

# XÂY DỰNG MÔ HÌNH MÔ PHÒNG TỐC ĐỘ GIÓ TỪNG GIỜ TỪ TỐC ĐỘ GIÓ TRUNG BÌNH HÀNG THÁNG

Nguyễn Thế Bảo<sup>(1)</sup>, Trần Hồng Hà<sup>(2)</sup>

(1) Trường Đại học Bách khoa, ĐHQG-HCM

(2) Trung tâm kiểm định kỹ thuật an toàn khu vực 2

(Bài nhận ngày 26 tháng 04 năm 2006, hoàn chỉnh sửa chữa ngày 30 tháng 07 năm 2006)

**TÓM TẮT:** Bài báo trình bày một phương pháp mô hình hóa thống kê tuyến tính và quy trình mô phỏng cho phép tạo ra chuỗi số liệu tốc độ gió từng giờ cho cả năm với số liệu đầu vào là tốc độ gió trung bình hàng tháng. Kết quả so sánh số liệu mô phỏng từ mô hình với số liệu quan sát thực tế cho thấy mô hình đảm bảo độ tin cậy để tạo ra chuỗi số liệu tốc độ gió đầu vào trong các bài toán mô phỏng kỹ thuật, nhất là các bài toán về năng lượng gió.

## 1. GIỚI THIỆU:

Số liệu tốc độ gió theo giờ là dữ kiện đầu vào không thể thiếu để giải rất nhiều bài toán trong các lĩnh vực nông nghiệp, môi trường, xây dựng, kỹ thuật nhiệt, kỹ thuật năng lượng gió v.v. Ở các nước phát triển, số liệu tốc độ gió được theo dõi một cách có hệ thống từ lâu, các trạm quan trắc khí tượng có thể cung cấp một cách chi tiết số liệu về tốc độ gió từng giờ trong thời gian dài ở từng vùng. Ở nước ta phần lớn các trạm quan trắc khí tượng chỉ có thể cung cấp số liệu tốc độ gió được đo một cách thủ công với tần suất đo 6 giờ/lần hay 3 giờ/lần. Hiện nay ở khu vực phía nam chỉ có trạm quan trắc khí tượng Nhà Bè – Tp. Hồ Chí Minh tiến hành đo tốc độ gió từng giờ bằng máy đo tự động, bắt đầu đo từ năm 2004. Sự hạn chế cả về số lượng và chất lượng của số liệu tốc độ gió ảnh hưởng rất lớn đến tính chính xác của lời giải các bài toán có liên quan đặc biệt là các bài toán mô phỏng trong kỹ thuật nhiệt, kỹ thuật năng lượng gió.

Bài viết này trình bày một phương pháp mô hình hóa thống kê nhằm xây dựng chương trình mô phỏng trên máy tính tạo ra chuỗi số liệu tốc độ gió theo giờ trong thời gian một năm với số liệu đầu vào là số liệu tốc độ gió trung bình hàng tháng.

### 1.1. Mô hình Box-Jenkins

Chuỗi số liệu về tốc độ gió trung bình theo giờ tại một địa điểm nào đó là một chuỗi thời gian (time series) với các đặc tính cơ bản:

- Là một dãy số thống kê, đặc tính thống kê của dãy số được biểu diễn qua các tham số đặc trưng như: dạng phân bố xác suất, giá trị trung bình, độ lệch chuẩn.
- Các giá trị của dãy số có liên hệ với nhau: giá trị hiện tại phụ thuộc vào các giá trị trước đó. Quan hệ tương quan này được phản ánh qua các hàm tự tương quan (autocorrelation function - acf) và tự tương quan riêng (partial autocorrelation function - pacf). Một trong những phương pháp mô hình hóa phổ biến nhất để mô phỏng đặc tính liên hệ tương quan của chuỗi số liệu là phương pháp mô hình hóa Box – Jenkins [4].

Mô hình Box-Jenkins hay mô hình ARMA (Autoregressive Moving Average) là mô hình tuyến tính có cấu trúc bao gồm 3 loại mô hình cơ bản:

- 1) Mô hình tự hồi quy (Autoregressive) AR(p): giá trị của chuỗi thời gian tại thời điểm hiện tại được xác định từ p giá trị trước đó và một xung nhiễu:

$$X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t$$

Với  $\alpha_i$  là các hệ số không đổi,  $\varepsilon_t$  là nhiễu trắng có trung bình bằng 0

- 2) Mô hình trung bình trượt (Moving Average) MA(q): giá trị của chuỗi thời gian tại thời điểm hiện tại được xác định từ trung bình trượt (có trọng số) của xung nhiễu tại thời điểm hiện tại và q thời điểm trước đó:

$$X_t = \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \dots + \beta_q \varepsilon_{t-q}$$

Với  $\beta_i$  là các hệ số không đổi,  $\varepsilon_i$  là các nhiễu trắng có trung bình bằng 0

3) Mô hình kết hợp ARMA(p,q):

$$X_t = \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q}$$

Mô hình này là sự kết hợp của hai mô hình trên.

Trong đó mô hình tự hồi quy bậc 2 AR(2) được sử dụng phổ biến nhất để mô phỏng số liệu tốc độ gió theo giờ.

Mô hình Box-Jenkins chỉ áp dụng được khi chuỗi thời gian có phân bố chuẩn và không có các thành phần biến đổi theo quy luật xác định (trend) và các thành phần biến đổi theo chu kỳ (seasonal components).

