Natural Resources Canada   Ressources naturelles Canada

GEOMATICS CANADA
OPEN FILE 65

Common hydrology features (CHyF) logical model

Version 1.0

M. Sondheim and C. Hodgson

2024

Canada

GEOMATICS CANADA
OPEN FILE 65

Common hydrology features (CHyF) logical model

Version 1.0

M. Sondheim[1] and C. Hodgson[2]

[1]Canada Centre for Mapping and Earth Observation, 506 Burnside Road West, Victoria, British Columbia (Deceased)
[2]Refractions Research, Inc., 1207 Douglas Street, Victoria, British Columbia

2024

Permanent link: https://doi.org/10.4095/328952

This publication is available for free download through GEOSCAN (https://geoscan.nrcan.gc.ca/).

Recommended citation
Sondheim, M. and Hodgson, C., 2024. Common hydrology features (CHyF) logical model; Geomatics Canada, Open File 65 (ver. 1.0), 72 p. https://doi.org/10.4095/328952

Publications in this series have not been edited; they are released as submitted by the author.

ISSN 2292-7875
ISBN 978-0-660-49927-7
Catalogue No. M103-3/65-2024E-PDF

Authorship

Mark Sondheim,
Canada Centre for Mapping and Earth Observation, Natural Resources Canada

Chris Hodgson,
Refractions Research, Inc.

Acknowledgements

The authors would like to thank the following individuals for their guidance and commentary on CHyF and its objectives.

David Blodgett,
Water Resources Mission Area, United States Geological Survey

James Craig,
Department of Civil and Environmental Engineering, University of Waterloo

Emily Gouge,
Refractions Research, Inc.

Marie-Eve Martin,
Canada Centre for Mapping and Earth Observation, Natural Resources Canada

The authors also wish to acknowledge the highly instructive input from other individuals, including:

Boyan Brodaric,
Geological Survey of Canada, Natural Resources Canada

Eric Boisvert,
Geological Survey of Canada, Natural Resources Canada

Kevin Shook,
Global Institute for Water Security, University of Saskatchewan

Alan Rea,
National Geospatial Program, United States Geological Survey

Peter Steeves,
National Geospatial Program, United States Geological Survey

Wayne Jenkinson,
International Joint Commission

Nick Mazany-Wright,
Canadian Wildlife Federation

Table of Contents

i. Abstract

The Open Geospatial Consortium has defined "OGC® WaterML 2: Part 3 - Surface Hydrology Features (HY_Features) - Conceptual Model", but not any particular implementation of it. The Common Hydrology Features (CHyF) model extends HY_Features and makes some minor changes to it required for implementation and the delivery of high performance services. HY_Features discusses catchment coverage and topological relations. In CHyF these are key ideas, as is the notion that hydrologically defined network components form elements of a mathematical graph, allowing for very fast network traversal.

HY_Features defines catchments and catchment networks, as well as rivers, channels, flowpaths and hydrographic networks. The CHyF logical model specifies a profile and some extensions to HY_Features, as required to implement topological and graph relations. This starts with the definition of elementary catchments and elementary flowpaths, which are treated as fundamental elements. They are tightly specified terms corresponding to basic catchments and flowpaths in HY_Features and the basic components in the standard reach-catchment model (Maidment and Clark, 2016). If they are subdivided, the result is simply more elementary catchments and elementary flowpaths. Consequently, they are the building blocks used to form complementary coverages as well as a graph structure referred to as a *hygraph*. Building the hygraph necessitates that connections between features be manifest through their geometry. Divergences and distributaries are supported in CHyF, as the hygraph need not be hierarchical. Nevertheless, CHyF does recognize hierarchical drainage basins and the value in identifying them explicitly (Blodgett, et al, 2021).

Different kinds of elementary catchments and elementary flowpaths are defined in CHyF. Of note is that polygonal waterbody features, or portions of such features, are treated as elementary catchments in their own right. In addition to these water catchments, several kinds of land-based elementary catchments are recognized. These model constructs are compatible with the higher level conceptual model in HY_Features, although they differ in detail from other popular implementation models. With the approach taken it becomes practical to handle very large lakes and rivers, as well as coastal ocean zones. CHyF also includes wetlands, glaciers and snowfields as kinds of hydro features; these features help complete the concept of a catchment coverage as put forward by HY_Features.

ii. Keywords

The following are keywords to be used by search engines and document catalogues.

CHyF; HY_Features implementation; implementation specifications; hydrologic features model; hydrologic services; hydrographic features.

iii. Preface

The CHyF (pronounced *chief*) logical model provides an implementation of the HY_Features conceptual model. It also borrows directly from hydrologic and topologic concepts, and from graph theory. It defines a geospatial data specification directly amenable to the development of open hydrologic services accessible over the World Wide Web. However, it is fully suitable for use on traditionally architected projects where the web is not a factor. In either case CHyF is designed to be performant and scalable.

CHyF impacts data management in several ways. It significantly reduces the amount of work that must be undertaken to maintain a database for a given area or for an entire continent. The general level of complexity of the data model is much less than that of some alternatives. Local adjustments to the data are not required to meet specific model limitations, such as the requirement by some coding systems for dendritic structures at all scales. Updates can be handled comparatively easily without the need to continuously support and conflate large numbers of feature identifiers.

The power of CHyF comes about because of its use of a mathematical graph, referred to here as a hygraph. This graph is similar in intent and design to graphs built for navigation through road networks. It is specialized in that it is tailored for hydrologically related networks composed of elementary catchments and flowpaths.

1. Scope

This document describes the CHyF logical model, a profile of the HY_Features conceptual model and also an implementation of it. CHyF is intended to meet the needs of hydrologists and environmental professionals concerned with hydrologic assessments or the health of watersheds and river systems.

CHyF pertains to the description and behaviour of hydrologic features found on the surface of the earth. It is relevant to such questions as: where does surface water exist, how much of it is there, how does it change over time, what is upstream or downstream of what, how are land features connected to water features, and how can other types of data be related to water features.

Groundwater and atmospheric phenomena must be taken into account with surface water features for many applications. However, they are out of scope for this document. Nevertheless, recent work demonstrating such links is described elsewhere.

Familiarity with HY_Features is helpful, but not mandatory. The model presented in HY_Features is similar in many respects to that used as the basis of the United States NHD (National Hydrographic Dataset) and its offshoots, the NHDPlus and the NHDPlus HR (High Resolution). So knowledge of any of these and their related developments (e.g., StreamStats) is also relevant, but again not necessary.

The structure of the document is similar to that of many Open Geospatial Consortium (OGC) publications. In all cases where relevant, the relationship to HY_Features is noted. In some cases, the relationship to the NHD suite of standards is also indicated. This is done to help foster semantic interoperability, and where differences exist, to help clarify why they do. Thus, reference is also made to the terminology found in the WMO/UNESCO "International Glossary of Hydrology".

2. Conformance

CHyF conforms to the HY_Features conceptual model. Appendix 1 clarifies the relationship through provision of a CHyF – HY_Features Crosswalk. Similar mappings are also included in Appendix 2 to Canada's National Hydro Network (NHN) and the NHDPlus model of the United States.

CHyF also conforms to a number of existing OGC and ISO standards, as listed in section 3 below. In particular, geometric representations follow the Simple Feature Access (ISO 19125) model and the specification in general aligns with the SQL/MM Spatial (ISO/IEC 13249-3) specification.

Because CHyF conforms with the standards and specifications noted above, popular open source and commercial geospatial software can work directly with CHyF compliant data, available in GeoJSON and GeoPackage formats.

CHyF also provides web services that follow a RESTful design, making use of HTML and GeoJSON. They are architected in line with common practice on the Web, and thus are in support of general interoperability. Currently, they are specified independently of OGC Web Services. In the future this may change as the OGC moves toward REST.

Formal modeling is expressed through UML diagrams, in line with OGC practices. However, for clarity, simple logic diagrams are used in the section on Terms and Definitions. These diagrams are similar to class and subclass definitions in UML but are more accessible to a broader audience.

3. References

This document contains references to the following:

A Land Use And Land Cover Classification System For Use With Remote Sensor Data; by James R. Anderson, Ernest E. Hardy, John T. Roach, and Richard E. Witmer; Geological Survey Professional Paper 964, United States Geological Survey (http://www.pbcgis.com/data_basics/anderson.pdf)

A new topological and hierarchical river coding method based on the hydrology structure. Journal of Hydrology. un Wang, Denghua Yan, Tianling Qin, Baisha Weng, Hao Wang, Wuxia Bi, Xiangnan Li, Batsuren Dorjsuren. 2020. Journal of Hydrology. (https://www.sciencedirect.com/science/article/abs/pii/S0022169419309783)

A topological system for delineation and codification of the Earth's river basins. K.L. Verdin and J.P. Verdin. 1999. Journal of Hydrology. (https://doi.org/10.1016/S0022-1694(99)00011-6)

An Operational Mean High Water Datum for Determination of Shoreline Position from Topographic Lidar Data. Kathryn M. Weber, Jeffrey H. List, and Karen L. M. Morgan. USGS. (https://pubs.usgs.gov/of/2005/1027/index.html)

Australian Hydrological Geospatial Fabric (Geofabric) (http://www.bom.gov.au/water/geofabric/index.shtml)

Coastal Wiki, Definitions of coastal terms; Coastal Wiki (http://www.coastalwiki.org/wiki/Definitions_of_coastal_terms)

DE-9IM (Dimensionally Extended nine-Intersection Model) (https://en.wikipedia.org/wiki/DE-9IM)

Guidance for Flood Risk Analysis and Mapping: Ice-Jam Analyses and Mapping. (FEMA) 2018. (https://www.fema.gov/media-library-data/1520964160270-7c49e1753d0b2634e0c5fb4999459374/Ice_Jam_Guidance_Feb_2018.pdf)

Ice, Snow, and Glaciers: The Water Cycle. United States Geological Service. (USGS) (https://water.usgs.gov/edu/watercycleice.html)

International Glossary of Hydrology / Glossaire International d'Hydrologie. World Meteorological Organization, United Nations Educational, Scientific and Cultural Organization (eds.). WMO (Series), no385. WMO, Geneva (2016). ISBN 978-92-63-03385-8. ISBN 978-92-3-001154-3 (https://hydrologie.org/glu/HINDEN.HTM)

Mainstems: A logical data model implementing *mainstem* and *drainage basin* feature types based on WaterML2 Part 3: HY_Features concepts. David Blodgett, J. Michael Johnson, Mark Sondheim, Michael Wieczorek, Nels Frazier. 2021; Environmental Modelling & Software. (https://www.sciencedirect.com/science/article/pii/S1364815220309841?via%3Dihub)

NHDPlus and the National Water Model. 2016. David Maidment and Edward Clark.
(https://www.epa.gov/sites/production/files/2018-01/documents/nhdplus_and_the_national_water_model_-_maidment.pdf)

NHDPlus Version 2: Users Guide (Data Model version 2.1); 2019; USGS (NHDPlus)
https://s3.amazonaws.com/nhdplus/NHDPlusV21/Documentation/NHDPlusV2_User_Guide.pdf

OGC Abstract Specifications
(http://www.opengeospatial.org/docs/as)

OGC® WaterML 2: Part 3 - Surface Hydrology Features (HY_Features) - Conceptual Model
(http://docs.opengeospatial.org/is/14-111r6/14-111r6.html)

Oxford Dictionaries
(https://en.oxforddictionaries.com/)

Principles of Design and Operations of Wastewater Treatment Pond Systems for Plant Operators, Engineers, and Managers; 2011; U.S. Environmental Protection Agency
(https://www.epa.gov/sites/production/files/2014-09/documents/lagoon-pond-treatment-2011.pdf)

The Ramsar Convention Manual, 6th edition; 2013; Ramsar Convention Secretariat
(https://www.ramsar.org/sites/default/files/documents/library/manual6-2013-e.pdf)

What is a wetland? 2017; U.S. Environmental Protection Agency
(https://www.epa.gov/wetlands/what-wetland)

Wikipedia; Wikipedia.org
(https://en.wikipedia.org/wiki/Main_Page)

WordNet Search; Princeton University
(https://wordnet.princeton.edu/)

Chapter 4. Terms and Definitions

The following modelling terms and definitions are used in this document. To ease understanding, they are provided within groups that correspond to the primary entities within the CHyF model. These concepts are taken from HY_Features and extended as needed to relate effectively to compatibility with a graph implementation and with specific geospatial concerns.

To aid understanding, for each of the high level concepts, a simple logic model is shown, depicting the further breakdown of the concept. This begins with CHyF as a domain, with the major types of objects encompassed by CHyF shown on the right in Figure 1. In subsequent chapters the CHyF model is fully explained and UML diagrams are provided. These figures should not be interpreted as equivalent to UML class diagrams; in some cases they do correspond to class – subclass relationships, but in other cases they do not.

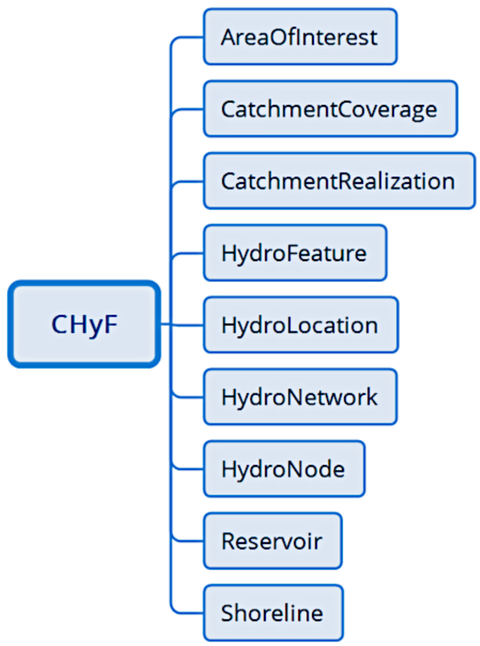

Figure 1: Primary concepts in the CHyF Hydro Fabric

Another term of interest is hydro fabric. In the diagram above, *CHyF* could be *CHyF Hydrological Geospatial Fabric*, or more simply, *CHyF Hydro Fabric*, where the term hydro fabric implies that the various elements are woven together to form a comprehensive terrain model consisting of features pertaining to water phenomena. This is comparable in intent to the Australian Hydrological Geospatial Fabric.

4.1 Area of Interest

The geographic extent of a set of hydrologic or hydrologically related features, as represented on a map or in a dataset or database, and typically corresponding to the area occupied by a drainage basin or drainage area. This area of interest (AOI) definition is useful in an operational context and is relevant to CHyF services.

Figure 2: Area of Interest (AOI)

4.2 Catchment Coverage

The concept of a geospatial coverage is key to the provision of various CHyF services. A coverage can be defined from elementary catchments but also from larger catchments, including major drainage basins or catchment aggregates. It can also be based on drainage areas composed of catchments, even though the drainage area does not necessarily meet the definition of a catchment.

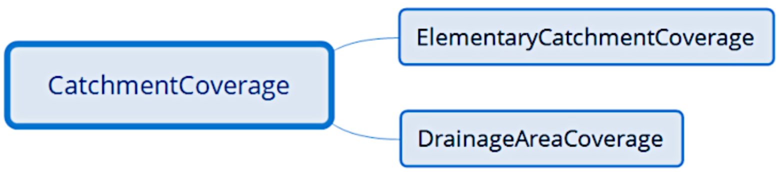

Figure 3: Drainage area coverage and elementary catchment coverage

4.2.1 Drainage Area Coverage

Large drainage areas are often defined to support data management and high level analysis. Often they are catchments, but for practical reasons this is not always the case. The USGS Hydrologic Units at any given Hydrologic Unit Code (HUC) level are an example of this situation. Where sufficient detail is available, the drainage areas within a drainage area coverage can be defined through aggregation of elementary catchments, with no gaps or overlaps.

Aggregations of small catchments along coastlines do not form catchments, but they do form drainage areas. Consequently, any continental level coverage must be defined in the context of drainage areas, even though the majority of them may be catchments.

4.2.2 Elementary Catchment Coverage

A geospatial coverage is formed from a set of elementary catchments with no gaps and no overlaps within the area of interest. Since polygonal waterbodies are treated as catchments, and since areas that do not include streams and areas that are internally drained also define catchments, 100% of the terrain is within the coverage.

4.3 Catchment Realization

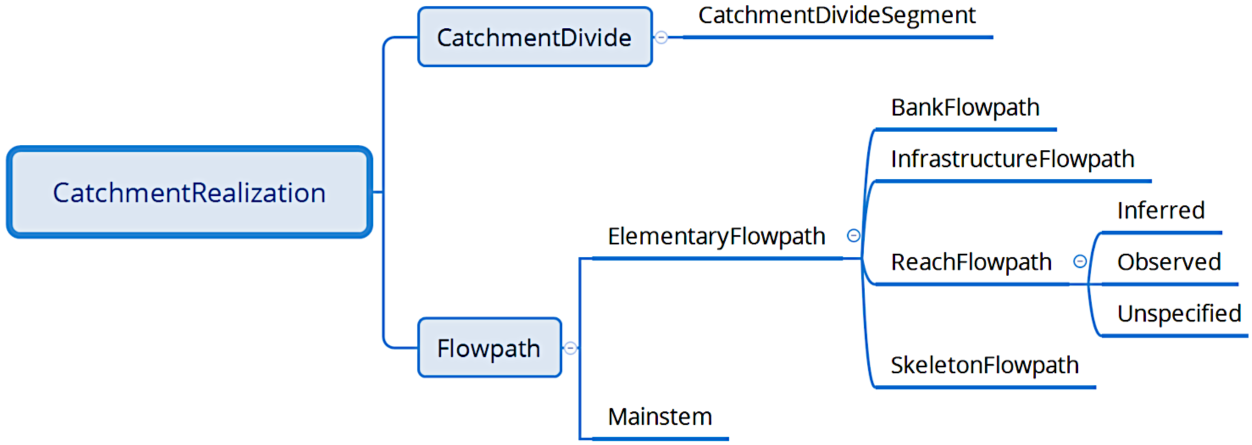

Figure 4: Catchment realizations

HY_Features introduces the general notion of catchment realization. In CHyF a catchment may be realized by: (i), a catchment divide, i.e., a boundary or portion of a boundary of a catchment, or (ii), a flowpath, representing the idealized one-dimensional drainage of water from a catchment.

4.3.1 Catchment Divide

A catchment divide forms all or a portion of the boundary of a catchment. It is a one-dimensional (linear) feature [after HY_F]. It is derived from the geometry of the waterbodies and an elevation model that may be based on a point cloud or a gridded dataset. A catchment divide may be instantiated as a single linestring, which may or may not form a linear ring.

4.3.1.1 Catchment Divide Segment

A portion of a catchment divide, represented by a linestring, that forms an edge between nodes, in a mesh consisting of all catchment boundaries.

4.3.2 Flowpath

A derived linear feature that realizes a catchment specifically as a path connecting the inflow or headwater start point with the outlet of the catchment [after HY_F]. A flowpath may be instantiated as a single, directed linestring or a series of connected, directed linestrings. The direction of a flowpath is always downstream.

4.3.2.1 Elementary Flowpath

A flowpath terminated at either end by a hydronode, such as a confluence point, a headwater start point, a terminal point where a river empties into a lake or the ocean, or the place on a flowpath where the data is terminated. An elementary flowpath has an attribute *rank*, which designates whether it represents a primary or secondary flow. It also has a series of order attributes (*strahlerOrder*, *hortonOrder* and *hackOrder*), and a *nameString* attribute.

4.3.2.1.1 Bank Flowpath

An elementary flowpath that connects a bank catchment to a skeleton flowpath in a waterbody with polygonal geometry. It is otherwise similar to a skeleton flowpath. The direction of flow is always away from the bank catchment, as it acts as a proxy for the drainage along the shoreline (i.e., the bank) from the land to the water.

4.3.2.1.2 Infrastructure Flowpath

An elementary flowpath representing: (i) a flow contained in a conduit, such as a storm drain or a sanitary sewer, or (ii) a flow through a dam or an industrial complex. Conduits may be buried, at ground level, or elevated.

4.3.2.1.3 Reach Flowpath

An elementary flowpath that corresponds to a segment of a single-line river (a river represented by linear geometry), ditch or canal. Three variants of a reach flowpath are recognized:

- Inferred: the reach flowpath corresponds to a presumed channel that appears to traverse the land but was not clearly visible or distinguishable when mapped.
- Observed: the reach flowpath represents the flow that is observed in a natural or anthropogenic channel represented by linear geometry.
- Unspecified: the type of reach flowpath is impossible to determine, or unrelated to the other existing subtypes.

4.3.2.1.4 Skeleton Flowpath

An elementary flowpath that represents the path of the flow of water through a waterbody with polygonal geometry. It is similar to a bank flowpath, but does not connect to a bank catchment. Ideally a skeleton flowpath aligns with the thalweg through a river or lake, or alternatively, it acts as a connector to the thalweg. In practice and without bathymetry, it is typically placed along the medial axis of a polygonal waterbody or along a connecting section to the medial axis. Skeleton flowpaths usually form dendritic patterns in their respective waterbodies; however, in the presence of islands, secondary flows around the islands are defined.

4.3.2.2 Mainstem

A directed linear feature that traces flow to the outlet of a drainage basin from its headwater source [Blodgett, et al, 2020]. The mainstem is a flowpath that can be constructed from elementary flowpaths. The particular path is chosen based on stream name, longest upstream length, largest cumulative upstream length, largest upstream area, estimated flow volume, or some other measure of importance. Each drainage basin (see section 4.3.1.3) has a corresponding mainstem, which is considered a realization of it.

4.4 Hydro Feature

“Feature of a type defined in the hydrology domain, whose identity can be maintained and tracked through a processing chain from measurement to distribution of hydrologic information.” [HY_F]. This is a high level construct that includes waterbodies, catchments, and other hydrologically related features. Ice Snow, Nearshore Zone, and Wetland are included in CHyF for reasons of semantic completeness and user requirements. Similarly, Buried Infrastructure is included to capture drainage through urban environments as well as dams and industrial complexes.

