GEOLOGICAL SURVEY OF CANADA
OPEN FILE 7357

Remote Predictive Mapping of Surficial Materials West of Repulse Bay, Nunavut (NTS 46M-SW, 46L-W and -S, 46K-SW)

U. Wityk, J.R. Harris, I. McMartin, J.E. Campbell,
M. Ross, and E. Grunsky

2013

Natural Resources Canada   Ressources naturelles Canada

Canada

GEOLOGICAL SURVEY OF CANADA
OPEN FILE 7357

Remote Predictive Mapping of Surficial Materials West of Repulse Bay, Nunavut (NTS 46M-SW, 46L-W and -S, 46K-SW)

U. Wityk[1], J.R. Harris[2], I. McMartin[2], J.E. Campbell[2], M. Ross[1], and E. Grunsky[2]

[1] Department of Earth and Environmental Sciences, University of Waterloo, 200 University Avenue, West Waterloo, ON N2L 3G1
[2] Geological Survey of Canada, 601 Booth St., Ottawa, ON K1A 0E8

2013

doi:10.4095/292578

Recommended citation
Wityk, U., Harris, J.R., McMartin, I., Campbell, J.E., Ross, M., and Grunsky, E., 2013. Remote Predictive Mapping of Surficial Materials West of Repulse Bay, Nunavut (NTS 46M-SW, 46L-W and -S, 46K-SW); Geological Survey of Canada, Open File 7357. doi:10.4095/292578

TABLE OF CONTENTS

Abstract .... p. 1
Introduction .... p. 1
Regional Setting .... p. 3
Methodology .... p. 3
    1 - Data acquisition, image preparation and masking .... p. 3
    2 - Initial field data collection and selection of ROIs .... p. 4
    3 - Final ROI selection .... p. 7
    4 - Evaluation of ROIs .... p. 7
    5 - Classification method and selection of “best” classification maps .... p. 9
    6 - Variability and Majority classification maps .... p. 12
Results and Discussion .... p. 12
    Maximum Likelihood Classification Maps .... p. 12
    Variability maps .... p. 16
Summary and Recommendations .... p. 17
Acknowledgements .... p. 19
References .... p. 19

Appendices

The Readme file provides the complete list of files and sub-directories containing the datasets:

1. Appendix_01_Data
2. Appendix_02_Classification maps
3. Appendix 03_Variability maps
4. Appendix 04_ Shape files

LIST OF TABLES

1. Summary of data characteristics .... p. 4
2. List of 21 surficial material classes .... p. 8
3. Summary of LANDSAT Bands and reflectance data used to produce classification maps of surficial materials .... p. 10
4. Calculated overall and individual class accuracies of the 4 maps GK1, GK2, Stat1, Stat2 .... p. 11
5. Class combinations used to produce the "best" maps in this study ………..… p. 13

LIST OF FIGURES

1. Location map of study area .... p. 2
2. Flow chart outlining steps to produce RPM of surficial materials .... p. 4
3. LANDSAT imagery presenting RGB bands 7, 4, 2 .... p. 5
4. LANDSAT 7 imagery cut-lines over the Wager Bay North area .... p. 6
5. Production of final mask to mask out water and cloudy regions .... p. 6
6. Location of ROIs used in the classification maps .... p. 9
7. Average Separability Statistic of 21 classes .... p. 9
8. Classifications maps .... p. 14-15
9. Variability maps .... p. 17-18

uwityk@uwaterloo.ca

REMOTE PREDICTIVE MAPPING OF SURFICIAL MATERIALS WEST OF REPULSE BAY, NUNAVUT (NTS 46M-SW, 46L-W AND -S, 46K-SW)

U. Wityk, J.R. Harris, I. McMartin, J.E. Campbell, M. Ross, and E. Grunsky

Abstract

Canada's vast northern territories require an efficient and timely method to create surficial geology maps. Due to the convenient availability of remotely sensed imagery, it is effective to develop and test this resource by using various automatic and remote approaches to assist with the production of surficial geology maps.

The goal of this remote predictive mapping (RPM) effort is to generate a map of surficial materials, in turn accelerating traditional mapping efforts, which use fieldwork, sampling and other resources. This report describes techniques used to create a remote predictive surficial materials map near Repulse Bay in Nunavut using classification algorithms applied to LANDSAT imagery, and tests two approaches in the iterative RPM process. The approaches are: 1) geological knowledge-based, incorporating expert knowledge and traditional mapping data, and 2) statistically-based, investigating statistical accuracies of the classifications. Based on variability maps of the most optimal classification maps, it was found that the geological knowledge-based approach is more suitable for remotely mapping materials in the study area.

Introduction

Quaternary geological mapping in the Wager Bay-Repulse Bay region of mainland Nunavut was initiated in 2009 at the Geological Survey of Canada within the framework of the Geo-mapping for Energy and Minerals (GEM) Program. The purpose of this activity was to address and fill in knowledge gaps with respect to the distribution and nature of Quaternary sediments, regional drift composition, and glacial and post-glacial histories. As part of this work, a Masters thesis research project was undertaken to assist the mapping of surficial earth materials by Remote Predictive Mapping (RPM) in an area covering parts of NTS sheets 46 K, L and M located near Repulse Bay, Nunavut (Fig. 1). The purpose of this Open File publication is to release a selection of RPM classification maps and accompanying datasets for the study area specific to the thesis project.

Remote Predictive Mapping is a semi-automated approach used to increase the efficiency of mapping bedrock and surficial geology over large regions of Canada's Far North (Harris, 2008a, 2008b). It is a useful tool, especially for producing maps of regions that will not be field mapped or have limited fieldwork in the foreseeable future due to logistical constraints, high costs and the size and remoteness of the map area. The purpose of this new mapping approach is not to replace traditional geological mapping i.e. field work and interpretation of air photos, but rather to enhance the mapping process by providing insight regarding surficial materials found in these regions. These constraints on fieldwork can, in part, be addressed through the use of remotely sensed imagery which offers a broad view of large inaccessible areas, providing a wealth of geologic information that can be enhanced and processed using image analysis and GIS technologies (Harris, 2008a, 2008b). RPM is a tool to aid fieldwork, streamline the mapping process, as well as enhance extrapolation and interpretation between field observation sites.

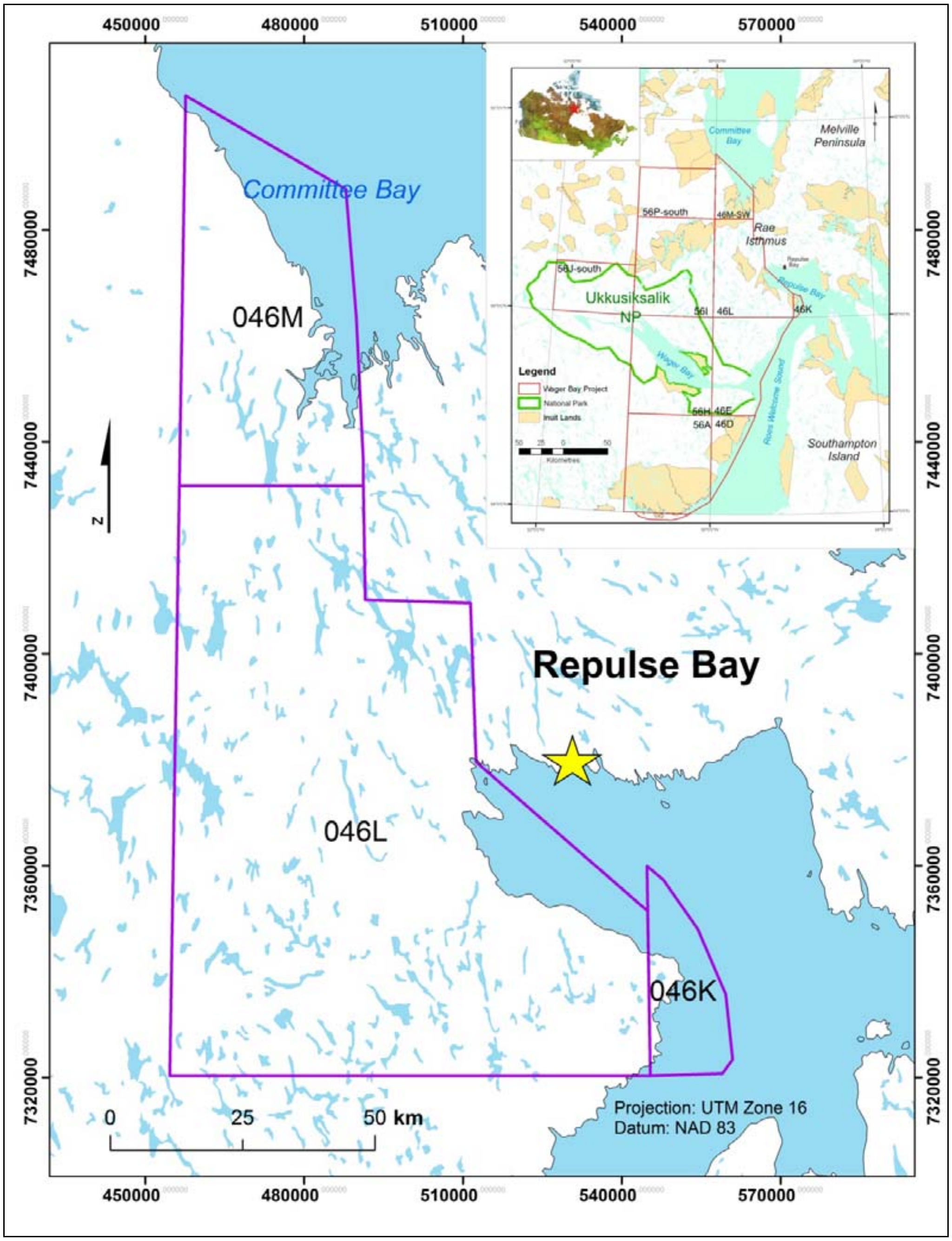

Figure 1: Location map of study area west of Repulse Bay, Nunavut. Inset map shows the project area around Wager Bay (from Campbell and McMartin, 2011).

