https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

A physical-consistency-based ensemble method for estimating boundary-layer height from radiosondes and its sensitivity to vertical resolution

Minquan Zhang[1,2,3], Yunjie Xia[4], Qinghui Sun[1,2,3], Disong Fu[2,3], Hongrong Shi[2,3], Xuebo Fan[4], Xiang'ao Xia[1,3]

[1]Laboratory of Middle Atmosphere and Global Environment Observation (LAGEO), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China
[2]State Key Laboratory of Atmospheric Environment and Extreme Meteorology, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, 100029, China
[3]College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing, 100049, China
[4] Beijing Meteorological Observation Center, China Meteorological Administration, Beijing, 100089, China

Correspondence to: Disong Fu (fudisong@mail.iap.ac.cn) and Hongrong Shi (shihrong@mail.iap.ac.cn)

Abstract. Planetary boundary layer height (PBLH) is an important parameter for understanding land–atmosphere interactions and for weather, climate, and air quality modeling. Radiosonde-based PBLH estimates can differ substantially across diagnostic variables and methods, particularly when stable boundary layers (SBLs) coexist with overlying residual layers (RLs). We develop an ensemble method that obtains candidate PBLHs from nine radiosonde-based algorithms and evaluates their physical support using six consistency scores. These scores assess thermodynamic structure, shear–buoyancy conditions, surface-layer similarity, temperature and humidity distributions, gradient contrasts, and wind structure. The resulting scores are combined with algorithm-specific quality and agreement among candidate heights to determine each candidate's contribution to the final PBLH estimate, while accounting for redundant information shared among related methods. Convective and neutral boundary layer heights are obtained by weighted averaging, while SBL heights are determined by first selecting a layer supported by multiple physical criteria and then averaging estimates within that layer. RL tops are detected separately. Applied to 12,266 radiosonde profiles at Beijing station during 2009–2023, the method achieves a retrieval rate of 79.41 % at approximately 5 m vertical spacing. Increasing the vertical grid spacing to 25 m raises the SBL retrieval rate from 66.55 % to 74.26 %, whereas further coarsening to 100 m reduces it to 14.48 %. At 100 m, retrieval rates for convective and neutral boundary layers remain above 99 %, but estimated heights shift upward overall. The method provides an interpretable framework for PBLH estimation under complex atmospheric conditions and highlights the importance of high-resolution observations for identifying shallow stable layers and distinguishing vertically separated structures. Such an approach could also strengthen the observational basis for understanding boundary layer processes and evaluating their representation in numerical models.

Keywords: planetary boundary layer height; high-resolution radiosondes; ensemble estimation; physical consistency scoring; vertical resolution

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

1 Introduction

The planetary boundary layer connects the surface to the free atmosphere and is the main region of momentum, heat, water vapor, and pollutant exchange (Sun et al., 2025). Planetary boundary layer height (PBLH) characterizes the vertical extent of surface influence and turbulent mixing. It is central to describing land–atmosphere interactions and developing weather, climate, and air quality models (Stull, 1988; Zhang et al., 2020). Boundary layer deepening or contraction also changes the volume available for vertical dilution of near-surface pollutants, directly affecting pollutant accumulation and dispersion (Su et al., 2018). Accurate characterization of PBLH and its variability therefore supports atmospheric process research, improved model performance, and the development of a global boundary layer observing system (Teixeira et al., 2025).

Near-surface thermal stability distinguishes the convective boundary layer (CBL), neutral boundary layer (NBL), and stable boundary layer (SBL) (Liu and Liang, 2010). Under typical clear-sky conditions, daytime surface heating drives convective mixing. After sunset, surface cooling stabilizes the near-surface atmosphere, while remnants of the daytime mixed layer may persist aloft as a residual layer (RL) (Smith and Carlin, 2024). Layer-top signatures vary with stratification, and heights inferred from thermodynamic, dynamic, and material distributions can differ (Kotthaus et al., 2023; Tan et al., 2025; Su et al., 2026). A pronounced change in a profile may therefore mark the top of the surface-connected boundary layer or a separate layer aloft. At night, identifying the SBL and RL tops separately distinguishes the depth of the near-surface stable layer from that of the overlying remnant mixed layer.

Radiosondes provide vertical profiles of temperature, humidity, pressure, and wind, supporting PBLH estimation through gradients, parcel ascent, the bulk Richardson number, surface-based inversions, and wind structure. However, these methods rely on different diagnostic variables and physical criteria, which may yield substantially different PBLH estimates for the same sounding. Seidel et al. (2010) compared multiple radiosonde-based methods and found systematic differences of several hundred meters, with relative humidity and potential temperature gradient methods generally producing higher estimates and the parcel method yielding lower values. Liu and Liang (2010) further highlighted the need to account for different physical mechanisms when diagnosing PBLH under varying stability conditions, particularly by considering both thermodynamic and wind-related criteria in SBLs. Li et al. (2021) compared four methods based on potential temperature gradients, relative humidity gradients, parcel ascent, and the bulk Richardson number at nine Chinese stations. They found that the bulk Richardson number method generally produced higher estimates under stable conditions, while the agreement among methods varied substantially with boundary layer stability. For example, the consistency rate of the potential temperature gradient method decreased from 83.96 % and 91.72 % under CBL and NBL conditions, respectively, to 58.35 % under SBL conditions. Similar discrepancies associated with different diagnostic variables and atmospheric structures have also been reported in remote-sensing studies (Wang et al., 2016; Wang et al., 2021; Duncan et al., 2022; Roldán-Henao et al., 2024). These findings indicate that individual diagnostic variables may capture different aspects of boundary layer structure, motivating the integration of complementary thermodynamic and dynamic information for more robust PBLH estimation.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Ensemble methods combine complementary estimates to improve PBLH retrieval. Zhang et al. (2022) merged daytime PBLH estimates using second-resolution radiosondes across China. Using high-resolution soundings collected at Beijing station, Chen et al. (2023) combined multiple methods, statistical threshold corrections, and clustering to reduce disagreement among individual estimates. Their ensemble achieved a success rate of 62.6 % against lidar-assisted manual interpretation, using a maximum deviation of 50 m from the reference height. This exceeded the 33.9 %–56.1 % achieved by four individual gradient methods. In remote sensing, Min et al. (2020) developed the Integrated System for Atmospheric Boundary Layer Height Estimation (ISABLE) using a ceilometer and microwave radiometer. Comparison with 148 radiosondes yielded a correlation coefficient of 0.72 and a root-mean-square error of 322 m, outperforming the individual methods considered. Differences in nocturnal stable-layer identification among instruments nevertheless affected the combined estimates. Smith and Carlin (2024) used fuzzy logic to combine complementary observations from multiple instruments. Recent studies also integrate remote sensing, surface meteorology, and radiation measurements using machine learning (Zhang et al., 2025; Caicedo et al., 2025). For ensemble retrievals from individual soundings, the remaining challenge is to assess the physical evidence supporting each height estimate and combine estimates of the same layer.

The multiple interfaces associated with SBL impose additional demands on fusion methods. A nocturnal SBL beneath an RL creates two distinct layer tops within the lower atmosphere. Virtual potential temperature can increase sharply near the surface, vary weakly through the RL, and increase again across an elevated capping inversion (Kotthaus et al., 2023). Wind profiles can exhibit similar vertical complexity, with strong shear within or near the shallow SBL and additional wind-speed transitions or shear maxima associated with elevated structures, such as low-level jets (Liu and Liang, 2010; Baas et al., 2009). Similarly, other variables used for PBLH identification may also exhibit multiple distinct features at different heights, causing individual diagnostic methods to respond to different vertical structures within the same sounding. Averaging these distinct candidate heights can place the combined estimate between recognizable layer boundaries. Chen et al. (2023) identified multilayer structures, common overestimation, and intermittent SBL turbulence as major challenges for ensemble methods. An ensemble method must therefore distinguish estimates associated with the near-surface stable layer from those associated with elevated structures before calculating a combined SBL height.

Vertical resolution directly affects the fine boundary layer structures resolved by radiosondes. High-resolution observations have supported multiyear height merging and continental-scale dataset development (Zhang et al., 2022; Guo et al., 2024). Long radiosonde records also support analyses of boundary layer changes and model comparisons (Yue et al., 2021). Seidel et al. (2010) compared high-resolution and standard-resolution reports from the same sounding at Koror. Between the surface and 4 km, the reports contained 127 and 19 observation levels, respectively. Relative humidity gradients yielded heights of 1114 and 3715 m, whereas specific humidity, potential temperature, and refractivity gradients all yielded 1114 m in both reports. Thus, resolution changes affect both the structures that can be resolved and the resulting heights, with impacts that depend on the method. Seidel et al. (2012) further noted that vertical-profile limitations have larger relative effects on shallow boundary layers. For ensemble methods with local physical scoring, coarser sampling can alter the

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

observations available around each estimated height and the agreement among methods. Retrieval rates and height changes therefore provide complementary measures of resolution effects.

To address discrepancies among PBLH estimates from different radiosonde-based methods, particularly when multiple vertical interfaces coexist under stable conditions, we develop an ensemble approach that combines physical consistency scoring with layer-aware fusion. We apply the method to 12,266 soundings collected at Beijing station during 2009–2023, with an original vertical spacing of approximately 5 m. The analysis evaluates retrieval rates and representative layered structures, then examines interannual, seasonal, and observation-time variations in PBLH and independently detected RL-top heights. Downsampling the same soundings quantifies how reduced resolution changes individual height estimates, physical scores, and final ensemble estimates. Section 2 describes the observations, method, and downsampling design. Section 3 presents retrieval performance, height distributions, and resolution responses. Section 4 summarizes the main conclusions.

2 Data and methods

2.1 Radiosonde observations and sample composition

We use second-resolution radiosonde observations from Beijing station (54511) during 2009–2023. The Beijing Meteorological Observatory is located north of the South Fifth Ring Road in Daxing District, Beijing, at an elevation of 31.5 m. Observations include air temperature, pressure, specific humidity, and horizontal wind components. The original vertical spacing is approximately 5 m. Minor interpolation between neighboring observation heights regularizes the profiles onto a 5 m grid. This grid derives from dense original observations rather than interpolation from sparse mandatory pressure levels. The analysis covers 20–3,500 m, and all heights are above ground level (AGL).

Original UTC timestamps are converted to Beijing Time (BJT, UTC+8). Nominal operational times of 08, 14, and 20 BJT correspond to typical launch times of approximately 07:15, 13:15, and 19:15 BJT. Additional summer afternoon soundings began around 2010. After selecting the study years and removing duplicate observation times, the dataset contains 12,266 profiles. These comprise 5,466, 1,307, and 5,475 profiles at 08, 14, and 20 BJT, respectively, plus 18 at other times. The 14 BJT soundings predominantly represent enhanced warm-season observations: 1,223 of 1,307 profiles occur in June–August (93.6 %), with none in 2009.

2.2 Overall workflow and boundary layer classification

Profiles are first classified by near-surface thermal stability, and applicable algorithms are then used to identify possible layer-top heights (Fig. 1). Each height returned by an individual algorithm is termed a candidate height and is accompanied by an internal quality measure and structural attributes. Six physical consistency scores evaluate the observed evidence supporting each candidate. CBLs and NBLs use direct weighted averaging of applicable primary candidates. For SBLs, nearby candidates are grouped and their physical support is evaluated to select a layer before averaging the primary

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

candidates within it. Once a valid SBL height is established, an independent search identifies the overlying RL, yielding separate SBL-top and RL-top heights.

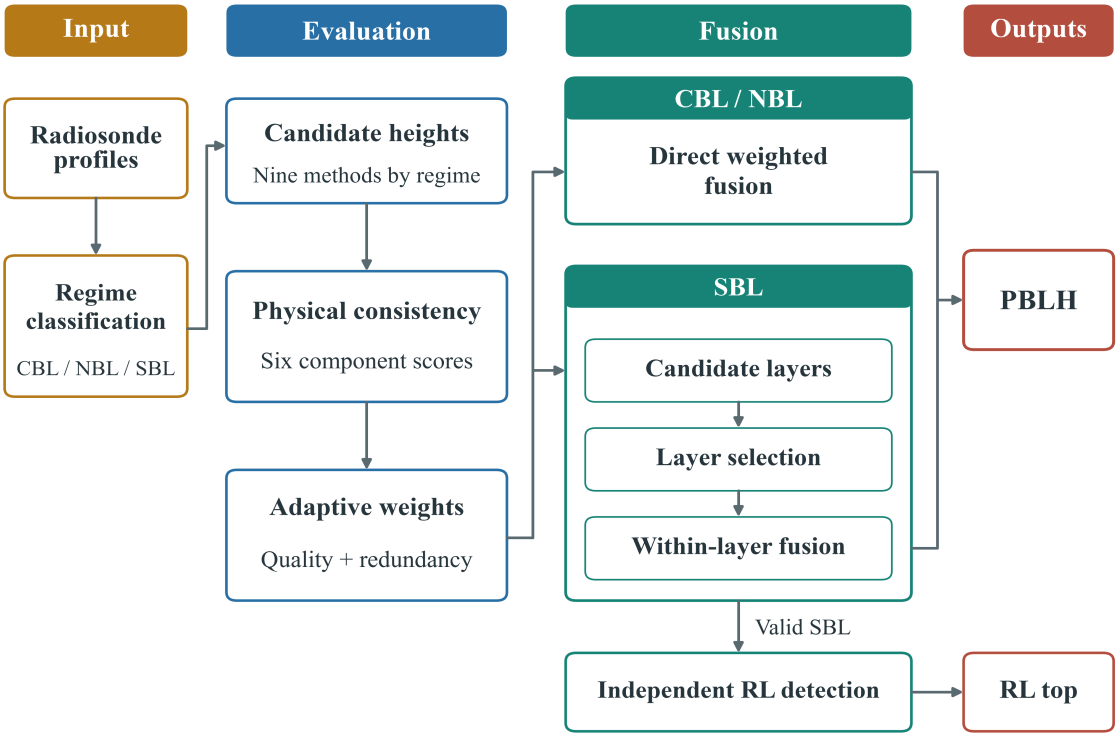


Figure 1. Workflow of the ensemble method.

Boundary layer classification uses virtual potential temperature to characterize the thermal stratification of moist air:

$$\theta_v = T\left(\frac{p_0}{p}\right)^{\kappa_d}(1 + 0.61q), \qquad p_0 = 1000\ \text{hPa}, \quad \kappa_d = 0.286. \qquad (1)$$

Here, T is absolute temperature, p is pressure, and q is specific humidity in kilograms per kilogram. The near-surface virtual potential temperature gradient is estimated using the Theil–Sen slope over 20–100 m. Slopes below −0.002 K m^{-1} indicate CBLs, slopes above 0.002 K m^{-1} indicate SBLs, and all remaining cases are NBLs. Organizing height diagnosis by near-surface stability follows Liu and Liang (2010) and Li et al. (2021). The height window and thresholds are specified for this study.

Conditional reassessment addresses two groups susceptible to shallow near-surface structures. Profiles initially classified as SBLs at 14 BJT are reassessed over 20–200 m using the same thresholds, better representing the broader afternoon low-level stratification. Initially neutral profiles at routine observation times are reassessed when solar elevation does not exceed 15°. Reclassification as SBL requires a surface-based air-temperature inversion and a shallow layer containing that inversion candidate. The layer must have support from at least two primary-candidate physical categories, satisfy the layer-dominance

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

criterion, and lie at or below 300 m. Other profiles retain NBL classification and direct fusion. Appendix A provides the thresholds and low-level data checks.

Each algorithm uses only valid levels of its required variables. Gradients are calculated using height-defined local smoothing and finite differences, without bridging gaps exceeding the permitted spacing. Routine smoothing on the native grid uses a second-order Savitzky–Golay filter (Savitzky and Golay, 1964), with diagnostic-specific scales retained. Local wind spikes and widespread anomalies are treated separately. Local anomalies invalidate only wind information used by affected candidates or scores; widespread anomalies invalidate wind-dependent diagnostics. Original wind values are not replaced. Appendix A specifies the windows, actual smoothing spans, and quality control rules.

2.3 Candidate height algorithms

Nine algorithms extract candidate heights from distinct layer-top signatures in temperature, humidity, and wind profiles. Let h_j denote candidate j and $Q_j \in [0,1]$ its internal structural quality. Depending on the algorithm, Q_j reflects peak prominence, reference-layer support, inversion depth and warming, inflection persistence, or jet completeness. Table 1 summarizes the diagnostic variables, structural criteria, and candidate roles; Appendix B (Table B1) provides the quality mappings. Primary candidates contribute to final height weighting. Diagnostic candidates assist structural identification and can contribute to SBL layer-support scores.

Table 1. Overview of the nine candidate algorithms.

