

Wireless Datasets for Aerial Networks

Amir Hossein Fahim Raouf, Donggu Lee, Mushfiquir Rahman, Saad Masrur, Gautham Reddy, Cole Dickerson, Md Sharif Hossen, Sergio Vargas Villar, Anil Gürses, Simran Singh, Sung Joon Maeng, Martins Ezuma, Christopher Roberts, Mohamed Rabeek Sarbudeen, Thomas J. Zajkowski, Magreth Mushi, Ozgur Ozdemir, Ram Asokan, Ismail Guvenc, Mihail L. Sichitiu, Rudra Dutta

Abstract—The integration of unmanned aerial vehicles (UAVs) into 5G-Advanced and future 6G networks presents a transformative opportunity for wireless connectivity, enabling agile deployment and improved LoS communications. However, the effective design and optimization of these aerial networks depend critically on high-quality, empirical data. This paper provides a comprehensive survey of publicly available wireless datasets collected from an airborne platform called Aerial Experimentation and Research Platform on Advanced Wireless (AERPAW). We highlight the unique challenges associated with generating reproducible aerial wireless datasets, and review the existing related works in the literature. Subsequently, for each dataset considered, we explain the hardware and software used, present the dataset format, provide representative results, and discuss how these datasets can be used to conduct additional research. The specific aerial wireless datasets presented include raw I/Q samples from a cellular network over different UAV trajectories, spectrum measurements at different altitudes, flying 4G base station (BS), a 5G-NSA Ericsson network, a LoRaWAN network, an radio frequency (RF) sensor network for source localization, wireless propagation data for various scenarios, and comparison of ray tracing and real-world propagation scenarios. References to all datasets and post-processing scripts are provided to enable full reproducibility of the results. Ultimately, we aim to guide the community toward effective dataset utilization for validating propagation models, developing machine learning algorithms, and advancing the next generation of aerial wireless systems.

Index Terms—air-to-ground, AERPAW, C-band, helikite, software-defined radio, spectrum monitoring.

I. INTRODUCTION

As the global demand for seamless and high-capacity wireless connectivity continues to grow, aerial platforms, particularly unmanned aerial vehicles (UAVs) have emerged as a promising complement to terrestrial infrastructure in 5G-Advanced and future 6G networks [1], [2]. UAVs offer unique opportunities for agile deployment, rapid coverage extension, and spectrum monitoring in disaster-stricken or hard-to-reach environments. Their altitude advantage over terrestrial transmitters and receivers enables line-of-sight (LoS) links, supporting applications such as aerial base stations (BSs),

edge caching, environmental monitoring, and emergency response [3]. Advances in sensing, perception, and decision-making have enabled UAVs with autonomous navigation capabilities to operate with minimal human intervention in complex and dynamic environments.

Recognizing this potential, regulatory and standardization bodies have begun laying the groundwork for the integration of aerial users into cellular networks. For instance, the Federal Communications Commission (FCC) has examined the use of licensed mid-band spectrum for UAV command-and-control (C2) links and emphasized the role of 5G in enabling unmanned aircraft systems (UAS) integration within cellular networks [4], [5]. In parallel, the 3rd Generation Partnership Project (3GPP) has introduced UAV-specific enhancements in Release 15 and subsequent releases to address challenges such as interference mitigation, mobility management, and flight-relevant key performance indicators (KPIs) for aerial user equipments (UEs) [6]. The International Telecommunication Union (ITU) has similarly recognized the potential of UAVs in advancing global broadband connectivity and enhancing disaster response capabilities [7]. In addition to these regulatory and standardization efforts, industry stakeholders, e.g., Ericsson, Samsung, and Qualcomm, have submitted joint contributions to 3GPP RAN4 aimed at evaluating interference scenarios between aerial and terrestrial users in the shared spectrum bands, highlighting the critical role of empirical measurements in guiding standards development and system design (e.g., see [8]).

This shift toward aerial connectivity has underscored the growing need for open, well-documented, and reproducible wireless datasets captured from airborne platforms. Such datasets serve as a critical enabler for validating theoretical models, training machine learning algorithms, and guiding system-level simulations and protocol design. In practice, producing these datasets requires capabilities beyond those of terrestrial campaigns. These include support for custom waveform generation, programmable radios, repeatable and controllable mobility patterns, and precise synchronization of location and time for each measurement sample.

To tackle these challenges, platforms such as the Aerial Experimentation and Research Platform for Advanced Wireless (AERPAW) have emerged as critical enablers of aerial wireless research. AERPAW integrates a diverse suite of aerial data collection systems, including drones, helikites, and software-programmable radios, equipped with synchronized GPS modules, wideband radio frequency (RF) sensing capabilities, and Signal Metadata Format (SigMF)-compliant data logging infrastructure. This multi-modal framework fa-

This work has been supported in part by the NSF award CNS-1939334 and CNS-2332835. Amir Hossein Fahim Raouf, Donggu Lee, Mushfiquir Rahman, Saad Masrur, Gautham Reddy, Cole Dickerson, Md Sharif Hossen, Sergio Vargas Villar, Anil Gürses, Simran Singh, Martins Ezuma, Christopher Roberts, Mohamed Rabeek Sarbudeen, Thomas J. Zajkowski, Magreth Mushi, Ozgur Ozdemir, Ismail Guvenc, Mihail L. Sichitiu, and Rudra Dutta are with the Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC. Sung Joon Maeng is with Hanyang University, Ansan 15588, S. Korea. Ram Asokan is with the Wireless Research Center, Wake Forest, NC. (Corresponding author: Amir Hossein Fahim Raouf).

enables reproducible and scalable experimentation across a broad range of wireless scenarios. These efforts have produced high-resolution datasets that span I/Q signals, received power, and commercial network KPIs across a wide altitude range, helping researchers model three-dimensional (3D) propagation environments and evaluate LoS and non-line-of-sight (NLoS) transitions at scale [9].

The remainder of this paper is organized as follows. Section II surveys prior aerial wireless data collection efforts and identifies key gaps and limitations in the existing literature. Section III describes the technical, logistical, and regulatory challenges involved in designing and deploying programmable aerial wireless platforms. Sections V through XIV present a comprehensive set of datasets collected using AERPAW equipment and infrastructure, detailing the measurement configurations, data formats, representative results, and application domains. Finally, Section XV concludes the paper with final remarks and discusses future directions.

II. LITERATURE REVIEW

There have been only some sporadic efforts on systematically capturing and disseminating aerial cellular datasets, with many studies constrained by narrow deployment scenarios, proprietary data formats, or limited reproducibility. For example, Mozy *et al.* conducted an extensive performance evaluation of Long-Term Evolution (LTE) and 5G networks for UAV services, highlighting significant degradation in downlink performance at higher altitudes, yet the corresponding data remains unavailable for public use [10]. Similarly, Braunkopf *et al.* presented controlled drone flight measurements over a commercial LTE network, analyzing KPIs such as Reference Signal Received Power (RSRP), Reference Signal Received Quality (RSRQ), and Signal-to-Interference-plus-Noise Ratio (SINR) as functions of altitude, but without providing standardized or reusable datasets [11].

Other experimental studies have explored software-defined radio (SDR)-enabled aerial testbeds for deploying LTE BSs on UAVs, demonstrating the feasibility of airborne infrastructure for enhanced connectivity [12]. Separately, Ruseno *et al.* analyzed 4G signal quality in the context of UAV Remote ID systems, leveraging machine learning techniques to model signal performance under varying conditions [13]. While these studies contribute valuable architectural and performance insights, they do not offer altitude-resolved, metadata-rich datasets necessary for reproducible research. Furthermore, Zulkifley *et al.* evaluated the feasibility of LTE-based connectivity for small UAVs, showing that increasing altitude leads to substantial degradation in signal quality and increased latency due to non-optimized terrestrial deployments [14]. In a complementary study, Kovács *et al.* performed aerial measurements over live LTE networks and analyzed interference patterns, highlighting the challenges posed by sidelobe reception and elevated interference levels in UAV operations [15].

In parallel, studies on terrestrial and regional spectrum occupancy, such as that by Chennamsetty *et al.*, have demonstrated the utility of passive spectrum monitoring across 4G and 5G bands, though such efforts remain disconnected from the

aerial domain [16]. While most prior efforts have focused on targeted aerial tests or single-purpose deployments, a few have investigated broader spectrum usage trends under real-world operational conditions. For instance, Kuester *et al.* presented a comprehensive study of radio spectrum occupancy during the COVID-19 pandemic, capturing temporal and spatial variations in 4G and 5G usage across multiple urban and suburban locations using passive monitoring equipment [17]. Although not UAV-specific, such work underscores the value of large-scale, reproducible spectrum datasets in understanding wireless dynamics and informs the design of future aerial deployments.

Ensuring future progress in aerial connectivity research requires sustained support for measurement platforms that emphasize openness, documentation, and multi-modal sensing capabilities. The community would benefit from initiatives that align with FAIR data principles (Findable, Accessible, Interoperable, and Reusable) [18], enabling standardized benchmarking and accelerating advancements toward 5G-Advanced and 6G network deployments.

III. CHALLENGES FOR GENERATING DATASETS WITH AERIAL WIRELESS SYSTEMS

There are substantial difficulties in collecting high-quality datasets for aerial wireless systems. In what follows, we will detail a few challenges that were encountered and overcome by the AERPAW team. Some of the challenges are common across all the platforms for advanced wireless research (PAWR), while some are unique to AERPAW.

A. Programmable Radios

Many of the existing datasets available for aerial wireless systems used commercial off-the-shelf (COTS) equipment to collect wireless data; for example, wireless phones, or even a small 4G/5G modem connected to a single-board computer (SBC) like a Raspberry Pi can be used to collect some KPIs like RSRP, RSRQ, and cell ID. However, while light and portable, such a setup can *only* collect those KPIs for those particular networking technologies. The situation is similar for Wi-Fi, LoRa, or other COTS equipment.

In contrast, in AERPAW we decided to build our radio system around some of the best SDRs available, namely the universal software radio peripherals (USRPs) from National Instruments (NI). The main advantage of a USRP setup is that they can impersonate *any* of the technologies that a COTS radio can; furthermore, the USRPs can transmit and receive custom waveforms for which there is no equivalent COTS radio, thus allowing an unprecedented degree of flexibility and programmability, and hence a broad range of wireless experiments.

The main challenge of using USRPs is the relatively large size of the resulting portable node: although the USRPs themselves can be relatively small and light (especially true for the B200 series of USRPs), the *supporting* hardware is large and heavy. In order to drive the USRPs a relatively powerful computer needs to be employed (a seventh-generation Intel NUC in our case), filters and power amplifiers on the front

ends, and even a custom-made GPS-DO for tight frequency and time synchronization. This increases the size of a portable node by almost an order of magnitude (from a few hundred grams for a COTS portable node to 3.5 kg for a B210 node which requires separate front ends for each of its channels).

B. Outdoor Radio Infrastructure and Spectrum

Supporting outdoor experiments with drones and wireless technologies requires access to towers equipped with diverse radio systems and to experimental spectrum bands for testing new waveforms and protocols. While some studies and datasets rely on commercial cellular networks, such approaches limit the scope and flexibility of experiments. The AERPAW platform addresses these challenges through the deployment of five towers, two rooftop sites, and one purpose-built light pole to host USRPs and other commercial off-the-shelf (COTS) wireless devices. All fixed nodes are fiber-connected and dedicated exclusively to AERPAW experimentation. Furthermore, AERPAW is designated as one of the four FCC Innovation Zones in the United States, providing access to specific frequency bands for wireless experimentation with drones [19]. Because airborne transmissions from drones can create significant interference on the ground and with incumbent spectrum users, obtaining experimental frequencies is particularly challenging. AERPAW has secured access to the 900 MHz Industrial, Scientific, and Medical (ISM) band, 1.7/2.1 GHz, and 3.3-3.45 GHz bands to support experiments involving USRPs, commercial 4G/5G equipment, and UAVs, and continues to expand its available experimental spectrum.

C. Programmable Drones

The immediate consequence of having large portable node based on USRPs is that the drones required to fly the portable nodes have to be much larger than a drone designed to carry portable nodes based on COTS UEs. In turn, the larger drones are more expensive, and more difficult to design and implement than smaller drones. In AERPAW, we designed and implemented our large drones from first principles. AERPAW could have used COTS drones (or at least COTS frames) for their drones, but instead chose to design and implement them from readily available materials like carbon fiber plates and carbon fiber tubes. The main advantage of this approach is the reproducibility of the AERPAW frames: all the COTS frames we initially considered (including the DJI Matrix 600) are currently discontinued.

Another important choice for the AERPAW drones is the open-source software stack employed: in the interest of a fully programmable drone, the AERPAW vehicle control software is fully open: the autopilot firmware is ArduPilot [20], and the software is based on the MAVLink open protocol [21]. The ground control station (GCS) used both in development and operations is QGroundControl [22]. This allows for relying on a large base of existing software while developing software that can be reused by other researchers. The software employed by the AERPAW drones allows experimenters both high-level (e.g., preplanned trajectories) as well as low-level (e.g., off-board control) of the AERPAW drones, allowing for highly customized experiments.

D. Development in Digital Twins

A unique requirement for AERPAW, among all other PAWR platforms, is its digital twin (DT). In particular, the use of autonomous vehicles in AERPAW and the safety requirements for these vehicles make the development of vehicle software in the testbed itself a major challenge. Instead, for all canonical experiments, all experimenters have to develop their experiments in a custom-made DT of the physical testbed. The AERPAW DT is deployed in the AERPAW data-compute store, which can host several hundred instances of the DT (the exact number depends on the complexity of the experiment instantiated). In the DT, all the software of the real testbed is preserved while simulating three main hardware components of the real testbed: the frames of the drones, the USRPs, and the propagation between the USRPs. The virtual USRPs operate at I/Q sample level, thus allowing for the development of realistic channel models, including antenna patterns, MIMO radios, reflections, and Doppler shifts. The drone emulation includes a virtual machine running the same firmware as the drones on the autopilots of the drones in the testbed, resulting in identical responses to commands of both the drones in the testbed and in the DT.

The use of AERPAW DT allows experimenters to develop all their radio and drone software fully remotely, at their own pace, without needing to access any radio or drone hardware. Once the experimenters develop and test their software in the AERPAW DT, they can be deployed quickly in the real-world testbed environment. AERPAW supports a large variety of sample experiments that are tested in the DT [23] and can be accessed by experimenters to quickly initiate baseline experiments.

E. Precise Localization and Roll/Yaw/Pitch Information

Finally, AERPAW developed an infrastructure that achieves high levels of precision for capturing high-quality datasets. For example, to achieve centimeter-level accuracy, we have deployed a Real-Time Kinematic (RTK) BS at one of the fixed nodes, and we performed the precise point positioning (PPP) procedure, resulting in an accuracy of a few millimeters for this BS. RTK updates are then fed online for each of out vehicles (drones, rovers, and the helikite), ensuring that all collected geographical information is captured with sub-centimeter accuracy.

Additionally, each fixed node and several portable nodes are equipped with GPS receivers providing both time and frequency corrections to both the USRPs as well as the fixed and portable nodes, allowing for tight time synchronization, which in turn results in testbed-wide synchronized logs. The logs are being generated from multiple sources at each node: the vehicles generate vehicle information (e.g., latitude, longitude, altitude, roll, pitch, yaw, velocities, etc.), the low-level radio software (e.g., srsRAN) generates radio KPIs (e.g., RSRP, RSRQ, I/Q samples), and traffic software generates traffic KPIs (e.g., throughput, delay, error rates). All these logs are time-stamped with a testbed-wide synchronized time-stamp, allowing for coherent post-processing.

TABLE I: Summary of Datasets by Section, Focus, Data Types, Platforms, Frequency Bands, and Prior Usage.

Name	Dataset Link	Focus of Dataset	Data Types	Platforms	Frequency Bands	Published Papers
Wireless I/Q Datasets (Section V)	[24], [25]	LTE I/Q sample collection for A2G propagation, channel estimation, spectrum occupancy, and UAV localization	SigMF raw I/Q samples; GPS logs; antenna pattern data	UAV with NI USRP B205mini SDR; fixed AERPAW radio node (LTE eNB)	LTE band	[26]–[28]
Wireless Spectrum Datasets (Section VI)	[29]–[33]	Wideband aerial spectrum monitoring for propagation analysis, model calibration, and spectrum allocation	Power measurements (dBm) with frequency tags; GPS logs; spectrum monitoring logs	Helikite with dual NI USRP B205mini-i and dual receive antennas; Intel NUC	87 MHz–6 GHz	[34]–[36]
5G NSA Wireless KPI Datasets (Section VII)	[37]–[40]	Aerial LTE/NR KPI measurement on a 5G-NSA network using UAV-mounted portable nodes	KPI logs (RSRP, RSRQ, SINR, CQI, RI, MCS, throughput); GPS and UAV telemetry	AERPAW SAM UAV carrying Quectel 5G modem; Nemo or PawPrints Android device; Ericsson 5G RAN (n77)	5G NR (n77)	[41]
LoRa Propagation Datasets (Section VIII)	[42]	Aerial and ground LoRaWAN propagation for IoT coverage, signal quality, and latency analysis	SigMF signal summaries (RSSI, SNR); transmitter and gateway logs (packet metadata, GPS/IMU, spreading factor, data rate); latency metrics	Tethered Helikite, UAV, and ground vehicles with USB-programmable LoRa devices; seven LoRaWAN gateways	915 MHz LoRaWAN	—
Multipath Propagation Datasets (Section IX)	[43]	A2G multipath channel characterization via synchronized channel sounding (delay spread, Doppler, path loss)	Raw I/Q with metadata (GPS, waveform parameters, sync); post-processed CIR and path-loss traces; notebooks and scripts	UAV and fixed node with USRP B210; RF front-end; GNSS-disciplined oscillator; Intel NUC	2.4 GHz	[44]
Wireless Localization Datasets (TDOA-Based) (Section X)	[45], [46]	UAV localization and tracking using TDOA from distributed RF sensors under mixed LoS/NLoS conditions; CRLB benchmarking	CSV logs of TDOA estimates; GPS ground truth; LoS/NLoS flags; error metrics; UAV telemetry	Four Keysight N6841A RF sensors on AERPAW towers; UAV-mounted SDR transmitter or DJI Inspire 2	2.4–3.3 GHz	[47], [48]
UAV-Collected RF Measurements for RF Source Localization (AFAR Challenge) (Section XI)	[49]	UAV and UGV RF localization challenge across DT and real-world testbed for locating an RF-emitting UGV	Timestamped RSS and RSQ; synchronized UAV navigation and orientation; CIR-derived quality metrics; experiment metadata	UAV and UGV with USRP B205mini SDRs; Intel NUC; DT (V-USRP, CHEM-VM, SITL); AERPAW testbed	Cellular spectrum	[50], [51]–[53]
UAV Signal Classification Datasets (Section XII)	[54]	RF-based detection and classification of UAV remote controller signals for EW and SIGINT	Raw RF captures (MAT files); time-domain samples; spectrograms; signal metadata; MATLAB scripts for preprocessing and database creation	Keysight MSOS604A oscilloscope; 24 dBi parabolic antenna; LNA; bandpass filter; MATLAB-based data acquisition system	2.4 GHz ISM band	[55]–[57]
UAV Trajectory, RSRP, and Throughput Dataset (Section XIII)	[58], [59]	Trajectory-aware KPI emulation and validation combining emulation and simulation for LTE SISO	Processed CSVs (time-stamped lat/lon/alt); RSRP; SNR; throughput; UAV telemetry	UAV with LTE SISO radio; AERPAW DT and MATLAB simulation; four BSs	LTE SISO	[60]
Ray Tracing Simulation and Measurement Comparison Dataset (Section XIV)	[61]	Direct comparison of measured versus Sionna RT-simulated RSS along UAV trajectories using OSM and Blender environments	Measured and simulated RSS; path coefficients; delays; MATLAB alignment scripts	UAV transmitter; fixed receivers; NVIDIA Sionna RT	3.3 GHz	[62]

Color code: **Signal-Level** datasets (raw I/Q), **Power-Level** datasets (spectrum/power measurements), **KPI-Level** datasets (4G/5G performance metrics).

IV. AERIAL WIRELESS DATASETS

In the rest of this paper, we will introduce various aerial wireless datasets collected at the NSF AERPAW platform, enabled by the features and capabilities discussed in Section III. For each dataset, we will provide a description of the hardware and software used to collect the dataset, the dataset format, and representative results. References are provided to access each dataset and associated post-processing scripts to replicate the results. Example uses of the dataset are also described.

Table I provides a unified summary of the aerial wireless datasets presented throughout the rest of this article, categorized by their focus, data types, experimental platforms, and prior usage. The color-coded classification, which includes signal-level (I/Q), power-level (RSS), and KPI-level (4G/5G performance) datasets, highlights the breadth and granularity of the measurements.

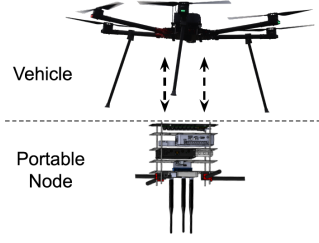


Fig. 1: Illustration of the AERPAW large multirotor-type UAV setup for the experiment, where the UAV carries a portable node [63].

V. WIRELESS I/Q DATASETS

I/Q datasets represent the most fundamental layer of wireless measurements, providing a versatile foundation for supporting research across a wide range of topics. From these raw signals, numerous KPIs can be derived, making them especially valuable for both modeling and experimental studies. In this section, we present I/Q datasets collected from an LTE network at different UAV flight heights over the Lake Wheeler Field Labs.

A. Description of Hardware and Software

The large multirotor-type UAV from AERPAW was used to collect I/Q samples during the experiment. As shown in Fig. 1, the UAV carries a portable payload that includes an NI USRP B205mini SDR. Python scripts run on the SDR and GPS module to collect I/Q samples at the desired center frequency, sampling rate, and interval, and to record the UAV's location and position. The UAV is equipped with a dipole-type antenna (SA-1400-5900). Before the experiment, the UAV's flight path, navigation speed, and position were pre-planned by placing waypoints on the map, enabling automatic control of the UAV and repeatable experiments. The collected I/Q dataset is stored in MATLAB (.mat) format and can be post-processed using MATLAB toolboxes. The GPS logs are also saved in a text file (.txt) for post-processing.

An AERPAW fixed radio node at the Lake Wheeler Road Field Labs (LWRFL) site (see Fig. 2a and Fig. 2b) is configured as an LTE Evolved Node B (eNB) to transmit the LTE downlink signal. The srsRAN open-source SDR software is used to realize the LTE eNB, where the transmitter antenna gain, center frequency, and number of resource blocks are configurable. A USRP B205mini SDR with a wideband antenna (RM-WB1) is installed at the fixed radio node.

