Properties of Coherent Structures over Paris: A Study Based on an Automated Classification Method for Doppler Lidar Observations

IOANNIS CHELIOTIS, SELSA DIEUDONNÉ, Hervé DELBARRE, ANTON SOKOLOV, EGOR DMITRIEV, PATRICK AUGUSTIN, MARC FOURMENTIN, FRANÇOIS RAVETTA, AND JACQUES PELON

Abstract: The studies related to the coherent structures in the atmosphere, using Doppler wind lidar observations, so far have relied on the manual detection and classification of the structures in the lidar images, making this process time-consuming. We developed an automated classification that is based on texture analysis parameters and the quadratic discriminant analysis algorithm for the detection of medium-to-large fluctuations and coherent structures recorded by single Doppler wind lidar quasi-horizontal scans. The algorithm classified a training dataset of 150 cases into four types of patterns, namely, streaks (narrow stripes), rolls (wide stripes), thermals (enclosed areas), and “others” (impossible to classify), with 91% accuracy. Subsequently, we applied the trained algorithm to a dataset of 4577 lidar scans recorded in Paris, atop a 75-m tower for a 2-month period (September–October 2014). The current study assesses the quality of the classification by examining the physical properties of the classified cases. The results show a realistic classification of the data: with rolls and thermals cases mostly classified concurrently with a well-developed atmospheric boundary layer and the streaks cases associated with nocturnal low-level jets events. Furthermore, rolls and streaks cases were mostly observed under moderate or high wind conditions. The detailed analysis of a 4-day period reveals the transition between the types. The analysis of the space spectra in the direction transverse to the mean wind, during these four days, revealed streak spacing of 200–400 m and roll sizes, as observed in the lower level of the mixed layer, of approximately 1 km.

Significance statement: We studied the physical properties of medium-to-large fluctuations and coherent structures in the atmosphere classified by a newly developed algorithm for single Doppler lidar horizontal scans. This study emphasizes the ability to analyze a large dataset by using such an automated classification method. The results show rolls and thermals occurrences during a well-developed atmospheric boundary layer and streaks formations associated with nocturnal low-level jets events. Rolls and streaks were observed under moderate/high wind conditions with horizontal sizes, in the transverse to the mean wind direction, equivalent to approximately 1 km and 300 m, respectively. The results of this study can be used in combination with atmospheric numerical model simulations to examine the impact of the structures on pollutant concentrations.

Keywords: Boundary layer; Wind; Lidars/lidar observations; Remote sensing; Classification; Pattern detection; Automated systems; Machine learning; Pattern recognition

1. Introduction

Coherent structures in the atmospheric boundary layer (ABL) have been the subject for a significant number of studies throughout the years. These structures are characterized by the organization of eddies in a quasi-periodic pattern in the two horizontal dimensions within a turbulent atmosphere (Tur and Levich 1992). The form of the coherent structures is maintained for a sufficient time period to apply temporal statistics calculation (Hussain 1983). Their role in the atmospheric processes is substantial as they are key factors in the transport of heat and moisture (e.g., Brient et al. 2019), the dispersion of pollutants (e.g., Han et al. 2019), and the dust transport (e.g., Zhang et al. 2018) throughout the ABL. For this purpose, various methods have been developed in order to observe and better understand these structures, each of them having advantages and disadvantages. More specifically, some studies have utilized aircraft measurements, which can cover large areas but are limited in short time periods and are not accessible to many researchers (e.g., Brümmer et al. 1985; Martin and Bakan 1991). The meteorological towers can take observations over large time periods; however, they are installed at specific locations with limited spatial coverage (e.g., Barthlott et al. 2007; Smedman 1991). The weather numerical simulations can overcome limitations in spatial and temporal coverages while still using observation as initial conditions (e.g., Bartello et al. 1994; Jiang et al. 2017). There are various studies that combine more than one methods in order to combine the advantages of each method, which requires expertise and...
accesses to all the different methods and can be limited to few researchers (e.g., Bernard-Trottolo et al. 2004; Drobinski et al. 2004; Lohou et al. 1998). On the other hand, remote sensing methods, such as radar (e.g., Eymard and Weill 1988; Rabin et al. 1982) and lidar (e.g., Lin et al. 2008; Lothon et al. 2006) can be installed in several areas onshore or offshore, they have the advantage to measure horizontally or vertically while operating for long time periods (Kim et al. 2019). Furthermore, the remote sensing methods can provide the visualization of coherent structures (Drobinski et al. 2004). Despite this asset of lidars, the studies so far have been limited to few cases covering only short periods of time (e.g., Träumner et al. 2015). This is mainly because the analysis of the lidar data relies on the subjective observation of the researchers, which can be a remarkably time-consuming process for a large dataset and furthermore it is substantially difficult to ensure consistency throughout the classification.

We have developed an automated method with the aim to achieve consistent medium-to-large fluctuations and coherent structures (mlf-cs) classifications for single Doppler wind lidar observations (Cheliotis et al. 2020). We focus on the streaks and rolls as presented by Young et al. (2002). The streaks occur in the surface layer with the possibility to extend in the mixed layer. The main characteristics of the streaks is their formation, which resembles stripes of stronger horizontal winds alternate with weaker ones in the direction aligned with the mean wind (Khanha and Brasseur 1998). The average space between the stronger horizontal winds is equivalent to hundreds of meters, and their life-span is no more than several tens of minutes (Drobinski and Foster 2003). One of the main driving factors for the formation of the streaks is the high wind shear near the surface. The buoyancy contributes also to the streaks formation with the physics varying with regard to the atmospheric conditions (Moeng and Sullivan 1994). They usually form under neutral or near-neutral stratification, but they may form under stable or unstable conditions as well (Khanha and Brasseur 1998). The streaks formation can be visible during low clouds or fog events (e.g., MacDonald et al. 2020; Nakanishi 2000).