ID	Variable or method	Structural criterion	CBL	NBL	SBL
A1	Virtual potential temperature gradient	Prominent positive-gradient peak; additional stable-layer termination transition retained for SBLs	P	P	D
A2	Specific humidity gradient	Prominent drying-gradient peak	P	P	D
A3	Bulk Richardson number	First persistent threshold crossing of the buoyancy-to-wind-difference ratio	P	P	P
A4	Parcel method	First intersection of near-surface reference and environmental virtual potential temperature	P	—	—
A5	Surface-based air-temperature inversion	Top of a continuous surface-based inversion	—	—	P
A6	Haar wavelet of virtual potential temperature	Prominent response to a positive jump across multiple scales	P	P	—
A7	Maximum vector wind shear	Midpoint of the plateau in maximum shear magnitude	P	P	P

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

ID	Variable or method	Structural criterion	CBL	NBL	SBL
A8	Wind-speed inflection	Curvature sign reversal with persistent structure on both sides	D	P	P
A9	Low-level jet	Lowest jet core satisfying core-speed and two-sided speed-decrease criteria	—	D	P

Note: P denotes a primary candidate, D denotes a diagnostic candidate, and — denotes not applicable. SBL primary candidates comprise thermodynamic (A5), combined thermodynamic–dynamic (A3), and dynamic (A7, A8, and A9) categories.

A1 and A2 identify positive virtual potential temperature gradient peaks and negative specific humidity gradient peaks, respectively. Peak prominence selects interfaces with distinct warming or drying signals. For SBLs, A1 additionally records stable-layer termination transitions as thermodynamic evidence for layer selection.

A3 characterizes the relative effects of buoyancy and wind shear using the bulk Richardson number:

$$\mathrm{Ri_b}(z) = \frac{g}{\overline{\theta}_{\mathrm{v,r}}} \frac{[\theta_{\mathrm{v}}(z) - \overline{\theta}_{\mathrm{v,r}}](z - \overline{z}_{\mathrm{r}})}{[u(z) - \overline{u}_{\mathrm{r}}]^2 + [v(z) - \overline{v}_{\mathrm{r}}]^2}. \qquad (2)$$

Here, g is gravitational acceleration, u and v are horizontal wind components, and subscript r denotes the near-surface reference layer. The reference layers are 20–100 m for CBL/NBL and 20–50 m for SBL, with critical values of 0.25 and 0.35, respectively. Searching upward from the reference layer identifies the first threshold exceedance lasting at least 15 m. The height is determined by linear interpolation between adjacent valid levels. Locations with squared wind differences below 0.01 $m^2\ s^{-2}$ are excluded. The reference layer and critical value govern the response to different stratification regimes (Vogelezang and Holtslag, 1996; Seidel et al., 2012).

A4 uses the mean virtual potential temperature over 20–50 m plus 0.05 K as the near-surface parcel reference. The first intersection with the environmental virtual potential temperature profile defines the candidate height. This reference supports mixed-layer diagnosis from individual profiles and follows the parcel concept used to estimate maximum mixing depth from daily maximum surface temperature (Holzworth, 1964). A5 identifies surface-based inversions using air temperature. The inversion must extend continuously from the lowest analysis level, span at least 20 m, and produce cumulative warming of at least 0.2 K. Its termination height provides the thermodynamic SBL candidate.

A6 uses a Haar response defined as the upper-half-window mean minus the lower-half-window mean:

$$W(a,h) = \frac{2}{a}\left[\int_{\mathrm{h}}^{h+a/2} \theta_{\mathrm{v}}(z)\,\mathrm{d}z - \int_{\mathrm{h-a/2}}^{h} \theta_{\mathrm{v}}(z)\,\mathrm{d}z\right]. \qquad (3)$$

Here, a is wavelet scale, h is window-center height, and z is the vertical coordinate. Scales of 50, 100, 200, and 350 m are evaluated only where both window halves have complete observational support. Positive peaks require a response of at least 0.1 K and prominence of at least 0.05 K. Qualifying peaks from all four scales are pooled, and the height of the most

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

prominent peak becomes the A6 candidate. Ties favor the smaller scale. This selects the interface with the strongest local contrast among jumps of different thicknesses, rather than averaging peak heights across scales. The zero-mean kernel removes constant background values, while thickness normalization permits comparison across scales. Following the multiscale structural approach of Davis et al. (2000) and Brooks (2003), we apply Haar responses to positive virtual potential temperature jumps.

A7, A8, and A9 identify maximum vector wind shear, wind-speed inflections, and low-level jets, respectively. A7 uses the midpoint of the maximum-shear plateau, with quality determined by prominence relative to the local background and proximity to search boundaries. A8 requires a sign reversal in the second wind-speed derivative across finite-depth regions, supported by first-derivative and wind-speed-change criteria. The selected inflection maximizes the product of first-derivative magnitude and curvature change. Its quality combines two-sided sign consistency and normalized wind-speed change. A9 selects the lowest complete jet core, requiring core speed of at least 8 m s^{-1} and decreases of at least 2 m s^{-1} both above and below. Low-level jets represent important nocturnal boundary layer dynamics (Baas et al., 2009). Together with shear and inflections, they complement temperature and humidity diagnostics.

2.4 Physical consistency scoring

The six scores assess complementary aspects of the structure supporting each candidate height h_j (Table 2). Each $C_{ij} \in [0,1]$ maps a thermal, dynamic, or distributional feature to a common support scale. Most scores compare local windows around the candidate; C_{SIM} instead combines a near-surface profile fit with the candidate's h/L. Higher values indicate stronger support under the specified mapping, rather than a probability that the height is correct. Three quantities have distinct roles: Q_j describes the quality of the structure identified by the candidate algorithm, C_{ij} evaluates the associated physical evidence, and q_{ij} describes the observational support for that evaluation. The computability indicator a_{ij} excludes scores with insufficient support; these are missing rather than zero. Section 2.5 combines the scores and quality measures while reducing shared information.

Table 2. Overview of the six physical consistency scores.

Symbol	Name	Information used	Structure receiving higher scores
C_E	Thermodynamic structure	Virtual potential temperature θ_v: median gradients and temperature differences between the lower and upper windows	CBL: weak gradient below and warming above; SBL: stable gradient weakening upward; NBL: weak gradient below and a gradient change
C_{TKE}	Shear–buoyancy environment	Horizontal wind u, v and virtual potential temperature θ_v: S^2 and stable/unstable components of N^2	CBL: shear/unstable-stratification contrast and stable cap; SBL: shear contrast and stable stratification; NBL: shear contrast
C_{SIM}	Surface-layer similarity	Near-surface wind speed U and virtual potential temperature differences: joint fit and candidate h/L	Good profile fit together with h/L within the regime-specific scoring range

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Symbol	Name	Information used	Structure receiving higher scores
C_{ENT}	Distributional information entropy	Virtual potential temperature θ_v and specific humidity q: normalized entropies using common bins above and below	Greater upper-window entropy, averaged across temperature and humidity
C_{GD}	Local gradient sharpness	Virtual potential temperature gradient: cross-interface change relative to gradient variability in a guarded local background	A gradient change that stands out from surrounding variability
C_{WS}	Wind structure	Vector shear from horizontal wind u, v: local/background magnitude ratio and upper–lower vector difference	Locally enhanced shear and/or a change in shear magnitude or direction across the candidate

Note: All scores range from 0 to 1. Appendix C provides the complete mappings, observational support requirements, and baseline weights (Table C1).

The thermodynamic structure score C_E tests the thermal transition expected for each regime using windows extending from 25 to 200 m below and above the candidate, bounded by the available profile. For CBLs, it favors a weak lower-window gradient and an upward increase in virtual potential temperature, representing a mixed-layer-to-cap transition. For SBLs, it favors positive gradients below and weaker gradients above, identifying the termination of a stable layer. For NBLs, it combines a weak lower-window gradient with a change across the candidate. These are local structural tests, not averages over the entire boundary layer.

The shear–buoyancy environment score C_{TKE} uses

$$S^2 = \left(\frac{\partial u}{\partial z}\right)^2 + \left(\frac{\partial v}{\partial z}\right)^2, \qquad N^2 = \frac{g}{\theta_v}\frac{\partial \theta_v}{\partial z}. \qquad (4)$$

Here, S^2 is squared vector wind shear magnitude and N^2 is squared buoyancy frequency, both in s^{-2}. Motivated by shear production and buoyancy effects in the turbulent kinetic energy (TKE) budget, the score evaluates mean-profile environments associated with turbulence. For CBLs, it favors stronger combined shear and unstable-stratification signals below the candidate and stronger stable stratification above. For SBLs, it combines a decrease in shear across the candidate with the presence of stable stratification in the two windows. For NBLs, it uses the shear contrast alone. The resulting indicators are mapped to [0,1] as detailed in Appendix C2. The score summarizes how well the observed shear and stratification match these regime-specific patterns. It is based on gradients of mean wind and virtual potential temperature, so it represents environmental conditions for turbulence rather than turbulent energy or its production rate (Mahrt, 2014; Grachev et al., 2013).

The surface-layer similarity score C_{SIM} first estimates near-surface similarity parameters, then evaluates each candidate through its h/L. Wind speed and virtual potential temperature differences are fitted jointly using integrated Monin–Obukhov

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

similarity theory (MOST). Only friction velocity u_* and $\lambda = 1/L$ are free parameters; the virtual potential temperature scale is derived from them:

$$U(z) = \frac{u_*}{\kappa}\left[\ln\frac{z}{z_0} - \psi_{\mathrm{m}}(z\lambda) + \psi_{\mathrm{m}}(z_0\lambda)\right], \qquad (5)$$

$$\theta_{\mathrm{v}}(z) - \theta_{\mathrm{v}}(z_{\mathrm{r}}) = \frac{\theta_{\mathrm{v}*}}{\kappa}\left[\ln\frac{z}{z_{\mathrm{r}}} - \psi_{\mathrm{h}}(z\lambda) + \psi_{\mathrm{h}}(z_{\mathrm{r}}\lambda)\right], \qquad \theta_{\mathrm{v}*} = \frac{u_*^2\overline{\theta}_{\mathrm{v}}\lambda}{\kappa g}. \qquad (6)$$

Here, $\kappa = 0.4$ is the von Kármán constant, z_0 is 0.1 m, and zero-plane displacement is zero. The functions ψ_{m} and ψ_{h} are integrated stability corrections for momentum and heat. Joint robust fitting normalizes residuals by 1 m s^{-1} for wind speed and 0.5 K for virtual potential temperature. Fitting uses 20–100 m, with one extension to 20–200 m after failure. Following residual, parameter-boundary, and condition-number checks, the same accepted fit supplies L and fit quality for the profile's candidates. The score is the geometric mean of candidate-specific h/L support and fit quality. Thus, near-surface similarity and candidate-height compatibility both contribute. Integrated relations follow Businger et al. (1971), Paulson (1970), and Dyer (1974), with stable functions from Cheng and Brutsaert (2005). Equation (6) links the fitted parameters and temperature scale; Sect. 2.5 reduces shared-variable redundancy (Baas et al., 2006).

The distributional information entropy score C_{ENT} measures how virtual potential temperature and specific humidity occupy common value bins in the lower and upper windows. Normalized Shannon entropy uses layer-thickness weighting (Shannon, 1948). Concentration in a few bins gives lower entropy, whereas more even occupancy gives higher entropy. The score favors an upward increase in entropy, averaged across the two variables. It supplies distributional evidence complementary to signed gradients, with the same directional mapping used for all three regimes.

C_{GD} tests whether a gradient change stands out from surrounding variability. It compares median virtual potential temperature gradients within 25 m on either side of the candidate and scales their absolute difference by robust background variability 50–250 m away on both sides. The intervening 25–50 m buffer zones limit contamination by the interface. Larger contrasts receive higher scores regardless of sign; C_{E} separately assesses the regime-specific thermal pattern.

C_{WS} combines two wind-structure signals: locally enhanced vector shear relative to its background, and a change in shear magnitude or direction between the lower and upper background regions. Each component increases with its respective contrast, and their mean gives the score. Unlike C_{TKE}, this score emphasizes local shear prominence and vector structure rather than regime-dependent shear–buoyancy conditions. Appendix C provides all mappings, support requirements, and empirical scales.

2.5 Score integration and ensemble height estimation

Score weights combine regime-specific baseline weights b_i, computability a_{ij}, support quality q_{ij}, and redundancy reduction ρ_{ij}:

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

$$r_{ij} = b_i a_{ij} q_{ij} \rho_{ij}, \qquad P_j = Q_j \sum_i \omega_{ij} C_{ij}. \qquad (7)$$

The values r_{ij} are first normalized across available scores, then subject to physical-family caps to obtain ω_{ij}. The thermodynamic family, comprising C_E, C_{ENT}, and C_{GD}, has a total weight cap of 0.50. The dynamic family, C_{TKE} and C_{WS}, has a cap of 0.40, and the similarity family, C_{SIM}, has a cap of 0.20. Excess weight is redistributed proportionally to other available families with remaining capacity. If available family caps sum to less than 1, the corresponding weight deficit is retained. Scores directly sharing a criterion with candidate generation receive a reduction factor of 0.5; all others receive 1. Multiplication by internal candidate quality Q_j then gives candidate support P_j.

CBLs and NBLs require at least two primary candidates with positive support. Their fused height is calculated as

$$w_j = \frac{P_j^2 \exp[-(h_j - m)^2/(2s^2)]}{\sum_k P_k^2 \exp[-(h_k - m)^2/(2s^2)]}, \qquad h = \sum_j w_j h_j. \qquad (8)$$

Here, m is the median candidate height, s = max(1.4826 MAD, 50 m), and MAD is the median absolute deviation. The exponential term represents height consensus, while P_j^2 represents physical structural support. Together, they determine each primary candidate's contribution to the fused height.

For SBLs, primary candidates are first sorted by height and grouped into candidate layers spanning no more than 150 m. Each layer must contain at least two primary candidates representing at least two physical categories: thermodynamic (A5), combined thermodynamic–dynamic (A3), and dynamic (A7, A8, A9). This requirement provides joint support from different physical information. Diagnostic candidates A1 and A2 can be attached to layers whose centers lie within 150 m. The layer-support score sums the maximum squared candidate support within each category:

$$F_\ell = \sum_f \max_{j \in \mathcal{L}_\ell \cap f} P_j^2. \qquad (9)$$

Here, ℓ indexes candidate layers, L_ℓ is the candidate set in layer ℓ, and f denotes physical category. Taking the maximum within each category controls redundant responses from similar methods. A single qualifying layer is selected directly. With multiple qualifying layers, the highest score must be at least 1.25 times the second highest; otherwise, no layer is selected. Primary candidates in the selected layer are combined using Eq. (8). Here, m and MAD become the P_j^2-weighted median and median absolute deviation. Selecting a layer before averaging prevents estimates of vertically separated structures from being combined into a single SBL height.

The spread of the ensemble height estimates is defined as

$$\sigma_h = \left[\sum_j w_j (h_j - h)^2\right]^{1/2}. \qquad (10)$$

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

where σ_h is the weighted dispersion of participating candidate heights around the final height, describing their agreement.

2.6 Residual layer identification

Under stable stratification, a near-surface SBL and an overlying RL can coexist with distinct top heights. The capping inversion above an RL can identify its top (Kotthaus et al., 2023). We therefore report the fused SBL-top height as the current PBLH and the independently detected RL-top height as a separate structural parameter. This distinguishes near-surface stable-layer depth from the vertical extent of the overlying residual layer.

RL detection is performed only when a valid SBL height is available. The complete virtual potential temperature profile over 20–3,500 m is first smoothed using an actual span of 100 m. Searching then begins 50 m above the SBL top, so changes in the search boundary do not alter smoothing at the full-profile boundary. A candidate RL must contain a near-neutral core at least 200 m thick, with absolute virtual potential temperature gradients no greater than 0.0015 K m^{-1}. Within 300 m above the core, a stable cap must have a median gradient of at least 0.002 K m^{-1} over at least 25 m. The cap must also lie at least 100 m above the SBL top. The search proceeds upward and records the first qualifying structure and its top height.

2.7 Vertical downsampling and algorithm adaptation

The same 12,266 profiles are sampled onto uniform grids of 25, 50, and 100 m and a surface-plus-mandatory-pressure-level grid. Uniform grids start at the lowest valid observation and select original observations near target heights. Mandatory levels are 1,000, 925, 850, 700, and 500 hPa, where they fall within the valid observation range. Values are interpolated in log-pressure coordinates, with the surface point retained. All levels remain subject to the 20–3,500 m analysis range. The four groups yield 49,064 coarse-grid records, all retaining the 5 m boundary layer classifications. This enables comparison of candidate diagnosis and height fusion responses to reduced resolution.

