Sensing Without Sensing: Energy-Efficient Virtual Sensing for IoT-Based Environmental Monitoring

Shadi Attarha, Anna Förster
Dept. Sustainable Communication Networks, University of Bremen, Germany
Corresponding author: sattarha@uni-bremen.de

***Abstract*—Environmental monitoring applications using IoT sensors play a crucial role across diverse domains such as agriculture, industrial facilities, and urban air quality monitoring. These systems typically rely on multiple heterogeneous sensors to collect environmental data and support informed decision-making. However, since many IoT devices are battery-powered with limited energy capacity, sustaining long-term operation becomes challenging, particularly when energy-intensive sensors are involved. To address these energy constraints, we propose a resource-efficient virtual sensing framework for real-time operation in multi-sensor IoT environments. Sensors or edge devices can be temporarily deactivated to reduce energy consumption while maintaining continuous data streams. Readings from deactivated sensors are reconstructed using a recursive prediction mechanism: each prediction uses the two most recent values from the sensor's stream, which may themselves be predicted, along with recent and current measurements from spatially correlated active sensors. The predicted values are continuously integrated into the data stream, enabling extended deactivation periods and strategic cycling of sensors between active and inactive states. We also propose a calibration phase that estimates the maximum safe deactivation period for each sensor to ensure reconstruction accuracy remains within acceptable bounds. We evaluate our method on real-world sensor datasets, demonstrating accurate reconstruction with low computational overhead. We showed that the model's lightweight design enables efficient edge deployment via TinyML implementation, reducing overall energy consumption without compromising data quality.**

***Index Terms*—Virtual sensing, TinyML, Internet of Things (IoT), sensor data quality.**

I. Introduction and Motivation

The Internet of Things (IoT) has become a key enabling technology for environmental monitoring across a wide range of applications, including smart homes, heating, ventilation, and air conditioning (HVAC) systems, as well as precision agriculture. By deploying distributed sensor networks that continuously acquire real-time data from the physical environment, IoT systems support data-driven decision-making and automated control, leading to improved operational efficiency [1]. For instance, heating and cooling energy demand contributes to approximately 40% of total building energy consumption, which in turn constitutes approximately one-third of global energy consumption [2]. In this context, IoT-based environmental monitoring has demonstrated substantial potential for energy savings. Studies have shown that spatial monitoring of temperature and humidity can facilitate HVAC control strategies that reduce energy consumption by 7% to 30% [3]. Similar benefits extend to other domains. For example, smart home systems optimize lighting and appliance usage based on occupancy and environmental conditions, while precision agriculture platforms reduce water and energy consumption through targeted irrigation and climate control.

However, this promising avenue for energy optimization introduces a critical paradox. While IoT systems enable energy savings in the monitored environments, the sensing infrastructure itself consumes energy that can undermine these benefits. Advances in low-cost sensing and wireless communication have made it technically feasible to deploy large numbers of sensors within buildings or fields to capture fine-grained spatial variations. However, scaling sensor density increases the overall sensor footprint, leading to higher battery consumption, increased electronic waste, and greater maintenance and operational overhead. In practice, many of these deployments also result in oversampled systems, where multiple sensors provide highly correlated measurements rather than additional information. This redundancy is particularly problematic given that many IoT devices are battery-powered and installed in locations where frequent battery replacement or recharging is costly or impractical.

Hence, in light of growing sustainability concerns and the imperative to develop energy-aware IoT systems, there is increasing interest in reducing the energy consumption of the sensing infrastructure itself without compromising the quality and reliability of the collected data. One promising direction toward this goal is *virtual sensing*, where the readings of a target sensor are inferred from other available sensors rather than obtained through direct measurement. By exploiting the spatial correlations that naturally exist among environmental variables, virtual sensors can provide high-level monitoring information at a considerably lower sensing and energy cost. Accordingly, for the sake of energy efficiency, numerous studies have proposed virtual sensing approaches for IoT-based environmental monitoring applications that focus on optimizing sensor placement during the system design phase. These methods typically identify a subset of physical sensor locations such that the values at unsensed locations can be estimated using other deployed sensors, thereby reducing deployment cost and the overall energy footprint [3]–[5].

