Freshwater Atlas
Mainstemming Technical Documentation

Submitted To: Mark Sondheim and Stephanie Forbes
Base Mapping and Geomatic Services Branch
Ministry of Agriculture and Lands
Province of British Columbia

Submitted By: Refractions Research Inc.
400 - 1207 Douglas Street
Victoria, BC, V8W-2E7
Phone: (250) 383-3022
Fax: (250) 383-2140

Date: December 2005

Table of Contents

Figures

Tables

1 MAINSTEMMING

1.1 Mainstem

A mainstem is a connected collection of flow edges which make up a river. Thus all the segments that make up the main flow of the Fraser River will be a mainstem. Mainstems only apply to primary flows therefore the side channels and distributaries of the Fraser River will not be a part of this mainstem, nor will they make up a mainstem; however they can be associated with the Fraser River mainstem.

This document outlines the rules and requirements for generating mainstems on the Freshwater Atlas dataset.

1.2 Program Requirements

The following data requirements must be met before mainstemming:

- Edges with correct codes
- directed data (upstream direction)
- non-cyclic data
- topologically clean data (properly noded)

The following tables in the PostGIS database are also required:

Table 1: Required PostGIS Tables

FWA_edges	Table containing all fwa edges and associated attributes including: - code - primary/secondary attribute - contained attribute
fwa_waterbodies	Table containing all waterbodies and associated attributes.
fwa_skel_wb_ids	Table linking edges to waterbodies (identifies skeletons and which waterbodies that are contained within)
fwa_named_streams	Table containing names matched to edges
fwa_fixed_mainstems	Table containing manually defined mainstems
fwa_mainstem_edges	Table containing mainstemming results
all_gnis_info	Table containing BC GNIS information
fwa.fwa_global_id_seq	Sequence for generating mainstem ids

1.3 Mainstemming Algorithm

1.3.1 General Algorithm

The general mainstemming algorithm is outlined below:

1. Find and flags all headwaters (glacial, river, lake).
2. Extend headwaters downstream.
3. Weight edges.
4. Find lake areas and associated with output edge.
5. Computer longest upstream path for each edge.
6. Generate mainstems.
7. QA/QC.

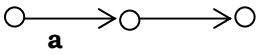
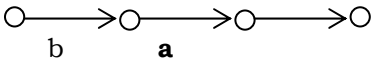
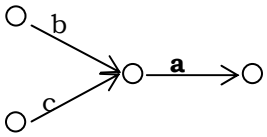
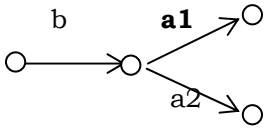
1.3.2 How Primary/Secondary Attributes are Computed

The computation of primary/secondary attribution is not completed during the mainstemming process, it is a pre-requisite. The following flow edge codes have well defined primary/secondary attributes:

Table 2: Flow Edge Codes

primary	secondary
1000	1100
1050	1150
1200	1300
1250	1350
1400	1475
2000	2300

Connector and Inferred edges (i.e. 1450, 1410, 1425) can be either primary or secondary depending on the edges around them. The following rules outline how the primary/secondary attribute is computed for these edges.

