ỨNG DỤNG MÔ HÌNH THỐNG KÊ DỰ BÁO DÒNG CHẢY THÁNG PHỤC VỤ QUY TRÌNH VẬN HÀNH LIÊN HỒ CHỨA SÔNG SÊ SAN TRONG MÙA CẠN

Trần Hồng Thái[1], Phùng Tiến Dũng[2], Đoàn Văn Hải[2], Đoàn Quang Trí[3], Dương Quốc Hùng[1]

Tóm tắt: *Trong Quy trình vận hành liên hồ chứa (QTVHLHC) sông Sê San quy định trách nhiệm của Trung tâm KTTV quốc gia phải thực hiện dự báo hạn dài (tháng) mùa cạn đối với các vị trí thủy văn khống chế trên lưu vực sông Sê San cung cấp cho các cơ quan chức năng để phục vụ QTVHLHC. Nhóm nghiên cứu ứng dụng mô hình thống kê Hồi quy đa biến để xây dựng phần mềm dự báo đặc trưng dòng chảy tháng trên sông Sê San. Số liệu đưa vào dùng để hiệu chỉnh mô hình từ 1990 - 2010 và kết quả kiểm định năm 2011 - 2014 cho kết quả khá tốt. Từ những kết quả đó nhóm nghiên cứu xây dựng phần mềm có khả năng hỗ trợ cho các dự báo viên trong quá trình tác nghiệp dự báo phục vụ VHLHC và đáp ứng các yêu cầu ngày càng cao của xã hội dưới sự phát triển không ngừng của hệ thống hồ chứa thủy điện trên sông Sê San. Đây có thể là một công cụ hiệu quả phục vụ tốt cho công tác dự báo nghiệp vụ trong tương lai.*

Từ khóa: *Đặc trưng dòng chảy tháng, Sông Sê San, Mô hình hồi qui bội, Công nghệ dự báo hạn dài.*

Ban Biên tập nhận bài: 15/3/2017 Ngày phản biện xong: 20/4/2017

1. Mở đầu

Hiện nay nhóm phương pháp thống kê và hồi quy được sử dụng trong dự báo thủy văn hạn vừa, hạn dài khá phổ biến phổ biến trên thế giới và cả ở nước ta. Mô hình thống kê dự báo dòng chảy mùa được phát triển trên cơ sở mối quan hệ trong quá khứ của dòng chảy và dự báo khí hậu. Các phương pháp thống kê để dự báo dòng chảy theo mùa đã được phát triển nhiều thập kỷ trước vì sự tương quan đáng kể giữa lượng mưa trong tương lai hoặc dòng chảy và chỉ số khí hậu [1, 2, 3], tiếp tục được phát triển gần đây. Phương pháp *Bayesian Joint Probability (BJP)*, kết hợp giữa chỉ số khí hậu và dòng chảy quá khứ để dự báo dòng chảy mùa trước đã được phát triển và thử nghiệm ở nhiều địa điểm ở Úc [4, 5, 6]. Một vài phương pháp trong dự báo dòng chảy theo mùa đã được nghiên cứu như: (1) Phương pháp hồi quy đa biến có thể được giới thiệu trong dự báo dòng chảy mùa có thể giải quyết với sự đa cộng tuyến của các biến dự báo. Phương pháp này hiện đang được sử dụng chủ yếu trong hồi quy và hồi quy *Z-score*, nó có hiệu quả trong việc loại bỏ các vấn đề cộng tuyến. Ngoài ra các phương pháp khác như phân tích tương quan chuẩn *(Canonical Correlation Analysis (CCA))*, hồi quy bình phương nhỏ nhất *(Least Squares Regression (LSR))* có thể được áp dụng để dự báo dòng chảy theo mùa để ứng phó với các vấn đề cộng tuyến. Tuy nhiên, việc áp dụng các phương pháp thay thế trong dự báo dòng chảy mùa chỉ áp dụng ở giai đoạn đầu; (2) Phương pháp chính phục vụ dự báo dòng chảy theo mùa ở miền Tây Hoa Kỳ vẫn dựa trên phương pháp hồi quy. Risley và các cộng sự (2005) [7] đã phát hiện việc ứng dụng của mạng thần kinh nhân tạo (ANN) phục vụ dự báo dòng chảy theo mùa. Mặc dù mục tiêu cải thiện độ chính xác dự báo bằng cách sử dụng mạng thần kinh không phải là kết luận trong nghiên cứu của họ, nhưng nó cung cấp một hướng mới cho việc áp dụng các phương pháp

[1]*Trung tâm Khí tượng Thủy văn quốc gia*
[2]*Trung tâm Dự báo Khí tượng Thủy văn Trung ương*
[3]*Tạp chí Khí tượng Thủy văn*
Email: ptdung77@mail.com
Email: doanquangtrikttv@gmail.com

phức tạp hơn trong dự báo dòng chảy theo mùa; (3) Có hàng trăm các biến đầu vào có sẵn cho dòng chảy theo mùa trong việc phát triển phương trình dự báo. Việc lựa chọn các biến quan trọng và xây dựng các phương trình dự báo đáng tin cậy luôn luôn đầy thách thức. Ngoài ra, tìm kiếm sự kết hợp các biến tối ưu hoặc gần tối ưu như đề xuất của Garen (1992) [8], sự phát triển các chỉ số hỗn hợp và sử dụng chúng làm đầu vào cho phương trình hồi quy đa biến cũng cần phải được kiểm tra trong quá trình xây dựng phương trình dự báo dòng chảy theo mùa.

Ở Việt Nam, dự báo hạn vừa, hạn dài mùa cạn đã được quan tâm từ những năm 60. Một số các phương pháp dự báo hạn vừa, hạn dài của nước ngoài đã được nghiên cứu ứng dụng vào dự báo dòng chảy 10 ngày, tháng, mùa, năm tại một số trạm cơ bản trên hệ thống sông Hồng, sông Thái Bình, như xây dựng quan hệ dòng chảy tháng với sự biến đổi của các dạng hoàn lưu khí quyển; phương pháp thống kê khách quan; phương pháp diễn biến lịch sử của Dương Giám Sơ; phương pháp Vine-Hop, khai triển chuỗi dòng chảy năm dưới dạng tổng của các hàm điều hoà; phương pháp động lực thống kê của M. Alôkhin; phương pháp phân tích phân lớp; phương pháp tương tự; phương pháp phân tích tổng hợp... Có thể tổng hợp các phương pháp dự báo hạn vừa, hạn dài vào 3 nhóm chính: Nhóm 1: Mô hình tương quan với hoàn lưu khí quyển, khí hậu dựa trên cơ sở xây dựng mối quan hệ của dòng chảy mùa cạn, tháng với chỉ tiêu hoạt động của mặt trời, các dạng hoàn lưu khí quyển, các yếu tố khí hậu mặt đất hoặc trên cao. Đối với các nước trong khu vực nhiệt đới gió mùa như nước ta, thường xây dựng các mối quan hệ dòng chảy với các chỉ tiêu hoạt động của mặt trời, các dạng hoàn lưu khí quyển, chỉ tiêu hoạt động của Dao động Nam bán cầu (SOI), về hiện tượng El-Nino, La Nina,... cũng như các yếu tố khí hậu mặt đất và trên cao; Nhóm 2: Các mô hình nhận thức được xây dựng dựa vào cơ sở vật lý của các mối quan hệ giữa dòng chảy và nhân tố ảnh hưởng. Nhóm này có thể dùng để dự báo dòng chảy hạn vừa cho sông tự nhiên cũng như có điều tiết của hồ chứa nước. Tuy nhiên, thời gian dự kiến của chúng còn hạn chế, phụ thuộc nhiều vào dự báo lượng mưa trong thời gian dự kiến; mức đảm bảo còn hạn chế trong các tháng giao thời giữa hai mùa. Mặc dù vậy, cùng với sự phát triển của dự báo mưa số trị với công nghệ, kỹ thuật mới, nhóm này có nhiều triển vọng ứng dụng tốt vào dự báo hạn vừa dòng chảy cạn trên các sông suối ở Việt Nam; Nhóm 3: Các phương pháp nhận dạng - tương tự và phương pháp thống kê xác suất, phương pháp thống kê khách quan cũng được sử dụng trong dự báo hạn dài dòng chảy sông. Phương pháp nhận dạng tương tự được dùng ở nhiều dạng khác nhau, từ đơn giản, với một hoặc hai nhân tố dự báo, đến phức tạp với hàng trăm nhân tố, hàng chục loại số liệu khác nhau. Phương pháp đã được nghiên cứu ứng dụng trong dự báo hạn vừa, dự báo tháng và phân phối dòng chảy các tháng trong năm.

