remote sensing

MDPI

Article

Sparse Height Sampling of Radar Wind-Profile Products: Criterion-Dependent Sufficiency and Temporal Stability

Yifang Li [1], Yiyun Guo [2] and Dongning Jia [1,*]

[1] Faculty of Information Science and Engineering, Ocean University of China, 238 Songling Road, Laoshan District, Qingdao 266100, China
[2] Qingdao Port International Co., Ltd., No. 7 Gangji Road, Shibei District, Qingdao 266011, China
* Correspondence: jiadn@ouc.edu.cn

Highlights

What are the main findings?

- Retained-level sufficiency was criterion-dependent: proper orthogonal decomposition with QR-based height selection (POD–QR) required 14, 10, 5 and 3 of 15 levels at all-six-fold root-mean-square error (RMSE) tolerances of 0.30, 0.40, 0.50 and 0.70 m s^{-1}, respectively.
- Across all folds, the selected POD–QR layouts ranked within the best 5.94% of endpoint-matched five-level layouts and the best 4.74% of seven-level layouts, with early-fold layouts remaining effective later in the same deployment.

What are the implications of the main findings?

- Dense reporting grids should not be equated with equally independent vertical information; useful sampling density should be tied to an explicit reconstruction objective.
- Historical profile structure can prioritise height channels for product and reporting design and inform prospective reduced-height sampling strategies.

Abstract

Wind-profile products are often reported on dense vertical grids, while neighbouring levels can share substantial covariance. This study quantifies the number and placement of reported heights needed to reconstruct a 15-level, 10 min mean horizontal wind-speed (WS_MEAN) radar product spanning 20–300 m under explicit root-mean-square error (RMSE) criteria. Historical profiles were used for proper orthogonal decomposition with QR-based height selection (POD–QR), and performance was evaluated in six forward-chaining windows against endpoint-matched layouts and a training-only exhaustive subset analysis. The all-six-fold POD–QR minimum was 14, 10, 5, and 3 retained levels at RMSE tolerances of 0.30, 0.40, 0.50 and 0.70 m s^{-1}, respectively. At 0.50 m s^{-1}, five levels achieved a mean RMSE of 0.395 m s^{-1} and ranked within the best 5.94% of endpoint-matched layouts in every fold. Seven levels reduced mean RMSE to 0.328 m s^{-1}, with 91.3% of profiles below 0.50 m s^{-1}. Early-fold layouts retained comparable performance in later periods, while profile curvature, level availability, and observation noise systematically affected recoverability. These findings support accuracy-constrained wind-profile product design and height-channel prioritisation, with potential applications in reduced-height sampling, missing-level reconstruction, data quality checks, and reduced transmission and storage requirements.

Check for updates

Academic Editor: Ismail Gultepe

Received: 2 September 2026
Revised: 6 November 2026
Accepted: 14 November 2026
Published: 17 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

Keywords: wind profile radar; vertical sampling; sampling sufficiency; sparse height selection; radar-product reconstruction; proper orthogonal decomposition

1. Introduction

Wind-profiler radar and Doppler lidar are widely used to describe low-altitude winds for boundary-layer research and operational monitoring. Early profiler studies quantified measurement precision and differences from rawinsonde observations [1,2], while lidar evaluations and marine boundary-layer applications demonstrated the value of vertical wind profiling [3,4]. Pulsed Doppler lidar has since become a common tool across a broad range of atmospheric applications [5]. Many wind-profile products are reported on dense, fixed height grids. In practice, neighbouring levels often vary smoothly and share substantial covariance because the atmosphere is vertically coherent and each remote-sensing retrieval represents a finite sampling volume. Retrieval theory treats reporting grid, effective vertical resolution and independent information as related quantities through concepts such as averaging kernels and degrees of freedom for signal [6,7]. Wind-profile measurements show the same practical feature: effective resolution can extend across several reported gates or levels [8,9]. This structure makes vertical sampling density a useful design variable for products that need to preserve profile shape within a defined reconstruction accuracy.

Observation design has been studied from both the measurement and reconstruction perspectives. Beam geometry, scan strategy and volume averaging shape the wind field represented by a retrieval [10–12]. Application-oriented lidar studies have examined scan-pattern optimisation and practical measurement requirements for turbine control [13], while wind-profiling studies have evaluated scanning and averaging effects on retrieval error and sampling requirements [14,15]. AtmoProKIT v1.1 supports retrievals from heterogeneous lidar systems and scan patterns [16]. Wind-field reconstruction uses dynamic state estimation [17], and learning models optimise sensor placement for reconstruction [18]. Reduced-order methods are especially suited to this setting. Proper orthogonal decomposition (POD) represents dominant spatial structure in a low-dimensional basis [19], and Gappy-POD estimates the corresponding modal coefficients from incomplete observations [20,21]. The discrete empirical interpolation method (DEIM) and pivoted-QR sparse-sensing approaches use low-dimensional structure to identify informative measurement locations [22,23]. Sensor selection has also been formulated through convex optimisation [24]. Wind applications include POD reconstruction of sparse scanning-lidar turbine inflow [25] and recent POD–QR sensor placement for urban wind-field reconstruction [26]. Recent rooftop-flow experiments have also assessed sparse-sensor reconstruction using POD–QR placement and compared interpolation with deep learning methods [27]. For a radar product on a fixed vertical grid, historical profile structure can guide the choice of retained heights, while an explicit error criterion sets the sampling density needed for the reconstruction objective.

The number of retained levels depends on both the reconstruction tolerance and the placement of the retained heights. Because layouts are derived from historical profiles, their performance also needs to be checked on later observations. The complete design space provides a direct benchmark for height placement, while chronological test periods show whether a selected geometry remains effective as the record advances.

In this study, we analyse a 15-level, 10 min mean horizontal wind-speed (WS_MEAN) radar product spanning 20–300 m. Historical profiles are used for POD–QR height selection, and six forward-chaining evaluation windows reserve later observations for reconstruction assessment. POD–QR layouts are compared with uniform and supervised greedy designs

and with the complete endpoint-matched layout space. A training-only exhaustive subset analysis provides an independent estimate of the retained-level transition. Sensitivity analyses cover profile complexity, interpolation choice, selected-level availability and observation perturbations. The objective is to quantify how reconstruction tolerance, retained-level count and height placement interact, and to provide a basis for criterion-based vertical product design, height-channel prioritisation and prospective reduced-height sampling strategies.

2. Materials and Methods

2.1. Radar Observations and Chronological Evaluation

The observations were obtained with a VORTRAD-L300 multi-height millimetre-wave wind-profiling radar (Suzhou Dufeng Technology Co., Ltd., Suzhou, China) installed at 36.103449°N, 120.328013°E at an unobstructed coastal industrial-port site in Qingdao, China. The radar base was approximately 0 m above mean sea level. The manufacturer specifies a 10–300 m measurement range, selectable 15- or 30-level configurations, selectable 10 or 20 m range resolution, and a four-beam geometry with each beam 15° from the centreline. The configuration analysed here contains 15 reported levels from 20 to 300 m at 20 m spacing, referenced to the radar installation level. Timestamps are in Beijing Standard Time and label the end of each 10 min product interval.

The primary variable was WS_MEAN, the 10 min mean horizontal wind-speed product. The dataset comprised 10,355 complete 15-level profiles collected from 11 May to 26 August 2026, spanning late spring and summer and capturing a range of wind conditions (Table 1). During the initial deployment, intermittent data-transmission instability contributed to gaps in the archived profiles, primarily before 18 June. Nevertheless, the available complete profiles from this period were retained to preserve temporal coverage and wind-condition diversity. A sensitivity analysis excluding the early observations showed comparable reconstruction performance for five- and seven-level layouts, without changing the minimum retained-level count at the 0.50 m s^{-1} RMSE criterion. Data availability and the sensitivity results are detailed in Supplementary Material S6 and S8, respectively.

Table 1. Radar product and key evaluation settings.

