Time Series Techniques for Wind Energy Conversion Site Surveys

T.N. Goh and G.K. Nathan

Faculty of Engineering
National University of Singapore
Singapore 0511

ABSTRACT

Installation of a wind energy conversion system must be preceded by surveys in which wind speed data at either the potential site of the system or nearby stations are analysed. A suitable approach to such analysis is that of statistical time series modelling, which not only is capable of capturing wind persistence characteristics, but also expresses such characteristics in parametric forms that can be used for comparisons involving different potential sites or different time periods at the same site. Some suitable time series techniques and their applications are described in this paper.

INTRODUCTION

Research, development and demonstration efforts in renewable energy entail frequent collection, analysis and transmission of information. Where information takes the form of numerical data, it is necessary to employ appropriate data analysis and representation techniques so as to ensure efficient information utilization. Many investigators of renewable energy devices and systems tend to approach the subject largely from a mechanistic point of view using deductive reasoning with engineering principles; their work would be much facilitated if, in processing large amounts of data, advantage is also taken of the inductive process of statistical analysis and modelling. Some applications of statistical techniques to wind energy conversion site studies are described in this paper and it is hoped that these techniques can be more widely adopted to complement existing procedures.

WIND ENERGY CONVERSION SITE INVESTIGATIONS

The technical feasibility of conversion of wind energy into electricity has generally been established. However, decision on actual installation of a wind energy conversion system cannot be made before operational feasibility has been carefully studied. One main factor in a feasibility analysis is wind energy availability in terms of magnitude and variability, which have direct bearings on the amount and stability of energy supply and, consequently, the economic viability of the system. Thus in site surveys for wind energy conversion systems, apart from considerations related to the locations and needs of the target service areas, both long and short term characteristics of wind dynamics constitute the predominant factors for investigation. The information needs in such studies should first be met before consideration of engineering design requirements.

Conventional approaches to wind analysis[1-4] tend to centre around relationships between wind speed and ground characteristics such as surface roughness, or between wind profile and

topographic configurations. Aerodynamic analyses of wind flow over terrains and objects then provide predictive equations to be verified by observations, simulations, or experiments. While such studies are useful in describing what to expect for a given set of environmental characteristics, they are limited by difficulties in modelling actual three dimensional terrain characteristics and in reflecting the probabilistic nature of wind speed variations within a given space and a given time span. In recent years, a common approach is one of empirical model building, i.e. using actual data from the past to construct mathematical expressions capable of describing enough essential data characteristics for prediction purposes.[5-8]

Empirical modelling may be preferable to mechanistic modelling in that it is based on actual data; the physical phenomena of a given environment are captured once a statistically adequate model is attained and no search for, or reference to, physical laws (which in turn are models of a more general nature) need be made. Since wind speed measurements are routine at meteorological stations, availability of data for preliminary analysis usually poses no problems. From the point of view of information utilization, wind measurements are useful in that they characterize past wind behaviour, and are necessary to the extent that such behaviour can be generalized and extrapolated in space or time through appropriate mathematical models: this is depicted in Fig. 1, where X denotes known information from the past and Y the projected information, and 1, 2 are locations where information is available and A, B, C are locations where information is lacking. The gaps between known and unknown are to be bridged by the models.
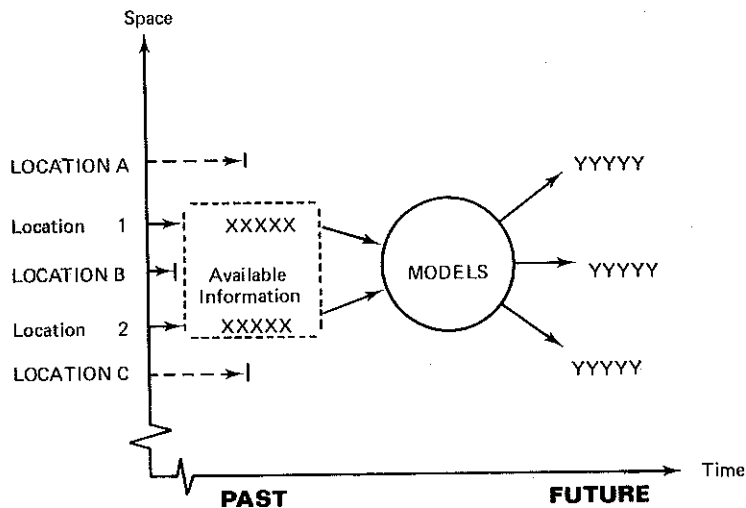

