Reference Model Supporting Documentation For CABIN Analytical Tools

MODEL NAME: Okanagan 2025

AUTHORS: Trefor B. Reynoldson, Sheena Pappas, Brennan Williams and Stephanie Strachan

DATE: March 2026

This document is based on the following technical report: Reynoldson, T.B. 2025. An Update of the Okanagan CABIN RCA Model for British Columbia (2006-2023). 76pp + appendices. Report submitted to Environment and Climate Change Canada March 2025.

In the CABIN database and analytical tools, reference groups are labelled 1, 2, 3, 4. These group numbers correspond to reference groups 1Aia, 1Aib, 1B, and 2 in this document.

1. CONTEXT

The Canadian Aquatic Biomonitoring Network (CABIN) applies the Reference Condition Approach (RCA) to assess stream health (Reynoldson et al., 1997). The RCA requires a large dataset of minimally disturbed sites to define the relationship between the habitat characteristics and the benthic macroinvertebrate communities for a given region.

The purpose of a reference condition model is to predict the macroinvertebrate community expected at a test site in the absence of a disturbance. Models do not determine cause of divergence from reference; therefore, habitat variables that are likely to have a causal relationship with the benthic invertebrate communities are excluded from the RCA model.

The purpose of this document is to provide a summary of the reference data used in the model and key information about the model development. It is intended for CABIN participants using CABIN analytical tools and aid in the interpretation of CABIN assessments.

2. STUDY DESIGN AND SITE SELECTION

2.1 Model Purpose

The Okanagan region, located in southern British Columbia is dominated by orchards, vineyards, rangeland, pasture, logging, urban and recreational development influencing stream and river environments. Basin concerns in the Okanagan include nutrient enrichment, sedimentation, erosion, stormwater runoff, invasive species and water quantity.

The Okanagan 2025 model is a revision of the Okanagan 2017 Model (Strachan and Bennet, 2018) and improvement on earlier modelling efforts. This revised model was developed following model building guidance provided as part of a suite of CABIN program resources found here: https://www.canada.ca/en/environment-climate-change/services/canadian-aquatic-biomonitoring-network.html.

The Okanagan 2025 model improves upon the previous model by extending spatial and temporal coverage, reexamining previously identified outlier sites, and providing an evaluation of Type II error rates for the model based on simulated dataset representing organic enrichment disturbance.

This revised Okanagan model enables CABIN participants in the Okanagan to compare test site data to a reference model for which the characteristics of Okanagan biological communities and habitat types have been specifically modeled. CABIN participants in the Okanagan study area can evaluate stream health to support informed decision-making and management actions. Results generated from this tool by CABIN participants will expand the science and knowledge of health, productivity and threats to aquatic ecosystems in the Okanagan watersheds.

2.2 Spatial and Temporal Scope

The Okanagan model includes three major sub basins: Similkameen, Okanagan, and Kettle river basins. It also includes the small Vernon Diversion drainage area. From west to east, this area extends from Manning Provincial Park to Granby and Gladstone Provincial Parks and from north to south it extends from Armstrong to Osoyoos. This area includes five ecoregions: Interior Transition Ranges, Okanagan Highlands, Okanagan Range, Selkirk-Bitterroot Foothills, and the Thompson-Okanagan Plateau. The Thompson-Okanagan Plateau comprises the largest area in this region. The Thompson–Okanagan Plateau has the highest number of reference sites (Table 1). A range of stream orders was sampled to ensure a variety of stream sizes was included. Third-order streams are the dominant reference site type (Table 1).

Table 1. Number of reference sites sampled by stream order and ecoregion.

	Stream Order						
Ecoregion	**1**	**2**	**3**	**4**	**5**	**6**	**Total**
Interior Transition Ranges		2	1				3
Okanagan Highlands			1				1
Okanagan Range		4	27	20	1		52
Selkirk-Bitterroot Foothills			9	1			10
Thompson-Okanagan Plateau	26	42	28	6	1	2	105
Total	**26**	**48**	**66**	**27**	**2**	**2**	**171**

The complete reference dataset includes samples spanning a 19-year period, from 2005 to 2023, extending the previous model by seven years. All samples were collected in the late summer to fall period (late August to mid-October) by CABIN certified field technicians. This model includes 171 samples, of which 83 are new reference samples. There are 73 unique sites, 28 of which are repeat reference sites across the study area's three sub-basins (Table 2).

Table 2. Number of reference site samples collected each year by sub-basin. Unique sites are indicated in parentheses.

	Sub-basin			
Year	Kettle	Okanagan	Similkameen	Total
2005		4 (4)		**4 (4)**
2006	1 (1)	14 (13)	3 (3)	**18 (17)**
2007	4 (4)	4 (1)	7 (4)	**15 (9)**
2008		2 (0)		**2 (0)**
2009	3 (0)	5 (2)	1 (0)	**9 (2)**
2010		5 (2)	2 (0)	**7 (2)**
2011		5 (2)		**5 (2)**
2012	2 (0)	4 (2)	1 (0)	**7 (2)**
2013		2 (0)	5 (1)	**7 (1)**
2014	9 (9)	2 (0)	5 (3)	**16 (12)**
2015		5 (2)	16 (14)	**21 (16)**
2016	2 (0)	5 (0)	3 (1)	**10 (1)**
2017	1 (0)	5 (0)	3 (0)	**9 (0)**
2018		6 (0)	2 (0)	**8 (0)**
2019	2 (0)	2 (0)	2 (0)	**6 (0)**
2020		3 (0)	4 (0)	**7 (0)**
2021		2 (0)	5(3)	**7 (3)**
2022	2 (0)	2 (0)	3 (0)	**7 (0)**
2023		4 (2)	2 (0)	**6 (2)**
Total	**26**	**81**	**64**	**171 (73)**

Fifteen sites were sampled on three or more occasions and three sites on more than ten occasions (Table 3) to capture temporal variation in local climate or hydrology affecting the reference communities. Refer to Table 3 for the list of repeated reference sites. This dataset provides a balance of spatial variation from all three watersheds, over five ecoregions and temporal variation over 19 years.

Table 3. Repeated reference sites sampled between 2005 and 2023 within each sub-basin.

Site	'05	'06	'07	'08	'09	'10	'11	'12	'13	'14	'15	'16	'17	'18	'19	'20	'21	'22	'23	Total
Similkameen sub-basin: 11 repeat reference sites																				
ASH01		1	1			1				1			1			1				**4**
ASH02		1	1		1				1		1		1			1				**5**
BUC01			1					1	1	1	1	1	1	1	1	1	1	1	1	**6**
COP01			1											1						**2**
COP02												1			1					**2**
GDF01			1						1										1	**3**
KER05		1	1			1			1											**4**
SIM03										1									1	**2**
SIM05										1		1				1				**3**
SIM06																	1		1	**2**
TUL05											1						1			**2**
Okanagan sub-basin: 13 repeat reference sites																				
BX01		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**18**
CHT01		1	1									1								**3**
ELL02	1	1				1						1								**4**
EQU01	1						1													**2**
EQU02		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**18**
HYD01								1						1						**2**
HYD02								1						1						**2**
MCL01						1						1								**2**
PEA01		1			1								1							**2**
ROB01			1											1						**2**
TRE01					1									1		1				**3**
WHI02		1					1						1							**3**
WHI04							1						1							**2**
Kettle sub-basin: 4 repeat reference sites																				
GRA01			1		1			1				1			1			1		**6**
GRA02			1		1			1				1			1			1		**6**
SND01		1											1							**2**
WKET01			1		1															**2**
Total	**2**	**10**	**13**	**2**	**8**	**6**	**5**	**7**	**6**	**6**	**5**	**10**	**9**	**8**	**6**	**7**	**5**	**7**	**4**	**126**

2.3 Reference Site Selection

Expanded spatial and temporal coverage was achieved through collective sampling efforts by Environment and Climate Change Canada (ECCC) and British Columbia Ministry of Environment (ENV). Samples were collected following standard CABIN methods as outlined in Environment Canada (2012). New potential reference sites were distributed across the three sub basins and ecoregions, apart from no additional potential reference sites in the Selkirk-Bitterroot ecoregion (Figure 1).

