Method LUME-1: Continuous In-Situ Screening for Fecal Contamination in Ambient Recreational Water by Tryptophan-Like Fluorescence (TLF)

A real-time screening method calibrated to fecal-indicator-bacteria criteria

Evan Thomas
University of Colorado Boulder
Virridy, Inc.

September 2026
Working Draft v0.3

Status and intended use. This document is written in U.S. EPA method-report format and is the method specification supporting adoption of TLF as an *alternative indicator/method* under the EPA Recreational Water Quality Criteria (RWQC) *Site-Specific Alternative Recreational Criteria Technical Support Materials for Alternative Indicators and Methods* (EPA-820-R-14-011, December 2014). Under that framework a state, tribe, or territory adopts the alternative indicator/method into site-specific water quality standards, which the EPA Region reviews and approves; the developer supplies the performance characterization, the paired-sample data, and the index-of-agreement (IA) demonstration. Acceptance rests on an IA ≥ 0.70 relationship to an EPA fecal-indicator method (Method 1603, 1600, 1611, or an approved equivalent) on paired *environmental* samples, and single-laboratory (Tier 1) validation is sufficient for site-specific use. This is deliberately *not* proposed as an EPA-numbered method (the numbered-method route requires measuring a defined target) and *not* as a 40 CFR 136 compliance method; the 40 CFR 136 Alternate Test Procedure is not available to a correlative indicator, because it requires measuring the same analyte as the reference. The first site-specific adoption in preparation is Boulder Creek Segment 2b (COSPBO02B) in Colorado, with the City of Boulder, the Colorado Water Quality Control Division, and a Regulation 38 rulemaking before the Water Quality Control Commission as the route to a revised standard. Bracketed items marked `[TBD]` require data from the site-specific validation dataset.

Contents

1.0 Scope and Application

1.1 This method describes a procedure for the continuous, in-situ *screening of fecal contamination* in ambient recreational waters using tryptophan-like fluorescence (TLF). TLF responds to microbial-associated organic matter that is elevated by fecal inputs and microbial activity; it is a broad indicator of fecal contamination rather than an organism-specific enumerator.

1.2 The method produces a *continuous* fecal-contamination screening index. This continuous estimate is the measured quantity: it is validated by its agreement with an EPA-approved fecal-indicator-bacteria (FIB) reference method (Section 14) and is applied by comparison to the applicable regulatory action limit to screen each reporting interval for likely exceedance. The method is *not* proposed as a certified quantitative enumeration of *E. coli*; the continuous output is a fecal-contamination index calibrated to relate to FIB criteria.

1.3 **The exceedance threshold is a configurable operating parameter, not a fixed property of the method.** Because the underlying measurement is continuous, the same calibrated method screens against whatever action limit the responsible jurisdiction applies. Examples include the *E. coli* 126 CFU/100 mL geometric-mean criterion (2012 RWQC / Colorado Regulation 93); other single-sample, statistical-threshold, or beach-action values used in other jurisdictions (e.g., on the order of 235–410 CFU/100 mL, or ∼300 CFU/100 mL as applied in the Chicago area); the 900 CFU/100 mL European bathing single-sample threshold; and enterococci limits in marine and estuarine waters. Selecting or changing the operating threshold does not re-validate the measurement; it only sets the decision point applied to the continuous index. The averaging period and the excursion-frequency provision are likewise properties of the adopting jurisdiction's criterion and not of the method. Colorado expresses the *E. coli* standard as a *two-month* geometric mean with no statistical threshold value (Regulation 31, Table I, footnote 7 [12]), where the 2012 RWQC recommend a 30-day geometric mean with excursions no more than 10% of the time and EPA-820-R-14-011 directs that a site-specific alternative criterion use the same duration and frequency as the 2012 RWQC. Which form a Segment 2b criterion would take is a question for the state and is listed as an open item in Section 14.4.

1.4 Intended use is **screening and prioritization**: real-time, high-frequency detection of likely fecal-contamination events to trigger public advisories, target confirmatory culture sampling, and provide temporal coverage that grab-sample culture methods cannot. It complements, and does not replace, organism-specific culture confirmation.

1.5 **Two modes of use.** The method is applied in two modes. They differ in what enters the estimate, and performance is reported separately for each throughout this document.

(a) *Grab-and-read* (mobile or laboratory). The instrument is carried to the water, or a captured sample is carried to the instrument. The reading is taken on that sample under the unit's own per-unit calibration, and the estimate comes from a regression on the instrument's own channels; nothing external to the instrument enters it. This is the conventional method-comparison setting and it is the mode that supports LUME-1 as a measurement method.

(b) *Continuous in-situ*. The instrument is deployed in the stream and reports on a fixed interval. The estimate combines the fluorescence signal with the independent optical channel that detects fouling, with water temperature and a temperature interaction, and

with a site-level term fitted to that waterbody's own prior FIB results. Ancillary data enters the same framework where it improves the prediction: antecedent rainfall does so at several deployments, and upstream sensors and streamflow have been evaluated for Boulder Creek. The coefficients are fitted on the paired record and refitted as paired data accumulate, rather than set by hand.

Correcting the measurement for temperature or for fouling (Section 10) and tiering an independent indicator such as antecedent rainfall alongside it are different operations, and a submission should keep the two distinct. Which covariates belong in the model that is submitted is settled by the calibration-model freeze in Section 14.4; the Segment 2b in-situ fit as it currently stands carries the sensor channels, the temperature terms and the site term, and no rainfall or streamflow term.

