Characterization of local wind patterns in complex mountain valleys

A. Pérez-Foguet*

Research Institute for Sustainability Science Technology IS.UPC, Laboratori de Càlcul Numèric LACAN, Applied Math Dept. MA3, Civil Engineering School ETSECCPB, Universitat Politècnica de Catalunya – BarcelonaTech (UPC), Jordi Girona 31, UPC Campus Nord C2-206 08034, Barcelona, Spain

ABSTRACT: In this work, the wind patterns in high mountain areas with complex orography are characterized using hourly data provided by a network of weather stations. The key novelty of the study is the methodology. Data are grouped separately by wind speed and wind direction using two cluster analyses. The groups are analysed and described according to measurements at key stations in the network and their hourly presence. Both classifications are subsequently compared using contingency tables, and the main wind patterns are identified. The uncertainties associated with the average values of each wind pattern are quantified by principal component analysis of the wind vectors at each station. One year of data from nine stations located in the area of La Oroya, Peru was used to validate the proposed method. The local wind behaviour was characterized, the wind patterns were compared with respect to the seasons, and the winter morning transitions were analysed in detail. The methodology allows quantitative description of the local wind patterns and their temporal dynamics in complex mountain valleys. Both wind speed and direction were found to be relevant in wind pattern characterization. In particular, both parameters have proven helpful in identifying and quantifying the prevailing winds during cold dawns and thermal inversion periods.

KEY WORDS wind regime; wind transition; thermal inversion; Andean region; cluster analysis; principal components analysis

1. Introduction

Wind patterns characterize the most current configurations of the wind field in a given area, and identifying and analysing these patterns is of upmost importance for air quality studies (Darby, 2005; Ratto *et al.*, 2010b) in the field of renewable energy (Gómez-Muñoz and Porta-Gandara, 2002; Burlando *et al.*, 2008; Jiménez *et al.*, 2008), and for characterization of potentially hazardous atmospheric conditions (Jiménez *et al.*, 2009; Ferragut *et al.*, 2010). Specific study of wind patterns in complex orography areas and mountainous regions has been addressed in several previous studies (Kastendeuch and Kaufmann, 1997; Kaufmann and Whiteman, 1999; Drobinski *et al.*, 2003; Ludwig *et al.*, 2004; Beaver and Palazoglu, 2006; Burlando *et al.*, 2008; Jiménez *et al.*, 2008, 2009).

In complex terrain, the airflow is modified on a wide range of scales depending on the nature of the forcing, which can be mechanical or thermal (Drobinski *et al.*, 2003). Dynamically and thermally driven wind systems and their interactions generate a wide variety of flow patterns (Jiménez *et al.*, 2008). Whiteman and Doran (1993), who studied the relationship between the synoptic scale and the flow within a valley, suggested that thermal forcing occurs when the large-scale dynamic force is weak. At smaller scales, pure thermally driven circulations are generated in the absence of large-scale forcing (Kaufmann and Weber, 1998). When regional winds are light and the skies are cloud-free during the night periods, the valley surfaces and the air in contact with these surfaces are cooled. This cool, stable and dense air drains downhill into the bottom of the valley inducing down-slope and down-valley winds. During the day, mountain slopes that face the sun are heated, and the air in contact with these slopes rises up the mountain slopes because of its buoyancy, thus manifesting as up-slope and up-valley wind (Doran *et al.*, 1990; Guardans and Palomino, 1995). Furthermore, thermally driven winds at the large scale can overwhelm those at the local scale (Stewart *et al.*, 2002). In contrast, the near-surface winds over complex terrain are often channelled by valleys (Kaufmann and Weber, 1998). The mechanical influence of topography produces a strong channelling effect on winds generated at the large scale by synoptic pressure gradients (Drobinski *et al.*, 2003).

Two multivariate statistical techniques are commonly used to classify wind patterns, namely, principal component analysis (PCA) and cluster analysis (CA). Both techniques are intensively applied in many scientific fields for

* Correspondence to: A. Pérez-Foguet, Research Institute for Sustainability Science and Technology IS.UPC, Laboratori de Càlcul Numèric LACAN, Applied Math Dept. MA3, Civil Engineering School ETSECCPB, Universitat Politècnica de Catalunya – BarcelonaTech (UPC), Jordi Girona 31, UPC Campus Nord C2-206, 08034 – Barcelona, Spain. Email: agusti.perez@upc.edu

data reduction, interpretation and classification (Muñoz-Díaz and Rodrigo, 2004; Esteban *et al.*, 2006; Burlando *et al.*, 2008). The PCA technique defines a linear approximation similar to the original dataset. In identifying wind patterns, this approximation corresponds to the average values in the stations plus the directions of the greatest variability of the data or the highest eigenvectors of the PCA (Ludwig *et al.*, 2004). Various reduction strategies using PCA have been implemented for hourly (Hardy and Walton, 1978) and daily data (Kalkstein *et al.*, 1987). In particular, PCA has been used successfully by Ludwig *et al.* (2004) to identify wind patterns in mountain valleys. In contrast, the CA technique is based on the assumption that data with similar measurements can be grouped into reduced sets of representative cases. The CA seeks to directly identify attractors in the space of the wind patterns, or in other words, configurations with additional probability density (Burlando *et al.*, 2008), and CA has been used with data sources to group them for regionalization and with the data itself to identify wind patterns. Kaufmann and Weber (1998) presented a CA that can identify regions using the correlation between wind directions. Burlando *et al.* (2008) characterized regions and wind patterns by comparing different measurements and aggregation schemes, thus applying CA to wind speeds. Using original data facilitates interpretation of the results. However, it is convenient at times to reduce the original variables using PCA and to subsequently perform CA on the sample expression according to the main factors if the domain areas are over-represented or the number of variables is too high (Kalkstein *et al.*, 1987; Esteban *et al.*, 2006; Jiménez *et al.*, 2008).

To identify synoptic behaviour, it is common to perform CA with only the wind speed (Gómez-Muñoz and Porta-Gandara, 2002; Burlando *et al.*, 2008; Jiménez *et al.*, 2008). Darby (2005) performed this analysis using wind vectors, thereby including both wind speed and direction. To prevent high speeds from overly influencing the results, Kaufmann and Weber (1996) and Kaufmann and Whiteman (1999) used standardized speeds to analyse areas of complex orography, and thus their data primarily reflect wind direction. Kaufmann and Weber (1998) ran an analysis using only the wind direction in mountain valleys, whereas Ratto *et al.* (2010a, 2010b) performed the analysis grouping the directions into eight categories. Aikawa *et al.* (2008) presented a CA that integrates diverse information sources, including wind speed and direction.

Several clustering techniques can be found in the related literature. One of the simplest techniques is the non-hierarchical k-means method. This approach has the advantage of easily processing large data pools and usually producing groups of similar sizes. It requires an input of the initial centres of groups and their number. Proposals for initial centres can be set by randomly choosing data samples, especially in cases involving a large number of samples, because the method allows group reassignment. The number of centres can be increased gradually to facilitate interpretation of the results. The goodness of the classifications can be evaluated with different measures and indices (Davies and Bouldin, 1979; Ray and Turi, 1999). These measures or practical reasons are used to fix the number of clusters with non-hierarchical methods. Although Darby (2005) also used a non-hierarchical method, Ratto *et al.* (2010a, 2010b) used a hierarchical approach. Two-step schemes are common, first applying a hierarchical method to obtain an initial proposal of groups followed by a non-hierarchical method used to classify the groups (Kaufmann and Weber, 1996; Kaufmann and Whiteman, 1999; Burlando *et al.*, 2008; Jiménez *et al.*, 2008).

In this article, CA is proposed for separate clustering of the samples by wind speed and wind direction to characterize the wind patterns in complex mountain valleys. The strategy proposed in this work follows a similar approach to that of Aikawa *et al.* (2008), except that the previous work focused on hourly data obtained solely for wind and treated the wind speed and direction separately. In this way, the influence for each condition in the wind data can be identified independently, aiding in the identification of wind patterns. The k-means method is used in the CA of both wind speed and wind direction.

The approach was applied to analyse and characterize local wind patterns in the La Oroya region in the Andes in Peru. For this purpose, hourly wind data were collected from a network of nine stations located along the two main valleys of the region and at different elevations. In this study, 9431 cases that occurred in consecutive sequences of at least 1 d (with 75% of the sequences lasting at least 10 d) were taken into account. The obtained results were compared with the general descriptions of typical wind patterns in the area (Culqui, 2001; McVehil, 2005, C.N.A., 2006, 2007; Clark *et al.*, 2006; D.E. & B.O.C., 2007). In particular, the proposed methodology of separate processing of wind speed and direction aided in the description of the transition from nighttime to daytime behaviours, although this is known as a highly complex dynamic (McVehil, 2005). The characterization of wind patterns aims to better describe the local wind dynamics, which are critical for local air quality modelling and simulation. Several authors have modelled the meteorological and air quality conditions in the region (McVehil, 2005; Clark *et al.*, 2006; Zarauz and Pasken, 2010). However, none of the previous studies have revealed the detailed behaviours in the boundary layer, which directly affect the immission values close to the emitter. The regional impact of industrial complex emissions has been also identified using satellite data (Carn *et al.*, 2007; Khokhar *et al.*, 2008; Zarauz, 2011).

A case study is presented in the following section, the proposed methodology is described in the third section, and the results are discussed in the fourth section. Overall, we show how the wind patterns were characterized and how the available references are reproduced.

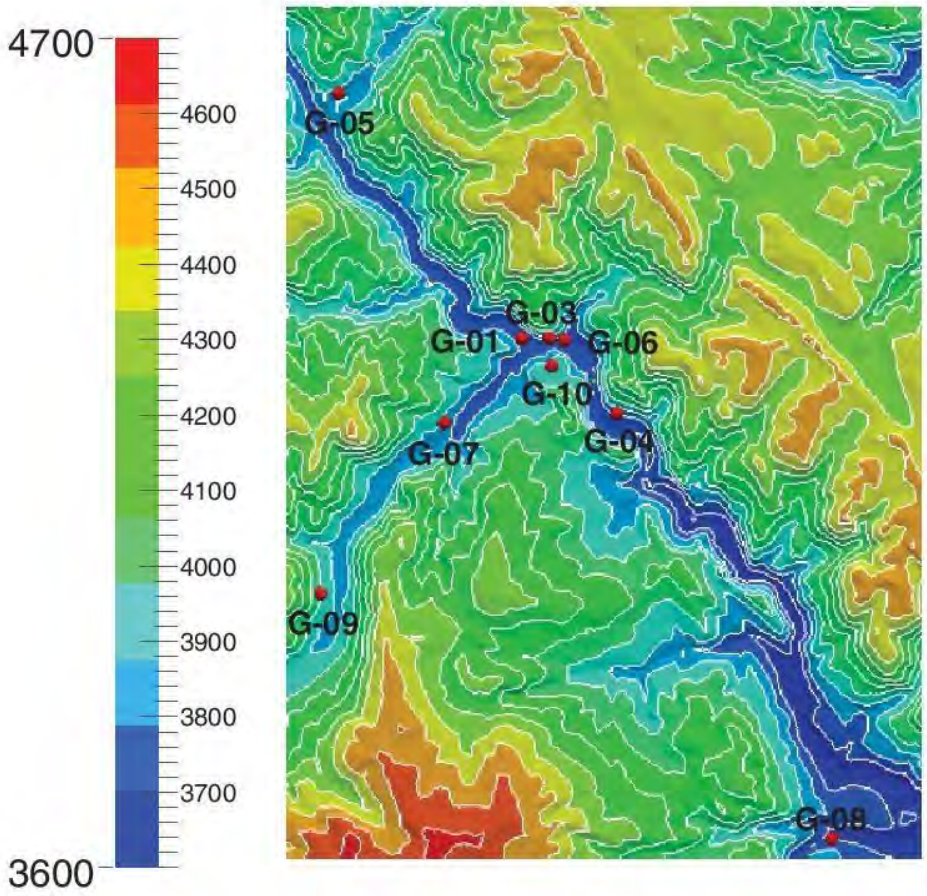

Figure 1. Orography of the La Oroya region. The nine stations are indicated G01 and G03 to G10. Region at 150 km from Lima between Cerro de Pasco and Huancayo.

