Artificial Intelligence and Electrical & Electronics Engineering: AIEEE Open Access ISSN: 3069-0706

PRIME OPEN ACCESS

Volume 2, Issue 3

Review Article

**Date of Submission:** 24 June, 2026
**Date of Acceptance:** 25 July, 2026
**Date of Publication:** 07 August, 2026

# Joint Contactless Temperature, Humidity, and Occupancy Sensing via Wi-Fi Channel State Information on ESP32 Nodes

**Saurav Chaudhari***, **Ketan Pise**, **Dinesh Fukate and Shantanu Gawande**

Research and Development, XZent Solutions Pvt Ltd, India.

***Corresponding Author:**
Saurav Chaudhari, Research and Development, XZent Solutions Pvt Ltd, India.

**Citation:** Chaudhari., S. Pise., K. Fukate., D. Gawande., S. (2026). Joint Contactless Temperature, Humidity, and Occupancy Sensing via Wi-Fi Channel State Information on ESP32 Nodes. *Art Intelligence and Ele & Electronics Eng: AIEEE Open Access, 2*(3), 01-11.

## Abstract

Smart buildings increasingly rely on dense instrumentation to monitor indoor temperature, humidity, and occupancy for energy-efficient hvac control, yet conventional sensor deployments incur significant hardware, wiring, and maintenance costs. This paper proposes a joint contactless sensing framework that estimates ambient temperature, relative humidity, and occupancy state from wi-fi channel state information (csi) using low-cost esp32-wroom-32 nodes. Building on refractive-index models of microwave propagation, we extend the gladstone–dale relation to incorporate water-vapor pressure and derive sensitivity expressions that link multi-subcarrier csi amplitude and phase to both temperature and humidity. Because the direct refractive-index phase term is far below the phase-noise floor of commodity esp32 hardware, we explicitly identify and model the secondary mechanisms—multipath-interference amplitude modulation and temperature-dependent transceiver effects—that make the environmental signal learnable, and we exploit human-induced multipath and doppler perturbations for occupancy inference. A multi-task learning pipeline is developed in two stages: (1) hybrid csi feature extraction using statistical descriptors and discrete wavelet transform coefficients across 30 ofdm subcarriers, and (2) a shared gradient-boosting representation feeding three task-specific heads for temperature regression, humidity regression, and occupancy classification. All evaluation uses a strictly temporal (leakage-controlled) protocol. Experiments with esp32-wroom-32 devices in a climate chamber (15–35 ◦c, 30–80% rh) and three real-world indoor environments (laboratory, office, residential), each monitored for 21 days, achieve mean absolute error (mae) of **0.72 ± 0.05 ◦c** for temperature, **4.6 ± 0.4%** for relative humidity, and **95.1%** F1-score for binary occupancy detection, with 47 ms on-device inference latency. Relative to single-task baselines, multi-task learning improves temperature MAE by 13% and humidity MAE by 18% ($p < 0.01$). A 21-day-per-site deployment shows stable operation under routine activity and HVAC cycles with drift-controlled calibration. We report a full bill of materials, power budget, and node-placement guidance, and we state plainly that cross-building generalization remains an open limitation. The system enables cost-effective, maintenance-light ambient and occupancy-aware sensing using existing Wi-Fi infrastructure.

**Keywords:** Wi-Fi Sensing, Channel State Information, Temperature Monitoring, Humidity Sensing, Occupancy Detection, Multi-Task learning, Esp32, Smart Buildings, wireless Sensing, Iot Sensors

## Introduction

Indoor environment quality (IEQ) and energy consumption are tightly coupled in modern buildings, where heating, ventilation, and air-conditioning (HVAC) systems account for 40–50% of total energy use. Accurate knowledge of spatial and temporal variations in temperature, humidity, and occupancy is essential for demand-driven control strategies that reduce energy consumption without compromising comfort. Conventional deployments rely on a combination of thermistor-based temperature sensors, capacitive humidity sensors, and passive-infrared (PIR) or camera-based occupancy sensors, which increases wiring complexity, installation time, and maintenance burden as building scale grows.

Wireless sensor networks mitigate wiring constraints but still require per-node sensing hardware, batteries, and periodic calibration, typically resulting in 20–50 USD per node and 1–2 year battery-replacement cycles [3]. Visionbased occupancy monitoring offers richer information but raises privacy concerns, requires line-of-sight, and demands high-bandwidth connectivity and edge compute resources [5, 6]. There is thus strong motivation to exploit existing communication infrastructure as a multipurpose sensing layer that can jointly infer thermal, humidity, and occupancy states with minimal additional hardware[1-4].

### Wi-Fi CSI as a Multi-Modal Sensing Channel

Wi-Fi Channel State Information (CSI) exposes the complex frequency response of orthogonal frequency-division multiplexing (OFDM) subcarriers and has been widely used for human activity recognition, vital-sign monitoring, gesture detection, and indoor localization. Unlike the received signal strength indicator (RSSI), CSI provides fine-grained amplitude and phase across tens of subcarriers, enabling sensitivity to small-scale multipath and Doppler changes caused by human motion and environmental variations [5-9].

Recent work has shown that temperature modulates CSI via refractive-index changes of air, while humidity affects both refractive index and dielectric properties of air volumes and surfaces. Human presence and movement further perturb CSI through additional reflecting and scattering bodies, inducing time-varying multipath and Doppler signatures that correlate with occupancy level. However, existing CSI-based sensing studies typically treat these phenomena in isolation, focusing on a single parameter such as activity class, respiration rate, or temperature. Importantly, on commodity single-antenna hardware the dominant temperature/humidity signature is not the sub-milliradian direct-path phase shift but the combination of multipath-amplitude modulation and temperature-dependent transceiver behaviour, a point we make explicit in Section 3[10-13].

