Naming Convention for ECCC’s Shoreline Segmentation

|  |  |
| --- | --- |
| **Created by** | Vincent Boulianne and Anne-Marie Demers |
| **Creation date** | 2021 June |
| **Last Modified (Author)** | 2021-09-24 |
| **Other Document** |  |
|  |  |

INDEX

[Definitions 1](#_Toc83390452)

[Introduction 1](#_Toc83390453)

[Objectives 1](#_Toc83390454)

[Existing literature and revisions of naming conventions in existing datasets 1](#_Toc83390455)

[Methodology 2](#_Toc83390456)

[Approach 2](#_Toc83390457)

[Sequential naming of segmentation options 2](#_Toc83390458)

[By ordering field: 2](#_Toc83390459)

[By proximity (nearest neighbour): 2](#_Toc83390460)

[By land mass 3](#_Toc83390461)

[Challenges and Issues 3](#_Toc83390462)

[Script “Name Segments” 4](#_Toc83390463)

[Method “By Ordering Field” 5](#_Toc83390464)

[Method “By Proximity” 6](#_Toc83390465)

[Set sector\_first\_segment 6](#_Toc83390466)

[Results and Discussion 7](#_Toc83390467)

[Method 1: By Ordering Field. 7](#_Toc83390468)

[Method 2: By Proximity 8](#_Toc83390469)

[Recommendations 8](#_Toc83390470)

[Annexes 8](#_Toc83390471)

[Previous work 12](#_Toc83390472)

[Code alphanumérique 12](#_Toc83390473)

# Definitions

**Segment:** Part of the shoreline feature class. The segments are multipart and can include

**Sector:** Individual cell of the reference grid.

**Reference Grid:** Polygon feature class covering the extent of a shoreline feature class.

# Introduction

[TO DO: INTRO]

## Objectives

The purpose of this work is to develop a naming convention for ECCC’S shoreline segmentation nationwide by comparing existing data and naming standards.

The qualities searched for in this convention are thus:

|  |
| --- |
| * Stable in time |
| * Based on existing systems |
| * Geographic references are known |
| * Geographic references can be applied to multiple scales |
| * Geographic references can be applied throughout Canada |
| * Sequential numbers |
| * Easy to update/modify |

## Existing literature and revisions of naming conventions in existing datasets

Across the years, there have been many naming conventions (see annexes; table 1) following different paradigms.

Most datasets in the annexe (table 2) were acquired from NEEC's pre-existing database. Other datasets were explored from this site <https://canadiangis.com/data.php#Canadian-Data> (source to multiple Canadian geospatial resources).

Some follow and accepted practice of devolution from the general to the specific. For example, you could name a segment *BC-VAN-BI-SP-1*. The meaning would be British Columbia – Vancouver – Burrard Inlet – Stanley Park- 1. There are many advantages to this approach. Others are numeral based but with the same geographical scale principle so that it provides a level of granularity to the naming structure. The main drawback of this nomenclature is that is it dependent on the presence of named geographical features. Those are not equally spatially distributed and may be subject to change.

Following discussions within the Shoreline workgroup, consensus was reached towards a grid-based and sequential naming of the shore segments within a sector. A grid system would be stable in time and not dependent on the presence and variability in space and time of named places. It can be applied uniformly throughout Canada and it would also provide as a reference system. Any update in a particular localized grid would necessitate at most of the renumbering of only the segments within this grid.

Two grid systems were evaluated: ECRC/SIMTEC and the national topographic system (NTS) from Natural Resources Canada. NTS was selected based on its national coverage, neutrality regarding to geographic names that might be subject to change and its stability. The number of segments would be manageable and should remain in the hundreds. Most importantly, as the grid already covers the whole of Canada and is maintained by Natural Resources Canada, the cost of using this grid is low.

