Autocorrelation Analysis of Meteorological Data from a RASS Sodar

ISIDRO A. PÉREZ, M. ÁNGELES GARCÍA, M. LUISA SÁNCHEZ, AND BEATRIZ DE TORRE

Department of Applied Physics, Faculty of Sciences, University of Valladolid, Valladolid, Spain

(Manuscript received 12 December 2003, in final form 25 March 2004)

ABSTRACT

Autocorrelation analysis is necessary in persistence studies and identification of cyclical processes. In this paper, autocorrelations of available wind speed and temperature data from a radio acoustic sounding system (RASS) sodar were calculated. This device was placed on flat terrain, and the measuring campaign extended over April 2001. Ten-minute averages were considered from 40 to 500 m in 20-m levels. The direction frequency rose indicated clear, prevailing directions in the east-northeast–west-southwest axis. Analysis of median temperatures revealed that east-northeast advections were 5°C colder than those from the west-southwest. A defined pattern was obtained for both autocorrelations, comprising deterministic and random parts. Noise became more relevant at the higher levels. The deterministic part could be considered as an initial fast-decaying term with the addition of two harmonic functions. The initial decay, linked to fast changes, increased with height for wind speed and decreased for temperature. A diurnal cycle was relevant at intermediate levels for wind speed and at lower temperature levels. The absence of the surface influence added to the horizontal movement associated with the stable night stratification and diurnal convection produced a sharp daily contrast in wind speed at intermediate levels. The influence of the surface decreased with height for temperature. The second cycle was linked to changes in the synoptic pattern and had a 5–6-day period. It was more relevant at lower levels for wind speed, and its amplitude decreased with height. For temperature, this second cycle was less significant. Following these assumptions, a model for the autocorrelation function was proposed and its coefficients are calculated by means of a simple method—a multiple linear regression beyond the first day and a simple linear regression for the first 12-h residuals. This model proved satisfactory, especially below 300 m. A rough height parameterization has been proposed that retained the relevant information and provided satisfactory fits. The influence of the number of observations was investigated. Only extremely high data reduction provided noticeable noise.

1. Introduction

The autocorrelation function is a simple and effective tool for data analysis. It may be observed as a depiction of persistence in the time series. In this sense, it provides information on data memory. For instance, a fast decrease is indicative of no correlated measurements. However, it may also be used to reveal periodicities in the data (Stull 1988). Knowledge of the autocorrelation function also aids in selecting time series models, such as autoregressive moving average (ARMA; Rehman and Halawani 1994), autoregressive integrated average model (ARIMA), and so on (Box and Jenkins 1970), and allows for the correct application of statistical methods in data analysis by aiding in the determination of whether the data are independent (Wilks 1995). Moreover, this function may inform as to the nature of noise when a variable may be considered as an addition of a group of frequencies. However, it has scarcely been used in the atmospheric sciences, perhaps because it sometimes fails to exhibit a defined pattern, which can prove discouraging.

Several studies have been performed from a meteorological point of view. Brett and Tuller (1991) calculated the autocorrelation function for a very extensive wind speed database at seven stations on the west coast of Canada. The strong influence of the surrounding topography was noticeable in this case: the more uniform the surrounding topography was, the higher the short-lag autocorrelations were. The values were successfully fitted by means of a model with a modified exponential term and cosine terms, which corresponded to strictly defined cyclic evolutions. Some of them appeared clearly (diurnal and annual), and others were linked to the stations (semidiurnal, seasonal, and semiannual). The autocorrelation function generally did not decay to 0, even at 2 months. Daoud et al. (2003) calculated autocorrelations corresponding to synoptic-scale flow from a broad trajectory database. The wavelike behavior for the zonal and meridional wind velocity components was attributed to the synoptic-scale eddies. Additionally, these functions were satisfactorily modeled using

Corresponding author address: Isidro A. Pérez, Department of Applied Physics, Faculty of Sciences, University of Valladolid, c/ Prado de la Magdalena s/n, 47071 Valladolid, Spain.
E-mail: iaperez@fa1.uva.es

Gaussian and second-order autoregressive autocorrelation models.

Having reviewed the basic uses of the autocorrelation function it is also useful to give a brief outline of some specific applications.

In turbulence analyses, the autocorrelation function is integrated until it reaches 0 to obtain the integral time (Kaimal and Finningan 1994), which is closely related with dispersion coefficients. In this case, the most widely used parameterization is the exponential function (Arya 1999). However, observations have shown that this did not seem to be a good approach under low-wind situations (Oettl et al. 2001).

Another possible application is for wind energy studies, where as precise as possible knowledge of wind stability is the major interest. Recently, Tomson and Hansen (2001) calculated wind autocorrelations on the Estonian coast as well as on the mainland, and they also investigated the influence of height within the range that is useful from the point of view of power engineering. Later, Pryor and Barthelmie (2002) considered autocorrelations with data from the Danish wind-monitoring network for three stations in near-shore and offshore environments. Only the land data showed clear evidence of a diurnal cycle, whereas in the coastal zone this cycle was not evident because of the thermal lag introduced by the water body in the absence of advection from land.

A third field of use for this function is in air pollution research, where the ability of autocorrelation to highlight data memory has also been successfully proven. Brunciak et al. (2001) presented autocorrelations for several meteorological variables and pollutants at two different sites: Baltimore, Maryland, and over the Chesapeake Bay. In this case the contrast between both places was clearly revealed. Hauck et al. (1999) have shown strong differences in autocorrelations for several pollutants in Linz, Austria, from different sources: traffic, industry, domestic heating, and photochemical origin. These latter examples, together with the previous ones, are encouraging signs for future use of this function, not only in basic but also applied research.

The purpose of this research was to analyze the autocorrelation function for data from a radio acoustic sounding system (RASS) sodar. One exclusive advantage of this device is that it can obtain vertical information of the low atmosphere. In particular, RASS sodar may be used to calculate the autocorrelation function of a time series for one variable and to analyze its vertical behavior with very detailed spatial resolution.