B. Dataset Format

The I/Q measurement campaign was conducted using an LTE BS and a UAV (see Fig. 2c), which followed a zigzag trajectory at five different fixed altitudes, as illustrated in Fig. 2d. The altitudes ranged from 30 m to 110 m, with increments of 20 m between each. The I/Q samples and the UAV's GPS logs datasets are available on IEEE Dataport [24], providing a valuable resource for in-depth research involving raw I/Q data. The dataset is organized into five main folders corresponding

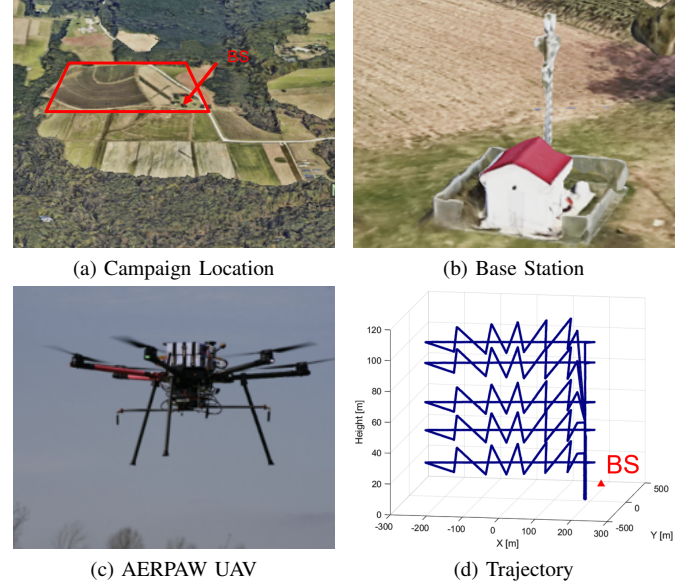


Fig. 2: Campaign environment and UAV trajectory for the I/Q measurement dataset: (a) Google Earth view of the site, (b) BS or transmitter, (c) pre-planned UAV trajectory, and (d) AERPAW UAV for I/Q signal reception.

to the I/Q measurements taken at each altitude. Each folder is named according to the respective altitude. For every altitude, the data is further divided into two subfolders: one for the I/Q samples and the other for the GPS logs. The raw data from both sources are converted into SigMF [64], accompanied by metadata that describes the recordings information. Scripts are provided to facilitate conversion into raw data formats compatible with MATLAB or Python (e.g., .mat or comma-separated values (CSV) files). For each altitude, the dataset is structured as follows:

- 1) **I/Q Sample Data:** These files are in SigMF format and are located in the `IQ_samples` folder. Each file contains I/Q samples collected over a duration of 20 milliseconds, and the files are named according to the timestamp of the I/Q sample measurement start time. Consecutive files are separated by 100 milliseconds, corresponding to the interval between measurements. With a 2 MHz sampling rate, each SigMF file contains approximately 40k I/Q samples. Additionally, the measurement collection time in Unix timestamp is included within the SigMF data.
- 2) **GPS Logs:** The GPS log file in SigMF format is stored in the `GPS_logs` folder. When converted to CSV format, each row includes the latitude, longitude, altitude (in meters), and the Unix timestamp of the GPS measurement. The record interval is one second.

To facilitate the conversion of SigMF-formatted files into more widely used formats, the repository also provides the following Python scripts:

- `sigMF2mat_IQ.py`: Converts I/Q sample data to MATLAB format. The output files are saved in the `./IQ_samples/matfile` directory.

Name	Value
IQ	1x40800 complex single
time_stamp	1x20 complex single

(a) I/Q samples of 20 milliseconds, converted to MATLAB file using provided script: `results_2022_04_02_12_13_18_008.mat`.

Latitude	Longitude	Altitude	Timestamp
-78.696544	35.727220	89.969000	1648918769.978881
-78.696545	35.727194	89.990000	1648918770.979600
-78.696542	35.727163	90.018000	1648918771.979996
-78.696539	35.727129	90.025000	1648918772.980303
-78.696536	35.727091	90.023000	1648918773.981240

(b) GPS Trajectory Data for at 90 m altitude, converted to CSV using provided script: `2022-04-02_12_57_03_vehicleOut.csv`.

Fig. 3: Snapshot of I/Q samples and GPS logs for I/Q Measurement Dataset.

- `sigMF2csv_GPS.py`: Converts GPS log file into CSV format. The output file is saved in the `./GPS_logs/csvfile` directory.

Fig. 3 shows a snapshot of a MATLAB file after conversion representing a 20 milliseconds I/Q measurement, along with five rows of the generated CSV file depicting the UAV trajectory over a five-second period.

Additionally, a `README` file is provided, offering an overview of dataset usage. The dataset is also equipped with a detailed description of the methodology, equipment, environment, and post-processing codes for air-to-ground (A2G) propagation modeling [65]. Post-processing scripts are available on CodeOcean [25] for extracting channel information, coherence bandwidth, RSRP, and RSRQ from the I/Q samples.

C. Representative Results

In this section, we present several representative results using the published post-processing scripts [25], [26]. Since the I/Q samples in this dataset are LTE waveforms and the measurement window is 20 milliseconds, we can decode the LTE frame start time and extract a full 10 millisecond LTE frame from each I/Q measurement MATLAB file. We can then plot the LTE resource grid and estimate the channel using the reference signals, as shown in Fig. 4a and Fig. 4b, respectively. From each I/Q measurement file, we can also obtain the RSRP value, and by matching the Unix timestamp with the GPS logs, we can determine the corresponding 3D location of the UAV and plot the RSRP over the UAV trajectory, as shown in Fig. 4c. Additionally, from the GPS logs, we can plot the 3D distance from the BS to the UAV over time and calculate the UAV's speed by differentiating the position, as shown in Fig. 4d and Fig. 4e, respectively. Finally, Fig. 4f accumulates the GPS logs and RSRP data along the Unix timestamp to plot the RSRP across the 3D distance, where the fitted free-

space and two-ray path loss models incorporating 3D antenna patterns are aligned with the measured RSRP behavior.

D. Possible Uses of Dataset

1) *A2G Propagation Modeling*: Artificial intelligence (AI)-based A2G propagation models can be developed using the provided I/Q datasets. AI-based training and testing models can incorporate multiple features, such as the 3D antenna pattern, UAV altitude and position, BS tower height, communication frequency and bandwidth, UAV speed. Note that our dataset [66] includes 3D radiation pattern measurements of both the transmit and receive antennas, obtained in an anechoic chamber.

2) *UAV Receiver Algorithm Design*: The provided I/Q dataset can facilitate the development of practical time and frequency synchronization algorithms, cell search, channel estimation, and decoding techniques optimized for UAV communication systems.

3) *Spectrum Occupancy and Interpolation*: For spectrum sharing and coexistence between terrestrial and aerial networks, UAV-based spectrum monitoring and interpolation techniques have been widely investigated. Our I/Q dataset includes a single zigzag trajectory at multiple altitudes, enabling altitude-dependent spectrum analysis and the study of spectrum interpolation techniques. In our preliminary works [27], [67], we propose spectrum interpolation approaches based on the 3D Kriging [27] and matrix completion [67] using the I/Q dataset.

4) *UAV Localization and Tracking*: The detection, localization, and tracking of signal sources by UAVs are key techniques for ensuring privacy and enabling network coexistence. By classifying malicious UAVs or incumbent signal sources, UAV-based systems can enhance situational awareness and support secure and reliable spectrum operations. Our I/Q dataset provides received signal strength (RSS) measurements along the UAV trajectory across the experiment site, enabling the development and evaluation of source localization and tracking algorithms. In our preliminary work, we propose a UAV localization technique based on the two-ray path loss model, incorporating 3D antenna radiation patterns [28].

VI. WIRELESS SPECTRUM DATASET

Empirical spectrum measurements are essential for analyzing wireless channel behavior under practical conditions. Such datasets offer insight into signal variations across frequency, altitude, and environment, and serve as a basis for validating analytical and simulation models. This section describes the measurement setup, including the hardware platform and data collection procedures used to obtain the reported results. Currently, five distinct spectrum measurement datasets are available on the AERPAAW dataset page [9] and Dryad [29]–[33]. In this work, we focus on the Packapalooza 2024 dataset [29]; however, the other datasets follow the same file format and directory structure.

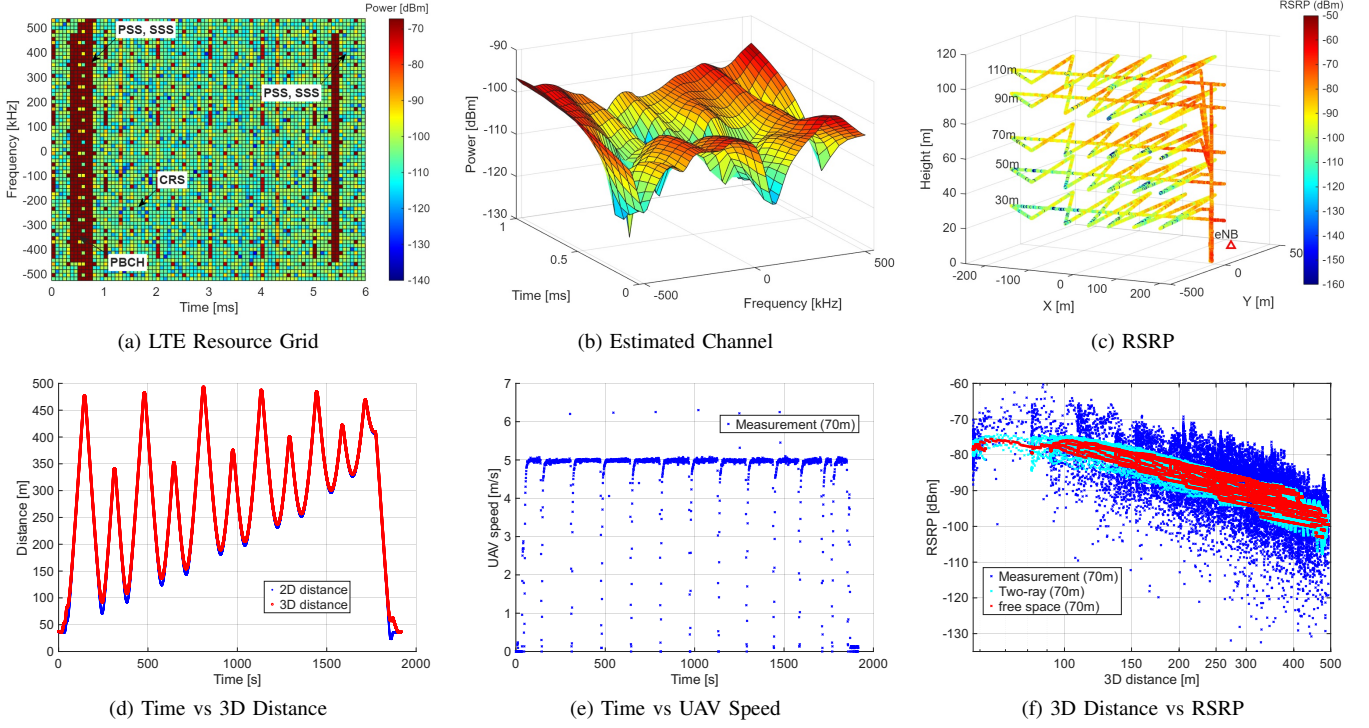


Fig. 4: Representative results from the Wireless I/Q dataset: (a) LTE resource grid, (b) estimated channel, (c) RSRP along UAV trajectory, (d) time vs. 3D distance (UAV altitude: 70 m), (e) time vs. UAV speed (UAV altitude: 70 m), and (f) 3D distance vs. RSRP with path loss fitting (UAV altitude: 70 m).

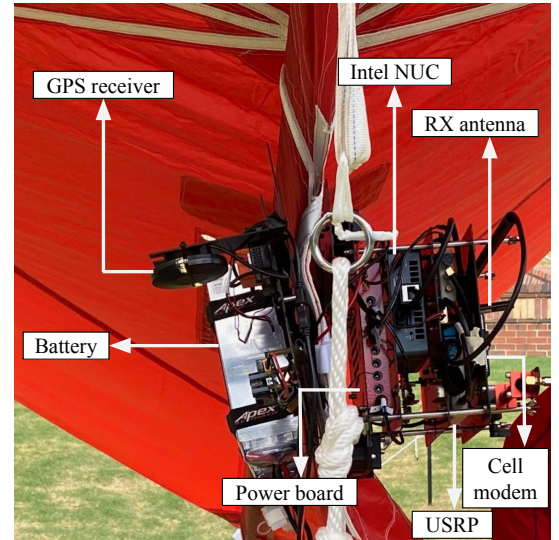
A. Description of Hardware and Software

The measurement setup consists of several components selected to ensure reliable data acquisition and analysis. The main platform is a Helikite equipped with an SDR system. The Helikite provides stable flight at altitudes up to 300 meters, enabling spectrum measurements over a wide area with limited obstruction. The SDR system comprises a USRP device and an antenna, forming the core of the data collection unit.

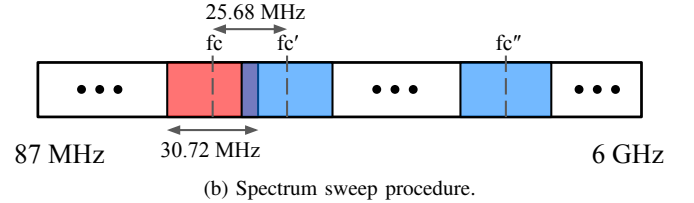
The portable nodes utilize the NI USRP B205mini-i [68], the smallest USRP featuring a Xilinx Spartan 6 XC6SLX150 FPGA [69] and an Analog Devices AD9361 RFIC direct-conversion transceiver [70]. This device supports frequencies from 70 MHz to 6 GHz and offers up to 56 MHz of instantaneous bandwidth¹ (61.44 MS/s quadrature) for full-duplex operation. As the core of our SDR platform, the USRP B205mini-i provides high sensitivity and selectivity across most commercial wireless bands, including LTE, 5G NR, and ISM.

Fig. 5a illustrates the equipment configuration attached to the helikite. On the left, the helikite adapter is shown, which houses the battery and the GPS logger. On the right, the adapter accommodates the portable node. The battery supplies power to both the portable node and the GPS logger. GPS logger is a very simple AERPAAW sample application that only logs the location of the vehicle (based on information from the vehicle's autopilot) but does not control the vehicle at all. The purpose of this sample application is to allow the monitoring

¹Instantaneous bandwidth is the maximum width of a frequency band that the device can receive or transmit without retuning.



(a) Setup of the portable node on the helikite.



(b) Spectrum sweep procedure.

Fig. 5: Measurement setup and procedure for spectrum data collection using the helikite-mounted portable node. (a) Experimental setup of the portable node on the tethered helikite. (b) Spectrum sweep procedure.

Longitude	Latitude	Altitude	Timestamp
-78.666617	35.7863918	0.977	1724514505.82774
-78.666617	35.7863911	0.997	1724514506.83163
-78.6666175	35.7863893	1	1724514507.83576
-78.666618	35.7863882	0.982	1724514508.83948
-78.6666183	35.7863862	0.971	1724514509.84351
-78.6666193	35.7863827	0.939	1724514510.84748
-78.6666204	35.7863802	0.879	1724514511.85172

(a) GPS Trajectory Data for the Packapalooza 2024 Event: 2024-08-24_11_48_21_vehicleOut.csv.

`struct` with fields:

```
freqs: [87.1600037 87.2200012 ... ] (1x98868 single)
powers: [-1.1142997e+02 -1.1111731e+02 ... ] (1x98868 single)
```

(b) Frequency and Power Data for the Packapalooza 2024 Event: spec_results_20240824_140512.mat.

Fig. 6: Snapshot of GPS Log, Frequency, and Power Data for the Packapalooza 2024 Dataset.

of the vehicle while the vehicle is controlled externally. This allows for drones and rover to be operated under manual control from the pilot, while still logging their position. The GPS logger is connected via a USB cable to a companion computer (Intel NUC) housed within the portable node. The Intel NUC manages two B205mini-i SDRs; one is equipped with a 3.5 GHz front end while the other operates without a specific front end.

The portable node is designed to operate on a 19 V supply from the battery, which powers not only the companion computer but also the active components of the front end, such as the Low Noise Amplifier (LNA). Both USRPs are connected to identical but separate receive antennas and are configured only to receive signals; there is no transmission functionality in this setup. The measured data for each sweep requires approximately 15 seconds to be stored. The spectrum sweep procedure used in the experiments is depicted in Fig. 5b, where the center frequency shift and sampling rate are 25.68 MHz and 30.72 MHz, respectively. This dataset presents results exclusively for the entire bandwidth, i.e., from 87 MHz to 6 GHz.

The primary software components consist of custom Python scripts that automate data collection and initial processing stages, ensuring consistency and efficiency in long-duration measurement campaigns. For additional information regarding the spectrum monitoring experiment, please refer to [71].

B. Dataset Format

The dataset available on the Dryad Digital Repository is comprehensively structured to support extensive research into spectrum analysis and wireless communications. It is organized into several key components aimed at providing a robust set of tools for data analysis and application.

The dataset comprises three primary types of files stored in designated folders:

- 1) **Power Spectrum Raw Data:** These are SigMF formatted files located in the `pow_spec` folder. Each MATLAB file contains frequency data (freqs in MHz) and

power measurements (powers in dBm), with filenames that correspond to timestamps indicative of the data collection time.

- 2) **GPS Logs:** SigMF formatted GPS logs are stored in the `GPS_logs` folder. These logs are convertible to CSV format using provided Python scripts. The converted CSV files include four columns: longitude, latitude, altitude (in meters), and Unix epoch timestamp.

Additionally, the repository provides Python scripts to facilitate the conversion of these SigMF formatted files into more widely used formats:

- `sigMF2mat_PW.py`: Converts power spectrum data to MATLAB format. The output files are saved in the `./pow_spec/matfile` directory.
- `sigMF2csv_GPS.py`: Converts GPS logs into CSV format. The output files are stored in the `./GPS_logs/csvfile` directory.

To use these scripts, users can run `sigMF2mat_PW.py` to process the power spectrum data, and `sigMF2csv_GPS.py` to extract the GPS logs. Each script ensures that data is accessible and easy to manipulate, enhancing the usability of the dataset for various research purposes.

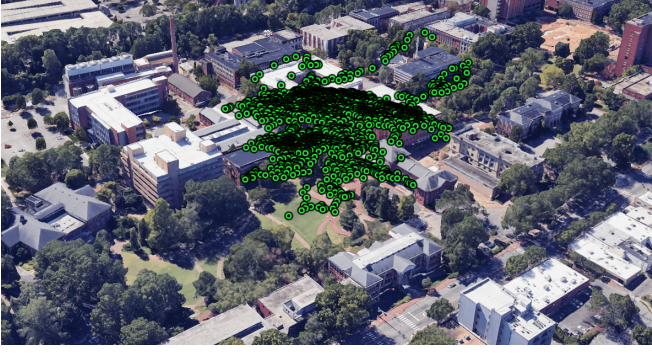
Fig. 6 shows a snapshot of the GPS trajectory data from Packapalooza 2024, along with a MATLAB file generated after conversion that represents the frequency and corresponding measured power for a single sweep.

This dataset is equipped with detailed metadata, which includes the methodology, equipment used, settings, and examples of post-processing. This comprehensive metadata ensures that users can reliably replicate measurements and fully understand the legal and ethical considerations for using the data. Additionally, a `README` file is included, providing a comprehensive overview of the dataset's scope, usage instructions, and curator contact information. This essential support is invaluable for effectively interpreting and applying the data to specific research needs.

C. Representative Results

This section presents representative results from spectrum monitoring experiments conducted in both urban and rural environments. Utilizing advanced aerial platforms such as helikites, these experiments offer valuable insights into how environmental factors and topographical features influence wireless signal propagation and distribution. By comparing the results from densely populated urban areas during the Packapalooza event with those from the more open and sparse rural areas near Lake Wheeler, we aim to highlight the distinct challenges and dynamics encountered in different settings.

1) **Helikite Trajectory Analysis:** In urban settings during the Packapalooza event, the helikite's trajectory, as depicted in 7a, showcases its path above a densely populated area with significant deviations caused by complex wind interactions with urban structures. This erratic movement, indicated by a red trace, potentially affects spectrum measurements due to variable altitudes and obstructions. Conversely, in the rural landscape near Lake Wheeler as shown in Fig. 7b, the helikite



(a) Packapalooza 2024.



(b) Lake Wheeler 2024.

Fig. 7: Helikite location for spectrum measurements in (a) Packapalooza 2024 and (b) Lake Wheeler 2024.

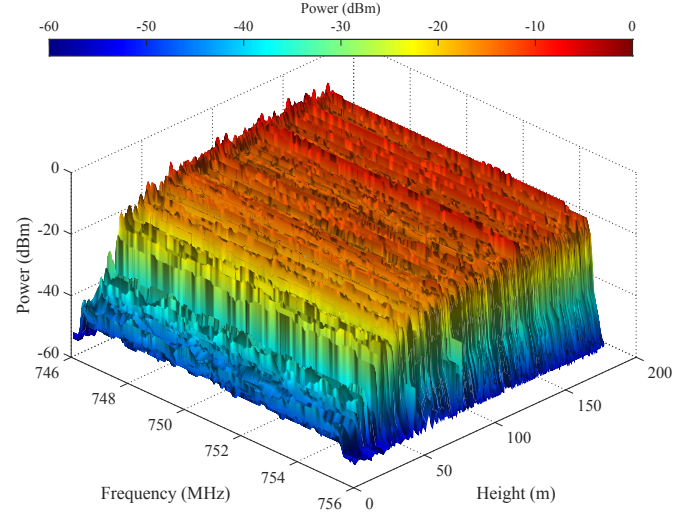
exhibits a more stable and elongated flight path across open fields, suggesting more consistent data collection due to fewer obstructions and a steadier altitude control.

2) *Spectrum Analysis*: In Fig. 8, the 3D plots for Band 13 downlink (746 - 756 MHz) reveal distinct variations in signal power across different environments. In urban settings, Fig. 8a, power levels fluctuate significantly with altitude due to multipath effects and obstructions, showing a trend toward stabilization as altitude increases. In contrast, rural settings in Fig. 8b display a more uniform increase in power levels at higher altitudes, indicating clearer signal paths and fewer obstructions. Fig. 9 illustrates the mean received power as a function of frequency across altitudes for the Packapalooza 2024 dataset, with several active United States radio bands also indicated. The results show that the mean received power is significantly higher below 1 GHz, where many LTE and NR network bands are located, compared to the remainder of the sub-6 GHz spectrum.

