CHyF Pilot Data Specification, v1.0

Prepared by: Mark Sondheim, Natural Resources Canada (mark.sondheim@canada.ca)
Marie-Eve Martin, Natural Resources Canada (marie-eve.martin@canada.ca)
Denis Boutin, Natural Resources canada (denis.boutin@canada.ca)

Date: 2018 March 6

This specification pertains to data that will be used to generate graph theoretic constructs and other structures needed to support hydrologic web services. The terms used herein generally follow CHyF terminology, but the details are specific to the input requirements. The specification is not appropriate for general distribution of CHyF compliant data.

1 Projection

All data to be input shall be in EPSG:4617, *NAD83 (CSRS)*. Calculation of lengths and areas shall be carried out in EPSG:6624, *NAD83(CSRS) / Quebec Albers*. Display will be in EPSG:3857, *Web Mercator*.

2 Format

For this initial pilot, all data is delivered in a shapefile format, with a shapefile provided for each feature type as noted below. In subsequent work, GeoPackage will be supported for both input and output.

3 Features Types

3.1 Elementary Flowpath

An elementary flowpath is a representation of the centerline of water flow through a river or lake, with start and endpoints defined by river confluences or other features of interest. Elementary flowpaths are alway oriented in a downhill direction such that the order of the vertices defining the elementary flowpath follows the flow of water. Elementary flowpaths can intersect one another or other features in the dataset only at flowpath start or endpoints.

Flowpaths form directed, acyclic flowpath networks, which may be fully represented topologically as a connectivity graph. Flowpaths through waterbodies with an areal extent are generated as skeletons. These skeletons are part of the elementary flowpath input dataset.

A series of attributes of elementary flowpaths are defined below for use on the pilot. The name of the attribute is of the form A/B/C, where A is the input attribute name, B is the attribute name to be used for the user interface display in the pilot, and C is the output attribute name, used in GeoJSON output in the pilot. For C names, the number of characters never exceeds 10, for compatibility with shapefiles. In the future, with GeoPackage support, this restriction may be lifted.

Note that `Internal Codes` are also provided. These are strictly for software coding purposes; they are not exposed to end-users.

TYPE / Type / type

- `Observed / Internal Code: 3010`
  If the elementary flowpath corresponds to a single line river segment mapped in the NHN as Observed.
- `Constructed / Internal Code: 3020`
  If the elementary flowpath corresponds to a single line river segment mapped in the NHN as Constructed.
- `Bank / Internal Code: 3030`
  If it corresponds to a section of a skeleton that connects to a bank catchment. Neither end of a bank type flowpath should connect to a pre-existing connection point. This has no equivalent in the NHN.
- `Inferred / Internal Code: 3040`
  If it corresponds to a section of a skeleton in a lake or double-line river that connects at both ends to other skeleton segments or to an observed flow. These are referred to as Inferred in the NHN.

RANK / Rank / type

- `Primary / Internal Code: 1`
  If the flowpath is considered to contain the principal water flow. It is equivalent to LevelPriority=1 in the NHN.
- `Secondary / Internal Code: 2`
  If the flowpath is considered a braid, a distributary or other flowpath encountered in a downstream direction that is of less importance than the primary flowpath. It is equivalent to LevelPriority=2 in the NHN.

- / Strahler Order / strahleror

- *Integer*
  Also referred to as stream order in the US, this value is not provided in the input dataset. It is generated by the CHyF software and available on the demo. This has no equivalent in the NHN. See Addendum 1 for a description of how it is calculated.

- / Horton Order / hortonor

- *Integer*
  This value is not provided in the input dataset. It is generated by the CHyF software and available on the demo. This has no equivalent in the NHN. See Addendum 1 for a description of how it is calculated.

- / Hack Order / hackor

- *Integer*
  Also referred to as stream level in the US, this value is not provided in the input dataset. It is generated by the CHyF software and available on the pilot. This has no equivalent in the NHN. See Addendum 1 for a description of how it is calculated.

DIRECTION / FD Certainty / certainty

- *Integer*
  A value of 1 or -1. 1 indicates that the flow direction is known, and that the order of vertices meaningfully conveys that direction. -1 indicates the flow direction is not known and that the implied direction is arbitrary.

- / Length (m) / length

- *Float*
  The planimetric length in metres of the elementary flowpath. No input name is given because the value is determined by the web service.

NAME / Primary Name / name

- *String*
  This is the primary name, which in the case of the pilot area will be in French. Note that in a CHyF compliant exported dataset, the facility to support multiple names in different languages will be supported, as described more generally in HY_Features.

NAMEID / - / nameID

- *String*
  This is a UUID corresponding to the Primary Name described above. It is included in the pilot input data. It is not available through the pilot's user interface, but it is provided in the GeoJSON output.

- / ID / ID

- *Integer*
  This identifier is not a persistent, immutable ID. Instead it is considered as transient, only relevant to a given dataset. It is not provided as part of the input dataset. It is useful on the pilot if the user wants to specify a particular ID. It is also of value if a connectivity table of feature relations is of interest; this is not an available output right now but it will be in the future. This has no equivalent in the NHN. No input name is given because the value is determined by the web service.

CCWS Watershed

- *String*
  The contracted watershed in which the feature resides, following the CHyF Contracted Watershed System (CCWS). If it spans more than one watershed, then the contracted watershed in which the lowest hydro nexus resides. The CCWS watersheds would be equivalent to HUC 10 or 12 in the US system. Details about them have not yet been

determined. They are not included in the current pilot, but should be considered for inclusion in future work.

The table below summarizes which attributes are included in the input data and which are included in the output data.

Input Attribute Name	CHyF Pilot Names	Input Data	Output Data
TYPE	Type / type	Yes	Yes
DIRECTION	FD Certainty / Certainty	Yes	Yes
RANK	Rank / rank	Yes	Yes
NAME	Primary Name / name	Yes	Yes
NAMEID	- / nameID	Yes	Yes (but not in UI)
	Strahler Order / strahleror	Yes	Yes
	Horton Order / hortonor	Yes	Yes
	Hack Order / hackor	No	No
	Length (m) / length	No	Yes
	ID / ID	No	Yes
	(CCWS Watershed)	No	No

Table 1: Elementary flowpath attributes, for input and output data

3.2 Elementary Catchment

The term elementary catchment refers to an areal feature with polygonal geometry that defines a fundamental drainage area. Input names are not provided in the list below because all attributes are calculated. Output attributes are determined algorithmically using correspondence with elementary flowpaths and/or geometric and topologic characteristics.

- / Type / type

- `Reach / Internal Code: 4010`
  If the catchment contains an observed or constructed flowpath.
- `Bank  / Internal Code: 4020`
  If the catchment does not contain an elementary flowpath and is a neighbour of a water catchment.

- `Water / Internal Code: 4030`
  If the catchment is a lake, a portion of a lake, a double-line river or a portion of a double-line river.
- `Empty / Internal Code: 4040`
  A catchment containing no waterbodies and with no flow connections to other catchments. An empty catchment corresponds to a depression in a DEM, where water has not been mapped.

- / SubType

This is completed only if the catchment is a water catchment. Otherwise it is null.

- `Lake / Internal Code: 4031 / Waterbody definition Code: 4`
  The water catchment is part of a lake.
- `Pond / Internal Code: 4032 / Waterbody definition Code: 9`
  The water catchment is part of a pond.
- `River  / Internal Code: 4033 / Waterbody definition Code: 6`
  The water catchment is part of a river.
- `Canal / Internal Code: 4034 / Waterbody definition Code: 1`
  The water catchment is part of a canal

These subtypes are not necessarily all CHyF compliant. In the current version of CHyF, `Pond` is not recognized. All four are recognized in the NHN specification and this has been carried forward for use in the pilot.

- / Terminal / terminal

- `true`
  If the catchment occupies the lowest part of an interior catchment network. This may be (i) an isolated lake, with no inflow or outflows, (ii) the catchment of a single river segment that has no connecting flowpaths, or (iii) the lowest catchment of a set of connected catchments, which may or may not be a lake. The second case is considered as terminal, as it does not drain into another catchment, even though it does flow into the contained single-line river.

