



DỰ BÁO ĐỈNH MẶN TẠI CÁC TRẠM ĐO CHÍNH CỦA TỈNH CÀ MAU BẰNG MÔ HÌNH CHUỖI THỜI GIAN MỜ

Dương Tôn Đàm¹, Võ Văn Tài², Phạm Minh Trực² và Đặng Kiên Cường³

¹Trường Đại học Công nghệ Thông tin, Đại học Quốc Gia Thành phố Hồ Chí Minh

²Khoa Khoa học Tự nhiên, Trường Đại học Cần Thơ

³Trường Đại học Nông Lâm, Thành phố Hồ Chí Minh

Thông tin chung:

Ngày nhận: 29/05/2016

Ngày chấp nhận: 22/12/2016

Title:

Forecasting crest of sanility at three main stations of Ca Mau province by fuzzy time series model

Từ khóa:

Chuỗi thời gian mờ, dự báo, tiêu chuẩn thống kê, AIC, đỉnh mặn

Keywords:

Fuzzy time series, forecast, statistical criterion, AIC, crest of sanility

ABSTRACT

The article proposes fuzzy time series model in forecasting crest of sanility at the three main stations of Ca Mau province: Ca Mau (Cua Lon river), Ganh Hao (Ganh Hao river) and Ong Doc (Ong Doc river). The result obtained from this method is compared with optimal non-fuzzy time series modes which are established from original data and fuzzy one by different methods. Based on statistical criterions and realistic data, the proposed time series model shows more advantageous than the existing ones. This model is used to forecast crest of sanility for each station till 2020.

TÓM TẮT

Bài báo đề xuất mô hình chuỗi thời gian mờ trong dự báo đỉnh mặn tại 3 trạm đo chính trên địa bàn tỉnh Cà Mau: Cà Mau (sông Cửa Lớn), Gành Hào (sông Gành Hào), và Ông Đốc (sông Ông Đốc). Kết quả thực hiện được so sánh với mô hình chuỗi thời gian không mờ tối ưu được thiết lập từ dữ liệu gốc, dữ liệu mờ hóa mà nó được thiết lập theo nhiều phương pháp khác nhau. Dựa trên các tiêu chuẩn thống kê và số liệu thực tế, mô hình chuỗi thời gian đề xuất được đánh giá có nhiều ưu điểm hơn các mô hình đã có. Mô hình này cũng được sử dụng để dự báo đỉnh mặn đến năm 2020 cho mỗi trạm.

Trích dẫn: Dương Tôn Đàm, Võ Văn Tài, Phạm Minh Trực và Đặng Kiên Cường, 2016. Dự báo đỉnh mặn tại các trạm đo chính của tỉnh Cà Mau bằng mô hình chuỗi thời gian mờ. Tạp chí Khoa học Trường Đại học Cần Thơ. 47a: 68-78.

1 GIỚI THIỆU

Cà Mau là tỉnh cực nam cuối cùng của nước ta, theo đánh giá là một trong những nơi chịu ảnh hưởng nặng nề của biến đổi khí hậu. Năm 2016, tình hình hạn mặn ở nước ta nói chung và Cà Mau nói riêng diễn ra rất trầm trọng, là biểu hiện rõ ràng cho sự thay đổi bất lợi của khí hậu. Tại tỉnh Cà Mau, nền nông nghiệp được phát triển theo hai hướng chính: trồng trọt và nuôi thủy sản. Cũng như một số tỉnh ven biển của Đồng bằng sông Cửu Long, phát triển nông nghiệp dựa vào nguồn nước ngọt, hay nguồn nước mặn là bài toán đã và đang

đặt ra nhưng chưa có kết luận cuối cùng. Tại hội thảo ở Thành phố Cần Thơ năm 2016, nhiều nhà khoa học, nhà quản lý đã có những ý kiến khác nhau về vấn đề này. Đề đề ra được một cơ cấu nông nghiệp hợp lý, hiệu quả phù hợp với địa phương trong điều kiện biến đổi khí hậu phức tạp, chúng ta phải kết hợp giải quyết nhiều bài toán phức tạp. Đối với Cà Mau, việc dự báo được độ mặn cũng như mức độ xâm nhập của nó có ý nghĩa quan trọng. Khi có được những dự báo này, chúng ta mới có cơ sở cho các chiến lược phát triển nông nghiệp hợp lý. Việc lập bản đồ cụ thể các vùng cho trồng trọt hay nuôi thủy sản để mang lại hiệu quả

cao nhất cũng dựa trên cơ sở này. Ngoài việc làm cơ sở cho nền nông nghiệp, dự báo liên quan đến độ mặn cũng làm cơ sở cho sự phát triển kinh tế xã hội của vùng, hạn chế tối đa những tác động tiêu cực, phát huy được lợi thế của tự nhiên. Theo địa lí, việc xâm nhập mặn của tỉnh Cà Mau chủ yếu qua ba con sông chính đổ ra biển: sông Gành Hào (GH), sông Ông Đốc (ÔĐ) và sông Cửa Lớn (CL). Khi biết đỉnh mặn tại 3 trạm này, chúng ta sẽ biết mức độ cũng như sự xâm nhập mặn bên trong các vùng tỉnh Cà Mau. Theo tìm hiểu của chúng tôi, so với việc dự báo đỉnh lũ, dự báo về mặn ít được quan tâm hơn. Hằng năm, Trung tâm Khí tượng Thủy văn Nam Bộ có đưa ra những dự báo ngắn hạn cho toàn vùng Đồng bằng sông Cửu Long, nhưng cụ thể cho các trạm đo của tỉnh Cà Mau thì chưa được quan tâm. Do đó, bài báo này quan tâm đến việc dự báo đỉnh mặn tại các trạm đo chính của tỉnh Cà Mau.

Dự báo nói chung và dự báo trong khí tượng thủy văn nói riêng là một khoa học phức tạp. Mặc dù chúng ta đã có những thành tựu quan trọng, tuy nhiên cho đến nay nó vẫn là thách thức lớn đối với các nhà khoa học trên toàn thế giới. Trong thống kê, chúng ta có hai phương pháp thực hiện chính: hồi qui và chuỗi thời gian. Với số liệu diễn biến phức tạp như đỉnh mặn, mô hình hồi qui không phù hợp để dự báo. Về chuỗi thời gian, mô hình ARIMA (Nguyễn Thanh Sơn, 2003, Phan Văn Tân, 2005, Tô Văn Trường, 2005) đã được sử dụng phổ biến. Mô hình này tỏ ra khá hiệu quả cho những dự báo ngắn hạn của nhiều vấn đề. Điều kiện quan trọng khi dự báo tốt bằng mô hình chuỗi thời gian là tính dừng của dữ liệu. Trong thực tế các dữ liệu thực hiện thường không có tính dừng, do đó chúng ta phải xử lý vấn đề này để có một mô hình hiệu quả. Phương pháp xử lý thông thường là lấy sai phân. Việc lấy sai phân có thể làm mất đi khuynh hướng của dữ liệu, làm cho mô hình xây dựng không còn phù hợp. Làm trơn dữ liệu theo các phương pháp trung bình trượt và hàm mũ là một cách làm cổ điển để làm giảm sự biến đổi phức tạp của số liệu (Võ Văn Tài, 2015). Một phương pháp có ý nghĩa tương tự như sự làm trơn dữ liệu là sự mờ hóa. Dựa trên lý thuyết tập mờ của Zadeh (Zadeh, 1965) nhiều nhà thống kê đã đề xuất các mô hình mờ hóa số liệu (Song và Chisom, 1993, Song và Chisom, 1994, Chen, 1996, Huarng, 2001, Chen và Hsu, 2004, Singh, 2009). Khi dữ liệu được mờ hóa, các số liệu đã có một sự liên kết xác suất nhất định, nên được đánh giá có nhiều ưu điểm hơn các phương pháp làm trơn. Tuy nhiên, các mô hình này thực chất chỉ để mờ hóa dữ liệu, không thể được sử dụng để dự báo. Abbasov và Mamedova (2003) đã đề xuất mô hình để dự báo dân số nước Áo. Mô hình này chỉ phát huy hiệu quả khi chọn

được những tham số thích hợp cho từng bộ số liệu cụ thể. Dựa trên mô hình này, với một sự điều chỉnh thích hợp, chúng tôi áp dụng vào dự báo đỉnh mặn tại các trạm đo chính ở tỉnh Cà Mau mà nó tỏ ra khá phù hợp.