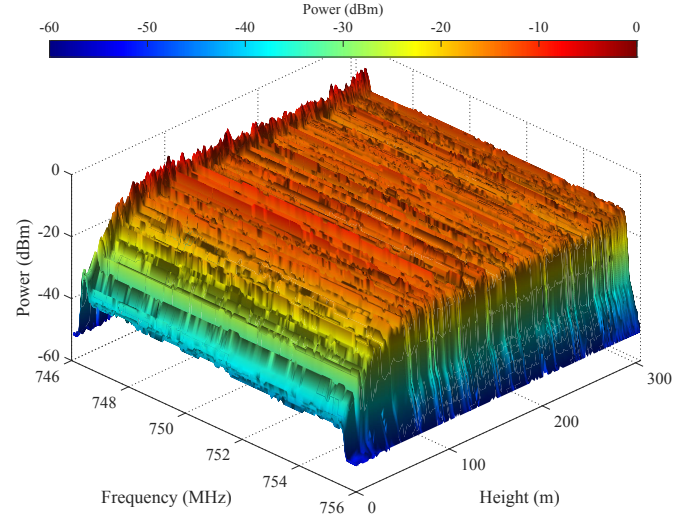
D. Possible Uses of Dataset

Our helikite-based spectrum monitoring dataset provides calibrated received power measurements across a wide frequency range (87 MHz–6 GHz) with corresponding GPS coordinates, altitude, and timestamp. While it does not include raw I/Q samples or power spectral density (PSD) estimates, the dataset remains valuable for many practical wireless research and regulatory applications.

1) *Spectrum Allocation Analysis*: The dataset enables spatial and altitudinal characterization of spectrum utilization



(a) Packapalooza 2024.



(b) Lake Wheeler 2024.

Fig. 8: Altitude dependent power for DL 13 in (a) Packapalooza 2024 and (b) Lake Wheeler 2024.

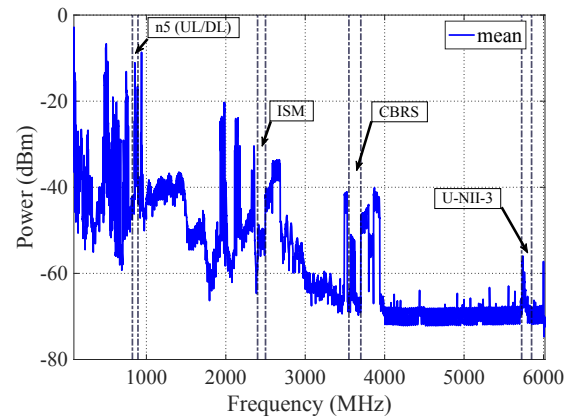


Fig. 9: Representative results on the mean received power versus frequency averaged across all altitudes for the Packapalooza 2024 dataset.

across urban and rural environments. For example, compar-

isons of received power in the Citizens Broadband Radio Service (CBRS) and Television White Spaces (TVWS) bands can reveal underutilized areas or high-demand regions. In urban settings, it allows researchers to measure outdoor signal levels in the 6 GHz unlicensed band to assess potential interference from indoor Wi-Fi 6E deployments. The availability of time-stamped data also permits the exploration of temporal usage trends, such as peak usage periods or band-specific congestion.

2) *Calibration of Analytical Models*: The dataset supports calibration of analytical and simulation models, including stochastic geometry and empirical path loss frameworks. The relationship between received power and altitude can be used to validate altitude-aware propagation assumptions. Differences observed between rural and urban measurements can help refine clutter loss models. In addition, LoS probability models can be empirically evaluated using elevation-dependent signal trends.

3) *Propagation Model Tuning*: Researchers can use the dataset to develop or refine radio propagation models. The location-tagged power measurements support construction of empirical path loss curves for a range of frequency bands. Comparing measurements taken in urban versus rural environments helps to characterize the impact of buildings, vegetation, and other obstructions. Moreover, differences between low-frequency bands like FM and higher bands like 3.5 GHz can be used to study frequency-dependent attenuation.

4) *Machine Learning for Signal Estimation*: The dataset is suitable for developing machine learning models that estimate received power from spatial and environmental features. Inputs such as latitude, longitude, altitude, and frequency can be used to train regressors for power prediction. The labeled nature of rural and urban environments supports classification tasks, such as identifying the type of environment based on observed signal levels. The data can also help delineate signal boundaries or approximate coverage maps through supervised learning.

5) *Anomaly and Interference Detection*: The received power measurements allow for basic anomaly detection techniques. Sudden spikes or dips in power levels may indicate unauthorized transmissions or interference events. Statistical properties such as skewness or variance can be used to detect deviations from normal signal patterns. When the data is tracked over time, researchers can analyze signal disruptions or temporal anomalies in specific frequency bands.

6) *Interpolation and Coverage Mapping*: Despite the absence of I/Q or PSD data, the dataset is well-suited for generating radio environment maps. The geolocation and altitude information associated with each measurement can be used for spatial interpolation techniques such as Kriging or inverse distance weighting. By interpolating the power values, researchers can generate two-dimensional (2D) or 3D signal coverage maps for individual bands. The multi-band nature of the dataset further allows for frequency-aware coverage visualizations across the monitored spectrum.

VII. 5G NSA WIRELESS KPI DATASET

As the demand for using advanced communications to support various use cases of UAVs rapidly increases, under-

standing the performance of 5G terrestrial networks in the 3D spatial domain becomes critical. In this section, we present datasets for 5G wireless KPIs and the corresponding experimental results from aerial measurements using a 5G-enabled UAV on a 5G non-standalone (5G-NSA) network in C-band in the AERPAW testbed. In particular, the datasets provide the field-measured RF and physical (PHY) layer parameters of LTE and NR carriers on the 5G-NSA network using three different types of small portable nodes (SPN) based on Quectel 5G modem, a modified Android phone with Nemo software, or a COTS Android device with PawPrints, a custom App that relies on open-source Android application programming interfaces (APIs).

A. Description of Hardware and Software

AERPAW testbed constitutes various types of wireless radio access platforms of 4G/5G. One of them is a commercial grade 5G network based on Ericsson equipment, which is the main focus of the experiment for these datasets. AERPAW infrastructure for these experiments involves a) Ericsson 5G network with RAN and Core, b) SPN, c) AERPAW fixed node (AFN) as application server of user plane traffic, and d) the UAV to carry the SPN during aerial experimentation (Fig. 10). The RAN of this 5G system is deployed at the AERPAW LWRFL, which is a rural agricultural area (see Fig. 2a), predominantly an open aerial field with some vegetation on the ground.

To characterize the aerial performance of a 5G system, we used a 5G NSA system with overlaid NR and LTE sectors. The LTE anchor carrier is in band 66 (1.7/2.1 GHz) with 5 MHz channel bandwidth and an NR carrier at 3.4 GHz in band n77 with 100 MHz of channel bandwidth. The sectors use a pair of dual +/-45 deg cross polarized directional antennas with 120 degrees of azimuth beam width facing the north-west direction from the BS tower. LTE employs 2x2 MIMO on the downlink, whereas NR uses 4x4 MIMO. For this experiment, both LTE and NR carriers are set at 5 watts of transmit power per antenna port.

For aerial experiments, we have used three types of portable nodes as 5G devices carried by UAVs based on a) 5G modem, b) Keysight Nemo device, and c) COTS Android device with PawPrints software. The portable node based on the 5G modem as shown in Fig. 11 uses a Quectel 5G module to connect to the Ericsson 5G network, a LattePanda module as a companion computer to interface with the UAV, and an ATT 5G modem for C2. The portable node that is used as UE is mounted on a small AERPAW Multi-rotor (SAM) UAV as shown in Fig. 11.

A high-level end-to-end software architecture is given in Fig. 10. There are three main software functions at the portable node, namely radio software, traffic software, and vehicle control software. These software modules run on the E-VM of the portable node and generate real-time radio, traffic, and vehicle logs during the experiment. The E-VM at the AFN provides the other end point of the server-client model for the user plane data through a traffic software and logging.

For Android device-based experiments, the measurement campaigns were conducted using SAM carrying a SPN with

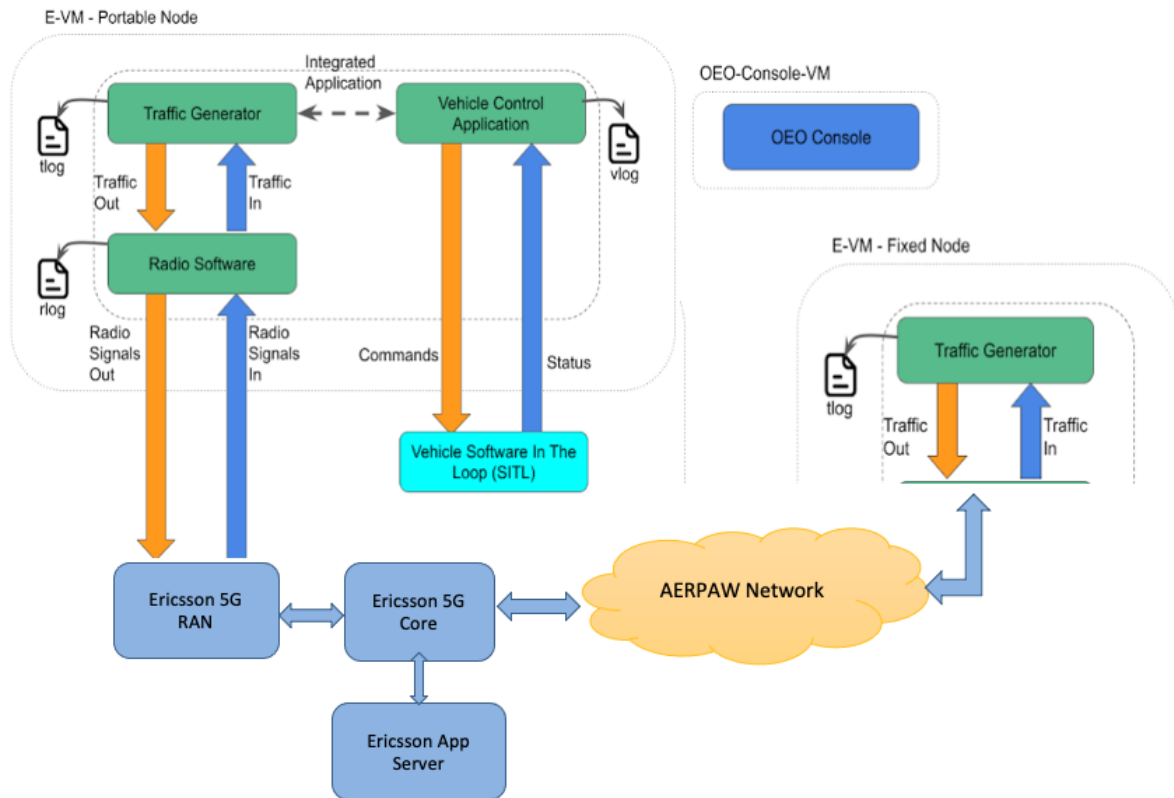
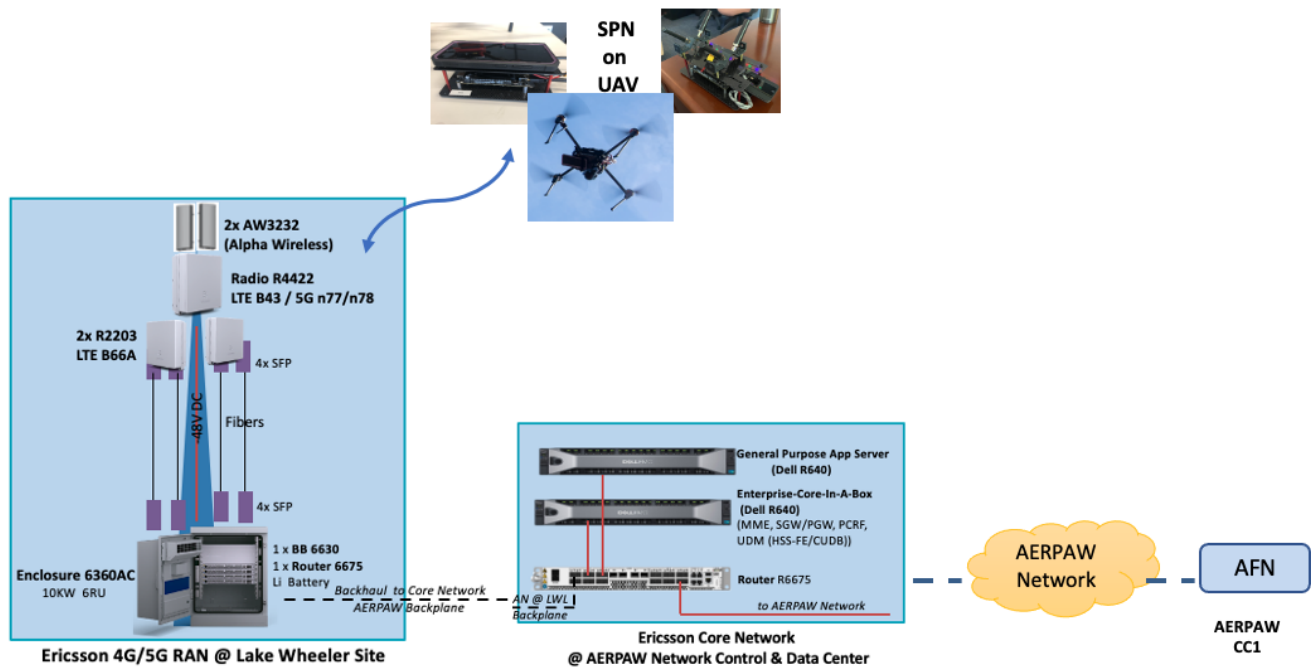


Fig. 10: Ericsson 5G-NSA AERPAW Infrastructure (top), and the SW workflow with 5G SPN on the 5G-NSA network (bottom).

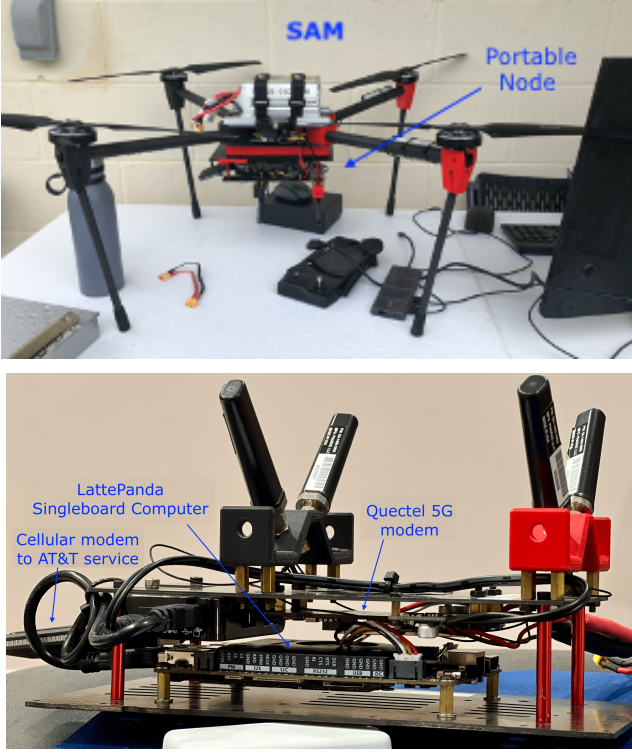


Fig. 11: SAM (top) mounted with a portable radio node (bottom) used in A2G measurements [41].

a Nemo and a PawPrints devices as the payload. Custom hardware mounts were designed for this SPN to carry two Android smartphones simultaneously during aerial flights. The two Android devices used were: a commercial standard Samsung S21 device with PawPrints, an internal AERPAW-built Android application using open source Android APIs to log radio KPIs, and a modified Samsung S23 containing Keysight Nemo software [72], with custom firmware, access to internal modem metric, and a wider range of KPIs. A Latte Panda on-board the SPN collected and logged GPS and UAV attitude measurements, obtained from GPS sensors and an Ardupilot, respectively, along with radio KPIs streamed by the Android phone over USB. These Android phones connected to the private 4G/5G Ericsson cell tower. These phones also connected as clients to the iPerf Server at the AFN, enabling throughput measurements.

B. Dataset Format

The 5G-NSA datasets contain two different categories based on the type of portable node used: a) 5G modem-based SPN and b) Android device-based SPN.

1) **5G modem datasets:** This dataset provides comprehensive RF/PHY layer and throughput measurements from an Ericsson 5G-NSA network on the AERPAW platform captured during an UAV flight mission conducted on September 15 and October 26, 2023. A Portable Node equipped with a Quectel 5G Modem mounted on an UAV was used for these experiments. The UAV was programmed to follow a predefined zigzag flight path at 30 meter altitude covering the north-west quadrant of the area from the BS tower, while the portable

node captures and records the logs. This data set specifically includes RF/PHY layer performance metrics such as RSRP, SINR, Channel Quality Indicator (CQI), Rank Indicator (RI), Modulation and Coding Scheme (MCS) on the LTE and NR carriers as well as the throughput, the application layer performance metric. The same experiment was repeated for different yaw positions of the UAV, 315 and 45 degrees, to understand the impact of angular orientation of portable node antennas with respect to the BS antennas. An iperf application with downlink traffic between the client at the portable node and server at the AERPAW fixed node was used for this experiment.

The data set comprises raw data collected at the portable node right from the experiment as well as a set of post-processed log files. The following four files contain the raw data logs from the sample experiment.

- `Serving_cell_Params_ENDC.csv`: Serving cell configuration and RF measurements such as Cell Id, RSRP, SINR, eARFCN, band, and subcarrier spacing on both the LTE and NR carriers of the 5G NSA network
- `Basic_and_Other_Params.csv`: PHY layer parameters such as downlink MCS values and the measured and reported channel state information CQI, PMI, and RI by the portable node.
- `<date>_vehicleOut.txt`: The vehicle telemetry log recorded during the UAV flight that includes its geolocation, orientation, speed with date and time.
- `<date>_iperfclient_log.txt`: The raw iperf3 log captured in real time during the flight contains the interval, transfer size in megabytes, and bandwidth in megabytes/sec.

The above raw log files are post-processed to generate the following CSV data files of individual RF and PHY parameters.

- `inputf1_cellid_with_header.csv`: Serving cell cell-id of LTE sector of 5G-NSA.
- `inputf2_cellid_with_header.csv`: Serving cell cell-id of NR sector of 5G-NSA.
- `inputf1_rsrp_with_header.csv`: RSRP values of LTE carrier, recorded in dBm.
- `inputf2_rsrp_with_header.csv`: RSRP values of the NR carrier, recorded in dBm. The RSRP values indicate the power level of the reference signals received from the cell.
- `inputf1_sinr_with_header.csv`: SINR data of the LTE carrier, provided in dB.
- `inputf2_sinr_with_header.csv`: SINR data of the NR carrier, provided in dB. The SINR values offer insights into the quality of the signal relative to background noise and interference.
- `inputf2_cqi_with_header.csv`: Measured CQI values on the downlink NR carrier.
- `inputf2_mcs_with_header.csv`: Downlink MCS values on the NR carrier.
- `inputf2_ri_with_header.csv`: Measured RI values on the downlink NR carrier.
- `input_throughput_with_header.csv`: Down-

link throughput, expressed in Mbps, the quantify the data transmission rate achieved during the UAV's flight.

The above CSV log data include the recorded date, time and the UAV geolocation (longitude, latitude, and altitude) data. These datasets from these experiments are publicly available at the AERPAW datasets webpage [9] as well as in Dryad repository, an open data publishing platform.

2) **Nemo and PawPrints datasets**: Depending on the setup, the Nemo and PawPrints datasets can contain up to three logs per measurement campaign: 4G/LTE logs, 5G/NR logs, and iPerf throughput logs. Within each log type, radio or throughput KPIs are merged with UAV location based on timestamps into a single CSV file. The PawPrints 4G KPIs, obtained from Android APIs, include LTE RSRP, RSRQ, Received Signal Strength Indication (RSSI), Physical Cell Identity (PCI), tracking area code (TAC), and cell id (CI). The PawPrints 5G KPIs consist of the 5G synchronization signal's RSRP, RSRQ and RSSI. Throughput measurements are also shared whenever applicable. Thus, the PawPrints dataset consists of three files:

- `pawprints_4G_LTE.csv` : PawPrints 4G log
- `pawprints_5G_NR.csv` : PawPrints 5G log
- `pawprints_iperf_throughput.csv` : iPerf client throughput.

The Nemo KPIs are more comprehensive and additionally include 5G MCS, CQI, and channel rank. The Nemo dataset consists of:

- `nemo_5G_NR.csv` : Nemo 5G log
- `nemo_4G_LTE.csv` : Nemo 4G log

C. Representative Results

This section presents a few representative results from the aerial experiments using the above 5G modem and Android devices based portable nodes on the Ericsson 5G-NSA network.

1) **5G modem dataset results**: As explained above, we used an UAV powered by a 5G Quectel modem to measure and collect data along a zigzag aerial path trajectory. From these experiments the RF/PHY parameters as well as application layer performance parameters were collected for analysis.

The RSRP and SINR are some of the RF parameters and CQI, RI and MCS are some of the PHY layer parameters measured and logged on the 5G-powered portable node presented here. Fig. 12a and Fig. 12b show the measured levels of downlink RSRP and SINR on LTE and NR carriers versus distance and time, whereas Fig. 12c and Fig. 12d show the same RSRP and SINR of NR carrier along the geo locations of the zigzag UAV flight trajectory. Similarly, Fig. 13 shows the reported CQI, RI and MCS versus the distance, time and geo location on the LTE and NR downlink carriers. An iPerf client-server app with downlink traffic was used between the portable node and the wired AFN, and Fig. 14 shows the downlink throughput achieved at the application layer in real-time.

2) **Nemo and PawPrints dataset results**: This section presents the representative results from the wireless KPI datasets collected from a UAV in two scenarios: measurements

of a private BS at the AERPAW LWRFL with controlled UAV trajectories (using PawPrints and Nemo), and measurements of commercial cell towers from a tethered Helikite during the Packapalooza 2023 festival (using PawPrints). Fig. 15 depicts some representative results from the first scenario, when the UAV traces sawtooth trajectories in the horizontal plane, at increasing distances from the private Ericsson BS. Fig. 15a and Fig. 15b show a heatmap of the iPerf throughput measured by the Nemo and PawPrints device, respectively, when both were operating as client simultaneously. The heatmaps confirm the reduction in throughput observed near handover regions, particularly at sector boundaries, as evidenced by the blue-shaded areas. Fig. 15c shows the channel rank of the physical downlink channel recorded by Nemo during the flight.

Fig. 16 depicts the changing RSRP of an LTE node with varying Helikite altitude as observed by a PawPrints device during the Packapalooza 2023 event. The PawPrints Packapalooza 2023 dataset also contains extensive records of other nearby commercial LTE cellular towers, operated by AT&T and their network KPI values near the NC State University campus. Both datasets can be processed using the data processing scripts in the previous sections to analyze network coverage and performance in the aerial dimension.