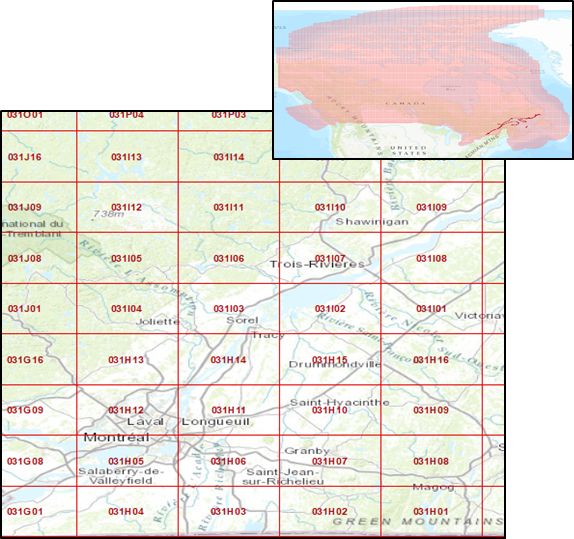


Figure 1. Exemple and coverage of the NTS Grid

# Methodology

## [TO DO: Intro section]

## Approach

* Grid-based
* Sequential numerical naming of segments within specific grid

## Sequential naming of segmentation options

Once the segments are grouped by sector. There are different options that can used to name the segments sequentially. It is essentially a way to establish a spatial relationship between each segments. Ideally, close segments would be numerically related. Different options were evaluated.

### By Ordering Field:

1. Use ordering field in which the segments have been already spatially sorted manually or automatically.

### By Proximity (nearest neighbour):

1. Find nearest neighbour within the grid and name it sequentially starting with the leftmost segment.

Options available are:

* 1. To find spatially the first segment through the proximity of segments in adjoining sectors
  2. To select the segment with the lowest value in an ordering field. Usually, segment should have been created in sequence. This does not preclude the possibility of an outlier.

### **By Landmass**:

1. List all land mass with a unique identifier
2. Group all segments by landmass
3. Identify each segment in a landmass sequentially starting from the segment to the bottom left.

### By Network/Topology:

This method would seek to ascertain spatial relationship through membership to a specify route. Multiple “route” would need to be created for each landmass. As is discussed below, this method would present multiple challenges.

## Challenges and Issues

* The first issue concerns the simple spatial relationship between landmasses. How can the segments be named according to a logical succession? Any script based approach would need to have a method to deal with the proximity of each landmass or it might start numbering segments in a runaway process further away from the main land.

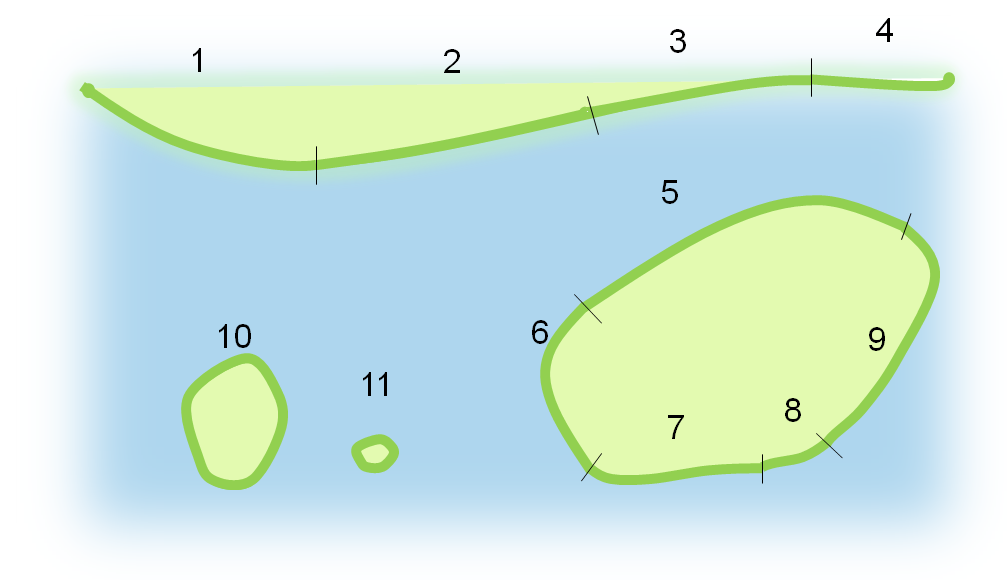


Figure 2. Sequential Naming

* One of the main issues is the uneven nature of cut-offs in the segments as they have been created. It cuts through landmasses and group shorelines according to nature of the sediments and morphology present. The subsequent problem is one of spatial relationship. What would be the closest segment? It is an issue that would particularly affect the idea of sorting segments by landmass or the topology/network approach.