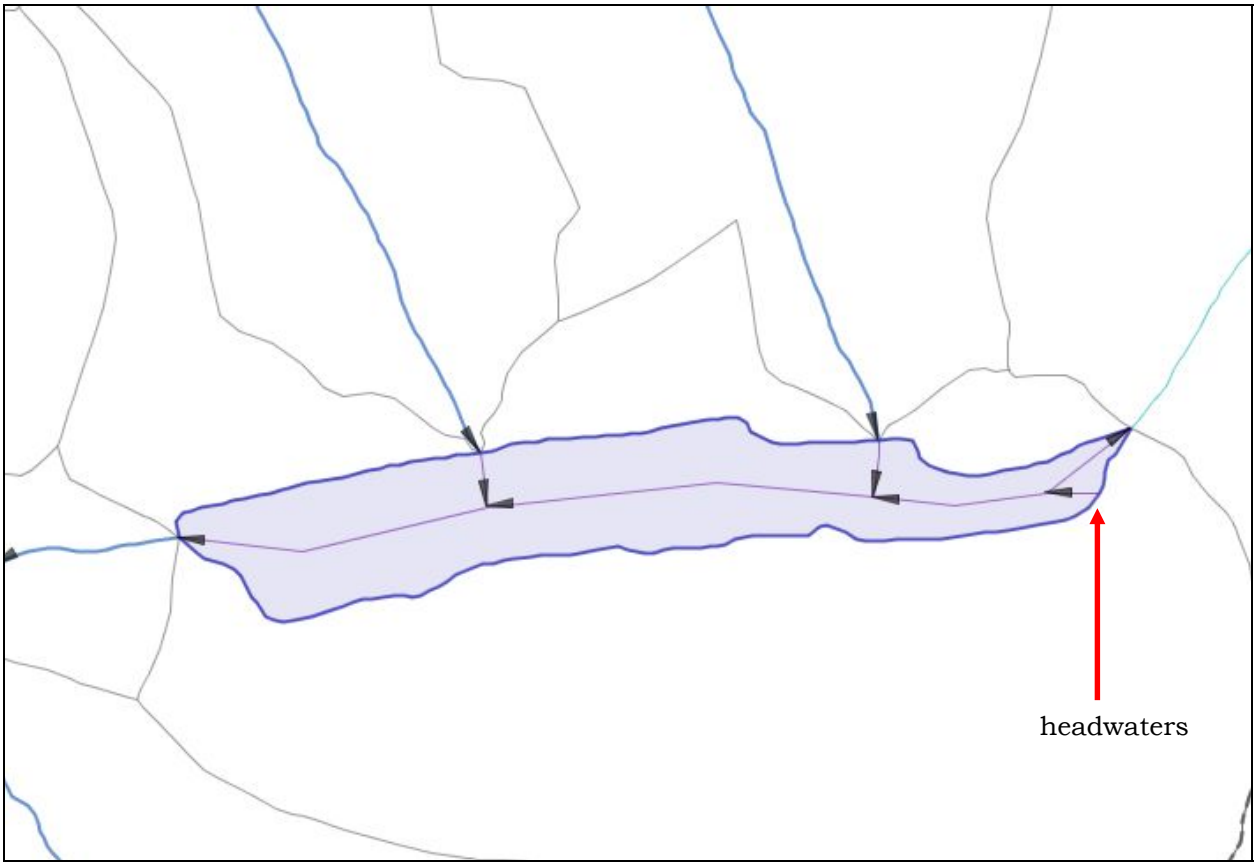
Source Edges 	a is primary
In-Line Edge 	a is given the primary/secondary coding as edge b
In Degree = 2 , Out Degree = 1 	if b or c is primary then a is primary else a is secondary
In Degree 1, Out Degree = 2 	If a1 is primary then a2 is secondary. If a2 is primary then a1 is secondary. Otherwise the edges are walked downstream and the connector edge (a1 or a2 or both) is assigned to the first primary/secondary edge downstream.

1.3.3 Headwaters

1.3.3.1 Lake Headwaters

Lake headwaters are identified as source edges inside a lake.