This Open File report includes information regarding the regional setting of the study area, and a description of the methodology used to produce RPM maps for the Repulse Bay study area, including data acquisition and preparation, field data collection, region of interest (ROI) selection and evaluation, and classification methods. The report also discusses four of the classification maps that offer the highest classification accuracies determined through analysis of a confusion matrix and associated variability maps. The datasets related to the RPM classification maps include: LANDSAT data, raster image files of the four classification maps and corresponding variability maps, ROIs (in the form of vector shape files in ENVI) used to produce the maps, and the base layers of the study area.

Regional Setting

The RPM study area is located on mainland Nunavut west of the northern community of Repulse Bay between latitudes 66°N and 67.5°N and longitudes 88°W and 86°W (Fig. 1). It is comprised of parts of NTS map sheets 46K, 46L and 46M. The region is located along the west coast of Hudson Bay, west of Committee Bay and Repulse Bay. It is underlain by Canadian Shield rocks of the western Churchill Geological Province including Archean through Paleoproterozoic intrusive and supracrustal rocks within the 2.7-2.6 Ga Rae Domain (Paul et al., 2002), and covered by fairly continuous glacial drift which is streamlined in the general direction of the regional ice flow, predominantly northward (Prest et al., 1968; Aylsworth and Shilts, 1989; McMartin et al., 2013). Marine limit elevations decrease from approximately 240 m asl to 140 m asl southward within the study area. Extensive areas of marine sand, silts and clays are exposed along the Committee Bay coastal plain in 46M-SW (e.g. Campbell and McMartin, 2010).

Methodology

The methodology used here is modified after the RPM approaches of Grunsky et al. (2006, 2009), Schetselaar et al. (2007) and Harris et al. (2008a; 2012) (Fig. 2). It uses data acquisition and image preparation, fieldwork and selection of training data, which are then applied to a maximum likelihood classification (MLC) algorithm to produce classification maps of surficial materials. Two approaches are used to determine the optimal class combination and resultant MLC maps: a geological knowledge-based approach and a statistical approach. The preferred optimal class combinations are then run through the robust classification method (RCM) to produce variability maps and evaluate the uncertainty of the classifications.

1- Data acquisition, image preparation and masking

LANDSAT TM-7 imagery used in this study (Table 1) was downloaded from GeoGratis (http://geogratis.cgdi.gc.ca), available from Natural Resources Canada, in GeoTIFF format. After collection, the individual scenes were "stitched" together by the Geological Survey of Canada and BlackBridge Geomatics (formally Iunctus Geomatics). The overlapping LANDSAT scenes were combined and levelled to produce a single visually, but unfortunately not spectrally seamless image of the study area (Fig. 3). LANDSAT images were projected to Universal Transverse Mercator (UTM), Zone 16, and referenced to the North American Datum (NAD83). The mosaic image was then clipped to the boundaries of the study area (See Appendix 01). Cutlines, and borders at which individual imagery scenes were "stitched" together, are shown in white on Figure 4.

Water was not classified in this study. Running and standing water, along with heavily saturated areas, were masked out using the masking tool in ENVI, or more specifically spectral principles. This was accomplished by utilizing the LANDSAT near-infrared (NIR) channel (band 4), which easily identifies water (Frazier and Page, 2000). The interactive stretching tool in the ENVI interface was used to discriminate pixels with digital numbers (DN) representing water. Clouds and cloud shadows were also masked by manually digitizing cloud polygons in the ENVI software package as ROIs. Then, using simple band math, water and cloud masks were merged together to create a final mask for the LANDSAT mosaic image (Fig. 5). The purpose of this masking process is to exclude water bodies (inclusive of lakes and saltwater and sea ice), and clouds from the classifications, which would artificially inflate overall accuracy measures.

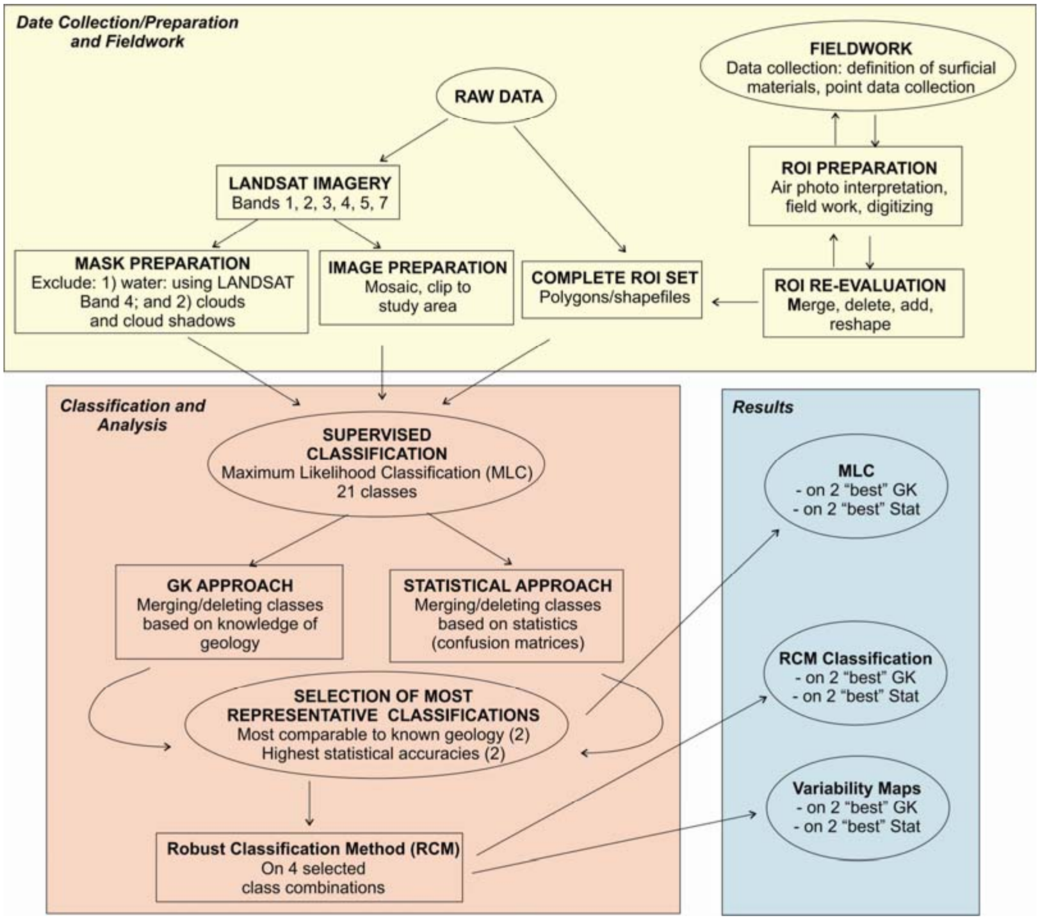

Figure 2: Flow chart outlining steps to produce RPM of surficial materials using supervised classification and selection (Geological Knowledge and Statistical approaches) of most representative classifications to arrive at final maps.

Table 1: Summary of data characteristics

Data	Source	Bands	Channels	Resolution
LANDSAT TM-7	GeoGratis (http://geogratis.cgdi.gc.ca)	1, 2, 3, 4, 5, 7	Visible, NIR, SWIR	30 meters

2- Initial field data collection and selection of ROIs

Initial fieldwork, conducted in the summer of 2010, involved determining the surficial materials present in the study area and acquiring field observations. This included making observations regarding surficial geological units as part of the mapping activity as well as collecting point data information for specific field sites. The observations collected included site location (latitude/longitude), general terrain descriptions and geomorphological conditions. Surficial material type, surface texture, boulder cover (%/presence/size/shape), vegetation cover (%), moisture content, geomorphology, topography, drainage and lithology were systematically noted. Photographs were taken and recorded at each site. Initial field data collection ensured that the classes used in the supervised classification process reflected the diversity of surficial materials present in the study area.

The first season of fieldwork and preliminary air photo interpretation (guided by the GSC field geologists) identified eleven classes of surficial materials, having distinctive physical (mainly textural) and geomorphological characteristics.

This initial working class list included:

1. Ap – Alluvial plain sediments
2. Mg - Marine, gullied fine-grained sediments
3. Ms – Marine sands
4. O – Organic
5. SG – Sand and gravel
6. R - Bedrock
7. Tb – Till blanket
8. Tm – Modified till
9. Tv – Till veneer
10. B - Boulders
11. Tr – Ribbed till

