MAP REALIZATIONS OF CATEGORICAL VARIABLES: ILLUSTRATIONS FROM LAKE AND TILL SURVEYS IN NEWFOUNDLAND AND LABRADOR

S.D. Amor and S. Duquet[1]
Geochemistry, Geophysics and Terrain Sciences Section
[1]Geosciences Data Management Section

ABSTRACT

Categorical or nominal field observations are an important component of many field geochemical programs, and they are also the final result of multivariate statistical methods, such as discriminant analysis and cluster analysis. Displaying such observations in map form presents special challenges.

A method has been developed in ArcGIS 10.2 to create a smoothed map of categorical variables based on a 'rolling-mode' principle. This has been applied to observations made by samplers during the collection of lake-sediment and water samples during the National Geochemical Reconnaissance (NGR) program, and subsequent programs carried out in Newfoundland and Labrador, as well as to the results of cluster analysis on Labrador lake samples, and till samples from the Burin Peninsula.

The subjectivity of some field observations is illustrated by sharp discontinuities in lake-sediment colours, and lakewater turbidity, between NTS map areas in Labrador: an apparent consequence of different areas being sampled by different crews.

Some field observations display significant regional control, and locally, close relationships to the underlying geology. Examples include the colour of sediments collected over carbonate rocks in southwestern and northwestern Newfoundland. The effects of forest fires are also apparent although in some cases, where the fieldwork was done more than 30 years ago, this is probably no longer the case.

The smoothing method has advantages in creating a 'big-picture' view of regional categorical variables and compensating for variable sample density, but has the undesirable effect of obscuring thin linear features including those described above.

INTRODUCTION

During the collection of geochemical samples it is normally recommended to record certain field observations in categorical or nominal form, usually (and most conveniently) in the form of numeric or alphanumeric codes (*e.g.*, Hoffman, 1986). This work takes up a significant proportion of the total time taken to complete a sampling program. With the exception of those that might be related to gross variations in the composition of the sediment, such as anthropogenic contamination or significant shoreline mineralization, these observations are rarely put to any use after the completion of the analytical program, although attempts to render them in symbolic form have been made (*e.g.*, Amor, 2013a-c). One reason for this is that such observa-

tions are considered to be 'noisy'; *i.e.*, the colour or texture of a lake-sediment at one site bears little or no relationship to that of a sample collected at another lake only one or two kilometres away.

Such inhomogeneity is also often observed in the results of statistical classification methods such as cluster or discriminant analysis. The classes to which geochemical samples are assigned by these methods, based on chemical analyses, are also categorical variables.

Even when numbers are used as identifiers of such parameters as sample colours or cluster allocations (usually for more economical computer storage) there is no implication that 'Class 2' has greater magnitude than 'Class 1'

(Davis, 1973, page 8). Consequently, it is not meaningful to subject such observations to mathematical manipulation.

The concept of 'smoothing' regionalized variables is widely known and even more widely applied, whether its purpose is to generate a contour map, or to separate 'signal' (in a geochemical context, case, a large-scale regional trend) from 'noise' (small-scale, local deviations from the trend). Smoothing methods normally involve the calculation of some form of weighted or unweighted average and the method has been termed 'moving average' or 'rolling mean' (*e.g.*, Howarth, 1983, pages 142-150). Local deviations from these regional trends may prove to be indicative of anomalous behaviour and as such, of interest in their own right. The samples that are used to calculate a moving average may be defined variously as all those that fall within a fixed radius of each sample site (in which case, the smoothed value is plotted at the location of the sample itself), or those that fall within a rectilinear, usually rectangular cell, which is then coded by colour or shading according to the smoothed value (*e.g.*, Howarth and Garrett, 2010). Most contouring algorithms also make use of a moving average, where a grid of regularly spaced nodes is laid out over the extent of the map area and an average value of its neighbouring samples is assigned to each node. The mean is normally weighted by the reciprocal of the distance between each sample and the node, or some power of it (Howarth, *op. cit.*, pages 136-138). Formerly, contour lines would then be threaded between nodes of constant value and drawn with a computer-driven pen plotter, but with the computer hardware and software facilities currently available, the nodes are more commonly represented by colour-coded pixels (often, with 'cool' colours assigned to low values and 'warm' colours assigned to high values) or assigned greyscale shades according to the magnitude of the averaged values.

Use of the moving average is dependent on the variable to be smoothed being interval- or ratio-scale (Davis, 1973, page 8); otherwise, the variable is not amenable to the calculation of a mean or weighted mean and, as stated above, cannot be applied to categorical variables. It is, however, possible to assign to each point, or cell, the characteristic that is predominant within the specified radius of the sample point, or within the cell. This has been termed the 'moving majority vote' or 'rolling mode' method, although the source of these terms is unknown. This method is applied to categorical variables in the current study.

DATA PROCESSING METHODOLOGY

The 'Fishnet' tool in ArcGIS is suitable for creating an overlay grid of rectangular cells and the procedure is as fol-

lows (it is advisable to disable the 'Background Processing' in 'Geoprocessing → Geoprocessing options'):

1. The 'Create Fishnet' tool is selected from 'Data Management Tools → Feature Class'.
2. The Output Feature Class is named and the location of the file specified.
3. The extent of the fishnet grid is specified. The default option consists of the northern, southern, eastern and western extremities of a specified feature class, but the user can also specify these bounds. The feature class should be selected from a pull-down menu and not dragged-and-dropped from the Table of Contents pane.
4. The cell size is specified, in coordinate units. The width and height of the cells need not be the same although this is the most common specification.
5. The Geometry Type parameter is set at 'POLYGON'.

With the grid established, the rolling modes are assigned to each cell, as follows:

1. The 'Spatial Join' tool is selected from 'Analysis Tools → Overlay'.
2. The fishnet created during the first step is selected as the Target Feature.
3. The output feature class for the rolling modes is named, and the location of the file specified.
4. The Join operation is specified as 'JOIN ONE TO MANY'.
5. The 'Keep All Target Features (optional)' box is unchecked.
6. The input feature class is specified, with the Merge Rule set to 'Mode' (other options exist, including mean and median; as stated above, these would not be suitable when smoothing categorical variables).
7. The Match Option is set at 'COMPLETELY_CONTAINS'.

Although the spatial join process automatically eliminates any cells that contain zero sample points, it may also be advisable to screen out cells in the fishnet that contain less than a minimum number that is greater than zero, because the mode of less than three values is essentially meaningless; this minimum number has been set at 3 in the current study. The process of creating point counts in each cell is as follows:

1. The 'Spatial Join' tool is selected from 'Analysis Tools" → Overlay.
2. The fishnet created during the previous step is selected as the Target Feature.
3. The output feature class for the point counts is named and the location of the file specified.
4. The Join operation is specified as 'JOIN ONE TO ONE' (this is at variance with the 'JOIN ONE TO MANY' option chosen when estimating modes, above).
5. No Join Features are specified.

6. The Match Option is set at 'COMPLETELY_CON-TAINS'.
7. The resulting feature class is merged with the fishnet feature class and the resulting file is queried on the basis of the 'Join_Count' attribute. This enables the elimination of those cells that contain less than three points.

As a result of running the spatial join, modal values of the categorical variables can then be assigned to each of the qualifying cells in the fishnet, and colour-coded appropriately.

Application of the method will be demonstrated with observational field data from three lake-sediment datasets in Newfoundland and Labrador, and multivariate cluster allocations of lake-sediment data from Labrador, and till data from the Burin Peninsula in Newfoundland. In most cases, maps of smoothed data will be compared with those with the same colour codes of the samples themselves, so that the usefulness of the rolling-mode smoothing method may be evaluated.

NGR SAMPLES, LABRADOR

The categorical observations recorded during the NGR program in Labrador (Friske and Hornbrook, 1991) consist of the following:
- Local relief (low, medium, high)
- Contamination (Camp, Camp Fuel, Camp Gossan, Fuel, Gossan, None, Work, Work Camp, Work Camp Fuel, Work Camp Fuel Gossan, Work Fuel)
- Sediment colour (brown, green, grey, black, tan, yellow), and
- The amount of suspended material in the water (heavy, light, none).

In addition, an estimate was made of the depth from which the sample was recovered; this can be taken as a measurement (though not a precise one) of the lake's maximum depth, and will be examined in conjunction with the observed local relief. Contamination is not considered in this study, because it is absent over most of the map area.

Creation of a 5 km fishnet resulted in a significant number of cells containing less than three sample points (Figure 1); consequently the areal distribution of local relief and sample colour is studied with respect to a 10 km fishnet, in which almost all of the occupied cells contain at least three points (Figure 2).

SUSPENDED MATERIAL IN LAKE WATER

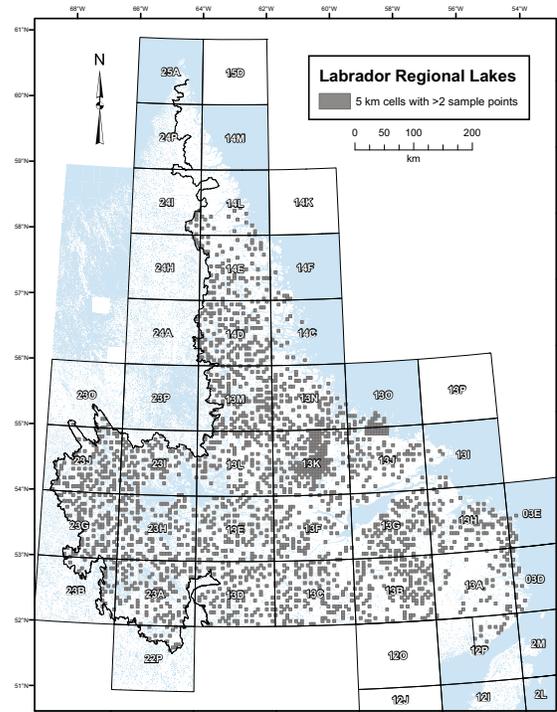Suspended material is recorded as 'none', 'light' or

Figure 1. *5 km cells containing more than 2 sample points. The large areas where there are no such cells militate against such a small cell size in smoothing the categorical data.*
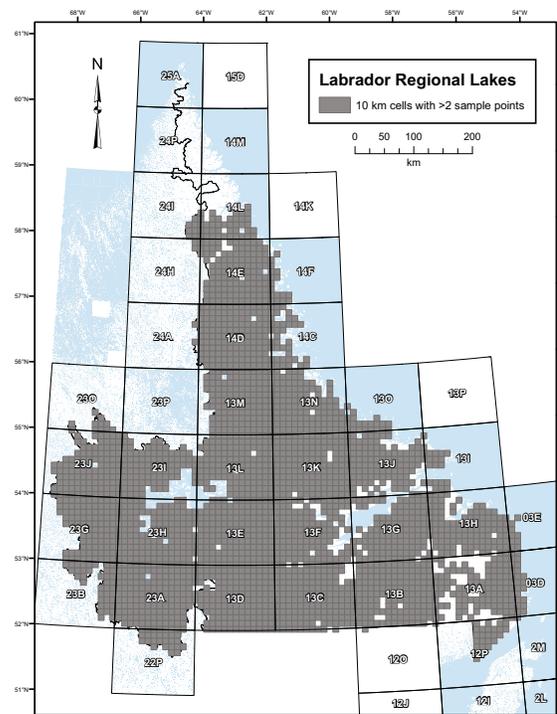