1.6 Applicable matrices: freshwater rivers, streams, and lakes (and, with enterococci calibration, marine and estuarine recreational waters). Performance is characterized over turbidity `[TBD]` NTU and temperature `[TBD]` °C. Waters with turbidity > 10 NTU or rapidly changing temperature require the corrections in Section 4 and site verification per Section 9. The limit of quantification applied to date is **10 CFU/100 mL** as *E. coli*: under EPA-820-R-14-011 the LOQ is the lowest quantity the assay reliably enumerates, that is, the bottom of a validated working range, and not a 10σ detection limit. Confirmation of the LOQ in the recreational matrix is one of the items to be settled before the Step 3 run (Section 14.4).

1.7 Because TLF is not organism-specific, the method is validated by its *agreement with* an EPA fecal-indicator method on paired environmental samples, not by organism confirmation. Under EPA-820-R-14-011 this is an *alternative indicator/method*: TLF is a fecal-associated indicator *substance* (fluorescent dissolved organic matter), a different analyte from the culture indicator, consistent with that framework's provision for a "different indicator organism (or substance) with a different method" (the framework names caffeine and detergent brighteners as example indicator substances). The basis of comparison ("method one") must be EPA Method 1603 (*E. coli* by membrane filtration), Method 1600, Method 1611, or an approved equivalent; agreement is evaluated by the index of agreement (IA) on $\log_{10}$ paired environmental samples (Section 14), and an IA ≥ 0.70 permits use of the unchanged numerical criterion.

1.8 **Adoption pathway.** Adoption is *site-specific*: the alternative indicator/method is adopted into a state's (or tribe's or territory's) water quality standards for a specified waterbody and approved by the EPA Region, supported by at least 30 paired environmental samples within the limits of quantification collected over the range of site conditions. Because the TLF–FIB relationship can be site-dependent, each waterbody requires its own agreement demonstration; the method, calibration, and analysis are fixed and only the site paired-data demonstration varies. The relationship is re-confirmed on the state's triennial water-quality-standards review. Broad use is achieved by accumulating site adoptions, not by a single national approval.

2.0 Summary of Method

An in-situ optical sensor excites the water sample at a center wavelength near 280 nm and measures emission near 350 nm, the excitation/emission pair characteristic of tryptophan-like fluorescence. TLF originates from intrinsic fluorophores of microbial cells and their metabolic byproducts; its

intensity rises with fecal contamination and microbial activity and is correlated with fecal indicator bacteria in natural waters. The raw fluorescence signal is corrected for temperature quench and turbidity attenuation, referenced to a clean-water baseline, and expressed as a continuous fecal-contamination screening index, which a per-unit calibration relates to the applicable FIB criterion for exceedance screening. No reagents, incubation, or sample culturing are required, and a result is produced within seconds of measurement.

3.0 Definitions

TLF
Tryptophan-like fluorescence; emission near 350 nm under excitation near 280 nm.

fDOM
Fluorescent dissolved organic matter; includes humic-like components that can interfere with TLF.

Action limit
The applicable recreational-water criterion for *E. coli* (e.g., 126 CFU/100 mL).

Exceedance
A sample or interval for which the true *E. coli* concentration equals or exceeds the action limit.

QCS
Quality control (stability) standard; a stable fluorophore (e.g., quinine sulfate) used to verify instrument response.

Calibration blank
Clean water of negligible fluorescence used to establish the instrument baseline.

Reproducibility
Agreement among readings of the same water by replicate units or replicate deployments.

Index of Agreement (IA)
The RWQC Alternative-Methods statistic quantifying the agreement between a candidate method and an EPA reference method.

4.0 Interferences

This section receives disproportionate attention because the published TLF literature identifies these interferences as the principal barrier to standardization.

4.1 **Dissolved and humic-like organic matter.** Non-fecal fDOM contributes to fluorescence in the tryptophan region. Controlled by a clean-water baseline (Section 10) and, where required, by site-specific calibration.

4.2 **Turbidity.** Suspended particles scatter and attenuate excitation and emission light, biasing the signal. Corrected using the on-board optical backscatter channel per Section 10.4. This effect is the dominant field interference for high-gain configurations and is corrected per unit.

4.3 **Temperature.** Fluorescence quench varies with temperature. Corrected to a 20 °C reference per Section 10.3.

4.4 **Biofouling.** Growth on the optical window causes signal drift. Controlled by the QC schedule and the fouling diagnostic in Sections 9 and 12.

4.5 **Bubbles and transient particles.** Produce short-duration spikes; removed by the data-quality screen in Section 12.

4.6 **Chlorine / oxidants (chlorinated matrices).** In chlorinated waters, oxidant residual can alter the fluorophore signal; the method has been characterized in paired pre- and post-chlorination samples (Section 14). For ambient recreational fresh water this is generally not applicable, but it is relevant to chlorinated-influent or reuse matrices and is flagged where present.

5.0 Safety

5.1 Standard field water-sampling and electrical-safety precautions apply. The excitation source is a low-power UV LED contained within the sensor housing; direct viewing of the source is avoided. No hazardous reagents are used.

6.0 Equipment and Supplies

6.1 In-situ TLF sensor (Lume; Virridy, Boulder, CO). Excitation: two UV LEDs, nominal peak 273 nm (>10% peak power 262–290 nm), ≈19 mW radiant power each, linear-current driver with 4,096 discrete levels and 100 μs excitation pulses on a 1 ms cycle.

