A Stochastic Model of Surface Wind Speed for Air Quality Control Purposes

G. FINZI

Dipartimento di Elettronica, Centro Teoria dei Sistemi CNR, Politecnico, Milano, Italy

P. BONELLI

ENEL, Centro Ricerche Termiche e Nucleari, Milano, Italy

G. BACCI

Centro Meteorologico Regionale, Linate, Italy

(Manuscript received 8 July 1983, in final form 2 June 1984)

ABSTRACT

Wind speed in the lower layers of the atmosphere is a relevant factor in pollutant dispersion and therefore its forecast is an essential datum for any reliable real-time predictor of ambient concentrations. Here a stochastic model for one-day forecasts of the surface daily wind speed at a site is described. According to the model, this speed is a linear combination of the surface wind speed of the previous day, the daily variation of the 500 mb wind speed over the site, the daily variation of the 500 mb geopotential over the site, and noise. Moreover, the coefficients of the linear combination are assumed to depend both on a "synoptic category" (defined in light of the prevailing synoptic circulation) and on the 500 mb wind direction over the site. The forecast performance is satisfactory for the two sites in the Po Valley (northern Italy) where the model has been tested. Furthermore, it is much better than the performance of a simple autoregressive model. Since such a model uses recent values of the forecast variable (surface daily wind speed) as its only information, the difference in performance shows the importance of using additional meteorological information (500 mb wind and geopotential, and "synoptic category") for obtaining a reliable forecast.

1. Introduction

As an alternative to deterministic models (*K*-type, Lagrangian, Gaussian), time series analysis techniques, namely the so-called ARMAX (autoregressive moving average with exogenous inputs) stochastic models (see Box and Jenkins, 1970) have been widely used to describe the dispersion of air pollutants on a local scale (many examples are shown in Benarie, 1980).

From the point of view of air quality management, the most relevant application of these models is the short-term (≤24 h) real-time prediction of future ambient concentrations. In fact, whenever the predicted concentrations exceed the pre-assigned standards, proper actions can be taken (e.g., load shifting, load reduction or fuel switching in industrial plants), or at least an alarm can be given (see, for example, Soeda and Omatu, 1982, Bolzern and Fronza, 1982).

Naturally, in view of such use for control or alarm purposes, an ARMAX concentration predictor is required to be particularly accurate in situations of forthcoming pollution "episodes." In turn, its performance in episode conditions depends almost entirely on the quality of the local meteorological forecasts and precisely on the availability of reliable forecasts of the meteorological variables which are inputs for the ARMAX models (see, for example, Bacci *et al.*, 1981, Finzi and Tebaldi, 1982).

In particular, site wind speed is a major explanatory factor in most cases. This is true not only for episodes due to low sources in calm and inversion situations, but also for episodes due to tall stacks, e.g., in calm situations with a sufficiently high fog layer, in the presence of a strong wind with plume breakdown, and even for summer fumigation due to an unstable convective layer.

In all these cases, the time evolution of site wind speed explains most, or at least part of the pollution event. Therefore, wind speed is the first variable which should be forecast in order to supply ambient concentration predictors with reliable meteorological information.

The following considerations must be taken into account when choosing a mathematical model for real-time forecasting of site wind speed:

- The complexity of the wind speed model must be comparable with the complexity of the concentration predictor to which the wind speed model supplies its forecasts. For instance, it would not be reasonable to supply an ARMAX concentration predictor, which gives only an aggregate description of the future pollution field, with an extremely detailed forecast of the vertical wind profile in the lower layers.

• Deterministic wind models do not offer satisfactory real-time forecasts of wind speed in the lower layers of the atmosphere (see, for example, Bengtsson, 1976), while their performance is good so far as the circulation aloft is concerned. In fact, because of their grid size, such models cannot take into account a very detailed profile of the terrain. In the case of the Po Valley, in particular, the complexity of the surrounding orography often makes the characteristics of the lower layers in two relatively close sites significantly different.

In view of these considerations, in order to obtain real-time forecasts of surface (30 m, say) wind speed for the Po Valley sites for ARMAX pollution prediction purposes, a stochastic model is used (see, for example, Bonivento *et al.,* 1980; Bacci and Finzi, 1980; Godfrey, 1982; Kau *et al.,* 1982; Roldan-Cañas *et al.,* 1982; McWilliams and Sprevack, 1982).

The structure of the model and the procedure for estimating its coefficients by recorded time series are described in detail in Section 3. It is sufficient to say here that the nature of model inputs allows the real-time forecast to be obtained with relative ease. In fact, such inputs are of two distinct types:

1) The last value of the variable to be predicted (surface wind speed at the site).

