

VIỆN HÀN LÂM KHOA HỌC VÀ CÔNG NGHỆ VIỆT NAM

VIỆN CƠ HỌC

BÁO CÁO KẾT QUẢ THỰC HIỆN ĐỀ TÀI CƠ SỞ

CẤP VIỆN CƠ HỌC NĂM 2020

ĐỀ TÀI

**NGHIÊN CỨU XÂY DỰNG MÔ HÌNH THỦY VĂN
THÔNG SỐ TẬP TRUNG TRONG DỰ BÁO LŨ
CHO CÁC LƯU VỰC SÔNG Ở VIỆT NAM**

Chủ nhiệm Đề tài: TS. Nguyễn Chính Kiên

HÀ NỘI – 2020

DANH SÁCH CÁN BỘ THAM GIA THỰC HIỆN ĐỀ TÀI

TT	Họ và tên	Phòng chuyên môn
1	TS. Nguyễn Chính Kiên	Thủy động lực và Giảm nhẹ thiên tai trong lưu vực
2	ThS. Dương Thị Thanh Hương	Thủy động lực và Giảm nhẹ thiên tai trong lưu vực
3	ThS. Nguyễn Tuấn Anh	Thủy động lực và Giảm nhẹ thiên tai trong lưu vực
4	PGS. TS. Trần Thu Hà	Thủy động lực và Giảm nhẹ thiên tai trong lưu vực
5	ThS. Nguyễn Hồng Phong	Thủy động lực và Giảm nhẹ thiên tai trong lưu vực
6	KS. Nguyễn Thị Hằng	Thủy động lực và Giảm nhẹ thiên tai trong lưu vực
7	CN. Trần Thị Thanh Huyền	Thủy động lực và Giảm nhẹ thiên tai trong lưu vực

MỤC LỤC

DANH MỤC KÝ HIỆU, CHỮ VIẾT TẮT THƯỜNG SỬ DỤNG	iii
DANH MỤC CÁC BẢNG	iv
DANH MỤC CÁC HÌNH VẼ VÀ ĐỒ THỊ	v
MỞ ĐẦU	1
CHƯƠNG I: TỔNG QUAN VỀ MÔ HÌNH THỦY VĂN	3
1.1 Mô hình thủy văn.....	3
1.1.1 Khái niệm và vai trò.....	3
1.1.2 Phân loại.....	4
1.1.3 Tiêu chuẩn lựa chọn mô hình	7
1.1.4 Một số mô hình thủy văn phổ biến	9
1.2 Mô hình thủy văn thông số tập trung Nielsen-Hansen.....	13
1.2.1 Giới thiệu	13
1.2.2 Cấu trúc của mô hình	14
1.2.3 Các thông số của mô hình.....	16
CHƯƠNG II: PHƯƠNG PHÁP SỐ GIẢI BÀI TOÁN ƯỚC TÍNH THÔNG SỐ TỐI ƯU CHO MÔ HÌNH THỦY VĂN THÔNG SỐ TẬP TRUNG.....	18
2.1 Tổng quan về các phương pháp giải bài toán ước tính thông số tối ưu	18
2.2 Phương pháp ô vuông.....	19
2.3 Phương pháp Rosenbrock	22
2.4 Phương pháp Nelder-Mead	26
2.5 Phương pháp Hooke-Jeeves	27
2.6 Giải thuật di truyền	28
2.7 Phương pháp SCE (Shuffled Complex Evolution)	30
CHƯƠNG III: XÂY DỰNG PHẦN MỀM DỰ BÁO LŨ BẰNG MÔ HÌNH THỦY VĂN THÔNG SỐ TẬP TRUNG CÓ SỬ DỤNG PHƯƠNG PHÁP ƯỚC TÍNH THÔNG SỐ TỐI ƯU	31
3.1 Sơ đồ khối mô hình Nielsen-Hansen	31
3.2 Sơ đồ khối phương pháp số giải bài toán ước tính thông số tối ưu	32
3.2.1 Sơ đồ khối của thuật toán ô vuông.....	32
3.2.2 Sơ đồ khối của thuật toán RosenBrock.....	34

3.2.3 Sơ đồ khối của thuật toán Nelder-Mead	35
3.2.4 Sơ đồ khối của thuật toán Hooke-Jeeves	38
3.2.5 Sơ đồ khối của giải thuật di truyền	39
3.2.6 Sơ đồ khối của thuật toán SCE	41
3.3 Xây dựng phần mềm	44
CHƯƠNG IV: THỬ NGHIỆM CÔNG NGHỆ TRONG TÍNH TOÁN DỰ BÁO LŨ	
TRÊN MỘT SỐ LƯU VỰC SÔNG	47
4.1 Thử nghiệm dự báo.....	47
4.2 Kết quả tính toán và đánh giá	49
KẾT LUẬN VÀ KIẾN NGHỊ	58
TÀI LIỆU THAM KHẢO	59

DANH MỤC KÝ HIỆU, CHỮ VIẾT TẮT THƯỜNG SỬ DỤNG

NST	Nhiệm sắc thể
SCE	Shuffled Complex Evolution
WMO	Tổ chức khí tượng thế giới
NSE	Nash–Sutcliffe Efficiency
SCE	Shuffled Complex Evolution
KTTV	Khí tượng – Thủy văn
LV	Lưu vực
CSDL	Cơ sở dữ liệu
BTNMT	Bộ Tài nguyên môi trường
KHCN	Khoa học công nghệ
VN	Việt Nam

DANH MỤC CÁC BẢNG

Bảng 1.1: Bảng tiêu chuẩn lựa chọn mô hình của WMO.....	8
Bảng 1.2: Các thông số hiệu chỉnh của mô hình Nielsen-Hansen.....	17
Bảng 4.1: Bảng kết quả hệ số NSE khi sử dụng các phương pháp tính toán dò tìm thông số tối ưu được thử nghiệm cho 12 lưu vực.....	49

DANH MỤC CÁC HÌNH VẼ VÀ ĐỒ THỊ

Hình 1.1: Sơ đồ phân loại mô hình dòng chảy	4
Hình 1.2: Sơ đồ mô phỏng của mô hình Nielsen-Hansen	14
Hình 2.1: Mô phỏng phương pháp giải bằng ô vuông	21
Hình 2.2: Các ánh xạ gương, phép co, phép dẫn	26
Hình 3.1: Sơ đồ khối mô hình Nielsen-Hansen	31
Hình 3.2: Sơ đồ khối thuật toán ô vuông	32
Hình 3.3: Sơ đồ khối của thuật toán RosenBrock.....	34
Hình 3.4: Sơ đồ khối của thuật toán Nelder-Mead	35
Hình 3.5: Sơ đồ khối của thuật toán Hooke-Jeeves	38
Hình 3.6: Sơ đồ khối của thuật toán giải thuật di truyền	40
Hình 3.7: Sơ đồ khối thuật toán SCE.....	41
Hình 3.8: Sơ đồ khối thuật toán CCE	43
Hình 3.9: Giao diện chung của phần mềm.....	45
Hình 3.10: Thông tin về các tham số của phương án tính toán	45
Hình 3.11: Một số phương pháp giải có trong phần mềm.....	46
Hình 3.12: Lựa chọn loại hình tính toán cho mô hình thủy văn.....	46
Hình 4.1: Vị trí các lưu vực nghiên cứu.	47
Hình 4.2: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Lâm Sơn..	51
Hình 4.3: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực hồ Bản Chát	51
Hình 4.4: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực sông Cầu..	52
Hình 4.5: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Hàm Yên .	52
Hình 4.6: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Hòa Bình .	53
Hình 4.7: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Huội Quảng	53
Hình 4.8: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Sơn La	54
Hình 4.9: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Thác Bà ...	54

Hình 4.10: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực sông Thương55

Hình 4.11: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Tuyên Quang55

Hình 4.12: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Yên Bái .56

Hình 4.13: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Lục Nam.56

MỞ ĐẦU

1. Tính cấp thiết của đề tài

Dự báo thủy văn, thủy lực là công việc dự báo trước cho tương lai một cách có khoa học về trạng thái biến đổi các yếu tố thủy văn, tuy nhiên sự biến đổi này là một quá trình tự nhiên phức tạp, chịu tác động của rất nhiều yếu tố. Tính biến động của các yếu tố này phụ thuộc vào cả không gian và thời gian nên gây khó khăn rất lớn cho quá trình dự báo. Thêm vào đó, do thiếu các trạm quan trắc cần thiết và thiếu sự kết hợp giữa các ngành liên quan cho nên dữ liệu quan trắc thực tế thường là không đầy đủ, không mang tính chất đại diện.

Dự báo thủy văn có ý nghĩa lớn trong điều tiết dòng chảy, khai thác nguồn nước cho phát điện, giao thông thủy, tưới (nhất là ở vùng khô hạn), cấp nước, quản lý nguồn nước, chất lượng nước. Dự báo thủy văn đặc biệt quan trọng trong đối phó với các hiện tượng nguy hiểm trên sông như lũ, lụt, hạn hán,... cũng như rất quan trọng đối với thiết kế và thi công, vận hành các công trình thủy lợi, các công trình thủy nói chung.

Hiện nay, có rất nhiều phương pháp dự báo đã được đưa ra dựa trên mô hình vật lý và toán học, kết quả của các mô hình nói trên đã đạt được một số thành công đáng ghi nhận. Tuy nhiên, các mô hình đều có những ưu và nhược điểm nhất định, vấn đề tìm kiếm phương pháp đủ tốt, đáp ứng các yêu cầu thực tế giải quyết bài toán dự báo thủy văn, thủy lực vẫn là nội dung nghiên cứu thời sự hiện nay.

Các mô hình thủy văn thường bao gồm hàng chục thông số đặc trưng của từng lưu vực. Thông qua việc sử dụng mô hình thủy văn với các thông số này, ta có được chuỗi giá trị lưu lượng chảy ra. Cần hiệu chỉnh các thông số của mô hình sao cho giá trị lưu lượng tính toán này gần với các giá trị lưu lượng thực đo đã biết trước. Trải qua nhiều lần hiệu chỉnh cho nhiều bộ số liệu (nhiều năm), ta sẽ có bộ thông số phù hợp để sử dụng dự báo cho lưu vực. Tuy nhiên, để hiệu chỉnh các thông số này cùng một lúc bằng kinh nghiệm của người dự báo viên là rất khó khăn, mất nhiều công sức vì kết quả đầu ra (lưu lượng dòng chảy) là một hàm phi tuyến của các thông số này. Ngoài ra, hàm quan hệ phi tuyến này không thể đạo hàm nên cần tìm phương pháp để dò tìm các tham số này là cần thiết.

Tìm kiếm phương pháp giải bài toán ước tính thông số tối ưu trong mô hình thủy văn giúp giải quyết được bài toán hiệu quả, tìm ra bộ thông số phù hợp nhất cho từng vùng phục vụ cho quá trình tác nghiệp dự báo thủy văn một cách chính xác hơn. Trong đề tài này, tác giả chọn lựa hướng đề xuất “*Nghiên cứu phát triển mô hình thủy văn thông số tập trung trong dự báo lũ cho các lưu vực sông ở Việt Nam*”.

2. Mục đích nghiên cứu

Nghiên cứu và xây dựng mô hình thủy văn thông số tập trung sử dụng phương pháp số giải bài toán ước tính thông số tối ưu.

3. Phương pháp nghiên cứu

- Phân tích, thống kê và tính toán các tư liệu thu thập được.
- Mô hình hóa số trị tính toán mô hình thủy văn thông số tập trung có sử dụng ước tính tối ưu bằng việc phát triển chương trình tính toán ngôn ngữ Fortran, C#.

CHƯƠNG I

TỔNG QUAN VỀ MÔ HÌNH THỦY VĂN

1.1 Mô hình thủy văn

1.1.1 Khái niệm và vai trò

a) Một số khái niệm

Dựa trên các nghiên cứu hiện nay, các nhà nghiên cứu thủy văn đã đưa ra những khái niệm khái quát về quá trình thủy văn, theo đó có thể khái quát đơn giản thủy văn là một tập hợp các quá trình vật lý, hóa học và sinh học tác động lên các yếu tố đầu vào, chuyển đổi thành các yếu tố đầu ra.

Dữ liệu cần thiết cho mô hình thủy văn bao gồm rất nhiều loại khác nhau và tùy thuộc vào mỗi mô hình. Mỗi mô hình thủy văn có yêu cầu về dữ liệu khác nhau và được định dạng theo quy định riêng của nó. Tuy nhiên về cơ bản, các mô hình thủy văn thông thường có yêu cầu về một số loại dữ liệu như: địa hình, sử dụng đất, thổ nhưỡng, khí tượng thủy văn, ...

Mô hình toán là một công cụ nghiên cứu khoa học bao gồm cả hệ thống trừu tượng (ý nghĩ) và hệ thống vật lý (vật chất) phản ánh hoặc tái hiện lại các hiện tượng hoặc quá trình đang nghiên cứu. Chúng cho phép thu nhận được lượng thông tin cần thiết để hiểu sâu hơn các hiện tượng đó, hoặc những ghi chép định lượng các quá trình đó.

Dự báo thủy văn là báo trước một cách có khoa học trạng thái (tình hình) biến đổi các yếu tố thủy văn sông, suối, hồ như lượng nước, mực nước. Dự báo thủy văn hạn cực ngắn là dự báo giá trị trong tương lai của yếu tố thủy văn hoặc của lũ, lụt sẽ xảy ra sau một thời khoảng dưới 6 giờ; Dự báo thủy văn hạn ngắn là dự báo giá trị trong tương lai của yếu tố thủy văn hoặc của lũ, lụt sẽ xảy ra sau một thời khoảng không quá 2 ngày kể từ khi phát báo; Dự báo thủy văn hạn vừa (dự báo mở rộng) - Là dự báo giá trị trong tương lai của yếu tố thủy văn hoặc của lũ, lụt sẽ xảy ra sau một thời khoảng từ 2 đến 10 ngày kể từ khi phát báo; Dự báo thủy văn hạn dài - Dự báo giá trị tương lai của yếu tố thủy văn hoặc của lũ, lụt có khả năng xảy ra sau một thời khoảng từ 1 tháng đến 1 năm: Dự báo tháng, mùa và năm; Dự báo thủy văn hạn siêu dài là dự báo giá trị trong tương lai của yếu tố thủy văn trước trên 1 năm: Dự báo biến đổi của nguồn nước trong tương lai 5, 10, 20 năm.

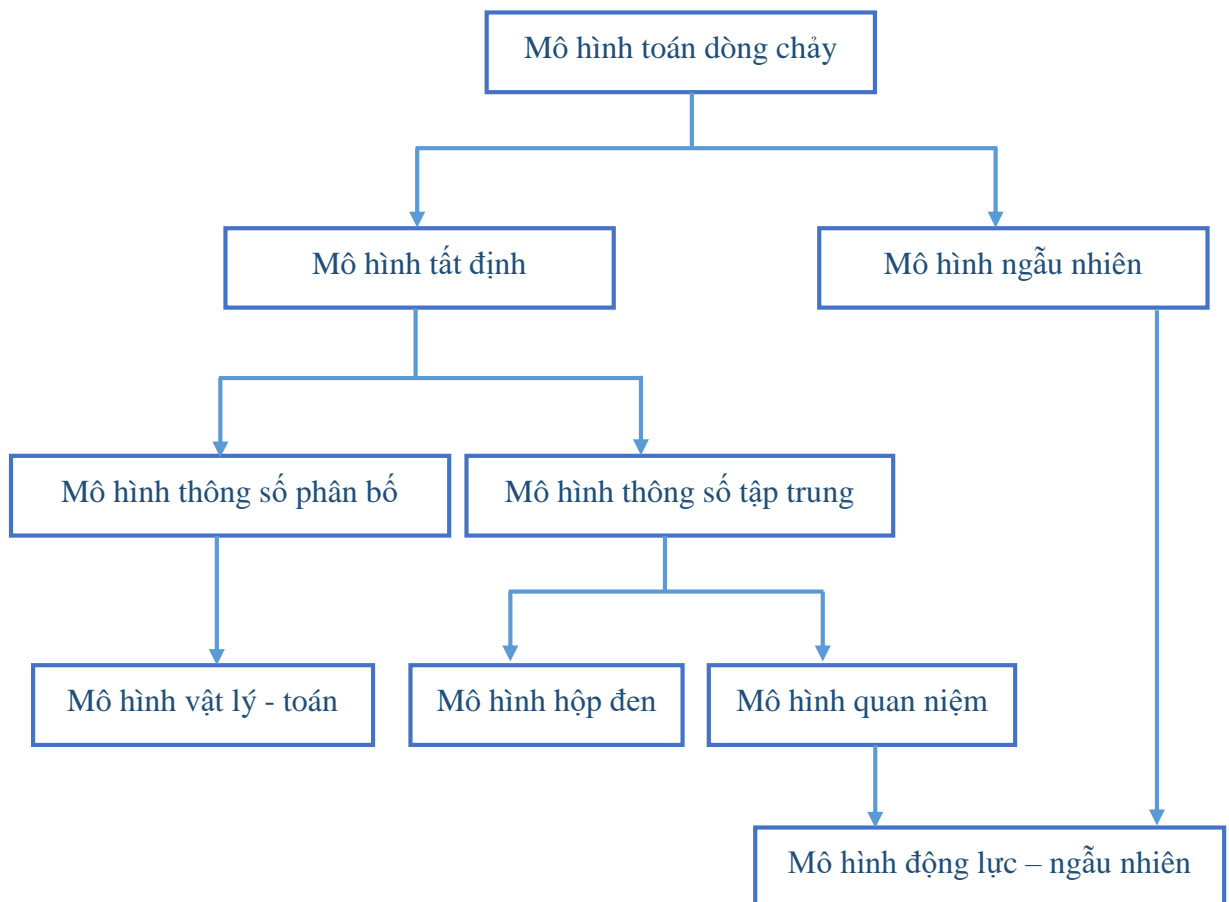
b) Vai trò

Dự báo thủy văn phục vụ rất nhiều ngành, lĩnh vực quan trọng trong cuộc sống như: phục vụ thi công và khai thác công trình thủy điện, phục vụ tưới tiêu, phục vụ giao thông đường thủy, phục vụ các hệ thống thủy nông và phòng chống thiên tai, lũ lụt.

Việc sử dụng mô hình thủy văn đã giúp các nhà thủy văn nói riêng, các nhà khoa học nói chung trong việc mô phỏng quá trình dòng chảy, các hiện tượng thủy văn, sự vận động phức tạp của nước trong tự nhiên dưới dạng các phương trình toán học, logic và giải trên máy tính điện tử.

1.1.2 Phân loại

Trên hàng trăm mô hình hình thành dòng chảy hiện hành, trong dự báo lũ mô hình được chia ra làm 3 loại phân biệt: mô hình tất định, mô hình ngẫu nhiên và mô hình tất định – ngẫu nhiên. Sự phân biệt này cũng nằm ngay trong mục đích mô hình hoá: Dự báo chuỗi dòng chảy trong tương lai phục vụ bài toán thiết kế hay phục vụ bài toán quản lý - điều khiển hệ thống thủy điện, thủy lợi.



Hình 1.1: Sơ đồ phân loại mô hình dòng chảy

a) Mô hình ngẫu nhiên

Quan niệm xác suất lần đầu được Hazen đưa vào trong thủy văn từ năm 1914. Ngày nay, dòng chảy được coi là một quá trình ngẫu nhiên. Với quan điểm này, trong cấu trúc các mô hình ngẫu nhiên không có các nhân tố hình thành dòng chảy và nguyên liệu để xây dựng mô hình chính là bản thân chuỗi dòng chảy quá khứ, phải đủ dài để có thể bộc lộ hết bản tính của mình. Sự thật, dòng chảy là hiện tượng nhiều nhân tố. Từng nhân tố dòng chảy đến lượt mình lại là hàm của vô vàn các nhân tố khác mà quy luật biến đổi của chúng con người chưa mô tả được. Do vậy, trong kết cục cuối cùng, tổng hợp của vô vàn các mối quan hệ tương hỗ phức tạp, dòng chảy biểu hiện là một hiện tượng ngẫu nhiên. Do tính ngẫu nhiên được thể hiện nhiều nhất ở dòng chảy năm và điều tiết nhiều năm dòng chảy, lớp mô hình này hoàn toàn không đánh giá được khả năng phát sinh cùng những diễn biến động lực của quá trình, mà chủ yếu là sản sinh ra những thể hiện mới đầy đủ hơn của một quá trình ngẫu nhiên.

Các mô hình ngẫu nhiên (các phương pháp phân tích chuỗi) thường được sử dụng để mô tả dao động dòng chảy tháng, mùa, năm và nhiều năm của các lưu vực vừa và lớn, ít sử dụng trong dự báo hạn ngắn.

b) Mô hình tất định

Quá trình hình thành dòng chảy - xác định hoặc nhận thức được qua đánh giá, phân tích bản chất vật lý của hiện tượng;

Mô hình tất định là mô hình mô phỏng quá trình hình thành dòng chảy, trong đó chủ yếu là dòng chảy lũ trên lưu vực;

Mô hình tất định được ứng dụng vào nhận định, nhận dạng lũ, tính toán, dự báo dòng chảy lũ hạn ngắn và hạn vừa.

- **Mô hình tất định phân bố**

Mô hình mô phỏng quá trình hình thành dòng chảy bằng những phương trình vi phân đạo hàm riêng chứa các hệ số thay đổi theo không gian và thời gian.

Các đặc trưng địa hình, thủy địa mạo lưu vực đóng vai trò các thông số của phương trình, quá trình mưa là điều kiện biên, trạng thái ban đầu của lưu vực là các điều kiện ban đầu, đều được mô phỏng bằng những phương trình vi phân với các hệ số thay đổi theo không gian.

Mô hình thông số phân bố đòi hỏi đo đạc chi tiết về địa hình, địa mạo, sử dụng đất lưu vực, mưa, dòng chảy,... theo không gian (dưới dạng DEM) để xác định các thông số khi giải hệ phương trình mô phỏng quá trình dòng chảy.

Hiện có một số mô hình kiểu này đang được nghiên cứu tính toán dự báo lũ hạn ngắn như mô hình MARINE (Pháp), TOPMODEL (Anh), mô hình USDAHL, mô hình WETSPA (Bi),...

- **Mô hình tất định tập trung**

Mô hình mô phỏng quá trình hình thành dòng chảy bằng những phương trình vi phân đạo hàm riêng hệ số không đổi theo không gian.

Các đặc trưng địa hình, địa mạo lưu vực,... đóng vai trò các thông số của phương trình và được xem là thông số trung bình cho toàn lưu vực, có thể xem như gắn với một điểm là trọng tâm của lưu vực.

Trong các mô hình tất định, mô hình thông số tập trung là mô hình có ít thông số nhất, dễ sử dụng và được ứng dụng rộng rãi nhất.

Hiện nay, mô hình này được sử dụng rộng rãi và rất hiệu quả trong tính toán và dự báo dòng chảy lũ trên các hệ thống sông ở Việt Nam.

Mô hình thông số tập trung thường được chia làm hai loại: Mô hình "*hộp đen*" và mô hình "*quan niệm*".

- *Mô hình hộp đen*

- + “Hộp đen” - không biết về cấu trúc, về mối quan hệ giữa các nhân tố tác động và về các thông số của nó.
- + Hệ thống kiểu này thường chỉ có số liệu đo đạc ở các đầu vào (lượng vào) và đầu ra (lượng ra) của hệ.
- + Hệ thống thủy văn có thể xem là hệ thống tuyến tính, chúng thoả mãn nguyên lý "cộng dồn".
- + Mô hình loại này có thể là các phương pháp: Đường đơn vị, mô hình diễn toán lũ trên đoạn sông kiểu Kalinin-Miliukov, mạng thần kinh nhân tạo (ANN), phương pháp hồi quy,...

- *Mô hình quan niệm*

- + Mô hình quan niệm là loại mô hình mô tả quá trình hình thành dòng chảy bằng phương pháp toán học dựa trên những hiểu biết, những kiến thức của con người về các quá trình đó.
- + Mô hình quan niệm được sử dụng rộng rãi trong việc mô tả quá trình hình thành dòng chảy từ mưa với hàng loạt các quá trình thành phần: Mưa, bốc hơi, điện trường, giữ nước ở thảm thực vật, xói mòn đất, thấm, tạo nước chảy trên mặt đất, nước chảy trong tầng thổ nhưỡng, nước chảy

trong các tầng ngầm,..., cuối cùng là tập trung nước vào mạng lưới sông, vận chuyển của sóng lũ trong sông,...

