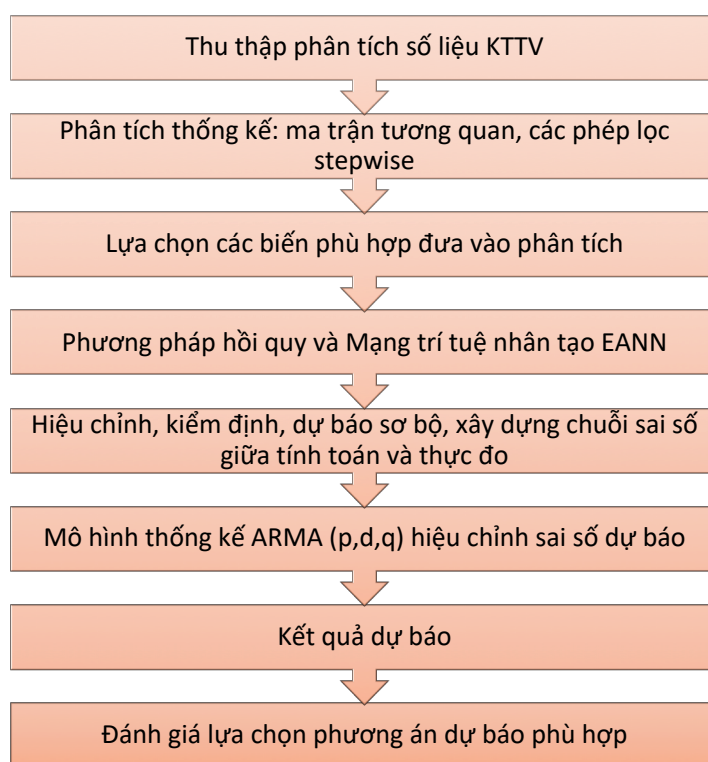


## CHƯƠNG 5. NGHIÊN CỨU XÂY DỰNG MÔ HÌNH DỰ BÁO DÒNG CHẢY HẠN DÀI TRÊN SÔNG MÊ CÔNG

### 5.1. Lựa chọn phương pháp dự báo dòng chảy hạn dài trên Sông Mê Công

Từ điều kiện số liệu đã thu thập được có thể thấy số liệu đo đặc khí tượng – thủy văn khu vực thượng lưu lưu vực sông Mê Công (trên lãnh thổ Trung Quốc) được thu thập và cung cấp rất hạn chế nên trạm đo dòng chảy chính trên sông Mê Công xa nhất về thượng nguồn có thể lấy là trạm Chiang Saen thuộc Lào. Chính vì vậy trước mắt (trong ngắn hạn) có thể sử dụng phương pháp thống kê và mạng trí tuệ nhân tạo (AI) để dự báo dòng chảy hạn dài (1 tháng, 3 tháng) cho một số vị trí trên dòng chính của sông Mê Công. Còn trong dài hạn cần xây dựng một mô hình thủy văn lưu vực cho phép tính toán dòng chảy từ mưa đến các lưu vực hồ chứa, kết hợp với việc giám sát và dự báo nguồn nước từ các hồ chứa trên lưu vực để dự báo dòng chảy hạn dài (1 tháng) về ĐBSCL khi đó không cần phải xem xét tác động của các hồ thủy điện điều tiết ngày đêm trên lưu vực.

Sơ đồ dưới đây minh họa quy trình xây dựng các mô hình, phương án dự báo dòng chảy hạn dài (tháng và mùa) sử dụng kết hợp phương pháp phân tích thống kê và mạng trí tuệ nhân tạo:



Hình 5.1: Quy trình sử dụng phương pháp phân tích thống kê và mạng trí tuệ nhân tạo AI xây dựng các phương án dự báo dòng chảy hạn dài trên lưu vực Sông Mê Công

Trong sơ đồ nghiên cứu trên, việc lựa chọn số liệu để phân tích là rất quan trọng, đặc biệt với phương pháp mạng trí tuệ nhân tạo (AI) khi có quá nhiều biến đầu vào mà không biết những số liệu nào có ảnh hưởng lớn đến biến đầu ra (biến cần dự báo). Đề tài sẽ sử

dụng thuật toán “Stepwise” để thực hiện việc này. Thuật toán này cho phép lần lượt đưa vào và đưa ra từng biến để đánh giá mức độ ảnh hưởng/đóng góp của từng biến với biến cần dự báo. Khi mỗi biến đưa vào, mô hình sẽ phân tích và đánh giá các chỉ tiêu thống kê (ví dụ như Fisher (F), Student (T),...hệ số tương quan bội ( $R^2$ ) nếu chỉ tiêu đạt thì giữ lại, không đạt thì loại ra.

Bên cạnh việc áp dụng thuật toán “Stepwise” để lựa chọn số liệu, đề tài sẽ sử dụng kết hợp mô hình mạng trí tuệ nhân tạo AI (EANN) + mô hình ngẫu nhiên ARIMA (p,d,q). Trong đó mô hình EANN là mô hình cho ra giá trị dự báo sơ bộ, còn mô hình ARIMA(p,d,q) cho phép hiệu chỉnh giá trị dự báo. Đề xuất này xuất phát từ bản chất các yếu tố KTTV vừa mang tính ngẫu nhiên vừa mang tính tất định.

Mô hình mạng trí tuệ nhân tạo AI được áp dụng trong dự báo dòng chảy hạn dài là mô hình EANN cũng là một trong những dạng lai ghép giữa mạng Nơ ron thần kinh với thuật toán quét ngược (BPNN – Back Propagation Neural Network) được áp dụng nhiều trong thủy văn với thuật toán giải đoán gen (GA - Genetic Algorithms). Nhiều nhà nghiên cứu đã khẳng định mô hình AI luôn cho kết quả tốt hơn so với mô hình hồi quy nhiều biến nhờ khả năng giải quyết bài toán có tính phi tuyến.

Mạng trí tuệ nhân tạo là một hệ thống xử lý thông tin được xây dựng trên cơ sở tổng quát hoá mô hình toán học của nơ-ron sinh học và phỏng theo cơ chế làm việc của bộ não con người. Mạng nơ-ron nhân tạo được thể hiện thông qua ba thành phần cơ bản: mô hình của nơ-ron, cấu trúc và sự liên kết giữa các nơ-ron, phương pháp học được áp dụng cho mạng nơ-ron.

Hiện nay trên thế giới mô hình ANN được phát triển theo các hướng căn cứ vào 2 thành phần nêu trên của mạng nơ ron đó là: i) cấu trúc của mạng nơ ron; và ii) phương pháp học được áp dụng cho mạng nơ ron (thuật toán tối ưu để xác định các trọng số liên kết, và cấu trúc mạng).

Nếu căn cứ vào “cấu trúc của mạng ANN” thì ta thấy có các loại mô hình ANN sau:

- Mạng nơ-ron truyền thẳng một lớp (Single-layer feedforward network): loại mạng này còn được gọi là mạng perceptron một lớp
- Mạng nơ-ron truyền thẳng nhiều lớp (Multil-layer feedforward network).
- Mạng phản hồi (Feedback network)
- Mạng nơ-ron hồi quy (Recurrent network)

Nếu căn cứ vào “cách học của mạng ANN” thì ta thấy có những cách học: học có giám sát (Supervised learning), học không có giám sát (Unsupervised learning), học tăng cường (reinforced learning). Từ đó có các loại mô hình ANN tương ứng với các thuật toán sau:

- Thuật toán quét ngược (Back Propagation Neural Network)
- Thuật toán giải đoán gen (Genetic Algorithm)

- Thuật toán tập mờ (Adaptive neuron – fuzzy inference system – ANFIS)

Đề tài lựa chọn sử dụng mô hình lai ghép giữa mạng Nơ ron thần kinh lan truyền ngược (BPNN) với thuật toán giải đoán gen GA (gọi tắt là EANN – Mạng Nơ ron tiến hóa) để dự báo dòng chảy hạn dài cho một số vị trí trên dòng chính sông Mê Công; Trong đó thuật giải đoán gen GA sẽ hỗ trợ việc tối ưu hóa sơ đồ mạng Nơ ron (số nút đầu vào, số lớp ẩn và số nút trong từng lớp ẩn). Trong quá trình dự báo tác nghiệp, khi số liệu được cập nhật thì mạng cũng được cập nhật theo nhờ thuật toán GA.

Ngoài ra để hướng tới mục tiêu lâu dài trong dự báo và cảnh báo dòng chảy hạn dài về ĐBSCL phục vụ phát triển kinh tế xã hội, đề tài đã bước đầu xây dựng mô hình thủy văn lưu vực trên toàn lưu vực sông Mê Công để kết nối giữa các sản phẩm dự báo mưa tháng, và kết quả giám sát nguồn nước các hồ chứa trên lưu vực sông từ ảnh vệ tinh được trình bày trong chương 2 và 3 của báo cáo.