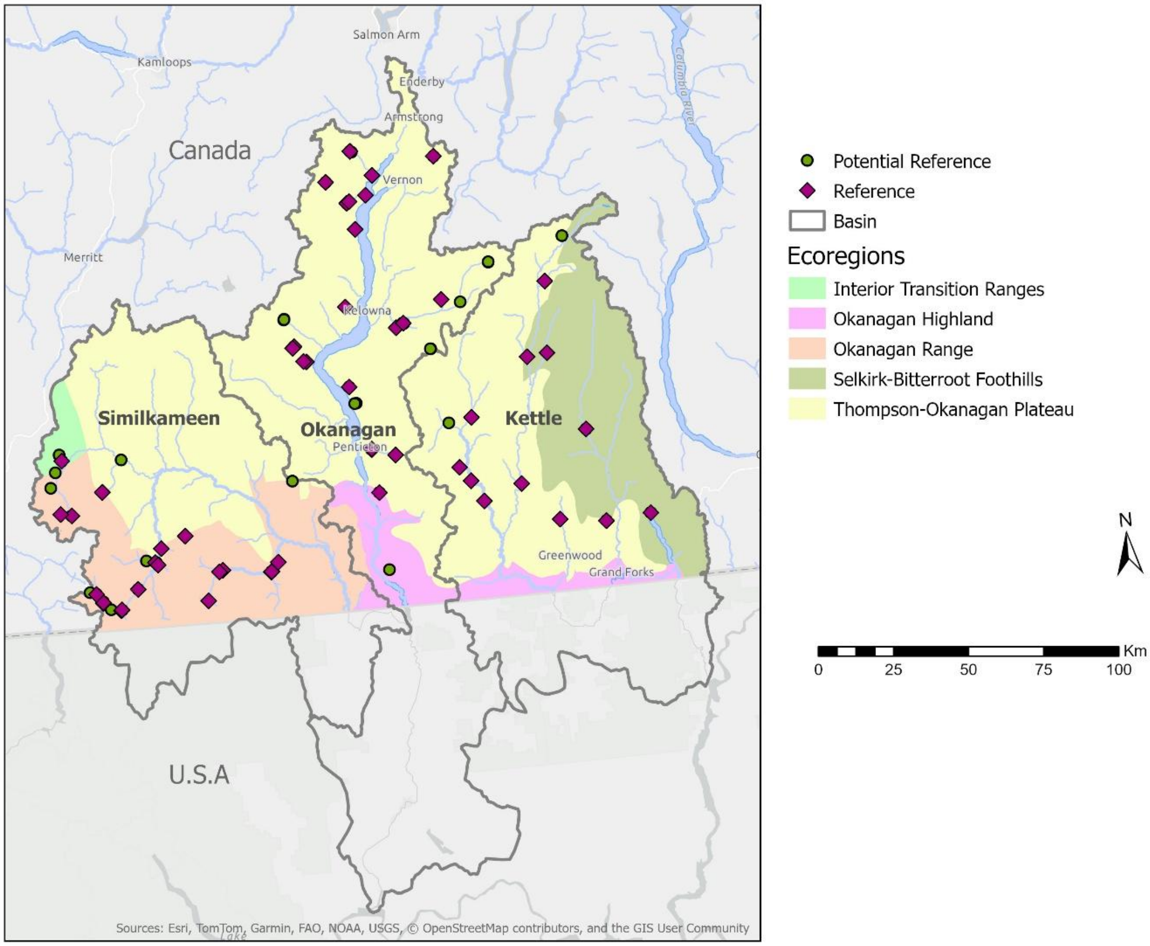

Figure 1. Distribution of reference and potential reference site locations for Okanagan model sub basins. Sites represented by a diamond are from the Okanagan 2017 model (Strachan and Bennett 2018) and potential (new) reference sites are represented by a circle.

The reference sites included the earlier versions of the Columbia-Okanagan model were selected by best professional judgement. The data gaps were filled for the Okanagan 2017 model after applying the Human Activity Gradient (HAG, Yates and Bailey 2010). The following paragraphs summarize the six steps of the HAG approach applied to the Okanagan study area. A fuller description of this approach is described in the Okanagan 2017 Reference Model Supporting Documentation (Strachan and Bennet, 2018).

Step 1: The Okanagan study area was subdivided into microbasins of 3rd order and higher using 1:50K BC Freshwater Atlas delineations from BC Ministry of Environment and Climate Change DataBC (https://www2.gov.bc.ca/gov/content/data/about-data-management/databc).

Step 2: Microbasins were grouped into similar landscape types based on natural habitat characteristics such as bedrock geology, landcover and climate. More than 400 microbasins were identified in the study area. Landscape types in the Okanagan study area were classified into five natural landscape groups (Figure 2). The natural habitat characteristics of these landscape groups are described in Table 4.

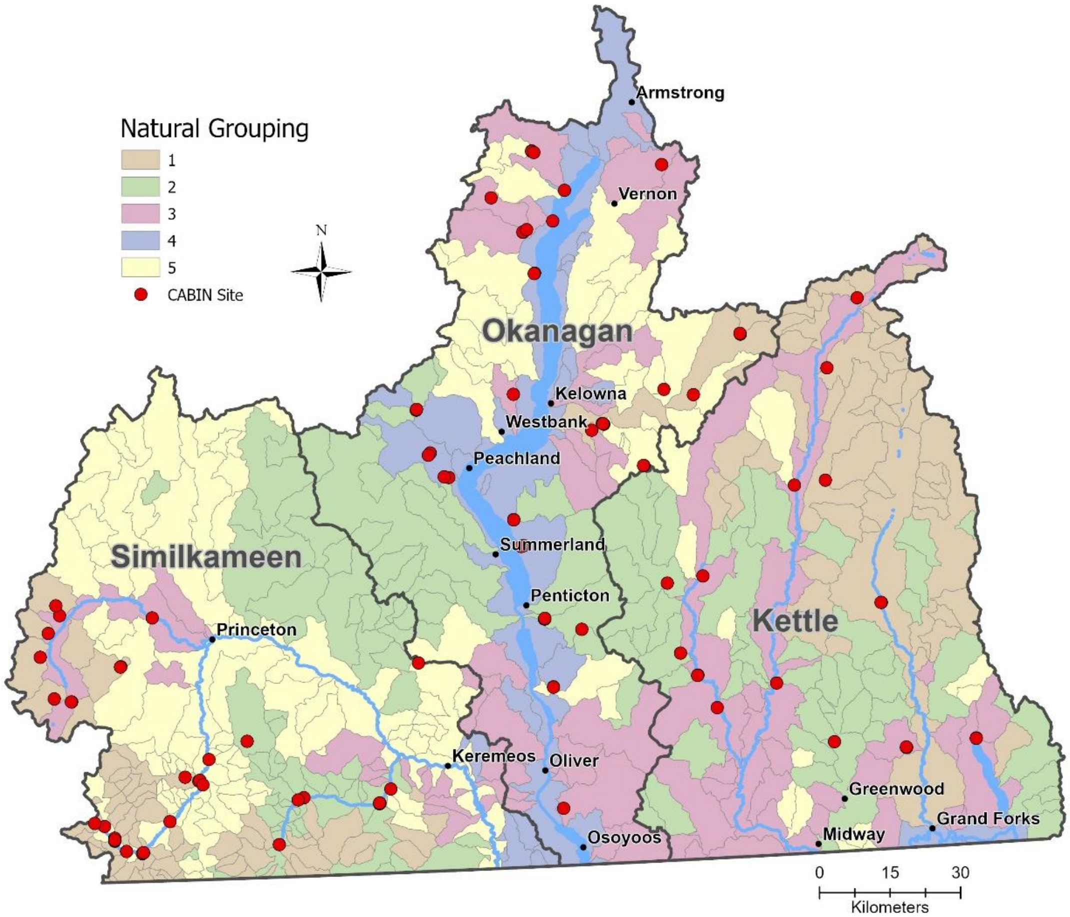

Figure 2. Geographical distribution of reference sites in the Okanagan, Similkameen, and Kettle basins distributed across five natural landscape groups.

Step 3: Potential reference microbasins were objectively identified by establishing a stressor gradient for each natural grouping and defining the best available reference sites within each natural grouping (Yates and Bailey 2010). Natural habitat characteristics and stressors in the Okanagan study area are described in Table 4.