Finally, the main conclusions are summarized in the final section.

2. Case study: La Oroya, Perú

Figure 1 shows the location of the study area in La Oroya, Peru. The zone is situated in the Andes at 3700–4500 m above sea level and east of the continental divide. The orography of the area is complex, with two main valleys crossing the area, one of which is oriented SE–NW and the other running SW–NE until it intersects with the first. The city and the metallurgical complex of La Oroya are situated at this valley intersection (C.N.A., 2006, 2007).

The locations of the nine stations are denoted in Figure 1. The position of G10 stands out because it is elevated and located at the junction of the main valleys, whereas G03 and G06 are situated at the valley bottom at approximately the same position, and G01 is located slightly to the north. The remaining stations are located at the ends and the central sections of the valleys (NW, SW, and SE). Neither G05 nor G08 (in the NW and SE ends, respectively) are located in the main valley but instead lie in perpendicular tributary valleys. Table 1 presents the names, altitudes, and the areas in which the stations are located. The stations were renovated at the beginning of 2007 (D.E. & B.O.C., 2007). The C.N.A. (2007) lists the exact location of the G01, G03, G07, and G08 stations. Two weather stations (G06 and G10) and the seven immission monitoring stations also exist (G01, G03, G04, G05, G07, G08, and G09). The data for wind speed and direction as well as temperature were available from all stations, and solar radiation and atmospheric pressure data were additionally available from G06.

Table 1. Label, name, altitude, and area of the monitoring stations (D.E. & B.O.C., 2007; C.N.A., 2007).

G01	Hotel Inca	3734 m Central Zone
G03	Sindicato de Obreros	3736 m Central Zone
G04	Huanchan	3790 m SE Valley
G05	Casaracra	3779 m NW Valley
G06	Fundición	3721 m Central Zone
G07	Marcavalle	3780 m SW Valley
G08	Huari	3684 m SE Valley
G09	Huaynacancha	3860 m SW Valley
G10	Cerro Sumi	3936 m Central Zone

A general overview of the characteristics of the local meteorology is available from Culqui (2001), McVehil (2005), Clark *et al.* (2006), C.N.A. (2006, 2007), and D.E. & B.O.C. (2007). These studies include the differences between the dominant winds during the day

and night periods, the seasonal characteristics, the presence of thermal inversions, and the altitude of the boundary layer elevation. Wind data for certain time slots at the approximate G06 location are presented in Culqui (2001) and Clark *et al.* (2006), and a general description of the transition from night-time to daytime behaviour is available in McVehil (2005); the morning behavioural data can be found in the C.N.A. (2007). Four meteorological seasons were considered (Culqui, 2001): (1) wet summer (January to March), (2) dry winter (May to August), (3) sub-humid spring (October to December), and (4) transition (April and September).

This study uses hourly data for the period from June 2007 to October 2008. A cold phase of the El Niño/Southern Oscillation (ENSO) was initiated in September 2007 and disappeared by June 2008, and thus the summer conditions included in the analysis is an example of the 'La Niña' phenomenon. The analysis focuses on winter wind characterization from May to August, which can be considered representative because there is no reason to indicate otherwise. However, systematic errors have been identified in measurements of selected stations in other time periods (Clark *et al.*, 2006). Although it is desirable to identify ex ante errors in the data, errors that are credible, systematic, and present throughout the entire time period do not invalidate the analysis and might be also useful in relative terms.

3. Methodology

The proposed methodology is summarized in Table 2. After a first step of data pre-processing, (CAs) are computed by wind speed and wind direction. Both classifications are compared using contingency tables, which allows the identification of wind patterns. The process ends with characterization of the wind pattern uncertainties and detailed characteristics.

Table 2. Algorithm for characterization of wind patterns.

1. Data pre-processing
   a. Quality requirements
   b. Continuity requirements
   c. Representativeness analysis
2. Cluster analyses by wind speeds
   a. Hourly distribution and description
3. Cluster analyses by wind directions
   a. Hourly distribution and description
4. Identification of wind patterns by contingency tables
5. Characterization of wind patterns
   a. Principal component analysis to quantify uncertainty
   b. Complementary analyses
      (i) Comparative analysis between seasons
      (ii) Identification of transitions
      (iii) Segmented analysis of other data

First, all available data were pre-processed. The functional status of each station was validated by revising the different available measurements for each one. Samples with erroneous data were discarded (negative speeds, pressures that were too low, etc.). Three or more consecutive samples without a certain measurement were also deleted. All gaps of one or two consecutive samples were completed via linear interpolation. The requirement for data presence in periods of at least 24 consecutive hours was imposed, and the hourly data that did not meet this constraint were discarded. Once the working data were selected, the sample representativeness was contrasted with respect to the distribution for hourly, monthly, and yearly seasons.

Next, the wind speed and the unit vector of the wind direction were used to define different sets of clusters. The wind speed is defined as the modulus of the wind velocity vector, $v = (v_x{}^2 + v_y{}^2)^{1/2}$, given the two components of the wind v_x and v_y. The unit vector of the wind direction is defined by the two components n_x and n_y given by $n_x = v_x/v$ and $n_y = v_y/v$, and its length is equal to one. The Euclidean distance is used to compare pairs of sets of wind speeds or wind directions.

Both aggregations were computed for four, six, and eight clusters with the k-means method. Links between clusters could be identified by comparing the results obtained from increasing the number of groups. The automatic labelling of the groups was recoded to improve the interpretability of the results (if there are several groups with notably few cases, which usually occurs when clustering anomalous data, and if this significantly reduces the number of valid groups, it could be necessary to remove the data involved and repeat the pre-processing step).

Both aggregations were analysed and described based on wind speed and direction at key stations of the network together with their hourly distributions, which were represented by heat-plots. Because the valleys winds show strong daily dynamics, hourly analysis allows the classifications to be interpreted and linked.

The classifications were directly compared using contingency tables (step four in Table 2), which were used to identify the relevant and singular groups. In this work, the contingency tables between classifications were constructed using equal numbers of groups in each: 4, 6, and 8. Up to 64 possibilities give an adequate number for results analysis (in the cases in which additional disaggregation is required, it may be more convenient to segment the sample prior to the analysis). Only the cells that displayed a higher number of cases than the total divided by the number of cells (i.e. those equivalent to a uniform distribution) are shown. Subgroups with a relatively large number of cases with respect to the total number (or the sum by column or row) were analysed as well as those with a singular distribution. The wind patterns were identified from these subgroups.

The wind patterns are presented using the wind average from each station and its associated uncertainty. To characterize the uncertainty, PCA was applied to the

wind data from each station and for each identified wind pattern. The cluster variances approximate the behaviour of each wind attractor at each station. The PCA results provide a description, with both direction and magnitude, of the intra-group variability. The square roots of the eigenvalues of the PCA represent the data variances and the eigenvectors represent their directions. The rotated ellipse described by this pair of vectors defines the area of uncertainty. If the speed probability distribution at each station was close to a bivariate normal distribution, approximately 39 per cent of the cases from each group were located in the ellipse areas (Crutcher and Baer, 1962). Representing the wind patterns graphically with the average winds and their uncertainty facilitates interpretation. In Figures 5 and 6 presented hereafter, the directions of the two eigenvectors are explicitly indicated on the ellipse with dots.

The last point in Table 2 includes complementary analyses that can help to identify the characteristics of the wind patterns and their dynamics. The differences between the contingency tables of different periods of the year were calculated to identify the seasonal variations. In this work, the configuration of the winter months (May to August) has been analysed in detail, with a special focus on the morning time slot from 6:00 to 11:00. Interest in the morning wind evolution stems from its possible connection with reported high immission episodes.

Once the wind patterns were identified, the transitions between them were analysed as a first approximation to the wind dynamics (Guardans and Palomino, 1995; Darby, 2005). To better identify these transitions, subgroups from the initial classification were re-defined by focusing on interesting combinations. The total sample was maintained, and cases were not excluded to retain the daily sequences. The transitions were quantified by comparing the wind patterns of each data with those from the following hour. The most relevant transitions were those of higher magnitude in the corresponding contingency table, with the exception of the diagonal values of the table, which represent the persistence of wind patterns (transitions within the same subgroup).

Finally, segmentation and analysis of other available data complemented the description of wind pattern characteristics. Specifically, temperature data from stations G06 and G10 (which appear close in the map but lie at different altitudes; see Figure 1 and Table 1) and solar radiation data from G06 were analysed with the aim of characterizing the thermal inversion situations.

4. Results

The initial dataset obtained from June 2007 to October 2008 with a total of 12 456 cases was reduced to 9431 cases after the initial pre-processing step. Data available after the quality requirements included 10 110 cases, which were reduced after imposing the 24-h continuity requirement. Data were consecutive over periods of at least 1 d, with 75% of the time covered by series of at least 10 consecutive days (7092 cases) and 56% by a series of at least 20 d (5268 cases). The hourly distribution was uniform (with 389–399 cases for every hour of the day), but the monthly distribution was not (with 376–1372 cases depending on the month). Cases grouped by seasons were distributed as follows: humid summer (3 months), 1464 cases; dry winter (4 months), 3952 cases; sub-humid spring (3 months), 2505 cases; and transition (2 months), 1510 cases. Thus, the monthly average ranges from 500 to 1000 cases/month (summer and winter). The bias of the sample was taken into consideration when interpreting the results. A comparative analysis between summer and winter aided in characterizing the wind patterns.

In Figure 2, the wind velocity vectors are plotted for the stations in the central zone (G06 and G10), the NW valley (G05), and the SW valley (G09). The ends of the wind velocity vectors are shown (and not the origin), and the positive x-axis is oriented in the east direction. Additional relevant directions and maximum speeds can be identified for each station, but it is clear that it is difficult to progress further without a method to analyse the data. The main results of the above-presented methodology are analysed in the following subsections. First, the wind patterns are identified from the wind speed and wind direction data. Next, the main seasonal differences are described. After that, work focuses on winter conditions, especially on the morning periods. The PCA representation of the wind patterns is used to improve the interpretability of the results. This section finishes with an analysis of the transitions between the wind patterns during the winter morning periods and an analysis of the thermal and solar radiation data segmented by wind pattern.