## 5.2. Phương pháp dự báo dòng chảy hạn dài trên Sông Mê Công

### 5.2.1. Phương pháp thống kê – ngẫu nhiên

Do số liệu đo đạc khí tượng – thủy văn trên lãnh thổ Trung Quốc được cung cấp rất hạn chế nên trạm đo dòng chảy chính trên sông Mê Công xa nhất về thượng nguồn có thể lấy tại trạm Chiang Saen thuộc Lào.

Để dự báo dòng chảy hạn dài (1 tháng, 3 tháng, 6 tháng) thời đoạn tháng, các phương pháp thống kê – ngẫu nhiên có thể được sử dụng để dự báo cho các vị trí trên dòng chính.

- *Phương pháp mô hình tự hồi quy AR(p)*

Mô hình tự hồi quy AR(p) có dạng:

$$Z_t = \sum_{i=1}^p a_i Z_{t-i} + \varepsilon_t$$

Trong đó:  $Z_t$ : Giá trị tính toán của chuỗi Z tại thời điểm t; p: Nguyên dương và là bậc của mô hình;  $a_1, a_2, \dots, a_p$ : các hệ số hằng số;  $Z_{t-1}, Z_{t-2}, \dots$ : là các giá trị Z thực đo tại t-1, t-2, ... ;  $\varepsilon_t$ : Thành phần ngẫu nhiên

- *Phương pháp hồi quy tuyến tính đa biến*

Mô hình hồi quy tuyến tính đa biến có dạng:

$$H_{t+dt} = \sum_{i=1}^k b_i X_i + b_0$$

Trong đó:  $H_t$ : Giá trị tính toán (biến phụ thuộc) của chuỗi H tại thời điểm t; k: Nguyên dương và là số biến độc lập của mô hình;  $b_1, b_2, \dots, b_k$ : các hệ số hằng số của phương trình;  $X_i$ : là các giá trị biến độc lập thứ i.

### 5.2.2. Phương pháp mạng trí tuệ nhân tạo EANN

### 5.2.2.1. Lựa chọn số liệu đầu vào dự báo dòng chảy trung bình tháng các trạm Tân Châu, Châu Đốc và Chiang Sean (Lào)

Với số liệu mực nước ngày các trạm trên dòng chính Sông Mê Công đã thu thập được từ năm 1980 đến năm 2020, cùng với việc phân tích dòng chảy tháng trên dòng chính Sông Mê Công, ta thấy để dự báo mực nước trung bình tháng tại trạm Tân Châu với thời gian dự kiến là 1 tháng ta có thể lựa chọn sơ bộ các nhóm biến đầu vào gồm:

1	Hchiengsean_t-1	mực nước trung bình trạm Chiang Sean i tháng trước tháng cần dự báo ( $i=0, 1, 2$ )
2	HTanchau_cuoitthang_i	mực nước trung bình trạm Tân Châu i ngày cuối tháng trước tháng cần dự báo ( $i = 1, 2, 3$ ngày)
3	HTanchau_t-i	mực nước trung bình tháng trạm Tân Châu i tháng trước tháng cần dự báo (với $i = 1, 2 \dots 11$ )

Áp dụng thuật toán “**Stepwise**” ta lựa chọn được 8 biến tốt nhất đưa vào xây dựng mạng EANN cho trạm Tân Châu

Tương tự như vậy với ta cũng lựa chọn được 8 biến ban đầu cho trạm Châu Đốc và 6 biến ban đầu cho trạm Chiang Sean (Lào – dự báo thử nghiệm thêm cho 1 trạm ngoài lãnh thổ Việt Nam)

### 5.2.2.2. Xây dựng mô hình EANN cho trạm cần dự báo

Sau khi lọc và lựa chọn được những biến thích hợp đưa vào mạng, thì việc xây dựng mạng được tiến hành theo các bước sau:

#### Bước 1 - Thiết lập mạng:

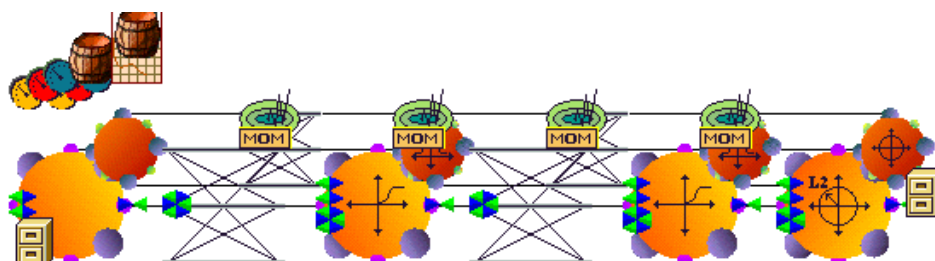
- Thiết đặt trước các thể hệ di truyền và các giá trị cho hàm mục tiêu để đặt ra tiêu chí cho việc dùng chạy mô hình.
- Thiết đặt số lượng ghen phân tích trong mạng.
- Lựa chọn dữ liệu cho đào tạo (learning): chuỗi số liệu từ tháng 1/1980 đến tháng 12/2010, kiểm tra chéo (crossover) : chuỗi số liệu từ tháng 1/2011 đến tháng 12/2014. Và để dự báo thử nghiệm: chuỗi dữ liệu từ 1/2015 đến 1/2020.

**Bước 2 - Khởi tạo mạng:** Bao gồm việc ngẫu nhiên hóa số lượng các nút ẩn trong các lớp ẩn khác nhau trong đó có việc lựa chọn hàm chuyển đổi (trong nguyên cứu này lựa chọn hàm Sigmoid), gán ngẫu nhiên các giá trị trọng số kết nối

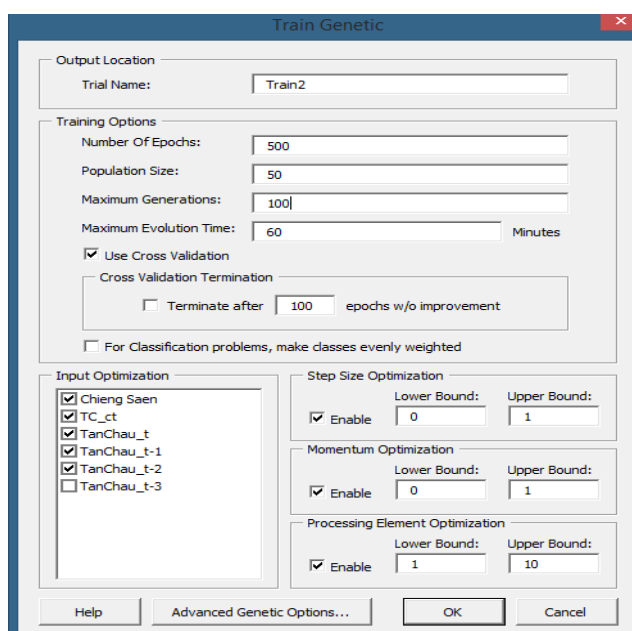
**Bước 3 - Mã hóa:** Bao gồm việc mã hóa các nút trong cấu trúc mạng như “1” có nghĩa là đầu vào được lựa chọn “0” là không được chọn, sau đó chuyển đổi các mã nhị phân này vào từng lớp ẩn và xây dựng mạng.

**Bước 4 - Xây dựng các mạng ANNs:** bao gồm việc đưa dữ liệu đào tạo và kiểm tra vào từng cấu trúc mạng đã xây dựng, tính toán các hàm mục tiêu đã đặt ra.

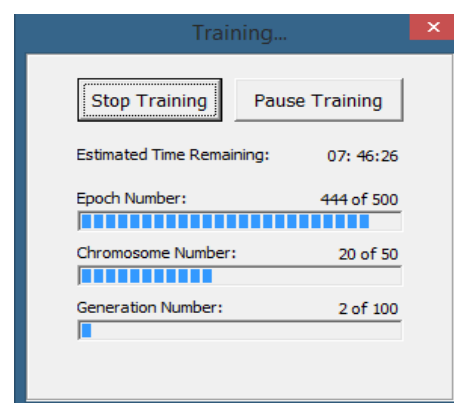
**Bước 5 - Áp dụng thuật giải đoán ghen GA vào tối ưu hóa mạng:** bao gồm việc kiểm tra các chỉ tiêu cho từng mạng, thay đổi các cấu trúc đã định sẵn để tìm ra mạng phù hợp nhất, tiếp đó thực hiện những thay đổi nhỏ, thực hiện việc kiểm tra chéo giữa các đoạn số liệu khác nhau xem có đạt không, nếu đạt chuyển sang các thể hệ tiếp theo. Chương trình sẽ dừng tính và cho mạng phù hợp nhất khi thực hiện hết các thể hệ đã định trước ở bước 1.



a) Cấu trúc mạng EANN



b) Các thông số chạy EANN



c) Quá trình luyện mạng EANN

*Hình 5.2: Minh họa quá trình xây dựng mô hình EANN dự báo mực nước trung bình tháng trạm Tân Châu*

Các thông số mạng thiết lập cho trạm Tân Châu bao gồm:

- Số biến đầu vào là 8 (i);
- Số lớp ẩn tối đa là 2, mỗi lớp có không quá 15 nút (thường chọn từ  $i \rightarrow 2i-1$ ) như vậy số ghen trong 2 lớp ẩn sẽ không quá 30;
- Mạng ban đầu được chọn để chạy có khoảng 30 nút (số nút sẽ dần dần được thay đổi để tìm mạng thích hợp);
- Lựa chọn 100 thể hệ;
- Số lần tính tối đa cho mỗi cấu trúc mạng là 500;

- Lựa chọn chỉ tiêu đánh giá và dùng tính toán là sự thay đổi của sai số quân phương sau mỗi lần tính (cũng có thể lựa chọn thời gian tối đa luyện mạng trong trường hợp không muốn tốn quá nhiều thời gian).