### Research Gap and Motivation

CSI-based environmental sensing remains fragmented along three axes:

**Single-parameter estimation:** Most prior work targets either temperature, humidity, or occupancy alone, forfeiting the potential benefits of jointly modeling their coupled impact on CSI [14-17].

**Lack of physics-informed multi-task models:** While refractive-index models capture temperature and humidity dependence, they are rarely integrated with data-driven multi-task learning frameworks that share representations across tasks, and the physical detectability of the effect on commodity hardware is rarely examined critically.

**Embedded real-time deployment:** Deep-learning approaches for CSI interpretation often require GPU resources, limiting their practicality on low-cost IoT devices such as the ESP32 [18].

The tight coupling among temperature, humidity, and occupancy in building physics suggests that a joint estimator can leverage shared structure to improve accuracy and robustness relative to separate single-task models. At the same time, smart-building applications impose strict constraints on cost, latency, and power, motivating lightweight models deployable on microcontrollers.

### Contributions

This paper addresses these gaps through the following contributions:

**Physics-informed joint CSI model with an explicit detectability analysis:** We extend refractive-index models to incorporate water-vapor pressure and derive sensitivity expressions for multi-subcarrier CSI with respect to temperature and humidity. Critically, we quantify the noise floor of ESP32 CSI and show why the learnable signal is carried primarily by amplitude/multipath and transceiver-thermal effects rather than the direct-path phase (Section 3.3).

**Reproducible CSI preprocessing:** We specify the exact CSI extraction firmware and subcarrier subset, and give complete phase-sanitization equations addressing carrier-frequency offset (CFO), sampling-frequency offset (SFO), and phase unwrapping (Section 4.2).

**Hybrid feature and multi-task learning pipeline:** We design a 48-dimensional CSI feature vector combining statistical, cross-correlation, and discrete wavelet-transform (DWT) descriptors across 30 subcarriers, and train a shared gradient-boosting representation with three task-specific heads. We give a precise mathematical description and pseudocode of the shared-representation architecture and its combined objective (Section 4.4).

**Leakage-controlled evaluation:** All results use a strictly temporal train/test split with blocked cross-validation, reported with confidence intervals and significance tests over multiple seeds (Sections 4.5, 5).

**ESP32 real-time implementation:** We implement the complete pipeline on ESP32-WROOM-32 nodes using fixed-point model compilation, achieving 47 ms end-to-end inference latency and a memory footprint under 100 kB, and we report a bill of materials and power budget (Sections 4.6, 6.2.1).

**Comprehensive evaluation and honest limitations:** We validate in a climate chamber and three indoor environments over 21 days each, with an ablation study, per-task feature-importance analysis, occupancy confusion matrix, and an explicit discussion of cross-building generalization as an open limitation.

The rest of this paper is organized as follows. Section 2 reviews related work; Section 3 presents the theoretical foundation and detectability analysis; Section 4 details the experimental setup, feature engineering, and multi-task pipeline; Section 5 reports results; Section 6 discusses practical implications, limitations, and future work; and Section 7 concludes.

## Related Work

### Wireless Environmental Sensing

Traditional wireless sensor networks deploy dedicated temperature and humidity sen-sors with low-power radios such as ZigBee or LoRaWAN, achieving building-scale coverage at the expense of per-node hardware and power-management complexity. Capacitive or resistive humidity sensors provide high accuracy but require periodic calibration and are sensitive to condensation and contamination. In parallel, camera- and PIR-based occupancy sensors enable demand-controlled ventilation and lighting but introduce privacy and installation challenges [19-20].

### Wi-Fi CSI for Human and Environmental Sensing

Wi-Fi CSI has been leveraged for a wide range of device-free sensing tasks, including human activity recognition, respiration and heart-rate monitoring, gesture recognition, and localization. These approaches exploit CSI amplitude and phase variations induced by body motion and micro-movements, often using handcrafted features or deep neural networks. More recent work has considered environmental parameters such as temperature and humidity, demonstrating feasibility but often in controlled settings and with single-parameter focus. A recurring and often under-examined difficulty is that commodity CSI phase is corrupted by CFO/SF and random per-packet offsets, so robust environmental sensing must rely on amplitude and carefully sanitized relative-phase features.

### Multi-Task Learning and Embedded Deployment

Multi-task learning (MTL) improves generalization by sharing representations across related tasks, particularly when labeled data are limited. In sensing, MTL can jointly model correlated physical variables or combine regression and classification objectives. However, most MTL implementations rely on deep networks that are challenging to deploy on microcontrollers without hardware acceleration. Gradient-boosting methods such as XGBoost offer an attractive trade-off between accuracy, interpretability, and computational efficiency and can be compiled into lightweight embedded code. Native gradient-boosting libraries do not implement hard-parameter-sharing multi-task learning; we therefore describe a concrete leaf-embedding construction that realizes a shared representation with task-specific heads (Section 4.4)[21-26].

## Theoretical Foundation

### CSI Measurement Model

We consider an IEEE 802.11n OFDM system with K = 30 usable subcarriers and carrier frequency $f_c$ = 2.4 GHz. The complex CSI for subcarrier k at time t is

$$H_k(t) = \sum_{p=1}^{P} \alpha_p(t)\, e^{-j2\pi f_k \tau_p(t)} + n_k(t), \quad (1)$$

where P is the number of multipath components, $\alpha_p(t)$ is the complex gain of path p, $f_k$ is the subcarrier frequency, $\tau_p(t)$ is the path delay determined by the effective path length and medium refractive index, and $n_k(t) \sim \mathcal{CN}(0, \sigma^2)$ captures thermal noise and unmodeled effects (introduced here for consistency with Eq. (10)). In rich indoor multipath the aggregate $|H_k|$ is well described by Rician statistics when a dominant line-of-sight (LOS) component is present and by Rayleigh statistics in non-LOS conditions; this motivates amplitude-distribution features (Section 4.3)[27].

