BỘ GIÁO DỤC VÀ ĐÀO TẠO

ĐẠI HỌC ĐÀ NẴNG

NGUYỄN VĂN HƯNG

NGHIÊN CỨU MÔ HÌNH HỒI QUY GAMMA BẬC 1 [GAR(1)] ỨNG DỤNG TRONG LĨNH VỰC THỦY VĂN

Chuyên ngành: Khoa học máy tính
Mã số: 62.48.01.01

TÓM TẮT LUẬN ÁN TIẾN SĨ KỸ THUẬT

Đà Nẵng - Năm 2016

Công trình được hoàn thành tại

ĐẠI HỌC ĐÀ NẴNG

Người hướng dẫn khoa học:

1. PGS.TSKH. Trần Quốc Chiến
2. GS.TS. Huỳnh Ngọc Phiên

Phản biện 1: GS.TS. Nguyễn Thanh Thủy,
Trường Đại học Công nghệ Hà Nội;

Phản biện 2: PGS.TS. Nguyễn Mậu Hân,
Trường Đại học Khoa học Huế;

Phản biện 3: TS. Phạm Minh Tuấn,
Trường Đại học Bách khoa Đà Nẵng.

Luận án đã được bảo vệ tại Hội đồng chấm Luận án Tiến sĩ Kỹ thuật cấp Đại học Đà Nẵng vào ngày 24 tháng 6 năm 2016 tại Đại học Đà Nẵng.

Có thể tìm hiểu Luận án tại:

- Thư viện Quốc gia;
- Trung tâm Thông tin – Học liệu, Đại học Đà Nẵng.

GIỚI THIỆU

Ngày nay, ngành khoa học máy tính có vai trò rất quan trọng trong sự phát triển của toàn cầu, đã tác động sâu sắc đến hầu hết các ngành, lĩnh vực kỹ thuật, kinh tế xã hội. Trên thế giới đã có nhiều công trình trong lĩnh vực khoa học máy tính nghiên cứu về Tin viễn thông, Tin y sinh học đã và đang mang lại hiệu quả to lớn cho đời sống con người, trong khi đó, các công trình nghiên cứu về Tin thủy văn vẫn còn nhiều hạn chế. Đề tài này có mục đích góp phần cho sự phát triển lĩnh vực Tin thủy văn hiện nay và trong tương lai. Để đạt được mục đích nêu trên, mục tiêu nghiên cứu của Luận án là:

- Nghiên cứu mô hình GAR(1), tổng quan các công trình liên quan về mô hình GAR(1), phương pháp mô phỏng ngẫu nhiên, các phương pháp sinh biến ngẫu nhiên, các mô hình biểu thị lưu lượng dòng chảy và bài toán ước lượng dung tích hồ chứa;
- Nghiên cứu các thuật toán sinh biến ngẫu nhiên GAR(1) bao gồm: đánh giá các thuật toán sinh biến ngẫu nhiên có phân phối đều, phân phối mũ, phân phối chuẩn, phân phối Poisson và phân phối gamma;
- Nghiên cứu các mô hình biểu thị lưu lượng dòng chảy hàng tháng, hàng năm với quá trình ngẫu nhiên GAR(1);
- Nghiên cứu bài toán tính dung lượng trung bình của hồ chứa có dung tích vô hạn với chuỗi lưu lượng dòng chảy vào hồ chứa như quá trình ngẫu nhiên GAR(1).

CHƯƠNG 1
CÁC VẤN ĐỀ CHUNG

Để đáp ứng mục tiêu nghiên cứu của đề tài: “***Nghiên cứu mô hình hồi quy Gamma bậc 1 [GAR(1)] ứng dụng trong lãnh vực thuỷ văn***”, Tác giả nghiên cứu các tài liệu, công trình đã được công bố trong và ngoài nước có liên quan đến những vấn đề sau:

- Về lý luận: Các nghiên cứu cơ bản về lý thuyết xác suất, các kết quả nghiên cứu về các thuật toán sinh biến ngẫu nhiên, các phương pháp, mô hình và thuật toán dùng để mô phỏng lưu lượng dòng chảy hàng tháng, hàng năm và các nghiên cứu về hồ chứa.

- Về thực tiễn: Các kết quả công bố liên quan đến việc thực nghiệm, mô phỏng lưu lượng dòng chảy tại các trạm đo thuỷ văn và dung tích hồ chứa.

1.1. Một số vấn đề cơ bản của lý thuyết xác suất

Trong phần này trình bày các nội dung cơ bản về lý thuyết xác suất bao gồm các khái niệm về đại lượng ngẫu nhiên, luật phân phối tích phân, hàm mật độ xác suất của đại lượng ngẫu nhiên và các đặc trưng số cơ bản của đại lượng ngẫu nhiên: kỳ vọng, phương sai, hệ số lệch và độ nhọn làm cơ sở cho các nghiên cứu ở các nội dung kế tiếp.

1.2. Phân phối Gamma

1.2.1. Hàm mật độ xác suất của phân phối gamma

Một biến ngẫu nhiên liên tục *X* được gọi là có phân phối gamma 3 tham số nếu hàm mật độ xác suất có dạng:

$$f(x) = \frac{(x-c)^{a-1} e^{-(x-c)/b}}{b^a \Gamma(a)}, \qquad (1.1)$$

trong đó $a > 0,\ b > 0,\ c \geq 0,\ x \geq c$; $a,\ b,\ c$ tương ứng là các tham số hình dạng, tỉ lệ và vị trí.

Hàm $\Gamma(a)$ được xác định bởi

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt, \quad a > 0$$

khi $c = 0$ ta có phân phối gamma 2 tham số, khi $c = 0$ và $b = 1$ ta có phân phối gamma 1 tham số. Bằng phương pháp đổi biến số, phân phối gamma với 2 hoặc 3 tham số có thể biến đổi về phân phối gamma 1 tham số: với phân phối gamma 3 tham số, đặt: $y = (x-c)/b$ hoặc $x = c + by$, với phân phối gamma 2 tham số, đặt: $y = x/b$ hoặc $x = by$. Với cách đổi biến như trên thì biến ngẫu nhiên *y* có phân phối gamma 1 tham số.

1.2.2. Các đặc trưng số của phân phối gamma

Các đặc trưng số cơ bản của phân phối gamma 1 tham số được tính như sau:

Kỳ vọng: $$E(X) = a, \tag{1.2}$$

Phương sai: $$Var(X) = a, \tag{1.3}$$

Hệ số lệch: $$g = \frac{2}{\sqrt{a}}. \tag{1.4}$$

1.3. Mô hình hồi quy gamma bậc 1 [GAR(1)]

1.3.1. Mô hình GAR(1)

Lawrance và Lewis (1981) đề xuất mô hình GAR(1) như sau:

$$X_i = \Phi X_{i-1} + e_i, \tag{1.5}$$

trong đó: X_i là biến ngẫu nhiên biểu diễn quá trình phụ thuộc ở thời điểm i; Φ là hệ số hồi quy; e_i là biến ngẫu nhiên độc lập cần được xác định; X_i có phân phối gamma 3 tham số và có hàm mật độ xác suất như ở phương trình (1.1). Quá trình được xác định bởi phương trình (1.5) được gọi là mô hình GAR(1), để mô phỏng quá trình này thì các tham số của mô hình phải được xác định và e_i được sinh theo các lược đồ thích hợp và có sự kết hợp với các thuật toán sinh biến ngẫu nhiên có phân phối đều, phân phối mũ và phân phối Poisson.

