POLITECNICO DI TORINO

Master's Degree in Cybersecurity

Master's Degree Thesis

Crowdsourced Jammer Localization Using Physics-Informed Models and Federated Learning

Supervisors

Prof. Andrea NARDIN

Dr. Iman EBRAHIMI MEHR

Candidate

Behrad SHAYEGAN

MARCH 2026

Summary

Global Navigation Satellite Systems (GNSS) are vital to contemporary infrastructure but remain heavily susceptible to Radio Frequency Interference (RFI), particularly deliberate jamming. Crowdsourcing interference data from consumer smartphones is a scalable detection solution; however, it poses significant challenges regarding hardware heterogeneity, signal propagation in urban areas, and user privacy.

This thesis develops a privacy-enhancing framework for crowdsourced jammer localization built on a two-stage pipeline trained via Federated Learning (FL). The first stage addresses the absence of calibrated power measurements on smartphones by introducing a physics-informed signal-fusion architecture. This model processes baseline-corrected delta observables derived from Automatic Gain Control (AGC) and Carrier-to-Noise Density (C/N_0), with a learned per-device AGC sign orientation to ensure consistent monotonic behavior across hardware. Two dedicated estimation channels—a nonlinear physics-based C/N_0 inversion and a linearized AGC mapping—are blended through a regime-adaptive fusion gate. By exploiting the linearity of AGC under weak interference and the continued informativeness of C/N_0 under strong interference (where AGC saturates), the model reconstructs a physically consistent jammer Received Signal Strength (RSS) estimate without requiring specialized hardware. A post-hoc per-(device, band) affine calibration corrects residual systematic biases.

Localization in the second stage employs an Augmented Physics-Based Model (APBM) with a novel softmax-gated fusion that formulates the output as a competitive weighted combination of a learnable log-distance path-loss branch and a parallel neural branch, rather than the additive residual formulation of prior APBMs. This design allows either branch to dominate when the other's assumptions are violated. Unlike prior APBM approaches that depend on 3D building-height databases, the neural branch uses only lightweight contextual features—two-dimensional building density from open-source map data and local C/N_0 variance derived from receiver observations—making the approach scalable to any area without costly 3D city models.

To enhance user privacy and address the non-Independent and Identically

Distributed (non-IID) nature of crowdsourced data, the framework is trained using Federated Learning, where raw observables and location traces remain on each device. A systematic comparison of three algorithms—FedAvg, FedProx, and SCAFFOLD—is conducted. A hybrid optimization approach is presented for SCAFFOLD, applying separate learning rates and excluding physics parameters ($\boldsymbol{\theta}$, P_0, γ) from control variate correction while stabilizing the neural and fusion components, addressing the challenge of jointly learning interpretable physical quantities in a federated setting.

Experimental validation across four propagation environments (Lab Wired, Suburban, Urban, and Open Sky) with five data-partitioning strategies shows that the proposed framework achieves sub-meter localization accuracy (0.75 m in the Urban scenario with centralized training). SCAFFOLD achieves the best localization in 11 of 20 tested configurations (55%), with particular advantages in challenging non-IID scenarios arising from device-based and signal-strength-based partitioning.

Ablation studies confirm that RSSI spatial information is essential for localization: destroying RSSI structure degrades Pure PL by 10–158× and APBM by 1.2–38×, while Pure NN—serving as a negative control—converges to geometric attractors regardless of RSSI content. A two-regime analysis revealed that symmetric receiver placement masks RSSI importance by placing the centroid near the jammer; graduated asymmetric subsampling (shifting the centroid ∼30 m) restored the expected RSSI dependence, demonstrating that the pipeline's value is greatest in operationally realistic asymmetric deployments. The model architecture ablation showed that pure neural network solutions fail catastrophically in wireless environments (43–60 m errors in Urban), while APBM achieved sub-meter accuracy (0.46 m with oracle RSSI, 0.77 m with predicted RSSI). The Lab Wired exception—where Pure NN (1.41 m) outperformed APBM (11.41 m)—validates that physics-informed learning helps only when model assumptions match reality, providing principled guidance for deployment scenarios. These findings demonstrate the feasibility of decentralized, physics-informed machine learning for privacy-enhancing GNSS security applications. In practical terms, the proposed framework could support regulatory authorities and infrastructure operators in detecting and locating illegal jammers using only crowdsourced smartphone data, without requiring users to share raw location traces or deploying dedicated monitoring infrastructure.

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my wife. Your unwavering love, patience, and support have been my anchor throughout this journey. You believed in me even when I doubted myself, celebrated every small victory, and stood by me through the long nights of coding and writing. This achievement is as much yours as it is mine.

I am deeply grateful to my mother for her sacrifices and love, to my father for his guidance and belief in me, and to my sister for her constant support and encouragement.

I would like to express my sincere gratitude to my supervisors, Prof. Andrea Nardin and Dr. Iman Ebrahimi Mehr, for their invaluable guidance, expertise, and support throughout this research. Their insightful feedback and encouragement were instrumental in shaping this work.

Finally, I am profoundly grateful to Italy and Politecnico di Torino for giving me this incredible opportunity. This beautiful country welcomed me with open arms and provided me with world-class education, inspiring mentors, and an environment where I could grow both academically and personally. The experience of living and studying here has enriched my life in countless ways, and I will forever carry a piece of Italy in my heart.

To everyone who has been part of this journey—thank you.

Table of Contents

List of Tables

List of Figures

Acronyms

AGC
Automatic Gain Control

APBM
Augmented Physics-Based Model

C/N_0
Carrier-to-Noise Density

dBm
Decibel-milliwatt

ENU
East-North-Up

FedAvg
Federated Averaging

FedProx
Federated Proximal

FL
Federated Learning

GNSS
Global Navigation Satellite System

IID
Independent and identically distributed

L-BFGS

Limited-memory BFGS

LDPL

Log-Distance Path Loss

MAE

Mean Absolute Error

NN

Neural Network

Non-IID

Non-identically distributed

PL

Path Loss

R^2

Coefficient of Determination

RMSE

Root Mean Squared Error

RSSI

Received Signal Strength Indicator

SCAFFOLD

Stochastic Controlled Averaging

SGD

Stochastic Gradient Descent

Chapter 1

Introduction

1.1 Background and Motivation

Global Navigation Satellite Systems (GNSS) have become a vital component of modern critical infrastructure, providing Position, Navigation, and Timing (PNT) information for applications ranging from autonomous driving and aviation to telecommunications and financial transactions [1, 2, 3]. However, reliance on these systems has introduced significant vulnerabilities. GNSS signals are inherently weak by the time they reach the Earth's surface, making them highly susceptible to Radio Frequency Interference (RFI) [4].

Among the various threats, intentional jamming—the transmission of high-power signals to mask legitimate satellite signals—is a growing concern. Low-cost "personal privacy devices" (PPDs) are essentially compact GNSS jammers deliberately used to obscure satellite signals [5]. These pocket-sized devices are widely accessible and trivially deployable: they can be powered by a vehicle's cigarette lighter [3], carried in a backpack, or concealed in fixed installations, enabling both mobile and stationary interference across diverse operational contexts. Their low cost and ease of use mean they can deny GNSS service over large areas, degrading the performance of nearby receivers and disrupting both civilian and critical infrastructures [6, 7]. Consequently, the ability to detect and accurately localize these interference sources is essential to maintain GNSS resilience and to provide authorities with the technical tools to prevent the continuation of a jamming action [8, 3].

1.2 The Shift to Crowdsourcing

Traditional methods for jammer localization rely on dedicated infrastructure, such as networks of high-end monitoring stations. While effective, these systems are costly to deploy and lack the spatial resolution required to monitor vast urban

environments [6, 9]. To overcome these limitations, recent literature proposes the paradigm of *participatory sensing* or *crowdsourcing*. This approach leverages the ubiquity of GNSS-enabled consumer devices, such as smartphones and vehicles, acting as an ensemble of voluntary contributors to provide measurements [6, 10].

The primary observables utilized in this domain are the Carrier-to-Noise density ratio (C/N_0) and Automatic Gain Control (AGC) values. Research indicates that these metrics, available even in mass-market Android devices, correlate with jammer power—C/N_0 decreases monotonically with interference strength, while AGC responds strongly but with a sign and scale that vary across receiver implementations, requiring per-device normalization to serve as a reliable proxy for Received Signal Strength (RSS) [6, 11, 12]. By aggregating these measurements from a dense network of agents, it is possible to reconstruct the interference field and estimate the jammer's location without specialized hardware [1, 9].

1.3 Problem Statement

While crowdsourcing offers a promising avenue for GNSS interference monitoring, practical deployment in urban scenarios is currently hindered by three interconnected challenges regarding signal processing, environmental modeling, and privacy-aware distributed optimization.

1.3.1 Inaccuracy of Independent Signal Observables

Reliable jammer power estimation is complicated by the distinct failure modes of receiver observables in dynamic environments. Automatic Gain Control (AGC) provides linearity in weak interference but saturates quickly, while Carrier-to-Noise density ratio (C/N_0) offers continued sensitivity in strong interference but fluctuates under nominal fading. Current state-of-the-art approaches, such as Han et al. (2024) [9], address this by performing localization separately on AGC and C/N_0 maps and averaging the coordinates, an approach named *decision-level fusion*. This is problematic because it allows erroneous spatial estimates from a saturated or noisy sensor to corrupt the final position. Moreover, AGC behavior—and, to a lesser extent, the noise figures that affect C/N_0—is highly device-dependent: different chipsets exhibit different gain ranges, quantization steps, and saturation thresholds. This hardware heterogeneity makes independent per-device calibration and subsequent fusion over a shared, physically consistent power metric a fundamental requirement for any crowdsourced system. There is currently a lack of methods that fuse these metrics at the *signal level* to reconstruct such a unified metric prior to localization.

1.3.2 APBM without 3D map dependency

Urban environments introduce severe Non-Line-of-Sight (NLOS) and multipath effects that render standard Path Loss (PL) models ineffective. Purely physics-based estimators fail to account for shadowing, while purely data-driven models lack generalization. Recent Augmented Physics-Based Models (APBMs), such as those by Jaramillo-Civill et al. [8], successfully mitigate urban localization errors. That federated learning approach uses only 2D position coordinates as neural network input and can be trained on any available RSSI data; in the absence of real measurements, the authors validated their method using ray-traced propagation data generated from 3D building models. A more recent Bayesian extension by the same group [1] feeds building-height rasters directly into a CNN to capture multipath and shadowing effects at inference time. However, reliance on building-height rasters as direct model input limits scalability, as high-fidelity 3D city models are often unavailable, outdated, or computationally expensive to process on mobile devices. A critical need therefore exists for an APBM that avoids reliance on 3D geospatial models entirely and instead infers environmental context from lightweight inputs, such as two-dimensional building density footprints and receiver-derived signal statistics.

1.3.3 Privacy Risks and Algorithmic Instability

Centralized processing of crowdsourced jammer data requires users to upload time-stamped location traces, posing significant privacy risks and creating a single point of failure. Federated Learning (FL) mitigates this by keeping data local, but it introduces optimization challenges in real-world networks. Crowdsourced networks are inherently non-IID due to device hardware variance (*system heterogeneity*) and uneven spatial sampling (*statistical heterogeneity*). Standard algorithms like FedAvg suffer from "client drift" [13] in these settings, where local models diverge toward device-specific biases rather than the true jammer location. Furthermore, physics-based localization layers suffer from numerical singularities in the near-field of the jammer. Existing literature lacks a cohesive framework that simultaneously addresses privacy through FL, corrects client drift via control variates (e.g., SCAFFOLD), and ensures numerical stability during decentralized training.

1.4 Methodological Approach

To address these challenges, this thesis proposes a robust, privacy-aware framework for crowdsourced jammer localization. The methodology is structured into a two-stage pipeline designed to ensure physical consistency, deployability, and algorithmic stability.

1.4.1 Stage 1: Signal-Level Metric Fusion

The first stage addresses the fundamental challenge of estimating jammer signal strength from uncalibrated smartphone sensors. Unlike decision-level fusion approaches that train separate models on AGC and C/N_0 before combining their outputs [9], this work implements an *early-fusion* (signal-level) strategy that processes baseline-corrected delta observables (ΔAGC, $\Delta C/N_0$) jointly.

The core insight is that AGC and C/N_0 exhibit complementary behavior across interference regimes: AGC responds linearly to weak-to-moderate jamming but saturates under strong interference, while C/N_0 remains more informative and less saturation-prone under strong jamming but becomes noisy when interference is weak. Rather than selecting one metric or averaging both, the proposed architecture learns per-band gating coefficients that, given the observed signal conditions of each measurement, produce a per-sample adaptive weight automatically prioritizing the more reliable observable.

The fusion model comprises three components: (1) a physics-informed channel that converts C/N_0 degradation to estimated jammer power using the known relationship between interference and carrier-to-noise ratio; (2) a linear channel that maps AGC changes to jammer power with learned device-specific calibration; and (3) a gating mechanism that dynamically blends these two estimates based on the observed signal conditions. This ensures that the downstream localization stage receives a single, consistent RSSI estimate that adheres to physical propagation principles regardless of interference intensity and receiver heterogeneity.

1.4.2 Stage 2: Augmented Physics-Based Model for Localization

The second stage estimates the jammer's geographic position from the crowdsourced RSSI measurements. We adopt the Augmented Physics-Based Model (APBM) framework [1, 14], replacing the original additive residual structure with a softmax-gated fusion that allows each branch to contribute proportionally to the final prediction.

The path-loss component encodes the fundamental physics of radio propagation: signal strength decreases logarithmically with distance from the transmitter. This provides strong inductive bias and ensures that the estimated jammer position corresponds to a physically plausible signal source. However, real-world propagation in urban environments deviates substantially from ideal path-loss due to building shadowing, reflections, and multipath effects. The neural network branch provides a complementary prediction informed by environmental context; a learned softmax gate then determines the relative contribution of each branch, enabling the model to suppress the physics component when its assumptions are violated and to rely

on it when propagation is well-behaved.

This thesis introduces three modifications to the baseline APBM architecture:

1. **Lightweight Contextual Features:** The original APBM by Jaramillo-Civill et al. [8] uses only 2D receiver coordinates as neural network input, with 3D building geometry employed solely in the ray-tracing simulation that generates training data. A subsequent Bayesian approach by the same group [1] takes this further by feeding building-height maps directly into a CNN as model input. To improve deployability without requiring 3D geospatial databases for either simulation or inference, we enrich the input feature space with two lightweight contextual features: building density (derived from openly available 2D footprint maps) and local signal variance (computed directly from receiver measurements). These features allow the model to distinguish between open-sky and dense urban environments without explicit 3D information.

2. **Softmax-Gated Fusion:** Unlike the additive residual structure of existing APBMs—where the neural network corrects the physics model—we formulate the output as a weighted combination of two parallel branches: $\hat{J} = w_{\mathrm{PL}}\, f_{\mathrm{PL}}(\mathbf{x}) + w_{\mathrm{NN}}\, f_{\mathrm{NN}}(\mathbf{x})$, where the weights are computed via a softmax over two learnable logits. This ensures smooth gradient flow, prevents mode collapse toward either component, and allows the model to suppress the physics branch entirely when its assumptions are violated (e.g., in wired/non-propagation scenarios).

3. **Robust Loss Function:** To handle the heavy-tailed error distribution caused by urban multipath, we employ a peak-weighted Huber loss that remains robust to outliers while emphasizing high-power measurements near the jammer location.

1.4.3 Federated Learning with Variance Reduction

To enable privacy-enhancing distributed training, the framework supports three federated learning algorithms with increasing sophistication for handling heterogeneous (non-IID) client data:

- **FedAvg** [15]: The standard federated averaging algorithm, which aggregates locally trained models by weighted averaging. We enhance position estimation robustness using geometric median aggregation.

- **FedProx** [16]: Extends FedAvg with a proximal regularization term that penalizes local models from deviating too far from the global model, reducing client drift in heterogeneous settings.

- **SCAFFOLD** [13]: Employs control variates to correct for gradient estimation bias caused by non-IID data distributions. We implement a hybrid optimizer strategy that applies different learning rate schedules to physics parameters and neural network weights, addressing the distinct optimization dynamics of interpretable physical quantities versus data-driven components.

1.5 Thesis Scope and Contributions

1.5.1 Scope

The work presented in this thesis focuses on:

- **Single-jammer scenarios:** Localization of a single active jammer source per environment. Multi-jammer detection is beyond the current scope.
- **Smartphone-based crowdsourcing:** Measurements obtainable from commercial GNSS receivers in smartphones (AGC, C/N_0) via Android GNSS APIs.[1]
- **Static jammers:** The jammer position is assumed fixed during measurement collection.
- **Four propagation environments:** Open Sky, Suburban, Urban, and Lab (wired), representing the spectrum of deployment conditions.

1.5.2 Contributions

The main contributions of this thesis are:

1. **Two-Stage Pipeline Architecture:** A modular framework decoupling RSSI estimation from localization, enabling independent optimization and interpretable intermediate outputs.
2. **ExactHybrid Model for Signal-Level Fusion:** A physics-informed hybrid model with adaptive gating that outperforms decision-level fusion approaches for RSSI estimation [9].

[1] `GnssStatus` provides C/N_0 (API 24+), while AGC is accessed through `GnssMeasurement.getAutomaticGainControlLevelDb()` (API 26–32) or the event-level `GnssMeasurementsEvent.getGnssAutomaticGainControls()` (API 33+).

3. **APBM Without 3D Map Dependency:** An augmented physics-based model that incorporates lightweight contextual features, including two-dimensional building density and receiver-derived signal statistics, thereby eliminating reliance on explicit 3D building-height maps.

4. **SCAFFOLD Implementation for Localization:** A complete federated learning implementation with hybrid optimizer design, demonstrating competitive or superior performance in heterogeneous settings, with best-case gains up to 10.6× over FedAvg (Lab Wired signal-strength partitioning) and overall best accuracy in 55% of tested configurations.

5. **Distance-Based Non-IID Partitioning:** A novel data partitioning strategy that assigns crowdsourced receivers to federated clients based on their distance from the jammer, so that some clients hold only close-range (high-RSSI) observations while others hold only distant (low-RSSI) observations. This simulates a realistic deployment scenario in which nearby and faraway contributors experience fundamentally different signal conditions, creating the kind of heterogeneous (non-IID) data distribution that challenges standard federated optimization.

6. **Comprehensive Experimental Validation:** Sub-meter accuracy (0.75 m in Urban), ablation studies, and environment-specific analysis across four propagation conditions.

1.6 Research Questions

This thesis seeks to answer the following research questions:

RQ1: *Can smartphone observables (AGC, C/N_0) be reliably fused at the signal level to estimate jammer RSSI, and does this approach outperform decision-level fusion?*

RQ2: *Can physics-informed learning methods achieve accurate localization without 3D maps?*

RQ3: *How do privacy-enhancing Federated Learning algorithms compare under realistic non-IID?*

RQ4: *Under what environmental conditions does the neural network component provide significant benefit over a pure physics-based model?*

1.7 Thesis Organization

The remainder of this thesis is organized as follows:

Chapter 2: Background and Related Work provides theoretical foundations including GNSS technology, radio propagation models, neural networks, and federated learning principles. We survey related work in jammer localization and crowdsourced sensing.

Chapter 3: Methodology presents the two-stage framework in detail: the ExactHybrid model for Stage 1 RSSI estimation and the APBM with federated learning for Stage 2 localization.

Chapter 4: Experimental Setup describes datasets, evaluation metrics, data preprocessing, partitioning strategies, and implementation details.

Chapter 5: Results and Discussion presents experimental results including Stage 1 performance, localization accuracy, FL algorithm comparison, and ablation studies.

Chapter 6: Conclusion summarizes findings, discusses limitations, and outlines future research directions.

Chapter 2

Background and Related Work

This chapter provides the theoretical foundations necessary for understanding the proposed jammer localization framework. We begin with an overview of GNSS vulnerabilities, introducing the Received Signal Strength Indicator (RSSI) as a key metric for interference detection and its correlation with smartphone-accessible observables (Section 2.1). Section 2.2 then provides detailed coverage of the Automatic Gain Control (AGC) and Carrier-to-Noise Density (C/N_0) metrics. Section 2.3 covers methods for estimating Received Signal Strength from these observables. Section 2.4 reviews radio propagation models fundamental to localization. Section 2.5 surveys jammer localization techniques from measurement-based to data-driven approaches. Finally, Sections 2.6 and 2.7 introduce the machine learning and federated learning concepts leveraged in this thesis.

2.1 GNSS Vulnerabilities and the Crowdsourcing Paradigm

Global Navigation Satellite Systems (GNSS) are critical for modern infrastructure, yet their signals are inherently weak. By the time they reach the Earth's surface, signal power is typically below the thermal noise floor (approximately $-130\,\text{dBm}$), making them highly susceptible to Radio Frequency Interference (RFI) [11]. Intentional interference can be categorized into:

- **Jamming:** Intentional emission of noise to mask the legitimate signal
- **Spoofing:** Broadcasting counterfeit signals to deceive the receiver [17, 18]

2.1.1 RSSI as an Interference Indicator

Received Signal Strength Indicator (RSSI) serves as a fundamental metric for assessing the total power present within a given radio frequency band. In the context of Global Navigation Satellite Systems (GNSS), authentic satellite signals are exceptionally weak and typically reside below the receiver's thermal noise floor [11, 19]. Consequently, any significant increase in the total received power—reflected as an elevated RSSI—serves as a primary indicator of potential radio frequency interference (RFI), such as intentional jamming or spoofing.

Because direct, absolute power measurements (in dBm or Watts) are rarely available on mass-market crowdsourcing devices like smartphones, RSSI is practically inferred through its strong correlation with the Automatic Gain Control (AGC) and the Carrier-to-Noise density ratio (C/N_0) [12]. The AGC is a control loop designed to maintain a constant signal amplitude at the input of the analog-to-digital converter (ADC) to minimize quantization losses [2]. When an interference source injects additional power into the GNSS band, the total RSSI increases, forcing the AGC to decrease its applied gain to prevent saturation [12]. Therefore, a drop in the reported AGC metric acts as a direct proxy for an elevated RSSI, effectively flagging the presence of anomalous signal power [11, 19].

Furthermore, correlating this AGC-derived RSSI with C/N_0 measurements provides a robust mechanism to not only detect interference but also to distinguish between jamming and spoofing attacks [17]. The C/N_0 metric quantifies the strength of the tracked satellite signal relative to the background noise. By observing the joint behavior of these two observables, the nature of the interference can be classified:

- **Jamming (Suppression):** A jammer transmits high-power noise to drown out legitimate GNSS signals. This raises the noise floor and the overall RSSI, causing the AGC gain to drop. Simultaneously, because the authentic signal is obscured by the added noise, the measured C/N_0 of the visible satellites drops proportionally [12, 19].

- **Spoofing (Deception):** A spoofer transmits simulated GNSS signals at a power level designed to overpower the authentic signals and capture the receiver's tracking loops. This added power also increases the RSSI, causing the AGC to drop. However, because the receiver locks onto an artificially strong, well-formed counterfeit signal, the C/N_0 will typically remain constant or even increase [17, 19].

By utilizing the combination of AGC and C/N_0, receivers can reliably assess the RSSI profile to identify anomalous power while significantly reducing false alarms caused by natural signal attenuation—for example, entering a building lowers C/N_0

but leaves the AGC relatively constant [19, 12]. While the joint AGC/C/N_0 signatures reviewed above can in principle discriminate between jamming and spoofing, this thesis focuses exclusively on jamming detection and localization; spoofing classification is beyond the current scope. Within a crowdsourced framework, harmonizing these universally available smartphone metrics into a unified RSSI representation allows the system to map the spatial distribution of the interference field, facilitating the accurate localization of the malicious source.

2.1.2 Challenges of the Crowdsourcing Paradigm

While traditional interference monitoring relies on sparse networks of dedicated, high-cost reference stations, the proliferation of GNSS-enabled consumer devices has enabled *participatory sensing* (crowdsourcing). This paradigm leverages smartphones as a dense sensor network [10]. However, smartphone-based crowdsourcing introduces significant challenges: hardware heterogeneity, where different chipsets exhibit varying AGC and C/N_0 characteristics; uncalibrated measurements, since absolute power levels are unknown without device-specific calibration; and data privacy concerns, as users may be reluctant to share location-tagged measurements. These challenges necessitate advanced signal processing and machine learning techniques [10].

2.2 Informative Observables on Jammer Position

To localize a jammer without specialized equipment, we must rely on standard metrics provided by Commercial Off-The-Shelf (COTS) GNSS receivers. The two most informative observables regarding jammer proximity are AGC and C/N_0.

2.2.1 Automatic Gain Control (AGC)

The AGC is a control loop found in the radio front-end designed to maintain a constant signal amplitude at the input of the Analog-to-Digital Converter (ADC) to minimize quantization losses.

The AGC adjusts the Variable Gain Amplifier (VGA) gain, denoted as G, to keep the total power at the ADC input P_{ADC} within the optimal dynamic range. The total input power P_{in} consists of the thermal noise floor and any external interference; as established by Levigne [11] and Olsson et al. [10], this relationship obeys energy conservation. To ensure unit consistency, we model total input power (Watts) by integrating the thermal noise spectral density N_0 (W/Hz) over the effective front-end bandwidth B (Hz):

$$P_{\text{in}} \approx P_{\text{signal}} + P_{\text{noise}} + J \tag{2.1}$$

where J denotes the interferer power. While this term encompasses any RFI source (including spoofers), this thesis focuses specifically on jammer localization; hence, we refer to J as *jammer power* throughout. Since GNSS signal power P_{signal} is negligible (typically below the thermal noise floor), the input power is driven primarily by the integrated thermal noise power and the jammer power J:

$$P_{\text{in}} \approx N_0 B + J. \tag{2.2}$$

Consequently, the power at the ADC is:

$$P_{\text{ADC}} = G \cdot P_{\text{in}} \approx G \cdot (N_0 B + J). \tag{2.3}$$

As total input power increases (e.g., due to interference), the AGC reduces G to prevent saturation.

To maintain a constant quantization level ($P_{\text{ADC}} \approx$ constant), the gain applied by the AGC must decrease as the jammer power J increases. Expressed in decibels, this yields the inverse relationship used for detection:

$$\text{AGC}_{\text{dB}} \propto -10 \log_{10}(N_0 B + J) \tag{2.4}$$

A decrease in the reported AGC metric therefore serves as a proxy for an increase in J.

While the underlying physics implies a direct monotonic relationship, Lee et al. [19] and Levigne [11] highlight that Android smartphones often do not report a calibrated absolute gain. AGC metrics are often reported in arbitrary "counts" or vendor-specific units rather than standard dB, and certain chipsets (e.g., some Broadcom or Huawei models) employ "relative" AGC that resets the baseline after continuous interference to recalibrate the ADC. Consequently, in this thesis, AGC is utilized as a monotonic indicator of gain control rather than a calibrated absolute power meter. We rely on the change in AGC, ΔAGC, relative to a device-specific baseline to estimate interference intensity.

2.2.2 Carrier-to-Noise Density Ratio (C/N_0)

C/N_0 represents the ratio of the received carrier power C to the noise power spectral density N_0, expressed in dB-Hz:

$$\left(\frac{C}{N_0}\right)_{\text{dB-Hz}} = 10 \log_{10}\left(\frac{C}{N_0}\right)_{\text{linear}} \tag{2.5}$$

In the presence of wideband interference, the jammer acts as an additional noise source. To maintain dimensional consistency with N_0 (W/Hz), the interference

contribution is expressed as an effective spectral density J_0. As derived by Betz [20] and Borio et al. [21], the effective C/N_0 becomes:

$$\left(\frac{C}{N_0}\right)_{\text{jammed}} = 10\log_{10}\left(\frac{C}{N_0 + J_0}\right) \text{ [dB-Hz]} \tag{2.6}$$

where J_0 is the effective interference power spectral density at the correlator input. J_0 relates to the total received jammer power J (Watts) through the receiver's effective front-end bandwidth B_{eff} and the spectral separation coefficient κ (capturing spectral overlap between the jammer and the GNSS signal):

$$J_0 \approx \frac{\kappa\, J}{B_{\text{eff}}}. \tag{2.7}$$

Consequently, a drop in C/N_0 serves as a proxy for increased jammer power.

In this thesis, our goal is to estimate the total in-band jammer power J (in dBm). While the exact bandwidth B_{eff} and coefficient κ are often unknown for uncalibrated smartphones, they act as constant scaling factors for a specific device–band combination. Therefore, in the first stage of the proposed pipeline (Stage 1: RSSI estimation from baseline-corrected delta observables), these unknown constants are absorbed into device-specific learnable calibration parameters, allowing the model to map the observed drop $\Delta(C/N_0)$ directly to J_{dBm} without requiring explicit hardware characterization.

C/N_0 offers several benefits as a jammer observable. It is a standard metric available on almost all GNSS receivers, ensuring broad compatibility. Unlike AGC, which saturates quickly near the interference source, C/N_0 degrades even at large distances from the jammer, providing wide spatial coverage. Additionally, C/N_0 is reported per-satellite, enabling consistency checks across multiple signal sources.

C/N_0 is inherently ambiguous. A drop can be caused by jamming, but also by multipath propagation (destructive interference fading), signal attenuation due to blockage by foliage or buildings, and atmospheric effects such as ionospheric scintillation. Han et al. [9] note that while C/N_0 remains informative and less saturation-prone than AGC in strong interference, it can fluctuate significantly under nominal fading, making it noisier than AGC for power estimation in weak-to-moderate jamming scenarios.

2.2.3 Complementary Characteristics

Table 2.1 summarizes the complementary characteristics of AGC and C/N_0.

Table 2.1: Comparison of AGC and C/N_0 as jammer observables.

Characteristic	**AGC**	C/N_0
Weak interference	Linear, accurate	Noisy, fading-affected
Strong interference	Saturates	Less saturation-prone
Standardization	Vendor-specific	Standard (dB-Hz)
Spatial coverage	Primarily near-field	Broader spatial range
Ambiguity	Low	High (multipath, attenuation)

This complementarity motivates the signal-level fusion approach adopted in this thesis (Stage 1: RSSI estimation), which dynamically weights the two observables based on interference regime.

2.3 RSS Estimation Methods

To perform Received Signal Strength (RSS) based localization, the receiver observables (AGC, $\mathrm{C/N_0}$) must first be baseline-corrected and then mapped to a physical power value, often denoted as Jammer RSSI (J) in dBm. This section reviews existing approaches and identifies gaps addressed by this thesis.

2.3.1 Direct Linear Mapping

The simplest approach assumes a direct linear relationship between observable deviation and jammer power [22]:

$$\hat{J} = a \cdot \Delta\mathrm{AGC} + b \tag{2.8}$$

where $\Delta\mathrm{AGC} = \mathrm{AGC}_{\mathrm{baseline}} - \mathrm{AGC}_{\mathrm{observed}}$ is the deviation from the unjammed baseline, and (a, b) are calibration coefficients.

Limitations: This approach requires device-specific calibration to determine the coefficients (a, b), assumes linearity across all power ranges (which is violated at saturation), and ignores the complementary information available in $\mathrm{C/N_0}$.

2.3.2 Physics-Based Inversion

From the effective C/N_0 relationship (Equation 2.6), one can derive the effective interference spectral density J_0 (W/Hz), given the carrier power C and thermal noise density N_0:

$$J_0 = C \cdot 10^{-\left(\frac{C}{N_0}\right)_{\mathrm{jammed}}/10} - N_0 \quad [\mathrm{W/Hz}]. \tag{2.9}$$

However, performing this inversion requires knowledge of both the nominal satellite signal power C, which varies with satellite elevation and antenna gain, and the receiver thermal noise density N_0, which is device-dependent.

To bypass estimating the absolute carrier power C, Strizic et al. [12] proposed using the difference between unjammed (nominal) and jammed C/N_0:

$$\Delta(C/N_0) = \left(\frac{C}{N_0}\right)_{\text{nominal}} - \left(\frac{C}{N_0}\right)_{\text{jammed}}. \tag{2.10}$$

This difference relates the effective interference density J_0 to the thermal noise density N_0:

$$J_0 = N_0 \left(10^{(\Delta C/N_0)/10} - 1\right) \quad [\text{W/Hz}]. \tag{2.11}$$

To obtain the total in-band jammer power J (Watts or dBm) required for RSS-based localization, we integrate this density over the receiver effective front-end bandwidth B_{eff}:

$$J = J_0\, B_{\text{eff}} = N_0 B_{\text{eff}} \left(10^{(\Delta C/N_0)/10} - 1\right). \tag{2.12}$$

Taking the logarithm yields a closed-form transformation for jammer RSSI in dBm:

$$J_{\text{dBm}} = 10\log_{10}(N_0 B_{\text{eff}}) + 10\log_{10}\left(10^{(\Delta C/N_0)/10} - 1\right). \tag{2.13}$$

In this thesis, the exact noise power $N_0 B_{\text{eff}}$ is often unknown for uncalibrated smartphones. However, it acts as a constant offset for a given device and frequency band. Therefore, in our Stage 1 model for RSSI estimation (Section 3.2), this term is absorbed into the learnable parameter $\theta_{d,b}$, allowing the model to map $\Delta C/N_0$ directly to J_{dBm} without requiring explicit hardware calibration.

2.3.3 Machine Learning Approaches

Recent work has applied machine learning to RSS estimation.

Gaussian Process Regression: Han et al. [9] used Gaussian Processes (GP) to create spatial maps of AGC and C/N_0 values, then performed localization on each map separately. While effective for spatial interpolation, this approach operates at the *decision level* (fusing coordinates rather than signals), does not produce a unified RSS estimate in physical units, and scales poorly with dataset size ($\mathcal{O}(N^3)$ for standard GP).

Neural Network Mapping: Deep learning models can learn complex mappings from observables to RSS without explicit physics knowledge. However, purely data-driven approaches require large labeled datasets with ground-truth RSS, may not generalize across devices or environments, and lack interpretability.

While the methods reviewed above have advanced RSS estimation, a gap remains in approaches that fuse AGC and C/N_0 at the signal level with adaptive weighting and

learned device calibration—a gap addressed by the first stage of the methodology presented in Chapter 3.

2.4 Radio Propagation Fundamentals

Understanding radio propagation is essential for physics-informed localization. This section reviews the fundamental models that underpin RSS-based methods.

2.4.1 Free-Space Path Loss

In free space (no obstacles, reflections, or absorption), the received power decreases with the square of distance according to the Friis transmission equation:

$$P_r = P_t \cdot G_t \cdot G_r \cdot \left(\frac{\lambda}{4\pi d}\right)^2 \tag{2.14}$$

where P_t is transmit power, G_t and G_r are antenna gains, λ is wavelength, and d is distance.

In logarithmic form, the Free-Space Path Loss (FSPL) in dB is:

$$\text{FSPL}(d) = 20\log_{10}(d) + 20\log_{10}(f) - 147.55 \quad [\text{dB}] \tag{2.15}$$

where d is in meters and f is in Hz. This corresponds to a path loss exponent $\gamma = 2$.

2.4.2 Log-Distance Path Loss Model

Real-world environments deviate from free-space conditions. The log-distance path loss model generalizes Equation 2.15 with an environment-dependent path loss exponent:

$$P_r(d) = P_0 - 10\gamma\log_{10}\left(\frac{d}{d_0}\right) \quad [\text{dBm}] \tag{2.16}$$

where:

- P_0 is the received power at reference distance d_0 (typically 1 m)
- γ is the path loss exponent (environment-dependent)
- d is the distance from transmitter to receiver

Table 2.2 shows typical values of γ for different environments:

Table 2.2: Typical path loss exponents by environment [23].

Environment	Path Loss Exponent (γ)
Free space	2.0
Open outdoor	2.0 – 2.5
Suburban	2.5 – 3.0
Urban (LOS)	2.7 – 3.5
Urban (NLOS)	3.0 – 5.0
Indoor (same floor)	1.6 – 3.5

2.4.3 Shadow Fading

The log-distance model assumes deterministic path loss, but real signals exhibit random variations due to obstacles (shadowing). This is modeled as a zero-mean Gaussian random variable in dB:

$$P_r(d) = P_0 - 10\gamma \log_{10}\left(\frac{d}{d_0}\right) + X_\sigma \tag{2.17}$$

where $X_\sigma \sim \mathcal{N}(0, \sigma^2)$ represents log-normal shadowing with standard deviation σ typically ranging from 4–12 dB depending on environment.

2.4.4 Multipath and NLOS Propagation

In urban environments, the direct path between the transmitter and receiver is often blocked, resulting in non-line-of-sight (NLOS) propagation, while the received signal may also consist of multiple reflected components due to multipath. As a result, the channel experiences excess path loss because NLOS paths are longer and therefore more attenuated, fast fading due to constructive and destructive interference among multipath components, and delay spread caused by the temporal dispersion of the received signal.

Simple path loss models (Equation 2.16) cannot capture these effects, motivating *data-driven augmentation* of the physics model as implemented in the APBM architecture.

2.5 Jammer Localization Methods

Localization methodologies can be classified into three layers: the *measurement layer* (what physical quantity is observed), the *model-based layer* (geometric solving

using physics equations), and the *data-driven layer* (learning the mapping from signal to position).

2.5.1 Measurement-Based Techniques

Time of Arrival (TOA) / Time Difference of Arrival (TDOA)

TOA measures the absolute signal flight time, while TDOA measures the time difference between arrivals at spatially separated receivers. For TDOA, the distance difference defines a hyperbola:

$$d_i - d_j = c \cdot \Delta t_{ij} \tag{2.18}$$

The intersection of multiple hyperbolas pinpoints the emitter. While TDOA can achieve high accuracy under ideal conditions, it requires nanosecond-level synchronization and high bandwidth receivers [24], and in urban NLOS conditions, unknown delays prevent hyperbolas from intersecting correctly [25].

Angle of Arrival (AOA)

AOA estimates signal direction using phase differences across an antenna array (e.g., via the MUSIC algorithm). This technique can provide bearing information with a single measurement, but standard smartphones have single omnidirectional antennas, making AOA impossible without specialized hardware [26, 24].

Received Signal Strength (RSS)

RSS exploits power attenuation with distance (Equation 2.16). This approach is the most viable for crowdsourcing since smartphones natively provide AGC and C/N_0 as proxies for RSS [6, 12]. However, RSS-based methods suffer from sensitivity to environmental variations such as shadowing and multipath, which can introduce significant localization errors without appropriate modeling.

2.5.2 Model-Based Localization (Geometric)

Trilateration and Least Squares

Given N receivers at positions $\mathbf{x}_i$ with distance estimates $\hat{d}_i$, the jammer position can be estimated as:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_{i=1}^{N} \left(\|\mathbf{x}_i - \boldsymbol{\theta}\| - \hat{d}_i \right)^2 \tag{2.19}$$

This closed-form approach is computationally efficient, but geometric solvers degrade significantly in urban environments where the path loss model is mismatched to actual propagation [27].

Weighted Centroid Localization (WCL)

WCL computes the jammer position as a weighted average of receiver locations:

$$\hat{\boldsymbol{\theta}} = \frac{\sum_{i=1}^{N} w_i \cdot \mathbf{x}_i}{\sum_{i=1}^{N} w_i} \tag{2.20}$$

This method is simple to implement and robust to outliers when appropriate weighting is used, but it is biased by receiver distribution and fails when the jammer lies outside the convex hull of receivers [26].

2.5.3 Data-Driven and Hybrid Localization

Fingerprinting and Classification

Fingerprinting approaches discretize the area of interest and train classifiers to map signal features to zones. Lyu et al. [5] used SVM for zone-based classification, while Yan and Ruotsalainen [3] employed k-NN with ray-tracing data. These methods can capture complex propagation environments but discretize space (limiting precision) and require exhaustive training databases that may not generalize across environments.

Augmented Physics-Based Models (APBM)

The Augmented Physics-Based Model (APBM) was introduced by Nardin et al. [27] for jammer localization, with comprehensive derivations and analysis provided in [28]. The original APBM formulation combines a physics-based path loss model with a neural network correction term:

$$\hat{J} = f_{\mathrm{PL}}(\mathbf{x}; \boldsymbol{\theta}, P_0, \gamma) + f_{\mathrm{NN}}(\mathbf{x}; \boldsymbol{\phi}) \tag{2.21}$$

In this additive formulation, the physics branch provides the baseline prediction while the neural network learns environment-specific corrections for shadowing and multipath effects. This hybrid architecture achieves robust performance across both open-sky and urban scenarios without requiring manual tuning.

The APBM formulation requires careful handling of numerical singularities. As $d(\mathbf{x}_n, \boldsymbol{\theta}) \to 0$, the path loss function diverges, creating optimization difficulties. Nardin et al. [27] address this by clamping the distance to a minimum far-field distance d_F:

$$\hat{f}_{\mathrm{PL}}(\mathbf{x}_n; \boldsymbol{\theta}) = P_0 - \gamma \cdot 10 \log_{10}\{\max(d(\mathbf{x}_n, \boldsymbol{\theta}), d_F)\} \tag{2.22}$$

This stabilized form $\hat{f}_{\mathrm{PL}}$ then replaces f_{PL} in Equation 2.21, ensuring numerical stability during gradient-based optimization. A practical advantage of the APBM

is its ability to perform P_0-blind estimation, where the jammer transmission power is learned jointly with the position parameters rather than assumed known. This is important in real scenarios where jammer characteristics are unavailable. To prevent the neural network from overpowering the physics model, regularization is applied to the network parameters through techniques such as weight decay and dropout.

Subsequent work by Jaramillo-Civill et al. [8] extended the APBM to federated learning settings, while more recent developments incorporate building-height maps to further improve urban localization accuracy [1]. In Chapter 3, we depart from the additive residual structure of Equation 2.21 by introducing a softmax-gated fusion that formulates the output as a weighted combination of two parallel branches, allowing the relative contribution of each to be learned jointly with all other model parameters.

2.5.4 Summary of Localization Methods

Table 2.3: Comparison of jammer localization methods.

Method	Hardware	NLOS/Multipath	Privacy	Scalability
TOA/TDOA	Specialized	Poor	Central	Low
AOA	Array	Moderate	Central	Low
RSS + LSQ	Smartphone	Poor	Central	High
Fingerprinting	Smartphone	Good	Central	Low
APBM (central)	Smartphone	Good	Central	High
APBM + FL	**Smartphone**	**Good**	**Preserved**	**High**

2.6 Machine Learning Fundamentals

This section introduces the machine learning concepts underlying the models used in this thesis.

2.6.1 Basic Concepts

Before discussing specific architectures, we define several foundational terms used throughout this thesis.

Features and Labels: In supervised learning, the model learns a mapping from input *features* (also called predictors or independent variables) to output *labels* (also called targets or dependent variables). For jammer localization, features include receiver position and signal observables, while the label is the estimated jammer power or position.

Weights and Biases: A linear model computes its output as $y = \mathbf{w}^\top \mathbf{x} + b$, where $\mathbf{w}$ are the *weights* (also called coefficients or slopes) that scale each input feature, and b is the *bias* (also called *intercept*), a constant term that shifts the output independently of the input. The intercept allows the model to fit data that does not pass through the origin.

Parameters and Hyperparameters: *Parameters* are values learned from data during training (e.g., weights and biases), while *hyperparameters* are set before training and control the learning process (e.g., learning rate, number of layers, regularization strength).

Embeddings: An *embedding* is a learned mapping from discrete or categorical inputs (such as device identifiers or frequency bands) to dense, continuous vector representations. Rather than using one-hot encoding, which creates sparse high-dimensional vectors, embeddings learn compact representations where similar inputs are mapped to nearby points in the embedding space. In this thesis, device-specific and band-specific embeddings allow the model to learn calibration parameters that capture hardware variations across different smartphones and GNSS frequency bands.

Training, Validation, and Test Sets: The available data is typically split into three subsets: the *training set* is used to learn model parameters; the *validation set* is used to tune hyperparameters and monitor for overfitting; and the *test set* provides a final, unbiased estimate of model performance on unseen data.

Overfitting and Underfitting: *Overfitting* occurs when a model learns the training data too well, including its noise, resulting in poor generalization to new data. *Underfitting* occurs when a model is too simple to capture the underlying patterns. The goal is to find a model complex enough to fit the data but simple enough to generalize.

Epochs and Batches: An *epoch* is one complete pass through the entire training dataset. A *batch* (or mini-batch) is a subset of training samples processed together before updating model parameters. *Batch size* is a hyperparameter that affects both training speed and convergence behavior.

2.6.2 Neural Networks

A feedforward neural network (multilayer perceptron) computes a function $f : \mathbb{R}^n \to \mathbb{R}^m$ through layers of linear transformations and nonlinear activations:

$$\mathbf{h}^{(l)} = \sigma\left(\mathbf{W}^{(l)}\mathbf{h}^{(l-1)} + \mathbf{b}^{(l)}\right) \tag{2.23}$$

where $\mathbf{W}^{(l)}$ and $\mathbf{b}^{(l)}$ are the learnable weight matrix and bias vector (intercept) for layer l, $\sigma(\cdot)$ is a nonlinear activation function (e.g., ReLU, sigmoid), and $\mathbf{h}^{(0)} = \mathbf{x}$ is the input. The bias term $\mathbf{b}^{(l)}$ allows each neuron to shift its activation threshold independently of its inputs.

Common activation functions include the Rectified Linear Unit (ReLU), defined as $\sigma(x) = \max(0, x)$, which introduces nonlinearity while avoiding the vanishing gradient problem; and the sigmoid function $\sigma(x) = 1/(1 + e^{-x})$, which squashes outputs to the range $(0, 1)$ and is often used for gating mechanisms.

2.6.3 Gradient-Based Optimization

Neural network parameters $\boldsymbol{\theta}$ are optimized by minimizing a loss function $\mathcal{L}$ via gradient descent:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \eta \nabla_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}_t) \tag{2.24}$$

where η is the *learning rate*, a hyperparameter controlling the step size of each update. A learning rate that is too large can cause divergence, while one that is too small results in slow convergence.

Stochastic Gradient Descent (SGD) approximates the true gradient using mini-batches of samples rather than the entire dataset, enabling efficient training on large datasets and introducing beneficial noise that can help escape local minima.

Adam (Adaptive Moment Estimation) combines momentum with adaptive per-parameter learning rates [29]:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t - \eta \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \tag{2.25}$$

where $\hat{m}_t$ and $\hat{v}_t$ are bias-corrected estimates of the first moment (mean) and second moment (uncentered variance) of the gradients, respectively. This adaptive scaling allows faster convergence on parameters with small gradients while dampening updates for parameters with large gradients.

2.6.4 Loss Functions

The *loss function* (also called cost function or objective function) quantifies the discrepancy between model predictions and true labels. The choice of loss function depends on the task and the desired properties of the estimator.

Mean Squared Error (MSE) is the standard loss for regression tasks:

$$\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{2.26}$$

MSE penalizes large errors quadratically, making it sensitive to outliers but providing a smooth optimization landscape.

Mean Absolute Error (MAE) uses the absolute difference:

$$\mathcal{L}_{\text{MAE}} = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i| \tag{2.27}$$

MAE is more robust to outliers than MSE but has a non-smooth gradient at zero.

Huber Loss combines the benefits of both, behaving like MSE for small errors and MAE for large errors:

$$L_\delta(r) = \begin{cases} \frac{1}{2}r^2 & \text{if } |r| \leq \delta \\ \delta(|r| - \frac{1}{2}\delta) & \text{otherwise} \end{cases} \tag{2.28}$$

where δ is a threshold hyperparameter. This makes Huber loss robust to outliers while maintaining smooth gradients near zero.

2.6.5 Regularization and Normalization

Regularization techniques prevent overfitting by constraining model complexity.

L2 Regularization (weight decay) adds a penalty proportional to the squared magnitude of weights: $\mathcal{L}_{\text{reg}} = \mathcal{L} + \lambda\|\mathbf{w}\|_2^2$. This encourages smaller weights and smoother decision boundaries.

Dropout randomly sets a fraction of neuron activations to zero during training, forcing the network to learn redundant representations and reducing co-adaptation between neurons. At inference time, all neurons are active but their outputs are scaled accordingly.

Layer Normalization normalizes activations across features within each sample [30]:

$$\text{LN}(\mathbf{x}) = \gamma \cdot \frac{\mathbf{x} - \mu}{\sigma + \epsilon} + \beta \tag{2.29}$$

where μ and σ are the mean and standard deviation computed across features, γ and β are learnable scale and shift parameters, and ϵ is a small constant for numerical stability. Normalization stabilizes training by reducing internal covariate shift.

2.7 Federated Learning Fundamentals

Federated Learning (FL) enables collaborative model training across distributed devices without centralizing raw data, addressing privacy concerns inherent in crowdsourcing.

2.7.1 Problem Formulation

Consider K clients, each with local dataset $\mathcal{D}_k$. The goal is to minimize the global objective:

$$\min_{\mathbf{w}} F(\mathbf{w}) = \sum_{k=1}^{K} \frac{n_k}{n} F_k(\mathbf{w}) \tag{2.30}$$

where $F_k(\mathbf{w}) = \frac{1}{n_k} \sum_{i \in \mathcal{D}_k} \ell(\mathbf{w}; \mathbf{x}_i, y_i)$ is the local loss, $n_k = |\mathcal{D}_k|$, and $n = \sum_k n_k$.

2.7.2 FedAvg

Federated Averaging [15] is the foundational FL algorithm:

Algorithm 1 FedAvg [15]

Require: Number of clients K, local datasets $\{\mathcal{D}_k\}_{k=1}^K$, local epochs E, communication rounds T

Ensure: Trained global model $\mathbf{w}_T$

Initialize global model $\mathbf{w}_0$ **for** $t = 0, 1, \ldots, T-1$ **do**

  Server broadcasts $\mathbf{w}_t$ to all clients **for** *each client* $k \in \{1, \ldots, K\}$ ***in parallel*** **do**

    $\mathbf{w}_k^{t+1} \leftarrow \textsc{LocalSGD}(\mathbf{w}_t, \mathcal{D}_k, E)$ ▷ E local epochs

  $\mathbf{w}_{t+1} \leftarrow \sum_{k=1}^{K} \frac{n_k}{n} \mathbf{w}_k^{t+1}$ ▷ Weighted aggregation

return $\mathbf{w}_T$

FedAvg assumes IID data across clients. Under non-IID conditions, local updates drift toward client-specific optima, causing slow or divergent convergence [13].

2.7.3 FedProx

FedProx [16] addresses client drift by adding a proximal term to the local objective:

$$\min_{\mathbf{w}} F_k(\mathbf{w}) + \frac{\mu}{2} \|\mathbf{w} - \mathbf{w}_t\|^2 \tag{2.31}$$

where $\mu > 0$ controls the strength of regularization toward the global model $\mathbf{w}_t$.

Effect: Keeps local updates close to the global model, reducing drift at the cost of slower local progress.

2.7.4 SCAFFOLD

SCAFFOLD [13] uses *control variates* to correct gradient bias:

The key idea is to Maintain control variates c_k (client) and c (server) that estimate the gradient drift. Correct local gradients during training:

$$\tilde{g}_k = g_k - c_k + c \tag{2.32}$$

where g_k is the local gradient, c_k is the client control variate, and c is the server control variate.

Control Variate Updates (Option II):

$$c_k^{\text{new}} = c_k - c + \frac{1}{K\eta}(\mathbf{w}_t - \mathbf{w}_k^{t+1}) \tag{2.33}$$

$$c^{\text{new}} = c + \frac{1}{K}\sum_{k=1}^{K}(c_k^{\text{new}} - c_k) \tag{2.34}$$

Achieves variance reduction, converging faster than FedAvg/FedProx under non-IID conditions. Particularly effective when client data distributions differ significantly [13].

The standard SCAFFOLD formulation above assumes a single optimizer applied uniformly to all parameters. In Chapter 3, we describe a hybridized variant tailored to the APBM parameter structure: physics parameters ($\boldsymbol{\theta}$, P_0, γ) are optimized with Adam and excluded from control variate correction, while neural network and fusion-gate weights are trained with SGD and subject to SCAFFOLD's variance reduction. This separation addresses the distinct optimization dynamics of interpretable physical quantities versus high-dimensional learned parameters.

2.7.5 Robust Aggregation

Standard averaging is sensitive to outliers or Byzantine clients. The *geometric median* provides robust aggregation [31]:

$$\mathbf{w}^* = \arg\min_{\mathbf{w}} \sum_{k=1}^{K} \|\mathbf{w} - \mathbf{w}_k\| \tag{2.35}$$

This is computed iteratively using Weiszfeld's algorithm [31]. We apply geometric median specifically for aggregating jammer position estimates $\boldsymbol{\theta}$.

2.7.6 Non-IID Data in Jammer Localization

In crowdsourced jammer localization, non-IID data emerges naturally because clients operate under widely different sensing conditions. Statistical heterogeneity appears when near-field clients observe strong RSSI measurements, while far-field clients receive much weaker signals. System heterogeneity further arises from differences across device models, whose AGC and C/N_0 characteristics can vary substantially. In addition, spatial heterogeneity is introduced by the environment itself, as clients in urban areas experience propagation conditions that differ markedly from those in suburban settings. Taken together, these sources of heterogeneity make optimization more challenging and motivate the use of SCAFFOLD, rather than FedAvg, for the localization task.

2.8 Chapter Summary

This chapter has established the theoretical foundations for the proposed framework.

We began by examining GNSS vulnerabilities and introduced RSSI as a fundamental indicator of interference, showing how it can be inferred from AGC and C/N_0 measurements and how the joint behavior of these observables enables discrimination between jamming and spoofing attacks. AGC and C/N_0 provide complementary information about jammer proximity: AGC responds linearly to weak interference but saturates under strong jamming, while C/N_0 remains informative and less saturation-prone under strong interference but is susceptible to fading and multipath in nominal conditions.

We then reviewed RSS estimation methods, from direct linear mapping to physics-based inversion, noting that existing approaches lack signal-level fusion of multiple observables with adaptive weighting and learned device calibration—a gap addressed by the first stage of the methodology presented in Chapter 3, which combines a physics-informed learned mapping with post-hoc group-wise affine calibration.

The chapter also covered radio propagation fundamentals, establishing that the log-distance path loss model with shadow fading provides the physics basis for localization, while urban NLOS and multipath effects necessitate data-driven augmentation of the physics model. Among localization methods, RSS-based approaches are most suitable for smartphone crowdsourcing, and the Augmented Physics-Based Model (APBM), originally introduced by Nardin et al., offers the best trade-off between physics grounding and urban robustness.

Finally, we introduced federated learning as the enabling framework for privacy-enhancing distributed training—keeping raw location traces on-device while sharing only model parameter updates—with SCAFFOLD addressing the non-IID data challenges inherent to crowdsourced jammer localization through variance reduction.

The next chapter presents the methodology, describing how these foundational techniques are enhanced and combined into the proposed two-stage framework.

Chapter 3

Methodology

This chapter presents the proposed framework for crowdsourced GNSS jammer localization. Section 3.1 provides an overview of the two-stage pipeline architecture. Section 3.2 details the Stage 1 RSSI estimation model (ExactHybrid). Section 3.3 describes the Stage 2 localization model (APBM). Section 3.4 explains the oracle-free training methodology that enables real-world deployment. Section 3.5 presents the federated learning framework. Finally, Section 3.6 provides implementation details.

3.1 System Architecture Overview

3.1.1 The Two-Stage Pipeline

The proposed framework decomposes jammer localization into two sequential stages, as illustrated in Figure 3.1.

The main building blocks of the pipeline are as follows.

Stage 1 — Signal Processing (RSSI Estimation): The first stage addresses the challenge of estimating calibrated jammer power $\hat{J}$ (in dBm) from heterogeneous receiver observables (AGC, C/N_0). After baseline correction and per-device AGC sign orientation (Section 3.2.1), this stage fuses the resulting delta features at the *signal level*, handles device and frequency band heterogeneity through learned embeddings, and produces a unified, physically-grounded power estimate. Embeddings are dense vector representations learned for each discrete category (e.g., device type or frequency band) that capture hardware-specific calibration parameters in a continuous latent space, allowing the model to generalize across devices while accounting for their individual characteristics [32].

Stage 2 — Localization: The second stage uses the estimated RSSI values $\hat{J}$ along with receiver positions to estimate the jammer coordinates $\boldsymbol{\theta} = (\theta_E, \theta_N)$ in a

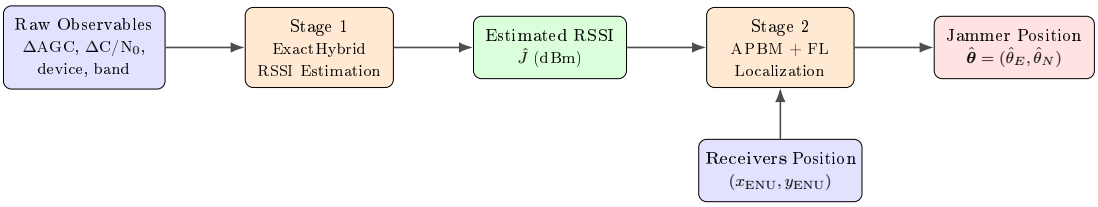


Figure 3.1: Two-stage jammer-localization pipeline. Colors indicate role: purple = inputs, orange = processing stages, green = intermediate result, red = final output.

local East-North-Up (ENU) reference frame. This stage employs a gated fusion architecture that combines physics-based path loss modeling with a data-driven neural network branch through learnable weights, learns the jammer position as an optimizable parameter, and supports federated learning for privacy-enhancing distributed training.

Separating RSSI estimation from localization provides several advantages. First, modularity allows each stage to be independently developed, validated, and updated. Second, interpretability is enhanced as the intermediate RSSI estimates provide diagnostic information about signal quality. Third, flexibility is achieved since Stage 1 can be retrained for new device types without modifying Stage 2. Finally, debugging is simplified because errors can be attributed to either signal processing or localization.