### 1.2.Áp dụng mô hình Box-Jenkins mô phỏng tốc độ gió tại từng địa điểm cụ thể

Brown B. G. et al (1984) [5] là người đầu tiên đề xuất một quy trình đầy đủ trong việc áp dụng mô hình Box-Jenkins vào mô phỏng tốc độ gió:

- 1) Chia chuỗi số liệu về tốc độ gió đo được từng giờ trong năm thành 12 chuỗi số biểu thị tốc độ theo giờ cho từng tháng nhằm loại bỏ thành phần biến đổi theo chu kỳ mùa.
- 2) Áp dụng phương pháp mũ hóa nhằm biến đổi chuỗi số từ phân bố Weibull thành phân bố chuẩn.

$$u'(t) = [u(t)]^m$$

Với  $m = k/3,6$  trong đó k là thông số hình dạng của phân bố Weibull

3) Sử dụng phép biến đổi:

$$u^*(t) = \frac{u'(t) - \mu(h)}{\sigma(h)}$$

Với giả thiết rằng  $\mu(h)$  và  $\sigma(h)$  là các hàm biến đổi có chu kỳ 24h tức là  $\mu(1) = \mu(25)$ ,  $\mu(2) = \mu(27)$ , v.v. để loại bỏ các thành phần biến đổi theo chu kỳ ngày và đưa chuỗi số về dạng phân bố chuẩn hóa.

Có thể thấy  $\mu(h)$  và  $\sigma(h)$  chính là các vec tơ 24 phần tử đặc trưng cho giá trị trung bình và độ lệch chuẩn của thành phần biến đổi theo chu kỳ ngày:

$$\mu(h) = \frac{1}{M} \sum_{i=1}^M u'(h, i) \quad \text{và} \quad \sigma^2(h) = \frac{1}{M} \sum_{i=1}^M [u'(h, i) - \mu(h)]^2$$

Với:  $h = 1 \div 24$  : chỉ số tương ứng với số giờ trong ngày, M: tổng số ngày của tháng,  $i = 1 \div M$  : chỉ số tương ứng với số ngày trong tháng.

4) Sau đó áp dụng mô hình AR(2) cho chuỗi  $u^*(t)$  đã được biến đổi có phân bố chuẩn hóa.

Quy trình này tiếp tục được nhiều tác giả (Daniel và Chen (1991) [8], Nfaoui et al (1996) [11], Torres et al (2004) [14]) áp dụng cho nhiều địa điểm khác nhau để dự báo ngắn hạn tốc độ gió phục vụ cho vận hành các trạm phát điện bằng sức gió. Tuy nhiên việc áp dụng trực tiếp phương pháp này vào điều kiện Việt Nam gặp rất nhiều khó khăn:

- Chúng ta không có sẵn các chuỗi số liệu quan sát tại từng vùng trong khi mô hình chỉ sử dụng được cho một địa điểm cụ thể.
- Việc phân tích với các số liệu hiện có cho thấy số liệu tốc độ gió tại Việt nam phần lớn không tuân theo phân bố Weibull.

### 1.3.Mô hình chung áp dụng được cho nhiều địa điểm :

Trên cơ sở khảo sát chuỗi số liệu của 10 địa điểm khác nhau ở Australia với nhận xét rằng trong những vùng có địa hình và đặc tính khí hậu tương đối thuần nhất, đặc tính thống kê của số liệu tốc độ gió theo giờ có tính quy luật khá ổn định, Nguyễn Thế Bảo (1996) [13] đề xuất phương pháp xây dựng mô hình chung (Common model) có thể áp dụng cho toàn vùng trong điều kiện chỉ có số liệu đầu vào là các giá trị trung bình theo tháng. Các bước mô phỏng được thực hiện như sau:

1) Xây dựng mô hình tự hồi quy bậc 2 AR(2) cho các địa điểm khác nhau. Lấy trung bình giá trị các tham số của các mô hình này để có một mô hình AR(2) mang tính đại diện.

2) Khi chạy chương trình mô phỏng, bước đầu tiên là tạo ra chuỗi số liệu  $U^*(t)$  có phân bố chuẩn hóa thông qua mô hình AR(2) đại diện đã có.

3) Thực hiện phép biến đổi ngược (destandardizing) để đưa chuỗi  $U^*(t)$  thành chuỗi  $U'(t)$ :

$$U'(t) = U^*(t)(\sigma)^m + (\mu)^m$$

-  $\sigma$  và  $\mu$  là độ lệch chuẩn và trung bình tốc độ gió trong tháng. Trong đó độ lệch chuẩn  $\sigma$  được xác định dựa vào công thức thực nghiệm có được trên cơ sở khảo sát các chuỗi số liệu:

$$\sigma = 0,948 + 0,2003\mu - 0,0501\sigma_{yr}$$

Với  $\sigma_{yr}$  là độ lệch chuẩn của số liệu gió cả năm.

-  $m$  là trung bình số mũ biến đổi có được bằng cách tra bảng theo  $\mu$ .

4) Xác định hệ số hình dạng  $k$  của phân bố Weibull từ các giá trị  $\sigma$  và  $\mu$  bằng phương pháp đồ thị.

5) Xác định lại số mũ biến đổi  $m = k/3,6$  và thực hiện biến đổi chuỗi  $U'(t)$  thành chuỗi  $U(t)$  có phân bố Weibull.

Việc áp dụng trực tiếp quy trình mô hình hóa và mô phỏng này vào điều kiện Việt Nam gặp các khó khăn cơ bản:

- Công thức xác định độ lệch chuẩn của chuỗi số liệu sau khi biến đổi bằng phương pháp mũ hóa theo độ lệch chuẩn của chuỗi số liệu chưa biến đổi  $\sigma' = (\sigma)^m$  không hoàn toàn đúng với số liệu tại Việt Nam.

- Mô hình chưa mô phỏng được yếu tố biến đổi có quy luật chu kỳ ngày của chuỗi số liệu tốc độ gió.

- Để xác định độ lệch chuẩn của chuỗi số liệu hàng tháng  $\sigma$  cần biết cả độ lệch chuẩn của số liệu gió cả năm  $\sigma_{yr}$ . Trong điều kiện số liệu hạn chế tại nước ta, rất khó xác định giá trị  $\sigma_{yr}$ .