On the other hand, rolls occur in the mixed layer, extending from the surface to the capping inversion and they are also aligned with the mean wind (LeMone 1973). The main factors for the formation and maintenance of rolls are the low-level wind shear, the spatial variations of surface-layer heat flux and the homogeneity of the surface (Weckwerth and Parsons 2006). The formation of the rolls resemble also stripes patterns, but these are wider when compared with the streaks’ structures. It is possible to identify the rolls in the atmosphere during convective conditions. The rotation of the rolls in the vertical plane generates ascending and descending motions that lead to the formation of cloud rows and clear sky areas. This phenomenon is known as “cloud streets” (Lohou et al. 1998). However, the convective conditions are not a requirement for the formation of rolls, as they can form during low relative humidity conditions, in a phenomenon known as “dry streets” (Kuettner 1971). The horizontal sizes of the rolls vary from a few kilometers to a few tens of kilometers, and their life-span ranges from hours to days (Drobinski and Foster 2003). For their width, Young et al. (2002) separate the rolls in two categories: narrow and wide. The ascending air masses inside the narrow rolls are one thermal wide (Weckwerth et al. 1999), whereas in the wide rolls each ascending area contains multiple thermals (Brümmer 1999).

The remote sensing methods can provide the mlf-cs fields, by applying the velocity azimuth display (VAD) method in the radial wind observations (Browning and Wexler 1968; Drobinski et al. 2004; Lhermitte 1962). In a previous article (Cheliotis et al. 2020), we presented our automated mlf-cs classification method based on statistical texture feature extraction (Haralick parameters: Haralick et al. 1973; Chaki and Dey 2020) and the quadratic discriminant analysis (QDA) (Hastie et al. 2009). We assessed its performance on a training dataset of 150 samples by means of tenfold cross validation. The training dataset consisted of four mlf-cs types: streaks (narrow elongated patterns associated with high wind shear near the surface), rolls (wide elongated patterns associated with cloud streets), thermals (large enclosed patterns associated with fair weather conditions), and “others” (any other pattern that does not belong in one of the other types). Our algorithm revealed five texture analysis parameters that minimized the classification error at approximately 9% for the training dataset. In the current study, these five texture analysis parameters were utilized for the classification, into one of the four mlf-cs types, of the 4577 lidar scans recorded in the urban area of Paris, France, for the period 4 September–6 November 2014. The aim of the study is to assess the quality of the classification in terms of physicsmeteorology for a large dataset. In this regard, we examined the physical parameters of the classified structures for the whole ensemble. These parameters are the time of the day, the ABL height and the horizontal wind speed during the occurrence of the structures. Furthermore, we investigated a 4-day period during which streaks, rolls, and thermals cases were all detected so as to retrieve information about the transition between the types of structures, their life cycle, and the sizes of the structures. Up to this point, large datasets of mlf-cs classifications have been analyzed only visually (e.g., Yagi et al. 2015). This is the first time an automated classification has been used for a large dataset.

The main body of the paper is composed of three sections. In section 2, we present the method of our study, which includes the experimental setup, the mlf-cs fields of our classification, the methods to estimate the ABL height and the horizontal size of the structures as well as the relevant weather conditions in the area during the study period. Section 3 contains the results, separated into two subsections, one for the classification of the whole ensemble and the other focusing on a peculiar 4-day period. Section 4 summarizes all of the main points of the study.

2. Method

2a. Experimental setup

The current study relies on radial wind observations derived from a 2-month campaign (4 September–6 November 2014) performed in Paris in the framework of the project Vegetation et Ilot de Chaleur Urbain (VEGILOT, i.e., vegetation and urban heat island) (Klein et al. 2019). The aim of the VEGILOT campaign was to study the atmospheric chemistry
and transport processes inside the urban boundary layer in Paris. For the radial wind measurements, the Leosphere WLS100 scanning Doppler lidar (https://www.leosphere.com) was settled atop a 75-m building in the Jussieu Campus, located in the center of Paris (48.82°N, 2.34°E), as showcased in Fig. 1. The Cloud and Aerosol Microlidar (CAML), manufactured by the Cimel company (https://www.cimel.fr), was installed on the roof of an adjacent building (20-m height) to provide aerosol measurements. The maximum range for the Doppler lidar was 5 km, but the effective range did not exceed 4 km for the vast majority of the data as backscatter signals with carrier-to-noise ratio (CNR) lower than ~27 dB was filtered out, along with outliers in wind speed values. The red circle in Fig. 1 represents the maximum horizontal scanning range for the Doppler lidar. It is important to note that despite Paris being a metropolitan city, there is a regulation with respect to the building heights in the center, and therefore these do not exceed 50 m, with few exceptions including the building in the Jussieu Campus (Saint-Pierre et al. 2010). Specifically, for the scanning area, the ground altitude ranges between 30 and 60 m with some exceptions near the boundaries where there are some hills. Hence the scanned urban area is mostly homogeneous. The significant lidar properties for the Doppler and the aerosol lidar during the VEGILOT campaign are displayed in Table 1.

