Okanagan Stream Temperature Risk Tool — Methods Document

Project status, research questions, station inventory, QA/QC confidence, and data gaps

Okanagan Basin Water Board (OBWB)

22 May 2026

Document status

This is a living methods document for the OBWB Okanagan Stream Temperature Risk Tool (project code `oktemp`). It is the operational companion to EFN_MethodsPaper_v2_20260514.docx (the formal methods paper, on OBWB OneDrive), and is updated as the project moves forward.

This v0.12 records the project state as of 2026-05-22, adding to the v0.11 state:

- ERA5-Land warming-trend cross-check run and confirmed; new Appendix H details both satellite analyses. `inst/dev/23_era5land_xcheck.R` was executed (output `inst/dev/out/23_era5land_xcheck_20260522.txt`): ERA5-Land Okanagan-box 2 m temperature, 1950–2024 → +0.324 °C/dec (Mann–Kendall $p = 2.9\times10^{-11}$); 1975–2024 +0.408; 1975–2020 +0.382 vs the AHCCD basin's +0.260 — an independent reanalysis line that corroborates the homogenized-station trend and reproduces the manuscript §3.4 numbers exactly. Earlier v0.10/v0.11 text calling this cross-check "shelved / never run" was wrong (it had been run locally at manuscript time but the artifact was not committed); §E.10 and Appendix G are corrected. The MODIS-LST predictor evaluation was also committed as a reproducible script (`inst/dev/24_lst_predictor_xcheck.R` → `inst/dev/out/24_lst_predictor_xcheck_20260522.txt`). New Appendix H records the methods, results, and provenance of both. No model, app, or data changes.

This v0.11 records the project state as of 2026-05-22, adding to the v0.10 state:

- New Appendix G — a synthesis of the satellite temperature datasets. Pulls the project's findings on each evaluated or cataloged satellite/reanalysis temperature dataset into one place: where each is strong, where it is weak for Okanagan stream-temperature work, and the specific role each could play in a future phase. Consolidates the empirical MODIS-LST result (§E.10), the candidate notes (§E.9), and the product catalog (Appendix F). Headline: the binding constraint is the length of the observed water-temperature record, not satellite availability — so the genuinely promising future temperature uses are spatial (filling ungauged/headwater gaps in an SSN v2) and mechanistic (Landsat lake-surface temperature as a direct driver of the lake-outflow mainstem), not same-day prediction. No model or data changes in v0.11.

This v0.10 records the project state as of 2026-05-22, adding to the v0.9 state:

- MODIS LST evaluated as an M1 covariate and dropped on skill grounds — same outcome as snow. The satellite land-surface-temperature series (`lst_mean`, AppEEARS MOD/MYD11A1) had already been fetched and joined into the M1 panel but had never actually entered the model formula. A held-out 10-fold nowcast cross-validation (2026-05-22, on the LST-available rows) found that adding `s(lst_mean)` as a global smooth gives zero predictive gain (out-of-fold RMSE 0.574 → 0.574 °C; deviance explained 99.3 % unchanged; ΔAIC +1.5), because air `surface_temp` plus AR(1) persistence already absorb essentially all explainable nowcast variance. As a standalone thermal driver (replacing air temp) LST was ≈16 % worse (RMSE 0.668), reflecting cloud-gappy coverage (≈35 % of stream-days) and the ~10:30 / 13:30 overpass-time skin-temperature bias. The term is now an explicit, reversible `fit_m1(use_lst = FALSE)` switch (not a deletion), mirroring `use_snow`; MODIS LST is retained as the app's viz-only thermal-grid context layer. The full evaluation — tests, limitations, and future-use options — is the new §E.10. No change to the deployed model.

This v0.9 records the project state as of 2026-05-19, adding to the v0.8 state — incorporating expert-review feedback (Sheena, Denise Neilsen):

- Management-action framing tightened (§1, this version). v0.8's "controllable management action (Fyke & Weaver 2023)" sentence read as implying thermal management (e.g., reservoir releases). On reviewer feedback (Denise Neilsen, 2026-05-19): in this basin the Okanagan mainstem is the epilimnetic outflow of Okanagan Lake — releases are already warm-surface water with no hypolimnetic outlet capacity, so reservoir release is not an effective thermal lever for the mainstem; small irrigation-storage on tributaries is similarly ineffective (limited storage, warms quickly). The v1 user/decision (locked) is advance to the next drought stage in the next two weeks — i.e. demand-side curtailment is the realistic management action the tool informs (preserving low-flow / cold-baseflow integrity by reducing withdrawals when thermal stress is forecast). v1.x roadmap adds groundwater storage / managed aquifer recharge / baseflow protection as the right long-term lever for cool lower-reach baseflow, plus reach-specific shading/riparian options. §1 "Reporting layer" updated accordingly.
- App entry-tab copy plain-languaged. v0.7/0.8 used "M1 hierarchical GAM (year-block CV RMSE ~0.56 °C)" on the Thermal Risk landing tab — appropriate for the Methods reader, not for the v1 audience (local-government water managers, ONA technical staff, fisheries/limnology reviewers). The Thermal Risk tab now describes the model in plain language; the technical specification and skill numbers remain in §3 / §3.6 of this document and the Methods tab.
- Thermal-grid (MODIS LST) caveat clarified. Reviewer (Denise) flagged surprise at the displayed thermal gain. Cause: ambiguity in the existing caveat. The Thermal grid tab now states explicitly that the layer is the radiometric skin temperature at the satellite's ~10:30 / 13:30 local overpass (the peak diurnal moment, not a daily mean, never the night minimum), and that dry/bare/south-facing surfaces routinely run 10–20 °C above ambient air at that moment — i.e. the apparent thermal gain is expected, not anomalous. Positioned as a thermal-landscape context layer, not a temperature map of the valley.
- Polygon hover-priority fix (UI only). Sheena's review: large Okanagan mainstem watershed was overtaking smaller tributary polygons on hover.

Fixed by rendering named-watershed polygons in area-descending order (largest first / smallest last) so smaller polygons sit on top in event handling. No model/data change.

This v0.8 records the project state as of 2026-05-19, adding to the v0.7 state:

- Water-temperature data architecture verified, and the record now banks (accumulates) instead of mirroring. Re-verified against current HYDAT (2026-05-19): WSC stream temperature is a real-time-only product. Across all of 08NM / 08NL / 08LG, exactly one station (08NM160 Vernon Creek near the mouth, 1976–1980, 5 yr) has any water temperature in the HYDAT historical archive — vs. 263 flow and 92 level station-records. The public real-time endpoint serves a fixed rolling window hard-coded at `days_back = 577L` (`_targets.R`); every cached station shared the identical first observation 2024-10-18, the signature of a rolling buffer, not a record start. Until v0.8 the pipeline mirrored that window — `write_daily_ts_cache()` overwrote the cache every run, so the oldest day rolled off permanently and the record was pinned at ~1.6 yr (and slid forward, losing its front). v0.8 adds `merge_banked_daily()` and a `daily_banked` target: each run unions the fresh QAQC'd pull with the prior bot-committed cache, deduped on (station, date) with the fresh value winning in-window, preserving rolled-off history. Both the time-series cache and the M1 model panel now train on the banked series, so the record accumulates ~1 day/day from 2024-10-18 forward (idempotent across re-runs; schema-guarded; verified by smoke test). This is the fixable half of the §E.1 / §3.6 verification ceiling — see the reframed §E.1.
- Threshold restructured into two explicit tiers (supersedes the v0.7 single "19 °C rainbow-trout" framing). The v0.7 label was imprecise. v0.8 adopts a literature-grounded two-tier scheme (§E.3): Tier 1 (primary, basin-wide, regulatory) = 19 °C general salmonid thermal stress — the BC provincial Approved Water Quality Guideline for Temperature Maximum Daily value (Oliver & Fidler 2001; BC MWLAP 2001), companion MWMT 18 °C; this is the value M4 exceedance, the AR6 likelihood statement, and the alert system compute against (unchanged operationally — 19 °C was already the deployed value). Tier 2 (explicit secondary, Okanagan-specific) = 21 °C sockeye en route migration barrier — Hyatt, Stockwell & Rankin 2003 (CWRJ 28(4):689–713), 30–40 % warm-year mortality; reported and labelled throughout (doc, app UI, station-panel reference lines, map legend) and independently skill-quantified by the §3.6 backtest, but scoped to the mainstem migration corridor and not the primary v1 trigger. OBWB's role is adoption of published values, not derivation. A separately computed 21 °C exceedance probability in the live cache (a schema bump) is the clean v1.x increment, deferred so it can be validated against the full pipeline + tripwire rather than blind-shipped. Decision: OBWB Water Stewardship Director, 2026-05-19.
- Climate context added to §1 (motivation). A homogenized multi-decade

Okanagan air-temperature warming trend (AHCCD) and a single-season two-tier stream thermal-regime baseline (JJAS 2025) now anchor why this tool, why now. See §1 "Climate context".

- Data-infrastructure direction set (v1.x). The banked record is now the authoritative long-term Okanagan stream-temperature dataset; there is no archive to recover it from if lost. Persistence backend is pluggable (the `prior_path` seam): durable today via the bot-committed cache, with the planned v1.x backend an OBWB Google Cloud Storage bucket in **`northamerica-northeast1`** (Montréal — Canadian data residency) with object versioning, and the app moving from shinyapps.io (AWS US) to Cloud Run in the same region via keyless Workload Identity Federation. Container scaffolding is in progress; the move is sequenced (bank→GCS first, app→Cloud Run second) and is additive — it does not block the v1 expert share, which can ship on the current app.

This v0.7 records the project state as of 2026-05-18, adding to the v0.6 state:

- EFN/CEFT threshold decision — resolves the long-standing §E.3 gap. OBWB has adopted a single 19 °C rainbow-trout thermal-stress threshold for v1 (decision by the OBWB Water Stewardship Director, 2026-05-18), consistent with the published BC-guidance value already noted in §E.3. It is no longer a "placeholder pending review": the v1 reporting layer (M4 exceedance + AR6 calibrated likelihood) is built against an adopted threshold. Per-species / per-stream multi-threshold refinement (kokanee ~17 °C, general thermal stress ~21 °C) is deferred to v1.x. App UI + model-state card drop the threshold DRAFT/placeholder framing; the broader tool DRAFT status is retained for the forecast-skill and single-open-water-season verification caveats (unchanged).
- Observed-Tw recency decoupled from the M1 nowcast (cache schema v0.11 → v0.12). For M1 stations `latest_temp`/`latest_date` are the covariate-gated M1 nowcast, which lags the live WSC feed by 1–3 d when the matched ECCC air-temp covariate is late/NA. `build_cache` now also surfaces true observed `latest_temp_obs`/`latest_obs_date` from `stats`; the Thermal-state map colour and the stale ring/badge read the observation, while risk/forecast keep the nowcast. Fixes M1 stations with current data spuriously reading "stale".

This v0.6 records the project state as of 2026-05-15, adding to the v0.5 state:

- Operational warning-skill validated (new §3.6). A rolling-origin backtest over the full verifiable open-water season — refitting M1 on data ≤ origin at each of 40 weekly origins, reducing every forecast to the exact binary the live alert system emits — gives 94 % detection (POD) of 19 °C exceedances at a 26 % false-alarm ratio, Peirce/Heidke skill 0.86 / 0.78. The honest caveat: realized advance notice is short (mean 2.2 d, median 1 d) — a reliable near-term trigger, not a fortnight-ahead planning forecast. Harness at `inst/dev/16_warning_skill.R`; multi-step RMSE harness at `inst/dev/14_forecast_skill.R`.

