Dresden University of Technology

Paul Schwarzbach, Muhammad Ammad
Chair of Transport Systems Information Technology

HYMN

From Least Squares to Deep Learning

Benchmarking Indoor Positioning on the HYMN Multi-Technology Dataset

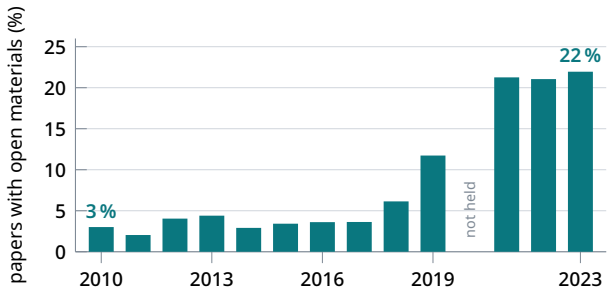
IPIN 2026 · Session 10: Datasets, Benchmarks & Standards · 7 October 2026 | Rome

TUD | Faculty of Transport and Traffic Sciences "Friedrich List"

iTVS Informationstechnik für Verkehrssysteme

Open research data

Most indoor positioning results rest on data that nobody else can access



IPIN papers per year that use or share open data or materials:
4 of 133 in 2010, 27 of 123 in 2023 [1, Fig. 1].

> 3000

indoor positioning papers per year in Scopus [1]

23

open datasets in the best year (2022)

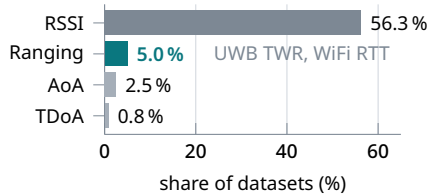
"Our community cannot disregard the danger of facing a reproducibility crisis as most publications do not share the data [...]." [1]

[1] G. G. Anagnostopoulos, P. Barsocchi, A. Crivello, C. Pendão, I. Silva, J. Torres-Sospedra, "ORDIP: Principle, practice and guidelines for open research data in indoor positioning," *Internet of Things*, vol. 30, 101485, 2025.

What open datasets lack

Few ranging measurements and few defined splits make learned methods hard to compare

Radio measurements in 119 datasets [1]



WiFi: 4 of 46 datasets carry round-trip ranges.
BLE: none of 24 has ranging, phase-based or other (own count in the ORDIP data [1]).

Metadata needed to compare results [1]

6.7 % of datasets define a train / validation / test split

50 % leave the size of the collection area undefined

21 % leave the ground-truth protocol undefined

Without a predefined split, each study draws its own partition and error figures stop being comparable. Fair comparison of learned methods needs defined splits, tasks and metrics [2].

[1] G. G. Anagnostopoulos, P. Barsocchi, A. Crivello, C. Pendão, I. Silva, J. Torres-Sospedra, "ORDIP: Principle, practice and guidelines for open research data in indoor positioning," *Internet of Things*, vol. 30, 101485, 2025.
[2] G. Pan et al., "AI-driven wireless positioning: Fundamentals, standards, state-of-the-art, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 28, 2026.

Research gaps and contributions

One multi-technology dataset, four method families, two evaluation protocols, all shared

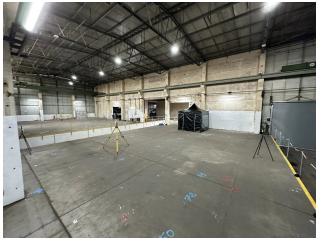
Gaps	Contributions
G1 Studies evaluate one radio technology or one method family [1]. **G2** Classical and learned methods are seldom compared on the same data.	**C1** ILS, RLS, BGF and a ResNet on fused BLE + UWB + WiFi ranges from HYMN, under one measurement model.
G3 Random train / test splits hide how learned models behave at unseen locations.	**C2** RandomSplit vs. SpatialHoldout protocol quantifies the spatial-generalisation gap.
G4 Data and code are rarely released together [2].	**C3** End-to-end reproducible benchmark: data, pipeline and training recipes with DOIs.

