

ƯỚC LƯỢNG CÁC THAM SỐ PHÂN BỐ WEIBULL CHO NGHIÊN CỨU CHẾ ĐỘ GIÓ PHÍA NAM VỊNH NHA TRANG

Trần Văn Chung, Võ Văn Quang, Nguyễn Văn Long

Viện Hải dương học

Tóm tắt: Bộ dữ liệu gió 10 m trên mặt biển được cung cấp bởi dữ liệu khí tượng NCEP CFSR tương ứng với khoảng thời gian 42 năm (1979-2020), với tần suất hàng giờ. Mục đích của bài báo này là tìm ra phương pháp số hiệu quả nhất cho hàm phân bố Weibull đối với chế độ gió ở vùng biển Nha Trang trên cơ sở phân tích và so sánh hiệu quả của 10 phương pháp số, cụ thể là phương pháp thực nghiệm của Justus (EMJ), phương pháp thực nghiệm của Lysen (EML), phương pháp Mô-moment (MoM), phương pháp đồ họa (GM), phương pháp Mabchour (MMab), phương pháp tham số dạng năng lượng (EPFM), phương pháp khả năng xảy ra tối đa (MLM), phương pháp khả năng xảy ra tối đa đã sửa đổi (MMLM), phương pháp năng lượng tương đương (EEM) và phương pháp khả năng xảy ra tối đa thay thế (AMLM). Theo kết quả phân tích, phương pháp MMab cho kết quả phù hợp nhất với chế độ gió ở phía Nam vịnh Nha Trang.

Từ khóa: Trường gió; hoa gió; phân bố xác suất gió; NCEP CFSR.

1. Đặt vấn đề

Gió trên bề mặt biển là một tham số điều khiển quan trọng để xác định các thông số của nhiều quá trình tương tác giữa khí-biển như thông lượng ẩn và cảm nhiệt biển-khí và tốc độ chuyển giao biển – khí của carbon dioxide, dòng động lượng và ứng suất gió trên lớp bề mặt biển. Trong các mô hình động lực học biển, tham số ứng suất gió được đưa vào và là nhân tố chính quyết định khả năng trao đổi động lực học trên bề mặt biển. Đặc biệt chế độ gió phản ánh đúng tính chất cục bộ khu vực nghiên cứu quyết định tính thành bại của kết quả mô phỏng. Như đã biết, chế độ gió tại Vịnh Nha Trang khá đặc biệt, chịu ảnh hưởng nhiều của đặc tính địa phương với ảnh hưởng của nhiều đảo che chắn lớn nhỏ. Do vậy, để nắm bắt được quy luật đúng cho gió tác động đến Vịnh Nha Trang không phải là công việc dễ dàng. Nhằm tận dụng tốt nguồn dữ liệu hiện có và cập nhật các phương pháp số tối ưu trên thế giới cho việc xác định các tham số (hình dạng và tỉ lệ) cho hàm phân bố Weibull. Đây là một việc làm có ý nghĩa, bước quan trọng cho đánh giá các tác động của gió đến quá trình thủy động lực trong vùng biển Nha Trang. Theo cập nhật mới nhất từ công trình nghiên cứu của

Kapen và cs., 2020 và theo các công trình công bố trước đây, tối đa cũng chỉ có 10 phương pháp số có thể được sử dụng để xác định các tham số trong phân tích tốc độ gió cho hàm Weibull.

2. Đối tượng và phương pháp nghiên cứu

2.1. Đối tượng nghiên cứu

Nguồn số liệu gió dùng cho phân tích được cập nhật từ cơ sở dữ liệu phân tích lại của mô hình khí hậu toàn cầu CFSR (Climate Forecast System Reanalysis) thuộc trung tâm dự báo môi trường NCEP (National Centers for Environmental Prediction) (NCEP CFSR) (Saha và cs., 2006). Theo các công trình công bố, kết quả dự báo hồi cổ và dự báo thời gian của NCEP (CFSv2) (Saha và cs., 2014) đã giúp cho các nhà quản lý khi đưa ra các quyết định phù hợp trong các lĩnh vực như quản lý nước của các lưu vực sông, nông nghiệp, giao thông vận tải, năng lượng, khai thác nguồn năng lượng sạch (gió...), các nguồn năng lượng bền vững khác, cũng như dự báo tai biến thiên nhiên như dự báo mùa mưa, bão.

Khu vực nghiên cứu số liệu hiện có cho vùng biển Nha Trang là khu vực phía Nam thể hiện trên hình 1, thời gian cho chuỗi dữ liệu gió được sử dụng từ 1979 – 2020 (42 năm), tần suất số

liệu là 1 giờ/số liệu. Cụ thể, chuỗi số liệu theo 02 nguồn chính:

(1) NCEP Climate Forecast System Reanalysis (CFSR) Selected Hourly Time-Series Products, January 1979 to December 2010 (ds093.1): số liệu được sử dụng từ 1979 – 2010 với tần suất 01 giờ/số liệu, có độ phân giải theo phương ngang 0,3°.

