation to the binomial distribution. Accordingly, confidence intervals for the overall accuracy of individual habitat parameters were calculated as:

$$C_l = P - \left(z \cdot \sqrt{\frac{P \cdot (1-P)}{n}} \right) \qquad (4\text{-}6)$$

$$C_u = P + \left(z \cdot \sqrt{\frac{P \cdot (1-P)}{n}} \right) \qquad (4\text{-}7)$$

where

P = proportion of correctly classified pixels
C_l = lower confidence limit
C_u = upper confidence limit
z = 1.96 (corresponding to a probability of 95 % in a normal distribution)

4.7 Suitability Mapping

This study focused primarily on the extraction of individual salmon habitat parameters. Two of these parameters, i.e. *Substrate Type* and *Channel Pattern*, were used in the assessment of suitable spawning habitat for Atlantic salmon. According to Dubois and Gosselin (1989), spawning habitat priority was calculated as:

$$P = B * C \tag{4-8}$$

where

P = spawning habitat priority
B = priority weight of substrate type
C = priority weight of channel pattern

The corresponding priority weights are presented in Table 4-5. A description of different spawning habitat types resulting from this analysis is given in Table 4-6. Spawning habitat priority can range from 0 to 4. A value of 4 indicates good spawning habitat suitability, while values of 1 or 2 indicate intermediate suitability. A value of 0 designates areas unsuitable for spawning.

Table 4-5: Priority Weights for Substrate Type and Channel Pattern

Substrate Type (Priority Weight)				**Channel Pattern (Priority Weight)**				
Gravel (1)	Rubble (2)	Boulder (2)	Bedrock (0)	Run (1)	Riffle (2)	Steady (1)	Flat (1)	Rapid (2)

Table 4-6: Salmon Spawning Habitats (after Dubois and Gosselin, 1989)

Type of Spawning Habitat	**Spawning Habitat Priority (P)**	**Suitability**
Habitat I	4	Good
Habitat II	1,2	Fair
Habitat III	0	Not Suitable

Chapter 5.0: Results

5.1 Introduction

In the first part of this chapter, the data pre-processing results are presented. Next, the process of selecting predictor variables for *Substrate Type*, *Channel Pattern* and *Land Cover* is described, followed by a presentation of decision tree classifications for each habitat parameter. The accuracy of the respective classifications is assessed. Finally, it is shown how the habitat parameters *Substrate Type* and *Channel Pattern* were used to model spawning habitat suitability for Atlantic salmon.

5.2 Pre-Processing

The following sections describe the results of substrate data categorization as well as image correction and transform procedures required in order to subject the data to further analysis.

5.2.1 Substrate Data

The number of natural substrate groups inherent in the data was estimated by plotting the cluster solutions obtained from hierarchical clustering against the distance associated with the merging of groups at each level. Natural groups are indicated by a distinct visual separation from one cluster solution to the next. In Figure 5-1, a sharp increase in distance is observed for the transition from seven to six and from four to three terminal clusters. The first of these transitions reflects the fact that the type of bottom substrate is represented by seven individual grain size variables. The second transition suggests the presence of four natural substrate classes.

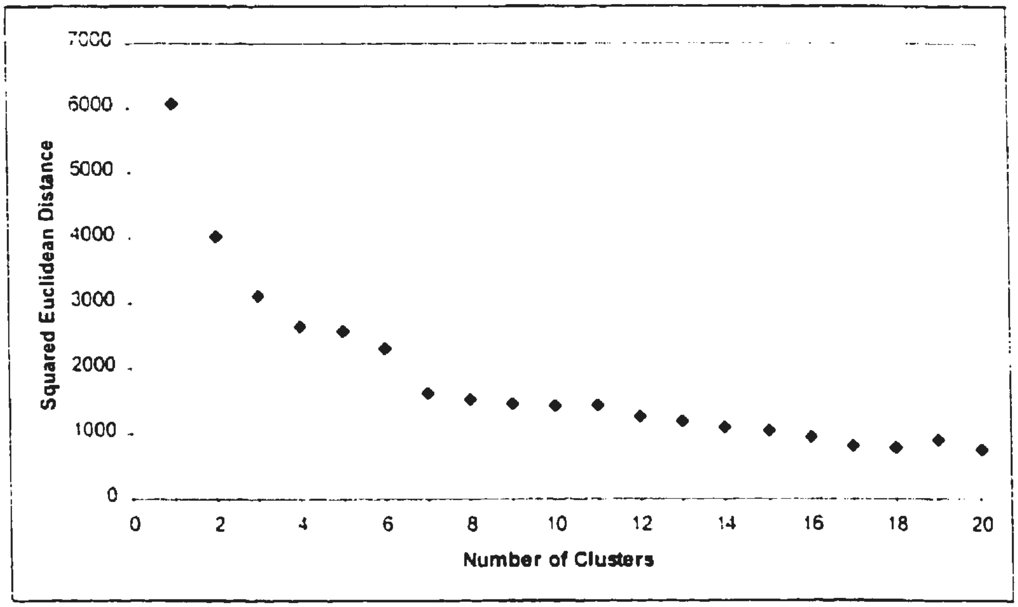

Figure 5-1: Cluster Solutions of Median Method

In order to verify that the substrate classes obtained from hierarchical clustering were not spurious, clustering was repeated using the non-hierarchical k-means procedure. The resulting cluster centers of both approaches are presented in Table 5-1. They represent average proportions of the original grain size variables for each cluster. Average proportions greater than 10 % are highlighted. Both cluster solutions are similar in the characterization of individual clusters. Cluster 1 is dominated by the variables Gravel, Pebble and Cobble. Cluster 2 is characterized by the grain size classes Rubble, Cobble and Gravel. In Cluster 3, the average proportions are largest for Boulder and Rubble, while Cluster 4 represents the grain size variables

Bedrock and Rubble. Although the observed clusters are not mutually exclusive due to overlap in average grain size proportions, the resulting categories reflect a clear trend from finer to coarser substrate.

Table 5-1: Cluster Centers of Median and K-Means Methods

	Cluster Centers (Average Proportion [%])							
Grain Size Variable	Cluster 1 (Median)	Cluster 1 (K-Means)	Cluster 2 (Median)	Cluster 2 (K-Means)	Cluster 3 (Median)	Cluster 3 (K-Means)	Cluster 4 (Median)	Cluster 4 (K-Means)
Bedrock	2.11	1.96	2.72	3.31	1.90	4.42	**66.92**	**69.58**
Boulder	2.31	1.40	3.85	3.55	**44.52**	**42.92**	9.28	8.22
Rubble	4.27	2.11	**40.64***	**30.72***	**35.24**	**32.75**	**13.64**	**12.97**
Cobble	**20.58***	**10.38***	**40.00**	**41.72**	9.62	9.75	7.69	7.17
Pebble	**21.95**	**28.89**	4.38	4.26	4.24	5.04	1.00	1.00
Gravel	**45.86**	**49.89**	**11.13***	**19.22***	6.29	7.00	4.87	4.53
Fine	7.00	9.53	1.00	1.00	1.86	1.75	1.00	1.00

* = cluster centers are significantly different at a significance level of 0.05

Significant differences between both cluster solutions exist only for Clusters 1 and 2. In Cluster 1, this difference is observed for the grain size variable Cobble. The average proportion for the median method is with 20.58 % twice as large as for the k-means procedure. In Cluster 2, the average proportions of both Rubble and Gravel differ for the median and k-means methods by approximately 10 %. These differences become more apparent from the cross-tabulation of both procedures in Table 5-2. Most of the discrepancy is accounted for by 18 river segments that are classified as Cluster 1 in the median method and as Cluster 2 in the k-means method.

Table 5-2: Cross-Tabulation of Median and K-Means Clustering Results

Median Method	K-Means Method: Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	45	18	1	0
Cluster 2	0	39	0	0
Cluster 3	0	0	21	0
Cluster 4	0	1	2	36

When examined for their grain size composition, the segments showed average proportions for Cobble and Gravel of 45.56 % and 37.78 %, respectively. This indicates that river segments with relatively high proportions of both Cobble and Gravel substrate cause the disagreement between both clustering algorithms. Therefore, difference between clustering techniques reflects confusion between grain size classes inherent in the data. The median cluster solution featuring four terminal groups was selected to represent substrate information in all subsequent analysis. According to the highest average grain size proportions presented in Table 5-1, the final substrate categories were labelled *Gravel*, *Rubble*, *Boulder* and *Bedrock*.

5.2.2 Image Corrections and Transforms

Pre-processing of image data involved atmospheric and geometric correction, radiometric calibration and the calculation of image transforms. Atmospheric correction involved the estimation of a DN value equivalent to path radiance in each spectral band. The corresponding values are shown in Table 5-3. The effect of path radiance was corrected by subtracting these values from the respective spectral bands.

In order to investigate the need for radiometric calibration due to variations in solar elevation, the mean spectral response of high (paved highway) and low (water) reflectance areas after correcting for path radiance was compared. The result of this comparison is presented in Table 5-4. As expected, the differences are larger between the paved highway test areas than between the water-covered locations.

Table 5-3: DN Values Equivalent to Path Radiance and Time of Acquisition

	Band 1	Band 2	Band 3	Band 4	Local Time
Swath 1	26	15	9	6	11 : 21
Swath 2	30	17	11	7	11 : 28
Swath 4	30	17	11	7	11 : 38
Swath 6	32	18	11	8	11 : 45
Swath 7	32	18	11	8	11 : 48
Swath 8	32	18	11	8	11 : 53
Swath 9	33	18	12	8	11 : 58

Table 5-4: Comparison of High and Low Reflectance Areas

	Mean Spectral Response [DN] of High Reflectance Target (Paved Highway)							
	Band 1		**Band 2**		**Band 3**		**Band 4**	
	Swath 1	**Swath 8**	**Swath 1**	**Swath 8**	**Swath 1**	**Swath 8**	**Swath 1**	**Swath 8**
Mean	171.5	168.1	189.1	189.7	172.5	177.8	148.3	153.9
Standard Deviation	4.4	10.3	9.1	12.4	16.6	19.2	11.8	12.2
Minimum	160	143	169	139	142	122	123	125
Maximum	184	179	210	208	210	216	174	180

	Mean Spectral Response [DN] of Low Reflectance Target (Water)							
	Band 1		**Band 2**		**Band 3**		**Band 4**	
	Swath 1	**Swath 8**	**Swath 1**	**Swath 8**	**Swath 1**	**Swath 8**	**Swath 1**	**Swath 8**
Mean	9.1	7.7	8.8	9.4	8.1	6.9	15.2	14.4
Standard Deviation	3.9	2.7	4.3	4.4	4.0	4.1	2.9	4.3
Minimum	3	0	3	3	3	3	10	9
Maximum	17	11	18	18	18	20	22	26

In all instances, no significant differences in mean spectral response were observed between corresponding test pixels of Swaths 1 and 8 at a significance level of 0.05. These results indicate that correction for differences in solar elevation was not required.

The imagery was geometrically corrected for two reasons. First, resampling to a common coordinate system ensures a constant pixel resolution for all images. In uncorrected images the scale would vary within as well as between scenes due to sensor movement. Second, geometric registration was required for the integration of image and digital elevation data. Results of the geometric registration of image data to UTM coordinates are presented in Table 5-5. Given the lack of distinct features suitable for use as ground control points on the maps used in the correction process, a first-order polynomial transform was applied to register all images. Extreme geometric distortions in parts of the image data caused comparatively large registration errors of two to three times the resolution. However, since the resolution of the elevation data is 30 m, observed the image registration errors are well within the bounds required for data integration.

Table 5-5: Image Registration Characteristics

	Number of GCP's	Resolution [m]	Locational Accuracy [m]
Swath 1	10	2	5.3
Swath 2	15	2	5.1
Swath 4	11	2	4.7
Swath 6	8	2	5.8
Swath 7	8	2	6.3
Swath 9	5	2	4.4

Image transforms applied to training data of the habitat parameters *Substrate Type* and *Channel Pattern* included principal component analysis (PCA) of original and log-transformed spectral bands. The characteristics of these transforms are presented in Tables 5-6 and 5-7.

Table 5-6: Characteristics of PCA Applied to Original Spectral Bands

	Eigenvalue	Variance Explained [%]	Cumulative Variance [%]	
Component 1	968.226	86.4	86.4	
Component 2	105.169	9.4	95.8	
Component 3	42.484	3.8	99.6	
Component 4	4.714	0.4	100	
Component Loadings				
	Component 1	**Component 2**	**Component 3**	**Component 4**
Band 1	0.343	0.265	0.719	0.544
Band 2	0.507	0.315	0.225	-0.770
Band 3	0.527	0.427	-0.657	0.329
Band 4	0.590	-0.806	-0.024	0.052

Table 5-7: Characteristics of PCA Applied to log-Transformed Spectral Bands

	Eigenvalue	Variance Explained [%]	Cumulative Variance [%]	
Component 1	1.527	88.8	88.8	
Component 2	0.097	5.7	94.5	
Component 3	0.085	4.9	99.4	
Component 4	0.0097	0.6	100	
Component Loadings				
	Component 1	**Component 2**	**Component 3**	**Component 4**
ln(Band 1)	0.493	0.563	-0.568	0.344
ln(Band 2)	0.515	-0.043	-0.111	-0.849
ln(Band 3)	0.567	-0.730	-0.001	0.381
ln(Band 4)	0.413	0.385	0.816	0.125

In both PCA transforms, the proportions of variance explained by each of the components are very similar. When PCA is applied to the original spectral bands, 86 % of the variance in the data is explained by the first component. The second component accounts for approximately 10 % of the total variation. The proportion of variance explained by the third principal component is 3.8 %.. The fourth component accounts for less than 1 % of the variability in the data. Likewise, the first principal component resulting from PCA applied to log-transformed bands explains 88.8 % of the variability. However, the second and third components account for an almost equal amount of variance with 5.7 % and 4.9 %, respectively, while less than 1 % is explained by the fourth component. Both approaches do not differ significantly in accounting for the overall variability in the data.

The result of PCA applied to land cover training data is presented in Table 5-8. In this case, the first component accounts for only 75.1 % of data variability, while more than 20 % of the total variability is explained by the second principal component. The third and fourth components account for 2.1 and 0.3 % of the total variance, respectively.

Table 5-8: Characteristics of PCA Using Land Cover Training Data

	Eigenvalue	**Variance Explained [%]**	**Cumulative Variance [%]**	
Component 1	7847.43	75.1	75.1	
Component 2	2352.58	22.5	97.6	
Component 3	223.82	2.1	99.7	
Component 4	29.25	0.3	100	
Component Loadings				
	Component 1	**Component 2**	**Component 3**	**Component 4**
Band1	0.400	-0.326	0.714	0.473
Band2	0.512	-0.226	0.149	-0.815
Band3	0.599	-0.274	-0.677	0.328
Band4	0.468	0.876	0.093	0.068

5.3 Selection of Predictor Variables

5.3.1 Substrate Type

Group means, standard deviations and coefficients of variation of individual substrate categories are presented in Figure 5-2 and in Appendix A. In the original spectral bands, the observed differences between substrate categories were low with respect to the variability in each group. In *Band 1*, mean values ranged from 14.51 for the class *Boulder* to 23.07 for *Rubble*. Variability was generally high in all categories. The class *Rubble* showed the lowest coefficient of variation with a value of 51.02 %, while the corresponding values in the remaining categories ranged from 60.58 % for *Gravel* to 64.02 % for *Boulder*. The substrate category *Boulder* showed the lowest mean response in *Band 2* with an average DN value of 20.6. It is clearly separated from the classes *Gravel* and *Rubble*, which showed a nearly identical mean response of 28.46 and 28.87, respectively. All substrate categories in *Band 2* showed considerable spectral overlap with standard deviations ranging from 12.65 for *Boulder* to 16.56 in the class *Bedrock*. Variability remains high with coefficients of variation from 44.52 to 61.98 %. Mean response of substrate categories in *Band 3* were similar to the spectral behaviour observed in *Band 2*. However, group variability in this variable is highest for the classes *Boulder* (77.10 %) and *Bedrock* (70.35 %) and lowest for *Gravel* (44.98 %). The variable *Band 4* showed the highest spectral response for all substrate groups. Mean values ranged from 30.25 in the class *Boulder* to 37.32 for *Rubble*. Group variability in this variable was lowest in the categories *Gravel* and *Rubble*, with coefficients of variation of 53.87 and 55.17 %, respectively. The highest variability was observed in the classes *Boulder* (58.81 %) and *Bedrock* (60.52 %).

The principal components of log-transformed spectral bands also showed considerable

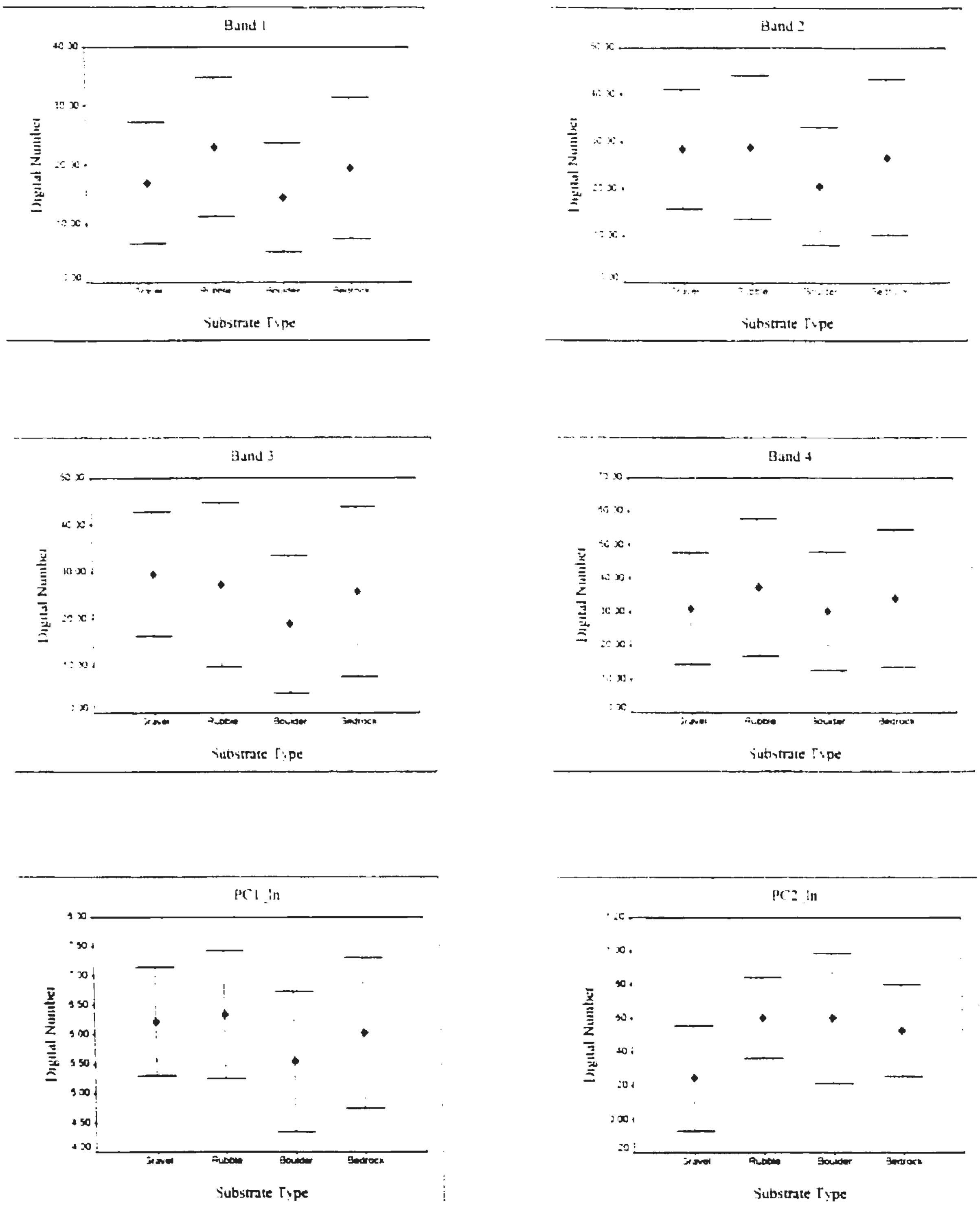

Figure 5-2: Group Means and Standard Deviations of Substrate Categories

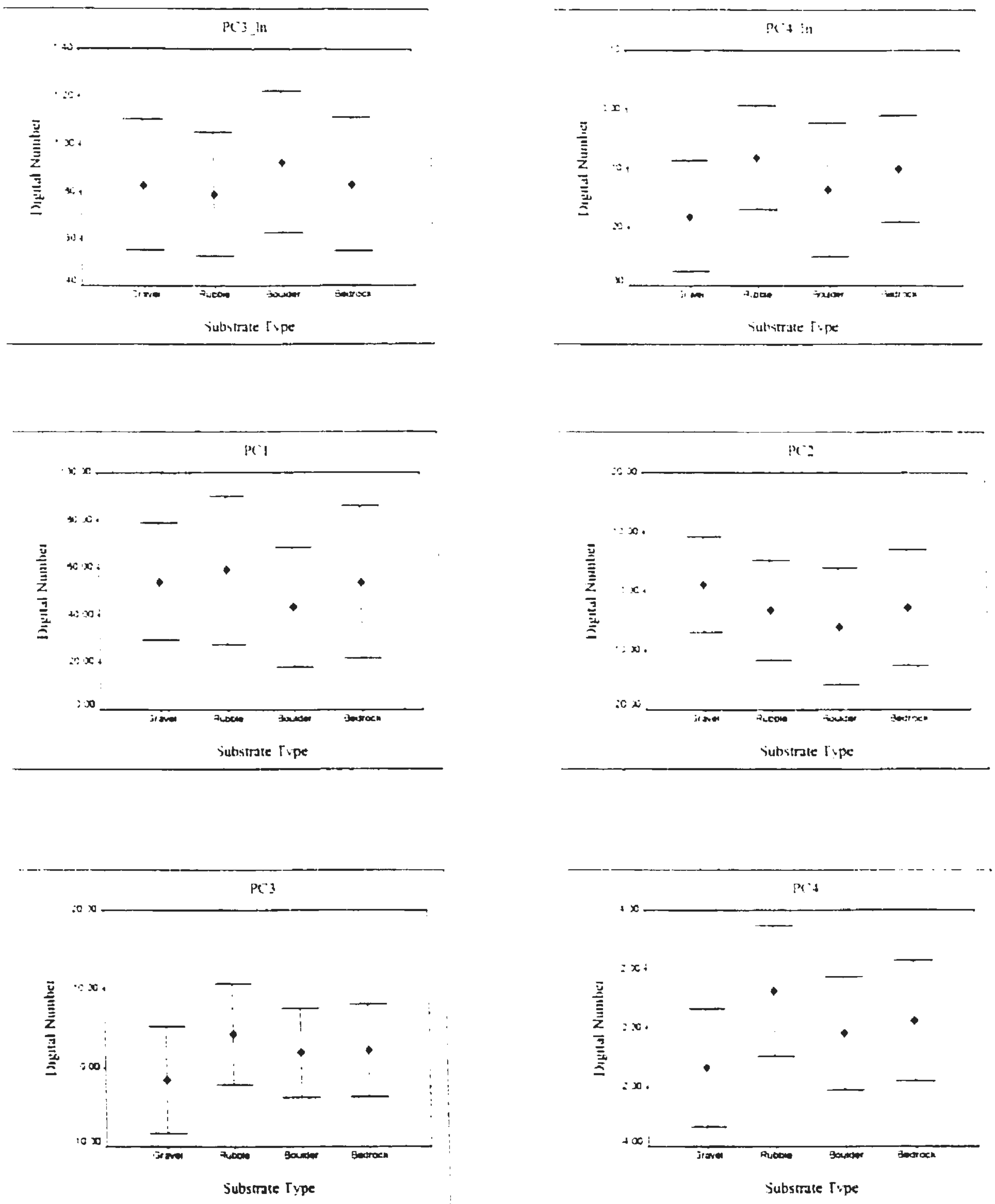

Figure 5-2: Group Means and Standard Deviations of Substrate Categories (continued)

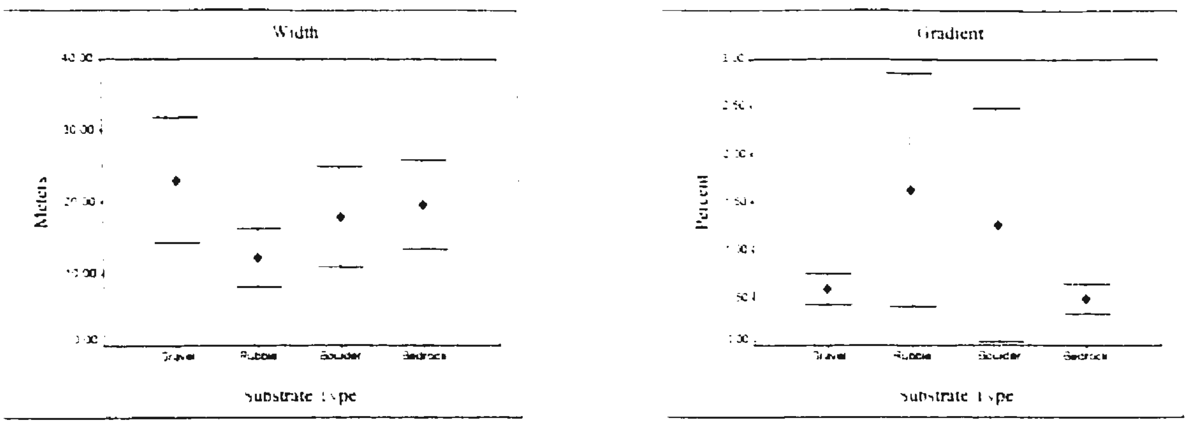

Figure 5-2: Group Means and Standard Deviations of Substrate Categories (continued)

overlap between all substrate categories. *PC1_ln* largely resembled the response in the original spectral bands. The category *Boulder* showed the lowest mean response with a value of 5.54. Conversely, the highest response occurred in the category *Rubble* with a mean DN of 6.34. Group variability in this variable was characterized by coefficients of variation ranging from 14.79 % for *Gravel* to 21.48 % in the category *Bedrock*. In *PC2_ln*, the category *Gravel* showed a low mean response of 0.24. Mean values in the remaining classes ranged from 0.53 to 0.60. Variability was highest in the category Gravel with 31.63 %, while the groups *Rubble*, *Bedrock* and *Boulder* showed coefficients of variation of 17.91 %, 22.05 % and 29.10 %, respectively. Spectral overlap of substrate classes was also dominant in *PC3_ln*. Observed mean response values ranged from 0.79 to 0.83, with standard deviations from 0.26 to 0.30. Variability was similar in all categories with coefficients of variation from 23.85 to 24.78 %. In the variable *PC4_ln*, the substrate category *Gravel* was set apart from the classes *Rubble*, *Boulder* and *Bedrock* through a low mean response of –0.18 and a standard deviation of 0.09. The highest response occurred in the substrate type *Rubble* with a mean value of -0.08. Group variabilities were low and ranged from 9.57 % in the category *Rubble* to 12.50 % in the class *Boulder*.