1.3.2. Ước lượng các tham số của mô hình GAR(1)

Bằng phương pháp moment, Fernandez và Salas (1990) đề xuất lược đồ điều chỉnh độ lệch để ước lượng các tham số của mô hình GAR(1). Quá trình ngẫu nhiên tuyến tính dừng GAR(1) ở phương trình (1.5) có 4 tham số là a, b, c và Φ. Sử dụng phương pháp moment, các tham số này và các moment của biến ngẫu nhiên X_i có mối liên hệ sau:

$$M = c + ab, \tag{1.6}$$

$$S^2 = ab^2, \tag{1.7}$$

$$G = 2/\sqrt{a}, \tag{1.8}$$

$$R = \Phi. \tag{1.9}$$

Trong đó M, S^2, G, R là trung bình mẫu, phương sai, độ lệch và hệ số tương quan bậc 1. Các tham số đặc trưng này có thể được ước lượng dựa trên mẫu thống kê $\{X_1, X_2, ..., X_N\}$. bằng cách tính:

$$m = \frac{1}{N}\sum_{i=1}^{N} X_i, \tag{1.10}$$

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N} (X_i - m)^2, \tag{1.11}$$

$$g = \frac{N}{(N-1)(N-2)s^3}\sum_{i=1}^{N} (X_i - m)^3, \tag{1.12}$$

$$r = \frac{1}{(N-1)s^2}\sum_{i=1}^{N-1} (X_i - m)(X_{i+1} - m), \tag{1.13}$$

trong đó *m*, *s*, g và *r* là ước lượng của M, S, G và R tương ứng, và *N* là kích thước mẫu thống kê. Khi các biến ngẫu nhiên là phụ thuộc và không chuẩn, các ước lượng này thường bị lệch vì vậy cần phải điều chỉnh độ lệch và sau khi điều chỉnh độ lệch ta thu được các ước lượng không lệch của M,R, S và G và các công thức (1.6) - (1.9) được sử dụng để ước lượng tập các tham số của mô hình: a, b, c và Φ tương ứng.

1.4. Sinh biến ngẫu nhiên theo mô hình GAR(1)

Sinh biến ngẫu nhiên theo mô hình GAR(1) cần phải kết hợp các thuật toán sinh các biến ngẫu nhiên có phân phối đều đơn vị, phân phối mũ, phân phối chuẩn, phân phối Poisson và phân phối gamma. Có nhiều công trình nghiên cứu đề xuất các thuật toán để sinh biến ngẫu nhiên có phân phối gamma và được phân chia ra hai trường hợp: (1) *Trường hợp tham số hình dạng a*≤1, và, (2) *Trường hợp tham số hình dạng* a>1. Trong những năm gần đây có một số tác giả nghiên cứu đề xuất các thuật toán để sinh biến ngẫu nhiên gamma với tham số *a* là bất kỳ như trong công trình của Marsaglia và Tsang (2000), và gần đây Hong LiangJie (2012) đánh giá thuật toán do Marsaglia và Tsang (2000) đề xuất là một trong các thuật toán dễ cài đặt, có tốc độ nhanh nhất hiện nay và được cài đặt trong thư viện GSL và phần mềm Matlab "gamrnd".

1.5. Bài toán mô phỏng lưu lượng dòng chảy

Bài toán mô phỏng lưu lượng dòng chảy đặt ra vấn đề là trên cơ sở chuỗi lưu lượng lịch sử hàng năm hoặc hàng tháng quan trắc được tại các trạm đo thuỷ văn, áp dụng các phương pháp, mô hình để sinh chuỗi số liệu có độ dài *n* đủ lớn sao cho chuỗi số liệu sinh bảo toàn được các đặc trưng số thống kê gồm *giá trị trung bình, độ lệch chuẩn, hệ số lệch và hệ số tương quan* của chuỗi lưu lượng lịch sử.

Các đặc trưng số thống kê của chuỗi lưu lượng dòng chảy lịch sử hàng tháng: giá trị trung bình, độ lệch chuẩn, hệ số lệch được tính bởi các phương trình:

$$m_j = \frac{1}{N}\sum_{i=1}^{N} X_{i,j}\ , \qquad j = 1..12$$

$$s_j^2 = \frac{1}{N-1}\sum_{i=1}^{N} (X_{i,j} - m_j)^2\ , \qquad j = 1..12$$

$$g_j = \frac{N}{(N-1)(N-2)s_j^3}\sum_{i=1}^{N} (X_{i,j} - m_j)^3, j = 1..12.$$

Các mô hình và phương pháp được đề xuất dùng để mô phỏng lưu lượng dòng chảy được phân thành *nhóm mô hình có tham số* và *nhóm mô hình phi tham số*. Nhóm mô hình có tham số được chia thành các loại mô hình *độc lập và phụ thuộc* cuả chuỗi lưu lượng lịch sử. Với giả thiết chuỗi lưu lượng lịch sử là độc lập có liên quan đến kiểu phân phối xác suất thì nhiều mô hình được đề xuất và trong đó, mô hình Thomas-Fiering (1962) biểu thị lưu lượng dòng chảy với bất kỳ kiểu phân phối xác suất được sử dụng phổ biến. Với sự đa dạng về khí hậu, nhiều công trình nghiên cứu xác định kiểu phân phối của lưu lượng dòng chảy thường không có phân phối chuẩn, có độ lệch và phụ thuộc, và đối với trường hợp này, theo Fernandez và Salas (1990) thì áp dụng mô hình GAR(1) là rất hiệu quả để mô phỏng lưu lượng dòng chảy hàng năm.

1.6. Bài toán ước lượng dung tích hồ chứa

Trong các nghiên cứu về hồ chứa, nhiều bài toán được đặt ra như bài toán quy hoạch, thiết kế, bài toán vận hành hồ chứa hoặc vận hành liên hồ chứa. Đối với lớp các bài toán quy hoạch, thiết kế hồ chứa, vấn đề quan trọng là xác định được dung tích của hồ chứa trên cơ sở các nguồn nước chảy vào và điều tiết dòng chảy ra khỏi hồ chứa. Các nghiên cứu về dung tích hồ chứa tuỳ thuộc vào các trường hợp *hồ chứa có dung tích hữu hạn, bán hữu hạn* hoặc *vô hạn*. Một hồ chứa hữu hạn có thể có lượng nước trong hồ tràn đầy và cạn kiệt, hồ chứa bán hữu hạn chỉ có thể có một trong hai trường hợp hoặc tràn đầy hoặc cạn kiệt. Đối với hồ chứa có dung tích vô hạn thì giả thiết rằng hồ chứa không bao giờ tràn đầy hoặc kiệt nước trong khoảng thời gian hoạt động của nó là n năm, theo Salas-La Cruz (1972), giả thiết này phù hợp cho việc nghiên cứu quy hoạch, thiết kế các hồ chứa có dung tích lớn (hàng trăm triệu m^3 trở lên). Với sự biến đổi khí hậu toàn cầu hiện nay, mưa và khô hạn kéo dài dẫn đến lũ lụt và hạn hán phổ biến ở nhiều quốc gia, thực tế này đòi hỏi cần nghiên cứu xây dựng các hồ chứa có dung tích lớn để điều tiết nguồn nước hợp lý, vì vậy, việc nghiên cứu dung lượng hồ chứa để phục vụ việc thiết kế các hồ chứa có dung tích lớn cần được quan tâm.

KẾT LUẬN CHƯƠNG 1

Từ việc nghiên cứu có hệ thống theo chủ đề của các công trình đã công bố, Tác giả luận án phát hiện những hạn chế sau đây:

- Chưa có nghiên cứu, đánh giá, chọn lựa các thuật toán thích hợp để sinh biến ngẫu nhiên GAR(1), chưa có nghiên cứu đề xuất mô hình biểu thị lưu lượng dòng chảy hàng tháng với quá trình ngẫu nhiên GAR(1) và chưa có nghiên cứu xác định dung lượng trung bình của hồ chứa với dòng chảy vào hồ chứa là quá trình ngẫu nhiên GAR(1).

Từ những hạn chế nêu trên, định hướng nghiên cứu là nghiên cứu đánh giá và chọn lựa các thuật toán sinh biến ngẫu nhiên thích hợp để sinh biến ngẫu nhiên GAR(1), nghiên cứu các đặc trưng số cơ bản của tổng các biến ngẫu nhiên GAR(1), nghiên cứu bài toán mô phỏng lưu lượng dòng chảy hàng tháng, hàng năm với quá trình ngẫu nhiên GAR(1) và nghiên cứu mô phỏng dung lượng trung bình của hồ chứa với dòng chảy vào hồ chứa là quá trình ngẫu nhiên GAR(1).