Kết quả chạy mô hình EANN dự báo mực nước trung bình tháng cho trạm Tân Châu, Châu Đốc và Chiang Sean được trình bày chi tiết trong Báo cáo sản phẩm 4 của đề tài và được tóm tắt trong mục 5.3.5 của báo cáo.

### 5.2.3. Phương pháp mô hình mưa – dòng chảy

Phương pháp mô hình mưa – dòng chảy có ưu điểm là có thể mô tả đúng quá trình vật lý từ mưa sinh ra dòng chảy. Nếu kết hợp với các dữ liệu dự báo mưa hạn dài và quá trình tích, xả của các hồ chứa trên lưu vực thì phương pháp này sẽ giúp tính toán dự báo dòng chảy phù hợp hơn.

#### 5.2.3.1. Giới thiệu mô hình 2 thông số

Do lưu vực sông Mê Công có diện tích lớn, thời gian chảy truyền từ thượng nguồn về hạ lưu dài, dòng chảy lũ tại đồng bằng sông Cửu Long thường lên xuống rất chậm nên dòng chảy có sự thay đổi không nhiều trong bước thời gian ngắn. Vì thế, việc sử dụng các mô hình thủy văn thời đoạn dài (10 ngày, tháng) sẽ giúp tăng tốc độ tính toán nhưng không làm giảm độ chính xác trong mô phỏng của bài toán mô phỏng và dự báo dòng chảy thời hạn dài.

Trong nghiên cứu này, mô hình cân bằng nước 2 thông số thời đoạn tháng được lựa chọn để mô phỏng và dự báo dòng chảy thử nghiệm cho lưu vực sông Mê Công.

Mô hình được Shenglian Guo và Xiong Lihua xây dựng vào năm 1999 (Xiong and Guo, 1999) sử dụng 2 thông số để tính toán dòng chảy từ mưa.

#### 5) Cấu trúc của mô hình

Trong mô hình cân bằng nước thời đoạn tháng, lượng bốc thoát hơi thực sự đóng vai trò quan trọng. Trên thực tế, người ta sử dụng các thông tin về lượng quan trắc bốc hơi chậu ở trên cánh đồng. Từ đó, để chuyển đổi từ bốc hơi chậu sang bốc hơi thực sự, lượng bốc hơi chậu được nhân với một hệ số triết giảm là một hàm của lượng nước trong đất trong lưu vực. Một phương trình thường dùng để xác định lượng bốc thoát hơi thực sự từ bốc hơi chậu:

$$E_t = EP_t * \tanh[P_t / EP_t] \quad (5.1)$$

Trong đó  $E_t$  biểu thị lượng bốc thoát hơi thực sự,  $EP_t$  là giá trị bốc hơi chậu năm,  $P_t$  là mưa năm, và  $\tanh()$  là hàm tang hyperbol. Xiong and Guo (1999) kiến nghị sử dụng phương trình (5.1) để tính lượng bốc thoát hơi thực sự từ bốc hơi chậu với vẻ phải có thêm một hệ số nhân. Phương trình tính bốc thoát hơi thực sự dùng trong mô hình cân bằng nước thời đoạn tháng 2 thông số là:

$$E_t = c * EP_t * \tanh[P_t / EP_t] \quad (5.2)$$

Trong đó c chính là thông số đầu tiên của mô hình. Hệ số c không thứ nguyên này được dùng để tính đến ảnh hưởng của việc chuyển đổi tỷ lệ thời gian từ năm sang tháng.

#### 6) *Tính toán dòng chảy thời đoạn tháng*

Dòng chảy thời đoạn tháng Q có liên quan mật thiết tới lượng trữ nước trong đất S. Trong các mô hình nhận thức thủy văn, các tác động điều tiết của một lưu vực với mưa được giả thiết vận hành như một hồ chứa tuyến tính hoặc phi tuyến (Shaw, 1994). Ở mô hình này, dòng chảy Q được giả thiết là một hàm tang hyperbol phụ thuộc vào lượng ẩm trong đất S như sau:

$$Q_t = S_t * \tanh[S_t / SC] \quad (5.3)$$

Trong đó  $Q_t$  là dòng chảy tháng,  $S_t$  là lượng ẩm trong đất, và SC biểu thị lượng ẩm tối đa. SC là thông số thứ hai được sử dụng trong mô hình này, có đơn vị là mm.

#### 7) *Các bước tính trong mô hình*

Khi có các số liệu quan trắc mưa  $P_t$  và bốc hơi chậu  $EP_t$  thời đoạn tháng, lượng bốc thoát hơi thực sự  $E_t$  thời đoạn tháng có thể được xác định bằng phương trình (5.2). Sau lượng tổn thất bốc thoát hơi, lượng nước trữ còn lại trong đất sẽ là  $[S_{t-1} + P_t - E_t]$ , trong đó  $S_{t-1}$  là lượng trữ nước trong đất ở cuối của thời điểm tháng (t-1) và bắt đầu của tháng t. Sau đó, sử dụng phương trình (5.3) để tính toán dòng chảy  $Q_t$  của tháng thứ t:

$$Q_t = [S_{t-1} + P_t - E_t] * \tanh\{[S_{t-1} + P_t - E_t]/SC\} \quad (5.4)$$

Cuối cùng, lượng nước trong đất  $S_t$  ở thời điểm cuối tháng thứ t được tính dựa trên định luật cân bằng nước:

$$S_t = S_{t-1} + P_t - E_t - Q_t \quad (5.5)$$

#### 8) *Mô phỏng dòng chảy cho lưu vực lớn*

Để ứng dụng mô hình 2 thông số cho lưu vực sông Mê Công, nghiên cứu phân chia lưu vực thành nhiều lưu vực nhỏ. Mỗi tiểu lưu vực được mô phỏng dòng chảy bằng mô hình 2 thông số và kết nối với nhau bằng diễn toán dòng chảy trên sông. Các tiểu lưu vực được phân chia theo nguyên tắc tương tự về mặt địa hình, thảm phủ, khí hậu và tại các vị trí cần tính toán. Hình 5.2 thể hiện các tiểu lưu vực được phân chia.

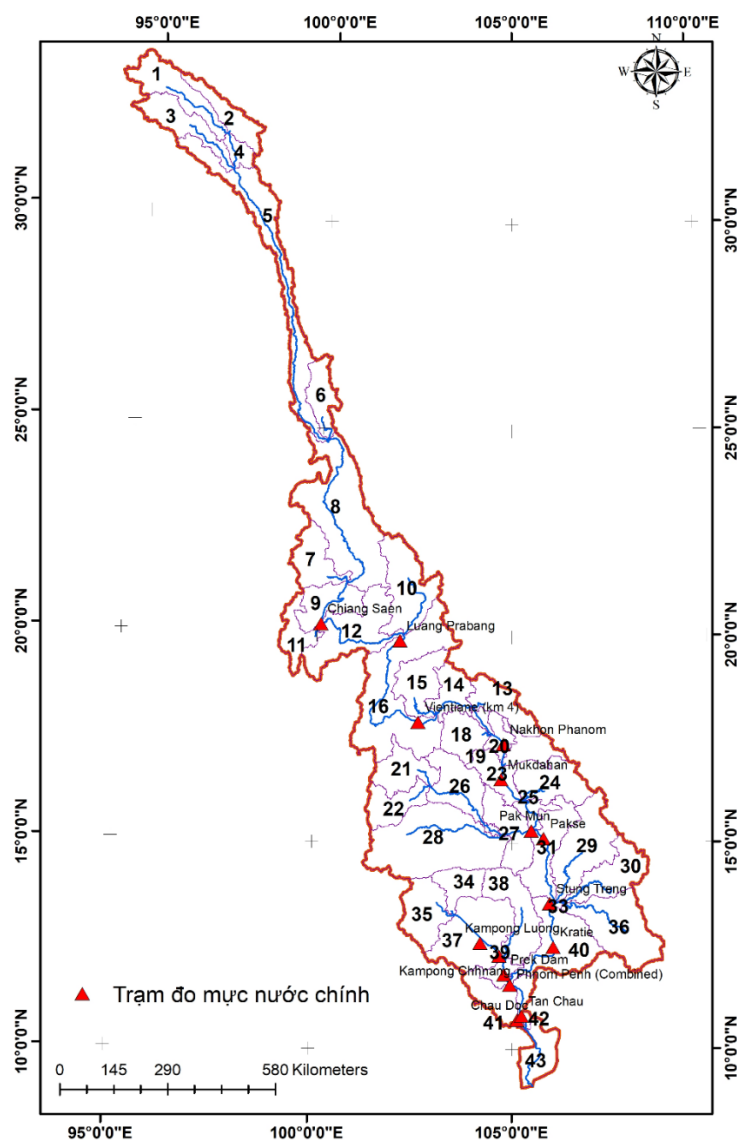
##### • *Mô phỏng vận hành hồ chứa*

Khi coi hồ chứa là một đoạn sông, có quá trình lưu lượng vào hồ  $Q(t)$  lưu lượng ra khỏi hồ là tổng của lưu lượng xả qua công trình đầu mối, lưu lượng qua công trình lấy nước hoặc lưu lượng qua tuốc bin đối với các hồ chứa phát điện và tổng lưu lượng tổn thất. Tính toán điều tiết lũ dựa trên nguyên lý cân bằng nước và phương trình động lực:

$$V_2 - V_1 = \frac{Q_1 + Q_2}{2} \Delta T - \frac{q_1 + q_2}{2} \Delta T$$

$$q = f(Z_t, Z_h, A)$$

Với  $Q_1$ ,  $Q_2$  là lưu lượng lũ đầu và cuối thời đoạn, các giá trị này đã biết;  $q_1$  là lưu lượng xả đầu thời đoạn đã xác định;  $q_2$  là lưu lượng xả cuối thời đoạn cần phải xác định.  $V_1$ ,  $V_2$  tương ứng là dung tích đầu và cuối thời đoạn tính toán.  $Z_t$ ,  $Z_h$  là mực nước thượng và hạ lưu hồ tại thời điểm tính toán.  $A$  là thông số hình thức biểu thị thông số công tác của công trình xả lũ.



Hình 5.3. Phân chia các tiểu lưu vực trong nghiên cứu

Trong hệ phương trình trên cần xác định 2 giá trị  $q_2$  và  $V_2$ , do vậy, tại thời đoạn bất kỳ các giá trị này được xác định bằng cách tính đúng dần.

- Vận hành theo quy trình

*Mực nước hồ dưới đường hccn (hạn chế cấp nước):* Khi mực nước hồ thấp hơn đường hccn của hồ thì hạn chế lượng nước xả ra khỏi hoặc ngưng xả để đảm bảo lượng nước trong thời gian của lại của vận hành công trình



*Mực nước hồ cao hơn đường phòng phá hoại (đường phòng lũ):* Khi mực nước hồ vượt đường phòng lũ của hồ cần phải xả lũ mở tất cả các cửa xả hoặc xả tối đa khả năng xả của công trình để đảm bảo an toàn cho hồ chứa.

*Khi lưu lượng xả ra khỏi hồ < q yêu cầu tối thiểu ở hạ du:* Nhìn chung các hồ chứa trên dòng chính đều phải duy trì xả một lượng nước nhất định nhằm đảm bảo duy trì dòng chảy tối thiểu. Nếu dòng chảy xả ở hai bước đầu nhỏ hơn q tối thiểu thì lượng xả thực tế của hồ phải bằng q tối thiểu.

#### **5.2.3.2. Mô hình mưa dòng chảy thời đoạn ngày**

Đối với bài toán dự báo dòng chảy cho lưu vực sông Mê Công, cách tiếp cận sử dụng mô hình mô phỏng thời đoạn ngày cũng có thể ứng dụng được. Tuy nhiên, các mô hình này cần phải có khả năng mô tả tốt diễn biến khí tượng – thủy văn theo không gian và thời gian.

Mô hình SWAT (Neitsch *et al.*, 2011) là một mô hình thông số bán phân bố có thể đáp ứng được các yêu cầu đặt ra. Số liệu đầu vào của mô hình SWAT bao gồm các dữ liệu bản đồ địa hình, thảm phủ, loại đất cũng như các số liệu mưa, nhiệt độ. Mô hình SWAT có khả năng mô phỏng diễn toán dòng chảy qua hồ chứa ở mức độ đơn giản.

#### **5.2.3.3. Các bước mô phỏng/dự báo dòng chảy bằng mô hình mưa – dòng chảy**

Mô hình mưa dòng chảy có lợi thế hơn hẳn các mô hình thống kê ở chỗ nó có thể mô tả tốt nhất bản chất vật lý của quá trình sinh dòng chảy từ mưa cũng như qua các công trình khai thác sử dụng nước. Do vậy, nó có khả năng xem xét được ảnh hưởng của các hồ chứa ở thượng nguồn đến dòng chảy ở hạ du, từ đó cho các kết quả dự báo/cảnh báo có độ tin cậy hơn. Khi kết hợp với dữ liệu mưa dự báo, mô hình mưa – dòng chảy có thể mô phỏng/dự báo dòng chảy với thời gian dự kiến dài hơn.

Trong nghiên cứu này, các bước ứng dụng mô hình mưa – dòng chảy dự báo dòng chảy trên sông Mê Công được thực hiện theo các bước sau:

**Bước 1:** Sử dụng công nghệ viễn thám mô phỏng/đánh giá diễn biến mực nước/thể tích của các hồ chứa lớn trên dòng chảy. Từ đó, xác định cơ bản dòng chảy đến hồ và dòng chảy ra khỏi hồ.

**Bước 2:** sử dụng mô hình mưa dòng chảy 2 thông số (hoặc SWAT) sử dụng kết hợp dữ liệu mưa thực đo và các nguồn mưa vệ tinh khác cho các khu vực không có trạm đo để mô phỏng dòng chảy đến các điểm kiểm soát (hồ chứa, trạm thủy văn...)

**Bước 3:** sử dụng mô hình WEAP kết nối với dữ liệu mô phỏng/dự báo dòng chảy tại các dòng nhánh, dòng chính để mô phỏng vận hành hệ thống liên hồ chứa chính có xét đến các quy trình vận hành của từng hồ chứa.

**Bước 4:** đánh giá kết quả mô phỏng/dự báo nhằm thực hiện hiệu chỉnh sau dự báo (hiệu chỉnh kết quả dự báo, hiệu chỉnh mô hình dự báo).



Hình 5.4: Sơ đồ hệ thống thủy văn trên lưu vực sông Mê Công và một số các hồ chứa chính (hình tam giác thể hiện vị trí nút hồ chứa)

Mô hình WEAP (Yates *et al.*, 2005) được sử dụng trong nghiên cứu này vì đây là một công cụ có khả năng tính toán mô phỏng dòng chảy, cân bằng nước tại các nút, các công

trình hồ chứa tốt. Mô hình có thể mô phỏng diễn toán dòng chảy qua hồ thông qua các quy trình của từng hồ được người sử dụng đưa vào. Mô hình rất thích hợp trong các bài toán đánh giá dự báo cho các hệ thống sông có nhiều các hoạt động liên quan như hồ chứa thủy lợi, thủy điện, chuyển nước cũng như số lượng các nút hoạt động lớn. Do vậy, mô hình WEAP rất phù hợp cho lưu vực sông Mê Công để thực hiện đánh giá tác động của hệ thống liên hồ cũng như các công trình lấy nước của lưu vực sông này.

### 5.3. Kết quả dự báo thử nghiệm dòng chảy hạn dài một số vị trí trên lưu vực sông Mê Công

#### 5.3.1. Thời gian chảy truyền

Thời gian chảy truyền trên sông là thời gian để chất điểm nước di chuyển từ đầu đoạn sông đến cuối đoạn sông. Việc phân tích thời gian chảy truyền giữa các vị trí (các trạm đo dòng chảy, các hồ chứa...) sẽ giúp ích cho việc hiệu chỉnh thời gian trễ cho các mô hình mô phỏng dòng chảy tại các vị trí bên dưới.

Để xác định thời gian chảy truyền giữa các trạm đo, người ta có thể dựa trên phương pháp hình trực quan, vẽ các quá trình dòng chảy, hoặc so sánh thời gian trễ giữa các trị số cực trị, hoặc phân tích tương quan giữa các chuỗi dòng chảy với các thời gian lệch khác nhau.

Từ các phương pháp phân tích ở trên, nhìn chung thì thời gian chảy truyền của dòng chảy giữa một số trạm thượng nguồn và hạ nguồn lưu vực sông Mê Công được thể hiện ở bảng dưới (trị số âm thể hiện thời gian xuất hiện trước).