Phần tiếp theo của bài báo được cấu trúc như phần 2 trình bày các mô hình dự báo chuỗi thời gian không mờ và các vấn đề liên quan. Phần 3 trình bày thuật toán xây dựng chuỗi thời gian mờ và tiêu chuẩn đánh giá. Phần 4 xây dựng mô hình dự báo chuỗi thời gian mờ và so sánh với các mô hình chuỗi thời gian không mờ cho từng trạm đo. Việc dự báo đỉnh mặn của các trạm đo đến năm 2020 cũng được thực hiện trong phần này. Cuối cùng là kết luận của bài viết.

2 MÔ HÌNH DỰ BÁO CHUỖI THỜI GIAN KHÔNG MỜ VÀ CÁC VẤN ĐỀ LIÊN QUAN

2.1 Các mô hình

i) Mô hình tự hồi qui bậc p (AR(p)):

$$y_t = \phi_0 + \sum_{i=1}^p \phi_i y_{t-i} + u_t, \quad (1)$$

trong đó ϕ_i là các hệ số ước lượng của mô hình, u_t là số hạng đảm bảo tính ồn trắng.

ii) Mô hình trung bình di động bậc q (MA(q)):

$$y_t = \beta_0 + \sum_{i=1}^q \beta_i u_{t-i}, \quad (2)$$

trong đó β_i cũng là các hệ số ước lượng của mô hình và u_i cũng là số hạng đảm bảo tính ồn trắng.

iii) Mô hình tự hồi qui và trung bình di động (ARMA(p,q)):

$$y_t = \varphi_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + u_t + \beta_1 u_{t-1} + \beta_2 u_{t-2} + \dots + \beta_q u_{t-q}. \quad (3)$$

Một quá trình ARMA(p,q) sẽ có quá trình tự hồi qui bậc p và quá trình trung bình di động bậc q.

iv) Mô hình trung bình di động tổng hợp với tự hồi qui ARIMA(p,d,q):

$$y_t = \delta + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p} + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q} + e_t, \quad (4)$$

Trong đó $\alpha_i, i=1,2,\dots,p$ là tham số tự hồi qui; $\varepsilon_{t-j}, j=1,2,\dots,q$ là tham số trung bình di động; $\delta = \mu(\beta_1 + \beta_2 + \dots + \beta_q)$; μ là giá trị trung bình của chuỗi thời gian; e_t là sai số dự báo ($e_t = \hat{y}_t - y_t =$ số liệu dự báo - số liệu thực tế); (Mô hình được xây dựng khi dữ liệu ở đây đã được lấy sai phân theo bậc d).

Mô hình ARIMA được đánh giá có nhiều ưu điểm khi dự báo số liệu theo thời gian, bởi vì nó không cần nhiều dữ liệu như các phương pháp khác và thực tế đã thành công trong việc dự báo ngắn hạn của nhiều vấn đề.

2.2 Phương pháp Box-Jenkin

Mục tiêu của phương pháp này là tìm trong số tất cả các dạng của mô hình ARIMA một mô hình thích hợp nhất với bộ số liệu của hiện tượng đang nghiên cứu. Các bước thực hiện của phương pháp này như sau:

Bước 1: Nhận dạng mô hình

Đây là bước quan trọng và khó nhất trong phương pháp này. Nhận dạng mô hình ARIMA(p,d,q) là tìm các giá trị thích hợp của p, d, q với p là bậc tự hồi qui, d là bậc lấy sai phân, q là bậc trung bình trượt. Các nguyên tắc sau đây nhằm xác định thông số p, d, q của mô hình ARIMA:

- Xác định p, q của mô hình ARIMA nhờ vào đồ thị tự tương quan riêng phần mẫu (SPAC) và tự tương quan(SAC).
- Chọn mô hình AR(p) nếu đồ thị SPAC có giá trị cao tại độ trễ 1, 2, ..., p và giảm nhiều sau p và dạng hàm SAC giảm dần.

Chọn mô hình MA(q) nếu đồ thị SAC có giá trị cao tại độ trễ 1, 2, ..., q và giảm nhiều sau q và dạng hàm SPAC giảm dần. Cụ thể chúng ta có bảng tổng kết sau:

Bảng 1: Xác định tham số trong mô hình ARIMA

| Loại mô hình | Dạng đồ thị SAC | Dạng đồ thị SPAC |
|---------------|-----------------|------------------|
| AR(p) | Giảm dần | Có đỉnh ở p |
| MA(q) | Có đỉnh ở q | Giảm dần |
| ARMA(p,q) | Giảm dần | Giảm dần |

Bước 2. Ước lượng các hệ số của mô hình

Các hệ số của mô hình sẽ được xác định theo phương pháp bình phương tối thiểu:

$$\sum (y_t - \hat{y}_t)^2 \rightarrow \min,$$

trong đó y_t và \hat{y}_t lần lượt là số liệu thực tế và dự báo ở thời điểm t .

Bước 3: Kiểm định mô hình

Sau khi các thông số của mô hình được xác định, ta sẽ thực hiện kiểm định trên các kết quả của ước lượng thu được. Cụ thể:

- Các hệ số của mô hình phải khác 0 (kiểm định t). Nếu có một hay nhiều hệ số không thỏa

mãn ta sẽ loại bỏ ra mô hình AR hay MA đang xét.

- Phần dư của mô hình là sai số giữa kết quả dự báo và số liệu thực tế phải là một ồn trắng. Trong thực tế việc này được thực hiện bằng cách vẽ đồ thị ACF của mô hình. Nếu ACF mẫu nằm trong giới hạn $\pm 1.96\sqrt{N}$ (khoảng tin cậy 95%), trong đó N là số các quan sát, thì phần dư là ồn trắng, chứng tỏ mô hình xây dựng phù hợp.

Bước 4: Dự báo

Khi mô hình đã được kiểm định, ta có thể dùng nó để dự báo cho tương lai.

3 CHUỖI THỜI GIAN MỜ

3.1 Thuật toán

Dựa vào mô hình của Abbasov-Mamenova chúng tôi có một sự cải tiến về sự tịnh tiến thời gian. Cụ thể mô hình đề nghị gồm 6 bước như sau:

Bước 1: Xác định tập nền U chứa đoạn thời gian giữa các biến đổi nhỏ nhất và lớn nhất của chuỗi dữ liệu khảo sát.

Bước 2: Chia tập U thành n đoạn có độ dài bằng nhau, chứa các giá trị biến đổi tương ứng với tỷ lệ biến đổi khác nhau của dữ liệu. Tính các giá trị trung bình của từng đoạn ($u_m^i, i = 1, \dots, n$).