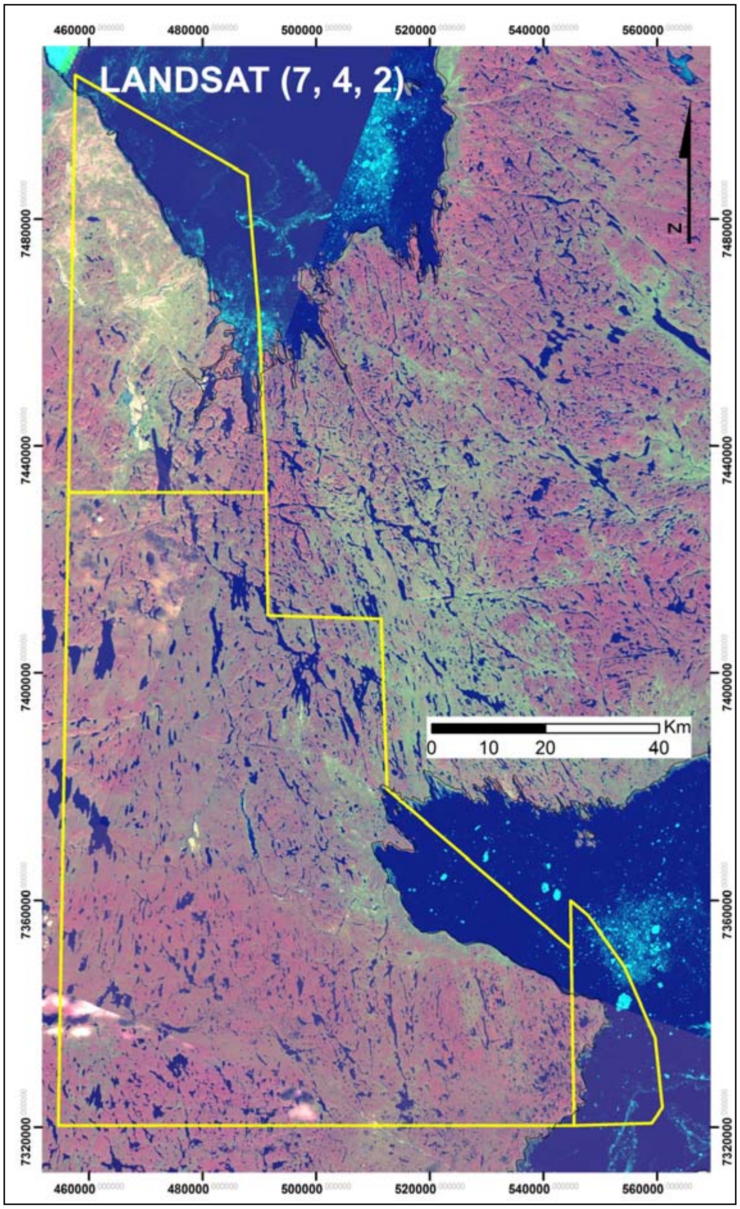

Figure 3: LANDSAT imagery presenting RGB bands 7, 4, 2. Study area outlined in yellow.

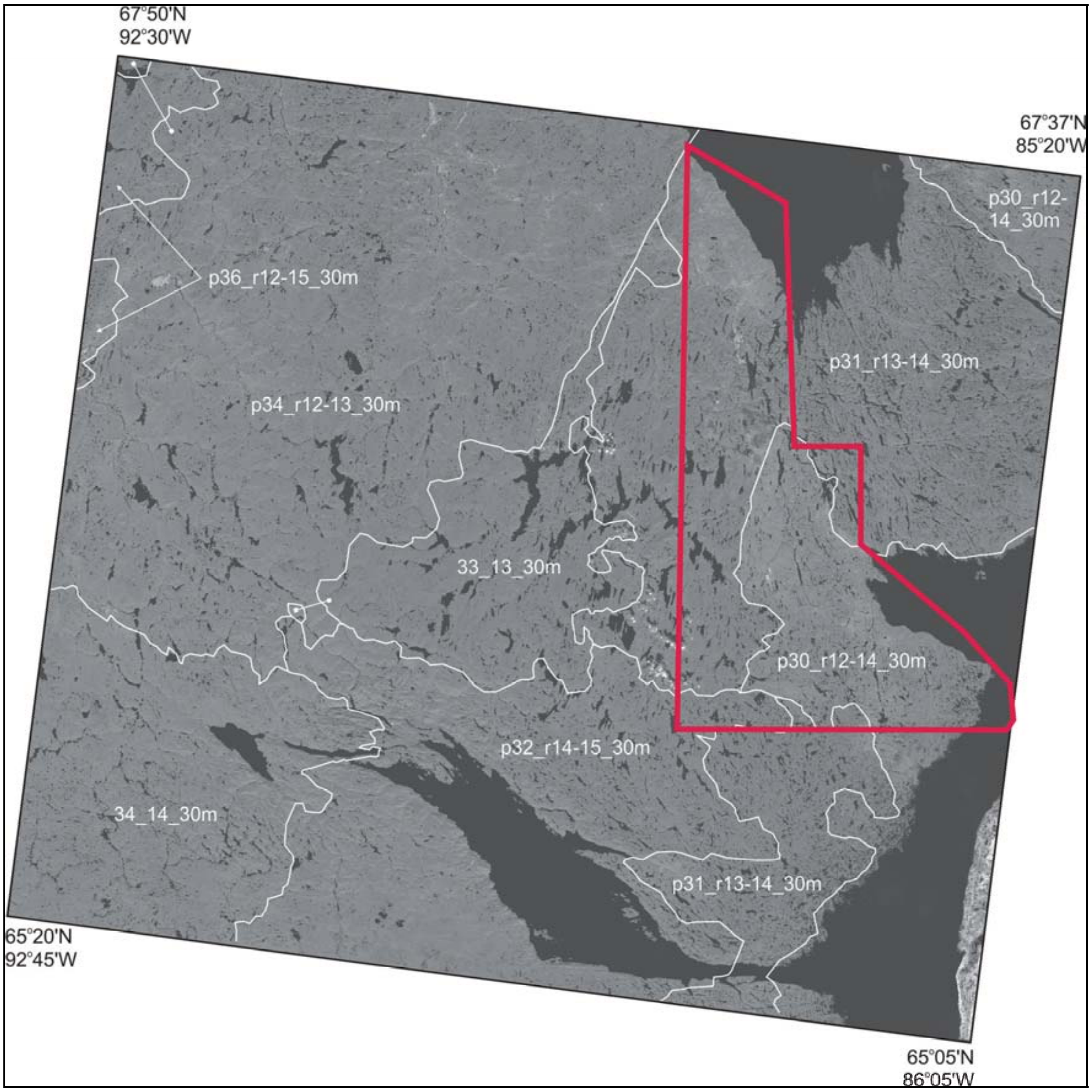

Figure 4: LANDSAT 7 (SWIR, band 5) imagery cut-lines over the Wager Bay North area. White lines indicate boundaries where LANDSAT scenes were merged together (Huntley et al., in prep.). The red line outlines the study area.

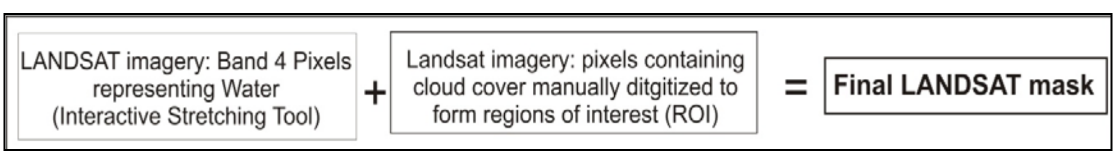

Figure 5: Production of final mask to mask out water and cloudy regions.

In the fall of 2010, multiple regions of interest (ROIs) were carefully selected to represent each of the 11 surficial material classes identified after the first field season. The ROIs were manually delineated as polygons on air photos, using a combination of field data, air photo interpretation and visual analysis of LANDSAT imagery. The polygons were then transferred on printed maps of LANDSAT imagery (bands 7, 4, 2) and then digitally captured in ENVI software as ROIs.

3 - Final ROI selection

Upon review of the satellite imagery in the winter 2010/spring 2011, two additional classes were added to the original 11 classes: ice-covered regions (ice/snow) as well as shallow water (Sw) thus resulting in 13 classes in total. After the second field season (July 2011) and a review of the original ROIs through visual exploration of the spectral response of LANDSAT imagery (bands 7, 4, 2), some classes were divided into sub-classes. A carbonate till unit (Ct) was also added based on the mapping of suspected carbonate till after a preliminary supervised classification in the spring of 2011 and confirmation in the field in 2011. Existing ROI polygons were assigned to sub-classes (i.e. Tb classes were divided into Tb1 and Tb2), and additional ROI polygons were defined to ensure each class had a sufficient number of ROIs to produce a robust classification. ROIs were added using air photo interpretation and visual analysis of spectral responses on LANDSAT (7, 4, 2) imagery which highlighted the spectral variation between surficial materials. It was apparent through visual analysis that the Ms, Tb, Tv, Tm, R, and SG classes could be further subdivided by differences in spectral response into subclasses (Ms1, Ms2, Tb1, Tb2, Tv1, Tv2, Tm1, Tm2, R1, R2, SG1, SG2 and SG3). Thus the final number of surficial material classes to be classified totalled twenty-one. As part of quality assurance/control, some original ROIs were manually clipped and/or reshaped while others eliminated in order to have the most homogeneous group of pixels per class based on their visual appearance (colour) using LANDSAT bands 7, 4, 2. The following table (Table 2) is a list of the division of classes. A further detailed description is available in Appendix 02 (Class descriptions). Figure 6 shows the location of the final regions of interest. The shape files for the ROIs are included in Appendix 04 (ROIs).

4- Evaluation of ROIs

The ROIs were evaluated through fieldwork, visual analysis/interpretation and statistics. After their initial delineation following the first round of fieldwork, the ROIs were visually assessed; and those that were spectrally (in a visual sense) different to the other ROIs of that class were investigated during the second season. As a result, a significant number of the field-verified ROIs were used for control and the basis for comparison included in the classification process. ROIs were also evaluated based on the Transform Divergence (TD) statistic which measures the statistical separability between classes (Fig. 7). The TD statistic is a number between 0 and 2 in which values <1.0 indicate very poor separation between classes, 1.0-1.9 indicates poor to moderate separation, and 1.9-2.0 good separation (Richards and Jia, 1999). Figure 7 shows that the average class separability ranges from poor/moderate (1.0-1.9) to good (1.9-2), with the majority of classes falling under the moderate category. Those classes that are spectrally separable from others include Ap, Sw, Mg, Ms1, Ice/Snow, O, SG3, and Ct. Till, bedrock, together with sand and gravel subclasses, are characterized by lower TD values (moderate or poor separation), suggesting that the potential for confusion between these subclasses will occur. Part of this potential confusion is captured in the variability maps that are generated as part of the classification process. Once identified, these areas of uncertain classification can be excluded from the classification map if desired.