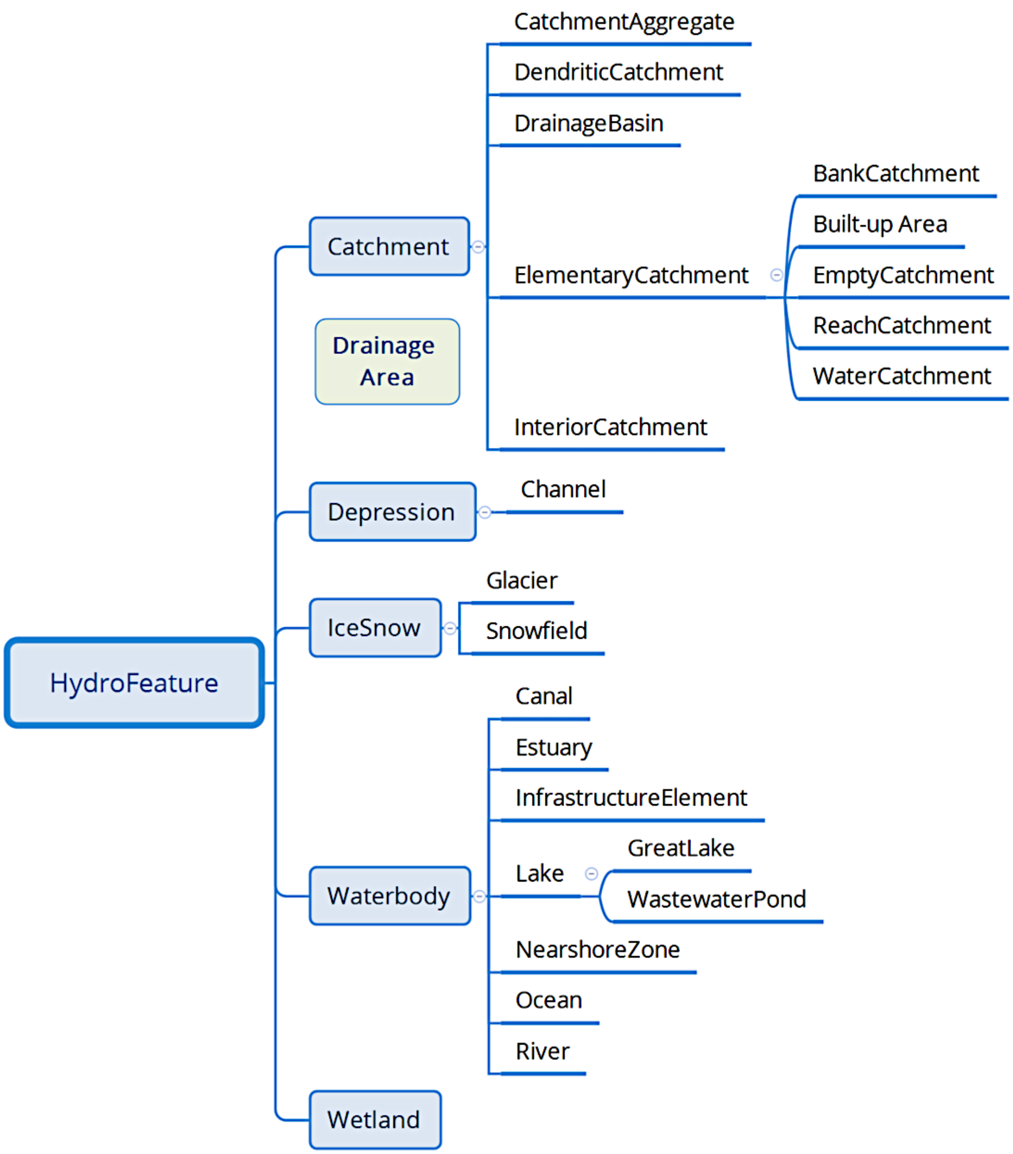

Figure 5: Hydro features

4.4.1 Catchment

“A physiographic unit where hydrologic processes take place. This class denotes a physiographic unit, which is defined by a hydrologically determined outlet to which all waters flow …” [HY_F]. A number of types of catchments are recognized, as shown in figure 4. Catchments (and drainage basins) may have nested, hierarchical relationships; in western North America, the Similkameen River drainage basin is contained in the Okanagan River basin, which in turn is contained in the Columbia River basin.

CHyF allows for a strict definition of a catchment, in which case interior catchments are excluded and all areas explicitly drain to a common outlet. It also allows for a looser definition in which all fully contained interior catchments are considered as part of a large catchment; in this case it should technically be referred to as a catchment aggregate. The Richelieu River Watershed in Quebec for example contains nearly 300 lakes that are not connected to nearby rivers through the mapped surface water network. Each of these lakes and its surrounding area constitutes an interior catchment. Depending upon the application, it may or may not be of interest to include those areas in the Richelieu catchment.

In HY_Features a catchment may have zero or one inflows and zero or one outlets. Because CHyF treats polygonal waterbodies (or subdivisions of them) as water catchments, which may have multiple inflows and multiple outlets, this rule is relaxed. It is replaced by a rule with a similar intent. In CHyF, the hydro nexus where one catchment meets another catchment must be an outlet for all catchments that drain into it. If River A drains into River B, then the catchment of River B contains the catchment of River A. The drainage of B minus the drainage of A can be considered a drainage area but not a catchment. This is described in more detail in section 6.1.2.6.

4.4.1.1 Catchment Aggregate

A catchment type defined "... as a set of non-overlapping dendritic and interior catchments arranged in an encompassing catchment." [HY_F]. It is not a general term for an aggregation of adjacent catchments. Instead it is intended to describe hierarchical systems based on dendritic catchments; it may also contain interior catchments.

4.4.1.2 Dendritic Catchment

"Catchment in which all waters flow to a single common outlet. A dendritic catchment is permanently connected to others in a dendritic (tree) network …" [HY_F]. If secondary flows around islands or across deltas exist, then the catchment may still be considered as dendritic, so long as the primary flows are identified and so long as they form a dendritic pattern.

4.4.1.3 Drainage Area

A drainage area is a generic term for a hydrologic unit, which may be compliant with the Hydrologic Unit Code system of the United States Geological Survey. Alternatively, it may be any area with a hydrologically determined, drainage-related boundary. A drainage area may be a catchment if it meets the specific criteria defining a catchment.

4.4.1.4 Drainage Basin

A catchment with no inflows and a single outflow, and with a corresponding mainstem. Other terms for drainage basin are total accumulated catchment or simply watershed. It generally refers to a medium to large catchment with a minimum size measured in tens of square kilometres. Drainage basins are generally hierarchically structured [Blodgett, et al, 2020]. All drainage basins are catchment aggregates, but the reverse is not true, as a catchment aggregate may not meet the criteria for a drainage basin.

The Tennessee River Drainage Basin is contained in the Ohio River Drainage Basin, which in turn is part of the Mississippi River Drainage Basin. A similar hierarchy exists with the basins associated with the Gatineau River, the Ottawa River and the Saint Lawrence River. These examples are three tiers deep, but they can be defined at more detailed levels as well. However, for practical reasons, the smallest drainage basins are still much larger than the available detail from contained catchments.

4.4.1.5 Elementary Catchment

A catchment defining a fundamental subdivision of the landscape in which water can be modelled as draining to a single outlet, to an adjacent waterbody, or internally to an area devoid of waterbody features. An elementary catchment is bounded by other elementary catchments; that is, elementary catchments compose a complete coverage. Five types of elementary catchments are recognized: reach catchments, bank catchments, water catchments, empty catchments, and built-up areas. An elementary catchment is generally equivalent to an incremental catchment in the US NHDPlus [NHDPlus].

4.4.1.5.1 Bank Catchment

An elementary catchment consisting of land that drains to a section of a river represented geometrically as a polygon in 2D. It does not contain a waterbody, although it is adjacent to one. For example, if two streams drain into a lake, the remnant area between the catchments for the two streams also drains into the lake; it defines a bank catchment. Another notion of a bank catchment is that it is a zero order hydro feature, in contrast to a first order hydro feature such as a headwater stream.

4.4.1.5.2 Built-up Area

Urban or industrial area, or area under construction. Such areas where constructed features may involve significant infrastructure developments that include conduits for the transfer of freshwater or waste water. Within a built-up area, infrastructure flowpaths may cross one another in 2D and not intersect because of being at different elevations. Also, such flowpaths will generally not drain the immediate area through which they pass.

4.4.1.5.3 Empty Catchment

An elementary catchment consisting of internally drained land that does not touch a waterbody. It is the limiting case of a bank catchment around an isolated lake that shrinks to the point where it no longer exists, leaving only an isolated depression. In general, depressions defining non-channelized, internally drained areas are empty catchments.

4.4.1.5.4 Reach Catchment

An elementary catchment consisting of land that drains to a section of a river represented geometrically as a linear element. The river feature is contained in the catchment.

4.4.1.5.5 Water Catchment

An elementary catchment consisting entirely of a waterbody of a portion of a waterbody, where the geometry of the feature is a polygon in 2D. A single small lake may be geometrically equivalent to a water catchment. A larger lake or a river sufficiently large to have polygonal geometry may be broken into a series of areas, each defined as a water catchment. For example, a lake with a significant lake arm may be subdivided such that the lake arm is treated as a water catchment and the remainder of the lake is treated as a second water catchment. Similarly, a river may be subdivided at a hydrometric station or at the location of the beginning or end of a built-up area such as a city or industrial complex.

CHyF services also have the ability to subdivide a water catchment on the fly as appropriate. A user may be interested in what is upstream of an arbitrary point along a river, such as a fish sampling site or a water quality site. The service can create a boundary across the river at that point in real time, with a catchment resulting above the boundary and another below the boundary. Unless such boundaries have hydrologic significance, it is not recommended to make such subdivisions as part of the catchment fabric.

A water catchment is a constituent part of a catchment network, whereas a waterbody (with either a polygonal or linear representation) participates in a hydrographic network. Water catchments are represented directly in a hygraph data structure, whereas waterbodies are not.

4.4.1.6 Interior Catchment

A "... feature type that specializes the general ... [catchment] class as a catchment that is generally not connected to other catchments." [HY_F]. An interior catchment consists of one or more elementary catchments that collectively define an area with no flowpath connections on the surface to flowpaths outside of the catchment.

4.4.2 Depression

"Landform lower than the surrounding land and partially or completely closed that is able to but does not necessarily contain water." [HY_F]. Waterbodies exist within depressions. Some depressions though do not contain waterbodies, either because the waterbody is ephemeral and not present at the time of mapping or because the soil or substrate is so porous that a waterbody never forms.

4.4.2.1 Channel

"Natural or artificial waterway, clearly distinguished, which periodically or continuously contains moving water, or which forms a connecting link between two bodies of water." [HY_F]. Channels are depressions that form the containers for rivers, but water may be present in a channel only during heavy rains or large floods.

4.4.3 IceSnow

Perennial cover of either ice or snow [after Anderson].

4.4.3.1 Glacier

Accumulation of ice with an atmospheric origin which usually moves slowly on land over a long period [WMO].

4.4.3.2 Snowfield

"Perennial Snowfields are accumulations of snow and firn that did not entirely melt during previous summers. Snowfields can be quite extensive ... or can be quite isolated and localized …" [Anderson].

4.4.5 Waterbody

"Mass of water distinct from other masses of water." [HY_F]. The term refers to lakes, rivers and other watercourses of any size, which may be permanent or ephemeral.

The geometry of a waterbody may be one-dimensional, with 2D or 3D coordinates, or it may be two-dimensional with 2D or 3D coordinates, and in either case if 3D coordinates are used the third coordinate refers to the elevation of the water surface. Bathymetric data may be available separately.

As water levels vary as a function of seasonal changes, drought or flood conditions, or other temporal effects, the geometric representation of a waterbody and its specific position will vary as well. CHyF can be used to support temporal changes, as discussed in section 6.1.12. However, in the general case the geometry is considered to represent, or to be close to, the mean 2D surface extent for a given waterbody.

4.4.5.1 Canal

A body of surface water, participating in a hydrographic network, special due to its artificial origin (man-made) and its permanent or temporary flow. More specifically, *canal* refers to the water contained in an artificial waterway built for navigation or the transport of water. Ditches and drainage channels that typically contain water are also included [after HY_F, Oxford, Ramsar].

4.4.5.2 Estuary

"... a body of surface water, participating in a hydrographic network, made special due to branching and its interaction with the open sea." [HY_F]. Estuaries are characterized by tidal waters and are often associated with deltas.

4.4.5.3 Infrastructure Element

Water contained in a conduit or buried tanks, participating in a hydrographic network. This relates to water flow through a dam or industrial complex or as part of an urban, agricultural or water supply infrastructure. Storm drains and sanitary sewers fall into this class. All of these features are of note because they may cross one another in 2D yet not intersect and because they typically do not drain the area through which they pass.

4.4.5.4 Lake

A body of surface water, participating in a hydrographic network. It is special due to its considerable size and the lack of significant observable flow except at inflows and outflows. A lake may or may not be anthropogenic in origin and may or may not be regulated [after HY_F and Anderson]. It usually contains freshwater, but may also contain salt water, as with the Dead Sea, Lake Assal and the Great Salt Lake.

4.4.5.4.1 Great Lake

One of the Laurentian Great Lakes of North America or other very large lakes. They are conceptualized as a specialization of the general notion of a lake, with potentially specific associated methods and attributes that may for example be derived from ocean models.

4.4.5.4.2 Wastewater Pond

A pond or lagoon designed to contain wastewater for treatment [after EPA-2011].

4.4.5.5 Nearshore Zone

The zone extending from the edge of an ocean or a large lake or a large river or a large estuary, where the zone is defined as the limit of land to an arbitrary distance into the water that may be related by bathymetry, littoral characteristics including wave activity, coastal currents, or a buffer of a given width. A river emptying into the ocean or a large lake may be said to be emptying into the nearshore zone. In the context of an ocean, the following definition applies: “The zone extending seaward from the low water line well beyond the surf zone; it defines the area influenced by the nearshore or longshore currents. ...” [Coastal Wiki].

4.4.5.6 Ocean

A large body of saline water that composes much of the earth’s hydrosphere and that is not situated inland [After Wikipedia and Princeton]. Seas and bays that extend to the open ocean are classed as ocean. The Caribbean and Mediterranean Seas and the Bay of Bengal are considered as ocean, whereas the Caspian Sea in Asia is a lake.

4.4.5.7 River

“A body of surface water, participating in a hydrographic network; it is special due to its property of permanent or temporary flow.” [HY_F]. In common parlance, streams and rivers of any size fall under this class.

4.4.6 Wetland

“Wetlands are areas where water covers the soil, or is present either at or near the surface of the soil all year or for varying periods of time during the year, including during the growing season.” [EPA 2017]. Bogs, fens, marshes, shallow open water and swamps all fall into the wetland class. More generally, two kinds of wetlands are recognized in the EPA document: coastal/tidal wetlands and inland/non-tidal wetlands. However, with this version of CHyF, Wetland is not subdivided further. Most wetlands are permanent features, but some are ephemeral. Excluded from the wetland class are human-made wetlands [Ramsar], such as irrigated land, aquaculture ponds, farm ponds, water storage areas, salt exploitation sites, excavations, etc.

4.5 Hydro Location

“Any location of hydrologic significance located on a hydrologic network that is a hydrology-specific realization of a hydrologic nexus.” [HY_F]. The term “on” is interpreted to mean that a point representing the hydro location is on or near a flowpath.

In all cases a corresponding hydronode either already exists in the dataset or potentially could be defined, based on either projecting the geographic position of the hydro location onto a nearby flowpath, or additionally or alternatively, employing another assignment process to ensure that the chosen flowpath is the correct one.

CHyF recognizes a subset of the types of hydro locations that are specified in HY_Features. For practical reasons it also adds *arbitrary location, ice jam,* and *barrier*.

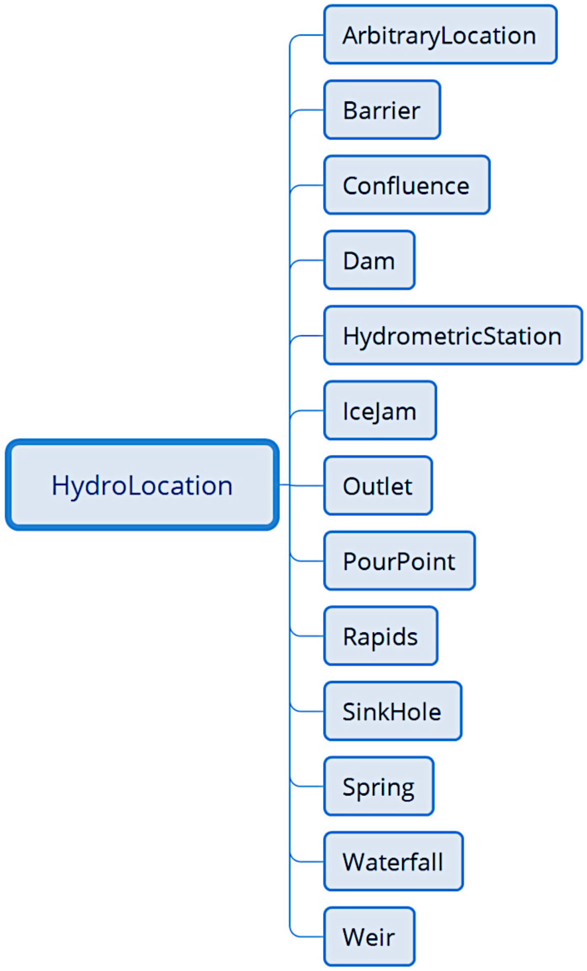

Figure 6: Hydro locations

4.5.1 Arbitrary Location

An arbitrary location is a hydro location of indeterminate type. It may be used to reference other phenomena not included in the other options described below.

4.5.2 Barrier

"Obstruction to the flow of surface water …" [WMO]. This is a general term that overlaps with some of the other hydro location types, including dam, rapids, waterfall, and weir.

4.5.3 Confluence

"Joining, or the place of junction, of two or more streams." [HY_F, WMO].

4.5.4 Dam

"Barrier constructed across a valley for impounding water or creating a reservoir." [HY_F, WMO].

4.5.5 Hydrometric Station

“Station at which data on water in rivers, lakes or reservoirs are obtained on one or more of the following elements: stage, streamflow, sediment transport and deposition, water temperature and other physical properties of water, characteristics of ice cover and chemical properties of water.” [HY_F, WMO]. Another term for a hydrometric station is a gauging station.

4.5.6 Ice Jam

“... an accumulation of ice in a river, stream or other flooding source that reduces the cross-sectional area available to carry the flow and increases the water-surface elevation.” [FEMA, similar to WMO]. A number of types of ice jams have been identified; however, this is beyond the current scope of CHyF.

4.5.7 Outlet

The farthest location downstream in a catchment, and geospatially coincident with a hydro nexus. The term is most meaningfully applied when associated with a drainage basin and its contained mainstem. In CHyF an outlet is always represented geometrically by a point.

4.5.8 Pour Point

“Specified catchment outlet defined to delineate a catchment upslope from that point.” [HY_F]. Pourpoint (or pour point) is not listed in the WMO glossary, but it is a very useful concept. CHyF provides a service that allows for various ways of specifying a pourpoint.

4.5.9 Rapids

“Reach of a stream where the flow is very swift and shooting, and where the surface is usually broken by obstructions, but has no actual waterfall or cascade.” [HY_F, WMO].

4.5.10 Sinkhole

“Place where water disappears underground in a limestone region. It generally implies water loss in a closed depression or blind valley.” [HY_F, WMO].

4.5.11 Spring

“Place where water flows naturally from a rock or soil onto land or into a body of surface water.” [HY_F, WMO].

4.5.12 Waterfall

“Vertical fall or the very steep descent of a stream of water.” [HY_F, WMO].

4.5.13 Weir

“Overflow structure which may be used for controlling upstream water level or for measuring discharge or for both.” [HY_F, WMO].

4.6 Hydro Network

In HY_Features the hydro network type realizes a catchment as a network of connected hydrologic features. The term network implies: (i), that the set of connections can be modelled as nodes connected by edges, and (ii), that the features can all be connected through the network. Where features cannot all be connected, then two or more networks exist; as an implementation detail, they may all be included within the same mathematical graph.

A number of different kinds of networks are recognized, as noted below. Since a hydro network is a network of hydrologic features, catchment network, channel network, and hydrographic network are considered as types of hydro networks. Flowpath network and catchment divide network are realizations of a catchment; however, they are first and foremost networks, which is why they are included here in this logic diagram.

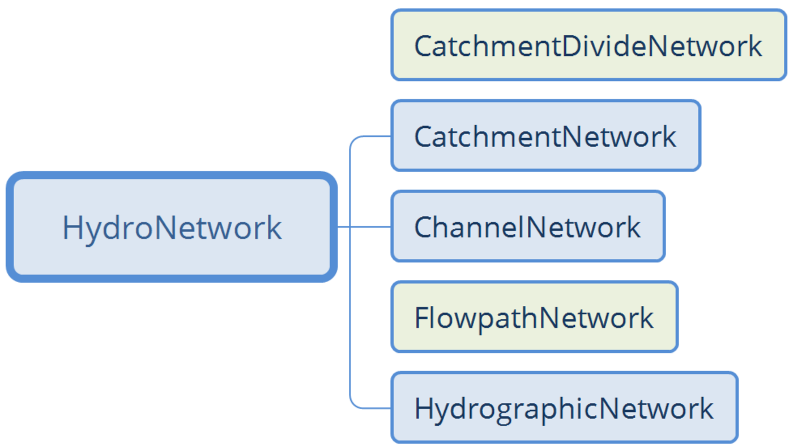

Figure 7: Hydro Networks

4.6.1 Catchment Divide Network

A catchment divide network consists of a set of catchment divides connected to one another at vertices. The resulting mesh implicitly includes the bounding segments of individual catchments within a larger catchment.

4.6.2 Catchment Network

A set of catchments connected through hydro nexuses [after HY_F]. Each catchment has one or more neighbouring upstream catchments and one or more neighbouring downstream catchments, with the exception of the highest and lowest catchments in the network.

4.6.3 Channel Network

“Connected set of depressions and channels that continuously or periodically contain water.” [HY_F]. Practically, a channel network is a superset of a hydrographic network, (i) since lakes and rivers can be considered as being contained by depressions and channels, and (ii) since other channels exist that are not considered as lakes or rivers.

4.6.4 Flowpath Network

The set of all flowpaths forming a connected network within a catchment and realizing that catchment. If interior catchments are included in a larger catchment, then the contained flowpaths may form more than one network. The network or networks contain no cycles whereby flow can move through flowpaths and return to its starting point. It needs to be acyclic and directed downstream throughout. Divergences are allowed, such as with flows around islands, flows through braided channels, and flows on deltas. However, if only primary flowpaths are considered through a network, then the network is dendritic in nature.

4.6.5 Hydrographic Network

"Aggregate of rivers and other permanent or temporary watercourses, and also lakes and reservoirs, over any given area." [WMO]. The WMO is treating reservoirs as a water feature in this definition, whereas HY_Features and CHyF do not. Nevertheless, the intent is clear. Using the terminology in this document (and consistent with HY_Features), a hydrographic network consists of a directed, acyclic network formed by linear and polygonal waterbody features.

4.7 Hydronode

A hydronode is a network construct equivalent to a node or a vertex, and existing at each endpoint of a flowpath.

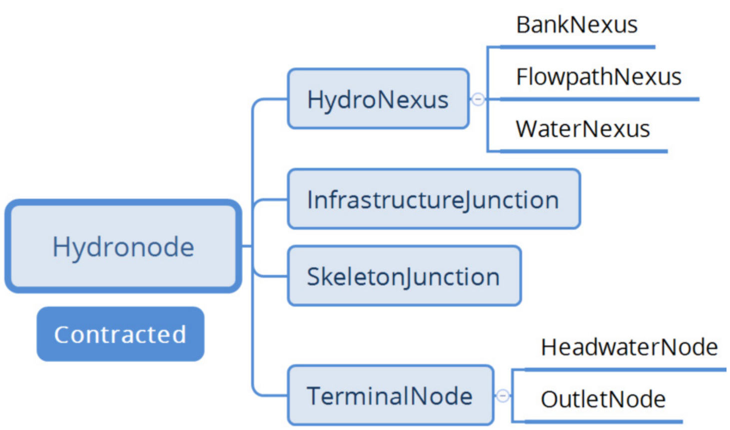

Figure 8: Hydronodes

A hydronode that is treated as a permanent feature is said to be *contracted*. It is a reference location defined to support interoperability [after AHGFNode in AHGF]. Any type of hydronode potentially could be contracted, but the most likely candidates are outlet nodes. Typically these would be outlets, represented as points, of drainage basins or major drainage areas. However, to support interoperability and connectivity directly, water nexuses, that by design are coincident with gauging stations and large dams, could be treated as contracted. Note that "contracted" as used here has nothing to do with the notion of a contraction hierarchy as used in graph implementations in computer science.

