Received: 3 December 2025 | Revised: 10 March 2026 | Accepted: 1 April 2026 | Published online: 8 April 2026

RESEARCH ARTICLE

Archives of Advanced Engineering Science
2026, Vol. 00(00) 1–8
DOI: 10.47852/bonviewAAES62028681

BON VIEW PUBLISHING

Distributed Estimation of Ambient Temperature with Missing Data

Miguel A. Vazquez-Olguin[1], Oscar G. Ibarra-Manzano[1] and Yuriy S. Shmaliy[1,*]

[1] *Department of Electronics, University of Guanajuato, Mexico*

Abstract: Air quality and weather monitoring over large geographical areas is typically performed using expensive equipment, which limits spatial resolution and coverage. Wireless sensor networks (WSN) technology enables the large-scale deployment of affordable sensing nodes. However, such networks operate under challenging conditions, including noisy measurements, missing data, faulty sensor readings, and the absence of reliable noise statistics. These factors significantly degrade the performance of conventional distributed estimators that rely on accurate stochastic models. Because low-cost sensors are typically not very reliable and inherently noisy, distributed estimators exploit spatial redundancy to mitigate individual sensor imperfections. In this work, we develop a distributed unbiased finite impulse response (UFIR) filtering framework to conduct robust ambient temperature estimation in WSNs under realistic data-quality constraints. The proposed method combines local UFIR estimation, which does not require prior knowledge of noise statistics, with a consensus-based information fusion mechanism that enables cooperative estimation across the network. This structure enhances robustness against missing measurements and allows the evaluation of performance in the presence of faulty sensor readings and the absence of reliable knowledge about noise statistics while preserving distributed operation. The solution is validated using real environmental measurements containing missing data, together with an analysis of the influence of network connectivity on estimation performance. Results demonstrate that the proposed estimator achieves accurate and stable temperature estimation across the network even under significant data degradation, confirming its suitability for real-world WSN monitoring applications.

Keywords: wireless sensor networks, distributed estimation, missing data, environmental monitoring

1. Introduction

Large-scale environmental and weather monitoring systems rely on measurements collected over wide geographic areas using networks of sensing stations. In many practical deployments, the high cost of precision instrumentation limits the number of stations that can be installed, which in turn restricts spatial resolution and redundancy. To overcome this limitation, wireless sensor network (WSN) technologies have emerged as an effective alternative, enabling the use of low-cost sensors deployed in large numbers. However, the reduced reliability of such sensors, together with harsh environmental conditions, frequently leads to missing data, corrupted measurements, and uncertain noise characteristics.

From an estimation standpoint, these conditions pose significant challenges to classical centralized processing schemes. Decentralized estimation strategies have gained relevance as they allow cooperative processing among neighboring nodes while reducing communication and computation demands by exploiting information redundancy. Among existing estimators, Kalman filter (KF)–based solutions are widely adopted due to their optimality under ideal assumptions. Nevertheless, their performance may degrade severely when noise statistics are unknown, measurements are intermittent, or sensor faults occur, which are common situations in real-world WSN applications.

To address these limitations, finite-horizon estimators have been shown to provide improved robustness by avoiding reliance on precise statistical information and long-term model assumptions. In particular, the unbiased finite impulse response (UFIR) filtering framework has demonstrated strong resilience to modeling uncertainties, missing observations, and non-Gaussian disturbances. When combined with distributed consensus mechanisms, UFIR-based methods become especially attractive for WSN scenarios, where local processing and limited communication are essential.

In this work, we develop and analyze a distributed implementation of the UFIR filter based on an estimate-consensus protocol (distributed UFIR [dUFIR]) for distributed ambient temperature monitoring. The proposed framework integrates local finite-horizon unbiased estimation with a matrix-weighted consensus correction mechanism, enabling fully distributed operation without requiring prior knowledge of noise covariance statistics. The formulation explicitly accounts for intermittent measurements and constrained connectivity, and its behavior in the presence of persistently faulty readings is evaluated experimentally rather than addressed through explicit fault-detection mechanisms. Using real-world environmental data, the dUFIR estimator is systematically assessed in terms of noise attenuation, data reconstruction capability, and consensus behavior among neighboring nodes. The results demonstrate that reliable estimation can be achieved

*__Corresponding author:__ Yuriy S. Shmaliy, Department of Electronics, University of Guanajuato. Email: shmaliy@ugto.mx

under realistic operating conditions, highlighting the practical applicability of the proposed dUFIR framework.

2. Literature Review

The use of WSN technologies for weather monitoring has become popular in recent years [1–5]. This imposes important restrictions on the data acquisition techniques, which are due to limited battery life and processing power [6–9]. This requires the development of appropriate algorithms to ensure the efficient use of WSN resources [10–13]. From this perspective, distributed filtering contributes to energy efficiency by reducing the local computational load while still enabling real-time estimation. Using distributed methods, each node estimates a quantity Q, and a consensus procedure is used to combine the estimates, measurements, or information across the network.

The KF is most widely used to fuse sensor data [14–17] primarily due to its optimality and low computational complexity. Accordingly, we find many efficient KF-based state estimation algorithms addressing the consensus problem in WSNs. A solution proposed in Olfati-Saber's study [18] requires each node to locally aggregate data and the covariance matrices taken from the neighbors. At the posterior phase, an estimate is computed using a KF with a consensus term; for example, in the study conducted by Carli et al. [19], the fusion KF has been developed for local estimation using a consensus matrix. More recent distributed observer-based frameworks have extended these ideas to scenarios involving switching communication topologies, unknown inputs, and quantized information exchange. For instance, the work conducted by Yang et al. [20] establishes convergence guarantees under jointly connected communication graphs for linear time-invariant (LTI) systems, while the study by Yu et al. [21] derives estimator gains by minimizing upper bounds on the error covariance in the presence of quantized communication constraints. Additionally, distributed estimation under stochastic communication failures and energy-aware transmission mechanisms has been investigated in the study by Chen et al. [22]. These contributions provide rigorous mathematical guarantees under explicit stochastic modeling assumptions and well-defined system dynamics. Despite their strong theoretical foundation, such model-based distributed observer and covariance-driven estimation frameworks rely on explicit system modeling assumptions and structured statistical descriptions.