Table 2: List of 21 surficial material classes, with code and colour as per classified maps, and short description. The complete description of the classes are provided in Appendix 02 (Class descriptions).

Surficial Material Class	**Code**	**Colour Classification**	**General Description**
Exposed alluvial sediments	Ap		Alluvial sands and minor silts; exposed
Marine gullied fine sediments	Mg		Marine silts and clays; exposed sediments; gullied
Marine fine sediments	Ms1		Marine fine sands, silts and clays; some surface runoff features
Marine sands and silts	Ms2		Marine sands and silts; nearshore deposits; coarser than Ms1
Ice/snow	Ice/Snow		Frozen water
Organics	O		Thin organic deposits
Vegetated coarse sand and gravel	SG1		Glaciofluvial and marine sands and gravels
Sand and gravel	SG2		Marine fine-grained sands, silty-sands
Exposed sand and gravel	SG3		Glaciofluvial and marine sands and gravels; exposed
Bedrock (bare)	R1		Bedrock; exposed
Bedrock	R2		Bedrock with some discontinuous material cover; lichen covered
Boulder fields	B		Broken bedrock; continuous boulder cover
Till blanket	Tb1		Thick drift cover with little boulder cover or exposed bedrock
Bouldery till blanket	Tb2		Thick drift cover; more boulders and less vegetated than Tb1
Modified till	Tm1		Modified till; eroded in places; may include sand and gravel; bouldery
Modified till	Tm2		Modified till; less bouldery than Tm1
Till veneer	Tv1		Thin drift cover; mixed with bedrock and boulders or bedrock and sand; contains more boulder/bedrock terrain than Tv2
Till veneer	Tv2		Thin drift cover; contains more moisture and vegetation than Tv2
Carbonate till	Ct		Till with carbonate clasts and calcareous matrix
Ribbed till	Tr		Till mixed with sand, gravel and boulders; eroded, disorganized, gravelly ridges, terraces and hummocks
Shallow water	Sw		Heavily sediment laden water

#

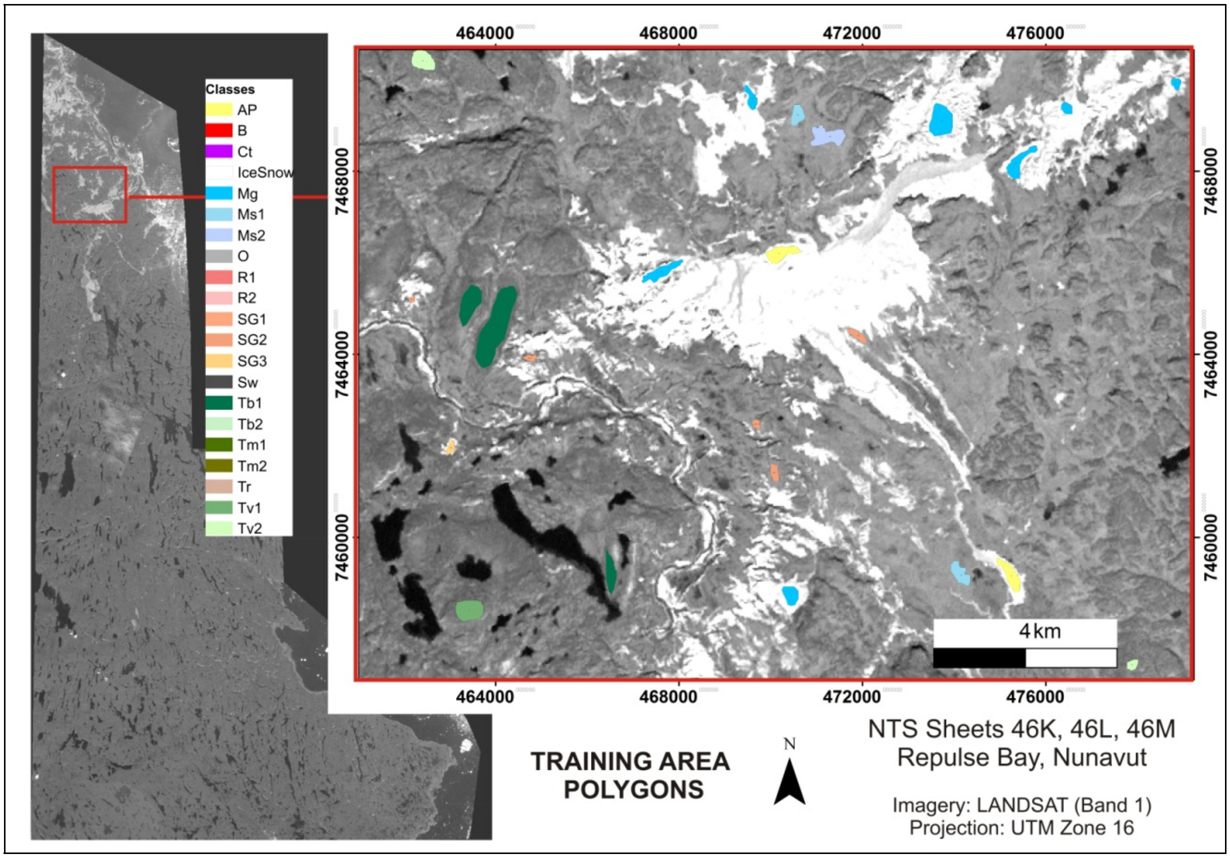

Figure 6: Location of ROIs used to produce classification maps. Inset map shows example of ROI polygons and their locations overlain on LANDSAT imagery (Band 1).

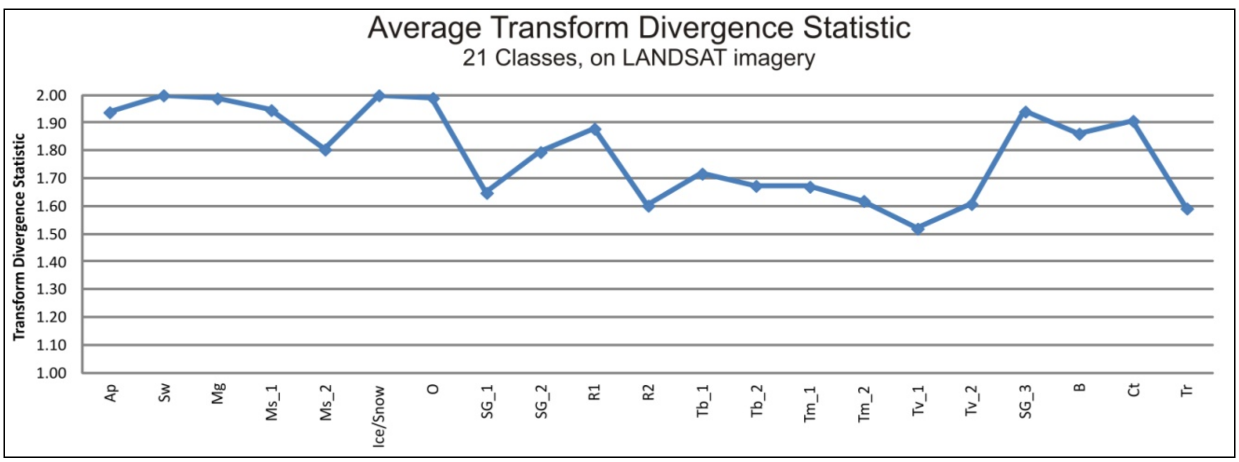

Figure 7 : Average Separability Statistic of 21 classes for the Repulse Bay Study area. Calculated on LANDSAT imagery for the region.

5- Classification method and selection of "best" RPM classification maps

A supervised classification using single pass maximum likelihood classification (MLC) algorithms were run on all 6 available LANDSAT bands (Table 3), using all 21 surficial materials classes (Table 2).

Table 3: Summary of LANDSAT Bands and reflectance data used to produce MLC classification maps of surficial materials.

Band	Reflectance Data Recorded
1	Blue (B)
2	Green (G)
3	Red (R)
4	Near-Infrared (NIR)
5	Short Wave Infrared (SWIR)
7	Short Wave Infrared (SWIR)

To produce the four "best" MLC classification maps, a geological knowledge-based approach to earth material classification was evaluated against a statistically based approach involving computer-assisted numerical analyses. Each approach yielded a selection of two predictive surficial materials maps of the study area, totaling four "best" RPM maps. For purposes of this Open File, the term "best" refers to the most appropriately mapped training area combination (geological knowledge-based approach) and highest accuracies (statistical approach) of the run classification iterations.

The geological knowledge-based approach involved comparison of RPM maps with geological mapping results for this region, based on expert Quaternary knowledge, air photo interpretation and field observations. It included a careful visual interpretation of the satellite imagery by the Quaternary geologists, specifically redefining ROIs spectrally and merging/deleting classes, an assessment and comparison of a series of predictive maps produced from using different combinations of the 21 established classes, and the selection of two of these maps most consistent with the interpreted geology (See Appendix 02; Geological/Knowledge map 1: GK1 and Geological/Knowledge map 2: GK2). These maps are provided in jpg, pdf and raster formats. Table 4 presents the overall accuracies and the individual class accuracies for these two maps.

Two other "best" MLC classification maps were chosen based on an unbiased statistical approach, which investigated the statistics of the confusion matrix of a classification using all 21 classes. To produce the two most statistically accurate maps, both general (overall accuracy) and detailed (individual class accuracy – i.e. user's and producer's accuracy) data were considered. The producer's accuracy measures the number of pixels within a particular class that have been classified appropriately. The user's accuracy measures how many pixels of a class were classified properly over the total number of pixels assigned to that class (Grunsky et al., 2009). This approach produced an unbiased decision *without* expert interpretation. The goal was to select the most accurate maps through various iterations; and in the process, increase accuracy within each individually mapped surficial material class. This was accomplished by using all earth materials in the classification to produce a confusion matrix that was used to assess which classes were mapped reasonably (70% accurate), "moderately" (35-70% accurate), and "poorly" (<35%). Classes that were mapped with an accuracy below 70% were investigated and decisions were made for further classification modifications to improve class accuracies and/or overall accuracies of the map: for example, removing or merging classes used in the original classification for a subsequent re-classification. Accuracy was obtained from the confusion matrices calculated using the ROI set employed in each classification (see Appendix 02\Confusion_matricies).

