Online Resource 5 Exploratory development and source diagnosis

Label Budget Trade-offs and Calibration Transfer in GNSS Spoofing Detection

Annals of Telecommunications
Bo Xie and Yi Zhao
Wuhan Digital Engineering Institute, Wuhan, China
Corresponding author: Bo Xie (xiebo@mail.hfut.edu.cn)

This supplement describes post-confirmation development, not unseen confirmation. All four cohorts were performance-exposed before the first development round. Four-source leave-one-out folds use 40 class-labeled support rows (20/class total across three sources), 160 disjoint normal calibration rows, and 4,000 validation-source rows. Both seeds (0,1), both nominal levels (5%,1%) and all attempts are retained. No target-source adaptation is performed. The exposed research pool contains 16,000 labeled rows.

Chronology and settings

Round 1: 80 models across ten configurations. Round 2: 72 models across nine additional configurations. Round 3: 32 models across four configurations, including 16 exact second-round controls. Thus 184 model instances do not represent 184 distinct designs or independent acquisitions. Each round was fixed before its own scores; the sequence was adaptive.

The original metric uses the primary BiLSTM settings, with 40-step and 200-step development controls. Aligned candidates use a 64-hidden/16-embedding residual encoder, normalized embeddings, leave-one-out class prototype CE, and either 0.1 supervised contrastive loss or its CE-only ablation; AdamW lr=0.001, weight_decay=0.001, 200 full-support steps. Raw, unit-normalized, complex Gram, and Gram-plus-log-energy representations are separate configurations.

Round-2 neighbor candidates use class log-mean-exp cosine similarities at temperature 0.2, equal class priors and self-excluded training, with the same encoder and AdamW settings in float64. Multiple prototypes are class-by-source means; a single-prototype CE control and fixed non-learning RBF controls are included. Componentwise asinh(x/s) uses the support-only median absolute component as s (zero median replaced by 1); it preserves signed magnitude. StandardScaler uses support only. Round 3 changes only the training retrieval mask to exclude the whole query source. All support rows still fit preprocessing and train the model; inference requires no source ID.

TabM uses official tabm 0.0.3: two blocks, width 128, 16 ensemble branches, dropout 0.1, AdamW lr=0.002 and weight_decay=0.0003, 200 steps. Prediction averages member probabilities. This is a budget adaptation, not the paper-default network or tuned benchmark recipe. ModernNCA uses unmodified TALENT core at commit 1b973adffa4203c3f62c4949957b1ba3efbe60d3: dim128, dropout0.1, d_block512, n_blocks0, temperature1, sample_rate0.5; PLR77 frequencies, frequency_scale0.04431360576139521, embedding34, lite=True. Full-support training has no external candidates and masks the internal self diagonal; float64 avoids distance cancellation. AdamW lr=0.01, weight_decay=0.0002, 200 steps are declared harness settings. Neither method uses validation early stopping. XGBoost retains depth3, eta0.05 and 200 rounds from the separate extension.

Selection ranks only the declared new candidates by worst-source mean-seed balanced error at nominal5%, then mean error, then name. Round2/3 engineering screening requires every source mean FPR<=10%, FNR<=50%, BE<=40%; this is a descriptive development criterion, not population risk control. No candidate passes. Round3 raw source masking ranks ahead of asinh by the declared worst-source criterion despite worse

mean error; the criterion is not changed after observing results.

Complete method means

Each number averages two seeds within source, then four sources equally. FPR/FNR/BE are percentages. Round3 control rows are retained explicitly.

Table OR5.1 is provided as Table_OR5_1.csv: 46 data rows. Column headings and displayed precision are preserved from the archived supplement. The table is not filtered.

Source-level results

Table OR5.2 is provided as Table_OR5_2.csv: 184 data rows. Column headings and displayed precision are preserved from the archived supplement. The table is not filtered.

Post-hoc component diagnosis

Ten source MAT hashes and selected-row values were rechecked. All 96,000 imported components match their source values exactly; Genuine/Counterfeit condition labels match source directories. PRN-count quotas match across conditions within each of the four cohorts. This is numerical provenance verification, not proof of common physical units, integration interval, phase convention or independent acquisition. No observations are filtered.