[1] G. M. Mendoza-Silva, J. Torres-Sospedra, J. Huerta, "A meta-review of indoor positioning systems," *Sensors*, vol. 19, no. 20, 4507, 2019.
[2] G. G. Anagnostopoulos, P. Barsocchi, A. Crivello, C. Pendão, I. Silva, J. Torres-Sospedra, "ORDIP: Principle, practice and guidelines for open research data in indoor positioning," *Internet of Things*, vol. 30, 101485, 2025.

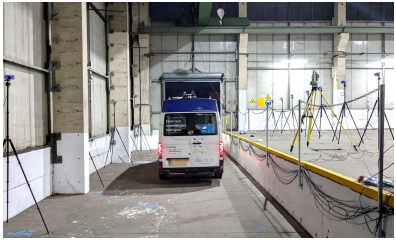
Preliminary work

HYMN

An industrial hall with open gates mixes indoor, outdoor and transition-zone propagation



Hall floor and lowered driveway (2025)



Setup along the driveway [2, Fig. 3]



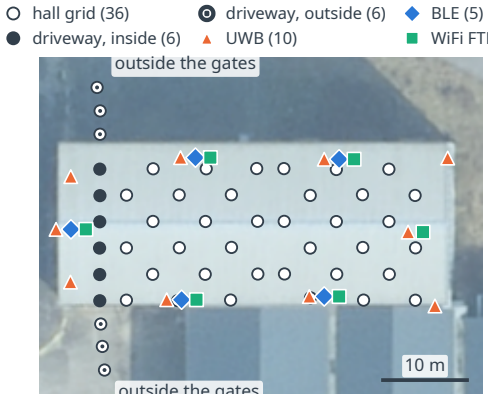
Open gate: the transition zone

- Industrial hall, 44 m × 18 m
- Gates on both sides open throughout
- Recorded on 23 October 2024
- Open data, CC BY 4.0 [1]

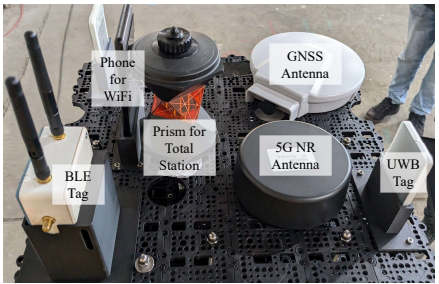
[1] M. Ammad et al., “Descriptor: A hybrid indoor and indoor–outdoor positioning multitechnology dataset (HYMN),” *IEEE Data Descriptions*, vol. 3, pp. 379–387, 2026. Dataset: doi 10.5281/zenodo.17979434.
[2] A. Michler, P. Schwarzbach, J. Ninnemann, M. Ammad, O. Michler, “An empirical assessment of indoor-outdoor localization based on signals of opportunity from multiple systems,” in *Proc. IEEE/ION PLANS*, 2025, pp. 948–959, doi 10.1109/PLANS61210.2025.11028349.

Preliminary work

48 static reference points, co-located anchors, one measurement plate



Local frame of the dataset, co-located anchors drawn side by side.
Orthophoto: GeoSN, dl-de/by-2-0.

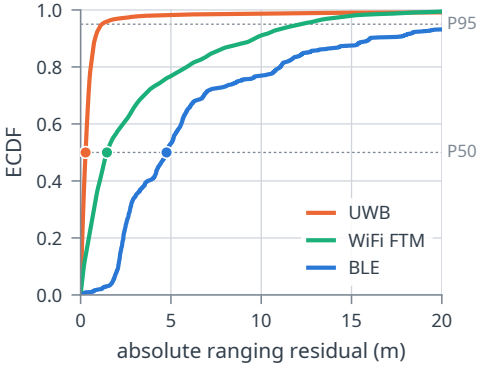


Measurement plate, 3 min static recording per point. A Leica TS16 total station tracks the prism as ground truth.

5G NR and GNSS were recorded too but are not fusion inputs here.

Ranging quality

Three ranging technologies, with different ranging quality characteristics



	Ranging principle	Median (m)	P95 (m)
▲ UWB (10)	two-way ranging ZigPos	0.27	1.18
■ WiFi (6)	round-trip time (FTM) Aruba 630, Pixel 6a	1.46	12.25
♦ BLE (5)	phase-based ranging Metirionic DMK-215	4.76	26.77

Anchor count in brackets. Absolute range residual against the total-station geometry, 50–63 k ranges per technology.