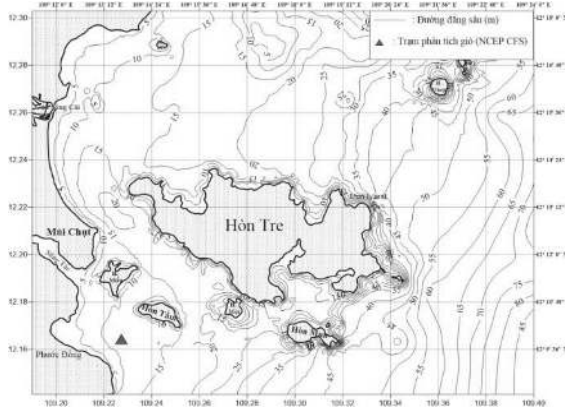
(2) NCEP Climate Forecast System Version 2 (CFSv2) Selected Hourly Time-Series Products (ds094.1): sử dụng chuỗi số liệu từ năm 2011 – 2020 với tần suất 01 giờ/số liệu, có độ phân giải theo phương ngang 0,2°.

Bộ số liệu NCEP CFSR đã được chúng tôi sử dụng trong các công trình nghiên cứu về chế độ gió cho Vịnh Vân Phong (Khánh Hòa) (Trần Văn Chung và Ngô Mạnh Tiên, 2019); cho vùng biển Cù Lao Chàm (Quảng Nam) (Trần Văn Chung và Nguyễn Văn Long, 2020) và khu vực đảo Lý Sơn (Quảng Ngãi) (Trần Văn Chung và cs., 2020).

Để phân tích đặc điểm gió khu vực phía Nam của Nha Trang, chúng tôi trích xuất bộ số liệu NCEP CFSR đã được đồng hóa về tính địa phương từ kết quả nghiên cứu của 02 đề tài:

Đề tài cấp Viện Hàn lâm KH và CN Việt Nam (VAST06.06/18-19): Nghiên cứu cơ chế lưu giữ lại và quá trình phát tán của nguồn giống cá (trứng cá-cá bột) trong vịnh Nha Trang.

Đề tài cấp nhà nước (KC.09.41/16-20): Nghiên cứu cơ chế phát tán nguồn giống và tính liên kết quần thể nguồn lợi nâng cao hiệu quả quản lý các khu bảo tồn vùng biển từ Quảng Trị đến Kiên Giang.



Hình 1. Vị trí trích xuất gió cho khu vực phía Nam Nha Trang

2.2. Hàm phân bố Weibull

Hàm phân bố dữ liệu được sử dụng phổ biến cho phân tích gió là phân bố Weibull, do nó mô tả và cung cấp các trường hợp dữ liệu gió khá tốt so với các hàm phân bố khác. Tùy thuộc vào

số lượng tham số hàm Weibull sử dụng có thể trên hai biến. Tuy nhiên, đối với dữ liệu gió, hai tham số chính được sử dụng. Nếu v là tốc độ gió (m/s), hàm phân bố xác suất Weibull $f(x)$ được mô tả:

$$f(v) = \frac{k}{c} \left(\frac{v}{c}\right)^{k-1} \exp\left(-\left(\frac{v}{c}\right)^k\right) \quad (1)$$

Trong đó, c (m/s) và k lần lượt là các tham số tỉ lệ và hình dạng phân bố Weibull. Vùng dưới đường cong của hàm mật độ xác suất được gọi là hàm phân bố tích lũy. Vì vậy, hàm phân bố tích lũy Weibull có thể đạt được bằng cách lấy một tích phân của $f(v)$, ký hiệu là $F(v)$ và cho dưới dạng:

$$F(v) = 1 - \exp\left(-\left(\frac{v}{c}\right)^k\right) \quad (2)$$

Dưới đây là 10 phương pháp số để xác định 02 tham số (c , k) trong phân bố Weibull

2.2.1. Phương pháp khả năng tối đa (MLM)

Phương pháp này là một biểu thức toán học được gọi là khả năng xảy ra của dữ liệu tốc độ gió theo chuỗi thời gian định dạng. Phương pháp MLM được giải thông qua các lần lặp số để xác định các tham số của phân bố Weibull. Các tham số hình dạng (k) và tỷ lệ (c) được tính bằng:

$$k = \left[\frac{\sum_{i=1}^n v_i^k \ln(v_i)}{\sum_{i=1}^n v_i^k} - \frac{\sum_{i=1}^n \ln(v_i)}{n} \right]^{-1}; c = \left(\frac{1}{n} \sum_{i=1}^n v_i^k \right)^{1/k} \quad (3)$$

n là số của các điểm số liệu khác không; i : khoảng đo; v_i : tốc độ gió đo tại khoảng đo i (m/s)

