https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

Bias corrections and prediction intervals for time-series measurements with fouling and calibration drift

Timothy O. Hodson[1] and Gregory E. Schwarz[1]

[1]U.S. Geological Survey Water Mission Area

Correspondence: Timothy O. Hodson (thodson@usgs.gov)

Abstract. Environmental time-series measurements are subject to systematic errors from fouling and calibration drift, requiring operators to routinely check, clean, and recalibrate their sensors. Information from these checks can then be used to estimate bias corrections and prediction intervals for the measurements. If the checks are exact, the corrections and intervals have direct solutions. More realistically, the checks are noisy, and the error state is estimated with a Kalman smoother. Check information can also be pooled across sensors in a monitoring network to improve estimates for newly deployed sensors. The method relies on checks that operators already perform, and its computation and storage demands are modest, so it integrates readily into existing monitoring programs.

1 Introduction

Environmental monitoring relies on *in situ* sensors deployed over extended periods. Over time, these sensors accumulate systematic errors from fouling and calibration drift. Consequently, they must be periodically checked against known references, cleaned, and recalibrated. Operators can then use this check information to correct the raw measurements. A standard practice is to linearly interpolate the errors between checks (Wagner et al., 2006), which is optimal for exact check measurements but not for noisy ones.

First, we define the three sources of measurement error—fouling, calibration drift, and measurement noise—along with standard procedures for assessing their magnitudes and uncertainties. We then derive bias corrections and prediction intervals, beginning with the idealized case of exact checks. If the error evolves as a generalized Wiener process, conditioning on the checks yields a Brownian bridge (Taylor and Karlin, 1998) with direct expressions for the corrections and intervals. More realistically, the checks are noisy measurements, so we reformulate the error as a state-space model and apply a Kalman smoother (Kalman, 1960; Rauch et al., 1965). We also use empirical Bayes with conjugate priors to pool information from similar sensors across a monitoring network, improving accuracy for newly deployed sensors. Finally, we discuss the computational, storage, and data requirements, which are kept minimal to facilitate real-time, high-frequency applications.

This paper focuses on single-point checks, where a sensor is evaluated against a single reference value. Some sensors exhibit nonlinear behavior that requires multiple reference measurements—a procedure known as a multipoint check. Our approach could be adapted to multipoint checks, but doing so is beyond the scope of this paper.

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

2 Preliminaries

This section defines the three types of measurement error and establishes notation. Let X be the true value of the measurand, and let Y be its measured value; then the measurement error is

$$E = Y - X. \tag{1}$$

We assume measurement errors have three independent components—fouling E_f, calibration drift E_d, and noise E_n—that combine additively to the total error

$$E = E_f + E_d + E_n. \tag{2}$$

Each error component is isolated and assessed using specific check procedures. Fouling is any physical, chemical, or biological buildup on the sensor that causes systematic error. Under stable conditions, fouling error is estimated by comparing measurements taken immediately before and after cleaning the sensor,

$$\tilde{E}_f = Y^b - Y^a, \tag{3}$$

where Y^b and Y^a are the measurements before and after cleaning, respectively, and the tilde indicates a noisy estimator of the true fouling error E_f, modeled as

$$\tilde{E}_f = E_f + \eta_f, \qquad \eta_f \sim \mathcal{N}(0, R_f), \tag{4}$$

where $R_f := \mathrm{Var}[\tilde{E}_f \mid E_f]$ is the noise variance of the fouling check, derived in Section 3. Under changing conditions, a reference sensor is used to account for any environmental change during the cleaning process,

$$\tilde{E}_f = (Y^b - Y^a) - (Z^b - Z^a), \tag{5}$$

where Z^b and Z^a are the before-and-after measurements from the co-located reference sensor (following Wagner et al., 2006, but with the opposite sign convention).

Any systematic bias remaining after cleaning is attributed to calibration drift, which is estimated by comparing a clean measurement against a known standard. Let Z be the standard value and let Y be the clean measurement of Z; then the calibration drift error is given by

$$\tilde{E}_d = Y - Z \qquad \text{(Wagner et al., 2006, but with the opposite sign convention)}, \tag{6}$$

where $\tilde{E}_d$ is the noisy estimator of the true drift error E_d, modeled as

$$\tilde{E}_d = E_d + \eta_d, \qquad \eta_d \sim \mathcal{N}(0, R_d), \tag{7}$$

where $R_d := \mathrm{Var}[\tilde{E}_d \mid E_d]$ is the noise variance of the drift check, derived in Section 3.

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

Finally, any residual error after cleaning and recalibrating the sensor is attributed to measurement noise, which is modeled as a Gaussian random variable with zero mean,

$$E_n \sim \mathcal{N}(0, v_n(x)), \tag{8}$$

where $v_n(x)$ is the variance of the noise as a function of the measurand x. Because x is unknown, we substitute our best estimate given the measurement, $\hat{x} = \mathbb{E}[X \mid Y = y]$. Unlike the fouling and drift variances, the noise variance $v_n(x)$ has no explicit dependence on t. Ideally, $v_n(x)$ is estimated from repeated measurements of a known standard under controlled conditions. This is impossible when sensors truncate readings to a fixed number of significant digits; in that case, the noise can be estimated from the manufacturer's specifications (Appendix A).

Fouling and calibration checks also differ in their reset behavior. During a fouling check, the sensor is cleaned, resetting the fouling error to zero. We assume no residual fouling after cleaning, so this constitutes a *hard reset*. In contrast, recalibration after a drift check is optional and is decided per check; the operator may recalibrate at some drift checks and skip others. When recalibration is performed, the operator applies a correction equal to the measured drift $\tilde{E}_d$, leaving a residual $E_d - \tilde{E}_d = -\eta_d$. This is a *soft reset* because the residual variance is nonzero. The distinction affects how uncertainty propagates through the Kalman filter and smoother (Section 5).

3 Check-Measurement Uncertainty

Noise in the check measurements contributes additional uncertainty that affects the bias corrections and prediction intervals. The next subsections quantify this uncertainty for fouling and calibration checks, deriving the check-noise variances R_f and R_d.

3.1 Fouling-Check Uncertainty

For a fouling check, the sensor is checked against a reference sensor before and after cleaning. Let the principal sensor m and reference sensor r have independent noise,

$$E_{m,n} \sim \mathcal{N}(0, v_n(x)), \qquad E_{r,n} \sim \mathcal{N}(0, v_r(x)), \qquad E_{m,n} \perp E_{r,n}. \tag{9}$$

Let X be the measurand; then the principal and reference measurements (Y, Z) are

$$Y = X_m + E_m, \qquad Z = X_r + E_r. \tag{10}$$

The subscripts on X_m and X_r reflect that the local value of the measurand at each sensor may differ from the true measurand X. Substituting into the fouling-check estimator (Equation 5) gives

$$\begin{aligned}\tilde{E}_f &= (Y^b - Y^a) - (Z^b - Z^a) \\ &= \big((X_m^b + E_m^b) - (X_m^a + E_m^a)\big) - \big((X_r^b + E_r^b) - (X_r^a + E_r^a)\big) \\ &\overset{*}{=} (E_m^b - E_m^a) - (E_r^b - E_r^a),\end{aligned} \tag{11}$$

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

where the equality $(*)$ holds under the assumption that the change in the measurand is the same at both sensors, $X_m^b - X_m^a = X_r^b - X_r^a$, so the two differences cancel.

Next, we expand the remaining measurement errors into their components. Before cleaning, the measurement error includes fouling, drift, and noise ($E_m^b = E_{m,f} + E_{m,d} + E_{m,n}^b$), whereas after cleaning, the fouling component is removed ($E_m^a = E_{m,d} + E_{m,n}^a$). The reference sensor is not cleaned, and we assume its fouling is constant during the check. Expanding $\tilde{E}_f$ and canceling terms gives

$$\begin{aligned}\tilde{E}_f &= (E_{m,f} + \cancel{E_{m,d}} + E_{m,n}^b - \cancel{E_{m,d}} - E_{m,n}^a) - (\cancel{E_{r,f}} + \cancel{E_{r,d}} + E_{r,n}^b - \cancel{E_{r,f}} - \cancel{E_{r,d}} - E_{r,n}^a) \\ &= E_{m,f} + (E_{m,n}^b - E_{m,n}^a) - (E_{r,n}^b - E_{r,n}^a), \end{aligned} \tag{12}$$

which leaves the true fouling error $E_{m,f}$ plus four independent noise terms. The noise variance of the fouling-check estimator is therefore

$$R_f = v_n(\hat{x}^b) + v_n(\hat{x}^a) + v_r(\hat{x}^b) + v_r(\hat{x}^a), \tag{13}$$

where $\hat{x}^b$ and $\hat{x}^a$ are the best estimates of the measurand before and after cleaning, respectively. Thus, for identical sensors with constant noise variance, the fouling-check variance simplifies to $R_f = 4v_n$.

