

RESEARCH ARTICLE

OPEN ACCESS

Ground Monitoring Console with Causal Link-Quality Prediction

A. Aswini¹ and Prof. S. Varadarajan²

¹ M. TECH Student, Department of Electronics and Communication Engineering, Sri Venkateswara University College of Engineering, Tirupati, 517502, India

² Professor, Department of Electronics and Communication Engineering, Sri Venkateswara University College of Engineering, Tirupati, 517502, India

Abstract— Ground monitoring becomes more useful when aircraft state, communication condition, time, and location are interpreted together rather than as independent streams. This paper presents an integrated ground-monitoring workflow that combines historical flight-telemetry visualization with a causal multi-horizon machine-learning pipeline for predicting future wireless-link change. The main prediction target is future change in signal-to-interference-plus-noise ratio (Δ SINR) at horizons of 1, 2, 5, 10, 15, 20, and 30 s. The quantitative study uses three physical flight records drawn from two public AERPAW measurement sources: two Ericsson 5G NSA traces at yaw orientations of 45° and 315°, and one Lake Wheeler Android-based 4G LTE/5G NR trace. The records contain LTE/NR radio indicators, throughput, position, altitude, cell information, and motion/orientation fields, with 2376 raw observations across the three physical flights. The processing uses a causal 1-s grid, flight-isolated temporal features, chronological evaluation with a horizon-sized purge, training-only mutual-information feature selection, and a persistence baseline. Three classical regressors—HistGradientBoosting (HGB), Random Forest, and Extra Trees—are evaluated. An engineering link-quality index (LQI) in the range 0–100 summarizes 12 available cellular and throughput indicators. At the predeclared 30-s operating point, HGB obtains a Δ SINR MAE of 3.110 dB versus 3.669 dB for persistence, a 15.23% reduction. A separate NASA Tail-654 ground-monitoring implementation replays a historical MATLAB flight record, presents map-based trajectory and primary telemetry, and applies phase-aware statistical anomaly monitoring using robust envelopes, eight-second persistence, and recovery hysteresis. The NASA component is kept experimentally separate from the wireless-link prediction metrics. The resulting system connects telemetry observability with forward-looking communication-state information while preserving a strict separation between information available at prediction time and future observations.

Keywords— UAV telemetry, ground monitoring console, causal machine learning, link-quality prediction, SINR prediction, 5G, LTE, link quality index, anomaly detection, aerial communications.

I. INTRODUCTION

Unmanned aerial vehicles increasingly depend on wireless links for command, control, telemetry, navigation support, and payload communication. Consequently, the ground segment must do more than display a vehicle position: it must transform heterogeneous telemetry and radio measurements into information that can be interpreted rapidly by an operator. Recent ground-station research has demonstrated integrated UAV monitoring and mission-management interfaces, emphasizing modularity, real-time telemetry, visualization, and operator usability [1]. Related

work on integrated satellite ground services likewise shows the value of combining telemetry reception, operational control, visualization, and data management in a unified ground architecture [2]. Earlier telemetry-processing research established the importance of converting received telemetry into engineering quantities and synchronizing measurements before analysis [4], while work on portable command-and-telemetry databases emphasized reusable interfaces and reduced coupling to a particular ground-control system [5].

A complementary requirement is timely awareness of communication degradation. Traditional aircraft recording systems are principally designed for post-event analysis, whereas continuous transmission of aircraft information to ground systems can support earlier intervention and off-board situational awareness [6]. In wireless-network research, machine-learning-based link-quality estimation has been extensively studied, but the design choices surrounding target definition, temporal features, data leakage, evaluation procedures, and application requirements strongly affect reported performance [3]. For UAV cellular links specifically, measurement studies have shown that aerial mobility, altitude, antenna configuration, and network geometry materially affect radio and application-level performance [9,10]. Recent AERPAW work further demonstrates the value of synchronized aerial KPI datasets for data-driven cellular analysis [7,8].

The work was therefore organized around two complementary functions. The first is a ground-monitoring console that makes a historical aircraft record easy to inspect by combining a map, timeline controls, primary telemetry, flight phase, and persistent-anomaly information. The second is a causal link-prediction pipeline that forecasts future Δ SINR, maps the result to an engineering LQI, and produces a compact future link state suitable for presentation at the ground station. The final quantitative experiment deliberately keeps the model set small—HGB, Random Forest, and Extra Trees, with persistence as the reference—so that the effect of causal feature construction can be examined without mixing fundamentally different model families.

The main contributions are:

1. an integrated ground-monitoring architecture that connects telemetry acquisition, temporal preprocessing, causal link prediction, future LQI calculation, visualization, and operator alerts;
2. a leakage-controlled multi-horizon formulation for future Δ SINR prediction with explicit chronological splitting and horizon-sized purging;

3. a weighted, component-aware 0–100 LQI for translating multiple heterogeneous cellular indicators into an operator-facing engineering quantity;
4. an evaluation of HGB, Random Forest, and Extra Trees against persistence, including multi-horizon results, feature ablations, confidence intervals, event detection, and leave-one-flight-out analysis; and
5. a telemetry-replay console that complements communication prediction with aircraft-state visualization and phase-aware persistent-anomaly monitoring.

II. RELATED WORK

A. Ground Monitoring and Telemetry Processing

Júnior et al. developed an integrated UAV ground station for telemetry, mission planning, and mission management, emphasizing modular monitoring and operator interaction [1]. Liu et al. proposed an integrated satellite ground service platform combining telemetry, command, operational control, storage, and data services [2]. Kim and Lee described structured telemetry data processing and synchronization for flight-test measurements [4]. Segal investigated reusable command-and-telemetry processing tools designed to reduce dependence on a particular ground-control system [5]. Alqodah and Matalgah discussed the requirements and challenges of transferring aircraft information to ground-based systems for monitoring and analysis [6].