- Phương pháp chỉ phù hợp nếu các chuỗi số liệu gió hàng tháng thực sự có phân bố Weibull

Trong phạm vi nghiên cứu của đề tài này, chúng tôi cố gắng xử lý các khó khăn nêu trên để áp dụng vào việc khảo sát các chuỗi số liệu hiện có nhằm xây dựng một mô hình chung áp dụng được cho bài toán mô phỏng số liệu tốc độ gió tại Việt Nam.

## 2. PHƯƠNG PHÁP MÔ HÌNH HOÁ

### 2.1. Dữ liệu phục vụ xây dựng mô hình

Dữ liệu dùng làm cơ sở để xây dựng mô hình bao gồm:

- Số liệu gió từng giờ cả năm 2004 tại trạm khí tượng Nhà Bè.

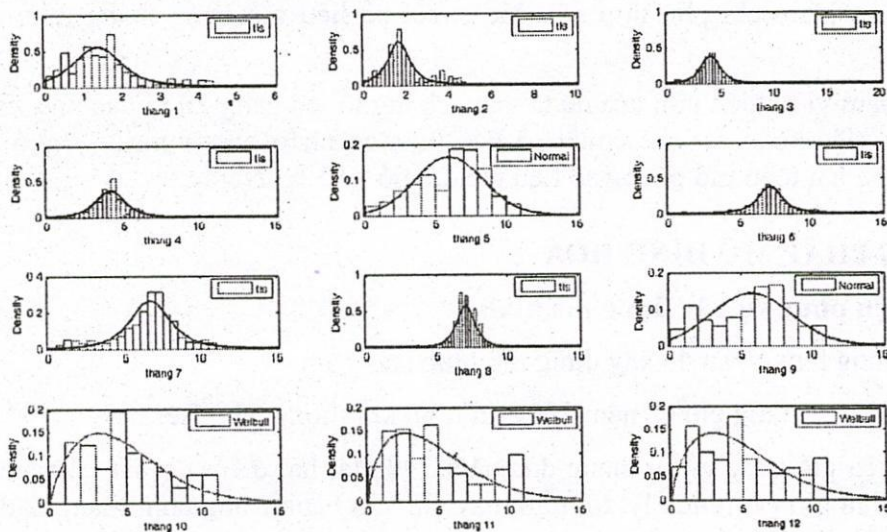
- Số liệu gió từng giờ 9 tháng đầu năm 1993 tại hai điểm đo đặt tại trạm phát điện gió Cần Thạnh, Cần giờ (vì nhiều lý do, trạm này chỉ vận hành trong chín tháng đầu năm 1993 rồi không tiếp tục hoạt động)

Số liệu tốc độ gió tại từng địa điểm được tách thành 30 chuỗi số liệu bao gồm các giá trị tốc độ gió từng giờ trong một tháng tại một địa điểm. Có thể thấy, so với cơ sở dữ liệu sử dụng để xây dựng mô hình chung tại Australia [13], số lượng dữ liệu có được hiện tại là hết sức hạn chế. Mặc dù vậy căn cứ trên sự đồng nhất về địa hình (đồng bằng) và chế độ gió (ven biển) của khu vực lân cận Thành phố Hồ Chí Minh và các tỉnh Nam Bộ cũng như một số cải tiến thích hợp trong phương pháp mô hình hóa, có cơ sở để tin rằng mô hình được xây dựng sẽ có đủ độ tin cậy cần thiết để áp dụng cho khu vực này.

Để minh họa, đặc tính thống kê của 12 chuỗi số liệu tại Nhà bè năm 2004 và kết quả ghép (fit) các hàm phân bố xác suất vào các chuỗi số liệu được thể hiện ở bảng 1 và hình 1

**Bảng 1:** Đặc trưng thống kê của số liệu tốc độ gió tại trạm quan trắc Nhà bè năm 2004

Tháng	Trung bình (m/s)	Trung vị (m/s)	Tốc độ cực đại (m/s)	Độ lệch chuẩn (m/s)	Độ bất đối xứng	Hệ số chuẩn hóa	Phân bố phù hợp nhất
1	1,51	1,4	4,5	0,96	1,06	0,37	t location-scale
2	1,94	1,7	6,7	1,14	1,12	0,40	t location-scale
3	4,01	4,1	10,4	1,44	0,52	0,83	t location-scale
4	4,19	4,1	10,6	1,40	0,64	0,66	t location-scale
5	5,97	6,5	10,9	2,42	-0,35	1,18	Phân bố chuẩn
6	6,86	7,1	10,8	1,79	-1,17	1,91	t location-scale
7	6,34	6,7	10,7	2,10	-0,77	1,56	t location-scale
8	6,98	7,0	10,7	0,94	-0,87	2,50	t location-scale
9	5,72	6,3	11	2,79	-0,29	1,01	Phân bố chuẩn
10	4,75	4,4	12,4	2,82	0,45	0,47	Weibull
11	4,68	4,1	11,0	3,09	0,58	0,30	Weibull
12	4,85	4,4	10,9	3,04	0,42	0,41	Weibull



**Hình 1:** Phân bố mật độ xác suất số liệu tốc độ gió tại trạm quan trắc Nhà bè năm 2004

## 2.2. Nguyên lý xây dựng mô hình

Mô hình được xây dựng trên cơ sở áp dụng các giả thuyết đã đề ra khi xây dựng mô hình cho số liệu tại Australia [13]:

- Có thể dùng một mô hình ARMA duy nhất để mô phỏng đặc tính tự tương quan của chuỗi số liệu của từng tháng trong năm ở các địa điểm khác nhau trong mỗi vùng khí hậu.
- Các tham số đặc trưng cho dạng phân bố cũng như các thành phần biến đổi theo chu kỳ ngày của chuỗi số liệu như số mũ biến đổi  $m$ , các vec tơ  $\mu(h)$  và  $\sigma(h)$  trong mỗi vùng khí hậu có thể xác định được khi biết tốc độ gió trung bình tháng mà không phụ thuộc vào địa điểm.

