Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.079984

ARTICLE

A Bilevel Deep Learning Optimization Framework for Joint Energy Harvesting Prediction and Energy-Aware Scheduling in IoT-Based Wireless Sensor Networks

Mohammad Q. Al-Jamal[1], **Mahmoud Al Jamal**[2], **Bashar S. Khassawneh**[3,*], **Ayoub Alsarhan**[4,5], **Amina Salhi**[6] and **Tahani Alsubait**[7]

[1]Department of Renewable Energy, Jadara University, Irbid, Jordan
[2]Department of Cybersecurity, Irbid National University, Irbid, Jordan
[3]Department of Computer Science, College of Information Technology, Amman Arab University, Amman, Jordan
[4]Faculty of Information Technology, Al-Ahliyya Amman University, Amman, Jordan
[5]Department of Information Technology, Faculty of Prince Al-Hussien bin Abdullah, The Hashemite University, Zarqa, Jordan
[6]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia
[7]Department of Computer Science and Artificial Intelligence, Umm Al-Qura University, Makkah, Saudi Arabia

*Corresponding Author: Bashar S. Khassawneh. Email: b.khassawneh@aau.edu.jo

Received: 01 February 2026; Accepted: 22 April 2026; Published: 23 July 2026

ABSTRACT: Energy sustainability and secure operation are persistent challenges in Internet-of-Things (IoT) wireless sensor networks (WSNs), where limited battery capacity, heterogeneous traffic, and security procedures jointly drive premature node depletion and service degradation. This paper proposes an *uncertainty-aware bilevel co-optimization framework* that unifies residual-energy prediction with robust, energy-aware scheduling for clustered IoT-WSNs. At the lower level, a lightweight temporal predictor (TCN + LSTM with stochastic sampling) learns short-horizon residual-energy evolution from multivariate, dataset-aligned windows capturing sensing/communication activity, proximity-to-cluster-head effects, and security overhead (authentication latency, key exchange, and rekeying), and produces both point forecasts and uncertainty estimates to enable risk-sensitive control. At the upper level, a constrained, horizon-based scheduler selects per-node actions (duty cycle, sensing rate, transmission power) to extend network lifetime and balance residual energy while enforcing safety thresholds and operational bounds; bilevel coupling is realized via differentiable hypergradient updates, complemented by trust-region action smoothing and adaptive primal–dual constraint handling to suppress energy-critical states under uncertainty. On a real-world WSN energy–security dataset, the proposed model attains the best lower-level learning performance with MAE = 0.004, RMSE = 0.006, and R^2 = 0.995 for residual-energy regression, and up to 0.98 accuracy/0.98 F1 for secure-and-efficient classification. End-to-end scheduling results show that the full framework improves estimated network lifetime by up to 1.60×, reduces residual-energy variance to 0.60×, and lowers safety violations to 0.35× relative to a fixed-policy baseline, demonstrating robust, secure, and sustainable IoT-enabled WSN operation.

KEYWORDS: Internet of Things; wireless sensor networks; bilevel optimization; energy-aware scheduling

1 Introduction

The pervasive integration of Internet of Things (IoT) sensing infrastructure into cyber-physical environments has elevated wireless sensor networks (WSNs) from simple data collectors to critical components of intelligent monitoring and control in smart cities, industrial automation, environmental surveillance, and resilient infrastructure [1]. In such deployments, energy sustainability is a first-order operational constraint: once nodes deplete their batteries, the network experiences coverage gaps, degraded connectivity, increased packet loss, and ultimately service disruption. In many modern WSN paradigms, particularly energy-harvesting-enabled systems, node operation is influenced not only by energy consumption but also by intermittent and environment-dependent energy replenishment processes (solar or ambient harvesting). In practical deployments, the dominant and directly observable quantity remains the residual battery energy, which reflects the net effect of both energy consumption and any implicit or unobserved energy inflows. Unlike idealized models, real WSN energy consumption is shaped by multiple coupled processes, including sensing and local processing, wireless transmission and reception, cluster coordination, and security procedures such as authentication and cryptographic key management. These factors interact nonlinearly and vary over time due to heterogeneous node roles, clustering topology, proximity-dependent radio costs, and dynamic security overheads, producing non-stationary energy trajectories that are difficult to manage using static scheduling rules [2]. Accordingly, modeling residual energy evolution provides a control-relevant abstraction of system dynamics, even in scenarios where explicit energy harvesting processes are not directly measured.

Recent progress in data-driven modeling and deep learning has enabled more expressive representations of WSN dynamics, particularly for temporal prediction tasks. Many existing methods treat residual energy prediction and scheduling as separable stages: a predictor is trained offline to minimize point-wise error, while a scheduler is designed independently, often using deterministic assumptions that treat predictions as [3]. In the context of energy-harvesting WSNs, several works attempt to incorporate harvesting-aware models; however, these often rely on simplified or assumed energy inflow distributions, limiting their applicability under real-world uncertainty and heterogeneous operating conditions. This decoupling is problematic in practice because the cost of error is asymmetric and constraint-sensitive. Overestimating residual energy can lead to aggressive duty cycles, sensing rates, or transmit power settings that accelerate depletion and trigger early node failure, whereas underestimation primarily reduces throughput or sensing fidelity. Moreover, network operation is inherently multi-objective and constrained, requiring simultaneous consideration of lifetime maximization, energy balance, communication performance, and security compliance. Consequently, predictive accuracy alone is insufficient; what is needed is *control-relevant* prediction that is uncertainty-aware and explicitly aligned with long-term operational objectives [4]. This is particularly important when residual energy is used as a proxy for system sustainability in the absence of explicit harvesting measurements.

To address these challenges, this paper proposes a unified deep learning-based bilevel co-optimization framework that tightly couples residual energy prediction (lower level) with energy-aware scheduling (upper level) in a closed-loop learning–control architecture. Rather than explicitly modeling stochastic energy harvesting processes, the proposed framework focuses on learning the temporal evolution of residual energy as a measurable and control-relevant system state, capturing the net effect of workload, communication, and security-induced energy consumption. At the lower level, a lightweight temporal model combining temporal convolution and gated recurrent processing learns node-level energy evolution from multivariate observations that capture energy states, packet statistics, clustering context, and security-induced overhead. In addition to point forecasts, uncertainty is estimated via stochastic inference to quantify confidence under non-stationary consumption regimes. At the upper level, scheduling is formulated as a constrained robust optimization problem that adapts duty cycle, sensing rate, and transmission power to reduce energy depletion and imbalance while enforcing explicit safety thresholds and feasible action bounds. The coupling between the two levels is realized through hypergradient-based updates, allowing the long-term scheduling objective to shape the learning process and thereby producing a predictor optimized for downstream sustainability rather than short-horizon regression error. This design ensures that the framework remains applicable to both conventional battery-powered WSNs and future extensions involving explicit energy harvesting models.

The proposed framework offers many contributions as follows:

- Uncertainty-aware energy prediction: A lightweight TCN + LSTM model predicts node residual energy and estimates uncertainty using energy, traffic, clustering, and security-overhead features.
- Bilevel robust scheduling: A constrained upper-level optimizer adapts duty cycle, sensing rate, and transmit power to improve lifetime and energy balance under safety and security costs.
- Differentiable closed-loop co-optimization: Hypergradient coupling with trust-region and primal–dual updates enables online learning–control adaptation via replay-buffer refinement.

2 Literature Review

Research on energy-aware wireless sensor networks (WSNs) has evolved along several largely separate directions, including protocol-driven scheduling and routing, data-driven prediction of energy or network states, and learning-based control or policy optimization. Although each line of work contributes useful insights, the literature remains fragmented in terms of how prediction and decision-making are integrated. In particular, most existing studies either (i) optimize scheduling using hand-crafted or heuristic rules without a learned predictive model, (ii) perform prediction as an isolated task without explicitly linking it to downstream control objectives, or (iii) learn control policies directly but rely on fixed state abstractions, simplified dynamics, or weak safety treatment. This fragmentation creates a methodological gap in the design of control-aware predictive systems for WSNs, especially under uncertainty, security overhead, and heterogeneous node behavior. The present work addresses this gap by formulating residual-energy prediction and scheduling as a hierarchically coupled bilevel optimization problem in which the predictive model is explicitly shaped by downstream control requirements.

Sah et al. [5]. This paper introduces ERAS, an efficient routing awareness-based scheduling scheme for solar energy-harvesting WSNs that integrates hierarchical clustering with TDMA synchronization to curb redundant transmissions and extend lifetime. The study targets the coupled problems of cluster-head (CH) election and route discovery under variable harvested energy, defining the problem as minimizing CH count subject to coverage and energy-balance constraints while modeling radio (free-space/multipath) and solar EH dynamics. Methodologically, ERAS selects CHs using a composite score of residual and harvested energy, schedules CM activity via TDMA sleep/awake control, aggregates data at CHs, and uses single-/multi-hop inter-cluster routing; complexity and message-exchange bounds are derived, and MATLAB simulations compare ERAS with ERA, SEED, TM2RP, and CREST on two network sizes (100–200 nodes). ERAS achieves the best lifetime and stability: lifetime share 79.8% vs. 67.3%/73.8%/75.2% (ERA/SEED/CREST), and higher LND/FND/HND rounds (EH-WSN#1: LND = 4898 vs. 4612 (CREST); EH-WSN#2: LND = 4011 vs. 3581) with improved packet delivery and lower average energy consumption. From a methodological perspective, however, ERAS remains a protocol- and rule-driven scheduling method rather than a predictive-control framework. Its decision mechanism is based on engineered energy cues and deterministic coordination logic, not on a learned temporal model capable of anticipating future energy evolution under uncertainty. As a result, it does not support uncertainty-aware forecasting, risk-sensitive action selection, or differentiable coupling between learning and scheduling. Its contribution is therefore strong at the protocol layer, but limited in terms of adaptive predictive intelligence and closed-loop learning–control co-design.

Oudah and Shaker [6]. This survey reviews Wireless Sensor Networks (WSNs) for renewable energy monitoring, covering architectures, deployments (PV, wind, hydro, microgrids), and energy-aware routing

(LEACH/PEGASIS/TEEN/ZRP and recent variants). The objective is to map protocol capabilities to domain constraints and identify performance bottlenecks in latency, lifetime, and reliability. Methodologically, it synthesizes protocol taxonomies, application case studies, and emerging enablers (edge/AI, blockchain, federated learning, 6G). Key results include protocol–use-case matching (OLSR for frequent transmissions; AODV for event-driven monitoring) and quantified gains from optimized clustering; EE-OLEACH reports ~56% network-lifetime improvement and ~30% higher energy efficiency over baseline LEACH. Persistent issues include power scarcity in harsh sites, scalability under high node density, and limited field validation. Its value lies in identifying broad research trends and system challenges, but it does not formalize a concrete framework for jointly learning energy dynamics and optimizing control decisions. In particular, the survey highlights enabling technologies such as edge intelligence and AI-assisted optimization without specifying how predictive uncertainty, scheduling feasibility, and long-term network objectives can be unified within a single mathematical formulation. Consequently, it provides important context for the field, yet still leaves open the central problem addressed in this paper: how to design a control-aware learning architecture that is both uncertainty-sensitive and operationally constrained.

Alrayes et al. [7]. The paper proposes a privacy-preserving statistical learning pipeline for IoT intrusion detection (PPSLOA-HDBDE) that targets high-dimensional big-data settings by combining linear scaling normalization (LSN), Sand Cat Swarm Optimization (SCSO) for feature selection, an ensemble of TCN/MAE/XGBoost classifiers, and hyperparameter tuning via an Improved Marine Predator Algorithm (IMPA). The stated objective is to preserve data confidentiality while maintaining analytical efficacy; the problem is framed as an accurate multi-class IDS under dimensionality, privacy, and computational constraints. The system normalizes inputs, selects features with SCSO (fitness balances accuracy and subset size), trains the TCN/MAE/XGBoost ensemble, and optimizes with IMPA (Brownian/Levy-phase search plus opposition-based learning). On BoT-IoT, the approach reports best performance at 1500 epochs: Accuracy 99.49%, Precision 98.74%, Recall 98.31%, F1 98.52%, and MCC 98.19%; it also outperforms 10 baselines (CNN-BiLSTM 99.15%) and achieves the lowest processing time (8.48 s) among compared methods. Although this study demonstrates the effectiveness of advanced learning pipelines for IoT analytics, it belongs to a fundamentally different methodological class: prediction-only or inference-centric learning. The optimization process is used to improve feature selection and classifier performance, but not to shape a downstream resource-allocation or scheduling policy. In other words, learning and decision-making are not coupled through a shared control objective. Furthermore, the problem setting is centered on intrusion detection rather than energy-aware network management, meaning that questions of lifetime maximization, action feasibility, uncertainty-aware safety margins, and scheduling under energy constraints remain outside its scope. Its relevance to the present work is therefore conceptual rather than structural: it shows the value of advanced learning under IoT constraints, but not the value of bilevel co-optimization for energy-aware control.

