A Stochastic Model for Wind Occurrence

JOSE ROLDAN-CAÑAS, ADELA GARCIA-GUZMAN AND ALBERTO LOSADA-VILLASANTE

Escuela T.S. Ingenieros Agronomos, Universidad de Cordoba, Apartado 246, Cordoba, Spain

19 May 1981 and 8 December 1981

ABSTRACT

Knowledge of wind occurrence in any location is a valuable input in most agricultural studies. A bivariate stochastic process is proposed here as a model of daily wind occurrence. Two wind variables have been considered: 1) horizontal direction, fitted to a first-order Markov chain; and 2) speed, relative to the given direction, fitted to a gamma distribution. Performance of the model is tested by comparing the statistics of generated and experimentally observed data.

1. Introduction

Wind effects on agriculture have been recognized throughout history. Other than their mechanical influence on vegetation, they affect insect life, contribute to spreading plant diseases, and are related to important terms of the water balance equation, namely, precipitation and evaporation.

Because of the present oil crisis, wind energy availability and, therefore, wind speed have been a major concern in most recent statistical studies of wind occurrence (Corotis *et al.*, 1978; Justus *et al.*, 1979; Lund and Grantham, 1979). It is well known that aerodynamics of evaporation are influenced by the vertical wind profile; but wind direction is also a valuable characteristic to be considered when water balance conditioning factors are studied. In fact, under a given set of geomorphological and hydrological conditions, both air humidity and heat energy horizontal transfer (advection) can be associated with the place where the air is coming from.

Our purpose is to develop a joint probabilistic model of daily wind speed and wind direction occurrence. It should provide a forecast guidance to deal with wind-affected phenomena, and it could eventually be incorporated into more complete simulation models of water balance.

2. Model formulation

The wind event for the kth day of a period t, with wind speed $X_k(t)$ and wind direction $D_k(t)$, can be formulated as a bivariate stochastic process $\{X_k(t), D_k(t)\}$.

The process $D_k(t)$, assumed to be stationary during the tth period, is a first-order Markov chain of nine states described by the transition probability matrix

$$P_{ij}(t) = P(D_k(t) = j | D_{k-1}(t) = i); \quad i, j = 0, 1, \ldots, 8, \tag{1}$$

and by the probability of being in the jth state at the beginning of the tth period

$$P(D_0(t) = j) = p_j(t); \quad j = 0, 1, \ldots, 8. \tag{2}$$

State 0 marks the no wind event and states ranging from 1 to 8 represent the eight main directions on the compass card, i.e., 1 for the northern (N), 2 for the northeastern (NE), etc.

The Markov chain has been chosen as a simple model to account for the day-to-day persistence of wind direction. Markovian models seem to fit adequately to daily weather events (Gringorten, 1966). More specifically, the first-order Markov chain has been used for modeling daily precipitation by several authors (Todorovic and Woolhiser, 1975; Haan *et al.*, 1976; Katz, 1977).

The process $X_k(t)$, under the hypothesis of unique dependence on $D_k(t)$, is described by its conditional distribution function

$$P(X_k(t) \leq x | D_k(t) = j) = F_j(x, t); \quad j = 1, 2, \ldots, 8, \tag{3}$$

where $F_j(x, t)$ is the gamma distribution with parameters α_{jt} and λ_{jt}, i.e.,

$$F(x; \alpha, \lambda) = \lambda^\alpha [\Gamma(\alpha)]^{-1} \int_0^x x^{\alpha-1} e^{-\lambda x} dx. \tag{4}$$

The gamma distribution was used for modeling wind speed by Neumann (1977) and by Romanenko (1976) among others.

Although the dependence between wind speed and wind direction is not a general assumption, it seems

0021-8952/82/050740-05$05.25

a reasonable one for areas close to mountainous systems. In fact, it is widely supported by our observed data (see Section 3 and Table 4).

The unconditional distribution of $X_k(t)$ and $D_k(t)$ can be obtained by using (1), (2) and (3), i.e.,

$$P[X_k(t) \leqslant x, D_k(t) = j] = P[D_k(t) = j]F_j(x, t) = \sum_{i=0,8} p_i(t)P_{ij}^k(t)F_j(x, t), \quad (5)$$

where $P_{ij}^k(t)$ is the (i, j) element of $(\mathbf{P}(t))^k$ and $(\mathbf{P}(t))^k$ is the transition probability matrix given in (1) raised to the kth power.