3.2 Calibration-Check Uncertainty

For a calibration check, the sensor is compared against a standard with known precision. Let E_s be the error of the standard, which is independent of the sensor noise E_n,

$$E_s \sim \mathcal{N}\left(0, \sigma_s^2\right), \qquad E_n \sim \mathcal{N}\left(0, v_n(z_s)\right), \qquad E_s \perp E_n, \tag{14}$$

where σ_s^2 is the variance of the standard. Often, σ_s^2 is negligible relative to sensor noise, but not always. Let X be the true value of the measurand in the standard and let E_d be the true drift error of the (clean) sensor; then the measurement Y and standard Z are

$$Y = X + E_d + E_n \qquad \text{and} \qquad Z = X + E_s. \tag{15}$$

Substituting these into the drift-check estimator (Equation 6) gives

$$\tilde{E}_d = (X + E_d + E_n) - (X + E_s) = E_d + (E_n - E_s), \tag{16}$$

and by independence ($E_n \perp E_s$), the noise variance of the drift check is

$$R_d = \mathrm{Var}[E_n] + \mathrm{Var}[E_s] = v_n(z_s) + \sigma_s^2. \tag{17}$$

Reusing the same batch of standards across checks also introduces covariance between checks. Let $\eta_k := \tilde{E}_{d,k} - E_{d,k} = E_{n,k} - E_s$ denote the noise of the kth drift check. Successive checks against the same standard covary according to

$$\mathrm{Cov}[\eta_{k-1}, \eta_k] = \sigma_s^2. \tag{18}$$

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

Standards can also exhibit heterogeneity from mixing effects or gradual degradation, which contributes additional variance or covariance. We do not account for these effects here.

4 Generalized Wiener-Process Error Model

Having defined the measurement procedures and uncertainties, we can now model the time evolution of the error as a stochastic process. When the checks are exact and the error follows a generalized Wiener process, the bias corrections and prediction intervals have direct solutions.

Let $E(t)$ denote the random measurement error at time t, and let $e(t)$ denote its realization, which is measured at $K+1$ check times, $t_0 < t_1 < \cdots < t_K$. We assume the sensor is always checked pre-deployment at t_0, so enumerating from t_0 makes K equal to the number of inter-check intervals, which simplifies notation.

For state variables, the subscript k denotes evaluation at time t_k, so that $E_k := E(t_k)$ and $e_k := e(t_k)$. When needed, we let $k \in \{f, d\}$ to distinguish between fouling and calibration drift. Although the parameters (μ, σ) also take different values for fouling and drift, we omit their subscripts to keep the notation compact.

We assume fouling and calibration drift evolve as generalized Wiener processes, each of the form

$$E_t = \delta_{k-1} E_{k-1} + \mu(t - t_{k-1}) + \sigma\left(W_t - W_{t_{k-1}}\right), \qquad t > t_{k-1}, \tag{19}$$

where δ_{k-1} indicates whether the sensor was cleaned or recalibrated at the previous check, thereby resetting the error process,

$$\delta_{k-1} = \begin{cases} 0, & \text{reset at } t_{k-1} \\ 1, & \text{otherwise} \end{cases}, \tag{20}$$

μ is the drift rate (rate of change of the mean), σ is the diffusion coefficient (with σ^2 the rate of change of the variance), and W_t is a standard Wiener process (without drift) whose increments satisfy

$$W_t - W_{t_{k-1}} \sim \mathcal{N}(0, t - t_{k-1}) \qquad \text{(Taylor and Karlin, 1998).} \tag{21}$$

Note that for a fouling-mode process, $\delta_{k-1} = 0$ for all k, because every check includes physical cleaning.

The checks divide the time series into intervals, each bounded by one or two checks. An interval bracketed by a single check is called "open." For example, the interval after the last check ($t > t_K$) remains open until another check occurs. In contrast, a "closed" interval is bracketed by checks at both ends, denoted as $t \in [t_{k-1}, t_k]$. In real-time monitoring, the present interval is open. Once the next check is recorded, that interval closes and a new open interval begins.

For an open interval, the error distribution is conditionally Gaussian,

$$p(e_t \mid E_{k-1} = e_{k-1}) = \mathcal{N}(m_{t|K}, V_{t|K}), \qquad t > t_{k-1}, \tag{22}$$

with mean and variance given by

$$m_{t|K} := \mathbb{E}[E_t \mid E_{k-1} = e_{k-1}] = \delta_{k-1} e_{k-1} + \mu(t - t_{k-1}) \qquad \text{and} \tag{23}$$

$$V_{t|K} := \mathrm{Var}[E_t \mid E_{k-1} = e_{k-1}] = \sigma^2 (t - t_{k-1}). \tag{24}$$

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

Once the error is observed at time t_k, the conditional distribution of E_t is updated and the interval closes. A generalized Wiener process conditioned to start and end at specified values forms a Brownian bridge (Taylor and Karlin, 1998). To simplify notation, let $\lambda_k(t)$ be the linear-interpolation factor,

$$\lambda_k(t) = \frac{t - t_{k-1}}{t_k - t_{k-1}}, \qquad t \in [t_{k-1}, t_k]. \tag{25}$$

Then the bridge is given by

$$E_t = (1 - \lambda_k(t))\delta_{k-1}E_{k-1} + \lambda_k(t)E_k + \sigma B_t, \qquad B_t \sim \mathcal{N}(0, \phi_k(t)), \tag{26}$$

where $\phi_k(t)$ is the bridge factor,

$$\phi_k(t) = \frac{(t - t_{k-1})(t_k - t)}{t_k - t_{k-1}}. \tag{27}$$

Closed intervals are also conditionally Gaussian,

$$p(e_t \mid E_{k-1} = e_{k-1}, E_k = e_k) = \mathcal{N}\left(m_{t|K}, V_{t|K}\right), \qquad t \in [t_{k-1}, t_k], \tag{28}$$

with mean and variance given by

$$m_{t|K} := \mathbb{E}[E_t \mid E_{k-1} = e_{k-1}, E_k = e_k] = (1 - \lambda_k(t))\,\delta_{k-1}e_{k-1} + \lambda_k(t)e_k \qquad \text{and} \tag{29}$$

$$V_{t|K} := \mathrm{Var}[E_t \mid E_{k-1} = e_{k-1}, E_k = e_k] = \sigma^2\phi_k(t). \tag{30}$$

We can unify the open- and closed-interval expressions by extending the interpolation, bridge, and drift factors as

$$\lambda_k^+(t) := \begin{cases} 1, & t < t_0 \\ \frac{t - t_{k-1}}{t_k - t_{k-1}}, & t \in [t_{k-1}, t_k] \\ 0, & t > t_K \end{cases}, \tag{31}$$

so $\lambda_k^+(t) = 1$ or 0 on open intervals; extending the bridge factor (Equation 30) as

$$\phi_k^+(t) := \begin{cases} (t_0 - t), & t < t_0 \\ \frac{(t - t_{k-1})(t_k - t)}{(t_k - t_{k-1})}, & t \in [t_{k-1}, t_k] \\ (t - t_K), & t > t_K \end{cases}; \tag{32}$$

and extending the drift factor as

$$\zeta_k^+(t) := \begin{cases} (t - t_0), & t < t_0 \\ 0, & t \in [t_{k-1}, t_k] \\ (t - t_K), & t > t_K \end{cases}, \tag{33}$$

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

so the drift term is zero on closed intervals; we can then express the unified mean and variance as

$$m_{t|K} = \left(1 - \lambda_k^+(t)\right)\delta_{k-1}e_{k-1} + \lambda_k^+(t)e_k + \mu\zeta_k^+(t) \qquad \text{and} \tag{34}$$

$$V_{t|K} = \sigma^2\phi_k^+(t). \tag{35}$$

The index k is set by the interval containing t: k is the unique index with $t \in [t_{k-1}, t_k]$ for a closed interval, $k = K+1$ for the open interval $t > t_K$ (so that $e_{k-1} = e_K$ propagates from the most recent check), and $k = 0$ for $t < t_0$ (so that $\lambda_0^+ = 1$ selects e_0). The $t < t_0$ branch is included only for generality; we assume the time series always begins at t_0.