Coarse grids retain the main physical thresholds and fusion architecture. Minimum observation counts, finite-difference spacing, and finite-depth support for A8 are adapted to grid spacing (Appendix D, Table D1). On uniform coarse grids, routine smoothing requires at least three points and uses a first-order polynomial when only three points are used. A5 retains a second-order filter with at least five points, giving actual smoothing spans of 100, 200, and 400 m on complete 25, 50, and 100 m grids, respectively. RL smoothing spans 100 m on the 25 and 50 m grids and 200 m on the 100 m grid, where at least three points are required. If sequential SBL clustering yields no qualifying layer, complete primary-candidate windows spanning 150 m are enumerated to identify combinations omitted by sequential partitioning. Existing qualifying layers retain the original selection rule. Overlapping windows that share primary candidates and do not identify a unique layer are recorded as unresolved. Mandatory-level records retain individual algorithm candidates and subsequent scores; RL detection is restricted to uniform grids.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

3 Results and analysis

3.1 Overall results and representative cases

3.1.1 Retrieval rates and algorithm availability

Table 3 lists annual observation and retrieval counts. Annual inputs at 08 and 20 BJT each range from approximately 362 to 367 profiles, providing stable coverage throughout the study period. The 14 BJT sample varies with enhanced warm-season sounding schedules. The 18 profiles at other times are included in overall retrieval statistics, while height distributions are analyzed separately for the three main observation times.

Table 3. Annual sample counts during 2009–2023.

Year	08 BJT	14 BJT	20 BJT	Other times	All times
2009	364/304	0/0	365/296	0/0	729/600
2010	364/279	82/81	364/265	0/0	810/625
2011	363/294	77/77	365/265	5/3	810/639
2012	366/299	78/78	365/287	3/3	812/667
2013	365/301	92/92	365/288	0/0	822/681
2014	365/264	87/87	365/262	0/0	817/613
2015	365/263	99/99	365/276	2/2	831/640
2016	366/273	92/90	366/249	2/1	826/613
2017	362/259	93/92	362/260	4/3	821/614
2018	365/265	92/91	364/282	0/0	821/638
2019	362/261	110/108	365/281	1/1	838/651
2020	366/302	95/94	366/276	0/0	827/672
2021	364/313	92/91	365/293	0/0	821/697
2022	364/304	123/123	367/293	1/0	855/720
2023	365/310	95/93	366/268	0/0	826/671
Total	5466/4291	1307/1296	5475/4141	18/13	12266/9741

Note: Each cell gives the number of soundings/PBLH retrievals. A value of 0/0 indicates no observations at that time in the given year. Years are calendar years.

The 5 m observations yield 9,741 PBLH estimates from 12,266 input profiles, an overall retrieval rate of 79.41 % (Table 4). CBL and NBL retrievals number 2,417 and 2,301, respectively. For SBLs, 5,023 of 7,548 profiles yield heights, a retrieval rate of 66.55 %. Among the 2,525 SBL profiles without a result, 2,501 (99.05 %) lack a qualifying candidate layer. The remaining 24 do not yield a selected layer because the highest layer-support score is less than 1.25 times the second highest.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

SBL retrieval therefore depends primarily on joint support from different candidate categories within the same layer; unresolved competition among multiple layers accounts for only a small fraction.

Table 4. PBLH retrievals and RL detections from 5 m observations.

Grouping	Group	Inputs	PBLH retrievals	Retrieval rate (%)	RL detections
By regime	CBL	2417	2417	100.00	Not performed
By regime	NBL	2301	2301	100.00	Not performed
By regime	SBL	7548	5023	66.55	3246
By time	08 BJT	5466	4291	78.50	1066
By time	14 BJT	1307	1296	99.16	2
By time	20 BJT	5475	4141	75.63	2173
By time	Other	18	13	72.22	5
Overall	All	12266	9741	79.41	3246

CBLs and NBLs have abundant candidate availability (Fig. 2a). Availability exceeds 95 % for A1, A2, A3, and A6, with A1 and A6 exceeding 99 %. A4 reaches 98.97 % for CBLs and A8 reaches 95.09 % for NBLs; A7 reaches 79.19 % and 78.66 %, respectively. For SBLs, A3 and A8 remain available in 96.41 % and 95.65 % of profiles, whereas A5 and A9 reach 68.19 % and 49.67 %. Diagnostic candidates A1 and A2 reach 99.48 % and 98.56 %. These contrasts reflect the different structural requirements of individual methods. Bulk Richardson numbers and wind-speed inflections apply to a broad range of profiles, whereas continuous surface-based inversions and complete jets require more specific structures. Complementarity among methods allows fusion to exploit valid information under different stratification conditions.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

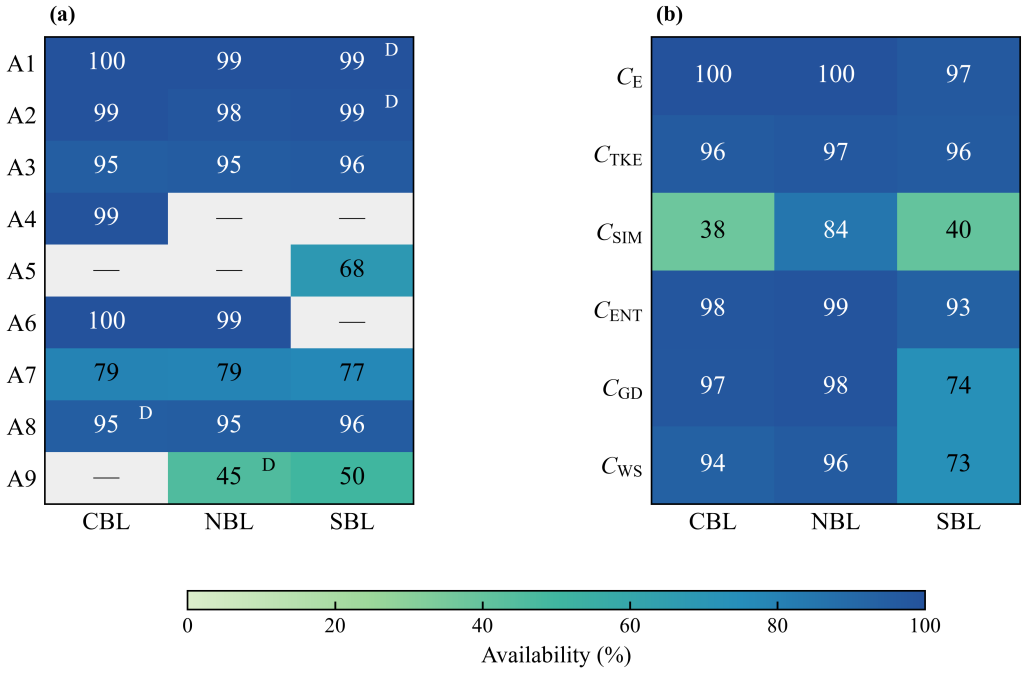


Figure 2. Candidate availability and score computability by boundary layer regime. (a) Availability of the nine candidate algorithms. (b) Computability of the six physical consistency scores. Candidate availability counts profiles with positive-quality candidates; score computability counts primary-candidate records with positive quality. D denotes diagnostic candidates; — denotes not applicable or no available result.

Score computability further distinguishes the regimes (Fig. 2b). In CBLs and NBLs, all scores except C_{SIM} exceed 93 % computability, providing broad support for candidate weighting. For SBLs, C_{GD} and C_{WS} reach 73.74 % and 73.24 %, respectively. Limited observational depth below shallow candidates makes local interface-to-background comparisons more difficult. C_{SIM} is computable for 37.59 %, 84.45 %, and 39.98 % of CBL, NBL, and SBL records, respectively. Near-neutral profiles more commonly support the integrated similarity relations. Weight integration across available scores accommodates these differences in observational support.

3.1.2 Representative convective and neutral boundary layers

At 14 BJT on 16 July 2019, virtual potential temperature is relatively uniform from near the surface to approximately 1.3 km, with stronger gradients above. Wind speed also changes slope near this height (Fig. 3a). A1, A3, and A4 cluster at 1,323–1,329 m, with final weights of 32.2 %, 31.7 %, and 36.0 %, respectively. Together, they determine a CBL height of 1,325.8 m (Fig. 3b, c). Specific humidity gradients, wavelet responses, and wind shear also identify higher interfaces. Fusion concentrates weight on the mixed-layer top jointly identified by thermodynamic gradients, parcel ascent, and the bulk Richardson number, preserving its distinct physical position.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

At 08 BJT on 31 May 2019, virtual potential temperature and specific humidity change near 0.4 km, while low-level wind speed increases with height (Fig. 3d). A1, A2, A3, and A6 cluster at 410–440 m. Their respective weights of 19.1 %, 32.8 %, 29.4 %, and 18.7 % yield a fused height of 424.9 m (Fig. 3e, f). A7 shear and A8 wind-speed inflection candidates lie at 1,460 and 2,740 m, respectively. Both are NBL primary candidates, but the height-consensus term reduces their final weights to nearly zero. These cases show how physical scoring combined with height consensus emphasizes a layer top supported by multiple criteria and moderates contributions from distant primary candidates.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

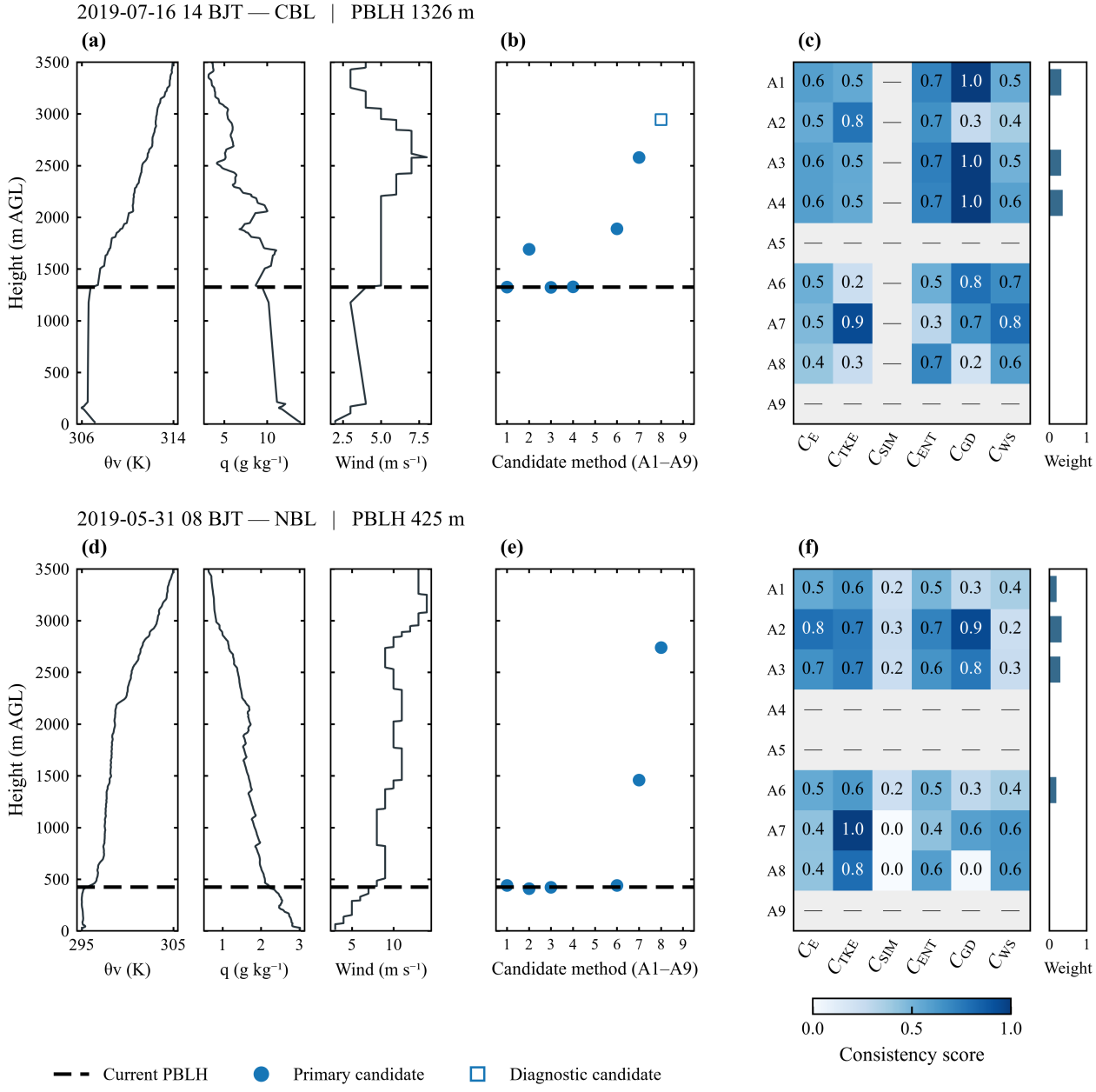


Figure 3. Height fusion for representative CBL and NBL cases. CBL at 14 BJT on 16 July 2019: (a) virtual potential temperature, specific humidity, and wind-speed profiles; (b) candidate heights; (c) physical scores and final weights. NBL at 08 BJT on 31 May 2019: (d) three-variable profiles; (e) candidate heights; (f) physical scores and final weights. Black horizontal lines indicate fused PBLH. Filled circles and open squares indicate primary and diagnostic candidates, respectively. Shading denotes scores, right-hand bars show normalized height weights, and — denotes missing scores.

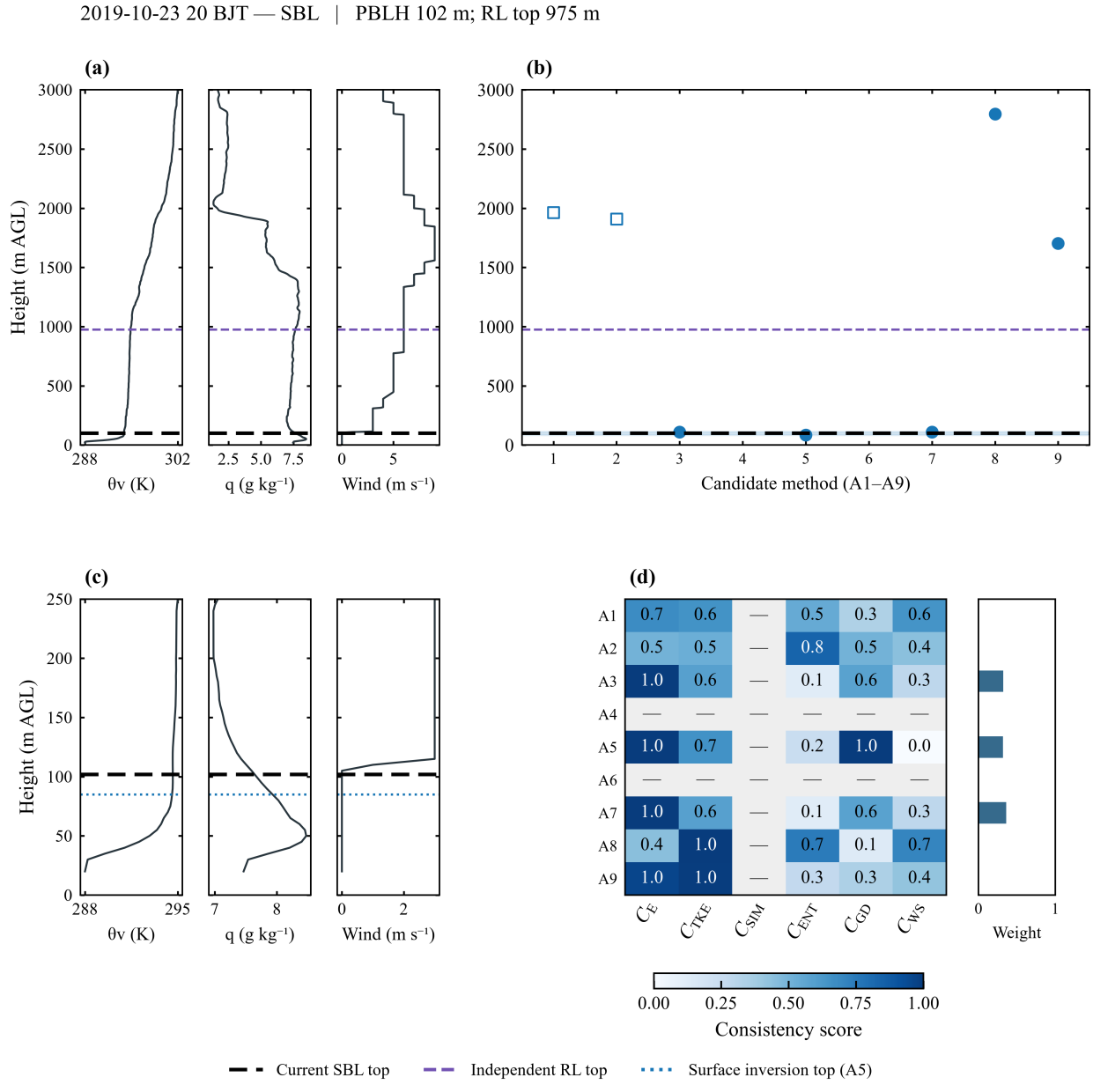
https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

3.1.3 A representative stable boundary layer and residual layer identification

The profile at 20 BJT on 23 October 2019 contains a shallow stable layer and distinct structures aloft (Fig. 4). A3, A5, and A7 yield 110, 85, and 110 m, respectively, forming the only qualifying layer within a 150 m span. They provide combined thermodynamic–dynamic, air-temperature inversion, and wind-shear support. Within-layer fusion yields 102.0 m. The A5 height of 85 m more closely represents surface-based inversion termination, whereas A3 and A7 are slightly higher. These differences have a clear physical interpretation within the same shallow layer and can be meaningfully combined through weighting.