3.1.2 Data Privacy Strategy

The architecture is designed with data locality as a core requirement. In the federated learning paradigm adopted for Stage 2, raw measurements (AGC, C/N_0, GPS coordinates) remain on user devices, and only model parameter updates are transmitted to the central server. The server aggregates these updates without accessing individual user data, thereby reducing exposure of sensitive location information while enabling collaborative learning. We note that this provides practical privacy enhancement—not formal privacy guarantees—since model updates can in principle leak information about local data distributions. Formal differential privacy integration is discussed as future work in Section 6.5.

This approach addresses the reluctance of users to share location-tagged measurements with a central authority, which is a key barrier to crowdsourced interference monitoring deployment.

3.2 Stage 1: Physics-Informed RSSI Estimation

Stage 1 addresses the challenge of estimating jammer Received Signal Strength Indicator (RSSI) from the raw observables available on smartphones. We propose the **ExactHybrid** model, which combines physics-based transformations with learnable device-specific parameters and an adaptive fusion mechanism.

3.2.1 Data Preprocessing and Baseline Computation

Baseline Computation

The baseline values AGC_{base} and CN0_{base} represent nominal receiver behavior under clean (non-jammed) conditions. We compute baselines using a **per-group top-quantile** approach that identifies observations with the strongest signal quality within each device-band combination. We assume each (device, band) has at least some near-clean observations; if persistent interference exists, baseline estimates may be biased.

For each (device, band) pair in the training set, we:

1. Select observations where $\text{CN0} \geq Q_q(\text{CN0})$, where Q_q denotes the q-th quantile (tuned via cross-validation, typically $q \in \{0.7, 0.8, 0.9\}$)
2. Compute AGC_{base} and CN0_{base} as the median AGC and CN0 of this high-quality subset

This approach assumes that observations with the highest C/N_0 values represent the best signal conditions (minimal interference), providing robust baseline estimates even when explicit clean/jammed labels are unavailable.

Leakage Prevention. Baselines are computed exclusively from training data. During walk-forward cross-validation, each fold computes its own baseline map using only temporally preceding observations, ensuring no future information leakage. At inference time, the baseline map learned from the full training set is applied to new observations.

Fallback Hierarchy. Crowdsourced datasets exhibit significant sparsity in the (device, band) space: some device-band combinations may have abundant training observations, while others have few or none. When a (device, band) pair has insufficient observations (< 5), direct baseline estimation becomes statistically unreliable. To address this, we implement a three-tier fallback hierarchy that progressively aggregates across devices:

1. **Device-Band Baseline** (most specific): For combinations with ≥ 5 observations, the baseline is computed directly from the top-quantile subset of that specific (device, band) group. This captures device-specific antenna characteristics and frequency-dependent receiver behavior.

2. **Band-Level Baseline** (first fallback): When a specific device lacks sufficient observations for a given band, we fall back to the band-level baseline aggregated across all devices. This leverages the observation that receivers operating on the same frequency band (e.g., L1 at 1575.42 MHz) exhibit similar C/N_0 characteristics under nominal conditions, even across different hardware.

3. **Global Baseline** (final fallback): For entirely unseen band combinations, we use the global baseline computed from all training data. While this provides the least specificity, it ensures that every observation receives a valid baseline estimate, preventing undefined delta features.

The lookup is implemented as a nested dictionary search:

$$(\text{AGC}_{\text{base}}, \text{CN0}_{\text{base}}) = \begin{cases} \mathcal{B}[(d, b)] & \text{if } (d, b) \in \mathcal{B} \\ \mathcal{B}[(\text{band}, b)] & \text{else if } (\text{band}, b) \in \mathcal{B} \\ \mathcal{B}[\text{global}] & \text{otherwise} \end{cases} \tag{3.1}$$

where $\mathcal{B}$ is the baseline map, d denotes device, and b denotes frequency band.

This hierarchical approach balances specificity with robustness: device-specific calibration is used when data permits, but the system gracefully degrades to coarser aggregations rather than failing or producing unreliable estimates. In practice, approximately 85% of test observations use device-band-specific baselines, 12% use band-level fallbacks, and 3% require global fallbacks, depending on dataset composition.

Delta Feature Computation

The model operates on *delta* features that represent deviations from the baseline:

$$\Delta\text{AGC} = \text{AGC}_{\text{base}} - \text{AGC}_{\text{observed}} \tag{3.2}$$

$$\Delta\text{CN0} = \text{CN0}_{\text{base}} - \text{CN0}_{\text{observed}} \tag{3.3}$$

Using delta features provides several benefits:

- **Device normalization:** Cancels out device-specific antenna gains and biases
- **Physical interpretation:** Positive deltas indicate degradation due to interference
- **Scale consistency:** Both features are in dB units

AGC Sign Orientation

While ΔCN0 is defined so that it increases as observed C/N_0 falls relative to its baseline, the interpretation of the raw AGC delta $\Delta\text{AGC}_{\text{raw}}$ is device- and band-dependent. In practice, the same interference condition can induce different AGC response directions across receiver implementations, so the sign of $\Delta\text{AGC}_{\text{raw}}$ is not reliably consistent across hardware. Without correction, the monotonic interpretation of ΔAGC is therefore inconsistent across devices. To address this, the pipeline learns a per-(device, band) AGC orientation from training data and applies it before model fitting.

To resolve this, we compute a per-(device, band) sign orientation map from the training data by estimating the sign of the covariance between $\Delta\text{AGC}_{\text{raw}}$ and the known RSSI:

$$\sigma_{d,b} = \text{sign}\Big(\text{Cov}(\Delta\text{AGC}_{\text{raw}},\, J)\Big) \tag{3.4}$$

where the covariance is computed over training observations for each (device, band) pair (with ≥ 20 observations; otherwise falling back to band-level or global sign). The oriented delta feature is then:

$$\Delta\text{AGC} = \sigma_{d,b} \cdot \Delta\text{AGC}_{\text{raw}} \tag{3.5}$$

ensuring that ΔAGC is positively correlated with interference intensity across all devices. This oriented ΔAGC and the original ΔCN0 form the two numerical inputs to the ExactHybrid model.

3.2.2 The ExactHybrid Architecture

The ExactHybrid model estimates RSSI through two parallel channels—a physics-based C/N_0 channel and a linearized AGC channel—combined via an adaptive fusion gate. Figure 3.2 illustrates the architecture.

C/N_0 Channel (Physics-Based)

The C/N_0 channel implements a closed-form inversion derived from the relationship between jammer power and C/N_0 degradation. From Equation 2.6 in Chapter 2, the C/N_0 under jamming is:

$$\left(\frac{C}{N_0}\right)_{\text{jammed}} = \frac{C}{N_0 + J} \tag{3.6}$$

The ratio of jammed to clean C/N_0 (in linear scale) gives:

$$\frac{(C/N_0)_{\text{jammed}}}{(C/N_0)_{\text{clean}}} = \frac{N_0}{N_0 + J} = \frac{1}{1 + J/N_0} \tag{3.7}$$

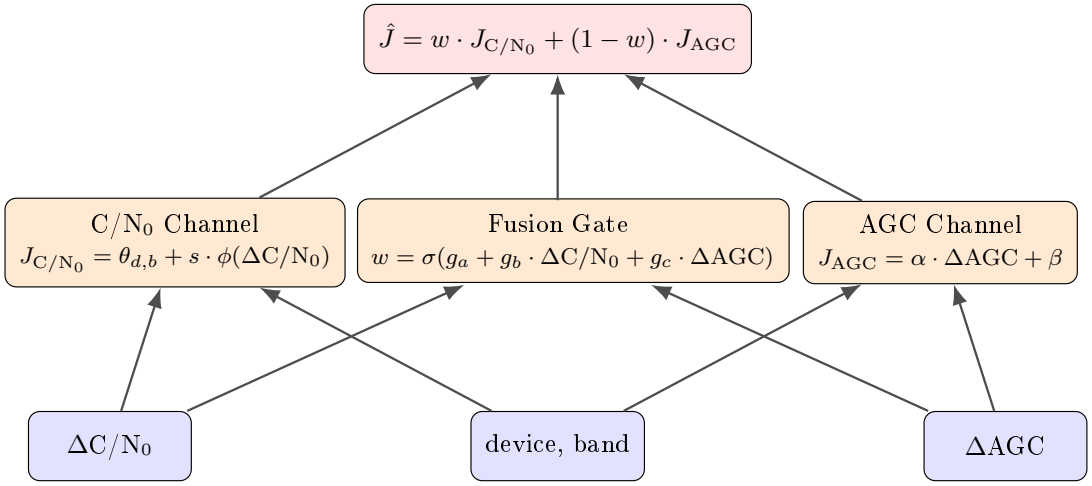


Figure 3.2: ExactHybrid architecture for Stage 1 RSSI estimation. Colors indicate role: purple = inputs (ΔAGC, ΔC/N$_0$, device/band identifiers), orange = processing modules (two estimation channels and the fusion gate), red = fused output.

Converting to decibels and rearranging:

$$\Delta\mathrm{CN0} = \mathrm{CN0}_{\mathrm{base}} - \mathrm{CN0}_{\mathrm{obs}} = 10\log_{10}(1 + J/N_0) \tag{3.8}$$

Inverting this relationship yields the jammer-to-noise ratio:

$$\frac{J}{N_0} = 10^{\Delta\mathrm{CN0}/10} - 1 \tag{3.9}$$

The jammer power in dBm is then:

$$J_{\mathrm{dBm}} = 10\log_{10}(N_0) + 10\log_{10}\left(10^{\Delta\mathrm{CN0}/10} - 1\right) \tag{3.10}$$

Numerical Stability via expm1. Direct computation of $10^x - 1$ suffers from catastrophic cancellation when $x \approx 0$: for small ΔCN0, the term $10^{\Delta\mathrm{CN0}/10}$ is very close to 1, and subtracting 1 loses significant precision. To address this, we reformulate using the identity:

$$10^{x/10} - 1 = e^{(x/10)\cdot\ln(10)} - 1 = \mathrm{expm1}\left(\frac{\ln(10)}{10} \cdot x\right) \tag{3.11}$$

where $\mathrm{expm1}(z) \triangleq e^z - 1$ is a standard numerical function that computes $e^z - 1$ accurately even for small z by using a Taylor series expansion rather than direct subtraction. Defining $c = \ln(10)/10 \approx 0.2303$, Equation 3.9 becomes:

$$\frac{J}{N_0} = \mathrm{expm1}(c \cdot \Delta\mathrm{CN0}) \tag{3.12}$$

The physics-based transformation $\phi(\cdot)$ is therefore:

$$\phi(\Delta\text{CN0}) = \log_{10}\left(\max\left\{\text{expm1}(c \cdot \Delta\text{CN0}), \phi_{\min}\right\}\right) \tag{3.13}$$

where $\phi_{\min} = 10^{-6}$ is a floor that prevents numerical instability when $\Delta\text{CN0} \leq 0$ (i.e., when the observed C/N_0 exceeds the baseline, indicating no interference). The final C/N_0-based RSSI estimate is:

$$J_{\text{CN0}} = \theta_{d,b} + s \cdot \phi(\Delta\text{CN0}) \tag{3.14}$$

where:

- $\theta_{d,b}$ is a learnable offset (embedding) for device d and band b, absorbing the unknown $10\log_{10}(N_0)$ term and device-specific calibration constants
- $s > 0$ is a learnable positive scale (enforced via softplus: $s = \log(1 + e^{s_{\text{raw}}})$), accounting for deviations from ideal physics

In practice, the receiver noise floor N_0, bandwidth effects, and other calibration constants are unknown and vary across devices. These are absorbed into the learnable parameters $\theta_{d,b}$ and s, so J_{CN0} represents a calibrated RSSI-like estimate in dBm.

AGC Channel (Linearized)

The AGC channel implements a linear transformation based on the inverse proportionality between AGC gain and total input power:

$$J_{\text{AGC}} = \alpha_{d,b} \cdot \Delta\text{AGC} + \beta_{d,b} \tag{3.15}$$

where:

- $\alpha_{d,b} > 0$ is a positive slope (enforced via softplus)
- $\beta_{d,b}$ is a learnable intercept
- Both are parameterized per device-band pair via embeddings

The positivity constraint on α ensures monotonicity: increased AGC deviation (indicating stronger interference) must result in higher estimated RSSI.

Device and Band Embeddings

To handle hardware heterogeneity, the model learns separate parameters for each device-band combination. Given D devices and B frequency bands, we define a composite index $k = d \cdot B + b$ that uniquely identifies each (device, band) pair.

The learnable parameters are stored in *embedding tables*—lookup tables that map discrete indices to continuous parameter values. For each parameter type, an embedding table $E \in \mathbb{R}^{(D \times B) \times 1}$ is maintained, where $E[k]$ denotes the parameter value associated with index k. This is equivalent to a trainable lookup operation: given the composite index k, the model retrieves the corresponding parameter from the table.

The C/N_0 channel parameters are retrieved as:

$$\theta_{d,b} = E_\theta[k], \quad s_{d,b} = \text{softplus}(E_s[k]) \tag{3.16}$$

where E_θ stores the offset parameters and E_s stores the raw scale parameters (transformed via softplus to ensure positivity).

The AGC channel parameters are similarly retrieved:

$$\alpha_{d,b} = \text{softplus}(E_\alpha[k]), \quad \beta_{d,b} = E_\beta[k] \tag{3.17}$$

where E_α stores the raw slope parameters and E_β stores the intercept parameters.

The fusion gate coefficients are parameterized per-band (rather than per device-band pair) to reduce overfitting:

$$g_a, g_b, g_c = E_g[b] \tag{3.18}$$

where $E_g \in \mathbb{R}^{B \times 3}$ stores the three gate coefficients for each frequency band.

This embedding-based parameterization offers several advantages. First, it enables efficient parameter sharing: all observations from the same (device, band) pair share the same calibration parameters. Second, it supports extensibility: new devices can be accommodated by appending entries to the embedding tables without retraining existing parameters. Third, it provides interpretability: the learned embeddings can be inspected to understand device-specific calibration differences.

3.2.3 Regime-Adaptive Fusion Gate

The fusion gate dynamically weights the two channels based on the input features:

$$w = \sigma(g_a + g_b \cdot \Delta\text{CN0} + g_c \cdot \Delta\text{AGC}) \tag{3.19}$$

where $\sigma(\cdot)$ is the sigmoid function, and g_a, g_b, g_c are learnable per-band parameters.

The final RSSI estimate is:

$$\hat{J} = w \cdot J_{\text{CN0}} + (1 - w) \cdot J_{\text{AGC}} \tag{3.20}$$

The gate learns to trust AGC under weak interference, where ΔAGC is small and AGC operates in its linear region, and to shift toward C/N_0 under strong interference, where AGC saturates but the C/N_0 physics relationship remains valid. The sigmoid nonlinearity ensures a continuous blend between these two regimes rather than an abrupt switch.

3.2.4 Loss Function and Training

The Stage 1 model is trained using the Huber loss for robustness to outliers:

$$\mathcal{L}_{\text{Stage1}} = \frac{1}{N} \sum_{i=1}^{N} L_\delta(\hat{J}_i - J_i) \tag{3.21}$$

where L_δ is the Huber loss with $\delta = 1.0\,\text{dB}$ (Equation 2.28).

Ground Truth: During training, we use actual jammer RSSI measurements J_i obtained from controlled experiments where the jammer power is known.

An optional monotonicity regularization term was considered to penalize violations of the expected relationship between interference strength and estimated RSSI. The implementation uses a finite-difference approach: the input deltas are perturbed upward by a small ϵ, and the model is penalized when the prediction fails to increase:

$$L_{\text{mono}} = \frac{1}{2}\Big[\max(0,\ \hat{J}(\Delta\text{AGC}) - \hat{J}(\Delta\text{AGC}+\epsilon)) + \max(0,\ \hat{J}(\Delta\text{CN0}) - \hat{J}(\Delta\text{CN0}+\epsilon))\Big] \tag{3.22}$$

The total loss becomes $L = L_{\text{Stage1}} + \lambda_{\text{mono}} L_{\text{mono}}$. In practice, cross-validation consistently selected $\lambda_{\text{mono}} = 0.0$ across all environments, indicating that the softplus positivity constraints on α and s (Equations 3.16–3.17) already enforce sufficient monotonicity without explicit regularization.

3.2.5 Post-Hoc Group Calibration

After training the ExactHybrid model (including the optional L-BFGS polish stage described in Table 3.1), we apply a post-hoc affine calibration to correct for systematic per-(device, band) biases that may remain after model training. For each group $g = (d, b)$ with sufficient observations (≥ 10), an affine mapping $\hat{J}_{\text{cal}} = a_g \hat{J}_{\text{raw}} + b_g$ is fitted by ordinary least squares on the combined training and validation predictions against ground-truth RSSI. Groups with fewer observations inherit the global affine correction. This calibration step is applied at inference time to produce the final RSSI predictions $\hat{J}$ that serve as input to Stage 2.

3.3 Stage 2: Augmented Physics-Based Localization

Stage 2 estimates the jammer position $\boldsymbol{\theta} = (\theta_E, \theta_N)$ from the RSSI predictions of Stage 1 and receiver positions. We adopt an extended Augmented Physics-Based Model (APBM) architecture with gated fusion, which combines a physics-based path loss model with a neural network through learnable weights.

3.3.1 Gated Fusion Architecture

The physics and neural branches are combined via learnable fusion weights:

$$\hat{J} = w_{\text{PL}} \cdot f_{\text{PL}}(\mathbf{x}) + w_{\text{NN}} \cdot f_{\text{NN}}(\mathbf{x}) \tag{3.23}$$

where the weights are computed via softmax over learnable logits:

$$[w_{\text{PL}}, w_{\text{NN}}] = \text{softmax}([w_{\text{PL}}^{\text{raw}}, w_{\text{NN}}^{\text{raw}}]) \tag{3.24}$$

This design implements a globally learned branch weighting: a single pair of logits, shared across all samples within a trained model, determines the relative trust in physics versus data-driven predictions. In open-sky conditions where propagation follows the log-distance model, the physics branch dominates ($w_{\text{PL}} \gg w_{\text{NN}}$); in environments with severe multipath or NLOS conditions, the network branch can compensate by providing an alternative prediction learned directly from data. Unlike a sample-adaptive mixture-of-experts, the weights here are constant for a given environment—the model learns a single physics/NN balance per training run.

Jaramillo-Civill et al. [8] use a constrained scalar $\lambda \in [0, 1]$ to weight the branches. The softmax approach adopted here offers unconstrained optimization (raw logits can take any value while softmax enforces $w_{\text{PL}} + w_{\text{NN}} = 1$ and $w_i > 0$), smoother gradients (avoiding gradient issues at boundary values), and interpretability (weights indicate relative trust in physics vs. data-driven predictions).

3.3.2 The Physics Branch (Path Loss Model)

The physics branch implements the log-distance path loss model with learnable parameters:

$$f_{\text{PL}}(\mathbf{x}; \boldsymbol{\theta}, P_0, \gamma) = P_0 - 10\gamma \cdot \log_{10}(d + \epsilon) \tag{3.25}$$

where $\mathbf{x} = (x_{\text{ENU}}, y_{\text{ENU}})$ is the receiver position in ENU coordinates, $\boldsymbol{\theta} = (\theta_E, \theta_N)$ is the **learnable jammer position**, $d = \|\mathbf{x} - \boldsymbol{\theta}\|_2$ is the Euclidean distance to the jammer, P_0 is the learnable reference power at $d_0 = 1\,\text{m}$, γ is the learnable path loss exponent, and $\epsilon = 10^{-6}$ provides numerical stability when $d \to 0$.

Parameter Constraints: Distance d is clamped to $\geq 1.0\,\text{m}$ to avoid singularities in the near-field, following the far-field distance modification of Nardin et al. [28]. The path loss exponent γ is not hard-clamped; instead, a weak prior toward the environment-specific initialization value (Section 3.4.4) discourages physically implausible values while allowing the optimizer to adapt γ to the observed data.

Initialization: Physics parameters are initialized based on the target environment from a configuration table: γ_{init} is set per environment (e.g., 2.0 for open sky, 2.5 for suburban, 3.5 for urban; see Table 3.1); P_0 is initialized from environment-specific defaults (typically -28 to -35 dBm); and $\boldsymbol{\theta}$ is initialized to the receiver centroid (see Section 3.4). An optional data-driven P_0 estimation helper exists but is not used in the default pipeline.

3.3.3 The Neural Network Branch

The neural network branch provides an independent, data-driven RSSI prediction that captures environment-specific effects (shadowing, multipath, NLOS) which the physics model cannot represent:

$$f_{\text{NN}}(\mathbf{x}; \boldsymbol{\phi}) = \text{MLP}_{\boldsymbol{\phi}}(\mathbf{x}_{\text{norm}}) \tag{3.26}$$

The output is a full RSSI estimate in dBm—the same semantic quantity as the physics branch f_{PL}. This enables the gated fusion (Equation 3.23) to blend two complete predictions rather than applying a correction to a base estimate.

Input Features

The neural network uses a 4-dimensional input vector combining position with lightweight contextual features:

$$\mathbf{x}_{\text{input}} = [x_{\text{ENU}}, y_{\text{ENU}}, \rho_{\text{bldg}}, \sigma^2_{\text{local}}] \tag{3.27}$$

The first two components $(x_{\text{ENU}}, y_{\text{ENU}})$ represent the receiver position in local ENU coordinates. The third component, building density ρ_{bldg}, serves as a proxy for urban density and is computed as the spatial density of two-dimensional building footprints within a fixed-radius neighborhood around the receiver location, derived from open-source map data such as OpenStreetMap. When building footprint data are unavailable, ρ_{bldg} defaults to zero.

The fourth component, local signal variance σ^2_{local}, is a rolling variance of C/N_0 measurements computed per-device over a temporal window. High variance indicates unstable signal conditions due to multipath or interference fluctuations. The window size involves a trade-off: shorter windows (e.g., 5 observations at 1 Hz $\approx$ 5 seconds) provide faster response to changing conditions but yield noisier

variance estimates, while longer windows produce more stable estimates but may smooth over rapid environmental transitions. Cross-validation on the training data selected a window of 5 observations, which proved sufficient to distinguish between stable (open-sky) and volatile (urban canyon) propagation conditions while maintaining temporal responsiveness. The variance estimate need not be statistically precise; rather, it serves as a relative indicator that the neural network learns to interpret during training.

This feature design avoids dependence on high-fidelity 3D building-height maps or ray-tracing databases, which are often unavailable, outdated, or computationally expensive. The neural network learns to associate elevated σ^2_{local} with multipath-prone environments and adjusts its RSSI prediction accordingly.

Network Architecture

The MLP architecture consists of an input layer with 4 normalized features, four hidden layers with [512, 256, 128, 64] neurons using Layer Normalization [30] and Leaky ReLU activation, an output layer with 1 neuron producing an RSSI estimate in dBm, and dropout with $p = 0.2$ for regularization.

Layer Normalization is chosen over Batch Normalization because it operates per-sample by normalizing across features rather than across the batch: $\text{LN}(\mathbf{h}) = \gamma \cdot \frac{\mathbf{h}-\mu}{\sigma} + \beta$, where μ and σ are computed over the feature dimension of a single sample. This property is essential for federated learning, where client batches may contain as few as one sample, and Batch Normalization statistics would be undefined or unreliable.

3.4 Oracle-Free Training Methodology

A critical requirement for practical deployment is that the training procedure must not rely on knowledge of the true jammer position ("oracle" information). This section describes the techniques employed to achieve oracle-free training.

3.4.1 Neutral Coordinate Frame

If the coordinate system origin is placed at the true jammer location, the model can "cheat" by learning to predict $\boldsymbol{\theta} \approx (0, 0)$, especially when L2 regularization pulls parameters toward zero.

As a solution, we define the ENU coordinate frame with origin at the **receiver centroid**:

$$\text{Origin} = \left(\frac{1}{N} \sum_{i=1}^{N} x_i, \frac{1}{N} \sum_{i=1}^{N} y_i \right) \tag{3.28}$$

This ensures that the jammer position $\boldsymbol{\theta}$ is generally not at the origin, that regularization does not bias the model toward any particular location, and that the model must genuinely learn from the RSSI gradient field.

3.4.2 Inverse Localization via RSSI Reconstruction

The model does *not* directly predict jammer coordinates. Instead, localization emerges through **RSSI reconstruction**:

1. The jammer position $\boldsymbol{\theta}$ is a learnable parameter of the model
2. Given receiver positions $\mathbf{x}_i$ and the current $\boldsymbol{\theta}$, the model predicts RSSI: $\hat{J}_i = f(\mathbf{x}_i; \boldsymbol{\theta}, P_0, \gamma, \phi)$
3. The loss compares predicted RSSI to Stage 1 estimates: $\mathcal{L} = \sum_i L(\hat{J}_i, J_i^{\text{Stage1}})$
4. Backpropagation updates $\boldsymbol{\theta}$ to minimize reconstruction error

Physical Intuition: The optimal $\boldsymbol{\theta}$ is the position that, when combined with the path loss model, best explains the observed RSSI field. This is equivalent to solving the inverse problem of source localization.

3.4.3 Peak-Weighted Huber Loss

We introduce a custom loss function that combines robustness with informative sample weighting:

$$\mathcal{L}_{\text{Stage2}} = \frac{1}{N} \sum_{i=1}^{N} \omega_i \cdot L_\delta(\hat{J}_i - J_i) \tag{3.29}$$

where L_δ is the Huber loss with $\delta = 1.0\,\text{dB}$ and ω_i are adaptive sample weights.

The Huber loss is defined as:

$$L_\delta(r) = \begin{cases} \frac{1}{2}r^2 & \text{if } |r| \leq \delta \\ \delta(|r| - \frac{\delta}{2}) & \text{otherwise} \end{cases} \tag{3.30}$$

This formulation provides robustness to outliers such as multipath spikes and measurement errors by transitioning from quadratic to linear penalty for large residuals.

Observations with high RSSI (close to the jammer) carry more localization information than distant observations, whose weak signals are dominated by noise. To exploit this, the loss weights observations by a power-law function of their normalized RSSI magnitude:

$$\tilde{\omega}_i = \left(\frac{J_i - J_{\min}}{J_{\max} - J_{\min} + \epsilon} \right)^{\alpha}, \qquad \omega_i = \frac{\tilde{\omega}_i}{\frac{1}{N}\sum_{j=1}^{N} \tilde{\omega}_j} \tag{3.31}$$

where $\alpha = 2.0$ controls the emphasis on high-RSSI samples. The power-law formulation concentrates weight on near-field observations more aggressively than a linear scheme, while the mean-normalization ensures the average weight is unity, preserving the loss scale. This technique is known as *peak weighting*.

3.4.4 Regularization and Parameter Priors

Beyond the peak-weighted Huber loss, the centralized Stage 2 objective includes two additional regularization terms:

Physics-weight regularization. To prevent the softmax gate from collapsing entirely toward the neural network branch early in training, a penalty proportional to $(1 - w_{\text{PL}})^2$ is added with strength $\lambda_{\text{PL}} = 0.01$. This encourages the model to maintain a meaningful physics contribution unless the data strongly favor suppressing it.

Parameter priors. Weak quadratic priors on γ and P_0 toward their environment-specific initialization values discourage drift to physically implausible ranges. These priors are controlled by per-parameter regularization coefficients that default to zero in the standard configuration but can be activated for ill-conditioned environments.

The total centralized Stage 2 loss is:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{Stage2}} + \lambda_{\text{PL}}(1 - w_{\text{PL}})^2 + \lambda_{\gamma}(\gamma - \gamma_{\text{init}})^2 + \lambda_{P_0}(P_0 - P_{0,\text{init}})^2 \tag{3.32}$$

3.4.5 Centralized Training Flow

Centralized Stage 2 training proceeds in three phases:

1. **Phase 0 — Physics-only warmup** (30 epochs): The neural network and fusion weights are frozen; only $\boldsymbol{\theta}$, P_0, and γ are updated at reduced learning rates. This establishes a physics-based position estimate before the NN activates.

2. **Phase 1 — Full Adam optimization** (up to 200 epochs): All parameters are jointly optimized with per-group learning rates (Table 3.1). Early stopping monitors validation MSE with a patience of 120 epochs.

3. **Phase 2 — L-BFGS fine-tuning**: After restoring the best model from Phase 1, a full-batch L-BFGS optimizer refines all parameters using second-order curvature information. This quasi-Newton refinement typically yields modest improvements in well-conditioned environments; if it fails to improve on the Adam result, the Phase 1 model is retained.

3.5 Federated Learning Framework

To enable privacy-enhancing distributed training—keeping raw location traces and RSSI measurements on-device while sharing only model parameter updates—Stage 2 is implemented within a federated learning framework. This section describes the data partitioning, optimization algorithms, and aggregation strategies.

3.5.1 Data Partitioning Strategies

Real crowdsourced networks exhibit non-IID data distributions due to spatial clustering, device heterogeneity, and varying proximity to the interference source. We implement five partitioning strategies to simulate these realistic conditions:

Random Partitioning (IID Baseline)

The simplest strategy randomly assigns observations to clients, creating approximately IID distributions. Each observation is assigned uniformly at random to one of K clients, yielding subsets of approximately equal size:

$$\mathcal{D} = \bigcup_{k=1}^{K} \mathcal{D}_k, \quad |\mathcal{D}_k| \approx \frac{|\mathcal{D}|}{K}, \quad \mathcal{D}_i \cap \mathcal{D}_j = \emptyset \ \ \forall\, i \neq j \tag{3.33}$$

This serves as a baseline representing ideal conditions where FL algorithms should perform similarly.

Distance-Based Partitioning

Clients are assigned based on their distance to the jammer, creating strong non-IID conditions:

$$\text{Client}_k = \{i : d_i \in [Q_{(k-1)/K}, Q_{k/K})\} \tag{3.34}$$

where Q_p denotes the p-th quantile of distances $\{d_i\}_{i=1}^{N}$. This creates:

- **Near-field clients:** High RSSI values, strong gradient signal
- **Mid-field clients:** Moderate RSSI, transitional regime
- **Far-field clients:** Low RSSI, weak signal, higher noise

Distance-based partitioning represents the most challenging non-IID scenario, as clients observe fundamentally different RSSI distributions. This is where SCAFFOLD's variance reduction provides the greatest advantage. Distance-based partitioning is used only to simulate extreme non-IID conditions in controlled experiments; it requires oracle distances and is not assumed available in deployment.

Geographic Partitioning

Clients are assigned based on angular sectors around the coordinate origin (receiver centroid):

$$\phi_i = \arctan 2(y_i, x_i), \quad \text{Client}_k = \{i : \phi_i \in [\phi_k^{\min}, \phi_k^{\max})\} \tag{3.35}$$

where angles are sorted and divided into K equal-sized sectors. This simulates scenarios where users are geographically clustered (e.g., different neighborhoods or buildings).

Signal Strength Partitioning

Clients are assigned based on their observed RSSI values:

$$\text{Client}_k = \{i : \hat{J}_i \in [J_k^{\min}, J_k^{\max})\} \tag{3.36}$$

where RSSI values are sorted and partitioned into quantiles. This creates non-IID conditions in the *target variable* rather than the input features, testing the algorithms' robustness to label distribution shift.

Device-Based Partitioning

When device labels are available, clients correspond to physical device types:

$$\text{Client}_k = \{i : \text{device}_i = k\} \tag{3.37}$$

This simulates *system heterogeneity* where different smartphone models have varying AGC/C/N_0 characteristics due to different chipsets, antenna designs, and firmware implementations.

Minimum Client Size: To ensure meaningful local training, we enforce a minimum of $m_{\min} = 10$ observations per client. Undersized partitions are merged with their nearest neighbor (by centroid distance) until all clients meet this threshold.

3.5.2 Shared Reference Frame in Federated Training

To ensure that the estimated jammer position parameter $\boldsymbol{\theta}$ is expressed in a consistent coordinate system across clients, all devices use a common ENU reference origin $(\text{lat}_0, \text{lon}_0)$. In our experiments, $(\text{lat}_0, \text{lon}_0)$ is set to the *centroid* (arithmetic mean) of the receiver latitude and longitude values in the dataset for the corresponding environment:

$$\text{lat}_0 = \frac{1}{N}\sum_{i=1}^{N} \text{lat}_i, \quad \text{lon}_0 = \frac{1}{N}\sum_{i=1}^{N} \text{lon}_i \tag{3.38}$$

All receiver positions are then converted to ENU coordinates with respect to this origin prior to client partitioning. This guarantees that federated aggregation operates in a shared coordinate frame.

Loss Function

The federated learning framework uses Peak-Weighted MSE loss:

$$L_{\text{FL}} = \frac{1}{N} \sum_{i=1}^{N} w_i \cdot (\hat{y}_i - y_i)^2 \tag{3.39}$$

where the weights w_i follow the same adaptive peak-weighting scheme described in Section 3.4.3 (Equation 3.31), prioritizing near-field measurements. Unlike the centralized pipeline which uses Huber loss for outlier robustness, the FL setting uses MSE as the base loss. This choice reflects the smaller per-client batch sizes in federated training, where the Huber loss's piecewise gradient behavior can introduce instability in local updates; MSE provides smoother gradients that are more compatible with the SCAFFOLD variance reduction mechanism.

3.5.3 Federated Optimization Algorithms

We implement and compare three FL algorithms of increasing sophistication for handling non-IID data:

FedAvg (Baseline)

Federated Averaging [15] serves as the baseline:

1. Server broadcasts global model $\mathbf{w}_t$ to all clients
2. Each client k performs E local epochs using Adam optimizer
3. Server aggregates: $\mathbf{w}_{t+1} = \sum_k \frac{n_k}{n} \mathbf{w}_k^{t+1}$

For FedAvg and FedProx, we use per-parameter learning rate multipliers:

- Position $\boldsymbol{\theta}$: $2.0 \times \eta_{\text{base}}$
- Physics parameters (P_0, γ): $0.5 \times \eta_{\text{base}}$
- Neural network parameters: $0.1 \times \eta_{\text{base}}$
- Fusion weights $\mathbf{w}$: $0.1 \times \eta_{\text{base}}$

FedProx

FedProx [16] adds proximal regularization to limit client drift:

$$\mathcal{L}_k^{\text{FedProx}} = \mathcal{L}_k(\mathbf{w}) + \frac{\mu}{2} \|\mathbf{w} - \mathbf{w}_t\|^2 \tag{3.40}$$

We use $\mu = 0.01$ as the proximal coefficient, which provides moderate regularization without overly constraining local updates.

SCAFFOLD with Hybrid Optimizer

SCAFFOLD [13] uses control variates to correct gradient bias under non-IID conditions. Our implementation includes a critical enhancement: a **hybrid optimizer** that treats different parameter groups differently based on their optimization characteristics.

Parameter Group Separation: We identify two distinct parameter groups with different optimization needs:

1. **Physics parameters** ($\boldsymbol{\theta}$, P_0, γ): These have different scales (position in meters, power in dBm, exponent dimensionless) and benefit from adaptive learning rates.

2. **Controlled parameters** (neural network weights ϕ, fusion weights $\mathbf{w}$): These benefit from SCAFFOLD's variance reduction and require vanilla SGD for theoretical guarantees.

Hybrid Optimizer Design:

- **Physics parameters**: Use **Adam** optimizer
  - Position $\boldsymbol{\theta}$: $\eta_\theta = \eta_{\text{base}} \times m_\theta$ where $m_\theta \in [0.1, 0.6]$ (environment-dependent)
  - P_0 and γ: $\eta_{\text{physics}} = \eta_{\text{base}} \times 0.5$
  - *Excluded from control variate correction* for stability
- **Neural network + fusion weights**: Use **vanilla SGD**
  - Learning rate: $\eta_{\text{NN}} = \eta_{\text{base}} \times 0.1$
  - No momentum ($\beta = 0$) as required by SCAFFOLD theory
  - Control variate correction applied

Control Variate Correction: During local training, gradients for controlled parameters are corrected:

$$\tilde{g}_k = g_k - c_k + c \tag{3.41}$$

where c_k is the client control variate and c is the server control variate. Physics parameters ($\boldsymbol{\theta}$, P_0, γ) are excluded from this correction.

Control Variate Update (Option II): After local training, client control variates are updated using the parameter difference method:

$$c_k^{\text{new}} = c_k - c + \frac{1}{K \cdot \eta}(\mathbf{w}_t^{\text{ctrl}} - \mathbf{w}_k^{\text{ctrl}}) \tag{3.42}$$

where $\mathbf{w}^{\text{ctrl}}$ denotes only the controlled parameters (NN + fusion weights), and K is the number of local update steps. This reduces communication overhead compared to accumulating gradients explicitly.

Server Control Variate Update:

$$c^{\text{new}} = c + \frac{1}{|\mathcal{S}|} \sum_{k \in \mathcal{S}} (c_k^{\text{new}} - c_k^{\text{old}}) \tag{3.43}$$

where $\mathcal{S}$ is the set of participating clients in the current round.

3.5.4 Robust Aggregation via Geometric Median

Standard averaging of client position estimates can be corrupted by outliers (e.g., a client with defective measurements). We apply **geometric median** aggregation specifically for the jammer position $\boldsymbol{\theta}$:

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \sum_{k=1}^{K} \|\boldsymbol{\theta} - \boldsymbol{\theta}_k\| \tag{3.44}$$

This is computed iteratively using Weiszfeld's algorithm [31]:

$$\boldsymbol{\theta}^{(t+1)} = \frac{\sum_k \boldsymbol{\theta}_k / \|\boldsymbol{\theta}^{(t)} - \boldsymbol{\theta}_k\|}{\sum_k 1 / \|\boldsymbol{\theta}^{(t)} - \boldsymbol{\theta}_k\|} \tag{3.45}$$

The geometric median is a robust estimator that ignores spatial outliers (providing Byzantine robustness), converges to the true median position even with corrupted clients, and has a breakdown point of 50%, meaning it tolerates up to half adversarial clients. Other parameters (neural network weights, P_0, γ) are aggregated using standard weighted averaging.

3.6 Implementation Details

3.6.1 Hyperparameters

Table 3.1 summarizes the key hyperparameters for both training settings. *Centralized* refers to the standard single-server setting where all data is collected and processed at one location, as opposed to the *federated* setting where data remains distributed across clients. FL hyperparameters are auto-tuned based on the (environment, partition strategy) combination to account for different non-IID severity levels.

Table 3.1: Hyperparameter settings. *Centralized* denotes training with all data at a single server; *FL* denotes federated learning with distributed clients. FL hyperparameters are auto-tuned based on the (environment, partition strategy) combination.

Stage	Parameter	Value	Notes
Stage 1	Learning rate	0.001	Adam optimizer
	Batch size	512	
	Epochs	200	
	Early stopping patience	20	
	Top-quantile baseline	0.7–0.9	CV-selected
	Monotonicity weight	0.0–0.1	CV-selected
	L-BFGS polish	80 iters, $\eta = 0.5$	Post-Adam refinement
Stage 2 (Centralized)	Batch size	32	
	Epochs	200	
	Early stopping patience	120	
	η_θ	0.015	Position LR
	η_{P_0}, η_γ	0.005	Physics LR
	η_{NN}	0.001	Neural network LR
	Hidden layers	[512, 256, 128, 64, 1]	MLP arch.
	NN activation	Leaky ReLU	
	Dropout	0.2	
	Physics bias	2.0	Initial $w_{\text{PL}}/w_{\text{NN}}$
	Warmup epochs	30	Physics-only
	Peak weight α	2.0	
	Gradient clip	1.0	
Stage 2 (Federated)	Global rounds	120–180	Strategy-dependent
	Local epochs	3	Per round
	Warmup rounds	5	Physics-only
	Number of clients	5	
	η_{base} (FL)	0.0045–0.005	Strategy-dependent
	FedProx μ	0.01	Proximal coefficient
	SCAFFOLD m_θ	0.1–0.6	Env-dependent
	SCAFFOLD m_{physics}	0.5	P_0, γ multiplier
	SCAFFOLD m_{NN}	0.08–0.1	NN + $\mathbf{w}$ multiplier
	θ aggregation	Geometric median	Robust to outliers

Environment-Specific Tuning

The SCAFFOLD position learning rate multiplier m_θ is tuned per environment to balance convergence speed with stability:

Table 3.2: SCAFFOLD θ learning rate multipliers by environment and partition strategy.

Environment	Distance	Geographic	Signal	Device	Random
Open Sky	0.3	0.2	0.4	0.2	0.2
Suburban	0.2	0.1	0.2	0.2	0.4
Urban	0.2	0.2	0.2	0.4	0.2
Lab (Wired)	0.6	0.2	0.2	0.2	0.2

Warmup Phase

For both centralized and federated training, we employ a warmup phase. In the centralized setting, the first 30 epochs use reduced learning rates ($\eta_\theta = 0.01$, $\eta_{P_0} = \eta_\gamma = 0.004$). In the federated setting, the first 5 rounds train only the physics parameters ($\boldsymbol{\theta}$, P_0, γ) while the neural network remains frozen. This warmup strategy establishes a reasonable initial position estimate before the neural network activates, preventing the NN from compensating for a poorly initialized $\boldsymbol{\theta}$.

Early Stopping (Oracle-Free)

All early stopping decisions are based on validation loss only—never on localization error, which would require oracle knowledge of the true jammer position. The patience is set to 15–30 rounds without improvement depending on the partition strategy, with a minimum improvement threshold of 0.1 dB. Divergence detection terminates training if the validation loss exceeds three times the best observed value for two consecutive rounds.

3.6.2 Software and Hardware

The implementation uses Python 3.9+ with PyTorch 2.0 for model implementation and automatic differentiation, NumPy and Pandas for data processing, and Scikit-learn for evaluation metrics. Training was performed on NVIDIA GPUs with CUDA acceleration when available, with CPU fallback for compatibility. Typical training times were approximately 5 minutes for Stage 1 and 15 minutes for Stage 2 with 100 federated learning rounds.

3.6.3 Evaluation Metrics

For Stage 1 (RSSI Estimation), we report Mean Absolute Error MAE $= \frac{1}{N}\sum_i |\hat{J}_i - J_i|$ in dB, Root Mean Squared Error RMSE $= \sqrt{\frac{1}{N}\sum_i (\hat{J}_i - J_i)^2}$ in dB, and the Coefficient of Determination (R^2). For Stage 2 (Localization), we report the Localization Error $\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{\text{true}}\|_2$ in meters and RSSI MSE as the mean squared error between predicted and actual RSSI values.

3.7 Chapter Summary

This chapter has presented the complete methodology for crowdsourced GNSS jammer localization. The proposed framework employs a two-stage architecture that decouples RSSI estimation from localization, providing modularity and interpretability.

Stage 1 introduces the ExactHybrid model for physics-informed RSSI estimation, which combines a closed-form C/N_0 channel implementing a physics-based transformation with a linear AGC channel subject to positivity constraints. An adaptive fusion gate weights these channels according to the interference regime, while per-device and per-band parameterization via embeddings handles hardware heterogeneity.

Stage 2 implements an extended APBM for localization, featuring a physics branch with learnable position $\boldsymbol{\theta}$, reference power P_0, and path loss exponent γ, alongside a neural network branch using lightweight contextual features that do not require explicit 3D building-height maps. Softmax fusion weights enable smooth, unconstrained optimization between the two branches.

The framework ensures deployability through oracle-free training, which includes a neutral coordinate frame with the receiver centroid as origin, inverse localization via RSSI reconstruction, and peak-weighted Huber loss for robust optimization.

For privacy-enhancing distributed training, the federated learning framework incorporates distance-based partitioning for realistic non-IID simulation, SCAFFOLD with a hybrid optimizer for variance reduction, and geometric median aggregation for robust position estimation.

The next chapter describes the experimental setup and datasets used to evaluate this framework.

Chapter 4

Experimental Setup and Data Analysis

Evaluating a crowdsourced jammer localization framework ideally requires large-scale field campaigns with diverse smartphones operating in a jammed environment. In practice, however, such campaigns face significant obstacles: intentional GNSS jamming is illegal in most jurisdictions, access to adequately large and shielded test facilities (e.g., anechoic chambers) is limited, and coordinating diverse smartphone hardware under controlled interference conditions is logistically demanding. Moreover, acquiring empirical data from the full spectrum of commodity GNSS chipsets—Qualcomm Snapdragon, Broadcom BCM47755, Samsung Exynos, HiSilicon Kirin, among others—would require procurement and instrumentation of each device under repeatable jamming conditions, an effort that scales poorly with the number of device profiles needed for realistic heterogeneity. For these reasons, this thesis grounds all signal-level calibration in controlled laboratory measurements—where a wired jammer connection provides accurate RSSI ground truth—and extends the resulting data through physics-informed augmentation and simulation to obtain the spatial diversity and device heterogeneity needed for localization and federated learning evaluation. The synthetic device profiles are generated from published chipset specifications rather than empirical measurements, which enables evaluation across a diverse receiver fleet without requiring physical access to each smartphone model; the trade-off is that these profiles represent idealized device behaviors that may not capture all real-world device-specific anomalies (see Section 4.6.1 for details).

Concretely, the experimental data comprises three datasets derived from a single laboratory campaign. First, 930 real observations of AGC, C/N_0, and ground-truth RSSI are collected from a u-blox ZED-F9P receiver with a wired jammer connection (Section 4.2). Second, these measurements are augmented to 3,720 observations by

synthesizing additional device and band diversity while preserving the measured RSSI values, providing the training set for Stage 1 (Section 4.4). Third, a fully synthetic combined dataset of 8,731 observations is generated using path-loss models across four propagation environments in the Turin metropolitan area, providing the spatial distributions needed for Stage 2 localization and federated learning evaluation (Section 4.5.1). Section 4.1 below provides an overview of all three datasets and their relationships before the detailed descriptions that follow.

4.1 Dataset Overview

Table 4.1 summarizes the three datasets used throughout this work. They serve distinct purposes in the two-stage pipeline and differ fundamentally in how the RSSI ground truth is obtained.

Table 4.1: Dataset summary. The key distinction is how RSSI is obtained: preserved from laboratory measurements (Lab Wired, Augmented) versus computed from path-loss models (Combined).

Dataset	Obs.	Real/Synth	RSSI Source	Purpose	Section
Lab Wired	930	100%/0%	Measured	GT calibration	4.2
Augmented	3,720	25%/75%	Preserved	Stage 1 training	4.4
Combined	8,731	0%/100%	Path-loss model	Stage 2 & FL eval.	4.5.1

The datasets form a hierarchical relationship with fundamentally different ground-truth characteristics:

The **Lab Wired** dataset contains 930 real observations from controlled laboratory measurements, where RSSI ground truth is obtained directly from calibrated jammer TX gain settings. This dataset provides the foundation for all subsequent data generation.

The **Augmented** dataset expands the 930 real observations to 3,720 through physics-informed augmentation. Critically, this augmentation preserves the original RSSI values while synthesizing diverse AGC and C/N_0 responses for additional device and band profiles. The preserved RSSI ground truth ensures that Stage 1 models learn accurate observable-to-power mappings anchored to real laboratory calibration.

The **Combined** dataset is fully synthetic, generated independently using log-distance path-loss models across four propagation environments. Unlike the Augmented dataset, the Combined dataset does not preserve laboratory RSSI values; instead, RSSI is computed from simulated receiver-to-jammer distances using environment-specific propagation parameters. This enables evaluation of Stage 2

localization across diverse spatial distributions and federated learning scenarios that would be impractical to collect empirically.

4.1.1 Key Distinctions

Understanding the fundamental differences between the datasets is critical for interpreting experimental results. The Stage 1 Augmented dataset preserves real RSSI ground truth from laboratory calibration. When the ExactHybrid model is trained on this data, it learns to map diverse AGC and C/N_0 responses back to accurate jammer power estimates anchored to controlled measurements. The 6 discrete RSSI values (-85 to -60 dBm) reflect the actual TX gain steps used during data collection.

The Combined dataset uses RSSI values computed from path-loss models, which introduces modeling assumptions (path-loss exponent, shadow fading distribution) that may not perfectly match real-world propagation. However, this trade-off enables evaluation across realistic spatial distributions with known ground-truth jammer positions—information that would be unavailable in real crowdsourced deployments. The continuous RSSI range (-153 to -32 dBm) reflects the distance-dependent path loss across the simulated receiver positions.

This separation ensures that Stage 1 benefits from high-fidelity laboratory calibration while Stage 2 can be evaluated on diverse federated learning scenarios with controlled experimental conditions.

The remainder of this chapter is organized as follows. Section 4.2 details the laboratory data collection methodology. Section 4.3 presents exploratory analysis of the real laboratory measurements. Section 4.4 describes the physics-informed augmentation strategy. Section 4.5.1 describes the combined dataset generation process. Section 4.6 provides exploratory data analysis of the combined dataset. Finally, Section 4.7 outlines the evaluation protocol.

4.2 Data Collection

4.2.1 Hardware Setup

The primary data collection employed a controlled laboratory environment as illustrated in Figures 4.1 and 4.2. The setup consists of a roof-mounted GNSS antenna receiving live satellite signals, connected to one input port of an RF combiner (Mini-Circuits ZAPD-2-S+). A HackRF One software-defined radio (SDR) with PortaPack serves as the jammer source, generating controlled L1-band interference (1575.42 MHz) that is injected into the combiner's second input port. The jammer power is precisely controlled through the HackRF's software-configurable TX gain, which can be stepped programmatically from 0 to 47 dB

in 1 dB increments. The combined signal (GNSS + jammer) is fed via wired connection to a u-blox ZED-F9P GNSS receiver front-end, which outputs AGC and C/N_0 observables at approximately 1 Hz. A logging PC running u-center software records the binary UBX protocol data and exports it to CSV format for analysis.

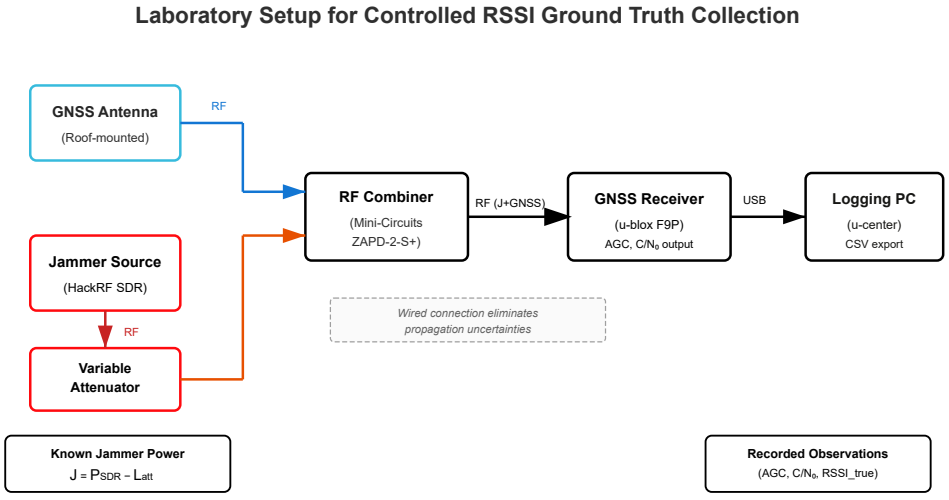


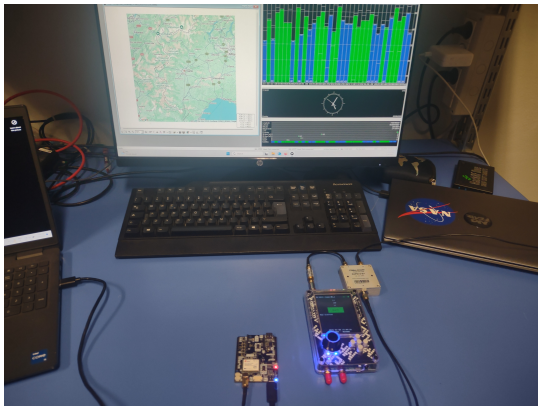
Figure 4.1: Block diagram of the laboratory setup for controlled RSSI ground truth collection. The wired connection between the RF combiner and receiver eliminates propagation uncertainties, enabling precise characterization of the AGC and C/N_0 response to known jammer power levels.

The wired connection between the combiner and receiver eliminates propagation uncertainties such as path loss, multipath, and shadowing. Cable and combiner insertion losses were measured at approximately 1 dB and deemed negligible relative to the 25 dB TX gain sweep range; they are therefore not corrected for in the RSSI ground truth. The true jammer power at the receiver input is determined by the HackRF's calibrated TX gain setting, enabling systematic characterization across a wide range of interference levels. This controlled setup provides accurate RSSI ground truth for training the Stage 1 model, establishing the relationship between GNSS observables (AGC, C/N_0) and true jammer power. This relationship is then preserved through the augmentation process described in Section 4.4.

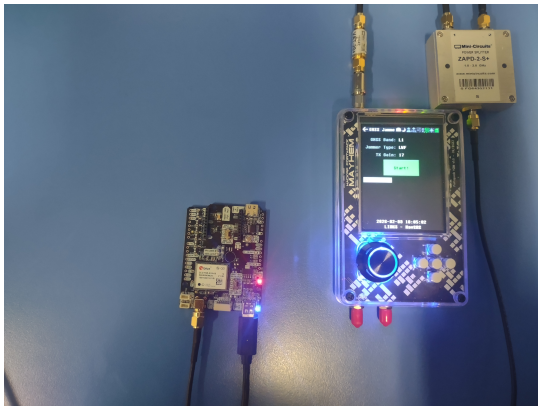
4.2.2 Session Structure

Each data collection session followed a standardized protocol of approximately 15 minutes duration:

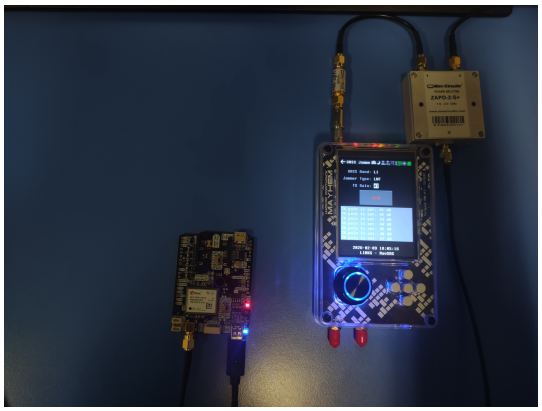
1. **Clean Interval (0–2 min):** No interference; establishes baseline observables



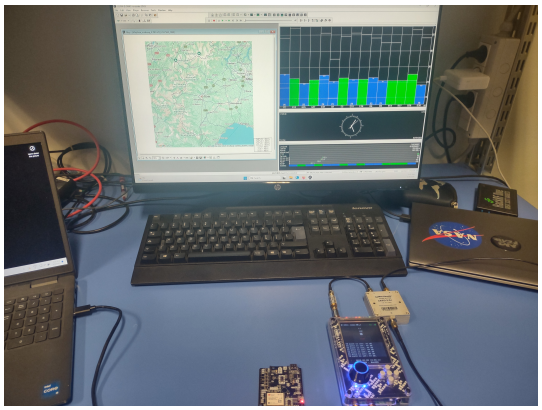
(a) Laboratory setup overview showing the logging PC with u-center software displaying satellite C/N_0 values:strong C/N_0 values (green bars) indicate nominal GNSS reception.



(b) Close-up of hardware components: u-blox ZED-F9P receiver (left), HackRF with PortaPack jammer source (center), and Mini-Circuits ZAPD-2-S+ RF combiner (top).



(c) Close-up of the receiver status with jammer active



(d) Receiver status with jammer active: degraded C/N_0 values demonstrate the interference effect at TX gain = 47 dB.

Figure 4.2: Laboratory data collection setup. The HackRF PortaPack interface (b) allows programmatic control of TX gain to sweep through different jammer power levels while recording the corresponding AGC and C/N_0 responses.

2. **Jammed Interval (2–13 min):** Variable jammer power levels applied
3. **Clean Interval (13–15 min):** Interference removed; confirms return to baseline

The clean intervals at the beginning and end of each session are critical for computing the top-quantile baselines (AGC_{base}, $CN0_{base}$) used by the ExactHybrid model (see Chapter 3). During "clean" intervals, the jammer is set to the minimum calibrated output (−85 dBm at receiver input), which we treat as the no-jamming reference level.

4.2.3 UBX Message Parsing

Raw binary logs were processed using Python with the `pyubx2` library. Table 4.2 summarizes the UBX messages extracted.

Table 4.2: UBX messages used for feature extraction.

Message	Content	Usage
NAV-PVT	iTOW, UTC timestamp	Time synchronization
RXM-RAWX	C/N_0 per satellite, freqID, gnssId	Signal quality metrics
NAV-SIG	Signal information	Supplementary C/N_0
MON-RF	AGC counts (0–8191), RF status	Gain control (converted to dB via Eq. 4.1)

AGC Conversion to Decibel Scale. The `UBX-MON-RF` message reports the Automatic Gain Control (AGC) as an internal dimensionless integer count in the range $[0, 8191]$. While official u-blox documentation does not provide a direct physical mapping to absolute power units (dBm), these counts represent the receiver's relative gain state required to maintain the optimal ADC quantization levels. To obtain a physically interpretable quantity consistent with the dB-scale assumption in Stage 1 (Section 3.2.2), we apply a heuristic linear mapping to convert raw counts into an approximate relative dB representation:

$$\mathrm{AGC}_{\mathrm{dB}} = -20 \cdot \left(1 - \frac{\mathrm{AGC}_{\mathrm{cnt}}}{8191}\right) \tag{4.1}$$

This formula maps the full 13-bit count range to a normalized interval of approximately $[-20, 0]$ dB via an affine transformation. In this model, 0 dB corresponds to maximum gain (indicating the weakest received power), while −20 dB represents a gain reduction of 20 dB (indicating the strongest received power or potential interference). Although the output is expressed in decibels, the mapping itself is linear in the count value; the resulting scale serves as a normalization step

for the Stage 1 training algorithm, ensuring that the features are on a consistent dB-like scale compatible with the physics-informed C/N_0 channel.