Beside wind speed, temperature was also investigated. Autocorrelations for both of these variables were calculated and compared. Several explanations are proposed in order to link this function to boundary layer profiles and their time evolution.

A simple model was built with the purpose of checking our assumptions, and a rough parameterization is suggested to eliminate height dependence while maintaining essential information.

Last, the influence of the number of observations on the autocorrelation noise is investigated, and a dependence relationship is proposed.

2. Experimental description

The equipment used in this paper is a DSDPA.90-24 sodar equipped with a RASS extension and built by Meteorologische Messtechnik (METEK) GmbH. It was installed at the Centro de Investigación de la Baja Atmósfera (CIBA), 41°49′2″N, 4°56′15″W, about 30 km northwest of Valladolid in northern Spain. The location is on a very extensive plateau at 840 m MSL, which does not present relief elements, thus, guaranteeing horizontal homogeneity. For this reason, orographic influences on the temporal evolution of the boundary layer vertical structure are excluded. Nonirrigated crops and a group of low trees between the west and southwest directions make up the surrounding vegetation.

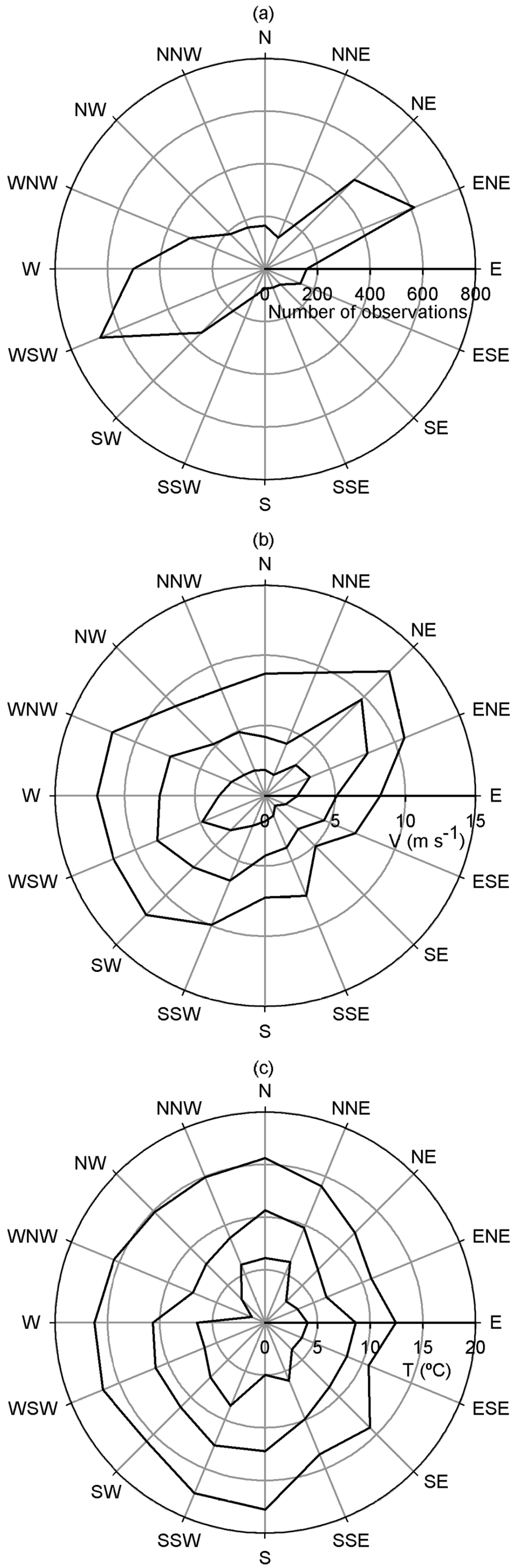
The measuring period considered in this paper was April 2001. This was a short enough interval to avoid annual trends and long enough to provide a representative number of data. The equipment worked correctly and interruptions were below 0.8% and were due to external reasons. Ten-minute averages were considered. The minimum height from which data were obtained was 40 m and the maximum was 500 m. A 20-m vertical resolution was used.

The data files yielded plenty of information, although in the present analysis only wind speed and temperature were considered. The data were subjected to quality control that consisted of rejecting those data whose signal-to-noise ratio was below -3 dB.

This database has been used by the authors in previous papers (Pérez et al. 2003a,b). The availability of data was significantly influenced by height. At 500 m, it was slightly below 4% of observations for the sodar and above 20% for the RASS. Additionally, a diurnal cycle was detected. For the sodar, minimum in availability was observed at 1800 UTC from the 140-m level that expanded to the morning hours at higher levels. The behavior of data availability was opposite for the RASS, with a zone of maximum availability clearly observed between 0800 and 1700 UTC between a 280- to 440-m height.

Vertical profiles were also studied. The boundary layer presented a well-defined structure with a daily cycle. The median wind speed profile in the lowest 200 m was an exponential profile during the night. During the day, strong convection was responsible for nearly flat profiles. However, the median temperature profile was linear during the day and early hours of the night, because of intensive surface heating. Outside of this period, nocturnal stable stratification was dominant.

Figure 1a shows the wind direction frequency rose at the 100-m level for the measuring period. This level was selected for its high data availability (near 100%) and also its distance from the ground surface. Sixteen



wind direction sectors were used in order to obtain a more detailed description. Very few observations are from directions between the southeast and south, and two prevailing sectors are observed. The first is clearly defined by the east-northeast and northeast directions, whereas the second is wider and is centered in the west-southwest. The data appear equally distributed between both of these prevailing sectors. Because orographic effects are excluded, it may be concluded that the airflow at the synoptic scale is the only factor responsible for this wind distribution. Following Tuller and Brett (1984), cumulative frequencies have been calculated in each sector. The 10th, 50th, and 90th percentiles have been drawn in Figs. 1b and 1c for wind speed and temperature, respectively. The median wind speeds indicate that the fastest winds come from the northeast and the slowest from the southeast, with a 6 m s^{-1} difference between the fastest and the slowest. Warmest advection corresponds to the south-southwest, where median temperature is 6°C hotter than east-northeast winds. Looking at the prevailing direction axis from east-northeast to west-southwest wind speeds are very similar, although a sharp 5°C contrast is also observed in median temperatures. As a consequence, the synoptic flow consisted of alternation between cold (east-northeast) and warm (west-southwest) advection. The position of the percentiles is indicative of the wind speed distribution shape. The wind speed is slightly left skewed in the east-northeast and is very symmetrical in the west-southwest. However, the temperature distribution is more symmetrical with a wide distribution in the west-southwest.