Figure 1: Lake Headwaters



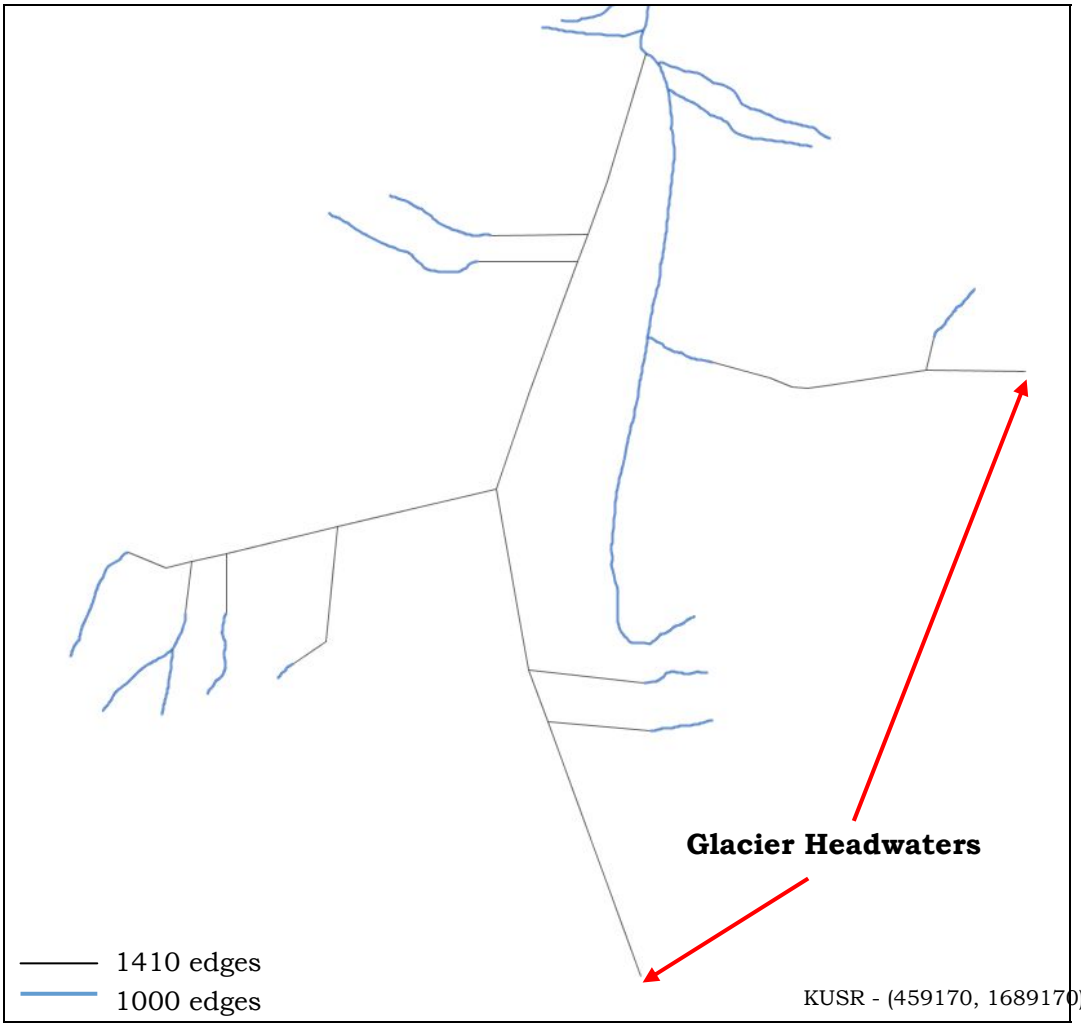
1.3.3.2 River Headwaters

River headwaters are identified as source edges inside a double line river.

1.3.3.3 Glacial Headwaters

Glacial headwaters are identified by finding source 1410's. Each source 1410 edge is assumed to be a glacier headwater.

Figure 2: Glacier Headwaters

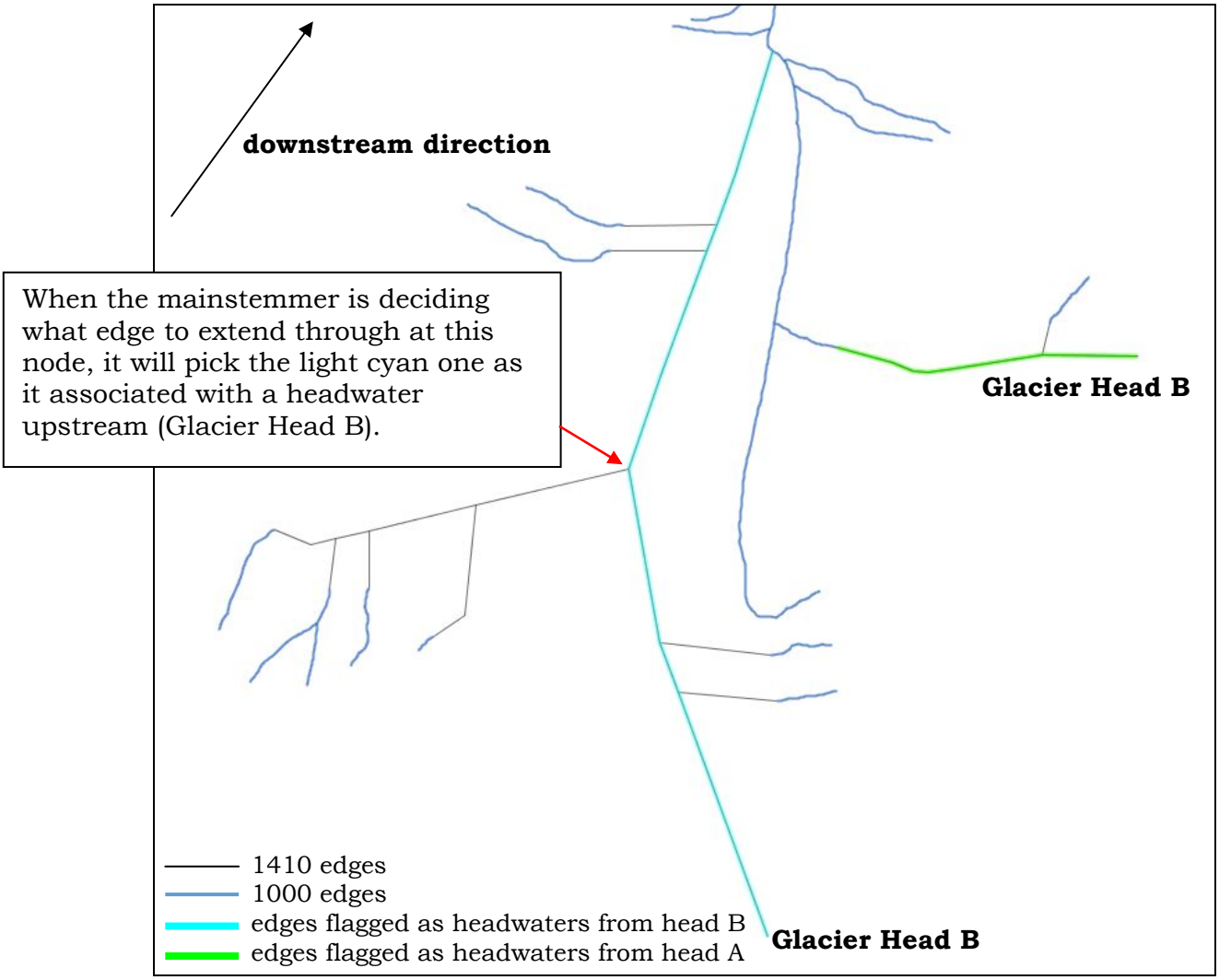


1.3.3.4 Extending Headwaters Downstream

To ensure mainstems push up to headwaters once they enter a lake, river, or glacier the edges downstream of the headwaters are also flagged as headwater edges. The headwater is pushed downstream until there are no downstream edges with the same code as the current headwater edge. (Lake headwaters are pushed down to the lake output, river headwaters are pushed down to the river output, and glacier headwaters push down to the edge of the glacier).