The independently detected RL top is 975 m, approximately 873 m above the SBL top, producing a clearly separated two-layer structure. Low-level A3, A5, and A7 jointly provide buoyancy–shear, surface-based inversion, and wind-shear information, whereas higher A8 and A9 candidates primarily provide dynamic information. Layer selection therefore places the current SBL near 100 m. The overlying near-neutral layer and its cap then identify the RL. Separate shallow-SBL and deep-RL outputs preserve both the nocturnal near-surface stable layer and remnants of the earlier mixed layer.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.



380 **Figure 4. SBL fusion and RL identification at 20 BJT on 23 October 2019. (a) Full-height virtual potential temperature, specific humidity, and wind-speed profiles. (b) Candidate heights. (c) Low-level enlargement of the three profiles. (d) Physical scores and final weights. Black dashed, purple dashed, and blue dotted lines indicate the fused SBL top, independently detected RL top, and A5 air-temperature inversion top, respectively. Shading and bars follow Fig. 3.**

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

3.2 Distributions and variability of boundary layer and residual layer heights

3.2.1 Interannual variability

CBLs, NBLs, and SBLs exhibit distinct height scales throughout 2009–2023 (Fig. 5). At 08 BJT, annual median CBL heights range from 425.3 to 609.5 m, fluctuating around 500 m. NBL medians range from 554.6 to 1,156.5 m, with greater interannual variability. SBL medians remain between 68.6 and 114.8 m, consistently representing shallow near-surface layers. Near-surface stability classification thus captures persistent structural differences, with each regime exhibiting a different amplitude of interannual height variability.

At 20 BJT, annual median NBL heights range from 968.2 to 1,672.3 m and exceed their 08 BJT counterparts in every year. SBL medians at the same time range from 67.1 to 116.1 m. At both times, SBL medians peak in 2012 and subsequently decrease overall. Most years after 2014 lie near 70–85 m. Annual median CBL heights at 20 BJT vary from 501.0 to 1,408.6 m, a much wider range than at 08 BJT. Small annual sample sizes may partly explain this variability. The annual 20 BJT CBL sample ranges from only 2 to 41 profiles, allowing a few deep mixed-layer events to strongly influence the median.

Annual median RL-top heights range from 1,350 to 2,045 m at 08 BJT and from 1,130 to 1,515 m at 20 BJT, with morning values higher every year. The separation between shallow SBLs and kilometer-scale RLs persists throughout the study period, demonstrating a consistent statistical distinction between near-surface stable and overlying residual structures. Interannual fluctuations also differ among boundary layer regimes and RLs. This may reflect different responses to surface cooling, mechanical mixing, and previous convective mixing. SBLs more directly reflect near-surface stable-layer development, whereas RLs retain information about earlier structures in the deeper atmosphere.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

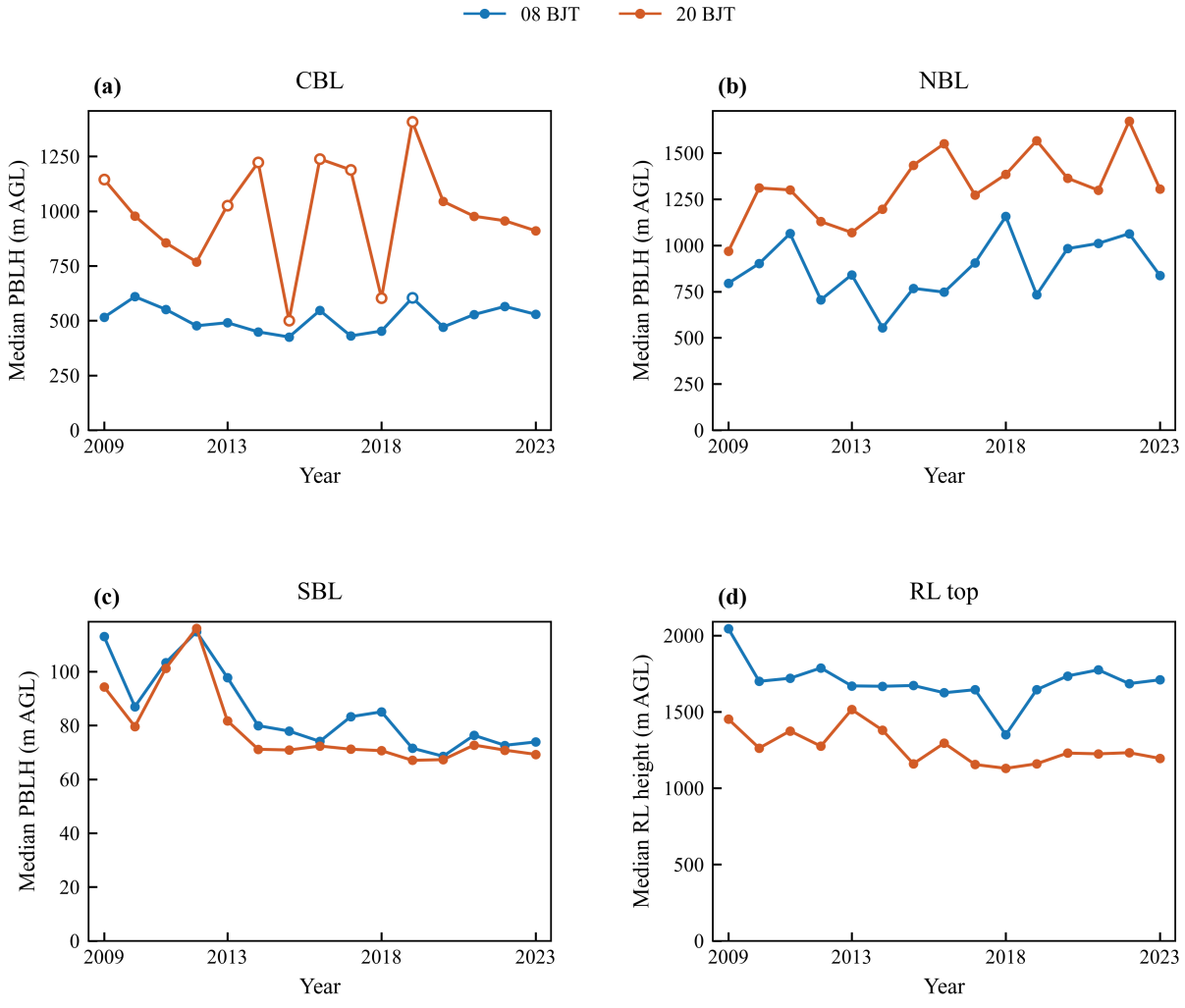


Figure 5. Annual median heights at 08 and 20 BJT. (a) CBL. (b) NBL. (c) SBL. (d) RL top. Blue and orange represent 08 and 20 BJT, respectively. Open symbols indicate fewer than 20 valid observations in the corresponding annual group.

3.2.2 Seasonal differences among boundary layer regimes

Seasonal statistics primarily use routine 08 and 20 BJT observations. Spring, summer, and autumn comprise March–May, June–August, and September–November, respectively. Winter combines December of the preceding year with January–February of the current year; complete winters during 2010–2023 are used. The lower row of Fig. 6 instead pools all seasons at each observation time. The 14 BJT sample predominantly comprises enhanced warm-season observations. Box plots include all valid heights and describe distributions using medians and interquartile ranges.

Pooling routine 08 and 20 BJT observations gives seasonal median SBL heights of approximately 74–80 m, far below CBL and NBL heights (Fig. 6a–c). SBL seasonality appears mainly in distribution width. The summer upper quartile exceeds its

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

winter counterpart, while winter boxes are more concentrated at shallow heights. Thus, typical SBL depths remain below 100 m throughout the year, whereas the occurrence of deeper stable layers varies seasonally.

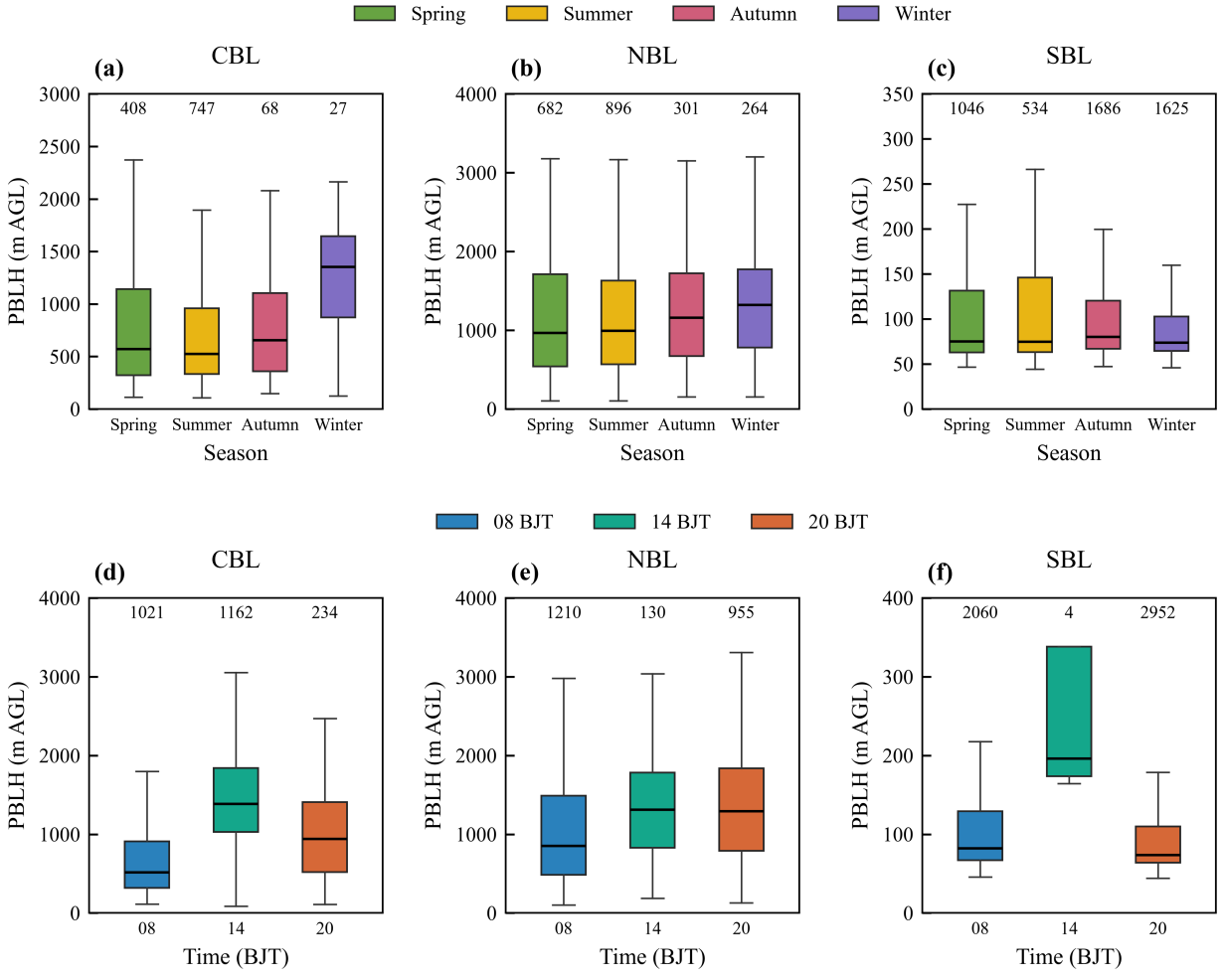


Figure 6. Seasonal and observation-time distributions of PBLH by regime. (a–c) Seasonal CBL, NBL, and SBL distributions, pooling 08 and 20 BJT. (d–f) Distributions at 08, 14, and 20 BJT, using all seasons at each time. Boxes show interquartile ranges, black lines show medians, and whiskers extend to the most distant observations within 1.5 interquartile ranges. Numbers indicate valid sample counts. All valid heights enter the box-plot statistics; outlier points are omitted. The 14 BJT sample predominantly represents the warm season.

Routine-time CBLs yield 408, 747, 68, and 27 valid results in spring, summer, autumn, and winter, with median heights of 571.4, 523.5, 654.4, and 1,353.1 m, respectively. CBLs occur mainly in spring and summer at routine observation times, but winter occurrences correspond to deeper mixed layers. Occurrence frequency and height conditional on occurrence therefore have different seasonal patterns. Median NBL heights also rise from 966.1 m in spring and 993.7 m in summer to 1,158.1 m in autumn and 1,320.3 m in winter. Seasonal median SBL heights vary much less. Deeper cold-season CBLs and NBLs may

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

reflect stronger mechanical mixing and different synoptic conditions. Near-surface stability grouping separates these structures from numerous shallow SBLs, revealing distinct vertical scales within the same season.

3.2.3 Differences among observation times

Height distributions across the three observation times differ sharply among regimes (Fig. 6d–f). Median CBL heights are 514.7, 1,385.7, and 939.4 m at 08, 14, and 20 BJT, respectively. The afternoon maximum and evening heights exceeding morning values are consistent with daytime mixed-layer growth and persistence of deeper mixed structures during evening decay. Median NBL heights are 852.6, 1,310.4, and 1,290.2 m, showing comparable afternoon and evening scales. SBL medians are 82.0 and 73.7 m at 08 and 20 BJT, respectively, concentrated near the surface. SBLs at 14 BJT are uncommon: only four valid heights are obtained, with a median of 196.2 m. Reduced surface heating under greater cloud cover offers one possible explanation for afternoon stable layers. Zhang et al. (2018) similarly found that approximately 70 % of afternoon SBLs in enhanced summer soundings across China occurred under overcast conditions.

Further seasonal grouping of routine observations shows higher NBL medians at 20 than at 08 BJT in every season. Evening versus morning medians are 1,240.8 versus 816.9 m in spring, 1,273.9 versus 732.7 m in summer, 1,386.3 versus 1,020.7 m in autumn, and 1,504.6 versus 1,150.4 m in winter (Fig. 7b). CBLs likewise have higher evening medians in all four seasons (Fig. 7a). This morning–evening contrast is consistent with persistence of deeper daytime mixed structures into the evening, supporting a continuing influence of earlier mixing on evening boundary layer structure.

SBLs show the opposite morning–evening contrast to CBLs and NBLs: 08 BJT medians exceed 20 BJT values in every season, most strongly in summer (Fig. 7c). Summer median and upper-quartile heights are 100.0 and 279.7 m at 08 BJT, compared with 72.1 and 122.6 m at 20 BJT. Morning SBLs span a wider height range and include more deep stable layers. This may reflect different developmental stages. Newly established evening stable layers are generally shallow, whereas overnight cooling and dynamic adjustment can produce thicker layers in profiles that remain stable in the morning.

