

## MIT Open Access Articles

*Pitfalls of modeling wind power using Markov chains*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Brokish, K., and J. Kirtley. "Pitfalls of modeling wind power using Markov chains." Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES. 2009. 1-6. © 2009 Institute of Electrical and Electronics Engineers.

**Published Version:** <http://dx.doi.org/10.1109/PSCE.2009.4840265>

**Publisher:** Institute of Electrical and Electronics Engineers

**Permanent Link:** <http://hdl.handle.net/1721.1/58935>

**Version:** Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

**Terms of use:** Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



# Pitfalls of Modeling Wind Power Using Markov Chains

Kevin Brokish, *Student Member, IEEE*, James Kirtley, *Fellow, IEEE*

**Abstract**—An increased penetration of wind turbines have given rise to a need for wind speed/power models that generate realistic synthetic data. Such data, for example, might be used in simulations to size energy storage or spinning reserve. In much literature, Markov chains have been proposed as an acceptable method to generate synthetic wind data, but we have observed that the autocorrelation plots of wind speeds generated by Markov chains are often inaccurate. This paper describes when using Markov chains is appropriate and demonstrates the gross underestimation of storage requirements that occurs at short time steps. We found that Markov chains should not be used for time steps shorter than 15 to 40 minutes, depending on the order of the Markov chain and the number of wind power states. This result implies that Markov chains are of limited use as synthetic data generators for small microgrid models and other applications requiring short simulation time steps. New algorithms for generating synthetic wind data at shorter time steps must be developed.

**Index Terms**—Energy storage, Markov processes, modeling, wind energy, wind power generation.

## I. INTRODUCTION

RECENT investment into wind power has led to speculation about what infrastructure will be needed to incorporate this nondispatchable generation reliably. How much storage or spinning reserve is necessary? If a massive amount of wind speed/power data and load data has been collected for a specific location, then simulations with real wind data and virtual storage might yield quantitative requirements, but most locations lack this wealth of data. Time series simulations may yield storage estimates, but the estimates will be accurate only if realistic synthetic time series data can be generated for wind turbines.

In much literature, Markov chains have been proposed as a reasonably acceptable generator of synthetic wind speed data. Authors have used various transition matrix sizes, various time steps, and various orders:

- Jones 1986 - Uses first-order  $11 \times 11$  transition matrix for eight-hour means.[1]
- Kaminsky 1991 - Uses first-order and second-order 21 wind speed Markov chain at 3.5Hz. Correctly points out that the Markov model does not contain enough low-frequency data.[2]
- Sahin 2001- Uses first order  $8 \times 8$  transition matrix for hourly time steps. States second and even third order autocorrelation coefficients are significant, and suggests higher order transition matrices for future work.[3]

- Ettoumi 2003 - Uses a highly discretized ( $3 \times 3$ ) Markov transition matrix for three-hour increments. Notes that measurements performed at h-6, h-9 are non-negligible.[4]
- Nfaoui 2004 - Uses first-order  $12 \times 12$  transition matrix for hourly means.[5]
- Shamshad 2005 - Uses first and second-order Markov chains with 12 wind speeds for hourly means. This paper notes the autocorrelation plots are a poor match.[6]
- Papaefthymiou 2008 - Uses 35-state first through third-order Markov chains for 30-minute intervals.[7]

Using months of power data from a 1kW wind turbine in Denmark, we trained Markov chains of various orders, various number of states, and various time steps to generate synthetic wind data. Although the probability distribution of the wind power was correct in each model, the generated synthetic data often lacked other characteristics of the original data. In particular, time evolution characteristics of the wind data characterized by autocorrelation plots of synthetic wind speeds or wind powers generated by Markov chains are often very different from the original data, especially for models with short time steps. Modeling wind with especially short time steps is vital for microgrid simulations, since microgrids have much lower inertia than larger grids, and will require agile automatic generation control or storage control. Additionally, power cannot be supplied to an islanded microgrid from elsewhere, so it is critical that the storage or backup generation be sized correctly.

This paper seeks to determine when Markov chains are appropriate for modeling wind, and demonstrates the danger of inappropriately applied Markov models.

