สำนักงานวิทยทรัพยากร จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

วิธีดำเนินการศึกษา

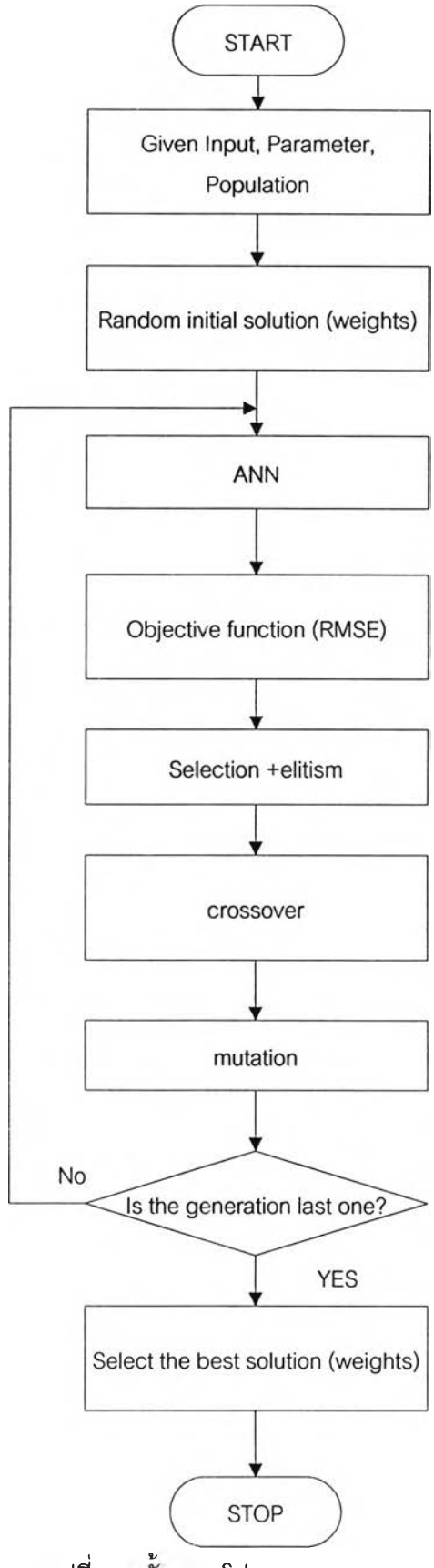
ในบทนี้จะกล่าวถึงรายละเอียดวิธีการดำเนินการศึกษาและขั้นตอนในการดำเนินงานวิจัย ประกอบด้วย 1) การพัฒนาโปรแกรม BPNN GA+ANN และ GA+BPNN 2) การตรวจสอบความถูกต้อง 3) การคัดเลือกโครงสร้างและพารามิเตอร์ GA ที่เหมาะสม และ 4) การเปรียบเทียบผลการพยากรณ์ โดยมีรายละเอียดดังต่อไปนี้