Fig. 1 Information analysis and utilization

WIND SPEED CHARACTERIZATION

Most wind data models may be categorised into two types: descriptive statistics and probabilistic distribution models. The first category are graphs or tables such as speed histograms, speed or power duration curves and speed-time plots, whereas the second takes the form of concise analytic expressions such as the Weibull distribution. Most wind surveys contain both representa-

tions.[9],[10] Such representations, while in themselves useful summaries of data values, consist of transformed information and do not contain the real time element; they are unable to express wind speed dynamics explicitly in the time domain, a factor essential to operational analysis of wind power systems. Furthermore, these representations, apart from lending themselves to visual assessments, may not be adequate if indices or parameters are called for in comparative studies of data sets from, for example, two sites or two periods of time. In fact, the need for a more appropriate model form arises in common cases where a potential site is so located that wind data are unavailable or insufficient: in such situations, techniques of data interpolation and extrapolation in space, in addition to forecasting in time, are necessary.

ANALYSIS OF TIME SERIES

Some of the deficiencies in conventional models discussed above may be overcome through a different approach based on methodologies of discrete time series analysis. To the extent that most available wind measurements are in the form of hourly averaged speeds, and that successive wind speed values are as a rule highly dependent, techniques of stochastic time series analysis can be directly employed to exploit the nature of data dependence, expressing dynamics of wind persistence and patterns of variation through parametric models in the time domain.[11-13] The collection of n successive hourly speeds $[x_t]$, $t = 1, 2, \ldots, n$, constitutes the primary data for analysis of wind characteristics. The most important intrinsic data property is expressed through the *autocorrelation function* $[\rho_k]$, defined by

$$\rho_k = \frac{E\left[(x_t - \mu)\ (x_{t+k} - \mu)\right]}{\sigma^2}\ , \qquad k = 0, 1, \ldots, \qquad (1)$$

which is a dimensionless measure of the degree of association of any pair of wind speeds that are k hours apart, where μ, σ^2 are the mean and variance of x_t respectively, and E denotes the expected (or mean) value of the mathematical expression. ρ_k may be estimated by the *sample autocorrelation function* $[r_k]$:

$$r_k = \frac{1}{n-k} \sum_{t=1}^{n-k} (x_t - \bar{x})\,(x_{t+k} - \bar{x})/ \frac{1}{n} \sum_{t=1}^{n} (x_t - \bar{x})^2\ , \quad k = 1, 2, \ldots, n, \qquad (2)$$

$$x = \frac{1}{n} \sum_{t=1}^{n} x_t. \qquad (3)$$

For a given value of k, the *sample autocorrelation coefficient* r_k falls between -1 and $+1$; the greater its magnitude, the stronger is the dependence of wind speeds at k hours apart. Since r_k is only an estimate of ρ_k, its statistical significance may be ascertained through an approximate test: if the magnitude of r_k is greater than $2n^{-\frac{1}{2}}$, it may be considered statistically significant (at 5% level of significance), and vice versa. A typical example of sample autocorrelation function, obtained from wind measurements recorded in a year at a meteorological station "R" in Singapore, is shown in Fig. 2, which exhibits an exponential decay pattern and a seasonal (i.e. recurring) fluctuation, both of which are statistically significant: the former reflecting wind persistence and the latter the recurrence of diurnal variations (r_k peaking at $k = 24$ and other multiples of 24).
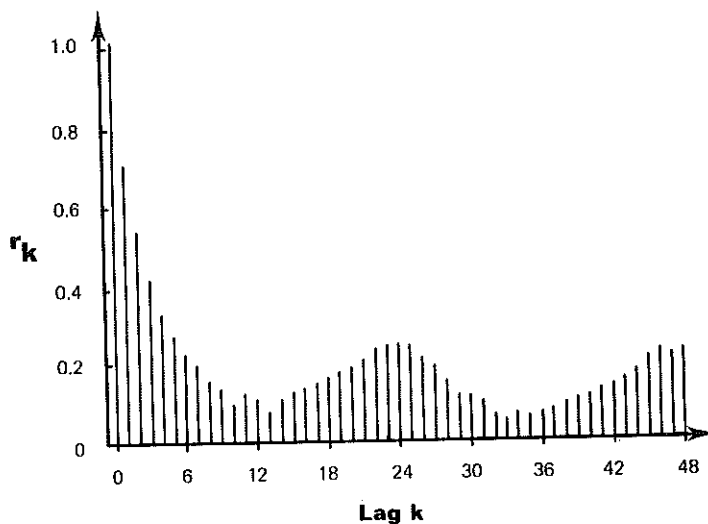
Fig. 2 Sample autocorrelation function from station "R"

