How BC Cycle Tourism Rates Routes

Methods and Limitations

BC Cycle Tourism Society

Published July 2026

Preliminary — working data, subject to revision

Researched and drafted with AI assistance (Claude models, Anthropic)

Suggested citation: BC Cycle Tourism Society (2026). How BC Cycle Tourism Rates Routes: Methods and Limitations [report]. Version July 2026. https://bccycletourism.ca/research/

Version	Date	Notes
1.0	July 2026	First published version

Heather Piwowar · heather@bccycletourism.ca
bccycletourism.ca/research/

1. Introduction and scope

BC Cycle Tourism Society lists 320 published cycling routes for British Columbia (BC) and describes each one with a difficulty tier and a set of descriptive labels: traffic comfort, fast-road presence, surface mix, ruggedness, climbing, and remoteness. The routes are published by route makers across the province; what we add is one consistent description of all of them. We are not the publisher of any of these routes. Nor are they all expedition rides: 111 of the 320 are under 30 km.

This document explains how every one of those numbers and labels is computed, what evidence the choices rest on, and where the system's limits are. It is written for researchers, planners, trail organizations, and anyone who wants to check our work or build on it.

What this document is not. It publishes methods and limitations, not reproducibility. These numbers are derived from OpenStreetMap. We do not yet record which OpenStreetMap extract produced any given number, so a route's labels can change when we rebuild, and the published figures cannot be exactly reproduced from a dated source. Recording the extract is planned work, not something already done.

Companion dataset. Every derived metric described here, one row per route, is published as the route metrics dataset (Excel and CSV with a plain-text column guide) at bccycletourism.ca/research/. A blank cell means we could not measure that thing, not that it is zero. A rider-facing summary of what the labels mean lives at bccycletourism.ca/routes/how-we-describe/.

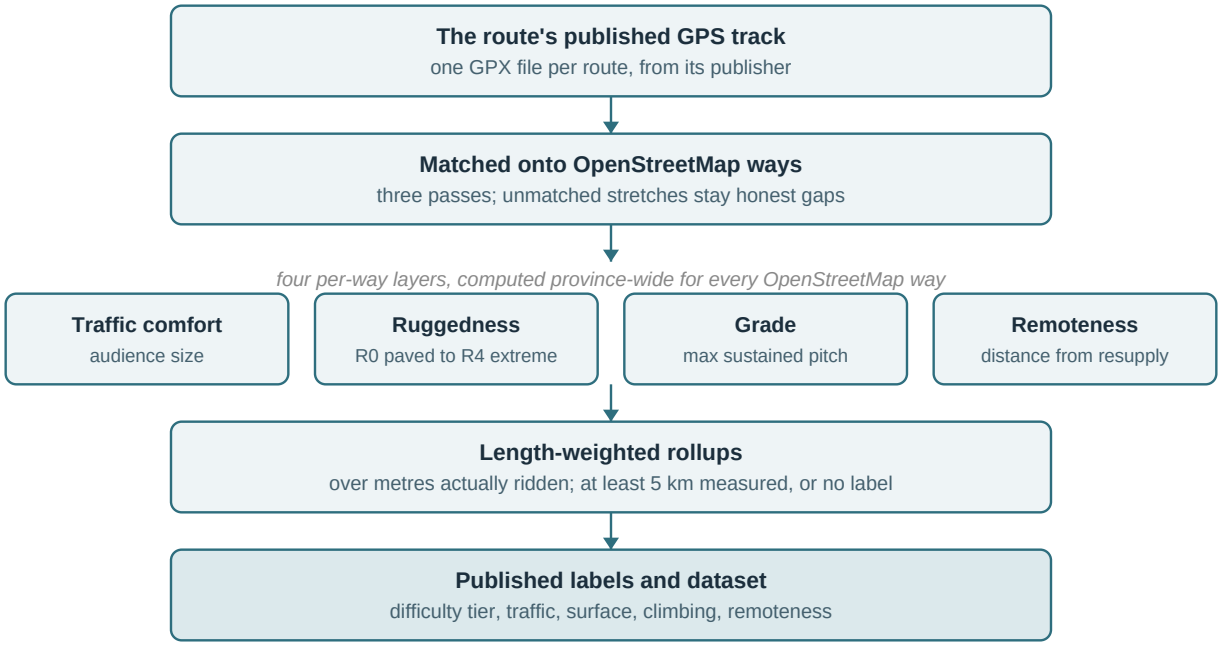

Figure 1. The pipeline, from a route's published GPS track to its published labels.

