

# Collection: *UAV-Based Wireless Multi-modal Measurements from AERPAW Autonomous Data Mule (AADM) Challenge in Digital Twin and Real-World Environments*

Md Sharif Hossen<sup>1</sup>, Cole Dickerson<sup>1</sup>, Ozgur Ozdemir<sup>1</sup>, Anil Gurses<sup>1</sup>, Mohamed Rabeek Sarbudeen<sup>1</sup>, Thomas Zajkowski<sup>1</sup>, Ahmed Manavi Alam<sup>1</sup>, Everett Tucker<sup>1</sup>, William Bjorndahl<sup>2</sup>, Fred Solis<sup>2</sup>, Sadaf Javed<sup>3</sup>, Anirudh Kamath<sup>4</sup>, Xiangyao Tang<sup>4</sup>, Joarder Jafor Sadique<sup>5</sup>, Kevin Liu Hermstein<sup>6</sup>, Kaies Al Mahmud<sup>7</sup>, Jose Angel Sanchez Vilorio<sup>8</sup>, Skyler Hawkins<sup>9</sup>, Yuqing Cui<sup>10</sup>, Annoy Dey<sup>10</sup>, Yuchen Liu<sup>1</sup>, Ali Gurbuz<sup>1</sup>, Joseph Camp<sup>2</sup>, Rizwan Ahmad<sup>3</sup>, Jacobus van der Merwe<sup>4</sup>, Ahmed Ibrahim Mohamed<sup>5</sup>, Gil Zussman<sup>6</sup>, Mehmet Kurum<sup>7</sup>, Namuduri Kamesh<sup>9</sup>, Zhangyu Guan<sup>10</sup>, Dimitris Pados<sup>8</sup>, George Sklivanitis<sup>8</sup>, Ismail Guvenc<sup>1</sup>, Mihail Sichitiu<sup>1</sup>, Magreth Mushi<sup>1</sup>, Rudra Dutta<sup>1</sup>

<sup>1</sup>North Carolina State University, Raleigh, NC, USA; <sup>2</sup>Southern Methodist University, TX, USA; <sup>3</sup>National University of Sciences and Technology, Pakistan; <sup>4</sup>University of Utah, UT, USA; <sup>5</sup>Florida International University, FL, USA; <sup>6</sup>Columbia University, NY, USA; <sup>7</sup>The University of Georgia, GA, USA; <sup>8</sup>Florida Atlantic University, FL, USA; <sup>9</sup>The University of North Texas, TX, USA; <sup>10</sup>University of Minnesota, MN, USA

CORRESPONDING AUTHOR: Md Sharif Hossen (e-mail: mhossen@ncsu.edu).

This work is supported in part by the NSF under Grant CNS-1939334, CNS-2450593, 2332661, OAC-2512931, 2346555, CNS-2144297, CNS-2148128, CNS-2332662, CNS-2117822, EEC-2133516, CNS-2440756, SWIFT-2229563, DGE-2137100, and by HEC, PK.

**ABSTRACT** In this work, we present an unmanned aerial vehicle (UAV) wireless dataset collected as part of the AERPAW Autonomous Aerial Data Mule (AADM) challenge, organized by the NSF Aerial Experimentation and Research Platform for Advanced Wireless (AERPAW) project. The AADM challenge was the second competition in which an autonomous UAV acted as a data mule, where the UAV downloaded data from multiple base stations (BSs) in a dynamic wireless environment. Participating teams designed flight control and decision-making algorithms for choosing which BSs to communicate with and how to plan flight trajectories to maximize data download within a mission completion time. The competition was conducted in two stages: Stage 1 involved development and experimentation using a digital twin (DT) environment, and in Stage 2, the final test run was conducted on the outdoor testbed. The total score for each team was compiled from both stages. The resulting dataset includes link quality and data download measurements, both in DT and physical environments. Along with the USRP measurements used in the contest, the dataset also includes UAV telemetry, Keysight RF sensors position estimates, link quality measurements from LoRa receivers, and Fortem radar measurements. It supports reproducible research on autonomous UAV networking, multi-cell association and scheduling, air-to-ground propagation modeling, DT-to-real-world transfer learning, and integrated sensing and communication, which serves as a benchmark for future autonomous wireless experimentation.

**IEEE SOCIETY/COUNCIL** IEEE Communications Society (ComSoc), IEEE Vehicular Technology Society (VTS), IEEE Aerospace and Electronic Systems Society (AESS)

**DATA DOI/PID** <https://doi.org/10.5061/dryad.7d7wm3898>

**DATA TYPE/LOCATION** Time-series; AERPAW Lake Wheeler Field Labs, Raleigh, North Carolina, USA; AERPAW Digital Twin (DT) for Lake Wheeler Field Labs

**INDEX TERMS** AERPAW, air-to-ground propagation modeling, digital twin, real-world wireless datasets, UAV-assisted data mule, UAV communications, USRP, Keysight RF sensor, LoRa, Fortem radar

## BACKGROUND

Ground-based sensing systems, such as environmental monitors, motion detectors, weather stations, and long-range (LoRa) receivers or radio frequency sensors (RFS), are often deployed in remote areas where network coverage is limited or unavailable. In these settings, data are commonly retrieved through in-person site visits, which are labor-intensive and slow. This motivates the aerial experimentation and research platform for the advanced wireless (AERPAW) [1] autonomous data mule (AADM) challenge, where unmanned aerial vehicles (UAVs) can autonomously collect data from distributed ground nodes without relying on fixed communication infrastructure. Although several sensing approaches are considered, USRP-based wireless nodes are selected as the main mechanism for data collection. The goal of the AADM challenge is to evaluate how autonomous UAVs can make real-time decisions about mobility and base-station selection to maximize data collection under realistic wireless constraints.

UAVs, also known as drones, offer a promising solution to data collection challenges by serving as autonomous data mules—airborne nodes that can visit remote sensor sites and collect data. Their ability to span large geographic distances, operate autonomously, and establish temporary communication links makes them well-suited to extend connectivity to remote locations. Compared to static or terrestrial infrastructure, UAV-based data mules offer on-demand and versatile connectivity, which is particularly important for applications such as environmental monitoring, disaster recovery, and wide-scale Internet of Things (IoT) networks [2].

However, most research work with UAV data mule systems has been conducted in simulation environments that oversimplify wireless link behavior, mobility constraints, and environmental dynamics. Such oversimplifications often fail to capture the actual interference effects, path loss variability, and channel fading caused by the motion of UAVs, altitude changes, or terrain variations. Hence, algorithms and scheduling methods fine-tuned via simulation may not be easily portable to real-world (RW) deployments, limiting their reliability and scalability to RW missions [3].

To bridge the gap, there is a growing need for integrated experimentation facilities that integrate virtual digital twin (DT) frameworks and physical testbeds to experiment with data collection systems enabled by UAVs. Such a platform facilitates researchers in designing, testing, and iteratively optimizing mobility and communication tactics—e.g., trajectory planning, link association, and offloading. AERPAW [1] delivers this two-in-one capability via its hybrid ecosystem that consists of a DT for repeatable, controlled experiments and an outdoor testbed for RW verification.