Figure 2. *10 km cells containing more than 2 sample points. Most cells of this size meet the minimal criterion for estimation of the rolling mode.*
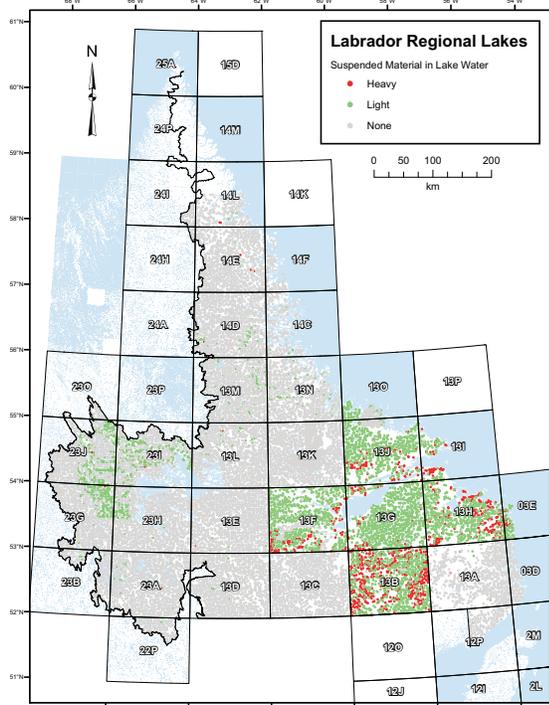
199

Figure 3A. *Suspended material in lake sediment, as described by samplers. The sharp contrasts between certain NTS sheets reflect an element of subjectivity in this observation.*

Figure 3B. *Suspended material in lake sediment. Reducing the description to two settings ('High' and 'Low') does not alleviate the subjectivity.*

'heavy'. Samples coded by these descriptions are shown in Figure 3A. Sharp boundaries separate NTS map areas 13B, 13F, 13G, 13H and 13J, where suspended material in the lake waters is invariably described as 'light' or 'heavy', from most of the rest of the survey area, where suspended material is recorded as 'none' at almost every site. This clearly reflects a strong element of subjectivity in describing the amount of suspended material, whereby the amount of suspended material in the same sample would have been described as 'none' by one observer, and 'light' by another, and another sample would have been described as 'light' and 'heavy' by the same two observers. An attempt has been made to compensate for this subjectivity by reclassifying the suspended material merely as 'high' or 'low'. The former description is applied to sites in the above five NTS map areas described as 'heavy', and to points in every other map area that were described as 'light'; the latter is applied to samples from the five map areas described as 'light', and to samples described as 'none'. The results of this classification are shown in Figure 3B and with the possible exception of some coastal areas of southeastern Labrador, where the turbidity of the lake waters might be attributable to recent marine incursion (NTS 3/E, 13H, I and O), the areas described as 'high' still reflect some subjectivity. This is particularly noticeable in NTS map area 13B, and also in western Labrador, where the boundaries of the area characterized
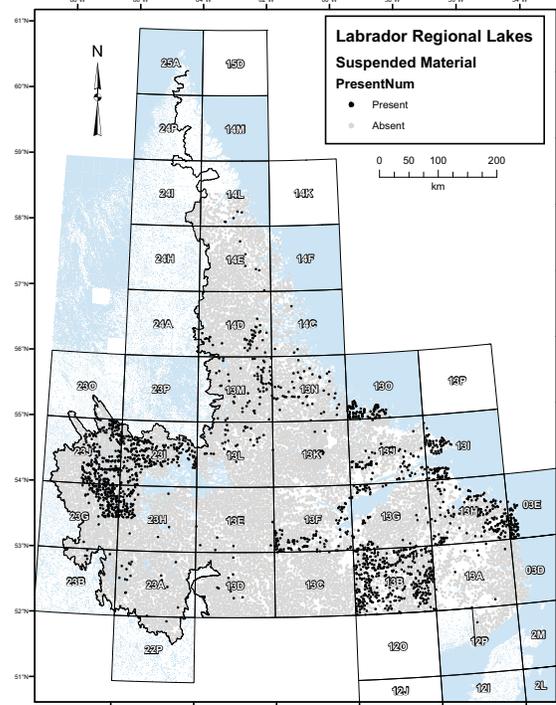
by 'high' suspended material is rectilinear, suggesting again that different individuals were responsible for describing the water samples on either side of it. In conclusion, the estimates of suspended material in the lake waters are of little value.

LOCAL RELIEF

Local relief was described by the samplers as either 'low', 'moderate' or 'high'. Both the point and rolling-mode plots (Figure 4A, B) show a number of conspicuous regional features. A linear feature, extending about 330 km from the Québec border near Lac Veron in the southwest of NTS map area 22P, to the valley of the Red Wine River in the northeast corner of NTS 13E, is defined by relief described as 'high'. The most southwesterly 100 km of this lineament corresponds to the contact between the troctolitic/anorthositic Atikomak River Massif, to the southeast, and granitic rocks to the northwest (Nunn *et al.*, 1986). Over the rest of its length the lineament crosscuts a variety of rock units and structures, and its northwestern boundary corresponds rather closely to the thrust faults that separate the Churchill Falls terrane, to the northwest, from the Lake Joseph and Wilson Lake terranes to the southeast (Gower and Krogh, 2002). The direction and extent of the linear are also defined by sampling depth (Figure 5).
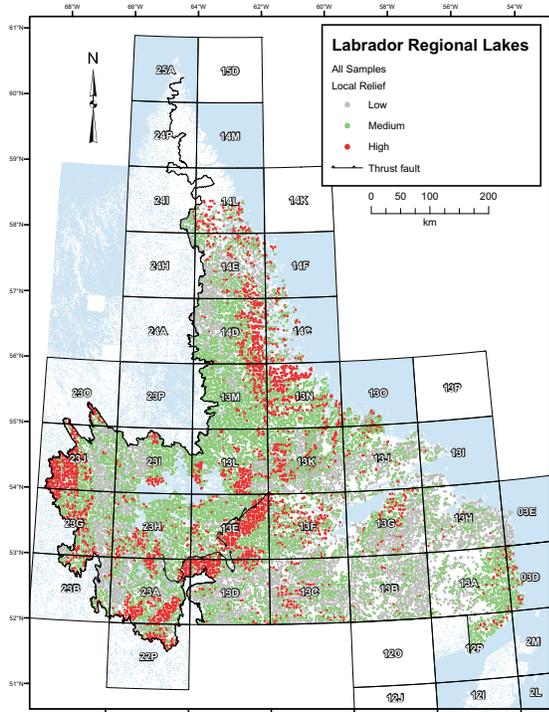
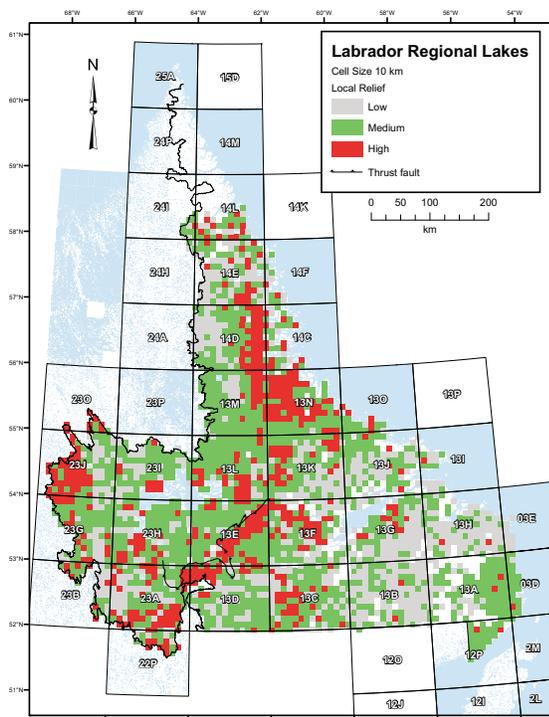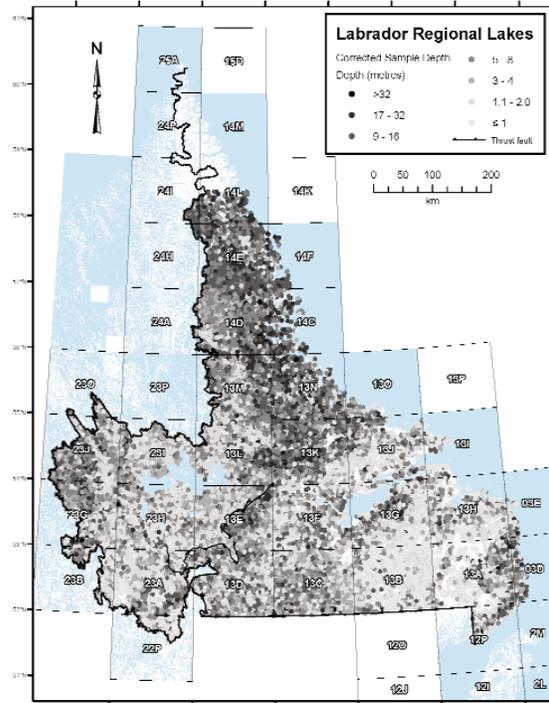**Figure 4A.** *Local relief, as described by lake samplers (point plot).*

Figure 5. *Depth of recovery of lake sediment, Labrador. The linear feature, defined by relatively deep lakes, extending from southwestern to central Labrador corresponds to a similar feature defined by high relief.*

Elsewhere, lakes located in areas of high relief define a zone extending from the north of NTS map area 13M/08 to the southwest of NTS 14E/08. This linearity is also associated with deep (greater than 9 m) samples, although not as precisely as the one described above. The southern termination of the linear corresponds to the northern contact of the Harp Lake Intrusive Suite; it crosses numerous geological contacts but is roughly parallel to the foliation of the gneisses to the west, and the series of alkaline intrusions that extends from the Voisey Bay–Notakwanon batholith in the south (NTS 13M) to the Umiakovik batholith in the north (NTS 14E).

Relief is also high over most of NTS map area 13N; that is, in the hinterland of Hopedale and Natuashish.