2) Some geostrophic variables, which can be directly obtained through routine information supplied by the meteorological service.

Thus, the resulting forecast turns out to combine information on both the surface situation at the site and the upper circulation over the site.

The forecast performance has been tested against real data from two sites in the Po Valley and has been found to be satisfactory in both cases (see Section 4). In particular, it has been compared with the performance of the following two simpler predictors:

1) The persistence predictor (future wind speed will be equal to the present), which does not require any collection of information.

2) The autoregressive (AR) predictor, which makes use only of information about the recent values of the forecast variable, but does not require collection of any data on the upper circulation.

Therefore, these comparisons point out the actual value of both the surface and the geostrophic meteorological information for the forecast.

Finally, it must be pointed out that the performance test has been carried out using domestic heating season data (November–March) because this is the season characterized by the most significant air pollution events.

2. The choice of model inputs

When analyzing the characteristics of the wind field in the lower layers of the Po Valley (Fig. 1), it is first necessary to distinguish between stationary and evolutionary situations.

During stationary situations (i.e., low baroclinicity), the surface wind at a given site is mainly driven by two phenomena:

1) The differential heating between plains and mountains and/or between land and sea produces breezes with a daily cycle (see, for example, Holton, 1979 and, for the region under consideration, Gandino, 1976).

2) The interaction between the upper flow and the surrounding orography produces a surface pressure gradient, dependent on the intensity and especially on the direction of such flow (see again Holton, 1979).

Evolutionary situations usually correspond to the rapid passage of a frontal system. In this case, strong surface winds (5–10 m s^{-1}) may occur. Again, for a given site, the orographic effect on the flow is strongly related to the flow direction. In fact, northerly and westerly flows meet the Alps (average height ~ 2500 m), southerly flows meet the Appennines (average height ~ 1500 m), and easterly flows enter directly across the Adriatic coast encountering no significant obstacles. Therefore, at a given site and for a given upper wind speed, the surface speed can be very different, according to the direction of the upper wind and, conversely, even a uniform upper circulation yields rather different surface winds at the various sites in the Po Valley.

These considerations suggest that information about the circulation aloft must be supplied in order to set up a reliable predictor of the surface wind speed at a

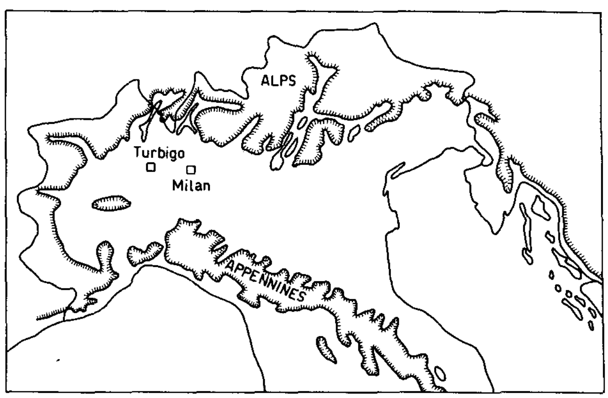

FIG. 1. Northern Italy and the Po Valley.