- + Hiện nay, trong tính toán và dự báo lũ, có một số mô hình quan niệm rất tiên tiến như mô hình SSARR (Mỹ), TANK (Nhật), bộ mô hình MIKE (Đan Mạch), STANFORD (Mỹ), SACRAMENTO (Mỹ),...

c) Mô hình tất định – ngẫu nhiên

Trong những năm gần đây đã xuất hiện những xu hướng liên kết cách tiếp cận tất định và ngẫu nhiên vào việc mô tả các hiện tượng thủy văn. Việc xét tính ngẫu nhiên của các quá trình trong mô hình tất định diễn ra theo 3 phương hướng:

- Xét sai số tính toán như một quá trình ngẫu nhiên và trở thành một thành phần trong các mô hình tất định.
- Sử dụng các mô tả xác suất - thống kê (luật phân bố) của các tác động khí tượng - thủy văn với tư cách là hàm vào của mô hình tất định.
- Xét các quy luật phân bố xác suất theo không gian của tác động khí tượng - thủy văn vào lưu vực. Với những ý tưởng này đã hình thành những mô hình động lực - ngẫu nhiên.

1.1.3 Tiêu chuẩn lựa chọn mô hình

Hiện có rất nhiều mô hình thủy văn được xây dựng và phát triển dựa trên các lý thuyết và nền tảng khác nhau. Những mô hình này có những điểm mạnh và những hạn chế khác nhau. Do đó, để lựa chọn được mô hình thích hợp cho mô phỏng chế độ thủy văn lưu vực, chúng ta cần căn cứ vào những tiêu chí nhất định. Do quá trình thủy văn lưu vực phụ thuộc vào nhiều yếu tố khác nhau nên các tiêu chí dùng để lựa chọn mô hình hình thủy văn cũng phụ thuộc vào rất nhiều yếu tố. Tuy nhiên các tiêu chí này cơ bản dựa vào mục đích nghiên cứu, tình hình thực tế khu vực nghiên cứu và thực trạng dữ liệu. Sau khi tổng kết nhiều nghiên cứu với các đặc trưng lưu vực và mô hình khác nhau, Cunderlik [9] cho rằng, các tiêu chí lựa chọn mô hình tập trung vào bốn vấn đề cơ bản sau:

- Mục đích của mô phỏng;
- Khả năng diễn tả các thành tố của chế độ thủy văn lưu vực vào mô hình;
- Khả năng tương thích với cơ sở dữ liệu sẵn có;
- “Chi phí” của mô phỏng.

Những tiêu chí trên cũng được tổ chức khí tượng thế giới WMO, cũng như là những chuyên gia đầu ngành về thủy văn, thống nhất đưa ra để làm căn cứ để so sánh

lựa chọn mô hình cho từng lưu vực sông. Các tiêu chí lựa chọn mô hình của WMO được liệt kê trong bảng 1.1 dưới đây.

Bảng 1.1: Bảng tiêu chuẩn lựa chọn mô hình của WMO

Loại mô hình	Yếu tố 1	Diện tích lưu vực			
			Lưu vực nhỏ	Lưu vực trung bình	Lưu vực lớn
		Loại mô hình	Mô hình tập trung	Mô hình bán phân bố	Mô hình phân bố
	Yếu tố 2	Định tuyến dòng chảy	Hầu như không cần	Mô phỏng thủy lực / thủy văn	Mô phỏng thủy lực / thủy văn, tương quan giữa các trạm đo
		Hình thái lưu vực			
			Đồng bằng, bằng phẳng	Trung du	Miền núi, địa hình dốc
Tính năng của mô hình	Yếu tố 3	Loại mô hình	Mô hình tập trung	Mô hình bán phân bố	Mô hình phân bố
		Ảnh hưởng của độ ẩm đất tới sự hình thành dòng chảy			
			Không ảnh hưởng	Ảnh hưởng vừa phải	Ảnh hưởng lớn
	Yếu tố 4	Mô phỏng nước dưới đất	Không yêu cầu	Khuyến nghị lựa chọn	Bắt buộc phải có
		Ảnh hưởng của tuyết tan đến dòng chảy			
			Không ảnh hưởng	Ảnh hưởng vừa phải	Ảnh hưởng lớn
	Yếu tố 5	Mô phỏng tuyết tan	Không yêu cầu	Khuyến nghị lựa chọn	Bắt buộc phải có
		Ảnh hưởng của các công trình (hồ chứa, đập dâng) đến dòng chảy			
			Không ảnh hưởng	Ảnh hưởng vừa phải	Ảnh hưởng lớn

		Mô phỏng trữ nước	Không yêu cầu	Khuyến nghị lựa chọn	Bắt buộc phải có			
Yêu cầu về số liệu	Yếu tố 6	Hình thức mưa ảnh hưởng tới lũ						
			Mưa theo mùa	Mưa bình lưu	Mưa đối lưu			
		Dữ liệu mưa khuyến nghị	Mưa ngày	Mưa ngày / mưa giờ	Mưa giờ / mưa thời đoạn nhỏ			
	Yếu tố 7	Thời đoạn mô phỏng						
			Ngắn		Trung bình		Dài	
		Diện tích lưu vực	Nhỏ	Lớn	Nhỏ	Lớn	Nhỏ	Lớn
		Dữ liệu mưa khuyến nghị	Mưa dự báo, rada hiện tại	Mưa thực đo	Mưa dự báo, rada hiện tại	Mưa dự báo và mưa đo	Mưa hiện tại và/hoặc dự báo từ mô hình địa phương	Mưa dự báo và mưa đo
	Yếu tố liên quan khác	Yếu tố 8	Sự sẵn có của dữ liệu phân bố về sử dụng đất và khí tượng					
				Không		Có		
			Loại mô hình	Khuyến nghị sử dụng mô hình tập trung	Khuyến nghị sử dụng mô hình Bán phân bố / Phân bố			
Yếu tố 9		Khả năng mô phỏng						
			Thấp		Trung bình		Cao	
		Loại mô hình	Mô hình đơn giản	Mô hình tập trung / mô hình kiểu hộp đen	Tất cả các tùy chọn sẵn có			

1.1.4 Một số mô hình thủy văn phổ biến

a) Mô hình TANK

Mô hình Tank là loại mô hình tất định, nhận thức. Mô hình do M.Surawara đưa ra từ năm 1956 tại Trung tâm Quốc gia Phòng chống lũ lụt Nhật và tới nay đã được tác

giả cải tiến nhiều lần, được thế giới công nhận là một trong số những mô hình ứng dụng có kết quả tốt.

Lưu vực được diễn tả như một chuỗi các bể chứa sắp xếp theo 2 phương thẳng đứng và nằm ngang. Giả thiết cơ bản của mô hình là dòng chảy cũng như dòng thấm và các hàm số của lượng nước trữ trong các tầng đất. Mô hình có hai dạng cấu trúc đơn và kép.

Do cấu trúc phi tuyến với các bể chứa sắp xếp theo chiều thẳng đứng, chưa có phương pháp toán học hữu hiệu nào để xác định các thông số của mô hình TANK, cách duy nhất là thử sai, tuy nhiên với mô hình TANK đơn có 24 thông số, phương pháp thử sai sẽ gây lúng túng và khó khăn cho người sử dụng do có những thông số cỡ phần nghìn (0.001). Hiện nay, một số phương pháp tối ưu hoá được sử dụng trên máy tính để tự động hoá quá trình thử sai này nhằm tìm bộ thông số tối ưu.

b) Mô hình SSARR

Mô hình SSARR (Streamflow Synthesis and Reservoir Regulation) tạm dịch là mô hình tổng hợp dòng chảy từ mưa và điều tiết hồ chứa, là mô hình có thể sử dụng cho các lưu vực có diện tích tương đối lớn.

Mô hình SSARR do Rocwood đề xuất năm 1956 và liên tục được hoàn thiện. Lúc đầu mô hình SSARR chỉ dùng để dự báo lũ và vận hành hồ chứa, sau được mở rộng phục vụ cho thiết kế hồ chứa trong hệ thống sông, mô hình được dùng cho vùng có lũ do mưa hoặc do tuyết tan. Khi xây dựng mô hình này người ta quan niệm rằng hệ thống sông ngòi dù phức tạp cũng chỉ gồm các thành phần cơ bản sau:

- Các lưu vực sông nhỏ,
- Các hồ chứa tự nhiên và nhân tạo,
- Các đoạn sông.

Do đó người ta xây dựng mô hình toán học cho từng loại, sau cùng tập hợp lại ta sẽ có mô hình toán học của cả hệ thống sông. Các mô hình toán học thành phần đều sử dụng hai phương trình cơ bản là phương trình liên tục và phương trình trữ lượng.

Các thông số được lựa chọn giá trị tối ưu thông qua việc tính thử dần sao cho sai khác giữa lưu lượng thực đo và lưu lượng tính toán theo mô hình là nhỏ nhất.

Cho tới nay việc điều chỉnh thông số của mô hình SSARR vẫn làm theo cách thử dần do đó phụ thuộc nhiều vào kinh nghiệm của người sử dụng mô hình.

c) Mô hình lũ đơn vị HEC-HMS

Mô hình HEC-HMS được nâng cấp từ HEC-1 công bố vào năm 2000. HEC-HMS sử dụng tài liệu mưa để tính toán quá trình mưa rào - dòng chảy trên một lưu vực cụ

thể. Chức năng các thành phần mô hình dựa trên các mối quan hệ toán học đơn giản mà các quan hệ này có xu hướng biểu thị các quá trình khí tượng, thủy văn, thủy lực bao gồm quá trình mưa rào - dòng chảy. Những quá trình này được phân ra thành mưa, tích đọng, thấm, chuyển lượng mưa hiệu quả thành dòng chảy của lưu vực, cung cấp nước cho dòng chảy cơ bản và diễn toán lũ. Kết quả tính toán dùng cho dự báo lũ hoặc đầu vào cho mô hình thủy lực. Mô hình HMS là mô hình có ít tham số và dễ sử dụng, không yêu cầu cao về tài liệu địa hình lưu vực, độ chính xác của mô hình cũng đã được kiểm nghiệm đối với các lưu vực từ 15 đến 1500km².

d) Mô hình Nielsen-Hansen

Mô hình Nielsen-Hansen, được công bố trong tạp chí “Nordic Hydrology” năm 1973 và sau này được Viện Thủy lực Đan Mạch phát triển và đổi thành NAM (từ 3 từ viết tắt tiếng Đan Mạch của mô hình mưa – dòng chảy). Mô hình gồm 4 bể chứa, nguyên lý tính toán trong mỗi bể chứa là giải phương trình cân bằng nước. Điều khác biệt so với mô hình TANK là dòng chảy từ các bể chứa vào sông, tính theo mô hình TANK là theo quy luật tuyến tính còn tính theo mô hình Nielsen - Hansen là theo quy luật phi tính (dạng đường cong nước rút).

Chi tiết các bước tính toán mô phỏng dòng chảy được Nielsen - Hansen trình bày cụ thể trong mục 1.2.

e) Mô hình MARRINE

Mô hình Marine - dựa trên phương trình Saint-Vernant, tính toán dự báo quá trình lưu lượng tại các tuyến hạ lưu. Mô hình MARINE đòi hỏi phải có số liệu địa hình, thổ nhưỡng, lớp phủ, mạng lưới trạm KTTV đủ dày, đặc biệt phải dự báo được mưa với độ phân giải cao.

Lưu vực được chia theo lưới ô vuông. Phương trình liên tục được sử dụng để tính giá trị mực nước trong mỗi ô. Tốc độ dòng chảy mặt trên lưu vực được xác định bằng phương trình sóng khuếch tán (bỏ thành phần gia tốc địa phương và gia tốc đối lưu chỉ giữ lại thành phần áp lực, trọng lực, ma sát trong phương trình Saint-Venant). Hệ số thấm là hàm số phụ thuộc vào mực nước trong từng ô.

Dòng chảy trong sông được xác định bằng hệ phương trình Saint-Vernant đầy đủ với các hàm Q gia nhập khu giữa được xác định từ các lưu vực liền kề.

f) Mô hình WETSPA

WetSpa (Water and Energy Transfer between Soil, Plant and Atmosphere) là một mô hình thủy văn phân bố mô tả quy luật vận động của nước trong tự nhiên dùng cho dự báo trao đổi nước và nhiệt giữa đất, thảm phủ thực vật, khí quyển trong phạm vi

một vùng, một lưu vực. Mô hình được Wang và nhóm cộng sự phát triển đầu tiên năm 1996.

Quá trình thủy văn được xem xét và mô hình hoá trong các tầng khí quyển, lớp tán lá của thảm phủ, đới rễ cây, tầng chuyển tiếp và tầng bão hoà. Mưa rơi từ khí quyển trước khi xuống mặt đất bị giữ lại bởi lượng ngưng tụ trên lá cây. Phần mưa còn lại rơi xuống mặt đất được chia thành hai phần phụ thuộc vào thảm phủ, loại đất, độ dốc, cường độ mưa và độ ẩm kì trước của đất. Thành phần đầu tiên làm đầy các vùng trũng trên mặt đất và đồng thời chảy tràn trên mặt đất trong khi phần còn lại ngấm vào đất. Phần mưa ngấm đó có thể giữ lại ở đới rễ cây, chảy sát mặt hay thấm sâu hơn xuống tầng nước ngầm, chúng phụ thuộc vào độ ẩm của đất. Nước tích tụ từ một ô lưới bất kì chảy sát mặt phụ thuộc vào lượng trữ nước ngầm và hệ số triết giảm. Dòng thấm từ lớp đất được giả định cung cấp cho lượng nước ngầm. Dòng chảy sát mặt trong đới rễ cây được giả định đóng góp vào dòng chảy tràn và diễn toán ra cửa ra của lưu vực cùng với dòng chảy tràn.

Tổng lượng dòng chảy ra từ mỗi ô lưới là tổng lượng dòng chảy tràn, sát mặt và dòng ngầm. Bốc thoát hơi diễn ra từ thực vật qua hệ thống rễ cây ở trong lớp đất và một phần nhỏ từ lượng nước ngầm. Cân bằng nước đối với lượng ngưng tụ gồm có mưa, bốc hơi và dòng chảy. Cân bằng nước cho các vùng trũng gồm có lượng mưa rơi, thấm, bốc hơi và chảy tràn. Cân bằng nước cho tầng đất gồm ngấm, bốc thoát hơi, thấm và chảy sát mặt. Cân bằng nước cho lượng nước ngầm gồm lượng cung cấp cho nước ngầm, bốc thoát hơi từ tầng sâu và dòng chảy sát dòng ngầm.

g) Mô hình DIMOSOP

DIMOSOP (DIstributed hydrological Model for the Special Observing Period) là mô hình toán thủy văn phân bố được sử dụng nhiều ở một số nước Châu Âu như Italia, Pháp, Thụy Sĩ.

Đặc điểm nổi bật của mô hình này là có thể sử dụng dữ liệu dạng điểm của các trạm đo mưa trong lưu vực hoặc sử dụng kết quả dự báo dưới dạng ô lưới (grid) là đầu ra của các mô hình dự báo thời tiết như MM5 và BOLAM để dự báo lũ.

Ở Việt Nam, mô hình DIMOSOP được chuyển giao cho trường Đại học Thủy lợi trong khuôn khổ dự án hợp tác song phương giữa Chính phủ Việt Nam và Chính phủ Italia trong việc xây dựng mô hình hỗ trợ dự báo lũ trung hạn cho hệ thống sông Hồng và sông Thái Bình. Khi chuyển giao cho Việt Nam, mô hình mang tên mới là DIMOSHONG. Cấu trúc chính của mô hình dựa trên quan điểm chia lưu vực tính toán ra thành một hệ thống các ô lưới. Mỗi một ô lưới trên lưu vực đều được đặc trưng bởi một yếu tố thủy văn nào đó, có thể là một phần tử của lưu vực, có thể là một phần tử

của sông, hay là một phần tử của hồ chứa. Nếu ô lưới là một phần tử của lưu vực thì mô hình sẽ tính toán ra dòng chảy từ mưa thông qua mô phỏng các quá trình vật lý trên lưu vực như mưa, tổn thất, tập trung dòng chảy trên sườn dốc cho ô lưới đó, nếu ô lưới là một phần tử của đoạn sông, mô hình sẽ mô phỏng quá trình diễn toán dòng chảy trên sông cho ô lưới, còn nếu ô lưới là một phần tử của hồ chứa, mô hình sẽ mô phỏng quá trình diễn toán dòng chảy qua hồ chứa cho ô lưới này.

Phương pháp tính dòng chảy từ mưa trong DIMOSOP dựa trên các lý thuyết: Phương pháp tính tổn thất: SCS-CN; đường quá trình SCS tính mưa hiệu quả, diễn toán trong sông Muskingum,... Kết quả tính toán của mô hình chính là lưu lượng hay mực nước lũ tại bất kỳ một ô lưới nào (điểm) trên lưu vực đã được đăng ký trước.

1.2 Mô hình thủy văn thông số tập trung Nielsen-Hansen

Dựa trên bảng tiêu chuẩn lựa chọn mô hình của WMO cùng với tình hình thực tế (số liệu thu thập được, năng lực tính toán mô phỏng, ...), chúng tôi lựa chọn mô hình Nielsen-Hansen làm cơ sở lý thuyết để xây dựng bộ phần mềm thủy văn dự báo cho các lưu vực nghiên cứu.

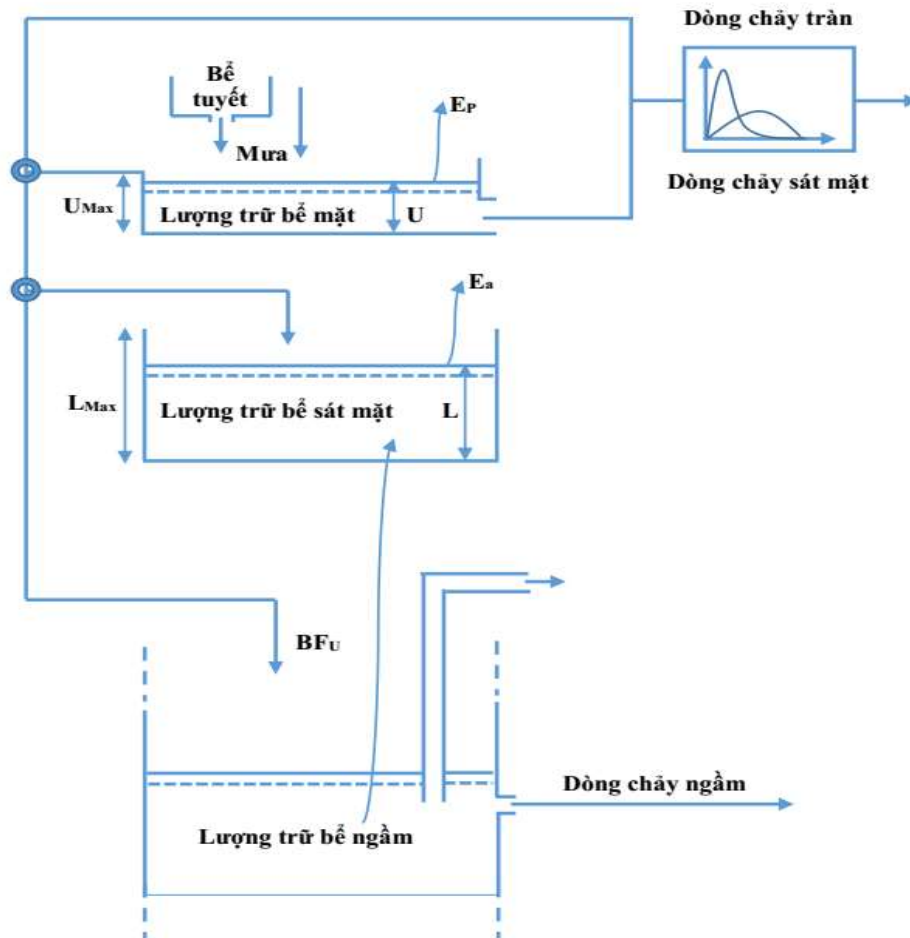
1.2.1 Giới thiệu

Mô hình Nielsen-Hansen là mô hình thủy văn mô phỏng quá trình mưa – dòng chảy diễn ra trên lưu vực. Là một mô hình toán thủy văn, mô hình Nielsen-Hansen bao gồm một tập hợp các biểu thức toán học đơn giản để mô phỏng các quá trình trong chu trình thủy văn. Mô hình Nielsen-Hansen mô phỏng quá trình mưa – dòng chảy một cách liên tục thông qua việc tính toán cân bằng nước ở bốn bể chứa thẳng đứng, có tác dụng qua lại lẫn nhau để diễn tả các tính chất vật lý của lưu vực. Các bể chứa đó gồm:

- Bể tuyết (chỉ áp dụng cho vùng có tuyết);
- Bể mặt;
- Bể sát mặt hay bể tầng rễ cây;
- Bể ngầm.

Dữ liệu đầu vào của mô hình là mưa, bốc hơi tiềm năng, và nhiệt độ (chỉ áp dụng cho vùng có tuyết). Kết quả đầu ra của mô hình là dòng chảy trên lưu vực, mực nước ngầm, và các thông tin khác trong chu trình thủy văn, như sự thay đổi tạm thời của độ

âm của đất và khả năng bổ sung nước ngầm. Dòng chảy lưu vực được phân một cách gần đúng thành dòng chảy mặt, dòng chảy sát mặt, dòng chảy ngầm.



Hình 1.2: Sơ đồ mô phỏng của mô hình Nielsen-Hansen

1.2.2 Cấu trúc của mô hình

a. Bể tuyết

Giáng thủy sẽ được giữ lại trong bể tuyết khi nhiệt độ dưới 0 độ C, còn nếu nhiệt độ lớn hơn 0 độ C thì nó sẽ chuyển xuống bể chứa mặt:

$$Q_{melt} = \begin{cases} CSNOW \cdot TEMP & \text{for } TEMP > 0 \\ 0 & \text{for } TEMP \leq 0 \end{cases}$$

Trong đó: CSNOW = 2 mm/day/K là hệ số tuyết tan trong ngày.

b. Bể chứa mặt

Lượng ẩm trữ trên bề mặt của thực vật, cũng như lượng nước điện trở trên bề mặt lưu vực được đặc trưng bởi lượng trữ bề mặt. U_{Max} đặc trưng cho giới hạn trữ nước tối đa của bể này.

Lượng nước, U - trong bể chứa mặt sẽ giảm dần do bốc hơi, do thất thoát theo phương nằm ngang (dòng chảy sát mặt). Khi lượng nước này vượt quá ngưỡng U_{\max} , thì một phần của lượng nước vượt ngưỡng, P_N này sẽ chảy vào suối dưới dạng dòng chảy tràn bề mặt, phần còn lại sẽ thấm xuống bề mặt đất và tầng ngầm.

c. Bể sát mặt hoặc bể tầng rễ cây

Bể này thuộc tầng rễ cây, là lớp đất mà thực vật có thể hút nước để thoát ẩm. L_{\max} đặc trưng cho lượng ẩm tối đa mà bể này có thể chứa.

Lượng ẩm của bể chứa này được đặc trưng bằng đại lượng L . L phụ thuộc vào lượng tổn thất thoát hơi của thực vật. Lượng ẩm này cũng ảnh hưởng đến lượng nước sẽ đi xuống bể chứa ngầm để bổ sung nước ngầm.

d. Bốc thoát hơi

Nhu cầu bốc thoát hơi nước trước tiên là để thỏa mãn tốc độ bốc thoát hơi tiềm năng của bể chứa mặt. Nếu lượng ẩm U trong bể chứa mặt nhỏ hơn nhu cầu này, thì nó sẽ lấy ẩm từ tầng rễ cây theo tốc độ E_a . E_a là tỷ lệ với lượng bốc thoát hơi tiềm năng E_p :

$$E_a = E_p \cdot L / L_{\max}$$

e. Dòng chảy mặt

Khi bể chứa mặt tràn nước, $U \geq U_{\max}$, thì lượng nước vượt ngưỡng P_N ($P_N = U - U_{\max}$) sẽ hình thành dòng chảy mặt và thấm xuống dưới. QOF là một phần của P_N , tham gia hình thành dòng chảy mặt, nó tỉ lệ thuận với P_N và thay đổi tuyến tính với lượng ẩm tương đối, L/L_{\max} , của tầng rễ cây:

$$QOF = \begin{cases} CQOF \frac{L/L_{\max} - TOF}{1 - TOF} & \text{for } L/L_{\max} > TOF \\ 0 & \text{for } L/L_{\max} \leq TOF \end{cases}$$

Trong đó $CQOF$ là hệ số dòng chảy mặt ($0 \leq CQOF \leq 1$).

TOF là ngưỡng của dòng chảy mặt ($0 \leq TOF \leq 1$).