## II. HOW MARKOV CHAINS WORK

A Markov chain is a model for representing a stochastic process whereby a state changes at discrete time steps. A finite set of states is defined, and the Markov chain is described in terms of its transition probabilities,  $p_{ij}$ , which determine the probability of transitioning from state  $i$  to state  $j$ , regardless of previous states that were visited.[8]

It is straightforward to represent wind data with Markov chains: each state is a wind speed [m/s] or a wind power [kW], and from any given speed/power, there is some probability distribution function of what the next speed/power will be.

### A. Mathematical Description

At its core, a Markov chain is a sequence of random variables ( $W_1, W_2, W_3, \dots$ ) such that future states

K. Brokish and J. Kirtley are with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, 02139 USA (e-mail: brokish@mit.edu)

$(W_{n+1}, W_{n+2}, \dots)$  are dependent only upon the current state  $W_n$  and are independent of all past states  $(W_1, W_2, \dots, W_{N-1})$ .

$$W_n = w_j | w_j \in \{w_1, w_2, \dots, w_{K-1}, w_K\} \quad (1)$$

Where  $\{w_1, w_2, \dots, w_{K-1}, w_K\}$  is the set of  $K$  discretized wind speeds or wind powers.

The transition probabilities between states can be represented by a transition matrix  $P$  such that the element  $p_{ij}$  is the probability of transitioning from state  $i$  to state  $j$ . Formally,

$$p_{ij} = P(W_{n+1} = w_j | W_n = w_i) \quad (2)$$

Since the transition probabilities from a given state must add to 1, it must be true that

$$\sum_j p_{ij} = 1 \quad (3)$$

### B. Higher Order Markov Chains

Though memoryless by mathematical definition since the current state solely determines the transition probability distribution, Markov chains can be created to have multiple time step memories. For a first order chain, each state represents a wind power value for a single time period, but it is possible to create an  $N$ -order Markov chain where each state is defined by a set of  $N$  wind powers. For example, a third order model ( $N = 3$ ) would include states with three elements:  $\{\omega_{n-2}, \omega_{n-1}, \omega_n\}$ , where  $\omega_n$  is the value of the wind power at time step  $n$ . In this higher order case, the next state's  $\{\omega_{n-2}, \omega_{n-1}\}$  must equal the current state's  $\{\omega_{n-1}, \omega_n\}$ . The problem with higher order Markov models is that there are  $K^N$  states, where  $K$  is the number of discretized wind powers and  $N$  is the order of the model, which is intractable for large  $N$ .

### C. Using Markov Chains to Model Wind

The process for creating a Markov chain from real data is as follows.

The continuous spectrum of wind power levels must be discretized into  $K$  states.

Begin with a zero matrix  $M$  of length equal to  $K$ , the number of discrete wind power states, and of dimension equal to one plus the order of the model. For example, a second order model with 32 discretized wind powers would begin with a zero matrix  $M$  of dimensions  $32 \times 32 \times 32$ .

Step through the real data, incrementing the "tally matrix"  $M$ . For example, if the wind power is  $w_i$  at time step  $n - 2$ ,  $w_j$  at time step  $n - 1$ , and  $w_k$  at time step  $n$ , then  $m_{ijk}$  would be incremented.

The probabilities for state transitions are calculated from the frequency of transitions, so it is necessary to normalize the matrix  $M$  into a probability transition matrix  $P$ . Each row of  $M$  along the highest dimension is divided by the sum of that row. Effectively, each highest dimension row then adds to 1, and is a valid probability mass function (PMF).

The transition matrix can then be used to simulate time series data. Given a number of 'seed' wind powers equal to the order of the model, the matrix provides a PMF of what the

next wind power will be. Using a random number generator in conjunction with the appropriate PMF from the matrix, the next wind power,  $\omega_{n+1}$ , is chosen. The process is continued indefinitely using the recently generated wind power states as inputs to the matrix to find the PMF of the next wind power.