2.2.2. Phương pháp khả năng tối đa đã sửa đổi (MMLM)

Phương pháp MMLM chỉ được sử dụng cho dữ liệu tốc độ gió có sẵn trong phân bố Weibull. Phương pháp MMLM được giải quyết thông qua các lần lặp số để xác định các tham số của phân bố Weibull. Hai tham số Weibull được tính như sau:

$$k = \left[\frac{\sum_{i=1}^n v_i^k \ln(v_i) f(v_i)}{\sum_{i=1}^n v_i^k f(v_i)} - \frac{\sum_{i=1}^n \ln(v_i) f(v_i)}{f(v \geq 0)} \right]^{-1};$$

$$c = \left[\frac{1}{f(v \geq 0)} \sum_{i=1}^n v_i^k f(v_i) \right]^{1/k} \quad (4)$$

Trong đó: $f(v_i)$: tần số Weibull với tốc độ gió rơi trong khoảng i ; và $f(v \geq 0)$: xác suất của tốc độ gió ($v \geq 0$).

2.2.3. Phương pháp Moment (MoM):

Các hệ số Weibull k và c cho phương pháp Moment được ước lượng từ tốc độ gió trung bình $\bar{v}$ và độ lệch chuẩn (σ) của tốc độ gió. Phương pháp mômen dựa trên phép lặp số của hai phương trình sau:

$$\bar{v} = c\Gamma\left(1 + \frac{1}{k}\right); \sigma = c \left[\Gamma\left(1 + \frac{2}{k}\right) - \Gamma^2\left(1 + \frac{1}{k}\right) \right]^{1/2} \quad (5)$$

Trong đó

$$\bar{v} = \frac{1}{n} \sum_{i=1}^n v_i \quad ; \quad \sigma = \left[\frac{1}{n-1} \sum_{i=1}^n (v_i - \bar{v})^2 \right]^{1/2} \quad (6)$$

Trong đó $\Gamma()$ là hàm Gamma biểu diễn bởi :

$$\Gamma(x) = \int_0^\infty t^{x-1} \exp(-t) dt \quad (7)$$

MoM là một sự thay thế tốt để tính toán các tham số hình dạng và tỷ lệ, theo phương pháp khả năng tối đa. Tham số hình dạng và tham số tỷ lệ được xác định như sau:

$$k = \left(\frac{0,9874}{\frac{\sigma}{\bar{v}}} \right) \quad ; \quad c = \frac{\bar{v}}{\Gamma(1+\frac{1}{k})} \quad (8)$$

2.2.4. Phương pháp hệ số hình dạng năng lượng (EPFM)

Trong phương pháp này, hệ số hình dạng năng lượng được sử dụng để tính toán các tham số Weibull. Hệ số này cần được xác định trước tiên bằng các sử dụng tốc độ gió trung bình được tính toán trên dữ liệu đo, được cho là tỷ số giữa tốc độ gió lập phương trung bình và trung bình lập phương. Cụ thể:

$$E_{pf} = \frac{\bar{v}^3}{\bar{v}^3} \quad (9)$$

Trong đó $\bar{v}^3$ là trung bình của bậc ba tốc độ gió. Tham số hình dạng Weibull có thể ước tính bằng phương trình sau:

$$k = 1 + \frac{3,69}{E_{pf}^2}; \quad c = \frac{\bar{v}}{\Gamma(1+1/k)} \quad (10)$$

2.2.5. Phương pháp thực nghiệm của Lysen (EML)

Được giới thiệu bởi Lysen, là phương pháp dựa trên phương pháp độ lệch chuẩn. Tham số hình dạng “k” và tham số tỉ lệ “c” có thể được xác định như:

$$k = \left(\frac{\sigma}{\bar{v}} \right)^{-1,086}; \quad c = \bar{v} \left(0,58 + \frac{0,433}{k} \right)^{-\frac{1}{k}} \quad (11)$$

Phương pháp thực nghiệm này có thể được coi là một trường hợp đặc biệt của phương pháp mô-ment (MoM).

2.2.6. Phương pháp đồ họa (GM)

Phương pháp đồ thị được thực hiện bằng cách khớp một đường thẳng với dữ liệu tốc độ gió bằng cách sử dụng khái niệm bình phương nhỏ nhất, trong đó dữ liệu theo chuỗi thời gian phải được sắp xếp vào các khoản phân loại tốc độ gió. Thực hiện phép biến đổi logarit kép, phương trình của hàm phân bố tích lũy có thể được viết lại như sau:

$$\ln\{-\ln[1-F(v)]\} = k \ln(v) - k \ln(c) \quad (12)$$

Vẽ đồ thị $\ln(v)$ so với $\ln\{-\ln[1-F(v)]\}$, độ dốc của đường thẳng phù hợp nhất với các cặp dữ liệu là tham số hình dạng, tham số tỷ lệ sau đó thu được bởi giao điểm với tung độ y. Dữ liệu tốc độ gió đo được được sử dụng để tính toán các giá trị cho x và y; a là độ dốc, b là hệ số chặn, phương pháp hồi quy bình phương nhỏ nhất chuẩn được sử dụng để tính toán chúng tương ứng.