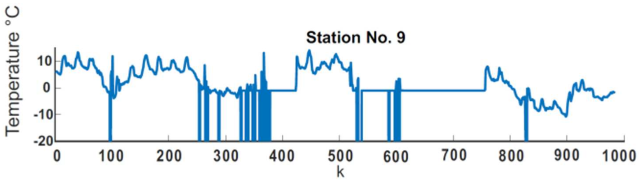
It is worth noting that the KF optimality is guaranteed only under complete knowledge of the Gaussian noise statistics, accurate process modeling, and initial conditions [23]. If these conditions are not met, the KF performance may significantly degrade and become unacceptable [24]. On the other hand, the filters operating over finite data horizons have better robustness. For example, an UFIR filter developed in Vazquez-Olguin et al.'s study [25] with consensus on measurements has demonstrated better robustness than KF for WSN.

According to Vazquez-Olguin et al. [24], a dUFIR filter has been developed, implementing the consensus on estimates and tested over a WSN for a rapidly maneuvering object. Better performance compared to the KF and H_∞ filter was also shown.

One of the most challenging issues in WSN is missing data. It can be addressed using KF and UFIR approaches augmented with prediction blocks [26]. For example, a consensus finite-horizon H_∞ filter was developed in the study by Shen et al. [27] under missing data.

It is worth noting that the real-life sophisticated applications, such as weather monitoring, often suffer from data loss and false packets as well as uncertain noise and disturbances. This issue is

Figure 1
Historical temperature measurements from a single node



illustrated in Figure 1 for temperature measurements taken from a weather station.

To cope with such problems, in this work, we develop and use the dUFIR filter under unknown statistics. We test this solution under minimum link requirements and evaluate the robustness and the ability to reconstruct data in scenarios involving missing or corrupted measurements.

While existing distributed Kalman-based approaches rely on explicit stochastic covariance modeling and often assume reliable communication among nodes, their performance may degrade significantly in the presence of missing or persistently faulty measurements. Robust filtering strategies such as H_∞ methods address model uncertainties but still require parameter tuning and may increase computational complexity. Previous dUFIR implementations and consensus-based extensions primarily emphasize the algorithmic decentralization of the finite impulse response (FIR) structure, with limited analysis of performance under real-world data imperfections and constrained connectivity conditions. In contrast, the proposed dUFIR-based distributed framework explicitly examines the behavior of the estimate-consensus structure under realistic data imperfections and constrained connectivity conditions, using real environmental measurements. This distinguishes the present contribution from prior distributed estimation studies focused primarily on idealized or statistically well-defined environments.

3. Formulation of the dUFIR Filter for Estimation with Missing Data

Let Q denote a variable of interest observed by a set of spatially distributed and interconnected sensing nodes. The discrete-time state-space representation of the system dynamics is given in Equation (1), where the n-dimensional state vector x_k evolves according to transition matrix F_k and is corrupted by process disturbances w_k.

$$x_k = F_k x_{k-1} + B_k w_k, \tag{1}$$

$$\bar{y}_k^{(i)} = H_k^{(i)}(F_k x_{k-1}), \tag{2}$$

$$y_k^{(i)} = \gamma_k \left(H_k^{(i)} x_k + v_k^{(i)} \right) + (1-\gamma_k)\, \bar{y}_k^{(i)}, \tag{3}$$

$$y_k = H_k x_k + v_k, \tag{4}$$

Each node performs local measurements of the system state through an observation model defined by Equations (2) and (3). In practical WSN deployments, measurements may be intermittently unavailable due to communication failures, sensor faults, or power limitations. To capture this behavior, a binary indicator γ_k is introduced, which distinguishes between available ($\gamma_k = 1$) and missing ($\gamma_k = 0$) measurements. If a measurement is unavailable, a predicted observation based on a previous estimation is used instead.

The sensing nodes are modeled as vertices of an undirected graph, where communication links define neighborhood relationships. Each node acquires a local measurement vector $y_k^{(i)}$, while the aggregated local observation vector y_k is formed by stacking the measurements received from the one-hop neighboring nodes $j \in \mathcal{N}_i$, where $\mathcal{N}_i$ is the set of one-hop neighbors of node i and $J_i \triangleq |\mathcal{N}_i|$ is the number of such neighbors. Measurement noise and process disturbances are assumed to be zero mean, not necessarily Gaussian, and mutually uncorrelated, which reflects the uncertainty typically encountered in low-cost sensing platforms.

This modeling framework allows the estimation problem to be formulated in a distributed manner while explicitly accounting for missing data and uncertain noise characteristics. The resulting structure serves as the basis for the development of a robust distributed estimation scheme that does not rely on precise statistical assumptions and is suitable for real-world WSN applications.