4.7.1 Hydro Nexus

A "... hydro nexus represents the place where a catchment interacts with another catchment, i.e. where the outflow of a contributing catchment becomes inflow into a receiving catchment." [HY_F]. A hydro nexus exists where adjacent flowpaths touch in a flowpath network and where that location acts as the interface between two or more catchments.

4.7.1.1 Bank Nexus

A bank nexus is a point representing the interface between a bank catchment and a water catchment. Because of its relationship to the two catchments, it is considered as a type of hydro nexus and not a hydro end node. A bank nexus is equivalent to the upstream endpoint of a bank flowpath.

4.7.1.2 Flowpath Nexus

A flowpath nexus is a point representing the interface between a reach catchment and either another reach catchment or a water catchment.

4.7.1.3 Water Nexus

A water nexus is a point representing the interface between two water catchments.

4.7.2 Infrastructure Junction

The location where two infrastructure elements, such as buried conduits, join. It is typically represented as a point. Because these elements do not drain the region through which they pass, infrastructure junctions are not equivalent to hydro nexuses.

4.7.3 Skeleton Junction

A skeleton junction (or simply junction) is a confluence point on a flowpath network that does not represent the interface between catchments. It occurs in lakes, rivers, and estuaries as a means of establishing network connectivity. It is always coincident with an endpoint of one or more skeleton flowpaths and it may be coincident with the downstream endpoint of a bank flowpath.

4.7.4 Terminal Node

A terminal node exists at the start points and endpoints of a flowpath network. These are referred to respectively as headwater nodes (or points) and outlet nodes (or points). A bank nexus is considered as equivalent to a flowpath nexus, connecting one elementary catchment to another, and thus is not considered as a terminal node.

4.7.4.1 Headwater Node

A hydronode used to specify the start of a flowpath, where there does not exist any inflowing flowpath. It corresponds to the start of a headwater (first order) stream. A flowpath exists downstream of the headwater point, but none is defined upstream of it. A headwater node usually has a valence of 1.

A headwater node can also be defined in the context of the boundary of the area of interest. If water flows into the AOI, then a headwater node exists on the boundary at the vertex which is also the upstream end of the flowpath extending into the AOI from that point.

4.7.4.2 Outlet Node

A hydronode used to specify the end of a flowpath that is not connected to any further downstream flowpath. It acts as the end of a flowpath network. An outlet node usually has a valence of 1.

A sink may exist within an area of interest. That sink is then treated as an outlet node. If an outflowing flowpath touches the boundary of the AOI, then an outlet node exists at the vertex representing where they intersect.

4.8 Reservoir

“... a concept of water storage … [allowing] any waterbody type to be considered a managed reservoir.” [HY_F]. Thus a lake or a section of a large river may act as a reservoir, a term that describes the use of the waterbody, as opposed to a type of waterbody. A dam or some other type of water control structure usually exists on the boundary of a reservoir.

Figure 9: Reservoir

4.9 Shoreline

The boundary segment or series of segments between land and water, where the water represents an ocean, a sea, or a very large lake. It often refers to the operational Mean High Water (MHW) elevation contour along an ocean [after Wikipedia and Weber, et al, 2005].

If estuaries or nearshore zones are present, then the shoreline is implemented as the edge of the open ocean, which is acting as the ultimate sink within the area of interest. The segments forming the shoreline must be continuous. In a typical situation where a river enters an ocean, the boundary between the two bodies of water also constitutes a shoreline segment.

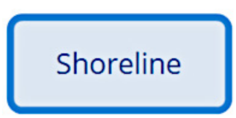

Figure 10: Shoreline

Chapter 5. Conventions

This section provides details that will be helpful in understanding this document.

5.1 URI References

The primary location for CHyF specifications is given by the following URL:
https://github.com/NRCan/chyf-pilot/tree/master/docs

The code repository for all CHyF software is on GitHub here:
https://github.com/cmhodgson/chyf-pilot

5.2 Abbreviations

DEM:	Digital Elevation Model (defined by a grid or a point cloud)
GWML2:	GroundwaterML 2
GPS:	Global Positioning System
HY-F:	HY_Features
HY_Features:	OGC® WaterML 2: Part 3 - Surface Hydrology Features (HY_Features) - Conceptual Model
ISO:	International Organization for Standardization
OGC:	Open Geospatial Consortium
UML:	Unified Modeling Language
WGS:	World Geodetic System
WMO:	World Meteorological Organization

5.3 Formal Model Notation

The modelling diagrams make use of UML, with class names specified in UpperCamelCase and property names in lowerCamelCase.

5.4 HY_Features and WMO Terminology

CHyF makes use of terminology from HY_Features and WMO as much as possible. In some cases the meaning is identical. In others it is sufficiently similar so that the same term is used. In some cases small differences may exist with some class properties.

CHyF also introduces a number of classes required for implementation, while ignoring others that are not required or where the CHyF and HY_Features models are intentionally divergent. For example, HY_Features supports linear referencing for those who wish to use it. CHyF intentionally does not support linear referencing, but instead is directly compatible with direct referencing through WGS coordinates (often referred to as GPS coordinates) and CHyF web services.

Like HY_Features, CHyF avoids the use of many common terms, such as *watershed* and *basin*. With some terms HY_Features describes the respective concepts, but does not include them in its model, whereas in CHyF the terms are included formally; *catchment coverage* and *mainstem* are examples.

As an implementation model, CHyF specifies the geometric representation of features. In all cases this is a necessary requirement for implementation. In some cases though, it is done to be able to take advantage of fast computation or navigation. A nexus in CHyF is always represented as a point, to simplify compatibility between geospatial representation and graph theoretic algorithms. In CHyF a dam may be located at a hydro nexus, whereas in HY_Features the dam might be defined as a hydro location of type dam, without reference to geometric representation.

5.5 Naming Conventions

As with HY_Features, formal class names are based on English. These class names make use of the prefix CA_, in line with OGC and ISO naming conventions. The relationship to HY_Features model elements, with the prefix HY_, is described where applicable. CA_ indicates common applications, for which CHyF is designed to be applicable.

5.6 Language Support

HY_Features recognizes the need to support feature names in multiple languages. CHyF takes a different approach. CHyF makes use of an attribute called *name_string,* if available. This attribute is of data type *Character*. It may be populated by a natural language name, by an arbitrary key or integer value, by a UUID, or by any other character string.

For example, the *Pouce Coupe River* flows in British Columbia and Alberta. In French it is called the *rivière Pouce Coupé*. Either of these strings could be used, but so could the Geographical Names in Canada key, *JDBYJ,* or the corresponding UUID *4a712959ba3611d892e2080020a0f4c9.* CHyF software only cares that the name_string is consistent. If keys are used, then all sections of the river should have the value *JDBYJ* for the name_string attribute.

If the intent is to support a human viewable display, then map labels and associated table values should be in French or English, at the user's discretion. CHyF does not support this directly. It would be necessary for the developer to make use of a look-up table in the background, whereby the data has the value *JDBYJ*, but the user sees *rivière Pouce Coupé* or *Pouce Coupe River*, depending upon her or his preferred language.

Chapter 6. The Logical Model

CHyF provides a practical implementation of the conceptual model presented in HY_Features. Both models define hydrologic features and different kinds of networks based on these features. Collectively, these concepts define a domain. That domain centres on a geospatial representation of the surface flow of water through river systems and through the drainage areas on the landscapes that contain them. A core concept in CHyF, which extends HY_Features, is that these features and their relationships can be represented in a mathematical graph, referred to as a hygraph. Features in CHyF are defined to make the benefits of the hygraph particularly useful to practicing hydrologists, environmental specialists, and water managers.

6.1 CHyF Overview and the CHyF Hydro Fabric

HY_Features establishes the concepts of catchments, flowpaths, and hydro nexuses, and CHyF extends each of these for operational purposes. In order to leverage graph theory, CHyF defines basic versions of these features that then become the objects (i.e., nodes) in the hygraph. These elemental objects are referred to as elementary catchments, elementary flowpaths, and hydronodes. An elementary catchment or an elementary flowpath can be subdivided, at a gauging station or barrier for example; this results in smaller elementary features (similar to cell division) above and below the point of interest. Thus the basic idea of a graph composed of the connections between elementary features does not change.

HY_Features does not indicate what a given catchment can contain. The NHDPlus allows reach catchments (the smallest catchments in its system) to include arbitrary amounts of land and water in the vicinity of polygonal waterbodies. By contrast, CHyF does not. Precipitation falling on a waterbody or on land is very different hydrologically, as are the processes within surface waters compared to within terrestrial surficial materials. Consequently with CHyF, catchments are either land-based areas or water-based areas, but never a mixture of the two. Just as with terrestrial catchments, polygonal waterbodies can be subdivided; this allows for example for the creation of river sections, each of which is treated as an elementary catchment.

The central concept in HY_Features is that of a catchment as an integral component of a hydrologic features fabric. Flowpaths and catchment divides (watershed boundaries) are modelled as realizations of catchments. As a profile of HY_Features, CHyF makes use of the same modelling constructs, but as the basis for an implementation model involving both data and services, its central concept extends the hydro fabric notion. CHyF emphasizes a series of intertwined, topologically harmonious networks that can all be related through graph theoretic constructs, and that collectively form a hydro fabric represented by a CHyF dataset.

This hydro fabric serves as a terrain representation that can be used as the common space on which entities of interest are located and related, ranging from rivers to gauging stations, to floodplains, to agricultural and urban developments. The starting point is often a hydrographic network formed from waterbody features. In Figure 11, an area is shown in the Richelieu Valley in southern Quebec. The hydrographic network includes the cyan waterbodies and the blue streams, and excludes the skeletons in the lake (Lac Hertel) and the river (Rivière Richelieu). The blue lines, including those shown as in the polygonal waterbodies, form a flowpath network.

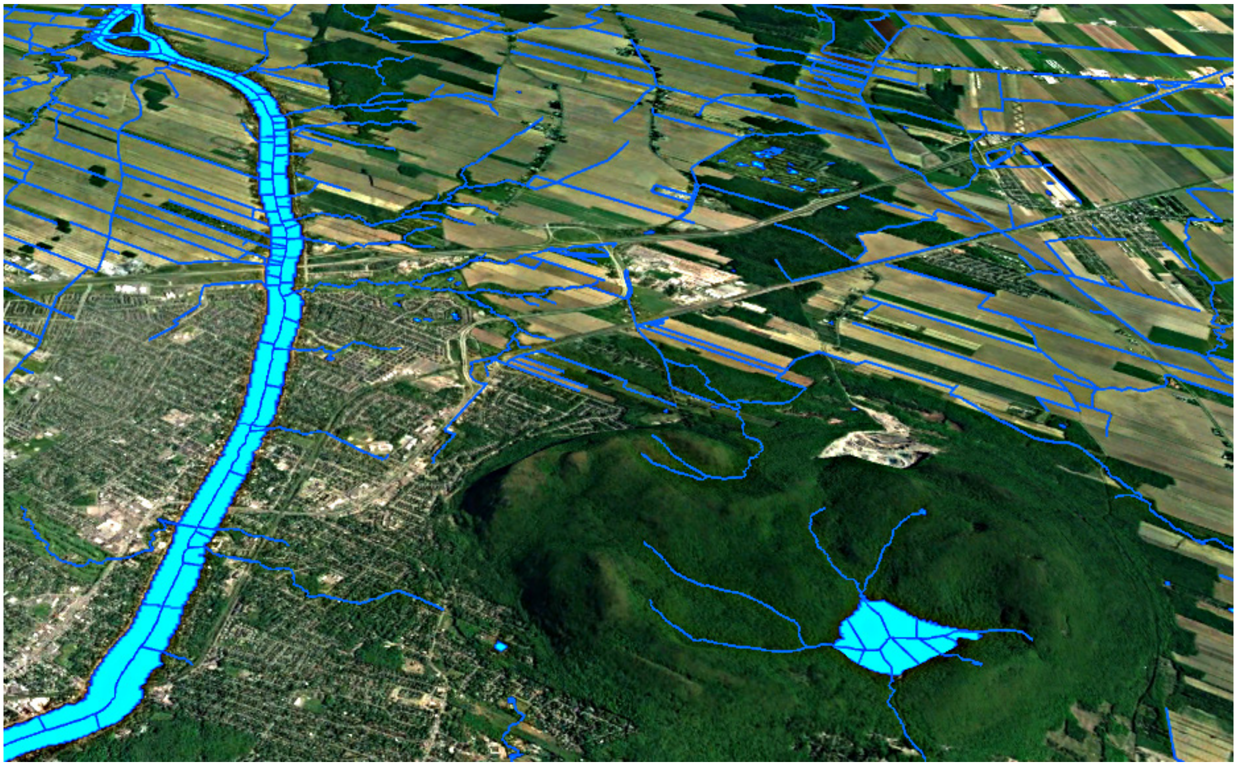

Figure 11: A flowpath network (rendered as blue lines) in the Richelieu Valley

From the hydrographic and flowpath networks, and a DEM if available, a catchment coverage can be created that in most respects is the dual of the flowpath network, as shown in Figure 12. The catchments form an associated network, as do the bounding linestrings that define the catchment divides. The hydronodes (points) at the beginning and end of every flowpath segment form yet another network. If channels were extracted from depressions in a DEM, then those channels in combination with the hydrographic rivers will form a channel network that ideally will be consistent with the other networks.

Figure 12: A catchment coverage (with boundaries rendered in red) extending across all of the terrain

The various networks are all composed of network elements. Network elements of the same type have relationships, as do those of different types. These relationships form the basis of navigating through the land and waterscape. The power of the CHyF model emanates from the interwoven topologies of these networks. If the geometry of the underlying features is well defined topologically, the hygraph can be quickly generated, and fast traversal through it becomes possible. This enables a number of web services to return hydrologic features data meeting specific connectivity criteria. It also allows for a simpler model in comparison to some that are currently in common use, as functions that are often embedded in data through coding systems can be generated through fast computation instead.

6.2 Classes and Relationships

In the subsections that follow, a series of UML diagrams are provided. Within each diagram the classes that correspond to a class in HY_Feature are depicted with a black boundary. A red boundary means that the class represents a newly introduced construct. For example, in the diagram under Hydrologic Features below, CA_Catchment is considered as equivalent to HY_Catchment from HY_Features, whereas *CA_ElementaryCatchment*, CA_Wetland, *CA_IceSnow*, CA_Snowfield, and CA_Glacier are all introduced.

Another convention used here concerns the text style. If it is in italics, the class is an abstract superclass, which means that it is never instantiated directly; instead, instances may be defined by its subclasses. *CA_ElementaryCatchment, CA_Waterbody* and *CA_IceSnow* are all examples of this. So CA_Glacier and CA_Snowfield may be instantiated directly, but *CA_IceSnow* cannot.

The UML diagrams emphasize class hierarchies and general relationships. Details about constraints are not included in the UML; instead they are enunciated as topological rules in a subsequent section. Additional geographic figures are provided to help convey details not provided previously. The same graphical conventions are used here as well.

6.2.1 Hydrologic Features

Hydrologic (or hydro) features consist of water features (waterbodies), containers for water (depressions), and specific areas of the terrain conceptualized as draining into flowpaths (catchments), as shown in figure 13 below. Also included are wetlands, glaciers and snowfields.

Although these concepts are all distinct, overlaps exist geographically. These details are not included in the UML. A depression may or may not contain a waterbody. A depression in its entirety or in part may also be considered a catchment. Catchments may be entirely terrestrially based, they may contain a stream, or they may correspond to polygonal waterbodies or portions of such waterbodies.

In general a hydrologic feature flows into another hydrologic feature through a hydro nexus. A catchment flows into another catchment, a waterbody flows into another waterbody, and a depression (including the notion of channel) may be connected to another depression. However, these flows are not constrained to the same type of hydro feature. For example, a channel may flow into a river, which may flow through a catchment. Many other combinations are also possible.

As noted in Section 5.6, CHyF does not explicitly support naming with a specific data structure; however, it does include a *name_string* attribute which can be any arbitrary character string, such as a common name, an official gazetteer key, or a UUID. Official names in one or more languages can be cross-referenced to the value for name_string, but that is outside of the scope for CHyF.

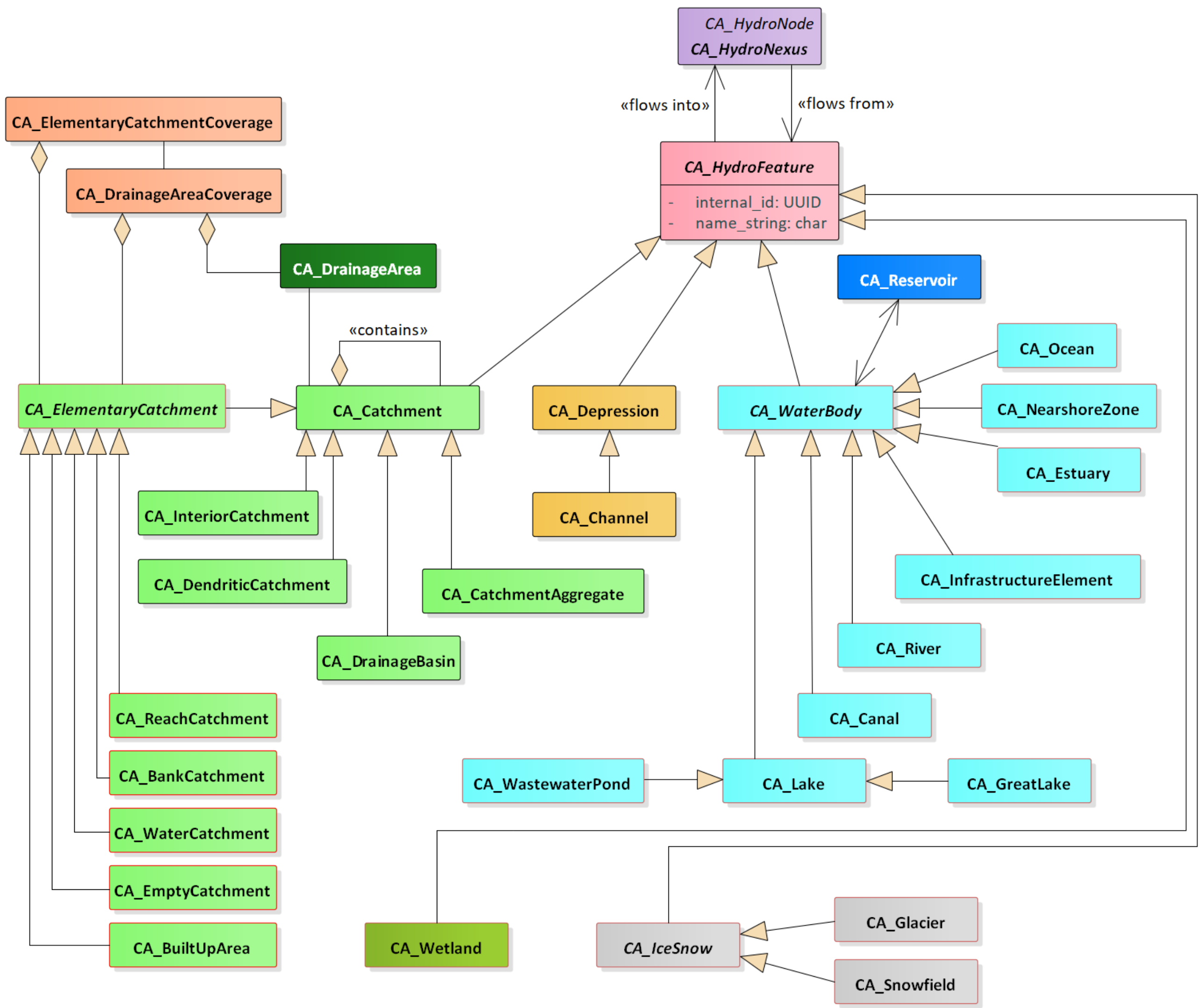

Figure 13: Hydrologic features and their relationships

6.2.1.1 Waterbodies and Depressions

The term waterbody applies to any body of water, whether it is actively moving or largely static. Thus rivers, lakes and estuaries are considered as types of waterbodies, as are subsurface contained flows, near shore zones, and oceans. In common parlance the terms rivers, streams, and creeks are frequently used. In HY_Features and CHyF they are considered together under the class HY_River and CA_River, respectively. Similarly, ponds are grouped with lakes under HY_Lake and CA_Lake. The class *infrastructure element* is included in CHyF so that connectivity between surface features and within urban landscapes can be modelled effectively. The most significant feature in this class is storm drain, although controlled flows through large dams are also of major interest.

Great lakes are large lakes and include the Great Lakes of North America and other lakes of notable size. The class is recognized in CHyF as a specialization of lake, since different kinds of hydrologic models may be applied to them, compared to more typically sized lakes. Wastewater ponds are included as a type of lake in the CHyF model because of their constraints and because they are often mapped separately from lakes. Nearshore zones represent the zones along the shorelines of estuaries, oceans, large lakes, and wide rivers. This class is introduced into CHyF because it helps resolve problems with how to include large waterbodies in hydrologic networks. Reservoirs are not considered as a type of hydrologic feature. Instead, a reservoir is treated as a control structure to which a hydro feature has a relationship (see Section 4.8).

Waterbodies sit in depressions in the land, but are recognized as constituting a separate class. A channel is a subclass of depression that may contain a river. Unconnected features may be generated through extraction efforts using lidar and other earth observation data. With CHyF they can remain unconnected, or alternatively, connections can be made to one another and to nearby waterbodies. In either case a relationship exists between a river and the channel in which it exists; this is not shown in Figure 13.

6.2.1.2 Glaciers, Snowfields and Wetlands

Other types of hydro features are included in CHyF because of their general importance. More than two-thirds of the world’s freshwater is estimated to be held in ice and snow [USGS]. Glaciers and snowfields represent important sources of water in many areas of the world and are also of consequence in the context of climate change. CHyF includes them as hydro features. Also included are wetlands, which hold considerable water globally and are of high significance to floodwater retention, biodiversity, aquaculture, carbon cycling and climate change.

6.2.1.3 Catchments

Precipitation falls indiscriminately on land and water. Like NHDPlus, CHyF explicitly states that it always falls on a catchment, an area that can be modelled as draining to a common outlet point. A catchment may be of continental scale, such as the catchments for the Saint Lawrence River or the Mississippi River, or may be very small, such as those for unnamed creeks. At the most detailed level, CHyF recognizes an elementary catchment. As described elsewhere, in CHyF each elementary catchment may contain an area of land or an area of water, but not both areas of land and water. This is in contrast to the NHDPlus, which does not recognize water catchments and consequently includes arbitrary sections of lakes, wide rivers and estuaries in the delineation of catchments along shorelines. HY_Features does not give guidance as to what may be contained in a given catchment, so either approach can be said to be compliant with HY_Features’s conceptual model.