4.1 Parameter Estimation for Exact Checks

If the check measurements are exact, then the drift and diffusion parameters have simple closed-form estimators,

$$\hat{\mu} = \frac{\sum_{k=1}^{K}(e_k - \delta_{k-1}e_{k-1})}{\sum_{k=1}^{K}(t_k - t_{k-1})}, \tag{36}$$

and

$$\hat{\sigma}^2 = \frac{1}{K-1}\sum_{k=1}^{K}\frac{\nu_k^2}{(t_k - t_{k-1})}, \tag{37}$$

where ν_k is the one-step-ahead residual,

$$\nu_k := e_k - \left(\delta_{k-1}e_{k-1} + \hat{\mu}\,(t_k - t_{k-1})\right). \tag{38}$$

When the checks are noisy, this estimator overestimates σ^2, because ν_k^2 incorporates the measurement noise in addition to the diffusion. In that case, the Kalman filter, which models this noise explicitly, should be used instead.

5 Continuous-Discrete Kalman Filter

This section introduces the Kalman filter and smoother for estimating the error-process parameters and states from noisy check measurements. Once the parameters and states are estimated, bias corrections and prediction intervals can be computed by interpolation. Because fouling and calibration drift are independent processes with separate checks and resets, a separate filter and smoother are applied to each, and the total error is the sum of the two.

Let $\tilde{e}_k = e_k + \eta_k$, with $\eta_k \sim \mathcal{N}(0, R_k)$, denote the noisy check measurement of e_k at time t_k, where the check variance R_k is set to R_f or R_d from Section 3. We assume the η_k are independent here; however, modeling correlation, such as that induced by reusing standards, is a straightforward extension of the state-space model (Appendix F).

At a drift recalibration, the operator's correction leaves a residual of $-\eta_{k-1}$. This residual is correlated with the check measurement $\tilde{e}_{k-1} = e_{k-1} + \eta_{k-1}$ because both involve η_{k-1}, so it cannot be treated as independent noise. Substituting $\eta_{k-1} = \tilde{e}_{k-1} - e_{k-1}$ expresses the residual as $-\eta_{k-1} = e_{k-1} - \tilde{e}_{k-1}$. A drift recalibration thus carries forward this estimation error

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

rather than the state e_{k-1} itself. This yields a standard linear-Gaussian state-space model (in the convention of Durbin and Koopman, 2012),

$$e_k = F_k\, e_{k-1} + B_k u_k + q_k, \qquad q_k \sim \mathcal{N}(0, Q_k) \tag{39}$$

$$\tilde{e}_k = H_k\, e_k + \eta_k, \qquad \eta_k \sim \mathcal{N}(0, R_k), \tag{40}$$

with transition matrix F_k encoding the reset type,

$$F_k = \begin{cases} I & \text{drift checks } (d) \\ 0 & \text{fouling checks } (f) \end{cases}, \tag{41}$$

observation matrix $H_k = I$, process-noise variance $Q_k = \sigma^2(t_k - t_{k-1})$, and control input

$$B_k u_k = \mu\,(t_k - t_{k-1}) - F_k\,(1 - \delta_{k-1})\,\tilde{e}_{k-1}, \tag{42}$$

where δ_{k-1} is the reset indicator from Section 4. Here I is the identity matrix, which is 1 for a scalar error process. For fouling, $F_k = 0$ enacts the hard reset (the state restarts at zero each interval), and the operator-correction term vanishes. For drift, $F_k = I$ carries the error forward; at a recalibration ($\delta_{k-1} = 0$) the term $-(1 - \delta_{k-1})\,\tilde{e}_{k-1}$ subtracts the measured check, so that the post-recalibration error reduces to the residual $-\eta_{k-1}$.

Following each check, the error state is updated by conditioning on the past and current checks $(k \mid k)$ using the Kalman (1960) filter,

$$\begin{aligned}
m_{k|k-1} &= F_k\, m_{k-1|k-1} + B_k u_k, \\
V_{k|k-1} &= F_k\, V_{k-1|k-1}\, F_k^\top + Q_k, \\
\nu_k &= \tilde{e}_k - H_k\, m_{k|k-1}, \\
S_k &= H_k\, V_{k|k-1}\, H_k^\top + R_k, \\
G_{k|k-1} &= V_{k|k-1}\, H_k^\top\, S_k^{-1}, \\
m_{k|k} &= m_{k|k-1} + G_{k|k-1}\, \nu_k, \qquad k = 1, \ldots, K \\
V_{k|k} &= V_{k|k-1} - G_{k|k-1}\, S_k\, G_{k|k-1}^\top.
\end{aligned} \tag{43}$$

The recursion is initialized by conditioning on the deployment check under a flat prior on e_0, $m_{0|0} = \tilde{e}_0$ and $V_{0|0} = R_0$. Next, the Rauch–Tung–Striebel smoother (Rauch et al., 1965) further conditions on the subsequent checks, giving $(k \mid K)$,

$$\begin{aligned}
m_{k+1|k} &= F_{k+1}\, m_{k|k} + B_{k+1} u_{k+1}, \\
V_{k+1|k} &= F_{k+1}\, V_{k|k}\, F_{k+1}^\top + Q_{k+1}, \\
G_{k|K} &= V_{k|k}\, F_{k+1}^\top\, V_{k+1|k}^{-1}, \\
m_{k|K} &= m_{k|k} + G_{k|K}(m_{k+1|K} - m_{k+1|k}), \qquad k = K-1, \ldots, 0 \\
V_{k|K} &= V_{k|k} + G_{k|K}(V_{k+1|K} - V_{k+1|k}) G_{k|K}^\top.
\end{aligned} \tag{44}$$

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

The smoother also yields the one-step cross-covariance,

$$V_{k,k+1|K} = G_{k|K} V_{k+1|K}, \tag{45}$$

which is needed for interpolation (Section 5.2). For drift, $G_{k|K} > 0$ at all steps, so information propagates across recalibration events. For fouling, $G_{k|K} = 0$ at all steps, because each check causes a hard reset, so no information propagates across checks.

5.1 Parameter Estimation for Noisy Checks

If the initial error is unknown, then drift and diffusion can be estimated with an exact initial-state Kalman filter (Koopman, 1997). However, by taking a check measurement at the start of each deployment, we can use the simpler prediction-error decomposition (Harvey, 1989).

The unknown parameters are the drift rate μ and the diffusion coefficient σ^2. A natural approach is to estimate both by jointly maximizing the likelihood; however, the maximum-likelihood estimator of σ^2 is biased downward because it does not account for the degree of freedom consumed by estimating μ. Restricted maximum likelihood (REML) corrects this by marginalizing out μ analytically before estimating σ^2.

Using the prediction-error decomposition, the log-likelihood is

$$\log \mathcal{L}(\mu, \sigma^2) = -\frac{1}{2} \sum_{k=1}^{K} \left[\log(2\pi) + \log|S_k| + \nu_k^\top S_k^{-1} \nu_k \right], \tag{46}$$

where ν_k and S_k are the innovation and innovation covariance from the Kalman filter. Because the predicted mean depends linearly on μ, we can write $m_{k|k-1}(\mu) = m^{(0)}_{k|k-1} + \mu\, c_k$ and $\nu_k(\mu) = \nu_k^{(0)} - \mu\, c_k$, where the superscript (0) denotes evaluation at $\mu = 0$ and $c_k := \partial m_{k|k-1}/\partial \mu$ is the sensitivity coefficient at step k. Differentiating the filter prediction step with respect to μ gives the recursion

$$c_k = F_k \left(I - G_{k-1|k-2} \right) c_{k-1} + (t_k - t_{k-1}), \tag{47}$$

with base case $c_0 = 0$ because $\tilde{e}_0$ is observed and μ-independent, so that $c_1 = t_1 - t_0$. Here $G_{k-1|k-2}$ is the gain from the previous update step (Appendix E; Francke et al., 2010). For fouling ($F_k = 0$), the recursion collapses to $c_k = t_k - t_{k-1}$: the hard reset eliminates any dependence on the previous posterior. For drift ($F_k = I$), c_k accumulates across intervals, including across recalibrations.

Maximizing the log-likelihood over μ yields the generalized least-squares drift estimate

$$\hat{\mu} = \frac{\sum_{k=1}^{K} c_k^\top S_k^{-1} \nu_k^{(0)}}{\sum_{k=1}^{K} c_k^\top S_k^{-1} c_k}, \tag{48}$$

which is the precision-weighted estimate of the drift rate. Substituting $\hat{\mu}$ back gives profiled innovations $\hat{\nu}_k = \nu_k^{(0)} - \hat{\mu}\, c_k$, and marginalizing μ under a flat prior (Appendix E; Francke et al., 2010) yields the REML log-likelihood, up to a constant

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

independent of σ^2,

$$\log \mathcal{L}_R(\sigma^2) = \underbrace{-\frac{1}{2}\sum_{k=1}^{K}\left[\log(2\pi) + \log|S_k| + \hat{\nu}_k^\top S_k^{-1}\,\hat{\nu}_k\right]}_{\text{profiled MLE}} \underbrace{-\frac{1}{2}\log\left(\sum_{k=1}^{K} c_k^\top S_k^{-1}\, c_k\right)}_{\text{REML correction}}. \tag{49}$$

The term inside the log, $\sum c_k^\top S_k^{-1} c_k$, is the Fisher information for μ and measures how precisely the data constrain the drift rate. The REML correction subtracts $\frac{1}{2}\log$ of this term, inflating $\hat{\sigma}^2$ to account for the degree of freedom consumed by estimating $\hat{\mu}$. The estimate $\hat{\sigma}^2$ is found by maximizing $\log \mathcal{L}_R$ subject to $\sigma^2 \geq 0$. Because both S_k and c_k depend on σ^2 through the Kalman recursion, $\hat{\mu}$ is then recovered by evaluating Equation 48 at $\hat{\sigma}^2$. Kalman smoothing then yields the posterior states and covariances at each check.