Phần còn lại của P_N sẽ thấm xuống tầng dưới. Một phần DL của phần nước thấm xuống này, ($P_N - QOF$), sẽ làm tăng lượng ẩm L của bể chứa tầng rễ cây này. Phần còn lại sẽ thấm thấu xuống tầng sâu hơn để bổ sung cho bể chứa tầng ngầm.

f. Dòng chảy sát mặt

Dòng chảy sát mặt, QIF, được giả thiết tỉ lệ thuận với U và biến đổi tuyến tính với độ ẩm tương đối của bề chứa tầng rễ cây:

$$QIF = \begin{cases} (CKIF)^{-1} \frac{L/L_{max} - TIF}{1 - TIF} U & \text{for } L/L_{max} > TIF \\ 0 & \text{for } L/L_{max} \leq TIF \end{cases}$$

Trong đó CKIF là hằng số thời gian của dòng chảy sát mặt

TIF là giá trị ngưỡng của dòng chảy sát mặt ($0 \leq TIF \leq 1$)

g. Bỏ sung dòng chảy ngầm

Lượng nước thấm xuống G, bỏ sung cho bề chứa ngầm phụ thuộc vào độ ẩm của đất ở tầng rễ cây:

$$G = \begin{cases} (P_N - QOF) \frac{L/L_{max} - TG}{1 - TG} & \text{for } L/L_{max} > TG \\ 0 & \text{for } L/L_{max} \leq TG \end{cases}$$

h. Lượng ẩm của đất

Bề chứa tầng sát mặt biểu thị lượng nước có trong tầng rễ cây. Lượng mưa hiệu quả sau khi trừ đi lượng nước tạo dòng chảy mặt, lượng nước bỏ sung cho tầng ngầm, sẽ bỏ sung và làm tăng độ ẩm của đất ở tầng rễ cây L bằng một lượng DL.

$$DL = P_N - QOF - G$$

i. Diễn toán dòng chảy mặt và dòng chảy sát mặt

Dòng chảy mặt và dòng chảy sát mặt sẽ được diễn toán thông qua 2 bề chứa tuyến tính theo chuỗi thời gian với cùng một hằng số thời gian CK12.

k. Diễn toán dòng chảy ngầm

Dòng chảy ngầm được diễn toán thông qua một bề chứa tuyến tính với hằng số thời gian CKBF.

1.2.3 Các thông số của mô hình

Mô hình Nielsen-Hansen bao gồm 9 thông số cần được hiệu chỉnh (xem Bảng 1.2).

Bảng 1.2: Các thông số hiệu chỉnh của mô hình Nielsen-Hansen.

Thông số mô hình	Mô tả	Khoảng giá trị
L_{\max} [mm]	Lượng nước tối đa trong bể chứa tầng rễ cây. L_{\max} có thể gọi là lượng ẩm tối đa của tầng rễ cây để thực vật có thể hút để thoát hơi nước.	50 - 400
U_{\max} [mm]	Lượng nước tối đa trong bể chứa mặt. Lượng trữ này có thể gọi là lượng nước để điền trũng, rơi trên mặt thực vật, và chứa trong vài centimet của bề mặt của đất.	5 - 35
CQOF	Hệ số dòng chảy mặt ($0 \leq CQOF \leq 1$). CQOF quyết định sự phân phối của mưa hiệu quả cho dòng chảy ngầm và thấm.	0 - 1
TOF	Giá trị ngưỡng của dòng chảy mặt ($0 \leq TOF \leq 1$). Dòng chảy mặt chỉ hình thành khi lượng ẩm tương đối của đất ở tầng rễ cây lớn hơn TOF.	0 - 0.9
TIF	Giá trị ngưỡng của dòng chảy sát mặt ($0 \leq TOF \leq 1$). Dòng chảy sát mặt chỉ được hình thành khi chỉ số ẩm tương đối của tầng rễ cây lớn hơn TIF.	0 - 0.9
TG	Giá trị ngưỡng của lượng nước bổ sung cho dòng chảy ngầm ($0 \leq TOF \leq 1$). Lượng nước bổ sung cho bể chứa ngầm chỉ được hình thành khi chỉ số ẩm tương đối của tầng rễ cây lớn hơn TG.	0 - 0.9
CKIF [giờ]	Hằng số thời gian của dòng chảy sát mặt. CKIF cùng với U_{\max} quyết định dòng chảy sát mặt. Nó chi phối thông số diễn toán dòng chảy sát mặt $CKIF \gg CK12$.	200 - 2000
CK12 [giờ]	Hằng số thời gian cho diễn toán dòng chảy mặt và sát mặt. Dòng chảy mặt và dòng chảy sát mặt được diễn toán theo các bể chứa tuyến tính theo chuỗi với cùng một hằng số thời gian CK12.	3 - 72
CKBF [giờ]	Hằng số thời gian dòng chảy ngầm. Dòng chảy ngầm từ bể chứa ngầm được tạo ra sử dụng mô hình bể chứa tuyến tính với hằng số thời gian CKBF.	500 - 5000

CHƯƠNG II

PHƯƠNG PHÁP SỐ GIẢI BÀI TOÁN ƯỚC TÍNH THÔNG SỐ TỐI ƯU CHO MÔ HÌNH THỦY VĂN THÔNG SỐ TẬP TRUNG

Như vậy sau khi nhập các dữ liệu đầu vào là lượng mưa, thông qua mô hình Nielsen-Hansen với các thông số đã được xác định trước, ta có được chuỗi giá trị lưu lượng chảy ra. Cần hiệu chỉnh các thông số của mô hình sao cho giá trị lưu lượng tính toán này gần với các giá trị lưu lượng thực đo đã biết trước. Trải qua nhiều lần hiệu chỉnh cho nhiều bộ số liệu (nhiều năm), ta sẽ có bộ thông số phù hợp để sử dụng dự báo cho lưu vực.

Tuy nhiên, mỗi lưu vực gồm nhiều trạm mưa, mỗi trạm mưa gồm nhiều thông số nên tổng các thông số của mô hình cần hiệu chỉnh là rất lớn (ví dụ bài toán lưu vực Sơn La bao gồm 121 thông số cần hiệu chỉnh). Từ đó đặt ra yêu cầu cần sử dụng các phương pháp tự động để ước tính thông số tối ưu.

2.1 Tổng quan về các phương pháp giải bài toán ước tính thông số tối ưu

Xét hàm mục tiêu là hàm bình phương tối thiểu:

$$F = \frac{1}{n} \sum_{i=1}^n (Q_i - \hat{Q}_i)^2$$

Trong đó

Q_i : Giá trị lưu lượng tính toán $Q_i = G(R_j, TS_k)$: hàm của giá trị mưa đầu vào R_j và các thông số TS_k bị chặn trong khoảng cho trước;

\hat{Q}_i : Giá trị lưu lượng thực đo.

Như vậy để giá trị tính toán gần sát với các giá trị thực đo thì giá trị hàm F phải tiến về cực tiểu $F = 0$.

Bài toán được viết dưới dạng tổng quát như sau:

Cho tập D xác định bởi:

$$D = \{x \in R^n : h_i(x) = 0 ; i = 1, 2, \dots, m; g_j(x) \leq 0 ; j = 1, 2, \dots, p\}$$

Cho hàm số $f : D \rightarrow R$, hàm số $h_i : R^n \rightarrow R$, hàm số $g_j : R^n \rightarrow R$

Xét bài toán tối ưu với ràng buộc tổng quát dạng:

$$\min \{f(x) : x \in R^n\} \text{ với các điều kiện: } \begin{cases} h_i(x) = 0, i = 1, 2, \dots, m \\ g_j(x) \leq 0, j = 1, 2, \dots, p \end{cases}$$

Đối với các hàm F , tùy thuộc vào tính chất của nó mà người ta sử dụng các cách khác nhau tìm cực trị tương ứng.

Các phương pháp giải bài toán tối ưu hóa hiện tại có thể được chia thành ba nhóm: phương pháp tìm kiếm (Heuristic methods), phương pháp quy hoạch toán học (Mathematical programming models) và các thuật toán tối ưu dựa trên nền tảng của sự tiến hóa.

Nhóm phương pháp tìm kiếm (Phương pháp khung, phương pháp độ dốc - gradient ...) xuất phát từ một nghiệm ban đầu từ đó tìm kiếm các nghiệm cho giá trị hàm mục tiêu tốt hơn trên cơ sở phân tích các hàm ràng buộc. Nhóm này yêu cầu phải có hàm ràng buộc tường minh và thường chỉ cho nghiệm tối ưu cục bộ, tốc độ hội tụ phụ thuộc nhiều vào nghiệm ban đầu.

Nhóm phương pháp quy hoạch toán học (Phương pháp đồ thị, phương pháp đơn hình,...) phù hợp cho bài toán tối ưu tuyến tính, đối với bài toán tối ưu hóa phi tuyến nhóm này chủ yếu cho nghiệm tối ưu cục bộ. Cũng như nhóm phương pháp tìm kiếm, nhóm này yêu cầu phải có hàm ràng buộc tường minh.

Nhóm các thuật toán dựa trên nền tảng tiến hóa (giải thuật di truyền - GA, tiến hóa - DA, mô phỏng quá trình ủ - SA...) có ưu điểm không cần các hàm ràng buộc tường minh nhưng để tìm được nghiệm tối ưu toàn cục thì cần số lần lặp rất lớn, thông thường nhóm phương pháp này được dùng để lựa chọn nghiệm gần với nghiệm tối ưu và dùng nghiệm đó làm nghiệm ban đầu cho các phương pháp khác.

Tuy nhiên đối với hàm F ở trên, phụ thuộc vào hàm G – được xác định qua mô hình Nielsen-Hansen là một hàm nhiều biến mà nó không có đạo hàm bậc nhất, không có đạo hàm bậc hai, không lồi, không lõm, không DC, không đơn điệu, không thỏa mãn điều kiện Lipchitz, do đó các phương pháp giảm gradien, phương pháp Newton, phương pháp gradien liên hợp ... đều không áp dụng được. Khi đó ta cần sử dụng các phương pháp tối ưu không dùng đạo hàm, ví dụ như phương pháp dò tìm theo các tọa độ ô vuông, phương pháp Rosen Brock, phương pháp giảm theo các đơn hình Nelder – Mead, phương pháp Hook-Jeeves, giải thuật di truyền, phương pháp SCE (Shuffled Complex Evolution), ... Trong khuôn khổ đề tài này, chúng tôi thử nghiệm một vài thuật toán để đánh giá các ưu nhược điểm của chúng và khuyến cáo sử dụng cho phù hợp với nhiều lưu vực có các đặc trưng khác nhau.

2.2 Phương pháp ô vuông

Có nhiều phương pháp lựa chọn giá trị tối ưu của thông số nhưng dễ hiểu hơn cả và tính toán vất vả hơn cả là phương pháp ô vuông (phương pháp lưới). Giả sử mô hình có hai thông số là a và b . Hàm mục tiêu chọn theo luật bình phương tối thiểu dạng:

$$F = \frac{1}{n} \sum_{i=1}^n (Q_i - \hat{Q}_i)^2$$

Miền xác định của thông số a là từ 0 đến 0.9, miền xác định của thông số b là từ 200 đến 2000. Để dễ trình bày, ta coi thông số a biến thiên theo trục nằm ngang Ox, còn thông số b biến thiên theo trục thẳng đứng Oy. Chọn bước dò tìm trong đợt tính lặp thứ nhất là: 0.1 vậy ta cần dò tìm hai thông số a, b trong hình vuông mỗi cạnh có 10 giá trị, tương ứng có $10 \times 10 = 100$ cặp điểm. Với mỗi cặp điểm xác định một bộ thông số của mô hình. Thực hiện n lần tính toán theo mô hình để tìm ra 1 giá trị của hàm mục tiêu.

Với 100 cặp điểm ứng với 100 nút lưới ta tìm được 100 giá trị của hàm mục tiêu. Ghi các giá trị này vào các nút lưới tương ứng rồi vẽ đường đồng mức. Tìm miền có giá trị nhỏ nhất của hàm mục tiêu. Giả sử tìm thấy miền nhỏ nhất là:

$$0.3 < a < 0.4$$

$$800 < b < 1000$$

Chọn lại bước dò tìm trong đợt tính lặp thứ hai bằng 1 phần 10 bước dò tìm trong đợt tính lặp thứ nhất. Bây giờ tiếp tục dò tìm trong miền:

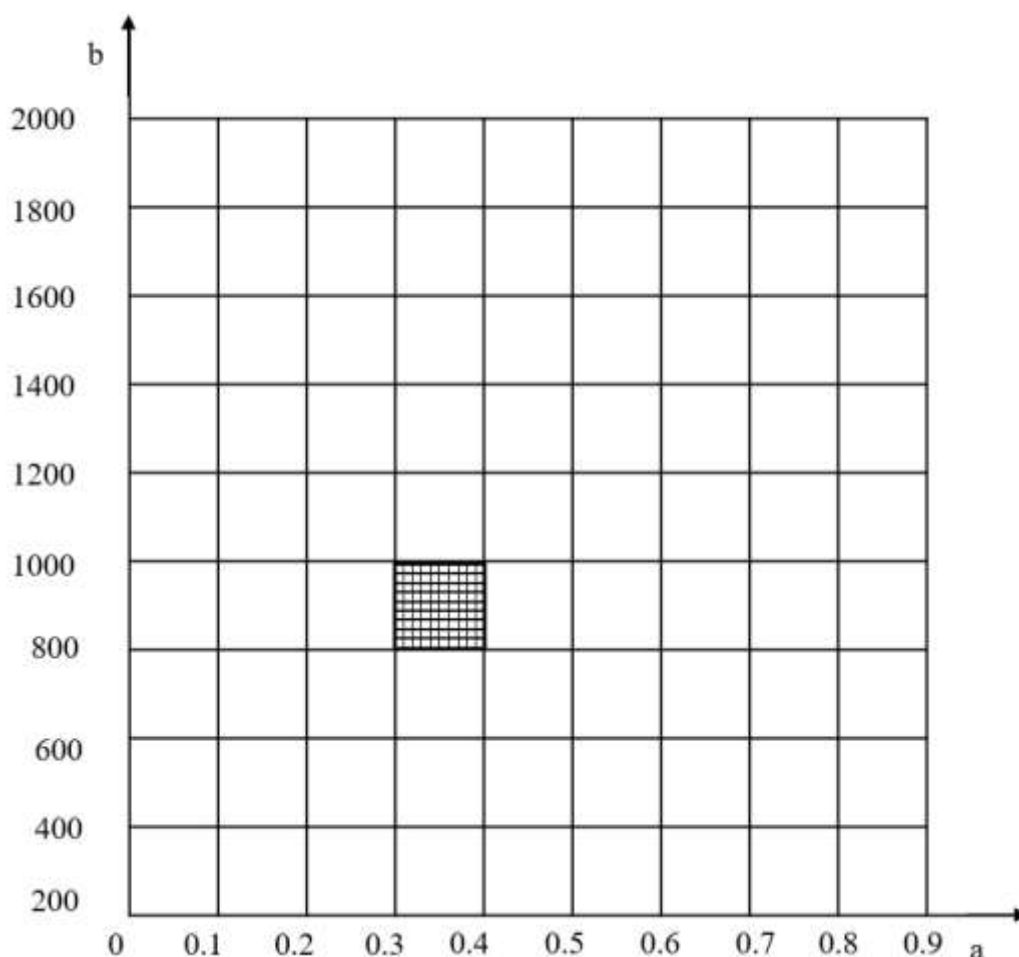
$$0.3 < a < 0.4 \text{ và } 800 < b < 1000$$

Bước dò tìm trong đợt tính lặp thứ hai là 0.01 nên phạm vi tìm kiếm hẹp hơn nhưng vẫn có đủ 100 cặp điểm. Các bước tiếp theo làm giống như đợt tính lặp thứ nhất. Kết quả tìm thấy miền nhỏ nhất là:

$$0.35 < a < 0.36$$

$$840 < b < 860$$

Chọn lại bước dò tìm trong đợt tính lặp thứ ba bằng 1 phần 10 bước dò tìm trong đợt tính lặp thứ hai. Quá trình cứ như thế tiếp diễn cho đến khi bước dò tìm nhỏ hơn một vô cùng bé chọn trước.



Hình 2.1: Mô phỏng phương pháp giải bằng ô vuông

Quá trình tính toán sẽ rất phức tạp và tốn nhiều thời gian nếu số thông số tăng lên. Chẳng hạn số thông số tăng từ 2 lên 3 thông số, số trường hợp phải tính thử cho 1 lần lặp tăng lên 10 lần: $10 \times 10 \times 10 = 1000$ cặp điểm.

Nếu áp dụng cách dò tìm này cho mô hình Nielsen-Hansen có 9 thông số, sẽ phải tính 10^9 trường hợp cho một lần tính lặp để chọn vùng cực tiểu của hàm mục tiêu. Mỗi trường hợp là một bộ thông số đưa ra lựa chọn để tìm giá trị của hàm mục tiêu, phải trải qua một số lần tính lặp mới chọn được bộ thông số tối ưu. Rõ ràng số phép tính quá lớn.

Cách dò tìm này vẫn có khả năng rơi vào cực trị địa phương, chẳng hạn sau lần tính lặp thứ nhất, xuất hiện hai miền cách biệt có giá trị hàm mục tiêu nhỏ hơn xung quanh gọi là miền 1 và miền 2. Miền 1 có giá trị hàm mục tiêu nhỏ hơn miền 2, theo cách dò tìm trên ta tiếp tục chia nhỏ bước dò tìm trong miền 1 mà bỏ qua miền 2, nhưng rất có thể nếu tiếp tục chia nhỏ bước dò tìm trong miền 2 ta sẽ tìm thấy nút mới có giá trị hàm mục tiêu nhỏ hơn trong miền 1. Nói cách khác phương pháp này chỉ cho phép phát hiện cực trị địa phương đủ rộng hơn mắt lưới, mặc dù đã chấp nhận khối lượng tính toán khổng lồ nhưng vẫn có khả năng rơi vào cực trị địa phương.

2.3 Phương pháp Rosenbrock

Phương pháp Rosenbrock công bố năm 1960 và đang được ứng dụng rộng rãi trong nhiều ngành khoa học khác nhau, có sử dụng mô hình. Toán học và tính đúng đắn. Phương pháp Rosenbrock là bước phát triển của phương pháp độ dốc, rất thích hợp với dạng hàm mục tiêu không tính được đạo hàm riêng phần.

Xét hàm mục tiêu phụ thuộc N thông số:

$$F = f(x_1, x_2, x_3, \dots, x_N)$$

Lần tính lặp thứ nhất: Chọn sơ bộ mỗi thông số một giá trị hợp lý nào đó, tính giá trị của hàm mục tiêu ứng với các giá trị của thông số được chọn lần đầu:

$$F = (X_1^0, X_2^0, \dots, X_N^0) = F(X^0)$$

Chỉ thay đổi giá trị của một thông số, giữ nguyên tất cả giá trị của các thông số còn lại, tính giá trị mới của hàm mục tiêu để xem xét phản ứng của hàm mục tiêu với sự thay đổi này thông qua kết quả tính thử:

$$\Delta F_0 = F(X_1^0, X_2^0, \dots, X_N^0) - F(X_1^0, X_2^0, \dots, X_N^0)$$

Trong đó $X_1^1 = X_1^0 + \lambda_1^1$

Nếu ΔF^0 nhỏ hơn không, phép thử có lợi vì đã chọn được vị trí mới của hàm mục tiêu có giá trị nhỏ hơn tại vị trí cũ, chọn ngay giá trị của thông số vừa tính thử làm giá trị chính thức:

$$x_1^1 = x_1^0 + \lambda_1^1$$

và chọn bước thay đổi thông số lần thứ nhất làm bước thay đổi thông số lần thứ hai:

$$\lambda_1^2 = \alpha \cdot \lambda_1^1$$

với α là hằng số lớn hơn 1 thường chọn $\alpha = 3$

Nếu ΔF^0 lớn hơn không, phép thử không thành công vì hàm mục tiêu không tiến thêm về phía cực tiểu, chứng tỏ ta đã tìm thấy điểm dừng của thông số nên giữ nguyên giá trị thông số trước lần thử thứ nhất:

$$X_1^1 = X_1^0$$

và bước thay đổi thông số lần thứ hai cũng được chọn giảm đi so với bước thay đổi thông số lần thứ nhất:

$$\lambda_1^2 = \beta \cdot \lambda_1^1$$

với β là hằng số nhỏ hơn 1 thường chọn $\beta = 0,5$

Quá trình được lặp lại cho đến khi tất cả N thông số đã thử xong và nhận giá trị mới x_i^1 thay cho giá trị cũ x_i^0 .

Lần tính lặp thứ hai: Tính biểu thức.

$$\Delta F^1 = F(X_1^1, X_2^1, X_3^1, \dots, X_N^1) - F(X_1^0, X_2^0, X_3^0, \dots)$$

Tiếp tục lặp lại các phép thử như lần tính lặp thứ nhất cho lần tính lặp thứ hai, thứ ba, ... mỗi khi thử không thành công bước thay đổi thông số giảm đi một nửa. Quá trình tính lặp dừng lại khi tất cả các thông số đều có bước thay đổi thông số nhỏ hơn một vô cùng bé ε chọn trước:

$$\lambda_1^2 < \varepsilon$$

Trường hợp ngược lại bước thay đổi thông số chưa nhỏ hơn một vô cùng bé ε chọn trước mà tất cả N phép thử ở lần tính lặp thứ $k+m$ đều không thành công, chúng ta véc tơ dò tìm đã rơi vào cực trị địa phương, cần tiến hành xoay trục tọa độ theo quy định riêng của phương pháp Rosenbrock, những quy định này được trình bày dưới dạng tổng quát sau:

Xét hàm mục tiêu phụ thuộc N thông số:

$$F = f(x_1, x_2, x_3, \dots, x_N)$$

tới lần tính lặp thứ k, các thông số đã nhận các giá trị tương ứng với véc tơ N chiều:

$$\vec{X}^k = \begin{bmatrix} x_1^k \\ x_2^k \\ x_3^k \\ \vdots \\ x_j^k \\ \vdots \\ x_N^k \end{bmatrix}$$

Giá trị của hàm mục tiêu ứng với các giá trị của thông số tới lần tính lặp thứ k là:

$$F(x_1^k, x_2^k, x_3^k, \dots, x_N^k) = F(X^k)$$

Chỉ thay đổi giá trị của một thông số, giữ nguyên tất cả giá trị của các thông số còn lại, tính giá trị mới của hàm mục tiêu để xem xét phản ứng của hàm mục tiêu với sự thay đổi này thông qua kết quả tính thử. Giả sử thông số thứ j được thay đổi theo quy luật sau:

$$x_j^{k+1} = x_j^k + \lambda_j^k \cdot D_j$$

Trong đó:

λ_j^k là hệ số hằng số thay đổi theo lần tính lặp thứ k và thông số thứ j ,

D_j là véc tơ đơn vị chuẩn hóa theo phương thông số x_j và chỉ số j chạy từ 1 đến N .

$$\vec{D}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}; \vec{D}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}; \vec{D}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}; \dots \dots \dots; \vec{D}_N = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Tính $\Delta F_j = F(X_j^{k+1}) - F(X^k)$ với:

$$\vec{x}_j^{k+1} = \begin{bmatrix} x_1^k \\ x_2^k \\ x_3^k \\ \vdots \\ \vdots \\ x_j^k \\ \vdots \\ x_N^k \end{bmatrix} + \lambda_j^k \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} x_1^k \\ x_2^k \\ x_3^k \\ \vdots \\ \vdots \\ x_j^k + \lambda_j^k \\ \vdots \\ x_N^k \end{bmatrix}; \vec{x}_j^k = \begin{bmatrix} x_1^k \\ x_2^k \\ x_3^k \\ \vdots \\ \vdots \\ x_j^k \\ \vdots \\ x_N^k \end{bmatrix}$$

Khi $\Delta F_j < 0$, phép thử thành công vì đã chọn được vị trí mới của hàm mục tiêu có giá trị nhỏ hơn tại vị trí cũ, chọn:

$$X_j^{k+1} = X_j^k + \lambda_j^k \cdot D_j$$

và chọn bước thay đổi thông số thứ j lần thứ sau bằng bước thay đổi thông số thứ j lần trước nhân với hệ số lớn hơn 1:

$$\lambda_j^{k+1} = \alpha \lambda_j^k \text{ với } \alpha \text{ là hằng số lớn hơn 1 thường chọn } \alpha = 3.$$

Nếu $\Delta F_j > 0$, phép thử không thành công vì hàm mục tiêu không tiến thêm về phía cực tiểu, chứng tỏ ta đã tìm thấy điểm dừng của thông số nên giữ nguyên giá trị thông số trước lần thử thứ nhất:

$$X_j^{k+1} = X_j^k$$

và bước thay đổi thông số lần sau cũng được chọn giảm đi so với bước thay đổi thông số lần trước:

$$\lambda_j^{k+1} = \beta \lambda_j^k \text{ với } \beta \text{ là hằng số } < 1, \text{ thường chọn } \beta = 0,5$$

Quá trình được lặp lại cho đến khi tất cả N thông số đã thử xong và nhận giá trị mới x_i^{k+1} thay cho giá trị cũ x_j^k . Khi tất cả N thông số đã thử xong lần tính lặp thứ k+1, ta lại tính lặp lần thứ k+2. Quá trình tính lặp dừng lại khi tất cả các thông số đều có bước thay đổi thông số nhỏ hơn một vô cùng bé ε chọn trước:

$$\lambda_j^n < \varepsilon$$

Trường hợp ngược lại bước thay đổi thông số chưa nhỏ hơn một vô cùng bé ε chọn trước mà tất cả N phép thử ở lần tính lặp thứ k+m đều không thành công, chứng tỏ véc tơ dò tìm đã rơi vào cực trị địa phương, cần tiến hành xoay trục tọa độ theo quy định riêng của phương pháp Rosenbrock.