Các bước xây dựng cụ thể được thay đổi cho phù hợp với điều kiện số liệu hiện có nhằm đảm bảo độ tin cậy của mô hình.

## 2.3. Xây dựng công thức xác định độ lệch chuẩn $\sigma$ và số mũ chuẩn hoá $m$ theo tốc độ trung bình tháng $\mu$

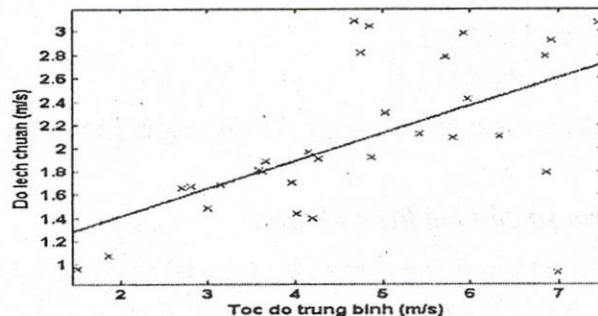
Khảo sát 30 chuỗi số liệu từng tháng cho thấy: phân bố xác suất phù hợp với các chuỗi số liệu không thuần nhất là phân bố Weibull mà chủ yếu là phân bố chuẩn hoặc phân bố  $t$  location-scale – một dạng phân bố xác suất ba tham số. Như vậy, không thể áp dụng phương pháp xác định số mũ biến đổi dựa trên giả thiết chuỗi số liệu là phân bố Weibull. Số mũ biến đổi sẽ được xác định theo phương pháp biến đổi Box-Cox dựa trên nguyên tắc tìm giá trị  $m$  làm cho thông số Log-Likelihood Function (LLF) đạt cực đại. Các ngôn ngữ mô phỏng như Matlab đều có sẵn các hàm số cho phép thực hiện biến đổi Box-Cox.

Việc khảo sát quan hệ phụ thuộc giữa số mũ chuẩn hóa  $m$  và giá trị độ lệch chuẩn  $\sigma$  (m/s) vào tốc độ trung bình tháng  $\mu$  (m/s) cho thấy không có sự khác biệt nhiều về sai số mô hình hóa giữa các mô hình tuyến tính bậc nhất, đa thức hay hàm mũ v.v. (sai khác về chỉ tiêu căn bậc hai của trung bình bình phương sai số (RMSE) giữa các mô hình không quá 5%). Vì vậy mô hình đơn giản nhất – mô hình tuyến tính bậc nhất được lựa chọn để xác định  $m$  và  $\sigma$  theo  $\mu$ . Áp dụng công cụ Curve fit của Matlab, hàm số biểu diễn sự phụ thuộc của  $m$  và  $\sigma$  vào  $\mu$  có dạng:

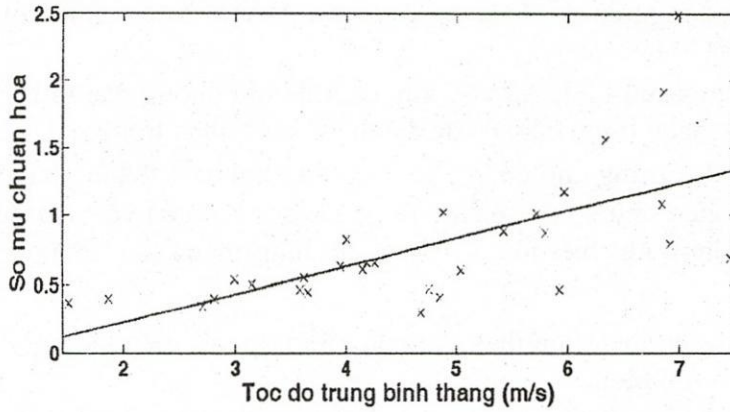
$$m = 0,202\mu - 0,176 \quad (1)$$

$$\sigma = 0,2379\mu + 0,9379 \quad (2)$$

Kết quả mô hình hóa bằng hàm tuyến tính bậc nhất được thể hiện ở hình 2 và 3 cho thấy đối với quan hệ  $\sigma=f(\mu)$  tồn tại các sai số khá lớn ở vùng tốc độ trung bình từ 5 – 7 m/s, nhược điểm này tồn tại ở tất cả các dạng mô hình (tuyến tính bậc nhất, đa thức, hàm mũ). Nhược điểm này chỉ có thể khắc phục khi có cơ sở dữ liệu về tốc độ gió đủ lớn – điều chưa thể đáp ứng được trong điều kiện hiện tại. Bảng số liệu cho thấy giá trị nhỏ nhất và lớn nhất của tốc độ gió trung bình tháng là 1,5104 m/s và 7,4544 m/s. Điều này có nghĩa là mô hình đang xây dựng chỉ có giá trị tin cậy với các giá trị tốc độ gió trung bình tháng từ 1,5 đến 7,5 m/s.



Hình 2: Sự phụ thuộc tuyến tính của độ lệch chuẩn  $\sigma$  vào tốc độ gió trung bình tháng  $\mu$



Hình 3: Sự phụ thuộc tuyến tính của số mũ chuẩn hóa m vào tốc độ gió trung bình tháng  $\mu$

**2.4. Xác định quan hệ giữa giá trị trung bình  $\mu'$ , độ lệch chuẩn  $\sigma'$  của chuỗi số liệu đã được chuẩn hóa theo tốc độ trung bình tháng  $\mu$**

Để thấy rằng  $\mu' = \mu^m$  (3)

Coi chuỗi số chuẩn hóa là giá trị của hàm số  $y=f(x)$  với x là các giá trị của chuỗi số chưa chuẩn hóa. Thực hiện khai triển Taylor bậc nhất xung quanh giá trị  $x = \mu$ , có thể rút ra quan hệ biểu diễn sự phụ thuộc của  $\sigma'$  vào  $\sigma$  và  $\mu$  như sau:

$$\sigma' = \sigma \left( \frac{\partial f}{\partial x} \Big|_{x=\mu} \right)$$

Đối với trường hợp này  $f(x) = x^m$  nên  $\frac{\partial f}{\partial x} = mx^{m-1}$ . Thay  $x=\mu$ :

$$\sigma' = m\sigma\mu^{m-1} \quad (4)$$

Ở đây m và  $\sigma$  được xác định qua  $\mu$  theo công thức (1) và (2)