Autocorrelation coefficients can either be used for comparison of wind characteristics at different stations or serve as a guide for further data modelling efforts.

TIME SERIES MODELLING

The dynamic nature of wind data can next be quantified through time series models, of which the formulation due to Box and Jenkins[13, 14] would be the most appropriate. If a series of measurements is such that the value at a given time is influenced by the value one time interval ago, then this characteristic is described by

$$x_t = \phi x_{t-1} + e_t , \tag{4}$$

where ϕ is a constant, and e_t is an error term whose values are independent of one another and are described by a normal distribution with zero mean and constant variance σ_e^2. Such data dependence would also be exhibited through the sample autocorrelation function, which shows an approximately exponentially decaying value of r_k (Fig. 3). This property is exploited in practice for the *identification* of the model form, i.e. Eq. (4). More complex patterns of the sample autocorrelation function would suggest correspondingly more complicated model forms: for example, sometimes it is possible that the error terms themselves are not independent, in which case x_t may be represented, for example, by the model

$$x_t = \phi x_{t-1} + e_t - \theta e_{t-1} , \tag{5}$$

where θ is a constant. The expression indicates that the value measured at a given time t is a function of the previous measurement, subject to a random fluctuation consisting of unpredictable factors which exert themselves at time t and t-1. Another autocorrelation function is shown in Fig. 4, which contrasts markedly with that shown in Fig. 3.
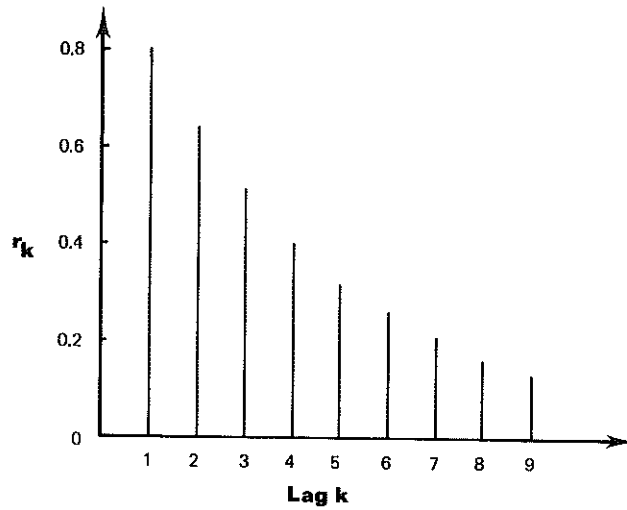
Fig. 3 Theoretical autocorrelation function of model
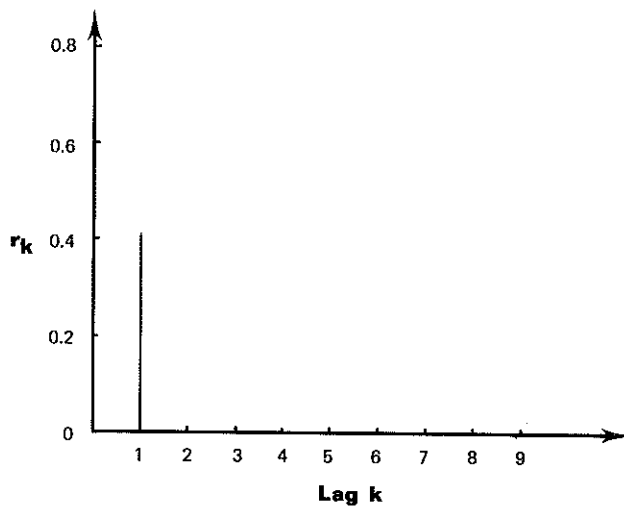$(1 - 0.8\ B)\ x_t = e_t$