$$k = a \text{ và } c = \exp(-b/k) \quad (13)$$

2.2.7. Phương pháp thực nghiệm của Justus (EMJ)

Phương pháp thực nghiệm này được đề nghị bởi Justus, “k” được tính toán như sau:

$$k = 1 + \frac{3,69}{E_{pf}^2} \quad (\text{giống phương pháp hệ số}$$

hình dạng năng lượng), và “c” (tham số tỉ lệ) được đề nghị bởi Justus được xác định:

$$c = \frac{\bar{v}}{\Gamma(1+\frac{1}{k})} \quad (14)$$

2.2.8. Phương pháp Mabchour (MMab)

Phương pháp này xác định các tham số như sau [Kapen và cs., 2020]

$$k = 1 + (0,483(\bar{v} - 2))^{0,51}; \quad c = \frac{\bar{v}}{\Gamma(1+1/k)} \quad (15)$$

2.2.9. Phương pháp bình phương tối thiểu (LSM)

Trong các vấn đề kỹ thuật và toán học, phương pháp này thường hữu ích, một tương quan tuyến tính giữa hai biến được giả định trong phương pháp này và sau một số tính toán ưu việt để tối thiểu hóa mối quan hệ, biểu thức để tính toán các tham số Weibull có thể được viết dưới dạng các phương trình:

$$k = \frac{n \sum_{i=1}^n \ln v_i \times \ln[-\ln\{1-F(v_i)\}] - \sum_{i=1}^n \ln v_i \times \sum_{i=1}^n \ln[-\ln\{1-F(v_i)\}]}{n \sum_{i=1}^n \ln v_i^2 - (\sum_{i=1}^n \ln v_i)^2}$$

;

$$c = \exp \left[\frac{k \sum_{i=1}^n \ln v_i - \sum_{i=1}^n \ln[-\ln\{1-F(v_i)\}]}{nk} \right] \quad (16)$$

2.2.10. Phương pháp thay thế khả năng tối đa (AMLM)

Các tham số của phương pháp này được lấy theo dạng:

$$k = \frac{\pi}{\sqrt{6}} \left[\frac{n(n-1)}{N(\sum_{i=1}^n \ln v_i^2) - (\sum_{i=1}^n \ln v_i)^2} \right]^{\frac{1}{2}};$$

$$c = \left[\frac{1}{n} \sum_{i=1}^n (v_i)^k \right]^{\frac{1}{k}} \quad (17)$$

2.3. Tiêu chí thống kê được sử dụng để đánh giá tính phù hợp

Một số tiêu chí phải được đưa ra để xác nhận tính phù hợp của các phép tính. Để đánh giá hiệu

quả của các phương pháp số nói trên, chúng được mô tả ngắn gọn như sau:

2.3.1. Sai số độ lệch tuyệt đối trung bình (MABE)

MABE cung cấp đại lượng trung bình của tổng sai số độ lệch tuyệt đối giữa tần số ước tính và quan sát của tốc độ gió được đưa ra bởi:

$$MABE = \frac{1}{n} \sum_{i=1}^n |E_i - O_i| \quad (18)$$

2.3.2. Căn bậc hai của sai số bình phương trung bình (RMSE)

RMSE đại diện cho độ chính xác của phân bố bằng cách đo sự không phù hợp trung bình giữa các giá trị của tần suất ước tính và quan sát của tốc độ gió. Nó được đưa ra bởi:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (E_i - O_i)^2} \quad (19)$$

Giá trị cao của RMSE cho biết phân bố ước tính có vấn đề và giá trị nhỏ cho thấy phân bố ước tính phù hợp với dữ liệu.

2.3.3. Hệ số tương quan (R²)

Hệ số xác định R² xác định mối quan hệ tuyến tính giữa các giá trị tính toán từ phân bố Weibull và các giá trị tính toán từ dữ liệu đo được. Giá trị hệ số tương quan (R²) cho thấy dữ liệu đủ gần để khớp với đường hồi quy. Mối tương quan có thể được viết là:

$$R^2 = \frac{\sum_{i=1}^n (E_i - O_i)^2 - \sum_{i=1}^n (E_i - \bar{O})(O_i - \bar{O})}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (20)$$

Tiêu chí này mô tả mối tương quan giữa các giá trị của tần suất ước tính và tần suất quan sát của tốc độ gió. Giá trị của R² tiến tới giá trị 1 (đạt đến tương quan hoàn hảo), trong khi giá trị 0 cho thấy hai tập dữ liệu hoàn toàn khác nhau.