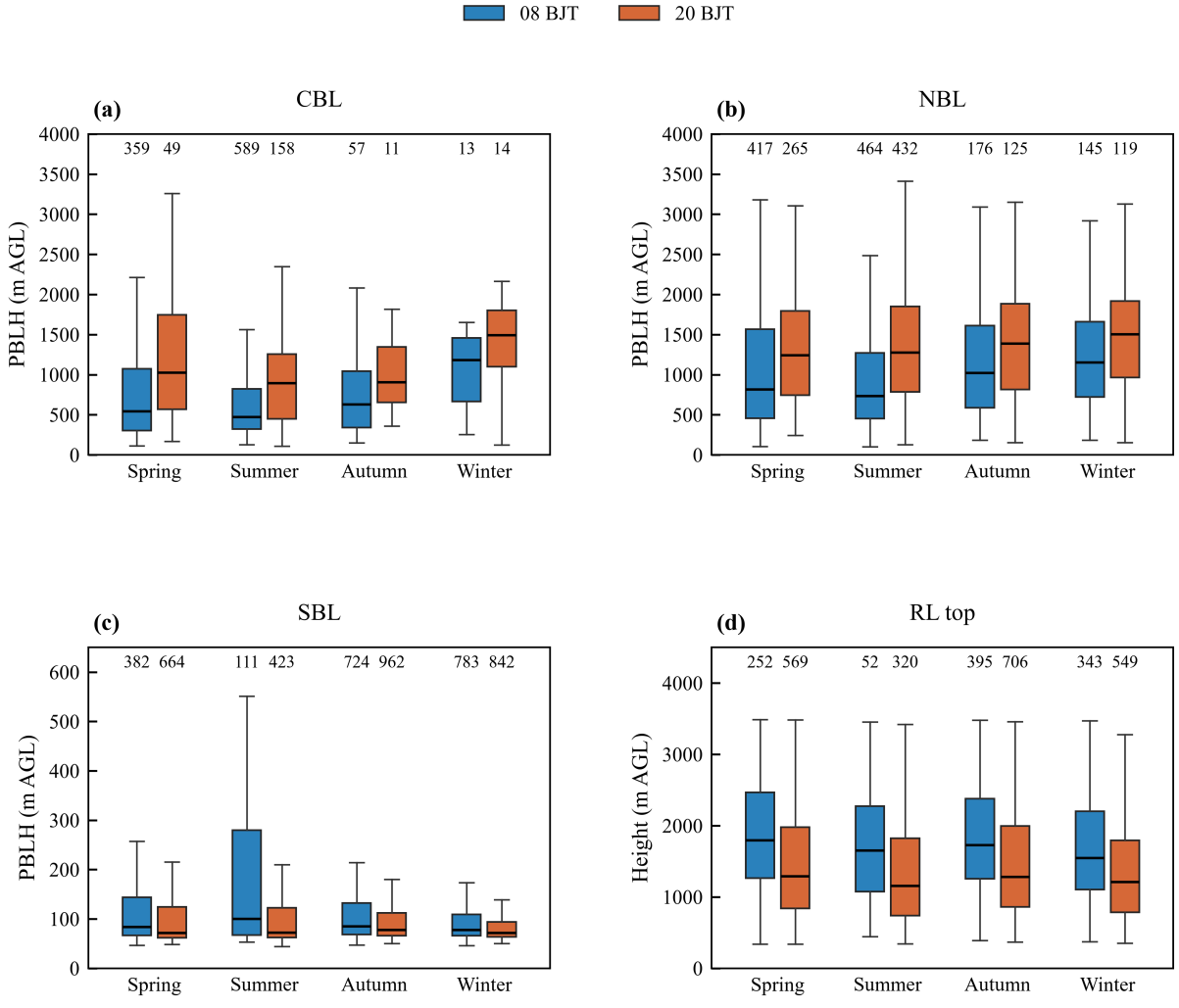
3.2.4 Residual layer heights and their relationship to stable boundary layers

The algorithm detects 3,246 RLs, equivalent to 64.62 % of the 5,023 valid SBL retrievals and 43.00 % of all 7,548 SBL input profiles. The former proportion describes RL detection after a successful SBL-height retrieval. The latter describes joint retrieval of both heights among all SBL inputs. Approximately two-thirds of valid SBL profiles also contain a detected RL, demonstrating broad coverage of multilayer structures through joint diagnosis. Routine-time observations yield 3,239 RLs, including 1,066 at 08 BJT and 2,173 at 20 BJT. Their median top heights are 1,700 and 1,250 m, respectively, far above the corresponding SBLs.

RL tops are higher at 08 than at 20 BJT in all seasons (Fig. 7d). Morning medians are 1,795, 1,650, 1,725, and 1,545 m in spring, summer, autumn, and winter, respectively. Corresponding evening medians are 1,290, 1,155, 1,282.5, and 1,210 m. Morning RL detections are less frequent but higher, showing that morning–evening differences involve both detection

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

frequency and vertical position. Separate SBL-top and RL-top statistics distinguish changes in near-surface stable-layer depth from changes in overlying residual structures, rather than conflating them as a single PBLH variation.



460 **Figure 7. Seasonal PBLH and RL-height distributions at 08 and 20 BJT. (a) CBL. (b) NBL. (c) SBL. (d) RL. Blue and orange boxes represent 08 and 20 BJT, respectively. Numbers indicate valid sample counts; box-plot definitions follow Fig. 6.**

3.3 Effects of reduced vertical resolution

3.3.1 Retrieval rates and retention of shallow stable layers

The 25, 50, and 100 m grids yield 10,323, 7,672, and 5,791 PBLH retrievals, corresponding to overall rates of 84.16 %,
465 62.55 %, and 47.21 %. Regime-specific rates use all input profiles in the corresponding regime as denominators. SBLs dominate changes in total retrieval counts (Fig. 8a). Their retrieval rate rises from 66.55 % at 5 m to 74.26 % at 25 m, then

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

falls to 39.14 % at 50 m and 14.48 % at 100 m. CBLs and NBLs retain complete retrievals at 25 and 50 m and reach 99.30 % and 99.87 %, respectively, at 100 m. Sampling at 50–100 m therefore weakens shallow stable-layer identification far more than CBL or NBL retrieval.

To track existing shallow-layer results, valid 5 m SBLs are grouped by height, and continued PBLH retrieval is assessed on each coarse grid (Fig. 8b). The 3,405 samples at or below 100 m represent 67.79 % of valid 5 m SBLs. Their retention rates are 91.89 %, 39.27 %, and 12.45 % at 25, 50, and 100 m, respectively. For the 100–200 m group, corresponding rates are 90.82 %, 53.47 %, and 20.16 %. At 100 m grid spacing, samples originally exceeding 500 m retain a rate of 28.04 %. Retention increases with original layer depth at both 50 and 100 m, identifying shallow SBLs as the structures most readily lost under coarse sampling. As layer depth approaches sampling spacing, observations rapidly lose the ability to resolve inversion tops, wind transitions, and backgrounds on both sides simultaneously.

The 25 m grid retains 4,531 valid 5 m SBL retrievals and adds 1,074 results from profiles without a 5 m retrieval, for a net increase of 582. Moderate coarsening changes local gradient fluctuations, actual smoothing spans, and candidate quality. Some profiles that previously lacked joint support consequently form qualifying layers, while 492 existing retrievals are lost. The initial increase and subsequent decline in retrieval counts reflect competing effects. Moderate coarsening can strengthen candidate consensus for some structures, whereas further coarsening makes shallow-layer information loss dominant.

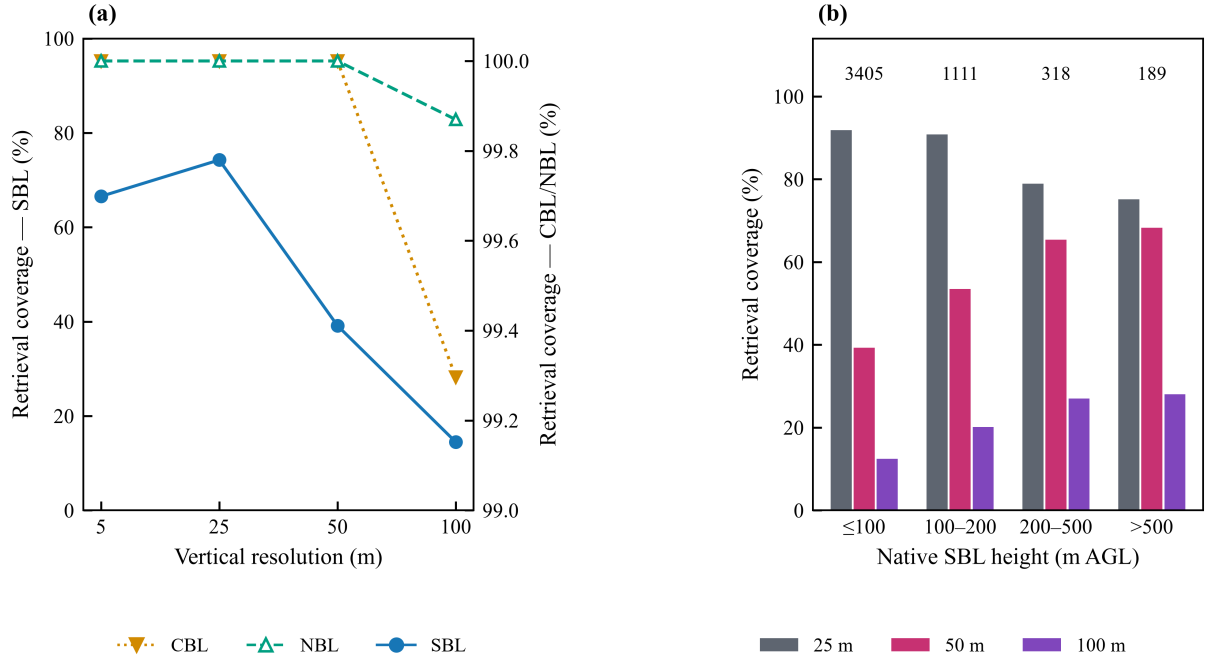


Figure 8. PBLH retrieval rates and SBL retention under different vertical resolutions. (a) Retrieval rates for CBL, NBL, and SBL; SBL uses the left axis and CBL/NBL the right axis. (b) Coarse-grid retention grouped by 5 m SBL height. Retrieval rates use all inputs within each regime as denominators. Retention rates use valid 5 m retrievals in each height group. All grids share the same boundary layer classifications.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Mandatory pressure levels provide some individual-algorithm heights, but observational support for local scoring becomes the bottleneck for complete fusion. A3 produces 1,946, 2,178, and 7,394 raw candidates for CBLs, NBLs, and SBLs, respectively; CBL A4 produces 2,383. None receives positive support in subsequent scoring, leaving no complete fusion retrievals. Thus, bulk Richardson number and parcel methods can estimate heights from sparse levels, whereas local physical assessment of interfaces and surrounding backgrounds requires denser observations.

3.3.2 Resolution response of fused heights

To compare heights for identical profiles, Fig. 9 uses 2,400 CBLs, 2,298 NBLs, and 531 SBLs with retrievals at all four resolutions. Differences are defined as coarse-grid height minus 5 m height. Median CBL differences are 35.4, 104.4, and 235.6 m at 25, 50, and 100 m, respectively; corresponding NBL differences are 17.3, 78.8, and 212.8 m. Both regimes shift upward overall as grids coarsen, with broader difference distributions. Together with retrieval rates above 99 %, these results show that the main CBL/NBL resolution response is a shift in height rather than widespread retrieval loss.

For the common SBL sample, median height differences are 21.6, 94.2, and 241.9 m at 25, 50, and 100 m, respectively. Median absolute differences are 36.8, 107.9, and 241.9 m. From 25 to 100 m grid spacing, the median upward shift increases by approximately an order of magnitude. Combined with sharply declining shallow-SBL retention, this reveals a dual effect of reduced resolution. Many shallow layers lose qualifying candidate support, while retained retrievals shift upward overall.

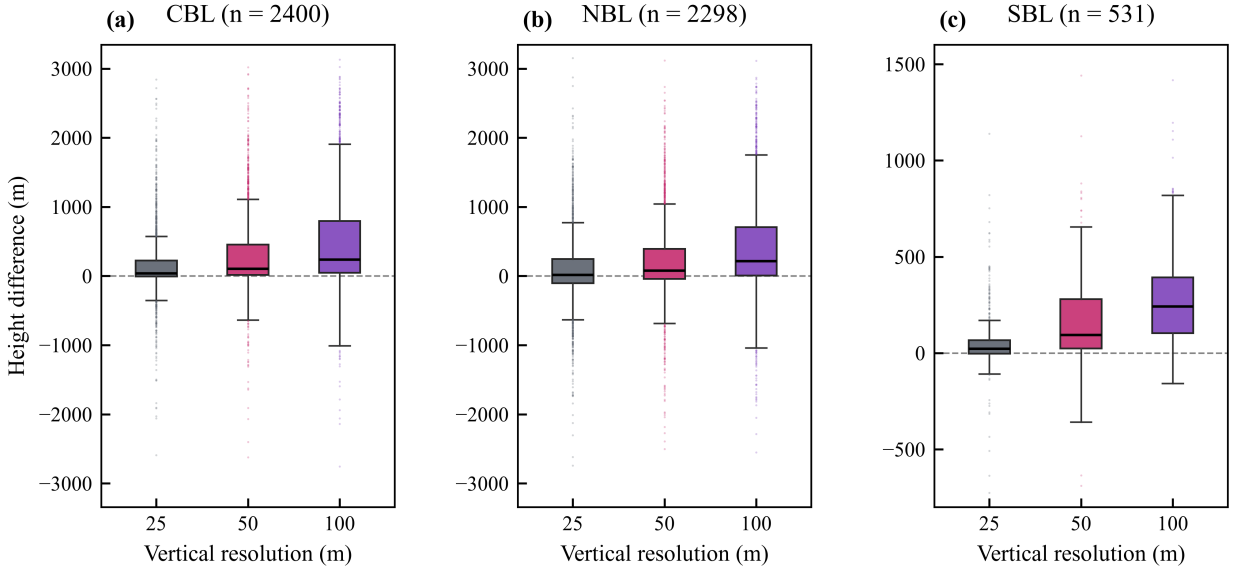


Figure 9. PBLH differences between coarse and 5 m grids for identical profiles. (a) CBL. (b) NBL. (c) SBL. Each regime uses the common set of profiles retrieved at all four resolutions; sample counts are shown. Differences are coarse-grid minus 5 m heights. Horizontal dashed lines indicate zero difference. Box-plot definitions follow Fig. 6; points show observations beyond the whiskers.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

This overall upward shift agrees with the high- and low-resolution radiosonde comparison of Seidel et al. (2010). Additional observation levels help capture parcel intersections and inversion transitions at lower heights, whereas sparse sampling can shift these features upward. Our regime-specific results further show that nearly complete CBL/NBL retrieval and sharply declining SBL retrieval can both accompany positive height responses. Different stratification regimes therefore respond differently to the same sampling change.

3.3.3 Resolution responses of candidate identification and physical scoring

Candidate availability responds to coarsening according to the structure diagnosed by each algorithm (Fig. 10a). SBL A5 availability falls from 68.19 % at 5 m to 64.73 %, 54.16 %, and 44.49 % at 25, 50, and 100 m, respectively. Diagnostic A2 availability decreases from 98.56 % to 91.12 %, 83.96 %, and 68.96 %, and A9 also declines. Surface-based inversions, local drying transitions, and complete jets thus depend more strongly on dense observations. In contrast, A3 and A7 each retain 99.75 % positive-quality candidate availability at 100 m, while A8 reaches 97.03 %. These methods preserve broad candidate availability, but their ability to combine with other evidence within the same layer weakens substantially.

Among 7,548 SBL profiles at 100 m, 3,230 lack qualifying layers because scorable primary candidates represent fewer than two categories. This group includes 957 profiles with fewer than two scorable primary candidates. Another 3,224 profiles contain different candidate categories, but their vertical separation exceeds the qualifying-layer span. Coarsening therefore causes losses through two main pathways: reduced observational support around candidates and increasing separation of interfaces identified by different methods. Increasing A7 availability alongside declining final SBL retrieval further demonstrates the distinction between a local shear extremum and convergence of multiple physical indicators at the same height.