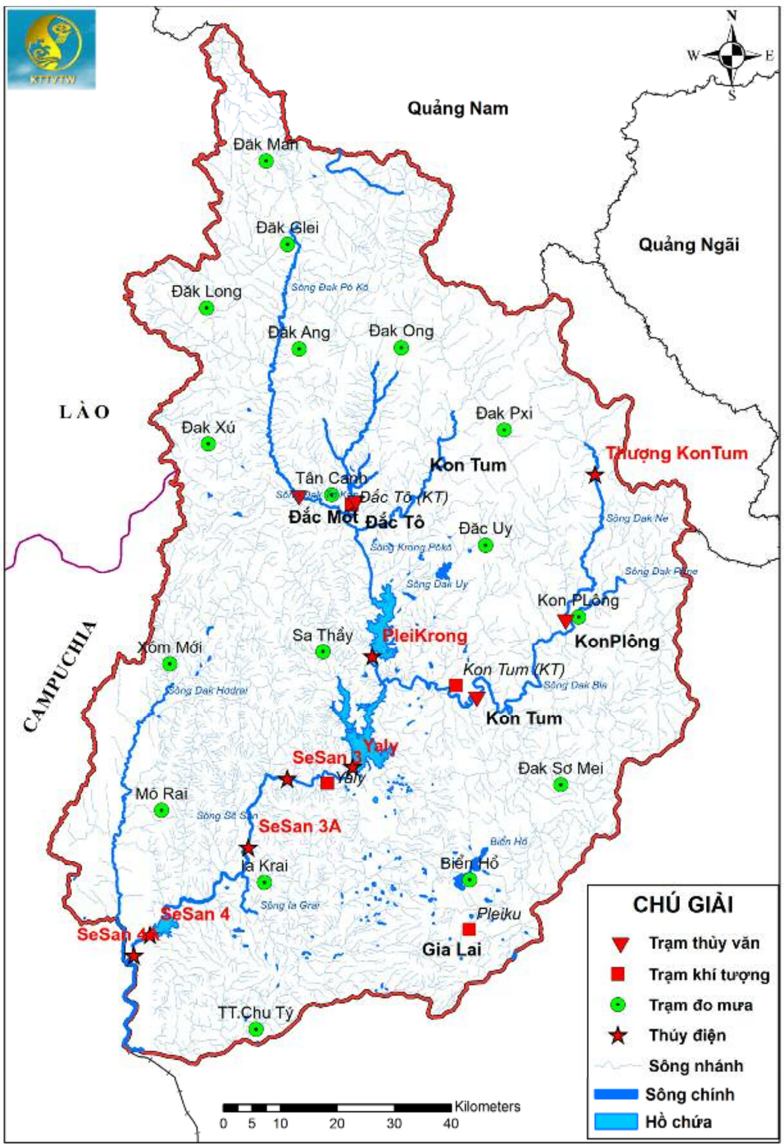
Mục tiêu của nghiên cứu này là: (1) Xây dựng được phương án dự báo phục vụ dự báo dòng chảy tháng trong mùa cạn trên sông Sê San; (2) Giới thiệu phần mềm phục vụ dự báo dòng chảy tháng trong mùa cạn trên sông Sê San, đây chính là một công cụ hỗ trợ hữu hiệu đối với các dự báo viên trong quá trình tác nghiệp dự báo phục vụ VHLHC.

2. Phương pháp nghiên cứu và tài liệu thu thập

2.1 Tổng quan khu vực nghiên cứu

Sông Sê San là một trong các nhánh lớn của lưu vực hạ du sông Mê Kông. Sông Sê San được bắt nguồn từ vùng núi cao Ngọc Linh tỉnh Kon Tum thuộc phía Bắc Tây Nguyên của Việt Nam, chảy sang Campuchia và sau nhập với hạ lưu các sông Srêpôk, SêKông sau đó nhập vào sông Mê Kông ở Strung Treng. Trên lãnh thổ Việt Nam, sông Sê San nằm trên hai tỉnh Kon Tum và Gia Lai với chiều dài 230 km,diện tích lưu vực 11.620 km^2. Lưu vực có tọa độ địa lý 13^{0}45’ - 15^{0}14’ vĩ độ Bắc; toạ độ 107^{0}10’ - 108^{0}24’ kinh độ Đông (Hình 1). Sông Sê San có mật độ lưới sông vào loại trung bình. Đổ vào dòng chính Sê San có 27 nhánh sông suối lớn nhỏ, nhỏ nhất là suối Đắc Mi có diện tích lưu vực là 20 km^2 và lớn nhất là lưu vực sông Đak Bla có diện tích lưu vực là 3507 km^2. Những nhánh đổ vào dòng chính Sê San phải kể đến là các nhánh: Đăk PSi, Đăk Bla, Krong PoCo, Sa Thầy, Đăk Tơ Kan [9].

Lưu vực sông Sê San nằm trên vùng khí hậu nhiệt đới gió mùa cao nguyên, theo đó khí hậu được chia làm hai mùa tách biệt với sự tương phản sâu sắc. Mùa lũ trong QTVHLHC quy định từ tháng 7 - 11, với thời tiết ẩm ướt, mưa nhiều, mùa khô từ tháng 12 - 6, với thời tiết khô, ít mưa. Lượng mưa trên lưu vực sông Sê San phân hóa sâu sắc và biến động khá phức tạp theo năm, mùa và theo không gian. Tổng lượng mưa mùa khô trên lưu vực chỉ chiếm khoảng 1/3 tổng lượng mưa năm.