6.2 Emission path: sapphire window, optical bandpass filter (OD ≥6 below 300 nm, >70% transmission near 350 nm), silicon photomultiplier (SiPM) array with tunable bias (single-photon to average-intensity regime), 12-bit ADC/DAC, sampled at 3.2 MHz with 16 excitation cycles averaged.

6.3 Integrated turbidity/submersion channel: time-of-flight (ToF) sensor, ≈940 nm, reporting signal-per-SPAD in kcps (in-water ≈28–140 kcps; air-exposed >200 kcps), which also serves as a submersion detector and UV-safety interlock.

6.4 Co-located water-temperature sensor in the optical head; IP67 housing with an anti-bubble surface treatment. Calibration is applied server-side following cellular upload.

6.5 Reference instrumentation for QC: calibrated thermometer and turbidimeter. Sampling equipment for paired reference analysis (Section 8).

7.0 Reagents and Standards

7.1 **L-tryptophan stock.** Dissolve certified L-tryptophan in reagent-grade (Type I) water to a stock of `[TBD: e.g., 1000 ppb]`; dissolution may require gentle warming and protection from light. Record lot and certificate of analysis for traceability.

7.2 **Tryptophan calibration series.** Dilute the stock to the working series (nominal 0, 0.1, 0.5, 1, 5, 10, 50 ppb) in reagent water. Prepare fresh each calibration session and protect from light and elevated temperature; discard after [TBD: stability/hold time].

7.3 **Stability check standard (QCS).** A stable secondary fluorophore (quinine sulfate in [TBD: 0.05--0.5 M H_2SO_4] at [TBD: concentration], or equivalent) used to verify instrument response between tryptophan calibrations and in continuing verification (Section 9).

7.4 **Turbidity standards.** Formazin or StablCal primary turbidity standards spanning 0–100 NTU, for the turbidity-correction characterization (Section 10.4). *Note: primary standards approximate but do not fully represent natural suspended sediment;* field turbidity performance is confirmed in the validation study.

7.5 **Calibration blank.** Reagent-grade water of negligible tryptophan-region fluorescence, used to establish and verify the clean-water baseline.

7.6 **Storage and expiry.** [TBD: storage conditions and expiry for tryptophan stock/series, QCS, and turbidity standards.]

8.0 Sample Collection, Handling, and Storage

8.1 **In-situ deployment.** The sensor is deployed at a representative location and depth in flowing or standing water, oriented to keep the optical window in continuous submersion and to minimize fouling and bubble entrainment. Siting criteria: [TBD: depth, flow, distance from bank/bed, avoidance of stagnant zones]. Submersion is confirmed by the ToF channel (Section 6) before data are reported.

8.2 **Paired reference samples.** For validation and continuing accuracy checks, grab samples are collected co-located with the sensor optical volume, at the sensor depth, within [TBD: co-location distance/time tolerance] of the sensor reading. Samples are collected in sterile [TBD: 250 mL polypropylene] containers, held on ice, and analyzed within the reference method holding time (culture methods: process within [TBD: 2--6 h]; $\leq$8 h total).

8.3 **Grab-and-read option.** The sensor may be operated on freshly collected samples read in a flow cell or cuvette under the same calibration, at controlled temperature, avoiding bubbles on the optical window.

8.4 **Field records.** Each reading and paired sample is logged with timestamp, location, depth, and operator (Section 8 record). [TBD: data-logging/telemetry and chain-of-custody detail.]

9.0 Quality Control

Quality control is the core of this method's reliability and mirrors the control-centric QC of EPA Method 1611. Acceptance criteria are consolidated in Table 1 (Section 15). Values shown as *(prov.)* are provisional engineering placeholders to be finalized from the validation-study distributions; the reproducibility limit is set from the validation distribution and is not fixed here.

9.1 **Initial calibration verification.** Prior to deployment, verify sensor response against the tryptophan standard curve (linearity $R^2 \geq 0.98$) and the QCS (within ±10% of nominal *(prov.)*). Frequency: at manufacture/pre-deployment and after any repair or firmware change.

9.2 **Continuing calibration verification.** Read the QCS pre- and post-deployment and every `[TBD: e.g., 7 d]` in service; response within ±10% of the unit's initial value *(prov.)*. Out-of-tolerance triggers recalibration and review of intervening data.

9.3 **Calibration blank / baseline.** The clean-water baseline is verified; drift beyond `[TBD]` $\log_{10}$ over `[TBD]` triggers recalibration.

9.4 **Reproducibility (precision) check.** Where feasible, a co-located replicate unit reads the same water; the paired $\log_{10}$ difference is recorded. Acceptance: $\leq$ `[set from validation]` $\log_{10}$ (target ≈0.10). See Section 14 for measured reproducibility and the associated performance limitation.

9.5 **Biofouling check.** The optical backscatter and dark-channel diagnostics are monitored; an out-of-range trend beyond `[TBD]` triggers cleaning and, if unresolved, data flagging.

9.6 **Interference/recovery check.** A spike-recovery procedure (Section 13) confirms recovery within `[TBD: e.g., 80--120%]` *(prov.)* and detects DOM/turbidity interference.

9.7 **Negative/field blanks.** Reagent-water blanks are run at `[TBD]` frequency to confirm no carryover or contamination in grab-and-read operation.