These studies establish the value of structured telemetry processing, reusable interfaces, and ground-side observability. The present project extends that ground-segment perspective by exposing a predicted communication state alongside current aircraft telemetry and trajectory information.

The communication-aware component is consistent with research on cellular-connected aircraft and aerial wireless propagation, where altitude, mobility, geometry, and propagation conditions influence link behaviour [11], [12].

B. Machine Learning for Wireless Link Quality

Cerar et al. surveyed machine-learning approaches for wireless link-quality estimation and highlighted the importance of data characteristics, feature design, target definitions, and evaluation methodology [3]. The present work follows this emphasis by defining an explicit future-change target and constraining temporal features to information available at prediction time. Prior link-quality prediction studies use historical context and online adaptation for future-link inference [13], [14], while spatial-temporal work explicitly models mobility-dependent future variation [15].

The experimental data are drawn from public AERPAW aerial cellular measurements. The curated AERPAW literature describes synchronized RF, KPI, throughput, and trajectory measurements suitable for data-driven aerial-network research [7], while the AERPAW Lake Wheeler study analyses 4G/5G air-to-ground KPIs using lightweight machine-learning models [8]. Because those studies use different targets and evaluation protocols from the present

causal 30-s Δ SINR forecasting problem, no numerical performance comparison is made here.

III. SYSTEM ARCHITECTURE

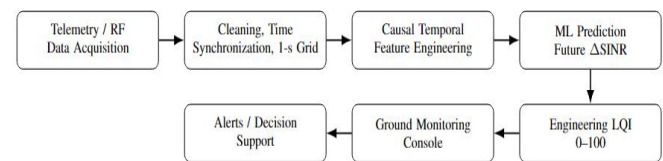


Fig. 1. Integrated ground-monitoring workflow. Aircraft telemetry and cellular measurements are cleaned and synchronized on a 1-s grid, followed by causal temporal feature engineering and future Δ SINR prediction.

The proposed system combines the telemetry-console and causal link-prediction functions in a single ground-monitoring workflow. The architecture is shown in Fig. 1. The data-acquisition stage can receive aircraft telemetry, cellular RF KPIs, and historical flight-record data used for replay. Data are then cleaned and synchronized onto a 1-s grid. Causal temporal feature engineering generates only information available at prediction time. The machine-learning stage predicts future Δ SINR, after which the result is converted to a future LQI. The ground console presents aircraft state, current link state, predicted link state, and decision-support information.

IV. DATA AND PREPROCESSING

A. Experimental Data

The quantitative experiment uses three independent physical flight traces from two public AERPAW measurement sources. Two traces come from the Ericsson 5G NSA network RF and throughput measurement release and correspond to yaw orientations of 315° and 45°. The third is the Lake Wheeler Android-based 4G LTE/5G NR and throughput measurement trace used as the primary native-flight experiment. The AERPAW literature describes these resources as synchronized aerial measurements containing cellular KPIs together with position, altitude, trajectory, and application-level measurements [7], [8]. The present study does not treat the sources as interchangeable: the Ericsson traces are used for cross-flight transfer analysis, whereas the Lake Wheeler trace is the primary chronological test environment.

The Ericsson traces provide modem-side RF and throughput indicators together with UAV geometry and orientation. The Lake Wheeler trace provides a different measurement context based on Android-derived LTE/NR observations and throughput. The resulting differences in measurement devices, logging pipelines, trajectories, and experimental conditions are retained because they are part of the cross-flight generalization problem [8]. The final feature matrix uses the available radio, mobility, throughput, and temporal information rather than assuming that every field is present in every source.

TABLE I. DATASETS AND PHYSICAL FLIGHT TRACES USED IN THE PROJECT.

Source / trace	Raw rows	Role	Principal information used
Ericsson 5G NSA, yaw 315°	769	Cross-flight evaluation	LTE/NR SINR, RSRP, RSRQ, CQI, MCS, RI, cell information, throughput, geolocation, altitude, and motion/orientation.
Ericsson 5G NSA, yaw 45°	798	Cross-flight evaluation	Same Ericsson measurement family; retained as a physically distinct flight for leave-one-flight-out analysis.
Lake Wheeler Android LTE/NR	809	Primary native-flight experiment	4G LTE/5G NR radio KPIs, throughput, latitude, longitude, altitude, and synchronized time information.
Total physical-flight rows	2376	Final quantitative study	Three independent physical traces; duplicate copies of the same physical traces are excluded from flight-level generalization analysis.

The project also contains a separate NASA Tail-654 flight-record source for the ground-monitoring console. The project documentation describes the source as a MATLAB flight-record collection with 186 parameter definitions, including sensor data, sampling rates, units, descriptions, and parameter identifiers. The supplied implementation targets the flight file 654200107020629.mat, corresponding to a start time of 2001-07-02 06:29:00 UTC. The notebook searches recursively for an existing copy and provides a NASA download fallback when the file is not already available. It dynamically discovers MATLAB parameter structures, converts the stored data to numeric arrays, treats the NASA invalid sentinel value -390 and nonphysical values above 10^{12} in magnitude as missing, and resamples parameters to a 1-Hz replay grid. Short missing intervals are linearly interpolated for dashboard continuity.

The console extracts latitude, longitude, altitude, ground speed, heading, true track (TRK), and magnetic track (TRKM). The implementation divides the replay into four operational phases—GROUND, TAKEOFF/CLIMB, CRUISE, and DESCENT/LANDING—using fixed time indices at 1316, 2821, and 9397 s. The dashboard provides a 50/50 map-and-telemetry layout, UTC timestamps, an aircraft symbol that follows heading, received and future trajectory views, parameter search and selection, and phase controls for takeoff, cruise, and landing. Progress graphs are based on received values rather than future trajectory samples.

