

UTILIZATION OF MARKOV CHAIN TECHNIQUE TO GENERATE HIGH-DENSITY WIND SPEED DATA FOR THE COASTAL BELT OF PAKISTAN

Zeeshan Alam Nayyar*¹, Nayyer Alam Zaigham² and Abdul Qadeer¹

¹Renewable Energy Research Group, Department of Applied Physics, University of Karachi, Karachi, Pakistan

²Department of Geology, University of Karachi, Karachi, Pakistan

Abstract

In order to estimate the reliable wind potential of any region, the high density wind speed data of 10-minutes or an hour interval is necessary. In most part of the Pakistan, available wind speed data is usually consists of monthly averages. In such a case, the Markov chain method based on transition matrix approach is used to synthesize the hourly wind speed data from the available averaged monthly wind speed data. The coastal belt of Pakistan is considered as a favorable regime for the development of wind energy activities and mainly targeted for the energy estimation. The first-order Markov chain technique has been used to synthesize and generate the wind speed data for the height of 10m from the monthly averaged wind data into the daily wind speed data for the coastal regions of Pakistan. Furthermore, the trends of both the daily original airport wind speed data and the synthesized wind speed data have been analyzed and significant compatibility is observed.

Keywords: Wind energy, wind potential, Pakistan coast, Markov chain technique

INTRODUCTION

In general, the coastal regions are considered favorable belt for the reliable and consistent wind resources. Similarly, huge wind potential is expected associated with the Pakistan coastal belt considering its length and apparent wind trends (Zaigham & Nayyar, 2005; Nayyar, 2009). Wind-mills were installed for shallow groundwater pumping in the region (Shah, 1991), but this technique has not been adopted on large scale energy generation due to lack of technological knowledge and limited incentives to small farmer by the government. However, realizing the better prospects of utilizing wind energy, different government organizations attempted to utilize wind-mill technology but they lacked mutual coordination among the organizations and the communities due to some unknown reasons and the projects could not be successful at large.

With the start of the new millennium, the private and the public sectors realized importance to install wind-farm(s) to generate energy from the wind. The Pakistan Meteorological Department (PMD) has a network of regular meteorological stations all over the country. The wind data are being recorded at these stations for the basic meteorological purposes. Attempts are also being made to utilize this data for the preliminary estimations of wind potential along the coastal belt of Pakistan (Chaudri & Farooqi, 2000; Khan, 2001), but these attempts could not practically achieve the required results to develop an adoptable wind database for the power generation.

The available wind data from PMD is usually consisted of monthly averages and are not enough to calculate any type of reliable wind potential. Therefore, there is a need to get the high resolution wind speed data on hourly or daily basis by processing from the monthly wind speed averages. The first-order Markov chain technique provides a systematic approach to interpolate mathematically the high density wind data out of low density dataset (Shamshad *et al.*, 2005; Sahin & Sen, 2001; Ettoumi *et al.*, 2003).

This paper describes the utilization of the first-order Markov chain technique to synthesize daily wind speed data from the available monthly averages of wind data and to validate the authenticity of the generated daily data with the available airport daily wind dataset.

DATA COLLECTION

The data set for the present study consists of wind speeds for the period of January 1, 2005 to December 31, 2005 at a height of 10 meters as measured from ground level. Daily and monthly average wind speed data have been collected from the Weather Underground Inc. (WUI) for the airport locations namely Badin, Hyderabad for the coastal region of Sindh province and Pasni for the coastal region of Balochistan province (Table 1). The WUI provides the online meteorological data of more than 12,000 cities and airports worldwide including Pakistan (WUI, 2009). The WUI received the data from the airport sources and generated output dataset from the AVN computer forecast model run at the U.S. National Weather Service's National Center for Environmental Prediction. These dataset do fit reasonably well for a large number of

*Corresponding author: E-mail: zanayyer@uok.edu.pk

Table 1: Location and elevation of the WUI airports in Pakistan

Name of Station	Latitude °N	Longitude °E	Elevation a.m.s.l. (m)
Badin	24.6°	68.9°	10
Hyderabad	25.3°	68.4°	40
Pasni	25.3°	53.5°	4

locations, but are not so in some locations, particularly mountainous areas.

METHODOLOGY

Efforts have been made to generate the synthesized daily wind data from the monthly average value by using computer model based on *first-order Markov chain* and compare it with the original WUI wind data by computing and comparing their trend equations, standard deviation and coefficient of variance.