Item	Configuration
Instrument and site	VORTRAD-L300 wind-profiling radar; ground-mounted coastal industrial-port site, Qingdao, China
Primary product	WS_MEAN, 10 min mean horizontal wind speed
Vertical configuration	15 reported levels, 20–300 m, 20 m spacing
Analysis record	10,355 complete profiles, 11 May–26 August 2026
Evaluation	Six forward-chaining five-day test folds
Retained-level counts	2–14 of 15
Reconstruction	Linear vertical interpolation

Evaluation strategies for temporally structured data should account for temporal dependence [28,29]. We therefore used a forward-chaining retrospective design in which each test period occurred later than the observations used for dimensionality analysis and height selection. The same fold geometry was applied throughout: each test block covered five consecutive days (720 ten-minute profiles), spanning several diurnal cycles while providing a compact, repeatable evaluation unit. A 72 h 10 min separation, equivalent to three full days plus one product interval, was placed between each fit cutoff and test start to reduce dependence on immediately adjacent observations. For each fold, all POD modes and endpoint-constrained selections based on proper orthogonal decomposition with QR

factorization (POD–QR) were derived from profiles at or before the fold-specific training cutoff; supervised greedy selections were likewise training-only, and the later test window did not contribute to selection.

Six five-day test windows were evaluated chronologically. Training cutoffs advanced through July and August 2026, and each test window began 72 h 10 min after its fold-specific cutoff. Training sample sizes increased from 2467 to 8500 profiles, while each test window contained 720 profiles. The exact cutoffs, test windows and sample counts are provided in Supplementary Material S1. Figure 1 summarises the timing of the expanding training histories, fixed separation intervals and subsequent held-out test windows for F1–F6.

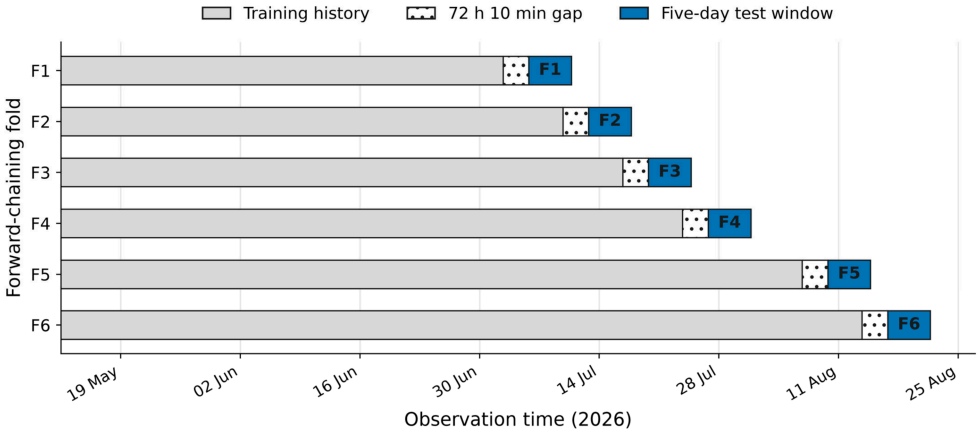


Figure 1. Chronological forward-chaining evaluation folds. Light-grey bars show the training history used for each fold; dotted blocks show the fixed 72 h 10 min separation between the fit cutoff and test start; and blue-grey bars show the subsequent five-day held-out test windows F1–F6. Each test window contains 720 consecutive 10 min profiles. Exact cutoffs, test windows and sample counts are provided in Supplementary Material S1.

2.2. Vertical Structure and Sparse Height Selection

For each fold, the training profile matrix X was centred by subtracting the training-period mean profile μ from every row, yielding X_c. The centred matrix was then decomposed by singular value decomposition (SVD) as shown in Equation (1), where U and V contain the left and right singular vectors, respectively, and Σ is the diagonal matrix of singular values. POD is used here as a statistical representation of vertical covariance structure [19,30]. Let s_j denote the j-th singular value in Σ. Its explained-variance fraction EV_j is defined by Equation (2), and K_{95} by Equation (3). Here j indexes POD modes, the sum in Equation (2) runs over all singular values, and K in Equation (3) denotes a cumulative mode count. The symbol ℓ in Equation (2) is a dummy summation index over all singular values.

$$X_c = U\Sigma V^{\mathrm{T}} \tag{1}$$

$$EV_j = \frac{s_j^2}{\sum_\ell s_\ell^2} \tag{2}$$

$$K_{95} = \min\left\{K : \sum_{j=1}^{K} EV_j \geq 0.95\right\} \tag{3}$$

Pairwise Pearson correlations between height levels were also calculated within each training period. K_{95} describes dimensionality but does not set the retained-level count. For a layout with a prescribed retained-level count, POD–QR uses the same number of empirical modes to construct the design. For sparse height selection, m denotes the number of retained height levels, with m = 2, ..., 14 in the evaluated configurations. The main comparison used a common boundary condition: 20 and 300 m were retained in every

POD–QR, uniform, greedy and reference layout. This removes the boundary-extrapolation advantage that would otherwise arise when one method retained the profile endpoints and another did not.

Uniform layouts were constructed from m approximately equally spaced indices over the 15 levels, with 20 and 300 m fixed as endpoints. The implementation used rounded equally spaced indices over 0, ..., 14; if rounding produced a duplicate, unused indices were added deterministically to preserve m unique levels. No observed data were used to choose the uniform layout.

For each m, every layout containing both 20 and 300 m was enumerated. The number of admissible endpoint-matched layouts is the corresponding binomial count; it equals 286 at m = 5 and 1287 at m = 7. The held-out errors were used only to locate the pre-test POD–QR, uniform and greedy layouts within this design space and did not feed back into layout selection.

POD–QR selection used the right singular vectors derived from the pre-test training profiles. The 20 and 300 m endpoint columns were treated as mandatory. Their span was orthogonalised first, after which the remaining m−2 levels were selected sequentially by the largest residual column norm of the first m POD modes. This endpoint-constrained QR selection adapts QR-based DEIM node selection [22], which extends the original DEIM framework [31]. It is also related to data-driven sparse sensor placement [23]. Reconstruction uses interpolation (Section 2.3).

A supervised comparator also started from the mandatory 20 and 300 m levels. The training profiles were divided into four contiguous, approximately equal-sized chronological blocks using deterministic index splitting. At each step, the remaining height that minimised the mean linear-interpolation root-mean-square error (RMSE) across those four training blocks was added. The process continued until m levels were retained. This comparator follows the general idea of objective-driven greedy sensor selection [32].

A separate training-only exhaustive analysis tested whether the inferred minimum retained-level count depended on POD–QR. For each fold and each m from 2 to 14, all m-of-15 height subsets were enumerated without forcing the endpoints. Candidates were scored on four chronological blocks of the pre-test training history using the same reconstruction rule, and the best training subset was then evaluated on the later test window. This allowed the sampling-density transition to be checked without relying on POD–QR.

2.3. Reconstruction, Evaluation Criteria, and Sensitivity Tests

Linear vertical interpolation was the primary reconstruction rule. Let t index test profiles and let $v_{t(z)}$ denote the observed WS_MEAN value at height z in profile t. For an omitted height z bracketed by the nearest retained heights $z_a < z < z_b$, the reconstructed value $\hat{v}_t(z)$ is given by Equation (4).

$$\hat{v}_t(z) = v_t(z_a) + \frac{z - z_a}{z_b - z_a}[v_t(z_b) - v_t(z_a)] \quad (4)$$

The endpoint-matched main comparison always contained 20 and 300 m and therefore required no extrapolation. In the unconstrained exhaustive analysis, a subset omitting one boundary used the nearest retained value outside the observed range. The piecewise cubic Hermite interpolating polynomial (PCHIP) [33] served as a secondary interpolation sensitivity test; full results are reported in Supplementary Material S4.