4.1. Cluster analyses: wind speed and direction influences

Groups were constructed using wind speed, and classifications were named in terms of wind patterns with n velocities (WPnV), with $n = 4$, 6, and 8 groups for each, whereas the groups were named as WPnV_j, with $j = 1 \ldots n$. Figure 3 shows the cumulative distribution function for the wind speed for the same stations as in Figure 2. The group WP8V_8 contains only three members. This group corresponds to data with an error not previously detected in G03 (zero speed values for 3 h in a row). The analysis was not repeated because the resulting breakdown was considered sufficient. The CA groups were renumbered to facilitate the analysis.

Most distributions are unimodal and symmetrical. The groups WP8V_1 to _3 are distinguished by high values in four stations (e.g. up to 12 m s^{-1} at G10 in 80% of WP8V_1 cases, which is the group with the highest speeds). The groups WP8V_5 to _7 display low average speeds in all cases (below 4 m s^{-1}) with similarities observed between them depending on the station. The WP8V_7 groups the lowest values of G10, approximately 2 m s^{-1}, and is the group with most homogeneous speed

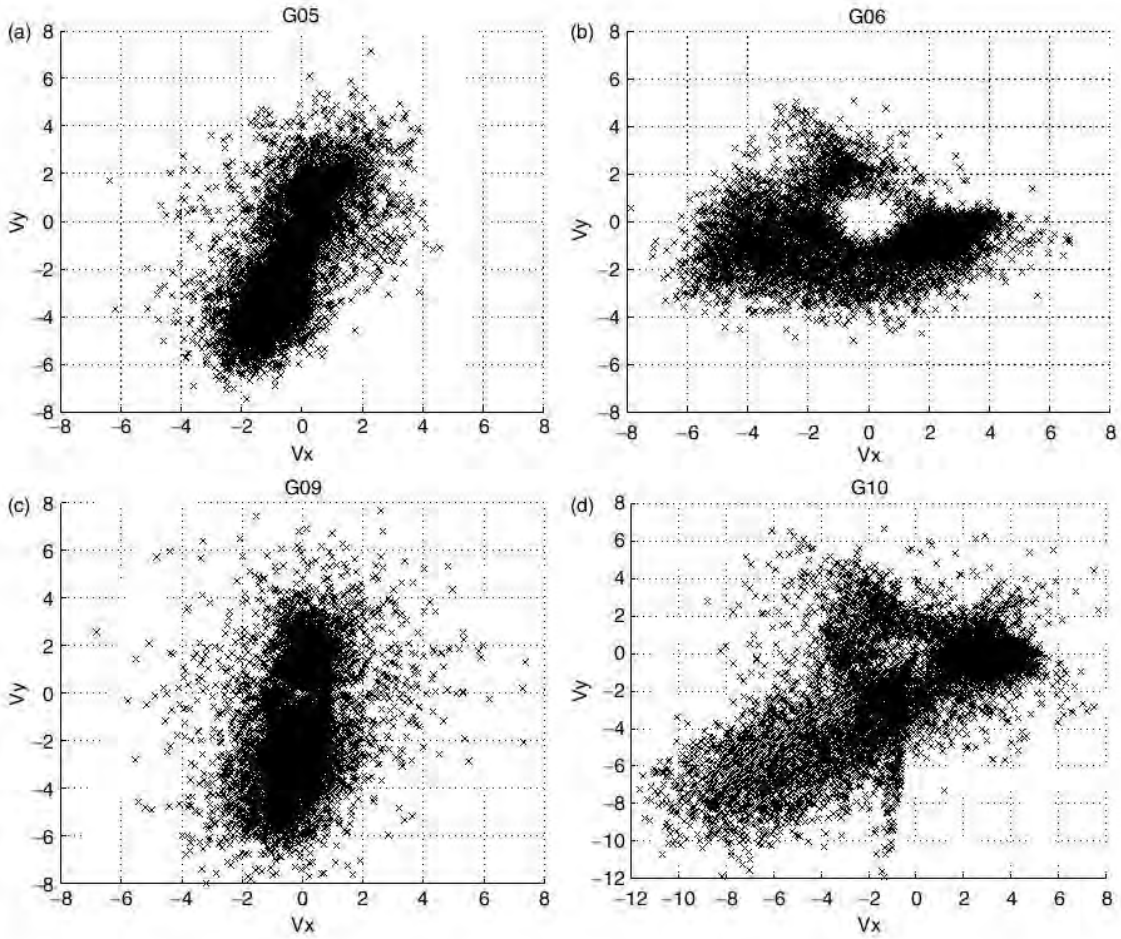

Figure 2. Wind velocity vectors: stations (a) G05, (b) G06, (c) G09, and (d) G10.

values among the stations (although in other stations, other groups had lower average speeds of $1-1.5\,\mathrm{m\,s^{-1}}$).

Table 3 shows the hourly distribution of the groups in the WP8V and WP4V classifications. In the WP4V classification, the day–night dynamics clearly stand out. The intermediate speed groups can be distinguished in WP8V, and the night winds can be separated into two groups. The groups WP4V_1 and _2 correspond to WP8V_1 and _2, WP4V_3 to WP8V_3, _4, and _5, and WP4V_4 to WP8V_6 and _7.

The following classification is proposed based on the wind speed clustering:

- Mostly day winds: Very strong, strong, light, and very light (WP8V_1 to _4, respectively). These winds appeared from 10:00 to 23:00 but were predominant from 11:00 to 20:00. The strong and very strong values were concentrated between 14:00 and 19:00.
- Mostly night winds with two main typologies: Non-uniform, with a significant presence from 21:00 to 7:00, which can be classified as non-uniform/light (WP8V_5) or non-uniform/very light (WP8V_6), and uniform/very light (WP8V_7; present from 21:00 to 12:00 but predominantly between 8:00 and 10:00).

The groups were also constructed based on wind direction. The classifications were denoted as wind patterns with n directions (WPnD) with $n=4$, 6, and 8 group numbers in each; the groups were named as WPnD_j, with $j=1\ldots n$. Figure 4 shows the cumulative distribution function for the azimuth for the same stations as in Figures 2 and 3. The reference angles can be distinguished in each case, with unimodal (WP8D_4 and _6 in G10, and WP8D_8, _7, and_6 in G09) or bimodal distributions (WP8D_1 in G10 and G05, and WP8D_4 in G09). The preferential directions for each station were also evident (west, 270° and NE, 45°, for G10 and G06; north, 0° and south, 180°, for G09; and NNE and SW for G05). The uniqueness of the WP8D_3 group stands out because it has an SE orientation in G10 and an SSE orientation in G06. However, the groups in G06 generally followed the same orientations as in G10, with the exception of WP8D_1, which presents a unimodal distribution in the west. In addition, except for WP8D_3 with an SSE orientation, the remainder of the G06 groups showed no events with a south orientation. This particular effect can be attributed to the location in the urban area of La Oroya. Finally, G05 and G09 show variations in orientation between each other and relative to the directions of G10 and G06, as discussed later (see Sections 4.3 and 4.4).

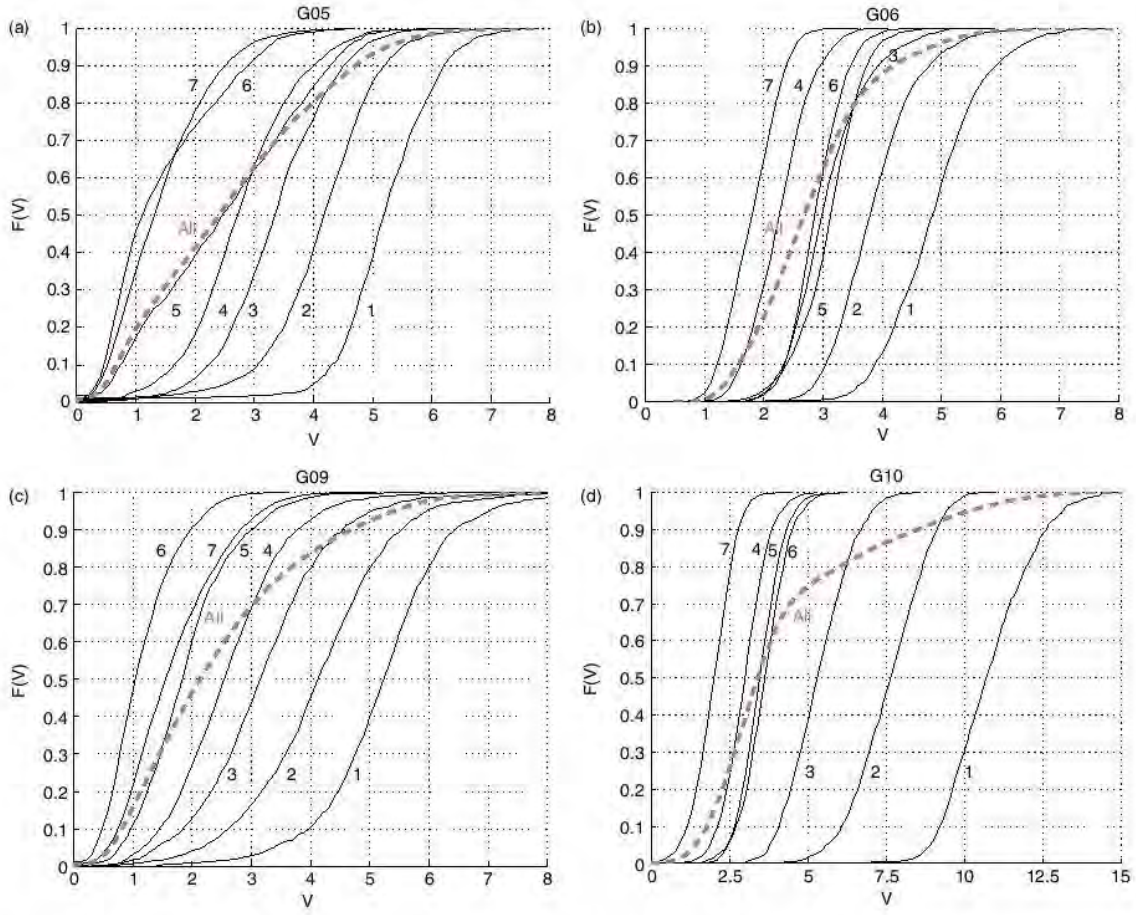

Figure 3. Cumulative distribution function for wind speed ($m\,s^{-1}$): groups WP8V_1 to _8 and stations (a) G05, (b) G06, (c) G09, and (d) G10.

Table 3 shows the distribution of the groups during the day with the resulting classifications for the WP8D and WP4D. The characterization of the morning transition stands out. The WP4D aggregation distinguishes the predominant winds from 7:00 to 10:00 and WP4D_1 and _2 from 10:00 to 12:00. Among the classifications, WP4D_1 can be identified with WP8D_1; WP4D_2 with WP8D_2 and _3; WP4D_3 with WP8D_4, _5 and _6; and WP4D_4 with WP8D_7 and _8.

The proposed classification of wind direction is summarized below:

- Morning wind (WP8D_1, from 8:00 to 10:00) and midday wind (WP8D_2, from 10:00 to 12:00): Local winds oriented along the main valley and the valley running SW.
- Day wind from NE (WP8D_4 and WP8D_6, from 12:00 to 21:00): Regional wind.
- Evening wind (WP8D_5, from 20:00 to 23:00): Wind with a nocturnal direction in G09 and G05 (descending winds through both valleys) but a diurnal direction in G10.
- Night and dawn wind (WP8D_8, from 21:00 to 7:00 and WP8D_7, from 1:00 to 8:00): Descending winds through both valleys.
- SE wind (WP8D_3, from 10:00 to 15:00). This wind was characterized in G10 and G06 by a differential direction with respect to that of the other groups.