Step 4: Unique stressor gradients were developed for each natural group to establish reference site criteria for each landscape group (Table 5). As a result of the different natural features in each group, land use is different in each natural group (i.e. logging vs agriculture vs rangeland vs development). The significant presence forestry related landcover stressors (i.e. active cutblock) and developed or urban landcover eliminated some microbasins from the list of potential reference sites. Rangeland occurs throughout most of the Okanagan. To capture the best available microbasins in the smaller natural groups 3 and 4, the reference site criterion for rangeland is relatively high. Similarly, Group 3 landscape type tends to be where communities are centered and the criteria for road density, developed and urban landcover is also relatively high.

Table 4. Median of natural habitat and stressor characteristics of microbasins in each of the five natural landscape groups (NL Gp) derived from National Land Cover data (refer to Natural Resources Canada (2009) for stressor definitions).

Natural Landscape Groups	NL Gp1 (n=109)	NL Gp2 (n=113)	NL Gp3 (n=71)	NL Gp4 (n=24)	NL Gp5 (n=111)
Natural habitat characteristic					
Basin size	18.0 km^2	37.2 km^2	29.4 km^2	17.0 km^2	26.7 km^2
Low Shrub landcover	2.8%	7.4%	6.9%	7.6%	9.1%
Herb landcover	6.1%	7.7%	11.3%	18.1%	7.7%
Grassland landcover	0%	0%	1.4%	6.2%	0%
Intrusive bedrock	52.5%	94.2%	13.0%	23.5%	4.8%
Metamorphic bedrock	0%	0%	26.0%	5.6%	1.1%
Volcanic bedrock	0%	0.4%	4.1%	8.9%	47.1%
Annual mean precipitation	956 mm	638 mm	525 mm	379 mm	666 mm
Annual mean temperature	1 ^oC	2 ^oC	4 ^oC	7 ^oC	2 ^oC
Stressor characteristic					
Road Density	0 km/km^2	0.003 km/km^2	0.249 km/km^2	0.569 km/km^2	0.0001km/km^2
Developed	0%	0%	0.01%	0.12%	0%
Urban	0%	0%	0%	0%	0%
Agricultural	0%	0%	1.45%	6.08%	0%
Rangeland	0%	0%	6.6%	35.88%	0%
Young Forest	30.87%	56.65%	46.83%	33.88%	44.86%
Recently Logged	0%	10.05%	3.49%	0%	12.32%
Active cut block	0.61%	5.35%	0.94%	0%	7.14%

Table 5. Reference criteria in each natural landscape group (NL Gp) for the Okanagan study area.

Reference Criteria	NL Gp 1	NL Gp 2	NL Gp 3	NL Gp 4	NL Gp 5
Road Density	0 km/km^2	0 km/km^2	0.030 km/km^2	0.168 km/km^2	0 km/km^2
Developed	0%	0.1%	1.5%	2%	0.1%
Urban	0%	0%	2%	0%	0%
Agricultural	0.2%	0.3%	15%	22%	0.5%
Rangeland	7%	0.6%	38%	62%	8.7%
Recently Logged	13%	32%	13%	0.1%	19%
Active cut block	5.5%	23%	13%	0%	19%

Step 5: A subset of potential reference microbasins identified through this desktop site selection exercise was verified through field reconnaissance by ENV and ECCC.

Step 6: Final reference site selection was determined by habitat suitability (i.e., presence of a riffle), safe access, and visual confirmation that surrounding land use was not affected by human disturbance.

3. REFERENCE DATA

Reference and potential reference data used to develop the final model was downloaded from the CABIN database from two studies: BC MOE - Okanagan Kick net and EC-Okanagan Basin. The complete dataset of 171 samples spanning the 19-year period were analyzed. Sites with incomplete data were removed prior to further analysis.

3.1 Biological description of the Okanagan dataset

The dataset was comprised of 69 families of invertebrates to be used for subsequent analyses. Three families were removed from further analysis as they were either not strictly aquatic fauna or were excluded according to laboratory and sample processing requirements for CABIN which included Lumbricidae, Planariidae, and Sparganophilidae. Refer to Environment Canada 2014 for further details on the taxa excluded from CABIN sample identifications. A fourth family, Momoniidae, was removed as it was a single unverified specimen at the time of this analysis.

Classification of family level reference data was performed using the Bray-Curtis similarity measure in PATN (Belbin, 1993). Non-metric multi-dimensional scaling (NMDS) analysis was used in PRIMER 6 (Clarke and Gorley, 2015) ordinations to explore biological patterns in the data based on the group structure generated in the initial clustering analysis. The cluster analysis produced a hierarchical structure with several groups labeled as shown in Figure 3. A four-group classification was determined to be appropriate for the final model: Group 1Aia, 1Aib, 1B, and 2. These group labels correspond to Groups 1,2,3 and 4, respectively, in the CABIN analytical tools. Eight outliers were identified in group 1Aii.

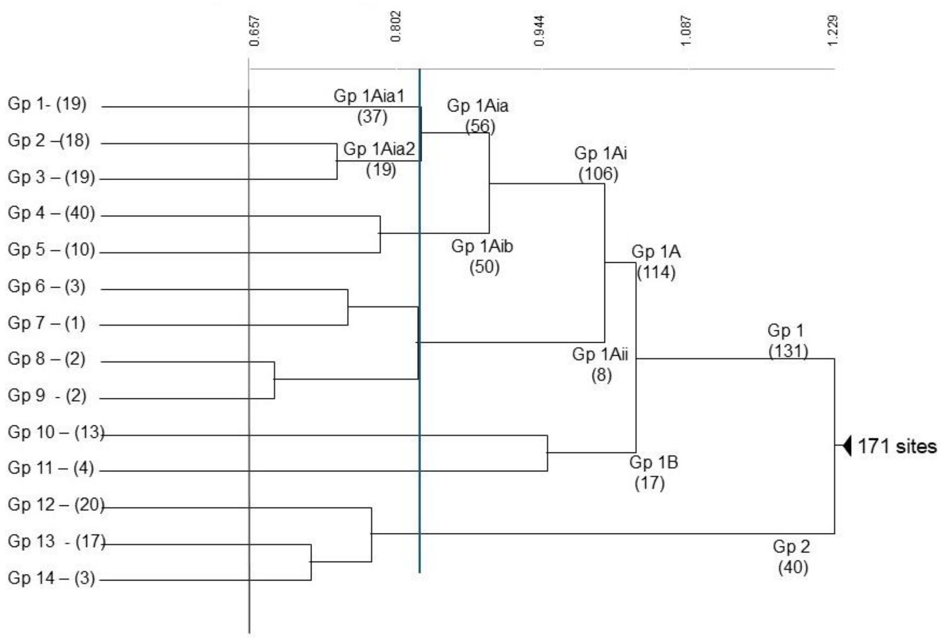

Figure 3. Dendrogram from PATN cluster classification of 171 reference and potential reference sites in the Okanagan. The four-group classification is identified as follows: group 1Aia, 1Aib, 1B, and 2. Eight outliers are identified in group 1Aii. The number of sites on each branch are shown in parentheses. Group 1Aia, 1Aib, 1B, and 2 labels correspond to Groups 1,2,3 and 4, respectively, in the CABIN analytical tools.

Ordination of the biological data is illustrated in Figure 4 based on Bray-Curtis similarities. Similar to Figure 3, Group 1Aia, 1Aib, 1B, and 2 labels correspond to Groups 1,2,3 and 4, respectively, in the CABIN analytical tools. Eight outliers were identified in Group 1Aii.

There is overlap in the reference group sites in the center of the reference cloud. The biological communities from eight sites identified as outliers (Group 1Aii, green triangles) are located outside the reference cloud along the left and upper sides of the ordination plot (Figure 4) and removed from further analysis. A fuller description (descriptive community metrics) of the groups is summarized with the description of the final model in Section 4.