Table 4: Calculated overall and individual class accuracies of the 4 "best" MLC classification maps GK1, GK2, Stat1, Stat2. Notation "m" indicates these classes have been merged with another class; "e" indicates class was eliminated and not used in the classification.

GK1		GK2		Stat1		Stat2	
Overall Accuracy	**60.54**		**60.42**		**62.17**		**60.64**
Ap	77.06	Ap	77.06	Ap	76.14	Ap	76.14
Mg	88.23	Mg	88.23	Mg	88.23	Mg	88.23
Ms1	91.88	Ms1	91.88	Ms1	91.94	Ms1	92.2
Ms2+SG2	59.18	Ms2+SG2	61.63	Ms2	62.99	Ms2	"m"
O	78.54	O	81.78	O	76.92	O	77.33
SG1	25.52	SG1	31.09	SG1	"e"	SG1+SG2+Ms2	45.21
SG2	"m"	SG2	"m"	SG2	"e"	SG2	"m"
SG3	61.88	SG3	61.88	SG3	63.39	SG3	62.03
R1+R2	48.6	R1+R2	64.09	R1	70.75	R1	66.23
R2	"m"	R2	"m"	R2+Tm1	45.62	R2	31.98
B	89.19	B	92.16	B	86.22	B	81.62
Tb1+Tb2	48.01	T	49.65	Tb1	56.35	Tb1	56.99
Tb2	"m"	Tb2	"m"	Tb2	48.8	Tb2	48.03
Tm1+Tm2	61.66	Tm1	"m"	Tm1	"m"	Tm1	64.77
Tm2	"m"	Tm2	"m"	Tm2	53.51	Tm2	50.69
Tv1+Tv2	49.12	Tv1	"m"	Tv1+Tv2	51.25	Tv1+Tv2	43.09
Tv2	"m"	Tv2	"m"	Tv2	"m"	Tv2	"m"
Ct	90.59	Ct	91.15	Ct	86.07	Ct	87
Sw	94.4	Sw	94.4	Sw	"e"	Sw	"e"
Ice	7.37	Ice	7.37	Ice	"e"	Ice	"e"

A scale for merging classes based on their separability was developed to assist in reducing the confusion between classes. The separability thresholds for this scale included: <1% for low, 1-10% for moderate and >10% for high. If confusion of an individual class was statistically high, and was higher than the accuracy of the class itself, the class was removed and subsequent classifications were run without that class. If there was a moderate or high confusion with another class, but the value was still lower than the class accuracy, these two classes were merged together to form a new class. For example, in a manual analysis of a confusion matrix calculated from a classification result, the confusion matrix revealed an accuracy of 31.79% for class R2. Since its confusion with another class (Tm1) was at 25.63%, R2 was not deleted, but merged with Tm1, producing a new class: R2_Tm1.

The various modifications to the class combinations were used in the initial classifications and applied to subsequent classifications, which included one modification to the total number of classes used (i.e. only one alteration, merging classes together, or deleting them), until adequate accuracies (>35%) were attained. The two "best" MLC classification maps selected based on statistics include: 1) a map with the highest overall accuracy (Statistical 1: STAT1), and 2) a map whose confusion matrix illustrates that individual class accuracies are all "moderately" (35-70% accurate) or "well" (+70% accurate) mapped, and none are “poorly” (<35%) mapped (Statistical 2: STAT2). Both maps are presented in Appendix 02 in jpg, pdf and raster formats.

6- Variability and majority classification maps

Once the four “best” supervised classification maps were chosen, the resulting class combinations, based on different test runs involving the merging of the original 21 classes, were run through the robust classification method (RCM). The RCM algorithm produces a majority classification map, a variability map, and more summary statistics (Harris et al., 2012). The method uses a training dataset that is repeatedly and randomly split into two groups of ROIs: one for classification and the second for evaluation. These two groups comprise random collections of the ROIs for each repetition which are specified by the user. For example, in a specification of a 50% sampling percentage, each repetition will involve a different combination of 50% of the ROIs (Harris et al., 2012). Outputs of the RCM are based from inputs specified by the user, including: input files, ROIs, ROI sampling percentage, sampling type, classification method, threshold, number of repetitions, and a root name for output files (Harris et al., 2012). The user inputs specified in this study to produce variability and majority classification maps included a sampling percentage of 50, sampling type based on polygon and a repetition of 40, meaning that 40 iterations using random combinations of 50% of the data for classification and the other 50% for validation. The RCM variability (uncertainty) map provides a summary of the number of different classes that have been classified on a pixel-to-pixel basis over the 40 iterations of the RCM algorithm (Harris et al., 2012). Variability maps were produced for each the surficial material majority classification maps to show their respective spatial variability in terms of the degree of uncertainty or reliability of pixel classification in the repetitive classification processes. Masked areas were not included in the classification and accuracy assessment. The variability maps are provided in Appendix 03 and are provided in jpg, pdf and raster formats. The RCM majority classification maps were produced but are not discussed in this report.

Results and Discussion

Maximum Likelihood Classification Maps

The four “best” Maximum Likelihood Classification (MLC) maps show similarities and differences with respect to the classification of surficial materials (Figs. 8a and 8b). General consistencies across the four maps include similar mapping of the classes Ap, and Mg/Ms-related sediments. These materials were mainly classified in the northernmost coastal region of the study area. Various combinations of till classes are consistently mapped in the mid to southern part of the study area. Rock and boulder units are generally mapped in similar areas across the four maps; however their classification densities in these areas differ. For example, in the central and western south of the study area, these units are more pronounced in the GK2 and Stat1 maps than in the GK1 and Stat2 maps although they appear in the same areas. This is the same scenario for the region in the North of the study area that occurs south of the dominated marine sediments (blue colours, as per the legend). Carbonate till is also mapped with some consistency across all four maps in the southeastern area, but with more variability within the central-eastern area. As discussed below, though repeated regional patterns are recognized, significant differences also prevail when comparing the four classified maps with one another. Table 5 below presents the class combinations used in the 4 "best" MLC maps that will be described in the following predictive maps. Note that the ribbed till (Tr) class was eliminated from all classifications as it disrupted classifications, as per visual analysis, and was clearly overrepresented for this region.

1) Map GK1: General Knowledge 1

Map GK1 (Fig. 8a), using 15 surficial classes (Table 5) classified the northern portion of the study area with dominant surficial materials comprising marine sediments and sand and gravel (Ms1, Ms2+SG2, Mg). Some of these sediments (Ms2+SG2) were also mapped along rivers and at low elevation in the south-central part of the study area along the coast, below the known limit of marine inundation. An extensive till cover (Tb1+Tb2 and Tm1+Tm2) was mapped (classified) in the central and southern region of the map, and carbonate tills were mapped in the southeast and central eastern regions.

The Ms2+SG2, Ms1 and Ct classes were mapped with reasonable accuracy based on the current surficial geology mapping activities undertaken at the GSC (Campbell and McMartin, 2010, 2011, in prep; Campbell et al., 2011; McMartin et al., 2012, 2013). Bedrock (R1 and R2) and boulder (B) regions were more sparsely distributed than evident during field mapping.

Table 5: Class combinations used to map the "best" MLC maps in this study; "m" indicates the class was merged with another class; "e" indicates removal of the entire class.

	GK1	**GK2**	**Stat1**	**Stat2**
Ap	Ap	Ap	Ap	Ap
Mg	Mg	Mg	Mg	Mg
Ms1	Ms1	Ms1	Ms1	Ms1
Ms2	Ms2+SG2	Ms2+SG2	Ms2	"m"
Ice/Snow	Ice/Snow	Ice/Snow	Ice/Snow	Ice/Snow
O	O	O	O	O
SG1	SG1	SG1	"e"	SG1+SG2+Ms2
SG2	"m"	"m"	"e"	"m"
SG3	SG3	SG3	SG3	SG3
R1	R1+R2	R1+R2	R1	R1
R2	"m"	"m"	R2+Tm1	R2
B	B	B	B	B
Tb1	Tb1+Tb2	T	Tb1	Tb1
Tb2	"m"	"m"	Tb2	Tb2
Tm1	Tm1+Tm2	"m"	"m"	Tm1
Tm2	"m"	"m"	Tm2	Tm2
Tv1	Tv1+Tv2	"m"	Tv1+Tv2	Tv1+Tv2
Tv2	"m"	"m"	"m"	"m"
Ct	Ct	Ct	Ct	Ct
Tr	"e"	"e"	"e"	"e"
Sw	Sw	Sw	"e"	"e"

2) Map GK2: General Knowledge 2

Similar to map GK1, map GK2 (Fig. 8a) classified the northern region with a dominance of marine sediments and sand and gravel (Ms1, Ms2+SG2, Mg). Marine sediments were also mapped similarly to map GK1 along the central coastal areas. A higher proportion of boulder fields and bedrock was mapped in the entire region, but are present in the same general areas as in map GK1. Though till was mainly classified in the central and southern parts of the map, its overall distribution is much less than that of Map GK1 because of the increased areas mapped as bedrock and boulders. Carbonate tills were mapped consistently between map GK1 and GK2, located in the southeast and central east portions of the region.

Similar to map GK1, Ms2+SG2 as well as Ms1 appear to be mapped reasonably well based on geological knowledge of the region, while the mapping of the boulder (B) class in map GK2 is more representative in the central/southern portion of the map in comparison to GK1. Bedrock is more pervasive in map GK2 than in map GK1, especially in the southeast portion of the study area along Repulse Bay and just west of southern Committee Bay. Although the three till units used in the previous classification were combined to form a single class in map GK2, the abundance of till classified here has significantly decreased.