Hình 1. Bản đồ vị trí lưu vực sông Sê San

2.2 Số liệu xây dựng phương án

Số liệu dòng chảy

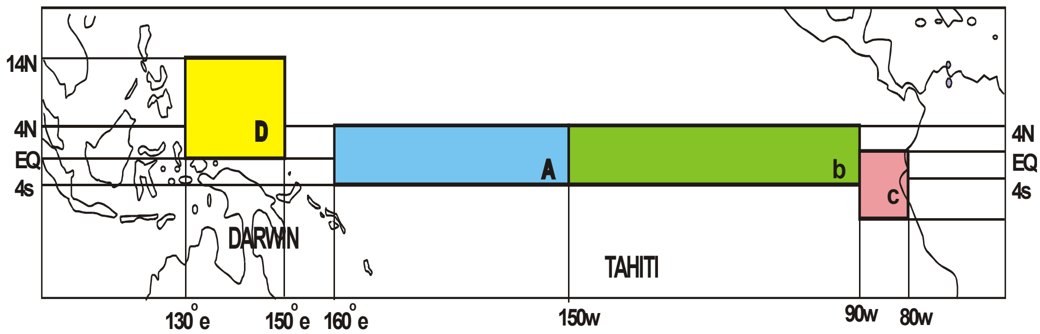
Để đặc trưng cho lượng trữ nước trong sông và khả năng điều tiết dòng chảy của lưu vực nhóm nghiên cứu sử dụng số liệu dòng chảy lưu lượng trung bình tháng và lưu lượng ngày cuối tháng trước tại bốn trạm thuỷ văn: Kon Plông, Kon Tum, Đăk Tô, Đăk Mốt và bốn vị trí hồ chứa: Pleikrông, Ialy, Sê San 4 và Sê San 4A.

Số liệu khí tượng

Để chọn các yếu tố khí tượng đưa vào cơ sở số liệu ban đầu, đã tiến hành đánh giá về chất lượng, độ dài và tính liên tục của số liệu quan trắc về các yếu tố khí tượng tại các trạm khí tượng mặt đất trên lưu vực sông Sê San và lân cận. Kết quả đã chọn được bốn trạm khí tượng: Đắk Tô, Kon Tum, Yaly, Pleiku có số liệu quan trắc đầy đủ, đặc trưng cho khí hậu khu vực nghiên cứu được trình bày trong bảng 1.

Bảng 1. Các nhân tố khí tượng dùng để dự báo đặc trưng thủy văn

1. Số ngày mưa trong tháng	6. Độ ẩm tương đối trung bình tháng
2. Tổng lượng mưa tháng	7. Độ ẩm nhỏ nhất tháng
3. Nhiệt độ trung bình tháng	8. Tổng lượng bốc hơi tháng
4. Nhiệt độ cao nhất tháng	9. Tổng số giờ nắng trong tháng
5. Nhiệt độ thấp nhất tháng	



Hình 2. Nhiệt độ nước biển lớp mặt

Số liệu về hoàn lưu

Số liệu về hoàn lưu dưới dạng chuẩn sai tháng của các đặc trưng sau đã được sử dụng:

- Chỉ số dao động nam bán cầu SOI;

- Nhiệt độ nước biển lớp mặt khu El Nino A, khu El Nino B, khu El Nino C và khu El Nino D

Vị trí áp cao Thái Bình Dương được biểu thị bằng đường đẳng áp trên bản đồ AT500mb bao gồm: trạm số 58847, số 91115, số 91131, số 91245, số 98327.

Để thống nhất với liệt số liệu của yếu tố dự báo, số liệu của các nhân tố dự báo cũng đã được thu thập từ năm 1990 đến nay. Bộ số liệu các nhân tố dự báo khá đầy đủ và liên tục. Những số liệu đã được tính toán, bổ sung theo số liệu các trạm tương tự và chỉnh biên theo các quy trình quy định của ngành khí tượng thủy văn.

2.2 Phân tích ảnh hưởng của ENSO đến dòng chảy tháng mùa cạn

a) Ảnh hưởng của ENSO đến thời gian xuất hiện đặc trưng dòng chảy cạn

Theo QTVHLHC sông Sê San, mùa cạn trên lưu vực sông Sê San bắt đầu vào tháng 12 và kết thúc vào tháng 6. Qua kết quả nghiên cứu cho thấy, hầu hết các đợt El Nino gây thâm hụt lượng mưa phổ biến từ 25 - 50% so với TBNN ở hầu hết các vùng, các đợt La Nina cũng làm cho lượng mưa bị thâm hụt ở các tỉnh Tây Nguyên.

El Nino luôn là tác nhân tạo ra mức độ cạn kiệt ở các sông suối. Thời gian cạn kiệt kéo dài, mực nước sông suối xuống thấp, thậm chí nhiều sông suối không còn dòng chảy dẫn đến tình trạng thiếu nước nghiêm trọng cho sinh hoạt, sản xuất nông nghiệp và phát điện. Rõ nét nhất là El Nino 1997 - 1998, mùa cạn 1997 - 1998 trên các sông ở Tây Nguyên kéo dài từ 7 - 10 tháng. Chẳng hạn, trên sông Đakbla tại Kon Tum kéo dài từ tháng 12/1997 - 8/1998. Xét quan hệ giữa chỉ số SSTA của những năm El Nino và La Nina với dòng chảy năm và dòng chảy mùa cạn trên các sông Krong Poko, Đăk Tơ Kan, Đakbla chỉ ra trong bảng 2 cho thấy:

Trong những năm El Nino:

- Các trạm đo dòng chảy trên hệ thống sông Sê san đều có hệ số tương quan lớn hơn 0;

- Đối với dòng chảy năm, xu thế và mức độ ảnh hưởng của El Nino khác nhau giữa các sông. Trên sông Đakbla hệ số tương quan giữa chỉ số SSTA với dòng chảy năm cao hơn các sông khác, tương quan của sông Đăk Tơ Kan là nhỏ nhất.

- Tương tự Q_n, xu thế và mức độ ảnh hưởng của El Nino đến dòng chảy mùa cạn khác nhau giữa các sông. Trên sông Đakbla hệ số tương quan giữa chỉ số SSTA với dòng chảy cạn cao hơn các sông khác đạt khoảng 0,3 - 0,5, sau đó đến sông Krong Poko đạt khoảng 0,3 - 0,4, tương quan của sông Đăk Tơ Kan là nhỏ nhất.

Trong những năm Lanina:

- Đối với dòng chảy năm, xu thế và mức độ ảnh hưởng của Lanina khác nhau giữa các sông. Hệ số tương quan giữa chỉ số SSTA với dòng chảy năm tại Konplong trên sông Đakbla và hồ Pleikrong trên sông Krong Poko cao hơn các sông khác đạt khoảng 0,5 - 0,6%, các vị trí còn lại hệ số tương quan đạt khoảng 0,25 - 0,45, riêng tại Kontum là nhỏ nhất.

- Tương tự Q_n, xu thế và mức độ ảnh hưởng của Lanina đến dòng chảy mùa cạn khác nhau giữa các sông. Trên sông Krong Poko hệ số tương quan giữa chỉ số SSTA với dòng chảy cạn đạt khoảng 0,15 - 0,4, sông Krong Ana đạt khoảng 0,05 - 0,17, sông Sê San quan hệ chặt nhất đạt 0,4 và sông Đăk Tơ Kan đạt 0,28.

Bảng 2. Hệ số tương quan giữa các đặc trưng dòng chảy với chỉ số SSTA trung bình năm trong các năm El Nino và La Nina

Trạm thủy văn	Sông	Tương quan hệ số SSTA với dòng chảy năm Elnino		Tương quan hệ số SSTA với dòng chảy năm Lanina	
		Năm	Mùa cạn	Năm	Mùa cạn
Kon tum	Đakbla	0.25	0.34	0.1	0.17
Konplong	Đakbla	0. 5	0.55	0.6	0.05
Đăk tô	Đăk Tơ Kan	0.1	0.08	0.33	0.28
Đăk môd	Krong Poko	0.16	0.4	0.27	0.15
Yali	Sê san	0.3	0.2	0.41	0.4
Pleikrong	Krong Poko	0.24	0.3	0.54	0.38

Trong những năm El Nino, phần lớn các trạm có dòng chảy năm và dòng chảy trung bình mùa cạn nhỏ hơn trung bình nhiều năm từ 5 - 15%, tại Yaly nhỏ hơn 28%, riêng tại Konplong và Đăk môd tương đương mức TBNN, năm Elnino mạnh như năm 2015 dòng chảy mùa cạn nhỏ hơn TBNN cao nhất tại PleiKrong lên đến 31% chủ yếu do lượng mưa giảm và lượng bốc hơi tăng. Trong những năm La Nina, dòng chảy năm các sông thường lớn hơn trung bình nhiều năm từ 3 - 16%, năm Lanina mạnh như năm 1999, tại PleiKrong dòng chảy mùa cạn lớn hơn cao nhất 41% so với TBNN (Bảng 3, Bảng 4).