While this design-time optimization paradigm is effective and well justified, it exhibits an important limitation. Existing approaches primarily assume static system configurations in which all deployed sensors remain continuously active during operation. As a result, they do not address runtime energy

ISBN: 978-3-903176-79-9

management, where already-deployed sensors could also be temporarily deactivated and their readings reconstructed from other active sensors. Therefore, it can be argued that energy savings can be further improved by extending virtual sensing beyond deployment-time decisions toward operational, run-time control. In this perspective, virtual sensing is used to support the dynamic operation of already-deployed sensing infrastructures. Under such a paradigm, selected sensors can be temporarily deactivated during operation, and their readings can be estimated from the remaining active sensing resources. Consequently, further energy conservation can be attained without compromising data quality.

In this paper, we propose a resource-efficient virtual sensing framework for multi-sensor IoT systems. The framework supports on-demand sensor deactivation while operating under strict computational and memory constraints. Our key contributions are summarized as follows:

- We propose a recursive prediction mechanism that continuously estimates deactivated sensor readings using short-term data from the target sensor and correlated active sensors, enabling efficient runtime operation suitable for TinyML deployment on resource-constrained devices.
- We introduce a multi-model strategy in which, for each target sensor, separate predictors are trained for all viable subsets of statistically correlated companion sensors. This ensures accurate and robust reconstruction even when the availability of companion sensors varies during operation.
- We conduct a calibration study for each target sensor and correlation subset, characterizing how reconstruction accuracy evolves as the deactivation duration increases. This analysis provides empirical guidance for selecting a safe dormant phase in real-world deployments.
- We show that the proposed virtual sensing approach can reduce energy consumption, as demonstrated in two distinct environmental monitoring scenarios.

The remainder of this paper is organized as follows. Section II details the proposed framework, including its key components and design choices. Section III presents the experimental setup and results used to evaluate model performance. Finally, Section IV summarizes the findings and outlines directions for future research.

II. Proposed Virtual Sensing Paradigm

In IoT applications, the energy consumption of edge devices is primarily driven by three main components: sensing, on-device computation, and wireless communication. The relative contribution of each component depends on deployment-specific factors such as sensor modality, sampling rate, and communication patterns. In many practical IoT systems, wireless transmission and power-intensive sensors (e.g., CO_2 and particulate matter sensors) dominate energy consumption, making continuous operation costly in long-term monitoring. To reduce this energy burden without disrupting data availability, we propose an adaptive virtual sensing paradigm. In this paradigm, individual sensors or, in some cases, entire edge devices, are intentionally deactivated for predefined periods, while a prediction-based virtual sensor module reconstructs the missing measurements. This allows the system to maintain continuous data streams while reducing energy consumption.

The proposed virtual sensing mechanism builds on our earlier fault-correction framework [6], which was originally developed to reconstruct isolated faulty or missing sensor readings. The key distinction in the present work lies in the operational setting. Rather than addressing sporadic faults, the virtual sensor is applied to intentional and potentially extended sensor deactivation. During these periods, missing measurements are continuously estimated and recursively fed back into the sensor stream, enabling uninterrupted virtualized outputs over extended durations. In addition, the framework is resilient to variations in companion sensor availability. Its lightweight design allows deployment on resource-constrained devices as well as fog or cloud platforms, depending on application requirements and available resources. Figure 1 shows these deployment options. The remainder of this section describes the virtual sensor module and the training and calibration procedures that govern its operation.

