Naming Convention for ECCC’s Shoreline Segmentation

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

Table des matières

[Abstract i](#_Toc81915513)

[Résumé i](#_Toc81915514)

[Introduction 1](#_Toc81915515)

[Objectives 1](#_Toc81915516)

[Existing literature and revisions of naming conventions in existing datasets 1](#_Toc81915517)

[Methodology 2](#_Toc81915518)

[Appproach 2](#_Toc81915519)

[Challenges and Issues 2](#_Toc81915520)

[Sequential naming of segmentation options 3](#_Toc81915521)

[Segmentation naming by land mass 3](#_Toc81915522)

[Segmentation naming by nearest neighbour 3](#_Toc81915523)

[Annexes 5](#_Toc81915524)

[Previous work 8](#_Toc81915525)

[Code alphanumérique 8](#_Toc81915526)

# Definition

Segment:

Sector:

Grid:

# Introduction

## Objectives

The purpose of this work is to develop a naming convention for ECCC’S shoreline segmentation nation-wide 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.

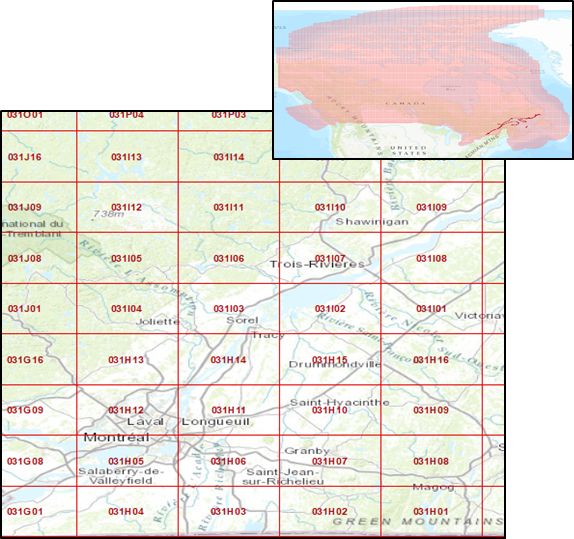
Most datasets in the annexe (table 2) were acquired from NEEC's pre-existing database. Other datasets were explored from this site https://canadiangis.com/data.php#Canadian-Data (source to multiple Canadian geospatial resources).

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 dependant 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. 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 localised grid would necessitate at most the renumbering of only the segments within this grid.

There were currently two grid systems that are being 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 it’s stability. The number of segments would be manageable and should remain in the hundreds. Most importantly as the grid is already existing and maintained by Natural Resources Canada, the cost of using this grid is low.



# Methodology

## Appproach

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

## Sequential naming of segmentation options

Segmentation naming by land mass.

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 segment to the bottom left.

### Segmentation naming by 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 though a the proximity of segments in adjoining sectors
  2. To select the segment with the lowest OID value. Usually, segment should have been created in sequence. This does not preclude the possibility of an outlier. 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.

## 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 registry of the proximity of each landmasses or it might start numbering segments in a runaway process further away from the main land.

