On-Device Federated Learning for Remote Alpine Livestock Monitoring

✉Sabtain Ahmad[1,*][0009−0006−3792−3249], Thomas Schneidergruber[2,*][0009−0009−7629−3042], Ivona Brandic[1][0009−0007−0661−5937], and Johannes Scholz[2][0000−0002−3212−8864]

[1] Vienna University of Technology, Vienna Austria
{sabtain.ahmad, ivona.brandic} @tuwien.ac.at
[2] University of Salzburg, Salzburg, Austria
{thomas.schneidergruber, johannes.scholz} @plus.ac.at
[*] These authors contributed equally to this work.

Abstract. Alpine livestock monitoring is critical for ecological preservation and agricultural efficiency. However, existing solutions struggle with energy constraints, limited network availability, and intermittent connectivity in remote environments. To address this, we propose an on-device federated learning framework tailored for PV-powered IoT sensors to optimize energy-communication tradeoffs. Our approach introduces staleness-aware aggregation and solar-aware training scheduling to address intermittent connectivity and PV variability in remote alpine environments. Deployed on a real-world testbed with collar sensors, the framework achieves 92% accuracy in time-series location prediction and 89% F1-score in anomaly detection while using 68% less energy than centralized baselines.

Keywords: Edge Intelligence · Federated Learning · PV Sensors · Livestock Monitoring.

1 Introduction

Livestock monitoring is crucial for sustainable agriculture, especially in mountainous regions, to support biodiversity and rural economies. However, continuous monitoring of free-grazing animals in these remote, resource-constrained environments faces critical challenges [1]. Traditional cloud-based approaches, which rely on centralized data aggregation requiring high-bandwidth communication, are impractical due to a lack of internet connectivity, energy constraints, and the need for real-time decision-making [2]. Emerging collar-mounted sensors equipped with GPS, accelerometers, and solar harvesting capabilities offer a promising alternative, but their potential is limited by on-device computational powers. Furthermore, transporting frequent and huge quantities of data not only requires significant network resources but also puts a strain on the battery of these resource-constrained IoT sensors [1].

For livestock monitoring, IoT sensors collect farm data from animals and send it to the cloud for ML model training. However, in remote alpine regions, limited energy and internet access pose major challenges. Alternatively, deploying ML models on sensors directly demands overcoming two key challenges: 1) intermittent connectivity, which requires adaptive communication across the computing continuum, and 2) energy-accuracy tradeoff, which arises due to the limited battery power of sensors and unstable solar harvesting in alpine regions [3].

Federated learning (FL) is a decentralized approach that allows model training through collaboration between devices/clients without requiring them to share their data as it moves learning away from the cloud to devices [4]. FL is particularly appealing for smart IoT applications as it reduces communication costs and ensures data privacy. Although FL has been extensively studied for domains such as smart cities [5], healthcare [6], and intelligent traffic management [7], its adoption in smart farming, especially for remote alpine regions, is still limited. Existing FL frameworks for IoT [8, 9] focus on urban and controlled environments, assuming stable connectivity and grid-powered devices. Furthermore, these frameworks are often tailored for devices with sufficient computational, network, and energy resources, making them unsuitable for resource-constrained IoT devices, such as PV-powered sensors used for livestock monitoring.

We propose an energy-efficient FL framework (EA-FL) specifically designed for resource-constrained PV-powered IoT sensors, with our major contributions listed as follows:

1. PV-aware FL architecture that aims to extend the lifetime of the system by adapting local training to the availability of solar energy.
2. Hybrid network optimization for enabling model aggregation over 4G/5G while maintaining fault tolerance via LoRaWAN and reducing communication costs under intermittent connectivity.
3. We conduct evaluations on a real testbed consisting of collar sensors designated as clients and a Raspberry Pi-based edge server acting as a central aggregator.

Beyond livestock monitoring, the proposed framework advances FL for extreme environments (e.g., disaster zones) and provides design principles for sustainable, decentralized IoT systems.

The rest of this paper is structured as follows: Section 2 reviews related work. Section 3 details the framework design, while Section 4 describes use cases and implementation details. Section 5 presents the results and discusses the system's scalability. We present our conclusion in Section 6.

2 Related Work

IoT technologies are increasingly being used for livestock monitoring, to improve animal welfare and optimize farm management [10, 11]. Prior works in this domain have utilized a variety of sensor modalities, including GPS, accelerometers, and environmental sensors to track animal behavior [11], location [12], and health [10]. However, traditional systems often rely on centralized data collection and processing, which results in critical challenges due to intermittent connectivity in remote environments [3]. Moreover, the high communication costs associated with continuous data transmission present another challenge that significantly impacts the systems' lifetime.

