

# Balance plot for visualizing and examining tradeoff between accuracy and data quantity of lidar water vapor measurement data

Atsushi Takahashi<sup>1</sup>, Tetsu Sakai<sup>2</sup>, Takuya Kawabata<sup>2</sup>, Satoru Yoshida<sup>2</sup>, and Naonori Ueda<sup>1</sup>

<sup>1</sup>*Riken Center for Advanced Intelligence Project,  
Nihonbashi 1-chome Mitsui Building, 15<sup>th</sup> floor, 1-4-1 Nihonbashi, Chuo-ku, Tokyo,  
Japan 103-0027*

<sup>2</sup>*Meteorological Research Institute,  
1-1 Nagamine Tsukuba, Ibaraki, Japan 305-0052*

(Received March 3, 2022; revised June 7, 2022; revised August 16, 2022; accepted August 17, 2022)

Water vapor lidar has made it possible to continuously observe water vapor in the atmosphere and is a prominent tool for studying the hydrological cycle, cloud formation, and atmospheric convection. However, a standardized quality control method for managing the accuracy and number of available data has not been established, as well as a method for determining the appropriate spatial resolution. We propose a concise, visual framework in which observed data is plotted to determine the uncertainty threshold and spatial resolution of lidar data. Application of the method to observed data showed that it is effective for determining the optimum uncertainty threshold and the finest vertical resolution for balancing the accuracy and data quantity at every altitude range.

**Key Words:** Water vapor Raman lidar, Quality Control, Radiosonde, Spatial Resolution

## 1. Introduction

The radiosonde has been widely used in atmospheric science for measuring water vapor as it provides an accurate and dependable vertical profile. Despite its effectiveness, however, it is difficult to observe water vapor continuously because of restrictions on the available frequency of receivers at ground stations. Moreover, it is costly to launch radiosondes continuously.

To deal with these shortcomings, lidar (i.e., Light Detection and Ranging) has been introduced for continuous water vapor observation. Previous studies demonstrated the effectiveness of lidar for measuring the water vapor profile of the troposphere continuously.<sup>1-3)</sup> Subsequently, lidar observation of water vapor has been expected to be utilized for numerical weather forecasts.<sup>4)</sup>

Despite such success, there are still obstacles to expanding the use of lidar observation. First, a principle of data quality control needs to be established. Experimental methods for ensuring data accuracy were proposed in previous studies. For example, Dinoev et al.<sup>2)</sup> used an uncertainty threshold of 10% to choose the optimum vertical resolution varying between 15 m and 300 m, while Sakai et al.'s case study<sup>5)</sup> used an empirical uncertainty threshold of 30% at vertical resolutions of 75 m and 150 m for data quality control. Both studies intended to make data as accurate as possible, but they did not aim for balance between accuracy and the number of available data. A too strict uncertainty threshold results in data that is too sparse. A loose threshold increases the number of measurements; however, it prevents further analysis by assimilating data into numerical weather models due to large noise levels.

Determining the optimum spatial resolution is another concern. The vertical resolution of lidar measurements needs to be as fine as possible. However, in practice, it must be coarser at higher altitudes in order to maintain an adequate

signal-to-noise ratio. If the resolution is too coarse, the data may not be usable even if it is precise, while data with too fine of a resolution and a large error will not be applicable to numerical weather modeling.

It is necessary to balance attaining a fine spatial resolution and minimizing the measurement uncertainty. In this paper, we propose an intuitive and concise method of visualizing resolution and measurement uncertainty to determine the most optimal balance. In the following section, we demonstrate how our diagram identifies the optimum uncertainty threshold and vertical resolution using the dataset obtained with the Raman lidar used by Sakai et al.'s case study<sup>5)</sup>.

## 2. Data and instrument

### 2.1 Water vapor Raman lidar

#### 2.1.1 Instrument

The water vapor Raman lidar used in our study was developed by the Meteorological Research Institute (MRI) of Japan. Details of the lidar system and the results of the comparison with collocated radiosonde, the global navigation satellite system (GNSS), and objective analysis data are given in Sakai et al. (2019). The lidar utilizes tripled harmonics of a Nd:YAG laser operating at 355 nm with 220 mJ/pulse, a 35-cm diameter Cassegrain telescope, receiving optics, and a data acquisition system. Raman backscattering is detected from water vapor (407.5 nm) and nitrogen (386.7 nm), and elastic backscattering from air molecules and particles (355 nm). The signals are acquired with a transient recorder operating in photon-counting mode (20 MHz) and analog mode (12 bit). The original data resolution in the temporal and vertical directions was 1 min and 7.5 m, respectively.<sup>5)</sup> The data analysis procedure for deriving water vapor mixing ratio ( $W$ ) and its uncertainty ( $dW$ ) is given by Sakai et al. (2019). The values of  $dW$  are estimated from the measured photon counts assuming Poisson statistics.<sup>6)</sup> Note that the value of  $dW$  estimated from the observed photon counts is composed of the error from the precision of the lidar instrument itself<sup>6)</sup> and the water vapor concentration in the air varying during the measuring time window (20 min.) (D. N. Whiteman, personal communication, 26 April 2021).

#### 2.1.2 Lidar observation and data

Lidar water vapor was observed from August 2016 to December 2016 in Tsukuba, Japan. We averaged the raw data every 20 minutes in the temporal direction. For the vertical direction, we averaged the raw data with six different vertical resolutions, 15 m, 75 m, 150 m, 300 m, 600 m, and 900 m, to examine which resolution would balance accuracy and number of available data at each altitude.

### 2.2 Radiosonde data

We used two types of radiosondes (RS-11G, Meisei Electric, Co., Japan and RS92, Vaisala, Finland) as reference to validate the lidar data. The radiosondes were launched twice daily at 08:30 and 20:30 LST and weekly at 14:30 JST from an aerological observatory located 80 m northeast of the lidar. The temporal resolutions of the measured data were 1 sec for RS-11G and 2 sec for RS92, which correspond to the vertical resolutions of  $\sim 6$  m and  $\sim 12$  m, respectively. For quality control, Global Climate Observing System (GCOS) Reference Upper-Air network (GRUAN) data processing algorithms were applied to the raw data.<sup>7)</sup> The measurement uncertainties of relative humidity by the radiosondes were reported to be 5% and 3–5% for RS-11G<sup>8)</sup> and RS92<sup>7)</sup>, respectively.