**2.5. Xác định mô hình chung**

**2.5.1. Xác định dạng của mô hình chung**

Sau khi chuẩn hóa và loại bỏ các thành phần biến đổi theo chu kỳ ngày, ghép (fit) lần lượt các mô hình AR(2), ARMA(1,2), ARMA(2,1), ARMA(2,2) vào 30 chuỗi số liệu. Mô hình phù hợp nhất với chuỗi số liệu là mô hình có chỉ số thông tin Akaike (Akaike information criterion-AIC) nhỏ nhất [14]. Trong 30 mô hình:

- 14 mô hình AR(2)
- 3 mô hình ARMA(1,2)
- 5 mô hình ARMA(2,1)
- 8 mô hình ARMA(2,2)

Mô hình xuất hiện nhiều nhất là mô hình AR(2) được lựa chọn để xây dựng mô hình chung.

**2.5.2. Xác định các tham số của mô hình chung:**

Mô hình AR(2) chung cho từng tháng được xây dựng bằng cách:

- Xác định các tham số của mô hình AR(2) cho từng tháng tại từng địa điểm

- Lấy trung bình các mô hình này theo từng tháng để tạo ra mô hình AR(2) chung bằng lệnh merge của Matlab, lệnh này lấy trung bình có trọng số theo tham số  $\sigma^2_\varepsilon$  của các mô hình.

Số liệu chi tiết về các tham số của mô hình AR(2) chung được thể hiện ở bảng 3

**Bảng 2:** Tham số mô hình AR(2) của từng chuỗi số liệu và của mô hình chung

Tháng	Nhà bè 2004			Cần giờ 1993 (điểm đo A)			Cần giờ 1993 (điểm đo B)			Mô hình chung		
	$\alpha_1$	$\alpha_2$	$\sigma^2_\varepsilon$	$\alpha_1$	$\alpha_2$	$\sigma^2_\varepsilon$	$\alpha_1$	$\alpha_2$	$\sigma^2_\varepsilon$	$\alpha_1$	$\alpha_2$	$\sigma^2_\varepsilon$
1	-0,3023	-0,1073	0,8376	-0,9035	0,0034	0,1819	-0,9356	0,0973	0,2607	-0,7445	-0,0927	0,1819
2	-0,4450	-0,1423	0,6985	-1,0840	0,1636	0,1246	-0,9735	0,0807	0,1813	-0,8516	-0,0316	0,1246
3	-0,2844	-0,2455	0,7633	-1,1590	0,2189	0,0879	-0,9048	0,0486	0,2455	-0,7958	-0,1012	0,0879
4	-0,3653	-0,1853	0,7411	-0,6924	-0,1976	0,2337	-0,5414	-0,2805	0,3745	-0,5593	-0,2619	0,2337
5	-0,4738	-0,1773	0,6290	-0,8341	0,0244	0,3266	-0,7042	-0,0743	0,4064	-0,6730	-0,0938	0,3266
6	-0,2745	-0,1736	0,8270	-0,7172	-0,0271	0,4421	-0,7952	0,0682	0,4301	-0,5960	-0,0914	0,4421
7	-0,3440	-0,0093	0,8574	-0,7491	0,0216	0,4480	-0,7660	0,0438	0,4465	-0,6376	-0,0248	0,4480
8	-0,3809	-0,1552	0,8026	-0,8250	0,0019	0,3157	-0,7931	0,0021	0,3617	-0,6779	-0,0901	0,3157
9	-0,3373	-0,1546	0,7942	-0,8185	0,1310	0,4468	-0,8055	0,1123	0,4485	-0,6106	-0,0304	0,4468
10	-0,3433	-0,1447	0,8041							-0,3433	-0,1447	0,8041
11	-0,2570	-0,1489	0,8509							-0,2570	-0,1489	0,8509
12	-0,3549	-0,2174	0,7642							-0,3549	-0,2174	0,7642

### 3. QUY TRÌNH MÔ PHỎNG

Với các tham số mô hình nói trên, quá trình mô phỏng để tạo ra chuỗi số liệu tốc độ gió cho 8760 giờ của 12 tháng trong năm khi biết 12 số liệu tốc độ trung bình của các tháng được thực hiện như sau:

1. Từ 12 mô hình AR(2) chung của 12 tháng cho ở bảng 2, tạo ra 12 chuỗi số liệu  $U_t^*$  có phân bố chuẩn cơ sở:

$$U_t^* = \alpha_1 U_{t-1}^* + \alpha_2 U_{t-2}^* + \varepsilon_t$$

2. Từ 12 giá trị tốc độ trung bình tháng  $\mu$ , xác định 12 giá trị số mũ chuẩn hóa  $m$  và 12 giá trị độ lệch chuẩn  $\sigma$  của các chuỗi số liệu tốc độ gió từng tháng từ công thức (1) và (2).
3. Xác định giá trị trung bình  $\mu'$  và độ lệch chuẩn  $\sigma'$  của các chuỗi số đã được chuẩn hoá  $U_t'$  từ các giá trị tốc độ trung bình tháng  $\mu$  bằng công thức (3) và (4).
4. Tạo ra 12 chuỗi số  $V_t'$  ngẫu nhiên thuần túy có phân số chuẩn, trung bình  $\mu'$ , độ lệch chuẩn  $\sigma'$ , độ dài bằng số giờ từng tháng. Từ các chuỗi số  $V_t'$  này, rút ra 12 vec tơ  $\mu'_h$  và  $\sigma'_h$  của các tháng.