FL for IoT and Edge Systems: FL has emerged as a key paradigm for collaborative learning in IoT applications [9]. Although vanilla FL [4] assumes stable connectivity and homogeneous data, recent advances adapt FL to handle various IoT constraints [13, 14]. For instance, Charles et al. [15] employed gradient compression to reduce communication overhead, and Lim et al. [16] addressed the challenge of stragglers in edge networks. In [17], a dynamic voltage scaling is proposed for FL devices. A client selection strategy with the objective of ensuring energy fairness is proposed in [18]. However, these frameworks focus on urban deployments (e.g., smart factories) and consequently fail to address the critical challenges of remote alpine environments.

FL is gaining traction in agriculture and wildlife monitoring [19, 20]. Manoj et al. [21] employed FL for crop yield prediction using drones, assuming continuous 5G connectivity. These preliminary works in agricultural applications of FL have demonstrated their potential to reduce communication overhead while preserving data privacy. However, these studies typically focus on relatively simpler environments and do not account for the compound challenges of remote alpine regions, which are subject to severe energy and communication constraints.

Our work differentiates itself by addressing these challenges through edge-device coordination. More specifically, we propose an on-device FL framework deployed on PV-powered collar sensors, combined with a hybrid communication strategy to ensure robust performance in remote alpine settings.

3 Framework Design

Our proposed framework (Fig. 1) is designed to address the energy and connectivity constraints of remote alpine livestock monitoring while leveraging on-device FL framework to learn a global model. This section describes the testbed (hardware) setup, the FL framework, and optimization strategies to ensure robust performance under remote challenging conditions.

3.1 Testbed setup

The proposed framework (Fig. 1) consists of three layers: PV-powered collar sensors, a hybrid 4G/5G-LoRaWAN communication layer, and a lightweight FL server. Each collar sensor integrates a GPS module, 6-axis IMU (accelerometer/gyroscope), 2W solar panel with LiPo battery (550 mAh), and a Raspberry Pi 4 compute module. Furthermore, the setup includes 4G/5G and LoRaWAN gateways to enable communication between sensors and the FL server. The FL server, hosted on a low-power edge device, coordinates training via adaptive aggregation and network orchestration.

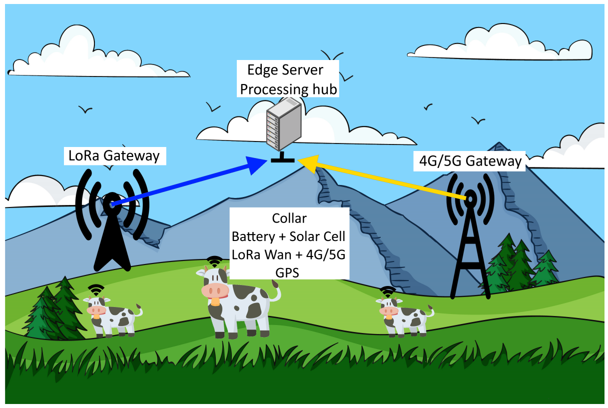


Fig. 1: High-level architecture: Alpine livestock monitoring using on-device FL.

3.2 Federated Learning Framework

The EA-FL framework consists of the standard client-server model, where each collar sensor serves as an FL client, and an edge server aggregates and coordinates model updates. The training process consists of three main steps: local model training, federated aggregation, and communication-efficient update strategy. Each sensor trains a local model using its private dataset D_i, optimizing a loss function $\mathcal{L}$ with respect to its local model parameters w_i;

$$w_i^{t+1} \leftarrow w_i^t - \eta \nabla \mathcal{L}(w_i^t, D_i) \tag{1}$$

where η is the learning rate and w_i^t denotes the local model update at iteration t.

After local training, sensors send their model updates to the edge server. Instead of sending full gradient updates, each sensor applies top-k sparsification to send only the most significant updates;

$$G_k = \{g_j | j \in arg\max_k |g_j|\} \tag{2}$$

where g_j is the gradient component.

Once the server has received the updates, it aggregates local models to generate the global model update;

$$w_g^{t+1} = \sum_{i \in S} \frac{|D_i|}{D} w_i^t \tag{3}$$

where S represents the number of active sensors and w_g denotes the global model.