CHƯƠNG 2
CÁC THUẬT TOÁN SINH BIẾN NGẪU NHIÊN GAR(1)

Nội dung chương này trình bày các thuật toán sinh biến ngẫu nhiên GAR(1). Bằng phương pháp nghiên cứu lý thuyết và phương pháp mô phỏng, các vấn đề lý luận cơ bản và các thuật toán sinh biến ngẫu nhiên GAR(1) được nghiên cứu, cài đặt và thử nghiệm.

2.1. Nghiên cứu một số thuật toán dùng để sinh biến ngẫu nhiên GAR(1)

Để áp dụng mô hình GAR(1) vào thực tế, cần phải sinh biến ngẫu nhiên GAR(1) dựa vào mẫu thống kê. Để sinh biến ngẫu nhiên GAR(1) cần kết hợp các thuật toán sinh biến ngẫu nhiên có phân phối đều đơn vị, phân phối mũ, phân phối chuẩn, phân phối Poisson và phân phối gamma.

2.2. Đề xuất thuật toán sinh biến ngẫu nhiên gamma với giá trị bất kỳ của tham số hình dạng *a*

Thuật toán do Minh (1988) đề xuất được sử dụng để sinh biến ngẫu nhiên có phân phối gamma với tham số hình dạng $a>1$. Dựa vào kết quả của Marsaglia và Tsang (2000), thuật toán cải tiến từ thuật toán Minh được đề xuất bởi Hung, Trang và Chien (2014) gọi là thuật toán IMGAG để sinh biến ngẫu nhiên gamma với giá trị bất kỳ của tham số a của phân phối gamma như sau:

(1) Nếu $a>1$ sử dụng thuật toán của Minh với tham số a để sinh X, chuyển đến bước (3);

(2) Nếu $1\geq a>0$ sử dụng thuật toán của Minh với tham số $a+1$ để sinh X^*, tính $X = X^*U^{1/a}$ với $U \sim$ U(0,1) (U có phân phối đều trong khoảng (0,1));

(3) Nhận được X;

(4) Kết thúc.

2.3. Đề xuất bổ sung tiêu chí đánh giá hiệu quả của thuật toán sinh biến ngẫu nhiên

Trong thực tế, việc đánh giá tính hiệu quả các thuật toán sinh biến ngẫu nhiên chủ yếu dựa vào các tiêu chí là độ phức tạp và tính dễ cài đặt của thuật toán. Ngoài các tiêu chí nêu trên; Hung, Trang và

Chien (2014) đề xuất bổ sung tiêu chí để đánh giá tính hiệu quả của các thuật toán khác nhau dùng để sinh biến ngẫu nhiên có kiểu phân phối xác suất xác định là sử dụng thuật toán sinh chuỗi số ngẫu nhiên độc lập và kiểm tra sự bảo toàn các đặc trưng số gồm: kỳ vọng, phương sai và hệ số lệch của chuỗi số phát sinh.

2.4. Mô phỏng thực nghiệm

2.4.1. Phương pháp mô phỏng

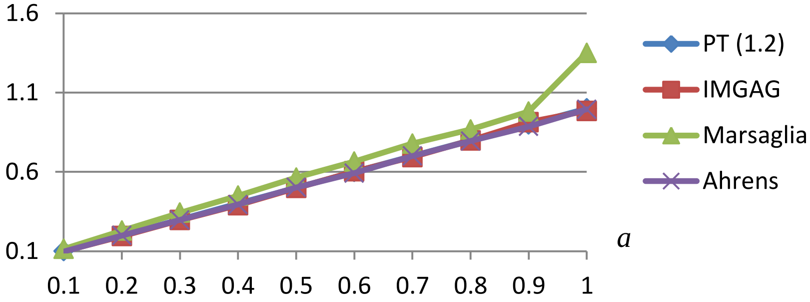
Sử dụng các thuật toán sinh biến ngẫu nhiên gamma: Thuật toán Ahrens (1974) sử dụng cho trường hợp tham số $a \leq 1$, thuật toán Tadikamalla (1978) sử dụng cho trường hợp tham số $a>1$, thuật toán IMGAG và thuật toán Marsaglia (2000) sử dụng cho mọi giá trị của tham số *a*. Các thuật toán được cài đặt bằng ngôn ngữ C và sử dụng mỗi thuật toán để sinh 10.000 số ngẫu nhiên có phân phối gamma với các tham số *a* khác nhau (từ 0.1 đến 500). Dựa vào mẫu các số ngẫu nhiên được sinh, các đặc trưng số thống kê gồm giá trị trung bình, phương sai và hệ số lệch được tính theo các công thức (1.10) - (1.12). Hệ số tương quan tính theo công thức (1.13).

2.4.2. Kết quả mô phỏng

Từ mô phỏng thử nghiệm, kết qủa được trình bày trong các bảng 2.1 - 2.3 và các hình vẽ 2.1 - 2.3 như sau:

Bảng 2.1. Giá trị trung bình của 10.000 số ngẫu nhiên gamma được sinh theo thuật toán IMGAG, thuật toán Marsaglia và thuật toán Ahrens

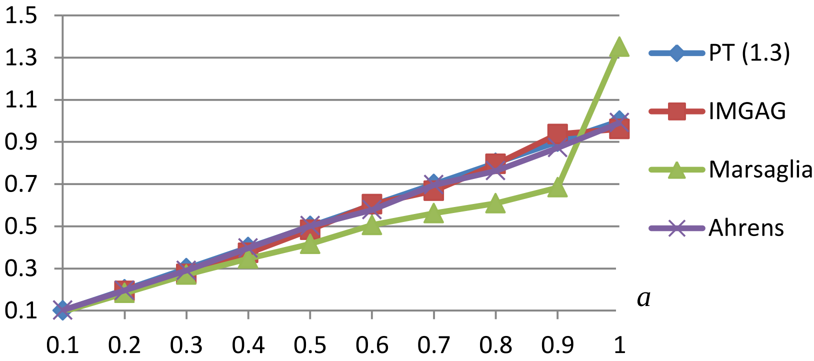
a	IMGAG		Marsaglia		Ahrens	
	TB sinh	% sai số	TB sinh	% sai số	TB sinh	% sai số
0.1	0.099	*0.78*	0.114	*14.32*	0.098	*2.13*
0.2	0.195	2.39	0.230	*15.02*	0.199	*0.55*
0.3	0.296	*1.27*	0.343	*14.38*	0.297	*1.09*
0.4	0.390	*2.57*	0.450	*12.67*	0.394	*1.54*
0.5	0.498	*0.41*	0.564	*12.79*	0.502	*0.34*
0.6	0.603	*0.58*	0.665	*10.90*	0.592	*1.26*
0.7	0.693	*1.04*	0.778	*11.14*	0.700	*0.00*
0.8	0.798	0.30	0.867	*8.43*	0.794	*0.78*
0.9	0.914	*1.55*	0.980	*8.94*	0.886	*1.54*
1.0	0.984	*1.60*	1.350	*35.03*	0.995	*0.53*



Hình 2.1: Giá trị trung bình với các tham số hình dạng $a \leq 1$

Bảng 2.2. Phương sai của 10.000 số ngẫu nhiên gamma được sinh theo thuật toán IMGAG, thuật toán Marsaglia và thuật toán Ahrens

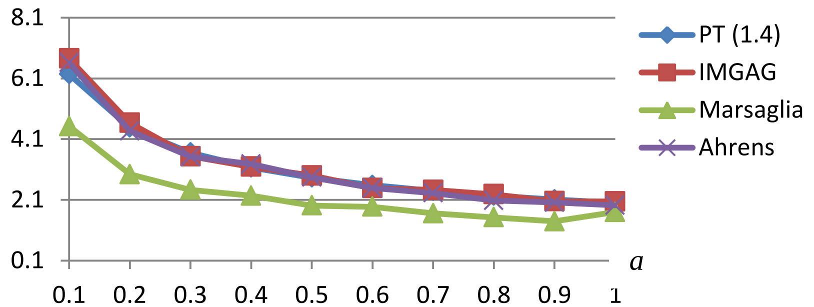
a	IMGAG		Marsaglia		Ahrens	
	PS sinh	% sai số	PS sinh	% sai số	PS sinh	% sai số
0.1	0.098	*1.79*	0.094	*6.44*	0.102	*2.13*
0.2	0.192	*4.18*	0.183	*8.54*	0.196	*2.25*
0.3	0.273	*8.03*	0.270	*10.08*	0.290	*3.34*
0.4	0.373	*6.78*	0.346	*14.89*	0.396	*1.01*
0.5	0.483	*3.42*	0.416	*16.71*	0.502	*0.36*
0.6	0.604	*0.70*	0.506	*15.59*	0.578	*3.67*
0.7	0.668	*4.53*	0.562	*19.74*	0.696	*0.52*
0.8	0.795	0.64	0.609	*23.92*	0.763	*4.60*
0.9	0.937	*4.12*	0.684	*23.99*	0.872	*3.09*
1.0	0.961	*3.86*	1.351	*35.06*	0.991	*0.86*