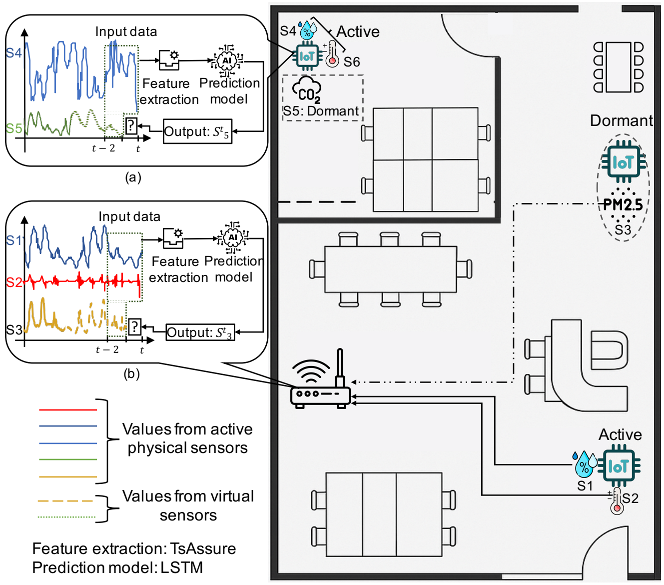


Fig. 1: Fig. 1: Flexible deployment architecture of the proposed virtual sensing framework. The model can be deployed either on the edge or in the cloud/fog, depending on application requirements. In the edge deployment (a), a lightweight TinyML model runs directly on the microcontroller, autonomously reconstructing readings from a dormant sensor (e.g., CO_2) using only an on-device companion sensor (e.g., humidity). In the cloud/fog deployment (b), prediction models run on a centralized server, enabling estimation of sensor readings (e.g., PM2.5) using data from multiple spatially distributed sensors (e.g., humidity and temperature from different devices).

A. Virtual Sensor Module

The virtual sensor module predicts the measurements of a physical sensor while it is intentionally deactivated. It leverages two complementary sources of information:

- Temporal patterns: By analyzing recent changes in a sensor's measurements, the module captures short-term temporal dynamics and extrapolates the sensor's likely future readings over a short time horizon.

- Cross-sensor relationships: By analyzing how readings from other sensors that are closely related to the target sensor change together, the module captures complementary aspects of the system's current state.

Both sources are necessary because temporal information alone is sufficient for short deactivation periods, whereas accurate predictions over longer intervals or under highly variable conditions require the inclusion of cross-sensor relationships.

For each target sensor s_i, the virtual sensing procedure begins by leveraging its recent measurements to capture temporal patterns. At each time step t, we consider the two most recent available values, $s_i(t-1)$ and $s_i(t-2)$, which may be either measured or previously predicted. From these values, we derive temporal features that explicitly encode short-term dynamics, including differences between consecutive readings and their rates of change. These features capture both the magnitude and direction of recent variations, enabling the model to extrapolate the sensor's likely value at the next time step. To complement temporal information, the module leverages measurements from related sensors in the network. We identify a set of *companion sensors* $\mathcal{C}_i$ whose historical measurements exhibit strong correlation with s_i. Correlations between the target sensor and potential companion sensors are quantified using the Pearson coefficient. Sensors with an absolute correlation value of $|r| \geq \tau_r$ (default $\tau_r = 0.7$) are selected as companions. Companion sensors may reside on the same device, on nearby edge nodes, or originate from external sources such as local weather stations. Measurements from these sensors are used to extract spatial features that show the inter-sensor deviation score and also the rate of change between their consecutive values, providing complementary context to improve predictions.

Both temporal and spatial features are computed using TsAssure [7], which transforms the raw sensor measurements into structured spatio-temporal input vectors. These feature vectors are then fed into a compact LSTM-based model with 100 hidden units, followed by a dense output layer. This architecture prioritizes simplicity and effectiveness while remaining lightweight enough for deployment on resource-constrained IoT edge devices. The model outputs the reconstructed sensor value $\tilde{s}_i(t)$ at each time step. To maintain robust predictions when companion sensors are intermittently unavailable, the system employs an adaptive multi-model strategy, which is detailed in the next subsection.

B. Training and Calibration Phases

In practical deployments, the availability of companion sensors can vary over time. The sensors themselves that were considered as companions may also be deactivated for various reasons, including temporary on-demand deactivation, transient communication failures, or sensor faults. To ensure reliable virtual sensing under these conditions, the proposed system adopts an adaptive multi-model strategy. It means rather than relying on a single predictor that assumes all companion sensors are simultaneously active, a small bank of lightweight LSTM-based predictors is trained for each target sensor. Each predictor corresponds to a representative subset of companion sensors expected to be available during operation. At runtime, the system selects the predictor whose input configuration matches the currently active sensors, enabling robust reconstruction under dynamic sensing availability while maintaining minimal computational overhead.