Asynchronous Staleness-Aware Aggregation: Conventional frameworks such as FedAvg [4] assume synchronous participation, which fails in alpine environments where sensors often disconnect due to limited network or energy constraints. Specifically, intermittent connectivity leads to staleness; local updates from some sensors may arrive late. Consequently, aggregating stale updates could cause model divergence, as outdated parameters may conflict with recent updates. To address this, we propose an asynchronous aggregation method that incorporates delayed updates rather than discarding them.

For a client i with a local model, w_i delayed by t_i rounds, its contribution to the global model is scaled as;

$$w_g^{t+1} = \frac{\sum_{i \in S} \alpha^{t_i} \cdot w_i^t}{\sum_{i \in S} \alpha^{t_i}} \tag{4}$$

where α^t is a staleness discount factor.

PV-Aware Training Scheduling: Operating in extreme alpine environments imposes strict constraints on the availability of energy to power the system to ensure uninterrupted operations. We propose an energy-aware training mechanism to dynamically optimize model training, aggregation, and model transmission to maximize efficiency and robustness.

More formally, each sensor operates under a strict energy budget and, therefore must balance energy consumption between model training, inference, and communication. Given an energy (battery) level E_i at sensor i, the sensor suspends training if the battery level drops below a threshold ($E_i < E_{thresh}$). Additionally, the number of local training epochs is dynamically adjusted based on current battery levels. Let ζ denote the epoch count, it can be computed as follows:

$$\zeta_i = \zeta_{max} \times (\frac{E_i - E_{min}}{E_{max} - E_{min}})^\gamma \tag{5}$$

where ζ_{max} is the maximum allowed training epochs, E_{min} and E_{max} define the operational battery range, and γ is a decay parameter that controls how aggressively training is reduced as the battery levels deplete.

Hybrid Network Orchestration: Considering the intermittent and bandwidth-limited nature of LoRaWAN and cellular 4G/5G networks in alpine regions, we implement an energy-latency tradeoff model to dynamically select the optimal communication strategy. The hybrid network orchestration (HNO_{comm}) evaluates whether a sensor should transmit model updates immediately or defer based on energy availability and network conditions:

$$HNO_{comm} = arg_{m \in 4G/5G, LoRaWAN}(\frac{d_m}{E_m} + \lambda L_m) \tag{6}$$

where d_m is the data transmission cost, E_m is the energy budget required for transmission, and L_m is the expected latency in mode m. λ is a latency penalty factor for prioritizing low-latency communication when immediate updates are required. Algorithm 1 outlines the steps involved in the training of EA-FL.

Additionally, we employ dynamic update transmission, where updates are locally aggregated and only sent when;

$$E_i > E_{thresh} \quad and \quad L_m < L_{max} \tag{7}$$

Energy Model: To optimize model transmission, we define an energy model that quantifies the cost of transmitting model updates over different networks. The total energy consumption $E_m(n)$ required to transmit a payload of n bytes over network mode m is given by:

$$E_m(n) = n \cdot E_{byte}(m) + T_{tx}(m) \cdot P_{idle}(m) \tag{8}$$

where $E_{byte}(m)$, $T_{tx}(m)$, and $P_{idle}(m)$ are the energy required to transmit one byte, transmission duration, and power consumption during idle states for network m, respectively. For example, transmitting a 100 KB model over 4G consumes $1.2J$ compared to $0.08J$ over LoRaWAN.

Algorithm 1: Federated Learning with Energy-Aware Optimization

Input: Communication rounds T, participating devices $\mathcal{S}$
Output: trained global model w_g
for *each global round* $t = 1, \ldots, T$ **do**
    Server selects subset of active devices $\mathcal{S}_t \subseteq \mathcal{S}$;
    for *each device* $i \in \mathcal{S}_t$ ***in parallel*** **do**
        if $E_i < E_{thresh}$ **then**
            Suspend training and perform inference only;
            continue;
        Adaptive epoch selection: using EQ (5)
        Local training using SGD: $w_i^{t+1} \leftarrow w_i^t - \eta \nabla \mathcal{L}(w_i^t, D_i)$
        Gradient pruning: Transmit top-k gradients;
        Decide communication mode using Eq. 6
        Transmit model update to edge server;
    Server aggregates updates using staleness-aware aggregation using Eq. 4

4 Implementation and Experiments

We deployed our framework in the Austrian Alps involving three collar sensors attached to free-grazing cows and evaluated it on two real-world use cases: location prediction and anomaly detection.