Bước 3: Biểu diễn cấp độ mờ các giá trị biến đổi dữ liệu theo ngôn ngữ, khi đó các tập mờ được thiết lập theo nguyên tắc sau:

$$A^t = [\mu_{A_i}(u_i)/u_i], u_i \in U, \mu_{A_i} \in [0,1],$$

$$\mu_{A_i}(u_i) = \frac{1}{1+[c \times (u - u_m^i)]^2}, \tag{5}$$

trong đó

A^t là mờ hóa các biến của năm t ,

C là hằng số tự chọn sao cho $\mu_{A_i}(u_i) \in [0,1]$,

U là các biến đổi của từng năm, hoặc là giá trị trung bình của từng đoạn thứ i ,

u_m^i là giá trị trung bình của từng đoạn thứ i .

Bước 4: Mờ hóa các dữ liệu đầu vào hay chuyển đổi các giá trị số vào các giá trị mờ. Việc này cho phép phản ánh sự tương ứng giá trị định lượng hay định tính của tỷ lệ biến đổi dữ liệu tiêu biểu trong giá trị của hàm quan hệ.

Bước 5: Lựa chọn tham số w ($1 < w < n, n$ là số năm của dữ liệu ban đầu) tương ứng với đoạn thời gian trước khi sang năm có liên quan, tính toán ma trận R các mối quan hệ mờ.

$$[i, j] = O^w[i, j] \cap K[1, j] = \begin{bmatrix} R_{11} & R_{12} & \dots & R_{1j} \\ R_{21} & R_{22} & \dots & R_{2j} \\ \dots & \dots & \dots & \dots \\ R_{i1} & R_{i2} & \dots & R_{ij} \end{bmatrix}$$

$$F(t) = \left[\max_{\dots} (R_{11}, R_{21}, \dots, R_{i1}) \max(R_{12}, R_{22}, \dots, R_{i2}) \dots \max(R_{1j}, R_{2j}, \dots, R_{ij}) \right]$$

trong đó

$$i = 1, 2, \dots, w; j = 1, 2, \dots, n,$$

O^w : là ma trận các tập mờ từ năm $t - w$ đến $t - 2$,

K : là tập mờ của năm $t - 1$,

\cap : là so sánh từng phần tử của hàng thứ $i = 1, \dots, n$ đối với từng phần tử tương ứng của K theo tiêu chuẩn min,

$F(t)$: là tập mờ tại thời điểm t .

Bước 6: Dự báo sự biến đổi của dữ liệu của năm tiếp theo $V(t)$ theo công thức:

$$V(t) = \frac{\sum_{i=1}^w \mu_t(u_i) \times u_m^i}{\sum_{i=1}^w \mu_t(u_i)} \quad (6)$$

trong đó $\mu_t(u_i)$ là phần tử của $F(t)$.

Kết quả dự báo cho năm thứ $t-1$ được tính theo công thức sau:

$$Y(t - 1) = N(t - 1) + V(t), \quad (7)$$

trong đó

$N(t - 1)$ là đỉnh mặn thực tế của năm $t - 1$,

$Y(t - 1)$ là đỉnh mặn dự báo của năm $t - 1$,

$V(t)$ là độ sai lệch từ năm $t - 1$ đến năm t .

Trong mô hình trên, việc chọn hàm thuộc $\mu_{A_i}(u_i)$ rất quan trọng vì nó sẽ ảnh hưởng đến kết quả dự báo. Trong xác định hàm thuộc bởi công thức (5) vai trò của hằng số C có ý nghĩa quyết định. C được chọn sao cho sai số trung bình của số liệu quá khứ và dự báo của mô hình là nhỏ nhất. Hiện tại, chúng ta chưa có một công thức tối ưu cho việc chọn C với tất cả bộ dữ liệu. Trong áp dụng của bài viết này, chúng tôi thử nhiều lần giá trị của C để chọn được hằng số phù hợp nhất.

Chúng tôi đã viết chương trình chi tiết để thực hiện mô hình trên với phần mềm R. Chương trình này đã được áp dụng một cách hiệu quả trong các ứng dụng của mục 4.

3.2 Đánh giá mô hình dự báo

Cho đến hiện tại có rất nhiều tiêu chuẩn để

đánh giá một mô hình dự báo như hệ số xác định, hệ số Schwarz, hệ số AIC, ... Có rất nhiều bình luận về vấn đề này, tuy nhiên theo Bozdogan (2000) và Abbasov (2002), chúng ta chưa thể khẳng định tiêu chuẩn nào được xem là tốt nhất. Trong bài viết này, chúng tôi sử dụng kết hợp các tiêu chuẩn sau để đánh giá các mô hình xây dựng:

i) Hệ số AIC (Akaike Information Criterion)

AIC được tính bởi công thức sau:

$$AIC = \frac{SSR}{n} \exp\left(\frac{2k}{n}\right), \quad (8)$$

trong đó k là số biến ước lượng (bao gồm cả hệ số chặn), n là số mẫu quan sát, SSR tổng bình phương sai số của phần dư. Khi so sánh hai hay nhiều mô hình, mô hình nào có AIC thấp nhất thì mô hình đó tốt hơn.

ii) Sai số tuyệt đối trung bình

Gọi y_i và \hat{y}_i lần lượt là giá trị thực tế và dự báo của biến y tại thời điểm $t_i, i = 1, 2, \dots, n$. Khi đó sai số tuyệt đối trung bình khi dự báo được xác định bởi:

$$M = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (9)$$

ME là một thước đo rất hữu ích khi người phân tích muốn đo lường sai số có cùng một đơn vị tính với dữ liệu gốc. ME càng nhỏ thì việc dự báo càng chính xác.

iii) Đồ thị phân tán và số liệu thực tế

Tiêu chuẩn được xem xét sau cùng, cũng như sự phù hợp của mô hình được đánh giá qua số liệu thực tế và nội suy từ mô hình đã xác định. Nó cũng được đánh giá qua số liệu thực tế mà mô hình lấy để làm minh chứng kiểm tra. Nếu kết quả có được từ mô hình và thực tế càng gần nhau thì mô hình càng được đánh giá phù hợp hơn.

4 DỰ BÁO ĐỈNH MẶN TẠI CÁC TRẠM ĐO CHÍNH CỦA TỈNH CÀ MAU

4.1 Giới thiệu

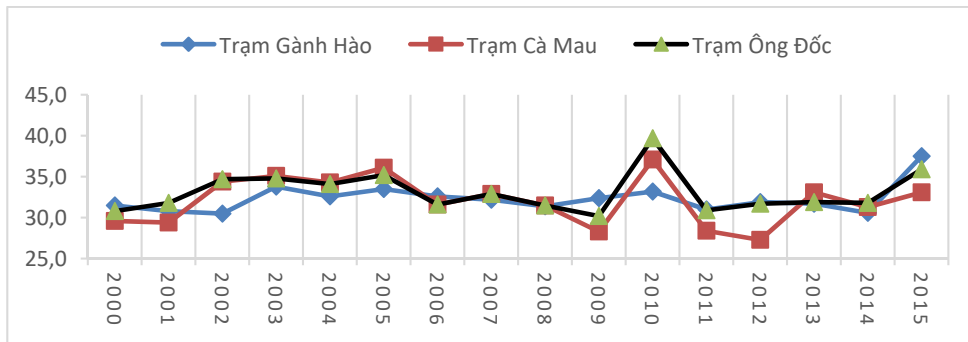
Như đã giới thiệu, trên địa bàn tỉnh Cà Mau có 3 con sông lớn đổ ra biển, đó là sông CL, sông GH và sông OĐ. Tương ứng trên 3 con sông này, chúng ta đã đặt các trạm CM, GH và OĐ quan trắc đo độ mặn. Đây là những trạm đo chính để đánh giá mức độ mặn trong tỉnh, bởi vì từ số liệu này chúng ta có thể biết được mức độ xâm nhập mặn những vùng khác nhau trong tỉnh.