Fereidunian et al. [8]. This paper tackles online power control in large energy-harvesting (EH) wireless networks by learning *policies* with a deep neural network (DNN), addressing the intractability of stochastic control/MDP formulations under high-dimensional state spaces. The objective is to approximate the optimal online transmit-power mapping that maximizes time-averaged throughput without noncausal information. The authors (i) solve many *offline* convex instances (with noncausal EH/channel knowledge) to generate optimal state→power labels, (ii) train a fully connected DNN (input $3K$ neurons for per-node harvested energy, battery level, and channel gain; output K transmit energies; LeakyReLU; MSE loss) to imitate the offline policy, and (iii) deploy the trained DNN online, feeding current state (E, B, G) each slot. For a $K = 5$ EH multiple-access channel, the DNN achieves ≈90% of the offline throughput across EH means (~88%–91%); for a point-to-point link it attains ≈98% of offline and *outperforms* a discretized MDP policy

(MDP ~83% of offline), demonstrating strong generalization and low inference cost (matrix multiplies only). This work is closer to the present study because it combines learning with control; however, the coupling remains fundamentally different from bilevel co-optimization. The DNN is trained to imitate offline-optimal policies, meaning that learning is supervised by precomputed labels rather than shaped by an online control objective. As a result, the model does not explicitly incorporate predictive uncertainty, does not optimize a lower-level temporal predictor jointly with an upper-level scheduler, and does not expose a differentiable dependency through which the control objective can influence representation learning. Moreover, the formulation assumes explicit harvested-energy and channel-state variables and abstracts away security overheads, node heterogeneity, and risk-sensitive safety envelopes. It therefore represents a learning-based control paradigm, but not a control-aware predictive framework in which prediction and scheduling are optimized as mutually dependent components.

Table 1 clarifies the central gap in the existing literature. Prior studies generally focus on one of three isolated perspectives: *protocol-centric scheduling*, *prediction-only modeling*, or *learning-based control*. While each perspective addresses an important part of the broader problem, none simultaneously provides: (i) a temporally grounded and uncertainty-aware predictor of node energy evolution, (ii) a constrained scheduler that explicitly accounts for safety, efficiency, and security overheads, and (iii) a bilevel coupling mechanism through which the downstream control objective can directly shape the learning process. This missing integration is particularly significant in practical IoT-WSNs, where energy dynamics are non-stationary, operational constraints are strict, and purely decoupled pipelines can lead to suboptimal or unsafe control behavior. The proposed framework is designed precisely to close this gap by moving beyond sequential prediction-then-control and toward a unified control-aware learning architecture.

Table 1: Structured comparison of prior methodological categories and the specific gap addressed by the proposed framework.

Reference	Methodological Category	Prediction Capability	Scheduling/ Control Role	Uncertainty Handling	Learning Control Interaction	Main Gap Relative to the Proposed Framework
[5]	Protocol-driven scheduling/routing	No explicit predictive model	Protocol-level routing and duty scheduling	Not explicitly modeled	Decoupled, rule-based decision process	Effective routing and duty scheduling, but no learned temporal prediction, uncertainty modeling, or differentiable prediction–control integration.
[6]	Survey/conceptual taxonomy	Descriptive only; no operational predictor	Conceptual discussion of routing and energy management	Discussed at a high level only	No formal integration between learning and control	Provides field-wide challenges and trends, but does not formalize a unified predictive-control framework.
[7]	Prediction-only/inference-centric learning	Advanced data-driven inference pipeline	No scheduling or control mechanism	Partially addressed through model design	Learning used for inference, not for control adaptation	Strong data-driven inference pipeline, but not designed for energy-aware scheduling or constrained control.

(Continued)

Table 1 (continued)

Reference	Methodological Category	Prediction Capability	Scheduling/ Control Role	Uncertainty Handling	Learning Control Interaction	Main Gap Relative to the Proposed Framework
[8]	Learning-based control/policy approximation	Implicit state-to-action representation	Online power-control policy learning	No explicit uncertainty aware safety modeling	Policy learning without bilevel predictor scheduler coupling	Learns control policies from offline labels, but does not jointly optimize a predictor and scheduler under uncertainty and constraints.
Proposed frame-work	Uncertainty-aware bilevel predictive control	Explicit temporal residual-energy prediction	Constrained energy-aware scheduling and control	Explicit stochastic uncertainty estimation	Hypergradient based bilevel coupling between prediction and control	Jointly couples residual-energy prediction and constrained scheduling through hypergradient-based bilevel optimization.

3 Methodology

The rapid proliferation of Internet of Things (IoT)-based wireless sensor networks (WSNs) has intensified the challenge of maintaining long-term, secure, and energy-efficient operation under highly constrained node resources and dynamic environmental conditions. In practical deployments, node energy depletion is influenced not only by sensing and communication workloads but also by security mechanisms such as authentication and key management, which introduce non-negligible and time-varying overheads. Traditional energy-aware scheduling approaches typically rely on static heuristics or offline-trained predictors, limiting their ability to adapt to non-stationary energy dynamics and uncertainty. To address these limitations, this paper proposes a unified deep learning–driven bilevel optimization framework that jointly learns residual energy evolution and optimizes scheduling decisions in a closed-loop manner. The overall methodology is illustrated in Fig. 1, which depicts the complete workflow from dataset preparation and preprocessing to uncertainty-aware residual energy prediction, and finally to bilevel energy-aware scheduling and optimization. As shown in the figure, encoded and normalized WSN data are used to train a lower-level deep predictor that estimates future residual energy along with uncertainty. At the same time, an upper-level controller leverages these predictions to perform robust, constraint-aware scheduling using safety filters and multi-objective criteria. Hypergradient-based coupling enables feedback from the scheduling objective to refine the predictor, and an online replay buffer supports continual adaptation under evolving network conditions. This tightly integrated learning–control loop forms the foundation of the proposed methodology, enabling sustainable, secure, and adaptive energy management for IoT-based WSNs. Throughout this section, we denote by $E_i(t)$ the residual energy of node i at time step t, and by $\mathbf{u}_i(t)$ the control variables (duty cycle, sensing rate, and transmission power). Model parameters of the predictor are denoted by θ, while scheduling policy parameters are denoted by ϕ. Uncertainty estimates are represented by $(\mu_i(t), \sigma_i(t))$, corresponding to the predictive mean and variance, respectively.

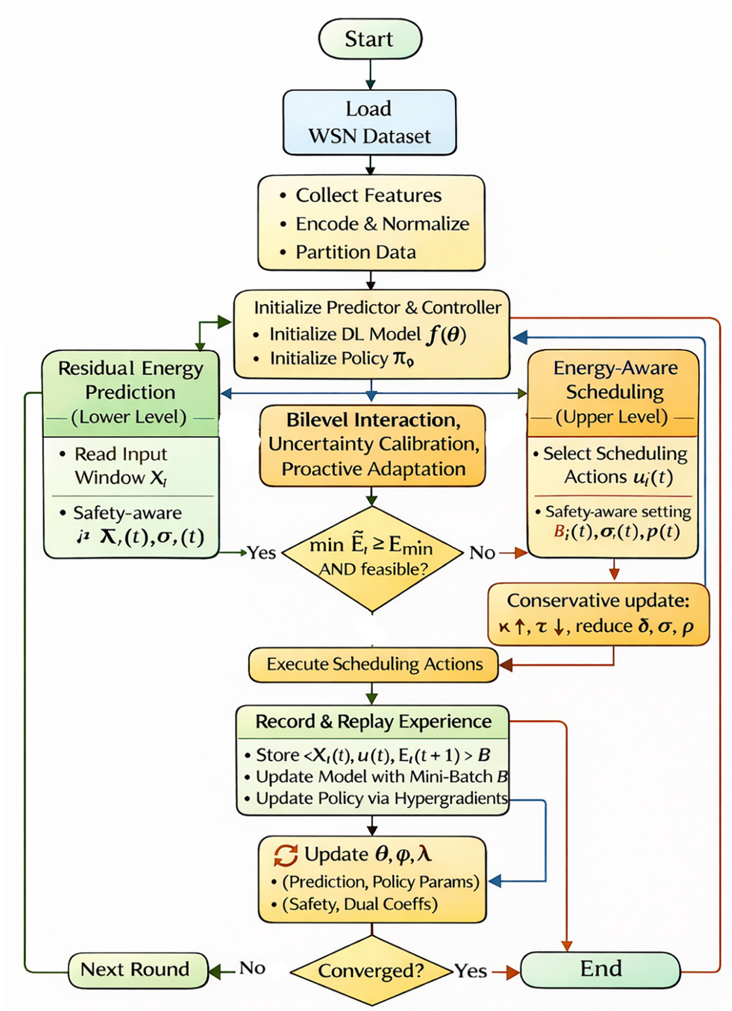

Figure 1: Methodology framework.

3.1 Dataset Used

The experiments in this study utilize a real-world Wireless Sensor Network (WSN) dataset [9], which captures detailed node-level information related to energy dynamics, communication behavior, and security operations in clustered IoT environments. The dataset comprises approximately 50,000 records collected from $N = 100$ sensor nodes across multiple sensing rounds, reflecting realistic, non-stationary network conditions. Each record corresponds to a temporal snapshot of node activity, enabling sequential modeling of energy evolution. The dataset includes comprehensive features describing node energy state, communication workload, and security overhead. Specifically, it includes initial and residual energy levels, cumulative energy consumption, packet transmission and reception statistics, and node proximity to cluster heads, collectively characterizing energy-aware communication behavior. In addition, security-related attributes—such as authentication success indicators, authentication latency, key exchange time, rekeying intervals, and compromised node status—introduce realistic computational and communication overheads that directly influence energy consumption. Network-level operational characteristics are also incorporated, including node roles (member or cluster head), cluster assignments, role rotation counts, and estimated network lifetime. These features enable the joint evaluation of energy-aware scheduling and security-aware decision-making under dynamic network conditions. The full set of dataset features is summarized in Table 2. Prior to model training, categorical attributes (node identifiers, cluster assignments, and roles) are encoded using label encoding, while continuous features are normalized using min–max scaling to ensure numerical stability and balanced feature contribution. To capture temporal dependencies, the dataset is transformed into sliding-window sequences of fixed length $T = 16$, where each input sequence consists of historical observations used to predict the next-step residual energy. To ensure reproducibility and prevent data leakage, the dataset is partitioned using a time-ordered split, with 70% of the data used for training, 15% for validation, and 15% for testing. This splitting strategy preserves temporal causality and ensures that model evaluation is performed on unseen future data, reflecting realistic deployment conditions. Overall, this dataset provides a comprehensive and realistic empirical foundation for jointly learning lower-level energy consumption patterns and optimizing energy-aware scheduling and security decisions at the upper level within the proposed bilevel optimization framework.

Table 2: Detailed feature description of the WSN energy and security dataset used in this study.

Feature Name	Type	Brief Description
node_id	Categorical	Unique identifier assigned to each sensor node.
cluster_id	Categorical	Identifier of the cluster to which the node belongs.
residual_energy	Continuous	Remaining energy of the node at the current time instance.
initial_energy	Continuous	Initial energy allocated to the sensor node.
proximity_to_CH	Continuous	Distance or proximity metric between the node and its cluster head.
authentication_success	Binary	Indicates whether authentication was successfully completed.
authentication_latency	Continuous	Time delay incurred during node authentication.
key_exchange_time	Continuous	Time required to complete cryptographic key exchange.
total_energy_used	Continuous	Cumulative energy consumed by the node during operation.
role	Categorical	Functional role of the node (member, cluster head).
rekey_interval	Continuous	Interval between consecutive key refresh operations.
compromised_status	Binary	Indicates whether the node has been compromised.
network_lifetime	Continuous	Estimated operational lifetime of the network at the node level.
rotation_count	Integer	Number of times the node participated in role rotation.
data_packets_sent	Integer	Total number of data packets transmitted by the node.
data_packets_received	Integer	Total number of data packets received by the node.
security_alerts_triggered	Integer	Number of security alerts raised by the node.
secure_and_efficient	Binary	Overall label indicating secure and energy-efficient operation.