10.0 Calibration and Standardization of Method-Related Instruments

10.1 **Tryptophan standard curve.** Each unit is characterized against the L-tryptophan series (Section 7) over the working range, capturing the full LED×SiPM-bias sweep per standard. The response metric (the bias-response signal at the reference operating point) is fit against concentration. Linearity acceptance $R^2 \geq 0.98$ (demonstrated $R^2 \approx 0.98$ over 0.1–50 ppb and > 0.997 over 0.1–1 ppb); record slope (gain), intercept, and limit of detection.

10.2 **Per-unit calibration.** Each sensor is individually characterized before deployment. Per-unit gain, SiPM bias-vs-temperature response, and clean-water baseline are recorded and are not shared across units. Normalized inter-unit responses differ by less than `[TBD]` after per-unit calibration.

10.3 **Temperature correction.** Fluorescence is normalized to a 20 °C reference using an exponential quench model,

$$s_{20} = s \cdot e^{\rho\,(20-T)},$$

where T is measured temperature and ρ is the per-unit temperature coefficient determined from a controlled temperature series (fluorescence-vs-temperature $R^2 \approx 0.99$ over the characterized range).

10.4 **Turbidity correction.** The temperature-corrected signal is corrected for optical attenuation using the on-board ToF backscatter channel, with a per-unit coefficient determined from a controlled turbidity series (formazin/StablCal; ToF-vs-turbidity $R^2 \approx 0.99$). *Field confirmation against natural suspended sediment is part of the validation study (Section 14).*

10.5 **Screening-index and FIB calibration.** The corrected, baseline-referenced signal is expressed as the continuous fecal-contamination screening index. A calibration model relates the index to FIB concentration of the form

$$\log_{10}(\widehat{\text{FIB}}) = f(\text{index}, \text{covariates}),$$

fit jointly with any residual temperature/turbidity covariate terms (Section 12) to avoid double-counting the explicit corrections above. The screening index is universal across sites via per-unit calibration; the exceedance operating threshold is set per deployment (Section 1). `[TBD: fix the model form and coefficients / reference the versioned calibration model.]`

11.0 Procedure

11.1 Deploy and allow the sensor to equilibrate (warm-up per manufacturer specification).

11.2 Record readings at the selected cadence, each comprising the fluorescence, temperature, and backscatter channels.

11.3 Insert QC readings (QCS, blank) at the scheduled intervals (Section 9).

11.4 Collect paired reference samples on the schedule required for the validation study or continuing accuracy assessment (Section 8).

11.5 Retrieve, clean, and re-verify the sensor at the end of the deployment interval.

12.0 Data Analysis and Calculations

12.1 **Correction chain.** Apply temperature correction (10.3), then turbidity correction (10.4), then baseline referencing.

12.2 **Data-quality screen.** Exclusion has three tiers, and only the first is mechanical. *(a) The install window*, applied automatically: a reading taken before the sensor entered the water at the site, after it was removed, or attributed to the wrong site is not a measurement of that water and is dropped without judgement. *(b) Recoverable effects are corrected, not excluded*: temperature quench (10.3), optical fouling detected on the independent ToF channel (10.4), the step following a physical cleaning, and auto-range transitions each have a correction, and a reading carrying one of them is kept and corrected. *(c) Everything else is a per-case judgement* against stated criteria, with the reasoning recorded and the count reported either way; the default is to keep. The criteria are whether an existing correction already handles the effect; whether the instrument was measuring the intended water at all; whether the channel is physically invalid (railed, pedestal-pinned, or a dark-channel fault) rather than merely reading high; whether the flag came from an operator or from an advisory auto-detect heuristic; and how much the exclusion removes and in which direction it moves the result. An exclusion removing more than about 10% of a calibration set, or one that improves the headline metric, is raised explicitly rather than taken as a default. No reading is dropped for disagreeing with the reference. Intervals under out-of-tolerance QC (Section 9) are flagged on the same basis. An automatic filter enforcing the instrument's own suspect-window flags was

tried and withdrawn: it removed 30% of the calibration set, because those windows open on every fouling flag and fouling is already corrected. The flags are surfaced for inspection and not enforced.

12.3 **Quantitation.** Convert the corrected, baseline-referenced signal to the continuous fecal-contamination screening index using the calibration model. The model may additionally carry temperature and turbidity as covariates to absorb residual, uncorrected dependence after the explicit corrections of §10; the explicit corrections and the covariate terms are fit jointly so the effect is not double-counted.

12.4 **Classification.** Classify each reporting interval by comparing the estimated FIB (or, equivalently, the screening index at the mapped operating point) to the jurisdiction's action limit: an estimate at or above the limit is flagged as likely exceedance. Where a probabilistic call is preferred, the exceedance probability is $P = \Phi\big((\log_{10}\widehat{\mathrm{FIB}} - \log_{10}\mathrm{limit})/\sigma\big)$ with σ the model residual; an operating point at or below the limit may be used as a protective buffer. Where the applicable criterion is expressed as a geometric mean over an averaging period rather than as a single-sample value, the interval estimates are aggregated over that period on the $\log_{10}$ scale before comparison, and any excursion-frequency provision is evaluated on the same record; the averaging period is the jurisdiction's (Section 1.3). `[TBD: fix the operating-point/buffer policy and the aggregation rule once the form of the criterion is settled.]`

12.5 **Reference uncertainty.** Where the method is compared to a culture reference, the reference result is treated as a value with its own measurement uncertainty (e.g., the Quanti-Tray 95% confidence interval), not as an error-free point, consistent with EPA-821-B-10-001 Section 8.4.3. On the paired grab record the tabulated Colilert 95% confidence interval straddles 126 CFU/100 mL for 8.1% of pairs (17/209), so the reference itself does not resolve their category; on the 145-pair Boulder Creek in-situ record the figure is 11.0%, or 13.1% counting the 1000 CFU/100 mL boundary as well. That is a property of the reference, and it is one reason the continuous index of agreement is a sounder basis for adoption than a classification accuracy computed at a single threshold.