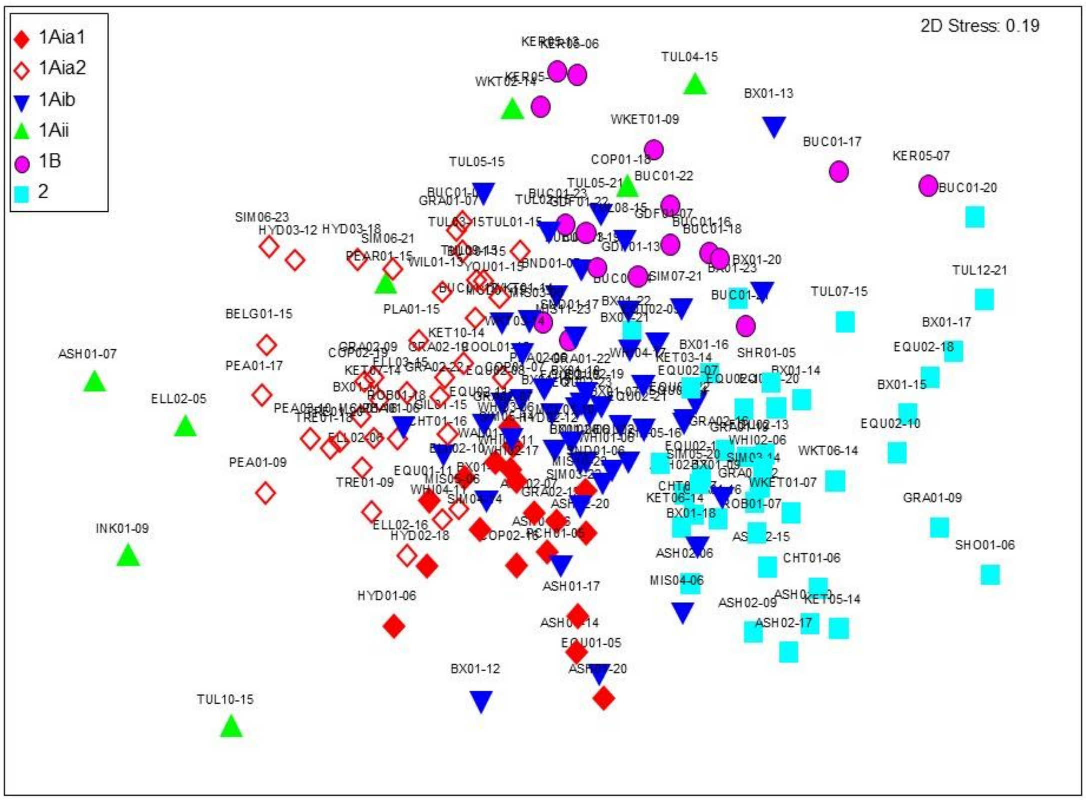

Figure 4. NMDS ordination of 171 Okanagan reference/potential reference sites showing six groups. The four-group classification is developed from this and identified as group 1Aia (displayed separately as 1Aia2 and 1Aia2) (diamonds), 1Aib (upside down triangles), 1B (circles), and 2 (squares). Eight outliers are identified in group 1Aii (triangles). Group 1Aia, 1Aib, 1B, and 2 labels correspond to Groups 1,2,3 and 4, respectively, in the CABIN analytical tools.

Figure 5 illustrates the median abundance of the most common taxa found in each reference group. While the common families present in each group are similar the abundance of the families differs among the groups.

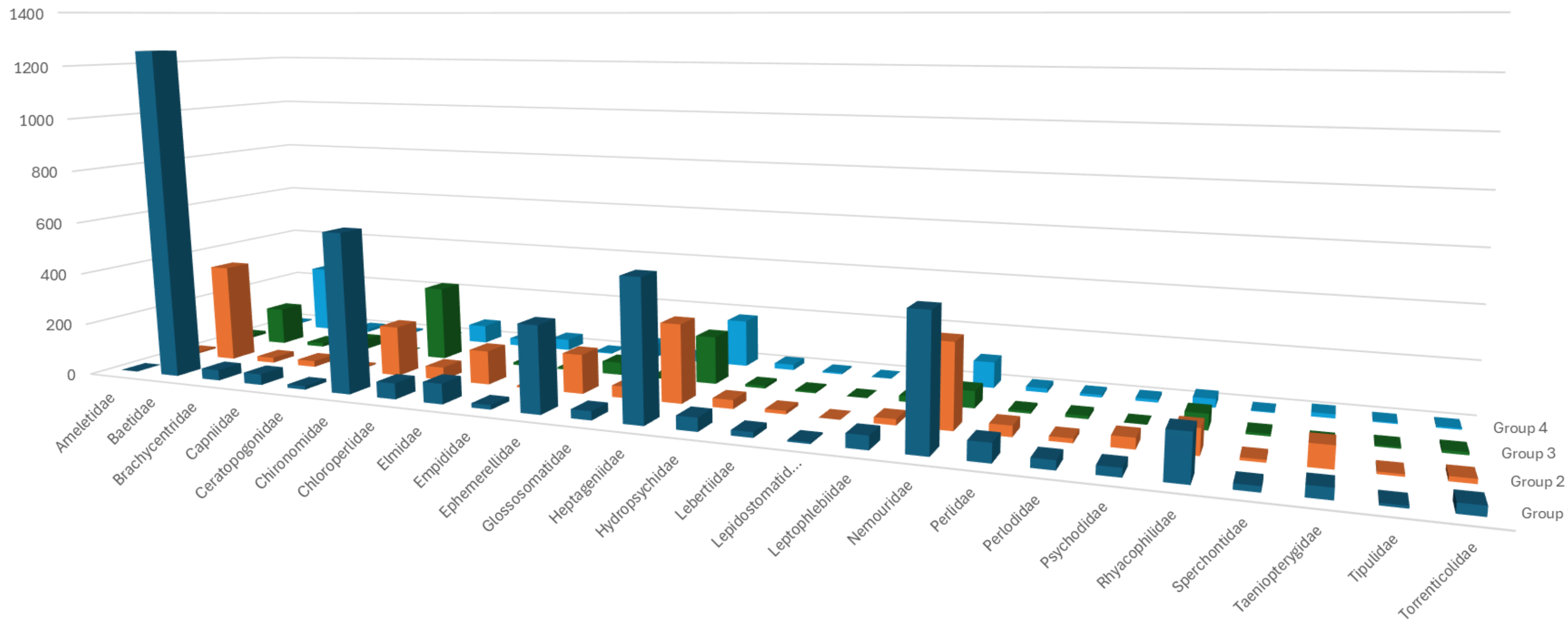

Figure 5. Median family abundance of each reference group for the common taxa.

The biological description of each of the reference groups is described below.

- Group 1 is the largest reference group. The highest abundance value of these sites is dominated by four taxa: the Baetidae mayflies, Chironomidae, Heptageniidae mayflies and Nemouridae stoneflies. It is the high numbers of baetid mayflies that best characterize this group.

- Group 2 is the second largest. This group has the second highest abundance and highest diversity. Although the community is characterized by the same families as Group 1, the Baetidae mayflies are less dominant, and the most abundant family is the Taeniopterygidae.

- Group 3 is the smallest group of reference sites. This group is dominated by chironomids and nemourid stoneflies. It is also characterized by the presence of Capniidae stoneflies, Lebertidae mites and Ameletidae mayflies that are characteristic of small fast flowing mountain streams.

- Group 4 is the second smallest group of reference sites and has the lowest abundance. Small Baetidae and Heptageniidae mayflies are most characteristic of this community, with almost a third of the organisms belonging to these families. Other community metrics are similar to the other groups (Refer to Section 4, Table 10).

3.2 Habitat Description of the Okanagan dataset

Habitat characteristics for reference sites were described for the following categories: location, basin morphology/hydrology, topography, climate, landcover, bedrock geology, stream channel, and water chemistry. The geographical description of the reference dataset is described below and in Table 6, with a detailed summary in the description of the final model in Section 4.

- Group 1 is the largest reference group where sites are generally located on deeper wider streams located further downstream (of higher stream order) with larger drainage basins. These sites are evenly dispersed in the Okanagan and Similkameen sub basins and the majority are in the Thompson Okanagan Plateau ecoregion.

- Group 2 is the second largest. Sites are primarily located in the Okanagan sub basin and Thompson Okanagan Plateau ecoregion. Sites in this reference group tend to be smaller shallower streams with more canopy coverage. Sites tend to experience drier and warmer conditions.

- Group 3 is the smallest group which tends to be the highest altitude sites with the smallest drainage basins and the shortest stream length (i.e. headwater streams). Sites tend to experience wetter and cooler conditions. The sites are almost entirely in the Similkameen sub basin and distributed evenly between the Thompson Okanagan Plateau and Okanagan Ranges ecoregions.