The high relief reported in western Labrador (over Archean rocks of the Ashuanipi Complex) may also be a feature of inconsistently applied criteria, as a zone of high relief, bounded to the west by the Labrador–Québec border, is bounded sharply to the east by a north−south line. This line is obscured by the smoothing process (Figure 4B); an undesirable effect of the latter.

In conclusion, observations of local relief indicate major linear features, some of which may be related to geo-
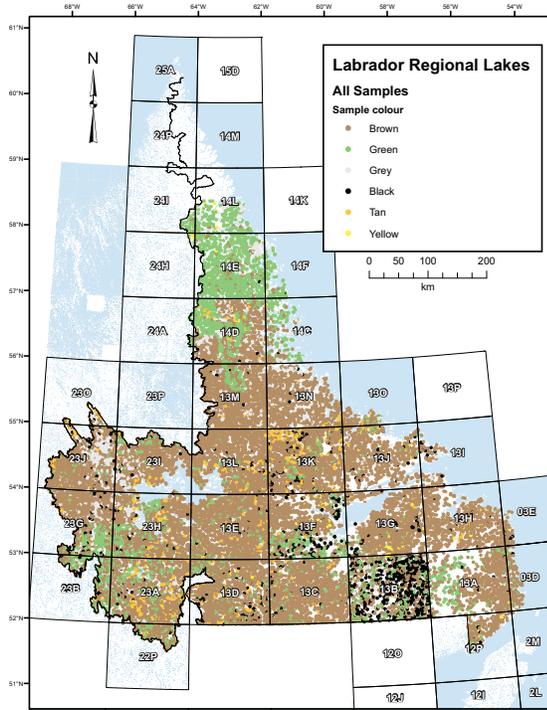

Figure 4B. *Local relief, as described by lake samplers (rolling mode, cell size 10 km).*

Figure 6A. *Lake-sediment colour recorded by samplers (point plot). An element of subjectivity is evident in the sharp boundaries between certain zones, which correspond in many cases to NTS map area boundaries.*
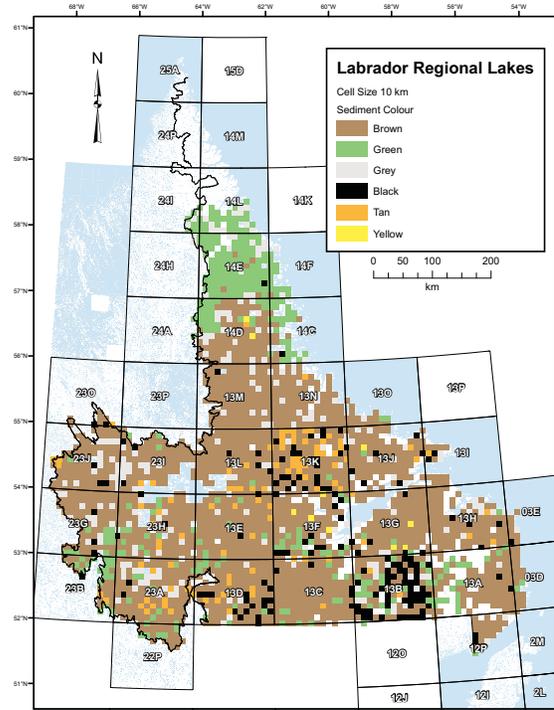
logical structures. The inconsistent definition of what constitutes 'high' or 'low' relief does, however, interfere to some extent with the overall visual impact of the map.

SAMPLE COLOUR

Lack of consistency between sample crews in the colour of the lake sediments (which, in common with the other lake-sediment surveys considered in this paper, was recorded without the aid of Munsell colour charts) is also apparent, with a significant concentration of samples recorded as 'black' sharply terminated by the boundary between NTS map area 13B and its neighbours (Figure 6A). As with the relief, this sharp boundary is made less clear by the rolling-mode smoothing process (Figure 6B). There is also some evidence, in the form of the loss-on-ignition content (normally, a first approximation of the organic content) of the samples (Geological Survey of Newfoundland and Labrador, 2015a; Figure 7) of an element of subjectivity in the almost universal description of samples from the north of the sampled area (primarily, in NTS 14E) as 'green' (normally indicative of a significant component of organic ooze). The subjectivity of colour descriptions may also be the reason for the reported presence of tan samples in the north of NTS 13K and green samples in the south of NTS 13D. However, the concentrations of green samples in
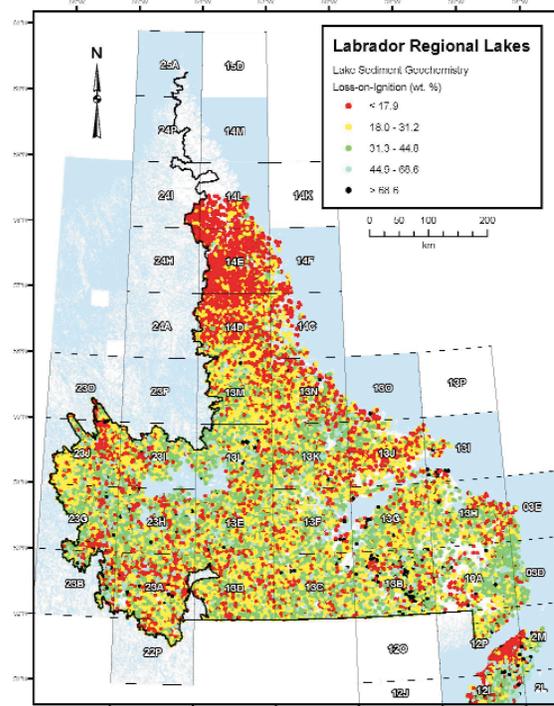

Figure 6B. *Lake-sediment colour recorded by samplers (rolling mode, cell size 10 km).*

Figure 7. *Loss-on-Ignition (LOI) in Labrador lake sediments. Class boundaries derived by Jenks' optimization.*
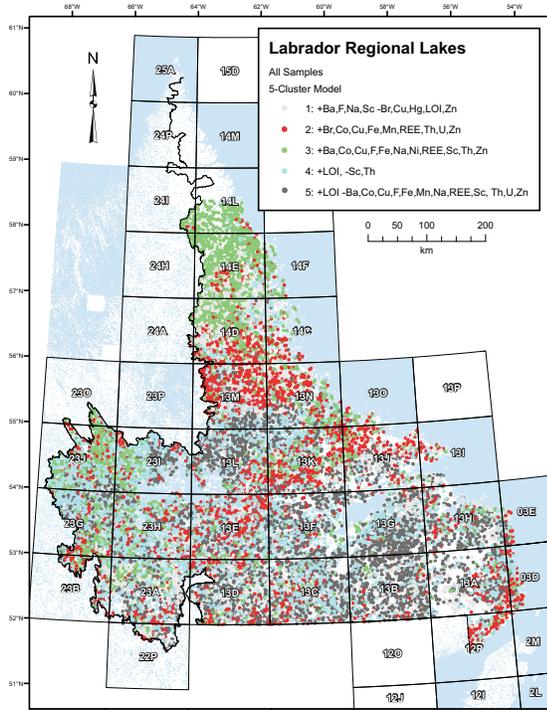
Figure 8A. *5-cluster allocations, Labrador lake sediments (point plot).*

Figure 8B. *5-cluster allocations, Labrador lake sediments (rolling mode, cell size 10 km).*

southwestern Labrador (NTS 23B, 23F and 23G) and grey samples in NTS 23I, over Labrador Trough rocks (Wardle *et al*., 1997), may be genuine; their convoluted boundaries do not suggest operator bias, and the disposition of the grey samples (Figure 6A) even shows a loose concordance with the regional strike, although this relationship is much less obvious when the data are smoothed (Figure 6B).

CLUSTER ANALYSIS

Figure 8A and B display the results of k-means clustering (Howarth and Sinding-Larsen, 1983, pages 218-220; Sinding-Larsen, 1975) of multi-element lake-sediment data; this is primarily for demonstration purposes and interpretation of the results, although mandatory in any full evaluation of the results, is kept to a minimum. The input variables consist of Ba, Br, Ce, Co, Cu, F, Fe, Hg, La, LOI, Mn, Na, Ni, Sc, Sm, Th, U and Zn, transformed so that their frequency distribution approximates a normal distribution, and normalized so that each has a mean of zero and a standard deviation of 1.0. Other elements were omitted because too many analyses were exceeded by the analytical detection limit; and for simplicity, the additional ICP-OES analyses of the same samples (McConnell and Finch, 2012) were also omitted.

Cluster analysis is an 'unsupervised' classification method in which samples (in a geochemical context) are grouped together as poin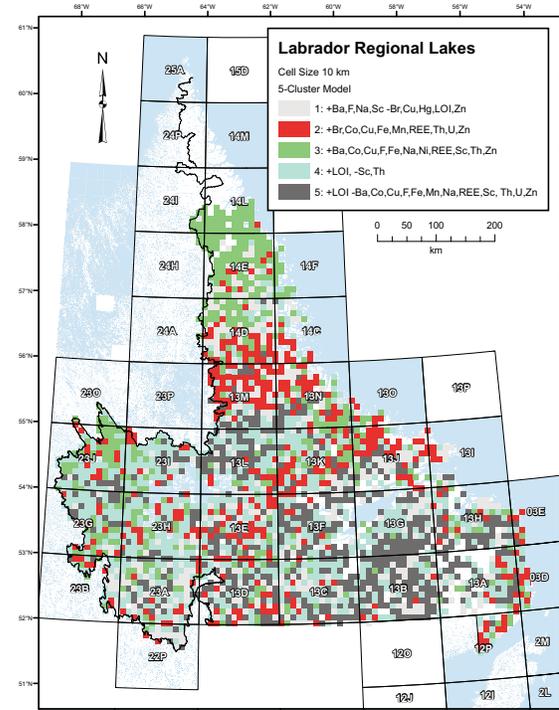ts in multi-element space, based on the points' proximity to a limited number of group centroids. Other than the number of centroids, no information is available, or provided, regarding group affinities. In this respect, cluster analysis differs from discriminant analysis, where *a priori* information regarding group assignments is available for certain samples and used to create a multi-element model that is used subsequently to classify other, unknown samples. Such classifications would also be amenable to the smoothing process but are not considered in the current study.

As well as being distinct in 18-element space, despite the apparent similarity of their element associations, the clusters identified in the 5-cluster model show a fair degree of spatial co-association, with samples assigned to Cluster 3 (the most strongly depleted in LOI) showing a concentration in the north (NTS map areas 14C, D, E, F and L), samples assigned to Cluster 2 (enriched in rare-earth elements) indicating the dispersion trains from the mineralized occurrences in the Red Wine Mountains (NTS 13E, L and K) and Mistastin Lake area (NTS 13M, N, 14D, E) and in southeastern Labrador (NT 12P, 13A), and samples assigned to Clusters 4 and 5, both LOI-enriched, which concentrate in the Mealy Mountains (NTS 13B, G).