2.3.4. Kiểm định Chi-bình phương (c²)

Kiểm định Chi-bình phương (c²) được sử dụng để xác định xem có tồn tại bất kỳ sự khác biệt nào giữa tần số mong đợi và tần số quan sát được hay không. Trong kiểm định giả thuyết thống kê này, phân bố là phân bố chi-bình phương khi giả thuyết rằng là đúng. Biểu thức chi-square như được hiển thị bên dưới trong phương trình:

$$\chi^2 = \sum_{i=1}^n \left[\frac{(E_i - O_i)^2}{O_i} \right] \quad (21)$$

Phân bố xác suất được cho là chính xác khi R² lớn với c² tối thiểu.

2.3.5. Kiểm định Kolmogorov-Smirnov

Kiểm tra Kolmogorov-Smirnov về độ phù hợp được coi là xếp hạng chính xác và thừa nhận các phương pháp phù hợp với dữ liệu gió có sẵn. Quy trình của thử nghiệm Kolmogorov-Smirnov áp dụng cho biểu đồ vận tốc n với các

khoảng thời gian, xác minh giả thuyết rằng tập dữ liệu được biểu diễn bằng phân bố Weibull với các tham số hình dạng và tỷ lệ đã biết. Sau đó, nó tính toán xác suất tích lũy kết hợp với phân bố Weibull F(v) và biểu đồ thực nghiệm Fn(v).

Kiểm định Kolmogorov – Smirnov là một kiểm tra phi tham số để so sánh mẫu với một hàm mật độ xác suất tham chiếu. Đối với một hàm tích lũy F(v) đã cho, phép kiểm định Kolmogorov-Smirnov (KOL) có thể được biểu thị như sau:

$$KOL = \max |F(v) - F_n(v)| \quad (22)$$

Trong nghiên cứu này, mức ý nghĩa đã chọn cho KOL đã được xác định là 10% (tức là khả năng được chọn với loại bỏ ban đầu là 10%). Các tham số tối hạn cho mức ý nghĩa được đưa ra:

$$KOL_{0.10} = 0.8324905 - \frac{0.199103}{\sqrt{n}} = 0.026511 \times KOL + 0.002725911 \times (KOL)^2 \quad (23)$$

Tham số KOL* cho số khoảng thời gian n của biểu đồ gió được cho bởi:

$$KOL^* = KOL \times \sqrt{n} \quad (24)$$

Nếu giá trị của KOL* lớn hơn giá trị của thông số quan trọng KOL_{0.10}, thì phương pháp đã sử dụng không phù hợp với dữ liệu gió xác định.

2.4. Các tốc độ gió được tính từ các tham số Weibull

2.4.1. Giá trị ngẫu nhiên được tạo bởi phân bố Weibull

Để tạo một tập dữ liệu của phân bố Weibull theo sau với các tham số cụ thể hình dạng và tỷ lệ, giả định rằng hàm phân bố tích lũy cho bất kỳ biến liên tục nào được phân bố đồng đều trong phạm vi [0, 1] được xem xét. Do đó, một biến ngẫu nhiên có phân bố Weibull với tham số hình dạng (k) và tỷ lệ (c) cho trước có thể được tạo ra chỉ bằng cách giải tốc độ gió trong phương trình (25) như sau:

$$v = c \left[\ln \left(\frac{1}{1 - R_n} \right) \right]^{1/k} \quad (25)$$

Trong đó R_n là số ngẫu nhiên trong phạm vi [0, 1].

2.4.2. Tốc độ gió có thể xảy ra nhiều nhất và năng lượng tối đa mang theo

Có hai giá trị tốc độ gió khá hữu ích được gọi là tốc độ gió có thể xảy ra nhất (V_{mp}) và tốc độ gió mang theo năng lượng lớn nhất (V_{me}) rất cần thiết để ước tính thế năng gió. Tốc độ gió có thể xảy ra nhất đại diện cho tốc độ gió xuất hiện thường xuyên nhất đối với một phân bố xác định. Sau khi tính toán các thông số hình dạng và tỷ lệ, V_{mp} có thể được xác định là:

$$V_{mp} = c(1 - \frac{1}{k})^{1/k} \quad (26)$$

V_{me} là tham số quan trọng của đo gió được xem xét cho một địa điểm. Khuyến nghị rằng đo gió nên được chọn với tốc độ gió định mức, gần với tốc độ gió cung cấp năng lượng tối đa để có được năng lượng đầu ra tối đa. V_{me} có thể được tính bằng biểu thức sau.