All prediction features are constructed independently within each physical flight. Measurements are aligned to a causal 1-s grid. Missing and invalid sentinel values are treated as missing observations; artificial zero imputation is disabled. Where a model pipeline requires imputation, median imputation is applied within the training pipeline rather than before the train/test separation.

B. For a prediction issued at time t , temporal features use samples at or before t . Lagged and rolling statistics are formed after a one-sample shift so that a rolling window cannot contain the current or future target sample inadvertently. Temporal features are never carried across flight boundaries.

V. CAUSAL PREDICTION METHODOLOGY

A. Prediction Target

The principal target is future change in NR SINR:

$$\Delta \text{SINR}_{t,h} = \text{SINR}_{t+h} - \text{SINR}_t,$$

where $h \in \{1,2,5,10,15,20,30\}$ s. The 30-s horizon is the predeclared primary operating point. Predicting a change rather than the absolute future SINR makes the output directly interpretable as link improvement or degradation relative to the current state.

A persistence baseline predicts

$$\widehat{\Delta \text{SINR}}_{t,h}^{\text{pers}} = 0,$$

which corresponds to the physically meaningful assumption that the link will not change over the prediction interval.

B. Temporal Feature Construction

The feature set combines radio, mobility, throughput, and temporal information. Candidate radio variables include NR/LTE SINR, RSRP, RSRQ, RSSI, CQI, MCS, and rank indicator; throughput and aircraft-position/mobility variables are also available. Lagged values, rolling statistics, exponentially weighted statistics, first differences, and short-term trends are computed causally within each flight. Training-only mutual-information selection limits the final feature set to at most 50 features.

To prevent temporal leakage, the data are split chronologically. For each prediction horizon, a purge gap at least equal to the horizon is removed around the train/test boundary. For the primary 30-s experiment, the purge is 30 s. Four walk-forward validation splits are used for model assessment inside the training portion. The use of forecasting-oriented temporal validation is also consistent with recent RAN time-series research, where forecasting and anomaly detection are treated as sequential learning tasks rather than i.i.d. classification problems [18].

C. Regression Models

Three classical regressors are evaluated:

- HistGradientBoosting (HGB), with 300 boosting iterations, learning rate 0.04, maximum 15 leaf nodes, and L_2 regularization;
- Random Forest with 300 trees, maximum depth 18, minimum leaf size 2, and square-root feature subsampling; and
- Extra Trees with 300 trees, maximum depth 18, minimum leaf size 2, and square-root feature subsampling.

The final primary 30-s HGB model is predeclared as the native-flight baseline model; the test set is not used to switch the primary model after observing test performance.

D. Engineering Link-Quality Index

The ground console requires a compact representation of multi-dimensional link state. Each available quality component x_j is mapped to a 0–100 score. For a higher-is-better variable,

$$s_j(t) = 100 \text{ clip} \left(\frac{x_j(t) - l_j}{u_j - l_j}, 0, 1 \right),$$

where l_j and u_j are engineering lower and upper bounds. A weighted LQI is then calculated only from components that are actually available:

$$\text{LQI}(t) = \frac{\sum_{j \in \mathcal{A}(t)} w_j s_j(t)}{\sum_{j \in \mathcal{A}(t)} w_j},$$

where $\mathcal{A}(t)$ denotes the set of available components. This renormalization prevents a missing measurement from being treated as a zero-quality observation. BER is included only if a real BER measurement is present; no BER value is fabricated or inferred in the final run.

The final experiment used 12 components: NR SINR, LTE SINR, NR RSRP, LTE RSRP, NR RSRQ, LTE RSRQ, LTE RSSI, NR CQI, LTE CQI, NR MCS, NR RI, and throughput. The component weights are renormalized over the measurements available at each time step; the detailed weight table is omitted here to keep the presentation focused on the predictive results.

The continuous index is clipped to 0–100 and quantized as

$$\text{LQI}_Q(t) = \text{round}[\text{clip}(\text{LQI}(t), 0, 100)],$$

producing 101 possible integer states. For operator display, the following descriptive bands are retained: poor < 40 , fair $40 \leq 60$, good $60 \leq 80$, and excellent ≥ 80 . These bands are not substitutes for the 101-state LQI.

E. Future LQI Construction

The primary model predicts future NR SINR. The predicted SINR is inserted into the LQI calculation while the other required quality components remain at their causal current values. Thus,

$$\widehat{\text{LQI}}_{t+h} = F(\widehat{\text{SINR}}_{t+h}, \mathbf{x}_{-\text{SINR},t}),$$

where $F(\cdot)$ is the engineering LQI function. A separate direct future-LQI regression is also evaluated as a secondary benchmark because it optimizes the LQI target directly rather than through the SINR-to-LQI engineering conversion.

F. Event Detection and Uncertainty

Two physical degradation definitions are used: $\Delta \text{SINR} \leq -2$ dB for link-degradation events and $\Delta \text{LQI} \leq -5$ points for LQI degradation. The continuous regression output is used as the event score. Decision thresholds are calibrated only on walk-forward out-of-fold training predictions using a prevalence-constrained maximum-F1 procedure; the test labels are not used to choose the threshold.

The second system function is a telemetry-oriented ground console that replays a historical aircraft record and presents the aircraft state in an operator-oriented dashboard. The supplied implementation is based on the NASA Tail-654 flight record described in Section 4.1. It is intentionally separate from the AERPAAW wireless prediction pipeline so that aircraft-state anomalies are not confused with cellular-link prediction performance.

Fig. 2. NASA Tail-654 ground monitoring console during telemetry replay, showing the map trajectory, replay controls, primary telemetry, flight phase, and persistent-anomaly monitoring.