5. Biến đổi các chuỗi số liệu  $U_t^*$  để tạo ra 12 chuỗi số liệu  $U_t'$  có phân bố chuẩn và đặc tính biến đổi theo chu kỳ ngày:

$$U_t' = \sigma'_h U_t^* + \mu'_h$$

6. Biến đổi để đưa các chuỗi số  $U_t'$  thành 12 chuỗi số phi chuẩn  $U_t$  cần mô phỏng:

$$U_t = (U_t')^{1/m}$$

7. Ghép 12 chuỗi số phi chuẩn nói trên ta có chuỗi số gồm 8760 đại diện cho tốc độ gió trong 8760 giờ của năm.

#### 4. ĐÁNH GIÁ ĐỘ TIN CẬY CỦA MÔ HÌNH

##### 4.1. Chuỗi số liệu so sánh:

Các chuỗi số liệu sau đây được sử dụng để so sánh, đánh giá độ tin cậy của mô hình:

- Chuỗi số liệu tốc độ gió từng giờ tại Trạm quan trắc Nhà Bè năm 2005.
- Chuỗi số liệu tốc độ gió đo theo tần suất 3 h/lần tại Trạm quan trắc Vũng Tàu năm 2004.
- Chuỗi số liệu tốc độ gió đo theo tần suất 6 h/lần tại Trạm quan trắc Phú Quốc năm 2005.

Trong đó số liệu tại Nhà bè được ghi tự động, lấy giá trị đến 1 số lẻ sau hàng đơn vị. Số liệu tại Vũng Tàu và Phú Quốc được ghi thủ công, lấy giá trị đến hàng đơn vị. Đây là các chuỗi số liệu độc lập, không được sử dụng khi xây dựng mô hình.

##### 4.2. Phương pháp đánh giá

- Từ các chuỗi số liệu quan sát, tính 12 giá trị tốc độ trung bình của 12 tháng trong năm. Sử dụng 12 giá trị này để mô phỏng ra các chuỗi 8760 số liệu tốc độ của từng giờ trong năm bằng mô hình chung. Để đảm bảo tính tương tự, chuỗi số liệu mô phỏng tại Vũng tàu và Phú Quốc được lược bớt, chỉ giữ lại các số liệu tương ứng với các thời điểm quan sát. Như vậy, chuỗi số liệu mô phỏng tại Vũng tàu gồm 2920 số liệu, chuỗi số liệu mô phỏng tại Phú Quốc gồm 1460 số liệu

- So sánh số liệu mô phỏng với số liệu quan sát.

Độ tin cậy của mô hình được đánh giá chủ yếu thông qua so sánh với số liệu tại Nhà Bè. Các chuỗi số liệu quan sát tại Vũng Tàu và Phú Quốc có độ tin cậy thấp hơn, việc so sánh chỉ mang tính tham khảo.

**Bảng 3:** Tốc độ gió trung bình tháng (m/s) tại các địa điểm so sánh

Số liệu	Tháng											
	1	2	3	4	5	6	7	8	9	10	11	12
Nhà bè	4,076	3,912	3,963	4,067	5,654	6,272	6,500	6,753	6,840	4,749	2,660	2,405
Vũng tàu	2,746	3,107	3,718	2,742	1,831	2,004	2,431	2,327	1,980	1,815	1,762	1,423
Phú Quốc	1,331	1,696	1,976	1,775	2,323	4,658	9,508	4,508	4,600	2,202	2,483	3,605



- Khả năng mô phỏng của mô hình được đánh giá thông qua các chỉ tiêu cơ bản:

1) Mật độ năng lượng gió trung bình:

$$p = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} d v_i^3 \text{ (W/m}^2\text{)}$$

$d = 1,2 \text{ kg/m}^3$  là khối lượng riêng của không khí.

N: độ dài của chuỗi số liệu.

2) So sánh khả năng lặp lại hàm phân bố xác suất của chuỗi số mô phỏng so với chuỗi số quan sát.

3) So sánh sản lượng điện năng của cùng một tuabin gió:

Chọn mô hình tua bin gió MOD 0 nêu trong Chou K. C. và R. B. Corotis (1981) [8] có các đặc tính sau:

- Công suất định mức:  $P_r = 100 \text{ kW}$
- Vận tốc bắt đầu hoạt động:  $v_0 = 3,1 \text{ m/s}$
- Vận tốc đạt công suất định mức:  $v_1 = 6,5 \text{ m/s}$
- Vận tốc tối đa cho phép:  $v_2 = 15,3 \text{ m/s}$

Tổng công suất phát của tua bin trong năm được xác định như sau:

$$P_{total} = \sum_{i=1}^N n P_i / 1000 \text{ (MWh)}$$

Trong đó  $n$  là khoảng cách thời gian tính bằng giờ giữa hai giá trị tốc độ gió,  $N$  là độ dài của chuỗi số liệu.

Theo Jutus C. G. et al (1976) [9], đặc tính công suất phụ thuộc vào tốc độ gió của tua bin được xác định bởi biểu thức:

$$P(v) = \begin{cases} 0, v \leq v_0 \\ A + Bv + Cv^2, v_0 < v \leq v_1 \\ P_r, v_1 < v \leq v_2 \\ 0, v > v_2 \end{cases} \text{ (kW)}$$

$v$ : Vận tốc gió (m/s)

$P_r$ : công suất định mức của tua bin (kW)

A, B, C là các hệ số xác định từ hệ phương trình:

$$\left. \begin{aligned} A + Bv_0 + Cv_0^2 &= 0 \\ A + Bv_1 + Cv_1^2 &= P_r \\ A + Bv_c + Cv_c^2 &= P_r \left( \frac{v_c}{v_1} \right)^3 \end{aligned} \right\}$$

Với  $v_c = (v_0 + v_1)/2$

4) So sánh khả năng lặp lại hàm số tự tương quan (acf) của chuỗi số mô phỏng so với chuỗi số quan sát. Chỉ tiêu này chỉ có thể thực hiện được trên chuỗi số liệu đo từng

giờ tại Nhà bè năm 2005 vì với tần suất đo 3 giờ/lần hay 6 giờ/lần, chuỗi số liệu không thể hiện đặc tính tự tương quan.