Equal anchor weights let BLE and WiFi outliers dominate UWB.

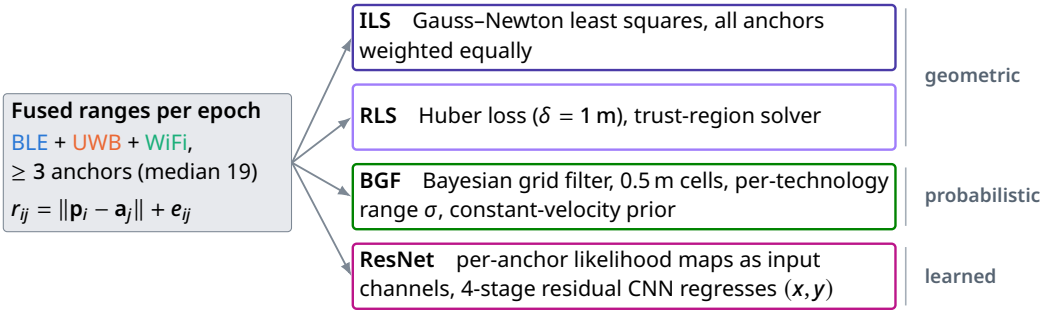
More detailed ranging analysis in [1].

[1] A. Michler, P. Schwarzbach, J. Ninnemann, M. Ammad, O. Michler, "An empirical assessment of indoor-outdoor localization based on signals of opportunity from multiple systems," in *Proc. IEEE/ION PLANS*, 2025, pp. 948–959, doi 10.1109/PLANS61210.2025.11028349.

Positioning methods

HYMN

Four method families share one range-based measurement model



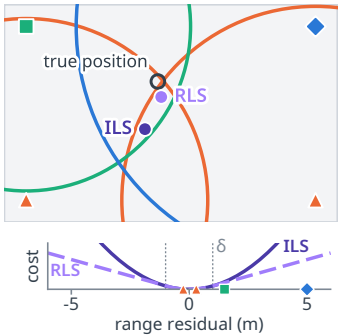
Output: 2-D horizontal error against the total-station ground truth.
ResNet architecture and hyperparameters come from prior work [1, 2] and are not tuned on HYMN.

[1] P. Schwarzbach et al., "Deep learning indoor positioning for connected aircraft cabins: A ResNet approach with real-world validation," *Sensors*, vol. 26, no. 5, 1569, 2026.
[2] M. Ammad et al., "Range-angle likelihood maps for indoor positioning using deep neural networks," arXiv:2508.12746, 2025.