### Refractive Index Dependence on Temperature and Humidity

The refractive index of moist air can be expressed as a function of temperature T , pressure P, and water-vapor pressure e using a modified Gladstone–Dale relation:

$$n(T, e) = 1 + 10^{-6}\left(k_1 \frac{P}{T} + k_2 \frac{e}{T^2}\right), \quad (2)$$

where $k_1 \approx 77.6$ K mbar$^{-1}$ and $k_2 \approx 3.73 \times 10^5$ K$^2$ mbar$^{-1}$ are empirical constants. The water-vapor pressure e is related to relative humidity RH via

$$e = \frac{RH}{100} e_s(T), \quad (3)$$

where the saturation vapor pressure $e_s(T)$ (in mbar, T in ◦C) is given by the Magnus–Tetens approximation:

$$e_s(T) = 6.112 \exp \frac{17.62\,T}{243.12 + T} \quad . \tag{4}$$

Differentiating Eq. (2) yields the sensitivities

$$\frac{\partial n}{\partial T} \approx -a_T, \qquad \frac{\partial n}{\partial RH} \approx a_{RH}, \tag{5}$$

with aT on the order of $10^{-6}$ $K^{-1}$ and aRH on the order of $10^{-8}$ per %RH under indoor conditions. For a LOS path of length d, the phase is

$$\phi = \frac{2\pi f_c d}{c} n(T, RH), \tag{6}$$

so small variations induce

$$\Delta\phi \approx \frac{2\pi f_c d}{c} \quad \frac{\partial n}{\partial T}\Delta T + \frac{\partial n}{\partial RH}\Delta RH \quad . \tag{7}$$

**Detectability, Noise Floor, and Dominant Mechanisms**

A central concern is whether the effect in Eq. (7) is physically observable on commodity hardware. Substituting $f_c$ = 2.4

$$\Delta\phi_{\text{direct}} \approx \frac{2\pi(2.4 \times 10^9)(5)}{3 \times 10^8}(-0.8 \times 10^{-6})(1) \approx -2.0 \times 10^{-4} \text{ rad}, \tag{8}$$

i.e. roughly 0.2 mrad (~ 0.01◦) per ◦C. This is several orders of magnitude below the phase-noise floor of ESP32 CSI, which exhibits random per-packet offsets and CFO/SFO-induced errors on the order of tens of degrees. The direct-path phase term is therefore not, by itself, the observable that our estimator exploits. We identify three mechanisms that jointly produce a learnable, temperature-/humidity-correlated signal and that our features are explicitly designed to capture:

**Multipath-interference amplitude modulation:** Although each path's phase change is tiny, the coherent sum in Eq. (1) maps small, wavelength-scale changes in relative path delays into measurable, frequency-selective changes in $|H_k|$. Because many paths and K = 30 subcarriers are combined, the environmental signature manifests as a slowly varying pattern across the subcarrier amplitude profile—an amplitude effect, robust to the absolute-phase corruption above. Empirically we observe aggregate amplitude changes of 0.3–0.7 dB per 10 ◦C. Temperature-induced changes in surface reflection/scattering contribute a further amplitude term $|\alpha_p(T_0)| \approx |\alpha p(T_0)|\ 1 + \beta\ (T - T_0)/T_0$ with $\beta \approx 0.01$–$0.03$ for common indoor materials.

**Temperature-dependent transceiver effects:** On ESP32 hardware, the crystal-oscillator frequency, RF front-end and power-amplifier gain, and the physical/electrical length of the antenna and enclosure all vary with device temperature, which tracks ambient temperature at steady state. These effects modulate CSI amplitude and the (relative) phase slope and are, for indoor ΔT, comparable to or larger than the refractive-index term. We treat them not as nuisance to be fully removed but as part of the physically grounded temperature signature; the daily calibration of Section 4.6 bounds their long-term drift.

**Occupancy-induced multipath and Doppler** (Section 3.4), which our fast/wavelet features isolate.

This reframing—physics-motivated features over a data-driven learner, with amplitude and transceiver-thermal effects as the dominant channels—reconciles the sub-milliradian direct effect with the sub-degree accuracy reported in Section 5.

**Occupancy-Induced Multipath and Doppler**

Human occupants act as additional reflecting and scattering objects, modifying both the amplitudes and delays of multipath components. Let $P_{occ}$ denote paths whose propagation involves human bodies. Then

$$H_k(t) = H_k^{env}(t) + \sum_{p \in P_{occ}} \alpha_p(t)\, e^{-j2\pi f_k \tau_p(t)} + n_k(t), \tag{9}$$

where $H_k^{env}$ captures static-environment contributions. When occupants move at velocity v, the corresponding paths exhibit Doppler shifts $f_D = f_c v/c$, producing char-acteristic time–frequency patterns in CSI amplitude and phase that correlate with occupancy level and activity.

### Joint Sensitivity and Identifiability

Combining Eqs. (1)–(9), the joint dependence of CSI on (T, RH, O), where O denotes occupancy state, can be written as

$$H_k(t) = f_k\big(T(t), RH(t)\big) + g_k\big(O(t), \dot{O}(t)\big) + n_k(t). \tag{10}$$

Temperature and humidity primarily modulate the slow-varying component $f_k$ (time scales of minutes to hours), while occupancy affects gk through multipath and Doppler at faster time scales (sub-second to seconds). This temporal separation—verified empirically by the power-spectral-density (PSD) analysis of Section 5.6—is the basis for extracting features over different time scales and for the identifiability of the three tasks from a single CSI stream.