Cluster assignments for a 10-cluster model are shown in Figure 9A and B, and the relationship between the 5- and 10-cluster models is shown as a confusion matrix in Table 1.
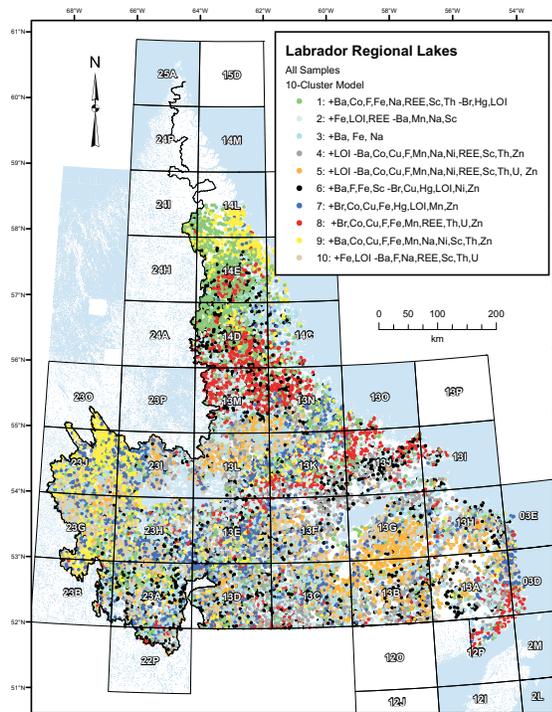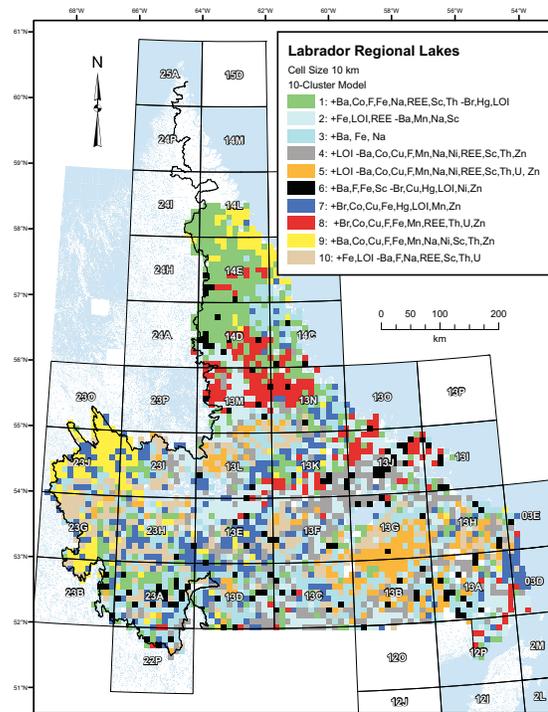
Figure 9A. *10-cluster allocations (point plot).*

Figure 9B. *10-cluster allocations (rolling mode, cell size 10 km).*

Table 1. Confusion matrix (5- and 10-cluster models)

Samples from Cluster 5/1 are mainly assigned to Clusters 10/6, 10/3 and 10/1

Samples from Cluster 5/2 are mainly assigned to Clusters 10/8, 10/7, 10/9 and 10/2

Samples from Cluster 5/3 are mainly assigned to Clusters 10/1 and 10/9

Samples from Cluster 5/4 are mainly assigned to Clusters 10/10, 10/2 and 10/7

Samples from Cluster 5/5 are mainly assigned to Clusters 10/4 and 10/5

		10-Cluster Assignment										
		1	2	3	4	5	6	7	8	9	10	**Total**
5-Cluster Assignment	1	545	80	783	32	0	1381	0	12	0	0	2833
	2	0	357	66	0	0	0	803	1318	420	0	2964
	3	1490	4	235	0	0	0	1	163	967	0	2860
	4	0	1244	1099	55	0	0	1234	0	50	1543	5225
	5	0	40	24	1546	1216	0	0	0	0	867	3693
	Total	2035	1725	2207	1633	1216	1381	2038	1493	1437	2410	17575

The rolling-mode smoothing method shows distinct advantages over the point plot in highlighting areas where most of the samples are assigned to a single cluster. The area north of latitude 56°30' (NTS map areas 14C, D, E, F) is dominated by Cluster 1, which appears to represent samples having a very high clastic and low organic content; several of the elements showing enrichment (Ba, F, Na, Sc), and depletion (LOI, Br) in these samples are associated with clastic deposition (Amor, 2014). Nevertheless, an isolated exception is the presence of samples assigned to Cluster 8 (enriched in Br, Co, Cu, F, Fe, Mn, REE, Th, U and Zn), which indicates the dispersion train from the possible REE mineralization at North River (Amor, 2011). Larger accumulations of samples assigned to Cluster 8 are present in north-central Labrador: along the Québec border, in the west (NTS 13M) and in NTS map areas 13M, N, 14C, D in the east. Both overlie Mesoproterozoic alkaline intrusive rocks; the former coincides with the Mistastin batholith, and the latter, the Arc Lake and Voisey Bay−Notakwanon batholiths (Wardle, 1993). They are separated by samples assigned to Cluster 2 (enriched in Fe, LOI and REE; deleted in Ba, Mn, Na and Sc) overlying Paleoproterozoic and/or Archean gneiss and migmatite (*ibid*.).

Other significant spatial concentrations of samples assigned to the same cluster are present in the Mealy Mountains (NTS map areas 13B, G) where Cluster 5, of what are probably organic-rich samples (enriched in LOI, depleted in Ba, Ce, Co, Cu, F, La, Mn, Na, Ni, Sc, Sm, Th, U and Zn) is present; and in NTS 23J where an accumulation of samples assigned to Cluster 9 (enriched in Ba, Co, Cu, F, Fe, Mn, Na, Ni, Sc, Th and Zn) appears to be indicative of the underlying, mostly sedimentary, rocks of the Labrador Trough (Wardle *et al.*, 1997; Geological Survey of Newfoundland and Labrador, 2015b). Samples assigned to this cluster also concentrate along the coast in the north (NTS 14E, F, L) in an area underlain by Archean gneiss and migmatite (Ryan, 1990); the Kiglapait (mafic) Intrusion is situated at the southern extremity of this zone but does not seem to be a plausible source of the elements that characterize the cluster, bearing in mind local ice-movement directions (*e.g.*, Batterson, 1989).

Although the significance of the clusters identified in the lake-sediment data has not been fully investigated, the spatial co-association of samples assigned to the clusters suggests that they represent valid subdivisions of the data. Furthermore, the rolling-mode smoothing process has the overall effect of facilitating the viewing of the cluster allocations.

DETAILED LABRADOR SURVEYS

More than 13 000 lake-sediment samples have been collected during detailed surveys conducted at selected locations in Labrador between 1978 and 2014 (Amor, 2013a-c; Amor and Ricketts, 2015; McConnell, 2009; McConnell and Ricketts, 2008, 2010, 2011; McConnell *et al.*, 2007). The field observations made during these surveys differ from those recorded during the NGR work as they exclude local relief, and include the following:

- Shoreline vegetation (Forest, Rock and Forest, Mixed Forest and Bog, Bog, Barren >25% rock, Tundra <25% rock, Burned), and
- Sample composition (Clastic Fine Grained, Clastic Coarse grained, Organic Ooze, Organic granular, Organic peaty).

In common with the NGR surveys in Labrador, and in the interests of productivity, sample colour was recorded without the aid of Munsell colour charts, but more options were available: tan yellow, brown, brown–lustrous (formerly referred to as 'brown – jelly-like'), chocolate brown, greenish brown, green, grey, black, peaty, orange, white, red and grey brown.

SHORELINE VEGETATION

The most conspicuous features in Figure 10A (point plot) and 10B (rolling mode; cell size 6.25 km) are extensive concentrations of bog in southeastern Labrador (NTS map area 13A) and various burned areas (for example, in the north of NTS map area 13A, in the southwest of NTS 13H, and at several locations in NTS 13J, 13K and 13N). In southeastern Labrador there is a distinct contrast between mixed bog and forest, to the north in NTS 013H, and a broader variety of classifications to the south in NTS 013A. The two map areas were not sampled by different crews; therefore, the recorded difference in the vegetation cover between them is probably genuine. The samples in the dispersion train from the Strange Lake rare-earth element/rare-metal deposit (NTS 14E, 24A) were all collected in terrain described as 'tundra'; this is consistent with the latitude, and the observations of other workers in the area.

The 'smearing' effect of the rolling-mode calculations causes a number of apparently rather prominent features to be dissipated. Examples include areas of bog in western Labrador (NTS map area 23J, west of Menihek Lake, and NTS 13L, between Fraser and Sail lakes).

SAMPLE COMPOSITION

Figure 11A and B show how the physical composition of the lake-sediment samples varies in the sampled areas. Organic ooze is the predominant sediment type, but notable exceptions to this generalization are present in southeastern Labrador, where organic granular and organic peaty samples are in the majority in the west of NTS map area 13A. Organic granular samples are predominant in the south of NTS 23H, whereas organic peaty characterize coastal areas in the west of NTS 13I, the east of NTS 13J, and several small sampled areas in NTS 13B and 13G. Areas where clastic inorganic samples predominate are rare; this is to be expected, because such material tends to concentrate in active watercourses that are avoided, whenever possible, when seeking a sample site. However, a concentration of such clastic fine-grained samples is present in NTS 23H, whereas a number of smaller concentrations of clastic coarse-grained samples are present in southeastern Labrador (NTS 12P, 13A and 13H). These concentrations are more apparent in the rolling-mode map (Figure 10B) than in the sample point map (Figure 10A).