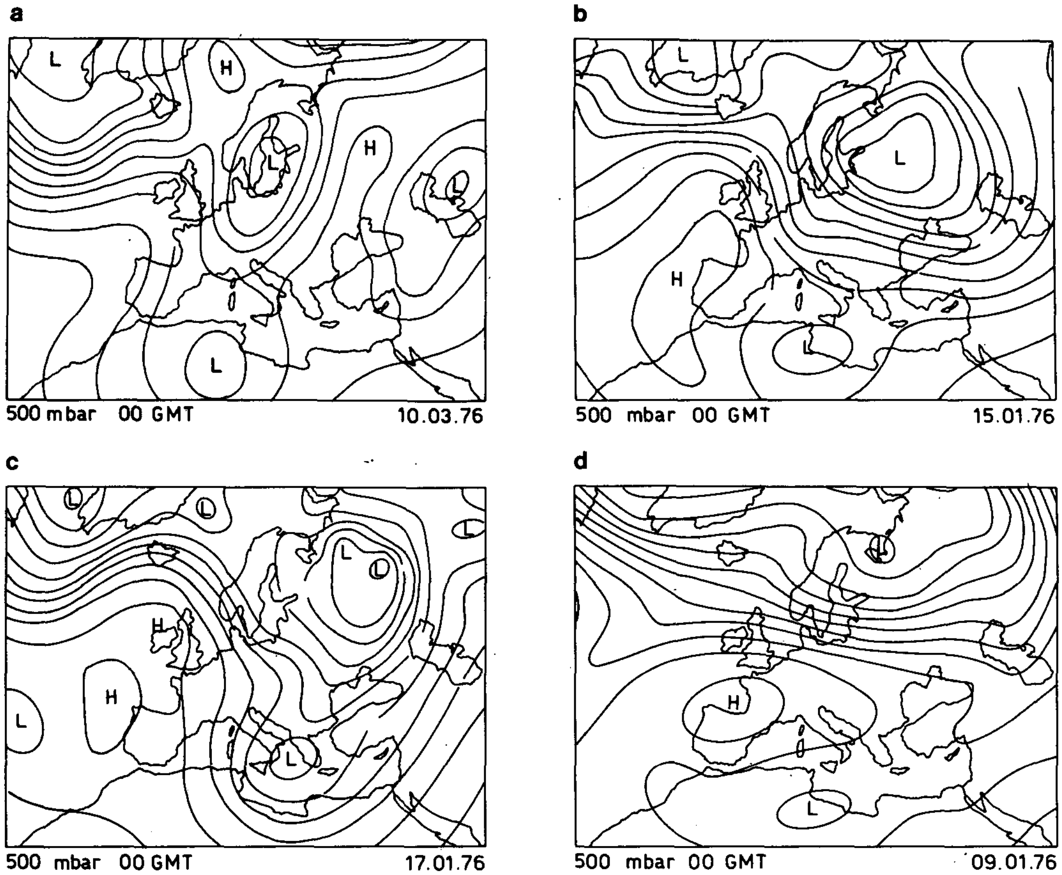

FIG. 2. Examples of synoptic circulation categories over Northern Italy for the times and dates (day, month, year) shown. (a) cyclonic circulation, (b) weak cyclonic circulation, (c) weak anticyclonic circulation, (d) anticyclonic circulation.

given site. As a matter of fact, the following "upper circulation" variables have been chosen as exogenous inputs for the daily site wind speed model (of the stochastic type) which is described in detail in the next section:

- An index (synoptic category) which gives an approximate picture of the 500 mb height field over Northern Italy. Four classes of this field have been empirically defined: I, cyclonic with strong curvature; II, cyclonic with weak curvature; III, anticyclonic with weak curvature; IV, anticyclonic with strong curvature. An example for each class is shown in Fig. 2.
- An index representing the class of the 500 mb wind direction over the site. Taking into account the different orographic effects on the flow mentioned above, four sectors of upper flow direction have been defined for every class of 500 mb circulation (Fig. 3).
- The 24 h change in 500 mb wind speed (between the day under consideration and the preceding day) over the site.
- The 24 h change in 500 mb geopotential over the site.

In particular, both the 500 mb speed and 500 mb geopotential variations between the day under consideration and the previous day have turned out, *a posteriori,* to be better inputs than simply the 500 mb speed and 500 mb geopotential of the day. In fact, the variations also contain useful information on whether the situation is stationary or evolutionary.

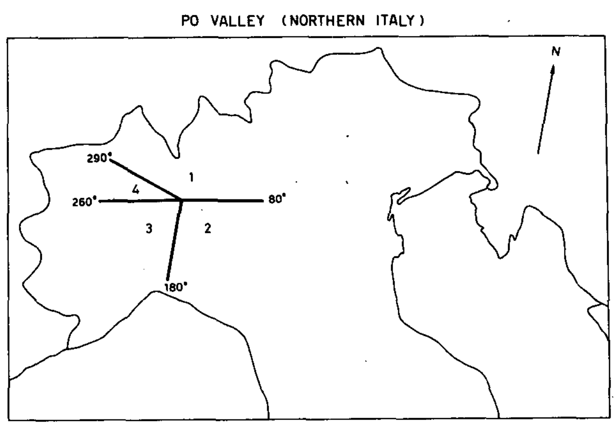

FIG. 3. The four sectors of geostrophic wind direction.

The actual role of all these inputs with respect to the overall reliability of the daily site speed model is discussed in Section 4. It is sufficient to say here that *a priori* there was no reason to discard, as a carrier of no additional information, any of the four variables listed above. In fact, the correlations between each pair of the four variables are rather weak (see Table 1 where an example of the cross-correlation matrix between the four variables is shown). Hence, *a priori,* the information given by the value of one of the variables is not "entirely contained" in the value of another variable. This is in accordance with physical intuition, which assigns to each of the abovementioned variables a different piece of information about the upper circulation over the site.

3. The stochastic model of daily site wind speed

Basically, stochastic models differ from deterministic models for three reasons:

1) Their structure is arbitrarily predetermined (though in a "reasonable way") instead of being derived from physical laws. Thus a stochastic model can be viewed as a sort of extreme simplification of physical mechanisms.