2. Data sources

Everything below is derived from open data. Dates are given where we have them; where we do not record one, that is stated rather than papered over.

- **OpenStreetMap (OSM).** All road and trail geometry and attributes: highway class, surface, speed limits, cycling facilities, access restrictions. This is the input every derived layer reads. We do not record which OSM extract a published number came from; see Limitations.

- **Shuttle Radar Topography Mission (SRTM) elevation**, 1 arc-second (roughly 30 m). All grade computations sample this terrain model.
- **North American Land Cover Monitoring System (NALCMS) 2020 land cover**, 100 m raster. Used to gate untagged-way defaults (urban versus forest) and to detect water crossings in the remoteness layer.
- **The We Welcome Cyclists directory, Groceries category** (about 2,600 listings province-wide when the remoteness layer was computed; the directory keeps growing underneath it). The resupply population behind the remoteness score. Restaurants are deliberately not counted: they are not resupply.
- **Publisher-supplied statistics.** Each route's own published difficulty rating where one exists, and one publisher's self-reported singletrack share. These are used as labelled pass-throughs, never silently mixed into measurements.
- **HUB Cycling's Bikeway Classification System (Appendix C, 2023)** as the base rule table for traffic comfort, and HUB Cycling's 17,622 labelled Metro Vancouver route segments as the validation set (Section 6). The Canadian Bikeway Comfort and Safety (Can-BICS) classification is used as a cross-check baseline only, never as an input. We thank HUB Cycling for publishing both the framework and the data.

3. Matching routes to the map

Every route arrives as a GPX (GPS Exchange Format) track from its publisher. Before anything can be measured, the track has to be matched onto the OpenStreetMap ways it actually rides. Every derived number in this document is a rollup over those matched ways, weighted by the metres of each way the route uses, so a 200 m rough patch on a 60 km ride counts as 200 m, not as "one segment."

Three matching passes

- **Spine cover.** The drawn line is densified to a point every 25 m. Candidate OSM ways within 30 m are gathered and the route is covered greedily: take the way that explains the longest stretch of route, credit it only that stretch, continue from where it ends. Near-ties prefer the way just credited, then a trail over a parallel road (where a path parallels a highway, the path gets the credit), then the nearest. Ferry ways are never matched or credited as riding: stretches of track within 120 m of a mapped ferry way are recognized as crossings and set aside.
- **Routed gap-fill.** Uncovered stretches of 200 m or more that stay near some mapped way (at least 60% of the gap within 300 m) are bad GPS fits, not backcountry. They are re-routed along the author's corridor and accepted only if the result hugs the drawn line: routed length at most 1.4 times the gap, mean deviation at most 200 m. Gaps that substantially follow a ferry corridor (more than 30% of the gap within 120 m of a ferry way) are never routed; they are crossings.
- **Drift retry.** What remains is retried at a 120 m radius with a ±25° bearing-alignment gate: a hand-drawn track drifting off a mapped way still matches, but cross streets cannot.

Coverage accounting

After matching, the route is re-sampled every 100 m and each point classified as ferry, outside BC, matched, or in-BC unmatched; points on open ocean beyond any ferry corridor fall outside all four classes and are excluded from the denominators. Two coverage figures are kept per route: the raw matched share of the whole line, and the match rate over the rideable portion alone (in BC, on land, off ferries), which is the honest one, since ferry and cross-border kilometres explain a low raw share. Spans matching no way stay honest gaps and are never force-matched; each route records the dominant cause of its gaps (ferry, cross-border, or unmapped backcountry), not a cause per span.

What partial coverage means for every label. A rollup sees only the matched portion of a route, and the 5 km measurement gate (Section 5) checks kilometres measured, not share of route measured, so a route with 40% coverage can still carry labels, computed on the covered part. Unmatched kilometres are typically the most obscure riding, unmapped resource roads and private connectors, which biases partially covered routes slightly toward the tamer side. Both coverage figures ship per route in the companion dataset so readers can judge this themselves.