3. Results

a. Change in height of the autocorrelation function

The autocorrelation function was calculated by means of the Pearson correlation coefficients of the time series against itself (Wilks 1995). Although the usual practice is to complete the data gaps in the time series under different considerations, in this paper, missing data were excluded, and so no hypotheses have been made regarding the behavior of the meteorological variables considered. This assumption at the same time maintains all of the relevant information. Proof will be presented at the end of this paper of the robustness of this method for data exclusion.

This function is drawn as a black line in Fig. 2 for wind speed and in Fig. 3 for temperature in the low atmosphere for 20 days. A defined pattern is observed and comprises two parts, one deterministic and another random that only becomes relevant at higher levels. For

←

FIG. 1. (a) Wind direction frequency rose and 10th, 50th, and 90th percentiles for (b) wind speed and (c) temperature at 100 m during Apr 2001 at CIBA.

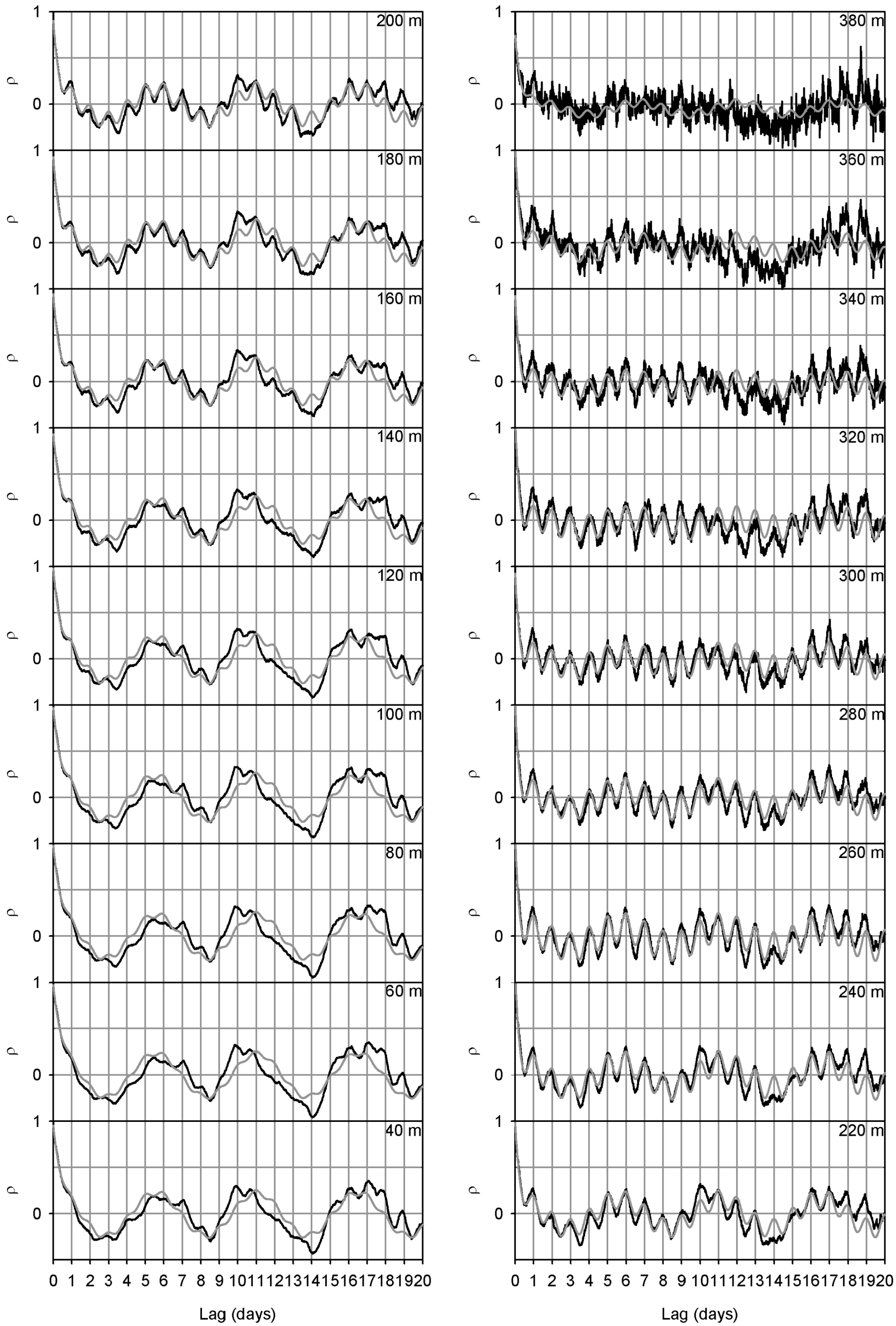


FIG. 2. Autocorrelation function for wind speed from experimental measurements (black) and modeled (gray) for heights of 40–380 m and lags of 0–20 days.