When the mainstemmer is making a decision as to what upstream edge to push the current mainstem, edges that are flagged as headwaters will take priority over other edge.

Figure 3: Extending Headwaters Downstream



1.3.4 Weighting Flow Edges

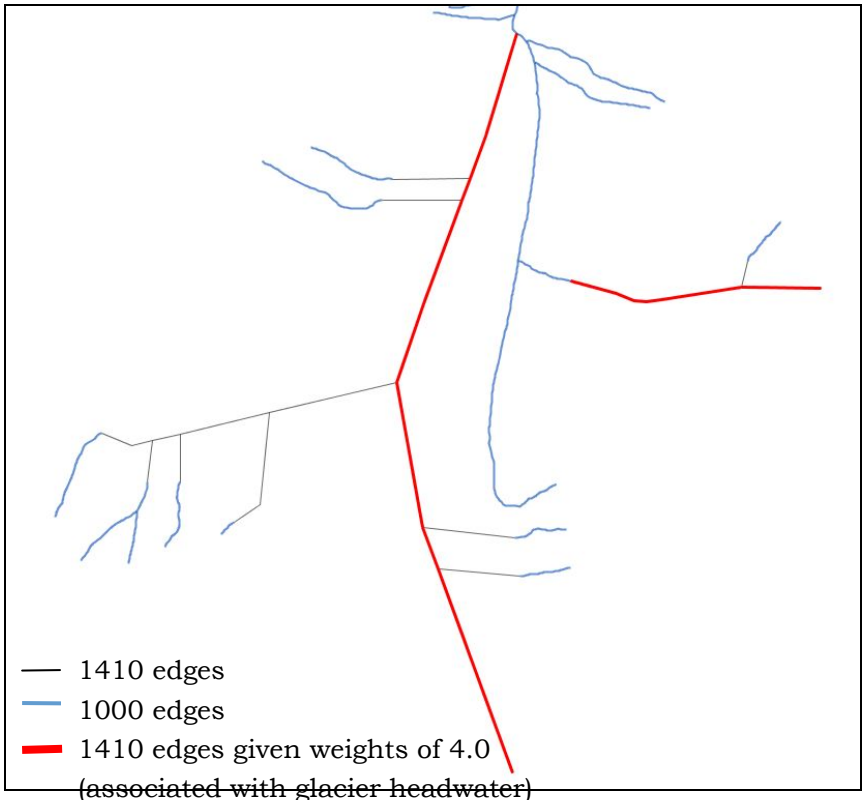
In order to more accurately estimate the upstream area using upstream length, certain edges are given higher weights than others. These weights are applied to the length of each edge.

Table 3: Edge Weights

Edge Code	Weight
1000	1.0
1050	1.0
1200	1.0
1250	3.0
1400	1.0
1425	1.0
1450	1.0
1410	4.0 for 1410 associated with glacier headwaters 1.0 for all other 1410
2000	0.5 for canals not inline with other flows 1.0 for all other canal edges

The following figure shows that the red arcs are given a weighting of 4 times there length. The red arcs are all coded as inferred flows (1410's), however these red arcs are associated with a glacier headwater; therefore they are weighted higher than other inferred connectors.

Figure 4: Glacier Weighting Example



The following figure shows both an in-line and non in-line canal.