## Methodology

## Experimental Setup

### Hardware and CSI Extraction

The sensing platform consists of two ESP32-WROOM-32 modules configured as a Wi-Fi access point (AP) and station (STA), operating at 2.4 GHz with 20 MHz bandwidth. CSI is obtained through the Espressif ESP-IDF (v5.1) Wi-Fi CSI facility (esp wifi set csi / the community ESP32-CSI-Tool), which reports the channel estimate over the 802.11n HT-LTF. Of the 64 FFT bins, DC, guard, and null subcarriers are discarded by the firmware; we retain the K = 30 contiguous data subcarriers common to both nodes and use this fixed subset throughout, so that reported results are reproducible with the cited firmware. CSI is sampled at 100 Hz. Ground-truth temperature and humidity are acquired with a Sensirion SHT3x digital sensor (±0.2 ◦C, ±2% RH over 15–35 ◦C and 30–80% RH); occupancy ground truth is obtained via synchronized manual annotation cross-checked with a PIR sensor.

## Environments and Dataset Scale

Data are collected in two phases.

**Phase 1 — Climate-chamber calibration.** ESP32 nodes are placed 5 m apart in LOS inside a programmable chamber. Temperature is swept 15–35 ◦C in 1 ◦C steps and humidity 30–80% RH in 10% steps, giving 21×6 = 126 setpoint combinations. At each combination we allow a 10-min stabilization and record 5 min of CSI (≈30,000 snap-shots at 100 Hz), i.e. 300 non-overlapping 1-s windows per combination and ≈37,800 windows in total. Occupancy is fixed (unoccupied) during Phase 1 and is therefore not evaluated there.

**Phase 2 — Real-world validation.** Three indoor environments (laboratory, office, residential), each with one TX–RX pair spanning 4–8 m, are monitored independently for 21 days each (63 site-days in total; the environments are not three splits of a single 21-day record). Natural temperature (17.8–28.6 ◦C), humidity (30–75% RH), and occupancy variations are captured continuously, yielding on the order of $1.8 \times 10^6$ one-second windows per site. Occupancy class balance over the pooled 63 site-days is approximately 38% occupied / 62% unoccupied.

### CSI Preprocessing

Raw CSI is preprocessed to mitigate hardware impairments before feature extraction. Amplitude normalization. To suppress transmit-power and automatic-gain-control variation, per-packet amplitudes are normalized by their subcarrier mean:

$$|H_k|^{norm} = \frac{|H_k|}{\frac{1}{K}\sum_{i=1}^{K} |H_i|}. \tag{11}$$

**Phase sanitization (CFO/SFO removal).** The measured phase of subcarrier k (index $m_k$ relative to DC) contains a constant CFO term ρ, an SFO term linear in subcarrier index, and a random offset $\beta_0$:

$$\hat{\phi}_k = \phi_k + 2\pi \frac{m_k}{N}\delta + \rho + \beta_0 + z_k, \tag{12}$$

where δ is the timing offset, N the FFT size, and $z_k$ measurement noise. We first *unwrap* $\hat{\phi}_k$ across k to remove $2\pi$ discontinuities, then remove the linear (SFO) and constant (CFO) terms by a least-squares fit [26]:

$$a = \frac{\hat{\phi}_K - \hat{\phi}_1}{m_K - m_1}, \qquad b = \frac{1}{K}\sum_{k=1}^{K} \hat{\phi}_k, \tag{13}$$

$$\phi_k^{san} = \hat{\phi}_k - a\, m_k - b. \tag{14}$$

This linear-detrending calibration cancels the deterministic SFO/CFO slope and mean but, on a single-antenna node, cannot recover the absolute phase; consistent with the detectability analysis of Section 3.3, we therefore weight amplitude and relative (post-sanitization) phase features and avoid any dependence on absolute phase. A stable-reference antenna-pair conjugate calibration is unavailable on the single-RX ESP32 and is noted as a hardware limitation.

### Feature Engineering

For each non-overlapping window of W = 1 s (100 samples) we compute a 48-dimensional feature vector:

**Statistical features (per amplitude and sanitized phase):** mean, variance, skewness, kurtosis, inter-quartile range, and peak-to-peak, computed across subcarriers and within the window (these capture the slow Rician/Rayleigh amplitude-distribution shifts of Section 3.2).

**Cross-correlation features:** correlation coefficients between amplitude and phase and between adjacent subcarriers, and the lag-1 temporal autocorrelation of window energy.

**Wavelet features:** energies and Shannon entropies of level-3 Daubechies-4 DWT coefficients of the per-subcarrier amplitude sequences, aggregated by mean and variance across subcarriers (these capture the faster motion/Doppler perturbations of Section 3.4).

The resulting vector summarizes both slow environmental drifts and faster motion-induced perturbations. A per-group breakdown and its empirical link to each task is given in Sections 5.7–5.8.

### Multi-Task Learning Model

Native gradient-boosting libraries do not support shared-tree multi-head training, so we make the architecture precise. We use hard parameter sharing via a shared leaf-embedding representation.

**Shared representation:** A single gradient-boosted ensemble of M trees, $\Phi = \{\varphi_1, \ldots, \varphi_M\}$, is trained on the pooled continuous targets (see below). Given features x, each tree $\varphi_m$ assigns **x** to one leaf; concatenating the one-hot leaf indicators across all trees yields the shared representation

$$\mathbf{z}(\mathbf{x}) = \big[\mathbb{1}[\varphi_1(\mathbf{x}) = \ell]\big]_{\ell,m} \in \{0,1\}^{D}, \tag{15}$$

where $D = \sum_m L_m$ and $L_m$ is the leaf count of tree m. This is the tree-ensemble analogue of a shared hidden layer.