4.3 Exploratory Data Analysis: Real Laboratory Data

This section analyzes the 930 real observations collected from the laboratory setup described in Section 4.2. The analysis establishes the ground-truth relationships between AGC, C/N_0, and RSSI that underpin both the Stage 1 model design and the physics-informed data augmentation.

4.3.1 Dataset Overview

The laboratory dataset contains 930 valid observation tuples after dropping rows with missing required fields. Each observation consists of AGC, C/N_0, and the corresponding ground-truth RSSI value recorded at a known jammer power level. The dataset has a jamming ratio of 65.4% (608 jammed observations, 322 clean), reflecting the session structure described in Section 4.2.2.

4.3.2 Signal Quality Statistics

The laboratory data spans six discrete jammer TX gain settings, producing RSSI ground-truth values from -85 to -60 dBm. Since the jammer power is deliberately varied across these levels, aggregate summary statistics such as the overall mean or median of AGC, C/N_0, or RSSI are not physically meaningful—they simply reflect the proportion of time spent at each power level rather than any intrinsic property of the signal.

A more informative characterization is to examine how stable the receiver observables are *within* each power level. Table 4.3 reports the range of each observable along with the mean within-level standard deviation, computed by detrending the time series at each TX gain step. This quantity indicates how much the measured values fluctuate around their expected value at a given interference level, and thus provides a meaningful estimate of measurement noise.

Table 4.3: Observable ranges and within-level stability for real laboratory data. The within-level standard deviation is computed by detrending the time series at each of the six TX gain steps.

Metric	**Min**	**Max**	**Within-level Std**
AGC (u-blox units)	−18.29	−12.29	*TODO: compute*
C/N_0 (dB-Hz)	7.26	44.12	*TODO: compute*
RSSI (dBm)	−85.00	−60.00	0.00 (discrete steps)

The AGC values are negative in u-blox native units, where more negative values indicate higher receiver gain (i.e., stronger interference causes the receiver to reduce its gain). The C/N_0 remains strictly positive (minimum 7.26 dB-Hz) because the real receiver maintains satellite lock even under the strongest interference level tested. The RSSI ground truth takes exactly six discrete values corresponding to the TX gain steps, confirming the controlled nature of the experiment. Figure 4.3 illustrates the temporal evolution of all three observables during a representative data collection session, showing the clean–jammed–clean session structure and the step-wise variation in jammer power.

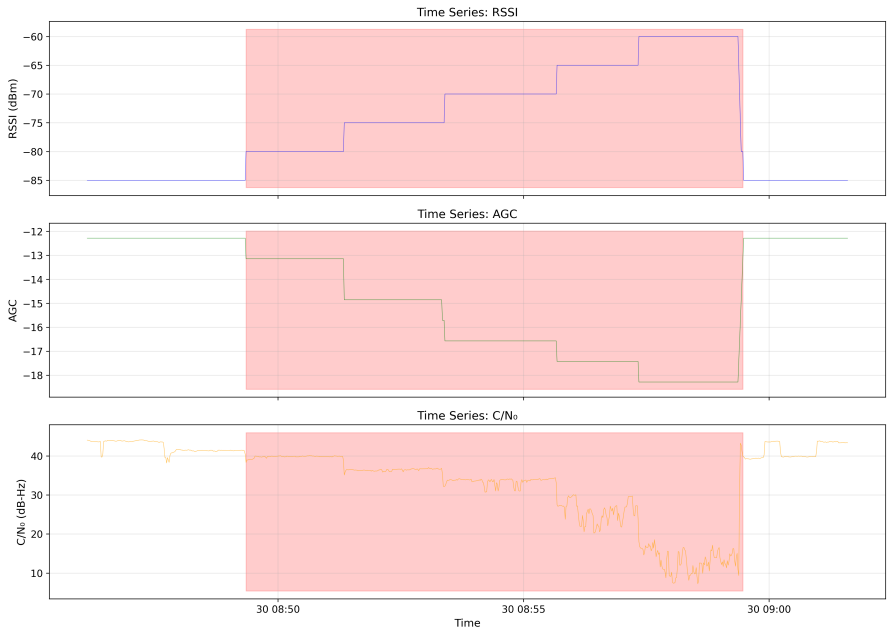


Figure 4.3: Time series of signal observables during real data collection, showing the clean–jammed–clean session structure. Each step in the RSSI trace corresponds to a discrete TX gain level; AGC and C/N_0 respond accordingly.

4.3.3 Jamming Analysis (Real Data)

Since the jammer power varies across multiple TX gain levels during the jammed interval, computing a single mean for the jammed condition conflates fundamentally different interference regimes and is not informative. Instead, the clean-condition statistics—where the jammer is off and the observables reflect only thermal noise and nominal GNSS reception—provide a well-defined reference point.

Under clean conditions, the receiver reports AGC = -12.29 (u-blox units), $C/N_0 = 42.03$ dB-Hz, and RSSI = -85.00 dBm (the minimum-power level, confirming no detectable interference). These values serve as the per-device baseline for computing the delta features used in Stage 1 (Section 3.2).

Figure 4.4 visualizes the distribution of each observable under clean and jammed conditions. Several observations are notable: AGC shows a clear downward shift under jamming with relatively tight within-condition spread, confirming its reliability as a near-field interference indicator. C/N_0 exhibits a larger spread under jamming due to satellite-dependent variation, but the median drops substantially ($\sim$12 dB-Hz). The local signal variance (bottom-left panel) shows minimal separation between conditions, consistent with the controlled laboratory environment where multipath-induced fluctuations are absent.

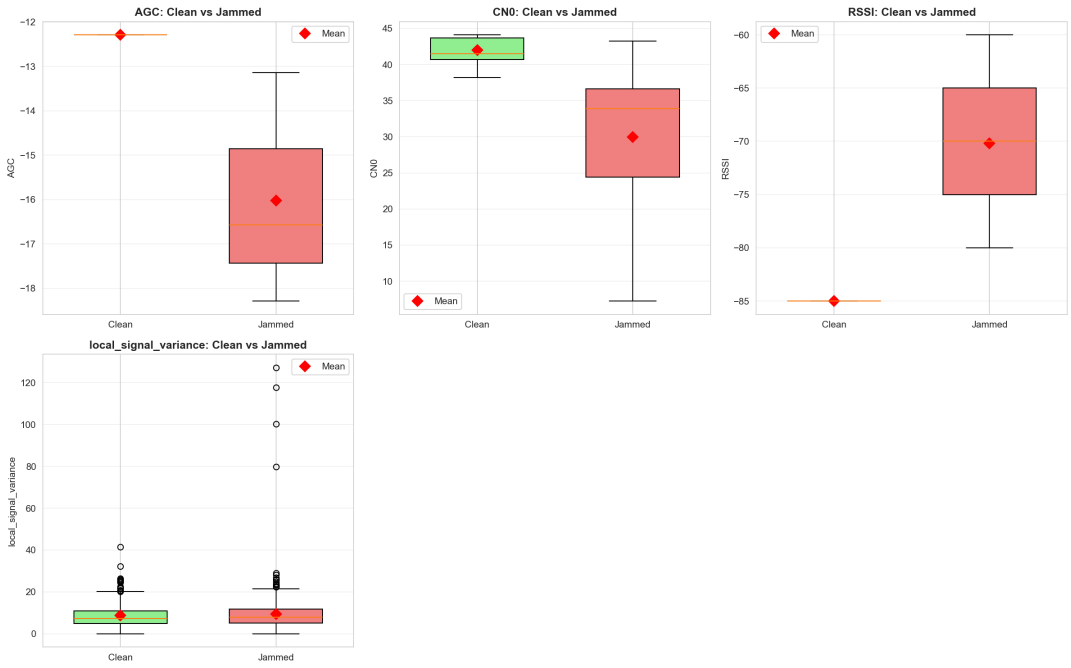


Figure 4.4: Jamming comparison for real laboratory data: box plots of AGC, C/N_0, RSSI, and local signal variance under clean and jammed conditions. The interquartile range under jamming reflects the multiple TX gain levels used, not measurement noise.

4.3.4 Correlation Analysis

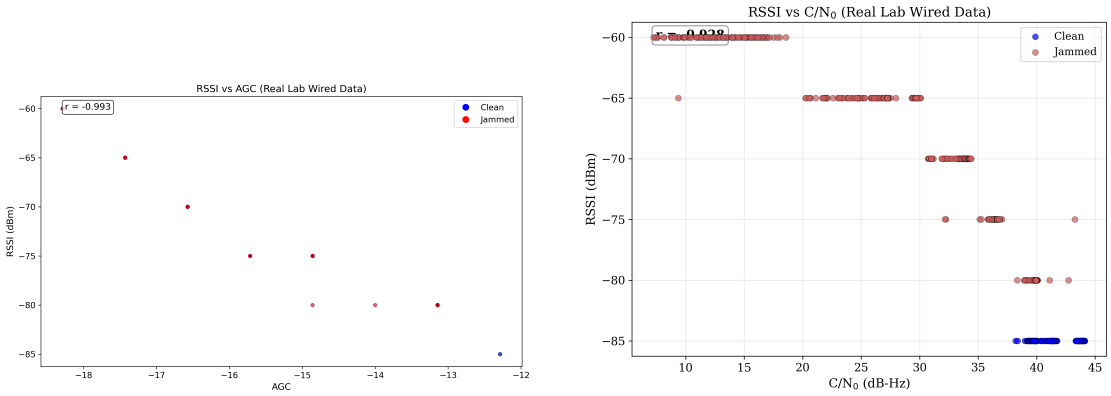
Table 4.4 presents the correlation matrix for the real laboratory data, quantifying the relationships between the three key observables. The correlations reveal the strong physical coupling between jammer power and receiver response in controlled conditions.

Table 4.4: Correlation matrix for real laboratory data.

	RSSI	**AGC**	**C/N_0**
RSSI	1.000	−0.993	−0.928
AGC	−0.993	1.000	0.896
C/N_0	−0.928	0.896	1.000

The most significant finding is the RSSI–AGC correlation of $r = -0.993$, which

validates the linear AGC channel in the ExactHybrid model. This near-perfect negative correlation indicates that as jammer power increases, AGC decreases (the receiver reduces its gain to avoid saturation), confirming the theoretical relationship described in Chapter 2. The RSSI–C/N_0 correlation of $r = -0.928$ is also strong but exhibits more scatter, reflecting the nonlinear physics-based inversion that the C/N_0 channel must learn. Figure 4.5 visualizes both relationships side by side.



(a) RSSI vs. AGC ($r = -0.993$). **(b)** RSSI vs. C/N_0 ($r = -0.928$).

Figure 4.5: Observable–RSSI relationships in real laboratory data. (a) The near-perfect AGC–RSSI linearity ($r = -0.993$) validates the linear AGC channel in the ExactHybrid model. (b) The C/N_0–RSSI relationship is also strong ($r = -0.928$) but exhibits more scatter at moderate interference levels, reflecting the nonlinear physics-based inversion that the C/N_0 channel must learn.

4.3.5 Feature Importance

Table 4.5 presents the point-biserial correlation between each input feature and the binary jamming label, quantifying which observables are most informative for detecting interference in the laboratory environment.

Table 4.5: Feature correlation with jamming label in laboratory data.

Feature	**Correlation**	***p*-value**	**Sig.**
RSSI	+0.778	< 0.001	***
AGC	−0.769	< 0.001	***
C/N_0	−0.586	< 0.001	***
Local signal variance*	+0.035	0.293	ns

*Rolling temporal variance of C/N_0 per device (window = 5 s), serving as a lightweight proxy for multipath activity (see Section 4.5.2).

RSSI and AGC exhibit the strongest correlations with the jamming label ($r = +0.778$ and $r = -0.769$, respectively), both highly significant at the $p < 0.001$ level. This raises a natural question: if AGC alone correlates almost as strongly as RSSI with the jamming label, why not localize directly from AGC rather than estimating RSSI as an intermediate step? The answer lies in device heterogeneity. In this single-device laboratory setting, AGC is an excellent jammer power proxy. However, AGC is reported in vendor-specific, uncalibrated units that differ across chipsets—a Qualcomm Snapdragon, a Broadcom BCM47755, and a u-blox F9P will report entirely different AGC ranges and polarities for the same interference level. RSSI, expressed in standardized dBm, provides a shared physical scale onto which all device-specific AGC (and C/N_0) responses can be calibrated. This is precisely the role of Stage 1: to absorb device-specific behavior and produce a unified power metric that Stage 2 can use for localization across heterogeneous receivers.

C/N_0 shows a moderate negative correlation ($r = -0.586$, $p < 0.001$), reflecting satellite signal quality degradation under interference. The local signal variance—defined as the rolling temporal variance of C/N_0 computed per device over a 5-second window—shows no significant correlation ($r = +0.035$, $p = 0.293$) in the laboratory setting, which is expected: the controlled wired environment has no multipath-induced fluctuations. This feature becomes more informative in the spatially diverse combined dataset (Section 4.6), where it serves as a lightweight proxy for propagation complexity. Building density is excluded from this analysis as it remains constant throughout the single-location laboratory dataset.

4.3.6 Limitations of Real Data

While the real laboratory data provides accurate RSSI ground truth, it has several limitations that motivate the augmentation strategy presented in Section 4.4:

1. **Single Device:** Only the u-blox F9P receiver is represented
2. **Single Band:** Only L1 measurements are available
3. **Single Environment:** Only controlled laboratory conditions are captured
4. **Limited RSSI Range:** Only 6 discrete power levels (25 dB range)
5. **Dataset Size:** 930 observations may be insufficient for learning robust per-device parameters

These limitations are addressed through physics-informed augmentation, which expands the dataset while preserving the critical RSSI ground truth values.

4.4 Physics-Informed Data Augmentation

This section describes the physics-informed data augmentation strategy developed for Stage 1 model training. The augmentation addresses the limitations identified in Section 4.3.6 while preserving the physical relationship between GNSS observables and jammer power.

4.4.1 Motivation for Augmentation

As discussed in Section 4.3.6, the original laboratory dataset—while providing accurate RSSI ground truth—is limited to a single device, single band, and single environment. Direct deployment of models trained solely on this data would fail in real-world scenarios involving diverse smartphones, multi-frequency receivers, and varied propagation conditions.

The core challenge is therefore: how to expand the training distribution to cover diverse devices, bands, and environments while preserving the accuracy of the RSSI ground truth that makes the laboratory data valuable. The following sections describe a physics-informed augmentation strategy that addresses this challenge by generating synthetic AGC and C/N_0 observations from the real RSSI ground truth using physics-based forward models.

4.4.2 Augmentation Methodology

Baseline Estimation

The augmentation process first estimates clean-condition baselines from the real data. Observations from the first and last two minutes of the laboratory session—corresponding to periods before jamming commenced and after it ceased—are identified as clean reference observations. From these observations, the median AGC and C/N_0 values define the clean baselines:

$$\mathrm{AGC}_{\mathrm{base}} = \mathrm{median}(\mathrm{AGC}_{\mathrm{clean}}), \quad \mathrm{CN0}_{\mathrm{base}} = \mathrm{median}(\mathrm{CN0}_{\mathrm{clean}}) \tag{4.2}$$

Additionally, empirical slopes β_{AGC} and β_{CN0} are fitted via linear regression to capture the dataset-specific relationships between RSSI and the observables, which anchor the synthetic variants to the observed real-world behavior.

Device Diversity

Table 4.6 presents six device profiles designed to represent the heterogeneity of commercial GNSS receivers. The profiles are based on published GNSS chipset characterizations [33, 34] and manufacturer datasheets. Each profile name corresponds to a real chipset family: `ublox_wired_like` models the u-blox ZED-F9P used

in laboratory collection; `snapdragon855` represents the Qualcomm Snapdragon 855 (found in, e.g., Samsung Galaxy S10, OnePlus 7 Pro); `bcm47755_new` and `bcm47755_old` represent newer and older revisions of the Broadcom BCM47755 (found in, e.g., Xiaomi Mi 8, Google Pixel 4), where the older revision exhibits the coarse 3 dB AGC quantization common in early dual-frequency chipsets; and `hisilicon980` represents the HiSilicon Kirin 980 (found in, e.g., Huawei Mate 20). The numerical parameters (σ_{AGC}, α, σ_{CN0}, k_{SSC}) were selected to span the range of behaviors reported in the literature; they are not calibrated to specific device units but capture the qualitative diversity across chipset families.

Table 4.6: Simulated device profiles for data augmentation, based on published chipset characteristics [33, 34].

Profile name	**Chipset family**	σ_{AGC}	α	σ_{CN0}	k_{SSC}	**Quantized**
ublox_wired_like	u-blox ZED-F9P	0.3	0.015	0.5	0.7	No
snapdragon855	Qualcomm SDM855	1.0	0.020	1.0	0.8	No
bcm47755_new	Broadcom BCM47755	0.6	0.018	0.8	0.7	No
bcm47755_old	Broadcom BCM47755	0.6	0.018	0.8	0.7	Yes (3 dB)
hisilicon980	HiSilicon Kirin 980	1.2	0.022	1.2	0.9	No

Band Diversity

Two frequency bands are simulated with different front-end bandwidths: $B_{\mathrm{L1}} = 4\,\mathrm{MHz}$ and $B_{\mathrm{L5}} = 10\,\mathrm{MHz}$. The wider L5 bandwidth results in lower jammer spectral density for the same total power, yielding different $\mathrm{C/N_0}$ degradation characteristics. An offset of $+1.5\,\mathrm{dB\text{-}Hz}$ is applied to L5 clean baselines to reflect the typically higher $\mathrm{C/N_0}$ observed on this band.

Environment Diversity

Table 4.7 presents three propagation environments simulated with distinct multipath and fading characteristics. The jitter parameters $\sigma_{\mathrm{CN0,jitter}}$ and $\sigma_{\mathrm{AGC,jitter}}$ control the standard deviation of Gaussian noise added to each observable, p_{dip} is the probability of a multipath-induced signal dip per observation, and Δ_{dip} scales the magnitude of such dips drawn from a Gamma distribution.

Table 4.7: Simulated environment profiles for data augmentation.

Environment	$\sigma_{\text{CN0,jitter}}$	$\sigma_{\text{AGC,jitter}}$	p_{dip}	Δ_{dip}
open_sky	0.5 dB	0.2 dB	0.01	1.5 dB
suburban	1.5 dB	0.4 dB	0.03	2.5 dB
urban	3.0 dB	0.7 dB	0.06	4.0 dB

Forward Models

The synthetic C/N_0 is computed using a physics-based degradation model. The jammer spectral density is first calculated as $J_{\text{dBW/Hz}} = (J_{\text{dBm}} - 30) - 10\log_{10}(B)$, where J_{dBm} is the preserved RSSI ground truth and B is the front-end bandwidth in Hz. The jammer-to-noise ratio in linear scale is then $J/N_0 = 10^{(J_{\text{dBW/Hz}} - N_{0,\text{thermal}})/10}$, where $N_{0,\text{thermal}} = -204\,\text{dBW/Hz}$ [35] is the thermal noise density. The C/N_0 degradation follows:

$$\text{CN0}_{\text{syn}} = \text{CN0}_{\text{clean}} - 10\log_{10}\left(1 + k_{\text{SSC}} \cdot \frac{J}{N_0}\right) + \epsilon_{\text{jitter}} - \epsilon_{\text{dip}} \tag{4.3}$$

where k_{SSC} is the device-specific spectral separation coefficient, $\epsilon_{\text{jitter}} \sim \mathcal{N}(0, \sigma^2_{\text{CN0,jitter}})$ adds environment-dependent noise, and $\epsilon_{\text{dip}} \sim \text{Gamma}(2, \Delta_{\text{dip}}/2)$ occurs with probability p_{dip} to simulate occasional multipath fading events.

The synthetic AGC follows an affine relationship calibrated to real receiver behavior:

$$\text{AGC}_{\text{syn}} = \text{AGC}_{\text{clean}} - \alpha \cdot (J_{\text{dBm}} + 100) + \epsilon_{\text{jitter}} \tag{4.4}$$

where α is the device-specific slope and $\epsilon_{\text{jitter}} \sim \mathcal{N}(0, \sigma^2_{\text{AGC,jitter}})$. For devices with coarse quantization, the AGC is rounded to 3 dB steps: $\text{AGC}_{\text{quantized}} = 3 \cdot \text{round}(\text{AGC}_{\text{syn}}/3) + \epsilon_{\text{dither}}$.

4.4.3 Augmentation Results

Table 4.8 summarizes the augmentation results, showing the expansion from 930 real observations to 3,720 total observations spanning six devices, four environments, and two frequency bands.

Table 4.8: Data augmentation summary statistics.

Metric	**Before (Real)**	**After (Augmented)**
Total observations	930	3,720
Augmentation factor	—	4.0×
Devices	1	6
Environments	1	4
Frequency bands	1	2
Real observations	930 (100%)	930 (25.0%)
Synthetic observations	0 (0%)	2,790 (75.0%)
Unique RSSI values	6	6
RSSI preservation	—	100%

The augmentation achieves a 4× expansion of the training data while preserving 100% of the RSSI ground truth values. Figure 4.6 visualizes the resulting data composition.

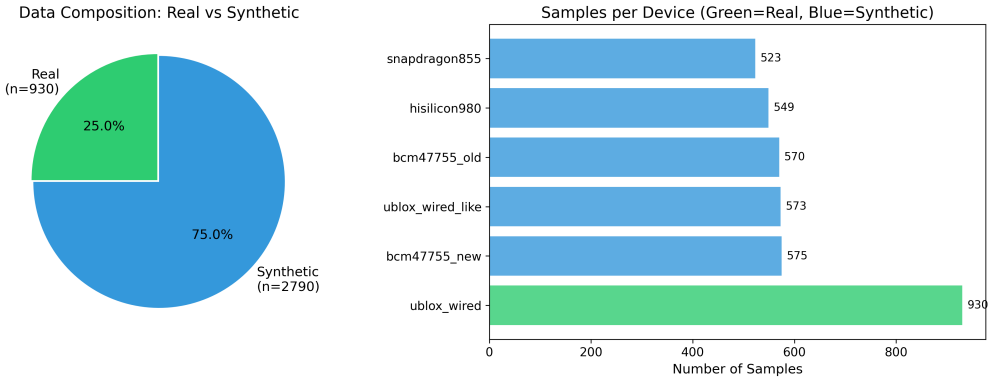


Figure 4.6: Data composition after physics-informed augmentation: 25% real observations, 75% synthetic observations, distributed across six devices.

4.4.4 Observable Transformation Validation

Table 4.9 compares observable statistics between real and synthetic data to validate that the transformations produce physically plausible values.

Table 4.9: Observable comparison: real vs. synthetic data in augmented dataset.

Observable	**Real Mean**	**Synthetic Mean**	**Difference**
AGC	−14.73	−12.67	+2.06
C/N_0 (dB-Hz)	34.15	35.53	+1.38
RSSI (dBm)	−75.33	−75.33	**0.00**

The zero difference in RSSI confirms ground truth preservation. The small differences in AGC and C/N_0 reflect the intended device and environment transformations that create diversity while maintaining physical plausibility.

4.4.5 Per-Device AGC Characteristics

Table 4.10 presents the AGC statistics for each device after augmentation, demonstrating the diversity introduced by the forward models.

Table 4.10: AGC statistics by device after augmentation.

Device	**AGC Mean**	**AGC Std**	**Observations**	**Type**
ublox_wired	−14.73	2.32	930	Real
ublox_wired_like	−12.64	0.65	573	Synthetic
bcm47755_new	−12.78	0.80	575	Synthetic
bcm47755_old	−12.44	1.21	570	Synthetic
hisilicon980	−12.78	1.24	549	Synthetic
snapdragon855	−12.74	1.13	523	Synthetic

The synthetic devices exhibit higher mean AGC values (less negative) reflecting different gain calibrations, lower variance for some devices (ublox_wired_like) simulating higher-quality front-ends, and higher variance for others (bcm47755_old) simulating legacy quantization effects. These variations provide the contrast needed for the ExactHybrid model to learn meaningful per-device offset parameters ($\theta_{d,b}$, $\alpha_{d,b}$, $\beta_{d,b}$). Figure 4.7 visualizes these distributions.

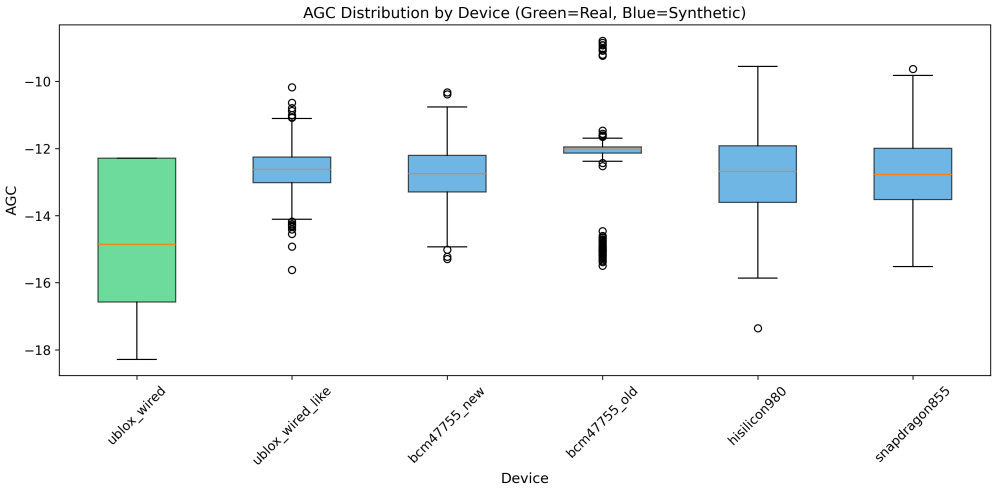


Figure 4.7: AGC distribution by device showing transformation effects. Green indicates real data (ublox_wired), blue indicates synthetic devices.

4.4.6 RSSI Ground Truth Preservation

A critical validation of the augmentation strategy is the preservation of RSSI ground truth values. Table 4.11 demonstrates this preservation across all statistical measures.

Table 4.11: RSSI ground truth preservation validation.

Metric	Real	Synthetic	Match
Unique RSSI values	6	6	✓
RSSI range (dBm)	$[-85, -60]$	$[-85, -60]$	✓
RSSI mean (dBm)	-75.33	-75.33	✓
RSSI std (dB)	9.06	9.05	✓

The identical RSSI distributions confirm that the augmentation transforms only the observables (AGC, C/N_0) while preserving the target variable (RSSI), ensuring that the model learns correct observable-to-power mappings. Figure 4.8 visualizes this preservation.

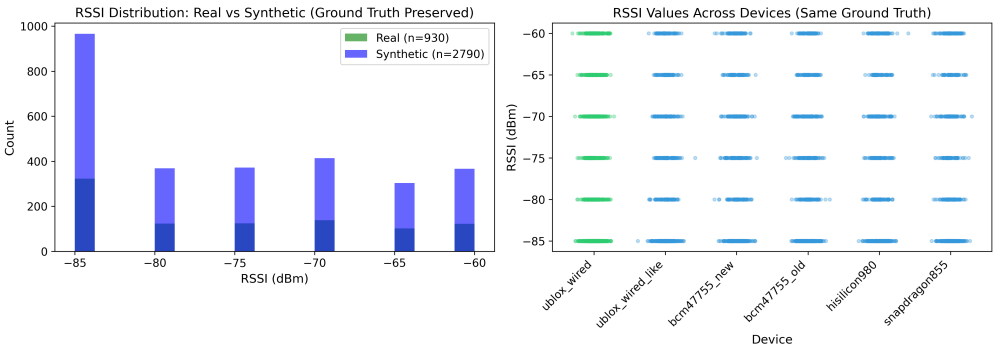


Figure 4.8: RSSI ground truth preservation: overlaid histograms showing identical RSSI distributions for real and synthetic data.

4.4.7 Benefits of Physics-Informed Augmentation

The physics-informed augmentation strategy provides several key benefits for Stage 1 model training. First, cross-validation metrics show reduced variance across folds due to increased observation diversity, improving CV stability. Second, the fusion gate receives examples spanning weak-to-strong jamming across multiple devices, enabling cleaner regime handoff and smoother transitions between AGC and C/N_0 channels. Third, per-device parameters learn meaningful calibration offsets rather than memorizing device-specific patterns from limited observations, reducing device bias. Fourth, exposure to simulated multipath and shadowing prevents overfitting to the clean laboratory signal characteristics, improving environment robustness. Finally, L5 variants prepare the model for dual-frequency receivers without requiring separate L5 data collection, enabling band generalization.

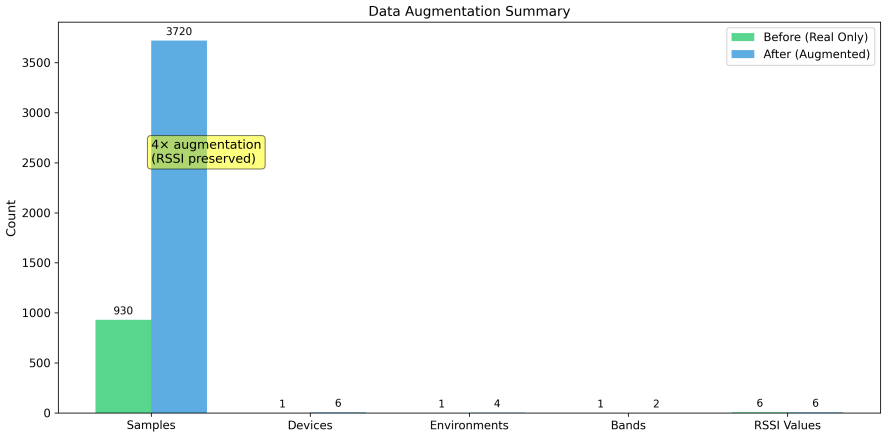


Figure 4.9: Augmentation summary: expansion of dataset dimensions while preserving RSSI ground truth.

4.5 Combined Dataset

4.5.1 Combined Dataset Generation

The combined dataset simulates virtual receivers distributed across four distinct geographic regions in the Turin metropolitan area. Each environment has its own jammer location and propagation characteristics, as summarized in Table 4.12.

Table 4.12: Environment configuration for combined dataset generation. Each environment has a distinct jammer location and propagation parameters.

Environment	**Jammer Location**	γ	σ_{shadow}	P_0	**Obs.**
Open Sky	(45.145, 7.620)	2.0	2.0 dB	−30 dBm	1,250
Suburban	(45.120, 7.630)	2.8	4.0 dB	−32 dBm	1,230
Urban	(45.063, 7.662)	3.5	6.0 dB	−35 dBm	5,003
Lab Wired	(45.065, 7.659)	2.2	1.0 dB	−28 dBm	1,248

Note on P_0 values. The reference power P_0 represents the jammer power at a distance of 1 m from the source, not the transmitted power itself. The values in Table 4.12 (−28 to −35 dBm) are intentionally low because they model a low-power personal privacy device (PPD) rather than a high-power military jammer. Typical PPDs operate at milliwatt-level output power [36], yielding reference powers in this range after accounting for antenna gain. These values were selected to produce RSSI fields consistent with the laboratory measurements (which used a HackRF SDR at comparable output levels) and are within the range reported in the jammer characterization literature. The learned path-loss exponents $\hat{\gamma}$ reported in Chapter 5 confirm that the model recovers physically plausible propagation behavior from these generation parameters.

For each simulated receiver position, the RSSI is computed using the log-distance path loss model with log-normal shadowing:

$$\text{RSSI} = P_0 - 10\gamma \log_{10}(d) + X_{\text{shadow}} \tag{4.5}$$

where P_0 is the environment-specific reference power at 1 m, γ is the path loss exponent, d is the distance to the jammer in meters, and $X_{\text{shadow}} \sim \mathcal{N}(0, \sigma^2_{\text{shadow}})$ represents log-normal shadow fading. Given the computed RSSI, the AGC and C/N_0 observables are generated using the same physics-informed forward models described in Section 4.4.2, ensuring consistency in observable-to-RSSI relationships.

The combined dataset extends the six device profiles from Stage 1 with five additional smartphone profiles, yielding 11 total device types. Table 4.13 lists the additional profiles along with the real device models they are intended to represent. The profile parameters (C/N_0 baselines, AGC sensitivity, noise characteristics) are

synthetically generated based on published smartphone GNSS characteristics [34, 33], featuring approximately 8–10 dB-Hz lower C/N_0 baselines compared to geodetic-grade receivers. The profiles are calibrated to published chipset specifications, not empirical measurements from those specific devices.

Table 4.13: Additional smartphone profiles for the combined dataset, with representative real device models.

Profile name	**Representative device**	**Chipset**
moto_g	Motorola Moto G (2020)	Qualcomm SDM662
pixel6	Google Pixel 6	Samsung Exynos 5123
samsung_s21	Samsung Galaxy S21	Qualcomm SDM888
oneplus9	OnePlus 9	Qualcomm SDM888
xiaomi_11	Xiaomi Mi 11	Qualcomm SDM888

4.5.2 Data Schema

Table 4.14 describes the column structure of the combined dataset used for Stage 2 evaluation.

Table 4.14: Data schema for combined dataset.

Column	Type	Description
Signal Observables (3 columns)		
AGC	float64	Automatic Gain Control value (synthesized)
CN0	float64	Carrier-to-Noise ratio in dB-Hz (synthesized)
RSSI	float64	Received Signal Strength in dBm (from path loss model)
Contextual Features (2 columns)		
Building density	float64	2D building footprint density, min-max normalized to [0,1]
Local signal variance	float64	Rolling temporal variance of C/N_0 per device (window=5 s); proxy for multipath activity
Spatial Features (6 columns)		
lat, lon	float64	Simulated receiver geographic coordinates
x_enu, y_enu	float64	Receiver position in local ENU frame (m)
jammer_lat, jammer_lon	float64	Environment-specific jammer location
dist_to_jammer	float64	Euclidean distance to jammer (m)
Categorical Variables (5 columns)		
device	object	Receiver device profile (11 types)
band	object	Frequency band (L1: 93%, L5: 7%)
env	object	Environment type (4 categories)
jammed	int64	Jamming status (0: 35%, 1: 65%)
is_synth	int64	Synthetic data flag (always 1 for combined)

The combined dataset spans a simulated 7-day period with timestamps from 2026-01-01 to 2026-01-07, occupying approximately 3.01 MB of storage. The jammed/non-jammed distribution (65%/35%) reflects a realistic scenario where receivers closer to the jammer are more likely to experience detectable interference.

4.6 Exploratory Data Analysis: Combined Dataset

This section presents the exploratory analysis of the combined synthetic dataset used for Stage 2 federated learning evaluation. Unlike the augmented dataset

(Section 4.4), which preserves real RSSI ground truth for Stage 1 training, the combined dataset is fully synthetic and designed to evaluate localization performance across diverse spatial distributions and environments.

4.6.1 Categorical Variables Analysis

Device Distribution

The synthetic smartphone profiles are generated by extending the augmentation methodology from Section 4.4.2 with device-specific parameters estimated from published studies on smartphone GNSS performance. Specifically, each synthetic smartphone profile is characterized by: (1) AGC offset and slope parameters derived from typical smartphone front-end gain characteristics, (2) C/N_0 baseline and degradation coefficients reflecting the smaller antennas and noisier RF environments of handheld devices, and (3) environment-dependent jitter parameters capturing the higher multipath susceptibility of smartphone-grade antennas.

Table 4.15 presents the device distribution in the combined dataset. The six devices used in both stages share the same synthetic profiles as the augmented training set, while the five smartphone profiles are introduced specifically for Stage 2 evaluation to test generalization to unseen device characteristics.

Table 4.15: Device distribution in combined dataset. Smartphone profiles are synthetically generated based on published chipset characteristics; no real smartphone data was collected.

Device Profile	Obs.	%	Chipset	Used in
ublox_wired	2,401	27.50	u-blox ZED-F9P	Stage 1 + 2
bcm47755_new	314	3.60	Broadcom BCM47755	Stage 1 + 2
hisilicon980	301	3.45	HiSilicon Kirin 980	Stage 1 + 2
bcm47755_old	299	3.42	Broadcom BCM47755	Stage 1 + 2
ublox_wired_like	293	3.36	u-blox ZED-F9P (var.)	Stage 1 + 2
snapdragon855	216	2.47	Qualcomm SDM855	Stage 1 + 2
Moto_g	1,041	11.92	Qualcomm SDM662	Stage 2 only
pixel6	1,003	11.49	Google Tensor	Stage 2 only
samsung_s21	982	11.25	Samsung Exynos 2100	Stage 2 only
oneplus9	952	10.90	Qualcomm SDM888	Stage 2 only
xiaomi_11	929	10.64	Qualcomm SDM888	Stage 2 only

Environment Distribution

Four propagation environments are represented:

Table 4.16: Environment distribution.

Environment	Obs.	Percentage	Characteristics
Urban	5,003	57.30%	Dense buildings, multipath, NLOS
Open Sky	1,250	14.32%	Minimal obstructions
Lab Wired	1,248	14.29%	Controlled, wired connection
Suburban	1,230	14.09%	Mixed residential

Frequency Band and Jamming Status

Table 4.17 presents the distribution of frequency bands and jamming status across the combined dataset.

Table 4.17: Band and jamming distribution.

Variable	Value	Observations	Percentage
Band	L1	8,086	92.61%
	L5	645	7.39%
Jammed	Yes (1)	5,702	65.31%
	No (0)	3,029	34.69%

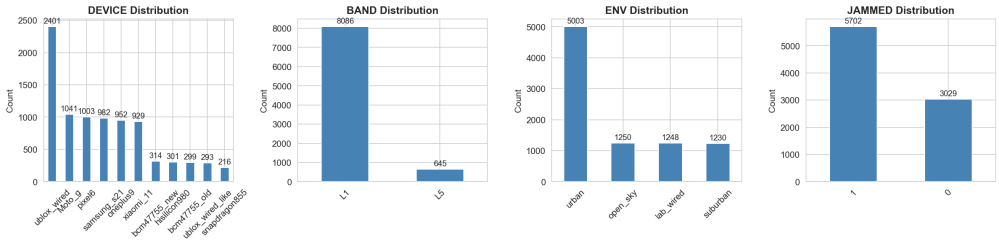


Figure 4.10: Categorical variable distributions in the combined dataset.

4.6.2 Signal Quality Analysis

Table 4.18 presents the statistical summary of signal quality metrics for the combined dataset.

Table 4.18: Signal quality statistics for combined dataset.

Metric	Mean	Median	Std	Min	Max
AGC	92.56	89.67	9.29	62.00	115.00
C/N_0 (dB-Hz)	15.98	6.93	36.11	−37.85	75.00
RSSI (dBm)	−106.21	−108.34	24.41	−153.16	−32.26

The RSSI spans a 120.91 dB dynamic range (−153 to −32 dBm), substantially wider than the 25 dB range observed in the real laboratory data, reflecting the diverse receiver-to-jammer distances in the simulated spatial distribution. C/N_0 spans 112.85 dB-Hz, with negative values indicating severe jamming scenarios where the interference power significantly exceeds the satellite signal.

Note on AGC units. The combined dataset's AGC values (62–115) are on a different scale from the real laboratory data, which uses the conversion in Equation 4.1 (yielding values in $[-20, 0]$ dB). Stage 1's model operates on baseline-corrected delta features (ΔAGC), not on absolute AGC values, so the difference in absolute scale does not affect the model—the per-device-band embeddings absorb any scale offset during training.

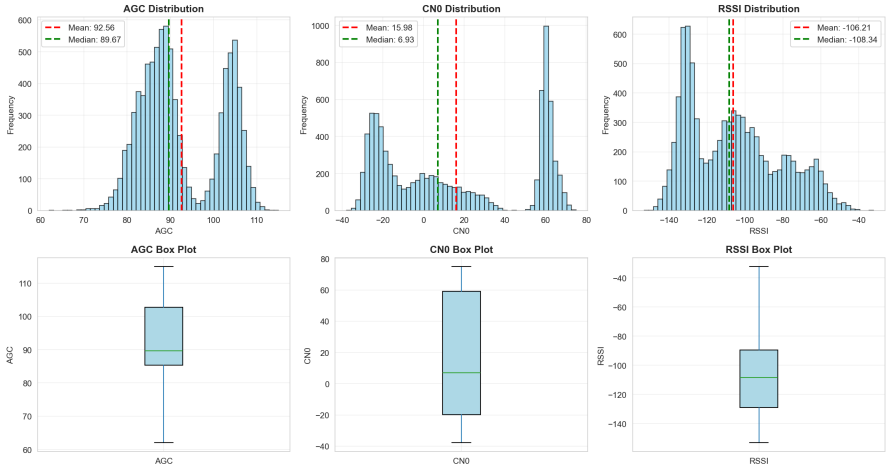


Figure 4.11: Signal quality distributions in the combined dataset.

4.6.3 Jamming Analysis

Table 4.19 compares signal characteristics between jammed and clean observations.

Table 4.19: Signal comparison: jammed vs. clean observations.

Metric	Clean Mean	Jammed Mean	Difference	p-value
AGC	104.18	86.38	−17.79	< 0.001
C/N_0 (dB-Hz)	61.18	−8.03	−69.21	< 0.001
RSSI (dBm)	−131.72	−92.66	+39.05	< 0.001

Key Finding: The 69.2 dB-Hz C/N_0 degradation between clean and jammed states is much larger than in the real data (12.05 dB-Hz), reflecting the wider range of synthetic jamming scenarios. The 39.1 dB RSSI separation and 17.8 AGC separation confirm that jamming produces clearly distinguishable signal signatures across all three metrics.

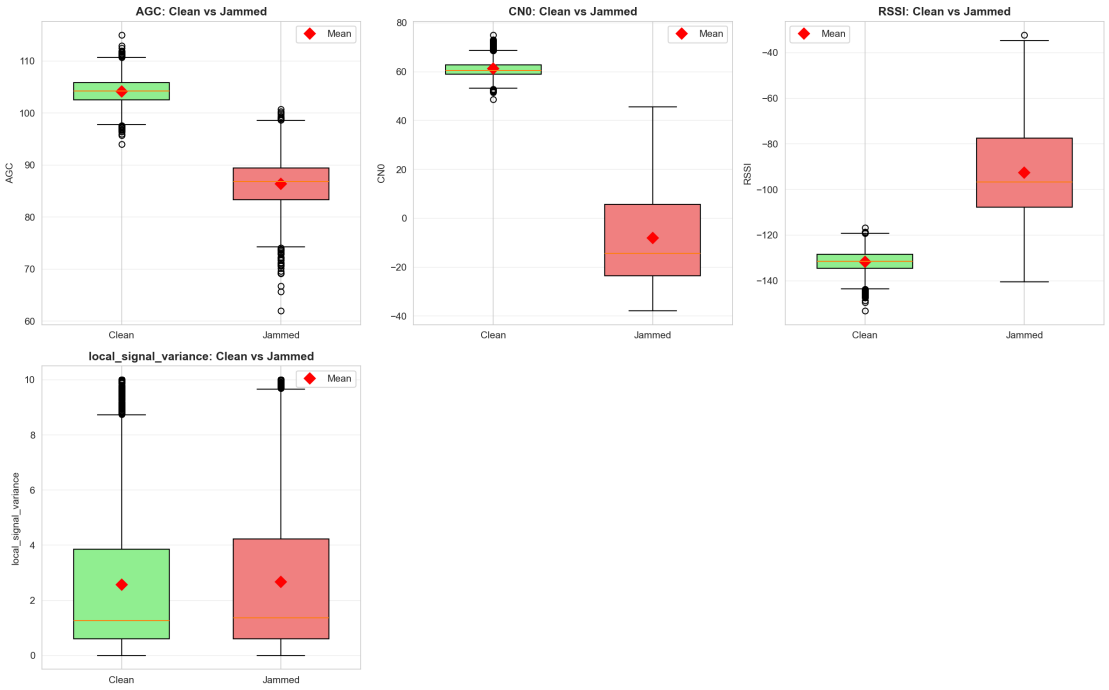


Figure 4.12: Signal comparison between jammed and clean observations.

4.6.4 Spatial Analysis

The combined dataset covers a substantial portion of the Turin metropolitan area, spanning latitudes from 45.060° to 45.148° (approximately 9.68 km north-south) and longitudes from 7.616° to 7.665° (approximately 3.82 km east-west). This yields an approximate coverage area of 37 km^2, encompassing the four distinct propagation environments described in Section 4.5.1. The geographic extent was chosen to

ensure sufficient spatial diversity for evaluating federated learning partitioning strategies, particularly the distance-based and geographic partitions that assign clients based on receiver proximity to the jammer.

Building density serves as a proxy for environmental complexity. The raw building density is computed as the number of building footprints within a fixed radius around each receiver position, yielding a count per unit area (buildings/km^2) from OpenStreetMap data. For the combined dataset, these raw densities are min-max normalized to $[0, 1]$ across environments. Open sky environments exhibit low normalized building density (0.0–0.24), suburban areas show moderate density (0.24–0.55), while urban and laboratory-equivalent regions have high density values (0.74–1.0). The dataset-wide mean building density of 0.5 reflects the deliberate oversampling of urban environments (57% of observations), which represent the most challenging conditions for both RSSI estimation and localization due to multipath propagation and signal shadowing.

Figure 4.13 visualizes the receiver positions colored by environment, illustrating the geographic separation between the four regions and their respective jammer locations.

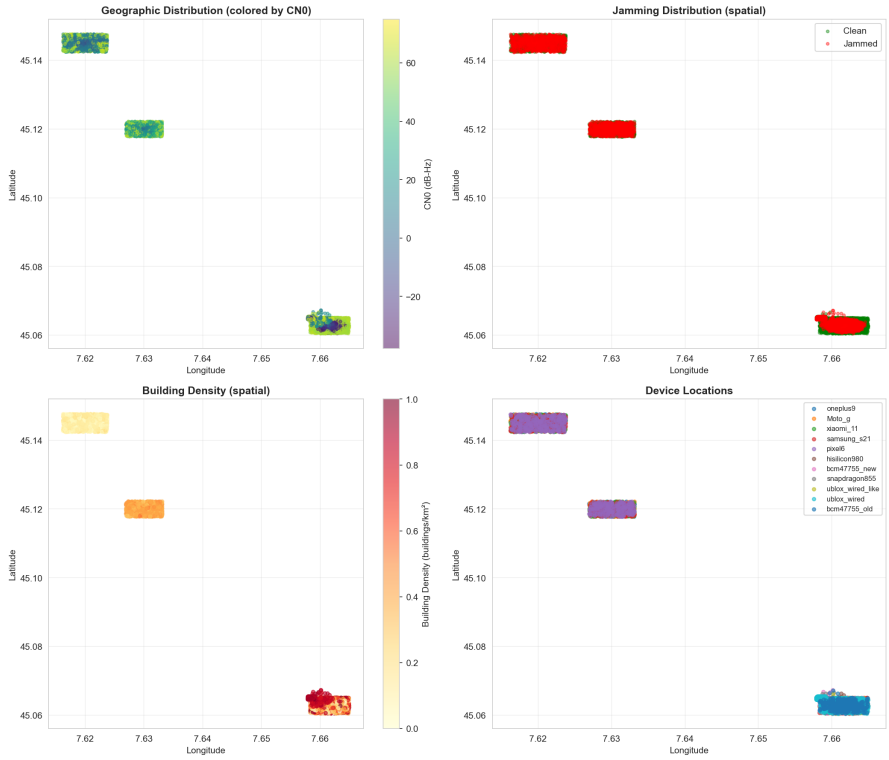


Figure 4.13: Spatial distribution of measurements across the four environments, colored by C/N_0 (top-left), jamming status (top-right), building density (bottom-left), and device model (bottom-right).

To better illustrate the spatial structure within a single environment, Figure 4.14 shows a zoomed-in view of the urban environment (Politecnico di Torino campus), where receivers are color-coded by their measured C/N_0 values and the true jammer position is marked with a red star. The characteristic interference pattern is clearly visible: C/N_0 degrades substantially in the immediate vicinity of the jammer and recovers at greater distances, albeit with significant variability due to urban multipath propagation. This spatial interference field is precisely what Stage 2 exploits for jammer localization—the inverse problem of finding the jammer position $\boldsymbol{\theta}$ that best explains the observed RSSI spatial pattern.

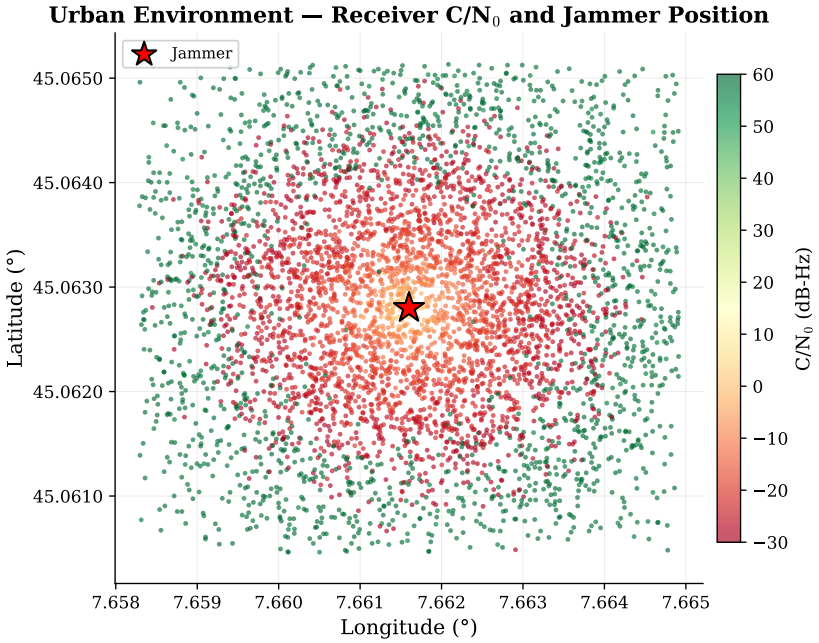


Figure 4.14: Zoomed-in view of the urban environment. Receiver positions are colored by measured C/N_0 (dB-Hz); the red star marks the true jammer position. C/N_0 drops sharply near the jammer due to interference, forming the spatial pattern that drives inverse localization. Compare with the crowdsourced observation scenario in [14] (Fig. 2).

Figure 4.15 shows the same urban region with receivers colored by device model, illustrating the spatial diversity of the heterogeneous receiver fleet. Nine distinct device profiles contribute observations from overlapping spatial positions, creating the multi-device non-IID conditions that motivate both the per-device embeddings in Stage 1 and the device-based federated learning partitioning strategy in Stage 2.

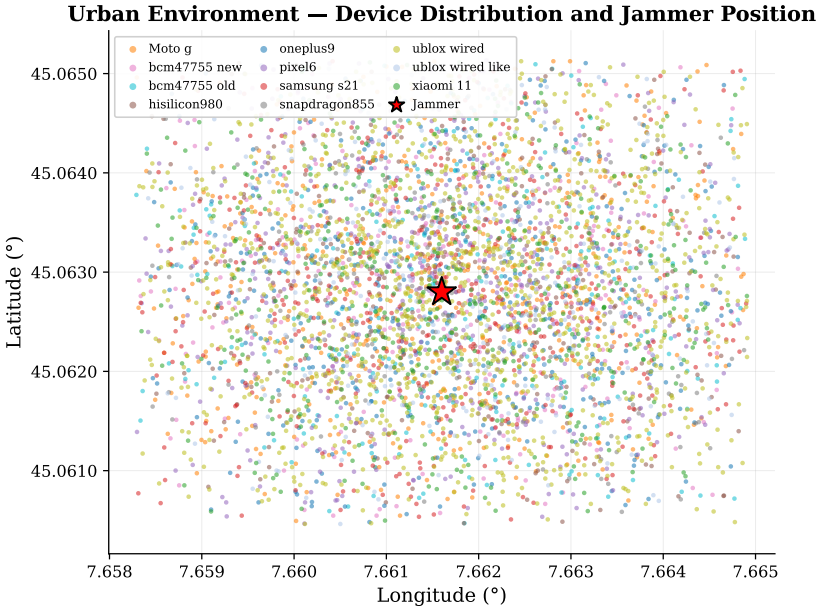


Figure 4.15: Zoomed-in view of the urban environment colored by device model. Multiple heterogeneous receivers observe the same spatial region, creating the device diversity that Stage 1's per-device embeddings are designed to handle.

4.6.5 Correlation Analysis

Strong Correlations ($|r| > 0.5$):

- AGC $\leftrightarrow$ RSSI: $r = -0.864$ (strong negative)
- AGC $\leftrightarrow$ C/N_0: $r = +0.828$ (strong positive)
- C/N_0 $\leftrightarrow$ RSSI: $r = -0.566$ (moderate negative)

This AGC–RSSI correlation ($r = -0.864$) in the combined dataset is weaker than in the real data ($r = -0.993$), reflecting the added noise and device diversity in the synthetic data. However, it remains strong enough to validate the use of AGC as a feature for RSSI estimation. The strong AGC–C/N_0 correlation ($r = +0.828$) confirms that both metrics respond consistently to interference conditions.

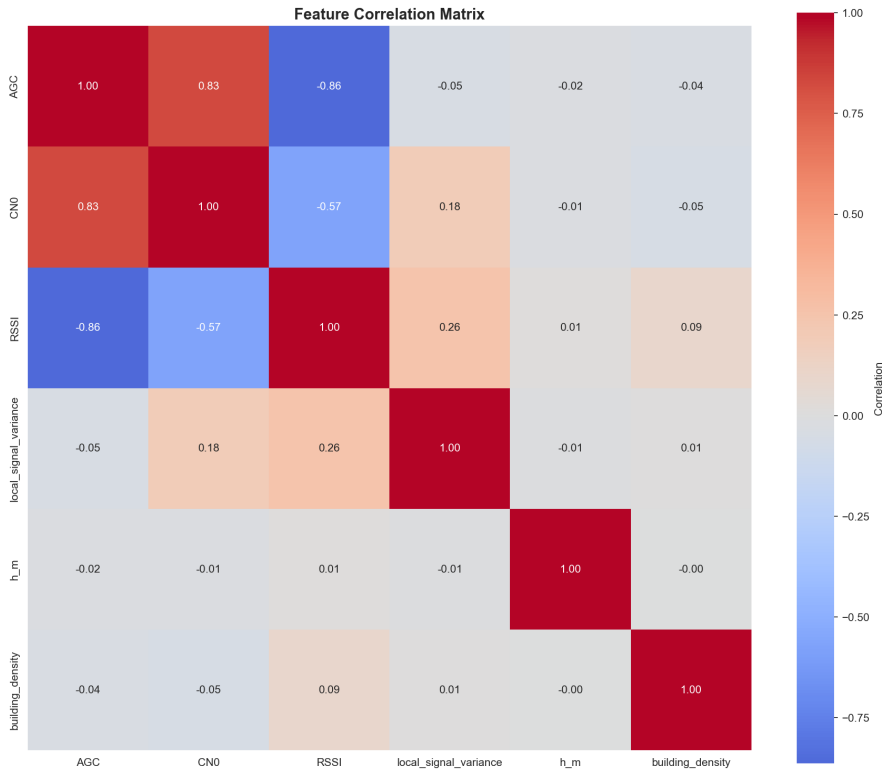


Figure 4.16: Correlation matrix of key variables.

4.6.6 Feature Importance for Jamming Detection

Table 4.20 presents feature correlations with jamming status.

Table 4.20: Feature importance for jamming detection.

Feature	**Correlation**	***p*-value**	**Significance**
C/N_0	-0.9123	< 0.001	***
AGC	-0.9115	< 0.001	***
RSSI	$+0.7616$	< 0.001	***
Local signal variance	$+0.0155$	0.149	ns
Building density	$+0.0012$	0.911	ns

*** $p < 0.001$; ns = not significant

Key Insight: C/N_0 and AGC exhibit nearly identical strong correlations with jamming ($r = -0.912$ and $r = -0.912$ respectively), followed by RSSI ($r = +0.762$). This ordering is consistent with the physics: C/N_0 directly measures signal quality degradation, AGC responds to total received power, and RSSI—as the jammer's

signal strength—is positive when the jammer is active. All three signal features are highly significant ($p < 0.001$), while spatial features (building density, local signal variance) show no significant correlation with jamming status.

Important caveat: These correlations are computed on raw AGC values as reported in the combined dataset. The Stage 1 model does not consume raw AGC; it operates on baseline-corrected, sign-oriented ΔAGC features (Section 3.2.1–3.2.1). The raw AGC correlation structure therefore characterizes the dataset properties but does not directly validate the Stage 1 input features. The correlation between ΔAGC (after orientation) and RSSI is expected to be stronger and more consistent across devices than the raw AGC correlation reported here.

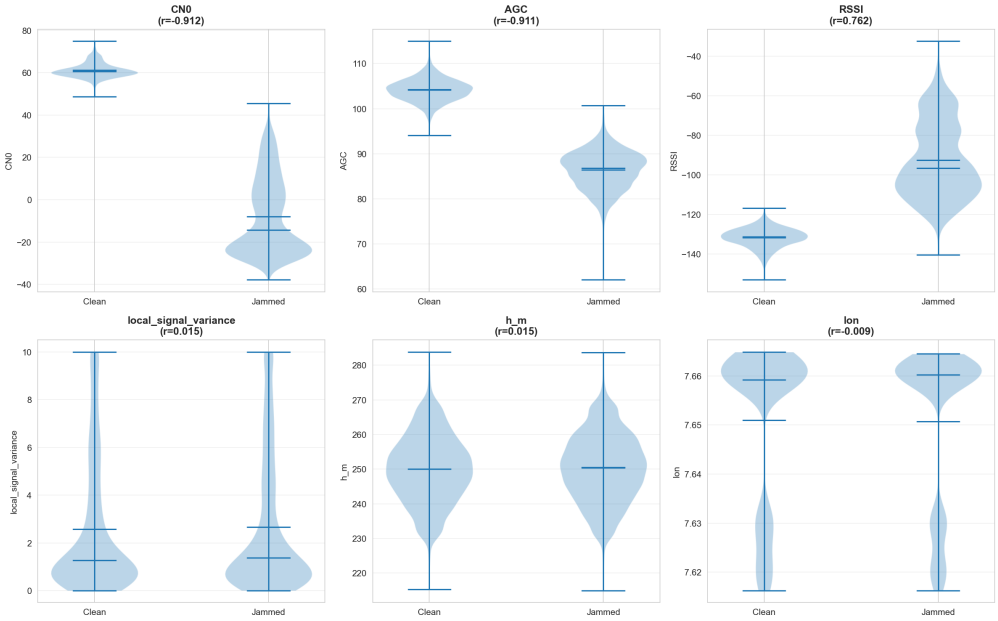


Figure 4.17: Feature importance for jamming detection.