Chọn :

$$D_i^{k+m} = \frac{B_i^{k+m-1}}{|B_i^{k+m-1}|}$$

Với

$$B_i^k = A_i^k - \sum_{j=1}^i [(A_j^k)^T \cdot D_i^{k+1}] \cdot D_i^{k+1}$$

$$A_1^k = C_1^k \cdot D_1^k + C_2^k \cdot D_2^k + \dots + C_N^k \cdot D_N^k$$

$$A_2^k = 0 + C_2^k \cdot D_2^k + \dots + C_N^k \cdot D_N^k$$

$$A_3^k = 0 + 0 + \dots + C_N^k \cdot D_N^k$$

.....

$$A_N^k = 0 + 0 + \dots + 0 + C_N^k \cdot D_N^k$$

C_i^k là tổng đại số của tất cả các số hạng λ_i^k khi phép thử thành công.

$$C_i^k = \sum_{j=1}^p \lambda_{i,j}^k$$

$j = 1, 2, 3 \dots p$, là số lần phép thử thành công.

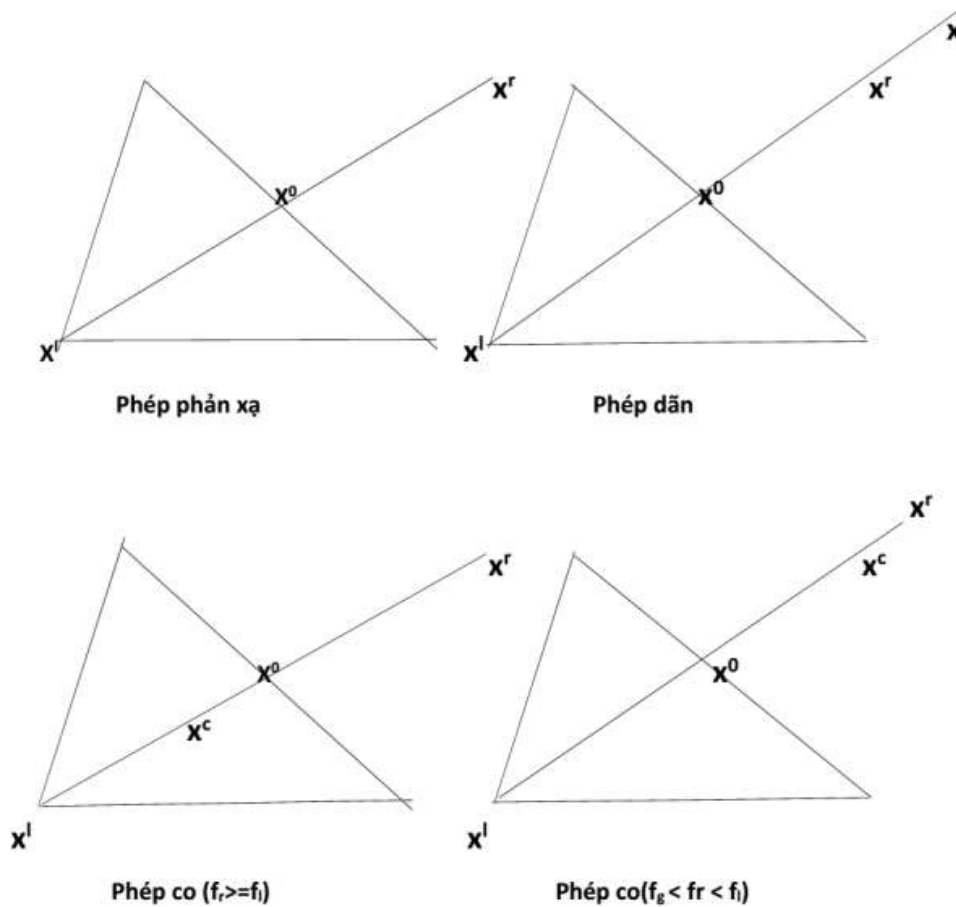
Theo phương i cứ một lần phép thử thành công lại kéo theo một lần phép thử không thành công thì lấy:

$$\lambda_j^{k+1} = 0$$

Về nguyên tắc phương pháp Rosenbrock được áp dụng cho hàm mục tiêu có số lượng thông số không hạn chế, nhưng với các máy tính cá nhân chỉ nên dùng phương pháp Rosenbrock khi mô hình có dưới 30 thông số.

2.4 Phương pháp Nelder-Mead

Thuật toán Nelder – Mead dùng để cực tiểu hàm số thực $f(x)$ với $x \in \mathbb{R}^n$. Thuật toán Nelder-Mead sử dụng một mô hình hình học gọi là simplex để di chuyển đi mò điểm tối ưu trong không gian tìm kiếm. Đó là lí do tại sao nó được gọi là “Simplex search method”. Các simplex n-chiều này được biến dạng nhờ 3 phép biến đổi : đối xứng gương, phép co, phép dãn.



Hình 2.2: Các ánh xạ gương, phép co, phép dãn

Có 4 tham số cần xác định trong thuật toán Nelder – Mead đó là: hệ số phản xạ ρ , hệ số dãn C , hệ số co D và hệ số thu hẹp E . Theo bài báo gốc của Nelder – Mead các tham số này cần thỏa mãn:

$$\rho > 0; C > 1; C > \rho; 0 < \gamma < 1 \text{ và } 0 < \sigma < 1$$

Cách chọn phổ biến nhất được dùng trong thuật toán Nelder – Mead chuẩn là:

$$\rho = 1, X = 2, \gamma = \frac{1}{2} \text{ và } \sigma = \frac{1}{2}$$

Các giá trị tham số này làm cho phương pháp trở nên hiệu quả, ngay cả khi làm việc trong những tình huống phức tạp.

Lúc bắt đầu bước lặp thứ k ($k \geq 0$) ta có đơn hình không suy biến Δ_k với $n+1$ đỉnh, mỗi một đỉnh là một điểm trong không gian R^n . Ta luôn luôn có thể giả thiết rằng bước lặp thứ k bắt đầu bằng việc sắp xếp và đánh nhãn các đỉnh này là x^1, x^2, \dots, x^{n+1} sao cho:

$$f_1^{(k)} \leq f_2^{(k)} \leq \dots \leq f_{n+1}^{(k)}$$

Trong đó $f_i^{(k)}$ kí hiệu cho $f(x^i)$. Bước lặp thứ k sinh ra một tập $n+1$ đỉnh xác định một đơn hình mới $\Delta_{k+1} \neq \Delta_k$. Vì ta cần tính cực tiểu của hàm f nên ta coi x^1 là điểm tốt nhất, x^{n+1} là đỉnh xấu nhất, x^n là điểm gần xấu nhất. Tương tự ta coi $f_{n+1}^{(k)}$ là giá trị hàm xấu nhất.

Kết quả của mỗi bước lặp là: hoặc tìm được một đỉnh mới thay thế x^{n+1} trong tập hợp các đỉnh trong bước lặp tiếp, hoặc nếu thực hiện việc thu hẹp thì một tập U đỉnh mới cùng với x^1 tạo nên đơn hình mới cho bước lặp tiếp theo.

2.5 Phương pháp Hooke-Jeeves

Thủ tục cơ bản của thuật toán Hooke – Jeeves là thủ tục dò tìm địa phương: xuất phát từ một điểm $x \in R^n$ theo $2n$ hướng dọc theo n trục tọa độ với độ dài bước h tìm điểm x' có giá trị hàm mục tiêu $f(x')$ tốt hơn $f(x)$. Mô tả chi tiết thủ tục dò tìm địa phương:

Bước 0. Có điểm x , bước dịch chuyển h , giá trị tốt nhất hàm mục tiêu hiện biết f^t ứng với phương án x^t . Biến $change = 0$ (biến $change$ để xác định gốc dò tìm có thay đổi không).

Bước 1. Xét tọa độ $j = 1$.

Bước 2. Tính điểm $z = x + h e^j = (x_1, \dots, x_{j-1}, x_j + h, x_{j+1}, \dots, x_n)$, $t = f(z)$ trong đó e^j là véc tơ đơn vị thứ j , $e^j = (0, \dots, 0, 1, 0, \dots, 0)$, số 1 ở vị trí j . Tính $t = f(z)$.

Bước 3. Nếu $t < f^t$ thì $f^t \leftarrow t$, $x \leftarrow z$, $change = change + 1$, và sang bước 6.

Bước 4. Tính điểm $z = x - h.e^j = (x_1, \dots, x_{j-1}, x_j - h, x_{j+1}, \dots, x_n)$, $t = f(z)$.

Bước 5. Nếu $t < f^t$ thì $f^t \leftarrow t$, $x \leftarrow z$, $change = change + 1$.

Bước 6. $j = j + 1$. Nếu $j \leq n$ thì chuyển lên bước 2.

Sau thử tục trên nếu $change = 0$ thì ta không tìm được điểm x' có giá trị hàm mục tiêu tốt hơn điểm xuất phát. Nếu $change > 0$ thì ta tìm được điểm x' (giá trị véc tơ này chứa trong biến x và x^t) có giá trị hàm mục tiêu tốt hơn điểm xuất phát, giá trị tốt nhất hàm mục tiêu f^t và x^t được cập nhật tốt hơn.

2.6 Giải thuật di truyền

Ý tưởng của giải thuật Di truyền là mô phỏng những gì mà tự nhiên đã thực hiện. Giải thuật Di truyền hình thành dựa trên quan niệm cho rằng: quá trình tiến hóa tự nhiên là quá trình hoàn hảo nhất, hợp lý nhất và tự nó đã mang tính tối ưu. Quá trình tiến hoá thể hiện tính tối ưu ở chỗ thế hệ sau bao giờ cũng tốt hơn thế hệ trước.

Giải thuật Di truyền áp dụng quá trình tiến hóa tự nhiên (vận dụng các nguyên lý của tiến hóa như di truyền, đột biến, chọn lọc tự nhiên, và trao đổi chéo) để giải các bài toán tối ưu trong thực tế (từ tập các lời giải có thể ban đầu thông qua nhiều bước tiến hóa hình thành các tập hợp mới với lời giải tốt hơn và cuối cùng sẽ tìm được lời giải gần tối ưu).

a) Các khái niệm cơ bản

Giải thuật Di truyền dựa vào quá trình tiến hoá trong tự nhiên nên các khái niệm và thuật ngữ của nó đều có liên quan đến các thuật ngữ của di truyền học.

Cá thể: Một cá thể trong giải thuật Di truyền, biểu diễn một giải pháp của bài toán. Tuy nhiên không giống với trong tự nhiên, một cá thể có nhiều nhiễm sắc thể (NST), có 1 thì gọi là thể đơn bội, còn nếu có nhiều thì là thể đa bội, ở đây để giới hạn trong giải thuật Di truyền ta quan niệm một cá thể có một nhiễm sắc thể. Do đó khái niệm cá thể và nhiễm sắc thể trong giải thuật Di truyền coi như là tương đương.

Nhiễm sắc thể: Một NST được tạo thành từ nhiều gen, mỗi gen có thể có các giá trị khác nhau để quy định một tính trạng nào đó. Trong giải thuật Di truyền, một gen được coi như một phân tử trong chuỗi NST.

Quần thể: Quần thể là một tập hợp các cá thể có cùng một số đặc điểm nào đấy. Trong giải thuật Di truyền ta quan niệm quần thể là một tập các lời giải của một bài toán.

Chọn lọc: Trong tự nhiên, quá trình chọn lọc và đấu tranh sinh tồn đã làm thay đổi các cá thể trong quần thể. Những cá thể tốt, thích nghi được với điều kiện sống thì có khả năng đấu tranh lớn hơn, do đó có thể tồn tại và sinh sản. Các cá thể không thích nghi được với điều kiện sống thì dần mất đi. Dựa vào nguyên lý của quá trình chọn lọc và đấu tranh sinh tồn trong tự nhiên, chọn lựa các cá thể trong giải thuật Di truyền

chính là cách chọn các cá thể có độ thích nghi tốt để đưa vào thế hệ tiếp theo hoặc để cho lai ghép, với mục đích là sinh ra các cá thể mới tốt hơn. Có nhiều cách để lựa chọn nhưng cuối cùng đều nhằm đáp ứng mục tiêu là các cá thể tốt sẽ có khả năng được chọn cao hơn.

Lai ghép: Lai ghép trong tự nhiên là sự kết hợp các tính trạng của bố mẹ để sinh ra thế hệ con. Trong giải thuật di truyền, lai ghép được coi là một sự tổ hợp lại các tính chất (thành phần) trong hai lời giải cha mẹ nào đó để sinh ra một lời giải mới mà có đặc tính mong muốn là tốt hơn thế hệ cha mẹ. Đây là một quá trình xảy ra chủ yếu trong giải thuật Di truyền.

Đột biến: Đột biến là một sự biến đổi tại một (hay một số) gen của nhiễm sắc thể ban đầu để tạo ra một nhiễm sắc thể mới. Đột biến có xác suất xảy ra thấp hơn lai ghép. Đột biến có thể tạo ra một cá thể mới tốt hơn hoặc xấu hơn cá thể ban đầu. Tuy nhiên trong giải thuật Di truyền thì ta luôn muốn tạo ra những phép đột biến cho phép cải thiện lời giải qua từng thế hệ.

b) Mô hình giải thuật di truyền

Bước đầu tiên của quá trình tìm lời giải cho vấn đề là khởi tạo quần thể, nghĩa là phát sinh ra một số lượng lớn, hữu hạn các cá thể. Sau đó, dựa trên một hàm số nào đó gọi là hàm thích nghi để xác định giá trị độ thích nghi, giá trị này có thể hiểu là “độ tốt” của cá thể. Vì được phát sinh một cách ngẫu nhiên nên độ tốt của một lời giải hay tính thích nghi của một cá thể trong quần thể là không xác định.

Hãy tưởng tượng một máy tính giải quyết bài toán theo kiểu leo đồi, trong không gian tìm kiếm của bài toán là một vùng đất gập ghềnh, có nhiều ngọn đồi cao thấp khác nhau. Một người leo núi sẽ bị kẹt ở một đỉnh đồi thấp. Như vậy, nếu có nhiều người cùng leo đồi ở nhiều địa điểm khác nhau thì có khả năng có một trong số các người này leo đến đỉnh đồi cao nhất sẽ cao hơn.

Đến đây chúng ta sẽ có thể nảy sinh ý tưởng: sử dụng nhiều thế hệ các người leo đồi. Nghĩa là toàn bộ những người leo đầu tiên (chẳng hạn 100 người) đều chưa gặp đỉnh đồi cao nhất thì sẽ cho 100 người khác tiếp tục leo. Nhưng có khả năng trong nhóm người leo đồi mới, không ai trong số họ leo được những ngọn đồi cao hơn nhóm người trước. Hay nói một cách tổng quát: phải làm cách nào để giữ lại những người leo cao nhất cho thế hệ sau. Tiến trình cứ tiếp tục cho đến khi có một thế hệ nào đó có một người leo đến đỉnh đồi cao nhất hoặc hết thời gian cho phép. Trong trường hợp hết thời gian thì trong toàn bộ các thế hệ, người nào leo cao nhất sẽ được chọn.

Như vậy, để cải thiện tính thích nghi của quần thể, người ta tìm cách tạo ra quần thể mới. Có hai thao tác thực hiện trên thế hệ hiện đại để tạo ra một thế hệ khác có độ thích nghi tốt hơn.

Thao tác 1: sao chép nguyên mẫu một nhóm các cá thể tốt từ thế hệ trước rồi đưa sang thế hệ sau bằng phép tái sinh và chọn lọc. Thao tác này đảm bảo độ thích nghi của thế hệ sau luôn luôn tốt hơn hoặc bằng thế hệ trước.

Thao tác 2: là tạo ra các cá thể mới bằng phép lai ghép và đột biến. Trong phép lai ghép, gen của hai cá thể tốt được chọn trong thế hệ trước sẽ được phối hợp với nhau (theo một quy tắc nào đó) để tạo ra hai cá thể mới. Phép đột biến là biến đổi ngẫu nhiên một hoặc nhiều thành phần gen của một cá thể ở thế hệ trước để tạo ra một cá thể hoàn toàn mới ở thế hệ sau. Phép đột biến có thể làm xáo trộn và làm mất đi tính thích nghi cao của cá thể được chọn.

2.7 Phương pháp SCE (Shuffled Complex Evolution)

Phương pháp SCE được xây dựng dựa trên: sự kết hợp giữa phương pháp Downhill Simplex của Neader Mead (1956) với các khái niệm của phương pháp tìm kiếm ngẫu nhiên có kiểm soát (Price, 1987); sự xáo trộn cạnh tranh (Holland, 1975); xáo trộn các phức hợp. Phương pháp này giúp có thể giải nhanh và tránh được các bẫy của tối ưu địa phương, phù hợp để giải các bài toán quy hoạch phi tuyến phức tạp. Qua một số nghiên cứu (Duan và cộng sự 1992, 1993, 1994, Giha Lee, Yasuto Tachikawa và Kaoru Takara 2006, ...) cho thấy phương pháp này có tính ổn định và hội tụ rất tốt.

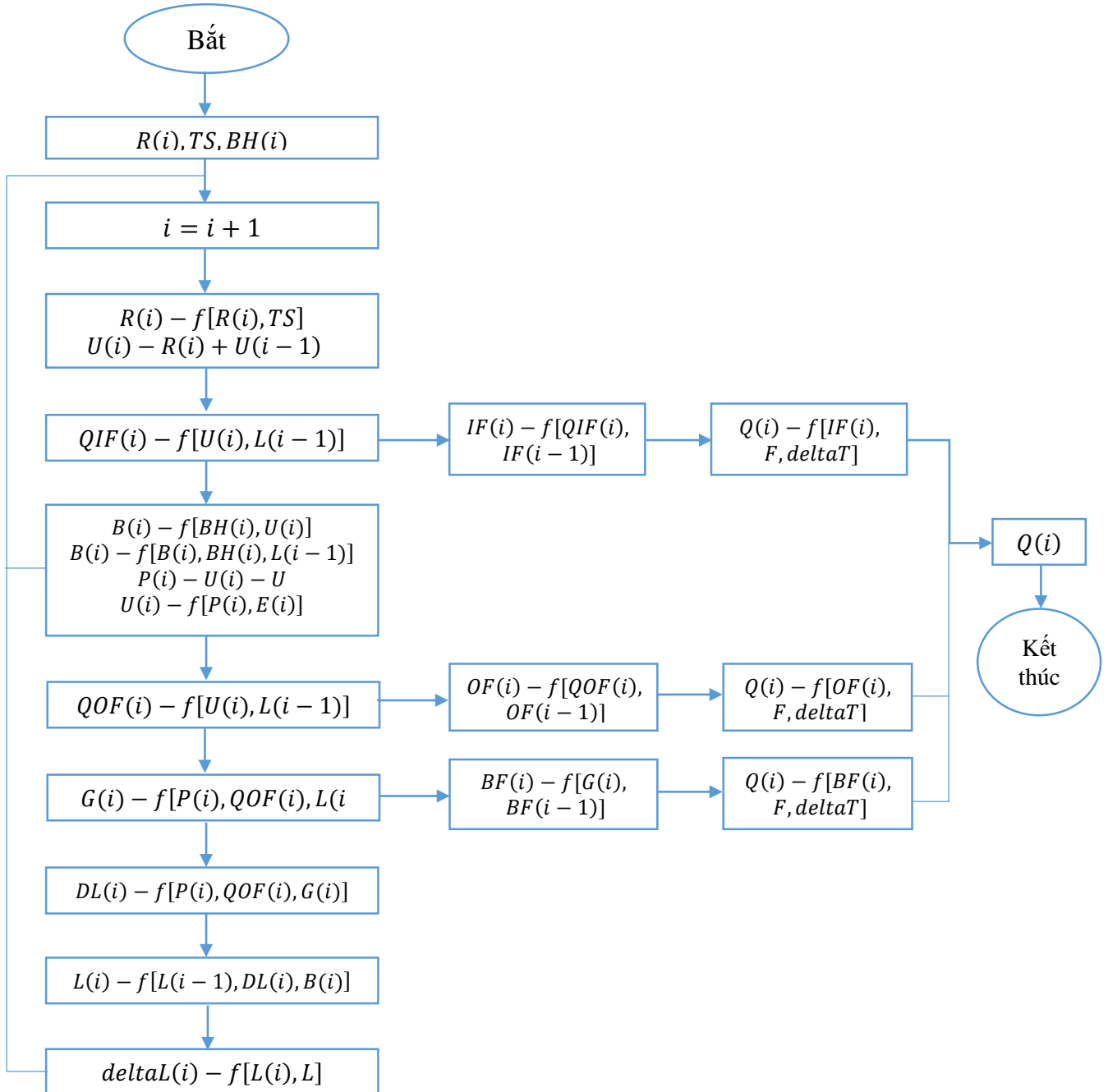
Phương pháp SCE được khởi tạo bằng cách chọn các tham số p và m với p là số phức hợp, m là số điểm trong mỗi phức hợp. Không gian mẫu s là mẫu lấy ngẫu nhiên trong không gian khả thi của thông số sử dụng một phân phối xác suất thống nhất và tính toán giá trị hàm mục tiêu tại mỗi điểm đó.

Sau đó, các điểm trong s đều được sắp xếp theo thứ tự giá trị của hàm mục tiêu tốt dần. Các điểm này sẽ được chia thành p phức hợp, mỗi phức hợp gồm m điểm. Mỗi phức hợp sẽ tiến hóa theo một cách độc lập theo phương pháp Downhill Simplex. Bước tiếp theo, xáo trộn, kết hợp các điểm trong các phức hợp đã phát triển thành một tập mẫu mới dựa trên thông tin của tập mẫu ban đầu. Sự phát triển và xáo trộn này sẽ được lặp đi lặp lại cho tới khi nào các tiêu chí hội tụ được thỏa mãn.

CHƯƠNG III

XÂY DỰNG PHẦN MỀM DỰ BÁO LŨ BẰNG MÔ HÌNH THỦY VĂN THÔNG SỐ TẬP TRUNG CÓ SỬ DỤNG PHƯƠNG PHÁP ƯỚC TÍNH THÔNG SỐ TỐI ƯU

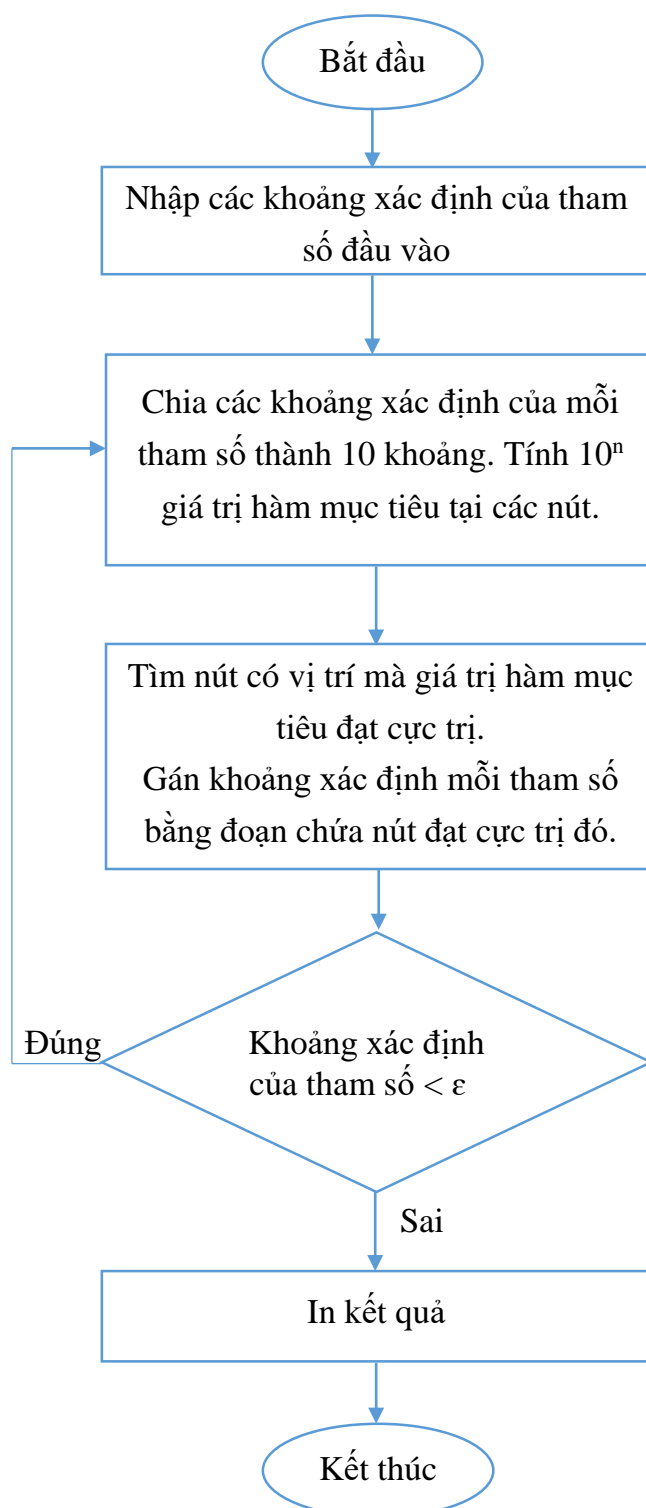
3.1 Sơ đồ khối mô hình Nielsen-Hansen



Hình 3.1: Sơ đồ khối mô hình Nielsen-Hansen

3.2 Sơ đồ khối phương pháp số giải bài toán ước tính thông số tối ưu

3.2.1 Sơ đồ khối của thuật toán ô vuông



Hình 3.2: Sơ đồ khối thuật toán ô vuông

Dưới đây là mô tả một bước lặp của thuật toán ô vuông:

Bước 1: Ứng với mỗi tham số của một khoảng xác định, nhập các khoảng xác định của các tham số đầu vào của mô hình.