4.2 Tổng quan việc thực hiện

a. Số liệu

Từ số liệu đỉnh mặn theo tháng, tại 3 trạm quan

sát, chúng tôi rút ra số liệu đỉnh mặn theo năm cho mỗi trạm. Số liệu này được cung cấp bởi Đài Khí tượng Thủy văn khu vực Nam Bộ, đặt tại Thành phố HCM, giai đoạn 2000 – 2015.



Hình 1: Đỉnh mặn tại 3 trạm chính của Cà Mau giai đoạn 2000- 2015

Số liệu cung cấp cho thấy đỉnh mặn xuất hiện tại trạm CM và OĐ đều tập trung vào tháng 4 và 5. Tại trạm CM, 75% đỉnh mặn xuất hiện ở tháng 4 và 25% còn lại xuất hiện vào tháng thứ 5. Tại trạm OĐ, 62.5% đỉnh mặn xuất hiện vào tháng tư, 18.8% xuất hiện vào tháng năm, phần còn lại chủ yếu xuất hiện vào tháng ba. Trong khi đó, tại trạm GH, đỉnh mặn thường xuất hiện muộn hơn và phân bố thời gian dài hơn. Tại đây đỉnh mặn có thể xuất hiện từ tháng tư đến tháng bảy, trong đó phần lớn tập trung ở tháng 5 (37.5%) và 6 (37.5%).

b. Các bước thực hiện

Từ số liệu, chúng tôi lần lượt dự báo độ mặn tại trạm CM, OĐ và GH. Tại mỗi trạm, các bước được thực hiện như sau:

i) Xác định hệ số tương quan của đỉnh mặn từng trạm với đỉnh lũ tại Tân Châu trên sông Hậu,

lượng mưa trung bình của tỉnh. Hệ số tương quan của đỉnh mặn từng đôi trạm và tần suất xuất hiện đỉnh mặn cũng được xem xét.

ii) Thực hiện cụ thể việc dự báo đỉnh mặn tại mỗi trạm bằng mô hình chuỗi thời gian mờ.

iii) Sử dụng số liệu gốc và số liệu mờ hóa tối ưu, tìm mô hình ARIMA phù hợp cho mỗi trường hợp.

iv) Nhận xét, so sánh mô hình chuỗi thời gian mờ với các mô hình khác đã thực hiện và tiến hành dự báo đỉnh mặn đến năm 2020.

4.3 Một số phân tích thống kê từ số liệu

Từ số liệu, chúng ta tính được hệ số tương quan giữa đỉnh mặn tại 3 trạm đo chính của Cà Mau, lượng mưa trung bình trong tỉnh và đỉnh lũ sông Hậu tại trạm đo Châu Đốc như sau:

Bảng 2: Hệ số tương quan giữa đỉnh mặn, lượng mưa TB và đỉnh lũ tại 3 trạm

| | | Đỉnh mặn | Lượng mưa TB | Đỉnh lũ |
|----|--------------|----------|--------------|---------|
| OĐ | Đỉnh mặn | 1 | | |
| | Lượng mưa TB | -0.120 | 1 | |
| | Đỉnh lũ | -0.375 | -0.509 | 1 |
| CM | Đỉnh mặn | 1 | | |
| | Lượng mưa TB | -0.217 | 1 | |
| | Đỉnh lũ | -0.238 | -0.153 | 1 |
| GH | Đỉnh mặn | 1 | | |
| | Lượng mưa TB | -0.247 | 1 | |
| | Đỉnh lũ | -0.571 | -0.056 | 1 |

Bảng 2 cho thấy, đỉnh mặn tại 3 trạm khảo sát có tương quan tỉ lệ nghịch với lượng mưa trung bình của tỉnh và đỉnh lũ sông Hậu. Dựa vào bảng trên, ta cũng thấy giữa đỉnh mặn với đỉnh lũ có sự tương quan lớn hơn so với lượng mưa trung bình của tỉnh. Đỉnh mặn tại GH và OĐ bị tác động mạnh hơn bởi lũ so với trạm CM, trong đó đỉnh mặn tại GH bị tác động nhiều nhất. Nhận xét này cũng phù

hợp với đặc điểm địa lí của các con sông. Tuy nhiên, do các hệ số tương quan không cao, nên chúng ta cũng kết luận: ngoài đỉnh lũ và lượng mưa trung bình, đỉnh mặn còn bị tác động bởi những nhân tố khác.

Chúng ta cũng có hệ số tương quan giữa đỉnh mặn tại 3 trạm như sau:

Bảng 3: Hệ số tương quan giữa các đỉnh mặn tại 3 trạm

| | GH | CM | OD |
|-----------|-----------|-----------|-----------|
| GH | 1 | | |
| CM | 0.386 | 1 | |
| OD | 0.532 | 0.841 | 1 |

Bảng 3 cho thấy, hệ số tương quan đỉnh mặn tại 3 trạm ở mức độ trung bình đến khá cao theo chiều thuận. Đặc biệt hệ số tương quan cao nhất giữa CM và OD (0.841). Điều này cũng có thể do về mặt địa

lý, vì 2 trạm này nằm trên 2 sông CL và OD có cửa đều hướng về bờ biển phía Tây Nam, trong khi cửa sông GH ở biển Đông.

4.4 Dự báo đỉnh mặn tại trạm đo Cà Mau

a. Mô hình chuỗi thời gian mờ

Bước 1. Ta có giá trị lớn nhất và nhỏ nhất của biến đổi chuỗi thời gian đỉnh mặn Cà Mau từ -8.7 đến 8.8. Do đó, tập nền U được xác định là giá trị trong đoạn $[-8.7; 8.8]$. Chọn hằng số $C=0.09$.