AERPAW facilitates seamless transition from simulation to deployment for end-to-end assessment of UAV-based wireless operations from the communication, mobility, and control perspectives [4]. Reliable and effective data mule performance requires RW datasets that reflect both wire-

less and mobility dynamics. These datasets should include key performance metrics (KPMs), for example, throughput, signal-to-noise ratio (SNR), and UAV trajectory, speed, and altitude collected across multiple heterogeneous ground sensors. These multi-modal data enable data-driven modeling of how UAV mobility and environmental factors affect link behavior, resulting in predictive and adaptive algorithms for future autonomous UAV data mule networks.

### AADM Challenge

The AADM challenge is the most recent of AERPAW's UAV student competitions, which shifts from RF source localization in the AFAR challenge [5]) to wireless data collection and mobility scheduling. The new challenge was designed to study how UAVs can serve as data mules—airial nodes that download data dynamically from a number of distributed ground stations within a finite time interval.

Student teams from different universities designed UAV algorithms that could autonomously download data from multiple base stations (BSs) spread over the AERPAW Lake Wheeler Field Labs. Each UAV operated as an autonomous data mule (ADM), dynamically determining which BS to associate with, for how long, and when to switch to the next BS, to maximize total data downloaded over a 500-second flight mission. Each flight departed at an altitude of 25 m, with four AERPAW BSs (BS-1 through BS-4) providing ground data for the Lake Wheeler Field Labs in Raleigh, as illustrated in Fig. 1. Serving as an ADM, the UAV flies between BSs to collect sensor data while continuously adapting its trajectory based on link quality and mission time constraints.

The competition consisted of two stages. In Stage 1 (DT development), participants deployed and tested their algorithms on AERPAW's high-fidelity DT platform, which replicated the topology of the real testbed and wireless channel conditions. In Stage 2 (Outdoor testbed evaluation), the same containerized deployments were executed on the real testbed during field trials without any modification of the participant's code. This DT testbed architecture allowed teams to remotely develop, test, and validate ADM solutions under real communication and mobility constraints.

In addition to the USRP-based measurements used for scoring, all RW AADM flights were also monitored by external sensing instruments on the AERPAW testbed, including a network of Keysight RFS, LoRa, and a Fortem R20 radar. Although not used by competition teams, these instruments provided independent UAV tracking used for dataset enrichment.

### Scoring and Ranking

Team performance was evaluated by a composite score that balanced data collection and mission time to completion. The overall score for each mission was calculated as:

$$S = S_1 + S_2 - P, \quad (1)$$

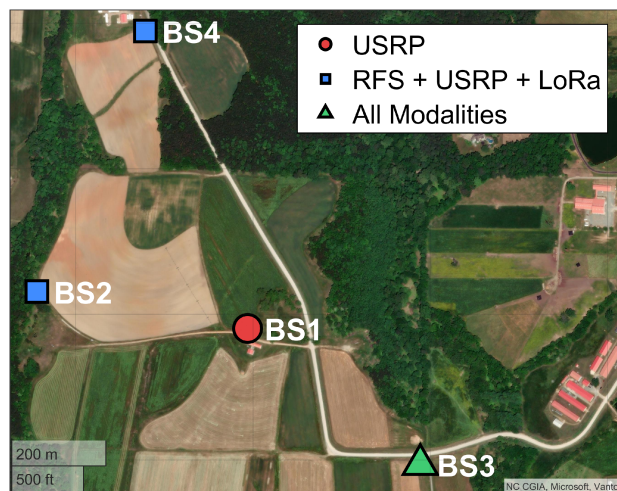


FIGURE 1: Experimental setup of the AADM Challenge, showing four BSs (BS1-BS4) and the sensing modalities colocated at each site across the Lake Wheeler Field Labs in Raleigh, NC. Distinct marker shapes denote the combination of sensing modalities deployed at each BS.

where  $S_1 = (500 - T) \times I(\text{Download\_Complete})$ ,  $S_2 = 100 \times \left(\frac{D_{\text{down}}}{D_{\text{total}}}\right)$ , and  $P$  represents the penalty if the UAV does not land within the specified time.  $D_{\text{down}}$  is the total downloaded data by the UAV from all the BSs, while  $D_{\text{total}}$  is the total data volume set across the BSs. The final ranking was based on a weighted average of DT and testbed outcomes, with 20% of the weight assigned to DT scores and 80% to actual testbed results. This motivated the teams to design algorithms that generalized well from DT to RW environments, highlighting their flexibility in handling dynamic link variations and environmental uncertainty.

### Dataset Overview and Uniqueness

In an outdoor operational ADM scenario, the AADM dataset is a unique, comprehensive set of time-synchronized logs from both a DT and an RW testbed. It includes comprehensive UAV state data on heading, altitude, velocity, and position. It also provides per-second wireless data about BS association status, link quality, and throughput. In contrast to previous open datasets, it provides the first publicly available measurements of a UAV performing dynamic BS association and data-transfer scheduling based on real-time link circumstances. Moreover, AADM is the first to integrate all the multi-cell topologies in contrast to the earlier single-link AFAR [5] dataset. Therefore, it allows for an unprecedented opportunity to benchmark the fidelity of DTs, evaluate the transferability of algorithms across domains, and validate models for cross-layer optimization and AI-enabled resource scheduling in next-generation wireless networks.

Another distinguishing feature of the RW portion of the AADM dataset is the inclusion of three external sensing streams that were collected independently of the UAV and BSs. First, multiple Keysight RFS extracted time-difference-of-arrival (TDOA) measurements from raw IQ data and

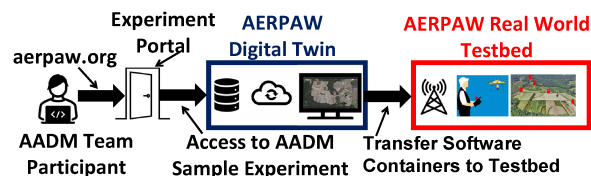


FIGURE 2: AADM competition workflow.

estimated the UAV’s position as it moved across the field. In parallel, a Fortem R20 radar unit continuously tracked the UAV’s position and radial velocity while reporting radar cross-section (RCS) and additional tracking parameters throughout the missions. Finally, LoRa uplink transmissions from the UAV were received by distributed gateways across the test site, with each gateway reporting packet-level reception metadata such as received signal strength indicator (RSSI), SNR, and gateway identifiers and geolocation information. Integrating these sensing modalities with ground-truth UAV trajectory, speed, and altitude information results in a uniquely comprehensive multi-modal dataset for UAV detection and localization research.

A total of 22 teams registered for the AADM challenge, but 15 teams submitted their solutions developed in the DT to participate in the challenge. Each team designed and implemented its own UAV algorithm for ADM operations, focusing on efficient BS association and trajectory planning.