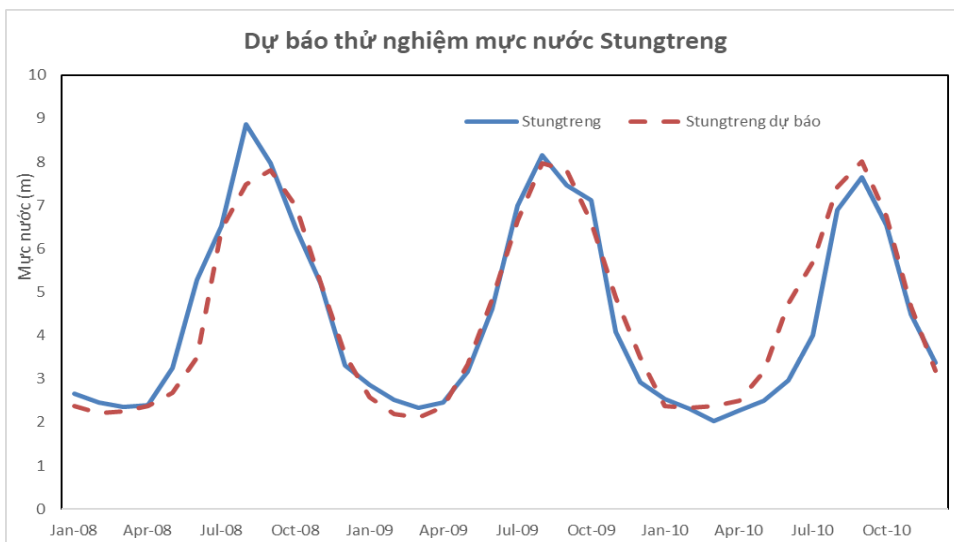
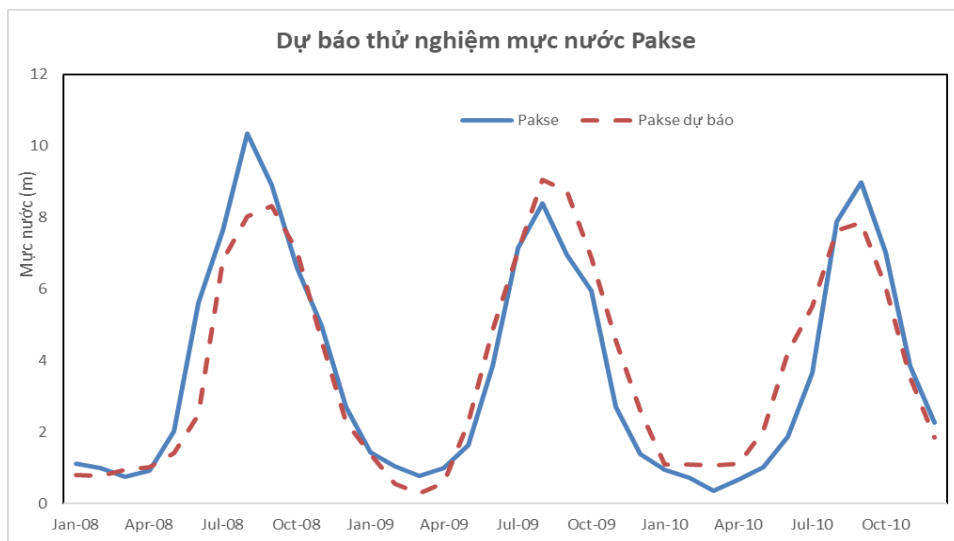
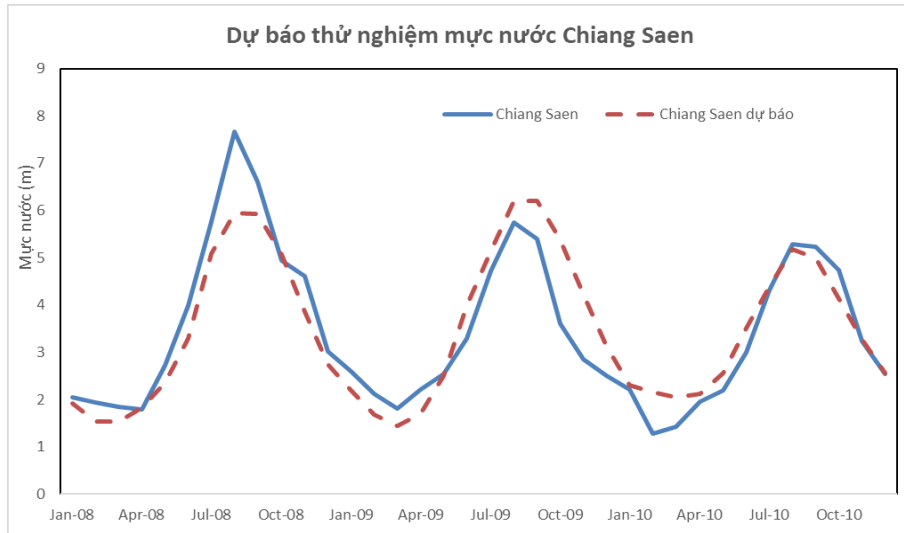
*Bảng 5.1: Thời gian chảy truyền trên sông trung bình*

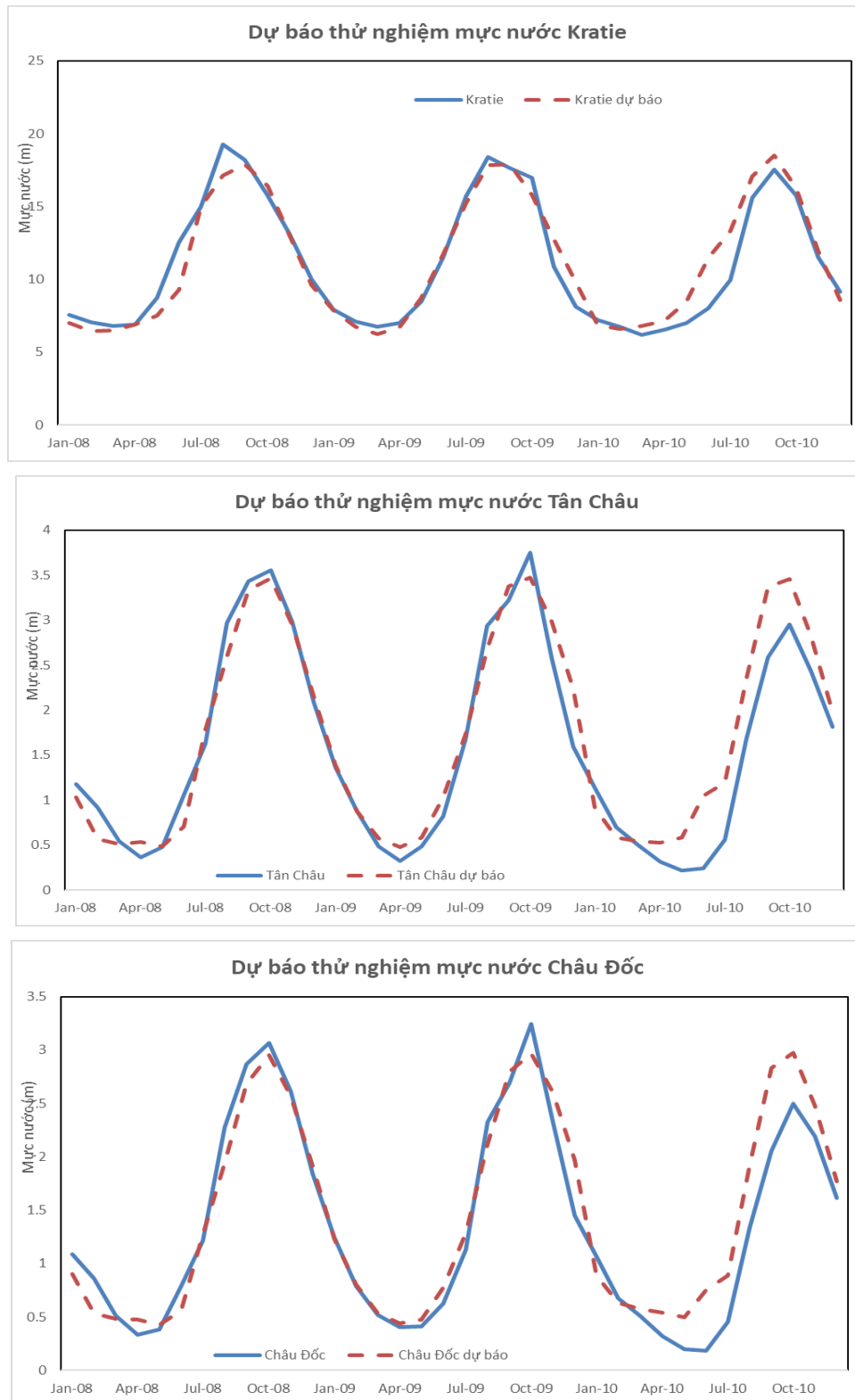
Trạm	Stung Treng	Kratie	Kampong Cham	Vientiane	Pakse	Chiang Saen	Nong Khai	Tân Châu
Thời gian trễ (ngày)	-7	-6	-5	-15	-9	-17	-15	0

#### 5.3.2. Dự báo thử nghiệm bằng phương pháp tự hồi quy AR(p)

Nghiên cứu tiến hành thử nghiệm dự báo mực nước trung bình tháng thời hạn 6 tháng sử dụng phương pháp tự hồi quy cho một số trạm trên dòng chính, đặc biệt là các trạm ở thượng nguồn vùng hạ lưu sông Mê Công (Chiang Saen) và trạm hạ nguồn (Tân Châu, Châu Đốc).