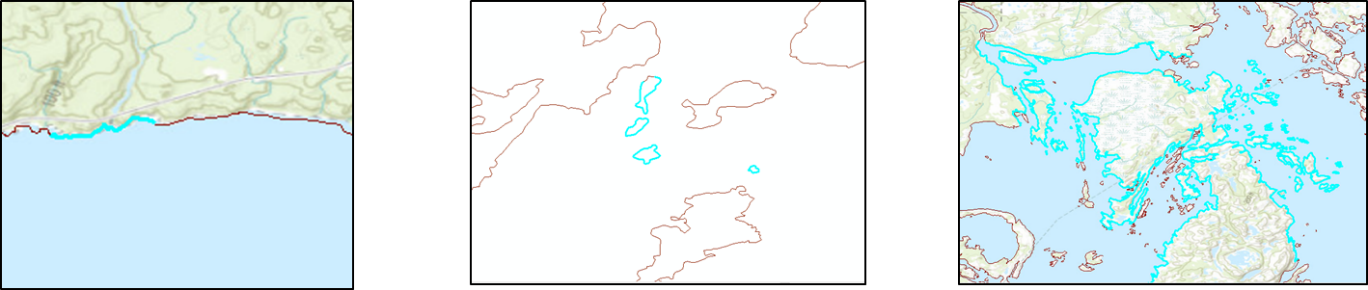


Figure 3. Individual segment

* An approach by grid may sometimes lead to a situation where a same landmass will have segments that are part of two or more sectors in a grid. This is an inevitable result of choosing to divide segments by grid unless it follows the shore closely and its cell size is small.

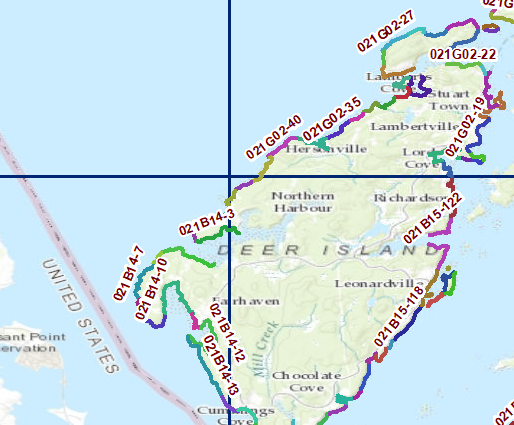


Figure 4. Division of Deer Island (NB) in three separate sectors of the NTS grid

* One problem that has cropped up in the development of the method is choosing the right starting point. Indeed, dividing segment by segment may break spatial relationships. In addition, the NTS grid is not strictly related to the shoreline.

# Script “Name Segments”

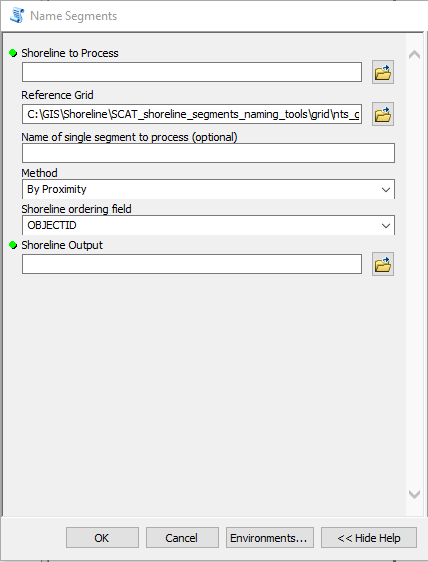
A python script has been created with different methods. It has been encapsulated in an ArcGIS geoprocessing tool for ease of use. So far, two methods have been implemented: “By Ordering Field” and “By Proximity”. The third method that would be a hybrid adding the benefit of “ordered island” may be arduous if not impossible to implement. The segments being multipart and on different landmasses would require a selection to be made on the landmass membership of that particular group. The fourth method would seem to be difficult to implement as the standard geometry of the shorelines features classes currently stand as segments.

