

Clustering of weak fluorescence spectra from bioaerosol in air using laser-induced fluorescence lidar

DAWEI TANG,¹ ZHEKAI LI,² AND HAIYUN XIA^{1,2,3,*} 

¹*School of Earth and Space Science, University of Science and Technology of China, Hefei 230026, China*

²*School of Atmospheric Physics, Nanjing University of Information Science and Technology, Nanjing 210044, China*

³*Institute of Lidar Technology, GuangZai Co., Ltd., Hangzhou 310005, China*

**hsia@ustc.edu.cn*

Abstract: Laser-induced fluorescence (LIF) lidar is the main method in bioaerosol remote detection. However, its current applications are limited to specific scenarios due to weak fluorescence signals in ambient environments. In this paper, a spectrally resolved lidar system based on LIF spectroscopy was developed for bioaerosol detection. This lidar system was deployed at an elevation angle of 30° in Nanping City, Fujian Province, China from April to July 2024. During observation, weak fluorescence was detected, and its spectral contamination risk by the N₂ Raman signal was discussed and resolved. In this paper, we introduced a clustering method to process and analyze the weak fluorescence spectra. This method successfully separates fluorescence spectra into 3 clusters under weak fluorescence intensities. The possible types of each cluster are discussed and inferred based on their spatiotemporal distribution combined with backtrajectory results and meteorological data. This method enables aerosol classification even under weak fluorescence intensities, providing a new approach to atmospheric aerosol analysis.

© 2025 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Bioaerosol is an important subset of atmospheric aerosol. It is defined as aerosols originating from biological materials [1]. At present, research on bioaerosols usually focuses on its subset, the primary biological aerosols (PBAs). The PBAs are defined as solid particles derived from biological organisms and directly emitted from the surface into the atmosphere [2,3]. Bioaerosols are ubiquitous and diverse, including bacteria, fungi, algae, viruses, pollen, and debris from biological tissue. Their diversity of origins results in a wide range of particle sizes, from a few nanometers to several hundred micrometers. Bioaerosols have significant impacts on the atmospheric environment. They can act as cloud condensation nuclei and ice nuclei to participate in atmospheric physical processes, and even affect regional climate [4]. In addition, the transport of bioaerosols can spread disease and allergy and even threaten human health as a biological weapon [5,6]. Therefore, the detection and monitoring of atmospheric bioaerosols are crucial for applications like pollen allergy warning, biological weapon warning, and the study of cloud-aerosol-precipitation interactions.

At present, some bioaerosol sampling and detection techniques based on laser-induced fluorescence (LIF) mechanisms have been widely applied [7–10]. LIF uses lasers to excite fluorescent components such as tryptophan and riboflavin in bioaerosol particles. Then the excited fluorescence spectra are detected to recognize between bioaerosols and non-bioaerosols [11]. LIF lidar techniques combine LIF with lidar technology to directly excite airborne bioaerosols using ultraviolet (UV) laser without sampling. Thus it enables long-term and large-scale standoff online detection of bioaerosols [12–14]. At present, multiple LIF lidar systems have been developed and tested. However, the intensities of fluorescence are much lower than that of elastic backscatter

and inelastic Raman signal. Even worse, the fluorescence spectra in the LIF lidar system have a continuous shape, usually covering a wide spectra range of about 200 nm. This property results in a relatively high background signal. Unlike Raman lidar [15–18], the background signals in LIF lidar cannot be filtered out by applying narrow bandwidth optical filters. By now, the performance of LIF lidar is still limited and developed lidar systems are mostly used in experimental conditions. In these conditions, bioaerosols of known categories are manually released and diffused in semi-enclosed chambers [19–22] or open fields [23–26]. Recently there have been reports about detecting airborne bioaerosols using LIF lidar under natural conditions [27–32]. However, their spectra analyses are limited to high-concentration events, such as pollen spreading in pollen season or dust transport. More reports from ordinary scenarios, especially in the most frequent cases of weak fluorescence signals, are still needed. These studies can help explore the potential of LIF lidar techniques in fluorescence detection, classification, and recognition.

In this paper, a spectral-resolving lidar system based on LIF techniques is developed and deployed. This system emits a 355-nm 3rd harmonic laser light from a Nd:YAG laser to excite the fluorescence from atmospheric aerosols. The backscattering fluorescence signal is received by a 12-inch telescope, then dispersed by a spectrometer, and finally detected by a 32-channel multi-anode PMT array. This system is deployed from April to the beginning of July 2024. During the observation, the relative intensity of fluorescence compared with Raman signal is discussed, and the contamination risk is addressed. Then the conventional spectra integration is used to visualize the fluorescence variation during our observations. Finally, we performed clustering analysis to the fluorescence spectra during the observation, and a 3-cluster solution is adopted. Combined with HYSPLIT backtrajectory and local meteorological data, we inferred the possible types of each cluster. The analysis results of fluorescence spectra using LIF lidar demonstrates the its potential for detection, information extraction, classification, and recognition in weak fluorescence scenarios.

2. Materials and methods

2.1. Site and meteorological data

The observation site was located on the top floor of the National Center of Carbon Metrology Building in Yanping District, Nanping City, Fujian Province, China (26.59°N, 118.27°E). Nanping City is located in the Mount Wuyi region in the north of Fujian Province, with a forest coverage rate of 80% [33]. The observation site is located on the bank of Minjiang River, the biggest River in Fujian Province, about 10 km away from the city center. The surrounding region is mainly covered with natural vegetation and greenery, as well as a small number of rural residential areas. The traffic flow during the measurement period was extremely low. According to the hourly air quality monitoring data released by the local government [34], the Air Quality Index (AQI) during the observation period was all lower than 50, with PM_{2.5} data all below 30 $\mu\text{g m}^{-3}$. According to Chinese standards of air quality, the air quality during the observation period is classified as excellent. The information above indicates that the observation is unlikely to be affected by anthropogenic aerosol which could be fluorescent interferences during observation.

2.2. Lidar setup

The lidar system is diagrammed as Fig. 1. The system was mainly composed of an injection-seeded 3rd harmonic Nd:YAG laser (Innolas Spitlight EVOIII), a telescope (Meade LX200 ACF12), and a 32-channel multispectral lidar detector (Licel SP32HR). The laser emitted a 355 nm laser at the frequency of 100 Hz, with a single pulse energy greater than 200 mJ, and an emission divergence angle of 1 mrad. The emitted laser was expanded by a 10× beam expander, reducing the laser divergence angle to 0.1 mrad. Then the expanded laser beam was redirected into the atmosphere

by a reflector at an elevation angle of 30° . Therefore, the laser detection heights were half of the detection range. The telescope is a Schmidt-Cassegrain reflector with a diameter of 12 inches (304.8 mm) and a focal length of 3048 mm. The backscatter signal was focused by the telescope to its rear, where an optic fiber bundle end was installed and connected with the detector. An optical filter was placed in front of the fiber bundle entrance. We used two types of filters in different observation periods: a 355 nm notch filter (Edmund #39-387) and a 400 nm long-pass filter (Thorlabs FELH0400). The 355 nm notch filter has an optic density over 6 at 355 nm to prevent elastic signal leakage. The 400 nm long-pass filter provides an optical density larger than 5 at 200–400 nm. Thus, it can suppress the signal leakage from elastic scattering signals at 355 nm, as well as the first-order vib-rotational Raman signals of O_2 and N_2 at 376 and 386 nm, respectively. In the spectral range of 400–600 nm, both filters have transmission above 94%. It should be mentioned that there is still a risk of elastic backscatter leakage when the backscatter is strong. At the entrance end of the fiber bundle, 47 fibers with 100 μm core diameter and numeric apertures of 0.12 are arranged in a circular shape. This arrangement forms an effective core diameter of 1.15 mm, which leads to a receiving field of view angle of 0.38 mrad. The distance between the emitted laser and the telescope, as well as the low divergence angle of the laser, and the low field of view angle of the telescope, creates a blind zone of about 500 m in the range. In addition, signals within the range of 800 m are still weak. As a result, the backscattered signal below the height of 250 m is undetectable in our observation results.