In the first stage (DT evaluation), each team’s proposed solution was run in the DT environment with three separate UAV missions. The overall competition workflow, including the DT and RW testbed stages, is illustrated in Fig. 2. Each mission corresponds to different data-volume configurations across AERPAAW BSs to make the challenge more complex and to optimize the trajectory with heterogeneous traffic demands. Four data-volume configurations, denoted as D1, D2, D3, and D4, are defined across four BSs: BS1, BS2, BS3, and BS4, to introduce heterogeneous traffic demands. The first flight uses the configuration (600, 600, 600, 200) Mbits, the second uses (100, 400, 400, 100) Mbits, and the third flight uses (20, 50, 400, 30) Mbits assigned to BS1 through BS4, respectively. In each case, the UAV aims to maximize the total downloaded data within 500 seconds. In total, 45 flights (15 teams times 3 flights) are run in stage 1.

Based on the performance in the DT, 11 finalist teams were selected for stage 2 (RW testbed evaluation). For the RW testing, the same three-flight configuration, similar to DT, is used in the physical testbed to evaluate each team’s solution, where 33 flights are run (11 teams with 3 flights each). The finalist teams were *InFlux* from the University of Utah (Utah), *Neural Flyers* from North Carolina State University (NC State-1), *IMPRESS@UGA* from the University of Georgia (UGA), *WiMNet* from Columbia University (Columbia), *NICE Lab* from North Carolina State University (NC State-2), *Credente* from Florida International University (FIU), *WINGS Lab* from the University at Buffalo (UB), *CAAI* from Florida Atlantic University (FAU), *FlyFetchers* from

National University of Sciences and Technology (NUST), *SMU ADT Lab* from Southern Methodist University (SMU), and *Eagles* from the University of North Texas (UNT).

### **How the Dataset can be Reused**

With an emphasis on multi-cell data collection, heterogeneous BS traffic, and autonomous association control, the AADM dataset offers a unique platform for investigating how UAV behavior is affected when communication and mobility decisions are tightly interconnected. This dataset can be reused for localization, air-to-ground (A2G) channel propagation modeling, and data-driven RW applications, as discussed in [5]. Additionally, the dataset can be further reused for the following purposes.

#### Multi-Cell Scheduling and Association Policies

In the AADM challenge, each experiment is evaluated in terms of three distinct BS configurations with different data volumes. This provides insights into how UAVs make dynamic scheduling and association decisions. Furthermore, this allows researchers the opportunity to explore new strategies for balancing link quality, distance, and remaining data volume to optimize total throughput and minimize flight time under RW dynamic logs. The dataset has been used for autonomous UAV trajectory design and evaluation for data muling [6] [7] [3].

#### Mission Planning Under Energy and Time Constraints

As in the challenge, each AADM flight lasts approximately 500 seconds and captures detailed UAV motion logs, allowing for studying energy-efficient trajectory planning and time-constrained mission scheduling. Even without direct data on battery usage, researchers can explore the energy consumption from speed, acceleration, and hovering periods to investigate trade-offs between travel distance and data collection. This is important for studying how UAVs can optimize their performance in large-scale and time-sensitive data-mule missions.

#### UAV Localization and Sensor Fusion-Based Tracking

The inclusion of ground-truth UAV position data, position estimates from the Keysight RFS, radar tracking from the Fortem R20 unit, and LoRa gateway reception measurements enables a variety of localization and tracking studies, including ISAC-based schemes [8]. Researchers can benchmark RF-based localization algorithms, evaluate tracking filters, and explore sensor fusion approaches that combine RFS, radar, and LoRa modalities. The dataset further supports analyses of bias and error characterization and the validation of theoretical bounds on localization under realistic flight geometries and channel dynamics, building on our previous work [9].

#### Fairness and Quality-of-Service Across Base Stations

Due to the heterogeneous BS traffic with the UAVs, the dataset allows the analysis of fairness and service distribution. For example, a BS might always get neglected

throughout the mission. This dataset can help explore the design of resource allocation policies similar to distributed edge computing or IoT networks.

#### Scalability and Multi-Agent Scenarios

It can be used to explore the scalability of data-mule operations in next-generation aerial networks. Researchers can simulate scenarios with multiple UAVs collecting data in parallel under different BS data-demand configurations and validate their results with the DT and RW testbed data logs.

## **COLLECTION METHODS AND DESIGN**

The data for the AERPAW 2nd AADM Challenge was collected using AERPAW's hybrid experimentation platform, which supports both a DT and a fully programmable outdoor testbed at the Lake Wheeler Field Labs. Fig. 1 presents the overall system architecture used in both environments, where a UAV-mounted portable node collects the data across four fixed BSs (BS1–BS4). The physical UAV was equipped with an Intel NUC-10 computer (i7-10710U, 64 GB RAM), an onboard real-time kinematic (RTK)-enabled global navigation satellite system (GNSS) receiver, and a USRP B205 software-defined radio with a 3.4 GHz antenna for probe transmission and data-rate estimation. Each BS node was also equipped with a USRP B210 for channel sounding and executed identical receiver software. All radios were configured to operate at a 3.4 GHz center frequency with SNR updates generated each second. These hardware components provided synchronized measurements of UAV position, orientation, link quality, and the derived data-rate estimates used to compute the total amount of downloaded data. The software stack was executed in the DT, where the UAV was replaced by a software-in-the-loop (SITL) emulator and the wireless channel was modeled with a line-of-sight (LOS) plus ground-reflection propagation model, without fading or hardware imperfections. Both environments use identical message formats, experiment containers, and mission rules to ensure comparability.

The challenge provided a template for fixed and autonomous trajectory guidance. Depending on the experimenter-designed algorithm, during each flight, the UAV autonomously took off, ascended to 25 m, and received the mission-specific data volumes D1, D2, D3, and D4 from the controller. After that, data collection was conducted through a coordinated exchange between the UAV and the BS controller. The UAV's GNU radio transmitter [1] periodically emitted probe signals and the four BSs' receivers detected that probe signal and estimated individual SNR. When the UAV sent a request for link quality measurements, the controller would send the SNR value for each BS. Based on the data volume at each BS and the link quality, the UAV then starts downloading the data under the coordination of the controller. The experimenter set up their algorithm, by which the UAV could select a BS to download data in the next time interval. While the controller accumulated these data-rate estimates

at each BS over time, simultaneously the UAV logged SNR, selected BS index, instantaneous rate, cumulative downloaded data, and UAV telemetry (position, velocity, attitude, and battery state). Once the 500 seconds had passed, the controller prevented the UAV from downloading data. This ensures consistency across all experiments by enforcing the download timing constraints.

