Autocorrelation Function Kalman Filter (ACF-KF) Design for Adaptive Measurement Scheduling in Multi-Channel LoRaWAN Sensor Networks

Soheyl Faghir Hagh
Electrical Engineering Department, University of Vermont, Burlington, VT, USA
sfaghirh@uvm.edu

***Abstract*—Frequent acquisition and wireless transmission can dominate the energy budget of autonomous wireless environmental sensing systems (including water quality monitoring), particularly when measurements activate pumps and high-power optical hardware. This paper presents an autocorrelation-aware Kalman filter (ACF–KF) for multichannel prediction and adaptive measurement scheduling. The method is evaluated using 458 records acquired at 15-min intervals during a nominal five-day deployment. Seven channels are analyzed: temperature, battery voltage, turbidity, low- and high-sensitivity phycocyanin (PC), and low- and high-sensitivity chlorophyll-a (CHL). After excluding the initial transient from model identification, 96 samples form a 24-h learning set and 361 samples are held out for evaluation. Each channel is modeled with a local-linear KF. Lag-one autocorrelation of recent innovations is used to adapt the process-noise covariance when persistent correlation indicates under-modeled dynamics. Relative to a fixed-covariance KF, ACF–KF reduces held-out one-step RMSE by 1.9–6.7% and MAE by 2.9–8.5% across all channels, while reducing pooled median recovery time for learning-defined upper-quintile changes from 45 to 30 min. A scheduler using a temperature sentinel, predicted uncertainty, predicted state change, and a maximum-gap constraint reduces full sensing cycles by 52.6%, increasing the mean full-cycle interval from 15 to 31.7 min with 5.83% mean reconstruction NRMSE. Fixed-interval analysis from 15 to 240 min further quantifies the sampling–accuracy tradeoff.**

***Index Terms*—autocorrelation, adaptive sampling, environmental monitoring, Kalman filter, LoRaWAN, predictive sensing, stochastic processes, water quality.**

I. Introduction

Battery lifetime remains a principal constraint in remotely deployed Internet-of-Things (IoT) and low-power wide-area network (LPWAN) sensing systems. Communication, sensor warm-up, optical excitation, pumping, and repeated acquisition can collectively dominate node energy consumption, making fixed-rate operation inefficient when the monitored process evolves slowly [1], [2]. Predictive sensing therefore offers a complementary strategy to hardware-level low-power design: instead of measuring every variable at every nominal sampling instant, a state estimator can determine when a new physical measurement is necessary.

This article is a Stochastic Processes course project paper and has been archived. All simulations were carried out using MATLAB.

Soheyl Faghir Hagh is with the Electrical and Biomedical Engineering Department, University of Vermont, Burlington, VT, USA.

Kalman filtering (KF) is well suited to this problem because it explicitly separates process uncertainty from measurement uncertainty and recursively propagates both the state estimate and its covariance. KFs and related ensemble methods have been applied to water-quality state estimation [3], while broad reviews summarize their use under linear, nonlinear, and uncertain dynamics [4]. Energy-aware implementations have also been studied at the computing level [5]. In parallel, low-power environmental sensing platforms have used energy harvesting, duty cycling, efficient radios, and low-power sensor interfaces to extend field lifetime [6]–[9]. These approaches reduce the cost of an individual acquisition or transmission, but they do not by themselves determine whether the acquisition is needed.

The sensing platform underlying this study is the LoRaWAN fluorometer–nephelometer developed in [10]. This device measures temperature, turbidity, phycocyanin (PC) fluorescence, and chlorophyll-a (CHL) fluorescence and transmits environmental data through a long-range IoT link. The optical instrument uses amber, blue, and near-infrared excitation for PC, CHL, and turbidity sensing, respectively, and circulates water through the optical cell using a peristaltic pump. The prior system was operated on a 15-min acquisition schedule, and the pump can draw approximately 600 mA at 12 V during operation [10]. Consequently, suppressing redundant full sensing cycles can produce substantial energy savings when the water-quality variables are temporally correlated.

The central statistical issue is that a conventional fixed-covariance KF assumes a dynamic model and noise description that remain adequate over time. In field data, however, model mismatch is often visible as temporal correlation in the KF innovation. Adaptive KF methods address this problem through nonlinear filtering, recursive covariance estimation, Bayesian adaptation, innovation-based correction, or maximum-likelihood noise estimation [12]–[17]. A standard predict/update interpretation of state estimation is also summarized in [11]. The remaining engineering problem for the present sensor is to connect such temporal structure directly to measurement scheduling and to quantify the resulting accuracy–sampling trade-off on a complete deployment record.

The main contributions of this work are summarized as follows:

1) The autocorrelation structure of seven simultaneously

recorded sensor channels is quantified, demonstrating substantial short-lag redundancy in the 15-min measurements.

2) An ACF-aware local-linear Kalman filter is formulated, in which the lag-one autocorrelation of the innovation sequence adaptively modifies the process-noise covariance while maintaining positive semi-definiteness.
3) An adaptive sensing scheduler is evaluated on the recorded dataset by combining a low-power temperature sentinel with Kalman-filter uncertainty and predicted state variation to determine when full multichannel measurements are required.
4) A controlled measurement-frequency analysis is performed over sampling intervals from 15 to 240 min to quantify the tradeoff between measurement reduction and reconstruction accuracy.

All model-selection and parameter-tuning operations are restricted to the initial learning window, while the remaining deployment data are held out exclusively for evaluation.

The remainder of this paper is structured as follows. Section II details the sensor platform and deployment data; Section III details the stochastic characterization and estimation methods; Section IV proposes the adaptive measurement scheduling module; Section V evaluates the protocol; Section VI presents the results and discussion; and Section VII concludes the study.

II. Sensor Platform and Deployment Data

A. Water-Quality Sensing Platform

The underlying instrument is a portable LoRaWAN optical fluorometer–nephelometer developed for real-time aquatic monitoring [10]. The platform includes temperature sensing, an 870-nm near-infrared turbidity channel, a 590-nm phycocyanin excitation channel, and a 465-nm chlorophyll-a excitation channel. Silicon photodiodes detect optical responses, a peristaltic pump moves water through a polymethylmethacrylate optical cell, local storage retains measurements, and LoRaWAN provides wireless telemetry. The deployment record used here additionally contains low- and high-sensitivity PC and CHL outputs, providing seven time-varying fields when battery voltage is included (Fig. 1).

The proposed scheduling layer does not alter the optical calibration or sensor electronics. Instead, it determines when the high-power sensing sequence must be executed. Temperature is treated as a low-power sentinel that can be sampled at the nominal 15-min clock without activating the entire optical sequence. This architecture is consistent with the two-tier strategy in which a low-cost environmental observation can trigger a complete measurement when rapid dynamics are detected.