## 3. Proposed method

### 3.1. Balance plot drawing

The value of  $dW/W$  represents the relative uncertainty of the measurements, and the conventional quality control method classifies data on the basis of whether  $dW/W$  is greater than or less than 0.3<sup>5)</sup>. However, this threshold value is given empirically and not objectively determined. In addition, the vertical resolution in the conventional method is uniform, but in reality, it needs to be adjusted for the observation altitude. Here, we introduce our method for determining the optimum threshold value of the measurement uncertainty for quality control and spatial resolution. First, we pair the lidar data with the radiosonde data recorded during the same measurement periods and at the same altitudes. These

pairs of data are sorted by  $dW/W$  values. Then we divide the dataset into sub-datasets on the basis of the altitude ranges. In other words, the data is in layers corresponding to altitude ranges. In our case study, we divide the data into 500-m layers below an altitude of 2,000 m. Above 2,000 m, we divide the data into 1,000-m layers. Then we use the root mean square error (RMSE) of the lidar and radiosonde data for the error metric. We introduce  $\varphi$  to denote the threshold value of  $dW/W$  and calculate the RMSE values using data with  $dW/W$  values below  $\varphi$ .

$$RMSE_i(\varphi) = \sqrt{\frac{1}{n_i(\varphi)} \sum_{j=1}^{n_i(\varphi)} (d_{j,lidar}(z, t, \varphi) - d_{j,sonde}(z, t))^2} \quad (1)$$

$z$  and  $t$  denote altitude and time, respectively.  $d_{lidar}(\varphi)$  denotes the water vapor mixing ratio measured by lidar with  $dW/W$  less than  $\varphi$ .  $d_{sonde}$  denotes the water vapor mixing ratio obtained by radiosonde observation.  $i$  is the index of altitude ranges.  $n(\varphi)$  is the number of available data whose  $dW/W$  are less than  $\varphi$ .

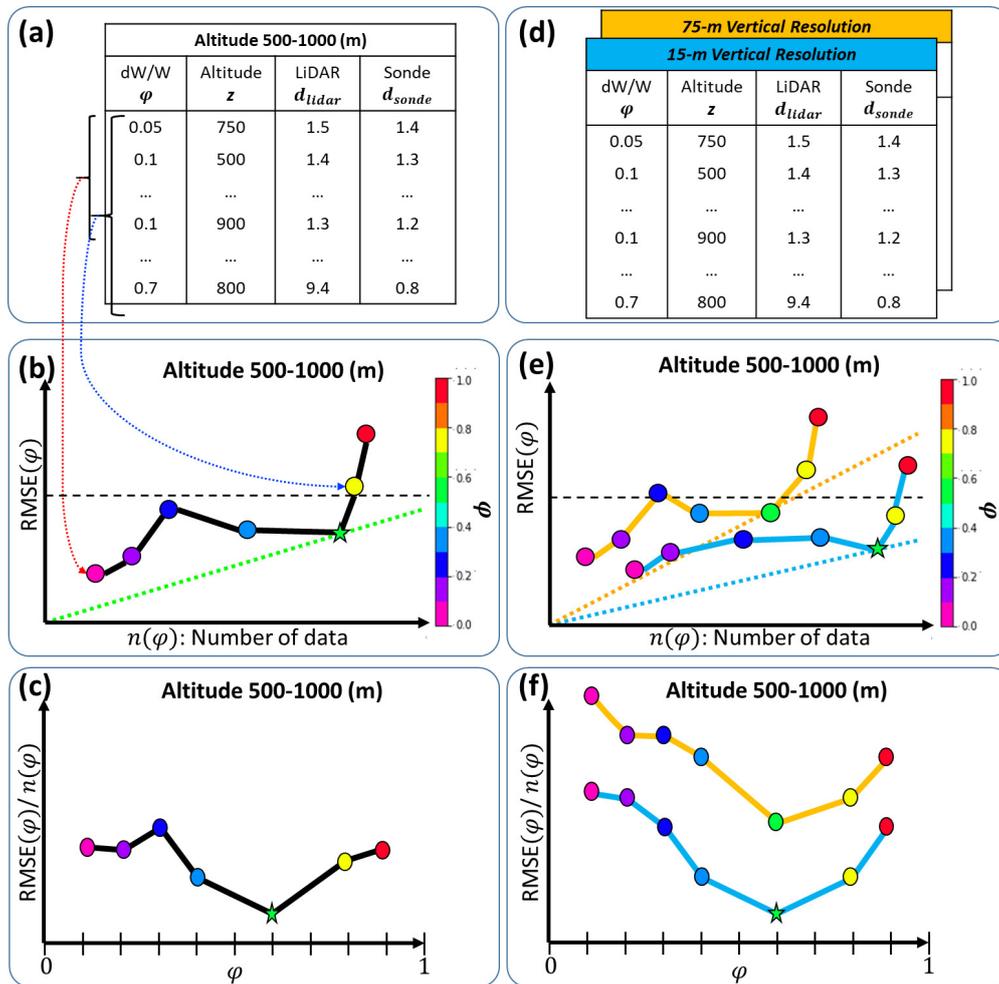


Fig. 1 Concept of proposed balance plot. (a): Table of dataset for all altitude ranges; lidar and radiosonde data recorded at the same altitude and time are paired. (b): The paired data are plotted in a scatter chart of RMSE values between lidar and radiosonde data vs. number of data. The symbols in the scatter chart are colored according to the color scale for the uncertainty threshold  $\varphi$ . The dashed black line represents the accuracy lower bound. A dashed colored line is a regression line that connects the origin of the coordinate and every datum. (c): The slope of the line corresponds to the ratio of RMSE value to the number of data. The optimum  $\varphi$  value corresponds to the bottom most green star, where the ratio of RMSE value to the number of data is minimum. (d): Application of proposed method to determine optimum vertical resolution. (e): Vertical resolution data is plotted and connected with corresponding colored lines. (f): The vertical resolution is determined by comparing their minimum slope value. The smallest slope value of the 15-m resolution is smaller than that of the other resolutions. Thus, the 15-m resolution should be the optimum vertical resolution in this altitude range.