- **`snow_present`** dropped from the operational M1 spec. The MODIS-snow indicator only joined for a subset of stream-days and halved the usable training panel (4,758 vs 10,662 rows) for zero/negative forecast skill (§E.9, §3.6). Now an explicit, reversible `fit_m1(use_snow = FALSE)` switch (not a deletion). Refit dev.expl 0.9955, mean per-stream RMSE 0.51 °C, tripwire green.
- Deploy hardening. `inst/dev/03_deploy_shinyapps.R` now fails loud and early with the exact missing repo-secret names instead of an opaque rsconnect error.

The v0.5 state, carried forward, was recorded as of 2026-05-14 and included:

- Project goal, research questions, and v1 user/decision
- Three live data sources wired into the `targets` pipeline (WSC real-time, WSC HYDAT historical, ECCC daily air temperature)
- One supplementary data source identified but pending credentials (BC MOE AQUARIUS — draft request email at `docs/correspondence/aquarius_credential_request.`
- One public-API provincial source wired (BC Streamflow Inventory FeatureServer — 39 Okanagan-bbox stations with summary stats)
- Station inventories (30 WSC, 9 ECCC, ≥85 AQUARIUS Okanagan-area, 39 BC SFI Okanagan-bbox)
- BC Freshwater Atlas sub-basin polygons (added v0.2) ingested via `bc–data`, simplified, cached as a static repo asset, wired into the Shiny basin map with polygon-to-station cross-selection
- M1 stream-temperature HGAM fits real Okanagan data with year-block CV RMSE 0.56 °C (v0.3); gaulss family default since v0.5 — row-specific residual SD lets M4's calibrated-likelihood statement spread across the AR6 lexicon instead of collapsing all stations to "exceptionally unlikely"
- M1 14-day forecast wired (v0.5) — `forecast_m1()` runs each pipeline refresh with a DOY-climatology surface_temp covariate; cache statement now legitimately reads "M1 14-day forecast … P(Tw > 19 °C) any day in next 14 …"
- Shiny app: thermal-state + thermal-risk maps, station panel with observed Tw trajectory, observed Q with HYDAT DOY P10–P50 envelope, and modelled forecast trajectory + 90% PI band (v0.5)
- App audit closure: 5-tier audit (Sev 1–5) worked end-to-end with one fix each per item; current state has model provenance + refresh-log table surfaced on the Methods tab, persistent disclaimer modal, observation-staleness markers, PNG download, name-fallback chain, test scaffolding
- Sub-basin polygons tint by AR6 likelihood word on the Risk tab (v0.5 polygon-shading map)
- Model QA/QC backed into `fit_m1()` and `predict_m1()`: physical-range gates per column, output clamps to [-0.5, 30] °C, fit-time diagnostics target with dev.expl / R^2 / per-stream RMSE flags
- Satellite predictor scaffolds: AppEEARS MODIS LST + MODIS Snow Cover tasks submitted, ERA5-Land daily-statistics fetcher chunked by year (latest retry in flight); bake-off harness in `04_model_diagnostics.R`

ready to score them against the ECCC baseline once any returns
- QA/QC confidence per station and data type
- Data and methodological gaps for future consideration

The Markdown source is the source of truth; the `.docx` and `.pdf` siblings are generated artefacts. Version is captured in the filename (`v0.1_20260514`) and at the top of the document.

Code repository: github.com/Okanagan-Basin-Water-Board/oktemp (private)

---

1 Project overview

Goal. Translate real-time and historic stream-temperature data into calibrated 7–14 day estimates of thermal risk for Okanagan streams, using discharge and demand as supporting covariates, to support near-term water-management decisions by local and First Nation governments.

v1 user / decision. A water manager deciding whether to advance to the next drought stage in the next two weeks.

Why a parallel tool. The existing Ecofish-built EFN Explorer reports current flow status against EFN/CEFT thresholds but does not forecast threshold risk, does not couple thermal stress to low flow, and does not represent withdrawal demand. The `oktemp` tool fills those gaps as an OBWB-owned product, complementary to the Explorer rather than replacing it.

Climate context (why now). Two independent lines establish that thermal stress is the rising, decision-relevant signal:

- Long-term warming (homogenized, multi-decade). A regional-anomaly composite of the long homogenized Adjusted & Homogenized Canadian Climate Data (AHCCD) records in the Okanagan — Kelowna, Vernon, Penticton, Summerland — warms at +0.167 °C/decade over 1908–2020 (Mann-Kendall $p \approx 1\times10^{-10}$; ≈ 1.9 °C total) and +0.260 °C/decade over the recent era 1975–2020 ($p \approx 7\times10^{-4}$). Window-matched against published IPCC AR6 global rates, the basin's recent-era warming is ≈ 1.4× the recent global-mean land-and-ocean rate and its full-period warming ≈ 2.1× the 20th-century global-mean rate. AHCCD is ECCC's homogenized, co-located-station-spliced product, so these are like-for-like observed trends (not attribution); the record ends 2020 (normal homogenized-product lag — immaterial to a ~110-yr trend). Reproduced by `inst/dev/22_ahccd_trend.R` (captured artifact `inst/dev/out/22_ahccd_trend_20260519.txt`).
- Present thermal regime (one fully-observed season). Over the single fully-observed open-water season in the record (June–September 2025), against the primary 19 °C tier (general salmonid stress, §E.3) 9 of 21 stream gauges (43 %) crossed at least once (median 68 exceedance days, max 113); against the 21 °C Okanagan sockeye-migration tier (Hyatt et al. 2003), 8 of 21 (38 %) crossed (median 52 days, max 93). Peak weekly mean Tw reached 24.6 °C; the basin-max single daily-mean was 25.3 °C. Lake-fed, regulated mainstem Okanagan River sites ran warm and buffered (mean ≈ 21.8 °C, ≈ 111 days ≥ 19 °C and ≈ 89 days ≥ 21 °C) against cooler tributaries (mean ≈ 13.6 °C, ≈ 13 and ≈ 4 days respectively). The flow–temperature coupling is negative at 16 of 20 gauges (median Spearman $\rho \approx -0.19$) — the drought → low-flow → thermal-stress pathway this tool exists to anticipate. This is a baseline, not a trend (one season; provisional real-time WSC data; daily means understate instantaneous peaks).

Reproduced by `inst/dev/21_stream_regime_baseline.R` (artifact `inst/dev/out/21_stream_regime_baseline_20260519.txt`).

Indicator hierarchy (locked in methods paper v2, 2026-05-14):

- Stream temperature is the analytical indicator. It governs aquatic habitat suitability and constrains cold-water fish, including kokanee and chinook, in their critical spawning windows. Thermal stress is rising under consecutive drought years.
- Discharge is a supporting covariate and contextual signal, not a co-equal indicator. M1 uses it as a predictor; M2 produces percentile/threshold-level context for the user.
- Demand is tertiary in v1, structured to admit incremental improvement.

Reporting layer. Per-stream calibrated exceedance statement (IPCC AR6 likelihood lexicon, after Fyke et al. 2026) paired with a controllable management action (after Fyke & Weaver 2023). The headline output is a sentence, not a risk-matrix cell.

Which management action — explicit (v0.9 clarification, after Denise Neilsen, 2026-05-19). "Controllable management action" in this basin must be honest about the available levers; thermal management by reservoir release is not one of them. The mainstem Okanagan River is the epilimnetic outflow of Okanagan Lake — the Penticton control structure releases warm surface water, with no hypolimnetic-outlet capacity to release cold water; adding more lake water cannot cool the mainstem. Tributaries are mostly unregulated or regulated by small irrigation-storage dams whose limited reservoirs warm quickly and cannot supply a cool-water signal downstream. The lever the v1 tool actually informs is therefore demand-side curtailment — exactly the locked v1 user/decision (advance to the next drought stage in the next two weeks) — which preserves low-flow and cold-baseflow integrity by reducing withdrawals when thermal stress is forecast. The v1.x roadmap adds groundwater storage, managed aquifer recharge, and baseflow protection as the right long-term lever for cool lower-reach baseflow, alongside reach-specific riparian shading. Reservoir-release thermal management is explicitly out of scope for the v1 reporting layer in this basin.

Related work and precedent. USGS Delaware River Basin near-term stream-temperature forecasting (Zwart et al. 2022; deep learning + data assimilation, ~7-day horizon, reservoir-release support — same purpose, different method); USFS NorWeST regional stream-temperature project (Isaak et al. 2017; Spatial Stream Network models on >1M km of stream, the principal methodological alternative retained as a v2 candidate); Pacific Northwest statistical and machine-learning approaches with sub-1 to 1.5 °C RMSE benchmarks (the v1 QA/QC pass criterion in `R/qaqc/pipeline_gates.R::qaqc_rmse_benchmark()`).

---

2 Research questions

The project is built around six research questions. Each maps to a specific element of the implementation.

1. Forecast skill. Can a hierarchical generalized additive model (HGAM) on paired water- and air-temperature data produce 7–14 day daily stream-temperature predictions with prediction-interval coverage close to nominal (90 % observed within nominal 90 % PI), and root-mean-square error within the published 1.0–1.5 °C band for Pacific Northwest stream-temperature models?
2. Coupled vs. flow-only signal. Does coupling thermal risk to flow status change the operational picture relative to the EFN-only nowcast the Explorer currently provides? Specifically, are there current-regime situations where flow status alone would not flag risk but the combined flow–temperature stress index would, and vice versa?
3. Calibrated communication. Does an exceedance probability paired with IPCC AR6 calibrated likelihood language (after Fyke et al. 2026) produce more actionable interpretation by water managers than a percentile or a status colour? Tested by feedback from OBWB and local-government water-staff review.
4. Demand sensitivity. What is the magnitude of difference between the v1 expected-demand and the licensed-full-demand exceedance statements? If demand levers are small relative to climatic variability, demand modelling is lower priority; if they are large, the v2 demand-side roadmap accelerates.
5. Spatial vs. partial-pooling tradeoff. As station density grows, when does the methodological case for Spatial Stream Network modelling (Isaak et al. 2017) overtake the hierarchical-GAM choice for v1? This question becomes answerable as the BC AQUARIUS supplementary stations (Appendix C) come online in the pipeline.
6. Data-gap consequences. Six of thirty WSC Okanagan stations return no real-time water-temperature data on the public endpoint. Which sub-basins lose coverage as a consequence, and does the OBWB logger network (and/or BC AQUARIUS, pending credentials) close those gaps?

3 Methods summary

This document records the methods paper v2 at a glance; consult `EFN_MethodsPaper_v2_20260514.docx` (OBWB OneDrive) for the full design rationale.

3.1 Models

Model	Role	Implementation
M1 Stream-temperature GAM	Analytical core. Hierarchical GAM with `mgcv::bam()`; factor-smooth `s(surface_temp, stream, bs="fs")` for partial pooling across streams; cyclic seasonal smooth; AR(1) thermal-memory term.	`R/models/M1_stream_temp.R`
M1 fallback Mohseni 1998	4-parameter nonlinear logistic air–water regression for data-sparse streams; also a validation benchmark.	`R/models/mohseni.R`
M2 Supporting flow component	Day-of-year percentile envelope (P05/P10/P20/P50) from HYDAT historical record; recession-curve forecast for late-summer low-flow window; SSI as v1.x extension.	`R/models/M2_flow_support.R`, `R/ingest/hydat_historical.R`
M3 Demand overlay (basic)	Licensed-allocation ceiling distributed by seasonal shape; forward-increment overlay, no naturalization; demand-level toggle in the app.	Scaffolded; not yet wired (deferred — see gaps).

Model	Role	Implementation
M4 Threshold-exceedance	Combines M1 with M2 and M3 context; per-stream probability of crossing temperature thresholds over the forecast horizon; mapped to AR6 calibrated likelihood language.	`R/models/M4_threshold.R`, `R/utils/likelihood_lexicon.R`
Baseline current-conditions (interim)	Local Gaussian climatology over the last 14 days, used while paired water+air history accumulates and M1 cannot yet fit.	`R/models/baseline_current_conditions.R`

3.2 Pipeline architecture

A scheduled `targets` ETL refreshes data and predictions; the Shiny application reads the cached outputs only and never pulls live data on load. Pipeline definition: `_targets.R`. Cache outputs: `data/cache/predictions.rds`, `data/cache/doy_envelope.rds`.