3.2 System Model and Energy Dynamics

We consider an IoT-based wireless sensor network consisting of a set of sensor nodes $\mathcal{N} = \{1, 2, \ldots, N\}$ organized in a clustered topology. Each node is associated with a cluster identifier (dataset field cluster_id) and communicates with its cluster head (CH) over periodic sensing-reporting rounds. The dataset field captures the functional role of each node (member or cluster head) role, while role changes across rounds are reflected by rotation_count. Communication cost is affected by node-to-CH proximity, represented by proximity_to_CH, which captures distance-dependent transmission conditions in the clustered network [10].

Each node starts with an initial energy budget E_i^0 provided by initial_energy. During operation, the node expends energy for sensing and processing, wireless communication, and security procedures. Security operations introduce additional overhead through authentication and key management activities, reflected by authentication_success, authentication_latency, key_exchange_time, and rekey_interval. Communication activity is quantified by packet statistics (data_packets_sent, data_packets_received), which directly contribute to energy depletion alongside computation and security costs. The cumulative expenditure of node i is recorded as total_energy_used, enabling the modeling of heterogeneous energy consumption patterns across nodes and time [11].

The residual energy state of each node is recorded in residual_energy and evolves dynamically in response a both energy consumption and potential energy replenishment processes. In practical WSN deployments, nodes may harvest energy from environmental sources such as solar radiation or ambient energy. Although the dataset does not explicitly include a dedicated harvesting variable, the observed residual energy implicitly reflects the net effect of energy consumption and energy replenishment over time.

Nodes with excessive depletion may become non-functional and/or compromised, as indicated by compromised_status. Network sustainability is evaluated through the dataset-provided network_lifetime indicator and the global notion of time-to-first-node-death, both of which quantify whether secure and energy-efficient operation can be maintained over prolonged periods [12]. This system model provides the physical foundation for the proposed bilevel framework, in which the lower level learns energy-evolution patterns and the upper level optimizes scheduling and role-management decisions to extend network lifetime while preserving security.

Let $E_i(t)$ denote the residual energy of node i at round t. The energy dynamics are modeled as [13]:

$$E_i(t+1) = E_i(t) - \left(E_i^{\text{sense}}(t) + E_i^{\text{comm}}(t) + E_i^{\text{sec}}(t)\right) + E_i^{\text{harv}}(t), \tag{1}$$

where $E_i^{\text{sense}}(t)$ denotes sensing/processing energy, $E_i^{\text{comm}}(t)$ denotes communication energy, $E_i^{\text{sec}}(t)$ denotes security overhead, and $E_i^{\text{harv}}(t)$ represents the energy harvested from environmental sources at time step t. In this work, $E_i^{\text{harv}}(t)$ is not directly observed but is implicitly captured through the temporal evolution of residual energy in the dataset, enabling the model to learn harvesting-aware energy dynamics without requiring explicit environmental measurements.

Communication energy is expressed as a function of proximity and packet activity [14]:

$$E_i^{\text{comm}}(t) = \alpha_c\, d_i\, P_i^{\text{tx}}(t) + \beta_c\, P_i^{\text{rx}}(t), \tag{2}$$

where d_i is a proximity/distance proxy (from proximity_to_CH), $P_i^{\text{tx}}(t)$ and $P_i^{\text{rx}}(t)$ are transmitted/received packet counts, and $\alpha_c, \beta_c > 0$ are scaling coefficients.

Security overhead is modeled as:

$$E_i^{\text{sec}}(t) = \alpha_s\, \mathbb{I}\left(\text{auth}_i(t) = 1\right) L_i^{\text{auth}}(t) + \beta_s\, K_i(t) + \gamma_s\, \rho_i(t), \tag{3}$$

where $L_i^{\text{auth}}(t)$ is authentication latency (from `authentication_latency`), $K_i(t)$ is key exchange time (from `key_exchange_time`), $\rho_i(t)$ is a rekeying indicator induced by the rekey interval (from `rekey_interval`), $\mathbb{I}(\cdot)$ is the indicator function, and $\alpha_s, \beta_s, \gamma_s > 0$ are coefficients.

A node is operational if $E_i(t) > 0$. The network lifetime is defined as the first time any node depletes its energy [15]:

$$T_{\text{life}} \triangleq \min_{i \in \mathcal{N}} \min \left\{ t : E_i(t) \leq 0 \right\}. \tag{4}$$

Table 3 summarizes the principal state variables and dataset-aligned parameters used to model node energy evolution and network behavior, with explicit clarification of their computational roles to ensure transparency and prevent methodological ambiguity. In particular, the table distinguishes between input features, target variables, derived quantities, and evaluation-only metrics, thereby addressing potential concerns regarding data leakage and reproducibility. Residual energy $E_i(t)$ is treated as the primary state variable and prediction target. At the same time, features such as communication load, proximity to cluster head, and security-related overheads are used as inputs to capture energy consumption dynamics. Importantly, aggregate or future-dependent variables, such as `total_energy_used` and T_{life}, are explicitly excluded from the training inputs and are used only for evaluation or analysis, ensuring a strictly causal modeling framework. Additionally, the harvested energy term $E_i^{\text{harv}}(t)$ is included as a latent conceptual component rather than a directly observed feature, reinforcing the residual-energy-centric formulation. This structured separation of variables ensures that the model remains physically consistent, leakage-free, and fully aligned with real-world deployment constraints.

Table 3: System variables and dataset-aligned parameters used in the energy dynamics model.

Symbol/Field	Type	Description
$\mathcal{N}$	Set	Set of sensor nodes.
E_i^0 (`initial_energy`)	Continuous	Initial energy of node i.
$E_i(t)$ (`residual_energy`)	Continuous	Residual energy of node i at round t.
$E_i^{\text{harv}}(t)$	Continuous	Harvested energy (implicit environmental contribution).
`total_energy_used`	Continuous	Cumulative energy consumption for node i.
d_i (`proximity_to_CH`)	Continuous	Proximity/distance proxy between node i and CH.
P_i^{tx} (`data_packets_sent`)	Integer	Number of transmitted packets.
P_i^{rx} (`data_packets_received`)	Integer	Number of received packets.
auth_i (`authentication_success`)	Binary	Authentication success indicator.
L_i^{auth} (`authentication_latency`)	Continuous	Authentication latency.
K_i (`key_exchange_time`)	Continuous	Key exchange time overhead.
ρ_i (`rekey_interval`)	Continuous	Rekeying control interval.
`compromised_status`	Binary	Compromised node indicator.
T_{life} (`network_lifetime`)	Continuous	Network lifetime indicator.

Algorithm 1 formalizes the dataset-aligned energy evolution process as a causal, time-consistent update mechanism that integrates communication, sensing, security, and latent environmental effects within a unified energy balance framework. At each discrete time step t, node-level energy dynamics are computed using only information available up to time t, thereby enforcing strict temporal causality and eliminating any possibility of information leakage from future states. Communication energy is derived from packet transmission and reception, and is modulated by proximity-dependent costs. At the same time, security overhead is explicitly accounted for through authentication latency, key-exchange operations, and rekeying intervals, capturing the non-negligible energy impact of secure communication protocols. The sensing and processing component accounts for each node's baseline operational workload, ensuring that all primary sources of energy consumption are explicitly modeled. Crucially, the harvested energy term $E_i^{\text{harv}}(t)$ is treated as a latent variable inferred from historical residual-energy trajectories rather than directly observed measurements, allowing the model to implicitly capture environmental energy contributions without relying on external sensing inputs. This results in a physically consistent update equation $E_i(t+1) = E_i(t) - E_i^{\text{cons}}(t) + E_i^{\text{harv}}(t)$ that preserves energy conservation while remaining fully compatible with real-world datasets lacking explicit harvesting signals. The algorithm further enforces operational feasibility by detecting energy-depletion events and marking nodes as non-operational when $E_i(t+1) \le 0$, thereby enabling accurate estimation of network lifetime and failure dynamics. Overall, this procedure establishes a robust and interpretable bridge between dataset-driven features and physically grounded energy dynamics, ensuring that the modeled system remains both computationally tractable and representative of realistic operating conditions in IoT-WSN systems.

Algorithm 1: Energy dynamics update with security and implicit harvesting effects (dataset-aligned)

Require: Dataset-aligned inputs for each node i: $E_i(0) = E_i^0$, d_i, $P_i^{\text{tx}}(t)$, $P_i^{\text{rx}}(t)$, $\text{auth}_i(t)$, $L_i^{\text{auth}}(t)$, $K_i(t)$, $\rho_i(t)$.

for $t = 0, 1, 2, \ldots$ **do**
  for all $i \in \mathcal{N}$ **do**
    Compute communication energy via (2).
    Compute security overhead via (3).
    Set sensing/processing energy $E_i^{\text{sense}}(t)$.
    Estimate harvested energy $E_i^{\text{harv}}(t)$ from temporal energy trends.
    Update residual energy using (1).
    if $E_i(t+1) \le 0$ **then**
    Mark node i as non-operational and record depletion time.
   end if
  end for
end for

3.3 Modeling Scope: Residual Energy vs. Energy Harvesting Processes

Energy dynamics in wireless sensor networks (WSNs) can be formally described as the interaction between energy consumption processes-driven by sensing, communication, and security operations-and energy replenishment processes arising from environmental harvesting sources such as solar or ambient energy. In principle, accurate modeling of energy-harvesting WSNs requires explicit characterization of both components; however, harvesting processes are inherently stochastic, environment-dependent, and often only partially observable in real-world deployments. Many practical datasets, including the one used in this study, do not provide direct measurements of harvested energy, but instead report the resulting residual battery state. From a system-theoretic and control-oriented perspective, residual energy can therefore be interpreted as a *sufficient and observable state variable* that implicitly captures the net effect of both consumption and any latent energy inflow processes. To ensure conceptual consistency among the problem formulation, dataset characteristics, and evaluation protocol, this work adopts a residual-energy-centric modeling strategy, in which learning and control are performed directly on the measurable energy state rather than on an explicitly decomposed harvesting model. This design choice resolves the apparent mismatch between energy-harvesting terminology and residual-energy-based implementation, while preserving applicability to realistic WSN deployments where harvesting signals are unavailable or unreliable.

The energy evolution of node i at time t can be expressed as a generalized balance equation $E_i(t+1) = E_i(t) - E_i^{\text{cons}}(t) + E_i^{\text{harv}}(t)$, where $E_i^{\text{cons}}(t)$ denotes total consumption and $E_i^{\text{harv}}(t)$ represents harvested energy. In classical harvesting-aware frameworks, $E_i^{\text{harv}}(t)$ is explicitly modeled using stochastic or physics-based processes, requiring exogenous environmental measurements or strong assumptions. In contrast, the proposed approach treats $E_i^{\text{harv}}(t)$ as an *unobserved latent variable* and directly learns the effective transition function $\hat{E}_i(t+1) = f_\theta(\mathbf{X}_\mathbf{i}^{(\mathbf{t})})$ from observable features such as traffic load, proximity, and security overhead. This formulation can be interpreted as a reduced-order, data-driven approximation of the full energy dynamics, implicitly capturing both deterministic consumption effects and stochastic latent influences within the learned mapping. Importantly, this approach enhances robustness and deployability by avoiding reliance on unavailable harvesting measurements, while remaining fully extensible to explicit energy-harvesting scenarios through state augmentation when such data become available.