2) The model coefficients are estimated by statistical techniques from recorded data instead of being assigned on the basis of physical considerations.

3) All the "errors" affecting the model are explicitly pointed out in the form of a noise term. The statistical properties of the noise, estimated from the so-called model residual (see, for example, Box and Jenkins, 1970), is taken into account when using the model for predictive purposes.

In view of the general considerations of the previous section, daily site surface wind speed has been given the following stochastic representation:

$$\begin{aligned} v_s(i) = {} & \alpha[c(i), d(i)]v_s(i-1) \\ & + \beta[c(i), d(i)]\{v_h(i) - v_h(i-1)\} \\ & + \gamma[c(i), d(i)]\{g(i) - g(i-1)\} \\ & + \eta[c(i), d(i)] + \epsilon(i), \quad (1) \end{aligned}$$

TABLE 1. Cross-correlation matrix between the meteorological variables (Milan, synoptic category I, 500 mb wind direction from sector 3).

Meteorological variables	1	2	3	4
1 (local wind speed of the day)	1.0	0.483	0.404	0.226
2 (local wind speed of the preceding day)	0.483	1.0	0.146	0.003
3 (500 mb wind speed daily variation)	0.404	0.146	1.0	0.051
4 (500 mb geopotential daily variation)	0.226	0.003	0.051	1.0

where $v_s(i)$ is the average surface site wind speed (m s^{-1}) on the ith day (the interval 1200 GMT of the $(i-1)$th day to 1200 of the ith day); $v_h(i)$ the 500 mb wind speed (m s^{-1}) at 0000 GMT of the ith day; $g(i)$ the 500 mb height (m) over the site at 0000 GMT of the ith day; $\epsilon(i)$ the purely random zero mean stochastic process (white noise); $d(i)$ the 500 mb wind direction category (1, 2, 3 or 4; see Fig. 3); $c(i)$ the synoptic circulation category (I, II, III or IV; see previous section and Fig. 2); and α, β, γ, η are model coefficients. Though arbitrary, form (1) can be given a sort of gross physical justification. Specifically, from Eq. (1) the surface daily wind speed at a site in the Po Valley turns out to be the sum of three contributions [respectively ascribed to the speed of the previous day, $v_s(i-1)$, the daily variation of the 500 mb wind speed, $v_h(i) - v_h(i-1)$, and the daily variation of the local geopotential over the site, $g(i) - g(i-1)$], plus a residual term $\eta[c(i), d(i)] + \epsilon(i)$.

Neglecting for a while the residual term which represents the unexplained part of the dependent variable, the three addenda in the right-hand side of (1) can be "interpreted" as follows. The term $\alpha v_s(i-1)$ is the "stationary component" of v_s, i.e., that part of v_s which is explained by the previously existing situation at the site. The other two addenda are the "evolutionary components" of v_s, i.e., those parts of v_s which can be ascribed to the change in the upper situation over the sites. All the contributions are weighted by coefficients which are assumed to depend on the upper circulation field, represented by the pair $c(i)d(i)$, namely the synoptic category and the 500 mb wind direction category. Therefore, each parameter can take sixteen possible values [one for each pair $c(i)$, $d(i)$] estimated by means of least-squares applied to the recorded meteorological series. Moreover, the average η of the residual term is made to depend on $c(i)$, $d(i)$.

Naturally, it can still be charged after this type of justification that model (1) is largely arbitrary. However, the structure (1) has, *a posteriori,* turned out to be the best among similar aggregate models of the physical phenomena. In other words, it is quite acceptable to assume that $v_s(i)$ is driven by variables like $v_s(i-1)$, $c(i)$, $d(i)$, $v_h(i)$, $v_h(i-1)$, $g(i)$ and $g(i-1)$ (see also the previous section), i.e., these variables contain most of the information required for determining $v_s(i)$. However, the specific role to be assigned to each variable, namely the most suitable form of the stochastic model, remains largely arbitrary and can be established in practice only through a trial and error procedure.

As for the term $\epsilon(i)$, the whiteness of the noise has been validated, *a posteriori,* through analysis of the cumulative periodogram of the residual (Box and Jenkins, 1970).

In view of noise whiteness, a real-time predictor of surface daily wind speed can be derived from model (1) simply by (see again Box and Jenkins, 1970):

1) Setting $\epsilon(i)$ to its mean value (zero).

2) Introducing for $v_h(i)$ and $g(i)$ their forecasts $\hat{v}_h$ and $\hat{g}$ obtained from the 500 mb 24 h predicted map at 0000 GMT of the ith day.