SAMPLE COLOUR

The most noticeable lake-sediment colour is chocolate brown (Figure 12A, B). This is predominant in all of east-
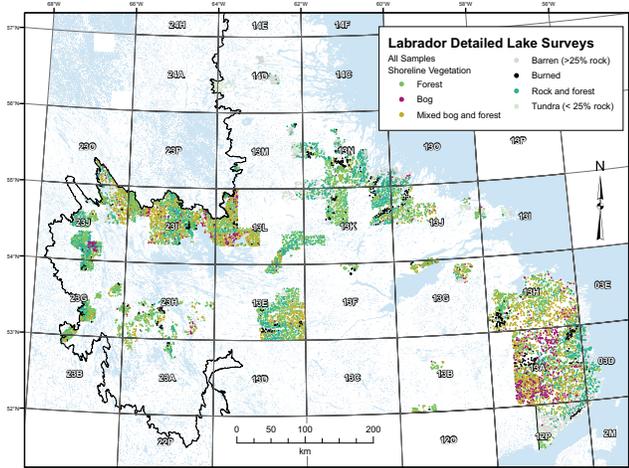
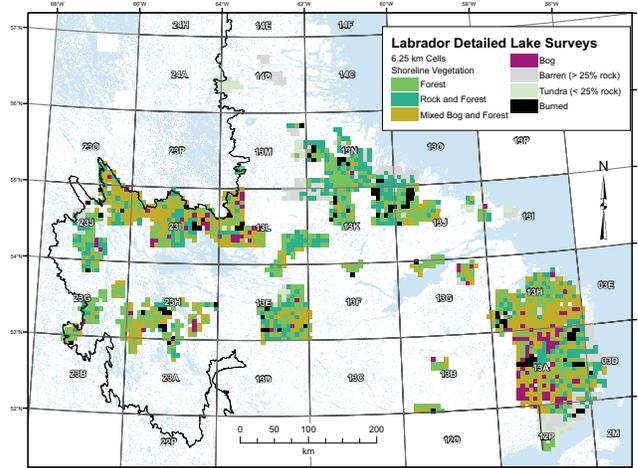**Figure 10A.** *Labrador detailed surveys – shoreline vegetation (point plot).*

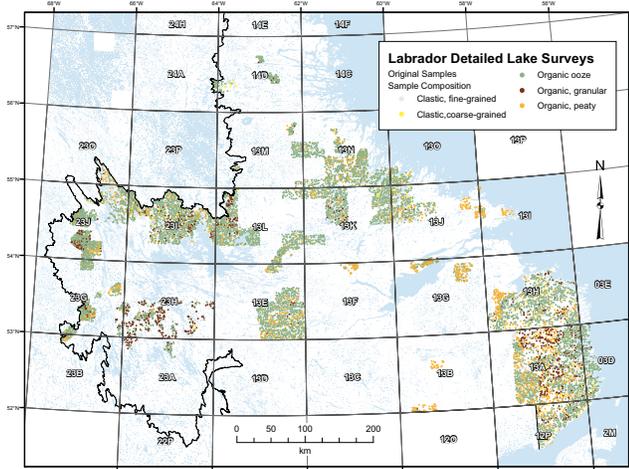**Figure 10B.** *Labrador detailed surveys – shoreline vegetation (rolling mode, cell size 6.25 km).*

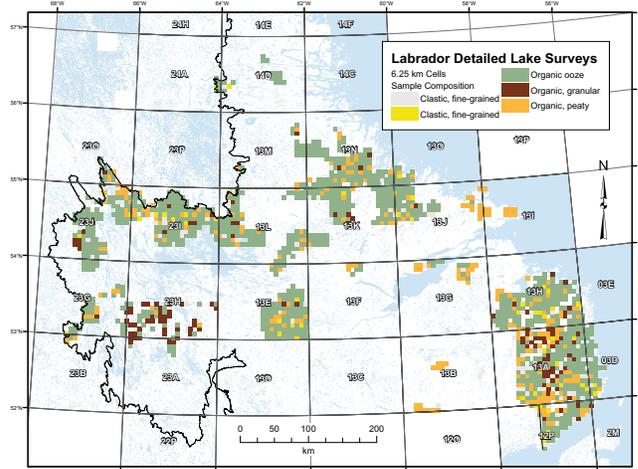**Figure 11A.** *Labrador detailed surveys – sample composition (point plot).*

Figure 11B. *Labrador detailed surveys – sample composition (rolling mode, cell size 6.25 km).*

ern, central and southern Labrador (although samples described merely as 'brown' characterize NTS map area 13H and part of 13A) as well as western Labrador over rocks of the Labrador Trough in NTS 23I, J and O. Areas characterized by other colours include that north of the Smallwood Reservoir in NTS 23I and 13L, where a variety of colours were reported including brown-lustrous, tan yellow, grey-green and green. A similar mix of colours is present in lakes over Superior Province rocks in NTS 23J, in the centre of NTS 13N northwest of Postville, and on the boundary between NTS 13E and 13L.

The colours do not correspond in any recognizable way with those observed during the NGR survey of Labrador, in which the great majority of samples, except in the north, were described as (plain) brown.

NEWFOUNDLAND

Although the density of sampling in Newfoundland, while variable (as well as the suite of elements analyzed), conformed more closely to the terms of reference of the NGR sampling program in Labrador, the field observations made during the sample collection follow the design of the detailed sampling programs in Labrador, described above.

SHORELINE VEGETATION

A number of large-scale controls on vegetation are present (Figure 13A, B). 'Barren (>25% rock)' terrain dominates a strip of land 50 km wide in the south of the Island, extending from the southwestern corner (NTS map area 11O/11) to the Burin Peninsula (NTS 1M/10) and resuming, in some-
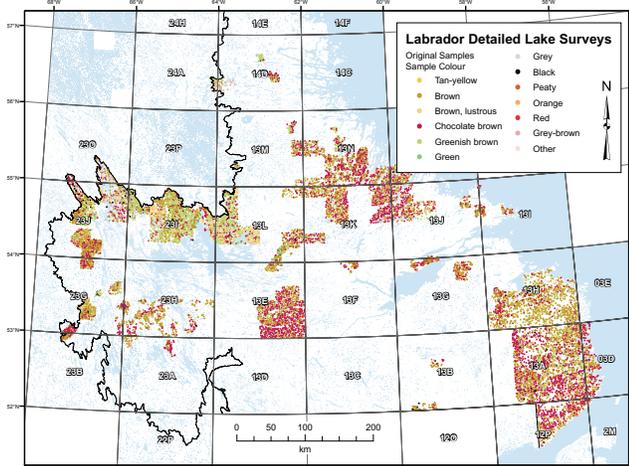
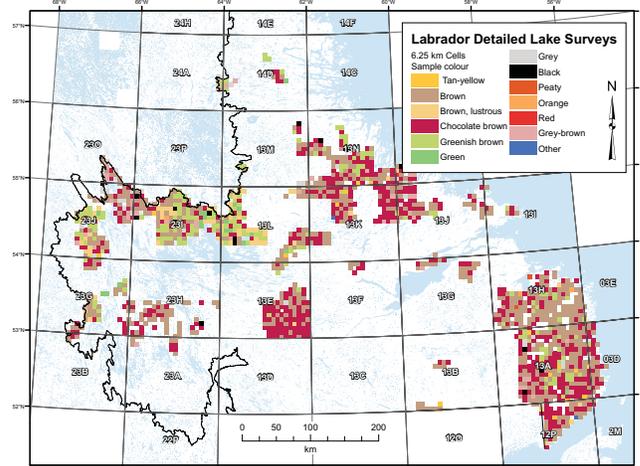**Figure 12A.** *Labrador detailed surveys – sample colour (point plot).*

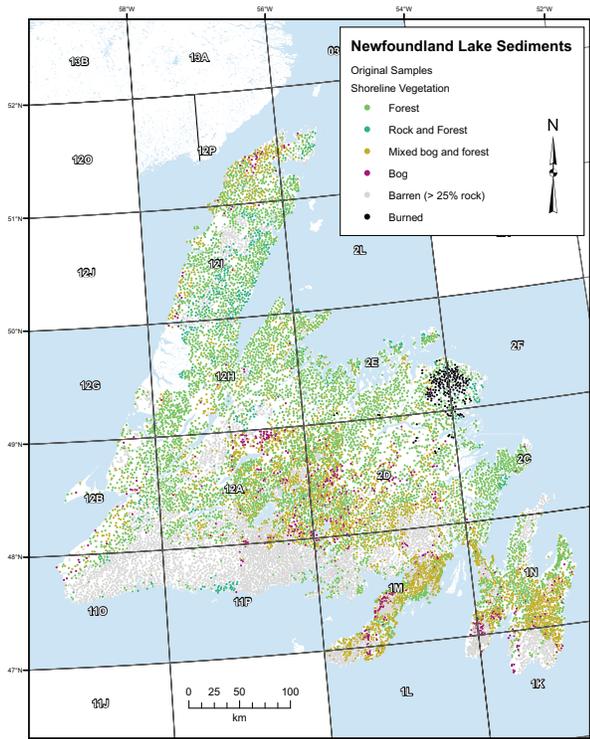**Figure 12B.** *Labrador detailed surveys – sample colour (rolling mode, cell size 6.25 km).*

Figure 13A. *Newfoundland regional survey – shoreline vegetation (point plot).*
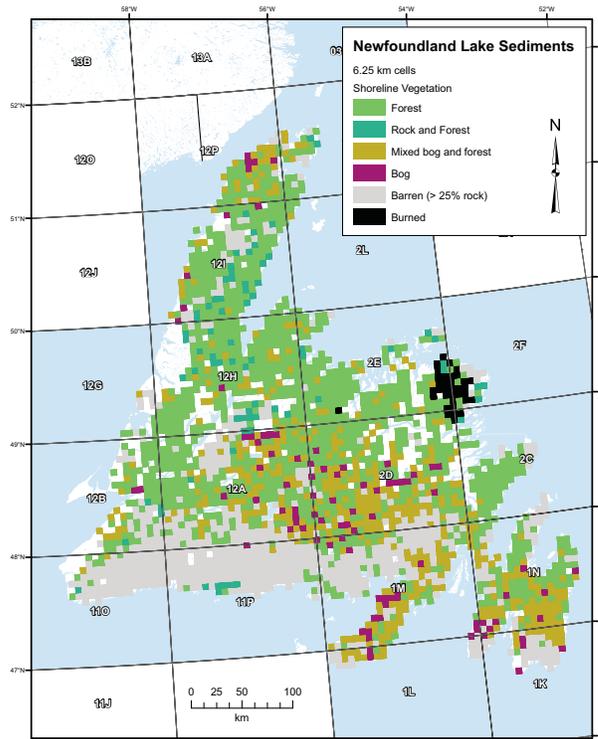

Figure 13B. *Newfoundland regional survey – shoreline vegetation (rolling mode, cell size 6.25 km).*

what attenuated form, in the south of the Avalon Peninsula (1K/11, 13, 14, 1L/16). Other barren areas are apparent between Red Indian and Grand lakes (NTS 12A/14, 12H/02 and 03) in the north of the Bay de Verde Peninsula (NTS 1N/14, 2C/02, 03) and along the spine of the Northern Peninsula (NTS 12I/06, 07).