5.1 การพัฒนาโปรแกรม BPNN

โปรแกรม BPNN ที่พัฒนาขึ้นมาเป็นโปรแกรมจำลองขั้นตอนการทำงานของโครงข่ายใยประสาทเทียมที่มีโครงข่ายหลายชั้น ได้แก่ ชั้นนำเข้า ชั้นแอบแฝง และชั้นผลลัพธ์ ในแต่ละชั้นประกอบด้วยโหนด (node) ที่ทำหน้าที่คำนวณผลลัพธ์จากข้อมูลนำเข้าที่โหนดได้รับมาจากโหนดที่อยู่ติดกันในชั้นก่อนหน้า โหนดแต่ละชั้นจะเชื่อมต่อกับโหนดในชั้นถัดไปแบบทั่วถึง เส้นเชื่อมต่อระหว่างโหนดในแต่ละชั้นด้วยค่าถ่วงน้ำหนัก (Weight, W) โปรแกรม BPNN มีวิธีหาค่าถ่วงน้ำหนักที่เหมาะสมด้วยการคำนวณซ้ำ (Iterative) จำนวนมาก เพื่อให้ได้ค่าถ่วงน้ำหนักที่ทำให้ผลลัพธ์จากการคำนวณมีค่าใกล้เคียงกับค่าข้อมูลที่วัดจริงมากที่สุด วิธีการดังกล่าวเรียกว่า วิธีการเรียนรู้แบบย้อนกลับ (Back Propagation, BP) คือ ค่าความผิดพลาดในการคำนวณจากโหนดชั้นผลลัพธ์จะส่งย้อนกลับไปให้โหนดในชั้นแอบแฝงเพื่อทำการปรับแก้ค่าถ่วงน้ำหนักให้มีค่าความผิดพลาดน้อยลง ขั้นตอนการคำนวณหลักของวิธีการเรียนรู้แบบย้อนกลับมี 2 ขั้นตอน ได้แก่ ขั้นตอนการคำนวณไปข้างหน้า (Feed forward) และ ขั้นตอนการคำนวณย้อนกลับ (Backward) นอกจากนี้ยังมีค่าพารามิเตอร์และวิธีการย่อยอื่นๆ รายละเอียดของขั้นตอนการพัฒนาโปรแกรม BPNN และข้อกำหนดและการเลือกใช้พารามิเตอร์ดังกล่าวมาแล้วในบทที่ 3 หัวข้อ 3.3

5.2 การพัฒนาโปรแกรม GA+ANN

โปรแกรม GA+ANN ที่พัฒนาขึ้นมาเป็นการบูรณาการการประยุกต์ใช้ GA ร่วมกับ ANN โดยใช้ GA ทำการเรียนรู้หาค่าเหมาะสมของค่าถ่วงน้ำหนักของ ANN แทนการเรียนรู้แบบย้อนกลับ (BP) ด้วยขั้นตอนของ GA ที่มีอยู่ 4 ขั้นตอน ได้แก่ การสร้างพันธุ์ คัดเลือกพันธุ์ (Selection) ผสมข้ามพันธุ์ (Crossover) และปรับปรุงพันธุ์ (Mutation) โดยใช้โครงสร้างของ ANN ให้ใช้ชุดเดียวกับโครงสร้างของ BPNN เพียงต่างกันตรงที่การเรียนรู้หาค่าถ่วงน้ำหนักสุดท้ายเท่านั้น ดังมีขั้นตอนการคำนวณหลักดังรูปที่ 5-1 และขั้นตอนในการพัฒนาโปรแกรมที่อธิบายได้ดังนี้