3) Using the set of parameters α, β, γ and η corresponding to the pair $c(i)$, $d(i)$ obtained from the same 500 mb 24 h predicted map.

In conclusion, the surface wind predictor takes the form

$$\hat{v}_s = \alpha v_s + \beta(\hat{v}_h - v_h) + \gamma(\hat{g} - g) + \eta. \quad (2)$$

4. Application of the model to Milan and Turbigo

Predictor (2) has been tested on two relatively close (~30 km) sites in the Po Valley: the city center of Milan and the Turbigo power plant (Fig. 1). The orographic effect is therefore quite similar for the two sites, both of which are located in the vicinity of the central Alps and far enough from the western Alps to make the western barrier effect scarcely relevant.

The meteorological time series considered for the two model applications cover the following winter periods:

- For Milan: January–March 1976, November 1976–March 1977, November 1979–March 1980, and November 1980–March 1981.
- For Turbigo: November 1977–March 1978, November 1978–March 1979, November 1979–March 1980, and November–December 1980.

In particular, the daily surface wind speed has been derived from 30-minute averages recorded at 40 m (Milan) and 20 m (Turbigo) above the ground. The values are rather low; in none of the sixteen meteorological classes $[c(i), d(i)]$ (see the previous sections) is the average speed greater than 2 m s^{-1}. This weak circulation is typical of the Po Valley during the winter season. Moreover, the average daily wind speed at Milan (1.7 m s^{-1}) is slightly higher than at Turbigo (1.2 m s^{-1}) due to both the different height of the anemometer, and to the city heat-island effect which induces a circulation cell during periods of relative calm (see Munn, 1966). On the contrary, the Milan standard deviation (0.72 m s^{-1}) is lower than Turbigo (1.31 m s^{-1}) due to the damping effect of city buildings on moderate and strong winds.

As recalled above, the parameters α, β, γ and η of model (1) have been estimated by means of standard least-squares applied to the abovementioned meteorological series. Their values are shown in Table 2 for both sites. The coefficients having low reliability (after a Student's test at the 5% significance level) have been marked by asterisks. An analysis of the table leads to the following conclusions:

- The signs of the coefficients turn out as expected by physical intuition in the large majority of the 16 cases. In particular, the coefficient of the 500 mb wind speed variation is nearly always positive, while the coefficient of the 500 mb geopotential variation is negative in most cases.
- The (absolute) values of the coefficients of the stationary component are generally larger in percentage for synoptic categories III and IV (anticyclonic), while the coefficients of the evolutionary component are generally larger for synoptic categories I and II (cyclonic). This result is also in agreement with physical intuition. In fact, in anticyclonic situations there is nearly always a subsidence inversion on the Po Valley, which "decouples" the 500 mb and surface circulations. On the contrary, in cyclonic situations upper-level and surface meteorological events are much more closely correlated, and hence the 500 mb

TABLE 2. Numerical coefficients of model (1).

	Milan				Turbigo			
c-d	α	β	γ (s^{-1})	η (m s^{-1})	α	β	γ (s^{-1})	η (m s^{-1})
I-1	0.450	0.146	−0.003	1.306	0.801	−0.104	−0.005	0.092
I-2	0.504	−0.018	0.0	0.830	0.322	0.104	−0.003	1.576
I-3	0.408	−0.048	−0.002	1.013	0.506	−0.012	0.0	0.694
I-4	1.146*	0.192*	−0.009*	−0.541*	0.811*	0.050*	−0.003*	0.384*
II-1	0.616	0.014	−0.002	0.869	1.185	0.016	−0.004	0.099
II-2	0.427	0.110	−0.017	1.892	0.617	0.150	0.0	0.524
II-3	0.420	0.0	0.001	0.876	1.145	0.016	0.0	−0.039
II-4	0.680	−0.044	−0.005	0.637	0.501	−0.044	−0.002	0.238
III-1	0.516	0.006	−0.003	0.888	0.545	−0.080	−0.006	0.763
III-2	0.850*	0.084*	−0.003*	0.223*	0.516*	0.054*	−0.004*	0.232*
III-3	0.592	0.028	−0.001	0.564	0.564	0.024	0.0	0.253
III-4	0.550	0.014	−0.001	0.580	0.479	0.012	0.003	0.305
IV-1	0.774	0.022	−0.001	0.324	0.231	0.026	0.0	0.548
IV-2	0.472*	0.078*	0.005*	0.635*	0.278	0.046	0.0	0.458
IV-3	0.884*	0.034*	−0.002*	0.061*	0.0*	0.0*	0.005*	0.107*
IV-4	0.559	−0.030	0.0	0.332	0.229	0.004	0.0	0.301

* Values with low reliability.