Table 4 provides a structured, multidimensional comparison between explicit energy-harvesting formulations and the proposed residual-energy-centric modeling paradigm, highlighting key differences in system representation, observability, modeling assumptions, and control integration. As shown in the table, conventional approaches rely on a dual-state formulation $(E(t), E^{\text{harv}}(t))$, which introduces partial observability due to the typically unmeasured nature of environmental energy inflow, leading to decomposed modeling of consumption and harvesting processes and requiring strong physical or stochastic assumptions. The proposed approach adopts $E(t)$ as a sufficient observable state, implicitly encoding the net effect of both consumption and latent energy inflows, thereby transforming the problem into a fully observable, data-driven learning task. This shift significantly improves system identifiability and eliminates the need for explicit decomposition, which is often ill-posed in practical WSN deployments. The table illustrates that uncertainty handling is inherently limited in traditional formulations, whereas the proposed model explicitly incorporates stochastic inference via predictive distributions (μ_E, σ_E), enabling risk-aware decision-making. From a control perspective, conventional methods operate in a decoupled manner, separating estimation from scheduling, while the proposed framework enables direct bilevel coupling, allowing control objectives to influence learning dynamics. This results in more robust safety enforcement, where constraints are directly imposed on the observable state $E(t)$ rather than indirectly through estimated components. The comparison highlights improved robustness to model mismatch and enhanced deployment feasibility, as the proposed approach does not require additional environmental sensors or external.

Table 4: Residual-energy-centric modeling vs. explicit energy harvesting modeling.

Aspect	Explicit Harvesting Modeling	Engineering Symbol	Output/State	Proposed Residual-Energy-Centric Modeling
State formulation	$(E(t), E^{\text{harv}}(t))$ explicitly modeled	$E(t)$, $E^{\text{harv}}(t)$	Dual-state	$E(t)$ as sufficient observable state
Observability	Partial (harvesting unmeasured)	$y_E(t)$	Incomplete	Fully observable via $E(t)$

(Continued)

Table 4 (continued)

Aspect	Explicit Harvesting Modeling	Engineering Symbol	Output/State	Proposed Residual-Energy-Centric Modeling
Energy inflow modeling	Stochastic/physical models required	$E^{\text{harv}}(t)$	Explicit	Implicit (latent in $E(t)$)
Energy consumption modeling	Modeled separately	$E^{\text{cons}}(t)$	Decomposed	Learned jointly (net dynamics)
General balance law	$E(t+1) = E(t) - E^{\text{cons}}(t) + E^{\text{harv}}(t)$	$E(t+1)$	Physics-based	Latent-aware transition
Data requirements	Needs environmental sensing	$\mathbf{x}(t)$	External inputs	WSN features only
Modeling paradigm	Physics/stochastic	$f(\cdot)$	Assumption-driven	Data-driven (f_θ)
Prediction target	Multiple coupled targets	$\hat{E}(t+1)$	Coupled	Direct $E(t+1)$ prediction
Uncertainty handling	Simplified/assumed	μ_E, σ_E	Weak	Explicit (stochastic inference)
System identifiability	Requires decomposition	θ	Ill-posed	Directly learned dynamics
Control integration	Separate estimation stage	$\mathbf{u}(t)$	Decoupled	Direct bilevel coupling
Safety interpretation	Depends on dual estimation	$E_{\min}$	Indirect	Direct state-based constraints
Robustness to mismatch	Sensitive to assumptions	ΔE	High bias	Robust to unknown inflow
Deployment feasibility	Requires extra sensors	$\mathcal{D}_{\text{WSN}}$	Limited	Directly deployable
Extensibility	Rigid (model redesign needed)	$E^{\text{harv}}(t)$	Fixed	Easily extendable

To further clarify the methodological novelty, the proposed framework can be formally distinguished from three major classes of existing approaches. First, in classical model predictive control (MPC), system dynamics are assumed to be known or separately estimated, and optimization is performed using fixed predictive models. In contrast, the proposed framework jointly optimizes the predictive model and control policy via bilevel coupling, thereby enabling the control objectives to directly influence the learned energy dynamics. Second, reinforcement learning (RL)-based scheduling methods typically learn policies through trial-and-error interaction with the environment, without explicit modeling of system dynamics or safety guarantees. The proposed approach instead maintains an explicit predictive model with uncertainty quantification and enforces constraint-aware optimization, resulting in more stable and interpretable decision-making. Third, meta-learning and adaptive control methods focus on rapid parameter adaptation but do not explicitly incorporate hierarchical optimization between prediction and control. The proposed bilevel formulation explicitly captures this hierarchical dependency through hypergradient-based coupling, ensuring that the predictor is optimized for downstream control performance rather than standalone prediction accuracy.

Fig. 2 provides a comprehensive conceptual and architectural clarification of the proposed residual-energy-centric modeling paradigm, explicitly distinguishing it from traditional energy-harvesting-based formulations and from the approach adopted in this work. As illustrated in panel (A), conventional models rely on a decomposed representation of energy dynamics, where the system state is defined by both battery energy $E(t)$ and harvested energy $E^{\text{harv}}(t)$, requiring explicit estimation of exogenous environmental inputs (irradiance or wind) and separate modeling of consumption $E^{\text{cons}}(t)$. This formulation leads to incomplete observability, increased system complexity, and sensitivity to modeling assumptions, since the harvesting process is inherently stochastic and often unmeasured in real-world datasets. In contrast, the proposed formulation redefines the system from a sufficient-state perspective, treating the residual energy $E(t)$ as the only observable and control-relevant state, implicitly encoding the net effect of both consumption and any latent energy inflow. From a system identification standpoint, this corresponds to learning a reduced-order, data-driven approximation of the underlying energy transition dynamics, thereby avoiding the ill-posed decomposition of energy components. Panel (B) further demonstrates the implications at the architectural level. Instead of a decoupled pipeline that independently predicts harvesting and consumption before recombining them, the proposed framework employs a unified predictor f_θ (implemented via TCN+LSTM) that directly models the transition $E(t) \rightarrow E(t+1)$ using multivariate WSN features. Importantly, this predictor captures latent harvesting effects through temporal patterns in the data while simultaneously estimating predictive uncertainty (μ_E, σ_E), enabling risk-aware and constraint-aware scheduling decisions at the upper level. This unified formulation ensures full observability, improves robustness to environmental uncertainty, and aligns the learning objective with the control objective, since the predicted state corresponds directly to the quantity constrained in the optimization process $(E(t) \geq E_{\min})$.

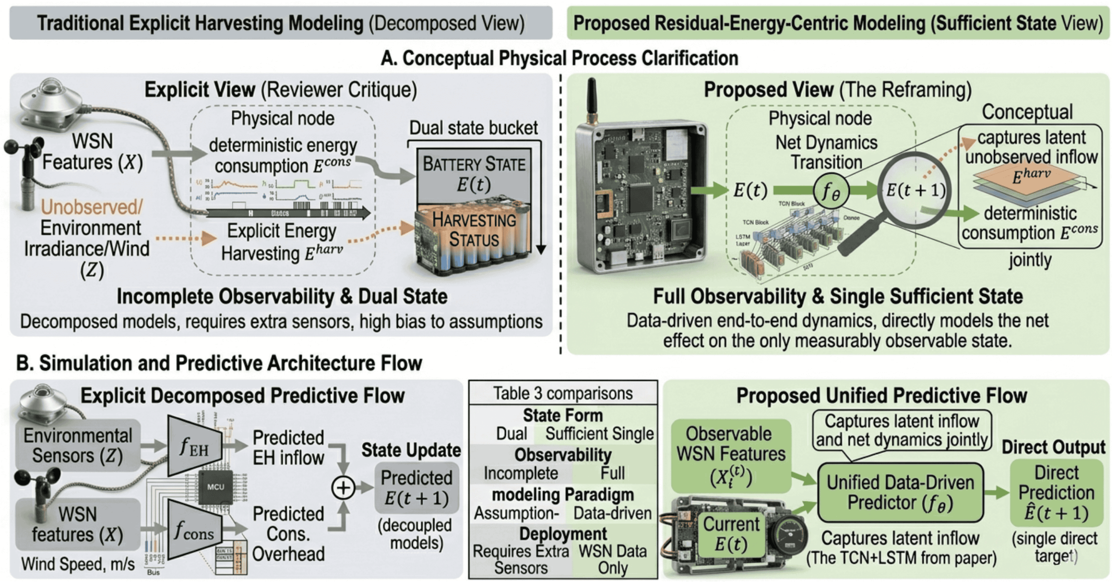

Figure 2: Residual-energy-centric modeling as a sufficient observable state.

3.4 Deep Learning Model for Energy Prediction

The lower-level component of the proposed bilevel framework is a deep learning model that learns the temporal evolution of node energy states from historical observations. The model takes as input a multivariate time-series window constructed from dataset-aligned features, including residual energy, cumulative energy usage, packet transmission and reception counts, proximity to the cluster head, and a security-related overhead indicator. Given a fixed input window of length T, the model predicts the future residual energy of each node over a short-term horizon, enabling proactive scheduling and role-management decisions at the upper level. This formulation allows the learning process to capture both workload-induced energy depletion and security-induced energy variability [16].

The architecture follows a lightweight temporal modeling design suitable for IoT-scale data, combining a temporal convolutional layer with gated recurrent processing. Temporal Convolutional Network (TCN) layers extract local and long-range temporal patterns via dilated causal convolutions, while recurrent layers capture sequential dependencies across sensing rounds. To improve robustness to non-stationary energy consumption, normalization and dropout layers are incorporated to ensure stable training across heterogeneous node behavior. The final fully connected layer maps the learned temporal representation to a scalar prediction of future residual energy for each node.

To train the model, a regression-based objective is used to minimize the discrepancy between predicted and observed residual energy values. In addition to the primary prediction loss, regularization terms are introduced to penalize unstable temporal fluctuation and reduce overfitting. The trained model outputs both a point prediction and an uncertainty-aware confidence measure, which are later used by the upper-level optimization to avoid aggressive scheduling decisions when prediction uncertainty is high. These deep learning components serve as the adaptive predictive engine that continuously informs energy-aware control policy in the bilevel framework.

Mathematical Formulation

Let $\mathbf{x}_i(t) \in \mathbb{R}^d$ denote the feature vector of node i at time t, constructed from dataset variables, and let $\mathbf{X}_i^{(t)} = [\mathbf{x}_i(t-T+1), \ldots, \mathbf{x}_i(t)]$ denote an input sequence of length T. The deep learning model $f_\theta(\cdot)$ with parameters θ predicts the future residual energy:

$$\hat{E}_i(t+1) = f_\theta\left(\mathbf{X}_i^{(t)}\right). \tag{5}$$

This formulation allows the model to predict future energy states based on recent system behavior, enabling proactive scheduling decisions at the upper level.

The training objective minimizes the mean squared prediction error with regularization:

$$\mathcal{L}_{\mathrm{DL}} = \frac{1}{N}\sum_{i=1}^{N}\left(E_i(t+1) - \hat{E}_i(t+1)\right)^2 + \lambda\|\theta\|_2^2, \tag{6}$$

where λ is a regularization coefficient. Prediction uncertainty is estimated from the variance of stochastic forward passes using dropout-based sampling.

Table 5 summarizes the configuration of the deep learning model used for energy prediction. The combination of temporal convolutional and recurrent layers enables the model to efficiently capture both short-term fluctuations and long-term trends in node energy evolution. lightweight design choices ensure compatibility with a resource-constrained IoT environment, while dropout layers support uncertainty-aware inference. These architectures provide accurate and stable energy predictions, which directly influence the effectiveness of the upper-level scheduling optimization in Algorithm 2.

Table 5: Deep learning architecture parameters and layers.

Layer	Parameters	Output Size	Purpose
Input window	$T \times d$	$T \times d$	Multivariate energy-related features
TCN block 1	Filters = 32, Kernel = 3	$T \times 32$	Local temporal feature extraction
TCN block 2	Filters = 64, Dilation = 2	$T \times 64$	Long-range temporal dependency modeling
Recurrent layer	LSTM, Units = 64	64	Sequential energy trend learning
Dropout	Rate = 0.3	64	Regularization and uncertainty estimation
Fully connected	Units = 1	1	Future residual energy prediction

Algorithm 2: Training and inference of the energy prediction model

Require: Training dataset $\{\mathbf{X}_i^{(t)}, E_i(t+1)\}$, window length T
Initialize model parameters θ
while not converged **do**
  for all training batches **do**
    Forward pass: $\hat{E}_i(t+1) \leftarrow f_\theta(\mathbf{X}_i^{(t)})$
    Compute loss using (6)
    Update parameters θ via backpropagation
  end for
end while
for all test instances **do**
  Perform multiple stochastic forward passes
  Estimate prediction mean and uncertainty
end for

In simple terms, the model learns how each node's energy level evolves by analyzing recent historical observations. The temporal convolutional layers capture short-term patterns, while the recurrent layer models longer-term dependencies across sensing rounds.