4.1 Use Case 1: Animal Location Prediction

Accurately tracking livestock movements in alpine regions is essential for monitoring grazing patterns and keeping animals safe. However, continuous GPS tracking is infeasible due to energy constraints and intermittent connectivity. We frame animal location prediction as a time-series forecasting problem, where we train a model to predict an animal's future location based on historical sensor data. More formally, given a sequence of past sensor measurements $X_t =$ [Latitude, Longitude, Altitude, ...], our goal is to predict next position;

$$\hat{X}_{t+1} = f(X_{t-l}, \ldots, X_t; \theta) \tag{9}$$

where X_t is the feature vector at time t, f is the prediction model (LSTM), θ are the global model parameters learned via FL, and l is the lookback window. Each sensor trains on its local movement data while periodically communicating updates with the edge server for global model training (Section 3.2)

Model: We developed an LSTM-based model to predict future animal locations using historical GPS trajectories. The model architecture includes input, output, and two LSTM layers (32 hidden units in each).

The model is trained using Adam optimizer (lr=0.001, weight decay=0.01), dropout (0.2), and MSE loss. For each FL client, data is partitioned by (80% train, 10% validation, 10% test) to preserve temporal integrity. Predictions within 5m of ground truth are labeled accurate.

Baselines: We compare our proposed framework with following baselines;

- Centralized: a conventional approach in which all training data is uploaded to a central server for model training and serves as an upper bound.
- Standalone: a non-collaborative baseline, where each sensor trains a model using only its own local data and acts as a lower bound.
- Standard FL: we compare our approach with the conventional synchronous FL with standard averaging and without dynamic aggregation.

We define three key metrics to measure energy consumption, communication efficiency, and prediction precision defined in Equations 10, 11, and 12, respectively.

$$E_{\text{total}} = \sum_{t=1}^{T} \sum_{i \in S} (E_{\text{train}}(i,t) + E_{\text{comm}}(i,t)) \tag{10}$$

where $E_{\text{train}}(i,t)$ is the energy required for training a local model for ζ_i epochs on i-th device, while $E_{\text{comm}}(i,t)$ (Eq. 8) represents the energy used for communicated model updates

$$Comm_{\text{total}} = \sum_{t=1}^{T} \sum_{i \in S} |w_i^t| \tag{11}$$

where $|w_i^t|$ is the size (in bytes) of the update of the i-th device model in round t.

$$MAE = \frac{1}{N}\sum_{j=1}^{N} ||X_{j+1} - \hat{X_{j+1}}|| \tag{12}$$

4.2 Use Case 2: Anomaly Detection

In livestock monitoring, anomalies such as irregular movements, prolonged inactivity, or excessive motion can indicate potential health problems, predator threats, or natural disasters. Moreover, due to their erratic nature and harsh alpine environments, wireless sensor networks deployed in such environments are more prone to experiencing outliers. This is primarily because these networks collect real-world data using imperfect sensors, which are susceptible to external factors such as aging and potential malfunctions [22]. We adopt an unsupervised autoencoder-based anomaly detection model [23] to detect abnormal movement patterns and trajectories. Given a sequence of past sensor measurements, our goal is to identify instances with abnormally high reconstruction error:

$$A_t = \mathbb{I}(||X_t - f(X_{t-n}, \ldots, X_t; \theta)|| > \delta) \tag{13}$$

where A_t represents the instance label ($A_t = 1$ if anomalous, 0 otherwise) and δ is the threshold defined as; $mean\ loss \times d.standard\ deviation$.

Data preparation: A trajectory is described by a chronologically ordered sequence of past sensor measurements, where each pair of consecutive data points constitutes a trajectory segment. Segments with a duration shorter than two minutes or exceeding 30 minutes are excluded to ensure data consistency. Additionally, to reduce GPS signal noise and improve reliability, a Kalman filter [24] is applied as a preprocessing step. Eight features were extracted using both individual trajectory segments and a four-hour sliding window (covering 12 segments, advancing one segment at a time):

- **Distance:** Segment length (km).
- **Average movement:** Euclidean distance between the most recent position and the average of prior starting points in the window.
- **Path-length:** Cumulative segment length in the window (km).
- **Summary statistic:** Mean, median, 25th/75th percentiles, and standard deviation of segment lengths (excluding min/max due to noise sensitivity).

We use the ensemble method to prepare the ground truth for the evaluation. More specifically, a segment is classified as anomalous using the majority vote based on the decision of the Optics [25], Abod [26], Isolation-Forest [27] models. Moreover, these identified abnormal segments should occur consecutively, reflecting the temporal continuity of meaningful anomalies. This assumption is aligned with scenarios such as a cow fleeing from a predator, sustaining an injury, or a sensor malfunction.