The script can be found at: [\\int.ec.gc.ca\sys\InGEO\GW\EC3140EnvEmer\_UrgEnv\SensitivityMapping\Analysis\Tools\SCAT\_shoreline\_segments\_naming\_tools](file:///\\int.ec.gc.ca\sys\InGEO\GW\EC3140EnvEmer_UrgEnv\SensitivityMapping\Analysis\Tools\SCAT_shoreline_segments_naming_tools)

It should also be available on **Github** at <https://github.com/vboulianne/ECCC-Shoreline-Segmentation>

The whole folder is part of the script. There is an example in the folder ‘/example’. The default grid in grid is the NTS grid projected to GRS80 NAD83. The python script is included in “/code”.

One added to the ArcToolBox, it should appear as “*SCAT Shoreline Segment Naming Tools*” in the tool list. There is at this time only the tool “Name Segments” in this toolbox.

**Shoreline to Process:** The shoreline feature class to process. Ideally, it should already be in the NEEC standard data model.

**Reference Grid:** The grid with which the segments will be divided. By default, the NRCAN NTS is used. Any grid covering the shoreline can be used as ownership is determined by a spatial join.

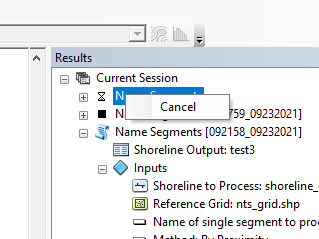
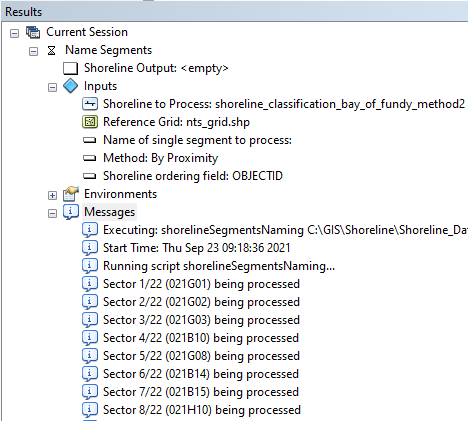
**Name of single segment to process:** This optional parameter enables the processing of a single sector. It will create a new feature class will only the written sector having been processed.

**Method:** There a two methods that can be used in this script. “By Ordering Field” will name segments sequentially in an ascending order of the data in the chosen field. “By Proximity” will name segments according to their respective proximity (see below for further information)

**Shoreline ordering field:** To use with method “By Ordering Field” to name sequentially or “By Proximity” to choose the first segment.

**Shoreline Output:** The feature class output.

Once running, the script should take about 3 minutes for 1500 segments. The progression can be followed in the Results Pannel under Current Session/Name Segments/Messages. Right-clicking on the running process in the results page allows the script running to be stopped.



## Method “By Ordering Field”

This is the simplest method. After the segments have been divided by sector, it will process the segments in ascending order according to the ordering field. The ordering field can be OBJECTID or a sequential number field populated by the provider or a manually edited intermediate number created to correct the sequence.

## Method “By Proximity”

This method names segments according to their respective proximity. Adjoining segments have the distance “0” so a continuous shoreline will be named sequentially. There are, however, two ways of choosing the first segment. One is to use the lowest value in the ordering field (OBJECTID, etc.). In that case, the result quality depends entirely on the appropriateness of the provided field: if the number are not spatially sequential, quality will suffer. The other is to use a custom field “sector\_first\_segment” added and populated by the operator through manual edits. Although it is more labour intensive, this approach deal with the difficult issue of determining what segments to use as a starting point in each sector. The following describe the method to set the first segment in each sector.

### Set sector\_first\_segment

The first step is to add a field (INT) named sector\_first\_segment.

An edit session has to be started with the reference grid (root folder/Reference.lyr) in the same window. For each sector, each segment that it is estimated to be the first of the sector is selected and the field *sector\_first\_segment* is modified to “1”.



The general rule would be to select a first sector including a landmass on the starting side of the shoreline feature class and then to choose the first segment in this sector. We would choose afterwards each segment according to logical sequence according to the first sector. In the figure below, it can be seen that the first segment would be to the left for the northern shore and to the right for the southern shore. It is important to note that first segments might not be entirely contained within a grid. Segments are part of a sector if their geometric centre is included in that sector.