In *PC1*, the observed pattern of mean response and group variability was similar to the spectral behaviour of substrate categories in the original spectral bands. Mean response was lowest in the category *Boulder* (43.26) and highest for *Rubble* (58.97). Group variability ranged from 45.52 % in the class *Gravel* to 59.71 % for *Bedrock*. Mean response in the variable *PC2* was characterized by mean values from 1.10 for *Gravel* to –5.95 for *Boulder*, while the observed standard deviations ranged from 8.08 to 9.95. Variability was lowest in the categories *Gravel* (13.22 %) and *Rubble* (14.88 %). The highest coefficients of variation in this variable were obtained for *Bedrock* (17.30 %) and *Boulder* (18.41 %). Pixels of the substrate type *Gravel* showed the lowest mean response in the variable *PC3* with a value of –1.58. The remaining substrate categories were characterized by mean values from 1.90 to 4.23. Group standard deviations were similar in all classes and ranged from 5.66 to 6.77, resulting in coefficients of variation from 15.52 % in the category *Boulder* to 20.52 % for *Gravel*. The spectral behaviour of substrate categories in *PC4* indicated separability of the class *Gravel* from *Rubble*, *Boulder* and *Bedrock* through a comparatively low mean response of –1.34. The remaining mean values ranged from –0.17 to 1.95, while standard deviations in all categories varied from 1.92 to 2.22. Variability was highest in the class *Gravel* with 28.99 %. Coefficients of variation in the remaining categories were similar and ranged from 23.39 to 24.03 %.

Differences between individual substrate classes were more apparent in the variable *Width*. This applied especially to the differentiation between the categories *Gravel* and *Rubble*. The respective mean values were 23.08 and 12.22, with corresponding standard deviations of 8.69 and 4.05. Similar mean response values of 17.95 and 19.62 were observed in the respective categories of *Boulder* and *Bedrock*. Variability in the variable *Width* ranged from 31.70 % in the class *Bedrock* to 38.83 % in the category *Boulder*. Mean response and variability in the variable

Gradient were similar for the categories *Gravel* and *Bedrock*. The observed mean values were 0.44 and 0.34, with standard deviations of 0.17 and 0.15, respectively. These categories were differentiated against the classes *Rubble* and *Boulder*, which show respective mean values of 1.48 and 1.11, and a standard deviation of 1.22 in both cases. Group variability was large for *Rubble* and *Boulder* with respective coefficients of variation of 82.43 and 109.91 %.

Standardized distances between the mean values of substrate categories in all predictor variables are presented in Table 5-9. The largest distance between means for each pair of substrate categories is underlined. Distances that are not significantly different from the largest distance for a given pair of categories are printed in bold letters. The significance level is set at 0.05. Standard error, significance level and confidence limits corresponding to each distance are listed in Appendix A. Rank correlation coefficients between independent variables are presented in Table 5-10.

The variable *Width* showed the largest difference in mean spectral response between the categories *Gravel* and *Rubble*. The observed value of 1.370 was not significantly different from the corresponding distances observed for the variables *Gradient*, *PC2*_ln and *PC4*. Class means of the substrate types *Gravel* and *Boulder* were furthest apart in the variable *PC2_ln* with a standardized distance of 1.052. The 95 % confidence interval associated with this value overlapped with the corresponding intervals of *PC2* and *Gradient*. Likewise, *PC2_ln* showed the largest difference between the categories *Gravel* and *Bedrock*. In this case, the observed distance of 0.839 was not significantly different from the corresponding values in *PC4_ln*, *PC3* and *PC4*. The largest difference between group means for the classes *Rubble* and *Boulder* appeared in *Band 1*, with a standardized distance of 0.756. This value was significantly different only from *PC2_ln* and *PC2*. The variable *Gradient* showed the largest difference between the substrate categories *Rubble* and *Bedrock* as well as between *Boulder* and *Bedrock*. In both instances, the observed values of 1.392 and 0.940

Table 5-9: Standardized Distances between Substrate Category Means

	Band 1	Band 2	Band 3	Band 4	PC1_ln	PC2_ln	PC3_ln	PC4_ln	PC1	PC2	PC3	PC4	Width	Gradient
Gravel vs. Rubble	0.539	0.028	0.139	0.330	0.108	**1.060**	0.135	0.969	0.166	0.458	0.879	**1.152**	**1.370**	**1.261**
Gravel vs. Boulder	0.217	0.532	0.650	0.040	0.587	**1.052**	0.338	0.433	0.375	**0.751**	0.526	0.521	0.648	**0.811**
Gravel vs. Bedrock	0.219	0.118	0.232	0.157	0.162	**0.839**	0.007	**0.776**	0.001	0.414	**0.576**	**0.707**	0.436	0.128
Rubble vs. Boulder	**0.756**	**0.560**	**0.512**	**0.371**	**0.694**	0.008	**0.473**	**0.536**	**0.542**	0.293	**0.353**	**0.631**	**0.723**	**0.452**
Rubble vs. Bedrock	0.321	0.145	0.093	0.176	0.270	0.221	0.142	0.193	0.176	0.044	0.304	0.445	0.934	**1.392**
Boulder vs. Bedrock	0.435	0.414	0.418	0.195	0.425	0.213	0.331	0.343	0.365	0.336	0.050	0.186	0.212	**0.940**

underlined: largest distance observed for each pair of categories; bold font: distances not significantly different form largest distance at a significance level of 0.05

Table 5-10: Bivariate Correlations between Predictor Variables

	Band 1	Band 2	Band 3	Band 4	PC1_ln	PC2_ln	PC3_ln	PC4_ln	PC1	PC2	PC3	PC4	Width
Band2	**0.887**												
Band 3	**0.732**	**0.932**											
Band 4	**0.768**	**0.836**	**0.762**										
PC1_ln	**0.909**	**0.988**	**0.929**	**0.879**									
PC2_ln	0.353	0.018	-0.260	0.244	0.062								
PC3_ln	-0.290	-0.060	0.047	0.324	-0.020	-0.140							
PC4_ln	0.172	-0.070	0.007	0.101	0.056	0.249	-0.040						
PC1	**0.872**	**0.972**	**0.916**	**0.928**	**0.989**	0.078	0.089	0.037					
PC2	0.239	0.327	0.379	-0.140	0.254	-0.450	-0.650	-0.240	0.172				
PC3	0.411	0.055	-0.250	0.074	0.065	**0.845**	-0.480	0.083	0.033	-0.120			
PC4	0.198	-0.110	-0.090	0.026	0.005	0.350	-0.180	**0.886**	-0.020	-0.160	0.235		
Width	-0.400	-0.160	-0.010	-0.240	-0.210	-0.540	0.186	-0.340	-0.190	0.161	-0.520	-0.430	
Gradient	-0.050	-0.070	-0.070	-0.110	-0.070	-0.010	-0.110	0.012	-0.090	0.062	0.023	0.023	0.070

bold font: correlation coefficients > 0.70

were significantly different from the distances obtained for all other predictor variables.

Variables discriminating between *Gravel* and *Rubble* include *PC2_ln*, *PC4*, *Width* and *Gradient*. Correlation among these variables ranges from -0.01 to -0.54. These values do not indicate redundancy in information. Therefore, the variables *PC2_ln*, *PC4* and *Width* were selected as predictor variables. The substrate classes *Gravel* and *Boulder* are differentiated by the variables *PC2_ln*, *PC2* and *Gradient*. In this case *PC2* is the only variables not yet selected as predictors. It was not found highly correlated with any other variable and therefore included in the analysis. The largest mean difference between the substrate categories *Gravel* and *Bedrock* was obtained in the variables *PC2_ln*, *PC4_ln*, *PC3* and *PC4*. A comparison of correlation coefficients showed that *PC3* and *PC2_ln* were highly correlated with r = 0.845. Therefore, *PC3* was not used in the classification of bottom substrate. Likewise, a correlation coefficient of 0.886 between *PC4_ln* and *PC4* indicated the capture of duplicate information and resulted in the exclusion of *PC4_ln* from all subsequent analysis. The distinction between the substrate types *Rubble* and *Boulder* involves all independent variables except *PC2_ln* and *PC2*. The variables *Band 2*, *Band 3*, *Band 4*, *PC1_ln*, and *PC1* were highly correlated with *Band 1*, with correlation coefficients ranging from 0.732 to 0.909. These variables were therefore excluded from further analysis.

The final predictor variables for the habitat parameter *Substrate Type* include *Band 1*, *PC2_ln*, *PC3_ln*, *PC2*, *PC4*, *Width* and *Gradient*.

5.3.2 Channel Pattern

Figure 5-3 contains the mean response and standard deviations of channel pattern categories. A tabular presentation of group means and variability is given in Appendix A. Throughout the original spectral bands, differences between category means were low with respect to the associated standard deviations. Likewise, high group variabilities were observed in all spectral bands. The mean response in *Band 1* ranged from 15.20 to 39.52, with standard deviations from 6.86 in the class *Flat* to 19.62 in the category *Rapid*. The lowest variability was observed in the category *Flat* with a coefficient of variation of 40.07 %, while the class *Steady* showed the highest variation (66.45 %). Spectral response increased progressively towards *Band 4*, where mean values varied from 29.43 for *Steady* to 48.96 in the class *Rapid*. Group variability in *Band 4* is similar to the coefficients of variation observed in *Band 1*, *Band 2* and *Band 3* and ranged from 44.44 % for *Rapid* to 65.07 % in the category *Steady*. The category *Rapid* is distinctly set apart from any other class in all spectral bands. The largest variability in all spectral bands was observed for the class *Steady* with coefficients of variation from 61.66 % in *Band 2* to 66.45 % in *Band 1*. Conversely, variability in the category *Flat* showed the lowest values throughout all spectral bands and varied from 34.64 % in *Band 2* to 47.47 % in *Band 4*.

Mean response observed in *PC1_ln* was similar to that in the original spectral bands. The category *Rapid* showed a mean value of 7.18, while average DN values from 5.74 to 6.35 were observed in the remaining categories. The group standard deviations varied from 0.65 to 1.21. Variability was lowest in the class *Flat* with 10.25 % and highest for *Steady* with 21.08 %. In *PC2_ln*, the categories *Run* and *Riffle* can be differentiated versus the classes *Steady* and *Flat*. The category *Rapid* showed the highest mean response with a value of 0.75. Group variability in *PC2_ln* is characterized by coefficients of variation ranging from 17.65 % in the class *Rapid* to 44.59 % in the category *Flat*. In *PC3_ln*, the channel pattern *Rapid* shows the lowest mean response with a value of 0.64. Mean response in the categories *Run*, *Riffle*, *Steady* and *Flat* ranged from 0.82 to 0.86. Similar group variabilities were observed in the classes *Run* and *Steady* with respective values of 26.73 % and 26.92 %. The class *Flat* showed the lowest coefficient of variation (19.80 %), while variability in

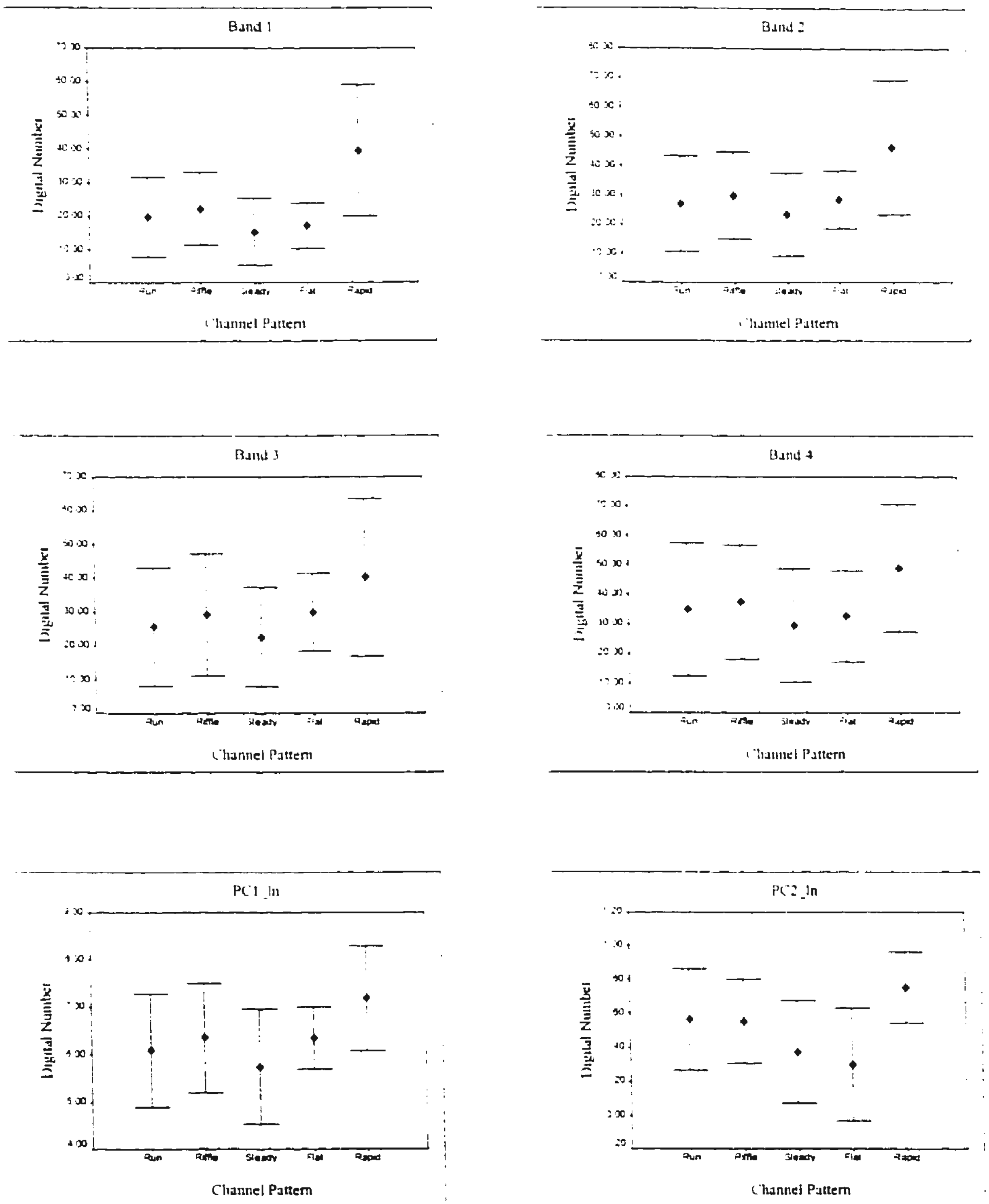

Figure 5-3: Group Means and Standard Deviations of Channel Pattern Categories

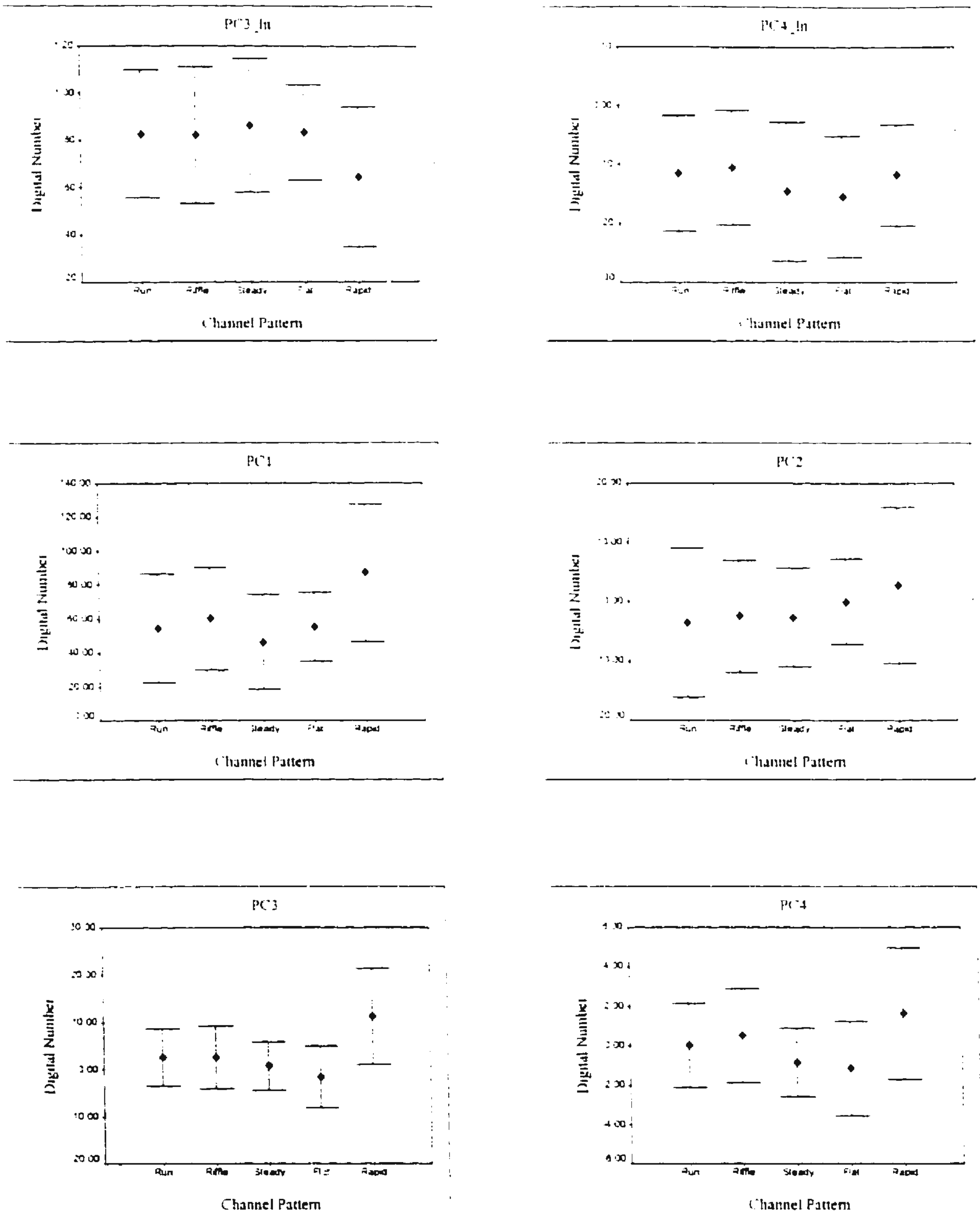

Figure 5-3: Group Means and Standard Deviations of Channel Pattern Categories (continued)

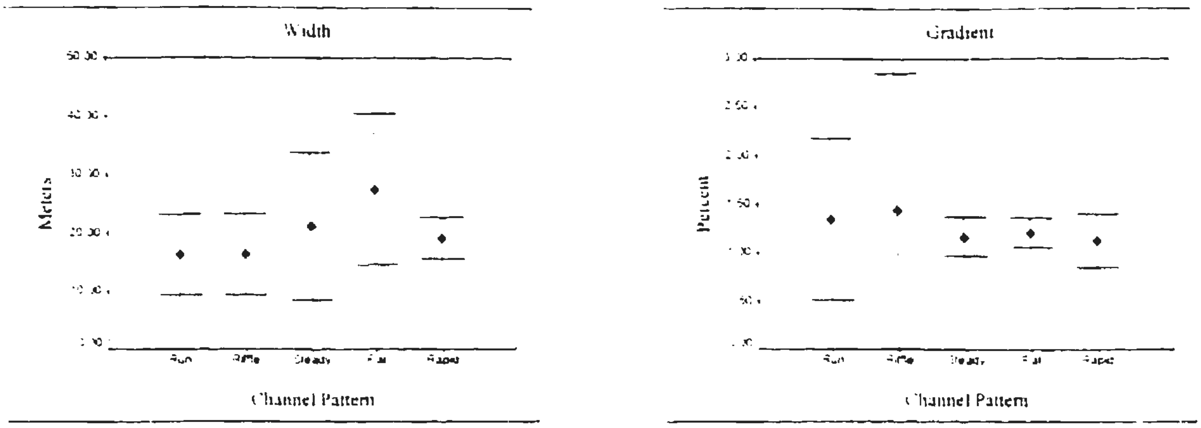

Figure 5-3: Group Means and Standard Deviations of Channel Pattern Categories (continued)

the categories of *Riffle* and *Rapid* was 30.00 % and 36.59 %, respectively. The channel patterns *Run*, *Riffle*, *Flat* and *Rapid* showed similar coefficients of variation in *PC4_ln* with values from 16.67 to 19.61 %. Group variability was highest in the class *Steady* with 23.53 %. Standard deviations in this variable varied from 0.09 to 0.10. Mean response values ranged from -0.15 in the class *Flat* to -0.11 in the category *Riffle*.

The observed mean values in *PC1*, *PC3* and *PC4* resembled the mean spectral response of channel patterns in the original spectral bands. The class *Rapid* showed mean response values of 87.09, 11.23 and 1.64, respectively. Mean response in the remaining classes ranged from 46.31 to 60.13 in *PC1*, from –1.64 to 2.60 in *PC3* and from –1.13 to 0.02 in *PC4*. In *PC2*, the class *Rapid* was less clearly separated from the remaining categories. Variability was approximately equal in the groups *Run* and *Rapid* with coefficients of variation of 24.52 and 22.84 %, respectively. The remaining classes were characterized by group variabilities from 13.09 % to 17.90 %. Mean response in *PC2* ranges from –3.50 in the class *Run* to 2.75 for *Rapid*.

In the variable *Width*, the categories *Run* and *Riffle* both showed a mean response of 16.34.

The respective coefficients of variation were 43.02 and 42.53 %, respectively. Mean response values in the remaining classes varied from 19.13 to 21.17. The highest variability was observed in the classes *Steady* and *Flat*, with respective coefficients of variation of 59.71 and 47.02 %. Group variability was lowest in the category *Rapid* with a value of 18.35 %.

The variable *Gradient* indicated separability of the categories *Run* and *Riffle* on one side, and *Steady*, *Flat* and *Rapid* on the other. Variability was highest in the classes *Riffle* and *Run* with coefficients of variation of 120.34 and 82.35 %, respectively. The lowest variability was obtained for the category *Flat* with 18.92 %.

Standard distances between channel pattern categories and corresponding rank correlation coefficients between predictor variables are presented in Table 5-11 and Table 5-12. The largest observed distance for a given pair of categories is underlined, while all distances which are not significantly different from this value are printed in bold letters. Overall, the largest distances occurred between the categories *Rapid* and *Run*, *Rapid* and *Riffle*, *Rapid* and *Steady*, as well as between *Rapid* and *Flat* in *Band 1*. The corresponding mean differences range from 1.299 to 1.809. The largest difference between the groups *Run* and *Flat* as well as between *Riffle* and *Flat* occurred in the variable *Width* with a standardized distance of 1.183. Differences between the remaining combinations of channel patterns range from 0 for *Run* versus *Riffle* to 0.671 for the mean distance between the classes *Steady* and *Flat*.

The variables *Band 1* and *PC3* show the most frequent occurrence of an overlap with the largest distance observed for each pair of categories. In both instances, the observed differences between were not significantly different form the respective largest distance seven out of ten times. This provides evidence that these variables contribute significantly to the differentiation between channel pattern categories. *Band 1* and *PC3* were consequently included as predictors in the analysis.

Table 5-11: Standardized Distances between Channel Pattern Category Means

	Band 1	Band 2	Band 3	Band 4	PC1_ln	PC2_ln	PC3_ln	PC4_ln	PC1	PC2	PC3	PC4	Width	Gradient
Run vs. Riffle	**0.180**	0.163	**0.199**	0.120	**0.227**	0.035	0.021	0.089	0.173	0.095	0.005	**0.205**	0	0.172
Run vs. Steady	**0.330**	0.220	0.170	**0.264**	*0.295*	**0.620**	0.128	**0.309**	**0.255**	0.074	**0.257**	**0.344**	**0.512**	**0.391**
Run vs. Flat	0.187	0.082	0.238	0.114	0.217	**0.870**	0.023	0.397	0.021	0.325	0.576	0.468	**1.183**	0.283
Run vs. Rapid	**1.479**	**1.137**	0.820	0.675	0.928	0.613	0.654	0.041	**1.008**	0.582	**1.173**	0.664	0.295	0.448
Riffle vs. Steady	**0.510**	**0.383**	**0.369**	**0.384**	**0.522**	**0.584**	0.148	**0.398**	**0.428**	0.021	**0.252**	**0.549**	**0.512**	**0.563**
Riffle vs. Flat	0.367	0.082	0.039	0.234	0.010	**0.835**	0.044	0.487	0.152	0.230	0.571	0.673	**1.183**	0.455
Riffle vs. Rapid	**1.299**	**0.974**	0.621	0.555	0.701	0.648	0.633	0.131	**0.835**	0.487	**1.178**	0.459	0.295	0.620
Steady vs. Flat	0.143	**0.302**	**0.408**	0.151	**0.512**	0.250	0.105	0.089	0.277	0.251	**0.319**	0.124	**0.671**	0.108
Steady vs. Rapid	**1.809**	**1.357**	0.990	0.939	1.223	1.330	0.782	0.268	**1.263**	0.508	**1.431**	1.008	0.217	0.058
Flat vs. Rapid	**1.666**	1.055	0.582	0.788	0.711	**1.484**	0.677	0.356	0.987	0.257	**1.749**	1.132	0.888	0.166

underlined: largest distance observed for each pair of categories; bold font: distances not significantly different form largest distance at a significance level of 0.05

Table 5-12: Bivariate Correlations between Predictor Variables

	Band 1	Band 2	Band 3	Band 4	PC1_ln	PC2_ln	PC3_ln	PC4_ln	PC1	PC2	PC3	PC4	Width
Band2	**0.915**												
Band 3	**0.777**	**0.932**											
Band 4	**0.785**	**0.858**	**0.788**										
PC1_ln	**0.923**	**0.989**	**0.940**	**0.895**									
PC2_ln	0.316	0.039	-0.251	0.214	0.054								
PC3_ln	-0.263	-0.047	0.043	0.333	-0.003	-0.142							
PC4_ln	0.070	-0.097	0.052	0.063	0.027	0.023	0.048						
PC1	**0.890**	**0.976**	**0.927**	**0.937**	**0.991**	0.069	0.097	0.010					
PC2	0.312	0.351	0.409	-0.089	0.289	-0.386	-0.686	-0.182	0.210				
PC3	0.403	0.105	-0.196	0.068	0.087	**0.834**	-0.515	-0.120	0.055	-0.007			
PC4	0.176	-0.067	0.015	0.033	0.040	0.186	-0.153	**0.875**	0.013	-0.061	0.084		
Width	-0.197	-0.031	0.059	-0.109	-0.068	-0.376	0.083	-0.265	-0.061	0.153	-0.321	-0.302	
Gradient	0.076	0.163	0.188	0.052	0.133	-0.182	-0.045	-0.193	0.126	0.207	-0.146	-0.194	0.133

bold font: correlation coefficients > 0.70

Conversely, all distances in the variables *PC3_ln* and *PC2* were significantly smaller than the corresponding largest distance observed. This indicates that neither of these variables contributes significantly to the discrimination between categories. Therefore, *PC3_ln* and *PC2* were excluded from further analysis. The variables *Band 2*, *Band 3*, *Band 4*, *PC1_ln* and *PC1* were highly correlated with *Band 1*. Bivariate correlation coefficients varied from 0.777 to 0.989. These variables were therefore considered redundant and removed from the analysis. Likewise, *PC2_ln* was highly correlated with *PC3* (r = 0.834) and not selected as a predictor variable. The variables *PC4* and *PC4_ln* were correlated with a bivariate correlation coefficient of 0.875. Standardized distances between channel pattern categories not significantly different from the largest distance were observed in three instances in *PC4*, and in two instances in *PC4_ln*. The former was therefore selected as predictor, while the latter was excluded from further analysis. *Width* and *Gradient* were uncorrelated with any other variable and therefore included as predictor variables in the analysis.

Predictor variables used in the classification of the habitat parameter *Channel Pattern* include *Band 1*, *PC3*, *PC4*, *Width* and *Gradient*.