B. Deployment Record and Modeling Split

The collected dataset contains $N = 458$ records at a nominal sampling period of $\Delta t = 15$ min, spanning 6855 min (114.25 h). The recorded vector is

$$\boldsymbol{y}_k = [T_k, V_{b,k}, U_{\mathrm{Turb},k}, U_{\mathrm{PC,L},k}, U_{\mathrm{PC,H},k}, U_{\mathrm{CHL,L},k}, U_{\mathrm{CHL,H},k}]^T. \quad (1)$$

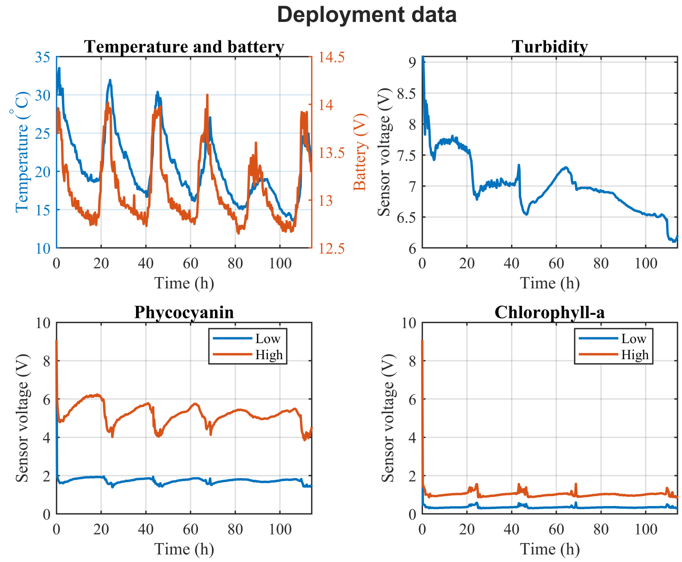

Fig. 1. Recorded field deployment data. The raw t=0 record is shown for completeness; the temperature and battery swings demonstrate diurnal patterns and battery charge/discharge via solar panels.

where T is temperature, V_b is battery voltage, and U denotes the recorded optical-channel output (i.e., turbidity, PC and CHL). Because the first acquisition contains a common transient high value across the optical channels, it is retained in the raw-data figure but excluded from model identification.

The next $N_L = 96$ samples constitute a 24-h learning set $\mathcal{L}$. All model scaling and controller parameters are determined using $\mathcal{L}$. The remaining $N_E = 361$ samples constitute the held-out evaluation set $\mathcal{E}$. For numerical conditioning, each channel is standardized using only the learning statistics,

$$\tilde{y}_{j,k} = \frac{y_{j,k} - \mu_{j,L}}{\sigma_{j,L}}, \quad (2)$$

where $y_{j,k}$ denotes the measured value of sensor channel j at time step k, and $\mu_{j,L}$ and $\sigma_{j,L}$ are the mean and standard deviation of channel j, respectively, computed exclusively from the learning set L.

III. Stochastic Characterization and Estimation Method

A. Autocorrelation Analysis

For a finite sequence $x_1, \ldots, x_N$, the sample autocorrelation at lag ℓ is

$$\hat{\rho}(\ell) = \frac{\sum_{k=1}^{N-\ell}(x_k - \bar{x})(x_{k+\ell} - \bar{x})}{\sum_{k=1}^{N}(x_k - \bar{x})^2}. \quad (3)$$

For the $N_L = 96$ learning points, an approximate 95% white-noise bound is $\pm 1.96/\sqrt{N_L} = \pm 0.200$. Table I shows strong lag-one correlation in every channel after the startup (transient) samples. The first lag at which the ACF enters the white-noise bound occurs between 2.25 and 4.0 h, indicating substantial short-term temporal redundancy at the 15-min sampling interval. See Fig. 2 for sample autocorrelation function plots.

TABLE I
AUTOCORRELATION OF THE 24-H LEARNING WINDOW

Channel	$\rho(1)$	$\rho(4)$	$\rho(8)$	First nonsig. lag	Time
Temperature	0.957	0.792	0.546	16	4.00 h
Battery	0.949	0.786	0.513	14	3.50 h
Turbidity	0.827	0.465	0.206	9	2.25 h
PC-Low	0.945	0.706	0.411	14	3.50 h
PC-High	0.953	0.749	0.438	13	3.25 h
CHL-Low	0.885	0.618	0.363	12	3.00 h
CHL-High	0.848	0.544	0.335	12	3.00 h

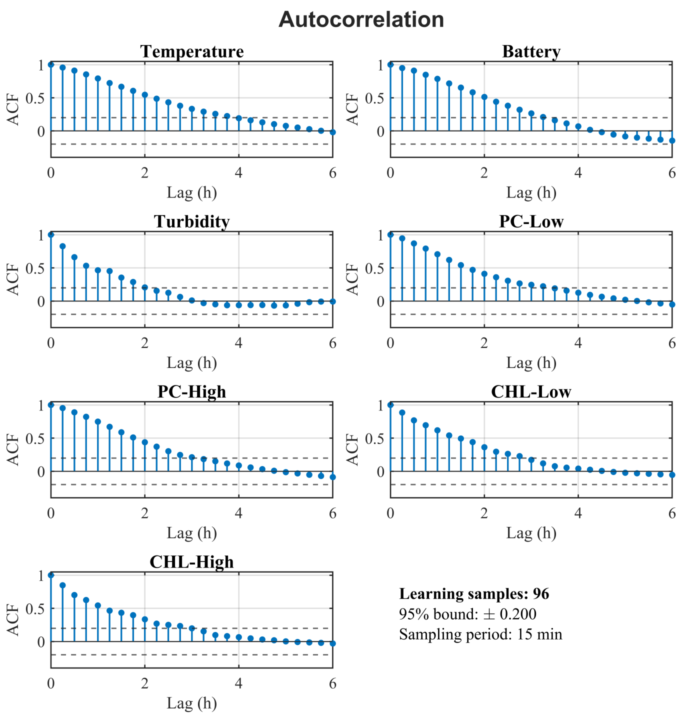

Fig. 2. Sample autocorrelation functions computed from the learning window. Dashed limits denote the approximate 95% white-noise bounds.

B. Local-Linear State Model

A separate two-state model is maintained for each standardized channel. The state contains the instantaneous level and its rate per sample,

$$\boldsymbol{x}_{j,k} = \begin{bmatrix} \tilde{y}_{j,k} & \dot{\tilde{y}}_{j,k} \end{bmatrix}^T. \tag{4}$$

Here, $\boldsymbol{x}_{j,k} \in \mathbb{R}^2$ is the state vector of sensor channel j at discrete time step k, $\tilde{y}_{j,k}$ is the standardized measured value of channel j at time step k, and $\dot{\tilde{y}}_{j,k}$ is the estimated rate of change of the standardized channel value per sampling interval. The superscript T denotes matrix transpose, j identifies the sensor channel, and k denotes the discrete sample index.

The discrete local-linear model is

$$\boldsymbol{x}_{j,k} = \boldsymbol{F}\boldsymbol{x}_{j,k-1} + \boldsymbol{w}_{j,k}, \tag{5}$$
$$\tilde{z}_{j,k} = \boldsymbol{H}\boldsymbol{x}_{j,k} + v_{j,k}, \tag{6}$$

where $\boldsymbol{x}_{j,k}$ is the state vector of channel j at sample k, $\boldsymbol{x}_{j,k-1}$ is its state at the preceding sample, $\boldsymbol{F}$ is the state-transition matrix, and $\boldsymbol{w}_{j,k}$ is the process-noise vector representing unmodeled changes in the channel dynamics. The quantity $\tilde{z}_{j,k}$ is the standardized physical measurement supplied to the Kalman filter for channel j at sample k, $\boldsymbol{H}$ is the measurement matrix that maps the state vector to the measured quantity, and $v_{j,k}$ is the scalar measurement-noise term.