For evaluation, Ω_m denotes the set of omitted heights when m levels are retained, $|\Omega_m|$ its size, and N the number of profiles in a test fold. The primary metric, missing-height RMSE, is defined in Equation (5). Mean absolute error (MAE) over the same omitted-height

set was also calculated as a secondary diagnostic, and profile-level RMSE was computed for each test profile to estimate the proportion below a specified tolerance.

$$\mathrm{RMSE}_{\mathrm{miss}} = \sqrt{\frac{1}{N|\Omega_m|}\sum_t \sum_{z\in\Omega_m} [\hat{v}_t(z) - v_t(z)]^2} \quad (5)$$

Three reference tolerances, 0.30, 0.50 and 0.70 m s^{-1}, were retained for direct operating-point comparisons. In the tolerance-sweep analysis, ε denotes the RMSE tolerance and the all-six-fold empirical minimum was evaluated over ε = 0.30–0.70 m s^{-1} in 0.01 m s^{-1} increments. The empirical minimum is the smallest m satisfying the stated criterion. The all-six-fold criterion requires every fold-level RMSE to remain below ε. A second, reliability-oriented criterion requires at least 90% of individual profiles, on average across folds, to have RMSE below 0.50 m s^{-1}. The 0.50 m s^{-1} tolerance serves as a representative accuracy criterion for examining the trade-off between reconstruction error and retained-height count. The 90% profile-satisfaction target provides a reliability-oriented assessment. In practical applications, accuracy criteria and retained-height counts can be selected according to specific reconstruction requirements, allowing a flexible balance between reconstruction accuracy, vertical sampling density, and data transmission and storage demands.

Stability between selected height sets was quantified with the pairwise Jaccard index (intersection size divided by union size), and selection frequency was recorded at each height for POD–QR and greedy layouts. To test within-deployment temporal stability rather than set similarity alone, the five- and seven-level POD–QR layouts selected from the earliest training fold (F1) were frozen and applied unchanged to all later test folds F2–F6. Their errors were compared with the corresponding fold-specific POD–QR selections; the frozen layouts were not re-optimised.

Vertical-profile complexity was quantified on each held-out test profile using the root-mean-square (RMS) second finite difference defined in Equation (6). Here i is the reported-height index, $v_{i,t}$ denotes the WS_MEAN value at height index i in profile t, and C_t is the resulting curvature index.

$$C_t = \sqrt{\frac{1}{13}\sum_{i=2}^{14}\left[v_{i+1,t} - 2v_{i,t} + v_{i-1,t}\right]^2} \quad (6)$$

The sum in Equation (6) uses the 13 interior levels, i = 2, …, 14. The 4320 non-overlapping held-out profiles were pooled after all height selections were fixed and stratified into quartiles of C_t. Reconstruction RMSE and the probability of profile RMSE below 0.50 m s^{-1} were then evaluated within each quartile; Spearman correlation quantified the monotonic association between C_t and profile-level error.

Reconstruction performance was also evaluated across wind-speed groups, day/night periods and vertical wind-speed gradients using the original fold-specific layouts. Low, medium and high wind-speed groups were defined by the training-period P50 and P90 of profile-mean WS_MEAN for each fold; daytime and nighttime were defined as 06:00–18:00 and 18:00–06:00 (Beijing time). The vertical-gradient index was the largest absolute WS_MEAN difference between adjacent 20 m levels divided by 20 m, with the training-period P90 defining the high-gradient group. Group RMSE pooled errors over omitted heights, while profile satisfaction was the proportion with profile RMSE below 0.50 m s^{-1}. Height-specific errors were assessed only when a height was omitted (Supplementary Material S7).

Observation-degradation sensitivity retained two complementary tests. First, each selected height in the five- and seven-level POD–QR configurations was removed in turn and reconstruction error was recomputed with the unavailable level included in the target set; single-level, double-level and contiguous-band loss results are summarised in the main text,

with numerical summaries retained in Supplementary Material S5. Second, independent Gaussian perturbations were added to retained observations while the pre-test-selected height sets were kept fixed. The perturbation model is given in Equation (7), where vobs is the perturbed observed value, v is the corresponding unperturbed value, η is the added zero-mean Gaussian perturbation, and σ is its standard deviation.

$$v_{\text{obs}} = v + \eta, \quad \eta \sim \text{Normal}(0, \sigma^2) \tag{7}$$

The tested values of σ were 0, 0.10, 0.25, 0.50 and 1.00 m s^{-1}, with 100 Monte Carlo realisations for each non-zero condition, fold and retained-level count. The empirical minimum retained-level count was recomputed at each perturbation level using the 0.50 m s^{-1} criterion. Monte Carlo settings are summarised in Supplementary Material S6.

A separate sensitivity test examined rare extreme adjacent-level discontinuities in the reported product. For each profile, the maximum absolute WS_MEAN difference between two adjacent 20 m reported levels was calculated. Profiles exceeding thresholds of 4, 5 or 6 m s^{-1} were flagged at each threshold. At each threshold, flagged profiles were removed from both the pre-test fitting history and the held-out test windows, and the complete endpoint-constrained POD–QR scan from m = 2 to 14 was repeated. Flagging was not treated as evidence of physical error; it tested whether a small number of extreme reported-product discontinuities drove the inferred operating points.

2.4. Software and Implementation

All analyses were implemented in Python 3.13.5 using NumPy 2.3.5, pandas 2.2.3, SciPy 1.17.0, and Matplotlib 3.10.8. Pivoted QR and related linear-algebra operations used SciPy and NumPy routines, and PCHIP used the SciPy interpolation module. Height selection, chronological fold construction, sensitivity analyses, and figure generation followed deterministic or pre-specified computational procedures.

3. Results

3.1. Vertical Structure and Temporal Stability

The 15-level WS_MEAN product exhibited strong vertical covariance and consistently low-dimensional structure in every chronological training fold. Median adjacent-height correlation ranged from 0.987 to 0.991. The first two POD modes explained 96.67–97.14% of training-period variance, giving $K_{95} = 2$ in all six folds despite the increase in available historical profiles from 2467 to 8500 (Figure 2).

The five-level endpoint-constrained POD–QR layouts had a mean pairwise Jaccard index of 0.889. The dominant layout was 20, 60, 120, 220 and 300 m: 20, 120, 220 and 300 m were selected in all six folds, while 60 m was selected in five. The seven-level layouts were even more stable, with mean Jaccard 0.917 and a dominant configuration of 20, 40, 100, 140, 200, 260 and 300 m; one fold replaced 260 m with 240 m. Endpoint-constrained greedy selection was less stable, with mean Jaccard values of 0.822 and 0.725 at five and seven retained levels, respectively (Figure 3).

Frozen-layout performance was also stable over later periods within the same deployment. Over future folds F2–F6, the frozen five-level layout had mean RMSE 0.4101 m s^{-1} versus 0.4124 m s^{-1} for fold-specific reselection; the corresponding seven-level values were 0.3483 and 0.3486 m s^{-1}. Worst-fold RMSE remained 0.4907 m s^{-1} for five levels and 0.4573 m s^{-1} for seven levels. The early layouts therefore retained comparable performance over the later evaluation period at this site (Table 2).

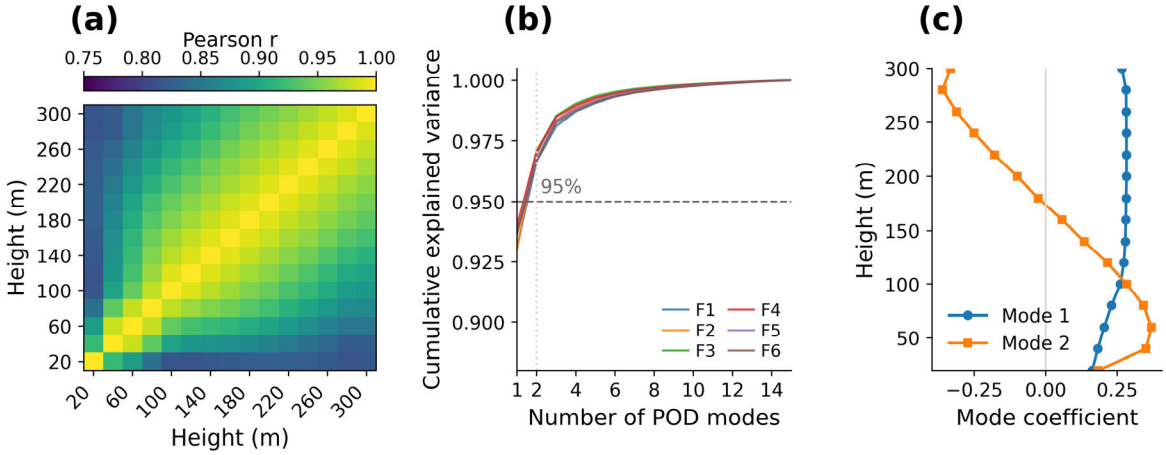


Figure 2. Vertical covariance and low-dimensional structure of the radar wind-speed product. (**a**) Height-to-height Pearson correlation matrix for the largest pre-test training set. (**b**) Cumulative explained variance across the six chronological training sets; the horizontal dashed line marks 95%, and the vertical dotted line marks two POD modes. (**c**) First two proper orthogonal decomposition (POD) modes from the largest pre-test training set. Mode signs are arbitrary and are shown only to describe empirical covariance structure.