5.3.3 Land Cover

Values of mean response, standard deviation and variability of each land cover category are presented in Figure 5-4 and in Appendix A. A progressively increasing mean response from *Band 1* to *Band 4* was observed for all categories. Spectral response of land cover classes was similar in the variables *Band 1*, *Band 2* and *Band 3*. The category *Water* showed the lowest mean response and highest variability in all spectral bands. Mean values ranged from 9.80 in *Band 1* to 17.65 in *Band 3*, while coefficients of variation varied from 56.96 % in *Band 2* to 63.80 % in *Band 3*. The category

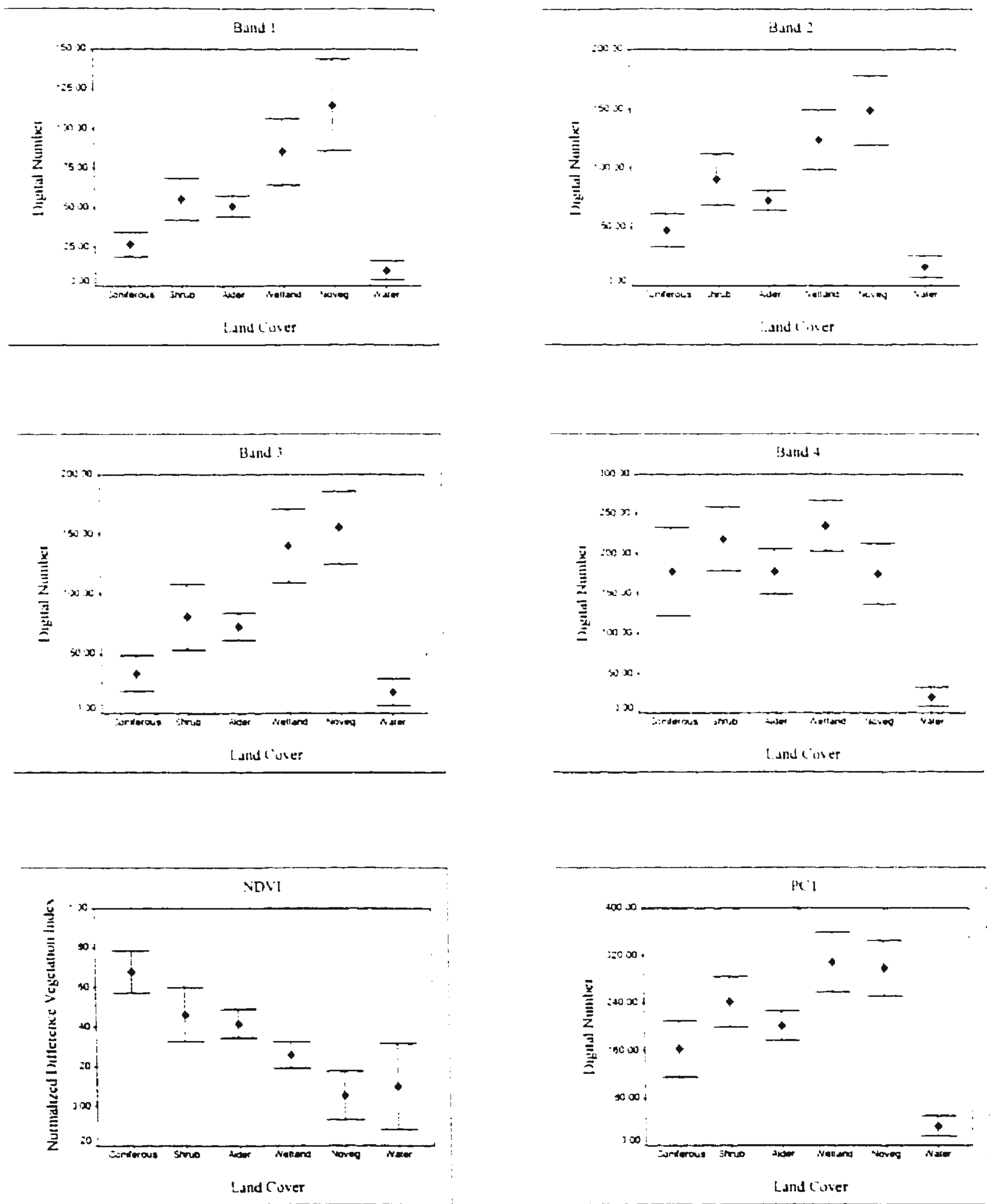

Figure 5-4: Group Means and Standard Deviations of Land Cover Categories

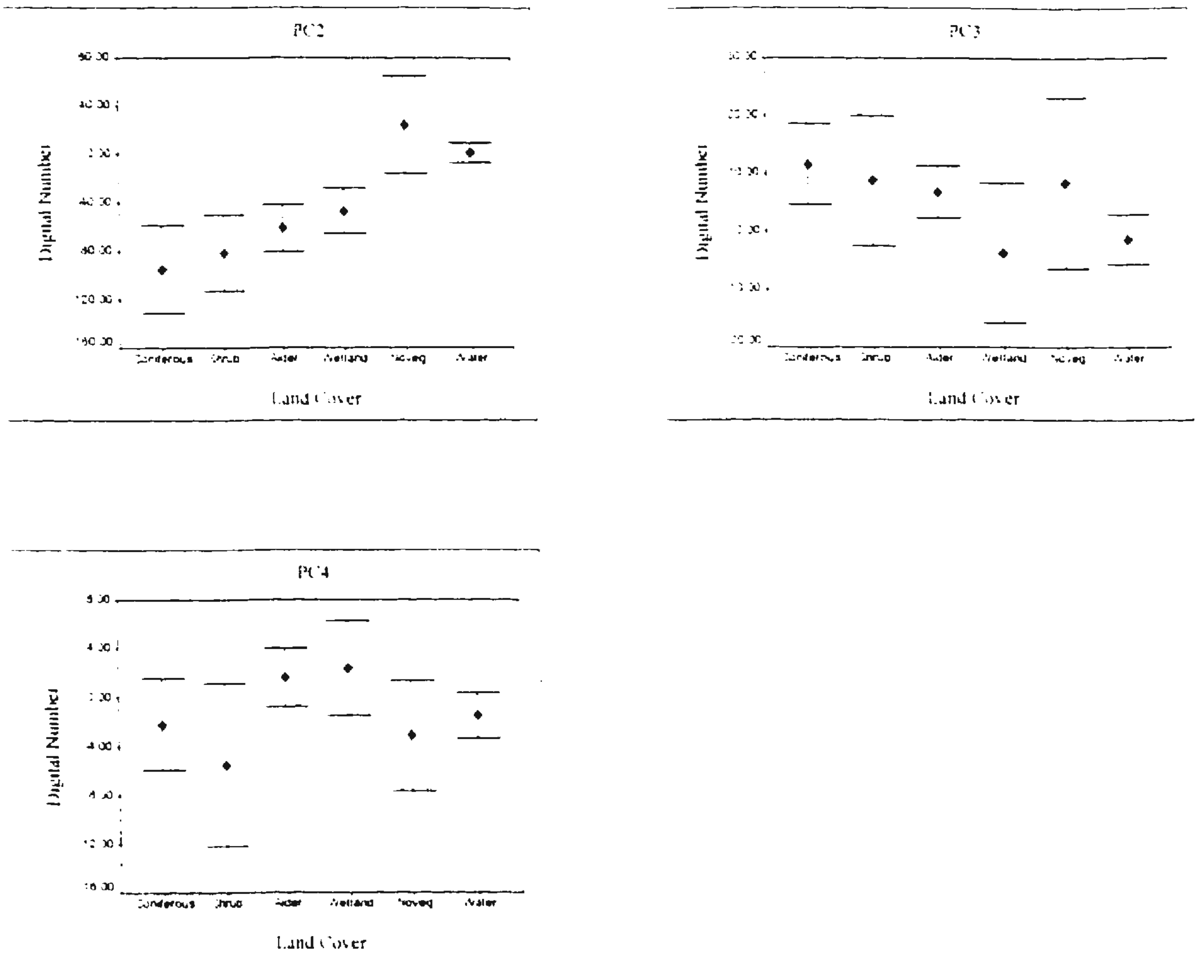

Figure 5-4: Group Means and Standard Deviations of Land Cover Categories (continued)

Coniferous showed the second lowest spectral response with average DN values from 26.58 in *Band 1* to 33.48 in *Band 3*. Variability of this class is similar in *Band 1* (29.16 %), *Band 2* (29.75 %) and *Band 4* (31.47 %), and is highest in *Band 3* with a coefficient of variation of 45.07 %. The classes *Shrub* and *Alder* showed mean response values of 55.16 and 50.88 in *Band 1*. In *Band 2*, these values increase to 90.70 and 72.40, respectively. The land cover category *Shrub* showed a slightly decreased mean of 80.71 in *Band 3*, while response in the class *Alder* remained nearly constant with a value of 72.31. Variability of the category *Shrub* in *Band 1* (24.15 %), *Band 2* (23.93 %) and *Band 3* (33.50 %) was approximately twice as high as the corresponding coefficients of variation of the class *Alder*. The respective standard deviations ranged from 13.32 and 6.68 in *Band 1* to 27.04 and 11.24 in *Band 3*. Group variability was lowest in all spectral Bands for the category *Alder*. Mean values and standard deviations in the class *Wetland* varied from 85.36 and 21.14 *in Band 1* to 140.23 and 30.96 in *Band 3*. Likewise, mean response in the category *Noveg* ranged from 114.90 to 155.24, with standard deviations from 28.90 to 30.50. Coefficients of variation of the class *Wetland* varied from 13.54 % in *Band 4* to 24.77 % in *Band 1*. Variability of the class *Noveg* ranged from 19.65 % in *Band 3* to 25.15 % in *Band 1*. Overall, the separability of individual land cover classes was less distinct in *Band 4*. Observed standard deviations ranged from 11.26 for *Water* to 55.82 in the class *Coniferous*. The category Water showed the lowest mean response with a value of 17.65. Mean values in the remaining categories ranged from 174.42 in the class *Noveg* to 218.26 for *Shrub*.

The normalized difference vegetation index (*NDVI*) differentiated between vegetated and not vegetated categories. The former were characterized by mean values from 0.22 in the class *Wetland* to 0.68 in the category *Coniferous*, while the latter showed mean response values of 0.05 and 0.09 for *Noveg* and *Water*, respectively. Variability was highest for *Water* and *Noveg*, with respective coefficients of variation of 66.67 % and 41.38 %. The remaining categories showed group variabilities from 10.61 to 20.00 %.

The first principal component, *PC1*, resembled the pattern of spectral response observed in *Band 1*, *Band 2* and *Band 3*. Mean response was lowest (32.64) and variability highest (53.62 %) for the category *Water*. The remaining categories were characterized by mean values ranged from 162.28 for *Coniferous*, to 308.93 for *Wetland*. Group variability was similar in the classes *Shrub* (17.81 %) *Wetland* (16.29 %) and *Noveg* (15.71 %). The largest coefficient of variation was observed in the category *Coniferous*, while *Alder* showed the smallest variability with 12.40 %. The variable *PC2* indicated separability between vegetated classes on one side, and the categories *Noveg* and *Water* in the other. Variability was lowest in the class *Water* (4.99 %) and highest for *Coniferous* (56.67 %). The highest spectral response was observed in the category *Noveg* with a mean value of 24.42, whereas the class *Water* showed an average DN value of 1.33. In the remaining classes, mean values varied from –95.21 to –47.13. In *PC3*, variability was lowest in the groups *Alder* (7.80 %) and *Water* (8.73 %) with mean values of 6.77 and –1.47, respectively. The highest variability in *PC3* occurred in the categories *Wetland* and *Noveg* with respective coefficients of variation of 25.56 and 24.95 %. In *PC4*, mean response varied from –5.52 in the category *Shrub* to 2.39 for *Wetland*. Variability was lowest in the groups *Water* and *Alder* with respective values of 7.21 and 8.12 %. The largest coefficient of variation was observed for *Shrub* with 30.64 %. Group variability of the remaining classes of *Coniferous*, *Wetland* and *Noveg* ranged from 13.05 to 18.56 %.

Standardized distances between land cover category means are presented in Table 5-13. The largest distance observed for a given pair of categories is underlined. Distances that are not significantly different from this value are printed in bold letters. Table 5-14 shows the bivariate rank correlation coefficients between predictor variables. Correlation coefficients exceeding a value of 0.70 are highlighted.

The variables *Band 3*, *Band 4*, *NDVI*, *PC2* and *PC4* showed mean differences between land cover categories that were significantly different from all other distance observed. This applied to the discrimination of *Coniferous* versus *Wetland* in *Band 3*, *Shrub* and *Alder* versus *Water* in Band 4, *Coniferous* versus *Water* in the *NDVI*, *Shrub* versus *Noveg* in *PC2* and *Shrub* versus *Alder* and *Wetland* in *PC4*. These variables were therefore considered important in the differentiation between land cover categories and included as predictors in all subsequent analysis. The standardized distance between the categories *Coniferous* and *Shrub* is largest in the variable *Band 2*. However, the observed

Table 5-13: Standardized Distances between Land Cover Category Means

	Band 1	Band 2	Band 3	Band 4	NDVI	PC1	PC2	PC3	PC4
Coniferous vs. Shrub	**0.739**	**0.884**	**0.861**	0.488	**0.801**	**0.764**	0.256	0.259	**0.678**
Coniferous vs. Alder	0.628	0.511	0.708	0.003	**0.972**	0.376	0.635	0.442	**0.763**
Coniferous vs. Wetland	1.520	1.562	**1.946**	0.690	1.557	1.401	0.902	1.391	0.916
Coniferous vs. Noveg	**2.284**	2.064	**2.220**	0.035	**2.318**	1.303	**2.244**	0.301	0.195
Coniferous vs. Water	0.434	0.629	0.288	1.879	**2.152**	1.239	1.811	1.185	0.137
Shrub vs. Alder	0.111	0.372	0.153	0.491	0.171	0.389	0.379	0.183	**1.442**
Shrub vs. Wetland	0.781	0.678	1.085	0.202	0.756	0.637	0.646	1.132	**1.595**
Shrub vs. Noveg	1.545	1.180	1.358	0.523	1.517	0.539	**1.988**	0.042	0.436
Shrub vs. Water	1.730	1.513	1.149	**2.368**	1.350	2.003	1.555	0.926	0.816
Alder vs. Wetland	0.892	**1.051**	**1.238**	0.693	0.585	1.026	0.267	**0.949**	0.153
Alder vs. Noveg	**1.655**	**1.553**	**1.511**	0.032	1.346	0.927	**1.609**	0.141	0.958
Alder vs. Water	1.062	1.140	0.996	**1.876**	1.179	1.614	1.176	0.743	0.626
Wetland vs. Noveg	0.764	0.502	0.273	0.725	0.761	0.098	**1.342**	**1.090**	**1.111**
Wetland vs. Water	1.954	2.191	2.234	**2.569**	0.594	**2.640**	0.909	0.206	0.779
Noveg vs. Water	**2.718**	**2.693**	2.508	1.844	0.166	2.542	0.433	0.884	0.332

underlined: largest distance observed for each pair of categories, bold font: distances not significantly different form largest distance at a significance level of 0.05

Table 5-14: Bivariate Correlations between Predictor Variables

	Band 1	Band 2	Band 3	Band 4	NDVI	PC1	PC2	PC3
Band 2	**0.987**							
Band 3	**0.965**	**0.973**						
Band 4	0.604	0.656	0.602					
NDVI	-0.219	-0.205	-0.317	0.383				
PC1	**0.936**	**0.965**	**0.936**	**0.798**	-0.087			
PC2	0.087	0.045	0.123	-0.646	**-0.878**	-0.118		
PC3	0.177	0.149	-0.033	0.264	0.551	0.138	-0.401	
PC4	-0.017	-0.087	0.045	0.0751	0.006	0.001	-0.087	-0.34

bold font: correlation coefficients > 0.70

value of 0.884 was not significantly different form the corresponding distance in *Band 3*. Moreover, *Band 2* and *Band 3* were highly correlated with r = 0.973. *Band 2* was therefore excluded from further analysis. Likewise, *Band 3* was highly correlated with *PC1* at a correlation coefficient of 0.936. *PC1* showed the largest distance between the land cover classes *Wetland* and *Water* with a mean difference of 2.569. This value was not significantly different from the corresponding distance in *Band 4*. The bivariate correlation of 0.602 between *Band 3* and *Band 4* did not indicate the capture of redundant information. Consequently, *Band 4* was selected as predictor variable, while *PC1* was excluded from the analysis. In the variable *PC3*, significant differences were observed between the categories *Alder* and *Wetland* and for *Wetland* versus *Noveg*. Since *PC3* was not highly correlated with any other independent variable it was selected as a predictor variable. *Band 1* showed the largest distance between the categories *Noveg* and *Water*. The observed value of 2.718 was significantly different from the corresponding distances in all other variables except *Band 2*. Therefore, *Band 1* was included as predictor variable for the type of land cover. The final predictor variables for the habitat parameter *Land Cover* include *Band 1*, *Band 3*, *Band 4*, *NDVI*, *PC2*, *PC3* and *PC4*. Among these variables, high correlation coefficients were observed between *NDVI* and *PC3*, as well as between *Band 1* and *Band 3*. The observed correlation coefficients were -0.878 and 0.965, respectively. It was required to adjust for the capture of duplicate information in these variables before the classification. Details of this adjustment are explained in Section 5.4.3.

5.4 Classification of Habitat Parameters

Classification of the habitat parameters *Substrate Type*, *Channel Pattern* and *Land Cover* was carried out using the exhaustive partitioning approach to decision tree analysis. At every node, the predictor variable which showed the most significant relationship with each habitat parameter was selected to split the data. In order to minimize the risk of detecting spurious relationships, all significance levels were adjusted for multiple comparisons using the Bonferroni procedure. The minimum level of significance was set at 0.05. The partitioning process was stopped if the criterion of minimum significance was not satisfied. Likewise, no further division occurred if the number of observations in a node dropped below a pre-defined Threshold. This value was related to the smallest sampling unit for each habitat parameter and was set to 150 for *Substrate Type* and *Channel Pattern*. The corresponding threshold value for the habitat parameter *Land Cover* was set to 60.

5.4.1 Substrate Type

The classification tree for the habitat parameter *Substrate Type* is shown in Figure 5-5. The number of observations in each node is shown in brackets. Terminal nodes are indicated by the substrate category which occurred most frequently in that node. A total of 11 terminal nodes were obtained. These nodes were interpreted as classification rules and used to classify the remainder of the data. Classification rules for all habitat parameters are presented in Appendix B.

Gradient was the first variable selected to split the data set. The resulting sub-sets were further divided by *Width* and *PC4*. The last level of the tree is formed by the variables *Band 1* and *PC2*. The substrate category *Gravel* occurred exclusively at intermediate valley gradients ranging from 0.29 to 0.42 %. Conversely, the class *Bedrock* was found at valley gradients of less than 0.29 % and more than 0.43 %. In both cases, this substrate type occurred predominantly at wider stream sections, while the categories *Rubble* and *Boulder* were found at stream widths of less than 20m. The substrate classes *Rubbie*, *Boulder* and *Bedrock* were further differentiated by the optical variables

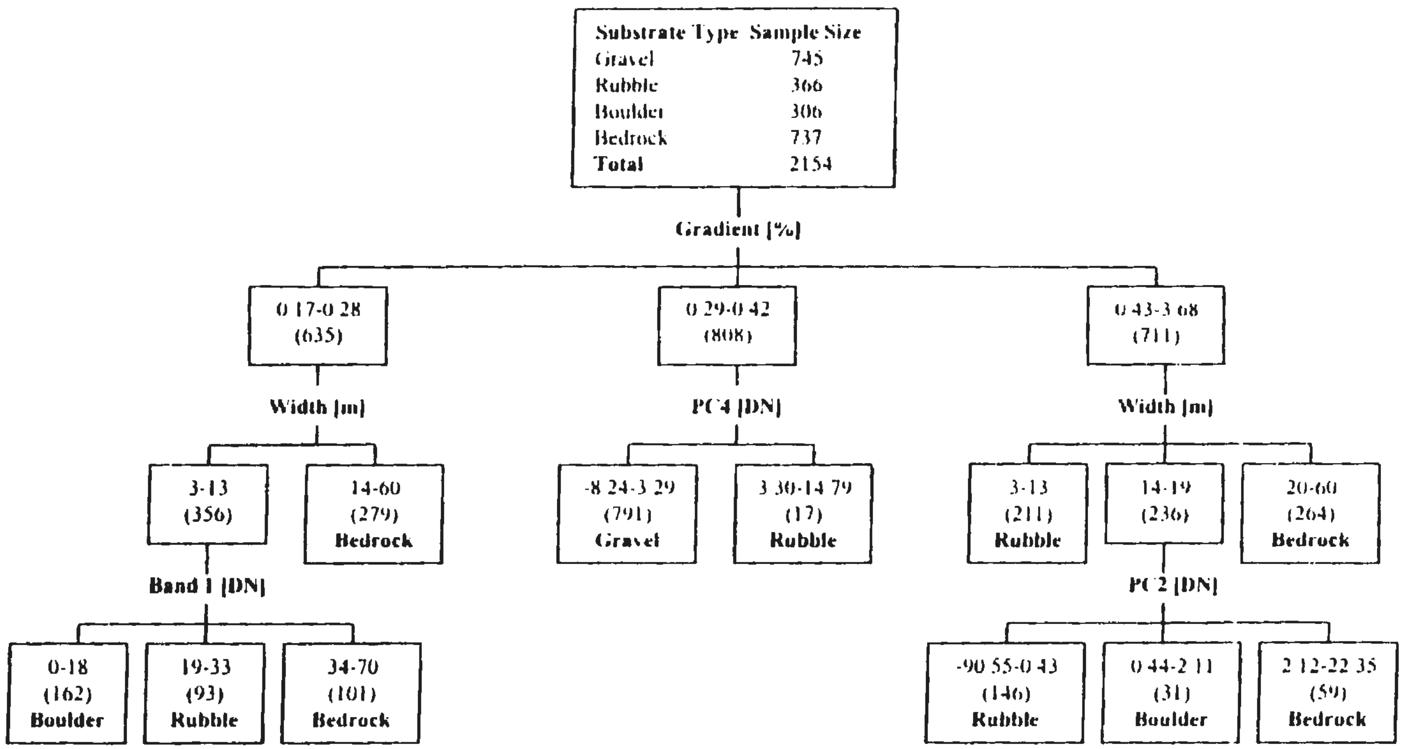

Figure 5-5: Decision Tree for Substrate Classification

Band1 and *PC2_ln* at stream widths of less than 14m and from 14 to 19m, respectively.

Classification accuracy was assessed by using a test sample which was not used in the derivation of the decision tree. A cross-tabulation of predicted and actual class membership is presented in Table 5-15. A total of 989 out of 991 test pixels were classified, leaving 0.2 % of the verification data unclassified. The estimated overall classification accuracy was 66.80 %. The associated kappa index of agreement suggests that this value was 55 % better than would be expected for pure chance assignment.

Table 5-15: Error Matrix for Substrate Classification - Original Categories

	REFERENCE DATA						
CLASSIFIED DATA	**Gravel**	**Rubble**	**Boulder**	**Bedrock**	**Total**	**CE [%]**	**UA [%]**
Gravel	**248**	25	50	0	323	23.22	76.78
Rubble	37	**117**	46	45	245	52.24	47.76
Boulder	7	23	**51**	21	102	50.00	50.00
Bedrock	10	52	11	**246**	319	22.88	77.12
Total	302	217	158	312	**989**	unclassified: 0.20 %	
OE [%]	17.88	46.08	67.72	21.15	**OA = 66.80 %; κ = 0.546**		
PA [%]	82.12	53.92	32.28	78.85	CI_{95} = [63.87 % ; 69.73 %]		

OE = Omission Error; CE = Commission Error; PA = Producer's Accuracy; UA = User's Accuracy;
OA = Overall Accuracy; CI_{95} = 95 % Confidence Interval; κ = Kappa Index of Agreement

The highest user's accuracies were observed in the categories *Gravel* and *Bedrock* with 76.78 % and 77.12 %, respectively. Considerably lower values of 47.76 % and 50.00 % were observed in the respective categories *Rubble* and *Boulder*. Likewise, the substrate classes *Gravel* and *Bedrock* showed the highest producer's accuracy with respective values of 82.12 % and 78.85 %. The corresponding value for *Rubble* was 53.92 %, while the lowest producer's accuracy was observed for the substrate type *Boulder* with 32.28 %.

Confusion between the substrate classes *Gravel* and *Bedrock* was minimal with only 10 cases misclassified as *Bedrock*. On the other hand, confusion was particularly strong between the categories *Rubble* and *Boulder*. The poor definition of the category *Boulder* and pronounced confusion with the class *Rubble* suggested that the overall accuracy could be improved by combining the substrate categories Rubble and Boulder. Collapsing of these classes is feasible since they share the same ecological significance as potential locations of spawning beds for Atlantic salmon (Dubois and Gosselin, 1989). The result of this operation is given in Table 5-16. The overall accuracy increased by 6.96 % to 73.76 %. This increase in overall classification accuracy is significant at the 95 % level of confidence. Accordingly, the agreement index κ shows that this accuracy is 61 % higher than would be expected under conditions of random assignment. User's and producer's accuracy of the combined category increased to 68.30 % and 63.20 %, respectively.

Table 5-16: Error Matrix for Substrate Classification - Collapsed Categories

	REFERENCE DATA					
CLASSIFIED DATA	**Gravel**	**Rubble/Boulder**	**Bedrock**	**Total**	**CE [%]**	**UA [%]**
Gravel	**248**	75	0	323	23.22	76.78
Rubble/Boulder	44	**237**	66	347	31.70	68.30
Bedrock	10	63	**246**	319	22.88	77.12
Total	302	375	312	**989**	unclassified: 0.20 %	
OE [%]	17.88	36.80	21.15	**OA = 73.76 %; κ = 0.608**		
PA [%]	82.12	63.20	78.85	CI_{95} = [71.02 % ; 76.50 %]		

OE = Omission Error; CE = Commission Error; PA = Producer's Accuracy; UA = User's Accuracy;
OA = Overall Accuracy; CI_{95} = 95 % Confidence Interval; κ = Kappa Index of Agreement

5.4.2 Channel Pattern

Figure 5-6 shows the decision tree for the channel pattern classification. A total of 18 terminal nodes was obtained and used to classify the remainder of the data. The corresponding classification rules are presented in Appendix B. The predictor *Width* showed the most significant relationship with the habitat parameter *Channel Pattern* and was therefore selected as the first variable to partition the data set. The resulting nodes were further divided by the variable *Gradient*. The last level of the tree was formed by *Band 1* and *PC3*. The channel pattern *Run* occurred in all stream widths from 3 to 60 m. The category *Riffle*, on the other hand, was found at stream widths of less than 27 m. Both categories were further differentiated by the variable *Gradient*. At a stream width of less than 16 m, the class *Run* occurred at a valley gradient of less than 0.45 %, while the category *Riffle* was observed at gradients greater than 0.45 %. The channel patterns *Steady* and *Flat* were largely discriminated in the image variables *Band 1* and *PC3*. At stream widths from 12 to 16 m, these categories occurred in *Band 1* at DN values from 0 to 26, while the classes *Riffle* and *Rapid* were observed at values of greater than 26.