The state-transition and measurement matrices are

$$\boldsymbol{F} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad \boldsymbol{H} = \begin{bmatrix} 1 & 0 \end{bmatrix}. \tag{7}$$

Here, $\boldsymbol{F} \in \mathbb{R}^{2\times 2}$ propagates the estimated channel level and rate from sample $k-1$ to sample k under the local-linear assumption. Its $(1,2)$ element is unity because the rate state is defined per sampling interval. The matrix $\boldsymbol{H} \in \mathbb{R}^{1\times 2}$ selects the first state component, such that the physical measurement directly observes the standardized channel level but does not directly observe its rate of change.

The stochastic terms are modeled as $\boldsymbol{w}_{j,k} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{Q}_{j,k})$ and $v_{j,k} \sim \mathcal{N}(0, r_0)$. Here, $\mathcal{N}(\cdot,\cdot)$ denotes a Gaussian distribution, $\boldsymbol{0}$ is the zero-mean vector for the process noise, $\boldsymbol{Q}_{j,k} \in \mathbb{R}^{2\times 2}$ is the process-noise covariance matrix for channel j at sample k, 0 is the scalar mean of the measurement noise, and r_0 is the scalar measurement-noise variance.

A constant-acceleration process-noise structure is used,

$$\boldsymbol{Q}_{j,k} = q_{j,k}\boldsymbol{G}\boldsymbol{G}^T, \qquad \boldsymbol{G} = \begin{bmatrix} 1/2 & 1 \end{bmatrix}^T. \tag{8}$$

Here, $\boldsymbol{Q}_{j,k}$ is the process-noise covariance matrix, $q_{j,k}$ is the scalar process-noise intensity assigned to channel j at sample k, and $\boldsymbol{G} \in \mathbb{R}^{2\times 1}$ is the discrete process-noise distribution vector that maps an unmodeled change in rate into the level and rate states. The factor $1/2$ represents the contribution of the rate change to the channel level over one normalized sampling interval, while the second element, 1, represents its direct contribution to the rate state. $\boldsymbol{G}\boldsymbol{G}^T$ produces a 2×2 positive semi-definite covariance structure.

The fixed-covariance benchmark uses $q_{j,k} = q_0 = 0.01$ and $r_0 = 0.9$, matching the noise magnitudes used in the prior fixed-KF implementation after channel standardization. Here, q_0 denotes the nominal fixed process-noise intensity used by the conventional KF, while r_0 denotes the fixed measurement-noise variance.

C. Kalman Prediction and Measurement Update

For each channel, the time update is

$$\hat{\boldsymbol{x}}^-_{j,k} = \boldsymbol{F}\hat{\boldsymbol{x}}_{j,k-1}, \tag{9}$$
$$\boldsymbol{P}^-_{j,k} = \boldsymbol{F}\boldsymbol{P}_{j,k-1}\boldsymbol{F}^T + \boldsymbol{Q}_{j,k}. \tag{10}$$

Here, $\hat{\boldsymbol{x}}^-_{j,k}$ is the *a priori* predicted state estimate for channel j at sample k before incorporating the current measurement, $\hat{\boldsymbol{x}}_{j,k-1}$ is the corrected state estimate from the preceding sample, and $\boldsymbol{F}$ is the state-transition matrix. The matrix $\boldsymbol{P}^-_{j,k} \in \mathbb{R}^{2\times 2}$ is the predicted state-estimation error covariance, $\boldsymbol{P}_{j,k-1}$ is the corrected error covariance from the previous sample, and $\boldsymbol{Q}_{j,k}$ is the process-noise covariance added during state propagation.

When a physical measurement is available, the innovation and innovation covariance are

$$\nu_{j,k} = \tilde{z}_{j,k} - \boldsymbol{H}\hat{\boldsymbol{x}}_{j,k}^{-}, \tag{11}$$

$$S_{j,k} = \boldsymbol{H}\boldsymbol{P}_{j,k}^{-}\boldsymbol{H}^{T} + r_0. \tag{12}$$

Here, $\nu_{j,k}$ is the scalar innovation, or measurement residual, for channel j at sample k, defined as the difference between the standardized physical measurement $\tilde{z}_{j,k}$ and its one-step prediction $\boldsymbol{H}\hat{\boldsymbol{x}}_{j,k}^{-}$. The quantity $S_{j,k}$ is the scalar innovation covariance, representing the expected uncertainty of the innovation. The matrix $\boldsymbol{H}$ maps the predicted state to the measurement domain, $\boldsymbol{P}_{j,k}^{-}$ is the predicted state-error covariance, $\boldsymbol{H}^{T}$ is the transpose of the measurement matrix, and r_0 is the measurement-noise variance.

The Kalman gain, corrected state, and covariance are

$$\boldsymbol{K}_{j,k} = \boldsymbol{P}_{j,k}^{-}\boldsymbol{H}^{T}S_{j,k}^{-1}, \tag{13}$$

$$\hat{\boldsymbol{x}}_{j,k} = \hat{\boldsymbol{x}}_{j,k}^{-} + \boldsymbol{K}_{j,k}\nu_{j,k}, \tag{14}$$

$$\boldsymbol{P}_{j,k} = (\boldsymbol{I} - \boldsymbol{K}_{j,k}\boldsymbol{H})\boldsymbol{P}_{j,k}^{-}. \tag{15}$$

Here, $\boldsymbol{K}_{j,k} \in \mathbb{R}^{2\times 1}$ is the Kalman gain for channel j at sample k, which determines the relative weighting of the model prediction and the new measurement. The quantity $S_{j,k}^{-1}$ is the inverse of the scalar innovation covariance. The vector $\hat{\boldsymbol{x}}_{j,k}$ is the *a posteriori* corrected state estimate after incorporating the current innovation $\nu_{j,k}$, while $\hat{\boldsymbol{x}}_{j,k}^{-}$ is the corresponding *a priori* state prediction. The matrix $\boldsymbol{P}_{j,k}$ is the corrected state-estimation error covariance, $\boldsymbol{I} \in \mathbb{R}^{2\times 2}$ is the identity matrix, $\boldsymbol{K}_{j,k}\boldsymbol{H}$ represents the measurement-induced reduction in state uncertainty, and $\boldsymbol{P}_{j,k}^{-}$ is the predicted covariance prior to measurement assimilation.

When a measurement is intentionally skipped, only (10) is executed and $\hat{\boldsymbol{x}}_{j,k} = \hat{\boldsymbol{x}}_{j,k}^{-}$, $\boldsymbol{P}_{j,k} = \boldsymbol{P}_{j,k}^{-}$. Here, $\hat{\boldsymbol{x}}_{j,k}$ and $\boldsymbol{P}_{j,k}$ denote the state estimate and its covariance carried forward to the next sample. Because no measurement is available for correction, they are set equal to the predicted state $\hat{\boldsymbol{x}}_{j,k}^{-}$ and predicted covariance $\boldsymbol{P}_{j,k}^{-}$, respectively.