Fig. 4 Theoretical autocorrelation function of model
$x_t = (1 + 0.5\ B)\ e_t$

Equations (4) and (5) may also be rewritten, respectively,

$$(1 - \phi B)\, x_t = e_t \tag{6}$$

and

$$(1 - \phi B)\, x_t = (1 - \theta B)\, e_t, \tag{7}$$

where B is a *backward shift operator* such that

$$Bx_t = x_{t-1}, \qquad B(B)x_t = B^2 x_t = x_{t-2} \tag{8}$$

or, in general,

$$B^m x_t = Bx_{t-m}, \qquad m = 1, 2, \ldots \tag{9}$$

A model with higher order terms may take the form

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)x_t = (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)e_t \tag{10}$$

where p is the number of *autoregressive* parameters $\phi_1, \phi_2, \ldots, \phi_p$, and q is the number of *moving-average* parameters $\theta_1, \theta_2, \ldots, \theta_q$. Appropriate modifications to the models are usually made for data exhibiting *seasonality* (i.e. recurring) characteristics, the most general form being

$$(1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p)(1 - \lambda_1 B^S - \lambda_2 B^{2S} S - \ldots - \lambda_p B^{PS})x_t$$

$$= (1 - \theta_1 B - \theta_2 B^2 - \ldots - \theta_q B^q)(1 - \gamma_1 B^S - \gamma_2 B^{2S} - \ldots - \gamma_Q B^{QS})e_t, \tag{11}$$

where S is a measure of the seasonality of data characteristics, and P, Q are orders of the seasonal autoregressive and moving average parameters respectively. Thus hourly wind speeds would invariably have a seasonality of 24 due to diurnal variations.

Once the model form is established through p, q, P, Q, S, the parameter, values are obtained by means of *least squares estimation,* for which special computer softwares are available (e.g. the PACK system), followed by model *checking* procedures which ensure that $[e_t]$ is indeed normally and independently distributed. Based on the resulting model, *forecasting* equations for future values of x can be readily derived if necessary. Further details on identification, estimation, checking, and forecasting and the use of computer programmes can be found in standard texts.[14, 15]

PRACTICAL INTERPRETATION OF MODELS

Time series modelling of hourly wind speed measurements has been used in exploratory wind energy conversion site studies in and around the island of Singapore, situated at the tip of Malay Peninsula.[10, 16, 17] For example, the general model form found to be adequate in describing the data from two stations "R" and "F" (Fig. 5) is of the autoregressive type

$$(1 - \phi_1 B - \phi_{24} B^{24} - \phi_{25} B^{25})x_t = e_t, \tag{12}$$

i.e. the average wind speed at any hour is statistically dependent on its value an hour ago, its values a day ago at the same hour and an hour ago. Any unpredictable short-term random deviations from the functional relationship are reflected through the $[e_t]$ series. It is aso possible to effect real-time forecasting through the equation

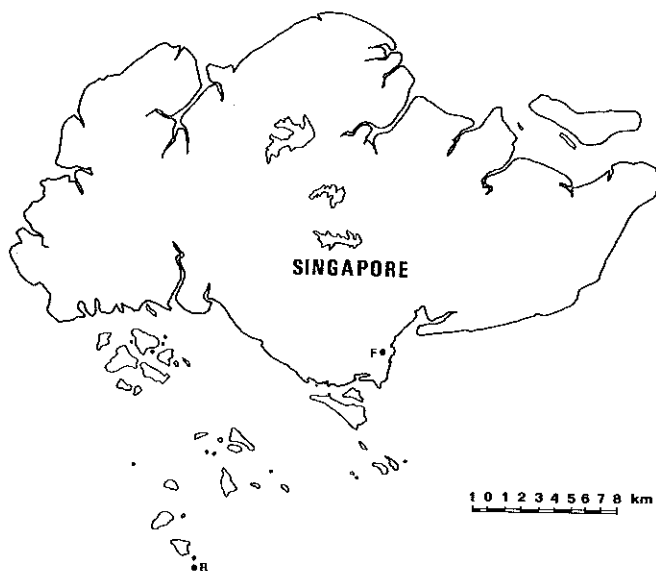$$\hat{x}_t = \phi_1 x_{t-1} + \phi_{24} x_{t-24} + \phi_{25} x_{t-25}, \tag{13}$$

Fig. 5 Location of stations "R" and "F"

since the expected value of e_t for the future is zero.