An assessment of the accuracy of this classification is presented in Table 5-17. Of the 1047 observations in the test data set only one pixel remained unclassified. The overall classification accuracy was extremely low with only 38.11 % of all cases correctly classified. The observed overall accuracy is 18 % higher than would be obtained for a random classification. User's accuracy values for individual channel patterns ranged from 0 % in the class *Rapid* to 63.44 % in the category *Riffle*. Producer's accuracies varied from 0 % for *Rapid* to 66.54 % in the class *Run*. Confusion is high between the categories *Run*, *Riffle* and *Steady*, where 312 out of 566 cases (55.12 %) classified as *Run* actually belonged to the classes *Riffle* or *Steady*. All 53 observations of the channel pattern *Rapid* were

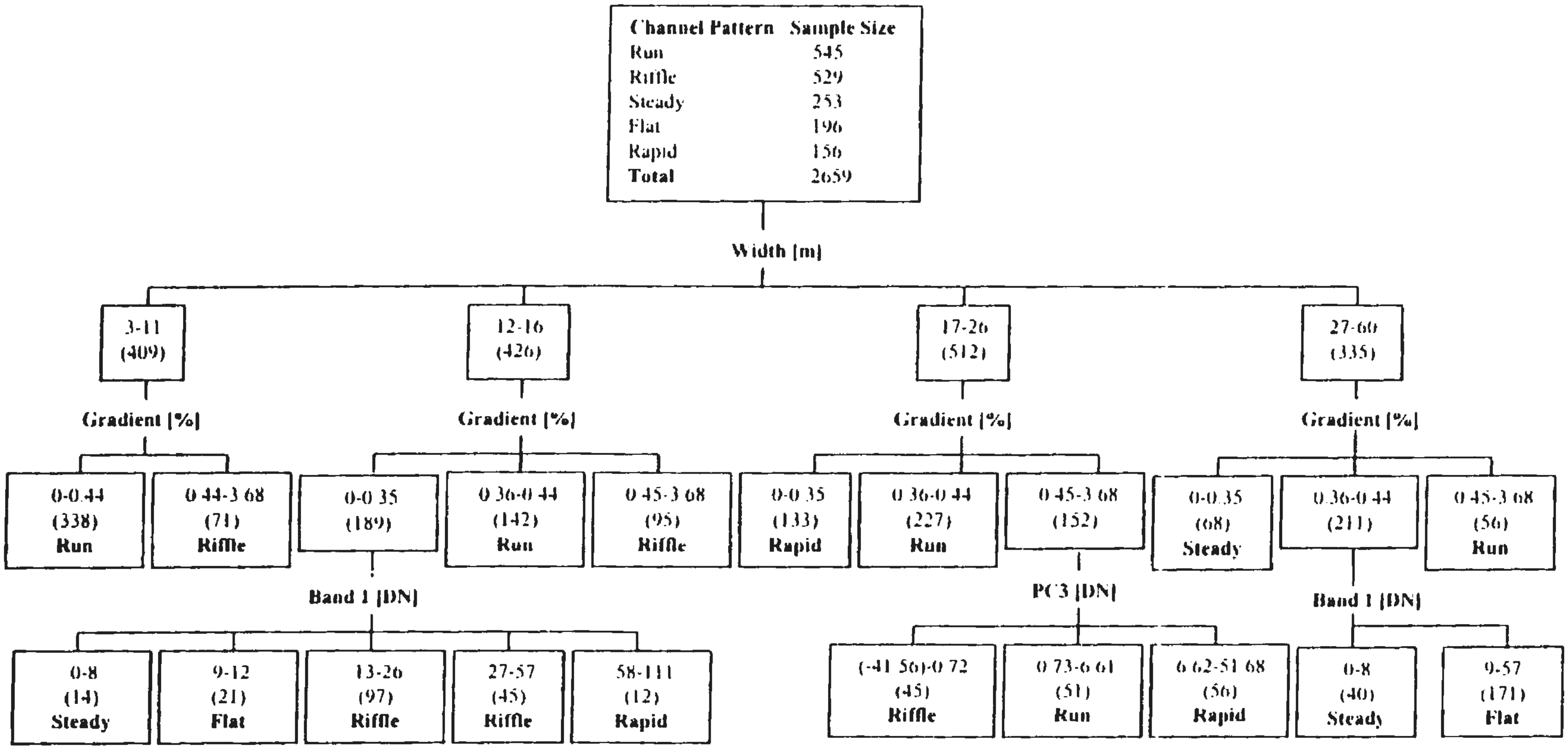

Figure 5-6: Decision Tree for Channel Pattern Classification

falsely classified as *Run*.

In order to improve the classification result, the original channel patterns were combined according to their respective importance as spawning habitat for Atlantic salmon. An assessment of classification accuracy using the collapsed categories is presented in Table 5-18.

Table 5-17: Error Matrix for Channel Pattern Classification – Original Categories

	REFERENCE DATA							
CLASSIFIED DATA	**Run**	**Riffle**	**Steady**	**Flat**	**Rapid**	**Total**	**CE [%]**	**UA [%]**
Run	**173**	165	147	28	53	566	69.43	30.57
Riffle	62	**144**	1	20	0	227	36.56	63.44
Steady	16	12	**33**	4	0	65	49.23	50.77
Flat	4	53	29	**49**	0	135	63.70	36.30
Rapid	5	48	0	0	**0**	53	100.00	0.00
Total	260	422	210	101	53	**1046**	unclassified: 0.10 %	
OE [%]	33.46	65.88	84.29	51.49	100.00	**OA = 38.11 %; κ = 0.176**		
PA [%]	66.54	34.12	15.71	48.51	0.00	CI_{95} = [35.17 %; 41.05 %]		

OE = Omission Error; CE = Commission Error; PA = Producer's Accuracy; UA = User's Accuracy;
OA = Overall Accuracy; CI_{95} = 95 % Confidence Interval; κ = Kappa Index of Agreement

The original five categories were collapsed into the new classes of *Riffle/Rapid* and *Run/Steady/Flat*, and the overall classification accuracy increased from 38.11 to 64.47 %. This increase was statistically significant at a significance level of 0.05. The corresponding value of kappa increased to 26 %. The observed user's accuracy values were 68.57 % in the class *Riffle/Rapid* and 63.05 % for *Run/Steady/Flat*, with producer's accuracy levels of 40.42 % and 84.59 %, respectively.

Table 5-18: Error Matrix for Channel Pattern Classification - Collapsed Categories

CLASSIFIED DATA	REFERENCE DATA Riffle/Rapid	Run/Steady/Flat	Total	CE [%]	UA [%]
Riffle/Rapid	**192**	88	280	31.43	68.57
Run/Steady/Flat	283	**483**	766	36.95	63.05
Total	475	571	**1046**	unclassified:0.10 %	
OE [%]	59.58	15.41	**OA = 64.47 %; κ = 0.259**		
PA [%]	40.42	84.59	CI_{95} = [61.57 %; 67.37 %]		

OE = Omission Error; CE = Commission Error; PA = Producer's Accuracy; UA = User's Accuracy;
OA = Overall Accuracy; CI_{95} = 95 % Confidence Interval; κ = Kappa Index of Agreement

5.4.3 Land Cover

The decision tree for the land cover classification is presented in Figure 5-7. High correlation coefficients between *Band 1* and *Band 3* (r = 0.965) as well as between *NDVI* and *PC2* (r = -0.878) indicated the capture of duplicate information. This increased the probability of detecting spurious relationships. Consequently, the Bonferroni procedure was also used to adjust the significance levels associated with the relationships between these variables and *Land Cover*. The Bonferroni adjustment was set to a value of 2 to equal the number of highly correlated variables.

The first variable selected to partition the data set was *Band 1*. Subsequent nodes were further split by *PC2*, *PC4*, *NDVI*, *Band 3* and *Band 4*. The remaining sub-sets were divided by the predictors *Band 1*, *PC2*, *PC4*, and *NDVI*. At DN values of less than 30 in *Band 1*, the land cover classes *Water* and *Coniferous* were differentiated by *PC2*. In addition to the category *Coniferous*, the classes *Shrub* and *Alder* occurred at DN values from 31 to 47 in *Band 1*. *Shrub* and *Alder* dominate the land cover types observed at DN values in *Band 1* from 48 to 71. They were further discriminated by *PC2* and

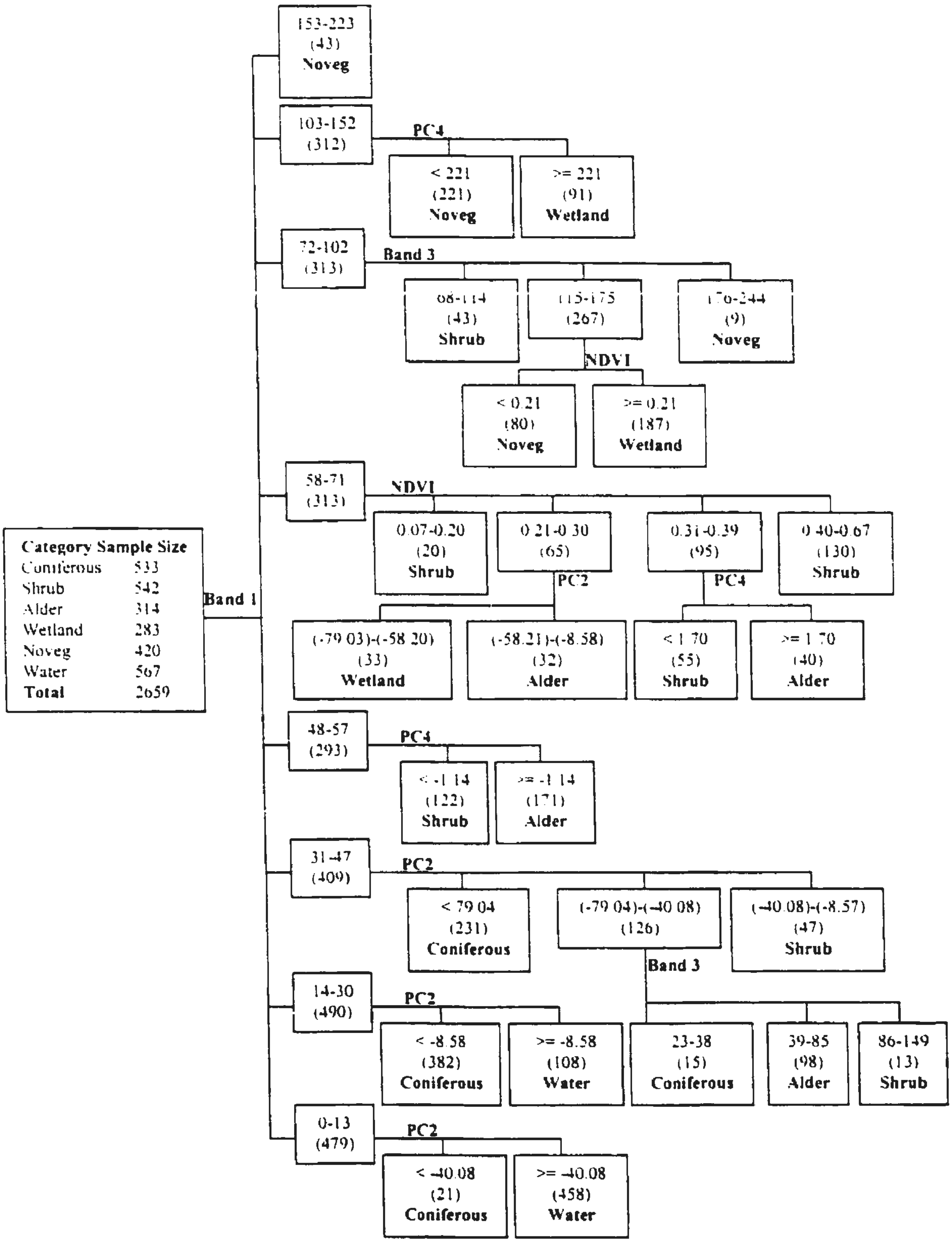

Figure 5-7: Decision Tree for Land Cover Classification

PC4. The land cover classes *Wetland* and *Noveg* occurred predominantly in *Band 1* at DN values of greater than 71.

Classification rules derived from the decision tree were applied to a test data set in order to evaluate the performance of the classification. Table 5-19 contains a cross-tabulation of predicted and actual land cover classes. Only one observation out of 2611 remained unclassified. The overall classification accuracy was 84.91 %. The κ value indicates an accuracy 81 % higher than expected for class assignment by chance. The lowest user's accuracies were observed in the categories *Shrub* and *Alder* with 74.25 % and 75.17 %, respectively. User's accuracies in the remaining classes ranged from 86.67 for *Wetland* to 94.50 for *Water*. Producer's accuracy levels were lowest in the categories *Shrub* and *Wetland* with respective values of 71.24 % and 69.55 %. The corresponding value for *Alder* was 77.51 %. Producer's accuracies in the remaining categories varied from 89.02 % in the class *Coniferous* to 100 % for *Noveg*. The categories *Noveg* and *Water* were least affected by confusion

Table 5-19: Error Matrix for Land Cover Classification

REFERENCE DATA

CLASSIFIED DATA	**Coniferous**	**Shrub**	**Alder**	**Wetland**	**Noveg**	**Water**	**Total**	**CE [%]**	**UA [%]**
Coniferous	**608**	47	35	0	0	1	691	12.01	87.99
Shrub	47	**369**	59	22	0	0	497	25.72	74.25
Alder	10	72	**324**	25	0	0	431	24.83	75.17
Wetland	0	26	0	**169**	0	0	195	13.33	86.67
Noveg	0	4	0	27	**438**	0	469	6.61	93.39
Water	18	0	0	0	0	**309**	327	5.50	94.50
Total	683	518	418	243	438	310	**2610**	unclassified: 0.04 %	
OE [%]	10.98	28.76	22.49	30.45	0.00	0.32	**OA = 84.91 %; κ = 0.815**		
PA [%]	89.02	71.24	77.51	69.55	100.00	99.68	CI_{95} = [83.54; 86.28]		

OE = Omission Error; CE = Commission Error; PA = Producer's Accuracy; UA = User's Accuracy; OA = Overall Accuracy; CI_{95} = 95 % Confidence Interval; κ = Kappa Index of Agreement between land cover types. On the other hand, confusion was most prominent between the categories *Shrub*, *Alder* and *Wetland*. The difference between user's and producer's accuracies were small in most land cover classes, with the exception of the category *Wetland*. In this case, the user's accuracy was by 17.12 % higher than the producer's accuracy. This difference was caused by test pixels which were falsely labelled as *Shrub*, *Alder* and *Noveg*.

5.5 Suitability Mapping

In order to demonstrate how the results of this study may be applied to the inventory and analysis of freshwater resources, the habitat parameters Substrate Type and Channel Pattern were subsequently combined to designate spawning habitat priority in the Come By Chance River according to Equation 5-1. The result of this analysis is presented in Table 5-20.

$$P = B * C \tag{5-1}$$

where

P = spawning habitat priority
B = priority weight of substrate type
C = priority weight of channel pattern

Table 5-20: Areal Extent of Spawning Habitat Classes

	Area [m^2]	Proportion [%]
Habitat I	24944	7.79
Habitat II	209964	65.56
Habitat III	85368	26.65
Total	320276	100

Approximately two thirds of the Come By Chance River is characterized by spawning habitat of Class II. Areas most suitable for spawning occupy only about 8 % of the river course, while one fourth of the total area is not suitable for spawning due to predominant bedrock substrate. Combining Habitat I and Habitat II yields a total area of 234,908 m^2, or 73.35 %, suitable for spawning. The distribution of salmon habitat over the whole length of the river is easily observed in Figure 5-8. Figure 5-9 shows the distribution of land cover categories. The majority of spawning habitats of type I are located in the upper reaches of the river, while areas unsuitable for spawning are concentrated in the middle section. The lower part of the river is dominated by Habitat II. For display purposes, a scale of 1:100,000 was selected. Given an original spatial resolution of 2 m, these results can be used at scales of up to 1:5000.

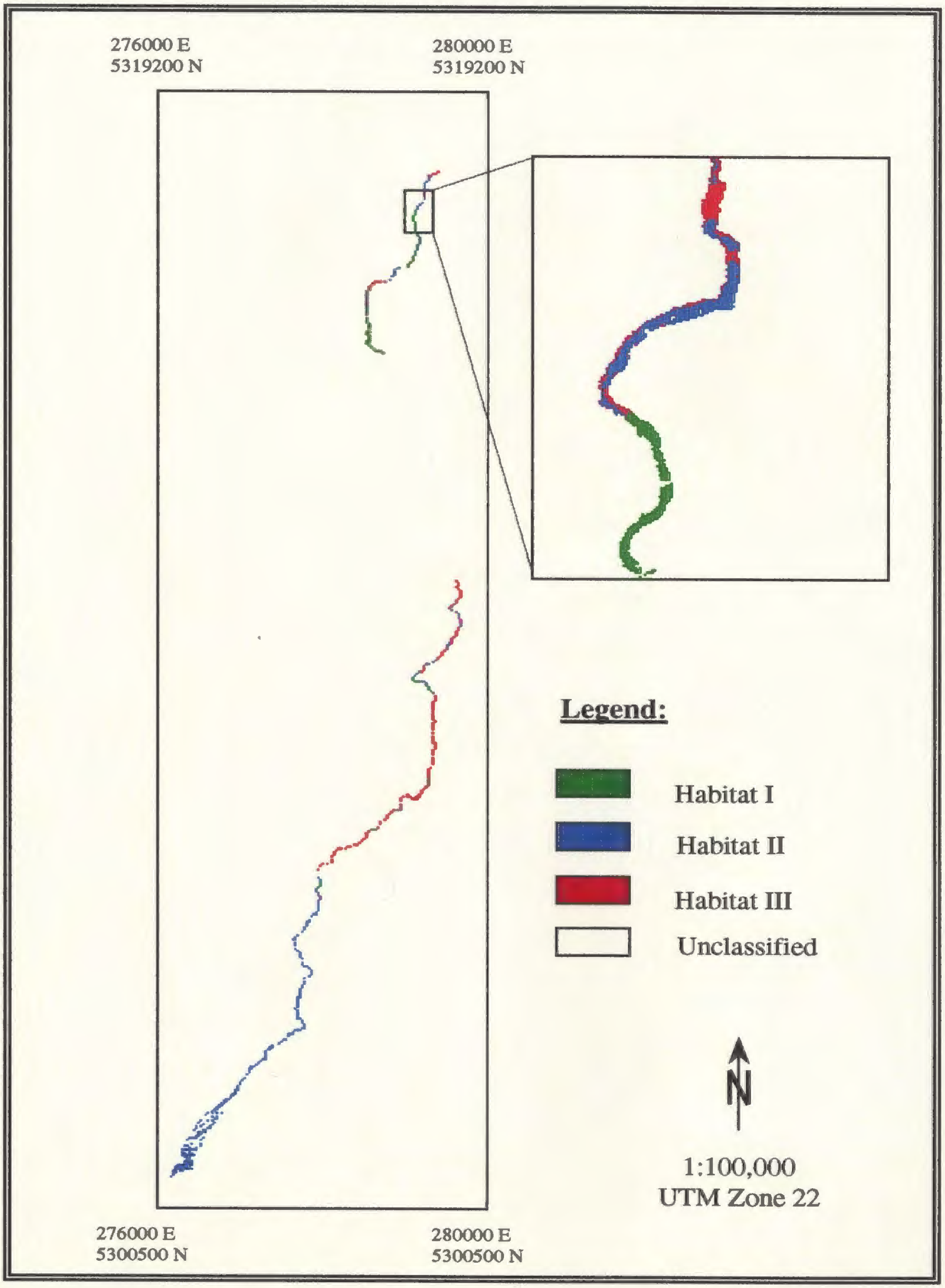

Figure 5-8: Spawning Habitat Suitability

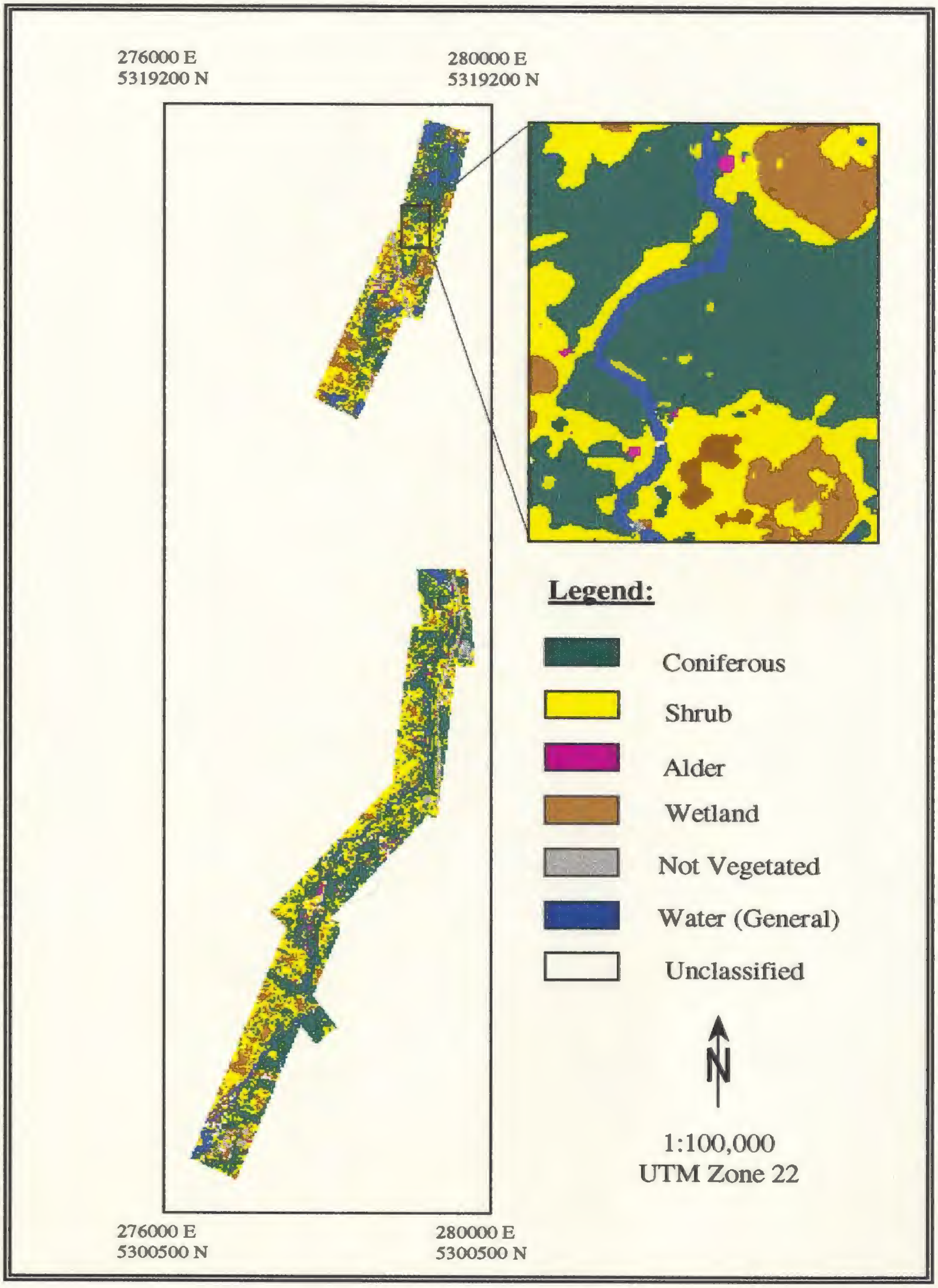

Figure 5-9: Land Cover Classification

Chapter 6.0: Discussion

6.1 Substrate Classification

The original spectral bands used in the classification of bottom substrate indicated limited discrimination between substrate categories. All bands showed considerable overlap in group variability and comparatively small differences in mean response. Standardized distances in these variables ranged from 0.040 to 0.756. An unexpected spectral response was observed in *Band 4*. At a central wavelength of 720 nm, the mean response was expected to be lower than in the visible spectral bands. However, the observed mean response values were highest in *Band 4* for all substrate classes. Most of the Come By Chance River is characterized by water depths in the decimeter range, resulting in comparatively little attenuation in the water column. Moreover, the response of photosynthetically active vegetation is higher in the near-infrared than in the visible spectral region. It is therefore likely that the high response in *Band 4* is caused by layers of microscopic algae covering the bottom substrate (Zacharias *et al.*, 1992). In areas of fast flowing water, sun glint due to surface roughness may also contribute to an uncharacteristically high response in *Band 4*. However, sun glint would result in a high reflectance for all spectral bands, whereas the spectral response in *Band 4* was observed to be consistently higher in all substrate and channel pattern categories. This response may in part also be caused by an inappropriate sensor gain setting. The same gain was used for all spectral bands although the expected amount of radiation reflected from water bodies is much smaller in the near-infrared.

Among all image variables, *PC2_ln* showed the strongest relationship with the type of bottom substrate with relatively large standardized distances between the substrate category *Gravel* and the respective classes *Rubble*, *Boulder* and *Bedrock*. The observed standardized distances ranged from 0.839 to 1.060. This variable was derived by transforming the original spectral bands so as to separate the signal into water depth dependent and bottom dependent components. It was therefore expected to carry information about the type of bottom substrate. The magnitude of correlation coefficients (from 0.845 to 0.989) between both PCA solutions suggests that both approaches performed similarly in describing the type of bottom substrate.

Separation of substrate classes was strongest in the non-spectral variables *Width* and *Gradient*. The variable *Gradient* was selected as first variable to split the data set, while the predictor *Width* further partitioned two of the three subsequent sub-sets. Conversely, the optical variables *Band 1*, and *PC2* appear only at the third level of the decision tree. The strong relationship between *Gradient* and *Substrate Type* reflects the presence of spatial correlation between these variables. Bottom substrates found in the study area are not uniformly distributed over the whole length of the river. Rather, they occur predominantly in certain river sections. For instance, all training pixels belonging to the substrate category *Gravel* were collected within 3 km from the estuary. A digital elevation model created from digitized contour lines was used to calculate valley gradient. Incidentally, very few contour intervals characterize the lower 3 km of the river course, resulting in a uniform gradient of 0.42 % for these sections. This coincides with the predominance of the substrate category *Gravel* in this part of the river.

Overall, the differentiation of substrate classes was comparatively weak in all predictor variables. This may be largely attributed to heterogeneity in the training and test data. Bottom substrate was recorded as proportions in seven grain size variables for river sections extending over 100 or more meters. Homogeneous bottom substrate, i.e. 100 % of the substrate in a given section belonging to one grain size class, was observed only in ten cases. Nine of these sections were characterized by bedrock substrate, while one showed a substrate composition of 100 % gravel. Consequently, the remaining river sections are characterized by more than one grain size variable. It was therefore required to group river sections of similar substrate composition together. The resulting substrate categories showed substantial overlap in grain size composition. For instance, the category *Gravel*, contains grain sizes from fines to rubble. The class *Rubble* contains equal proportions of 40 % of rubble and cobble substrates. An even greater overlap occurs between the classes *Rubble* and *Boulder*. A proportion of 30 % belonging to the category *Boulder* actually consist of the grain size class rubble. Finally, the class *Bedrock* consists of grain sizes ranging from cobble to bedrock substrate.

The heterogeneity inherent in substrate categories is reflected in the classification result. Per-class accuracy was highest in the categories *Gravel* and *Bedrock*. The corresponding values for user's and producer's accuracy ranged from 76.78 to 82.12 %. Omission and commission errors in these categories were predominantly caused by confusion with the substrate classes *Rubble* and *Boulder*. On the other hand, little confusion occurred between the categories *Gravel* and *Bedrock*. The definition of the substrate classes *Rubble* and *Boulder* was particularly poor. Classification errors of omission and commission ranged from 46.08 to 67.72 %. Combining the classes *Rubble* and *Boulder* resulted in a significant increase in overall classification accuracy from 66.80 to 73.76 %. Omission and commission errors in the combined categories were reduced to 31.70 and 36.80 %, respectively.

6.2 Channel Pattern Classification

The discrimination of channel pattern categories was weak in all predictor variables. In the original spectral bands, mean response was very similar for all categories with the exception of the channel pattern *Rapid*. This category showed consistently the highest mean response in all spectral bands. Moreover, the largest standardized distances were observed for a discrimination between *Rapid* and all other categories, ranging from 0.974 to 1.809. This type of channel pattern is characterized by turbulent flow and a broken water surface. Consequently, comparatively high spectral response characterized these areas. The same trend was observed in both principal component transformations. All spectral variables contributed little to a discrimination between channel pattern categories other than *Rapid*.