Next, we prepare a scatter chart that takes the number of available data for the horizontal axis and the corresponding RMSE value for the vertical axis (Fig. 1; hereinafter, “balance plot”). The horizontal dashed line represents the target accuracy, which our quality control must achieve. The lines were drawn at 1.5 g/kg for daytime and 1.0 g/kg for nighttime in our case study. These accuracy targets are based on the case study in Sakai et al. (2019). Then we plot the  $n(\varphi)$  and RMSE ( $\varphi$ ) in the scatter chart. The trade-off between  $n(\varphi)$  and RMSE ( $\varphi$ ) is visualized for every  $\varphi$ ;  $\varphi$  ranges from 0 to 1 in increments of 0.05. The data in the scatter chart are colored according to the  $\varphi$  value. If more data is available, the corresponding datum is on the right side of the diagram. The data under the dashed line satisfies the target accuracy.

### 3.2. Balance plot comprehension

The bottom right most data indicates the most optimal balance of accuracy and amount of data (Fig. 1b). In other words, the slope of the regression line connecting the origin of the coordinates and the datum is smallest. In order to uniquely identify which  $\varphi$  value yields the smallest slope, we introduce another plot (Fig. 1c). In this plot, we take the slope of the regression lines for the vertical axis and  $\varphi$  value for the horizontal axis. As illustrated by the example in Fig. 1c, we can identify the most optimal balance (the light green star) and the corresponding  $\varphi$  value is 0.6.

### 3.3. Application to spatial resolution comparison

With a slight extension, this chart can also be used to identify the appropriate spatial resolution for all altitude ranges (Fig. 1d–1f). we prepared paired data with different spatial resolutions (Fig. 1d) and plot the data together with those of other spatial resolutions (Fig. 1e). The slopes of the regression lines are also shown for each vertical resolution as in Fig. 1c. The vertical resolution that contains the bottom most data is the optimum vertical resolution, which is the 15-m resolution (denoted with blue lines) in this plot. We will apply this chart to the observed data to identify the optimal resolution and threshold value in Section 4.

## 4. Case study

We applied our method to the dataset during daytime and nighttime separately to determine the optimum values because the measurement altitude ranges of the Raman lidar differed depending on time of day. The range was usually between 200 m and 1,000–2,000 m during the day and between 200 m and 5,000–6,000 m at night in cloud free conditions. Daytime was from 6 a.m. to 6 p.m. local time, and the other twelve hours were nighttime. For the daytime dataset, we examined data for every 500-m altitude range below 2,000 m (experiment 1). For the nighttime dataset, we examined data for every 500-m altitude range below 2,000 m. From 2,000 m up to 8,000 m, we examined data for every 1,000-m altitude range (experiment 2).

### 4.1 Experiment 1: Daytime data

Fig. 2 shows the balance plot for daytime data for every 500 m up to 2,000 m. The corresponding  $dW/W$  is colored according to the color scale on the right hand side. The colored lines show the RMSE for each spatial resolution of the data. Fig. 3 shows the ratio of RMSE value to the number of available data for every  $dW/W$  in increments of 0.05. The minimum of the ratio is highlighted with a red dot as long as it satisfies the target accuracy. Fig. 4 is the balance plot in which the optimum thresholds for the vertical resolutions are highlighted with red dots.

Fig. 4a shows that a spatial resolution of 15 m is the most effective below the altitude of 500 m because it has the most available data and is sufficiently accurate. However, above 500 m, it is no longer accurate enough and the number of available data is not necessarily beneficial. Between 500 m and 1,000 m, the plot shows that 75 m is the ideal resolution as it yields one and a half times as many data as the 150-m resolution does. Above 1,000 m, there is very little data that is sufficiently accurate. We cannot say for certain what threshold of  $dW/W$  is the optimum and which spatial resolution is appropriate for that altitude. The optimum thresholds of uncertainty and spatial resolution for daytime data are summarized in Table 1.