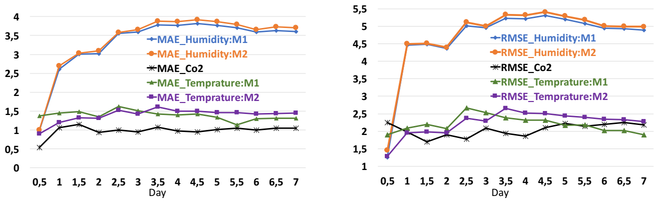
Following training, a calibration procedure is performed to estimate a safe continuous sensor deactivation period. For each trained predictor, sensor dormancy is progressively emulated over increasing time horizons within the available calibration dataset. During each emulated interval, the predictor reconstructs the sensor signal while the corresponding ground-truth measurements are retained to quantify reconstruction error. The resulting errors are evaluated as a function of deactivation duration and compared against predefined acceptability thresholds derived from application requirements or manufacturer-specified sensor error-margins. The longest interval that meets these criteria within the calibration data defines the maximum validated deactivation period for each sensor. This value serves as an operational configuration parameter to limit how long a sensor may remain continuously inactive.

III. Experimental Case Studies

In this section, we evaluate the proposed virtual sensing paradigm with respect to both prediction performance and energy efficiency. The evaluation is conducted using real-world datasets collected from operational IoT deployments. These datasets originate from two environmental monitoring applications: a smart farming deployment and a smart building deployment. This experimental setup enables a realistic assessment of the trade-offs between reconstruction accuracy and energy savings in practical IoT scenarios.

The smart building deployment used an ESP32 microcontroller equipped with an SCD30 CO_2 sensor (energy-intensive) and a DHT11 temperature–humidity sensor (low-power). To reduce energy consumption, the CO_2 sensor was periodically deactivated, and its readings were reconstructed using the virtual sensing model. Correlation analysis during training showed that CO_2 was strongly correlated only with humidity. Accordingly, the model uses short-term historical values of both CO_2 and humidity, along with the most recent humidity measurement, as input features to predict CO_2 levels. The trained LSTM predictor was converted to a TinyML format of 330 KB and deployed directly on the ESP32, enabling on-device inference and autonomous reconstruction of CO_2 readings whenever the physical sensor was inactive.

The smart farming deployment consisted of two edge devices, each based on an ESP32 microcontroller, positioned in close proximity. Edge device A was equipped with only a temperature sensor, while edge device B included both temperature and humidity sensors. In addition, a weather station located approximately 2.5 km away provided temperature and humidity data. To emulate energy-aware operation, edge device B was fully deactivated, and its temperature and humidity readings were predicted using correlated data from edge device A and the weather station. Because multiple correlated



(a) (b)

Fig. 2: Prediction error as a function of deactivation duration. (a) MAE and (b) RMSE for three target sensors: CO_2 (smart building, predicted from co-located humidity sensor), and humidity and temperature (smart farming, using M1 and M2 configurations). M1 uses the temperature from the nearby edge device, combined with temperature and humidity from the weather station; M2 uses only the temperature from the nearby edge device. Errors remain within sensor-specific tolerances across 7-day deactivation periods.

inputs were available, two input subsets were evaluated for each target sensor (temperature and humidity): a richer subset (M1) that included temperature from edge device A and both temperature and humidity from the weather station, and a lean subset (M2) that included only temperature from edge device A. Accordingly, four LSTM predictors were trained.

For both experiments, a 21-day training period was used to learn representative system dynamics and train the prediction models. This was followed by a calibration phase where selected sensors were intentionally deactivated for increasing durations, and their readings were reconstructed using correlated companion sensors. Figure 2 demonstrates the prediction accuracy for these two experiments across dormancy horizons using mean absolute error (MAE) and root mean square error (RMSE). It can be seen that across all datasets, prediction errors increase with longer dormancy phases but stabilize at extended horizons rather than diverging. In this study, due to deployment and data-collection constraints, the datasets used for calibration tests span only 7 days, which therefore defines the maximum dormancy horizon that can be empirically assessed in this study. It can be seen that reconstructed sensor readings remain within acceptable error bounds for deactivation periods of up to 7 days. These bounds are evaluated relative to manufacturer-specified sensor error margins (±5% for humidity, ±2 °C for temperature, and ±3% for CO_2), indicating that the reconstruction accuracy remains within limits. Another observation is that, in the smart farming deployment, both M1 and M2 models exhibit comparable error behavior across all tested horizons, demonstrating robustness to different companion-sensor configurations.