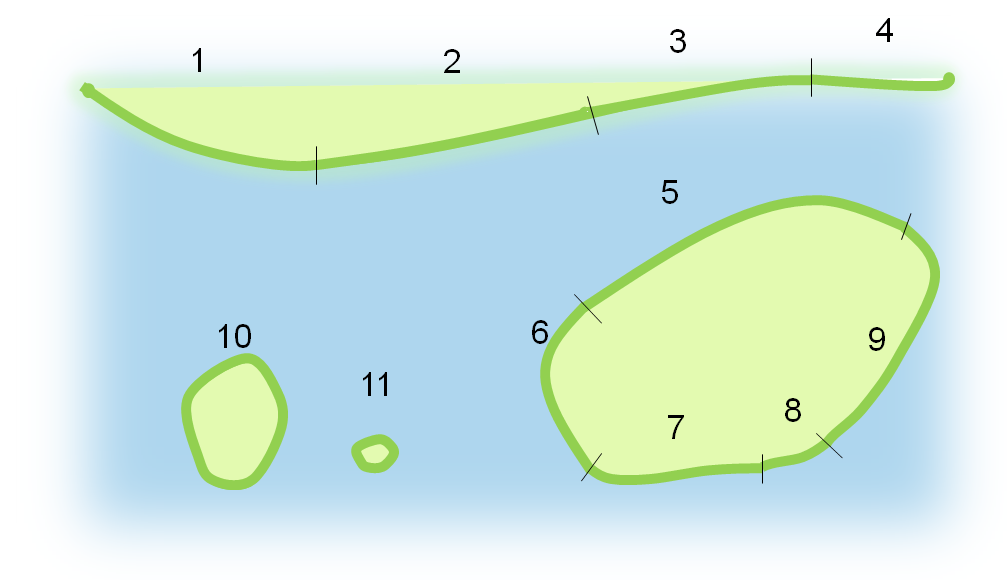


Figure 1. Sequential Naming

* One of the main issues with grids is the uneven nature of cut-offs in the segments as they have been created. It cuts through landmasses and group shorelines according to their respective nature. The subsequent problem is one of spatial relationship. What would be the closest segment?

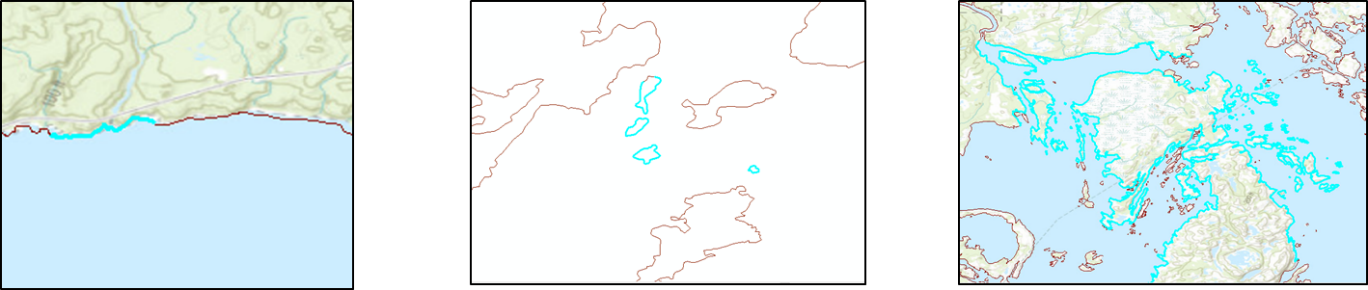


Figure 2. Individual segment

* An approach by grid may also sometimes lead to 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 cut by grid unless it follows the shore closely and its dimensions are small.

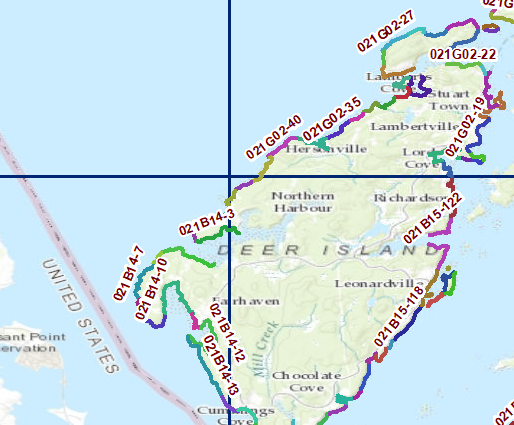


Figure 3. Division of Deer Island (NB) in three separate sectors of the NTS grid

To assuage the better technical proposition, a python script has been created with the different methods. It is encapsulated in an arctool for ease of use.

# Results and Discussion

The script has been included with this documentation. It should aslo be available on github at https://github.com/vboulianne/ECCC-Shoreline-Segmentation

## 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 work. The result is mostly 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.

## Method 2: By Proximity

Select a first segment according to an ordering field or an added field “sector\_first\_segment” with edited manually to “Yes”. The following segments are afterwards named based on proximity to the previous segment with the tool “*Create Near Table*”. An adjoining segment have the distance “0” ensuring that touching segments will be named one after the other. The result is of mixed quality. Again, when the shoreline is simple, the sequential naming is appropriate. The islands are different since the script will process a whole landmass before processing other landmasses (fig. 1). Since the continent can have a long interrupted shoreline, the island may have a number much different.

# 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)