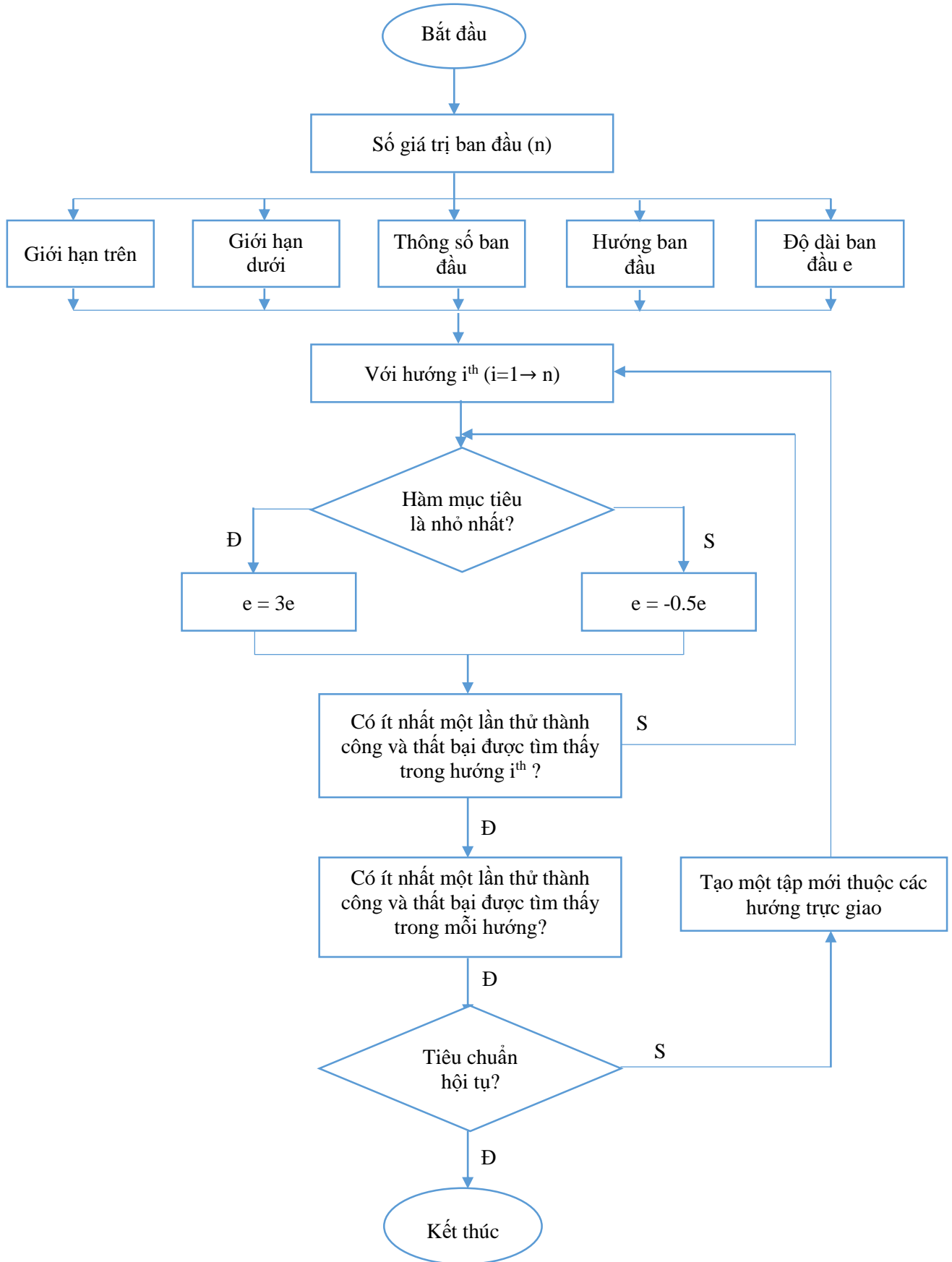
Bước 2: Khoảng xác định của mỗi tham số được chia thành 10 khoảng. Tính 10^n giá trị của hàm mục tiêu tại các nút chia thuộc khoảng đó.

Bước 3: Tìm nút có vị trí mà ở đó giá trị của hàm mục tiêu là cực trị. Tiến hành gán khoảng giá trị của mỗi tham số bằng đoạn chứa nút tại vị trí có giá trị hàm mục tiêu là cực trị vừa tìm được.

Bước 4: Xác định xem khoảng xác định của mỗi tham số vừa tìm được có nhỏ hơn ϵ không? Nếu đúng thì quay trở lại bước 2. Ngược lại, nếu sai thực hiện bước 5.

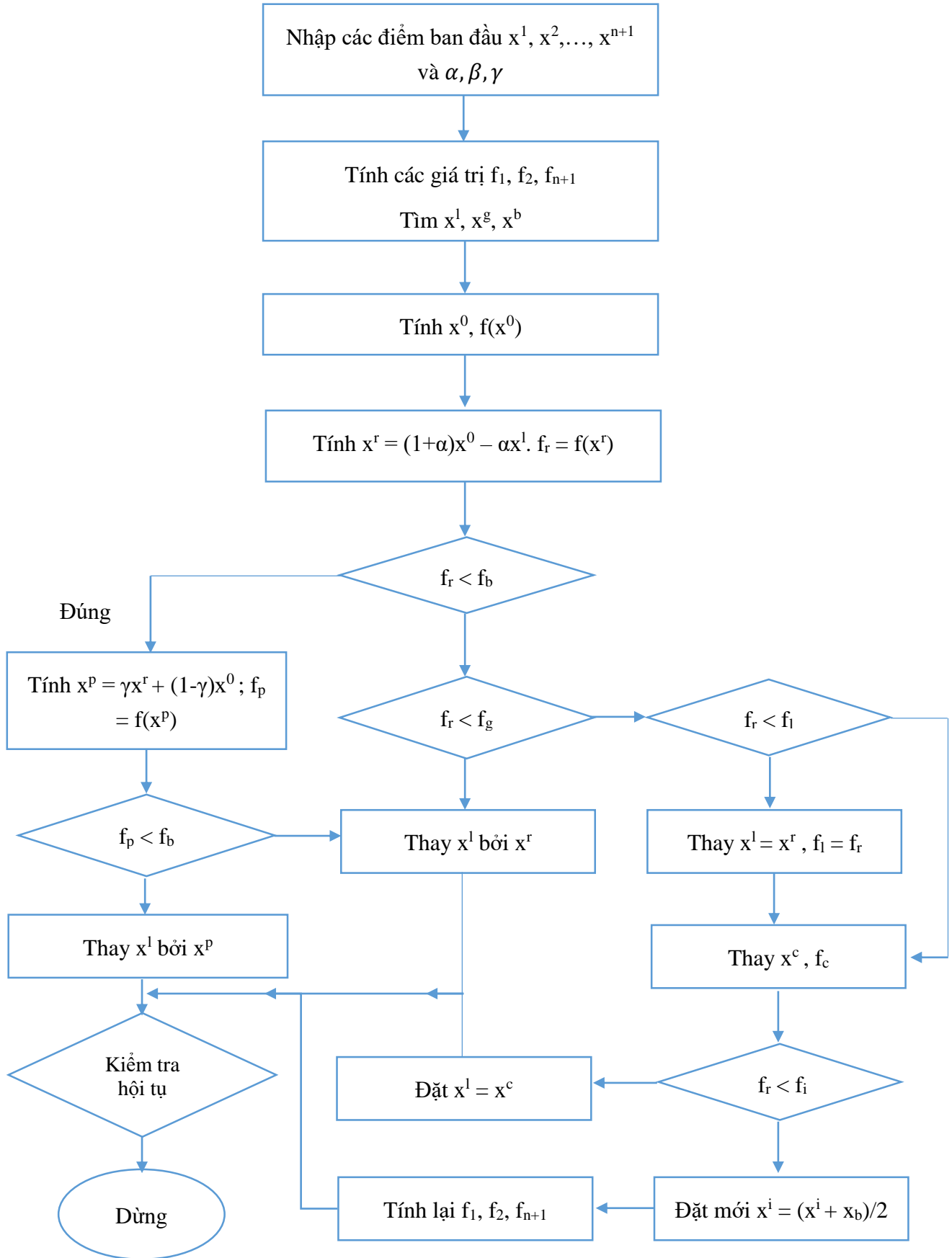
Bước 5: In kết quả giá trị của tham số vừa tìm được.

3.2.2 Sơ đồ khối của thuật toán RosenBrock



Hình 3.3: Sơ đồ khối của thuật toán RosenBrock

3.2.3 Sơ đồ khối của thuật toán Nelder-Mead



Hình 3.4: Sơ đồ khối của thuật toán Nelder-Mead

Bước 1: Tính giá trị hàm mục tiêu ở các đỉnh của đơn hình $f(x^1), f(x^2), \dots, f(x^{n+1})$

Bước 2: Xác định giá trị lớn nhất f_l giá trị tiếp theo f_g và giá trị nhỏ nhất f_b , các đỉnh tương ứng sẽ là x^l, x^g, x^b .

Bước 3: Xác định trọng tâm hình không kể đến đỉnh x^l theo công thức:

$$x^0 = \frac{1}{n} \sum_{i \neq l} x^i \quad \text{và} \quad f(x^0)$$

Bước 4: Thực hiện ánh xạ gương x^l qua x^0 nhận được x^r , tính $f_r = f(x^r)$.

Nếu hệ số ánh xạ $\alpha > 0$ thì vị trí x^r được xác định theo công thức:

$$x^r = x^0 + \alpha(x^0 - x^l), \quad x^r = (1 + \alpha)x^0 - \alpha x^l \quad \text{và} \quad \alpha = |x^r - x^0|/|x^0 - x^l|$$

Bước 5: So sánh f_r và f_b xảy ra:

- 1) Nếu $f_r < f_b$ thì ta nhận được giá trị nhỏ nhất của hàm. Hướng từ f_0 đến f_r là hướng dịch chuyển tốt nhất cho dịch chuyển. Kéo dẫn theo hướng này để tìm điểm x^p (hình 2.2b). Hệ số dẫn $\gamma > 1$ được tính từ:

$$x^p - x^0 = \gamma(x^r - x^0), \quad x^p = \gamma x^r + (1 - \gamma)x^0, \quad \gamma = |x^p - x^0|/|x^r - x^0|$$

Tính $f_p = f(x^p)$.

- Nếu $f_r < f_b$ ta thay đổi x^r bởi x^p và kiểm tra điểm thứ (n+1) của đơn hình về tính hội tụ đến cực tiểu. Nếu hội tụ thì dừng, nếu không thì quay lại bước 2.
 - Nếu $f_r > f_b$ thì bỏ điểm x^p vì đã đi quá xa. Thay x^l bởi x^r và kiểm tra tính hội tụ, nếu vẫn chưa hội tụ thì quay lại bước 2.
- 2) Nếu $f_r > f_b$ nhưng $f_r > f_g$ thì x^r tốt hơn so với 2 điểm còn lại của đơn hình. Thay x^l bằng x^r , kiểm tra nếu chưa hội tụ thì quay lại bước 2.
 - 3) Nếu $f_r > f_b$ và $f_r > f_g$ thì chuyển qua bước 6.

Bước 6: So sánh các giá trị f_r và f_l :

- 1) Nếu $f_r > f_l$ thì chuyển về mục 2 của bước 6.
 Nếu $f_r < f_l$ thay x^l bởi x^r và thay f_l bởi f_r
 Nếu $f_r > f_g$ thì thực hiện phép co.
- 2) Thực hiện phép co:
 - Nếu $f_r > f_l$ ta đã dịch quá xa theo hướng từ x^l đến x^0 nên cần sửa lại phép co để tìm x^c (hình 2.2c): $x^c = x^0 + \beta(x^l) - x^0$ với $0 < \beta < 1$ là hệ số co.
 - Nếu $f_r < f_l$ thay $f_l = f_r$ và $x^l = x^r$ sau đó thực hiện phép co và tìm x^c (hình 2.2d) theo công thức sau: $x^c = x^0 + \beta(x^r - x^0) = \beta x^r + (1 - \beta)x^0$.

Bước 7: Tính f_c . So sánh giá trị f_c và f_l .

- Nếu $f_c < f_l$ thì thay x^l bởi x^c , nếu chưa hội tụ thì quay về bước 2.
- Nếu $f_c > f_l$ thì không tìm được giá trị nhỏ hơn f_l , chuyển qua bước 8.

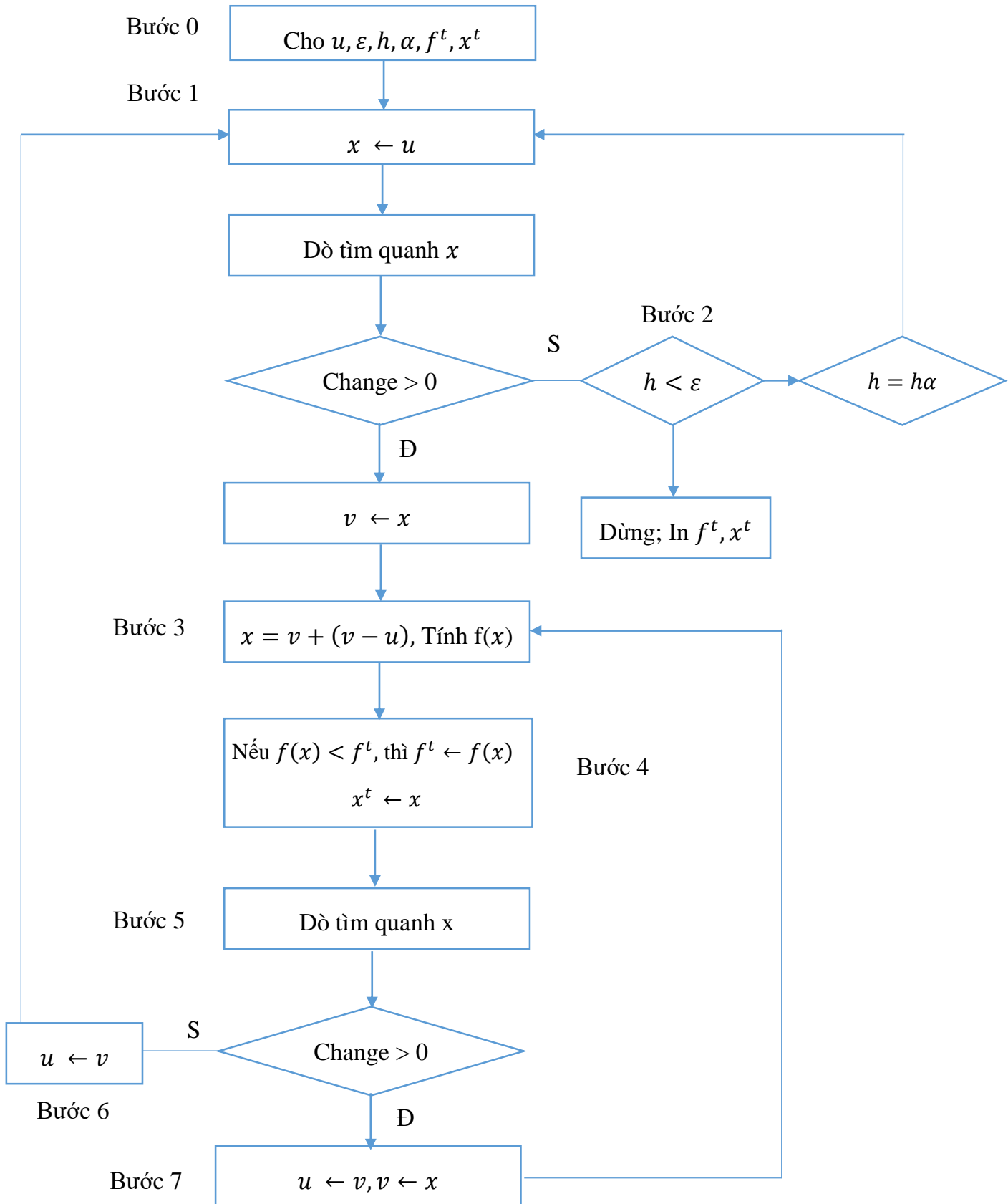
Bước 8: Thu nhỏ kích thước đơn hình còn một nửa, lấy x^b làm chuẩn. Nghĩa là x^l được thay bằng:

$$x^i = x^i - \frac{1}{2}(x^i - x^b) = \frac{1}{2}(x^i + x^b).$$

Tính các f_i , $i = 1, 2, \dots, n+1$. Kiểm tra tính hội tụ, nếu không được thì quay lại bước 2.

Bước 9: Tính $\sigma^2 = \sum_{i=1}^{n+1} (f_i - f)^2 / (n + 1)$, $f = f_i / (n + 1)$. Nếu σ nhỏ hơn độ chính xác ϵ nào đó thì giá trị của các hàm số rất gần nhau. Vì vậy điểm min rất gần với x^b .

3.2.4 Sơ đồ khối của thuật toán Hooke-Jeeves



Hình 3.5: Sơ đồ khối của thuật toán Hooke-Jeeves

Dưới đây là mô tả một bước lặp của thuật toán Hooke-Jeeves:

Bước 0. Cho điểm xuất phát $u \in \mathbb{R}^n$, cho bước dịch chuyển theo trục tọa độ h , cho hệ số giảm của h là α , cho số ε đủ nhỏ. Tính giá trị hàm mục tiêu tại điểm xuất phát $f(u)$, $f^t \leftarrow f(u)$, $x^t \leftarrow u$.

Bước 1. $x \leftarrow u$, tiến hành thủ tục dò tìm địa phương theo điểm gốc x . Nếu $change > 0$ thì $v \leftarrow x$, chuyển sang Bước 3.

Bước 2. Trường hợp kết quả dò tìm có điểm gốc không thay đổi. Nếu $h < \varepsilon$ thì dừng thuật toán, in giá trị tối ưu f^t và phương án tối ưu x^t , trái lại ta giảm bước dịch chuyển h : $h \leftarrow h \cdot \alpha$ và quay lại Bước 1.

Bước 3. Xác định điểm $x = (v-u)$, tính $f(x)$.

Bước 4. Nếu $f(x) < f^t$ thì $f^t \leftarrow f(x)$, $x^t \leftarrow x$.

Bước 5. Thực hiện thủ tục dò tìm địa phương theo điểm gốc x . Nếu $change > 0$ thì chuyển sang Bước 7.

Bước 6. Trường hợp kết quả dò tìm có điểm gốc không thay đổi, tiến hành $u \leftarrow v$, quay lại Bước 1.

Bước 7: Tiến hành $u \leftarrow v$, $v \leftarrow x$ và quay về Bước 3.

3.2.5 Sơ đồ khối của giải thuật di truyền

Một bước lặp của phương pháp giải thuật di truyền được mô tả như sau:

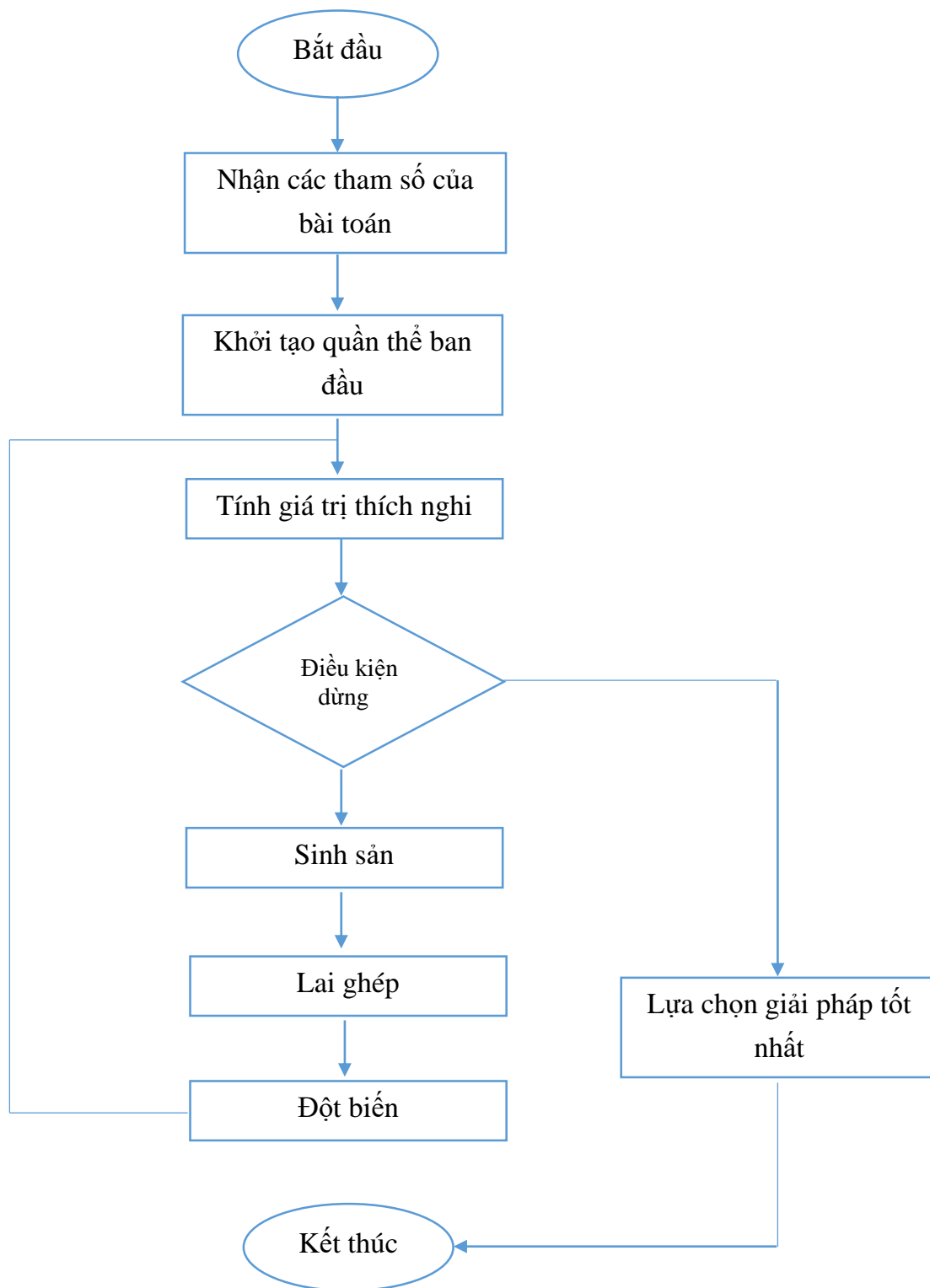
Bước 1. [Bắt đầu] Nhận các tham số cho thuật toán.

Bước 2. [Khởi tạo] Sinh ngẫu nhiên một quần thể gồm n cá thể (là n lời giải cho bài toán), mỗi cá thể là 1 bộ thông số trong các khoảng cho trước.

Bước 3. [Quần thể mới] Tạo quần thể mới bằng cách lặp lại các bước sau cho đến khi quần thể mới hoàn thành.