Hình 2.2: Phương sai với các tham số hình dạng $a \leq 1$

Bảng 2.3. Hệ số lệch của 10.000 số ngẫu nhiên gamma được sinh theo thuật toán IMGAG, thuật toán Marsaglia và thuật toán Ahrens

a	Hệ số lệch lý thuyết	IMGAG		Marsaglia		Ahrens	
		HSL sinh	% Sai số	HSL sinh	% Sai số	HSL sinh	% Sai số
0.1	6.235	6.752	*6.75*	4.524	*28.47*	6.614	*4.57*
0.2	4.472	4.633	*3.36*	2.938	*34.30*	4.363	*2.44*
0.3	3.651	3.530	*3.34*	2.429	*33.47*	3.521	*3.58*
0.4	3.162	3.187	*0.78*	2.235	*29.31*	3.276	*3.59*
0.5	2.828	2.898	*2.45*	1.912	*32.40*	2.840	*0.42*
0.6	2.582	2.480	*3.94*	1.872	*27.51*	2.486	*3.73*
0.7	2.390	2.422	*1.30*	1.653	*30.87*	2.323	*2.82*
0.8	2.236	2.283	*2.10*	1.525	*31.78*	2.074	*7.24*
0.9	2.108	2.048	*2.86*	1.393	*33.93*	2.011	*4.59*
1.0	2.000	2.046	*2.28*	1.698	*15.08*	1.917	*4.13*



Hình 2.3: Hệ số lệch với các tham số hình dạng $a \leq 1$

Đối với trường hợp tham số $a>1$, sử dụng các thuật toán IMGAG, thuật toán Marsaglia, thuật toán Tadikamalla và thu được các bảng và hình vẽ tương ứng.

KẾT LUẬN CHƯƠNG 2

Qua nghiên cứu ở chương 2, Tác giả đạt được các kết quả sau đây: nghiên cứu các thuật toán sinh biến ngẫu nhiên GAR(1) bao gồm các thuật toán sinh biến ngẫu nhiên có phân phối đều, phân phối chuẩn, phân phối mũ, phân phối Poisson và phân phối gamma. Tác giả nghiên cứu đề xuất thuật toán IMGAG để sinh biến ngẫu nhiên gamma với mọi giá trị của tham số hình dạng $a>0$ và đề xuất bổ sung tiêu chí để đánh giá tính hiệu quả của thuật toán sinh biến ngẫu nhiên là dựa vào kỹ thuật mô phỏng và sử dụng thuật toán để sinh một chuỗi số ngẫu nhiên, dựa vào chuỗi số ngẫu nhiên được sinh, kiểm tra tính độc lập và sự bảo toàn các đặc trưng số gồm kỳ vọng, phương sai và hệ số lệch của phân phối xác suất xác định. Các kết quả chi tiết sẽ được trình bày ở phần kết luận của Luận án.

CHƯƠNG 3
MÔ PHỎNG LƯU LƯỢNG DÒNG CHẢY VỚI QUÁ TRÌNH NGẪU NHIÊN GAR(1)

Nội dung chương này trình bày nghiên cứu về các mô hình và các thuật toán dùng để mô phỏng lưu lượng dòng chảy. Tác giả sử dụng mô hình GAR(1), nghiên cứu mô hình Thomas-Fiering, phương pháp Fragments và đề xuất mô hình GAR(1)-Monthly và mô hình GAR(1)-Fragments dùng để mô phỏng lưu lượng dòng chảy hàng tháng. Bằng phương pháp mô phỏng, các mô hình và các thuật toán được thử nghiệm và đánh giá sự bảo toàn các đặc trưng số thống kê gồm giá trị trung bình, độ lệch chuẩn và hệ số lệch của chuỗi lưu lượng dòng chảy lịch sử.

3.1. Bài toán mô phỏng lưu lượng dòng chảy

Trên cơ sở chuỗi lưu lượng lịch sử quan trắc được tại các trạm đo thuỷ văn, bài toán mô phỏng lưu lượng dòng chảy trở thành việc đánh giá tính bảo toàn các đặc trưng số của các chuỗi lịch sử quan trắc gồm giá trị trung bình, độ lệch chuẩn, hệ số lệch và hệ số tương quan khi sử dụng mô hình để sinh các chuỗi lưu lượng dòng chảy (theo hàng tháng, hàng năm tại các trạm đo thuỷ văn) có độ dài n đủ lớn.

3.2. Mô hình Thomas-Fiering (Th.Fiering)

Trên cơ sở mẫu thống kê lưu lượng dòng chảy hàng tháng qua N năm (N gọi là kích thước của mẫu thống kê) tại một trạm đo, mô hình

Thomas-Fiering dùng để diễn tả chuỗi lưu lượng dòng chảy này theo hàng tháng như sau:

$$Q_{i,j} = Q_j^* + b_j\left(Q_{i,j-1} - Q_{j-1}^*\right) + s_j(1 - r_j^2)^{\frac{1}{2}} t_{i,j}, \quad (3.1)$$

trong đó: $Q_{i,j}$ là lưu lượng dòng chảy tháng *j* của năm *i*; b_j là hệ số hồi quy để ước lượng lưu lượng dòng chảy tháng *j* từ tháng *j*-1; Q_j^* và s_j là trung bình và độ lệch chuẩn của chuỗi lịch sử của tháng *j*; r_j là hệ số tương quan giữa chuỗi lưu lượng dòng chảy lịch sử tháng *j* và tháng *j*-1 và $t_{i,j}$ là một biến ngẫu nhiên có trung bình là 0 và phương sai đơn vị.

3.3. Phương pháp Fragments

Trên cơ sở chuỗi lưu lượng lịch sử hàng tháng của *N* năm, Svanidze (1964) đề xuất phương pháp mô phỏng lưu lượng dòng chảy hàng tháng bằng cách sinh chuỗi lưu lượng hàng năm theo mô hình lưu lượng hàng năm và kết hợp với mảnh lưu lượng lịch sử hàng tháng theo từng năm một cách ngẫu nhiên để tính lưu lượng hàng tháng. Phương pháp này không bảo toàn tốt hệ số tương quan của chuỗi lưu lượng lịch sử giữa tháng 1 của năm hiện tại và tháng 12 của năm trước. Srikanthan và McMahon (1980) đề xuất phương pháp Fragments cải tiến để khắc phục hạn chế này bằng cách sắp xếp chuỗi lưu lượng lịch sử hàng tháng thành *N* lớp tăng dần theo lưu lượng hàng năm và lưu lượng sinh hàng năm sẽ được kết hợp với lớp lưu lượng hàng tháng phù hợp (đã được sắp xếp) để tính lưu lượng hàng tháng

3.4. Đề xuất các mô hình biểu thị lưu lượng dòng chảy hàng tháng với quá trình ngẫu nhiên GAR(1)

3.4.1. Mô hình GAR(1)-Monthly (GAR(1)-M)

Mô hình GAR(1) được sử dụng để mô phỏng lưu lượng dòng chảy hàng năm: $X_i = \Phi X_{i-1} + e_i$.