รูปที่ 5-1 ขั้นตอนโปรแกรม GA+ANN

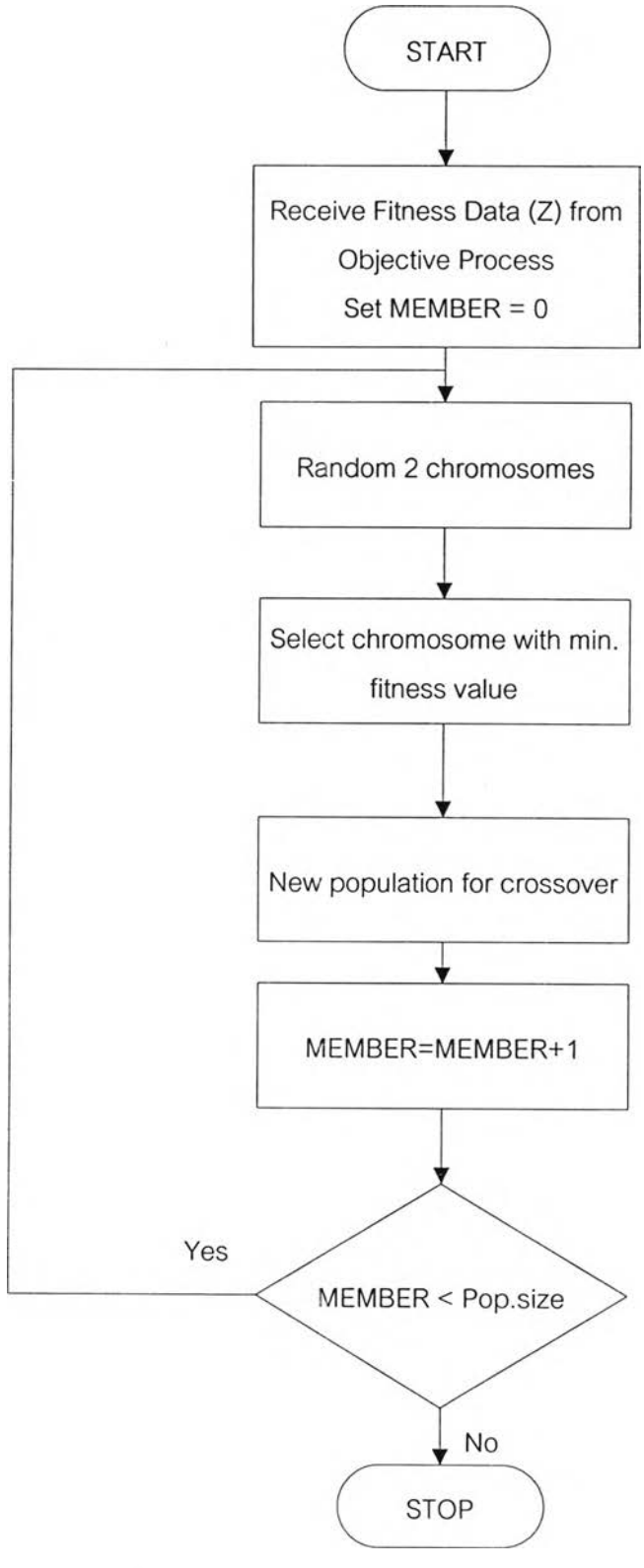
5.2.1 รูปแบบในการสร้างพันธุ์

การสร้างพันธุ์ในขั้นตอนของ GA เป็นการสร้างประชากรของค่าถ่วงน้ำหนักหรือการสร้างชุดของค่าถ่วงน้ำหนักของ ANN หลายชุดด้วยวิธีการสุ่มค่าถ่วงน้ำหนักเช่นเดียวกับ BPNN แต่จะมีชุดค่าถ่วงน้ำหนักมากกว่าชุดค่าถ่วงน้ำหนักของ BPNN ที่มีเพียงชุดเดียว โปรแกรม GA+ANN สามารถสร้างชุดค่าถ่วงน้ำหนักได้หลายชุดขึ้นอยู่กับขนาดของประชากร (Population Size) ที่กำหนด

โดยทั่วไปการสร้างประชากรของ GA สามารถสร้างประชากรได้ในหลายรูปแบบ สำหรับการพัฒนาโปรแกรม GA+ANN เป็นการสร้างประชากรของค่าถ่วงน้ำหนักซึ่งมีค่าเป็นเลขจำนวนจริง ซึ่งวิธีที่ใช้ในการสร้างพันธุ์ที่เป็นเลขจำนวนจริงตามวิธีของ GA สามารถสร้างได้ 2 รูปแบบ คือ รูปแบบเลขฐานสอง (Binary Coding) หรือ รูปแบบ 0 กับ 1 และรูปแบบที่เป็นเลขจำนวนจริง (Real Value Coding) แต่เนื่องจากที่ผ่านมาไม่สามารถกำหนดช่วงพิสัยของค่าถ่วงน้ำหนักสุดท้ายที่เหมาะสม ทำให้รูปแบบการสุ่มในรูปแบบเลขฐานสองที่จำกัดช่วงของพิสัยของค่าถ่วงน้ำหนักไม่เหมาะสมกับปัญหาที่ไม่ทราบขอบเขตของเซตคำตอบสุดท้าย ในการศึกษานี้จึงได้เลือกใช้รูปแบบการสุ่มค่าถ่วงน้ำหนักเป็นเลขจำนวนจริง ซึ่งมีขั้นตอนในการสุ่มเช่นเดียวกับการสุ่มค่าถ่วงน้ำหนักเริ่มต้นของ BPNN