Pairing raw candidates from the same profile and algorithm shows median A5 height differences increasing to 20, 60, and 145 m at 25, 50, and 100 m, respectively. A5 thus exhibits one of the strongest upward shifts among SBL candidates (Fig. 10b). At 100 m, the median A3 difference is approximately 64 m, A1 and A2 differences are approximately 40 m, and A7 differences are negative. Differences in both direction and magnitude show that diagnosed structures do not move uniformly under coarse sampling. Their changing relative positions subsequently alter the composition of within-layer candidates.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

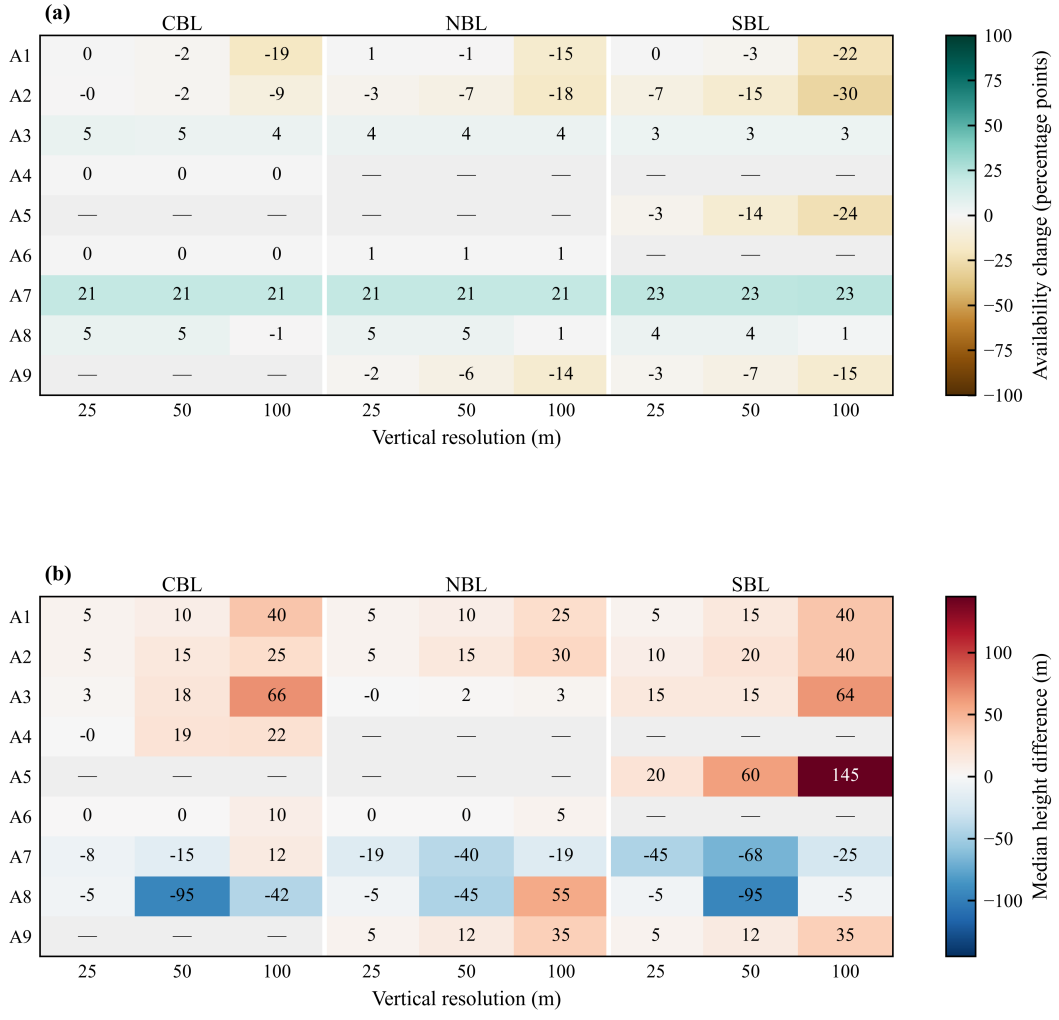


Figure 10. Resolution responses of candidate algorithms. (a) Changes in candidate availability relative to 5 m, in percentage points. (b) Median candidate height differences relative to 5 m, in meters. Within each boundary layer regime, columns represent 25, 50, and 100 m. Heights are paired using raw candidates from the same profile and algorithm. A dash (—) indicates not applicable or no paired results.

3.3.4 A shallow stable boundary layer case illustrating resolution effects

The case at 20 BJT on 25 January 2019 illustrates how these changes propagate into final heights (Fig. 11). On the 5 m grid, A3, A5, and A7 yield 55, 80, and 75 m, respectively. These three physical categories support the same shallow layer, producing a fused height of 70.0 m. At 25 m, all three candidates lie at 70 m, preserving the shallow-layer height. Weights nevertheless change: the final A5 weight increases from approximately 33.5 % to 85.0 %.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

At 50 m, A3 and A7 remain at 70 m, while A5 shifts to 120 m. The three candidates still occupy one qualifying layer, but reduced quality and scoring contributions from A3 and A7 raise the A5 weight to 99.84 %. The fused height consequently increases to 119.9 m. Weight concentration amplifies the influence of the upward-shifting inversion top on final PBLH, showing that candidate position and weighting jointly determine the resolution response.

At 100 m, only a few observations describe the low-level inversion, and A5 no longer provides a valid candidate. A7 shifts to 170 m and A3 to 562.8 m, while wind-structure candidates remain aloft. Vertical separation among scorable primary candidates leaves each candidate layer with only one physical category, eliminating the fused retrieval. This case traces a progression from shallow-height retention through dominant-weight concentration to qualifying-layer loss, illustrating how dense observations support shallow-SBL identification.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

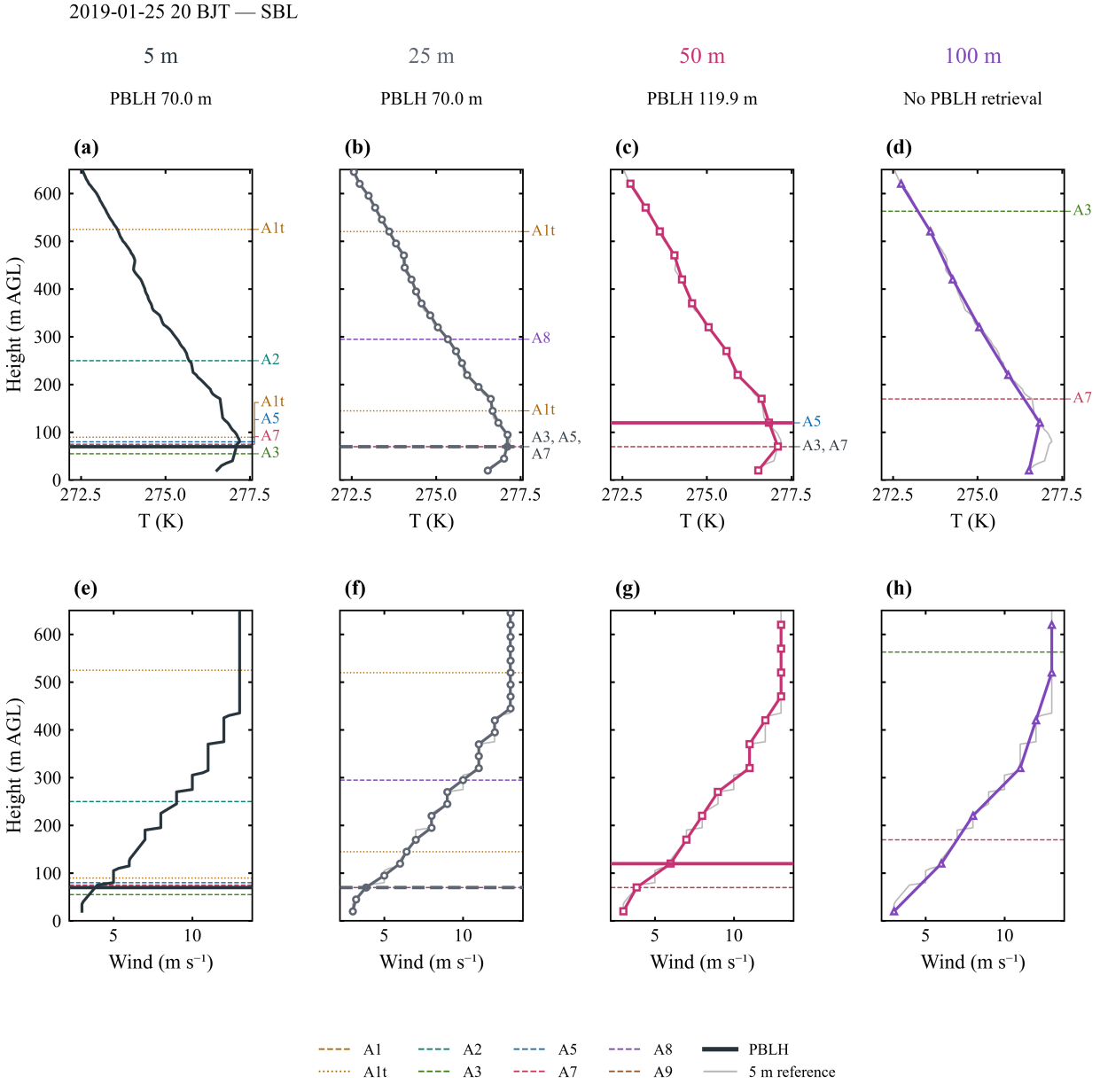


Figure 11. Downsampling of a shallow SBL at 20 BJT on 25 January 2019. (a–d) Air-temperature profiles. (e–h) Wind-speed profiles. Columns show 5, 25, 50, and 100 m resolution, respectively. Thin gray lines show 5 m reference profiles; symbols indicate coarse-grid observation points. Thin horizontal lines mark candidate heights, and thick horizontal lines mark fused PBLH. A1t denotes an additional thermodynamic transition. The plotted range is 0–650 m; Fig. E1 in Appendix E shows candidates over the full height range.

RL detections also initially increase and then decline: 3,246, 4,004, 2,051, and 593 RLs are identified at 5, 25, 50, and 100 m, respectively. Each grid searches independently above its final SBL height. RL counts therefore respond jointly to valid SBL availability, persistence of the near-neutral core, and resolution of the cap. At 25 m, more SBL retrievals and altered

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

representation of local structure increase RL detections. At 50–100 m, detections decline as observational support for near-surface layers and elevated caps weakens. Resolution effects thus extend across both near-surface SBL and overlying RL height diagnoses.

4 Conclusions

Using high-resolution radiosonde observations collected at Beijing station during 2009–2023, we developed an ensemble method that combines nine PBLH retrieval algorithms with six physical consistency scores. The method was applied to 12,266 soundings with approximately 5 m vertical spacing, and its performance and sensitivity to vertical resolution were examined. The main conclusions are as follows.

(1) The proposed method integrates candidate-level physical assessment with explicit vertical-layer selection. Six consistency scores evaluate the thermodynamic, dynamic, and vertical structural support for each candidate height. These scores are combined with candidate quality and height agreement to determine fusion weights, while accounting for overlapping information among related diagnostics. For CBLs and NBLs, physically supported candidates are combined through robust weighted averaging. For SBLs, candidate heights are first grouped into vertical layers, and layer selection requires support from multiple physical categories. Within-category support is consolidated to prevent several related methods from dominating the selection simply through their number. Only candidates within the selected layer contribute to the final SBL height. This framework addresses both the physical credibility of individual estimates and their attribution to distinct vertical structures, reducing the risk of averaging heights associated with different interfaces.

(2) The ensemble method provides PBLH estimates across different stability regimes while distinguishing shallow SBLs from elevated residual structures. The 5 m observations yield 9,741 PBLH estimates, corresponding to an overall retrieval rate of 79.41 %. All CBL and NBL profiles return results, while the SBL retrieval rate is 66.55 %. Representative cases demonstrate that the layer-selection strategy can identify shallow stable layers near 100 m without combining their heights with elevated candidates. A separate structural search identifies overlying RL tops near 1 km in these cases. Among SBL profiles without a valid result, 99.05 % fail to form a qualifying candidate layer, indicating that sufficient support from distinct physical categories within the same layer is the primary constraint on SBL retrieval.

(3) The retrieved heights exhibit distinct temporal variations across boundary layer regimes, with a clear separation between shallow SBLs and elevated RLs. Afternoon CBLs have a median height of 1,385.7 m. CBL and NBL median heights are consistently higher at 20 than at 08 BJT across all seasons, whereas SBLs show the opposite pattern. Seasonal SBL medians remain within 74–80 m, with broader height distributions during summer mornings. The method detects 3,246 RLs, accounting for 64.62 % of valid SBL profiles. Median RL-top heights reach 1,700 and 1,250 m at 08 and 20 BJT, respectively, compared with SBL medians of 82.0 and 73.7 m. These results demonstrate the value of separately characterizing near-surface stable layers and elevated residual structures in long-term boundary layer analyses.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

(4) Vertical resolution affects both the availability and the estimated positions of boundary layer heights, with particularly strong impacts on shallow SBLs. At 100 m resolution, CBL and NBL retrieval rates remain above 99 %, but their estimated heights shift upward overall. For soundings successfully retrieved at all four resolutions, median height differences relative to 5 m reach 235.6 and 212.8 m for CBLs and NBLs, respectively. The SBL retrieval rate initially increases from 66.55 % at 5 m to 74.26 % at 25 m, before declining to 39.14 % and 14.48 % at 50 and 100 m. Only 12.45 % of SBLs originally retrieved at or below 100 m retain valid estimates at 100 m resolution. Candidate-level analyses show that coarser resolution alters the detected interface positions, reduces observational support for physical scoring, changes the relative alignment of candidates, and redistributes fusion weights. These interacting effects explain why resolution changes propagate through the entire retrieval process rather than merely limiting the precision of individual height estimates.

Overall, the proposed method provides an interpretable framework for integrating multiple radiosonde-based PBLH estimates through physical consistency assessment and layer-aware fusion. The results highlight the importance of high-resolution observations not only for identifying shallow stable layers but also for distinguishing vertically separated structures. However, the present evaluation is limited to a single station, and the applicability of the empirical parameters and fusion criteria to different climatic and surface conditions remains to be established. Future studies involving multiple stations and complementary observations could further examine the generalizability of the framework and support broader investigations of boundary layer structures and their spatial and temporal variability.

Appendix A Preprocessing and classification parameters

A1 Observations and numerical support

Temperature, pressure, and specific humidity are expressed in kelvin, hectopascals, and kilograms per kilogram, respectively; horizontal wind components are in meters per second. Heights follow strictly increasing AGL coordinates. Native-grid gradients permit a maximum adjacent-level spacing of 25 m. Routine Savitzky–Golay smoothing uses a 25 m target window and a second-order polynomial. The point count is rounded upward to the nearest odd integer, subject to a minimum of 5 and limited by available valid levels. A complete 5 m grid uses 5 points, spanning 20 m from first to last. RL and A8 diagnostics use their specified actual spans of 100 m.

Wind quality control counts vector jumps exceeding 5 m s^{-1} between levels separated by no more than 15 m. A local spike comprises three points spanning at most 35 m. The middle vector must differ from the endpoint mean by more than 5 m s^{-1}, while the endpoint difference remains below 2 m s^{-1}. More than 10 large jumps, more than 3 isolated spikes, or a nonisolated wind speed above 60 m s^{-1} invalidate wind evidence for the entire profile. Otherwise, wind candidates and dynamic scores are excluded within 150 m of spikes. A3 additionally checks the 20–100 m reference layer and a 100 m neighborhood around the candidate; MOST checks its 20–200 m fitting region. Thermodynamic candidates are not removed wholesale by these checks.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

A2 Regime classification and local reassessment

Classification and conditional reassessment follow Sect. 2.2. The Theil–Sen fit requires at least three valid levels; slopes exactly at either classification threshold are assigned to NBL. An invalid afternoon reassessment leaves the original classification unchanged. NBL-to-SBL reassessment additionally requires the low-level anomaly checks below.

Low-level anomaly flags include an exactly linear segment of at least 300 m in a primary observed variable or a pressure decrease exceeding 5 hPa per 5 m. Flagged initial NBL profiles cannot pass the shallow-SBL reclassification criteria. These checks apply to classification reassessment rather than additional height screening of all profiles.

A3 RL parameters

For the RL search described in Sect. 2.6, both the smoothed column and search segment require at least 10 valid levels on the native grid. Core and cap depths are calculated from actual heights rather than fixed level counts.

Appendix B Candidate thresholds and internal quality

Define clip(x) = min(1, max(0, x)). Each Q_j describes only the internal structural quality of that candidate. Successful algorithm results with $Q_j = 0$ are retained for diagnosis but excluded from positive-weight fusion.

Table B1. Candidate thresholds and internal quality measures.