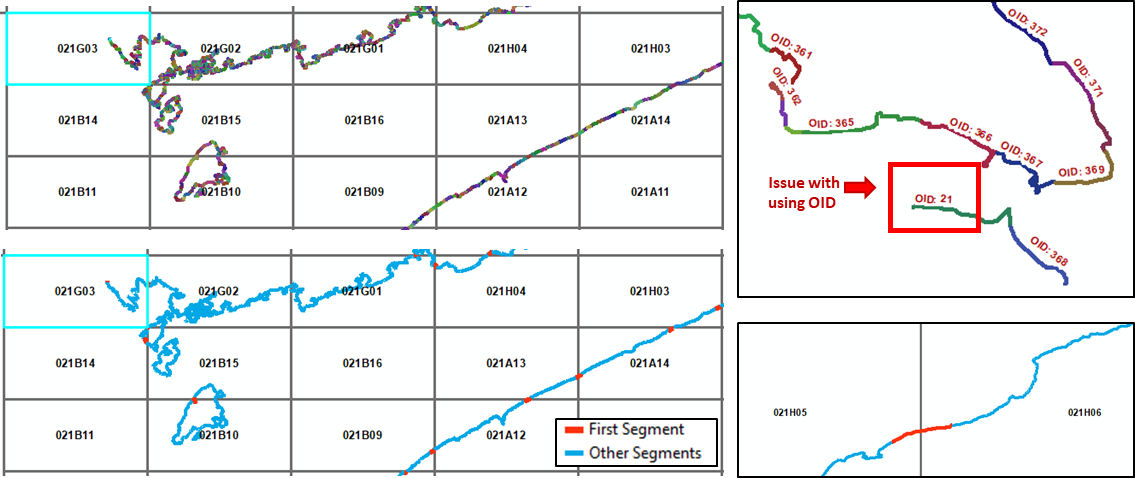


Figure 5. Select sector first segment

Once the edit has been made, it can be saved. The script may stop if there are more than two “first segment” in a same sector.

# Results and Discussion

## Method 1: By Ordering Field.

In many cases, the segments have been created in order and the OID of the segments along a shoreline are already sequential. After being bundled, they can be named in ascending order. There may be some odd segments named incorrectly but in many cases, the method seems to work. The result is adequate to the exception of a few outliers.

The heart of the script is rather simple and is fast:

  for sector in sector\_list: # Loop by sector

            # Select all segments within grid sector

            sql\_clause\_ord = (None, "ORDER BY " + shoreline\_order\_field + " ASC")

            # SQL query create a subgroup by sector and afterward named sequentially in ascending order of their "OBJECTID"

            num\_seq = 1

            with arcpy.da.UpdateCursor(shln\_to\_process, ["NAME\_EN"], where\_clause="NTS\_SNRC='" + sector + "'", sql\_clause=sql\_clause\_ord) as cursor:

                for row in cursor:

                    row[0] = sector + "-" +  str(num\_seq).zfill(4)

                    cursor.updateRow(row)

                    num\_seq += 1

The following are some a few screenshots of the first method on the segmentation of Bay of Fundy.



The ordering field could also be a manually an altered OBJECTID list. One approach would be to have this field populated at creation time by the team already creating the rest of the data. Provided that the process still involves manual editing, it would be rather simple to add the requirement to enter a sequential number for each segment at creation time.

## Method 2: By Proximity

One case has been found in a sector that the first segment was a small island in the middle of the sector whose OBJECTID was very different from close segments.

Select a first segment according to an ordering field or an added field “**sector\_first\_segment**” with edited manually to “1”. The following segments are afterwards named based on proximity to the previous segment with the tool “*Generate Near Table (Analysis*”. To ensure that a segment is not chosen twice, the script has a temporary feature class of segments remaining that will be processed one after the other until it is empty. Adjoining segments have the distance “0”; this ensures that touching segments will be named one after the other.

When the shoreline is simple, the sequential naming is appropriate. The script will process a whole landmass before processing other landmasses (fig. 1). This means, however, that since the continent can have a long interrupted shoreline, the segments on islands may have a number much different from “continental segments” close to it.