Elementary catchments in CHyF form a coverage such that they completely cover the land and waterscape with no gaps and no overlaps. This catchment coverage is an intrinsic part of the CHyF model. HY_Features discusses the general idea of a coverage composed of catchments, although it does not include it within the normative part of the document.

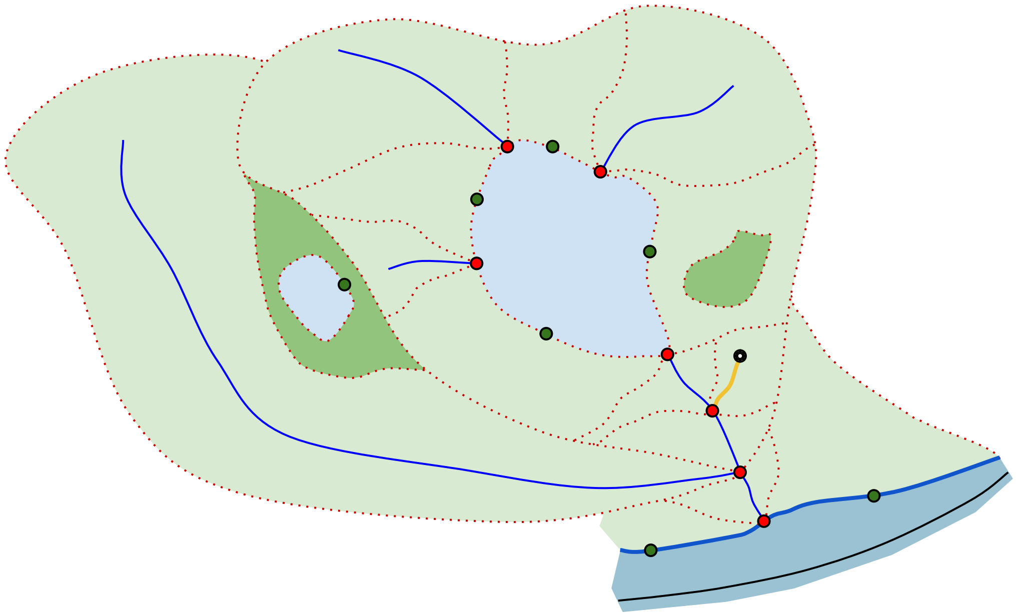

Figure 14: Different kinds of elementary catchments

CHyF recognizes five kinds of elementary catchments: reach catchments, bank catchments, water catchments, empty catchments and built-up areas. In figure 14 above, A, B, C, D, J, K, W and M are reach catchments. They each contain a single-line river segment. The elementary catchments E, F, G, H, and Y are referred to as bank catchments, since they drain directly into adjacent waterbodies with areal extent, which in these cases, are neighbouring lakes. L and N are also bank catchments, draining into the ocean. The two lakes, I and Z, and the ocean, O, are elementary catchments that are referred to as water catchments. Elementary catchment Q is an empty catchment, as it does not contain nor is it adjacent to any waterbodies.

One property of a catchment is that it may contain other catchments that in turn may include still other catchments. This recursive nature of catchments stops at elementary catchments. However, the container structure is not necessarily completely hierarchical. A large catchment may for example be broken down into an upper catchment and a lower catchment. As another example, an interior catchment, such as Y-Z or Q in Figure 14, may be included with adjacent catchments, even though a flow relationship between them does not explicitly exist.

An interior catchment may be very small, as with a depression in the terrain containing no water (an empty catchment). Alternatively, an interior catchment may contain a complex network of smaller catchments that ultimately do not drain to the ocean. The Great Basin in the United States is an excellent example of a very large interior catchment. An interior catchment may include other interior catchments. Consider the simple case of two nearby isolated lakes. The ring of land around each lake in combination with the lake forms an interior catchment, and the two resulting adjacent interior catchments can be considered together to form a single interior catchment. This is shown in the map snippet below from the Richelieu Valley in Quebec, where five interior catchments combine to form two, larger interior catchments.

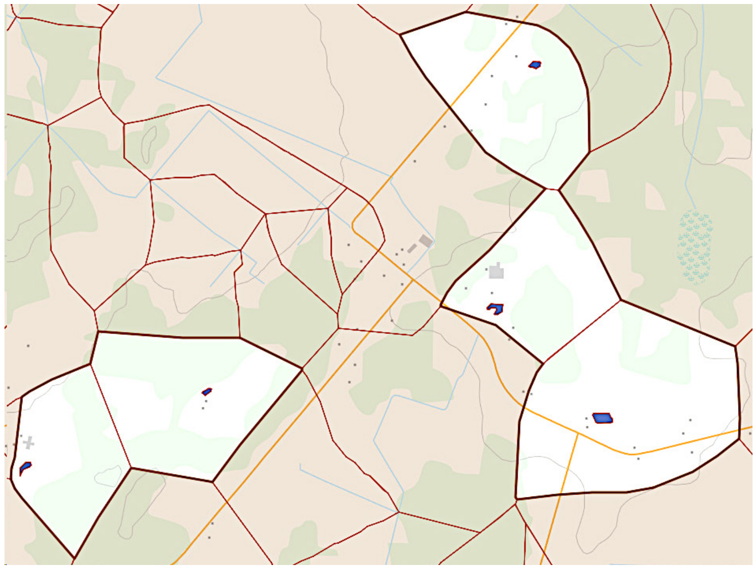

Figure 15: Interior catchments containing interior catchments containing interior catchments (isolated lakes)

Interior catchments are referred to as endorheic, as opposed to exorheic catchments, which drain externally. In many locales, small lakes may exist with no visible outlets; they and their immediate drainage areas constitute interior catchments. They can be allocated to larger, adjacent catchments if desired, based on proximity of the interior waterbodies to outside waterbodies that are part of exorheic catchments. So for example in the example above, depending upon the specific location of nearby waterbodies (not shown), all five interior catchments may be assigned to the same major, externally drained catchment, or alternatively, some could be assigned to one such catchment and others to another, depending entirely on proximity.

CHyF recognizes dendritic catchments as a class, CA_DendriticCatchment, as is also true with HY_Features, using HY_DendriticCatchment. A catchment is considered dendritic if it contains a corresponding flowpath network that is dendritic. Because secondary flowpaths are common, through braided stream systems, around islands, and on deltas for example, such a pattern may still be considered as dendritic so long as flowpaths are demarcated as either primary or secondary. In CHyF, dendritic is then taken as meaning that the primary flowpaths are identified and that they form a tree-like structure. Dendritic catchments are important because of their relationship to catchment aggregates (see below) and because they can help establish interoperability with data adhering to other models.

HY_Features introduces the very useful concept of a catchment aggregate, an aggregation of one or more dendritic catchments with associated interior catchments. CHyF has the class CA_CatchmentAggregate, defined as an abstract superclass. The intent is to treat large catchments in hierarchical systems as catchment aggregates. The areas defined by the Water Survey of Canada Sub-Sub-Drainage network or the HUC12 watersheds in the US NHDPlus are practical examples of catchment aggregates. More generally, this supports communication about for example the catchments defining the lower Fraser River or the upper Colorado River.

6.2.2 Catchment Realizations

HY_Features introduces the notion that catchments can be realized by other features, with catchment boundaries and flowpaths being examples that are included in CHyF. This has the interesting effect that, even though catchments are hydro features, their boundaries and associated flowpaths are not. This dichotomy is conceptually useful as divides and flowpaths exist only as a function of the existence of catchments; as well, they can both be algorithmically defined based on general notions of flow in and through catchments.

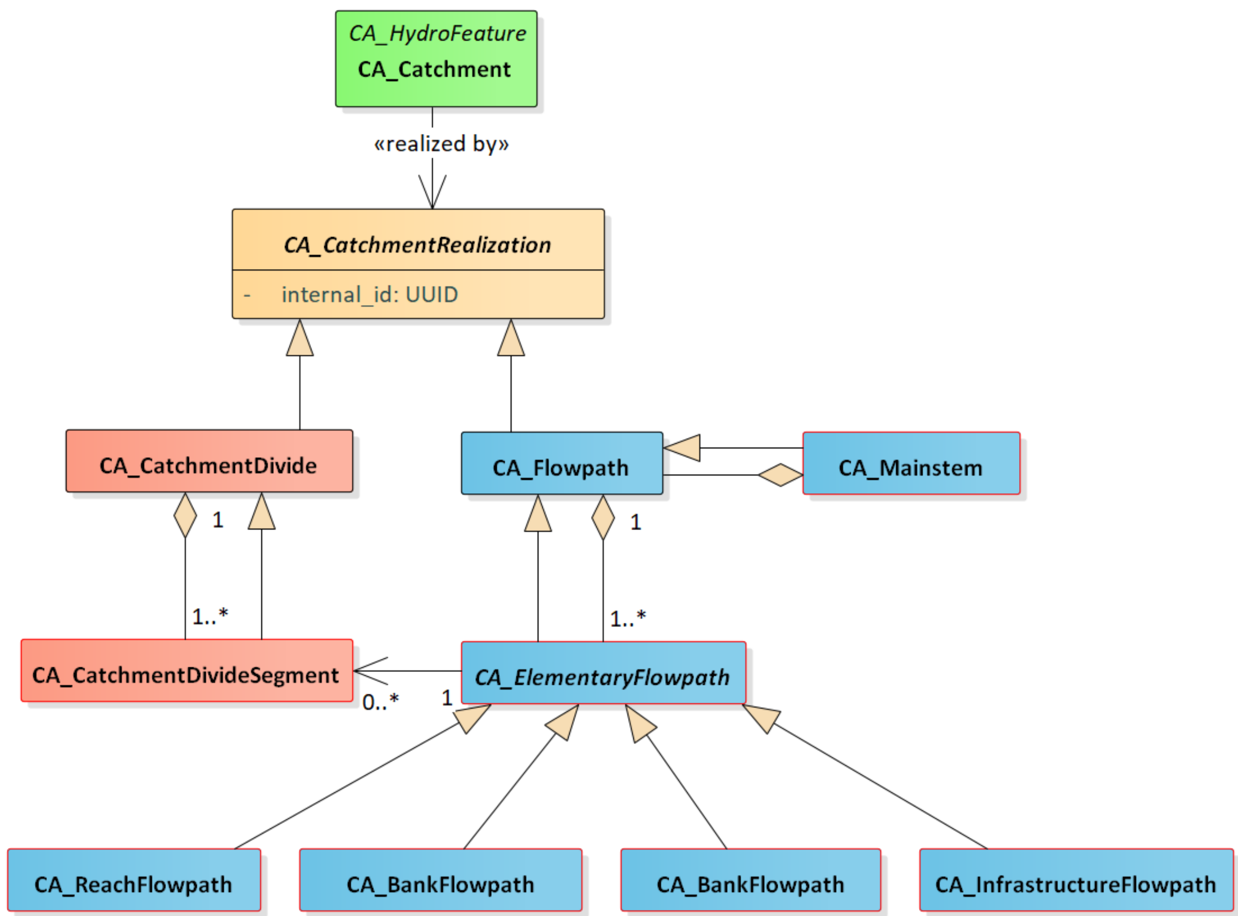

Figure 16: Catchment realizations and their relationships

6.2.2.1 Catchment Divides and Catchment Divide Segments

Figure 17 below is equivalent to a subset of figure 14 above, but with only the catchments and their boundaries showing. The endpoints of the bounding segments are indicated with an **x**. These bounding segments in CHyF are referred to as catchment divide segments that can be assembled to form catchment divides, which also exist in HY_Features. A catchment divide is any arbitrary path through the catchment divide segments, including but not restricted to the complete boundary of any given catchment. The catchment divide segments are comparable to elementary flowpaths, with particular paths through them of hydrologic significance.

Two approaches exist that can be used to generate catchment divide segments. The first involves the generation of boundaries approximating the medial axis between waterbody features, ignoring elevation data. This places the divides halfway between nearby waterbodies that may have either linear or polygonal geometries. This is appropriate on flat terrain where the elevation surface may not have sufficient resolution to indicate accurately the direction of overland flow. It may also be used where suitable elevation data is not available. The second approach takes elevation into account, such that flow direction and thus the position of the divide is consistent with downhill flow over the surface. This is required in areas of noticeable relief. The suite of CHyF Preprocessing Tools contains a catchment delineation routine that simultaneously takes both approaches into account. It should be noted however that other techniques are also possible.

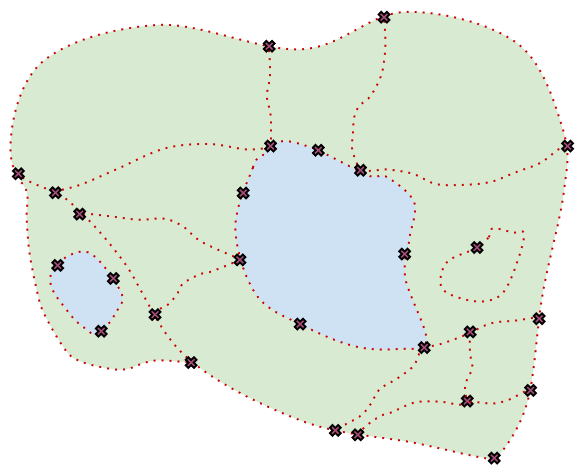

Figure 17: Catchment divide segments; these define catchment divides that form catchment boundaries

6.2.2.2 Flowpaths and Elementary Flowpaths

A flowpath is a theoretical construct representing the flow of water on the terrain. Geometrically, a flowpath is either a single elementary flowpath or a sequence of elementary flowpaths, all with a consistent direction. An elementary flowpath is a section of a flowpath bounded by a hydronode (e.g., a confluence) at either end and with no intermediary hydronodes. The elementary flowpaths form a Directed Acyclic Graph; consequently, no loops exist where water flows from an elementary flowpath and ultimately returns to it. In the case of a single-line river, the geometry of an elementary flowpath is equivalent to a segment of the river. In the case of any waterbodies with polygonal geometry, elementary flowpaths exists as part of a flowpath skeleton in the waterbody. These skeletal elements do not contribute to the delineation of catchments, so the number and exact locations of such elements are generally not important.

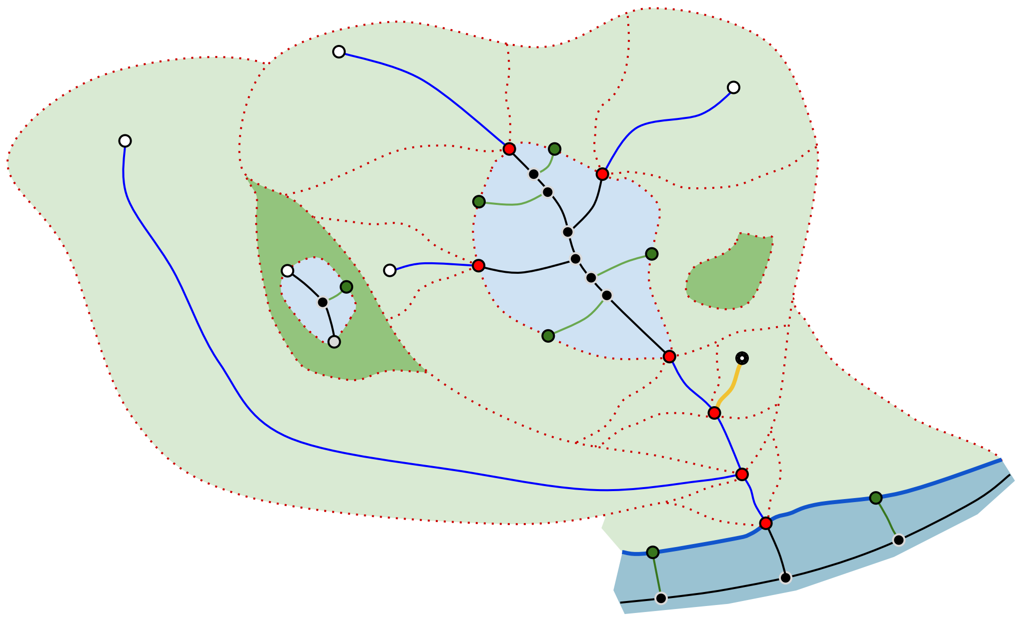

Figure 18: Elementary flowpaths and hydronodes

CHyF specifies four classes of elementary flowpath. Reach flowpaths are geometrically equivalent to single-line rivers segments; two subtypes are recognized, observed and inferred. Observed flowpaths (in blue in Figure 18) correspond to observed, single-line rivers. Inferred flowpaths (in orange) are assumed to be present in the given location as indicated by the presence of riparian vegetation for example or by other indicators. The sole example (flowpath d) connects a spring to the nearby river, following the path of steepest descent over the elevation surface. Bank flowpaths (in green) are elementary flowpaths that connect a bank catchment (without a contained stream) to the flowpath skeleton in an adjacent waterbody. They are of note topologically because each intersects its corresponding catchment only at its starting point. Skeleton flowpaths (in black) are elementary flowpaths added to create connectivity through polygonal waterbodies. Their respective positions are somewhat arbitrary, as is even true of how they are connected within the respective waterbodies. Not shown are infrastructure flowpaths. These represent contained flows through buried and/or constructed pipes, conduits, drains, sewers and other infrastructure elements through which water moves. Adjacent terrain generally does not drain into them and they may cross in 2D without intersecting, both factors making them different from the other types of flowpaths.

6.2.3 Hydronodes and Hydro Nexuses

CHyF recognizes different kinds of hydronodes in a three level hierarchy (figure 19 below). A hydronode may be a hydro nexus, acting as an interface between hydro features, a terminal node, indicating the start or end of a network, or either a skeleton junction in a polygonal waterbody or an infrastructure junction typically in a buried conduit complex.

A terminal node has an attribute, onAOIBoundary, that has a value of True if the hydronode is on the boundary of the area of interest, i.e., the area under consideration. In the more typical case, onAOIBoundary has a value of False, which is the default.

Two types of terminal nodes exist, headwater nodes and outlet nodes, which are rendered as white dots and gray dots respectively in the previous figures. Outlet nodes are found in two distinct situations. The first case is shown in figure 19, where the endpoint of a flowpath represents the end of a flowpath network in an interior catchment. In the second case, the outlet node lies on the boundary of the area under consideration, as described above. A headwater node may correspond to a spring, or more generally, the starting (upstream) point on a single-line, headwater stream. However, it may also exist on a boundary, like an outlet node.

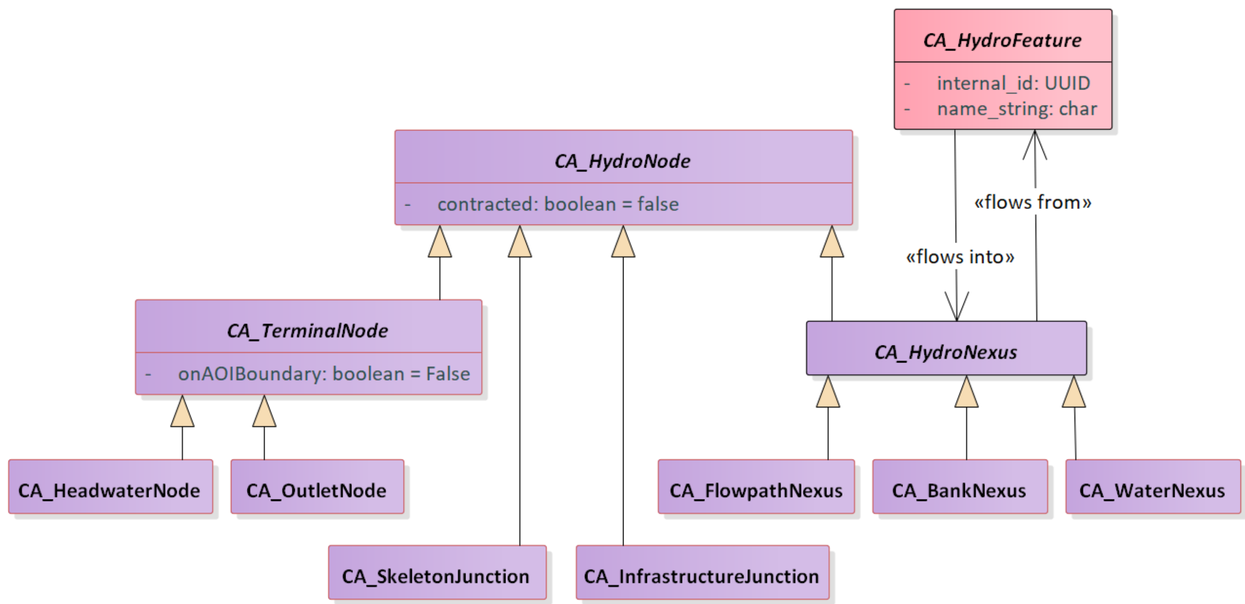

Figure 19: Hydronode hierarchy

A skeleton junction connects skeleton flowpaths to one another and to bank flowpaths, and consequently it always has a valence greater than one and it always is within the water away from the land. An infrastructure junction exists where infrastructure elements meet one another. However, where a buried conduit for example meets a surface stream, the point of intersection is a flowpath nexus.

A hydro nexus acts as an interface between hydro features of the same or different types. Three classes of hydro nexus are recognized. A flowpath nexus always serves as the outflow of a reach catchment and the inflow of the immediate downstream catchment. A bank nexus is a proxy for the shoreline that separates a bank catchment from the adjacent water catchment; it is represented as a point located on the land-water boundary and it serves as the upstream endpoint of a bank flowpath. A water nexus exists on the boundary of two water catchments. If a river or nearshore zone, for example, is subdivided into multiple catchments, or if a boundary is placed in the water between a lake and a wide (double line) river, then a water nexus exists along that bounding segment. A skeleton flowpath flows into a given water nexus, and another skeleton flowpath flows out of it.

6.2.4 Hydro Locations

Hydro locations are located along flowpaths and are represented as points that can be projected onto a flowpath. The projection may be to a flowpath endpoint, a nexus, or to an arbitrary point along the flowpath. In the latter case, the nearest point along the nearest elementary flowpath would be the default, although the projection could follow the path of steepest descent until intersecting with a flowpath.

A set of hydro location types has been defined, as shown in the enumerated list on the right side of figure 20 below. With the exception of arbitrary location, the types all come from HY_Features, and with the second exception of pour point, also from the International Glossary of Hydrology. These ten terms have been retained, whereas others have been dropped because of either redundancy or lack of use.

Arbitrary location has been introduced to satisfy the requirement for a general hydro location, with an unassigned type.