Bảng 4: Sự biến đổi và mờ hóa các biến đỉnh mặn CM giai đoạn 2000-2015

| Năm | Đỉnh mặn | Biến đổi | Mờ hóa các biến đổi |
|------------|-----------------|-----------------|--|
| 2000 | 29.6 | | |
| 2001 | 29.4 | -0.2 | $A^{2001} = \left(\frac{0.69}{u_1}, \frac{0.82}{u_2}, \frac{0.93}{u_3}, \frac{0.99}{u_4}, \frac{0.99}{u_5}, \frac{0.91}{u_6}, \frac{0.79}{u_7}, \frac{0.66}{u_8}\right)$ |
| 2002 | 34.4 | 5 | $A^{2002} = \left(\frac{0.44}{u_1}, \frac{0.53}{u_2}, \frac{0.65}{u_3}, \frac{0.77}{u_4}, \frac{0.89}{u_5}, \frac{0.98}{u_6}, \frac{1.00}{u_7}, \frac{0.94}{u_8}\right)$ |
| 2003 | 35.1 | 0.7 | $A^{2003} = \left(\frac{0.64}{u_1}, \frac{0.77}{u_2}, \frac{0.89}{u_3}, \frac{0.98}{u_4}, \frac{1.00}{u_5}, \frac{0.95}{u_6}, \frac{0.84}{u_7}, \frac{0.72}{u_8}\right)$ |
| 2004 | 34.3 | -0.8 | $A^{2004} = \left(\frac{0.73}{u_1}, \frac{0.85}{u_2}, \frac{0.95}{u_3}, \frac{1.00}{u_4}, \frac{0.97}{u_5}, \frac{0.88}{u_6}, \frac{0.76}{u_7}, \frac{0.63}{u_8}\right)$ |
| 2005 | 36.1 | 1.8 | $A^{2005} = \left(\frac{0.58}{u_1}, \frac{0.70}{u_2}, \frac{0.83}{u_3}, \frac{0.94}{u_4}, \frac{1.00}{u_5}, \frac{0.98}{u_6}, \frac{0.90}{u_7}, \frac{0.78}{u_8}\right)$ |
| 2006 | 31.6 | -4.5 | $A^{2006} = \left(\frac{0.93}{u_1}, \frac{0.99}{u_2}, \frac{0.99}{u_3}, \frac{0.91}{u_4}, \frac{0.79}{u_5}, \frac{0.67}{u_6}, \frac{0.55}{u_7}, \frac{0.45}{u_8}\right)$ |
| 2007 | 32.90 | 1.30 | $A^{2007} = \left(\frac{0.61}{u_1}, \frac{0.73}{u_2}, \frac{0.86}{u_3}, \frac{0.96}{u_4}, \frac{1.00}{u_5}, \frac{0.97}{u_6}, \frac{0.87}{u_7}, \frac{0.75}{u_8}\right)$ |
| 2008 | 31.50 | -1.40 | $A^{2008} = \left(\frac{0.76}{u_1}, \frac{0.88}{u_2}, \frac{0.97}{u_3}, \frac{1.00}{u_4}, \frac{0.95}{u_5}, \frac{0.85}{u_6}, \frac{0.72}{u_7}, \frac{0.60}{u_8}\right)$ |
| 2009 | 28.30 | -3.20 | $A^{2009} = \left(\frac{0.86}{u_1}, \frac{0.96}{u_2}, \frac{1.00}{u_3}, \frac{0.96}{u_4}, \frac{0.87}{u_5}, \frac{0.74}{u_6}, \frac{0.62}{u_7}, \frac{0.51}{u_8}\right)$ |
| 2010 | 37.10 | 8.80 | $A^{2010} = \left(\frac{0.31}{u_1}, \frac{0.38}{u_2}, \frac{0.46}{u_3}, \frac{0.56}{u_4}, \frac{0.68}{u_5}, \frac{0.80}{u_6}, \frac{0.92}{u_7}, \frac{0.99}{u_8}\right)$ |
| 2011 | 28.40 | -8.70 | $A^{2011} = \left(\frac{0.99}{u_1}, \frac{0.92}{u_2}, \frac{0.80}{u_3}, \frac{0.68}{u_4}, \frac{0.56}{u_5}, \frac{0.46}{u_6}, \frac{0.38}{u_7}, \frac{0.31}{u_8}\right)$ |
| 2012 | 27.30 | -1.10 | $A^{2012} = \left(\frac{0.74}{u_1}, \frac{0.87}{u_2}, \frac{0.96}{u_3}, \frac{1.00}{u_4}, \frac{0.96}{u_5}, \frac{0.86}{u_6}, \frac{0.74}{u_7}, \frac{0.61}{u_8}\right)$ |
| 2013 | 33.10 | 5.80 | $A^{2013} = \left(\frac{0.41}{u_1}, \frac{0.50}{u_2}, \frac{0.60}{u_3}, \frac{0.72}{u_4}, \frac{0.85}{u_5}, \frac{0.95}{u_6}, \frac{1.00}{u_7}, \frac{0.97}{u_8}\right)$ |
| 2014 | 31.30 | -1.80 | $A^{2014} = \left(\frac{0.79}{u_1}, \frac{0.90}{u_2}, \frac{0.98}{u_3}, \frac{1.00}{u_4}, \frac{0.93}{u_5}, \frac{0.82}{u_6}, \frac{0.70}{u_7}, \frac{0.58}{u_8}\right)$ |
| 2015 | 33.10 | 1.80 | $A^{2015} = \left(\frac{0.58}{u_1}, \frac{0.70}{u_2}, \frac{0.83}{u_3}, \frac{0.94}{u_4}, \frac{1.00}{u_5}, \frac{0.98}{u_6}, \frac{0.90}{u_7}, \frac{0.78}{u_8}\right)$ |

Bước 2. Chia tập nền U ra làm 8 đoạn có độ dài bằng nhau:

$$u_1 = [-8.7; -6.5], u_2 = [-6.5; -4.3], u_3 = [-4.3; -2.1], u_4 = [-2.1; 0.0]$$

$$u_5 = [0.0; 2.2], u_6 = [2.2; 4.4], u_7 = [4.4; 6.6], u_8 = [6.6; 8.8]$$

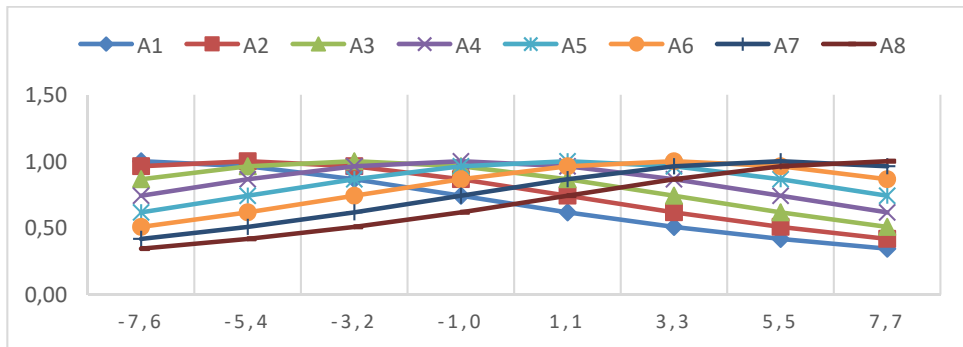
Tìm điểm giữa của các đoạn trên

$$u_m^1 = -7.6, u_m^2 = -5.4, u_m^3 = -3.2, u_m^4 = -1.0, u_m^5 = 1.1, u_m^6 = 3.3, u_m^7 = 5.5, u_m^8 = 7.7$$

Bước 3. Xác định các biến mờ:

$$\begin{aligned}
 A_2 &= \left\{ \left(\frac{0.96}{u_1} \right), \left(\frac{1.00}{u_2} \right), \left(\frac{0.96}{u_3} \right), \left(\frac{0.87}{u_4} \right), \left(\frac{0.74}{u_5} \right), \left(\frac{0.62}{u_6} \right), \left(\frac{0.51}{u_7} \right), \left(\frac{0.42}{u_8} \right) \right\} \\
 A_3 &= \left\{ \left(\frac{0.87}{u_1} \right), \left(\frac{0.96}{u_2} \right), \left(\frac{1.00}{u_3} \right), \left(\frac{0.96}{u_4} \right), \left(\frac{0.87}{u_5} \right), \left(\frac{0.74}{u_6} \right), \left(\frac{0.62}{u_7} \right), \left(\frac{0.51}{u_8} \right) \right\} \\
 A_4 &= \left\{ \left(\frac{0.74}{u_1} \right), \left(\frac{0.87}{u_2} \right), \left(\frac{0.96}{u_3} \right), \left(\frac{1.00}{u_4} \right), \left(\frac{0.96}{u_5} \right), \left(\frac{0.87}{u_6} \right), \left(\frac{0.74}{u_7} \right), \left(\frac{0.61}{u_8} \right) \right\} \\
 A_5 &= \left\{ \left(\frac{0.62}{u_1} \right), \left(\frac{0.74}{u_2} \right), \left(\frac{0.87}{u_3} \right), \left(\frac{0.96}{u_4} \right), \left(\frac{1.00}{u_5} \right), \left(\frac{0.96}{u_6} \right), \left(\frac{0.87}{u_7} \right), \left(\frac{0.74}{u_8} \right) \right\} \\
 A_6 &= \left\{ \left(\frac{0.51}{u_1} \right), \left(\frac{0.62}{u_2} \right), \left(\frac{0.74}{u_3} \right), \left(\frac{0.87}{u_4} \right), \left(\frac{0.96}{u_5} \right), \left(\frac{1.00}{u_6} \right), \left(\frac{0.96}{u_7} \right), \left(\frac{0.87}{u_8} \right) \right\} \\
 A_7 &= \left\{ \left(\frac{0.42}{u_1} \right), \left(\frac{0.51}{u_2} \right), \left(\frac{0.62}{u_3} \right), \left(\frac{0.74}{u_4} \right), \left(\frac{0.87}{u_5} \right), \left(\frac{0.96}{u_6} \right), \left(\frac{1.00}{u_7} \right), \left(\frac{0.96}{u_8} \right) \right\} \\
 A_8 &= \left\{ \left(\frac{0.34}{u_1} \right), \left(\frac{0.42}{u_2} \right), \left(\frac{0.51}{u_3} \right), \left(\frac{0.62}{u_4} \right), \left(\frac{0.74}{u_5} \right), \left(\frac{0.87}{u_6} \right), \left(\frac{0.96}{u_7} \right), \left(\frac{1.00}{u_8} \right) \right\}
 \end{aligned}$$

Bước 4. Mờ hóa các biến đổi được tính toán tại bước đầu tiên.