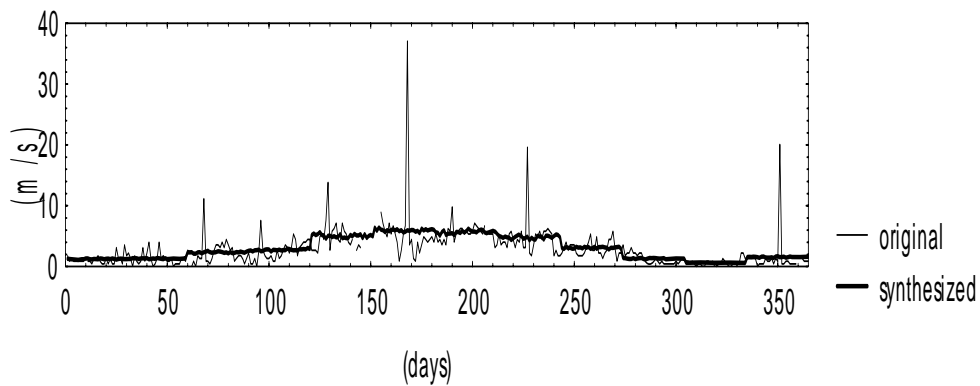
Markov chains' stochastic process facilitates

parameterization with the help of empirical estimation transition probabilities between discrete states in the observed systems (*Lun & Lam, 2000; Nfaoui et al., 2004*). In this study only the first order Markov chain technique has been used.

Let $X(t)$ be a stochastic process, possessing discrete states space $S=\{1,2,\dots,K\}$. According to Logofet & Lensnaya (2000), for a given sequence of time points $t_1 < t_2 < \dots < t_n$, the conditional probabilities should be as follows:

$$\Pr\{X(t_n) = i_n \mid X(t_1) = i_1, \dots, X(t_{n-1}) = i_{n-1}\}$$

(a) time-series plot of original and synthesized wind speed for Badin



(b) trends of original and synthesized wind speed for Badin

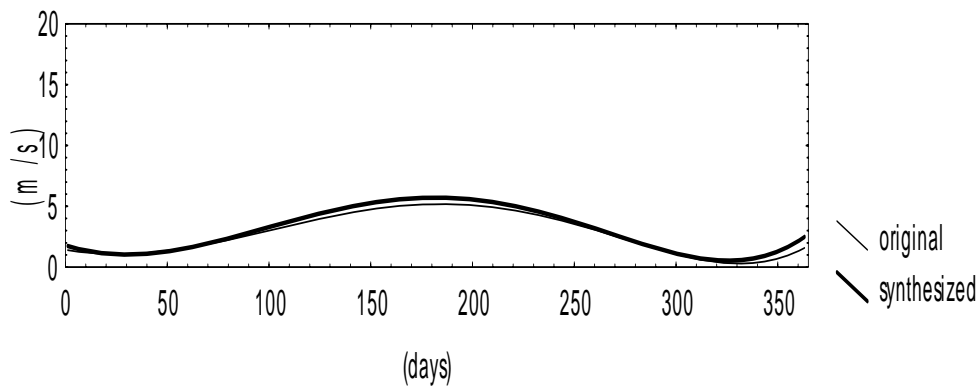


Fig. 1: Time-series and trend plots for Badin airport, showing graphical comparison between the original and the synthesized daily wind speed data and their trend-lines.