Accordingly, a 7-day deactivation period is considered a safe upper bound for sensor dormancy in this study, and energy consumption was evaluated over this interval. For both applications, all sensors sampled every 10 minutes and transmitted data via Wi-Fi to the cloud. Comparing scenarios with and without virtual sensing, we observed that in the smart building deployment, deactivating the CO_2 sensor for 7 days yielded approximately 30% energy savings, while in the smart farming scenario, putting edge device B into deep sleep for 7 days resulted in more than 90% energy savings. Table I summarizes the corresponding energy consumption values for both applications. These results demonstrate that the virtual sensing paradigm can significantly reduce energy usage while maintaining accurate reconstruction of deactivated sensors.

TABLE I: Energy consumption comparison with and without the virtual sensing framework.

Experiment		**Energy usage (mA)**	**Duration**
Smart building	Without virtual sensing	168	7 days
	Virtual sensing	119.04	
Smart farming	Without virtual sensing	110,88	7 days
	Virtual sensing	1	

IV. Conclusion

This paper introduced a virtual sensing framework that enables energy-efficient operation by reconstructing measurements during planned sensor deactivation. The approach relies on a recursive prediction mechanism that combines short-term temporal features of the target sensor with data from correlated active sensors, allowing efficient runtime execution on resource-constrained devices. To ensure robustness under varying sensor availability, a multi-model strategy was adopted in which separate predictors are trained for all viable subsets of correlated companion sensors. Also, a calibration study was conducted to quantify how reconstruction accuracy evolves with increasing deactivation duration, providing guidance for selecting safe dormant periods. The effectiveness of the proposed framework was validated in two real-world environmental monitoring scenarios, demonstrating its potential for reducing energy consumption. For future work, we plan to extend the framework to automatically select an optimal subset of correlated sensors in scenarios with many potential correlations, enhancing scalability in larger sensing deployments.

Acknowledgment

The authors acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG) for the Resilient Worlds program, specifically under project number 503199853.

References

[1] A. Narayan, B. H. Hassan, S. Attarha, C. Krüger, D. Babazadeh, and S. Lehnhoff, "Grid function virtualization for reliable provision of services in cyber-physical energy systems," in *2020 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2020, pp. 1–5.

[2] D. Ürge-Vorsatz, L. F. Cabeza, S. Serrano, C. Barreneche, and K. Petrichenko, "Heating and cooling energy trends and drivers in buildings," *Renewable and Sustainable Energy Reviews*, vol. 41, pp. 85–98, 2015.

[3] A. Kowli, V. Rani, and M. Sanap, "Data-driven virtual sensing for spatial distribution of temperature and humidity," *Journal of Building Engineering*, vol. 67, p. 105726, 2023.

[4] M.-S. Go, J. H. Lim, and S. Lee, "Physics-informed neural network-based surrogate model for a virtual thermal sensor with real-time simulation," *International Journal of Heat and Mass Transfer*, vol. 214, p. 124392, 2023.

[5] M. V. d. S. Lemos, R. H. Filho, R. d. A. L. Rabelo, C. G. N. d. Carvalho, D. L. d. S. Mendes, and V. d. G. Costa, "An energy-efficient approach to enhance virtual sensors provisioning in sensor clouds environments," *Sensors*, vol. 18, no. 3, p. 689, 2018.

[6] S. Attarha and A. Förster, "Building resilient iot systems through resource-efficient sensor fault correction," in *2025 IEEE Annual Congress on Artificial Intelligence of Things (AIoT)*. IEEE, 2025.

[7] S. Attarha and A. Förster, "Assuresense: A framework for enabling sensor fault detection in low-power iot edge devices," *IEEE Sensors Journal*, 2024.