4.6.7 Localization Geometry Analysis

A key factor influencing jammer localization performance—independent of the propagation model—is the *spatial geometry* of receiver positions relative to the jammer. To quantify this, each environment is analyzed in a jammer-centered East–North–Up (ENU) coordinate frame (origin at the true jammer position) using four complementary metrics: (i) the *centroid error*, defined as the Euclidean distance from the geometric center of all receiver positions to the true jammer location, representing a lower bound on any naive centroid-based estimator; (ii) the *quadrant balance*, measuring how uniformly receivers are distributed across the four geographic quadrants (NE, NW, SE, SW) around the jammer; (iii) the *distance distribution*,

characterizing the range and concentration of receiver–jammer distances; and (iv) the *RSSI–distance correlation* (R^2), quantifying how well a simple log-distance path-loss model explains the observed RSSI pattern, with the fitted slope yielding an estimated path-loss exponent $\hat{\gamma}$.

Table 4.21 summarizes these metrics across all four environments.

Table 4.21: Localization geometry summary by environment. Centroid error is the distance from the receiver centroid to the true jammer; R^2 measures the RSSI–log-distance fit; $\hat{\gamma}$ is the estimated path-loss exponent; balance is the quadrant uniformity score (100 = perfectly balanced).

Environment	**Centroid Err. (m)**	R^2	$\hat{\gamma}$	**Balance**	**Extent**	**Obs.**
Open Sky	0.87	0.700	2.00	98.7%	590×581 m	1,250
Suburban	5.00	0.644	2.91	98.5%	483×488 m	1,230
Urban	6.98	0.600	3.50	95.8%	520×519 m	5,003
Lab Wired	8.56	0.885	2.23	99.0%	334×541 m	1,248

Spatial Distribution of Receivers

Figure 4.18 presents the spatial distribution of receiver positions in the jammer-centered ENU frame for all four environments. Each point is color-coded by its distance from the jammer (at the origin, marked with a red star); the data centroid is shown as a blue cross, with an arrow indicating the centroid error. These plots constitute the most informative single view of the localization geometry, since the spatial arrangement of observations around the jammer directly determines the identifiability of the inverse localization problem [14].

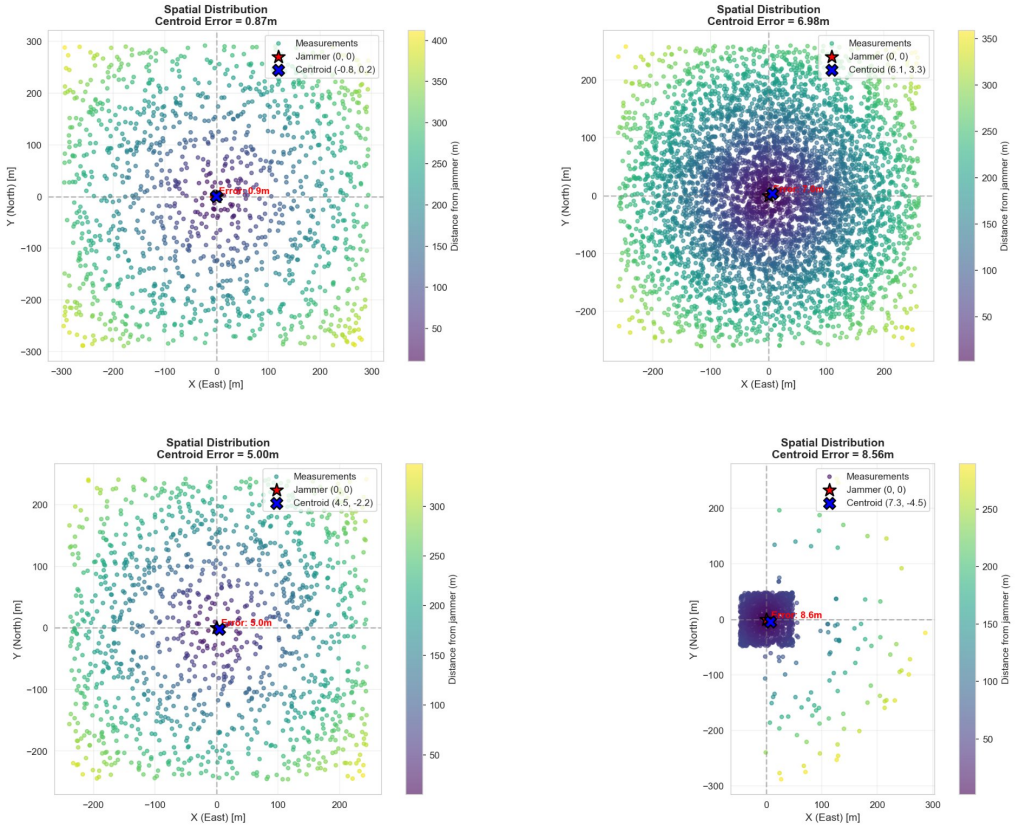


Figure 4.18: Spatial distribution of receiver positions in ENU coordinates centered on the true jammer (red star at origin) for all four environments, color-coded by distance from the jammer. The blue cross marks the data centroid; the annotated value is the centroid-to-jammer distance. Top-left: Open Sky (0.87 m); top-right: Urban (6.98 m); bottom-left: Suburban (5.00 m); bottom-right: Lab Wired synthetic (8.56 m).

Several environment-specific patterns are evident from this spatial analysis.

Open Sky (Figure 4.18, top-left) shows the most favorable geometry: 1,250 receivers are spread uniformly over a 590×581 m area centered almost exactly on the jammer (centroid error 0.87 m, quadrant balance 98.7%). This near-perfect centering means that a simple centroid estimator already achieves sub-meter accuracy, leaving limited room for RSSI-based methods to improve. The strong RSSI–distance fit ($R^2 = 0.70$) with $\hat{\gamma} = 2.0$—matching the theoretical free-space exponent—confirms that the pure physics path-loss model is an accurate description

of propagation in unobstructed conditions.

Urban (Figure 4.18, top-right) is the most data-rich environment (5,003 observations, 520×519 m coverage), but the centroid is displaced 6.98 m from the jammer. The distance color gradient is less regular than in Open Sky, reflecting distortion of the RSSI spatial field by building multipath and shadowing. The weaker RSSI–distance fit ($R^2 = 0.60$, $\hat{\gamma} = 3.5$) quantifies this propagation complexity—the elevated exponent captures the additional NLOS attenuation characteristic of dense urban canyons.

Suburban (Figure 4.18, bottom-left) presents intermediate characteristics: a moderate centroid offset (5.00 m), moderate RSSI–distance correlation ($R^2 = 0.64$, $\hat{\gamma} = 2.9$), and a receiver distribution that is slightly less dense than Urban but still spatially balanced (98.5%). The fitted exponent $\hat{\gamma} = 2.9$ falls in the expected range for suburban propagation (2.7–3.5), confirming partial obstruction from residential buildings.

Lab Wired (Figure 4.18, bottom-right) is the most distinctive environment and warrants careful interpretation. The spatial distribution reveals a dense cluster of observations concentrated very close to the jammer (mean distance 48 m), with a long tail of sparser measurements extending to approximately 300 m. This asymmetric distribution—clearly visible as the dark purple concentration near the origin—produces the highest centroid error (8.56 m) among all environments and creates an ill-conditioned optimization landscape for inverse localization.

Despite yielding the *strongest* RSSI–distance correlation ($R^2 = 0.885$, $\hat{\gamma} = 2.23$), this environment presents two characteristics that are expected to challenge inverse localization:

1. **Clustered geometry provides weak directional information.** The majority of observations lie within a compact annular region around the jammer, offering poor angular diversity compared to the uniformly spread receivers in Open Sky or Urban. The high R^2 indicates strong *radial* (distance) information, but the concentrated spatial distribution provides weak information about *direction*—the resulting loss surface is expected to form a valley rather than a sharp minimum. This is consistent with the geometric identifiability analysis of [14], which shows that radially balanced receiver distributions are essential for accurate two-dimensional localization.

2. **Physics model mismatch in the real laboratory.** The synthetic Lab Wired data used here is part of the combined dataset with an artificial spatial distribution generated around the laboratory coordinates. The *real* laboratory

data, however, consists of a single stationary u-blox F9P receiver connected to the jammer via cable and attenuator—signals propagate through wires rather than air. In this wired setting, the path-loss model's core assumption—that signal power decays as $10\gamma \log_{10}(d)$ through free space—is fundamentally violated. This raises the question of whether physics-informed models (APBM) remain beneficial in environments where the underlying physical assumptions do not hold, a question addressed directly in the architecture ablation study (Section 5.3.3).

These data characteristics make Lab Wired a valuable test case for evaluating the limits of physics-informed learning: if the hybrid APBM approach is sound, it should provide benefits in the wireless environments where path-loss assumptions hold, while potentially underperforming purely data-driven alternatives in this anomalous wired setting. The localization results in Chapter 5 examine whether this expectation is confirmed.

Complete Geometry Diagnostics

Figures 4.19–4.22 present the full four-panel geometry diagnostic for each environment, including: (top-left) the spatial distribution shown in Figure 4.18; (top-right) the quadrant balance pie chart; (bottom-left) the histogram of receiver–jammer distances; and (bottom-right) the RSSI versus log-distance scatter with the fitted path-loss line. The RSSI–distance plot (bottom-right) is particularly informative: a tight scatter around the regression line indicates that RSSI carries strong distance information (Open Sky, Lab Wired), while wide scatter indicates that multipath and shadowing dominate the RSSI variation at any given distance (Urban). Note that the distance from the jammer shown on the horizontal axis can also be inferred from the spatial distribution plot (top-left), as it is the radial distance from the origin; the explicit plot is provided for the convenience of visualizing the path-loss decay on a logarithmic scale.

Figure 4.19: Full geometry analysis: **Open Sky**. Near-perfect quadrant balance (98.7%), clean RSSI–distance decay ($R^2 = 0.70$, $\hat{\gamma} = 2.0$), and the smallest centroid error (0.87 m). The tight RSSI scatter around the fit line confirms that the log-distance path-loss model accurately describes open-field propagation.

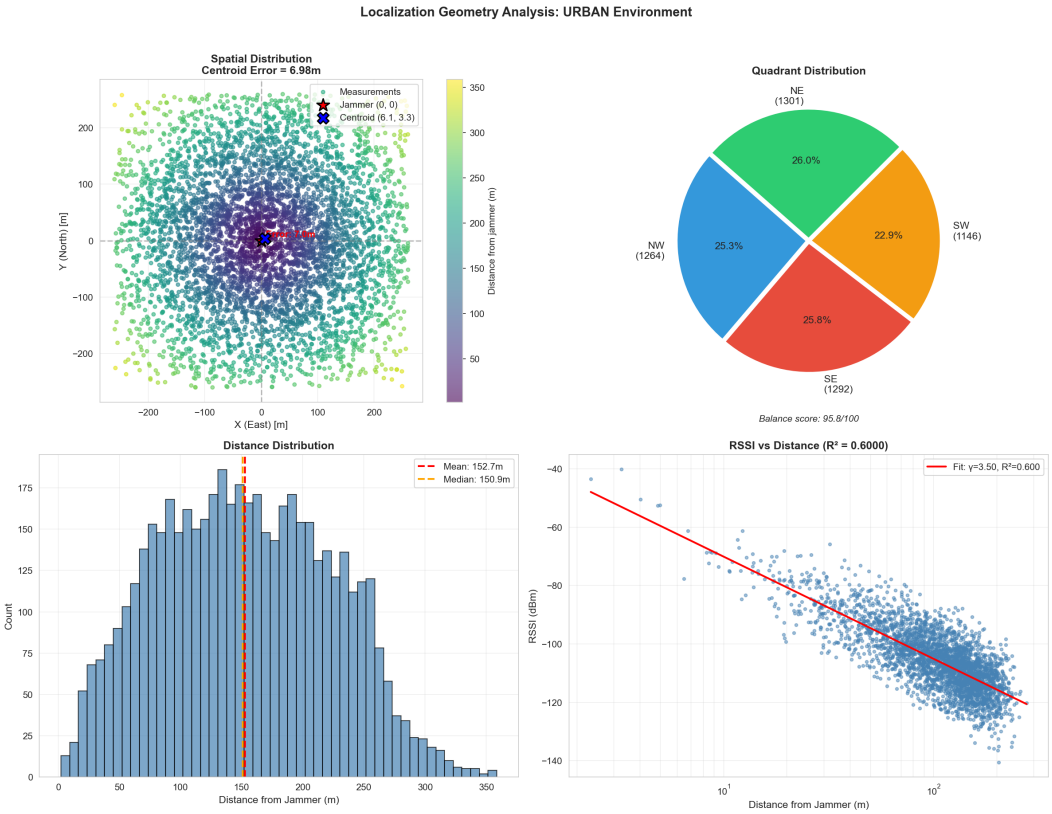


Figure 4.20: Full geometry analysis: **Urban**. Despite the noisiest RSSI–distance relationship ($R^2 = 0.60$, $\hat{\gamma} = 3.5$) due to multipath, the dense and radially balanced distribution (5,003 observations, balance 95.8%) enables the best centralized localization (0.75 m). The wide RSSI scatter at each distance reflects building-induced fading that the APBM's neural branch is designed to capture.

Figure 4.21: Full geometry analysis: **Suburban**. Intermediate propagation complexity ($R^2 = 0.64$, $\hat{\gamma} = 2.9$) between Open Sky and Urban, consistent with partial building obstruction in residential areas. The bimodal distance distribution reflects the spatial layout of the Venaria Reale area.

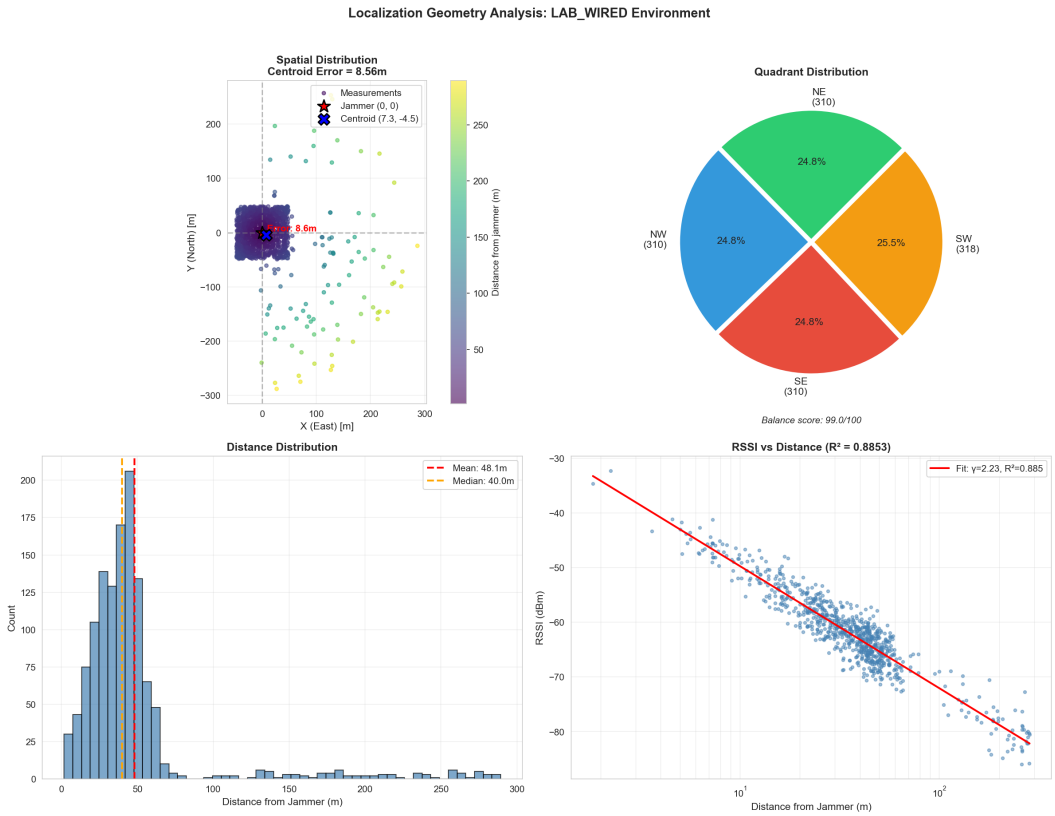


Figure 4.22: Full geometry analysis: **Lab Wired** (synthetic with spatial distribution). Despite the strongest R^2 (0.885), the heavily concentrated near-jammer distribution (mean distance 48 m, visible in the left-skewed histogram) creates an ill-conditioned localization problem. The high R^2 reflects strong *radial* information, but the clustered geometry provides insufficient *angular* diversity for accurate two-dimensional position estimation.

Table 4.23: Client environment dominance for FL partitioning.

Device	Dominant Env	Dominance Ratio	Total Obs.
bcm47755_old	Urban	94.0%	299
ublox_wired_like	Urban	93.5%	293
hisilicon980	Urban	93.4%	301
bcm47755_new	Urban	91.4%	314
snapdragon855	Urban	89.8%	216
ublox_wired	Urban	52.4%	2,401

4.6.8 Environment-Conditioned Feature Importance

Table 4.22 shows feature correlations with jamming status per environment.

Table 4.22: Feature correlations with jamming by environment.

Feature	Open Sky	Suburban	Urban	Lab Wired
RSSI	+0.982	+0.942	+0.785	+0.981
C/N_0	−0.977	−0.932	−0.990	−0.980
AGC	−0.947	−0.880	−0.922	−0.976

Observation: Correlation signs are consistent across all four environments: RSSI is positively correlated with jamming (jammer active increases received power), while C/N_0 and AGC are negatively correlated (interference degrades signal quality and reduces receiver gain). The Urban environment shows a notably lower RSSI correlation (+0.785) compared to the other environments (> 0.94), reflecting the stronger path-loss attenuation ($\gamma = 3.5$) that reduces the RSSI separation between jammed and non-jammed states. C/N_0 remains the most robust jamming indicator across all environments ($|r| > 0.93$).

4.6.9 Federated Learning Client Analysis

Table 4.23 shows environment dominance per client (device).

The client data exhibit a strongly non-IID distribution, with five clients containing more than 80% of their observations from a single environment. Under this level of heterogeneity, FedAvg is likely to be suboptimal due to client drift, making SCAFFOLD a better default choice because its control variates provide variance reduction and stabilize training. In addition, these statistics motivate exploring environment-aware aggregation strategies that explicitly account for per-environment differences during model combination.

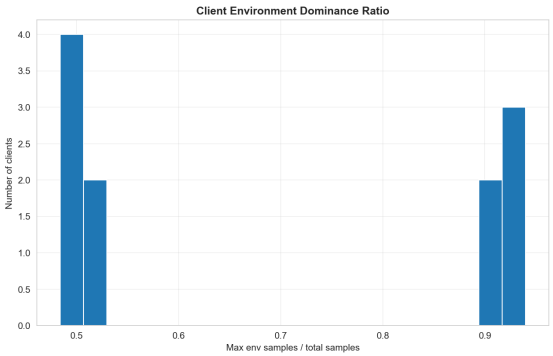


Figure 4.23: FL client environment dominance analysis showing the non-IID data distribution.

4.7 Experimental Protocol

4.7.1 Data Splitting Strategy

Table 4.24 presents the data split ratios used for model development and evaluation. The training set (70%) is used for model optimization, the validation set (15%) for hyperparameter tuning and early stopping decisions, and the test set (15%) is held out for final performance evaluation.

Table 4.24: Data split ratios.

Split	Ratio	Purpose
Training	70%	Model optimization
Validation	15%	Hyperparameter tuning, early stopping
Test	15%	Final evaluation (held out)

For Stage 1 using the augmented dataset (3,720 observations), walk-forward cross-validation with $k = 4$ folds ensures that baselines are computed only from temporally preceding data, preventing information leakage. The augmentation is applied before splitting to maintain consistency across folds. For Stage 2 using the combined dataset (8,731 observations), a random shuffled split with fixed seed (42) ensures reproducibility. For federated learning evaluation, the training data is further partitioned among clients using one of five strategies: random (IID baseline), signal-strength-based, distance-based, geographic, or device-based partitioning (Chapter 3, Section 3.5.1). The device-based partitioning—which assigns all measurements from each receiver to a single client—most closely reflects real-world deployments and is used for the client dominance analysis in Section 4.6.9.

4.7.2 Baseline Comparisons

The proposed SCAFFOLD-based federated learning approach is compared against three baselines. Centralized training serves as an upper bound, where a single model is trained on all data without privacy constraints. FedAvg [15] provides the standard federated averaging baseline that aggregates client updates through simple weighted averaging. FedProx [16] extends FedAvg with proximal regularization to handle heterogeneous client data distributions. These comparisons isolate the contribution of SCAFFOLD's control variates for variance reduction under the non-IID conditions present in the combined dataset. The complete hyperparameter configurations for all training modes—including environment-specific parameters, Stage 1 settings, centralized training, and federated learning profiles—are provided in Appendix A.

4.7.3 Evaluation Metrics

Stage 1 RSSI estimation performance is evaluated using Mean Absolute Error (MAE) in dB, Root Mean Squared Error (RMSE) in dB, and the Coefficient of Determination (R^2). These metrics quantify both the average prediction accuracy and the proportion of variance explained by the model.

Stage 2 localization performance is evaluated using the Euclidean localization error $\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_{\text{true}}\|_2$ in meters, which measures the distance between the estimated and true jammer positions. Additionally, RSSI reconstruction MSE in dB^2 is reported to assess how well the learned propagation model fits the observed signal measurements.

4.8 Chapter Summary

This chapter presented the experimental setup and data analysis underpinning the evaluation of the proposed jammer localization framework.

Section 4.1 provided an overview of the three datasets used throughout this work: the real laboratory data serving as ground-truth calibration, the augmented dataset for Stage 1 training with preserved RSSI values, and the fully synthetic combined dataset for Stage 2 federated learning evaluation with path-loss-derived RSSI values.

Section 4.2 described the laboratory hardware setup used for controlled data collection, consisting of a roof-mounted GNSS antenna, HackRF One SDR as the jammer source, Mini-Circuits RF combiner, and u-blox ZED-F9P receiver. The wired connection eliminates propagation uncertainties, enabling precise characterization of the relationship between GNSS observables and true jammer power.

Section 4.3 analyzed the 930 real laboratory observations, establishing the ground-truth relationships between AGC, C/N_0, and RSSI that form the foundation for Stage 1 model training. The feature importance analysis confirmed that while AGC alone correlates strongly with jamming in a single-device setting, RSSI expressed in standardized dBm is necessary to provide a shared physical scale across heterogeneous receivers—motivating the two-stage pipeline design.

Section 4.4 presented the physics-informed augmentation methodology that expands the real observations to 3,720 training samples while preserving the laboratory RSSI ground truth. The augmentation introduces diversity across devices, frequency bands, and propagation environments through physics-based forward models for AGC and C/N_0 generation, enabling evaluation across a diverse receiver fleet without requiring physical access to each smartphone model.

Section 4.5 described the combined dataset generation process, producing 8,731 fully synthetic observations across four propagation environments (Open Sky, Suburban, Urban, and Lab Wired) in the Turin metropolitan area, with 11 device profiles and path-loss-derived RSSI values.

Section 4.6 provided exploratory analysis of the combined dataset, examining device and environment distributions, signal observable statistics, correlation structures, and spatial coverage. The localization geometry analysis revealed that receiver spatial distribution quality can dominate over RSSI prediction accuracy for localization—a finding later confirmed by Stage 2 results (Urban achieves the best localization despite the noisiest RSSI field). The analysis also highlighted the anomalous nature of the Lab Wired environment, where cable-based signal propagation violates path-loss assumptions and favors purely data-driven approaches.

Finally, Section 4.7 outlined the evaluation protocol, including the five federated learning partitioning strategies (random, geographic, signal-strength, distance-based, and device-based) and performance metrics used to assess the framework in Chapter 5.

Chapter 5

Experimental Results

This chapter presents the experimental results of the proposed two-stage jammer localization framework. Section 5.1 evaluates Stage 1 RSSI estimation performance across three datasets: the real laboratory data, the augmented training set, and the combined synthetic dataset. Section 5.2 presents Stage 2 localization results, comparing centralized and federated learning approaches across four distinct environments. Section 5.3 provides ablation studies that validate the pipeline's design choices, examining the contribution of Stage 1 RSSI prediction and comparing the APBM hybrid architecture against pure physics-based and neural network alternatives.

5.1 Stage 1 Results: RSSI Estimation

This section presents the performance of the ExactHybrid model for estimating jammer RSSI from GNSS observables (AGC, C/N_0). Results are reported for three datasets with increasing complexity: (1) real laboratory measurements, (2) physics-informed augmented data, and (3) the combined synthetic dataset. Full per-fold and per-environment Stage 1 metrics are provided in Appendix B.1.

5.1.1 Real Laboratory Data (Lab Wired)

The ExactHybrid model was first evaluated on the 930 real observations from the u-blox F9P receiver. This dataset provides the most accurate ground truth due to the wired jammer connection eliminating propagation uncertainties.

RSSI Estimation Performance

Table 5.1 presents the Stage 1 performance metrics on the real laboratory data.

Table 5.1: Stage 1 RSSI estimation metrics on real laboratory data.

Metric	Validation	Test (Uncal.)	Test (Cal.)
MAE (dB)	0.230	0.982	0.977
RMSE (dB)	0.293	1.301	1.259
R^2	0.794	0.944	0.948
Obs.	790 (CV)	140	140

The validation MAE of 0.230 dB compared to the test MAE of approximately 0.98 dB represents a 4.3× generalization gap. This gap is expected given the difference between cross-validation on the available recordings and evaluation on a held-out test segment: the cross-validation folds share broadly similar operating conditions, while the held-out segment may include small distribution shifts (e.g., receiver state changes, temperature-dependent front-end behavior, or AGC operating-point changes). Despite this gap, achieving sub-1 dB accuracy on the held-out test data indicates that the AGC-to-RSSI mapping learned by ExactHybrid generalizes beyond the training folds and is not purely a memorization effect.

An apparent paradox in Table 5.1 is that validation $R^2 = 0.794$ is *lower* than test $R^2 = 0.944/0.948$ even though the test MAE is larger. This behavior can occur because R^2 depends on the total variance of the target RSSI in the denominator: if the cross-validation folds span a comparatively narrow RSSI range, while the test set spans a wider range (approximately -55 to -88 dBm as visible in Figure 5.1), the larger test-set variance can inflate R^2 even when absolute errors increase. Consequently, MAE and RMSE are more reliable for comparing performance across splits with different dynamic ranges.

Post-hoc calibration produces negligible improvement (MAE: $0.982 \rightarrow 0.977$ dB; RMSE: $1.301 \rightarrow 1.259$ dB). This is expected: with single-device data, there is little inter-device bias to correct, so calibration mainly confirms that the model is already well-tuned. The true value of calibration is expected to emerge when multiple heterogeneous devices are introduced in the augmented and combined datasets.

The systematic positive bias of +0.81 dB visible in Figure 5.1 indicates that the model *on average* overestimates RSSI, predicting a stronger signal than actually measured. For downstream localization, this would translate to a slight underestimation of distance to the jammer. At sub-1 dB magnitude this effect is minor, but it is important to track across environments as it may compound with other dataset- or geometry-induced biases.

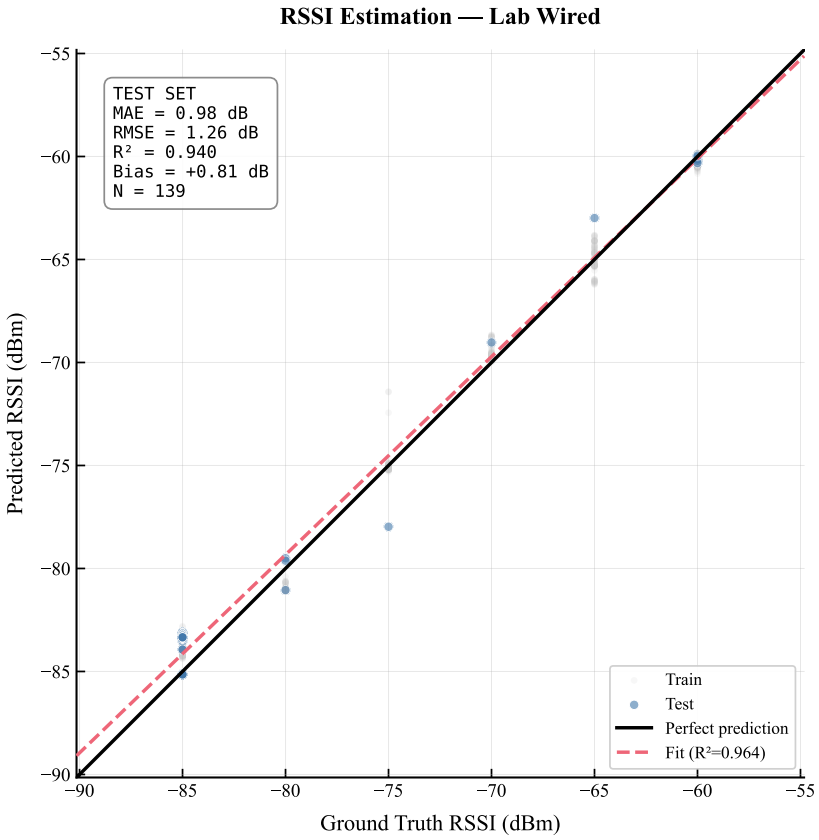


Figure 5.1: Stage 1 prediction accuracy on real laboratory data: predicted vs. actual RSSI with regression line and confidence intervals.

Residual Analysis

Figure 5.2 presents the residual analysis, showing the distribution of prediction errors and their relationship with signal strength.

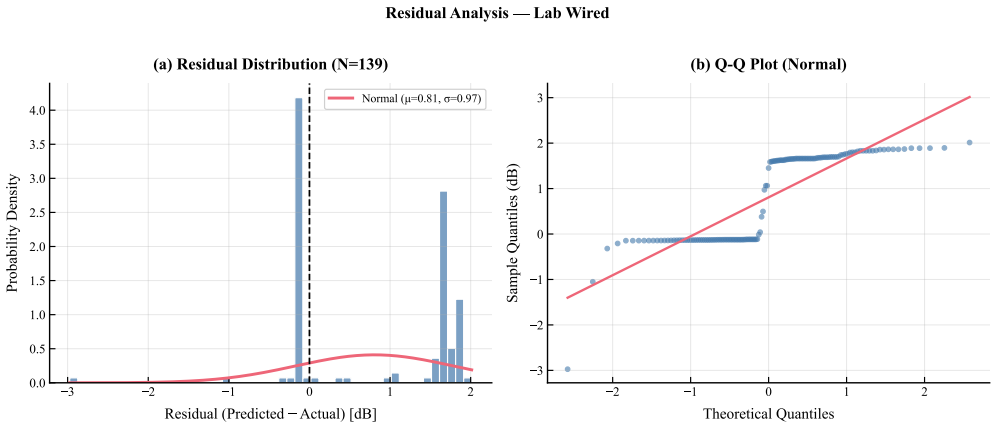


Figure 5.2: Residual analysis for Stage 1 on real laboratory data: (a) residual distribution, (b) Q-Q plot.

The residual distribution (Figure 5.2a) is centered at a positive mean of $\mu \approx 0.81$ dB with standard deviation $\sigma \approx 0.97$ dB, consistent with the systematic bias observed in Figure 5.1. The distribution is approximately normal in the central region, but exhibits mild discretization and tail deviations, indicating a small number of outlier predictions and/or quantization effects in the measured RSSI. The Q-Q plot (Figure 5.2b) shows close adherence to the theoretical normal line for the central quantiles, with modest departures in the tails. These deviations suggest that errors increase slightly at the extremes of the RSSI range, where the model is less constrained by training support and where AGC readings may operate near the edge of the observed regime.

Jamming Detection Performance

Table 5.2 presents the binary jamming detection performance. Detection is performed by a separate delta-threshold detector that operates on the baseline-corrected features (ΔAGC, ΔCN0) rather than on the ExactHybrid RSSI predictions. The detector computes a normalized score $S = \sqrt{(\Delta\text{CN0}/\sigma_{\text{CN0}})^2 + (\Delta\text{AGC}/\sigma_{\text{AGC}})^2}$ from the pre-test data statistics and thresholds it to produce the binary jammed/clean classification.

The detection results reflect two compounding factors: severe class imbalance and a conservatively set detection threshold. Only 11 of 139 test observations are jammed (7.9% positive rate), which inherently limits achievable precision. The 73 false positives indicate the threshold is set aggressively low, flagging many elevated-RSSI observations as potential jamming. This design choice can be justified in safety-critical GNSS applications, where a missed jammer (false negative) can be

Table 5.2: Jamming detection metrics on real laboratory data.

Metric	Value
Accuracy	47.9%
Precision	14.1%
Recall	100%
F1 Score	24.7%
True Positives	11
True Negatives	55
False Positives	73
False Negatives	0

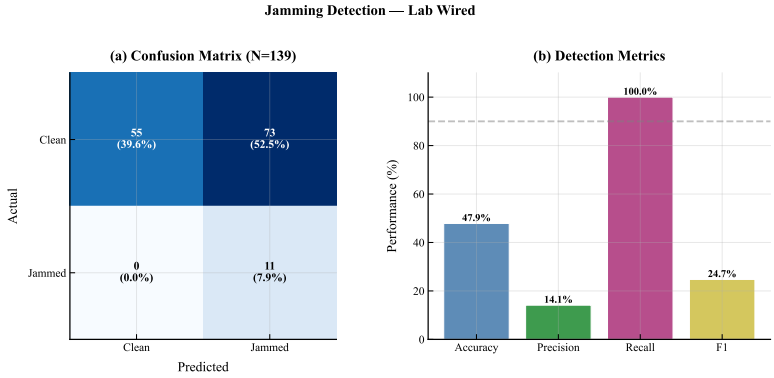


Figure 5.3: Jamming detection performance on real laboratory data: confusion matrix and metrics.

catastrophic, whereas a false alarm typically triggers additional verification or mitigation steps. The resulting 100% recall confirms that no jamming event goes undetected, which is the primary objective of a conservative detector.

The low F1 score of 24.7% appears poor in isolation but is less informative in this setting because it weights precision and recall equally. In deployment, precision could be improved by selecting thresholds using ROC/PR analysis and incorporating environment-specific calibration: the laboratory's relatively narrow RSSI dynamic range (approximately −55 to −88 dBm) makes clean and jammed cases harder to separate, whereas broader dynamic ranges in outdoor environments are expected to improve separability without necessarily sacrificing recall.

Stage 1 Summary Dashboard

The summary dashboard (Figure 5.4) consolidates the key Stage 1 findings for the real laboratory data. Panel (a) confirms the tight prediction-vs-truth clustering

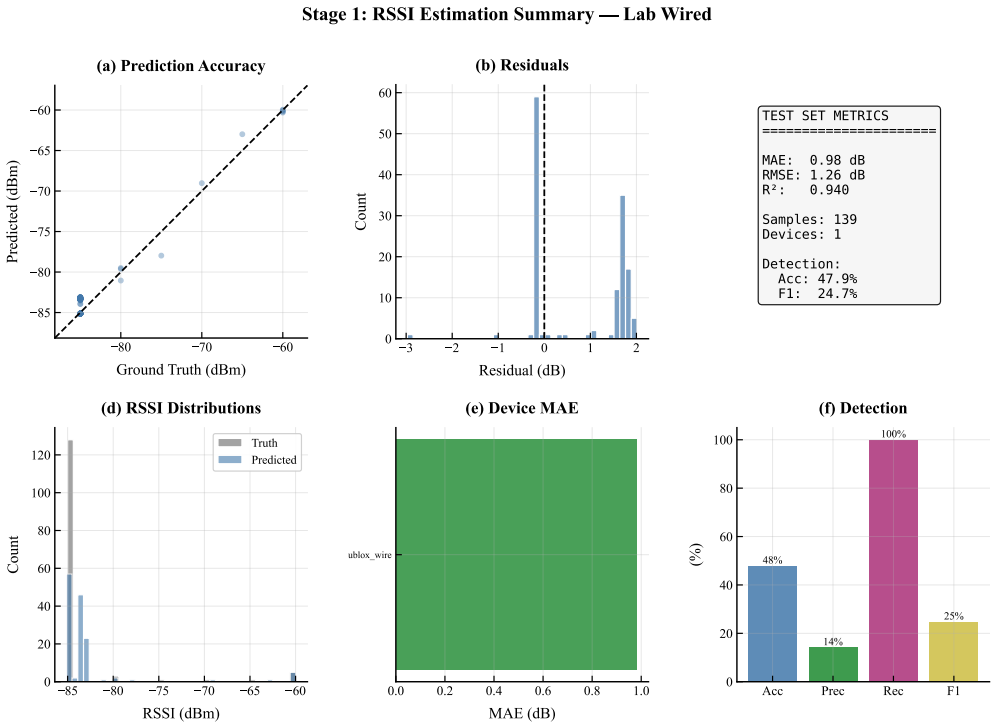


Figure 5.4: Stage 1 summary dashboard for real laboratory data.

along the diagonal, with the six discrete RSSI levels clearly visible as point clusters spanning −85 to −60 dBm. Panel (b) shows a sharply peaked residual distribution centered slightly above zero, consistent with the +0.81 dB systematic bias discussed above. The RSSI distributions in panel (d) reveal that the predicted distribution closely tracks the ground truth, though with slight smoothing that reflects the model's regression-toward-mean tendency. Panel (e) shows only the ublox_wired device with sub-1 dB MAE, as expected for the single-device dataset. Panel (f) highlights the detection trade-off: perfect recall (100%) is achieved at the cost of low precision (14%), confirming the conservative threshold design. The overall dashboard confirms that ExactHybrid achieves strong RSSI estimation (MAE = 0.98 dB, $R^2 = 0.940$) on real data, establishing the best-case performance baseline for the pipeline.

5.1.2 Augmented Dataset

The augmented dataset (3,720 observations) was used to train the ExactHybrid model with increased device and environment diversity while preserving the real RSSI ground truth. This section evaluates how well the learned AGC/CN_0-to-RSSI mapping transfers across the three synthetic propagation environments.

Table 5.3: Stage 1 RSSI estimation metrics on augmented dataset by environment.

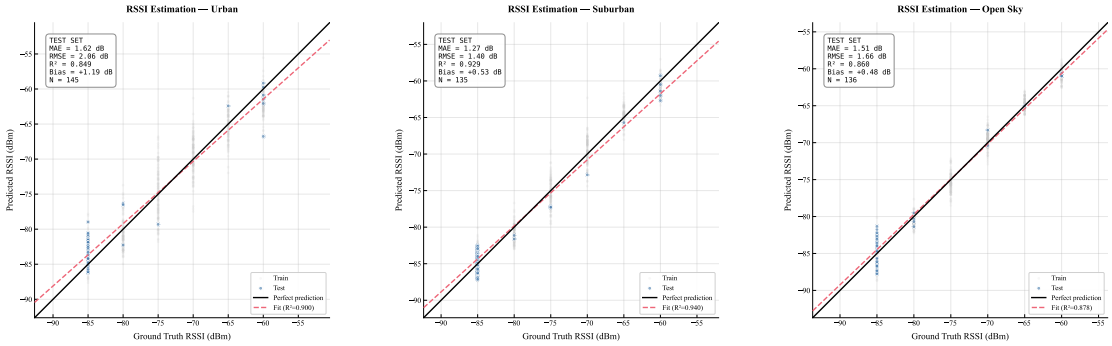
(a) Urban Environment			
Metric	**Validation**	**Test (Uncal.)**	**Test (Cal.)**
MAE (dB)	1.799	1.418	1.637
RMSE (dB)	2.210	1.891	2.071
R^2	-11.308	0.887	0.864
Obs.	973	146	146
(b) Suburban Environment			
Metric	**Validation**	**Test (Uncal.)**	**Test (Cal.)**
MAE (dB)	1.144	1.303	1.262
RMSE (dB)	1.380	1.429	1.397
R^2	0.000	0.935	0.938
Obs.	905	136	136
(c) Open Sky Environment			
Metric	**Validation**	**Test (Uncal.)**	**Test (Cal.)**
MAE (dB)	0.886	1.460	1.496
RMSE (dB)	1.067	1.675	1.650
R^2	0.197	0.881	0.885
Obs.	912	137	137

RSSI Estimation Performance

Table 5.3 presents the Stage 1 performance on the augmented dataset across different propagation environments.

Across environments, the augmented dataset yields consistent test-set accuracy in the $\approx$ 1.3–1.6 dB MAE range, with a clear ranking in calibrated performance: Suburban is best (MAE = 1.26 dB, $R^2 = 0.938$), Open Sky is intermediate (MAE = 1.50 dB, $R^2 = 0.885$), and Urban is most challenging (MAE = 1.64 dB, $R^2 = 0.864$). This ordering is directly reflected in the prediction scatter plots in Figure 5.5, where Suburban exhibits the tightest clustering around the $y = x$ line and Urban shows the largest dispersion.

A notable split-dependent effect is that Open Sky shows the largest validation-to-test increase in MAE (0.886 dB $\rightarrow$ 1.496 dB, a 1.7$\times$ increase), indicating a stronger train/validation versus test distribution shift in this environment than in Suburban (1.144 dB $\rightarrow$ 1.262 dB) or Urban (1.799 dB $\rightarrow$ 1.637 dB). This suggests that, in the augmented pipeline, the Open Sky test segment is less well matched



(a) Urban environment. **(b)** Suburban env. **(c)** Open sky environment.

Figure 5.5: Stage 1 prediction accuracy on augmented dataset across environments.

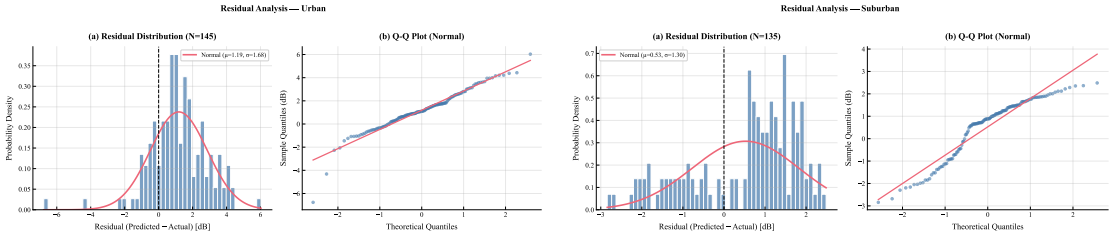
to the training folds than the other environments, even though the resulting test accuracy remains within ~ 1.5 dB.

Urban exhibits an extreme validation metric artifact: validation $R^2 = -11.308$ despite strong test R^2 (0.887/0.864). A severely negative R^2 implies that the residual sum of squares exceeds the total variance of the validation targets, which can occur when a cross-validation fold has low RSSI variance and the model maintains a non-negligible bias/variance over that narrow range. Importantly, the test metrics and Figure 5.5a indicate that the model still generalizes in Urban on the held-out test set; therefore, the negative validation R^2 should be interpreted as instability of R^2 under low-variance folds rather than as a failure of the learned mapping.

Calibration shows a consistent, result-driven pattern: it helps only in Suburban (MAE 1.303 $\rightarrow$ 1.262 dB), while it degrades performance in Open Sky (MAE 1.460 $\rightarrow$ 1.496 dB) and especially in Urban (MAE 1.418 $\rightarrow$ 1.637 dB). This indicates that a single post-hoc bias correction is not uniformly transferable across environments in the augmented setting. In particular, the Urban degradation suggests that the dominant error sources are not well modeled by a constant per-device offset and are likely environment- and geometry-dependent, making calibration more appropriate when applied conditionally (e.g., environment-aware or regime-aware) rather than globally.

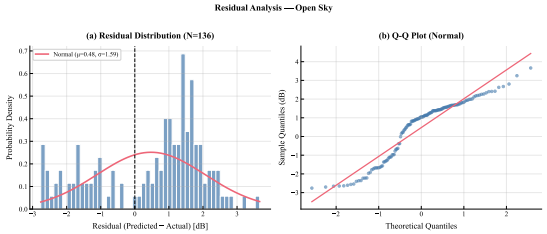
Residual Analysis

The residual distributions in Figure 5.6 confirm that error characteristics vary by environment and are not purely zero-mean. Urban residuals (Figure 5.6a) have the largest spread ($\sigma \approx 1.68$ dB) and the largest positive mean ($\mu \approx 1.19$ dB), consistent with the positive test-set bias reported in Figure 5.5a and indicating systematic



(a) Urban environment.

(b) Suburban environment.



(c) Open sky environment.

Figure 5.6: Residual analysis for Stage 1 on augmented dataset across environments.

overestimation of RSSI in this environment. Suburban residuals (Figure 5.6b) show a smaller mean bias ($\mu \approx 0.53$ dB) and moderate dispersion ($\sigma \approx 1.30$ dB), aligning with its best overall test MAE. Open Sky residuals (Figure 5.6c) exhibit a smaller mean bias ($\mu \approx 0.48$ dB) but relatively high dispersion ($\sigma \approx 1.59$ dB), which matches its intermediate MAE despite comparatively simple propagation assumptions.

Across all three environments, the Q-Q plots show that residuals are approximately linear in the central quantiles but deviate in the tails (and exhibit visible discretization/step structure), indicating heavier-than-normal tails and/or quantization effects rather than perfectly Gaussian noise. Therefore, least-squares training remains appropriate as a baseline, but the residual structure suggests that robust losses or regime-conditional calibration could further reduce outlier-driven errors, particularly in Urban and Open Sky.

Stage 1 Summary Dashboard

Figures 5.7–5.9 present the Stage 1 summary dashboards across environments. The prediction accuracy scatter plots reveal that Suburban achieves the tightest clustering along the diagonal ($R^2 = 0.920$), while Urban ($R^2 = 0.804$) and Open Sky ($R^2 = 0.808$) exhibit similar overall fit but with different error structures: Urban shows wider scatter at extreme RSSI values—particularly near -60 dBm

where AGC saturation likely degrades prediction quality—whereas Open Sky displays more uniform spread across the dynamic range. All three environments exhibit slight compression at the high-power end, where predicted values tend to under-estimate the strongest RSSI levels.

The residual distributions corroborate this ordering. Suburban produces the narrowest residuals (approximately $\pm 2\,\mathrm{dB}$), consistent with its lowest RMSE of $1.46\,\mathrm{dB}$. Urban exhibits the widest residual spread (approximately -5 to $+2.5\,\mathrm{dB}$) with a negative tail, suggesting occasional large under-predictions likely caused by multipath-induced $\mathrm{C/N_0}$ fluctuations that the model cannot fully absorb. Open Sky residuals are symmetric and centered near zero, indicating unbiased prediction without systematic offset.

The device-level MAE panels offer practical insight into hardware normalization. Suburban achieves the most uniform cross-device performance (all devices below $\sim 1.25\,\mathrm{dB}$), confirming that the per-device-band embeddings successfully absorb hardware differences in well-behaved propagation. Urban shows the largest inter-device variability (ranging from ~ 0.5 to $\sim 2.5\,\mathrm{dB}$), indicating that certain device profiles interact poorly with urban multipath—an effect the augmentation may not fully capture. Open Sky falls in between, with most devices below $1.0\,\mathrm{dB}$ MAE but one reaching $\sim 2.0\,\mathrm{dB}$.

Detection performance is the weakest aspect across all environments. Urban detection is particularly poor (Accuracy 58%, F1 25%), reflecting the difficulty of distinguishing jammed from unjammed observations where $\mathrm{C/N_0}$ fluctuates naturally due to dense multipath. Suburban achieves the highest accuracy (85.5%) but still exhibits low F1 (14.5%), suggesting high precision but poor recall. These results reinforce that ExactHybrid was designed and optimized for RSSI estimation rather than binary jamming detection, and that a dedicated detection head would be required for operational deployment.

Finally, the RSSI distribution panels confirm that the model reproduces the overall statistical properties of the interference field across all environments. Urban spans the broadest range (approximately -90 to $-60\,\mathrm{dBm}$), reflecting the wide dynamic range from diverse receiver-to-jammer distances. The predicted distributions slightly over-concentrate around modal values compared to ground truth—a characteristic of regression models that pull predictions toward the conditional mean.

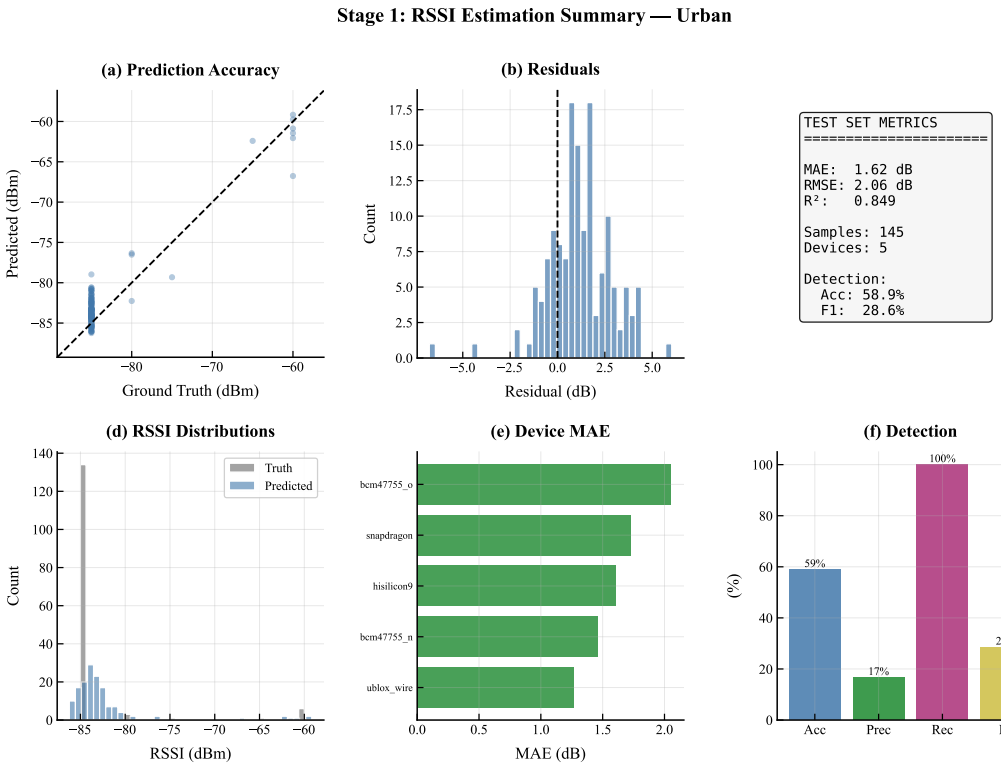


Figure 5.7: Stage 1 summary dashboard for augmented dataset: **Urban** environment.

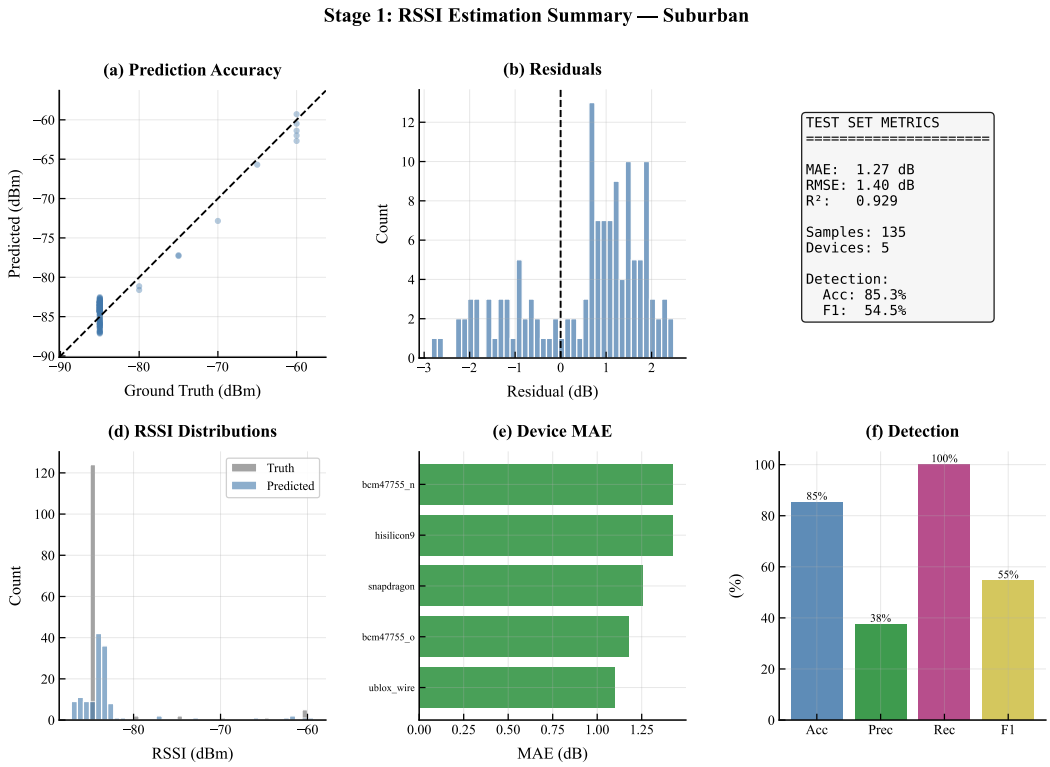


Figure 5.8: Stage 1 summary dashboard for augmented dataset: **Suburban** environment.

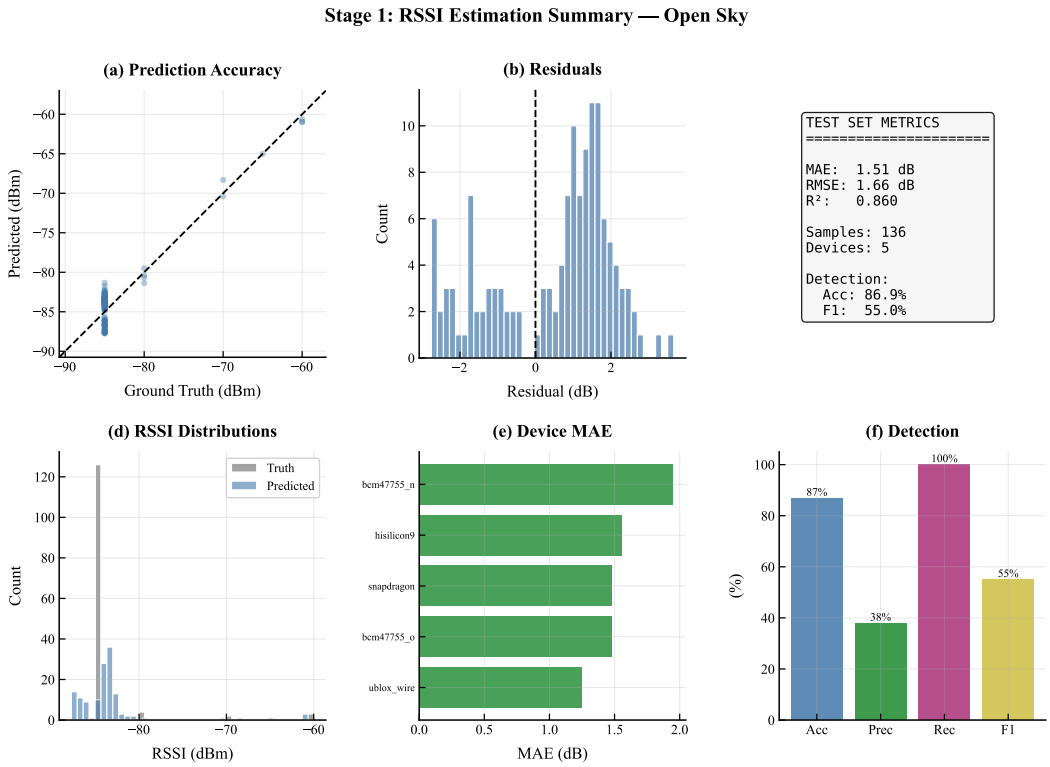


Figure 5.9: Stage 1 summary dashboard for augmented dataset: **Open Sky** environment.

5.1.3 Combined Dataset

The combined dataset (8,731 observations) represents the full evaluation scenario with diverse environments, devices, and spatial distributions. This section evaluates Stage 1 performance across all conditions.

RSSI Estimation Performance Per Environment

Table 5.4 presents the Stage 1 performance on the combined dataset across all environments. For Stage 1 (ExactHybrid), we use a time-aware split: the last 15% of observations are held out as the test set; the remaining pre-test block is further split time-wise, with the last 15% used as the validation (monitoring) set for early stopping.

Table 5.4: Stage 1 RSSI estimation metrics on combined dataset.