Theo kết quả của Hung, Phien và Chien (2014); với chuỗi dữ liệu hàng tháng của *N* năm, dữ liệu của mỗi tháng qua *N* năm tạo thành một chuỗi dữ liệu và có thể áp dụng mô hình GAR(1). Trường hợp này, mô hình GAR(1) áp dụng cho các chuỗi lưu lượng dòng chảy hàng tháng gọi là *mô hình GAR(1)-Monthly* được biểu diễn như sau:

$$X_{i,j} = \Phi_j X_{i-1,j} + e_i, j\text{=}1..12 \tag{3.2}$$

trong đó: $X_{i,j}$ là biến ngẫu nhiên biểu diễn quá trình phụ thuộc ở tháng j năm i; Φ_j là hệ số hồi quy của tháng j qua N năm; e_i là biến ngẫu nhiên độc lập cần được xác định. Mỗi chuỗi biến ngẫu nhiên gamma phụ thuộc biểu diễn cùng một tháng qua N năm có cấu trúc phân phối và hệ số hồi quy riêng, vì vậy hệ thống các phương trình (3.2) là mô hình thích hợp được áp dụng để mô phỏng dữ liệu hàng tháng.

Hung, Phien và Chien (2014) đề xuất: trong thực tế, hệ số tương quan r_j giữa cùng một tháng j qua các năm liên tiếp có thể có giá trị âm và điều này có thể dẫn đến hệ số hồi quy Φ_j có giá trị âm và mô hình GAR(1)-Monthly không thể áp dụng được. Để áp dụng được mô hình GAR(1)-Monthly cần phải khử giá trị âm của hệ số tương quan bằng cách tính: $r_j = -r_j$ nếu $r_j < 0$.

- ***Thiết kế thuật toán mô phỏng***

(1) Khởi tạo và cập nhật mảng lưu lượng lịch sử hàng tháng $A[N][12]$, N (số năm của chuỗi lịch sử), n (số năm của mẫu sinh);
(2) Khởi tạo mảng lưu lượng sinh hàng tháng $Q[n][12]$;
(3) Sử dụng các các công thức (1.6) - (1.13) và điều chỉnh độ lệch để tính 12 bộ tham số *a, b, c* và Φ của mô hình GAR(1)-Monthly (mỗi bộ tham số tương ứng với 1 chuỗi lịch sử theo từng tháng qua các năm);
(4) Với j = 1 đến 12: nếu $r_j < 0$ tính $r_j = -r_j$;
với i = 1 đến n: tính $Q_{i,j} = X_{i,j}$ (Sử dụng mô hình GAR(1) để sinh e_i và tính $X_{i,j}$);
(5) Kết thúc.

3.4.2. Mô hình GAR(1)-Fragments (GAR(1)-F)

Hung, Phien và Chien (2014) nghiên cứu áp dụng mô hình GAR(1) với lưu lượng dòng chảy hàng tháng bằng cách kết hợp mô hình GAR(1) với phương pháp Fragments và đề xuất mô hình gọi là *mô hình GAR(1)-Fragments* dùng để mô phỏng chuỗi lưu lượng dòng chảy hàng tháng. Trên cơ sở chuỗi lưu lượng lịch sử hàng tháng của *N* năm, mô hình GAR(1)-Fragments sinh các giá trị lưu lượng hàng tháng theo thuật toán sau:

Thiết kế thuật toán mô phỏng

(1) Khởi tạo và cập nhật mảng lưu lượng lịch sử hàng tháng $A[N][12]$, *N* (số năm của chuỗi lịch sử), *n* (kích thước mẫu sinh - số năm).

(2) Khởi tạo mảng lưu lượng sinh hàng tháng $Q[n][12]$;

(3) Phân chia chuỗi lịch sử thành *N* lớp, mỗi lớp là 1 năm lịch sử;

(4) Sắp xếp *N* lớp tăng dần theo lưu lượng lịch sử hàng năm ($A_i=\sum_{j=1}^{12} A_{i,j}$, $A_{i,j}$: lưu lượng lịch sử của tháng *j* năm i, sau khi sắp xếp A_1 ứng với lớp có lưu lượng hàng năm bé nhất, A_N ứng với lớp có lưu lượng hàng năm lớn nhất);

(5) Tính cận trên U_i của lớp *i*: $U_i = \frac{A_i+A_{i+1}}{2}$, i = 1,2,..*N*-1. U_N có giá trị lớn tuỳ ý;

(6) Tính các tham số hình dạng, tỉ lê, vị trí và hệ số hồi quy của mô hình GAR(1) dựa vào mẫu lưu lượng lịch sử hàng năm;

(7) Sinh số ngẫu nhiên X_1 có phân phối gamma 3 tham số: hình dạng, tỉ lệ và vị trí (tính ở bước 6);

(8) Chọn lớp có có cận trên bé nhất lớn hơn hoặc bằng X_1 (lớp *i*);

(9) Tính $Q_{1,j} = M_{i,j} \,.\, X_1$: $Q_{1,j}$ là lưu lượng sinh của tháng *j* năm 1, $M_{i,j} = \frac{A_{i,j}}{A_i}$, $M_{i,j}$ là fragment của lưu lượng lịch sử tháng *j* năm *i*.

(10) Tính $Q_{k,j}$, (*k* = 2,..,*n*, *n* là số năm cần sinh): sử dụng mô hình GAR(1) để sinh e_k và tính X_k , (*k* = 2,..,*n*), chọn lớp có cận trên bé nhất lớn hơn hoặc bằng X_k (gọi là lớp *i*) và: $Q_{k,j} = M_{i,j} \,.\, X_k$;

(11) Kết thúc.

3.5. Mô phỏng thực nghiệm

3.5.1. Số liệu và phương pháp mô phỏng

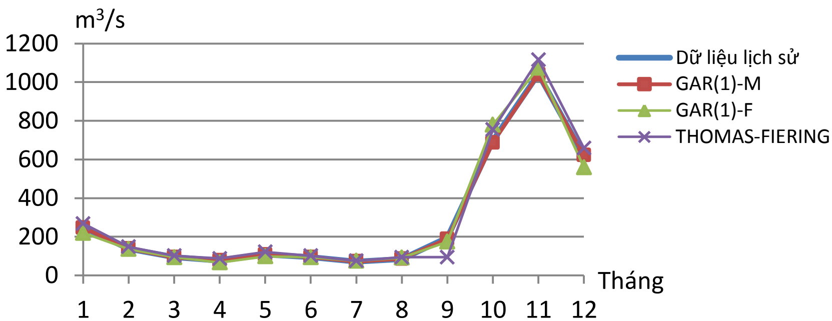
Từ kết quả nghiên cứu ở chương 2, sử dụng các thuật toán thích hợp để sinh các biến ngẫu nhiên trong mô hình Thomas-Fiering, mô hình GAR(1)-Monthly và mô hình GAR(1)-Fragments. Lưu lượng lịch sử hàng tháng (m^3/giây) của các trạm đo Thạnh Mỹ trên sông Vu Gia, trạm đo Nông Sơn trên sông Thu Bồn thuộc tỉnh Quảng Nam từ năm 1980 đến năm 2010 và trạm đo Yên Bái trên sông Thao từ năm 1958 đến năm 2011 được sử dụng. Các thuật toán được cài đặt bằng ngôn ngữ C. Để có được các ước tính chính xác cao, các chuỗi số liệu sinh sẽ được thực hiện với n = 1000 năm.

3.5.2. Kết quả mô phỏng

Kết quả của việc thí nghiệm được trình bày tóm lược trong các bảng 3.1 - 3.4 và các hình vẽ 3.1 - 3.3:

Bảng 3.1. Giá trị trung bình tại trạm đo Nông Sơn

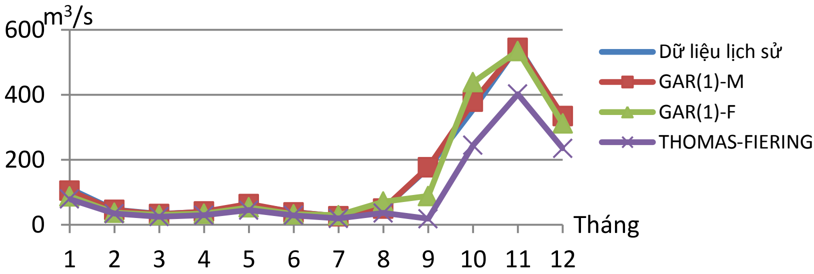
Tháng	Lịch sử	GAR(1)-M	GAR(1)-F	Th.Fiering
1	248.96	245.40	220.25	267.63
2	138.21	137.85	136.53	147.64
3	94.05	93.01	94.06	101.39
4	76.45	76.84	66.42	87.16
5	107.30	106.38	97.66	121.01
6	94.54	94.15	93.68	101.73
7	70.33	71.44	74.95	74.84
8	85.02	85.60	91.32	93.60
9	195.59	195.30	174.61	94.19
10	697.19	705.26	778.81	754.37
11	1041.81	1039.30	1074.54	1116.12
12	619.97	622.19	559.19	659.08