Figure 5: Canal Weighting Example

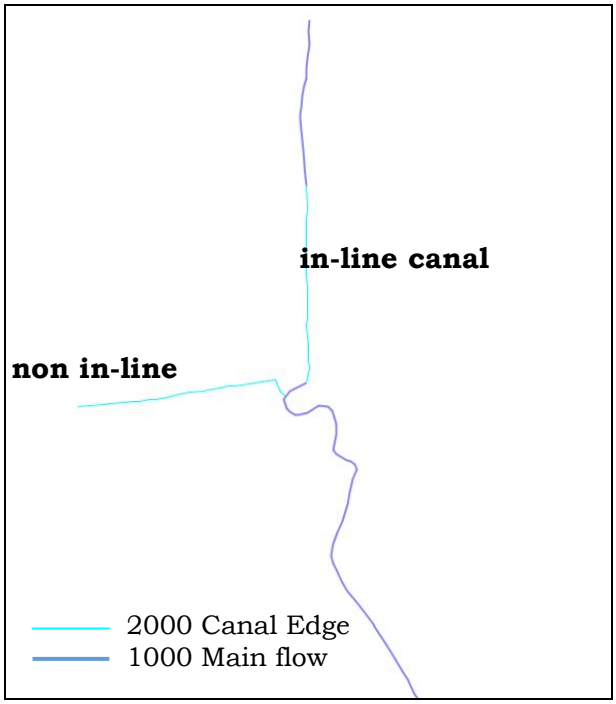
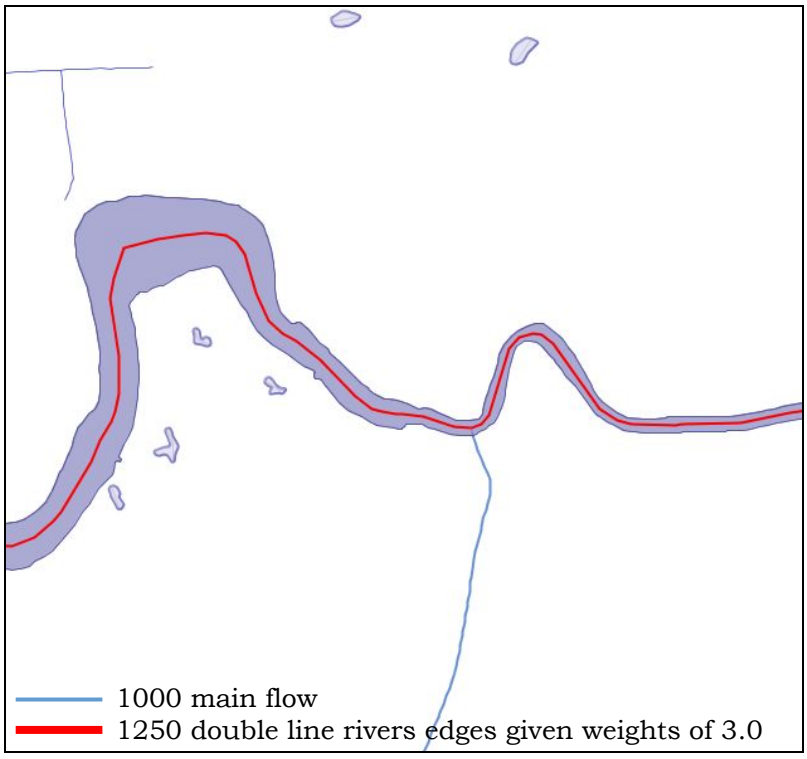


Figure 6 outlines the 1250 arcs that are given a weighting of 3.0.

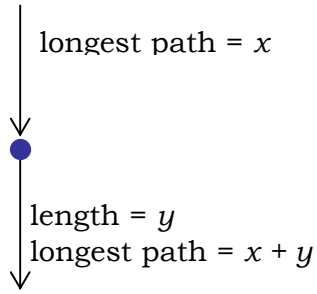
Figure 6: Double Line River Weighting Example



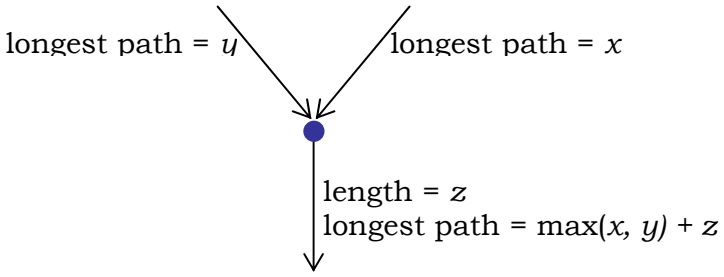
1.3.5 Calculating Longest Upstream Path

The longest upstream path is the sum of the length of the current edge with the maximum upstream length from the input edges.