13.0 Sample Spiking Procedure

Retained from the EPA molecular-method template as an interference/recovery control.

13.1 **Spike preparation.** A split of the field matrix is amended with a known increment of L-tryptophan (or a characterized fecal-derived standard) at `[TBD: e.g., low ≈1 ppb and mid ≈10 ppb]` above the native level; an unspiked split is read as the matrix background. Read both under the deployment calibration, in replicate (`[TBD:` $n \geq 3$`]`).

13.2 **Recovery.** Recovery $= 100\% \times (\mathrm{measured}_{\mathrm{spiked}} - \mathrm{measured}_{\mathrm{background}})/\mathrm{known\ increment}$. Recovery outside `[TBD: e.g., 80--120%]` *(prov.)* indicates matrix interference (DOM, turbidity, or, in chlorinated matrices, oxidant effects; see Section 4) and triggers site-specific review or recalibration.

13.3 **Coverage.** Spike recovery is characterized across the working range and across the applicable matrices (including at least one turbid and one high-DOM matrix), providing the recovery evidence required by the RWQC Alternative-Methods evaluation (Section 14).

14.0 Method Performance

Performance is reported against the acceptance framework of the RWQC *Alternative Indicators and Methods* Technical Support Materials (EPA-820-R-14-011): a demonstrated, consistent, predictable relationship to an EPA *E. coli* method (Method 1603, 1600, 1611, or an approved equivalent — "method one") on at least 30 paired *environmental* samples within the limits of quantification, quantified by the Index of Agreement (IA) on $\log_{10}$ values via the EPA Alternative Methods Calculator. An IA ≥ 0.70 permits the alternative indicator/method to carry the unchanged numerical criterion; an IA below 0.70 with $R^2 > 0.60$ permits deriving adjusted site-specific criteria by regression. Single-laboratory (Tier 1) validation is sufficient for site-specific adoption. The evaluation framework applied here — including the treatment of reference-method uncertainty (the approved culture references are themselves imprecise and non-equivalent) and the IA-based adoption test — is established and defended in the foundational peer-reviewed evaluation for this method [9]; the underlying sensor design, calibration, and environmental and drinking-water validation are reported in [10] and [11]. This section summarizes the results that bear directly on method acceptance.

The evidence base for the method covers **661 paired samples across 12 water bodies**, with Boulder Creek (Colorado Regulation 38 segment COSPBO02B, 303(d)-listed) as the lead freshwater matrix, complemented by independent field deployments (Seine and Marne, Paris) and laboratory dilution and grab studies. Validation is conducted with **Colorado State University**, with the **City of Boulder** and the Colorado Water Quality Control Division as the site-specific adoption partners for Segment 2b, and with additional wastewater-dilution work (San Diego State University).

For Step 2, applying the 10 CFU/100 mL limit of quantification (Section 1.6) and excluding non-detects leaves **326 eligible pairs**, 170 from the bench and 156 from the field, against the TSM minimum of 30 paired observations within the quantification limits of both assays. The Boulder Creek continuous in-situ branch is separately fitted on **145 paired grabs at six stations**.

The two records are kept separate throughout this section. The grab-and-read comparison and the continuous in-situ record come from the two modes of Section 1.5 and from different models: the first is a linear calibration on the instrument's own channels, the second a fitted model carrying a site term and covariates. Both clear the acceptance threshold, so the Step 3 conclusion holds either way, but the two describe different things and should not be averaged or quoted as a range. The figures below were current at the date of this draft and the paired record is still growing; the dataset of record for a submission is fixed by the freeze in Section 14.4.

14.1 Index of Agreement (EPA Alternative Methods Calculator) — primary

The AltCalc computes Willmott's index of agreement (IA) and R^2 on $\log_{10}$-transformed paired values; IA ≥ 0.70 permits an alternative method to use the same numerical criterion. Because IA and R^2 measure *continuous* agreement over the whole range and are independent of any single threshold, meeting them validates the continuous index for screening at *any* applicable action limit, not only the 126 CFU/100 mL example used for the classification statistics below. The screen met the threshold against *both* EPA reference methods, and agreed with each more closely than the two references agree with each other:

Comparison (reference vs. method)	n	IA	R^2
Colilert vs. Lume	209	**0.96**	0.86
Membrane filtration vs. Lume	206	**0.91**	0.70
Colilert vs. MF (two EPA methods)	153	0.79	0.52

Restricted to values within limits of quantification the IA values are 0.92 (Colilert), 0.87 (MF), and 0.72 (Colilert–MF): each still above 0.70 for the screen. This table is the *grab-and-read* comparison (Section 1.5a): a captured sample read under a per-unit calibration, with the estimate from a regression on the instrument's own channels. Under ISO 17994 the two EPA culture methods are not equivalent to each other, membrane filtration reading about 2.2 times higher than Colilert on the same samples; that is the level of agreement two approved methods reach with each other, and it is the context in which the rows above should be read.

