

MỘT CÁCH TIẾP CẬN DỰ BÁO CHỨA ĐỰNG THÔNG TIN BẤT ĐỊNH CHO CÔNG SUẤT PHÁT TỪ CÁC NGUỒN NĂNG LƯỢNG TÁI TẠO

AN UNCERTAINTY FORECASTING APPROACH FOR POWER GENERATION FROM RENEWABLE ENERGY SOURCES

Lê Đình Dương^{1*}, Ngô Văn Dương², Nguyễn Thị Ái Nhi¹, Huỳnh Văn Kỳ²

¹Trường Đại học Bách khoa - Đại học Đà Nẵng

²Đại học Đà Nẵng

*Tác giả liên hệ: ldduong@dut.udn.vn

(Nhận bài: 24/8/2021; Chấp nhận đăng: 21/10/2021)

Tóm tắt - Nguồn năng lượng tái tạo (NLTT) đóng một vai trò rất quan trọng trong hệ thống năng lượng quốc gia. Để phục vụ cho công tác quản lý vận hành hệ thống điện, việc dự báo chính xác công suất phát từ các nguồn NLTT là rất cần thiết. Tuy nhiên, do đặc tính biến đổi ngẫu nhiên theo điều kiện thời tiết nên việc dự báo nguồn NLTT gặp rất nhiều khó khăn thách thức. Ngoài ra, các cách tiếp cận dự báo phổ biến hiện nay không chứa đựng thông tin bất định về nguồn NLTT dẫn đến nhiều hạn chế trong việc sử dụng kết quả dự báo trong tính toán phân tích hệ thống điện. Bài báo trình bày cách tiếp cận dự báo chứa đựng thông tin bất định về công suất phát từ các nguồn NLTT cũng như khả năng ứng dụng thực tế của kết quả đạt được từ phương pháp tiếp cận đề xuất.

Từ khóa - Năng lượng tái tạo; bất định; dự báo

1. Đặt vấn đề

Với đặc điểm tự nhiên và khí hậu tương đối thuận lợi, Việt Nam có tiềm năng lớn về NLTT như điện mặt trời, điện gió và có thể đẩy mạnh phát triển các nguồn này để đáp ứng nhu cầu sử dụng điện ngày càng tăng nhằm vừa phục vụ phát triển kinh tế - xã hội của đất nước, vừa tạo ra lợi ích cho người dân, doanh nghiệp. Hiện nay, tỷ trọng nguồn NLTT chiếm tới gần 25% tổng nguồn đã góp phần bảo đảm nguồn cung cho hệ thống. Tính đến hết tháng 4 năm 2021, cả nước có 17.000 MW điện mặt trời trang trại và 7.700 MW điện mặt trời mái nhà. Điện gió hiện đang có 612 MW đã đi vào vận hành và dự kiến từ nay đến cuối năm 2021, hệ thống điện quốc gia sẽ có thêm gần 4000 MW điện gió [1].

Từ những thông tin trên có thể thấy rằng, NLTT đóng một vai trò rất quan trọng trong hệ thống năng lượng quốc gia. Tuy nhiên, nguồn NLTT tăng cao trong thời gian qua gây rất nhiều khó khăn và áp lực cho công tác quản lý vận hành hệ thống điện và trở thành một trong những chủ đề “nóng” thu hút nhiều sự quan tâm không chỉ của ngành điện mà của toàn xã hội. Nguyên nhân chính liên quan đến nguồn NLTT là do khả năng phát điện của nguồn này phụ thuộc hoàn toàn vào điều kiện khí hậu thời tiết và thay đổi thường xuyên nên việc đảm bảo cung cầu trong hệ thống vẫn còn phụ thuộc nhiều vào các nguồn năng lượng truyền thống.

Để phục vụ cho công tác quản lý vận hành hệ thống điện, dự báo công suất phát từ nguồn NLTT đóng một vai

Abstract - Renewable energy sources (RES) play a very important role in the national energy system. In order to serve the management and operation of the power system, it is necessary to accurately forecast the power generation from RES. However, due to the random variation according to weather conditions, the work of forecasting of RES faces many difficulties and challenges. In addition, the most widely used forecasting approaches do not contain information about uncertainty of RES, leading to many limitations in using forecasting results in power system computation and analysis. This article presents an uncertainty forecasting approach for power generation from RES as well as the practical applicability of the results obtained from the proposed approach.

Key words - Renewable energy; uncertainty; forecast

trò rất quan trọng. Tuy nhiên, do đặc điểm biến đổi ngẫu nhiên theo điều kiện khí hậu thời tiết nên việc dự báo nguồn NLTT gặp rất nhiều khó khăn thách thức. Một cách tổng quát, dự báo nguồn NLTT có thể phân loại theo miền thời gian dự báo và các ứng dụng chính của kết quả dự báo được mô tả trong Bảng 1 [2]. Trong các miền thời gian dự báo, miền dự báo cho ngày vận hành tiếp theo (day-ahead) rất được quan tâm và rất cần thiết hiện nay nên được lựa chọn trình bày và minh họa cho cách tiếp cận đề xuất trong bài báo này.