Model: The autoencoder architecture consists of five layers with the configuration: [8, 6, 5, 6, 8]. The model architecture was optimized using grid search.

Table 1: Predictive Performance comparison.

Method	Location Prediction			Anomaly Detection			
	MAE (m)	Accuracy	Energy (Wh)	F1-score	Precision	Recall	Energy (Wh)
Centralized	3.2	93.5	1700	91	93	89	1600
Standalone	8.1	72.4	32	65	68	62	27
FedAvg	4.5	89.1	81	83	81	85	67
EA-FL (ours)	3.8	92.2	36	89	91	87	30

8-bit integer quantization and quantization-aware training were applied to reduce memory usage and enhance speed. The model is trained using Adam optimizer (lr=0.001), dropout (0.05), and RMSE as reconstruction loss. We include additional metrics pertinent to anomaly detection, including F1-score, precision, and recall, to evaluate the proposed framework against the baselines.

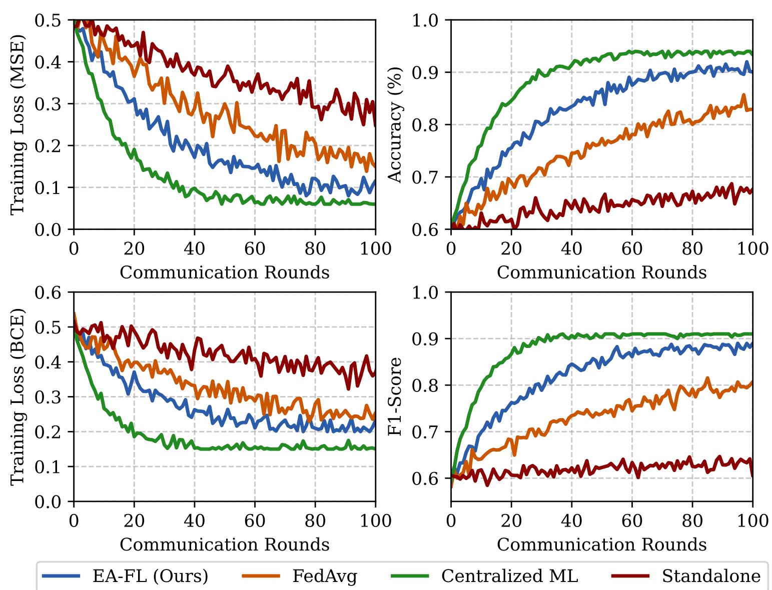


Fig. 2: Training and convergence comparison.

5 Numerical Results

This section presents a detailed analysis of our proposed framework evaluated on a real-world testbed consisting of three PV-powered collar sensors deployed in an alpine region in Austria. We analyze the framework's performance in terms of predictive performance, energy efficiency, and communication overhead for animal location prediction and anomaly detection.

5.1 Prediction Performance

Table 1 presents the performance and energy efficiency of EA-FL against the baselines for both use cases. EA-FL achieves a balance between accuracy and energy efficiency, outperforming standalone and FedAvg baselines while almost achieving the prediction accuracy of centralized training. More specifically, for location prediction, EA-FL achieves MAE of 3.8m and 92.2% accuracy, closely matching the $3.2m$ MAE and 93.5% accuracy of the centralized baseline while consuming 47× less energy, and significantly outperforming the FedAvg and standalone baselines. Similar results are obtained for the second use case (anomaly detection); EA-FL achieves an F1-score of 89 (91 precision, 87 recall), outperforming FedAvg and standalone by 6 and 24 F1 points, respectively. Although the standalone baseline consumes less energy (27 vs. 30 Wh), it struggles to attain a reasonable performance (F1-score: 65), demonstrating the impracticability of isolated training due to data and computational constraints. Furthermore, unlike FedAvg, EA-FL's ability to transmit pruned updates and staleness-aware aggregation at the edge reduces communication overhead by 60% while simultaneously achieving $3.5-7.2\%$ better performance compared to FedAvg.