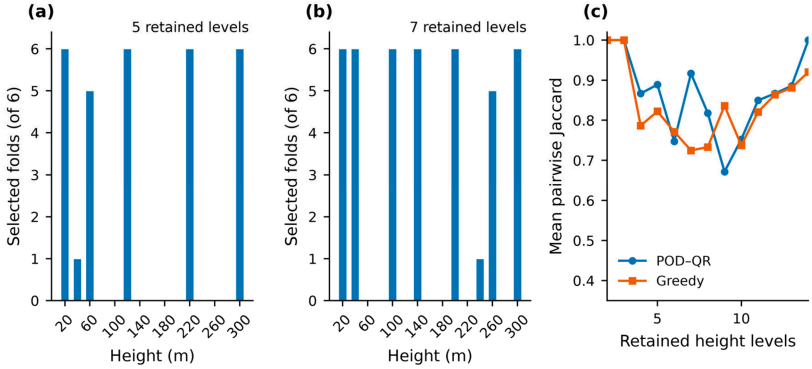


Figure 3. Chronological stability of endpoint-constrained height selection. (**a**,**b**) Selection frequency for proper orthogonal decomposition with QR-based height selection (POD–QR) across the six folds for five and seven retained levels. (**c**) Mean pairwise Jaccard similarity for POD–QR and supervised greedy layouts across retained-level counts.

Table 2. Within-deployment temporal-stability summary for future folds F2–F6. RMSE denotes root-mean-square error, and MAE denotes mean absolute error.

Retained Levels	Layout Mode	Mean RMSE ($m\ s^{-1}$)	Worst-Fold RMSE ($m\ s^{-1}$)	Mean MAE ($m\ s^{-1}$)	Profiles with RMSE < 0.50 $m\ s^{-1}$ (%)
5	Fold-specific	0.4124	0.4907	0.2774	82.9%
5	Frozen-F1	0.4101	0.4907	0.2757	83.6%
7	Fold-specific	0.3486	0.4586	0.2282	89.7%
7	Frozen-F1	0.3483	0.4573	0.2284	89.8%

The representative five- and seven-level configurations also maintained comparable reconstruction performance when the fitting histories were restricted to the high-completeness period (Supplementary Material S8).

3.2. Sampling Sufficiency and Layout Performance

With the endpoints matched across methods, the height-placement effect remained clear. At five retained levels, POD–QR achieved a mean missing-height RMSE of 0.3948 m s^{-1}, compared with 0.4860 m s^{-1} for the exact median of all 286 admissible layouts, an 18.8% reduction. Across the six test folds, the selected POD–QR layouts ranked between the best 0.35% and 5.94% of the complete layout distribution. Uniform sampling produced 0.4385 m s^{-1}, while endpoint-constrained greedy sampling produced 0.3937 m s^{-1} but was less stable across folds (Figure 4a).

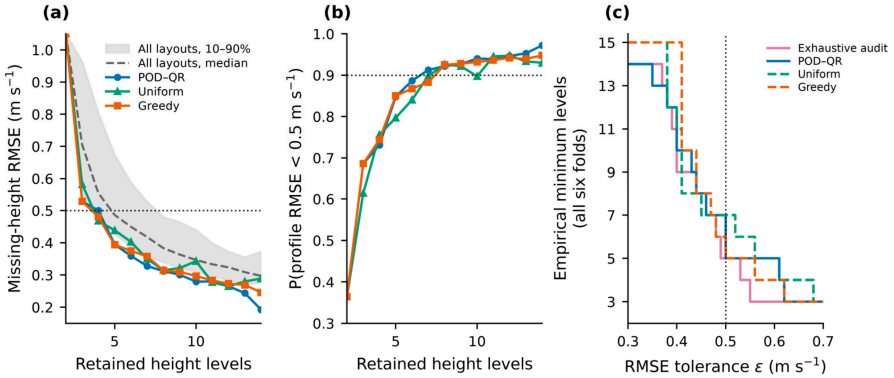


Figure 4. Endpoint-matched sampling density and conditional radar-product recoverability. (**a**) Mean missing-height root-mean-square error (RMSE) for proper orthogonal decomposition with QR-based height selection (POD–QR), uniform and supervised greedy layouts, together with the exact median and 10–90% range of all endpoint-matched layouts. (**b**) Mean empirical probability that an individual test profile has RMSE below 0.50 m s^{-1}. (**c**) Empirical minimum retained-level count as a function of the RMSE tolerance ε for the all-six-fold criterion; a plotted value of 15 denotes that no configuration from 2 to 14 retained levels met the criterion.

At seven retained levels, POD–QR achieved 0.3276 m s^{-1} compared with 0.4173 m s^{-1} for the median of all 1287 endpoint-matched layouts, a 21.5% reduction. Fold-specific POD–QR layouts ranked within the best 0.08–4.74% of the exact distribution. Uniform and greedy sampling produced 0.3499 and 0.3574 m s^{-1}, respectively. At four retained levels, uniform sampling performed better than POD–QR, showing that the placement advantage was sampling-density dependent. The exact fold-wise operating-point ranks are summarised in Table 3. A post hoc oracle-gap diagnostic showed that the residual distance from the fold-wise best admissible layout was only 9.0% and 10.0% of the median-to-best error range at five and seven levels, respectively (Table 4). The full endpoint-matched design-space distributions for m = 2–14 are provided in Supplementary Material S3 (Table S3).

Table 3. Fold-wise rank of proper orthogonal decomposition with QR-based height selection (POD–QR) within the complete endpoint-matched design space at the two representative operating points. Percentiles are measured from the best-performing side.

Fold	5-Level Rank (of 286)	5-Level Percentile	7-Level Rank (of 1287)	7-Level Percentile
1	4/286	1.40%	3/1287	0.23%
2	2/286	0.70%	1/1287	0.08%
3	17/286	5.94%	20/1287	1.55%
4	6/286	2.10%	2/1287	0.16%
5	1/286	0.35%	14/1287	1.09%
6	2/286	0.70%	61/1287	4.74%

Table 4. Post hoc oracle-gap diagnostic within the complete endpoint-matched design space. RMSE denotes root-mean-square error. The normalised gap is the selected-to-best residual error divided by the median-to-best error range.

Retained Levels	Selected RMSE ($m\ s^{-1}$)	Best RMSE ($m\ s^{-1}$)	Mean Normalised Gap	Maximum Normalised Gap	Median-to-Best Improvement Captured (%)
5	0.3948	0.3865	0.090	0.219	91.0%
7	0.3276	0.3186	0.100	0.304	90.0%

At the 0.50 m s^{-1} study criterion, the independent training-only exhaustive analysis showed a transition between four and five retained levels. Exhaustive selection gave mean test RMSE values of 0.4616, 0.4256 and 0.3865 m s^{-1} at three, four and five retained levels, respectively. Five of six folds were below 0.50 m s^{-1} at both three and four levels; five levels were the first count for which all six folds were below the tolerance, with a worst-fold RMSE of 0.4826 m s^{-1} (Table 5). The complete results for m = 2–14 are reported in Supplementary Material S2 (Table S2).

Table 5. Independent training-only exhaustive analysis around the 0.50 m s^{-1} sampling-density transition. RMSE denotes root-mean-square error.