The British Columbia clip

Each route is sampled roughly every 300 m against the provincial boundary; the share of land samples inside BC is stored per route (ocean samples are excluded from the denominator, because a coarse coastline would otherwise sink coastal rides). Routes that cross the border are clipped to their BC legs only where a curator has marked them for clipping; a short non-BC gap between two BC runs is bridged only when it contains no foreign land point, meaning it is coastline noise rather than a genuine excursion into Washington or Alberta. For clipped routes, distance, ascent, and every share-based label describe the BC leg only.

One honest exception. For 113 of the 320 routes, all from one BC guidebook series, the in-BC share is *asserted, not measured*: those routes arrived through a different ingestion path and the boundary sampling never ran on them. They are BC guidebook rides by definition, so the dataset asserts 1.0 and labels the source in its own column, and the underlying database keeps an honest null. This is Limitation 12.

4. The four per-way layers

Four attributes are computed for every rideable OpenStreetMap way in the province, independent of any route. Route labels are length-weighted rollups of these.

4.1 Traffic comfort

The comfort ladder answers one question per way: how much motor traffic would you share, and what stands between you and it? Six rungs, C1 to C6, each a condition, not a quality verdict: C1 is separated paths or slow streets with few cars; C4 is fast or heavy traffic with a painted line at most; a high-rung road is not a bad road, it is a road for experienced traffic riders. The base rule table is HUB Cycling's Bikeway Classification System (Appendix C, 2023), applied to OpenStreetMap tags only: highway class, speed limit, cycling facility, surface, and parking. The C1 to C6 labels are our own, not HUB's: HUB's classification of Metro Vancouver is field-reviewed in a way our automated province-wide model cannot match, so our classes carry our names, and HUB's names appear below only where we compare against HUB's own labels.

First, tags map to a facility class:

Facility	OpenStreetMap condition
Bike path	a cycling-only path, separate from any road
Protected bike lane	a physically separated on-road facility
Multi-use path (MUP)	a path shared with pedestrians, including rideable trails
Painted bike lane	paint on the road, no physical protection
Signed shoulder	a designated rideable shoulder, one band above paint
Shared roadway	everything else: a road with no cycling facility

Then facility, speed, and traffic volume look up a comfort class from the HUB chart: **C1** separated paths, or slow streets with few cars; **C2** sharing light traffic, or bike lanes on calm roads; **C3** sharing busier traffic, or bike lanes on busy roads; **C4** fast or heavy traffic, a painted line at most. We extend the ladder two rungs

below the chart, because touring routes ride roads the urban chart never contemplated: **C5** (highway-speed traffic, 60+ km/h) and **C6** (fast highway traffic, 70+ km/h), both with a painted line at most, existing only on true highway classes with no physical separation. Each rung of built infrastructure buys exactly one 10 km/h step of grace: a bare road reads C5 at 60 km/h and C6 at 70; a painted lane at 70 and 80; a signed shoulder at 80 and 90. Only physical separation exempts a way entirely, because paint is not protection at highway speed.

Two classes sit off this ladder. **BACKCOUNTRY** holds bare forestry tracks and undeveloped paths with no positive cycling signal: prime gravel-touring terrain, kept in the network, but not a family-friendly separated path. It is a traffic-comfort judgment only (car-free), so it counts as comfortable in the route traffic label while its roughness is the ruggedness layer's job. **FORBIDDEN** marks ways where cycling is legally excluded, so banned freeway mainlines render as "not allowed" instead of a map gap.

When tags are missing, we guess, and when guessing, we err scary. An untagged ordinary street is assumed 50 km/h. On true highway classes the fallback is each class's own tagged-speed distribution: 110 km/h for motorways down through 60 km/h for untagged secondary highways. That last one means roughly 2,000 km of remote untagged secondary highway reads C5 on a guessed speed. We chose to keep that, deliberately: a guessed road should show as a warning rather than quietly look fine. The bias this buys is conservative in the direction that protects riders.

Traffic volume is synthesized, and we name it as our weakest input. There is no open traffic-count feed for BC, so the chart's volume axis runs on OpenStreetMap proxies: a base volume per road class, bumped upward by validated signals (four or more lanes, a centre turn lane, scheduled bus service, a signed truck route). Each signal was validated by crosstab against HUB Cycling's labels before adoption. No traffic count appears anywhere in the pipeline. Section 8 describes the measurement that would fix this.