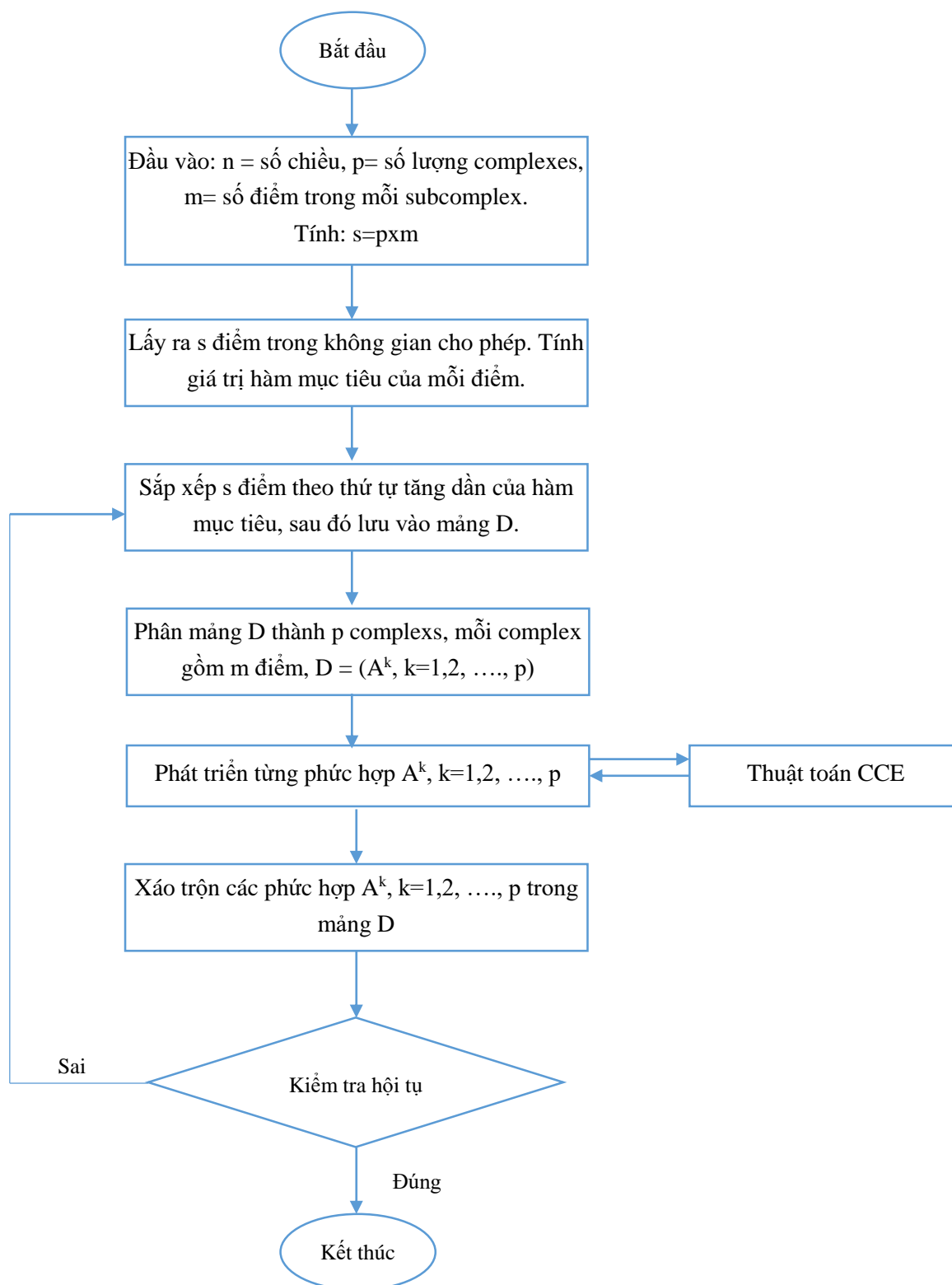
- [Thích nghi] Ước lượng độ thích nghi của mỗi cá thể (là hàm mục tiêu f).
- [Kiểm tra] Kiểm tra điều kiện kết thúc giải thuật.
- [Chọn lọc] Chọn hai cá thể bố mẹ từ quần thể cũ theo độ thích nghi của chúng (cá thể có độ thích nghi càng cao thì càng có nhiều khả năng được chọn) – chọn bộ thông số có hàm mục tiêu nhỏ theo thứ tự từ dưới lên trên.
- [Lai ghép] Với một xác suất lai ghép được chọn, lai ghép hai cá thể bố mẹ để tạo ra một cá thể mới.
- [Đột biến] Với một xác suất đột biến được chọn, biến đổi cá thể mới.

Bước 4. [Chọn kết quả] Nếu điều kiện dừng được thỏa mãn thì thuật toán kết thúc và trả về lời giải tốt nhất trong quần thể hiện tại.



Hình 3.6: Sơ đồ khối của thuật toán giải thuật di truyền

3.2.6 Sơ đồ khối của thuật toán SCE



Hình 3.7: Sơ đồ khối thuật toán SCE

Phương pháp SCE được mô tả trong sơ đồ khối như trên gồm các bước sau:

Bước 1: Khởi tạo tham số p, m với $p \geq 1, m \geq n+1$. Trong đó: p là số phức hợp, m là số điểm trong mỗi phức hợp, n là thông số hiệu chỉnh thì $s = p * m$ là không gian mẫu.

Bước 2: Tạo ra nhóm giá trị. Lấy ra ngẫu nhiên s điểm $x_i \{x_1, x_2, \dots, x_s\}$ trong không gian xác định của tham số. Sau đó tính giá trị f_i của hàm mục tiêu.

Bước 3: Sắp xếp s điểm x theo thứ tự tăng dần ý nghĩa giá trị của hàm mục tiêu và lưu trong mảng D , $D = \{(x_i, f_i), i = 1, 2, \dots, s\}$, ứng với $i=1$ giá trị của hàm mục tiêu là tệ nhất.

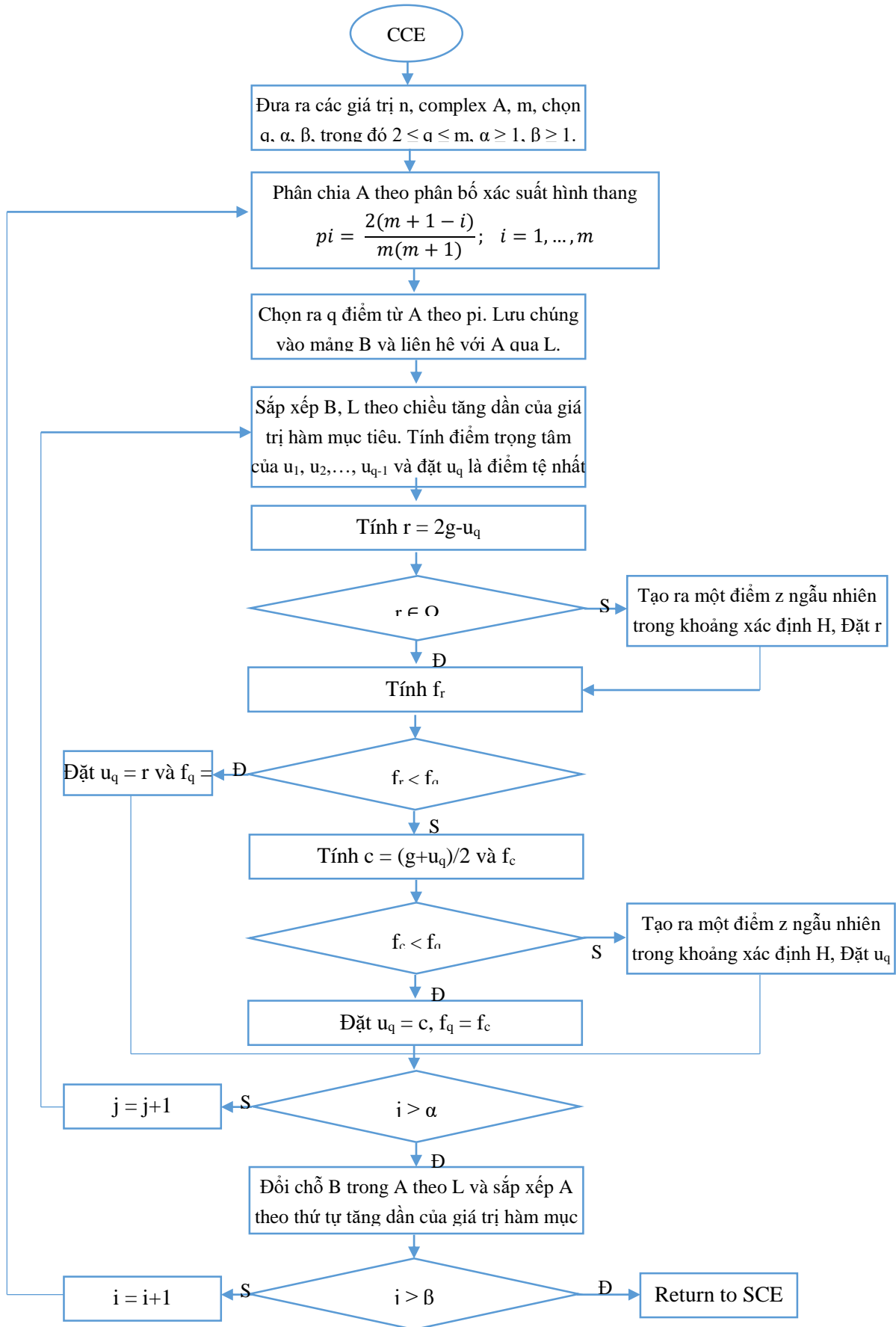
Bước 4: Phân nhóm, chia mảng D thành p phức hợp A_1, A_2, \dots, A_p với mỗi phức hợp gồm m điểm, sao cho: $A_k = [(x_j^k, f_j^k) \mid x_j^k = x_{k+p(j-1)}, f_j^k = f_{k+p(j-1)}], j = 1, 2, \dots, m$.

Bước 5: Phát triển từng phức hợp một, mỗi phức hợp $A_k, k=1, 2, \dots, p$ được phát triển theo thuật toán CCE. (Competitive Complex Evolution).

Bước 6: Xáo trộn các phức hợp, kết hợp các điểm trong các phức hợp đã được phát triển trong bước 5 thành một tập mẫu duy nhất và sắp xếp lại tập mẫu này theo thứ tự tăng dần ý nghĩa hàm mục tiêu. Phân nhóm lại tập mẫu và p phức hợp theo cách thức ở bước 4.

Bước 7: Kiểm tra điều kiện hội tụ, nếu thỏa mãn điều kiện thì dừng lại, nếu không thì tiếp tục.

Bước 8: Kiểm tra số lượng phức hợp, nếu số lượng nhỏ nhất của các phức hợp được đề nghị p_{\min} nhỏ hơn p , xóa phức hợp chứa điểm tệ nhất, đặt $p = p - 1, s = p \cdot m$. Quay trở về bước 4. Nếu $p_{\min} = p$ thì quay lại bước 4.



Hình 3.8: Sơ đồ khối thuật toán CCE

Bước 1: Xây dựng một phức hợp con bằng cách ngẫu nhiên chọn ra q điểm từ phức hợp đó theo phân phối xác suất hình trang.

Bước 2: Tìm ra điểm tệ nhất và tính điểm trọng tâm của phức hợp con không bao gồm điểm tệ nhất.

Bước 3: Lấy đối xứng điểm tệ nhất qua điểm trọng tâm, nếu điểm mới tạo ra không nằm trong khoảng cho phép thì ngẫu nhiên tạo ra một điểm trong không gian cho phép.

Bước 4: Nếu điểm mới tạo ra tốt hơn điểm cũ, thay thế điểm tệ nhất bằng điểm mới sau đó chuyển đến bước 7. Ngược lại chuyển đến bước 5.

Bước 5: Lấy trung tâm điểm của điểm tệ nhất và điểm trọng tâm, nếu điểm này tốt hơn điểm tệ nhất thì thay thế điểm tệ nhất bằng điểm này và chuyển đến bước 7, ngược lại chuyển đến bước 6.

Bước 6: Ngẫu nhiên tạo ra một điểm trong không gian cho phép, thay thế điểm tệ nhất bằng điểm ngẫu nhiên vừa tạo ra.

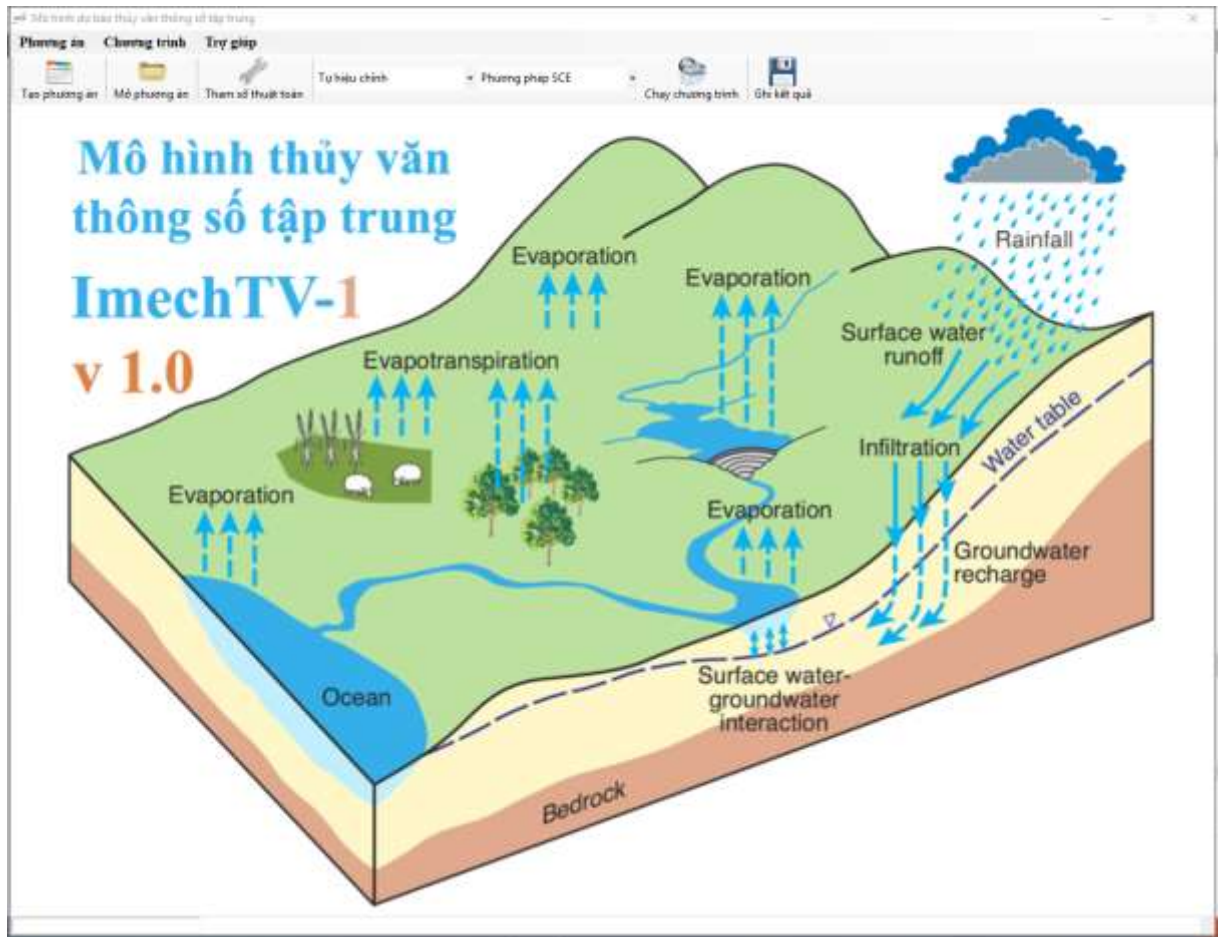
Bước 7: Lặp lại bước 2 – 6 α lần, với $\alpha \geq 1$ là số điểm mới mà mỗi phức hợp con tạo ra trước khi đưa vào phức hợp cũ.

Bước 8: Lặp lại bước 1 đến 7 với β lần, với $\beta \geq 1$ là số bước phát triển.

3.3 Xây dựng phần mềm

Phần mềm được xây dựng bằng ngôn ngữ C#. Phần mềm tính toán có giao diện đơn giản, dễ sử dụng, phù hợp cho nhiều nhóm đối tượng sử dụng.

Một số thông tin về thông tin lưu vực và các tham số của phương án tính toán bao gồm: Diện tích lưu vực (km^2), số trạm đo mưa của lưu vực, U_{\max}, L_{\max} , CQOF, CKIF, CK1, CK2, TOF, TIF, TG, CKBF. Người dùng nhập dữ liệu vào thư mục *Input* là đầu vào của mô hình thì sau khi lựa chọn phương án chạy phần mềm sẽ thể hiện các thông tin lưu vực như dưới đây:



Hình 3.9: Giao diện chung của phần mềm

Phương án tính

Thông tin lưu vực: Tham số phương án tính | Tham số thuật toán

Tên lưu vực: Lưu vực hồ Bàn Chát năm 2020

Diện tích: 1929 km² Lưu lượng thực đo: Qvao.txt

Tên trạm	File	Umax	Lmax	CQOF	CKIF	CK1	CK2	TOF	TIF	TG	CKBF	Weight
TamDuong	TamDuong_2020.txt	20.3631	260.85...	0.1370	683.5727	45.4217	40.0307	0.5394	0.4690	0.1706	689.9118	0.7565
BinhLu	BinhLu_2020.txt	11.6562	131.22...	0.3107	398.5553	15.5619	13.8147	0.2548	0.0931	0.0860	1649.5082	0.3212
ThanUyen	ThanUyen_2020.txt	7.6277	172.55...	0.0794	399.9988	7.3616	40.9732	0.4210	0.2354	0.1588	1317.5770	0.0236
MuCangChai	MuCangChai_2020...	8.8758	249.79...	0.1339	883.6403	37.3555	28.6563	0.3043	0.2533	0.1240	1521.8681	0.0321
TaGia	TaGia_2020.txt	6.9669	117.35...	0.1409	466.5619	6.5204	37.6889	0.1255	0.2750	0.2921	1734.9736	0.1540

20.0000 225.0000 0.5000 1100.0000 37.5000 37.5000 0.4500 0.4500 0.4500 2750.0000 0.5000

Thêm mới Xóa Cập nhật

Ghi lại

Hình 3.10: Thông tin về các tham số của phương án tính toán

Người dùng sau khi lựa chọn phương án tính toán cho lưu vực cần tính, có thể sử dụng một số thuật toán khác nhau để tính toán cho kết quả là đường quá trình dòng chảy tại lưu vực. Một số phương pháp số giải bài toán ước tính thông số tối ưu có thể lựa chọn như: Phương pháp chia ô, phương pháp RosenBrock, phương pháp Hooke-

Jeeves, phương pháp Nelder-Mead, phương pháp giải thuật di truyền và phương pháp SCE.

Phương pháp	Tham số	Giá trị	Đơn vị / Ghi chú
Phương pháp Ô vuông	Số khoảng chia	3	
	Sai số	0.001	
Phương pháp Rosen Block	beta	0.5	
	Sai số	0.001	
Phương pháp Hook-Jeeves	Sai số	0.001	
	Số vòng lặp	5000	
Phương pháp Nelder Mead	alpha	1.0	
	Sai số	0.001	
Thuật toán di truyền	Số cá thể	2000	
	Sai số	0.001	
Phương pháp SCE	ngs	10	
	Sai số	0.001	

Hình 3.11: Một số phương pháp giải có trong phần mềm

Tiếp đến, ngoài phương pháp dò tìm bộ thông số thủ công thì người dùng có thêm một lựa chọn khác là dò tìm thông số tự động. Hộp thoại ComboBox giúp người dùng dễ dàng lựa chọn và thay đổi kiểu tính toán cho mô hình thủy văn.

Hình 3.12: Lựa chọn loại hình tính toán cho mô hình thủy văn

CHƯƠNG IV

THỬ NGHIỆM CÔNG NGHỆ TRONG TÍNH TOÁN DỰ BẢO LŨ TRÊN MỘT SỐ LƯU VỰC SÔNG

4.1 Thử nghiệm dự báo

Nhóm tác giả thử nghiệm một số phương pháp giải bài toán ước tính thông số tối ưu cho mô hình thủy văn, phục vụ trong dự báo lũ cho một số lưu vực sông ở Việt Nam – nhóm tác giả là thành viên của 1 trong 6 nhóm tư vấn cho Cục Phòng tránh thiên tai về dự báo lũ lụt đồng bằng sông Hồng – Thái Bình hàng năm. Trong nghiên cứu này, chúng tôi thử nghiệm trên 12 lưu vực thủy văn bao gồm: Bản Chát, Huội Quảng, Sơn La, Hòa Bình, Thác Bà, Tuyên Quang, Hàm Yên, Yên Bái, Cầu, Thương, Lục Nam, Lâm Sơn. Trong 12 lưu vực trên, ngoài Lâm Sơn, thì 11 lưu vực này đều là các lưu vực chính trong hệ thống các lưu vực sông Hồng – Thái Bình, và được hiệu chỉnh dò tìm các thông số với bộ số liệu năm 2016 đến 2019 và lấy bộ thông số tốt nhất kiểm định lại với bộ số liệu năm 2020. Lâm Sơn là 1 lưu vực nhỏ bao gồm 1 trạm mưa và 1 trạm đo lưu lượng, được lựa chọn để thử nghiệm phương pháp ô vuông, được hiệu chỉnh dò tìm bằng bộ số liệu năm 2004 và kiểm định với bộ số liệu năm 2006.



Hình 4.1: Vị trí các lưu vực nghiên cứu.

- Lưu vực thủy văn Bản Chát có diện tích gần 1.929 km², phía trên giáp với biên giới Việt Nam – Trung Quốc, phía dưới giáp với lưu vực thủy văn Huội Quảng. Nước của lưu vực sẽ dồn về thủy điện Bản Chát, thủy điện bậc thang phía trên của thủy điện Huội Quảng, lưu lượng xả ra của hồ Bản Chát sẽ về hồ Huội Quảng trong một khoảng thời gian ngắn.

- Lưu vực thủy văn Thác Bà có tổng diện tích là 6.430 km², có nhánh chính là dòng sông Chảy thuộc lưu vực sông Lô. Lưu vực thủy văn Thác Bà là lưu vực thủy văn độc lập, không chịu sự ảnh hưởng bởi quá trình điều tiết của các hồ thủy điện khác.
- Lưu vực thủy văn Tuyên Quang có tổng diện tích là 5.868 km², có nhánh chính là dòng sông Gâm thuộc lưu vực sông Lô. Lưu vực thủy văn Tuyên Quang là lưu vực thủy văn độc lập, không chịu sự ảnh hưởng bởi quá trình điều tiết của các hồ thủy điện khác.
- Lưu vực thủy văn Huội Quảng có diện tích gần 2.825 km², phía trên giáp với đập thủy điện Bản Chát, phía dưới giáp với lưu vực thủy văn Sơn La. Lưu vực Huội Quảng sẽ chịu ảnh hưởng trực tiếp bởi thủy điện Bản Chát và gây ra ảnh hưởng tới lưu vực thủy văn Sơn La. Cần thu thập số liệu điều tiết của thủy điện Bản Chát để tính toán thủy văn cho lưu vực hồ Huội Quảng.
- Lưu vực hồ Sơn La có diện tích gần 43.760 km², phía trên giáp với đập thủy điện Lai Châu, phía dưới giáp lưu vực hồ Hòa Bình. Để tính toán thủy văn cho lưu vực Sơn La thì cần thu thập số liệu điều tiết thủy điện Lai Châu và thủy điện Huội Quảng do là các thủy điện bậc trên của thủy điện Sơn La. Lưu vực Sơn La sẽ nhận lưu lượng nước trực tiếp khi xả từ thủy điện Lai Châu và thủy điện Huội Quảng sau một khoảng thời gian chảy truyền.
- Lưu vực hồ Hòa Bình có diện tích gần 53.600 km², phía trên giáp với đập thủy điện Sơn La, phía dưới kết nối với hạ lưu sông Đà qua đập thủy điện Hòa Bình. Để tính toán thủy văn cho lưu vực Hòa Bình thì cần phải thu thập số liệu điều tiết hồ chứa Sơn La do là thủy điện bậc trên của thủy điện Hòa Bình vì lưu vực Hòa Bình sẽ nhận lưu lượng nước trực tiếp khi thủy điện Sơn La xả.
- Lưu vực Yên Bái có diện tích 27.306 km² thuộc một phần của lưu vực sông Thao. Lưu vực sông Thao là dòng chính của sông Hồng bắt nguồn từ dãy núi Ngụy Sơn cao trên 2000 m thuộc tỉnh Vân Nam Trung Quốc (ở Trung Quốc sông Thao còn có tên là sông Nguyên). Sông Thao có chiều dài sông là 902 km (tính đến Việt Trì) trong đó chiều dài trong nước là 332 km.
- Lưu vực Hàm Yên có diện tích 11.860 km² thuộc một phần của lưu vực sông Lô. Lưu vực Hàm Yên được tính từ thượng nguồn sông Lô – Hà Giang trải dài xuống Bắc Quang, Vĩnh Tuy cho đến Hàm Yên.
- Lưu vực sông Cầu có diện tích khoảng 2.760 km² và có dòng chính là sông Cầu bắt nguồn từ vùng núi Tam Tao (1.326 m), phía tây tỉnh Bắc Kạn, chảy

theo hướng tây bắc - đông nam, qua thị xã Bắc Kạn, rồi chuyển hướng gần bắc - nam chảy qua Chợ Mới, thành phố Thái Nguyên, Đáp Cầu rồi tiếp nhận sông Thương ở thượng lưu Phả Lại khoảng 2 km.