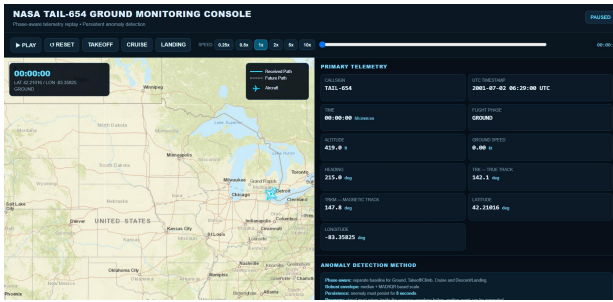
The console establishes a UTC timeline from the selected Tail-654 record and converts the discovered MATLAB parameters into a normalized display representation. The map displays the received trajectory and the future portion of the trajectory for replay context, while the telemetry panel presents latitude, longitude, altitude, ground speed, heading, true track, magnetic track, time, and flight phase. The aircraft symbol rotates with heading, parameter units are

displayed beside primary telemetry, and replay controls support play, reset, speed selection, and direct navigation to takeoff, cruise, and landing.

The notebook also provides parameter search and selection so that additional numerical parameters can be inspected without changing the underlying MATLAB parser. Progress graphs use received values only. The phase-aware anomaly monitor uses separate baselines for GROUND, TAKEOFF/CLIMB, CRUISE, and DESCENT/LANDING. For scalar parameters, the phase centre is the median and the robust scale is the larger of $1.4826 \times \text{MAD}$, $\text{IQR}/1.349$, and a small magnitude-dependent floor; the main envelope is $m \pm 5.5\sigma$ and the recovery envelope is $m \pm 4.0\sigma$. Heading and track use circular angular differences.

A candidate deviation must remain outside the phase-specific anomaly envelope for eight consecutive seconds before a persistent event is emitted. The event is cleared only after the signal returns inside the recovery envelope, providing hysteresis against rapid alarm toggling. Data-quality events are separately recorded when a parameter remains non-finite for at least ten seconds. These statistical envelopes are monitoring aids and are not certified aircraft operating limits. The console exports telemetry JSON, persistent-anomaly CSV/JSON files, phase-threshold JSON, and a standalone HTML dashboard package.

The console architecture follows the ground-system principles emphasized in previous telemetry and ground-station studies [1], [2], [4], [5], while adding the future link-quality output generated by the causal prediction module.



VI. EXPERIMENTAL RESULTS

A. Multi-Horizon Prediction

Fig. 3 summarizes the selected-model MAE across prediction horizons. As the horizon increases, the forecasting problem becomes more difficult. At 30 s, however, the predeclared HGB model achieves lower MAE than persistence. Fig. 4 presents the corresponding performance relative to persistence, normalized to 100%.

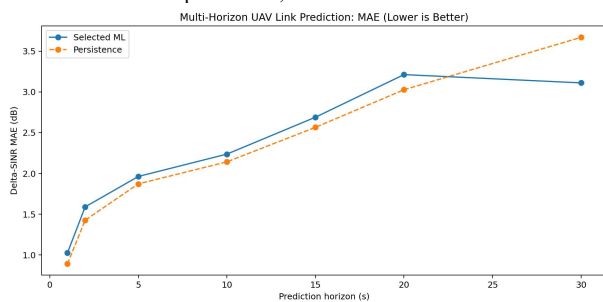


Fig. 3. Multi-horizon SINR prediction MAE.

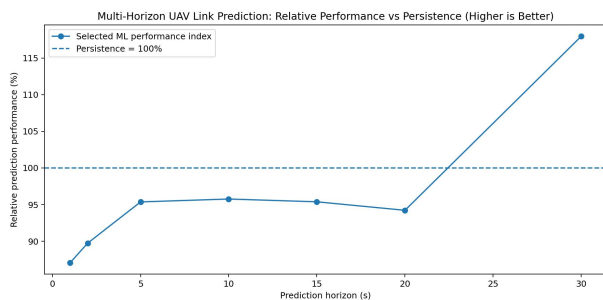


Fig. 4. Relative performance index with persistence normalized to 100%.

The multi-horizon results show a useful pattern rather than a uniform advantage for machine learning. At 1–20 s, the link often changes slowly enough that simply assuming no change is difficult to beat. The 30-s operating point behaves differently: HGB reaches 3.110 dB MAE compared with 3.669 dB for persistence. In other words, the benefit appears at the horizon for which the system was designed, not at every horizon. This is an important practical result because it prevents the model from being presented as a universal replacement for a simple short-term persistence forecast.

B. Primary 30-s Result

The primary 30-s comparison is shown in Fig. 5. The numerical results are given in Table II. The HGB model

obtains MAE 3.110 dB, RMSE 3.704 dB, and $R^2 = 0.258$. Random Forest obtains 3.380 dB MAE and Extra Trees obtains 3.694 dB. Persistence produces 3.669 dB MAE. Therefore, the HGB model reduces MAE relative to persistence by 15.23%.

TABLE II. 30-s Δ SINR PREDICTION ACCURACY AND ERROR METRICS.

Model	MAE	RMS E	R^2	≤ 1 dB	≤ 2 dB	≤ 3 dB	Δ MAE vs. P.
Persistence	3.669	4.346	-0.022	20.53%	37.75%	49.01%	0.00%
HGB	3.110	3.704	0.258	19.21%	33.77%	51.66%	15.23%
Random Forest	3.380	3.979	0.144	13.91%	33.77%	45.70%	7.88%
Extra Trees	3.694	4.361	-0.029	15.23%	31.79%	39.07%	-0.69%

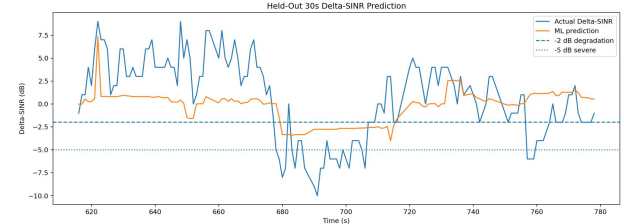


Fig. 5. Primary 30-s SINR prediction comparison.