Where we deviate from the HUB chart, the deviations are labelled and evidence-backed, and the faithful chart is kept separate underneath so the deviations stay auditable:

- A separated multi-use path is rated C1 even beside a busy road: the chart downgrades it, but HUB's own labels rate such paths most-comfortable in 94–100% of cases. We match what HUB implemented, not their printed rule.
- Shared-lane markings (sharrows) on an arterial rate C4: 70% of HUB's sharrowed arterial kilometres carry that label. Sharrows are paint on exactly the roads that need real facilities.
- Metal grating caps a way at C2 (86% of HUB's metal-surface labels).
- A way crossing a freeway on-ramp or off-ramp at grade is capped at C4, detected purely by shared geometry nodes: grade-separated crossings share none. Separation does not exempt this cap; it does not matter how good the path is if merging traffic crosses it.

4.2 Ruggedness

A tire-choice dial per way, five levels: paved plus the four gravel categories of the Industry Standard Guide to Gravel. Assigned from OpenStreetMap surface tags first, then track grade, then smoothness; a smoothness tag can roughen a level but never polish one, so broken pavement reads rough. Mountain-bike and alpine-hiking difficulty tags bump levels up and never down.

Level	Name	Meaning, tire ballpark
R0	Paved	pavement; road bike
R1	Champagne gravel	smooth hard-packed, rides much like pavement; any bike, 25–32 mm tires
R2	Standard gravel	maintained gravel or dirt, minor potholes and washboard; 28–35 mm

Level	Name	Meaning, tire ballpark
R3	Chunky gravel	poorly maintained, strawberry-sized stones, ruts, sand; 33–40 mm
R4	Extreme	unmaintained, jagged, roots, erosion; borders mountain biking; 40 mm and up

Most backcountry ways carry no surface tag at all, so defaults matter more than rules: roughly three-quarters of all track kilometres in BC are defaulted, not tagged. Defaults are gated on land cover, never on municipal boundaries: an untagged track in forest defaults to R3 (a forest service road), in farmland to R2, in built-up areas to R1. An untagged way that is part of a signed cycling route, or explicitly designated for bicycles, is capped at R1, because signed routes are maintained. Every way records whether its level was tagged, rescued, or defaulted; those confidence flags exist upstream but do not yet reach the published labels (Limitation 1).

4.3 Grade

The published grade per way is the worse of two measurements, taken in the worse direction, from SRTM terrain elevation sampled about every 20 m: a *grind* (ascent per kilometre over a smoothed profile, which catches long steady climbs) and a *wall* (the maximum sustained gradient over a 100 m window). Terrain elevation misreads engineered grades, so three protections guard the wall measurement against phantoms:

- **An adaptive window.** On a benched ledge or canyon rim, a 10 m horizontal slip lands the elevation sample on the cliff face. Where sidehill cross-slope is steep, or the way itself is long, the measurement window stretches from 100 m up to 500 m, diluting the phantom while a real sustained climb still reads steep.
- **A held rise.** A wall's rise must still be present 300 m past the window; a gully or spur hump recovers and is rejected, a real climb holds.
- **Bridges and tunnels are not scored.** The terrain model has no deck elevation, so a bridge's grade is unknowable, and we do not assert a grade we cannot measure.

Route-level distance comes from the track's planar length in the BC Albers equal-area projection; total ascent uses a 5 m hysteresis accumulator (climbs smaller than 5 m of continuous gain do not accumulate), which suppresses GPS elevation noise.

4.4 Remoteness

Remoteness is distance from resupply. Each way is sampled about every kilometre; at each sample we find the three nearest grocery listings from the We Welcome Cyclists directory and take the mean distance, so one wrong or closed listing cannot flip a way to remote. Distance is straight-line, with one correction: a store does not count if the straight line to it crosses a kilometre or more of water, because a store across an inlet is no help. Where fewer than three land-connected stores exist at all (about one way in nine, mostly on the coast), the water rule is dropped and the plain nearest three are used. The mean distance feeds a saturating score from 0 to 1 with its midpoint at 10 km; roughly 25 km from resupply scores about 0.8, the cut used by the route-level remoteness share. A way whose score hangs on a single shop is flagged upstream, another confidence signal that does not yet reach the labels.