- Lưu vực sông Thương có diện tích khoảng 2.330 km², có tên nhánh sông chính là sông Thương có chiều dài 157 km bắt nguồn từ dãy núi Na-Pa-Phước cao 600 m ở huyện Chi Lăng tỉnh Lạng Sơn, chảy theo hướng đông bắc - tây nam qua thành phố Bắc Giang, tiếp nhận thêm sông Lục Nam rồi chảy vào sông Cầu ở thượng lưu Phả Lại.
- Lưu vực sông Lục Nam có diện tích lưu vực vào khoảng 2.090 km². Dòng chính là sông Lục Nam dài khoảng 175 km bắt nguồn từ vùng núi Kham, cao 700 m ở huyện Đình Lập tỉnh Lạng Sơn, chảy theo hướng đông bắc - tây nam qua các huyện Sơn Động, Lục Ngạn, Lục Nam tỉnh Bắc Giang rồi đổ vào sông Thương ở Tứ Yên, cách Phả Lại 10 km về phía thượng lưu.
- Lưu vực Lâm Sơn có diện tích lưu vực khoảng 33 km². Dòng chính là sông Bùi có chiều dài 91 km, bắt nguồn từ xã Lâm Sơn, huyện Lương Sơn, Hòa Bình chảy qua các tỉnh Hà Nội, Hoà Bình và hợp lưu với sông Đáy tại xã Phúc Lâm, huyện Chương Mỹ, Hà Nội.

4.2 Kết quả tính toán và đánh giá

Ngoài phương pháp ô vuông, các phương pháp khác đều lựa chọn ngẫu nhiên 1 hay nhiều nghiệm ban đầu, và việc lựa chọn này có thể dẫn đến các cực trị địa phương, vì vậy tác giả thực hiện chạy mỗi phương pháp 1000 lần (100 lần với một số lưu vực lớn) để tính toán nhằm hạn chế việc kết quả rơi vào cực trị địa phương. Trong bảng 4.1 là kết quả hệ số NSE sau khi tính toán thử nghiệm các phương pháp số dò tìm thông số tối ưu cho các lưu vực.

Bảng 4.1: Bảng kết quả hệ số NSE khi sử dụng các phương pháp tính toán dò tìm thông số tối ưu được thử nghiệm cho 12 lưu vực.

Bảng kết quả NSE tính toán		RosenBrock		HookeJeeves		NelderMead		GA		SCE	
		HC	KĐ	HC	KĐ	HC	KĐ	HC	KĐ	HC	KĐ
Lâm Sơn	TBC	0.64	0.54	0.70	0.61	0.70	0.59	0.70	0.60	0.74	0.62
	Max	0.74		0.74		0.75		0.74		0.75	
	TGTB	5.6s		8.8s		6.3s		45.5s		29.3s	
Bản Chất	TBC	0.40	0.44	0.56	0.52	0.35	0.30	0.44	0.35	0.57	0.51
	Max	0.51		0.59		0.49		0.5		0.58	
	TGTB	187s		346s		198s		345s		755s	
Huội Quảng	TBC	0.28	0.19	0.25	0.21			0.30	0.25	0.50	0.5
	Max	0.34		0.28				0.36		0.52	

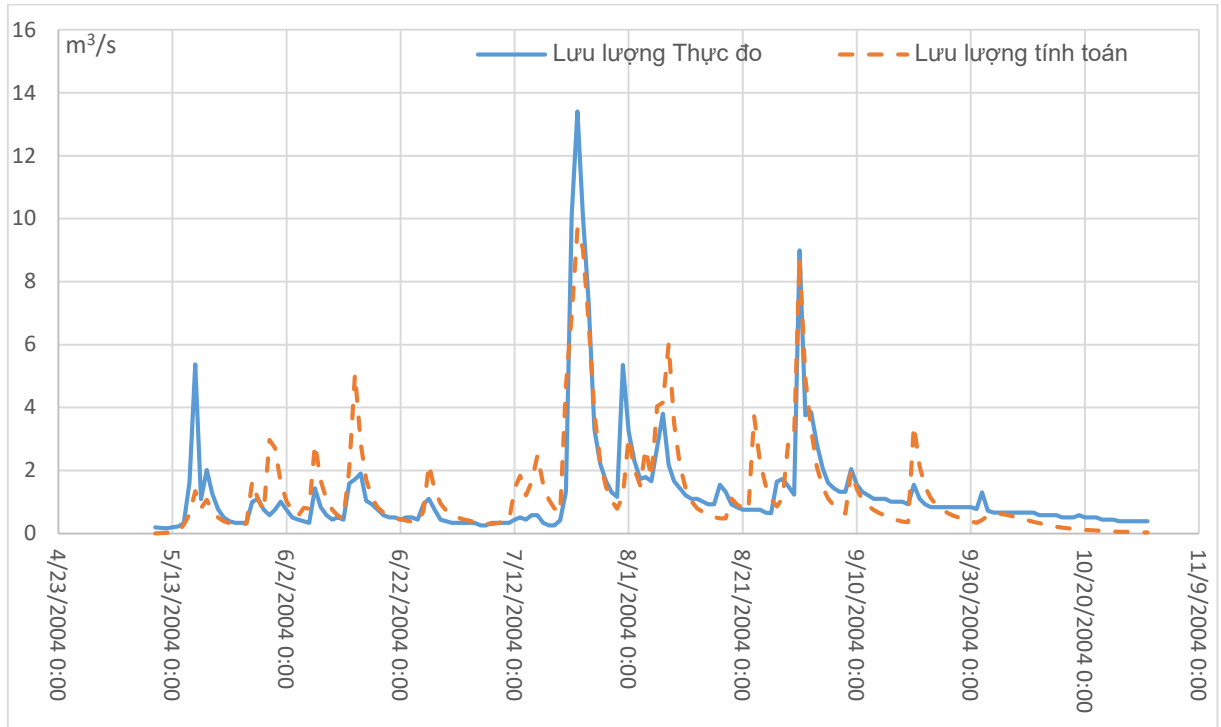
	TGTB	201s		355s				337s		724s	
Sơn La	TBC									0.50	0.51
	Max									0.54	
	TGTB									1484s	
Hòa Bình	TBC	0.19	0.24	0.43	0.40					0.48	0.50
	Max	0.33		0.44						0.52	
	TGTB	404s		693s						1005s	
Thác Bà	TBC	0.28	0.4	0.57	0.55					0.55	0.52
	Max	0.51		0.61						0.60	
	TGTB	235s		372s						760s	
Tuyên Quang	TBC	0.14	0.22	0.50	0.51					0.42	0.45
	Max	0.35		0.56						0.49	
	TGTB	329s		534s						987s	
Yên Bái	TBC	0.34	0.4	0.53	0.55					0.60	0.58
	Max	0.47		0.61						0.66	
	TGTB	335s		588s						1004s	
Hàm Yên	TBC	0.12	0.2	0.35	0.3					0.54	0.50
	Max	0.19		0.37						0.57	
	TGTB	135s		262s						623s	
Cầu	TBC	0.43	0.53	0.64	0.70	0.30	0.43	0.50	0.59	0.73	0.65
	Max	0.59		0.74		0.50		0.70		0.74	
	TGTB	85s		174s		95s		194s		493s	
Thương	TBC	0.4	0.51	0.52	0.53	0.10	0.11	0.26	0.35	0.56	0.52
	Max	0.55		0.56		0.16		0.4		0.57	
	TGTB	213s		357s		198s		335s		707s	
Lục Nam	TBC	0.43	0.5	0.57	0.61	0.22	0.33	0.44	0.51	0.61	0.59
	Max	0.56		0.64		0.39		0.56		0.62	
	TGTB	139s		215s		134s		281s		652s	

Trong đó:

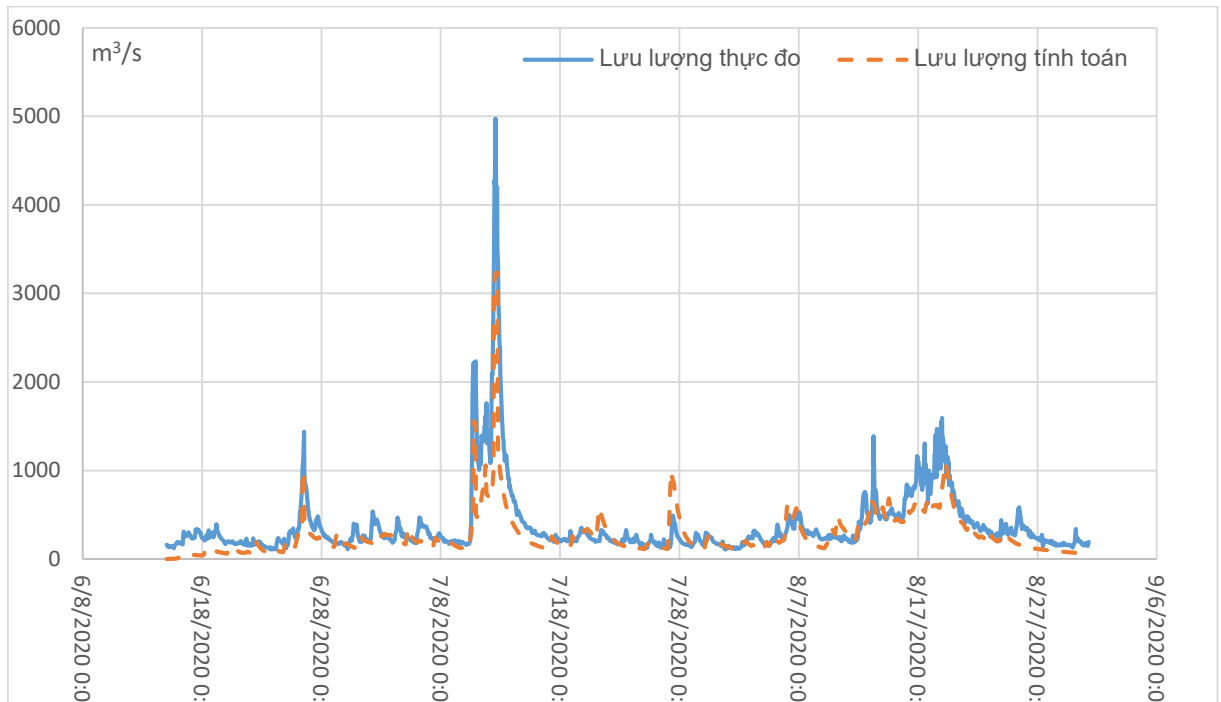
- TBC: là giá trị trung bình cộng NSE của nhiều lần tính.
- Max: Giá trị NSE lớn nhất đạt được.
- TGTB: Thời gian trung bình của 01 lần tính.
- Các ô có giá trị trống: $NSE < 0$, phương pháp không hội tụ.

Đối với lưu vực Lâm Sơn, phương pháp ô vuông đã được thử nghiệm bằng nhiều lần thay đổi số ô lưới được chia. Khi ô lưới tăng từ 3 đến 10, chỉ số NSE cũng chỉ đạt 0.39 nhưng thời gian tính toán đã rất lớn. Cần phải chia ô lưới mịn hơn nữa để cực trị địa phương đủ rộng hơn mắt lưới, tuy nhiên ứng dụng trong thực tế là khó vì thời gian tính toán tăng lên nhiều lần.

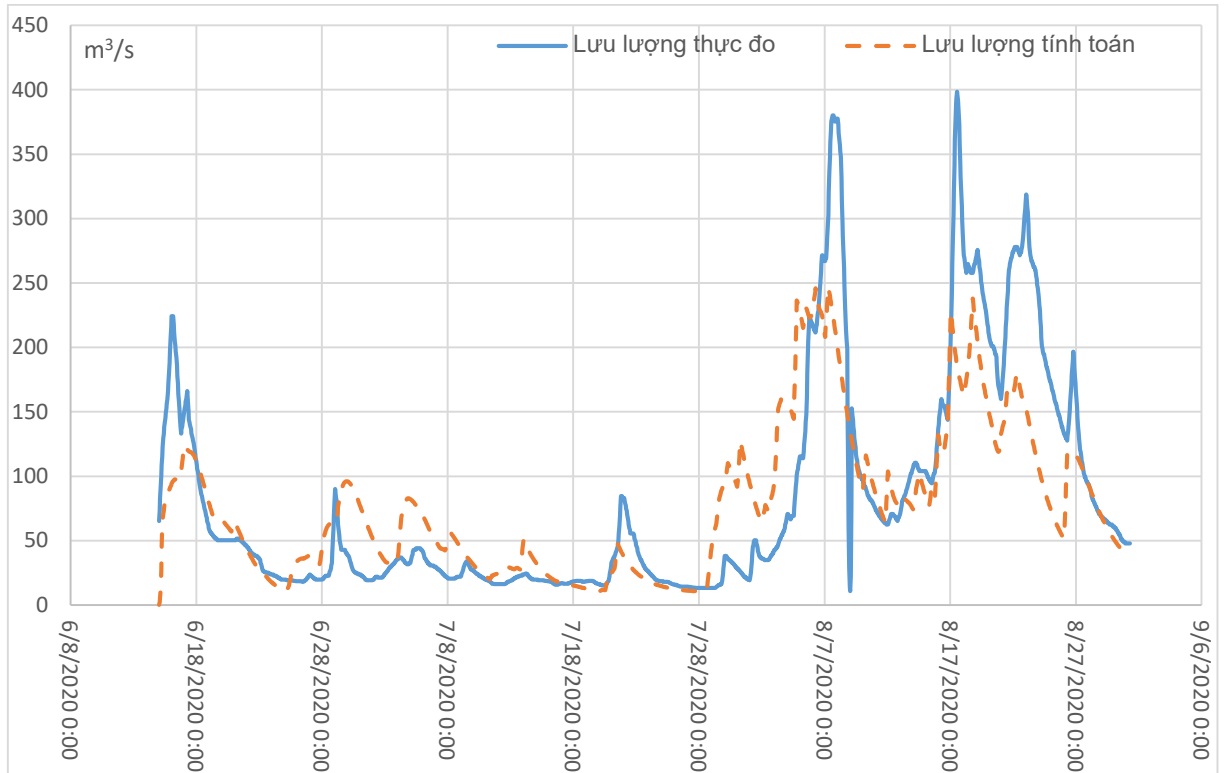
Từ hình 4.2 đến hình 4.13 dưới đây là đồ thị so sánh kết quả giữa lưu lượng thực đo và lưu lượng nước tính toán kiểm định cho 12 lưu vực thủy văn được lựa chọn để tính toán thử nghiệm.



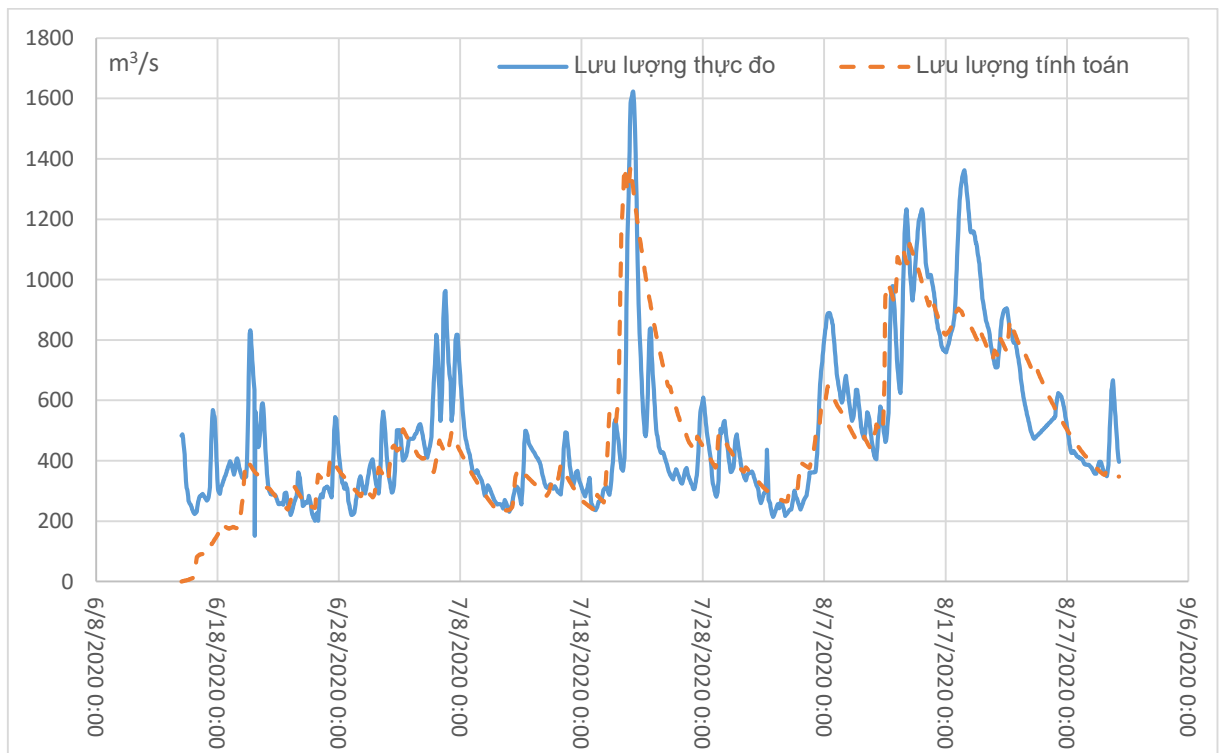
Hình 4.2: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Lâm Sơn



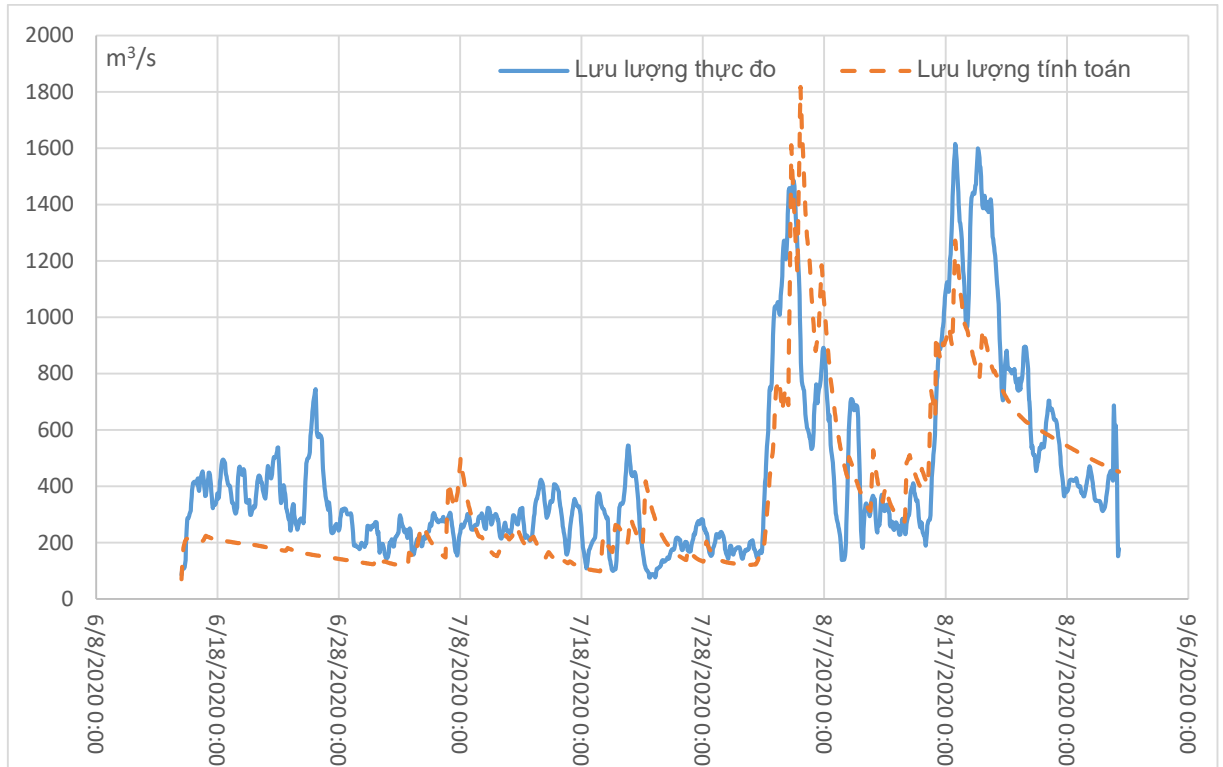
Hình 4.3: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực hồ Bản Chát



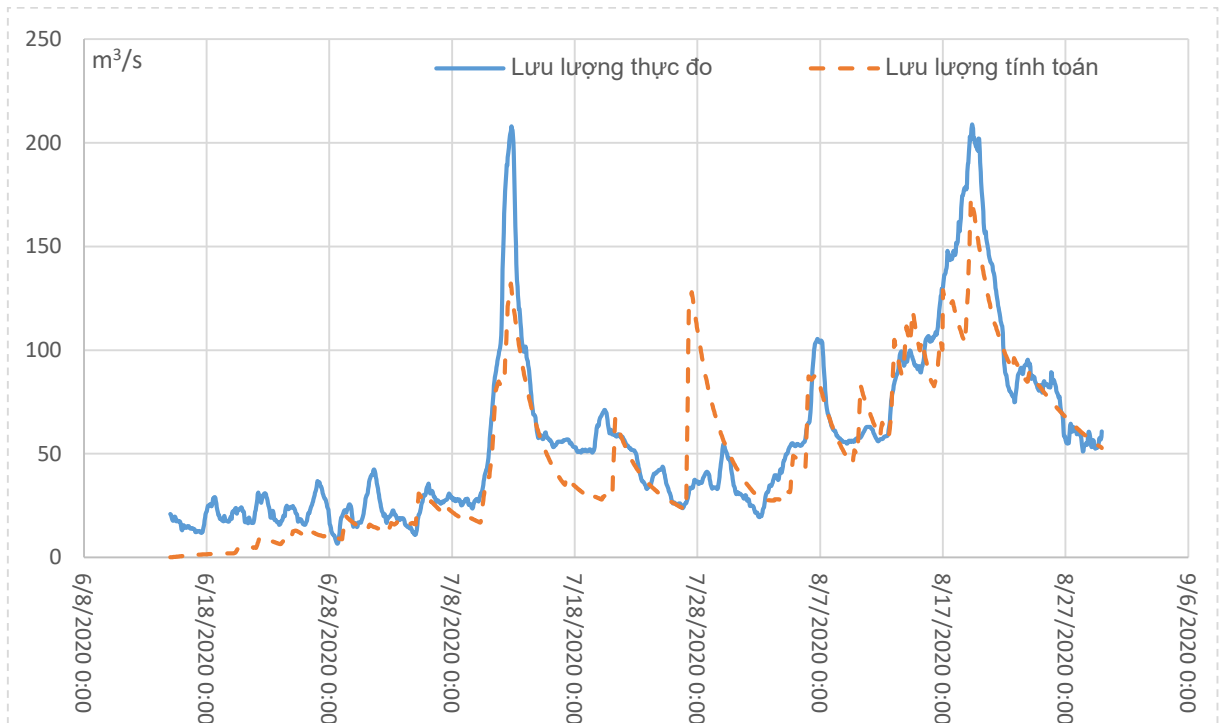
Hình 4.4: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực sông Cầu