After every flight, the raw logs from the UAV and the mission controller were merged and processed into the final dataset. All timestamps were converted to a common time base by using system-clock synchronization. Each mission contains time series data containing UAV GNSS position, altitude, velocity, roll/pitch/yaw angles, selected BS, instantaneous SNR for all four BSs, estimated data rate, and cumulative downloaded data. To compare DT and RW environments, log reports from both environments have the same column structure and features. The final dataset was validated with complete and consistent logs, ensuring valid geofence compliance, continuity of telemetry, and a fully executed mission.

In addition to the USRP-based communication infrastructure used for the core AADM task, Fig. 1 also illustrates the deployment of auxiliary sensing modalities used for independent observation of the UAV and its wireless environment. Keysight N6841A RFS co-located at BS2–BS4 passively monitored a channel-sounding waveform transmitted by the UAV at 3.32 GHz with a bandwidth of 5 MHz, serving as a proxy for control and telemetry links throughout AADM flights. A Fortem R20 radar system was mounted at a height of 10 m and co-located with BS3, operating at 16.3 GHz with a maximum bandwidth of 180 MHz to provide independent tracking measurements of the UAV. In addition, fixed LoRaWAN gateways were deployed at BS2–BS4 to enable low-power, long-range communications with a programmable USB-based LoRa end device carried by UAV.

### VALIDATION AND QUALITY

The dataset was validated to ensure that all missions in the AADM dataset contained reliable UAV telemetry, link quality measurements, and data download records. Only those flights were considered that had completed the full mission, a valid GPS fix, and intact logs from both the vehicle and the controller. For the retained missions, the UAV remained within the Lake Wheeler geofence, reached the commanded takeoff altitude of approximately 25 m, completed its traversal among the four BSs, and landed within the 500-second mission window. Following the trajectory shown in Fig. 3(a), the UAV downloads data from multiple BSs as illustrated in Fig. 3(b). We see from Fig. 3 that the controller allows the UAV to download data only during the specified fixed download time (500 seconds). After that, it stops issuing download commands, and the cumulative downloaded data remains constant for the rest of the flight. Similarly, in the controller log, a message indicates that no further actions are required, and the system is currently off.

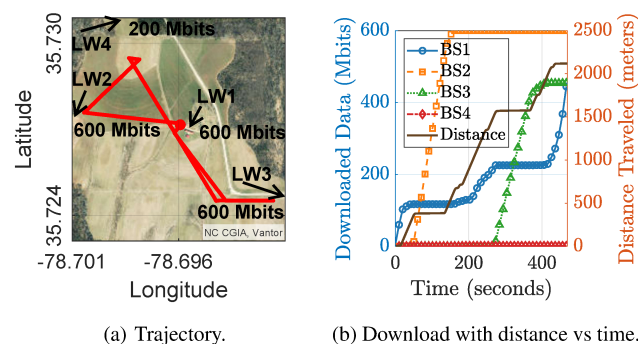


FIGURE 3: The controller enforced a 500-second download time limit for a given trajectory for Scenario 1, team 828.

TABLE 1: Statistics of SNR for four BSs across three AADM scenarios for Team 799 in the RW testbed.

Feature	BS	Stat.	Scenario-1	Scenario-2	Scenario-3
SNR (dB)	BS1	Mean	3.65	0.09	1.52
		Std	10.78	11.15	10.11
		Max	26.57	25.87	27.08
		Min	-38.17	-46.60	-52.11
	BS2	Mean	1.65	1.54	-2.85
		Std	11.52	11.21	5.51
		Max	32.21	32.29	8.92
		Min	-38.82	-19.20	-40.16
	BS3	Mean	-5.05	-8.62	-11.46
		Std	15.26	16.76	18.21
		Max	20.75	20.74	20.47
		Min	-70.52	-67.07	-66.61
BS4	Mean	-36.33	-33.27	-33.39	
	Std	13.82	15.65	15.22	
	Max	-4.52	-2.65	-1.93	
	Min	-70.43	-66.43	-72.10	
Speed (m/s)	-	Mean	3.621	5.205	3.622
		Std	4.405	4.632	4.460
		Max	10.267	10.361	10.320
		Min	0.004	0.011	0.006

Also, the controller outputs the final total downloaded data and the computed mission score. The cumulative download curve and the associated log messages in the controller log ensure that the 500-second download limit is strictly enforced for every mission. Table 1 summarizes the statistics of SNR and UAV speed for a representative team, 799, where the SNR values range from  $-70$  dB to more than 30 dB. These were expected for the geometry of the four BSs and variations in distance and antenna orientation during flight. BS1 and BS2 mostly reach higher average SNR values when the UAV is close to them, while BS3 and BS4 often exhibit lower average SNR due to longer ranges and the RW propagation conditions. The distribution of UAV speed is observed as expected based on mission rules, where the maximum values are close to the configured speed limit of 10 m/s, while the minimum speed is close to zero when the vehicle is hovering or taking a tight turn.

## RECORDS AND STORAGE

Table 2 shows an overview of the dataset across the different collection phases and how the experiments progressed from controlled digital twin (DT) studies to real-world (RW) testbed evaluations. The dataset includes early testbed flights designed to expose teams to RW channel behavior using USRP-based wireless measurements. It also contains DT flights used for algorithm development. Two testbed flights were conducted to study the effect of altitude on signal strength. Finally, the testbed evaluations with 11 finalist teams were conducted with all sensing modalities, including USRP measurements, Keysight RFS, LoRa receivers, and Fortem radar. In addition, 3 subsequent experiments were conducted using only the radar and RFS modalities, in which the UAV altitude was increased, and the transmitted signal bandwidth was expanded from 150 kHz to 5 MHz to improve localization performance [9].

The dataset is organized into three directories, namely USRP, LoRa, RFS, and Radar, under the main directory AADM. In the USRP folder, there are two subfolders: testbed and development. In the testbed folder (Sep), based on each team’s experiment, data volume number, and flight number, there are 33 folders. While inside the development folder (Jun), there are also 45 folders based on experiment number and data volume. Inside each folder, there are 5 sub-folders, namely, LW1, LW2, LW3, LW4, and UAV, representing BS1, BS2, BS3, BS4, and UAV, respectively. Inside the LW1 to LW4 folders, there are four text files, namely, *power\_log.txt*, *quality\_log.txt*, *responder\_log.txt*, and *snr\_log.txt*, which represent the received signal power, the quality of the signal, the SNR sent to the controller each second, and the SNR logs, respectively. In contrast, the LW1 folder contains an additional *controller\_log.txt* file, which includes all the exchanged messages between the UAV, all BSs, and the OEO console. The UAV folder contains *report\_log.txt* and *vehicleOut.txt* as discussed in Table 3 and the fields of the processed CSV files are discussed in Table 4.

However, for the SMU ADT Lab (experiment number 1050), the experimenter modified the vehicle logging script unintentionally, which resulted in stopping the storage of the vehicle log file. Hence, for this experiment, the *vehicleOut.txt* file is unavailable in both environments. In the dataset, there are also 12 (2+2+8) early testbed flights with the specific name of folders as mentioned in Table 2.