In practice, the process noise can be heavy-tailed: fouling can occur episodically, for instance during storms, producing large increments that dominate $\hat{\sigma}^2$ under a Gaussian fit. Modeling the process noise with a heavier-tailed Student-t distribution bounds their influence (Appendix B).

5.2 Kalman Interpolation

With exact check measurements, bias corrections and prediction intervals are interpolated between checks using a Brownian bridge (Equations 29 and 30). We can adapt the same approach to noisy check measurements by propagating check uncertainty through these equations. The interpolation is exact: the smoothed checks are already conditioned on all available information, and given its bracketing check states, the error within each interval is conditionally independent (Markovian) of the data outside. Each interval can therefore be interpolated from its bracketing checks, just as in the exact-check case.

Let the expectations, variances, and covariance of the smoothed checks be

$$m_{k-1|K}, \qquad m_{k|K}, \qquad V_{k-1|K}, \qquad V_{k|K}, \qquad V_{k-1,k|K}. \tag{50}$$

Applying linearity of the expectation to the bridge gives the interpolated mean,

$$m_{t|K} = \underbrace{(1-\lambda_k^+(t))\, m_{k-1|K}^+ + \lambda_k^+(t) m_{k|K}}_{\text{mean interpolation}} + \underbrace{\mu\zeta_k^+(t)}_{\text{open-interval drift}}, \tag{51}$$

where $m_{k-1|K}^+ = F_k\,(m_{k-1|K} - (1-\delta_{k-1})\,\tilde{e}_{k-1})$ is the smoothed post-reset mean at t_{k-1}, which is equal to $m_{k-1|K}$ on non-reset drift steps, $m_{k-1|K} - \tilde{e}_{k-1}$ at drift recalibrations, and 0 at fouling checks. The variance of a linear combination gives the interpolated variance,

$$V_{t|K} = \underbrace{\sigma^2\phi_k^+(t)}_{\text{bridge variance}} + \underbrace{(1-\lambda_k^+(t))^2 V_{k-1|K}^+ + (\lambda_k^+(t))^2 V_{k|K} + 2\lambda_k^+(t)(1-\lambda_k^+(t))V_{k-1,k|K}^+}_{\text{measurement variance}}, \tag{52}$$

with post-reset smoothed covariances $V_{k-1|K}^+ = F_k\, V_{k-1|K}\, F_k^\top$ and $V_{k-1,k|K}^+ = F_k\, V_{k-1,k|K}$. Both reduce to $V_{k-1|K}$ and $V_{k-1,k|K}$ for drift, and vanish for fouling. Figure 1 shows the interpolated variance between two uncertain check measurements.

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

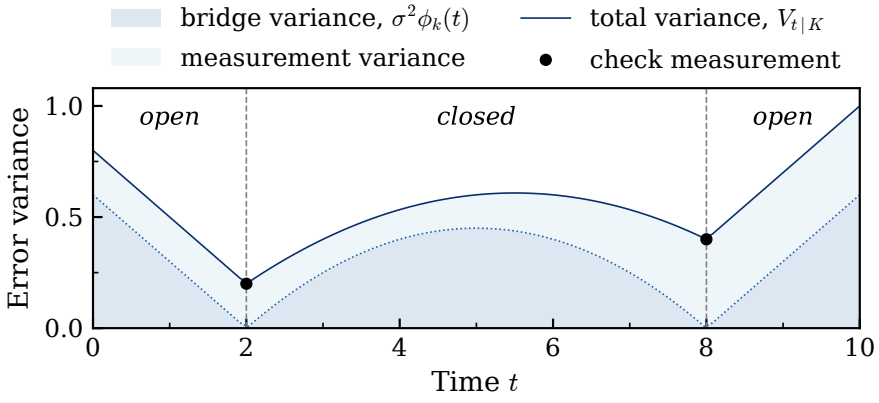

Figure 1. Error variance during open and closed intervals with diffusion coefficient $\sigma = \sqrt{0.3}$ and check variances $V_{t=2|K} = 0.2$ and $V_{t=8|K} = 0.4$. The two variance components represent uncertainty from the check measurements and from evolution of the generalized Wiener-process error.

6 Total Error

The mean and variance of the total error are sums of contributions from the independent components—fouling, drift, and noise—given by

$$m_{t|F,D} = m_{t|F} + m_{t|D} \quad \text{and} \tag{53}$$

$$V_{t|F,D} = V_{t|F} + V_{t|D} + v_n(\hat{x}_t), \tag{54}$$

where F and D denote the fouling and drift checks, respectively, $\hat{x}_t$ is the bias-corrected estimate of the measurand (Section 7), and the zero-mean noise term (Equation 8) contributes no bias to the mean but enters the variance through $v_n(\hat{x}_t)$.

7 Bias Corrections and Prediction Intervals

Bias corrections and prediction intervals are computed directly from the mean and variance. Given the measurement model $Y = X + E$, the bias-corrected estimate of the measurand X is given by

$$\hat{x}_t = y_t - m_{t|K}, \tag{55}$$

where $m_{t|K}$ is the expected error from fouling, drift, or both. Because the errors are conditionally Gaussian, the approximate prediction interval is

$$\text{PI} = \hat{x}_t \pm q_{1-\alpha/2}\sqrt{V_{t|F,D}}, \tag{56}$$

where $q_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution, and $V_{t|F,D}$ is the total error variance (Equation 54), which includes the measurement-noise variance $v_n(\hat{x}_t)$ in addition to the fouling and drift components.

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

8 Monitoring Networks

A monitoring network is a group of similar sensors, where checks from one sensor help estimate the parameters of the others. A newly deployed sensor has too few checks to estimate its parameters reliably. Pooling information across the network can give reasonable estimates during this initial period. For simplicity, we use empirical Bayes (Type-II maximum likelihood; Morris, 1983) with conjugate priors to shrink sensor parameters toward the network mean, an effect that diminishes as each sensor accumulates more checks. We keep the parameter index $k \in \{f, d\}$ implicit throughout this section.

Let $j \in \{1, \ldots, J\}$ index the sensors in the network. The parameters (μ_j, σ_j^2) are treated as random effects sampled from a conjugate Normal-Inverse-χ^2 distribution with hyperparameters $\theta_0 = (\mu_0, \kappa_0, \gamma_0, \sigma_0^2)$:

$$\sigma_j^2 \sim \text{Inv-}\chi^2(\gamma_0, \sigma_0^2), \qquad \mu_j \mid \sigma_j^2 \sim \mathcal{N}\left(\mu_0, \frac{\sigma_j^2}{\kappa_0}\right), \tag{57}$$

where μ_0 is the prior mean drift rate, and σ_0^2 is the prior scale for the diffusion σ_j^2.

The posterior distribution for sensor j is also Normal-Inverse-χ^2 with hyperparameters $\theta_j^* = (\mu_j^*, \kappa_j^*, \gamma_j^*, \sigma_j^{2*})$ given by

$$\kappa_j^* = \kappa_0 + T_j, \tag{58}$$

$$\gamma_j^* = \gamma_0 + K_j, \tag{59}$$

$$\mu_j^* = \frac{\kappa_0 \mu_0 + T_j \hat{\mu}_j}{\kappa_0 + T_j}, \tag{60}$$

$$\gamma_j^* \sigma_j^{2*} = \gamma_0 \sigma_0^2 + (K_j - 1)\hat{\sigma}_j^2 + \frac{\kappa_0 T_j}{\kappa_0 + T_j}(\hat{\mu}_j - \mu_0)^2, \tag{61}$$

where K_j is the number of check intervals for sensor j (consistent with Section 4, so sensor j has $K_j + 1$ check times $t_0, \ldots, t_{K_j}$), $T_j = \sum_{k=1}^{K_j}(t_k - t_{k-1})$ is the total elapsed check time, and $\hat{\sigma}_j^2$ uses a $K_j - 1$ denominator to account for a degree of freedom consumed by estimating $\hat{\mu}_j$ (Appendix C). Because T_j enters the update for κ_j^*, κ_0 has units of time and is interpreted as a prior elapsed time. If a sensor's error lacks a systematic trend, the drift rate μ is unnecessary and can be omitted, in which case the Normal-Inverse-χ^2 prior reduces to a Scaled Inverse-χ^2 prior on the diffusion σ_j^2 (Appendix D).