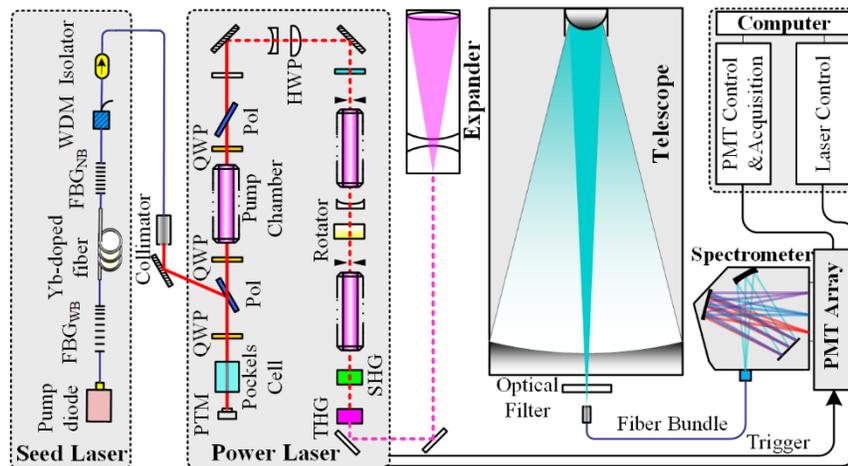


Fig. 1. Diagram of fluorescence lidar system in the paper. The magenta dashed lines refer to the 355 nm laser emitted, the solid red lines the emitted light by seed laser, the dashed red line the emitted 1064 nm light by power laser, the cyan area the collected backscattered signal, and the dark blue lines refer to fibers connecting seed laser and power laser or the fiber bundle that connects the telescope and detector. Black lines refer to the cables for data communication and device control. FBG: Fiber Bragg Grating; WB: Wideband; NB: Narrowband; WDM: Wavelength-Division Multiplexer; PTM: Piezo-Transducer-driven Mirror; QWP: Quarter-wave Plate; HWP: Half-wave Plate; SHG: Second-harmonic Generator; THG: Third-harmonic Generator; Pol: Thin film polarizer.

At another end of the fiber bundle, the multi-core fiber was linearly spaced and directed the receiving light into the entrance slit of a Czerny-Turner spectrometer (Oriel MS125). In the spectrometer, the backscattered light was dispersed by an optical grating with $1200 \text{ lines mm}^{-1}$ (Newport 77411). Then the dispersed light was detected by a multi-anode PMT array with 32 detection channels (marked with channel 0–31), and recorded in photon-counting mode. During

the observation, the center wavelength of the spectrometer was set to 475 nm, the grating provided a spectra resolution of 6.2 nm mm^{-1} . The 32 PMT anodes were linearly spaced at the interval of 1 mm. Hence, the spectral resolution and spectral detection width during the observation is 6.2 nm per channel. The overall detection spectral range of the spectrometer is from 378.9 nm (channel 31) to 571.1 nm (channel 0). The center wavelength of each channel is listed in Table 1.

Table 1. Channel numbers and their corresponding center wavelength in the PMT array

Channel	Center Wavelength (nm)						
0	571.1	8	521.5	16	471.9	24	422.3
1	564.9	9	515.3	17	465.7	25	416.1
2	558.7	10	509.1	18	459.5	26	409.9
3	552.5	11	502.9	19	453.3	27	403.7
4	546.3	12	496.7	20	447.1	28	397.5
5	540.1	13	490.5	21	440.9	29	391.3
6	533.9	14	484.3	22	434.7	30	385.1
7	527.7	15	478.1	23	428.5	31	378.9

During the observation, the range resolution was 40 m, and the range bin number was set to 1500. Before July, the accumulated lidar signal was averaged and recorded after 18000 laser pulses (corresponding to a time resolution of 3 min). In July, it was set to 60000 laser pulses (corresponding to 10 min). Before the observation, the sensitivity of the multi-channel detector was calibrated by placing a spectrally pre-calibrated LED in front of the telescope [35]. The raw lidar data were processed by correcting PMT array nonuniformity and dead time effects, followed by background signal subtraction. The background signal in this observation is defined as the mean signals in the last 300 range bins plus its one-fold standard deviation. This approach reduces the interference from background spectra. Due to the wavelength-dependent effects, the optical attenuation in the return path varied across channels. This difference is mainly caused by molecule and aerosol particles. Optical attenuation contributed by cloud and fog exhibited near zero Angstrom Exponent, resulting in nearly uniform attenuation across all channels [36]. Here, the attenuation difference induced by the molecule and aerosols are compensated. For molecule, it is compensated according to 1976 US standard atmosphere [37]. The aerosol properties near the observation site, including its optical depth and angstrom exponent, were derived from aerosol model product supplied by Japan Aerospace Exploration Agency (JAXA) [38]. Additionally, the vertical distributions of aerosols during the observation are derived from a coherent doppler lidar system located 30 m away from our LIF lidar system [39,40]. During the observation, the 355 nm notch filter was first used in 6 observations from April 19th to May 25th. Then it was replaced by a 400 nm long-pass filter in the other 6 observations from May 27th to July 3rd. All of these observations were performed at night to prevent solar radiation. The start and end times, as well as the average AQIs, temperature, and relative humidity of each observation, are recorded in Table 2.

2.3. Contaminations of N_2 Raman signal leakage on weak fluorescence detection

To exclude the effect of overlapping factor and range square attenuation, and to make the fluorescence signal intensities in different range comparable, the obtained spectra of each observation are normalized to the channel with the strongest signal. The normalized intensity in each channel is calculated as

$$I(\lambda_i)_{normalized} = \frac{I(\lambda_i)}{\max(I(\lambda))}$$

Table 2. Basic parameters during each observation period, including average air quality index (AQI), temperature, and relative humidity (RH). The observations using a notch filter in the table are marked with the prefix “S” in their observation number.

Observation Num	Start Time (UTC + 8)	End Time (UTC + 8)	Average AQI	Average Temperature (°C)	Average RH
S1	2024.4.19 23:19	2024.4.20 00:05	28	20.3	98%
S2	2024.4.25 03:31	2024.4.25 04:39	49	18.3	95.7%
S3	2024.4.27 02:55	2024.4.27 03:22	28.7	20.8	96.3%
S4	2024.5.07 21:29	2024.5.07 22:20	36.3	23.3	78.3%
S5	2024.5.11 01:41	2024.5.11 02:37	31.3	19.5	90.7%
S6	2024.5.23 21:27	2024.5.24 01:10	20.8	22.5	95.8%
1	2024.5.27 23:25	2024.5.28 02:57	16.4	23.3	97%
2	2024.6.20 23:34	2024.6.21 02:02	24.8	25.8	92.8%
3	2024.6.21 20:42	2024.6.22 00:02	21	27.2	81.8%
4	2024.6.28 21:04	2024.6.29 03:02	19.8	26.3	94.8%
5	2024.7.01 22:45	2024.7.02 04:27	16.5	26.0	87.9%
6	2024.7.02 23:08	2024.7.03 01:07	11.3	24.8	95%

where $I(\lambda_i)$ and $I(\lambda_i)_{normalized}$ is the signal intensity before and after normalization in each observation. $\max(I(\lambda))$ is the maximal signal intensity in all detection channels in each observation. For observations using the notch filter, they are normalized to channel 30, corresponding to the first-order vib-rotational N_2 Raman signal at 386 nm. For the long-pass filter, they are normalized to channel 27 corresponding to H_2O Raman Signal at 407 nm. The normalized spectra within the height range of 400–1000 m during each observation are accumulated and averaged. This height range provides strong signals and representative spectra for each observation. The average spectra are shown in Fig. 2. As illustrated, channel 30 and channel 27 exhibit the strongest signals in notch filter observations. In contrast, for long-pass filter observations, the signals in channel 30 are suppressed. When excluding the two channels mentioned above, channel 26 shows the strongest signals in both types of observations. This is because the wavelength of the H_2O Raman peak lies on the edge between channel 26 and 27. Besides, a small spectra peak is in channel 24 whichever filter is used. This peak is due to the N_2 Raman overtone at 424.4 nm [41]. For observations using the notch filter, Fig. 2(a) shows that during the first 6 observations, the H_2O Raman Signal in channel 27 is consistently about one order of magnitude lower than the N_2 Raman signal in channel 30, regardless of temperature and humidity variations. This result aligns with previous studies [28,42]. Besides, multiple periodic small peaks in channel 0–22 in Fig. 2(a) can be seen, including channel 21, 16, and 10 (marked with blue arrows). These peaks are most prominent in the observations from the nights of April 19th and May 23rd, when signals in channels 0–22 are low. These peaks are due to Lyman Ghosts in grating [43,44], and false signals from N_2 Raman signal leakage. This reveals that under weak fluorescence scenarios, even with suppressed elastic backscattered signals, the relatively weaker N_2 Raman signals