A large burned area is apparent south of Musgrave Harbour (NTS 2E/01, 08, 2F/04, 05); this appears to be related to a large fire that burned there in 1961. As the fieldwork was done in 1980, the forest cover will probably have been restored (M. Lawlor, Forestry Service of Newfoundland and Labrador, personal communication, 2015).
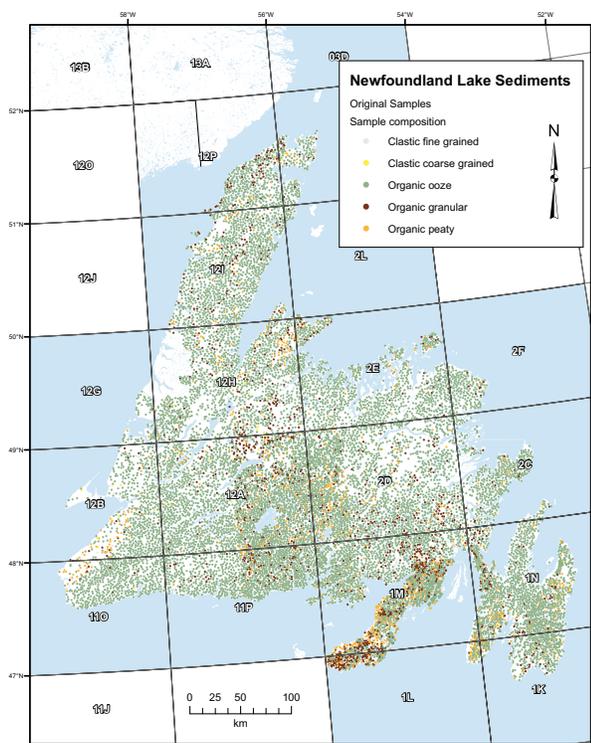
Figure 14A. *Newfoundland regional survey – sample composition (point plot).*

Figure 14B. *Newfoundland regional survey – sample composition (rolling mode, cell size 6.25 km).*

Boggy areas are apparent between Buchans and the Topsails (NTS 12A/15), in the Burin Peninsula (NTS 1M/03, 06) and in the St. Mary's Peninsula (NTS 1M/01, 01N/04). A large area dominated either by bog, or forest and bog, is also present in east-central Newfoundland (NTS 2D/2, 03, 04, 05, 06, 07, 11 and 12).

SAMPLE COMPOSITION

Organic ooze is the predominant sample type (Figure 14A, B). There are minor exceptions in NTS map areas 12A/15, 16, 12H/01 and 02, between Buchans and the Topsails, where the presence of organic granular sediments is apparent, and in the foot and northwest coast of the Burin Peninsula, where organic granular and organic peaty sediment are widespread. Organic peaty sediments are also conspicuous in a sharply defined zone in southwestern Newfoundland (NTS 11O/14, 15, 12B/02, 03, 07). The control on this pattern appears to be geological rather than topographical, as it encompasses both the Anguille Mountains and the Codroy Lowlands, and its boundary corresponds rather closely to the faulted boundary of the Mississippian and Pennsylvanian rocks of the St. George's sub-basin (Knight, 1983). The same area is also characterized by exceptionally shallow lakes (Figure 15).
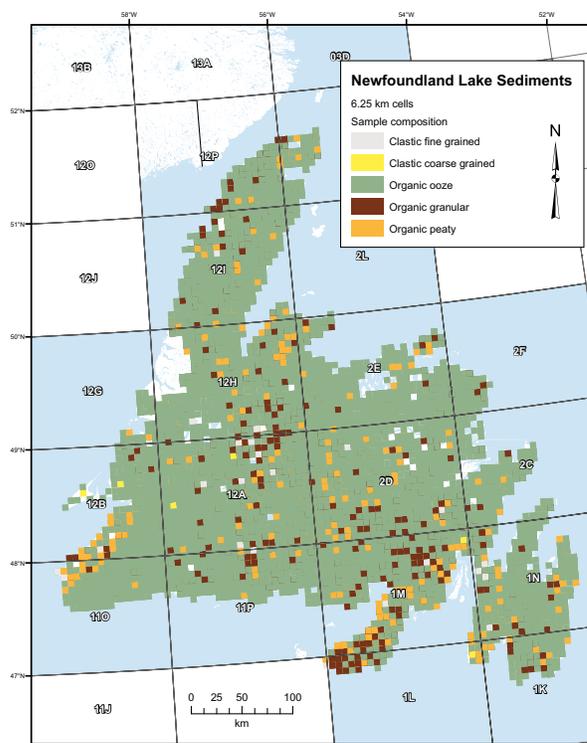
SAMPLE COLOUR

Brown or chocolate-brown sample material is predominant over most of Newfoundland (Figure 16A, B). A conspicuous exception is apparent in NTS map area 12H/05, where green lake sediment was reported at a number of localities. This area is underlain by Early to Middle Cambrian carbonate rocks, which are unlikely to have given rise to this colour. Even more prominent is an area of grey sediment in the north of the Northern Peninsula (NTS 12P/02, 07, 08 and 09), for which the evidence of geological control is strong; the area is underlain by Cambrian carbonate rocks and also displays unusual positive correlation between loss-on-ignition and certain elements like Cr (Figure 17A, B), with which it would be expected to be negatively correlated (*e.g.*, Amor, 2014).

BURIN PENINSULA TILLS

The smoothing process described here has also been applied to tills from the Burin Peninsula whose analyses have recently been subjected to k-means cluster analysis and the results investigated in detail (S. Amor, unpublished data, 2016). After experimentation, a 5-cluster model was adopt-
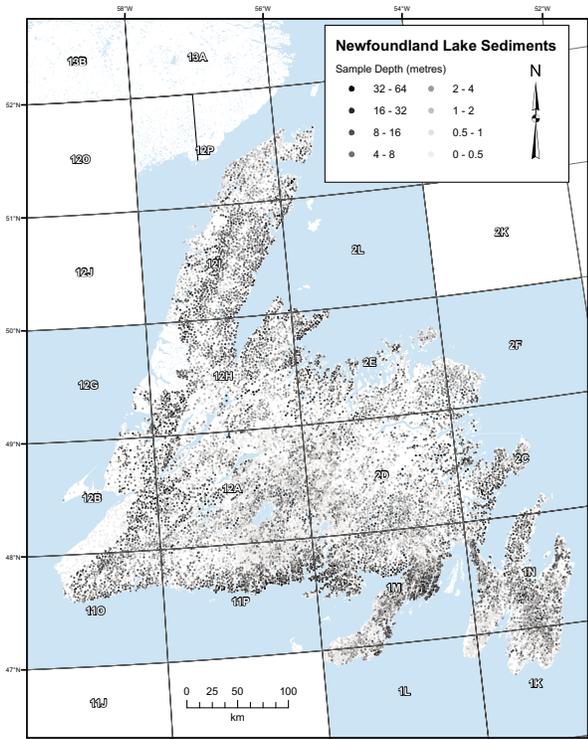
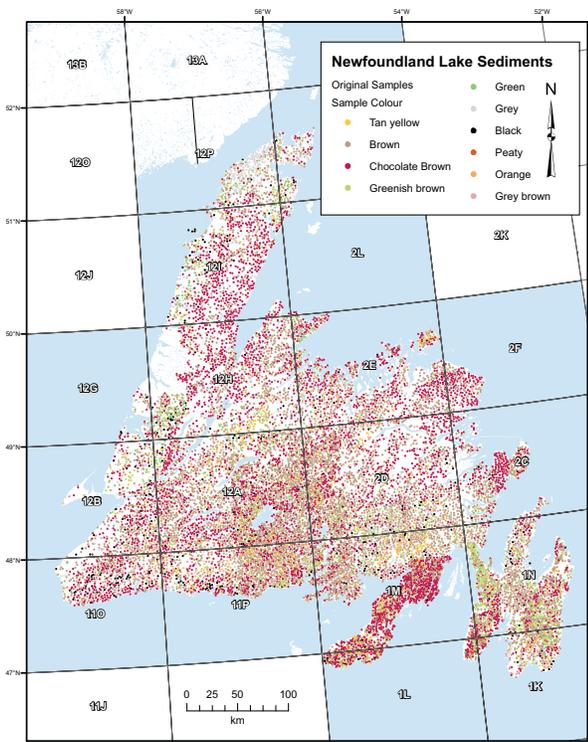**Figure 15.** *Newfoundland regional survey – depth of recovery of lake sediment.*

ed. Cluster assignments are shown as point plots and as rolling modes with a 3-km cell size, which was arrived at after experimentation; only about 40 2 km cells, out of about 2000 that cover the study area, are occupied by more than two sample points (Figures 18 and 19).

These data demonstrate both the advantages and disadvantages of the smoothing process. The samples assigned to Cluster 1 show a distinct preference for roadcuts and coastal sections, which accounts for their sinuous distribution in NTS map areas 1L/13, 1M/04 07 and 10 and 1L/1 (Figure 20A). The smoothing effect of the rolling mode might be better described, in this case, as a 'smearing' as the relationship of the point distribution to the roads and coastline is no longer apparent, even when the cell size is minimized (Figure 20B).

On the other hand, the large-scale features such as the St. Lawrence Granite and other felsic plutons are emphasized by the smoothing process, and the close (though not exclusive) spatial association of Cluster 2 with granites in the north of the peninsula, and of Cluster 3 with the St. Lawrence Granite in the south. The spatial relationship of these clusters with each other is also preserved.
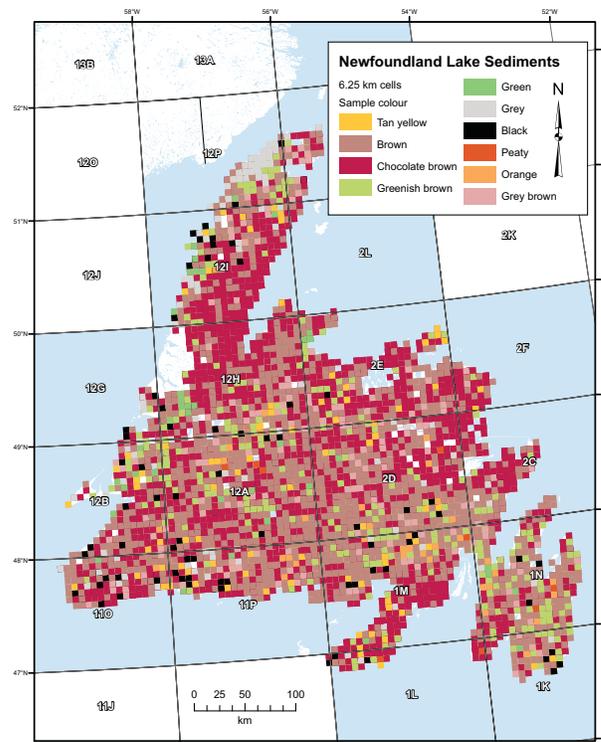