The point estimate should not be read in isolation. A 15-s moving-block bootstrap with 3000 resamples gives a 95% interval of 2.534–3.683 dB for HGB MAE and 2.894–4.483 dB for persistence. The estimated 15.23% improvement has a 95% bootstrap interval of 0.72–27.89%. The lower bound is close to zero, so the appropriate interpretation is that the experiment provides evidence of an improvement at the selected operating point, while the size of that improvement remains uncertain on this limited test set.

C. Feature Ablation

The ablation study gives a clearer explanation of where the predictive signal comes from. Using only radio measurements gives 5.215 dB MAE, whereas mobility alone gives 3.640 dB. Combining radio and mobility reduces the error to 3.094 dB. The largest improvement in this controlled ablation is obtained when temporal context is added to the radio features, giving 2.917 dB. Removing throughput from the full feature pool gives 2.943 dB, slightly below the 3.110 dB obtained when throughput is included. Thus, the result is not simply “more features are better”; the temporal structure of the flight appears to carry useful information, while the contribution of individual instantaneous variables depends on the feature combination.

TABLE III. 30-s HGB FEATURE ABLATION ON THE PRIMARY FLIGHT.

Feature set	Feature s	MAE (dB)	R^2
Radio only	13	5.215	-0.990
Mobility only	8	3.640	-0.008
Radio + Mobility	21	3.094	0.254
Radio + Temporal	50	2.917	0.331
All, no throughput	50	2.943	0.344
All features	50	3.110	0.258

D. Cross-Flight Generalization

Leave-one-flight-out (LOFO) evaluation is used to examine whether the learned relationship transfers between flight records. Fig. 6 shows the resulting MAE values. For HGB,

the held-out yaw-315 flight gives 2.597 dB MAE and the held-out yaw-45 flight gives 2.726 dB, while the held-out Lake Wheeler flight gives 4.292 dB. The corresponding persistence MAEs are 3.720, 3.619, and 3.571 dB. The difference among held-out flights demonstrates that generalization is affected by the underlying flight and measurement conditions.

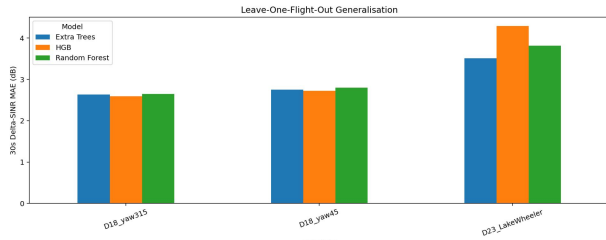


Fig. 6. Leave-one-flight-out 30-s SINR prediction performance.

The augmented-training experiment provides a useful caution. When the additional flights are pooled into training, HGB changes from 3.110 to 3.820 dB MAE, while Random Forest changes from 3.380 to 3.675 dB and Extra Trees from 3.694 to 3.521 dB. The extra data therefore do not produce a uniform gain. This is consistent with the LOFO results: flight records differ enough that simply increasing the number of samples can also increase distribution mismatch. For a deployed system, collecting more data should therefore be accompanied by checking whether the new flights represent the same operating domain.

E. Future LQI Prediction

Fig. 7 shows the future LQI output. The engineering LQI is useful because it maps heterogeneous radio indicators to an operator-facing 0–100 scale. In the 30-s direct future-LQI experiment, HGB obtains 3.183 LQI points MAE, Random Forest 3.200, and Extra Trees 3.998, compared with 4.651 for persistence. The direct HGB model therefore reduces LQI MAE by 31.56% relative to persistence.

random forest 3.200, and Extra Trees 3.998, compared with 4.651 for persistence. The direct HGB model therefore reduces LQI MAE by 31.56% relative to persistence.

In contrast, converting the primary Δ SINR prediction to future LQI gives an MAE of 4.515 points and $R^2 = -1.459$. This difference is important: the SINR predictor is optimized for future SINR change, not for the composite LQI. The direct future-LQI model is therefore retained as a separate engineering benchmark rather than being used to redefine the primary scientific target.

TABLE IV. FUTURE LQI PREDICTION ACCURACY AT 30 s.

Model	MAE (LQI)	RMSE	R^2	≤ 5 LQI	≤ 10 LQI
Persistence	4.651	5.764	-1.623	55.63%	89.40%
HGB (direct)	3.183	3.865	-0.180	78.15%	100.00%
Random Forest (direct)	3.200	4.083	-0.316	78.15%	98.01%
Extra Trees (direct)	3.998	4.887	-0.885	64.90%	95.36%
HGB via Δ SINR	4.515	5.581	-1.459	60.93%	92.72%

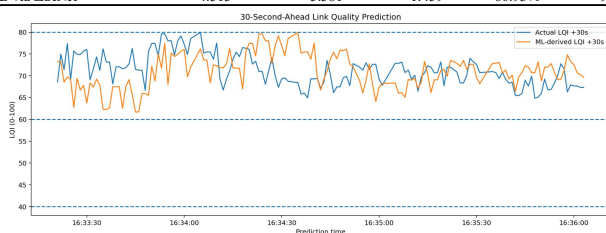


Fig. 7. Future LQI prediction from the causal link-quality pipeline.

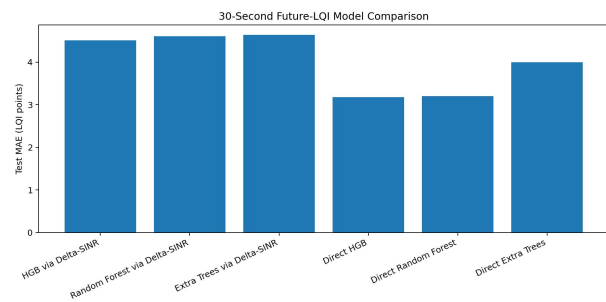


Fig. 8. Model comparison for the final prediction tasks.