3.1. Predictive dUFIR filter for lost data

In the dUFIR framework, a consensus mechanism is introduced to mitigate discrepancies among local estimates produced by neighboring nodes. This mechanism is governed by the optimal consensus weighting matrix $\lambda_k^{\text{opt}} \in \mathbb{R}^{n\times n}$, which regulates the influence of neighboring information on the final distributed estimate as follows:

$$\hat{x}_k^c = \hat{x}_k + \lambda_k^{\text{opt}}\Sigma_k, \tag{5}$$

where $\hat{x}_k^c$ is the consensus estimate at instant k. The centralized estimate $\hat{x}_k$ and the individual estimates $\hat{x}_k^{(i)}$ are obtained as

$$\hat{x}_k = K_{m,k} Y_{m,k}, \tag{6}$$

$$\hat{x}_k^{(i)} = K_{m,k}^{(i)} Y_{m,k}^{(i)}. \tag{7}$$

Using these estimates, a consensus protocol $\Sigma_k = \sum_{j\in\mathcal{N}_i} (\hat{x}_k^{(j)} - \hat{x}_k^{(i)})$ can be used to minimize the disagreement between the first-order neighbors [18]. The consensus weighting matrix λ_k^{opt} is chosen by minimizing the trace of the estimation error covariance $P(\lambda_k) = E\{\epsilon\epsilon^T\}$ where $\epsilon = x - \hat{x}^c$.

$$\lambda_k^{\text{opt}} = \underset{\lambda_k}{\operatorname{argmin}}\,\{\text{tr}\, P(\lambda_k)\}. \tag{8}$$

3.2. Design of the batch dUFIR filter

The main goal of the batch dUFIR filter design is to determine its gain $K_{m,k}$ and the individual gains $K_{m,k}^{(i)}$. To this end, we express the model Equations (1)–(4) in the extended state-space form over horizon N as described in Shmaliy and Zhao's study [23],

$$X_{m,k} = A_{m,k} x_m + D_{m,k} W_{m,k}, \tag{9}$$

$$Y_{m,k} = C_{m,k} x_m + M_{m,k} W_{m,k} + V_{m,k}, \tag{10}$$

$$Y_{m,k}^{(i)} = C_{m,k}^{(i)} x_m + M_{m,k}^{(i)} W_{m,k} + V_{m,k}^{(i)}, \tag{11}$$

where $X_{m,k} = \begin{bmatrix} x_m^T\; x_{m+1}^T \dots x_k^T \end{bmatrix}^T$, $Y_{m,k} = \begin{bmatrix} y_m^T\; y_{m+1}^T \dots y_k^T \end{bmatrix}^T$, $W_{m,k} = \begin{bmatrix} w_m^T\; w_{m+1}^T \dots w_k^T \end{bmatrix}^T$, $V_{m,k} = \begin{bmatrix} v_m^T\; v_{m+1}^T \dots v_k^T \end{bmatrix}^T$, $Y_{m,k}^{(i)} = \begin{bmatrix} y_m^{(i)T}\; y_{m+1}^{(i)T} \dots y_k^{(i)T} \end{bmatrix}^T$, $V_{m,k}^{(i)} = \begin{bmatrix} v_m^{(i)T}\; v_{m+1}^{(i)T} \dots v_k^{(i)T} \end{bmatrix}^T$, and the extended matrices are defined as follows:

$$A_{m,k} = [I\; F_{m+1}^T \dots (\mathcal{F}_k^{m+1})^T]^T, \tag{12}$$

$$D_{m,k} = \begin{bmatrix} B_m & 0 & \dots & 0 & 0 \\ F_{m+1}B_m & B_{m+1} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathcal{F}_{k-1}^{m+1}B_m & \mathcal{F}_{k-1}^{m+2}B_{m+1} & \dots & B_{k-1} & 0 \\ \mathcal{F}_k^{m+1}B_m & \mathcal{F}_k^{m+2}B_{m+1} & \dots & F_k B_{k-1} & Bk \end{bmatrix}, \tag{13}$$

with $C_{m,k} = \overline{C}_{m,k} A_{m,k}$, $M_{m,k} = \overline{C}_{m,k} D_{m,k}$, $C_{m,k}^{(i)} = \overline{C}_{m,k}^{(i)} A_{m,k}$, $M_{m,k}^{(i)} = \overline{C}_{m,k}^{(i)} D_{m,k}$, where

$$\overline{C}_{m,k} = \text{diag}\,(H_m\; H_{m+1} \dots\; H_k), \tag{14}$$

$$\overline{C}_{m,k}^{(i)} = \text{diag}\left(H_m^{(i)}\; H_{m+1}^{(i)} \dots\; H_k^{(i)}\right), \tag{15}$$

$$\mathcal{F}_l^r = \begin{cases} F_l F_{l-1} \dots F_r, & r < l+1 \\ I, & r = l+1\,. \\ 0, & r > l+1 \end{cases} \tag{16}$$

Equation (5) can now be formulated as

$$\hat{x}_k^c = K_{m,k} Y_{m,k} + J_i \lambda_k^{opt} K_{m,k} Y_{m,k} - J_i \lambda_k^{opt} K_{m,k}^{(i)} Y_{m,k}^{(i)}. \tag{17}$$

Because the robust UFIR filter does not rely on initial conditions, the unbiasedness requirement must be explicitly incorporated into the distributed, centralized, and local estimation processes, as expressed by the following relations:

$$E\{\hat{x}_k^c\} = E\{\hat{x}_k\} = E\{\hat{x}_k^{(i)}\} = E\{x_k\}, \tag{18}$$

where the state model is given by

$$x_k = \mathcal{F}_k^{m+1} x_m + \overline{D}_{m,k} W_{m,k}, \tag{19}$$

with $\overline{D}_{m,k} = [\mathcal{F}_k^{m+1} B_m\; \mathcal{F}_k^{m+2} B_{m+1} \dots\; F_k B_{k-1}\; Bk]$. The corresponding matrices can now be computed by

$$K_{m,k} = G_k C_{m,k}^T, \tag{20}$$

$$K_{m,k}^{(i)} = G_k^{(i)} C_{m,k}^{(i)T} \tag{21}$$

where $G_k = (C_{m,k}^T C_{m,k})^{-1}$ and $G_k^{(i)} = (C_{m,k}^{(i)T} C_{m,k}^{(i)})^{-1}$.