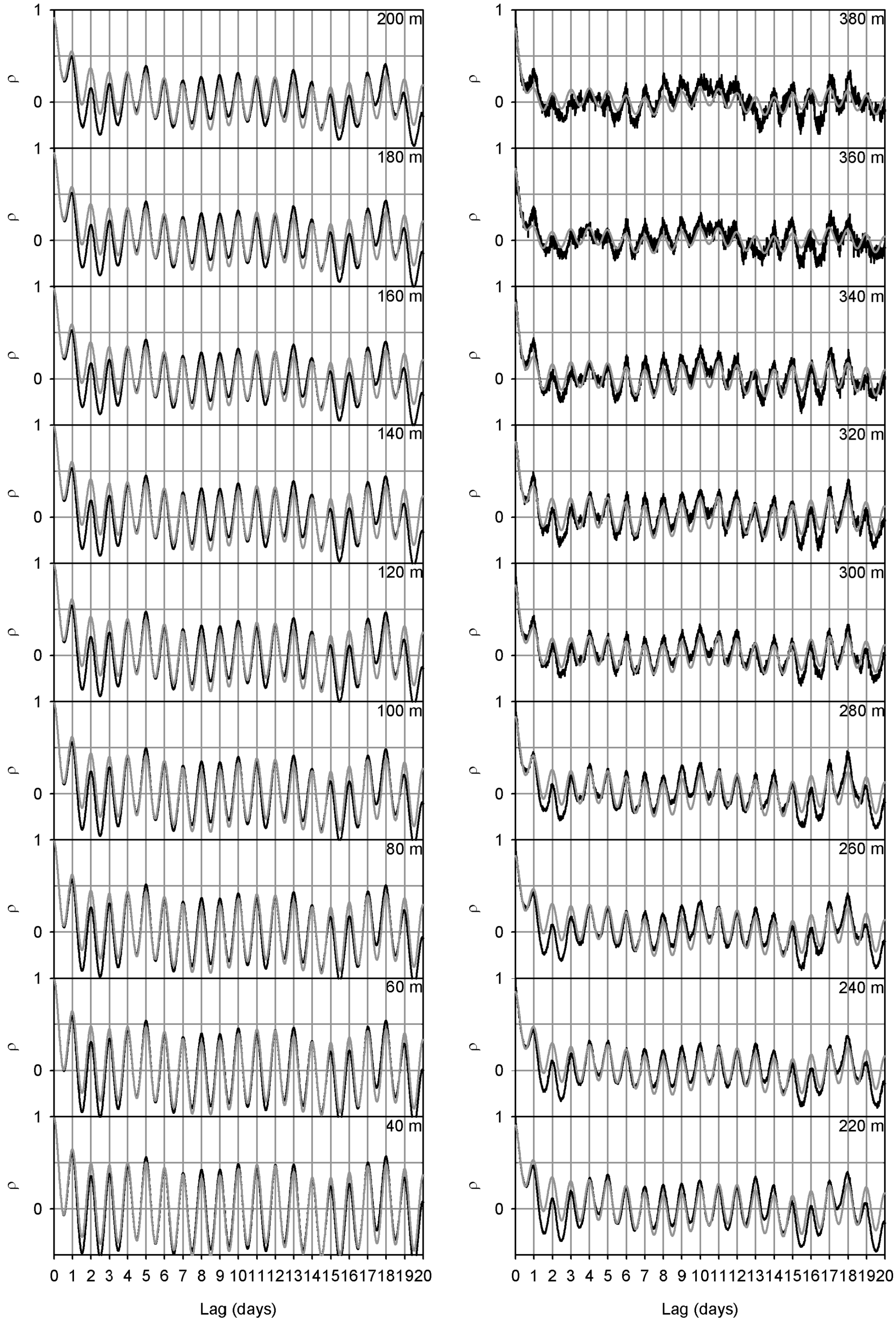


FIG. 3. Autocorrelation function for temperature from experimental measurements (black) and modeled (gray) as in Fig. 2.

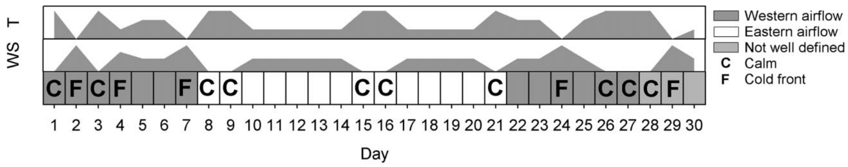


FIG. 4. Scheme corresponding to the synoptic pattern affecting the Iberian Peninsula in Apr 2001. (bottom) Dominent pattern by day: western flow (dark gray) with high pressure at low latitudes and warm winds; eastern flow (white) with displacement of the anticyclone to high latitudes and cold winds; flow that is not well defined (light gray). Here symbol C represents calm winds and increasing temperatures; F represents cold fronts, descending temperatures, and high winds. (middle) Time evolution of wind speed (WS) is qualitatively drawn, and (top) time evolution of temperature (T) is presented.

this reason, levels above 380 m were excluded because of the noise that appeared when the function is obtained with a small amount of data. Additionally, two components may be considered in the deterministic part—one transitory and another stationary, following an analogy with wave movement. The transitory part corresponds to a fast decrease from lag of 0 to 12 h, whereas the stationary part has two cycles: one daily and the other a multiday.

As a preliminary consideration, Pérez et al. (2003b) observed that wind speed and temperature 10-min averages from this sodar changed faster than those obtained from data taken on a meteorological mast. For this reason, it may be concluded that the data from this sodar clearly influence the autocorrelation function and, in this case, yield a faster initial fall in the autocorrelation function than data from a mast.

The absence of topographical features should make the autocorrelation function of the wind speed greater in the first 12 h. This should be even more so as height increases as the influence of the ground surface weakens. In this experiment, autocorrelation decreases more rapidly at higher levels where the wind speed is also highest and possible changes in wind speed become important. Initial behavior for temperature is quite the opposite, because of the decrease in the diurnal thermal wave as height increases.