Dự báo nguồn NLTT hiện nay có nhiều phương pháp và có thể phân loại thành các nhóm chính sau đây [2-4]:

- Phương pháp dựa trên đặc tính quán tính (Persistence): Phương pháp này dựa trên giả thuyết giá trị dự báo ở thời điểm t bằng với giá trị thực tế quan sát được ở thời điểm $t - 1$ trước đó. Phương pháp này thích hợp với miền dự báo cực ngắn từ vài phút đến vài chục phút;

- Phương pháp dựa trên mô hình thống kê (Statistical methods): Thích hợp với miền dự báo ngắn hạn (vài giờ), ngày tới, vài ngày tới;

- Phương pháp dựa trên mô hình vật lý (Physical methods) như mô hình dự báo thời tiết (Numerical Weather Prediction - NWP): Thích hợp với miền dự báo trung và dài hạn (vài tuần, tháng, mùa,...);

- Phương pháp dựa trên kỹ thuật học máy (Machine learning: ANN, AI,...): Thích hợp với miền dự báo ngắn hạn (vài giờ), ngày tới, vài ngày tới;

¹ The University of Danang - University of Science and Technology (Le Dinh Duong, Nguyen Thi Ai Nhi)

² The University of Danang (Ngo Van Duong, Huynh Van Ky)

- Phương pháp dựa trên mô hình lai (Hybrid methods): Thích hợp với miền dự báo ngắn hạn (vài giờ), ngày tới, vài ngày tới.

Bảng 1. Phân loại dự báo nguồn NLTT

Loại dự báo	Miền thời gian	Ứng dụng kết quả dự báo
Intra-hour/ Very short term (Cực ngắn hạn)	Từ vài phút đến 60 phút	Regulation (Điều khiển hệ thống), real-time dispatch (Điều độ thời gian thực)
Short term (Ngắn hạn)	Từ 1 đến 6 giờ tới	Short term scheduling (Lập kế hoạch vận hành ngắn hạn), congestion management (Quản lý tắc nghẽn)
Medium term (Trung hạn)	Ngày tới (day-ahead) hoặc vài ngày tới (cho tới 1 tuần)	Day-ahead scheduling (Lập kế hoạch vận hành ngày tới), unit commitment (Kế hoạch huy động công suất các tổ máy), reserve requirement (Yêu cầu dự trữ), market trading (Mua bán điện), congestion management (Quản lý tắc nghẽn), security assessment (Đánh giá an toàn hệ thống điện)
Long term (Dài hạn)	Nhiều tuần, mùa, năm, nhiều năm	Resource planning (Quy hoạch nguồn), contingency analysis (Phân tích chế độ đột biến), maintenance planning (Lập kế hoạch bảo trì), operation management (Quản lý vận hành), security assessment (Đánh giá an toàn hệ thống điện)

Tuy nhiên, các nhóm phương pháp trên đa phần áp dụng cho cách tiếp cận dự báo điểm (Point forecast) truyền thống. Với cách dự báo này, kết quả là chuỗi giá trị dự báo và mỗi giá trị (hàng số) ứng với từng thời điểm trong miền thời gian dự báo, không chứa đựng thông tin bất định về kết quả dự báo. Trong thực tế, kết quả dự báo luôn có một sai số nào đó và đối với một đại lượng biến đổi ngẫu nhiên như nguồn NLTT thì kết quả dự báo điểm không thể phản ánh hết bản chất bất định của đối tượng dự báo. Thông tin về sự bất định của công suất phát từ nguồn NLTT và các đại lượng khác trong hệ thống điện hiện nay rất cần thiết cho tính toán, quản lý vận hành hệ thống điện khi trong hệ thống tích hợp ngày càng nhiều nguồn NLTT vào. Mới đây Cục Điều tiết Điện lực (Bộ Công Thương) ban hành Quyết định số 67/QĐ-ĐLĐL ngày 10/08/2021 về Quy trình dự báo công suất, điện năng phát của các nguồn điện năng lượng tái tạo trong đó có nêu rõ yêu cầu đối với phương pháp dự báo là “Phương pháp có khả năng phân tích các yếu tố bất định”, do đó nội dung và cách tiếp cận đề xuất trong bài báo này vừa có tính khoa học vừa có tính thời sự, rất cần thiết hiện nay về vấn đề dự báo nguồn NLTT.

Bài báo có những đóng góp chính sau đây: (1) Cung cấp một “bức tranh” rõ nét về cách tiếp cận dự báo điểm (Point forecast) và cách tiếp cận dự báo chứa đựng thông tin bất định (Uncertainty forecast); (2) Đề xuất phương pháp thực hiện cụ thể về dự báo chứa đựng thông tin bất định cho công suất phát của nguồn NLTT; (3) Minh họa ứng dụng thực tế rất cần thiết hiện nay từ kết quả dự báo chứa đựng thông tin bất định trong công tác tính toán, quản lý vận hành hệ thống điện nói chung và các nguồn NLTT nói riêng.