The data from these experiments are publicly available at the AERPAW datasets webpage [9] and in Dryad research repository [37]–[40].

D. Possible Uses of Dataset

The KPIs available in this dataset can be utilized for the following representative research problems.

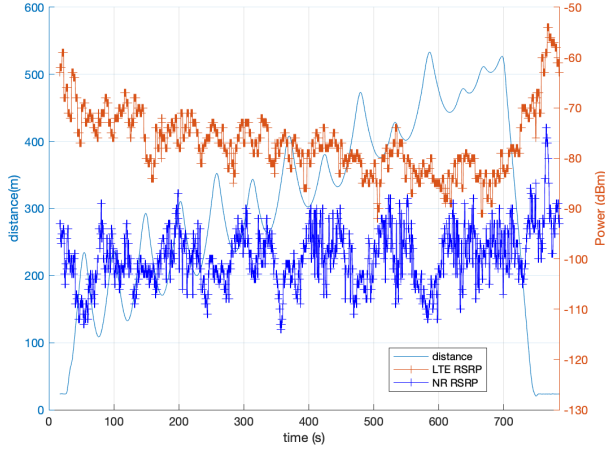
1) **Study of wireless channel propagation in rural aerial environments**: The empirical data on variations in signal strength, with distance from the BS and elevation and azimuth angles, can be used to validate existing theoretical propagation models and ray-tracing simulators, along with developing new mathematical or machine learning models. Moreover, deeper insights can be obtained by studying the variation in channel rank and channel quality index in aerial scenarios.

2) **Application layer throughput prediction in rural aerial scenarios**: Throughput prediction models for rural aerial scenarios can be created by studying the relation between physical layer KPIs such as channel rank and received SINR, and iPerf throughput.

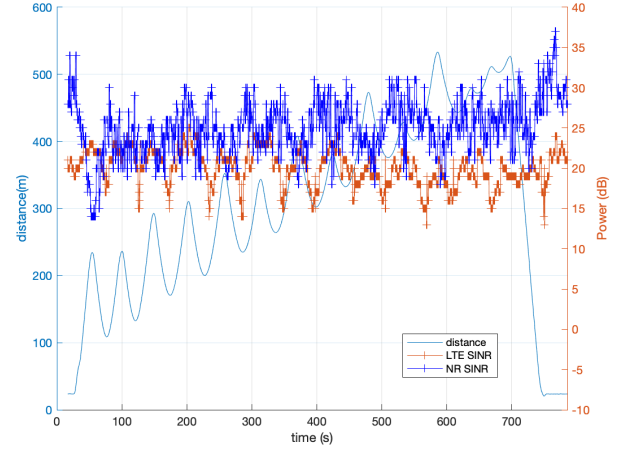
3) **Analysis of commercial cellular network coverage in aerial urban environments**: The Packapalooza dataset contains received signal strength and quality values at various altitudes, which can be used to study the suitability of commercial cellular networks, designed to serve ground users, for aerial operations.

VIII. LoRA PROPAGATION DATASET

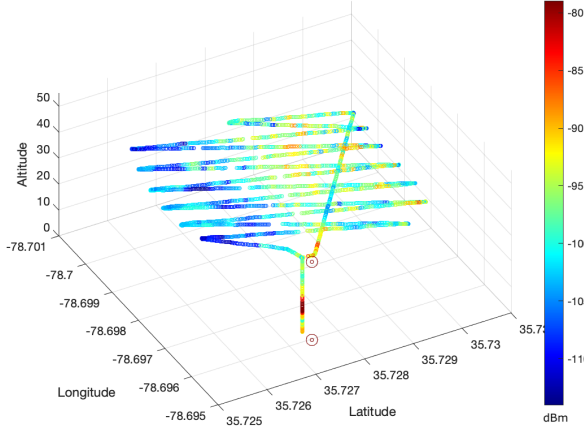
The LoRaWAN technology utilizes chirp modulation techniques to support long range, low power communications, exhibiting unique propagation characteristics for measurement. In this section, we describe AERPAW's infrastructure for such measurements and present representative results from LoRaWAN measurement campaigns.



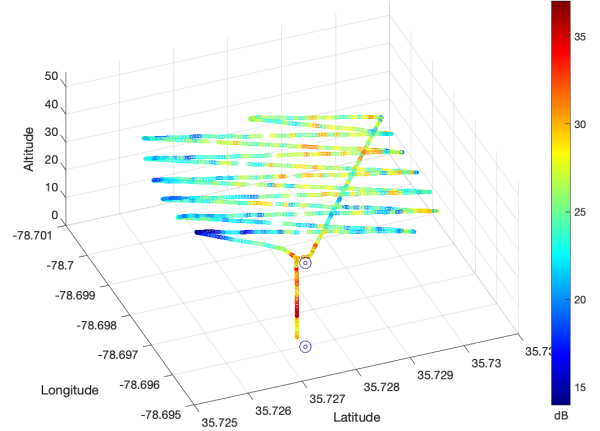
(a) RSRP of LTE and NR carriers



(b) SINR of LTE and NR carriers



(c) RSRP of NR carrier



(d) SINR of NR carrier

Fig. 12: (a) RSRP and (c) SINR of LTE and NR carriers vs. distance and time, (b) RSRP and (d) SINR of NR Carrier with UAV geo-location [41]

A. Description of Hardware and Software

The LoRa infrastructure consists of USB-compliant programmable LoRa devices and seven LoRaWAN gateways (see Fig. 17), which relay LoRa communications over the AERPAAW backplane to tenant-dedicated Docker containers for executing application-specific data processing tasks. Standard software stacks are installed on the containers, including Prometheus and PostGRES storage, which feed data for Grafana-based visualization. These LoRa devices can be driven by ground vehicles, or carried as a payload by UAVs or a Helikite, shown in Fig. 18, thus supporting ground, aerial, and static experiments. The LoRa devices allow the experiments to configure the spreading factor between 7 to 12, and accordingly set the transmission data rate from DR3 (5.47 kbps) to DR0 (0.25 kbps).

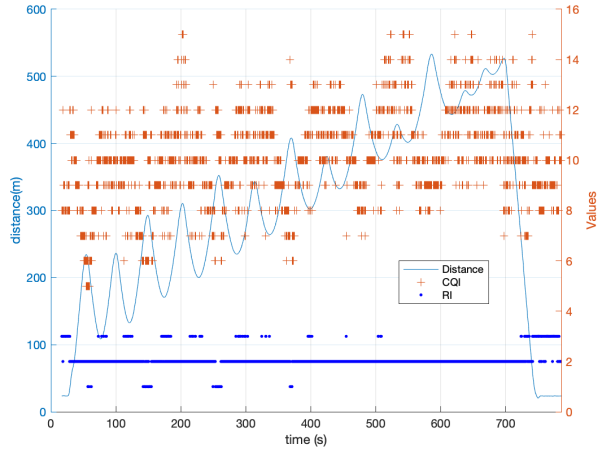
B. Dataset Format

The dataset includes logs from the LoRa transmitter device and the LoRaWAN gateways. The transmitter logs include the following measurements:

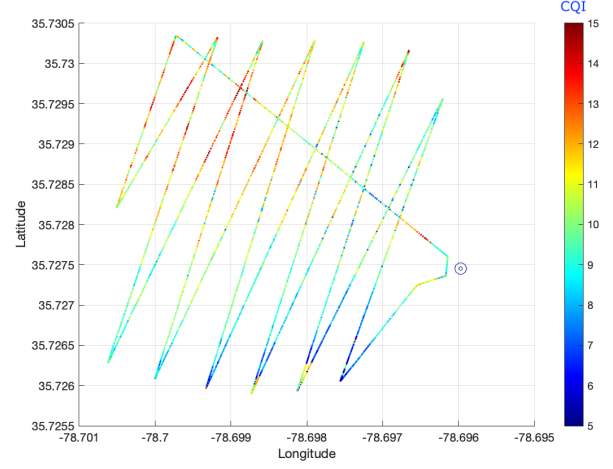
- Packet metadata such as packet ID and packet sequence number.
- LoRaWAN datarate
- Transmission bandwidth.
- Code rate of transmission.
- Timestamp of packet transmission.
- Frequency of transmission
- Geographic location (latitude, longitude, altitude) and orientation (yaw, pitch, roll) of the vehicle carrying the LoRa transmitter.
- Velocity of the vehicle carrying the LoRa transmitter.
- Transmission spreading factor.
- GPS metadata, such as the number of satellites locked to.

The logs shared from the LoraWAN gateway include:

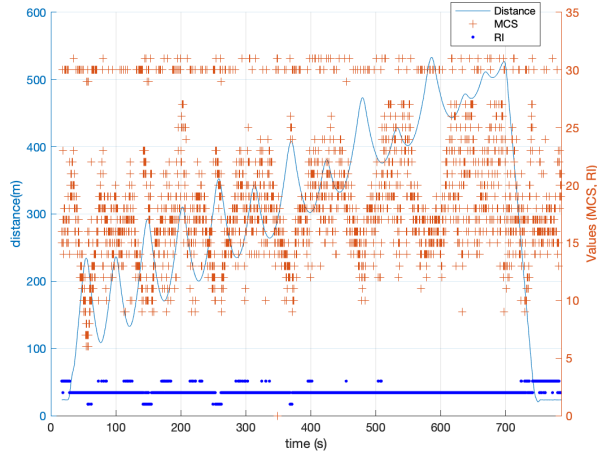
- Received signal strength of the packet, as RSSI.
- Received signal to noise ratio of the packet.
- Timestamp of packet reception.
- Frequency channel in which the packet was received.
- ID of the RF chain used to receive the packet.
- Geographic location and ID of the receiving LoRaWAN



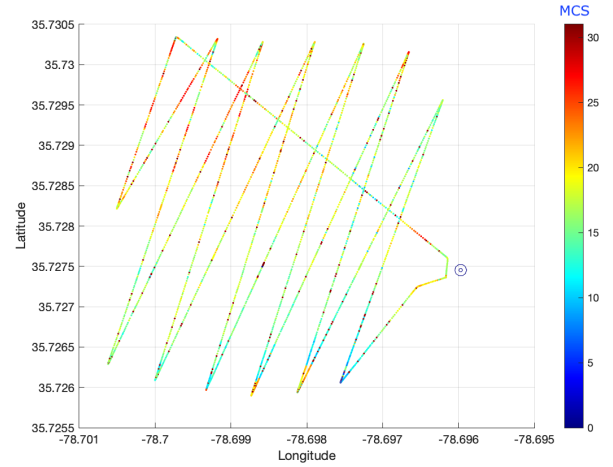
(a) CQI & RI of NR carrier



(b) CQI of NR carrier

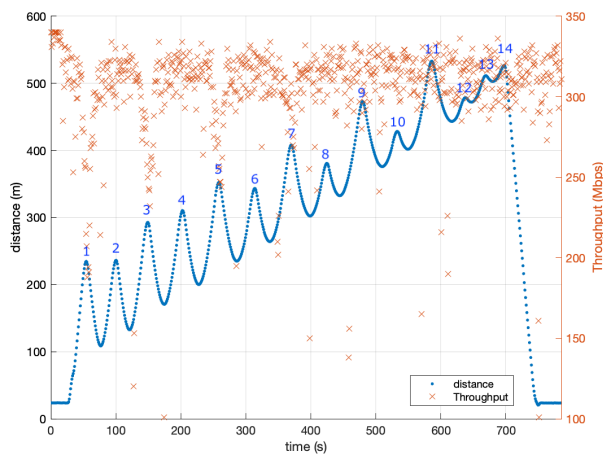


(c) MCS & RI of NR carrier

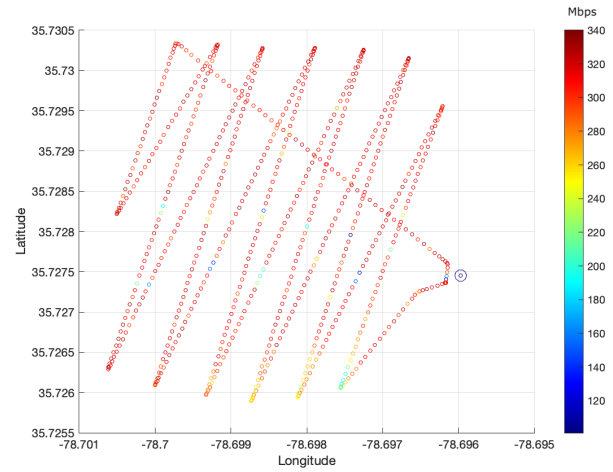


(d) MCS of NR carrier

Fig. 13: (a) CQI and (c) MCS of NR carrier with respect to distance and time, (b) CQI and (d) MCS of NR Carrier with UAV geo-location



(a) Throughput vs distance & time



(b) Throughput vs geo location

Fig. 14: (a) Throughput with respect to distance and time, (b) Throughput with UAV geo-location [41].

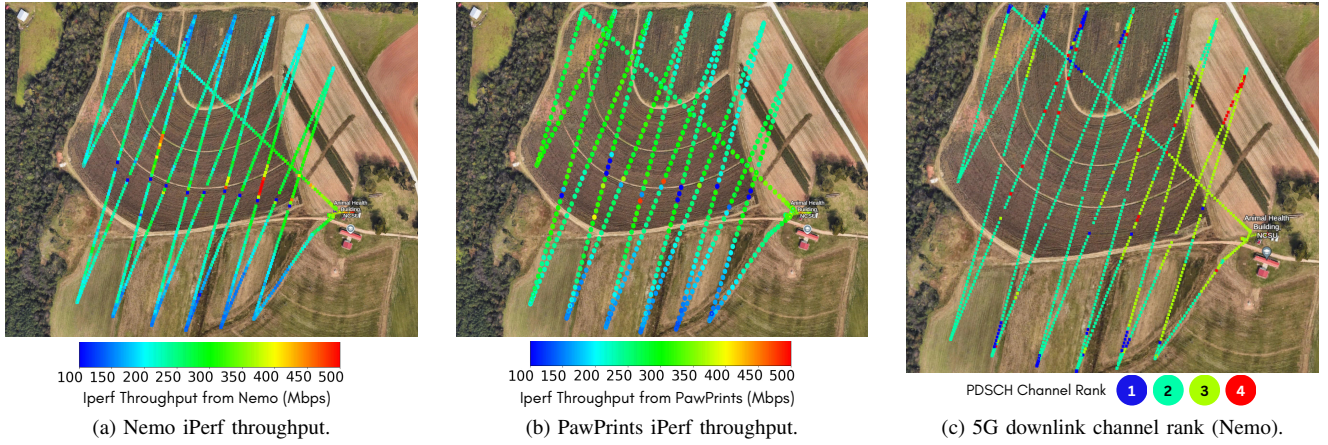


Fig. 15: Wireless KPI data gathered at AERPAW Lake Wheeler tower 1: (a) and (b) show iPerf throughput observed by Nemo and PawPrints, respectively, while simultaneously sending traffic as iPerf clients; (c) shows the 5G downlink channel rank as recorded by Nemo.

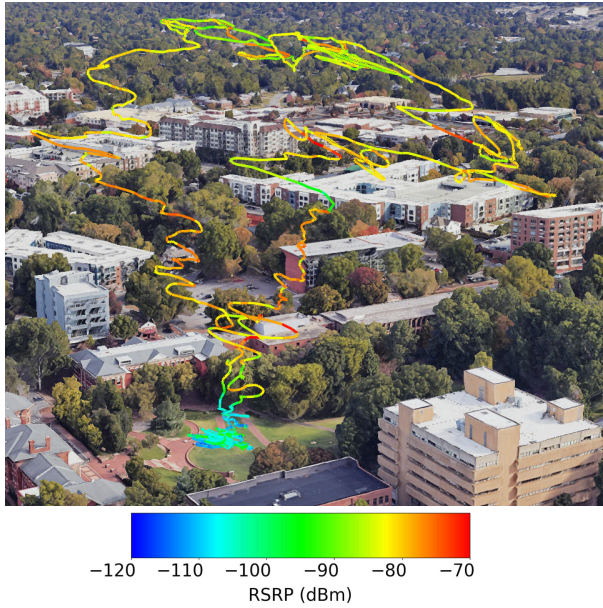


Fig. 16: RSRP variation against altitude observed by a Paw-Prints device on a Helikite at Packapalooza 2025.



Fig. 17: LoRaWAN gateway and a USB-compliant programmable LoRa device, used in the AERPAW system for mobile LoRa experiments.



Fig. 18: Ground and aerial vehicles that can carry LoRa devices for mobile experiments.

gateway.

Additionally, logs of failed transmissions are included in a

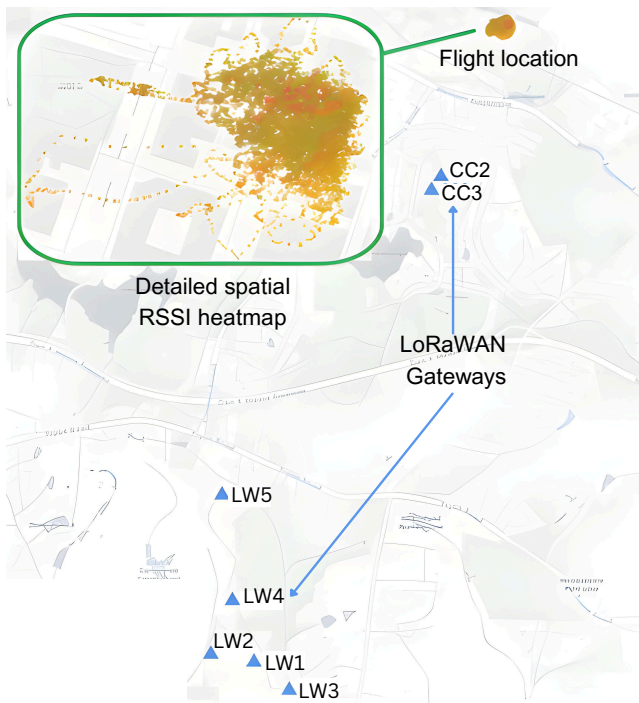


Fig. 19: Location of the LoRa tethered Helikite flight, showing location of the six LoRaWAN gateways and the top view of the Helikite locations.

CSV file, named `failed_tx_packages.csv`. The summary of data rates, grouped by the gateway, are also included in a CSV file, named `gateway-dataRate-Table.csv`. Raw signal data, containing signal-to-noise ratio (SNR) and RSSI, is shared in sigMF format. Examples of Python scripts are provided to convert from sigMF format to CSV format and to generate plots from the resultant CSV file.

C. Representative Results

The LoRa transmitter device was carried as a payload by a tethered HeliKite, which traced a free-floating trajectory over the North Campus of NC State University, Raleigh. The LoRa device was controlled by a Latte Panda mini computer, which also logged packet transmission details. The LoRa device transmitted packets, containing a sequence number and timestamp, at intervals of 1.5 s to six LoRaWAN gateways at multiple locations, as shown in Fig. 19. These gateways recorded the received RSSI, received SNR, and the timestamp of reception. The CDF of received RSSI at the gateways is shown in Fig. 20a. The variation in received SNR with received RSSI, over all the LoRaWAN gateways, is shown in Fig. 20b, which indicates that higher RSSI results in a smaller range of SINR values. In contrast, lower RSSI values can cause SIR fluctuations as noise and interference dominate. The dataset not only provides physical layer signal strength parameters but also includes packet-level latency metrics, the distribution of which is shown for each transmitted packet in Fig. 20c, aggregated over all LoRaWAN gateways.

Data from LoRaWAN experiments are publicly available at the AERPAW datasets webpage [9] and in Dryad research

repository [42].

D. Possible Uses of Dataset

Use cases of this data set include:

- 1) *LoRaWAN Optimization*: Analyze gateway performance to optimize deployment strategies for Internet of Things (IoT) networks.
- 2) *Data Reliability Studies*: Assess the reliability of transmitted data under varying conditions (e.g., noise levels, signal strength).
- 3) *IoT-Based Navigation Systems*: Support development of IoT-enabled vehicle navigation and telemetry systems.
- 4) *Energy Efficiency Analysis*: Explore power consumption patterns in LoRaWAN devices.

IX. MULTIPATH PROPAGATION DATASET

The behavior of radio signals as they propagate through the environment is a key factor in the design and performance of various wireless systems, ranging from radars to cellular networks. In this context, multipath propagation datasets provide a crucial pathway to a deeper understanding of wireless communication principles, which ultimately contributes to robust wireless systems. This section describes the hardware and software components of some of a propagation dataset acquired with an open-source channel sounder [43], [44] from the AERPAW testbed platform.

A. Description of Hardware and Software

The experimental setup involved a UAV and a fixed node located at the AERPAW Lake Wheeler testbed site in Raleigh, North Carolina. The UAV carries a portable node equipped with a USRP B210, RF front-end, Intel NUC, and a custom-designed Global Navigation Satellite System disciplined oscillator (GNSSDO), as shown in Fig. 21a. The fixed node is configured with identical equipment.

Accurate characterization of A2G multipath wireless channels requires precise temporal alignment between the transmitting and receiving devices. Multipath propagation inherently introduces time-varying delays and phase shifts, and without stringent synchronization, these effects can be misinterpreted as genuine channel behavior. The custom-designed GNSSDO system, as shown in Fig. 21b, employed by AERPAW ensures a stable and common timebase for both the UAV and fixed node, minimizing phase ambiguity and enabling reliable quantification of multipath components and time-domain channel impulse responses (CIRs). The GNSSDO system achieves 2.5 ns pulse-per-second (PPS) accuracy between the nodes, facilitating high-fidelity channel measurements.