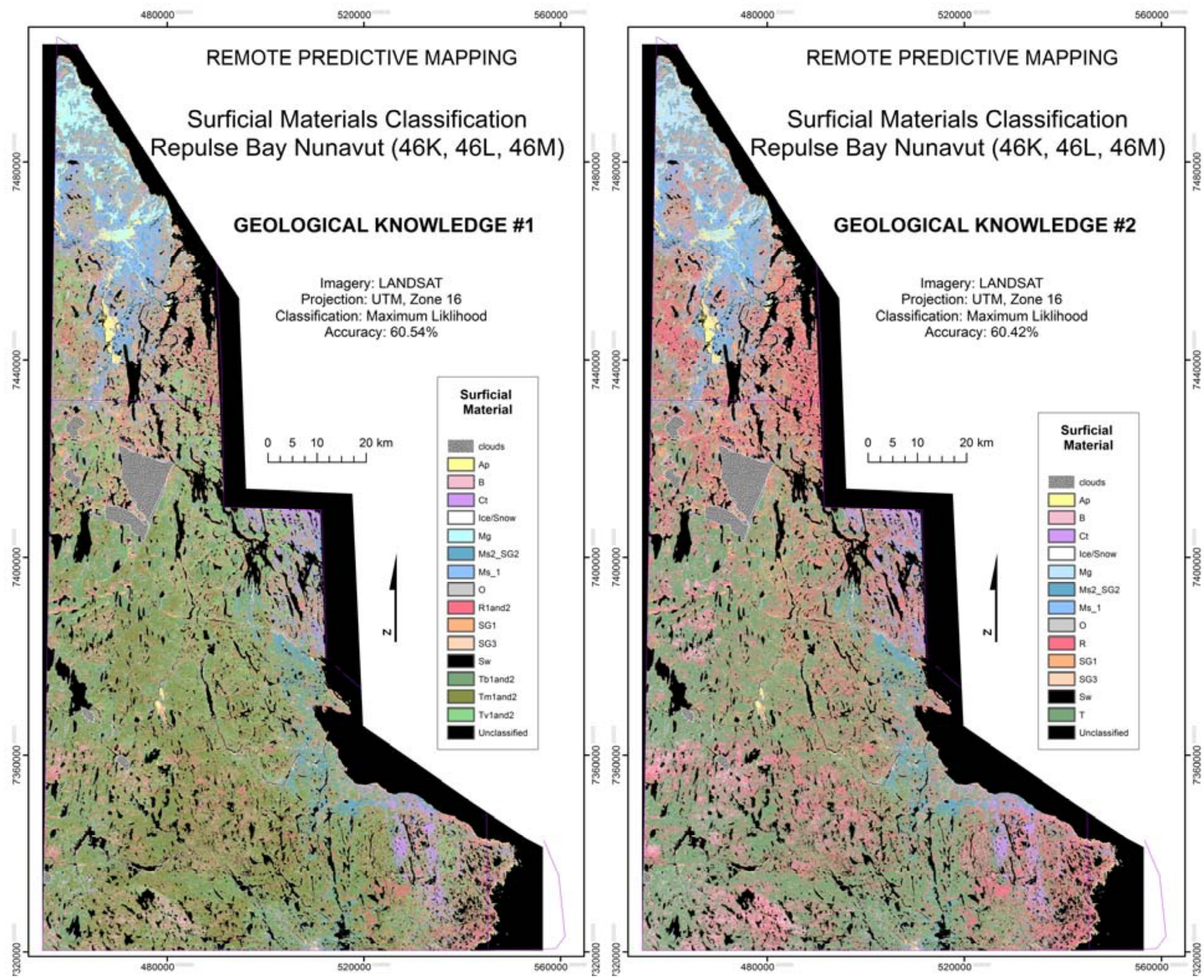

Figure 8a: Best predictive surficial materials based on Geological Knowledge (GK1 and GK2). Classifications produced by Maximum Likelihood Classification algorithm on LANDSAT imagery.

3) Map STAT1: Statistical 1

This classification resulted in widespread abundance of marine sediments in the north - Ms and Mg and some Ap (Fig. 8b). Till classes are dominant across most of the central and southern parts of the map; however, the southern areas contain a fair amount of bedrock as well as modified till. Marine sediments also occur near the central-east coast, with carbonate till in the southeast.

Based on geological field knowledge, marine sediments appear to be classified appropriately. The second marine sediment subset (Ms2) does not show much difference when combined with SG2. As per field observations, this map showed better predictions for the classification of boulder fields, which were mapped in eastern corridors of 46L, and the central portion of the study area. Carbonate till was not classified as abundantly in NTS 46M (northern study area), which is a more accurate representation when compared to field knowledge.

Although the statistical method (i.e. TD value) suggested that R2 and Tm1 are highly confused with one another and thus, could be combined, both units are not the same surficial material: R2 is bedrock while Tm1 is a bouldery till. It is possible that spectral characteristics of classes such as Tm1 and R2 (or R1) will vary with changes in bedrock lithology, but more likely with changes in boulder cover (percentage cover and boulder lithology) and with moisture content. These variations affects spectral signature within like-classes. The STAT1 map did not classify as much exposed bedrock (R1) as seen in the field, and assigned more areas as lichen-covered bedrock (R2). Although R1 is more pervasive in the central region and corresponds more closely to the surficial field mapping observations (Campbell and McMartin 2010, 2011, in prep.), it is under-mapped in the north.

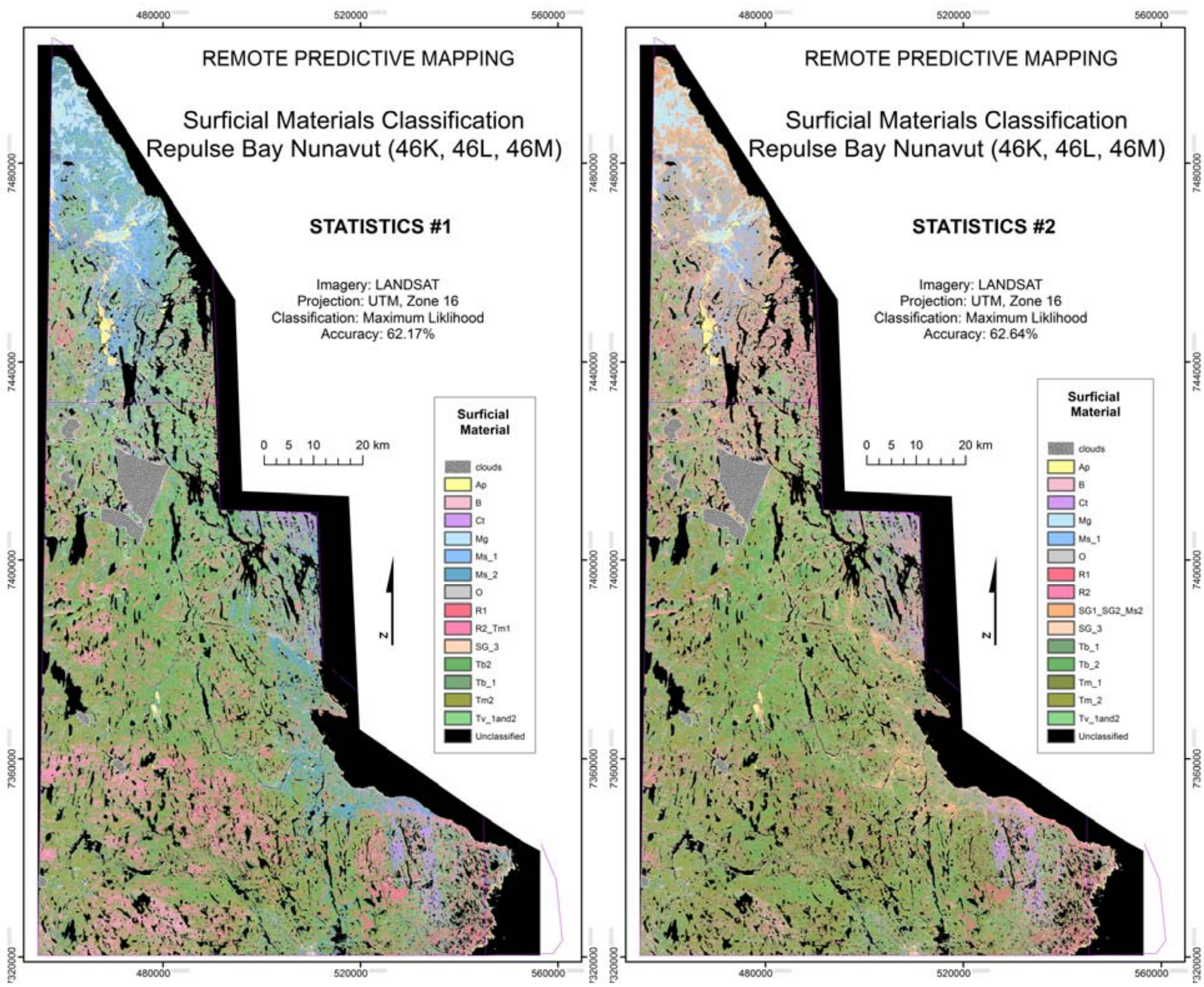

Figure 8b: Best predictive surficial materials maps based on Statistical approaches (Statistics 1, Statistics 2). Classifications produced by Maximum Likelihood Classification algorithm on LANDSAT imagery.

4) Map STAT2: Statistical 2

Unlike the other three maps, this map classified a sizable portion of the north as the merged class SG1+SG2+Ms2 (Fig. 8b). This classification indicates that sands and gravels are also dominant in the north, and along the central coastal areas; while the other maps suggested marine sediments were the dominant classes. This map shows more bedrock and boulder fields in the north when compared to STAT1 and comparable distribution to GK1; while it classified significantly less of these units in the north than GK2. Though less boulders and bedrock are mapped here, they do occur in generally the same regions on all maps; however their abundance is much less pronounced. The carbonate till is classified consistently in the southeastern portion of the map.