3.3. Optimum consensus matrix λ_k^{opt}

The optimal consensus factor λ_k^{opt} can be found by minimizing the trace of the error covariance $P_k = E\{\varepsilon_k \varepsilon_k^T\}$ as follows:

$$\frac{\partial}{\partial \lambda_k} trP_k = 0, \tag{22}$$

where $\varepsilon_k = x_k - \hat{x}_k^c$ is the estimation error. This gives

$$\begin{aligned}\lambda_k^{opt} = &-\tfrac{1}{J_i}(K_{m,k}\overline{R}_{m,k}K_{m,k}^T - G_k G_k^{(i)^{-1}} K_{m,k}^{(i)} \\ &\times \overline{R}_{m,k}^{(i)} K_{m,k}^{(i)^T})(K_{m,k}\overline{R}_{m,k}K_{m,k}^T - 2G_k G_k^{(i)^{-1}} \\ &\times K_{m,k}^{(i)} \overline{R}_{m,k}^{(i)} K_{m,k}^{(i)^T} + K_{m,k}^{(i)} \overline{R}_{m,k}^{(i)} K_{m,k}^{(i)^T})^{-1}\end{aligned} \tag{23}$$

where the matrices are defined by

$$\overline{R}_{m,k} = E\left\{v_{m,k} v_{m,k}^T\right\} = diag\,(R_m \dots R_k),$$

$$\overline{R}_{m,k}^{(i)} = E\left\{v_{m,k}^{(i)} v_{m,k}^{(i)^T}\right\} = diag\left(R_m^{(i)} \dots R_k^{(i)}\right),$$

$$\tilde{R}_{m,k}^{(i)} = E\left\{v_{m,k} v_{m,k}^{(i)^T}\right\} = diag\left(\tilde{R}_m^{(i)} \dots \tilde{R}_k^{(i)}\right).$$

Although the resulting expression for λ_k^{opt} involves covariance-related terms associated with both centralized and individual estimates, it can be computed offline when time-invariant models are considered and subsequently embedded into the sensor nodes. This avoids additional computational burden during online operation and is consistent with the resource constraints typical of WSNs.

In practical implementations, directly applying Equation (17) at the sensor nodes may be impractical due to the high dimensionality of the involved matrices and the limited memory resources available in WSNs. For this reason, more efficient recursive formulations of Equation (17) are required. The corresponding iterative algorithm is developed next.

3.4. Iterative dUFIR filter using recursions

Recursive forms for the batch dUFIR filter were derived in Shmaliy and Zhao's study [23]. Based on this result, the iterative centralized estimator algorithm can be stated as follows. Consider a n-dimensional state-space model over the horizon $[m, k]$, where $m = k - N + 1$. Introduce index variable l, starting at $l = k - N + n + 1$, and ending up at $l = k$. Use the following recursions:

$$G_l = [H_l^T H_l + (F_l G_{l-1} F_l^T)^{-1}]^{-1}, \tag{24}$$

$$\hat{x}_l = F_l \hat{x}_{l-1}, \tag{25}$$

$$\hat{x}_l = \hat{x}_l^- + G_l H_l^T \left(y_l - H_l \hat{x}_l^-\right). \tag{26}$$

Compute the starting values G_{l-1} and $\hat{x}_{l-1}$ at $s = k - N + n$ by

$$G_s = (C_{m,s}^T C_{m,s})^{-1}, \tag{27}$$

$$\hat{x}_s^c = G_s C_{m,s}^T Y_{m,s}. \tag{28}$$

For the individual estimate $\hat{x}_k^{(i)}$, accordingly write

$$G_l^{(i)} = [H_l^{(i)^T} H_l^{(i)} + (A_l G_{l-1}^{(i)} A_l^T)^{-1}]^{-1}, \tag{29}$$

$$\hat{x}_l^{(i)^-} = A_l \hat{x}_{l-1}^{(i)} \tag{30}$$

$$\hat{x}_l^{(i)} = \hat{x}_l^{(i)^-} + G_l^{(i)} H_l^{(i)^T} \left(y_l^{(i)} - H_l^{(i)} \hat{x}_l^{(i)^-}\right), \tag{31}$$

computing the initial values by

$$G_s^{(i)} = (C_{m,s}^{(i)^T} C_{m,s}^{(i)})^{-1}, \tag{32}$$

$$\hat{x}_s^{(i)} = G_s^{(i)} C_{m,s}^{(i)^T} Y_{m,s}^{(i)}. \tag{33}$$

By integrating Equations (24)–(33), we can summarize the resulting steps of the iterative dUFIR algorithm with estimate consensus in Algorithm 1.

The above formulation emphasizes the central contribution of this work. The conventional finite-horizon UFIR estimator is extended through the incorporation of a matrix-weighted consensus correction term, resulting in a fully distributed update mechanism. In contrast to scalar consensus gains commonly adopted in distributed filtering schemes, the proposed weighting matrix is obtained by minimizing the trace of the distributed estimation error covariance. This matrix-valued optimization enables state-dependent adjustment of the consensus contribution while preserving the unbiased finite-horizon structure of the original UFIR estimator. Consequently, the distributed formulation retains the robustness properties of FIR estimation while allowing cooperative information fusion over locally connected networks.