Bảng 3. Lượng dòng chảy năm ảnh hưởng Elnino và Lanina trên lưu vực sông Sê San

Trạm	Q TB năm	Q TB năm năm Elnino	Q TB năm năm Lanina	% Q TB năm năm Elnino	% Q TB năm năm Lanina
Kon tum	96.1	87.7	109	-9	13
Konplong	44.5	44.5	48.4	0	9
Đăk môd	73.9	75.8	76.3	3	3
Yali	263	247	282	-6	7
Pleikrong	130	119	145	-8	12

Bảng 4. Lượng dòng chảy mùa cạn của năm ảnh hưởng Elnino và Lanina trên lưu vực sông Sê san

Trạm	Q TB mùa cạn	Q mùa cạn năm Elnino	Q TB mùa cạn năm Lanina	% Q TB mùa cạn năm Elnino	% Q TB mùa cạn năm Lanina
Kon tum	60.1	55.2	69.5	-8	16
Konplong	26.4	29.7	27.5	13	4
Đăk môd	39.2	37.1	42.8	-5	9
Yali	141	102	153	-28	9
Pleikrong	59.3	52.4	68.2	-12	15

Đối với dòng chảy mùa cạn, trong những năm El Nino và Lanina, lượng dòng chảy 1 tháng nhỏ nhất và 3 tháng liên tục nhỏ nhất ở hầu hết các trạm đều lớn hơn trị số trung bình nhiều năm cùng thời kỳ tương ứng khoảng 5 - 20%, riêng trạm Đăk Môd lượng dòng chảy 1 tháng nhỏ nhất và 3 tháng liên tục nhỏ nhất thấp hơn TBNN 5% (Bảng 5, bảng 6).

Bảng 5. Lượng dòng chảy 3 tháng liên tục nhỏ nhất của năm ảnh hưởng El Nino và La Nina trên lưu vực sông Sê san

Trạm	Q TB 3 tháng min	Q TB 3 tháng min năm Elnino	Q TB 3 tháng min năm Lanina	% Q TB 3 tháng min năm Elnino	% Q TB 3 tháng min năm Lanina
Kon tum	39.3	39.5	44.1	1	12
Konplong	21.3	24.9	22.3	17	5
Đăk môd	22.1	23.5	20.9	6	-5
Yali	293.0	317.6	310	8	6
Pleikrong	109	115	121	6	11

Bảng 6. Lượng dòng chảy tháng nhỏ nhất của năm ảnh hưởng Elnino và Lanina trên lưu vực sông Sê san

Trạm	Q TB 1 tháng min	Q TB 1 tháng min năm Elnino	Q TB 1 tháng min năm Lanina	% Q TB 1 tháng min năm Elnino	% Q TB 1 tháng min năm Lanina
Kon tum	35.1	34.2	41.2	-3	17
Konplong	19.2	21.7	20.9	13	9
Đăk môd	20.4	21.8	19.3	7	-5
Yali	80.8	90.6	84.9	12	5
Pleikrong	31.0	32.8	35.3	6	14

2.3 Cơ sở phương pháp hồi quy bội

Mô hình hồi qui bội dựa trên số liệu quan trắc trong quá khứ thiết lập một phương trình tuyến tính, mô tả mối quan hệ giữa yếu tố dự báo với các nhân tố ảnh hưởng. Mô hình hồi qui bội có dạng tổng quát sau:

$$Y_i = a_0 + \sum_{j=1}^{m} a_j X_{i,j} + \varepsilon_i \qquad (2)$$

Trong đó: a_j (j = 1- m) là các tham số chưa biết; ε_i là sai số ngẫu nhiên. Các tham số a_j có thể xác định bằng phương pháp tối thiểu hàm mục tiêu S(a):

$$S(a) = \frac{1}{n}\sum_{i=1}^{n}\left[Y_i - \left(a_0 + \sum_{i=1}^{n} a_j X_{i,j}\right)\right]^2 \qquad (3)$$

Lấy đạo hàm thành phần hàm S(a) theo từng tham số aj và cho bằng 0 ta được hệ phương trình có m+1 ẩn:

$$\frac{dS}{da_j} = 0 \qquad (4)$$

$$a_0 + a_1 \sum_{i=1}^{n} X_{1,i} + a_2 \sum_{i=1}^{n} X_{2,i} + \ldots + a_m \sum_{i=1}^{n} X_{m,1} = \sum_{i=1}^{n} Y_i \qquad (5)$$

$$a_0 \sum_{i=1}^{n} X_{1,i} + a_1 \sum_{i=1}^{n} X^2_{1,i} + a_2 \sum_{i=1}^{n} X_{2,i} X_{1,i} + \ldots + a_m \sum_{i=1}^{n} X_{m,1} X_{1,i} = \sum_{i=1}^{n} Y_i X_{1,i} \qquad (6)$$

$$a_0 \sum_{i=1}^{n} X_{m,i} + a_1 \sum_{i=1}^{n} X_{1,i} X_{1,i} + a_2 \sum_{i=1}^{n} X_{m,i} X_{1,i} + \ldots + a_m \sum_{i=1}^{n} X_{2m,i} = \sum_{i=1}^{n} Y_1 X_{1,i} \qquad (7)$$

X_j=0 =1 đồng thời ký hiệu:

$$Y'_j = \sum_{i=1}^{n}\left(Y_i X_{j,i}\right);\quad X'_{j,k} = \sum_{i=1}^{n}\left(X_{j,i} X_{k,i}\right) \tag{8}$$

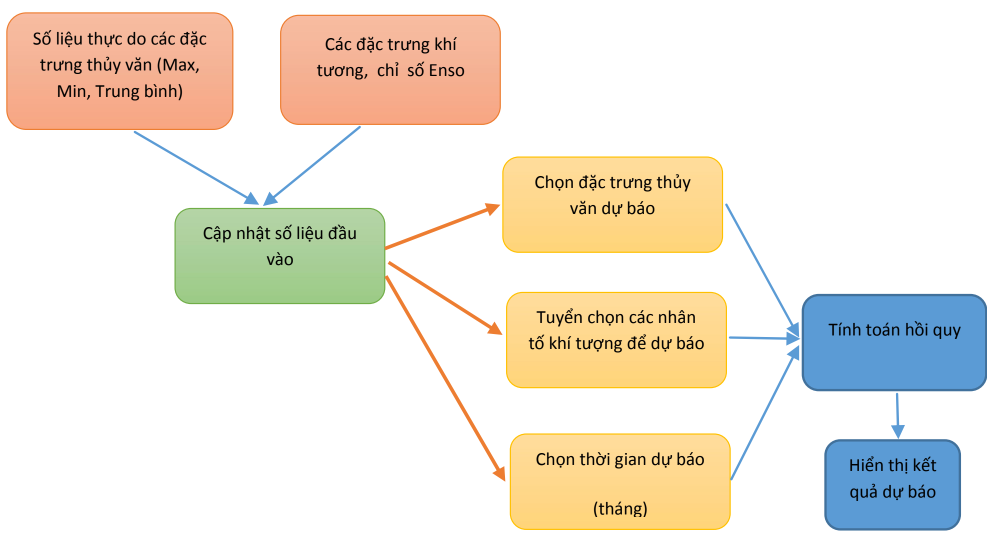
Và viết hệ phương trình trên dưới dạng ma trận ta được:

$$C = \begin{vmatrix} 1 \quad X'_{0,1} X'_{0,2} \ldots X'_{0,m} \\ X'_{0,1} X'_{1,1} X'_{1,2} \ldots X'_{1,m} \\ \ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots \\ X'_{m,1} X'_{m,1} X'_{m,2} \ldots X'_{m,m} \end{vmatrix};\quad B = \begin{vmatrix} Y'0 \\ Y'_1 \\ . \\ . \\ . \\ Y'_m \end{vmatrix};\quad A = \begin{vmatrix} a_0 \\ a_1 \\ . \\ . \\ . \\ a_m \end{vmatrix} \tag{9}$$

hay C x A = B; và véc tơ tham số hồi qui A sẽ bằng tích của ma trận nghịch đảo C^{-1} với véctơ B. ->A = C^{-1} x B.

2.4 Cơ sở phương pháp xây dựng phần mềm phục vụ dự báo

Phần mềm được xây dựng bằng ngôn ngữ lập trình C# phục vụ dự báo hạn dài lưu vực sông Sê San. Khái quát hóa sơ đồ khối quá trình xây dựng phần mềm được trình bày trong hình 3.



Hình 3. Sơ đồ khối xây dựng phần mềm dự báo

3. Phân tích kết quả và đánh giá

3.1 Nghiên cứu xây dựng phương án

Để tuyển chọn bộ nhân tố tối ưu, phương pháp phân tích tương quan với mô hình hồi qui bội đã được sử dụng nhằm tìm được nhóm các nhân tố mà tổ hợp của chúng theo nghĩa hồi quy tuyến tính có chứa đựng nhiều thông tin nhất từ các Cơ sở số liệu 1. Các bộ nhân tố tối ưu này phải đảm bảo đáp ứng đầy đủ các chỉ tiêu thống kê về phương sai, trung bình, chỉ tiêu thông tin... và mục tiêu cuối cùng là có mức đảm bảo dự báo lớn nhất có thể. Các phương trình dự báo tối ưu được xây dựng riêng biệt cho dòng chảy tháng.

Trên cơ sở tuyển chọn các nhân tố trong cơ sở số liệu, sử dụng phần mềm hồi quy đa biến đã xây dựng được các phương trình tối ưu dự báo dòng chảy tháng của 4 vị trí thủy văn kể trên. Bảng 7 sẽ trình bày tóm tắt về phương trình tối ưu và các chỉ tiêu chất lượng thống kê. Đánh giá kết quả tính toán theo các phương trình tại 8 vị trí cho kết quả khá tốt với chỉ tiêu đánh giá s/□ dao động từ 0,7 - 0,89; theo P% dao động từ 75 - 95% (Bảng 8).

Bảng 7. Các phương trình tính toán lưu lượng trung bình tháng đến các trạm thủy văn và hồ chứa trên lưu vực sông Sesan.

Thứ tự	Tên trạm/hồ	Phương trình tính toán
1	Đắc Tô	$Q_{tb}(t+1) = 42,15 + 0,35Qc(t) - 0,99T_{max}(t) - 0.004S(t) - 0,003E(t) - 0,041H_{min}(t) + 0,011X(t) - 0,135Nx(t) - 0,578SSTC(t) + 0,15H500A(t)$
2	Đắc Mod	$Q_{tb}(t+1) = 42,15 + 0,35Qc(t) - 0,99T_{max}(t) - 0.004S(t) - 0,003E(t) - 0,041H_{min}(t) + 0,011X(t) - 0,135Nx(t) - 0,578SSTC(t) + 0,15H500A(t)$
3	Kon Tum	$Q_{tb}(t+1) = 563,9 + 0,155Qc(t) - 9,72T_{max}(t) - 0,19S(t) + 0,025E(t) - 0,165H_{tb}(t) + 0,18X(t) - 3,79Nx(t) - 11,09SSTC(t) + 1,88H500C(t)$
4	Kon Plông	$Q_{tb}(t+1) = 387,6 + 0,21\ Qc(t) - 5,36\ T_{max}(t) - 0.067\ S(t) + 0,064\ E(t) - 1,87\ H_{min}(t) + 0,019\ X(t) - 5,8\ SSTC(t) + 1,21\ H500C(t)$
5	Plây krông	$Q_{tb}(t+1) = 988 - 14,6T_{max}(t) - 0.289S(t) + 0,069E(t) - 2,9Htb(t) + 0,44X(t) - 3,2Nx(t) - 17,99SSTC(t) + 3,47H500A(t)$
6	Yaly	$Q_{tb}(t+1) = 2703 - 30,5\ T_{max}(t) - 0,16\ S(t) - 0,82\ E(t) - 1,64\ H_{tb}(t) + 0,36\ X(t) + 2,1\ Nx(t) - 24,6\ SSTC(t) - 7,69\ H500E(t)$
7	Sesan 4	$Q_{tb}(t+1) = 1744 - 35,2\ T_{max}(t) - 0,15\ S(t) - 0,46\ E(t) + 2,15\ H_{tb}(t) + 0,79\ X(t) - 7,03\ Nx(t) - 31,8\ SSTC(t) + 4,65\ H500E(t)$
8	Sesan 4A	$Q_{tb}(t+1) = 2341 - 30,9\ T_{max}(t) - 0,12\ S(t) - 1,48\ E(t) + 0,5\ H_{tb}(t) + 0,62\ X(t) - 4,42\ Nx(t) - 31,7\ SSTC(t) - 1,9\ H500E(t)$

Trong đó: $Q_{tb}(t+1)$: Lưu lượng trung bình tháng thứ t +1; $Q_c(t)$: Lưu lượng ngày cuối tháng thứ t; X(t): Tổng lượng mưa tháng thứ t; $N_x(t)$: Số ngày mưa tháng thứ t; $H_{min}(t)$: Độ ẩm thấp nhất tháng thứ t; E(t): Tổng lượng bốc hơi tháng thứ t; S(t): Tổng số giờ nắng tháng thứ t; $T_{max}(t)$: Nhiệt độ không khí lớn nhất tháng thứ t; SSTC(t): Chuẩn sai nhiệt độ nước biển khu El-Nino C tháng thứ t; H500A(t): Vị trí áp cao Thái Bình Dương trạm số 58847 tháng thứ t, H500E(t): Vị trí áp cao Thái Bình Dương trạm số 98327 tháng thứ t (tại thời điểm làm dự báo).