From a quantitative point of view, it may be interesting to calculate the integral time scale, that is, the lag for an autocorrelation value of 0.37 (e^{-1}). This is a measurement of data "memory" and may be related to dispersion processes in the low atmosphere. With regard to wind speed, the integral time scale decreases from 11 h at 40 m to 5 h at 300 m. This shorter memory at intermediate–high levels is due to sharp wind speed changes and strong turbulent fluxes. In regard to temperature, the integral time scale increases from 7 h at 40 m to 8 h at 140 m and then remains essentially constant up to 260 m. This increase is due to the proximity of the ground surface, and the long memory observed is due to the slower temperature changes as height increases.

Above 180 m, the diurnal cycle is clearly evident for wind speed. The amplitude of this cycle increases with height but this trend is disturbed by the presence of noise at the higher levels. This behavior is consistent with the wind speed pattern: high wind speeds during the night from the stable stratification that favored horizontal movement, occasionally very clear inertial oscillations, and low horizontal winds during daylight hours because of convection. Therefore, a high contrast in wind speeds is visible at the higher levels, although not near the surface where friction prevented the development of high wind speeds. For temperature, the diurnal cycle is evident from the first level and becomes weakest at the highest level. Therefore, surface proximity is clearly relevant for temperature.

Last, a second kind of cycle is observed, mainly in wind speed. The periodicity in wind speed is evident below 300 m, although its amplitude decreases with height. For temperature, this cycle was less clear because of the more noticeable diurnal pattern. The series of synoptic structures affecting the Iberian Peninsula is responsible for this behavior. Near the ground the diurnal cycle is slightly present only for wind speed, and the only relevant item to be noticed is the oscillation between two prevailing wind directions that is due to the synoptic pattern and changes every 5 or 6 days. These two wind directions also have different thermal structures, although the strength of the diurnal oscillation conditioned this second cycle, yielding a slightly wider period for the temperature wave than for the wind speed.

The multiday cycle of wind speed and temperature schematically presented in Fig. 4 may be explained by a combination of three factors. The first is the time history of the high pressure system usually located west of the Iberian Peninsula. When this system is below 40°N, approximately, wet and warm westerly winds coming from the Atlantic Ocean affect the peninsula; whereas easterly airflow prevails and dry, cold winds from the continent appear when this system is located to the north (or northwest). In April 2001, this system was at low latitudes during the first 7 days of April. It then moved northward in the following 7 days and was again at low latitudes on 22 April. Finally, it began a slow ascent on 28 April, although this last stage is not well defined. Thus, a three-step cycle of around a 21-

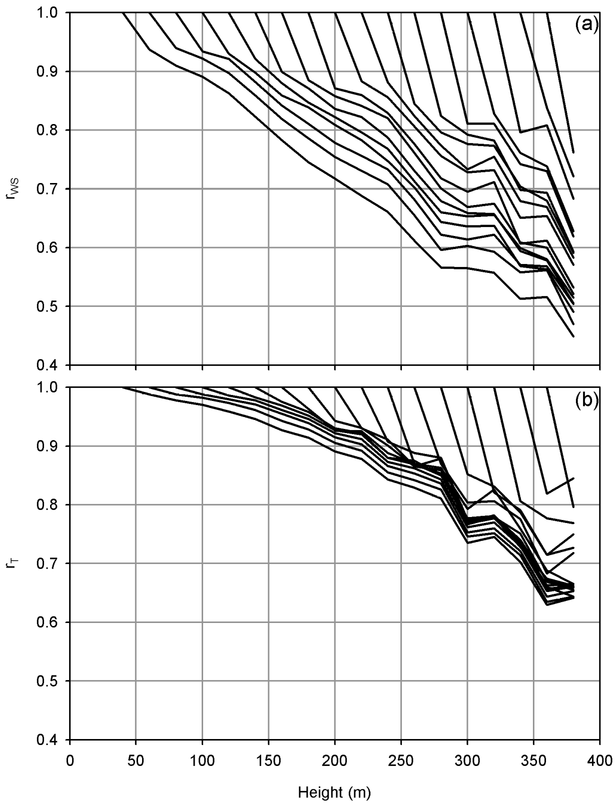


FIG. 5. Pearson correlation coefficient calculated between height levels for (a) wind speed r_{WS} and (b) temperature r_T. The reference level is the lowest for each line that connects correlations between levels above it. Correlations between one reference level and those below it are obtained by vertical downward displacement.

day period is inferred. The second factor is the cold fronts that penetrated from the west when the high pressure system lay at low latitudes. As a consequence, wind speed increases and temperature decreases in around 5 days. These fronts were less frequent under the eastern continental flow. The last but by no means least significant factor occurs in situations with very low wind speeds (i.e., calm conditions) and increasing temperatures observed under the influence of anticyclones, which occurs every 6 days at the beginning of every step in evolution of the high pressure system.

Correlations between height levels are presented in Fig. 5 where each line corresponds to upward correlation between its lowest (reference) level and the levels above. The main feature is the slope of the lines where changes reflect evolution of the vertical gradient of the correlation coefficient. The correlation between one reference level and those below it, is observed by downward vertical displacement. The distance between lines is the most relevant feature in this case; close lines imply light vertical gradients of correlation.

The change of the relative contribution from the two cycles for wind speed produces noticeable decay in correlations between the different levels as is depicted in Fig. 5a. The initial trend between the two first values shows an increasing decrease of correlation with height between the adjacent levels. However, if this initial trend is excluded, the lines are nearly straight with similar slopes and distance between them, indicating height independence in the correlation decay. However, Fig. 5b, drawn with the same scale for comparison, shows the strong influence of the diurnal cycle of temperature, whose correlations were higher than that of wind speed. The slightly curved trend of each line and increasing distances between the lines in downward displacement indicates height dependence in the correlation decay, which is stronger at higher distances from one reference level.

b. Autocorrelation function model

Figures 2 and 3 suggested that the transitory part was only relevant in the first 12 h, whereas the stationary part should only be considered from the first 24 h. For this reason, we have tried a simple function with the purpose of observing its variation with height.