Algorithm 1: Iterative Implementation of the Distributed UFIR Filter

Data: y_k, $R_k^{(i)}$, R_k, λ_k^{opt}
Result: $\hat{x}_k$
begin
For $k = N-1, N, N+1, \dots$ **do**

$$m = k - N + 1, \; s = m + n - 1;$$

$$G_s = (C_{m,s}^T C_{m,s})^{-1};$$

$$G_s^{(i)} = \left(\left(C_{m,s}^{(i)}\right)^T C_{m,s}^{(i)}\right)^{-1};$$

If $\gamma_k = 0$ **then**

$$y_k^{(j)} = H_k^{(j)} F_k \hat{x}_{k-1}^{(j)}, \; \forall j \in \mathcal{N}_i;$$

End if

$$\tilde{x}_s = G_s C_{m,s}^T Y_{m,s}.$$

$$\tilde{x}_s^{(i)} = G_s^{(i)} C_{m,s}^{(i)T} Y_{m,s}^{(i)}.$$

For $l = s+1, s+2, \ldots, k$ **do**

$$\hat{x}_l^- = F_l \hat{x}_{l-1};$$

$$\hat{x}_l^{(i)-} = F_l \hat{x}_{l-1}^{(i)};$$

$$G_l = \left[H_l^T H_l + \left(F_l G_{l-1} F_l^T\right)^{-1}\right]^{-1};$$

$$G_l^{(i)} = \left[\left(H_l^{(i)}\right)^T H_l^{(i)} + \left(F_l G_{l-1}^{(i)} F_l^T\right)^{-1}\right]^{-1};$$

$$\hat{x}_l = \hat{x}_l^- + G_l H_l^T \left(y_l - H_l \hat{x}_l^-\right).$$

$$\hat{x}_l^{(i)} = \hat{x}_l^{(i)-} + G_l^{(i)} \left(H_l^{(i)}\right)^T \left(y_l^{(i)} - H_l^{(i)} \hat{x}_l^{(i)-}\right);$$

End for

$$\hat{x}_k^{c} = \left(I + J_i \lambda_k^{opt}\right) \hat{x}_k - J_i \lambda_k^{opt} \hat{x}_k^{(i)}.$$

End for

End

Note: First data $y_0,\ y_1, \ldots, y_{N-1}$ must be available.

4. dUFIR Algorithm for Real Environment Monitoring

This section evaluates the proposed approach using a real-world dataset of ambient temperature measurements collected at the Grand St. Bernard pass (2400 m altitude), located at the border between Switzerland and Italy. The dataset was acquired during a field campaign conducted in 2007 within the framework of the Sensorscope project [28, 29], which aimed at developing large-scale, distributed environmental monitoring systems based on WSNs. The measurements, gathered independently by low-cost sensing nodes, are publicly available and were used here to assess the performance of the proposed filter. The present work does not address the physical deployment or hardware configuration of the sensing infrastructure. Instead, the proposed distributed estimation framework is evaluated offline using recorded real-world measurements. To test the filter, we follow the scenarios shown in Figures 2 and 3, where the node locations are indicated as red markers. For the purposes of analysis, the network connectivity considered in this paper is not physically implemented but modeled based on geographical distances between stations. Also, the raw measurements were aggregated into one-hour averages, and the associated error variances were computed for each sensor.

Although the temperature profiles reported by the 11 sensing nodes exhibit comparable overall trends, several stations present unstable behavior characterized by extended intervals of missing data. In addition, certain sensors produce erroneous readings, such as constant values near $-1°C$, which cannot be interpreted as data losses and therefore pose additional challenges for the estimation process. Figure 4 illustrates the one-hour averaged temperature measurements corresponding to selected stations, showing both missing-data segments and abnormal readings. This behavior exhibits the different impact of missing and faulty measurements on the estimation process. While missing data activate the prediction mechanism of the dUFIR filter, persistent faulty values are treated as valid observations.

Figure 2
WSN connectivity graph (modeled based on geographical distance) for a communication range of 350 m

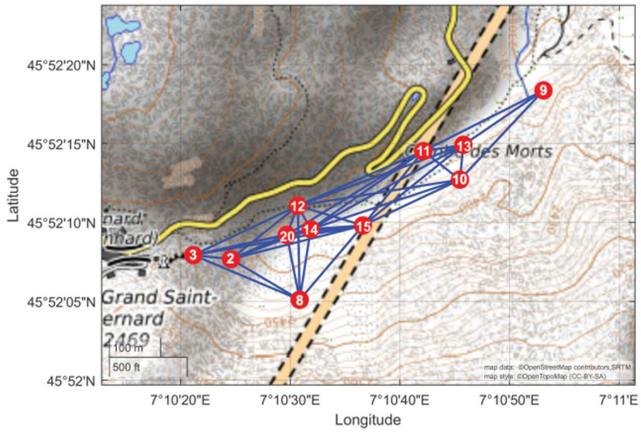


Figure 3
WSN connectivity graph (modeled based on geographical distance) for a communication range of 200 m

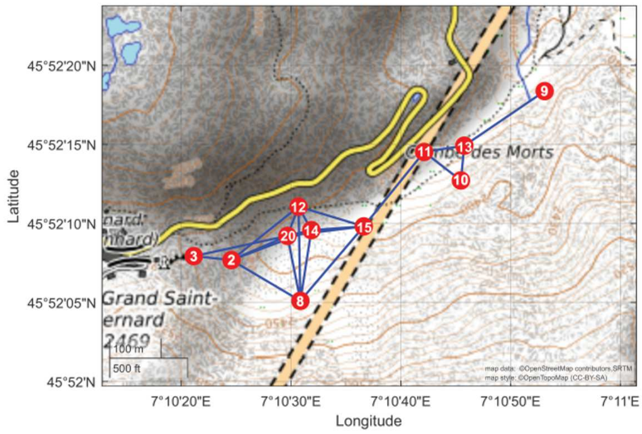


Figure 4
One-hour averaged temperature measurements illustrating missing-data segments and persistent faulty readings