- Group 4 sites are equally distributed among the three basins (Kettle, Okanagan and Similkameen) and the sites are largely located in the Thompson Okanagan Plateau ecoregion.

Table 6. Maximum and minimum values for the full set of potential predictor variables examined for the Okanagan Model.

Potential Predictor Variable	Minimum	Maximum
Latitude	49.0530	50.3662
Longitude	-121.014	-118.29
Altitude	326	1706
Stream Order	1	6
Drainage Area (km^2)	2.2452	952.3762
Stream Density (m/km^2)	960.4851	3478.731
Stream Length (m)	3291.77	1802266
Elevation-maximum watershed elevation (m)	1475	2614
Elevation-minimum watershed elevation (m)	358	1684
Gradient-watershed area with hillslope 30-50% (%)	1.8151	43.6327
Gradient-watrshed area with hillslope 50-60% (%)	0.0822	14.3745
Gradient-average watershed hillslope (%)	13.1826	55.3082
Gradient-watershed area with hillslope > 60% (%)	0.0776	43.3487
Slope < 30% (%)	17.2992	98.0251
Gradient-maximum watershed hillslope (%)	69.57	435.61
Gradient-minimum watershed hillslope Min. (%)	0	1.1866
Precip.-March 30 yr average (mm)	32.8361	125.89
Precip.-August 30 yr average (mm)	22.4	63.7033
Precip.-November 30 yr average (mm)	48.7343	181.93
Precip.-Total ANNUAL 30 yr average (mm)	513.49	1216.76
Temp.-January Min 30 yr average (°C)	-11.5681	-6.94
Temp.-February Max 30 yr average (°C)	-3.3433	1.16
Temp.-August Max 30 yr average (°C)	14.5386	24.1
Temp.-November Min 30 yr average (°C)	-7.5086	-4.33
Temp.-ANNUAL max 30 yr average (°C)	4.0114	10.55
Temp.-ANNUAL mean 30 yr average (°C)	0.3919	5.24

Potential Predictor Variable	Minimum	Maximum
Avg Channel Depth (cm)	3.4	62.4
Max Channel Depth (cm)	5	78
Pools in Reach (Binary)	0	4
Rapid in Reach (Binary)	1	4
Riffle in Reach (Binary)	0	1
Straight Run in Reach (Binary)	0	1
Slope (m/m)	0	1
Avg Velocity (m/s)	0	1
Max Velocity (m/s)	0	0.448
Bank Full Width (m)	0	1
Wetted Width (m)	0	1
% Canopy Coverage (Cat 0-4)	0	1
Dominant Streamside Vegetation (Cat 0-4)	0	1
Presence of Coniferous Trees (Binary)	0.05	1
Presence of Deciduous Trees (Binary)	0.12	1.79
Presence of Grasses (Binary)	1.7	34.9
Presence of Shrubs (Binary)	1.1	21
% Bedrock (%)	0	6
% Boulder (%)	0	44
% Cobble (%)	4	83
% Gravel (%)	0	20
% Pebble (%)	4	69
% Sand (%)	0	17
% Silt+Clay (%)	0	8
Median Particle Size D50 (cm)	2.2	23
Dom. Substrate (Cat 0-9)	4	8

Table 6. (Con't)

Potential Predictor Variable	Minimum	Maximum
Temp.-ANNUAL min 30 yr average (°C)	-3.2276	-0.08
Temp.-ANNUAL min – coldest months 30 yr average(°C)	-11.8625	-7.56
Temp.-ANNUAL max – warmest months 30 yr average (°C)	15.3091	25.17
Intrusive (%)	0	100
Metamorphic (%)	0	100
Sedimentary (%)	0	100
Ultramafic (%)	0	12.42
Volcanic (%)	0	100
Landcover -Snow/Ice (%)	0	0.53
Landcover-Water (%)	0	4.04

Potential Predictor Variable	Minimum	Maximum
2nd Dom. Substrate (Cat 0-9)	4	8
Embeddedness (Cat 1-5)	1	5
Surrounding Material (Cat 0-9)	1	5

The habitat variable list was initially reduced to 99 variables for model development based on data completeness, other identified issues, and whether a variable may be influenced by human activity. This list consisted of 69 Geographic Information System (GIS) derived and 30 site observation variables.

The list of potential predictor variables was reduced by excluding variables if they were missing data, were expected to vary with anthropogenic disturbance (i.e. vegetation landcover, water chemistry), were highly correlated with another variable (i.e. climate variables), or had no known relationship with invertebrate communities (i.e. statistical variables such as "Elevation standard deviation in the upstream watershed"). The reduced list consisted of 65 possible predictor variables (Table 6).

4. THE MODEL

The full reference dataset was divided into a training dataset to optimize the model and a validation dataset to evaluate model performance (Table 7). From the full reference dataset, 30% of sites were randomly selected from the small group (<50 sites) and 20% from the larger groups for validation dataset as per CABIN Guidance for Reference Condition Models (Environment and Climate Change Canada, 2024). A training dataset of 129 sites was used to derive the final model through refinement of the predictor variables. The validation dataset of 34 sites was used to evaluate the final model.

Table 7. Distribution of reference sites from each reference group for the validation and training datasets.

Reference Group	Full Dataset No. of sites	Proportion removed	Validation No. of Sites	Training No. of Sites
1 (1Aia)	56	20%	11	45
2 (1Aib)	50	20%	10	40
3 (1B)	17	30%	5	12
4 (2)	40	20%	8	32
Total	**163**	**22%**	**34**	**129**

4.1 Biological structure of the training dataset used for the final model

Table 8 provides SIMPER analysis values for the overall model and between reference groups illustrating the taxa that contribute to the similarity within each reference group. Table 9 provides SIMPER analysis values for the overall model and between reference groups illustrating the dissimilarity among groups. The same four families contribute to the within group similarity across the groups, and it is the change in relative abundance of these families that contributes to the differentiation of the groups, that and the contribution of the less frequently occurring families. Group 4 has the strongest within group similarity. The strongest between group dissimilarity is between Group 1 and Group 4.

Table 8. SIMPER analysis of the biological group similarity and top five contributing taxa in the Okanagan 2025 Model.

Group Similarity	Taxon	Avg. Abundance	% Contribution	Cumulative %
Group 1	Baetidae	1353.66	32.25	32.25
48.1%	Chironomidae	772.66	14.8	47.05
	Heptageniidae	686.21	13.56	60.61
	Nemouridae	543.45	10.87	71.49
	Ephemerellidae	482.64	8.3	79.79
Group 2	Nemouridae	394.58	16.98	16.98
50.44%	Baetidae	415.26	16.7	33.68
	Heptageniidae	358.72	14.64	48.32
	Ephemerellidae	223.54	8.53	56.85
	Taeniopterygidae	430.09	7.44	64.29
Group 3	Chironomidae	558.19	35.4	35.4
48.73%	Nemouridae	245.29	12.62	48.02
	Heptageniidae	181.28	9.69	57.71
	Capniidae	173.81	9.48	67.19
	Rhyacophilidae	122.52	5.51	72.7
Group 4	Baetidae	292.22	33.39	33.39
51.12%	Heptageniidae	171.23	19.29	52.68
	Nemouridae	90.11	7.44	60.12
	Chironomidae	93.36	7.06	67.18
	Ephemerellidae	61.37	6.92	74.1

Table 9. SIMPER analysis of the biological group dissimilarities and the top five taxa contributing to the difference between each group of the Okanagan 2025 Model.

	Group 1	Group 2	Group 3
Group 2	**59.18%**		
	Baetidae		
	Chironomidae		
	Heptageniidae		
	Taeniopterygidae		
	Nemouridae		
Group 3	**67.36%**	**62.4%**	
	Baetidae	Chironomidae	
	Chironomidae	Taeniopterygidae	
	Heptageniidae	Baetidae	
	Nemouridae	Nemouridae	
	Ephemerellidae	Heptageniidae	
Group 4	**71.71%**	**61.15%**	**65%**
	Baetidae	Taeniopterygidae	Chironomidae
	Chironomidae	Nemouridae	Baetidae
	Heptageniidae	Baetidae	Nemouridae
	Nemouridae	Heptageniidae	Capniidae
	Ephemerellidae	Elmidae	Heptageniidae

In terms of biological characteristics, sites in Group 1 have the highest overall abundance and lowest richness and diversity values. Lowest abundance values were seen in Group 4, with the range in Group 1 (5744 average individuals) to Group 4 (1210 average individuals). Highest diversity and % EPT values were seen in Groups 2 and 4 (Table 10). Despite the specific differences, it should be noted that there is not a great deal of variation in richness and diversity across the reference groups. In fact, this reference dataset had much less difference among the biological groups than some of the other models developed for British Columbia, such as the Fraser River and the BC Central North Coast.