This classified map shows the least similarity to the present knowledge of the surficial geology of the region. Carbonate till is classified in the north where in fact it does not occur; very little of the boulder class has been classified; there is no improvement with the mapping of R1 than in the previous map; and the statistically suggested class combination of SG1+SG2+Ms2 is not a sound combination of surficial materials as it does not differentiate marine fine sands from glaciofluvial sand and gravel. The Ms1 class is mapped appropriately, and not much of SG3 class has been classified. The STAT1 class combination produces a map with more till in the main central and southern portions of the study area, when compared to the other 3 maps.

Variability Maps

The four variability maps produced through the RCM algorithm have a maximum of 4 classes that form contiguous areas of variability. A much smaller number of pixels have variability of 5 to 9 classes (Figs. 9a and 9b). Although general trends and areas of relative variability are similar across all four maps, the degree of variability differs. Spatially, less variability occurs in the northern region of all four maps. This suggests the area has a higher classification accuracy. The greatest variability in all four maps occurs in the central part of the study area, and extends into the south for the STAT1 and STAT2 maps (Fig 9b).

Maps GK1 and GK2 show less variability overall when compared to STAT1 and STAT2 maps (Figs. 9a and 9b). These latter maps show more uncertainty in the southern third of the study area as well as in the western part. For example the STAT2 map shows the central area to have more variability than the STAT1, GK1 and GK2. The map showing the least variability in classification is GK2, followed by GK1, STAT1 and STAT2. The low variability of map GK2 suggests this RPM map is the most robust or reproducible of the four classified maps. Interestingly, while the scene boundary is highlighted in maps GK1, GK2, and STAT2, it is much less pronounced in map STAT1. This may be due to the particular class combination used for this classification.

In general, higher variability is present in till and bedrock-dominated areas (as per classification maps), while less variability is evident in areas mapped as marine sediments and/or sands and gravels, as well as carbonate tills. Regions dominantly classified as till and bedrock are mapped as being more variable and less certain in maps STAT1 and STAT2. It is important to note that the variability maps clearly indicate the effects of an unbalanced mosaic as linear discrepancies on these maps parallel that of the LANDSAT cutlines to create the mosaic (ref. Fig. 4). This will be further discussed in the following section.

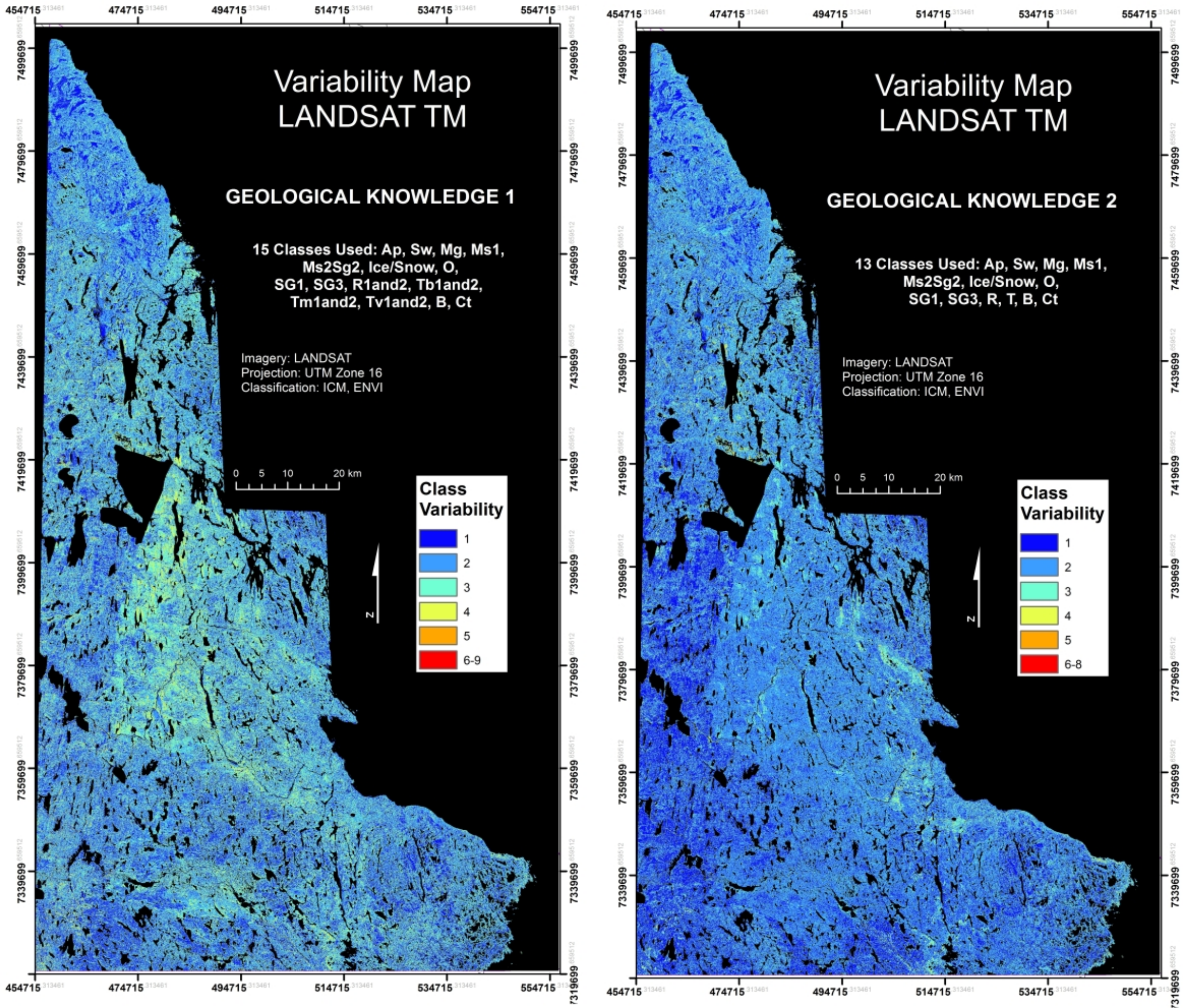

Figure 9a: Variability maps for 2 predictive maps based on Geological Knowledge (GK1 and GK2), indicating areas of more class variability and confusion (warmer red tones) and regions of less variability and more certainty (cooler, bluer hues).

Summary and recommendations

An initial 21 surficial materials classes were used in a supervised classification (single pass maximum likelihood classification - MLC) to classify LANDSAT imagery of the study area west of Repulse Bay, Nunavut. From this, classes were removed and/or combined based on two approaches: 1) geological knowledge, which compared MLC predictive maps to Quaternary knowledge of the area, and 2) statistical, which strictly analyzed statistics of confusion matrices. Two maps, representing the most optimal results from each approach were selected and variability maps were produced on those class combinations using the Robust Classification Method (RCM) to determine which class combination and resultant classification was less variable, and in turn more accurate.

Based on the variability maps produced for each of the 4 “best” MLC classification maps, it appears that a geological knowledge-based approach to produce remote predictive maps is more suitable for mapping surficial materials in the Repulse Bay area. The variability maps based on a statistical approach (STAT1 and STAT2) show a generally greater degree of variability for the mapped region, in comparison to the variability maps based on a geological knowledge-based approach (GK1 and GK2). As demonstrated in the geological knowledge-based approach, the RPM process must include a direct input by Quaternary geologists. This proved to be most useful in merging, splitting, or removing classes from classifications, using geological criteria, and in turn, coming up with classifications more representative of the actual geology. The statistical method to create these maps suggested merges of classes that were not similar by geological standards, i.e. merging R2 and Tm1 – a bedrock unit with surrounding till combined with modified till containing sand and gravel.