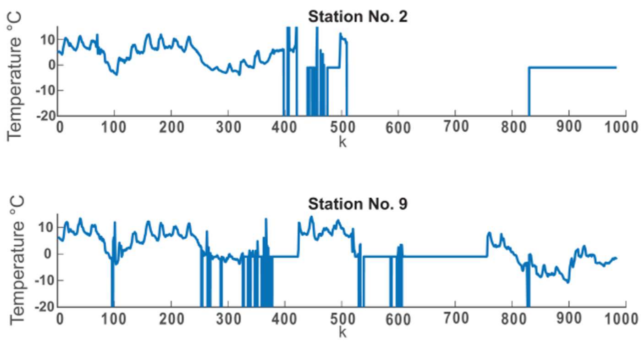


Table 1
Sensor-specific variance estimates used in the experiments

Node	2	3	8	9	10	11	12	13	14	15	20
Variance	0.15	0.16	0.11	0.19	0.13	0.17	0.18	0.20	0.13	0.17	0.15

To apply the proposed dUFIR filtering scheme, the system dynamics described in Equations (1)–(4) are considered with the following matrices:

$$A = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix}, \quad H^{(i)} = [1 \;\; 0],$$

where $\tau = 1$ and B is the identity. The original measurements exhibit irregular time stamps at the minute and second levels. To obtain a consistent characterization of sensor variability, measurements were first averaged over hourly intervals. For each sensor, the empirical variance was computed at the hourly level and subsequently averaged over the full observation period to obtain a representative variance value. These values, reported in Table 1, are used in the offline computation of the optimal consensus matrix λ_k^{opt} according to Equation (23). The matrix λ_k^{opt} is determined under time-invariant structural assumptions and remains fixed during online operation. It is important to emphasize that, although sensor dispersion is considered in the consensus weighting computation, the recursive UFIR estimation structure itself does not rely on prior stochastic noise covariance modeling.

Since ground-truth temperature measurements are not available, direct estimation error evaluation is not feasible. Therefore, the optimal horizon length N_{opt} was selected following the mean square value (MSV)-based optimal memory estimation principle proposed in Shmaliy and Zhao's work [23]. The trace of the MSV was used as an internal consistency metric, and the MSV curves obtained for each node were averaged across the network to produce a representative aggregated behavior. The selected horizon $N_{\text{opt}} = 37$ corresponds to the minimum of this aggregated MSV curve.

4.1. Network configuration with a 350 m communication range

Now, we test the scenario shown in Figure 2. The results related to the 9th and 2nd stations are shown in Figure 5. As we can see, the filter provides efficient noise reduction and robustness in the presence of long data gaps.

A relevant distinction between the behavior of the 9th and 2nd nodes can be observed within the interval $540 < k < 780$. During this period, the measurements from the 2nd station are entirely unavailable, whereas the 9th station reports spurious readings fixed at -1°C. It is important to emphasize that the proposed algorithm activates its prediction mechanism exclusively when data losses are detected, while erroneous measurements such as -1°C are treated as valid observations. Despite this limitation, the distributed nature of the dUFIR filter prevents the estimate of the 9th station from diverging significantly from those of neighboring nodes. This behavior is illustrated in Figure 6(a), where the state estimates of all stations are shown. Furthermore, Figure 6(b) reports the corresponding estimation variances, which serve as a quantitative measure of disagreement among nodes. As expected, substantially lower disagreement levels are observed when the measurements are reliable.

Figure 5
Temperature measurements and corresponding dUFIR estimates under 350 m connectivity for (a) station 9 and (b) station 2

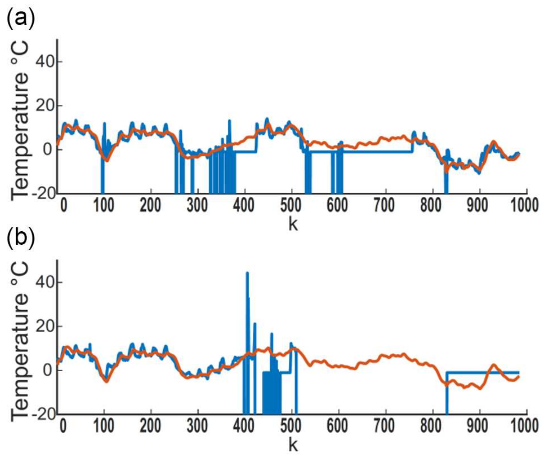
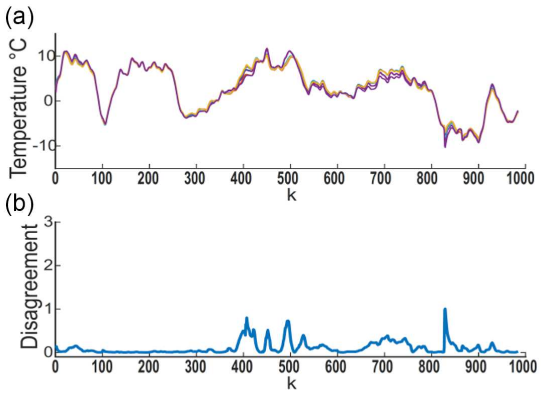


Figure 6
(a) Distributed state estimates under high connectivity (range = 350 m). (b) Inter-node disagreement metric illustrating divergence due to persistent faulty measurements

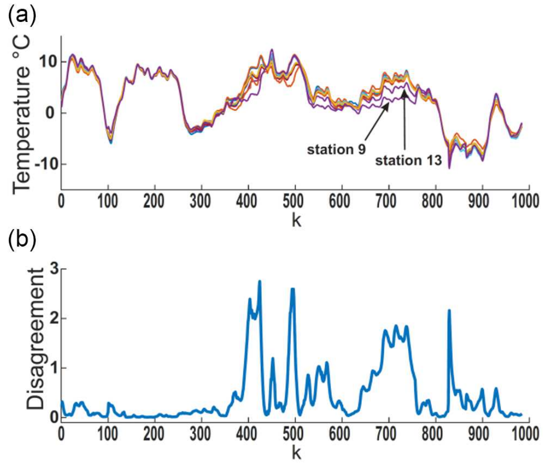


4.2. Network configuration with a 200 m communication range

The performance of the dUFIR filter is influenced by the amount of information available from neighboring nodes. In particular, reducing the transmission power will reduce the number of communication links, increasing the disagreement in the estimation errors. To examine this effect, we analyze the scenario corresponding to a 200 m communication range, illustrated in Figure 3. Under this configuration, the network connectivity is significantly reduced, and the 9th station is connected through only a single link, in contrast to the three links available in the previous case shown in Figure 2.