The way the Doppler operates has been described extensively by (Cariou et al. 2007). It took measurements continuously during the 2-month campaign, following a sequence of scanning methods that included quasi-horizontal scans [plan position indicator (PPI)], vertical beams [line of sight LOS]), and vertical profiling using the Doppler beam swinging (DBS). The full measurement cycle was repeated every 18 min. The LOS beams consisted of 10 vertical observations during 30 s, whose average was used for the estimation of the nocturnal ABL height (section 2c). A direct average over 5 min would have been preferable, in statistical terms; however, the scanning sequence was not optimized for turbulence observation, and the lidar raw files did not keep track of the single-shot profiles. For the PPI method, quasi-horizontal scans at a 1° elevation angle were carried out for azimuth angles ranging from 0° to 360° with a 2° resolution. This 1° elevation created a height difference between the central point and the boundaries of the scans equivalent to 87 m. The duration of a PPI scan was 3 min, which is adequate for the observation of the mlf-cs with an equivalent or longer life-span. The DBS method was used to identify the meteorological phenomena relevant to interpret the classification result, such as low-level jets. The DBS method was a combination of LOS beams in the north, east, south, and west directions with a 75° elevation and a total duration of 15 s. This method was applied two times consecutively in each cycle. For the aerosol lidar, its operation has been demonstrated by Pelon et al. (2008). For the current study, the aerosol lidar observations were used exclusively for the daytime estimation of the ABL height during the 2-month campaign (section 2c). The aerosol lidar was emitting beams vertically and data were averaged on 10-min periods. Table 2 summarizes the different scanning methods selected during the VEGILOT campaign. The measurement cycle included other scans that are not relevant for the current study, and thus they are not presented in Table 2.
b. The mlf-cs fields

As we stated in section 1, we have developed an automated method to detect and classify mlf-cs near the surface, based on single horizontal Doppler lidar scans by applying the VAD method. The principle of this method is based on the sinusoidal shape of the radial wind speed observations as a function of azimuth for a ring (observations at same distance from the lidar) when the wind field is homogeneous over the scanning area. In that case it is possible to detect the mean wind speed and the wind direction at the maximum/minimum of the cosine fitting curve. By subtracting the mean wind speed from the radial wind speed observations, the mlf-cs field is retrieved. A detailed description of this process is presented in Cheliotis et al. (2020). A characteristic example for each of the identified mlf-cs fields that we aimed to classify is displayed in Fig. 2.

Each field showcases some distinctive characteristics. In Fig. 2a, representing the streaks case, it is possible to see some narrow stripes aligned with the horizontal wind direction. Similarly, in Fig. 2b, there are some wide stripes aligned with the horizontal wind direction, representing the rolls case. Figure 2c showcases a more chaotic pattern where it was still possible to identify some organized enclosed areas of positive and negative values corresponding to the thermals case. Finally, Fig. 2d portrays a type that it would not be possible to classify into one of the aforementioned types and as a result it was identified as a separate type named “others.”

TABLE 1. Properties of the Doppler and the aerosol lidar utilized during VEGILOT.

<table>
<thead>
<tr>
<th>Lidar</th>
<th>Doppler lidar (Leosphere WLS100)</th>
<th>Aerosol lidar (CAML)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alt of lidar</td>
<td>75 m AGL</td>
<td>20 m AGL</td>
</tr>
<tr>
<td>Min range</td>
<td>100 m</td>
<td>200 m</td>
</tr>
<tr>
<td>Max range</td>
<td>5 km</td>
<td>20 km</td>
</tr>
<tr>
<td>Spatial resolution</td>
<td>50 m</td>
<td>15 m</td>
</tr>
<tr>
<td>Radial wind speed range</td>
<td>From −30 to 30 m s⁻¹</td>
<td>—</td>
</tr>
<tr>
<td>Laser wavelength</td>
<td>1.543 μm</td>
<td>532 nm</td>
</tr>
<tr>
<td>Radial wind accuracy</td>
<td>±0.1 m s⁻¹</td>
<td>—</td>
</tr>
<tr>
<td>Accumulation time</td>
<td>1 s per beam</td>
<td>10 min</td>
</tr>
</tbody>
</table>

c. Estimation of the ABL height

The Doppler lidar vertical LOS beams and the aerosol lidar observations were used for the estimation of the ABL height during the night and during the day, respectively. Tucker et al. (2009) showcased a method for the ABL height estimation by applying a threshold to the vertical wind variance near the surface equal to 0.3 or 0.4 m² s⁻², depending on the conditions. The ABL height is equivalent to the height where the vertical wind variance falls below that threshold. Several studies applying the same method have been performed since then, with different threshold values depending on the conditions.

TABLE 2. Scanning methods and their characteristics during VEGILOT.