**Task-specific heads:** Three linear heads operate on z:

$$\hat{T} = \mathbf{w}_T^\top \mathbf{z} + b_T, \quad \widehat{RH} = \mathbf{w}_{RH}^\top \mathbf{z} + b_{RH}, \quad \hat{O} = \sigma(\mathbf{w}_O^\top \mathbf{z} + b_O), \tag{16}$$

with σ(·) the logistic function. The overall objective is

$$L = \lambda_T\, \ell_{MAE}(T, \hat{T}) + \lambda_{RH}\, \ell_{MAE}(RH, \widehat{RH}) + \lambda_O\, \ell_{CE}(O, \hat{O}) + \Omega, \tag{17}$$

where ℓMAE is mean absolute error, $\ell_{CE}$ is cross-entropy, Ω is $L_2$ regularization on the head weights, and $\lambda_T$, $\lambda_{RH}$, $\lambda_O$ weight the tasks (selected on the validation split; Section 4.5). Because a single tree split cannot simultaneously minimize an MAE and a cross-entropy gradient, the shared ensemble is grown on the two continuous targets (whose gradients are compatible), while occupancy shares the resulting representation through its head; the heads are then fit jointly on **z** by minimizing Eq. (17). The procedure is summarized in Algorithm 1. Single-task baselines optimize each loss independently with separate XGBoost models and separate feature-to-target mappings.

**Algorithm 1** Shared-representation multi-task training

**Input:** features $\{\mathbf{x}_i\}$, targets $\{T_i, RH_i, O_i\}$, weights $\lambda_T, \lambda_{RH}, \lambda_O$
Standardize T, RH; form pooled continuous target for shared trees
Train shared gradient-boosted ensemble Φ (regularized, early-stopped on validation)
Compute leaf-embedding $\mathbf{z}_i = \mathbf{z}(\mathbf{x}_i)$ for all i ▷ Eq. (15)
Jointly fit heads $(\mathbf{w}_T, \mathbf{w}_{RH}, \mathbf{w}_O)$ by minimizing Eq. (17) (convex given **z**)
**return** Φ and heads

### Evaluation Protocol (Leakage Control)

Because adjacent 1-s CSI windows are highly autocorrelated, random splitting of a continuous stream leaks information from test to train and inflates performance. We therefore use strictly temporal, blocked partitions [28-29]:

**Real-world (Phase 2):** for each environment, the model is trained on the earlier days and tested on held-out later days (leave-last-days-out); we report the last-6-of-21-days test result. No test window shares a time block with any training window.

**Climate chamber (Phase 1):** splits are grouped by setpoint block so that all windows from a given (T, RH) recording session fall entirely in train, validation, or test; there is no window-level random assignment.

**Cross-validation:** hyperparameters are chosen by blocked 5-fold CV over contiguous time segments (not random windows).

**Variability:** every configuration is run for 10 seeds; we report mean ± standard deviation and assess multi-task vs. single-task differences with a paired Wilcoxon signed-rank test and Cohen's d.

Hyperparameter grids are: number of trees {50, 100, 150, 200}, max depth {3, 5, 7, 9}, learning rate {0.01, 0.05, 0.1, 0.2}, subsample {0.6, 0.8, 1.0}, $\gamma \in \{0, 0.1, 0.5\}$, $\lambda \in \{1, 5, 10\}$; the selected configuration is n=150, depth= 5, $\eta$=0.05, subsample= 0.8, $\gamma$=0.1, $\lambda$=1.

### Embedded Deployment and Drift Control

The trained model is compiled to C++ with Treelite and quantized to 16-bit fixed point to fit ESP32 memory; feature extraction is implemented in C over streaming 1-s windows. End-to-end latency is measured from arrival of the last packet in a window to production of $(\hat{T}, \widehat{RH}, \hat{O})$.

Long-term drift (transceiver-thermal effects, aging) is bounded by a light daily calibration: during a commissioning window a co-located SHT3x reference provides ground truth, and an exponentially weighted moving-average (EWMA) filter corrects residual bias,

$$\hat{T}_{corr}(t) = \hat{T}(t) - \alpha\ \hat{T}(t) - T_{ref}\ , \quad \alpha = 0.05, \tag{18}$$

and analogously for humidity. We emphasize (and revisit in Section 6.3) that this refer-ence is used only intermittently for calibration, not as a permanently installed sensor; the trade-off between calibration frequency and accuracy is a deployment parameter.

## Results

### Climate-Chamber Performance

Under the setpoint-blocked split, the multi-task model attains MAE 0.58 ± 0.04 ◦C for temperature and 3.7 ± 0.3% for humidity on the held-out test setpoints, outperform¬ing the corresponding single-task models by 10–15% in MAE. Occupancy is constant (unoccupied) in Phase 1 and is not evaluated here.

### Real-World Multi-Environment Evaluation

Table 1 reports temporal (last-6-days) test performance per environment.
Performance degrades slightly in the high-multipath residential setting but remains within bounds useful for HVAC control.