$$V_{me} = c(1 + \frac{2}{k})^{1/k} \quad (27)$$

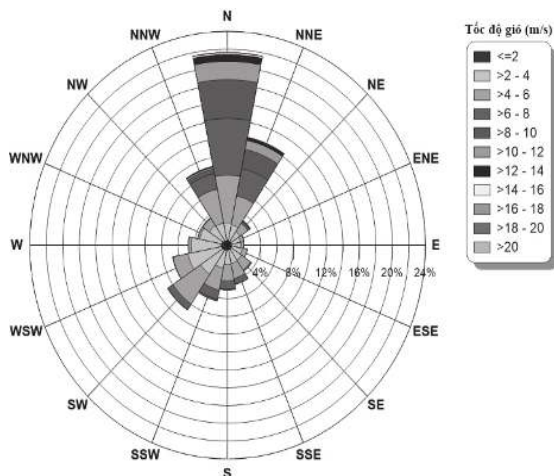
3. Kết quả nghiên cứu

Để phân tích bộ số liệu gió, sử dụng 02 nguồn dữ liệu, NCEP CFSR (01/01/1979 – 31/12/2010) và NCEP CFSv2 (01/01/2011 – 31/12/2020) cho ôp 01 giờ/số liệu (tương ứng với 42 năm). Toàn bộ sản phẩm phân tích về hướng và tốc độ gió được thể hiện khá chi tiết trên bảng 1. Để có thể nhìn thấy trực quan hơn về kết quả phân tích, chúng tôi đã thể hiện kết quả phân tích này ra dạng hoa gió (hình 1).

Bảng 1. Bảng phân cấp chế độ gió theo 16 hướng và 11 cấp độ gió tại trạm phân tích (1979 – 2020)

Hướng	$0 \leq ws < 2$	$2 \leq ws < 4$	$4 \leq ws < 6$	$6 \leq ws < 8$	$8 \leq ws < 10$	$10 \leq ws < 12$	$12 \leq ws < 14$	$14 \leq ws < 16$	$16 \leq ws < 18$	$18 \leq ws < 20$	$ws \geq 20$	Trung bình (m/s)	Tổng số liệu	Tỷ lệ (%)
N	1859	7353	19837	23403	15769	7392	2738	703	118	16	3	7,1	79191	21,5
NNE	1760	6873	12161	11462	7829	3695	1090	178	29	3	1	6,5	45081	12,3
NE	1791	4959	4165	913	267	104	29	7	1	2	0	3,9	12238	3,3
ENE	1901	4038	1553	113	25	5	2	1	0	0	0	3	7638	2,1
E	1859	4068	1279	51	22	1	0	0	0	0	3	2,9	7283	2,0
ESE	2036	4792	2306	41	4	1	0	0	0	0	0	3,1	9180	2,5
SE	1932	5642	5019	565	3	0	0	0	0	0	0	3,6	13161	3,6
SSE	1983	5853	6290	2388	115	0	0	0	0	0	0	4,1	16629	4,5
S	2047	5953	6654	3305	467	14	0	0	0	0	2	4,4	18442	5,0
SSW	2338	7461	9097	3797	657	54	0	0	0	0	0	4,4	23404	6,4
SW	2530	12121	14501	2781	104	26	6	0	0	0	0	4,1	32069	8,7
WSW	2731	15389	6717	130	21	6	4	0	0	0	0	3,4	24998	6,8
W	2836	13499	1214	61	19	0	1	0	0	0	3	2,9	17633	4,8
WNW	2652	10212	867	83	11	2	0	0	0	0	0	2,8	13827	3,8
NW	2332	7575	3554	222	47	7	4	0	0	0	0	3,3	13741	3,7
NNW	1933	7410	14482	6561	1992	705	267	100	22	5	0	5,2	33477	9,1
													367992	100,0

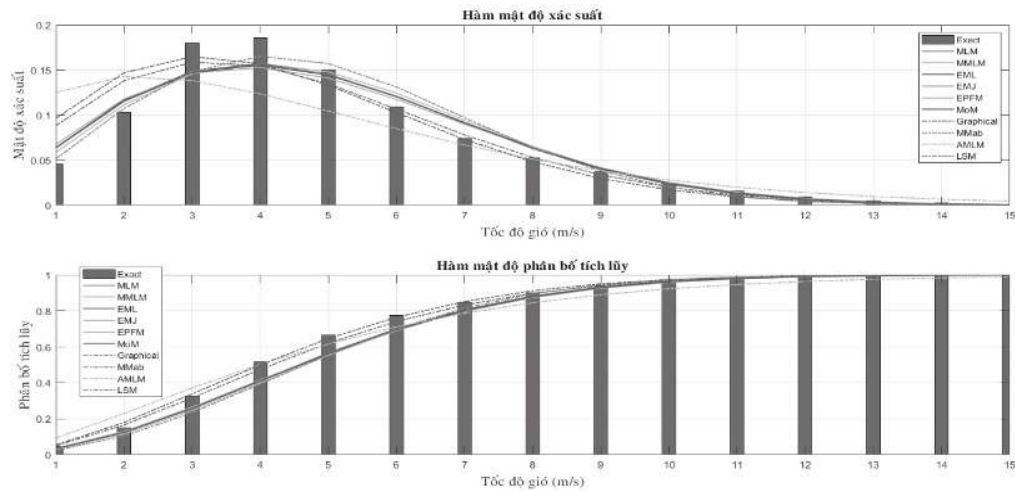
Tổng số liệu	34520	123198	109696	55876	27352	12012	4141	989	170	26	12	
Tỷ lệ (%)	9,4	33,5	29,8	15,2	7,4	3,3	1,1	0,3	0,0	0,0	0,0	100,0