Differentiation between the channel pattern classes *Run* and *Riffle* versus *Flat* was indicated in the non-spectral variable *Width*. In both instances the corresponding standardized distance was 1.183. In addition, this variable contributed to the discrimination between *Run* and *Steady*, *Riffle* and *Steady*, as well as between *Steady* and *Flat*. The variable *Gradient*, on the other hand, showed standardized distances which were not significantly different from the largest distance observed only in two instances. The distance between the group means of the categories *Run* and *Steady* was 0.391, while the corresponding value for the classes *Riffle* versus *Steady* was 0.512.

In the classification of channel pattern, *Width* was the first variable selected to partition the data set. At the following level, all sub-sets were further divided by the variable *Gradient*. The importance of this variable in discriminating between classes of channel pattern was not indicated in the analysis of group separability using the standardized distance measure. The third and final level of the classification tree was formed by the spectral variables *Band 1* and *PC3*, indicating the relatively subordinate significance of the spectral variables in discriminating among channel pattern categories.

The overall classification result for the type of channel pattern was extremely poor. Only 38.11 % of all cases were correctly classified. This result was mainly caused by a large commission error in the class *Run*. More than two thirds of all observations classified as *Run* actually belonged to other categories, mostly to *Riffle* and *Steady* with 29.15 % and 25.97 %, respectively. In view of the large classification error, individual channel pattern categories were combined according to their ecological significance in providing suitable spawning habitat for Atlantic salmon. The resulting categories were *Riffle/Rapid* and *Run/Steady/Flat*. The overall classification accuracy was significantly improved with 64.47 % of all cases correctly classified. Both classes showed similar user's accuracy values of 68.57 and 63.05 %, respectively. However, the error of omission for the class *Riffle/Rapid* was still very large with a value of 59.58 %, while the combined category *Run/Steady/Flat* showed an error of omission of only 15.41 %.

These classification results indicate poor correspondence between training and verification data. During field data acquisition, the type of channel pattern was recorded for river sections extending over 100 or more meters, over which it was assumed that the type of channel pattern was constant. Nonetheless, visual analysis of image data and aerial photographs indicated that type of channel pattern may well vary within river sections. An example of this situation is presented in Figure 3-3. The entire section between Transects 6175 and 6300 was assigned the channel pattern *Riffle*, while the type of bottom substrate was *Bedrock*. Closer inspection of this section, however, showed clear tonal variations over the whole section. Three distinct tonal features can be observed at dark, light and very light tones. Tonal variation of this kind are used as key indicators for the identification of channel pattern from aerial photography (Dubois and Gosselin, 1989). This indicates that the type of channel pattern cannot be assumed constant over the surveyed river sections.

The cause for the discrepancy between assumed and actual consistency of channel pattern lies in the initial objective of the field based data collection effort. This survey was carried out following the guidelines for small stream inventories (Scruton *et al.*, 1992). These guidelines were established to ensure adequate data collection procedures for conventional, field based habitat inventories and to facilitate the ecological interpretation of the collected data. The procedures were not designed for the acquisition of reference data in support of an analysis of remotely sensed data. This also applies to the recording of substrate data which were collected at the same time.

6.3 Land Cover Classification

In the visible spectral bands, the lowest spectral response and variability was observed in the category *Water*. Among the vegetated land cover types, the category *Coniferous* showed the lowest mean response in *Band 1*, *Band 2* and *Band 3*. Both *Water* and *Coniferous* were well set apart from the remaining categories through standard distances from 0.739 to 2.718. Conversely, the categories *Shrub* and *Alder* were characterized by similar mean response values. The highest mean response and variability was observed in the classes *Wetland* and *Noveg*. In *Band 4*, the distinction between *Water* and all other categories is particularly strong with standardized distances ranging from 1.876 to 2.569. Mean response was similar in the categories *Coniferous*, *Alder* and *Noveg* with respective values of 177.35, 177.11 and 174.42. The classes *Shrub* and *Wetland* showed the highest response with mean values of 218.26 and 235.19, respectively. Overall, separability of land cover categories was limited in *Band 4*. In particular, the unvegetated class was barely differentiated from the vegetated categories. Generally the spectral response of photosynthetically active vegetation is much higher in the near-infrared than the amount of radiation reflected from unvegetated surfaces. The similar response of vegetated and not vegetated surfaces can be explained by the radiometric setting of all spectral bands. That is, during the image acquisition the maximum amount of reflected radiation to be registered in each spectral band was selected so as to maximize the information content over water covered areas. As a result, a relatively low peak SRU of 2.5 $\mu W \cdot cm^{-2} \cdot sr^{-1} \cdot nm^{-1}$ was selected for all bands. While preserving variability of low reflectance features, a low peak SRU can result in saturation and , consequently, loss of information over high reflectance targets. This is observed for the spectral response of vegetation in *Band 4*.

Likewise, water showed an uncharacteristically high response in the *NDVI*. Generally, water covered areas are expected to have negative values. However, the mean response of the land cover category *Water* was with 0.09 slightly higher in *Band 4* than in *Band 3*. Moreover, the category *Water* showed the highest variability (66.67 %) of all land cover classes. Most of the training pixels for *Water* were collected in the Come By Chance River. Therefore, the combination of shallow water and algal cover may produce a stronger signal in *Band 4* than in the visible spectral bands. However, both unvegetated categories were clearly separated from the vegetated classes in the *NDVI*.

The first principal component, *PC1*, largely reflected the spectral response of land cover categories in the visible spectral bands, whereas *PC2* showed the same general pattern as the *NDVI*. This corresponds well with the observed correlation coefficients, ranging from -0.878 to 0.936. An important feature of the principal components relates to the differentiation between the categories *Shrub* and *Alder*. The largest standardized distance between these categories was observed in the variable *PC4* with a value of 1.442. Likewise, the distance between *Shrub* and *Wetland* was the highest in this variable with 1.595. Both values were significantly larger than the corresponding distances observed in other variables.

The smallest classification errors were observed in the categories *Water* and *Noveg*. In both cases, user's and producer's accuracy values ranged from 93.39 to 100 %. Of the vegetated land cover classes, the category *Coniferous* showed the smallest classification errors with a user's and producer's accuracy of 87.99 % and 89.02 %, respectively. Errors of omission and commission were related to confusion with the classes *Shrub*, *Alder* and *Water*. The land cover categories most seriously affected by confusion were *Shrub*, *Alder* and *Wetland*, with per-class accuracies ranging from 69.55 to 86.67 %.

The confusion between the categories *Coniferous* and *Water* is likely to have occurred in areas of deep shadow. Figure 3-3 clearly shows the difference between sunlit and shaded areas. Training areas for the class *Alder* were located at distinctly identifiable alder swamps along the river. However, occurrence of this land cover category was not limited to these specific locations. Rather, a certain amount of mix between the classes *Coniferous*, *Alder* and *Shrub* was observed throughout the study area. While mixed stands were not prominent enough to have warranted inclusion as separate land cover category, confusion between *Coniferous*, *Shrub* and *Alder* is likely to be related to these stands. The comparatively large error of omission of 69.55 % in the land cover class *Wetland* was related to the fact that this category represented any open vegetation in the study area. Training and test areas of this type could be easily identified on photographs due to their location within closed stands of trees or shrubs, the actual species composition on these sites is very heterogeneous. It includes various types of vegetation such as moss, grass and shrub, all of which have different spectral characteristics.

6. 4 Suitability Mapping

This study was undertaken with the objective to examine the applicability of airborne remote sensing data and ancillary digital information to the mapping of the salmon habitat parameters *Substrate Type*, *Channel Pattern* and *Land Cover*. Until recently, salmon habitat mapping and modeling was conducted almost exclusively by relying on field surveys and air photo interpretation. Therefore, this investigation fulfills an important step in exploring the potential of digital remote sensing and data analysis for riverine fish habitat monitoring.

The full potential of digital databases lies in the ease of data manipulation and modelling. An example is provided of how to combine individual habitat parameters to model a particular aspect of salmon freshwater ecology, such as the quality of spawning habitat. Figure 5-8 contains a composite map of spawning habitat suitability. These results agree with previous studies undertaken at Come By Chance River (Harmon, 1966). Most of Habitat I, i.e. habitat most suitable for spawning, is located in the river sections north of Goobies Pond (Figure 3-2). Areas not suitable for spawning are predominantly found in the central part of the river where the bottom substrate is dominated by bedrock. The lower part of the river is almost exclusively characterized by habitats of intermediate suitability for spawning. Gravel is the main type of bottom substrate found in this area. The estuary is designated as Habitat II in Figure 5-8. However, this area was excluded from the analysis as no field data could be recovered for these sections.

Including a classification of land cover categories would make it possible to include potential sources of disturbance, such as areas of excessive erosion, proximity of dumps, proximity of roads and the provision of cover by riparian vegetation in the habitat modelling process. While these features cannot be discerned at the scale present in Figure 5-9, they can easily incorporated at larger scales since the spatial resolution of the data is 2 m.

Chapter 7.0: Conclusion

It was the objective of this study to evaluate the potential of multispectral remote sensing for the mapping of substrate type, channel pattern and land cover as important freshwater habitat parameters of Atlantic salmon . To this effect, digital imagery was acquired in October 1993 using a CASI sensor at the spectral wavelengths of 510, 590, 660 and 730 nm. The findings of a stream survey conducted in the same month served as a database for reference purposes together with aerial photography from May 1992. In addition, elevation data extracted from the digital map sheet NTS 1/N13 were used to calculate valley gradient. Next was the stage of pre-processing involving the atmospheric and geometric correction of the imagery, as well as the categorization of substrate data contained in the reference database.

A set of potential predictor variables was defined for each habitat parameter. Besides the original, atmospherically and geometrically corrected images, these included principal component transformations to separate depth dependent bottom type dependent signals. Moreover, the non-spectral variables of valley gradient and stream width were used as potential predictors. Valley gradient was calculated from a DEM built with the digital elevation data, while stream width was obtained from the imagery. The potential predictors for land cover consisted of the original spectral bands together with an NDVI and the components of a PCA applied to pixels over land areas.

A classification method was desired that permitted the integration of data from different sources, such as remotely sensed imagery and digital map data. This method should be statistically robust while maintaining the ability to obtain statistically significant relationships between variables.

These requirements are intrinsic properties of statistical decision tree analysis, which was therefore selected to classify image and ancillary data. Training and verification data were collected at random for each habitat parameter. From the training data, statistical properties for each predictor variable were derived in the form of standardized distances between group means, estimates of group variability, and bivariate correlation of predictor variables. These figures were subsequently used to determine the final predictor variables for each habitat parameter. Classification was carried out using exhaustive partitioning DTA. Separate decision trees were grown for each habitat parameter. Tree size was controlled by selecting appropriate stop sizes related to the smallest sampling units, and the partitioning process was stopped if no more splits were found to be significant at the 0.05 level of significance. The resulting decision trees were interpreted as classification rules and applied to the entire data sets.

Classification performance was evaluated using verification data and confusion matrices. Initial classification accuracy was increased for the habitat parameters substrate type and channel pattern by combining categories according to their ecological significance. The improved overall accuracies were 73.76 and 64.47 %, respectively. The classification accuracy for the type of land cover was 84.91 %. The individual habitat parameters as extracted in this research were combined to model spawning habitat suitability throughout the river course. Qualitatively, the result of this procedure was found to correspond well with the established knowledge about spawning habitat locations at Come By Chance River.

The current study illuminates the potential of digital remote sensing and image processing strategies for the mapping, inventory and modelling of freshwater salmon habitat. The important habitat parameters of bottom substrate and channel pattern were extracted so as to demonstrate the applicability of this method in terms of data integration and robust statistical classification. While the classification results for the habitat parameter substrate type do not indicate immediate operational use, the value of the approach followed in this study was confirmed in principle. Of particular interest was the emergence of the non-spectral variables *Width* and *Gradient* as the most important predictors for the type of bottom substrate. The contribution of the spectral variables to the discrimination between substrate categories in the present investigation is comparatively minor. On the other hand, application of the present method to the extraction of channel pattern categories did not yield entirely satisfactory results.

Error rates in the classification of substrate type and channel pattern were linked to the inadequate collection of supporting field data. Significant improvements in classification accuracy are therefore likely to be observed if special attention is devoted to the creation of reference data sets compatible with the objectives of remote sensing oriented investigations. In particular, emphasis should be placed upon the selection of homogeneous training and verification sites, as well as on proper geo-referencing of these areas. Moreover, shallow water bathymetry information should be included in the extraction of both bottom substrate and channel pattern. Previous studies suggest this may lead to significant improvements in classification results (e.g. Acornley *et al.*, 1995). The computation of stream width as introduced in the current investigation is deemed sufficient. However, a more reliable and accurate method of calculating the valley gradient is highly desirable. This could be achieved by using stereo-plotted vector data from digital map sheets as primary source rather than digitized contour lines, since stereo-plotted data contain an elevation value for every location. The type of land cover was identified with high accuracy. Since principles of land cover mapping using digital imagery are well established it can be extended to include more categories or specific vegetation communities, for example.

The results of this analysis suggest that the most beneficial future course in developing more efficient habitat inventory strategies for Atlantic salmon would include both remotely sensed and non-spectral information. Further development pertaining to the inclusion of geomorphic parameters in the analysis should be directed at the derivation of appropriate digital elevation models. With respect to remotely sensed data, improvement of the current results could be achieved by the inclusion of bathymetry information and by adopting appropriate sampling schemes in the collection of field data. Finally, research effort should be directed at issues of efficient integration of both spectral and non-spectral spatial data sources.

References

Acornley, R.M., Cutler, M.E.J., Milton, E.J. and Sear, D.A. (1995) *Detection and mapping of salmonid spawning habitat in chalk streams using airborne remote sensing.* In: Remote Sensing in Action. Proceedings of the 21st Annual Conference of the Remote Sensing Society, Southampton, pp. 267-274.

Agriculture Canada (1991) *Soils of the Sunnyside area, Newfoundland.* Report no.24, Newfoundland Soil Survey.

Amiro, P.G. (1983) *Stream gradients and Atlantic salmon parr densities.* Canadian Atlantic Fisheries Scientific Advisory Committee Research Document 84/89.

Bailey, T.C. and Gatrell, A.C. (1995) *Interactive spatial data analysis.* Longman, Essex, 413 pp.

Barila, T.Y., Williams, R.D. and Stauffer, J.R. (1981) *The influence of stream order and selected stream bed parameters on fish diversity in Raystown Branch, Susquehanna River Drainage, Pennsylvania.* Journal of Applied Ecology, Vol. 81, No. 18, pp. 125-131.

Basu, J.P. and Odell, P.L. (1974) *Effects of intraclass correlation among training samples on the misclassification probabilities of Bayes' procedure.* Pattern Recognition, Vol. 6, pp. 13-16.

Belward, A.S. and de Hoyos, A. (1987) *A comparison of supervised maximum likelihood and decision tree classification for crop cover estimation from multitemporal Landsat MSS data.* International Journal of Remote Sensing, Vol. 8, No. 2, pp. 229-235.

Benda, L., Beechie, T.J., Wissmar, R.C. and Johnson, A. (1992) *Morphology and evolution of salmonid habitats in a recently deglaciated river basin, Washington State, USA.* Canadian Journal of Fisheries and Aquatic Sciences, Vol. 49, pp. 1246-1256.

Bierwirth, P.N., Lee, T.J. and Burne, R.V. (1993) *Shallow sea-floor reflectance and water depth derived by unmixing multispectral imagery.* Photogrammetric Engineering and Remote Sensing, Vol. 59, No. 3, pp. 331-338.

Biggs, D., De Ville, B. and Suen, E. (1991) *A method of choosing multiway partitions for classification and decision trees.* Journal of Applied Statistics, Vol. 18, No. 1, pp. 49-62.

Bisson, P.A., Nielsen, J.L. and Palmason R.A. (1981) *A system of naming habitat types in small streams, with examples of habitat utilization by salmonids during low stream flow.* In: Armantrout, N. (Ed.): Proceedings of the symposium on acquisition and utilization of aquatic habitat inventory information, Portland 1981, pp. 62-73.

Borstad, G.A. (1992) *Ecosystem surveillance and monitoring with a portable airborne imaging spectrometer system.* In: Proceedings of the First Thematic Conference on Remote Sensing for Marine and Coastal Environments, New Orleans, pp. 883-892.

Borstad, G.A., Hill, D.A. and Kerr, R.C. (1992) *Direct digital remote sensing of herring schools.* International Journal of Remote Sensing, Vol. 13, No. 12, pp. 2191-2198.

Breiman, L., Friedman, J.K., Olshen, R.A. and Stone, C.J. (1984) *Classification and regression trees.* Wadsworth International Group, Belmont, 358pp.

Campbell, J.B. (1981) *Spatial correlation effects upon accuracy of supervised classification of land cover.* Photogrammetric Engineering and Remote Sensing, Vol. 47, No. 3, pp. 355-363.

Chavez, Jr., P.S. (1988) *An improved dark-object technique for atmospheric scattering correction of multispectral data.* Remote Sensing of Environment, 24, pp. 459-479.

Chou, Y-H., Minnich, R.A., Salazar, L.A., Power, J.D. and Dezzani, R.J. (1990) *Spatial autocorrelation of wildfire distribution in the Idyllwild quadrangle, San Jacinto Mountain, California.* Photogrammetric Engineering and Remote Sensing, Vol. 56, No. 11, pp. 1507-1513.

Clavet, D. (1980) *Approche géomorphologique dans la détermination du potentiel d'accueil salmonique des rivières des principales régions physiographiques du Québec.* Bulletin de Recherche Numéro 52-53, Département de Géographie, Université de Sherbrooke, Sherbrooke, Québec, 83 pp.

Cliff, A.D. and Ord, J.K. (1981) *Spatial Processes. Models and Applications.* Pion, London, 266 pp.

Congalton, R.G. and Green, K. (1993) *A practical look at the sources of confusion in error matrix generation.* Photogrammetric Engineering and Remote Sensing, Vol. 59, No. 5, pp. 641-644.

Côté, Y., Clavet, D., Dubois, J-M. and Boudreault, A. (1987) *Inventaire des habitats à saumon et estimation de production par photographie aérienne.* In: Thibault, M. and Billard, R. (Ed.): Restauration des rivières a saumons, INRA, Paris, pp. 85-93.

Cracknell, A.P. and Hayes, L.W.B. (1991) *Introduction to remote sensing.* Taylor and Francis, London, 293 pp.

Curran, P.J. (1983) *Multispectral remote sensing for the estimation of green leaf area index.* Philosophical Transactions of the Royal Society of London, A 309, pp. 257-270.

Damman, A.W.H. (1983) *An ecological subdivision of the island of Newfoundland.* In: South, G.R. (Ed.): Biogeography and ecology of the island of Newfoundland, The Hague, pp. 163-206.

Davis, J.C. (1986) *Statistics and data analysis in geology.* Wiley, New York, 646 pp.

Dekker, A.G., Malthus, M.M. and Seyhan, E. (1992) *The effect of spectral bandwidth and positioning on the spectral signature analysis of inland waters.* Remote Sensing of Environment, 41, pp. 211-225.

Dubois, J-M.M and Clavet, D. (1979) *Télédétection de l'hydromorphologie et du potentiel de frayères à saumon de la rivière Saint-Jean en Gaspésie: rapport final.* Département de Géographie, Université de Sherbrooke, Québec.

Dubois, J-M. M. and Gosselin A. (1989) *Airphoto interpretation and salmon river restoration.* In: IGARSS '89, 12th Canadian Symposium on Remote Sensing, Vancouver, Vol.4, pp. 2434-2438.

Dubois, J-M. M. and Gosselin A. (1994) *La restauration des rivières à saumons et la photo-interprétation.* In: Bonn, F. (Ed.) Télédétection de l'environnement dans l'espace francophone, pp. 479-501.

Duda, R.O. and Hart, P.E. (1973) *Pattern classification and scene analysis.* Wiley, New York, 482 pp.

Dymond, J.R. and Luckman, P.G. (1994) *Direct induction of compact rule-based classifiers for resource mapping.* International Journal of Geographical Information Systems, Vol. 8, No. 4, pp. 357-367.

Eastman, J.R. (1997) Idrisi for windows version 2.0 user's guide. Clark Labs for Cartographic Technology and Geographic Analysis, Worcester, MA, USA.

Edgington, J., Alexandersdottir, M., Burns, C. and Cariello, J. (1987) *Channel type classification as a method to document anadromous salmon streams.* Alaska Department of Fish and Game Informational Leaflet No. 260.

Fabricius, C. and Coetzee, K. (1992) *Geographic information systems and artificial intelligence for predicting the presence or absence of mountain reedbuck.* South African Journal of Wildlife Research, Vol. 22, No. 3, pp. 80-86.

Fisheries and Oceans Canada (1994) *Come-By-Chance River fish habitat survey.* Project L-633.

Fitzgerald, R.W. and Lees, B.G. (1994) *Assessing the classification accuracy of multisource remote sensing data.* Remote Sensing of Environment, 47, pp. 362-368.

Friedl, M.A. and Brodley, C.E. (1997) *Decision tree classification of land cover from remotely sensed data.* Remote Sensing of Environment, 61, pp. 399-409.

Frissell, C.A., Liss, W.J., Warren, C.E. and Hurley, M. (1986) *A hierarchical framework for stream habitat classification: viewing streams in a watershed context*. Environmental Management, Vol. 10, No. 2, pp. 199-214.

Gibson, R.J. (1993) *The Atlantic salmon in fresh water: spawning, rearing and production*. Reviews in Fish Biologies and Fisheries, 3, pp. 39-73.

Gonzalez, R.C. and Woods, R.E. (1992) *Digital image processing*. Addison-Wesley, Reading, Massachusetts, 716 pp.

Goodin, D.G., Luoheng, H., Fraser, R.N., Rundquist, D.C., Stebbins, W.A., Schalles, J.F. (1993) *Analysis of Suspended Solids in Water Using Remotely Sensed High Resolution Derivative Spectra*. Photogrammetric Engineering and Remote Sensing, Vol. 59, No. 4, pp. 505-510.

Green, E.J., Strawderman, W.E. and Airola, T.M. (1993) *Assessing classification probabilities for thematic maps*. Photogrammetric Engineering and Remote Sensing, Vol. 59, No. 5, pp. 635-639.

Griffith, D.A. (1987) *Spatial autocorrelation. A Primer*. Resource Publications in Geography, Washington D.C., 82 pp.

Griffith, D.A. and Amrhein, C.G. (1997) *Multivariate statistical analysis for geographers*. Prentice Hall, Upper Saddle River, New Jersey, 345 pp.

Günther, O., Hess, G., Mutz, M., Riekert, W.-F. and Ruwwe, T. (1993) *RESEDA: A knowledge-based advisory system for remote sensing*. Journal of Applied Intelligence, Vol. 1, No. 3, pp. 317-341.

Hamilton, M.K., Davis, C.O, Rhea, W.J., Pilorz, S.H. and Carder, K.L. (1993) *Estimating Chlorophyll Content and Bathymetry of Lake Tahoe Using AVIRIS Data*.
In: Remote Sensing of Environment, 44, pp. 217-230.

Harmon, T.J. (1966) *A proposal for the integrated development of the Black and Come By Chance River basins*. MS Report, Fisheries Services, St. John's, Newfoundland.

Hart, A. (1986) Knowledge *acquisition for expert systems*. McGraw-Hill, New York, 180 pp.

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*. In: Fisheries, Vol. 18, No. 6, pp. 3-12.

Hawkins, D.M. and Kass, G.V. (1982) *Automatic interaction detection*. In: Hawkins, D.M. (Ed.): Topics in applied multivariate analysis. Cambridge, pp. 269-302.

Heller, D.A. and Hohler, D.B. (1981) *A hierarchical inventory system and its application in a national forest fish habitat management program.* In: Proc. 7th N. Am. Trout Stream Improvement Workshop, Waterloo, Ont., 12-14 Sept. 1990. Southern Ontario Chapter, American Fisheries Society, pp. 149-155.

Hess, L.L., Melack, J.M., Filoso, S. and Wang, Y. (1995) *Delineation of inundated area and vegetation along the Amazon Floodplain with SIR-C synthetic aperture radar.* IEEE Transactions on Geoscience and Remote Sensing, Vol. 33, No. 4, pp. 896-903.

Janssen, L.L.F. and Van der Wel, F.J.M. (1994) *Accuracy assessment of satellite derived land-cover data: a review.* Photogrammetric Engineering and Remote Sensing, Vol. 60, No. 4, pp. 419-426.

Jensen, J.R. (1986) *Introductory digital image processing : a remote sensing perspective.* Prentice-Hall, Englewood Cliffs, New Jersey, 379 pp.

Jupp, D.L.B., Kirk, J.T.O. and Harris, G.P. (1994) *Detection, identification and mapping of cyanobacteria - using remote sensing to measure the optical quality of turbid inland waters.* Australian Journal of Marine and Freshwater Research, 45, pp. 801-828.

Kass, G.V. (1980) *An exploratory technique for investigating large quantities of categorical data.* Applied Statistics, Vol. 29, No. 2, pp. 119-127.

Kelly, R.V. (1991) *Practical knowledge engineering.* Digital Press, Hamilton, 212 pp.

Kershaw, C.D. and Fuller, R.M. (1992) *Statistical problems in the discrimination of land cover from satellite images: a case study in lowland Britain.* International Journal of Remote Sensing, Vol. 13, No. 16, pp. 3085-3104.

Khan, M.A., Fadlallah, Y.H. and Al-Hinai, K.G. (1992) *Thematic mapping of subtidal coastal habitats in the western Arabian Gulf using Landsat TM data - Abu Ali Bay, Saudi Arabia.* International Journal of Remote Sensing, Vol. 13, No. 4, pp. 605-614.

Knighton, D. (1984) *Fluvial forms and processes.* Edward Arnold, London, 218 pp.

Labovitz, M.L. and Masuoka, E.J. (1984) *The influence of autocorrelation in signature extraction - an example from a geobotanical investigation of Cotter Basin, Montana.* International Journal of Remote Sensing, Vol. 5, No. 2, pp. 315-332.

Lachenbruch, P.A. (1975) *Discriminant analysis.* Hafner Press, New York, 128 pp.

Lambert, É. (1994) *La localisation des peuplements de macrophytes immergés à l'aide de la télédétection et de données connexes.* In: Bonn, F. (Ed.) Télédétection de l'environnement dans l'espace francophone, pp. 129-146.

Lanka, R.P., Hubert, W.A. and Wesche, T. (1987) *Relations of geomorphology to stream habitat and trout standing stock in small Rocky Mountain streams.* Transaction of the American Fisheries Society 116, pp. 21-28.

Lark, R.M. (1995) *Components of accuracy of maps with special reference to discriminant analysis on remote sensor data.* International Journal of Remote Sensing, Vol. 16, No. 8, pp. 1461-1480.

Lathorp, R.G.Jr. and Lillesand, T.M. (1989) *Monitoring Water Quality and River Plume Transport in Green Bay, Lake Michigan with SPOT-1 Imagery.* Photogrammetric Engineering and Remote Sensing, Vol. 55, No. 3, pp. 349-354.

Lees, B.G. and Ritmann, K. (1991) *Decision-tree and rule induction approach to integration of remotely sensed and GIS data in mapping vegetation in disturbed or hilly environments.* Environmental Management, Vol. 15, No. 6, pp. 823-831.

Luczkovich, J.J, Wagner, T.W., Michalek, J.L. and Stoffle, R.W. (1993) *Discrimination of coral reefs, seagrass meadows, and sand bottom types from space: a Dominican Republic case study.* Photogrammetric Engineering and Remote Sensing, Vol 59, No 3, pp 385-389.