In general, time series models for different data sets may be compared in respect of model form and model parameter values. For regions with distinct climatic changes in a year, annual data may be divided into suitable sections and modelled separately for such comparisons, thus providing insights into changes in the dynamics of winds from one time period to another. In longer terms, such comparisons can be extremely useful: for example, for the two Singapore stations, parameter estimation was carried out for data from two years, 1973 and 1979, with the numerical results summarized in Table 1. It is seen immediately that wind speeds at station R are described by essentially the same time series model even at periods six years apart. On the other hand, at station "F", considerable changes in the model are apparent. Since such changes cannot be ascribed to climatic changes, which if they existed would have affected the model for station "R" as well, they are explained by the fact that high-rise structures erected between

Table 1
Wind speed model parameters at two locations

Parameter	Station R			Station F		
	1973	1979	% change	1973	1979	% change
ϕ_1	0.7374	0.7071	−4.1	−0.7659	−0.7192	− 6.1
ϕ_{24}	0.0601	0.0615	−2.3	−0.1219	−0.1882	− 54.4
ϕ_{25}	*	*	—	−0.0249	−0.0545	−118.9

1973 and 1979 in the proximity of station "F" have changed the wind patterns around this station. It follows, therefore, that wind characteristics reflected by past data from the station should not be used in future studies. Such evaluations are particularly important in assessing the quality of past data that were not collected under the control of present investigators. An important advantage of studies via time series techniques is that model changes are not caused by instrumental drifts or other sources of systematic errors, which would only affect static parameters such as averages and speed duration measurements.

STUDY OF MORE THAN ONE SITE

Time series modelling techniques may also be used to facilitate analysis of inter-site wind dynamics. The identification of a stochastic-dynamic model relating simulataneously measured wind speed time series from two sites calls for more involved procedures,[14] but the result offers ease of physical interpretation unavailable from other conventional methods such as the distance-correlation analysis.[18] The general model form which describes the interdependence in time between two series $[x_{1,t}]$ and $[x_{2,t}]$ is

$$(1 - \delta_1 B - \delta_2 B^2 - \ldots - \delta_r B^r) x_{2,t}$$

$$= (\omega_0 - \omega_1 B - \omega_2 B^2 - \ldots - \omega_s B^s) B^b x_{1,t} + E_t, \tag{14}$$

where $x_{1,t}$ and $x_{2,t}$ are values measured simultaneously at time t at stations or sites 1 and 2 respectively, r, s, b are *orders* of the model to be identified, δ_1, δ_2, . . . and ω_0, ω_1, . . . are parameters to be estimated from the data, and $[E_t]$ is an error series. The relationship between $[x_{1,t}]$ and $[x_{2,t}]$ can also be expressed as

$$x_{2,t} = (\gamma_0 + \gamma_1 B + \gamma_2 B^2 + \ldots) x_{1,t} + e_t \tag{15}$$

in which $[\gamma_i]$, $i = 0, 1, 2, \ldots$, constitutes the equivalent of an *impulse response function* so that $[x_{2,t}]$ is now a function of the data $[x_{1,t}]$ from a reference site. This is a convenient expression because the value of $x_{2,t}$ can be forecast through

$$\hat{x}_{2,t} = \gamma_0 x_{1,t} + \gamma_1 x_{1,t-1} + \gamma_2 x_{2,t-1} + \ldots \tag{16}$$

so that if $[x_{1,t}]$ is obtained from an established measuring station, its values can be used to generate expected values of $[x_{2,t}]$ at the second site. The farther apart the two sites are, the weaker would be the correlation between $[x_{1,t}]$ and $[x_{2,t}]$, but such data extrapolation techniques are useful in exploratory site studies. For a given pair of sites, the model form and parameter values of Eq. (14) serve as indicators whose variations from one data measurement period to another would reveal and quantify any changes in data characteristics.