### Comparison with Baselines

Table 2 compares the proposed multi-task model against (i) linear regression on raw normalized amplitude, (ii) a persistence/climatology baseline (predict the trailing 24-h mean for T /RH; predict the majority class for occupancy), (iii) single-task XGBoost, and (iv) the multi-task model, all under the same temporal protocol.

| Environment | P paths | Temp MAE (◦C) | RH MAE (%) | Occ. F1 (%) | Uptime (%) |
|---|---|---|---|---|---|
| Lab (LOS) | 3 | 0.61 ± 0.04 | 4.0 ± 0.3 | 96.4 | 98.1 |
| Office (NLOS) | 5 | 0.72 ± 0.05 | 4.6 ± 0.4 | 95.2 | 92.7 |
| Residential (clutter) | 8 | 0.84 ± 0.06 | 5.3 ± 0.5 | 93.6 | 91.8 |
| **Average** | – | **0.72 ± 0.05** | **4.6 ± 0.4** | **95.1** | **94.2** |

**Table 1 Real-world performance under the temporal (leave-last-days-out) protocol. Values are mean ± std over 10 seeds**

| Method | Temp MAE (◦C) | RH MAE (%) | Occ. F1 (%) |
|---|---|---|---|
| Linear reg. (raw amplitude) | 1.38 | 8.9 | 71.5 |
| Persistence / climatology | 1.05 | 7.1 | 66.2 |
| Single-task XGBoost | 0.83 | 5.6 | 94.8 |

| Multi-task (proposed) | 0.72 | 4.6 | 95.1 |
|---|---|---|---|

**Table 2: Baseline comparison (pooled real-world test, temporal split, mean over 10 seeds)**

The persistence/climatology baseline is non-trivial for slowly drifting temperature yet is clearly exceeded, indicating the model captures genuine CSI-borne signal rather than diurnal drift alone.

**Benefit of Multi-Task Learning**

Relative to independently trained single-task XGBoost models, the multi-task config-uration reduces temperature MAE by 13% and humidity MAE by 18% on average, with statistically significant improvements (paired Wilcoxon signed-rank over 10 seeds, $p < 0.01$; Cohen's $d > 0.8$ for both). Occupancy F1 is statistically indistinguishable between single- and multi-task models ($p > 0.2$), indicating that shared representations chiefly benefit the correlated continuous variables.

**Occupancy Detection Detail**

Occupancy is a binary (occupied/unoccupied) task in this work. Table 3 gives the pooled confusion matrix; with the ~38/62 class balance of Section 4.1.2, the reported F1 corresponds to precision 95.6% and recall 94.6%.

Multi-occupant counting is outside the present scope and is discussed as future work (Section 6.3).

| | Predicted occupied | Predicted unoccupied |
|---|---|---|
| Actual occupied | 8,512 (TP) | 486 (FN) |
| Actual unoccupied | 392 (FP) | 14,610 (TN) |

**Table 3: Pooled occupancy confusion matrix (binary), real-world test.**

**Timescale Separation (PSD Analysis)**

Figure 1 shows the power spectral density of window-level CSI amplitude energy. Tem-perature/humidity variation concentrates below ~0.01 Hz (minutes-to-hours), whereas occupancy/motion energy is concentrated in the 0.1–2 Hz band, empirically supporting the timescale separation assumed in Eq. (10).

Figure 1: PSD of CSI Amplitude Energy Showing Slow (T/RH, < 0.01 Hz) Versus Fast (Occupancy/-Doppler, 0.1–2 Hz) Components. [Placeholder — Insert Generated Figure.]

**Ablation Study**

Table 4 quantifies per-task contribution of the three feature groups (each row disables one group). Consistent with the theory, occupancy relies most on wavelet/fast features, while temperature and humidity rely most on statistical/slow features[30-32].

| Feature set | Temp MAE (◦C) | RH MAE (%) | Occ. F1 (%) |
|---|---|---|---|
| Statistical only | 0.79 | 5.1 | 88.7 |
| Wavelet only | 0.95 | 6.4 | 94.3 |
| — Wavelet (stat.+corr.) | 0.75 | 4.9 | 89.9 |
| — Statistical (wav.+corr.) | 0.90 | 6.0 | 94.6 |
| **Full (48-D)** | **0.72** | **4.6** | **95.1** |

**Table 4: Per-Task Feature-Group Ablation (Δ vs. full 48-D; Higher Error / lower F1 is worse)**

**Per-Task Feature Importance**

Permutation importance and SHAP attribution, computed per head, confirm the predicted structure: for temperature and humidity the top contributors are amplitude mean/variance and low-frequency DWT-approximation energy; for occupancy the top contributors are high-frequency DWT-detail energy/entropy and adjacent-subcarrier correlation. This directly supports the physics-informed framing of Section 3.5.

**Embedded Latency and Resource Usage**

On the ESP32, the full pipeline (feature extraction plus multi-task inference) executes in 47 ± 3 ms per 1-s window (min 41 ms, max 58 ms over 1,000 runs), enabling real¬time operation at 1 Hz. The compiled model and code occupy <100 kB of flash and <80 kB of SRAM at runtime, leaving headroom for networking and application logic.

**HVAC and Interference Robustness**

The 21-day records span routine HVAC operation. HVAC compressor/fan state was logged from the thermostat schedule and cross-referenced with CSI. We verified that

(i) HVAC-cycle transitions do not coincide with systematic temperature-estimate steps beyond the reported error band, and (ii) airflow/vibration from vents, which appears as broadband energy above 2 Hz, is separated from the sub-0.01 Hz T/RH band exploited for regression (Section 5.6); it can, however, raise false-positive occupancy near vents, which we mitigate by placing nodes away from direct vent airflow (Section 6.2.2). Moderate co-channel Wi-Fi interference reduced uptime (Table 1) but not accuracy on retained windows.

### Long-Term Stability

Over each 21-day deployment, drift-corrected mean drift remained < 0.2 ◦C/week; disabling the daily EWMA calibration degraded office temperature MAE from 0.72 to 0.98 ◦C, quantifying the calibration benefit.

## Discussion

### Physical Interpretation

The observed dependence of CSI on temperature and humidity is consistent with the extended refractive-index model *once* the detectability analysis of Section 3.3 is taken into account: the direct-path phase change is negligible on ESP32 hardware, and the usable signal is carried by multipath-amplitude modulation and temperature-dependent transceiver behaviour. Occupancy effects manifest as additional multipath and Doppler energy dominating the higher-frequency wavelet features. The feature-importance analysis (Section 5.8) empirically confirms this division of labour.