Lyon, J.G., Lunetta, R.S. and Williams, D.C. (1992) *Airborne multispectral scanner data for evaluating bottom sediment types and water depths of the St. Marys River, Michigan.* Photogrammetric Engineering and Remote Sensing, Vol. 58, No. 7, pp. 951-956.

Lyon, J.G. and Hutchinson, W.S. (1995) *Application of a radiometric model for evaluation of water depths and verification of results with airborne scanner data.* Photogrammetric Engineering and Remote Sensing, Vol. 61, No. 2, pp. 161-166.

Lyzenga, D.R. (1978) *Passive remote sensing techniques for mapping water depth and bottom features.* Applied Optics, Vol. 17, No. 3, pp. 379-383.

Lyzenga, D.R. (1981) *Remote sensing of bottom reflectance and water attenuation parameters in shallow water using aircraft and Landsat data.* International Journal of Remote Sensing, Vol. 2, No. 1, pp. 71-82.

Lyzenga, D.R. (1983) *Shallow-water bathymetry using combined lidar and passive multispectral scanner data.* International Journal of Remote Sensing, 6, pp. 115-125.

MacLeod, W., Stanton-Grey, R.A., Dyk, A. and Farrington, G.A. (1992) *Discrimination of substrate type using airborne remotely sensed data, Bay of Quinte, Ontario, Canada.* In: Proceedings of the 15th Canadian Symposium on Remote Sensing, Toronto, pp. 13-16.

Mather, P.M. (1976) *Computational methods of multivariate analysis in physical geography.* Wiley, London, 532 pp.

Mather, P.M. (1987) *Computer processing of remotely-sensed images.* Wiley, Chichester, 352 pp.

Mather, P.M. (1990) *Theoretical problems in image classification.* In: Steven, M.D. and Clark, J.A. (Ed. 1990): Applications of remote sensing in agriculture. London.

Moller-Jensen, L. (1990) *Knowledge based classification of an urban area using texture and context information in Landsat TM imagery.* Photogrammetric Engineering and Remote Sensing, Vol. 56, No. 6, pp. 899-904.

Morgan, J.A. and Sonquist, J.N. (1963) *Problems in the analysis of survey data: and a proposal.* Journal of the American Statistical Association 58, pp. 415-434.

Neter, J., Wasserman, W. and Whitmore, G.A. (1993) *Applied statistics.* Allyn and Bacon, Boston, 989 pp.

Nichol, J.E. (1993) *Remote sensing of tropical blackwater rivers: a method for environmental water quality analysis.* Applied Geography 13, pp. 153-168.

Odland, J. (1988) *Spatial Autocorrelation.* Sage, London, 87 pp.

Philpot, W.D. (1989) *Bathymetric mapping with passive multispectral imagery.* Applied Optics, Vol. 28, No. 8, pp. 1569-1579.

Quinlan, J.R., Compton, P.J., Horn, K.A. and Lazarus, L. (1987) *Inductive knowledge acquisition: a case study.* In: Quinlan, J.R. (Ed.): Applications of Expert Systems, Tuning Institute Press, Sydney, pp. 157-173.

Reddy, R.K.T. and Bonham-Carter, G.F. (1991) *A decision-tree approach to mineral potential mapping in Snow Lake area, Manitoba.* Canadian Journal of Remote Sensing, Vol. 17, No. 2, pp. 191-200.

Richards, J.A. (1987) *Remote sensing digital image analysis: an introduction.* Springer, New York, 281 pp.

Rimmer, J.C., Collins, M.B. and Pattiaratchi, C.B. (1987) *Mapping of water quality in coastal waters using Airborne Thematic Mapper.* International Journal of Remote Sensing, Vol. 8, No. 1, pp. 85-102.

Roberts, A., MacDonald, J.P. and Williams, I.V. (1992) *Digital classification of salmon spawning habitat: an evaluation of airborne multispectral video imagery.* In: Proceedings of the 15th Canadian Symposium on Remote Sensing, Toronto, pp. 54-57.

Rubin, T. (1992) *Multitemporal aerial photography to guide restoration of spawning habitat for anadromous fish.* In: Proceedings of the First Thematic Conference on Remote Sensing for Marine and Coastal Environments, New Orleans, pp. 893-896.

Safavian, S.R. and Landgrebe, D. (1991) *A survey of decision tree classifier methodology.* IEEE Transactions on Systems, Man, and Cybernetics, Vol. 21, No. 3, pp. 660-675.

Scruton, D.A., Anderson, T.C., Bourgeois, C.E. and O'Brien, J.P. (1992) *Small stream surveys for public sponsored habitat improvement and enhancement projects*. Canadian Manuscript report of Fisheries and Aquatic Sciences, No. 2163, 49 pp.

Scruton, D.A. and Gibson, R.J. (1993) *The development of habitat suitability curves for juvenile Atlantic salmon (Salmo salar) in riverine habitat in insular Newfoundland, Canada*. In: Gibson, R.J. and Cutting, R.E. (Ed.): Production of juvenile Atlantic salmon, Salmo salar, in natural waters. Canadian Special Publications in Fisheries and Aquatic Sciences 118, pp. 149-161.

Shearer, W.M. (1992) *The Atlantic salmon*. Wiley, New York, 244 pp.

Slater, P.N. (1980) *Remote sensing: optics and optical systems*. Wiley, London, 575 pp.

Stehman, S.V. (1997) *Selecting and interpreting measures of thematic classification accuracy*. Remote Sensing of Environment, 62, pp. 77-89.

Tabachnik, B.G. (1996) *Using multivariate statistics*. HaperCollins, New York, 880 pp.

Van Stokkam, H.T.C., Stokman, G.N.M. and Hovenier, J.W. (1993) *Quantitative use of passive optical remote sensing over coastal and inland water bodies*. International Journal of Remote Sensing, Vol. 14, No. 3, pp. 541-563.

Velleman, P.F. and Hoaglin, D.C. (1981) *Applications, basics, and computing of exploratory data analysis*. Duxbury Press, Boston, 354 pp.

Walford, N. (1995) *Geographical data analysis. Wiley*, Chichester, 446.

Walker, P.A. and Moore, D.M. (1988) *An inductive modelling and mapping tool for spatially-oriented data*. International Journal of Geographical Information Systems, Vol. 2, No. 4, pp. 347-363.

Wharton, S.W. (1987) *Spectral knowledge based approach for urban land cover discrimination*. IEEE Transactions on Geoscience and Remote Sensing, GE 25, No. 3, pp. 272-282.

Zacharias, M., Niemann, O. and Borstad, G.A. (1992) *An assessment and classification of a multispectral bandset for the remote sensing of intertidal seaweeds*. Canadian Journal of Remote Sensing, Vol. 18, No. 4, pp. 263-274.

Data Sources

CASI imagery: acquired by Provincial Airlines Ltd., St. John's, for Fisheries and Oceans, Canada. Date: October, 1993. Location: Come By Chance River, Newfoundland. Data acquisition mode: spatial. Nominal spatial resolution: 1.5 m. Number of swaths: 10. Number of spectral bands: 4. Spectral band configuration: 499.5 to 521.1 nm, 579.0 to 600.7 nm, 648.3 to 671.9 nm, 718.1 to 741.8 nm.

Aerial photographs: Newfoundland and Labrador Department of Natural Resources. Date: May 1992. Roll: 92208. Frames: 142, 144, 146, 152, 154, 156, 158, 166, 168, 170. Altitude: 1341 m. Focal Length: 152.7 mm. Approximate scale: 1:8,800.

Map NTS 1N/13: Natural Resources, Canada. Date: 1971. Scale: 1:50,000. Projection: UTM Zone 22. Type: topographic, digital.

Map 2C4-11: Newfoundland and Labrador Department of Forestry and Agriculture. Date: 1975. Scale: 1:12,500. Projection: UTM Zone 22. Type: ortho-photo, paper.

Map 1N/13-41: Newfoundland and Labrador Department of Forestry and Agriculture. Date: 1975. Scale: 1:12,500. Projection: UTM Zone 22. Type: ortho-photo, paper.

Map 1N/13-31: Newfoundland and Labrador Department of Forestry and Agriculture. Date: 1975. Scale: 1:12,500. Projection: UTM Zone 22. Type: ortho-photo, paper.

Map 1N/13-160: Newfoundland and Labrador Department of Forest Resources and Lands. Date: 1985. Scale: 1:5,000. Projection: modified 3° transverse mercator. Type: topographic, paper.

Map 1N/13-170: Newfoundland and Labrador Department of Forest Resources and Lands. Date: 1985. Scale: 1:5,000. Projection: modified 3° transverse mercator. Type: topographic, paper.

Appendix A

Group Variability and Standardized Distances

Table A-1: Descriptive Statistics of Substrate Categories

Predictor Variable	Gravel (n = 745)			Rubble (n = 366)			Boulder (n = 306)			Bedrock (n = 737)		
	Mean	SD	CV [%]	Mean	SD	CV [%]	Mean	SD	CV [%]	Mean	SD	CV [%]
Band1 [DN]	16.97	10.28	60.58	23.07	11.77	51.02	14.51	9.29	64.02	19.44	11.99	61.68
Band2 [DN]	28.46	12.67	44.52	28.87	15.16	52.51	20.61	12.65	61.38	26.72	16.56	61.98
Band3 [DN]	29.59	13.31	44.98	27.32	17.53	64.17	18.95	14.61	77.10	25.80	18.15	70.35
Band4 [DN]	31.02	16.71	53.87	37.32	20.59	55.17	30.25	17.79	58.81	33.97	20.56	60.52
PC1_ln [arb]	6.22	0.92	14.79	6.34	1.09	17.19	5.54	1.19	21.48	6.03	1.28	21.23
PC2_ln [arb]*	0.98	0.31	31.63	1.34	0.24	17.91	1.34	0.39	29.10	1.27	0.28	22.05
PC3_ln [arb]*	1.13	0.27	23.89	1.09	0.26	23.85	1.22	0.30	24.59	1.13	0.28	24.78
PC4_ln [arb]*	0.84	0.09	10.71	0.94	0.09	9.57	0.88	0.11	12.50	0.92	0.09	9.78
PC1 [arb]	54.15	24.65	45.52	58.97	31.35	53.16	43.26	25.34	58.58	53.86	32.16	59.71
PC2 [arb]*	61.10	8.08	13.22	56.80	8.45	14.88	54.05	9.95	18.41	57.21	9.90	17.30
PC3 [arb]*	33.00	6.77	20.52	38.81	6.41	16.52	36.48	5.66	15.52	36.80	5.88	15.98
PC4 [arb]*	6.90	2.00	28.99	9.49	2.22	23.39	8.07	1.92	23.79	8.49	2.04	24.03
Width [m]	23.08	8.69	37.65	12.22	4.05	33.14	17.95	6.97	38.83	19.62	6.22	31.70
Gradient [%]	0.44	0.17	38.64	1.48	1.22	82.43	1.11	1.22	109.91	0.34	0.15	44.12

* mean values adjusted for negative values: data minimum added to original mean to enable calculation of coefficient of variation

Table A-2: Descriptive Statistics of Channel Pattern Categories

Predictor Variable	Run (n = 545)			Riffle (n = 529)			Steady (n = 253)			Flat (n = 196)			Rapid (n=159)		
	Mean	SD	CV [%]	Mean	SD	CV [%]	Mean	SD	CV [%]	Mean	SD	CV [%]	Mean	SD	CV [%]
Band1 [DN]	19.63	11.98	61.03	22.05	10.88	49.34	15.20	10.10	66.45	17.12	6.86	40.07	39.52	19.62	49.65
Band2 [DN]	27.12	16.40	60.47	29.87	14.97	50.12	23.42	14.44	61.66	28.49	9.87	34.64	46.27	22.97	49.64
Band3 [DN]	25.54	17.39	68.09	29.10	18.00	61.86	22.51	14.61	64.90	29.80	11.57	38.83	40.20	23.33	58.03
Band4 [DN]	34.93	22.48	64.36	37.43	19.31	51.59	29.43	19.15	65.07	32.57	15.46	47.47	48.96	21.76	44.44
PC1_ln [arb]	6.09	1.20	19.70	6.35	1.15	18.11	5.74	1.21	21.08	6.34	0.65	10.25	7.18	1.11	15.46
PC2_ln [arb]*	1.00	0.30	30.00	0.99	0.25	25.25	0.82	0.30	36.59	0.74	0.33	44.59	1.19	0.21	17.65
PC3_ln [arb]*	1.01	0.27	26.73	1.00	0.30	30.00	1.04	0.28	26.92	1.01	0.20	19.80	0.82	0.30	36.59
PC4_ln [arb]*	0.55	0.10	18.18	0.55	0.10	18.18	0.51	0.12	23.53	0.51	0.10	19.61	0.54	0.09	16.67
PC1 [arb]	54.55	31.93	58.53	60.13	30.03	49.94	46.31	28.15	60.79	55.24	20.43	36.98	87.09	40.54	46.55
PC2 [arb]*	51.50	12.63	24.52	52.51	9.40	17.90	52.29	8.30	15.87	54.99	7.20	13.09	57.75	13.19	22.84
PC3 [arb]*	44.16	6.09	13.79	44.12	6.63	15.03	42.26	5.17	12.23	39.92	6.53	16.36	52.79	10.16	19.25
PC4 [arb]*	9.80	2.14	21.84	10.30	2.38	23.11	8.95	1.75	19.55	8.65	2.40	27.75	11.42	3.32	29.07
Width [m]	16.34	7.03	43.02	16.34	6.95	42.53	21.17	12.64	59.71	27.52	12.94	47.02	19.13	3.51	18.35
Gradient [%]	0.51	0.42	82.35	0.59	0.71	120.34	0.32	0.10	31.25	0.37	0.07	18.92	0.29	0.14	48.28

* mean values adjusted for negative values: data minimum added to original mean to enable calculation of coefficient of variation

Table A-3: Descriptive Statistics of Land Cover Categories

Predictor Variable	Coniferous (n = 533)			Shrub (n = 542)			Alder (n = 314)			Wetland (n = 283)			Noveg (n = 420)			Water (n=567)		
	Mean	SD	CV [%]	Mean	SD	CV [%]	Mean	SD	CV [%]	Mean	SD	CV [%]	Mean	SD	CV [%]	Mean	SD	CV [%]
Band1 [DN]	26.58	7.75	29.16	55.16	13.32	24.15	50.88	6.68	13.13	85.36	21.14	24.77	114.90	28.90	25.15	9.80	6.01	61.33
Band2 [DN]	47.29	14.07	29.75	90.70	21.70	23.93	72.40	8.06	11.13	124.02	25.28	20.38	148.68	29.30	19.71	16.38	9.33	56.96
Band3 [DN]	33.48	15.09	45.07	80.71	27.04	33.50	72.31	11.24	15.54	140.23	30.96	22.08	155.24	30.50	19.65	17.65	11.26	63.80
Band4 [DN]	177.35	55.82	31.47	218.26	40.50	18.56	177.11	28.12	15.88	235.19	31.85	13.54	174.42	37.94	21.75	19.88	11.95	60.11
NDVI [arb]*	0.92	0.11	11.96	0.70	0.14	20.00	0.66	0.07	10.61	0.50	0.07	14.00	0.29	0.12	41.38	0.33	0.22	66.67
PC1 [arb]*	162.28	47.31	29.15	242.27	43.15	17.81	201.59	25.00	12.40	308.93	50.33	16.29	298.65	46.91	15.71	32.64	17.50	53.62
PC2 [arb]*	63.65	36.07	56.67	77.32	31.37	40.57	97.50	19.39	19.89	111.73	19.07	17.07	183.28	40.31	21.99	160.19	8.00	4.99
PC3 [arb]*	62.73	7.00	11.16	59.86	11.27	18.83	57.83	4.51	7.80	47.30	12.09	25.56	59.39	14.82	24.95	49.59	4.33	8.73
PC4 [arb]*	25.26	3.75	14.85	21.90	6.71	30.64	29.05	2.36	8.12	29.81	3.89	13.05	24.30	4.51	18.56	25.94	1.87	7.21

* mean values adjusted for negative values: data minimum added to original mean to enable calculation of coefficient of variation

Table A-4: Standardized Distances - Substrate Type

Dependent Variable	(I) Substrate Type	(J) Substrate Type	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Zscore(BAND1)	Gravel	Rubble	-.5391907*	.062	.000	-.7130026	-.3653787
		Boulder	.2165287*	.066	.013	3.164282E-02	.4014147
		Bedrock	-.2186346*	.051	.000	-.3601011	-7.7168138E-02
	Rubble	Gravel	.5391907*	.062	.000	.3653787	.7130026
		Boulder	.7557194*	.075	.000	.5447960	.9666428
		Bedrock	.3205560*	.062	.000	.1464336	.4946785
	Boulder	Gravel	-.2165287*	.066	.013	-.4014147	-3.1642824E-02
		Rubble	-.7557194*	.075	.000	-.9666428	-.5447960
		Bedrock	-.4351634*	.066	.000	-.6203412	-.2499855
	Bedrock	Gravel	.2186346*	.051	.000	7.7168314E-02	.3601011
		Rubble	-.3205560*	.062	.000	-.4946785	-.1464336
		Boulder	.4351634*	.066	.000	.2499855	.6203412
Zscore(BAND2)	Gravel	Rubble	-2.7581112E-02	.063	.979	-.2033392	.1481770
		Boulder	.5319919*	.067	.000	.3450358	.7189480
		Bedrock	.1176768	.051	.152	-2.5373667E-02	.2607273
	Rubble	Gravel	2.7581112E-02	.063	.979	-.1481770	.2033392
		Boulder	.5595730*	.076	.000	.3462880	.7728581
		Bedrock	.1452579	.063	.150	-3.0814181E-02	.3213300
	Boulder	Gravel	-.5319919*	.067	.000	-.7189480	-.3450358
		Rubble	-.5595730*	.076	.000	-.7728581	-.3462880
		Bedrock	-.4143151*	.067	.000	-.6015664	-.2270639
	Bedrock	Gravel	-.1176768	.051	.152	-.2607273	2.5373667E-02
		Rubble	-.1452579	.063	.150	-.3213300	3.081418E-02
		Boulder	.4143151*	.067	.000	.2270639	.6015664
Zscore(BAND3)	Gravel	Rubble	.1386640	.062	.178	-3.6101203E-02	.3134292
		Boulder	.6503393*	.066	.000	.4644394	.8362392
		Bedrock	.2318951*	.051	.000	8.965278E-02	.3741375
	Rubble	Gravel	-.1386640	.062	.178	-.3134292	3.610120E-02
		Boulder	.5116753*	.076	.000	.2995951	.7237554
		Bedrock	9.323110E-02	.063	.528	-8.1846319E-02	.2683085
	Boulder	Gravel	-.6503393*	.066	.000	-.8362392	-.4644394
		Rubble	-.5116753*	.076	.000	-.7237554	-.2995951
		Bedrock	-.4184442*	.067	.000	-.6046376	-.2322507
	Bedrock	Gravel	-.2318951*	.051	.000	-.3741375	-8.9652782E-02
		Rubble	-9.3231101E-02	.063	.528	-.2683085	8.184632E-02
		Boulder	.4184442*	.067	.000	.2322507	.6046376
Zscore(BAND4)	Gravel	Rubble	-.3303386*	.063	.000	-.5075843	-.1530929
		Boulder	4.029903E-02	.067	.949	-.1482394	.2288375
		Bedrock	-.1546945*	.052	.030	-.2989557	-1.0433248E-02
	Rubble	Gravel	.3303386*	.063	.000	.1530929	.5075843
		Boulder	.3706376*	.077	.000	.1555473	.5857279
		Bedrock	.1756441	.063	.054	-1.9182632E-03	.3532065
	Boulder	Gravel	-4.0299033E-02	.067	.949	-.2288375	.1482394
		Rubble	-.3706376*	.077	.000	-.5857279	-.1555473
		Bedrock	-.1949935*	.067	.040	-.3838297	-6.1573705E-03
	Bedrock	Gravel	.1546945*	.052	.030	1.043325E-02	.2989557
		Rubble	-.1756441	.063	.054	-.3532065	1.918263E-03
		Boulder	.1949935*	.067	.040	6.157370E-03	.3838297
Zscore(PC1_LN)	Gravel	Rubble	-.1076420	.062	.396	-.2822497	6.696568E-02
		Boulder	.5867198*	.066	.000	.4009874	.7724521
		Bedrock	.1619881*	.051	.017	1.987395E-02	.3041022
	Rubble	Gravel	.1076420	.062	.396	-6.6965679E-02	.2822497
		Boulder	.6943617*	.076	.000	.4824728	.9062507
		Bedrock	.2696300*	.063	.000	9.471046E-02	.4445496
	Boulder	Gravel	-.5867198*	.066	.000	-.7724521	-.4009874
		Rubble	-.6943617*	.076	.000	-.9062507	-.4824728
		Bedrock	-.4247317*	.066	.000	-.6107573	-.2387061
	Bedrock	Gravel	-.1619881*	.051	.017	-.3041022	-1.9873947E-02
		Rubble	-.2696300*	.063	.000	-.4445496	-9.4710457E-02
		Boulder	.4247317*	.066	.000	.2387061	.6107573
Zscore(PC2_LN)	Gravel	Rubble	-1.0598448*	.057	.000	-1.2188065	-.9008830
		Boulder	-1.0520336*	.060	.000	-1.2211231	-.8829440
		Bedrock	-.8385917*	.046	.000	-.9679715	-.7092119

Table A-4: Standardized Distances - Substrate Type (cont.)

Dependent Variable	(I) Substrate Type	(J) Substrate Type	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Zscore(PC2_LN)	Rubble	Gravel	1.0598448*	.057	.000	.9008830	1.2188065
		Boulder	7.811213E-03	.069	1.000	-.1850912	.2007136
		Bedrock	.2212531*	.057	.002	6.200738E-02	.3804988
	Boulder	Gravel	1.0520336*	.060	.000	.8829440	1.2211231
		Rubble	-7.8112133E-03	.069	1.000	-.2007136	.1850912
		Bedrock	.2134419*	.061	.006	4.408534E-02	.3827984
	Bedrock	Gravel	.8385917*	.046	.000	.7092119	.9679715
		Rubble	-.2212531*	.057	.002	-.3804988	-6.2007378E-02
		Boulder	-.2134419*	.061	.006	-.3827984	-4.4085337E-02
Zscore(PC3_LN)	Gravel	Rubble	.1347893	.063	.209	-4.2249736E-02	.3118283
		Boulder	-.3379679*	.067	.000	-.5262865	-.1496493
		Bedrock	6.9462739E-03	.052	.999	-.1510393	.1371467
	Rubble	Gravel	-.1347893	.063	.209	-.3118283	4.224974E-02
		Boulder	-.4727572*	.077	.000	-.6875966	-.2579177
		Bedrock	-.1417356	.063	.172	-.3190908	3.561972E-02
	Boulder	Gravel	.3379679*	.067	.000	.1496493	.5262865
		Rubble	.4727572*	.077	.000	.2579177	.6875966
		Bedrock	.3310216*	.067	.000	.1424057	.5196375
	Bedrock	Gravel	6.946274E-03	.052	.999	-.1371467	.1510393
		Rubble	.1417356	.063	.172	-3.5619719E-02	.3190908
		Boulder	-.3310216*	.067	.000	-.5196375	-.1424057
Zscore(PC4_LN)	Gravel	Rubble	-.9688539*	.059	.000	-1.1335655	-.8041423
		Boulder	-.4331593*	.063	.000	-.6083651	-.2579535
		Bedrock	-.7759015*	.048	.000	-.9099612	-.6418418
	Rubble	Gravel	.9688539*	.059	.000	.8041423	1.1335655
		Boulder	.5356946*	.071	.000	.3358146	.7355746
		Bedrock	.1929524*	.059	.014	2.794650E-02	.3579583
	Boulder	Gravel	.4331593*	.063	.000	.2579535	.6083651
		Rubble	-.5356946*	.071	.000	-.7355746	-.3358146
		Bedrock	-.3427422*	.063	.000	-.5182246	-.1672598
	Bedrock	Gravel	.7759015*	.048	.000	.6418418	.9099612
		Rubble	-.1929524*	.059	.014	-.3579583	-2.7946502E-02
		Boulder	.3427422*	.063	.000	.1672598	.5182246
Zscore(PC1)	Gravel	Rubble	-.1662362	.063	.074	-.3427914	1.031905E-02
		Boulder	.3753194*	.067	.000	.1875155	.5631233
		Bedrock	1.004079E-02	.051	.998	-.1336584	.1537400
	Rubble	Gravel	.1662362	.063	.074	-1.0319052E-02	.3427914
		Boulder	.5415556*	.077	.000	.3273032	.7558079
		Bedrock	.1762770	.063	.051	-5.9365087E-04	.3531476
	Boulder	Gravel	-.3753194*	.067	.000	-.5631233	-.1875155
		Rubble	-.5415556*	.077	.000	-.7558079	-.3273032
		Bedrock	-.3652786*	.067	.000	-.5533791	-.1771781
	Bedrock	Gravel	-1.0040794E-02	.051	.998	-.1537400	.1336584
		Rubble	-.1762770	.063	.051	-.3531476	5.936509E-04
		Boulder	.3652786*	.067	.000	.1771781	.5533791
Zscore(PC2)	Gravel	Rubble	.4579292*	.062	.000	.2854076	.6304509
		Boulder	.7506781*	.066	.000	.5671648	.9341915
		Bedrock	.4141918*	.050	.000	.2737755	.5546080
	Rubble	Gravel	-.4579292*	.062	.000	-.6304509	-.2854076
		Boulder	.2927489*	.075	.002	8.339138E-02	.5021064
		Bedrock	-4.3737492E-02	.062	.919	-.2165673	.1290923
	Boulder	Gravel	-.7506781*	.066	.000	-.9341915	-.5671648
		Rubble	-.2927489*	.075	.002	-.5021064	-8.3391379E-02
		Bedrock	-.3364864*	.066	.000	-.5202895	-.1526833
	Bedrock	Gravel	-.4141918*	.050	.000	-.5546080	-.2737755
		Rubble	4.373749E-02	.062	.919	-.1290923	.2165673
		Boulder	.3364864*	.066	.000	.1526833	.5202895
Zscore(PC3)	Gravel	Rubble	-.8793683*	.060	.000	-1.0482617	-.7104749
		Boulder	-.5259130*	.064	.000	-.7055670	-.3462591
		Bedrock	-.5758233*	.049	.000	-.7132865	-.4383600
	Rubble	Gravel	.8793683*	.060	.000	.7104749	1.0482617
		Boulder	.3534553*	.073	.000	.1485007	.5584099
		Bedrock	.3035450*	.060	.000	.1343500	.4727401

Table A-4: Standardized Distances - Substrate Type (cont.)