Table 10. Average biological characteristics of reference groups for the Okanagan 2025 Model.

Biological Characteristic	Group 1 (N=45)	Group 2 (N=40)	Group 3 (N=12)	Group 4 (N=32)
Richness	23.8	24.0	24.6	24.2
Abundance	5744	3247	2131	1210
Simpsons Diversity	0.82	0.86	0.84	0.84
Simpsons Evenness	0.70	0.76	0.73	0.73
% EPT	68%	74%	53%	75%
% Chironomidae	11%	6%	20%	6%

4.2 Discriminant function analysis results of the final model

A stepwise (forward and backward) discriminant function analysis (DFA) was performed on the training dataset based on the 4-group assignments defined with the complete dataset. Additional DFA models were examined by iteratively adding and removing variables to improve the accuracy of model prediction. A total of 54 different models were run in the technical report (Reynoldson, 2025) examining predictor variables based on:

- Examining correlation among variables within habitat categories (e.g., climate, channel, etc.)
- Examining correlation between biological data (NMDS ordination scores) and habitat matrices
- Using matching procedures[1] to identify a subset of habitat variables that best reflect the pattern observed in the biological dataset
- Minimizing variable redundancy identified through tolerance scores
- Identifying variable entry and removal based on F scores
- High overall classification rate and jackknifed classification rate
- Balanced individual group rates of the jackknifed classification and consistency in the classification of each group.

The final model, based on the training dataset with 129 sites, uses four groups and ten predictor variables. Ten predictor variables resulted in a fairly low Wilks Lambda of 0.295 (compared with other reported values for CABIN models) which indicates good discrimination among the groups (Table 11).

[1] matching of similarity matrices was completed using BEST analysis in PRIMER 6

Table 11. Final model variables with F scores and tolerance values.

Variable	F score	Tolerance
Latitude	3.007	0.601
Altitude	9.702	0.452
Draining Area	11.941	0.002
Stream length	11.288	0.002
Gradient-watershed area with hillslope > 60%	6.588	0.374
Gradient-maximum watershed hillslope	6.537	0.384
Landcover-Water	13.266	0.605
Max Channel Depth	7.820	0.315
% Canopy Coverage	5.381	0.515
Straight Run in Reach	3.906	0.803
Wilks Lambda = 0.295		

The classification results are provided in Table 12. The final model had an overall classification rate of 67% and an overall jackknifed classification rate of 61%. The individual group classification rates were on average 2.7 times better than a random prediction (random prediction of a four-group model is 25%) with good consistency for all groups (average rate was 68.5% ± 15.2%).

Table 12. DFA classification matrices for reference site prediction to reference groups (n=129).

Variables (10 variable)	Latitude, Altitude, Drainage Area, Stream length, Gradient-watershed area with hillslope > 60%, Gradient-maximum watershed hillslope, Landcover-Water, Max Channel Depth, % Canopy Cover, Straight Run in Reach				
Group	**1**	**2**	**3**	**4**	**Total**
No. of Sites	45	40	12	32	129
Predicted Group	Resubstitution Classification Matrix (no. of sites)				**% Correct**
1	**32**	4	6	3	**71**
2	4	**29**	4	3	**73**
3	0	0	**10**	2	**83**
4	2	12	3	**15**	**47**
Total	**38**	**45**	**23**	**23**	**67**
Predicted Group	Jackknifed Classification Matrix (no. of sites)				**% Correct**
1	**31**	5	6	3	**69**
2	4	**27**	7	14	**68**
3	1	0	**7**	4	**58**
4	3	12	3	**14**	**44**
Total	**39**	**44**	**20**	**26**	**61**
Precision (Jackknifed) CV	80.6%				
F ratio	5.905				

The final set of ten predictor variables with their descriptions for the final model are listed below:

1. **Latitude:** sampling location in decimal degrees.
2. **Altitude:** height in meters above sea level.
3. **Drainage Area:** delineated basin area in km2 upstream of sample location.
4. **Stream length:** total length of stream in upstream drainage area.
5. **Gradient-watershed area with hillslope > 60%:** the proportion of the drainage area with a gradient of more than 60%.
6. **Gradient-maximum watershed hillslope:** maximum slope of the drainage area.
7. **Landcover-Water:** % upstream drainage area with water including lakes, reservoirs and rivers.
8. **Max Channel Depth:** maximum recorded water depth in stream.

9. **% Canopy Coverage:** category based on range of canopy coverage percentage.
10. **Straight Run in Reach:** presence (1) or absence (0) of straight run in sampling reach.

4.3 Habitat description of the training dataset used for the final model

A summary of selected habitat characteristics of the reference model groups is provided in Table 13. Predictor variables are indicated in bold font.

Table 13. Average habitat characteristics of reference groups for the Okanagan 2025 model (predictor variables are in bold).

Habitat Variable	Group 1 (n=45)	Group 2 (n=40)	Group 3 (n=12)	Group 4 (n=32)
Latitude ºN	**49.483**	**49.919**	**49.256**	**49.708**
Longitude ºW	-119.803	-119.641	-120.276	-119.418
Altitude	**927.3**	**849.6**	**1246.2**	**831.1**
Gradient-watershed area with hillslope > 60 % (%)	**8.55**	**6.56**	**7.2**	**8.65**
Gradient-maximum watershed hillslope (%)	**164.7**	**125.4**	**133.5**	**198.3**
Stream Order	2.9	2.2	2.5	2.4
Stream length (km)	**287.7**	**90.9**	**29.0**	**310.6**
Drainage Area (km^2)	**160.8**	**49.1**	**17.3**	**168.3**
Precip.-Total ANNUAL 30 yr average (mm)	809	798	901	855
Temp.-ANNUAL mean 30 yr average (°C)	2.72	2.97	2.28	2.48
Temp.-ANNUAL min – coldest months 30 yr average (°C)	-10.72	-10.83	-10.17	-11.06
Temp.-ANNUAL max - warmest months 30 yr average (°C)	20.22	20.78	19.02	20.07
Landcover-Water (%)	**0.68**	**0.34**	**0.01**	**0.15**
% Canopy Coverage (category)	**1.69**	**2.80**	**2.42**	**2.25**
Straight Run in Reach (binary)	**0.91**	**0.70**	**0.42**	**0.72**
Bank Full Width (m)	10.4	7.3	5.5	11.4
Avg Channel Depth (cm)	17.5	11.8	10.8	19.5
Max Channel Depth (cm)	**25.3**	**16.7**	**15.1**	**27.9**
Avg Velocity (m/s)	0.372	0.318	0.320	0.366

4.4 Model evaluation

Validations sites are used as an independent test of the model classification success to assess model accuracy. The variation in the accuracy among the groups within a model provides an estimate of the precision of the predictive model, where a more precise model would have similar accuracy across the groups.

These sites were classified in one of the four reference groups prior to refining the model with the training dataset. The overall validation site accuracy of 44% is less than that reported in Table 12 (> 60%) and less than the model accuracy prior to the removal of validation dataset (Reynoldson 2025). This is generally the case when predicting validation sites. However, the precision of the model is relatively good at 72%.

The power of the various models in detecting divergence from reference can be described by quantifying Type I and Type II error rates. Validation sites provide an objective evaluation of the model's ability to detect a known disturbance to the reference condition. The validation data are altered to simulate an impact (referred to as *simpact* data).