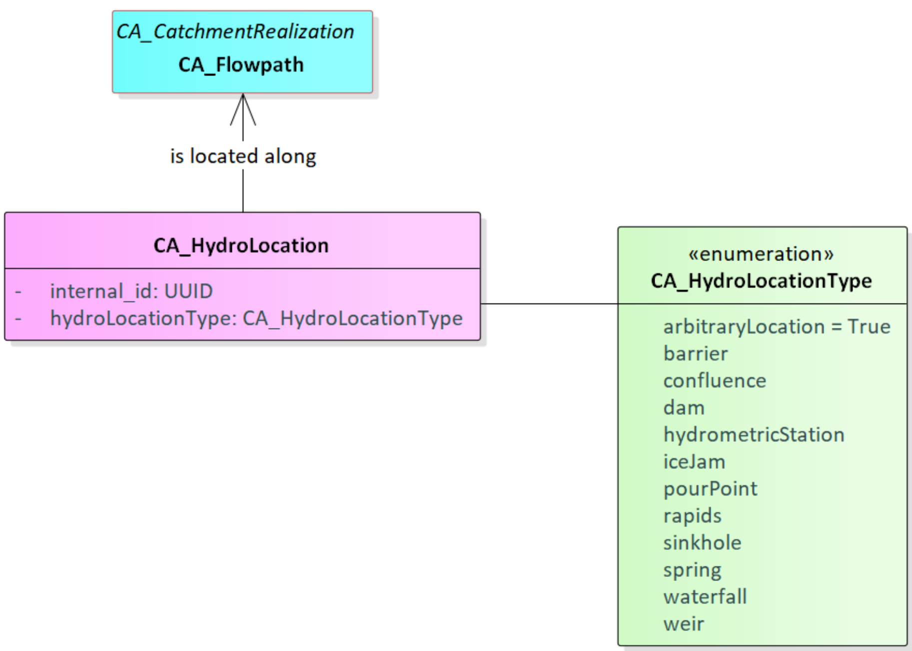

Figure 20: Hydro Locations

6.2.5 Hydro and Hydrologically Related Networks

Following HY_Features, CHyF recognizes three types of hydro networks (figure 21): hydrographic networks containing waterbodies (including instances of linear and polygonal rivers, lakes, estuaries, etc.), catchment networks containing catchments, and channel networks containing depressions (including instances of the subclass channel).

CHyF also includes two types of hydrologically related networks: flowpath networks containing flowpaths, and catchment divide networks containing catchment divides. Flowpath networks are discussed in HY_Features but are not part of the formal specification. Of note in figure 21 is that for every hydro network a corresponding flowpath network and catchment divide network exist. Similarly, each flowpath network has a one-to-one relationship with a catchment divide network.

Catchment divide networks are introduced but they currently are not used in CHyF's implementation.Flowpath networks and catchment networks are essentially embedded in the notion of the hygraph, so they are of fundamental importance to the CHyF model.

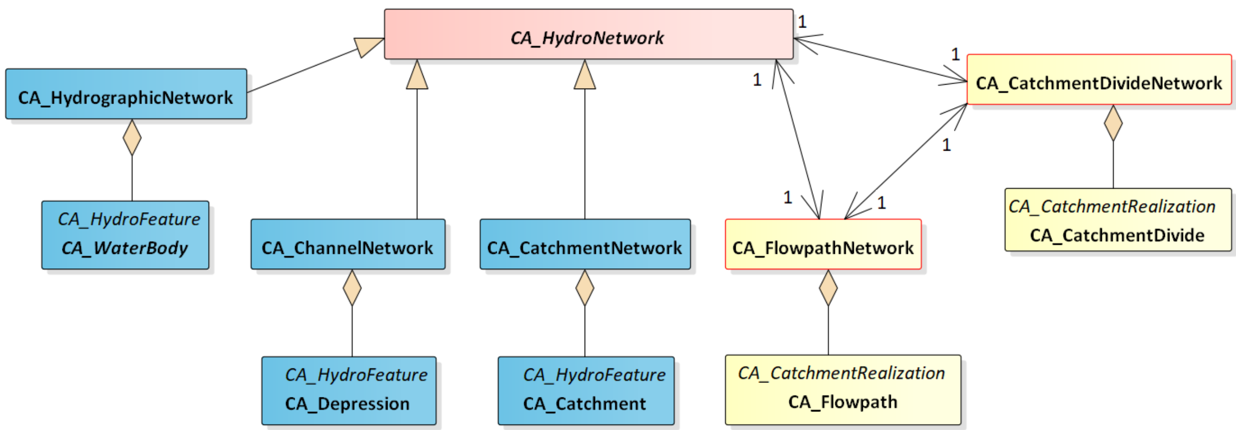

Figure 21: Hydro and hydrologically related networks

6.3 Structure and Relations of Elementary Features

The two central constructs of CHyF are elementary catchments and elementary flowpaths, as shown in figure 22.

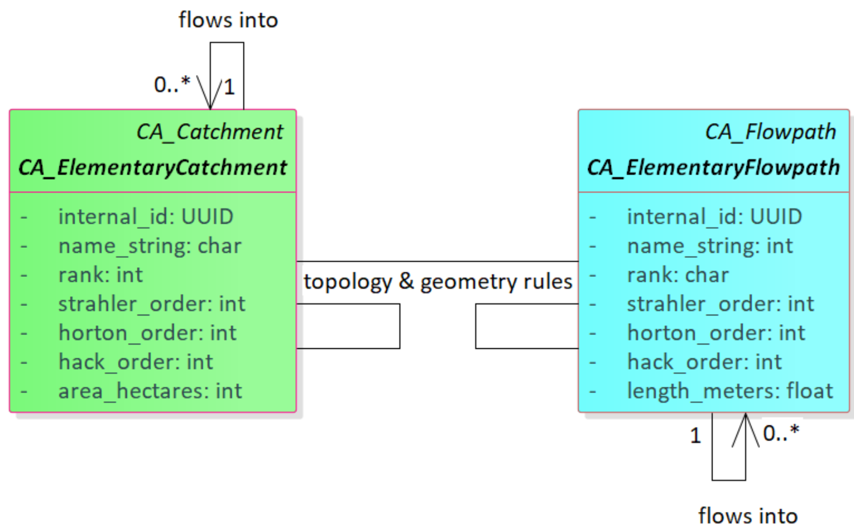

Figure 22: Elementary catchments and elementary catchments

These two are core to the current CHyF implementation. As described in other documents in this series, their respective attributes are generated algorithmically by CHyF tools.

Figure 23 below shows the relations between elementary catchments, elementary flowpaths and catchment divide segments. The centre part of the diagram shows that elementary flowpaths act to drain elementary catchments, with the details dependent on the flowpath and catchment types. Every reach catchment is drained by either an observed or inferred reach flowpath. Every bank catchment is drained by a bank flowpath. Water catchments are considered to be drained by potentially many skeleton flowpaths as well as by potentially many bank flowpaths. Catchments defined as built up areas are drained by potentially many infrastructure flowpaths; in these cases though the individual flowpaths do not necessarily correspond to individual catchments. Empty catchments do not have a relationship to elementary flowpaths since they do not contain and are not adjacent to surface water.

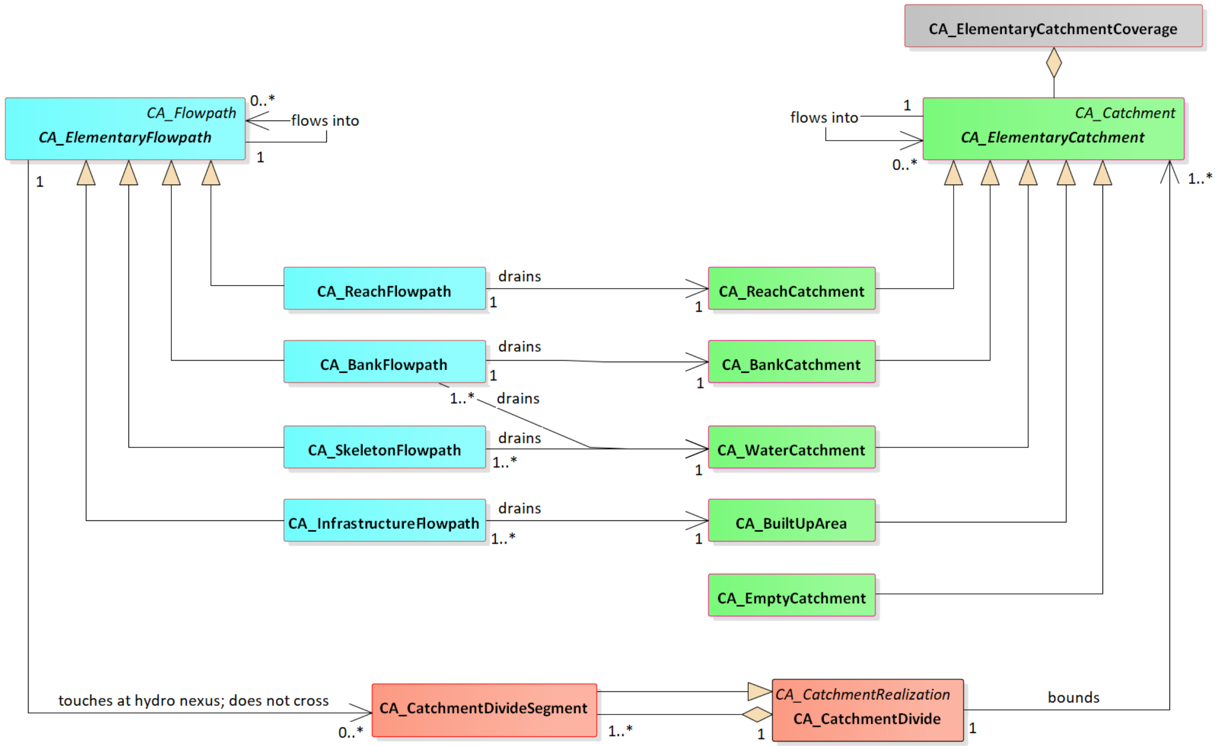

Figure 23: Elementary catchment, elementary flowpath and catchment divide relationships

Figure 23 also shows that every elementary catchment is bounded by a catchment divide; that is, the boundary of the elementary catchment defines a catchment divide. Other catchment divides may be specified for arbitrarily large catchments, but their constituent catchment divide segments will also be part of the boundary of local elementary catchments.

Shown in figure 24 are the relationships among four types of elementary catchments, three types of elementary flowpaths, and six instantiable types of hydronodes. Like those displayed above in figure 23, these relationships can contribute to a set of topological rules that can be applied to test data integrity (see section 6.3.2). Figure 24 takes into account naturally occurring elementary catchments, flowpaths and hydronode; built-up areas are addressed separately.

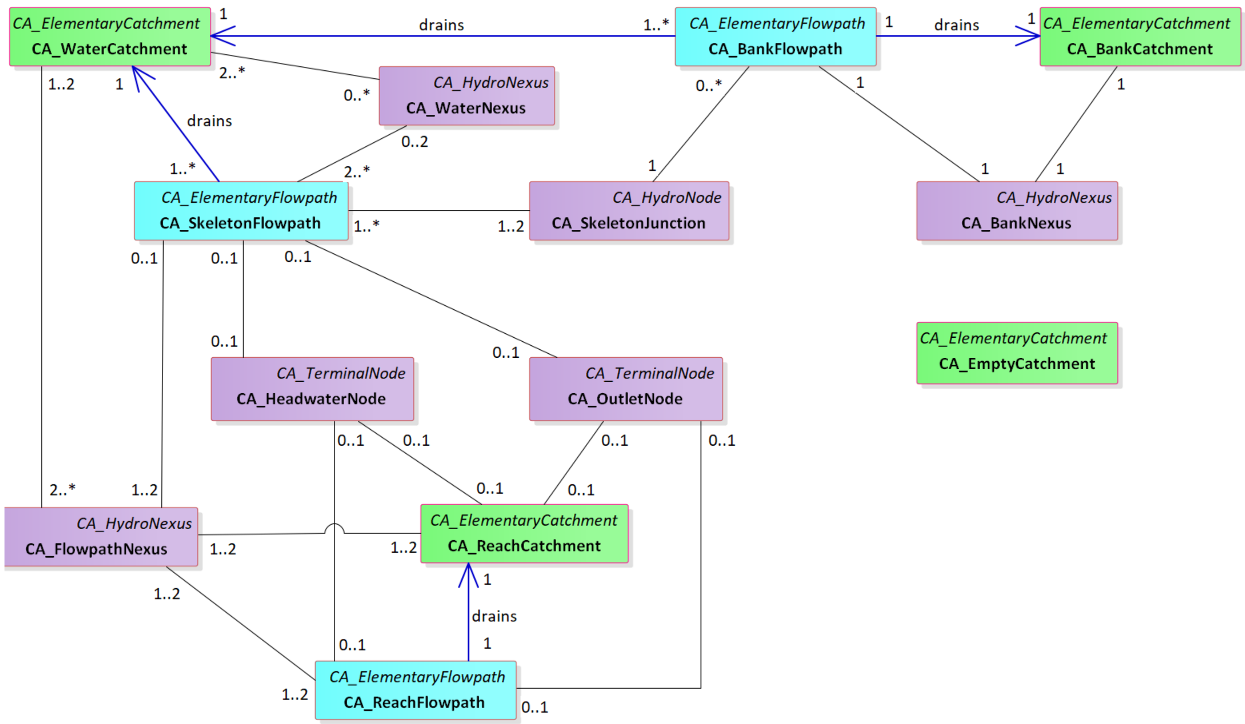

Figure 24: Elementary catchment, elementary flowpath and hydronode relationships

Figure 25 below considers built-up areas, infrastructure flowpaths, and infrastructure junctions. The infrastructure flowpaths represent infrastructure elements, just as reach flowpaths represent rivers. Note that the built-up areas are drained entirely by infrastructure flowpaths and that their connection to the outside network is through flowpath nexuses.

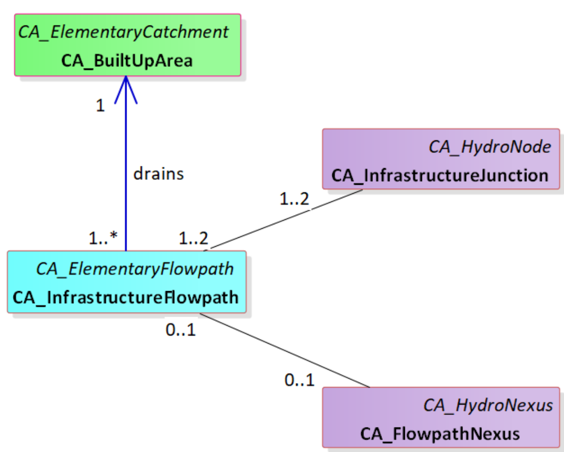

Figure 25: Built-up area, infrastructure element and hydronode relationships

6.4 Elementary Feature Topology Rules

6.4.1 Primary Topology Rules

The rules are based on the topological relationships between features and can be expressed using the Dimensionally Extended nine-Intersection Model (DE-9IM) model, which is recognized by the OGC Simple Feature Access specification. 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 may have slightly different semantics depending upon use.

The following rules apply to the topological relationships between elementary catchments and elementary flowpaths.

The following types of elementary catchments (LC) exist:

1. ReachCatchment (RC)
2. BankCatchment (BC)
3. WaterCatchment (WC)
4. EmptyCatchment (EC)
5. Built-up Area (BA)

In all cases they are represented geometrically as polygons, with an interior area and a boundary consisting of one outer ring and zero or more inner rings.

The following types of elementary flowpaths (LF) exist:

1. ReachFlowpath (RF)
2. BankFlowpath (BF)
3. SkeletonFlowpath (SF)
4. InfrastructureFlowpath (IF)

In all cases they are represented geometrically as linestrings, with an open-ended line as the interior of a line string and a boundary consisting of two endpoints.

The rules between these elementary features are as follows:

1. Every RF is *contained by* an RC.
2. Every RC *contains* one and only one RF.
3. Every BC does <u>not</u> *contain* an LF.
4. Every BC *touches* one and only one BF.
5. Every BF is *contained by* one and only one WC.
6. Every SF is *contained by* one and only one WC.
7. Every IF is *contained by* one and only one BA.
8. Every EC is *disjoint* from all LF.
9. Every EC *touches* one or more LC.
10. Every LF does <u>not</u> *cross* any other LF.
11. The interior of every LF does <u>not</u> *touch* the boundary of an LC.
12. An endpoint of every LF does <u>not</u> *touch* the other endpoint of the LF.

13. Every BC *touches* one or more WC and the intersection of each such BC–WC pair is a linestring.
14. All LC form a continuous coverage with no gaps and no overlaps, and with matching vertices along the boundaries of adjacent LC.

Rules for LF endpoints could be defined, but there is no reason to do so. If the rules above are all validated, so would rules about the endpoints.

More detailed rules covering topology and geometry can be found in section 4 in one of the companion documents, *Volume 3: CHyF Data for CHyF Services Implementation Specification.*

6.4.2 Additional Topology Rules

An additional set of rules applies to features that typically are composed of a number of elementary features. These include catchments covering more than an individual elementary catchment, as well as catchment aggregates, dendritic catchments, and interior catchments. Arbitrary flowpaths, mainstems and and the various kinds of hydro networks are other higher level features of interest here.

The three most important rules are as follows:

1. Every hydronode that is not an outlet node must have at least one outflowing flowpath
2. Every hydronode that is not a headwater node or bank nexus must have at least one inflowing flowpath.
3. An arbitrary flowpath through a flowpath network must never form a cycle, whereby flow from a flowpath endpoint flows back into the same flowpath endpoint through the flowpath network.

Additional rules of interest are listed here:

1. Every catchment network *contains* a corresponding flowpath network.
2. In areas where linear waterbodies exist and where empty catchments do not exist, a one-to-one relationship exists between elementary catchments and corresponding elementary flowpaths.
3. In areas where one or more polygonal waterbodies exist, or where empty catchments exist, a one-to-one relationship does not exist between elementary catchments and elementary flowpaths.
4. Every dendritic catchment *contains* a dendritic flowpath network formed by elementary flowpaths that are attributed as primary flows.
5. Every drainage basin has a corresponding mainstem.
6. Every catchment aggregate is composed of elementary catchments that form a coverage that *covers* the area of the catchment aggregate.

6.5 Feature Identity

All features within CHyF have a feature identifier, *internal_id*, implemented as a UUID that is required to define their respective relationships. They are not considered as persistent, immutable identifiers that must be maintained through update cycles for example. Nevertheless, the feature identifiers of the elementary flowpaths and elementary catchments are exposed intentionally to the user to support analysis and quality assurance functions. Large basins and major mainstems are however considered to remain through updates, so their identifiers can be treated as persistent.

The general position of CHyF is that fast computation can often replace complex data structures. Maintaining identifiers indefinitely into the future is considered as unrealistic, especially as updates increasingly involve more detailed data, often with different drainage patterns, derived from lidar and high resolution imagery. As described in section 6.1.4, CHyF recommends using geospatial coordinates (potentially including time) as the basis of referencing specific point or linear locations on a stream network. CHyF services will help make this practical. If an area is updated, then the user can determine whether the coordinates that they maintain need to be adjusted.

On the other hand, features with unique identifiers can participate in Linked Open Data (LOD) - and Internet of Water (IoW) - infrastructures. Features of interest in this context could range from elementary flowpaths and catchments to very large features such as mainstems and drainage basins. The ways in which LOD and IoW can be supported are outside the scope of this document, but potentially of significant interest.

What is clear is that fully automated processes and linked data techniques can lead to vastly simpler data models. Whether feature identity is based on an immutable identifier, or position in time and space, or an on-the-fly calculation, interoperability and geospatial techniques exist to relate hydrological features to other kinds of data with comparatively little effort. A large, inherently complicated model need not be implemented.

6.6 Particular Cases

6.6.1 Islands and Secondary Flows

The lake with the island and the double-line river are broken into water catchments, with the blue dots representing the hydro nexuses where each water catchment flows into its downstream neighbour. The separation of the lake from the river is marked by a hydro nexus and a catchment divide segment that meets similar segments on the two opposing banks. The river could continue for many kilometres before it meets another polygonal waterbody. Alternatively and more appropriate for some applications, it could be broken into a series of water catchments. The subdivision of large waterbodies into smaller ones can be arbitrary, e.g., every two kilometres, or could be meaningful as is the case with the two boundaries across the water as shown. The separation of H and M is based on the former being considered a lake and the latter a river. The separation of M and N is based on continuing the boundary on the other side. Note that the boundary between J and K does not continue through the water and across the opposite side. So long as the flowpath connections are made, any of these subdivisions are acceptable.

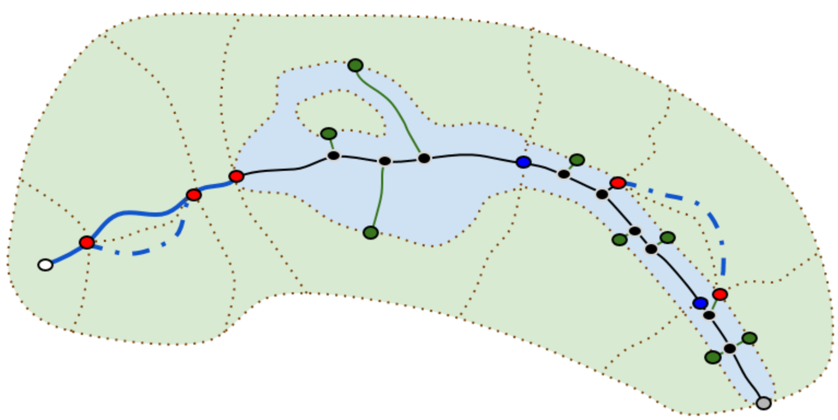

Figure 26: Islands and secondary flows

A few other characteristics evident in the figure are of note. (i) In figure 14 the red and green dots alternate around the lake. This is a common pattern. However, in figure 26 neither I nor its neighbour P contains a stream, so drainage from each is represented by a green dot, resulting in the two green dots being neighbours along the bank of the river. (ii) Catchment F is an island that does not contain a stream. Nevertheless the island is connected to the river through a bank flowpath extending from its bank to the skeleton flowpath in the middle of the river. (iii) Along the skeleton in the lake, most intersection points have a valence of three (three edges meet at the junction). The most downstream junction (black dot) in water catchment N has a valence of four. The two blue hydro nexuses have a valence of two. The CHyF model does not have a required or preferred value for the valence of such intersection points.

6.6.2 Distributaries

In figure 26 above, less important braids are represented as secondary flows (rendered as patterned blue lines). The same distinction also applies to flows below any divergence, including distributaries, as found on deltas (figure 27 below).

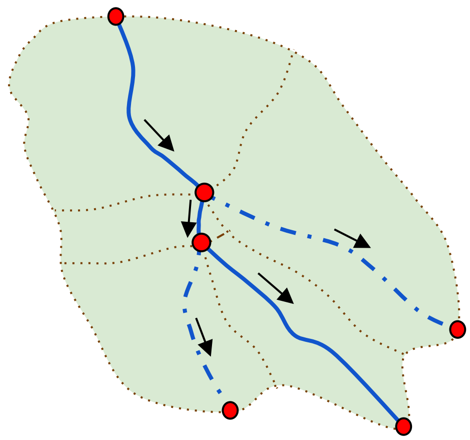

Figure 27: Distributaries below points of divergence

6.6.3 Subdivisions of Flowpaths and Catchments

Flowpaths and hydro nexuses may be defined somewhat arbitrarily if a point of interest exists along a stream. In the upper half of figure 28 below, this is demonstrated with a hydro nexus of valence 2 (two segments touch the common vertex) placed in the middle section of the flowpath network (the red dot with a white x). It may correspond to a gauging station or a proposed culvert location for example. Adding this hydro nexus leads to the original elementary flowpath at that location being broken into two elementary flowpaths and the original elementary catchment being broken into two corresponding elementary catchments.