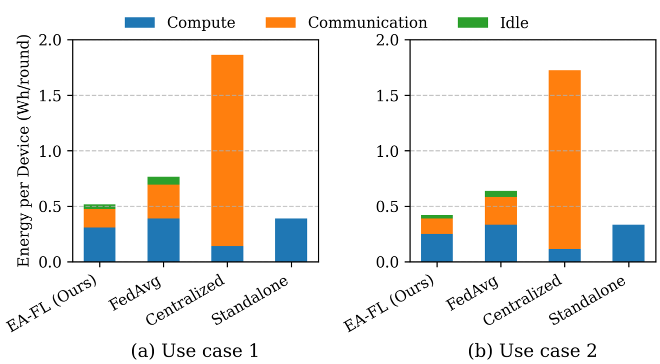


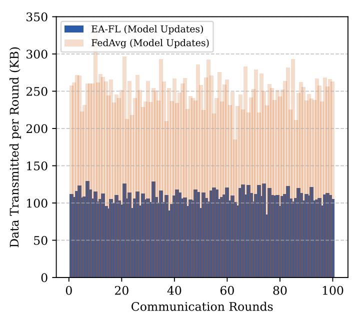
Fig. 3: Energy consumption breakdown of each method.

Figure 2 discusses the convergence and performance of these methods under alpine network conditions. As highlighted in the figure, EA-FL converges (80% Accuracy/F1) within 30 and 25 rounds for location prediction and anomaly detection, respectively. This is considerably faster than FedAvg (60, 78 rounds) due to staleness-aware weighted aggregation, which prioritizes fresh updates from sensors with heterogeneous data.

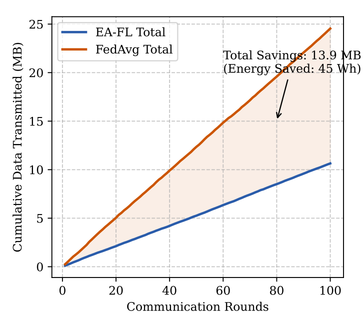
5.2 Energy Efficiency

Figure 3 compares the average energy consumption per sensor per training round, divided into compute, communication, and idle energy consumption. As evident from the figure, communication overhead is the most dominant factor influencing overall energy consumption, particularly for the Centralized baseline, where all sensor data is transmitted to a remote server. Our EA-FL framework optimizes energy consumption for both component computation and communication by employing dynamic training round adjustment and model quantization coupled with an optimal communication strategy. On the other hand, a Centralized baseline incurs the highest consumption due to high communication overhead, rendering this approach infeasible for resource-constrained environments.

For location prediction (left), EA-FL consumes $0.52Wh$, resulting in a 48% reduction compared to FedAvg. Although FedAvg reduces communication and thus the communication cost compared to Centralized. However, the frequent model synchronization leads to an increased overall energy drain ($0.77Wh$). Similar results are obtained for anomaly detection (right). The lightweight autoencoder and hybrid network strategy result in 52% reduction compared to FedAvg.



(a) Per-Round Data Overhead



(b) Cumulative Data & Energy Savings

Fig. 4: EA-FL vs. FedAvg Communication & Energy Comparison.

5.3 Communication Overhead

In this experiment, we compare the communication overhead of EA-FL with the FedAvg baseline. Figure 4 analyzes the per-round data transmission and the cumulative energy consumption of both approaches. As shown in Figure 4 (left), EA-FL reduces communication overhead by 56%; it transmits 100KB/round on average compared to 250KB/round transmitted by FedAvg. The major reason for the higher communication overhead for FedAvg can be attributed to the full model updates, along with the frequent communication with the server. In contrast, EA-FL minimizes redundant updates while prioritizing the most informative parameters through dynamic update transmission.

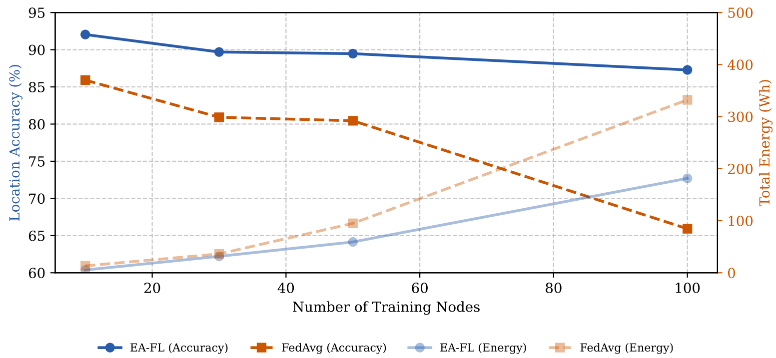


Fig. 5: Scalability analysis: impact on accuracy and energy consumption.

