International Journal of Applied Mathematics and Numerical Research

Cyberattack-Resilient Data Assimilation for Smart IoT Infrastructures: Sparse Attack Reconstruction and Certified State Estimation

Kaarina Nakale
International University of Management, Namibia

* Corresponding Author: **Kaarina Nakale**

Article Info

Impact Factor (RSIF): 8.62
ISSN (Online): 3107-7110
Volume: 02
Issue: 04
Received: 11-05-2026
Accepted: 13-06-2026
Published: 15-07-2026
Page No: 62-72

Abstract
Smart Internet-of-Things (IoT) infrastructures infer physical states from heterogeneous, delayed and noisy sensor streams. Their dependence on networked measurements creates a structural vulnerability: a false data injection adversary can alter a small subset of observations while preserving the superficial statistical consistency required by conventional residual tests. This article develops a cyberattack-resilient graph-sparse data-assimilation framework, abbreviated CR-GSDA, that separates genuine physical variation from malicious measurement corruption. The method combines a robust Kalman forecast, graph-frequency screening, sparse attack reconstruction, graph-regularised state analysis, ensemble-based nonlinear propagation, Bayesian covariance inflation and distributed edge consensus. For the observation equation $y_t = H_t x_t + v_t + a_t$, the attack is recovered through a weighted proximal programme and removed before the state update. Across 20 reported synthetic trials with 10% attacked sensors, CR-GSDA records a state RMSE of 0.1916 ± 0.0021 and an attack-localisation F1 score of 0.9204 ± 0.0062. These findings are limited to the stated synthetic setting.

Keywords: cyberattack resilience, data assimilation, false data injection, Internet of Things, robust Kalman filtering, sparse recovery, graph signal processing, Bayesian uncertainty, edge computing, distributed sensor fusion
MSC 2020: 62M20, 65C60, 68M14, 68M25, 93E11

Introduction

Smart infrastructures increasingly couple physical processes to dense layers of sensors, wireless links, edge processors and automated actuators. State estimates derived from these components influence traffic routing, voltage regulation, irrigation, industrial control and emergency response. Dynamic state estimation therefore functions as a safety mechanism as well as a numerical inference procedure. A recent review of cyber-physical power systems classifies resilient estimation into prevention, detection and mitigation and identifies the joint treatment of unknown, asynchronous and coordinated attacks as an unresolved requirement [1]. The same requirement extends beyond electrical networks because the underlying ICT pattern—distributed sensing, network transport and feedback control—is shared by most IoT infrastructures.

The most consequential integrity threat is false data injection (FDI). An adversary changes selected measurements so that a trusted estimator produces a biased state. The classical construction of stealth attacks in power-grid estimation showed that knowledge of the measurement map can be used to evade residual-based bad-data detectors [2]. Secure estimation studies subsequently established that the number and placement of compromised sensors, system observability and temporal redundancy jointly determine whether the true state remains recoverable [3,4]. These results provide essential identifiability conditions, but an operational IoT estimator must additionally handle non-Gaussian noise, uncertain models, changing communication graphs, limited edge resources and a need to quantify confidence after attack removal.

Kalman filtering remains a central tool for online state estimation because it combines a dynamical forecast with new observations through a covariance-weighted correction [5]. Ensemble Kalman methods extend the principle to nonlinear and

high-dimensional settings by propagating samples rather than a full covariance matrix [6]. Recent non-asymptotic work has strengthened the mathematical basis of ensemble filters through resampling and effective-dimension analysis [7]. Nevertheless, an unmodified Kalman update interprets a malicious innovation as physical information. Robust loss functions limit the influence of extreme residuals, but a stealthy or spatially coordinated attack can remain within a robust threshold. Down-weighting also discards information about which sensors were compromised and provides no reconstruction of the injected vector.

Sparse recovery offers a complementary mechanism because practical attackers are frequently constrained to a subset of sensors or communication channels. Sliding-window sparse recovery combined with an unknown-input observer has recently been used to identify control-channel disturbances in distribution networks [8]. Joint detection and localisation methods model falsification as non-Gaussian contamination [9], while distributed secure estimators address sparse sensor attacks over networked sensing architectures [10]. These contributions indicate a movement from binary detection towards reconstruction. However, sparse penalties alone can confuse a legitimate local event—a traffic incident, irrigation leak or voltage transient—with an attack when both appear as abrupt innovations.

The topology of the infrastructure supplies information needed to resolve this ambiguity. Graph signal processing represents sensor values as signals supported on a graph [11] and separates smooth, physically coherent variation from high-frequency inconsistency. Graph-based FDI detection has demonstrated that spectral residuals can expose attacks that are difficult to identify with conventional norms [12,13]. Yet graph detection is commonly applied as a pre-processing classifier rather than embedded in the state-analysis objective. Its threshold is also rarely propagated into the uncertainty assigned to the final state.

Bayesian inversion provides the language for that uncertainty. Modern analysis clarifies how discretisation, prior structure and ensemble approximations affect posterior accuracy [14]. Edge computing, in turn, makes low-latency local assimilation feasible but introduces resource, security and orchestration constraints [15,16]. A resilient design must therefore connect numerical estimation, sparse reconstruction, graph topology, probability and ICT deployment.

This article develops that connection. The proposed CR-GSDA method operates in five linked stages: forecast; graph-frequency screening; weighted sparse attack reconstruction; robust graph-regularised analysis; and uncertainty inflation with a computable certificate. A distributed alternating-direction implementation allows edge clusters to exchange sufficient statistics rather than raw observations. The framework is designed to distinguish genuine physical changes from FDI by asking whether an innovation is dynamically plausible, graph coherent, persistent across neighbouring nodes and statistically compatible with the forecast uncertainty.

The article contributes (i) a unified graph-sparse robust assimilation objective for linear and ensemble-based nonlinear systems; (ii) a certificate separating forecast, noise, attack-reconstruction and model-mismatch errors; (iii) sparse-support, stability, uncertainty-inflation and distributed-consensus guarantees; (iv) an edge deployment architecture with explicit computational and communication costs; and (v) a fully specified synthetic benchmark whose reported values can be reproduced without access to sensitive infrastructure data.

The work is methodological. It does not claim field validation or replace network-layer access control, cryptographic authentication or incident response. Instead, it adds an estimation-layer defence for situations in which apparently valid sensor packets have already entered the assimilation pipeline.

Related Work and Research Gap

Attack-resilient state estimation spans control, signal processing, optimisation and cybersecurity. Foundational secure estimators exploit temporal observability and sensor redundancy. If at most s of m sensors are corrupted, exact noiseless recovery typically requires a form of $2s$-sparse observability: deleting any $2s$ sensors must leave the system observable. In noisy systems, subset-based filters and convex relaxations replace exact consistency with bounded error. Attack-resilient H_2, H_∞ and ℓ_1 estimators formalise different trade-offs between average noise, worst-case disturbance and sparse corruption [17]. Coding-based approaches introduce additional redundancy so that manipulated measurements can be detected or corrected [18]. These techniques provide strong guarantees, but exhaustive subset search scales poorly and fixed coding assumptions can be restrictive in changing IoT deployments.