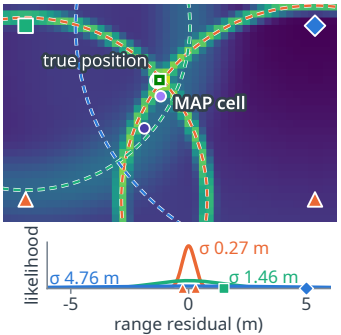
From ranges to a position

Four ranges of one schematic epoch, turned into a position by each method family

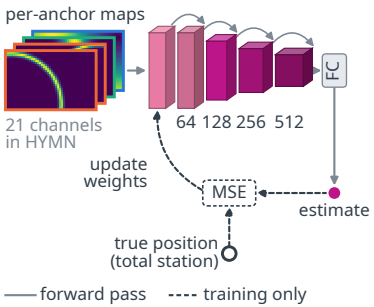
ILS and RLS — geometric



BGF — probabilistic



ResNet — learned

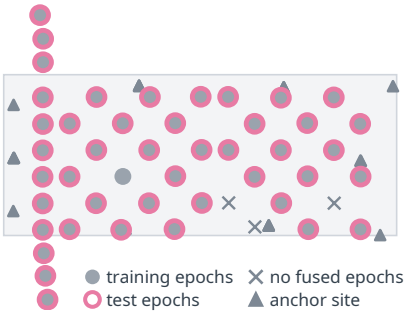


Evaluation protocols

Two ways to split the same epochs answer two different questions

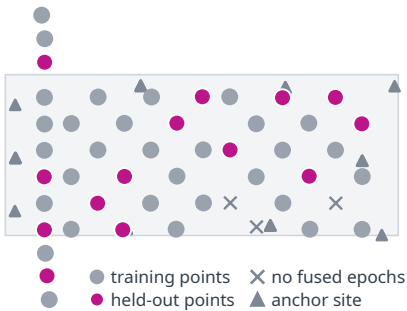
RandomSplit measures interpolation

70% / 30% of all epochs, same points in train and test



SpatialHoldout measures generalisation

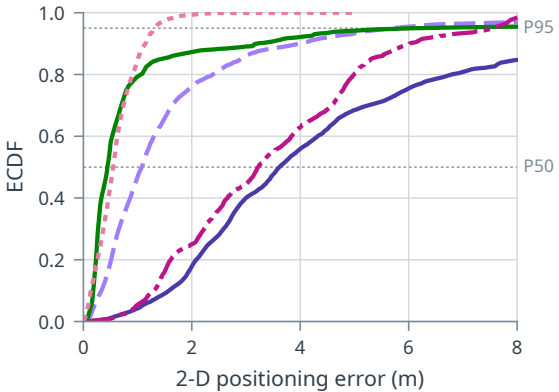
31 training points, 14 held-out points



Same architecture, hyperparameters and seed in both protocols.

Positioning accuracy

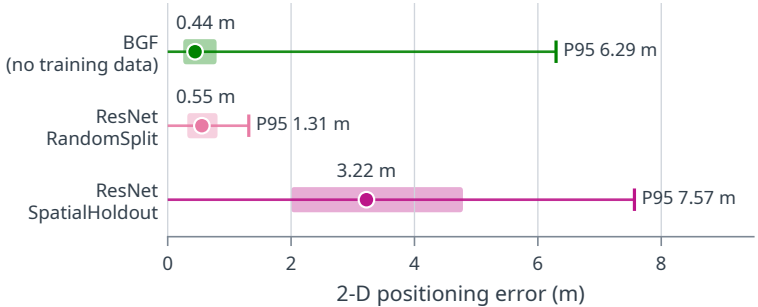
The grid filter matches the ResNet at the median, the ResNet wins the tail



Method	Median (m)	P95 (m)
BGF	**0.44**	6.29
ResNet RandomSplit	0.55	**1.31**
RLS	1.08	5.68
ResNet SpatialHoldout	3.22	7.57
ILS	3.62	18.66

Spatial generalisation

At held-out locations the ResNet error grows almost six-fold



Box: P25–P75. Dot: median. Whisker end: P95.

- Median 0.55 m to 3.22 m, P95 1.31 m to 7.57 m.
- RandomSplit measures interpolation between surveyed points
- BGF needs no training data and keeps its 0.44 m median

Conclusions and outlook

Report spatial-holdout results, and ship split definitions with the data

Take-aways

1. Ranging quality differs by technology. Robust or probabilistic weighting recovers most of the loss.
2. A Bayesian grid filter matches deep learning at the median, without training data.
3. Random splits measure interpolation. Report learned positioning on held-out locations too.
4. Datasets that ship split definitions make learned methods comparable.

Limitations and next steps

1. **Static points only**
   → moving receivers
2. One site
   → cross-dataset evaluation
3. BLE, UWB and WiFi only
   → add 5G NR and GNSS
4. Spatial-generalisation gap
   → train for unseen locations

Open data and open methods

HYMN documents what ORDIP often finds undefined, and every method is released

HYMN Dataset [1]

- ✓ Zenodo DOI, versioned releases, CC BY 4.0
- ✓ Peer-reviewed data descriptor
- ✓ Ground-truth protocol with total station
- ✓ Coordinate frames: local, UTM 33N, ECEF
- ✓ Environment: area, gates, photos, layout
- ✓ Raw logs, preprocessing code, data dictionary

This paper and its code release [2]

- ✓ ILS, RLS, BGF and ResNet implementations
- ✓ Evaluation pipeline, pinned packages, seed
- ✓ Code-to-paper map for figures and tables
- ✓ HYMN data bundled unchanged and cited
- ✓ Zenodo DOI and GitHub, MIT licence

The paper also opted into the IPIN 2026 Reproducibility Review.