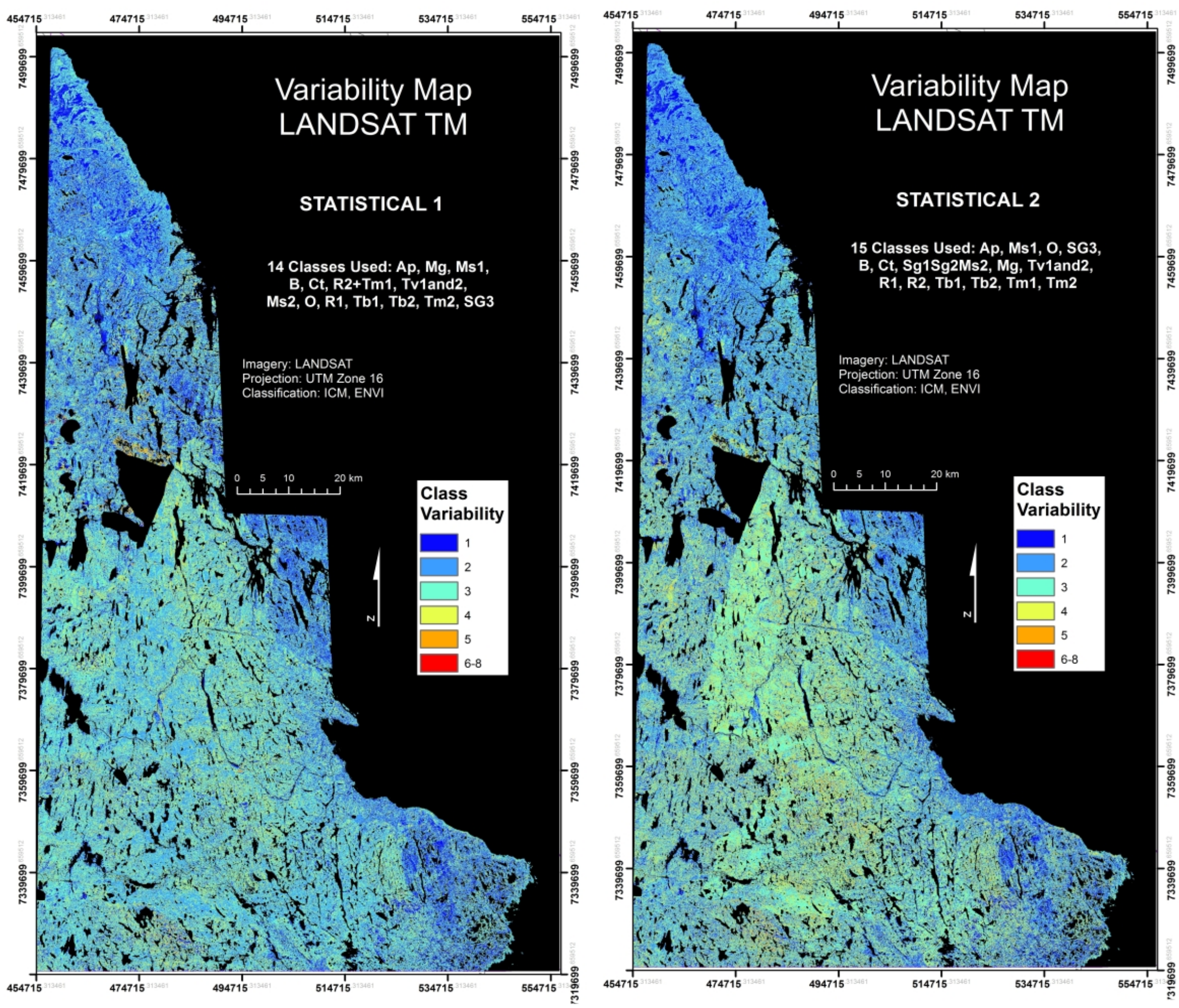

Figure 9b: Variability maps for 2 predictive maps based on Statistical approaches (STAT1 and STAT2), indicating areas of more class variability and confusion (warmer red tones) and regions of less variability (cooler, bluer hues).

It is important to note however that a direct comparison between predictive maps derived from remotely sensed data and geological maps produced through fieldwork and air photo interpretation will not always be conclusive because the processes behind producing the RPM and geological maps are quite different. LANDSAT data capture information at much higher spatial detail and high spectral variability, on a pixel-to-pixel basis. Some of this variability is noise and some is signal that represents the underlying bedrock (when exposed) as well as cover (surficial materials and vegetation).

A recommendation for further work is to use a well-balanced LANDSAT mosaic. The effects of an unbalanced mosaic were not as apparent in the classification maps versus corresponding variability maps produced on the same LANDSAT imagery. The scene boundaries are the result of seasonality and/or atmospheric differences between neighbouring images, which were not acquired during consistent conditions, therefore causing ill-correspondence for similar materials on either side of the boundaries. It is unlikely that differences in surficial materials are directly correlated to where scene boundaries occur. This issue is being addressed by the Canadian Centre for Remote Sensing (CCRS) and the Geological Survey of Canada (GSC).

Another recommendation is to use other parameters such as terrain (digital elevation models), texture, or other remotely sensed data, since LANDSAT provides one element of the surficial environment (e.g, RADARSAT). These data would provide additional parameters to classify the surficial materials and not only rely on spectral information from LANDSAT imagery.

Finally, although the Robust Classification Method (RCM) was used to produce the variability maps based on the combination of classes defined for the 4 “best” MLC maps, majority classification maps derived from the RCM were not discussed in this report. RCM is a technique that helps to present a more robust estimate of overall classification accuracy and to level out statistical variance in the training areas (Harris et al., 2012). The RCM should form an integral part of the supervised classification, and be used at the beginning of the classification process.

In conclusion, RPM will not substitute field mapping by the Geological Survey of Canada. Rather, this new approach is intended to be a supportive tool to direct, optimize, and perhaps enhance conventional geological knowledge. The supervised classification maps presented in this Open File are based on a physical parameter of the surface (spectral reflectance), while geological maps synthesize many parameters including photo-geologic variables (i.e. tone, texture, shape, pattern, context and association), field observations (e.g., of earth materials and geomorphic processes) and expert knowledge. The use of remotely sensed data for surficial materials mapping thus compliments the geological mapping process. The maps presented in this Open File and the methodology used to produce them have aided the GSC’s mapping efforts of the greater Wager Bay region of Nunavut, which is one of the outputs for the Geo-mapping for Energy and Minerals (GEM) program.

Acknowledgements

This research was conducted as part of an MSc thesis by U. Wityk at University of Waterloo and supported by a Research Affiliate Program grant from Natural Resources Canada, NSERC, the Northern Scientific Training Program of Aboriginal Affairs and Northern Development Canada, the University of Waterloo as well as by the Canada Foundation for Innovation and the Ontario Research Fund. Data was collected as part of the Geo-mapping for Energy and Minerals (GEM) Wager Bay Surficial Geology Activity under the Melville Peninsula Project at the Geological Survey of Canada. The authors would like to acknowledge the work of Sam Leiff of Blackbridge Geomatics Corp. for mosaicking satellite scenes into usable formats. Dave Huntley (GSC-Vancouver) is thanked for providing a careful review.

References

Aylsworth, J. M. and Shilts, W. W. 1989. Glacial features around the Keewatin Ice divide: Districts of MacKenzie and Keewatin; Geological Survey of Canada, Paper 88-24, 21 p.

Campbell, J.E. and McMartin, I. 2010. Surficial geological mapping in the Wager Bay area, Nunavut - filling in the gaps; *in* Palmer, E. (ed.), 38th annual Yellowknife Geoscience Forum, Abstracts of talks and posters, Northwest Territories Geoscience Office, Volume 2010, p. 6-7.

Campbell, J.E. and McMartin, I. 2011. Regional glacial patterns in the Repulse Bay - Wager Bay area, Nunavut; Geohydro 2011, Proceedings of the joint meeting of the Canadian Quaternary Association and the Canadian Chapter of the International Association of Hydrogeologists, Quebec City, Quebec, August 2011, Doc. 2156.

Campbell, J.E., McMartin, I., Tremblay, T., Wityk, U. and Dredge, L.A. 2011. New insights on the surficial geology of the Repulse Bay area, Nunavut: implications for mineral exploration; *in* Fischer, B.J. and Watson, D.M. (eds), 39th annual Yellowknife Geoscience Forum, Abstracts of talks and posters, Northwest Territories Geoscience Office, Volume 2011, p. 24-25.

Campbell, J.E. and McMartin, I. *in prep*. Surficial geology, SW Lefroy Bay, Nunavut; Geological Survey of Canada Canadian Geoscience Map (preliminary, scale 1:100 000).

Frazier, P.S. and Page, K.J. 2000. Water Body Detection and Delineation with LANDSAT TM Data; Photogrammetric Engineering and Remote Sensing, Vol. 66, No. 12, p. 1461-1467.

Grunsky, E., Harris, J.R., and McMartin, I. 2006. Predictive mapping of surficial materials, Shultz Lake area (NTS 66A), Nunavut, Canada; Geological Survey of Canada, Open File 5153, 49 p.

Grunsky, E., Harris, J.R., and McMartin, I. 2009. Predictive Mapping of Surficial Materials, Schultz Lake Area (NTS 66A), Nunavut, Canada; Society of Economic Geologists, v. 16., p. 177-198.

Harris, J.R. 2008a. Introduction; Chapter 1 *in* Remote Predictive Mapping: An Aid for Northern Mapping, (ed.) J. R. Harris, Geological Survey of Canada, Open File 5643, p. 1-4.

Harris, J.R. 2008b. Remote Predictive Mapping: A Strategy for Geological Mapping of Canada's North; Chapter 2 *in* Remote Predictive Mapping: An Aid for Northern Mapping, (ed.) J. R. Harris, Geological Survey of Canada, Open File 5643, p. 5-27.

Harris, J., Grunsky, E.C., He, J., Gorodetzky, D., and Brown, N. 2012. A robust, cross-validation classification method (RCM) for improved mapping accuracy and confidence metrics; Canadian Journal of Remote Sensing, Vol. 38, No. 1, p. 69-90.

Huntley, D., Campbell, J., Harris, J., McMartin, I., Wityk,U., and Chow, W. *in prep.* Remote Predictive Mapping of Surficial Earth Materials: Wager Bay North Area, Nunavut; Geological Survey of Canada, Open File 7118.

McMartin, I., Campbell, J.E., Dredge, L.A., Corrigan, D. and Huntley, D. 2012. Glacial history and dispersal patterns north of Wager Bay, Nunavut; *in* Watson, D.M. (compiler), 40th annual Yellowknife Geoscience Forum Abstracts of talks and posters, Northwest Territories Geoscience Office, Volume 2012, p. 29.

McMartin, I., Campbell, J.E., Dredge, L.A. and McCurdy, M.W. 2013. Till composition and ice-flow indicators West of Repulse Bay: 2010 and 2011 results from the GEM Wager Bay Surficial Geology Activity; Geological Survey of Canada, Open File 7288.

Paul, D., Hanmer, S., Tella, S., Peterson, T.D., and LeCheminant, A.N. 2002. Compilation , bedrock geology of part of the western Churchill Province, Nunavut – Northwest Territories; Geological Survey of Canada, Open File 4236, scale 1:100000

Prest, V.K., Grant, D.R. and Rampton, V.N., 1968. Glacial Map of Canada, Geological Survey of Canada, Map 1253A, Scale 1:5 000 000.

Richards, J.A., and Jia, X., 1999, Remote sensing digital image analysis: An introduction, 3rd ed., revised and enlarged: New York, Springer-Verlag, 363 p.

Schetselaar, E.M., J.R. Harris, T. Lynds and E.A. de Kemp, 2007. Remote Predictive Mapping (RPM): A Strategy for Geological Mapping of Canada's North. Geoscience Canada. Vol 34. No 3 and 4, 93-111.