Fig. 21c presents the overall architecture of the channel sounder software. The software, developed in Cython, enables real-time processing of the USRP B210's full bandwidth (56 MHz) and beyond, as described in [43]. The USRP B210 is synchronized using GNSSDO-generated PPS and 10 MHz reference signals, ensuring accurate timing control and reliable data acquisition. The system supports multiple configurable sounding waveforms, including Zadoff-Chu, pseudo-noise (PN), and chirp sequences. For the measurements in

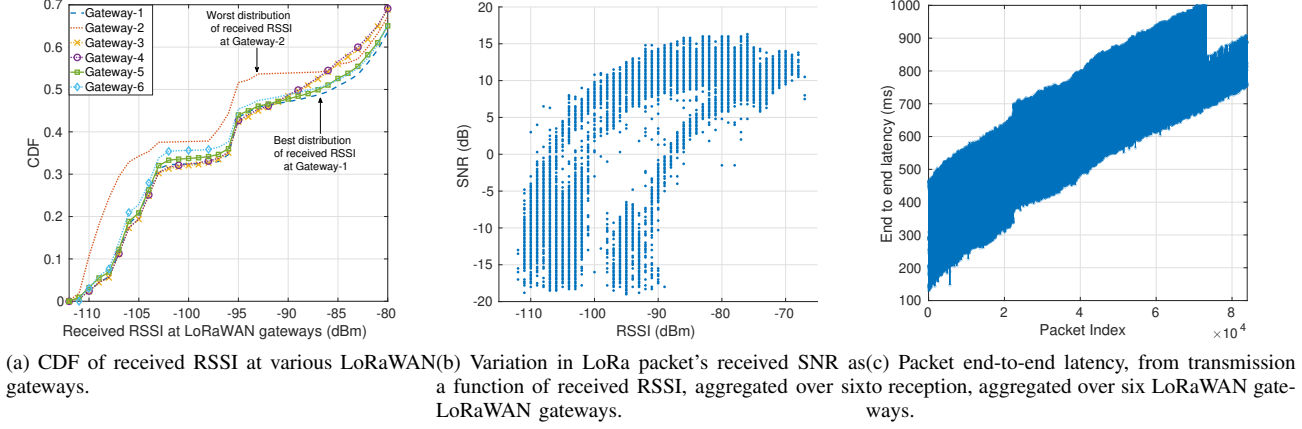


Fig. 20: The KPIs of LoRaWAN links measured during a tethered Helikite experiment, where the LoRa transmitter was elevated, and data were collected across six gateways. **(a)** Empirical CDF of received RSSI values observed at multiple gateways. **(b)** Relationship between received RSSI and corresponding SNR, highlighting link quality variation. **(c)** End-to-end packet latency measurements, from transmission to reception, aggregated across all gateways.

this study, a Zadoff-Chu sequence of length 401 and root index 200 was used, with each sequence repeated four times. The sounding was conducted at a measurement frequency of 4 Hz, with a center frequency of 3564, 3620, or 3686 MHz, and a transmit power of 19 dBm. A total of nine flight experiments were performed at three altitudes (30, 60, and 90 meters) along a 500-meter flight path, with a flight speed of 5 m/s. This configuration supports a wide range of channel sounding scenarios with high temporal and spatial resolution. For additional information regarding software, please refer to [43].

B. Dataset Format

The raw channel sounding data is stored in SigMF [64]. Each dataset consists of two files:

- `.sigmf-data`: a binary file containing complex I/Q samples for a given timestamp in 32-bit float format (little-endian).
- `.sigmf-meta`: a JSON file that provides metadata, including sample rate, center frequency, timestamp, and capture details such as the transmit waveform and sequence configuration.

The metadata file follows the SigMF specification and includes additional custom fields to describe:

- GPS coordinates and altitude of the UAV during each capture,
- sequence length and root index of the Zadoff-Chu waveform,
- measurement frequency and timestamp synchronization parameters,
- flight and experimental configuration details.

This format ensures reproducibility, interoperability with third-party tools, and ease of integration with standard signal processing pipelines. The GitHub repository also provides post-processing software, including Python scripts and a Jupyter notebook that operate on either the compressed `.npz`

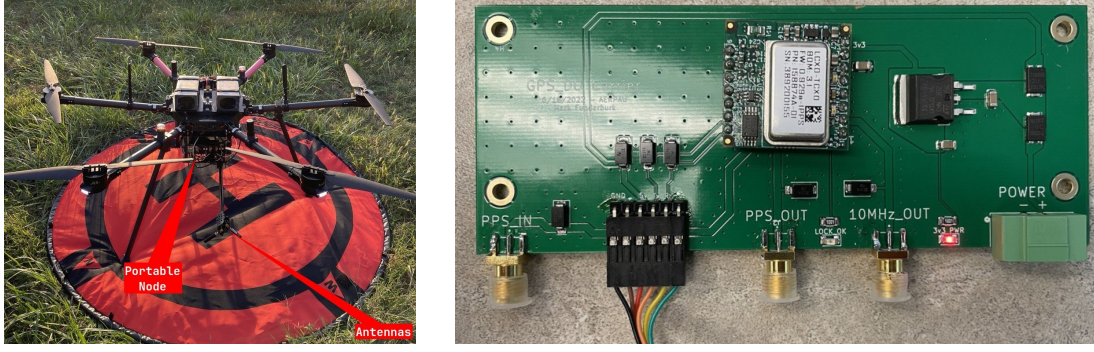
format or directly on the `.sigmf-data` files. Users can reproduce the published figures and analysis by running the provided `PostProcess.ipynb` notebook, which implements correlation-based channel response extraction and other visualization tasks.

C. Representative Results

This section presents representative results from channel sounding experiments conducted at the AERPAW LWRFL, a rural environment designed for wireless experimentation. Fig. 22a illustrates a representative CIR obtained using correlation-based processing of the received Zadoff-Chu sequences. The three-dimensional plot shows the evolution of multipath components over time during a UAV flight. The horizontal axis represents the delay (in microseconds), the vertical axis indicates the experiment time (indexed per snapshot), and the color scale denotes the received signal magnitude in dB. The black trace projected onto the back plane indicates the UAV's GPS distance from the fixed transmitter as a function of time. The variation in delay spread and path power over time reflects the dynamic nature of the A2G propagation channel, including the impact of UAV motion and altitude variation. Fig. 22b shows the corresponding path loss measurement over the full flight trajectory. The UAV's altitude profile is overlaid to highlight different flight phases, including takeoff, flight, and landing. The results demonstrate a clear relationship between the received power and UAV position, consistent with expected large-scale path loss behavior.

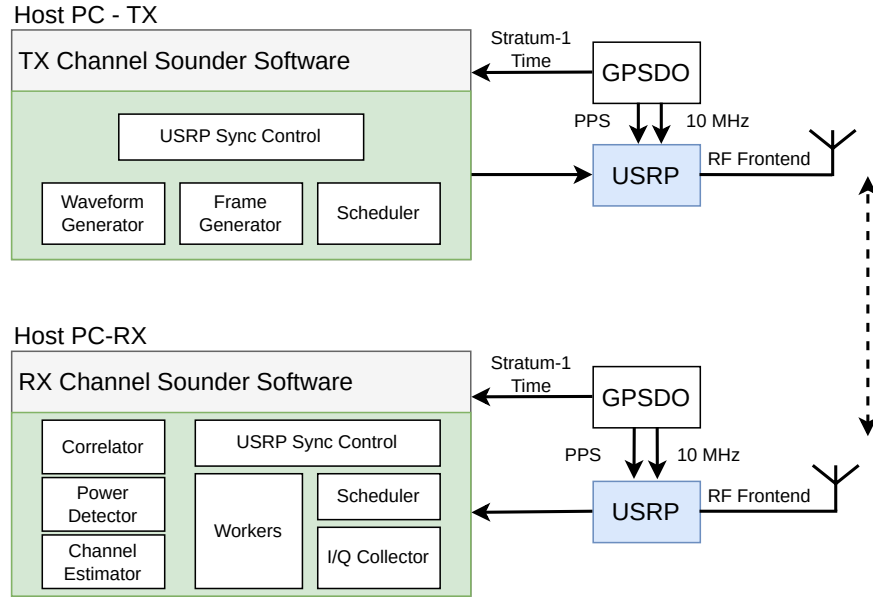
D. Possible Uses of Dataset

The provided dataset enables a wide range of research opportunities in the study and modeling of A2G wireless communication channels. Given the synchronized high-resolution measurements, as well as the availability of UAV position data, the dataset is well-suited for the following applications:



(a) Experiment setup for USRP B210 equipped portable node and UAV.

(b) Custom-designed GNSSDO system.



(c) Overall system architecture of channel sounder.

Fig. 21: Hardware and system architecture used for A2G multipath channel sounding measurements. (a) Field deployment of the portable measurement node based on USRP B210, mounted on a UAV platform for aerial data collection. (b) Custom-designed GNSSDO providing high-precision timing and synchronization. (c) Block diagram of the overall channel sounder system, highlighting key components and signal flow used during the A2G experiments [44].

1) *A2G Channel Modeling*: Researchers can extract key propagation characteristics such as delay spread, Doppler profiles, path loss exponents, and coherence bandwidth to develop statistical or geometry-based A2G channel models.

2) *Machine Learning for Wireless*: The dataset can be used to train and evaluate machine learning models for tasks such as channel estimation or link quality prediction in UAV-based networks.

3) *Impact of Altitude and Mobility*: With measurements conducted at multiple altitudes and along a defined flight trajectory, the dataset supports analysis of how UAV height and speed affect propagation conditions and coverage.

4) *Waveform and system Design*: The raw I/Q data and metadata can be used to simulate and evaluate the performance of waveform designs under realistic conditions.

5) *Validation of Ray-Tracing or Analytical Models*: The dataset provides empirical benchmarks that can be used to validate and calibrate ray-tracing simulations or analytical propagation models in rural environments.

X. WIRELESS LOCALIZATION DATASET

The decreasing cost of UAVs has led to rapid adoption across sectors such as defense, precision agriculture, aerial communications, search-and-rescue, and spectrum monitoring and enforcement. However, their growing presence introduces new challenges for security and airspace management, particularly as UAV activity increases around critical infrastructure. Effective UAV detection and tracking are essential to developing a secure UAV traffic management (UTM) ecosystem [73]. A range of studies have explored the use of RF signals

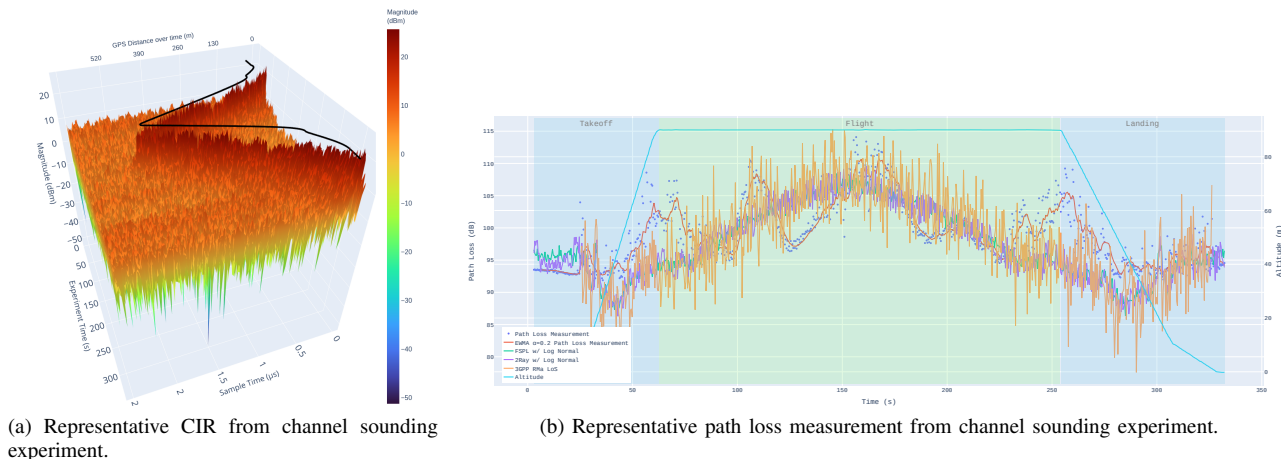


Fig. 22: Representative results from the A2G channel sounding campaign using a UAV-mounted transmitter. **(a)** Representative CIR illustrating multipath characteristics observed during flight. **(b)** Corresponding path loss measurement as a function of time, recorded at an altitude of 90 meters [44].



Fig. 23: Overview of the AERPAW testbed showing the RF sensor, UAV node, and LWRFL site layout [48].

for UAV detection, classification, localization, and tracking—including in scenarios involving non-cooperative or potentially malicious drones. Among the various RF-based approaches, Time Difference of Arrival (TDOA)—a multilateration technique that estimates UAV positions by measuring differences in signal arrival times at spatially separated sensors—has been shown to be particularly effective for passive RF sensing and UAV localization [74], [75]. To support further research, this section introduces two UAV localization and tracking datasets collected using the AERPAW testbed platform, providing researchers with valuable resources for evaluating TDOA-based tracking methods in real-world scenarios.

A. Description of Hardware and Software

UAV flight experiments were conducted in AERPAW's controlled environment to generate both datasets, using Keysight N6841A RF sensors to collect TDOA measurements. The N6841A (Fig. 23a) is capable of detecting, recording, and precisely time-stamping RF signals across a frequency range

of 20 MHz to 6 GHz, with a maximum bandwidth of 20 MHz. Equipped with a broadband omnidirectional antenna and GPS-based timestamping, the system uses Keysight's N6854A Geolocation Software and Sensor Management Tool to support TDOA, RSS, and hybrid localization methods for tracking RF sources within approximately a 2 km radius. However, the software is limited to 2D localization and does not estimate altitude.

Fig. 23c shows the deployment of these sensors at the LWRFL, where a single N6841A unit is mounted on each of the four towers labeled LW2 through LW5, approximately 10 meters above ground level. The sensors are deployed in a rural environment with mixed LoS conditions due to tree cover and building obstructions. All RF sensors within AERPAW are synchronized using a shared GPS-disciplined clock infrastructure and operate on the same local network, with individual IP addresses assigned within a common subnet. Centralized management via the Keysight Geolocation Server ensures network-level time synchronization, which is essential

for accurate TDOA-based localization. During experiments, the N6841A units capture I/Q data from RF signals, which are subsequently processed to estimate UAV positions. However, the specific algorithms used for TDOA extraction and position estimation are proprietary to Keysight and not publicly available.

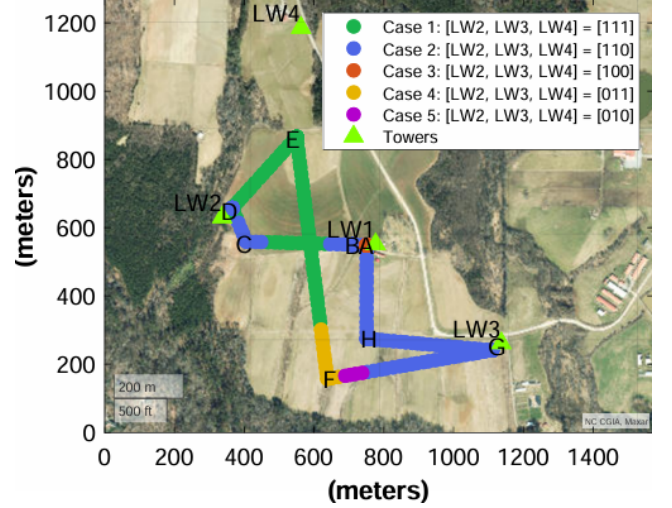
In our previous work [48], which focused on evaluating TDOA-based UAV localization accuracy under varying altitude and signal bandwidth conditions, measurements were compared to the Cramér-Rao Lower Bound (CRLB) to assess localization performance limits. In this study, a 3.32 GHz channel sounding waveform was transmitted from a UAV-mounted SDR (Fig. 23b) and RF sensors localized the UAV along multiple repeated trajectories. The dataset associated with this study is publicly available at [45]. Similarly, in a separate study by our group [47], which proposed an extended Kalman filter (EKF) framework to improve UAV tracking performance, TDOA-based localization was performed using downlink control signals from a DJI Inspire 2 UAV operating in the 2.400-2.483 GHz ISM band with a 20 MHz bandwidth. The dataset corresponding to this experiment is also publicly available [46].

B. Dataset Format

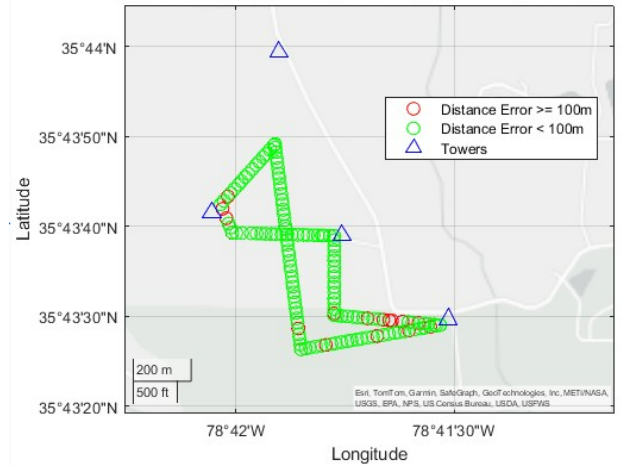
The dataset associated with [45], [48] is provided in the file `TDOA_UAV_Localization_Data.zip`. This archive contains one CSV file for each flight, with the altitude, signal bandwidth, and recording date indicated in each filename. For example, the file `40m_1.25MHz_7.15.24.csv` corresponds to a flight conducted at an altitude of 40 meters, using a 1.25 MHz signal bandwidth, on July 15th, 2024. The archive includes a total of five CSV files, each representing an individual flight. Each file contains the following recorded information:

- Center frequency of the signal.
- Latitude and longitude estimates for the UAV.
- Ground truth coordinates (latitude, longitude, and altitude) labeled with “GT”.
- RHO (degree of cross-correlation between received signals ranging from 0 to 1, where 1 indicates perfect correlation) and CEP (circular error probability) metrics for localization performance.
- Timestamps corresponding to each position estimate.
- A binary variable `LoStoLW2-5` indicating LoS status to AERPAW towers LW2 through LW5:
 - A value of `1` indicates LoS to the respective tower.
 - A value of `0` indicates NLoS to the respective tower.

If both the latitude and longitude values for a given measurement are recorded as `0`, this indicates that the geolocation software failed to produce a valid position estimate. The dataset archive also includes a MATLAB helper script, `KeysightRTDOALocalizationforFlights.m`, which assists in processing TDOA-based UAV localization data from a selected CSV file and generates visualizations and performance metrics for evaluating UAV position estimation accuracy.



(a) UAV trajectory at 40 meter altitude color-coded by LoS conditions to towers LW2-4 [48].



(b) Ground truth UAV trajectory color-coded by localization error.

Fig. 24: UAV trajectory and localization performance at 40 m altitude. (a) UAV flight trajectory color-coded by LoS visibility to cellular towers LW2 through LW4, based on geometry and terrain data. (b) Ground truth trajectory of the UAV color-coded by the magnitude of localization error, highlighting spatial variation in positioning accuracy.

The dataset associated with [46], [47] follows the same general format but differs in two key ways. First, the dataset for [46], [47] does not include LoS indicator variables. Second, the data is split into two separate CSV files: `Inspiron_backup.csv`, containing the field measurements collected by the Keysight RF sensor, and `GPS_Flight1_backup.csv`, containing the ground truth trajectory data recorded using a GPS application.

C. Representative Results

Representative examples from the collected datasets are presented in this section to demonstrate typical localization behavior and characteristics.

Fig. 24a illustrates the UAV’s 40-meter altitude trajectory overlaid on the LWRFL site map, showing LoS conditions relative to three RF sensor towers (LW2, LW3, and LW4),

which are indicated by green triangles. The trajectory is color-coded based on five distinct LoS scenarios, where each case indicates LoS or NLoS status relative to the three towers. For example, Case 1 (green) represents segments where the UAV maintained LoS to all three towers, while Case 5 (purple) corresponds to areas where only LW3 maintained LoS. The color-coding of the trajectory clearly highlights transitions between LoS conditions as the UAV moves through the field, illustrating the mixed LoS/NLoS environment characteristic of the LWRFL site. This visualization provides important context for interpreting localization performance, as LoS availability directly influences TDOA measurement quality and positioning accuracy.

Fig. 24b presents a color-coded visualization of localization error along the UAV's trajectory. Each point represents a ground truth position of the UAV, with circles indicating the corresponding localization error relative to the nearest estimated coordinate. Green circles mark locations where the distance error is less than 100 meters, while red circles identify points where the error exceeds 100 meters. Blue triangles indicate the locations of the RF sensor towers used for TDOA-based localization.

As shown in the figure, the majority of the trajectory is associated with low localization error (green), suggesting consistent and accurate TDOA-based position estimates along most of the UAV's flight path. Higher error regions (red) are concentrated near certain trajectory segments corresponding to NLoS conditions or poor geometric dilution of precision relative to the sensor towers. This visualization highlights how mixed LoS/NLoS conditions and sensor geometry influence localization performance throughout the trajectory.

D. Possible Uses of Dataset

The UAV TDOA localization datasets serve as a valuable resource for advancing research in RF-based localization, UAV tracking, and passive sensing systems. Its real-world measurements, mixed LoS/NLoS conditions, and ground-truth references enable exploration of the following research directions:

1) *Tracking Filter Development/Evaluation*: The dataset can support the development and evaluation of tracking filters, such as Kalman or particle filters, by providing real-world UAV trajectory and localization measurement sequences for state estimation under mixed LoS/NLoS conditions.

2) *TDOA Localization Algorithm Benchmarking*: Supports performance evaluation and comparison against Keysight's industry-standard TDOA-based localization methods using real-world RF measurements collected under mixed LoS/NLoS conditions.

3) *Sensor Fusion*: Enables the development and evaluation of multi-sensor fusion algorithms by combining TDOA-based RF localization data with complementary modalities such as inertial, visual, or radar measurements for improved UAV tracking and state estimation.

4) *NLoS Modeling and Analysis*: Supports characterization and modeling of NLoS conditions by providing real-world localization errors and LoS/NLoS labels, enabling research

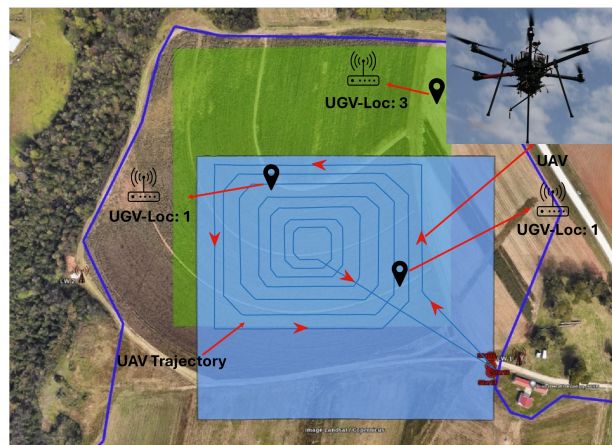


Fig. 25: AFAR Challenge setup highlighting the UAV flight zone (blue) and the designated area where the UGV could be hidden (green). The UAV used during the experiment is shown in the upper-right corner [49].

into NLoS detection, mitigation, and bias-aware localization techniques.

5) *CRLB Modeling and Validation*: Enables theoretical performance analysis through CRLB modeling and supports empirical validation by comparing theoretical bounds against real-world TDOA-based localization errors under varying sensor geometries and environmental conditions.

Together, these research directions highlight the dataset's value as a resource for advancing localization, tracking, and sensing technologies in real-world UAV applications.