D. ACF-Guided Process-Covariance Adaptation

A correctly specified KF should produce an approximately white innovation sequence. Persistent positive innovation autocorrelation therefore indicates that the state model is not tracking a time-correlated component quickly enough. The proposed method computes the lag-one ACF over the most recent $L = 24$ available innovations (6 h at the nominal rate),

$$\rho_{j,k}^{\nu} = \mathrm{ACF}_1(\nu_{j,k-L:k-1}), \tag{16}$$

where $\rho_{j,k}^{\nu}$ is the lag-one autocorrelation coefficient of the innovation sequence for channel j at time step k, $\mathrm{ACF}_1(\cdot)$ denotes the autocorrelation function evaluated at lag one, $\nu_{j,k-L:k-1}$ denotes the sequence of the most recent L innovations from samples $k-L$ through $k-1$, j denotes the sensor channel, k denotes the current discrete time index, and $L = 24$ is the innovation-window length corresponding to 6 h at the 15-min sampling interval.

and the corresponding innovation variance $s_{\nu,j,k}^{2}$. The scalar process intensity is then updated as

$$q_{j,k} = \mathrm{clip}\Big(q_0 + \lambda\,\big[\rho_{j,k}^{\nu}\big]_{+}\,s_{\nu,j,k}^{2}, q_{\min}, q_{\max}\Big), \tag{17}$$

where $q_{j,k}$ is the adaptive process-noise intensity for channel j at time step k, $\mathrm{clip}(a, q_{\min}, q_{\max})$ constrains a to the interval $[q_{\min}, q_{\max}]$, $q_0 = 0.01$ is the nominal process-noise intensity, λ is the ACF adaptation gain, $\left[\rho_{j,k}^{\nu}\right]_{+} = \max(\rho_{j,k}^{\nu}, 0)$ retains only positive innovation autocorrelation, $\rho_{j,k}^{\nu}$ is the lag-one innovation autocorrelation, $s_{\nu,j,k}^{2}$ is the variance of the innovations within the current L-sample window, $q_{\min} = 10^{-6}$ is the minimum allowable process-noise intensity, and $q_{\max} = 5$ is the maximum allowable process-noise intensity.

where $[a]_{+} = \max(a, 0)$, $q_{\min} = 10^{-6}$, and $q_{\max} = 5$. This construction retains $\boldsymbol{Q}_{j,k} \succeq 0$ and avoids assigning a potentially negative ACF value directly to a covariance. Measurement covariance r_0 is held fixed so that process-model mismatch and measurement noise are not adjusted by the same statistic.

The adaptation gain λ is selected using the learning set only. Candidate values $\{0.25, 0.5, 1, 2, 4, 8, 16\}$ are tested after a 24-sample (6-h) warm-up. A minimax non-degradation rule selects the largest mean improvement subject to non-increasing learning RMSE for every channel; for the present data this gives $\lambda = 0.25$. The value is then fixed for all held-out analyses.

E. Calibrated 95% Prediction Bands

The nominal one-step predictive standard deviation in standardized coordinates is

$$s_{j,k}^{-} = \sqrt{\boldsymbol{H}\boldsymbol{P}_{j,k}^{-}\boldsymbol{H}^{T} + r_0}. \tag{18}$$

where $s_{j,k}^{-}$ is the *a priori* one-step predictive standard deviation for channel j at time step k, $\boldsymbol{H}$ is the measurement matrix, $\boldsymbol{P}_{j,k}^{-}$ is the *a priori* state-estimation error covariance, and $r_0 = 0.9$ is the fixed measurement-noise variance.

To compensate for covariance-model mismatch without using evaluation data, a learning-only calibration factor is computed from the normalized absolute residual score

$$a_{j,k} = \frac{|\tilde{y}_{j,k} - \hat{\tilde{y}}_{j,k}^{-}|}{s_{j,k}^{-}}, \qquad k \in \mathcal{L}. \tag{19}$$

where $a_{j,k}$ is the normalized absolute one-step prediction error for channel j at time step k, $\tilde{y}_{j,k}$ is the standardized observed sensor value, $\hat{\tilde{y}}_{j,k}^{-}$ is its *a priori* standardized one-step prediction, $s_{j,k}^{-}$ is the corresponding predictive standard deviation, k is the sample index, and $\mathcal{L}$ denotes the set of samples belonging to the learning interval.

Let $c_{j,0.95}$ be the finite-sample 95th percentile of $a_{j,k}$ after warm-up. The plotted 95% predictive band in physical units is

$$\hat{y}_{j,k}^{-} \pm c_{j,0.95}\,\sigma_{j,L}s_{j,k}^{-}. \tag{20}$$

where $\hat{y}_{j,k}^{-}$ is the *a priori* one-step prediction of channel j expressed in its original physical units, $c_{j,0.95}$ is the learning-derived finite-sample 95% calibration factor for channel j,

$\sigma_{j,L}$ is the standard deviation of channel j calculated from the learning set, $s^{-}_{j,k}$ is the predictive standard deviation in standardized coordinates, j denotes the sensor channel, k denotes the time index, and $\pm$ defines the lower and upper prediction-band limits.

This interval expands automatically during open-loop prediction because $\boldsymbol{P}^{-}_{j,k}$ grows when measurements are skipped.

IV. Adaptive Measurement Scheduler

A. Decision Variables

After the 24-h learning interval, the recorded data are processed sequentially in chronological order. At each 15-min candidate time, the estimator uses only information available at or before the current time step and first predicts all seven channels. Temperature is then acquired as a low-power sentinel. A full sensing cycle is requested when any of four conditions is satisfied.

The normalized temperature innovation is

$$D_{T,k} = |\tilde{T}_k - \hat{\tilde{T}}^{-}_k|. \tag{21}$$

where $D_{T,k}$ is the absolute normalized temperature innovation at time step k, $\tilde{T}_k$ is the standardized measured temperature, $\hat{\tilde{T}}^{-}_k$ is the *a priori* standardized temperature prediction, T identifies the temperature channel, and k denotes the current time index.

The maximum non-temperature state-uncertainty score is

$$U_k = \max_{j \neq T} \left(1.96\sqrt{P^{-}_{j,k}(1,1)}\right), \tag{22}$$

where U_k is the maximum predicted uncertainty among all non-temperature channels at time step k, $\max_{j \neq T}$ denotes the maximum over all channels j except the temperature channel T, 1.96 is the standard-normal multiplier associated with an approximate two-sided 95% interval, $P^{-}_{j,k}(1,1)$ is the first diagonal element of the predicted covariance matrix corresponding to uncertainty in the channel-level state, j denotes the sensor channel, and k denotes the time index.

and the largest predicted deviation from the most recent full measurement is

$$D_k = \max_{j \neq T} \left|\hat{\tilde{y}}^{-}_{j,k} - \hat{\tilde{y}}_{j,k_f}\right|, \tag{23}$$

where D_k is the largest predicted change among the non-temperature channels at time step k, $\hat{\tilde{y}}^{-}_{j,k}$ is the *a priori* standardized prediction for channel j at the current time step, $\hat{\tilde{y}}_{j,k_f}$ is the standardized channel estimate associated with the most recent full sensing cycle, $j \neq T$ denotes all channels except temperature, and k_f is the index of the most recent full sensing cycle.

where k_f is the index of the latest full sensing cycle. The trigger is

$$\begin{aligned} \gamma_k = 1 \quad \text{if} \quad & D_{T,k} \geq \theta_T \ \vee \ U_k \geq \theta_U \\ & \vee \ D_k \geq \theta_D \ \vee \ k - k_f \geq G_{\max}, \end{aligned} \tag{24}$$

where γ_k is the binary full-sensing decision variable at time step k, with $\gamma_k = 1$ indicating execution of a full sensing cycle, $D_{T,k}$ is the normalized temperature innovation, θ_T is the temperature-innovation threshold, U_k is the maximum non-temperature state-uncertainty score, θ_U is the uncertainty threshold, D_k is the maximum predicted deviation from the most recent full measurement, θ_D is the predicted-deviation threshold, $\vee$ denotes the logical OR operation, $k - k_f$ is the number of base sampling intervals elapsed since the most recent full sensing cycle, and $G_{\max}$ is the maximum permitted number of intervals between full sensing cycles; otherwise $\gamma_k = 0$.