**4.3.Kết quả so sánh**

Tiến hành mô phỏng 30 lần, lấy giá trị trung bình từ 30 lần mô phỏng. Kết quả so sánh mật độ năng lượng và tổng sản lượng điện năng được thể hiện ở các bảng 4 và 5. Kết quả so sánh hàm CDF của một lần mô phỏng thể hiện ở hình 4, 5, 6. Đồ thị hàm tự tương quan của một chuỗi số liệu mô phỏng và chuỗi số liệu quan sát thể hiện ở hình 7.

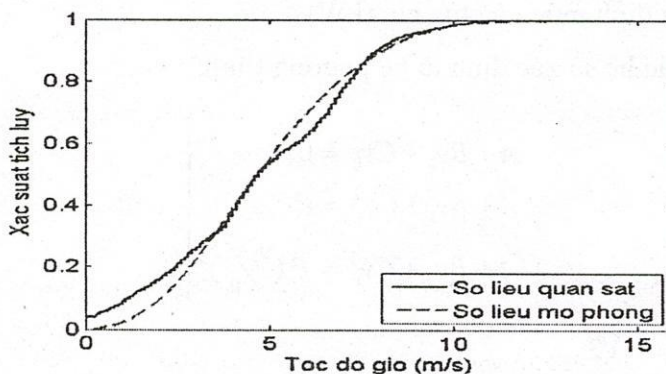
Kết quả mô phỏng thể hiện trên các bảng và đồ thị cho thấy mô hình lặp lại khá tốt các đặc tính thống kê và đặc tính tự tương quan của chuỗi số liệu. So sánh với số liệu quan sát tại Nhà bè năm 2005 sai số về mật độ năng lượng là 2,6%, sai số về tổng sản lượng của turbine gió là 7,1%. Kết quả này là hoàn toàn chấp nhận được trong điều kiện các giá trị này là hàm số bậc 3 và bậc 2 của tốc độ gió. Sai số mô phỏng tại Vũng tàu và Phú Quốc có thể lên đến 30%, tuy nhiên sai số này không thực sự là sai số của bản thân mô hình mà còn là sai số của số liệu quan sát. Việc chỉ lấy đến phần nguyên khi ghi các giá trị đo làm cho chuỗi số liệu có một phần bất định khá lớn, ảnh hưởng nhiều đến độ tin cậy của số liệu nhất là với các giá trị vận tốc gió nhỏ như ở Vũng tàu và Phú Quốc.

**Bảng 4:** So sánh mật độ năng lượng gió (W/m<sup>2</sup>) giữa số liệu quan sát và số liệu mô phỏng

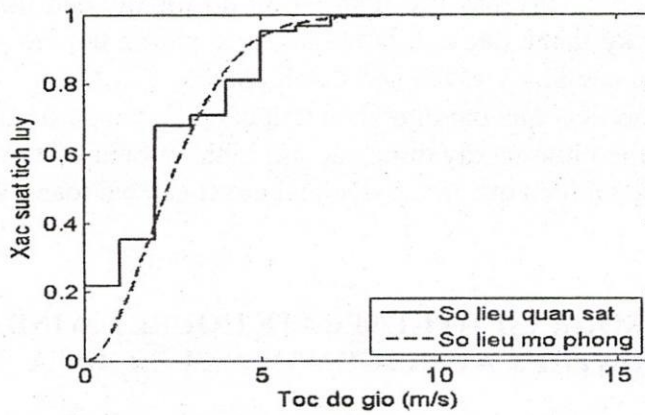
	Nhà bè 2005	Vũng Tàu 2003	Phú Quốc 2005
Tính theo số liệu quan sát	123,8958	26,4222	59,0967
Tính theo số liệu mô phỏng	120,6783	25,8576	44,7349
Sai số	2,6 %	2,14 %	24,30%
Độ lệch chuẩn của chuỗi 30 số liệu mô phỏng	1,4558	2,3700	2,4630

**Bảng 5:** So sánh tổng sản lượng của turbine gió trong 1 năm (MWh) giữa số liệu quan sát và số liệu mô phỏng

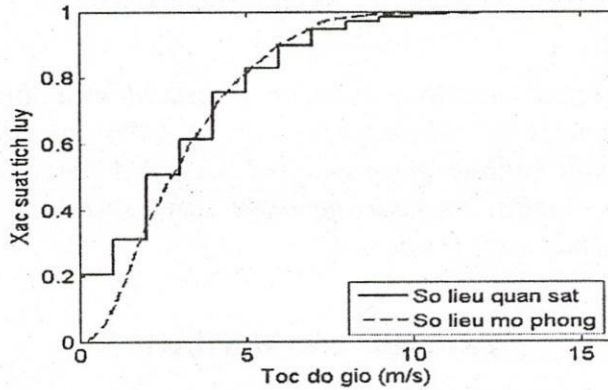
	Nhà bè 2005	Vũng Tàu 2003	Phú Quốc 2005
Tính theo số liệu quan sát	422,2715	113,4311	189,5590
Tính theo số liệu mô phỏng	392,3578	82,4908	170,9384
Sai số	7,1 %	27,28 %	9,82 %
Độ lệch chuẩn của chuỗi 30 số liệu mô phỏng	2,7870	4,1287	7,3831