XI. UAV-COLLECTED RSS MEASUREMENTS FOR RF SOURCE LOCALIZATION

The AERPAW Find A Rover (AFAR) Challenge [76] was a national-level competition designed to promote research in UAV-assisted RF localization. Organized under the AERPAW testbed, the challenge aimed to accelerate innovation by providing a standardized experimental environment for evaluating RF localization algorithms using UAVs. In this competition, UAVs were deployed to locate an RF-emitting unmanned ground vehicle (UGV) based solely on signal measurements. In the AFAR Challenge, the UGV could be placed anywhere within a designated search area (marked in green in Fig. 25), while the UAV was restricted to fly in the flight zone (marked in blue in Fig. 25). Teams were free to design either autonomous or fixed waypoint-based UAV trajectories to locate the UGV, with flight constraints of 20-110 meters altitude and speeds up to 10 m/s. The challenge consisted of two phases: development in a DT environment, and deployment in a real-world testbed at LWRFL, NC, as detailed in [49]. Each of the five finalist university teams independently devised UAV flight trajectories and localization algorithms as part of the competition.

A. Description of Hardware and Software

The AFAR dataset includes data collected from both a DT simulation environment and a real-world wireless testbed, both deployed via the AERPAW platform. Each experimental

run involved a UAV serving as the receiver and a UGV operating as the RF signal transmitter. The UAV and UGV were equipped with portable SDR nodes based on the USRP B205mini, capable of full-duplex operation over 70 MHz to 6 GHz.

Each SDR was connected to an Intel NUC 10 mini-PC equipped with an i7-10710U processor, 64 GB of RAM, and a 1 TB SSD, enabling real-time onboard signal processing. Transmissions employed a GNU Radio-based channel sounder that used a degree-12 Galois LFSR to generate a pseudo-random bit sequence (PRBS) of length 4095. The sequence was interpolated, pulse-shaped using a root-raised cosine filter, and transmitted at 2 MHz sampling rate over 3.0–4.2 GHz using a wideband antenna. On the receive side, frequency offset correction and correlation with the original PRBS enabled the extraction of CIR, from which RSS and received signal quality (RSQ) values were derived.

The DT environment mirrored the physical setup using containerized software emulation. A virtual USRP (V-USRP) and a channel emulator VM (CHEM-VM) simulated RF propagation based on real-time UAV-UGV position updates. Experiment logic ran in Experiment VMs (E-VMs), and UAV mobility was emulated using Software-In-The-Loop (SITL) vehicles, all orchestrated through AERPAW’s geofencing and control interfaces.

B. Dataset Format

The AFAR dataset is organized to reflect the structure of the experiments conducted during the challenge. For each of the five finalist teams, data was collected across three distinct UGV placements (Loc-1, Loc-2, Loc-3), and in two environments: the DT simulation and the real-world AERPAW testbed. This results in a total of 30 experiments (5 teams $\times$ 3 locations $\times$ 2 environments), each stored in a well-defined folder hierarchy.

The dataset is structured as follows:

- Top-level directory: Team identifier (e.g., `288`, `300`, `301`, etc.).
- Subfolders: `development` (for DT data) and `testbed` (for real-world data).
- Location-specific folders: `loc-1`, `loc-2`, `loc-3`.

Each location folder within `testbed` contains the following core files:

- `power_log.txt`: Table II shows a snippet from the `power_log.txt` file, where each row represents a RSS measurement captured by the UAV receiver. The first column contains precise timestamps (in microseconds), and the third column provides the measured RSS in dB. Although the raw file also includes a middle column for sample indexing, it is of no use for signal analysis.
- `quality_log.txt`: Has the same format as Table II, but contains RSQ values instead of RSS.
- `log.csv`: Contains time-synchronized UAV navigation data, including GPS coordinates (latitude, longitude, altitude), speed, heading, and satellite metadata. A sample snippet of this file is illustrated in Fig. 26.

TABLE II: Sample Format of `power_log.txt`

Timestamp	Index	RSS (dB)
2023-12-13 13:45:34.041027	0000000	-34.9675
2023-12-13 13:45:34.072347	0000004	-40.7695
2023-12-13 13:45:34.105068	0000010	-48.0318

TimeUS	Status	GMS	GWK	NSats	HDop	Lat	Lng	Alt	Spd	GCrs	VZ	Yaw	U
2166080919	5	3.27E+08	2292	14	0.77	35.727371	-78.6962127	112.65	0.03985	72.47443	0.007	0	1
2166260903	5	3.27E+08	2292	14	0.77	35.727371	-78.6962128	112.66	0.02624	261.8699	-0.059	0	1
2166460908	5	3.27E+08	2292	14	0.77	35.7273709	-78.6962131	112.66	0.04639	172.5686	0.007	0	1
2166660908	5	3.27E+08	2292	14	0.77	35.727371	-78.6962132	112.65	0.061294	5.61758	0.21	0	1
2166860908	5	3.27E+08	2292	14	0.77	35.727371	-78.6962131	112.66	0.023067	252.3409	-0.189	0	1
2167060890	5	3.27E+08	2292	14	0.77	35.7273709	-78.6962132	112.68	0.011662	329.0363	-0.148	0	1
2167260929	5	3.27E+08	2292	14	0.77	35.7273708	-78.6962132	112.69	0.008485	315	-0.13	0	1

Fig. 26: Sample entries from `log.csv` showing UAV navigation and GPS metadata.

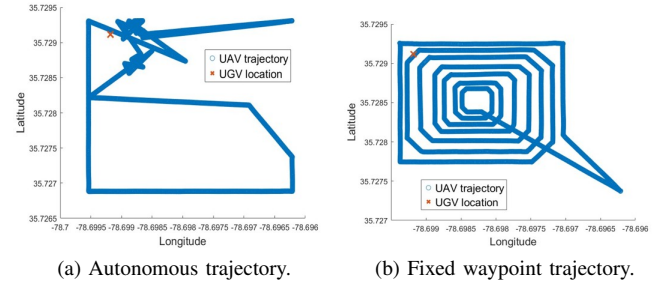


Fig. 27: Comparison of UAV trajectories: autonomous (Team-300) vs. fixed waypoint (Team-309).

- `angles.mat`: Provides the UAV’s orientation in terms of roll, pitch, and yaw.

C. Representative Results

The AFAR dataset reflects diverse RF and mobility dynamics, with teams employing distinct UAV trajectories—three teams using autonomous trajectories and two adopting fixed waypoint trajectories. Example trajectories for Team-300 (autonomous) and Team-309 (fixed) are shown for comparison in Fig. 27.

The measurements in the AFAR dataset exhibit considerable variability resulting from differences in UAV trajectories, UGV placements, and environment-specific propagation effects. KPIs such as RSS, RSQ, and UAV motion parameters (e.g., speed and altitude) vary significantly across locations. For instance, the UAV speed and altitude profile for Team-309, who employed a fixed waypoint trajectory, are shown in Fig. 28a. As the UAV approaches each waypoint, its speed increases, then decreases upon arrival, before accelerating again toward the next target. This cyclical speed pattern is characteristic of waypoint-based navigation. Additionally, the UAV maintains a relatively constant altitude of approximately 30 meters throughout the mission.

Fig. 28b presents RSS heatmaps overlaid on the UAV flight paths. The diamond marker indicates the RF source location. As expected, RSS is strongest when the UAV is in close proximity to the source and weakens with increasing distance. However, the spatial distribution of received power is not uniform, reflecting the influence of multipath and shadowing effects.

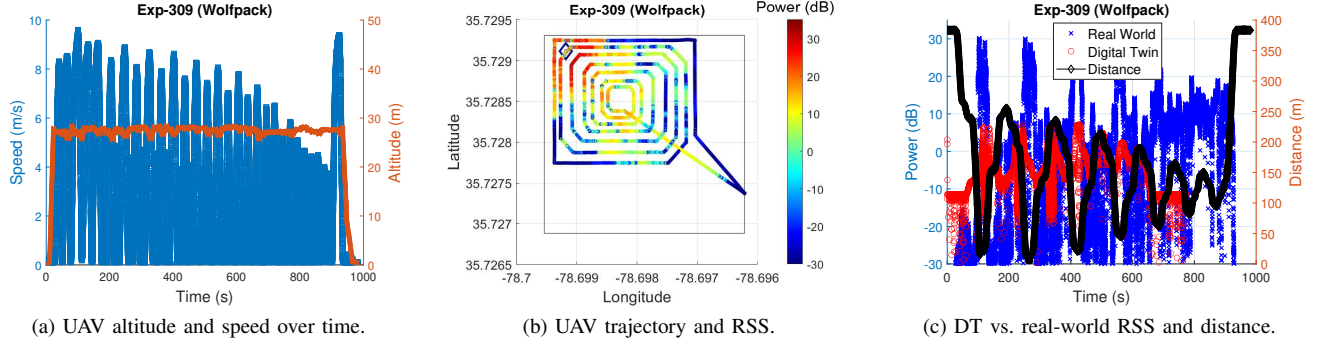


Fig. 28: Representative signal and mobility characteristics from the AFAR dataset across different teams and environments [49].

To highlight the contrast between simulated and real-world signal behavior, Fig. 28c plots RSS and distance against time for both DT and real-world environments. In the DT environment, RSS remains relatively smooth and predictable. In contrast, the real-world data exhibits significant fluctuations, even at similar distances, due to dynamic factors such as fading, body blockage, and environmental clutter. This discrepancy underscores the importance of accounting for real-world propagation effects when developing and validating RF localization algorithms.

D. Possible Uses of Dataset

The AFAR dataset serves as a comprehensive resource for advancing research in wireless communications, RF-based localization, and UAV-enabled signal intelligence. Its rich content and dual-environment structure (DT and real-world) enable the following research directions:

1) *A2G Channel Propagation Modeling*: The dataset enables realistic modeling of A2G wireless channels, accounting for UAV-specific factors such as altitude, elevation/azimuth angles, velocity, and orientation (roll, pitch, yaw). These measurements help characterize the propagation environment under mobility and elevation diversity [50].

2) *Antenna Gain and Shadowing Analysis*: The dataset facilitates the evaluation of directional antenna performance and gain variations due to UAV-body shadowing. This is particularly relevant for understanding signal attenuation in NLoS conditions and UAV maneuvers.

3) *Performance Benchmarking*: Researchers can perform comparative analysis of UAV-assisted RF localization algorithms. The dataset allows assessment of trajectory efficiency, signal quality, and overall localization accuracy under controlled and real-world constraints.

4) *Data-Driven Localization Algorithms*: The time-synchronized RF and positional data provide a robust basis for training and evaluating machine learning models for RF source localization, including regression, classification, or hybrid approaches.

5) *Signal Strength and Fading Prediction*: Deep learning models can be trained to forecast RSS/RSQ values under varying mobility conditions. This supports proactive planning in UAV-assisted sensing and communication tasks.

6) *Flight Path Optimization*: Using the dataset, UAV trajectories can be optimized for better link reliability, minimal energy use, or improved localization precision, either via reinforcement learning or optimization-based techniques.

7) *Simulation-to-Reality Transfer Learning*: The paired DT and real-world measurements enable transfer learning strategies that improve model robustness across synthetic and physical environments.

In summary, the AFAR dataset bridges multiple disciplines, offering a reproducible platform to study wireless localization, adaptive mobility strategies, and signal-aware autonomy in UAV networks.

XII. UAV SIGNAL CLASSIFICATION DATASET

In recent times, malicious UAVs have become a global threat to society. Even in modern warfare, the use of UAVs has altered the dynamics of traditional military operations, providing strategic advantages to state actors as well as established military forces. Besides, low cost, low altitude and low speed consumer UAVs (or micro-UAVs) pose a unique threat to both military assets and civilians. As a result, researchers have been investigating different techniques for UAV identification. Some of these techniques include RF, radars, computer vision (optical and infra-red cameras), high-energy lasers, and acoustic techniques [77], [78]. Each of these techniques has its own advantages and challenges.

To achieve long range detection, identification, and improved localization of a wide-range of UAVs, RF-based techniques are commonly preferred. In addition, RF-based techniques for detecting and identifying UAVs can operate in all weather conditions. As a result, UAV detection, identification (classification), and neutralization using RF-based electronic warfare (EW) and signals intelligence (SIGINT) systems are becoming popular. These techniques exploit electromagnetic spectrum or directed energy to detect, identify, and interdict an incoming drone. Consequently, these systems are comparatively more effective than alternative detection approaches (e.g. camera and acoustics) because they can operate in all weather conditions and achieve long detection ranges [77]. However, due to the ubiquitousness of electronics and communication systems, especially in the ISM band in urban centers, it could be difficult to accurately detect and identify specific

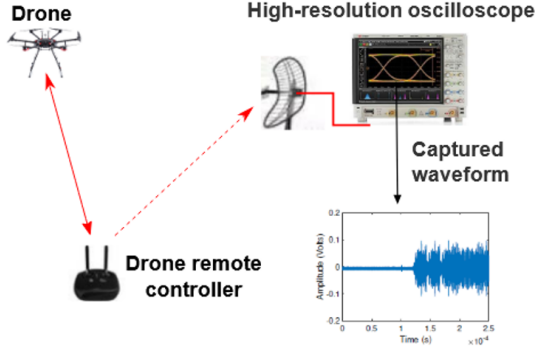


Fig. 29: The experimental setup for detecting and capturing RF signals from UAV controllers. The input of the receiver has an LAN, a bandpass filter and a parabolic antenna receiver [55].

UAV signals in the presence of intentional jammers and non-intentional EM radiators/interference. To mitigate this issue, researchers at AERPAW recognized the need to develop a dataset of unique UAV RF signals extracted from popular commercial UAVs. This dataset could be used to develop signal processing algorithms and machine learning models that can improve the detection and identification of specific consumer UAVs in the presence of interference.

A. Description of Hardware and Software

For the data collection, we designed a simple passive RF surveillance receiver that continuously listens to RF signals in the environment and saves the data for further processing. The RF signals captured are time-varying modulated signals from popular commercial UAV controllers.

The experimental setup is shown in Fig. 29. In the figure, a 24 dBi parabolic antenna operating in the 2.4 GHz ISM band listens to RF signals from UAV controllers, which are used to send control and navigation commands to an incoming UAV. The choice of the parabolic antenna is to further increase the range of the receiver because of the improved directivity (gain) of the antenna. The output signal from the antenna is fed to the receiver through an RF chain that combines a LNA and an RF bandpass filter. The LNA is used to amplify weak signals, while the bandpass filter ensures the input signal is band-limited so as not to saturate the receiver. Receiver saturation will cause non-linearity effects like intermodulation distortions (IMD), gain compression, and loss of signal integrity after amplification. For the experiment, the receiver is a high-resolution mixed signal Keysight oscilloscope (6 GHz Keysight MSOS604A) capable of sampling the captured signal at 20 GSa/s. This high sampling rate ensures the detection system captures all the transient features or fingerprints of the captured UAV RF controller signals. Also, within the receiver in Fig. 29, there is a custom MATLAB script for signal detection, data preprocessing and transformation of the raw signal.

Fig. 30 shows samples of UAV RF remote controller (RC) signals and UAV emitted signals captured using the detection systems in Fig. 29 [55]. From this figure, it is obvious that each of the UAV RF signal have some unique

features/waveform that can be exploited by an intelligent signal recognition system to identify the UAV or its controller. Moreover, using advanced signal localization algorithms with multiple RF receivers, the UAV and its controller can be localized and monitored.

TABLE III: Drone catalogue for the classification signal dataset.

Make	Model	Make	Model
DJI	Inspire 1 Pro	Spektrum	DX5e
	Matrice 100		DX6e
	Matrice 600		DX6i
	Phantom 4 Pro		JR X9303
	Phantom 3		
Futaba	T8FG	Graupner	MC-32
HobbyKing	HK-T6A	FlySky	FS-T6
Turnigy	9X	Jeti Duplex	DC-16

TABLE IV: Metadata of UAS signal classification dataset.

Description	Value
Number of drone controllers	17
Sampling frequency	20 GSa/s
Center frequency	2.4 GHz
Number of signals/drone RC	~1000
Number of samples/signal	5 million
Time duration/signal	0.25 ms
Average data size/signal	7 MB
Dataset size	124 GB
Data format	.mat

B. Dataset Format

The UAS signal classification dataset consists of RF signals from 17 popular commercial UAV controllers (15 of which are unique) from eight different manufacturers [54]. The UAV controller catalogue includes DJI, Futaba, Hobby King, and Turnigy as shown in Table III. For each UAV RF signal category in the dataset, Table IV provides basic information or metadata of each UAV RF samples in the UAS signal classification dataset. To extract and visualize a specific UAV RF signal and its metadata from the dataset, we need to create a MATLAB object instance, as in object-oriented programming. Also, to perform any experiment using the dataset, we can extract all or part of the dataset by creating a custom database.

The dataset is organized into processed CSV files for wire-less network analysis and comprises three primary directories:

- 1) **Creating a UAV Object to Visualize RF signals:** To extract UAV RF signal metadata from a .mat file in the dataset, we use a MATLAB script (*droneRC.m*) that is provided with the UAS signal classification dataset. The *droneRC.m* script defines a MATLAB class, and several methods/functions are defined for creating a drone RF object (specific UAV RF signal) as shown in Fig. 31.
- 2) **Creating a Database for Experiment:** To create a database of UAV RF signal from the UAS signal classification dataset, we use the createDatabase function. This function can generate a database of drone controller RF signals in both matrix and table format. Using the function, you can specify the number of signals, features

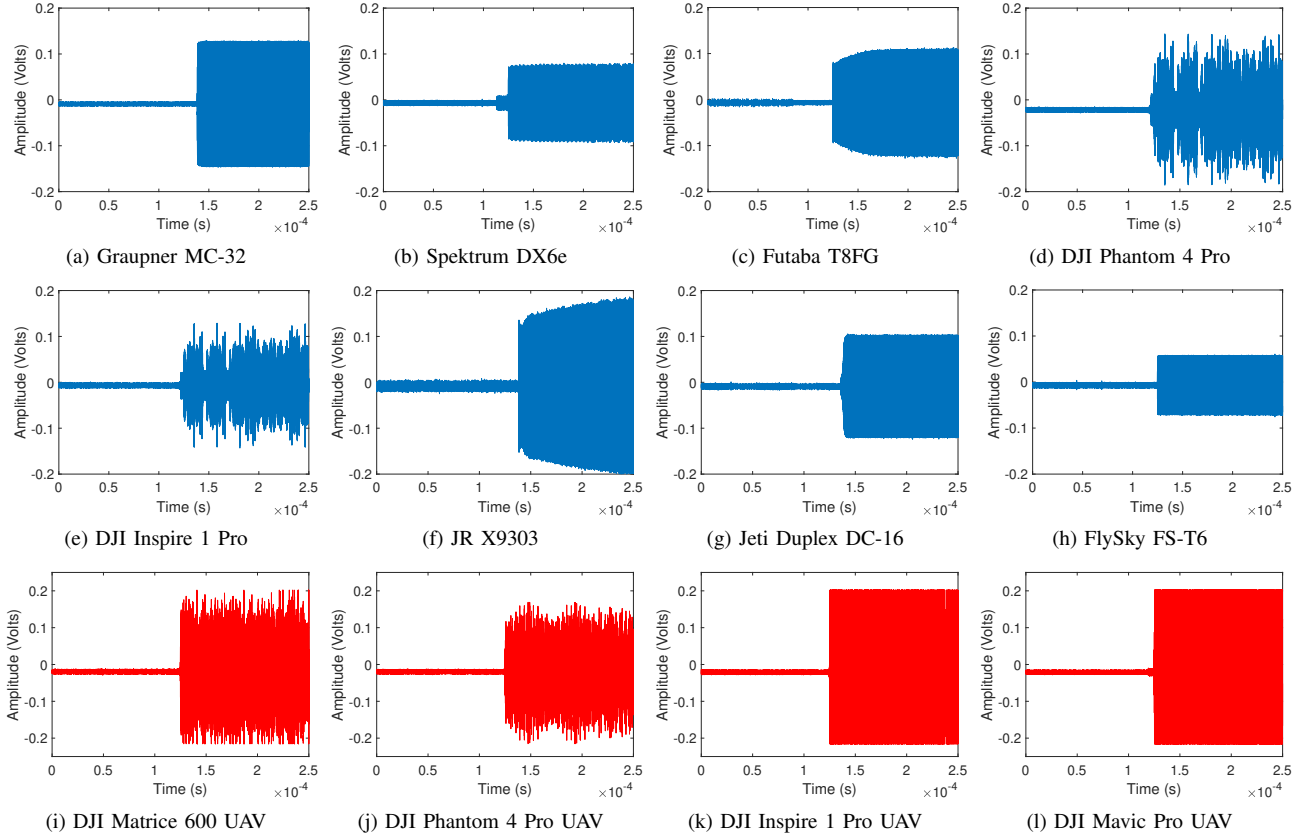


Fig. 30: This shows a sample of RF signal captured from eight different UAV controllers and four different UAVs while on flight: (a) Graupner MC-32, (b) Spektrum DX6e, (c) Futaba T8FG, (d) DJI Phantom 4 Pro, (e) DJI Inspire 1 Pro, (f) JR X9303, (g) Jety Duplex DC-16, (h) FlySky FS-T6, (i) DJI Matrice 600 UAV, (j) DJI Phantom 4 Pro UAV, (k) DJI Inspire 1 Pro UAV, (l) DJI Mavic Pro [55].

droneRC

```

droneRC
droneRC is used to create a drone remote controller (RC) object from a
.mat file in the MPACT drone RC RF database.

droneRC metadata properties:
Make
Model
Index          - Index of the RC in the corresponding folder
RawData        - Digitized RF signal
ScaleFactor    - Scale factor to convert the digitized signal to volts
NumSamples     - Number of samples in the captured signal
Duration       - Time span of the captured raw signal
Fs             - Sampling frequency
CroppedData    - A portion of the raw signal that includes the signal transient
                (to be used instead of raw signal to avoid memory problems)
Features       - A structure array of extracted features

droneRC methods and functions:
droneRC        - Create a droneRC object from the database
getEnergyTrajectory - Extract normalized energy trajectory from the
                    spectrogram of the cropped RF signal
findTransientStart - Find the start point using Higuchi's method
plot           - Plot the captured RF signal (or the cropped
                portion) and the spectrogram

```

Fig. 31: The *droneRC.m* class is used to extract both RF signal data and associated metadata from specific .mat files in the UAS signal classification dataset [54].

to include, and the format for the database as shown in Fig. 32.

C. Representative Results

The UAS signal classification dataset has been used to validate several UAV detection and classification algorithms. The results from these works have been published in the

createDatabase

```

createDatabase creates a database of drone remote controllers (RCs)

db = createDatabase(databasePath,Fs,NumSignals,'OnlyRawData') creates a
database db of drone RC objects that includes only the raw data and the
basic properties of each drone RC in the subfolders of the specified
path. Fs is the sampling frequency. NumSignals is the desired number of
signals from each drone RC folder.

db = createDatabase(...,'OnlyCroppedData') creates a database db that
includes only the cropped data and the basic properties to save space.

db = createDatabase(...,'IncludeFeatures') creates a database db that
contains all the available data.

db = createDatabase(...,'Table') creates a database db in table format.