In empirical Bayes, the prior hyperparameters θ_0 are estimated by maximizing the marginal likelihood of the check data. Let E_j denote the check errors for sensor j. Then the marginal likelihood under Normal-Inverse-χ^2 conjugacy is

$$p(E_j \mid \theta_0) = \frac{1}{\pi^{K_j/2}} \prod_{k=1}^{K_j} (t_k - t_{k-1})^{-1/2} \sqrt{\frac{\kappa_0}{\kappa_j^*}} \frac{\Gamma(\gamma_j^*/2)}{\Gamma(\gamma_0/2)} \frac{(\gamma_0 \sigma_0^2)^{\gamma_0/2}}{(\gamma_j^* \sigma_j^{2*})^{\gamma_j^*/2}}. \tag{62}$$

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

Taking the logarithm and summing over all J sensors gives

$$\begin{aligned}\log \mathcal{L}(\theta_0) &= \sum_{j=1}^{J} \log p(E_j \mid \theta_0) \\ &= \sum_{j=1}^{J} \Bigg[-\frac{K_j}{2}\log(\pi) - \frac{1}{2}\sum_{k=1}^{K_j}\log(t_k - t_{k-1}) + \frac{1}{2}\log\frac{\kappa_0}{\kappa_j^*} \\ &\quad + \log\Gamma\left(\frac{\gamma_j^*}{2}\right) - \log\Gamma\left(\frac{\gamma_0}{2}\right) + \frac{\gamma_0}{2}\log\left(\gamma_0\sigma_0^2\right) - \frac{\gamma_j^*}{2}\log\left(\gamma_j^*\sigma_j^{2*}\right) \Bigg],\end{aligned} \tag{63}$$

which is maximized to estimate $\hat{\theta}_0 = (\hat{\mu}_0, \hat{\kappa}_0, \hat{\gamma}_0, \hat{\sigma}_0^2)$.

As an approximation, the sensor- and network-level parameters need not be updated at the same frequency. The posterior sensor parameters (μ_j^*, σ_j^{2*}) are recomputed after each check, whereas the population hyperparameters $\hat{\theta}_0$ are updated less frequently, because they change slowly relative to the sensor parameters.

The conjugate update is exact when check measurements are exact: each check interval then supplies one degree of freedom for σ_j^2, which the update counts through $\gamma_j^* = \gamma_0 + K_j$. When checks are noisy, each innovation is part process increment and part check noise, so it informs σ_j^2 less than a noise-free interval would, and the effective degrees of freedom fall below K_j. Because the update still counts the full K_j, it slightly overweights the sensor-level estimate relative to the network prior. The approximation is tightest when the process variance $\sigma^2(t_k - t_{k-1})$ dominates the check variance R_k, so that each innovation is mostly signal.

The update likewise assumes $\hat{\sigma}_j^2$ is the Gaussian estimate carrying $K_j - 1$ degrees of freedom. When the diffusion is fit with the Student-t model (Appendix B), there are fewer than $K_j - 1$ effective degrees of freedom, so the update slightly overweights that sensor relative to the network prior.

9 Data Requirements

The method requires the following information, which should be available for scientific-grade equipment.

1. A unique identifier (serial number) for each sensor.
2. The sensor type (make and model) for grouping similar sensors in a monitoring network.
3. The time of each cleaning and recalibration (when the error process resets).
4. The time and value of each check measurement, including its associated references and standards.
5. The precision (noise) of each sensor and standard.
6. A unique identifier for each batch of standards to account for cross-covariance between checks against the same batch.

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

Optionally, if the precision of the standard is negligible relative to that of the sensor, the standard precision and batch information may be omitted. If the precision of the sensor is also negligible relative to the fouling and calibration drift, then the sensor precision and serial number may also be omitted, in which case fouling and drift are modeled as Brownian bridges.

10 Computational Requirements

Storing the raw measurements, check information, parameters, and states makes the bias corrections and prediction intervals computable in $\mathcal{O}(1)$ time. Therefore, corrections and intervals can be computed on demand rather than precomputed and stored for the entire time series. Processing proceeds in three main stages. With a slight overloading of notation, let K_j denote the combined number of drift and fouling checks for sensor j.

Occasionally, all sensor and network parameters are re-estimated: first, by estimating the sensor parameters, which takes $\mathcal{O}(\sum_{j=1}^{J} K_j)$ time (Section 5.1), and storing the $4J$ drift and fouling parameters; next, by estimating the network hyperparameters, which takes $\mathcal{O}(J)$ time (Section 8), and storing θ_0.

After each check, the latent error states are updated in two steps: the posterior sensor parameters are recomputed using the network hyperparameters θ_0, and the states are then obtained by Kalman filtering and smoothing. This takes $\mathcal{O}(K_j)$ time per check and stores the $2(K_j + 1)$ state moments (means and variances) and the optional drift-check cross-covariances. As an approximation, we do not re-estimate the network hyperparameters after each check.

11 Demonstration

As a proof of concept, we apply the method to the U.S. Geological Survey's network of continuous water-temperature sensors, excluding sensors with fewer than three check measurements. The records do not identify each sensor's make and model, so we assume temperature is measured with a single type of thermistor of accuracy $\pm 0.01\,^{\circ}\mathrm{C}$, which fixes the measurement-noise variance v_n (Section 3; Appendix A). Each drift check then differences two thermistors, giving $R_d = 2v_n$, and each fouling check differences the *in situ* and reference thermistors before and after cleaning, giving $R_f = 4v_n$ (Section 3).

Table 1 reports the number of sensors fit, the mean number of checks per sensor, the per-sensor wall time, and the median across sensors of the fitted drift magnitude $|\hat{\mu}_j|$ and diffusion $\hat{\sigma}_j^2$. The last three columns are diagnostics on the standardized one-step prediction errors, ν_k divided by $\sqrt{S_k}$, pooled over all checks and sensors (Harvey, 1989; Durbin and Koopman, 2012): their standard deviation, the coverage of the nominal $95\,\%$ prediction interval, and their excess kurtosis (kurtosis minus 3). A Jarque–Bera test rejects normality for both processes, and the large positive excess kurtosis identifies the departure as heavy tails (Appendix B). This is unsurprising as fouling and drift often accrue in episodic bursts, in response to environmental conditions, rather than through continuous diffusion.

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

Table 1. Fouling and drift results for U.S. Geological Survey water-temperature sensors. The timing is wall-clock and excludes the snapshot read and library imports. The drift rate $|\hat{\mu}|$ and diffusion $\hat{\sigma}^2$ columns are medians across sensors, in °C day^{-1} and °C^2 day^{-1}. The innovation standard deviation, coverage, and excess kurtosis equal 1, 0.95, and 0 under a correctly specified fit; coverage is the fraction of standardized errors within ± 1.96.

Process	Sensors J	Avg. checks $\bar{K}$	Milliseconds per sensor	Drift rate median $\lvert\hat{\mu}\rvert$	Diffusion median $\hat{\sigma}^2$	Innovation SD	95 % PI coverage	Excess kurtosis
Drift	1357	23	1.01	4×10^{-5}	1×10^{-4}	0.98	0.94	8.1
Fouling	2632	58	2.28	6×10^{-4}	8×10^{-4}	1.00	0.96	53.4

12 Conclusions

Environmental sensors are periodically checked for fouling and calibration drift, cleaned, and recalibrated. Information from these checks can then be used to estimate bias corrections and prediction intervals for the time series. For exact checks and generalized Wiener-process error, this can be done using direct interpolation between checks. For noisy checks, the error state is updated using a Kalman smoother prior to interpolation. Check information can also be pooled across multiple sensors in a monitoring network to improve the accuracy of newly deployed sensors. All necessary information is routinely available, and the computation and storage requirements are modest, so the method can be integrated into existing monitoring programs. Although the current approach is restricted to single-point checks, it provides a foundation for developing a multipoint method.

Appendix A: Deriving Noise from an Accuracy Specification

Ideally, the variance of the measurement noise $v_n(x)$ is determined under controlled conditions. However, operators may not be able to perform such tests, so they must rely on the "accuracy" reported in sensor specifications. This appendix describes how to estimate measurement noise from an accuracy specification, assuming the specification was determined from a standard calibration check. In practice, accuracy may be defined differently, so it is best to consult the manufacturer.

Consider a sensor with a reported accuracy of ± 2 units, so $a = 2$, which represents a 2σ (approximately 95%) prediction interval. If the manufacturer determined this specification by testing a clean, calibrated sensor against known standards, then the specification reflects uncertainty from both the calibration check and noise in the subsequent measurements.