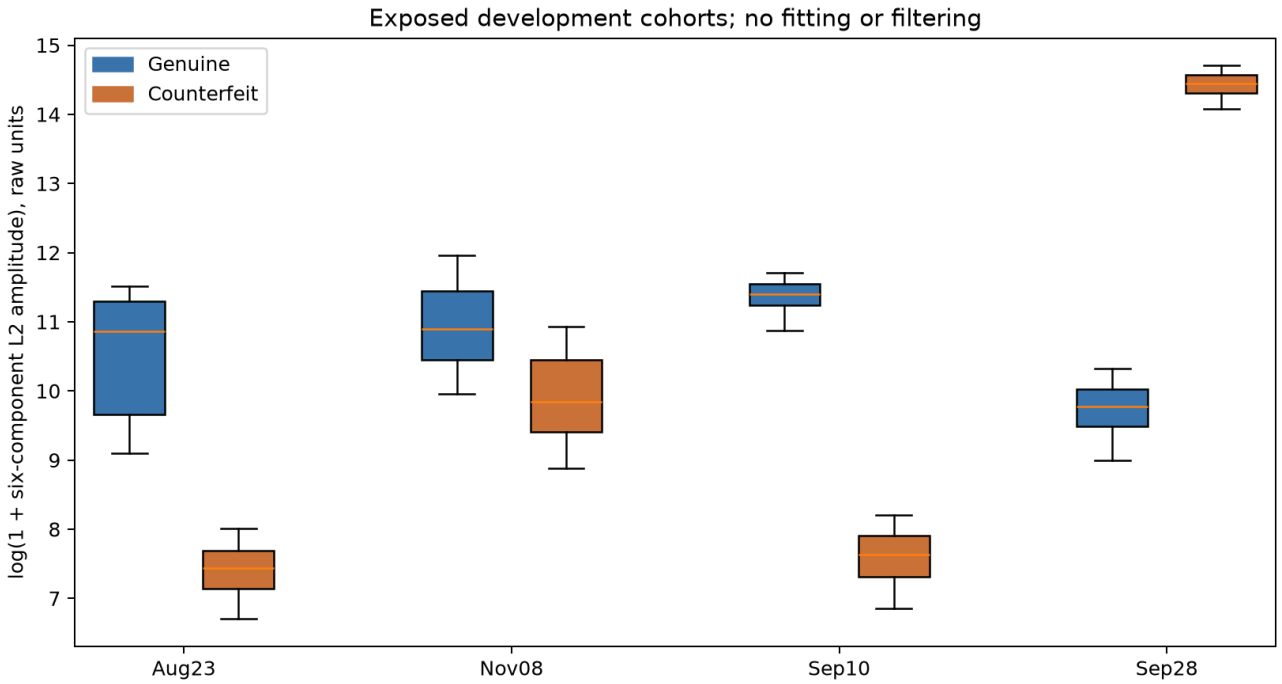


Figure S4. Six-component amplitude in exposed cohorts. Natural log(1+L2 amplitude), unverified raw units. Boxes show quartiles/median; whiskers mark the 5th/95th percentiles and outliers are omitted from display only, not from statistics. Each condition/cohort has 2,000 selected rows. This is not physical RF power or a fitted detector.

With increasing amplitude fixed as attack-oriented, per-source univariate AUROCs are Aug23 0.000034, Nov08 0.139750, Sep10 0.000000 and Sep28 0.999500. The association reverses across these conditions. It does not identify which variable drives each network or isolate a physical cause; location, generation mechanism, PRN and processing context co-vary. Direction is not flipped to improve detection, and these labels do not recalibrate deployed thresholds.

Verification and amendments

Round1/2 validation reconstructs identities, fitted preprocessing/reference states, saved-model scores and all metric rows; it does not independently retrain all models. Round3 independently retrains all32 models from support and matches all16 self-only controls exactly to round2. A first round3 validator used a different mathematically equivalent reduction layout; rounding divergence during optimization caused rejection. The retained r1 amendment matches reduction order without changing tolerances. All32 network states, losses and scores are identical before/after that validator-only amendment; the failed attempt is retained. Rehashed semantic corruption probes are archived separately. No reserved source is scored by any development round.

Machine-readable tables accompany the settings documented above.