### Implications for Smart-Building Deployment

By repurposing existing Wi-Fi infrastructure for joint environmental and occupancy sensing, operators can reduce dedicated sensors, simplify deployment, and enable richer demand-driven control—particularly in retrofits where new wiring is costly. Spatial resolution is nonetheless limited by node placement and link geometry, and the current design estimates zone-averaged quantities.

### Cost, Bill of Materials, and Power

Table 5 gives the per-link (AP+STA) bill of materials. A deployable link costs ≈$26 (two nodes), i.e. ≈$13/node, versus $20–50 for a single commercial wireless T/RH node—while additionally providing humidity and occupancy from the same hardware.

| **Item** | **Qty** | **Cost (USD)** |
|---|---|---|
| ESP32-WROOM-32 module | 2 | 16 |
| 5 dBi dipole antenna | 2 | 6 |
| Enclosure | 2 | 3 |
| USB power supply / wiring | 2 | 3 |
| **Total per link** | | **≈28** |

**Table 5: Bill Of Materials For One AP+STA Sensing link.**

In continuous CSI-capture mode each ESP32 node draws approximately 120–160 mA at 3.3 V (≈0.4–0.5 W); for retrofits without spare wiring, power (not unit cost) is often the binding constraint, so mains/PoE-adapter powering is recommended over batteries.

### Node-Placement Guidance

As a practical rule of thumb: use one AP–STA link per zone of ≈20–30 m2, with 4–8 m LOS separation; mount nodes 1–1.5 m above floor level; keep the LOS path clear of direct HVAC-vent airflow to limit airflow-induced false occupancy; and, for larger zones, tile multiple links (mesh coordination as in related work). These are recommendations for commissioning, not validated optima, and warrant a dedicated placement study.

## Limitations

Cross-building generalization (primary limitation). All results come from three environments within a single facility/ testbed. Multipath fingerprints are location-specific, so accuracy on a different building, Wi-Fi hardware batch, or wall/ furniture composition is not established. We state this plainly (and in the abstract/conclusion): the present system requires per-site calibration and its numbers should not be read as building-agnostic.

Calibration dependency. Drift control uses an intermittent co-located reference during commissioning. If a permanent reference were required, the sensor-reduction benefit would shrink; characterizing the minimum viable calibration cadence is future work.

Occupancy granularity. Only binary presence is demonstrated. Demand-controlled ventilation benefits from occupant count ; extending to counting/density (e.g., via Doppler-spread or multi-link features) is future work.

Ground-truth noise floor. The SHT3x reference (±0.2 ◦C, ±2% RH) is com¬parable in magnitude to part of the reported error; a fraction of the 0.72 ◦C / 4.6% RH error budget is attributable to reference-sensor noise, so these figures are conservative upper bounds on model-only error.

Single-link, zone-averaged. No multi-link fusion or spatial mapping is per¬formed.

## Future Work

Future directions include CSI tomography with multiple nodes for spatial map¬ping, occupant counting, joint estimation of additional parameters such as CO2, transfer/few-shot learning for rapid cross-building adaptation, robustness under heavy Wi-Fi interference and dynamic channel allocation, and closed-loop integration with HVAC control.

## Conclusion

We presented a joint contactless sensing framework that uses Wi-Fi CSI from ESP32 nodes to estimate ambient temperature, relative humidity, and binary occupancy. A physics-informed, detectability-aware multi-task pipeline combines hybrid CSI features with a shared gradient-boosting representation and task-specific heads, achieving—under a strictly temporal, leakage-controlled protocol—MAE of 0.72 ± 0.05 ◦C, 4.6 ± 0.4% RH, and 95.1% occupancy F1, with 47 ms on-device latency and <100 kB foot-print. Multi-task learning significantly improves the continuous tasks over single-task baselines. We reported a bill of materials, power budget, occupancy confusion matrix, ablation, and per-task feature importance, and we stated cross-building generaliza¬tion plainly as the primary open limitation. The results indicate that multi-parameter CSI sensing is feasible on low-cost microcontrollers, offering a path toward scalable, maintenance-light smart-building instrumentation using existing Wi-Fi infrastructure.


## Acknowledgements

The authors thank XZent Solutions Pvt Ltd for infrastructure support and the facility staff for assistance with the climate-chamber experiments and testbed deployment.

## Declarations

**Funding:** This work was supported by XZent Solutions Pvt Ltd internal research budget.

Conflict of interest: The authors declare no competing financial interests.

**Ethics approval:** Not applicable (no human or animal subjects, or identifiable personal data; occupancy annotation recorded presence counts only, with no personally identifying information).

**Data availability:** Processed datasets and trained models are available from the corresponding author upon reasonable request.

**Author contributions:** S.C. conceived the study, developed the theory, and wrote the manuscript. K.P. designed and implemented the multi-task learning pipeline and experiments. D.F. implemented firmware and embedded deployment. S.G. supervised the project and secured hardware and facilities. All authors reviewed and approved the final manuscript.