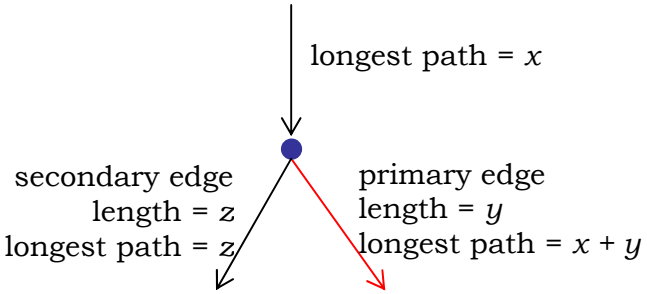
1.3.5.1 General Rules



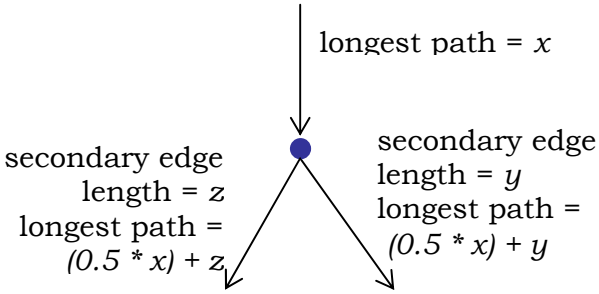
In Degree 1, Out Degree 1



In Degree 2, Out Degree 1

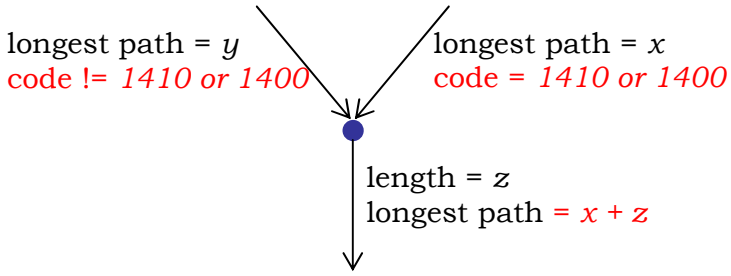


In Degree 1, Out Degree 2
(secondary & primary flow edge)



In Degree 1, Out Degree 2
(two secondary flow edges)

1.3.5.2 Exceptions



In Degree 2, Out Degree 1

If one in edge is coded 1410 or 1400 and the other is not coded 1410/1400 then the upstream length of the 1410/1400 is not pushed downstream

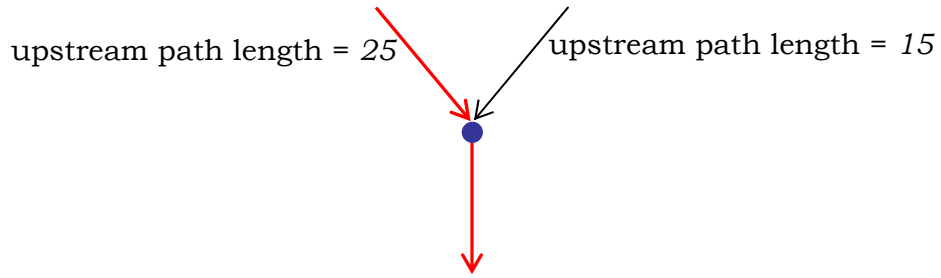
1.3.6 Upstream Lake Area

To approximate the upstream lake area, the edges associated with the primary exit from the edge have the lake size attached to it. This information is pushed downstream using the same rules and pushing down upstream length.

1.3.7 Computing Mainstems

1.3.7.1 *General Rule*

The general rule for computing mainstems is whenever there are two upstream paths to select from the one with the longest upstream length.



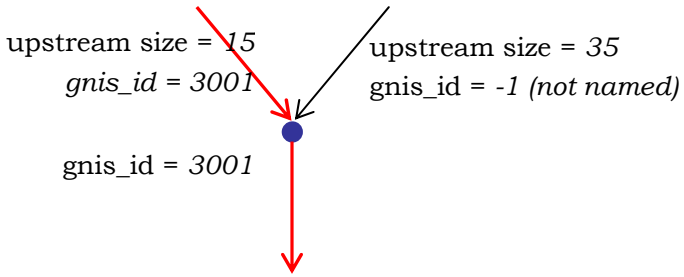
General Rule: The upstream edge with the larger upstream length is selected.

1.3.7.2 Exceptions

The exceptions to the longest upstream path rule are outlined below.

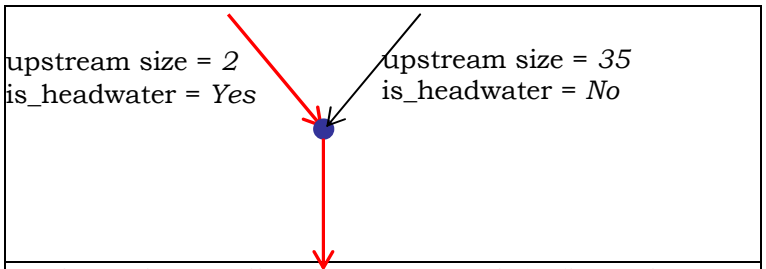
1. Manually Defined Mainstems
   - If the current edge is part of a manually defined mainstem and one of the upstream edges is part of the same manually defined mainstem, then this edge is selected.

2. Named Streams
   - If there is an upstream edge with the same name as the current edge, the mainstemmer must follow up this edge.



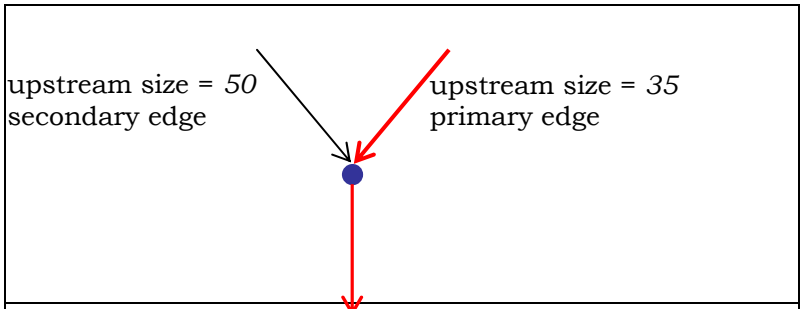
An upstream edge with the same gnis_id as the downstream edge exists therefore the mainstemmer must select the named upstream edge.