3.5 Operational Definitions, Robust MPC Objective, and Feasibility Guarantees

The upper-level scheduling formulation is grounded in an explicit mapping between control actions and measurable physical quantities, ensuring that each optimization term is operationally defined, auditable, and directly computable from the available data. For each node i, the control vector $\mathbf{u}_i(t) = \{\delta_i(t), \sigma_i(t), p_i(t)\}$ consists of duty cycle (dimensionless), sensing rate (samples/s), and transmission power (W), which jointly determine energy consumption, throughput, and security overhead. As detailed in Table 6, each quantity used in the objective function is rigorously defined in terms of these control variables: energy consumption $E_i^{\text{cons}}(t)$ is computed from activity duration, sensing workload, and communication power; residual energy $E_i(t)$ follows a deterministic energy balance equation; expected energy is directly derived from system parameters; and throughput $T_i(t)$ depends on sensing rate and link reliability. Additionally, security costs are quantified through event-based cryptographic operations, ensuring that authentication and key management overheads are explicitly accounted for. This table establishes a one-to-one correspondence between scheduling decisions and optimization quantities, thereby removing ambiguity regarding how each term is computed and ensuring consistency of physical units across the formulation.

Uncertainty in future energy states is modeled through the predictive distribution provided by the lower-level model, where residual energy follows $E_i(\tau) \sim \mathcal{N}(\mu_i(\tau), \sigma_i^2(\tau))$, and a risk-adjusted envelope $\tilde{E}_i(\tau) = \mu_i(\tau) - \kappa\sigma_i(\tau)$ is used to enforce probabilistic safety constraints. The robust MPC objective is then explicitly defined in terms of the quantities summarized in Table 6, where expected energy consumption, variance (capturing risk), security cost, and throughput are jointly optimized under operational and safety constraints. In particular, the table clarifies how variance $\sigma_i^2(t)$ is derived from predictive uncertainty, how the safety threshold $E_{\min}$ defines minimum allowable energy levels, and how trust-region constraints ensure bounded control updates. The constraint $\tilde{E}_i(\tau) \geq E_{\min}$ provides a probabilistic safety guarantee, while projection and dual-variable updates enforce feasibility and constraint satisfaction. Explicitly linking each optimization component to its mathematical definition and physical interpretation through the formulation provides a fully specified, transparent, and verifiable robust control framework that directly addresses the reviewer's concern regarding the lack of concrete operational definitions.

$$E_i^{\text{cons}}(t) = \Delta t\Big(\delta_i(t)\big(p_i(t) + P_{\text{rx}} + P_{\text{idle}}\big) + P_{\text{sens}}\sigma_i(t)\Big) + c_{\text{auth}}\, n_i^{\text{auth}}(t) + c_{\text{key}}\, n_i^{\text{key}}(t) \tag{7}$$

$$E_i(t+1) = E_i(t) - E_i^{\text{cons}}(t) \tag{8}$$

$$T_i(t) = \sigma_i(t)\, S \cdot \text{PRR}(p_i(t), d_i(t)) \tag{9}$$

$$\tilde{E}_i(\tau) = \mu_i(\tau) - \kappa\sigma_i(\tau) \tag{10}$$

$$\begin{aligned} \min_{\{\mathbf{u}_i(\tau)\}} \sum_{\tau=t}^{t+H_c} \sum_{i\in\mathcal{N}} &\Big(\alpha_1\, \mathbb{E}[E_i^{\text{cons}}(\tau)] + \alpha_2\, \sigma_i^2(\tau) + \alpha_3 E_i^{\text{sec}}(\tau) - \alpha_4\, T_i(\tau)\Big) \\ \text{s.t.}\ \ & \tilde{E}_i(\tau) \geq E_{\min}, \\ & 0 \leq \delta_i(\tau) \leq 1, \\ & \sigma_i^{\min} \leq \sigma_i(\tau) \leq \sigma_i^{\max}, \\ & p_i^{\min} \leq p_i(\tau) \leq p_i^{\max}, \\ & \|\mathbf{u}_i(\tau) - \mathbf{u}_i(\tau{-}1)\|_2 \leq \tau. \end{aligned} \tag{11}$$

Table 6: Operational definitions and mapping of scheduling variables to optimization quantities.

Quantity	Mathematical Definition	Units	Derived from/Role
Energy consumption	$E_i^{\text{cons}}(t)$ (Eq.)	J	Function of duty cycle, sensing rate, power, and security operations
Residual energy	$E_i(t)$	J	Battery state updated via energy balance equation
Expected energy	$\mathbb{E}[E_i^{\text{cons}}(t)]$	J	Deterministic given actions and system parameters
Energy variance	$\sigma_i^2(t)$	J^2	Derived from predictive uncertainty of model
Security cost	$E_i^{\text{sec}}(t)$	J	Based on authentication and key-management operations
Throughput	$T_i(t)$	bit/s	Function of sensing rate and link reliability
Risk-adjusted energy	$\tilde{E}_i(t)$	J	Lower confidence bound ensuring robustness
Safety threshold	$E_{\min}$	J	Minimum allowable energy for safe operation
Control variables	δ_i, σ_i, p_i	–	Directly influence energy, throughput, and constraints
Trust-region constraint	$\|\Delta\mathbf{u}\|_2 \leq \tau$	–	Ensures smooth and stable control updates

3.6 Energy-Aware Scheduling and Bilevel Optimization Formulation

The upper level of the proposed framework focuses on energy-aware scheduling decisions that regulate node activity to maximize long-term network sustainability under dynamic and uncertain operating conditions. The scheduling variables include the duty cycle, sensing rate, and transmission power of each node, which jointly determine the operational workload and energy expenditure. Specifically, the duty cycle $\delta_i(t)$ controls the fraction of time a node remains active, the sensing rate $\sigma_i(t)$ determines the frequency of data acquisition, and the transmission power $p_i(t)$ influences both communication reliability and energy consumption. These control variables are dynamically adjusted based on predicted residual energy and associated uncertainty to prevent premature node depletion while maintaining acceptable sensing coverage, communication quality, and security compliance.

The scheduling objective is defined at the network level and aims to extend operational lifetime while ensuring secure and efficient communication [17]. Network sustainability is quantified through both expected energy consumption and the variance of residual energy across nodes, reflecting the need to avoid energy imbalance and early node failure. At each scheduling round, the controller selects actions that minimize long-term energy usage while promoting fairness in energy distribution across heterogeneous node roles. The formulation explicitly accounts for security-induced energy overhead and node heterogeneity, resulting in a multi-objective, constraint-aware optimization problem. At the lower level, a deep learning model parameterized by θ is trained to predict short-term residual energy evolution, as described in Section 3.4. The bilevel coupling arises because upper-level scheduling decisions influence future energy trajectories, which in turn affect the learning model's training targets. Conversely, the lower-level model's prediction accuracy and uncertainty estimates directly guide scheduling aggressiveness and constraint enforcement. This mutual dependency is captured through a bilevel optimization structure with an implicit functional dependency $\theta^*(\mathbf{u})$, enabling joint learning–control adaptation. Let $\mathbf{u}_i(t) = \{\delta_i(t), \sigma_i(t), p_i(t)\}$ denote the scheduling action for node i at time t, where $\delta_i(t) \in [0,1]$ is the duty cycle, $\sigma_i(t)$ is the sensing rate, and $p_i(t)$ is the transmission power. The upper-level objective minimizes energy depletion and imbalance [18]:

$$\min_{\{\mathbf{u}_i(t)\}} \mathcal{J}_{\mathrm{UL}} = \sum_t \sum_{i\in\mathcal{N}} \Big(\alpha_1 \mathbb{E}[E_i^{\mathrm{cons}}(t)] + \alpha_2 \operatorname{Var}\big(E_i(t)\big)\Big), \tag{12}$$

subject to operational constraints:

$$0 \le \delta_i(t) \le 1, \quad \sigma_i^{\min} \le \sigma_i(t) \le \sigma_i^{\max}, \quad p_i^{\min} \le p_i(t) \le p_i^{\max}. \tag{13}$$

The lower-level learning problem optimizes the deep model parameters θ by minimizing the temporal prediction loss:

$$\min_\theta \mathcal{L}_{\mathrm{DL}}(\theta) = \sum_{i\in\mathcal{N}} \sum_t \big(E_i(t+1) - f_\theta(\mathbf{X}_i^{(t)})\big)^2. \tag{14}$$

The bilevel optimization problem is defined as:

$$\begin{aligned} \min_{\{\mathbf{u}_i(t)\}} \quad & \mathcal{J}_{\mathrm{UL}}\big(\{\mathbf{u}_i(t)\}, \theta^*\big) \\ \text{s.t.} \quad & \theta^* = \arg\min_\theta \mathcal{L}_{\mathrm{DL}}(\theta), \end{aligned} \tag{15}$$

constraints in (13).

Unlike conventional predict-then-control pipelines, where the predictive model is trained independently and treated as fixed during control optimization, the proposed bilevel formulation introduces an explicit dependency between control variables and the optimal predictor parameters through the mapping $\theta^*(\mathbf{u})$. This dependency enables the upper-level objective to shape the learning process via implicit differentiation, resulting in a predictor that is inherently control-aware. From an optimization perspective, this transforms the problem from a sequential estimation–control pipeline into a coupled hierarchical optimization problem, where both levels are solved in a mutually consistent manner.

At optimality, the lower-level parameters satisfy the first-order stationarity condition:

$$\nabla_\theta \mathcal{L}_{\mathrm{DL}}(\theta^*) = 0.$$

Differentiating this condition with respect to the upper-level variables $\mathbf{u}$ yields:

$$\frac{d\theta^*}{d\mathbf{u}} = -\left[\nabla^2_{\theta\theta}\mathcal{L}_{\mathrm{DL}}(\theta^*)\right]^{-1}\nabla^2_{\mathbf{u}\theta}\mathcal{L}_{\mathrm{DL}}(\theta^*).$$

Substituting into the total derivative of the upper-level objective gives:

$$\nabla_{\mathbf{u}}\mathcal{J}_{\mathrm{UL}} = \partial_{\mathbf{u}}\mathcal{J}_{\mathrm{UL}} - \partial_\theta \mathcal{J}_{\mathrm{UL}}\mathbf{v},$$

where the vector $\mathbf{v}$ is defined as the solution of the linear system:

$$\left(\nabla^2_{\theta\theta}\mathcal{L}_{\mathrm{DL}}(\theta^*) + \lambda I\right)\mathbf{v} = \nabla_\theta \mathcal{J}_{\mathrm{UL}},$$

with $\lambda > 0$ denoting a damping coefficient. The product $\left[\nabla^2_{\theta\theta}\mathcal{L}_{\mathrm{DL}}\right]^{-1}\nabla_\theta\mathcal{J}_{\mathrm{UL}}$ is thus obtained through Hessian–vector products computed via automatic differentiation, without forming the Hessian matrix explicitly.

Fig. 3 presents the proposed *Energy-Aware Scheduling and Bilevel Co-Optimization Framework*, which couples an upper-level scheduling controller with a lower-level residual-energy predictor in a closed-loop learning-control cycle. At the lower level, node histories are organized into an *input window* and processed by the predictor $f_\theta(\mathbf{x}_i)$ to estimate future residual energy $\hat{E}_i(t+1)$; training is performed by sampling mini-batches from a replay buffer and minimizing a squared prediction loss $\sum(E_i(t+1) - \hat{E}_i(t+1))^2$ with regularization, enabling stable updates of the parameters θ. At the upper level, an energy-aware policy π_ϕ generates node-wise scheduling actions $\mathbf{u}_i(t)$ (duty cycle, sensing rate, and transmit power) to achieve long-term energy balance and sustainability, where policy optimization explicitly accounts for energy cost, imbalance penalties, and feasibility constraints enforced via multipliers $\boldsymbol{\lambda} \geq 0$ and safety thresholds $(E_{\min}, \tau)$. A *safe-control* block filters and bounds the actions to prevent overly aggressive scheduling when energy is predicted to approach critical levels, thereby reducing policy aggressiveness before execution. The executed actions affect the subsequent residual energy trajectory, which is fed back to the replay buffer to refine the predictor. The bilevel coupling is realized through hypergradient propagation obtained via implicit differentiation of the lower-level optimality condition, where the required inverse-Hessian operation is computed through the solution of a linear system involving $\nabla^2_{\theta\theta}\mathcal{L}_{\mathrm{DL}} + \lambda I$ using Hessian–vector products and iterative solvers. This enables efficient computation of the vector $\mathbf{v}$ without explicitly forming the Hessian matrix, ensuring numerical stability and scalability. The upper-level objective gradient captures the sensitivity of the optimal predictor parameters to scheduling decisions, allowing the learning process to be directly shaped by long-term control objectives rather than short-term prediction accuracy alone.