Phân tích tương quan của các biến độc lập với nhau nhìn chung là rất lớn, điều này chứng minh mối quan hệ giữa các biến độc lập này không có mối tương quan và nó càng khẳng định tính "độc lập" tốt giữa các biến độc lập. Do đó đối với mỗi phương trình cho từng vị trí trạm thì mối quan hệ giữa yếu tố khí tượng mà chặt chẽ thì các yếu tố sẽ không được chọn vào phương trình hồi quy bội. Ví dụ đối với trạm Kon Tum: Mối quan hệ giữa yếu tố Tb với Ti và Tb với Hb có hệ số tương quan Pearson cao, cho thấy các yếu tố này có tính tương quan rất mạnh mẽ giữa các biến, như vậy yếu tố nhiệt độ trung bình, nhiệt độ tối thấp và độ ẩm trung bình là 3 yếu tố không được chọn chung khi đưa yếu tố khí tượng trạm Kon Tum vào trong phương trình hồi quy bội, trong khi đó mối quan hệ giữa các nhân tố khác có chỉ số tương quan Pearson thấp nhưng lại có mối tương quan chặt chẽ hơn đối với dòng chảy trung bình tháng trạm Kon Tum, nên phương trình tính toán theo phương pháp hồi quy bội chỉ nên chọn 9 nhân tố là đủ. Tương tự đối với các điểm trạm và hồ khác thì ta sẽ có bộ nhân tố được chọn lọc tối ưu nhất tránh hiện tượng đa cộng tuyến.

Bảng 8. Đánh giá kết quả tính toán dựa trên phương trình đã xây dựng

Thứ tự	Tên trạm/hồ	Chuỗi số liệu	Chỉ tiêu đánh giá		
			Sai số cho phép (m^3/s)	s/б	P (%)
1	Đắc Tô	264	6	0,8	95
2	Đắc Mod	264	43	0,84	82
3	Kon Plông	264	246	0,89	86
4	Kon Tum	264	391	0,79	87
5	Plây krông	264	78	0,82	90
6	Yaly	250	142	0,77	84
7	Sesan 4	247	202	0,70	75
8	Sesan 4A	247	179	0,76	78

Bảng 9. Tiêu chí đánh giá chất lượng cho các chỉ số

Xếp loại	RSR	NSE	PBIAS (%)
Rất tốt	0 ≤ RSR ≤ 0.5	0.75 < NSE ≤ 1	PBIAS < ± 10
Tốt	0.5 ≤ RSR ≤ 0.6	0.65 < NSE ≤ 0.75	±10 ≤ PBIAS < ±15
Đạt yêu cầu	0.6 ≤ RSR ≤ 0.7	0.5 < NSE ≤ 0.65	±15 ≤ PBIAS < ±25
Không đạt	RSR > 0.7	NSE ≤ 0.5	PBIAS ≥ ±25

Bảng 10. Đánh giá chất lượng hiệu chỉnh và kiểm định mô hình

Quá trình	Năm	Chỉ tiêu	Đăk Mốt	Đăk Tô	Kon Plong	Kon Tum
Hiệu chỉnh	1990-2010	NSE	0,84	0,78	0,71	0,76
		RSR	0,55	0,62	0,70	0,65
		PBIAS	-3,67	-3,18	2,55	-0,21
Kiểm định	2011-2014	NSE	0,84	0,87	0,77	0,81
		RSR	0,53	0,48	0,64	0,58
		PBIAS	12,37	12,3	-1,57	0,84

Bảng 11. Đánh giá chất lượng hiệu chỉnh, kiểm định lưu lượng đến các hồ

Quá trình	Năm	Chỉ tiêu	Plei Krong	Ialy	Sê San 4	Sê San 4A
Hiệu chỉnh	1994-2010	NSE	0,8	0,75	0,69	0,77
		RSR	0,59	0,66	0,72	0,63
		PBIAS	-0,77	0,34	-1,2	-0,37
Kiểm định	2011-2014	NSE	0,88	0,83	0,72	0,7
		RSR	0,46	0,55	0,69	0,71
		PBIAS	4,3	1,08	7,0	1,18

3.2 Phân tích và đánh giá kết quả

Để đánh giá chất lượng của phương án dự báo, nhóm nghiên cứu đã vận hành phương án với số liệu độc lập bốn năm từ năm 2011 - 2014 để đánh giá chất lượng phương trình dự báo lưu lượng trung bình tháng tại các trạm thủy văn, các hồ chứa. Kết quả tính toán và thực đo quá trình lưu lượng trung bình tháng tại các trạm và lưu lượng đến các hồ trên lưu vực sông Sê San được thể hiện trên hình 4. Chỉ số NSE *(Nash-Sutcliffe efficiency)*, PBIAS *(Percent bias)* và RSR *(RMSE-observations standard deviation ratio)*, được sử dụng để so sánh kết quả tính toán và thực đo (Bảng 9). Căn cứ vào chỉ tiêu đánh giá chất lượng cho các chỉ số (Bảng 8, Bảng 10 và Bảng 11) cho thấy kết quả tính toán tại hầu hết các vị trí đều đạt yêu cầu.

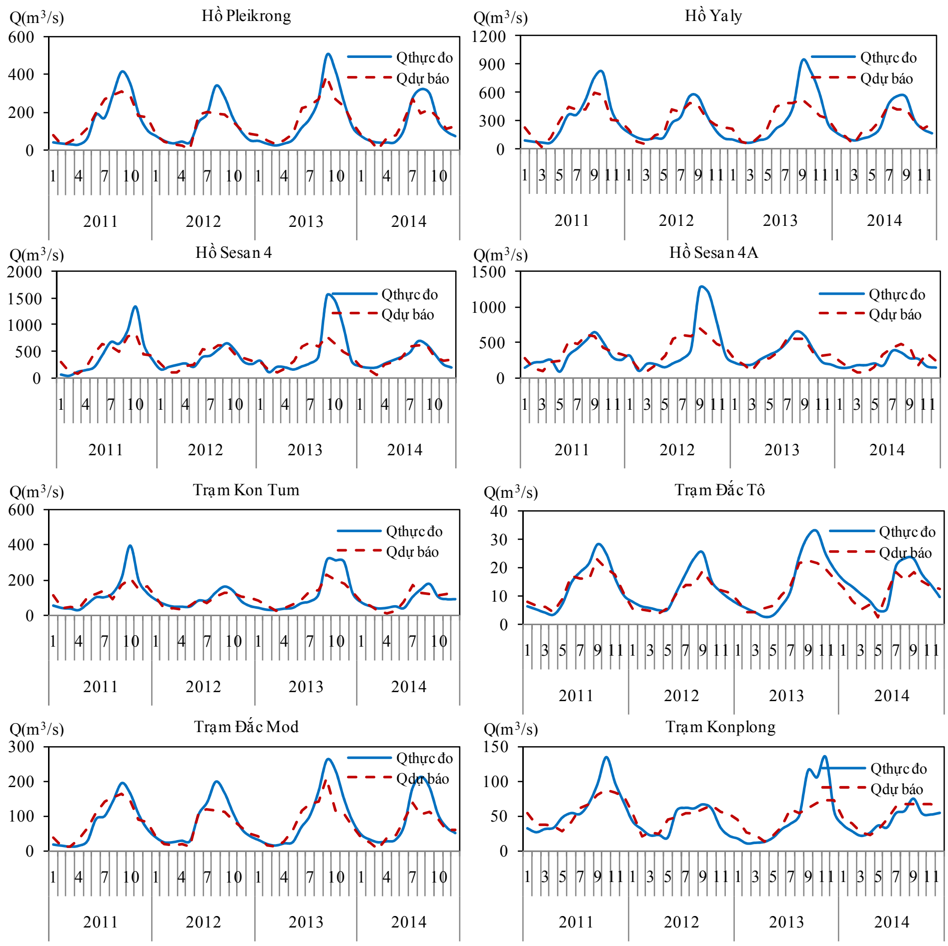
Dựa trên các đường quá trình, ta có thể rút ra một số nhận xét như sau đối với phương án dự báo lưu lượng trung bình đến các trạm theo phương pháp hồi quy bội:

- Về mặt xu thế, phương án có khả năng dự báo tương đối tốt về mặt xu thế;
- Về mặt trị số dự báo, phương án cho kết quả dự báo nhìn chung tốt đối với thời kỳ mùa cạn, thời kỳ mùa lũ trị số dự báo cho các trạm ở một số tháng thiên thấp so với giá trị thực tế.