<table>
<thead>
<tr>
<th>Lidar</th>
<th>Scanning area</th>
<th>Purpose</th>
<th>Elev and azimuth angles</th>
<th>Scan duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doppler-LOS</td>
<td>Vertical profile</td>
<td>Estimation of the ABL height (night)</td>
<td>Elevation 90°; azimuth 0°</td>
<td>10 × 30 s</td>
</tr>
<tr>
<td>Doppler-PPI</td>
<td>Almost horizontal scans near surface</td>
<td>Identification of structures</td>
<td>Elevation 1°; azimuth 0°–360° with 2° resolution</td>
<td>3 min</td>
</tr>
<tr>
<td>Doppler-DBS</td>
<td>Combination of LOS</td>
<td>Identification of low-level jet cases</td>
<td>Elevation 75°; azimuth 0°, 90°, 180°, and 270°</td>
<td>2 × 15 s</td>
</tr>
<tr>
<td>Aerosol</td>
<td>Vertical profile</td>
<td>Estimation of the ABL height (day)</td>
<td>Elevation 90°; azimuth 0°</td>
<td>10-min avg</td>
</tr>
</tbody>
</table>
on the conditions, including studies over metropolitan cities like Beijing, China (Huang et al. 2017), or London, England (Barlow et al. 2011). In Paris, we used the vertical wind speed dispersion provided by the wind lidar, a quantity based on the Doppler spectral width that is related to the wind standard deviation, but should be calibrated using reference measurement, for instance, from a sonic anemometer. However, no such observation was available in the vicinity, so that the value of the threshold we set ($1.1 \text{ m s}^{-1}$) is not comparable with other studies that used Tucker’s method. This value was selected because it provided the most realistic values for the ABL height. Lower threshold values would have resulted in nocturnal ABL heights above 500 m.

Moreover, this technique is limited to the cases when the wind profile extends above the ABL, hence it is applicable during the night and early morning, when the aerosols’ concentrations in the residual layer provide a sufficient CNR for wind retrieval. The building height combined with the lidar minimum range allowed to retrieve the ABL height only from 200 m, which might seem rather high for a stable ABL. However, previous observations from the same site (Dieudonné 2012), using 20 months of aerosol lidar measurements with a low-altitude channel, showed that the minimum height of the nocturnal ABL over Paris is around 200 m in all seasons so that the limitation of the method used in the present study should only lead to a limited overestimation of the nocturnal ABL height.

During the day, the aerosols’ concentrations are low in the free troposphere, thus the CNR is significantly lower than the $-27$ dB threshold. Therefore, we estimated the ABL height during the day using the maximum of the variance of the range corrected signal of the aerosol lidar, a principle described by Flamant et al. (1997) and Menut et al. (1999). This method takes advantage of the abundance of aerosol in the boundary layer in contrast to the free troposphere (Stull 1988). For the current study, the wind speed dispersion threshold method was used for the time periods between 0000 and 0800 UTC and between 1800 and 2359 UTC, while from 0800 to 1800 UTC we considered the ABL height to be the level where the minimum derivative of the standard deviation of the range corrected signal occurred because it gave better results than the wind dispersion threshold. Figure 3 illustrates the

---

**FIG. 2.** The mlf-cs types of our classification along with the horizontal wind direction (black arrow). (a) Streaks observed at 2122 UTC 8 Sep. (b) Rolls observed at 1324 UTC 9 Oct. (c) Thermals observed at 1313 UTC 4 Sep. (d) Others observed at 0308 UTC 10 Oct.
estimated ABL height along with the vertical wind speed LOS observations on 9 September 2014. During the day, the ABL height ranged between 1 and 1.2 km, and during the night, it ranged between 250 and 300 m. The ABL height values agree with previous observations on the site (Pal et al. 2012; Dieudonné 2012). Nevertheless, we acknowledge that given the height of the building and the minimum scanning range of the lidar, this method cannot be applied for a shallow ABL below 200 m.

It is important to remark that low clouds, such as stratus and stratocumulus, could inhibit the lidar signal. In such cases, the abovementioned method for the estimation of the ABL height was ineffective. Hence the estimation of the ABL height was performed solely for low clouds free cases. It was possible to estimate the ABL height for the entire day (during 24 h or so) for 18 of the 64 days of the study (noted in Fig. 5b). The nocturnal stable boundary layer (SBL) height values vary between 200 and 350 m during this period. The values of the ABL height during daytime vary between 1.2 and 1.8 km.

d. Estimation of the horizontal sizes for the coherent structures

In Figs. 2a and 2b, characteristic cases of coherent structures are displayed. The periodicity of the alternating positive and negative stripes is evident. It is possible to estimate the sizes of these periodicities by the means of the spectral analysis. The spectrum is computed by applying the fast Fourier transform analysis for physical space data (Stull 1988). Flamant et al. (1997) indicated the existence of turbulence in the spectrum peaks of the wind speed. Several studies were carried out since then, which take advantage of the spectrum peaks in order to estimate the horizontal sizes of coherent structures (e.g., Calaf et al. 2013; Drobinski et al. 2004, 1998). The spectrum was calculated for the radial wind speed measured along the direction transverse to the mean wind, after adjoining, the opposite pair of beams closest to this direction. To reduce the noise, the spectrum was computed separately for several pairs of beams, then the spectra were averaged. We observed that averaging five spectra reduced the noise without blurring the peaks appearing on a single spectrum. On the other hand, an average on only three spectra did not reduce the noise enough, while using seven spectra smoothened the peaks significantly. As five pairs of adjacent beams span over an 8°-wide cone, the point spacing in the transverse direction is not exactly the same for the side beams, which would shift slightly the peak position and intensity in the spectra. To avoid this, the radial wind speed values were interpolated to a regular 50 m grid in the transverse direction before computing the spectra.