CONCLUSIONS

Study of renewable energy systems frequently entails collection of large amounts of environmental data. Statistical modelling permits reduction of raw data to parametric forms which give concise descriptions of data characteristics. To express such characteristics in the time domain, which is necessary for purposes of real-time analysis, discrete time series modelling techniques

have been found to be most suitable. Although the mathematical foundation of such techniques could be quite involved, it would not impose undue difficulties in actual applications, thanks to the recent availability of computer software packages. Inasmuch as the most common data recording format used by established wind stations is time series of hourly measurements, the modelling approach described in this paper can be readily applied in many situations. Insights into relative dynamics of wind speeds at different time periods or different sites, either at the site survey stage or after the erection of the wind energy conversion systems, can be obtained almost on a routine basis once the basic procedures are set up; the various wind speed and environmental properties revealed through the models would eventually justify the modelling efforts required.

REFERENCES

1.   GOLDING, E.W. (1955), *The Generation of Electricity by Wind Power,* Spon, London.

2.   World Meteorological Organisation (1964), *Sites for Wind-Power Installations,* WMO Pub. No. 156, TP. 76, Geneva.

3.   PENNELL, W.T., W.R. BARCHET, D.L. ELLIOTT, L.L. WENDELL, and T.R. HIESTER (1980), Meteorological Aspects of Wind Energy: Assessing the Resource and Selecting the Sites, *Journal of Industrial Aerodynamics,* Vol. 5, pp. 223-246.

4.   BHUMRALKAR, C.M., R.L. MANCUSO and F.L. LUDWIG (1980), A Practical and Economic Method for Estimating Wind Characteristics at Potential Wind Conversion Sites, *Solar Energy,* Vol. 25, pp. 55-65.

5.   CHERRY, N.J. (1980), Wind Energy Resource Survey Methodology, *Journal of Industrial Aerodynamics,* Vol. 5, pp. 247-280.

6.   COROTIS, R.B., A.B. SIGL, and J. KLEIN (1978), Probability Models of Wind Velocity Magnitude and Persistence, *Solar Energy,* Vol. 20, pp. 483-493.

7.   JUSTUS, C.G., W.R. HARGRAVES, A. MIKHAIL and D. GRABER (1978), Methods for Estimating Wind Speed Frequency Distributions, *Journal of Applied Meteorology,* Vol. 17, pp. 350-353.

8.   SWIFT-HOOK, D.T. (1979), Describing Wind Data, *Wind Engineering,* Vol. 3, pp. 167-186.

9.   EXELL, R.H.B., S. THAVAPALACHANDRAN, and P. MUKHIA (1981), *The Availability of Wind Energy in Thailand,* AIT Research Report No. 134, Renewable Energy Resources Information Centre, Bangkok.

10.   NATHAN, G.K. and T.N. GOH (1981), Wind Measurements in an Equatorial Region (Singapore), *Solar Energy,* Vol. 26, pp. 275-278.

11.   PANOFSKY, H.A., and G.W. BRIER (1965), *Some Applications of Statistics to Meteorology,* Pennsylvania State University, University Park, Pennsylvania.

12.   SNEYERS, R. (1975), *Sur l'Analyse Statistique des Series d'Observations,* WMO Pub. No. 415, Geneva.

13.   ANDERSON, O.D. (1976), Some Methods of Time Series Analysis, *Math. Scientist,* Vol. 1, pp. 27-41.

14.   BOX, G.E.P., and G.M. JENKINS (1976), *Time Series Analysis, Forecasting and Control* (revised edition), Holden-Day, San Francisco.

15. MAKRIDAKIS, S., and S.C. WHEELWRIGHT (1978), *Interactive Forecasting* (2nd ed.), Holden-Day, San Francisco.

16. GOH, T.N., and G.K. NATHAN (1979), A Statistical Methodology for Study of Wind Characteristics from a Close Array of Stations, *Wind Engineering,* Vol. 3, pp. 197-206.

17. GOH, T.N., and G.K. NATHAN (1981), A Comparative Study of Wind Characteristics from Different Time Periods, *Wind Engineering,* Vol. 5, pp. 60-65.

18. BALLESTER, M. (1963), Speculative Methods in Wind Surveying, *Proceedings of UN Conference on New Sources of Energy,* UN Pub. 63.1.41, Vol. 7, pp. 49-59.