5. From ways to route labels

Every route label is a length-weighted rollup over the route's matched ways, weighted by the metres of each way the route actually uses. One gate applies everywhere, and it is the reason some routes carry no label for some dimension. We require at least 5 km of a route to be measured before we publish a figure or a label for it, so short or poorly matched routes are deliberately blank in places. A blank cell means we could not measure that thing, not that it is zero. On a route page the same rule appears as a label that simply is not there. We never guess a label.

5.1 The difficulty tier

Five tiers: Easy, Moderate, Ambitious, Challenging, Epic. A score is summed from four measured terms, and three floors can then raise the result but never lower it.

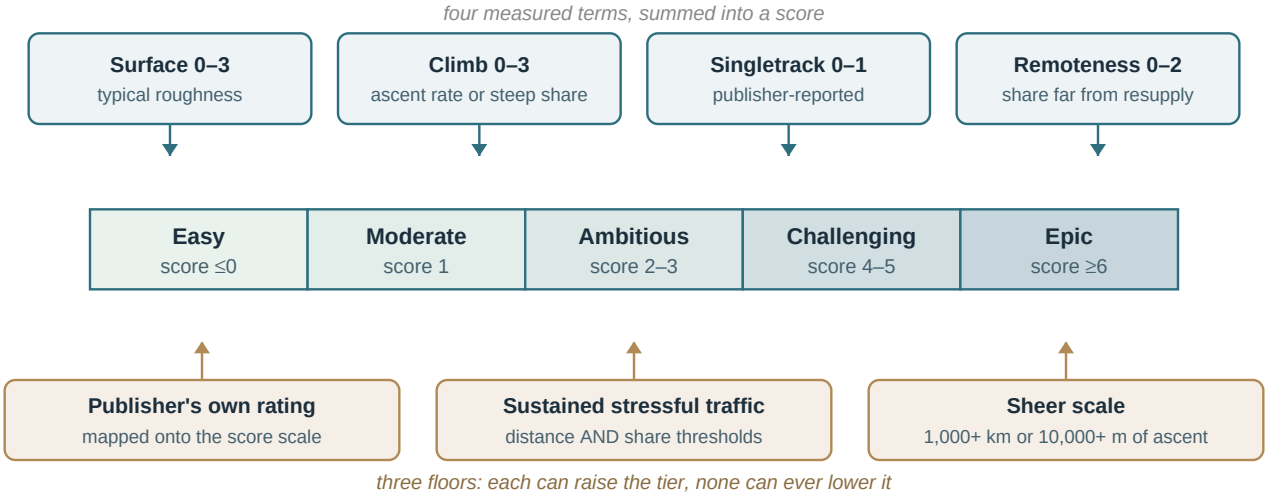

Figure 2. How a difficulty tier is assembled. Floors raise; nothing ever lowers.

- **Surface burden (0–3)** from the route's "typical rough" level: the roughest ruggedness level covering at least 25% of measured distance. For this term only, a smooth, flat, physically separated gravel way (a compacted dyke or rail trail) is demoted one level first, because that rides easy; the descriptive surface labels use the raw level, and the two deliberately differ.
- **Climbing (0–3)**: the worse of average ascent rate (thresholds at 6, 12, and 20 m of ascent per km) and steep share (the share of measured distance on ways whose maximum sustained pitch is 8% or steeper; thresholds at 10%, 25%, and 45%). The average alone would let short brutal walls hide.
- **Singletrack (0–1)**: one point when the publisher reports 30% or more singletrack. Publisher-supplied only; see Limitation 8.
- **Remoteness (0–2)**: one point when 40% or more of measured distance is far from resupply, two points at 60%.

Score to tier: 0 or less is Easy, 1 Moderate, 2–3 Ambitious, 4–5 Challenging, 6 or more Epic. The three raise-only floors:

- **The publisher's own rating**, mapped onto the same score scale, and only when the publisher gave a single unambiguous rating for the exact track.
- **Sustained stressful traffic.** Weighted stressful distance (full weight for the lowest comfort rung and legally excluded ways, half weight for C4 and C5) floors the tier at Challenging when it reaches both 150 km and 45% of the route, and at Ambitious at both 40 km and 25%. Gating on distance and share together means a short urban ride with two bad crossings never trips it.
- **Sheer scale**: 1,000 km or more, or 10,000 m or more of total ascent, floors the tier at Epic. Total ascent is used only at this extreme gate; elsewhere it is distance in disguise. A short curated list covers cross-border expeditions whose BC-clipped distance understates their real length.