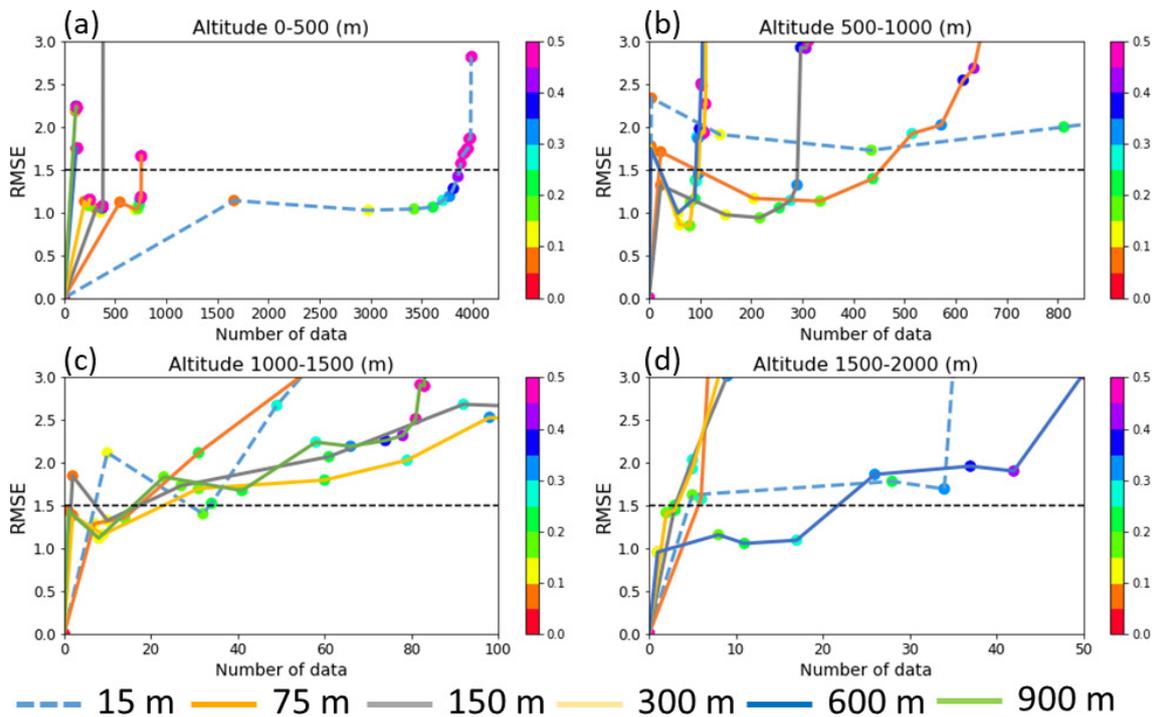


Fig. 2 Balance plot of daytime data analysis at every 500 m up to 2,000 m. The circles represent the threshold of  $dW/W$  corresponding to the color scale on the right side of the figure. The spatial resolution is denoted by the colored lines (e.g., the solid orange line denotes 75-m vertical resolution).

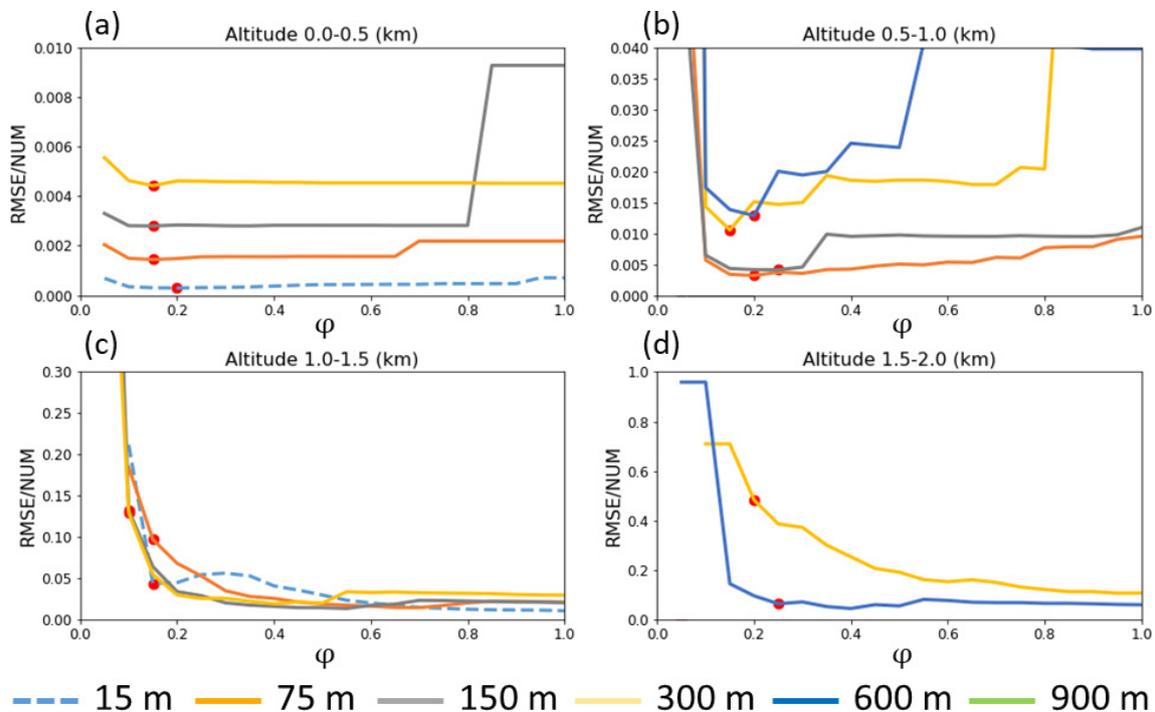


Fig. 3 Horizontal axis denotes the value of  $dW/W$  and the vertical axis is the ratio of RMSE value to the number of available data. The red dots indicate the minimum of every vertical resolution as long as the accuracy lower bound is satisfied.