From Equation 17, the calibration-check noise variance is $R_d = v_n(x) + \sigma_s^2$. A subsequent measurement adds another $v_n(x)$, so the accuracy specification a is

$$a = 2\sqrt{(v_n(x) + \sigma_s^2) + v_n(x)}. \tag{A1}$$

Solving for the measurement-noise variance $v_n(x)$ gives

$$v_n(x) = \frac{a^2}{8} - \frac{\sigma_s^2}{2}, \tag{A2}$$

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

where a is the 2σ accuracy specification, and σ_s is the standard deviation of the calibration standard, which may also be reported as $2\sigma_s$. Because variance must be nonnegative, this implies $\sigma_s^2 \le a^2/4$.

As a second example, consider a sensor with reported accuracy of ± 0.5 units or 5% of reading, whichever is greater, interpreted as a 2σ bound. Let $a(x)$ denote the half-width of that bound,

$$a(x) = \max\left(0.5,\, 0.05|x|\right) = \begin{cases} 0.5, & |x| \le 10 \\ 0.05|x|, & |x| > 10 \end{cases}. \tag{A3}$$

Substituting $a(x)$ into the same derivation gives

$$v_n(x) = \frac{a(x)^2}{8} - \frac{\sigma_s^2}{2} = \begin{cases} \frac{0.5^2}{8} - \frac{\sigma_s^2}{2}, & |x| \le 10 \\ \frac{(0.05|x|)^2}{8} - \frac{\sigma_s^2}{2}, & |x| > 10 \end{cases}. \tag{A4}$$

If standard uncertainty is negligible ($\sigma_s^2 \approx 0$), then $v_n(x) = \max\left(0.03125,\, 0.0003125\, x^2\right)$.

Appendix B: Robust Diffusion and the Student-*t* Smoother

The error model assumes Gaussian increments and is therefore sensitive to heavy-tailed outliers. When large, infrequent increments are expected, a Student-t distribution can be substituted for the Gaussian process noise. In the standard Gaussian scale-mixture representation (Lange et al., 1989), each increment k carries a latent weight $w_k \sim \mathrm{Gamma}(\psi/2, \psi/2)$, and conditional on w_k the increment $e_k - e_{k-1}^+$ is Gaussian with mean $\mu\,(t_k - t_{k-1})$ and variance $\sigma^2\,(t_k - t_{k-1})/w_k$, where $e_{k-1}^+ := F_k\,(e_{k-1} - (1-\delta_{k-1})\,\tilde{e}_{k-1})$ is the post-reset error at t_{k-1}. Marginalizing over w_k yields a Student-t increment with ψ degrees of freedom. A reasonable practice is to fix the degrees of freedom at $\psi = 4$, because ψ is weakly identified when there are few checks per sensor.

The parameters (μ, σ^2) are fit by expectation-maximization (Dempster et al., 1977; Lange et al., 1989). Because the checks are noisy, the latent increments are not observed, so the E-step replaces them with their smoothed posterior estimates $m_{k|K}$ from the Kalman smoother (Section 5). It sets each weight to its posterior mean (Lange et al., 1989, Eq. 3),

$$w_k = \frac{\psi + 1}{\psi + u_k^2}, \qquad u_k = \frac{m_{k|K} - m_{k-1|K}^+ - \hat{\mu}\,(t_k - t_{k-1})}{\sqrt{\sigma^2\,(t_k - t_{k-1})}}, \tag{B1}$$

where u_k is the smoothed increment standardized by the process standard deviation, and $m_{k-1|K}^+$ is the smoothed post-reset mean (Section 5.2). The M-step is the same maximum-likelihood (or REML) diffusion fit as in the Gaussian model, with the per-increment process variance $\sigma^2\,(t_k - t_{k-1})$ replaced by $\sigma^2\,(t_k - t_{k-1})/w_k$ to down-weight outlying increments; the drift term $\mu\,(t_k - t_{k-1})$ and the sensitivity recursion retain the physical gap, because the weight rescales variance, not time.

Carrying the converged weights into the smoother sets the per-increment process variance to $Q_k = \sigma^2\,(t_k - t_{k-1})/w_k$, which lets the smoother absorb outliers rather than propagate them across neighboring states (Meinhold and Singpurwalla, 1989; Roth et al., 2013). The same down-weighting keeps σ^2 calibrated to the bulk of the data, while allowing the posterior mean to follow

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

the check measurements, whose noise we still assume to be Gaussian. The heavy tail also propagates into the prediction interval, where the standard-normal quantile $q_{1-\alpha/2}$ of Equation 56 is replaced by the Student-t quantile $t_{\psi,1-\alpha/2}$, with $V_{t|F,D}$ taken as the squared scale. Because the check noise remains Gaussian and only the increments are heavy-tailed, this substitution is only approximate.

Appendix C: Hierarchical Posterior Derivation

This appendix derives the conjugate posterior of a generalized Wiener process with exact checks, which is used in Section 8 (Normal-Inverse-χ^2 updates follow Murphy, 2007); we drop the sensor index j for clarity. Under this model, the (reset-aware) increments are independent Gaussian,

$$e_k - \delta_{k-1} e_{k-1} \sim \mathcal{N}\big(\mu(t_k - t_{k-1}), \sigma^2(t_k - t_{k-1})\big)\,. \tag{C1}$$

Let $T = \sum_{k=1}^{K}(t_k - t_{k-1})$ be the total elapsed time. The likelihood is

$$p(e_{1:K} \mid \mu, \sigma^2) \propto (\sigma^2)^{-K/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{k=1}^{K}\frac{(e_k - \delta_{k-1}e_{k-1} - \mu(t_k - t_{k-1}))^2}{t_k - t_{k-1}}\right). \tag{C2}$$

Expanding the sum and completing the square in μ gives

$$\sum_{k=1}^{K}\frac{(e_k - \delta_{k-1}e_{k-1} - \mu(t_k - t_{k-1}))^2}{t_k - t_{k-1}} = T(\mu - \hat{\mu})^2 + (K-1)\hat{\sigma}^2, \tag{C3}$$

where $\hat{\mu}$ is the maximum-likelihood drift rate,

$$\hat{\mu} = \frac{1}{T}\sum_{k}(e_k - \delta_{k-1}e_{k-1}), \tag{C4}$$

and

$$(K-1)\hat{\sigma}^2 = \sum_{k}\frac{(e_k - \delta_{k-1}e_{k-1} - \hat{\mu}(t_k - t_{k-1}))^2}{t_k - t_{k-1}}, \tag{C5}$$

which is the sum of squared standardized residuals, with $K-1$ degrees of freedom because $\hat{\mu}$ is estimated from the same data.

Combining the likelihood with the Normal-Inverse-χ^2 prior,

$$p(\mu, \sigma^2 \mid \theta_0) \propto (\sigma^2)^{-(\gamma_0/2+3/2)} \exp\left(-\frac{\gamma_0\sigma_0^2 + \kappa_0(\mu - \mu_0)^2}{2\sigma^2}\right), \tag{C6}$$

the joint posterior is proportional to

$$(\sigma^2)^{-((\gamma_0+K)/2+3/2)} \exp\left(-\frac{\gamma_0\sigma_0^2 + (K-1)\hat{\sigma}^2 + \kappa_0(\mu - \mu_0)^2 + T(\mu - \hat{\mu})^2}{2\sigma^2}\right). \tag{C7}$$

The two quadratics in μ combine via the standard precision-weighted-mean identity,

$$\kappa_0(\mu - \mu_0)^2 + T(\mu - \hat{\mu})^2 = (\kappa_0 + T)(\mu - \mu^*)^2 + \frac{\kappa_0 T}{\kappa_0 + T}(\hat{\mu} - \mu_0)^2, \tag{C8}$$

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

with $\mu^* = (\kappa_0\mu_0 + T\hat{\mu})/(\kappa_0 + T)$. Substituting back, the posterior is again Normal-Inverse-χ^2 with hyperparameters

$$\kappa^* = \kappa_0 + T, \tag{C9}$$

$$\gamma^* = \gamma_0 + K, \tag{C10}$$

$$\mu^* = \frac{\kappa_0\mu_0 + T\hat{\mu}}{\kappa_0 + T}, \tag{C11}$$

$$\gamma^*\sigma^{2*} = \gamma_0\sigma_0^2 + (K-1)\hat{\sigma}^2 + \frac{\kappa_0 T}{\kappa_0 + T}(\hat{\mu} - \mu_0)^2, \tag{C12}$$

which are the formulas used in Section 8; the traditional Normal-Inverse-χ^2 update (Gelman et al., 2013) uses K in place of T throughout. The two coincide when check intervals are uniform across sensors, because T_j is then proportional to K_j and the constant of proportionality is absorbed into κ_0.