### D. Wind Speed or Wind Power?

Markov chains can be used to represent either wind speed or wind power. Many of the mentioned studies model wind speed instead of wind power. Luckily, it is straightforward to convert a wind speed Markov chain to a turbine power output Markov chain.

$$P \propto v^3 \quad (4)$$

Wind turbines have a minimum cut-in speed, a maximum power output, and a cut-out speed, which can all be incorporated using the function:

$$P = \begin{cases} 0 & \text{if } v < v_{cutin} \\ Cv^3 & \text{if } v_{cutin} \leq v \leq \sqrt[3]{P_{max}/C} \\ P_{max} & \text{if } \sqrt[3]{P_{max}/C} < v < v_{cutout} \\ 0 & \text{if } v_{cutout} \leq v \end{cases} \quad (5)$$

Where  $C$ ,  $P_{cutin}$ ,  $P_{max}$ , and  $P_{cutout}$  are all properties of the wind turbine.

### E. Appeal of Markovian Wind

Markov chains are intuitively appealing for modeling wind because given the current wind speed or power, one can guess the possibilities for the value a short while later: it will probably be slightly windier, slightly less windy, or about the same. Markov chains are able to model this because from every state, there is a set of probabilities of transitions to other states. The output of Markovian wind models are a giant improvement over a simple Monte Carlo approach with no temporal correlation.

A less intuitive appeal of Markov chains is that they nearly perfectly reproduce the PDF of the original data. What follows is a proof for the first order case.

Recall the "tally matrix"  $M$  from Section II-C. In stepping through the data, the transition from  $\omega_{n-1} = w_i$  to  $\omega_n = w_j$  is tallied by incrementing  $m_{ij}$ , and then the transition from  $\omega_n = w_j$  to  $\omega_{n+1} = w_k$  is tallied by incrementing  $m_{jk}$ , and so on. So row  $j$  and column  $j$  of matrix  $M$  are both incremented because the wind transitioned to state  $j$  on one time step, and then from state  $j$  on the next time step. Also note that if the wind power remains the same for consecutive time steps ( $\omega_{n-1} = w_j$  and  $\omega_n = w_j$ ),  $m_{jj}$  is incremented, and still both row  $j$  and column  $j$  of matrix  $M$  are incremented. The first datum and last datum are not tallied in both the respective row and the column, but for large amounts of input data, the effect of these two individual data is diminished. Hence, for a large amount of input data,

$$\sum_j m_{ij} \approx \sum_i m_{ij} \quad (6)$$

Define  $c_i$  and  $c_j$  as

$$c_i = \sum_j m_{ij} \quad (7)$$

$$c_j = \sum_i m_{ij} \quad (8)$$

From Equation 6,

$$c_i \approx c_j \quad (9)$$

Since row  $i$  of matrix  $M$  is tallied for every wind power state transition from state  $i$ , the observed distribution from the real data is the

$$\pi_i = \frac{c_i}{\sum_i c_i} \quad (10)$$

Next, it is shown that  $\pi$  of the actual data matches the stationary distribution of the simulated data, referred to as  $\hat{\pi}$ . Recall transition matrix  $P$  is a normalized version of  $M$ , so each element of  $M$  must be divided by its row's sum,  $c_i$ .

$$p_{ij} = m_{ij}/c_i \quad (11)$$

$$p_{ij}c_i = m_{ij} \quad (12)$$

$$\sum_i p_{ij}c_i = \sum_i m_{ij} = c_j \quad (13)$$

Substituting from Equation 9, we find that  $C^T$  is a left Eigen vector of  $P$ .

$$\sum_i p_{ij}c_i \approx c_i \quad (14)$$

$$C^T \approx C^T P \quad (15)$$

The stationary probabilities of the Markov simulation,  $\hat{\pi}$ , [8] are also the left Eigen vector of  $P$  with Eigen value of 1, such that

$$\hat{\pi} = \hat{\pi} P \quad (16)$$

So  $C^T$  is approximately a scaled version of  $\hat{\pi}$ .  $\hat{\pi}$  is already scaled to sum to 1. Equation 10 defines  $\pi$  as  $C^T$  but scaled to sum to 1. Since  $\pi$  is a unique solution to the stationary distribution of irreducible, aperiodic, recurrent Markov chains,

$$\pi \approx \hat{\pi} \quad (17)$$

### III. RESULTS

The difference between real data and synthetic data generated by a Markov chain using an overly short time step is clearly visible in Fig. 1. This synthetic data was generated from a first order Markov model with a time step of 1 minute that was constructed from the historical data above it. As shown by the proof in Section II-E, this synthetic data has the same PMF as the real time series data. However, the simulated data lacks the persistence of the historical data and would clearly predict a radically different storage requirement.