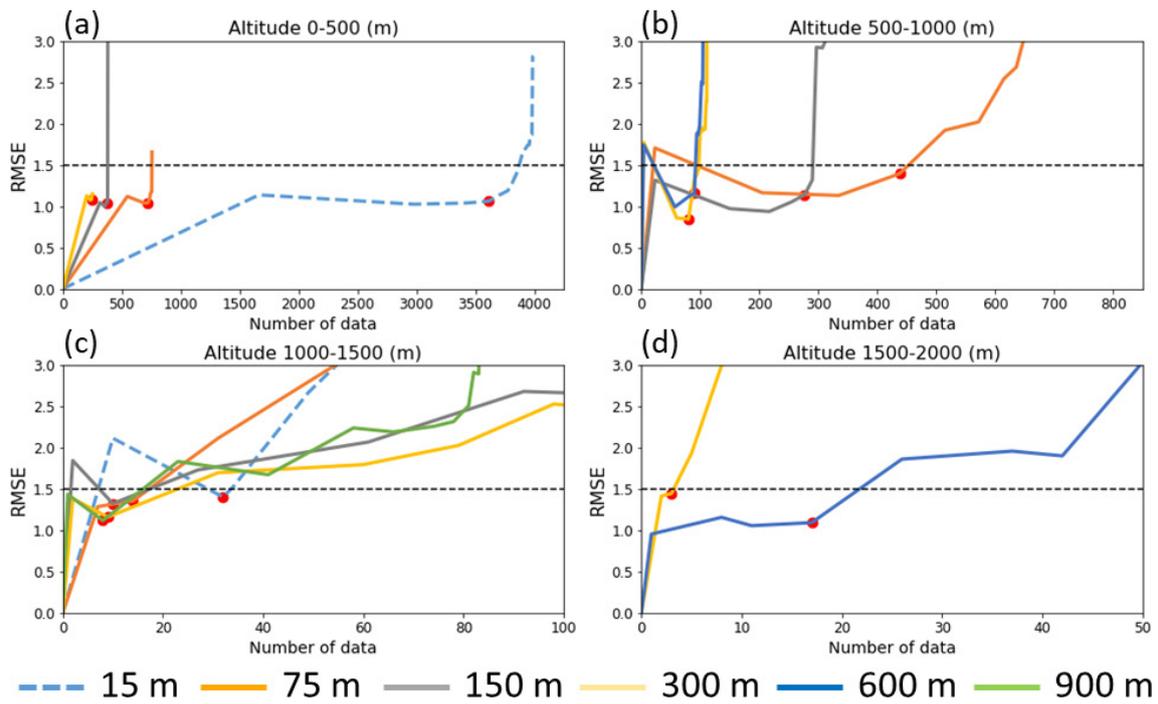


Fig. 4 Plots in Fig. 2 with optimum locations indicated with red dots for every vertical resolution.

Table 1 Optimal spatial resolution and threshold balance for every altitude range in daytime data.

Altitude Range (m)	Vertical Resolution (m)	Threshold $\varphi$ (dW/W)
0-500 (m)	15	0.2
500-1000 (m)	75	0.2
1000-1500 (m)	15	0.15
1500-2000 (m)	600	0.25

#### 4.2 Experiment 2: Nighttime data

Table 2 summarizes the optimal  $\varphi$  values and spatial resolutions for the nighttime data. Below an altitude of 1,000 m, the 15-m vertical resolution data is optimal in terms of accuracy and amount of available data (Fig. 5a–b, Fig. 6a–b). The 15-m resolution is also effective in the altitude range between 1,000 m and 1,500 m. Above 1,500 m, the 75-m reso-

Table 2 Optimal spatial resolution and uncertainty threshold balance for every altitude range in nighttime data.

Altitude Range (m)	Vertical Resolution (m)	Threshold $\varphi$ (dW/W)
0-500 (m)	15	0.1
500-1000 (m)	15	0.1
1000-1500 (m)	15	0.15
1500-2000 (m)	75	0.2
2000-3000 (m)	75	0.15
3000-4000 (m)	150	0.25
4000-5000 (m)	300	0.4
5000-6000 (m)	600	0.45

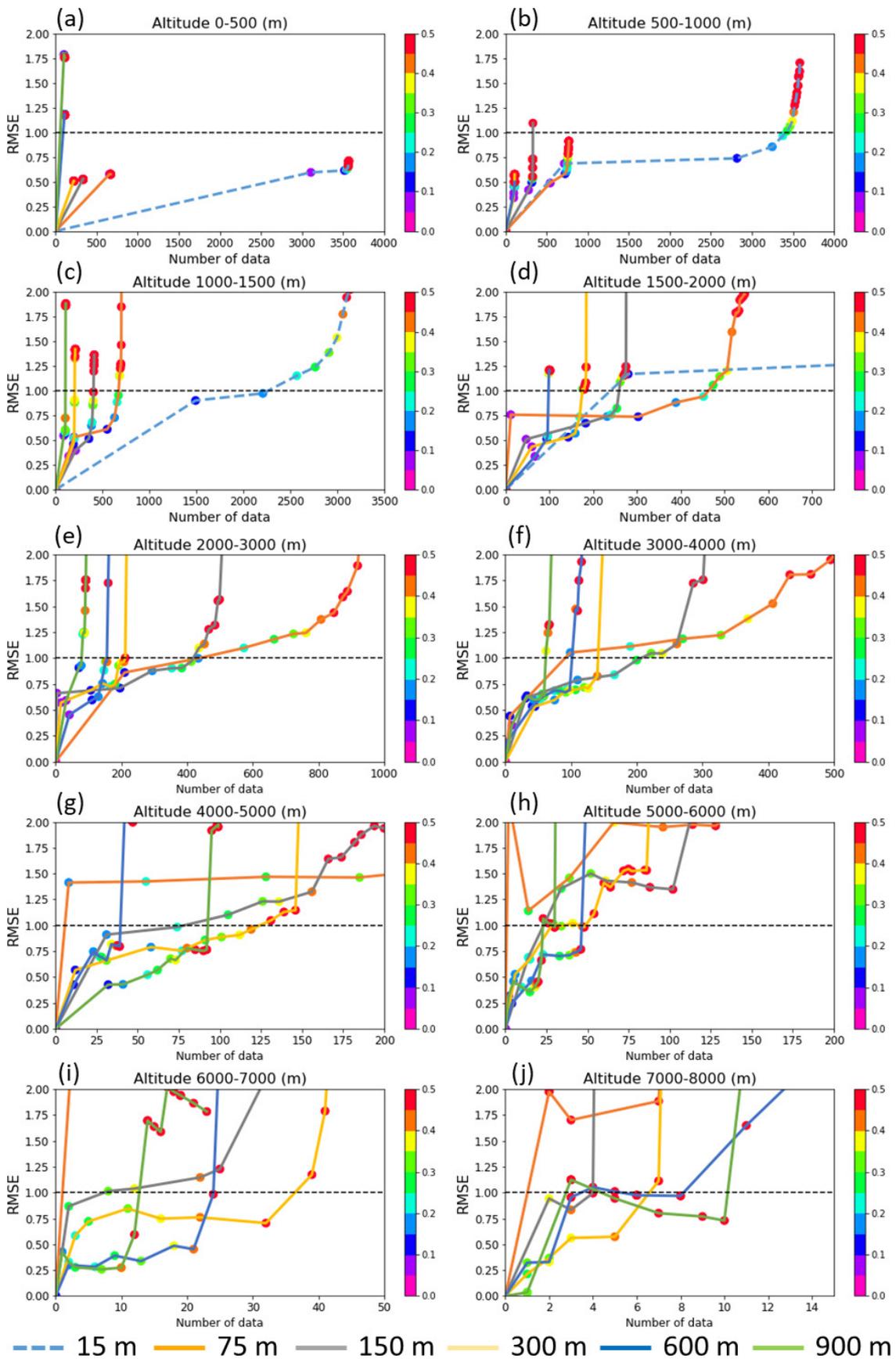


Fig. 5 Plots in Fig. 2 for nighttime data up to 8000-m altitude.