3.3 Data sources (live as of 2026-05-14)

1. WSC real-time (`tidyhydat::realtime_ws()`). Public inline CSV endpoint, no authentication. Verified depth: 577 days (2024-10-14 hard cutoff). Water temperature and discharge.
2. WSC HYDAT historical (`tidyhydat::hy_daily_flows()`). Decades of QA'd daily discharge; water-temperature record essentially absent.
3. ECCC daily air temperature (`weathercan::weather_dl()`). Nine active Okanagan stations as the air-temp predictor.

3.4 Spatial reference layer (added v0.2)

The BC Freshwater Atlas (FWA) provides authoritative sub-basin polygons for the Okanagan, ingested via `bcdata` and cached as a static repo asset.

- Okanagan Watershed Group (**`OKAN`**), FWA code `51f20b1a-ab75-42de-809d-bf415a0f9c62`. 1 polygon, 795,417 ha — the basin boundary.
- FWA Named Watersheds INTERSECTS(OKAN), FWA code `ea63ea04-eab0-4b83-8729-f8a93ac688a1`. 402 sub-basin polygons. 30 of these match the named tributaries hosting WSC stations: Mission Creek

(84,510 ha, largest), Vernon Creek (75,245), Trout Creek (74,645), Vaseux, Lambly, Mill, Coldstream, Equesis, Whiteman, and others.

Processing: polygons fetched in BC Albers (EPSG:3005), simplified at 50 m tolerance for sub-basins and 100 m for the basin outline using `sf::st_simplify(preserveTopology = TRUE)`, then reprojected to WGS84 for the leaflet map. Cached size: 1.1 MB `.rds` and 3.5 MB GeoJSON (down from 27 MB unsimplified — vertex reduction of ~90 % with no visible change at zoom 8–12).

Spatial join. Each WSC station is point-in-polygon assigned to a named watershed (`R/ingest/fwa_subbasins.R::join_stations_to_subbasins()`); when a station falls inside multiple overlapping polygons, the smallest enclosing polygon is selected (most specific tributary). Five Okanagan-mainstem stations resolve to the broad Okanagan River polygon (1,552,080 ha) rather than a tributary, which is the correct topological assignment.

Refresh script: `inst/dev/02_refresh_fwa.R` regenerates the cache from `bcdata` when FWA versions change.

3.4 QA/QC framework

Pipeline gates on every refresh (per methods paper §6.6 and `R/qaqc/pipeline_gates.R`):

- Water-temperature plausibility (−0.5 to 30 °C); flagged rows preserved with `flag_t_implausible`.
- Discharge non-negative and below station historic maximum within tolerance.
- Timestamp-gap detection with offline-station flagging.
- Model checks (when M1 is fit): year-block cross-validation, residual autocorrelation, prediction-interval coverage, RMSE benchmarked against sub-1.5 °C PNW reference band, residual-drift monitoring.

3.5 Forecast architecture (added v0.5)

The headline product is a 14-day probability that stream temperature will exceed the configured threshold at each station. Two architectural decisions:

1. Forecast air-temperature input. M1 takes daily air temperature as a covariate. For a 14-day forecast the future air-temperature signal is required. v1 uses a persisted DOY climatology of the matched ECCC station's mean air temperature (rolled per-day across the 14-day horizon). This is the methods-paper §E.7 fallback path: cheap, defensible, no external forecast feed required. The v1.x upgrade target is ECCC GDPS / HRDPS via the MSC GeoMet datamart (delivers a real ensemble forecast rather than a climatological assumption).

2. Heteroscedastic prediction intervals. M1's `family` is `gaulss` (Gaussian location-scale, mgcv general family). The second linear predictor models

`log(sd)` as a smooth of `surface_temp + doy`, so the residual standard deviation varies by row — warm summer days carry wider PIs than cold winter days, which is what M4's calibrated likelihood statement requires to be defensible. (Previous homoscedastic gaussian fit produced PIs so tight that every station collapsed to "exceptionally unlikely" at mid-May, regardless of actual conditions.)

Iterative rollout. `forecast_m1()` runs day-by-day per stream, feeding the previous day's predicted Tw forward as the AR(1) lag. The 7-day trailing mean of air temperature is recomputed in-place as the climatology unfolds.

Probability aggregation. Per-day P(Tw > threshold) is combined into a horizon-max P under independence: `1 - prod(1 - p_day)`. This is a conservative upper bound on the true any-day-in-window probability — true daily Tw is autocorrelated, so independence over-estimates the chance of crossing. v1.x can tighten this by computing the exceedance probability from a joint posterior simulation, which is straightforward once `predict_m1()`'s `n_sims` argument is plumbed through to the horizon.

What renders. The Shiny station panel draws the forecast trajectory as a purple line with a shaded 90 % PI band over the next 14 days, alongside the observed Tw trace and the threshold horizontal line. The dashed vertical line marks the observed/modelled boundary.

3.6 Forecast & warning skill (added v0.6)

This is the decision-maker question, scored head-on: if OBWB had been running this tool, how often would it have given an accurate 14-day advance warning that a modelled stream was about to exceed the threshold — and how often would it have cried wolf or missed?

Data-horizon constraint (stated up front). A verified multi-decade warning-accuracy backtest is impossible, and the tool does not claim one — this is the unfixable half of §E.1 (no deep historical Tw exists to score against). "Accurate" requires observed Tw, and the observed record is a real-time feed (`tidyhydat::realtime_ws`, WSC parameter 5) on a fixed `days_back =` 577L rolling window — every station starting 2024-10-18, ≈ 1.6 yr. HYDAT's deep archive carries essentially no water temperature for these stations: across all of 08NM / 08NL / 08LG exactly one station (08NM160, 1976–1980) has any, verified 2026-05-19. So the skill below is measured over ~one full open-water season: the honest ceiling on what the record can demonstrate today. Note this is now a clock, not a permanent wall — v0.8 banking (§E.1) makes the record accumulate ≈ 1 day/day from 2024-10-18, so a held-out independent open-water season (not multi-decade) becomes verifiable through ordinary calendar time.

Method. Rolling-origin, 40 weekly origins (2025-03-01 → 2026-03-26). At each origin M1 is refit on data ≤ origin (no future leak), the exact operational +1..+14 grid `build_forecast_grid()` serves is built (train-only DOY-climatology covariates, persisted discharge), `forecast_m1()` rolls the AR(1) posterior forward, and each forecast is reduced to the identical binary the live alert system emits (`inst/dev/15_threshold_alerts.R`): warned ⇔ P(any day Tw > T in window) ≥ 0.50, with P computed per-day Gaussian then `1 − ∏(1 − p_day)`. Truth = any observed day in the same window with Tw > T (case skipped if < 7 observed days fall in the window — unverifiable). Scored as a weather warning: hit / miss / false-alarm / correct-negative. Harness: `inst/dev/16_warning_skill.R` (cases → `data/cache/warning_skill_cases.rds`). The companion multi-step RMSE harness is `inst/dev/14_forecast_skill.R` (operational +1 d 0.59 °C, +7 d ≈ 2.0 °C, +14 d ≈ 1.84 °C; 90 % PI coverage 94–100 %, slightly conservative).

Scored on the no-snow operational spec (§E.9), 802 verifiable origin × stream cases per threshold, 21 modelled streams:

Threshold (°C)	Event base rate	POD (detection)	FAR (false-alarm)	CSI	PSS (Peirce)	HSS (Heidke)	Mean lead (d)
16	0.29	0.97	0.35	0.64	0.76	0.66	1.8
18	0.23	0.94	0.28	0.68	0.83	0.75	2.0
19 (primary tier — general salmonid stress, BC WQG)	0.20	0.94	0.28	0.69	0.85	0.76	2.2
20	0.18	0.92	0.27	0.69	0.85	0.77	2.5
21 (sockeye-migration tier — Hyatt et al. 2003)	0.15	0.90	0.35	0.60	0.81	0.70	2.5

At the primary 19 °C tier (general salmonid thermal stress, BC WQG — the value M4/alerts compute against): of 161 cases where a stream actually crossed 19 °C within the next 14 days, the tool issued an advance warning in 152 (94 % detection), missing 9; it raised 59 false alarms across 211 total warnings (28 % FAR). PSS 0.85 / HSS 0.76 confirm genuine skill, not chance. The 9 misses cluster on four mainstem Okanagan River sites (08NM002 / 08NM050 / 08NM116 / 08NM158). At the explicit 21 °C Okanagan sockeye-migration tier (Hyatt et al. 2003) the same backtest gives POD 0.90 / FAR 0.35 / PSS 0.81 / HSS 0.70, mean lead 2.5 d (rarer event, base rate 0.15) — i.e. the tool is independently skill-quantified at both tiers from the same rolling-origin run, even though only the primary tier drives the live exceedance/alert machinery in v1.

Honest caveat — short realized lead time. Mean lead on hits is 2.2 d (median 1 d); only ~11 % of hits were caught ≥ 7 d ahead (max 13 d). This is structural: with the §3.5 DOY-climatology covariate fallback plus AR(1) mean-reversion, the confident threshold-crossing signal mostly emerges once the stream is already near the threshold. The 14-day horizon supplies the window; the actionable warning typically lands 1–2 days out. Recommended framing for the expert share: "~94 % of 19 °C primary-tier (general salmonid stress) exceedances detected at ~28 % false alarms over the one verifiable open-water season — and the 21 °C sockeye-migration tier independently scores ~90 % / ~35 %; advance notice is short (1–2 days typically) — a reliable near-term trigger, not a two-week planning forecast." The two open levers on lead time (real ECCC GDPS/HRDPS air-temperature forecast input per §E.7; a lower trigger probability trading FAR for lead) are v1.x work.

---

4 Current implementation status

Component	Status	Notes
Repo + CI scaffold	Done	`github.com/Okanagan-Basin-Water-Board/oktemp`, private; `renv` lockfile present; `gh auth setup-git` configured for OBWB account.
WSC real-time ingest	Wired	24/30 stations returning live water-temp + discharge; 4 stations skipped gracefully.
WSC HYDAT historical ingest	Wired	26/30 stations with daily Q; median 54-year record.
ECCC air-temp ingest + nearest-station match	Wired	9 Okanagan stations; haversine nearest match per stream.
M2 day-of-year percentile envelope	Wired	P05/P10/P20/P50/P75/P95 with ±7-day smoothing.
M2 recession discharge forecast	Wired (new v0.8)	Brutsaert–Nieber `Q_t = Q_last·e^{-k t}` over the +1..+14 d horizon, re-anchored at the last observed Q; emitted only on a clean late-summer recession (DOY ~182–304, k>0, $R^2 \geq 0.5$). `m2_forecast` target → `data/cache/m2_forecast.rds`; station panel draws the line + 90 % band and the "Forecast (pending)" placeholder is removed. Off-season/non-recession streams show no Q forecast (honest). Skill backtest is a v1.x item (recession needs a 2nd accumulated low-flow season — see §E.1 banking).