3. Headwaters
   - If one of the upstream edges is flagged as a headwater and the other is not, the one flagged as a headwater must be followed. If they are both flagged as headwaters the other rules are used.



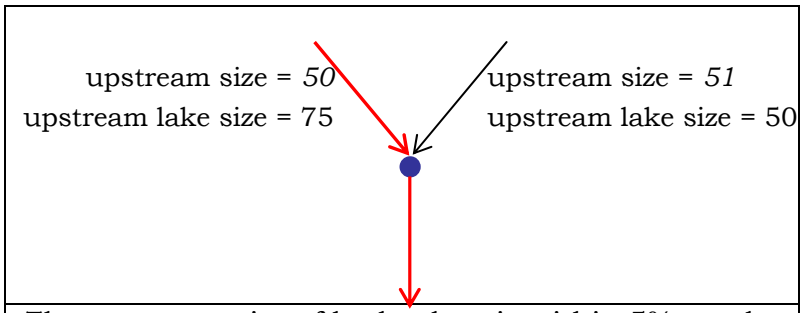
There the smaller upstream path is flagged as a headwater so this edge is selected.

4. Edge Type
   - Only primary flows are included in mainstems, thus if there is one primary edge and one secondary edge the primary edge is selected.



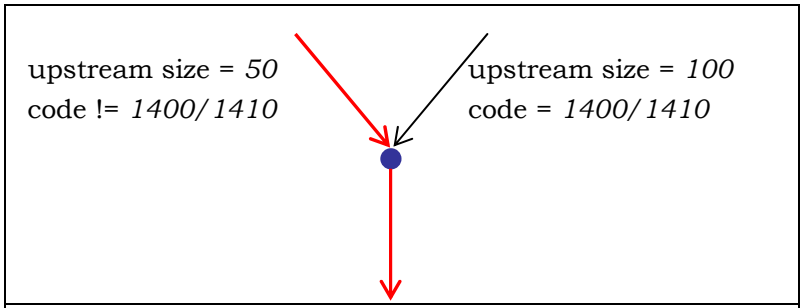
The secondary flow edge has a larger upstream size, but secondary flow edges are not mainstemed.

5. Similar Upstream Size
   - If both upstream edges are primary edges and the upstream size is within 5% of each other, the edge with the larger upstream lake size is selected (unless they are the same in which case the larger upstream length is used).



The upstream size of both edges is within 5%, so the upstream lake size is used to determine which edge to select.

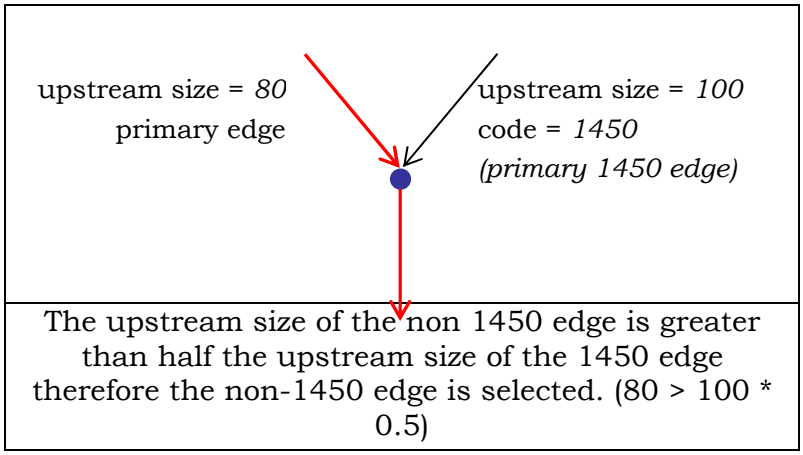
6. Edge Codes
   - If one edge is a 1410/1400 and the other is not a 1410/1400, then the non 1400/1410 edge is selected.



The larger upstream size edge is a 1400/1410 edge, but another non 1400/1410 upstream edge exists therefore this edge is selected.

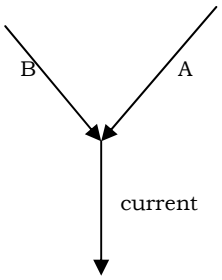
7. 1450

- If a primary 1450 edge is selected and the other edges upstream size is less than 2 times current edges upstream size than the other upstream edge is selected (this is to keep mainstems inside double line rivers as much as possible).