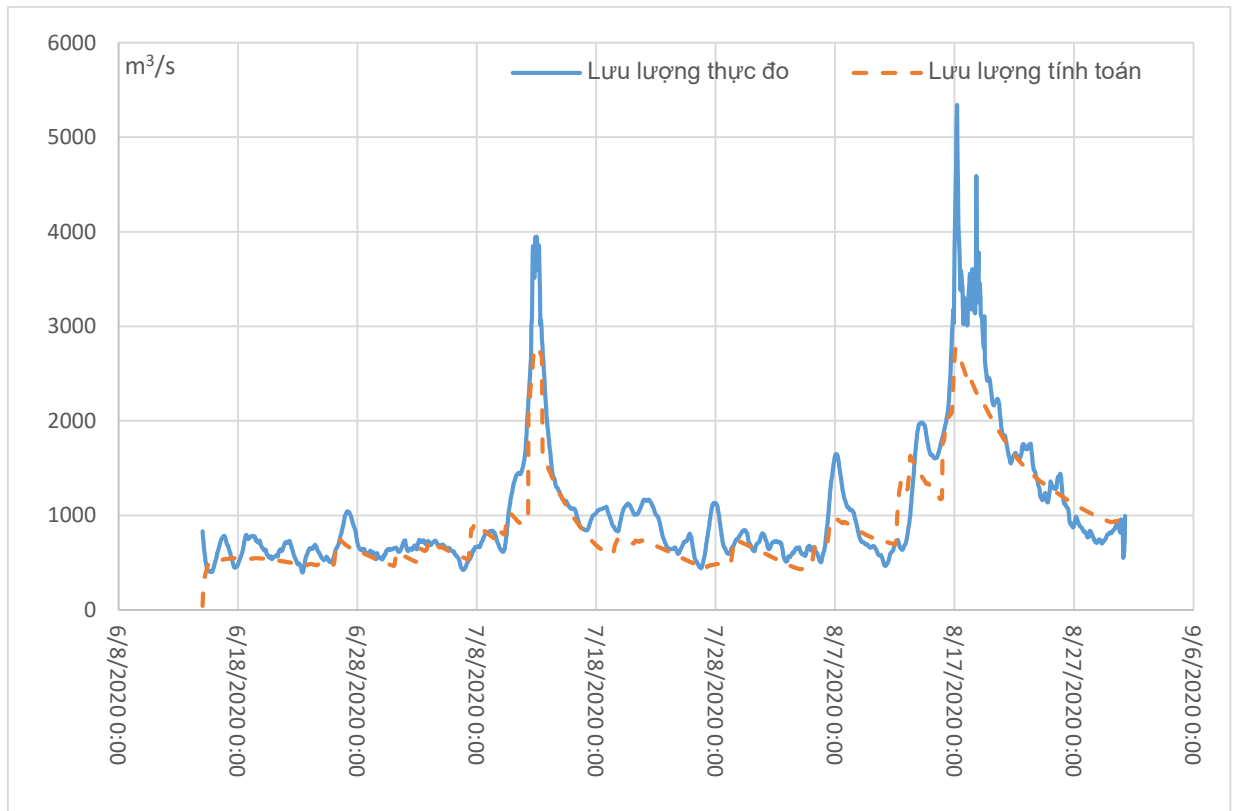
Hình 4.5: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Hàm Yên



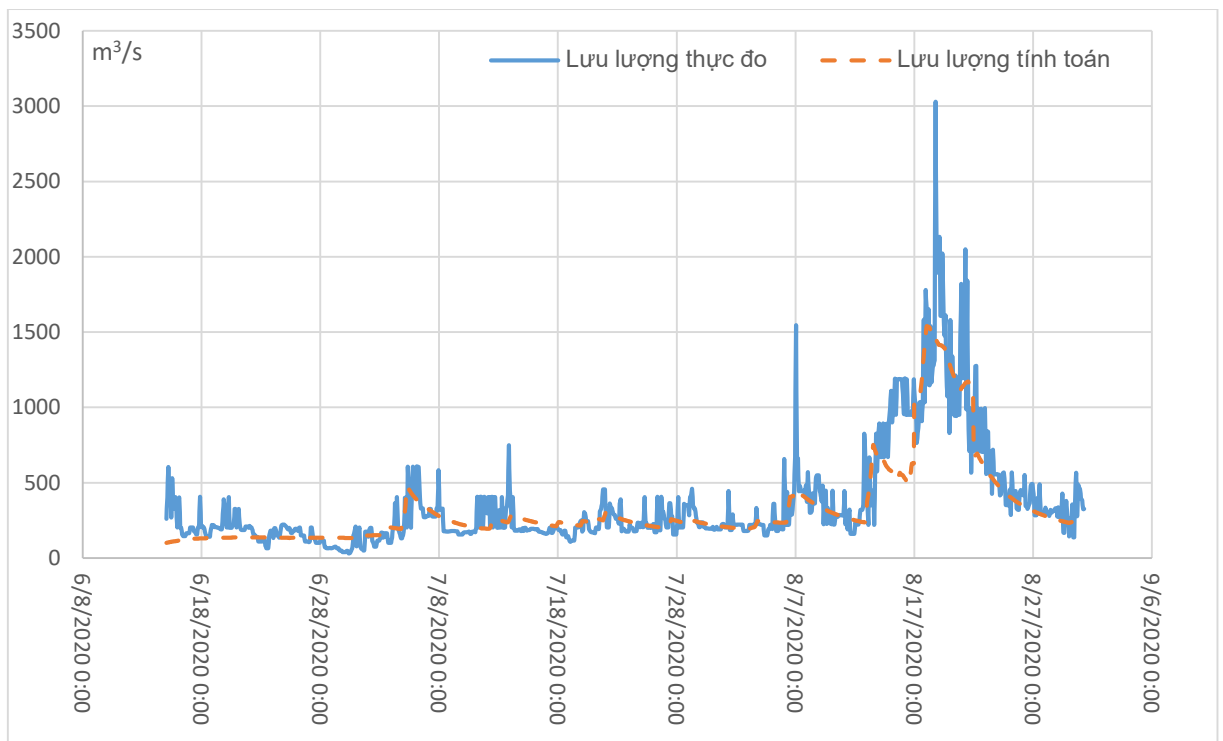
Hình 4.6: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Hòa Bình



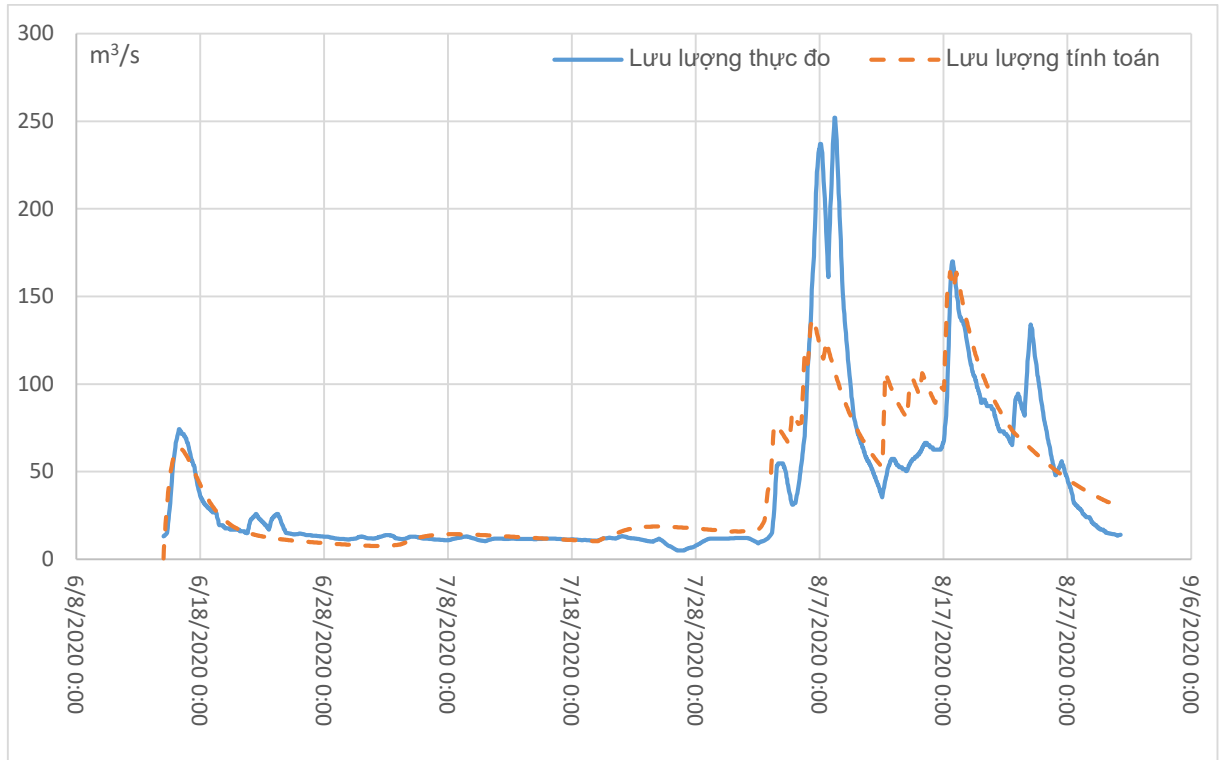
Hình 4.7: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Huội Quảng



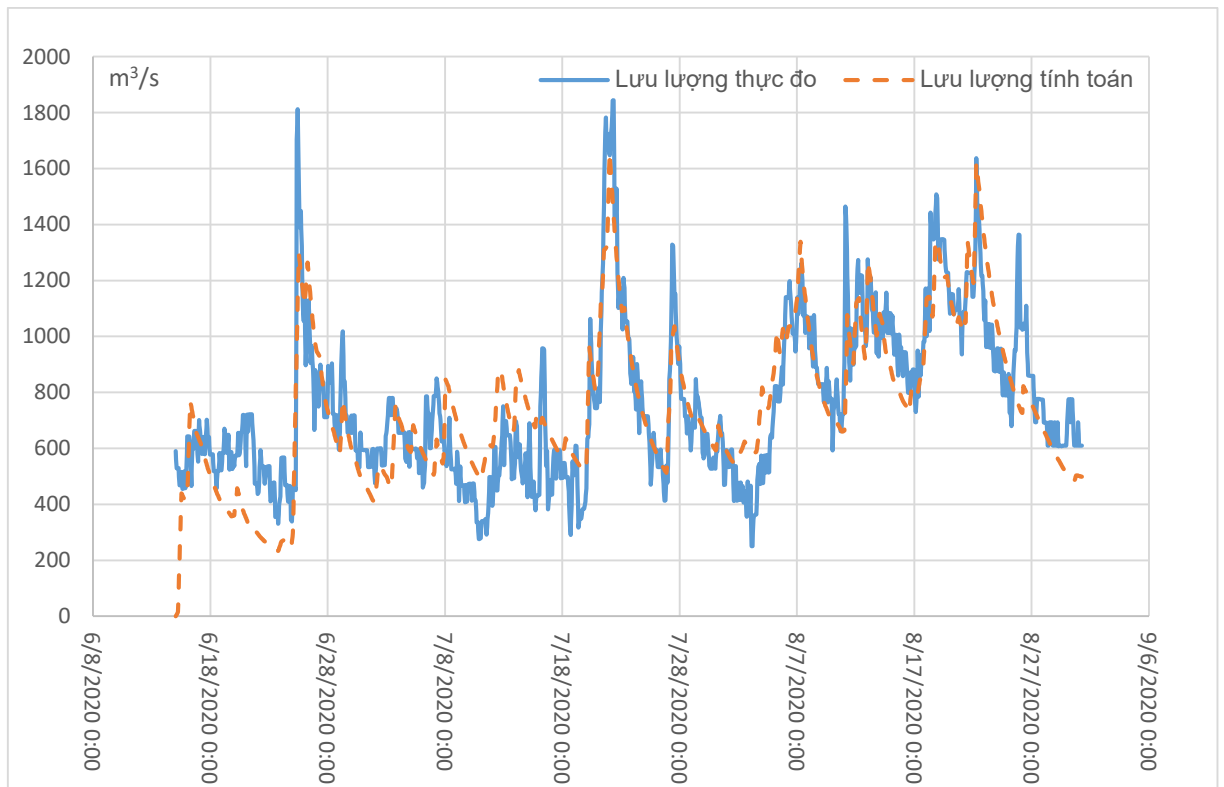
Hình 4.8: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Sơn La



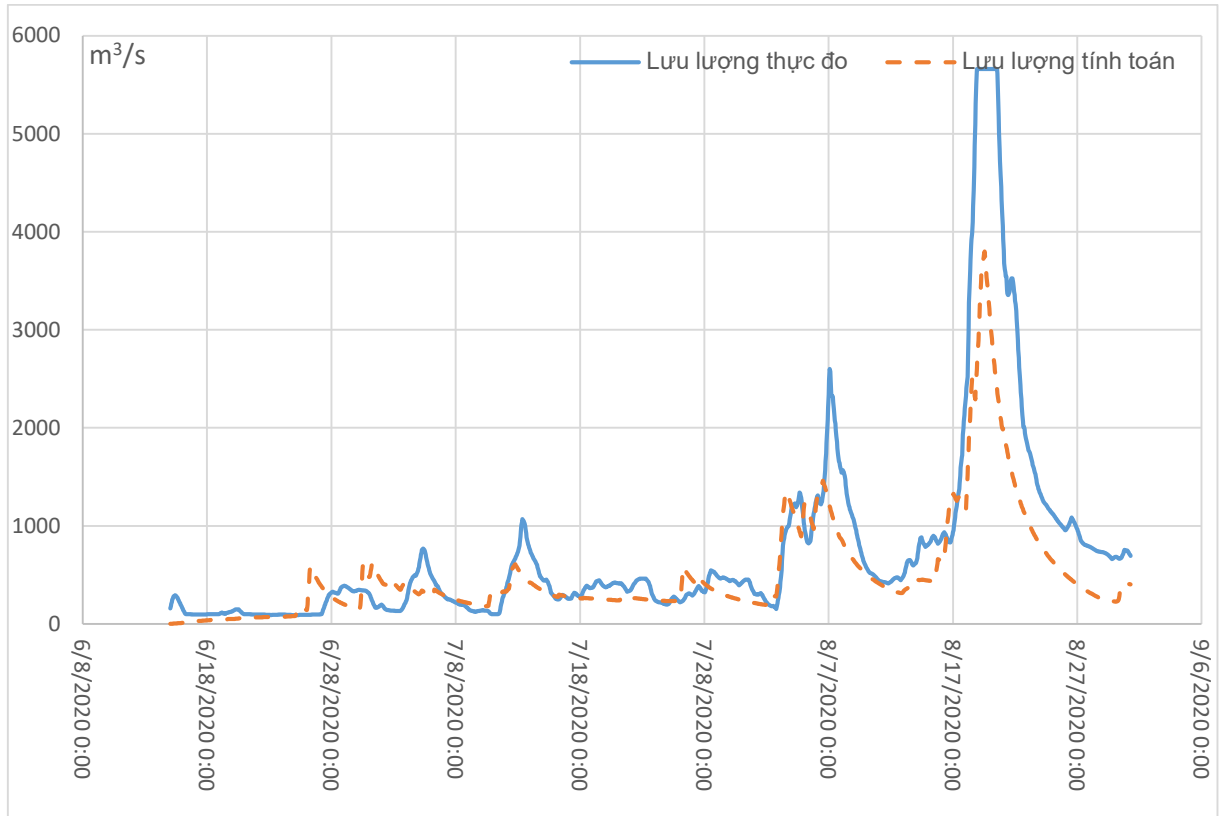
Hình 4.9: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Thác Bà



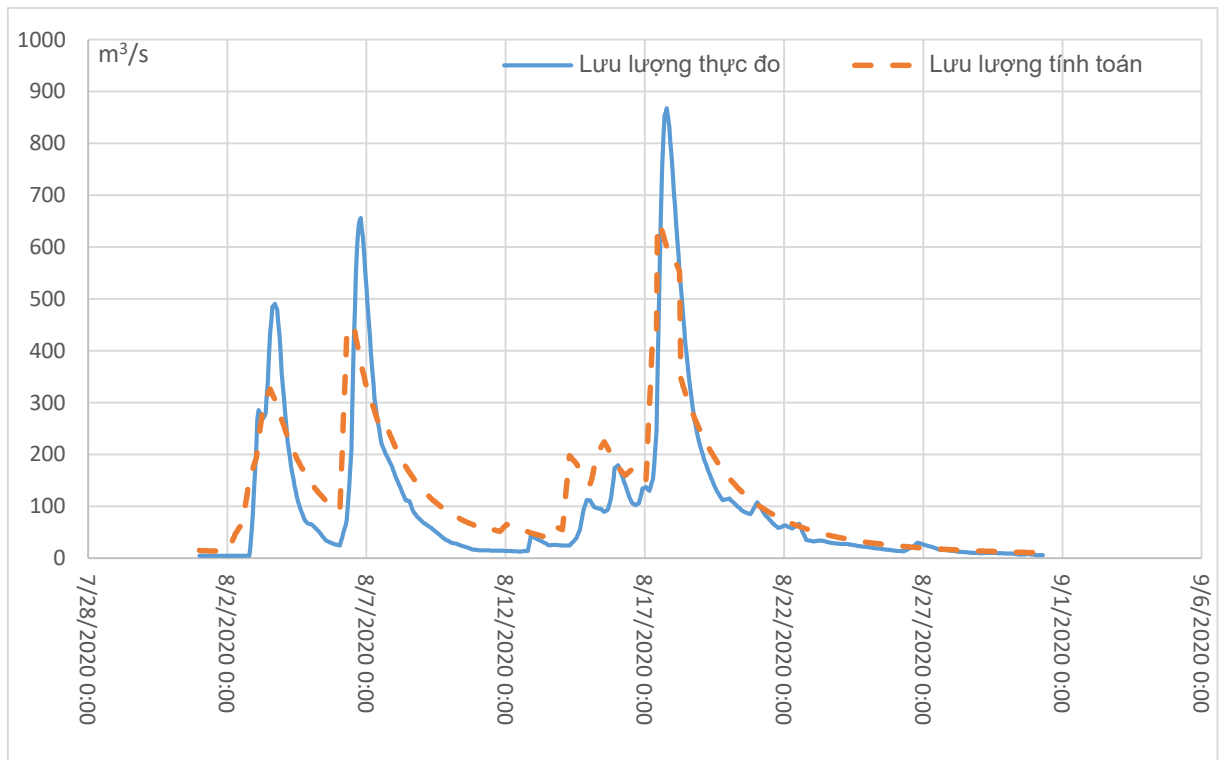
Hình 4.10: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực sông Thương



Hình 4.11: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Tuyên Quang



Hình 4.12: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Yên Bái



Hình 4.13: Đồ thị lưu lượng thực đo và lưu lượng tính kiểm định của lưu vực Lục Nam.

Đánh giá kết quả:

- Dựa trên tiêu chí thời gian tính: Phương pháp RosenBrock nhanh nhất và phương pháp SCE thường lâu nhất.
- Dựa trên tiêu chí kết quả tốt nhất: Trong 12 lưu vực, phương pháp SCE thường cho kết quả tính toán tốt nhất (chỉ số NSE cao nhất), đặc biệt tại lưu vực tiếp nhận lượng nước từ các lưu vực thượng nguồn đổ vào như Huội Quảng, Sơn La, Hoà Bình, khi các phương pháp khác không hội tụ (chỉ số NSE <0) thì phương pháp SCE cho kết quả chấp nhận được, điều này cũng nói lên việc phân tách nước từ mưa và từ nguồn vào khác chưa tốt.
- Dựa trên tiêu chí “chi phí tính toán”: Các lưu vực như Cầu, Thương, Lục Nam, Thác Bà, Tuyên Quang, Yên Bái có kết quả NSE của phương pháp Hooke Jeeves cũng gần tương đương với SCE, nhưng thời gian tính toán là thấp hơn đáng kể (thường chỉ mất một nửa thời gian so với khi tính SCE).
- Xét độ hội tụ thì phương pháp SCE có kết quả giữa các lần tính khá sát nhau và gần với các giá trị tốt nhất tìm được.

KẾT LUẬN VÀ KIẾN NGHỊ

Nội dung nghiên cứu của đề tài được định hướng vào nghiên cứu phát triển mô hình thủy văn thông số tập trung trong dự báo lũ cho một số lưu vực sông ở Việt Nam bằng việc áp dụng các phương pháp số vào giải bài toán ước tính thông số tối ưu. Sử dụng các kỹ thuật tính toán song song tiên tiến giúp giảm thời gian xuống hàng chục lần, thử nghiệm nhiều lần với nhiều thuật toán dò tìm cho nhiều lưu vực khác nhau, nghiên cứu này đã giúp cho quá trình tác nghiệp dự báo được tốt hơn khi có thể tận dụng các ưu nhược điểm của từng phương pháp để áp dụng cho từng lưu vực cụ thể.

Các đóng góp khoa học của đề tài

- Tổng quan các nội dung cơ bản về một số mô hình thủy văn phổ biến hiện nay.
- Giới thiệu một số phương pháp số giải bài toán ước tính thông số tối ưu cho mô hình thủy văn thông số tập trung.
- Xây dựng phần mềm dự báo lũ bằng mô hình thủy văn thông số tập trung có sử dụng phương pháp ước tính thông số tối ưu.
- Thử nghiệm dự báo lũ cho các lưu vực thủy văn đạt kết quả tốt, bộ thông số sau kiểm định đảm bảo độ tin cậy để tính toán mô phỏng, dự báo trong tương lai của một số lưu vực tính toán.

Hướng phát triển

- Áp dụng phần mềm tính toán trên toàn bộ các lưu vực lớn của Việt Nam từ Bắc vào Nam lần lượt theo số liệu địa hình thu thập được.
- Từ kết quả đầu ra của mô hình thủy văn kết hợp làm đầu vào cho mô hình thủy lực để ứng dụng dự báo lũ toàn diện từ thượng lưu xuống hạ lưu cho từng lưu vực cụ thể.
- Việc chỉ có số liệu vài trạm mưa trên lưu vực có diện tích hàng chục ngàn km^2 nên kết quả tính toán mới chỉ ở mức chấp nhận được, mới đưa ra được bức tranh đặc tính cơ bản, do đó cần phải bổ sung thêm số liệu đo đạc, khi đó mới đủ số liệu phản ánh đặc trưng lưu vực, từ đó, các kết quả tính theo các phương pháp sẽ được nâng lên.

TÀI LIỆU THAM KHẢO

- [1] Ngô Huy Cận, Trần Thu Hà (2002), “Mô hình tính toán điều tiết hồ Hòa Bình”, Tuyển tập công trình khoa học Hội nghị Cơ học toàn quốc lần thứ 7, 30-35.
- [2] Nguyễn Văn Điệp, Ngô Huy Cận và nnk (2002), “Mô hình tính toán quá trình lan truyền lũ trên hệ thống sông Hồng – Thái Bình”, Tuyển tập công trình khoa học Hội nghị Cơ học toàn quốc lần thứ 7, 79-87.
- [3] Nguyễn Văn Điệp, Ngô Huy Cận và nnk (2002), “Công nghệ mô phỏng số đánh giá các đặc trưng thủy lực của bài toán vỡ đập trong hệ thống hồ chứa trên sông Đà”, Tuyển tập công trình khoa học Hội nghị Cơ học toàn quốc lần thứ 7, 88-104.
- [4] Trần Thu Hà, Dương T.T.Hương (2004), “Mô hình mưa rào dòng chảy cho hệ thống sông Thái Bình”, Tuyển tập công trình khoa học Hội nghị Cơ học thủy khí toàn quốc, Hà Tiên.
- [5] Nguyễn Chính Kiên (2017), “Thử nghiệm ứng dụng mạng neuron nhân tạo trong dự báo thủy văn và thủy lực”, Hội nghị Cơ học toàn quốc lần thứ X, Hà Nội.
- [6] H.T.M.Linh và N.Đ.Hạnh (2013), “Tối ưu hóa một số thông số của mô hình mưa dòng chảy bằng phương pháp SCE”, Tạp chí khoa học ĐHQGHN, Khoa học Tự nhiên và Công nghệ, Tập 29 (2S), 113-220.
- [7] Lê Xuân Toàn (2015), “Một số phương pháp tối ưu không dùng đạo hàm”, Luận văn thạc sĩ toán học.
- [8] Lê Văn Quế (2017), “So sánh lựa chọn mô hình thích hợp để mô phỏng chế độ thủy văn lưu vực sông Vu Gia - Thu Bồn”, Luận văn thạc sĩ.
- [9] Juraj Cunderlik (2003), Project report “Hydrologic model selection for the CFCAS project”.
- [10] Nielsen S. A. và Hansen E. (1973). Numerical simulation of the rainfall-runoff process on a daily basis. *Nordic Hydrology* 4, 171-190.
- [11] WMO (2013), Intercomparison Of Flood Forecasting Models.
- [12] Mô hình toán thủy văn, Giáo trình, Đại học Thủy Lợi, 2005.
- [13] Nguyễn Chính Kiên, Nguyễn Thị Hằng. “Thử nghiệm một số phương pháp số giải bài toán ước tính thông số tối ưu cho mô hình thủy văn”, Hội nghị 45 năm thành lập Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

- [14] DHI (2014), *Reference Manual MIKE 11*
- [15] DHI (2014), *User's Manual MIKE 11*
- [16] US Army Corps of Engineers, Hydrological Model systems HEC-HMS User's Manual Version 4.2.1, 2015.
- [17] US Army Corps of Engineers, HEC-RAS river Analysis system User's Manual Version 5.0.7, 2016