TABLE 3. Forecast performance indices for Milan and Turbigo.

	ρ	$\sigma_\epsilon^2/\sigma^2$
Milan:		
Persistent	0.54	0.92
AR (3)	0.57	0.50
Predictor (2)	0.72	0.46
Turbigo:		
Persistent	0.48	1.04
AR (5)	0.53	0.90
Predictor (2)	0.70	0.52

variables are given a more relevant role in the explanation of the surface wind speed by the stochastic model.

The "reasonable" parameter estimates of Table 2 are confirmed by the real-time forecast performance of the model. Such performance has been evaluated by two indices (Table 3):

1) The correlation ρ between forecast and observed surface wind speed.

2) The percentage of unpredicted variance $\sigma_\epsilon^2/\sigma^2$, the ratio between the variance of the forecast error (equal to $\epsilon(i)$ in this case) and the variance of the forecast variable.

For completeness, Table 3 (first and second rows) shows the performance indices for the persistent predictor [$\hat{v}_s(i) = v_s(i-1)$] and for the best purely autoregressive (of order p) predictor

$$\mathrm{AR}(p)\text{: } \hat{v}_s(i) = \phi_1 v_s(i-1) + \phi_2 v_s(i-2) + \cdots + \phi_p v_s(i-p) + \omega.$$

For both sites, the resulting ρ and $\sigma_\epsilon^2/\sigma^2$ are significantly worse than those supplied by predictor (2). Therefore, the comparison demonstrates the value of the exogenous meteorological inputs (500 mb wind speed and direction, synoptic circulation class, 500 mb geopotential), for a reliable forecast of surface site wind speed.

Table 4 shows the results for each of the 16 classes in more detail, giving the number of cases N, the principal statistics (mean value μ and standard deviation σ), and the two performance indices defined above, both for the predictor (2) and for the following simpler predictor with the stationary component only:

$$\hat{v}_s = \delta v_s + \theta, \tag{3}$$

where the parameters δ and θ assume one of the 16 possible pairs of values, according to the forecast of the synoptic and 500 mb wind direction categories.

Looking at Table 4, the following remarks can be made, neglecting the results (marked by an asterisk) of the classes whose data set size was too small (less than 15). As for predictor (2), Table 4 shows that the best performance ($\rho > 0.65$ and $\sigma_\epsilon^2/\sigma^2 < 0.65$ for both Milan and Turbigo) has been obtained with pairs I-1, IV-1 and III-4. In particular, IV-1 and I-1 correspond to calm conditions and northerly dry wind (foehn), respectively, and represent the most typical anticyclonic and cyclonic situations. These equally satisfactory results in two contrasting situations are explained by the distinction in model structure

TABLE 4. Statistical and performance indices of the predictors for every meteorological data class.

	Milan							Turbigo						
				Predictor (2)		Predictor (3)					Predictor (2)		Predictor (3)	
c-d	N	μ	σ	ρ	$\sigma_\epsilon^2/\sigma^2$	ρ	$\sigma_\epsilon^2/\sigma^2$	N	μ	σ	ρ	$\sigma_\epsilon^2/\sigma^2$	ρ	$\sigma_\epsilon^2/\sigma^2$
I-1	27	2.29	1.40	0.67	0.62	0.52	0.72	25	1.30	1.10	0.70	0.59	0.44	0.81
I-2	21	1.70	0.80	0.59	0.56	0.59	0.65	18	2.18	1.54	0.41	1.00	0.36	0.87
I-3	36	1.80	0.73	0.62	0.66	0.49	0.76	36	1.34	1.11	0.61	0.69	0.61	0.63
I-4	6	1.73	0.89	0.83*	0.77*	0.66*	0.56*	8	1.78	0.99	0.84*	0.50*	0.81*	0.34*
II-1	59	1.86	0.84	0.51	0.77	0.45	0.80	51	1.63	1.68	0.71	0.52	0.69	0.52
II-2	10	2.51	1.94	0.76*	0.59*	0.26*	0.93*	18	1.05	1.31	0.51	0.83	0.36	0.87
II-3	45	1.53	0.61	0.63	0.62	0.60	0.64	48	1.07	1.16	0.91	0.18	0.91	0.17
II-4	11	1.91	0.93	0.92*	0.21*	0.67*	0.55*	12	0.78	0.42	0.89*	0.30*	0.77*	0.41*
III-1	108	1.84	0.84	0.63	0.58	0.57	0.68	85	1.33	1.77	0.58	0.67	0.43	0.82
III-2	11	1.28	0.71	0.75*	0.46*	0.77*	0.41*	9	0.73	0.43	0.82*	0.52*	0.50*	0.75*
III-3	32	1.49	0.51	0.51	0.81	0.45	0.80	19	0.78	0.68	0.83	0.35	0.81	0.34
III-4	29	1.46	0.58	0.67	0.59	0.64	0.59	15	0.71	0.47	0.68	0.62	0.56	0.69
IV-1	84	1.52	0.73	0.76	0.42	0.74	0.45	49	0.85	0.73	0.67	0.59	0.64	0.59
IV-2	9	1.43	0.59	0.85*	0.44*	0.67*	0.55*	10	0.50	0.32	0.75*	0.66*	0.55*	0.70*
IV-3	6	1.18	0.75	0.99*	0.00*	0.98*	0.04*	5	0.44	0.32	0.56*	0.90*	0.46*	0.79*
IV-4	12	0.98	0.39	0.70*	0.62*	0.65*	0.58*	10	0.48	0.20	0.74*	0.64*	0.71*	0.50*