The design values $\theta_T = 0.5$, $\theta_D = 0.5$, and $G_{\max} = 8$ base intervals are fixed in standardized units. The uncertainty threshold is selected without evaluation data by processing the final 18 h of the learning window after a 6-h filter warm-up and choosing the closest value in $\{0.5, 0.6, \ldots, 4.0\}$ to a 55% full-cycle-reduction target. This procedure selects $\theta_U = 2.2$, corresponding to a 54.17% reduction within the learning-window analysis. The maximum-gap rule limits open-loop operation to 2 h.

Algorithm 1: ACF–KF adaptive sensing at time k

1: Predict $\hat{\boldsymbol{x}}^{-}_{j,k}$ and $\boldsymbol{P}^{-}_{j,k}$ for all channels.
2: Acquire the low-power temperature sentinel T_k.
3: Update the temperature KF and compute $D_{T,k}$.
4: Compute U_k and D_k from the non-temperature predictors.
5: **if** $D_{T,k} \geq \theta_T$ OR $U_k \geq \theta_U$ OR $D_k \geq \theta_D$ OR gap $\geq G_{\max}$ **then**
6: Set $\gamma_k \leftarrow 1$ and execute the full sensing cycle.
7: Update all channel KFs with the new measurements.
8: **else**
9: Set $\gamma_k \leftarrow 0$; retain predicted states for non-temperature channels.
10: **end if**
11: Update each available innovation window and evaluate (17) for the next step.

Sequential measurement-scheduling logic used in the offline evaluation.

B. Energy Interpretation

Let E_s be the energy required for the always-available sentinel operation at one base interval and E_f the incremental energy of a full optical/pump/telemetry cycle. For N_E candidate times and N_f executed full cycles,

$$E_{\text{fixed}} = N_E(E_s + E_f), \tag{25}$$

$$E_{\text{adapt}} = N_E E_s + N_f E_f. \tag{26}$$

where E_{fixed} is the total energy required by the fixed-sampling strategy, E_{adapt} is the total energy required by the adaptive strategy, N_E is the total number of candidate sensing times, N_f is the number of full sensing cycles actually executed by the adaptive scheduler, E_s is the energy required for the sentinel operation during one base interval, and E_f is the incremental energy required to execute one full optical measurement, pumping, and telemetry cycle.

Hence, the energy reduction is

$$\eta_E = 1 - \frac{N_E E_s + N_f E_f}{N_E(E_s + E_f)}. \tag{27}$$

where η_E is the fractional energy reduction achieved by adaptive sensing relative to fixed sensing, E_s is the energy consumed by one sentinel operation, and E_f is the incremental energy consumed by one full sensing cycle.

When the pump and optical sequence dominate ($E_f \gg E_s$), η_E approaches the reduction in full sensing cycles. This work reports the full-cycle reduction directly rather than assigning a joule value. The pump requirements reported for the underlying platform [10] motivate this duty-cycle metric.

V. Evaluation Protocol

A. Predictive-Accuracy Comparison

The fixed-covariance KF and ACF–KF are both run with a physical measurement at every held-out 15-min slot. Performance is calculated using the *a priori* one-step prediction before the measurement update; therefore the comparison evaluates true predictive tracking rather than post-update smoothing. For errors $e_k = y_k - \hat{y}_k^-$,

$$\mathrm{RMSE} = \sqrt{\frac{1}{N_E}\sum e_k^2}, \tag{28}$$

$$\mathrm{MAE} = \frac{1}{N_E}\sum |e_k|, \tag{29}$$

$$R^2 = 1 - \frac{\sum e_k^2}{\sum (y_k - \bar{y})^2}, \tag{30}$$

$$\mathrm{NRMSE} = 100\frac{\mathrm{RMSE}}{y_{\max} - y_{\min}}. \tag{31}$$

where e_k is the one-step prediction error at sample k, y_k is the observed sensor value, $\hat{y}_k^-$ is the corresponding *a priori* one-step prediction, N_E is the number of held-out evaluation samples, RMSE is the root-mean-square error, MAE is the mean absolute error, R^2 is the coefficient of determination, $\bar{y}$ is the mean of the observed evaluation values, NRMSE is the normalized root-mean-square error expressed as a percentage, and $y_{\max}$ and $y_{\min}$ are the maximum and minimum observed values, respectively, over the evaluated dataset.

The percentage RMSE improvement is $100(\mathrm{RMSE}_{KF} - \mathrm{RMSE}_{ACF})/\mathrm{RMSE}_{KF}$, where RMSE_{KF} is the held-out RMSE of the fixed-covariance Kalman filter and RMSE_{ACF} is the held-out RMSE of the ACF–KF.

B. Change-Tracking Response

To quantify how quickly the estimator recovers after a change, a channel-specific event threshold is formed from the learning data as the 80th percentile of $|\tilde{y}_{j,k} - \tilde{y}_{j,k-1}|$. Held-out events above this threshold are separated by at least five samples to avoid repeatedly counting one transient. The tracking tolerance is the fixed-KF learning RMSE after warm-up. Recovery time is the first 15-min horizon at which the one-step prediction error is below this tolerance for two consecutive predictions, capped at 120 min. Because the event threshold and tolerance are learned before the holdout period, the response-time comparison does not tune on evaluation data.

C. Adaptive Scheduling Evaluation

For scheduling evaluation, all non-temperature measurements are withheld from the estimator whenever $\gamma_k = 0$ but retained as reference measurements for performance assessment. At $\gamma_k = 1$, the recorded measurement is revealed and the KF is corrected. Reconstruction metrics are computed at every held-out 15-min time, including skipped instants. The full-cycle reduction is

$$\eta_M = 100\left(1 - \frac{N_f}{N_E}\right)\%. \tag{32}$$

where η_M is the percentage reduction in executed full sensing cycles relative to measurement at every candidate time.

D. Fixed Measurement-Frequency Study

To isolate the effect of measurement frequency, the complete held-out sequence is evaluated using deterministic full-measurement intervals of 15, 30, 45, 60, 90, 120, 180, and 240 min. Between measured points, the ACF–KF operates in an open loop. The hidden 15-min data are used only to score reconstruction accuracy. This experiment therefore estimates how aggressively the acquisition period can be lengthened before the temporal model becomes inadequate.

VI. Results and Discussion

A. Temporal Redundancy

Figure 2 and Table I show that all seven channels are strongly correlated at one 15-min lag. Lag-one coefficients range from 0.827 for turbidity to 0.957 for temperature. PC and CHL channels also remain correlated over multiple lags. The first statistically nonsignificant ACF lag occurs no earlier than 135 min, and for temperature it occurs at 240 min. These results do not imply that the sensing process can safely wait several hours under all conditions; they show that the process contains sufficient temporal structure to justify predictive scheduling rather than unconditional 15-min full sensing.

B. Fixed KF Versus ACF–KF

The held-out one-step predictions for all seven channels are shown in Fig. 3. Table II summarizes the held-out one-step predictions. The ACF–KF improves RMSE and MAE for every channel. RMSE reduction ranges from 1.90% for CHL-High to 6.72% for temperature, while MAE reduction ranges from 2.87% for battery voltage to 8.49% for turbidity. The greatest absolute improvement occurs for temperature, whose RMSE decreases from 0.784 to 0.732 °C. For the optical channels, improvements are smaller because the 15-min observations are already highly persistent, but the proposed adaptation remains non-degrading on all held-out channels. R^2 also increases for every signal.