From (5) we can obtain the marginal distribution of $X_k(t)$:

$$P[X_k(t) \leqslant x] = \sum_{j=0,8} \{ \sum_{i=0,8} [p_i(t)P_{ij}^k(t)F_j(x, t)]\}, \quad (6)$$

or using matrix notation

$$P(X_k(t) \leqslant x) = \mathbf{p}(t)(\mathbf{P}(t))^{k-1}\mathbf{Q}(x, t)\mathbf{e}, \quad (7)$$

where $\mathbf{e}$ is a column vector of 1's and $\mathbf{Q}(x, t)$ is given by

$$\mathbf{Q}(x, t) = \begin{bmatrix} P_{00}(t) & P_{01}(t)F_1(x, t) \cdots P_{08}(t)F_8(x, t) \\ \vdots & \\ P_{80}(t) & P_{81}(t)F_1(x, t) \cdots P_{88}(t)F_8(x, t) \end{bmatrix}. \quad (8)$$

Similarly, the multivariate distribution of wind speed is

$$P(X_1(t) \leqslant x_1; X_2(t) \leqslant x_2; \ldots ; X_k(t) \leqslant x_k) = \mathbf{p}(t) \prod_{r=1,k} [\mathbf{Q}(x_r, t)]\mathbf{e}. \quad (9)$$

FIG. 1. Location of the meteorological station used as an example in this note.

TABLE 1. χ^2 test for independence versus first-order Markov chain.

Month	Statistic	Month	Statistic	Month	Statistic
January	191.25	May	118.46	September	139.45
February	179.33	June	138.23	October	179.02
March	168.02	July	159.65	November	151.35
April	152.97	August	91.18	December	204.40

Degrees of freedom: 64. Critical value for rejecting the null hypothesis is $W_{1-\alpha} = 83.66$ ($\alpha = 0.05$).

The persistence of wind velocity below and above a fixed value is taken into account by (9).

As can be noted (7) is a particular case of (9) with $x_r = \infty$, $r = 1, k - 1$, and $\mathbf{Q}(\infty, t) = \mathbf{P}(t)$.

A corollary of the distribution function (9) is the distribution function of extreme events. For instance, the maximum speed distribution in a n-day period is

$$P(X(t)_{\max} \leqslant x) = P(X_1(t) \leqslant x; X_2(t) \leqslant x; \ldots ; X_n(t) \leqslant x) = \mathbf{p}(t)(\mathbf{Q}(x, t))^n\mathbf{e}, \quad (10)$$

while the minimum speed distribution is

$$P[X(t)_{\min} \leqslant x] = 1 - P[X(t)_{\min} > x] = 1 - \mathbf{p}(t)[\mathbf{R}(x, t)]^n\mathbf{e}. \quad (11)$$

$\mathbf{R}(x, t)$ is a square matrix of ninth order whose elements are

$$\left.\begin{array}{ll} R_{ij} = 0, & j = 0 \\ R_{ij} = P_{ij}(t)[1 - F_j(x, t)], & j = 1, 2, \ldots, 8 \\ & i = 0, 1, 2, \ldots, 8 \end{array}\right\}. \quad (12)$$

3. Application of the model

Data from the meteorological station of Cordoba, Spain, were used for testing the model. Collected data are instantaneous wind speed, as recorded by an anemometer installed at a height of 2 m above the ground level, and wind direction. The record length is 20 years (1959–78). Available data refer

TABLE 2. χ^2 test for stationarity versus non-stationarity.

Group*	Statistic	Degrees of freedom	Critical value ($\alpha = 0.05$)
1	239.49	356	401.00
2	396.34	352	396.75

* Group 1: From October to March.
Group 2: From April to September.

TABLE 3. Estimated parameters for gamma distribution and χ^2 goodness of fit tests.

	Group 1				Group 2			
Direction**	α	λ	Statistic	Degrees of freedom	α	λ	Statistic	Degrees of freedom
N	4.98	0.32	4.76	2	4.22	0.36	0.41	2
NE	5.92	0.38	2.11	4	5.06	0.32	6.72	5
E	6.94	0.48	15.30*	3	5.12	0.37	5.83	3
SE	4.66	0.37	2.20	1	7.36	0.65	0.90	1
S	2.97	0.16	2.44	5	4.50	0.33	4.72	4
SW	3.58	0.14	12.56	6	5.79	0.27	8.45	5
W	4.56	0.24	2.00	4	6.67	0.35	5.32	5
NW	3.08	0.13	8.40	5	4.34	0.28	10.85	5

* The computed value is over the critical value $W_{1-\alpha} = 7.82$ ($\alpha = 0.05$).
** See Section 2 for symbol explanation.

to 0700, 1200 and 1800 GMT. A similar behavior at noon and at 1800 GMT was observed. Records at 0700 have not been considered as being required to describe the diurnal situation as compared to the other two measurements (Roldan, 1979). Therefore, it has been assumed that diurnal wind fields are described by daily readings at noon.