Retained Levels	Mean Test RMSE ($m\ s^{-1}$)	Worst-Fold RMSE ($m\ s^{-1}$)	Folds with RMSE < 0.50 m s^{-1}	Criterion Status
3	0.4616	0.5409	5/6	One fold fails
4	0.4256	0.5216	5/6	One fold fails
5	0.3865	0.4826	6/6	First all-fold minimum

A complementary fixed-layout comparison supported the same 0.50 m s^{-1} operating point. Among all endpoint-matched layouts, none of the 13 three-level or 78 four-level layouts achieved the criterion in all six test folds, whereas 4 of 286 five-level layouts did. Five levels were therefore the empirical minimum at this operating point for both the exhaustive training-only analysis and POD–QR; uniform sampling required seven levels, while endpoint-constrained greedy sampling also reached the criterion at five. More generally, the continuous tolerance sweep showed that the required count was not fixed: for POD–QR it decreased from 14 levels at ε = 0.30 m s^{-1} to 10 at 0.40, 5 at 0.50 and 3 at 0.70 m s^{-1} (Figure 4c).

At the individual-profile level, five POD–QR levels produced RMSE below 0.50 m s^{-1} for 84.8% of test profiles. Seven retained levels increased this fraction to 91.3%, so seven was the first POD–QR configuration to exceed the 90% profile-level criterion. The five- and seven-level cases are therefore used as representative operating points, not universal sampling requirements (Figure 4b; Table 6).

The sampling-density transition was also insensitive to the interpolation rule. Using the same pre-test-selected endpoint-constrained POD–QR layouts, PCHIP gave a mean RMSE of 0.3884 m s^{-1} at five retained levels, compared with 0.3948 m s^{-1} for linear interpolation; at seven levels the corresponding values were 0.3363 and 0.3276 m s^{-1}. The all-six-fold minimum remained 14, 5 and 3 retained levels at the 0.30, 0.50 and 0.70 m s^{-1} reference tolerances for both reconstruction rules. The full m = 2–14 sensitivity curve is retained in Supplementary Material S4.

Table 6. Clean-observation operating points for the 10 min mean horizontal wind-speed (WS_MEAN) product. RMSE denotes root-mean-square error, and POD–QR denotes proper orthogonal decomposition with QR-based height selection. Performance values are aggregated across fold-specific POD–QR selections; the listed heights are the dominant configurations across folds, not one fixed layout used for every fold.

Configuration	Dominant Retained Heights	Mean RMSE ($m\ s^{-1}$)	Profiles with RMSE $< 0.50\ m\ s^{-1}$ (%)	Interpretation
5 levels	20, 60, 120, 220, 300 m	0.395	84.8%	Representative efficiency-oriented point under the study-specific all-six-fold 0.50 $m\ s^{-1}$ criterion
7 levels	20, 40, 100, 140, 200, 260, 300 m	0.328	91.3%	Representative reliability-oriented point; first POD–QR retained-level count with ≥90% of profiles below 0.50 $m\ s^{-1}$

3.3. Profile Complexity and Observation Sensitivity

Recoverability depended strongly on vertical-profile structure. For five retained levels, missing-height RMSE increased from 0.163 $m\ s^{-1}$ in the lowest-curvature quartile to 0.648 $m\ s^{-1}$ in the highest, while the fraction of profiles below 0.50 $m\ s^{-1}$ fell from 100.0% to 45.9%. Seven retained levels reduced the corresponding highest-quartile RMSE to 0.572 $m\ s^{-1}$ and increased the success fraction to 65.1%. Across all held-out profiles, Spearman correlations between the curvature index and profile-level RMSE were 0.893 and 0.882 for five and seven levels, respectively (Figure 5 and Table 7).

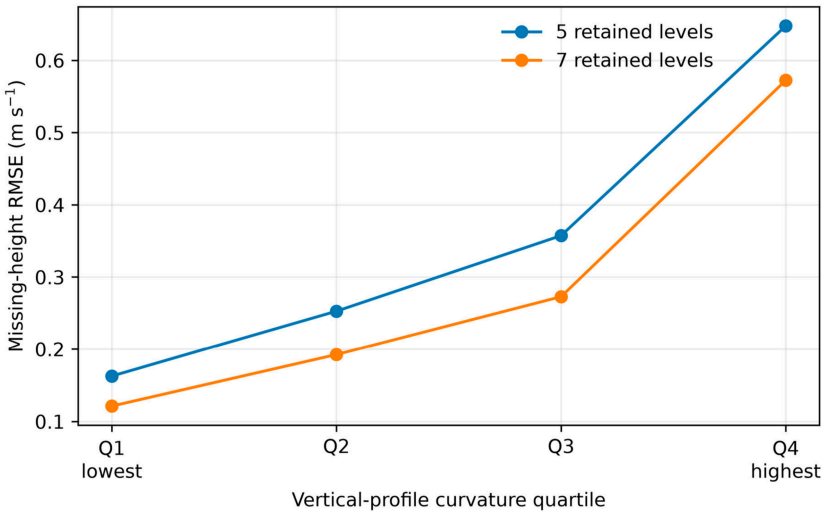


Figure 5. Vertical-profile complexity and sparse-reconstruction performance. Held-out profiles are stratified by the root-mean-square (RMS) second finite difference after all height selections are fixed; the figure shows the dependence of reconstruction error on profile complexity for the five- and seven-level configurations.

Selected-level loss exposed a clear efficiency–redundancy trade-off. Removing one selected level increased mean RMSE from 0.3948 to 0.5125 $m\ s^{-1}$ for the five-level configuration and from 0.3276 to 0.3901 $m\ s^{-1}$ for seven levels; 60.0% and 92.9% of fold-by-loss cases, respectively, remained below 0.50 $m\ s^{-1}$. With two selected levels unavailable, mean RMSE increased to 0.6975 and 0.4773 $m\ s^{-1}$. Contiguous loss produced the same ordering: for unavailable spans of 40, 80 and 120 m, mean RMSE was 0.5908, 0.6608 and 0.7657 $m\ s^{-1}$

for five retained levels, compared with 0.4662, 0.5701 and 0.7060 m s^{-1} for seven (Figure 6). Numerical summaries are provided in Supplementary Material S5.

Table 7. Reconstruction performance by vertical-profile complexity quartile. RMSE denotes root-mean-square error; the curvature index is the root-mean-square second finite difference in the reported wind-speed profile.

Retained Levels	Complexity Quartile	Mean Curvature Index (m s^{-1})	Profiles, *n*	RMSE (m s^{-1})	Profiles with RMSE < 0.50 m s^{-1} (%)
5	Q1 lowest	0.214	1080	0.1626	100.0%
5	Q2	0.356	1080	0.2527	99.7%
5	Q3	0.510	1080	0.3574	93.6%
5	Q4 highest	0.966	1080	0.6477	45.9%
7	Q1 lowest	0.214	1080	0.1212	100.0%
7	Q2	0.356	1080	0.1924	100.0%
7	Q3	0.510	1080	0.2728	99.9%
7	Q4 highest	0.966	1080	0.5724	65.1%

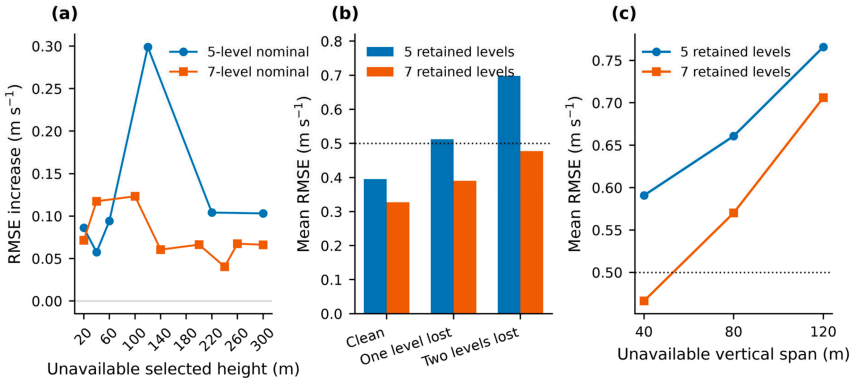


Figure 6. Selected-height availability sensitivity. RMSE denotes root-mean-square error. (**a**) Mean increase in missing-height RMSE when one selected height is unavailable. (**b**) Mean RMSE for clean, single-level-loss and double-level-loss scenarios; the dotted line marks 0.50 m s^{-1}. (**c**) Mean RMSE when contiguous 40, 80 or 120 m vertical spans are unavailable. Numerical summaries are provided in Supplementary Material S5.