Robust filtering replaces the quadratic innovation loss by Huber, Student-t or mixture likelihoods. It performs well for impulsive noise and unknown outliers because large residuals receive bounded influence. The limitation is semantic: a robust filter treats an outlier as an untrusted value but does not explicitly infer an adversarial action. Sparse reconstruction instead introduces a latent a_t and penalises $\|a_t\|_1$. Under restricted-eigenvalue or mutual-coherence conditions, the support of a_t becomes identifiable. Recent distribution-network research confirms the value of combining sparse recovery with observer-based estimation. The remaining difficulty is that physical anomalies can also be sparse in the sensor domain.

Graph priors encode neighbourhood dependence. If $G = (V, E, W)$ is a sensor graph with Laplacian L, a smooth physical state has small graph variation $x^\top Lx = \frac{1}{2}\sum_{i,j} W_{ij}\left(x_i - x_j\right)^2$. Graph Fourier components associated with large Laplacian eigenvalues represent rapid changes across connected nodes. High graph frequency is informative for attack screening, but it is not a universal attack signature: genuine discontinuities, moving fronts and topology errors also create high-frequency energy. Accordingly, CR-GSDA uses graph information as a soft regulariser whose strength decreases when physical-event evidence is high; it never declares an attack from graph frequency alone.

Distributed sensor fusion offers matrix-weighted fusion, covariance intersection and consensus-based filtering. Consensus estimation under hybrid attacks and missing observations has recently received attention [19]. Privacy-aware wireless sensor estimation and secure distributed updates are also expanding [20]. Most distributed schemes assume that local innovations are either trusted or screened by a separate detector. This separation can create contradictory decisions across clusters and underestimate the correlation introduced by shared forecasts.

Ensemble assimilation is attractive when f and h are nonlinear, the state dimension is large, or derivatives are

unavailable. It approximates forecast covariances from particles, but finite ensembles underestimate uncertainty and create spurious long-range correlations. Localisation, inflation and resampling address these effects. Ensemble methods should be used cautiously as complete Bayesian approximations because their analysis variable is generally closer to a linear Bayes estimator than to the full posterior [21]. For cyber-resilient assimilation, a further source of approximation is attack removal. If $\hat{a}_t$ is uncertain, treating $y_t - \hat{a}_t$ as an ordinary clean observation produces overconfident credible intervals.

Table 1 locates the proposed method relative to the principal families. The gap is not the absence of any one component; it is the absence of a mathematically connected pipeline that reconstructs attacks, respects infrastructure topology, quantifies the residual attack uncertainty, certifies the state error and can be partitioned across edge nodes.

Table 1: Methodological gap addressed by CR-GSDA. A filled circle denotes an explicit component; an open circle denotes partial or indirect support.

Method	Robust	Attack	Graph	UQ	Bound	Edge
KF/EnKF	∘	–	–	•	∘	∘
Robust filter	•	–	–	∘	∘	∘
Sparse estimator	∘	•	–	–	•	∘
GSP detector	∘	∘	•	–	–	∘
Distributed filter	∘	∘	∘	∘	∘	•
CR-GSDA	•	•	•	•	•	•

Smart IoT System and Threat Model

Consider a smart infrastructure observed at discrete times $t = 1,2,\ldots$. Its state $x_t \in \mathbb{R}^n$ may contain traffic densities, queue lengths, voltages, soil moisture, radio-link loads, device temperatures or other latent physical variables. The process and observation models are

$$x_t = f_t(x_{t-1}, u_t) + w_t + d_t.$$

$$y_t = h_t(x_t) + v_t + a_t.$$

where u_t is a known input, w_t is process noise, d_t is model discrepancy, v_t is measurement noise and a_t is an integrity attack. In the linear case,

$$x_t = F_t x_{t-1} + B_t u_t + w_t + d_t,$$

$$y_t = H_t x_t + v_t + a_t.$$

Nominal noises satisfy

$$\mathbb{E}w_t = 0, \mathbb{E}v_t = 0, \mathrm{cov}(w_t) = Q_t$$

and

$$\mathrm{cov}(v_t) = R_t.$$

Gaussianity is useful for interpretation but is not required by the deterministic certificate derived later.

Sensors form a weighted graph

$$G_t = (V_t, E_t, W_t).$$

An edge $W_{ij,t} > 0$ indicates physical, statistical or communication proximity. The combinatorial Laplacian is $L_t = D_t - W_t$, where $D_{ii,t} = \sum_j W_{ij,t}$. Normalised Laplacians may be used when degrees are highly heterogeneous. The physical state is expected to be approximately smooth within stable regimes, but abrupt genuine events are allowed through an adaptive graph weight.

The adversary may read system documentation and historical data, compromise at most s_t sensors at time t, coordinate attack values and vary the support over time. Thus

$$\|a_t\|_0 \le s_t \ll m_t.$$

The attacker may construct small-amplitude, ramp, replay, pulse or topology-aware injections. Cryptographic compromise is outside the estimator: CR-GSDA assumes the packet has passed transport security and asks whether its content is physically credible. Denial-of-service is represented by a time-varying selection matrix H_t and missing-data mask. Actuator attacks require augmentation of d_t or u_t and are discussed as an extension.

The defender knows f_t or an approximate simulator, h_t, nominal covariance models and a graph estimate. Edge nodes may hold only local rows of H_t, local observations and neighbouring graph blocks. The defender does not know the attacked support. The primary outputs are a state estimate $\hat{x}_t$, attack estimate $\hat{a}_t$, attacked-sensor probability vector π_t, analysis covariance P_t^a and a certificate radius $\mathcal{B}_t$.

Assumption 3.1 (Forecast regularity). On the operational set $\mathcal{X}$, f_t is Lipschitz with constant ℓ_f and its approximation error obeys $\|d_t\| \le \eta_{\mathrm{model},t}$.

Assumption 3.2 (Information coercivity). The regularised information matrix

$$M_t = (P_t^-)^{-1} + H_t^\top R_t^{-1} H_t + \gamma_t L_t$$

is symmetric positive definite with $\lambda_{\min}(M_t) \ge \mu_t > 0$.

Assumption 3.3 (Sparse identifiability). After projection away from the predictable state subspace, the effective attack dictionary satisfies a restricted eigenvalue condition of order $2s_t$ with constant $\kappa_t > 0$.

The coercivity condition is mild when P_t^- is positive definite; numerical jitter may be incorporated into P_t^- before inversion. Sparse identifiability is the substantive security condition. It states that two sufficiently sparse attacks cannot generate indistinguishable projected innovations. Sensor placement and moving-target defence can be used to improve κ_t.