Với phương án dự báo lưu lượng trung bình đến các hồ chứa theo phương pháp hồi quy bội:

- Về mặt xu thế, phương án có khả năng dự báo tương đối tốt về mặt xu thế;
- Về mặt trị số dự báo, phương án cho kết quả dự báo cho kết quả khá tốt. Trị số dự báo lưu lượng trung bình tháng cho bốn vị trí hồ chứa ở một số tháng mùa lũ cũng thiên thấp so với giá trị thực tế.

Phương án được xây dựng phục vụ dự báo trong thời kỳ mùa cạn nên trong các biểu đồ đường quá trình lưu lượng có sự sai khác trong thời kỳ mùa lũ là không thể tránh khỏi. Phương án phục vụ dự báo lưu lượng trung bình tháng tại vị trí các trạm thủy văn và lưu lượng đến các hồ hiện đã được xây dựng và đang được áp dụng phục vụ dự báo nghiệp vụ tại Trung tâm Dự báo Khí tượng thủy văn Trung ương.



Hình 4. Quá trình dự báo lưu lượng trung bình tháng tại các trạm và đến hồ trên lưu vực sông Sê San

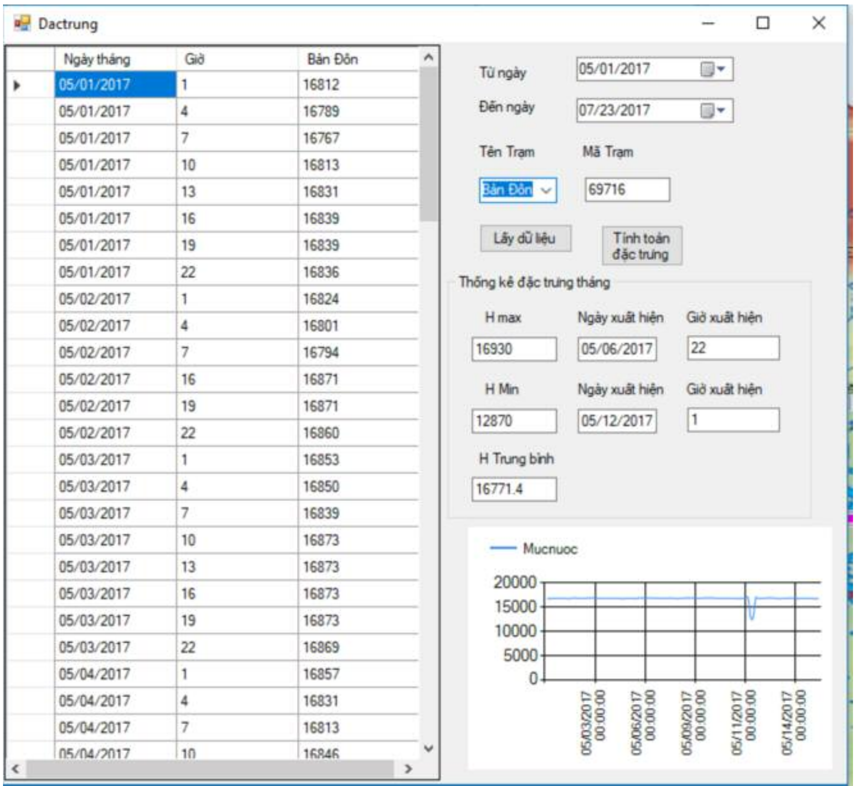
3.3 Xây dựng phần mềm phục vụ dự báo

Trên cơ sở bài toán đưa ra nhóm nghiên cứu lựa chọn ngôn ngữ lập trình C# để xây dựng phần mềm dự báo hạn dài lưu vực sông Sê San bởi các ưu điểm sau: (1) ngôn ngữ đơn giản; (2) ngôn ngữ hiện đại; (3) ngôn ngữ hướng đối tượng; (4) ngôn ngữ mạnh mẽ và mềm dẻo; (5) ngôn ngữ ít từ khóa; (6) ngôn ngữ hướng module; (7) ngôn ngữ phổ biến.

Phần mềm dự báo hạn dài lưu vực sông Sê San bao gồm nhiều module trong đó có module dự báo theo phương pháp hồi quy bội. Giao diện chính của phần mềm được thể hiện trên hình 5.

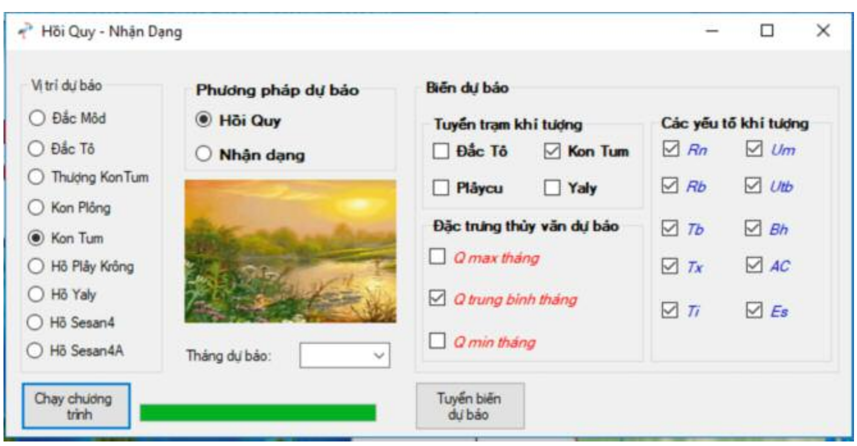
Hình 5. Giao diện chính của phần mềm dự báo hạn dài sông Sesan

Modul cập nhật số liệu: đây là module cho phép cập nhật số liệu thực đo theo thời gian thực, tính toán các đặc trưng thủy văn, cập nhật các nhân tố khí tượng đến tháng hiện tại để dự báo cho tháng tiếp theo (Hình 6).



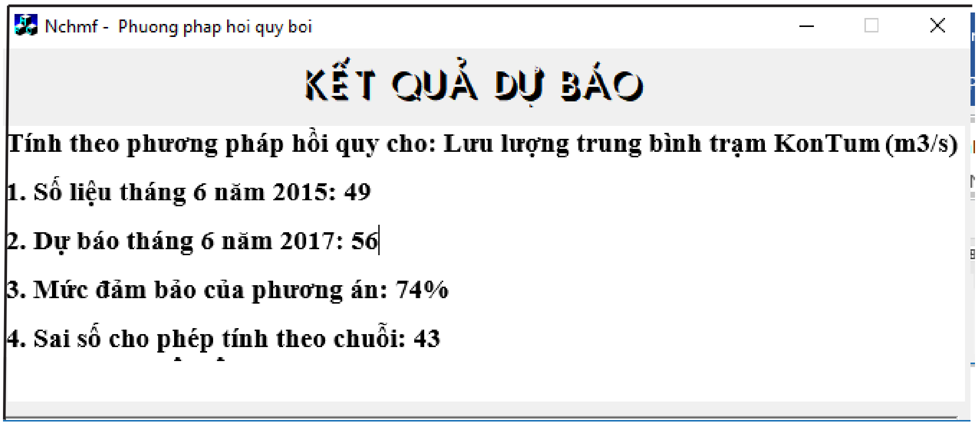
Hình 6. Giao diện module cập nhật số liệu theo thời gian thực

Module tính toán bằng phương pháp hồi quy: Được xây dựng với nhiều *radio button, check-box, combo box, button* dùng để lựa chọn các nhân tố khí tượng, đặc trưng thủy văn, phương pháp tính toán (Hình 7).