Added observation noise increased reconstruction error monotonically. At σ = 0.50 m s^{-1}, the five-level mean RMSE rose to 0.5501 m s^{-1}, whereas the seven-level configuration remained at 0.4959 m s^{-1}; at σ = 1.00 m s^{-1} the corresponding values were 0.8585 and 0.8084 m s^{-1}.

The retained-level requirement increased with perturbation amplitude: the all-six-fold 0.50 m s^{-1} criterion required 5, 5, 7 and 13 levels at σ = 0, 0.10, 0.25 and 0.50 m s^{-1}, respectively; no configuration from 2 to 14 levels met the criterion at σ = 1.00 m s^{-1} (Figure 7). The combined sensitivity results indicate that five levels are an efficiency-oriented clean-data point, whereas seven levels provide additional margin but do not remove the penalty associated with strong vertical curvature.

Removing rare extreme adjacent-level discontinuities did not change the operating-point counts. Thresholds of 4, 5 and 6 m s^{-1} flagged 77 (0.74%), 34 (0.33%) and 19 (0.18%) of the 10,355 profiles, respectively. After excluding flagged profiles from both fitting and held-out periods and repeating the full POD–QR scan, five levels remained the minimum satisfying the all-six-fold 0.50 m s^{-1} criterion and seven remained the minimum satisfying the 90% profile criterion at every threshold. At the central 5 m s^{-1} flag, only six held-out profiles were excluded; five-level mean and worst-fold RMSE were 0.3858 and 0.4694 m s^{-1},

while seven-level mean RMSE was 0.3207 m s^{-1} with 90.9% of profiles below 0.50 m s^{-1} (Table 8).

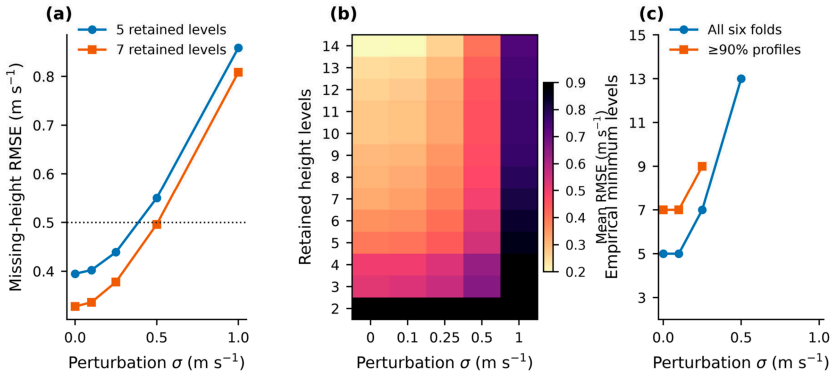


Figure 7. Observation-quality sensitivity. RMSE denotes root-mean-square error, and POD–QR denotes proper orthogonal decomposition with QR-based height selection. (**a**) Missing-height RMSE of the five- and seven-level POD–QR configurations under Gaussian perturbations. (**b**) Mean RMSE across the full 2–14 retained-level scan. (**c**) Empirical minimum retained-level count under the all-six-fold and ≥90%-of-profiles criteria. Monte Carlo settings and numerical summaries are provided in Supplementary Material S6.

Table 8. Sensitivity of the representative operating points to conservative adjacent-level discontinuity flags. RMSE denotes root-mean-square error. The last column reports the minimum retained-level counts for the all-six-fold/≥90%-of-profiles criteria.

Jump Threshold (m s^{-1})	Excluded Profiles, *n* (%)	Excluded Held-Out Profiles, *n*	5-Level RMSE, Mean/Worst (m s^{-1})	7-Level Mean RMSE (m s^{-1})	7-Level Profiles < 0.50 m s^{-1} (%)	Minimum Retained Levels, m (All-Six/≥90%)
4	77 (0.74%)	16	0.3810/0.4554	0.3149	91.3%	5/7
5	34 (0.33%)	6	0.3858/0.4694	0.3207	90.9%	5/7
6	19 (0.18%)	2	0.3901/0.4817	0.3254	90.9%	5/7

3.4. Reconstruction Performance Under Different Observational Conditions

Across low-, medium- and high-wind groups and both day/night periods, the seven-level POD–QR layouts had lower omitted-height RMSE than the five-level layouts (Table 9). The same ordering held in all 12 fold-by-day/night comparisons, each containing 360 profiles.

Table 9. Reconstruction performance across wind-speed and day/night groups. RMSE is pooled over omitted heights (m s^{-1}); profile success is the percentage of profiles with profile RMSE below 0.50 m s^{-1}.

Test Group	Profiles, *n*	5-Level RMSE (m s^{-1})	5-Level Success (%)	7-Level RMSE (m s^{-1})	7-Level Success (%)
Low wind	2648	0.3629	89.27	0.3166	93.05
Medium wind	1223	0.4488	78.41	0.3860	86.59
High wind	449	0.4563	75.95	0.3075	93.32
Daytime	2160	0.4553	79.03	0.3937	86.85
Nighttime	2160	0.3339	90.60	0.2682	95.65

The vertical-gradient analysis showed substantially higher errors for high-gradient profiles (Table 10). At five and seven levels, high-gradient RMSEs were 0.7565 and 0.6723 m s^{-1}, compared with 0.3101 and 0.2481 m s^{-1} for other-gradient profiles. At 100 m, omitted in all

six five-level layouts ($n = 4320$), RMSE and absolute-error P95 were 0.4620 and 0.9467 m s^{-1}, respectively; the seven-level layouts retained this height (Supplementary Material S7, Table S9).

Table 10. Gradient-stratified and representative height-specific reconstruction errors. RMSE and absolute-error P95 are evaluated at omitted heights (m s^{-1}); profile success applies only to gradient groups.

Condition	Retained Levels	Profiles, *n*	RMSE (m s^{-1})	Absolute Error P95 (m s^{-1})	Profile Success (%)
Other gradient (≤P90)	5	3746	0.3101	0.6500	92.71
Other gradient (≤P90)	7	3746	0.2481	0.5233	97.46
High gradient (>P90)	5	574	0.7565	1.5402	33.28
High gradient (>P90)	7	574	0.6723	1.3282	50.70
100 m (omitted)	5	4320	0.4620	0.9467	—

4. Discussion

4.1. Sampling Sufficiency and Design-Space Performance

The central result is the dependence of required sampling density on the reconstruction criterion. For POD–QR, the all-six-fold minimum changed from 14 retained levels at $\varepsilon = 0.30$ m s^{-1} to 10 at 0.40, 5 at 0.50 and 3 at 0.70 m s^{-1}. Five levels therefore provide an efficiency-oriented operating point at the 0.50 m s^{-1} criterion. Seven levels instead mark the first configuration for which at least 90% of individual profiles, on average, fall below the same error level. Relating sampling density to an explicit evaluation objective is consistent with lidar scan-design studies that consider spatial resolution and statistical error [15]. Required density consequently depends on reconstruction tolerance, observation quality and availability, reconstruction method and reported-profile complexity.

The endpoint-matched comparison addresses which levels should be retained. Once 20 and 300 m were fixed as common boundaries, the pre-test-selected POD–QR layouts remained in the upper tail of the complete admissible design space at both five and seven retained levels. The post hoc oracle-gap diagnostic showed that these layouts captured about 91% and 90% of the possible improvement from the median layout to the fold-wise best layout at five and seven levels, respectively. The selected layouts therefore approached the best admissible performance, rather than merely outperforming a single comparator. This near-best admissible performance is consistent with reduced-order sensor-placement studies showing that empirical field structure can guide informative placement, while the preferred layout remains dependent on the selection objective [22,23,32].

Freezing the earliest POD–QR layout produced almost the same mean RMSE as reselecting the layout for each later fold. This supports within-deployment temporal stability over the later F2–F6 periods and shows that the selected geometry did not require continual re-optimisation during this record. Performance was also not uniformly best at every sampling density: uniform spacing was better with four retained levels, and greedy selection was nearly tied with POD–QR at five. The practical value is therefore the combination of high design-space rank, chronological selection and future-period stability within the observed deployment.