The stations shown in Figure 3 all displayed the same behaviour type in the WP8D_6 and WP8D_4 groups (daytime, from NE) and in the WP8D_7 and WP8D_8 groups (night and dawn). Differences within each pair of groups are apparent in G08 and are discussed below (see Section 4.3).

4.2. Contingency tables: wind pattern classification

Table 4 shows the contingency tables between the cluster classifications based on wind speed and wind direction. Only values higher than the total number of cases divided by the possible subgroups are displayed (the thresholds used for this analysis are indicated in each sub-table). The most prevalent and singular groups stand out. The groups WP8D_2, _3, and _5 present the lowest percentages of occurrence (6, 7, and 4%, respectively), whereas the groups WP8D_1, _4, and _8 present the highest (16, 20, and 24%, respectively). The distribution based on speed is more homogeneous (with rates of 8, 11, and 17%, respectively). An exception is the group WP8V_7, with 26% of cases. This classification groups the night winds

Table 3. Heat maps displaying the hourly distribution of the classification groups WP4V, WP8V, WP4D, and WP8D.

N Cases	392	392	391	389	389	390	389	389	389	392	393	395	396	395	398	397	399	397	395	394	394	392	392	392
	Hours																							
WP4V	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1												4%	9%	23%	40%	52%	54%	43%	11%	2%				
2											6%	23%	38%	46%	44%	31%	28%	32%	51%	27%	11%	5%	2%	
3	48%	45%	38%	32%	31%	24%	24%	13%	4%	7%	23%	36%	39%	25%	12%	12%	14%	20%	27%	50%	52%	52%	52%	52%
4	51%	55%	62%	68%	69%	75%	75%	87%	95%	92%	70%	38%	14%	7%	4%	4%	5%	6%	11%	21%	37%	43%	46%	47%
WP8V	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1												2%	5%	14%	29%	44%	46%	32%	5%					
2											2%	9%	24%	37%	39%	35%	27%	31%	36%	9%	3%			
3	2%					2%	2%			2%	10%	24%	30%	27%	19%	10%	12%	16%	31%	36%	15%	10%	3%	2%
4	12%	8%	9%	6%	4%	4%	7%	5%	10%	29%	52%	52%	34%	17%	8%	7%	6%	8%	9%	25%	30%	21%	19%	16%
5	28%	25%	16%	12%	12%	8%	8%	6%				2%	3%	2%	3%	3%	7%	6%	7%	12%	20%	27%	32%	31%
6	28%	32%	41%	49%	50%	51%	46%	31%	5%										5%	7%	12%	13%	21%	22%
7	30%	34%	32%	32%	33%	35%	38%	57%	84%	68%	34%	11%	4%	2%	2%	2%	2%	5%	6%	11%	19%	28%	25%	29%
WP4D	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	17%	17%	16%	17%	19%	19%	25%	47%	71%	54%	32%	24%	15%	11%	8%	7%	6%	5%	6%	8%	10%	13%	14%	11%
2	5%	5%	4%	5%	4%	6%	6%	10%	19%	39%	54%	45%	28%	17%	12%	8%	8%	9%	6%	6%	6%	6%	5%	6%
3	5%	5%	5%	5%	3%		2%	2%	3%	6%	13%	31%	55%	69%	79%	86%	78%	76%	72%	61%	40%	23%	16%	12%
4	73%	73%	75%	73%	75%	74%	67%	42%	7%	2%				3%	2%	5%	9%	11%	16%	25%	44%	58%	66%	71%
WP8D	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	12%	13%	12%	11%	13%	15%	18%	36%	64%	48%	25%	21%	13%	9%	6%	5%	6%	4%	4%	6%	6%	10%	10%	9%
2								3%	14%	35%	45%	27%	9%	3%										
3	4%	4%	4%	3%	3%	5%	5%	6%	6%	9%	15%	19%	18%	13%	9%	7%	7%	7%	5%	5%	6%	6%	5%	6%
4	2%			2%					3%	5%	11%	26%	47%	57%	69%	67%	62%	53%	37%	11%	4%	4%	3%	2%
5	3%	3%	3%	3%	2%	3%	2%	2%					3%	3%	3%	6%	7%	8%	8%	13%	11%	9%	5%	3%
6	5%	3%	2%	2%							3%	6%	9%	12%	9%	11%	10%	17%	29%	43%	31%	17%	11%	9%
7	15%	23%	28%	39%	42%	42%	41%	37%	8%								2%	3%	3%	3%	5%	3%	8%	8%
8	60%	52%	50%	40%	37%	34%	32%	15%	4%					2%	2%	3%	7%	7%	14%	19%	37%	52%	57%	63%

Percentage w.r.t. daily cases indicated in the cells (for values >1%).

with uniform speed and is related to all directions except those of the day from the NE (WP8D_4 and _6). The very light day wind (WP8V_4) can be also associated with most groups according to the wind direction. The day groups with strong or very strong wind (WP8V_1 and _2, respectively) relate to the day group with the NE orientation (WP8D_4) and the light night group (WP8V_5) with the night direction group (WP8D_8). The very light night group (WP8V_6) was related mostly to the direction groups of dawn and night (WP8D_7 and _8, respectively). Therefore, the results of both classifications are consistent.

Table 5 summarizes the wind patterns found in the area. The main patterns are described and distinguished by the direction classification and predominant presence time:

- Morning winds aligned in the valleys with wind speeds mainly uniform/very light, which were predominant from 8:00 to 10:00 and had average speeds in the range of 1–2 $m\,s^{-1}$.
- Day wind from the NE, regional, with wind speed groups from very strong (over 12 $m\,s^{-1}$ in G10 and 6 $m\,s^{-1}$ in the other stations) to light and predominant from 13:00 to 19:00.
- Night winds, descending in both valleys with wind speed groups of night/light, night/very light, and uniform/very light; these winds were quite predominant from 21:00 to 4:00 with average speeds below 4 $m\,s^{-1}$ in G10 and 3 $m\,s^{-1}$ in the other stations.

In addition, the following secondary patterns were identified in the transitions between the main patterns: dawn wind, midday wind, and evening wind. Finally, the SE wind was identified with the speed groupings of uniform/very light and day/very light.

The two classifications have proven complementary. As a general rule, the main distinctions between the regional winds and the local winds are based on both wind speed and direction. The differences between the wind patterns associated with the local orography and valley dynamics are primarily based on wind direction. In the range of very light to light winds, the wind direction classification provides a refinement of the wind patterns into two main categories (night and morning) and three transitional categories. In contrast, in the range of strong to very strong winds, the more detailed description is given by wind speed, which presents only two direction patterns.

These results are consistent with the observations found in the literature, which refer to predominantly day winds from the NE (Culqui, 2001; McVehil, 2005; Clark *et al.*, 2006; C.N.A., 2006, 2007). The dynamics of the diurnal NE wind (7:00 to 18:00) and the nocturnal winds descending through the valleys (18:00–7:00) match the qualitative graphic description presented by Culqui (2001) for the Morococha and La Oroya areas from 1978 to 1980. Culqui (2001) also reported variations in the area wind roses between the diurnal and nocturnal periods. The diurnal results are consistent with the descriptions, primarily an NE wind, although SSW

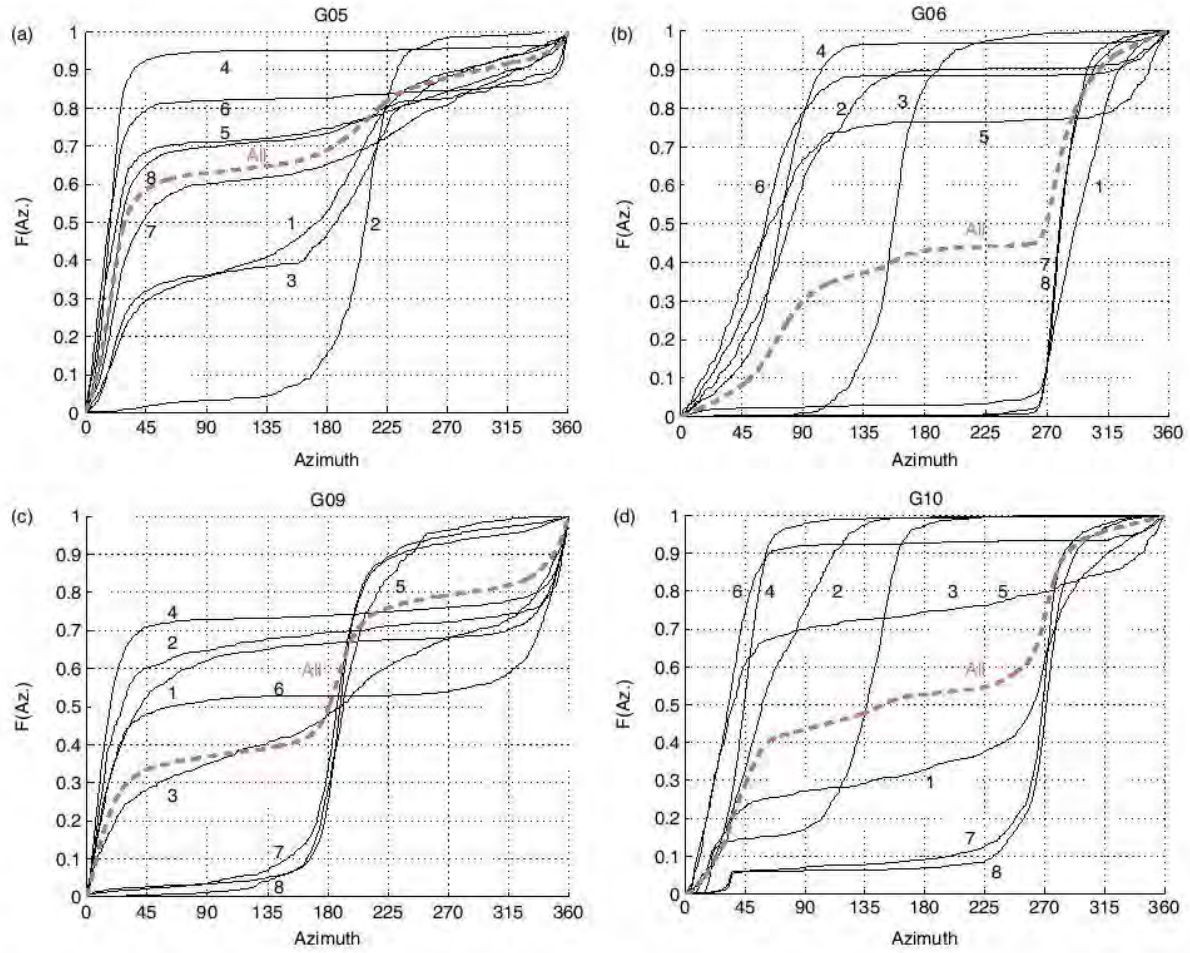

Figure 4. Cumulative distribution function of azimuth (°): Groups WP8D_1 to _8 and stations (a) G05, (b) G06, (c) G09, and (d) G10.

winds are also present. These winds may be linked to the SE wind that was identified in this work. Culqui (2001) presented further average results for Pachachaca for 1940–1980, which showed a preponderance of east winds from 13:00 to 19:00, which is consistent with the position in the valley with the orientation of E–W. No mention was found of south winds in this case. With respect to nocturnal winds, Culqui (2001) presented orientations of NNE–NE and more than 30% calm winds. The NE orientations can be related to the evening wind, which appears in the time slot of 18:00 to 7:00. The dominant nocturnal orientation detected in this work is the west; although this was not the case in Culqui (2001), it may have been masked by calm winds. Although not shown in the figures, the general winds descending in the valleys are present in the qualitative description provided by the same author.