Figure 4 (right) presents the cumulative data transmitted over 100 training rounds. FedAvg accumulates 24.4 MB of transmitted data, while EA-FL requires only 11 MB, resulting in a 55% reduction. This directly translates into energy efficiency; EA-FL consumes 36 Wh over the training period, compared to 83.4 Wh for FedAvg. The resultant energy efficiency extends the battery life of the sensor by 58 days, an improvement of more than 21 days over FedAvg.

5.4 Scalability Analysis

To validate EA-FL's applicability in large-scale environments, we emulated a 100 node network using the NS-3 simulator under realistic bandwidth constraints, reflecting alpine network conditions (50% 4G availability, 5km LoRaWAN range). We scaled the system from 10 to 100 nodes, and present the findings from this experiment in Figure 5.

As evident from Figure 5, EA-FL consistently outperforms FedAvg both in terms of accuracy and energy consumption. As more and more nodes are added, the accuracy of both EA-FL and FedAvg declines, although EA-FL exhibits robust performance and experiences a lower drop in performance (7.6% vs. 23.7%). Furthermore, similar results are obtained for energy consumption as we scale the system. EA-FL's energy consumption grows sublinearly, increasing from $6.8Wh$ at 10 nodes to $181Wh$ at 100 nodes. On the other hand, FedAvg follows a steeper trajectory, reaching $332Wh$ at 100 nodes (83% more than EA-FL). This significant drop in performance can be attributed to increased communication and the inability to handle stragglers in the face of intermittent connectivity. Unlike FedAvg, EA-FL maintains robust performance throughout due to quantized updates, staleness-aware aggregation, and hybrid network optimization.

6 Conclusion

This paper presents EA-FL, an on-device FL framework tailored for PV-powered, resource-constrained alpine livestock monitoring. EA-FL overcomes the limitations of existing FL approaches in extreme environments by introducing statelessness-aware aggregation and hybrid network orchestration to ensure communication and energy efficiency without compromising much on prediction performance. Extensive evaluations conducted on a real-world testbed for two use cases, location prediction and anomaly detection, show that EA-FL achieves comparable performance to centralized learning while significantly reducing communication and energy overhead by 56% and 68%, respectively. These findings highlight EA-FL's practical applicability for harsh, remote environments, where communication and energy constraints pose significant challenges.

Acknowledgment This study was funded through the Austrian Research Promotion Agency (FFG) and the Virtual Shepherd Project (FO999910627)

Disclosure of Interests The authors have no competing interests to declare that are relevant to the content of this article.

References

1. H. K. Kondaveeti, G. B. Sai, S. A. Athar, V. K. Vatsavayi, A. Mitra, and P. Ananthachari, "Federated learning for smart agriculture: Challenges and opportunities," in *2024 Third International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)*, 2024, pp. 1–7.
2. V. M. Aleluia, V. N. Soares, J. M. Caldeira, and A. M. Rodrigues, "Livestock monitoring: Approaches, challenges and opportunities," *International Journal of Engineering and Advanced Technology*, vol. 11, no. 4, pp. 67–76, 2022.
3. S. Ahmad and A. Aral, "Hierarchical federated transfer learning: A multi-cluster approach on the computing continuum," in *2023 International Conference on Machine Learning and Applications (ICMLA)*, 2023, pp. 1163–1168.
4. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Artificial intelligence and statistics*. PMLR, 2017, pp. 1273–1282.
5. S. Pandya, G. Srivastava, R. Jhaveri, M. R. Babu, S. Bhattacharya, P. K. R. Maddikunta, S. Mastorakis, M. J. Piran, and T. R. Gadekallu, "Federated learning for smart cities: A comprehensive survey," *Sustainable Energy Technologies and Assessments*, vol. 55, p. 102987, 2023.
6. D. C. Nguyen, Q.-V. Pham, P. N. Pathirana, M. Ding, A. Seneviratne, Z. Lin, O. Dobre, and W.-J. Hwang, "Federated learning for smart healthcare: A survey," *ACM Computing Surveys (Csur)*, vol. 55, no. 3, pp. 1–37, 2022.
7. S. Ahmad and A. Aral, "Fedcd: Personalized federated learning via collaborative distillation," in *2022 IEEE/ACM 15th International Conference on Utility and Cloud Computing (UCC)*, 2022, pp. 189–194.
8. D. C. Nguyen, M. Ding, P. N. Pathirana, A. Seneviratne, J. Li, and H. V. Poor, "Federated learning for internet of things: A comprehensive survey," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 3, pp. 1622–1658, 2021.
9. T. Zhang, L. Gao, C. He, M. Zhang, B. Krishnamachari, and A. S. Avestimehr, "Federated learning for the internet of things: Applications, challenges, and opportunities," *IEEE Internet of Things Magazine*, vol. 5, no. 1, pp. 24–29, 2022.