4.2. Conditions of Recoverability

Formal retrieval information content, empirical product dimensionality and sampling sufficiency describe different properties of the measurement. Retrieval information content is characterised in inverse theory through quantities such as averaging kernels and degrees of freedom for signal [6]. Empirical product dimensionality describes covariance structure in the reported product through POD [19,30]. In this study, two POD modes account for about 97% of variance throughout the expanding training histories (Figure 2). Sampling sufficiency instead asks how many actual reported height values are required, under a specified reconstruction operator and error criterion, to recover the omitted product levels. Wind-profiler and Doppler-lidar studies showing that effective vertical resolution can differ materially from gate or reporting spacing reinforce the need to keep these concepts separate [8,9].

The profile-complexity and sensitivity analyses define when that enabling structure can be exploited. Linear interpolation was highly accurate for vertically smooth reported profiles but degraded as local curvature increased, while selected-level loss (Figure 6) and observation perturbation (Figure 7) increased the sampling requirement. Because the analysis table does not contain a per-level retrieval-quality flag, high curvature cannot be assigned exclusively to atmospheric structure; it may also include product-level retrieval anomalies. The adjacent-level discontinuity test addresses this concern: removing the small fraction of profiles with extreme 20 m jumps across three conservative thresholds did not change the five- or seven-level operating points. The test cannot determine whether every flagged or unflagged structure is physically resolved. Doppler-wind-lidar studies provide related measurement context, showing that retrieval error depends on flow structure, observation height and sampling-volume effects [34]. PCHIP produced the same POD–QR minima at the 0.30, 0.50 and 0.70 m s^{-1} reference tolerances in this dataset, although its error curve differed from linear interpolation (Supplementary Material S4).

From an engineering perspective, the proposed criterion-based approach provides a practical basis for configuring wind-profile data products according to application requirements. For routine wind monitoring, data transmission, and archiving, retaining five of the original 15 height levels reduces the number of wind-speed values by 66.7% while satisfying the 0.50 m s^{-1} RMSE criterion across all six test folds. Retaining seven levels reduces the reported values by 53.3% while increasing the mean profile-level satisfaction rate to 91.3%. These findings support differentiated wind-profile data services, in which compact height configurations can be used for routine reporting and storage, while applications requiring detailed vertical wind information can adopt more stringent accuracy criteria and retain additional height levels.

4.3. Relation to Remote-Sensing Information and Sampling Design

The interpretation draws on retrieval-information studies as well as sparse measurement design. Retrieval theory distinguishes reporting grids from independently resolved information [6,7], and wind-profile studies show that effective vertical resolution need not match reporting spacing [8,9]. Optimal-estimation wind retrievals can combine prior information with observations from multiple instruments to address limited observational information and vertical data gaps [35]. Application-oriented lidar optimisation shows how scan patterns and instrument configurations can be tailored to specific measurement objectives [13]. Scan design and temporal averaging affect retrieval error and sampling requirements [14,15], while modular retrieval workflows accommodate heterogeneous lidar systems and scan patterns [16]. Sparse-sensing studies address the complementary question of which measurement locations carry the most useful information for reconstruction [18,22,23]. Recent urban-wind reconstruction work has likewise combined reduced-

order POD representations with constraint-aware QR pivoting to place sparse sensors for high-resolution field recovery [26]. That setting targets spatial wind-field reconstruction from deployment-constrained sensor networks; here we evaluate how many reported radar heights are needed to meet an explicit vertical reconstruction criterion, linking retrieval-information concepts with sparse measurement design at the product level.

The selected height layouts provide a practical basis for optimising existing radar wind-profile products. Informative heights can be identified from historical observations, allowing selected wind-speed values to be retained or transmitted and omitted levels to be reconstructed by linear interpolation when complete profiles are needed. This product-level approach can be implemented through suitable output or downstream processing interfaces, with potential benefits for reducing data transmission and storage requirements. The reconstruction framework may also support missing-level recovery and data quality checks where additional observations are available for comparison.

The VORTRAD-L300 provides 15- or 30-level products with selectable 10 or 20 m range resolution. Beyond these existing configurations, the identified height priorities may guide future improvements in radar sampling strategies and instrument design. Acquisition-level implementation could involve adaptations to the sampling geometry, hardware or control system, with configurations tailored to practical measurement requirements. Further field evaluation would help establish the achievable reconstruction accuracy, data availability and actual transmission and storage benefits.

4.4. Limitations

These findings concern WS_MEAN profiles from one coastal deployment during May–August 2026; their transferability to other sites, seasons, instruments and wind-direction or vertical-shear products requires further assessment. The frozen-layout result supports future-period stability within this deployment only. The profile-complexity index is an internal structural diagnostic of the reported product, and the adjacent-level discontinuity analysis is a sensitivity check rather than an instrument-level quality-control system. Independent instrument-error characterisation would still require calibrated measurements.

5. Conclusions

The required vertical sampling density was criterion-dependent. For POD–QR, the all-six-fold empirical minimum changed from 14 retained levels at an RMSE tolerance of 0.30 m s^{-1} to 10 at 0.40, 5 at 0.50 and 3 at 0.70 m s^{-1}. Under the representative 0.50 m s^{-1} criterion, five levels gave a mean RMSE of 0.395 m s^{-1}, while seven levels reduced mean RMSE to 0.328 m s^{-1} and were the first to exceed the study-specific 90% profile-level target. These two operating points were unchanged by a conservative sensitivity analysis that removed rare profiles with extreme adjacent-level discontinuities. The five-level POD–QR layouts ranked within the best 5.94% of endpoint-matched alternatives in every fold, and the earliest selected layout retained comparable performance over later periods of the same deployment. The first two POD modes explained about 97% of variance, but low rank alone did not determine the sampling requirement: reported-profile curvature, selected-level loss, observation noise and the reconstruction rule all mattered. The study therefore provides a criterion-based framework for vertical reporting-grid design, channel prioritisation and prospective sparse-acquisition testing.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs18223932/s1, S1, chronological fold cutoffs, test windows and sample counts; S2, training-only exhaustive subset-selection summary for m = 2–14; S3, endpoint-matched design-space distributions; S4, PCHIP reconstruction sensitivity; S5, observation-availability

sensitivity; S6, observation-quality perturbation and nominal-grid availability; S7, meteorological stratification and local errors; S8, missing-data episodes and training-history sensitivity.

Author Contributions: Conceptualization, Y.L. and D.J.; methodology, Y.L. and D.J.; software, Y.L.; validation, Y.L.; formal analysis, Y.L.; investigation, Y.L. and Y.G.; resources, Y.G.; data curation, Y.L. and Y.G.; visualisation, Y.L.; writing—original draft preparation, Y.L.; writing—review and editing, Y.L., Y.G. and D.J.; supervision, D.J.; project administration, D.J. and Y.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Qingdao Science and Technology Demonstration Project (Project No. 26-1-1-sfgc-2-hy).

Data Availability Statement: The data and reproduction materials supporting the findings of this study are not publicly available due to restrictions associated with data ownership and project use. A complete reproduction package, including the analysis data and code required to verify the reported results, is provided confidentially to the editors and peer reviewers for peer-review verification. Access beyond peer review is subject to the applicable data-use restrictions and permission of the data owner.