The G06 behaviour coincides with the wind roses presented for the 2000–2005 period with time slots from 24:00 to 8:00, 9:00 to 12:00, 13:00 to 18:00, and 19:00 to 23:00 by Clark *et al.* (2006) using data from McVehil (2005). In the percentages of the total, the nocturnal WNW winds stand out, as detected in this work. The C.N.A. (2006) references the general midday wind and light winds from the south using 2003 data. Finally, it should be noted out that the C.N.A. (2007) presented the predominant orientations for the period 2004–2008 from 5:00 to 7:00 at the stations G01, G03, G04, G05, and G06; the results are consistent with those revealed in this work for the night and dawn winds that descend through the valleys.

4.3. Seasonal winds

Table 6 shows the classification of the WP8D and WP8V contingency tables for the summer (from January to March) and winter cases (from May to August). Note the following two main differences:

- Characteristic summer winds (rarely present in winter) include the SE wind (WP8D_3) and the evening wind (WP8D_5). In general, the day and night winds were lighter than in winter time.
- Characteristic winter winds (rarely present in summer) include the day/very strong wind (WP8V_1) and the night/light (WP8V_5). These were the strongest typologies of winds for both day and night, respectively.

Table 4. Contingency tables between WPnV and WPnD with $n = 4$, 6, and 8.

WP4V	WP4D[a]				Total	
	1	2	3	4		
1			914		953	10%
2			1093		1379	15%
3			599	1600	2874	30%
4	1480	690		1725	4225	45%
Total	1837	1252	2936	3406	9431	
	19%	13%	31%	36%		100%

WP6V	WP6D[b]						Total	
	1	2	3	4	5	6		
1			718				764	8%
2			772	205			1158	12%
3		326	466	342			1402	15%
4						1314	1768	19%
5	1291	647		312	1050	921	4336	46%
6							3	0%
Total	1540	1154	2087	999	1264	2387	9431	
	16%	12%	22%	11%	13%	25%		100%

WP8V	WP8D[c]								Total	
	1	2	3	4	5	6	7	8		
1				645					710	8%
2				632		186			1007	11%
3				396		291			1025	11%
4	289	254	222			197		297	1567	17%
5								884	1057	11%
6	235						625	616	1628	17%
7	844	263	151				468	429	2434	26%
8									3	0%
Total	1481	549	697	1868	397	910	1222	2307	9431	
	16%	6%	7%	20%	4%	10%	13%	24%		100%

Cells show the number of cases. Total number and percentage are given at the end of each row and column. Thresholds for each sub-table are included.
[a]Threshold:590.
[b]Threshold:262.
[c]Threshold:148.

Culqui (2001) presented data from 1978 to 1980 in which the winter and summer seasons were segregated and observed that NE–NNE winds predominated in both cases. The difference is found in the SSW winds, which occur mostly in summer, and their result is consistent with the south wind detected in this work.

4.4. Winter characterization

This analysis focused on the winter months from May to August. Figure 5 shows the averages of the wind velocity vectors for all hours of the day according to the direction classification. The same stations as in Figures 2, 3 and 4 were considered in addition to stations G04 and G08. The marker size is proportional to group size, and the uncertainty associated with each average wind is also depicted. The main groups that stand out are identified in Tables 3 and 4, namely, the day winds from the NE and the night winds. The dynamics of the wind patterns in the different areas are described as follows:

- Central zone (G10): Night wind from W passes to E at low speeds. Directions are consistent with the position in the division between the SE–NW and SW valleys. Strong and very strong NE regional wind subsequently develop. The dynamics of night and morning can be distinguished from the regional winds as well as from the transitions from day to night.
- SW valley area (G07 and G09): Winds were aligned with the valley. In G09, winds were observed from S to N by W (with very low values), and in G07, wind was observed from SW to NE. In G07, the wind ascending in the valley can be distinguished from the NE wind. Minor winds primarily occur oriented according to the side of the valley in which the station was located and descending with the slope of the terrain.
- NW valley area (G05): Descending wind from the NE during the night (WP8D_7 and _8) ascends from the SW (WP8D_2) until the NE day wind develops. Local wind dynamics in the valley overlap with regional wind. The dynamics are consistent with those of the central and SW areas. At night, the wind speed is equal

Table 5. Wind patterns in La Oroya region: main characteristics.

Morning winds (WP8D_1)	Aligned in the valleys Wind speed groups: uniform (WP8V_7), in a lesser extend non-uniform/very light (WP8V_6) and day/very light (WP8V_4) [*] Average speeds in the range of 1 to 2 $m\,s^{-1}$	Predominant, from 8:00 to 10:00	Less presence in winter [*] Thermal inversion, more in winter time
Midday wind (transition) (WP8D_2)	Wind speed groups: uniform and day/very light	Present from 10:00 to 12:00	Less presence in summer
Day winds (WP8D_4, _6)	Regional wind from NE, Wind speed groups: day/light (WP8V_3), day/strong (WP8V_2) and day/very strong (WP8V_1) Speed up to over 12 $m\,s^{-1}$ in G10, and 6 $m\,s^{-1}$ in the other stations	Predominant, from 12:00 to 19:00	Lighter and less presence in summer
Evening wind (transition) (WP8D_5)	Wind speed group: uniform	Present from 19:00 to 21:00	Mostly summer
Night winds (WP8D_8)	Winds from W at most stations Wind speed groups: non-uniform/light (WP8V_5), non-uniform/very light, and uniform Average speeds below 4 $m\,s^{-1}$ in G10, and 3 $m\,s^{-1}$ in the other stations	Predominant, from 21:00 to 4:00, and significant until 7:00	Lighter and less presence in summer
Dawn wind (transition) (WP8D_7)	Wind speed groups: uniform and non-uniform/very light Characteristic behaviour of G08 and G05	Significant, from 4:00 to 8:00	Lighter and less presence in summer Colder than night winds
SE- wind (WP8D_3)	Wind speed uniform and day/very light	Present from 11:00 to 14:00	Mostly summer

or greater to that of the central zone, including the wind from the station at the highest altitude (G10).

- SE valley area (G04 and G08): Winds at station G04 are primarily oriented with the valley, descending from the NW at night and turning to the ENE during the morning. These winds are dispersed in all directions from E to W by the south during the day and evening. The morning transition occurred from the N through the valley slope where the station is located. The G08 station recorded an ascending wind throughout the valley (from the smaller side of the lateral valley in which it is located) at night. During the day and afternoon, the station reported SW winds that were lighter than the nocturnal winds.

The G04 station reported a rather high variability during the day, when the regional wind developed, with winds flowing in a direction opposite to the regional one. Given the exact location of the station and the origin of regional winds, such variable values could be due to an area of local turbulence. During the night and the morning, the winds follow the orientations of the valley.

In the SE end station, G08, the wind displayed an ascending behaviour with relatively high speed through the tributary valley that hosts the station (aligned in the E–W direction) during the nocturnal period, in agreement with the data from G05. One diurnal descending behaviour was recorded from the W–SW, which was lighter than the nocturnal wind; therefore, it is strongly influenced by local conditions. For the dawn winds (WP8D_7), those that are lighter and from the W stand out. In G08, the wind exhibits a main difference from the night wind (WP8D_8) throughout all stations. The presence of the dawn winds indicates the beginning of the morning transition, with a variation in the wind orientation to descending in the valley. Finally, the day winds WP8D_6, which appear in the evening, show an ascending orientation throughout the valley in G08, similar to the night wind; this is the main difference between the WP8D_4 and _6 groups.

Thus, the dynamics of the local winds in the valleys have been identified and are well developed in central zone and the SW and NW valleys. These winds are subsequently overlapped with the regional NE winds (notably, at G07 and G10; see Figure 5). Local influences invalidated the use of G04 and G08 in this later case. Four main behaviours can be distinguished over a typical winter day, and these occur in succession: night wind, descending valleys; morning wind, ascending valleys; NE regional day wind; and evening wind, which re-orients to the descending winds of the valleys. A morning wind that ascends in the valleys previous to the NE regional wind

Table 6. Contingency tables between WP8V and WP8D.

WP8V	WP8D[a]								Total	
	1	2	3	4	5	6	7	8		
1				28					37	3%
2				71		41			163	11%
3			31	59		53			184	13%
4	39		54						195	13%
5								26	80	5%
6	68		39				45	78	248	17%
7	197	44	62		48	31	87	74	557	38%
Total	338	69	226	194	106	169	151	211	1464	
	23%	5%	15%	13%	7%	12%	10%	14%		100%
WP8V	WP8D[b]								Total	
	1	2	3	4	5	6	7	8		
1				349					368	9%
2				313		75			411	10%
3				193		143			421	11%
4	134	135				105		139	708	18%
5								576	617	16%
6							346	231	646	16%
7	258	118					163	146	781	20%
Total	443	263	127	925	109	412	561	1112	3952	
	11%	7%	3%	23%	3%	10%	14%	28%		100%

Cells show the number of cases. Total number and percentage are given at the end of each row and column. Thresholds for each sub-table are included. Upper table shows summer data (January to March); bottom table shows winter data (May to August).
[a]Threshold: 23.
[b]Threshold: 62.

has not been explicitly referenced in either Culqui (2001) or Clark *et al.* (2006), but its orientation matches with the NE wind in the stations of the central zone and the SW valley, and its presence is identified in the NW valley of G05 by morning and midday winds WP8D_1 and _2.

4.5. Morning transition

This section provides a detailed discussion of the period from 6:00 to 11:00 during winter days. Table 7 presents the contingency table between classifications based on wind speed and direction. The wind patterns in this segment were concentrated in only 12 of the 64 options, and therefore, the subgroups were reclassified to facilitate analysis. Table 7 indicates the boxes that changed classification when they were regrouped: WP_ij, where $i = 1$ (day/light), 5 (night/non-uniform), or 7 (night/uniform), and $j = 1$ (morning), and 2 (midday), 4 (d), 7 (dawn), or 8 (night). Table 8 shows the hourly distribution of classifications WP8D, WP8V, and the new WP. Four major wind patterns are evident: WP_58, _57, _71, and _72, with speeds of night/non-uniform (_5j) and uniform (_7j), and directions of night, dawn, morning, and midday, respectively. The period ends with WP_11 and _12 (the latter changing to _14), all with speeds of day/very light increasing, and directions all aligned with the NE regional wind.

The transitions between the WP_*ij* groups are characterized with a contingency table between the groups presented in each case (rows) and the group for the next hour (columns), as presented in Table 9. The rows show the number of beginnings (mainly WP_71, _57, _58, and _72), and columns indicate the destinations (mainly WP_71, _12, _72, and _57). The WP_58 group (with a night direction and non-uniform speed) is mainly present at the beginning of the time interval, and WP_12 (midday direction and light speed) is present at the end of the time interval. The WP_14 group (day direction and light speed) appears at 12:00. The following transitions stand out: day/light cases (WP_1j, leaving free the direction j) remained in their category; night/non-uniform cases (WP_5j) evolved to either WP_71 (night/uniform–morning) and to a lesser extent, to WP_77 (night/uniform–dawn); and night/uniform cases (WP_7j) moved from WP_12 (day/light–midday), and to a lesser extent, to WP_11 (day/light–morning), always from WP_71 and WP_72 (night/uniform–morning and midday).