Component	Status	Notes
Climatology baseline (interim for M1)	Wired	14-day local Gaussian climatology; will be retired when M1 fits.
M4 threshold-exceedance + likelihood lexicon	Wired	Implemented and unit-tested (15/15 passing).
Shiny app skeleton	Built	Basin map + station panel modules; reads `predictions.rds`; baseline-mode banner explicit.
Sub-basin polygons on the basin map	Done (new in v0.2)	FWA `OKAN` group outline + 402 named-watershed polygons via `bcdata`; simplified to 50 m, cached as static asset; station-to-polygon spatial join produces cross-selection (polygon click selects the contained station, station click selects the dot).
M1 fit on real data	Done (v0.3; no-snow spec v0.6)	Hierarchical GAM, 21 streams. `snow_present` dropped from the operational spec (§E.9) — reversible `fit_m1(use_snow = FALSE)`; training panel recovers 4,758 → 10,662 daily obs (2024-10-15 → 2026-05-15). dev.expl 0.9955, mean per-stream RMSE 0.51 °C (passes §6.6 sub-1.5 °C). M1 beats Mohseni on 21 / 26 streams. Cache flipped from climatology baseline to M1 predictions.

Component	Status	Notes
Forecast & warning skill	Validated over available window (new in v0.6)	Rolling-origin operational backtest (§3.6): 19 °C POD 0.94, FAR 0.28, PSS 0.85, HSS 0.76; mean lead 2.2 d. Scored on ~1 open-water season — honest ceiling for a ~1.6-yr real-time record, not a held-out year (§E.1). `inst/dev/16_warning_skill.R`, `inst/dev/14_forecast_skill.R`.
Shiny station panel time-series plots	Done (new in v0.3)	Click-through fixed (missing `id="navbar"` on `page_navbar()`); panel renders observed Tw trajectory + threshold line + observed Q + HYDAT DOY P10-P50 envelope.
Threshold curves (EFN/CEFT/temperature)	Governance step	OBWB-internal review; sourced from published BC guidance + literature (replaces earlier ONA co-development scoping).
BC MOE AQUARIUS supplementary ingest	Auth-gated	API requires credentials (Appendix C). Catalog is public.
Demand overlay (M3)	Scaffolded	Code path planned; not in v1 pipeline.

Appendix A — WSC hydrometric station inventory

Authorized access set: 30 Okanagan core stations (`08NM*`) on the OBWB-authenticated wateroffice account (provisioned 2026-03-05). Note: as of 2026-05-14, the public inline endpoint (`tidyhydat::realtime_ws()`) returns data for 24 of these 30 stations without authentication; the 6 no-data stations are candidates for the authenticated path once an R client is available.

Sources: `inst/queries/wsc_ws_station_access.csv`, `inst/queries/v1_streams_realtime_temp_`
`inst/queries/v1_streams_realtime_depth.csv`, `tidyhydat::hy_stations()`, `tidyhydat::hy_stn_data_range()`.

Confidence scale.

- High — long HYDAT record (≥30 yr) AND active real-time temperature feed AND recent QA pass.
- Medium — either HYDAT short (<30 yr) OR real-time gap detected OR <90 days realtime depth.
- Provisional — real-time only, no significant HYDAT depth.
- Limited — partial coverage or known data quality concern.
- Missing — no current data on the public endpoint.

Station	Stream / Location	Real-time Tw	Real-time Q	HYDAT Q from	RT depth (days)	Confidence (Tw)	Confidence (Q)
08NM002	OKANAGAN RIVER AT PENTICTON	Y	Y	(see HYDAT)	577	High	High
08NM037	SHATFORD CREEK NEAR PENTICTON	Y	Y	(see HYDAT)	577	High	High
08NM050	OKANAGAN RIVER ABOVE McINTYRE DAM	Y	Y	(see HYDAT)	577	High	High
08NM065	VERNON CREEK AT outlet of Kalamalka Lake	Y	Y	(see HYDAT)	577	High	High
08NM083	OKANAGAN LAKE AT KELOWNA	Y	Y	1943 (H)	577	High	High
08NM084	OKANAGAN LAKE (auxiliary site)	N	—	1943 (H)	0	Missing	Provisional
08NM085	OKANAGAN RIVER NEAR OLIVER	Y	Y	1944	577	High	High
08NM116	MISSION CREEK NEAR EAST KELOWNA	Y	Y	1949	577	High	High

Station	Stream / Location	Real-time Tw	Real-time Q	HYDAT Q from	RT depth (days)	Confidence (Tw)	Confidence (Q)
08NM123	(08NM123 — historic only)	N	—	1959 (1978 end)	0	Missing	Limited
08NM134	CAMP CREEK AT MOUTH	N	Y	1965	0	Missing	High
08NM139	ESPERON CREEK NEAR (Vernon area)	Y	Y	1965	577	High	High
08NM143	KALAMALKA LAKE AT VERNON	Y	Y*	1967 (H)	577	High	Provisional (level)
08NM146	CLARK CREEK NEAR WINFIELD	N	Y	1968	0	Missing	High
08NM158	TROUT CREEK AT TROUT CREEK ROAD	Y	Y	1969	577	High	High
08NM160	VERNON CREEK NEAR (auxiliary)	Y	Y	1969 (Q to 1999)	577	High	Limited (Q ends 1999)
08NM165	LAMBLY CREEK ABOVE intake	Y	Y	1970	577	High	High
08NM171	VASEUX CREEK ABOVE Vaseux Lake	Y	Y	1970	577	High	High
08NM172	PEARSON CREEK NEAR (Westside)	Y	Y	1970	577	High	High
08NM173	(08NM173 — historic only)	N	Y*	1970 (Q to 2023)	0	Missing	Medium
08NM174	WHITEMAN CREEK AT MOUTH	Y	Y	1971	577	High	High
08NM200	INKANEEP CREEK NEAR mouth	Y	Y	1973	577	High	High
08NM232	BELGO CREEK BELOW intake	Y	Y	1976	577	High	High
08NM240	TWO FORTY CREEK NEAR (East Kelowna area)	Y	Y	1983	577	High	High
08NM241	TWO FORTY-ONE CREEK NEAR	Y	Y	1983	577	High	High
08NM242	DENNIS CREEK NEAR	Y	Y	1985	577	High	High
08NM243	(08NM243 — recent install, level only)	N	—	n/a (H only 1991+)	0	Missing	Provisional
08NM247	OKANAGAN RIVER BELOW McINTYRE DAM	Y	Y	2012	577	High	Medium (13-yr)
08NM248	COLDSTREAM CREEK NEAR (recent)	Y	Y	2021	577	High	Provisional
08NM249	B.X. CREEK AT B.X. (recent)	Y	Y	2023	577	High	Provisional

08NM250	VERNON CREEK AT SWALWELL PARK (recent)	Y	Y	2023	577	High Provisional

Notes. (a) The 577-day realtime depth is a hard endpoint cutoff (`tidyhydat::realtime_ws()` returns data back to 2024-10-14 regardless of requested `start_date`). (b) Several headwater stations (08NM248/249/250) have very short HYDAT records and so contribute little to the long-record percentile envelopes used by M2. (c) 08NM134 Camp Creek and 08NM146 Clark Creek report current discharge in real-time but no water-temperature; their thermal context relies on the M1 partial-pooling factor smooth from neighbouring streams once M1 fits.

Appendix B — ECCC climate stations (air-temp predictor)

Curated set: 9 Okanagan daily-resolution ECCC stations, end ≥ 2025 active. Used as the v1 air-temperature predictor for M1 (per methods paper §6.1).

Source: `R/ingest/eccc_air_temp.R::eccc_okanagan_stations()`, refresh via `weathercan::stations()`.

station_id	Name	Climate ID	Elev (m)	Start	End	Confidence
51117	KELOWNA UBCO	1123996	456	2013	2025	Medium (12-yr)
50269	PENTICTON A	1126146	344	2012	2025	Medium (13-yr)
48369	KELOWNA	1123939	433	2009	2025	High (16-yr)
46987	VERNON AUTO	1128582	482	2005	2025	High (20-yr)
979	SUMMERLAND CS	112G8L1	454	1990	2025	High (35-yr)
1041	OSOYOOS CS	1125852	283	1990	2025	High (35-yr)
1046	PEACHLAND	1126070	345	1971	2025	High (54-yr)
1070	VERNON SILVER STAR LODGE	1128584	1586	1970	2025	High (55-yr; alpine)
1039	OLIVER STP	1125766	297	1924	2025	High (101-yr)

Nearest-match assignments (haversine). Each WSC stream station is matched to the nearest ECCC daily station by great-circle distance; assignments are regenerated by the pipeline target `stream_match` in `_targets.R`. Note that the alpine station (Vernon Silver Star Lodge, 1586 m) provides a representative air-temp predictor for high-elevation headwater streams that would be poorly served by valley-floor stations.

Methodological note (v1.x refinement). The methods paper §6.1 specifies “best-fitting per stream by cross-validation”; v1 uses nearest by haversine, with the CV-based selector deferred to v1.x once paired history extends.

Appendix C — BC MOE AQUARIUS supplementary stations

The BC Ministry of Environment AQUARIUS WebPortal (`bcmoe-prod.aquaticinformatics.net`) carries the Provincial Hydrometric Network — separate from the federal WSC network — along with collaborating third-party stations. The OBWB stream-temperature logger network (Project 150) publishes to this portal.

Verified 2026-05-14 by direct probe of the dashboard catalog:

- Total locations in AQUARIUS catalog: 1,903 (province-wide).
- Okanagan-area matches (by keyword): ≥ 85.
- Station-ID prefixes relevant to the Okanagan: **`08NM`** 56 stations, **`3BNM`** 60 stations.
- Mission Creek alone has ≥ 10 monitoring locations in AQUARIUS vs. 1 in the federal WSC active set; one rich station (`2F05P`) carries 21 datasets.
- Other multi-site streams: Middle Vernon Creek (≥ 5), Mill Creek (≥ 4), Equesis Creek (≥ 3).

Access status (blocking). The AQUARIUS Publish REST API (`/AQUARIUS/Publish/v2/`) returns HTTP 401 Unauthorized to anonymous requests. The web UI catalog (`/Data/GetDropDownAll`) is public after disclaimer acceptance. Programmatic ingest requires:

1. A BC MOE AQUARIUS Publish-API user account, or
2. CSV exports via the web UI (one-shot, not pipeline-friendly), or
3. Confirmation that an alternative public endpoint exists.

Action item. Contact BC MOE `hydrometric@gov.bc.ca` to request programmatic-read access for the OBWB-related stations. Once unblocked, this becomes the primary source for OBWB logger water-temperature history, materially closing the data-depth gap on which M1 fitting is presently waiting (see Appendix E).

Appendix D — QA/QC confidence assessment

Confidence assignments above (Appendices A, B) are working judgments based on:

- Record length — HYDAT or weathercan start-year-to-present.
- Recent activity — most recent reporting date relative to today.
- Data completeness — fraction of expected daily readings present in the trailing window.
- Known historical issues — explicit flags in `tidyhydat/hy_stn_remarks()` or notes in the OBWB hydrometric program retrospective.

Pipeline-level checks running on every refresh (`R/qaqc/pipeline_gates.R`):

Check	Function	Threshold	Action on flag
Water-temperature plausibility	`qaqc_water_temp_plausible`	−0.5 °C ≤ Tw ≤ 30 °C	Row flagged, retained for inspection
Discharge plausibility	`qaqc_discharge_plausible`	Q ≥ 0 and Q ≤ 1.5 × historic max	Row flagged, retained
Timestamp-gap detection	`qaqc_timestamp_gaps`	> 24 h gap	Row flagged; station offline if > 48 h
Prediction-interval coverage (when M1 fit)	`qaqc_pi_coverage`	0.85–0.95 nominal	Coverage report attached to fit object
RMSE benchmark (when M1 fit)	`qaqc_rmse_benchmark`	≤ 1.5 °C “acceptable”; ≤ 1.0 °C “competitive” with PNW peer models	Tier returned; fail blocks publication

Drift monitoring (planned for v1.x). Compare recent residual envelope against the long-record envelope; flag stations whose residual distribution shifts beyond a 2-σ band. Relevant given the fourth consecutive drought year — published thermal regressions trained on pre-2020 data may already be drifting.