  If the elementary catchment is of type EmptyCatchment, then *Is Terminal* is necessarily True.
- `false`
  The more typical situation where the elementary catchment flows into another elementary catchment, and thus does not meet the conditions stated above for `TRUE`.

- / Rank / rank

- `1 / Internal Code: 1`
  A rank of 1 means a primary flow. If the contained flowpath is considered to contain the principal water flow. It is equivalent to the contained flowpath attributed as LevelPriority=1 in the NHN.

- `2 / Internal Code: 2`
  A rank of 2 means a secondary flow. If the contained flowpath is considered a braid, a distributary or other flowpath encountered in a downstream direction that is of less importance than the primary flowpath. It is equivalent to the contained flowpath attributed as LevelPriority=2 in the NHN.
- `null / Internal Code: 0`
  If the catchment contains no flowpath. The NHN does not have an equivalent. For `Primary` and `Secondary`, the value is transferred from the contained flowpath. Bank catchments and empty catchments all have a rank of `NoRank`.

  In the case of `Primary` and `Secondary`, the value is transferred from the corresponding flowpath; in the case of `NoRank`, the value is assigned, based on the catchment type being BankCatchment or EmptyCatchment.

  For water catchments, the rank is `Primary` if at least one contained flowpath has a rank of `Primary`. If no primary flowpath exists in the lake, but one or more secondary flowpaths do, then the water catchment has a rank of `Secondary`.

- / Strahler Order / strahleror

- *Integer*
  The Strahler Order of the corresponding elementary flowpath. In the case of a lake catchment or river catchment, the value is that of the contained elementary flowpath with the largest Strahler Order value. The value is transferred from the contained flowpath.

- / Horton Order / hortonor

- *Integer*
  The Horton Order of the corresponding elementary flowpath. In the case of a lake catchment or river catchment, the value is that of the contained elementary flowpath with the largest Horton Order value. This attribute is not in the input dataset. The value is transferred from the contained flowpath.

- / Hack Order / hackor

- *Integer*
  The Hack Order of the corresponding elementary flowpath. In the case of a lake catchment or river catchment, the value is that of the contained elementary flowpath with the smallest Hack Order value. The value is transferred from the contained flowpath. For the pilot this attribute is not considered, but it will be included at a later date.

- / Area (ha) / area

- *Float*
  The planimetric area in hectares of the elementary catchments.

- / Primary Name / name

- *String*
  This is the primary name, which for the pilot is taken from the name of the contained flowpath.

- / ID / ID

- *Integer*
  This identifier is not a persistent, immutable ID. Instead it is considered as transient, only relevant to a given dataset. It is not provided as part of the input dataset. It is of value if a connectivity table of feature relations is of interest.

- / CCWS Watershed

- *String*
  The contracted watershed in which the feature resides, following the CHyF Contracted Watershed System (CCWS). If it spans more than one watershed, then the contracted watershed in which the lowest hydro nexus resides. The CCWS watersheds would be equivalent to HUC 10 or 12 in the US system. Details about them have not yet been determined. They are not included in the current pilot, but should be considered for inclusion in future work.

The table below lists the data to be included as output. Note that no input attributes are required.

Output Attribute Name	**Output Data**
Type / type	Yes
SubType / subtype	Yes
Is Terminal / isTerminal	Yes
Rank / rank	Yes
Strahler Order / strahleror	Yes
Horton Order / hortonor	Yes
Hack Order / hackor	Yes
Area (ha) / area	Yes
Primary Name / name	Yes
ID / ID	Yes
(CCWS Watershed)	No

Table 2: Elementary catchment attributes, for output data

The elementary catchments form a coverage, such that no gaps or overlaps exist across the area of interest. Lakes, double-line rivers, and islands are all included in this coverage, as will be clear through an examination of the examples shown in a later section.

3.3 Waterbody

A waterbody file will be provided as input that will include all waterbodies large enough to define with polygon geometry. The waterbody polygons are used to identify and classify water catchments. The waterbody polygon geometry should be identical to the polygon geometry of the corresponding water catchment, except for allowing the waterbody polygon to be divided into smaller parts for ease of processing.

The following types of (polygonal) waterbodies are recognized in the pilot:

- `Lake / Internal Code: 1011`
- `Pond / Internal Code: 1012`
- `River  / Internal Code: 1021`
- `Canal / Internal Code: 1022`

3.4 Nexus

Since nexuses are not provided as part of the input dataset, they are listed here for reasons of clarity. Following the HY_Features conceptual model, the geometry of a (hydro) nexus is not restricted. In CHyF and the pilot implementation however, the geometry is always represented as a point.

In HY_Features a hydro nexus is a point of connection on a flowpath network. CHyF recognizes three kinds of hydro nexuses: flowpath nexus, bank nexus, and water nexus, as described below. CHyF also includes the idea of a unary nexus, which occurs at the end of flowpath where no other flowpath connections exist. Two types of unary nexuses are described below, headwater flowpath start point and terminal flowpath endpoint. Finally, CHyF also includes an inferred junction as a type of nexus, which is also described in what follows.

The following nexus types are defined and are shown as follows, with the type attribute also indicated on the pilot in a UI pop-up box.

- Headwater `/ Internal Code: 2011`
  (white in the figures that follow) - Headwater flowpath start point. The start of an elementary flowpath, for which the start point does not intersect other flowpaths, lake boundaries, or double-line river boundaries.
- Terminal `/ Internal Code: 2012`
  (gray) - Terminal flowpath endpoint. The endpoint of an elementary flowpath that is not connected to any other flowpath at the endpoint. It may be a true sink, or just the limit of what was visible when the hydrography was compiled, or the endpoint of a flowpath across a terminal lake.

- Flowpath / `Internal Code: 2021`
  (red) - Flowpath nexus. The junction of a single-line river with another single-line river or with the boundary of a double-line river or lake. These are defined by the endpoints of observed flowpaths.
- Water / `Internal Code: 2022`
  (blue) - Water nexus. The junction of a double-line river or lake with another double-line river or lake, as represented by the intersection of the incoming and outgoing flowpaths at that location.
- Bank / `Internal Code: 2023`
  (green) - Bank nexus. A type of hydro nexus that is arbitrarily placed on the boundary of adjacent catchments to show flow or potential flow from one to the other, where no observed flowpaths exist to convey that information, and where the upper catchment is land based and the lower is water based (lake or double-line river).
- Inferred / `Internal Code: 2030`
  (black) - Inferred junction. These are required for the purpose of connectivity only. They are not hydro nexuses.

4 Topology

4.1 Node and Vertex Topology

Where a flowpath meets a flowpath, the point of intersection is always either a start point or an endpoint for each flowpath.

Where a flowpath meets an elementary catchment of any type, it must meet at a vertex on the polygon representing the boundary of the catchment. That vertex may be any of the vertices that compose the polygon. A flowpath never crosses another flowpath, and similarly a flowpath never crosses a catchment boundary.

4.2 Detailed Rules

The data as input is to conform to a series of rules as described below.

The following types of elementary flowpaths (LF) exist:

1. Observed Flowpath (OF)
2. Constructed Flowpath (CF)
3. Inferred Flowpath (IF)
4. Bank Flowpath (BF)

The following types of elementary catchments (LC) exist:

1. Reach Catchment (RC)
2. Water Catchment (WC)
3. Bank Catchment (BC)

4. Empty Catchment (EC)

The rules are based on their topological relationships and can be expressed using the Dimensionally Extended nine-Intersection Model (DE-9IM) model. The term *contained by* is introduced; if x *contains* y then y is *contained by* x. The other term for *contained by* is *within*. This latter term is avoided here because *within* may be equated with *inside*, which is not quite the same thing.