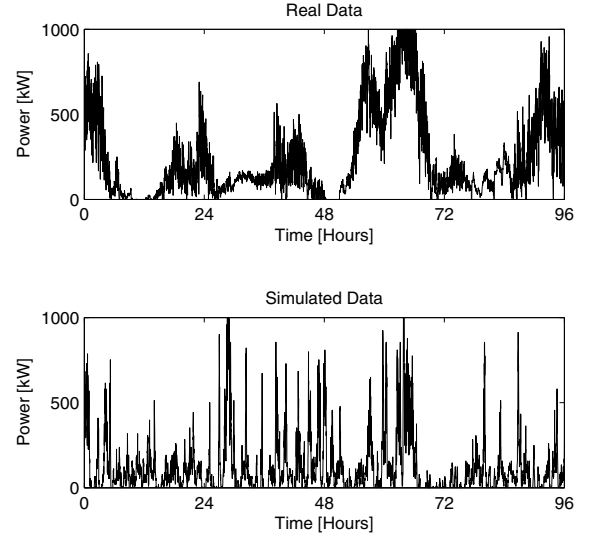


Fig. 1. Markov chains at small time steps clearly differ from the original data

#### A. Autocorrelation

Autocorrelation is a good metric for the dynamics of the Markov model. It is calculated by

$$R[j] = \frac{\sum_{k=1}^{l-j} (\omega_k - \bar{\omega})(\omega_{k+j} - \bar{\omega})}{l} \quad (18)$$

where  $\bar{\omega}$  is the mean of  $\omega$ . [9] Conceptually, the autocorrelation is a measure of the correlation between one data point and another that is  $j$  time steps ahead or behind it.

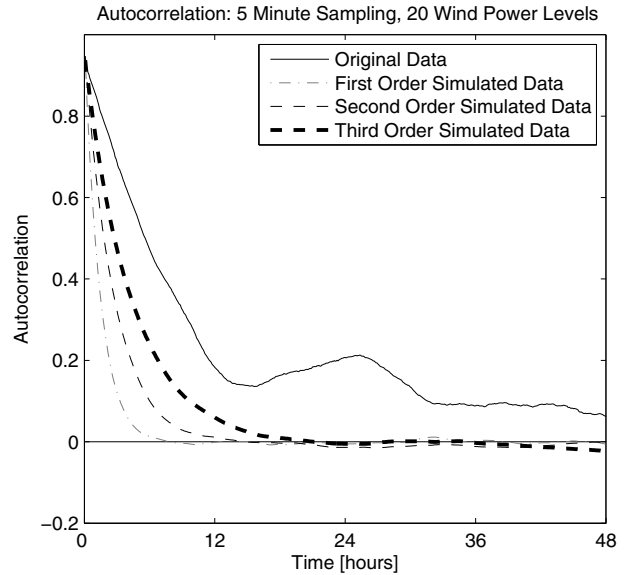


Fig. 2. Autocorrelation coefficients are too small at 5-minute time steps. (20 wind power levels)

The autocorrelation of synthetic data drops off too quickly when short time steps are used, as shown in Fig. 2. Clearly,

the problem is worse for first order Markov chains than it is for higher-order ones. The reason for the steep decline of the first order model is that even if current wind is affected by the wind five minutes ago, it has little to do with the wind several hours earlier.