The upstream size of the non 1450 edge is greater than half the upstream size of the 1450 edge therefore the non-1450 edge is selected. (80 > 100 * 0.5)

1.3.7.3 Processing Outline



Is the *current* edge part of a fixed mainstem*?

yes → Is A or B part of the same fixed mainstem?

yes → Select edge (A or B) with the same fixed mainstem as *current*

no

no → Is the *current* edge named?

yes → Does A or B have the same name?

yes → Select edge (A or B) with same name

no

no → Is A primary and B secondary (or visa versa)

yes → Select primary edge (A or B)

no

Are A and B both secondarv

yes → Stop Mainstem

no

Is A inferred (1410) and B observed (or visa versa)

yes → Pick non-inferred edge

no

Is A connector (1450), input to A inferred (1410), and B observed (or visa

yes → Pick observed edge

no

Is A headwater (glacial/lake/river) and B not a headwater (or visa

yes → Pick Headwater edge

no

Is A's largest upstream path within 5% of the B's longest upstream path?

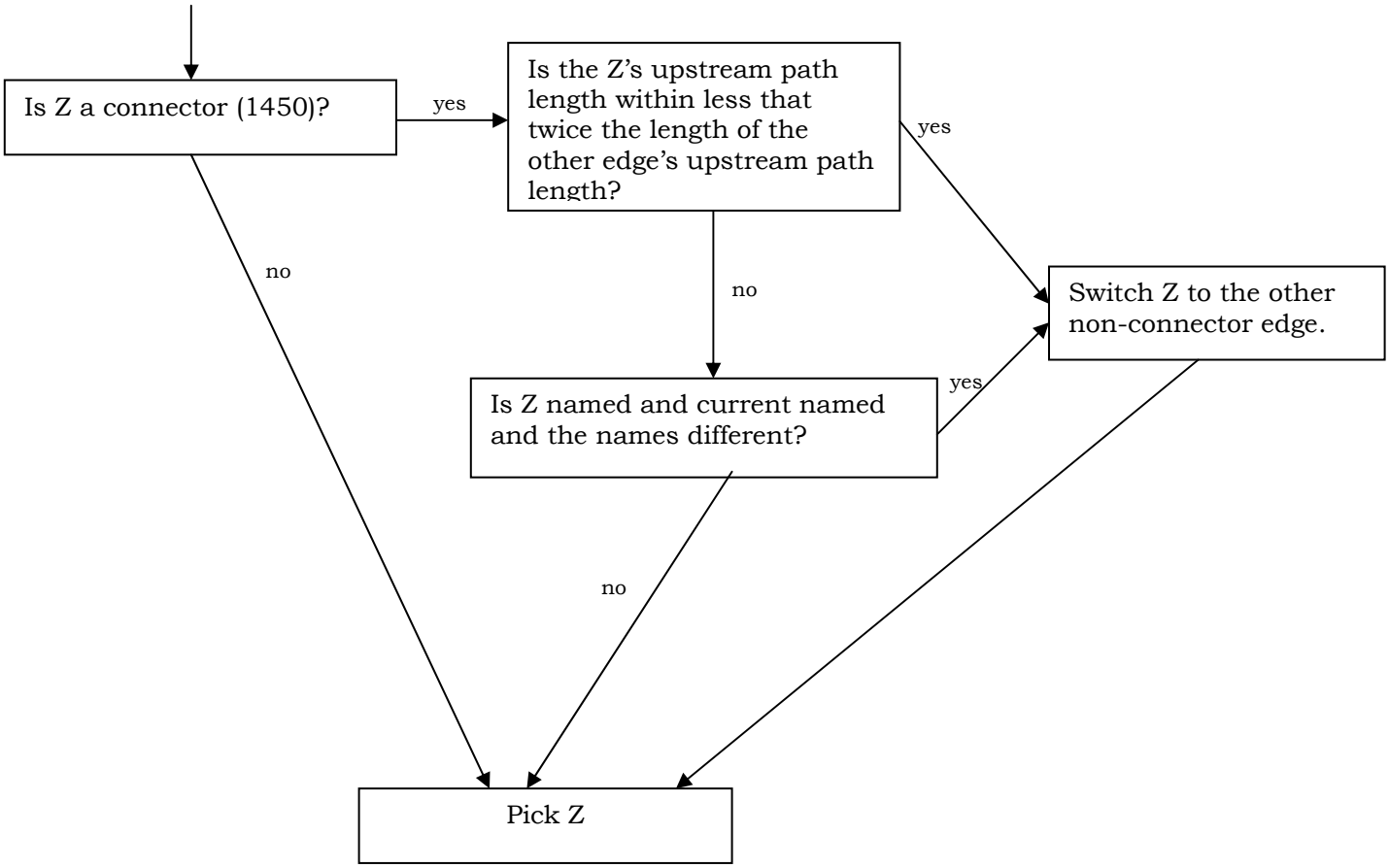
yes → Does A and B have different upstream lake areas

yes → Pick edge (A or B) with the largest upstream lake area

no

no

Pick edge 'Z' (A or B) with longest upstream path



* fixed mainstems ⇔ manually defined mainstems (must follow a given path)