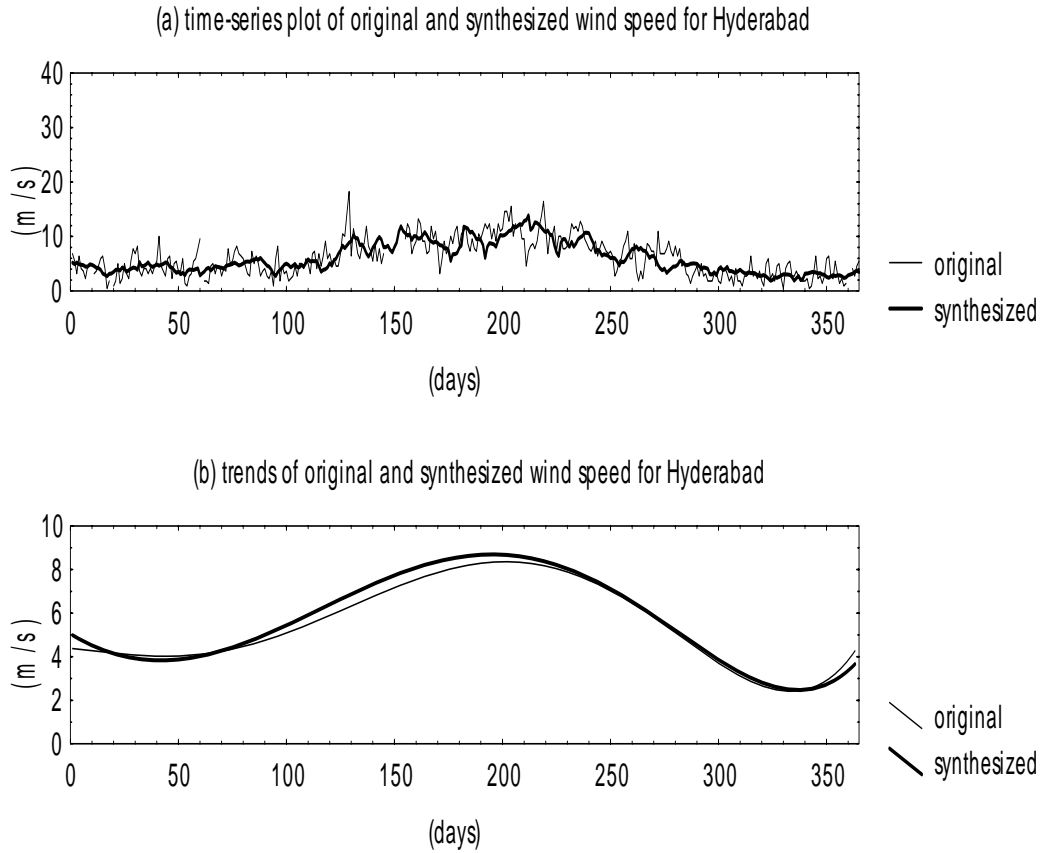


Fig. 2: Time-series and trend plots for Hyderabad airport, showing graphical comparison between the original and the synthesized daily wind speed data and their trend-lines.

$$= \Pr\{X(t_n) = i_n | X(t_{n-1}) = i_{n-1}\} \quad 1$$

The conditional probabilities $\Pr\{X(t) = j | X(s) = i\} = P_{ij}(s, t)$ generally are called transition probabilities of order $r=t-s$ from state i to state j for all indices $0 \leq s < t$, with $1 \leq i$ and $j \leq k$ as suggested by Heiko (2000). They are denoted as the transition matrix P . For k states, the first order transition matrix P has a size of $k \times k$ and takes the form as follows in Equation 2:

$$P = \begin{bmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,k} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,k} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ P_{k,1} & P_{k,2} & \cdots & P_{k,k} \end{bmatrix} \quad 2$$

If n_{ij} is the number of transitions from state i to state j in the sequence of speed data, the maximum likelihood estimates of the transition probabilities are:

$$p_{ij} = n_{ij} / \sum_j n_{ij} \quad 3$$

Where the transition probabilities of any state vary between 0 and 1

RESULTS

The synthesized wind data for Badin, Hyderabad and Pasni airports of the coastal regions of Pakistan have been generated for the period of January 1, 2005 to December 31, 2005 on daily basis by utilizing the monthly averaged wind data into the stochastic mechanism as described by the equation 1 to 3. The generated synthesized wind data have been plotted along with the original wind data and with their trends (Figures 1 to 3)

The trends have been represented in the polynomial equations 4 to 9 as follows:

$$T(\text{Badin,original}) = 1.693 - 0.052y + 0.001y^2 - 3.7 \times 10^{-6}y^3 \quad 4$$

$$T(\text{Badin,synthesized}) = 1.8 - 0.05y + 1.08 \times 10^{-3}y^2 - 3.91 \times 10^{-6}y^3 \quad 5$$

$$T(\text{Hyderabad,original}) = 4.772 - 0.033y + 4.599 \times 10^{-5}y^2 + 7.179 \times 10^{-6}y^3 \quad 6$$