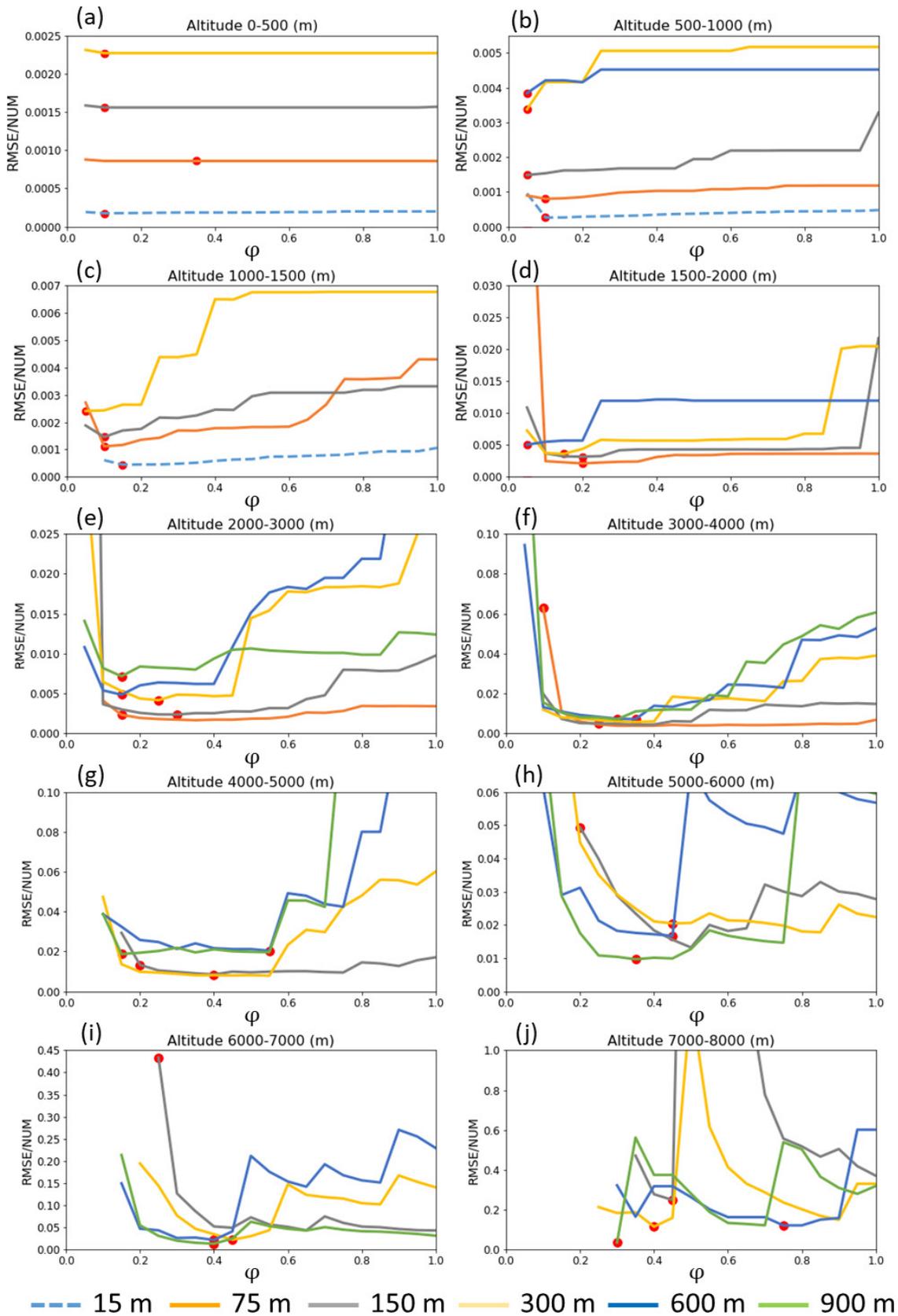


Fig. 6 Plots in Fig. 3 for nighttime data up to 8,000-m altitude. The most balanced cases in each resolution are highlighted with red dots.