Our proposal is the addition of two contributions, according to the two parts described above: two harmonic functions for the stationary part added to a simple exponential function for the transitory part. One of the harmonic functions must have a short period (1 day), and the other a longer period to consider the changing synoptic pattern. Our suggested expression is

$$\rho_i(j) = a_{0i} + a_{1i}e^{bj} + a_{2i}\cos\left(\frac{2\pi j}{T_1} + \varphi_{1i}\right) + a_{3i}\cos\left(\frac{2\pi j}{T_2} + \varphi_{2i}\right) + \varepsilon, \quad (1)$$

where the subscript i refers to the level, j corresponds to the lag, and ε considers the presence of a random noise (more relevant at higher levels).

Although the autocorrelation function sometimes decays more slowly than a simple exponential function, we mentioned above that our measurements do not have such a long ''memory,'' which suggested trying this function (Conradsen et al. 1984). An exponential fall of the autocorrelation in the first few hours has also been considered in ocean surface wind measurements (Sarkar et al. 2002).

Whereas the period T_1 was fixed and corresponded to the daily oscillation, the period T_2 must be obtained from the experimental function shown in Figs. 2 and 3. The lowest level was selected to establish this period. The values considered for T_2 were 787 lags (5.5 days) for wind speed and 936 lags (6.5 days) for temperature.

Our fitting procedure was a simple two-step method, using only linear regressions. The first step was to establish a multiple linear regression for each level and beyond the first day according to

$$\boldsymbol{\rho} = \mathbf{Xb}, \quad (2)$$

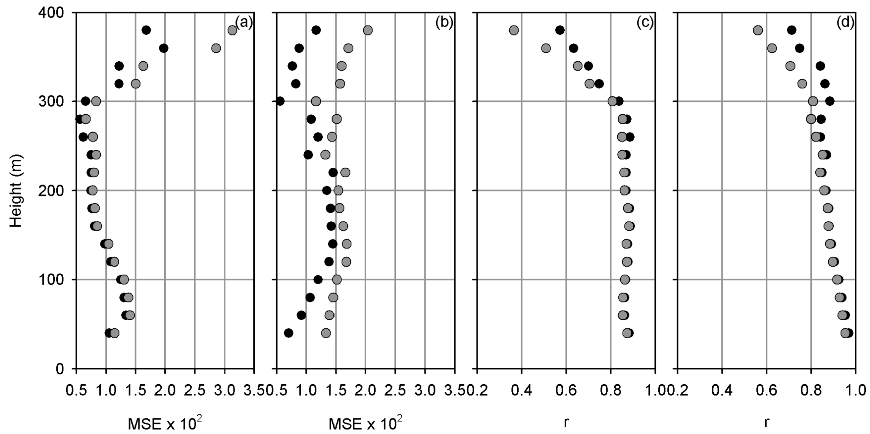


FIG. 6. MSE and Pearson correlation coefficients (r) between experimental autocorrelation and modeled by the method based on multiple linear regressions in black or calculated with the parameterized model from Eq. (5) in gray for (a), (c) wind speed and (b), (d) temperature.

where $\boldsymbol{\rho}$ was the matrix of correlation coefficients and the matrices $\mathbf{X}$ and $\mathbf{b}$ were

$$\mathbf{X} = \begin{bmatrix} 1 & \cos\left(\frac{2\pi k}{T_1}\right) & \sin\left(\frac{2\pi k}{T_1}\right) & \cos\left(\frac{2\pi k}{T_2}\right) & \sin\left(\frac{2\pi k}{T_2}\right) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \cos\left(\frac{2\pi l}{T_1}\right) & \sin\left(\frac{2\pi l}{T_1}\right) & \cos\left(\frac{2\pi l}{T_2}\right) & \sin\left(\frac{2\pi l}{T_2}\right) \end{bmatrix}$$

and

$$\mathbf{b} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \end{pmatrix} \tag{3}$$

and must be used between the lags k and l, in our case 144 and 2880, respectively. The matrix $\mathbf{b}$ was obtained as

$$\mathbf{b} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\boldsymbol{\rho}, \tag{4}$$

where $\mathbf{X}^{\mathrm{T}}$ is the $\mathbf{X}$ matrix transposed.

Once b_i is obtained, the values of the amplitudes and phase constants may be easily calculated. Natural logarithms of residuals for the first half-day were then considered and exponential function coefficients were calculated by a simple linear regression. The number of lags was lower when a few residuals were negative at higher levels.

The function fitted was drawn as a gray line in Figs. 2 and 3. The mean-squared error (MSE) was selected as a measurement of the goodness of fit (Wilks 1995), and the correlation coefficients have also been calculated as a measure of the coincidence of shape. Both quantities corresponding to the function fitting previously presented are drawn in Fig. 6 as black points. Figures 6a and 6c refer to wind speed, and Figs. 6b and 6d refer to temperature. The MSE values are very similar for the two variables, although contrasting behavior in height is observed; the higher values are reached at higher and lower levels for wind speed and at intermediate levels for temperature. The agreement suggested by the correlation coefficient below 300 m may be considered satisfactory. It is nearly constant for wind speed and decreases slightly for temperature. Lower values are obtained for wind speed above this level.

A quantitative indicator of the relative contribution of synoptic and diurnal components at different heights may be obtained by the relationship between the amplitude of their respective harmonics a_{3i}/a_{2i} from Eq. (1), which is presented in Table 1. The synoptic cycle of wind speed is highly pronounced in the lower levels where the diurnal cycle is only scarcely observed. However, the influence of the latter increases with height, whereas the first decays rapidly, with both cycles having

TABLE 1. Relationship between amplitudes calculated for the synoptic and daily cycles a_{3i}/a_{2i} from Eq. (1).