Hình 3.1: Giá trị trung bình tại trạm đo Nông Sơn

Bảng 3.2. Độ lệch chuẩn tại trạm đo Nông Sơn

Tháng	Lịch sử	GAR(1)-M	GAR(1)-F	Th.Fiering
1	110.97	104.54	87.42	79.22
2	46.07	45.50	37.07	34.23
3	33.30	32.67	30.37	24.61
4	39.32	40.82	34.25	29.29
5	60.89	63.72	53.22	45.05
6	39.63	38.2	32.01	29.01
7	25.65	26.07	29.32	19.35
8	48.82	49.52	71.14	36.02
9	174.70	178.68	88.39	18.56
10	354.16	376.42	438.79	244.56
11	549.65	544.42	534.59	401.98
12	329.72	334.52	311.34	235.41

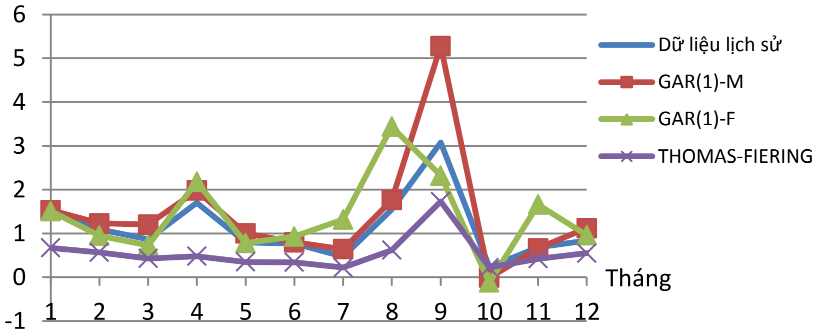


Hình 3.2: Độ lệch chuẩn tại trạm đo Nông Sơn

Bảng 3.3. Hệ số lệch tại trạm đo Nông Sơn

Tháng	Lịch sử	GAR(1)-M	GAR(1)-F	Th.Fiering
1	1.54	1.53	1.51	0.67
2	1.09	1.23	0.95	0.57
3	0.87	1.20	0.73	0.43
4	1.70	1.98	2.18	0.48
5	0.79	1.00	0.78	0.35
6	0.77	0.80	0.93	0.34
7	0.47	0.64	1.32	0.22
8	1.55	1.76	3.44	0.62

9	3.08	5.17	2.32	1.73
10	0.23	-0.01	-0.12	0.22
11	0.68	0.66	1.66	0.42
12	0.84	1.12	0.96	0.55



Hình 3.3 Hệ số lệch tại trạm đo Nông Sơn

Bảng 3.4. Các đặc trưng số thống kê hàng năm tại trạm đo Nông Sơn

Đặc trưng số	Lịch sử	GAR(1)-M	GAR(1)-F	Th.Fiering
Giá trị trung bình	3469.72	3454.17	3467.92	3588.66
Độ lệch chuẩn	1030.77	729.03	1025.29	664.64
Hệ số lệch	0.76	0.32	0.78	0.08

Tương tự tại các trạm đo Thạnh Mỹ và Yên Bái, Tác giả cũng thu được các bảng và các hình vẽ tương ứng.

KẾT LUẬN CHƯƠNG 3

Trong chương 3, Tác giả đã thực hiện nghiên cứu và đạt được kết quả như sau: nghiên cứu và đề xuất các mô hình biểu thị lưu lượng dòng chảy hàng tháng là mô hình GAR(1)-Monthly và mô hình GAR(1)-Fragments. Bằng mô phỏng thực nghiệm, kết quả thu được là mô hình GAR(1)-Monthly bảo toàn các đặc trưng số thống kê gồm giá trị trung bình, độ lệch chuẩn và hệ số lệch tốt hơn các mô hình GAR(1)-Fragments và mô hình Thomas-Fiering, và, trên cơ sở dữ liệu hàng tháng để tính dữ liệu hàng năm thì mô hình GAR(1)-Fragments bảo toàn các đặc trưng số thống kê gồm giá trị trung bình, độ lệch chuẩn và hệ số lệch tốt hơn so với mô hình GAR(1)-Monthly và mô hình Thomas-Fiering.

CHƯƠNG 4
DUNG LƯỢNG TRUNG BÌNH CỦA HỒ CHỨA VỚI QUÁ TRÌNH NGẪU NHIÊN GAR(1)

Nội dung chương này trình bày nghiên cứu về bài toán tính dung lượng trung bình của hồ chứa. Bằng phương pháp phân tích lý thuyết, các biểu thức giải tích về kỳ vọng và phương sai của tổng các biến ngẫu nhiên GAR(1) được đề xuất. Kết hợp công thức của Phien (1978) với biểu thức giải tích về phương sai của tổng các biến ngẫu nhiên có phân phối GAR(1) đã đạt được, Tác giả đề xuất biểu thức xấp xỉ dùng để tính dung lượng trung bình của hồ chứa với dòng vào là các biến ngẫu nhiên GAR(1). Bằng kỹ thuật mô phỏng, sử dụng mô hình GAR(1) phát sinh lưu lượng hàng năm chảy vào hồ chứa và thu được các giá trị về dung lượng trung bình của hồ chứa với các tham số khác nhau và được so sánh với các giá trị theo biểu thức xỉ.

4.1. Dung lượng của hồ chứa

4.1.1. Phương trình tính dung lượng hồ chứa tổng quát

Xem $\{z_i\}$ là một chuỗi các biến ngẫu nhiên với $E(z_i) = 0$ khi đó tổng tích luỹ hay tổng riêng gọi là S_i, cực đại của tổng riêng hay lượng dư thừa M_n, cực tiểu của tổng riêng hay lượng thiếu hụt m_n, và biên độ dao động của tổng riêng R_n của dãy gồm *n* biến ngẫu nhiên z_i được định nghĩa như sau:

$$S_i = z_1 + z_2 + \dots + z_i \, , i = 1,2, \dots , n \quad (4.1)$$

$$M_n = max(0, S_1, S_2, \dots , S_n), \quad (4.2)$$

$$m_n = \min(0, S_1, S_2, \dots , S_n), \quad (4.3)$$

$$R_n = M_n - m_n, \quad (4.4)$$

dễ thấy rằng $M_n \geq 0$, $m_n \leq 0$ và $E(S_i) = 0$.

4.1.2. Dung lượng trung bình của hồ chứa với dòng chảy vào là các biến ngẫu nhiên độc lập

Dung lượng trung bình của hồ chứa được nghiên cứu với giả thiết rằng các dòng chảy vào hồ chứa (I_k) là chuỗi các biến ngẫu nhiên độc lập. Để loại bỏ sự phụ thuộc của dung lượng trung bình của hồ chứa vào các kiểu phân phối khác nhau, một biến ngẫu nhiên mới được sử dụng bằng cách chuẩn hoá I_k:

$$Z_k = \frac{I_k - \mu}{\sigma} = \frac{z_k}{\sigma},$$

ở đây σ là độ lệch chuẩn của I_k. Biến ngẫu nhiên đã được chuẩn hoá Z_k có trung bình bằng 0 và phương sai đơn vị.