Graph-Sparse Robust Assimilation Model

Let $\hat{x}_t^-$ and P_t^- denote the forecast mean and covariance. Define the innovation $r_t = y_t - h_t(\hat{x}_t^-)$. A first diagnostic separates it into graph-frequency bands. With $L_t = U_t \Lambda_t U_t^\top$, the high-frequency score is

$$g_{i,t} = |[U_t\, \psi_\tau(\Lambda_t)\, U_t^\top r_t]_i|,$$

where $\psi_\tau(\lambda)$ is a smooth high-pass transfer function. Explicit eigendecomposition is unnecessary: a Chebyshev polynomial approximates $\psi_\tau(L_t)r_t$ using sparse matrix–vector products. The screening score also uses a normalised innovation and temporal persistence:

$$s_{i,t} = \omega_1 \frac{|r_{i,t}|}{\sqrt{S_{ii,t}}} + \omega_2 \frac{g_{i,t}}{\mathrm{MAD}(g_t)+\varepsilon} + \omega_3 p_{i,t-1} - \omega_4 q_{i,t},$$

$$J_t(x,a) = 1/2 \parallel x - \hat{x}_t^- \parallel^2_{(P_t^-)^{-1}} +$$
$$\sum_i \rho_{\delta_i}\left(y_{i,t} - h_{i,t}(x) - a_i/\sigma_{i,t}\right) +$$
$$\lambda_t \sum_i \vartheta_{i,t}\,|a_i| + \beta_t/2\; a^\top L_{a,t} a + \gamma_t/2\; x^\top L_t x.$$

where $S_t = H_t P_t^- H_t^\top + R_t$, $p_{i,t-1}$ is attack persistence and $q_{i,t}$ is evidence of a genuine physical event. The event evidence increases when neighbouring sensors change coherently and when known exogenous inputs explain the innovation. It reduces false alarms during actual physical transitions.

The core analysis jointly estimates state and attack:

$$(\hat{x}_t, \hat{a}_t) = \underset{x,a}{\operatorname{argmin}} J_t(x,a)$$

Here ρ_δ is the Huber loss, $\vartheta_{i,t} = 1/\left(s_{i,t} + \varepsilon\right)$ is an adaptive sparsity weight, $L_{a,t}$ is an optional attack-correlation graph and $\|z\|_A^2 = z^\top A z$. The attack graph is distinct from the physical graph: it can connect sensors sharing gateways, credentials or software. Setting $\beta_t = 0$ permits isolated attacks; positive β_t favours coordinated patterns on the cyber graph.

The physical regularisation parameter must adapt to real events. We use

$$\gamma_t = \gamma_{\min} + \frac{\gamma_{\max} - \gamma_{\min}}{1 + \exp\{c_\gamma(\bar{q}_t - q_0)\}},$$

so that strong event evidence weakens graph smoothing. This prevents a graph prior from erasing genuine local dynamics. In nonlinear systems, h_t is linearised only within the state block, or the state block is updated with an ensemble cross-covariance.

For fixed x, the attack block is convex. With a quadratic local approximation to the robust likelihood, it becomes

$$\hat{a}_t = \underset{a}{\operatorname{argmin}} \frac{1}{2}\|z_t - a\|^2_{\Omega_t} + \lambda_t \|V_t a\|_1 + \frac{\beta_t}{2} a^\top L_{a,t} a,$$

where $z_t = y_t - h_t(x)$, Ω_t contains robust precision weights and $V_t = \mathrm{diag}(\vartheta_t)$. When $\beta_t = 0$ and Ω_t is diagonal,

$$\hat{a}_{i,t} = \mathrm{soft}\left(z_{i,t}, \frac{\lambda_t \vartheta_{i,t}}{\Omega_{ii,t}}\right).$$

Graph coupling is handled by proximal gradient or ADMM. Warm starts make the online cost small because attack supports often persist over several sampling intervals.

After attack correction, $\tilde{y}_t = y_t - \hat{a}_t$. In the linear-quadratic approximation the state update solves

$$M_t \hat{x}_t = (P_t^-)^{-1}\hat{x}_t^- + H_t^\top \tilde{R}_t^{-1} \tilde{y}_t,$$

where $\tilde{R}_t$ includes robust weights and attack-reconstruction uncertainty. A sparse Cholesky, preconditioned conjugate-gradient or distributed consensus solver may be used.

Algorithm 1: Centralised CR-GSDA update Input: $\hat{x}_{t-1}$, P_{t-1}^a, model (f_t, h_t), graph L_t, observation y_t.
1. Forecast: compute $\hat{x}_t^- = f_t(\hat{x}_{t-1}, u_t)$ and $P_t^- = F_t P_{t-1}^a F_t^\top + Q_t$, or propagate an ensemble.
2. Screen: calculate normalised innovation, graph high-frequency energy, temporal persistence and genuine-event evidence using the screening equation above.
3. Alternate: initialise $a^{(0)} = \hat{a}_{t-1}$; alternate the robust state solve defined above and attack proximal step defined above until the primal change and KKT residual fall below tolerances.
4. Quantify: estimate attack covariance $\Sigma_{a,t}$, set $\tilde{R}_t = R_t + \Sigma_{a,t} + R_{\mathrm{rob},t}$, and recompute the final analysis covariance.
5. Certify: evaluate the bound $\mathcal{B}_t$ in Theorem 6.1; raise a safe-mode flag if it exceeds the application tolerance.
Output: $\hat{x}_t$, P_t^a, $\hat{a}_t$, attack probabilities π_t and certificate $\mathcal{B}_t$.

Ensemble and Bayesian Extension

For nonlinear f_t and h_t, let $\{x_{t-1}^{a,(j)}\}_{j=1}^N$ be the analysis ensemble. Forecast particles satisfy

$$x_t^{-,(j)} = f_t\left(x_{t-1}^{a,(j)}, u_t\right) + w_t^{(j)}.$$

Their sample mean and covariance are $\bar{x}_t^-$ and C_t^{xx}. Predicted observations $z_t^{(j)} = h_t\left(x_t^{-,(j)}\right)$ give C_t^{xz} and C_t^{zz}. Graph localisation multiplies long-range sample correlations by a compactly supported matrix $\mathcal{T}_t$,

$$\tilde{C}_t^{xz} = \mathcal{T}_t \circ C_t^{xz},$$

where $\circ$ denotes the Hadamard product. The ensemble analysis uses the attack-corrected observation:

$$x_t^{a,(j)} = x_t^{-,(j)} + \tilde{C}_t^{xz}\left(C_t^{zz} + \tilde{R}_t\right)^{-1}\left(\tilde{y}_t + \varepsilon_t^{(j)} - z_t^{(j)}\right).$$

The attack estimate is itself uncertain. A local Laplace approximation around $\hat{a}_t$ gives

$$\Sigma_{a,t} \approx \left(\Omega_t + \beta_t L_{a,t} + D_{\lambda,t}\right)^{-1},$$

where $D_{\lambda,t}$ is a smoothed curvature approximation to the weighted ℓ_1 penalty. For coordinates classified as attacked, a conservative diagonal alternative is obtained from bootstrap or ensemble variability. The effective observation covariance is

$$\tilde{R}_t = R_t + \Sigma_{a,t} + \zeta_t \mathrm{diag}(r_t \odot r_t),$$

where ζ_t increases when robust weights are small or the attack optimisation has not converged. Consequently, uncertain attack removal reduces the Kalman gain rather than creating false confidence.