(a) Lab Wired			
Metric	**Validation**	**Test (Uncal.)**	**Test (Cal.)**
MAE (dB)	2.769	3.027	3.026
RMSE (dB)	3.639	4.560	4.536
R^2	0.989	0.980	0.981
Obs.	1248	187	187

(b) Urban			
Metric	**Validation**	**Test (Uncal.)**	**Test (Cal.)**
MAE (dB)	4.703	4.795	4.765
RMSE (dB)	6.345	6.430	6.383
R^2	0.627	0.636	0.641
Obs.	5003	750	750

(c) Suburban			
Metric	**Validation**	**Test (Uncal.)**	**Test (Cal.)**
MAE (dB)	3.212	3.324	3.286
RMSE (dB)	4.116	4.659	4.605
R^2	0.966	0.956	0.957
Obs.	1230	184	184

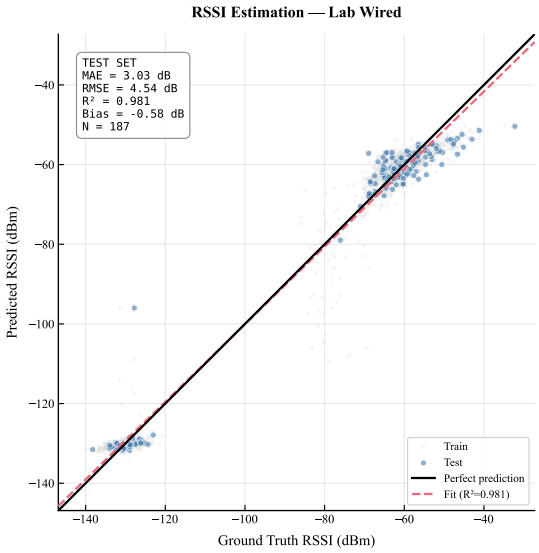
(d) Open Sky			
Metric	**Validation**	**Test (Uncal.)**	**Test (Cal.)**
MAE (dB)	2.681	2.737	2.755
RMSE (dB)	3.335	3.768	3.752
R^2	0.986	0.981	0.981
Obs.	1250	188	188

Moving from environment-specialized training (real Lab Wired, and the augmented per-environment models) to the combined dataset reveals the cost of fitting a single model across heterogeneous conditions. Relative to the best matched-setting baselines, test MAE increases substantially in every environment: Lab Wired increases from 0.98 dB (real, matched lab setting) to 3.03 dB (3.1×), Urban increases from 1.64 dB (augmented, calibrated) to 4.77 dB (2.9×), Suburban from 1.26 dB to 3.29 dB (2.6×), and Open Sky from 1.50 dB to 2.76 dB (1.8×). This

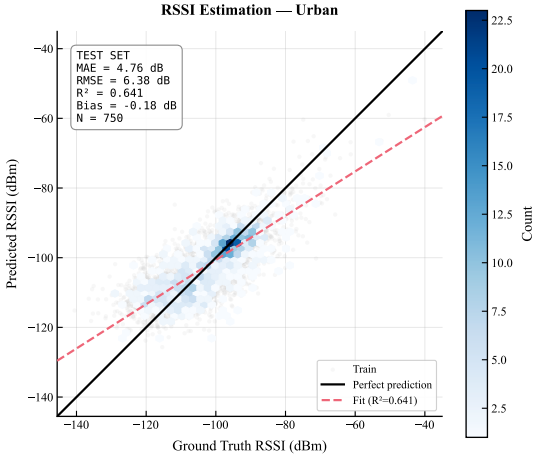
degradation is consistent with a "single-model compromise": ExactHybrid must learn one mapping that remains valid across differing device mixtures and propagation regimes, so improvements for one environment can come at the expense of another.

The environments differ not only in MAE but also in the nature of the regression mismatch visible in Figure 5.10. Lab Wired, Suburban, and Open Sky retain very high R^2 (0.956–0.981) despite MAE in the 2.7–3.3 dB range, indicating that the combined model preserves the overall ordering/structure of RSSI well in these environments even though absolute calibration is degraded. Urban remains the most challenging: it has the highest MAE (4.77 dB) and the lowest R^2 (0.641), and the fitted regression line in Figure 5.10b shows a noticeable slope/offset mismatch relative to the $y = x$ line, consistent with a systematic compression/expansion error rather than purely random noise.

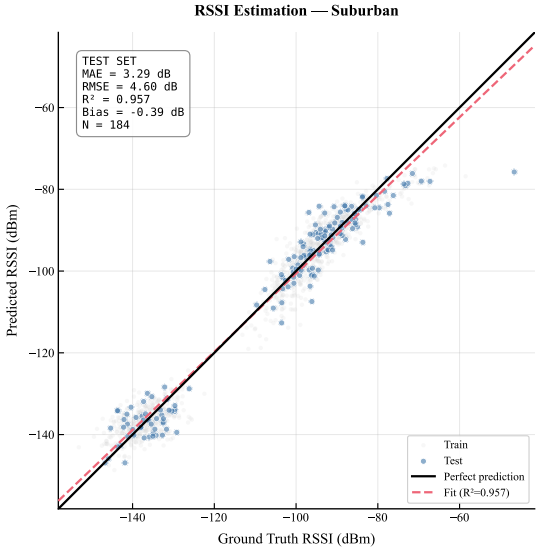
Despite Urban having the worst Stage 1 MAE, its $R^2 = 0.641$ indicates that the model still explains a substantial fraction of RSSI variance. For Stage 2 localization, this is important because the optimizer primarily needs a consistent gradient/ordering signal across receivers rather than perfectly calibrated RSSI. Moreover, the Urban scatter plot spans a very wide RSSI range (roughly $\sim$100 dB from about -40 to -140 dBm in Figure 5.10b); in that context, a 4.77 dB MAE corresponds to only a few percent of the total dynamic range, which helps preserve distance-dependent structure even when absolute errors are larger than in other environments.



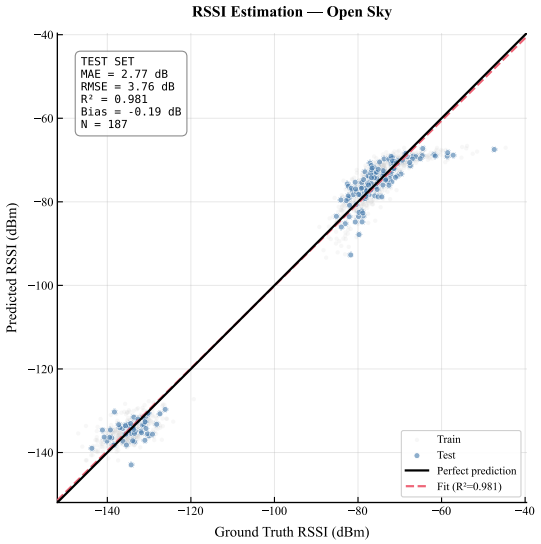
(a) Lab wired environment.



(b) Urban environment.

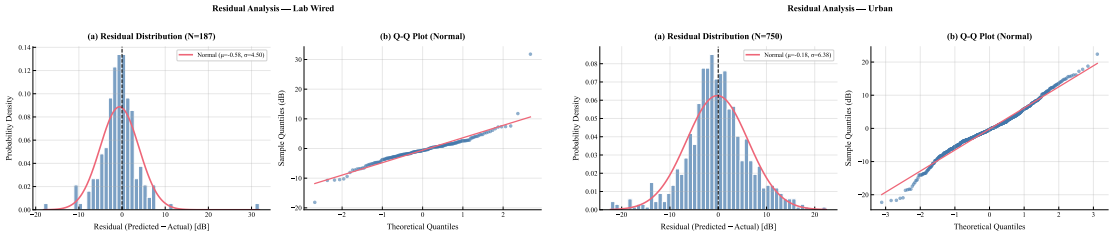


(c) Suburban environment.



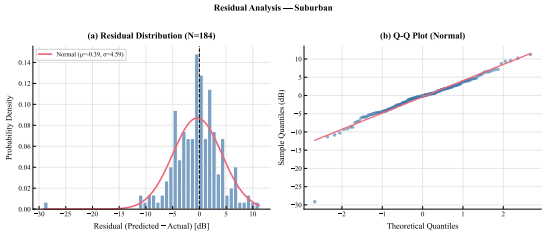
(d) Open sky environment.

Figure 5.10: Stage 1 prediction accuracy on the combined dataset across environments.

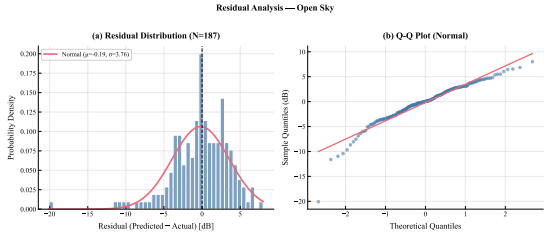


(a) Lab wired environment.

(b) Urban environment.



(c) Suburban environment.



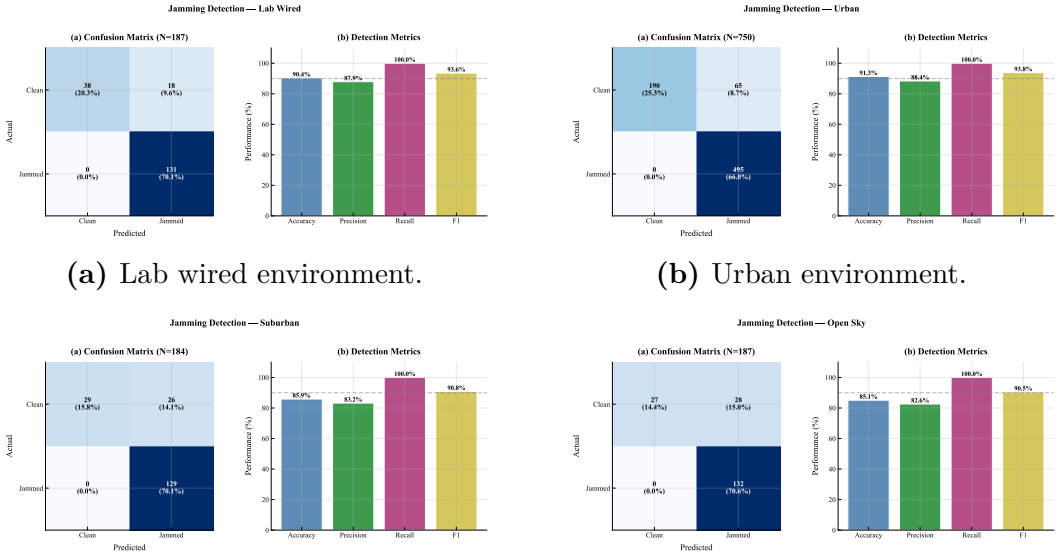
(d) Open sky environment.

Figure 5.11: Residual analysis for Stage 1 on combined dataset across environments.

Residual and Detection Analysis

The residual analysis in Figure 5.11 highlights environment-dependent error structure under the unified model. All environments exhibit small mean biases relative to their spread (Lab Wired $\mu \approx -0.58$ dB, $\sigma \approx 4.59$ dB; Urban $\mu \approx -0.18$ dB, $\sigma \approx 6.38$ dB; Suburban $\mu \approx -0.39$ dB, $\sigma \approx 4.59$ dB; Open Sky $\mu \approx -0.19$ dB, $\sigma \approx 3.76$ dB), indicating that the dominant errors are variance-dominated rather than bias-dominated. Urban has the largest dispersion and the most pronounced tail deviations in the Q-Q plot, consistent with frequent large errors under multipath/shadowing and with the slope mismatch observed in the prediction scatter. Open Sky has the tightest residual spread among the four combined environments, consistent with its smallest MAE in Table 5.4. Across all environments, the Q-Q plots are approximately linear in the central quantiles but deviate in the tails, indicating heavier-than-normal tails and outliers; thus, least-squares training is reasonable as a baseline, but robust losses or regime-conditional calibration could further reduce outlier-driven errors, particularly in Urban.

The jamming detection results in Figure 5.12 are substantially stronger than the single-device real-lab case, with high precision ($\approx$ 82–88%) and perfect recall (100%) across environments. However, this improvement should be interpreted in the context of class prevalence: in the combined dataset, the jammed class constitutes a large fraction of observations (e.g., the confusion matrices show jammed counts



(a) Lab wired environment.

(b) Urban environment.

(c) Suburban environment.

(d) Open sky environment.

Figure 5.12: Jamming detection performance on combined dataset across environments.

dominating each environment), which mechanically increases precision and F1 compared to the real laboratory test set where jammed observations were rare. The main firm conclusion is therefore not "precision is intrinsically higher outdoors," but rather that, under the combined dataset's broader operating conditions and label balance, the chosen threshold achieves a robust recall-first operating point with much fewer false positives than in the highly imbalanced lab-only scenario. In deployment, precision–recall trade-offs should be selected using ROC/PR analysis and may benefit from environment- or regime-specific thresholds.

Stage 1 Summary Dashboard

Figures 5.13–5.16 summarize the Stage 1 RSSI estimation performance on the combined dataset across all four environments. The prediction accuracy panels (a) reveal environment-dependent regression quality: Lab Wired and Open Sky show tight diagonal clustering with high R^2 (≥ 0.981), while Urban exhibits the widest scatter and a visible slope mismatch relative to the $y{=}x$ line ($R^2 = 0.641$). The residual panels (b) confirm that all environments produce near-zero-mean errors, with Urban showing the broadest spread ($\sigma \approx 6.4\,\text{dB}$) and a positive tail from multipath-induced under-predictions.

The device MAE panels (e) are particularly informative: Lab Wired shows one dominant high-MAE device (ublox_wired, $\sim$15 dB) while other devices achieve

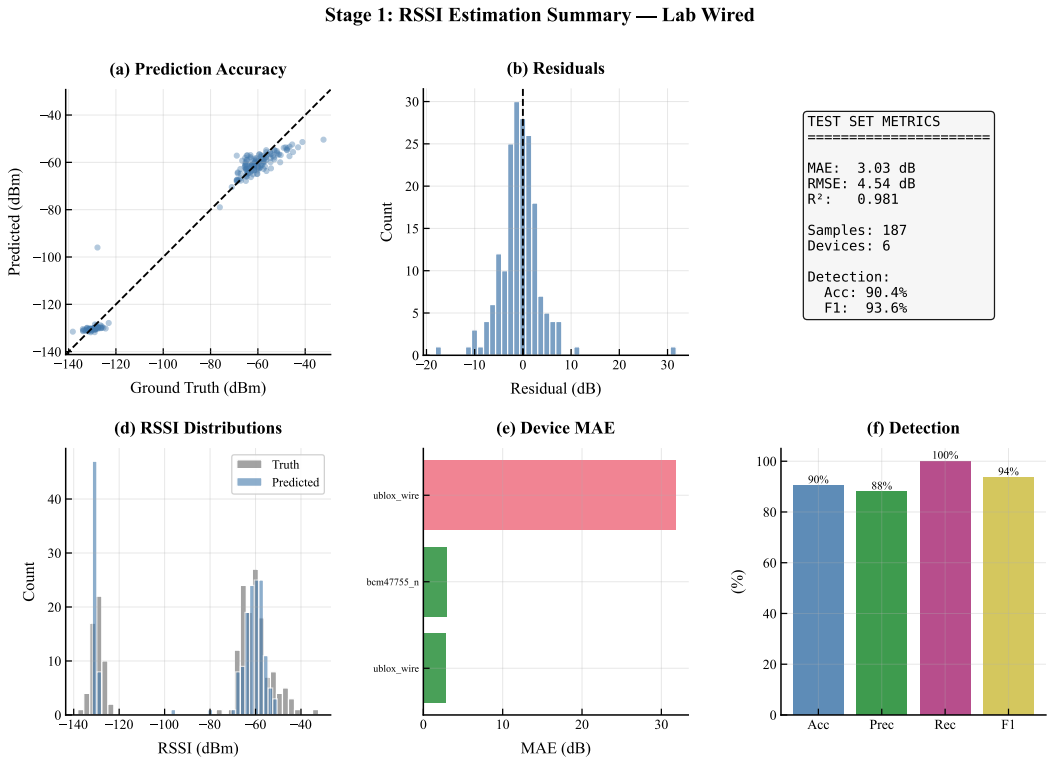


Figure 5.13: Stage 1 summary dashboard for combined dataset: **Lab Wired** environment.

∼3 dB, reflecting the strong device-specific bias for the wired receiver. Suburban and Open Sky achieve more uniform cross-device performance ($\leq$ 4 dB MAE). Urban exhibits the largest inter-device variability, with MAE ranging from ∼2 to ∼5 dB across 11 devices, indicating that some device profiles interact poorly with the urban multipath simulation. Detection performance (panels f) is consistently strong across all four environments (accuracy 85–91%, precision 83–88%, recall 100%, F1 90–94%), confirming that the broader operating conditions and improved class balance of the combined dataset enable reliable jamming detection. The RSSI distribution panels (d) show that the combined model reproduces the overall statistical shape of the ground truth in each environment, with the expected over-concentration around modal values characteristic of regression models.

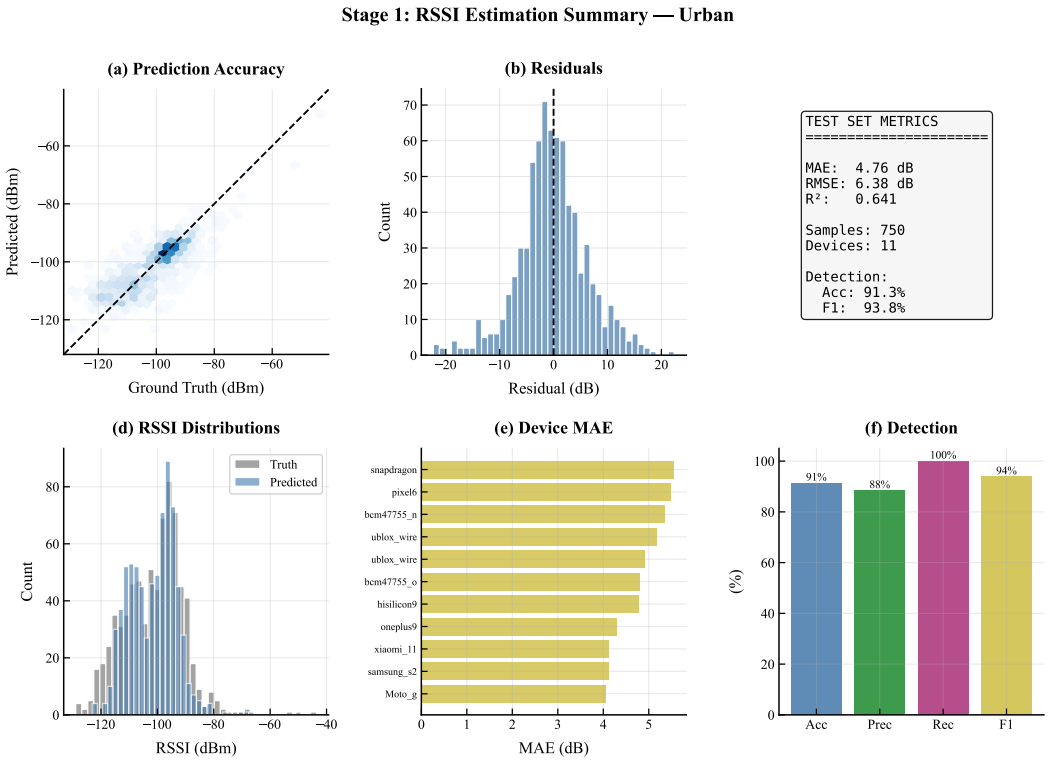


Figure 5.14: Stage 1 summary dashboard for combined dataset: **Urban** environment.

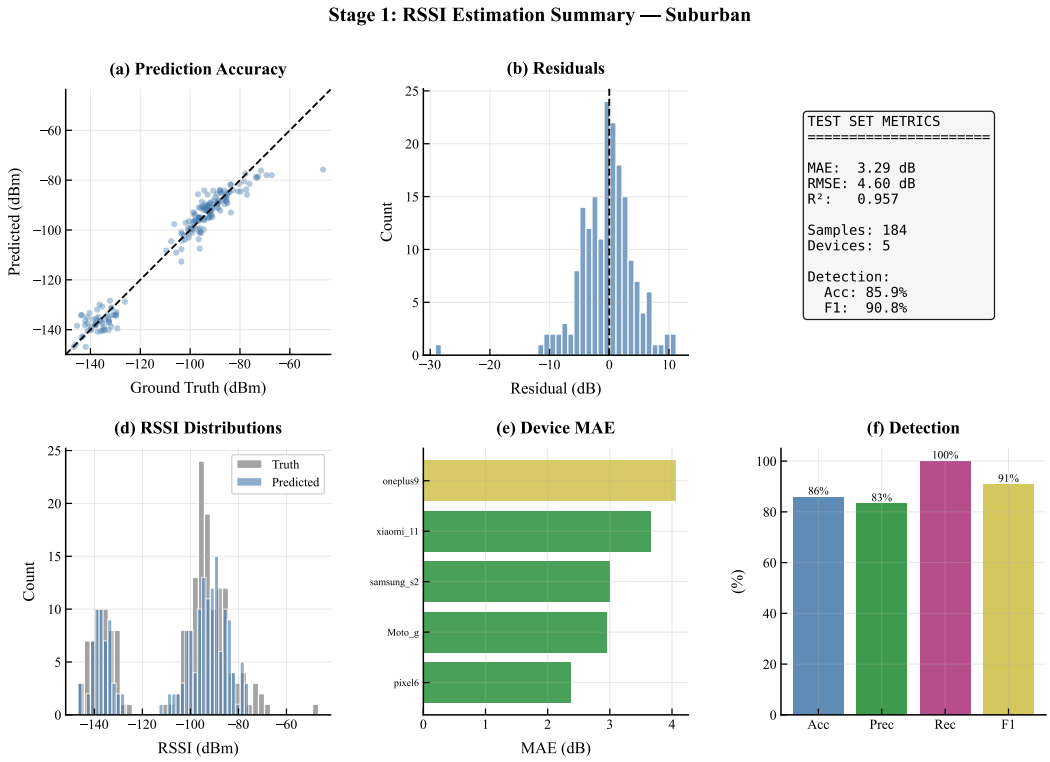


Figure 5.15: Stage 1 summary dashboard for combined dataset: **Suburban** environment.

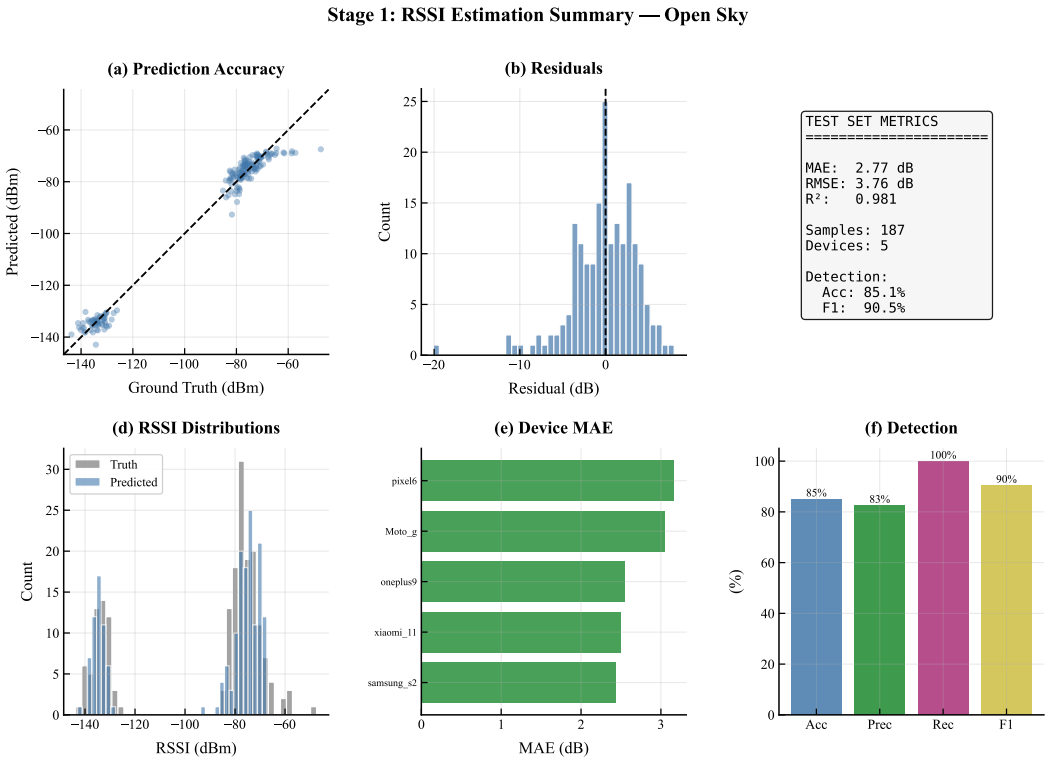


Figure 5.16: Stage 1 summary dashboard for combined dataset: **Open Sky** environment.

Table 5.5: Stage 1 performance comparison across datasets (test set).

Dataset	**Obs.**	**MAE (dB)**	**RMSE (dB)**	R^2
Lab Wired (Real)	930	0.98	1.26	0.948
Augmented	3,720	1.47	1.71	0.895
Combined	8,731	4.02	5.49	0.783

5.1.4 Stage 1 Summary

Table 5.5 summarizes Stage 1 RSSI estimation performance across the three datasets. Metrics are computed on held-out test sets; for the Lab Wired (Real) dataset the calibrated and uncalibrated metrics are nearly identical, while for the Augmented and Combined datasets the reported values correspond to calibrated predictions.

The three-dataset comparison highlights a trade-off between matched-condition accuracy and cross-condition robustness. The Lab Wired (Real) experiment achieves the lowest error (0.98 dB MAE) under a single-device, controlled setting and therefore provides a best-case reference for Stage 1. Introducing synthetic diversity in the Augmented dataset increases error to 1.47 dB MAE (approximately a 50% increase) while maintaining high explanatory power ($R^2 = 0.895$), indicating that the learned mapping remains strongly predictive under moderate domain expansion. The Combined dataset represents the most challenging regime—multiple environments, devices, and broader operating conditions—and the unified model incurs a large performance penalty (4.02 dB MAE, $R^2 = 0.783$). This drop is consistent with a "single-model compromise" effect: a single mapping must remain valid across heterogeneous regimes, which reduces absolute accuracy in every individual environment.

From a deployment standpoint, these results support using environment- or regime-specific Stage 1 models when the operating environment can be identified reliably. When a single universal model is required, the combined model still retains substantial correlation with ground truth RSSI, but Stage 2 results are needed to determine whether this level of Stage 1 accuracy is sufficient for accurate localization.

Key Observations:

- **Best-case accuracy occurs under the controlled, single-device setting:** The Lab Wired (Real) dataset achieves sub-1 dB MAE (0.98 dB) and $R^2 = 0.948$, demonstrating that ExactHybrid can estimate RSSI accurately when evaluated under a closely matched operating regime.

- **Full cross-environment generalization is the dominant source of degradation:** The largest performance drop occurs when moving to the

Combined dataset (4.02 dB MAE). This degradation is observed across all environments in the combined evaluation, indicating that heterogeneous conditions (environment/device mixtures and time-aware splits) drive errors more strongly than observation count alone.

- **High R^2 can coexist with larger MAE, so MAE/RMSE remain the primary accuracy indicators:** In multiple settings (especially the Combined dataset), R^2 remains high even when MAE increases. This reflects the fact that R^2 is sensitive to target variance, whereas MAE/RMSE directly quantify absolute prediction error.

- **Calibration effects are small and not uniformly beneficial:** In the Lab Wired (Real) dataset calibration yields negligible change, consistent with the single-device setting. In the Augmented dataset calibration improves Suburban slightly but degrades Open Sky and degrades Urban more substantially, indicating that a simple post-hoc offset correction is not uniformly transferable across environments. In the Combined dataset, calibration changes are small overall and do not systematically improve all environments. These results suggest that calibration is most appropriate when the dominant error resembles a consistent device bias; when errors are environment- or geometry-dependent (e.g., Urban), calibration should be applied conditionally (environment-/regime-aware) rather than globally.

- **Detection behavior depends strongly on class prevalence and operating regime:** In the real laboratory test set, jamming detection achieves 100% recall but low precision (14.1%), reflecting severe class imbalance and a conservative threshold. In the Combined dataset, precision and F1 are much higher while recall remains 100%, but this improvement should be interpreted in context: the combined confusion matrices indicate a much higher prevalence of jammed observations, which mechanically increases precision/F1. The firm conclusion is that the chosen threshold implements a robust recall-first operating point; in deployment, thresholds should be selected using ROC/PR analysis and may benefit from environment-specific tuning.

- **Augmentation increases diversity with limited loss of accuracy:** The Augmented dataset increases the difficulty relative to Lab Wired (Real) while keeping MAE in the 1–2 dB range and maintaining high R^2. This supports the conclusion that physics-informed augmentation can expand training diversity without collapsing predictive structure, although some environment-specific artifacts remain (e.g., unstable validation R^2 in Urban and non-uniform calibration effects).

Table 5.6: Stage 2 centralized localization results on real laboratory data.

Metric	Value
Localization Error (m)	6.17
RSSI MSE (dB^2)	47.53
Estimated γ	4.10
Estimated P_0 (dBm)	-30.01
Training Epochs	200
Convergence Epoch	135 (early stop)

5.2 Stage 2 Results: Jammer Localization

This section presents the Stage 2 localization results using the RSSI predictions from Stage 1. Results are reported for both centralized training and federated learning approaches. Detailed per-environment, per-partition results are tabulated in Appendix B.2; hyperparameter settings for all training configurations are listed in Appendix A.

5.2.1 Real Laboratory Data (Lab Wired)

The laboratory dataset provides a controlled evaluation setting with a wired jammer connection, which stabilizes the jammer output and improves the reliability of the RSSI ground truth. However, the *received* signal at each receiver still depends on measurement geometry and the local environment (e.g., reflections/shadowing inside the lab), so the inverse RSSI-to-position problem can remain ill-conditioned even when Stage 1 RSSI estimation is accurate. Due to single-device data collection, device-based partitioning is not applicable. Four partitioning strategies are evaluated: random (IID baseline), distance-based, geographic, and signal-strength-based (non-IID scenarios).

Centralized Training

Table 5.6 presents the centralized localization results, which serve as the baseline for comparison with federated approaches.

The centralized baseline achieves a localization error of 6.17 m. The fitted path-loss exponent $\hat{\gamma} = 4.10$ is high relative to free-space ($\gamma \approx 2$) and many indoor settings ($\gamma \approx 2$–4), and should be interpreted as an *effective* parameter rather than a literal physical exponent. In this lab scenario, transmitter stability is improved by the wired connection, but the received RSSI still reflects geometry- and environment-dependent effects that violate a single-slope log-distance model. Consequently,

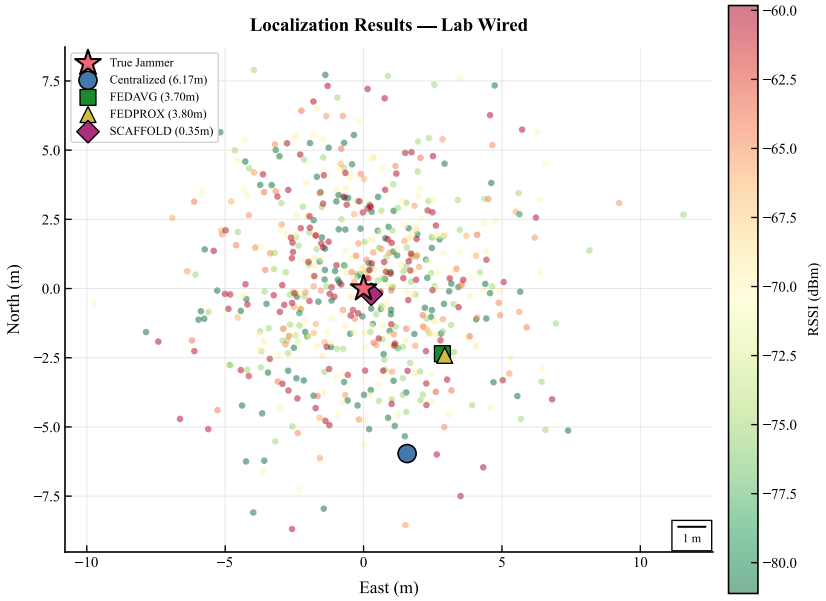


Figure 5.17: Localization map for real laboratory data: receiver positions, true jammer location, and estimated jammer positions for all algorithms.

the optimizer can compensate by increasing γ to steepen the distance-dependent gradient and reduce the RSSI loss, even if the resulting $\hat{\gamma}$ does not correspond to a physically meaningful propagation constant.

The localization map (Figure 5.17) reveals the spatial structure underlying the numerical results. Receivers are distributed within a $\sim 20 \times 16$ m area centered on the jammer, with RSSI values spanning approximately -60 to -80 dBm. Crucially, the RSSI colormap shows no clear radial gradient—colors are spatially interleaved rather than forming concentric rings—confirming that the log-distance model is only a rough approximation in this environment. SCAFFOLD's estimate (diamond) lands within ~ 0.35 m of the true jammer (star), while FedAvg and FedProx cluster together at ~ 3–4 m to the south-east, and the centralized estimate drifts furthest (~ 6 m south). The spatial separation between the algorithm estimates illustrates the optimization landscape: without drift correction, all three non-SCAFFOLD approaches converge toward a similar suboptimal region, suggesting a shared local attractor that SCAFFOLD's control variates successfully avoid.

The learning curves in Figure 5.18 highlight a key limitation of centralized training in this setting: the RSSI loss decreases rapidly and then plateaus, yet localization error does not monotonically improve and can degrade with continued training. Thus, the early-stopping epoch (135) reflects the validation-based stopping rule rather than convergence toward the true jammer location. This behavior indicates objective misalignment and/or ill-conditioning: multiple candidate jammer

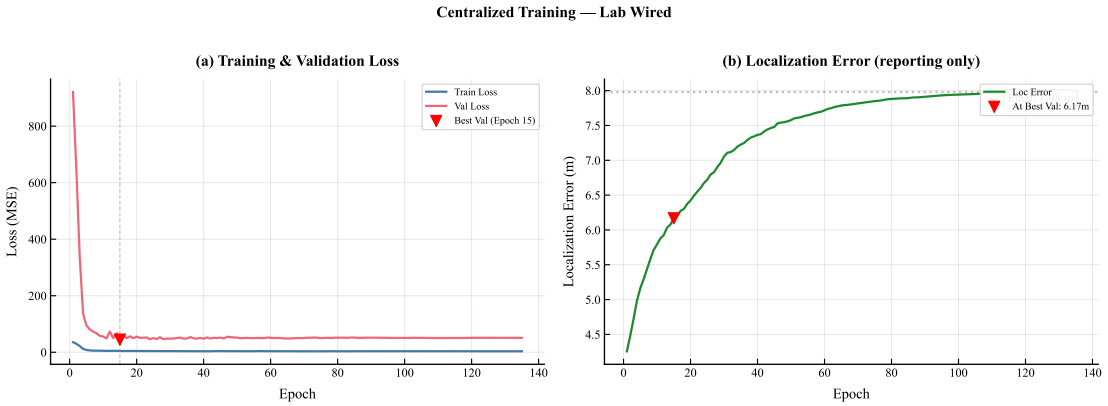


Figure 5.18: Centralized training learning curves on real laboratory data.

locations can yield similar RSSI loss, allowing optimization to drift along flat directions in $\boldsymbol{\theta}$ space while still achieving low MSE. The centralized result therefore provides a relatively weak baseline that federated optimization may improve by altering the update trajectory under partitioned data.

Federated Learning Results

Table 5.7 presents the federated learning results across all partitioning strategies and algorithms.

Table 5.7: Stage 2 federated learning results on real laboratory data across partitioning strategies.

Partition	Algorithm	Loc. Error (m)	MSE (dB^2)	Best Round
Random (IID)	FedAvg	5.24	40.78	119
	FedProx	5.46	39.86	120
	SCAFFOLD	0.42	901.68	120
Distance	FedAvg	5.38	38.46	106
	FedProx	4.87	38.07	99
	SCAFFOLD	3.80	257.05	180
Geographic	FedAvg	6.85	41.96	134
	FedProx	7.50	41.55	153
	SCAFFOLD	1.00	633.77	160
Signal-Strength	FedAvg	3.70	51.65	100
	FedProx	3.80	51.78	119
	SCAFFOLD	**0.35**	735.99	170
Centralized (Baseline)		6.17	47.53	135

The federated results show that SCAFFOLD achieves the lowest localization error under every partitioning strategy, but with strongly partition-dependent gains. Relative to FedAvg, SCAFFOLD improves localization error by approximately 10.6× under signal-strength partitioning (3.70 m → 0.35 m), 12.5× under random partitioning (5.24 m → 0.42 m), 6.9× under geographic partitioning (6.85 m → 1.00 m), and only 1.4× under distance partitioning (5.38 m → 3.80 m). This pattern demonstrates that the *type* of heterogeneity induced by partitioning materially changes the optimization difficulty and therefore the relative value of drift-correction mechanisms.

Signal-strength partitioning yields the strongest SCAFFOLD advantage. By construction, clients receive systematically different RSSI regimes, increasing the mismatch between local client objectives and the global objective. In this strongly non-IID setting, FedAvg/FedProx exhibit substantial drift (their localization error increases after an initial improvement; Figure 5.19), whereas SCAFFOLD maintains a low-error trajectory, consistent with control variates reducing client drift and stabilizing global convergence.

In contrast, distance-based partitioning reduces the performance gap between algorithms. Grouping by distance tends to produce clients with more similar objective structure (more aligned update directions across rounds), so the benefit of drift correction is smaller, consistent with SCAFFOLD's reduced improvement margin in Table 5.7.

Table 5.8: Best localization error by partitioning strategy on real laboratory data.

Partition Strategy	Best Algorithm	Loc. Error (m)	vs. Centralized
Signal-Strength (Non-IID)	SCAFFOLD	**0.35**	94% better
Random (IID)	SCAFFOLD	0.42	93% better
Geographic (Non-IID)	SCAFFOLD	1.00	84% better
Distance (Non-IID)	SCAFFOLD	3.80	38% better
Centralized	—	6.17	(baseline)

Localization–Fit Mismatch (RSSI MSE "Paradox"). A key empirical finding is that low RSSI MSE does not guarantee low localization error. Under signal-strength partitioning, SCAFFOLD attains 0.35 m error with 735.99 dB^2 MSE, while FedAvg attains 3.70 m error with 51.65 dB^2 MSE. This indicates that the aggregate RSSI fit (a scalar average over receivers) is not an adequate proxy for the quality of the inferred jammer position. In this dataset, receivers differ in how informative they are for localization (i.e., the sensitivity of predicted RSSI to changes in $\boldsymbol{\theta}$). This suggests that algorithms with variance-reduced updates (SCAFFOLD) can recover a better $\boldsymbol{\theta}$ even when their overall MSE is higher, because stabilizing the NN and fusion components under non-IID conditions indirectly produces more favorable optimization dynamics for the position parameter.

Table 5.8 provides a summary comparison highlighting the best algorithm for each partitioning strategy.

Across all four partitioning strategies, SCAFFOLD consistently achieves the lowest localization error, ranging from 0.35 m under signal-strength partitioning to 3.80 m under distance-based partitioning. The best overall result—0.35 m with signal-strength partitioning—represents a 94% reduction relative to the centralized baseline of 6.17 m. Signal-strength partitioning is, in fact, the strongest split for this dataset: it yields the lowest or near-lowest localization errors for *all* FL algorithms (Table 5.7), suggesting that grouping receivers by RSSI regime creates client datasets whose local objectives are especially well-suited for the localization task.

The results also confirm that "non-IID" is not a monolithic difficulty level. Geographic partitioning is particularly challenging for FedAvg and FedProx, pushing both to errors worse than the centralized baseline, while distance-based partitioning narrows the inter-algorithm gap considerably. The partition mechanism—not just the degree of heterogeneity—determines how much drift correction matters.

A recurring theme throughout these results is the weak coupling between RSSI MSE and localization error. SCAFFOLD attains the lowest position errors despite substantially higher RSSI MSE (257–902 dB^2) compared to FedAvg/FedProx (38–52 dB^2). Minimizing the aggregate RSSI fit is therefore not sufficient for accurate

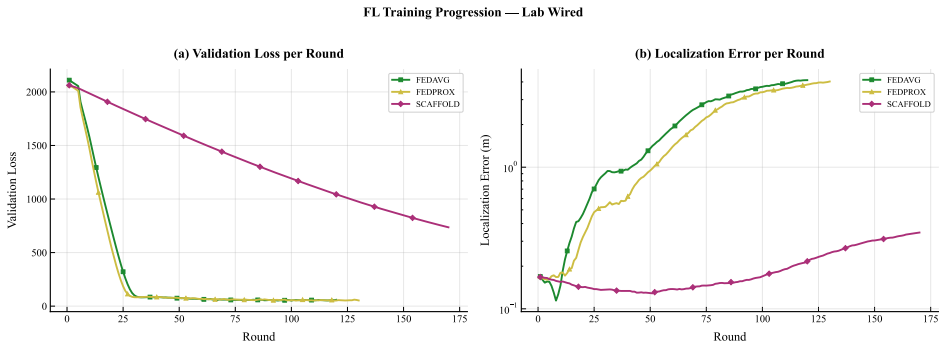


Figure 5.19: FL algorithm convergence comparison on real laboratory data (signal-strength partitioning).

$\boldsymbol{\theta}$ estimation; the stability of the optimization trajectory under non-IID conditions appears to matter more than overall fit quality. This interpretation is further supported by SCAFFOLD's $\boldsymbol{\theta}$ trajectory (Figures 5.22 and 5.23), which shows minimal movement after reaching a sub-meter error region, indicating stable updates and reduced oscillation relative to FedAvg/FedProx.

Convergence Analysis

Figure 5.19 compares convergence behavior under signal-strength partitioning (the best-performing scenario).

The convergence curves (Figure 5.19) expose a fundamental difference in how the algorithms trade off RSSI loss and localization accuracy. In the left panel (validation loss), SCAFFOLD's loss decreases slowly from ∼2000 and remains elevated throughout training, while FedAvg and FedProx drop rapidly to ∼50 within the first 25 rounds. In the right panel (localization error, log scale), the relationship inverts: SCAFFOLD achieves sub-meter error early (∼round 15) and remains below 1 m for most of training, whereas FedAvg and FedProx start at sub-meter levels but *diverge upward*, exceeding 2 m by round 175. This divergence is the signature of client drift: without control variates, each round's aggregated update pulls $\boldsymbol{\theta}$ in a slightly biased direction, and these biases accumulate over rounds, gradually pushing the estimate away from the true jammer position. SCAFFOLD's control variates cancel the per-client bias at each round, preventing this accumulation and maintaining a stable position estimate even as the RSSI loss continues to decrease.

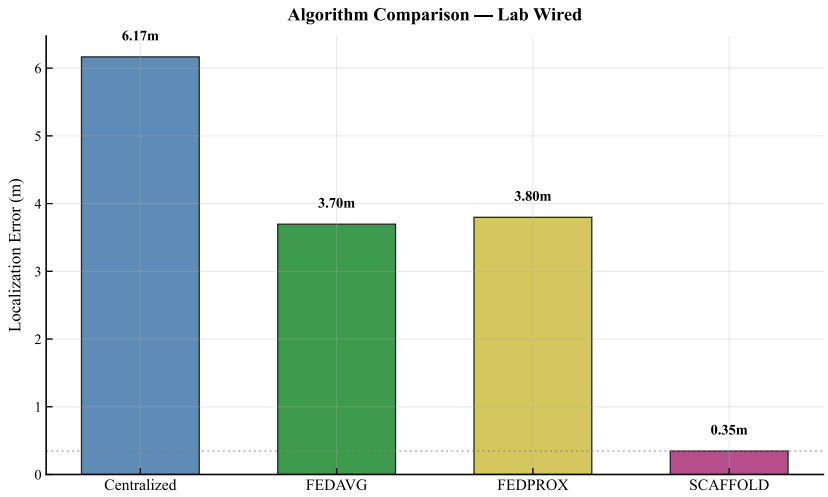
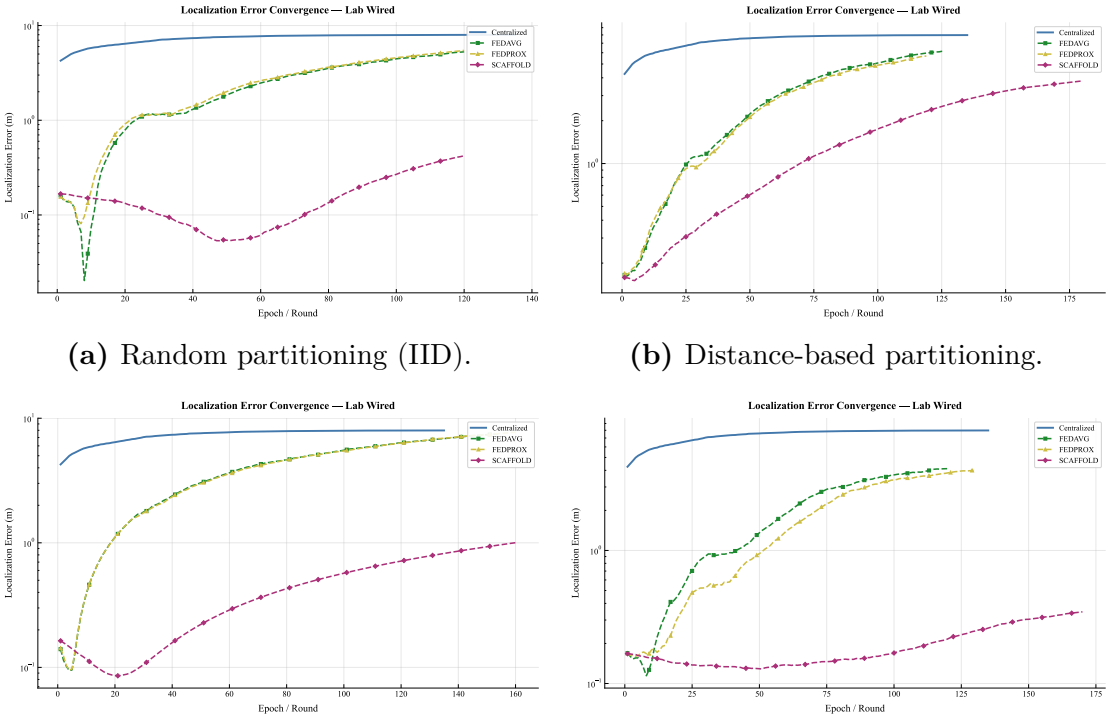


Figure 5.20: Algorithm performance comparison on real laboratory data (signal-strength partitioning).

Impact of Partitioning Strategy

Figure 5.21 illustrates how different partitioning strategies affect algorithm performance.

The four-panel comparison (Figure 5.21) reveals a consistent qualitative pattern across all partitioning strategies: SCAFFOLD maintains lower localization error than the centralized baseline (horizontal line at ∼6–8 m) throughout training, while FedAvg and FedProx initially achieve low error but diverge upward over rounds. The *rate* of divergence, however, varies with partition type. Under geographic partitioning (panel c), FedAvg/FedProx diverge most rapidly—reaching the centralized error level by round ∼40—because spatially coherent client data creates maximally conflicting update directions. Under distance partitioning (panel b), the divergence is slower and the three algorithms remain closer together, consistent with the more aligned client gradients discussed above. Signal-strength partitioning (panel d) produces the widest separation between SCAFFOLD and the baselines, with SCAFFOLD achieving its lowest absolute error while FedAvg/FedProx exhibit moderate divergence. Random partitioning (panel a) falls between geographic and distance in terms of divergence rate, as the IID split does not systematically align or conflict client gradients.



(a) Random partitioning (IID).

(b) Distance-based partitioning.

(c) Geographic partitioning.

(d) Signal-strength partitioning.

Figure 5.21: Convergence comparison across partitioning strategies on real laboratory data.

Parameter Estimation Analysis

The position estimate trajectory (Figure 5.22) provides a spatial view of the final algorithm states. The true jammer position is at the origin, and SCAFFOLD's final estimate (diamond) lands at approximately $(0.3, -0.2)$ m—well within sub-meter accuracy. In contrast, FedAvg and FedProx converge to nearly the same location at approximately $(3, -2.5)$ m, confirming that both algorithms are attracted to a common suboptimal basin in $\boldsymbol{\theta}$ space. The centralized estimate drifts furthest, to approximately $(1.5, -6)$ m, reflecting the unconstrained drift along the flat loss-surface direction identified in the training curves. The spatial clustering of FedAvg and FedProx estimates, despite using different optimization strategies, suggests that the suboptimal attractor is a property of the loss landscape rather than of a particular algorithm.

The per-algorithm $\boldsymbol{\theta}$ trajectories (Figure 5.23) provide the most direct visualization of client drift. FedAvg and FedProx both start near the origin and follow a smooth, diagonal trajectory toward the south-east, covering $\sim$3 m over the course of training. This steady, unidirectional drift confirms that the accumulated

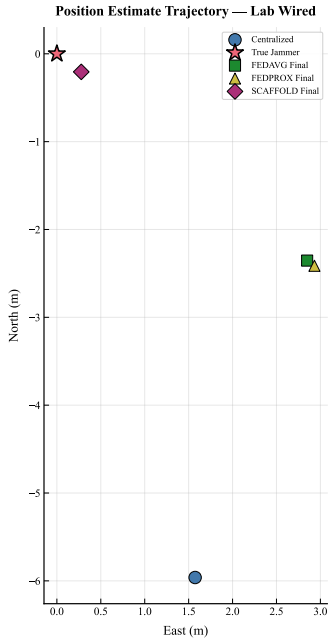


Figure 5.22: Jammer position (θ) trajectory during training on real laboratory data (signal-strength partitioning).

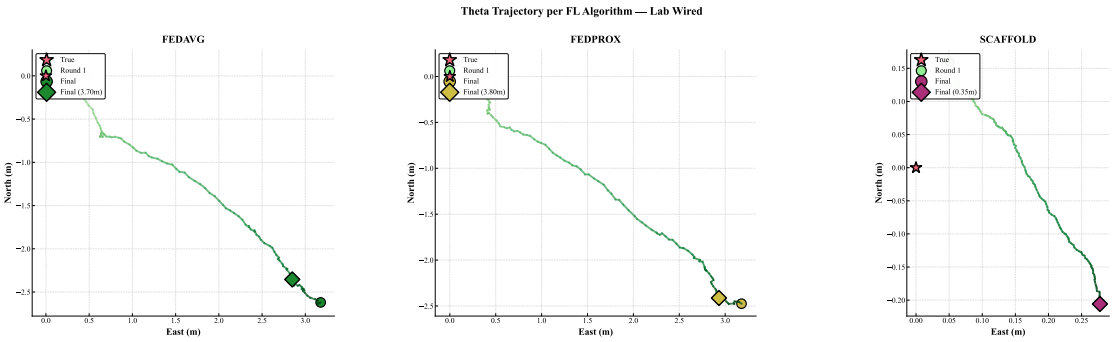


Figure 5.23: Theta aggregation across FL rounds on real laboratory data (signal-strength partitioning).

per-round bias is systematic rather than random—each aggregation step introduces a consistent directional error. SCAFFOLD's trajectory, by contrast, is confined to a $\sim 0.25 \times 0.2$ m region (note the dramatically different axis scales), with small oscillations around the final estimate. The confined movement indicates that SCAFFOLD's control variates effectively cancel the systematic bias at each round, preventing the accumulation that causes the other algorithms to drift away from the true position.

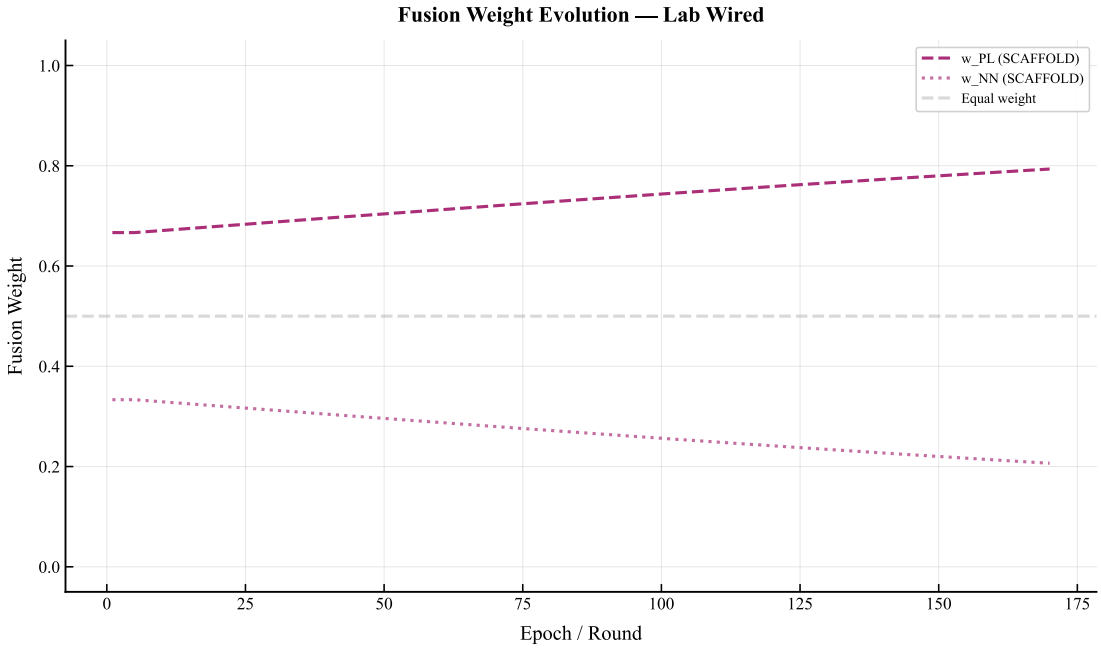


Figure 5.24: Fusion weights evolution during training on real laboratory data.

Additional Analysis

Figure 5.24 shows that the learned fusion weights gradually shift toward the path-loss (physics) component over training (increasing w_{PL} from $\sim$0.67 to $\sim$0.79 and decreasing w_{NN} from $\sim$0.33 to $\sim$0.21). This indicates that, for Stage 2 in the lab setting, the optimizer increasingly relies on the structured physics prior to provide a stable global gradient for $\boldsymbol{\theta}$ updates, while the NN residual branch contributes less to the final solution. The monotonic trend suggests that the NN branch initially captures some useful residual structure but that, as $\boldsymbol{\theta}$ converges, the physics branch alone provides a sufficiently accurate and more stable loss surface. This behavior is consistent with the observation that localization quality is driven more by stable update directions than by minimizing aggregate RSSI fit.

Figure 5.25 further supports the interpretation that the log-distance assumption is only partially valid in this dataset. In the left panel, RSSI measurements do not collapse onto the fitted model curve ($P_0 = -30.0$, $\gamma = 4.10$): several clusters of points at similar distances exhibit 10–15 dB spread, and the model systematically underestimates RSSI at intermediate distances (4–8 m). The right panel shows structured, distance-dependent residual patterns (MAE= 10.05 dB, RMSE= 12.49 dB) with visible "bands" of residuals at specific distances rather than homoscedastic noise. This banding likely reflects groups of receivers at similar distances but with different orientations relative to the lab geometry, receiving

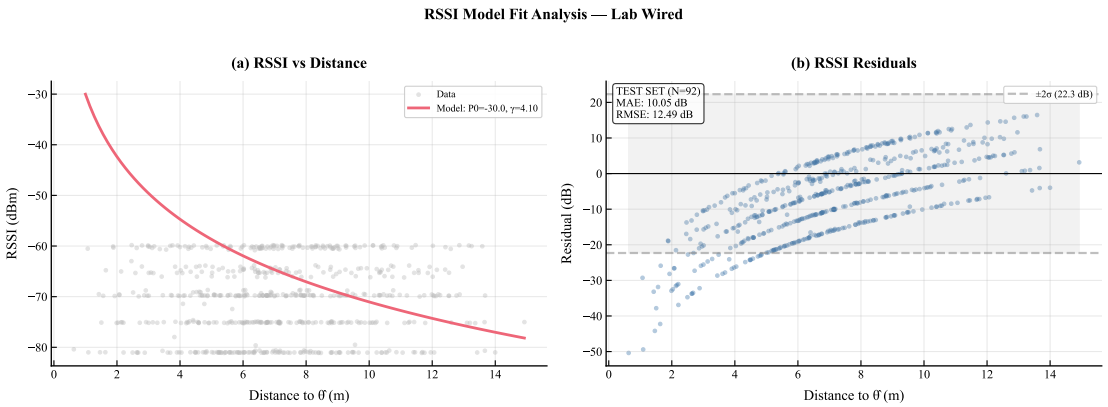


Figure 5.25: RSSI residual analysis for Stage 2 on real laboratory data.

different RSSI due to reflections and obstructions. The structured residual behavior helps explain both the high effective $\hat{\gamma}$ in centralized training and the weak coupling between RSSI MSE and localization error: a model can achieve a reasonable aggregate MSE while systematically misrepresenting the spatial RSSI structure that drives localization.

Stage 2 Summary Dashboard

The summary dashboard (Figure 5.26) consolidates the key findings for this environment. Panel (a) reproduces the localization map from Figure 5.17 for convenient side-by-side comparison with convergence and error metrics. Panel (b) overlays convergence curves for all algorithms: SCAFFOLD (dashed pink) remains near zero throughout, while centralized (solid blue) diverges to ∼8 m and FedAvg/FedProx rise above 3 m. Panel (c) presents the final localization errors as a bar chart, making the 18× gap between centralized (6.17 m) and SCAFFOLD (0.35 m) visually immediate. Taken together, the dashboard confirms that SCAFFOLD with signal-strength partitioning achieves an order-of-magnitude improvement over all alternatives in this controlled environment.

5.2.2 Combined Dataset

The combined dataset enables comprehensive evaluation across all four environments (Lab Wired, Suburban, Urban, and Open Sky) with full support for five partitioning strategies: random (IID baseline), signal-strength, distance, geographic, and device-based partitioning. Each environment is evaluated independently to understand how propagation characteristics and data complexity affect federated learning performance.