2. Phương pháp đề xuất

2.1. Các bước thực hiện

Các bước tổng quát của cách tiếp cận dự báo công suất phát của các nguồn NLTT (điện gió, điện mặt trời) có chứa đựng thông tin bất định như sau:

- *Bước 1:* Nhập dữ liệu công suất phát thu thập được từ nguồn NLTT;

- *Bước 2:* Xác định và loại bỏ các phần tử ngoại lai (outliers) [5-7];

- *Bước 3:* Xác định và xử lý các dữ liệu bị thiếu (missing data) [8, 9];

- *Bước 4:* Chuỗi dữ liệu quan sát được từ thực tế (sau khi thực hiện Bước 2 và 3) thường không có tính dừng vì chứa đựng đặc tính ngày và mùa, các đặc tính này được xác định, trích xuất và tách ra khỏi chuỗi dữ liệu để đạt được chuỗi dữ liệu có tính dừng [10];

- *Bước 5:* Xây dựng mô hình chuỗi thời gian (mô hình ARMA) cho chuỗi dữ liệu có tính dừng thu được ở Bước 4 [10];

- *Bước 6:* Dùng mô hình từ Bước 5 để phát N_s (N_s thường rất lớn) kịch bản (scenario). Trong đó, mỗi scenario là một chuỗi dữ liệu dự báo cho miền thời gian cần dự báo, các kịch bản có xác suất bằng nhau. Các chuỗi này được tích hợp lại đặc tính ngày và mùa đã được tách ra ở Bước 4 để đạt được tập hợp N_s chuỗi dữ liệu dự báo vừa mang đặc tính của đối tượng cần dự báo vừa chứa đựng thông tin bất định của kết quả dự báo.

Mô hình ARMA và các kỹ thuật xử lý ở các bước được mô tả trong các Mục 2.2 và 2.3. Việc minh họa kết quả và bàn luận về khả năng ứng dụng thực tế của kết quả đạt được từ phương pháp tiếp cận đề xuất được trình bày ở Mục 3.

2.2. Mô hình chuỗi thời gian

Chuỗi thời gian là một chuỗi các giá trị quan sát $X = \{x_1, x_2, \dots, x_n\}$ được xếp thứ tự diễn biến thời gian với x_1 là các giá trị quan sát tại thời điểm đầu tiên, x_2 là quan sát tại thời điểm thứ 2 và x_n là quan sát tại thời điểm thứ n . Trong chuỗi thời gian thường các giá trị ở những thời điểm khác nhau có mối tương quan với nhau.

Một chuỗi thời gian có tính dừng (stationary) nếu kỳ vọng (expected value) và phương sai (variance) của nó không đổi theo thời gian và giá trị hiệp phương sai (covariance) giữa hai giai đoạn chỉ phụ thuộc vào khoảng cách giữa hai giai đoạn đó chứ không phụ thuộc vào thời gian thực sự tại đó hiệp phương sai được tính [10].

Các chuỗi thời gian trong thực tế thường không có tính dừng (nonstationary), tuy nhiên các mô hình chuỗi thời gian phổ biến thường được áp dụng cho các chuỗi thời gian có tính dừng. Do đó, khi xây dựng mô hình cho chuỗi thời gian trong thực tế thường có 2 cách tiếp cận là dùng trực tiếp mô hình cho chuỗi không có tính dừng hoặc biến đổi một chuỗi thời gian không dừng về thành chuỗi thời gian dừng và áp dụng mô hình đơn giản, dễ thực hiện hơn.

Mô hình phổ biến áp dụng cho chuỗi thời gian có tính dừng là mô hình hồi quy trung bình trượt (Auto-Regressive Moving Average, viết tắt là ARMA), mô hình này có thể dùng cho mục đích dự báo. Mô hình ARMA bao gồm hai

thành phần là hồi quy (AR) và trung bình trượt (MA), thường được ký hiệu là ARMA(p, q). Trong đó, p là bậc của AR và q là bậc của MA. Mô hình ARMA(p, q) được biểu diễn như sau [10]:

$$x_t = \sum_{i=1}^p \alpha_i x_{t-i} + \varepsilon_t - \sum_{j=1}^q \beta_j x_{t-j} \quad (1)$$

Trong đó:

$\alpha_1, \alpha_2, \dots, \alpha_p$ là tham số của thành phần AR;

$\beta_1, \beta_2, \dots, \beta_q$ là tham số của thành phần MA;

x_t là giá trị dự báo tại thời điểm t ;

$\{\varepsilon_t\}$ là chuỗi ngẫu nhiên thuần túy (purely random) hay còn gọi là nhiễu trắng (white noise), đó là chuỗi thời gian có kỳ vọng bằng 0, phương sai không đổi (đồng nhất), không có tính tương quan (uncorrelated).