A Bayesian interpretation augments the posterior with the latent attack:

$$p(x_t, a_t \mid y_{1:t}) \propto p(y_t \mid x_t, a_t)\, p\big(a_t \mid a_{t-1}, G_{a,t}\big)\, p(x_t \mid x_{t-1}, u_t, G_t)\, p(x_{t-1}, a_{t-1} \mid y_{1:t-1}).$$

A spike-and-slab prior provides explicit attack probabilities,

$$a_{i,t} \mid z_{i,t} \sim (1 - z_{i,t})\delta_0 + z_{i,t}\mathcal{N}(0, \tau_a^2), \qquad z_{i,t} \sim \text{Bernoulli}(\pi_{i,t}).$$

The weighted ℓ_1 implementation is a computationally economical MAP approximation. Periodic particle or variational recalibration can correct its probability estimates. High-dimensional Bayesian computation can be accelerated using reduced representations, but posterior approximation error must be reported rather than hidden [22].
Coverage is assessed using the normalised estimation error squared,

$$\text{NEES}_t = (x_t - \hat{x}_t)^{\top}(P_t^a)^{-1}(x_t - \hat{x}_t).$$

Under correct linear-Gaussian calibration, NEES follows a χ_n^2 distribution. During attacks, the target is not exact Gaussianity but conservative empirical coverage. Inflation parameters are selected on held-out attack scenarios and then frozen for evaluation.

Certified State Estimation

The certificate follows from the first-order condition of the corrected regularised analysis. It is deterministic and therefore applies when noise is non-Gaussian.

Theorem 6.1 (One-step state-error certificate). *Let the true linear observation satisfy $y_t = H_t x_t + v_t + a_t$ and let $\hat{x}_t$ solve the state-update equation with graph penalty $\gamma_t L_t$. Suppose $M_t \succcurlyeq \mu_t I$. Define $e_t = x_t - \hat{x}_t$, $e_t^- = x_t - \hat{x}_t^-$ and $\Delta a_t = a_t - \hat{a}_t$. Then $\| e_t \|_2 \le \mu_t^{-1}[\| (P_t^-)^{-1} \|_2 \| e_t^- \|_2 + \| H_t^{\top}\tilde{R}_t^{-1} \|_2 (\| v_t \|_2 + \| \Delta a_t \|_2) + \gamma_t \| L_t x_t \|_2]$. If $\|e_t^-\| \le \ell_f \|e_{t-1}\| + \eta_{\text{model},t} + \|w_t\|$, substitution yields a recursive online certificate.*

Proof. Substitute $\tilde{y}_t = H_t x_t + v_t + \Delta a_t$ into the state-update equation and subtract the identity

$$M_t x_t = (P_t^-)^{-1} x_t + H_t^{\top}\tilde{R}_t^{-1} H_t x_t + \gamma_t L_t x_t.$$

Rearrangement gives

$$M_t e_t = (P_t^-)^{-1} e_t^- - H_t^{\top}\tilde{R}_t^{-1}(v_t + \Delta a_t) + \gamma_t L_t x_t.$$

Taking Euclidean norms, using submultiplicativity and $\|M_t^{-1}\|_2 \le 1/\mu_t$ proves the stated certificate. The forecast recursion follows from Lipschitz continuity of f_t and the triangle inequality. □

The term $\gamma_t \|L_t x_t\|$ is graph-prior bias. It is small for smooth states and explicitly warns against excessive smoothing during genuine discontinuities. Unlike a certificate that silently assumes the prior is correct, the stated certificate exposes this cost. In practice x_t is unknown; an upper estimator is

$$\hat{\eta}_{G,t} = \|L_t \hat{x}_t\| + \|L_t\| \mathcal{B}_t^{\text{pilot}},$$

where $\mathcal{B}_t^{\text{pilot}}$ is the preceding iteration's radius.

Lemma 6.2 (Attack-reconstruction error). *Let the projected attack regression be $b_t = \Phi_t a_t + \xi_t$ and let $\hat{a}_t$ minimise a weighted Lasso objective. Under the restricted eigenvalue condition with constant κ_t, if $\lambda_t \ge 2\|\Phi_t^{\top}\xi_t\|_{\infty}$, then*

$$\|a_t - \hat{a}_t\|_2 \le \frac{4\lambda_t\sqrt{s_t}}{\kappa_t^2}\frac{\vartheta_{\max,t}}{\vartheta_{\min,t}}.$$

Proof. The basic inequality for the weighted Lasso bounds the prediction error plus the weighted ℓ_1 difference by the noise inner product. The chosen λ_t gives the cone condition on the error vector. Applying the restricted eigenvalue inequality on that cone and converting the support ℓ_1 norm to $\sqrt{s_t}$ times the ℓ_2 norm yields the stated bound. The weight ratio accounts for the rescaling from weighted to ordinary coordinates. □

Combining Theorem 6.1 and Lemma 6.2 produces the requested decomposition

$$\| x_t - \hat{x}_t \| \le C_{1,t} \| v_t \| + C_{2,t} \| a_t - \hat{a}_t \| + C_{3,t}\eta_{\text{model},t} + C_{4,t} \| e_{t-1} \| + C_{5,t}\eta_{G,t}.$$

The constants are computable from M_t, H_t, P_t^-, $\tilde{R}_t$, the forecast Lipschitz constant and the graph weight.

Proposition 6.3 (Recursive boundedness). *Suppose the recursive certificate can be written $\|e_t\| \le \alpha_t\|e_{t-1}\| + b_t$, with $0 \le \alpha_t \le \bar{\alpha} < 1$ and $b_t \le \bar{b}$. Then $\|e_t\| \le \bar{\alpha}^t\|e_0\| + \frac{1-\bar{\alpha}^t}{1-\bar{\alpha}}\bar{b}$, and hence the asymptotic error is bounded by $\bar{b}/(1-\bar{\alpha})$.*

Proof. Iterate the scalar inequality and sum the resulting geometric series. □

The condition $\bar{\alpha} < 1$ is an assimilation observability requirement, not automatic stability. When too many informative sensors are attacked or disconnected, μ_t decreases and the certificate expands. The estimator should then enter safe mode rather than report an unjustifiably precise value.

Distributed Edge Implementation

Partition the IoT graph into K edge regions. Region k stores local state x_k, observation y_k, attack vector a_k and copies of boundary variables. Let $E_{k\ell}x_k = E_{\ell k}x_\ell$ impose agreement across neighbouring regions. The distributed objective is

$$\min_{\{x_k, a_k\}} \sum_{k=1}^{K} J_{k,t}(x_k, a_k) \quad \text{subject to} \quad E_{k\ell}x_k = E_{\ell k}x_\ell.$$

ADMM introduces consensus variables $z_{k\ell}$ and scaled multipliers $u_{k\ell}$. Each edge node alternates a local robust graph-sparse solve, boundary averaging and dual update. Raw sensor streams remain local; exchanged messages contain boundary-state estimates, information matrices, certificate contributions and attack alarms.

Algorithm 2: Distributed edge CR-GSDA At each edge region k: forecast the local ensemble and construct local innovation statistics.
Local security update: recover a_k with gateway-aware weights; quarantine only when posterior attack probability and certificate contribution both exceed policy thresholds.
Consensus round: solve the local augmented objective, transmit boundary sufficient statistics to neighbouring regions, average consistent copies, and update dual variables.