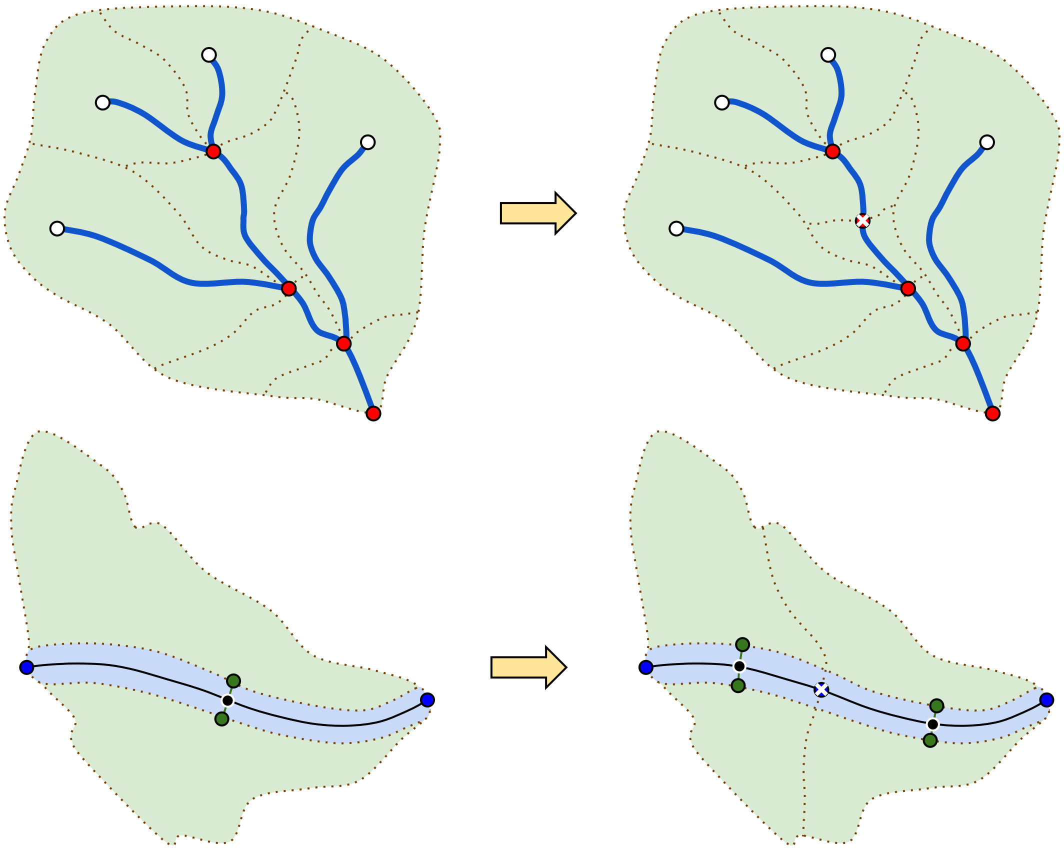

Figure 28: Subdividing elementary flowpaths and elementary catchments

If the new nexus is placed in a double-line river as shown in the lower half of the figure, the situation is more complicated, but the same principles apply. The original three elementary catchments consist of two opposing bank catchments and one water catchment. These three are each broken in two, with six elementary catchments then resulting. The black line segments are flowpaths, which double in number with the introduction of the new catchment divide segments (depicted as dotted, brown lines). The same is true of the number of nexuses, if the shared (blue) water nexuses at either end are each considered as contributing a half to the area in question.

6.6.4 Depressions and Channels

Depressions are concavities that contain or potentially contain water. They may be small pits that can be removed by software processes; however, in some cases they may be of significant size and could represent interior catchments that do not generally contain a waterbody. Depressions may also take the form of channels and support or potentially support water flow. Headwater ephemeral streams are good examples that may be mapped simply as channels. Complex gully systems and arroyos could be mapped similarly.

With the advent of lidar mapping, elevation point clouds or grids may be analyzed with the intention of defining waterbodies, and additional depressions and channels. This may result in isolated depressions and discontinuous channels. In some cases it may be evident that two channels are connected by a culvert through a roadbed for example. In other cases the topography may suggest that they must be connected. In both situations connections may be made to form a continuous network by adding inferred sections. If connections are made, then the derived channel can be combined with other hydro features to form continuous hydro networks. If not, then they can still be modelled by CHyF, with interior catchments resulting from the discontinuities. Depending upon the objectives of a given use case, this may or may not limit the applicability of the data.

6.6.5 Hierarchical Catchments and Partitioned Catchments

Large catchments may be composed of smaller catchments in different ways. Shown below in the upper right of figure 29 is a hierarchical structure with a large green catchment containing two catchments plus a lower remnant catchment. Each of the two upper catchments can be further subdivided in a recursive fashion. With such a hierarchy of order n, a given elementary catchment may be contained in from 1 to n catchments, including itself. As depicted, the four headwater catchments are in first, second and third order catchments; the lowest elementary catchment is only in the third order catchment. Outlets, depicted as red circles, are shared by adjacent catchments; however, the lowest outlet may or may not be shared in this fashion. Hierarchically specified catchments by definition overlap one another. These relationships are all part of the dendritic catchment – catchment aggregate model defined in HY_Features.

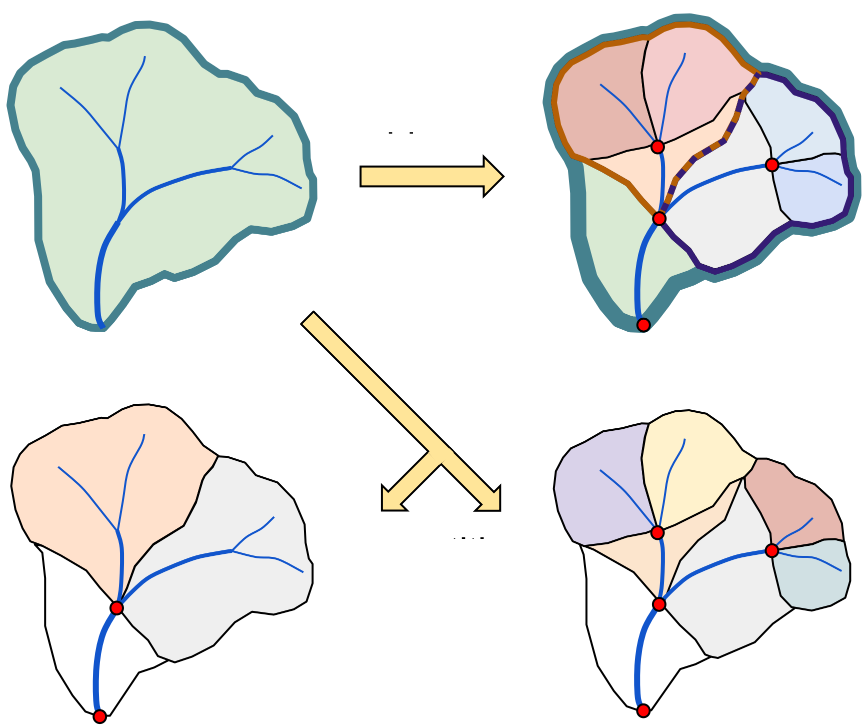

Figure 29: Hierarchical (overlapping) and partitioned (non-overlapping) catchments

Another way of considering the larger catchment is simply as a set of non-overlapping drainages, as shown in the lower two figures of figure 29 above. In the lower left, each of the three drainages is a catchment in its own right with zero or one inflows, and together they partition the overall watershed. The set of all elementary catchments in a given catchment always hydrologically partition it, as shown in the lower right. Hydrological partitioning means that a large catchment can be subdivided into non-overlapping catchments that form a coverage of the large catchment. The large catchment can be said to be a hydrologically partitioned catchment. Major drainage basins are large catchments typically broken into a set of drainage areas, with each meeting the criteria of a catchment; in such cases the basins can be said to be partitioned. The term partition as used here follows from the notion of a partition in set theory [Wolfram, 2019], in which the division of a set (e.g., all integers) results in multiple sets (all even integers and all odd integers), with the union of these sets covering the original set; modifications have been made to accommodate geographic and hydrologic considerations.

At any given level of detail, the outlets may all fall along the mainstem of the large catchment or they may fall along the mainstem and its tributaries. Assume that *numc* space-filling, non-overlapping catchments exist within a partitioned catchment, and that they drain into *numo* outlets, then in a classic dendritic pattern with every outlet, except the most downstream one, of valence 3 (i.e., two inflows and one outflow), the following holds.

$$numo \leftarrow floor(numc/2) + 1 \qquad [Eq\ 1]$$

The catchments shown in the lower half of Figure 29 above are in contrast to the subdivisions shown on the left side of Figure 30 below. In the case below, three non-overlapping drainage areas are defined, with C as the area remaining from the overall catchment after removing A and B. A and B both flow into C and each meets the criteria for a catchment. C does not meet the criteria for a catchment because it is partly downstream and partly upstream of one or more inflowing catchments (A and B), which also is not allowed. CHyF services can work with such drainage areas, but the results returned will not necessarily be hydrologically ideal.

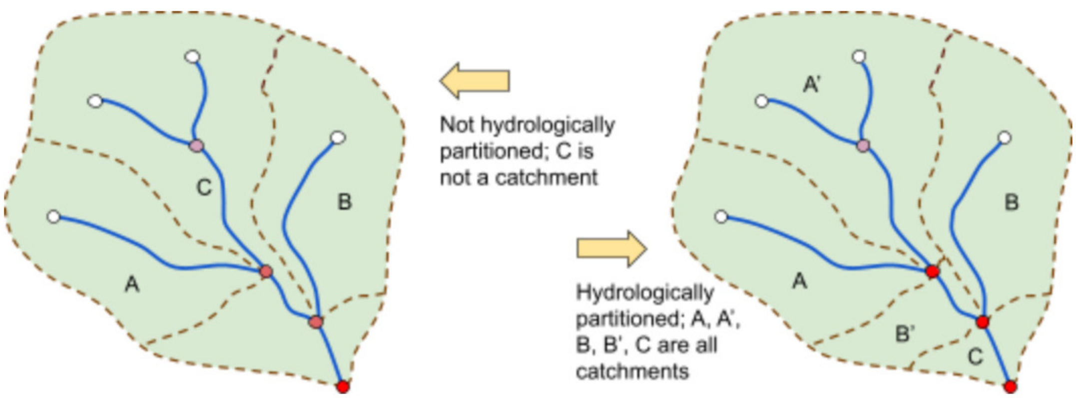

Figure 30: Hydrologically partitioning compared to an arbitrary subdivision of a large catchment

This situation is remedied by the further breakdown shown on the right side of Figure 30. A' drains into the same hydro nexus as A, and they both flow into B'; similarly, B' and B flow into C through a common hydro nexus. These four catchments plus C as depicted do meet the condition of being partitioned components of the overall catchment.

For a real world example, consider the Susquehanna River Basin, shown in figure 31 below. It is broken into six non-overlapping drainages. N1 is a hydro nexus serving as the common outlet for the Chemung and Upper Susquehanna Subbasins. N1 is also the inflow for the Middle Susquehanna. Further downstream at N2, the West Branch and the Middle Susquehanna drainages flow into the Lower Susquehanna drainage. So far these drainages taken together meet the partitioned definition described above.

At N3 the situation is different. The Juniata River joins the Susquehanna River, with part of the Lower Susquehanna Subbasin above that confluence and part of it below the confluence. If a catchment divide were drawn, approximately as shown by the white dashed line, then the Lower Susquehanna drainage could be split into a northern area and a southern area. The net result would be a set of seven catchments partitioning the Susquehanna River Basin. As shown though, without the white dashed catchment divide, this is not the case. The northern and southern areas of the Lower Susquehanna were likely combined in an effort to keep all subdivisions of the Susquehanna Basin of roughly comparable size, in line with the intentions of hydrologic units as defined by the US Watershed Boundary Dataset.

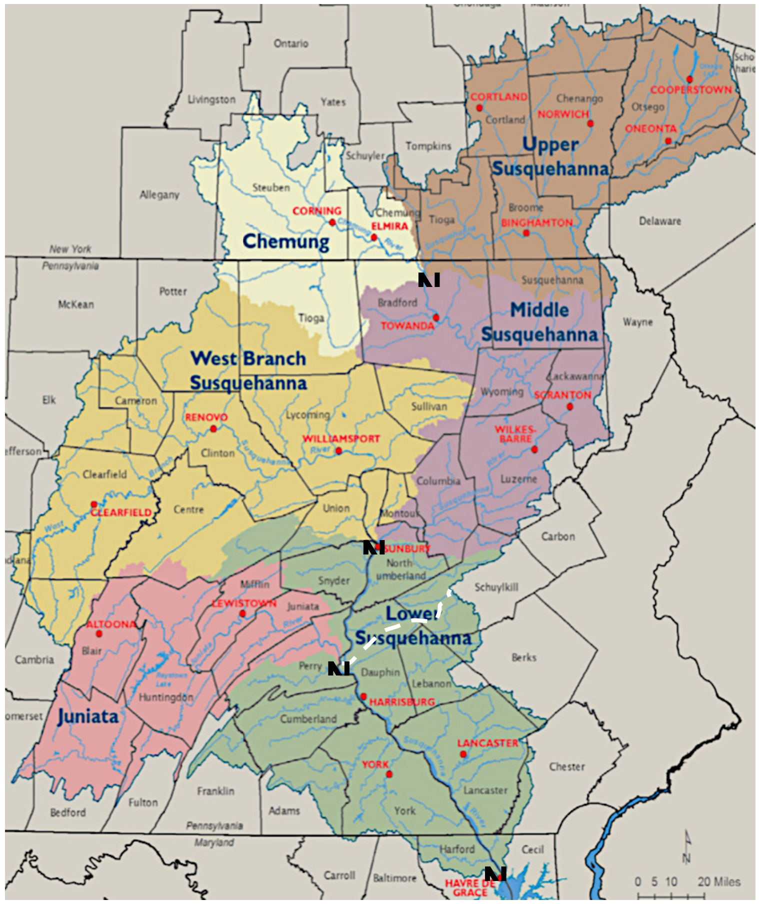

Figure 31: Subdivisions of the Susquehanna River Basin

6.6.6 Nearshore Zones and Large Waterbodies

Figure 32 below is part of the drainage for a large lake with four inflowing rivers and one outflowing river. The corresponding reach catchments for these rivers are rendered in pink with the bank catchments between them in light green. The flowpaths are shown by arrows indicating the direction of flow, which is always directed and acyclic, but not dendritic. The nearshore zone is depicted and is subdivided into 10 water catchments, rendered with dotted brown boundaries. The three enlargements relate to bank flowpath, skeleton flowpath, and catchment boundary segment, respectively. The blue, black and green hydro nexus points are as before.

Whether the nearshore zone is subdivided or how it is subdivided is up to the user. The inner boundary may for example be based on bathymetric data, on the location of the littoral zone, or on a buffer of a given dimension. The breaks along the zone could be based on influence areas from inflows, on the locations of towns or industrial areas, or on set lengths along the shore. CHyF only imposes the requirement that one or more flowpaths exist in each waterbody and that these flowpaths can be constructed into a directed acyclic flowpath network that meaningfully represents the general flow of water. As shown, each of the nearshore zone units plus the middle of the lake are treated as water catchments. Flow is assumed along the shore toward the outlet and away from the shore to deeper water.

The kind of treatment shown in this figure could be applied to the Great Lakes, other large lakes, large rivers, estuaries or along the coastline. This flexibility increases the potential application space to which CHyF data and services can be applied.

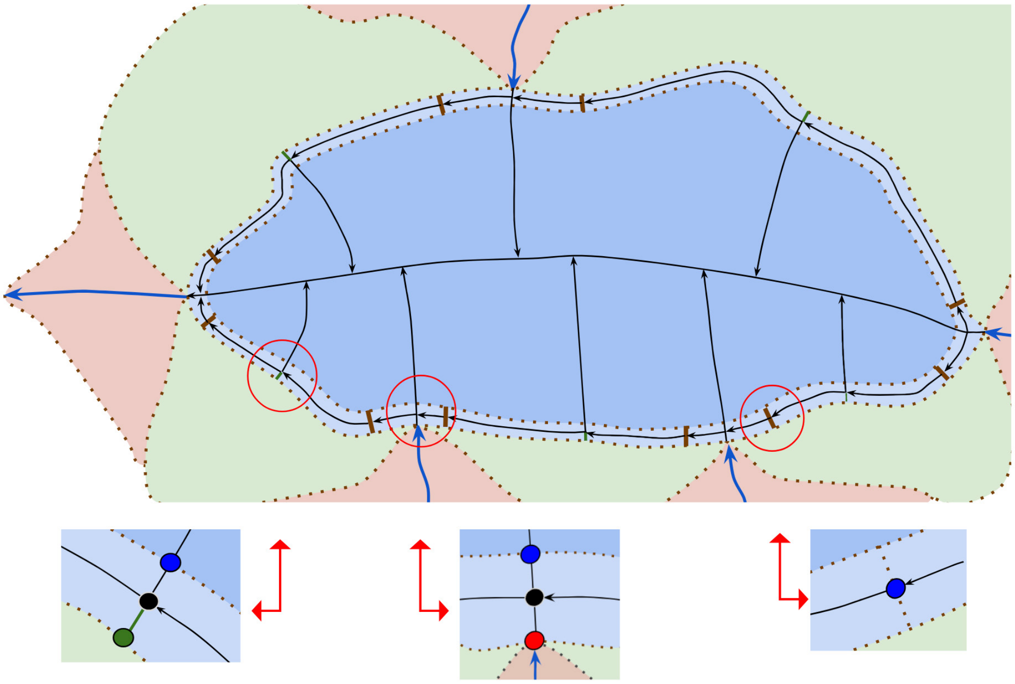

Figure 32: Nearshore zone of large lake broken into a series of catchments (enlargements show, left to right, bank flowpath, skeleton flowpath, and catchment boundary segment)

6.6.7 Hydro Locations

Locations on or near a flowpath network may be of particular interest. If they are of direct hydrological significance, they are referred to as hydro locations. This term is used in both HY_Features and CHyF. HY_Features provides a list of possible hydro locations, but allows for other lists to be used instead. CHyF provides an alternative list, which includes the following: `Dam`, `Hydrometric Station`, `Pour Point`, `Rapids`, `Sinkhole`, `Spring`, `Waterfall`, and `Weir`. The members of this list are commonly included in government mapping programs or are of direct interest to hydrologists and environmental specialists, which is why they are included here.

Other features could have been added to the list, including many that are not necessarily of interest to the hydrological modelling community. CHyF services related to location can be applied to hydro locations, but they can also be applied to other locations that can be represented geometrically by a point on or near a stream. The upstream catchment of an arbitrary pour point on a stream may be of interest. However, the same is true of a fish inventory site or even of the location where a cadastral or administrative boundary intersects a stream.

Such sites are not considered to be hydro locations, even though from a services perspective they are treated no differently.

7. The Hygraph

Mathematical graph theory can be used to navigate through relationships between elementary catchments, elementary flowpaths and nexuses. This can be implemented in different ways, but the basic ideas are presented here. Figure 33 below adds Q, an empty catchment to figure 14; it represents a depression without any contained waterbody at the time of mapping. As well, another catchment, W, is added to the left, and the ocean is ignored.

Two large catchments are evident, one consisting of the elementary catchment W and the other of a set of elementary catchments, A … K. Y-Z exists along the boundary of both and is not contained by either. However, they both have a common outlet at the confluence in the lower right. An upstream query from that hydro nexus would include A through K, plus W, but exclude Q and Y-Z. Alternatively, the query could specify that everything within its outer boundary be included, in which case A...K, Q, W, and Y-Z would all be included.

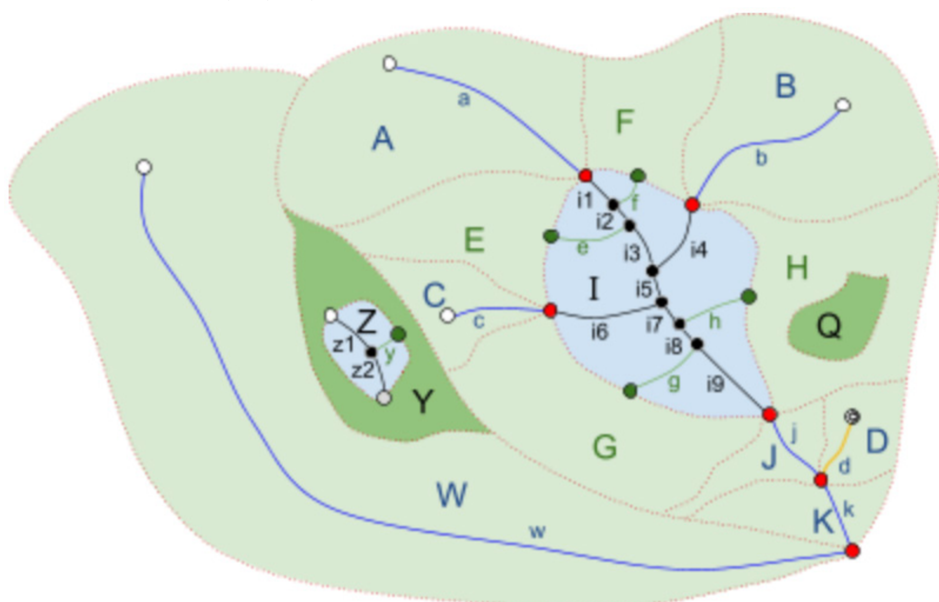

Figure 33: Elements of a hygraph shown geographically

The hydronodes are not identified individually in Figure 33. However, they are in an actual implementation, where the flow from one catchment to another or from one flowpath to another is always through a hydro nexus. For example, in Figure 34 catchments J, D, and K have flow relationships through hydro nexuses (labelled in Figure 34 below as *m, n,* and *o*). Collectively these features and their relationships can be represented as part of a Directed Acyclic Graph defined by alternating hydro feature - hydronode sequences.

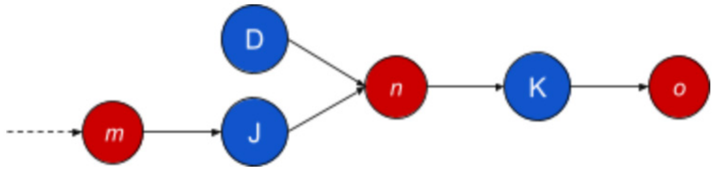

Figure 34: Elements (elementary catchments and hydronodes in this case) of a hygraph shown schematically

Flowpaths d and j (in D and J) flow into flowpath k (in K) through *n* as well. For reasons of clarity, the hydronodes are not shown in the diagrams that follow, but should be assumed to be present throughout.