Conflicts of Interest: Author Yiyun Guo was employed by the company Qingdao Port International Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Strauch, R.G.; Weber, B.L.; Frisch, A.S.; Little, C.G.; Merritt, D.A.; Moran, K.P.; Welsh, D.C. The Precision and Relative Accuracy of Profiler Wind Measurements. *J. Atmos. Ocean. Technol.* **1987**, *4*, 563–571. [CrossRef]
2. Weber, B.L.; Wuertz, D.B. Comparison of Rawinsonde and Wind Profiler Radar Measurements. *J. Atmos. Ocean. Technol.* **1990**, *7*, 157–174. [CrossRef]
3. Päschke, E.; Leinweber, R.; Lehmann, V. An Assessment of the Performance of a 1.5 µm Doppler Lidar for Operational Vertical Wind Profiling Based on a 1-Year Trial. *Atmos. Meas. Tech.* **2015**, *8*, 2251–2266. [CrossRef]
4. Pichugina, Y.L.; Banta, R.M.; Brewer, W.A.; Sandberg, S.P.; Hardesty, R.M. Doppler Lidar–Based Wind-Profile Measurement System for Offshore Wind-Energy and Other Marine Boundary Layer Applications. *J. Appl. Meteorol. Climatol.* **2012**, *51*, 327–349. [CrossRef]
5. Liu, Z.; Barlow, J.F.; Chan, P.-W.; Fung, J.C.H.; Li, Y.; Ren, C.; Mak, H.W.L.; Ng, E. A Review of Progress and Applications of Pulsed Doppler Wind LiDARs. *Remote Sens.* **2019**, *11*, 2522. [CrossRef]
6. Rodgers, C.D. *Inverse Methods for Atmospheric Sounding: Theory and Practice*; World Scientific: Singapore, 2000. [CrossRef]
7. von Clarmann, T.; Degenstein, D.A.; Livesey, N.J.; Bender, S.; Braverman, A.; Butz, A.; Compernolle, S.; Damadeo, R.; Dueck, S.; Eriksson, P.; et al. Overview: Estimating and reporting uncertainties in remotely sensed atmospheric composition and temperature. *Atmos. Meas. Tech.* **2020**, *13*, 4393–4436. [CrossRef]
8. Merceret, F.J. The Vertical Resolution of the Kennedy Space Center 50-MHz Wind Profiler. *J. Atmos. Ocean. Technol.* **1999**, *16*, 1273–1278. [CrossRef]
9. Bedka, K.M.; Nehrir, A.R.; Kavaya, M.; Barton-Grimley, R.; Beaubien, M.; Carroll, B.; Collins, J.; Cooney, J.; Emmitt, G.D.; Greco, S.; et al. Airborne lidar observations of wind, water vapor, and aerosol profiles during the NASA Aeolus calibration and validation (Cal/Val) test flight campaign. *Atmos. Meas. Tech.* **2021**, *14*, 4305–4334. [CrossRef]
10. Sathe, A.; Mann, J.; Vasiljević, N.; Lea, G. A Six-Beam Method to Measure Turbulence Statistics Using Ground-Based Wind Lidars. *Atmos. Meas. Tech.* **2015**, *8*, 729–740. [CrossRef]
11. Newman, J.F.; Klein, P.M.; Wharton, S.; Sathe, A.; Bonin, T.A.; Chilson, P.B.; Muschinski, A. Evaluation of Three Lidar Scanning Strategies for Turbulence Measurements. *Atmos. Meas. Tech.* **2016**, *9*, 1993–2013. [CrossRef]
12. Meyer Forsting, A.R.; Troldborg, N.; Borraccino, A. Modelling Lidar Volume-Averaging and Its Significance to Wind Turbine Wake Measurements. *J. Phys. Conf. Ser.* **2017**, *854*, 012014. [CrossRef]
13. Simley, E.; Fürst, H.; Haizmann, F.; Schlipf, D. Optimizing Lidars for Wind Turbine Control Applications—Results from the IEA Wind Task 32 Workshop. *Remote Sens.* **2018**, *10*, 863. [CrossRef]
14. Rahlves, C.; Beyrich, F.; Raasch, S. Scan strategies for wind profiling with Doppler lidar—An large-eddy simulation (LES)-based evaluation. *Atmos. Meas. Tech.* **2022**, *15*, 2839–2856. [CrossRef]
15. Letizia, S.; Zhan, L.; Iungo, G.V. LiSBOA (LiDAR Statistical Barnes Objective Analysis) for optimal design of lidar scans and retrieval of wind statistics—Part 1: Theoretical framework. *Atmos. Meas. Tech.* **2021**, *14*, 2065–2093. [CrossRef]
16. Erdmann, A.; Gasch, P. Modular Wind Profile Retrieval Software for Heterogeneous Doppler Lidar Measurements (AtmoProKIT v1.1). *Geosci. Model. Dev.* **2026**, *19*, 2497–2529. [CrossRef]

17. Towers, P.; Jones, B.L. Real-Time Wind Field Reconstruction from LiDAR Measurements Using a Dynamic Wind Model and State Estimation. *Wind. Energy* **2016**, *19*, 133–150. [CrossRef]
18. Kalinić, H.; Ćatipović, L.; Matić, F. Optimal Sensor Placement Using Learning Models—A Mediterranean Case Study. *Remote Sens.* **2022**, *14*, 2989. [CrossRef]
19. Berkooz, G.; Holmes, P.; Lumley, J.L. The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows. *Annu. Rev. Fluid. Mech.* **1993**, *25*, 539–575. [CrossRef]
20. Everson, R.; Sirovich, L. Karhunen–Loève Procedure for Gappy Data. *J. Opt. Soc. Am. A* **1995**, *12*, 1657–1664. [CrossRef]
21. Willcox, K. Unsteady Flow Sensing and Estimation via the Gappy Proper Orthogonal Decomposition. *Comput. Fluids* **2006**, *35*, 208–226. [CrossRef]
22. Drmač, Z.; Gugercin, S. A New Selection Operator for the Discrete Empirical Interpolation Method—Improved A Priori Error Bound and Extensions. *SIAM J. Sci. Comput.* **2016**, *38*, A631–A648. [CrossRef]
23. Manohar, K.; Brunton, B.W.; Kutz, J.N.; Brunton, S.L. Data-Driven Sparse Sensor Placement for Reconstruction: Demonstrating the Benefits of Exploiting Known Patterns. *IEEE Control Syst. Mag.* **2018**, *38*, 63–86. [CrossRef]
24. Joshi, S.; Boyd, S. Sensor Selection via Convex Optimization. *IEEE Trans. Signal Process.* **2009**, *57*, 451–462. [CrossRef]
25. Kidambi Sekar, A.P.; van Dooren, M.F.; Rott, A.; Kühn, M. Lower Order Description and Reconstruction of Sparse Scanning Lidar Measurements of Wind Turbine Inflow Using Proper Orthogonal Decomposition. *Remote Sens.* **2022**, *14*, 2681. [CrossRef]
26. Nav, F.M.; Snaiki, R.; Ding, F. A Hierarchical Machine Learning Framework for Real-Time Reconstruction of Urban Wind Fields from Sparse Sensor Networks. *Build. Environ.* **2026**, *302*, 114875. [CrossRef]
27. Zhou, Y.; Lin, C.; Kikumoto, H.; Ooka, R.; Cheng, S. Rooftop wind field reconstruction using sparse sensors: From deterministic to generative learning methods. *Build. Environ.* **2026**, *296*, 114480. [CrossRef]
28. Bergmeir, C.; Benítez, J.M. On the Use of Cross-Validation for Time Series Predictor Evaluation. *Inf. Sci.* **2012**, *191*, 192–213. [CrossRef]
29. Roberts, D.R.; Bahn, V.; Ciuti, S.; Boyce, M.S.; Elith, J.; Guillera-Arroita, G.; Hauenstein, S.; Lahoz-Monfort, J.J.; Schröder, B.; Thuiller, W.; et al. Cross-Validation Strategies for Data with Temporal, Spatial, Hierarchical, or Phylogenetic Structure. *Ecography* **2017**, *40*, 913–929. [CrossRef]
30. Sirovich, L. Turbulence and the Dynamics of Coherent Structures. *I. Coherent Struct. Q. Appl. Math.* **1987**, *45*, 561–571. [CrossRef]
31. Chaturantabut, S.; Sorensen, D.C. Nonlinear Model Reduction via Discrete Empirical Interpolation. *SIAM J. Sci. Comput.* **2010**, *32*, 2737–2764. [CrossRef]
32. Nakai, K.; Yamada, K.; Nagata, T.; Saito, Y.; Nonomura, T. Effect of Objective Function on Data-Driven Greedy Sparse Sensor Optimization. *IEEE Access* **2021**, *9*, 46731–46743. [CrossRef]
33. Fritsch, F.N.; Butland, J. A Method for Constructing Local Monotone Piecewise Cubic Interpolants. *SIAM J. Sci. Stat. Comput.* **1984**, *5*, 300–304. [CrossRef]
34. Robey, R.; Lundquist, J.K. Behavior and mechanisms of Doppler wind lidar error in varying stability regimes. *Atmos. Meas. Tech.* **2022**, *15*, 4585–4622. [CrossRef]
35. Gebauer, J.G.; Bell, T.M. A Flexible, Multi-Instrument Optimal Estimation Retrieval for Wind Profiles. *J. Atmos. Ocean. Technol.* **2024**, *41*, 605–620. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.