Appendix E — Data and methodological gaps for future consideration

Numbered so they can be tracked across versions of this document.

E.1 Water-temperature history depth (M1 data-bound)

Gap (verified 2026-05-19). WSC stream temperature is a real-time-only product. The public inline-CSV endpoint serves a fixed rolling window hard-coded at `days_back = 577L` in `_targets.R`; every cached station shares the identical first observation 2024-10-18 — the signature of a rolling buffer, not a record start. The deep, QA/QC'd HYDAT archive cannot backfill it: across all of 08NM / 08NL / 08LG, exactly one station (08NM160 Vernon Creek near the mouth, 1976–1980, 5 yr) has any water temperature in HYDAT, vs. 263 flow and 92 level station-records. So the usable record spans ≈ 1.6 yr — one full seasonal cycle — which does not support honest year-block cross-validation against a fully held-out independent year.

This gap has two halves — separate them.

- Unfixable (data physics). There is no deep historical Tw archive to recover. The lone 08NM160 1976–80 stub is a ~44-yr-detached orphan; no amount of engineering retrieves multi-decade in-stream temperature that was never measured or archived for these stations. Depth behind 2024-10-18 is gone.
- Fixable — and fixed in v0.8 (engineering). Until v0.8 the pipeline mirrored the rolling window: `write_daily_ts_cache()` overwrote the cache each run, so even data the project had already seen rolled off and was discarded — the record could never grow past ≈ 1.6 yr and was sliding forward, shedding its oldest day daily. v0.8 adds `merge_banked_daily()` + the `daily_banked` target: every run unions the fresh QAQC'd pull with the prior bot-committed cache, deduped on (station, date), fresh winning in-window, rolled-off history preserved. Both the time-series cache and the M1 model panel consume the banked series. From the current floor (2024-10-18) the record now accumulates ≈ 1 day per day. A second independent open-water season therefore becomes verifiable through ordinary calendar time — a held-out-calendar-year backtest is reachable in the natural course of operating the tool, not blocked in principle.

Status (v0.8): half-closed by construction, on a clock for the rest. §3.6 already reports honest within-record out-of-sample skill via rolling-origin backtesting (no leakage): 19 °C primary-tier warning POD 0.94 / FAR 0.28 / PSS 0.85 (21 °C sockeye tier POD 0.90 / FAR 0.35 / PSS 0.81); multi-step RMSE +1 d 0.59 °C / +14 d ≈ 1.84 °C; 90 % PI coverage 94–100 %. As of 2026-05-19 the record still contains only one verifiable open-water season, so the held-out-year test is not yet possible — banking changes the trajectory, not today's evidence. The one-season provenance caveat still travels with every skill number quoted now; the difference from v0.7 is that this caveat now has an expiry, not a permanent ceiling.

Why it still matters today. A single-season backtest cannot show inter-annual transfer (a hot-drought year vs. a cool-freshet year). Until ≥ 2 independent open-water seasons have accumulated, calibrated likelihood statements should be presented with the one-season provenance attached.

Resolution paths.

- (a) BC MOE AQUARIUS unblocking (Appendix C). The OBWB logger network on AQUARIUS likely carries multi-year water-temperature records for at least some `08NM*` and `3BNM*` stations — the only path that adds historical depth rather than waiting for it. Highest-leverage resolution.
- (b) Time-based accumulation — implemented (v0.8). No longer prospective: `merge_banked_daily()` banks each daily pull; ≥ 2 paired open-water seasons accrue by ordinary operation. Persistence is durable today via the bot-committed cache; the pluggable `prior_path` seam moves to an OBWB Google Cloud Storage bucket (`northamerica-northeast1`, object versioning) as the v1.x hardening — Canadian data residency for what is now the authoritative long-term Okanagan stream-temperature record.
- (c) OBWB direct logger ingest. If logger CSVs are available outside AQUARIUS (e.g., OneDrive shared with the Project 150 team), direct file ingest is fast and also adds historical depth.

E.2 Six WSC stations report no real-time temperature

Gap. 08NM084, 08NM123, 08NM134, 08NM146, 08NM173, 08NM243 return no data on the public endpoint as of 2026-05-14.

Sub-basin consequence. Camp Creek (08NM134) and Clark Creek (08NM146) are otherwise active (returning discharge) but missing on the thermal axis. Their thermal context in v1 falls to the M1 factor-smooth (partial pooling from neighbours).

Resolution path. Authenticated wateroffice account (provisioned to OBWB 2026-03-05, currently held in reserve in `.Renviron`). Awaits a maintained R client or a custom `httr2` wrapper.

E.3 EFN / CEFT / temperature threshold curves

Status (resolved for v1 as a two-tier scheme, 2026-05-19; supersedes the v0.7 single "19 °C rainbow-trout" framing). The decision layer M4 evaluates exceedance probability against an explicit, literature-grounded two-tier threshold. Both values are adopted from published sources — OBWB derives no number.

- Tier 1 — primary, basin-wide, regulatory: 19 °C, general salmonid thermal stress. This is the BC provincial Approved Water Quality Guidelines for Temperature (BC Ministry of Water, Land and Air Protection 2001, ISBN 0-7726-4624-4; consultant's technical report Oliver, G.G. & L.E. Fidler 2001, Towards a Water Quality Guideline for Temperature in the Province of British Columbia, Aspen Applied Sciences Ltd. for BC MELP) Maximum Daily Temperature = 19 °C for freshwater aquatic life in streams (companion MWMT = 18 °C). It coincides with the same guideline's rainbow-trout rearing-optimum ceiling (Table 2: 16.0–18.0 °C) plus the +1 °C fish-bearing-stream tolerance. This is the value the M4 exceedance layer, the AR6 calibrated-likelihood statement, and the alert system compute against in v1 — unchanged operationally, since 19 °C was already the deployed value. It is the right basin-wide choice: applicable to all 21 modelled gauges, chronic-sublethal/habitat-protective, and the strongest institutional footing for a public water-board tool that informs curtailment (it is the provincial regulatory guideline). The v0.7 label "rainbow-trout thermal stress" is corrected to "general salmonid thermal stress (BC WQG)" — the guideline is salmonid-assemblage-protective, not species-specific.
- Tier 2 — explicit secondary, Okanagan-specific, mainstem-corridor: 21 °C, sockeye en route migration barrier. Hyatt, Stockwell & Rankin (2003, Canadian Water Resources Journal 28(4): 689–713, doi:10.4296/cwrj2804689; with Stockwell & Hyatt 2007, OBWB Okanagan Lake Water Science Forum) document that adult Okanagan River sockeye migration halts above ≈ 21 °C and resumes below it, with en route delay and 30–40 % mortality in warm years. This tier is labelled and reported throughout — methods doc, app UI, station-panel reference line, map legend — and is independently skill-quantified by the same §3.6 rolling-origin backtest (POD 0.90 / FAR 0.35 / PSS 0.81; mean lead 2.5 d). Scoping caveat (stated honestly): Hyatt's 21 °C is an adult-sockeye en route migration block in the mainstem Okanagan River corridor, not a general-stream sublethal level; it is biologically on-target on the mainstem migration route and informational (not a stress proxy) on the small tributaries sockeye do not use. It is therefore an explicit secondary tier — surfaced for the basin's flagship anadromous species and the controllable mainstem flow/temperature management action — and deliberately not the primary v1 trigger.

Why two tiers, not one. The literature converges on two distinct management bands: ~16–19 °C is the chronic-sublethal/habitat-protection band that BC WQG, EPA Region 10 (2003, 7-DADM 16 °C core rearing / 18 °C migration), and Oregon/Washington standards share; ~21 °C is the acute migration-impairment / approaching-incipient-lethal band (McCullough 1999; McCullough et al. 2001 — daily max < 19–20 °C to avoid direct lethality, incipient lethal 21–22 °C for adult Columbia chinook/steelhead; Hyatt et al. 2003 for Okanagan sockeye specifically). A single number cannot serve both a precautionary basin-wide drought trigger and a species-specific migration-barrier signal; the two-tier scheme keeps the regulatory/precautionary primary while making the Okanagan sockeye consequence explicit.

v1.x refinement path. (a) Add a separately computed 21 °C (and finer per-species/per-life-stage) exceedance probability to the live prediction cache — a cache-schema bump, deferred to v1.x so it is validated against the full pipeline + tripwire rather than blind-shipped. (b) Extend to per-life-stage curves from BC WQG Table 2 (sockeye migration optimum 7.2–15.6 °C; sockeye/kokanee spawning 10–13 °C; kokanee ≈ 17 °C) plus the Okanagan thermal-barrier literature. Sourced from published BC guidance and peer-reviewed Okanagan literature (no longer routed through ONA co-development, as of v0.5).

E.4 ECCC station coverage at high elevation

Gap. The ECCC set has a single alpine station (Vernon Silver Star Lodge, 1586 m). High-elevation headwater streams in the south basin (e.g., upper Trout Creek, headwaters of Mission Creek) may be poorly served by valley-floor air-temp matches.

Resolution paths.

- Add ClimateBC gridded output as a per-stream extracted predictor.
- Add the Trinity-/Ecofish-operated weather stations if available.
- Add additional ECCC stations once `weathercan::stations_dl()` is refreshed.

E.5 No demand-side data in v1

Gap. v1 predicts the observed (current-withdrawal-regime) stream condition. It does not model demand. Forecasts are conditioned on status-quo withdrawals.

Resolution path. v1.x: licensed-allocation ceiling + seasonal demand shape (M3). v2: AWDM integration (Agriculture Water Demand Model, provincial), return-flow accounting, scenario exploration across withdrawal levels.

E.6 Spatial Stream Network modelling deferred

Gap. Network autocorrelation (flow-connected and flow-unconnected) is not represented in M1 in v1. Isaak et al. (2017) NorWeST shows SSN consistently outperforms nonspatial models on correlated stream networks.

Resolution path. v2 candidate per methods paper §6.3 and §9. Pre-requisite is a denser station set, partially closed by unblocking BC AQUARIUS (Appendix C). R implementation via the `SSN2` package.

E.7 Forecast air-temperature input

Status (v0.5): partially closed. The persisted-DOY-climatology fallback is wired in production (see §3.5). For each ECCC station, the daily-air-temperature DOY mean is computed from the available history (currently ~1.5 yr from the WSC realtime window) and rolled out 14 days forward as M1's surface_temp covariate. The cache statement legitimately reads "M1 14-day forecast" instead of "current-conditions estimate". One row per (stream, day t+1..t+14) is written to `data/cache/m1_forecast.rds` and rendered as the modelled trajectory + 90 % PI band on the Shiny station panel.

Remaining gap. DOY climatology assumes future air temp ~ historical-DOY-mean, which is a strong assumption during marine-heatwave / drought years. The v1.x upgrade path is a real numerical-weather-prediction feed:

- ECCC GDPS / HRDPS via MSC GeoMet (recommended, ~15 km global / ~2.5 km regional, deterministic + ensemble) — `httr2` client against `dd.weather.gc.ca` MSC datamart.
- NOAA GFS as an external cross-check / fallback.
- Copernicus ERA5-T preliminary for retrospective skill testing (5-day operational lag).

Once a real NWP feed lands, the climatology fallback stays as the graceful-degradation path when the live feed is unavailable.

E.8 Live data assimilation

Gap. The pipeline currently refreshes the full prediction cache on each `tar_make()` call; there is no live data-assimilation step that nudges M1 toward the most recent observations between fits (as the USGS DRB framework does, Zwart et al. 2022).

Resolution path. v2 candidate. A Kalman-filter update or rolling-origin refit can be added once the fit cadence is decided.