Due to the reduced information exchange among nodes and the lack of an explicit mechanism to identify and discard persistent faulty measurements, the estimates produced by the 9th

Figure 7
(a) Distributed state estimates under reduced connectivity (range = 200 m). (b) Inter-node disagreement metric illustrating divergence due to persistent faulty measurements



station progressively deviate from those of the remaining nodes, as illustrated in Figure 7(a). This deviation propagates through the network via the consensus process, affecting the estimation performance of neighboring nodes, such as the 13th station. As a result, the combined consensus and prediction mechanisms described in Algorithm 1 are unable to fully mitigate the impact of persistent faulty data under low-connectivity conditions, leading to an increasing level of disagreement among node estimates, as quantified in Figure 7(b).

All simulations were carried out in MATLAB®R2019a on a PC equipped with an AMD Ryzen 7 processor (2.3 GHz) and 16.0 GB of RAM. The total execution time for the full network simulation was 34 s.

It is worth noting that the parameters analyzed in this study are determined by physical and network constraints of the WSN. The communication range and network connectivity were explicitly evaluated, while the horizon length was selected using a data-driven optimization criterion. Other parameter variations outside these physically meaningful ranges are therefore beyond the scope of the present analysis.

5. Conclusions

The dUFIR filtering algorithm developed in this paper and applied to ambient temperature monitoring has demonstrated reliable estimation performance in the presence of missing and incorrect data. The results indicate that, even under low-connectivity conditions with as few as three links, the dUFIR filter provides satisfactory estimation accuracy and effective data reconstruction capabilities. Beyond its empirical validation, this study contributes to the understanding of distributed FIR estimation by providing a systematic evaluation of the dUFIR framework under realistic data imperfections and constrained connectivity conditions.

A key advantage of the proposed method is its independence from process noise statistics, which makes it particularly attractive in scenarios where such information is unavailable or difficult to model. Compared to conventional distributed Kalman-based estimators, which rely on accurate stochastic descriptions of noise processes, the dUFIR filter exhibits enhanced robustness under uncertain statistical conditions. Its finite-horizon formulation enables stable estimation performance under intermittent measurements and limited connectivity.

Despite its robustness, some limitations should be acknowledged. The algorithm does not include an explicit mechanism to detect persistent faulty measurements that are not classified as missing data, and the consensus factor is computed offline under time-invariant assumptions, which may limit adaptability in rapidly changing network conditions. These aspects define the scope of applicability of the proposed method and motivate future research on handling delayed observations and colored noise in distributed estimation frameworks.

Acknowledgment

This work was supported by the Mexican SEP-CONACyT Project under Grant A1-S-10287.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available from Guillermo Barrenetxea. (2019). Sensorscope Data [Data set]. Zenodo. https://doi.org/10.5281/zenodo.2654726 [29]

Author Contribution Statement

Miguel A. Vazquez-Olguin: Conceptualization, Software, Formal analysis, Writing – original draft. **Oscar G. Ibarra-Manzano:** Methodology, Validation, Investigation. **Yuriy S. Shmaliy:** Conceptualization, Methodology, Formal analysis, Supervision, Project administration.

References

[1] Sharma, P., & Prakash, S. (2021). Real time weather monitoring system using IoT. *ITM Web of Conferences*, *40*(3), 01006. https://doi.org/10.1051/itmconf/20214001006

[2] Mohapatra, D., & Subudhi, B. (2022). Development of a cost-effective IoT-based weather monitoring system. *IEEE Consumer Electronics Magazine*, *11*(5), 81–86. https://doi.org/10.1109/MCE.2021.3136833

[3] Wellyantama, P., & Soekirno, S. (2021). Temperature, pressure, relative humidity and rainfall sensors early error detection system for automatic weather station (AWS) with artificial neural network (ANN) backpropagation. *Journal of Physics: Conference Series*, *1816*(1), 012056. https://doi.org/10.1088/1742-6596/1816/1/012056

[4] Trigka, M., & Dritsas, E. (2025). Wireless sensor networks: From fundamentals and applications to innovations and future trends. *IEEE Access*. https://doi.org/10.1109/access.2025.3572328

[5] Vishwas, H. N., & Ramesh, T. K. (2025). Recent trends in localization, routing, and security for wireless sensor networks. *IEEE Access*. https://doi.org/10.1109/ACCESS.2025.3555280

[6] Chen, C., Zhu, S., Guan, X., & Shen, X. S. (2014). *Wireless sensor networks: Distributed consensus estimation*. Springer.