F. Quantized LQI and Link-State Representation

The integer LQI interface provides 101 discrete states while retaining four descriptive quality bands for operator interpretation. Fig. 9 illustrates the quantized future-LQI output, and Fig. 10 summarizes agreement between predicted and actual integer LQI. In the final data, the two Ericsson flights have mean LQI values of 67.31 and 68.54, while the Lake Wheeler flight has a mean of 67.90. Most observations fall in the good band; the Lake Wheeler flight has a larger fair-band fraction than the two Ericsson flights.

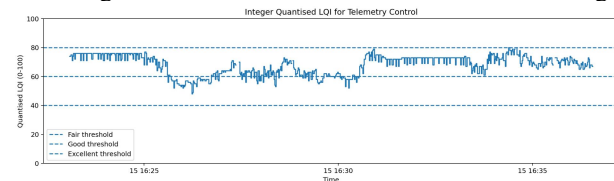


Fig. 9. Integer LQI representation on the 0–100 scale.

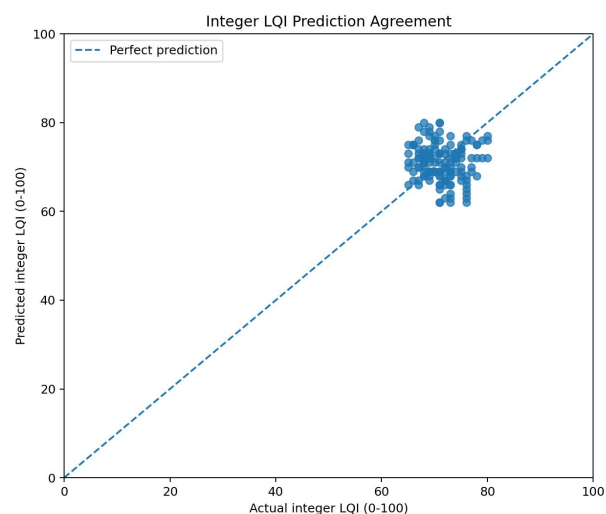


Fig. 10. Agreement of predicted and actual integer LQI states.

G. Tolerance-Based Accuracy

Exact integer agreement is a stringent metric because the LQI contains 101 discrete states. For the direct HGB future-LQI model, 78.15% of test predictions are within 5 LQI points and 100% are within 10 points. The corresponding exact-integer accuracy is 7.28%. Fig. 11 shows the future-SINR tolerance analysis. The descriptive LQI bands remain available as operator labels, but the standalone band-accuracy figure is omitted from the revised paper.

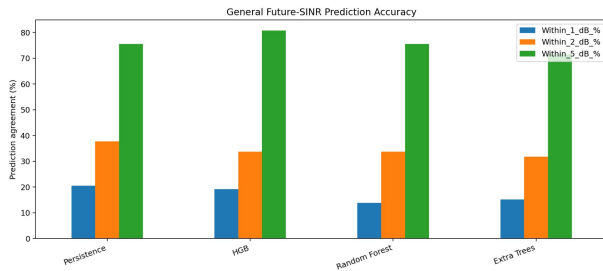


Fig. 11. Tolerance-based accuracy for future SINR.

H. Degradation Event Detection

For the 30-s HGB regression-derived SINR degradation event with threshold $\Delta \text{SINR} \leq -2$ dB, the test-set ROC-AUC is 0.710 and PR-AUC is 0.670. The calibrated detector reaches 77.48% accuracy, 88.24% precision, 31.91% recall, and 46.88% F1, with 98.08% specificity. The relatively low recall indicates that the calibrated detector is conservative and should not be interpreted as a complete failure-warning mechanism.

For the LQI degradation definition $\Delta \text{LQI} \leq -5$ points, the HGB detector has ROC-AUC 0.900 and PR-AUC 0.706, with 86.09% accuracy, 71.43% precision, 37.04% recall, and 48.78% F1. Fig. 12 and Fig. 13 show the calibrated event analyses.

TABLE V. 30-s DEGRADATION-EVENT DETECTION METRICS FOR HGB.

Target	Accuracy	Precision	Recall	F1	ROC-AUC	PR-AUC
SINR degradation	77.48%	88.24%	31.91%	46.88%	0.710	0.670
LQI degradation	86.09%	71.43%	37.04%	48.78%	0.900	0.706

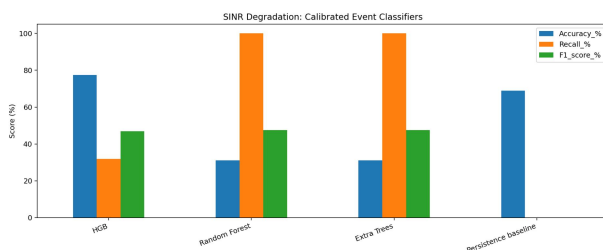


Fig. 12. Calibrated 30-s SINR-degradation detection.

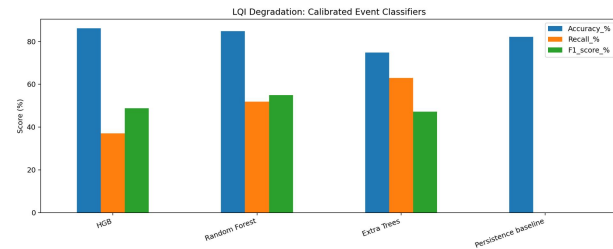


Fig. 13. Calibrated 30-s LQI-degradation detection.

I. Raw Fixed-Threshold LQI Audit

The calibrated degradation metrics above use decision thresholds selected only from walk-forward out-of-fold training predictions. To provide a separate audit of the physical event definition itself, a raw fixed-threshold evaluation was also performed directly on the 30-s future-LQI predictions using the physical threshold $\Delta \text{LQI} \leq -5$ points. No test-set threshold optimization is involved in this audit. Table VI reports the resulting confusion-matrix-derived metrics for all three regressors. This table should therefore be interpreted as a fixed engineering-threshold audit rather than as a replacement for the calibrated detector.