ID	Main thresholds and quality definition
A1	Search positive-gradient peaks over 50–3,500 m with gradients ≥0.0005 K m^{-1}. The prominence threshold is max(half the gradient threshold, twice the robust background scale). Q = clip[peak prominence/(2 × prominence threshold)].
A2	Search drying peaks over 50–3,500 m with negative specific humidity gradient magnitudes ≥10^{-6} kg kg^{-1} m^{-1}. Prominence and quality follow A1.
A3	Reference layer: 20–50 m for SBLs and 20–100 m otherwise. Critical Ri: 0.35 and 0.25, respectively. Persistent threshold exceedance: 15 m. Q = min(1, reference-level count/5) multiplied by the fraction of search levels with nondegenerate wind differences.
A4	Near-surface reference: mean virtual potential temperature over 20–50 m plus 0.05 K. Search the first environmental intersection over 50–3,500 m. Q = min(1, reference-level count/5).
A5	Use air temperature over 20–1,000 m. A local majority of valid near-surface gradients must be positive. Inversion depth ≥20 m and warming ≥0.2 K. Q is the mean of clip(depth/40 m) and clip(warming/0.4 K). Retain a boundary flag if the inversion persists to the search limit.
A6	Scales: 50, 100, 200, and 350 m. Positive response ≥0.1 K and peak prominence ≥0.05 K. Q = clip(prominence/response). Use only fully supported windows; equal maximum prominence favors the smaller scale.
A7	Maximum vector-shear plateau over 50–3,500 m. Background: 50–250 m from the candidate. Q = clip((peak/background median − 1)/2). Multiply by 0.25 if the plateau touches a search boundary.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

ID	Main thresholds and quality definition
A8	Search range: 100–3,500 m. Actual smoothing span: 100 m. Complete support depth: 100 m. Wind-speed change across the support ≥0.5 m s^{-1}. Two-sided curvature-sign consistency f ≥0.6. $Q = f \times$ clip(wind-speed change/1.0 m s^{-1}).
A9	Lowest complete jet core over 50–2,000 m. Core speed ≥8 m s^{-1}; decreases above and below each ≥2 m s^{-1}. Plateau tolerance: 0.2 m s^{-1}. Q = clip[min(core speed/12, lower-side decrease/4, upper-side decrease/4)], with speeds in m s^{-1}.

For SBLs, A1 retains its original gradient peak and can additionally record stable-layer termination transitions. These use a 75 m target smoothing window and 75 m cross-layer gradients. Assessment windows extend 100 m below and 180 m above the candidate, with 20 m buffer distances on both sides of the interface. The lower gradient must be at least 2 K km^{-1}, the upper gradient at most 1.5 K km^{-1}, and their decrease at least 2 K km^{-1}. Cumulative virtual potential temperature increase below must be at least 0.2 K. The search covers 60–2,000 m, requires at least 3 valid levels per side, and merges nearby transitions within 150 m. Additional qualifying transition candidates receive Q = 1 and subsequently undergo all six scores. Connection to near-surface stability is retained as diagnostic information, not an additional prerequisite for layer selection. Additional transitions remain diagnostic candidates and never directly enter primary-height weighting.

A1/A2 quality is further multiplied by 0.25 for peaks at search boundaries. The local check at the lowest end of A5 uses at least 3 levels. At least 60 % of finite gradients must be positive, with at least 2 positive levels. Searching upward from the minimum-depth position identifies the first persistent segment whose median gradient is nonpositive; each segment requires at least 2 levels. If no termination occurs, the upper search boundary is used and flagged. These discrete support requirements, warming, and depth criteria jointly determine candidate validity.

Appendix C Detailed mappings of the six scores

C1 Common windows and thermodynamic scoring

Define $\sigma(x) = 1/(1 + \exp(-x))$. Below candidate height h, the window is [max(z_{min}, h − 200), h − 25] m; above it, [h + 25, min(z_{max}, h + 200)] m. Let g_i and g_a denote median virtual potential temperature gradients in these windows, respectively. $\Delta\theta$ is the upper-window median virtual potential temperature minus the lower-window median. The gradient scale is G_0 = 0.001 K m^{-1}, and the temperature-difference scale is Θ_0 = 0.5 K. These common windows are used for C_E, C_{TKE}, and C_{ENT}; C_{SIM}, C_{GD}, and C_{WS} use the windows specified below. The following functions are empirical score mappings of the diagnosed physical features:

$$C_E^{CBL} = 1/2\,[\exp(-|g_i|/G_0) + \sigma(\Delta\theta/\Theta_0)]. \quad \text{(C1)}$$

$$C_E^{SBL} = 1/2\,[\sigma(g_i/G_0) + \sigma((g_i - g_a)/G_0)]. \quad \text{(C2)}$$

$$C_E^{NBL} = 1/2\,[\exp(-|g_i|/G_0) + 1 - \exp(-|g_a - g_i|/G_0)]. \quad \text{(C3)}$$

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Each side requires at least 2 valid gradient levels; q_E = min(1, smaller level count across the two sides/5). In Eq. (C1), the exponential rewards weak lower-window gradients, while the sigmoid rewards upward warming. In Eq. (C2), the two terms reward a positive lower-window gradient and its upward decrease. In Eq. (C3), the first term rewards weak lower-window gradients and the second increases with the magnitude of the gradient change. The scales G_0 and Θ_0 control the smoothness of these responses; they are not hard acceptance thresholds.

C2 Shear–buoyancy environment scoring

Let s_i and s_a denote the median S^2 in the two windows. Likewise, b_i and b_a are medians of max(N^2, 0), and c_i and c_a are medians of max($-N^2$, 0). Set $\varepsilon = 10^{-8}$ s^{-2} to avoid singular logarithmic ratios:

$$R_C = 1/2\left[\ln\frac{s_i + c_i + \varepsilon}{s_a + c_a + \varepsilon} + \ln\frac{b_a + \varepsilon}{b_i + \varepsilon}\right]. \qquad \text{(C4)}$$

$$R_S = 0.7\ln\frac{s_i + \varepsilon}{s_a + \varepsilon} + 0.3\ln\frac{b_i + b_a + \varepsilon}{\varepsilon}, \qquad R_N = \ln\frac{s_i + \varepsilon}{s_a + \varepsilon}. \qquad \text{(C5)}$$

CBLs, SBLs, and NBLs use $C_{TKE} = \sigma(R_C)$, $\sigma(R_S)$, and $\sigma(R_N)$, respectively. The logarithmic mapping scale is 1. In Eq. (C4), the first ratio rewards stronger combined shear and unstable-stratification signals below, while the second rewards stronger stable stratification above. In Eq. (C5), the SBL terms weight the lower-to-upper shear contrast by 0.7 and the presence of stable stratification by 0.3. The latter uses the sum of the two positive N^2 components, not their vertical contrast. The NBL expression retains only the shear contrast. A zero combined indicator maps to 0.5; positive and negative values map above and below 0.5. Each side requires at least 2 levels, with support quality defined as for C_E.

C3 Integrated MOST scoring

The fitting bounds in Eqs. (5)–(6) are $u_* \in [0.02, 2]$ m s^{-1} and $\lambda \in [-0.2, 0.2]$ m^{-1}. The upper λ bound is -10^{-7} for CBLs, the lower bound is 10^{-7} for SBLs, and NBLs use [−0.02, 0.02]. Fitting uses a soft-L1 loss with at most 500 function evaluations. Each native-grid fit requires at least 8 levels. Parameters must lie more than 0.5 % of the absolute boundary value away from each boundary. Normalized root-mean-square error (RMSE) must not exceed 3, and the condition number of $\mathbf{J}^T\mathbf{J}$ must not exceed 10^8. After an initial failure, the height window is expanded for one retry. Scoring always uses joint wind-speed and virtual potential temperature fitting.

Unstable integrated functions use $x = (1 - 16\zeta)^{1/4}$ and $y = (1 - 16\zeta)^{1/2}$:

$$\psi_m = 2\ln\frac{1+x}{2} + \ln\frac{1+x^2}{2} - 2\arctan x + \frac{\pi}{2}, \qquad \psi_h = 2\ln\frac{1+y}{2}. \qquad \text{(C6)}$$

Stable functions use

$$\psi_m = -6.1\ln[\zeta + (1 + \zeta^{2.5})^{1/2.5}], \qquad \psi_h = -5.3\ln[\zeta + (1 + \zeta^{1.1})^{1/1.1}]. \qquad \text{(C7)}$$

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

The stable-function parameters in Eq. (C7) are also discussed by Cheng et al. (2005) in relation to stable boundary layer similarity.

Let $f = \exp(-\mathrm{RMSE}^2/2)$. Stability support s equals 1 when h/L lies within $[-100, -0.1]$ for CBLs or $[0.05, 100]$ for SBLs. A sign inconsistent with the regime gives $s = 0$. For the correct sign outside the interval, distance d to the nearest endpoint gives $\exp(-d/0.1)$ for CBLs and $\exp(-d/0.05)$ for SBLs. For NBLs, $s = \exp[-\max(|h/L| - 0.5, 0)/0.5]$. Finally, $C_{\mathrm{SIM}} = \sqrt{(sf)}$, with support quality $q_{\mathrm{SIM}} = \min(1, \text{valid-level count}/12)f$. These h/L intervals are empirical scoring ranges adopted in this method. Within a profile and regime, candidates share the fitted L and fit factor f; their heights enter through s. The geometric mean requires both factors to be large for a high score. The additional factor f in q_{SIM} reduces the weight assigned to a less well-supported fit in Eq. (7).

C4 Distributional information entropy and local gradient sharpness

For virtual potential temperature and specific humidity separately, 20 equal-width bins span the pooled value range of the common upper and lower windows. Observations are weighted by adjacent height spans to obtain bin probabilities p_b. Both windows use identical bin boundaries. Normalized entropy is

$$H = -\frac{\sum_b p_b \ln p_b}{\ln 20}, \qquad C_{\mathrm{ENT}} = \sigma\left(\frac{(H_{\theta,\mathrm{a}} - H_{\theta,\mathrm{i}}) + (H_{q,\mathrm{a}} - H_{q,\mathrm{i}})}{2 \times 0.25}\right). \qquad \text{(C8)}$$

Constant distributions have zero entropy. Equal upper- and lower-window entropies, after averaging across the two variables, give $C_{\mathrm{ENT}} = 0.5$; an upward increase gives a larger score and an upward decrease a smaller score. The scale 0.25 controls how rapidly entropy differences change the score. This mapping favors relatively concentrated lower-window distributions and more dispersed upper-window distributions; it measures their bin occupancy rather than turbulent fluctuations or a difference in mean temperature or humidity. SBLs retain this directional distributional test alongside the regime-specific thermal and dynamic scores. Each window requires at least 3 levels; $q_{\mathrm{ENT}} = \min(1, \text{smaller level count across the two sides}/10)$.

C_{GD} uses J, the absolute difference between median gradients within $[h - 25, h]$ and $[h, h + 25]$. Background regions are $[h - 250, h - 50]$ and $[h + 50, h + 250]$. The background scale B is 1.4826 MAD. This score is unavailable if B is zero or cannot be calculated:

$$C_{\mathrm{GD}} = 1 - \exp[-J/(2B)]. \qquad \text{(C9)}$$

Each of the four regions requires at least 2 levels; $q_{\mathrm{GD}} = \min(1, \text{smallest level count across the four regions}/5)$. The ratio J/B describes interface-gradient change relative to local background variability. The absolute contrast treats gradient increases and decreases alike. The factor 2 sets the response scale: larger background variability reduces the score for the same gradient change. The buffer regions isolate the local contrast from its background estimate.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

C5 Wind structure scoring

The local candidate window is $h \pm 50$ m, with background regions 50–250 m from h on both sides. Let S_ℓ and S_b denote median vector shear magnitudes in the local and pooled background regions, respectively. D is the magnitude of the difference between upper and lower background median shear vectors, divided by $S_b + \varepsilon$. The score is

$$C_{WS} = 1/2\left[\sigma\left(\ln\frac{S_\ell}{S_b+\varepsilon}\right) + 1 - \exp(-D)\right]. \qquad \text{(C10)}$$

The logarithm argument has a numerical lower bound of 10^{-8}, and the background denominator uses $\varepsilon = 10^{-8}$ s^{-1}. The local and both background regions each require at least 2 levels. q_{WS} = min(1, smallest level count across the three regions/5). The first term in Eq. (C10) measures local shear prominence relative to the background. The second increases from zero as the upper and lower median shear vectors diverge in magnitude or direction. Their equal-weight mean combines these distinct signals without imposing a preferred direction of rotation.

C6 Baseline weights and information redundancy control

Table C1. Baseline weights of the six physical consistency scores by boundary layer regime.

Regime	C_E	C_{TKE}	C_{SIM}	C_{ENT}	C_{GD}	C_{WS}
CBL	0.24	0.16	0.20	0.12	0.16	0.12
NBL	0.16	0.24	0.15	0.10	0.15	0.20
SBL	0.20	0.20	0.15	0.10	0.15	0.20

Directly shared criteria are downweighted for the following algorithm–score pairs: A1–C_E/C_{GD}, A2–C_{ENT}, A3–C_{TKE}/C_{SIM}, A4–C_E, A5–C_E/C_{GD}, A6–C_{GD}, A7–C_{TKE}/C_{WS}, A8–C_{WS}, and A9–C_{TKE}/C_{WS}. These pairs use $\rho = 0.5$; all others use $\rho = 1$. This controls redundancy between candidate generation and scoring.

Appendix D Coarse-grid parameters

Table D1. Parameter settings for the coarse-resolution experiments.

Parameter	25 m	50 m	100 m	Mandatory pressure levels
Minimum points for routine smoothing	3	3	3	5
Configured adjacent-level gap limit for differentiation (m)	37.5	75	150	25
Minimum A8 support depth (m)	100	200	400	100
Minimum A8 wind-speed change (m s^{-1})	0.5	1.0	2.0	0.5

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Parameter	25 m	50 m	100 m	Mandatory pressure levels
Minimum valid levels for MOST	8	4	3	8
Minimum valid levels for RL detection	10	5	5	Not performed
Minimum RL-cap persistence depth (m)	25	50	100	Not performed

Except for A5, routine smoothing uses the target window and grid spacing to determine an odd point count no smaller than the tabulated minimum. Polynomial order does not exceed the point count minus 2. A5 uses a second-order filter with at least five points. RL smoothing uses five points at 25 m and at least three points at 50 and 100 m. The resulting A5 and RL smoothing spans are reported in Sect. 2.7 and recorded with the height results for each grid.

The tabulated value is the configured gap-limit parameter. For gradient diagnostics that adapt to valid-level spacing, the effective maximum permitted adjacent-level gap is the larger of this parameter and 1.5 times the median height spacing of the levels passed to the calculation. Gradients are not evaluated across gaps exceeding that effective limit. Thus, the 25 m parameter for mandatory pressure levels is not a fixed 25 m upper limit on their usable spacing. Two-sided windows and minimum observation counts follow the requirements of individual scores.

The 150 m SBL qualifying-layer span, requirement for at least two primary-candidate physical categories, within-layer P_j^2-weighted consensus, and 1.25 dominance threshold remain unchanged across resolutions. Omitted windows are enumerated only when original sequential clustering produces no qualifying layer. Overlapping qualifying windows with shared primary candidates remain unresolved when a unique layer cannot be selected. RL detection proceeds independently above the final primary height.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Appendix E Full-height view of the resolution case

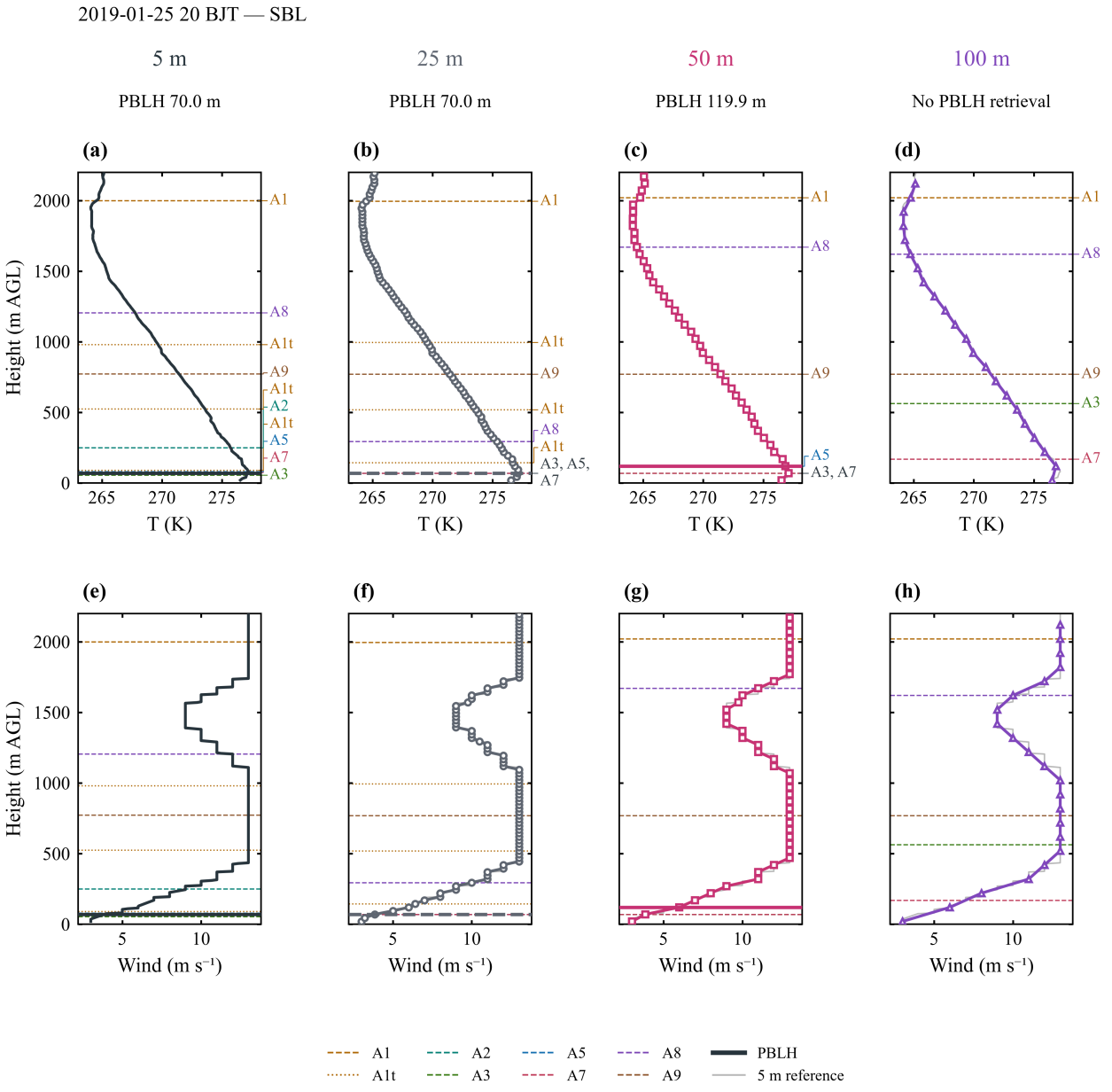


Figure E1. Full-height structure of the shallow-SBL case at 20 BJT on 25 January 2019. (a–d) Air-temperature profiles. (e–h) Wind-speed profiles. Columns show 5, 25, 50, and 100 m resolution, respectively. Thin gray lines show 5 m reference profiles; symbols indicate coarse-grid observation points. Thin horizontal lines mark candidate heights, and thick horizontal lines mark fused PBLH. A1t denotes an additional thermodynamic transition. The plotted range is 0–2,200 m, extending the view in Fig. 11.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Code and data availability