1. Every LF is contained by exactly one LC, based on the type:
    a. Every OF is *contained by* an RC.
    b. Every CF is *contained by* an RC.
    c. Every IF is *contained by* one and only one WC.
    d. Every BF is *contained by* one and only one WC.
2. An LC may contain 0, 1, or many LFs depending on the type:
    a. Every RC *contains* one and only one LF, which may be of type OF or CF.
    b. Every WC *contains* one or more IF.
    c. Every BC does not *contain* an LF.
    d. Every EC does not *contain* or *touch* any LF.
3. The interaction between the geometries of LFs and LCs is limited in the following ways:
    a. For every RC, its boundary *intersects* the boundary of the contained LF at either one or two points.
    b. The interior of every LF is *disjoint* from all other LF and from the boundary of all LC.
    c. Every BC *touches* one and only one BF.
    d. Every BC *touches* one or more WC and the intersection of each such BC - WC pair is a line segment.
4. All LF together form a directed, acyclic graph.
5. All LC together form a continuous coverage with no gaps and no overlaps. (This is not expressed here using DE-9IM terminology, but it is closely related to these concepts.)

We could have a series of rules for all the endpoints, but there is no reason to do so. If the rules above are all validated, so would rules about the endpoints. Also, they are not required input.

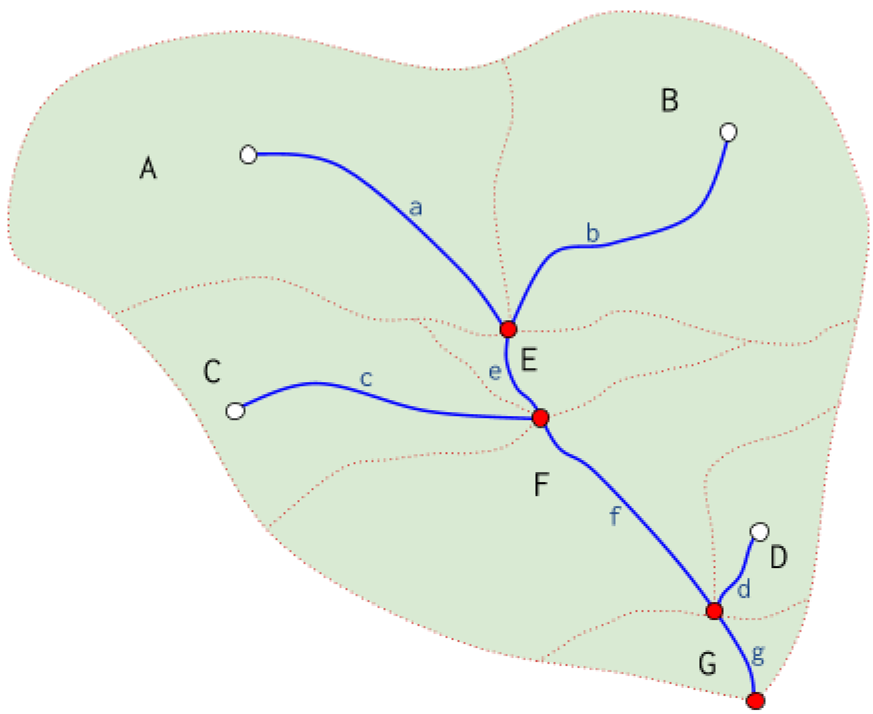
5 Examples

Some examples are provided below. Only some of the attribute - value pairs are listed in each case.

5.1 A Basic Catchment

A catchment consists of seven elementary catchments labelled “A” through “G” in Figure 1. No lakes or double-line rivers are involved. Note that in this example every elementary catchment contains an elementary flowpath, with the seven elementary flowpaths labelled “a” through “g”. Where flowpath nexuses bound the elementary flowpaths, the catchment is bisected by the flowpath. Where they do not, the elementary flowpaths are first order (Strahler Order = 1) rivers, that is, headwaters, with the start point of each defined as a unary nexus.

The endpoint of the flowpath network in this figure (in G) and in the figures to follow is shown as a red dot instead of a grey dot because it is assumed to be a flowpath nexus (see section 3.3 above).



Example 1: Elementary catchments and elementary flowpaths

Elementary Flowpath Data

The input data for the flowpath network displayed in Example 1 is shown in the table below. In this hypothetical example data no values are provided for NAME for elementary flowpaths d and e because none exists in the NHN source.

Elementary Flowpath	TYPE	RANK	STRAHLEROR	HORTONOR	HACKOR	NAME
a	Observed	Primary	1	1	2	Rivière aux Brochet
b	Observed	Primary	1	2	1	Rivière Magog
c	Observed	Primary	1	1	2	Rivière Magog
d	Observed	Primary	1	1	2	

e	Observed	Primary	2	2	1	
f	Observed	Primary	2	2	1	Rivière Saint-François
g	Observed	Primary	2	2	1	Rivière Yamaska

Table 3: Input flowpath data, for Example 1

The first column in Table 3 is provided as a means of relating the values in the table to the graphic in Figure 1.

The output data from the web service for the flowpath network is as shown in Table 4 below. The output is the same as the input with the addition of two attribute, *Flowpath ID* and *CCWS Watershed*. For the pilot this data will be visible when a feature is highlighted on the viewer. In a future pilot the data could be output in a simple table as shown. As well, a table of flow relationships could be output using the values for *Flowpath ID*.

Elementary Flowpath	Type	Rank	Strahler Order	Horton Order	Hack Order	Length (m)	Primary Name	ID	CCWS Watershed
a	Observed	1	1	1		113.0	Rivière au Brochet	3864905	
b	Observed	1	1	2		18.2	Rivière Magog	6582583	
c	Observed	1	1	1		110.4	Rivière Magog	7089992	
d	Observed	1	1	1		21.6		5538709	
e	Observed	1	2	2		18.2		1690373	
f	Observed	1	2	2		72.5	Rivière Saint-François	6439677	
g	Observed	1	2	2		25.1	Rivière Yamaska	4196783	

Table 4: Input flowpath data, for Example 1

Elementary Catchment Data

The input data for the catchment network displayed in Example 1 consists of polygonal data with no attributes.

The output data is as described in the table below.

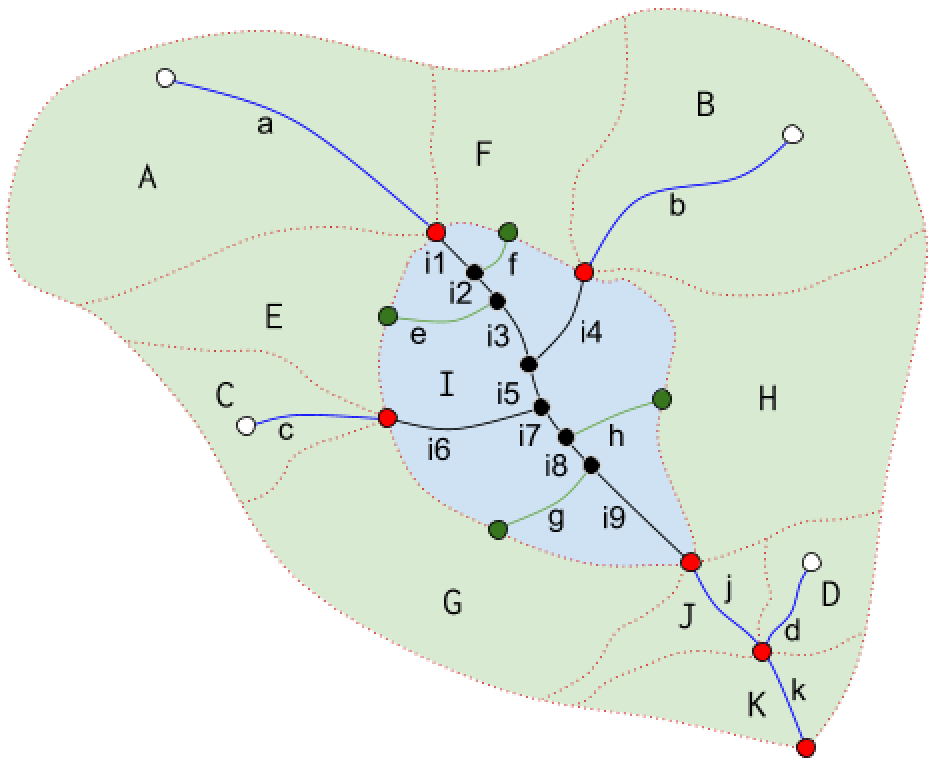
For the pilot this data will be visible when a feature is highlighted on the viewer. In a future pilot the data could be output in a simple table as shown. As well, tables of flow and containment relationships could be output using the values for *ID*.