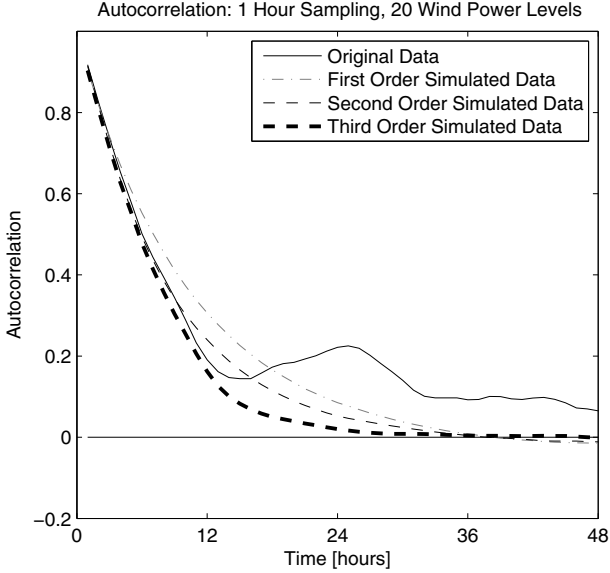


Fig. 3. Surprisingly, autocorrelation coefficients can be too large at longer time steps. (Hourly time steps, 20 wind power levels)

Interestingly, the autocorrelation for short lags under 12 hours is too large when longer time steps are used, as is shown in Fig. 3. Here the third order Markov chain does a better job of mimicking the early curve of the actual autocorrelation data.

In both cases, the third order model is generally better at replicating the original data than the first or second order models. Because the size of transition matrices is  $K^{(N+1)}$  where  $K$  is the number of wind powers and  $N$  is order, orders higher than  $N = 3$  were not tried.

Since Markov chains have only a short order/memory (one to three time steps in this study), it is impossible to model daily trends with time steps shorter than several hours. The local maximum around 24 hours in Fig. 2 and Fig. 3 occurs because afternoons are typically windier than mornings or nights at this wind turbine's location. As a result, the model is not valid for producing time series data for longer lengths of time. Interestingly, the first order model in Fig. 3 is actually a better match in the hourly case if the averaging occurs over 24 hours because of its abnormally high autocorrelation, which passes through the daily local maximum of the real data.

### B. Autocorrelation Error

In order to clearly determine the effect of various Markov chain parameters on the autocorrelation, a metric was developed for judging the accuracy of the autocorrelation function of the synthetic data: the RMS error of the autocorrelation

values between 0 and 12 hours. Formally,

$$\epsilon = \sqrt{\frac{\sum_{k=1}^T (R_{synthetic}[i] - R_{real}[i])^2}{T}} \quad (19)$$

where  $T$  is the number of time steps in the 12-hour window and  $R$  is the autocorrelation data.

Fig. 4, Fig. 5, and Fig. 6 display error as a function of time step and wind power resolution (the number of discrete wind power states) for first, second, and third order Markov models.