Stopping rule: stop when primal residual r_p, dual residual r_d and certificate change are below application tolerances, or when the latency budget is reached.
Fallback: if communication fails, use covariance intersection for boundary fusion and inflate uncertainty according to the age of information.

Proposition 7.1 (Finite-consensus perturbation). *Assume the centralised objective is μ-strongly convex in the state block and the distributed iterate has KKT residual at most $\varepsilon_{\mathrm{ADMM}}$. Then its distance from the centralised analysis is bounded by $\left\|\hat{x}_t^{\mathrm{edge}} - \hat{x}_t^{\mathrm{central}}\right\| \le \frac{\varepsilon_{\mathrm{ADMM}}}{\mu}$. This term is added to the operational certificate.*

Proof. Strong monotonicity of the state-gradient map gives $\mu\|x - x^\star\|^2 \le \langle \nabla J(x) - \nabla J(x^\star), x - x^\star \rangle$. Apply Cauchy–Schwarz and the KKT residual bound, then cancel one factor of $\|x - x^\star\|$. □

For a sparse graph with $|E|$ edges, graph filtering costs $O(J|E|)$ for a degree-J Chebyshev approximation. A diagonal or localised Kalman update is $O(N|E| + Nm\ell)$ for ensemble size N and localisation width ℓ, rather than the dense $O(n^3)$ covariance cost. Attack proximal iterations cost $O\big(I_a(|E_a| + m)\big)$. Each consensus round communicates $O(|\partial V_k|)$ boundary values per region. These costs permit an explicit latency budget:

$$T_{\mathrm{forecast}} + T_{\mathrm{screen}} + I_a T_{\mathrm{prox}} + I_c(T_{\mathrm{local}} + T_{\mathrm{network}}) \le T_{\max}.$$

The implementation should use square-root covariance factors, symmetric positive-definite solvers and guarded Cholesky updates. Mixed precision is acceptable for graph filtering and ensemble propagation, but certificate accumulation and final information solves should use sufficient precision to avoid a false security margin. Model, graph and firmware versions must be included in the audit record.

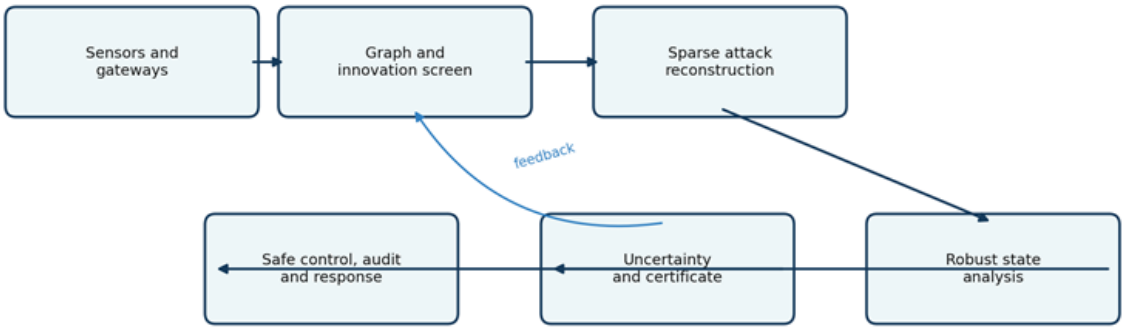

Fig 1: Edge-resident CR-GSDA pipeline. Cybersecurity decisions depend on the reconstructed attack and the state-error certificate, not on a classifier score alone.

Parameter Selection and Operational Decision Rules

Parameters are divided into statistical, structural and security groups. Q_t and R_t are estimated from attack-free calibration windows, with robust covariance estimators used when clean periods are uncertain. The graph may be built from physical adjacency, mutual information, hydraulic or electrical coupling, road connectivity, radio handover relations or a learned sparse precision matrix. Graph learning must be frozen or strongly regularised during suspected attacks; otherwise the attacker can poison the topology used for detection.

The sparsity penalty follows a noise-aware rule

$$\lambda_t = c_\lambda \hat{\sigma}_t \sqrt{2\log(m_t)/N_{\mathrm{eff},t}},$$

then is adjusted to satisfy an application-specific false-alarm constraint on attack-free validation data. The Huber threshold starts near 1.345 standard deviations for Gaussian efficiency but may be larger where heavy-tailed physical disturbances are common. γ_t is selected by blocked temporal cross-validation, not random shuffling, because random folds leak future state information.

Four decision levels are recommended. Level 0 reports normal operation. Level 1 records a statistical anomaly but leaves control unchanged. Level 2 marks sensors as suspect, uses corrected measurements and increases uncertainty. Level 3 enters safe mode when the certificate exceeds a state-risk tolerance, observability deteriorates, or consensus fails.

Automatic quarantine is avoided when it would remove the redundancy required for secure estimation.

An alarm is issued for sensor i only if

$$\pi_{i,t} \ge \tau_\pi, \quad |\hat{a}_{i,t}| \ge \tau_a, \quad \text{and} \quad \Delta\mathcal{B}_{i,t} \ge \tau_B,$$

where $\Delta\mathcal{B}_{i,t}$ measures the sensor's contribution to the certificate. Requiring agreement across probability, magnitude and estimation consequence reduces nuisance alarms. A low-amplitude attack may still be prioritised if its cumulative state impact is large.

Synthetic Experimental Design

The numerical study is deliberately synthetic. It demonstrates algorithmic behaviour and reproducibility; it is not evidence of operational readiness in a city, utility or telecommunications network. A 60-node ring sensor graph is used so that all matrices and random seeds are fully controllable. The Laplacian is $L = D - W$, and the stable process is

$$x_t = (I - 0.055L)x_{t-1} + w_t, \qquad w_t \sim \mathcal{N}(0, 0.018I).$$

All states are observed with $H = I$ and $v_t \sim \mathcal{N}(0, 0.08I)$. Each run has 300 time steps. At every step, 10% of sensors are sampled without replacement and receive signed attacks

$$a_{i,t} = \xi_{i,t}\, 2.5\big(0.65 + 0.7U_{i,t}\big),$$

where $\xi_{i,t} \in \{-1,1\}$ and $U_{i,t} \sim \text{Uniform}(0,1)$. Twenty independent seeds, 0 through 19, define the replications.

Four estimators are compared. KF is the conventional diagonal-covariance Kalman filter. Huber-KF clips the innovation influence at 2.5 predicted standard deviations. Sparse-KF estimates attacks by soft thresholding with threshold 0.62. CR-GSDA uses threshold 0.56, an attack-graph solve $(I + 0.008L)^{-1}$, and a state graph solve $(I + 0.025L)^{-1}$. These constants are fixed across the 20 evaluation seeds. Attack localisation declares $|\hat{a}_{i,t}| > 0.28$.

Evaluation metrics are

$$\text{RMSE} = \left[\frac{1}{nT}\sum_{t=1}^{T} \| x_t - \hat{x}_t \|^2\right]^{1/2}.$$

$$\text{MAE} = \frac{1}{nT}\sum_{t=1}^{T} \| x_t - \hat{x}_t \|_1.$$

$$\text{Precision} = \frac{TP}{TP+FP}, \qquad \text{Recall} = \frac{TP}{TP+FN}.$$

$$F_1 = \frac{2\,\text{Precision}\,\text{Recall}}{\text{Precision}+\text{Recall}}.$$

The primary endpoint is state RMSE because detection has value only insofar as it protects physical-state inference. Secondary endpoints are MAE and localisation F1. Mean and sample standard deviation are reported across seeds.