With that understanding, the flow relationships between the elementary catchments and the elementary flowpaths are shown in the following figure. The relationships in the larger catchment draining into the ocean are shown by the large green, blue, and brown subdiagrams on the left. The relationships in the interior catchment Y-Z are shown on the right. The empty catchment Q has no flow relationships and is found in the upper right.

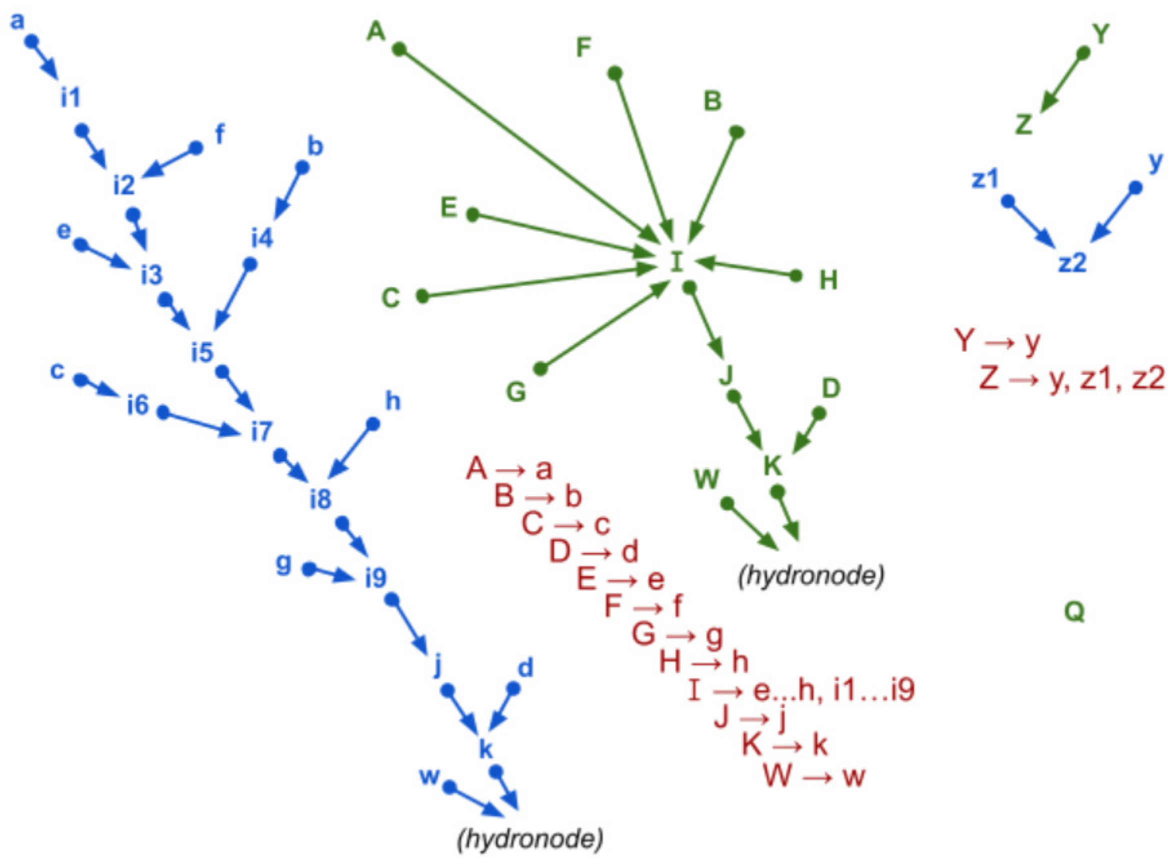

Figure 35: Different groups of connections forming subgraphs

Using the catchment to flowpath relationships in dark red, the seven subfigures above can be reduced to three, as shown in the next figure. G_0 is large, containing many connections. G_1 is much smaller but still contains several connections. G_2 , corresponding to an empty catchment, includes a null set of flow connections.

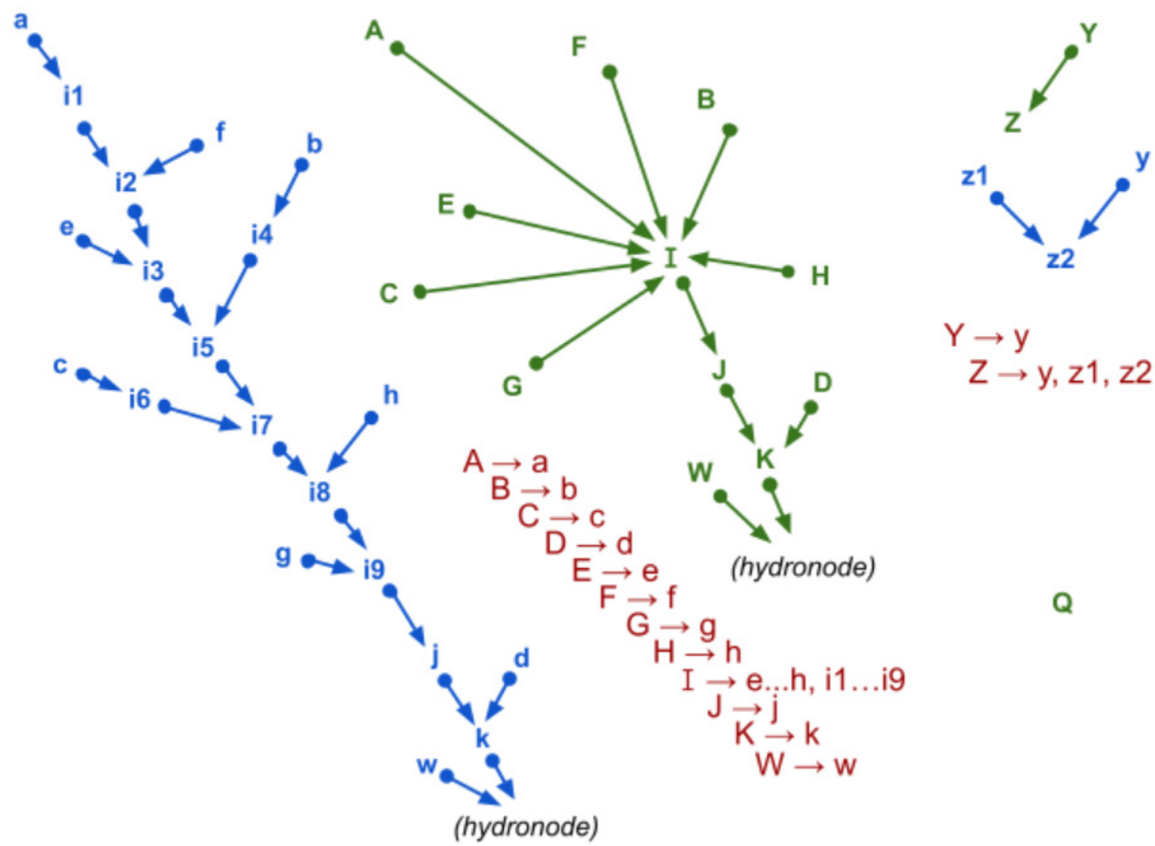

Figure 36: Different groups of connections structured as subgraphs of a single hygraph

Each of the three subfigures represents a connected Directed Acyclic Graph (DAG). These three DAGs are associated with three separate catchments. However, it is also reasonable to amalgamate the three into a single catchment identified as the area of interest, or the universe U under consideration. In this more general case, U is defined as a set of DAGs as per the following expression.

$$U \leftarrow \{G_0, G_1, G_2, \ldots\} \qquad \text{[Eq 2]}$$

Each DAG, G_i , is a component of the graph U. The hygraph is an implementation of this expression for any arbitrarily large catchment or set of catchments.

As noted earlier, the hydronodes are taken into account in the construction of the hygraph. They serve as interfaces as well as start and endpoints and are required when implementing this graph approach. Each flowpath graph G can be defined as a tuple of flowpaths, catchments, hydronodes, flow relationships and containment relationships. This expression can apply to the entire graph or to each individual DAG.

$$\text{Network Flow Graph} \leftarrow (\text{Flowpaths, Catchments, hydroNodes, Flow Relationships, Containment Relationships}) \qquad \text{[Eq 3]}$$

Using abbreviations, this can be written more compactly.

$$G \leftarrow (F, C, N, FR, CR) \qquad [Eq\ 4]$$

The expression below states that the flow relationships are defined by a set of relationships between x and y, where x flows into y if either of the following two conditions holds: (i) x is a flowpath or catchment and y is a hydronode, or (ii) x is a hydronode and y is either a flowpath or catchment. This has the same meaning as stating that a flowpath flows through a hydronode into a catchment or another flowpath, or a catchment flows through a hydronode into a flowpath or another catchment. Another way of considering this expression is that it indicates what is happening at every hydronode.

$$FR \leftarrow \{(x,y) \mid (x \in F \cup C \text{ and } y \in N) \text{ or } (x \in N \text{ and } y \in F \cup C)\} \qquad [Eq\ 5]$$

The next expression defines containment relationships as a set of relationships between catchment c and flowpath f.

$$CR \leftarrow \{(c,f) \mid c \in C \text{ and } f \in F\} \qquad [Eq\ 6]$$

We also define that each flowpath, f, is within a catchment, c. This does <u>not</u> imply that these relationships are of a one-to-one nature; similarly, this does <u>not</u> imply that every catchment contains a flowpath. An examination of the previous map figures demonstrates these situations.

$$\forall\ f \in F:\ \exists\ c \mid (c,f) \in CR \qquad [Eq\ 7]$$

These expressions give rise to a number of topological rules that are the subject of a later section in this document.

As demonstrated here, the hygraph may contain connected or disconnected elements. Because the feature connections are all within the same graph, navigating through the features is straightforward. Downstream and upstream relationships are readily determined, as are interior catchments. Braided streams, lakes with multiple outlets, and deltas can all be handled in the same way as classic dendritic patterns.

8 Network Traversal

8.1 Referencing Point and Linear Locations

The NHDPlus supports linear referencing based on meaningful codes. It allows the determination of what is upstream or downstream from what, along a river system without any need to interrogate the geometry. The importance of a section of river can be inferred from coding as well. HY_Features describes linear referencing of positions along a river based on a nominal mainstem, although it does not delve into implementation details.

CHyF does not support the linear referencing model used in the NHDPlus. Two reasons exist for this omission. The first and most important is that given the nearly instantaneous speed of graph operations, leveraging traversal techniques through a network provides similar if not much better performance than determining paths based on codes. The second reason concerns maintenance. If codes are not needed then maintenance is significantly reduced, especially as related to updating with its attendant conflation from old to new river segments.

8.2 Point Locations on a Flowpath Network

With CHyF it is sufficient to provide coordinates, specified in a coordinate reference system, for a given location of interest. That location may refer to an event or a general position of interest along a hydro network. The location may be on (e1 in figure 37 below) or near a flowpath (e2); in the latter case, a position on the flowpath can be determined by direct projection onto the nearest point on the nearest flowpath or potentially by calculating a downhill path to the flowpath. Relative position and network distances among points located along the network can be determined as described in the following subsections.

In cases where the position of a specific location cannot confidently be assigned to the nearest stream (e3), then a proxy location can be used (e3’), implemented as another point, as shown in Figure 37. This secondary point must then be maintained with the original point. This approach does involve some overhead, but it neatly circumvents the problem of snapping to the wrong segment.

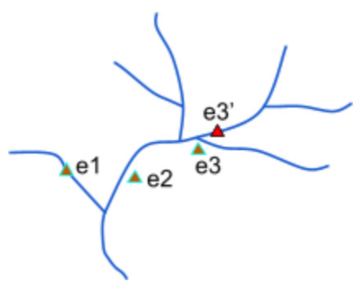

Figure 37: Point locations on a flowpath network

8.3 Linear Locations on a Flowpath Network

In addition to point locations, a location can also be defined as a section of a flowpath that may be contained in a single elementary flowpath or may instead consist of a series of elementary flowpaths. The linear event can always be mapped onto a flowpath, but it differs from a flowpath in two important respects. Flowpaths indicate flow and thus always have a direction, whereas a linear location does not. A linear location may be entirely within a single elementary flowpath (s1↔t1 in figure 38), with both a starting point and a terminal point on the same segment. This is the simplest case. Alternatively, the location could include all of a number of intermediary elementary flowpaths as well as all or part of each of the two elementary flowpaths on either end (s2↔t2), along a mainstem. Finally, the location may be defined by any arbitrary, directed path through the flowpath network (s3↔t3).

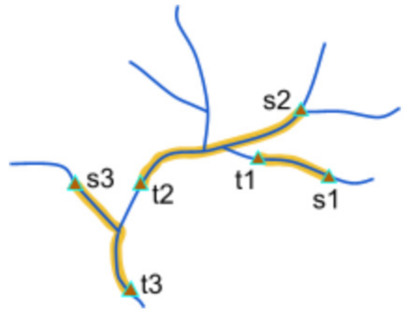

Figure 38: Linear locations on a flowpath network

In all cases, all that is required to reference a linear path is the following:

1. A start point on or very near an elementary flowpath, corresponding to a point location on the flowpath network,
2. A terminal point on or very near an elementary flowpath, corresponding to a point location on the flowpath network,
3. Assurance that a directed path exists between the start point and the terminal point.

The third point is important as it indicates for example that a single linear location cannot be defined between s1 and s2. If it were of interest to denote sections of the network that did not have a simple, directed flow relationship, then multiple linear locations could be referenced. These linear locations could form a contiguous linear space or could instead form discrete, unconnected linear spaces.

As described above, CHyF supports referencing of point and linear locations on a network, but with a different model than has been used previously. Linear referencing in the NHDPlus model allows for upstream - downstream relationships to be determined by comparing the values of the references along a given stream. With CHyF such addressing is replaced by using network traversal services. If it were of interest to know what the network relationships are among a number of point or linear locations, in relative or absolute terms, such services can be used to provide answers extremely quickly. The following subsections provide further background.

8.4 Network Distances between Locations

Consider a number of points of interest in a large watershed. A network distance can be determined from one location to another through the network. In Figure 29 below points p1 through p7 are depicted with triangle icons. These may be pour points intended for hydrologic analysis, hydrometric station locations, fish inventory sites or general points of interest. Assume the following: p2 and p5 are snapped to the nearest point on the nearest elementary flowpath; p6 is treated as the outflow point for both inflowing flowpaths; and where a pour point falls along a flowpath and not at an endpoint, the flowpath is broken there and the corresponding elementary catchment is broken as well.

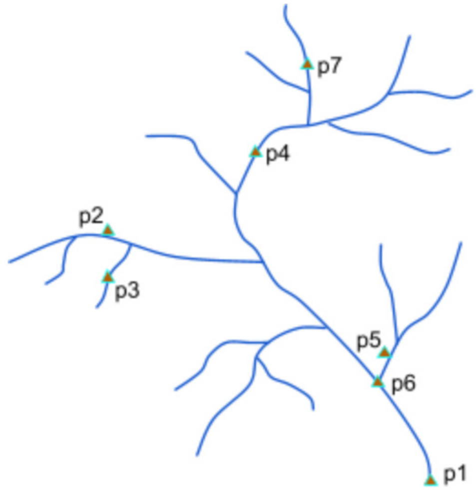

Figure 39: Point locations with implied paths and distances (see text)

A distance matrix, $d_{i,j}$, can be defined, with downstream distance indicated as a positive delta (∆) and upstream distance as a negative delta (-∆). The values below the main diagonal are equivalent to those above it, but with the sign reversed because of the difference in flow direction. The matrix could, if desired, be represented more compactly as a triangular matrix with no loss of information.

From\To	p1	p2	p3	p4	p5	p6	p7
p1	0	-Δ(p1,p2)	-Δ(p1,p3)	-Δ(p1,p4)	-Δ(p1,p5)	-Δ(p1,p6)	-Δ(p1,p7)
p2	Δ(p2,p1)	0	--	--	--	Δ(p2,p6)	--
p3	Δ(p3,p1)	--	0	--	--	Δ(p3,p6)	--
p4	Δ(p4,p1)	--	--	0	--	Δ(p4,p6)	-Δ(p4,p7)
p5	Δ(p5,p1)	--	--	--	0	Δ(p5,p6)	--
p6	Δ(p6,p1)	-Δ(p6,p2)	-Δ(p6,p3)	-Δ(p6,p4)	-Δ(p6,p5)	0	-Δ(p6,p7)
p7	Δ(p7,p1)	--	--	Δ(p7,p4)	--	Δ(p7,p6)	0

From this matrix it is easy to determine the complete set of major paths, encompassing all other paths, assuming that the network on which the points lie is dendritic. Starting points are those for which the deltas in their respective rows are all positive or zero (a negative delta implies that another point can be found upstream). The major paths, each starting with an uppermost point, are as follows:

```
p2 → p6 → p1
p3 → p6 → p1
p5 → p6 → p1
p7 → p4 → p6 → p1
```

Using the hygraph and spatial geometry, CHyF services can provide such information very quickly.

8.5 Network Flow Relationships between Locations

It may be of interest simply to state what pour point, or associated catchment, flows immediately into its downstream neighbour. This is shown in the table below for the example above. Note that this matrix amounts to a subset of the data defined in the table above. For some purposes this matrix is sufficient.

From\To	p1	p2	p3	p4	p5	p6	p7
p1	0	--	--	--	--	-1	--
p2	--	0	--	--	--	1	--
p3	--	--	0	--	--	1	--
p4	--	--	--	0	--	1	-1
p5	--	--	--	--	0	1	--
p6	1	-1	-1	-1	-1	0	--
p7	--	--	--	1	--	--	0

8.6 Order

Various ways of defining stream order have been developed. These all depend on a dendritic network that can be generated by either ignoring secondary flows or modifying their geometry such that a dendritic pattern results. CHyF provides direct support for the first technique since modifying the geometry means permanently changing the network connectivity information.

The most common stream ordering system is that defined by Arthur Strahler. Shown in the left-hand diagram of figure 40 below (after Wikipedia), it is often referred to simply as stream order. Robert Horton is responsible for a related ordering system, as shown on the middle diagram. Strahler order is readily determined moving from the headwater flowpath segments downstream. Horton order can then be derived moving in the opposite direction, as described later. John Hack worked on another ordering scheme, which is similar to Horton order, but beginning at the outlet and moving upstream. It is sometimes referred to as stream level or Gravelius Order. A useful artifact of determining Horton or Hack order is the definition of the mainstem, shown in red on the middle and right-hand diagrams. CHyF supports these three kinds of stream ordering, all of which can be algorithmically generated.

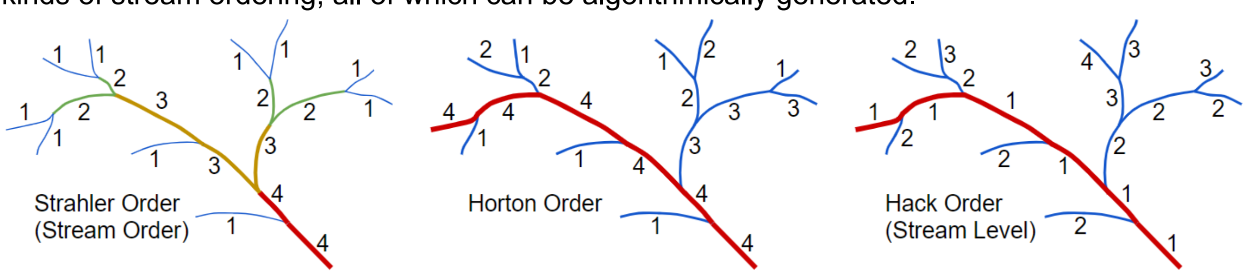

Figure 40: Stream ordering systems

Once either Horton or Hack Order is defined, the mainstem upstream from any arbitrary location on the network is defined. Horton order has been used to support generalization. For example in Figure 30, flowpaths could be removed with a Horton order of 1 or 2, and all others could be retained. The result would be a much simpler stream network that would also have much simpler catchments. The geometry of both the resulting network and the catchments could be simplified further through suitable algorithms (such as Douglas-Peucker) if desired. Other approaches to generalization exist, but this one can be easily implemented. Hack order has the property that order number is relatively invariant under different degrees of generalization. For example, a stream emptying into the ocean will always be of order 1, and branches that remain after pruning based on Horton Order, will have the same Hack Order as they did before pruning.

The determination of both Horton Order and Hack Order requires moving upstream following the most important stream at each confluence. Importance can be defined by: (i) name, (ii) actual flow, (iii) total upstream area, (iv) path length, (v) total upstream cumulative path length, or (vi) number of inflowing tributaries. Arguments exist for all of these. Name is often correlated with average flow magnitude, but flow data may be available or may be estimated. Cumulative length and path length are simple to measure. With current CHyF routines, name and total upstream cumulative path length are used as the default basis for such ordering.

Other ordering systems may be of interest. For example, Shreve Order provides the number of headwater nodes (equivalent to the number of first order streams) for a given stream segment. The Pfafstetter Coding System (Verdin and Verdin, 1999) embeds topologic relationships in the codes for each flowpath segment and corresponding primary catchment. A similar system has been put forward by Wang, et al (2020). CHyF does not support these, but they could be added as services if there were sufficient interest.

8.7 Point Relationship Tree

Through observation or calculation (using a CHyF preprocessing tool for example), the data may distinguish primary flows from secondary flows that may exist through braided channels or around islands. Similarly, the mainstem may be indicated as well. Horton Order can also be generated as a measure of importance for every flowpath: the higher the Horton Order value, the more important the flowpath. In Figure 41, Horton Order is indicated by the thickness of the line, with the heaviest line segments corresponding to the mainstem. Given the primary flowpaths, Horton Order values for each flowpath, and the hygraph, a Point Relationship Tree can be generated from the order of the points, from downstream to upstream.

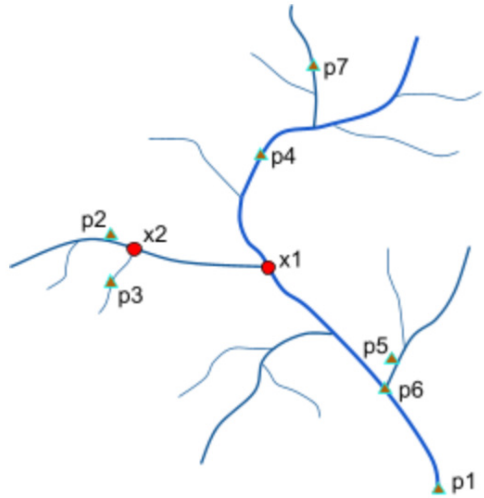

Figure 41: Point locations with additional x locations implied by the Point Relationship Tree, with line thickness corresponding to Horton Order (1, 2, or 3)

The same locational assumptions about the pour points described in the previous section apply here.

Consider just p6 and its upstream neighbour p5, a tree can be defined as `p6(p5)`. The round brackets contain between them what can be traversed upstream from the element appearing before the brackets.