Hình 1. Hoa gió được phân tích từ nguồn dữ liệu NCEP CFSR (1979 – 2020)

Phân tích trên bảng 1 và hình 1, đã thể hiện đầy đủ số lượng số liệu cho các tần suất xuất hiện của hướng và tốc độ gió. Có 2 phân loại tốc độ gió có tần suất xuất hiện nhiều nhất, đó là trong khoảng [2, 4) (m/s) với 123.198 số liệu tìm thấy chiếm 33,5 %; tiếp theo là tốc độ gió trong khoảng [4, 6) với 109.696 số liệu, chiếm 29,8 %.

Chế độ gió phía Nam Vịnh Nha Trang chịu ảnh hưởng tính khu vực thể hiện khá rõ qua các hướng gió thường xuyên xuất hiện (hình 1). Cụ thể, khi khu vực chịu tác động của mùa gió Đông Bắc, phía Nam vịnh Nha Trang chiếm ưu thế với 03 hướng gió là gió Bắc (21,5%), Bắc Đông Bắc (12,3%) và Bắc Tây Bắc (9,1%); trong khi tác động của mùa gió Tây Nam thể hiện qua 03 hướng gió, Tây Nam (8,7%), Tây Tây Nam (6,8%) và Nam Tây Nam (6,4%).



Hình 2. Các dạng hàm phân bố của tốc độ gió tại khu vực nghiên cứu

Tuy nhiên, để có thể phản ánh đúng quy luật gió, các phân tích cho từng phương pháp trích lọc tham số Weibull đã được thực hiện. Kết quả của các loại phân bố thể trên mật độ xác suất và

phân bố tích lũy so với số liệu quan sát được trình bày trên hình 2. Để đánh giá cho tính phù hợp của các phương pháp sử dụng, các tiêu chí kiểm định đã được thực hiện trên bảng 2.

Bảng 2. Đánh giá tính phù hợp qua các hàm phân bố theo các tiêu chí thống kê

Loại thông tin	Số liệu quan sát	ML M	MM LM	EML	EMJ	EPF M	MoM	GM	MMa b	AML M	LSM
v_{mean} (m/s):	4,874	4,885	4,916	4,840	4,874	4,877	4,875	4,520	4,871	4,681	4,348
V_{mp} (m/s):	3,907	3,905	4,014	3,876	3,903	3,800	3,883	3,306	4,153	2,194	3,140
V_{me} (m/s):	7,784	7,785	7,721	7,705	7,759	7,905	7,787	7,609	7,411	9,541	7,372
Phương sai	6,545	6,493	6,278	6,353	6,443	6,816	6,515	6,539	5,544	11,137	6,172
Độ lệch chuẩn (m/s)	2,565	2,548	2,506	2,521	2,538	2,611	2,552	2,557	2,354	3,337	2,484
k (tham số hình dạng):		2,005	2,057	2,008	2,008	1,948	1,996	1,832	2,182	1,423	1,812
c (tham số tỉ lệ (m/s)):		5,512	5,549	5,462	5,500	5,500	5,500	5,086	5,500	5,148	4,891
MABE		0,010	0,010	0,010	0,010	0,011	0,010	0,012	0,010	0,022	0,013
RMSE		0,056	0,055	0,054	0,055	0,059	0,056	0,069	0,053	0,128	0,078
R^2		0,947	0,948	0,951	0,948	0,941	0,947	0,918	0,953	0,721	0,896
Chi-squared		0,003	0,003	0,003	0,003	0,004	0,003	0,005	0,003	0,020	0,006
KOL0,1-KOL*		0,370	0,334	0,391	0,374	0,388	0,378	0,607	0,308	0,472	0,665

Từ kết quả tính trên bảng 2, chúng ta thấy rằng, nếu chỉ đánh giá trên 02 kiểm định sai số độ lệch tuyệt đối trung bình (MABE) và Chi-bình phương (χ^2), thì có 06 hàm phân bố thỏa mãn với giá trị tương đồng như phân bố MLM, MMLM, EML, EMJ, MoM và MMa_b; tuy nhiên nếu mở rộng phân tích thêm các tiêu chí khác thì thấy rằng hàm phân bố MMa_b thỏa mãn hầu hết các tiêu chí kiểm tra đã được đưa ra

kiểm định. Ngược lại, hàm phân bố AMLM thường không thỏa mãn các tiêu chí được kiểm định, nếu xét thêm kiểm định Kolmogorov-Smirnov tìm thấy có sự xuất hiện thêm 02 hàm phân bố LSM và GM có giá trị khá chênh lệch với các phân bố còn lại.