[1] M. Ammad et al., “Descriptor: A hybrid indoor and indoor–outdoor positioning multitechnology dataset (HYMN),” *IEEE Data Descriptions*, vol. 3, pp. 379–387, 2026. Dataset: doi 10.5281/zenodo.17979434.
[2] P. Schwarzbach, M. Ammad, “HYMN multi-technology indoor positioning evaluation,” companion code, Zenodo, 2026, doi 10.5281/zenodo.20058106.

Dresden University of Technology

HYMN

Thank you.

Project page

HYMN Dataset

Paper Code

Contact

Paul Schwarzbach · TUD Dresden University of Technology

TUD | Faculty of Transport and Traffic Sciences "Friedrich List"

iTVS Informationstechnik für Verkehrssysteme

Backup: full results

HYMN

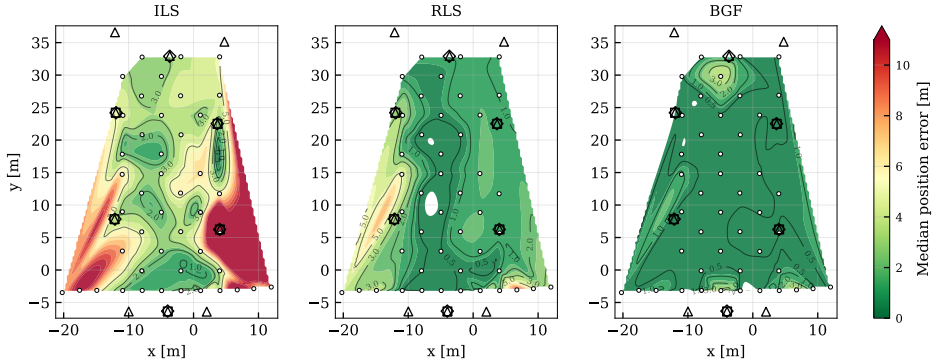
Fused positioning error per method, all statistics

Method	*n*	Mean (m)	Median (m)	Std (m)	P75 (m)	P95 (m)
ILS	6075	5.519	3.616	5.974	5.951	18.659
RLS	6075	2.013	1.075	3.744	1.930	5.684
BGF	6075	1.586	**0.444**	4.291	**0.793**	6.295
ResNet SpatialHoldout	1833	3.504	3.222	1.915	4.786	7.565
ResNet RandomSplit	1823	**0.605**	0.554	**0.403**	0.808	**1.314**

- Classical methods run on every fused epoch, the ResNet rows cover the test partitions only. The comparison is indicative, not strictly paired.
- RandomSplit: 70 / 30 random partition of all epochs (4256 train, 1823 test).
- SpatialHoldout: 31 training points, 14 held-out points (1833 test epochs).
- Trained checkpoints are not redistributed. They regenerate from the released inputs and the pinned seed.

Backup: spatial error pattern

ILS errors grow towards the hall perimeter, where anchor geometry degrades



Interpolated per-point median error of the classical methods, shared colour scale, anchors marked (paper Fig. 4).

Backup: method parameters

HYMN

BGF range widths follow the median ranging residual per technology

ILS	Gauss–Newton, anchor-centroid start, $\leq$ **7** steps, tolerance $\mathbf{10^{-6}}$
RLS	Huber $\delta = \mathbf{1.0\,m}$, trust-region reflective solver, $\leq \mathbf{50}N_i$ function evaluations
BGF	0.5 m grid, 2 m padding around the anchor hull, range σ of 0.27, 1.46 and 4.76 m (UWB, WiFi, BLE), constant-velocity model with mean speed 0.1 m/s and speed σ 0.2 m/s, estimate: weighted mean within 2 m of the MAP cell
ResNet	per-anchor likelihood maps as channels, 4 residual stages with 64/128/256/512 filters, batch norm, ReLU, dropout 0.2, MSE loss, Adam ($\mathbf{10^{-3}}$), batch 32, 50 epochs, seed 42

All values as in `config.py` of the companion code. ResNet architecture and hyperparameters reused from prior work, not tuned on HYMN.