The improvement mechanism is consistent with innovation-whiteness theory. When several consecutive residuals have

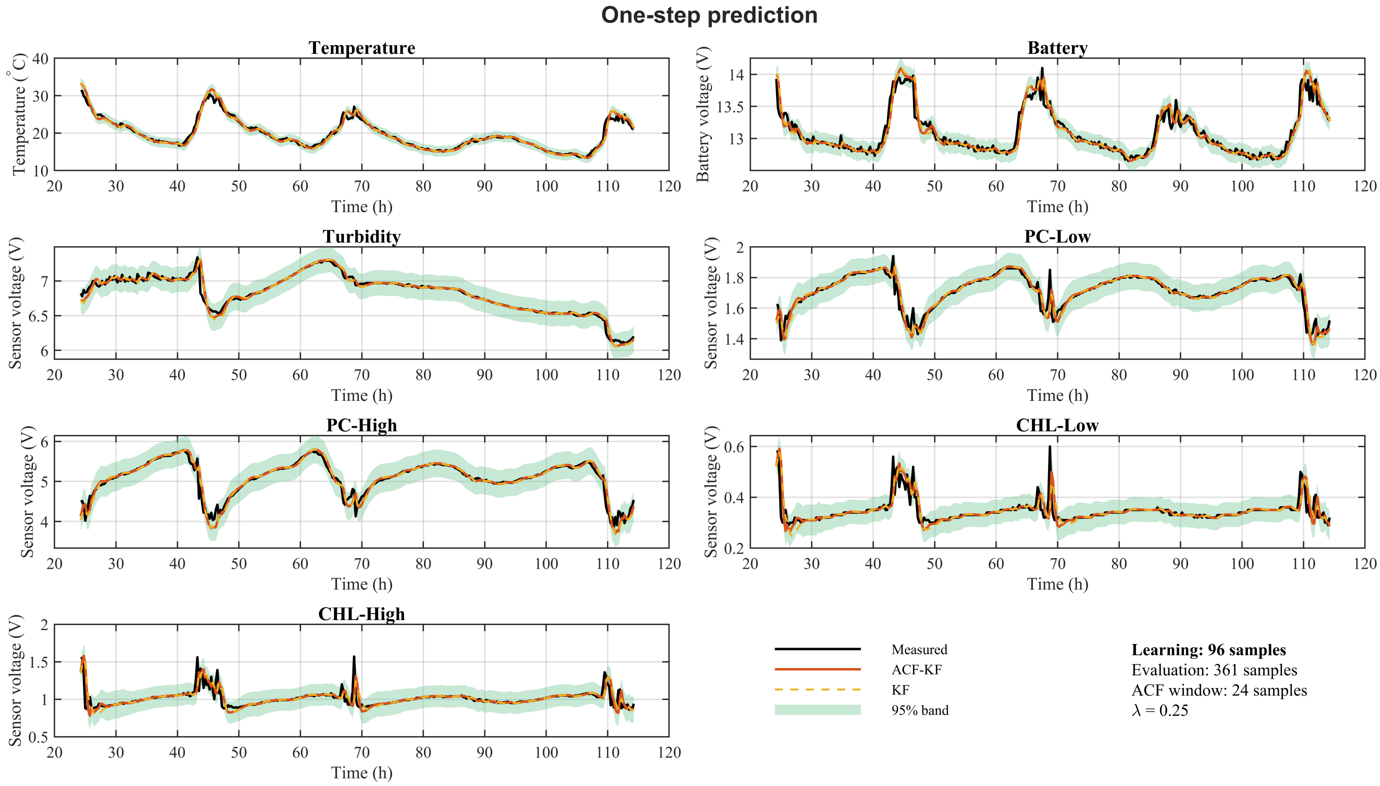

Fig. 3. Held-out one-step prediction for all channels. The proposed ACF–KF is compared with the fixed-covariance KF; shaded regions are learning-calibrated 95% prediction bands for ACF–KF.

the same sign, positive $\rho^{\nu}_{j,k}$ increases $q_{j,k}$, enlarging the predicted covariance and the Kalman gain at the next available correction. The estimator therefore responds faster to sustained changes than the fixed-covariance model. When the innovation is approximately white or negatively correlated, $q_{j,k}$ returns toward q_0 and unnecessary responsiveness is avoided.

The channel-wise prediction improvement and change-tracking response are summarized in Fig. 4. The change-response analysis identifies 127 non-overlapping upper-quintile events across the seven channels. The pooled median recovery time decreases from 45 min for the fixed KF to 30 min for ACF–KF. Median recovery is lower for temperature, battery, both PC channels, and both CHL channels, and unchanged for turbidity. This result directly quantifies the faster adaptation visible during sustained transients. The nominal 95% prediction bands obtain held-out empirical coverage between 88.9% and 98.9% across channels, with a mean coverage of 93.8%. The bands are therefore useful uncertainty indicators, but they are not interpreted as an exact finite-sample coverage guarantee under serial dependence.

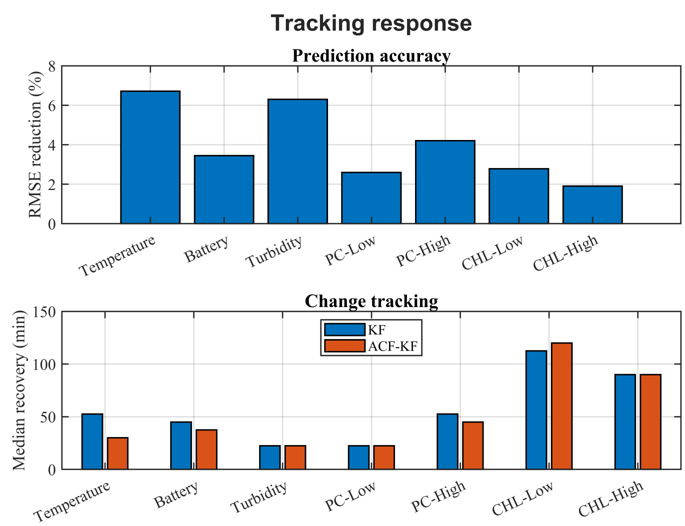

Fig. 4. Prediction and change-tracking response. Top: held-out RMSE reduction of ACF–KF relative to the fixed KF. Bottom: channel-wise median recovery time after learning-defined upper-quintile changes.

C. Adaptive Sensing Evaluation

The resulting adaptive measurement schedule and representative reconstructions for turbidity, phycocyanin, and chlorophyll-a are shown in Fig. 5. The scheduler executes 171 full sensing cycles over the 361-point evaluation set instead of 361 fixed 15-min cycles. This corresponds to a 52.63% reduction in full measurements. The mean and median full-cycle intervals are 31.7 and 30 min, respectively. The realized intervals are 15, 30, and 45 min because the dynamic and uncertainty triggers fire before the maximum-gap condition becomes active.