Figures 4a and 4c showcase the radial wind speed observations from the PPI scans at the same time as Figs. 2a and 2b. The corresponding averaged spectrums (Figs. 4b,d) in the transverse direction, depicted by the black lines, are displayed alongside the respective PPI scans. Furthermore, dashed curves highlighted an adjacent pair of a stronger (positive mlf-cs) and a weaker (negative mlf-cs) radial wind speed stripes, which is the examined periodicity. The effective scanning range for these cases is between 3 and 4 km, thus we can only observe periodicities within this range. The signal of the spectrum below 300 m it is very noisy and therefore excluded from the study. The peaks representing the structures are also highlighted.

For the streaks case (Figs. 4a,b), a prominent peak of periodicity is detected slightly above 400 m. From Fig. 4a, it is visible that the dashed curve, equivalent to a stronger and a weaker radial wind speed stripes, contains an area of an approximate size of 400 m. This value is comparable to the approximate 300-m horizontal spacing of the streaks (distance between the streaks, weaker radial wind area in Figs. 4a,b) observed in other Doppler lidar studies (e.g., Drobinski et al. 2004; Weckwerth et al. 1997) as well as large eddy simulation studies (e.g., Deardorff 1972; Moeng and Sullivan 1994). For the rolls case (Fig. 4), the most prominent peak is detected at approximately 1.3 km. This size is comparable to the highlighted area in Fig. 4b by the dashed curve. So, if we assume that there is half-space between the rolls as well, then the size of the individual rolls should be smaller than 1 km. According to previous studies, the width of the rolls may vary (Young et al. 2002). The horizontal size of narrow mixed-layer rolls could be equivalent to one thermal wide of hundreds of meters (Weckwerth et al. 1999). Although this is a comparable size to the abovementioned case, it is complicated to compare our results with studies focusing on mixed-layer rolls since the observations of the current study were retrieved near the surface. It is also important to mention the occurrence of secondary peaks in the spectrums. Therefore, it is of the utmost importance to handpick the peaks of the spectrums by detecting equivalent size patterns in the mlf-cs fields.

e. Weather conditions during the period under study

For a better understanding of the classification results, it was fundamental to examine some relevant meteorological parameters during the period under study (4 September 2014–6 November 2014). The nearest meteorological station with available results for that period was the Météo France station.
of Paris-Montsouris, located in a park, at the southern edge of inner Paris, 2.9 km south-southwest from the Jussieu site. Thermal patterns are expected to occur during fair cumuli weather; hence the sunshine duration, the atmospheric pressure, and the light wind conditions are appropriate parameters to indicate their development. Furthermore, the detection of precipitation events reveals the cases when the estimation of the ABL height via the lidar observations were not applicable. Figure 5 illustrates the minimum and maximum temperatures, the daily values of sunshine in hours, the daily accumulated precipitation, and the daily average atmospheric pressure for the period under study. The days when the estimation of the ABL height was feasible are pointed out with the magenta bars in Fig. 5b. It is apparent from Figs. 5b and 5d that during the period under study, there were many hours when fair cumuli weather conditions prevailed. So, the thermals type was expected to occur frequently during our observations. It is also evident from Fig. 5c that there was precipitation recorded for 30 of the 64 days of VEGILOT. The ABL height was estimated mostly during days with an abundance of sunshine hours, except 5 September and 23 October where the total sunshine hours did not exceed two hours. During these days the aerosol lidar signal was not interrupted by clouds and therefore it was still possible to estimate the ABL height.

In section 1, we mentioned the weather phenomena related to coherent structures. More particularly, the high wind shear near the surface is a driving factor for the development of streaks, and the cloud streets are an evidence for the formation of rolls. With regard to the wind shear, the horizontal wind profile was built from the DBS soundings. According to Stull (1988) a nocturnal low-level jet (nllj) is characterized by a low-altitude maximum of the horizontal wind speed (jet core), with an increment of at least 2 m s\(^{-1}\) relative to the wind minimum located above it. Figure 6 demonstrates a DBS profile of the horizontal wind speed with a jet core at 300 m altitude and for which it is evident that the difference in wind speed is more than 2 m s\(^{-1}\). We identified nllj events for 20 of the 63 nights based on this description. As a result, we expected to observe several cases of streaks during the nights. The days when the nllj events were observed are displayed in Fig. 7a. MODIS satellite images were used for the identification of cloud streets. Cloud-street formations were observed for 7 of the 64 days of VEGILOT. This is a confirmation for the formation of rolls during these days. However, it is important to emphasize that the analyzed observations were...
recorded near the surface, thus we observed the lower part of the structures during rolls formation. Figure 7b displays the days when cloud streets were observed.

3. Results

a. Classification of the 2-month dataset

In section 2b, we briefly summarized the main points of our classification and indicated the classification results for the training dataset made by the QDA algorithm. The classification error was minimized at approximately 9% for five texture analysis parameters (Cheliotis et al. 2020). We thereupon utilized the QDA algorithm using these five texture analysis parameters in order to classify the whole ensemble, consisted of 4577 PPI scans. Table 3 displays the outcome of the classification. More specifically, the total number of occurrences for each type of structures and the corresponding percentage over the whole ensemble are presented. As expected, the others category was dominant with almost one-half of the ensemble (46%) classified as such, the second most frequent type was the streaks, with 25%, the thermals type accounted for 20% of the total cases and finally 9% of the cases were classified as rolls.