* Results of classes whose data set size was too small (<15).

(1) between the contributions of the stationary and the evolutionary components, which in the two situations (I-1 and IV-1) play different roles (see also the comment above on the values of the coefficient estimates).

Moreover, the comparison with predictor (3) performance demonstrates the actual role of the evolutionary component (generally the smaller one) in the right-hand side of the predictor (2). As was to be expected, the performance difference is higher in the classes corresponding to evolutionary situations. Hence, the good predictability of wind speed obtained for Turbigo with the pairs II-1, II-3 and III-3 can be ascribed to the fact that the stationary component has a higher percentage; in fact, the two predictors give quite similar results.

Nevertheless, one may wonder if the 16 stratifications of the data were really necessary. Hence, predictor (2) also has been applied in the following three different ways:

1) Stratifying the time series of data on the basis of the synoptic category (I, II, III, IV) only, getting four sets of parameters, α, β, γ, η, for every case study.

2) Stratifying the time series by the wind direction category (1, 2, 3, 4) only, again obtaining four sets of parameters.

3) Without any stratification, i.e., estimating one set of parameters on the basis of the whole set of data.

Table 5 shows the results of all these analyses as follows:

• The first four rows show, for a fixed synoptic category, the number of cases corresponding to the different wind direction categories, the total and the two performance indices of predictor (2) when used as in 1) above.

• The first four columns show in a similar way the same information for every wind direction category

TABLE 5a. Number of cases of each of the 16 classes, of each of the synoptic categories (rows I, II, III, IV), of each of the wind direction categories (columns 1, 2, 3, 4), and totals; performance indices of predictor (2) computed for each synoptic category, for each wind direction category, and for the total series for Milan.

	1	2	3	4	Row total	ρ	σ_e^2/σ^2
I	27	21	36	6	90	0.53	0.73
II	59	10	45	11	125	0.50	0.76
III	108	11	32	29	180	0.64	0.60
IV	84	9	6	12	111	0.77	0.41
Column total	278	51	119	58	506		
ρ	0.64	0.51	0.58	0.75		0.60	
σ_e^2/σ^2	0.60	0.78	0.68	0.47			0.62

TABLE 5b. As in Table 5a but for Turbigo.

	1	2	3	4	Row total	ρ	σ_e^2/σ^2
I	25	18	36	8	87	0.52	0.75
II	51	18	48	12	129	0.71	0.51
III	85	9	19	15	128	0.57	0.70
IV	49	10	5	10	74	0.66	0.58
Column total	210	55	108	45	418		
ρ	0.53	0.55	0.75	0.77		0.56	
σ_e^2/σ^2	0.72	0.73	0.45	0.43			0.69

and for predictor (2) performance when used as in 2) above.

• Finally, on the right lower diagonal, the total number of data examined and the two performance indices of predictor (2) when used as in 3) above are reported.

First, comparing these last results with the third row of Table 3, it is evident that the stratification significantly improves the forecast. Moreover, comparing Tables 4 and 5, the following observations can be made:

• For synoptic categories I, II and III, the further splitting of data looking at the different wind directions seems worthwhile for forecast purposes.

• The same is true for wind direction categories 1, 2 and 3, with the exception of class III3.

• The comparison between the Tables 4 and 5 is not significant for synoptic category IV and wind direction category 4, due to the scarcity of data in the respective eight classes.

A final consideration concerns the reliability of the performance figures given in Table 2, because the forecast test has been carried out on the same set of data used for estimating the model coefficients by least-squares. In many of the 16 classes the length of the data set did not allow for subdivision into two subsets, one for parameter estimation and one for forecast testing. However, for the few classes characterized by a relatively abundant data set, this subdivision did not significantly decrease the forecast performance.