Hình 2: Hình vẽ minh họa các cấp độ mờ hóa

Bước 5. Chọn $w=7$, thiết lập ma trận mờ cho đỉnh mận 2008 như sau:

$$O^7 = \begin{bmatrix} \text{Biến đỉnh mận mờ năm 2001} A^{2001} \\ \text{Biến đỉnh mận mờ năm 2002} A^{2002} \\ \text{Biến đỉnh mận mờ năm 2003} A^{2003} \\ \text{Biến đỉnh mận mờ năm 2004} A^{2004} \\ \text{Biến đỉnh mận mờ năm 2005} A^{2005} \\ \text{Biến đỉnh mận mờ năm 2006} A^{2006} \end{bmatrix}$$

$$O^7 = \begin{bmatrix} 0.69 & 0.82 & 0.93 & 0.99 & 0.99 & 0.91 & 0.79 & 0.66 \\ 0.44 & 0.53 & 0.65 & 0.77 & 0.89 & 0.98 & 1.00 & 0.94 \\ 0.64 & 0.77 & 0.89 & 0.98 & 1.00 & 0.95 & 0.84 & 0.72 \\ 0.73 & 0.85 & 0.95 & 1.00 & 0.97 & 0.88 & 0.76 & 0.63 \\ 0.58 & 0.70 & 0.83 & 0.94 & 1.00 & 0.98 & 0.90 & 0.78 \\ 0.93 & 0.99 & 0.99 & 0.91 & 0.79 & 0.67 & 0.55 & 0.45 \end{bmatrix}$$

$$K = [0.61 \quad 0.73 \quad 0.86 \quad 0.96 \quad 1.00 \quad 0.97 \quad 0.87 \quad 0.75] A^{2007}$$

$$R = \begin{bmatrix} 0.61 & 0.73 & 0.86 & 0.96 & 0.99 & 0.91 & 0.79 & 0.66 \\ 0.44 & 0.53 & 0.65 & 0.77 & 0.89 & 0.97 & 0.87 & 0.75 \\ 0.61 & 0.73 & 0.86 & 0.96 & 1.00 & 0.95 & 0.84 & 0.72 \\ 0.61 & 0.73 & 0.86 & 0.96 & 0.97 & 0.88 & 0.76 & 0.63 \\ 0.58 & 0.70 & 0.83 & 0.94 & 1.00 & 0.97 & 0.87 & 0.75 \\ 0.61 & 0.73 & 0.86 & 0.91 & 0.79 & 0.67 & 0.55 & 0.45 \end{bmatrix}$$

$$F(2008) = [0.61 \quad 0.73 \quad 0.86 \quad 0.96 \quad 1.00 \quad 0.97 \quad 0.87 \quad 0.75]$$

Bước 6. Từ kết quả của bước 5, ta ước tính đỉnh mặn cho năm 2007.

$$V(2008) = \frac{0.61(-7.6) + 0.73(-5.4) + \dots + 0.87.5.5 + 0.75.7.7}{0.61 + 0.73 + \dots + 0.87 + 0.75} = 0.39$$

$$Y(2007) = N(2007) + V(2008) = 32.9 + 0.39 = 33.29.$$

Vậy đỉnh mặn dự báo cho năm 2007 là 33.29(%), các năm tiếp theo dự báo tương tự ta được kết quả sau:

Bảng 5: Kết quả nội suy đỉnh mặn CM giai đoạn 2008- 2015

| Năm | Thực tế | | Dự báo | | Sai số (%) | Sai số trung bình (%) |
|------|----------|----------|----------|----------|------------|-----------------------|
| | Đỉnh mặn | Biến đổi | Đỉnh mặn | Biến đổi | | |
| 2007 | 32.90 | 1.30 | 33.29 | 1.69 | 1.2 | |
| 2008 | 31.50 | -1.40 | 31.16 | -2.13 | 1.1 | |
| 2009 | 28.30 | -3.20 | 27.51 | -3.65 | 2.8 | |
| 2010 | 37.10 | 8.80 | 38.72 | 11.22 | 4.4 | |
| 2011 | 28.40 | -8.70 | 26.68 | -12.04 | 6.0 | 2.8 |
| 2012 | 27.30 | -1.10 | 27.04 | 0.36 | 1.0 | |
| 2013 | 33.10 | 5.80 | 34.49 | 7.45 | 4.2 | |
| 2014 | 31.30 | -1.80 | 30.86 | -3.63 | 1.4 | |
| 2015 | 33.10 | 1.80 | 33.60 | 2.75 | 1.5 | |

Bảng 6: Kết quả dự báo đỉnh mặn tại trạm CM giai đoạn 2016- 2020

| Năm | Dự báo | Biến đổi | Mờ hóa các biến đổi | | | | | | | |
|------|--------|----------|---|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| | | | $\left(\frac{0.65}{u_1}\right)$ | $\left(\frac{0.78}{u_2}\right)$ | $\left(\frac{0.90}{u_3}\right)$ | $\left(\frac{0.98}{u_4}\right)$ | $\left(\frac{1.00}{u_5}\right)$ | $\left(\frac{0.94}{u_6}\right)$ | $\left(\frac{0.83}{u_7}\right)$ | $\left(\frac{0.70}{u_8}\right)$ |
| 2016 | 33.60 | 0.50 | $A^{2016} = \left(\frac{0.65}{u_1}\right), \left(\frac{0.78}{u_2}\right), \left(\frac{0.90}{u_3}\right), \left(\frac{0.98}{u_4}\right), \left(\frac{1.00}{u_5}\right), \left(\frac{0.94}{u_6}\right), \left(\frac{0.83}{u_7}\right), \left(\frac{0.70}{u_8}\right)$ | | | | | | | |
| 2017 | 33.78 | 0.17 | $A^{2017} = \left(\frac{0.67}{u_1}\right), \left(\frac{0.80}{u_2}\right), \left(\frac{0.91}{u_3}\right), \left(\frac{0.99}{u_4}\right), \left(\frac{0.99}{u_5}\right), \left(\frac{0.93}{u_6}\right), \left(\frac{0.81}{u_7}\right), \left(\frac{0.69}{u_8}\right)$ | | | | | | | |
| 2018 | 33.86 | 0.08 | $A^{2018} = \left(\frac{0.68}{u_1}\right), \left(\frac{0.80}{u_2}\right), \left(\frac{0.92}{u_3}\right), \left(\frac{0.99}{u_4}\right), \left(\frac{0.99}{u_5}\right), \left(\frac{0.92}{u_6}\right), \left(\frac{0.81}{u_7}\right), \left(\frac{0.68}{u_8}\right)$ | | | | | | | |
| 2019 | 33.92 | 0.06 | $A^{2019} = \left(\frac{0.68}{u_1}\right), \left(\frac{0.80}{u_2}\right), \left(\frac{0.92}{u_3}\right), \left(\frac{0.99}{u_4}\right), \left(\frac{0.99}{u_5}\right), \left(\frac{0.92}{u_6}\right), \left(\frac{0.81}{u_7}\right), \left(\frac{0.68}{u_8}\right)$ | | | | | | | |
| 2020 | 33.97 | 0.05 | $A^{2020} = \left(\frac{0.68}{u_1}\right), \left(\frac{0.80}{u_2}\right), \left(\frac{0.92}{u_3}\right), \left(\frac{0.99}{u_4}\right), \left(\frac{0.99}{u_5}\right), \left(\frac{0.92}{u_6}\right), \left(\frac{0.81}{u_7}\right), \left(\frac{0.68}{u_8}\right)$ | | | | | | | |