Apart from the total number of occurrences, Table 3 displays the number and fraction of occurrences for the following time periods 0000–0800 UTC (late night/early morning), 0800–1800 UTC (daytime), and 1800–2359 UTC (evening/early night) in order to examine whether the results are realistic. It is evident that during the periods 0000–0800 UTC and 1800–2359 UTC, the others class is dominant, representing 67.4% and 55.6% of the scans, respectively. As we stated in section 1, the estimation of the mlf-cs fields via the VAD method requires a homogeneous mean wind field resulting in a symmetrical radial wind field. A nonsymmetric radial wind field will lead to a bad case of a mlf-cs field that is included in the others type. The most common cause for a nonsymmetric radial wind field is the varying wind direction during calm or low wind conditions (Wilson et al. 1976; Sharan and Yadav 1998). The streaks are the second most frequent type for these time periods with 24.4% and 32.7%, respectively. The most striking result for these time periods is the scarcity of the rolls cases. This is a remarkable result as the five texture analysis parameters were the only classifiers used by the algorithm and not the time of the day. The same outcome ensues for the thermals type to a lesser extent. The number of cases classified as thermals during daytime outweighs the one classified during night time. However, the number of “nighttime thermals” is not negligible. As we explained in section 2b, this was due to the random nature of the patterns in thermals that in some cases resembles the others category. Concerning the time

FIG. 5. (a) Minimum and maximum temperatures, (b) daily sunshine duration (ABL height estimation only for the magenta colored bars; the days with no ABL height estimation refer to days with long-lasting cloud cover), (c) daily accumulated precipitation, and (d) daily average pressure for the period 4 Sep 2014–6 Nov 2014. Data are from the Paris-Montsouris Météo France station.
period 0800–1800 UTC, the number of occurrences for streaks, rolls, and others were similar, around 20% each, with the category thermals being the most frequent (36.7% of the cases). This was an expected outcome as well, since the weather over Paris was rather sunny during the 2-month period under study, as portrayed in the sunshine hours (Fig. 5b).

It was important to examine more physical parameters in order to assess the automatic classification results. The horizontal wind speed is a parameter that is also distinct for the four types of structures. Figure 8 demonstrates the horizontal mean wind speed distribution (as obtained from the VAD method) associated with the different types of structures. It is apparent that the others and thermals cases were associated with low wind speeds, whereas streaks and rolls occurred during moderate to high wind speeds conditions. More particularly, Fig. 8 confirms that the vast majority of the cases classified as “others,” more than 1500 (75%), were associated with calm (<2 m s⁻¹) or low wind conditions (2–4 m s⁻¹) at the height of the lidar. For wind conditions higher than 4 m s⁻¹, the number of occurrences started to decay drastically, reaching zero for wind speed values higher than 9 m s⁻¹. This result further confirms that the nonsymmetrical radial wind fields caused by the variable winds are correctly classified by the algorithm as “others.” The thermal cases were also associated with light winds, as approximately 80% of the cases (720) occurred when the mean wind speed was lower than 5 m s⁻¹ at the height of the lidar. As other studies have shown, typical wind speed values near the surface during fair weather conditions are below 4 m s⁻¹ (e.g., Kossmann et al. 2002; Mondon and Redelsperger 1998). The number of thermal cases decreases greatly above 5 m s⁻¹ with no cases above 9 m s⁻¹. On the contrary, there are very few cases with calm wind conditions classified as streaks and none classified as rolls. The streaks and rolls were mainly found during moderate or high wind conditions (4–8 m s⁻¹) with approximately 700 (61%) and 330 (79%) cases, respectively. Yagi et al. (2015) found similar wind speed values for the coherent structures measured by sonic anemometers 25 m above ground during a three month study (1 November–31 December) in Tokyo, Japan. Träumner et al. (2015) also retrieved similar mean wind speed values associated with the structures, by using the dual-Doppler technique on horizontal scans recorded over a rural area in Germany for a period of several weeks in spring. Both studies, however, relied on a subjective classification made by experts.

Another useful parameter in order to assess the performance of the classification is the ABL height. As explained in section 2c, only 18 days of the campaign, have almost complete time series of ABL height observations. The time difference between the PPI scans for the estimation of the mlf-cs field and the LOS vertical scans was approximately 10 min. For the PPI and the aerosol lidar scans, the time difference was a few minutes. These time differences are relatively short, and therefore the ABL height was expected to remain steady. In Fig. 9, the distribution of the ABL height in relation to the classified pattern is portrayed via histograms. The ABL height was divided into bins of 200 m each, with the histograms for

<table>
<thead>
<tr>
<th>Type</th>
<th>0000–0800 UTC</th>
<th>0800–1800 UTC</th>
<th>1800–2359 UTC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Streaks</td>
<td>376 (24.4%)</td>
<td>394 (20.8%)</td>
<td>375 (32.7%)</td>
<td>1145 (25%)</td>
</tr>
<tr>
<td>Rolls</td>
<td>27 (1.8%)</td>
<td>365 (19.3%)</td>
<td>28 (2.4%)</td>
<td>420 (9.0%)</td>
</tr>
<tr>
<td>Thermals</td>
<td>98 (6.4%)</td>
<td>695 (36.7%)</td>
<td>107 (9.3%)</td>
<td>900 (20%)</td>
</tr>
<tr>
<td>Others</td>
<td>1036 (67.4%)</td>
<td>438 (23.2%)</td>
<td>638 (55.6%)</td>
<td>2112 (46.0%)</td>
</tr>
<tr>
<td></td>
<td>1537 (100%)</td>
<td>1892 (100%)</td>
<td>1148 (100%)</td>
<td>4577 (100.0%)</td>
</tr>
</tbody>
</table>
each type of structures stacked. There were only 8 cases of the 124 (6%) classified as rolls concurrently with the existence of the SBL (200–400 m). Almost all of the cases classified as rolls correspond to the developing (400–1000 m) or, mainly, to the well-developed ABL (>1000 m). This is in agreement with the physical behavior of rolls formed in the mixed layer of the ABL (Young et al. 2002).