Trường hợp đặc biệt khi $q = 0$, mô hình trở thành mô hình hồi quy AR(p). Khi $p = 0$, mô hình trở thành mô hình trung bình trượt MA(q). Mô hình AR thể hiện chuỗi thời gian như là sự kết hợp tuyến tính của các giá trị quá khứ. AR(p) cho ta biết cần bao nhiêu giá trị quá khứ để bao gồm vào mô hình. Mô hình MA bao gồm các thành phần bị trễ của quá trình nhiễu. Mô hình ARMA cho chuỗi thời gian có tính dừng có thể được ước lượng theo phương pháp Box-Jenkins được mô tả chi tiết trong tài liệu [10].

2.3. Kỹ thuật tiền xử lý dữ liệu

Phần này trình bày các kỹ thuật xử lý chuỗi dữ liệu thu thập được trong thực tế để đạt được chuỗi dữ liệu tốt và có tính dừng để có thể áp dụng được mô hình ARMA.

2.3.1. Xử lý dữ liệu bị thiếu

Dữ liệu bị thiếu (missing data) [8, 9] gây ra có thể do nguyên nhân khách quan lẫn chủ quan. Nguyên nhân khách quan có thể gây ra do khi thu thập dữ liệu thì dữ liệu không tồn tại hoặc dữ liệu bị lỗi, quá trình truyền tin và lưu trữ bị mất dữ liệu. Nguyên nhân chủ quan do tác nhân con người. Để xử lý dữ liệu bị mất thì có một số giải pháp sau: Bỏ qua không dùng dữ liệu đó, xử lý bằng tay, thay thế dữ liệu bị mất bằng giá trị trung bình của biến ngẫu nhiên, dùng các phương pháp hoặc thuật toán để suy ra dữ liệu thay thế cho dữ liệu bị mất. Khi dữ liệu bị mất không nhiều lắm thì dữ liệu có thể được suy ra từ các phương pháp nội suy để suy ra giá trị bị mất từ các dữ liệu sẵn có lân cận. Khi dữ liệu bị mất tương đối nhiều, các phương pháp dự báo có thể được dùng để suy ra dữ liệu bị mất bằng cách dùng dữ liệu sẵn có trước đó (training data) để xây dựng mô hình dự báo và dùng mô hình đó để suy ra dữ liệu bị mất. Các phương pháp xử lý dữ liệu bị mất được trình bày chi tiết trong các tài liệu [8, 9].

2.3.2. Loại bỏ các phần tử ngoại lai

Phần tử ngoại lai hay phần tử kỳ dị (outliers) [5-7] là những dữ liệu không tuân theo đặc tính chung của tập dữ liệu. Outliers gây ra bởi nguyên nhân khách quan như từ công cụ thu thập dữ liệu, lỗi trên đường truyền, giới hạn công nghệ, ... và nguyên nhân chủ quan của con người.

Các giá trị ngoại lai thường được xem như các mẫu dữ liệu đặc biệt, cách xa khỏi phần lớn dữ liệu khác trong tập dữ liệu. Các phần tử này có thể được nhận biết dựa vào

phân bố thống kê (statistical distribution-based), khoảng cách (distance-based), mật độ (density-based), độ lệch (deviation-based) [5-7].

Các phần tử ngoại lai có ảnh hưởng lớn đến độ chính xác của các mô hình mô phỏng và tính toán trong xác suất và thống kê. Phát hiện và xử lý các điểm ngoại lai là một bước quan trọng trong quá trình chuẩn bị dữ liệu.

Các phương pháp xác định và loại trừ outliers được trình bày trong [11, 12] bao gồm các nhóm phương pháp: Phân tích giá trị cực biên (Extreme Value Analysis), dựa trên các mô hình xác suất thống kê (Probabilistic and Statistical Models), dùng các mô hình tuyến tính (Linear Models), các mô hình dựa trên sự tiệm cận (Proximity-based Models), các mô hình lý thuyết thông tin (Information Theoretic Models) ..., trong đó mỗi phương pháp có những ưu nhược điểm riêng.

Phương pháp được sử dụng hiệu quả trong thực tế được trình bày trong tài liệu [13] được áp dụng trong bài báo, trong đó outliers được phát hiện và loại ra khỏi tập dữ liệu dựa vào hàm phân bố của tập dữ liệu và giá trị significance level (mức ý nghĩa thống kê).

2.3.3. Xử lý chuỗi dữ liệu không có tính dừng

Yêu cầu của quá trình ngẫu nhiên được mô tả bằng mô hình ở (1) là phải có tính dừng. Tuy nhiên, các nguồn NLTT như điện gió, điện mặt trời phụ thuộc vào điều kiện thời tiết, khí hậu và thường có đặc tính ngày và đặc tính mùa làm cho các chuỗi dữ liệu quan sát được thường không có tính dừng. Các đặc tính này được xác định và trích xuất ra khỏi chuỗi dữ liệu ban đầu có thể tạo ra chuỗi dữ liệu mang tính dừng như sau [14]:

$$x'_t = (x_t - \mu_{t,m}) / \sigma_{t,m} \quad (2)$$

Trong đó, $\mu_{t,m}$ và $\sigma_{t,m}$ là giá trị kỳ vọng (expected value) và độ lệch chuẩn (standard deviation) ứng với thời điểm t và m , với m là khoảng thời gian khác nhau trong năm như tháng, mùa... thể hiện tính chu kỳ, tính mùa vụ của chuỗi dữ liệu; x_t và x'_t là giá trị ban đầu và giá trị sau khi xử lý tại thời điểm t .