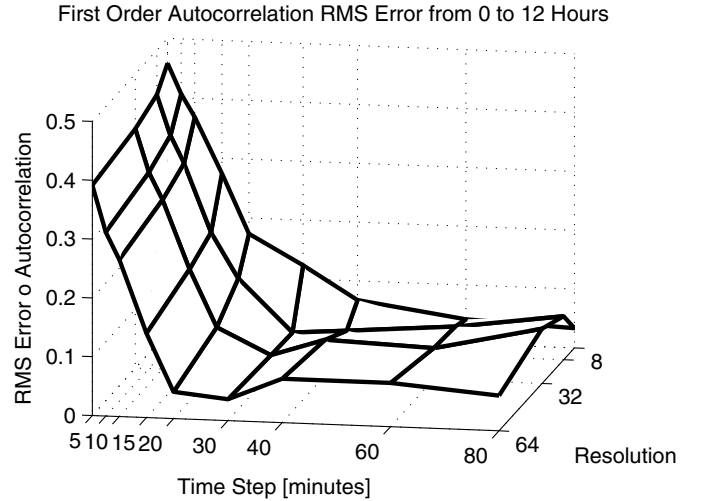


Fig. 4. First order autocorrelation error

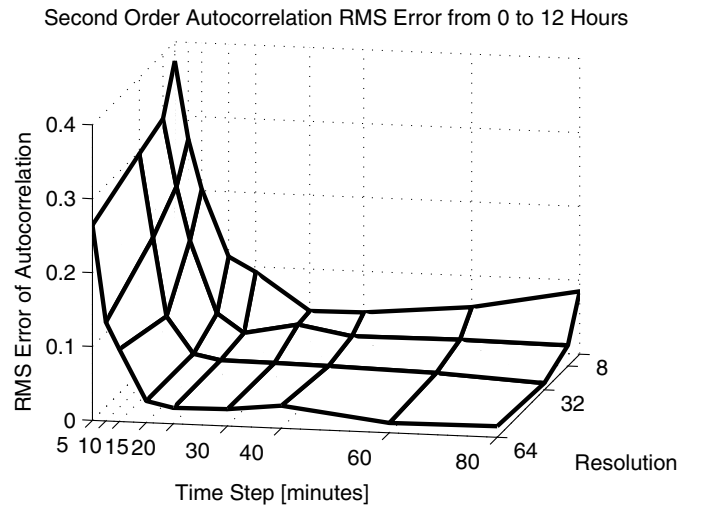


Fig. 5. Second order autocorrelation error

The important result is that higher order models reproduced the autocorrelation characteristics of real data down to the lower time step limit of 15 minutes.

In general, more states results in less error, but this is not always the case. In Fig. 4, the 80-minute model with eight wind states has less error than the model with 16. Also, small time steps are obviously inaccurate, but steps larger than an hour are not necessarily optimal, either. This is shown in Fig. 5, where the 8-state, 80-minute Markov model has more

Third Order Autocorrelation RMS Error from 0 to 12 Hours

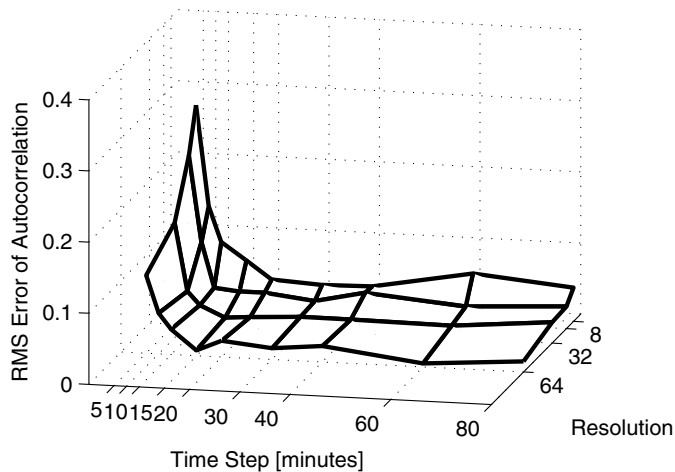


Fig. 6. Third order autocorrelation error

error than the 8-state 30-minute Markov model. Despite these exceptions, it is fairly clear that the best model is the third order model with many wind power states.

#### IV. WHY MARKOV CHAIN WIND MODELS ARE DANGEROUS

Dynamic wind data has multiple applications, including reliability studies as well as sizing storage and spinning reserve for grids with a high penetration of wind power. In order to determine whether or not the accuracy of the model actually matters, a simplified storage simulation was created.

##### A. Underestimated Storage

A hypothetical situation was modeled, in which a purely wind-powered microgrid was islanded for 12 hours. It is assumed that the average power generated by the wind model during this period perfectly matches the microgrid's perfectly flat load profile. The storage required by this situation was calculated by first integrating the generation minus the load to yield the energy stored as a function of time. Then the storage size was determined by subtracting the minimum energy from the maximum energy. Using the storage necessary to cover the worst case would have been highly susceptible to outliers, so the 95% tile was used instead—the amount of storage that would be able to supply the load in 95% of cases. This process was tried on both real and synthetic data. The synthetic storage values were then divided by the real storage value, yielding a storage fraction where 1 is a perfect estimate.

Fig. 7 is a plot of the storage fraction and the RMS error of the autocorrelation. Interestingly, all models underestimated the amount of storage necessary. Clearly the models with a low 12-hour autocorrelation RMS error yielded more accurate storage estimates. The higher error models grossly underestimated the storage necessary.