Table III reports reconstruction accuracy at every 15-min

TABLE II
HELD-OUT ONE-STEP PREDICTION PERFORMANCE: FIXED KF VS. ACF–KF

Channel	RMSE KF	RMSE ACF–KF	RMSE impr.	MAE KF	MAE ACF–KF	MAE impr.	R^2 KF	R^2 ACF–KF
Temperature (°C)	0.7842	0.7316	6.72%	0.4873	0.4503	7.59%	0.9595	0.9648
Battery (V)	0.1156	0.1116	3.45%	0.0722	0.0701	2.87%	0.8843	0.8922
Turbidity output	0.0516	0.0484	6.30%	0.0267	0.0245	8.49%	0.9625	0.9671
PC-Low output	0.0395	0.0385	2.60%	0.0198	0.0186	6.27%	0.8674	0.8742
PC-High output	0.1210	0.1159	4.20%	0.0651	0.0604	7.16%	0.9181	0.9248
CHL-Low output	0.0335	0.0326	2.78%	0.0150	0.0139	7.42%	0.4091	0.4415
CHL-High output	0.0818	0.0803	1.90%	0.0349	0.0329	5.86%	0.3488	0.3734

TABLE III
ADAPTIVE-SCHEDULER RECONSTRUCTION PERFORMANCE

Channel	RMSE	MAE	NRMSE	R^2
Temperature	0.4426	0.2730	2.49%	0.9871
Battery	0.0986	0.0583	6.80%	0.9159
Turbidity	0.0348	0.0196	2.80%	0.9830
PC-Low	0.0328	0.0160	5.96%	0.9089
PC-High	0.1042	0.0500	5.40%	0.9393
CHL-Low	0.0253	0.0112	8.16%	0.6638
CHL-High	0.0654	0.0273	9.21%	0.5842
Full-cycle reduction: 52.63%; mean interval: 31.7 min; mean NRMSE: 5.83%.				

ground-truth instant. The mean NRMSE across all seven channels is 5.83%. Temperature has the smallest NRMSE (2.49%) because it is the sentinel and is corrected at every base interval. Turbidity remains below 3% NRMSE. The low-amplitude CHL channels have the largest normalized errors because their evaluation ranges are small, although their absolute RMSE remains 0.0253 and 0.0654 in recorded-output units.

D. Effect of Measurement Frequency

Table IV and Fig. 6 quantify the fixed-interval accuracy trade-off using the proposed ACF–KF. At the native 15-min cadence, the mean reconstruction NRMSE is 4.18%. Doubling the interval to 30 min removes approximately half of the full measurements and increases mean NRMSE to 5.62%. A 45-min interval gives 6.33% mean NRMSE, while 60 and 90 min produce 8.10% and 9.95%, respectively. Beyond 120 min the error increases rapidly, reaching 18.87% at 180 min and 23.02% at 240 min.

The result demonstrates why a purely fixed longer interval is undesirable. A 30–45-min interval is often adequate for this record, but a fixed schedule cannot know when a fast transient occurs. The adaptive controller instead retains the 15-min decision clock and lengthens the full-measurement interval only when the temperature sentinel, predicted dynamics, and covariance indicate that open-loop prediction remains acceptable. At each 15-min decision step, the temperature channel is first predicted by the ACF–KF and then physically measured. The difference between the measured temperature and its *a priori* prediction forms $D_{T,k}$, while the temperature KF is subsequently corrected using the new measurement.

E. Scalability and Embedded Implementation

For M channels, the method maintains M independent 2×2 KFs and an innovation buffer of length L per channel. The per-step matrix operations are constant size, so computation grows linearly as $\mathcal{O}(M)$ and memory as $\mathcal{O}(ML)$. This structure is suitable for edge execution on a microcontroller and does not require cloud connectivity for the scheduling decision. In a larger LoRaWAN deployment, each node can execute the

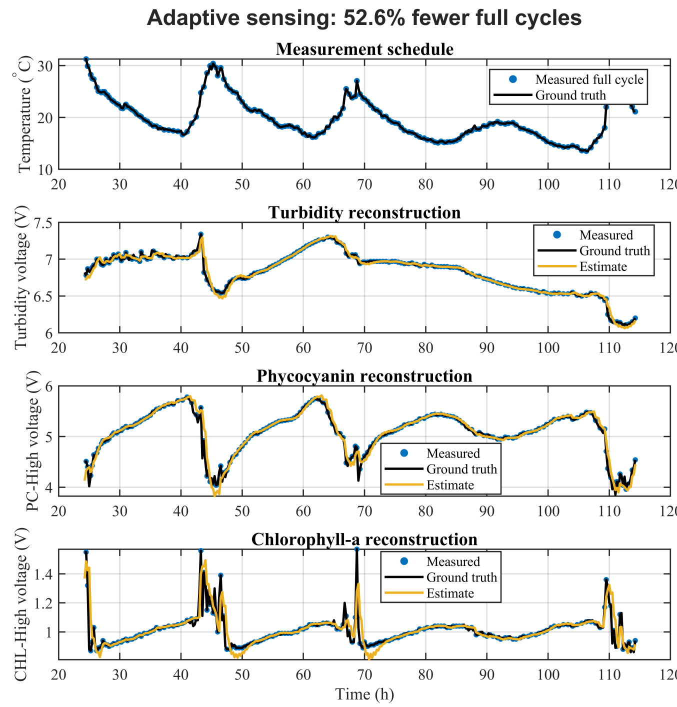

Fig. 5. Adaptive sensing evaluated on the recorded dataset. The panels show the temperature-based measurement schedule and representative reconstructions for turbidity, phycocyanin (PC-High), and chlorophyll-a (CHL-High). Full-cycle measurements are triggered by the temperature innovation, predicted state change, estimator uncertainty, or maximum-gap condition; measurements at skipped instants are retained only as reference data for performance evaluation.

TABLE IV
FIXED MEASUREMENT-INTERVAL SENSITIVITY WITH ACF–KF

Full-measurement interval	Full-cycle reduction	Mean NRMSE
15 min	0.00%	4.18%
30 min	49.86%	5.62%
45 min	66.48%	6.33%
60 min	74.79%	8.10%
90 min	83.10%	9.95%
120 min	87.26%	12.29%
180 min	91.41%	18.87%
240 min	93.63%	23.02%

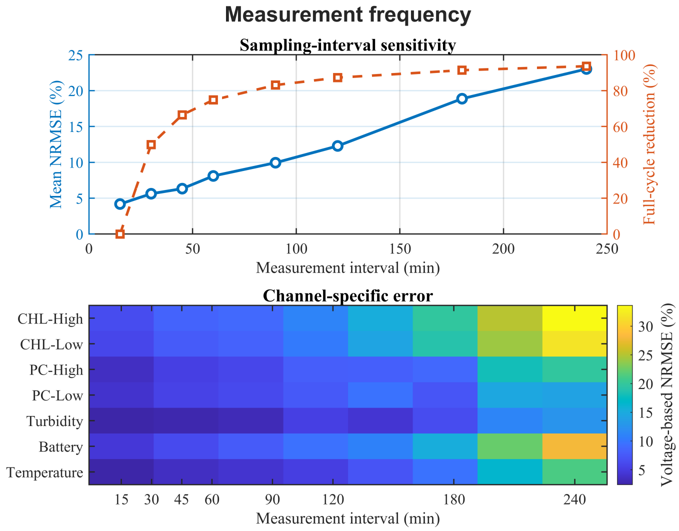

Fig. 6. Measurement-frequency sensitivity. Mean reconstruction NRMSE increases as the full sensing interval grows; the accompanying heat map shows channel-specific sensitivity.

same local estimator using node-specific scaling and covariance statistics, while the gateway receives only the selected measurements and status information. Node-local adaptation also avoids forcing a single noise model onto sensors with different optical gains or deployment conditions.