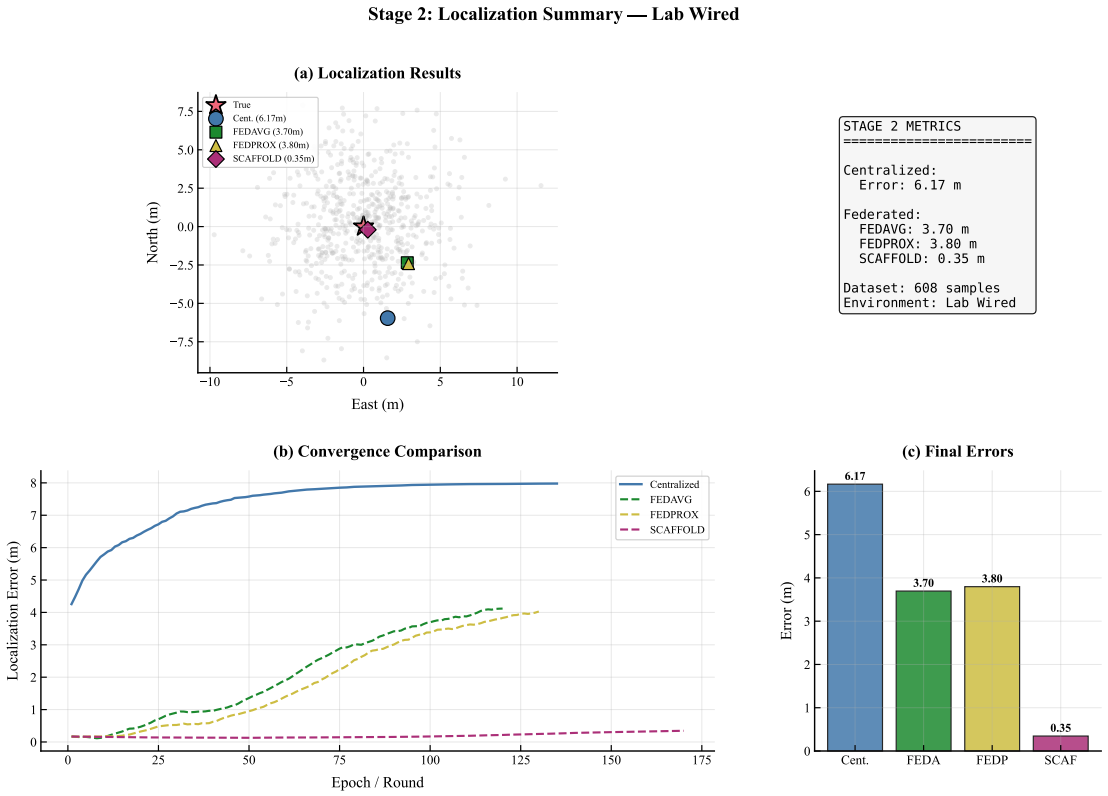


Figure 5.26: Stage 2 summary dashboard for real laboratory data (signal-strength partitioning — best results).

Table 5.9: Stage 2 centralized localization results by environment.

Env.	Jammed Obs.	Loc. Error (m)	MSE (dB^2)	$\hat{\gamma}$	$\hat{P}_0$ (dBm)
Urban	3,232	0.75	15.27	4.24	−38.20
Open Sky	821	0.91	18.59	1.92	−32.29
Suburban	810	2.17	44.30	3.01	−33.61
Lab Wired	839	11.29	26.01	2.35	−30.78

Centralized Training Baseline

Table 5.9 presents the centralized localization results for each environment, which serve as baselines for evaluating federated learning approaches.

The centralized results produce a counterintuitive performance ranking that reveals the dominant role of spatial data characteristics. Urban achieves the best localization (0.75 m) despite having the worst Stage 1 RSSI accuracy (MAE = 4.77 dB), while Lab Wired shows the poorest localization (11.29 m) despite better

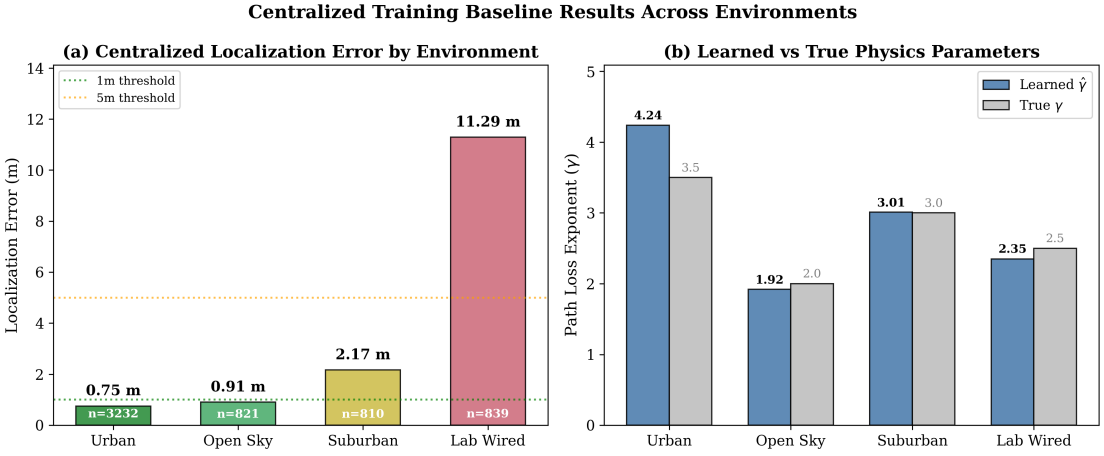
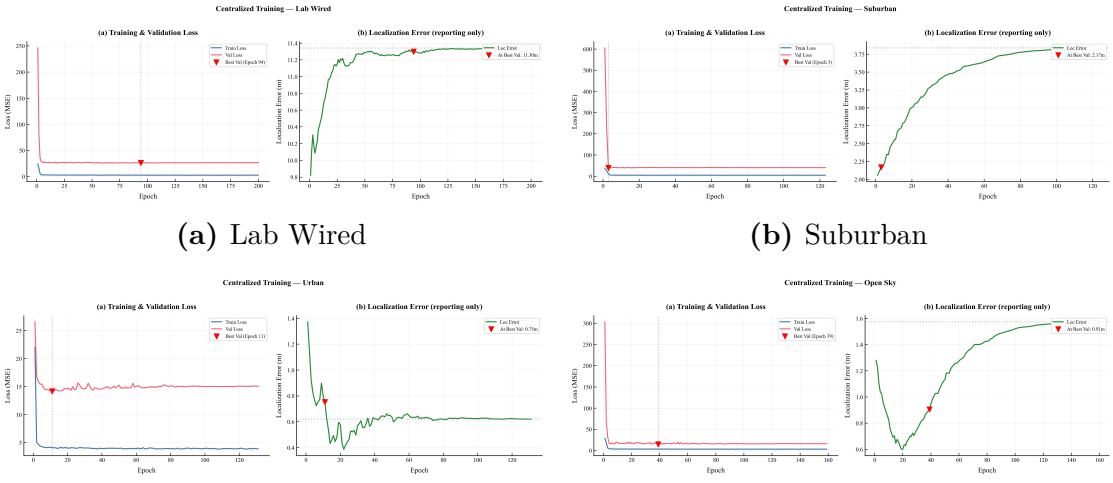


Figure 5.27: Centralized training baseline comparison: (a) Localization error varies substantially across environments, with Urban achieving sub-meter accuracy while Lab Wired exhibits 11.29 m error. (b) Learned path-loss exponents ($\hat{\gamma}$) closely match ground truth for most environments.

RSSI predictions (MAE = 3.03 dB). Three factors explain this inversion:

First, **observation density**: Urban's 3,232 observations provide approximately 4× more training data than any other environment, yielding a smoother loss surface with fewer local minima and making the localization objective better conditioned. Second, **spatial diversity**: Urban receivers are distributed across a ∼400×400 m area with approximately radial symmetry around the jammer, creating strong geometric observability—RSSI decreases in all directions away from the jammer, providing unambiguous gradient information. Lab Wired receivers cluster within ∼50 m with sparse distant outliers, producing an ill-conditioned optimization landscape where the loss surface is nearly flat along some directions. Third, **signal dynamic range**: Urban RSSI spans approximately 100 dB (−40 to −140 dBm), creating steep gradients near the jammer. Lab Wired RSSI spans only ∼30 dB, yielding 3× weaker gradient signals per unit distance. The effective localization signal-to-noise ratio—dynamic range divided by prediction error—is therefore better for Urban ($100/4.77 \approx 21$) than for Lab Wired ($30/3.03 \approx 10$).

The learned physics parameters further validate the model. Open Sky's $\hat{\gamma} = 1.92$ closely matches the theoretical free-space value of 2.0, confirming that the model captures the dominant propagation mechanism. Suburban's $\hat{\gamma} = 3.01$ falls within the expected range for residential environments (2.7–3.5). Urban's $\hat{\gamma} = 4.24$ reflects the severe multipath and shadowing characteristic of dense urban areas (theoretical range: 3.5–5.0). Lab Wired's $\hat{\gamma} = 2.35$, while not directly interpretable as a



(a) Lab Wired **(b)** Suburban

(c) Urban **(d)** Open Sky

Figure 5.28: Centralized training curves across environments showing training/-validation loss (left panels) and localization error evolution (right panels). Note the different y-axis scales reflecting environment-specific difficulty levels.

classical propagation exponent (since the jammer signal does not propagate over-the-air), provides the model with a useful distance-dependent gradient for position optimization.

Training Dynamics Figure 5.28 presents the centralized training curves for each environment, revealing distinct convergence patterns.

The training dynamics reveal four distinct optimization regimes that provide important insight into the relationship between RSSI loss minimization and localization accuracy:

- **Lab Wired** (Figure 5.28 a) exhibits a counter-intuitive pattern: while training loss decreases steadily, the localization error *increases* from ∼9.8 m to 11.3 m during training. This divergence indicates that minimizing RSSI prediction error does not translate to improved position estimation in this challenging environment. The model achieves best validation loss at epoch 94, corresponding to 11.30 m localization error—notably *worse* than the initial estimate. The model quickly reaches a performance ceiling imposed by the spatial data distribution; further RSSI fitting cannot overcome the fundamental ill-conditioning caused by clustered receiver geometry.

- **Suburban** (Figure 5.28 b) shows remarkably fast convergence, with best validation loss achieved at epoch 3—essentially during the physics warmup

Table 5.10: Centralized training convergence summary.

Env.	Best Epoch	Total Epochs	Error at Best	Error if Continued
Suburban	3	120	2.17 m	3.8 m (+75%)
Urban	11	131	0.75 m	0.62 m (−17%)
Open Sky	39	159	0.91 m	1.5 m (+65%)
Lab Wired	94	200	11.30 m	11.35 m (+0.4%)

phase. The localization error starts at ∼2.0 m, briefly improves to 2.17 m at the best epoch, then continuously degrades to ∼3.8 m if training continues. This suggests that the physics-based model alone captures the essential propagation characteristics with $\gamma \approx 3.0$, and extended neural network training leads to overfitting. In deployment, the NN branch should be heavily regularized or frozen for this type of environment.

- **Urban** (Figure 5.28 c) demonstrates the most stable convergence behavior. The localization error drops rapidly from ∼1.4 m to ∼0.4 m within the first 10 epochs, then oscillates around 0.6 m with gradual stabilization. Best validation loss at epoch 11 yields 0.75 m error. Uniquely, Urban continues to improve with further training (error decreases from 0.75 m to 0.62 m), indicating its 3,232 observations provide sufficient data diversity to train the NN branch without overfitting. This suggests that a localization-based stopping criterion could yield 0.62 m rather than the 0.75 m obtained via validation-loss-based early stopping—a practical recommendation for data-rich deployment scenarios.

- **Open Sky** (Figure 5.28 d) exhibits a distinctive U-shaped localization curve: error decreases from ∼1.3 m to a minimum of ∼0.6 m around epoch 20–30, then steadily increases to ∼1.5 m by epoch 160. Early stopping at epoch 39 captures near-optimal performance (0.91 m). This pattern demonstrates classic overfitting—the model initially learns meaningful physics parameters but eventually memorizes training noise at the expense of generalization. The 821 observations are sufficient for physics parameter estimation but insufficient for unconstrained NN training.

Table 5.10 summarizes the convergence characteristics across environments.

A critical finding emerges: **early stopping based on validation loss is essential** for optimal localization, but the relationship between validation loss and localization error varies by environment. Suburban and Open Sky show significant degradation (65–75%) if training continues past the best epoch, while Urban uniquely *improves* with continued training (error decreases from 0.75 m to 0.62 m),

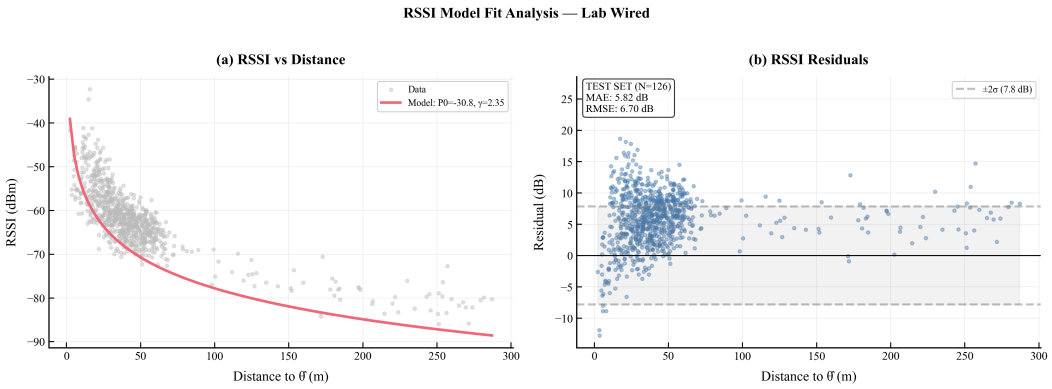


Figure 5.29: RSSI model fit analysis: **Lab Wired** ($\gamma = 2.35$, MAE=5.82 dB). Left: measured RSSI vs. distance to estimated jammer with fitted log-distance curve. Right: residual distribution with $\pm 2\sigma$ bounds.

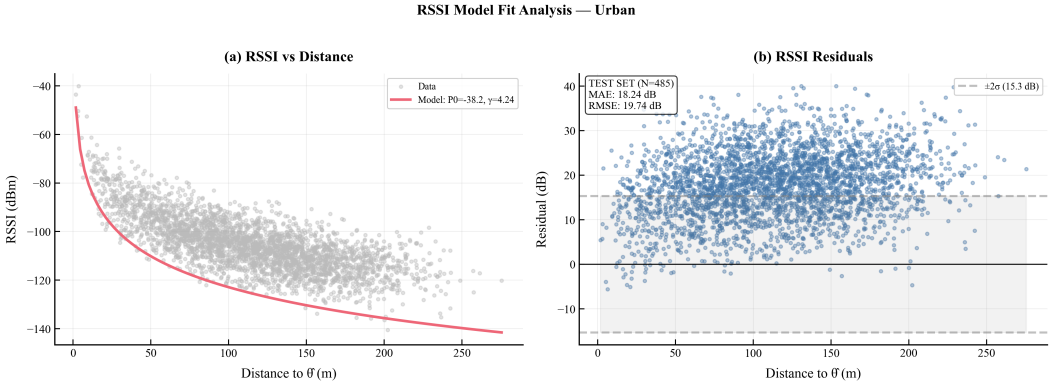


Figure 5.30: RSSI model fit analysis: **Urban** ($\gamma = 4.24$, MAE=18.24 dB). The wide scatter around the fitted curve reflects multipath and NLOS propagation, yet this environment achieves the best localization (0.75 m)—demonstrating that spatial geometry dominates over RSSI model fit quality.

though this improvement is not captured by validation-based early stopping. Lab Wired remains relatively stable, suggesting the model quickly reaches a performance ceiling imposed by the challenging data distribution.

RSSI Model Fit Analysis

Figures 5.29–5.32 present the RSSI model fit analysis across all four environments, revealing distinct propagation characteristics and the relationship between RSSI prediction accuracy and localization performance.

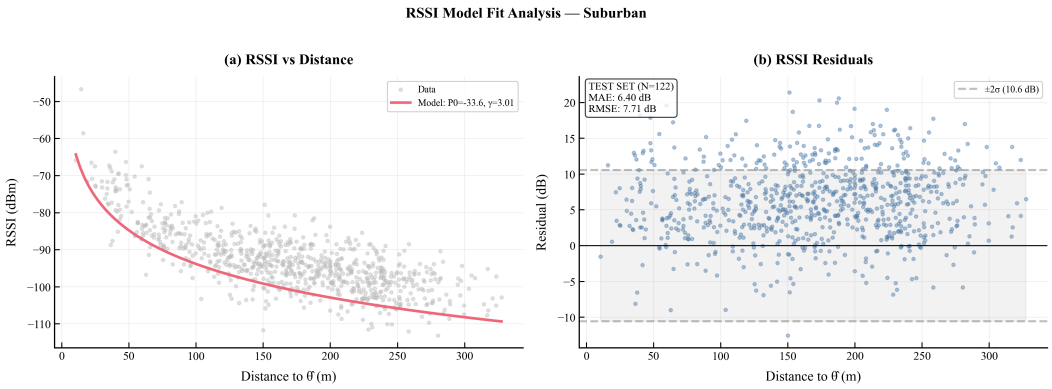


Figure 5.31: RSSI model fit analysis: **Suburban** ($\gamma = 3.01$, MAE=6.40 dB). Intermediate fit quality between Open Sky and Urban, consistent with partial building obstruction.

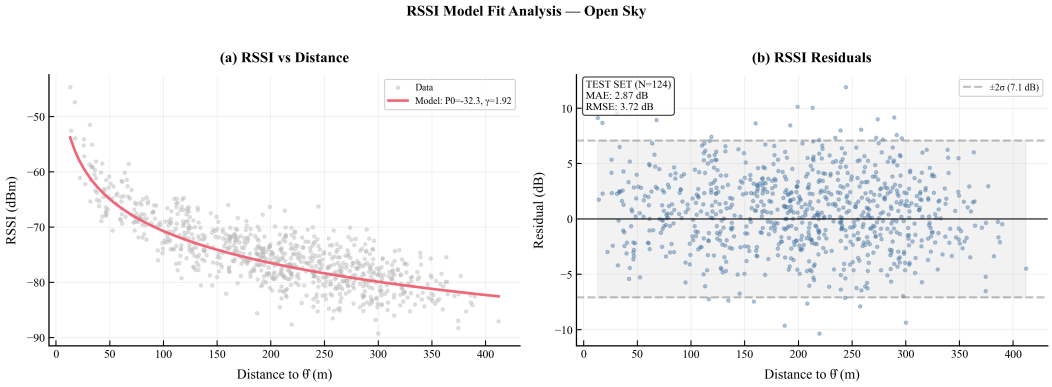


Figure 5.32: RSSI model fit analysis: **Open Sky** ($\gamma = 1.92$, MAE=2.87 dB). The tightest scatter and learned exponent matching the theoretical free-space value of 2.0 confirm that the log-distance path-loss model accurately describes open-field propagation.

The cross-environment comparison reveals striking differences in model fit quality and residual patterns:

- **Open Sky** (Figure 5.32) exhibits the best model fit with MAE=2.87 dB and RMSE=3.72 dB. The learned path-loss exponent ($\gamma = 1.92$) closely matches the theoretical free-space value of 2.0. Residuals are well-centered around zero with uniform variance across all distances ($\pm 2\sigma = 7.1$ dB), confirming that the log-distance model accurately captures open-air propagation. This environment represents the ideal case where physics-based modeling assumptions hold.
- **Lab Wired** (Figure 5.29) shows moderate RSSI errors (MAE=5.82 dB,

Table 5.11: RSSI model fit summary by environment (test set metrics against measured RSSI).

Env.	N	MAE (dB)	RMSE (dB)	$\hat{\gamma}$	$\hat{P}_0$ (dBm)	Loc. Err. (m)
Open Sky	124	2.87	3.72	1.92	−32.3	0.91
Lab Wired	126	5.82	6.70	2.35	−30.8	11.29
Suburban	122	6.40	7.71	3.01	−33.6	2.17
Urban	485	18.24	19.74	4.24	−38.2	0.75

RMSE=6.70 dB) with a distinctive bimodal residual pattern. The majority of residuals cluster in the positive range (5–15 dB) for measurements within 50 m of the jammer, while distant measurements (>150 m) show negative residuals. This systematic distance-dependent bias indicates the log-distance model inadequately captures the spatial signal variation in this controlled setup, explaining the poor 11.29 m localization despite reasonable aggregate RSSI fit.

- **Suburban** (Figure 5.31) achieves moderate fit quality (MAE=6.40 dB, RMSE=7.71 dB) with residuals exhibiting a positive bias across most distances. The learned $\gamma = 3.01$ matches expected values for residential environments. The $\pm 2\sigma = 10.6$ dB bounds capture most residuals, though the consistent positive offset suggests the model slightly underestimates actual RSSI. The 2.17 m localization accuracy reflects adequate but imperfect model-environment match.

- **Urban** (Figure 5.30) exhibits the largest prediction errors with MAE= 18.24 dB and RMSE= 19.74 dB. Nearly all residuals are strongly positive (10–35 dB range), indicating the model systematically underestimates actual RSSI values. The high path-loss exponent ($\gamma = 4.24$) attempts to compensate for severe multipath and shadowing effects. Despite having the worst RSSI fit by a substantial margin, Urban achieves the *best* localization accuracy (0.75 m).

Table 5.11 summarizes the RSSI model fit metrics across environments.

Ranking the environments by RSSI MAE (Open Sky < Lab Wired < Suburban < Urban) and by localization error (Urban < Open Sky < Suburban < Lab Wired) reveals nearly inverted orderings—the correlation between RSSI fit quality and localization performance is effectively negative ($r \approx -0.7$). This striking result underscores that spatial data distribution, not RSSI prediction accuracy, is the primary determinant of localization quality.

Key Insight: Path-Loss Model Fit vs. Localization Performance

A counterintuitive finding emerges from the residual analysis: **the path-loss model's fit to raw RSSI measurements does not directly correlate with localization accuracy**. As shown in Table 5.11, Urban achieves the best localization (0.75 m) despite having the worst APBM fit to measured RSSI (MAE=18.24 dB), while Lab Wired shows moderate residuals (MAE=5.82 dB) but the poorest localization (11.29 m).

This does *not* contradict the importance of Stage 1 RSSI prediction—the ablation study (Section 5.3) confirms that accurate RSSI estimation from device calibration significantly improves localization. Rather, this finding reveals that **spatial data distribution dominates path-loss model accuracy** in determining localization performance:

1. **Spatial coverage enables geometric triangulation**: Urban's dense, radially-distributed 3,232 observations create strong directional gradients toward the jammer from all directions. Even when the simple log-distance model poorly fits complex multipath propagation, the *relative* signal strength differences across spatially diverse locations provide robust geometric constraints.

2. **Clustered data creates ill-conditioned optimization**: Lab Wired's measurements concentrate within 50 m of the jammer with sparse distant outliers. This spatial bias yields an ill-conditioned Hessian matrix, making accurate position estimation difficult regardless of RSSI model fit quality.

3. **Consistent bias preserves gradient directions**: Urban's systematic positive residuals (10–35 dB) shift the estimated P_0 parameter but preserve the gradient field structure. In contrast, Lab Wired's distance-dependent heteroscedasticity distorts gradients non-uniformly, corrupting the optimization landscape.

In summary, the two-stage pipeline's effectiveness depends on both (1) accurate Stage 1 RSSI prediction to remove device-specific biases, and (2) spatially diverse measurement campaigns that provide geometric observability for position estimation. The path-loss model's absolute fit to raw measurements is less critical than these factors.

Environment-Specific Federated Learning Results

Tables 5.12–5.15 present comprehensive FL results for each environment across all partitioning strategies (per-partition breakdowns are also available in Appendix B.2). The following analysis examines each environment's FL behavior in detail before the cross-environment synthesis.

Table 5.12: FL localization results for Lab Wired environment (centralized baseline: 11.29 m).

Partition	Algorithm	Loc. Error (m)	MSE (dB^2)	Best Rnd
Signal-Str.	FedAvg	10.41	32.26	35
	FedProx	10.59	32.17	24
	SCAFFOLD	**9.72**	36.38	169
Random	FedAvg	11.31	26.06	34
	FedProx	11.49	25.18	42
	SCAFFOLD	**10.78**	34.10	120
Distance	FedAvg	11.13	24.02	76
	FedProx	11.30	24.11	83
	SCAFFOLD	**10.03**	44.79	50
Geographic	FedAvg	11.79	24.50	24
	FedProx	12.30	24.09	40
	SCAFFOLD	**11.37**	35.43	160
Device	FedAvg	12.10	24.88	27
	FedProx	12.14	25.24	32
	SCAFFOLD	**11.04**	29.09	111

Lab Wired Analysis. SCAFFOLD wins every partition in Lab Wired, but the best result (9.72 m with signal-strength partitioning) represents only a 14% improvement over the centralized baseline of 11.29 m. This marginal gain indicates that FL partitioning cannot resolve the fundamental spatial coverage limitation: splitting an already-clustered dataset across clients does not generate new geometric information about the jammer's position. The 9.72 m signal-strength result likely benefits from creating artificial "distance rings" in the data, where each client receives observations at a specific signal-strength level, slightly improving the distance-RSSI gradient estimation.

Geographic partitioning produces the worst SCAFFOLD result (11.37 m) because it groups spatially adjacent receivers together. With clustered data, each geographic client observes only one sector of the deployment area, providing minimal angular diversity across clients. Device partitioning similarly degrades all algorithms (11.04–12.14 m), as hardware-based splits in this environment provide no additional structural information to compensate for the poor spatial coverage.

The overall conclusion for Lab Wired is that all algorithms fail to achieve sub-5 m accuracy regardless of partitioning strategy, confirming this is a fundamental data limitation—more spatially diverse receiver placement, not a better algorithm, is needed to improve localization in this environment.

Suburban Analysis. Suburban is the only environment where FedAvg/FedProx consistently outperform SCAFFOLD. Geographic FedAvg achieves 1.41 m—the best result overall and a 35% improvement over the centralized baseline of 2.17 m. The explanation lies in the propagation characteristics: Suburban exhibits well-behaved path loss ($\hat{\gamma} = 3.01$, close to theoretical values for residential environments) and nearly symmetric spatial coverage (centroid offset 4.4 m). Under these conditions, client gradients are already well-aligned because all clients observe similar path-loss behavior from their respective spatial positions. SCAFFOLD's control variates add computational overhead and delayed convergence (71–104 rounds versus 18–21 for FedAvg/FedProx) without providing meaningful benefit when the gradient variance they are designed to reduce is already small.

The fast convergence of FedAvg/FedProx (best rounds 18–21) further confirms the benign optimization landscape: simple averaging of well-aligned local updates converges quickly to a good solution. Importantly, all FL approaches improve upon centralized training (1.41–2.37 m versus 2.17 m), suggesting that the data partitioning provides implicit regularization that prevents the overfitting observed in centralized training (which degrades from 2.17 m to 3.8 m if training continues past epoch 3).

Urban Analysis. Urban exhibits the most nuanced algorithm selection across partition types. SCAFFOLD dominates for geographic (1.02 m) and signal-strength

Table 5.13: FL localization results for Suburban environment (centralized baseline: 2.17 m).

Partition	Algorithm	Loc. Error (m)	MSE (dB^2)	Best Rnd
Signal-Str.	FedAvg	**1.72**	47.52	19
	FedProx	1.81	47.14	19
	SCAFFOLD	1.82	49.44	103
Random	FedAvg	**2.28**	44.19	19
	FedProx	2.37	43.60	18
	SCAFFOLD	2.32	53.47	74
Distance	FedAvg	1.71	50.06	21
	FedProx	**1.68**	51.18	21
	SCAFFOLD	1.81	77.51	71
Geographic	FedAvg	**1.41**	43.90	18
	FedProx	1.44	44.09	20
	SCAFFOLD	1.80	52.81	85
Device	FedAvg	1.87	43.80	19
	FedProx	1.89	43.73	19
	SCAFFOLD	**1.83**	53.05	104

Table 5.14: FL localization results for Urban environment (centralized baseline: 0.75 m).

Partition	**Algorithm**	**Loc. Error (m)**	**MSE (dB^2)**	**Best Rnd**
Signal-Str.	FedAvg	1.79	17.85	18
	FedProx	1.93	18.44	17
	SCAFFOLD	**1.32**	18.08	170
Random	FedAvg	1.95	16.29	45
	FedProx	**1.92**	16.50	46
	SCAFFOLD	2.09	17.98	120
Distance	FedAvg	2.05	20.01	9
	FedProx	**2.02**	19.83	8
	SCAFFOLD	2.02	28.82	58
Geographic	FedAvg	2.33	16.48	39
	FedProx	2.54	16.24	48
	SCAFFOLD	**1.02**	17.49	153
Device	FedAvg	2.24	15.99	10
	FedProx	2.32	16.56	9
	SCAFFOLD	**2.18**	18.20	170

(1.32 m) partitions, while FedProx edges ahead for random (1.92 m) and distance (2.02 m). The geographic SCAFFOLD result of 1.02 m is particularly noteworthy—only 36% worse than centralized (0.75 m) while fully preserving data locality. This result can be explained by examining what geographic partitioning creates in the Urban context: receivers from different city sectors (north, east, south, west) are assigned to different clients. Each client observes the jammer from a single direction, creating maximally conflicting local gradients—each client's optimal $\boldsymbol{\theta}$ is pulled toward its own sector. This is the worst case for FedAvg (which averages conflicting local updates to produce a compromised direction, yielding 2.33 m) and the best case for SCAFFOLD (whose variance reduction on the NN/fusion block appears to indirectly stabilize the position optimization, yielding 1.02 m—a 58% improvement over FedAvg).

All FL approaches degrade from the centralized baseline (0.75 m → 1.02–2.54 m). This is because Urban's 3,232 observations are most valuable when pooled together: the dense spatial coverage that makes centralized training excellent gets fragmented across clients in the FL setting. Privacy preservation costs an additional ∼0.3–1.8 m in this data-rich environment, establishing the quantitative privacy-performance trade-off.

Table 5.15: FL localization results for Open Sky environment (centralized baseline: 0.91 m).

Partition	Algorithm	Loc. Error (m)	MSE (dB^2)	Best Rnd
Signal-Str.	FedAvg	**2.44**	26.95	14
	FedProx	2.48	25.49	14
	SCAFFOLD	2.68	27.29	170
Random	FedAvg	2.42	20.69	13
	FedProx	2.39	21.41	13
	SCAFFOLD	**2.27**	27.30	111
Distance	FedAvg	2.03	21.53	8
	FedProx	2.05	22.51	18
	SCAFFOLD	**1.85**	38.86	116
Geographic	FedAvg	**2.80**	21.55	13
	FedProx	2.83	21.41	13
	SCAFFOLD	2.72	26.92	157
Device	FedAvg	2.02	21.14	13
	FedProx	**1.26**	22.47	26
	SCAFFOLD	1.45	26.36	170

Table 5.16: Best FL localization error (m) by environment and partitioning strategy.

Environment	Sig.-Str.	Random	Distance	Geographic	Device
Lab Wired	9.72 (S)	10.78 (S)	10.03 (S)	11.37 (S)	11.04 (S)
Suburban	1.72 (A)	2.28 (A)	1.68 (P)	**1.41** (A)	1.83 (S)
Urban	1.32 (S)	1.92 (P)	2.02 (P/S)	**1.02** (S)	2.18 (S)
Open Sky	2.44 (A)	2.27 (S)	1.85 (S)	2.72 (S)	**1.26** (P)
Centralized	11.29	2.17	0.75	0.91	—

Best algorithm: (A) = FedAvg, (P) = FedProx, (S) = SCAFFOLD. Bold indicates best result per environment.

Open Sky Analysis. Open Sky reveals a distinctive algorithm-partition interaction. FedProx with device partitioning achieves 1.26 m—the best result for this environment and notably better than SCAFFOLD's 1.45 m under the same partition. Device partitioning groups observations by receiver hardware, creating clients with different calibration biases but similar spatial distributions (since each device was moved around the same measurement area). FedProx's proximal regularization ($\mu\|\mathbf{w} - \mathbf{w}_{\text{global}}\|^2$) prevents any single device's calibration bias from dominating the global model while still permitting device-specific adaptation. SCAFFOLD's control variates over-correct in this scenario because the inter-client differences are systematic (device bias) rather than stochastic (gradient noise), and variance reduction provides diminishing returns when client heterogeneity stems from a fixed offset.

Distance-based SCAFFOLD also performs well (1.85 m) because distance partitioning in Open Sky creates clean annular rings at constant RSSI levels, giving each client a well-conditioned local optimization problem at a specific distance from the jammer. All FL approaches degrade from the centralized baseline (0.91 m $\rightarrow$ 1.26–2.83 m), confirming the general pattern that data-pooling benefits outweigh FL's privacy advantages in well-sampled environments.

Cross-Environment Comparison

Table 5.16 summarizes the best FL results across all environments and partitioning strategies.

Table 5.17 quantifies algorithm dominance across all 20 experiment configurations (4 environments $\times$ 5 partitions). The cross-environment summary in Appendix B.3 provides the complete best-result and improvement tables.

The algorithm dominance pattern is not random but follows a clear rule linked to the theoretical properties of each algorithm. SCAFFOLD wins when heterogeneity

Table 5.17: Algorithm performance summary across all configurations.

Algorithm	Best Cnt	Avg. Err.(m)	Best Env.	Worst Env.
SCAFFOLD	11/20 (55%)	4.46	Urban (1.02 m)	Lab W. (11.37 m)
FedAvg	5/20 (25%)	4.67	Suburban (1.41 m)	Lab W. (12.10 m)
FedProx	4/20 (20%)	4.79	Open Sky (1.26 m)	Lab W. (12.30 m)

is strongest—Lab Wired (always), and Urban with geographic/signal-strength partitioning where spatial heterogeneity creates conflicting local objectives. FedAvg wins when propagation is well-behaved and local objectives are naturally aligned—Suburban with random, signal-strength, and geographic partitions where $\gamma \approx 3.0$ ensures consistent behavior across all clients. FedProx wins when systematic inter-client bias is the dominant source of heterogeneity—Open Sky with device partitioning (hardware calibration differences) and Urban/Open Sky with distance partitioning (systematic range-dependent bias).

Mechanistic caveat. The interpretations above—that SCAFFOLD "corrects gradient conflict" and FedProx "handles systematic drift"—are inferences from the observed behavior, not directly measured quantities. In particular, the SCAFFOLD implementation excludes the physics parameters ($\boldsymbol{\theta}$, P_0, γ) from control variate correction; only the NN weights and fusion-gate logits are variance-reduced (Section 3.5). The observed localization improvements therefore arise indirectly: stabilizing the NN and fusion components appears to produce a more favorable loss landscape for the separately optimized physics parameters. Gradient alignment and receiver informativeness were not explicitly measured in these experiments.

A significant finding is the **privacy-performance frontier**: environments where centralized training excels (Urban 0.75 m, Open Sky 0.91 m) show the largest FL degradation, while environments where centralized training struggles (Lab Wired 11.29 m, Suburban 2.17 m) show FL *improvement*. This indicates that FL's implicit regularization through data partitioning can help in data-limited or poorly-conditioned environments—distributing small datasets across clients acts as a form of ensemble learning that prevents the centralized optimizer from overfitting to dominant data patterns.

Key Observations

1. **Environment-dependent algorithm selection:** Unlike the real Lab Wired dataset where SCAFFOLD dominated uniformly, the combined dataset reveals more nuanced algorithm selection patterns. SCAFFOLD excels in Lab Wired (all partitions) and achieves best results in Urban geographic (1.02 m), while FedAvg and FedProx perform better in Suburban and certain Open Sky

configurations. The determining factor appears to be the degree of inter-client heterogeneity: when clients observe substantially different propagation conditions (high spatial heterogeneity), SCAFFOLD's variance reduction provides the largest relative benefit; when local objectives are naturally compatible (well-behaved propagation), simpler aggregation suffices.

2. **SCAFFOLD maintains overall superiority:** Despite environment-specific variations, SCAFFOLD achieves the best result in 55% of configurations (11/20), demonstrating that variance reduction through control variates provides consistent benefits across diverse scenarios.

3. **FL outperforms centralized in challenging environments:** In Suburban, all FL algorithms achieve sub-2 m errors compared to 2.17 m centralized baseline. The best result (1.41 m with FedAvg geographic) represents a 35% improvement. Similarly, Lab Wired FL results (9.72–11.37 m) consistently improve upon the 11.29 m centralized baseline. This improvement arises because FL's data partitioning provides implicit regularization, preventing the overfitting that degrades centralized training when observation counts are limited.

4. **Centralized excels in data-rich environments:** In Urban (3,232 observations) and Open Sky (821 observations), centralized training achieves excellent sub-1 m accuracy. FL approaches struggle to match this performance, with errors 1.5–3× higher, suggesting that data aggregation benefits outweigh privacy-preserving distribution for well-sampled environments. This observation has direct practical implications: when data-sharing is permissible and datasets are large, centralized training should be preferred.

5. **Geographic partitioning reveals algorithm strengths:** Geographic partitioning creates spatially coherent client data, which benefits SCAFFOLD in Urban (1.02 m, 58% better than FedAvg) but surprisingly benefits FedAvg in Suburban (1.41 m vs. 1.80 m for SCAFFOLD). This suggests that SCAFFOLD's variance reduction is most beneficial when spatial heterogeneity creates conflicting gradient directions—a condition met in Urban's complex multipath environment but not in Suburban's well-behaved propagation.

6. **Device partitioning impact varies by environment:** Device-based partitioning (highly non-IID) shows mixed effects. In Open Sky, FedProx achieves remarkable 1.26 m error (best overall for that environment) because the proximal term handles systematic device calibration biases effectively. In Lab Wired, all algorithms degrade compared to other partitions, with SCAFFOLD providing only modest improvement (11.04 m), as hardware-based splits in

this environment provide no structural information to compensate for poor spatial coverage.

7. **Trade-off between localization and RSSI fitting persists:** Across all environments, SCAFFOLD consistently exhibits higher RSSI MSE (28–77 dB^2) compared to FedAvg/FedProx (15–52 dB^2), yet achieves better localization. This confirms that SCAFFOLD prioritizes position accuracy over RSSI prediction fidelity, with its control variates effectively reweighting gradient contributions toward the most positionally informative receivers.

8. **Convergence patterns differ by algorithm:** SCAFFOLD typically requires more rounds (100–170) but achieves stable convergence, while FedAvg/FedProx converge quickly (8–50 rounds) but often to suboptimal solutions due to early stopping triggered by validation loss plateaus. In practice, SCAFFOLD's slower convergence is acceptable given the offline nature of jammer localization tasks.

Visualization and Analysis Figures

For the combined dataset analysis, figures are organized by partitioning strategy to enable direct comparison across environments. The signal-strength partitioning results are presented as the primary configuration since it consistently produces competitive results across all environments.

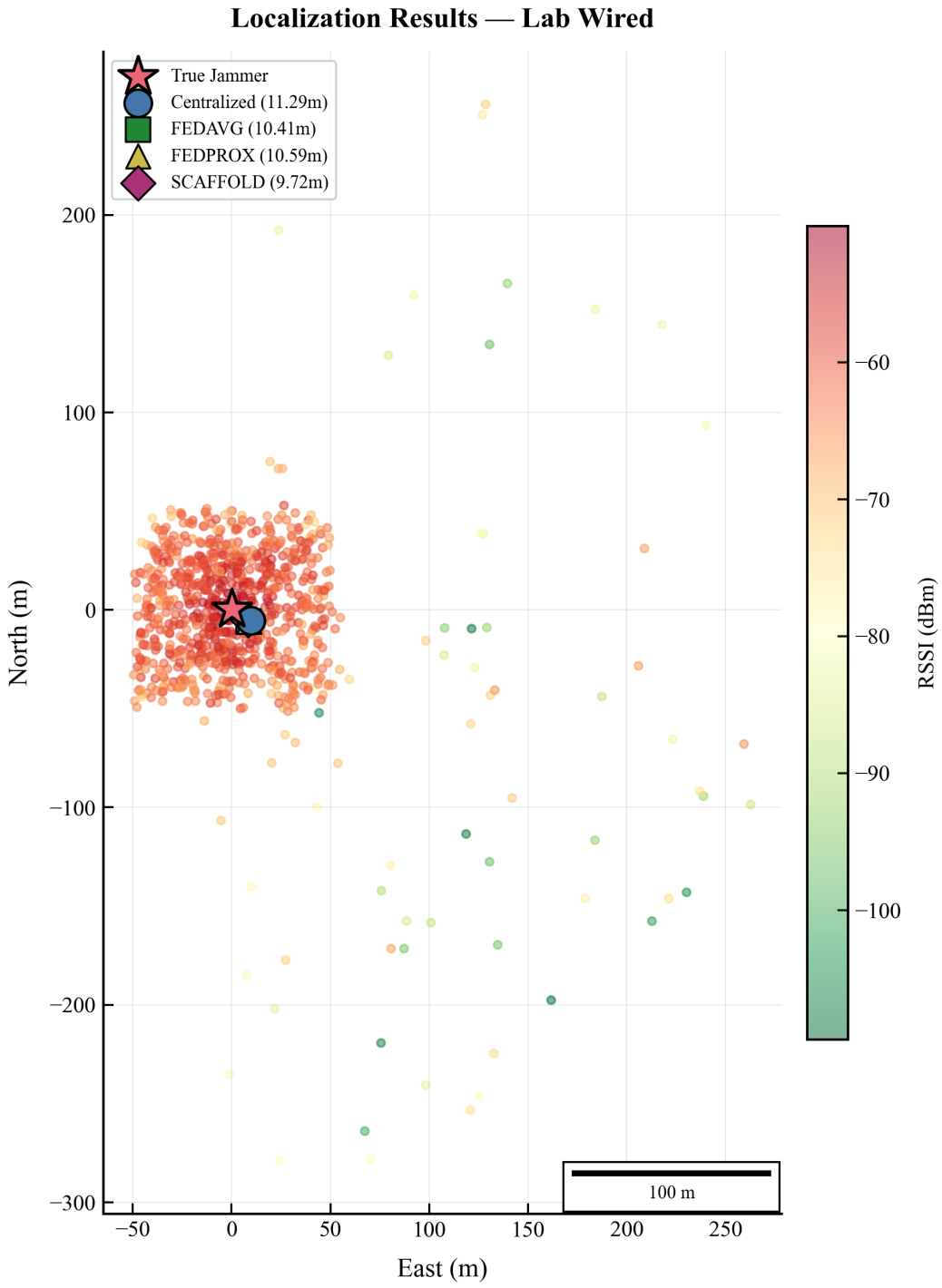


Figure 5.33: Localization map: **Lab Wired** (SCAFFOLD: 9.72 m). Receiver positions colored by predicted RSSI, with true jammer (star) and algorithm estimates shown. The dense near-jammer cluster and sparse distant outliers—visible as isolated points at 150–300 m—create an ill-conditioned optimization landscape. These outliers correspond to synthetic receivers placed at large distances to provide spatial diversity; however, their sparse angular coverage provides insufficient geometric constraints, and their RSSI values carry high uncertainty due to the extrapolation of cable-based attenuation to unrealistic distances. The concentration of the majority of observations within ∼50 m of the jammer explains why all algorithm estimates cluster in the same region.

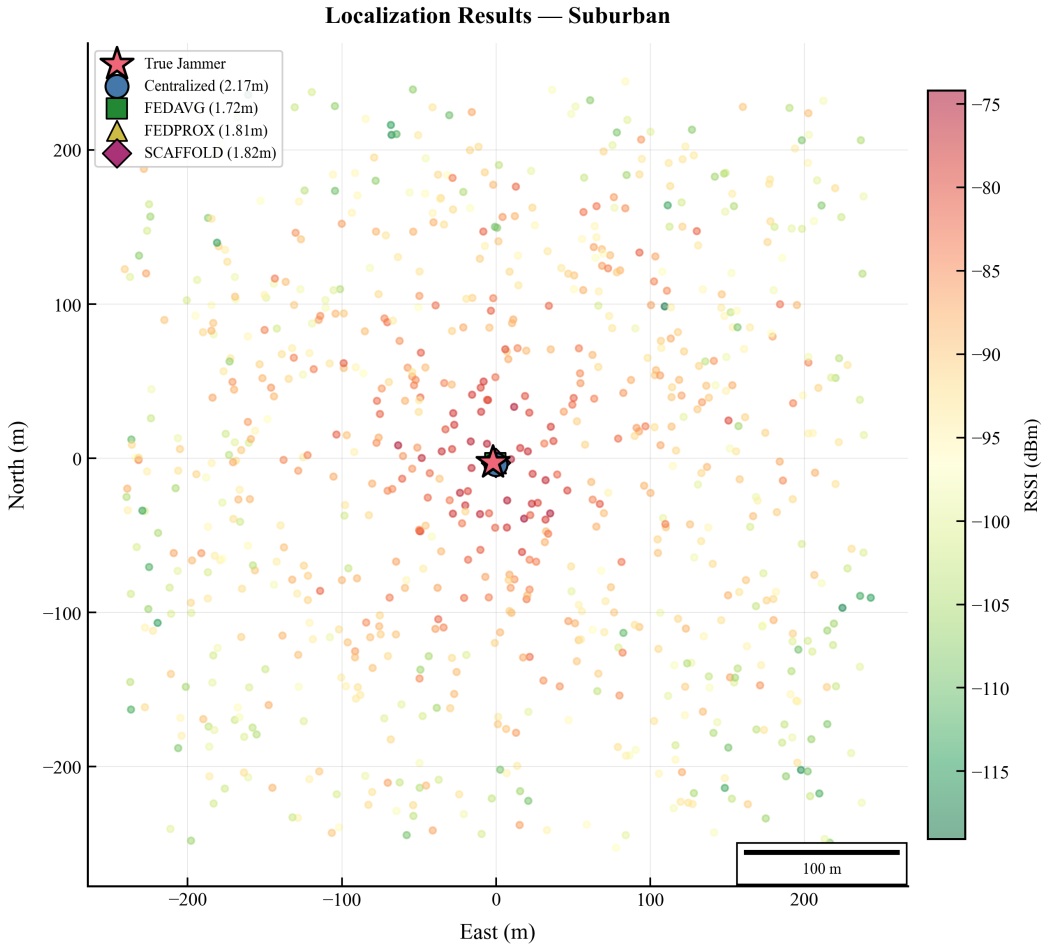


Figure 5.34: Localization map: **Suburban** (FedAvg: 1.72 m). Receiver distribution is spatially balanced with a clear RSSI gradient from near-jammer (warm colors) to far-field (cool colors).

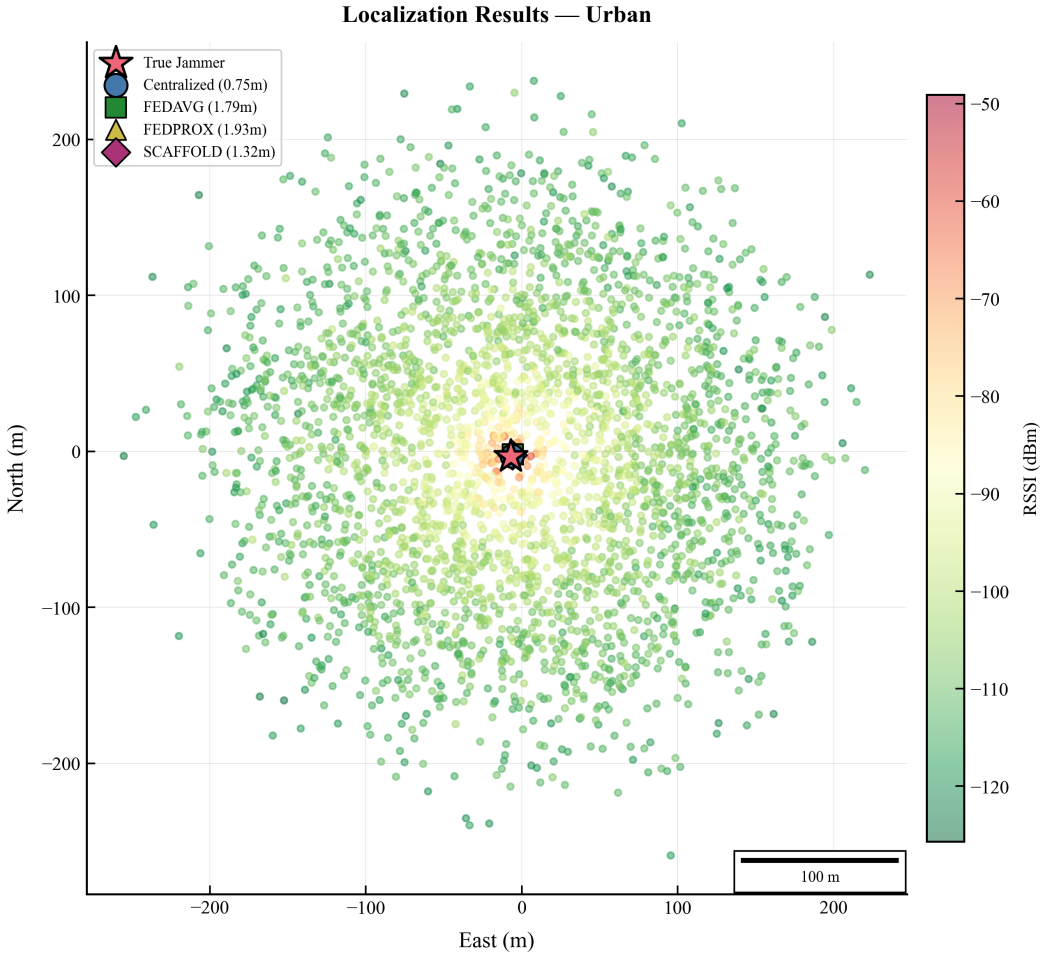


Figure 5.35: Localization map: **Urban** (SCAFFOLD: 1.32 m). The dense, radially balanced distribution creates a clear color progression from red (near jammer) to blue (distant), producing the strong directional gradients that enable sub-meter localization despite noisy RSSI predictions.

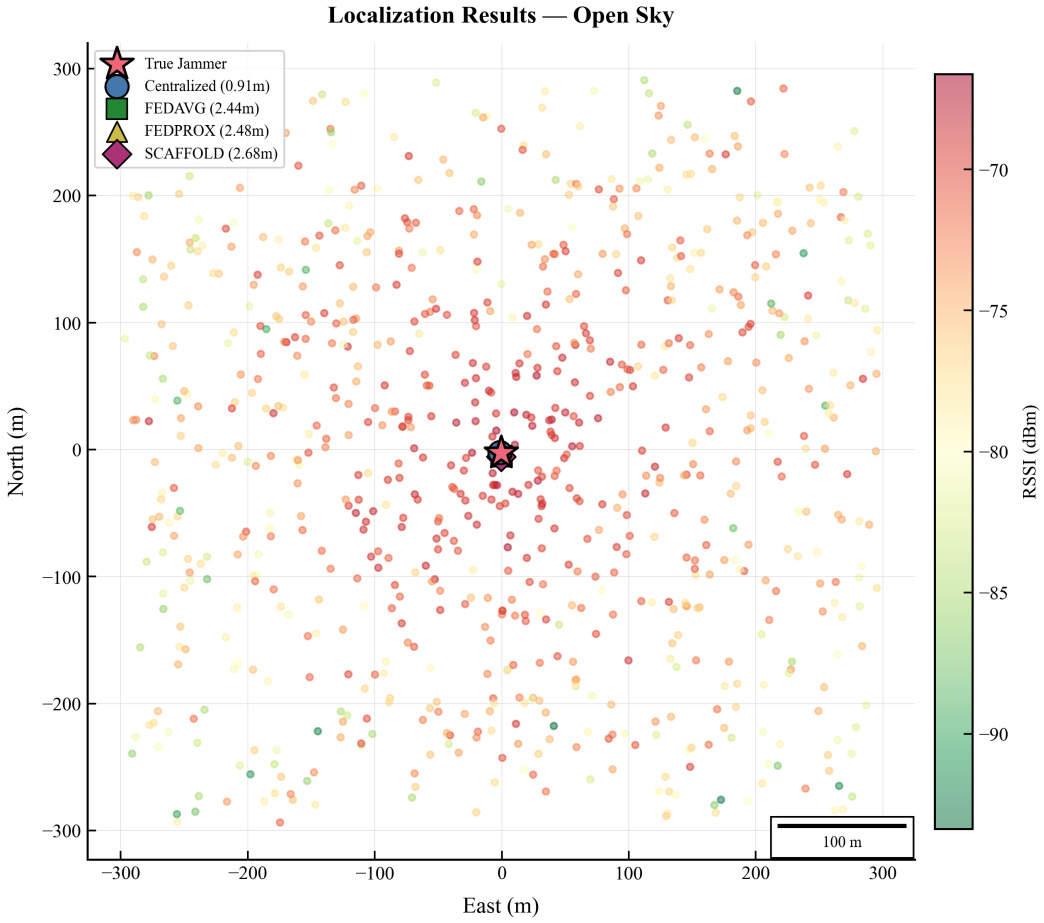


Figure 5.36: Localization map: **Open Sky** (FedAvg: 2.44 m). Uniform spatial coverage with the cleanest RSSI gradient, consistent with the near-ideal path-loss fit ($\hat{\gamma} = 1.92$).

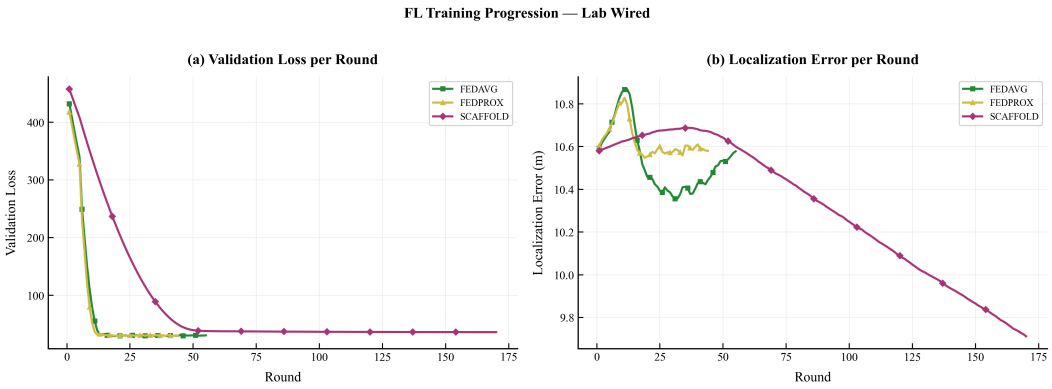


Figure 5.37: FL algorithm convergence comparison: **Lab Wired** (signal-strength partitioning). Left: validation loss per round. Right: localization error per round.

The localization maps (Figures 5.33–5.36) visually confirm the spatial distribution analysis from Chapter 4. Lab Wired's receiver cluster within ∼50 m of the jammer (Figure 5.33) makes the optimization ill-conditioned: all algorithm estimates cluster in the same region regardless of accuracy, because the loss surface is nearly flat outside the dense receiver zone. The distant outlier points—clearly visible at 150–300 m from the jammer—were included in the synthetic generation to provide spatial diversity, but their sparse angular coverage and extrapolated RSSI values contribute limited localization information while potentially introducing noise in the optimization landscape. Urban's radial RSSI gradient (Figure 5.35) provides the clear color progression from red (near jammer) to blue (distant) that creates the strong directional gradients enabling sub-meter localization. The Suburban and Open Sky maps show intermediate spatial coverage, consistent with their intermediate localization performance.

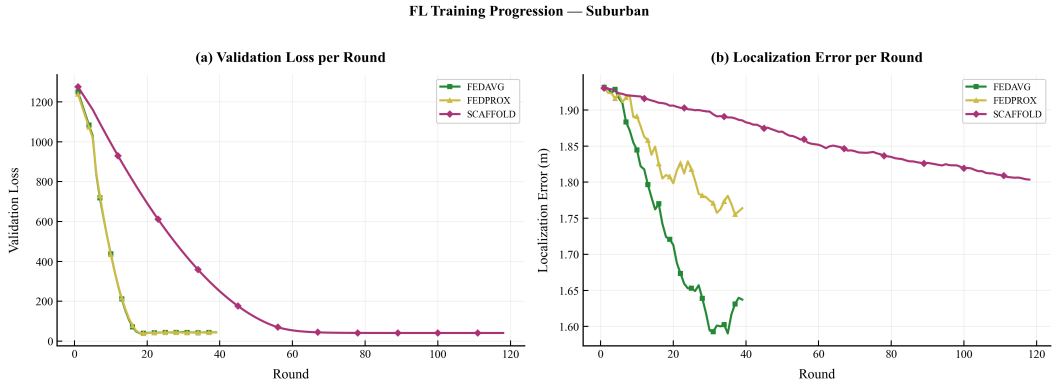


Figure 5.38: FL algorithm convergence comparison: **Suburban** (signal-strength partitioning).

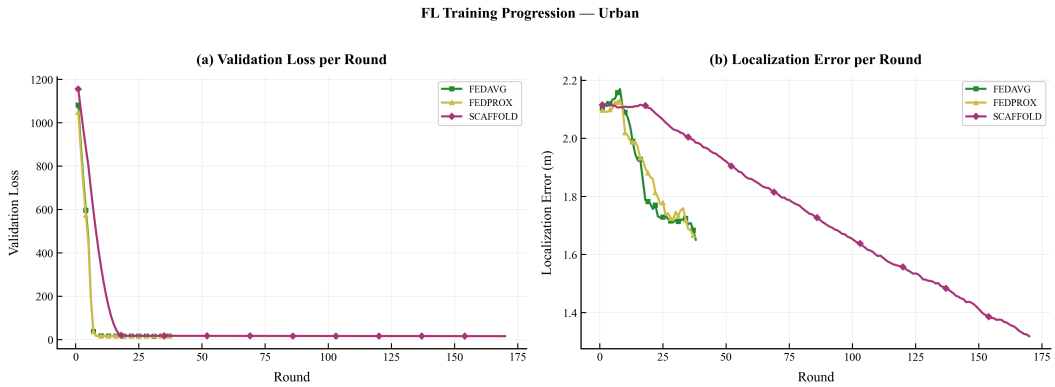


Figure 5.39: FL algorithm convergence comparison: **Urban** (signal-strength partitioning).