Held-out (field-deployment) subset. On field grabs alone, outside the bench calibration range ($n = 33$), agreement with Colilert remains above threshold (IA 0.91, R^2 0.71) but agreement with MF falls below it (IA 0.52, R^2 0.19). The pooled MF result is therefore buoyed by the wide bench dilution range; on field data alone, *Colilert is the robust reference and MF is not.* This is why the method is validated against **both references with Colilert primary**, and it identifies additional exceedance-range field data as a priority. ($n = 33$ is small; intervals are wide.)

Continuous in-situ operation (Boulder Creek Segment 2b). Scored as the method operates in the water, on 145 paired Boulder Creek observations at six stations with no per-sample handling and no site refitting, the deployed model of Section 1.5b gives:

Comparison (reference vs. method)	n	IA	R^2
Colilert vs. Lume in situ, Boulder Creek	145	**0.79**	0.47

This clears the 0.70 acceptance threshold, and it is about the agreement the two EPA culture methods reach with each other (IA 0.79). It is the harder of the two tests and it is the one that matches the continuous-monitoring use case. Two limitations are stated with it. First, the R^2 of 0.47 is below the TSM's secondary path of 0.60, so the in-situ arm qualifies by index of agreement alone and would not qualify by the regression route; anyone computing R^2 from the workbook will get 0.47. Second, the in-situ figures are in-sample, and leave-one-out and forward-in-time validation are a later pass. Out-of-sample evidence from a different water body is available in the interim: the independent Seine and Marne (Paris) recreational deployment classified the 900 CFU/100 mL bathing threshold at 96.8% overall accuracy on a forward-in-time split (Section 14.2.2).

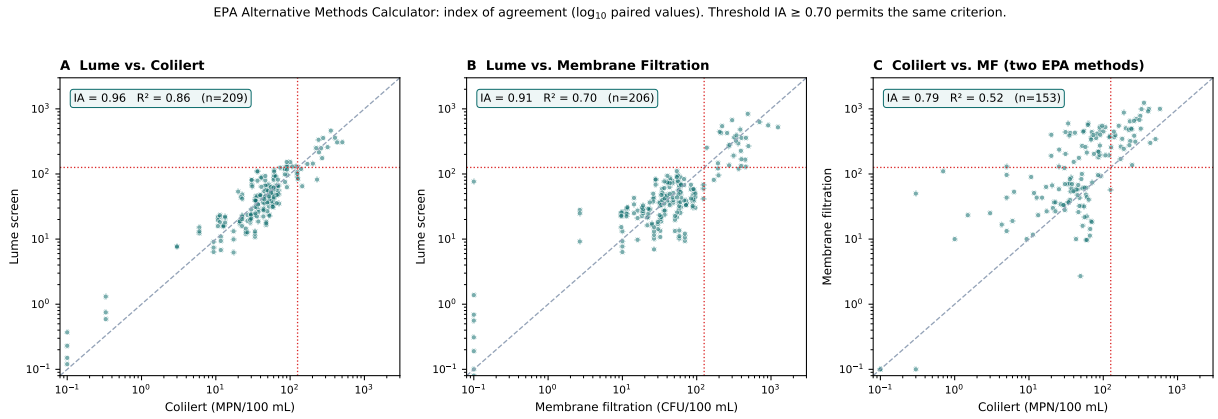

Figure 1. AltCalc agreement ($\log_{10}$ paired values; dashed 1:1, dotted 126 CFU crosshairs). The screen agrees with both references above IA = 0.70 (A, B) and more closely than the two EPA methods agree with each other (C).

14.2 Supporting single-laboratory characterization

14.2.1 **Regression agreement.** Screen vs. Colilert: $R^2 \approx 0.86$, mean bias 0.00 $\log_{10}$, 95% limits of agreement $[-0.42, +0.42]$ $\log_{10}$ ($n = 209$). Screen vs. MF: LoA $[-0.75, +0.75]$ ($n = 206$). Two EPA references vs. each other: $R^2 \approx 0.52$, bias +0.34 $\log_{10}$ (MF reads $\sim$2.2$\times$ higher); under ISO 17994 the two are not equivalent. All three are grab-and-read comparisons (Section 1.5a).

14.2.2 **Exceedance screening (illustrative operating points).** At the 126 CFU/100 mL *E. coli* criterion, the continuous index screened exceedance vs. Colilert ($n = 209$) at sensitivity 0.94, specificity 0.86, balanced accuracy 0.90, $\kappa = 0.47$, and vs. MF ($n = 206$) at sensitivity 0.97, specificity 0.99, balanced accuracy 0.98, $\kappa = 0.94$. The independent Seine and Marne (Paris) deployment, screened at the 900 CFU/100 mL European bathing threshold on a forward-in-time split, gave overall accuracy 96.8%, balanced accuracy 94%; this is out-of-sample evidence from continuous in-situ operation (Section 1.5b) in a different water body. These are two operating points on the same continuous index; screening at other jurisdictional limits (Section 1) uses the same calibration.

14.2.3 **Precision (reproducibility).** Two dimensions are distinguished. *Replicate reproducibility*: the screen's duplicate relative percent difference is $\approx$14%, better than the culture reference ($\geq$26% RPD for Colilert duplicates), and over 75% of predictions fall within the reference's own analytical-uncertainty bounds. *Inter-unit reproducibility* (different co-located units, laboratory bucket water): $\sigma \approx 0.16$ $\log_{10}$, versus reference measurement precision $\sigma \approx 0.03$–0.10 $\log_{10}$ (Colilert) and ≈ 0.056 $\log_{10}$ (MF). *So the screen is at least as reproducible as culture within a unit, but inter-unit reproducibility currently exceeds the reference; reducing it is the primary open performance item and a prerequisite for multi-laboratory validation.*

14.3 Supplementary performance (extended thresholds and matrices)

The following demonstrate robustness beyond the primary recreational fresh-water scope; they are supplementary and do not define the method's intended use (Section 1).