Dependent Variable	(I) Substrate Type	(J) Substrate Type	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Zscore(PC3)	Boulder	Gravel	.5259130*	.064	.000	.3462591	.7055670
		Rubble	-.3534553*	.073	.000	-.5584099	-.1485007
		Bedrock	-4.9910232E-02	.064	.896	-.2298479	.1300274
	Bedrock	Gravel	.5758233*	.049	.000	.4383600	.7132865
		Rubble	-.3035450*	.060	.000	-.4727401	-.1343500
		Boulder	4.991023E-02	.064	.896	-.1300274	.2298479
Zscore(PC4)	Gravel	Rubble	-1.1523635*	.058	.000	-1.3148307	-.9898963
		Boulder	-.5213138*	.062	.000	-.6941321	-.3484954
		Bedrock	-.7073378*	.047	.000	-.8395707	-.5751049
	Rubble	Gravel	1.1523635*	.058	.000	.9898963	1.3148307
		Boulder	.6310498*	.070	.000	.4338935	.8282061
		Bedrock	.4450257*	.058	.000	.2822683	.6077831
	Boulder	Gravel	.5213138*	.062	.000	.3484954	.6941321
		Rubble	-.6310498*	.070	.000	-.8282061	-.4338935
		Bedrock	-.1860241*	.062	.029	-.3591153	-1.2932877E-02
	Bedrock	Gravel	.7073378*	.047	.000	.5751049	.8395707
		Rubble	-.4450257*	.058	.000	-.6077831	-.2822683
		Boulder	.1860241*	.062	.029	1.293288E-02	.3591153
Zscore(WIDTH)	Gravel	Rubble	1.3703350*	.056	.000	1.2124137	1.5282563
		Boulder	.6475368*	.060	.000	.4795541	.8155196
		Bedrock	.4359332*	.046	.000	.3074002	.5644661
	Rubble	Gravel	-1.3703350*	.056	.000	-1.5282563	-1.2124137
		Boulder	-.7227981*	.068	.000	-.9144379	-.5311584
		Bedrock	-.9344018*	.057	.000	-1.0926052	-.7761984
	Boulder	Gravel	-.6475368*	.060	.000	-.8155196	-.4795541
		Rubble	.7227981*	.068	.000	.5311584	.9144379
		Bedrock	-.2116037*	.060	.006	-.3798517	-4.3355631E-02
	Bedrock	Gravel	-.4359332*	.046	.000	-.5644661	-.3074002
		Rubble	.9344018*	.057	.000	.7761984	1.0926052
		Boulder	.2116037*	.060	.006	4.335563E-02	.3798517
Zscore(GRADIENT)	Gravel	Rubble	-1.2630902*	.054	.000	-1.4138187	-1.1123617
		Boulder	-.8113247*	.057	.000	-.9716564	-.6509930
		Bedrock	.1284238*	.044	.036	5.745118E-03	.2511025
	Rubble	Gravel	1.2630902*	.054	.000	1.1123617	1.4138187
		Boulder	.4517655*	.065	.000	.2688543	.6346767
		Bedrock	1.3915140*	.054	.000	1.2405163	1.5425117
	Boulder	Gravel	.8113247*	.057	.000	.6509930	.9716564
		Rubble	-.4517655*	.065	.000	-.6346767	-.2688543
		Bedrock	.9397485*	.057	.000	.7791636	1.1003334
	Bedrock	Gravel	-.1284238*	.044	.036	-.2511025	-5.7451191E-03
		Rubble	-1.3915140*	.054	.000	-1.5425117	-1.2405163
		Boulder	-.9397485*	.057	.000	-1.1003334	-.7791636

* The mean difference is significant at the .05 level.

Table A-5: Standardized Distances - Channel Pattern

Dependent Variable	(I) FLOW	(J) FLOW	Mean Difference (I-J)	Std Error	Sig	95% Confidence Interval Lower Bound	Upper Bound
Zscore(BAND1)	run	riffle	- 1803173*	054	025	- 3465210	-1 4113631E-02
		steady	3298561*	067	000	1226953	5370169
		flat	1870295	074	167	-3 9772773E-02	4138318
		rapid	-1 4794044*	080	000	-1 7248494	-1 2339594
	riffle	run	1803173*	054	025	1 411363E-02	3465210
		steady	5101735*	068	000	3020218	7183251
		flat	3673469*	074	000	1396391	5950546
		rapid	-1 2990870*	080	000	-1 5453689	-1 0528052
	steady	run	- 3298561*	067	000	- 5370169	- 1226953
		riffle	- 5101735*	068	000	- 7183251	- 3020218
		flat	- 1428266	084	577	- 4019458	1162925
		rapid	-1 8092605*	089	000	-2 0848448	-1 5336762
	flat	run	- 1870295	074	167	- 4138318	3 977277E-02
		riffle	- 3673469*	074	000	- 5950546	- 1396391
		steady	1428266	084	577	- 1162925	4019458
		rapid	-1 6664339*	094	000	-1 9570718	-1 3757960
	rapid	run	1 4794044*	080	000	1 2339594	1 7248494
		riffle	1 2990870*	080	000	1 0528052	1 5453689
		steady	1 8092605*	089	000	1 5336762	2 0848448
		flat	1 6664339*	094	000	1 3757960	1 9570718
Zscore(BAND2)	run	riffle	- 1632923	057	088	- 3400173	1 343283E-02
		steady	2200350	071	050	-2 3991489E-04	4403100
		flat	-8 1564177E-02	078	896	- 3227240	1595957
		rapid	-1 1369220*	085	000	-1 3979047	- 8759393
	riffle	run	1632923	057	088	-1 3432832E-02	3400173
		steady	3833273*	072	000	1619987	6046558
		flat	8 172808E-02	079	897	- 1603945	3238506
		rapid	- 9736298*	085	000	-1 2355023	- 7117572
	steady	run	- 2200350	071	050	- 4403100	2 399149E-04
		riffle	- 3833273*	072	000	- 6046558	- 1619987
		flat	- 3015992*	089	023	- 5771217	-2 6076728E-02
		rapid	-1 3569570*	095	000	-1 6499869	-1 0639271
	flat	run	8 156418E-02	078	896	- 1595957	3227240
		riffle	-8 1728076E-02	079	897	- 3238506	1603945
		steady	3015992*	089	023	2 607673E-02	5771217
		rapid	-1 0553578*	100	000	-1 3643943	- 7463214
	rapid	run	1 1369220*	085	000	8759393	1 3979047
		riffle	9736298*	085	000	7117572	1 2355023
		steady	1 3569570*	095	000	1 0639271	1 6499869
		flat	1 0553578*	100	000	7463214	1 3643943
Zscore(BAND3)	run	riffle	- 1991761*	059	023	- 3812116	-1 7140589E-02
		steady	1697852	074	256	-5 7108796E-02	3966792
		flat	- 2382720	081	068	- 4866784	1 013450E-02
		rapid	- 8200181*	087	000	-1 0888431	- 5511931
	riffle	run	1991761*	059	023	1 714059E-02	3812116
		steady	3689613*	074	000	1409820	5969406
		flat	-3 9095861E-02	081	994	- 2884940	2103023
		rapid	- 6208420*	087	000	- 8905836	- 3511004
	steady	run	- 1697852	074	256	- 3966792	5 710880E-02
		riffle	- 3689613*	074	000	- 5969406	- 1409820
		flat	- 4080572*	092	001	- 6918588	- 1242555
		rapid	- 9898033*	098	000	-1 2916385	- 6879681
	flat	run	2382720	081	068	-1 0134497E-02	4866784
		riffle	3 909586E-02	081	994	- 2103023	2884940
		steady	4080572*	092	001	1242555	6918588
		rapid	- 5817461*	103	000	- 9000689	- 2634234
	rapid	run	8200181*	087	000	5511931	1 0888431
		riffle	6208420*	087	000	3511004	8905836
		steady	9898033*	098	000	6879681	1 2916385
		flat	5817461*	103	000	2634234	9000689

Table A-5: Standardized Distances - Channel Pattern (cont.)

Dependent Variable	(I) FLOW	(J) FLOW	Mean Difference (I-J)	Std Error	Sig	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Zscore(BAND4)	run	riffle	- 1200515	059	394	- 3030129	6 290991E-02
		steady	2642196*	074	013	3 617157E-02	4922677
		flat	1136571	081	741	- 1360129	3633270
		rapid	- 6746487*	088	000	- 9448410	- 4044564
	riffle	run	1200515	059	394	-6 2909907E-02	3030129
		steady	3842711*	074	000	1551323	6134100
		flat	2337086	081	083	-1 6958078E-02	4843752
		rapid	- 5545972*	088	000	- 8257107	- 2834836
	steady	run	- 2642196*	074	013	- 4922677	-3 6171571E-02
		riffle	- 3842711*	074	000	- 6134100	- 1551323
		flat	- 1505625	093	618	- 4358077	1346826
		rapid	- 9388683*	098	000	-1 2422387	- 6354979
	flat	run	- 1136571	081	741	- 3633270	1360129
		riffle	- 2337086	081	083	- 4843752	1 695808E-02
		steady	1505625	093	618	- 1346826	4358077
		rapid	- 7883057*	104	000	-1 1082476	- 4683639
	rapid	run	6746487*	088	000	4044564	9448410
		riffle	5545972*	088	000	2834836	8257107
		steady	9388683*	098	000	6354979	1 2422387
		flat	7883057*	104	000	4683639	1 1082476
Zscore(PC1_LN)	run	riffle	- 2267288*	058	004	- 4058282	-4 7629525E-02
		steady	2953054*	072	002	7 207119E-02	5185397
		flat	- 2166579	079	113	- 4610576	2 774185E-02
		rapid	- 9277552*	086	000	-1 1922440	- 6632663
	riffle	run	2267288*	058	004	4 762952E-02	4058282
		steady	5220343*	073	000	2977323	7463363
		flat	1 007098E-02	080	1 000	- 2353044	2554464
		rapid	- 7010263*	086	000	- 9664170	- 4356356
	steady	run	- 2953054*	072	002	- 5185397	-7 2071188E-02
		riffle	- 5220343*	073	000	- 7463363	- 2977323
		flat	- 5119633*	091	000	- 7911873	- 2327393
		rapid	-1 2230606*	096	000	-1 5200272	- 9260940
	flat	run	2166579	079	113	-2 7741853E-02	4610576
		riffle	-1 0070979E-02	080	1 000	- 2554464	2353044
		steady	5119633*	091	000	2327393	7911873
		rapid	- 7110973*	102	000	-1 0242856	- 3979090
	rapid	run	9277552*	086	000	6632663	1 1922440
		riffle	7010263*	086	000	4356356	9664170
		steady	1 2230606*	096	000	9260940	1 5200272
		flat	7110973*	102	000	3979090	1 0242856
Zscore(PC2_LN)	run	riffle	3 533271E-02	056	983	- 1373456	2080111
		steady	6196596*	070	000	4044287	8348906
		flat	8705459*	076	000	6349082	1 1061835
		rapid	- 6130621*	083	000	- 8680687	- 3580556
	riffle	run	-3 5332715E-02	056	983	- 2080111	1373456
		steady	5843269*	070	000	3680665	8005874
		flat	8352132*	077	000	5986348	1 0717915
		rapid	- 6483949*	083	000	- 9042709	- 3925188
	steady	run	- 6196596*	070	000	- 8348906	- 4044287
		riffle	- 5843269*	070	000	- 8005874	- 3680665
		flat	2508862	087	083	-1 8327175E-02	5200997
		rapid	-1 2327218*	093	000	-1 5190417	- 9464018
	flat	run	- 8705459*	076	000	-1 1061835	- 6349082
		riffle	- 8352132*	077	000	-1 0717915	- 5986348
		steady	- 2508862	087	083	- 5200997	1 832717E-02
		rapid	-1 4836080*	098	000	-1 7855680	-1 1816480
	rapid	run	6130621*	083	000	3580556	8680687
		riffle	6483949*	083	000	3925188	9042709
		steady	1 2327218*	093	000	9464018	1 5190417
		flat	1 4836080*	098	000	1 1816480	1 7855680

Table A-5: Standardized Distances - Channel Pattern (cont.)

Dependent Variable	(I) FLOW	(J) FLOW	Mean Difference (I-J)	Std Error	Sig	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Zscore(PC3_LN)	run	riffle	2 065137E-02	060	998	- 1638857	2051885
		steady	- 1276344	075	570	- 3576464	1023776
		flat	-2 2829371E-02	082	999	- 2746495	2289908
		rapid	6539171*	088	000	3813978	9264363
	riffle	run	-2 0651366E-02	060	998	- 2051885	1638857
		steady	- 1482858	075	418	- 3793980	8 282644E-02
		flat	-4 3480737E-02	082	991	- 2963062	2093447
		rapid	6332657*	089	000	3598173	9067141
	steady	run	1276344	075	570	- 1023776	3576464
		riffle	1482858	075	418	-8 2826437E-02	3793980
		flat	1048050	093	968	- 1828967	3925068
		rapid	7815515*	099	000	4755684	1 0875345
	flat	run	2 282937E-02	082	999	- 2289908	2746495
		riffle	4 348074E-02	082	991	- 2093447	2963062
		steady	- 1048050	093	968	- 3925068	1828967
		rapid	6767464*	105	000	3540492	9994436
	rapid	run	- 6539171*	088	000	- 9264363	- 3813978
		riffle	- 6332657*	089	000	- 9067141	- 3598173
		steady	- 7815515*	099	000	-1 0875345	- 4755684
		flat	- 6767464*	105	000	- 9994436	- 3540492
Zscore(PC4_LN)	run	riffle	-8 9397282E-02	060	698	- 2749121	9 611754E-02
		steady	3087390*	075	002	7 750832E-02	5399697
		flat	3974894*	082	000	1443350	6506437
		rapid	4 116112E-02	089	995	- 2328020	3151242
	riffle	run	8 939728E-02	060	698	-9 6117539E-02	2749121
		steady	3981363*	075	000	1657996	6304730
		flat	4868866*	082	000	2327217	7410516
		rapid	1305584	089	709	- 1443388	4054556
	steady	run	- 3087390*	075	002	- 5399697	-7 7508319E-02
		riffle	- 3981363*	075	000	- 6304730	- 1657996
		flat	8 875034E-02	094	925	- 2004757	3779764
		rapid	- 2675779	100	126	- 5751822	4 002636E-02
	flat	run	- 3974894*	082	000	- 6506437	- 1443350
		riffle	- 4868866*	082	000	- 7410516	- 2327217
		steady	-8 8750338E-02	094	925	- 3779764	2004757
		rapid	- 3563282*	105	022	- 6807352	-3 1921297E-02
	rapid	run	-4 1161120E-02	089	995	- 3151242	2328020
		riffle	- 1305584	089	709	- 4054556	1443388
		steady	2675779	100	126	-4 0026360E-02	5751822
		flat	3563282*	105	022	3 192130E-02	6807352
Zscore(PC1)	run	riffle	- 1727593	058	064	- 3512802	5 761570E-03
		steady	2553730*	072	014	3 285977E-02	4778863
		flat	-2 1207075E-02	079	999	- 2648174	2224033
		rapid	-1 0080789*	085	000	-1 2717135	- 7444442
	riffle	run	1727593	058	064	-5 7615698E-03	3512802
		steady	4281323*	073	000	2045548	6517099
		flat	1515522	079	456	-9 3030667E-02	3961351
		rapid	- 8353196*	086	000	-1 0998531	- 5707860
	steady	run	- 2553730*	072	014	- 4778863	-3 2859772E-02
		riffle	- 4281323*	073	000	- 6517099	- 2045548
		flat	- 2765801	090	052	- 5549023	1 742085E-03
		rapid	-1 2634519*	096	000	-1 5594594	- 9674444
	flat	run	2 120707E-02	079	999	- 2224033	2648174
		riffle	- 1515522	079	456	- 3961351	9 303067E-02
		steady	2765801	090	052	-1 7420851E-03	5549023
		rapid	- 9868718*	101	000	-1 2990485	- 6746951
	rapid	run	1 0080789*	085	000	7444442	1 2717135
		riffle	8353196*	086	000	5707860	1 0998531
		steady	1 2634519*	096	000	9674444	1 5594594
		flat	9868718*	101	000	6746951	1 2990485

Table A-5: Standardized Distances - Channel Pattern (cont.)

Dependent Variable	(I) FLOW	(J) FLOW	Mean Difference (I-J)	Std Error	Sig	95% Confidence Interval Lower Bound	Upper Bound
Zscore(PC2)	run	riffle	-9 4571436E-02	060	650	- 2801525	9 100960E-02
		steady	-7 3584842E-02	075	915	- 3048981	1577284
		flat	- 3247597*	082	004	- 5780044	-7 1514952E-02
		rapid	- 5815885*	089	000	- 8556493	- 3075276
	riffle	run	9 457144E-02	060	650	-9 1009603E-02	2801525
		steady	2 098659E-02	075	999	- 2114330	2534062
		flat	- 2301882	082	100	- 4844439	2 406747E-02
		rapid	- 4870170*	089	000	- 7620124	- 2120217
	steady	run	7 358484E-02	075	915	- 1577284	3048981
		riffle	-2 0986594E-02	075	999	- 2534062	2114330
		flat	- 2511748	094	128	- 5405041	3 815449E-02
		rapid	- 5080036*	100	000	- 8157177	- 2002896
	flat	run	3247597*	082	004	7 151495E-02	5780044
		riffle	2301882	082	100	-2 4067473E-02	4844439
		steady	2511748	094	128	-3 8154494E-02	5405041
		rapid	- 2568288	105	203	- 5813515	6 769393E-02
	rapid	run	5815885*	089	000	3075276	8556493
		riffle	4870170*	089	000	2120217	7620124
		steady	5080036*	100	000	2002896	8157177
		flat	2568288	105	203	-6 7693933E-02	5813515
Zscore(PC3)	run	riffle	5 155252E-03	055	1 000	- 1655971	1759076
		steady	2574824*	069	008	4 465211E-02	4703128
		flat	5759981*	076	000	3429887	8090074
		rapid	-1 1732801*	082	000	-1 4254423	- 9211178
	riffle	run	-5 1552517E-03	055	1 000	- 1759076	1655971
		steady	2523272*	069	010	3 847886E-02	4661755
		flat	5708428*	076	000	3369032	8047824
		rapid	-1 1784353*	082	000	-1 4314574	- 9254133
	steady	run	- 2574824*	069	008	- 4703128	-4 4652114E-02
		riffle	- 2523272*	069	010	- 4661755	-3 8478862E-02
		flat	3185156*	086	009	5 230498E-02	5847263
		rapid	-1 4307625*	092	000	-1 7138889	-1 1476361
	flat	run	- 5759981*	076	000	- 8090074	- 3429887
		riffle	- 5708428*	076	000	- 8047824	- 3369032
		steady	- 3185156*	086	009	- 5847263	-5 2304982E-02
		rapid	-1 7492781*	097	000	-2 0478701	-1 4506861
	rapid	run	1 1732801*	082	000	9211178	1 4254423
		riffle	1 1784353*	082	000	9254133	1 4314574
		steady	1 4307625*	092	000	1 1476361	1 7138889
		flat	1 7492781*	097	000	1 4506861	2 0478701
Zscore(PC4)	run	riffle	- 2049948*	058	014	- 3839755	-2 6014153E-02
		steady	3440545*	072	000	1209681	5671409
		flat	4682928*	079	000	2240549	7125306
		rapid	- 6635434*	086	000	- 9278571	- 3992297
	riffle	run	2049948*	058	014	2 601415E-02	3839755
		steady	5490493*	073	000	3248959	7732027
		flat	6732876*	080	000	4280747	9185004
		rapid	- 4585486*	086	000	- 7237635	- 1933337
	steady	run	- 3440545*	072	000	- 5671409	- 1209681
		riffle	- 5490493*	073	000	- 7732027	- 3248959
		flat	1242383	090	757	- 1548008	4032773
		rapid	-1 0075979*	096	000	-1 3043678	- 7108280
	flat	run	- 4682928*	079	000	- 7125306	- 2240549
		riffle	- 6732876*	080	000	- 9185004	- 4280747
		steady	- 1242383	090	757	- 4032773	1548008
		rapid	-1 1318362*	101	000	-1 4448169	- 8188554
	rapid	run	6635434*	086	000	3992297	9278571
		riffle	4585486*	086	000	1933337	7237635
		steady	1 0075979*	096	000	7108280	1 3043678
		flat	1 1318362*	101	000	8188554	1 4448169

Table A-5: Standardized Distances - Channel Pattern (cont.)

Dependent Variable	(I) FLOW	(J) FLOW	Mean Difference (I-J)	Std Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Zscore(WIDTH)	run	riffle	3 139466E-04	056	1 000	- 1732141	1738420
		steady	- 5117881*	070	000	- 7280781	- 2954981
		flat	-1 1831034*	077	000	-1 4199006	- 9463063
		rapid	- 2949684*	083	014	- 5512297	-3 8707043E-02
	riffle	run	-3 1394664E-04	056	1 000	- 1738420	1732141
		steady	- 5121021*	070	000	- 7294266	- 2947775
		flat	-1 1834174*	077	000	-1 4211598	- 9456749
		rapid	- 2952823*	083	014	- 5524175	-3 8147210E-02
	steady	run	5117881*	070	000	2954981	7280781
		riffle	5121021*	070	000	2947775	7294266
		flat	6713153*	088	000	- 9418535	- 4007772
		rapid	2168197	093	249	-7 0909121E-02	5045485
	flat	run	1 1831034*	077	000	9463063	1 4199006
		riffle	1 1834174*	077	000	9456749	1 4211598
		steady	6713153*	088	000	4007772	9418535
		rapid	8881351*	098	000	5846892	1 1915809
	rapid	run	2949684*	083	014	3 870704E-02	5512297
		riffle	2952823*	083	014	3 814721E-02	5524175
		steady	- 2168197	093	249	- 5045485	7 090912E-02
		flat	- 8881351*	098	000	-1 1915809	- 5846892
Zscore(GRADIENT)	run	riffle	- 1719568	059	079	- 3552111	1 129747E-02
		steady	3907460*	074	000	1623329	6191591
		flat	2828710*	081	016	3 280143E-02	5329406
		rapid	4483976*	088	000	1777728	7190223
	riffle	run	1719568	059	079	-1 1297470E-02	3552111
		steady	5627028*	074	000	3331972	7922084
		flat	4548278*	081	000	2037599	7058957
		rapid	6203544*	088	000	3488068	8919019
	steady	run	- 3907460*	074	000	- 6191591	- 1623329
		riffle	- 5627028*	074	000	- 7922084	- 3331972
		flat	- 1078750	093	852	- 3935768	1778267
		rapid	5 765156E-02	099	987	- 2462044	3615076
	flat	run	- 2828710*	081	016	- 5329406	-3 2801426E-02
		riffle	- 4548278*	081	000	- 7058957	- 2037599
		steady	1078750	093	852	- 1778267	3935768
		rapid	1655266	104	638	- 1549274	4859805
	rapid	run	- 4483976*	088	000	- 7190223	- 1777728
		riffle	- 6203544*	088	000	- 8919019	- 3488068
		steady	-5 7651557E-02	099	987	- 3615076	2462044
		flat	- 1655266	104	638	- 4859805	1549274

* The mean difference is significant at the 05 level

Table A-6: Standardized Distances - Land Cover

Dependent Variable	(I) Land Cover	(J) Land Cover	Mean Difference (I-J)	Std Error	Sig	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Zscore(BAND1)	Coniferous	Shrub	- 7390535*	024	000	- 8205321	- 6575748
		Alder	- 6284221*	029	000	- 7234424	- 5334018
		Wetland	-1 5200664*	030	000	-1 6183070	-1 4218257
		Noveg	-2 2838214*	026	000	-2 3709702	-2 1966727
		Water	4339247*	024	000	3533415	5145078
	Shrub	Coniferous	7390535*	024	000	6575748	8205321
		Alder	1106313*	028	010	1 590400E-02	2053587
		Wetland	- 7810129*	029	000	- 8789703	- 6830556
		Noveg	-1 5447680*	026	000	-1 6315972	-1 4579387
		Water	1 1729781*	024	000	1 0927406	1 2532157
	Alder	Coniferous	6284221*	029	000	5334018	7234424
		Shrub	- 1106313*	028	010	- 2053587	-1 5903999E-02
		Wetland	- 8916443*	033	000	-1 0011236	- 7821649
		Noveg	-1 6553993*	030	000	-1 7550457	-1 5557529
		Water	1 0623468*	028	000	9683886	1 1563050
	Wetland	Coniferous	1 5200664*	030	000	1 4218257	1 6183070
		Shrub	7810129*	029	000	6830556	8789703
		Alder	8916443*	033	000	7821649	1 0011236
		Noveg	- 7637550*	031	000	- 8664769	- 6610332
		Water	1 9539910*	029	000	1 8567773	2 0512048
	Noveg	Coniferous	2 2838214*	026	000	2 1966727	2 3709702
		Shrub	1 5447680*	026	000	1 4579387	1 6315972
		Alder	1 6553993*	030	000	1 5557529	1 7550457
		Wetland	7637550*	031	000	6610332	8664769
		Water	2 7177461*	026	000	2 6317566	2 8037356
	Water	Coniferous	- 4339247*	024	000	- 5145078	- 3533415
		Shrub	-1 1729781*	024	000	-1 2532157	-1 0927406
		Alder	-1 0623468*	028	000	-1 1563050	- 9683886
		Wetland	-1 9539910*	029	000	-2 0512048	-1 8567773
		Noveg	-2 7177461*	026	000	-2 8037356	-2 6317566
Zscore(BAND2)	Coniferous	Shrub	- 8835925*	024	000	- 9626410	- 8045441
		Alder	- 5111583*	028	000	- 6033445	- 4189722
		Wetland	-1 5617619*	029	000	-1 6570724	-1 4664515
		Noveg	-2 0636941*	025	000	-2 1482435	-1 9791447
		Water	6291824*	023	000	5510028	7073620
	Shrub	Coniferous	8835925*	024	000	8045441	9626410
		Alder	3724342*	028	000	2805323	4643362
		Wetland	- 6781694*	029	000	- 7732050	- 5831338
		Noveg	-1 1801016*	025	000	-1 2643411	-1 0958621
		Water	1 5127749*	023	000	1 4349306	1 5906192
	Alder	Coniferous	5111583*	028	000	4189722	6033445
		Shrub	- 3724342*	028	000	- 4643362	- 2805323
		Wetland	-1 0506036*	032	000	-1 1568175	- 9443897
		Noveg	-1 5525358*	029	000	-1 6492101	-1 4558615
		Water	1 1403407*	027	000	1 0491850	1 2314964
	Wetland	Coniferous	1 5617619*	029	000	1 4664515	1 6570724
		Shrub	6781694*	029	000	5831338	7732050
		Alder	1 0506036*	032	000	9443897	1 1568175
		Noveg	- 5019322*	030	000	- 6015902	- 4022742
		Water	2 1909443*	028	000	2.0966301	2 2852585
	Noveg	Coniferous	2 0636941*	025	000	1 9791447	2.1482435
		Shrub	1 1801016*	025	000	1 0958621	1 2643411
		Alder	1 5525358*	029	000	1 4558615	1 6492101
		Wetland	5019322*	030	000	4022742	6015902
		Water	2 6928765*	025	000	2.6094518	2.7763012
	Water	Coniferous	- 6291824*	023	000	- 7073620	- 5510028
		Shrub	-1 5127749*	023	000	-1 5906192	-1 4349306
		Alder	-1 1403407*	027	000	-1 2314964	-1 0491850
		Wetland	-2.1909443*	028	000	-2 2852585	-2 0966301
		Noveg	-2 6928765*	025	000	-2 7763012	-2 6094518
Zscore(BAND3)	Coniferous	Shrub	- 8610877*	025	000	- 9426682	- 7795072
		Alder	- 7080432*	029	000	- 8031822	- 6129042
		Wetland	-1 9460703*	030	000	-2 0444337	-1 8477069
		Noveg	-2.2195273*	026	000	-2 3067849	-2 1322697
		Water	2882596*	024	000	2075758	3689435

Table A-6: Standardized Distances - Land Cover (cont.)