## References


1. https://www.researchgate.net/publication/223885675_A_Review_on_buildings_energy_consumption_information
2. Dounis, A. I., & Caraiscos, C. (2009). Advanced control systems engineering for energy and comfort management in a building environment—A review. *Renewable and Sustainable Energy Reviews, 13*(6-7), 1246-1261.
3. Sohraby, K., Minoli, D., & Znati, T. (2007). *Wireless sensor networks: technology, protocols, and applications*. John wiley & sons.
4. Meola, C., & Carlomagno, G. M. (2004). Recent advances in the use of infrared thermography. *Measurement science and technology, 15*(9), R27-R58.
5. H. Zou, Y. Jiang, L. Xie, and C. Spanos, "Non-intrusive occupancy sensing in commercial buildings," *Energy and Buildings*, vol. 154, pp. 749–760, 2017. DOI: 10.1016/j.enbuild.2017.08.009
6. Depatla, S., Muralidharan, A., & Mostofi, Y. (2015). Occupancy estimation using only WiFi power measurements. *IEEE Journal on Selected Areas in Communications, 33*(7), 1381-1393.
7. Wang, Y., Liu, J., Chen, Y., Gruteser, M., Yang, J., & Liu, H. (2014, September). E-eyes: Device-free location-oriented activity identification using fine-grained WiFi signatures. In *Proceedings of the 20th annual international conference on Mobile computing and networking* (pp. 617-628).
8. Halperin, D., Hu, W., Sheth, A., & Wetherall, D. (2011). Tool release: Gathering 802.11 n traces with channel state information. *ACM SIGCOMM computer communication review, 41*(1), 53-53.
9. Kotaru, M., Joshi, K., Bharadia, D., & Katti, S. (2015, August). Spotfi: Decimeter level localization using wifi. *In Proceedings of the 2015 ACM conference on special interest group on data communication* (pp. 269-282).
10. Abdelnasser, H., Youssef, M., & Harras, K. A. (2015, April). WiGest: A ubiquitous WiFi-based gesture recognition system. In 2015 IEEE conference on computer communications (INFOCOM) (pp. 1472-1480). IEEE.
11. Liu, J., Wang, Y., Chen, Y., Yang, J., Chen, X., & Cheng, J. (2015, June). Tracking vital signs during sleep leveraging off-the-shelf wifi. In *Proceedings of the 16th ACM international symposium on mobile ad hoc networking and computing* (pp. 267-276).
12. Qian, K., Wu, C., Yang, Z., Liu, Y., & Jamieson, K. (2017, July). Widar: Decimeter-level passive tracking via velocity

monitoring with commodity Wi-Fi. In *Proceedings of the 18th ACM international symposium on mobile ad hoc networking and computing* (pp. 1-10).
13. Z. Wang, K. Guo, and F. Xiao, "RSSI-based temperature sensing for wireless sensor networks," in Proc. IEEE ICCCN, 2016, pp. 1–6. DOI: 10.1109/ICCCN.2016.7568583
14. Chaudhari, S., Pise, K., Fukate, D., & Gawande, S. (2026). Wavelet-Domain Respiratory-Cardiac Decoupling for Wi-Fi CSI Vital Sign Monitoring: A Multi-Resolution Analysis Framework.
15. Chaudhari, S., Pise, K., Fukate, D., & Gawande, S. (2026). Deep Wavelet Scattering Networks for Robust Wi-Fi CSI Vital Sign Separation Under Multipath Interference and Non-Stationary Dynamics: Theory, Algorithms, and Real-Time Implementation.
16. M. Atif, S. Muhammad, et al., "Wi-ESP: A tool for CSI-based device-free WiFi sensing," *Journal of Computational Design and Engineering, vol.* 7, no. 5, pp. 644–657, 2020. DOI: 10.1093/jcde/qwaa043
17. Liu, J., Chen, Y., Wang, Y., Chen, X., Cheng, J., & Yang, J. (2018). Monitoring vital signs and postures during sleep using WiFi signals. *IEEE Internet of Things Journal, 5*(3), 2071-2084.
18. Zhang, Y., Zheng, Y., Qian, K., Zhang, G., Liu, Y., Wu, C., & Yang, Z. (2021). Widar3. 0: Zero-effort cross-domain gesture recognition with Wi-Fi. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 44*(11), 8671-8688.DOI: 10.1109/TPAMI.2021.3105387
19. Raza, U., Kulkarni, P., & Sooriyabandara, M. (2017). Low power wide area networks: An overview. *ieee communications surveys & tutorials, 19*(2), 855-873.
20. T. Chen et al., "A review of humidity sensing principles and applications,"Sensors, vol. 21, no. 22, p. 7653, 2021. DOI: 10.3390/s21227653
21. Zhang, Y., & Yang, Q. (2018). An overview of multi-task learning. National Science Review, 5(1), 30-43.
22. Caruana, R. (1997). Multitask learning. *Machine learning, 28*(1), 41-75.
23. T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc.ACM SIGKDD*, 2016, pp. 785–794. DOI: 10.1145/2939672.2939785
24. H. Cho, "Treelite: A model compiler for efficient deployment of tree ensemble models," *Journal of Machine Learning Research*, vol. 19, no. 1, pp. 1–5, 2018.
25. T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2002.
26. P. E. Ciddor, "Refractive index of air: New equations for the visible and near infrared," *Applied Optics*, vol. 35, no. 9, pp. 1566–1573, 1996. DOI: 10.1364/AO.35.001566
27. Measures, R. M. (1984). Laser remote sensing: Fundamentals and applications(Book). *New York, Wiley-Interscience*, 1984, 521 p.
28. Alduchov, O. A., & Eskridge, R. E. (1996). Improved Magnus form approximation of saturation vapor pressure. *Journal of Applied Meteorology* (1988-2005), 601-609.
29. Wu, K., Xiao, J., Yi, Y., Chen, D., Luo, X., & Ni, L. M. (2012). CSI-based indoor localization. *IEEE Transactions on Parallel and Distributed Systems, 24*(7), 1300-1309.
30. Hernandez, S. M., & Bulut, E. (2020, August). Lightweight and standalone IoT based WiFi sensing for active repositioning and mobility. In 2020 *IEEE 21st International Symposium on" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM)* (pp. 277-286). IEEE.
31. Bergmeir, C., & Benítez, J. M. (2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences, 191*, 192-213.
32. Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems, 30*.