14.3.1 **Three-class categorical** (<10 / 10–100 / >100 MPN/100 mL; laboratory, $n = 334$, Colilert reference): overall accuracy 92.2%, balanced accuracy 95.0%, Cohen's $\kappa = 0.84$. Per class: <10 sensitivity 99.5%, 10–100 sensitivity 80.6%, >100 sensitivity 100%; misclassifications are predominantly in the conservative (risk-over-reporting) direction.

14.3.2 **Lower (drinking-water) thresholds** ($n = 361$, chlorinated and unchlorinated supplies): at 1 CFU/100 mL, balanced accuracy 91.0%, $\kappa = 0.82$ (sensitivity 93.9%, specificity 86.6%); at 10 CFU/100 mL, balanced accuracy 92.0%, $\kappa = 0.84$ (sensitivity 95.3%, specificity 89.0%).

14.3.3 **Chlorination tolerance.** In pre-chlorinated supply water ($n = 38$, 3–200 CFU/100 mL) the screen tracks Colilert with strong positive correlation; in post-chlorinated water ($n = 19$, culture below detection) it slightly over-predicts (residual TLF from inactivated cellular

material), i.e., in the conservative/protective direction. For ambient recreational water this matrix is generally not encountered.

14.3.4 **Chlorine-residual detection (supplementary capability).** Binary detection of chlorine residual (0 vs. >0 ppm; $n = 66$): overall accuracy 85%, $\kappa = 0.70$. Not a method target; noted as a multi-parameter capability relevant to treatment/reuse monitoring.

14.3.5 **Cross-instrument (wastewater dilution, SDSU).** In an eight-step raw-wastewater dilution series (0–100%; San Diego State University), the screen was compared against a benchtop excitation–emission-matrix fluorometer (inner-filter-corrected, as ground truth) and a commercial in-situ CDOM/TLF fluorometer (Turner C3), characterizing response linearity and cross-instrument agreement. `[TBD: add the SDSU comparison statistics.]`

14.4 Remaining steps for a site adoption

For site-specific adoption under EPA-820-R-14-011, single-laboratory (Tier 1) validation is sufficient where that laboratory is the only one analyzing samples for the water quality standards monitoring; multi-laboratory validation is required only where several laboratories or operators do so. A fleet of serialized field instruments operating under one method, one calibration and QC system, and one organization is read here as a single laboratory for that purpose. The TSM separates validation from accreditation and does not, on our reading, require that the reference analysis be run by an independent or accredited laboratory; where it asks that "an independent method should confirm what the method detects in environmental samples," the word independent attaches to the method and not to the laboratory. Both readings are put to the adopting state for confirmation, together with the question of how many laboratories producing reference results would move the work out of Tier 1.

The Step 3 run itself has not been executed. The index-of-agreement figures reported above reproduce the arithmetic the AltCalc tool performs, Willmott's d on $\log_{10}$ pairs. The official workbook has been populated but not run: the Segment 2b copy holds the 145 in-situ pairs, the two indicator names, and 126 CFU/100 mL set as the threshold on both axes, with the output cells empty. Running the tool is Step 3 proper, and it is to be run once the dataset of record is fixed rather than repeatedly while that record is still growing. The workbook and the pairs as a plain CSV are available for the reviewing agency to run or check.

Four things are settled before that run, so that each choice is made before the result of the run is known. Neither the dataset of record nor the calibration model is fixed at the date of this draft, and fixing both is a commitment carried in the site Sampling and Analysis Plan.

14.4.1 **Which reference method is "method one"** (Section 1): whether Colilert/Quanti-Tray, an ATP-approved equivalent, is sufficient as the reference for the comparison, or whether EPA Method 1603 membrane filtration is required. The answer decides whether parallel membrane-filtration sampling has to run at the site, and with it the field and laboratory design in the plan.

14.4.2 **The limit of quantification in the recreational matrix** (Section 1.6), which decides which pairs are eligible to reach the worksheet. The value applied to date is 10 CFU/100 mL.

14.4.3 **The frozen calibration model**, including which covariates it carries. Correcting for temperature and fouling and tiering an independent indicator such as antecedent rainfall are

different operations and the submission keeps them distinct (Section 1.5); the Segment 2b fit as it stands carries no rainfall or streamflow term.

14.4.4 **The choice between the grab-and-read and the continuous in-situ dataset** as the record of the comparison, or the decision to submit both. The two are reported separately in Section 14.1 and are not to be pooled.

Two further items are open but do not gate the run:

14.4.5 **Inter-unit precision reduction** (Section 14.2.3) toward the reference, required before any multi-laboratory or multi-operator deployment.

14.4.6 **The form of the criterion and the assessment unit**, which are the adopting state's to decide: whether a site-specific alternative criterion keeps Colorado's two-month geometric mean or takes the 30-day geometric mean and statistical threshold value that the TSM directs (Section 1.3); whether it applies seasonally, given that a continuous record produces data in every two-month period where a May-to-October grab program does not; and whether the recreation-use assessment runs on the listed reach, the whole segment, or individual stations. These change how the record is assessed rather than how the measurement is made, but they change the monitoring design that follows adoption.