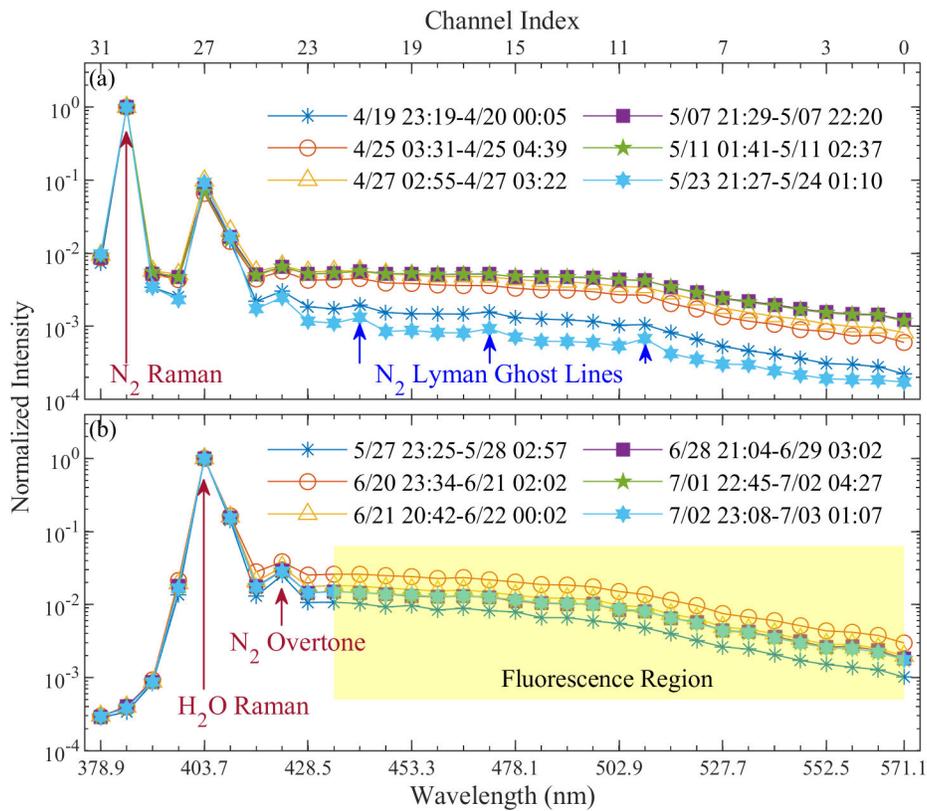


Fig. 2. The average of normalized spectra at the height of 400–1000 m during each observation using notch filter (a) or long-pass filter (b). Before averaging, the spectra are normalized to the channel with the strongest signal, which is channel 30 in (a) and channel 27 in (b). The x-axis refers to the channel number (on the top of (a)) and its corresponding center wavelength (bottom of (b)). The y-axis in (a) and (b) refers to the normalized intensity. The fluorescence region (channel 0–22) in (b) is marked with yellow shading. The normalized intensity in each channel is marked with a point. The observation time shown in this figure is in the time zone of UTC + 8. The Lyman Ghost Lines are marked with blue arrows.

can still contaminate the fluorescence spectra. Although N₂ Raman signals provide additional atmospheric aerosol information, they should be suppressed due to their contamination in weak fluorescence detection. In subsequent observations, a long-pass filter was applied to suppress potential contamination sources, the elastic, N₂ and O₂ Raman signal was all suppressed to ensure clean fluorescence detection. Therefore, due to the N₂ Raman signal leakage, notch filter observations were excluded from further fluorescence analysis.

From the observation results shown in Fig. 2(b), the 400 nm long-pass filter successfully suppressed the vib-rotational Raman signal of N₂ at 386 nm which would have appeared in channel 30, as well as the derived Lyman Ghosts. Therefore, channel 0–22 are defined as the fluorescence region as they exclude the interference of Raman signals. In all of the observation periods using the long-pass filter, the fluorescence shows a similar trend: fluorescence intensity decreases with increasing wavelength. However, the relative fluorescence intensities vary across observation periods: the intensity during the second observation from Jun 20th night to the 21st early morning shows the highest relative fluorescence intensity, while the first observation from May 27th night to the 28th early morning shows the lowest. In summary, the intensity in a

single channel in the fluorescence region is about 2 orders of magnitude lower than that of the H₂O Raman signal. Given the similar temperature and humidity of each observation shown in Table 2. We assumed that the N₂ Raman signal can be at least one order of magnitude higher than H₂O Raman signals in observations using the long-pass filter. Based on this assumption, we estimate that the intensity of a signal fluorescence channel is roughly 3 orders of magnitude lower than the N₂ Raman signal. In previous observations using similar lidar system reported in northwest China [32], the polluted dust aerosol has the highest fluorescence intensity during their observations. And its fluorescence intensity in each detection channel is more than 2% of N₂ Raman signal. The dust aerosol has the lowest fluorescence intensity which is more than about 1% of N₂ Raman signals. And the background and air pollutant aerosol during the observation has the fluorescence intensity of between these two types of aerosols mentioned above. During another similar observation in Japan [28], the low clouds has the maximum fluorescence intensity in each channel, which is near 1 order of magnitude lower than N₂ Raman Signal. The second and third highest fluorescence intensity are Asian dust aerosols and air pollution aerosols, which are both more than 1% of N₂ Raman signal in each of fluorescence channel. In their observations, the observed minimum relative fluorescence intensity is about 3 orders of magnitude lower than N₂ Raman signal, which is similar to observations introduced in this paper. In the observations in Japan, aerosols exhibit such level of fluorescence intensity are classified as non-fluorescent aerosols. Besides, in the two observations mentioned above, the dust aerosols, air polluted aerosols and low clouds can be distinguished from each other by their distinctive fluorescence spectra shapes. However, in this paper, the observed fluorescence spectra in each observation period show similar fluorescence shapes and cannot be distinguished simply from their fluorescence spectra shapes.

2.4. Fluorescence spectra processing

In some LIF lidar systems, the backscattered fluorescence is detected by the method of spectra integration [22,41,45]. Although this method reduces the spectra resolution, it can give a better visualization of spectra variation and an increased detection range. Therefore, the spectra integration method is applied in this study. The fluorescence spectra are divided into 2 parts: channel 0–16 is area 1 and channel 16–22 is area 2. Then the proportions of the integrated area of each part in the total fluorescence area (channel 0–22) are calculated as:

$$P_1 = \frac{\int_{\lambda_{16}}^{\lambda_0} I(\lambda_i) d\lambda}{\int_{\lambda_{22}}^{\lambda_0} I(\lambda_i) d\lambda}$$

and

$$P_2 = \frac{\int_{\lambda_{22}}^{\lambda_{16}} I(\lambda_i) d\lambda}{\int_{\lambda_{22}}^{\lambda_0} I(\lambda_i) d\lambda}$$

where $I(\lambda_i)$ is the signal intensity in channel number i ; λ_i is the corresponding central wavelength; P_1 and P_2 are the proportion of each part in total fluorescence area. The integration procedure uses trapezoidal integration method. This division method is based on the maximum variation range of proportion values, defined in this paper as the differences between the 90% and 10% percentiles. In addition, fluorescence data with zero intensity in any of its channels are excluded to eliminate low signal intensity data.

To explore the potential lost spectral information, and evaluate the system's classification potential for fluorescence spectra, further analysis of the observation data is needed. For this purpose, the data clustering method is introduced in the analysis of the obtained fluorescence signals. Due to their low signal-to-noise ratio, fluorescence signals at longer distances are more susceptible to noise, and even their feature distribution can be obscured. So, firstly, the

fluorescence data for clustering are limited to the range of 800–2000 m (corresponding to the height of 400–1000 m), and fluorescence spectra with zero intensity in any of its channels are excluded like before. Secondly, to exclude the influence of total fluorescence intensity, area normalization is applied to all fluorescence spectra. The normalized intensity of each channel in the fluorescence region is calculated as:

$$I(\lambda_i)_{normalized} = \frac{I(\lambda_i)}{\int_{\lambda_{22}}^{\lambda_0} I(\lambda_i) d\lambda}$$

In the last step, Hierarchical Agglomerative Clustering (HAC) is performed on the normalized fluorescence spectra. This method initially treats each spectrum as an independent cluster and iteratively merges the pair of clusters with the highest similarity until all data are grouped into a single cluster. During this process, a dendrogram is generated [46–48]. The HAC is an unsupervised clustering approach with the advantage of a single clustering dendrogram and no randomness in clustering results. In this study, inter-cluster similarity is measured using Euclidean distance and Ward linkage, which are commonly used metrics.

The optimal number of clusters is determined using the Calinski-Harabasz index (CHI) and Davies-Boulding index (DBI) [49,50]. The CHI evaluates the clustering results by comparing intra-cluster variance with inter-cluster variance, while the DBI evaluates them by comparing the intra-cluster distance with the inter-cluster distance. A good clustering result should have a high CHI and a low DBI.