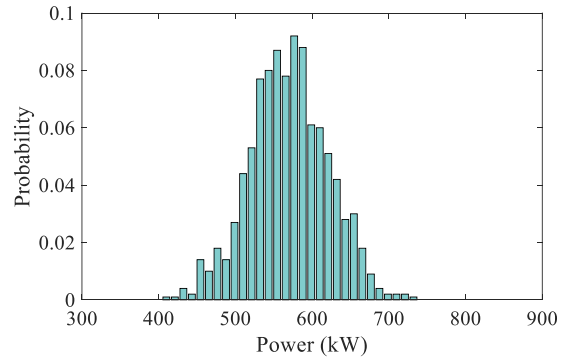
Chuỗi dữ liệu thu được $\{x'_t\}$ có thể được kiểm tra tính dừng bằng cách áp dụng các phương pháp kiểm tra phổ biến trong lĩnh vực xác suất thống kê như Augmented Dickey - Fuller (ADF) [15], Kwiatkowski - Phillips - Schmidt - Shin (KPSS) [16]. Nếu chuỗi $\{x'_t\}$ thỏa mãn thì quá trình loại bỏ đặc tính ngày và mùa đạt yêu cầu. Ngược lại, cần xác định lại đặc tính mùa cho đúng và kiểm tra lại cho đến khi chuỗi thu được có tính dừng.

3. Thử nghiệm và bàn luận về khả năng ứng dụng thực tế của kết quả đạt được từ phương pháp tiếp cận đề xuất

Để thử nghiệm cho phương pháp tiếp cận của bài báo và minh họa ứng dụng kết quả dự báo chứa đựng thông tin về sự bất định, số liệu thu thập theo giờ trong một năm từ tháng 7/2019 đến tháng 6/2020 của hệ thống điện mặt trời thực tế có công suất lắp đặt 1 MW ở Ninh Thuận được sử dụng. Chuỗi số liệu thu thập được xử lý tuần tự thông qua các bước trình bày ở Mục 2. Dữ liệu thu được có một

số giá trị ngoại lai và dữ liệu bị mất do lỗi của việc đo đếm và thu thập thông tin. Dữ liệu thiếu tương đối ít và không liên tiếp nhau trong một khoảng thời gian dài nên chỉ cần xử lý bằng phương pháp nội suy. Đối với nguồn năng lượng mặt trời ở khu vực lấy dữ liệu thử nghiệm, công suất phát của hệ thống ở các khoảng thời gian từ 0 đến 5 giờ sáng và từ 18 đến 23 giờ tối có giá trị bằng 0, các số liệu thu thập được từ 6 giờ sáng đến 17 giờ được tiếp tục xử lý trích xuất loại bỏ đặc tính ngày và mùa. Ở khu vực này, đặc tính mùa khá rõ rệt giữa mùa nắng (thường từ tháng 1 đến tháng 9 hàng năm) và mùa mưa (từ tháng 10 đến tháng 12) do đó chuỗi số liệu có thể chia ra làm 2 khoảng tương ứng với 2 mùa ứng với $m = 2$ ở công thức (2). Đối với những khu vực nói chung khó xác định đặc tính mùa hoặc để tăng tính chính xác của việc loại trừ đặc tính mùa thì có thể chia chuỗi thời gian của năm thành các khoảng ngắn hơn như chia theo từng tháng ứng $m = 12$ (được sử dụng trong bài báo này). Chuỗi số liệu sau các bước trên được kiểm tra tính dừng như trình bày ở Mục 2.3.3. Mô hình ARMA được xây dựng cho chuỗi số liệu có bậc $p = 1$ và $q = 1$ (mô hình ARMA(1,1)) và có chứa đựng thành phần ngẫu nhiên thuần túy. Mô hình được tạo mẫu và phát ra tập hợp rất lớn các kịch bản dự báo (scenario), trong đó mỗi scenario là một chuỗi dữ liệu dự báo cho miền thời gian dự báo (24 giờ của ngày tiếp theo) và xác suất đồng nhất cho các scenario. Số lượng scenario N_s thường rất lớn để đảm bảo kết quả dự báo là một tập hợp bao gồm tất cả các trường hợp có thể xảy ra trong thực tế, đây chính là thông tin về sự bất định của đối tượng dự báo (ở đây là công suất phát của nguồn điện mặt trời).

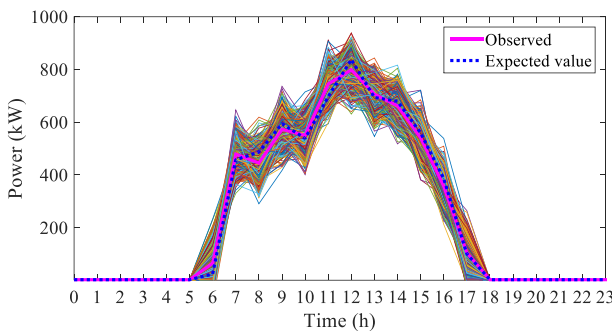
đơn vị đầu tư, quản lý nhà máy, đơn vị quản lý hệ thống điện...) trong hệ thống điện.