b. Mô hình từ dữ liệu mờ hóa

Mờ hóa từ dữ liệu gốc theo mô hình của Chen, Singh, Heuristic và Chen-Hsu ta có bảng tổng hợp Bảng 7.

Vì số liệu có được từ mô hình Chen-Hsu cho ta ME nhỏ nhất, nên chúng ta sử dụng phương pháp mờ hóa này cho dự báo.

Bảng 7: Số liệu mờ hóa đỉnh mặn tại CM theo các phương pháp

| Năm | Thực tế | Chen | Singh | Heuristic | Chen-Hsu |
|------|-----------|-------|-------|-----------|-------------|
| 2000 | 29.6 | | | | |
| 2001 | 29.4 | 32.20 | | 29.40 | 29.40 |
| 2002 | 34.4 | 32.20 | | 32.20 | 34.65 |
| 2003 | 35.1 | 34.30 | 35.27 | 35.00 | 35.00 |
| 2004 | 34.3 | 34.30 | 33.53 | 34.30 | 33.25 |
| 2005 | 36.1 | 33.60 | 36.40 | 36.40 | 36.40 |
| 2006 | 31.6 | 30.10 | 32.20 | 30.10 | 32.20 |
| 2007 | 32.9 | 31.50 | 32.18 | 32.20 | 32.20 |
| 2008 | 31.5 | 31.50 | 31.05 | 31.50 | 30.80 |
| 2009 | 28.3 | 30.80 | 28.00 | 28.00 | 28.00 |
| 2010 | 37.1 | 32.67 | 36.40 | 32.67 | 36.40 |
| 2011 | 28.4 | 30.10 | 28.00 | 30.10 | 28.00 |
| 2012 | 27.3 | 32.67 | 28.37 | 28.00 | 28.00 |
| 2013 | 33.1 | 32.67 | 33.60 | 32.67 | 33.60 |
| 2014 | 31.3 | 33.60 | 30.78 | 30.80 | 30.80 |
| 2015 | 33.1 | 30.80 | 33.45 | 33.60 | 33.60 |
| | ME | 2.02 | 0.53 | 0.89 | 0.49 |

c. Chọn mô hình và dự báo

Sử dụng số liệu gốc, mô hình tối ưu khi mờ hóa, áp dụng phương pháp Box-Jenkin, ta có mô hình tối ưu cho các trường hợp sau:

Bảng 8: Các mô hình tối ưu trong dự báo đỉnh mặn tại CM

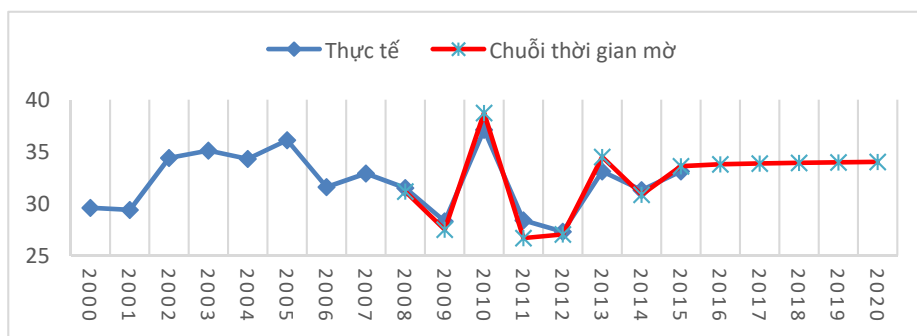
| Dữ liệu | Mô hình tối ưu | ME | AIC |
|---------|-----------------------|-------------|--------------|
| Gốc | ARIMA(1,1,1) | 2.28 | 186.89 |
| Mờ hóa | ARIMA-Chen-Hsu(1,1,1) | 2.56 | 202.75 |
| Gốc | Chuỗi thời gian mờ | 0.88 | 20.00 |

Bảng này cho thấy mô hình chuỗi thời gian mờ

và chỉ số ME và AIC nhỏ hơn các mô hình khác, do đó mô hình này được đánh giá phù hợp nhất. Sử dụng mô hình này, đỉnh mặn tại CM được dự báo đến năm 2020 như sau:

Bảng 9: Kết quả dự báo đỉnh mặn tại CM giai đoạn 2016 – 2020

| Năm | Kết quả dự báo |
|------|----------------|
| 2016 | 33.78 |
| 2017 | 33.86 |
| 2018 | 33.92 |
| 2019 | 33.97 |
| 2020 | 34.02 |



Hình 3: Đồ thị dự báo đỉnh mặn tại CM giai đoạn 2000 – 2020

4.5 Dự báo đỉnh mặn tại trạm Gành Hào

Thực hiện tương tự như mục 4.4, cho số liệu đỉnh mặn tại trạm Gành Hào, ta có bảng tổng kết sau:

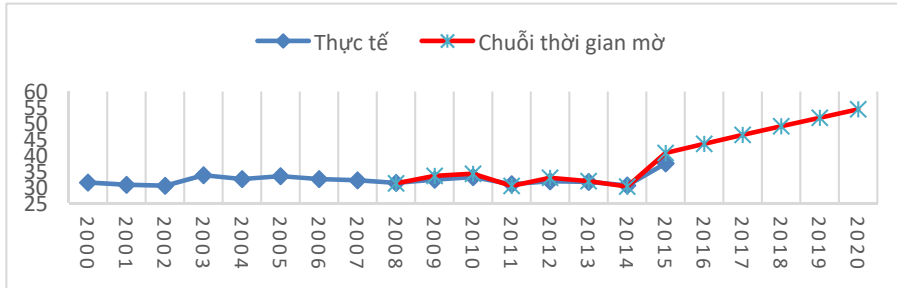
Bảng 10: Các mô hình tối ưu trong dự báo đỉnh mặn tại GH

| Dữ liệu | Mô hình tối ưu | ME | AIC |
|---------|-----------------------|-------------|--------------|
| Gốc | ARIMA(1,1,1) | 1.25 | 68.51 |
| Mờ hóa | ARIMA-Chen-Hsu(1,1,1) | 1.08 | 57.17 |
| Gốc | Chuỗi thời gian mờ | 0.98 | 13.20 |

Bảng 10 cho thấy, mô hình chuỗi thời gian mờ cho đỉnh mặn tại GH có ME và AIC nhỏ hơn hai trường hợp còn lại. Do đó, mô hình này phù hợp nhất để dự báo. Với mô hình đã chọn, ta dự báo đỉnh mặn tại trạm GH đến năm 2020 với kết quả tổng hợp sau:

Bảng 11: Kết quả dự báo đỉnh mặn tại GH giai đoạn 2016 – 2020

| Năm | Kết quả dự báo |
|------|----------------|
| 2016 | 43.65 |
| 2017 | 46.41 |
| 2018 | 49.13 |
| 2019 | 51.80 |
| 2020 | 54.44 |



Hình 4: Đồ thị kết quả dự báo đỉnh mặn tại GH giai đoạn 2000 – 2020.