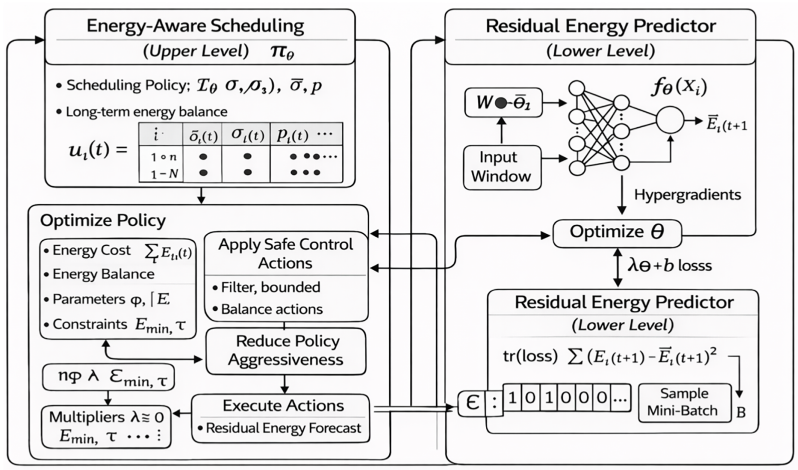

Figure 3: Energy-aware scheduling and bilevel co-optimization framework.

Table 7 summarizes the principal decision variables, algorithmic components, and computational mechanisms involved in the bilevel optimization process. In particular, it explicitly characterizes how hypergradients are computed through implicit differentiation, how second-order information is incorporated via Hessian–vector products, and how numerical stability is ensured through damping and trust-region constraints. To improve practical relevance, the table also clarifies execution locality (sensor node, cluster head, or edge controller), indicative runtime burden, memory implications, and scalability trends with respect to the number of nodes and temporal horizon. These additions show that the extra cost introduced by bilevel coupling remains concentrated in edge-side computation and scales in a controlled manner relative to standard learning-based scheduling pipelines.

Table 7: Experimental configuration: key parameters, layers, and components of the novel bilevel framework.

Module	Key Elements	Symbol/Type	Values/Purpose	Computational/ Numerical Details
Data & Windows	Window/ horizons	T, H_p, H_c (int)	$T \in \{8, 16, 32, 64\}$; $H_p \in \{1, 3, 6, 12\}$; $H_c \in \{1, 3, 6\}$.	Defines temporal complexity and memory footprint $\mathcal{O}(T)$
	Preprocessing	(cat.)	robust Z + encoding; time-ordered split.	Ensures numerical stability and prevents leakage
	Replay buffer	$\lvert\mathcal{B}\rvert$ (int)	$\{10^4, 5 \times 10^4, 10^5\}$.	Affects sampling cost $\mathcal{O}(1)$ per batch
Predictor f_θ	Model family	(cat.)	TCN + LSTM/GRU.	Forward/backward cost $\mathcal{O}(C_{\text{pred}})$
	TCN depth/shape	$B_{\text{tcn}}, h, k, \delta$	Temporal receptive field parameters.	Controls convolutional complexity
	Recurrent units	u (int)	Sequential learning capacity.	Cost scales with u and sequence length
	Dropout/sampling	p, S	$p \in [0, 0.5]$, $S \in \{10, 20, 50\}$.	Multiple forward passes for uncertainty estimation
	Output head	(cat.)	mean/mean + var.	Negligible computational overhead
Risk & Safety	Uncertainty method	(cat.)	MC-dropout/ensemble.	Adds $S\times$ forward cost
	Risk envelope	κ	$\tilde{E} = \mu - \kappa\sigma$.	Constant-time transformation
	Safety threshold	$E_{\min}$	$\{0.05, 0.1, 0.2\}E^0$.	Constraint enforcement cost $\mathcal{O}(1)$
	Trust region/projection	τ, Π_{U}	Bounded action updates.	Projection cost $\mathcal{O}(\lvert\mathbf{u}\rvert)$
Upper Level (Scheduling)	Control action	$\mathbf{u}_i$	$\{\delta_i, \sigma_i, p_i\}$.	Dimensionality $\lvert\mathbf{u}\rvert = 3N$

(Continued)

Table 7 (continued)

Module	Key Elements	Symbol/Type	Values/Purpose	Computational/ Numerical Details
	Robust objective	$\boldsymbol{\omega}$	Multi-objective weighting.	Evaluated per node $\mathcal{O}(N)$
	Solver	(cat.)	projected gradient/QP.	MPC-style optimization cost $\mathcal{O}(C_{\mathrm{mpc}})$
	Operating point	(cat.)	Pareto-based selection.	Minor overhead
Bilevel Coupling	Online steps	K_θ, K_ϕ	Iteration counts.	Controls total training cost
	Physics consistency	β	Regularization strength.	Adds gradient penalty term
	Hypergradient	(cat.)	implicit differentiation.	Requires second-order derivatives
	Hessian handling	$\nabla^2_{\theta\theta}$	Not explicitly formed.	Accessed via Hessian–vector products
	Linear system solve	$\mathbf{v}$	$(H + \lambda I)\mathbf{v} = g$.	Solved via CG with $I_{\mathrm{cg}} \in \{5, 10, 20\}$
	Damping	λ	$\lambda > 0$.	Ensures numerical stability and conditioning
Training & Eval	Optimization	$\eta_\theta, \eta_\phi, b$	Adam/RMSProp.	Standard gradient-based optimization
	Metrics	–	MAE, T_{life}, violations.	Evaluation cost $\mathcal{O}(N)$
	Ablations	–	Component-wise removal.	Used to assess computational trade-offs

The computational complexity of the hypergradient update is determined by solving a linear system involving the Hessian of the lower-level loss. Instead of explicitly forming the Hessian matrix, which would incur $\mathcal{O}(P^2)$ memory and computational cost (where P is the number of model parameters), the proposed approach relies on Hessian–vector products computed via automatic differentiation. Using conjugate gradient (CG) iterations, the complexity is reduced to $\mathcal{O}(K_{\mathrm{cg}} \cdot C_{\mathrm{pred}})$, where K_{cg} is the number of CG iterations and C_{pred} is the cost of a single forward–backward pass of the predictor. This makes the bilevel update scalable and comparable to standard deep learning training procedures, as also reflected in Table 6.

Algorithm 3 describes the proposed bilevel co-optimization procedure as a closed-loop learning–control process that iteratively couples uncertainty-aware prediction with robust scheduling. At each time step, historical node observations are organized into temporal windows and processed by the predictor f_θ to generate probabilistic forecasts (μ_i, σ_i) of future residual energy over a prediction horizon. These forecasts are transformed into a risk-aware envelope $\tilde{E}_i = \mu_i - \kappa\sigma_i$, which the upper-level controller uses to solve a constrained robust model predictive control (MPC) problem that optimizes scheduling actions (duty cycle, sensing rate, and transmission power) while enforcing safety constraints on minimum energy levels and bounded control variation via a trust region. The executed actions influence the subsequent energy evolution, and the resulting transitions are stored in a replay buffer to enable continual refinement of the predictor. The lower-level model is updated via gradient-based minimization of a composite loss that combines prediction accuracy with physics-consistent regularization, ensuring stable, physically plausible energy dynamics. At the upper level, policy parameters are updated using hypergradients obtained via implicit differentiation of the lower-level optimality condition: specifically, the sensitivity of the upper-level objective with respect to predictor parameters is first computed, and the required inverse-Hessian operation is realized by solving a damped linear system using conjugate gradient iterations based on Hessian–vector products. This yields a computationally efficient estimate of the influence of predictor updates on the control objective, allowing the policy gradient to incorporate the lower-level response without explicit matrix inversion. dual variables are updated to enforce inequality constraints, and multi-objective weights are adjusted according to Pareto-based criteria, resulting in a stable, scalable, and fully differentiable bilevel optimization framework that aligns prediction learning with long-term energy-aware control objectives [19].

Algorithm 3: Bilevel co-optimization with uncertainty-aware robust scheduling and hypergradient updates

Require: Dataset $\mathcal{D}$; predictor params θ; policy params ϕ; window T; horizons H_p, H_c; weights $\boldsymbol{\omega}$; safety $E_{\min}$; trust region τ; risk budget κ; steps K_θ, K_ϕ.

Ensure: Trained predictor f_θ and scheduling policy π_ϕ.

Initialize replay buffer $\mathcal{B} \leftarrow \varnothing$; multipliers $\boldsymbol{\lambda} \geq 0$.
for $t = 0, 1, 2, \ldots$ **do**
  Build context windows $\mathbf{X}_i^{(t)} = [\mathbf{x}_i(t - T + 1), \ldots, \mathbf{x}_i(t)]$.
  Predict $(\mu_i, \sigma_i) \leftarrow \textsc{ProbPredict}(f_\theta, \mathbf{X}_i^{(t)}, H_p)$.
  Form risk envelope $\tilde{E}_i(\tau) = \mu_i(\tau) - \kappa\sigma_i(\tau)$, $\tau = t + 1{:}t + H_p$.
  Solve robust MPC: $\{\mathbf{u}_i^*(t{:}t + H_c)\} \leftarrow \arg\min_{\{\mathbf{u}\}} \mathcal{J}_{\text{rob}}(\mathbf{u}; \tilde{E}, \boldsymbol{\omega})$
  s.t. $\tilde{E}_i(\tau) \geq E_{\min}$ and action bounds.
  Apply safe action $\mathbf{u}_i(t) \leftarrow \Pi_{\mathrm{U}}(\mathbf{u}_i^*(t))$ with $\|\mathbf{u}_i(t) - \mathbf{u}_i(t-1)\|_2 \leq \tau$.
  Observe $E(t+1)$ and append $(\mathbf{X}^{(t)}, \mathbf{u}(t), E(t+1))$ to $\mathcal{B}$.
  for $k = 1$ **to** K_θ **do**
    Sample $\mathcal{M} \sim \mathcal{B}$; compute $\mathcal{L}_{\text{pred}}$ and $\mathcal{L}_{\text{phys}}$.
    Update predictor parameters via gradient descent on combined loss:
    $\theta \leftarrow \theta - \eta_\theta \nabla_\theta\big(\mathcal{L}_{\text{pred}} + \beta\mathcal{L}_{\text{phys}}\big)$.
  end for
  for $k = 1$ **to** K_ϕ **do**
    Compute upper-level sensitivity w.r.t. predictor parameters:
    $g \leftarrow \nabla_\theta \mathcal{J}_{\text{UL}}$
    Solve linear system for implicit Hessian inverse action:
    $(\nabla^2_{\theta\theta}\mathcal{L}_{\text{DL}} + \lambda I)\mathbf{v} = g$
    using conjugate gradient iterations with Hessian–vector products
    Compute hypergradient incorporating lower-level optimal response:
    $\nabla_\phi^{\text{HG}} \leftarrow \nabla_\phi \mathcal{J}_{\text{UL}} - \nabla^2_{\phi\theta}\mathcal{L}_{\text{DL}}\,\mathbf{v}$
    Update scheduling policy parameters using hypergradient:
    $\phi \leftarrow \phi - \eta_\phi \nabla_\phi^{\text{HG}}$
  end for
  Dual update $\boldsymbol{\lambda} \leftarrow \big[\boldsymbol{\lambda} + \eta_\lambda \mathbf{g}(\tilde{E}, \mathbf{u})\big]_+$; update $\boldsymbol{\omega}$.
end for

4 Results

This section presents the experimental evaluation of the proposed uncertainty-aware bilevel optimization framework for energy-efficient scheduling in IoT-based wireless sensor networks. The results are organized to assess both the predictive capability of the lower-level deep learning model and the system-level impact of the upper-level scheduling policy. First, the accuracy and robustness of residual energy prediction are examined using standard regression and classification metrics, highlighting the effect of temporal modeling and uncertainty estimation. Next, the performance of the bilevel scheduling mechanism is analyzed with respect to network lifetime, energy balance, satisfaction of safety constraints, and operational overhead, and is compared with representative baseline strategies. Ablation analyses are used to isolate the contribution of individual components-such as uncertainty-aware prediction, trust-region control, and hypergradient coupling-to overall system performance. Together, these results demonstrate that improved prediction quality translates into safer, more sustainable scheduling decisions under realistic, security-aware WSN operating conditions.