Dữ liệu “học” được lấy từ năm 1980 đến 2007, dự báo thử nghiệm cho năm 2008, 2009, 2010 với thời gian dự kiến dự báo là 6 tháng. Dữ liệu học sẽ được cập nhật liên tục đến thời điểm phát báo. Kết quả dự báo thử nghiệm cho một số vị trí trên dòng chính sông Mê Công được trình bày tóm tắt trong các hình dưới đây:





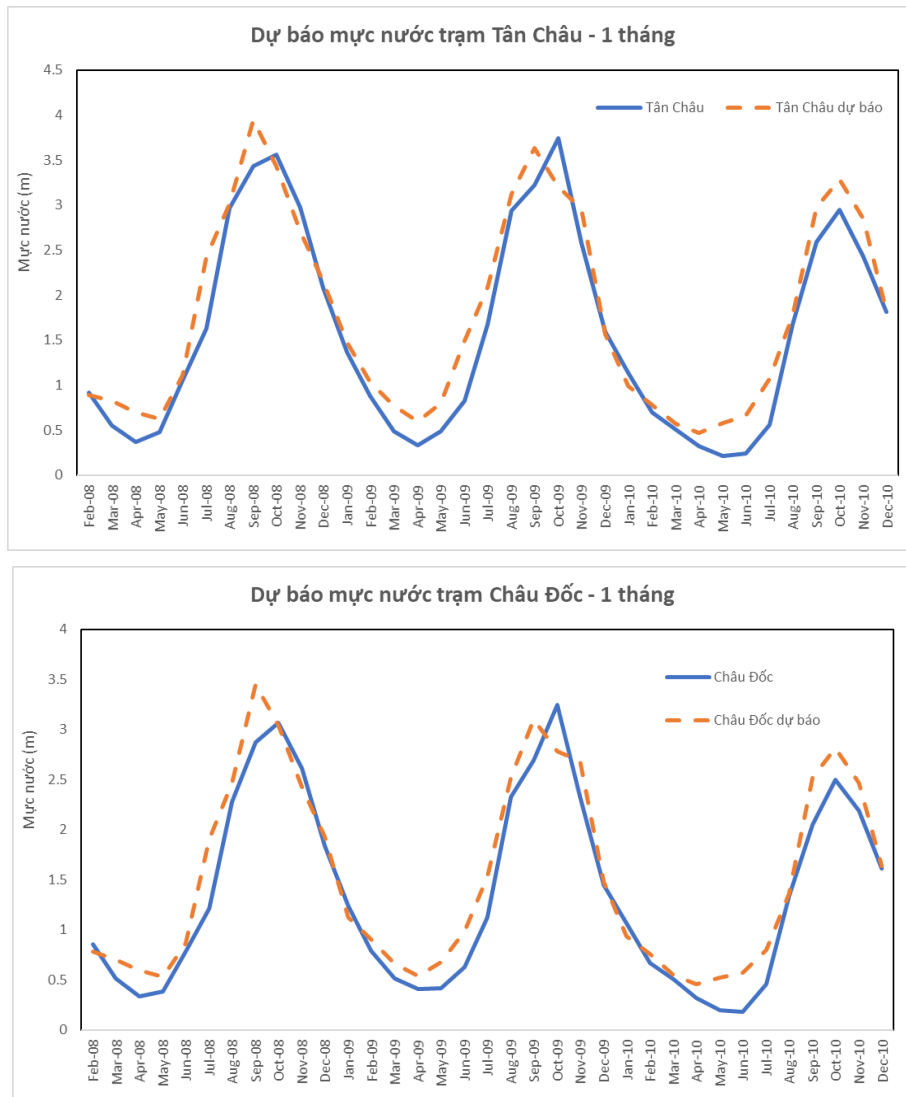
Hình 5.5: Đường quá trình mực nước dự báo và thực đo tại một số vị trí trên dòng chính sông Mê Công

Bảng 5.2: Tổng hợp kết quả dự báo dòng chảy thời hạn tháng thử nghiệm cho một số vị trí trên dòng chính sông Mê Công

Trạm	$R^2$	$\Delta H(m)$
Chiang Saen	0.91	0.49
Pakse	0.93	0.81
Stungcheng	0.95	0.44
Kratie	0.96	0.34
Tân Châu	0.96	0.25
Châu Đốc	0.96	0.21

Nhận xét: nhìn chung, phương pháp tự hồi quy có thể dự báo với kết quả chấp nhận được với hệ số tương quan tốt. Tuy nhiên sai số mực nước trung bình vẫn còn cao.

### 5.3.3. Dự báo thử nghiệm bằng phương pháp hồi quy đa biến



Hình 5.6: Đường quá trình mực nước dự báo và thực đo tại Tân Châu và Châu Đốc

Phương pháp hồi quy đa biến phân tích mối quan hệ giữa mực nước tại các trạm Tân Châu và Châu Đốc với dữ liệu mực nước các trạm thượng nguồn với thời gian lệch là 1 tháng. Hàm hồi quy có dạng:

$$H_{\text{Tân Châu}}(t+1) = b_0 + \sum_i^k H_{TL}(t) \cdot b_i$$

Trong đó:  $H_{\text{Tân Châu}}$  là mực nước ở hạ lưu ;  $H_{TL}$  là mực nước các trạm thượng nguồn ;  $k$  số biến độc lập ;  $b_i$  là các hệ số hằng số. Các hệ số này được cập nhật lại sau từng tháng.

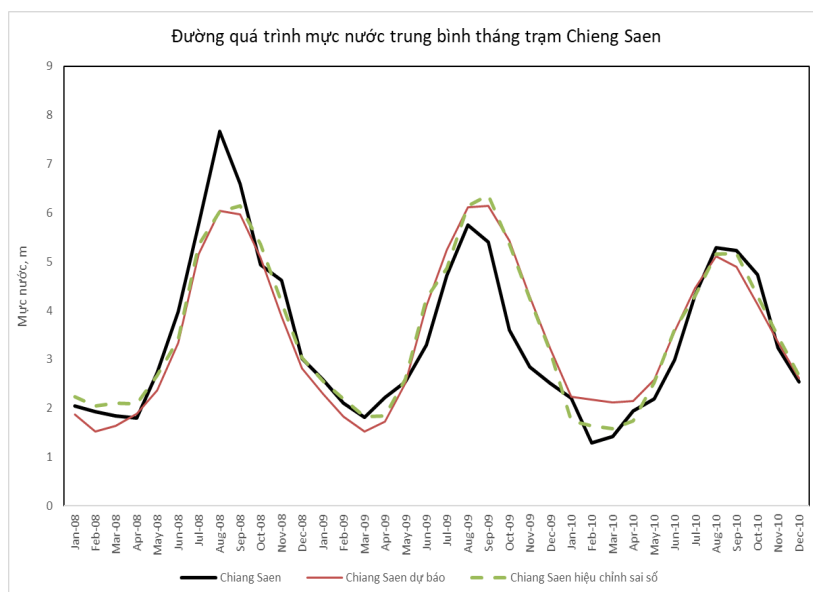
Dựa trên số liệu thu thập và phân tích, kết quả dự báo dự kiến 1 tháng tại Tân Châu và Châu Đốc được thể hiện ở hình 5.5.