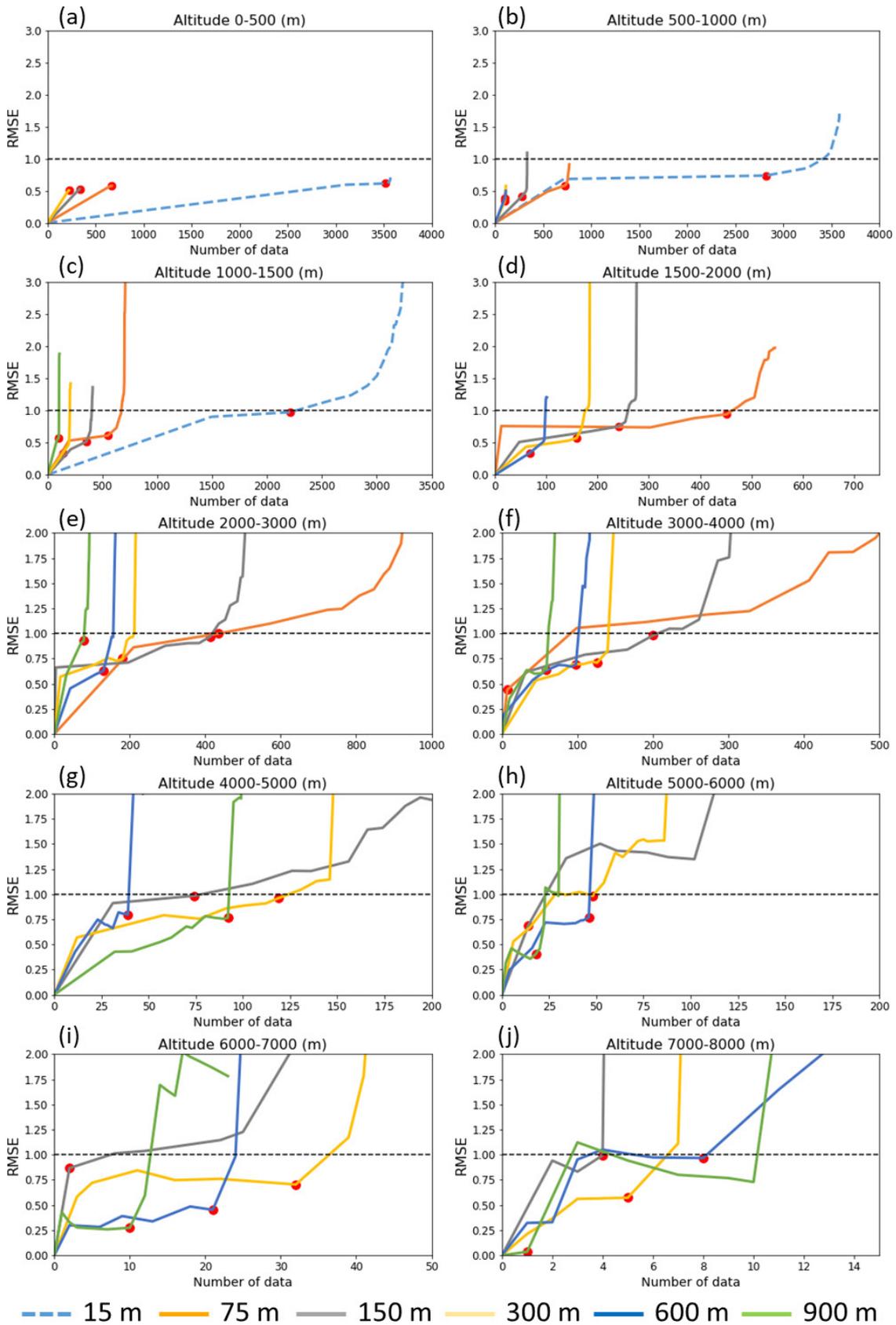


Fig. 7 Plots in Fig. 5 with optimum locations indicated with red dots for every vertical resolution.

lution outperforms the 15-m resolution (Fig. 6d). It requires a smaller uncertainty threshold than the empirical threshold, 0.3, to keep the RMSE value below the target accuracy. Between 2,000 m and 3,000 m, the 150-m resolution data is almost comparable in accuracy and amount of available data to those of the 75-m resolution data (Fig. 7e). The 150-m resolution yields more data than the 75-m resolution does in the 3,000–4,000-m altitude range (Fig. 7f). One layer above (4,000–5,000 m), the 300-m resolution data provides more available data than the 150-m resolution (Fig. 7g). At the next layer (5,000–6,000 m), the graph shows that the data of 300-m and 600-m resolutions are comparable (Fig. 7h). Above 6,000 m, we cannot conclude the optimal vertical resolution because the observation is limited (Fig. 5i-j).

## 5. Discussion

### 5.1 Optimum thresholds and spatial resolution

As seen in Fig. 2 and 5, the curves generally have a characteristic J-shape. The RMSE value is nearly constant as the number of available data increases as long as  $\varphi$  is low. By increasing the  $\varphi$  value, it reaches a tipping point, at which the accuracy suddenly decreases (RMSE value sharply increases). The J-shaped curves illustrate the trade-off between accuracy and the number of available data. The corner of the J-shaped curve indicates the most balanced situation. The most balanced spatial resolutions and corresponding uncertainty thresholds are summarized in Tables 1 and 2.

In addition, the optimum vertical resolution becomes coarser as the altitude increases (Tables 1 and 2). This indicates the trade-off between altitude and vertical resolution. In addition, judging from the decreasing number of available data (Fig. 5), the measurement limitation of our instruments may be 6,000 m. In other words, the characteristics of the instrument can be visualized by the balance plot.

Finally, we can see relatively larger (looser) uncertainty thresholds above 3,000 m at night (Table 2). We used the uniform target accuracy (RMSE < 1.0 g/kg) at all altitudes despite the fact that the amount of water vapor decreases as altitude increases. We will have to reflect this effect in the target accuracy in practice. However, the desirable accuracy target depends on the surroundings of the observation sites and the subsequent applications. It is beyond the scope of this paper. Nevertheless, our proposed framework can be adjusted to any target accuracy and will work robustly as demonstrated in the case study.

### 5.2 Comparison with conventional method

The conventional quality control method in Sakai et al. (2019) was based on empirical evidence. It yielded data of uncertainty,  $dW/W$ , of less than 0.3 with 75-m resolution below 1,000m altitude and 150-m resolution above that. Our experiments show that the conventional method did not achieve the accuracy lower bound, particularly in daytime (Fig. 3). In contrast, our method requires a slightly smaller uncertainty threshold (Table 1). Furthermore, our diagram indicates that a finer spatial resolution, 15 m, is available in low altitude ranges. The number of available data increases due to the finer spatial resolution in addition to the strict threshold (Fig. 4). Specifically, the 15-m resolution is suitable below an altitude of 500 m during the daytime and below 1500 m at night. The empirical vertical resolution, 150 m, is less effective than the 75-m resolution below 3,000 m. From these results, we can conclude that the optimum thresholds and vertical resolution determined by our method improve the quality of data more effectively than the conventional method.