The retrieved planetary boundary-layer height (PBLH) dataset is publicly available on Zenodo at https://doi.org/10.5281/zenodo.23556264 (Zhang et al., 2026b). The retrieval and analysis code is available at https://doi.org/10.5281/zenodo.23554993 (Zhang et al., 2026a). The original radiosonde observations were provided by the Beijing Meteorological Observation Center, China Meteorological Administration, Beijing, China. These observations are not publicly available owing to the data provider's data-sharing policy. Requests for access may be directed to Minquan Zhang at zhangminquan25@mails.ucas.ac.cn and are subject to approval by the data provider.

Author contributions

XX, YX, and QS provided the initial ideas for the methodology of the study; YX and XF provided data support; MZ and XX completed the refinement of the methodology, the specific code implementation, and data processing; MZ, DF, and HS jointly prepared the figures and drafted the manuscript. All authors contributed to the interpretation of the results and revision of the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

The authors gratefully acknowledge the Beijing Meteorological Observatory (station 54511) for providing the radiosonde data. The authors acknowledge the use of AI-assisted tools for language editing and proofreading of this manuscript. All scientific content, data analysis, and conclusions were conceived and written entirely by the authors, who take full responsibility for the final work.

Financial support

This research was supported by the National Natural Science Foundation of China (NSFC; 42275091) and the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB0760402).

References

Baas, P., Steeneveld, G. J., van de Wiel, B. J. H., and Holtslag, A. A. M.: Exploring Self-Correlation in Flux–Gradient Relationships for Stably Stratified Conditions, J. Atmos. Sci., 63, 3045–3054, https://doi.org/10.1175/jas3778.1, 2006.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Baas, P., Bosveld, F. C., Klein Baltink, H., and Holtslag, A. A. M.: A Climatology of Nocturnal Low-Level Jets at Cabauw, J. Appl. Meteorol. Climatol., 48, 1627–1642, https://doi.org/10.1175/2009jamc1965.1, 2009.

Brooks, I. M.: Finding Boundary Layer Top: Application of a Wavelet Covariance Transform to Lidar Backscatter Profiles, J. Atmos. Ocean. Technol., 20, 1092–1105, https://doi.org/10.1175/1520-0426(2003)020<1092:fbltao>2.0.co;2, 2003.

Businger, J. A., Wyngaard, J. C., Izumi, Y., and Bradley, E. F.: Flux-Profile Relationships in the Atmospheric Surface Layer, J. Atmos. Sci., 28, 181–189, https://doi.org/10.1175/1520-0469(1971)028<0181:fprita>2.0.co;2, 1971.

Caicedo, V., Sedlar, J., Riihimaki, L. D., Angevine, W., Turner, D. D., Zhang, D., and Lantz, K.: Enhanced Boundary Layer Height Detection Using Ceilometer, Surface Meteorology, and Radiation Products With a Random Forest Ensemble Method, J. Geophys. Res.-Atmos., 130, e2025JD043385, https://doi.org/10.1029/2025jd043385, 2025.

Chen, X., Yang, T., Wang, Z., Wang, F., and Wang, H.: An ensemble method for improving the estimation of planetary boundary layer height from radiosonde data, Atmos. Meas. Tech., 16, 4289–4302, https://doi.org/10.5194/amt-16-4289-2023, 2023.

Cheng, Y., and Brutsaert, W.: Flux-profile relationships for wind speed and temperature in the stable atmospheric boundary layer, Bound.-Lay. Meteorol., 114, 519–538, https://doi.org/10.1007/s10546-004-1425-4, 2005.

Cheng, Y., Parlange, M. B., and Brutsaert, W.: Pathology of Monin–Obukhov similarity in the stable boundary layer, J. Geophys. Res.-Atmos., 110, D06101, https://doi.org/10.1029/2004JD004923, 2005.

Davis, K. J., Gamage, N., Hagelberg, C. R., Kiemle, C., Lenschow, D. H., and Sullivan, P. P.: An Objective Method for Deriving Atmospheric Structure from Airborne Lidar Observations, J. Atmos. Ocean. Technol., 17, 1455–1468, https://doi.org/10.1175/1520-0426(2000)017<1455:aomfda>2.0.co;2, 2000.

Duncan Jr., J. B., Bianco, L., Adler, B., Bell, T., Djalalova, I. V., Riihimaki, L., Sedlar, J., Smith, E. N., Turner, D. D., Wagner, T. J., and Wilczak, J. M.: Evaluating convective planetary boundary layer height estimations resolved by both active and passive remote sensing instruments during the CHEESEHEAD19 field campaign, Atmos. Meas. Tech., 15, 2479–2502, https://doi.org/10.5194/amt-15-2479-2022, 2022.

Dyer, A. J.: A review of flux-profile relationships, Bound.-Lay. Meteorol., 7, 363–372, https://doi.org/10.1007/bf00240838, 1974.

Grachev, A. A., Andreas, E. L., Fairall, C. W., Guest, P. S., and Persson, P. O. G.: The Critical Richardson Number and Limits of Applicability of Local Similarity Theory in the Stable Boundary Layer, Bound.-Lay. Meteorol., 147, 51–82, https://doi.org/10.1007/s10546-012-9771-0, 2013.

Guo, J., Zhang, J., Shao, J., Chen, T., Bai, K., Sun, Y., Li, N., Wu, J., Li, R., Li, J., Guo, Q., Cohen, J. B., Zhai, P., Xu, X., and Hu, F.: A merged continental planetary boundary layer height dataset based on high-resolution radiosonde measurements, ERA5 reanalysis, and GLDAS, Earth Syst. Sci. Data, 16, 1–14, https://doi.org/10.5194/essd-16-1-2024, 2024.

Holzworth, G. C.: Estimates of mean maximum mixing depths in the contiguous United States, Mon. Weather Rev., 92, 235–242, https://doi.org/10.1175/1520-0493(1964)092<0235:eommmd>2.3.co;2, 1964.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Kotthaus, S., Bravo-Aranda, J. A., Collaud Coen, M., Guerrero-Rascado, J. L., Costa, M. J., Cimini, D., O'Connor, E. J., Hervo, M., Alados-Arboledas, L., Jiménez-Portaz, M., Mona, L., Ruffieux, D., Illingworth, A., and Haeffelin, M.: Atmospheric boundary layer height from ground-based remote sensing: a review of capabilities and limitations, Atmos. Meas. Tech., 16, 433–479, https://doi.org/10.5194/amt-16-433-2023, 2023.

Li, H., Liu, B., Ma, X., Jin, S., Ma, Y., Zhao, Y., and Gong, W.: Evaluation of retrieval methods for planetary boundary layer height based on radiosonde data, Atmos. Meas. Tech., 14, 5977–5986, https://doi.org/10.5194/amt-14-5977-2021, 2021.

Liu, S., and Liang, X.-Z.: Observed Diurnal Cycle Climatology of Planetary Boundary Layer Height, J. Climate, 23, 5790–5809, https://doi.org/10.1175/2010jcli3552.1, 2010.

Mahrt, L.: Stably Stratified Atmospheric Boundary Layers, Annu. Rev. Fluid Mech., 46, 23–45, https://doi.org/10.1146/annurev-fluid-010313-141354, 2014.

Min, J.-S., Park, M.-S., Chae, J.-H., and Kang, M.: Integrated System for Atmospheric Boundary Layer Height Estimation (ISABLE) using a ceilometer and microwave radiometer, Atmos. Meas. Tech., 13, 6965–6987, https://doi.org/10.5194/amt-13-6965-2020, 2020.

Paulson, C. A.: The Mathematical Representation of Wind Speed and Temperature Profiles in the Unstable Atmospheric Surface Layer, J. Appl. Meteorol., 9, 857–861, https://doi.org/10.1175/1520-0450(1970)009<0857:tmrows>2.0.co;2, 1970.

Roldán-Henao, N., Su, T., and Li, Z.: Refining Planetary Boundary Layer Height Retrievals From Micropulse-Lidar at Multiple ARM Sites Around the World, J. Geophys. Res.-Atmos., 129, e2023JD040207, https://doi.org/10.1029/2023jd040207, 2024.

Savitzky, A., and Golay, M. J. E.: Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Anal. Chem., 36, 1627–1639, https://doi.org/10.1021/ac60214a047, 1964.

Seidel, D. J., Ao, C. O., and Li, K.: Estimating climatological planetary boundary layer heights from radiosonde observations: Comparison of methods and uncertainty analysis, J. Geophys. Res.-Atmos., 115, D16113, https://doi.org/10.1029/2009jd013680, 2010.

Seidel, D. J., Zhang, Y., Beljaars, A., Golaz, J.-C., Jacobson, A. R., and Medeiros, B.: Climatology of the planetary boundary layer over the continental United States and Europe, J. Geophys. Res.-Atmos., 117, D17106, https://doi.org/10.1029/2012jd018143, 2012.

Shannon, C. E.: A Mathematical Theory of Communication, Bell Syst. Tech. J., 27, 379–423, https://doi.org/10.1002/j.1538-7305.1948.tb01338.x, 1948.

Smith, E. N., and Carlin, J. T.: A multi-instrument fuzzy logic boundary-layer-top detection algorithm, Atmos. Meas. Tech., 17, 4087–4107, https://doi.org/10.5194/amt-17-4087-2024, 2024.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, Kluwer Academic Publishers, https://doi.org/10.1007/978-94-009-3027-8, 1988.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.

Su, T., Li, Z., and Kahn, R.: Relationships between the planetary boundary layer height and surface pollutants derived from lidar observations over China: regional pattern and influencing factors, Atmos. Chem. Phys., 18, 15921–15935, https://doi.org/10.5194/acp-18-15921-2018, 2018.

Su, T., Zhang, Y., Zheng, Y., Roldán-Henao, N., Kahn, R., and Li, Z.: Remote sensing of planetary boundary layer from ground and space: Structure, thermodynamics, and boundary layer clouds, Remote Sens. Environ., 344, 115544, https://doi.org/10.1016/j.rse.2026.115544, 2026.

Sun, H., Wang, L., Shi, H., Ran, L., Zhang, J., Bai, Z., Liu, B., Zhen, X., Fan, X., and Chen, H.: Observational Study of Local and Non-local Effects on Aerosols Vertical Distribution in the Boundary Layers, Aerosol Sci. Eng., https://doi.org/10.1007/s41810-025-00356-7, 2025.

Tan, Y., Xin, J., Qin, X., Ma, Y., Xu, J., Du, Y., Shang, C., Zhao, L., Tian, Y., Fan, Q., Zhou, X., Ren, Y., Wang, P., Hao, F., Yin, N., Ren, X., Peng, K., Jia, D., Pan, X., Li, J., Wang, Z., and Butt, A. Q.: Differences Among Thermal, Dynamic, and Material Boundary Layers Over Complex Terrain, J. Geophys. Res.-Atmos., 130, e2024JD043043, https://doi.org/10.1029/2024jd043043, 2025.

Teixeira, J., Piepmeier, J. R., Nehrir, A. R., Ao, C. O., Chen, S. S., Clayson, C. A., Fridlind, A. M., Lebsock, M., McCarty, W., Salmun, H., Santanello, J. A., Turner, D. D., Wang, Z., and Zeng, X.: Toward a Global Planetary Boundary Layer Observing System: A Summary, B. Am. Meteorol. Soc., 106, E1566–E1579, https://doi.org/10.1175/bams-d-23-0228.1, 2025.

Vogelezang, D. H. P., and Holtslag, A. A. M.: Evaluation and model impacts of alternative boundary-layer height formulations, Bound.-Lay. Meteorol., 81, 245–269, https://doi.org/10.1007/bf02430331, 1996.

Wang, C., Shi, H., Jin, L., Chen, H., and Wen, H.: Measuring boundary-layer height under clear and cloudy conditions using three instruments, Particuology, 28, 15–21, https://doi.org/10.1016/j.partic.2015.04.004, 2016.

Wang, F., Yang, T., Wang, Z., Chen, X., Wang, H., and Guo, J.: A comprehensive evaluation of planetary boundary layer height retrieval techniques using lidar data under different pollution scenarios, Atmos. Res., 253, 105483, https://doi.org/10.1016/j.atmosres.2021.105483, 2021.

Yue, M., Wang, M., Guo, J., Zhang, H., Dong, X., and Liu, Y.: Long-Term Trend Comparison of Planetary Boundary Layer Height in Observations and CMIP6 Models over China, J. Climate, 34, 8237–8256, https://doi.org/10.1175/jcli-d-20-1000.1, 2021.

Zhang, D., Comstock, J., Sivaraman, C., Mo, K., Krishnamurthy, R., Tian, J., Su, T., Li, Z., and Roldán-Henao, N.: Best estimate of the planetary boundary layer height from multiple remote sensing measurements, Atmos. Meas. Tech., 18, 3453–3475, https://doi.org/10.5194/amt-18-3453-2025, 2025.

Zhang, H., Zhang, X., Li, Q., Cai, X., Fan, S., Song, Y., Hu, F., Che, H., Quan, J., Kang, L., and Zhu, T.: Research progress on estimation of atmospheric boundary layer height, Acta Meteorol. Sin., 78, 522–536 (in Chinese with English abstract), https://doi.org/10.11676/qxxb2020.044, 2020.

https://doi.org/10.5194/egusphere-2026-6592
Preprint. Discussion started: 9 November 2026
© Author(s) 2026. CC BY 4.0 License.



Zhang, J., Guo, J., Li, J., Zhang, S., Tong, B., Shao, J., Li, H., Zhang, Y., Cao, L., Zhai, P., Xu, X., and Wang, M.: A Climatology of Merged Daytime Planetary Boundary Layer Height Over China From Radiosonde Measurements, J. Geophys. Res.-Atmos., 127, e2021JD036367, https://doi.org/10.1029/2021jd036367, 2022.

Zhang, M., Xia, Y., Sun, Q., Fu, D., Shi, H., Fan, X., and Xia, X.: Code for a physical-consistency-based ensemble method for estimating boundary-layer height from radiosondes (version 1.0.0), Zenodo [code], https://doi.org/10.5281/zenodo.23554993, 2026a.

Zhang, M., Xia, Y., Sun, Q., Fu, D., Shi, H., Fan, X., and Xia, X.: Derived boundary-layer and residual-layer heights from radiosondes at Beijing station, 2009–2023 (version 1.0.0), Zenodo [data set], https://doi.org/10.5281/zenodo.23556264, 2026b.

Zhang, W., Guo, J., Miao, Y., Liu, H., Song, Y., Fang, Z., He, J., Lou, M., Yan, Y., Li, Y., and Zhai, P.: On the Summertime Planetary Boundary Layer with Different Thermodynamic Stability in China: A Radiosonde Perspective, J. Climate, 31, 1451–1465, https://doi.org/10.1175/jcli-d-17-0231.1, 2018.