[7] Soni, M. S., Jisan, M., Kaustav, B., & Ghosh, R. (2025). Advancements in weather monitoring system: A comprehensive review. In *2025 8th International Conference on Electronics, Materials Engineering & Nano-Technology (IEMENTech)*, 1–5. https://doi.org/10.1109/IEMENTech65115.2025.10959560

[8] Nurlan, Z., Zhukabayeva, T., Othman, M., Adamova, A., & Zhakiyev, N. (2021). Wireless sensor network as a mesh: Vision and challenges. *IEEE Access*, *10*, 46–67. https://doi.org/10.1109/access.2021.3137341

[9] Wang, H., Yemeni, Z., Ismael, W. M., Hawbani, A., & Alsamhi, S. H. (2021). A reliable and energy efficient dual prediction data reduction approach for WSNs based on Kalman filter. *IET Communications*, *15*(18), 2285–2299. https://doi.org/10.1049/cmu2.12262

[10] Guyeux, C., Haddad, M., Hakem, M., & Lagacherie, M. (2020). Efficient distributed average consensus in wireless sensor networks. *Computer Communications*, *150*, 115–121. https://doi.org/10.1016/j.comcom.2019.11.006

[11] Banara, S., Singh, T., & Chauhan, A. (2022). IoT based weather monitoring system for smart cities: A comprehensive review. In *2022 International Conference for Advancement in Technology (ICONAT)*, 1–6. https://doi.org/10.1109/iconat53423.2022.9726106

[12] Jia, X. C. (2021). Resource-efficient and secure distributed state estimation over wireless sensor networks: A survey. *International Journal of Systems Science*, *52*(16), 3368–3389. https://doi.org/10.1080/00207721.2021.1998843

[13] Pundir, M., & Sandhu, J. K. (2021). A systematic review of quality of service in wireless sensor networks using machine learning: Recent trend and future vision. *Journal of Network and Computer Applications*, *188*, 103084. https://doi.org/10.1016/j.jnca.2021.103084

[14] Farag, W. A. (2021). Kalman-filter-based sensor fusion applied to road-objects detection and tracking for autonomous vehicles. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, *235*(7), 1125–1138.

[15] Montañez, O. J., Suarez, M. J., & Fernandez, E. A. (2023). Application of data sensor fusion using extended Kalman filter algorithm for identification and tracking of moving targets from LiDAR–radar data. *Remote Sensing*, *15*(13), 3396. https://doi.org/10.3390/rs15133396

[16] Park, G. (2024). Optimal vehicle position estimation using adaptive unscented Kalman filter based on sensor fusion. *Mechatronics*, *99*, 103144. https://doi.org/10.1016/j.mechatronics.2024.103144

[17] Narayanasami, S., Butta, R., Govindaraj, R., Choudhary, S. S., Sharma, D. K., Poonia, A., . . . , & Verma, R. (2022). An enhanced trust-based Kalman filter route optimization technique for wireless sensor networks. *Wireless Personal Communications*, *127*(2), 1311–1329. https://doi.org/10.1007/s11277-021-08578-x

[18] Olfati-Saber, R. (2007). Distributed Kalman filtering for sensor networks. In *2007 46th IEEE conference on decision and control*, 5492–5498. https://doi.org/10.1109/CDC.2007.4434303

[19] Carli, R., Chiuso, A., Schenato, L., & Zampieri, S. (2008). Distributed Kalman filtering based on consensus strategies. *IEEE Journal on Selected Areas in Communications*, *26*(4), 622–633. https://doi.org/10.1109/JSAC.2008.080505

[20] Yang, G., Rezaee, H., Alessandri, A., & Parisini, T. (2023). State estimation using a network of distributed observers with switching communication topology. *Automatica*, *147*, 110690. https://doi.org/10.1016/j.automatica.2022.110690

[21] Yu, D., Xia, Y., Zhai, D. H., & Zhan, Y. (2024). Distributed simultaneous state-input estimation over sensor networks under quantized communication. *Automatica*, *163*, 111552. https://doi.org/10.1016/j.automatica.2024.111552

[22] Chen, W., Wang, Z., Ding, D., Yi, X., & Han, Q. L. (2022). Distributed state estimation over wireless sensor networks with energy harvesting sensors. *IEEE Transactions on Cybernetics*, *53*(5), 3311–3324. https://doi.org/10.1109/TCYB.2022.3179280

[23] Shmaliy, Y. S., & Zhao, S. (2022). *Optimal and robust state estimation: Finite Impulse Response (FIR) and Kalman approaches*. John Wiley & Sons.

[24] Vazquez-Olguin, M., Shmaliy, Y. S., & Ibarra-Manzano, O. G. (2019). Distributed UFIR filtering over WSNs with consensus on estimates. *IEEE Transactions on Industrial Informatics*, *16*(3), 1645–1654. https://doi.org/10.1109/TII.2019.2930649

[25] Vazquez-Olguin, M., Shmaliy, Y. S., & Ibarra-Manzano, O. G. (2017). Distributed unbiased FIR filtering with average consensus on measurements for WSNs. *IEEE Transactions on Industrial Informatics*, *13*(3), 1440–1447. https://doi.org/10.1109/TII.2017.2653814

[26] Nakamori, S. (2020). Robust recursive least-squares finite impulse response predictor in linear discrete-time stochastic systems with uncertain parameters. *WSEAS Transactions on Systems*, *19*, 86–101. https://doi.org/10.37394/23202.2020.19.13

[27] Shen, B., Wang, Z., & Hung, Y. S. (2010). Distributed H∞-consensus filtering in sensor networks with multiple missing measurements: the finite-horizon case. *Automatica*, *46*(10), 1682–1688. https://doi.org/10.1016/j.automatica.2010.06.025

[28] Barrenetxea, Guillermo. (2019). Sensorscope Data [Data set]. *Zenodo*. https://doi.org/10.5281/zenodo.2654726

[29] Barrenetxea, G., Ingelrest, F., Schaefer, G., Vetterli, M., Couach, O., & Parlange, M. (2008). Sensorscope: Out-of-the-box environmental monitoring. In *2008 International Conference on Information Processing in Sensor Networks*, 332–343. https://doi.org/10.1109/IPSN.2008.28

How to Cite: Vazquez-Olguin, M. A., Ibarra-Manzano, O. G., & Shmaliy, Y. S. (2026). Distributed Estimation of Ambient Temperature with Missing Data. *Archives of Advanced Engineering Science.* https://doi.org/10.47852/bonviewAAES62028681