Hình 2. Kết quả dự báo công suất phát tại thời điểm 10 giờ sáng của ngày tiếp theo

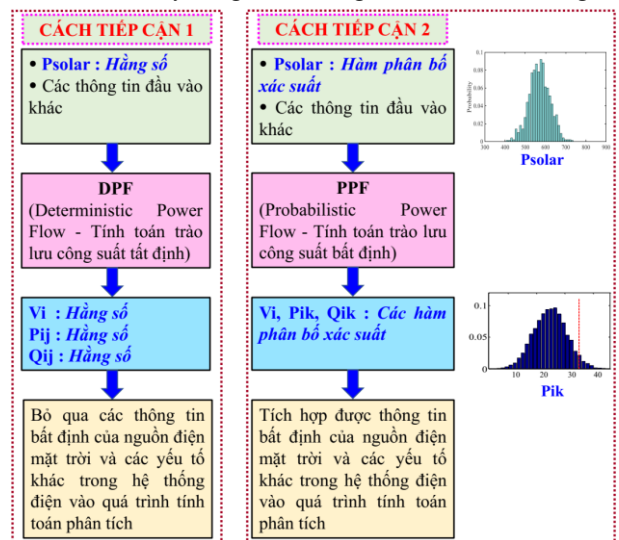
Với dạng kết quả dự báo thu được như trên, tại mỗi thời điểm bất kỳ trong miền dự báo, thông tin có được không chỉ là một giá trị cụ thể nào đó giống như các phương pháp dự báo điểm truyền thống mà là một hàm phân bố xác suất chứa đựng đầy đủ thông tin đúng với bản chất ngẫu nhiên của đại lượng dự báo. Hình 2 minh họa thông tin trích xuất ra từ kết quả dự báo tại thời điểm 10 giờ sáng. Thông tin về hàm số này có thể được tích hợp vào bài toán tính toán và phân tích hệ thống điện.

Hình 3 minh họa hai cách tiếp cận khác nhau trong tính toán phân tích hệ thống điện có sử dụng kết quả dự báo công suất phát từ nguồn điện mặt trời nói riêng, trong ví dụ minh họa ở bài báo này cũng như cho nguồn NLTT nói chung.



Hình 1. Kết quả dự báo công suất phát cho ngày tới chứa đựng thông tin bất định được biểu diễn bằng tập hợp scenario

Hình 1 vẽ kết quả dự báo cho một ngày tiếp theo bao gồm số lượng scenario $N_s = 1000$. Giá trị kỳ vọng (Expected value) của từng thời điểm cũng như số liệu đo đếm thực tế hay còn gọi số liệu quan sát thực tế (Observed) được thể hiện trên hình vẽ cho thấy, hai chuỗi số liệu này rất gần nhau chứng tỏ phương pháp ARMA rất thích hợp để sử dụng cho số liệu của bài toán dự báo. Như đã đề cập ở phần trước, mục tiêu của bài báo không đặt trọng tâm vào việc kiểm chứng tính chính xác của phương pháp dự báo chọn làm ví dụ minh họa (đó là mô hình ARMA, đây là mô hình dùng tương đối phổ biến để dự báo ngắn hạn và dự báo cho ngày tới và đã được kiểm chứng ở nhiều nghiên cứu đã công bố) mà muốn đề xuất một cách tiếp cận hiệu quả để cung cấp thêm thông tin bất định cần thiết về kết quả dự báo cho người dùng (cá nhân,



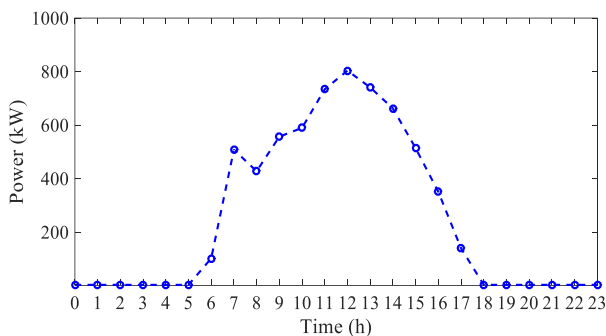
Hình 3. So sánh hai cách tiếp cận tính toán phân tích hệ thống điện liên quan đến việc sử dụng kết quả dự báo công suất phát của nguồn điện mặt trời

Cách tiếp cận 1, tương ứng với phương pháp tính toán phân tích hệ thống điện theo cách truyền thống (DPF) [17]. Thực hiện cho bài toán vận hành, ở đó sử dụng thông tin đầu vào của bài toán như thông số hệ thống, công suất từ phụ tải, kết quả từ dự báo công suất phát của nguồn điện mặt trời Psolar (Psolar ở đây thu được từ phương pháp dự báo điểm)... là các giá trị cố định. Do đó, cách tiếp cận này không tích hợp các yếu tố bất định từ thông tin đầu vào trong đó có nguồn điện mặt trời. Kết quả đầu ra của bài

toán như điện áp nút (Vi), công suất truyền trên các nhánh (công suất tác dụng và phản kháng P_{ik} và Q_{ik} trên nhánh ik) là những giá trị cố định. Do đó, không có đầy đủ thông tin về sự bất định trong hệ thống để có thể đánh giá đúng tình trạng hệ thống điện. Đó là nhược điểm lớn nhất của phương pháp này. Nhược điểm này ngày càng bộc lộ rõ hơn khi trong hệ thống điện tích hợp nhiều yếu tố bất định đặc biệt là từ nguồn NLTT.