E.9 Satellite remote-sensing predictor candidates

Gap. v1's air-temperature predictor is the inverse-distance-weighted air temperature from up to three ECCC valley-floor stations (Penticton, Vernon, Kelowna airports + a handful of secondary sites). At higher elevations and away from the valley floor this is the weakest part of M1: the nearest met station can be 30 km horizontally and 800 m vertically away from the catchment centroid (most acutely for Mission Creek, Trout Creek, Vernon Creek headwaters). The hierarchical GAM partially masks this with the factor-smooth on `(surface_temp, stream)`, but the underlying covariate is still a long-distance extrapolation.

Resolution paths (work in progress, 2026-05-14). See Appendix F for the full satellite catalog. Active wiring as of this version:

- MODIS Land Surface Temperature (**`MOD11A1.061`** + **`MYD11A1.061`**) — AppEEARS task `7ed938e0-9d4c-4d36-9b86-a6f3764739bd`, ingested by `inst/dev/06_ingest_modis_lst.R` into `data/cache/modis_lst_daily.rds` (15,002 station-days, 2024-10-14 → 2026-05-13) and joined into the M1 panel as `lst_day_mean` / `lst_night_mean` / `lst_mean`. Evaluated and dropped from the operational M1 spec (v0.10) — it added no nowcast skill and was worse as a standalone driver. Reversible `fit_m1(use_lst = FALSE)` switch; retained as the app's viz-only thermal-grid layer. Full evaluation, limitations, and the next-iteration worth-it verdict are in §E.10.
- ERA5-Land 2 m air temperature (**`reanalysis-era5-land`**, Copernicus C3S) — scaffolded at `inst/dev/09_fetch_era5_land.R`. ~9 km grid, hourly back to 1950. Submission path needs a free CDS account + Personal Access Token in `.Renviron` as `CDS_KEY`. Provides `t2m_mean / max / min`, `tp`, `sd` (SWE), and `ssrd` daily; pulled monthly into `data/cache/era5_land_daily.rds`.
- MODIS Snow Cover (**`MOD10A1.061`** + **`MYD10A1.061`**) — scaffolded at `inst/dev/07_submit_modis_snow.R / 08_ingest_modis_snow.R`. Provides daily NDSI snow cover + a `snow_present` (NDSI ≥ 40) indicator. Evaluated and dropped from the operational M1 spec (v0.6). Although snowmelt timing is physically a strong control on May–July discharge → stream temperature, the MODIS-snow series only joined cleanly for a subset of stream-days; including `snow_present` as an M1 covariate halved the usable training panel (4,758 vs 10,662 rows) for zero-to-negative forecast skill in the §3.6 rolling-origin backtest (at +14 d, no-snow was equal or better). The term is retained as a reversible `fit_m1(use_snow = FALSE)` switch, not deleted, so it can be re-evaluated once the snow series is denser or gap-tolerantly joined. The physical signal is not disputed — the operational data trade did not pay off on this record.

The selection rule (methods paper §6.1): each candidate is added to the M1 panel and the year-block leave-one-year-out CV is recomputed. The predictor selected for v1 deployment will be the one that minimises CV RMSE on the most stations (≥ 11 of 21), with ties broken in favour of the lower-latency / lower-cost product (ECCC < MODIS LST < ERA5-Land in operational fragility). Outcome (v0.10): under this rule ECCC air temperature is retained — MODIS LST was evaluated and added no skill (§E.10), and ERA5-Land's daily covariate columns remain joined-but-untested in the panel (its 2 m air temperature is ~9 km gridded reanalysis, expected to be largely redundant with the ECCC air predictor already in use).

E.10 Satellite surface-temperature data: evaluation, limitations, and future use

This section records the formal evaluation of satellite surface-temperature data as an M1 predictor — the data, the tests run, the limitations that drove the result, and a verdict on whether it is worth adding in the next iteration. It is the companion to the candidate list in §E.9.

Data and intent. MODIS Land Surface Temperature (`MOD11A1.061` Terra + `MYD11A1.061` Aqua) was acquired via AppEEARS as a 1 km point-sample at all 21 modelled stations, 2024-10-14 → 2026-05-13 (15,002 station-days; `data/cache/modis_lst_daily.rds`), with per-product day/night columns and a combined `lst_mean`. The motivating question (§E.9 gap) was whether satellite skin temperature improves M1 over the ECCC valley-floor air-temperature predictor — most plausibly at headwater stations where the nearest met station is ~30 km / ~800 m distant. The series had been joined into the M1 panel but had never entered the model formula.

Tests run (2026-05-22).

1. Panel-coverage cost. Because mgcv drops any row missing a model term, adding `lst_mean` to the formula cut the complete-case training panel from 10,681 → 3,662 rows (−66 %). All 21 baseline streams survived (85–262 LST rows each); day-only and lagged columns cut harder (`lst_day_mean` −78 %, `lst_mean_lag7` −72 %).
2. Held-out nowcast skill. A 10-fold cross-validation (gaussian `bam`, every spec scored on the same LST-available test points):

M1 spec	Out-of-fold RMSE (°C)	vs air-only
air-only (`surface_temp`)	0.574	—
air + `s(lst_mean)` (augment)	0.574	+0.1 % (nil)
LST as driver (replace air)	0.668	+16.4 %
air-only, full 10.7k training	0.575	+0.2 %

3. Full-data significance. `s(lst_mean)` resolved to edf 1.00 with p = 0.025 — weakly significant — but deviance explained was unchanged at 99.3 % and AIC rose +1.5. The term is statistically detectable yet carries no predictive information.

Interpretation. Air `surface_temp` together with AR(1) day-to-day persistence (`water_temp_lag1`) already absorb essentially all explainable nowcast variance, leaving no residual for LST to capture; and as a standalone driver LST is markedly worse than air temperature. This mirrors the `snow_present` outcome (§E.9): a physically-sensible covariate whose operational data trade does not pay off on this record. LST stays out of M1 behind a reversible `fit_m1(use_lst = FALSE)` switch, and remains the app's viz-only thermal-grid context layer.

Limitations driving the result.

- Cloud gaps / coverage. `lst_mean` is present on only ~35 % of station-days; clear-sky-biased sampling.

- Skin temperature at a fixed overpass. MODIS LST is radiometric skin temperature at the ~10:30 / 13:30 local Terra/Aqua overpass — the peak diurnal moment, not a daily mean and never the night minimum; dry/bare/south-facing surfaces routinely run 10–20 °C above ambient air. It is a poor direct analogue for in-stream water temperature.
- No forecast covariate. The 14-day rollout carries `surface_temp` forward via ECCC DOY climatology (§E.7); there is no LST forecast, so LST cannot contribute to the operational forecast unless an LST climatology is constructed.
- Record length. ~1.6 yr — the same single-open-water-season verification ceiling as the rest of the record (§E.1).

Verdict — is it worth adding in the next iteration? Not yet. Three findings converge:

1. The nowcast is saturated. Persistence + air temperature explain 99.3 % of deviance; there is no headroom for any additional thermal covariate (LST or ERA5-Land) to improve same-day prediction. This is proven, not assumed.
2. The forecast — where covariates could matter — cannot yet be improved or even validated. Covariates carry more weight at +7…+14 d (where recent observed Tw is unavailable to the AR(1) term), but exploiting that needs (a) a forecast LST/ERA5 input (new climatology infrastructure) and (b) more open-water seasons of observed Tw to validate against. The binding constraint is the water-temperature record (~1.6 yr, real-time only), not the satellite data — no covariate fixes a short truth record.
3. ERA5-Land is largely redundant for the predictor question. Its 2 m air temperature (~9 km reanalysis) correlates strongly with the ECCC air predictor already used, so it is unlikely to clear the same 99.3 % bar; its genuine, non-redundant value is continuity (no missing days) and length (1950–) — which matter for the warming-trend cross-check, not the nowcast.

Where satellite data IS worth the effort (outside M1): (i) ERA5-Land 1950–2024 as an AHCCD-independent basin warming-trend corroboration — run 2026-05-22 (`inst/dev/23_era5land_xcheck.R`; output `inst/dev/out/23_era5land_xcheck_20260522.txt`): +0.324 °C/dec 1950–2024, +0.382 1975–2020 (vs AHCCD +0.260), Mann–Kendall $p \leq 1\times10^{-5}$, reproducing the manuscript §3.4 numbers (full detail in Appendix H.2); a paper deliverable, not a model input; (ii) MODIS LST as the app's thermal-landscape context layer — already shipped.

Conditions that would flip the verdict to "worth it":

- A second-plus open-water season of banked Tw (the §E.1 clock), enabling a real held-out forecast test of an enriched covariate set.
- A move to a spatial stream-network model (SSN/SSN2) in v2, where a spatially-continuous covariate (LST or ERA5-Land) earns its place by filling spatial gaps at ungauged/headwater reaches rather than competing

on temporal nowcast variance (see the M1 source header and methods paper §5).

- Gap-tolerant LST (per-stream temporal interpolation or an air-temp regression backfill) that preserves the ~10.7k-row panel — worth building only after one of the above establishes that LST carries forecast-relevant signal.

Recommended next-iteration priority: spend effort on the binding constraint — continue banking Tw and widen station coverage (the six no-data WSC stations via the authenticated wateroffice account) — rather than on satellite covariates the current data cannot yet reward.

Appendix F — Satellite & gridded remote-sensing data catalog

Free / open remotely-sensed and gridded products evaluated for the M1 predictor set. The catalog is intended as a working reference: in-flight products are flagged; the rest are documented for future consideration so subsequent versions of this work can pick up the trail without re-doing the scoping.

F.1 In flight (2026-05-14)

Product	Resolution	Variable	Status / location	Value for OkTemp
MODIS LST `MOD11A1.061` / `MYD11A1.061`	1 km daily	Land surface temp (day + night)	AppEEARS task `7ed938e0-9d4c-4d36-9b86-a6f3764739bd` submitted; ingest script `inst/dev/06_ingest_modis_lst.R`	Independent surface-temperature predictor at the catchment scale; complements valley-floor ECCC stations at headwater elevations
ERA5-Land (Copernicus C3S)	~9 km hourly → daily	2 m T, dewpoint, precip, SWE, SW radiation	Scaffolded `inst/dev/09_fetch_era5_land.R`; needs free CDS PAT in `.Renviron`	Continuous gridded met record; never has missing days; eliminates “nearest-station-is-30-km-away” problem for headwater stations

Product	Resolution	Variable	Status / location	Value for OkTemp
MODIS Snow Cover `MOD10A1.061` / `MYD10A1.061`	500 m daily	NDSI fractional snow + QA	Scaffolded `inst/dev/07_submit_modis_snow.R` / `08_ingest_modis_snow.R`	Snowmelt timing as covariate; physical proxy for the spring discharge cooling effect M1 currently absorbs into the DOY smooth

F.2 Tier 2 — moderate-lift adds for v1.x

Product	Resolution	Why it would help	Access
Daymet v4 (NASA ORNL)	1 km daily	Higher-resolution gridded met than ERA5-Land for the BC domain; same predictor menu (T min/max, precip, SWE) at finer scale	`daymetr` pkg, free
Landsat 8/9 TIRS	100 → 30 m, ~16 d revisit	Lake surface temperature of Okanagan, Kalamalka, Wood, Skaha. Outlet temperature of these lakes drives downstream Penticton / Vernon mainstem stations directly; not just a proxy.	Earth Engine / USGS M2M / AppEEARS

Product	Resolution	Why it would help	Access
VIIRS LST `VNP21A1D`	750 m daily	Cross-check / continuation of MODIS LST past the MODIS retirement window (~2025); same AppEEARS workflow	AppEEARS
GPM IMERG Final	0.1° (~10 km) half-hourly	Basin-mean precipitation independent of the ~5 ECCC valley-floor gauges; especially mid-elevation precip that drives summer baseflow	NASA GES DISC / AppEEARS
ECOSTRESS LST + ET	70 m, irregular diurnal revisit	High-resolution thermal for spatial validation of the MODIS-LST → stream-T scaling; not for routine ingestion	AppEEARS
SNODAS (NOHRSC)	~1 km daily SWE	US-only operationally but extends a few km into BC at the Boundary / Kettle headwaters; could improve drainage-area-weighted SWE for Similkameen-adjacent stations	NSIDC