```

Fig. 32: The *createDatabase.m* function is used to create a table or matrix of droneRC objects from the UAS signal classification dataset. The database generated can be used to benchmark different UAV RF signal detection and classification algorithms in the presence or absence of interference [54].

literature [55]–[57]. In [56], the authors described a process for detecting and identifying Micro-UAVs using data extracted from the UAS signal classification dataset. The detection phase is based on the naïve Bayes approach using Markov models. Once the UAV signals have been detected, the identity of the UAV is determined (or classified) using classical machine learning algorithms. The authors showed that using features selection techniques such as neighborhood component analysis

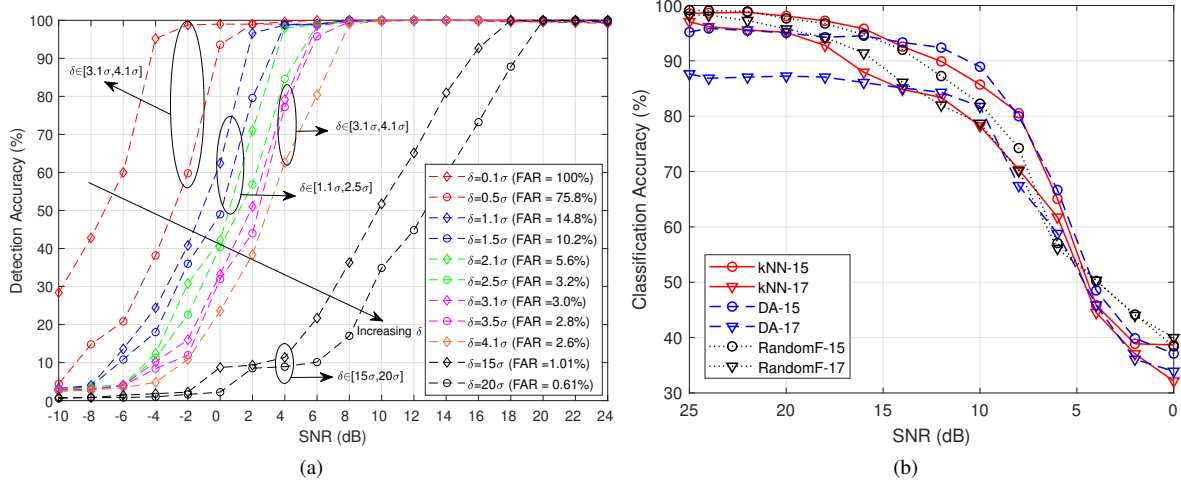


Fig. 33: Representation Result for UAV Detection and Classification Using the UAS signal classification dataset: (a) The performance of the proposed detection systems as a function of the SNR, the detection threshold (δ), and the fixed FAR, (b) The performance of the ML-based model as a function of the SNR, the number of UAV RF controller signal category extracted from the UAS signal classification dataset [55].

(NCA), the kNN machine learning algorithm achieved the highest classification accuracy of 96.3% at an SNR of 25 dB. In addition, the study showed that as the SNR reduces, simulating a drone moving farther away from the detection system, the detection and classification of the machine learning algorithms reduces. The limitation of this study is the absence of interference signal from the environment. In [55], the authors extended the study in [56] to include the presence of wireless interference signals from Wi-Fi and Bluetooth enabled devices. Fig. 33 shows the performance of the detection and classification system as a function of the SNR.

From Fig. 33a, we see that for a fixed false alarm rate (FAR), increasing the detection threshold (δ) will reduce the detection accuracy. Also, from Fig. 33b, we see that the accuracy of the classification system depends on the machine learning model deployed, the SNR, and the number of UAV classes in the database. Once again, the study shows that it is possible to achieve an accuracy of 98.13% in classifying 15 different UAV controllers using classical machine learning models like kNN and random forest. In [57], the authors investigated the impact of using the convolutional neural network (CNN) to classify/identify UAV RF controller signals. The CNN models are trained using spectrogram images representation of the raw UAV controller RF signals. The CNN model achieved an accuracy of about 92%.

D. Possible Uses of Dataset

Our UAS signal classification dataset can be used for many practical applications. They include the following:

1) *Benchmarking New Detection and Classification Algorithms*: As researchers continue to work in this domain, it is important to benchmark or compare the performance of their algorithms with alternative classical models in literature. In such a study, it is important to ensure the UAV RF data used were reliably captured in the same environment. Moreover, the

performance of different algorithms as a function of thermal noise or antenna temperature can be studied comparatively.

2) *Investigating the Impact of Intentional or Non-intentional Radiators*: Researchers can use the UAS signal classification dataset to investigate the impact of intentional interferences like jamming signal on the performance of UAV RF-based detection and classification algorithms. This is important because of the array of new and existing UAV countermeasure EW systems. Also, as described in [55], researchers can use the UAS signal classification dataset to investigate the impact of in-band and out-of-band non-intentional interference signals like Wi-Fi, BLE, Zigbee, medical equipment, microwave oven, and others. This is important for electromagnetic compliance (EMC) certification especially in shared unlicensed spectrum.

3) *Generating Synthetic UAV RF Dataset*: Using the UAS signal classification dataset, researchers can generate synthetic UAV RF signals. For example, using generative adversarial networks (GANs) or variational autoencoders (VAEs), we can create large synthetic data using the UAS signal classification dataset as input. The synthetic data can be used to train more robust machine learning models, increasing the generalization and accuracy of such models. Other methods that can be used to create synthetic UAV controller RF data include data augmentation techniques (translation and rotation), kernel density estimation, parametric statistical models, random sampling and Monte Carlo simulation. As a first principle, all these methods will need an input dataset like the UAS signal classification dataset for training.

XIII. UAV TRAJECTORY, RSRP, AND THROUGHPUT DATASET IN EMULATED AND SIMULATED ENVIRONMENTS

UAV flight path data, including throughput and RSRP readings, are crucial to developing, experimenting with, and evaluating next-generation wireless networks and DT solutions. Our

datasets capture critical A2G propagation and link characteristics based on trajectory, as well as interactions with emulated and simulated environments. The advantage of the emulation environment is that it implements full-stack software as it runs in real-world SDRs (such as USRPs) and UAVs; on the other hand, it can be relatively slow and computationally intensive for development purposes. Simulation abstractions offer the advantage of accelerating the development of UAS and radio algorithms that can subsequently be evaluated in emulation and real-world testbeds, though this comes at the expense of reduced realism.

A. Description of Hardware and Software

The measurements were obtained through a combination of MATLAB simulation and DT emulation [60], utilizing AERPAW's cutting-edge experimental facilities at Lake Wheeler. The major components are as follows.

The UAV testbed includes an LTE SISO radio-equipped UAV, GPS receivers, and flight controllers that ensure accurate trajectory tracking. Field experiments were conducted in the Lake Wheeler area, where high-fidelity RSRP measurements were collected from four BSs (LW1–LW4). A MATLAB-based simulation environment was developed to model UAV flights in virtual settings, with field measurements used to validate the simulations under realistic conditions. The AERPAW DT further supports emulation of UAV-to-base-station communications, providing parameterized control over mobility patterns, radio environments, and network conditions.

There are custom MATLAB and Python scripts for data processing, analysis, and visualization. The full dataset, including all processing and simulation code, is made publicly available at [58], [59] to facilitate reuse as easily and transparently as possible.

B. Dataset Format

The dataset is organized into processed CSV files for wireless network analysis and comprises three primary directories:

- 1) **Emulated RSRP Measurement Using Fixed Trajectory:** Emulated data for the UAV flights on designed trajectories are found in the `fixed_trajectory_rsrp_emulation` folder. The dataset contains timestamps, altitude, latitude, longitude, and RSRP measurements, along with pitch, roll, and yaw.
- 2) **Simulated RSRP Measurement Using Fixed Trajectory:** Simulated data for the UAV flights on designed trajectories are found in the `fixed_trajectory_rsrp_simulation` folder. The dataset contains timestamps, altitude, latitude, longitude, and RSRP measurements.
- 3) **Simulated RSRP Measurement Using Autonomous Trajectory:** Simulated data for the UAV flights on autonomous trajectory are found in the `autonomous_trajectory_rsrp_simulation` folder. The dataset contains timestamps, altitude, latitude, longitude, and RSRP measurements.

TABLE V: UAV measurement dataset sample with RSRP, SNR, and data rate.

Time (s)	Lon. (deg)	Lat. (deg)	Alt. (m)	RSRP (dBm)	SNR (dB)	Rate (Mbps)
25-03-28 03:36:50	-78.69627	35.72748	0.00	-53.00	37.00	7.35
25-03-28 03:36:51	-78.69627	35.72748	0.47	-53.00	37.00	7.35
...	...	...	...	...	...	...
25-03-28 03:37:05	-78.69627	35.72748	30.0	-54.00	36.00	7.29
25-03-28 03:37:06	-78.69622	35.72748	30.0	-53.00	37.00	7.35
25-03-28 03:37:07	-78.69615	35.72745	30.0	-52.00	38.00	7.40
...	...	...	...	...	...	...

Table V provides the structure of the dataset used in this section. Each row of the table corresponds to a specific timestamp in a UAV flight and records the UAV's position (altitude, longitude, and latitude), RSRP, SNR, and data rate.

To enable effective visualization and analysis, the dataset includes a collection of post-processing scripts. For scenarios with fixed UAV trajectories, scripts such as `plot_rsrp_emulation.m` and `plot_rsrp_distance_emulation.m` are used to graphically represent RSRP measurements as a function of UAV location and distance from a specified BS, respectively. The same set of plotting conventions is followed for the simulated dataset as well, with names or identifiers of scripts distinguishing simulation and emulation results. For the autonomous path scenario, `plot_rsrp_simulation.m` generates RSRP plots in dynamic UAV trajectories, and the Python scripts `plot_throughput_distance_vs_time_lw34.py` and `plot_throughput_distance_vs_time_lw12.py` provide distance vs. time plots and throughput for each BS pair, respectively. These post-processing scripts streamline the workflow, enabling reproducible preparation of all key figures and statistical analysis presented in this paper.

C. Representative Results

This part discusses analysis and results based on the dataset, particularly in terms of comparing simulation with emulation and the effect of trajectory design:

RSRP for Individual BSs: Here, we present the emulation RSRP measurements for the fixed trajectory in Fig. 34. When the UAV is closer to a BS, the signal strength tends to be stronger; e.g., we observe a strong signal strength in Fig. 34(a) because the UAV is close to the LW1 compared to other towers. Results in simulated environment (not provided here due to space reasons) show a high similarity in signal strength with the emulation results.

Fig. 35 shows the RSRP measurement by UAV to individual BSs for autonomous trajectory in simulation. We see a strong RSRP near all BSs except LW4 because UAV goes near LW1, LW2, and LW3, but LW4 is far away, and the UAV is restricted to the geofence. The trajectory details (fixed and autonomous) and other results, such as distance versus RSRP measurements and throughput measurements, are presented in [60].

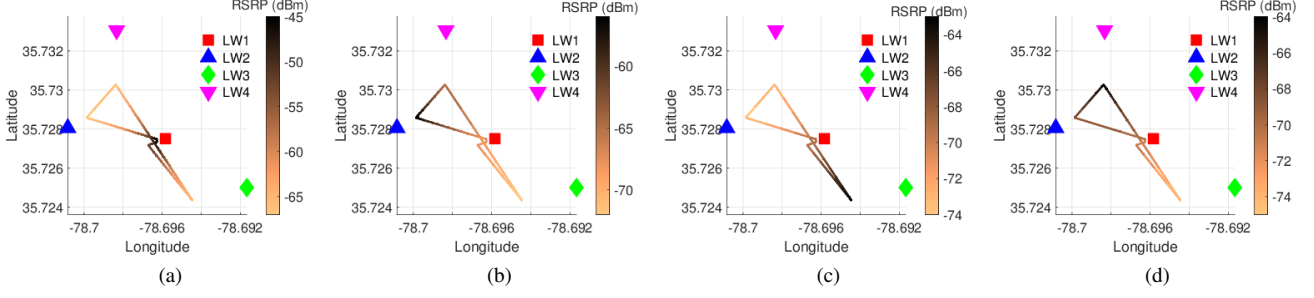


Fig. 34: RSRP measurement with respect to (a) LW1, (b) LW2, (c) LW3, and (d) LW4 for fixed trajectory in emulation [60].

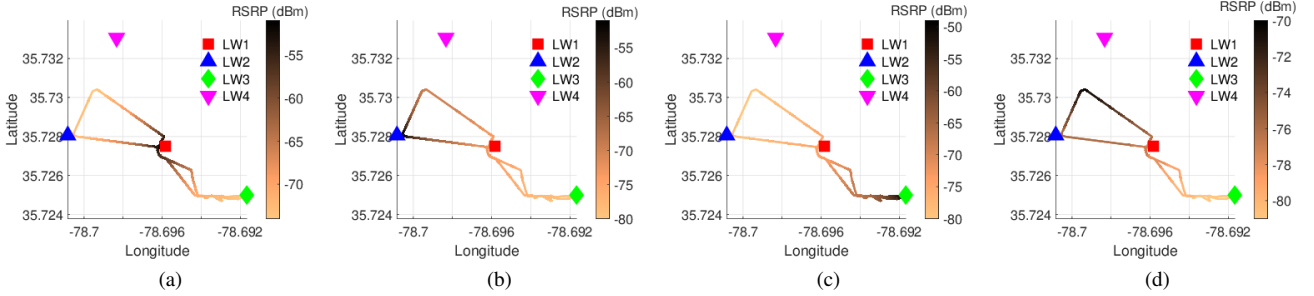


Fig. 35: RSRP measurement with respect to (a) LW1, (b) LW2, (c) LW3, and (d) LW4 for autonomous trajectory in simulation [60].

D. Possible Uses of Dataset

The AERPAW Lake Wheeler UAV dataset is a useful tool for a range of wireless networking and DT research tasks:

1) *Simulation Environment Calibration:* Use emulation data to calibrate, validate, and tune simulation environments, enabling them to be more effective predictive tools for real-world deployments.

2) *A2G Signal Propagation Analysis:* Investigate altitude, trajectory, and environmental effects on RSRP and throughput for UAV-ground links.

3) *DT Performance Evaluation:* Benchmark and validate the fidelity of network emulators of DTs by direct comparison with field data.

4) *Trajectory-Aware Algorithm Benchmarking:* Develop and test path planning, handover, and resource allocation algorithms using realistic UAV trajectory and signal data.

5) *Empirical Propagation Model Fitting:* Construct and calibrate A2G path loss models and spatial radio maps from measured RSRP and SNR data.

6) *Machine Learning for Predicting Link Quality:* Train the classifiers or regressors to predict RSRP or throughput using spatial and mobility attributes.

7) *Coverage and Connectivity Mapping:* Generate 2D/3D coverage maps for representing and assessing the radio network connectivity along complex UAV flight trajectories.

XIV. RAY TRACING SIMULATION AND MEASUREMENT COMPARISON DATASET

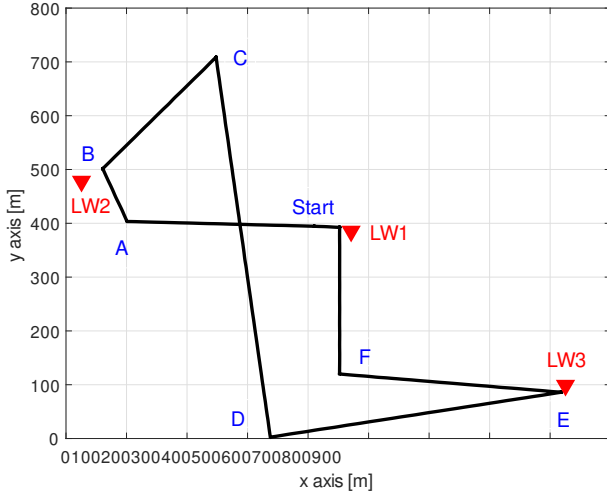
In this section, we present a dataset that enables direct comparison between ray tracing (RT) simulation and real-world measurement of RSS measured at the AERPAW testbed.

The dataset includes UAV trajectory and altitude information, and RSS data collected at each tower. To investigate realistic propagation characteristics in the RT simulation, we implement forest areas with a simple tree model and also incorporate geographical information, including buildings.

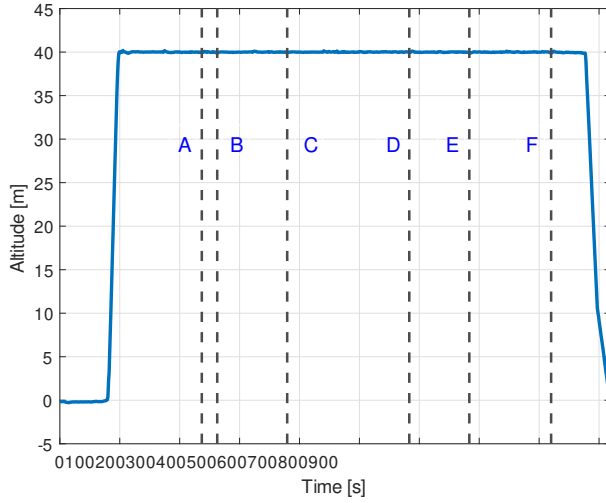
A. Description of Hardware and Software

The predefined trajectory and altitude over time of the UAV are demonstrated in Fig. 36. The UAV takes off near the LW1 tower. After takeoff, the UAV sweeps LW2 and LW3 towers. The UAV returns to LW1 and lands on the ground. Each waypoint of the trajectory is highlighted with a letter, which corresponds to vertical lines in Fig. 36b. The UAV and LW towers have SISO antenna setups with 3.3 GHz carrier frequency. The RSS is recorded by a dual-channel USRP B210 and GNU Radio at each tower while the UAV transmits signals. The RSS is measured for 20 ms for every 100 ms intervals.

For the RT, NVIDIA Sionna [79] is used with support of the Open Street Map (OSM) database [80] for geographic and building information and Blender [81] for 3D modeling. To consider the realistic effects of the trees in the LWRFL areas, we implement a simple tree model and populate it in the Lake Wheeler area, as shown in Fig. 2 of [62]. We adopt predefined material settings for RT simulation from Sionna. Specifically, the surface materials of the buildings are set to concrete and medium dry ground, which are defined as “itu_concrete” and “itu_medium_dry_ground”, respectively. On the other hand, a tree model consists of a wooden cylinder with “itu_wood” and a cone on top of the cylinder with custom foliage material constants, which is calculated under [82].



(a) Trajectory



(b) Altitude

Fig. 36: Trajectory and altitude of the signal coverage measurements and RT simulation (reused from [62]).

The RT simulation is conducted at each GPS coordinate of the UAV along the predefined trajectory to allow for direct comparison with the measurement data. For calibration purposes, offsets that have minimum root mean squared error (RMSE) are searched within the range of $[-50:50]$ dB by the unit of 0.1 dB. Moreover, the altitude over the predefined trajectory below 0.5 m is rounded up to 0.5 m for the RT simulation.

B. Dataset Format

The dataset and post-processing scripts are publicly available at [61]. The dataset consists of three MATLAB scripts and three data folders or .mat files, labeled with the prefixes “C” and “D”, respectively. The structure of the dataset and post-processing scripts can be summarized as follows.

- `D1_RSS_Measurements`: This folder has 5 subfolders RSS measurements in dB scale at 5 different LW tow-

ers (LW1-LW5). The GPS information of the predefined trajectory of the UAV is also included in each subfolder.

- `D2_RT_RSS_results_data.mat`: This file includes the RT simulation results over the predefined trajectory of the UAV. The path coefficient and propagation delays are stored for each UAV position.
- `D3_RT_Measurement_RSS_postprocessed_data.mat`: This file contains the time duration of the measurement campaign, RSS results from RT simulation, and calibrated RSS results from the measurements.
- `C1_trajectory_altitude.m`: This script plots the predefined trajectory and altitude over time as shown in Fig. 36.
- `C2_RSS_postprocessing.m`: This script is for post-processing and calibration of RT and measured data. This script is optional for reproducing the results because the output of this script, which is `D3_RT_Measurement_RSS_postprocessed_data.mat`, is already included in the dataset.
- `C3_RSS_comparison.m`: The post-processed RSS results from RT and the measurement campaign can be plotted by this script, which is demonstrated in Fig. 37. The output of this script shows the 5 figures of results for each tower.

C. Representative Results

The RT simulation results of RSS and measurement from each tower are shown in Fig. 37, where Z indicates out-of-coverage area. Here, the measurements from different antennas of the dual-channel USRP at each tower are labeled as USRP1 and USRP2, respectively. The RT simulation results are consistent with the measurements at all tower cases. Fluctuations are observed in the measurement over the trajectory due to changes in the direction (roll/yaw/pitch) of the UAV at each waypoint, fading effects, and other factors. It is also worthwhile to note that a 10 dB variation between USRP1 and USRP2 is observed in the LW1 and LW2 cases, which are highlighted in the figure with arrows. This variation can be attributed to LoS blockage and channel conditions due to the antenna orientations of USRP1 and USRP2, facing the UAV. Moreover, altitude-dependent out-of-coverage areas can be found in the RT simulation for LW3 at lower altitudes during takeoff and landing (time interval before 100 s and after 850 s).

D. Possible Uses of Dataset

Given the RT-based RSS results and dual-channel USRP measured dataset at the towers, this dataset can be used for the following purposes or analysis scenarios.

- 1) *Benchmarking RT Algorithms*: The dataset provides RSS measurements from the towers and simulated results from the NVIDIA Sionna RT. Thus, the dataset can be used for validation and benchmarking of different RT approaches by comparing the simulation results.

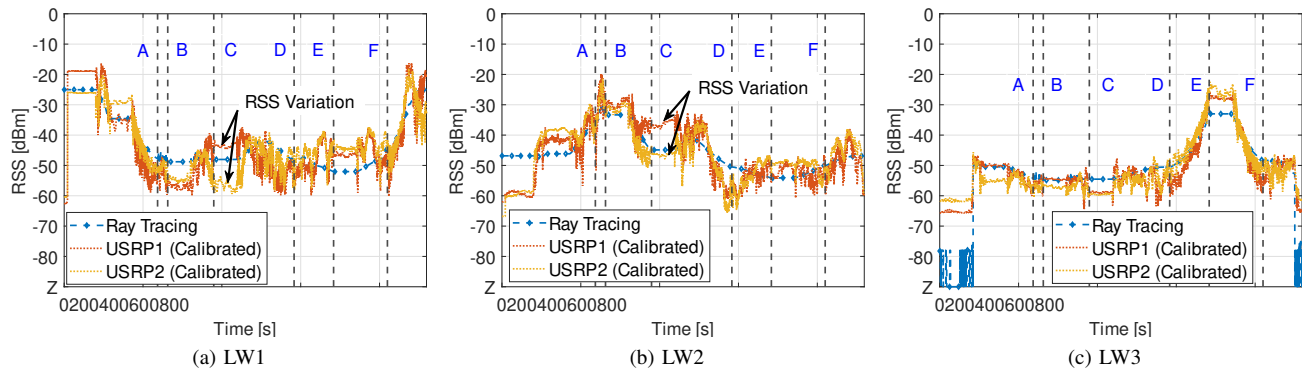


Fig. 37: RSS measurement and RT simulation results with the predefined trajectory for the specific UAV locations A-F in Fig. 36 (reused from [62]).