Elementary Catchment	Type	Is Terminal	Catchment Rank	Strahler Order	Horton Order	Hack Order	Area (ha)	Primary Name	ID	CCWS Watershed
A	Reach	False	1	1	1	2	423.5	Rivière au Brochet	675472	
B	Reach	False	1	1	2	1	392.6	Rivière Magog	309454	
C	Reach	False	1	1	1	2	223.5	Rivière Magog	549989	
D	Reach	False	1	1	1	2	93.0		214287	
E	Reach	False	1	2	2	1	83.8		673905	
F	Reach	False	1	2	2	1	446.7	Rivière Saint-François	389755	
G	Reach	False	1	2	2	1	71.1	Rivière Yamaska	297434	

Table 5: Output catchment data, for Example 1

5.2 Connectivity in a Catchment with a Simple Lake

In this example the catchment consists of 11 elementary catchments, each of which has a polygonal boundary. Collectively, they form a perfect coverage of the larger catchment. Only a limited number of attributes are shown, as pertinent to the example.



Example 2: Connectivity in a catchment with a simple lake

One of these is the lake (`I`); the other ten are land areas. The values for a select number of attributes for all features is as follows:

1. Elementary catchments
   - `I`: `Type =` `Water`
     `Is Terminal =` `false`
     `Strahler Order =` `0`
   - A - D: `Type =` `Reach`
     `Is Terminal =` `false`
     `Strahler Order =` `1`
   - E - H: `Type =` `Bank`
     `Is Terminal =` `false`
     `Strahler Order =` `0`

- J, K: `Type = Reach`
  `Is Terminal = false`
  `Strahler Order = 2`

Elementary flowpaths

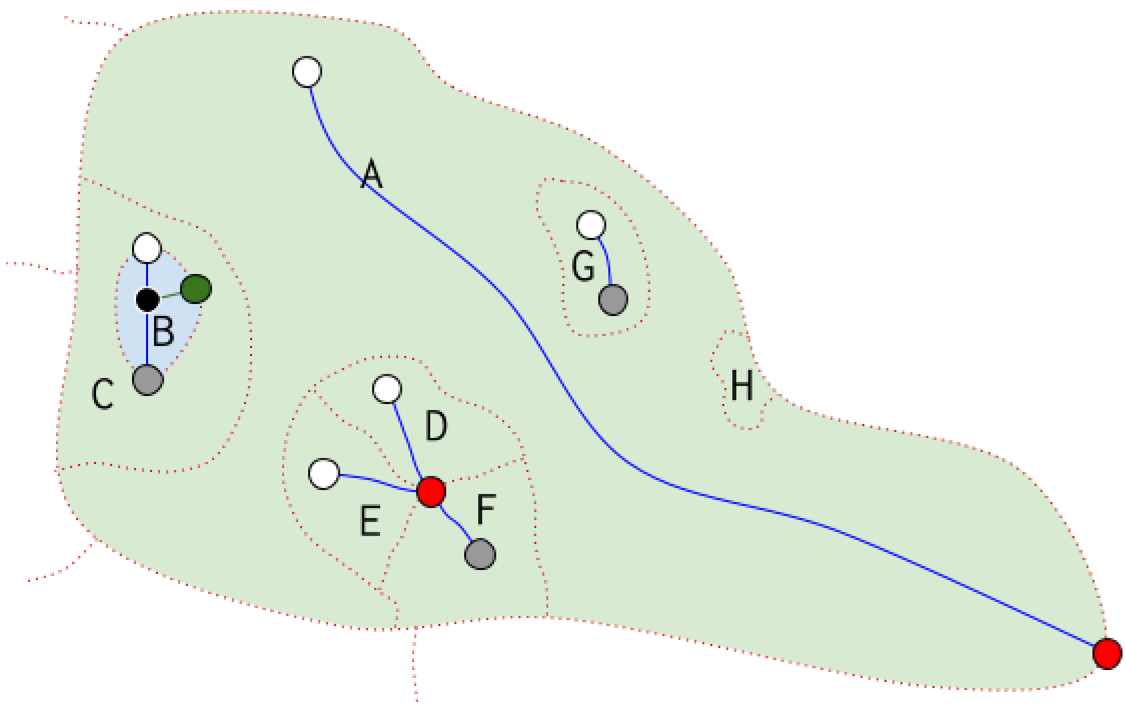
For all flowpaths: `Rank = Primary`

- Blue: `Type = Observed`
- Green: `Type = Bank`
- Black: `Type = Inferred`

The green hydro nexuses in Figure 2 must be determined first, as they do not exist in the NHN dataset. They are needed to create the green line segments. None of the points of any colour are to be provided in the input dataset, as they can all be determined algorithmically.

5.3 Interior Catchments

In this example, four interior catchments are present. In the first example C drains into B and an inferred elementary flowpath exists across the lake, connected a bank nexus to a terminal flowpath endpoint. The outer boundary of C represents the limit of the interior catchment; that outer boundary is also the sole boundary of the union of B and C. In the second example, E and D drain into F. The boundary of the union of E, D and F defines the limits of that interior catchment. In the third example, G contains an observed elementary flowpath that does not connect to other flowpaths. The interior catchment in this case is simply G. In the fourth example, H is an interior catchment with no contained waterbodies. Using a Voronoi approach, catchments like H cannot exist (except along the edge of the extent of the data); however, they can with a DEM approach.



Example 3a: Different forms of terminal catchments (B, F, G, H)

With sufficient rainfall, F could become completely flooded and drain into A. Similarly, G would flow into A given sufficient rainfall. Because the shared boundary between C and A is longer than the shared boundary of C with any of its neighbours, it is assumed in the absence of DEM analysis that C would also drain into A with sufficient flooding. However, because no evidence exists to indicate that these scenarios are in fact correct, the interior catchments are treated as true isolations.

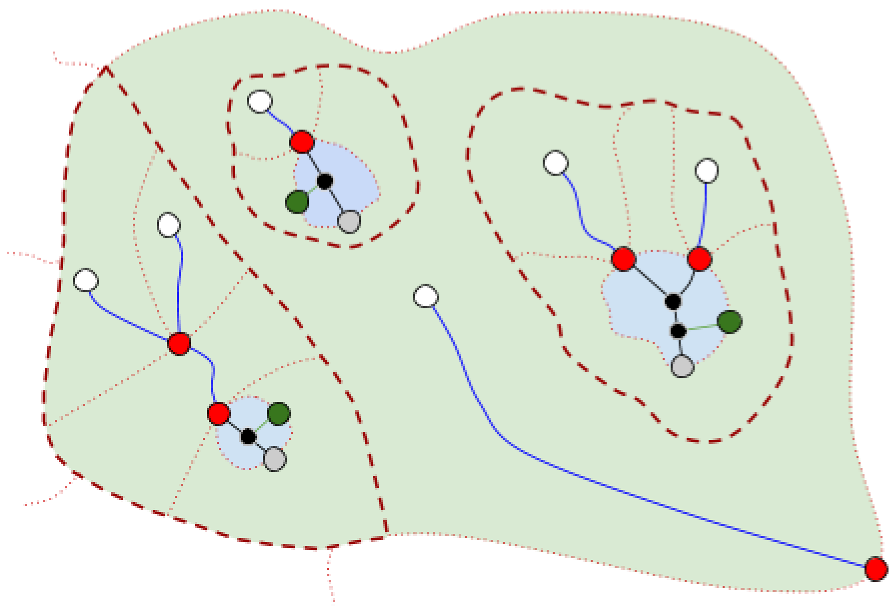
The various features are coded as follows:

2. Elementary catchments
   - `A, E, D`: `Type = Reach`
     `Is Terminal = false`
     `Strahler Order = 1`
   - B: `Type = Water`
     `Is Terminal = true`
     `Strahler Order = 1`
   - C: `Type = BankCatchment`
     `Is Terminal = false`
     `Strahler Order = 1`
   - F: `Type = Reach`
     `Is Terminal = true`
     `Strahler Order = 2`
   - G: `Type = Reach`
     `Is Terminal = true`
     `Strahler Order = 1`
   - H: `Type = Empty`
     `Is Terminal = true`
     `Strahler Order = null`

3. Elementary flowpaths

   For all flowpaths: `Rank = Primary`
   - `Blue:` `Type = Observed`
   - `Green:` `Type = Bank`
   - `Black:` `Type = Inferred`

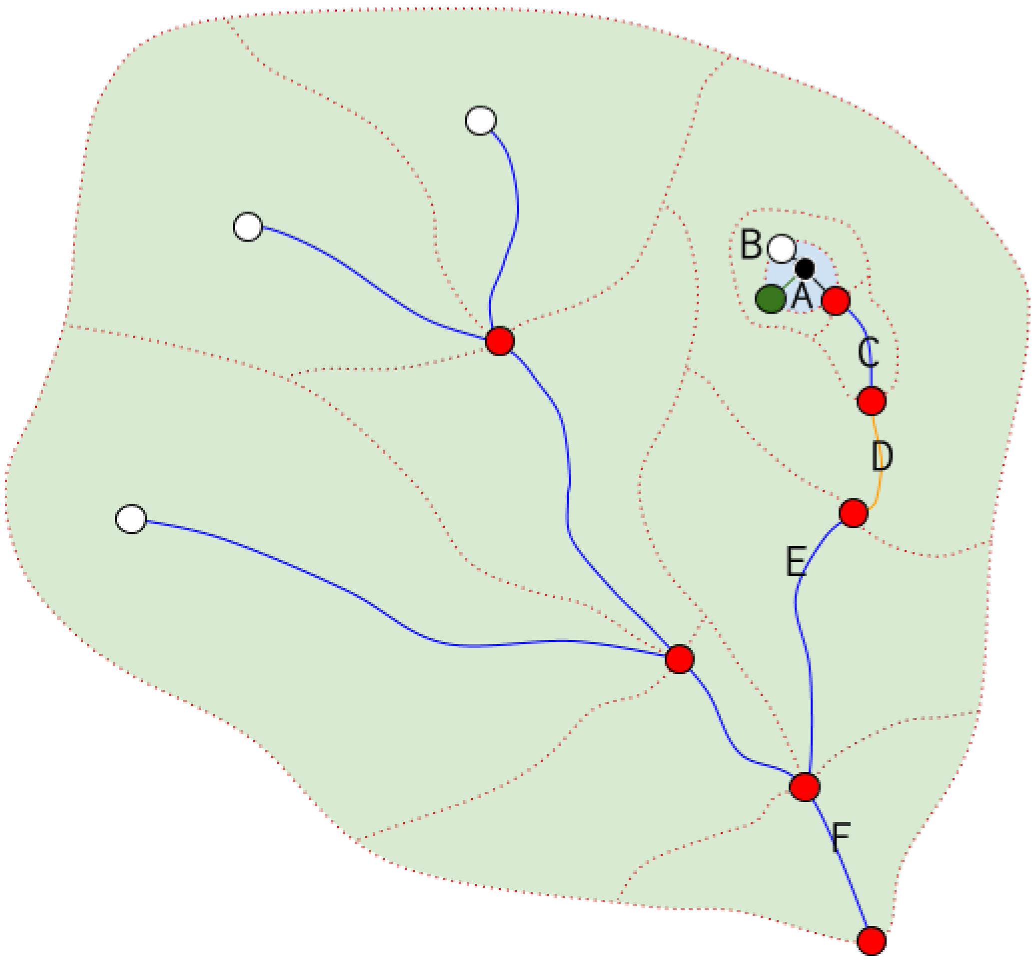
In the next example three lakes are terminal elementary catchments and each is within a larger interior catchment demarcated with a dashed brown boundary. Each of these larger catchments is created by a union operation of all elementary catchments that drain into the contained terminal elementary catchment. Interior catchments like these three are to be determined as part of the pilot.



Example 3b: Three interior catchments, each defined by a union of elementary catchments

5.4 Catchments with Constructed Flowpaths

In the next figure (below), an orange elementary flowpath depicts a constructed flowpath. Without it, the union of A, B and C represents an interior catchment and A would probably be a terminal catchment, with C flowing into it. With the addition of the constructed flowpath, the flowpaths for A and C are re-oriented in a downstream direction, with flow directions of these catchments now understood as: B⇒A⇒C⇒D⇒E⇒F



Example 4: Constructed flowpath connecting an interior catchment

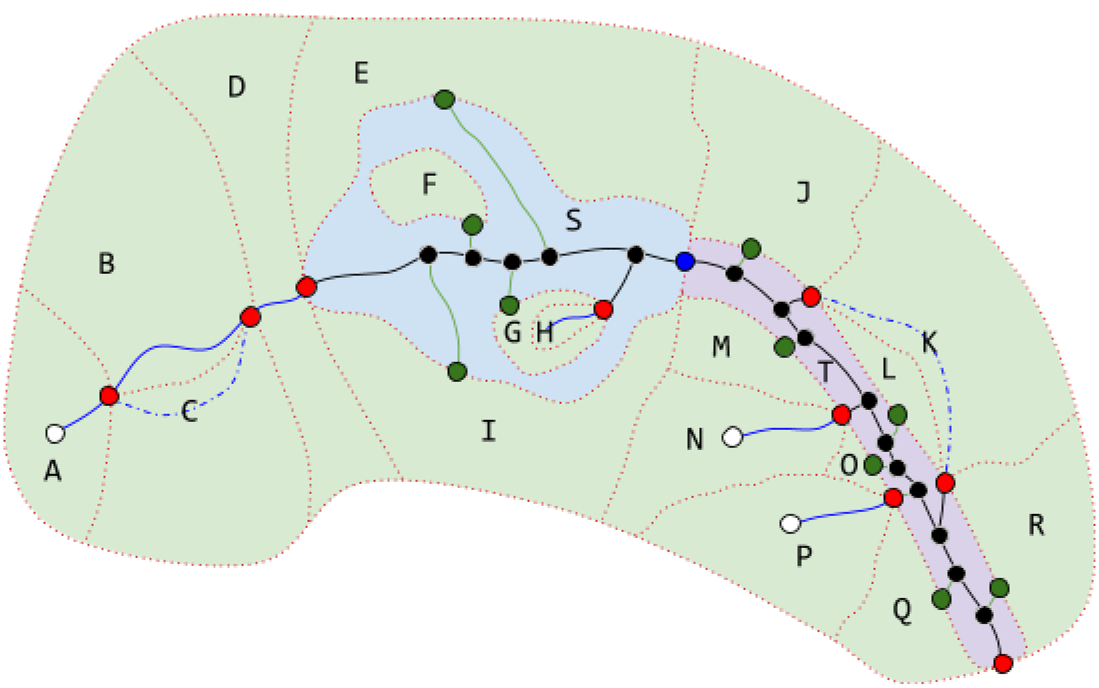
5.5 Braids, Islands, and a Hydro Nexus

In the next example several features are introduced: braided rivers, islands, and a double-line river.

C and K are elementary catchments defined in each case by a braid corresponding to a secondary elementary flowpath. By the definitions given previously, they have have a value of “secondary” for the attribute RANK. It is not always possible to carry out a union operation of such secondary catchments with one another or with neighbours to create catchments as they would have been generated directly from primary flowpaths. Thus, for purposes here it is sufficient to indicate that a given elementary catchment is or is not associated with a secondary flow, by considering the value of RANK.