F.3 Tier 3 — niche / experimental

Product	Use case	Why not in v1
Sentinel-2 NDWI / MNDWI	Lake area changes for small mid-basin impoundments (Mahoney L., Yellow L.)	Effort outweighs likely signal at the FWA polygon scale used here
Sentinel-1 SAR	Wet-snow detection at sub-daily revisit when cloud blocks optical	Niche cloud-gap-fill role; MODIS-A+T combined already covers most days
GEDI / ICESat-2 canopy	Riparian shading	Aggregating canopy height to FWA polygon scale loses the signal
SMAP	9 km L-band soil moisture	Resolution too coarse for the Okanagan tributary scale
GRACE-FO	Basin-scale water storage anomalies	Footprint (~100,000 km²) far too coarse for an Okanagan-only product
Copernicus CMIP6 / CORDEX	Future-climate scenarios	Belongs in M5 future-climate overlay, not in current-state predictor menu

F.4 Known gaps and limits

- Cloud cover on optical thermal products. MODIS LST point sample drops ~30–50 % of days at Okanagan latitude. The combined Terra + Aqua average partially mitigates this; routine v1 use requires a gap-fill rule (e.g., persisted last-good-value or carry-forward of the ECCC predictor).
- Spatial scale mismatch. Most products are scalar grids (ERA5-Land 9 km, MODIS 1 km). For small headwater stations the grid cell is much larger than the actual catchment; the extracted value averages a footprint that includes off-catchment terrain. Catchment-averaged extracts via `terra::extract(weights = TRUE)` over the FWA named-watershed polygon are a future refinement.
- Topographic bias in reanalysis air temperature. ERA5-Land 2 m temperature is interpolated onto a smoothed topography. For deep valleys (entirety of the Okanagan) the 9 km grid underestimates relief; a lapse-rate correction against the actual station elevation is required before this becomes a drop-in replacement for ECCC.
- Latency. ERA5-Land has a ~2-month operational lag for the final product (1-week lag for the preliminary “ERA5T” product); MODIS LST is ~24 h. For a forecast tool that runs daily, MODIS LST is operationally viable;

ERA5-Land is best for retrospective model training plus the ERA5T preliminary for near-real-time.

- Account / licence fragility. Each external endpoint adds an operational dependency: Earthdata Login (NASA AppEEARS), Copernicus CDS (ERA5-Land), USGS / EROS (Landsat). Credentials live in `.Renviron` (gitignored locally) and as GitHub Actions repo secrets for CI; rotation and renewal cadence is undocumented and a v1.x housekeeping task.

Appendix G — Satellite temperature datasets: strengths, weaknesses, and future roles

This appendix consolidates, in one place, what the project has learned about each satellite (and reanalysis) temperature dataset it evaluated or cataloged — where each is strong, where it is weak for Okanagan stream-temperature work, and the specific role each could play in a future phase. It draws together the empirical MODIS-LST evaluation (§E.10), the predictor-candidate notes (§E.9), and the product catalog (Appendix F).

Framing. The binding constraint on this project is the length of the observed water-temperature record (~1.6 yr, real-time only; §E.1), not the availability of satellite covariates. The operational nowcast is already saturated — air temperature plus day-to-day persistence explain ~99.3 % of deviance (§E.10) — so no temperature covariate can improve same-day prediction. The genuine future value of satellite temperature data is therefore spatial (filling ungauged / headwater gaps a station network cannot reach) and mechanistic (lake-surface temperature as a direct driver of the lake-outflow mainstem), realised only in a spatial stream-network model (v2) and/or once the observed record lengthens enough to validate the forecast horizon.

G.1 Dataset-by-dataset summary

Dataset	Resolution (space / time)	Strengths	Weaknesses for stream-T	Status here	Future role
MODIS LST point-sample (`MOD/MYD11A1`)	1 km / daily (day + night)	Independent of the ground network; catchment-scale; ~24 h latency (forecast-viable); free	Cloud gaps (~35 % coverage); skin temperature at a fixed ~10:30/13:30 overpass (biased high on dry surfaces, not a daily mean); poor direct water-T analogue; worse than air as a driver	Tested in M1 (§E.10) — no nowcast skill; dropped (`use_lst = FALSE`)	Gap-filled spatial covariate in an SSN v2; otherwise viz only
MODIS LST grid composite (`lst_grid.tif`)	1 km / ~weekly composite	Intuitive valley-wide thermal-landscape context	Same skin / overpass caveats; not a water-temperature map (user-confusion risk)	In use — the app's Thermal-grid viz tab (the only live satellite consumer)	Keep as context layer; VIIRS as the successor when MODIS retires

Dataset	Resolution (space / time)	Strengths	Weaknesses for stream-T	Status here	Future role
ERA5-Land (Copernicus C3S)	~9 km / hourly→daily, 1950–	Continuous (never a missing day); long record enables the warming-trend cross-check; gridded everywhere (no “nearest station 30 km away”)	Coarse 9 km smooths deep-valley relief (topographic bias; needs lapse correction); it is air temp, largely redundant with the ECCC predictor; ~2-month final latency; CDS account	Joined to the panel but untested as a predictor; daily refresh paused; trend cross-check run 2026-05-22 (App H.2)	(a) AHCCD-independent warming-trend corroboration — done (+0.324 / +0.382 °C/dec; output in `inst/dev/out/`); (b) lapse-corrected gridded met for headwaters in an SSN v2
Daymet v4 (NASA ORNL)	1 km / daily	Finer gridded met than ERA5-Land over the BC domain	Still air temp (same ECCC redundancy); retrospective only	Cataloged, not fetched	Finer gridded-air alternative to ERA5 if a gridded air predictor is ever wanted

Dataset	Resolution (space / time)	Strengths	Weaknesses for stream-T	Status here	Future role
Landsat 8/9 TIRS	100→30 m / ~16 d	High spatial resolution; lake-surface temperature of Okanagan / Kalamalka / Wood / Skaha — and because the mainstem is the epilimnetic outflow of these lakes, this is a direct driver of downstream Penticton / Vernon stations, not a proxy	~16-day revisit too sparse for daily forecasting; cloud-limited; needs the lake-outlet linkage modeled	Cataloged (F.2), not fetched	Most promising temperature use: lake-outlet temperature predictor for mainstem stations (v1.x / v2)
VIIRS LST (`VNP21A1D`)	750 m / daily	MODIS-equivalent daily LST; continuity past MODIS retirement (~2025+); same AppEEARS workflow	Same cloud / skin / overpass caveats; slightly coarser	Cataloged (F.2), not fetched	Drop-in successor to the MODIS LST viz grid

Dataset	Resolution (space / time)	Strengths	Weaknesses for stream-T	Status here	Future role
ECOSTRESS LST (ISS)	70 m / irregular diurnal	Very high resolution; samples different times of day → can probe the overpass-time bias and map fine-scale thermal diversity	Sparse / irregular revisit; not for routine daily ingestion	Cataloged (F.2), not fetched	High-res cold-water-refugia mapping; spatial validation of the LST → stream-T scaling; diel-bias characterisation

G.2 Cross-cutting themes

Where they are weak (shared failure modes). (i) Cloud gaps on every optical thermal product (~30–50 % of days at this latitude); (ii) spatial-scale mismatch — a 1–9 km grid cell is far larger than a small headwater catchment, so the extract averages off-catchment terrain; (iii) skin-vs-water and time-of-day bias — LST is a radiometric skin temperature at a fixed overpass, not a daily-mean water temperature; (iv) topographic bias in coarse reanalysis air temperature over deep valleys; (v) latency and account fragility for retrospective / credentialed products.

Where they are strong (what ground stations cannot do). (i) Spatial continuity / coverage where the ECCC network is sparse — the headwater elevations that are M1's weakest covariate; (ii) independence from the ground network (a genuine cross-check); (iii) uniquely for Landsat on the big lakes, a physically direct driver of the lake-outflow mainstem rather than a proxy.

The decisive lesson. None of these beats the saturated nowcast (air + persistence). Their value is spatial and mechanistic, not temporal — so it is unlocked by a spatial model (SSN v2) and a longer observed record, not by adding another covariate to the current M1.

G.3 Future-use roadmap (by purpose)

1. Trend corroboration (done 2026-05-22 — paper): ERA5-Land 1950–2024 as an AHCCD-independent check on the basin warming rate — run via `inst/dev/23_era5land_xcheck.R` (output

`inst/dev/out/23_era5land_xcheck_20260522.txt`): +0.324 °C/dec (1950–2024), +0.382 (1975–2020), corroborating AHCCD (Appendix H.2).
2. Viz continuity (v1.x): VIIRS LST as the thermal-grid successor when MODIS retires.
3. Mainstem mechanism (v1.x / v2): Landsat TIRS lake-surface temperature → lake-outlet driver for the Penticton / Vernon mainstem stations.
4. Spatial extension (v2): gap-filled MODIS / VIIRS LST + lapse-corrected ERA5-Land / Daymet as spatially-continuous covariates in an SSN / SSN2 model for ungauged headwaters.
5. Refugia / diel research: ECOSTRESS 70 m for fine-scale thermal diversity and overpass-bias characterisation.

Each is gated on the same precondition: a longer observed water-temperature record and/or the v2 spatial-model architecture (methods paper §5). Until then, satellite temperature data stays a context and corroboration resource, not a model input.

Appendix H — Satellite-data analysis: methods, results, and provenance

This appendix records, in reproducible detail, the two empirical satellite-data analyses behind the conclusions summarised in §E.10 and Appendix G: (H.1) the MODIS LST predictor evaluation, and (H.2) the ERA5-Land warming-trend cross-check. Both are ad-hoc analyses (not part of the `targets` pipeline); each has a committed script and writes a captured artifact to `inst/dev/out/`.

H.1 MODIS LST as an M1 predictor — held-out nowcast cross-validation

Question. Does satellite land-surface temperature improve the M1 stream-temperature nowcast over the ECCC air-temperature covariate, or work better as a replacement for it?

Data. `data/cache/modis_lst_daily.rds` — AppEEARS `MOD11A1.061` + `MYD11A1.061` 1 km point-samples at the 21 modelled stations, 2024-10-14 → 2026-05-13 (15,002 station-days), combined into `lst_mean`. Joined to the M1 panel; `lst_mean` is non-missing on only ~35 % of station-days (cloud gaps).

Method. Three model specifications were compared by 10-fold cross-validation, every fold scored on the same LST-available test rows (gaussian `bam`, fREML; `inst/dev/24_lst_predictor_xcheck.R`):

- air-only (the operational mean structure): `water_temp ~ s(surface_temp, stream, bs="fs") + s(log_Q) + s(doy, bs="cc") + s(surface_temp_lag7) + s(precip) + water_temp_lag1`, where `surface_temp` is the ECCC matched-station air temperature;
- air + LST (augment): the above plus a global smooth `s(lst_mean)`;
- LST as driver (replace): the per-stream factor-smooth moved onto `lst_mean`.

The complete-case panel cost of adding `lst_mean` was also measured: 10,681 → 3,662 rows (−66 %) (all 21 streams retained, 85–262 LST rows each), because `mgcv` drops any row missing a model term.

Results. Out-of-fold RMSE on the common LST-available test rows:

M1 specification	OOF RMSE (°C)	vs air-only
air-only (`surface_temp` = ECCC air)	0.574	—
air + `s(lst_mean)` (augment)	0.574	+0.1 % (nil)
LST as driver (replace air)	0.668	+16.4 %
air-only, full 10.7k training	0.575	+0.2 %

Full-data fit: `s(lst_mean)` edf 1.00, $p = 0.025$; deviance explained 99.3 % unchanged; AIC +1.5.