To better visualize and evaluate the clustering results, the principal component analysis (PCA) [51] is applied to the original normalized spectra. PCA is a feature extraction and dimensionality reduction method. In PCA, the original feature space is rotated, and a new set of variables called principal components (PCs) is generated. These PCs are created through a linear transformation of the original variables and are sorted in descending order based on the variance of their corresponding scores in the PC space. By discarding the last few PCs, the original data can be reduced in dimensionality while retaining most of their feature information.

3. Results and discussion

3.1. Variation of fluorescence spectra during each observation

Figure 3 shows lidar results during long-pass filter observations. It includes the range-corrected H₂O Raman signal in channel 27 (Fig. 3(a)), the total fluorescence signal (sum of channel 0–22, Fig. 3(b)), atmospheric transmission caused by aerosol and molecule in the fluorescence region at the height of 2 km (Fig. 3(e)), as well as the proportion of each fluorescence part (Fig. 3(c, d)). It can be seen from Fig. 3(b) that the maximum detection range for fluorescence during observations is generally at the height of 2–3 km. Sometimes it reaches up to 3.5 km (in the fourth and sixth observations) and sometimes it is limited by strong optical attenuation by clouds and near surface fogs. During observations, the strong fluorescence signal is mainly distributed below 1.5–2 km, which is a typical summertime nocturnal residual layer height. Fluorescence signal intensity above this height will rapidly weaken to near the level of the background signal. Figure 3(e) shows that the atmospheric transmission difference between channel 0 and channel 22 at the height of 2 km varies from 6% to 9%.

During the first observation from May 27th to 28th, the fluorescence intensity exhibits a two-layer structure. After 0:00 on the 28th, a fluorescent aerosol layer appeared at a height of 1.5–2 km. Then this layer gradually descends and eventually merges with near-surface fluorescence layer by the end of this observation. The upper layer shows a higher proportion in area 1. Even after the two layers merge, the spectral difference between the two layers remains. For the fluorescence spectra under 1 km height, the first observation has an obvious difference with the later 5 observations for its highest proportion of area 2. In the second and

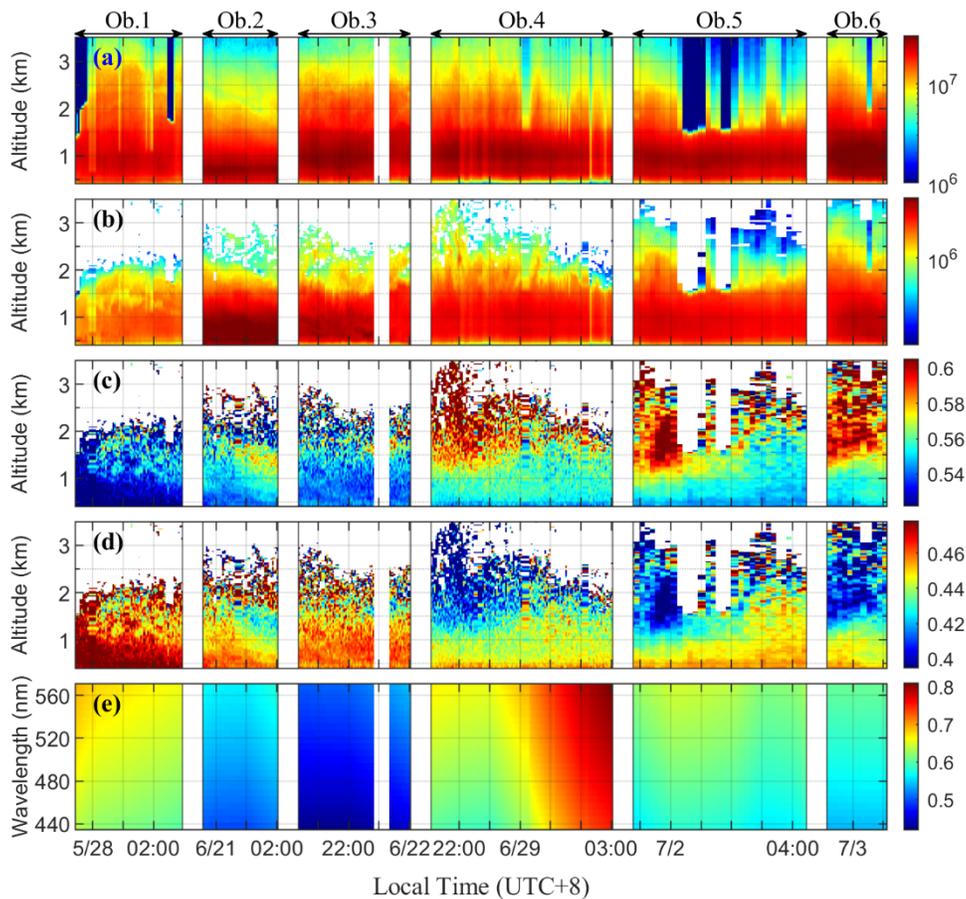


Fig. 3. Lidar data obtained during the observations. (a): The range-corrected H_2O Raman signal intensity in channel 27. (b): the sum of range-corrected signal intensity in channel 0–22 (fluorescence region). (c): proportion of the integrated area of channel 0–16 (area 1, 471.9–571.1 nm) in the total integrated area of the fluorescence region. (d): proportion of the integrated area of channel 16–22 (area 2, 434.7–471.9 nm) in the total integrated area of the fluorescence region. (e) the atmospheric transmission caused by aerosol and molecule in channel 0–22 at the height of 2 km during observation.

third observations from Jun 20th to 22nd, no two-layer structures are observed in the observation of the total fluorescence signal. However, the spectral difference among different heights still exists. Especially in the second observation, three layers can be found in about 0–1 km, 1–2 km, and 2–2.5 km. Among them, the middle layer has the highest proportion of about 0.57 in area 1, while the other two layers show similar spectra fingerprints. From the observation of the spectral, the middle layer appears at about 1700 m height and keeps downward transporting during the observation. By the end of the second observation, the bottom of the middle layer reaches its lowest height at about 1000 m.

Although the second observation shows the strongest fluorescence intensity, the last three observations have the maximum detection height. The extended detection range during these observations can be attributed to the existence of fluorescent aerosols at high altitudes. These aerosols are characterized by enhanced fluorescence signals and the highest area 1 proportion among all observations. During the fourth and fifth observations, the aerosol layers at high

altitudes show similar behavior: they occur at the beginning of observations, and gradually thin out, becoming difficult to identify by the end. During this period the upward diffusion of low-altitude aerosols can be identified by decreased area 1 proportion at 1.5–2 km height. This is particularly evident in the fifth observation, where the fluorescence intensity also increases. After 0:30 on July 1st during the fifth observation, lidar signals are strongly attenuated, which should be attributed to the cloud layer. After 2:00, the attenuation disappears but the fluorescence intensity at 1.5–2 km height along with the H₂O Raman signal decreases significantly compared to the beginning of the observation. The appearance of clouds following the high fluorescence area suggests that fluorescent bioaerosols can contribute to cloud formation as cloud condensation nuclei. The last observation on July 2nd – 3rd shows similar behavior to the fourth and fifth observations, including long detection ranges, a two-layer fluorescence structure and gradually upward-moving low-altitude fluorescent aerosols. Although the spectra integration method demonstrates its capability in fluorescence detection and classification, the loss of detailed spectral information limits the potential of aerosol classification.

3.2. Clustering results of low altitude fluorescence spectra

To choose to optimal number of clusters, the cluster results with cluster number ranges from 2 to 10 are evaluated using CHI and DBI, and the evaluation results are presented in Fig. 4(a). For both CHI and DBI evaluation, the 2-cluster solution is optimal, 3-cluster solution is the second optimal solution and have close CHI and DBI value with 2-cluster solution. Given the close DBI and CHI values and the need for more detailed clustering analysis, the 3-cluster solution is adopted for further analysis. Figure 4(b) presents the clustering dendrogram of the 3-cluster solution. It can be seen from the dendrogram that cluster 2 and cluster 3 have smaller linkage, indicating higher similarity between these two clusters among the three clusters. In the 2-cluster solution using HAC, the clustering results will be cluster 1, and a new cluster formed by merging cluster 2 and cluster 3.

In this study, PCA results show that the first two principal components explain 42.61% and 24.44% of the variance in original data, respectively, while the third principal component interprets 4.25% of the variance in original data. The PCA loadings of the first two PCs are shown in Fig. 4(c). From the figure, it can be seen that at this height range, the long wavelength part of the fluorescence (channel 0–10) plays a minor role in PCA. Besides, PC1 is primarily associated with channel 16–22, while PC2 is linked with channel 10–15. This means that the higher score in a certain principal component will have higher portion in its corresponding wavelength region. It should be mentioned that these conclusions are specific to current fluorescence spectra, and may not apply to spectra at other heights. Figure 4(d, e, f) shows the score distribution of the three clusters in the first two PCs (PC1 and PC2) after PCA. The three clusters are well separated. For PC1 scores, cluster 1 and 3 are distributed with centers of 0.0023 and -0.0011, respectively, while cluster 2 is likely to have bimodal distribution with center of -0.0011 and 0.0001. For PC2 scores, three clusters are distributed with centers of -0.0003, 0.0005 and -0.0017, respectively.