E, F, I J, L, M, O, Q, and R are all bank catchments connected by bank flowpaths. Islands containing no waterbodies are elementary catchments, as with F. If waterbodies do exist on an island, then the island will be subdivided into two or more catchments (G-H).

The example also shows a lake in the middle of the larger catchment draining into a double-line river on the right, with the blue dot representing a water nexus for connectivity purposes. If a double-line river is broken into two or more areas, then connections through them would be represented by a water nexus in the same fashion.

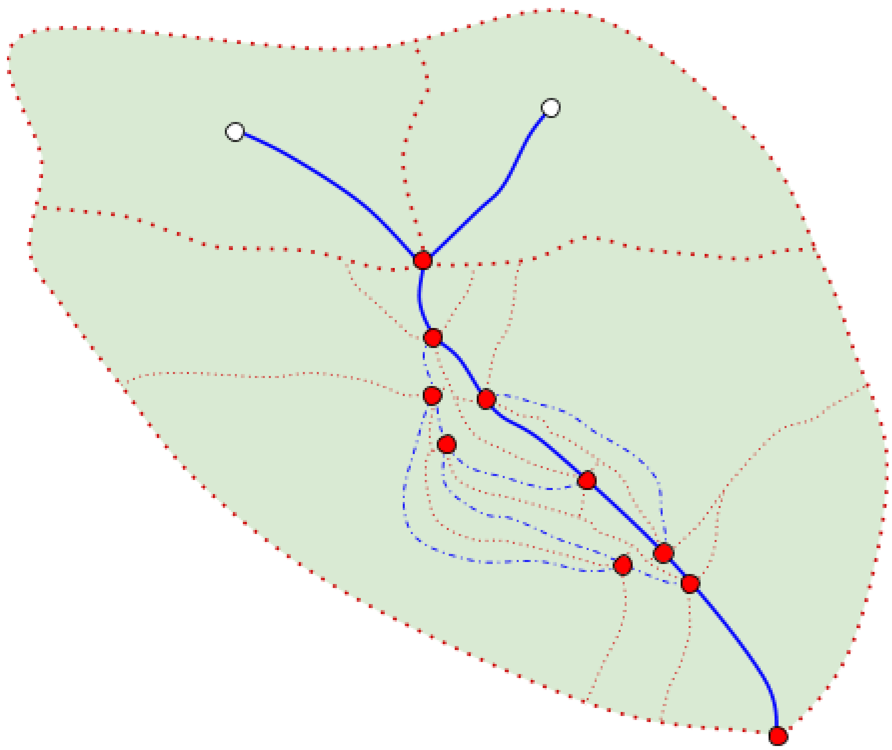


Example 5: Braided rivers, islands, water nexus

In Example 5, no dots are shown where the common boundary of the lake and the double-line river intersects the shoreline. This is intentional, as all areas are already accounted for with connections to the flowpath network.

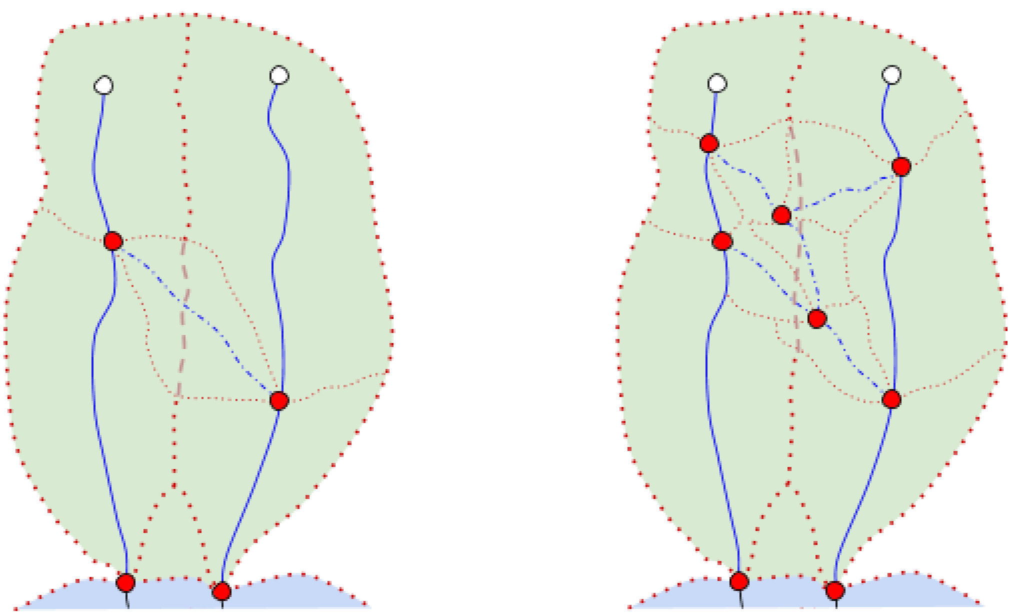
5.6 Secondary Flows

This example displays a set of complex braids along a primary flowpath. The heavier dotted lines show the three elementary catchments that will result if the secondary flows are ignored. The heavier and lighter dotted lines together give the elementary catchment boundaries if both primary and secondary flows are included. The large lower catchment based on the primary flows equates to the union of 12 contained catchments based on all flows.



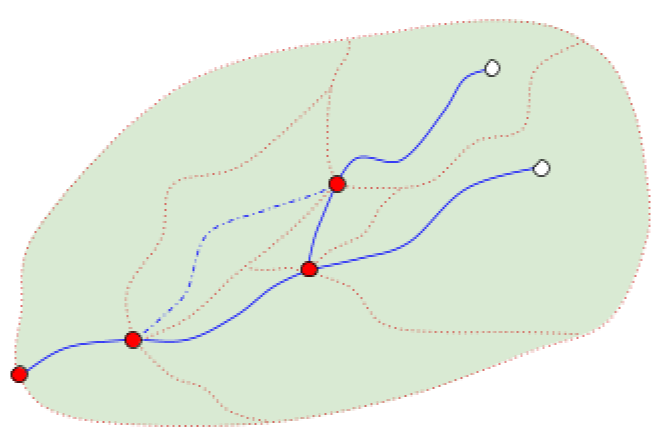
Example 6a: A set of complex braids and the associated catchments

In the two diagrams in Example 6b, the situation is more complicated. The catchments associated with the secondary flows cannot be amalgamated to form the catchments defined on the primary flows alone. The boundaries of the elementary catchments based just on the primary flows are shown by the heavy brown dotted lines as well as the dashed brown lines. As before, the boundaries of the elementary catchments based on all flows are shown by the heavy and light brown dotted lines. It is clear that at this finer level of detail the dashed brown lines are not part of the catchment boundary network.



Example 6b: Secondary connections between primaries

Example 6c below shows another example of a secondary flow connecting two primaries, which in this case are themselves connected. Here as well the catchments of the braids do not nest in the catchments that would result if they were not considered.



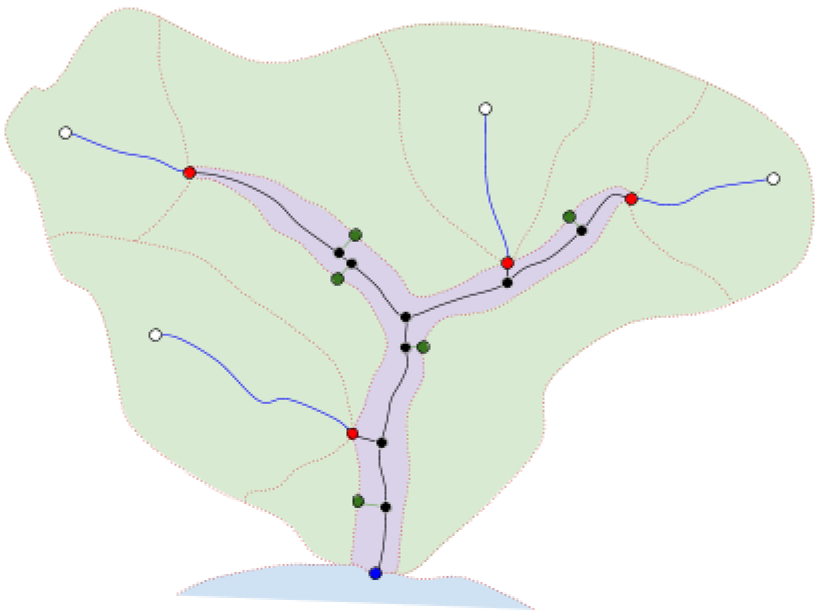
Example 6c: Secondary connections between connected primaries

These examples show that the relationship of catchments based just on primary flows to catchments based on all flows depends on how many primary flowpaths are involved. With a single primary flowpath, the introduction of braids along it leads to a straightforward containment relationship, as shown in Example 6a. Adding braids between primary flowpaths results in a more complex situation in which containment only holds at a higher level, as shown in the three diagrams in Example 6b and 6c.