Figure 16A. *Newfoundland regional survey – sample colour (point plot).*

Figure 16B. *Newfoundland regional survey – sample colour (rolling mode, cell size 6.25 km).*
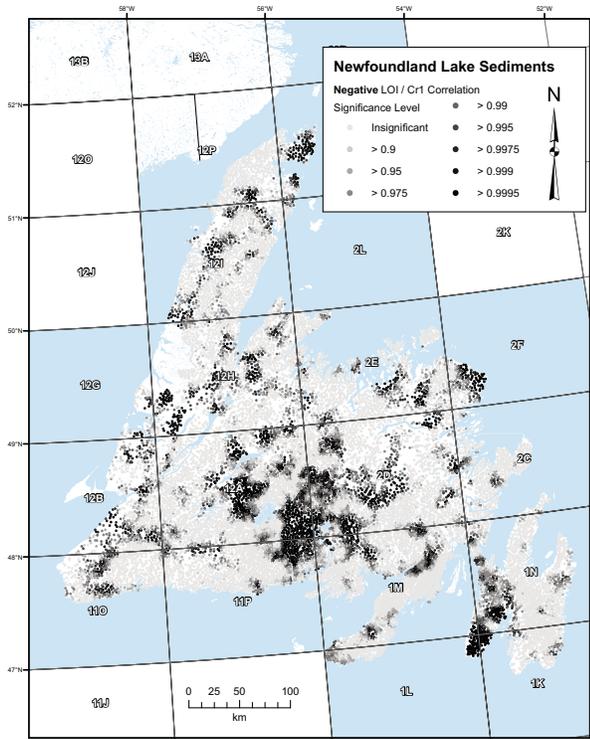
Figure 17A. *Correlation coefficients between LOI and Cr1; emphasis on negative correlations. Strong negative correlations between Cr1 and LOI are characteristic of much of the Island.*
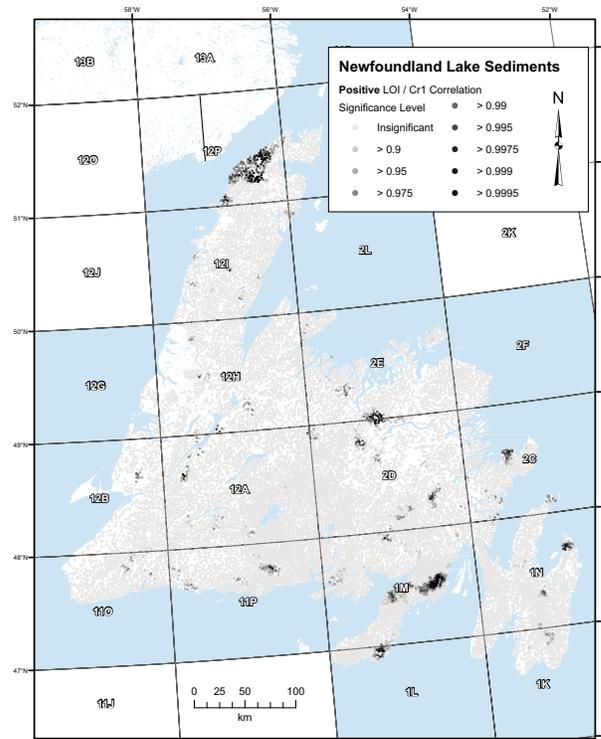

Figure 17B. *Correlation coefficients between LOI and Cr1; emphasis on positive correlations. Strong positive correlations between Cr1 and LOI are rare, but noticeable over Cambrian carbonate rocks in the Northern Peninsula.*
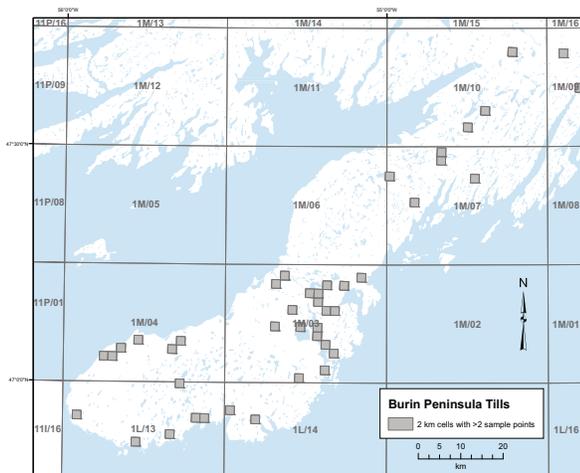
Figure 18. *Burin Peninsula tills – 2 km cells containing more than 2 sample points. The large areas where there are no such cells militate against such a small cell size in smoothing the categorical data.*
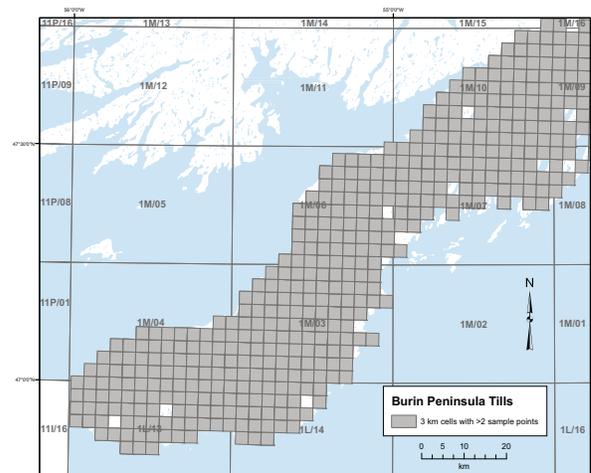
Figure 19. *3 km cells containing more than 2 sample points. Most cells of this size meet the minimal criterion for estimation of the rolling mode.*

DISCUSSION

THE VALUE OF FIELD OBSERVATIONS

The taking of notes to accompany the collection of geo-chemical samples in the field is widely advocated and there are certain observations, in certain branches of exploration geochemistry, where it is of critical importance. An example of a critical observation would be the recording of the horizon from which a soil sample is collected, because most elements occur in widely different concentrations in the A and

Figure 20A. *K-means cluster allocations of Burin Peninsula tills (point plot; from S. Amor, unpublished data).*

Figure 20B. *K-means cluster allocations of Burin Peninsula tills (rolling mode; cell size 3 km). Cluster characteristics are the same as in Figure 17A.*

(to a lesser extent) the B and C horizons (*e.g.*, Rose *et al.*, 1979, pages 150-151).

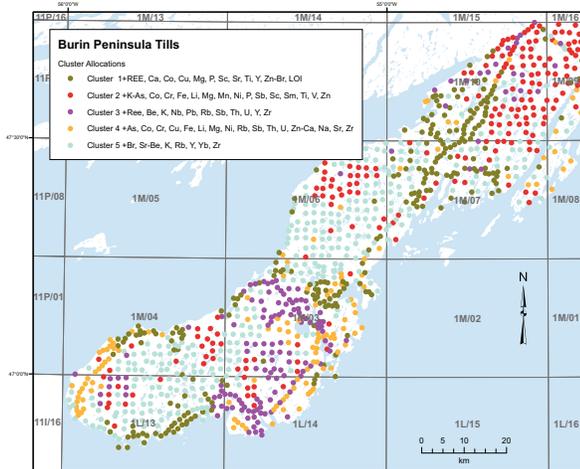Since the taking of field observations during lake-sampling programs occupies a significant proportion of the time taken to collect each sample, and if the information is considered to be of limited value, sampling productivity could probably be substantially improved if observations were reduced or eliminated. Certainly, in the NGR Labrador dataset the observations of suspended material in water, and probably, the observations of sediment colour, are too subjective as to be of application, at least on a regional scale. This is, however, less apparent in the regional Newfoundland data.

It has already been demonstrated (Amor, 2014) that observations regarding the physical composition of Labrador lake-sediment samples proved useful in identifying the relationship between the content of inorganic clastic material, and the content of certain elements. However, samples described as 'clastic – fine-grained' or 'clastic – coarse-grained' tend to be rather rare; in fact, the policy (*e.g.* McConnell, 2009) has been to avoid, wherever possible, parts of a lake where the sediment is likely to be dominated by such material.

The nature of the vegetation cover, as observed during lake-sampling work, which normally precedes till sampling and detailed geological mapping, may under certain circumstances facilitate the planning of the later work. It is unnecessary, for example, to budget the same amount of mapping work to an area reported to be predominantly boggy, as to an equivalent area characterized by exposed bedrock or till.
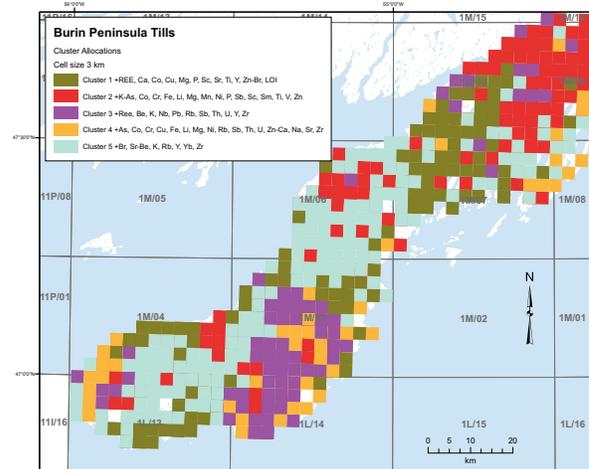
ADDRESSING SUBJECTIVITY

The problem of subjectivity in field observations giving rise to gross discontinuities in the map realizations of certain parameters is exemplified by suspended material and sediment colour in the regional Labrador dataset. It is, however, much less in evidence in the other datasets, notably the Newfoundland regional survey whose scope and terms of reference were similar to those of its Labrador counterpart.

Devices for measuring colour automatically would eliminate the problem of subjectivity and state-of-the-art models are sufficiently compact and portable that they could probably be used effectively, even in the restricted work environment of a helicopter float, and more rapidly than a Munsell colour chart. However, it is unlikely that they have been tested on lake sediments, and they are expensive (starting at almost $10 000). Perhaps as the price of such devices comes down, the feasibility of their use in lake-sediment sampling programs can be investigated further.

Water turbidity has been measured in some surveys with the use of Secchi disks (*e.g.*, Hamilton *et al.*, 1995) but this is also extremely time-consuming and a decision was made not to use such devices in programs conducted by the GSNL (J. McConnell, personal communication, 2009).

RELATIONSHIP OF FIELD OBSERVATIONS TO BEDROCK GEOLOGY

In a few instances, field observations have a relationship to certain bedrock-geology features. Examples include

the association of high relief, and deep lakes, with thrusting in rocks of the Grenville Province in southwestern Labrador, and sediments of a distinctive colour associated with carbonate rocks in southwestern and northwestern Newfoundland.

It appears that whereas none of the field observations studied in this paper can be used as a general geological mapping tool in unmapped areas, a group of spatially associated lakes with distinctive lake-sediment sample colour may (among many other causes) indicate an area of unusual geology, with respect to its surroundings, and as such, worthy of investigation.

IS THE SMOOTHING PROCESS JUSTIFIED OVER-ALL?