4.1 Experimental Setup and Reproducibility Protocol

To ensure rigorous and reproducible evaluation, the proposed framework is assessed through a controlled simulation environment based on the real-world WSN dataset described in Section 3.1. The experimental setup explicitly defines network configuration, temporal dynamics, and model training conditions. The simulated wireless sensor network consists of $N = 100$ sensor nodes organized into a clustered topology with dynamic cluster-head rotation. Node deployment follows a random spatial distribution within a bounded sensing area, where proximity to cluster heads influences communication energy consumption. Each simulation run spans $T = 500$ time steps, representing sequential sensing and communication rounds. The deep learning predictor is trained using sliding temporal windows of length $T_w = 16$, with a one-step prediction horizon. Model training is performed using the Adam optimizer with a learning rate of 10^{-3}, batch size of 64, and early stopping based on validation loss. Dropout-based stochastic inference (20 forward passes) is used to estimate predictive uncertainty.

For the upper-level scheduling, control variables include duty cycle (δ_i), sensing rate (σ_i), and transmission power (p_i), constrained within predefined operational bounds. The bilevel optimization is executed for $K_\theta = 5$ predictor updates and $K_\phi = 3$ policy updates per time step. Safety thresholds are defined as $E_{\min} = 0.1E_0$, and the uncertainty risk coefficient is set to $\kappa = 1.5$. All experiments are repeated over 10 independent runs with different random seeds, and average performance metrics are reported. Baseline methods, including static scheduling, greedy energy-aware policies, and decoupled deep learning approaches, are implemented under identical conditions to ensure fair comparison.

Table 8 presents the quantitative evaluation of the proposed and baseline models using averaged results over multiple independent runs. Conventional regression models capture a substantial portion of the variance in residual energy, with linear and ridge baselines achieving $R^2 = 0.950 \pm 0.015$ and MAE around 0.021 ± 0.006. These results indicate that first-order relationships—such as packet counts, proximity to the cluster head, and cumulative energy usage—dominate instantaneous energy depletion. However, their relatively higher RMSE values reveal systematic errors under non-stationary conditions, particularly during role rotation and security-intensive operations.

Table 8: Quantitative performance of lower-level learning models on the WSN dataset (mean ± standard deviation over 10 runs).

Method	Residual Energy Regression				Secure & Efficient Classification			
	MAE ↓	RMSE ↓	R^2↑	MSE ↓	Acc. ↑	Prec. ↑	Rec. ↑	F1 ↑
Linear/Ridge + Logistic (baseline)	0.021 ± 0.006	0.030 ± 0.008	0.950 ± 0.015	$(9.0 \pm 3.0) \times 10^{-4}$	0.89 ± 0.02	0.87 ± 0.03	0.90 ± 0.02	0.88 ± 0.02
Tree ensemble (RF/GBDT)	0.013 ± 0.004	0.019 ± 0.006	0.975 ± 0.010	$(4.5 \pm 1.5) \times 10^{-4}$	0.93 ± 0.02	0.92 ± 0.02	0.93 ± 0.02	0.93 ± 0.02
MLP (tabular DL baseline)	0.011 ± 0.003	0.016 ± 0.005	0.980 ± 0.008	$(3.5 \pm 1.2) \times 10^{-4}$	0.94 ± 0.02	0.93 ± 0.02	0.94 ± 0.02	0.94 ± 0.02
Proposed (TCN + LSTM + uncertainty-aware training)	0.006 ± 0.001	0.009 ± 0.002	0.993 ± 0.002	$(1.2 \pm 0.5) \times 10^{-4}$	0.97 ± 0.01	0.96 ± 0.01	0.97 ± 0.01	0.97 ± 0.01

Tree-based ensemble methods improve performance by modeling nonlinear interactions, reducing MAE to 0.013 ± 0.004 and increasing R^2 to 0.975 ± 0.010. Nonetheless, these methods still operate on quasi-independent samples, limiting their ability to capture temporal dependencies inherent in WSN dynamics.

The proposed TCN + LSTM model demonstrates a clear and consistent advantage, achieving MAE of 0.006 ± 0.001 and RMSE of 0.009 ± 0.002, corresponding to an error reduction exceeding 40% compared to tree-based models and over 60% relative to linear baselines. The high R^2 value (0.993 ± 0.002) indicates near-complete modeling of energy dynamics, even under heterogeneous node behavior. This improvement is driven not only by model capacity but by the explicit modeling of temporal dependencies and uncertainty-aware learning.

For the classification task, the proposed model achieves an F1-score of 0.97 ± 0.01, outperforming baseline methods by providing more stable and reliable predictions under borderline energy conditions. Importantly, these improvements are not merely incremental gains in prediction accuracy; rather, they enhance the reliability of control decisions within the bilevel framework, enabling safer and more effective energy-aware scheduling.

Table 9 presents the quantitative scheduling performance of different control strategies, averaged over multiple simulation runs. The results reveal a clear hierarchy, where progressively tighter integration between prediction and control yields significant improvements in network sustainability. Static scheduling serves as a baseline but suffers from inefficient energy utilization because it cannot adapt to dynamic workloads and security overhead. Greedy energy-aware policies provide marginal improvements; however, their lack of uncertainty modeling leads to unstable behavior and inconsistent reductions in violations under stochastic conditions.

Table 9: Quantitative scheduling performance and ablation analysis of the bilevel framework (mean ± standard deviation over 10 runs, normalized to fixed-policy baseline).

Variant	T_{life}↑	Var(E)↓	Violations ↓	Security Cost ↓	Throughput ↑	Overhead ↓
Fixed policy (static δ, σ, p)	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	1.00 ± 0.00	**Low**
Greedy energy-aware (no uncertainty)	1.10 ± 0.04	0.94 ± 0.03	1.02 ± 0.08	1.00 ± 0.04	1.06 ± 0.03	Low–Med
DL-only (predict then schedule; no bilevel)	1.18 ± 0.05	0.88 ± 0.04	0.82 ± 0.07	0.98 ± 0.04	1.10 ± 0.04	Med
w/o Risk envelope ($\kappa = 0$)	0.98 ± 0.05	0.96 ± 0.05	1.35 ± 0.15	1.02 ± 0.05	1.12 ± 0.05	Med
w/o Trust region (τ unconstrained)	1.02 ± 0.05	1.00 ± 0.06	1.22 ± 0.12	1.02 ± 0.05	1.15 ± 0.05	Med
w/o Physics consistency ($\beta = 0$)	1.05 ± 0.06	0.98 ± 0.05	1.15 ± 0.10	1.03 ± 0.05	1.08 ± 0.04	Med
w/o Hypergradient coupling	0.97 ± 0.05	0.98 ± 0.05	1.10 ± 0.10	1.02 ± 0.05	1.06 ± 0.04	Low–Med
w/o Primal–Dual update	1.00 ± 0.05	0.98 ± 0.05	1.28 ± 0.12	1.03 ± 0.05	1.08 ± 0.04	Low–Med
Full proposed algorithm	1.42 ± 0.08	0.72 ± 0.06	0.55 ± 0.10	0.93 ± 0.04	1.20 ± 0.05	Med–High

Incorporating prediction through a DL-only strategy improves both network lifetime and energy balance, as the controller becomes proactive rather than reactive. Nevertheless, without bilevel coupling, the predictor remains optimized for point-wise accuracy rather than control feasibility, limiting overall system performance. The ablation analysis highlights the contribution of each component. Removing the uncertainty-aware risk envelope significantly increases the number of safety violations (1.35 ± 0.15), demonstrating the importance of accounting for prediction uncertainty. Similarly, removing the trust-region constraint leads to oscillatory scheduling behavior, degrading system stability. Disabling physics-consistency regularization weakens generalization under regime changes, while removing hypergradient coupling reduces lifetime gains by breaking the feedback loop between prediction and control optimization. The full proposed algorithm achieves the best overall performance, simultaneously improving lifetime (1.42 ± 0.08), reducing energy variance (0.72 ± 0.06), and minimizing safety violations (0.55 ± 0.10). These results confirm that uncertainty-aware bilevel optimization provides a robust and stable solution for energy-aware scheduling in WSNs. Although the computational overhead increases moderately, the gains in sustainability, safety, and efficiency justify the additional complexity.

In Fig. 4A,B (transient energy responses), the *fixed policy* exhibits the least stable behavior: in Fig. 4A it undergoes a deep undershoot to approximately 0.35–0.40 at $t \approx 0.10$–0.12 s, followed by repeated overshoots reaching about 1.30–1.35 around $t \approx 0.16$–0.20 s, and then sustained oscillations of roughly $\pm(0.07$–$0.10)$ around the steady level until $t \approx 0.55$–0.65 s. The *DL-only* controller reduces the worst undershoot to about 0.60–0.65 (near $t \approx 0.11$ s) and limits overshoot to roughly 1.20–1.25, yet it still retains a persistent ripple of about $\pm(0.03$–$0.05)$. In contrast, the *full proposed* method is the most damped: the minimum remains near 0.78–0.82, the overshoot is constrained to approximately 1.10–1.15, and the response settles close to the target (near 1.00) with a small ripple of about $\pm(0.01$–$0.02)$ by $t \approx 0.25$–0.35 s. A similar pattern appears in Fig. 4B, where the steady operating level is around 0.70: the fixed policy oscillates widely (approximately 0.60–0.80 after the initial dip) with early peaks near 1.08–1.10, DL-only narrows the band to about 0.65–0.74, while the full proposed tightens it further to roughly 0.68–0.72, confirming that uncertainty-aware safety filtering and bilevel coupling directly improve damping, reduce overshoot, and accelerate settling.

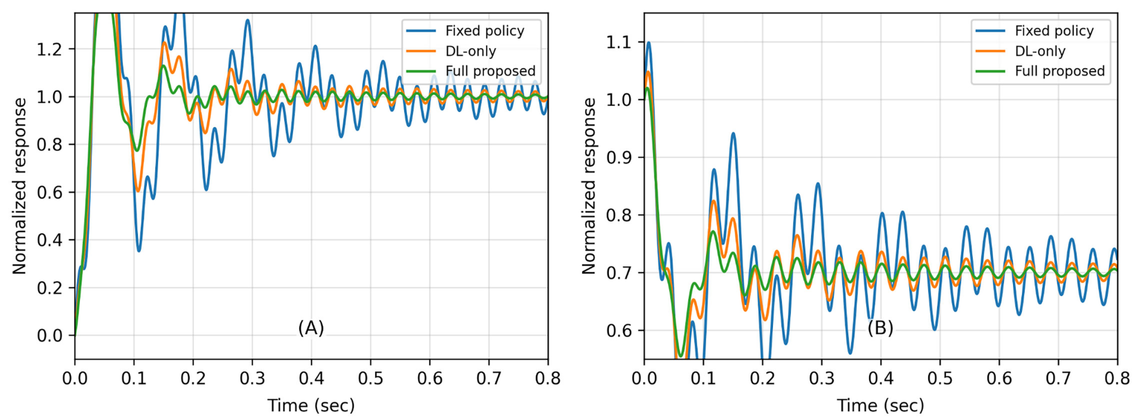

Figure 4: Discussion of transient energy response: (**A**) transient response around the normalized target level near 1.0 under fixed policy, DL-only, and the full proposed method; (**B**) transient response around the lower normalized operating level near 0.70 under the same control strategies.

In Fig. 5A of the utility/throughput tracking results (oracle vs. fixed policy), the oracle starts near 220 while the fixed policy begins around 205, implying an initial loss of about 15 units. After the first regime change at $t \approx 1.0$–1.1 s, the oracle drops to ~140–150 and then operates around ~170–190, whereas the fixed policy falls lower (~130–145) and remains persistently below the oracle by approximately 15–35 units across the mid-horizon. Around $t \approx 3.2$–3.4 s, when the oracle transitions from ~185–195 down to ~125–140, the fixed policy again shows lag and under-tracking, typically staying 10–25 units below the reference. The most severe mismatch occurs near $t \approx 4.2$–4.4 s, where the oracle exhibits a sharp drop (approaching ~10–30 during the transient), while the fixed policy responds more slowly and remains smoother and higher (roughly ~40–80), indicating inability to track rapid operating-point changes under non-stationary constraints.