Với việc sử dụng biến ngẫu nhiên mới Z_k, nếu $E(R_n^0)$ và $E(R_n)$ là các giá trị kỳ vọng của biên độ dao động của dung lượng tương ứng với z và Z, do đó ta có:

$$E(R_n^0) = \sigma E(R_n).$$

Bằng phương pháp sử dụng hàm đa biến, với giả thiết dòng chảy vào hồ chứa là chuỗi các biến ngẫu nhiên Z_k có phân phối chuẩn, Salas-La Cruz (1972) cho kết quả dung lượng trung bình của hồ chứa như sau:

$$E(R_n) = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} \frac{[Var(S_i)]^{\frac{1}{2}}}{i}.$$

Với trường hợp chuỗi các biến ngẫu nhiên Z_k có phân phối gamma độc lập, theo Phien (1978) thì hệ số lệch g của phân phối gamma cần được tính đến và cho kết quả là biểu thức xấp xỉ tính dung lượng trung bình của hồ chứa là:

$$E(R_n) = \sqrt{\frac{2}{\pi}} \sum_{i=1}^{n} \frac{[Var(S_i)]^{\frac{1}{2}}}{i} . e^{-\frac{0.0475 g^{1.65}}{0.7(n-1)^{0.6}+2}}. \qquad (4.5)$$

4.2. Các đặc trưng số cơ bản của tổng các biến ngẫu nhiên GAR(1)

Các biến ngẫu nhiên X_i theo mô hình GAR(1) được biểu diễn bởi phương trình: $X_i = \Phi X_{i-1} + e_i$.

Khi đó tổng của n biến ngẫu nhiên GAR(1) là một biến ngẫu nhiên gọi là S_n được tính theo phương trình :

$$S_n = \sum_{i=1}^{n} X_i,$$

trong đó: X_i, $i = 1, 2, \ldots, n$ là các biến GAR(1).

Bằng phân tích lý thuyết, Hung va Chien (2013) đạt được các biểu thức giải tích về các đặc trưng số cơ bản: kỳ vọng và phương sai của tổng các biến ngẫu nhiên GAR(1) với phân phối gamma 1 tham số như sau:

Kỳ vọng của tổng của *n* biến ngẫu nhiên GAR(1) gọi là $E(S_n)$ và $E(S_n) = na$.

Phương sai của tổng của *n* biến ngẫu nhiên GAR(1) được ký hiệu là $Var(S_n)$ và: $Var(S_n) = na + 2a\sum_{k=1}^{n-1}(n-k)\Phi^k$. (4.6)

4.3. Biểu thức xấp xỉ của dung lượng trung bình hồ chứa với dòng chảy vào là các biến ngẫu nhiên GAR(1)

Biên độ dao động của dung lượng hồ chứa được xem xét là tổng luỹ tích

$$S_i = \sum_{k=1}^{i} z_k = \sum_{k=1}^{i}(I_k - \mu),$$

trong đó z_k là sự dao động $I_k - \mu$ của I_k xung quanh giá trị trung bình dài hạn μ của I_k và I_k là một biến ngẫu nhiên có phân phối gamma phụ thuộc tuân theo mô hình GAR(1):

$$I_k = \Phi I_{k-1} + e_k.$$

Theo kết quả của Phien(1978), hệ số lệch được tính đến, và theo kết quả của Hung và Chien(2013), thay thế phương sai của tổng các biến ngẫu nhiên GAR(1) ở phương trình (4.6) vào phương trình (4.5) ta thu được biểu thức xấp xỉ dùng để tính dung lượng trung bình của hồ chứa với các biến ngẫu nhiên theo GAR(1) đã được chuẩn hoá là:

$$E(R_n) = \sqrt{\frac{2}{\pi}}\sum_{i=1}^{n}\frac{\left[i + 2\sum_{k=1}^{i-1}(i-k)\emptyset^k\right]^{\frac{1}{2}}}{i}.e^{-\frac{0.0475g^{1.65}}{0.7(n-1)^{0.6}+2}}. \quad (4.7)$$

4.4. Mô phỏng thực nghiệm

4.4.1. Số liệu và phương pháp mô phỏng

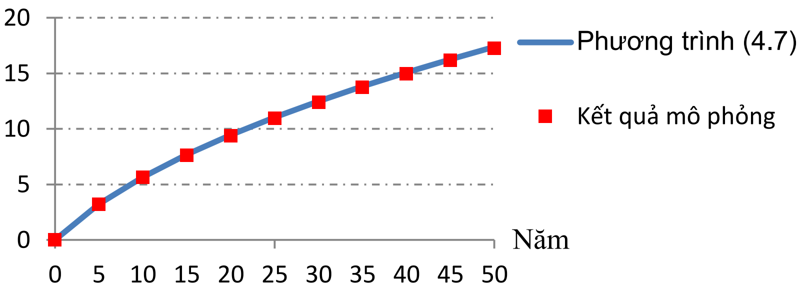
Với mỗi giá trị về hệ số lệch của phân phối gamma và giá trị về hệ số hồi quy của mô hình GAR(1), một mẫu gồm *n* = 100.000 chuỗi các biến ngẫu nhiên GAR(1) được sinh, mỗi chuỗi gồm *N* = 50 giá trị, mỗi giá trị tương ứng với một biên độ dao động của dung lượng của hồ chứa và được sử dụng để tính dung lượng trung bình của hồ chứa có độ dài (tuổi thọ) 50 năm. Tương tự, tính cho các hồ chứa có tuổi thọ (*l* năm) ngắn hơn (*l* < 50), mỗi chuỗi gồm *l* giá trị được sử dụng để tính dung lượng trung bình của hồ chứa có tuổi thọ *l* năm tương ứng. Trong thực nghiệm này, Tác giả sử dụng hệ số lệch của phân phối gamma có giá trị trong khoảng [0.5,3.0], theo Phien (1993); điều này phù hợp với hầu hết các trường hợp dòng chảy vào hồ chứa trong thực tế. Sử dụng biểu thức xấp xỉ tính dung lượng trung bình của hồ chứa ở công thức (4.7) và bằng phương pháp mô phỏng, số liệu được sinh tương ứng với biên độ dao động của dung lượng của hồ chứa và giá trị trung bình của dung lượng của hồ chứa được tính với các giá trị khác nhau của n, Φ và g.

4.4.2. Kết quả mô phỏng

Kết quả được cho ở bảng 4.1 và hình vẽ 4.1 như sau :

Bảng 4.1. Giá trị dung lượng trung bình của hồ chứa với trường hợp hệ số hồi quy $\Phi = 0.6$ và hệ số lệch $g = 2.0$

l Năm	Dung lượng trung bình hồ chứa		
	Phương trình (4.7)	Kết quả mô phỏng	% sai số
5	3.225	3.197	0.875
10	5.663	5.624	0.693
15	7.688	7.635	0.694
20	9.452	9.380	0.767
25	11.034	10.945	0.813
30	12.482	12.392	0.726
35	13.823	13.727	0.699
40	15.079	14.979	0.667
45	16.264	16.152	0.693
50	17.389	17.265	0.718



Hình 4.1: Hệ số hồi quy Φ = 0.6, hệ số lệch g = 2.0

Tác giả cũng thu được các bảng và các hình vẽ tương tự với các hệ số lệch trong khoảng từ 0.5 đến 3.0 và các hệ số hồi quy trong khoảng từ 0.2 đến 0.8.

KẾT LUẬN CHƯƠNG 4

Ở chương 4, các kết quả đạt được như sau: Phân tích lý thuyết và đạt được biểu thức giải tích về kỳ vọng và phương sai của tổng các biến ngẫu nhiên GAR(1), trên cơ sở biểu thức giải tích về phương sai của tổng các biến ngẫu nhiên GAR(1), Tác giả đề xuất biểu thức xấp xỉ tính dung lượng trung bình của hồ chứa với chuỗi lưu lượng dòng chảy vào hồ chứa như quá trình ngẫu nhiên GAR(1) và được so sánh với kết quả mô phỏng, các giá trị thu được là tương tự với nhau.

KẾT LUẬN LUẬN ÁN

1. Kết quả đạt được

Qua quá trình nghiên cứu ở các chương: Các vấn đề chung, các thuật toán sinh biến ngẫu nhiên GAR(1), mô phỏng lưu lượng dòng chảy với quá trình ngẫu nhiên GAR(1) và dung lượng trung bình của hồ chứa với quá trình ngẫu nhiên GAR(1) được trình bày trong Luận án, những kết quả sau đây đã đạt được:

1.1. Về lý thuyết

- Nghiên cứu đề xuất thuật toán cải tiến từ thuật toán của Minh(1988) gọi là thuật toán IMGAG để sinh biến ngẫu nhiên gamma với mọi giá trị của tham số hình dạng $a > 0$ của phân phối gamma. Đề xuất bổ sung tiêu chí để đánh giá tính hiệu quả của thuật toán sinh biến ngẫu nhiên có kiểu phân phối xác định là dựa vào kỹ thuật mô phỏng và sử dụng thuật toán để sinh một chuỗi số ngẫu nhiên. Trên cơ sở chuỗi số ngẫu nhiên được sinh, kiểm tra tính độc lập (dựa vào hệ số tương quan) và sự bảo toàn các đặc trưng số gồm kỳ vọng, phương sai và hệ số lệch của phân phối xác suất;

- Nghiên cứu đề xuất 2 mô hình: GAR(1)-Monthly và GAR(1)-Fragments dùng để mô phỏng lưu lượng dòng chảy hàng tháng.