The RF Sensor and Radar directory contains 37 subfolders: 33 corresponding to testbed flights collected with all sensing modalities, three additional testbed flights conducted to improve UAV localization and tracking performance, and one folder containing post-processing scripts. Each of the 36 data subfolders includes three files: (i) a copy of the *vehicleOut.txt* file described above, provided again for convenience; (ii) a CSV file named *AADMx.csv*, where *x* denotes the flight number, containing all Keysight RFS localization measurements, as detailed in [10]; and (iii) a JSON file containing the radar tracking measurements.

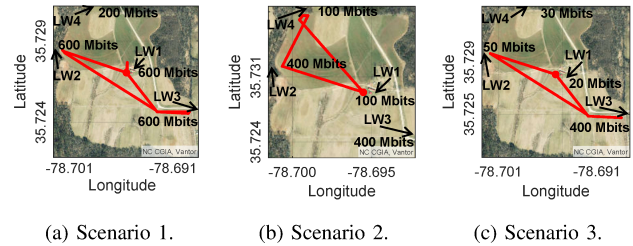


FIGURE 4: Representative (team-1047) autonomous UAV trajectories for data download under three distinct scenarios in the development environment.

The radar files include parameters such as track ID, range, azimuth, elevation, radar cross section (RCS) in dBsm, radial velocity, latitude and longitude, timestamps, as detailed in Table 5. The post-processing folder includes a Jupyter notebook for converting the radar JSON files into MATLAB *.mat* format, along with additional MATLAB scripts/helper files for loading the data and generating example plots for the RFS and radar datasets, including Fig. 10.

The LoRa directory contains 35 subfolders. Of these, 33 correspond to individual AADM flights, each containing a *LORALog.csv* file with LoRa measurements. These files record key PHY- and packet-level parameters, including carrier frequency, signal bandwidth, spreading factor, data rate index, RSSI, SNR, packet reception timestamps, and gateway geolocation information, as summarized in Table 5. An additional subfolder, *AADM\_Collections\_Fig\_11*, contains the data and the MATLAB code used to generate Fig. 11. The directory also includes a consolidated file, *All\_Logs.csv*, which aggregates LoRa measurements from all flights into a single CSV. Two MATLAB scripts, *extract\_packets\_vars\_usage.m* and the helper function *extract\_packets\_vars.m*, are provided to extract relevant variables from the CSV files and load them into the MATLAB workspace. The final subfolder, *separate\_vehicle\_logs*, contains scripts used to organize LoRa measurements by flight. Since a single *vehicleOut.txt* ground-truth file was recorded across all flights, these scripts align timestamps between *All\_Logs.csv* and the vehicle log to partition the LoRa data into the 33 per-flight subdirectories described above.

## INSIGHTS AND NOTES

The following representative observations are provided to guide users in interpreting the dataset and understanding common trends across UAV missions.

Fig. 4 shows example autonomous UAV trajectories for representative team 1047 to download data under three different scenarios in the development environment. Based on the data volume assigned to each BS for different scenarios, the UAV follows a planned path that brings it close to multiple BSs to download data with priority. The red line indicates the flight trajectory, and the annotated values show the total data to be downloaded from each BS. It is clear from Fig. 4 that the UAV spends more time near certain BSs to maximize

TABLE 2: Dataset summary across the various AADM development and experimentation phases in 2025. “All” denotes the use of all sensing modalities, including USRP, LoRa, Fortem Radar, and Keysight RFS data.

Phase	Timeline	Flights	Altitude	Modalities	Purpose
Testbed	Mar-Apr	2	Fixed	USRP	Guide the team to prepare their approach to how the RW environment changes.
Testbed	May	2	Fixed	USRP	Demonstrated the challenge during the ACW workshop.
Development	Jun	45	Varied	USRP	Teams’ proposed approaches were evaluated in the DT.
Testbed	Jun	8	Varied	USRP	Evaluate altitude effects, channel sounding, LTE transmission behavior, and the impact of different transmit antennas at the LW1.
Testbed	Sep	33	Varied	All	Teams’ proposed approaches were finally evaluated in the testbed.
Testbed	Sep	3	Varied	Radar + RFS	Improved UAV localization and tracking.

TABLE 3: Selected raw and processed files with descriptions used in the AADM challenge.

File	Description
controller_log.txt	Initial data assignments, periodic SNR, BS selection, and control data download when to start and stop.
report_log.txt	Controller-UAV message exchanges, including periodic SNR reporting, BS selection, and cumulative downloaded data.
vehicleOut.txt	Time-stamped UAV telemetry, including latitude, longitude, altitude, orientation, battery level, and control states.
vehicleOut_snr.csv	All fields from <i>vehicleOut.txt</i> and each BS SNR value from <i>controller_log.txt</i> after converting them to CSV.
vehicleOut_down.csv	Fields from <i>vehicleOut.txt</i> and each BS cumulative download from <i>controller_log.txt</i> after converting them to CSV.

TABLE 4: Fields used in *vehicleOut\_snr.csv* and *vehicleOut\_down.csv*.

Field	Description
time	Timestamp of the record (approx. 1 sec interval).
Altitude	UAV altitude in meters.
BatteryVolts	UAV battery voltage in volts.
GPSFix	GPS fix type/quality indicator.
Latitude, Longitude	UAV latitude and longitude in degrees.
NumberOfSatellites	Number of visible satellites.
Pitch, Roll, Yaw	UAV pitch, roll, and heading angles in radians.
VelocityX, VelocityY, VelocityZ	UAV velocity components along <i>X</i> , <i>Y</i> , and <i>Z</i> (m/s).
dI_lwN (USRP)	Cumulative data download from BS <i>N</i> (Mbits).
snr_lwN (USRP)	SNR for the BS <i>N</i> link in dB.

the data download, while in others it follows a broader path that leads to smaller data collection.

Fig. 5 shows the UAV altitude and speed over time for representative teams operating under the same outdoor scenario. Across all teams, the UAV ascends from its starting position at the beginning of the mission and descends near the end. Also, the UAV maintains a nearly constant altitude, with differences in the selected operating altitude reflecting distinct trajectory design choices made by each team. These altitude selections affect both LOS conditions and the effective distance to the BSs, which in turn affect the download performance.

Fig. 5 also shows the UAV speed as a function of time, revealing a consistent pattern of acceleration and deceleration throughout the mission. The UAV accelerates while traveling toward a waypoint and decelerates upon arrival, sometimes briefly hovering before proceeding to the next waypoint. This is observed across all teams, which indicates waypoint-based navigation. Differences in the frequency and duration

TABLE 5: Key fields used in *radar\_data\_N.json* and *LORA-log.csv*.