Figure 5 shows the spatiotemporal distribution of clustering results. It can be seen that in the first observation on May 27th – 28th, most of the fluorescence spectra are classified as cluster 1. During the second observation on Jun 20th – 21st, almost all fluorescence spectra are classified as cluster 2. Besides, clusters 2 and 3 dominate the fluorescence spectra in other observations. The spatiotemporal distribution difference between cluster 1 and cluster 2/3 can also be seen in our results using the spectra integration method. From the perspective of height, the majority of cluster 3 appears at the height of 400–600 m, while the fluorescence spectra at higher altitudes are mainly cluster 2.

To identify the origin of different clusters, the 24-hour backward trajectories are performed for the air masses above the observation site during different observation periods. The backward trajectories use the HYSPLIT model [52] and GDAS 1° meteorological data [53]. Trajectory

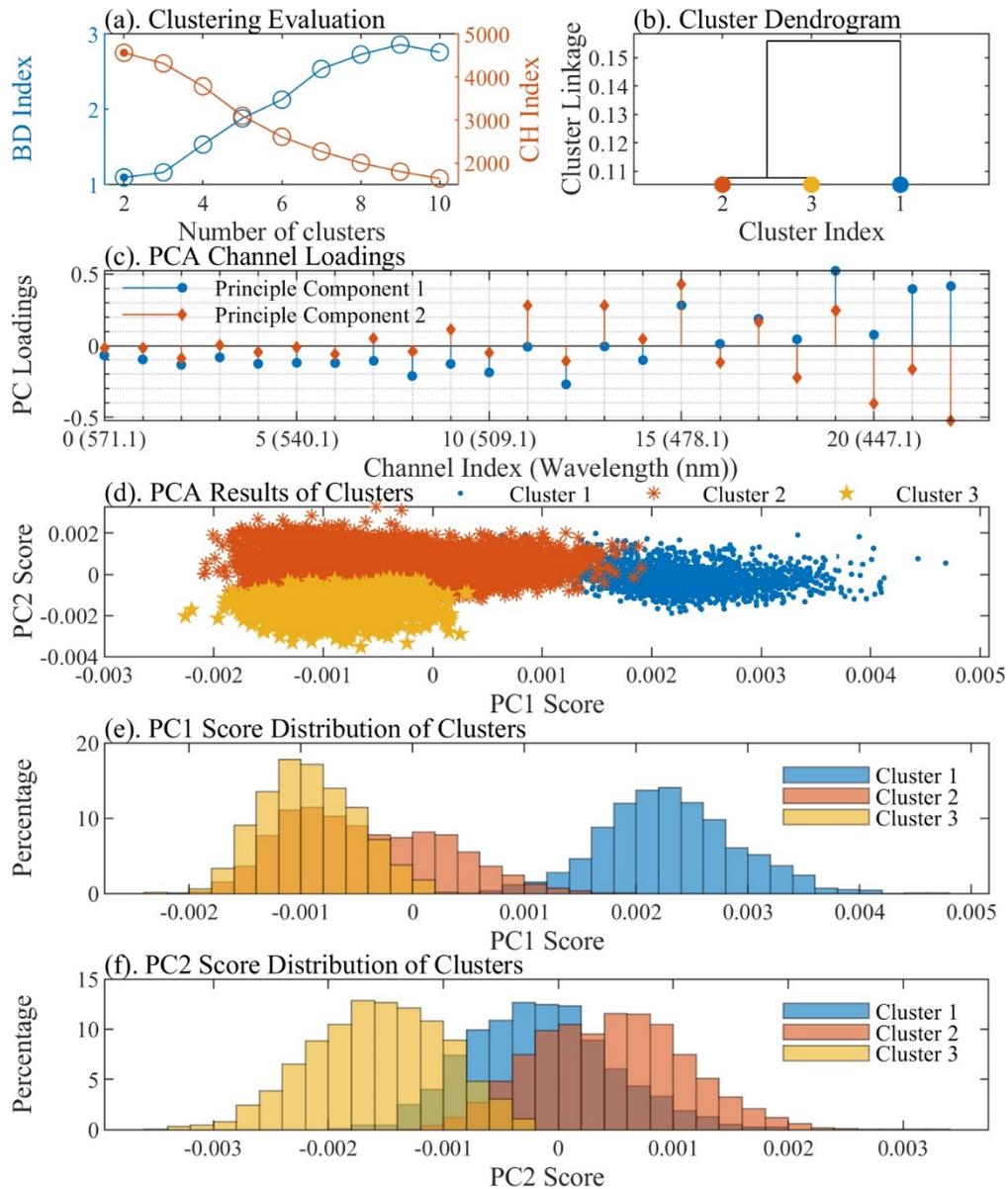


Fig. 4. (a): Davies-Bouldin Index (blue line) and Calinski-Harabasz Index (red line) when number of clusters ranges from 2 to 10. (b): dendrogram of clusters, blue, yellow, and red dots refer to cluster 1, 2, and 3 respectively. The vertical axis refers to the linkage between clusters. (c): the PCA loadings of the first two principal components in channels and corresponding wavelengths. (d): two-dimension scatter plots of PC1 and PC2 scores for three clusters (blue for cluster 1, red for cluster 2, and yellow for cluster 3). (e) and (f) refer to PC1 and PC2 scores histograms for cluster 1 (blue), cluster 2 (red), and cluster 3 (yellow), respectively. The vertical axes in the two histograms refer to the percentage of each bin in the total data numbers of the clusters.

starting times are set to the midpoint of each observation period. The starting heights are set to 400 m and 700 m. The trajectory results shown in Fig. 6 reveal that the air masses during

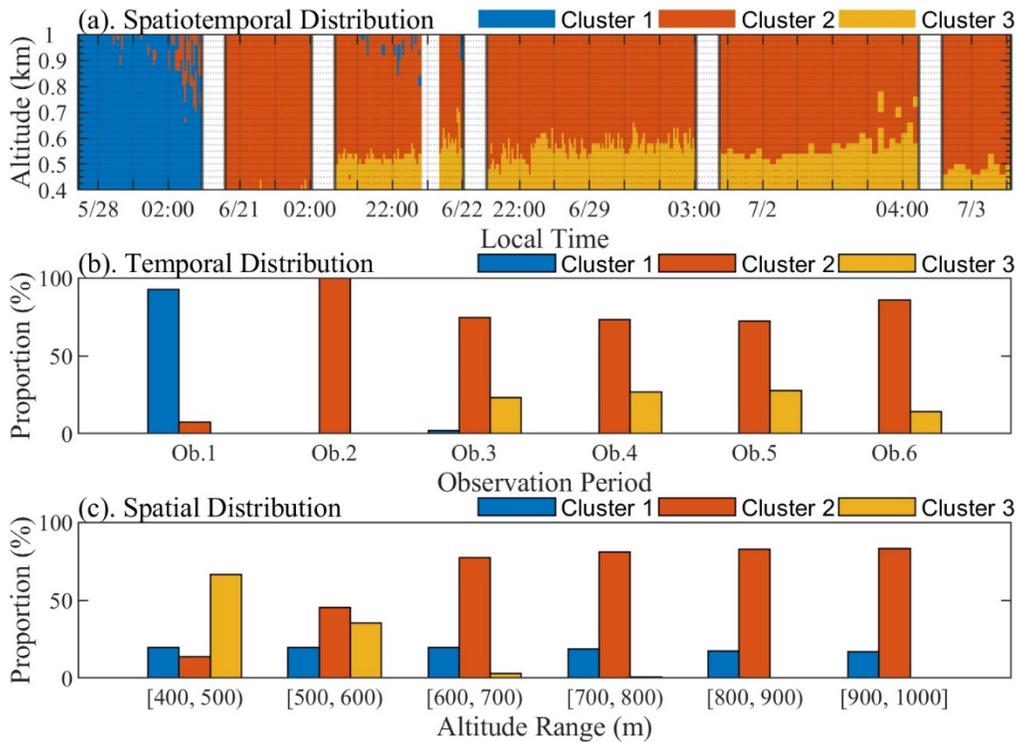


Fig. 5. (a): clustering results for normalized fluorescence spectra and their spatiotemporal distribution. (b) and (c) refer to the histogram of temporal (b) and spatial (c) distribution for cluster 1 (blue), cluster 2 (red) and cluster 3 (yellow), respectively. The vertical axes in the two histograms refer to the percentage of each bin in the total data numbers of the clusters.

most observations have similar transport paths, originating from coastal regions in Fujian and Guangdong Province in the south and southwest of the observation sites. The exception is the first observation on May 27th – 28th, during which the airmass above the sites originates from the inland Mount Wuyi region in the northwest of the site. Trajectories that start at the same time but at different heights still have similar transporting paths. The different sources of air masses are consistent with different time distributions of cluster 1 and cluster 2/3. This suggests that the fluorescence spectra in cluster 1 are related to aerosols from the inland region in the northwest of the site, while cluster 2/3 are related to aerosols originating from oceanic and coastal areas in the south and southwest of the site. Regarding the difference between cluster 2 and 3, their vertical distributions are notable: cluster 2 consistently appears above cluster 3 when both are present, with a boundary line between them at 400–600 m. This height is close to the height of the nocturnal stable boundary layer in southeastern China [54,55] and in other regions [56]. Previous lidar studies [57–59] have reported aerosol-derived nocturnal stable boundary layer heights of 400–600 m, particularly in Xiamen City [59], which is also located in Fujian Province. However, the two-layer structure is absent during the first two observations.