10. O. Unold, M. Nikodem, M. Piasecki, K. Szyc, H. Maciejewski, M. Bawiec, P. Dobrowolski, and M. Zdunek, "Iot-based cow health monitoring system," in *International Conference on Computational Science*. Springer, 2020, pp. 344–356.
11. N. Arago, R. R. Robles, C. Alvarez, A. Mabale, C. Legista, N. Repiso, T. Amado, A. Thio-ac, L. K. Tolentino, J. Velasco *et al.*, "Smart dairy cattle farming and in-heat detection through the internet of things (iot)," *International Journal of Integrated Engineering*, vol. 14, no. 1, pp. 157–172, 2022.
12. P. Cousin, C. Dupont, S. Fatnassi, C. Pham, O. Thiare, A. Wussah, and S. Koffi, "Iot, an affordable technology to empower africans addressing needs in africa," in *2017 IST-Africa Week Conference (IST-Africa)*. IEEE, 2017, pp. 1–8.
13. T. T. Sari, S. Ahmad, A. Aral, and G. Seçinti, "Collaborative smart environmental monitoring using flying edge intelligence," in *GLOBECOM 2023 - 2023 IEEE Global Communications Conference*, 2023, pp. 5336–5341.
14. S. Ahmad, H. Uyanık, T. Ovatman, M. T. Sandıkkaya, V. D. Maio, I. Brandic, and A. Aral, "Sustainable environmental monitoring via energy and information efficient multi-node placement," *IEEE Internet of Things Journal*, 2023.
15. Z. Charles, K. Bonawitz, S. Chiknavaryan, B. McMahan *et al.*, "Federated select: A primitive for communication-and memory-efficient federated learning," *arXiv preprint arXiv:2208.09432*, 2022.
16. W. Y. B. Lim, J. S. Ng, Z. Xiong, J. Jin, Y. Zhang, D. Niyato, C. Leung, and C. Miao, "Decentralized edge intelligence: A dynamic resource allocation framework for hierarchical federated learning," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 3, pp. 536–550, 2021.
17. Y. Li, Y. Chen, K. Zhu, C. Bai, and J. Zhang, "An effective federated learning verification strategy and its applications for fault diagnosis in industrial iot systems," *IEEE Internet of Things Journal*, vol. 9, no. 18, pp. 16 835–16 849, 2022.
18. T. Nishio and R. Yonetani, "Client selection for federated learning with heterogeneous resources in mobile edge," in *ICC 2019-2019 IEEE international conference on communications (ICC)*. IEEE, 2019, pp. 1–7.
19. K. R. Žalik and M. Žalik, "A review of federated learning in agriculture," *Sensors*, vol. 23, no. 23, p. 9566, 2023.
20. A. Mao, E. Huang, H. Gan, and K. Liu, "Fedaar: A novel federated learning framework for animal activity recognition with wearable sensors," *Animals*, vol. 12, no. 16, p. 2142, 2022.
21. T. Manoj, K. Makkithaya, and V. Narendra, "A federated learning-based crop yield prediction for agricultural production risk management," in *2022 IEEE delhi section conference (DELCON)*. IEEE, 2022, pp. 1–7.
22. A. Ayadi, O. Ghorbel, A. M. Obeid, and M. Abid, "Outlier detection approaches for wireless sensor networks:," *Computer Networks*, vol. 129, pp. 319–333, 2017.
23. S. Ahmad, K. Styp-Rekowski, S. Nedelkoski, and O. Kao, "Autoencoder-based condition monitoring and anomaly detection method for rotating machines," in *IEEE International Conference on Big Data (Big Data)*, 2020, pp. 4093–4102.
24. R. E. Kalman *et al.*, "A new approach to linear filtering and prediction problems [j]," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
25. M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, "Optics: Ordering points to identify the clustering structure," *ACM Sigmod record*, vol. 28, pp. 49–60, 1999.
26. H.-P. Kriegel, M. Schubert, and A. Zimek, "Angle-based outlier detection in high-dimensional data," in *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2008, pp. 444–452.
27. F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation forest," in *2008 eighth ieee international conference on data mining*. IEEE, 2008, pp. 413–422.