Height (m)	Wind speed	Temperature
40	7.58	0.18
60	9.28	0.20
80	7.35	0.22
100	6.54	0.23
120	5.31	0.25
140	4.38	0.26
160	3.59	0.28
180	2.75	0.27
200	1.98	0.28
220	1.57	0.31
240	1.07	0.36
260	0.67	0.39
280	0.77	0.42
300	0.55	0.39
320	0.40	0.39
340	0.55	0.48
360	0.97	0.89
380	1.43	0.71

the same contribution at 240 m. Above this level, the diurnal cycle prevails, although of little significance. The behavior for temperature is quite different than for wind speed. The synoptic cycle amplitude is, in fact, very small and nearly constant with height, and the diurnal cycle amplitude gradually decays with height. This controls the ratio values, which slowly increase with height.

The coefficients of Eq. (1) have been represented as a function of height and have been fitted to linear expressions that must be considered as a very rough approximation and only tentatively suggested for data below 380 m:

$$\rho_h(j) = 0.7e^{-0.03j} + (0.0005h - 0.01)\cos\left(\frac{2\pi j}{144}\right) + (-0.0006h + 0.28)\cos\left(\frac{2\pi j}{787}\right) \quad \text{and} \tag{5a}$$

$$\rho_h(j) = 0.7e^{-0.004j} + (-0.0009h + 0.42)\cos\left(\frac{2\pi j}{144}\right) + 0.07\cos\left(\frac{2\pi j}{936} + 1.6\right). \tag{5b}$$

Equations (5a) and (5b) are valid for wind speed and temperature autocorrelation, respectively. The faster height-independent exponential decrease for the wind speed should be noted in these expressions. For both harmonic wind amplitudes, the change rate is similar but opposite, indicating that the multiday cycle amplitude decrease is compensated by the daily cycle amplitude increase. For temperature, the multiday cycle amplitude is assumed to be constant and very small, whereas the emphasis is placed on the diurnal cycle amplitude decrease.

As before, the MSE and correlation coefficients, corresponding to this model whose coefficients have been parameterized in Eq. (5), were calculated again and drawn as gray points in Fig. 6. The new MSE values for wind speed autocorrelations are very similar to those previously obtained by direct fitting (black points), and only two prominent outliers appear at the highest levels analyzed. However, the new MSE values for temperature autocorrelations differed more noticeably, mainly at higher and lower levels, although they are very similar to those previously presented at intermediate levels. For the correlation coefficient, the agreement is slightly lower below 300 m and worse above this level.

For wind speed, the model suggested by Eq. (5) is approximated at lower levels where the multiday harmonic function only indicates the trend of the experimental function, whereas at higher levels the amplitude of the daily cycle is similar to noise, which is dominant. For temperature, the greatest discrepancies are found during the second and third days, although agreement is good in general.

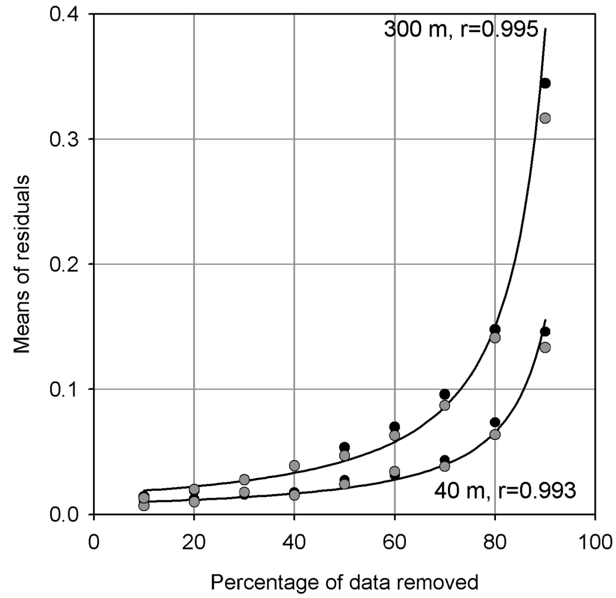


FIG. 7. Means of absolute differences (residuals) between autocorrelations with data removed and the original values as a function of percentage data removed for wind speed (black) and temperature (gray) at (bottom) 40 and (top) 300 m.

c. Influence of the available number of observations

In the following we analyzed for the presence of noise in the autocorrelation function arising from missing data. The 40- and 300-m levels were selected because of the well-defined pattern at these levels, and their different periodic behavior and autocorrelation values for both variables. Some initial data ranging in amount from 10% to 90% were selected and removed by generating random uniformly distributed numbers. New autocorrelations were calculated with these new samples. As an expected result, the greater was the amount of data removed, the more jittered was the resulting trace. Taking the initial autocorrelation as a reference, the absolute values of differences between new and old autocorrelation for each lag were calculated. The averages of these differences were calculated as a measurement of the noise introduced in the autocorrelation by the lack of data. The results are drawn in Fig. 7 for wind speed (black) and temperature (gray). Behavior is very similar for both variables and levels. For the noise–autocorrelation ratio, this figure suggests that there is a wide interval with a slight influence, and the presence of noise is only relevant at high data removal percentages. Temperature is less affected than wind speed at 40 m. However, data removal becomes more important at 300 m, where autocorrelations are low. As a final consequence, the greater importance of the wind speed diurnal cycle is responsible for the multiday cycle strength decrease with height, not the decrease in the number of data. The values depicted suggest using the following expression for the curves:

$$y = \frac{a}{(100 - x)^b}, \tag{6}$$

where y is the absolute difference in means and x is the removal percentage. The coefficients calculated by linear fitting are $a = 2.7$ and $b = 1.24$ for the 40-m level and $a = 9.1$ and $b = 1.37$ for the 300-m level. The behavior of this equation is satisfactory in the interval analyzed.

4. Conclusions

A RASS sodar placed on an extended plateau was used to obtain wind speed and temperature data during April 2001. Twenty-four levels were considered from 40 to 500 m. An analysis of wind direction revealed two clearly defined prevailing directions corresponding to cold and warm advections from the east-northeast and west-southwest, respectively. The wind blew alternatively between these two sectors from a synoptic point of view.

The autocorrelation function was calculated for the two variables. A well-defined pattern of autocorrelation was obtained, except at higher levels where low data availability introduced noise into the autocorrelation function.