The morning transition follows a wind speed sequence of night/non-uniform, night/uniform, and day/very light, with the wind directions of night, dawn, morning, midday, and day. Most of the periods begin with a night wind (WP_*j*8) that turns into a dawn wind (WP_*j*7), which is the strongest from 7:00 to 8:00 and is characterized by the wind in G08 moving from ascending to descending orientation in its tributary valley. These winds subsequently turn into morning winds (WP_*j*1), which predominate from 9:00 to 10:00, with very light winds in G09 and with transitional directions in G10 and G06, also with very light winds (slope winds). The transition occurs primarily in the WP_71 group with a wind

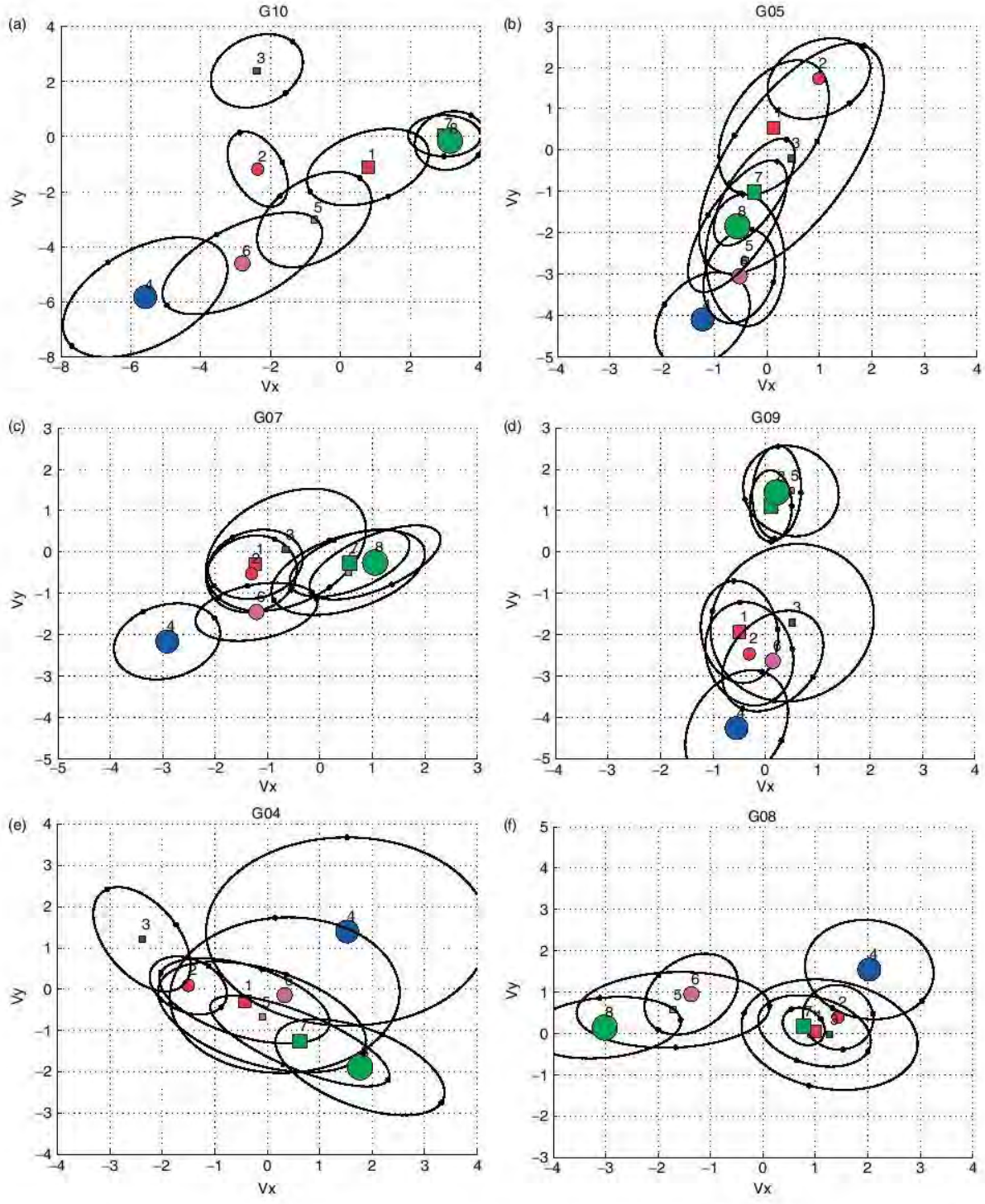

Figure 5. Averages of wind velocity vectors and uncertainties for full days during winter: Groups WP8D_1 to _8, and stations (a) G04, (b) G05, (c) G07, (d) G08, (e) G09, and (f) G10 (marker size is proportional to group size).

speed of night/uniform. The midday wind and the NE regional wind follow the morning wind.

Figure 6 shows the averages of the wind velocity vectors for the WP_*ij* groups, where $ij = \{11, 12, 57, 58, 71, 72, 77, 78\}$, together with their uncertainties for the period from 6:00 to 11:00 and for the same stations as in Figure 5. The passage of night wind from a west to NE regional wind was identified in G10, including cases with wind from the east before regional winds developed. The wind descends at night through the intersection of the valleys towards the SW valley, resulting in wind from the west in G10; this wind subsequently ascends through the valley, resulting in an east wind in G10, and finally turns to a NE orientation. At the G05 and G09

Table 7. Contingency table between WP8V and WP8D.

	WP8D									
WP8V	1	2	3	4	5	6	7	8	Total	
1									0	0%
2									1	0%
3									23	2%
4	28	59						31	174	18%
5								65	65	7%
6	260						3	56	251	26%
7	3	104					79	48	463	47%
Total	298	165	28	22	10	9	263	182	977	
	31%	17%	3%	2%	1%	1%	27%	19%		100%

Threshold:15

Cells show the number of cases. Total number and percentage are shown at the end of each row and column. Threshold is included. Data from May to August, from 6:00 to 11:00. Re-codification WP is indicated in boldface and is underlined.

stations, the wind direction turned from descending to ascending according to the valley, as already indicated. In stations G07 and G04, the morning transition origin is identified in the WP_58 group and ends in the WP_12 group, also with descending to ascending wind passage in the associated valleys. The G08 station behaviour corresponds to that described in the previous section.

Figure 7 shows the same results as in Figure 6 but for stations G01, G03, and G06 in the central zone, and the above analysis carried out in the more remote stations is confirmed. The G10, G01, G03, and G06 stations follow a similar behaviour, each with its specific local influences.

4.6. Thermal inversion

Thermal inversion situations can be characterized by the difference in the temperatures observed at the G10 and G06 stations. The two stations are located at approximately the same position except that G10 is situated at an altitude 215 m higher than that of G06. Positive values indicate that the lowest temperature occurred in the valley bottom. Figure 8 shows the relationships of the temperature difference (y-axis) with the G10 temperature and with the G06 solar radiation (x-axis) from 6:00 to 11:00 for the winter period and for all samples (averages and uncertainties of each variable and for each wind pattern are indicated).

The temperatures in the G10 transition from cold situations at night (WP8D_7 and _8) to the more temperate situations during the day (WP8D_2, _3, and _4) via the morning wind group (WP8D_1), and the warm day conditions transition to cold nights through the WP8D_6 and WP8D_5 groups. During the morning transition, the difference in temperatures is between -2 and $+1\,^{\circ}$C. The morning wind is characterized by the increased presence of thermal inversions, and specifically, with wind speeds of day/light and night/uniform (in the WP_11 and WP_71 groups, especially the first, which corresponds to higher temperatures in G10), both of which are affected by a high level of solar radiation. At night, the temperature difference ranges from $-1.5\,^{\circ}$C to $-2\,^{\circ}$C, coinciding with the adiabatic lapse rate, and during the day, it descends from higher (and even positive) values to $-4\,^{\circ}$C. The evening conditions proceed through intermediate temperatures and thermal differences. The thermal inversion phenomenon in the area has been described in the literature as occurring more frequently in winter and developing from 19:00 to 10:00 with highs during the dawn period (Culqui, 2001; C.N.A., 2006). The results obtained in the central zone classify the stronger thermal inversion periods with the morning winds of WP_11 and WP_71 occurring from 9:00 to 10:00.

Both classifications of wind data (by speed and by direction) are useful for characterizing these wind patterns. Neither of them alone segregates the winds for large intensity thermal inversions. Comparing the central and bottom sub-figures in Figure 8, all samples (including the upper ones) show this observation for mornings in winter. Remarkably, the seasonal variations do not help us to better identify these cases because both have an equivalent presence in Table 6 of 2.7 and 3.4% (39 and 134 cases, respectively) in summer and winter for the subgroups WP8V_4 and WP8D_1 recoded within WP_11. However, eight clusters were necessary to separate the WP8V_4 and WP8V_7 speeds (Table 4), and the WP8D_1 and WP8D_7 directions are separated in the classification with the four directions of WP4D.

In Figure 8, the dawn wind with a speed of night/non-uniform, WP_57, is distinguished by grouping the cases with lower G10 temperatures. The G08 data present a singular behaviour for these cases, which are predominant until 8:00 but are significantly present starting from 4:00. Cold dawn winds evolve to speeds of night/uniform WP_77 and are subsequently transformed to morning winds that are also uniform WP_71 and finally to a day/very light speed profile WP_11 with stronger thermal inversions or to midday WP_72 and WP_12. This identification is consistent with that given in the transition table in Table 9. The morning wind WP_71 is the same as that identified previously as the slope wind, or the transition between descending to ascending valley winds, and it no longer shows special low temperature characteristics. Finally, the SE wind stands out in the group due