Appendix D: Omitting the Drift Rate

Depending on the characteristics of a sensor, the drift-rate parameter can be omitted. In that case, the Normal-Inverse-χ^2 prior reduces to a Scaled Inverse-χ^2 prior on the diffusion parameter,

$$\sigma_j^2 \sim \text{Inv-}\chi^2(\gamma_0, \sigma_0^2). \tag{D1}$$

Omitting (μ, κ) from the Normal-Inverse-χ^2 gives a Scaled Inverse-χ^2 posterior for sensor j with parameters

$$\gamma_j^* = \gamma_0 + K_j, \tag{D2}$$

$$\sigma_j^{2*} = \frac{\gamma_0\sigma_0^2 + K_j\hat{\sigma}_j^2}{\gamma_0 + K_j}, \tag{D3}$$

where $\hat{\sigma}_j^2$ uses a K_j denominator rather than $K_j - 1$ because no degree of freedom is consumed by estimating μ. Similarly, the log marginal likelihood for (γ_0, σ_0^2) is

$$\log\mathcal{L}(\gamma_0, \sigma_0^2) = \sum_{j=1}^{J}\Bigg[-\frac{K_j}{2}\log(\pi) - \frac{1}{2}\sum_{k=1}^{K_j}\log(t_k - t_{k-1}) + \log\Gamma\left(\frac{\gamma_j^*}{2}\right) - \log\Gamma\left(\frac{\gamma_0}{2}\right) + \frac{\gamma_0}{2}\log\left(\gamma_0\sigma_0^2\right) - \frac{\gamma_j^*}{2}\log\left(\gamma_j^*\sigma_j^{2*}\right)\Bigg]. \tag{D4}$$

To test for the presence of drift, fit models with and without drift parameters, then compare their maximized log-likelihoods with a penalized criterion such as the Akaike information criterion (AIC; Akaike, 1974),

$$\text{AIC} = 2p - 2\log\hat{\mathcal{L}}, \tag{D5}$$

where p counts the parameters maximized in $\hat{\mathcal{L}}$. For a single sensor, $\hat{\mathcal{L}}$ is the maximized prediction-error likelihood (Section 5.1), so $p = 2$ with drift (μ and σ^2) and $p = 1$ without (σ^2 only); use the maximum likelihood rather than REML, because restricted likelihoods are not comparable between models with different mean structures. For a network, $\hat{\mathcal{L}}$ is the maximized marginal likelihood, so $p = 4$ with drift (μ_0, κ_0, γ_0, σ_0^2) and $p = 2$ without (γ_0 and σ_0^2).

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

Appendix E: REML Derivation

This appendix derives the sensitivity recursion (Equation 47), the generalized least-squares drift estimate (Equation 48), and the REML correction term of Section 5.1. It specializes the marginal-likelihood treatment of Francke et al. (2010) to a single scalar drift parameter, reducing their matrix recursion to a scalar recursion for c_k.

E1 Sensitivity Recursion

The prediction step of the filter is

$$m_{k|k-1} = F_k\, m_{k-1|k-1} + \mu\,(t_k - t_{k-1}) - F_k\,(1-\delta_{k-1})\,\tilde{e}_{k-1}, \tag{E1}$$

where $B_k u_k$ has been expanded as $\mu\,(t_k - t_{k-1}) - F_k\,(1-\delta_{k-1})\,\tilde{e}_{k-1}$. Differentiating both sides with respect to μ, and noting that $\tilde{e}_{k-1}$ is observed and so μ-independent (the operator-correction term contributes zero to the derivative), yields

$$c_k = F_k\,\frac{\partial m_{k-1|k-1}}{\partial \mu} + (t_k - t_{k-1}). \tag{E2}$$

To evaluate $\partial m_{k-1|k-1}/\partial\mu$, rewrite the filter update (with $H_{k-1} = I$) as

$$\begin{aligned} m_{k-1|k-1} &= m_{k-1|k-2} + G_{k-1|k-2}\,(\tilde{e}_{k-1} - m_{k-1|k-2}) \\ &= (I - G_{k-1|k-2})\,m_{k-1|k-2} + G_{k-1|k-2}\,\tilde{e}_{k-1}. \end{aligned} \tag{E3}$$

Differentiating, and again using that $\tilde{e}_{k-1}$ is μ-independent,

$$\frac{\partial m_{k-1|k-1}}{\partial \mu} = (I - G_{k-1|k-2})\,c_{k-1}. \tag{E4}$$

Substituting into Equation E2 gives the recursion

$$c_k = F_k\,(I - G_{k-1|k-2})\,c_{k-1} + (t_k - t_{k-1}). \tag{E5}$$

At $k = 1$, the filter is initialized at $m_{0|0} = \tilde{e}_0$, so $\partial m_{0|0}/\partial\mu = 0$ and Equation E2 gives $c_1 = t_1 - t_0$ directly. The recursion then runs for $k = 2, \ldots, K$.

E2 Generalized Least-Squares Drift Estimate

Substituting the innovation $\nu_k(\mu) = \nu_k^{(0)} - \mu\, c_k$, where the superscript (0) denotes evaluation at $\mu = 0$ (Section 5.1), into the log-likelihood gives

$$-2\log\mathcal{L}(\mu,\sigma^2) \mathrel{\hat{=}} \sum_{k=1}^{K}\left[\log|S_k| + (\nu_k^{(0)} - \mu\, c_k)^\top S_k^{-1}\,(\nu_k^{(0)} - \mu\, c_k)\right], \tag{E6}$$

where $\mathrel{\hat{=}}$ denotes equality up to a constant. Expanding the quadratic,

$$(\nu_k^{(0)} - \mu\, c_k)^\top S_k^{-1}\,(\nu_k^{(0)} - \mu\, c_k) = \nu_k^{(0)\top} S_k^{-1}\nu_k^{(0)} - 2\mu\, c_k^\top S_k^{-1}\nu_k^{(0)} + \mu^2\, c_k^\top S_k^{-1} c_k. \tag{E7}$$

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

Differentiating the sum with respect to μ and setting the derivative to zero,

$$-2\sum_{k=1}^{K} c_k^\top S_k^{-1}\nu_k^{(0)} + 2\mu\sum_{k=1}^{K} c_k^\top S_k^{-1}c_k = 0. \tag{E8}$$

Moving the μ-independent sum to the right side and dividing by 2 gives

$$\mu\sum_{k=1}^{K} c_k^\top S_k^{-1}c_k = \sum_{k=1}^{K} c_k^\top S_k^{-1}\nu_k^{(0)}. \tag{E9}$$

Solving for μ gives the generalized least-squares drift estimate $\hat{\mu}$ (Equation 48).

E3 REML Correction

Let

$$A := \sum_{k=1}^{K} c_k^\top S_k^{-1}c_k, \qquad B := \sum_{k=1}^{K} c_k^\top S_k^{-1}\nu_k^{(0)}, \tag{E10}$$

so that $\hat{\mu} = B/A$. Completing the square in μ,

$$\sum_{k=1}^{K}(\nu_k^{(0)} - \mu\, c_k)^\top S_k^{-1}(\nu_k^{(0)} - \mu\, c_k) = A\,(\mu-\hat{\mu})^2 + \sum_{k=1}^{K}\hat{\nu}_k^\top S_k^{-1}\hat{\nu}_k, \tag{E11}$$

where $\hat{\nu}_k = \nu_k^{(0)} - \hat{\mu}\, c_k$. Exponentiating to obtain the likelihood,

$$\mathcal{L}(\mu,\sigma^2) \propto \left(\prod_{k=1}^{K}|S_k|^{-1/2}\right)\exp\left(-\tfrac{1}{2}\,A\,(\mu-\hat{\mu})^2 - \tfrac{1}{2}\sum_{k=1}^{K}\hat{\nu}_k^\top S_k^{-1}\hat{\nu}_k\right). \tag{E12}$$

Marginalizing μ under a flat prior uses the Gaussian integral $\int \exp(-\frac{1}{2}A(\mu-\hat{\mu})^2)\,d\mu = \sqrt{2\pi/A}$, which contributes $-\frac{1}{2}\log A$ to the log marginal likelihood. Dropping constants,

$$\log\mathcal{L}_R(\sigma^2) \mathrel{\hat{=}} -\tfrac{1}{2}\sum_{k=1}^{K}\left[\log|S_k| + \hat{\nu}_k^\top S_k^{-1}\hat{\nu}_k\right] - \tfrac{1}{2}\log A, \tag{E13}$$

which is the REML log-likelihood of Section 5.1.