An analysis of this function suggested that it could be considered as the sum of two parts, one deterministic and the other random. The random part became more relevant at higher levels, and the deterministic part had two separate zones—an initial fast transitory decrease, followed by a nearly stationary harmonic pattern.

Behavior with height for the initial fall in autocorrelation was opposite for the two variables. Autocorrelation initial fall increased with height for wind speed, whereas it decreased for temperature. This was due to the greater variability of wind speed that produced a ''short memory'' in the data.

The stationary harmonic pattern had two components: a diurnal and a multiday cycle. The diurnal cycle was important at intermediate levels for wind speed, because of the sharp contrast between day and night, where convection and horizontal movements respectively were relevant. The diurnal cycle for temperature decreased from the lower to higher levels as a result of the decrease with height of the thermal wave.

The multiday cycle was very visible at lower levels for wind speed and may be attributed to airflow on the synoptic scale. The first level (40 m) was used to establish a period of 5.5 days for this cycle that was consistent with the permanence time of weather systems affecting the Iberian Peninsula. This multiday cycle was less significant for temperature because of the strong influence of the diurnal pattern.

A simple model has been proposed, which consisted of adding an exponential function for the transitory part of the cycle and two harmonic functions for the stationary part. The model coefficients have been fitted by means of a simple method consisting of a multiple linear regression for beyond the first day and a simple linear regression for the first 12-h residuals. Satisfactory agreement between observed and modeled data was obtained, especially below 300 m, indicating the validity of our assumptions.

A further step was to make the proposed model height-independent. This objective was achieved by a rough approximation to constant terms or linear fits for the model coefficients. In this case, good agreement was once again obtained.

The influence of reducing the amount of data was also investigated. The resulting noise was only significant at higher data removal percentages. As a consequence, the autocorrelation is clearly robust.

The autocorrelation function has proven useful for determining synoptic airflow and establishing its range, the diurnal cycle range, and the memory of these variables. The method of model building suggested in this paper may be easily used for other stations, because airflow is closely linked to the specific conditions of a location. Although this paper presents only an initial statistical characterization of wind speed and temperature, the extension of autocorrelation to other atmospheric variables, such as atmospheric pollution, will undoubtedly prove promising.

Acknowledgments. The authors acknowledge the financial support from the Interministerial Commission of Science and Technology and the Regional Government of Castile and Leon.

REFERENCES

Arya, S. P., 1999: *Air Pollution Meteorology and Dispersion.* Oxford University Press, 310 pp.

Box, G. E. P., and G. M. Jenkins, 1970: *Time Series Analysis: Forecasting and Control.* Holden-Day, 553 pp.

Brett, A. C., and S. E. Tuller, 1991: The autocorrelation of hourly wind speed observations. *J. Appl. Meteor.,* **30,** 823–833.

Brunciak, P. A., J. Dachs, T. P. Franz, C. L. Gigliotti, E. D. Nelson, B. J. Turpin, and S. J. Eisenreich, 2001: Polychlorinated biphenils and particulate organic/elemental carbon in the atmosphere of Chesapeake Bay, USA. *Atmos. Environ.,* **35,** 5663–5677.

Conradsen, K., L. B. Nielsen, and L. P. Prahm, 1984: Review of Weibull statistics for estimation of wind speed distributions. *J. Climate Appl. Meteor.,* **23,** 1173–1183.

Daoud, W. Z., J. D. W. Kahl, and J. K. Ghorai, 2003: On the synoptic-scale Lagrangian autocorrelation function. *J. Appl. Meteor.,* **42,** 318–324.

Hauck, H., H. Kromp-Kolb, and E. Petz, 1999: Requirements for the completeness of ambient air quality data sets with respect to derived parameters. *Atmos. Environ.,* **33,** 2059–2066.

Kaimal, J. C., and J. J. Finningan, 1994: *Atmospheric Boundary Layer Flows.* Oxford University Press, 289 pp.

Oettl, D., R. A. Almbauer, and P. J. Sturm, 2001: A new method to estimate diffusion in stable, low-wind conditions. *J. Appl. Meteor.,* **40,** 259–268.

Pérez, I. A., M. L. Sánchez, M. A. García, and M. J. López, 2003a: Profiles of wind speed and temperature in the low atmosphere. *Proc. Third Int. Conf. on Experiences with Automatic Weather*

Stations, Torremolinos, Málaga, Spain, Ministerio de Medio Ambiente, NIPO 310-03-003-0.

——, ——, ——, and B. de Torre, 2003b: Comparison between measurements obtained with a meteorological mast and a RASS sodar. *Proc. Third Int. Conf. on Experiences with Automatic Weather Stations,* Torremolinos, Málaga, Spain, Ministerio de Medio Ambiente, NIPO 310-03-003-0.

Pryor, S. C., and R. J. Barthelmie, 2002: Statistical analysis of flow characteristics in the coastal zone. *J. Wind Eng. Ind. Aerodyn.,* **90,** 201–221.

Rehman, S., and T. O. Halawani, 1994: Statistical characteristics of wind in Saudi Arabia. *Renewable Energy,* **4,** 949–956.

Sarkar, A., S. Basu, A. K. Varma, and J. Kshatriya, 2002: Autocorrelation analysis of ocean surface wind vectors. *Proc. Indian Acad. Sci. Earth Planet. Sci.,* **111,** 297–303.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology.* Kluwer Academic, 666 pp.

Tomson, T., and M. Hansen, 2001: Seasonal wind stability on the west Estonian coast. *Proc. Est. Acad. Sci. Eng.,* **7,** 212–221.

Tuller, S. E., and A. C. Brett, 1984: The characteristics of wind velocity that favor the fitting of a Weibull distribution in wind speed analysis. *J. Climate Appl. Meteor.,* **23,** 124–134.

Wilks, D. S., 1995: *Statistical Methods in the Atmospheric Sciences.* Academic Press, 467 pp.