F. Limitations

The present results are based on an offline sequential analysis of one approximately five-day dataset; the adaptive scheduler was not active during acquisition. Therefore, the reported reduction in full sensing cycles is a counterfactual estimate rather than a direct field measurement of battery-life extension. The optical fields are analyzed in their recorded output units rather than converted concentrations, and the learned thresholds may require retuning when water bodies, seasons, optical gains, or sensor fouling conditions change. The model treats channels independently and uses temperature as the always-available sentinel; future implementations can additionally exploit cross-channel covariance or substitute another low-power trigger. Finally, Eq. (27) shows that actual energy savings depend on the relative energies E_s and E_f, which should be measured on hardware in a future deployment.

VII. Conclusion

This work developed an autocorrelation-aware Kalman method for adaptive measurement scheduling in a multichannel LoRaWAN water-quality sensor. A 24-h learning window revealed strong lag-one correlation across all seven recorded fields, motivating predictive suppression of redundant full sensing cycles. Unlike a heuristic that directly assigns ACF values to covariance matrices, the proposed method uses positive innovation autocorrelation as a statistically interpretable signal of under-modeled dynamics and increases only the process covariance while preserving positive semidefiniteness. On 361 held-out samples, ACF–KF reduced one-step RMSE and MAE for every channel relative to a fixed-covariance KF. When evaluated on the recorded data, the scheduler reduced full sensing cycles by 52.6% with a mean 31.7-min full-cycle interval and 5.83% mean reconstruction NRMSE. The fixed-frequency analysis further showed a clear increase in reconstruction error between approximately 45 and 120 min, supporting adaptive rather than permanently lengthened sampling intervals. Future work will implement the controller on the sensor node, measure real energy consumption and battery lifetime, evaluate seasonal and multi-site datasets, and investigate joint multichannel state models and online measurement-noise adaptation.

Acknowledgment

The author would like to thank Tian Xia of the University of Vermont for providing access to laboratory facilities, equipment, and institutional resources that supported this work. The author also thanks Clayton J. Williams of Saint Michael's College for providing a portion of the experimental field dataset used in this study. The author acknowledges the University of Vermont for providing the facilities and resources associated with the development and evaluation of this work.

References

[1] R. K. Singh, P. P. Puluckul, R. Berkvens, and M. Weyn, "Energy consumption analysis of LPWAN technologies and lifetime estimation for IoT application," *Sensors*, vol. 20, no. 17, pp. 1–22, 2020, doi: 10.3390/s20174794.

[2] G. Callebaut, G. Leenders, J. Van Mulders, G. Ottoy, L. De Strycker, and L. Van der Perre, "The art of designing remote IoT devices—technologies and strategies for a long battery life," *Sensors*, vol. 21, no. 3, pp. 1–37, 2021, doi: 10.3390/s21030913.

[3] A. G. Rajakumar, M. S. Mohan Kumar, B. Amrutur, and Z. Kapelan, "Real-time water quality modeling with ensemble Kalman filter for state and parameter estimation in water distribution networks," *J. Water Resour. Plan. Manag.*, vol. 145, no. 11, pp. 1–12, 2019, doi: 10.1061/(ASCE)WR.1943-5452.0001118.

[4] M. Khodarahmi and V. Maihami, "A review on Kalman filter models," *Arch. Comput. Methods Eng.*, vol. 30, no. 1, pp. 727–747, 2023, doi: 10.1007/s11831-022-09815-7.

[5] J. Kern, E. Dupraz, A. Aïssa-El-Bey, and F. Leduc-Primeau, "Improving the energy-efficiency of a Kalman filter using unreliable memories," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2021, pp. 5345–5349, doi: 10.1109/ICASSP39728.2021.9413430.

[6] A. I. Sunny, A. Zhao, L. Li, and S. Kanteh Sakiliba, "Low-cost IoT-based sensor system: A case study on harsh environmental monitoring," *Sensors*, vol. 21, no. 1, pp. 1–12, 2021, doi: 10.3390/s21010214.

[7] M. A. M. Yazid, A. Jazlan, M. Z. M. Rodzi, M. A. Husman, and A. R. Afif, "A method for preserving battery life in wireless sensor nodes for LoRa based IoT flood monitoring," *J. Commun.*, vol. 17, no. 4, pp. 230–238, 2022, doi: 10.12720/jcm.17.4.230-238.

[8] S. C. Folea and G. Mois, "A low-power wireless sensor for online ambient monitoring," *IEEE Sensors J.*, vol. 15, no. 2, pp. 742–749, Feb. 2015, doi: 10.1109/JSEN.2014.2351420.

[9] M. Hulea, G. Mois, S. Folea, L. Miclea, and V. Biscu, "Wi-sensors: A low power Wi-Fi solution for temperature and humidity measurement," in *Proc. 39th Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, 2013, pp. 4011–4015, doi: 10.1109/IECON.2013.6699777.

[10] S. F. Hagh, P. Amngostar, W. Chen, J. Sheridan, C. J. Williams, A. M. Morales-Williams, D. Huston, and T. Xia, "A low-cost LoRa optical fluorometer–nephelometer for wireless monitoring of water quality parameters in real-time," *IEEE Sensors J.*, vol. 24, no. 13, pp. 21511–21519, Jul. 2024, doi: 10.1109/JSEN.2024.3403416.

[11] "State estimation: Kalman filters," Reality Bytes, Aug. 2017. [Online]. Available: https://realitybytes.blog/2017/08/15/state-estimation-kalman-filters

[12] X. Chen, Z. Gao, R. Ma, and X. Huang, "Hybrid extended-unscented Kalman filters for continuous-time nonlinear fractional-order systems involving process and measurement noises," *Trans. Inst. Meas. Control*, vol. 42, no. 9, pp. 1618–1631, 2020, doi: 10.1177/0142331219893788.

[13] J. Mu, F. Tian, C. Wang, and C. Jian, "Adaptive Masreliez–Martin fractional interpolatory cubature Kalman filter with recursive noise estimation," *J. Vib. Control*, vol. 29, no. 17–18, pp. 3907–3924, 2022, doi: 10.1177/10775463221106835.

[14] V. Mussot, G. Mercère, T. Dairay, V. Arvis, and J. Vayssettes, "Noise covariance matrix estimation with subspace model identification for Kalman filtering," *Int. J. Adapt. Control Signal Process.*, vol. 35, no. 4, pp. 591–611, 2021, doi: 10.1002/acs.3213.

[15] C. Shan, W. Zhou, Y. Yang, and Z. Jiang, "Multi-fading factor and updated monitoring strategy adaptive Kalman filter-based variational Bayesian," *Sensors*, vol. 21, no. 1, p. 198, 2020, doi: 10.3390/s21010198.

[16] R. Tripathi, A. Singh, and P. Gangwar, "Innovation-based fractional order adaptive Kalman filter," *J. Electr. Eng.*, vol. 71, no. 1, pp. 60–64, 2020, doi: 10.2478/jee-2020-0009.

[17] Z. Wang, J. Cheng, B. Qi, S. Cheng, and S. Chen, "An adaptive Kalman filtering algorithm based on maximum likelihood estimation," *Meas. Sci. Technol.*, vol. 34, no. 11, p. 115114, 2023, doi: 10.1088/1361-6501/ace9ef.