2) *RF Coverage Analysis and Trajectory Planning*: Based on a rural area with dense foliage, the dataset enables the prediction and design of reliable RF coverage by comparing with the desired performance requirements, such as required SNR levels, etc.

3) *UAV Communication Link Analysis*: The dataset captures fluctuation in RSS over the predefined trajectory, allowing for detailed analysis of link performance, e.g., throughput, under realistic propagation conditions.

4) *Statistical Propagation Channel Modeling*: Since the RT results include propagation delay and path coefficient information, the dataset can be used for propagation channel characterization by statistically analyzing this information.

XV. CONCLUDING REMARKS

Aerial wireless connectivity is becoming an essential enabler of next-generation communication systems, including 5G-Advanced and 6G. This paper presented a diverse collection of open and well-documented datasets from the NSF AERPAW testbed, covering various radio technologies such as 5G, Wi-Fi, and LoRa, and captured using UAVs, helikites, programmable SDR nodes, and commercial UE. We discussed the technical and regulatory challenges associated with developing programmable aerial wireless platforms to collect such datasets, including the integration of SDRs, real-time localization with centimeter-level precision, testbed-wide time synchronization, and compliance with FAA and FCC requirements. The resulting datasets offer high spatial and temporal resolution, supporting a wide range of research activities in wireless communications, signal processing, and machine learning. These datasets are curated according to FAIR principles and are intended to support the academic and industrial research community. They enable rigorous evaluation of propagation models, data-driven algorithm design, and performance benchmarking in altitude-varying environments. Future work will involve expanding the dataset scope to include cooperative and mobile scenarios, incorporating advanced networking features, and aligning with emerging standards to inform data-driven regulatory policy.

REFERENCES

- [1] G. Geraci, A. Garcia-Rodriguez, M. M. Azari, A. Lozano, M. Mezzavilla, S. Chatzinotas, Y. Chen, S. Rangan, and M. Di Renzo, "What will the future of UAV cellular communications be? A flight from 5G to 6G," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 3, pp. 1304–1335, 2022.
- [2] A. Guidotti, A. Vanelli-Coralli, V. Schena, N. Chuberre, M. El Jaafari, J. Puttonen, and S. Cioni, "The path to 5G-advanced and 6G non-terrestrial network systems," in *IEEE Advanced Satellite Multimedia Systems Conference and the Signal Processing for Space Communications Workshop*, 2022, pp. 1–8.
- [3] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2334–2360, 2019.
- [4] Federal Communications Commission, "The role of 5G for UAS," FCC Technological Advisory Council, Tech. Rep., Sep. 2021, accessed: Jul. 30, 2025. [Online]. Available: <https://www.fcc.gov/sites/default/files/tac-5g-uas-white-paper-2021.pdf>
- [5] —, "Use of the 5030–5091 MHz band for unmanned aircraft communications," Federal Communications Commission, Tech. Rep. FCC-22-101, Dec. 2022, accessed: Jul. 30, 2025. [Online]. Available: <https://docs.fcc.gov/public/attachments/FCC-22-101A1.pdf>
- [6] 3GPP, "Enhanced LTE support for aerial vehicles," 3rd Generation Partnership Project (3GPP), TR 36.777, Dec. 2017, available: https://www.3gpp.org/ftp/Specs/archive/36_series/36.777/.
- [7] International Telecommunication Union, "The use of International Mobile Telecommunications (IMT) for broadband public protection and disaster relief (PPDR) applications," International Telecommunication Union, Radiocommunication Sector (ITU-R), Recommendation Report ITU-R M.2291-2, 2021, accessed: 2025-07-31. [Online]. Available: https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2291-2-2021-PDF-E.pdf
- [8] Ericsson, Samsung, and Qualcomm, "Clarification of RSRP measurement triggering for number of cells for UAVs," 3GPP TSG RAN Working Group 4 (RAN4), LS / Technical Contribution R2-2203658, 2022.
- [9] AERPAW Project, "Aerpaw wireless datasets for public use," [Online]. Available: <https://aerpaw.org/experiments/datasets/>, accessed: Aug. 5, 2025.
- [10] R. Mozy, P. Masek, M. Stusek, K. Molnar, M. Palenska, D. Moltchanov, and J. Hosek, "Experimental quality assessment of cellular networks and their utilization for UAV services," in *IEEE Vehicular Technology Conference*, 2023, pp. 1–6.
- [11] J. Braunfelds, G. Jakovels, I. Murans, A. Litvinenko, U. Senkans, R. Rumba, A. Onzuls, G. Valters, E. Lidere, and E. Plone, "Experimental study on LTE mobile network performance parameters for controlled drone flights," *Sensors*, vol. 24, no. 20, p. 6615, 2024.
- [12] J. Zhang, K. Lu, Y. Wan, J. Xie, and S. Fu, "Empowering UAV-based airborne computing platform with SDR: Building an LTE base station for enhanced aerial connectivity," *IEEE Trans. Veh. Technol.*, vol. 73, no. 10, pp. 15 513–15 524, 2024.
- [13] N. Ruseno, H. V. Ongkowiyo, and C.-Y. Lin, "Analysis of 4G signal quality in the UAS network remote ID using machine learning methods,"

- Journal of Aeronautics, Astronautics and Aviation*, vol. 56, no. 2, pp. 555–568, 2024.
- [14] M. A. Zulkifley, M. Behjati, R. Nordin, and M. S. Zakaria, “Mobile network performance and technical feasibility of LTE-powered unmanned aerial vehicle,” *Sensors*, vol. 21, no. 8, p. 2848, 2021.
 - [15] I. Kovacs, R. Amorim, H. C. Nguyen, J. Wigard, and P. Mogensen, “Interference analysis for UAV connectivity over LTE using aerial radio measurements,” in *IEEE Vehicular Technology Conference*, 2017, pp. 1–6.
 - [16] S. Chennamsetty, S. Boddu, and P. Chandhar, “Real time measurement and analysis of 4G and 5G spectrum occupancy in Andhra Pradesh and Telangana - India,” in *IEEE India Council International Conference*, 2023, pp. 31–36.
 - [17] D. Kuester, X. Lu, D. Gu, A. Kord, J. Rezac, K. Carson, M. L. Dowell, E. Eysen, A. Feldman, K. Forsyth *et al.*, “Radio spectrum occupancy measurements amid COVID-19 telework and telehealth,” *National Institute of Standards and Technology, Tech. Rep. TN-2240*, 2022.
 - [18] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne *et al.*, “The FAIR guiding principles for scientific data management and stewardship,” *Scientific data*, vol. 3, no. 1, pp. 1–9, 2016.
 - [19] AERPAW User Manual, “FCC and FAA Regulations in AERPAW,” [Online]. Available: <https://sites.google.com/ncsu.edu/aerpaw-user-manual/3-user-resources-and-policies/3-5-fcc-faa-regulations>, accessed: Oct. 1, 2025.
 - [20] “ArduPilot website,” <https://ardupilot.org>.
 - [21] “MAVLink specifications,” <https://mavlink.io>.
 - [22] “QGroundControl,” <https://qgroundcontrol.com/>.
 - [23] AERPAW User Manual, “AERPAW Sample Experiments,” [Online]. Available: <https://sites.google.com/ncsu.edu/aerpaw-user-manual/6-sample-experiments-repository>, accessed: Oct. 1, 2025.
 - [24] S. J. Maeng, O. Ozdemir, I. Guvenc, M. L. Sichertiu, and R. Dutta, “LTE I/Q measurement by AERPAW platform for air-to-ground propagation modeling,” *IEEE Dataport, Dataset*, 2022. [Online]. Available: <https://doi.org/10.21227/0p43-0d72>
 - [25] S. J. Maeng, O. Ozdemir, I. Guvenc, M. L. Sichertiu, M. Mushi, and R. Dutta, “LTE I/Q measurement by AERPAW platform for air-to-ground propagation modeling,” <https://codeocean.com/capsule/8597224>, 2023.
 - [26] S. J. Maeng, O. Ozdemir, İ. Güvenç, M. Sichertiu, R. Dutta, and M. Mushi, “AERIQ: SDR-based LTE I/Q measurement and analysis framework for air-to-ground propagation modeling,” in *IEEE Aerospace Conference*, 2023, pp. 1–11.
 - [27] S. J. Maeng, O. Ozdemir, İ. Güvenç, and M. L. Sichertiu, “Kriging-based 3-D spectrum awareness for radio dynamic zones using aerial spectrum sensors,” *IEEE Sensors Journal*, vol. 24, no. 6, pp. 9044–9058, 2024.
 - [28] S. J. Maeng, H. Kwon, O. Ozdemir, and İ. Güvenç, “Impact of 3-D antenna radiation pattern in UAV air-to-ground path loss modeling and RSRP-based localization in rural area,” *IEEE Open Journal of Antennas and Propagation*, vol. 4, pp. 1029–1043, Oct. 2023.
 - [29] A. H. F. Raouf, O. Ozdemir, I. Guvenc, and M. Sichertiu, “AERPAW helikite spectrum measurements at Packapalooza Festival in Aug. 2024,” Dryad, Dataset, 2024. [Online]. Available: <https://doi.org/10.5061/dryad.0zpc8676n>
 - [30] —, “AERPAW helikite spectrum measurements at Lake Wheeler site in August 2024,” Dryad, Dataset, 2024. [Online]. Available: <https://doi.org/10.5061/dryad.pk0p2ngz8>
 - [31] S. J. Maeng, O. Ozdemir, I. Guvenc, and M. Sichertiu, “AERPAW helikite spectrum measurements at Lake Wheeler site in May 2022,” Dryad, Dataset, 2023. [Online]. Available: <https://doi.org/10.5061/dryad.jm63xsjhw>
 - [32] —, “AERPAW helikite spectrum monitoring at Packapalooza festival in Aug. 2022,” Dryad, Dataset, 2023. [Online]. Available: <https://doi.org/10.5061/dryad.qbzkh18qd>
 - [33] —, “AERPAW helikite spectrum measurements at Packapalooza Festival in Aug. 2023,” Dryad, Dataset, 2023. [Online]. Available: <https://doi.org/10.5061/dryad.mw6m9063f>
 - [34] A. H. F. Raouf, S. J. Maeng, I. Guvenc, Ö. Özdemir, and M. Sichertiu, “Cellular spectrum occupancy probability in urban and rural scenarios at various UAS altitudes,” in *IEEE Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2023, pp. 1–6.
 - [35] S. J. Maeng and İ. Güvenç, “Altitude-dependent cellular spectrum occupancy: From measurements to stochastic geometry models,” *IEEE Internet of Things Journal*, 2025.
 - [36] A. H. F. Raouf, S. J. Maeng, I. Guvenc, Ö. Özdemir, and M. Sichertiu, “Spectrum monitoring and analysis in urban and rural environments at different altitudes,” in *IEEE Vehicular Technology Conference (VTC)*, 2023, pp. 1–7.
 - [37] S. Singh, O. Ozdemir, M. Sichertiu, and I. Guvenc, “AERPAW: Aerial Android-based 4G measurements of public AT&T cell towers from a tethered Helikite in Aug 2023,” Dryad, Dataset, 2023. [Online]. Available: <https://doi.org/10.5061/dryad.02v6vwwqgt>
 - [38] S. Singh, R. Asokan, O. Ozdemir, M. Sichertiu, and I. Guvenc, “AERPAW: Aerial Android-based cellular measurements for horizontal sweeps across a private base station,” Dryad, Dataset, 2024. [Online]. Available: <https://doi.org/10.5061/dryad.6wwpzgnbd>
 - [39] —, “AERPAW: Aerial Android-based 4G measurements for semi circle trajectories around a private base station,” Dryad, Dataset, 2024. [Online]. Available: <https://doi.org/10.5061/dryad.gmsbcc30x>
 - [40] —, “AERPAW: Aerial android-based cellular measurements for two sweeps across a private base station,” Dryad, Dataset, 2024. [Online]. Available: <https://doi.org/10.5061/dryad.02v6vwwqgt>
 - [41] R. Asokan, O. Ozdemir, İ. Güvenç, and M. L. Sichertiu, “Aerial RF and throughput measurements on a non-standalone 5G wireless network,” in *2024 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2024, pp. 120–123.
 - [42] S. Vargas Villar, A. H. F. Raouf, O. Ozdemir, I. Guvenc, and M. Sichertiu, “LoRaWAN gateway performance and vehicle tracking data in AERPAW testbed,” Dryad, Dataset, 2025. [Online]. Available: <https://doi.org/10.5061/dryad.w0vt4b939>
 - [43] A. Gürses, “USRP Channel Sounder,” <https://github.com/anilgurses/USRP-Channel-Sounder>, 2025.
 - [44] A. Gürses and M. L. Sichertiu, “Air-to-ground channel modeling for UAVs in rural areas,” in *IEEE Vehicular Technology Conference*, 2024, pp. 1–6.
 - [45] C. Dickerson, O. Ozdemir, M. Sichertiu, and I. Guvenc, “AERPAW Keysight N6841A RF Sensor UAV Localization Measurements with LOS/NLOS Information, July 2024,” Dryad [Dataset], 2025, accessed: March 2025. [Online]. Available: <https://doi.org/10.5061/dryad.vq83bk44h>
 - [46] U. Bhattacharjee, I. Guvenc, O. Ozdemir, and M. L. Sichertiu, “Experimental study of outdoor UAV localization and tracking using passive RF sensing,” *IEEE Dataport, Dataset*, 2022. [Online]. Available: <https://doi.org/10.21227/5rrb-9y02>
 - [47] U. Bhattacharjee, E. Ozturk, O. Ozdemir, I. Guvenc, M. L. Sichertiu, and H. Dai, “Experimental study of outdoor UAV localization and tracking using passive rf sensing,” in *Proceedings of the 15th ACM Workshop on Wireless Network Testbeds, Experimental evaluation & CHaracterization*, 2022, pp. 31–38.
 - [48] C. Dickerson, S. Masrur, J. Dickerson, O. Ozdemir, and I. Guvenc, “Impact of altitude, bandwidth, and NLOS bias on TDOA-based 3D UAV localization: Experimental results and CRLB analysis,” in *Proc. IEEE ICC Workshops, Wkshp. on Positioning and Sensing over Wireless Netw.*, Montreal, Canada, June 2025, to appear.
 - [49] S. Masrur, O. Ozdemir, A. Gurses, I. Guvenc, M. L. Sichertiu, R. Dutta, M. Mushi, C. Dickerson, G. Reddy, S. V. Villar *et al.*, “Collection: Datasets from AFAR Challenge,” *arXiv preprint arXiv:2505.06823*, 2025.
 - [50] S. Masrur and I. Guvenc, “Bridging simulation and reality: A 3D clustering-based deep learning model for UAV-based RF source localization,” in *Proc. IEEE ICC Workshops, Wkshp.*, Montreal, Canada, June 2025, to appear.
 - [51] P. S. Kudyba, Q. Lu, and H. Sun, “Bayesian optimization for fast radio mapping and localization with an autonomous aerial drone,” in *IEEE Vehicular Technology Conference*, Washington, DC, USA, Sep. 2024, pp. 1–5.
 - [52] P. Kudyba, J. S. Mandapaka, W. Wang, L. McCorkendale, Z. McCorkendale, M. Kidane, H. Sun, E. Adams, K. Namuduri, F. Fund *et al.*, “A UAV-assisted wireless localization challenge on AERPAW,” *arXiv preprint arXiv:2407.12180*, 2024.
 - [53] B. Chatterjee, S. Chaudhari, L. Zhizhen, Y. Liu, and R. Dutta, “Wireless signal source localization by unmanned aerial vehicle using AERPAW digital twin and testbed,” in *Proc. IFIP Networking Conf.*, Thessaloniki, Greece, Jun. 2024, pp. 666–671.
 - [54] M. Ezuma, F. Erden, C. K. Anjinappa, O. Ozdemir, and I. Guvenc, “Drone remote controller RF signal dataset,” *IEEE Dataport doi:10.21227/ss99-8d56*, Nov. 2020.
 - [55] —, “Detection and classification of UAVs using RF fingerprints in the presence of Wi-Fi and Bluetooth interference,” *IEEE open j. commun. soc.*, pp. 60–76, Nov. 2019.

- [56] —, “Micro-UAV detection and classification from RF fingerprints using machine learning techniques,” in *IEEE Aerospace conference*, Big Sky, Montana, Mar. 2019.
- [57] E. Ozturk, F. Erden, and I. Guvenc, “RF-based low-SNR classification of UAVs using convolutional neural networks,” *arXiv preprint arXiv:2009.05519*, pp. 1–18, Sep. 2020, preprint.
- [58] “UAVFlexSimFramework,” 2025. [Online]. Available: <https://github.com/mhossenece/UAVFlexSimFramework.git>
- [59] M. S. Hossen, A. Gurses, M. Sichitiu, and I. Guvenc, “Accelerating development in UAV network digital twins with a flexible simulation framework,” <https://doi.org/10.5061/dryad.zw3r228k9>, 2025, dryad.
- [60] M. S. Hossen, A. Gurses, M. Schicitu, and I. Guvenc, “Accelerating development in UAV network digital twins with a flexible simulation framework,” in *Proc. IEEE ICC Workshops: Workshop on Digital Twins over NextG Wireless Networks*, Montreal, Canada, 2025.
- [61] “Sionna ray tracing vs. real-world measurements: Signal strength comparison in rural scenarios with foliage,” [Online]. Available: <https://datadryad.org/dataset/doi:10.5061/dryad.ttdz08m8r>, accessed: 2025.
- [62] D. Lee, O. Ozdemir, R. Asokan, and I. Guvenc, “Analysis and prediction of coverage and channel rank for uav networks in rural scenarios with foliage,” *IEEE Open Journal of Vehicular Technology*, vol. 6, pp. 1943–1962, 2025.
- [63] M. Funderburk, J. Kesler, K. Sridhar, M. L. Sichitiu, I. Guvenc, R. Dutta, T. Zajkowski, and V. Marojevic, “AERPAW vehicles: Hardware and software choices,” in *Proceedings of the Eighth Workshop on Micro Aerial Vehicle Networks, Systems, and Applications*, 2022, pp. 37–42.
- [64] The GNU Radio Foundation, Inc., “The signal metadata format (SigMF),” <https://sigmf.org>, 2018, cC-BY-SA-4.0. [Online]. Available: <https://sigmf.org>
- [65] S. J. Maeng, O. Ozdemir, I. Guvenc, M. L. Sichitiu, M. Mushi, and R. Dutta, “LTE I/Q data set for UAV propagation modeling, communication, and navigation research,” *IEEE Communications Magazine*, vol. 61, no. 9, pp. 90–96, 2023.
- [66] S. J. Maeng, O. Ozdemir, I. Guvenc, and M. Sichitiu, “3D antenna pattern measurement for AERPAW experiments,” IEEE Dataport, Dataset, 2023. [Online]. Available: <https://doi.org/10.21227/mma2-0t93>
- [67] M. Rahman, S. J. Maeng, İ. Güvenç, and C.-W. Wong, “3D spectrum awareness for radio dynamic zones using Kriging and matrix completion,” in *IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, Washington, DC, USA, May 2024, pp. 439–446.
- [68] Ettus Research, “USRP B205mini-i,” [Online]. Available: <https://www.ettus.com/all-products/USRP-B205Mini-I/>, accessed: Feb. 19, 2025.
- [69] AMD, “Spartan-6 FPGA Product Page,” [Online]. Available: <https://www.amd.com/en/products/adaptive-socs-and-fpgas/fpga/spartan-6.html>, accessed: Feb. 19, 2025.
- [70] Analog Devices, “AD9361 RF Agile Transceiver,” [Online]. Available: <https://www.analog.com/en/products/ad9361.html>, accessed: Feb. 19, 2025.
- [71] AERPAW User Manual, “UHD1: Spectrum Monitoring,” [Online]. Available: <https://sites.google.com/ncsu.edu/aerpaw-user-manual/6-sample-experiments-repository/6-1-radio-software/6-1-4-uhd-python-api-experiments/uhd1-spectrum-monitoring?authuser=0>, accessed: Feb. 19, 2025.
- [72] Keysight, “Nemo Handy Handheld Measurement Solution,” [Online]. Available: <https://www.keysight.com/us/en/product/NTH50047B/nemo-handy-handheld-measurement-solution.html>, accessed: 2025.
- [73] Federal Aviation Administration, “Unmanned aircraft system (UAS) traffic management (UTM): Concept of operations v2.0,” FAA NextGen Office, Washington, DC, USA, Tech. Rep., Mar. 2020.
- [74] I. Guvenc, F. Koohifar, S. Singh, M. L. Sichitiu, and D. Matolak, “Detection, tracking, and interdiction for amateur drones,” *IEEE Commun. Mag.*, vol. 56, no. 4, pp. 75–81, 2018.
- [75] M. M. Azari, H. Sallouha, A. Chiumento, S. Rajendran, E. Vinogradov, and S. Pollin, “Key technologies and system trade-offs for detection and localization of amateur drones,” *IEEE Commun. Mag.*, vol. 56, no. 1, pp. 51–57, 2018.
- [76] AERPAW Team, “AERPAW AFAR Challenge,” [Online]. Available: <https://aerpaw.org/aerpaw-afar-challenge/>, accessed: Aug. 19, 2025.
- [77] W. Khawaja, M. Ezuma, V. Semkin, F. Erden, O. Ozdemir, and I. Guvenc, “A survey on detection, classification, and tracking of UAVs using radar and communications systems,” *IEEE Commun. Surveys Tuts.*, pp. 1–40, Mar. 2025.
- [78] M. Ezuma, O. Ozdemir, C. K. Anjinappa, W. A. G. Khawaja, and I. Guvenc, “Micro-UAV detection with a low-grazing angle millimeter wave radar,” in *IEEE Radio and Wireless Symposium*, Orlando, FL, Jan. 2019.
- [79] NVIDIA, “NVIDIA Sionna,” [Online]. Available: <https://developer.nvidia.com/sionna>, accessed: 2022.
- [80] OpenStreetMap, “OpenStreetMap,” [Online]. Available: <https://www.openstreetmap.org>, accessed: 2004.
- [81] Blender Foundation, “Blender,” [Online]. Available: <https://www.blender.org/>, accessed: 1994.
- [82] International Telecommunication Union, “ITU-R P.833-10,” [Online]. Available: <https://www.itu.int/pub/R-REC/en>, accessed: 2021.