5.2 Traffic comfort label

Over the route's classified distance, the comfortable share counts the C1 and C2 classes plus car-free BACKCOUNTRY. At least 75% comfortable reads **"Low-traffic riding"**, 45–75% **"Some traffic sharing"**, under 45% **"Shares roads with traffic"**.

5.3 Fast-road presence

The traffic label above is a share average, so an absolute hazard washes out on a long route: at adoption (July 2026), ten routes labelled "Low-traffic riding" carried more than 10 km of riding on the lowest comfort rung, one of them 92 km of it. The fast-roads label reports presence and commitment instead of composition, and the two labels deliberately appear together: "Low-traffic riding" plus "Fast-road sections, 130 km" is the honest description of such a route.

Fast distance is the two lowest comfort rungs combined. **"Mostly fast roads"** fires at 50% or more of classified distance. **"Fast-road sections"**, with the distance shown, fires at 2 km on the lowest rung alone, or 20 km fast, or 5 km fast when that is at least 30% of the route. The 2 km hair trigger is deliberate, and we audited it before adopting it: almost every route holding 2–10 km of lowest-rung riding held it on a real named highway, exactly the short highway links a cautious reader wants to know about. We chose to err on giving people lots of warning. At adoption (24 July 2026), of 322 then-rated routes, 25 read "Mostly fast roads", 90 "Fast-road sections", and 207 carried no fast-road label; these are figures from that adoption audit, not live counts, and the companion dataset always carries the current ones.

5.4 Surface mix and ruggedness labels

The unpaved share is computed the same way for every route, so the label is consistent across publishers: over the matched ways, weighted by metres used, a way counts as unpaved unless its surface tag is in the paved set. Untagged constructed roads are inferred paved and untagged backcountry tracks unpaved, because paved roads commonly go untagged in OpenStreetMap while gravel gets tagged; every other untagged class is left uncounted, an honest partial. Only when OpenStreetMap coverage is under 5 km do we fall back to a publisher's own figure, and the source ships in the dataset. Bands: under 15% unpaved reads **"Mostly paved"**, 15–49% **"Mixed surface"**, 50–84% **"Mostly gravel"**, 85% and up **"Almost all gravel"**.

Two flags cover roughness the surface bands cannot: the roughest ruggedness level covering at least 10% of measured distance (so a 500 m rough patch on a 60 km ride does not brand the whole route) reads **"Chunky gravel"** at R3 and **"Rough, near-MTB"** (near mountain-bike) at R4. **"Lots of singletrack"** fires at a publisher-reported 30%; it is the one label taken entirely on faith, with no measurement, gate, or cross-check of ours (Limitation 8).

5.5 Climbing, remoteness, and length

- **Climbing**, from ascent per kilometre: under 6 m/km "Gentle grades", 6–12 "Some climbing", 12–20 "Lots of climbing", over 20 "Max climbing". This is the average; short walls are the tier's steep-share term.
- **"Remote stretches"** when at least 20% of measured distance scores 0.8 or higher on the remoteness scale (roughly 25 km or more from grocery resupply).
- **Length classes** over exact distance: under 30 km "Afternoon", 30–90 "Full day", 90–300 "Weekend or more", over 300 "Expedition".

6. Validation

The traffic-comfort classifier is the one layer with a gold standard to test against: HUB Cycling's 17,622 labelled route segments in Metro Vancouver, rating each segment on its own four-class scale (Comfortable for Most, Some, Few, or Very Few). We sample HUB's lines every 25 m, snap each sample to the nearest OpenStreetMap way within 10 m, and compare HUB's label with ours, length-weighted. We publish the crosstabs rather than a single headline accuracy figure; the per-class rows are the honest shape of the evidence.

One thing to know before reading the table: the shipped rules deliberately diverge from HUB's chart above 60 km/h. Our C5 and C6 rungs sit below HUB's scale, so every kilometre we place there scores as a near-miss or worse against a scale that bottoms out at Very Few. Exact agreement fell when those rungs were added, by design; the pre-divergence row shows the same system scored just before they existed. Do not read the gap between the two rows as a regression.