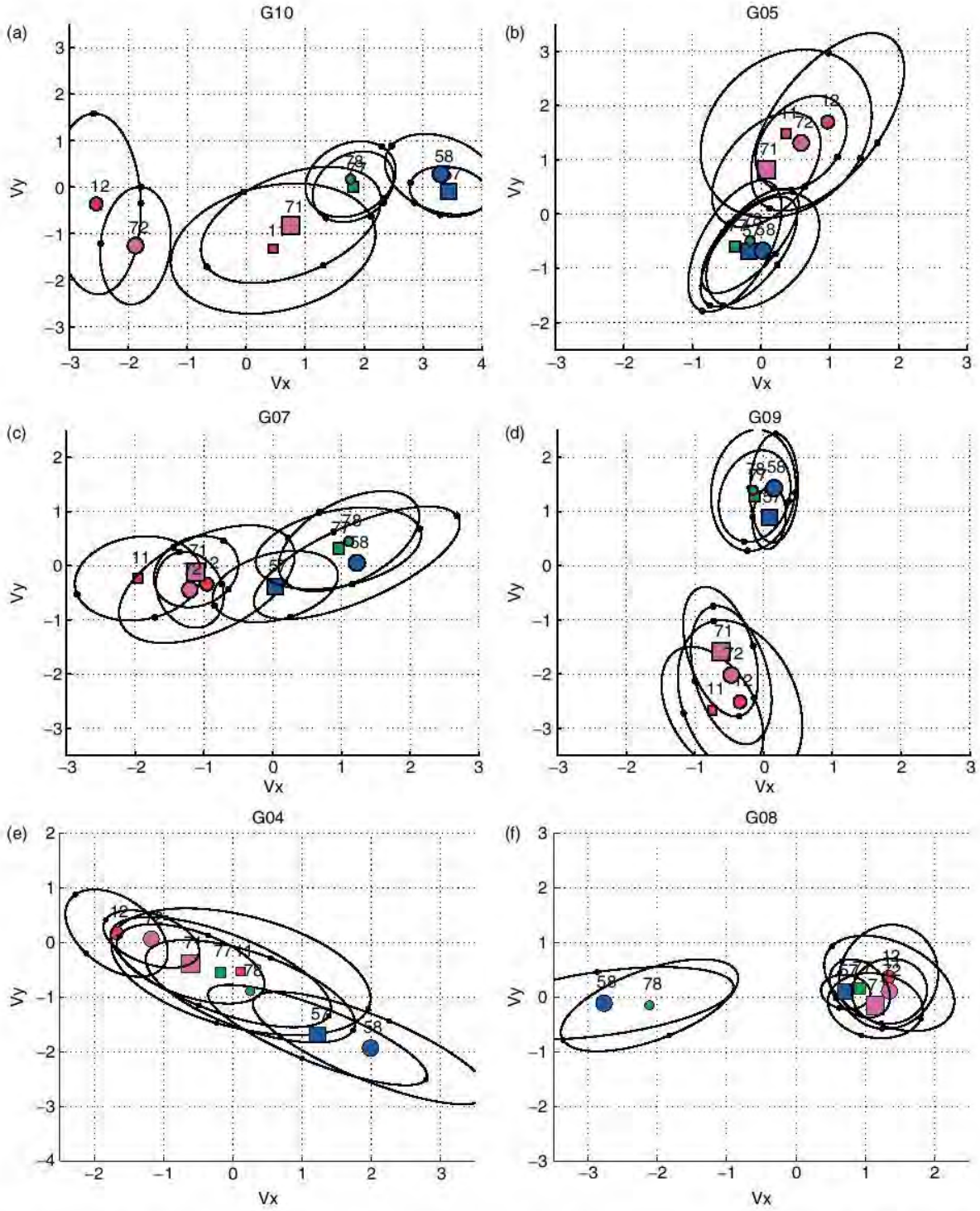

Figure 6. Averages of wind velocity vectors and uncertainties from 6:00 to 11:00 during winter: groups WP_11, _12, _57, _58, _71, _72, _77, and _78 and stations (a) G04, (b) G05, (c) G07, (d) G08, (e) G09, and (f) G10 (marker size is proportional to group size).

to its milder temperatures, which are consistent with its identification as a group characteristic for summer.

Both classifications of wind data are also useful for characterizing cold wind patterns, although both separately segregate the winds for cold dawn wind in this case. A comparison of the central and bottom sub-figures with the upper sub-figures in Figure 8 illustrates this observation. This comparison also shows that speed clustering is more helpful in this case given that it provides a more specific group (comparison of WP8D_7 and WP8V_6). Remarkably, seasonal variations aid in identifying these cases: 3.1 and 8.8% in summer and winter for subgroups WP8V_6 and WP8D_7 (recoded within WP_57) in Table 6. However, similar to the

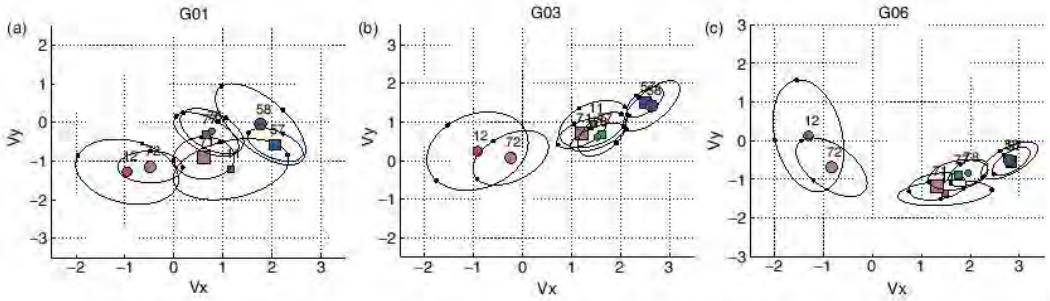

Figure 7. Averages of wind velocity vectors and uncertainties from 6:00 to 11:00 during winter: groups WP_11, _12, _57, _58, _71, _72, _77, and _78 and stations (a) G01, (b) G03, and (c) G06 (marker size is proportional to group size).

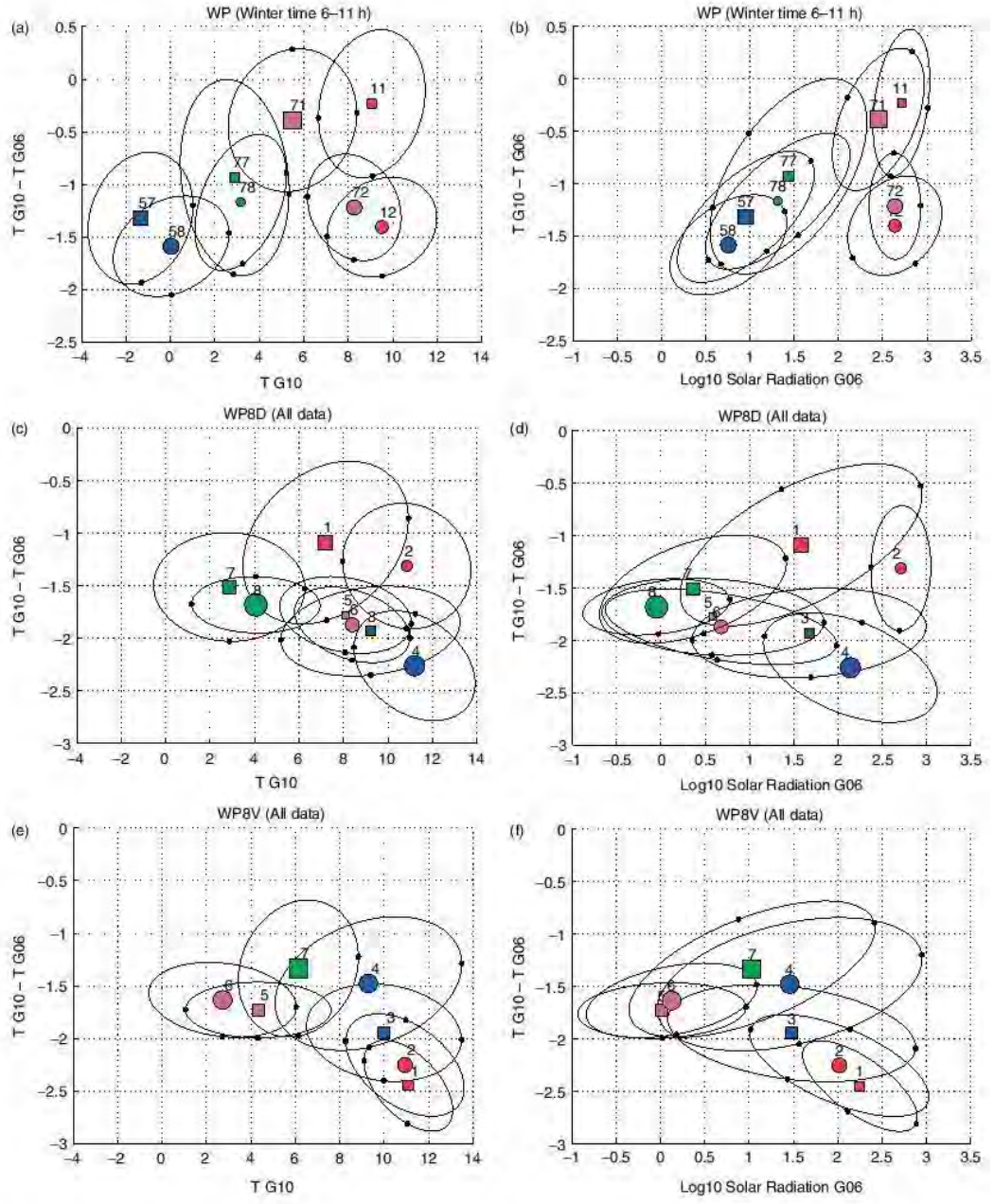

Figure 8. Relationships between differences of temperature in G10–G06 (°C) and temperature in G10 (°C), and the decimal logarithms of the solar radiation in G06 (w*m^2). Winter data from 6:00 to 11:00 (for groups WP_11, _12, _57, _58, _71, _72, _77 and _78), and data from all hours of all time periods for groups WP8D_1 to _8 and WP8V_1 to _7, (marker size is proportional to group size).

Table 8. Heat maps displaying the hourly distribution of WP8V, WP8D, and WP classifications.

N Cases	162	161	161	163	164	166
	Hours					
WP8V	6	7	8	9	10	11
1						
2						
3	3%	4%				3%
4	6%	9%	6%	9%	23%	53%
5	14%	14%	13%			
6	56%	52%	45%			
7	22%	22%	34%	88%	75%	43%
WP8D	6	7	8	9	10	11
1	7%	5%	14%	68%	60%	28%
2				6%	32%	61%
3						5%
4				3%	4%	3%
5						
6						
7	49%	45%	55%	13%		
8	38%	43%	24%	6%		
WP	6	7	8	9	10	11
11					14%	17%
12					6%	36%
14		3%			4%	4%
18	3%	6%				
51			6%			
57	36%	33%	36%			
58	29%	30%	16%			
71	4%	4%	8%	64%	45%	11%
72				7%	29%	30%
77	11%	10%	16%	11%		
78	6%	6%	7%	4%		

Data from May to August, from 6:00 to 11:00. Percentage w.r.t. daily cases indicated in the cells (for values >2%).

thermal inversion analysis, eight clusters were needed separate the speeds WP8V_6 and WP8V_7 (Table 4) and separate the directions WP8D_7 and WP8D_8 in the classification with six directions of WP6D.

The characterization of cold winds and thermal inversion using wind patterns can be complemented with an analysis of the wind evolution at each station. Wind patterns WP_57 (direction night, speed non-uniform/very light) and WP_11 (direction morning, speed day/very light) stand out at selected stations (G01, G07, G08) in Figures 6 and 7.

5. Conclusions

This article presents a methodology used to characterize the wind patterns in complex mountainous regions. The main objective is to classify data by wind speeds and wind directions and to subsequently compare the classifications. By complementing each other, these comparisons allow identification of wind patterns that are conditioned by both groups. CAs are separately performed for the hourly data of wind speed and the unit vectors of wind direction, and the aggregations are compared using contingency tables. The generated groups are characterized using the two aggregation schemes and are identified as wind patterns. Segmentation of the classifications by season allows the characteristic wind patterns for each case to be identified, whereas analysis of the hourly distribution allows the main wind patterns throughout the day to be identified. The interpretation of the transitions is facilitated by pattern visualization using the average values for wind and their uncertainties at each station.

The statistical methods included in the proposed method (i.e. CA, PCA, and contingency tables) are common in standard practice and provide easily interpretable results. This approach is robust, accurate, and easy to apply and interpret, and if necessary, it can be combined with other aggregation or data reduction techniques.