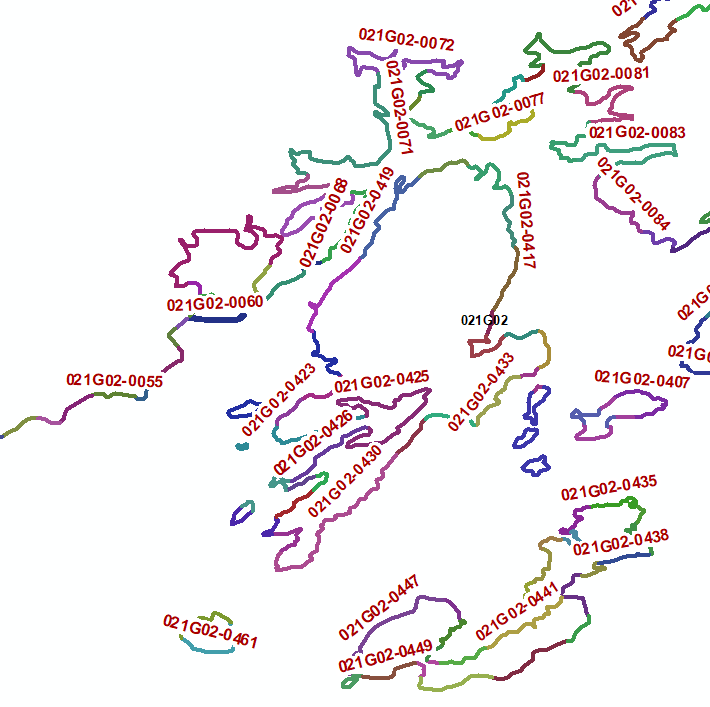


Figure 6. Islands have numbers in the 400s

# Recommendations and remaining work

[TO DO]

# Annexes

Table 1. Comparaison des identifiants de segments entre différentes sources de données (french)

| **Source de la donnée** | **Représentation graphique** | **Description du label** |
| --- | --- | --- |
| Le Manuel TERR-2ème édition  **Page 1.10**  Environnement Canada | AH-1  AH-2  AH-4  AH-5  …  AH-3 | AH-24 où « AH » est lié à un lieu géographique (Arichat Harbour) et où « 24 » est issue d’une représentation séquentielle le long de la rive. |
| Atlantic Canada  **\\GW\EC3140EnvEmer\_UrgEnv\Application\Data\EEMAP.gdb**  O14Oceans\_ShoreCharacter\_AR\_1  O14Oceans\_ShoreCharacter\_AR\_1\_20140211  O14Oceans\_ShorelineClass\_AR\_1  O14Oceans\_ShorelineClass\_AR\_1\_20140211  Environment Canada - National Environmental Emergencies Centre |  | [Champs : Segment\_ID]  BA-01,BB-01,BC-01,… . l’attribution des lettres semble être faite de façon arbitraire selon un ordre alphabétique.  Les ID vont de AB-01 à ZZ-09  La numérotation consécutive varie. Parfois, cela va de 1 à 10, de 1 à 50, etc. |
| Québec  **Q:\GW\EC3140EnvEmer\_UrgEnv\Service\Data\EC\_QC\_Shoreline\Shoreline\QC\_Shoreline.gdb**  O14Oceans\_ShorelineClass\_QC\_l  Shoreline classifications are prepared by Environment Canada  to categorize marine and freshwater shore types |  | [Champs : Name]  AO-41,POC-02,…  Les lettres sont liées à un lieu géographique. Par exemple, pour un segment autour de l’île de Laval, on a LAV-38. Pour Montréal, l’acronyme MTL est utilisé. Pour les segments dans l’arrondissement de Verdun à Montréal, l’abréviation VDN apparait. La numérotation consécutive peut aller jusqu’à 100, dépendant de la grandeur de la zone. |
| Ontario  **Q:\GW\EC3140EnvEmer\_UrgEnv\Application\Data\EEMAP.gdb**  O14Oceans\_ShorelineClass\_ON\_1  O14Oceans\_ShorelineClass\_ON\_1\_20140211 |  | [Champs : Name]  Les lettres semblent être liées à un lieu géographique. La structure du code alphanumérique n’est pas fixe. Le code est parfois composé de 2 lettres uniquement (LH-504) ou bien avec des lettres en plus qui sont délimitées par des tirets supplémentaires (LH-EASI-1). C’est probablement pour apporter des spécifications supplémentaires sur le lieu géographique. La numérotation consécutive peut aller jusqu’à 1500 (Ls-1 à Ls-1450). |
| Colombie Britannique  **Q:\GW\EC3140EnvEmer\_UrgEnv\Application\Data\EEMAP.gdb**  O14Oceans\_ShorelineClassKitimat\_BC\_1  Environnement Canada - Centre national des urgences environnementales | 4  …  1  5  2  3 | [Champs : Objectid]  Aucune identification alphanumérique |
| Amérique du Nord  [**http://response.restoration.noaa.gov/maps-and-spatial-data/download-esi-maps-and-gis-data.html#Canada**](http://response.restoration.noaa.gov/maps-and-spatial-data/download-esi-maps-and-gis-data.html#Canada)  ESI maps NOAA | 4  …  1  5  2  3 | [Champs : Objectid]  Selon les PDF visionnés et le shapefile ESIL représentant les rives de San Francisco Bay, il ne semble pas y avoir de codes alphanumériques attribués aux segments. |