Next consider just points p2, p3 and p4. The tree would be `x1(x2(p3,p2),p4)`. x2 is introduced as it represents the confluence of the branches that correspond to p3 and p2. Traversal from either p3 or p2 to x2 is possible, so p3 and p2 are in brackets preceded by x2. Since p3 and p2 do not have a traversal relationship, brackets are not used between them. p3 precedes p2 because the branch on which it resides is of less importance. x2 precedes p4 for the same reason, its corresponding flowpath is of less importance. x1 is introduced as the confluence of the branch associated with x2 and the branch associated with p4.

Now the entire Point Relationship Tree can be constructed. Following the same logic, the expression on the left below results. The two confluences x1 and x2 are additional points unspecified by the user. They are present to support branching in the tree. The tree can also be represented as a schematic as shown on the right, which makes it easier to understand. The upstream approach taken here to define the tree is sometimes referred to as *depth first* traversal, as the (upstream) children of each node in the tree are listed before the node's siblings.

Expression	Schematic
`p1(p6(p5,x1(x2(p3,p2),p4(p7))))`	p3 p2 p7 ↘ ↙ ↓ x2 p4 ↘ ↙ p5 x1 ↘ ↙ p6 ↓ p1

One important difference between the schematic and the actual geography is that with the schematically adjacent points at any level of the tree are located on flowpaths of increasing order of importance from left to right. p2 is to the right of p3 because its corresponding flowpath has a higher Horton Order value compared to that for p2. Such relationships are not necessarily the case with the actual geography; p5 for example is on the left side of the schematic but on the right side in the real world.

So long as the network on which the points are located is dendritic, a Point Relationship Tree can always be generated, which shows the downstream relationships and the branching structure. Reading from right to left, it is also possible to generate the paths noted earlier under the network distance matrix: `p7→p4→p6→p1, p2→p6→p1, p3→p6→p1,` and `p5→p6→p1.`

8.8 Catchment Relationships

Just as flowpath flow relationships can be represented by matrices, the same can be done for elementary catchments and for catchments in general. The two examples in figure 42 below are taken from earlier sections.

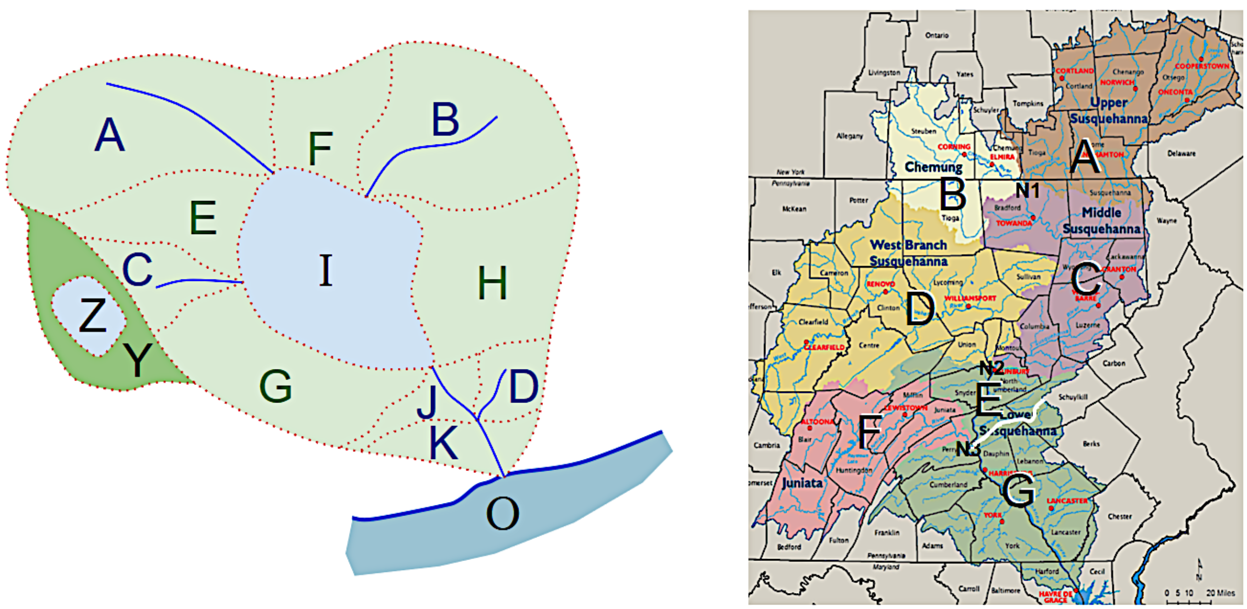

Figure 42: Catchment geography pertinent to simple flow relationships

The left example in Figure 42 involves elementary catchments. The flow relationships for it are shown below, and as was also the case earlier, the matrix is symmetric around the main diagonal with the exception of the sign reversal.

From\To	A	B	C	D	E	F	G	H	I	J	K	Y	Z	O
A	0	--	--	--	--	--	--	--	1	--	--	--	--	--
B	--	0	--	--	--	--	--	--	1	--	--	--	--	--
C	--	--	0	--	--	--	--	--	1	--	--	--	--	--
D	--	--	--	0	--	--	--	--	--	--	1	--	--	--
E	--	--	--	--	0	--	--	--	1	--	--	--	--	--
F	--	--	--	--	--	0	--	--	1	--	--	--	--	--
G	--	--	--	--	--	--	0	--	1	--	--	--	--	--
H	--	--	--	--	--	--	--	0	1	--	--	--	--	--
I	-1	-1	-1	--	-1	-1	-1	-1	0	1	--	--	--	--
J	--	--	--	--	--	--	--	--	-1	0	1	--	--	--
K	--	--	--	-1	--	--	--	--	--	-1	0	--	--	1
Y	--	--	--	--	--	--	--	--	--	--	--	0	1	--
Z	--	--	--	--	--	--	--	--	--	--	--	-1	0	--
O	--	--	--	--	--	--	--	--	--	--	-1	--	--	0

A similar matrix can be constructed for the large drainage areas shown on the right side of figure 42.

Code	Name	From\To	A	B	C	D	E	F	G
A	Upper Susquehanna	A	0	--	1	--	--	--	--
B	Chemung	B	--	0	1	--	--	--	--
C	Middle Susquehanna	C	-1	-1	0	--	1	--	--
D	West Branch Susquehanna	D	--	--	--	0	1	--	--
E	Northern Lower Susquehanna	E	--	--	-1	-1	0	--	1
F	Juniata	F	--	--	--	--	--	0	1
G	Southern Lower Susquehanna	G	--	--	--	--	-1	-1	0

9 Large Structures

9.1 Mainstems and Drainage Basins

Although the mainstem concept can be applied at any level of detail, it is most meaningfully applied with large catchments, i.e., recognized drainage basins. Each drainage basin has a mainstem which starts with a headwater and ends at an outlet, where water flows out of the drainage basin. Hierarchical and containment relationships exist between large drainage basins and smaller basins within them.

In figure 43 below, the mainstem of the Susquehanna River Basin is shown in light blue. Its headwater is the hydronode near the northeast corner of the figure, which is represented by a white circle. Its outlet is the hydronode along the southern border; it is represented by a blue circle with an embedded +. Four of the subbasins are also drainage basins in their own right: the Upper Susquehanna (A), the Chemung (B), the West Branch Susquehanna (D) and the Juniata (F). Each of these has a mainstem with a headwater as depicted by the white circle and an outlet rendered as a circle with an X. The Upper Susquehanna shares a headwater with the whole Susquehanna and its mainstem is the uppermost section of the mainstem for the parent basin. Neither the Middle Susquehanna (C) nor the Lower Susquehanna (E & G) should be considered as basins since they have more than one incoming flow. As discussed earlier, the Lower Susquehanna does not meet the criteria of a catchment, although subdividing it into northern (E) and southern (G) areas addresses that.

These mainstem and drainage basin relationships form organizing principles of surface water at a broad scale (Blodgett, et al, 2020). In an operational context, CHyF recognizes the value in having immutable identifiers for such basins and mainstems. The hydro nexuses associated with them, and implemented with point geometry, are good candidates to be considered as contracted hydronodes, i.e., major confluence points recognized as permanent features.

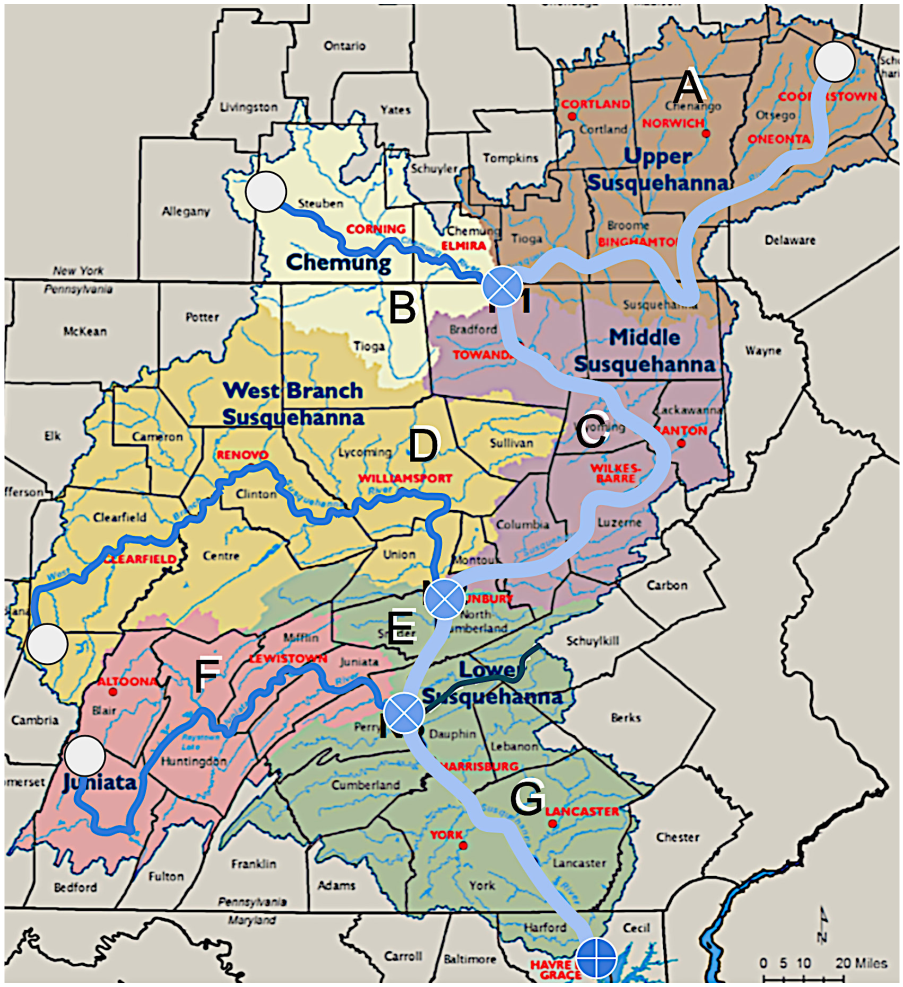

Figure 43: Susquehanna Basin with four contained basins (A, B, D, F), all with headwaters

9.2 Dynamic Hierarchical Catchment Framework

CHyF only requires that elementary flowpaths and catchments form a DAG, not that they be hierarchically organized. Nevertheless, a large number of applications can benefit from having a hierarchical framework available. Such a framework is defined by the HUC in the United States NHD, as well as by a comparable system defined in British Columbia. Many applications make use of such codes. In many cases they allow for upstream/downstream relationships to be determined without looking at the feature geometry. In the US the code also allows for summaries to be made at six different levels of detail, ranging from from very general (e.g., Pacific Northwest) to comparatively detailed (e.g., North Fork Imnaha River) (examples from Wikipedia).

CHyF includes *rank*, an attribute of both elementary flowpaths and elementary catchments with a value of primary or secondary. The primary flowpaths and catchments form dendritic networks. Consequently, it would be possible using CHyF to create a similar structure to the HUC system in the US. The relationships between features could be embedded directly in a code, as with the HUC, or could be made available through a table indicating by feature identifier what flows into what. Creating a six level hierarchical system would be possible, but it may be more detailed than required by most users.

One option would be to define a system with only three or four levels of depth and below that to consider another technique. This would involve designating some hydro nexuses as contracted, i.e., some hydronodes would be considered as of particular importance and worth defining as persistent. Another option would be to allow users to define such points for their particular purposes. In either case, upstream flowpaths and catchment networks could be calculated on the fly as necessary from these hydronodes, thus making it unnecessary to maintain them as features in their own right.

In Canada the starting point for a hierarchical catchment framework should be the sub-sub-drainages established by Environment and Climate Change Canada (ECCC) and further refined as working units by Natural Resources Canada (NRCan). This initial work would include the creation of flow relationship tables comparable in content to those shown in section 6.1.9 above. Such tables could be defined at different levels of the hierarchy, all the way down to elementary catchments. Hierarchical codes could be generated from these tables and used to populate a *hierarchicalCode* attribute if there were sufficient user interest.

Because new mapping occurs from time to time, the flow relationship tables will need to be updated to reflect the new population of features in the newly mapped area. Within a given drainage area, some features with their identifiers may be retired, whereas others may remain valid. This mix will need to be addressed if Linked Open Data or Internet of Water infrastructures are of interest. Overall, the hierarchical framework must be considered as dynamic, subject to change as updating occurs.

9.3 Drainage Areas Defining National Coverages

ECCC has specified approximately 1400 sub-sub-drainage areas that cover Canada. These have been refined along the boundaries by NRCan, with the results called working units or basins with watershed boundaries. In keeping with the terminology used here, in HY_Features, and in Blodgett, et al (2020), these units are best referred to as drainage areas, each with an identifier and name. This is also in keeping with ECCC usage. Note that some of these meet the criteria for a drainage basin or a catchment aggregate, but some do not. For example, coastal drainage areas exist that consist of many comparatively small catchments draining directly into the ocean. In the United States a similar situation exists with the Hydrologic Unit system. This hierarchical system has about 2300 units defined as subbasins covering the U.S. However, as with the Canadian situation, and for the same reasons, these are best referred to as drainage areas. In both countries these units are defined to support management and some kinds of analysis, but as a single entity they do not meet the criteria for a catchment

10. Temporal Considerations

10.1 Hygraphs and Time Series

Although outside of this release of the CHyF model and services, the hygraph concept can be applied to dynamic behaviour in different ways. Assume that three snapshots of a given area exist. These different versions may be based on set time intervals, given events, or simply the availability of applicable data. For each version, a hygraph U can be calculated as described above. This situation is depicted in Figure 44.

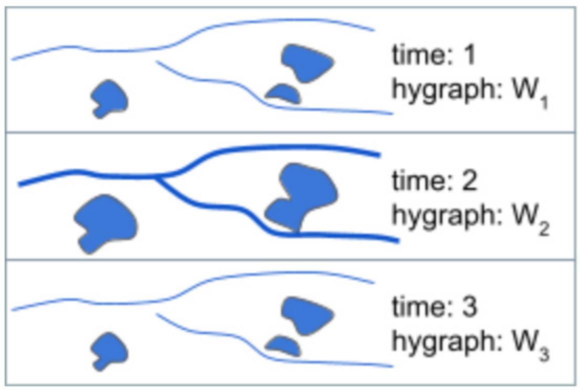

Figure 44: Steps in a time series, each modelled by a hygraph

For the area in question, a hygraph time series can be established for time t1, t2, t3, etc.:

$$T \leftarrow \{U_{t1}, U_{t2}, U_{t3}, \ldots\} \qquad [\text{Eq } 8]$$

For example, t1, t2, t3, and t4 could refer to the years 1980, 2000, 2020, and 2040, or to any particular times of interest. The underlying assumption here is that a different hygraph, U, applies to the entire area for each time.

10.2 Hygraphs and Recurring Floods (experimental)

The different times may represent different return periods. For example, the 5 year, 10 year and 100 year return periods can be shown as a hygraph series as follows:

$$R \leftarrow \{U_{r5}, U_{r10}, U_{r100}\} \qquad [\text{Eq } 9]$$

This approach assumes that separate, unrelated hygraphs are modelled for each return period. An alternative approach that integrates the geography of the recurring floods into a single hygraph is as follows.

Within the universe U there exists an arbitrary number of non-overlapping Maximum Areas of Potential Flooding (MAPFs) for which floodplain analyses have been performed. A particular MAPF refers to a given location subject to flooding, such as a specific town or agricultural area. It encompasses all observed or estimated flooded terrain for that particular location. The largest of these areas, which might for example correspond to the 200 year return period flood, is used as follows. The MAPF is the merged area of all elementary catchments that are completely or partially underwater, based on that largest flood.

Let a given $MAPF_i$ be represented by the subgraph S_i of U. Multiple versions of the subgraph can be defined, one for each value of x, where x represents the x year return period flood. The general case is denoted as $S_{i(rx)}$, with $S_{i(r0)}$ being the subgraph for the not-flooded state. For the $MAPF_i$ a collection of subgraphs can be defined, each of which represents the entire $MAPF_i$ under a different flooding extent.

$$C_i \leftarrow \{S_{i(ro)}, S_{i(rx1)}, S_{i(rx2)}, \ldots\} \quad \text{[Eq 10]}$$

C_i is a collection of subgraph options for the area within the APF_i. For the area outside of the $MAPF_i$ a subgraph is necessary, referred to as $S_{!i}$, i.e., $S_{\text{not } i}$. The entire graph, U, can then be defined as the subgraph for the area outside of the $MAPF_i$ combined with the subgraph for the area inside the $MAPF_i$, with the latter a member of C_i. In the expressions below, the + sign means combined into a common graph. $U_{i(rx)}$ is the entire graph taking into account flooding in a given location, with the maximum extent of the flooding defined by $MAPF_i$, and with a flood recurrence interval of rx.

$$U_{i(rx)} \leftarrow \{S_{!i} + S_{i(rx)}\} \text{ as a general statement, and} \quad \text{[Eq 11]}$$

$$U_{i(ro)} \leftarrow \{S_{!i} + S_{i(ro)}\} \text{ for the not-flooded state} \quad \text{[Eq 12]}$$

The representations of the MAPF, associated with r0 and the other versions of rx, share a common outer boundary that follows catchment divide segments. All hydronodes on the boundary of $MAPF_i$ must be shared by the subgraphs for the $MAPF_i$ for the different return periods ($S_{i(r0)}$, $S_{i(rx1)}$, etc.) and also by the subgraph for the area outside of the MAPF, ($S_{!i}$). In other words, the connections and boundary between the area inside of $MAPF_i$ and the area outside of $MAPF_i$ must be identical for all $S_{i(rx)}$. Flow analysis can be performed on the overall U with a given MAPF at a specified level of potential flooding by swapping out $S_{i(r0)}$ for $S_{i(rx)}$ in the graph. Different versions of $S_{i(rx)}$ with the associated geography can be retained, allowing for $U_{i(rx)}$ to be reconstituted as required. Figure 45 shows a simple case.

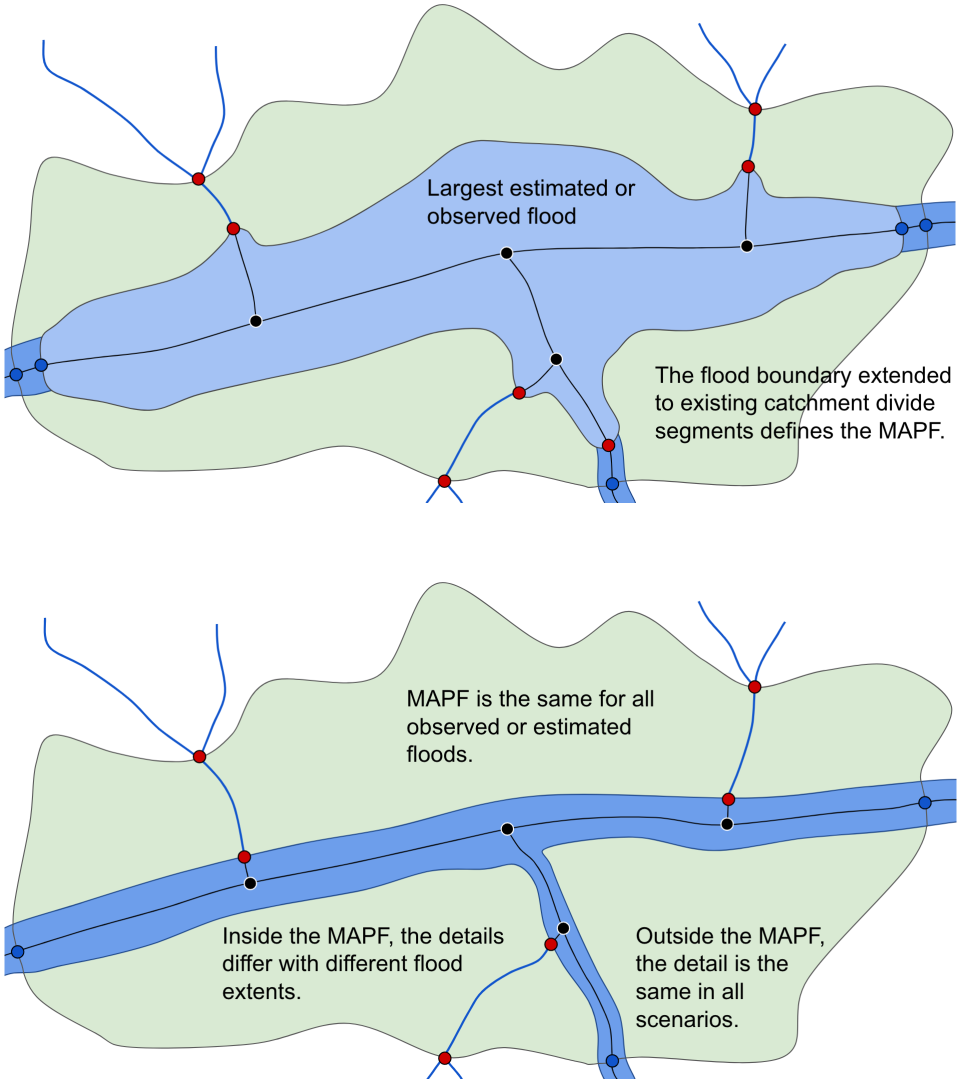

Figure 45: Maximum Area of Potential Flooding and different flood stages

Given that the MAPFs do not overlap, this approach can support an arbitrary number of simultaneous flood scenarios. The expression above can be extended to the following.

$$U_{i1(rx1),i2(rx2),...} \leftarrow \{S_{!i1,!i2,...} + S_{i1(rx1)} + S_{i2(rx2)} + ...\} \qquad [Eq\ 13]$$

The first S in brackets refers to the subgraph for the area outside of all MAPFs. The subsequent subgraphs refer to the areas $MAPF_{i1}$, $MAPF_{i2}$, etc. The return periods rx1, rx2, etc. would likely all be the same, but they could be different if desired.

10.3 Graphs and Coalescing Waterbodies (experimental)

Another way of considering time relates to the likelihood of two nearby waterbodies coalescing over a given time period. Figure 46 shows a number of lakes in dark blue and several streams in light blue. The red line segments show connections between nearby waterbodies. With sufficient data or modelling results, return periods could be estimated for each red segment, indicating the frequency of the two waterbodies merging. The proximity network formed by the red line segments can be represented by a graph separate from the hygraph. Using the two together could allow for powerful analyses.

Figure 46: Graph structure indicating potential coalescing waterbodies