Field	Description
id (Radar)	Identification number of the track.
range, azimuth, elevation (Radar)	Target range (m), azimuth, and elevation angle (deg).
rcsDbsm (Radar)	Radar cross section (dBsm, relative to 1 m <sup>2</sup> ).
radialVelocity (Radar)	Radial velocity (m/s); positive values indicate receding targets.
lla (Radar)	Track latitude, longitude (deg), altitude (m).
gpsX (Radar)	GPS time at CPI center.
tx_info_modulation_bandwidth (LoRa)	LoRa signal bandwidth (Hz).
tx_info_frequency	LoRa carrier frequency (Hz).
tx_info_modulation_spreadingFactor	LoRa spreading factor (SF7-SF10).
dr (LoRa)	LoRa data rate index.
rx_info_rssi (LoRa)	Received signal strength indicator (dBm).
rx_info_snr (LoRa)	Signal-to-noise ratio (dB).
rx_info_time (LoRa)	Packet reception time (UTC).
rx_info_N (LoRa)	N is latitude, longitude (deg), or altitude (m).

of speed fluctuations indicate variations in waypoint density and maneuvering strategies. Since UAV speed determines how long the UAV can associate with a BS based on the link quality, these mobility patterns play an important role in shaping data collection opportunities during the mission.

Fig. 6 shows the UAV trajectories for the representative teams under Scenario 1 in the outdoor environment. Although all teams operate within the same geographic area and mission constraints, their trajectories differ noticeably in shape and coverage. Each team adopts a distinct strategy for visiting BSs based on annotated downloaded data volumes, which is reflected in the different flight paths.

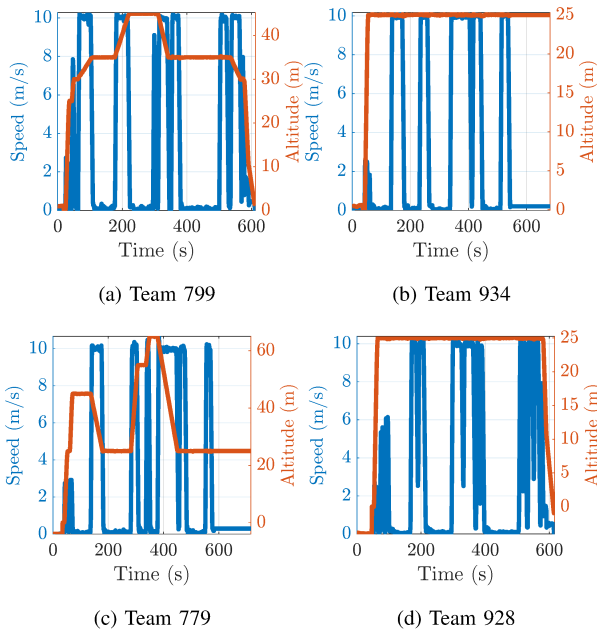


FIGURE 5: UAV altitude and speed over time for representative teams for Scenario 1 in the outdoor environment.

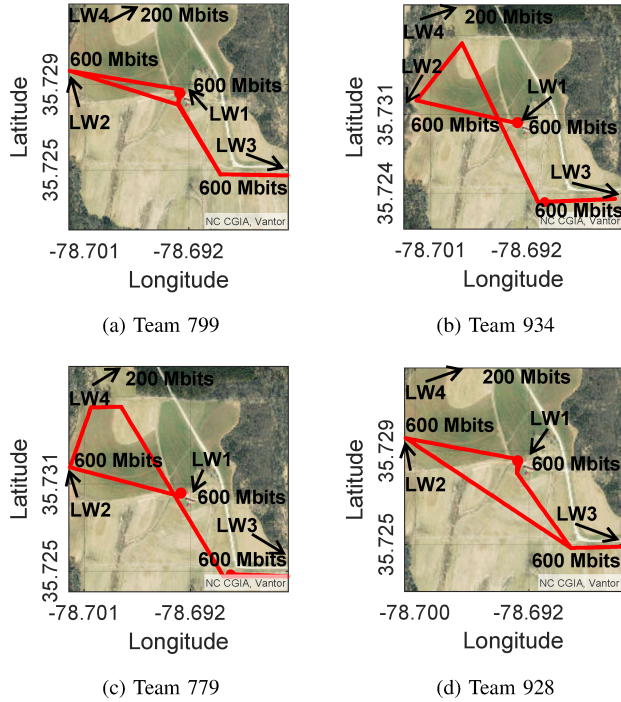


FIGURE 6: UAV trajectory for representative teams for Scenario 1 in the outdoor environment.

Fig. 7 compares the received signal power and the UAV-BS distance over time for representative teams in Scenario 1, shown for both the DT and RW environments. In all cases, received power generally decreases as the UAV moves farther from the serving BS and increases as the UAV approaches it. However, the plots also show noticeable deviations from a simple monotonic relationship. In the DT environment,

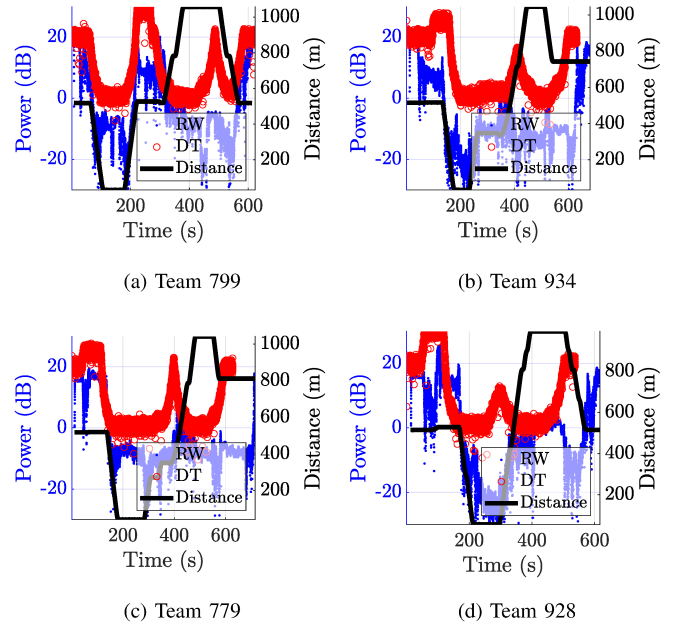


FIGURE 7: Comparison of received signal power and distance over time with respect to BS1 in DT and RW for representative teams for Scenario 1 in the outdoor environment.