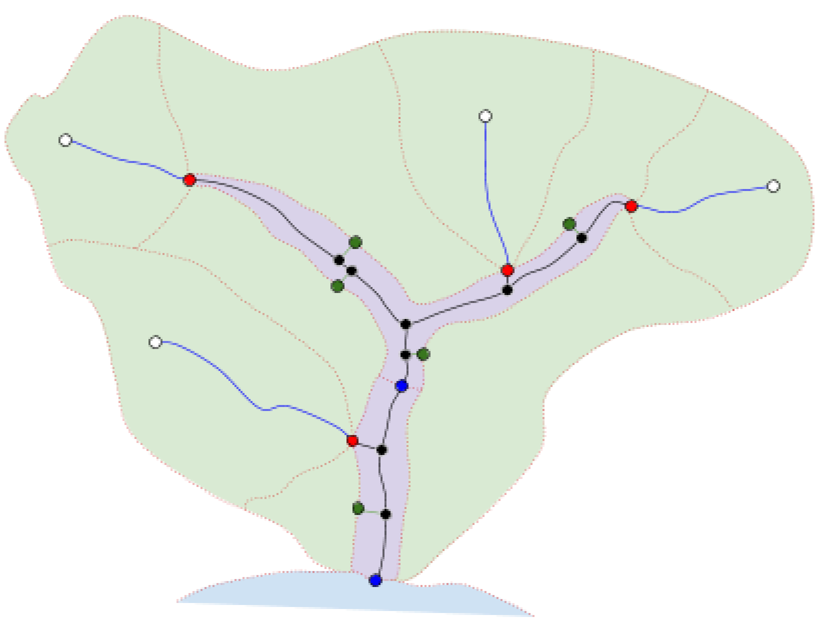
The pilot will make use of both primary and secondary flows and their corresponding catchments defined taking all flows into account. An option for another time would be to create two catchment input datasets and to let the user decide whether all flows are of interest or only primary flows.

5.7 Double-line Streams

Example 7a shows a double line river draining into a lake. Four bank catchments are depicted, each with a bank nexus (green dot) on its boundary. The double-line river forms a single catchment as shown. In some situations this may work well, but in others the size of some of these elementary catchments may be larger or longer than desired. In Example 7b the double-line river is subdivided arbitrarily in the lower section.



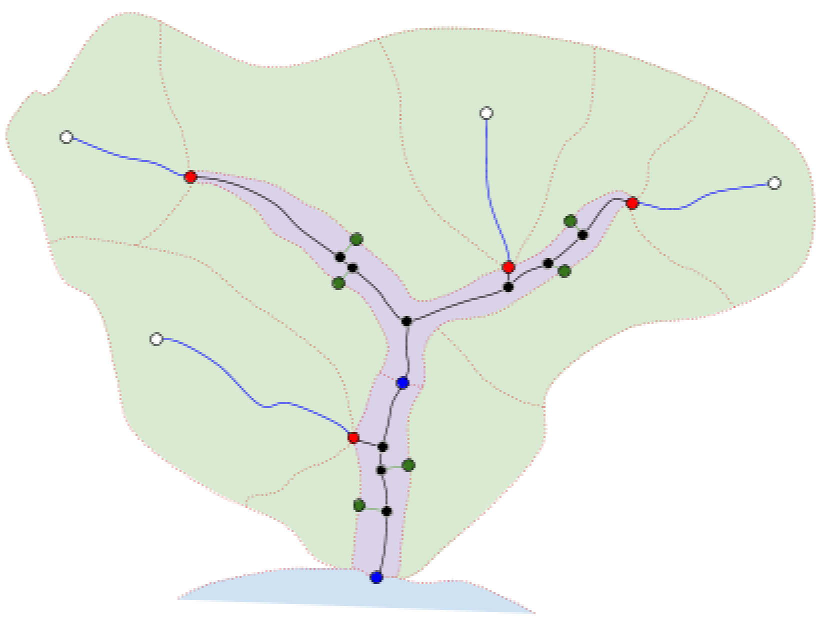
Example 7a: A double-line river with a bifurcation



Example 7b: A double-line river subdivided into two double-line rivers

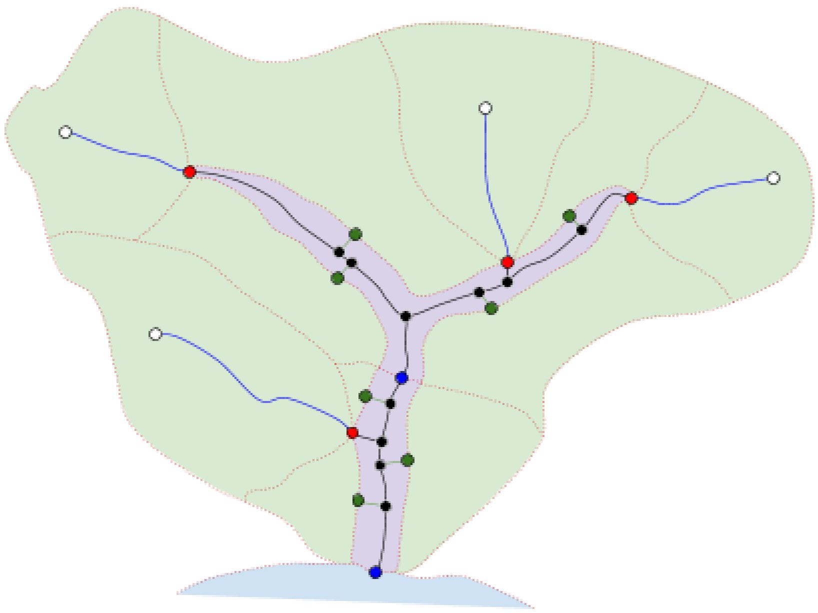
The location of the water nexus in Example 7b could be based on distance upstream or it could be based on some other factor. In any case, it is easily handled as shown. Potentially a web service could be provided that would allow for such water nexuses to be entered interactively; alternatively, processing could insert such breaks at every kilometre or whatever is deemed appropriate for the given scale.

A similar approach could be taken with large land-based elementary catchments. Example 7c shows one of the bank catchments of Example 7b subdivided into two adjacent bank catchments, each connected to the flowpath network through a bank nexus.



Example 7c: A bank catchment subdivided arbitrarily into two bank catchments

An alternative, provided in Example 7d below and similar to what is shown in Figure 5, involves extending the dotted line separating separating the two sections of the double-line river through the land on both sides. This works equally well. This latter case would be preferred if for example the water nexus represented a point of specific interest. In the pilot, should any of the input include subdivisions of double-line streams or bank catchments,than the model in either Example 7b or Example 7d will be followed.



Example 7d: A bank catchment subdivided into two bank catchments consistent with a water boundary

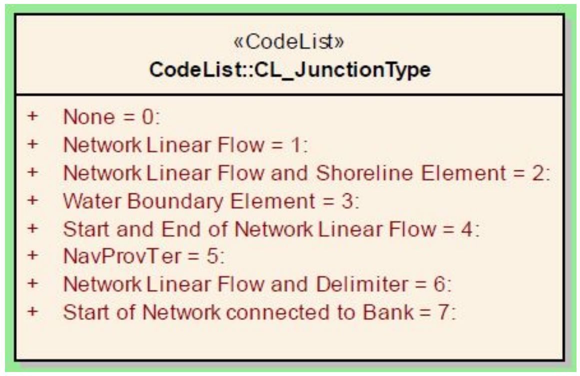
The examples shown in 7a through 7d demonstrate that the general approach to catchment delineation taken with CHyF is highly flexible and has the ability to control the size of elementary catchments if desired.