Results of applying a rolling-mode smoothing method to categorical variables are mixed; although the same could justifiably be said of any smoothing method. The method certainly has the undesirable effect of 'smearing' certain fine features so that their aerial extent is either exaggerated or obscured; examples include the Cluster 1 samples in the Burin Peninsula tills, and the sharp contrast in parameters such as suspended material in lake water, and lake-sediment sample colour, along the borders between some NTS sheets, in the last case to the extent that the effects of subjectivity in observation are almost completely obscured.

The rolling-mode process appears to work best as a 'big-picture' method for observing very large map areas. The most effective use, of those tried in this study, appears to be the cluster allocations in the Burin Peninsula; however, the potential for expanding the study into the rest of Newfoundland is limited, not only by the sample coverage (at least for the present) but also by Newfoundland's diverse geology and the unlikelihood that the tills derived from it can be divided into a manageable number of clusters.

One advantage of the rolling-mode smoothing technique lies in its capacity to compensate for variable sample density and the visual bias that this causes. The sample density was more than twice as high, in south-central Newfoundland (1M/12, 13, 2D/04, 05, 12, 13, 11P/09, 10, 15, 16, 12A/01, 02, 07, 08, 09, 10), as well as in the Burin and southern Avalon peninsulas, as much of the rest of the Island (sample density was not as variable in Labrador, although there is an increase of the sample density in western and west-central Labrador). Areas where the samples are more closely spaced have the effect of drawing the viewer's attention, whether or not their colour-coding indicates samples of a distinctive class. This undesirable effect is entirely alleviated by smoothing the data, since the cell size and density are the same over the entire map.

CONCLUSIONS

Map representations of various categorical variables, including field observations and multivariate cluster allocations of geochemical analyses, have been created to assess their value and investigate the feasibility of smoothing such observations to improve their visual impact.

Observations of sample colour, and suspended material in lake water, taken during the collection of lake-sediment and water samples in Labrador, as part of the National Geochemical Reconnaissance program, are highly subjective. Observations of local relief appear to be more meaningful, and some large-scale linear features have been identified, some of which appear to be related to thrust faults.

In Newfoundland, and at selected localities in Labrador, the field observations appear to be less subjective. Evidence for this lies in the absence of sharp discontinuities in parameters such as sample colour along NTS map boundaries. This may have application in planning future detailed bedrock and surficial mapping, and geochemical sampling, as well as providing clues about unsuspected geological units in unmapped areas.

Although categorical data, of which field observations and certain statistical classifications are an example, cannot be smoothed by calculated weighted or moving averages, they can be used to create a 'big-picture' map by assigning modal values to the cells in a square grid. This method, which has been termed 'rolling mode', was implemented in ArcGIS 10.2. It improves the visual impression imparted by maps of certain categorical variables although it has a 'smearing' effect on fine linear features, and considerable experimentation is necessary to arrive at an optimal cell size.

ACKNOWLEDGMENTS

REFERENCES

Amor, S.

2011: An untested rare-earth element target in northern Labrador. Government of Newfoundland and Labrador, Department of Natural Resources, Geological Survey, Open File 14E/0229, 45 pages.

2013a: A high-density lake-sediment and water survey in the Fraser Lake region, western Labrador (NTS map areas 13L/05, 06, 12, 13, 23I/08 and 09). Government

of Newfoundland and Labrador, Department of Natural Resources, Geological Survey, Open File LAB/1616, 215 pages.

2013b: A high-density lake-sediment and water survey in the Knox Lake region, western Labrador (NTS map areas 023I/06, 07, 10, 11, 12, 13, 14 and 15). Government of Newfoundland and Labrador, Department of Natural Resources, Geological Survey, Open File 23I/0101, 239 pages.

2013c: A high-density lake-sediment and water survey in the Alexis River region, southeastern Labrador (NTS map areas 13A/02, 13A/07, 13A/08, 13A/10, 13A/14 and 13A/15). Government of Newfoundland and Labrador, Department of Natural Resources, Geological Survey, Open File 13A/0089, 259 pages.

2014: Geochemical quantification of the clastic component of Labrador lake sediments and applications to exploration. Government of Newfoundland and Labrador, Department of Natural Resources, Geological Survey, Open File LAB/1625, 169 pages.

Amor, S. and Ricketts, M.J.
2015: Lake-sediment and water-sampling survey in the Kyfanan Lake region, southeastern Labrador. *In* Current Research. Government of Newfoundland and Labrador, Department of Natural Resources, Geological Survey, Report 15-1, pages 243-252.

Batterson, M.J.
1989: Quaternary geology and glacial dispersal in the Strange Lake area, Labrador. Government of Newfoundland and Labrador, Department of Mines, Geological Survey, Report 89-03, 63 pages.

Davis, J.C.
1973: Statistics and Data Analysis in Geology. Wiley, 550 pages.

Friske, P.W.B. and Hornbrook, E.H.W.
1991: Canada's National Geochemical Reconnaissance program. Institution of Mining and Metallurgy, Transactions, Section B: Applied Earth Sciences, Volume 100, pages B47-B56.

Geological Survey of Newfoundland and Labrador
2015a: Newfoundland and Labrador Geological Survey. Regional Lake Sediment Database. Newfoundland and Labrador GeoScience Atlas OnLine. Last update: 1995. http://geoatlas.gov.nl.ca. Downloaded November 2015.

2015b: Detailed Bedrock Geology. Newfoundland and Labrador GeoScience Atlas OnLine. Last update: January 2013. http://geoatlas.gov.nl.ca. Downloaded November 2015.

Gower, C.F. and Krogh, T.E.
2002: A U–Pb geochronological review of the Proterozoic history of the eastern Grenville Province. Canadian Journal of Earth Sciences, Volume 39, pages 795-829.

Hamilton, S.M., Fortescue, J. A.C. and Hardy, A.S.
1995. A zinc-cadmium-copper anomaly: preliminary results of the Cow River geochemical mapping project, Batchawana greenstone belt. Ontario Geological Survey, Open File Report 5917, 31 pages.

Hoffman, S.J.
1986: Chapter 3: Soil Sampling. *In* Exploration Geochemistry: Design and Interpretation of Soil Surveys. Society of Economic Geologists, Reviews in Economic Geology, Volume 3, 180 pages.

Howarth, R.J.
1983: Chapter 5: Mapping. *In* Statistics and Data Analysis in Geochemical Prospecting. *Edited by* R.J. Howarth. Handbook of Exploration Geochemistry, Volume 2, Elsevier, pages 111-205.

Howarth R.J. and Garrett, R.G.
2010: Statistical analysis and display at the Geochemical Prospecting Research Centre and Applied Geochemistry Research Group, Imperial College, London. Geochemistry: Exploration, Environment, Analysis, Volume 10, pages 289-315.

Howarth, R.J. and Sinding-Larsen, R.
1983. Chapter 6: Multivariate Analysis. *In* Handbook of Exploration Geochemistry. *Edited by* G.J.S. Govett. Volume 2, Statistics and Data Analysis in Geochemical Prospecting, Elsevier Scientific Publishing Company, 437 pages.

Knight, I.
1983: Geology of the Carboniferous Bay St George sub-basin, western Newfoundland. Government of Newfoundland and Labrador, Department of Mines and Energy, Mineral Development Division, Memoir 1, 382 pages.

McConnell, J.W.
2009: Complete geochemical data for detailed-scale Labrador lake surveys, 1978-2005. Government of

Newfoundland and Labrador, Department of Natural Resources, Geological Survey, Open File LAB/1465, 25 pages.

McConnell, J.W and Finch, C.
  2012: New ICP-ES geochemical data for regional Labrador lake-sediment and lake-water surveys. Government of Newfoundland and Labrador, Department of Natural Resources, Geological Survey, Open File LAB/1602, 2012, 56 pages.

McConnell, J.W. and Ricketts, M.J.
  2008: A high-density lake-sediment and water survey in eastern Labrador [NTS map areas 13J/11, 13J/12, 13J/13, 13K/9, 13K/16, 13N/1, and 13O/4]. Government of Newfoundland and Labrador, Department of Natural Resources, Geological Survey, Open File LAB/1427, 238 pages.

  2010: A high-density lake-sediment and water survey in southeastern Labrador [NTS maps areas 2M/13, 3D/4-5, 12-13, 12P/6-7, 9-11,14-16, 13A/1,9,16, and 13H/1-11]. Government of Newfoundland and Labrador, Department of Natural Resources, Geological Survey, Open File LAB/1538, 269 pages.

  2011: A high-density lake-sediment and water survey in two areas of central and western Labrador [NTS maps areas 13E/1, 2 and 8 and Labrador portions of 23I/12 (north half), 23I/13, 23J/9, 15 and 16 and 23O/1, 2 and 7]. Government of Newfoundland and Labrador, Department of Natural Resources, Geological Survey, Open File LAB/1585, 295 pages.

McConnell, J.W., Ricketts, M.J. and McCuaig, S.J.
  2007: A high-density lake-sediment and water survey in NTS areas 13K/11 and 13K/14, central Labrador. Government of Newfoundland and Labrador, Department of Natural Resources, Geological Survey, Open File 13K/0292, 104 pages.

Nunn, G.A.G., Emslie, R.F., Lefebvre, C.E., Noel, N. and Wells, S.
  1986: Geology of the Atikonak River Massif - north part (23A/SE), southwestern Labrador and Quebec. Government of Newfoundland and Labrador, Department of Mines and Energy, Mineral Development Division, Map 86-67, Scale: 1:100 000.

Rose, A.W., Hawkes, H.E. and Webb, J.S.
  1979: Geochemistry in Mineral Exploration, 2nd Edition. Academic Press, 657 pages.

Ryan, B. (compiler)
  1990: Preliminary geological map of the Nain Plutonic Suite and surrounding rocks (Nain-Nutak, NTS 14 S.W.). Map 90-44. Scale: 1:500 000. Government of Newfoundland and Labrador, Department of Mines and Energy, Geological Survey Branch. Colour map. GS# LAB/0964

Sinding-Larsen, R.
  1975: A computer method for dividing a regional geochemical survey into homogeneous areas prior to statistical interpretation. *In* Geochemical Exploration 1974. *Edited by* I.L. Elliott and W.K. Fletcher. Elsevier, pages 191-217.

Wardle, R.J. (compiler)
  1993: Geology of the Naskaupi River region, central Labrador (13NW). Map 93-16. Scale: 1:500 000. Government of Newfoundland and Labrador, Department of Mines and Energy, Geological Survey Branch. Colour map. GS# LAB/0981.

Wardle, R.J., Gower, C.F., Ryan, B., Nunn, G.A.G., James, D.T. and Kerr, A.
  1997: Geological map of Labrador, Scale: 1:1 000 000. Government of Newfoundland and Labrador, Department of Mines and Energy, Geological Survey, Open File GS# LAB/1226.