The proposed method has been applied in data analyses for nine stations in the area of La Oroya in Peru, an area with complex patterns of wind dynamics. The model has

Table 9. Contingency table between WP at time 'n' and time '$n+1$'.

WP "n"	WP "$n+1$"[a]									Total
	11	12	14	57	58	71	72	77	78	
11	**30**	*15*	5							56
12	6	**56**	*12*							77
14			**18**							28
57				**71**	16	*50*		23		170
58				*31*	**46**	12		5	10	121
71	28	23	6			**88**	*65*			222
72	12	*47*	8		9		**36**			112
77				6		*25*	10	**19**		79
78						*14*			**7**	38
Total	93	150	57	111	74	221	124	61	28	977

Cells show the number of cases. Total number and percentage are given at the end of each row and column. Threshold is included. Data from May to August, from 6:00 to 11:00. The diagonal is shown in bold, and the highest off-diagonal value is shown in italics.
[a]Threshold: 4.

been applied to a total of 9431 h of observation over periods of at least one consecutive day between June 2007 and October 2008. The methodology provides a classification that includes the main local wind pattern description. Indications in the literature with respect to the prevailing winds in the area were identified. Importantly, the analysis of wind transitions during the morning periods has contributed to identification of weather conditions that are more favourable to high immission values. Furthermore, the characteristics of wind, speed, and direction describe the thermal inversion winds in the valleys.

Acknowledgements

This research was conducted within the framework of projects CGL2008-06003-C03-02 and CLG2011-29396-C03-00 from the Spanish Ministry of Education and Science. Altac S.L. provided data to UPC in December 2008 under contract N° C07585. The author thanks the anonymous referees for useful comments.

References

Aikawa M, Hiraki T, Eiho T. 2008. Grouping and representativeness of monitoring stations based on wind speed and wind direction data in urban areas of Japan. *Environmental Monitoring and Assessment* **136**: 411–418. DOI: 10.1007/s10661-007-9696-0

Beaver S, Palazoglu A. 2006. Cluster analysis of hourly wind measurements to reveal synoptic regimes affecting air quality. *Journal of Applied Meteorology and Climatology* **45**: 1710–1726. DOI: 10.1175/JAM2437.1

Burlando M, Antonelli M, Ratto CF. 2008. Mesoscale wind climate analysis: identification of anemological regions and wind regimes. *International Journal of Climatology* **28**: 629–641. DOI: 10.1002/joc.1561

Carn SA, Krueger AJ, Krotkov NA, Yang K, Levelt PF. 2007. Sulfur dioxide emissions from Peruvian copper smelters detected by the Ozone Monitoring Instrument. *Geophysical Research Letters* **34**: L09801. DOI: 10.1029/2006GL029020

Clark, S., Partelpoeg, E., Young, J.W.S. (2006), *Expert Comments on Exceptional Fulfillment Extension Request for the Sulfuric Acid Plant Project of La Oroya Metallurgical Complex PAMA*, Ministerio de Energía y Minas, República del Perú, May 10.

C.N.A. (2006). *Plan de Acción para la Mejora de la Calidad del Aire en la Cuenca Atmosférica de La Oroya*, Decreto del Consejo Directivo, Consejo Nacional del Ambiente, República del Perú, N° 020-2006-CONAM/CD and 026-2006-CONAM/CD, August 2 and October 10.

C.N.A. (2007). *Plan de Contingencia Episodios de Emergencia para estados de alerta por contaminación del aire en la cuenca atmosférica de La Oroya*, Decreto del Consejo Directivo, Consejo Nacional del Ambiente, República del Perú, No 015-2007-CONAM/CD, July 18.

Culqui, E. (2001) Estudio de la contaminación potencial del aire en la zona de Morococha y La Oroya. *Anales Cientrficos UNALM*, Universidad Nacional Agraria La Molina: Lima, Perú, 217–237.

Crutcher HL, Baer L. 1962. Computations from elliptical wind distribution statistics. *Journal of Applied Meteorology* **1**: 522–530. DOI: 10.1175/1520-0450(1962)001<0522:CFEWDS>2.0.CO;2

D.E.&B.O.C. (2007). Examen Especial al Complejo Metalúrgico La Oroya. Desarrollo y Ecología S.A.C. and B. O. Consulting S.A. for Programa de Adecuación y Manejo Ambiental del Complejo Metalúrgico de La Oroya. Ministerio de Energía y Minas, República del Perú, No 002-2007-MEM-DGM/V, January 3.

Darby LS. 2005. Cluster analysis of surface winds in Houston, Texas, and the impact of wind patterns on ozone. *Journal of Applied Meteorology* **44**: 1788–1806. DOI: 10.1175/JAM2320.1

Davies DL, Bouldin DW. 1979. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2**: 224–227. DOI: 10.1109/TPAMI.1979.4766909

Drobinski P, Dabas AM, Haeberli C, Flamant PH. 2003. Statistical characterization of the flow structure in the Rhine valley. *Boundary-Layer Meteorology* **106**: 483–505. DOI: 10.1023/A:1021262321679

Doran JC, Horst TW, Whiteman CD. 1990. The development and structure of nocturnal slope winds in a simple valley. *Boundary-Layer Meteorology* **52**(1–2): 41–68. DOI: 10.1007/BF00123177

Esteban P, Martin-Vide J, Mases M. 2006. Daily atmospheric circulation catalogue for western Europe using multivariate techniques. *International Journal of Climatology* **26**: 1501–1515. DOI: 10.1002/joc.1391

Ferragut L, Montenegro R, Montero G, Rodríguez E, Asensio MI, Escobar JM. 2010. Comparison between 2.5-D and 3-D realistic models for wind field adjustment. *Journal of Wind Engineering and Industrial Aerodynamics* **98**(10–11): 548–558. DOI: 10.1016/j.jweia.2010.04.004

Gómez-Muñoz VM, Porta-Gándara MA. 2002. Local wind patterns for modeling renewable energy systems by means of cluster analysis techniques. *Renewable Energy* **25**: 171–182. DOI: 10.1016/S0960-1481(01)00013-1

Guardans R, Palomino I. 1995. Description of wind filed dynamic patterns in a valley and their relation to mesoscale and synoptic-scale meteorological situations. *Journal of Applied Meteorology* **34**: 49–67. DOI: 10.1175/1520-0450-34.1.49

Hardy DM, Walton JJ. 1978. Principal components analysis of vector wind measurements. *Journal of Applied Meteorology* **17**: 1153–1162. DOI: 10.1175/1520-0450(1978)017<1153:PCAOVW>2.0.CO;2

Jiménez PA, González-Rouco JF, Montávez JP, García-Bustamante E, Navarro J. 2009. Climatology of wind patterns in the northeast of the Iberian Peninsula. *International Journal of Climatology* **29**: 501–525. DOI: 10.1002/joc.1705

Jiménez PA, González-Rouco JF, Montávez JP, Navarro J, García-Bustamante E, Valero F. 2008. Surface wind regionalization in complex terrain. *Journal of Applied Meteorology and Climatology* **47**: 308–325. DOI: 10.1175/2007JAMC1483.1

Kalkstein LS, Tan G, Skindlov JA. 1987. An evaluation of three clustering procedures for use in synoptic climatological classification. *Journal of Applied Meteorology* **26**: 717–730. DOI: 10.1175/1520-0450(1987)026<0717:AEOTCP>2.0.CO;2

Kastendeuch PP, Kaufmann P. 1997. Classification of summer wind fields over complex terrain. *International Journal of Climatology* **17**: 521–534. DOI: 10.1002/(SICI)1097-0088(199704)17:5<521::AID-JOC143>3.0.CO;2-Q

Kaufmann P, Weber RO. 1996. Classification of mesoscale wind fields in the MISTRAL field experiment. *Journal of Applied Meteorology* **35**: 1963–1979. DOI: 10.1175/1520-0450(1996)035<1963:COMWFI>2.0.CO;2

Kaufmann P, Weber RO. 1998. Directional correlation coefficient for channeled flow and application to wind data over complex terrain. *Journal of Atmospheric and Oceanic Technology* **15**: 89–97. DOI: 10.1175/1520-0426(1998)015<0089:DCCFCF>2.0.CO;2

Kaufmann P, Whiteman CD. 1999. Cluster-Analysis Classification of Wintertime Wind Patterns in the Grand Canyon Region. *Journal of Applied Meteorology* **38**: 1131–1147. DOI: 10.1175/1520-0450(1999)038<1131:CACOWW>2.0.CO;2

Khokhar MF, Platt U, Wagner T. 2008. Temporal trends of anthropogenic SO2 emitted by non-ferrous metal smelters in Peru and Russia estimated from Satellite observations. *Atmospheric Chemistry and Physics Discussions* **8**: 17393–17422. DOI: 10.5194/acpd-8-17393-2008

Ludwig FL, Horel J, Whiteman CD. 2004. Using EOF analysis to identify important surface wind patterns in mountain valleys. *Journal of Applied Meteorology* **43**: 969–983. DOI: 10.1175/1520-0450(2004)043<0969:UEATII>2.0.CO;2

McVehil GE. (2005) *Air quality dispersion modeling for human health risk assessment. La Oroya metallurgical complex*, McVehil-Monnett Associates, Inc., for Integral Consulting, Inc. MMA Project Number 1875-0.

Muñoz-Díaz D, Rodrigo FS. 2004. Spatio-temporal patterns of seasonal rainfall in Spain (1912–2000) using cluster and principal component analysis: comparison. *Annales Geophysicae* **22**: 1435–1448. DOI: 10.5194/angeo-22-1435-2004

Ratto GE, Maronna R, Berri G. 2010a. Analysis of wind roses using hierarchical cluster and multidimensional scaling analysis at La Plata, Argentina. *Boundary-Layer Meteorology* **137**: 477–492.

Ratto GE, Videla F, Maronna R, Flores A, Dávila FP. 2010b. Air pollutant transport analysis based on hourly winds in the city of

La Plata and surroundings, Argentina. *Water, Air, and Soil Pollution* **208**: 243–257. DOI: 10.1007/s11270-009-0163-0

Ray S, Turi RH. (1999). Determination of number of clusters in K-means clustering and application in colour image segmentation. In Proceedings of the 4th International Conference on Advances in Pattern Recognition and Digital Techniques (ICAPRDT'99), Pal NR, De AK, Das J (eds). Calcutta, India.

Stewart JQ, Whiteman CD, Steenburgh WJ, Bian X. 2002. A climatological study of thermally driven wind systems of the U.S. Intermountain West. *Bulletin of the American Meteorological Society* **83**: 699–708.

Whiteman CD, Doran JC. 1993. The relationship between overlying synoptic-scale flows and winds within a valley. *Journal of Applied Meteorology* **32**: 1669–1682. DOI: 10.1175/1520-0450(1993)032<1669:TRBOSS>2.0.CO;2

Zarauz J. (2011) *Sulfur Dioxide Estimations in the Planetary Boundary Layer Using Dispersion Models and Satellite Retrievals*, PhD thesis, Saint Louis University, ProQuest LLC. ISBN: 9781124774817.

Zarauz J, Pasken R. (2010). WRF model validation for complex terrain and integration with pollutant dispersion models, *11th Annual WRF Users' Workshop*, National Center for Atmospheric Research, June 21–25, Boulder.