6 Relationship to NHN

This section describes specifically how nexuses, flowpaths, and catchments related to entities in the National Hydro Network. The order of the sub-sections differs from what was used above by intention.

6.1 Nexuses

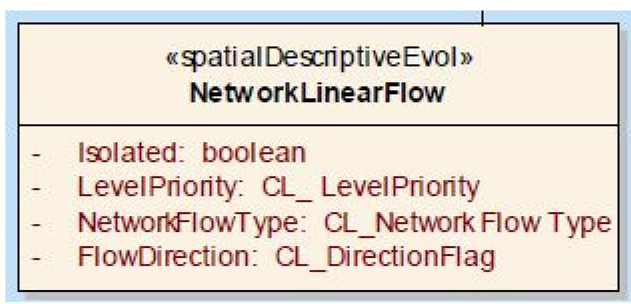
The NHN has rules specifying that where two flowlines intersect, the corresponding vertices must have identical coordinates. In the NHN formal model, the intersection points are referred to as hydro junctions, with a code list specifying the types of junctions recognized:



Some of these values relate to what is used in CHyF, but only the break at provincial or territorial boundaries provides additional information. This may be of interest in the context of hydro location, as defined in HY_Features, and in CHyF with a modified list of hydro location types. For the pilot, the NHN hydro junctions are ignored and all nexuses, including their respective types, are determined by algorithm.

6.2 Flowpaths

CHyF's elementary flowpaths are derived from the NHN's network linear flows, which contain the attributes shown in the box below, as well as two inherited properties: geometry and an identifier. The geometry of a network linear flow, and of the corresponding elementary flowpath, is always a linestring. The identifier is not used.



Various detailed changes have been made as described below.

1. The isolated attribute is not used.
2. LevelPriority has been renamed Rank for the pilot, and contains only two possible values, "Primary" and "Secondary" instead of the numeric codes used in the NHN.
3. NetworkFlowType has been renamed Type for the pilot. Type has four possible values: "Observed", "Inferred", "Constructed", and "Bank". The first three are taken from existing NHN data. Bank flowpaths are additions that have been generated.
4. FlowDirection is not used. Instead, two major differences have been applied.
    a. For all flowpaths the orientation of the vertices is always in a downhill direction. If the direction is not known, then the existing orientation is not altered.

b. A FDCertainty (Flow Direction Certainty) attribute is added, equal to 1 if the direction is known and -1 if it is unknown. In the pilot data, two situations exist where a value of -1 is found. Nearly all cases of -1 are associated with isolated lakes. A small number relate to flowpath network cross connections, threading between islands in the Richelieu River. In both these cases no special treatment is provided, as no significant difference exists in the context of hydrological queries. The isolated lakes are of interest, but other means are used to determine that they are isolated. They can be found using the catchment attribute *isTerminal* with a value of *True*.

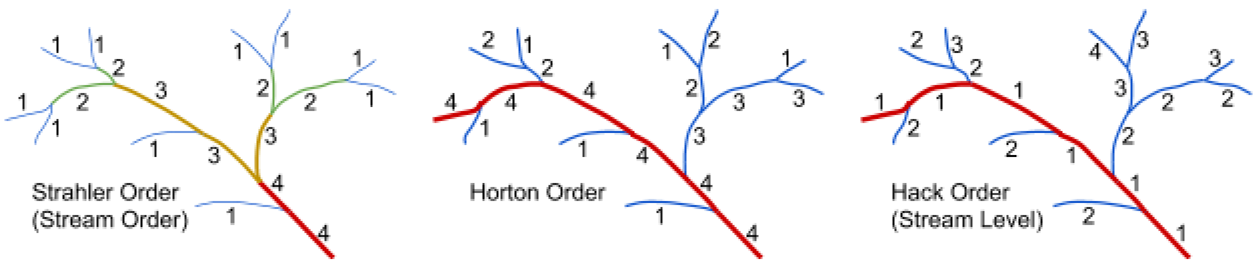
5. Additional flowpaths are required by the CHyF model. Observed, constructed, and inferred eflowpaths are readily constructed from NHN network linear flows. The bank eflowpaths however must be added. In the NHN network linear flows exist around islands, which simplifies the creation of bank eflowpaths, since they are unlikely to cross land. However, no guarantee exists that this will not occur. In the future, a new skeletonizer will be built, which will also include the generation of bank eflowpaths following topologically rules.

6.3 Catchments

The NHN does not include catchments. The ecatchments on the pilot were created using a line Voronoi technique available with the FME. A future development will be a new catchment definition technique that makes use of both elevation and space partitioning approaches, based on local conditions.

Addendum 1: Calculation of Stream Ordering Schemes

Three ordering schemes are defined here, Strahler Order, Horton Order, and Hack Order. They are depicted in the diagram below.



These examples are very idealized. Details are provided below. However, in all three cases, the following two rules apply:

1. Only eflowpaths with rank = 1 are considered. All secondary flows are given the value *null* for all three orders.
2. Bank eflowpaths are ignored and are also given the value *null* for all three orders.

The software code that implements the three orders varies somewhat from what is outlined below; however, what is described here is a fair description of the results. Developers are welcome to review the code.

Strahler Order

Strahler Order is determined entirely by the geometry of the eflowpaths with rank defined as primary.

1. Headwater stream segments (eflowpaths with one endpoint coincident with a headwater nexus) are assigned a value of 1.
2. Moving in a downstream direction, when two or more inflowing eflowpaths meet at a hydro nexus and are of different orders, then the outflowing eflowpath has a value equal to that of the inflowing eflowpath with the largest order.
3. When two or more inflowing eflowpaths meet with at least two of them of the same order and with that order higher than that of any other inflowing eflowpaths, then the outflowing eflowpath has a value equal to that order + 1.

Horton Order

Horton Order requires mainstems to be defined. As implemented here, this is done based on both name and length. CHyF uses a recursive definition, working upstream on primary flows (eflowpaths with rank = 1) that are not of type = bank.

1. For each eflowpath that ends at a terminal nexus (including those at the boundary of the area in question), the horton order is assigned the same value as the strahler order.

2. Moving upstream, if there are any outflowing eflowpaths at a hydro nexus with the same name as this inflowing eflowpath, then the one with the longest flowpath to a headwater nexus is selected as the mainstem and it is assigned the same horton order as the inflowing eflowpath.
3. Otherwise, if there are any outflowing eflowpaths with any name at all, the one with the longest flowpath to a headwater nexus is selected as the mainstem and it is assigned the same horton order as the inflowing eflowpath.
4. Otherwise, if there are any outflowing eflowpaths at all, the one with the longest flowpath to a headwater nexus is selected as the mainstem and it is assigned the same horton order as the inflowing eflowpath.
5. Otherwise, this is the top of this mainstem and it is assigned the same horton order as the inflowing eflowpath.
6. This process is repeated starting at the bottom again. Every unassigned eflowpath connecting to the newly defined mainstem is assigned a horton order of n-1, where n is the order of the mainstem. The process in steps two through six is then followed iteratively until all eflowpaths are assigned.

Hack Order

The logic behind the hack order is the same as that behind the horton order with two specific exceptions. In Step 1 above, the eflowpaths that end at a terminal nexus are assigned a hack order of 1. In Step 6, the hack order is incremented (n+1 instead of n-1).

Another minor change that is useful analytically is as follows. For all eflowpaths whose corresponding mainstem ends at a terminal flowpath, 1000 is added to the hack order. So for example the inferred eflowpaths in isolated lakes have a hack order of 1001. The exception to this rule is that the eflowpath at the mouth of the Richelieu keeps its calculated value of 1.