Ngược lại, cách tiếp cận 2 sử dụng các kỹ thuật trong xác suất và thống kê để tích hợp các thông tin bất định liên quan đến bài toán tính toán phân tích hệ thống điện. Ví dụ, khi tính toán cho miền thời gian vận hành của ngày tiếp theo thông tin đầu vào của bài toán trong đó có công suất phát dự báo từ nguồn điện mặt trời được biểu diễn bằng các hàm phân phối xác suất (kết quả của cách tiếp cận dự báo xác suất trong bài báo này) chứa đựng thông tin đầy đủ về tính bất định của thông tin đầu vào. Sau đó, các thông tin này được tích hợp vào công cụ tính toán tra cứu công suất bất định (PPF) [17], cho ra kết quả là các hàm phân bố xác suất của điện áp và công suất truyền trên các nhánh. Các hàm phân bố xác suất của biến ngẫu nhiên đầu ra cho phép nhà vận hành hệ thống điện đánh giá được đúng tình trạng hệ thống và những yếu tố rủi ro có thể dẫn đến mất an toàn hệ thống để từ đó có giải pháp xử lý hiệu quả.

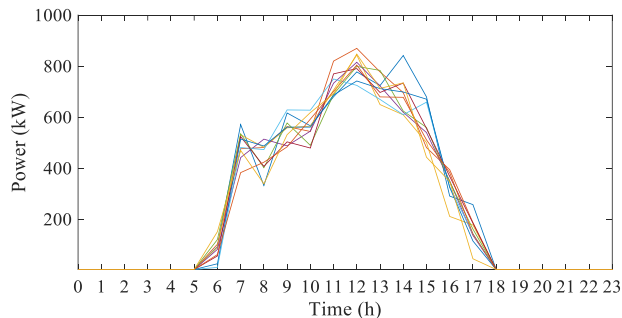
Ngoài bài toán tính toán phân tích hệ thống điện như trên, kết quả dự báo dạng tập hợp scenario có thể dễ dàng sử dụng cho công tác tính toán lập kế hoạch vận hành tối ưu ngày tới (day-ahead optimal scheduling) có tích hợp thông tin bất định của nguồn NLTT nói chung và nguồn điện mặt trời nói riêng.



Hình 4. Dạng kết quả dự báo từ phương pháp dự báo điểm cho ngày vận hành tiếp theo

Hình 4 minh họa kết quả dự báo điểm cho công suất phát từ nguồn điện mặt trời cho ngày vận hành kế tiếp, với kết quả này, các bài toán lập kế hoạch vận hành tối ưu ngày tới sẽ bỏ qua thông tin bất định. Thay vì sử dụng kết quả đầu vào từ dự báo điểm như các phương pháp truyền thống, các phương pháp lập kế hoạch tối ưu ngày tới sử dụng các kỹ thuật xác suất (stochastic day-ahead optimal scheduling) [18], có thể tích hợp yếu tố bất định đầu vào từ kết quả dự báo dạng tập hợp scenario cho công suất phát từ nguồn điện mặt trời được trình bày trong bài báo này. Ngoài ra, để giảm thời gian tính toán cho các phương pháp lập kế hoạch tối ưu ngày tới sử dụng các kỹ thuật xác suất, các kỹ thuật rút gọn tập hợp scenario (scenario reduction) [19] có thể được áp dụng. Hình 5 mô tả minh họa kết quả rút gọn số lượng scenario từ tập hợp

ban đầu rất lớn ($N_s = 1000$) xuống còn một tập hợp rút gọn bao gồm 10 scenario đại diện cho 1000 scenario trong tập hợp ban đầu nhưng thông tin về bất định suy giảm không đáng kể.



Hình 5. Kết quả dự báo được biểu diễn bằng một tập hợp đại diện

Ngoài ra, dạng kết quả dự báo nguồn NLTT có chứa đựng thông tin bất định được trình bày trong bài báo này, có thể sử dụng cho các công việc khác trong hệ thống điện như quản lý tắc nghẽn (congestion management), tính toán yêu cầu dự trữ (reserve requirement), công tác mua bán điện (market trading), quy hoạch nguồn (resource planning), phân tích an toàn hệ thống (security assessment)... dùng các thuật toán xác suất.