For HGB, the fixed-threshold audit gives 86.09% accuracy, 71.43% precision, 37.04% recall, and 48.78% F1, with 96.77% specificity and an MCC of 0.447. Random Forest gives 84.77% accuracy and a higher recall of 51.85%, while its precision is 58.33%. Extra Trees gives 74.17% accuracy, 59.26% recall, and 45.07% F1. The different precision–recall trade-offs illustrate why the physical threshold and the calibration procedure should be reported separately: a detector can identify more physical degradation events while also producing more false alarms. The raw audit is consequently useful for transparency about the fixed event definition, whereas the calibrated results quantify the behavior of the selected operator-oriented decision rule.

TABLE VI. RAW FIXED-THRESHOLD LQI AUDIT AT 30 s USING $\Delta \text{LQI} \leq -5$ POINTS.

Model	Acc.	Prec.	Rec.	F1	Spec.	Bal. Acc.	MCC	TP/FP/FN/TN
HGB	86.09%	71.43%	37.04%	48.78%	96.77%	66.91%	0.447	10/4/17/120
Random Forest	84.77%	58.33%	51.85%	54.90%	91.94%	71.89%	0.459	14/10/13/114
Extra Trees	74.17%	59.26%	45.07%	52.07%	77.42%	68.34%	0.309	16/28/11/96

J. Uncertainty Quantification

The 90% conformal prediction interval has a mean width of 10.99 dB and achieves 88.08% empirical coverage on the test data. The interval is therefore reasonably informative, but it does not reach the nominal 90% level on this particular test sample. For an operational console, this

uncertainty information is arguably as important as the point forecast: a predicted degradation accompanied by a wide interval should be treated differently from a similar point forecast with a narrow interval.

TABLE VII. UNCERTAINTY QUANTIFICATION FOR THE PRIMARY 30-s FORECAST.

Method	Nominal coverage	Test coverage	Summary
90% conformal interval	90.00%	88.08%	Mean width = 10.99 dB
Moving-block bootstrap	95%	—	Improvement CI = 0.72–27.89%

VII. DISCUSSION

The experimental results indicate three important engineering observations. First, the value of causal temporal context is evident in the feature ablation results. Radio measurements alone are insufficient for the 30-s task, whereas adding causal temporal features substantially reduces HGB error. Second, persistence remains difficult to beat at short horizons, which is expected for slowly varying link states. The main improvement occurs at the selected 30-s operating point, where the HGB model reduces MAE by 15.23%.

Third, future link quality should not be conflated with a single radio metric. The LQI combines multiple available measurements, while the primary predictor is optimized for future SINR change. The difference between the SINR-derived LQI and direct future-LQI prediction demonstrates that the engineering interface and the scientific prediction target answer different questions. For operational deployment, a multi-output or multi-task model could be investigated so that future SINR and future LQI are jointly optimized. Such an extension is consistent with broader AI-enabled UAV communication research, which identifies predictive and adaptive intelligence as an important direction for beyond-5G aerial networking [16].

The cross-flight results also caution against assuming that additional flight records automatically improve performance. The augmented experiment degraded the predeclared HGB result at 30 s, while LOFO performance varied substantially by held-out flight. This suggests that future work should explicitly model domain shift caused by base-station geometry, measurement hardware, flight orientation, trajectory, and network configuration.

The present system should be interpreted as a ground-monitoring and decision-support prototype rather than an autonomous flight-control system. The LQI control profiles are engineering guidance only. Actual modulation, coding, bandwidth, or reliability actions must be mapped to the capabilities and policies of the specific modem and network.

No direct numerical comparison with the metrics of the cited ground-station, cellular-measurement, or link-prediction studies is made because their tasks, datasets, targets, and evaluation protocols are not identical to the present causal 30-s Δ SINR prediction problem. The comparisons in this paper are therefore internal, using the same test protocol and persistence reference for all evaluated models.

VIII. LIMITATIONS AND FUTURE WORK

The main limitation is the size and diversity of the evaluation set. The quantitative experiment contains 2376 observations from three flight records, and the LOFO results show that the held-out flight can materially change the error. Stronger claims will require substantially more flights spanning different base stations, trajectories, altitudes, environments, weather conditions, and radio configurations.

Second, the primary LQI is an engineering index with manually specified normalization ranges and weights. It is useful for an operator interface, but it is not a standardized universal measure of wireless quality. Future work should investigate learned or task-optimized weighting while preserving interpretability and ensuring that missing measurements do not create artificial quality penalties.

Third, the present prediction target is scalar future Δ SINR. Future research can extend the framework to probabilistic forecasting, multi-target prediction of SINR/RSRP/CQI/throughput, domain adaptation across testbeds, and uncertainty-aware alerting. The conformal interval already provides a starting point for uncertainty-aware visualization.

Finally, the telemetry replay and the cellular prediction pipeline are currently coupled at the information level rather than through a deployed live radio interface. The next engineering step is therefore not simply another model: it is the integration of live telemetry ingestion, model-health checks, uncertainty-aware alerts, and latency/resource monitoring on the ground computer.

IX. CONCLUSION

This paper presents a ground-monitoring workflow that connects aircraft telemetry replay, causal temporal processing, future cellular-link prediction, an engineering LQI, and operator-facing monitoring outputs. The main quantitative experiment uses two public AERPAW measurement sources and three independent physical flight traces. At the selected 30-s operating point, HGB achieves 3.110 dB MAE versus 3.669 dB for persistence. The separate NASA Tail-654 console adds map-based replay, primary aircraft telemetry, phase-aware statistical monitoring, eight-second persistence, and recovery hysteresis without mixing aircraft-state anomaly monitoring with wireless-link prediction metrics. The resulting system is a practical ground-monitoring prototype with clear separation between the predictive communication experiment and the aircraft telemetry observability branch.