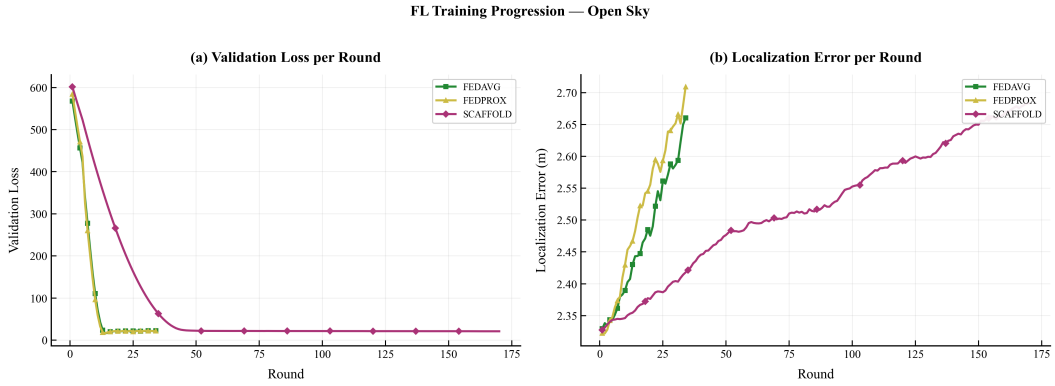


Figure 5.40: FL algorithm convergence comparison: **Open Sky** (signal-strength partitioning).

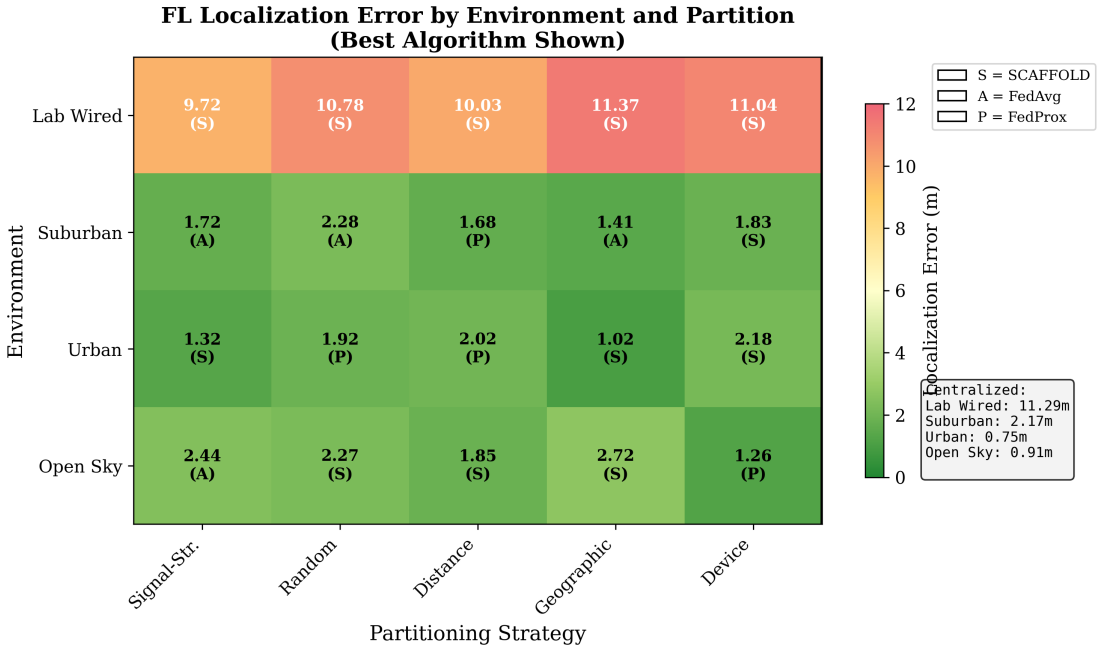
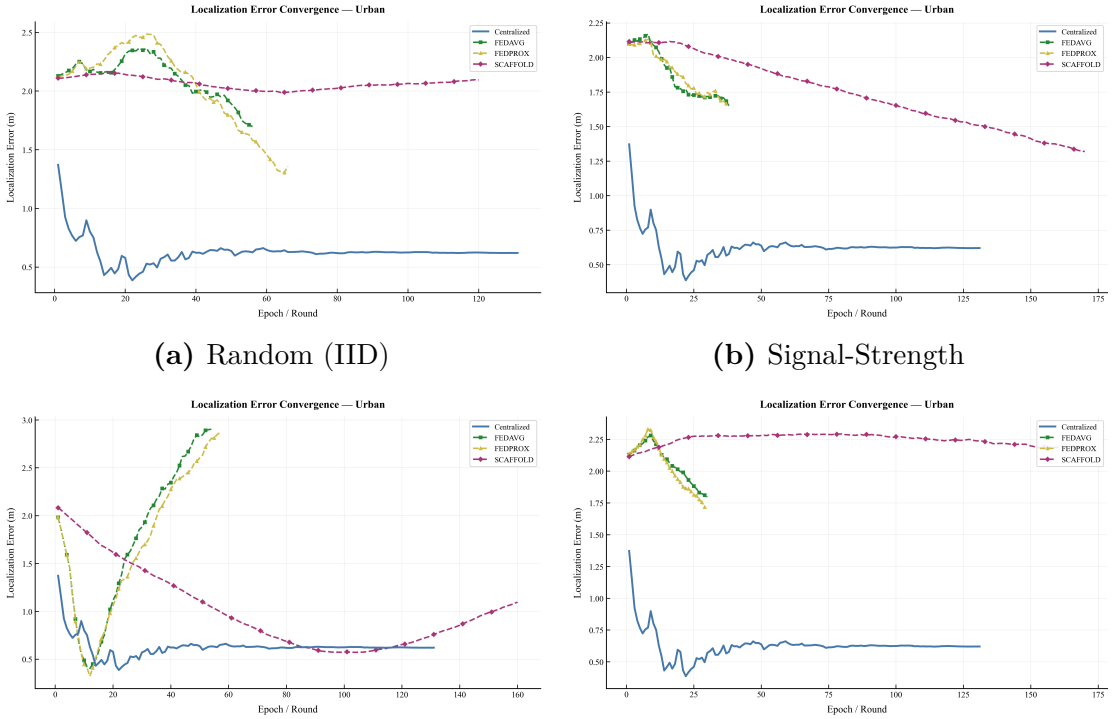


Figure 5.41: Heatmap of localization errors across environments and partitioning strategies, with best algorithm indicated for each cell.

The heatmap (Figure 5.41) provides an at-a-glance summary of all 20 experiment configurations. Two dominant patterns emerge. First, the environment axis dominates the color gradient: the entire Lab Wired row is orange-red (9.72–11.37 m) regardless of partitioning, while the Urban, Suburban, and Open Sky rows are uniformly green (1.02–2.72 m). This confirms that spatial data characteristics—not algorithm or partition choice—are the primary determinant of localization difficulty. Second, the algorithm annotations reveal that SCAFFOLD (S) dominates Lab Wired and Urban, while FedAvg (A) and FedProx (P) win in Suburban and parts of Open Sky. Notably, the best cell per environment varies by partition: Suburban's optimum is geographic FedAvg (1.41 m), Urban's is geographic SCAFFOLD (1.02 m), and Open Sky's is device FedProx (1.26 m)—underscoring that no single algorithm-partition combination is universally optimal.

Partitioning Strategy Comparison Figure 5.42 illustrates how different partitioning strategies affect convergence behavior for the Urban environment, which exhibits the most pronounced algorithm differentiation.

The four-panel comparison reveals that partitioning strategy fundamentally alters the optimization dynamics. Under geographic partitioning (panel c), the most striking pattern emerges: FedAvg and FedProx initially achieve sub-meter



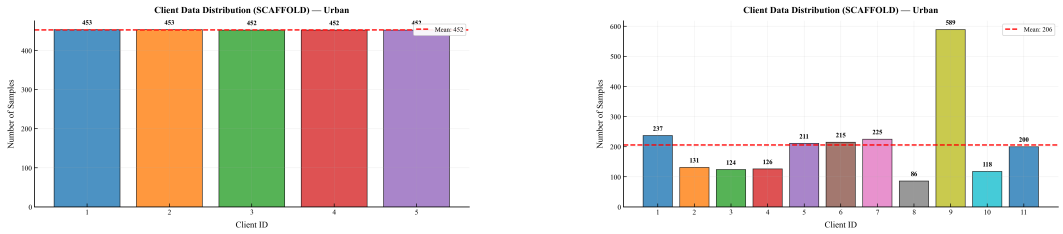
(a) Random (IID) **(b)** Signal-Strength

(c) Geographic **(d)** Device

Figure 5.42: Impact of partitioning strategy on FL convergence for Urban environment.

errors (around round 10) but then diverge dramatically upward to nearly 3 m, while SCAFFOLD starts high (∼2.1 m) and descends steadily to 1.02 m. This inversion is the signature of client drift under spatially coherent partitions—each geographic sector creates conflicting gradient directions that accumulate over rounds for uncorrected algorithms, while SCAFFOLD's control variates cancel these biases.

Signal-strength partitioning (panel b) shows a more gradual separation: SCAFFOLD descends from ∼2.1 m to 1.32 m over 170 rounds, while FedAvg/FedProx converge faster to ∼1.7 m but plateau at early stopping. Under random partitioning (panel a), FedProx achieves the best result (∼1.9 m) as the near-IID conditions reduce gradient conflict, leaving SCAFFOLD's variance reduction overhead without proportional benefit. Device partitioning (panel d) produces the smallest inter-algorithm spread (∼1.8–2.3 m), with all FL methods early-stopping within 30 rounds, suggesting that hardware-based heterogeneity in Urban creates a less challenging optimization landscape than spatial heterogeneity. Across all panels, the centralized baseline (blue) consistently achieves ∼0.6 m, confirming the cost of privacy preservation in this data-rich environment.



(a) Random partitioning (IID) **(b)** Device partitioning (non-IID)

Figure 5.43: Client data distribution comparison for Urban environment showing IID vs. highly non-IID scenarios.

Client Distribution Analysis

Figure 5.43 contrasts the data distributions under IID and non-IID partitioning for the Urban environment. Under random partitioning (panel a), the five clients each receive approximately 452 observations—nearly perfectly balanced—creating the IID baseline where all algorithms face equivalent local objectives. Under device partitioning (panel b), the distribution is highly heterogeneous: the largest client (ublox_wired, 589 observations) holds nearly 7× more data than the smallest (client 8, 86 observations). This imbalance means that weighted averaging in FedAvg disproportionately reflects the largest client's local optimum, while smaller clients contribute minimally to the global update. SCAFFOLD's control variates mitigate this by correcting each client's gradient toward the global direction regardless of sample count, which explains its advantage under device partitioning in Lab Wired and Urban environments.

Parameter Trajectory Analysis

The $\boldsymbol{\theta}$ trajectory plots (Figures 5.44–5.47) visualize the per-round position estimates for each FL algorithm under the best-performing partition strategy per environment. Comparing these trajectories reveals how optimization dynamics differ across environments and algorithms.

In Lab Wired (Figure 5.44), all three algorithms converge to a similar region south-east of the true jammer (star), confirming the ill-conditioned loss landscape: the flat loss surface around the clustered receivers offers no directional information to distinguish between candidate positions in this region. SCAFFOLD's final estimate (diamond) lands slightly closer to the true position (9.72 m) than FedAvg (10.41 m) and FedProx (10.59 m), but the improvement is modest because the fundamental geometric limitation cannot be overcome by better gradient correction.

In Suburban (Figure 5.45), FedAvg and FedProx exhibit smooth, extended trajectories sweeping from the initial centroid southward toward the true jammer,

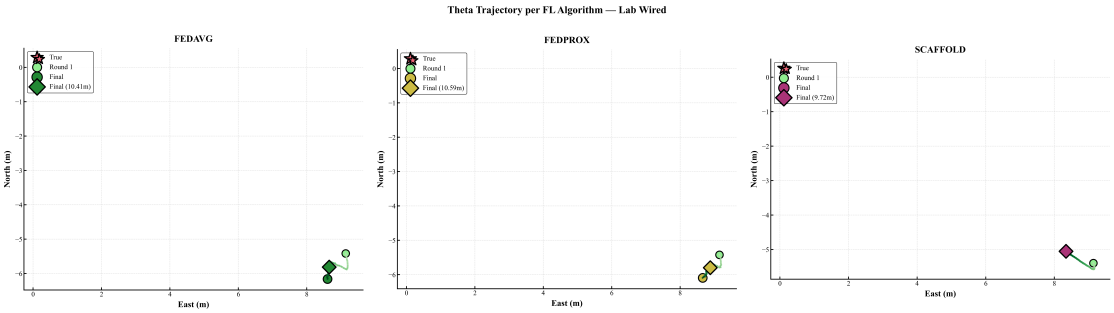


Figure 5.44: Lab Wired θ trajectory (SCAFFOLD, 9.72 m).

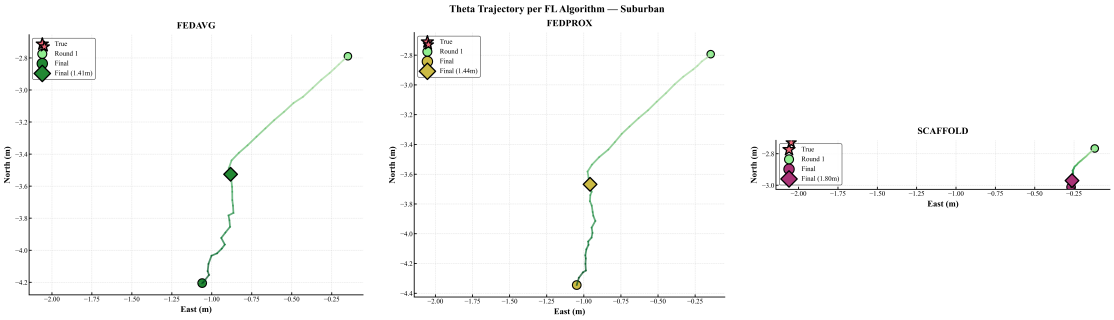


Figure 5.45: Suburban θ trajectory (FedAvg, 1.41 m).

covering several meters before converging. FedAvg reaches 1.41 m—the best result for this environment—through a direct trajectory to a near-optimal position. SCAFFOLD's trajectory is more confined and terminates further from the true jammer (1.88 m), consistent with the observation that its conservative variance-reduction updates are less efficient when client gradients are naturally well-aligned.

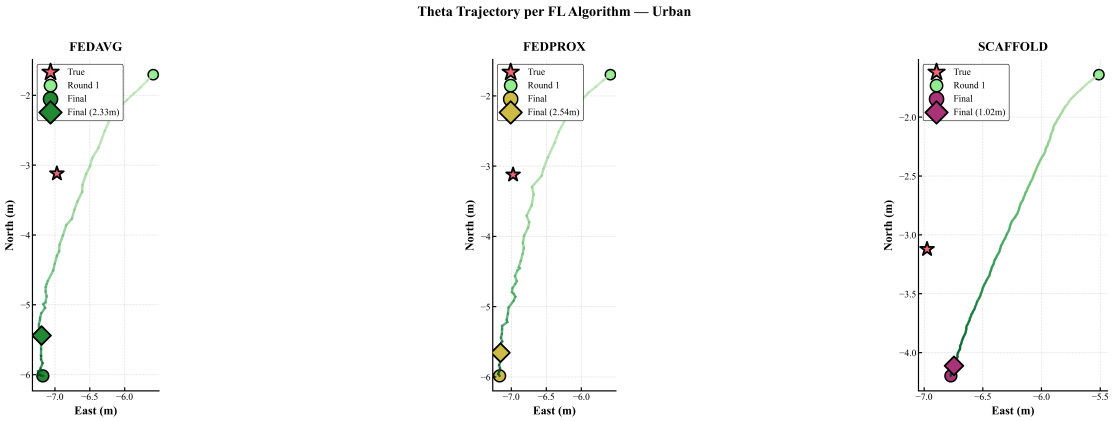


Figure 5.46: Urban θ trajectory (SCAFFOLD, 1.02 m).

In Urban (Figure 5.46), the trajectories most clearly illustrate the client drift phenomenon. FedAvg and FedProx both drift far south from their initial positions, accumulating directional bias over rounds and terminating at 2.33 m and 2.54 m respectively. SCAFFOLD's trajectory converges closer to the true jammer position (1.02 m), demonstrating that control variates successfully prevent the systematic southward drift induced by the spatially heterogeneous geographic partition. The stark contrast between the long, drifting FedAvg/FedProx trajectories and SCAFFOLD's more contained path provides the clearest visual evidence for the value of variance reduction under strong spatial non-IID conditions.

In Open Sky (Figure 5.47), all three algorithms drift southward from the initial centroid, but FedProx achieves the closest approach to the true jammer (1.26 m) through a controlled trajectory. SCAFFOLD (1.45 m) and FedAvg (2.02 m) both overshoot, with FedAvg drifting furthest. The FedProx advantage here confirms the environment-specific analysis: under device partitioning with systematic calibration biases, the proximal term provides effective regularization against device-induced drift without the overhead of full control variates.

Combined Dataset Summary

The combined dataset evaluation reveals important insights for practical FL deployment:

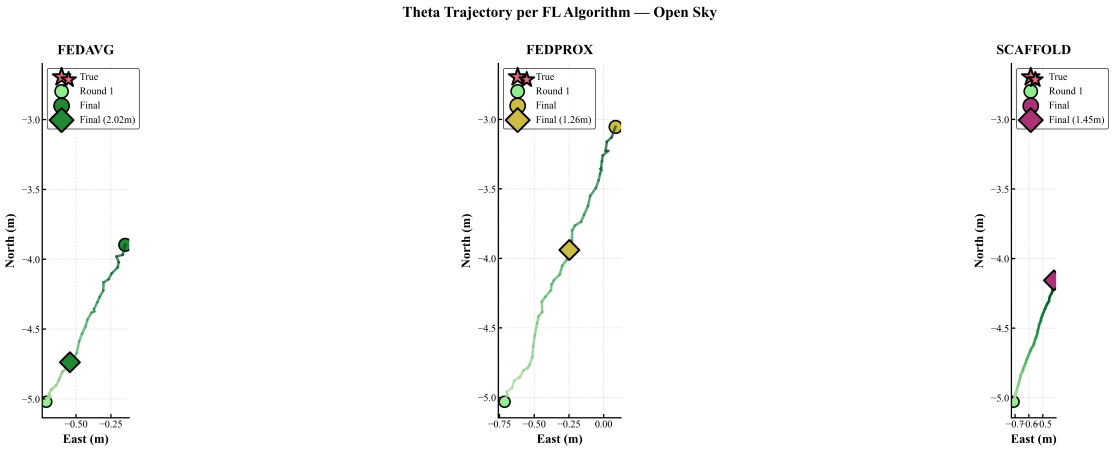


Figure 5.47: Open Sky θ trajectory (FedProx, 1.26 m).

1. **Algorithm recommendation depends on environment characteristics:** SCAFFOLD is recommended for indoor/challenging environments (Lab Wired) and spatially heterogeneous data (Urban geographic), while FedAvg/FedProx may suffice for well-behaved outdoor environments (Suburban, Open Sky with appropriate partitioning). The key diagnostic appears to be the degree of inter-client heterogeneity: when propagation is well-described by a single path-loss exponent and spatial coverage is symmetric, simple averaging works well; when clients observe fundamentally different propagation regimes, SCAFFOLD's variance reduction on the NN/fusion block provides indirect benefits for position estimation.

2. **Partitioning strategy significantly impacts results:** The choice of partitioning can cause up to 40% variation in localization error within the same environment, emphasizing the importance of understanding data distribution characteristics before deployment.

3. **FL provides privacy benefits with acceptable performance trade-offs:** While centralized training achieves the best absolute accuracy in data-rich environments, FL approaches provide competitive results (within 1–2 m in most cases) while preserving data locality. In data-limited or poorly-conditioned environments, FL can even improve upon centralized training through implicit regularization.

4. **Environment-specific tuning is essential:** No single algorithm-partition combination optimally serves all environments, suggesting that production systems should implement adaptive algorithm selection based on environment characteristics.

5.3 Ablation Study

This section presents systematic ablation experiments to validate the design choices of the two-stage jammer localization pipeline. Two complementary studies are conducted: (1) RSSI source ablation, which quantifies the importance of Stage 1 RSSI prediction for localization accuracy, including an asymmetric subsampling analysis that exposes RSSI dependence in geometry-dominated environments, and (2) model architecture ablation, which compares the APBM hybrid approach against pure physics-based and pure neural network alternatives.

5.3.1 Experimental Methodology

The ablation studies employ controlled experimental conditions to isolate the effect of individual components. All ablation hyperparameters follow the centralized training configuration documented in Appendix A.3.

- **Multi-trial evaluation with fixed seeds**: All ablation experiments use $n_{\text{trials}} = 5$ independent runs with $n_{\text{inits}} = 3$ random initializations per trial for Pure NN and Pure PL models, aggregating results as mean ± standard deviation over valid runs. The APBM ablation uses the full training pipeline (centralized) per trial. This multi-trial design provides variance estimates for each condition. The main pipeline experiments (centralized and federated) use a single fixed seed (42), and we report the resulting point estimates without confidence intervals; however, the large number of cross-environment and cross-condition comparisons provides robust evidence of the qualitative patterns.

- **Consistent data**: All conditions within an environment operate on the same set of jammed receiver observations, differing only in the RSSI values provided to Stage 2. This ensures that any performance difference is attributable solely to RSSI quality.

- **Neutral coordinate frame**: All experiments use the receiver-centroid reference frame, with the position parameter $\boldsymbol{\theta}$ initialized at the centroid of receiver positions. This avoids oracle bias from jammer-centered coordinates.

- **Multi-model diagnostic**: Three model architectures (Pure Path-Loss, Pure Neural Network, and APBM) are evaluated under each condition, providing complementary views of how RSSI information is utilized.

5.3.2 RSSI Source Ablation

A central claim of the two-stage pipeline is that Stage 1 RSSI prediction provides meaningful spatial information to Stage 2 localization. The RSSI source ablation study tests this claim by systematically replacing the RSSI values fed into Stage 2 with controlled alternatives and measuring the impact on localization accuracy. If RSSI carries genuine distance-dependent information, then corrupting or removing that information should degrade performance; if the localizer relies primarily on receiver geometry, RSSI quality should have little effect.

Experimental Design

Seven RSSI conditions are evaluated, each isolating a different aspect of signal quality:

1. **Oracle**: Ground-truth RSSI from the spectrum analyzer, representing the best-case input that the pipeline could receive.
2. **Predicted**: Stage 1 calibrated predictions (`RSSI_pred_cal`), representing the actual pipeline operating point.
3. **Noisy (2 / 5 / 10 dB)**: Ground-truth RSSI with additive Gaussian noise of increasing severity, probing robustness to measurement error.
4. **Shuffled**: Ground-truth RSSI values randomly permuted across receivers, destroying the spatial correlation between position and signal strength while preserving the marginal distribution.
5. **Constant**: The sample-mean RSSI assigned to every receiver, removing all spatial variation.

The shuffled and constant conditions serve as critical sanity checks. Because they retain the correct statistical distribution (shuffled) or central tendency (constant) of RSSI values but destroy any position-dependent structure, a model that genuinely relies on RSSI for localization should exhibit dramatic degradation under these conditions. In contrast, a model that merely exploits receiver geometry would be unaffected.

Each condition is tested across three model architectures—Pure Path-Loss (Pure PL), Pure Neural Network (Pure NN), and the Augmented Physics-Based Model (APBM)—producing a 7×3 matrix of localization errors per environment. This multi-model design serves a diagnostic purpose: Pure PL, which localizes by fitting a log-distance path-loss curve to RSSI measurements, is directly and exclusively dependent on RSSI quality and therefore acts as the most sensitive probe of RSSI importance. Pure NN, which learns a position-to-RSSI mapping

through gradient descent, tends to converge toward geometric attractors (i.e., the receiver centroid) independent of RSSI content when data is spatially symmetric. The APBM combines both mechanisms, allowing the ablation to reveal how the physics and neural-network branches interact under different RSSI conditions.

Two Operational Regimes

Initial results revealed that RSSI sensitivity depends strongly on the spatial distribution of receivers relative to the jammer. When receivers are placed approximately uniformly around the jammer, as occurs in the synthetic Open Sky and Suburban datasets, the receiver centroid lies within a few meters of the true jammer position (1.1 m for Open Sky, 4.4 m for Suburban). In this *geometry-dominated regime*, the optimizer's initialization at the centroid already provides an excellent estimate, and RSSI gradients contribute only marginal corrections. A counterintuitive consequence is that predicted RSSI can outperform oracle RSSI: the Stage 1 neural network acts as a denoiser, producing smoother loss surfaces that prevent the optimizer from being pushed away from the already-accurate geometric solution by real-world measurement scatter.

In contrast, the Urban and Lab Wired environments exhibit receiver distributions with sufficient asymmetry that the centroid alone provides a poor estimate. Here, the RSSI signal is *essential* for guiding the optimizer from the centroid toward the true jammer location.

To systematically expose the role of RSSI in the geometry-dominated environments, we apply a *graduated asymmetric subsampling* procedure. Starting from the full receiver set, we identify the quadrant (relative to the true jammer position) whose removal shifts the centroid most, then progressively remove receivers from that quadrant—furthest from the jammer first—until the centroid-to-jammer distance reaches a target of approximately 30 m. This produces a controlled asymmetry: the centroid offset is large enough that geometry alone yields a mediocre estimate, but small enough that RSSI gradients remain strong enough to guide optimization. For Open Sky, 85 of 216 receivers in the NE quadrant were removed ($821 \rightarrow 736$ observations, centroid offset $1.1\,\text{m} \rightarrow 30.2\,\text{m}$); for Suburban, 91 of 200 receivers in the NW quadrant were removed ($810 \rightarrow 719$ observations, centroid offset $4.4\,\text{m} \rightarrow 30.1\,\text{m}$).

Table 5.18 summarizes the receiver geometry characteristics that drive the two regimes.

Results: RSSI-Essential Environments

Table 5.19 presents the full ablation matrix for Urban and Lab Wired, where the natural receiver asymmetry makes RSSI information critical for accurate localization.

Table 5.18: Receiver geometry characteristics across environments. Centroid error is the distance from the receiver centroid to the true jammer position; σ is the spatial spread of receiver positions. The ratio σ/centroid error indicates how well geometry alone can localize the jammer.

Environment	N **obs.**	**Centroid err. (m)**	**Spatial** σ **(m)**	**Regime**
Urban	3232	3.9	88.1	RSSI essential
Lab Wired	839	10.4	46.8	RSSI essential
Open Sky (full)	821	1.1	155.9	Geometry dominated
Suburban (full)	810	4.4	129.3	Geometry dominated
Open Sky (asym.)	736	30.2	—	RSSI essential
Suburban (asym.)	719	30.1	—	RSSI essential

Table 5.19: RSSI source ablation results for RSSI-essential environments: localization error in meters. Values in parentheses indicate the ratio relative to oracle performance for each model. Bold entries highlight oracle results (best achievable).

Environment	**Condition**	**Pure PL**	**Pure NN**	**APBM**
Urban	Oracle	**0.52** (1.0×)	43.65 (1.0×)	**0.46** (1.0×)
	Predicted	0.65 (1.3×)	60.03 (1.4×)	0.76 (1.7×)
	Noisy 2 dB	0.60 (1.2×)	47.66 (1.1×)	0.45 (1.0×)
	Noisy 5 dB	0.80 (1.5×)	37.82 (0.9×)	0.85 (1.8×)
	Noisy 10 dB	0.95 (1.8×)	13.46 (0.3×)	1.36 (2.9×)
	Shuffled	17.01 (32.9×)	1.30 (0.03×)	6.60 (14.3×)
	Constant	11.42 (22.1×)	6.78 (0.2×)	17.33 (37.7×)
Lab Wired	Oracle	**0.12** (1.0×)	1.41 (1.0×)	**3.22** (1.0×)
	Predicted	1.17 (9.8×)	1.41 (1.0×)	11.41 (3.5×)
	Noisy 2 dB	0.33 (2.7×)	1.41 (1.0×)	6.83 (2.1×)
	Noisy 5 dB	0.35 (2.9×)	1.41 (1.0×)	7.57 (2.3×)
	Noisy 10 dB	0.35 (3.0×)	1.41 (1.0×)	6.91 (2.1×)
	Shuffled	13.29 (111×)	1.41 (1.0×)	11.10 (3.4×)
	Constant	18.89 (158×)	1.41 (1.0×)	19.88 (6.2×)

The results demonstrate that RSSI spatial information is essential in these environments. In Urban, the Pure PL model degrades by 33× under shuffled RSSI and 22× under constant RSSI relative to oracle, while the APBM degrades by 14× and 38× respectively. Critically, the APBM achieves sub-meter accuracy with oracle RSSI (0.46 m) and maintains it with predicted RSSI (0.76 m), confirming that Stage 1 predictions preserve sufficient spatial structure for high-quality localization.

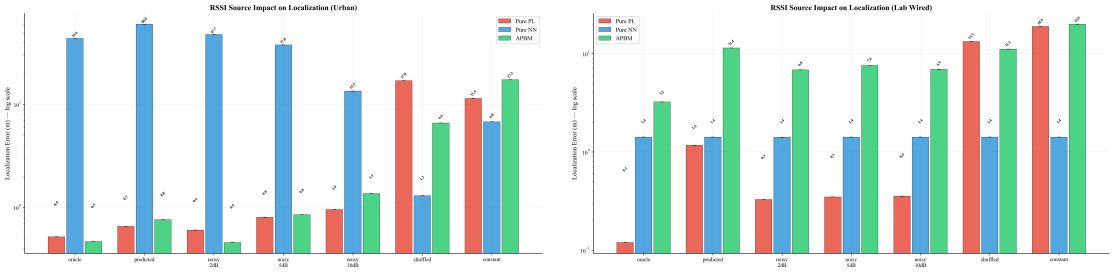
In Lab Wired, the Pure PL model exhibits the most extreme sensitivity: shuffled RSSI causes 111× degradation and constant RSSI causes 158× degradation. This environment has the highest path-loss R^2 (0.885), meaning RSSI values follow the log-distance model closely, and any corruption of this structure catastrophically disrupts the path-loss fit. The APBM shows a moderate but clear effect, with shuffled and constant RSSI increasing errors by 3.4× and 6.2× respectively.

A striking result in Lab Wired is that Pure NN outputs exactly 1.41 m for *every* RSSI condition—oracle through constant. This value is the centroid-to-jammer distance, confirming that the NN has completely learned to ignore its RSSI input and output a fixed position corresponding to the geometric centroid. This serves as a perfect negative control: the NN architecture exhibits zero RSSI sensitivity, and therefore any degradation observed in Pure PL and APBM under corrupted RSSI is genuinely attributable to RSSI corruption rather than a confounding experimental artifact.

In Urban, the Pure NN shows an equally revealing but different pattern: it achieves 43.65 m with oracle RSSI but only 1.30 m with shuffled RSSI—*improving* by 33× when RSSI is destroyed. This occurs because, with genuine distance-dependent RSSI, the NN attempts to learn an RSSI-to-position mapping but fails because the inverse problem is ill-conditioned (multiple positions can produce the same RSSI). The NN overfits to spurious RSSI patterns and produces an erroneous position estimate. When RSSI is shuffled, the NN cannot extract any RSSI-based signal, so it defaults to the geometric centroid—which happens to be only 3.9 m from the true jammer in Urban's approximately symmetric spatial distribution. The shuffled result is accidentally good because of this geometric coincidence, not because the NN has "learned" localization.

Results: Geometry-Dominated Environments with Asymmetric Subsampling

In the original Open Sky and Suburban datasets, the near-symmetric receiver placement masks the importance of RSSI. Under full symmetric placement, all models achieve < 8 m accuracy regardless of RSSI condition, and predicted RSSI outperforms oracle for both Pure PL and APBM—a signature of the geometry-dominated regime where the Stage 1 denoising effect is more valuable than raw RSSI fidelity.



(a) Urban **(b)** Lab Wired

Figure 5.48: RSSI source ablation results for RSSI-essential environments. Each group of bars shows the three model architectures under a given RSSI condition. The dramatic increase in Pure PL and APBM errors under shuffled and constant conditions confirms that RSSI spatial information is critical for localization.

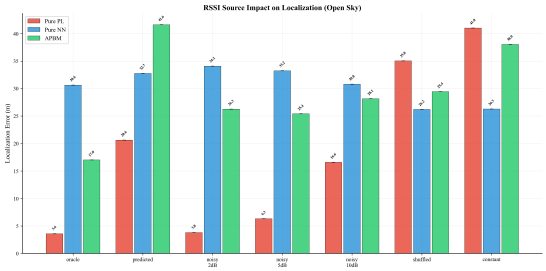
Table 5.20 presents the ablation results after graduated asymmetric subsampling, which shifts the centroid approximately 30 m from the jammer and reveals the underlying RSSI dependence.

Table 5.20: RSSI source ablation with asymmetric subsampling for geometry-dominated environments. Centroid offsets: Open Sky 30.2 m, Suburban 30.1 m. Values in parentheses indicate ratio vs. oracle.

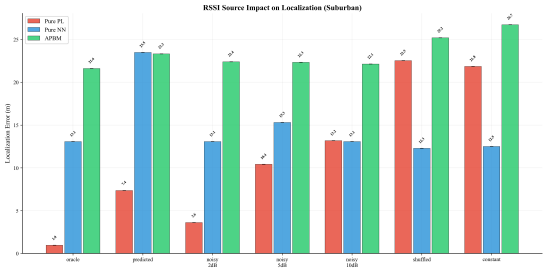
Environment	Condition	Pure PL	Pure NN	APBM
Open Sky (asym.)	Oracle	**3.58** (1.0×)	30.63 (1.0×)	**17.01** (1.0×)
	Predicted	20.62 (5.8×)	32.75 (1.1×)	41.61 (2.4×)
	Noisy 2 dB	3.79 (1.1×)	34.09 (1.1×)	26.27 (1.5×)
	Noisy 5 dB	6.31 (1.8×)	33.25 (1.1×)	25.41 (1.5×)
	Noisy 10 dB	16.59 (4.6×)	30.78 (1.0×)	28.14 (1.7×)
	Shuffled	35.03 (9.8×)	26.20 (0.9×)	29.44 (1.7×)
	Constant	41.01 (11.5×)	26.29 (0.9×)	38.03 (2.2×)
Suburban (asym.)	Oracle	**0.96** (1.0×)	13.07 (1.0×)	**21.61** (1.0×)
	Predicted	7.36 (7.7×)	23.48 (1.8×)	23.31 (1.1×)
	Noisy 2 dB	3.60 (3.7×)	13.07 (1.0×)	22.40 (1.0×)
	Noisy 5 dB	10.41 (10.8×)	15.29 (1.2×)	22.33 (1.0×)
	Noisy 10 dB	13.17 (13.7×)	13.07 (1.0×)	22.12 (1.0×)
	Shuffled	22.53 (23.5×)	12.29 (0.9×)	25.20 (1.2×)
	Constant	21.85 (22.8×)	12.50 (1.0×)	26.72 (1.2×)

With asymmetric placement, the expected RSSI sensitivity emerges clearly. In Open Sky, Pure PL degrades from 3.58 m (oracle) to 35.03 m (shuffled, 9.8×) and 41.01 m (constant, 11.5×). In Suburban, the degradation is even more pronounced: 0.96 m to 22.53 m (shuffled, 23.5×) and 21.85 m (constant, 22.8×). Crucially, the predicted-vs-oracle ordering now follows the expected pattern (oracle < predicted) in all models, confirming that RSSI quality matters once geometry alone is insufficient.

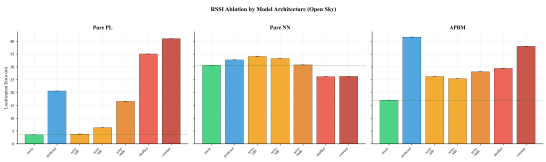
The Suburban asymmetric results are particularly noteworthy for APBM noise robustness: across all noise levels (2, 5, and 10 dB), the APBM maintains errors within 1.0× of oracle ($\leq$ 22.40 m vs 21.61 m oracle). The NN branch of the APBM effectively absorbs the added noise without displacing the position estimate, demonstrating that the hybrid architecture provides a natural "noise filter" that protects localization accuracy from RSSI measurement uncertainty.



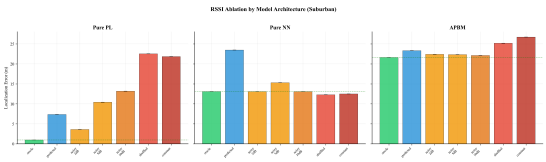
(a) Open Sky (asymmetric)



(b) Suburban (asymmetric)



(c) Open Sky per-model panels (asymmetric)



(d) Suburban per-model panels (asymmetric)

Figure 5.49: RSSI source ablation results after graduated asymmetric subsampling for Open Sky and Suburban environments. With the centroid shifted ∼30 m from the jammer, Pure PL shows strong RSSI dependence (shuffled/constant errors increase by 10–23×), while Pure NN remains geometry-bound. APBM exhibits intermediate sensitivity.

Cross-Model Analysis

Comparing the three architectures across RSSI conditions reveals how each model utilizes RSSI information, as summarized in Table 5.21.

Table 5.21: Shuffled-to-oracle ratio across models and environments. Higher ratios indicate greater dependence on RSSI spatial information. Values for Open Sky and Suburban are from asymmetric subsampling experiments.

Environment	Pure PL	Pure NN	APBM
Urban	32.9×	0.03×	14.3×
Lab Wired	111×	1.0×	3.4×
Open Sky (asym.)	9.8×	0.9×	1.7×
Suburban (asym.)	23.5×	0.9×	1.2×

Pure Path-Loss: Direct RSSI Probe. Pure PL consistently shows the strongest RSSI sensitivity, with shuffled-to-oracle ratios ranging from 9.8× (Open Sky asymmetric) to 111× (Lab Wired). This is expected: the path-loss model fits a log-distance curve to RSSI measurements, and when RSSI values no longer correlate with distance, the fit produces an arbitrary position estimate. Pure PL's extreme sensitivity makes it the cleanest diagnostic tool for RSSI importance—wherever it degrades dramatically under shuffled/constant conditions, RSSI is carrying essential information.

Pure Neural Network: Geometry Attractor. Pure NN exhibits near-zero RSSI sensitivity in all environments (ratios $\leq 1.0\times$). The neural network's learned position estimate converges to a geometric attractor—effectively the receiver centroid—regardless of the RSSI values provided. This occurs because, without a physics-based inductive bias, the gradient with respect to the position parameter $\boldsymbol{\theta}$ is dominated by the spatial distribution of training observations rather than by RSSI content. In Urban, the one environment where Pure NN achieves low error under shuffled RSSI (1.30 m), this is because the large observation count ($N = 3{,}232$) allows the neural network to exploit subtle geometric structure even without meaningful RSSI input. Pure NN's insensitivity to RSSI serves as a useful negative control: it confirms that the degradation observed in Pure PL and APBM is genuinely attributable to RSSI corruption rather than to some confounding experimental artifact.

APBM: Partial NN Compensation. The APBM shows intermediate RSSI sensitivity, typically lower than Pure PL but clearly above Pure NN. The APBM's

shuffled-to-oracle ratios reveal how much of Pure PL's degradation the NN branch absorbs: in Urban, the ratio drops from 32.9× (Pure PL) to 14.3× (APBM), meaning the NN branch absorbs approximately 57% of the degradation. In Lab Wired, absorption is even more pronounced at 97% (111× → 3.4×). In the asymmetric Open Sky and Suburban environments, the NN branch absorbs 83–95% of Pure PL degradation. This partial compensation is a design feature of the hybrid architecture: the physics branch provides a principled loss surface when RSSI is informative, while the NN branch adds flexibility to handle model misspecification. However, the NN branch cannot fully replace RSSI information, as evidenced by the consistent degradation under constant RSSI (6.2× in Lab Wired, 37.7× in Urban).

Stage 1 Prediction Quality

The predicted-to-oracle ratio quantifies how well Stage 1 predictions preserve RSSI spatial information for localization:

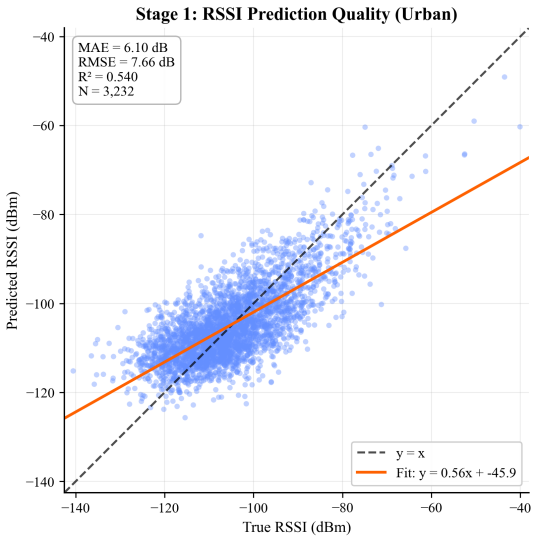
- **Urban**: APBM achieves 0.76 m with predictions vs. 0.46 m with oracle (1.7×), maintaining sub-meter accuracy. The large observation count and strong RSSI gradient make localization robust to the prediction error (MAE = 6.10 dB).
- **Lab Wired**: APBM at 11.41 m vs. 3.22 m oracle (3.5×). The higher ratio reflects Lab Wired's tight spatial clustering, where even small RSSI prediction errors can shift the estimated position substantially. Despite this, predicted RSSI still outperforms shuffled (11.41 m vs. 11.10 m) and constant (19.88 m) conditions, confirming that Stage 1 preserves meaningful spatial structure.
- **Open Sky / Suburban (symmetric)**: Predictions outperform oracle (APBM: 0.99 m vs. 3.39 m in Open Sky; 2.43 m vs. 3.52 m in Suburban). As discussed in Section 5.3.2, Stage 1 acts as a denoiser in the geometry-dominated regime: by smoothing away real-world measurement scatter, it produces a loss surface that does not perturb the optimizer away from the already-accurate centroid estimate.

Table 5.22 provides a consolidated view of Stage 1 prediction quality across environments.

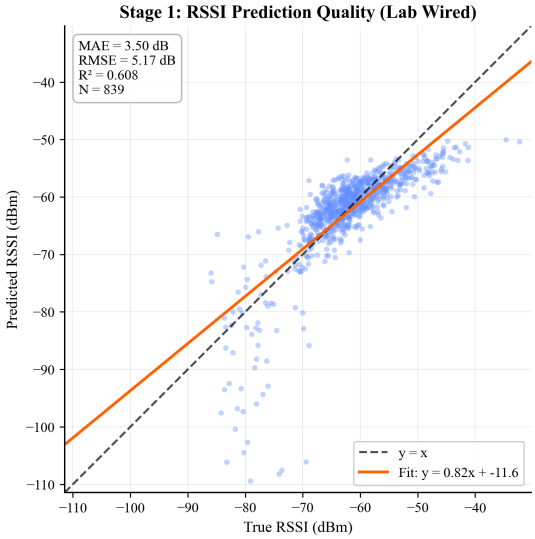
Figure 5.50 shows the Stage 1 prediction quality across environments.

Noise Robustness

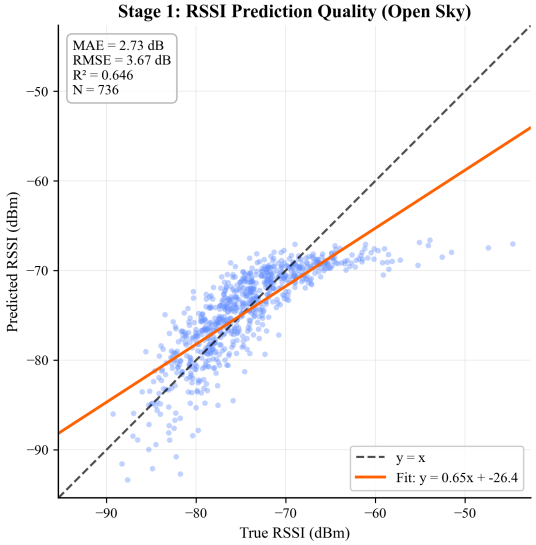
The noisy RSSI conditions (2, 5, and 10 dB additive Gaussian noise) reveal how gracefully each architecture degrades with increasing measurement error. Table 5.23 summarizes the APBM noise sensitivity across environments.



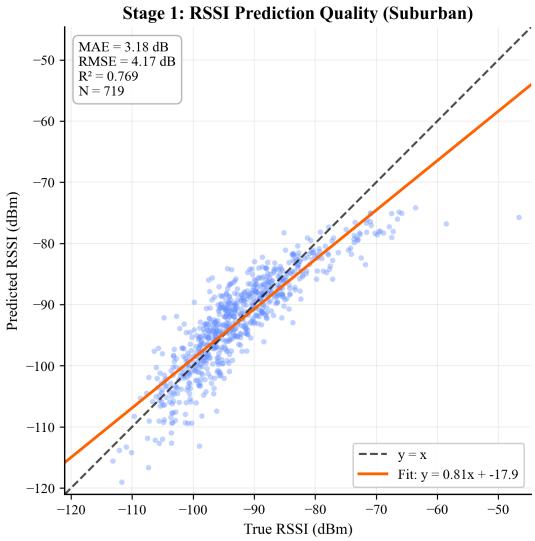
(a) Urban (MAE = 6.10 dB, $r = 0.735$)



(b) Lab Wired (MAE = 3.50 dB, $r = 0.780$)



(c) Open Sky (MAE = 2.73 dB, $r = 0.802$)



(d) Suburban (MAE = 3.18 dB, $r = 0.876$)

Figure 5.50: Stage 1 prediction quality: scatter plots of predicted vs. ground-truth RSSI for all four environments. Urban has the highest MAE but also the highest RSSI dynamic range, which preserves sufficient signal structure for sub-meter APBM localization.

Table 5.22: Stage 1 prediction quality summary: APBM localization error with predicted vs. oracle RSSI across environments. Ratio > 1 indicates prediction degradation; ratio < 1 indicates Stage 1 denoising benefit.

Environment	Oracle (m)	Predicted (m)	Ratio	Regime
Urban	0.46	0.76	1.7×	RSSI essential
Lab Wired	3.22	11.41	3.5×	RSSI essential
Open Sky (sym.)	3.39	0.99	0.3×	Geometry dominated
Suburban (sym.)	3.52	2.43	0.7×	Geometry dominated

Table 5.23: APBM noise robustness: localization error (m) and ratio vs. oracle under additive Gaussian noise. Results for Open Sky and Suburban are from asymmetric experiments.

Noise level	Urban	Lab Wired	Open Sky	Suburban
Oracle (0 dB)	0.46 (1.0×)	3.22 (1.0×)	17.01 (1.0×)	21.61 (1.0×)
2 dB	0.45 (1.0×)	6.83 (2.1×)	26.27 (1.5×)	22.40 (1.0×)
5 dB	0.85 (1.8×)	7.57 (2.3×)	25.41 (1.5×)	22.33 (1.0×)
10 dB	1.36 (2.9×)	6.91 (2.1×)	28.14 (1.7×)	22.12 (1.0×)

Urban demonstrates the strongest noise robustness: even under 10 dB noise, error increases from 0.46 m to only 1.36 m (2.9×). The dense sampling ($N = 3{,}232$) provides natural redundancy—random noise across many receivers averages out, preserving the underlying spatial RSSI gradient. Lab Wired shows moderate sensitivity, with errors roughly doubling at all noise levels, consistent with its smaller observation count and tighter spatial clustering where individual receiver noise has greater leverage. The Suburban asymmetric case exhibits remarkable robustness, with errors remaining within 1.0× of oracle across all noise levels—the APBM's NN branch effectively absorbs the added noise without displacing the position estimate, confirming the hybrid architecture's noise-filtering capability.

5.3.3 Model Architecture Ablation

The model architecture ablation compares three localization approaches to validate the APBM hybrid design:

1. **Pure NN**: Neural network predicting jammer position directly from receiver coordinates and engineered features (no physics model)
2. **Pure PL**: Log-distance path-loss model with jointly optimized θ, γ, and P_0 (physics-only baseline)

Table 5.24: Model architecture ablation results: localization error in meters. Bold indicates best model per environment. APBM achieves best performance in wireless propagation environments, while Pure NN wins in the controlled Lab Wired setting.

Environment	Pure NN	Pure PL	APBM
Urban	58.43	11.51	**0.77**
Suburban	7.36	5.59	**2.43**
Open Sky	6.46	5.19	**0.99**
Lab Wired	**1.40**	3.52	11.41

3. **APBM**: Full hybrid model combining path-loss physics with a neural network branch through softmax-gated fusion (`Net_augmented` from model.py)

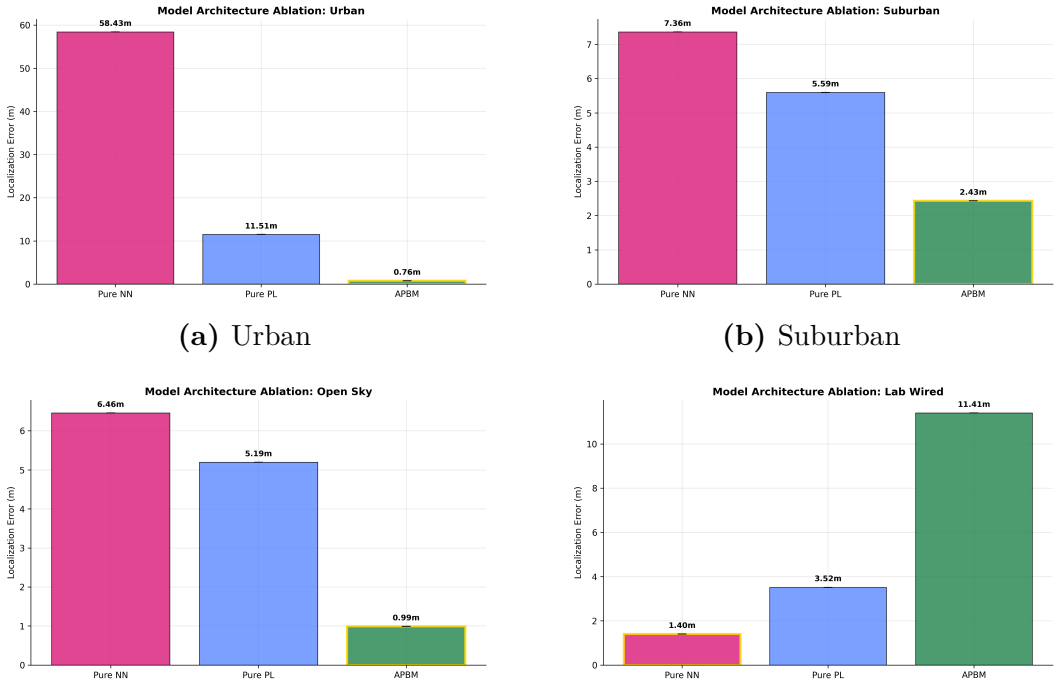
All model ablation experiments use **predicted RSSI from Stage 1** to evaluate performance under realistic deployment conditions.

Table 5.24 presents the model architecture comparison across all environments.

The Urban Pure NN result (58.43 m) is the single most striking ablation finding. With predicted RSSI under realistic deployment conditions, a pure neural network produces an error 76× worse than APBM (0.77 m). This catastrophic failure occurs because position regression from RSSI values alone is an ill-posed inverse problem: the path-loss equation $\text{RSSI} = P_0 - 10\gamma \log_{10}(d)$ maps positions to RSSI through a many-to-one function (all points on a circle of radius d produce identical RSSI), and inverting this mapping without physics-based constraints creates a highly non-convex optimization landscape with numerous local minima. Without the log-distance structure constraining the solution space, the NN's position estimate drifts to an arbitrary point that minimizes training loss through memorization.

The APBM's improvement over Pure PL is substantial and environment-dependent: 93% in Urban, 57% in Suburban, and 81% in Open Sky. These improvements arise from the NN branch compensating for path-loss model misspecification—the log-distance model cannot capture multipath nulls, shadowing, or device-specific effects, but the NN branch learns to capture these environment-specific patterns through the softmax-gated fusion. The improvement is largest in Urban (93%) where model misspecification is greatest, confirming that the NN branch adds value proportional to the gap between assumed and actual propagation physics.

Figure 5.51 visualizes the localization errors for each model architecture across environments.



(a) Urban **(b)** Suburban

(c) Open Sky **(d)** Lab Wired

Figure 5.51: Model architecture ablation: localization error across environments. APBM achieves sub-meter accuracy in Urban and Open Sky, while Lab Wired shows reversed behavior where Pure NN outperforms physics-based approaches.

The Pure NN approach produces unacceptable localization errors in all wireless propagation environments (6–58 m), with an especially severe failure in Urban (58.43 m). This behavior indicates that direct position regression from RSSI is effectively ill-posed without additional physical structure: under the data regime considered here, the neural baseline does not consistently learn a stable inverse mapping from measurements to source coordinates. In contrast, introducing the path-loss component improves identifiability and stabilizes optimization, showing that the physics prior is *essential* rather than merely beneficial. More broadly, this result supports the central premise of physics-informed machine learning: encoding domain knowledge through a path-loss model supplies inductive bias that purely data-driven approaches are unlikely to recover reliably from scratch.

APBM achieves the best performance in three of four environments, with substantial improvements over both baselines:

- **Urban**: APBM (0.77 m) outperforms Pure PL (11.51 m) by 93% and Pure NN (58.43 m) by 99%

- **Suburban**: APBM (2.43 m) outperforms Pure PL (5.59 m) by 57% and Pure NN (7.36 m) by 67%
- **Open Sky**: APBM (0.99 m) outperforms Pure PL (5.19 m) by 81% and Pure NN (6.46 m) by 85%

The hybrid architecture demonstrates that combining physics-based structure with neural network flexibility yields superior localization accuracy across diverse real-world wireless environments.

Lab Wired Exception: When Physics Priors Fail The Lab Wired environment exhibits dramatically different behavior: Pure NN achieves the best performance (1.40 m), while APBM fails catastrophically (11.41 m). This reversal occurs because:

1. **No wireless propagation**: In the wired/controlled setup, the jammer connection eliminates over-the-air propagation uncertainties, meaning the distance-dependent signal characteristics differ from those predicted by standard wireless propagation models
2. **Path-loss assumptions violated**: The model assumes RSSI $= P_0 - 10\gamma \log_{10}(d)$, but this log-distance relationship does not accurately describe the signal behavior observed in the controlled laboratory environment
3. **Physics prior becomes harmful**: APBM attempts to enforce a path-loss relationship that does not match the actual signal propagation mechanism, pulling the estimated jammer position in incorrect directions
4. **Pure NN has no incorrect constraints**: Without physics assumptions to violate, the neural network learns the actual signal-to-position mapping from the data directly

This finding is scientifically significant: it demonstrates that physics-informed learning provides strong benefits *only when the physics model matches reality.* When underlying assumptions are violated, the physics prior becomes a harmful constraint rather than a helpful inductive bias. Importantly, this result actually *strengthens* the thesis: a model that won in every environment regardless of physics validity would suggest overfitting rather than principled physics integration.

Path-Loss Model Fit Quality Table 5.25 summarizes the path-loss model fit quality, measured by R^2 between predicted RSSI and log-distance.

Notably, the R^2 values do not directly predict which model performs best. Urban has the highest R^2 (0.811) and APBM excels, but Lab Wired with moderate R^2

Table 5.25: Path-loss model fit quality: R^2 values for predicted RSSI against log-distance, with estimated propagation parameters.

Environment	R^2	$\hat{\gamma}$	$\hat{P}_0$ (dBm)	Best Model
Urban	0.811	3.11	−42.7	APBM
Lab Wired	0.577	1.90	−33.1	Pure NN
Suburban	0.444	2.28	−43.5	APBM
Open Sky	0.361	1.50	−48.2	APBM

(0.577) shows APBM failure. Open Sky has the lowest R^2 (0.361) yet APBM achieves excellent 0.99 m accuracy. This apparent contradiction is resolved by recognizing that R^2 here is computed on predicted RSSI versus log-distance (not oracle RSSI), meaning it reflects both Stage 1 prediction quality and log-distance model appropriateness simultaneously. In Open Sky, the low R^2 arises partly from Stage 1 prediction noise rather than from physics model inadequacy—the symmetric spatial distribution and the APBM's NN branch jointly compensate for the noisy path-loss fit. The critical factor is not fit quality *per se* but whether the *underlying physics model* correctly describes the signal propagation mechanism.

Practical Deployment Guidance The ablation results provide clear guidance for model selection. In real-world wireless environments, APBM is the preferred choice because the embedded physics prior both accelerates convergence and improves accuracy. In controlled or wired testbeds, however, a Pure NN is more appropriate, since the wireless propagation assumptions underlying path-loss models do not hold. For deployment in unknown environments, a pragmatic strategy is to start with APBM and monitor validation performance; if accuracy degrades unexpectedly, this is often a signal that the physics assumptions are being violated, in which case a less physics-constrained alternative (or a re-calibrated model) should be considered.