To investigate the reason for the absence of the two-layer structure, the local meteorological data from ERA5 hourly data on single levels [60] are analyzed and plotted in Fig. 7. The figure shows that the surface sensible heat flux during all observations propagates downwards. Since sensible heat flux propagates from high temperature to low temperature, this indicates the presence of temperature inversion layers above the surface during all observations. Its presence is a typical characteristic of the nocturnal stable boundary layer [61]. During the first two

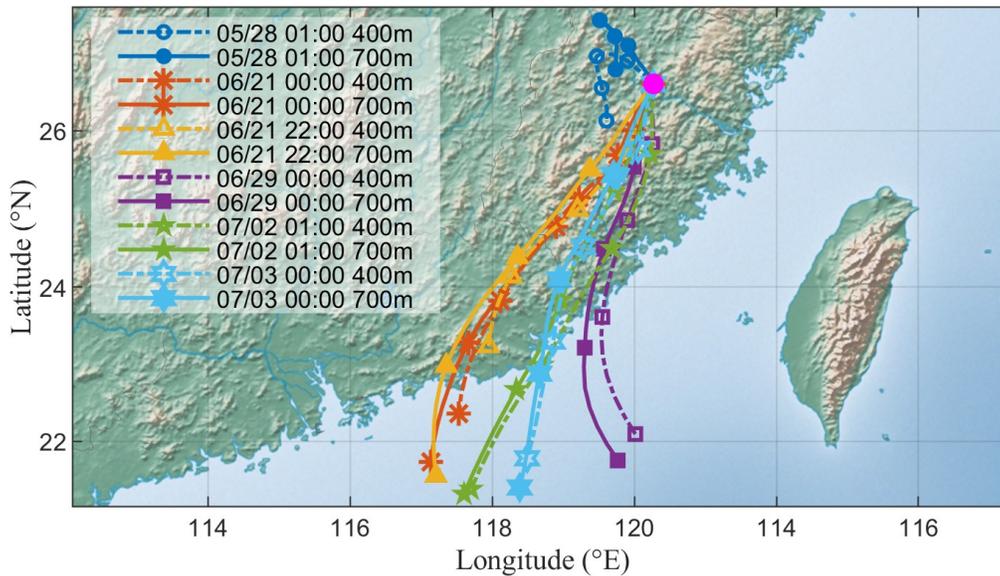


Fig. 6. 24-h backward trajectory using HYSPLIT model and GDAS1 datasets. The trajectory starts at the observation site and is marked as a magenta point. Different line colors refer to different trajectory starting times. The trajectory time during each observation is set to the middle time of each observation. Trajectory which starts at 700 m and 400 m are marked with solid and dashed lines, respectively. Terminal points are marked every 6 hours. The base map is made with Natural Earth, free vector, and raster map data at <http://naturalearthdata.com>.

observations without two-layer structures, the upward net thermal radiation fluxes are the lowest (Fig. 7(a)), and the volumetric water contents in the near-surface soil layer highest (Fig. 7(d)). As Fig. 7(c) illustrate, a large amount of precipitation was generated during the daytime before the observation. The precipitation clouds reduced the solar radiation received by the surface during the day. Combined with high surface water content from the precipitation, this leads to decreased surface thermal radiation at night. In the second observation, no strong precipitation was recorded. However, a flood peak had emerged in the nearby Minjiang River a few days earlier. This flood peak caused high water runoff and soil water content, further inhibiting the surface thermal radiation [62]. Surface thermal radiation is the major driver of the nocturnal stable boundary layer [63]. Reduced thermal radiation hindered the development of the stable boundary layer at night, leading to a decreased boundary layer height. Therefore, the boundary layer during the first two observations was likely too low, placing it in the blind zone of the lidar. This explains the absence of the two-layer structure during those periods.

In summary, the cluster 1 originate from inland aerosols. Considering it origins from Mount Wuyi region and its observation time is between spring and summer, the fluorescence in cluster 1 are likely to be attributed to pollen in Mount Wuyi region. Cluster 2 and 3 have more similar fluorescence spectra comparing with cluster 1, and have same origins from oceanic and coastal areas in the south and southwest of the site. There is a microlayer on the surface of the ocean with higher abundance of microorganisms and other biological materials compared with subsurface waters. Breaking waves, rain and boat traffic can help form droplets from this sublayer. These biomaterials can be further concentrated in these droplets and form bioaerosols [64]. So, cluster 2 are inferred to be the oceanic aerosols, whose fluorescence is produced by bacteria, fungi, and algae live in the ocean surface. Considering the similar fluorescence spectra and vertical distribution difference between cluster 2 and 3. The cluster 3 are inferred to be the mixture of

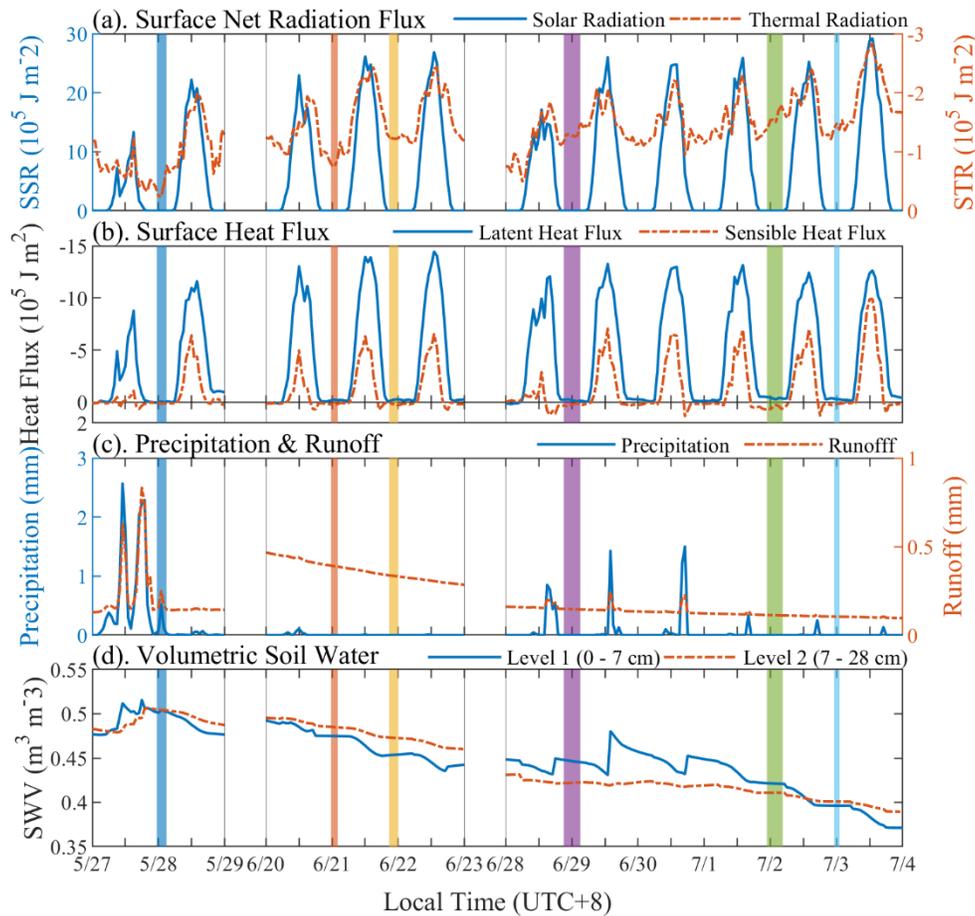


Fig. 7. Meteorological Data from ERA5 hourly data on single levels. (a): hourly accumulated net solar radiation flux received by the ground (blue line) and hourly accumulated net thermal radiation flux emitted by the ground (red line). (b): hourly accumulated surface latent heat flux (blue line) and latent surface sensible heat flux (red line). In (a) and (b) the positive value means the downward propagation direction. (c): hourly accumulated total precipitation (blue line) and total runoff (red line) including water drains away over the surface or under the ground. (d): the volume of water in soil layer 1 (0–7 cm, blue line) and layer 2 (7–28 cm, red line), the surface is at 0 cm. The shading region in the figure indicates each observation period.

transported external oceanic bioaerosols and local emitted bioaerosols below the nocturnal stable boundary layer.