### 5.3 Suggestions and implications

We propose two extensions of our method. The first is an application in the temporal direction. In practice, the desirable temporal resolution depends on the subsequent numerical weather models. Ideally, an optimum temporal resolution can be identified by simply replacing vertical resolutions with temporal resolutions. Tradeoffs of data observed with different spatial and temporal resolutions can also be compared and optimized. However, we need to keep in mind that results from temporal resolutions shorter or longer than 20-minutes should be evaluated by more frequent radiosonde observation. Another extension is using a different error metric because our method is compatible with other error metrics such as the mean absolute error.

Given that the amount of water vapor in the air varies with region and season, whether the same threshold is applica-

ble to observed data in different regions (or seasons) is a reasonable concern. Because the amount of data we used in this study is limited, it is difficult to assess exact issues that arise from the difference in seasons. However, previous studies have provided insight into this concern. Yoshida et al.<sup>9)</sup> conducted their observation in Nagasaki (western Japan) using the lidar instruments and empirical quality control method identical to those of Sakai et al. Even though the study area in Yoshida et al. was geographically distant from that of Sakai et al.'s, the results did not suffer from any significant quality control problems that arise in the conventional method. This indicates that rather than regional variation, the characteristics of lidar instruments may have more impact on measurement quality. The characteristics of lidar instruments, such as the power of the laser, area of the receiver telescope, and the detector's sensitivity, can vary greatly. Thus, differences in instruments should be assessed prior to lidar observation. By using the balance plot to show data obtained from parallel observation with two lidar instruments, we can visually identify and compare the performance of instruments. Our proposed diagram is robust and can potentially be expanded for lidar data quality control.

## 6. Concluding remarks

We devised a diagram that handles multiple information in a 2D plot to visualize the contradictory demands of accuracy and number of available data for lidar water vapor measurement. The optimum vertical resolution is also determined by using the diagram for comparison. A case study with real data shows the trade-off between the optimum vertical resolution and available data at various altitude ranges. The trade-off reflects the characteristics of our lidar instrument. This suggests our balance plot is useful for comparing the performance of lidar instruments. Lastly, we demonstrated that our robust and expansive balance plot can improve the quality control of geophysical data.

## Acknowledgements

We are grateful to Dr. Naoto Yokoya and Dr. Tatsumi Uezato for their valuable discussion on our study. We used radiosonde data measured by the Japan Meteorological Agency. The high resolution radiosonde data were provided by Mr. Shunsuke Hoshino at Aerological Observatory (current affiliation is numerical prediction division, information infrastructure department of Japan Meteorological Agency).

## References

- 1) D. N. Whiteman, K. Rush, S. Rabenhorst, W. Welch, M. Cadirola, G. McIntire, R. Felicita, A. Mariana, V. Demetrius, C. Rasheen, V. Igor, F. Ricardo, M. Bernd, S. Bernhard, L. Thierry, M. Stuart, and H. Vömel, "Airborne and ground-based measurements using a high-performance Raman lidar," *Journal of Atmospheric and Oceanic Technology*, **27** (11), 1781–1801 (2010). <https://doi.org/10.1175/2010JTECHA1391.1>
- 2) T. Dineev, V. Simeonov, Y. Arshinov, S. Bobrovnikov, P. Ristori, B. Calpini, M. Parlange, and H. Bergh, "Raman lidar for meteorological observations, RALMO—Part 1: Instrument description," *Atmospheric Measurement Techniques*, **6** (5), 1329–1346 (2013).
- 3) T. Leblanc, I. S. McDermid, and T. D. Walsh, "Ground-based water vapor Raman lidar measurements up to the upper troposphere and lower stratosphere for long-term monitoring," *Atmospheric Measurement Techniques*, **5** (1), 17–36 (2012).
- 4) V. Wulfmeyer, R. M. Hardesty, D. D. Turner, A. Behrendt, M. P. Cadeddu, P. Di Girolamo, P. Schlüssel, J. Van Baelen, and F. Zus, "A review of the remote sensing of lower tropospheric thermodynamic profiles and its indispensable role for the understanding and the simulation of water and energy cycles," *Rev. Geophys.* **53**, 819–895, doi:10.1002/2014RG000476, 2015.
- 5) T. Sakai, T. Nagai, T. Izumi, S. Yoshida, and Y. Shoji, "Automated compact mobile Raman lidar for water vapor measurement: instrument description and validation by comparison with radiosonde, GNSS, and high-resolution objective analysis," *Atmospheric Measurement Techniques*, **12** (1), 313–326 (2019).
- 6) D. N. Whiteman, "Examination of the traditional Raman lidar technique. II. Evaluating the ratios for water vapor and aerosols," *Applied Optics*, **42**, 2593–2608 (2003).
- 7) R. J. Dirksen, M. Sommer, F. J. Immler, D. F. Hurst, R. Kivi, and H. Vömel, "Reference quality upper-air measurements: GRUAN data processing for the Vaisala RS92 radiosonde," *Atmos. Meas. Tech.*, **7**, 4463–4490, doi:10.5194/amt-7-4463-2014 (2014).
- 8) N. Kizu, T. Sugidachi, E. Kobayashi, S. Hoshino, K. Shimizu, R. Maeda, and M. Fujiwara, Technical characteristics and GRUAN data processing for the Meisei RS-11G and iMS-100 radiosondes, GRUAN Technical document, 5. [https://www.gruan.org/gruan/editor/documents/gruan/GRUAN-TD-5\\_MeiseiRadiosondes\\_v1\\_20180221.pdf](https://www.gruan.org/gruan/editor/documents/gruan/GRUAN-TD-5_MeiseiRadiosondes_v1_20180221.pdf), 2018. (Accessed

2022.8.16)

- 9) S. Yoshida, T. Sakai, T. Nagai, Y. Ikuta, Y. Shoji, H. Seko, and K. Shiraishi, Lidar Observations and Data Assimilation of Low-Level Moist Inflows Causing Severe Local Rainfall Associated with a Mesoscale Convective System, *Monthly Weather Review*. **150** (7), 1781–1798 (2022).

### Author Introduction

Corresponding author: Atsushi Takahashi

Doctor of Science, September, 2019, Kyoto University, Kyoto Japan. RIKEN AIP Postdoctoral researcher (2019–present)