Moreover, the vast majority of the cases classified as others (~400 cases; i.e., 86%) inside the SBL, as expected from the analysis of Table 3. An even larger share of streak cases (240 cases, i.e., 90%) occurred in the SBL. The DBS time–height cross sections revealed that, during 13 of the 18 days for which we could estimate the ABL height, a nllj was present similar to the ones presented in this section (3b) during the case study (Fig. 11a, only the dates for the case study are shown). This confirms that the wind shear under the jet is an important mechanism supporting streak formation. Even if the majority of the cases classified as thermals occurred in the developing or the well-developed ABL (ABL deeper than 400 m), a nonnegligible amount of cases was found in the SBL (more than 50, i.e., 20% of the thermal cases). These misclassifications were rather associated with the random patterns, as we stated before, and not with the atmospheric physics. The possible addition of more/other texture analysis parameters could improve the classifications of the thermals.

b. mlf-cs during a peculiar 4-day period

We examined further the results of the classification by focusing on a period when the three mlf-cs cases (streaks, rolls, and thermals) have been detected (7–10 September). These days were characterized by moderate wind conditions and only few low wind cases. This was important, since as we have...
previously stated, the mlf-cs fields were constructed by utilizing the VAD method that is not always applicable under low wind conditions when the wind field is not homogeneous, for example, during variable winds. Figure 10 presents the wind rose for the period 7–10 September computed from the PPI scans using the VAD method. The wind speed values ranged between 4 and 8 m s\(^{-1}\) for more than 70% of the cases. The wind direction was northeasterly during most of the period.

During this period there were anticyclonic conditions over Paris. Additionally, as we can see in Fig. 5c, there was no precipitation during this 4-day period, but instead plenty of sunshine hours (Fig. 5b). Except on the 7 September when the sunshine duration was approximately 6 h, for 8–10 September, the sunshine duration exceeded 12 h. This allowed a reliable estimation of the ABL height for these days. In Fig. 11a, the ABL height is displayed along with the time–height cross sections of the horizontal wind speed from the DBS soundings. To examine the classification, each type is represented with different color and symbol. The first element to notice in Fig. 11a was the accurate representation of the ABL height development throughout the 4-day period. The ABL height ranged between 200 and 400 m during the night and from 1.2 to 1.8 km during the day. Furthermore, the classification of the cases was realistic with respect to the time and the ABL height for almost the entire 4-day period. Only four cases were misclassified in that regard with three cases of thermals found on 7 September between 2230 and 2330 UTC and one case of thermals around 0300 UTC 9 September. For rolls, no cases were found on 7 September, which is the day with the least sunshine hours. Most of the rolls cases were found on 10 September. Since there was no weather system affecting the region at this date, these structures developed mainly due to the thermal instabilities. The classification of the streaks cases was particularly interesting. It is evident in all four days that the streaks are associated with a high wind shear near the surface. The streaks appear for the first time around 1900 UTC simultaneously with the wind shear presence near the surface. The missing ABL height values on 8 September between 0000 and 0800 UTC and between 1800 and 2000 UTC were due to the absence of LOS vertical wind speed observations and therefore the structure type is plotted at the 0 m altitude. Other missing points were due to the lack of PPI observations.

Additionally, there was a shift in the types of patterns in the early morning on 9 and 10 September from streaks to rolls as...
soon as the wind shear dissipated and the ABL height began to develop. We examined the meteorological parameters for the period 7–10 September 2014, in order to identify the possible changes during the shift in the types of patterns. The temperature, relative humidity, visibility and solar radiation observed at the Paris-Montsouris station are displayed in Figs. 11b and 11c. An interesting aspect to encounter was the reduced visibility of 5 and 7 km during the early morning on 9 and 10 September, respectively (Fig. 11c). The Paris-Montsouris station recorded fog events only at 0200 and 0400 UTC. The reduced visibility signifies possible convections near the surface with the time of the occurrence indicating radiation fog events (Roach et al. 1976). In section 1 we referred to studies that observed the streaks formed during fog events. As the surface heat increased in the morning the smaller structures were replaced by larger ones. The temperature increase during the 3-h period from 0600 to 0900 UTC was at the magnitude of 6°C, which resulted in a drop of relative humidity of 40% for both days.

To reveal possible periodicities for these structures and their corresponding sizes, we computed the spectrums of the radial wind speed in the transverse direction as explained in section 2d. Figure 11d contains the values of the highest peaks of the spectrums representing the sizes of the periodicities for the streaks and rolls cases. The sizes of the periodicities differ between the streaks and rolls cases by a magnitude of at least 500 m. In section 2d, we have already stated that the size of the periodicity referred to the combined red (positive) and blue (negative) stripes (Figs. 2a,b). For the streaks cases, the size ranged from 400 to 800 m, so the width of the stripes would be approximately 200–400 m, respectively. These sizes, as showcased in section 2d, agree with other studies where the spacing was estimated to be around 300 m (Drobinski et al. 2004; Weckwerth et al. 1997). Since the life-span of the streaks is several minutes and the quasi-horizontal PPI scans in our study had an approximate 18-min time difference, most of the streaks we observed dissolved and were replaced by new ones before the next scan occurred, thus we did not observe the evolution of the same structures. For the rolls cases, the sizes ranged from 1.3 to 2 km; therefore, the size of the structures would probably be from close to 600 m to 1 km. Since we observed the structures only near the surface, it is complicated to compare these results with other studies focusing on the rolls structures. The low values of relative humidity indicated that these were dry rolls cases. For the rolls, it is also important to note that, for 3 cases of the 31, it was not possible to estimate the size of the periodicities from the spectrum, which also led to a misclassification of these cases by the algorithm.