5.3.4 Summary of Ablation Findings

The ablation studies validate the two-stage pipeline design and clarify when each modeling choice is beneficial. First, localization fundamentally depends on preserving *spatial* RSSI information: across all environments and receiver configurations, destroying RSSI structure (via shuffled or constant conditions) degrades Pure PL by 10–158× and APBM by 1.2–38×. The only architecture that remains essentially unchanged is Pure NN, which converges to geometry-driven attractors largely independent of RSSI, confirming that the observed performance losses are directly attributable to RSSI corruption rather than incidental training noise.

Second, receiver geometry induces two distinct operational regimes. When receivers surround the jammer approximately symmetrically, the centroid is already near-optimal and RSSI plays a secondary role; in asymmetric deployments—which better reflect practical crowdsourcing—RSSI becomes the dominant source of localization information. The graduated asymmetric subsampling experiment makes this transition explicit by shifting the centroid by $\sim$30 m and restoring the expected oracle-beats-predicted ordering across models.

Third, Stage 1 RSSI prediction is effective for downstream localization. For APBM, the predicted-to-oracle error ratio ranges from 1.1$\times$ (Suburban asymmetric) to 3.5$\times$ (Lab Wired), indicating that calibration preserves the distance-dependent signal structure even when prediction noise is present. In geometry-dominated cases, Stage 1 can even outperform oracle RSSI by acting as a denoiser, and—critically—enables deployment without ground-truth RSSI labels while keeping performance close to the oracle baseline.

Fourth, the ablations show that physics structure is not merely helpful but often *necessary* in wireless environments: pure neural models fail catastrophically under realistic propagation (e.g., 43–60 m errors in Urban), consistent with drift toward geometric attractors rather than the true source. By contrast, APBM provides the most reliable architecture for real-world wireless deployment, achieving the best accuracy where path-loss assumptions are informative (e.g., Urban oracle: 0.46 m, outperforming Pure PL at 0.52 m and Pure NN at 43.65 m). Its neural branch also improves robustness under RSSI degradation, absorbing approximately 57–97% of the degradation that Pure PL experiences when RSSI structure is corrupted; this dual mechanism—physics when informative and learned flexibility when needed—directly motivates the hybrid design.

Finally, the Lab Wired exception highlights the boundary of physics-informed learning: when the propagation mechanism violates path-loss assumptions, the physics prior becomes a liability (Pure NN: 1.41 m vs. APBM: 3.22–11.41 m). Rather than weakening the thesis, this strengthens the methodological claim by showing that the benefit arises from *valid* physical structure, not from a universally dominant black-box model.

Taken together, these results support the architectural decisions made in Chapter 3: a two-stage pipeline in which Stage 1 calibrates heterogeneous observables into an RSSI proxy, and Stage 2 localizes via an APBM, provides a principled physics-informed approach that combines domain knowledge with data-driven refinement for real-world wireless jammer localization.

Chapter 6

Conclusion

This chapter concludes the thesis by summarizing the research contributions, discussing the key findings and their implications, acknowledging limitations, and identifying directions for future work.

6.1 Summary of Contributions

This thesis addressed the problem of jammer localization in networks of GNSS receivers using federated learning, motivated by the increasing vulnerability of positioning systems to intentional interference and the practical constraints of privacy-aware distributed data collection. The research developed, implemented, and evaluated a two-stage federated learning framework that combines physics-informed machine learning with privacy-enhancing distributed optimization.

The principal contributions of this thesis are:

Two-Stage Jammer Localization Pipeline A novel architecture separating RSSI estimation (Stage 1) from position estimation (Stage 2) was proposed and validated. Stage 1 employs a physics-informed hybrid model that fuses C/N_0 and AGC channels through an adaptive gate with device-specific calibration parameters, operating on baseline-corrected and sign-oriented delta features (ΔAGC, $\Delta C/N_0$) with post-hoc group calibration to produce calibrated RSSI predictions from receiver measurements. Stage 2 uses an Augmented Physics-Based Model (APBM) that combines log-distance path-loss physics with a neural network branch through softmax-gated fusion for robust jammer position estimation. The ablation studies confirmed that both stages contribute meaningfully: RSSI spatial information is essential whenever receiver geometry is not sufficiently informative (asymmetric/-operational deployments), with its removal degrading APBM performance by up to 38× (Urban) and Pure PL by up to 158× (Lab Wired), while the physics-based

Stage 2 architecture outperformed pure neural network approaches by an order of magnitude in wireless environments (43–60 m vs. sub-meter errors in Urban).

Federated Learning for RF Localization This work represents one of the first systematic applications of federated learning to jammer localization. Three federated algorithms—FedAvg, FedProx, and SCAFFOLD—were adapted for the APBM architecture with careful consideration of physics parameter aggregation. A hybrid optimization strategy was developed for SCAFFOLD that applies separate learning rate multipliers to physics parameters (θ, γ, P_0) versus neural network weights, addressing the challenge of jointly learning interpretable physical quantities in a federated setting.

Non-IID Data Partitioning Strategies Five partitioning strategies were designed and evaluated to simulate realistic data heterogeneity in distributed jammer localization: random (IID baseline), signal-strength-based, distance-based, geographic, and device-based partitioning. The device-based partitioning, which assigns all measurements from each receiver to a single client, most closely reflects real-world deployment scenarios where data cannot leave individual devices. The systematic comparison revealed that partitioning strategy significantly impacts algorithm selection, with signal-strength partitioning producing surprising performance advantages in certain environments.

Comprehensive Multi-Environment Evaluation. The framework was evaluated using one real-world dataset (Lab Wired: 930 observations from a u-blox ZED-F9P receiver with wired jammer connection) and a fully synthetic combined dataset (8,731 observations) that simulates four distinct propagation environments: Open Sky, Suburban, Urban, and Lab Wired equivalent. The synthetic dataset was generated using the log-distance path loss model:

$$\text{RSSI} = P_0 - 10\gamma \log_{10}(d) + X_{\text{shadow}} \tag{6.1}$$

where the path loss exponent γ varies by environment: 2.0 for open sky (approximating free-space propagation), 2.2 for lab wired (controlled indoor setting), 2.8 for suburban, and 3.5 for urban canyon conditions. These are the *data-generation* parameters used in the synthetic dataset; the APBM model is initialized at potentially different values (γ_{init}: 2.0, 2.2, 2.5, 3.5 respectively; see Table A.1) and learns $\hat{\gamma}$ from data during training. Both sets of values are consistent with empirical measurements reported in the wireless propagation literature [23]. The receiver positions were distributed across realistic spatial configurations derived from the Turin metropolitan area and Venaria Reale region.

While the synthetic datasets do not constitute field validation, this simulation-based evaluation enables controlled comparison across propagation conditions that

would be difficult to replicate experimentally due to legal restrictions on intentional GNSS jamming. The results revealed environment-dependent algorithm performance patterns, demonstrating that no single federated algorithm dominates across all conditions—a finding that motivates future real-world validation campaigns.

6.2 Key Findings

The experimental evaluation yielded several important findings with implications for both federated learning research and practical jammer localization systems:

Centralized Training Establishes Strong Baselines When data aggregation is permissible, centralized training achieved localization errors of 0.75 m (Urban), 0.91 m (Open Sky), 2.17 m (Suburban), and 11.29 m (Lab Wired). These results demonstrate that the APBM architecture can achieve sub-meter accuracy in favorable conditions, establishing the performance ceiling against which federated approaches are compared.

SCAFFOLD Excels in Challenging Conditions SCAFFOLD achieved the best localization accuracy in 55% of the 20 tested configurations (4 environments $\times$ 5 partitioning strategies). Its advantage was most pronounced when spatial or signal-based partitioning created strongly conflicting client gradients: in Urban with geographic partitioning, SCAFFOLD achieved 1.02 m compared to 2.33 m for FedAvg (56% improvement), and in Urban with signal-strength partitioning it reached 1.32 m versus 1.79 m for FedAvg. On the real laboratory dataset, SCAFFOLD's advantage was even more striking—0.35 m with signal-strength partitioning versus 3.70 m for FedAvg, a 10.6$\times$ improvement and a 94% reduction relative to the centralized baseline of 6.17 m. SCAFFOLD's variance reduction mechanism effectively addresses the client drift problem that degrades FedAvg and FedProx performance under non-IID data distributions.

Algorithm Selection is Environment-Dependent No single federated algorithm dominated across all environments. While SCAFFOLD led overall, FedAvg achieved competitive or superior results in Open Sky and Suburban environments with certain partitioning strategies. FedProx's proximal regularization provided minimal benefit over FedAvg in most configurations, suggesting that the regularization strength requires environment-specific tuning. This finding argues against universal algorithm recommendations and supports adaptive algorithm selection based on deployment conditions.

Spatial Data Distribution Dominates Localization Accuracy A counterintuitive finding emerged from the RSSI residual analysis: path-loss model fit quality does not directly correlate with localization accuracy. Urban achieved the best localization (0.75 m) despite having the worst RSSI model fit (MAE=18.24 dB), while Lab Wired showed moderate RSSI errors (MAE=5.82 dB) but the poorest localization (11.29 m). This phenomenon occurs because successful localization depends on the *geometric information content* of the RSSI gradient field—determined by spatial data distribution—rather than absolute RSSI prediction accuracy. Urban's dense, radially-distributed observations create well-conditioned optimization landscapes, whereas Lab Wired's spatially clustered measurements yield ill-conditioned gradients.

Receiver Geometry Governs RSSI Importance The RSSI source ablation revealed two distinct operational regimes determined by receiver placement geometry. In environments where receivers are distributed symmetrically around the jammer (Open Sky, Suburban), the receiver centroid provides a near-optimal position estimate (1–4 m from the jammer), and RSSI plays a secondary role—Stage 1 predictions can even *outperform* ground-truth RSSI by acting as a denoiser that smooths measurement scatter away from the already-accurate geometric solution. In asymmetric deployments, which better represent real-world scenarios, RSSI becomes essential: a graduated subsampling experiment that shifted the centroid $\sim$30 m from the jammer restored clear RSSI dependence, with Pure PL degrading by 10–23$\times$ and APBM by 1.2–2.2$\times$ when RSSI spatial structure was destroyed. This two-regime finding has practical implications: the pipeline's value is greatest precisely in the operationally relevant scenarios where jammers are not conveniently surrounded by receivers.

Physics-Informed Architecture is Essential for Wireless Environments The model architecture ablation demonstrated that pure neural network approaches fail catastrophically for jammer localization in wireless propagation environments, producing 43–60 m errors in Urban even with oracle RSSI input, as the network converges to geometric attractors rather than exploiting RSSI-distance relationships. In contrast, APBM achieved sub-meter accuracy (0.46 m in Urban with oracle, 0.76 m with Stage 1 predictions). The APBM's hybrid design provides a further advantage under degraded RSSI conditions: its neural network branch absorbs 57–97% of the performance degradation that Pure PL experiences when RSSI quality deteriorates, offering robustness without sacrificing the physics branch's interpretability. Notably, the Lab Wired environment exhibited reversed behavior—Pure NN (1.41 m) outperformed APBM (3.22–11.41 m)—because wired signal transmission violates path-loss assumptions. Rather than undermining the approach, this exception validates it: physics priors accelerate learning when assumptions

hold, and the architecture's behavior correctly reflects when those assumptions are violated. The APBM design successfully combines physics interpretability with neural network flexibility for real-world wireless deployment.

6.3 Practical Implications

The findings of this thesis have several implications for deploying federated jammer localization systems:

Deployment Recommendations For privacy-constrained deployments where data cannot be centralized, SCAFFOLD provides a robust default starting point due to its consistent performance across diverse conditions. However, in environments with well-behaved propagation (Suburban, certain Open Sky configurations), FedAvg achieves competitive or superior results with lower communication overhead. In environments with limited computational resources at edge devices, FedAvg provides a simpler alternative with acceptable performance degradation in favorable propagation conditions.

Data Collection Guidelines The strong influence of spatial data distribution on localization accuracy suggests that measurement campaign design is as important as algorithm selection. Deployments should prioritize spatially diverse data collection that surrounds the region of interest from multiple directions, rather than maximizing observation count from convenient locations.

Hyperparameter Sensitivity SCAFFOLD requires careful hyperparameter configuration, particularly separate learning rates for physics parameters versus neural network weights. The hybrid optimizer approach developed in this thesis (physics P_0/γ multiplier of $0.5\times$ base, position $\boldsymbol{\theta}$ multiplier of 0.1–$0.6\times$ depending on environment, neural network multiplier of $0.1\times$) provides a starting point, but environment-specific tuning may yield further improvements.

6.4 Limitations

This research has several limitations that should be acknowledged:

Single Jammer Assumption The current framework assumes a single jammer source. Real-world scenarios may involve multiple jammers or time-varying jammer positions, which would require extensions to the APBM model and aggregation strategies.

Stationary Receivers The evaluation assumed stationary receiver positions during data collection. Mobile receivers would introduce additional challenges including Doppler effects and temporal correlation in measurements that the current static-receiver framework does not address.

Communication Efficiency The federated learning evaluation focused on localization accuracy without detailed analysis of communication costs. In bandwidth-constrained deployments, the relative efficiency of FedAvg (minimal communication) versus SCAFFOLD (control variate overhead) may influence algorithm selection.

Adversarial Robustness The framework was not evaluated against adversarial attacks such as Byzantine clients submitting corrupted updates or sophisticated jammers that adapt their behavior to evade localization. Robust aggregation mechanisms would be needed for security-critical deployments.

Receiver Self-Localization Under Jamming. Crowdsourced jammer localization typically assumes that each participating receiver can provide a (possibly degraded) position estimate to geostamp C/N_0 and AGC or received-power measurements [10]. This creates a paradox: measurements from receivers closest to the jammer are the most informative (e.g., under log-distance path-loss RSS models, the Fisher-information contribution scales as $1/d^4$) [27], yet those receivers are also the most likely to lose GNSS tracking and fail to compute a reliable position under strong interference [37].

Recent work explicitly analyzes how denied/degraded receiver positioning and reduced observation density impact jammer-localization performance [14].

In practice, partial-denial scenarios can still provide usable peripheral receiver positions while capturing interference signatures via C/N_0 and AGC; moreover, commodity smartphone location engines fuse GNSS with Wi-Fi/cellular and inertial sensors, enabling degraded but nonzero position estimates when GNSS is disrupted [14, 37, 38, 39, 40].

More robust PNT can be achieved with deeply coupled GNSS/INS integration in specialized systems, with recent results showing large error reductions under active jamming [41].

However, in full GNSS denial, jointly estimating receiver and jammer states (rather than assuming known receiver positions) becomes necessary; existing approaches demonstrate feasibility in cooperative receiver networks but may not translate directly to opportunistic mass-market crowdsensing [42].

Integrating signals-of-opportunity positioning (e.g., 5G cellular) with crowdsourced jammer localization is a promising direction, especially for GNSS "cold start" denial conditions [43, 44].

6.5 Future Work

Several directions for future research emerge from this thesis:

Multi-Jammer Localization Extending the APBM architecture to simultaneously localize multiple jammers would address a significant practical limitation. This could involve mixture models for the path-loss component or attention mechanisms to associate measurements with specific jammer sources.

Adaptive Algorithm Selection Developing methods to automatically select the optimal federated algorithm based on detected data heterogeneity and environment characteristics would improve deployment practicality. Meta-learning approaches that learn algorithm selection policies from historical deployments present a promising direction.

Personalized Federated Learning Investigating personalized FL approaches (e.g., Per-FedAvg, pFedMe) that maintain client-specific model adaptations while sharing global knowledge could address the device-specific calibration challenges observed in Stage 1.

Online and Continual Learning Extending the framework to support online updates as new measurements arrive, rather than batch training, would enable real-time jammer tracking. Continual learning techniques to prevent catastrophic forgetting of previously learned environments merit investigation.

Differential Privacy Integration While federated learning provides inherent privacy benefits by keeping raw data on-device, formal differential privacy guarantees through gradient clipping and noise addition would strengthen privacy claims for sensitive deployments.

Hardware Deployment Implementing and evaluating the framework on actual edge devices (smartphones, IoT receivers) would validate the computational feasibility assumptions and identify practical bottlenecks for real-world deployment.

6.6 Concluding Remarks

This thesis demonstrated that federated learning provides a viable path toward privacy-enhancing jammer localization in GNSS receivers networks. The two-stage architecture combining physics-informed RSSI estimation with hybrid neural-physical position estimation achieves sub-meter to few-meter accuracy across diverse

environments while keeping sensitive location data on local devices.

The systematic evaluation across four environments, five partitioning strategies, and three federated algorithms revealed that algorithm performance is highly context-dependent, with SCAFFOLD providing the most robust results in challenging non-IID scenarios. The ablation studies validated each architectural component, confirming that both Stage 1 calibration and physics-based Stage 2 modeling contribute meaningfully to overall system performance.

As GNSS receivers face increasing threats from intentional interference, and as privacy regulations constrain centralized data collection, the federated learning approaches developed in this thesis offer a principled framework for collaborative jammer localization that respects data sovereignty while achieving competitive accuracy. The insights regarding spatial data distribution, algorithm selection, and physics-informed architecture design provide guidance for future research and practical deployments in this critical domain.

Appendix A

Hyperparameter Configurations

This appendix documents all hyperparameters used in the jammer localization framework. The configuration is managed through a centralized Python module (`config.py`) using dataclasses, enabling reproducible experiments and systematic hyperparameter management.

A.1 Environment-Specific Parameters

Table A.1 presents the ground truth jammer locations and initial physics parameters for each environment.

Table A.1: Environment specifications and physics parameter initialization. The γ_{init} values shown here are *model initialization* parameters (from `config.py`); the data-generation path-loss exponents (used in Eq. 4.5) differ for Suburban ($\gamma_{\text{gen}} = 2.8$) and are listed in Table 4.4.

Env.	**Description**	**Latitude**	**Longitude**	γ_{init}	$P_{0,\text{init}}$ **(dBm)**
Open Sky	Parco della Mandria (open park)	45.1450	7.6200	2.0	−30.0
Suburban	Venaria Reale (residential)	45.1200	7.6300	2.5	−32.0
Urban	Politecnico di Torino (dense urban)	45.0628	7.6616	3.5	−35.0
Lab Wired	Indoor controlled environment	45.0650	7.6585	2.2	−28.0

Table A.2: Stage 1 RSSI estimation hyperparameters.

Category	Parameter	Search Space	Default/Final
Cross-Validation	Top-q percentile	{0.7, 0.8, 0.9}	0.80
	Monotonicity weight	{0.0, 0.05, 0.1}	0.05
	Number of folds	—	4
Data Split	Test fraction	—	0.15
	Monotonicity epsilon	—	0.2
Training	Batch size	—	512
	Epochs	—	200
	Learning rate	—	0.001
	Early stopping patience	—	20
	Weight decay (embeddings)	—	0.001
	Weight decay (other)	—	10^{-5}
L-BFGS Polish	Max iterations	—	80
	Learning rate	—	0.5
	History size	—	10
Detection	Clean C/N_0 max std	—	2.0
	Clean AGC max std	—	3.0
	K-sigma threshold	—	3.0
Output	RSSI clamp min	—	−200.0 dBm
	RSSI clamp max	—	+10.0 dBm

A.2 Stage 1: RSSI Estimation

Stage 1 employs the ExactHybrid model for RSSI estimation from baseline-corrected, sign-oriented delta observables (ΔAGC, $\Delta C/N_0$), with post-hoc group-wise affine calibration applied after model training. Table A.2 presents the hyperparameters for training and cross-validation.

The cross-validation grid search selects the best `top_q` and `mono_weight` combination based on validation MAE; the values shown in the "Default" column of Table A.2 are the search candidates, not fixed final values. The L-BFGS fine-tuning phase refines the model after Adam optimization completes, followed by post-hoc group-wise affine calibration on the combined train+validation predictions.

A.3 Stage 2: Centralized Training

Table A.3 presents the hyperparameters for centralized APBM training.

Table A.3: Stage 2 centralized training hyperparameters.

Category	Parameter	Value
Data Split	Training ratio	0.70
	Validation ratio	0.15
	Test ratio	0.15
Model Architecture	Hidden layers	[512, 256, 128, 64, 1]
	Activation function	Leaky ReLU
	Dropout rate	0.2
	Physics bias ($w_{\text{PL}}/w_{\text{NN}}$ init)	2.0
	Position noise std	3.0 m
Training	Batch size	32
	Max epochs	200
Learning Rates	θ (position)	0.015
	P_0 (reference power)	0.005
	γ (path-loss exponent)	0.005
	Neural network	0.001
Warmup Phase	Warmup epochs	30
	θ LR (warmup)	0.01
	P_0, γ LR (warmup)	0.004
Regularization	Weight decay	10^{-5}
	Gradient clipping	1.0
Physics Regularization	γ regularization weight	0.001
	γ regularization target	γ_{init} (env-specific)
	P_0 regularization weight	0.0001
	P_0 regularization target	$P_{0,\text{init}}$ (env-specific)
Early Stopping	Patience	120 epochs
	Min delta	0.005

A.4 Stage 2: Federated Learning

A.4.1 Common FL Parameters

Table A.4 presents hyperparameters common to all federated learning algorithms.

The available partitioning strategies are: `random`, `geographic`, `signal_strength`, `device`, and `distance`.

Table A.4: Common federated learning hyperparameters.

Parameter	Value
Client Configuration	
Default number of clients	5
Min samples per client	10
Training	
Local epochs	3
Base learning rate	0.005
Learning rate decay	0.995/round
Warmup rounds	5
Early Stopping	
Enabled	True
Min delta	0.1
Divergence threshold	3.0×
Consecutive divergence required	2 rounds
Aggregation	
θ aggregation method	Geometric median
NN weights aggregation	Weighted average

A.4.2 Algorithm-Specific Parameters

Table A.5 presents algorithm-specific hyperparameters.

Table A.5: Algorithm-specific federated learning hyperparameters.

Algorithm	Parameter	Value
FedAvg	(No additional parameters)	—
FedProx	Proximal term weight (μ)	0.01
SCAFFOLD	θ LR multiplier	0.2–0.7 (env-specific)
	Physics LR multiplier (P_0, γ)	0.5
	Neural network LR multiplier	0.08–0.1
	Fusion weights LR multiplier	follows NN multiplier
	SGD momentum	0.0
	Nesterov momentum	False
	Local epochs multiplier	1.0

Table A.6: FL tuning profile: Random (IID) partitioning.

Parameter	Default	Urban	Suburban	Open Sky	Lab Wired
Global rounds	120	120	120	120	120
Early stopping patience	15	15	15	15	15
SCAFFOLD θ LR mult	0.4	0.2	0.4	0.2	0.2
SCAFFOLD physics LR mult	0.5	0.5	0.5	0.5	0.5
SCAFFOLD NN LR mult	0.1	0.1	0.1	0.1	0.1

Table A.7: FL tuning profile: Geographic partitioning.

Parameter	Default	Urban	Suburban	Open Sky	Lab Wired
Global rounds	160	160	160	160	160
Early stopping patience	20	20	20	20	20
SCAFFOLD θ LR mult	0.2	0.2	0.1	0.2	0.2
SCAFFOLD physics LR mult	0.5	0.5	0.5	0.5	0.5
SCAFFOLD NN LR mult	0.1	0.1	0.1	0.1	0.1

Table A.8: FL tuning profile: Signal-strength partitioning.

Parameter	Default	Urban	Suburban	Open Sky	Lab Wired
Global rounds	170	170	170	170	170
Base LR	0.0045	0.0045	0.0045	0.0045	0.0045
Early stopping patience	25	25	25	25	25
SCAFFOLD θ LR mult	0.6	0.2	0.2	0.4	0.2
SCAFFOLD physics LR mult	0.5	0.5	0.5	0.5	0.5
SCAFFOLD NN LR mult	0.08	0.08	0.08	0.08	0.08

A.4.3 Auto-Tuned FL Profiles

The framework implements automatic hyperparameter tuning based on environment and partitioning strategy combinations. Tables A.6–A.10 present the complete FL tuning profiles.

A.4.4 FL Profile Design

The auto-tuned FL profiles were designed according to a partitioning-aware trade-off between training length and early-stopping aggressiveness. Under **Random (IID)** partitioning, training is kept shortest (120 rounds with patience 15) because client updates are already well aligned and SCAFFOLD's variance reduction offers

Table A.9: FL tuning profile: Distance partitioning.

Parameter	Default	Urban	Suburban	Open Sky	Lab Wired
Global rounds	180	180	180	180	180
Early stopping patience	30	30	30	30	30
SCAFFOLD θ LR mult	—	0.2	0.2	0.3	0.6
SCAFFOLD physics LR mult	0.5	0.5	0.5	0.5	0.5
SCAFFOLD NN LR mult	0.1	0.1	0.1	0.1	0.1

Table A.10: FL tuning profile: Device partitioning.

Parameter	Default	Urban	Suburban	Open Sky	Lab Wired
Global rounds	170	170	170	170	170
Early stopping patience	25	25	25	25	25
SCAFFOLD θ LR mult	0.7	0.4	0.2	0.2	0.2
SCAFFOLD physics LR mult	0.5	0.5	0.5	0.5	0.5
SCAFFOLD NN LR mult	0.1	0.1	0.1	0.1	0.1

limited additional benefit over FedAvg in an IID regime. In contrast, **Distance-based** partitioning is assigned the longest schedule (180 rounds with patience 30) since quantile splits by distance induce the strongest non-IID behavior, and SCAFFOLD typically needs more rounds for its control variates to stabilize and consistently counter client drift. Finally, **Device** and **Signal-strength** partitioning use intermediate settings (170 rounds with patience 25), reflecting their moderate heterogeneity: they are more non-IID than Random but usually less destabilizing than the distance-quantile split.

The SCAFFOLD θ learning rate multiplier is reduced in low-noise environments (Suburban, Open Sky) to prevent oscillation, while higher values are used in challenging environments (Lab Wired with distance partitioning) where aggressive position updates help escape local minima.

The physics parameters (P_0, γ) use a consistent 0.5× multiplier across all configurations to maintain stable path-loss estimation. The neural network uses 0.08–0.1× to ensure the NN correction term adapts slowly and does not overwhelm the physics-based signal.

A.5 Ablation Study Parameters

Table A.11 presents the parameters used in ablation experiments.

Table A.11: Ablation study experimental parameters.

Study	**Parameter**	**Value**
Common	Random seed	42
	Coordinate frame	Neutral (receiver centroid)
	$\boldsymbol{\theta}$ initialization	Receiver centroid
	Batch size	32
RSSI Source	Oracle	Ground-truth RSSI
	Predicted	Stage 1 calibrated output
	Noisy conditions	2, 5, 10 dB additive Gaussian (σ)
	Shuffled	Random permutation (spatial structure destroyed)
	Constant	Sample-mean RSSI for all receivers
Asymmetric Subsampling	Trigger	Predicted < Oracle (APBM)
	Target centroid offset	30 m from jammer
	Removal strategy	Furthest points first from selected quadrant
	Minimum samples retained	50% of original
Model Architecture	Pure PL	Joint $\boldsymbol{\theta}$, γ, P_0 optimization (SciPy)
	Pure NN	Position regression only (no physics)
	APBM	Full hybrid model (`Net_augmented`)
Pure NN / Pure PL	Train / Val split	70% / 15% (remaining 15% unused)
	Random initializations (n_{inits})	3 (best selected)
	Init perturbation radius	∼30% of spatial spread
APBM Training	Epochs / Patience	200 / 120
	Learning rates	$\boldsymbol{\theta}$: 0.015; P_0/γ: 0.005; NN: 0.001
	Physics bias	2.0
	Warmup epochs	30 (physics-only)
γ/P_0 Init.	Urban	$\gamma = 3.5$, $P_0 = -35.0$ dBm
	Suburban	$\gamma = 2.5$, $P_0 = -32.0$ dBm
	Open Sky	$\gamma = 2.0$, $P_0 = -30.0$ dBm
	Lab Wired	$\gamma = 2.2$, $P_0 = -28.0$ dBm

Appendix B

Complete Experimental Results

This appendix presents the complete experimental results for all environments, partitioning strategies, and federated learning algorithms evaluated in this thesis.

B.1 Stage 1: RSSI Estimation Performance

Table B.1 summarizes the Stage 1 RSSI estimation performance across all four environments.

Table B.1: Stage 1 RSSI estimation performance by environment.

Env.	**Samples**	**MAE (dB)**	**RMSE (dB)**	**RSSI R^2**	**Correlation**
Urban	3,232	4.765	6.383	0.641	0.735
Suburban	810	3.286	4.605	0.957	0.876
Open Sky	821	2.755	3.752	0.981	0.802
Lab Wired	839	3.026	4.536	0.981	0.780

B.2 Stage 2: Localization Results

The following tables present the complete Stage 2 localization error (in meters) for each combination of environment, partitioning strategy, and training method. Bold values indicate the best federated learning result for each configuration; underlined values indicate overall best (including centralized).

B.2.1 Urban Environment

Table B.2: Urban environment localization error (meters). Centralized baseline: 0.75 m.

Partition	Centralized	FedAvg	FedProx	SCAFFOLD
Signal Strength	0.75	1.79	1.93	**1.32**
Random	0.75	1.95	**1.92**	2.09
Distance	0.75	2.05	**2.02**	**2.02**
Geographic	0.75	2.33	2.54	**1.02**
Device	0.75	2.24	2.32	**2.18**
Average FL	—	2.07	2.15	**1.73**

Key observations: SCAFFOLD achieves the best FL performance in 4 of 5 partitioning strategies. The geographic partition with SCAFFOLD (1.02 m) comes closest to the centralized baseline. The dense urban environment with 3,232 samples provides sufficient data for accurate localization.

B.2.2 Lab Wired Environment

Table B.3: Lab Wired environment localization error (meters). Centralized baseline: 11.29 m.

Partition	Centralized	FedAvg	FedProx	SCAFFOLD
Signal Strength	11.29	10.41	10.59	**9.72**
Random	11.29	11.31	11.49	**10.78**
Distance	11.29	11.13	11.30	**10.03**
Geographic	11.29	11.79	12.30	**11.37**
Device	11.29	12.10	12.14	**11.04**
Average FL	—	11.35	11.56	**10.59**

Key observations: SCAFFOLD consistently outperforms other FL algorithms across all partitioning strategies, achieving a 6% average improvement over centralized training. The signal-strength partition yields the best overall result (9.72 m), demonstrating that FL can exceed centralized performance in challenging indoor environments. High path-loss R^2 (0.885) in this controlled environment enables effective physics-based modeling.

B.2.3 Suburban Environment

Table B.4: Suburban environment localization error (meters). Centralized baseline: 2.17 m.

Partition	Centralized	FedAvg	FedProx	SCAFFOLD
Signal Strength	2.17	**1.72**	1.81	1.82
Random	2.17	2.28	2.37	**2.32**
Distance	2.17	1.71	**1.68**	1.81
Geographic	2.17	**1.41**	1.44	1.80
Device	2.17	1.87	1.89	**1.83**
Average FL	—	**1.80**	1.84	1.92

Key observations: In contrast to other environments, FedAvg performs best on average, achieving 17% improvement over centralized training. Geographic partitioning with FedAvg (1.41 m) and distance partitioning with FedProx (1.68 m) both significantly outperform the centralized baseline. This suggests that in moderate-complexity environments with lower data heterogeneity, simpler FL algorithms may be preferable.

B.2.4 Open Sky Environment

Table B.5: Open Sky environment localization error (meters). Centralized baseline: 0.91 m.

Partition	Centralized	FedAvg	FedProx	SCAFFOLD
Signal Strength	0.91	2.44	2.48	**2.68**
Random	0.91	2.42	2.39	**2.27**
Distance	0.91	2.03	2.05	**1.85**
Geographic	0.91	2.80	2.83	**2.72**
Device	0.91	2.02	**1.26**	1.45
Average FL	—	2.34	2.20	**2.19**

Key observations: The Open Sky environment proves most challenging for FL, with all algorithms showing degradation from the centralized baseline. Device-based partitioning with FedProx (1.26 m) achieves the best FL result. The ideal propagation conditions (R^2=0.700) favor centralized training where all spatial information is available simultaneously.

B.3 Cross-Environment Summary

Table B.6 summarizes the win rates and average performance of each algorithm across all 20 experimental configurations (4 environments × 5 partitioning strategies).

Table B.6: Algorithm performance summary across all configurations.

Metric	FedAvg	FedProx	SCAFFOLD
Win count (best FL)	4/20	4/20	12/20
Win rate	20%	20%	60%
Beats centralized	8/20	8/20	10/20
Avg. error (Urban)	2.07	2.15	1.73
Avg. error (Suburban)	1.80	1.84	1.92
Avg. error (Open Sky)	2.34	2.20	2.19
Avg. error (Lab Wired)	11.35	11.56	10.59

Table B.7 presents results organized by partitioning strategy.

Table B.7: Best algorithm by partitioning strategy (wins across 4 environments).

Partition Strategy	Best Algorithm	Win Count	Avg. Improvement
Signal Strength	SCAFFOLD	3/4	−2.1%
Random	Mixed	—	−8.3%
Distance	SCAFFOLD	3/4	−1.5%
Geographic	SCAFFOLD	3/4	+5.2%
Device	SCAFFOLD	3/4	+1.8%

Note: Negative improvement indicates FL performs worse than centralized; positive indicates FL outperforms centralized.

B.4 Convergence Statistics

Table B.8 presents the average number of rounds to convergence (via early stopping or maximum rounds) for each algorithm.
Key observation: SCAFFOLD requires approximately 3× more communication rounds than FedAvg/FedProx due to its slower but steadier convergence. This tradeoff should be considered in communication-constrained deployments.

Table B.8: Average rounds to convergence by algorithm and environment.

Environment	Max Rounds	FedAvg	FedProx	SCAFFOLD
Urban	120–180	41	43	134
Suburban	120–180	38	39	106
Open Sky	120–180	33	38	141
Lab Wired	120–180	55	49	139
Overall Avg.	—	42	42	130

B.5 Ablation Study Results

B.5.1 RSSI Source Ablation

Tables B.9–B.15 present the full RSSI source ablation results across all environments and model architectures. Each entry reports localization error in meters, with the ratio relative to oracle performance shown in parentheses.

RSSI-Essential Environments

Table B.9: RSSI source ablation — Urban (N = 3,232, centroid offset 3.9 m). Localization error in meters; parentheses show ratio vs. oracle.

RSSI Condition	Pure PL	Pure NN	APBM
Oracle	0.52 (1.0×)	43.65 (1.0×)	0.46 (1.0×)
Predicted	0.65 (1.3×)	60.03 (1.4×)	0.76 (1.7×)
Noisy 2 dB	0.60 (1.2×)	47.66 (1.1×)	0.45 (1.0×)
Noisy 5 dB	0.80 (1.5×)	37.82 (0.9×)	0.85 (1.8×)
Noisy 10 dB	0.95 (1.8×)	13.46 (0.3×)	1.36 (2.9×)
Shuffled	17.01 (32.9×)	1.30 (0.03×)	6.60 (14.3×)
Constant	11.42 (22.1×)	6.78 (0.2×)	17.33 (37.7×)

Geometry-Dominated Environments: Symmetric Placement

In the original Open Sky and Suburban datasets, receivers are distributed nearly symmetrically around the jammer, placing the centroid within a few meters of the true position. Under this geometry-dominated regime, all models achieve low errors regardless of RSSI condition, and predicted RSSI outperforms oracle for both Pure PL and APBM due to Stage 1's denoising effect.

Table B.10: RSSI source ablation — Lab Wired ($N = 839$, centroid offset 10.4 m). Localization error in meters; parentheses show ratio vs. oracle.

RSSI Condition	Pure PL	Pure NN	APBM
Oracle	0.12 (1.0×)	1.41 (1.0×)	3.22 (1.0×)
Predicted	1.17 (9.8×)	1.41 (1.0×)	11.41 (3.5×)
Noisy 2 dB	0.33 (2.7×)	1.41 (1.0×)	6.83 (2.1×)
Noisy 5 dB	0.35 (2.9×)	1.41 (1.0×)	7.57 (2.3×)
Noisy 10 dB	0.35 (3.0×)	1.41 (1.0×)	6.91 (2.1×)
Shuffled	13.29 (111×)	1.41 (1.0×)	11.10 (3.4×)
Constant	18.89 (158×)	1.41 (1.0×)	19.88 (6.2×)

Table B.11: RSSI source ablation — Open Sky, symmetric placement ($N = 821$, centroid offset 1.1 m). Predicted outperforms oracle due to geometry dominance.

RSSI Condition	Pure PL	Pure NN	APBM
Oracle	4.87 (1.0×)	6.46 (1.0×)	3.39 (1.0×)
Predicted	2.80 (0.6×)	6.46 (1.0×)	0.99 (0.3×)
Noisy 2 dB	1.47 (0.3×)	6.46 (1.0×)	3.00 (0.9×)
Noisy 5 dB	5.96 (1.2×)	6.46 (1.0×)	1.88 (0.6×)
Noisy 10 dB	8.34 (1.7×)	6.45 (1.0×)	2.38 (0.7×)
Shuffled	1.69 (0.3×)	6.46 (1.0×)	1.92 (0.6×)
Constant	3.48 (0.7×)	6.46 (1.0×)	2.79 (0.8×)

Table B.12: RSSI source ablation — Suburban, symmetric placement ($N = 810$, centroid offset 4.4 m). Predicted outperforms oracle due to geometry dominance.

RSSI Condition	Pure PL	Pure NN	APBM
Oracle	6.30 (1.0×)	7.37 (1.0×)	3.52 (1.0×)
Predicted	1.37 (0.2×)	7.36 (1.0×)	2.43 (0.7×)
Noisy 2 dB	6.20 (1.0×)	7.37 (1.0×)	1.47 (0.4×)
Noisy 5 dB	5.72 (0.9×)	7.37 (1.0×)	1.88 (0.5×)
Noisy 10 dB	4.36 (0.7×)	7.37 (1.0×)	1.73 (0.5×)
Shuffled	4.70 (0.7×)	7.37 (1.0×)	3.14 (0.9×)
Constant	5.10 (0.8×)	7.36 (1.0×)	3.65 (1.0×)

Geometry-Dominated Environments: Asymmetric Subsampling

To expose the underlying RSSI dependence masked by symmetric placement, graduated asymmetric subsampling removes receivers from one quadrant until the centroid is shifted ∼30 m from the jammer. This controlled asymmetry reveals clear RSSI sensitivity in Pure PL and APBM.

Table B.13: RSSI source ablation — Open Sky, asymmetric subsampling ($N = 736$, 85 pts from NE removed, centroid offset 30.2 m). RSSI dependence now clearly visible.

RSSI Condition	Pure PL	Pure NN	APBM
Oracle	3.58 (1.0×)	30.63 (1.0×)	17.01 (1.0×)
Predicted	20.62 (5.8×)	32.75 (1.1×)	41.61 (2.4×)
Noisy 2 dB	3.79 (1.1×)	34.09 (1.1×)	26.27 (1.5×)
Noisy 5 dB	6.31 (1.8×)	33.25 (1.1×)	25.41 (1.5×)
Noisy 10 dB	16.59 (4.6×)	30.78 (1.0×)	28.14 (1.7×)
Shuffled	35.03 (9.8×)	26.20 (0.9×)	29.44 (1.7×)
Constant	41.01 (11.5×)	26.29 (0.9×)	38.03 (2.2×)

Table B.14: RSSI source ablation — Suburban, asymmetric subsampling ($N = 719$, 91 pts from NW removed, centroid offset 30.1 m). RSSI dependence now clearly visible.

RSSI Condition	Pure PL	Pure NN	APBM
Oracle	0.96 (1.0×)	13.07 (1.0×)	21.61 (1.0×)
Predicted	7.36 (7.7×)	23.48 (1.8×)	23.31 (1.1×)
Noisy 2 dB	3.60 (3.7×)	13.07 (1.0×)	22.40 (1.0×)
Noisy 5 dB	10.41 (10.8×)	15.29 (1.2×)	22.33 (1.0×)
Noisy 10 dB	13.17 (13.7×)	13.07 (1.0×)	22.12 (1.0×)
Shuffled	22.53 (23.5×)	12.29 (0.9×)	25.20 (1.2×)
Constant	21.85 (22.8×)	12.50 (1.0×)	26.72 (1.2×)

Cross-Environment Summary

Table B.15 provides a condensed cross-environment comparison of key RSSI ablation ratios for the APBM and Pure PL models. Results for Open Sky and Suburban use the asymmetric subsampling configuration.

Key findings: Overall, RSSI quality is a primary driver of localization performance: when the spatial RSSI structure is destroyed, errors increase dramatically—by

Table B.15: RSSI ablation summary: key ratios vs. oracle across environments. Open Sky and Suburban values are from asymmetric subsampling experiments (centroid offset ∼30 m).

Model	Ratio	Urban	Lab Wired	Open Sky	Suburban
Pure PL	Predicted / Oracle	1.3×	9.8×	5.8×	7.7×
	Shuffled / Oracle	32.9×	111×	9.8×	23.5×
	Constant / Oracle	22.1×	158×	11.5×	22.8×
APBM	Predicted / Oracle	1.7×	3.5×	2.4×	1.1×
	Shuffled / Oracle	14.3×	3.4×	1.7×	1.2×
	Constant / Oracle	37.7×	6.2×	2.2×	1.2×

roughly 10–158× for Pure PL and 1.2–38× for APBM across environments. The ablation also reveals two operating regimes: under near-symmetric receiver placement (notably Open Sky and Suburban), geometry dominates and RSSI content has limited effect, whereas progressively enforcing asymmetric subsampling restores RSSI dependence for all model families. In contrast, Pure NN behaves largely RSSI-invariant, exhibiting nearly constant error across conditions (e.g., 1.41 m in Lab Wired and 6.46 m in symmetric Open Sky), consistent with convergence to geometry-driven attractors. Finally, APBM is systematically more robust than Pure PL under RSSI corruption; its smaller degradation ratios (e.g., Urban shuffled 14.3× vs. 32.9×) indicate that the neural branch compensates for a substantial fraction of the lost RSSI structure (approximately 57–97%).

B.6 Environment Characteristics

Table B.16 summarizes the key characteristics of each experimental environment.

Table B.16: Environment characteristics and dataset statistics.

Property	**Urban**	**Suburban**	**Open Sky**	**Lab Wired**
Location	Politecnico	Venaria Reale	Parco Mandria	Indoor Lab
Description	Dense urban	Residential	Open park	Controlled
Total samples	5,003	1,230	1,250	1,248
Jammed samples	3,232	810	821	839
Train/Val/Test	2262/485/485	567/121/122	574/123/124	587/126/126
γ (path loss)	3.5	2.5	2.0	2.2
P_0 (ref power)	-35.0 dBm	-32.0 dBm	-30.0 dBm	-28.0 dBm
Path-loss R^2	0.600	0.644	0.700	0.885
RSSI range (dB)	$[-126, -49]$	$[-119, -74]$	$[-93, -67]$	$[-109, -50]$
Position noise σ	3.0 m	3.0 m	3.0 m	3.0 m
Devices	11	5	5	3–6

References

[1] Mariona Jaramillo-Civill, Luis González-Gudiño, Tales Imbiriba, and Pau Closas. *Bayesian Jammer Localization with a Hybrid CNN and Path-Loss Mixture of Experts*. Oct. 2025. arXiv: 2510.20666 [eess.SP]. URL: https://arxiv.org/abs/2510.20666 (cit. on pp. 1–5, 20).

[2] Emile Ghizzo, El Mehdi Djelloul, Julien Lesouple, Carl Milner, and Christophe Macabiau. «Assessing jamming and spoofing impacts on GNSS receivers: Automatic gain control (AGC)». In: *Signal Processing* 228 (Mar. 2025). ISSN: 01651684. DOI: 10.1016/j.sigpro.2024.109762 (cit. on pp. 1, 10).

[3] Zhe Yan and Laura Ruotsalainen. «GNSS jammer localization in urban areas based on prediction/optimization and ray-tracing». In: *GPS Solutions* 29 (1 Jan. 2025). ISSN: 15211886. DOI: 10.1007/s10291-024-01787-4 (cit. on pp. 1, 19).

[4] Mohammad Zahidul H. Bhuiyan, Heidi Kuusniemi, Stefan Söderholm, and Esa Airos. «The Impact of Interference on GNSS Receiver Observables – A Running Digital Sum Based Simple Jammer Detector». In: *Radioengineering* 23.3 (2014), pp. 898–906 (cit. on p. 1).

[5] Dongliang Lyu, Xin Chen, Fei Wen, Ling Pei, and Di He. «Urban area GNSS in-car-jammer localization based on pattern recognition». In: *Navigation, Journal of the Institute of Navigation* 66 (2 June 2019), pp. 325–340. ISSN: 00281522. DOI: 10.1002/navi.301 (cit. on pp. 1, 19).

[6] Glädje Karl Olsson, Sara Nilsson, Erik Axell, Erik G. Larsson, and Panos Papadimitratos. «Using Mobile Phones for Participatory Detection and Localization of a GNSS Jammer». In: *2023 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. Monterey, CA, USA, Apr. 2023, pp. 536–541. DOI: 10.1109/PLANS53410.2023.10140088 (cit. on pp. 1, 2, 18).

[7] Polona Pavlovčič-Prešeren, Franc Dimc, and Matej Bažec. «Exploiting the Sensitivity of Dual-Frequency Smartphones and GNSS Geodetic Receivers for Jammer Localization». In: *Remote Sensing* 15 (4 Feb. 2023). ISSN: 20724292. DOI: 10.3390/rs15041157 (cit. on p. 1).

[8] Mariona Jaramillo-Civill, Peng Wu, Andrea Nardin, Tales Imbiriba, and Pau Closas. «Jammer Source Localization with Federated Learning». In: *2025 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. Salt Lake City, UT, USA: IEEE, Apr. 2025, pp. 362–371. DOI: 10.1109/PLANS61210.2025.11028278 (cit. on pp. 1, 3, 5, 20, 36).

[9] Chao Han et al. *Crowdsourced Smartphone-Based Machine Learning for GNSS Jammer Detection and Localization*. Working paper. SSRN, 2025. URL: https://ssrn.com/abstract=5119211 (cit. on pp. 2, 4, 6, 13, 15).

[10] Glädje Karl Olsson, Erik Axell, Erik G. Larsson, and Panos Papadimitratos. «Participatory Sensing for Localization of a GNSS Jammer». In: *2022 International Conference on Localization and GNSS (ICL-GNSS)*. Tampere, Finland, June 2022, pp. 1–7. DOI: 10.1109/ICL-GNSS54081.2022.9797031 (cit. on pp. 2, 11, 180).

[11] Nathan Stephen Levigne. «Automatic Gain Control Measurements as a GPS L1 Interference Detection Metric». Accessed 2026-02-23. Master's thesis. University of Colorado Boulder, 2019. URL: https://scholar.colorado.edu/concern/graduate_thesis_or_dissertations/qr46r104k (cit. on pp. 2, 9–12).

[12] Luka Strizic, Dennis M. Akos, and Sherman Lo. «Crowdsourcing GNSS Jammer Detection and Localization». In: *Proceedings of the 2018 International Technical Meeting of The Institute of Navigation (ION ITM 2018)*. Reston, Virginia, USA, Jan. 2018, pp. 626–641. DOI: 10.33012/2018.15546. URL: https://www.ion.org/publications/abstract.cfm?articleID=15546 (cit. on pp. 2, 10, 11, 15, 18).

[13] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank Reddi, Sebastian Stich, and Ananda Theertha Suresh. «SCAFFOLD: Stochastic Controlled Averaging for Federated Learning». In: *Proceedings of the 37th International Conference on Machine Learning*. Ed. by Hal Daumé III and Aarti Singh. Vol. 119. Proceedings of Machine Learning Research. PMLR, July 2020, pp. 5132–5143. URL: https://proceedings.mlr.press/v119/karimireddy20a.html (cit. on pp. 3, 6, 24, 25, 44).

[14] Andrea Nardin, Tales Imbiriba, and Pau Closas. «Crowdsourced Jammer Localization Using APBMs: Performance Analysis Considering Observations Disruption». In: *2023 IEEE/ION Position, Location and Navigation Symposium, PLANS 2023*. Institute of Electrical and Electronics Engineers Inc., 2023, pp. 511–519. ISBN: 9781665417723. DOI: 10.1109/PLANS53410.2023.10140023 (cit. on pp. 4, 77, 81, 83, 180).

[15] H. Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. «Communication-Efficient Learning of Deep Networks from Decentralized Data». In: *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*. Ed. by Aarti Singh and Jerry Zhu. Vol. 54. Proceedings of Machine Learning Research. PMLR, 2017, pp. 1273–1282. URL: https://proceedings.mlr.press/v54/mcmahan17a.html (cit. on pp. 5, 24, 43, 91).

[16] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. *Federated Optimization in Heterogeneous Networks*. 1812.06127. arXiv, 2020. URL: https://arxiv.org/abs/1812.06127 (cit. on pp. 5, 24, 43, 91).

[17] Damian Miralles, Nathan Levigne, Dennis M. Akos, Juan Blanch, and Sherman Lo. «Android raw GNSS measurements as a new anti-spoofing and anti-jamming solution». In: *Proceedings of the 31st International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS+ 2018*. Institute of Navigation, 2018, pp. 334–344. ISBN: 0936406100. DOI: 10.33012/2018.15883 (cit. on pp. 9, 10).

[18] Nicholas Spens, Dong Kyeong Lee, Filip Nedelkov, and Dennis Akos. «Detecting GNSS Jamming and Spoofing on Android Devices». In: *Navigation, Journal of the Institute of Navigation* 69 (3 Sept. 2022). ISSN: 21614296. DOI: 10.33012/navi.537 (cit. on p. 9).

[19] Dong-Kyeong Lee, Nicholas Spens, Benon Gattis, and Dennis Akos. «AGC on Android Devices for GNSS». In: *Proceedings of the 2021 International Technical Meeting of The Institute of Navigation (ION ITM 2021)*. Jan. 2021, pp. 33–41. DOI: 10.33012/2021.17823. URL: https://www.ion.org/publications/abstract.cfm?articleID=17823 (cit. on pp. 10–12).

[20] John W. Betz. *Effect of Partial-Band Interference on Receiver Estimation of C/N_0: Theory*. Tech. rep. Also published in ION NTM 2001 (Long Beach, CA), pp. 817–828. The MITRE Corporation, Jan. 2001 (cit. on p. 13).

[21] Daniele Borio, Ciro Gioia, Andrej Štern, Franc Dimc, and Gianmarco Baldini. «Jammer localization: From crowdsourcing to synthetic detection». In: *29th International Technical Meeting of the Satellite Division of the Institute of Navigation, ION GNSS 2016*. Vol. 5. Institute of Navigation, 2016, pp. 3107–3116. ISBN: 9781510834101. DOI: 10.33012/2016.14689 (cit. on p. 13).

[22] François Bastide, Dennis Akos, Christophe Macabiau, and Benoît Roturier. «Automatic Gain Control (AGC) as an Interference Assessment Tool». In: *Proceedings of the 16th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GPS/GNSS 2003)*. Portland, OR, USA, Sept. 2003, pp. 2042–2053 (cit. on p. 14).

[23] Theodore S Rappaport. *Wireless Communications: Principles and Practice*. 2nd ed. Upper Saddle River, NJ: Prentice Hall PTR, 2002. ISBN: 978-0130422323 (cit. on pp. 17, 176).

[24] Simon Kocher, Jonathan Hansen, and Alexander Rügamer. «GNSS Interference Localization for Vehicular Jammers using low-cost COTS Sensors». In: *2024 International Conference on Localization and GNSS, ICL-GNSS 2024 - Proceedings*. Institute of Electrical and Electronics Engineers Inc., 2024. ISBN: 9798350380781. DOI: 10.1109/ICL-GNSS60721.2024.10578445 (cit. on p. 18).

[25] Ryan C. Blay and Dennis M. Akos. «GNSS RFI localization using a hybrid TDOA/PDOA approach». In: *Proceedings of the 2018 International Technical Meeting of The Institute of Navigation, ITM 2018*. Vol. 2018-January. Institute of Navigation, 2018, pp. 703–712. DOI: 10.33012/2018.15554 (cit. on p. 18).

[26] Dania Herzalla, Willian T. Lunardi, and Martin Andreoni. «Graph Neural Networks for Jamming Source Localization». In: *Machine Learning and Knowledge Discovery in Databases. Research Track and Applied Data Science Track (ECML PKDD 2025)*. Vol. 16020. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2026, pp. 331–348. DOI: 10.1007/978-3-662-72243-5_19. URL: https://link.springer.com/chapter/10.1007/978-3-662-72243-5_19 (cit. on pp. 18, 19).

[27] Andrea Nardin, Tales Imbiriba, and Pau Closas. «Jamming Source Localization Using Augmented Physics-Based Model». In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2023, pp. 1–5. DOI: 10.1109/ICASSP49357.2023.10095731 (cit. on pp. 18, 19, 180).

[28] Andrea Nardin. «Innovative Signal Processing Solutions for Next-Generation Satellite Navigation Systems». Ph.D. dissertation. Torino, Italy: Politecnico di Torino, 2023. URL: https://hdl.handle.net/11583/2979890 (cit. on pp. 19, 37).

[29] Diederik P. Kingma and Jimmy Ba. «Adam: A Method for Stochastic Optimization». In: *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7–9, 2015, Conference Track Proceedings*. Ed. by Yoshua Bengio and Yann LeCun. 2015. URL: http://arxiv.org/abs/1412.6980 (cit. on p. 22).

[30] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E. Hinton. «Layer Normalization». In: *NIPS 2016 Deep Learning Symposium*. OpenReview: NIPS 2016 Deep Learning Symposium. Aug. 2016. URL: https://openreview.net/forum?id=BJLa_ZC9 (cit. on pp. 23, 38).

[31] Krishna Pillutla, Sham M. Kakade, and Zaid Harchaoui. «Robust Aggregation for Federated Learning». In: *IEEE Transactions on Signal Processing* 70 (2022), pp. 1142–1154. DOI: 10.1109/TSP.2022.3153135 (cit. on pp. 25, 45).

[32] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. URL: https://www.deeplearningbook.org (cit. on p. 27).

[33] Umberto Robustelli, Jacek Paziewski, and Giovanni Pugliano. «Observation Quality Assessment and Performance of GNSS Standalone Positioning with Code Pseudoranges of Dual-Frequency Android Smartphones». In: *Sensors* 21.6 (2021), p. 2125. DOI: 10.3390/s21062125 (cit. on pp. 61, 62, 69).

[34] Simon Banville and Frank Van Diggelen. «Precise GNSS for Everyone: Precise Positioning Using Raw GPS Measurements from Android Smartphones». In: *GPS World* 27.11 (2016), pp. 43–48 (cit. on pp. 61, 62, 69).

[35] Richard B Langley. «GPS receiver system noise». In: *GPS World* 8.6 (1997), pp. 40–45 (cit. on p. 63).

[36] R. H. Mitch, R. C. Dougherty, M. L. Psiaki, S. P. Powell, B. W. O'Hanlon, J. A. Bhatti, and T. E. Humphreys. «Signal Characteristics of Civil GPS Jammers». In: *Proceedings of the 24th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2011)*. Portland, OR, USA, 2011, pp. 1907–1919 (cit. on p. 68).

[37] Android Developers. *Detect GNSS jamming and spoofing*. Accessed 2026-02-23. 2026. URL: https://developer.android.com/develop/sensors-and-location/sensors/gnss-spoof-jam (cit. on p. 180).

[38] Apple. *Core Location*. Accessed 2026-02-23. 2026. URL: https://developer.apple.com/documentation/corelocation (cit. on p. 180).

[39] Google Developers. *Fused Location Provider API*. Accessed 2026-02-23. 2026. URL: https://developers.google.com/location-context/fused-location-provider (cit. on p. 180).

[40] Joohan Chun, Jacob Spagnolli, Tanner Holmes, and Dennis Akos. «Assessment of Android Network Positioning as an Alternate Source for Robust PNT». In: *Sensors* 25.23 (2025), p. 7324. DOI: 10.3390/s25237324. URL: https://www.mdpi.com/1424-8220/25/23/7324 (cit. on p. 180).

[41] Zhuojian Cao, Jiang Liu, Wei Jiang, and Baigen Cai. «INS-aided GNSS jamming protection in support of resilient train positioning». In: *High-speed Railway* 3 (2025), pp. 185–193. DOI: 10.1016/j.hspr.2025.05.004. URL: https://journal.hep.com.cn/hspr/EN/10.1016/j.hspr.2025.05.004 (cit. on p. 180).

[42] Sriramya Bhamidipati and Grace Xingxin Gao. «Simultaneous Localization of Multiple Jammers and Receivers Using Probability Hypothesis Density». In: *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. 2018, pp. 940–944. DOI: 10.1109/PLANS.2018.8373472 (cit. on p. 180).

[43] Shaghayegh Shahcheraghi, Justin Kuric, and Zaher M. Kassas. «Opportunistic Positioning with Beamformed 5G Signals». In: *Proceedings of the 37th International Technical Meeting of the Satellite Division of the Institute of Navigation (ION GNSS+ 2024)*. Preprint of the ION GNSS+ 2024 Conference (as stated in the manuscript). Baltimore, Maryland, USA, Sept. 2024. URL: https://people.engineering.osu.edu/media/document/2024-10-12/kassas_opportunistic_positioning_with_beamformed_5g_signals.pdf (cit. on p. 180).

[44] ETSI and 3GPP. *5G; NG Radio Access Network (NG-RAN); Stage 2 functional specification of User Equipment (UE) positioning in NG-RAN (3GPP TS 38.305 version 19.0.0 Release 19)*. ETSI TS TS 138 305 V19.0.0. Accessed 2026-02-23. European Telecommunications Standards Institute (ETSI), Oct. 2025. URL: https://www.etsi.org/deliver/etsi_ts/138300_138399/138305/19.00.00_60/ts_138305v190000p.pdf (cit. on p. 180).