Source : Laurent Saint-Arnaud (2015). « Appendix\_NamingConvention\_LSA-TableauLabel.docx »

Tableau page suivante : il s’agit du document « ShorelineClassification\_DataIdentifierComparison\_NamingConvention.xlsx » ici Z:\SensitivityMapping\ShorelineData\DOC\ShorelineClassification

(Ce document a été archivé)

Table 2. Segmentation exemples

| **Shoreline Classifcation (Name)** | **Owner** | **Segment ID Type** | **Creation Year (year of overflight)** | **Update Available (Date)** |
| --- | --- | --- | --- | --- |
| ShorelineClassification\_AR\_EC\_20150707 | ECCC | Alphanumerical Unclear what the letters (2) stand for… there seems to be no relation with administrative separations as foor ON or QC data. |  | Maybe (2016) | Potential overflight by Jason Duffe team? |
| ShorelineClassification\_BC\_GovBC\_20150708 | SHRUNT\_LN\_line | WHSE\_ENVIRONMENTAL\_MONITORING\_SHZN\_SHORE\_UNIT\_CLASS\_LINES\_SV | Gov BC | ***PHYIDENT* is a 13-character slash-delimited (/) concatenation of the shoreunits REGION (13), SECTION, UNIT and SUBUNIT that uniquely identifies the shoreunit for use in the shorezone mapping system**, e.g., 01/02/0345/01. These are unique numbers assigned to a particular region, and section used for administrative purposes. UNIT is a number assigned to the primary unit of measure in the shorezone coverage. This main subdivision of a SECTION forms the "along-shore" physical unit. There are a maximum of 9,999 UNITS within a SECTION. SUBUNIT is a number used to show the across-shore component of the shoreline. In the PHYIDENT, the SUBUNIT portion is at the end of the phyident and typically has a leading "0". A subunit of 00 signifies a homogeneous unit while subunits identified with 01, 02, etc. signify variants or anomalies in the unit, usually features that have only a short alongshore length, e.g., 0,1,2.  PROJECT CODE is a unique three-to-six letter acronym code assigned to a particular very large administrative project area in the province of BC. Project codes are used for administrative purposes for the identification and grouping of regions and sections, e.g., WCVI (for West Coast Vancouver Island), MIDC (for MidCoast). | 1991-1996 | Yes (2014) | SHRUNT\_LN\_line from BC's data catalogue seem to hold the latest version of Shore Zone's BC Shoreline - ID information found from this layer. |
| ShorelineClassKitimat\_BC\_2015 | ECCC | Numerical | 2013-2015 | Yes (2018) | See BurrardInlet\_LCC\_Final\_20180531 (unknown) | See Kitimat\_HaidaGwaii\_LCC |
| ShorelineClassification\_ON\_EC\_20150706 | OntarioShorelineSegmentation | ECCC | Alphanumerical 1- Seems like (i.e. based on comparative research) first letters ID the hydrological area (e.g. GB = Georgian Bay; HH = Hamilton Harbour, SLR = Saint-Lawrence River). If the land is separated (e.g. island) there is a second/third set of letters (e.g. GB-BADI = Badgeley Island). Exceptions for some areas where first letters represent administrative division (e.g. LAN = Lancaster). 2- Numbers complete the unique ID (starting at 01). |  |  |
| ShorelineClassification\_QC\_EC\_20160812 | O14Oceans\_ShorelineClass\_QC\_l | ECCC | Alphanumerical 1- Letter ID that may be a reference to the an administrative separation (e.g. ANM = Sainte-Anne-des-Monts, MSP = Mont-Saint-Pierre…); 2- Numbers complete the unique ID (starting at 01). |  |  |
| BurrardInlet\_LCC\_Final\_20180531 | ECCC |  | 2017 |  |
| Kitimat\_HaidaGwaii\_LCC | ECCC | Numerical | 2015 |  |
| O14Oceans\_ShorelineClass\_NO\_1\_Edited20170424 | ECCC | Numerical | 2010-2014 | Yes (2015-2016) | Nunavik overflown by Jason Duffe team |
| Queen Charlotte Straight | ECCC |  | 2018 |  |
| Régions Hydrographiques  (Il existe 16 régions, les ZIP sont téléchargés dans le dossier ShorelineData, mais seul un fichier a été décompressé puisqu'ils sont très lourds.) | MERN | Alphanumerical Seems like a random ID system (e.g. f34ece8472d14acbb27dcd8a0979de50) | (Date Source) 2006 | (Mise-à-jour) 2017 |
| bcmca\_eco\_physical\_coastalclasses\_marxan\_data | BCMCA | Numerical |  |  |
| Lake\_Simcoe\_Shoreline.gdb | Ontario MNRF | N/A | (Publication) 2015 |  |
| What is the source dataset (i.e. same as BC MoE - ShoreZone)? | WCMRC | Numerical (e.g. 01/05/0220/00, 01/05/0221/00, etc.) | Must validate data source |  |