Cách tiếp cận trên có thể áp dụng cho các nguồn NLTT khác như điện gió và cả cho phụ tải điện. Tuy nhiên, đối với nguồn điện gió thì độ chính xác của bài toán dự báo có thể thấp hơn do độ bất định của nguồn điện này tương đối cao hơn so với nguồn điện mặt trời. Trong phạm vi bài báo, dự báo công suất phát của nguồn điện mặt trời được lựa chọn trình bày nhằm mục đích minh họa cho cách tiếp cận của phương pháp dự báo.

4. Kết luận

Bài báo trình bày cách tiếp cận dự báo chứa đựng thông tin bất định về công suất phát từ các nguồn NLTT. Bài báo cung cấp thông tin rõ nét về cách tiếp cận dự báo chứa đựng thông tin bất định và đề xuất phương pháp thực hiện cụ thể trong miền thời gian dự báo của ngày vận hành tiếp theo của hệ thống điện. Ngoài ra, ứng dụng thực tế của kết quả dự báo từ cách tiếp cận trình bày trong bài báo cho việc tính toán và công tác quản lý vận hành hệ thống điện nói chung và các nguồn NLTT nói riêng cũng được bàn luận.

Bài báo mở ra hướng tiếp cận phù hợp trong việc chuyển từ bài toán “tắt định” sang bài toán “bất định” phù hợp với tình hình hiện nay, khi các nguồn NLTT với bản chất biến đổi ngẫu nhiên được tích hợp càng nhiều và chiếm tỷ trọng lớn trong hệ thống điện.

Lời cảm ơn: Nghiên cứu này được tài trợ bởi Bộ Giáo dục và Đào tạo trong đề tài có mã số B2020-DNA-02.

TÀI LIỆU THAM KHẢO

- [1] Website: <http://baochinhphu.vn/Doanh-nghiep/Van-hanh-he-thong-dien-trong-boi-can-hung-no-nang-luong-tai-cao/430180.vgp>.
- [2] C. Monteiro, K. Keko, R. Bessa, et al., *Quick Guide to Wind Power Forecasting: State-of-the-Art*, Argonne National Laboratory, ANL/DIS-10-2, 2009.
- [3] R. Ahmed, V. Sreeram, Y. Mishra, M.D. Arif, “A review and

- evaluation of the state-of-the-art in PV solar power forecasting: Techniques and optimization”, *Renewable and Sustainable Energy Reviews*, 124, 2020, 1-26.
- [4] S. Sobri, S. Koohi-Kamali, N. A. Rahim, “Solar photovoltaic generation forecasting methods: A review”, *Energy Conversion and Management*, 156, 2018, 459-497.
- [5] J. Han, M. Kamber, J. Pei, *Data Mining: Concepts and Techniques*, 3rd Edition, The Morgan Kaufmann Series in Data Management Systems Morgan Kaufmann Publishers, July 2011.
- [6] D. L. Olson, D. Delen, *Advanced Data Mining Techniques*, Springer-Verlag, 2008.
- [7] G. J. Williams, S. J. Simoff, *Data Mining: Theory, Methodology, Techniques, and Applications*, Springer-Verlag, 2006.
- [8] D. B. Rubin, R. J. A. Little, *Statistical Analysis with Missing Data*, 2nd ed., New York: Wiley, 2002.
- [9] C. K. Enders, *Applied Missing Data Analysis*, 1st ed., New York: Guildford Press, 2010.
- [10] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, San Francisco, CA, USA: Holden-Day, 1976.
- [11] C. C. Aggarwal, *Outlier Analysis*, 2nd ed., Springer, 2016.
- [12] I. Ben-Gal, *Outlier detection, Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*, Kluwer Academic Publishers, 2005.
- [13] F. E. Grubbs, “Procedures for Detecting Outlying Observations in Samples”, *Technometrics*, vol. 11, no. 1, 1969, 1-21.
- [14] A. Papavasiliou, S. S. Oren, “Stochastic modeling of multi-area wind power production”, in *Proc. 12th Int. Conf. Probab. Methods Appl. Power Syst.*, Istanbul, Turkey, Jun. 10-14, 2012, 1-6.
- [15] J. B. Cromwell, W. C. Labys, M. Terraza, *Univariate Tests for Time Series Model*, Newbury Park, CA, USA: Sage, 1994.
- [16] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin, “Testing the null hypothesis of stationarity against the alternative of a unit root”, *J. Econometrics*, vol. 54, 1992, 159-178.
- [17] D. D. Le, A. Berizzi, C. Bovo, “A probabilistic security assessment approach to power systems with integrated wind resources”, *Renewable Energy*, 85, 2016, 114-123.
- [18] M. Nick, R. Cherkaoui, M. Paolone, “Stochastic Day-ahead Optimal Scheduling of Active Distribution Networks with Dispersed Energy Storage and Renewable Resources”, *IEEE Conference on Technologies for Sustainability (SusTech)*, 2014, 91-96.
- [19] J. Liang, W. Tang, “Scenario Reduction for Stochastic Day-Ahead Scheduling: A Mixed Autoencoder Based Time-Series Clustering Approach”, *IEEE Transactions on Smart Grid*, Vol. 12, No. 3, May 2021, 2652-2662.