Hình 7. Giao diện module tính toán theo phương pháp hồi quy

Module tính toán theo phương pháp hồi quy được xây dựng tại 8 vị trí dự báo đã được tính toán, kiểm định như trên. Trong module cho phép lựa chọn các yếu tố khí tượng thủy văn dùng để tính toán. Kết quả tính toán theo phương pháp hồi quy tại trạm Kon Tum dự báo tháng 6/ 2017, với các nhân tố khí tượng, yếu tố thủy văn là lưu lượng trung bình tháng 6 cho kết quả khá tốt với mức đảm bảo phương án 74% (Hình 8).



Hình 8. Kết quả tính toán theo phương pháp hồi quy tại trạm Kon Tum

3.4 Đánh giá phần mềm

Phần mềm được xây dựng bằng ngôn ngữ lập trình C# nên có nhiều ưu điểm: (1) Cài đặt dễ dàng; (2) Giao diện thân thiện, dễ sử dụng; (3) Chạy ổn đỉnh trên PC với các phiên bản hệ điều hành khác nhau; (4) Kết quả tính toán rất nhanh.

4. Kết luận kiến nghị

Trong những năm vừa qua do ảnh hưởng của biến đổi khí hậu nên chế độ dòng thủy văn trên các lưu vực sông nói chung và lưu vực sông Sê San nói riêng có nhiều biến đổi bất thường. Vì vậy, dự báo hạn dài chế độ thủy văn trên lưu vực sông hết sức quan trọng. Việc xây dựng được phương án dự báo tại tác trạm thủy văn và lưu lượng đến các hồ chứa trên sông Sên San là rất cần thiết. Trong nghiên cứu này, nhóm nghiên cứu muốn giới thiệu một phần mềm dự báo hạn dài trong thời kỳ mùa cạn kết hợp với thời kỳ quan trắc là hết sức cần thiết và cấp bách, đóng vai trò quan trọng trong việc đưa ra cảnh báo, dự báo tình hình thiếu hụt dòng chảy trên lưu vực sông Sê San. Kết quả từ việc xây dựng mô hình hồi quy đa biến được tích hợp xây dựng thành phần mềm dự báo hạn dài cho thấy:

- Chế độ thủy văn tháng tại các trạm, đến hồ vẫn tuân theo quy luật chung với mức đảm bảo từ 70% trở lên là khá cao.

- Quá trình biến đổi của các đặc trưng thủy văn theo tháng diễn ra mạnh (có thể là do điều tiết của hồ chứa) nhưng quy luật theo mùa vẫn được đảm bảo.

- Xây dựng phần mềm dự báo thủy văn là rất cần thiết rút ngắn được thời gian tính toán, có độ ổn định cao.

Lời cảm ơn: *Bài báo được sự hỗ trợ của đề tài nghiên cứu khoa học và phát triển công nghệ "Nghiên cứu xây dựng công nghệ dự báo thủy văn hạn vừa hạn dài mùa cạn phục vụ quy trình vận hành liên hồ chứa cho các sông chính ở khu vực Tây Nguyên", Mã số TNMT.2016.05.01. Kết quả đạt được trong bài báo có sự tham gia tích cực của các thành viên trong nhóm nghiên cứu của đề tài.*

Tài liệu tham khảo

1. Chiew, F.H., and T.A. McMahon (2002), *Modelling the impacts of climate change on Australian streamflow*, Hydrological Processes, vol. 16, 1235-1245.

2. Chowdhury, S., and A. Sharma (2009), *Long-Range Nino-3.4 Predictions Using Pairwise Dynamic Combinations of Multiple Models*, American Meteorological Society, vol. 22, 793-805.

3. Garen, D.C. (1992), *Improved techniques regression-based streamflow volume forecasting*, J. Water Resour. Plann. Manage, Vol. 118, 654 - 670, https://doi.org/10.1061/(ASCE)0733-9496(1992)118:6(654).

4. Maurer, E.P., and D.P. Lettenmaier (2003), *Predictability of seasonal runoff in the Mississippi river basin*, J. Geophys. Res., vol. 108, 8607, doi:10.1029/2002JD002555.

5. Phùng Tiến Dũng, Đoàn Quang Trí, Đào Ngọc Hiếu (2016), *Nghiên cứu ứng dụng mô hình thủy*

văn trong mô phỏng dự báo quá trình dòng chảy cạn cho các hồ chứa trên lưu vực sông Sê San, Tạp chí KTTV, số 670, tr. 13-19.

6. Risley J., M. Gannett, J. Lea, and E. Roehl (2005), *An analysis of statistical methods for seasonal flow forecasting in the Upper Klamath river basin of Oregon and California*, USGS Scientific Investigations Report, 2005-5177.

7. Wang, Q.J., D.E. Robertson, and F.H.S. Chiew (2009), *A Bayesian joint probability modelling approach for seasonal forecasting of streamflows at multiple sites*, Water Resources Research, vol. 45, W05407, doi:10.1029/2008WR007355.

8. Wang, E.Y. Zhang, J. Luo, F.H.S. Chiew, Q.J. Wang (2011), *Monthly and seasonal streamflow forecasts using rainfall - runoff modelling and historical weather data*, Water Resources Research, in press.

9. Wang, Q.J., and D. Robertson (2011), *Multisite probabilistic forecasting of seasonal flows for streams with zero value occurrences*, Water Resources Research., vol. 47, W02546, doi:10.1029/2010WR009333.

APPLICATION OF STATISTIC MODEL FOR MONTHLY FLOW FORECASTING TO SERVE THE INTER-RESERVOIR TO OPERATE PROCEDURES IN DRY SEASON IN SE SAN RIVER

Tran Hong Thai[1], Phung Tien Dung[2], Doan Van Hai[2], Doan Quang Tri[3], Duong Quoc Hung[1]

[1]National Hydrometeorology Service

[2]National Center for Hydrometeorological Forrecasting

[3]Scientific and Technical Hydrometeorological Journal

Abstract: *The responsibility of the National Center for Hydrometeorological Forecasting (NCHMF), prescribed in the Se San River Inter-reservoir Operation Procedures (IROP), is to make a long-term forecasting (for months) in dry season for controlled hydrological stations in the Se San River basin to help authorities for the purposes of Inter-reservoir Operation Procedures. The study was applied the multivariate statistical regression model to develop a software for predicting monthly flow characteristics in Se San River. The data used to calibrate the model from 1990 to 2010 and the results valid from 2011 to 2014 were sufficiently qualified. From the results, the research group did research to set up an application which could support forecasters during the forecasting process to regulate the reservoir and meet the increasing society's demands along with ceaseless development of hydropower reservoir system in Sesan River. This can be an effective tool to serve the forecast work in the future.*

Keywords: *Monthly flow characteristic, Se San River, Multiple regression model, Long-term forecasting technology.*