The threat suite for a larger validation campaign should add pulse, ramp, replay, slowly drifting, topology-aware and clustered gateway attacks; missing packets; graph misspecification; model drift; and genuine abrupt events. Each hyperparameter must be selected on disjoint seeds before final testing. Real-world evaluation should use a cyber range or authorised digital twin rather than attacks on operational infrastructure.

Table 2: Reproducible synthetic benchmark configuration.

Component	Specification	Value
Graph	Undirected ring, two neighbours per node	n = 60
Dynamics	Diffusive transition F = I − 0.055L	T = 300
Process noise	Independent Gaussian	Q = 0.018I
Observation noise	Independent Gaussian	R = 0.08I
Attack support	Uniform without replacement per time	s = 6
Attack magnitude	Random sign and bounded random scale	base 2.5
Replications	Deterministic NumPy random generators	seeds 0: 19
CR-GSDA	Soft threshold, attack and state graph weights	0.56,0.008,0.025
Detection	Magnitude threshold	0.28

Results

Table 3 reports the computed synthetic results. The conventional KF has RMSE 0.3717 because it incorporates the injected values as if they were trustworthy measurements. Robust clipping reduces RMSE by 44.1%. Explicit sparse correction reduces it further. CR-GSDA records the lowest mean RMSE, 0.1916, a 48.5% reduction relative to KF and a 7.8% reduction relative to Huber-KF. The difference from Sparse-KF is deliberately modest because the benchmark graph is simple and the physical process is already well matched by a diagonal forecast.

Table 3: State-estimation performance over 20 deterministic synthetic trials. Values are mean ± sample standard deviation.

Method	RMSE	MAE
Conventional KF	0.3717 ± 0.0034	0.2569 ± 0.0025
Huber-KF	0.2078 ± 0.0023	0.1600 ± 0.0016
Sparse-KF	0.1929 ± 0.0021	0.1508 ± 0.0015
CR-GSDA	0.1916 ± 0.0021	0.1503 ± 0.0015

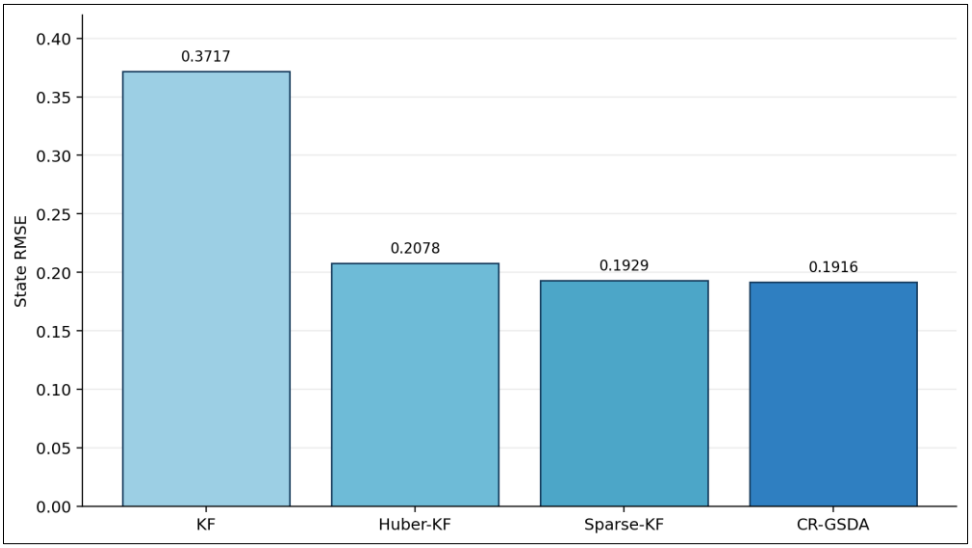

Fig 2: Mean state RMSE in the synthetic 60-node experiment.

Attack localisation by CR-GSDA achieves precision 0.8535 ± 0.0106, recall 0.9989 ± 0.0010 and F1 0.9204 ± 0.0062. High recall is expected because attack magnitudes substantially exceed measurement noise. Precision is lower because occasional large innovations from process and observation noise cross the fixed attack threshold. In an operational deployment, posterior probability, event evidence and persistence should replace a single magnitude threshold.

Table 4: CR-GSDA attack-localisation performance in the synthetic benchmark.

Metric	Precision	Recall	F1 score
Mean	0.8535	0.9989	0.9204
Sample standard deviation	0.0106	0.0010	0.0062

The results support three limited conclusions. First, robustification is essential: a standard Gaussian update is not resilient to sparse integrity attacks. Second, explicit attack reconstruction slightly outperforms influence clipping because it reuses corrected observations rather than merely suppressing them. Third, graph regularisation can improve estimation when the state is genuinely graph smooth, but the gain depends on topology accuracy and should not be generalised beyond this benchmark.

No significance test is emphasised because the seeds are paired algorithmic replications rather than a random sample of field conditions. A future comparative study should report paired bootstrap confidence intervals across heterogeneous scenarios and compute cost, latency, energy and coverage alongside RMSE.

Ablation, Sensitivity and Stress-Testing Protocol

A credible evaluation must determine which component produces each benefit. The ablation sequence removes: graph screening; adaptive attack weights; Huber likelihood; attack covariance inflation; graph state regularisation; event evidence; and distributed consensus. The expected outcome is not that every component lowers RMSE in every regime. For example, graph regularisation should be neutral or harmful when the physical state contains a legitimate sharp boundary. Its value is a lower error in smooth regimes without unacceptable bias in event regimes.

Attack sparsity is varied from 0 to 40% of sensors. The theoretical phase transition occurs when sparse observability fails, but its empirical location depends on sensor placement and dynamics. Attack magnitude is swept from below the noise standard deviation to ten times that value. Low-magnitude attacks test cumulative bias, whereas high-magnitude attacks test robust influence. Attack duration distinguishes isolated pulses from persistent drifts. Support mobility measures whether temporal priors become misleading.

Graph misspecification is created by randomly deleting, adding and reweighting edges. The spectral distance $\|L - \tilde{L}\|_2$ is recorded. The certificate must increase with graph mismatch through the model term rather than remain falsely narrow. Genuine-event scenarios impose local state jumps that then diffuse according to the physical model. A successful method should assign high event evidence, reduce γ_t and avoid labelling the entire transition as malicious.

The edge stress test varies packet loss, delay, processor throttling and the number of permitted consensus rounds. Accuracy is plotted against end-to-end latency and transmitted bytes. Safe mode is evaluated as a selective-prediction problem: among time points at which the estimator accepts its output, what is the maximum error and empirical coverage? A useful certificate should permit routine states while abstaining under unobservable or severely attacked conditions.

Table 5: Recommended stress-test matrix for a publication-scale validation.