Configuration	Exact	Within one class	HUB Most recall	HUB Some recall	HUB Few recall	HUB Very Few recall	No match
Shipped rule set, measured 26 July 2026	59.5%	82.1%	85%	41%	42%	19%	2.1%
Pre-divergence rule set (2 July 2026), aligned to HUB's scale	71.9%	91.7%	85%	41%	72%	56%	2.4%
HUB chart alone, no labelled deviations	65.1%	93.3%	69%	45%	74%	51%	2.1%
Can-BICS baseline, mapped to the 4-level scale	34.3%	67.4%	22%	3%	69%	35%	8.2%

The shipped rule set's confusion matrix (measured 26 July 2026; rows are HUB's labels, columns ours, row percentages; BACKCOUNTRY sits off HUB's scale, so a prediction there is shown but can never count as agreement):

HUB label \ ours	C1	C2	C3	C4	C5	C6	BACKCOUNTRY	(none)	row km
Most	85%	5%	2%	2%	0%	0%	1%	4%	1,350
Some	15%	41%	35%	8%	0%	0%	0%	0%	436
Few	6%	7%	42%	27%	7%	10%	0%	0%	873
Very Few	7%	2%	14%	19%	17%	39%	0%	1%	277

Reading it honestly: the comfortable end is strong (85% of HUB's Most kilometres agree exactly), and the middle is soft (HUB's Some splits almost evenly between our C2 and C3, mostly a resolution limit of synthesized traffic volume). The bottom two rows are where the deliberate divergence lives: over half of HUB's Very Few kilometres, and a sixth of its Few kilometres, land on our C5 and C6 rungs — fast roads split out by posted speed and facility, which the urban chart never contemplated. HUB's scale cannot validate that split; those rungs were checked instead against known highway corridors (the Fraser Canyon, the Coquihalla, the Malahat and their kin), a weaker form of evidence, and the fast-roads audit in Section 5.3.

The Can-BICS row deserves a fair note: mapped onto a comfort scale it scores poorly, but Can-BICS is a facility-inventory scheme, not a comfort model. It is good at "there is a painted lane here" and structurally unable to describe the comfortable half of the network: on HUB's network it writes off roughly a third of kilometres as non-conforming, and HUB rates just over half of that written-off distance Most (gravel paths, dyke trails, greenways). Where Can-BICS does state a comfort-relevant opinion, we usually agree with it.

The validation boundary matters more than the score. Every gold-standard label is urban and suburban Metro Vancouver. Rural BC, where tags are sparsest and the defaults fire most, has no validation set at all, and province-wide comfort claims are extrapolation, deliberately tuned to err scary when guessing. Applied to every rideable way in BC, the shipped classifier yields: C1 3%, C2 33%, C3 5%, C4 2%, C5 1%, C6 6%, and BACKCOUNTRY 50% of about 254,000 rideable kilometres, plus roughly 560 km marked as legally closed to cycling. That distribution is unscored by design; no labels exist to score it against.

7. Limitations

These are the known limits of the rating system itself, stated without softening. Anything we publish that hedges should draw from this list, not shrink it.