In this case, the abundance and richness of certain taxa based on known tolerances to organic pollution (Hilsenhoff, 1988) were altered to simulated enrichment disturbance. Three levels of disturbance (mild, moderate, and severe) were imposed on the validation sites to increase the degree of departure from the reference taxonomic composition (Bailey *et al.* 2014, Reynoldson and Raggett 2019). For example, at the' Mild' level of enrichment disturbance 'Sensitive' taxa abundance was reduced to 50% and richness by 20%. For 'Insensitive' taxa, abundance was decreased to 75% while richness remained stable. Tolerant taxa saw a doubling of abundance at this disturbance level. At the 'Severe' level of disturbance, 'Sensitive' taxa were reduced to 15% of the original abundance and 25% of the original richness and 'Tolerant' taxa were decreased to 50% and richness was decreased to 85%. Table 14 outlines the changes made to the communities at the three levels to simulate an organic enrichment disturbance.

Table 14. Changes to validation (reference) data to simulate three levels of disturbance.

Enrichment Disturbance Level	Taxa Sensitivity Type		
	Sensitive	Insensitive	Tolerant
Mild (SIM 1)	50% of abundance 80% of richness	75% of abundance No change to richness	200% of abundance No change to richness
Moderate (SIM 2)	25% of abundance 50% of richness	50% of abundance 80% of richness	75% of abundance 90% of richness
Severe (SIM 3)	15% of abundance 25% of richness	25% of abundance 60% of richness	50% of abundance 85% of richness

4.4.1 Type I Error

Standard CABIN assessments use a 90% confident ellipse to detect mild divergence from predicted reference condition. To determine the Type I Error, each validation site (using actual data, not *simpacted* data) was assessed against the reference sites from the group to which it is known to belong. Using the 90% confidence ellipses, 10% of sites in reference condition (in this case three validation sites) would be expected to fall outside the 90% ellipse. However, seven validation sites fell outside the 90% ellipse resulting in a Type I Error of 21% (Table 15). In this case, the model finds sites to be divergent from reference when they are actually in reference condition.

4.4.2 Type II Error

To determine the Type II Error, *simpact* sites were assessed against their known reference group to determine if the disturbance would be detected. As shown in Table 15, the model's Type II Error indicates that moderate disturbance was detected in 85% of cases, and severe disturbance was detected in 97% of cases. This results in a Type II Error of 15% and 3% for moderate and severe disturbances. In these cases, the model is very effective at detecting impacts where changes to abundance and richness occur for all taxa.

Overall, a higher Type I Error indicates that more sites will tend to be assessed as disturbed when they are not. This is in favour of the environment. The low Type II Error indicates that the model has a high capacity to detect moderate to severe disturbance.

Table 15. Type I and II error for proposed model and individual reference groups for a simulated enrichment disturbance of the validation dataset (n=34).

	Type I	Type II		
Level of Disturbance	**Null[2] (SIM 0)**	**Mild (SIM 1)**	**Moderate (SIM 2)**	**Severe (SIM 3)**
No. Sites in error	7	25	5	1
% Error rate	21%	74%	15%	3%
By Reference Group				
Group 1 (n=11)	0%	91%	18%	9%
Group 2 (n=10)	18%	64%	0%	0%
Group 3 (n=5)	0%	100%	60%	0%
Group 4 (n=8)	63%	38%	0%	0%

Overall, a higher Type I Error indicates that more sites will tend to be assessed as disturbed when they are not. This is in favour of the environment. The low Type II Error indicates that the model has a high capacity to detect moderate to severe disturbance.

4.4.3 Comparison with other models

Comparison of error rates with other models for this degree of disturbance shows both the models perform as well or better than other CABIN RCA models for which error rates have been reported (Table 16). The Type I errors are within the range shown by other models and almost the same as the Okanagan 2017 model. The Type II error rates are similar to the other models for mild disturbance but are markedly lower for moderate and severe disturbances. Comparison with the current model shows the Type I errors to be similar, except for the mild disturbance, and the Type II errors are very low (Table 16). Type II Error was not determined for the Okanagan 2017 model.

[2] In the *simpact*ing process, the starting point (null) was the reference condition (*i.e.,* the actual assemblage of organisms at each of the validation sites).

Table 16. Comparison the final Okanagan 2025 model error rates based on simulated enrichment disturbance with other models in the CABIN analytical tools.

Model	Type 1	Type 2 (mild)	Type 2 (moderate)	Type 2 (severe)
Fraser 2021	22%	56%	14%	2%
Peace River 2019	9%	61%	52%	0%
BC Central North Coast 2015	22%	72%	59%	53%
BC Central North Coast 2023	10%	89%	60%	16%
Okanagan 2017	19%		Not measured	
Okanagan 2025	21%	74%	15%	3%

4.4.4 Comparison with Okanagan 2017 model

A comparison of the proposed Okanagan 20025 model with the Okanagan 2017 model (Table 17) shows the 2017 model to have slightly higher accuracy and precision and requires one additional predictor variable. However, it is generally the case that as the models incorporate more sites and thus capture more of the biological variability, accuracy tends to decline. The corollary is that the models have built in redundancy as the degree of overlap between the communities tends to be greater. The two models have almost half (five) variables in common, suggesting a strong relationship between biological structure and the physical drivers.

The two models also show strong concordance in the reference groups. While the group labels differ, three of the groups to be very similar in the composition of the reference sites. Overall, there is 75% concordance between the two classifications. Comparison of the models are summarized in Table 16 below.

The same test sites assessed in the supporting document for the Okanagan 2017 model were assessed with this Okanagan 2025 model (Reynoldson 2025). Overall, the 2025 model was slightly more powerful in detecting difference from reference as more sites were detected outside of the 99% (Divergent) or the 99.9% (Highly divergent) confidence ellipse compared with the 2017 model (Table 17).

The 2025 model is recommended for ongoing regional assessments due to its sensitivity in detecting difference from reference condition, despite the modest reduction in classification accuracy from the earlier version.

Table 17. Comparison of Okanagan 2017 and Okanagan 2025 models.

Parameter	Okanagan 2017 Model	Okanagan 2025 Model
Reference datasets	121 total reference sites (2005-2016)	171 total reference sites (2005-2023)
Number of families	65	69
Outliers	15	8
Groups	4	4
Group size	23,28,30,20	32, 45, 40, 12
Number of Predictors	11	10
Predictor Variables	Latitude; Drainage Area; Maximum Slope; % Water; Channel Maximum Depth; % Intrusive Bedrock; % Metamorphic Bedrock; April Precipitation; % Wetland Treed; % Pebble Substrate; Minimum Elevation;	Altitude; Latitude; Drainage Area; Stream length; Gradient-watershed area with hillslope > 60%; Gradient-maximum watershed hillslope; Landcover-Water; Max Channel Depth; % Canopy Coverage; Straight Run in Reach
Training dataset	85	129
Validation dataset	21	34
F ratio	5.735	5.905
Classification Accuracy	74% resubstitution; 65% jackknifed	67% resubstitution, 61% jackknifed
Within Model Precision (CV)	89.1%	80.6%
Validation Accuracy	90%	44%
Validation Precision (CV)	88%	72%
Type I Error	19%	21%
Type II Error (Moderate)	Not reported	15%
Type II Error (Severe)	Not reported	3%
Test sites assessed as divergent or highly divergent	8/16	10/16
Model Strength	Good performance but limited representation of high-slope / headwater conditions; smaller dataset	Performance within acceptable range; broader geomorphic and climatic coverage; larger dataset

5. CONCLUSIONS AND RECOMMENDATIONS

The Okanagan 2025 model was constructed from an original reference dataset of 171 sites. With the removal of eight outliers and a validation dataset, the final model was based on 129 reference sites. It comprises four reference groups representing different assemblages of organisms and requires ten habitat variables for model predictors. This is also the first time the power of the model to detect disturbance has been determined for the Okanagan.

Examination of the range of biological variation shows that the same invertebrate families tend to characterize the reference groups in different ways. It is the change in relative abundance of these families that contributes to the differentiation of the groups, as well as the contribution of the less frequently occurring families. It is notable the actual range in the biological gradient for the Okanagan reference communities is much less than in other models and reference datasets in BC. The habitat characteristics of the reference groups are associated with the composition aspects of the benthic invertebrate communities and less associated with the absolute abundance.

Compared to 2017, the 2025 model uses a larger dataset, updated classification standards, an updated predictor set and captures greater biodiversity with four additional families included in this revision (Table 17). The 2017 model reported slightly higher classification accuracy and precision, which is expected when fewer sites are included. The strength of the 2025 model is that it has a strong ability to detect difference from reference condition, an important aspect for ongoing regional assessments. The Okanagan 2025 CABIN RCA model provides a reliable framework for ongoing stream health assessment in southern British Columbia.