The meteorological station of Cordoba is part of the network of the Spanish National Meteorological Service (Servicio Meteorologico Nacional). It is located in the middle of the Guadalquivir River Valley, in the southwestern part of Spain and its geographical position corresponds to 4°51′W longitude, 37°51′N latitude and altitude of 110 m MSL. The triangular shaped valley, bounded abruptly by a sierra to the northwest, and by hilly country with a wavy boundary in the southeast, is open to the Atlantic Ocean toward the southwest (see Fig. 1).

Preliminary statistical studies on wind patterns in this area revealed both a strong dependence between wind speed and its direction and a seasonal variation in the wind direction occurrence (Roldan, 1979). A summary of the results is shown in Table 4. It can be seen that the average wind speed shows a range from 24.69 (SW direction) to 12.63 km h^{-1} (SE direction) in group 1 and from 21.27 (SW direction) to 11.28 km h^{-1} (SE direction) in group 2. SW wind direction showed not only the highest monthly average speed throughout the year but also the highest occurrence frequency. Nevertheless, some correspon-

TABLE 4. Average and variance of wind speed[a] and percentage of occurrence for each wind direction.

		Average		Variance		Percent of occurrence	
Group[b]	Direction[c]	Simulated	Observed	Simulated	Observed	Simulated	Observed
1	N	16.09	15.72	58.90	49.63	0.58	0.65
	NE	15.88	15.75	43.79	41.89	22.00	22.18
	E	14.71	14.56	30.18	30.55	17.56	18.11
	SE	12.51	12.63	37.61	34.21	0.83	0.80
	S	19.91	18.73	128.79	118.31	2.02	2.16
	SW	24.69	24.69	165.79	170.20	20.28	19.92
	W	18.71	18.63	74.32	76.10	7.32	7.32
	NW	23.82	24.05	185.77	187.58	2.40	2.43
	calm	—	—	—	—	27.01	26.43
2	N	11.34	11.88	29.67	33.44	2.57	2.50
	NE	15.77	15.83	30.23	49.55	9.68	9.81
	E	13.66	13.86	33.01	37.49	5.33	5.58
	SE	11.34	11.28	15.25	17.29	0.97	1.01
	S	13.64	13.77	40.69	42.08	3.24	3.19
	SW	21.26	21.27	84.74	78.04	27.98	28.04
	W	18.85	18.88	56.00	53.38	19.29	19.01
	NW	15.53	15.36	58.94	54.41	6.68	6.68
	calm	—	—	—	—	24.26	24.19

[a] Wind speed is measured in km h^{-1}.
[b] Group 1: from October to March. Group 2: from April to September.
[c] See Section 2 for symbol explanation.

TABLE 5. Simulated and observed frequency of runs.

	Length of the runs (days)	1	2	3	4	5	6	7	8	9	10	11
Group 1												
calm	Observed	366	124	48	16	9	1	1	0	0	0	1
	Simulated	366	132	53	19	7	4	1	0	0	0	0
N	Observed	20	1									
	Simulated	17	1									
NE	Observed	332	110	35	15	2	1	1	1			
	Simulated	348	108	39	11	5	2	1	0			
E	Observed	303	72	32	11	1	2	1				
	Simulated	307	84	26	8	3	0	1				
SE	Observed	27										
	Simulated	29										
S	Observed	67	3									
	Simulated	66	2									
SW	Observed	259	76	32	14	9	7	2	1			
	Simulated	255	104	39	14	8	3	1	1			
W	Observed	159	35	6	0							
	Simulated	160	31	8	2							
NW	Observed	77	1	1								
	Simulated	78	3	0								
Group 2												
calm	Observed	352	120	39	18	8	1	2				
	Simulated	333	125	48	16	6	2	1				
N	Observed	78	3	1								
	Simulated	80	5	0								
NE	Observed	210	44	13	1							
	Simulated	213	43	10	2							
E	Observed	130	23	3	1	1						
	Simulated	125	22	5	1	0						
SE	Observed	31	2									
	Simulated	30	2									
S	Observed	95	8									
	Simulated	98	7									
SW	Observed	347	139	39	29	15	2	2	2			
	Simulated	347	133	60	26	7	4	1	1			
W	Observed	326	92	29	7	6	1	0				
	Simulated	324	100	28	9	4	1	1				
NW	Observed	197	11	3	1							
	Simulated	187	20	1	0							

dence does exist between wind direction and the season: the NE and E winds, together with the SW, are the most frequent in fall and winter (group 1) while the W winds, together with the SW, are the most frequent in the remaining seasons (group 2). In any case, the average wind speed is higher in fall and winter.