4. Conclusion

In this paper, a spectra-resolved fluorescence lidar system was constructed and deployed at an elevation angle of 30° . This system was used to conduct range-resolved observations of fluorescent aerosols over Nanping City, Fujian Province from April to early July 2024. Firstly, we found that in weak atmospheric fluorescence detection, even if the elastic backscattered signals are suppressed, the leakage of first-order vib-rotational N_2 Raman signals is high enough to

contaminate fluorescence detection, and thus interfere with subsequent spectra analysis. To address this issue, in the last 6 observations, the N₂ Raman signals leakage is eliminated by switching to a 400 nm long-pass filter. After switching, we make a rough comparison in magnitude orders of the fluorescence intensities during each observation based on previous observation results. And results showed that fluorescence signals in each channel are estimated about 3 orders of magnitude lower than that of the N₂ Raman signal. This fluorescence intensity is previously thought to be non-fluorescent. The fluorescence spectra during these observations showed similar shapes with decreasing intensity as wavelength increases.

Despite the weak signal, the maximum detection range of the fluorescent aerosols is 7 km (corresponding to a height of 3.5 km). The strong fluorescence signal is primarily distributed below the height of 1.5–2 km, a typical summertime residual layer height. In the spectral analysis, the conventional spectra integration method was first applied to reduce noise and visualize spectral variations. Using this method, several high-altitude fluorescent aerosol transport events were identified by enhanced fluorescence signals and increased proportions of long-wavelength fluorescence. In addition, the downward movement of high-altitude fluorescent aerosols and the upward diffusion of low-altitude fluorescent aerosols were observed.

To enhance fluorescence spectra analysis, we introduced a clustering method and combined it with backward trajectory analysis and meteorological data. This approach addresses the limitations of the conventional spectral integration method and evaluates the classification potential of LIF lidar. The clustering results for fluorescence signals below 1 km height demonstrate effective classification into distinct 3 clusters. These clusters showed significant correlations with aerosol transport pathways and nocturnal stable boundary layer development. The possible related types of bioaerosols are discussed according to their origins and spatiotemporal distributions. Therefore, these findings highlight the impact of aerosol transport pathways and nocturnal stable boundary layer developments on local fluorescence spectra.

This study advances LIF lidar capabilities in fluorescence detection and classification, particularly under weak fluorescence intensities and similar spectral shapes. By introducing a novel data analysis approach, we demonstrate that even in low fluorescence conditions, atmospheric aerosol fluorescence retains valuable information about aerosol activity. This method enables the extraction of such information, extending LIF lidar applications to ordinary scenarios where high-concentration fluorescent bioaerosols are absent. This extension provides a more comprehensive understanding of atmospheric physical and chemical processes. In the subsequent work, the system will be upgraded to integrate the detection of other atmospheric parameters, such as backscattering coefficient and color ratio. Joint observations with our coherent Doppler wind lidar which is capable of depolarization detection will be conducted in the future [65,66].

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. M. Yao, "Reprint of bioaerosol: A bridge and opportunity for many scientific research fields," *J. Aerosol Sci.* **119**, 91–96 (2018).
2. V. R. Després, J. Alex Huffman, S. M. Burrows, *et al.*, "Primary biological aerosol particles in the atmosphere: A review," *Tellus, Series B: Chemical and Physical Meteorology* **64**(1), 15598 (2012).
3. J. Fröhlich-Nowoisky, C. J. Kampf, B. Weber, *et al.*, "Bioaerosols in the Earth system: Climate, health, and ecosystem interactions," *Atmos. Res.* **182**, 346–376 (2016).
4. A. M. Jones and R. M. Harrison, "The effects of meteorological factors on atmospheric bioaerosol concentrations - A review," *Sci. Total Environ.* **326**(1–3), 151–180 (2004).
5. X. Li, Y. Shi, Q. Liu, *et al.*, "Airborne transmission of a novel Tembusu virus in ducks," *J. Clin. Microbiol.* **53**(8), 2734–2736 (2015).
6. A. I. Donaldson and S. Alexandersen, "Predicting the spread of foot and mouth disease by airborne virus," *Rev. Sci. Tech.* **21**(3), 569–575 (2002).

7. M. J. Fennelly, G. Sewell, M. B. Prentice, *et al.*, “Review: The use of real-time fluorescence instrumentation to monitor ambient primary biological aerosol particles (PBAP),” *Atmosphere* **9**(1), 1 (2017).
8. G. Lieberherr, K. Auderset, B. Calpini, *et al.*, “Assessment of real-time bioaerosol particle counters using reference chamber experiments,” *Atmos. Meas. Tech.* **14**(12), 7693–7706 (2021).
9. P. H. Kaye, W. R. Stanley, E. Hirst, *et al.*, “Single particle multichannel bio-aerosol fluorescence sensor,” *Opt. Express* **13**(10), 3583 (2005).
10. D. Tang, T. Wei, J. Yuan, *et al.*, “Observation of Bioaerosol Transport Using Wideband Integrated Bioaerosol Sensor and Coherent Doppler Lidar,” *Atmos. Meas. Tech.* **15**(9), 2819–2838 (2022).
11. J. R. Lakowicz, *Principles of Fluorescence Spectroscopy* (Springer US, 2006), **287**(2).
12. O. Owoicho, C. O. Olwal, and O. Quaye, “Potential of laser-induced fluorescence-light detection and ranging for future stand-off virus surveillance,” *Microb. Biotechnol.* **14**(1), 126–135 (2021).
13. D. Joshi, D. Kumar, A. K. Maini, *et al.*, “Detection of biological warfare agents using ultra violet-laser induced fluorescence LIDAR,” *Spectrochim. Acta, Part A* **112**, 446–456 (2013).
14. J. Gelbwachs and M. Birnbaum, “Fluorescence of Atmospheric Aerosols and Lidar Implications,” *Appl. Opt.* **12**(10), 2442 (1973).
15. M. Weng, F. Yi, F. Liu, *et al.*, “Single-line-extracted pure rotational Raman lidar to measure atmospheric temperature and aerosol profiles,” *Opt. Express* **26**(21), 27555 (2018).
16. X. Pan, F. Yi, F. Liu, *et al.*, “Diurnal temperature variations in the lower troposphere as measured by an all-day-operational pure rotational Raman lidar,” *Appl. Opt.* **59**(28), 8688 (2020).
17. F. Liu, R. Wang, F. Yi, *et al.*, “Pure rotational Raman lidar for full-day troposphere temperature measurement at Zhongshan Station (69.37°S, 76.37°E), Antarctica,” *Opt. Express* **29**(7), 10059 (2021).
18. F. Liu and F. Yi, “Lidar-measured atmospheric N₂ vibrational-rotational Raman spectra and consequent temperature retrieval,” *Opt. Express* **22**(23), 27833 (2014).
19. L. Fellner, M. Kraus, F. Gebert, *et al.*, “Multispectral LIF-Based Standoff Detection System for the Classification of CBE Hazards by Spectral and Temporal Features,” *Sensors* **20**(9), 2524 (2020).
20. J. Wojtanowski, M. Zygmunt, M. Muzal, *et al.*, “Performance verification of a LIF-LIDAR technique for stand-off detection and classification of biological agents,” *Optics and Laser Technology* **67**, 25–32 (2015).
21. Ø. Farsund, G. Rustad, and G. Skogan, “Standoff detection of biological agents using laser induced fluorescence—a comparison of 294 nm and 355 nm excitation wavelengths,” *Biomed. Opt. Express* **3**(11), 2964 (2012).
22. S. Chen, Y. Chen, Y. Zhang, *et al.*, “Dual-channel mobile fluorescence lidar system for detection of tryptophan,” *Appl. Opt.* **59**(3), 607 (2020).
23. S. Buteau and J.-R. Simard, *JBSDS Increment II Tech Demo III SINBAHD Performances* (2010), (October).
24. J. R. Simard, G. Roy, P. Mathieu, *et al.*, “Standoff sensing of bioaerosols using intensified range-gated spectral analysis of laser-induced fluorescence,” *IEEE Trans. Geosci. Remote Sensing* **42**(4), 865–874 (2004).
25. S. Buteau, J.-R. Simard, and S. Rowsell, “Bioaerosols standoff detection simultaneously refereed with particle concentration (ppl) and viability units (ACPLA),” in *Optically Based Biological and Chemical Detection for Defence V*, J. C. Carrano and C. J. Collins, eds. (2009), 7484, p. 748408.
26. S. Buteau, J. Ho, P. Lahaie, *et al.*, *Laser Based Stand-Off Detection of Biological Agents (Détection à Distance Des Agents Biologiques à l’aide Du Laser)* (2010), **AC/323(SET(SET-098/RTG-55))**.
27. I. Veselovskii, Q. Hu, P. Goloub, *et al.*, “Mie-Raman-fluorescence lidar observations of aerosols during pollen season in the north of France,” *Atmos. Meas. Tech.* **14**(7), 4773–4786 (2021).
28. N. Sugimoto, Z. Huang, T. Nishizawa, *et al.*, “Fluorescence from atmospheric aerosols observed with a multi-channel lidar spectrometer,” *Opt. Express* **20**(19), 20800 (2012).
29. Y. Saito, K. Ichihara, K. Morishita, *et al.*, “Remote Detection of the Fluorescence Spectrum of Natural Pollens Floating in the Atmosphere Using a Laser-Induced-Fluorescence Spectrum (LIFS) Lidar,” *Remote Sens.* **10**(10), 1533 (2018).
30. S. C. Richardson, M. Mytilinaios, R. Foskinis, *et al.*, “Bioaerosol detection over Athens, Greece using the laser induced fluorescence technique,” *Sci. Total Environ.* **696**, 133906 (2019).
31. I. Veselovskii, Q. Hu, P. Goloub, *et al.*, “Combined use of Mie-Raman and fluorescence lidar observations for improving aerosol characterization: Feasibility experiment,” *Atmos. Meas. Tech.* **13**(12), 6691–6701 (2020).
32. Y. Wang, Z. Huang, T. Zhou, *et al.*, “Identification of fluorescent aerosol observed by a spectroscopic lidar over northwest China,” *Opt. Express* **31**(13), 22157 (2023).
33. “Natural Resources,” <https://www.np.gov.cn/cms/html/npszf/2024-04-30/1745898444.html>.
34. “Release Platform of Environmental Air Quality Information in Fujian Province,” <https://fjaqi.fjemc.org.cn/fb/>.
35. J. Reichardt, O. Behrendt, and F. Lauermann, “Spectrometric fluorescence and Raman lidar: absolute calibration of aerosol fluorescence spectra and fluorescence correction of humidity measurements,” *Atmos. Meas. Tech.* **16**(1), 1–13 (2023).
36. G.-J. Roelofs and V. Kamphuis, “Cloud processing, cloud evaporation and Angström exponent,” *Atmos. Chem. Phys.* **9**(1), 71–80 (2009).
37. X. Shang, H. Xia, X. Dou, *et al.*, “Adaptive inversion algorithm for 1 . 5 μ m visibility lidar incorporating in situ angstrom wavelength exponent,” *Opt. Commun.* **418**, 129–134 (2018).
38. K. Yumimoto, T. Y. Tanaka, N. Oshima, *et al.*, “JRAero: the japanese reanalysis for aerosol v1.0,” *Geosci. Model Dev.* **10**(9), 3225–3253 (2017).