X. ACKNOWLEDGMENT

The author acknowledges the guidance and academic support associated with the project work at Sri Venkateswara University College of Engineering. The public AERPAW and NASA datasets used in the study are acknowledged as external data sources.

XI. REFERENCES

- [1] H. S. Júnior, M. A. Santos, T. F. Lúcio, and G. V. Raffo, "Developing an Integrated Ground Station for UAV Operations: Telemetry and Mission Planning," in Proc. 2025 Brazilian Conference on Robotics (CROS), 2025, pp. 1–6, doi: 10.1109/CROS66186.2025.11066147.
- [2] C. Liu, B. Shangguan, P. Yue, and M. Zhang, "Towards an Integrated Service Platform in Remote Sensing Satellite Ground Station for Telemetry, Operational Control, and Data Application," in Proc. 10th Int. Conf. Agro-Geoinformatics, 2022, pp. 1–6, doi: 10.1109/Agro-Geoinformatics 55649.2022.9859217.
- [3] G. Cerar, H. Yetgin, M. Mohorčič, and C. Fortuna, "Machine Learning for Wireless Link Quality Estimation: A Survey," IEEE Communications Surveys & Tutorials, vol. 23, no. 2, pp. 696–728, 2021, doi:10.1109/COMST.2021.3053615.
- [4] I. J. Kim and S. Lee, "Development of Telemetry Data Processing Program," in Proc. 2013 Int. Conf. ICT Convergence (ICTC), 2013, pp. 468–473, doi: 10.1109/ICTC.2013.6675398.
- [5] L. Segal, "Development of a Command and Telemetry Processing Interface to be Independent of Ground Control Systems," in Proc. AUTOTESTCON, 2011, doi: 10.1109/AUTEST.2011.6058770.
- [6] M. A. Alqodah and M. M. Matalgah, "Ground-Based Aircraft Recorders: Requirements, Challenges, Implementations, and Future Perspectives," IEEE Communications Magazine, vol. 63, no. 3, pp. 62–68, 2025, doi: 10.1109/MCOM.001.2300779.
- [7] A. H. Fahim Raouf, D. Lee, M. Rahman, et al., "Curated wireless datasets for aerial network research," npj Wireless Technology, vol. 2, Art. no. 41, 2026, doi: 10.1038/s44459-026-00050-1.
- [8] S. Singh, A. Gürses, O. Ozdemir, R. Asokan, M. L. Sichitiu, I. Guvenc, R. Dutta, and M. Mushi, "Modeling and Analysis of Air-to-Ground Cellular KPIs in a 5G Testbed using Android Smartphones," in Proc. IEEE Aerospace Conference, 2026, doi: 10.1109/AERO66936.2026.11520074.
- [9] A. Festag, S. Udupa, L. Garcia, R. Wellens, M. Hecht, and P. Ulfing, "End-to-End Performance Measurements of Drone Communications in 5G Cellular Networks," in Proc. 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall), 2021, doi:10.1109/VTC2021-Fall 52928.2021.9625429.
- [10] W. Khawaja, I. Guvenc, D. W. Matolak, U.-C. Fiebig, and N. Schneckenberger, "A Survey of Air-to-Ground Propagation Channel Modeling for Unmanned Aerial Vehicles," IEEE Communications Surveys & Tutorials, vol. 21, no. 3, pp. 2361–2391, 2019, doi: 10.1109/COMST.2019.2915069.
- [11] D. Mishra and E. Natalizio, "A survey on cellular-connected UAVs: Design challenges, enabling 5G/B5G innovations, and experimental advancements," Computer Networks, vol. 182, Art. no. 107451, 2020, doi: 10.1016/j.comnet.2020.107451.
- [12] A. Alzenad, A. El-Keyi, F. Lagum, and H. Yanikomeroglu, "UAV-Enabled Ultra-Reliable Low-Latency Communications for 6G: A Comprehensive Survey," IEEE Access, vol. 9, pp. 137338–137352, 2021, doi: 10.1109/ACCESS.2021.3117902.
- [13] Z. Li, M. Chen, G. Li, X. Lin, and Y. Liu, "Map-Driven mmWave Link Quality Prediction With Spatial-Temporal Mobility Awareness," IEEE Transactions on Mobile Computing, vol. 23, no. 12, pp. 12161–12175, 2024, doi: 10.1109/TMC.2024.3406728.
- [14] T. Nishio, K. Yamamoto, M. Morikura, and S. Itahara, "Online Trainable Wireless Link Quality Prediction System using Camera Imagery," in Proc. IEEE Globecom Workshops, 2020, doi: 10.1109/GCWkshps 50303.2020.9367396.
- [15] M. L. Bote-Lorenzo, E. Gómez-Sánchez, C. Mediavilla-Pastor, and J. I. Asensio-Pérez, "Online machine learning algorithms to predict link quality in community wireless mesh networks," Computer Networks, vol. 132, pp. 68–80, 2018, doi: 10.1016/j.comnet.2018.01.005.
- [16] A. O. Hashesh, S. Hashima, R. M. Zaki, M. M. Fouda, K. Hatano, and A. S. T. Eldien, "AI-Enabled UAV Communications: Challenges and Future Directions," IEEE Access, vol. 10, pp. 92048–92066, 2022, doi: 10.1109/ACCESS.2022.3202956.
- [17] C. Xu and Y. Xie, "Conformal Prediction for Time Series," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 10, pp. 11575–11587, 2023, doi: 10.1109/TPAMI.2023.3272339.
- [18] I. Panitsas and L. Tassiulas, "A Family of Open Time-Series Foundation Models for the Radio Access Network," arXiv preprint arXiv:2604.04271, 2026.