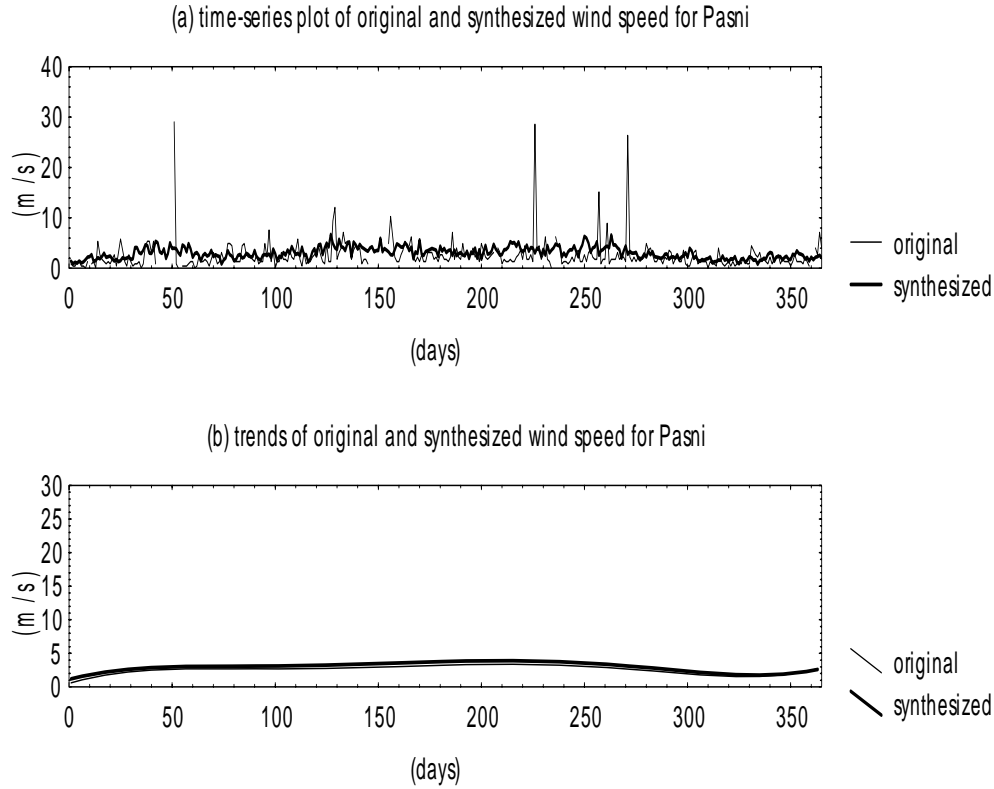


Fig. 3: Time-series and trend plots for Pasni airport, showing graphical comparison between the original and the synthesized daily wind speed data and their trend-lines.

Table 2: Statistical parameters computed for the original and the synthesized wind speed data.

Badin Airport							
	n (no. of days)	v_{mean} (m/s)	v_{min} (m/s)	v_{max} (m/s)	σ (m/s)	CV	R 95% Confid.
Data _{Original}	350	2.826	0.000	37.101	3.104	1.098	0.561
Data _{Synthesized}	365	3.000	0.596	6.453	1.847	0.616	
Hyderabad Airport							
	n (no. of days)	v_{mean} (m/s)	v_{min} (m/s)	v_{max} (m/s)	σ (m/s)	CV	R 95% Confid.
Data _{Original}	350	6.048	0.457	18.280	3.408	0.564	0.721
Data _{Synthesized}	365	6.000	1.889	13.901	2.799	0.467	
Pasni Airport							
	n (no. of days)	v_{mean} (m/s)	v_{min} (m/s)	v_{max} (m/s)	σ (m/s)	CV	R 95% Confid.
Data _{Original}	342	2.568	0.000	29.055	3.018	1.175	0.134
Data _{Synthesized}	365	3.000	0.788	6.738	1.139	0.380	

n =number of observations, v =wind speed, σ =standard deviation, CV=coefficient of variance, R =correlation value

$$T(\text{Hyderabad, synthesized}) = 5.2 - 0.07y + 6.55 \times 10^{-4}y^2 + 2.58 \times 10^{-6}y^3 \quad 7$$

$$T(\text{Pasni, original}) = 0.507 + 0.095y - 0.002y^2 + 1.096 \times 10^{-5}y^3 \quad 8$$

$$T(\text{Pasni, synthesized}) = 1.12 + 0.08y - 1.33 \times 10^{-3}y^2 + 9.79 \times 10^{-6}y^3 \quad 9$$

The statistical parameters computed for original and synthesized wind data of the three airport locations have been summarized in Table 2. The mean wind speeds of the original and the synthesized data for Badin, Hyderabad and Pasni airports show compatible results. Moreover, the behaviors of the trend lines also show very

close coincidence with each other for the three airport locations.