The relationship is re-confirmed on the state's triennial water quality standards review (Section 1), which the TSM requires in order to confirm that the relationship between the indicator/methods has remained valid.

15.0 Tables, Figures, and Validation Data

Table 1 consolidates the quality-control acceptance criteria (Section 9). Values marked *(prov.)* are provisional pending the validation-study distributions; the reproducibility limit is set from that distribution. The Alternative Methods index-of-agreement figure appears in Section 14; calibration curves, confusion matrices, and precision data are to be added from the validation dataset.

Table 1: Quality-control acceptance criteria (Section 9).

QC element	Acceptance	Frequency
Tryptophan linearity	$R^2 \geq 0.98$	at calibration
Limit of detection	≤ 0.1 ppb tryptophan	at calibration
Initial calibration verification (QCS)	within ±10% of nominal *(prov.)*	pre-deployment / post-service
Continuing calibration verification (QCS)	within ±10% of initial *(prov.)*	pre/post + every `[TBD]`
Calibration blank / baseline drift	$\leq$ `[TBD]` $\log_{10}$	`[TBD]`
Inter-unit reproducibility	$\leq$ `[from validation]` $\log_{10}$ (target ≈0.10)	when co-deployed
Biofouling trend (backscatter/dark)	within `[TBD]`	continuous
Spike recovery	80–120% *(prov.)*	per matrix / periodic
ToF in-water backscatter	28–140 kcps; interlock functions	continuous

16.0 Pollution Prevention

The method is reagentless and generates no chemical waste in normal operation, reducing the environmental burden relative to culture-based analysis.

17.0 Waste Management

No hazardous waste is generated. Spent calibration standards are managed per laboratory practice. End-of-life electronics are recycled per applicable regulation.

18.0 References

[1] Sorensen JPR et al. (2015) In-situ tryptophan-like fluorescence: a real-time indicator of faecal contamination in drinking water supplies. *Water Research.*

[2] Khamis K, Sorensen JPR, Bridgeman J et al. (2015) In situ tryptophan-like fluorometers: assessing turbidity and temperature effects for freshwater applications. *Environ. Sci.: Processes & Impacts.*

[3] Sorensen JPR et al. (2018) Tryptophan-like fluorescence as a measure of microbial contamination risk in groundwater. *Sci. Total Environ.*

[4] U.S. EPA (2012) *Recreational Water Quality Criteria*, EPA 820-F-12-058.

[5] U.S. EPA (2014) *Site-Specific Alternative Recreational Criteria Technical Support Materials for Alternative Indicators and Methods*, EPA-820-R-14-011, December 2014. *Primary adoption pathway for this method.* Companion tool: *Alternative Methods Calculator*, EPA 821-B-21-002, with user guide EPA 821-B-21-001 (November 2021).

[6] U.S. EPA Method 1603, *E. coli in Water by Membrane Filtration* (candidate "method one").

[7] U.S. EPA Method 1611.1 / 1609.1, *Enterococci in Water by qPCR* (EPA method-report format model).

[8] U.S. EPA (2010) *Microbiological Alternate Test Procedure (ATP) Protocol*, EPA-821-B-10-001. *Cited for comparability statistics and reference-fallibility concepts only; the 40 CFR 136 ATP is not the adoption pathway, as it requires measuring the same analyte as the reference.*

[9] **Thomas E.** (2026) Evaluating rapid fecal-contamination screening against imprecise, non-unique reference methods: a reference-uncertainty framework and regulatory pathway, applied to tryptophan-like fluorescence. *(in review, ACS ES&T Water). Foundational peer-reviewed evaluation and RWQC adoption evidence for this method.*

[10] Knopp et al. (2026) Advancing continuous in-situ quantification of microbial contamination in environmental waters using tryptophan-like fluorescence: sensor design and validation. *Water Research (in review). Primary sensor design, calibration, and environmental validation.*

[11] Knopp et al. Validating continuous in-situ quantification of microbial contamination in drinking water using tryptophan-like fluorescence. *Water Research X. Drinking-water validation.*

[12] Colorado Water Quality Control Commission. *Regulation No. 31, The Basic Standards and Methodologies for Surface Water*, 5 CCR 1002-31 (effective 31 December 2024), Table I footnote 7; *Regulation No. 38*, 5 CCR 1002-38 (segment COSPBO02B); *Regulation No. 93*.

[13] Colorado Water Quality Control Division. *Colorado's Section 303(d) Listing Methodology, 2026 Listing Cycle*. March 2024.

[14] Colorado Water Quality Control Commission. *Policy 25-1: Advancing External Proposals for Revised Water Quality Classifications and Standards (in Regulations Nos. 31–38) before the Water Quality Control Commission*. Adopted 12 May 2025.

[15] GEI Consultants for the City of Boulder. *E. coli TMDL Implementation Plan*, June 2019, for the Boulder Creek Segment 2b TMDL approved by EPA 27 September 2011.

[16] ISO 17994, *Water quality — Requirements for the comparison of the relative recovery of microorganisms by two quantitative methods*.

19.0 Acronyms

CFU, colony-forming unit; DOM, dissolved organic matter; fDOM, fluorescent dissolved organic matter; IA, Index of Agreement; MF, membrane filtration; MPN, most probable number; NTU, nephelometric turbidity unit; QCS, quality control standard; RWQC, Recreational Water Quality Criteria; TLF, tryptophan-like fluorescence.