Factor	Levels	Primary diagnostic
Attacked fraction	0, 5, 10, 20, 30, 40%	RMSE and recoverability transition
Attack type	pulse, ramp, drift, replay, coordinated	Detection delay and cumulative bias
Graph error	0–30% edge perturbation	Bias and certificate calibration
Model mismatch	parameter drift and unmodelled input	False alarms versus model term
Physical event	local jump, moving front, regime switch	Attack/event discrimination
Communication	delay, loss, disconnection	Consensus error and availability
Edge budget	CPU cap, memory cap, round limit	Latency–accuracy frontier
Uncertainty	Gaussian, heavy-tailed, heteroscedastic	Coverage and NEES

ICT Application Profiles

In a smart city, x_t may describe traffic density, air quality, water pressure and crowd flow on interdependent graphs. A genuine event is likely to propagate through spatial neighbours and respond to exogenous variables such as weather or a scheduled gathering. An attack may instead align with a shared gateway or administrative domain. The two-graph construction—physical L_t and cyber $L_{a,t}$—is especially valuable because physical proximity and common compromise pathways are different relations.

For intelligent transportation, connected vehicles and roadside units produce asynchronous measurements with strict latency constraints. The forecast model can be a cell-transmission model, a graph neural surrogate or a traffic PDE discretisation. Event evidence includes accident reports and coherent upstream/downstream shockwaves. The certificate can gate automated signal timing: if the estimated queue state is insufficiently reliable, the controller reverts to a conservative plan.

In telecommunications, the state includes link load, interference, handover demand, queue occupancy and equipment health. The graph may combine physical cells and logical routing. False telemetry can provoke harmful load balancing or conceal congestion. Edge CR-GSDA can run at multi-access edge nodes; only boundary statistics need cross-region exchange. Network slicing requires separate covariance and risk thresholds because a safety-critical slice should not share the tolerance of a best-effort service.

Smart agriculture uses soil-moisture, weather, pump and crop sensors that are resource constrained and intermittently

connected. Genuine spatial heterogeneity is common, so graph smoothing must follow soil zones, elevation and irrigation topology rather than Euclidean distance alone. The estimator can identify a compromised moisture sensor without suppressing an actual leak. Certificate-driven control is important because excessive irrigation and missed drought stress both have costs.

Smart energy remains the most established state-estimation application. PMU and SCADA streams create multirate data, while topology changes and renewable intermittency create model mismatch. Dynamic state-estimation research increasingly combines prevention, detection and mitigation [23]. CR-GSDA complements these efforts by reconstructing the injection and including its uncertainty in the corrected state. Application-specific AC or distribution-system models should replace the generic linear benchmark before any operational claim.

Security, Safety and Governance

CR-GSDA is one layer in defence in depth. Device identity, secure boot, signed firmware, encrypted transport, access control, network segmentation and logging remain necessary. The estimator addresses semantic integrity: authenticated data can still be false when a trusted endpoint is compromised. Its outputs should be forwarded to a security information and event management system with sensor identifiers, attack probabilities, estimated magnitudes, graph context and certificate history.

The attack estimate is sensitive security information. Publishing real-time support locations could help an adversary learn detection thresholds. Access must be role based, retained according to policy and separated from public operational dashboards. Where sensors relate to people or vehicles, graph construction and cross-region fusion must comply with privacy requirements. Sufficient-statistic exchange reduces raw-data exposure but is not automatically differentially private.

Human oversight is required for high-consequence isolation. A sensor should not be permanently quarantined solely because a model disagrees with it; the model may be wrong. The audit trail should preserve the raw packet hash, model version, graph version, parameters, optimisation residual, uncertainty inflation and decision rule. This information supports incident investigation and protects operators against opaque automated decisions.

The safe-mode threshold must be derived from physical risk. A telecommunications estimator may tolerate a moderate queue error, whereas a medical or grid-protection application may require immediate abstention. The certificate is valuable because it translates multiple uncertainties into a state-error scale, but it remains conditional on stated bounds and cannot prove safety when model mismatch is unbounded.

Adversarial evaluation must be authorised. Synthetic data, digital twins and isolated cyber ranges should be used before any live pilot. Responsible disclosure procedures are necessary if evaluation identifies vulnerabilities in commercial devices or public infrastructure.

Discussion

The central design choice is to estimate the attack rather than merely classify the observation. Classification separates security from numerical inference: a detector emits a label and a filter then proceeds as though that label were correct. Joint reconstruction keeps uncertainty visible. If $\hat{a}_t$ is inaccurate, the state certificate and inflated covariance reflect the remaining risk.

Graph information helps distinguish physical and cyber structure, but it can also encode a dangerous prior. A road closure, burst pipe or localised voltage disturbance may be graph nonsmooth. The adaptive event term and explicit graph-bias component are therefore not cosmetic; they prevent the method's strongest structural assumption from becoming an attack on genuine information. Two graphs are preferable: one for physical propagation and one for common cyber compromise.

The theoretical bound is conservative because it uses operator norms and worst-case addition. Conservatism is acceptable for safe-mode gating but may limit availability. Tighter certificates could use directional norms, local observability Gramians, concentration inequalities and conformal calibration. Bayesian credible regions answer an average probabilistic question, while the deterministic certificate answers a bounded perturbation question. Reporting both is more informative than treating either as universally valid.

The synthetic results show a clear failure of the conventional KF and a smaller advantage of graph-sparse correction over a sparse filter. They do not establish superiority across infrastructure types. The benchmark favours graph methods because the process is diffusive on the same graph used by the estimator. Graph perturbation and genuine-event tests are essential next steps. The contribution of the present results is transparency: parameters, distributions, seeds and metrics are stated, and the experiment does not masquerade as field evidence.

Computational deployment requires co-design. A theoretically sophisticated estimator that exceeds the sampling interval is unusable. Polynomial graph filters, localisation, warm-started proximal updates and bounded consensus rounds make the method compatible with edge processors. The latency cutoff introduces optimisation error, which the distributed perturbation term places inside the certificate.

Limitations and Future Research

The present formulation assumes sparse sensor attacks. Dense low-amplitude corruption, collusion across most informative sensors or an attacker that simultaneously poisons the model and graph may violate identifiability. Group sparsity, low-rank-plus-sparse decompositions and moving-target sensing are promising extensions. Actuator attacks require unknown-input modelling because they alter the state trajectory rather than only the observation.

The graph is treated as known within each update. Online graph learning under attack is vulnerable to poisoning. A robust bilevel method should learn topology from long clean windows while limiting the influence of recent data. Temporal multilayer graphs could represent physical, communication and organisational relations without collapsing them into one Laplacian.

The Laplace approximation to attack uncertainty is imperfect near zero because the ℓ_1 penalty is nonsmooth and the posterior can be multimodal. Spike-and-slab variational inference, sequential Monte Carlo or conformal prediction may improve calibration at greater cost. Future theory should connect support-selection uncertainty directly to credible-set coverage.

The certificate assumes usable bounds for noise and model error. Learning these bounds without leaking attack energy

into the noise estimate is challenging. Robust online conformal calibration with change-point protection is a promising direction. Another priority is certificate-aware sensor scheduling: measurements should be requested where they reduce the bound most per unit of energy and bandwidth. The benchmark should be extended to nonlinear cyber-physical simulators, public traffic and environmental data with injected attacks, smart-grid test cases and hardware-in-the-loop edge platforms. A preregistered evaluation should freeze hyperparameters, attacks and endpoints before comparison. Runtime, peak memory, transmitted bytes, energy, false-alarm burden and selective coverage should accompany accuracy.