39. M. Elshora, H. Xia, L. Su, *et al.*, "Observation and study of consecutive dust storms in the taklimakan desert from March 16 to 27, 2022, using reanalysis models and lidar," *Atmos. Res.* **305**, 107431 (2024).
40. T. Wei, M. Wang, P. Jiang, *et al.*, "Retrieving aerosol backscatter coefficient using coherent doppler wind lidar," *Opt. Express* **33**(4), 6832 (2025).
41. I. Veselovskii, N. Kasianik, M. Korenskii, *et al.*, "Multiwavelength fluorescence lidar observations of smoke plumes," *Atmos. Meas. Tech.* **16**(8), 2055–2065 (2023).
42. U. Wandinger, "Raman Lidar," in *Lidar* (Springer-Verlag, 2006), pp. 241–271.
43. W. F. Meggers and C. C. Kiess, "False Spectra from Diffraction Gratings Part I Secondary Spectra," *J. Opt. Soc. Am.* **6**(5), 417 (1922).
44. C. Runge, "False Spectra from Diffraction Gratings Part II Theory of Lyman Ghosts," *J. Opt. Soc. Am.* **6**(5), 429 (1922).
45. Y. Zhang, Z. Sun, S. Chen, *et al.*, "Classification and source analysis of low-altitude aerosols in Beijing using fluorescence-Mie polarization lidar," *Opt. Commun.* **479**, 126417 (2021).
46. F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: an overview, II," *WIREs Data Min. & Knowl.* **7**(6), 1219 (2017).
47. F. Murtagh and P. Contreras, "Algorithms for hierarchical clustering: An overview," *WIREs Data Min. & Knowl.* **2**(1), 86–97 (2012).
48. A. Lukasová, "Hierarchical agglomerative clustering procedure," *Pattern Recognition* **11**(5-6), 365–381 (1979).
49. D. L. Davies and D. W. Bouldin, "A Cluster Separation Measure," *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-1**(2), 224–227 (1979).
50. T. Calinski and J. Harabasz, "A dendrite method for cluster analysis," *Comm. in Stats. - Theory & Methods* **3**(1), 1–27 (1974).
51. M. Verleysen and M. Verleysen, "Principal Component Analysis (PCA) Principal Component Analysis (PCA) Principal Component Analysis (PCA)," *Statistics* **9**, 1–8 (2001).
52. A. F. Stein, R. R. Draxler, G. D. Rolph, *et al.*, "NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System," *Bull. Am. Meteorol. Soc.* **96**(12), 2059–2077 (2015).
53. NOAA Air Resources Laboratory (ARL), "Global data assimilation system (GDAS1) archive information," <http://ready.arl.noaa.gov/gdas1.php>.
54. W. Wu, H. Chen, J. Guo, *et al.*, "Regionalization of the Boundary-Layer Height and its Dominant Influence Factors in Summer over China," *Chinese Journal of Atmospheric Sciences* **48**(3), 1201–1216 (2024).
55. J. Guo, Y. Miao, Y. Zhang, *et al.*, "The climatology of planetary boundary layer height in China derived from radiosonde and reanalysis data," *Atmos. Chem. Phys.* **16**(20), 13309–13319 (2016).
56. M. Collaud Coen, C. Praz, A. Haeefe, *et al.*, "Determination and climatology of the planetary boundary layer height above the Swiss plateau by in situ and remote sensing measurements as well as by the COSMO-2 model," *Atmos. Chem. Phys.* **14**(23), 13205–13221 (2014).
57. Y. Jiang, J. Xin, D. Zhao, *et al.*, "Analysis of differences between thermodynamic and material boundary layer structure: Comparison of detection by ceilometer and microwave radiometer," *Atmos. Res.* **248**, 105179 (2021).
58. Z. Zhu, H. Li, X. Zhou, *et al.*, "A Cluster Analysis Approach for Nocturnal Atmospheric Boundary Layer Height Estimation from Multi-Wavelength Lidar," *Atmosphere* **14**(5), 847 (2023).
59. Y. Ren, S. Chen, A. Yu, *et al.*, "Reliability Analysis of Wind Field and Boundary Layer Height," *Journal of Arid Meteorology* **39**(3), 514–523 (2021).
60. J.-N. Hersbach, H. Bell, B. Berrisford, *et al.*, "ERA5 hourly data on single levels from 1940 to present,"
61. J. Kondo and S. Ishida, "Sensible Heat Flux from the Earth's Surface under Natural Convective Conditions," *J. Atmos. Sci.* **54**(4), 498–509 (1997).
62. E. E. Small and S. A. Kurc, "Tight coupling between soil moisture and the surface radiation budget in semiarid environments: Implications for land-atmosphere interactions," *Water Resources Research* **39**(10), 1–14 (2003).
63. J. C. Wyngaard, "The stable boundary layer," in *Turbulence in the Atmosphere* (Cambridge University Press, 2010), pp. 267–294.
64. J. Löndahl, *Bioaerosol Detection Technologies* (Springer New York, 2014), **53**.
65. C. Wang, H. Xia, M. Shangguan, *et al.*, "1.5 μm polarization coherent lidar incorporating time-division multiplexing," *Opt. Express* **25**(17), 20663 (2017).
66. J. Qiu, H. Xia, M. Shangguan, *et al.*, "Micro-pulse polarization lidar at 15 μm using a single superconducting nanowire single-photon detector," *Opt. Lett.* **42**(21), 4454 (2017).