4.6 Dự báo đỉnh mặn tại trạm Ông Đốc

Cũng thực hiện tương tự như mục 4.4, ta có các kết quả sau:

Bảng 12: Các mô hình tối ưu trong dự báo đỉnh mặn tại OD

| Dữ liệu | Mô hình tối ưu | ME | AIC |
|---------|-----------------------|-------------|--------------|
| Gốc | ARIMA(1,1,1) | 2.05 | 134.59 |
| Mờ hóa | ARIMA-Chen-Hsu(1,1,1) | 2.20 | 142.89 |
| Gốc | Chuỗi thời gian mờ | 0.60 | 17.12 |

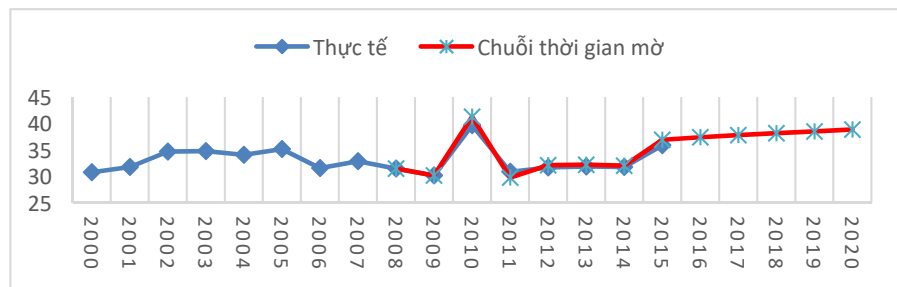
Bảng này một lần nữa cho thấy, các tham số đánh giá mô hình (ME và AIC) của chuỗi thời gian mờ nhỏ hơn rất nhiều so với hai trường hợp còn lại.

Như vậy, mô hình chuỗi thời gian mờ phù hợp hơn trong dự báo.

Sử dụng mô hình chuỗi thời gian mờ, ta dự báo đỉnh mặn OD đến năm 2020 như sau:

Bảng 13: Kết quả dự báo đỉnh mặn tại OD giai đoạn 2016 - 2020

| Năm | Kết quả dự báo |
|------|----------------|
| 2016 | 37.44 |
| 2017 | 37.82 |
| 2018 | 38.17 |
| 2019 | 38.52 |
| 2020 | 38.87 |



Hình 5: Đồ thị cho dự báo đỉnh mặn tại OD giai đoạn 2000 – 2020

4.7 Nhận xét

Từ các phân tích ta rút ra được một số nhận xét sau:

Giá trị AIC, ME theo mô hình chuỗi thời gian mờ nhỏ hơn nhiều so với các mô hình tối ưu đã thiết lập khi sử dụng dữ liệu gốc và dữ liệu mờ hóa cho cả 3 nơi OD, CM và GH. Như vậy, cả 3 nơi sử dụng mô hình này phù hợp nhất.

So sánh dữ liệu thực tế và nội suy từ mô hình, cùng với các tham số đánh giá cho thấy, mô hình chuỗi thời gian mờ phù hợp để dự báo.

Kết quả dự báo cho thấy, đỉnh mặn có khuynh hướng tăng với các năm tiếp theo, trong đó tại GH mức độ tăng nhiều nhất. Với đặc điểm dự báo khuynh hướng có độ tin cậy khá tốt trong ngắn hạn của mô hình thì đây là kết quả mà chúng ta cần quan tâm.

5 KẾT LUẬN

Bài báo đã đề xuất mô hình chuỗi thời gian mờ, trong dự báo đỉnh mặn tại 3 trạm đo chính của tỉnh Cà Mau: Gành Hào, Cà Mau và Sông Đốc. Các tiêu chuẩn thống kê cho thấy mô hình đề nghị có ưu điểm hơn các mô hình chuỗi thời gian không mờ xây dựng từ số liệu gốc và số liệu mờ hóa. Các hình vẽ cho thấy, dữ liệu nội suy từ mô hình này khá sát với dữ liệu thực tế của quá khứ cho thấy tính tương đối hợp lý của mô hình xây dựng. Sử dụng mô hình đề xuất, chúng tôi đã dự báo đỉnh mặn tại 3 trạm đến năm 2020. Các kết quả dự báo cho thấy đỉnh mặn có khuynh hướng tăng trong những năm tiếp theo. Do đặc điểm của dữ liệu, mô hình dự báo xây dựng chưa thực hiện việc kiểm tra, tuy nhiên với những kết quả đã xây dựng mô hình ở trên, chúng tôi nghĩ rằng kết quả dự báo là một thông tin đáng quan tâm. Mô hình chuỗi thời gian mờ với ưu điểm nổi bật dựa vào sự liên kết xác suất của dữ liệu, không đòi hỏi nhiều dữ liệu quá khứ, có thể áp dụng trong dự báo cho nhiều vấn đề thực tế khác.

TÀI LIỆU THAM KHẢO

- Abbasov, A.M., 2002. Fuzzy relational model for knowledge processing and decision making. *Advances in Mathematics*. 1: 1991–223.
- Abbasov, A.M. and Mamedova, M.H., 2003. Application of fuzzy time series to population forecasting, *Proceedings of 8th Symposium on Information Technology in Urban and Spatial Planning*, Vienna University of Technology, February 25-March 1, 545–552.
- Bozdogan, H., 2000. Akaike's information criterion and recent developments in information complexity. *Journal of mathematical psychology*. 44: 62–91.
- Chen, S.M., 1996. Forecasting enrollments based on fuzzy time series. *Fuzzy Sets and Systems*. 81: 311–319.
- Chen, S.M. and Hsu, C.C., 2004. A New method to forecast enrollments using fuzzy time series. *International Journal of Applied Science and Engineering*, 12: 234–244.
- Huang, H., 2001. Huarng models of fuzzy time series for forecasting. *Fuzzy Sets and Systems*. 123: 369–386.
- Nguyễn Thanh Sơn, 2003. *Tính toán thủy văn*. Nhà xuất bản Đại học Quốc gia Hà Nội. Hà Nội. 187 trang.
- Phan Văn Tân, 2005. *Các phương pháp thống kê trong khí hậu*. Nhà xuất bản Đại học Quốc gia Hà Nội. Hà Nội. 162 trang.
- Singh, S.R., 2008. A computational method of forecasting based on fuzzy time series. *Mathematics and Computers in Simulation*. 79: 539–554.
- Singh, S.R., 2009. A computational method of forecasting based on high-order fuzzy time series. *Expert Systems with Applications*. 36: 10551–10559.
- Song, B. and Chisom, B.S., 1993. Forecasting enrollments with fuzzy time series (Part I), *Fuzzy Sets and Systems*. 54: 1–9.
- Song, S. and Chisom, B.S., 1994. Forecasting enrollments with fuzzy time series (Part II), *Fuzzy Sets and Systems*. 62: 1–8.
- Tô Văn Trường, 2005. *Phương pháp và công nghệ dự báo lũ Đồng bằng sông Cửu Long*. Phân viện khảo sát qui hoạch thủy lợi. 112 trang.
- Võ Văn Tài, 2015. Phân tích thống kê đỉnh lũ tại trạm đo Châu Đốc trên sông Hậu. *Tạp chí Khoa học Trường Đại học Cần Thơ*. 34: 33 – 44.
- Zadeh, L. A., 1965. Fuzzy sets. *Information and Control*. 8(3): 338–353.