#### V. CONCLUSION

Markov models, while properly reflecting the probability density function of wind data, are not necessarily appropriate

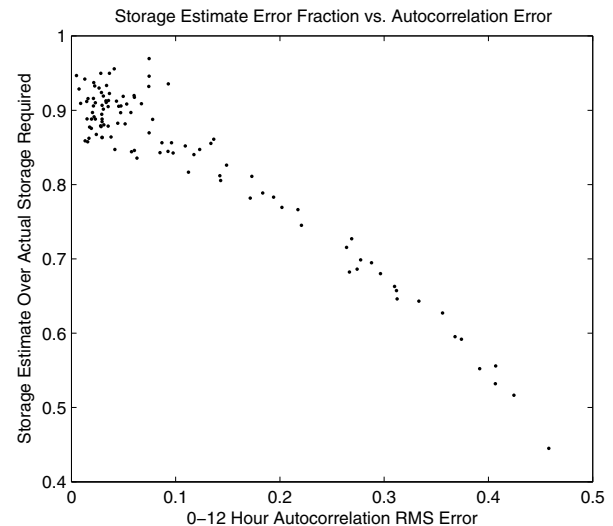


Fig. 7. Storage estimates and RMS error are strongly correlated

in generating synthetic data for simulations in which time evolution of the data is important for determining artifacts such as the need for storage. This is particularly true for short simulation time steps. While our study used anecdotal data from a single turbine, it was determined that the limit was about fifteen minutes. Even well fitted Markov models cause simulations to underestimate energy storage requirements.

New methods need to be developed for the generation of short time step synthetic wind speeds and powers: methods that can replicate an autocorrelation function while simultaneously retaining the correct probability distribution of the original data. ARMA models have been used in the literature, but these do not necessarily retain the probability distribution of the original data. Further study is also needed for large wind farms with many turbines, which will likely have smoother characteristics.

#### ACKNOWLEDGMENT

The authors would like to thank Dr. Scott Kennedy and Ozan Candogan for their help, as well as the MIT Portugal program for the funding and support of this research.

#### REFERENCES

- [1] D. Jones and M. Lorenz, "An application of a Markov chain noise model to wind generator simulation," *Mathematics and Computers in Simulation*, vol. 28, pp. 391–402, 1986.
- [2] F. Kaminsky, R. Kirchoff, C. Syu, and J. Manwell, "A comparison of alternative approaches for the synthetic generation of a wind speed time series," *Journal of Solar Energy Engineering*, vol. 113, pp. 280–289, 1991.
- [3] A. D. Sahin and Z. Sen, "First-order Markov chain approach to wind speed modelling," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 89, pp. 263–269, 2001.
- [4] F. Y. Ettoumi, H. Sauvageot, and A. Adane, "Statistical bivariate modelling of wind using first-order Markov chain and Weibull distribution," *Renewable Energy*, vol. 28, pp. 1787–1802, 2003.
- [5] H. Nfaoui, H. Essiarab, and A. Sayigh, "A stochastic Markov chain model for simulating wind speed time series at Tangiers, Morocco," *Renewable Energy*, vol. 29, pp. 1407–1418, 2004.
- [6] A. Shamshad, M. Bawadi, W. W. Hussin, T. Majid, and S. Sanusi, "First and second order Markov chain models for synthetic generation of wind speed time series," *Energy*, vol. 30, pp. 693–708, 2005.

- [7] G. Papaefthymiou and B. Klockl, "MCMC for wind power simulation," *IEEE Transactions on Energy Conversion*, vol. 23, no. 1, pp. 234–240, Mar. 2008.
- [8] D. P. Bertsekas and J. N. Tsitsiklis, *Introduction to Probability*. Belmont, Massachusetts: Athena Scientific, 2002, ch. 6, pp. 314–315,332.
- [9] C. Chatfield, *The Analysis of Time Series: An Introduction*, 4th ed. London, England: Chapman & Hall, 1989, ch. 4, p. 49.



**Kevin Brokish** is a graduate student at MIT with the Laboratory for Electromagnetic and Electronic Systems. He received a BS in Electrical and Computer Engineering from the University of Colorado at Boulder in 2007. Kevin is on the executive committee of the MIT Energy Club and is also a member of the MIT Energy Initiative. He has several patents pending.



**James Kirtley, Jr.** is Professor of Electrical Engineering at MIT. He holds a Ph.D. from MIT in Electrical Engineering. Dr. Kirtley is a Fellow of IEEE, a member of the National Academy of Engineering and a Registered Professional Engineer in Massachusetts. He was awarded the Third Millennium Medal of the IEEE in 2000 and the Nikola Tesla award in 2002.