*Bảng 5.3: Tổng hợp kết quả dự báo dòng chảy thời hạn tháng thử nghiệm cho trạm Tân Châu và Châu Đốc*

Trạm	R <sup>2</sup>	ΔH(m)
Tân Châu	0.97	0.26
Châu Đốc	0.97	0.23

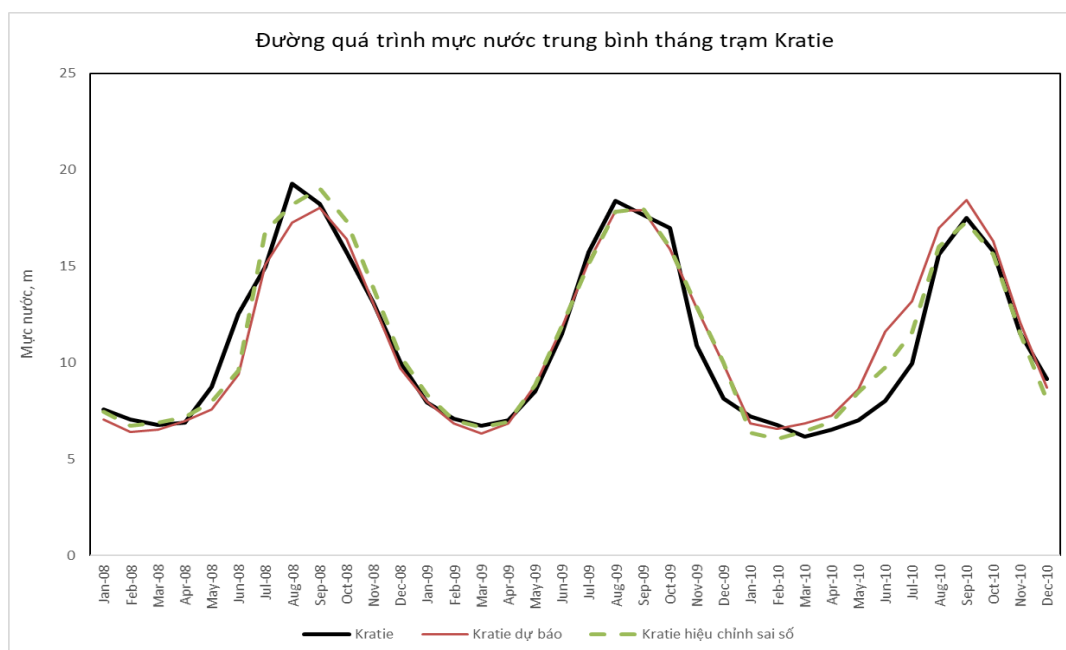
Nhận xét: Nhìn chung, phương pháp hồi quy tuyến tính đa biến cho kết quả dự báo không có sự khác biệt đáng kể về chỉ tiêu R<sup>2</sup> và sai số tuyệt đối trung bình cho 2 trạm Tân Châu và Châu Đốc so với phương pháp tự hồi quy. Tuy nhiên, phương pháp dự báo hồi quy tuyến tính đa biến có ưu điểm có thể xem xét các kịch bản khác nhau, các trường hợp khác nhau, đặc biệt là có thể kết hợp với mưa dự báo trong việc kéo dài thời gian dự kiến dự báo.

#### 5.3.4. Dự báo thử nghiệm bằng phương pháp kết hợp hồi quy nhiều biến với mô đun hiệu chỉnh sai số ARMA

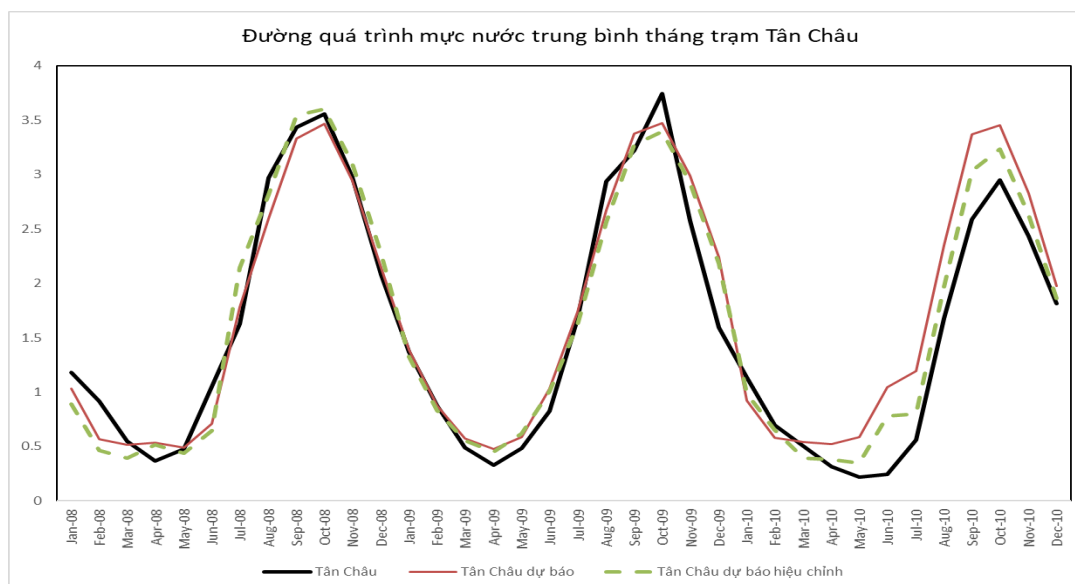


*Hình 5.7: Dự báo mực nước trung bình tháng thử nghiệm tại trạm Chiang Saen*

Từ chuỗi sai số trong mô phỏng trong quá khứ so sánh với thực đo, ta có được chuỗi sai số dự báo. Chuỗi sai số dự báo này là chuỗi số dừng có tổng bằng 0 và phương sai bằng 1 nên có thể sử dụng mô hình thống kê ARMA để dự báo sai số dự báo. Đề tài đã thực hiện dự báo thử nghiệm kết hợp hiệu chỉnh sai số. Các dữ liệu hiệu chỉnh sai số luôn được cập nhật trong từng bước dự báo. Kết quả dự báo thử nghiệm kết hợp với hiệu chỉnh sai số bằng mô hình ARMA được minh họa trong ở các hình 5.6 đến 5.9.

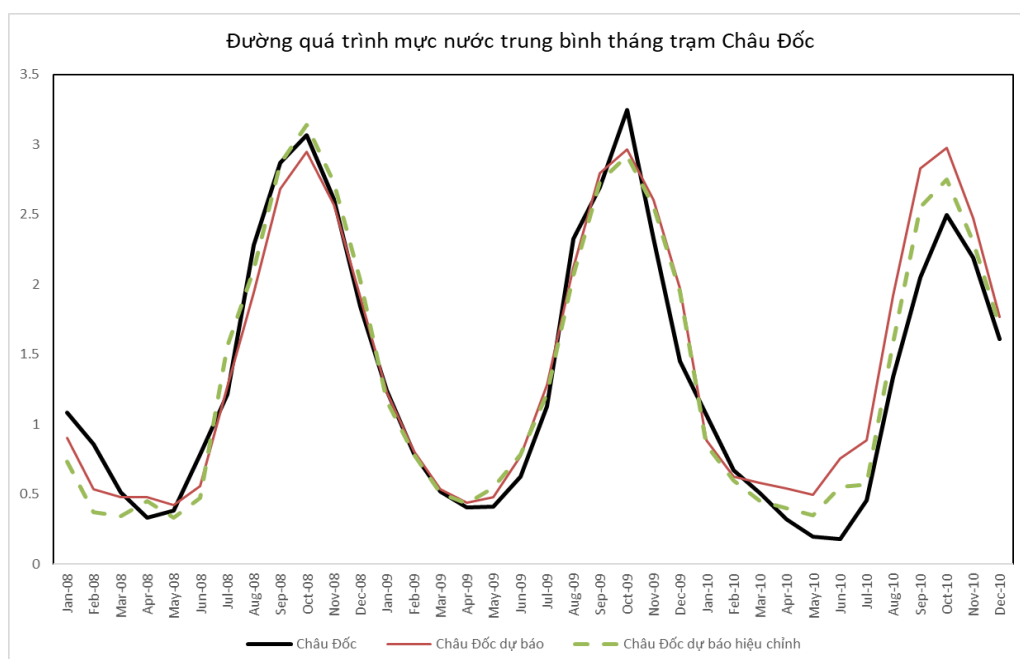


Hình 5.8: Dự báo mực nước trung bình tháng thử nghiệm tại trạm Kratie



Hình 5.9: Dự báo mực nước trung bình tháng thử nghiệm tại trạm Tân Châu





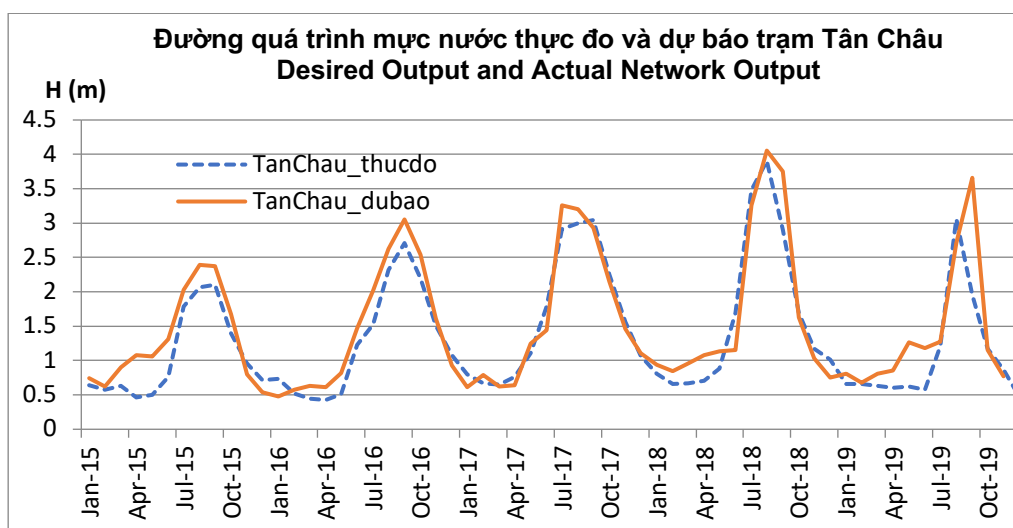
Hình 5.10: Dự báo mực nước trung bình tháng thử nghiệm tại trạm Châu Đốc