Appendix F: Correlated Check Noise via State Augmentation

When the same batch of standards is used for multiple drift checks, the check noise terms share a common error component E_s (Section 3). The scalar-R_k filter in Section 5 treats each check as independent, ignoring this cross-covariance. State augmentation, the standard treatment for correlated measurement noise (Bryson and Henrikson, 1968), avoids this approximation by including E_s in the state vector.

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

Let b index the active batch. Augment the state to $\mathbf{x}_k = (e_k, E_{s,b})^\top$, with process model

$$\mathbf{x}_k = \begin{pmatrix} F_k & 0 \\ 0 & 1 \end{pmatrix} \mathbf{x}_{k-1} + \begin{pmatrix} \mu\,(t_k - t_{k-1}) - F_k\,(1-\delta_{k-1})\,\tilde{e}_{k-1} \\ 0 \end{pmatrix} + \mathbf{q}_k,$$
$$\mathbf{Q}_k = \begin{pmatrix} \sigma^2(t_k - t_{k-1}) & 0 \\ 0 & 0 \end{pmatrix}, \tag{F1}$$

and observation model

$$\tilde{e}_k = \begin{pmatrix} 1 & -1 \end{pmatrix} \mathbf{x}_k + E_{n,k}, \qquad E_{n,k} \sim \mathcal{N}(0, v_n). \tag{F2}$$

The filter is initialized by conditioning on the deployment check $\tilde{e}_0 = e_0 - E_{s,0} + E_{n,0}$ under a flat prior on e_0 and an $\mathcal{N}(0, \sigma_s^2)$ prior on $E_{s,0}$, giving

$$\mathbf{m}_{0|0} = \begin{pmatrix} \tilde{e}_0 \\ 0 \end{pmatrix}, \qquad \mathbf{V}_{0|0} = \begin{pmatrix} v_n + \sigma_s^2 & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 \end{pmatrix}, \tag{F3}$$

whose upper-left entry recovers the scalar filter's $V_{0|0} = R_0$ from Section 5. The off-diagonal σ_s^2 is the cross-covariance between e_0 and $E_{s,0}$ induced by their joint observation in $\tilde{e}_0$. The batch state $E_{s,b}$ has zero process noise, so it persists unchanged until a new batch is introduced, at which point $E_{s,b}$ is removed from the state and a new $E_{s,b'}$ is appended with prior $\mathcal{N}(0, \sigma_s^2)$. The observation noise reduces from $v_n + \sigma_s^2$ to v_n alone, because the error of the standard is captured by the batch state. This adds one state dimension per active batch and is unnecessary when σ_s^2 is small relative to v_n.

Appendix G: Implementation Notes for Reproducibility

As a reproducibility check, we provided this manuscript as a prompt to Claude (an AI assistant), asked it to implement the model and compare its implementation against ours, and then updated the manuscript based on the discrepancies. Both implementations were validated by four diagnostic tests that map to specific claims in the manuscript. The descriptions below are intended as prompts for reimplementation.

Test 1 (Kalman diagnostic). Simulate a single sensor under both drift and fouling modes with noisy checks, fit parameters by REML, and assert that the standardized filter innovations $\nu_k/\sqrt{S_k}$ are approximately $\mathcal{N}(0,1)$ and that the smoother's 95% prediction interval covers the true latent error at approximately the nominal rate, both at check times and dense interpolation times.

Test 2 (REML bias). Sweep K from small to large and run Monte Carlo replicates fitting both the joint MLE and REML. Assert that at small K the MLE of σ^2 is biased downward, that REML has smaller absolute bias than MLE at every K, and that both biases decay as K grows.

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

Test 3 (Bridge vs. Kalman). For exact checks, compute the analytical Brownian-bridge variance $\sigma^2\phi_k(t)$ at dense times, then add synthetic check noise at decreasing R and run the Kalman smoother with true parameters. Assert that the maximum absolute deviation between the smoothed interpolation variance and the bridge variance decreases monotonically and vanishes as $R \to 0$.

Test 4 (Augmented state). Simulate a sensor whose drift checks share a common error component $E_{s,b}$ within each batch (Appendix F). Sweep the batch variance σ_s^2 and compare the scalar filter (which treats each check as independent with $R = v_n + \sigma_s^2$) against the augmented 2-state model that includes $E_{s,b}$ in the state vector. Assert that the augmented estimator's RMSE in $\hat{\sigma}^2$ is at most that of the scalar estimator at every positive σ_s^2, and that the two agree at $\sigma_s^2 = 0$.

Code availability. A reference implementation of the methods described here is available in Python as a U.S. Geological Survey software release (Hodson, 2026). The package implements the parameter estimation, Kalman smoothing, and bias-correction routines together with the diagnostic tests reported in this paper.

Author contributions. Both authors contributed equally. G.E.S. developed the initial concept, but retired before completing the manuscript.

Competing interests. Authors T.O.H. and G.E.S. are employees of the U.S. Geological Survey.

Disclaimer. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Acknowledgements. This work was supported by the Groundwater and Streamflow Information Program of the U.S. Geological Survey. During the preparation of this manuscript, the authors used Claude (Anthropic) to assist with derivations and editing. All content was subsequently reviewed and revised by the authors.

https://doi.org/10.5194/egusphere-2026-4028
Preprint. Discussion started: 17 August 2026

References

Akaike, H.: A New Look at the Statistical Model Identification, IEEE Transactions on Automatic Control, 19, 716–723, https://doi.org/10.1109/TAC.1974.1100705, 1974.

Bryson, Jr, A. E. and Henrikson, L. J.: Estimation Using Sampled Data Containing Sequentially Correlated Noise, Journal of Spacecraft and Rockets, 5, 662–665, https://doi.org/10.2514/3.29327, 1968.

Dempster, A. P., Laird, N. M., and Rubin, D. B.: Maximum likelihood from incomplete data via the EM algorithm, Journal of the Royal Statistical Society: Series B, 39, 1–22, https://doi.org/10.1111/j.2517-6161.1977.tb01600.x, 1977.

Durbin, J. and Koopman, S. J.: Time Series Analysis by State Space Methods, Oxford Statistical Science Series, Oxford University Press, 2 edn., https://doi.org/10.1093/acprof:oso/9780199641178.001.0001, 2012.

Francke, M. K., Koopman, S. J., and De Vos, A. F.: Likelihood functions for state space models with diffuse initial conditions, Journal of Time Series Analysis, 31, 407–414, https://doi.org/10.1111/j.1467-9892.2010.00673.x, 2010.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B.: Bayesian Data Analysis, Chapman and Hall/CRC, 3 edn., https://doi.org/10.1201/b16018, 2013.

Harvey, A. C.: Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press, https://doi.org/10.1017/CBO9781107049994, 1989.

Hodson, T.: singlepoint-correction: A reference implementation for single-point fouling and calibration drift corrections, https://doi.org/10.5066/P1F4XS7V, 2026.

Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering, 82, 35–45, https://doi.org/10.1115/1.3662552, 1960.

Koopman, S. J.: Exact Initial Kalman Filtering and Smoothing for Nonstationary Time Series Models, Journal of the American Statistical Association, 92, 1630–1638, https://doi.org/10.1080/01621459.1997.10473685, 1997.

Lange, K. L., Little, R. J. A., and Taylor, J. M. G.: Robust statistical modeling using the t distribution, Journal of the American Statistical Association, 84, 881–896, https://doi.org/10.1080/01621459.1989.10478852, 1989.

Meinhold, R. J. and Singpurwalla, N. D.: Robustification of Kalman Filter Models, Journal of the American Statistical Association, 84, 479–486, https://doi.org/10.1080/01621459.1989.10478794, 1989.

Morris, C. N.: Parametric Empirical Bayes Inference: Theory and Applications, Journal of the American Statistical Association, 78, 47–55, https://doi.org/10.2307/2287098, 1983.

Murphy, K. P.: Conjugate Bayesian analysis of the Gaussian distribution, Tech. rep., https://www.cs.ubc.ca/~murphyk/Papers/bayesGauss.pdf, 2007.

Rauch, H. E., Tung, F., and Striebel, C. T.: Maximum likelihood estimates of linear dynamic systems, AIAA Journal, 3, 1445–1450, https://doi.org/10.2514/3.3166, 1965.

Roth, M., Özkan, E., and Gustafsson, F.: A Student's t filter for heavy tailed process and measurement noise, in: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5770–5774, https://doi.org/10.1109/ICASSP.2013.6638770, 2013.

Taylor, H. and Karlin, S.: An Introduction to Stochastic Modeling, Elsevier Science, 1998.

Wagner, R. J., Boulger Jr., R. W., Oblinger, C. J., and Smith, B. A.: Guidelines and Standard Procedures for Continuous Water-Quality Monitors: Station Operation, Record Computation, and Data Reporting, https://doi.org/10.3133/tm1D3, 2006.