5. Concluding remarks

A stochastic predictor of surface site wind speed for air quality control purposes has been described, and the impact of both local and synoptic meteorological information on the reliability of the forecast has been demonstrated.

This work can be extended mainly in two directions:

1) By developing similar predictors for other local meteorological variables (wind direction, vertical temperature profile, etc.) which are relevant factors in pollutant dispersion.

2) By developing more complex stochastic predictors of wind speed, in order to obtain a spatially more detailed forecast. For instance, a predictor able to supply simultaneously the wind speed forecast at one or two levels besides the surface would give useful information for pollutant dispersion models in many situations. More specifically, for some types of pollution "episodes," the surface (20 or 30 m, say) value generally provides sufficient information about wind speed. This is true, for example, in episodes resulting from plume breakdown by strong wind in neutral conditions, when the vertical wind profile more or less follows a regular pattern (e.g., a power law). On the contrary, in other episode situations the wind profile can be more complex and cannot be easily extrapolated from the surface value. This is true, for example, for summer fumigations around tall stacks in the Po Valley, which are generally characterized by breeze layers of different depth. In this case, the vertical profile follows the power law only above 500 or 600 m, and at least one or two intermediate values would be required in order to account correctly for the effect of wind on the plume. This would require the development of a two- or three-variate stochastic model. Obviously, it would not be reasonable to consider stochastic models more complex than that, because of the nature of these mathematical representations themselves.

Acknowledgments. The authors wish to thank Prof. G. Fronza of Dipartimento di Elettronica, Politecnico (Milan) and Dr. P. Bacci of ENEL, CRTN (Milan) for their collaboration and useful criticism.

REFERENCES

Bacci, G., and G. Finzi, 1980: A statistical predictor of surface wind speed in Milan city. *Proc. WMO-RA VI Meeting on Forecasting of Conditions Leading to High Levels of Air Pollution Occurrence,* Leningrad, WMO.

Bacci, P., P. Bolzern and G. Fronza, 1981: A stochastic predictor of air pollution based on short-term meteorological forecasts. *J. Appl. Meteor.,* **20,** 121–129.

Benarie, M. M., 1980: *Urban Air Pollution Modelling.* MacMillan, 405 pp.

Bengtsson, L., 1976: *Proc. WMO Symp. on the Interpretation of Broad-Scale NWP Products for Local Forecasting Purposes.* WMO No. 450, 250 pp.

Bolzern, P., and G. Fronza, 1982: Cost-effectiveness analysis of real-time control of SO_2 emission from a power plant. *J. Environ. Management,* **14,** 253–263.

Bonivento, C., G. Fronza and A. Tonielli, 1980: Real time prediction of local wind by means of stochastic models. *Proc. 14th Int. Colloquium of the IRCA,* M. Benarie, Ed., Elsevier, 105–108.

Box, G. E. P., and G. M. Jenkins, 1970: *Time Series Analysis, Forecasting and Control.* Holden-Day, 544 pp.

Finzi, G., and G. Tebaldi, 1982: A mathematical model for air pollution forecast and alarm in an urban area. *Atmos. Environ.,* **16,** 2055–2059.

Gandino, C., 1976: The influence of the Alps on the diurnal winds. *Riv. Ital. Geofis.,* **3,** 150–152.

Godfrey, R. A., 1982: An application of model output statistics to the development of a local wind regime forecast procedure. *J. Appl. Meteor.,* **21,** 1786–1791.

Holton, J. R., 1979: *An Introduction to Dynamic Meteorology.* Academic Press, 391 pp.

Kau, W. S., N. Leeth and S. K. Kao, 1982: A statistical model for wind prediction at a mountain and valley station near Anderson Creek, California. *J. Appl. Meteor.,* **21,** 18–21.

McWilliams, B., and D. Sprevak, 1982: The simulation of hourly wind speed and direction. *Math. Comput. Simulation,* **24,** 54–59.

Munn, R. E., 1966: *Descriptive Micrometeorology.* Academic Press, 245 pp.

Roldan-Cañas, J., A. Garcia-Guzman and A. Losada-Villasante, 1982: A stochastic model for wind occurrence. *J. Appl. Meteor.,* **21,** 740–744.

Soeda, T., and S. Omatu, 1982: Real-time control of emissions in Japanese cities. *Proc. IIASA Workshop on Mathematical Models for Planning and Controlling Air Quality,* G. Fronza and P. Melli, Eds., Pergamon, 215–231.