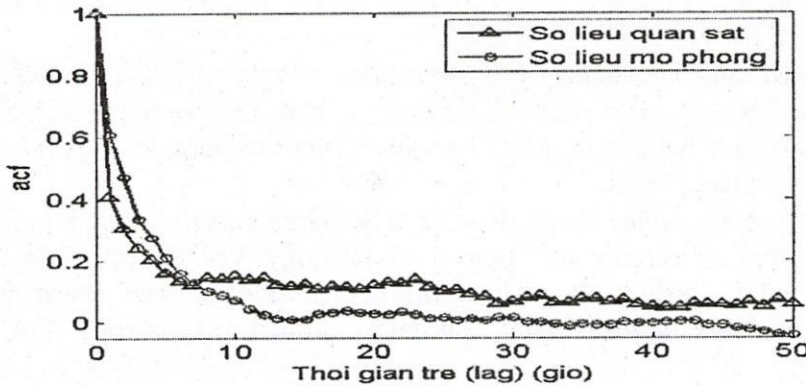
**Hình 4:** So sánh đồ thị hàm phân bố xác suất của chuỗi số quan sát tại Nhà bè năm 2005 và của chuỗi số mô phỏng



Hình 5: So sánh đồ thị hàm phân bố xác suất của chuỗi số quan sát tại Vũng Tàu năm 2003 và của chuỗi số mô phỏng



Hình 6: So sánh đồ thị hàm phân bố xác suất của chuỗi số quan sát tại Phú Quốc năm 2005 và của chuỗi số mô phỏng



Hình 7: So sánh đặc tính tự tương quan của chuỗi số quan sát tại Nhà bè năm 2005 và của chuỗi số mô phỏng

## 5. KẾT LUẬN

Một mô hình mô phỏng thống kê tuyến tính chung cho phép tạo ra chuỗi số liệu gió từng giờ trong một năm khi biết giá trị trung bình hàng tháng đã được xây dựng để đáp ứng nhu cầu số liệu về tốc độ gió theo giờ trong điều kiện thiếu các số liệu quan sát. Điểm còn tồn tại lớn nhất của mô hình là chưa mô phỏng được thông tin về hướng gió cũng như chưa xem xét tốc độ gió trong mối quan hệ với các thông số khí hậu khác như bức xạ mặt trời, nhiệt độ, độ ẩm v.v.

Mặc dù vậy kết quả kiểm tra cho thấy mô hình có đủ độ tin cậy cần thiết để áp dụng trong nhiều bài toán mô phỏng kỹ thuật, đặc biệt là bài toán mô phỏng tua bin gió và đánh giá tiềm năng phát điện sức gió tại các khu vực lân cận thành phố Hồ Chí Minh và các vùng ven biển, các đảo ở khu vực Nam bộ. Kết quả đạt được ban đầu cũng là căn cứ để tiếp tục xây dựng các mô hình hoàn thiện hơn theo hướng xây dựng các mô hình đa biến có thể mô phỏng tổng hợp các thông số khí hậu, phục vụ hiệu quả trong việc giải quyết các bài toán kỹ thuật phức tạp hơn.

## A SIMULATION PROGRAM TO GENERATE HOURLY WIND SPEED FROM MONTHLY AVERAGE WIND SPEED DATA

Nguyen The Bao<sup>(1)</sup>, Tran Hong Ha<sup>(2)</sup>

(1)University of technology, VNU-HCM

(2)Center of Industrial Safety Registration Zone

**ABSTRACT:** *This article describes a linear statistical modeling method to generate hourly wind speed sequences of whole year from monthly average wind speed data. Comparisons have been made between generated and measured data. The results show that the model proves sufficient reliability to generate input wind speed data series for technical simulation, especially in wind power studies.*

### TÀI LIỆU THAM KHẢO

- [1]. Akpınar E. Kavak and Akpınar S., *A statistical analysis of wind speed data used in installation of wind energy conversion systems*, Energy Conversion and Management Vol. 46, pp. 515-532, 2005.
- [2]. Aksoy H. et al, *Stochastic generation of hourly mean wind speed data*, Renewable Energy, Vol. 29, pp. 2111-2131, 2004.
- [3]. Blanchard M. and Desrochers G., *Generation of autocorrelated wind speeds for wind energy conversion system study*, Solar Energy, Vol. 33, No. 6, pp. 571-579, 1984.
- [4]. Box G. E. P. and Jenkins G. M., *Time series analysis forecasting and control*, Holden-day, San Francisco, 1970.
- [5]. Brown B. G. et al., *Time Series Models to Simulate and Forecast Wind Speed and Wind Power*, Journal of climate and applied meteorology, Vol. 23, pp. 1184-1194, 1984.
- [6]. Carlin J. and J. Haslett, *The probability distribution of wind power from a dispersed array of wind turbine generators*, Journal of applied meteorology, Vol. 2, pp. 303-312, 1982.
- [7]. Chou K. C. and R. B. Corotis, *Simulation of hourly wind speed and array wind power*, Solar Energy, Vol. 26, No. 3, pp. 199-212, 1981.
- [8]. Daniel A. R. and Chen A. A., *Stochastic simulation and forecasting of hourly average wind speed sequences in Jamaica*, Solar Energy, Vol. 46, No. 1, pp. 1-11, 1991.
- [9]. Jutus C. G. et al, *Nationwide assessment of potential output from wind-powered generators*, Journal of applied meteorology, Vol. 15, pp. 673-678, 1976.
- [10]. McWilliam B. and Sprevak D., *Time series models for horizontal wind*, Wind Engineering, Vol. 6, pp. 219-228, 1982.
- [11]. Nfaoui H. et al, *Stochastic simulation of hourly average wind speed sequences in Tangiers (Morocco)*, Solar Energy, Vol. 56, No. 3, pp. 301-314, 1996.

- [12]. Nguyễn Hồ Quỳnh, *Chuỗi thời gian – Phân tích và nhận dạng*, NXB Khoa học và kỹ thuật, Hà Nội, 2004.
- [13]. Nguyễn Thế Bảo, *The generation of daily and hourly wind speed sequences for Australia*, Research project No. 175, Murdoch University – Energy research institute, 1996.
- [14]. Torres J. L. et al, *Forecast of hourly average wind speed with ARMA models in Navarre (Spain)*, Solar Energy, Vol. 79, No. 1, pp. 65-77, 2004.

Faint, illegible text at the top of the page, possibly a header or title.