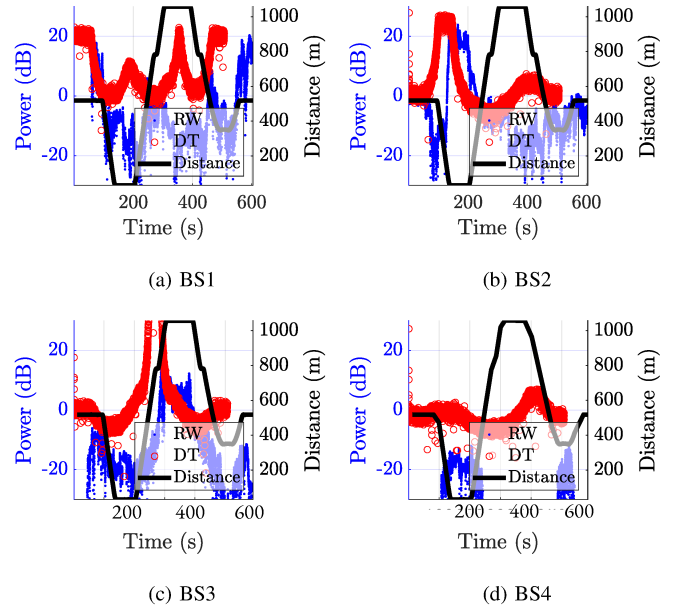


FIGURE 8: UAV trajectory and the received signal power with respect to BS1, BS2, BS3, and BS4 for team 779 for Scenario 2 in the outdoor environment.

received power follows a smoother trend with distance. In contrast, in the RW environment, it exhibits stronger fluctuations due to RW effects such as multipath, shadowing, antenna orientation, and environmental obstructions. These differences highlight the gap between DT and RW propagation and demonstrate the value of the dataset for studying DT-to-RW discrepancies.

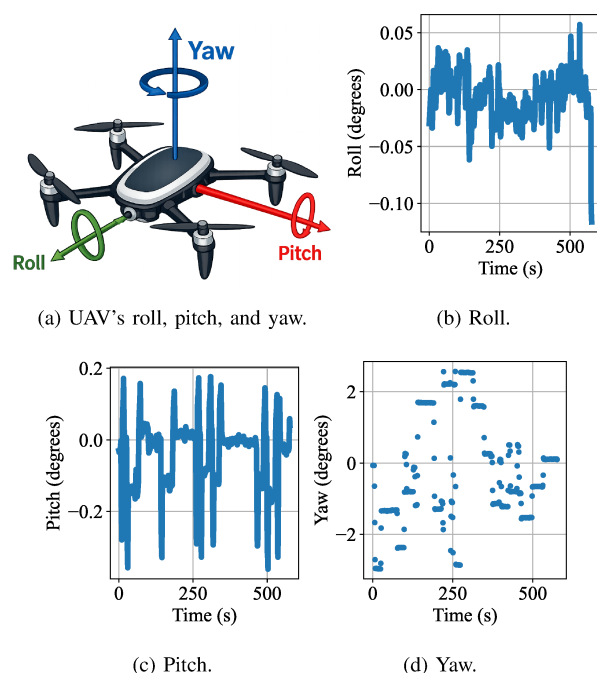


FIGURE 9: Time-series plots of the UAV’s roll, pitch, and yaw angles for Team 799 in Scenario 1.

Fig. 8 shows the received signal power and UAV–BS distance over time for the representative team 779 with respect to BS1 through BS4 under Scenario 2 in the outdoor environment. For each BS, the received power increases as the UAV approaches the station and decreases as it moves away. We observe strong received power for a BS when the UAV remains nearby, but weaker power as the UAV moves away due to increased distance and geometry. However, RW measurements exhibit noticeable fluctuations compared to the smoother DT traces due to environmental effects such as multipath and shadowing. Fig. 9 shows the UAV’s orientation, i.e., its roll, pitch, and yaw angles for Team 799 in Scenario 1 in an outdoor environment. The roll and pitch correspond to the UAV’s tilting motion, and the yaw represents changes in heading. Fig. 9(b) and Fig. 9(d) show that the roll and pitch remain relatively small for most of the mission, showing short spikes that may occur during turns or speed changes. However, as shown in Fig. 9(d), the yaw angle changes more noticeably over time, which captures the UAV’s frequent reorientation. Although the roll and pitch angles are small in magnitude, they can still affect antenna alignment and link quality. These measurements enable the dataset to be used to study how UAV dynamics influence wireless performance in RW conditions.

Fig. 10 illustrates RFS and radar localization over the duration of the UAV’s trajectory for a representative flight. RF-based 2D position estimates exhibit variability and intermittent localization failures, but also demonstrate periods of moderate accuracy when favorable LOS and sensor geometry conditions are present, enabling wide-area, passive localization. In contrast, the radar system provides continuous 3D

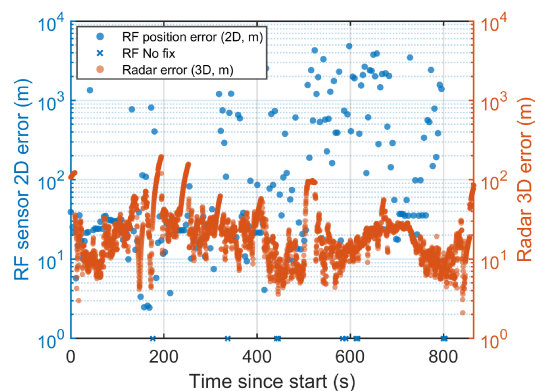


FIGURE 10: Localization performance comparison between the Fortem R20 radar and the Keysight N6841A RFS over time for improved localization and tracking Flight 2.

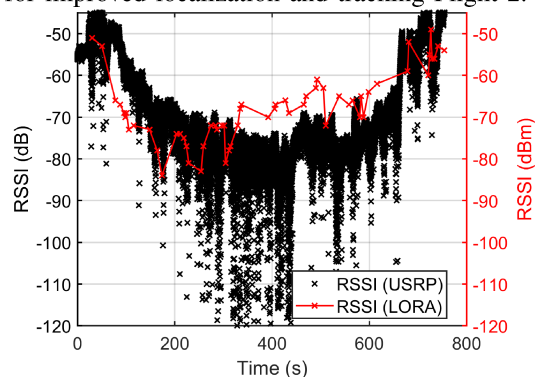


FIGURE 11: Normalized LoRa and USRP RSSI time series measured from AERPAW tower LW1 during a test flight conducted in September before the AADM Challenge.

localization with consistently lower and more stable error. Together, these results highlight the complementary roles of RF sensing for passive, wide-area detection and radar for robust, high-accuracy tracking [8].

Fig. 11 shows the RSSI time series for a representative flight measured at AERPAW tower LW1, which is equipped with both a USRP and a LoRa gateway. The LoRa RSSI exhibits relatively smooth, slowly varying trends over time that are consistent with large-scale path loss and UAV trajectory geometry, in contrast to the higher-rate fluctuations and deeper fades observed in the co-located USRP RSSI. These differences are consistent with the higher carrier frequency of the USRP measurements, which are more susceptible to foliage blockage and small-scale fading, whereas the lower-frequency LoRa signals experience improved penetration through vegetation and reduced vulnerability to deep fades. Although the LoRa and USRP RSSI magnitudes are not directly comparable, and a fixed offset of -65 dB was applied to the USRP measurements for visualization purposes, the contrast in temporal behavior is robust across the trajectory.