Both the geographical position and the geomorphology of this area help to explain its climatic behavior. In the valley, general trends are dominated by the proximity of the Azores anticyclone, bordered on the north by a track of fronts. On the other hand, since the valley is wider and its topography smoother in the southwestern part, wind entrance is easier from the SW direction. Depending on the season, however, the air temperature gradient between the land and water surfaces generates air movement which may disturb this pattern of wind direction occurrence. The higher speed of SW winds is another consequence of the lack of geomorphological obstructions in the southwestern area of the valley and of its shape, which narrows as extends further away from the southwest.

4. Parameter estimation

As a preliminary study, a transition probability matrix was estimated for each month and each tested for independence against the hypothesis of a first-order Markov chain. Details of the testing procedure can be found in Anderson and Goodman (1957). Results of the test are given in Table 1.

Having assumed the first-order Markov chain, the hypothesis of stationarity for transition probabilities was subsequently checked. Stationarity was not accepted for a period of a year, but was accepted for two six-month periods: one from October to March (group 1), the other from April to September (group 2). Table 2 gives the results of the stationarity tests. Degrees of freedom for these tests are: $9(9-1)(6-1)-d$, where d is the number of zeros in the transition probability matrices (see Bhat, 1972).

Parameters of the gamma distribution [α and λ in (4)] for wind speed in km h^{-1}, were estimated from data, and goodness of fit to the gamma distribution was tested. Results of this part are summarized in Table 3. Because the statistics are below the critical value in all cases, except for the easterly direction in group 1, the gamma distribution is accepted. Degrees of freedom for the goodness of fit test are: $k-3$, where k is the number of intervals in which the sample is divided. In Table 3 less frequent wind directions have fewer degrees of freedom, because k is chosen in such a way that none of the intervals has less than five observations.

5. Model performance

In order to ascertain the model capability in reproducing observed data, five series of the same length of the historical record were simulated.

Table 4 shows a good agreement between some parameters of the simulated and observed data. Simulated parameters stand for the average of those calculated from the five synthetic records.

Table 5 confirms the model ability to represent the wind direction occurrence pattern. A run of length n for the ith direction is defined here as a period of $n+2$ consecutive days, such that

$$J_k = i; \quad k = 1, n \quad \text{and} \quad J_0 \neq i, \quad J_{n+1} \neq i,$$

and the number of simulated occurrences in each length and direction is the average of five simulations rounded to the closest integer.

These results demonstrate the capability of the model to reproduce the behavior of the real data.

6. Conclusions

A stochastic model for daily wind is proposed with the following basic assumptions: 1) daily occurrence of wind direction fits a first-order Markov chain; 2) wind speed on the kth day is conditioned by the wind direction on the same day; and 3) wind speed is fitted by a gamma distribution. The model gives reasonably good predictions of the behavior of such meteorological phenomena and appears to be of interest for agricultural and civil engineering purposes.

Acknowledgments. The authors wish to thank Dr. J. V. Giraldez for his many valuable suggestions and comments as well as the Meteorological Station of the Cordoba airport of the Spanish National Meteorological Service for providing the data.

REFERENCES

Anderson, T. W., and L. A. Goodman, 1957: Statistical inference about Markov chains. *Ann. Math. Stat.*, **28,** 89–110.

Bhat, U. N., 1972: *Elements of Applied Stochastic Processes.* Wiley, 414 pp.

Corotis, R. B., A. B. Sigl and J. Klein, 1978: Probability models of wind velocity magnitude and persistence. *Solar Energy,* **20,** 483–493.

Gringorten, I. I., 1966: A stochastic model of the frequency and duration of weather events. *J. Appl. Meteor.,* **5,** 606–624.

Haan, C. T., C. M. Allen and G. O. Strett, 1976: A Markov chain model of daily rainfall. *Water Resour. Res.,* **12,** 443–449.

Justus, C. G., K. Mani and A. S. Mikhail, 1979: Interannual and month-to-month variations of wind speed. *J. Appl. Meteor.,* **18,** 913–920.

Katz, R. W., 1977: Precipitation as a chain-dependent process. *J. Appl. Meteor.,* **16,** 671–676.

Lund, I. A., and D. D. Grantham, 1979: Estimating recurrence probabilities of weather events. *J. Appl. Meteor.,* **18,** 921–930.

Neumann, J., 1977: Averaging for wind speed in a given direction in the concentration equation for pollutants. *J. Appl. Meteor.,* **16,** 1097–1100.

Roldan, J., 1979: Caracterizacion del regimen hidrologico de una region: sistemas de vientos y lluvias. Ph.D. dissertation, Universidad de Cordoba, Spain, 216 pp.

Romanenko, T. P., 1976: Statistical structure of surface-wind time series. *Soviet Hydrol. Meteor.,* **4,** 24–27.

Todorovic, P., and D. A. Woolhiser, 1975: A stochastic model of n-day precipitation. *J. Appl. Meteor.,* **14,** 17–24.