DISCUSSION

In connection with the trend behavior analysis, statistical analysis including multiple regression analysis have also performed on the original versus synthesized wind data for the airport locations. In comparison to the original data, the coefficient of variance and standard deviation of the synthesized data for the Hyderabad Airport shows low value to variability. Furthermore, the Pearson correlation analysis based regression value (with 95% confidence), i.e. 0.721 for the Hyderabad Airport, also conclude the higher level of similarity between original and synthesized wind speed data. If we match up the values of standard deviations and coefficient of variances for original and synthesized data of Badin Airport, we have observed some variability. Also, the value of Pearson correlation analysis is just 0.561, which indicate lower level of data uniformity. This is because of the gusting wind speed behavior, which is the abnormal events and do occur any time and hence affect the systematic wind speed model. The standard deviation and coefficient of variance values for the wind data of Pasni Airport are scattered and the regression analysis shows a poor level of data similarity. This is because of the complex terrain around Pasni causing intense gusting frequently, which eventually influences the wind model heavily.

The above analyses and discussion illustrate that the Markov chain technique can also be replicated for the other meteorological stations having low-resolution wind speed data located in other parts of the Pakistan especially having flat terrains. For the wind speed data of the complex terrain meteorological stations, this technique may yield high coefficient of variance. Thus, it is recommended that for the complex terrain wind speed data interpolation, a more specific Markov chain model with control variables may be needed.

CONCLUSION

Considering the results of the study, it is inferred that the computer model based for the first-order Markov chain technique can be used to generate the synthesized high-resolution (daily) wind speed data from the available low-resolution (monthly) wind speed data for the coastal regions of Pakistan having relatively flat terrain. It is also recommended that for the coastal areas of Pakistan having complex or mix-type terrains and having abnormal gusting wind feature, the second-order Markov chain method should be attempted to increase the certainty in the synthesized wind speed data.

REFERENCES

- Zaigham, NA and Nayyar ZA. 2005. Prospects of renewable energy sources in Pakistan. *In: Renewable energy technologies and sustainable development*, Khan, HA, Qurashi, MM, Hussain, T and Hayee, I (eds.). COMSATS. 65-86.
- Nayyar, ZA. 2009. Investigation & development of GIS-linked aerodynamic wind potential models of Sindh and Balochistan coastal areas of Pakistan. Ph.D. dissertation (unpublished).
- Shah, MA. 1991. Wind Power an answer for socio economic uplift of rural areas. *In: Wind Energy for Rural Areas*, Jong, I and Hule, FV (eds.). 89-102.
- Chaudri, QUZ and Farooqi, AB. 2000. Study on wind power generation in the coastal areas of Balochistan. Project report of Pakistan Meteorological Department, Islamabad. p.37.
- Khan, NA. 2001. Project report on wind mapping of Pakistan. Report of Pakistan Council for Appropriate Technology, Islamabad. p.6.
- Shamshad, A, Bawadi, MA, Hussin, WMAW, Majid, TA and Sanusi, SAM. 2005. First and second order Markov chain models for synthetic generation of wind speed time series. *Energy*. 30 (5): 693-708.
- Sahin, AD and Sen, Z. 2001. First-order Markov chain approach to wind speed modeling. *J Wind Eng. Ind. Aerodyn*. 89(3-4): 263-269.
- Ettoumi, FY, Sauvageot, H and Adane, AEH. 2003. Statistical bivariate modeling of wind using first-order Markov chain and Weibull distribution. *Renewable Energy*. 28(11):1787-1802.
- Weather Underground Inc, WUI. 2009. Information about the Data - Weather Underground [*Online*]. Available: <http://www.underground.com/about/data.asp>.
- Lun, IYF and Lam, JC. 2000. A study of Weibull parameters using long-term wind observations. *Renewable Energy*. 20(2):145-153.
- Nfaoui, H, Essiarab, H and Sayigh AAM. 2004. A stochastic Markov chain model for simulating wind speed time series at Tangiers, Morocco. *Renewable Energy*. 29(8):1407-1418.
- Logofet, DO and Lensnaya, EV. 2000. The mathematics of Markov models - what Markov chains can really predict in forest successions. *Ecol. Model*. 126(2-3):285-298.
- Heiko, B. 2000. Markov chain model for vegetation dynamics. *Ecol. Model*. 126(2-3):139-154.

Received: 11-01-2011

Accepted: 31-03-2011