- Phân tích lý thuyết và đạt được biểu thức giải tích về kỳ vọng và phương sai của tổng các biến ngẫu nhiên GAR(1). Trên cơ sở biểu thức giải tích về phương sai của tổng các biến ngẫu nhiên GAR(1), kết hợp với kết quả lý thuyết của Salas-La Cruz (1972) và kết quả thực nghiệm của Phien (1978), đề xuất biểu thức xấp xỉ tính dung lượng trung bình của hồ chứa với chuỗi lưu lượng dòng chảy vào hồ chứa như quá trình ngẫu nhiên GAR(1).

1.2. Về mô phỏng thực nghiệm

- Trường hợp tham số hình dạng a<1: Thuật toán IMGAG và thuật toán AHRENS bảo toàn rất tốt các đặc trưng số gồm kỳ vọng, phương sai và hệ số lệch của phân phối gamma trong khi đó thuật toán MARSAGLIA bảo toàn không tốt các đặc trưng số của phân phối gamma. Trường hợp tham số hình dạng 1<a<5, thuật toán TADIKAMALLA và thuật toán IMGAG bảo toàn các đặc trưng số: kỳ vọng, phương sai và hệ số lệch của phân phối gamma tốt hơn thuật toán MARSAGLIA;

- Các mô hình GAR(1)-Monthly, mô hình GAR(1)-Fragments và mô hình Thomas-Fiering bảo toàn tốt các đặc trưng số thống kê hàng tháng: giá trị trung bình và độ lệch chuẩn của các trạm đo được thử nghiệm. Mô hình GAR(1)-Fragments và mô hình Thomas-Fiering không bảo toàn tốt hệ số lệch;

- Mô hình GAR(1)-Monthly bảo toàn các đặc trưng số thống kê gồm giá trị trung bình, độ lệch chuẩn và hệ số lệch tốt hơn các mô hình GAR(1)-Fragments và mô hình Thomas-Fiering;

- Trên cơ sở dữ liệu hàng tháng để tính dữ liệu hàng năm thì mô hình GAR(1)-Fragments bảo toàn các đặc trưng số thống kê gồm giá trị trung bình, độ lệch chuẩn và hệ số lệch tốt hơn so với mô hình GAR(1)-Monthly và mô hình Thomas-Fiering;

- So sánh, đánh giá kết quả thu được bằng phương pháp mô phỏng và biểu thức xấp xỉ tính dung lượng trung bình hồ chứa, biểu thức xấp xỉ và phương pháp mô phỏng cho kết quả tương tự với nhau. Vì vậy, biểu thức xấp xỉ ở phương trình (4.7) có thể được sử dụng trong thực tế để tính dung lượng trung bình của hồ chứa có dung tích lớn.

Với những kết quả đạt được nêu trên, thể hiện: thuật toán IMGAG, biểu thức xấp xỉ, mô hình GAR(1)-Monthly và mô hình GAR(1)-Fragments do Tác giả nghiên cứu đề xuất được kiểm chứng tính hiệu quả bằng mô phỏng thử nghiệm với các số liệu thực tế. Kết quả cho thấy thuật toán IMGAG dùng để sinh biến ngẫu nhiên gamma, biểu thức xấp xỉ dùng để tính dung lượng trung bình hồ chứa, mô hình GAR(1)-Monthly và mô hình GAR(1)-Fragments dùng để mô phỏng lưu lượng dòng chảy hàng tháng với quá trình ngẫu nhiên GAR(1) có thể được ứng dụng hiệu quả trong lãnh vực thủy văn.

2. Hướng nghiên cứu tiếp tục

Bên cạnh những kết quả đã đạt được, hướng nghiên cứu tiếp tục của Luận án bao gồm:

- Để sinh biến ngẫu nhiên GAR(1) cần phải sử dụng các thuật toán sinh các biến ngẫu nhiên có các phân phối đều, phân phối mũ, phân phối chuẩn, phân phối Poisson và phân phối gamma. Nội dung nghiên cứu của Luận án chỉ mới đánh giá các thuật toán sinh biến ngẫu nhiên gamma, vì vậy, sẽ nghiên cứu đánh giá tính hiệu quả của các thuật toán sinh biến ngẫu nhiên có phân phối chuẩn và phân phối Poisson để có thể áp dụng vào thực tế tốt hơn;
- Với mỗi mô hình sinh lưu lượng dòng chảy hàng tháng, phân tích một số đặc trưng số thống kê của số liệu lịch sử hàng tháng tại một số trạm đo thử nghiệm chưa được bảo toàn tốt. Nghiên cứu đánh giá việc bảo toàn hệ số tương quan của các mô hình đã đề xuất;
- Nghiên cứu về biểu thức giải tích của dung lượng trung bình hồ chứa với dòng vào là quá trình ngẫu nhiên GAR(1).

Trên đây là những vấn đề nên được tiếp tục nghiên cứu và giải quyết trong tương lai ./.

DANH MỤC CÔNG TRÌNH ĐÃ ĐƯỢC CÔNG BỐ CỦA TÁC GIẢ

1. Nguyen Van Hung and Tran Quoc Chien, (2013), "*Computer simulation and approximate expression for the mean range of reservoir storage with GAR (1) inflows*" In Proceedings of the Fourth Symposium on Information and Communication Technology, SOICT 2013, DaNang, VietNam, Dec. 5 - 6, 2013 (ACM Digital Library, New York, NY, USA), pp. 11 - 17.
2. Nguyen Van Hung, Huynh Ngoc Phien, Tran Quoc Chien, (2014), *"Computer Simulation of Streamflows with GAR(1)-Monthly and GAR(1)-Fragments Models*", The World of Computer Science and Information Technology Journal (WCSIT, USA), Volume 4, No.11, pp. 150 - 156.
3. Nguyen Van Hung, Ngo Thi Thanh Trang, Tran Quoc Chien, (2014), *"An Improvement of Minh's Algorithm for Generating Gamma Variates with Any Value of Shape Parameter",* Indian Journal of Computer Science and Engineering (IJCSE), Volume 5, No 6, pp. 199 - 205.
4. Nguyễn Văn Hưng, Trần Quốc Chiến, Võ Đình Nam, (2012) *"Nghiên cứu đánh giá các thuật toán sinh biến ngẫu nhiên có phân phối gamma"*, Tạp chí Khoa học và Công nghệ - Đại học Đà Nẵng, Số 10(59), trang 58 - 63.
5. Nguyen Van Hung, Tran Quoc Chien, (2013), *"Computer Simulation of Monthly Streamflows with Thomas-Fiering model and Gar(1)-Fragments model*", Tạp chí Khoa học và Công nghệ - Đại học Đà Nẵng, Số.12(73), (Số tiếng Anh). trang. 46 - 51.
6. Nguyễn Văn Hưng, Ngô Thị Thanh Trang, (2014), *"Mô phỏng lưu lượng dòng chảy hàng tháng với mô hình FGAR(1) và mô hình MGAR(1)",* Tạp chí Khoa học và Công nghệ - Đại học Đà Nẵng, Số 1(74), Quyển 2 (Số đặc biệt dành cho hội nghị RAIT), trang 25 - 29.
7. Nguyễn Văn Hưng, Phan Văn Sơn, Trần Quốc Chiến, (2014), *"Nghiên cứu sự bảo toàn các tham số đặc trưng của mô hình FGAR(1)*", Kỷ yếu Hội thảo Khoa học Hệ thống thông tin, Trường Đại học Sư phạm - Đại học Đà Nẵng, trang 110 - 116.