Kết quả dự báo thử nghiệm cho các trạm Chiang Saen, Kratie, Tân Châu và Châu Đốc sử dụng phương pháp tự hồi quy và phương pháp tự hồi quy kết hợp mô đun hiệu chỉnh sai số ARMA được trình bày ở hình trên cho thấy: phương pháp dự báo kết hợp với hiệu chỉnh sai số có thể cải thiện chất lượng dự báo. Trong năm đầu tiên dự báo thử nghiệm (2008), do chuỗi số liệu sai số được “học” còn chưa nhiều nên việc hiệu chỉnh sai số còn có sự chênh lệch đáng kể (thậm chí còn cho kết quả xấu hơn là không sử dụng phương pháp hiệu chỉnh). Tuy nhiên đến các năm dự báo sau, kết quả dự báo có hiệu chỉnh đã cải thiện dần kết quả dự báo. Đường xu thế của dự báo có hiệu chỉnh nhìn chung khá tương đồng với đường dự báo không hiệu chỉnh nhưng kết quả dự báo có xu thế sát với đường thực đo hơn.

Điều này cho thấy tiềm năng ứng dụng các phương pháp hiệu chỉnh sai số thực hiện sau dự báo nhằm cải thiện dự báo dòng chảy hạn dài.

### 5.3.5. Dự báo thử nghiệm bằng phương pháp Trí tuệ nhân tạo EANN

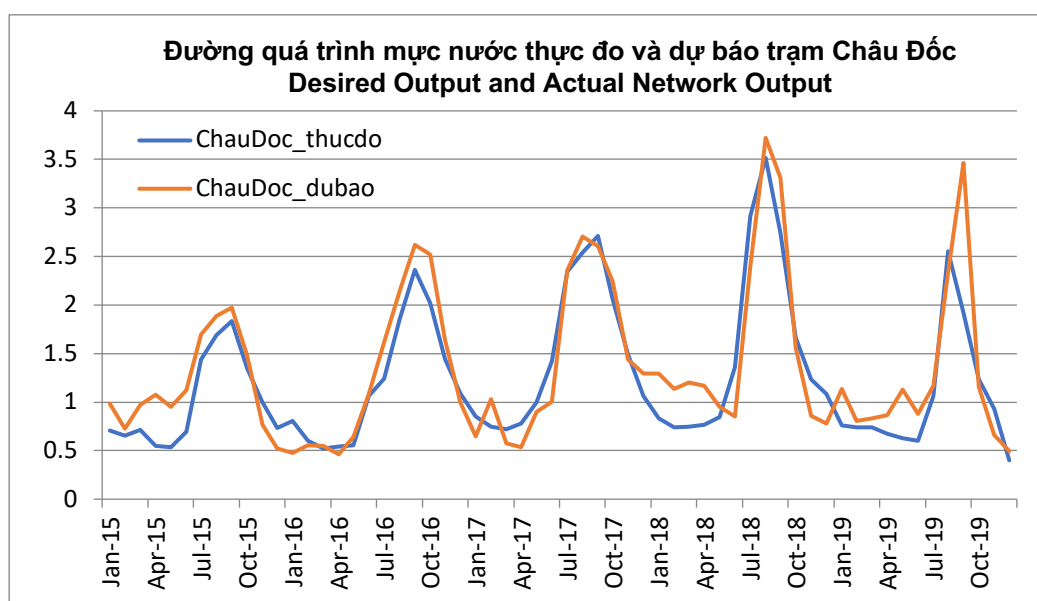
Kết quả xây dựng mạng EANN và dự báo thử nghiệm dòng chảy trung bình tháng cho các trạm Tân Châu, Châu Đốc và Chiang Sean được trình bày trong các hình 5.11 đến 5.13 và các bảng 5.4 đến 5.6.



Hình 5.11: Dự báo mực nước trung bình tháng thử nghiệm tại trạm Chieng Sean bằng mô hình EANN

Bảng 5.4: Đánh giá kết quả dự báo dòng chảy thời hạn tháng thử nghiệm cho trạm Tân Châu bằng mô hình EANN

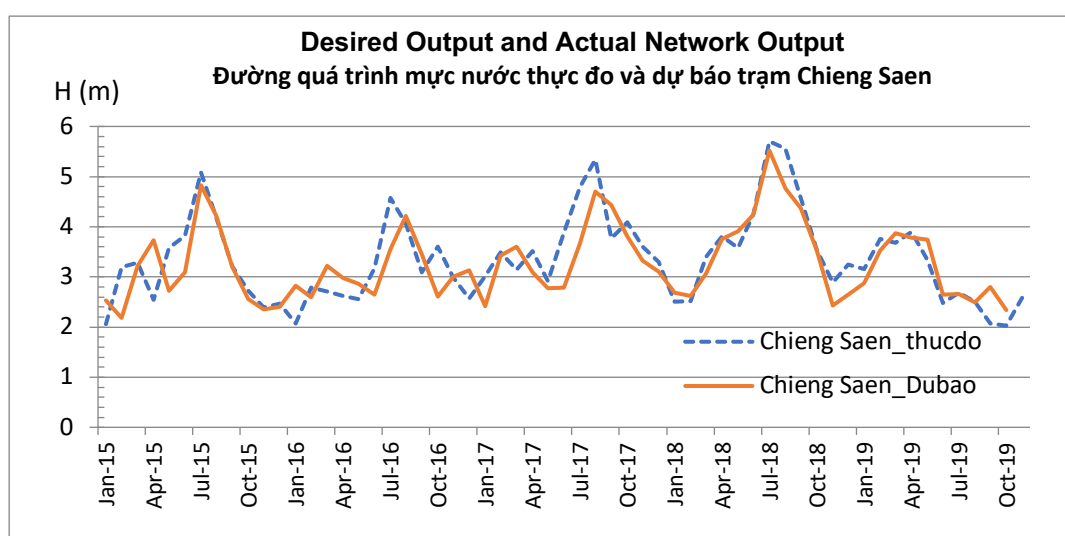
Chỉ số đánh giá	Dự báo trạm Tân Châu
MSE (sai số toàn phương trung bình)	0.14
NMSE (sai số tuyệt đối trung bình hiệu chỉnh)	0.18
MAE (sai số tuyệt đối trung bình)	0.27
Min Abs Error (sai số dự báo nhỏ nhất)	0.02
Max Abs Error (sai số dự báo lớn nhất)	1.70
R (hệ số tương quan giữa thực đo và dự báo)	0.93



Hình 5.12: Dự báo mực nước trung bình tháng thử nghiệm tại trạm Châu Đốc bằng mô hình EANN

*Bảng 5.5: Đánh giá kết quả dự báo dòng chảy thời hạn tháng thử nghiệm cho trạm Châu Đốc bằng mô hình EANN*

<b>Chỉ số đánh giá</b>	<b>Dự báo trạm Châu Đốc</b>
MSE (sai số toàn phương trung bình)	0.12
NMSE (sai số tuyệt đối trung bình hiệu chỉnh)	0.24
MAE (sai số tuyệt đối trung bình)	0.27
Min Abs Error (sai số dự báo nhỏ nhất)	0.01
Max Abs Error (sai số dự báo lớn nhất)	1.54
R (hệ số tương quan giữa thực đo và dự báo)	0.91



*Hình 5.13: Dự báo mực nước trung bình tháng thử nghiệm tại trạm Chieng Sean bằng mô hình EANN*

*Bảng 5.6: Đánh giá kết quả dự báo dòng chảy thời hạn tháng thử nghiệm cho trạm Chieng Sean bằng mô hình EANN*

<b>Chỉ số đánh giá</b>	<b>Chieng Sean du bao</b>
MSE (sai số toàn phương trung bình)	0.25
NMSE (sai số tuyệt đối trung bình hiệu chỉnh)	0.34
MAE (sai số tuyệt đối trung bình)	0.38
Min Abs Error (sai số dự báo nhỏ nhất)	0.01
Max Abs Error (sai số dự báo lớn nhất)	1.19
R (hệ số tương quan giữa thực đo và dự báo)	0.82

### 5.3.6. Dự báo thử nghiệm bằng mô hình mưa - dòng chảy

#### 5.3.6.1. Dữ liệu đầu vào

Dữ liệu Aphrodite bao gồm mưa và nhiệt độ được sử dụng cho các vùng không có số liệu ở lưu vực sông Mê Công. Hình 5.10 mô tả vị trí các ô lưới của bộ dữ liệu này. Đầu vào của mô hình thủy văn mưa dòng chảy là các dữ liệu về mưa và bốc hơi. Các dữ liệu