4. Kết luận

Kết quả phân tích chế độ gió cho khu vực Nha Trang trong 42 năm, thể hiện tốc độ gió

điển hình trong khu vực thường nằm trong khoảng < 6 m/s chiếm đến 63,32% và hướng gió chủ đạo chiếm ưu thế Bắc chiếm đến 21,5% hướng gió toàn khu vực.

Chế độ gió tại khu vực Nha Trang mang tính địa phương khá rõ ràng và thể hiện độ chính xác cao nhất theo phương pháp Mabchour (MMab). Để tránh ảnh hưởng có thể dẫn đến sai số khi đánh giá quy luật gió địa phương tại khu vực Nha Trang, không nên xác định tham số phân bố Weibull theo 03 phương pháp: phương pháp thay thế khả năng tối đa (AMLM), phương pháp bình phương tối thiểu (LSM) và phương pháp đồ họa (GM).

Từ nguồn dữ liệu đo gió 42 năm với tần suất 01 giờ/số liệu, chúng tôi đã cung cấp được bộ thông số cho các phân tích gió cho phía Nam vịnh Nha Trang. Dữ liệu này là cơ sở quan trọng là nguồn dữ liệu đầu vào cho các mô hình số trong nghiên cứu các tác động qua lại trong tương tác biển - khí trong khu vực Nha Trang./.

Tài liệu tham khảo

1. Kapen, P.T., M.J. Gouajio, D. Yemélé (2020), *Analysis and efficient comparison of ten numerical methods in estimating Weibull parameters for wind energy potential: Application to the city of Bafoussam*, Cameroon. Renewable Energy, 159, 1188-1198.

2. Saha, S., S. Moorthi, X. Wu, J. Wang, S. Nadiga, P. Tripp, D. Behringer, Y.-T. Hou, H.-y. Chuang, M. Iredell, M. Ek, J. Meng, R. Yang, M. P. Mendez, H. van den Dool, Q. Zhang, W. Wang, M. Chen and E. Becker (2014), *The NCEP Climate Forecast System Version 2*. J. Climate, 27, 2185–2208.

3. Saha, S., S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H.M. van den Dool, H.-L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Peña, S. Lord, G. White, W.P. Ebisuzaki, W. Peng and P. Xie (2006), *The NCEP Climate Forecast System*, Accepted J. Climate.

4. Trần Văn Chung và Ngô Mạnh Tiến (2019), *Đặc trưng của chế độ gió Vịnh Vân Phong qua 40 năm (1979-2018) từ nguồn số liệu NCEP CFSR*, Khoa học & Công nghệ Khánh Hòa (ISSN 1859-1981), số 4, 16 – 19 (2019).

5. Trần Văn Chung và Nguyễn Văn Long (2020), *Ước lượng các phân bố tốc độ gió cho nghiên cứu chế độ gió vùng biển Cù Lao Chàm*, Tạp chí Khoa học Đại học Khánh Hòa, ISSN: 2588-1353, tập 1, số 2 (2020), 89 – 96.

6. Trần Văn Chung, Ngô Mạnh Tiến và Nguyễn Văn Long (2020), *Nghiên cứu các đặc trưng phân bố của trường tốc độ gió tại vùng biển đảo Lý Sơn, tỉnh Quảng Ngãi*, Tạp chí Khoa học và Công nghệ biển, ISSN 1859 – 3097, 20(4A), 11-20.

ESTIMATE WEIBULL DISTRIBUTION PARAMETERS FOR WIND SPEED WITH SUITABLE FOR STUDYING WIND REGIME FOR A SITE IN THE SOUTHERN AREA OF NHA TRANG BAY

Tran Van Chung, Vo Van Quang, Nguyen Van Long

Institute of Oceanography

Abstract: The wind data set of 10 m on the sea surface is provided by meteorological data NCEP CFSR corresponding to a period of 42 years (1979–2020), with hourly frequency. The purpose of this paper is to find the most effective numerical method for the Weibull distribution function for the wind regime in Nha Trang waters on the basis of analyzing and comparing the efficiency of 10 (ten) numerical methods namely, the empirical method of Justus (EMJ), the empirical method of Lysen (EML), the method of moments (MoM), the graphical method (GM), the Mabchour's method (MMab), the energy pattern factor method (EPFM), the maximum likelihood method (MLM), the modified maximum likelihood method (MMLM), the equivalent energy method (EEM), and the alternative maximum likelihood method (AMLM). According to the analysis results, the MMab method gives the most suitable result for the wind regime in the southern area of Nha Trang bay.

Keywords: Wind field, Wind rose; Wind probability distribution; NCEP CFSR