Finally, Fig. 12 illustrates a comparison of the AADM scores of all final participating teams. It demonstrates how each team performed across the three scenarios (denoted as Scen in Fig. 12) in both the DT and RW testbeds under the same mission rules. We discuss all teams’ details

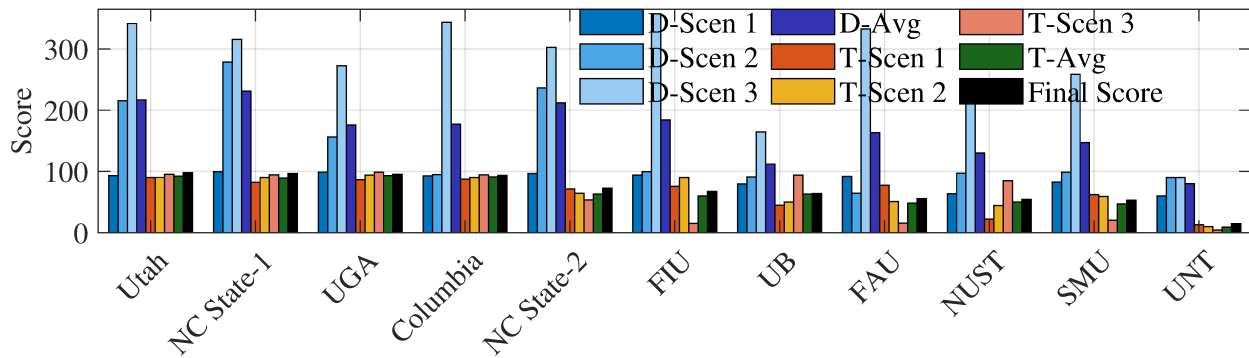


FIGURE 12: Scenario-wise, average and final AADM scores across teams. D and T denote DT and RW experiment modes.

earlier in the *Dataset Overview and Uniqueness* subsection. Some teams performed well across all scenarios, while others showed strong results only in specific cases and weaker performance. The figure also shows that success in a single scenario is not sufficient to achieve a high final score. A strong DT score does not guarantee a good score overall. Teams that maintained stable performance across all scenarios generally achieved higher overall rankings. This highlights the importance of designing autonomous trajectories that remain reliable across different operating conditions rather than being tuned to a single scenario.

### SOURCE CODE AND SCRIPTS

All source code and post-processing scripts employed in the cleaning, merging, and analysis of the dataset are made publicly available alongside the dataset at Dryad (DOI: 10.5061/dryad.7d7wm3898). The dataset includes MATLAB source code to merge log files to generate datasets in CSV format, compute SNR and cumulative download statistics, and recreate the trajectory, altitude, speed, and power-versus-distance plots. Other scripts are used to process the LoRa gateway logs to obtain RSSI and SNR statistics, convert the Fortem radar JSON data to a MATLAB format, and combine the Fortem, Keysight RF sensor data, and UAV ground truth data while also computing localization errors. Other LTE processing scripts are included to generate the RSRP versus distance plots from the raw IQ recordings.

### ACKNOWLEDGEMENTS AND INTERESTS

Special thanks to NI, Keysight Technologies, AUVSI NC, and 6GNC for sponsoring the challenge awards.

This work was supported by the NSF under Grant CNS-1939334 and in part by CNS-2450593 (AERPAAW). Additional support for participating teams was provided in part by NSF Grants 2332661, OAC-2512931, 2346555, CNS-2144297, CNS-2148128, CNS-2332662, CNS-2117822, EEC-2133516, CNS-2440756, SWIFT-2229563, DGE-2137100, and by HEC, PK.

M. S. Hossen prototyped the AADM challenge software framework, led the dataset curation process, processed and analyzed the USRP measurements, and prepared the initial manuscript draft. C. Dickerson processed and validated the

radar and RFS datasets and assisted with manuscript preparation. O. Ozdemir processed and analyzed the LoRa and LTE measurements and validated the dataset.

The AERPAAW platform development and the execution of the field experiments were carried out by O. Ozdemir, A. Gurses, M. R. Sarbudeen, T. Zajkowski, I. Guvenc, M. Sichitiu, M. Mushi, and R. Dutta.

All remaining authors participated in the AADM challenge and contributed to algorithm development in the digital twin environment, experimental execution, data validation, and critical review of the manuscript.

The authors declare no conflicts of interest.

### REFERENCES

- [1] "AERPAAW: Aerial experimentation and research platform for advanced wireless," <https://aerpaw.org/>, NSF Platforms for Advanced Wireless Research (PAWR), accessed: 2025-12-12.
- [2] F. Betti Sorbelli, "UAV-based delivery systems: A systematic review, current trends, and research challenges," *ACM J. Auton. Transport. Syst.*, vol. 1, no. 3, May 2024.
- [3] M. S. Hossen, O. Ozdemir, M. L. Sichitiu, and I. Guvenc, "Resilient UAV data mule via adaptive sensor association under timing constraints," in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, USA, Mar. 2026.
- [4] A. Gurses, G. Reddy, S. Masrur, O. Ozdemir, I. Guvenc, M. Sichitiu, A. Sahin, A. Alkhateeb, M. Mushi, and R. Dutta, "Digital twins and testbeds for supporting AI research with autonomous vehicle networks," *IEEE Comm. Mag.*, 2024.
- [5] S. Masrur, O. Ozdemir, A. Gurses, I. Guvenc, M. Sichitiu, R. Dutta, M. Mushi, T. Zajkowski, C. Dickerson, G. Reddy, S. V. Villar, C.-W. Wong, B. Chatterjee, S. Chaudhari, Z. Li, Y. Liu, P. Kudyba, H. Sun, J. S. Mandapaka, K. Namuduri, W. Wang, and F. Fund, "Collection: UAV-based RSS measurements from the AFAR challenge in digital twin and real-world environments," *IEEE Data Descr.*, pp. 1–10, 2025.
- [6] J. J. Sadique, M. R. Khan, and A. S. Ibrahim, "UAV-aided fast data collection via machine learning using aerpaw's digital twin," in *Proc. ACM Worksh. Wireless Netw. Testbeds*, ser. WiNTECH '25, New York, NY, USA: Assoc. Comput. Mach. (ACM), 2025, p. 89–96.
- [7] S. Javed, A. Hassan, R. Ahmad, M. M. Alam, M. S. Hossen, A. Gurses, and O. Ozdemir, "Autonomous uav trajectory design and evaluation for data muling in aerpaw digital twin," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Glasgow, United Kingdom, May 2026.
- [8] C. Dickerson, S. Kearney, S. Manjur, I. Guvenc, S. Gurbuz, A. Gurbuz, O. Ozdemir, and M. Sichitiu, "Fusion of cellular ISAC and passive RF sensing for UAV detection and tracking," 2025. [Online]. Available: <https://arxiv.org/abs/2512.14608>
- [9] 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 Int. Conf. Commun. Wkshps. (ICC Workshops)*, Montreal, QC, Canada, June 2025, pp. 671–677.
- [10] A. H. F. Raouf et al., "Wireless datasets for aerial networks," 2025. [Online]. Available: <https://arxiv.org/abs/2510.08752>