Any future reference site sampling in the Okanagan study area should be focused on sites anticipated to be in Group 3. Distribution of sites by ecoregion suggest that sampling should be focused in the Similkameen basin in the Okanagan Ranges ecoregion (41% of Group 3 sites) and the Thompson Okanagan Plateau ecoregion (58% of Group 3 sites)

6. OTHER RELEVANT LITERATURE

Bailey, R.C., S. Linke, A.G. Yates. 2014. Bioassessment of Freshwater Ecosystems using the Reference Condition Approach: comparing established and new methods with common data sets. Freshwater Sc. 33:1204-1211.

Belbin, L. 1993. PATN, pattern analysis package. Division of Wildlife and Ecology, CSIRO, Canberra, Australia.

Clarke, K.R. and Gorley, R.N. 2015. "PRIMER v7: User Manual/Tutorial". PRIMER-E.

Environment Canada. 2012. Canadian Aquatic Biomonitoring Network field manual – wadeable streams. 57pp ISBN 978-1-100-20816-9 https://publications.gc.ca/collections/collection_2012/ec/En84-87-2012-eng.pdf

Environment Canada. 2020. CABIN Laboratory Methods: Processing, Taxonomy and Quality Control of Benthic Macroinvertebrate Samples. 36 pp. ISBN 978-0-660-37046-0 https://publications.gc.ca/site/eng/9.895039/publication.html

Environment and Climate Change Canada, 2024. Canadian Aquatic Biomonitoring Network Guidance for Reference Condition Models V2.0. November 2024. 18p.

Hilsenhoff, W.L. 1988. Rapid Field assessment of Organic Pollution with a Family Level Biotic Index. J. N. Am. Benth. Soc. 7:65-68.

Natural Resources Canada, 2009. GeoBase – Landcover, circa 2000 http://geogratis.gc.ca/api/en/nrcan-rncan/ess-sst/643c4911-475b-4765-b730-2dde9be50d5b.html)

Norris, S. 2012. British Columbia's Provincial Stream Biomonitoring Program. Technical Documentation. GIS Tools for Reference Site Selection and Upstream Watershed Analysis. Prepared for B.C. Ministry of Environment. March 2012.

Reynoldson, T.B. 2025. An Update of the Okanagan CABIN RCA Model for British Columbia (2006-2023). 76pp + appendices. Report submitted to Environment and Climate Change Canada March 2025.

Reynoldson, T.B., Rosenberg, D.R., Day, K.E., Norris, R.H., Resh, V.H. 1997. The reference condition: a comparison of multimetric and multivariate approaches to assess water-quality impairment using benthic macroinvertebrates. J. N. Am. Benth. Soc. 16:833-852.

Reynoldson T.B. and Raggett, J. 2019. Development of a CABIN Model for the Peace River. Report submitted to British Columbia Ministry of Environment and Climate Change Strategy by GHOST Environmental Consulting. 65pp.Strachan, S. and S. Bennett. 2018. Reference model supporting documentation for CABIN Analytical Tools: Okanagan 2017. Prepared for CABIN web application by Environment and Climate Change Canada and Limnotek Research and Development Inc. 23 pg.

Yates, A. and R.C. Bailey. 2010. Selecting objectively defined reference sites for stream bioassessment programs. Environmental Monitoring and Assessment 170:129-140.

APPENDIX: DATA COLLECTION, ANALYSIS AND QUALITY ASSURANCE

1. Field Collection

CABIN Study Name	BC MOE-Okanagan Kicknet	EC-Okanagan Basin
Agencies involved	BC Ministry of Environment	Environment and Climate Change Canada
Date range	2005-2023	2005-2023
Sampling season	Late Aug-Mid Oct	Late Aug-Mid Oct
# reference samples	89	74
Certified samplers (Y or N)	Y	Y
Certified team leader (Y or N)	Y	Y
400 um kicknet (Y or N)	Y	Y
Other sampling device	--	--
Field Preservative	Ethanol	Formalin

2. Macroinvertebrate Identification

CABIN Study Name	BC MOE-Okanagan Kicknet	EC-Okanagan Basin
Taxonomist	Fraser Environmental/ Cordillera Consulting	Cordillera Consulting/ EcoAnlaysts
Marchant Box used	N	Y
Other subsampling device	Caton Tray	--
Subsample count	300	300
10% of reference samples sent to National Lab for QA	N	Y
Reference Collection maintained	N	Y

3. GIS analyses

All basins were delineated using ArcGIS Pro 3.3.5 Spatial Analysis tools (ESRI 2024). Delineations were based on a 20 metre resolution digital elevation model (DEM) and 1:20,000-1:50,000 scale hydrological network. The DEM was hydrologically reconditioned with the stream network and artificial sinks filled to improve flow modeling. The resulting corrected DEM was then used to calculate flow direction, flow accumulation, and stream definitions, which enabled the delineation of basin areas. Using the original DEM and ESRI Spatial Analyst tools, slope and elevation statistics were calculated for each delineated basin. The basins were further characterized using the publicly available GIS layers summarized in the table below.

Description of data layers and sources:

Data layer	Scale/ Resolution	Source and method
Basin Morphometry	20 m	Area and perimeter were calculated from delineated basin as described above

Bedrock	1:50,000 to 1:250,000	BC Geological Survey, BC Data Catalogue, Bedrock Geology https://catalogue.data.gov.bc.ca/dataset/bedrock-geology The percentage of each bedrock category within each delineated basin was calculated using Spatial Analyst Intersect tool (ESRI 2024)
Climate	60arc - sec/ 2 km	Natural Resources Canada, Open Government Portal, 30 Year Spatial Climate Averages (1971-2000, 1981-2010, 1991-2020) https://ftp.maps.canada.ca/pub/nrcan_rncan/Climate-archives_Archives-climatologiques/NAM_30_Year_Averages/ Long-term monthly and annual temperature and precipitation normals were summarized for each basin using the Spatial Analyst Zonal Statistics as Table tool (ESRI 2024). Sampling dates were temporally matched to long-term normal periods as follows: 1971-2000 data were used for pre-2010 samples, 1981-2010 data for 2010-2019 samples, and 1991-2020 data for all sites sampled in 2020 or later.
Hydrology	1:50,000	Natural Resources Canada, Open Government Portal, National Hydro Network /CanVec https://natural-resources.canada.ca/science-data/science-research/geomatics/download-directory-documentation Stream length and drainage density for each delineated basin was calculated using Spatial Analyst Intersect tool (ESRI 2024).
Land Cover	30m	Natural Resources Canada , Open Government Portal Land Cover of Canada (2010, 2015, 2020) 2010 - Open data -Record ID: c688b87f-e85f-4842-b0e1-a8f79ebf1133 https://open.canada.ca/data/en/dataset/c688b87f-e85f-4842-b0e1-a8f79ebf1133 2015 - Open data -Record ID: 4e615eae-b90c-420b-adee-2ca35896caf6 https://open.canada.ca/data/en/dataset/4e615eae-b90c-420b-adee-2ca35896caf6 2020 - Open data -Record ID: ee1580ab-a23d-4f86-a09b-79763677eb47 https://open.canada.ca/data/en/dataset/ee1580ab-a23d-4f86-a09b-79763677eb47 Both vector and raster-based methods were used, both methodologies yielded highly comparable results. For the vector-based approach, spatial data were converted to polygons and processed via the Spatial Analyst Intersect tool (ESRI 2024). The raster-based approach quantified zonal distributions using the Spatial Analyst Tabulate Area tool (ESRI 2024). Sampling dates were temporally matched to the landcover layers as follows: 2010 landcover data were used for pre-2015 sites, 2015 landcover data for the 2015–2019 period, and 2020 landcover data for all sites sampled in 2020 or later.

4. Statistical Analyses

Statistical programs used:

- Excel – data manipulation and storage
- PATN - clustering

- PRIMER 6 – clustering, SIMPROF, MDS ordination, ANOSIM, SIMPER for biological data
- SYSTAT 11 – discriminant analysis and test site analysis (plotting BEAST assessments with probability ellipses)

Model was reviewed by the CABIN Science Team November 2025. Comments provided and recommendations applied and final copy submitted January 2026.