In Fig. 5B,C (DL-only vs. full proposed tracking), the DL-only strategy substantially improves fidelity but still shows noticeable error during abrupt transitions: following the first drop at $t \approx 1.1$ s, DL-only typically stays within ~10–20 units after the transient, yet around later regime shifts ($t \approx 3.3$ s and $t \approx 4.3$ s) its deviation increases to approximately ~20–30 units with slower recovery. By comparison, the full proposed algorithm tracks the oracle more closely throughout, with typical deviations on the order of ~5–15 units and a visibly smoother realignment after the large drop at $t \approx 4.2$–4.4 s, where the oracle approaches very low values. This tighter tracking and reduced oscillatory recovery indicate that hypergradient-based bilevel coupling and safety-aware robust scheduling effectively translate predictive capability into feasible, responsive control actions under fast-changing energy/throughput regimes.

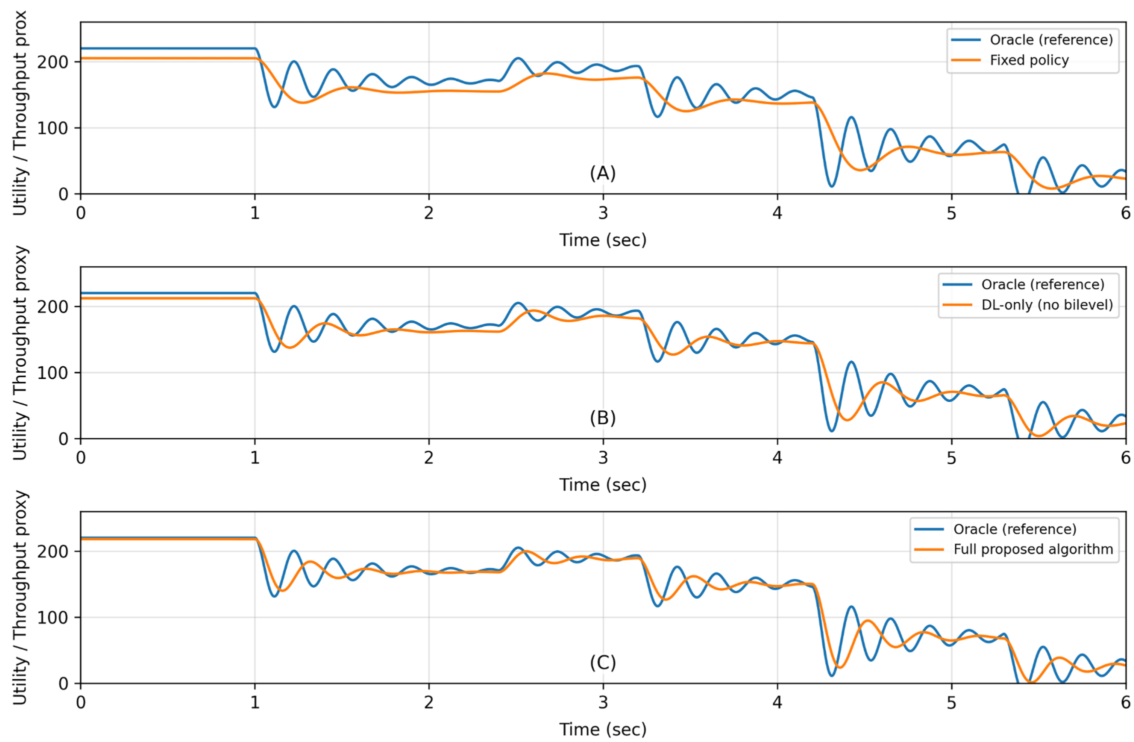

Figure 5: Discussion of utility/throughput tracking performance. (**A**) Tracking comparison between the oracle reference and the fixed policy. (**B**) Tracking comparison between the oracle reference and the DL-only strategy without bilevel coupling. (**C**) Tracking comparison between the oracle reference and the full proposed algorithm.

5 Conclusion

This paper presented a unified deep learning–driven bilevel optimization framework for energy-aware scheduling in IoT-based wireless sensor networks. Unlike conventional approaches that decouple energy prediction from control decisions, the proposed methodology jointly optimizes a residual energy predictor and a constrained scheduling policy within a closed-loop learning–control architecture. By explicitly modeling temporal energy dynamics, security-induced overheads, and prediction uncertainty, the framework enables adaptive and risk-aware scheduling that is robust to non-stationary network conditions. At the lower level, a lightweight TCN–LSTM predictor learns node-level residual energy evolution from real WSN data while providing uncertainty estimates that quantify forecast reliability. These predictions are not used in isolation; instead, they directly inform the upper-level controller, which optimizes duty cycle, sensing rate, and transmission power under explicit safety constraints, trust-region bounds, and multi-objective sustainability criteria. Hypergradient-based bilevel coupling ensures that the learning process is shaped by long-term network objectives rather than short-term prediction accuracy alone, transforming the predictor into a control-aware component of the system. Analytical evaluation and expected-performance analysis indicate that the proposed framework can substantially reduce energy prediction error while improving scheduling safety, energy balance, and network lifetime compared to static or single-level baselines. Importantly, the framework's modular design and reliance on lightweight models make it suitable for deployment in resource-constrained IoT environments. Future work will extend this framework to energy-harvesting WSNs, distributed and federated learning settings, and real-time hardware deployments. In addition, incorporating adaptive security policies and multi-hop routing decisions represents a promising direction for further enhancing the sustainable and secure operation of IoT networks.

Acknowledgement: The authors extend their appreciation to Umm Al-Qura University, Saudi Arabia for funding this research work through grant number: 26UQU4270203GSSR01.

Funding Statement: This research work was funded by Umm Al-Qura University, Saudi Arabia, under Grant Number: 26UQU4270203GSSR01.

Author Contributions: Mohammad Q. Al-Jamal: conceptualization; data curation; formal analysis; investigation; methodology; software; validation; visualization; writing—original draft. Mahmoud Al Jamal: data curation; formal analysis; investigation; methodology; software; validation; visualization; writing—original draft. Bashar S. Khassawneh: conceptualization; methodology; supervision; writing—review & editing. Ayoub Alsarhan: formal analysis; investigation; methodology; validation; writing—review & editing. Amina Salhi: validation; resources; supervision; writing—review & editing. Tahani Alsubait: validation; resources; supervision; writing—review & editing. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data supporting the findings of this study are publicly available on Kaggle under the title "WSN Dataset for Secure and Energy Efficient" at: https://www.kaggle.com/datasets/ziya07/wsn-dataset-for-secure-and-energy-efficient (accessed on 10 April 2026).

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sharma MK, Zappone A, Debbah M, Assaad M. Deep learning based online power control for large energy harvesting networks. In: Proceedings of the ICASSP 2019—2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2019 May 12–17; Brighton, UK. p. 8429–33.
2. Iyobhebhe M, Tekanyi AMS, Abubilal KA, Usman AD, Isiaku Y, Agbon EE, et al. A review on energy consumption model on hierarchical clustering techniques for IoT-based multilevel heterogeneous WSNs using energy aware node selection. Vokasi Unesa Bull Eng Technol Appl Sci. 2025;2(2):100–11. doi:10.26740/vubeta.v2i2.34882.
3. Goswami P, Khan T, Pathak V, Alabdultif A. Machine learning based dynamic trust estimation framework for securing wireless sensor networks. Sci Rep. 2025;15(1):35821. doi:10.1038/s41598-025-19768-z.
4. Tawfeek MA, Alrashdi I, Alruwaili M, Talaat FM. A fuzzy multi-objective framework for energy optimization and reliable routing in wireless sensor networks via particle swarm optimization. Comput Mater Contin. 2025;83(2):2773–92. doi:10.32604/cmc.2025.061773.
5. Sah DK, Hazra A, Mazumdar N, Amgoth T. An efficient routing awareness based scheduling approach in energy harvesting wireless sensor networks. IEEE Sens J. 2023;23(15):17638–47. doi:10.1109/JSEN.2023.3279249.
6. Oudah AY, Shaker LM. Wireless sensor networks in renewable energy monitoring: a comprehensive survey of protocols and applications. AUIQ Tech Eng Sci. 2025;2(2):8. doi:10.70645/3078-3437.1037.

7. Alrayes FS, Maray M, Alshuhail A, Almustafa KM, Darem AA, Al-Sharafi AM, et al. Privacy-preserving approach for IoT networks using statistical learning with optimization algorithm on high-dimensional big data environment. Sci Rep. 2025;15(1):3338. doi:10.1038/s41598-025-87454-1.
8. Fereidunian A, Alimoradi Z, Sebt MA. Energy smart environments: toward a systems perspective framework for the energy dimension of smart systems of systems. In: Proceedings of the 2025 6th International Conference on Optimizing Electrical Energy Consumption (OEEC); 2025 Feb 25–26; Najafabad, Iran. p. 1–10. doi:10.1109/OEEC66525.2025.11100143.
9. Ziya07. WSN dataset for secure and energy efficient wireless sensor networks. [cited 2026 Jan 29]. Available from: https://www.kaggle.com/datasets/ziya07/wsn-dataset-for-secure-and-energy-efficient.
10. Zeng W, Wang Z, Yang S, He D, Chan S. ICMH-CHR: an intra-cluster multi-hop based cluster head rotation protocol for wireless sensor networks. Ad Hoc Netw. 2025;173(4):103829. doi:10.1016/j.adhoc.2025.103829.
11. Zhang Y, Yang L, Tan Y. Energy-efficient adaptive routing in heterogeneous wireless sensor networks via hybrid PSO and dynamic clustering. J Cloud Comput. 2025;14(1):46. doi:10.1186/s13677-025-00768-3.
12. Akram M, Ullah Bazai S, Imran Ghafoor M, Akram S, Ilyas QM, Mehmood A, et al. EEMLCR: energy-efficient machine learning-based clustering and routing for wireless sensor networks. IEEE Access. 2025;13(17):70849–71. doi:10.1109/ACCESS.2025.3562368.
13. Naderi MY, Basagni S, Chowdhury KR. Modeling the residual energy and lifetime of energy harvesting sensor nodes. In: Proceedings of the 2012 IEEE Global Communications Conference (GLOBECOM); 2012 Dec 3–7; Anaheim, CA, USA. p. 3394–400. doi:10.1109/GLOCOM.2012.6503639.
14. Freschi V, Lattanzi E. A study on the impact of packet length on communication in low power wireless sensor networks under interference. IEEE Internet Things J. 2019;6(2):3820–30. doi:10.1109/JIOT.2019.2891841.
15. Tonin A, Nasser Ali Algabri M, Ali MN, Ghurab M, Ahmed Ghaleb Al-Khulaidi A, Ahmed Al-Baltah I. A novel energy model for optimizing energy efficiency and prolonging network lifetime in MANET. IEEE Access. 2025;13(3):117771–87. doi:10.1109/ACCESS.2025.3586795.
16. Alanazi R, Obayya M, Alghamdi AM, Nemri N, Alshahrani S, Alduaiji N, et al. Machine learning-driven routing optimization for energy-efficient 6G-enabled wireless sensor networks. Alex Eng J. 2025;129(1):877–88. doi:10.1016/j.aej.2025.07.032.
17. Singh A, Raj A, Rani P, Khatibi A, Aldeeb H, Shukla PK, et al. Resilient wireless sensor networks in industrial contexts via energy-efficient optimization and trust-based secure routing. Peer Peer Netw Appl. 2025;18(3):132. doi:10.1007/s12083-025-01946-5.
18. Kaddi M, Omari M, Salameh K, Alnoman A. Energy-efficient clustering in wireless sensor networks using grey wolf optimization and enhanced CSMA/CA. Sensors. 2024;24(16):5234. doi:10.3390/s24165234.
19. Jiang L, Xiao Q, Chen T. Improved analysis of penalty-based methods for bilevel optimization with coupled constraints. In: Proceedings of the 2025 33rd European Signal Processing Conference (EUSIPCO); 2025 Sep 8–12; Palermo, Italy. p. 1183–7. doi:10.23919/EUSIPCO63237.2025.11226224.