Conclusion. LST adds no nowcast skill — air temperature plus AR(1) day-to-day persistence already explain ~99.3 % of deviance, leaving no residual for LST — and it is materially worse as a standalone driver (cloud-gappy ~35 % coverage; ~10:30 / 13:30 overpass skin-temperature bias). Dropped from M1 behind the reversible `fit_m1(use_lst = FALSE)` switch; retained as the app's thermal-grid visualization layer. Verdict in §E.10; cross-dataset context in Appendix G.

H.2 ERA5-Land independent warming-trend cross-check

Question. Does an independent reanalysis corroborate the AHCCD homogenized-station Okanagan warming trend reported in §3.4?

Data + method. Copernicus ERA5-Land monthly-mean 2 m air temperature (`reanalysis-era5-land-monthly-means`), Okanagan box [N 50.7, W −120.2, S 49.0, E −118.9], 1950–2024, retrieved via the `ecmwfr` R package (`inst/dev/23_era5land_xcheck.R`). Box spatial-mean per month → annual mean (complete years only) → Theil–Sen slope + Mann–Kendall significance per window. Independent of AHCCD by construction (model + data assimilation vs a homogenized station network). Run 2026-05-22; captured output `inst/dev/out/23_era5land_xcheck_20260522.txt`.

Results.

Window	ERA5-Land trend	Mann–Kendall p
1950–2024 (full)	+0.324 °C/dec	2.9×10^{-11}
1975–2024 (recent)	+0.408 °C/dec	4.4×10^{-7}
1950–2020 (AHCCD-matched full)	+0.304 °C/dec	1.4×10^{-9}
1975–2020 (AHCCD-matched recent)	+0.382 °C/dec	9.4×10^{-6}

Comparison. AHCCD basin (homogenized stations): FULL 1908–2020 +0.167 °C/dec; RECENT 1975–2020 +0.260 °C/dec. The two independent lines agree in sign, significance, and order of magnitude, and bracket the recent-era Okanagan rate at ≈ 0.26–0.38 °C/dec (≈ 1.4–2× the contemporaneous global mean). ERA5-Land runs warmer than the valley-floor AHCCD record over the matched window (+0.382 vs +0.260), as expected for a grid-box mean over mixed terrain versus homogenized townsite stations. These values reproduce the manuscript §3.4 figures exactly.

Caveats. ERA5-Land is a grid-box mean (not a point); reanalysis carries its own structural and observing-system trend uncertainty; this is observed change, not formal attribution.

H.3 Reproducibility

Analysis	Script	Captured output	External dependency
MODIS LST predictor CV (H.1)	`inst/dev/24_lst_predictor_xcheck.R`	`inst/dev/out/24_lst_predictor_xcheck_20260522.txt`	None (uses the cached M1 panel)
ERA5-Land trend (H.2)	`inst/dev/23_era5land_xcheck.R`	`inst/dev/out/23_era5land_xcheck_20260522.txt`	Copernicus CDS account + PAT (`CDS_KEY`); ERA5-Land licence

Both scripts print a provenance header (timestamp, git SHA, R version) to their captured output. Neither is in the `targets` pipeline; re-run manually to regenerate.

Changelog

Version	Date	Author	Notes
v0.1	2026-05-14	OBWB + Claude Code	Initial draft. Records project state at end of 2026-05-14 development session: three live data sources wired, BC AQUARIUS catalog probed (access blocked pending credentials), M1 fit deferred pending paired-history depth, Shiny app in baseline-climatology mode.

Version	Date	Author	Notes
v0.2	2026-05-14	OBWB + Claude Code	Added FWA sub-basin spatial reference layer (1 watershed group + 402 named watersheds via `bcdata`, simplified to 50 m, cached static). New §3.4 documents the layer and processing. Shiny basin map now renders polygons with two-way cross-selection between polygon and station marker. Implementation-status table flipped from “Planned” to “Done” for sub-basin polygons. No model or data-source changes.

Version	Date	Author	Notes
v0.3	2026-05-14	OBWB + Claude Code	M1 milestone. Hierarchical GAM fits real Okanagan data (10,644 obs, 21 streams) with year-block CV RMSE 0.56 °C — inside the methods paper §6.6 sub-1.5 °C pass criterion. Predictors added: 7-day rolling-mean surface_temp and daily total precipitation. Critical QAQC fix: WSC's 99999 sentinel-value-for-missing leaked through into daily means, polluting all M1 fits until trapped at `daily_stats_per_station()`. M1 beats Mohseni on 21 / 26 streams. Cache flipped from climatology baseline to M1; cache statement now leads with "M1 stream-temp HGAM prediction (year-block CV RMSE 0.56 °C)". Shiny station panel click-through fixed (missing `id="navbar"` on `page_navbar()` was suppressing `nav_select`); panel renders real Tw trajectory + HYDAT Q envelope. Still deferred: 7-14 day forecast

Version	Date	Author	Notes
v0.4	2026-05-14	OBWB + Claude Code	Satellite & gridded-data scoping. New Appendix F catalogs evaluated remote-sensing products in three tiers, with value, gaps, and access for each. New §E.9 documents three in-flight predictor candidates: (1) MODIS LST AppEEARS task `7ed938e0...` submitted; (2) ERA5-Land Copernicus C3S fetcher scaffolded at `inst/dev/09_fetch_era5_land.R`; (3) MODIS Snow Cover NDSI scaffolded at `inst/dev/07_submit_modis_snow.R` + `08_ingest_modis_snow.R`. Selection rule (best per-stream CV RMSE; ties to lower-fragility product) recorded. Shiny app gained a second “Thermal risk” map tab coloured by IPCC AR6 calibrated likelihood word (Fyke palette) alongside the existing thermal-state map. FWA polygon cache tightened: centroid-within-OKAN filter dropped 93 adjacent / parent

Version	Date	Author	Notes
v0.5	2026-05-14	OBWB + Claude Code	Forecast wired + scope reset to OBWB-only + model & app QA work. ONA co-development language retired from app UI and forward methods copy; project is now solely OBWB-owned. M1 default family flipped from `gaussian` to `gaulss` (mgcv general family — fits via `gam()` since `bam()` rejects general families). The second LP models `log(sd)` as a smooth of `surface_temp + doy`, so M4's calibrated likelihood spread across the AR6 lexicon instead of collapsing all stations to "exceptionally unlikely". M1 14-day forecast wired: `forecast_m1()` runs each pipeline refresh against a DOY-climatology surface-temp covariate (methods §E.7 fallback path); cache statement now leads with "M1 14-day forecast (HGAM, year-block CV RMSE 0.56 °C) … P(Tw > 19 °C) any day in next 14 …" and

Version	Date	Author	Notes
v0.12	2026-05-22	OBWB + Claude Code	ERA5-Land warming-trend cross-check run + confirmed; new Appendix H; corrected the premature "shelved/never-run" wording in §E.10 + Appendix G. Ran `inst/dev/23_era5land_xcheck.R` (output `inst/dev/out/23_era5land_xcheck_20260` ERA5-Land Okanagan-box 2 m temperature 1950–2024 → +0.324 °C/dec (MK p=2.9e-11), 1975–2024 +0.408, 1975–2020 +0.382 vs AHCCD +0.260 — an independent reanalysis corroboration of the AHCCD homogenized-station trend that reproduces the manuscript §3.4 numbers exactly. The MODIS-LST predictor evaluation was also committed as a reproducible script (`inst/dev/24_lst_predictor_xcheck.R` → `inst/dev/out/24_lst_predictor_xcheck_` New Appendix H documents the methods, results, and provenance of both satellite analyses. v0.10/v0.11 had described the ERA5 cross-check as

Version	Date	Author	Notes
v0.11	2026-05-22	OBWB + Claude Code	New Appendix G — satellite temperature dataset synthesis. Consolidates the project's findings on each evaluated / cataloged satellite + reanalysis temperature dataset (MODIS LST point-sample + grid, ERA5-Land, Daymet v4, Landsat TIRS, VIIRS LST, ECOSTRESS): a dataset-by-dataset strengths/weaknesses/status/future-role table (G.1), cross-cutting failure modes vs. unique strengths (G.2), and a purpose-ordered future-use roadmap (G.3). Headline: the binding constraint is the observed water-temperature record length, not satellite availability, so the promising future temperature uses are spatial (un-gauged/headwater gaps in an SSN v2) and mechanistic (Landsat lake-surface temperature as a direct driver of

Version	Date	Author	Notes
v0.10	2026-05-22	OBWB + Claude Code	MODIS LST evaluated as an M1 covariate and dropped on skill grounds — **`use_lst`** switch + new §E.10. The satellite LST series (`lst_mean`, AppEEARS MOD/MYD11A1) was fetched and joined into the M1 panel but had never entered the model formula. A held-out 10-fold nowcast cross-validation (on the LST-available rows) found adding `s(lst_mean)` as a global smooth gives zero predictive gain (out-of-fold RMSE 0.574 → 0.574 °C; deviance explained 99.3 % unchanged; ΔAIC +1.5) — air `surface_temp` + AR(1) persistence already absorb ~all explainable nowcast variance; as a standalone driver replacing air temp, LST was ≈16 % worse (RMSE 0.668), reflecting ≈35 % cloud-gappy coverage and overpass-time skin-temperature bias. Same

Version	Date	Author	Notes
v0.9	2026-05-19	OBWB + Claude Code	Expert-review patches (Sheena + Denise Neilsen). Management-action framing in §1 “Reporting layer” tightened — reservoir-release thermal management is explicitly out of scope for the v1 reporting layer in this basin (epilimnetic-outflow mainstem; no hypolimnetic outlet at Penticton; tributary irrigation storage too small/warm); the lever the tool actually informs is the locked v1 user/decision = demand-side curtailment; v1.x roadmap adds groundwater storage / managed aquifer recharge / baseflow protection + riparian shading. App entry-tab (Thermal Risk) copy plain-languaged — drops “M1 hierarchical GAM / year-block CV RMSE” from the landing tab (kept in Methods §3, §3.6) for the v1 audience. Thermal-grid (MODIS LST) caveat

Version	Date	Author	Notes
v0.8	2026-05-19	OBWB + Claude Code	Data architecture verified + cache banking + climate-context motivation + infra direction. Re-verified against current HYDAT (2026-05-19): WSC Tw is real-time only; across all 08NM/08NL/08LG exactly one HYDAT Tw record exists (08NM160, 1976–1980) vs. 263 flow / 92 level; the real-time feed is a hard-coded `days_back = 577L` rolling window (every station starting 2024-10-18 — a buffer signature, not a record start). The pipeline previously mirrored that window (overwrote the cache each run, discarding even data already seen). New `merge_banked_daily()` + `daily_banked` target union each fresh QAQC'd pull with the prior bot-committed cache, deduped on (station,date), fresh-wins-in-window, rolled-off history preserved; both the time-series cache and the

Version	Date	Author	Notes
v0.7	2026-05-18	OBWB + Claude Code	EFN/CEFT threshold decision + observed-Tw recency fix + CI auto-deploy repair. OBWB adopted a single 19 °C rainbow-trout thermal-stress threshold for v1 (Water Stewardship Director decision); §E.3 reframed from “placeholder pending review” to an adopted v1 threshold (per the BC-guidance value the section already cites), per-species/per-stream multi-threshold deferred to v1.x; app UI + model-state card drop the threshold DRAFT/placeholder framing (broader tool DRAFT status retained for forecast-skill + single-season-verification caveats). Cache schema v0.11→v0.12: `build_cache` surfaces true observed `latest_temp_obs`/`latest_obs_date` from `stats` so the Thermal-state map colour + stale ring/badge read the observation, not the

End of document.