Conclusion

This article formulated cyberattack resilience as a numerical data-assimilation problem rather than an isolated intrusion-classification task. CR-GSDA combines robust forecasting, graph-frequency evidence, sparse attack reconstruction, graph-regularised analysis, ensemble nonlinear propagation, Bayesian uncertainty inflation and distributed edge consensus. The resulting certificate makes explicit how measurement noise, residual attack error, forecast error, model mismatch, graph bias and incomplete consensus contribute to the final state uncertainty.

In a reproducible synthetic 60-node experiment, CR-GSDA reduced state RMSE from 0.3717 for a conventional Kalman filter to 0.1916 and achieved an attack-localisation F1 score of 0.9204. These values demonstrate the mechanism under controlled assumptions, not field validation. The most important operational principle is that a smart infrastructure should not silently trust a corrected measurement: it should retain the uncertainty of that correction and abstain when observability or the certificate becomes unacceptable.

The framework is applicable to smart cities, intelligent transportation, telecommunications, agriculture and energy, provided that each deployment supplies a defensible process model, physical and cyber graphs, authorised validation and risk-based safe-mode rules. Future work should address dense attacks, poisoned topology learning, actuator corruption, conformal certificate calibration and hardware-in-the-loop evaluation.

Declarations

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflict of Interest

The author declares no conflict of interest.

Data Availability

No operational infrastructure data were used. All numerical results are based on the synthetic model, parameters, distributions and deterministic random seeds reported in the manuscript. The accompanying implementation can be reconstructed directly from the experimental specification.

Ethics Approval

The study uses synthetic computational data and does not involve human participants, identifiable personal data or animals.

Author Contributions

Conceptualisation, methodology, mathematical formulation, analysis, software design, writing and validation: Kaarina Nakale.

Use of Artificial Intelligence

Any language or computational assistance used during manuscript preparation should be disclosed to the target journal in accordance with its current editorial policy. The author remains responsible for the accuracy, originality, citations and final text.

References

1. Zhou Z, Zhang X, Zhang J, Taylor G. Comprehensive review on dynamic state estimation techniques with cybersecurity applications. IET Smart Grid. 2024;7(4):370-385. doi:10.1049/stg2.12168.
2. Liu Y, Ning P, Reiter MK. False data injection attacks against state estimation in electric power grids. In: Proceedings of the 16th ACM Conference on Computer and Communications Security; 2009. p. 21-32. doi:10.1145/1653662.1653666.
3. Fawzi H, Tabuada P, Diggavi S. Secure estimation and control for cyber-physical systems under adversarial attacks. IEEE Trans Autom Control. 2014;59(6):1454-1467. doi:10.1109/TAC.2014.2303233.
4. Pajic M, Lee I, Pappas GJ. Attack-resilient state estimation for noisy dynamical systems. IEEE Trans Control Netw Syst. 2017;4(1):82-92. doi:10.1109/TCNS.2016.2607420.
5. Kalman RE. A new approach to linear filtering and prediction problems. J Basic Eng. 1960;82(1):35-45. doi:10.1115/1.3662552.
6. Evensen G. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dyn. 2003;53:343-367. doi:10.1007/s10236-003-0036-9.
7. Al-Ghattas O, Sanz-Alonso D. Ensemble Kalman filters with resampling. SIAM/ASA J Uncertain Quantif. 2024;12(3):1036-1071. doi:10.1137/23M1594935.
8. Li H, Li W, Yang L, Zheng F, Wang Y. Research on attack detection and state estimation for distribution networks based on sparse recovery and unknown input observer. Discov Appl Sci. 2026;8:5. doi:10.1007/s42452-025-07979-y.
9. Zhang G, *et al*. Joint detection and localization of false data injection attacks in smart grids: an enhanced state estimation approach. Comput Electr Eng. 2024.
10. Xiao Y, *et al*. Distributed state estimation for cyber-physical systems against sensor attacks based on Isolation Forest. Inf Sci. 2025:122258. doi:10.1016/j.ins.2025.122258.
11. Shuman DI, Narang SK, Frossard P, Ortega A, Vandergheynst P. The emerging field of signal processing on graphs: extending high-dimensional data analysis to networks and other irregular domains. IEEE Signal Process Mag. 2013;30(3):83-98. doi:10.1109/MSP.2012.2235192.
12. Drayer E, Routtenberg T. Detection of false data injection attacks in smart grids based on graph signal processing. IEEE Syst J. 2020;14(2):1886-1896. doi:10.1109/JSYST.2019.2927469.
13. Lai X, *et al*. Detection of false data injection attacks in smart grid based on graph signal processing. IEEE Access. 2024.
14. Sanz-Alonso D, Waniorek N. Analysis of a computational framework for Bayesian inverse

problems: ensemble Kalman updates and MAP estimators under mesh refinement. SIAM/ASA J Uncertain Quantif. 2024;12. doi:10.1137/23M1567035.
15. Satyanarayanan M. The emergence of edge computing. Computer. 2017;50(1):30-39. doi:10.1109/MC.2017.9.
16. Nencioni G, Garroppo RG, Gonzalez AJ. 5G multi-access edge computing: a survey on security, dependability, and performance. IEEE Access. 2023.
17. Nakahira Y, Mo Y. Attack-resilient H2, H∞, and l1 state estimator. IEEE Trans Autom Control. 2018;63(12):4353-4360.
18. Miao F, Zhu Q, Pajic M, Pappas GJ. Coding schemes for securing cyber-physical systems against stealthy data injection attacks. IEEE Trans Control Netw Syst. 2017;4(1):106-117.
19. Cheng Z, *et al*. Distributed consensus estimation for networked multi-sensor systems subject to hybrid attacks and missing measurements. Sensors. 2024;24.
20. Hu X, *et al*. An enhanced privacy protection distributed state estimation in wireless sensor networks against eavesdropping attacks. IEEE Internet Things J. 2025;12.
21. Ernst OG, Sprungk B, Starkloff HJ. Analysis of the ensemble and polynomial chaos Kalman filters in Bayesian inverse problems. SIAM/ASA J Uncertain Quantif. 2015;3(1):823-851. doi:10.1137/140981319.
22. Lan S, Li S, Shahbaba B. Scaling up Bayesian uncertainty quantification for inverse problems using deep neural networks. SIAM/ASA J Uncertain Quantif. 2022;10(4):1684-1713. doi:10.1137/21M1439456.
23. Paudel S, Dhungana D. False data injection attacks against state estimation in the smart grid. In: Proceedings of IEEE EUROCON 2025; 2025. p. 1-6. doi:10.1109/EUROCON64445.2025.110731.

How to Cite This Article

Nakale K. Cyberattack-resilient data assimilation for smart IoT infrastructures: sparse attack reconstruction and certified state estimation. Int J Appl Math Numer Res. 2026;2(4):62-72.

Creative Commons (CC) License