1. **Stacked defaults are invisible in the labels.** A single label can rest simultaneously on a defaulted surface (roughly three-quarters of track kilometres), a guessed speed limit, and a synthesized traffic volume, and it displays identically to a label backed by real tags. The per-way confidence flags that would let us hedge exist upstream but are dropped before the labels.
2. **Calibration is Metro Vancouver; application is all of BC.** The comfort classifier's gold-standard labels are entirely urban and suburban. Rural BC, where tag completeness is worst and the fallbacks fire most, has no validation set. Province-wide comfort claims are extrapolation, deliberately tuned to err scary when guessing.
3. **Threshold cliffs.** Every band is a hard cut: a route at 74.9% comfortable reads "Some traffic sharing" and at 75.0% "Low-traffic riding"; the tier sums coarse buckets, so one OpenStreetMap edit can move a route a full tier. There is no hysteresis and no confidence interval, and no sensitivity analysis has been run on how many routes sit near a boundary. That analysis is planned; nothing false is published without it, but until it runs we cannot say how fragile the labels are.
4. **Coverage bias leans tame.** Rollups see only the matched portion of a route, and the 5 km gate checks kilometres measured, not share measured. Unmatched kilometres skew toward unmapped, wild riding, and untagged-way defaults skew conservative; the errors correlate in the direction of underselling remote routes rather than cancelling out.
5. **Traffic volume is synthesized, not measured.** There is no traffic-count input anywhere in the pipeline; the comfort chart's volume axis runs entirely on OpenStreetMap proxies. This is the weakest input feeding the most rider-facing dimension.
6. **Terrain elevation misreads engineered grades.** A 30 m terrain model cannot see rail beds, benched paths, or bridge decks. The anti-phantom machinery contains the damage; the real fix is LiDAR (light detection and ranging) or recorded-GPS elevation. Bridges and tunnels are unscored by design.
7. **Remoteness is straight-line distance to groceries.** No network distance (a store across an unbridged river can count when the water gap is under a kilometre), no store hours, and no water sources, and water is the resupply axis that matters most exactly where remoteness is highest. And where fewer than three land-connected stores exist at all, the water rule is silently dropped for that way — exactly the coastal places where it matters most.
8. **Singletrack is a pass-through.** One publisher's self-reported number, with no derivation, no measurement gate, and no cross-check of ours.
9. **The mountain-bike exclusion is a name test.** Routes are excluded from the collection when their names match mountain-biking terms. A gravel route with "singletrack" in its title gets flagged; a mountain-bike route with a neutral name does not. Flags are advisory and human-confirmed per route, but the automated net is name-only.

10. **Ratings drift with OpenStreetMap.** Every derived number is a function of one OpenStreetMap extract; edits, including fixes we prompt ourselves, change labels on rebuild. No extract date or pipeline version is currently stamped on published numbers, which is why these figures cannot be exactly reproduced from a dated source.
11. **Publisher floors import unvalidated scales.** The mappings from publishers' own difficulty scales onto our score scale are our reading of their scales, checked against a handful of cases, not systematically validated.
12. **The in-BC share is asserted, not measured, for 113 routes.** Every route from one BC guidebook series bypassed the boundary sampling; the dataset asserts an in-BC share of 1.0 for them, labels the assertion in its own source column, and the underlying database keeps an honest null.

8. What would improve this

The limitations above are also a work plan. In rough order of payoff:

- **Real traffic counts.** The BC Ministry of Transportation and Infrastructure publishes traffic count data covering numbered highways best, exactly the roads where touring happens and where our comfort calls matter most. Snapping those counts to our network would replace the synthesized volume, our weakest input, with measurement on the highest-stakes roads.
- **Shoulder width.** The provincial road features inventory likely includes shoulder attributes. A 90 km/h highway with a two-metre paved shoulder and one with none are different sports; today only an explicit no-shoulder tag distinguishes them. This is the single change that would most improve credibility with touring cyclists.
- **LiDAR elevation for engineered grades**, starting with the marquee rail trails, where terrain elevation is most wrong and coverage is only a few hundred kilometres of named trail.
- **Water sources for remoteness.** Distance to groceries is a good resupply proxy in the populated south and a poor one exactly where remoteness matters most.
- **The boundary sensitivity analysis** (Limitation 3): for every route, how close each label sits to its nearest band edge, and which single input would flip it. If the answer is a handful of routes, the hard cuts are fine; if it is a third of them, the bands need rethinking or hedged display.
- **Better OpenStreetMap, induced rather than awaited.** The pipeline already knows which ways' missing tags cost the most confidence. Published as a fix-list for the mapping community, that turns our weakest inputs into a concrete community project: the several hundred BC ways where one surface tag most improves cycle-touring data.

9. Data and code availability

The companion dataset, one row per route with every derived metric and the labels actually shown, is freely downloadable as Excel and CSV with a plain-text column guide at bccycletourism.ca/research/. A blank cell means we could not measure that thing, not that it is zero.

The code behind all of this is not yet packaged for published release. If you are trying to reproduce a number, check a method, or build on a piece of it, tell us what you need and we will gladly send the relevant piece: heather@bccycletourism.ca.

Riders looking for the plain-language version of what the labels mean should read bccycletourism.ca/routes/how-we-describe/ instead of this document.

The research here uses AI assistance (Claude models, Anthropic) for data collection, analysis, and writing. Feel free to ask for more details if you are interested.