4. Conclusions

In the current study we examined the performance of an automated classification method that we developed in order to classify mlf-cs based on Doppler lidar observations recorded during two months over Paris. We intended to identify and classify streaks (narrow stripes), rolls (wide stripes), thermals (enclosed areas), and “others” (impossible to classify). In our previous work, we had presented the algorithm itself and assessed its performance on the training dataset: a classification error of approximately 9% for 150 cases with five texture analysis parameters as classifiers. We then applied these five parameters to the full dataset of 4577 quasi-horizontal scans and in this paper, we analyze the results of the classification from the physical/meteorological point of view.

The results of the classification showed that the most frequent type was the others category with 46%, and then second was the streaks type with 25%. The third most frequent type was the thermals with 20%, and fourth was the rolls with 9%. By examining the occurrence of each type for three different time periods of the day, in particular, 0000–0800 UTC, 0800–1800 UTC, and 1800–2359 UTC, it was possible to examine the distribution of the types during the day. The results showed a realistic distribution as rolls and thermals cases were mainly classified during the day. Streaks cases were mostly classified during the night, which is an expected result as nlj events occurred during 20 of the 63 nights of the study period. For the category “others,” the results were also expected, because the majority of the cases occurred during the night when the low winds result in asymmetric radial wind fields and the VAD method cannot be applied. This observation was confirmed by the distribution of the classified cases with regard to the horizontal wind speed: for 75% of the others category occurred when the horizontal wind speed was lower than 4 m·s⁻¹. On the contrary, for streaks and rolls that are driven by wind shear, the majority of the cases were characterized by moderate or high winds (4–10 m·s⁻¹) for approximately 67% and 89% of the cases, respectively. The accurate representation of the types as a function of the time of day was further supported by the distribution of the estimated ABL height. For the 18 days, during which it was possible to estimate the ABL height for the majority of the 24h, the vast majority of the rolls cases occurred in a developing or developed ABL, whereas the vast majority of streaks and others cases were found in the SBL or the nocturnal boundary layer. The thermals had the worst performance, but nevertheless 80% of the cases were found in the developed or developing ABL.

The investigation for a 4-day period, when the three types of structures were classified, showed that the streaks occur at the same time as nlj and dissolve during the morning when the thermal instabilities prevail. During the two days with more than 12 sunshine hours, which resulted in strong thermal instabilities, rolls cases were classified. The spectrum of the radial wind speed in the direction transverse to the horizontal wind vector was estimated in order to detect periodicities during the two days when the streaks cases were followed by rolls. The spectra exhibited prominent peaks for almost all but three rolls cases. The periodicities of the streaks and rolls were clearly separated, with at least a 500-m difference. For the streaks, the sizes of periodicities ranged between 400 and 800 m; for rolls, they were between 1.3 and 2.0 km, indicating realistic sizes of the structures that would be around 200 and 400 m for the streak spacing and more than 1 km for the rolls.

Overall, the performance of the algorithm with regard to the classification of the structures based solely on texture analysis parameters is in agreement with the physical parameters. The results were as expected, especially for the coherent structures rolls and streaks. However, there is room for improvement, especially for the category of thermals.
Acknowledgments. The authors thank François Raveteta, Jacques Pelon, Gilles Plattner, and Amelie Klein of the LATMOS, Sorbonne University, Paris, for organizing and carrying out the VEGILOT campaign. We acknowledge the use of imagery from the NASA Worldview application (https://worldview.earthdata.nasa.gov/), last accessed 2 December 2020, part of the NASA Earth Observing System Data and Information System (EOSDIS). Experiments presented in this paper were carried out using the CALCULCO computing platform, supported by the Service Commun du Système d’Information de l’Université du Littoral Côte d’Opale (SCoSUL ULCO). This work is a contribution to the Contrat de Plan Etat-Région (CPER) research project Innovation et Recherche en Environnement (IReN) and Climibo. The work is supported by the French Ministère de l’Enseignement Supérieur, de la Recherche et de l’Innovation, the region Hauts-de-France, and the European Regional Development Fund. The work is also supported by the Chemical and Physical Properties of the Atmosphere (CaPPA) project. The CaPPA project is funded by the French National Research Agency (ANR) through the Programme d’Investissement d’Avenir (PIA; Contract ANR-11-LABX-0005-01) and by the regional council of Nord-Pas-de-Calais and the European Regional Development Fund. The work was carried out with the financial support of the project by the Russian Federation represented by the Ministry of Science and Higher Education of the Russian Federation, Agreement 075-15-2020-776.

Data availability statement. All lidar data used in the study are property of the Laboratoire Atmosphères, Milieux, Observations Spatiales (LATMOS), Paris, France, and of the Laboratoire de Physico-Chimie de l’Atmosphère (LPCA), Dunkirk, France, and are not publicly available.

REFERENCES