Dependent Variable	(I) Land Cover	(J) Land Cover	Mean Difference (I-J)	Std. Error	Sig	95% Confidence Interval	
						Lower Bound	Upper Bound
Zscore(BAND3)	Shrub	Coniferous	8610877*	025	000	7795072	9426682
		Alder	1530445*	028	000	5 819880E-02	2478902
		Wetland	-1 0849826*	029	000	-1 1830624	-9869029
		Noveg	-1 3584396*	026	000	-1 4453773	-1 2715018
		Water	1 1493473*	024	000	1 0690096	1 2296851
	Alder	Coniferous	7080432*	029	000	6129042	8031822
		Shrub	-1530445*	028	000	-2478902	-5 8198803E-02
		Wetland	-1 2380271*	033	000	-1 3476432	-1 1284110
		Noveg	-1 5114841*	030	000	-1 6112549	-1 4117132
		Water	9963028*	028	000	9022273	1 0903784
	Wetland	Coniferous	1 9460703*	030	000	1 8477069	2 0444337
		Shrub	1 0849826*	029	000	9869029	1 1830624
		Alder	1 2380271*	033	000	1 1284110	1 3476432
		Noveg	-2734569*	031	000	-3763071	-1706068
		Water	2 2343300*	029	000	2.1369948	2 3316652
	Noveg	Coniferous	2 2195273*	026	000	2 1322697	2 3067849
		Shrub	1 3584396*	026	000	1 2715018	1 4453773
		Alder	1 5114841*	030	000	1 4117132	1 6112549
		Wetland	2734569*	031	000	1706068	3763071
		Water	2 5077869*	026	000	2 4216900	2 5938838
	Water	Coniferous	-2882596*	024	000	-3689435	-2075758
		Shrub	-1 1493473*	024	000	-1 2296851	-1 0690096
		Alder	-9963028*	028	000	-1 0903784	-9022273
		Wetland	-2.2343300*	029	000	-2 3316652	-2 1369948
		Noveg	-2 5077869*	026	000	-2 5938838	-2 4216900
Zscore(BAND4)	Coniferous	Shrub	-4882049*	027	000	-5794951	-3969148
		Alder	2 917375E-03	032	1 000	-1035450	1093798
		Wetland	-6903353*	033	000	-8004059	-5802647
		Noveg	3 495632E-02	029	922	-6 2686682E-02	1325993
		Water	1 8793817*	027	000	1 7890949	1 9696685
	Shrub	Coniferous	4882049*	027	000	3969148	5794951
		Alder	4911223*	032	000	3849881	5972565
		Wetland	-2021304*	033	000	-3118835	-9 2377197E-02
		Noveg	5231613*	029	000	4258762	6204463
		Water	2 3675866*	027	000	2 2776871	2.4574862
	Alder	Coniferous	-2 9173747E-03	032	1 000	-1093798	1035450
		Shrub	-4911223*	032	000	-5972565	-3849881
		Wetland	-6932527*	037	000	-8159152	-5705901
		Noveg	3 203895E-02	034	969	-7 9606622E-02	1436845
		Water	1 8764643*	032	000	1 7711919	1 9817367
	Wetland	Coniferous	6903353*	033	000	5802647	8004059
		Shrub	2021304*	033	000	9 237720E-02	3118835
		Alder	6932527*	037	000	5705901	8159152
		Noveg	7252916*	035	000	6102003	8403830
		Water	2 5697170*	033	000	2.4607970	2 6786370
	Noveg	Coniferous	-3 4956324E-02	029	922	-1325993	6 268668E-02
		Shrub	-5231613*	029	000	-6204463	-4258762
		Alder	-3 2038949E-02	034	969	-1436845	7 960662E-02
		Wetland	-7252916*	035	000	-8403830	-6102003
		Water	1 8444254*	029	000	1 7480812	1 9407695
	Water	Coniferous	-1 8793817*	027	000	-1 9696685	-1 7890949
		Shrub	-2.3675866*	027	000	-2 4574862	-2 2776871
		Alder	-1 8764643*	032	000	-1 9817367	-1 7711919
		Wetland	-2 5697170*	033	000	-2 6786370	-2.4607970
		Noveg	-1 8444254*	029	000	-1 9407695	-1 7480812
Zscore(NDVI)	Coniferous	Shrub	8013281*	032	000	6950764	9075798
		Alder	9724737*	037	000	8485632	1 0963842
		Wetland	1 5574409*	038	000	1 4293309	1 6855510
		Noveg	2.3180049*	034	000	2.2043592	2.4316506
		Water	2 1517850*	032	000	2 0467011	2 2568688
	Shrub	Coniferous	-8013281*	032	000	-9075798	-6950764
		Alder	1711457*	037	001	4 761713E-02	2946742
		Wetland	7561129*	038	000	6283723	8838535
		Noveg	1 5166769*	034	000	1 4034478	1 6299060
		Water	1 3504569*	031	000	1 2458237	1 4550901

Table A-6: Standardized Distances - Land Cover (cont.)

Dependent Vanable	(I) Land Cover	(J) Land Cover	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
Zscore(NDVI)	Alder	Coniferous	- 9724737*	037	000	-1 0963842	- 8485632
		Shrub	- 1711457*	037	001	- 2946742	-4 7617132E-02
		Wetland	5849672*	043	000	4422015	7277329
		Noveg	1 3455312*	039	000	1 2155881	1 4754744
		Water	1 1793113*	037	000	1 0567858	1 3018367
	Wetland	Coniferous	-1 5574409*	038	000	-1 6855510	-1 4293309
		Shrub	- 7561129*	038	000	- 8838535	- 6283723
		Alder	- 5849672*	043	000	- 7277329	- 4422015
		Noveg	7605640*	040	000	6266103	8945177
		Water	5943440*	038	000	4675731	7211149
	Noveg	Coniferous	-2 3180049*	034	000	-2 4316506	-2 2043592
		Shrub	-1 5166769*	034	000	-1 6299060	-1 4034479
		Alder	-1 3455312*	039	000	-1 4754744	-1 2155881
		Wetland	- 7605640*	040	000	- 8945177	- 6266103
		Water	- 1662199*	034	000	- 2783539	-5 4085972E-02
	Water	Coniferous	-2 1517850*	032	000	-2.2568688	-2 0467011
		Shrub	-1 3504569*	031	000	-1 4550901	-1 2458237
		Alder	-1 1793113*	037	000	-1 3018367	-1 0567858
		Wetland	- 5943440*	038	000	- 7211149	- 4675731
		Noveg	1662199*	034	000	5 408597E-02	2783539
Zscore(PC1)	Coniferous	Shrub	- 7643400*	023	000	- 8415879	- 6870920
		Alder	- 3755895*	027	000	- 4656759	- 2855031
		Wetland	-1 4011516*	028	000	-1 4942911	-1 3080120
		Noveg	-1 3029673*	025	000	-1 3855910	-1 2203437
		Water	1 2385917*	023	000	1 1621928	1 3149906
	Shrub	Coniferous	7643400*	023	000	6870920	8415879
		Alder	3887505*	027	000	2989418	4785592
		Wetland	- 6368116*	028	000	- 7296826	- 5439406
		Noveg	- 5386274*	025	000	- 6209481	- 4563066
		Water	2 0029317*	023	000	1 9268604	2 0790029
	Alder	Coniferous	3755895*	027	000	2855031	4656759
		Shrub	- 3887505*	027	000	- 4785592	- 2989418
		Wetland	-1 0255621*	031	000	-1 1293568	- 9217674
		Noveg	- 9273779*	028	000	-1 0218502	- 8329056
		Water	1 6141812*	027	000	1 5251017	1 7032606
	Wetland	Coniferous	1 4011516*	028	000	1 3080120	1 4942911
		Shrub	6368116*	028	000	5439406	7296826
		Alder	1 0255621*	031	000	9217674	1 1293568
		Noveg	9 818423E-02*	029	047	7 961607E-04	1955723
		Water	2 6397433*	028	000	2 5475773	2 7319093
	Noveg	Coniferous	1 3029673*	025	000	1 2203437	1 3855910
		Shrub	5386274*	025	000	4563066	6209481
		Alder	9273779*	028	000	8329056	1 0218502
		Wetland	-9 8184228E-02*	029	047	- 1955723	-7 9616075E-04
		Water	2 5415590*	024	000	2 4600345	2 6230836
	Water	Coniferous	-1 2385917*	023	000	-1 3149906	-1 1621928
		Shrub	-2 0029317*	023	000	-2 0790029	-1 9268604
		Alder	-1 6141812*	027	000	-1 7032606	-1 5251017
		Wetland	-2 6397433*	028	000	-2 7319093	-2 5475773
		Noveg	-2 5415590*	024	000	-2 6230836	-2 4600345
Zscore(PC2)	Coniferous	Shrub	- 2564303*	033	000	- 3652029	- 1476576
		Alder	- 6350077*	038	000	- 7618582	- 5081572
		Wetland	- 9019321*	039	000	-1 0330817	- 7707824
		Noveg	-2.2443535*	035	000	-2.3606957	-2 1280114
		Water	-1 8111883*	032	000	-1 9187655	-1 7036112
	Shrub	Coniferous	2564303*	033	000	1476576	3652029
		Alder	- 3785774*	038	000	- 5050369	- 2521180
		Wetland	- 6455018*	039	000	- 7762732	- 5147304
		Noveg	-1 9879233*	035	000	-2 1038389	-1 8720076
		Water	-1 5547580*	032	000	-1 6618738	-1 4476423
	Alder	Coniferous	6350077*	038	000	5081572	7618582
		Shrub	3785774*	038	000	2521180	5050369
		Wetland	- 2669244*	044	000	- 4130774	- 1207713
		Noveg	-1 6093458*	040	000	-1 7423721	-1 4763196
		Water	-1 1761806*	038	000	-1 3016132	-1 0507480

Table A-6: Standardized Distances - Land Cover (cont.)

Dependent Variable	(I) Land Cover	(J) Land Cover	Mean Difference (I-J)	Std. Error	Sig	95% Confidence Interval Lower Bound	Upper Bound
Zscore(PC2)	Wetland	Coniferous	9019321*	039	000	7707824	1 0330817
		Shrub	6455018*	039	000	5147304	7762732
		Alder	2669244*	044	000	1207713	4130774
		Noveg	-1 3424215*	041	000	-1 4795534	-1 2052895
		Water	- 9092563*	039	000	-1 0390350	- 7794775
	Noveg	Coniferous	2 2443535*	035	000	2 1280114	2 3606957
		Shrub	1 9879233*	035	000	1 8720076	2 1038389
		Alder	1 6093458*	040	000	1 4763196	1 7423721
		Wetland	1 3424215*	041	000	1 2052895	1 4795534
		Water	4331652*	034	000	3183707	5479597
	Water	Coniferous	1 8111883*	032	000	1 7036112	1 9187655
		Shrub	1 5547580*	032	000	1 4476423	1 6618738
		Alder	1 1761806*	038	000	1 0507480	1 3016132
		Wetland	9092563*	039	000	7794775	1 0390350
		Noveg	- 4331652*	034	000	- 5479597	- 3183707
Zscore(PC3)	Coniferous	Shrub	2590764*	053	000	8 306001E-02	4350928
		Alder	4417397*	062	000	2364698	6470097
		Wetland	1 3911738*	064	000	1 1789469	1 6034007
		Noveg	3010636*	057	000	1127983	4893289
		Water	1 1851956*	052	000	1 0111138	1 3592774
	Shrub	Coniferous	- 2590764*	053	000	- 4350928	-8 3060007E-02
		Alder	1826634	061	116	-2 1973825E-02	3873005
		Wetland	1 1320974*	064	000	9204825	1 3437123
		Noveg	4 198719E-02	056	990	- 1455880	2295624
		Water	9261192*	052	000	7527840	1 0994544
	Alder	Coniferous	- 4417397*	062	000	- 6470097	- 2364698
		Shrub	- 1826634	061	116	- 3873005	2 197383E-02
		Wetland	9494340*	071	000	7129285	1 1859395
		Noveg	- 1406762	065	449	- 3559398	7 458748E-02
		Water	7434559*	061	000	5404803	9464314
	Wetland	Coniferous	-1 3911738*	064	000	-1 6034007	-1 1789469
		Shrub	-1 1320974*	064	000	-1 3437123	- 9204825
		Alder	- 9494340*	071	000	-1 1859395	- 7129285
		Noveg	-1 0901102*	067	000	-1 3120177	- 8682027
		Water	- 2059782	063	059	- 4159867	4 030298E-03
	Noveg	Coniferous	- 3010636*	057	000	- 4893289	- 1127983
		Shrub	-4 1987195E-02	056	990	- 2295624	1455880
		Alder	1406762	065	449	-7 4587484E-02	3559398
		Wetland	1 0901102*	067	000	8682027	1 3120177
		Water	8841320*	056	000	6983710	1 0698930
	Water	Coniferous	-1 1851956*	052	000	-1 3592774	-1 0111138
		Shrub	- 9261192*	052	000	-1 0994544	- 7527840
		Alder	- 7434559*	061	000	- 9464314	- 5404803
		Wetland	2059782	063	059	-4 0302983E-03	4159867
		Noveg	- 8841320*	056	000	-1 0698930	- 6983710
Zscore(PC4)	Coniferous	Shrub	6784147*	053	000	5034974	8533319
		Alder	- 7634593*	061	000	- 9674475	- 5594711
		Wetland	- 9162776*	063	000	-1 1271793	- 7053760
		Noveg	1948514*	056	035	7 761676E-03	3819411
		Water	- 1372122	052	223	- 3102069	3 578256E-02
	Shrub	Coniferous	- 6784147*	053	000	- 8533319	- 5034974
		Alder	-1 4418740*	061	000	-1 6452333	-1 2385146
		Wetland	-1 5946923*	063	000	-1 8049858	-1 3843988
		Noveg	- 4835633*	056	000	- 6699672	- 2971594
		Water	- 8156269*	052	000	- 9878797	- 6433741
	Alder	Coniferous	7634593*	061	000	5594711	9674475
		Shrub	1 4418740*	061	000	1 2385146	1 6452333
		Wetland	- 1528183	071	455	- 3878470	8 221033E-02
		Noveg	9583107*	064	000	7443912	1 1722301
		Water	6262471*	061	000	4245390	8279552
	Wetland	Coniferous	9162776*	063	000	7053760	1 1271793
		Shrub	1 5946923*	063	000	1 3843988	1 8049858
		Alder	1528183	071	455	-8 2210331E-02	3878470
		Noveg	1 1111290*	066	000	8906072	1 3316508
		Water	7790654*	063	000	5703683	9877625

Table A-6: Standardized Distances - Land Cover (cont.)

Dependent Variable	(I) Land Cover	(J) Land Cover	Mean Difference (I-J)	Std Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Zscore(PC4)	Noveg	Coniferous	-.1948514*	.056	.035	-.3819411	-7.7616755E-03
		Shrub	.4835633*	.056	.000	.2971594	.6699672
		Alder	-.9583107*	.064	.000	-1.1722301	-.7443912
		Wetland	-1.1111290*	.066	.000	-1.3316508	-.8906072
		Water	-.3320636*	.055	.000	-.5166646	-.1474625
	Water	Coniferous	.1372122	.052	.223	-3.5782558E-02	.3102069
		Shrub	.8156269*	.052	.000	.6433741	.9878797
		Alder	-.6262471*	.061	.000	-.8279552	-.4245390
		Wetland	-.7790654*	.063	.000	-.9877625	-.5703683
		Noveg	.3320636*	.055	.000	.1474625	.5166646

* The mean difference is significant at the .05 level.

Appendix B

Decision Rules for the Classification of Habitat Parameters

Table B1: Decision Rules for Substrate Type Classification

Rule 1 If
- Gradient = [0.17,0.29[
- Width = [0,14[
- Band1 = [1 19[

Then
- Substrate Type = Rubble 8.0%
- Substrate Type = Boulder 59.3%
- Substrate Type = Bedrock 32.7%

Rule 2 If
- Gradient = [0.17,0.29[
- Width = [0,14[
- Band1 = [19,34[

Then
- Substrate Type = Rubble 39.8%
- Substrate Type = Boulder 24.7%
- Substrate Type = Bedrock 35.5%

Rule 3 If
- Gradient = [0.17,0.29[
- Width = [0,14[
- Band1 = [34,70]

Then
- Substrate Type = Gravel 25.7%
- Substrate Type = Rubble 25.7%
- Substrate Type = Boulder 12.9%
- Substrate Type = Bedrock 35.6%

Rule 4 If
- Gradient = [0.17,0.29[
- Width = [14,60]

Then
- Substrate Type = Bedrock 100.0%

Rule 5 If
- Gradient = [0.29,0.43[
- Pc4 = [-8.24,3.3[

Then
- Substrate Type = Gravel 87.2%
- Substrate Type = Rubble 4.4%
- Substrate Type = Boulder 8.3%

Rule 6 If
- Gradient = [0.29,0.43[
- Pc4 = [3.3,14.79]

Then
- Substrate Type = Gravel 11.8%
- Substrate Type = Rubble 82.4%
- Substrate Type = Boulder 5.9%

Rule 7 If
- Gradient = [0.43,3.68]
- Width = [0,14[

Then
- Substrate Type = Gravel 12.8%
- Substrate Type = Rubble 67.3%
- Substrate Type = Bedrock 19.9%

Rule 8 If
- Gradient = [0.43,3.68]
- Width = [14,20[
- Pc2 = [-90.55,0.44[

Then
- Substrate Type = Rubble 54.1%
- Substrate Type = Boulder 6.8%
- Substrate Type = Bedrock 39.0%

Rule 9 If
- Gradient = [0.43,3.68]
- Width = [14,20[
- Pc2 = [0.44,2.12[

Then
- Substrate Type = Rubble 29.0%
- Substrate Type = Boulder 58.1%
- Substrate Type = Bedrock 12.9%

Rule 10 If
- Gradient = [0.43,3.68]
- Width = [14,20[
- Pc2 = [2.12,22.35]

Then
- Substrate Type = Rubble 18.6%
- Substrate Type = Boulder 16.9%
- Substrate Type = Bedrock 64.4%

Rule 11 If
- Gradient = [0.43,3.68]
- Width = [20,60]

Then
- Substrate Type = Boulder 26.1%
- Substrate Type = Bedrock 73.9%

Table B2: Decision Rules for Channel Pattern Classification

Rule 1 If
Width = [0,12[
Gradient = [0,0.45[
Then
Channel Pattern = Run 63.6%
Channel Pattern = Riffle 11.5%
Channel Pattern = Steady 17.2%
Channel Pattern = Flat 7.7%

Rule 2 If
Width = [0,12[
Gradient = [0.45,3.68]
Then
Channel Pattern = Run 1.4%
Channel Pattern = Riffle 98.6%

Rule 3 If
Width = [12,17[
Gradient = [0,0.36[
Band1 = [1,9[
Then
Channel Pattern = Run 14.3%
Channel Pattern = Riffle 21.4%
Channel Pattern = Steady 64.3%

Rule 4 If
Width = [12,17[
Gradient = [0,0.36[
Band1 = [9,13[
Then
Channel Pattern = Run 19.0%
Channel Pattern = Riffle 9.5%
Channel Pattern = Steady 23.8%
Channel Pattern = Flat 42.9%
Channel Pattern = Rapid 4.8%

Rule 5 If
Width = [12,17[
Gradient = [0,0.36[
Band1 = [13,27[
Then
Channel Pattern = Run 8.2%
Channel Pattern = Riffle 34.0%
Channel Pattern = Steady 4.1%
Channel Pattern = Flat 30.9%
Channel Pattern = Rapid 22.7%

Rule 6 If
Width = [12,17[
Gradient = [0,0.36[
Band1 = [27,58[
Then
Channel Pattern = Run 2.2%
Channel Pattern = Riffle 66.7%
Channel Pattern = Flat 2.2%
Channel Pattern = Rapid 28.9%

Rule 7 If
Width = [12,17[
Gradient = [0,0.36[
Band1 = [58,111]
Then
Channel Pattern = Rapid 100.0%

Rule 8 If
Width = [12,17[
Gradient = [0.36,0.45[
Then
Channel Pattern = Run 38.0%
Channel Pattern = Riffle 27.5%
Channel Pattern = Steady 34.5%

Rule 9 If
Width = [12,17[
Gradient = [0.45,3.68]
Then
Channel Pattern = Run 11.6%
Channel Pattern = Riffle 88.4%

Rule 10 If
Width = [17,27[
Gradient = [0,0.36[
Then
Channel Pattern = Riffle 30.1%
Channel Pattern = Steady 27.1%
Channel Pattern = Rapid 42.9%

Table B2: Decision Rules for Channel Pattern Classification (cont.)

Rule 11 If
Width = [17,27[
Gradient = [0.36,0.45[
Then
Channel Pattern = Run 49.8%
Channel Pattern = Riffle 46.3%
Channel Pattern = Steady 4.0%

Rule 12 If
Width = [17,27[
Gradient = [0.45,3.68]
Pc3 = [-41.56,0.73[
Then
Channel Pattern = Run 40.0%
Channel Pattern = Riffle 55.6%
Channel Pattern = Rapid 4.4%

Rule 13 If
Width = [17,27[
Gradient = [0.45,3.68]
Pc3 = [0.73,6.82[
Then
Channel Pattern = Run 56.9%
Channel Pattern = Riffle 17.6%
Channel Pattern = Rapid 25.5%

Rule 14 If
Width = [17,27[
Gradient = [0.45,3.68]
Pc3 = [6.82,51.68]
Then
Channel Pattern = Run 30.4%
Channel Pattern = Riffle 3.6%
Channel Pattern = Rapid 66.1%

Rule 15 If
Width = [27,60]
Gradient = [0,0.36[
Then
Channel Pattern = Riffle 47.1%
Channel Pattern = Steady 52.9%

Rule 16 If
Width = [27,60]
Gradient = [0.36,0.45[
Band1 = [1,9[
Then
Channel Pattern = Riffle 7.5%
Channel Pattern = Steady 62.5%
Channel Pattern = Flat 30.0%

Rule 17 If
Width = [27,60]
Gradient = [0.36,0.45[
Band1 = [9,58[
Then
Channel Pattern = Run 10.5%
Channel Pattern = Riffle 7.6%
Channel Pattern = Steady 12.9%
Channel Pattern = Flat 69.0%

Rule 18 If
Width = [27,60]
Gradient = [0.45,3.68]
Then
Channel Pattern = Run 96.4%
Channel Pattern = Rapid 3.6%

Table B3: Decision Rules for Land Cover Classification

Rule1 If
- Band1 = [0,14[
- Pc2 = [-158.86,-40.08[

Then
- Land Cover = Coniferous 90.5%
- Land Cover = Water 9.5%

Rule2 If
- Band1 = [0,14[
- Pc2 = [-40.08,126.48]

Then
- Land Cover = Coniferous 1.3%
- Land Cover = Water 98.7%

Rule3 If
- Band1 = [14,31[
- Pc2 = [-158.86,-8.58[

Then
- Land Cover = Coniferous 95.3%
- Land Cover = Shrub 4.7%

Rule4 If
- Band1 = [14,31[
- Pc2 = [-8.58,61.14[

Then
- Land Cover = Water 100.0%

Rule5 If
- Band1 = [31,48[
- Pc2 = [-158.86,-102.37[

Then
- Land Cover = Coniferous 51.4%
- Land Cover = Shrub 48.6%

Rule6 If
- Band1 = [31,48[
- Pc2 = [-102.37,-79.04[

Then
- Land Cover = Coniferous 34.5%
- Land Cover = Shrub 32.8%
- Land Cover = Alder 32.8%

Rule7 If
- Band1 = [31,48[
- Pc2 = [-79.04,-40.08[
- Band3 = [23,39[

Then
- Land Cover = Coniferous 73.3%
- Land Cover = Shrub 26.7%

Rule8 If
- Band1 = [31,48[
- Pc2 = [-79.04,-40.08[
- Band3 = [39,86[

Then
- Land Cover = Coniferous 9.2%
- Land Cover = Shrub 17.3%
- Land Cover = Alder 73.5%

Rule9 If
- Band1 = [31,48[
- Pc2 = [-79.04,-40.08[
- Band3 = [86,149[

Then
- Land Cover = Coniferous 23.1%
- Land Cover = Shrub 61.5%
- Land Cover = Alder 15.4%

Rule10 If
- Band1 = [31,48[
- Pc2 = [-40.08,-8.58[

Then
- Land Cover = Coniferous 6.4%
- Land Cover = Shrub 44.7%
- Land Cover = Alder 17.0%
- Land Cover = Wetland 31.9%

Rule11 If
- Band1 = [48,58[
- Pc4 = [-27.42,-1.14[

Then
- Land Cover = Coniferous 5.7%
- Land Cover = Shrub 76.2%
- Land Cover = Alder 13.1%
- Land Cover = Wetland 4.9%

Table B3: Decision Rules for Land Cover Classification (cont.)

Rule12 If
- Band1 = [48,58[
- Pc4 = [-1.14,13.11]

Then
- Land Cover = Coniferous 0.6%
- Land Cover = Shrub 15.2%
- Land Cover = Alder 81.3%
- Land Cover = Wetland 2.9%

Rule13 If
- Band1 = [58,72[
- Ndvi = [-0.051,0.065[

Then
- Land Cover = Noveg 100.0%

Rule14 If
- Band1 = [58,72[
- Ndvi = [0.065,0.213[

Then
- Land Cover = Shrub 75.0%
- Land Cover = Noveg 25.0%

Rule15 If
- Band1 = [58,72[
- Ndvi = [0.213,0.312[
- Pc2 = [-79.04,-58.21[

Then
- Land Cover = Shrub 18.2%
- Land Cover = Wetland 81.8%

Rule16 If
- Band1 = [58,72[
- Ndvi = [0.213,0.312[
- Pc2 = [-58.21,-8.58[

Then
- Land Cover = Shrub 53.1%
- Land Cover = Alder 9.4%
- Land Cover = Wetland 12.5%
- Land Cover = Noveg 25.0%

Rule17 If
- Band1 = [58,72[
- Ndvi = [0.312,0.402[
- Pc4 = [-27.42,1.7[

Then
- Land Cover = Shrub 72.7%
- Land Cover = Alder 10.9%
- Land Cover = Wetland 12.7%
- Land Cover = Noveg 3.6%

Rule18 If
- Band1 = [58,72[
- Ndvi = [0.312,0.402[
- Pc4 = [1.7,13.11]

Then
- Land Cover = Shrub 7.5%
- Land Cover = Alder 90.0%
- Land Cover = Wetland 2.5%

Rule19 If
- Band1 = [58,72[
- Ndvi = [0.402,0.669[

Then
- Land Cover = Shrub 87.7%
- Land Cover = Alder 8.5%
- Land Cover = Wetland 0.8%
- Land Cover = Noveg 3.1%

Rule20 If
- Band1 = [72,103[
- Band3 = [39,68[

Then
- Land Cover = Water 100.0%

Rule21 If
- Band1 = [72,103[
- Band3 = [68,115[

Then
- Land Cover = Shrub 58.1%
- Land Cover = Alder 4.7%
- Land Cover = Wetland 11.6%
- Land Cover = Noveg 25.6%

Rule22 If
- Band1 = [72,103[
- Band3 = [115,176[
- Ndvi = [-0.244,0.213[

Then
- Land Cover = Shrub 3.8%
- Land Cover = Wetland 8.8%
- Land Cover = Noveg 87.5%

Table B3: Decision Rules for Land Cover Classification (cont.)

Rule23 If
Band1 = [72.103[
Band3 = [115.176[
Ndvi = [0.213.0.402[
Then
Land Cover = Shrub 15.0%
Land Cover = Wetland 70.6%
Land Cover = Noveg 14.4%

Rule24 If
Band1 = [72.103[
Band3 = [176.244]
Then
Land Cover = Shrub 11.1%
Land Cover = Noveg 88.9%

Rule25 If
Band1 = [103.153[
Band4 = [2.221[
Then
Land Cover = Noveg 100.0%

Rule26 If
Band1 = [103.153[
Band4 = [221.249]
Then
Land Cover = Wetland 80.2%
Land Cover = Noveg 19.8%

Rule27 If
Band1 = [153.223]
Then
Land Cover = Noveg 100.0%