Source: « ShorelineClassification\_DataIdentifierComparison\_NamingConvention.xlsx » in Z:\SensitivityMapping\ShorelineData\DOC\ShorelineClassification

# Previous work

### Code alphanumérique

(Laurent St-Arnaud – PDI, Hiver 2015)

(source : St-Arnaud, L. (2015). *Rapport de stage*, présenté à Caroline Rousseau et revu par Anne-Marie Demers (CNUE), Université de Sherbrooke, Géomatique appliquée à l’environnement, T2.)

“D’autre part, durant mon stage, j’ai eu à me pencher sur la manière d’attribuer d’un code alphanumérique aux segments constituants le rivage de la province de la Colombie-Britannique. Le but étant d’assigner un code unique à chaque segment afin de faciliter leur gestion. Évidemment, un code alphanumérique est composé d’une partie alphabétique et d’une partie numérique. Ce travail s’est alors divisé en fonction de ces deux parties du code afin d‘explorer les différentes possibilités.

Avant tout, il a fallu faire un travail d’exploration et d’observation au tout début afin d’analyser ce qui a déjà été fait pour les autres provinces. J’ai alors regardé les codes de données existantes pour en tirer certaines conclusions et offrir des suggestions quant à la manière de représenter le code alphabétique. La plupart utilisaient des acronymes comme code alphabétique. Par exemple, pour les segments de rivage de l’île de Montréal, un segment peut se nommer MTL-1, MTL-2, MTL-3, etc. Une option était d’utiliser des codes arbitraires tels AA-1, AA-2, AB-1, etc.

Le travail concernant la partie alphabétique était alors essentiellement de l’exploration et de l’analyse. Pour ce qui de la partie numérique, je me suis intéressé à automatiser la numérotation séquentielle de segments. On pourrait ainsi attribuer des numéros à des segments se trouvant dans une zone d’une manière rapide et simple.

Pour ce faire, j’ai développé un modèle dans le logiciel Arcgis à l’aide de la fonctionnalité « Model builder ». « Model builder » permet de créer des modèles en rassemblant différentes fonctionnalités du logiciel. Mon modèle utilise entre autres des requêtes du langage SQL et quelques lignes de programmation du langage Python. Ce modèle est essentiellement un prototype réalisé dans le but d’explorer la possibilité d’automatiser une telle procédure.”

Les fichiers du modèle créé par Laurent sont dans ce répertoire ci-bas. Inclut notamment le guide de l’usager « Outil de numérotation automatique et séquentielle de segments; Documentation et marche à suivre »

[\\int.ec.gc.ca\sys\InGEO\GW\EC3140EnvEmer\_UrgEnv\SensitivityMapping\Analysis\Tools\code\_Alphanumerique](file:///\\int.ec.gc.ca\sys\InGEO\GW\EC3140EnvEmer_UrgEnv\SensitivityMapping\Analysis\Tools\code_Alphanumerique)