5.2.2 การคัดเลือกพันธุ์

จากรูปที่ 5-1 แสดงขั้นตอนการทำงานหลักของ GA+ANN หลังจากสร้างประชากรที่เป็นชุดค่าถ่วงน้ำหนักจำนวนเท่ากับขนาดของประชากร ชุดค่าถ่วงน้ำหนักทุกชุดจะนำไปใช้กับโครงข่ายใยประสาทเทียม (ANN) เพื่อคำนวณผลลัพธ์จากชุดข้อมูลนำเข้า ผลลัพธ์ทุกชุดที่ได้จาก ANN นำไปเปรียบเทียบกับค่าที่วัดจริงและคำนวณค่าฟังก์ชั่นวัตถุประสงค์ที่กำหนดให้เท่ากับ ค่า RMSE ระหว่างข้อมูลผลลัพธ์กับข้อมูลจริงในช่วงการเรียนรู้ ชุดค่าถ่วงน้ำหนักที่ให้ผลลัพธ์มีความถูกต้องมากสุดจากจำนวนชุดค่าถ่วงน้ำหนักทั้งหมด หรือ มีค่า RMSE ต่ำสุด มีโอกาสที่ถูกคัดเลือกมากหรือถูกคัดเลือกซ้ำหลายครั้ง เพื่อนำไปใช้ในการผสมข้ามพันธุ์ต่อไป วิธีการคัดเลือกที่ใช้ในโปรแกรม GA+ANN คือ วิธีการคัดเลือกที่เรียกว่า Tournament Selection โดยมีวิธีดังนี้คือ สุ่มเลือกชุดของค่าถ่วงน้ำหนักขึ้นมา 2 ชุด มาเปรียบเทียบกัน ค่าถ่วงน้ำหนักชุดที่มีความถูกต้องมากกว่า หรือมีค่า RMSE ต่ำกว่าจะถูกคัดเลือก ไปใช้สร้างชุดคำตอบชุดใหม่ในขั้นต่อไป ทำการสุ่มจับคู่คัดเลือกซ้ำจนได้จำนวนชุดค่าถ่วงน้ำหนักเท่ากับจำนวนชุดคำตอบที่กำหนดไว้หรือขนาดประชากร ดังรูป 5-2



รูปที่ 5-2 ขั้นตอนการทำงานของโปรแกรมย่อย Selection

5.2.3 <u>การสร้างชุดคำตอบชุดใหม่</u>

ชุดค่าถ่วงน้ำหนักที่ได้รับการคัดเลือกถูกนำไปใช้สร้างชุดค่าถ่วงน้ำหนักชุดใหม่ ด้วยขั้นตอน 2 ขั้นตอน คือ การผสมข้ามพันธุ์ (Crossover) และการปรับปรุงพันธุ์ (Mutation) ซึ่งมีขั้นตอนและวิธีที่เลือกใช้ ดังนี้

(1) วิธีผสมข้ามพันธุ์

วิธีผสมข้ามพันธุ์เป็นวิธีการสลับพันธุ์ระหว่างชุดค่าถ่วงน้ำหนักคู่หนึ่งที่ถูกเลือกมาจับคู่กัน โดยการแลกเปลี่ยนค่าถ่วงน้ำหนักระหว่างค่าถ่วงน้ำหนักชุดหนึ่งกับอีกชุดหนึ่งในการศึกษานี้ได้ทดสอบหาวิธีผสมข้ามพันธุ์ที่เหมาะสมที่สุดจาก 5 วิธี ประกอบด้วย 1-point, 2-point, Uniform, Average, Heuristic โดยชุดค่าถ่วงน้ำหนักแต่ละคู่จะได้มาจากการสุ่มเลือกดังรูปที่ 5-3 และอธิบายได้ดังนี้

(ก) สุ่มค่า Prob_X ระหว่าง 0 ถึง 1 ให้กับทุกโครโมโซม

(ข) ถ้าโครโมโซมตัวใดมีค่า Prob_X < Pc เลือกโครโมโซมตัวนั้นเก็บไว้จับคู่

(ค) จับคู่โครโมโซมเป็นคู่ๆ

(ง) นำโครโมโซมแต่ละคู่เข้าสู่วิธีการ Crossover เพื่อสลับค่ายีนระหว่างคู่

(จ) นำโครโมโซมทั้งหมดเข้าสู่กระบวนการ Mutation ต่อไป

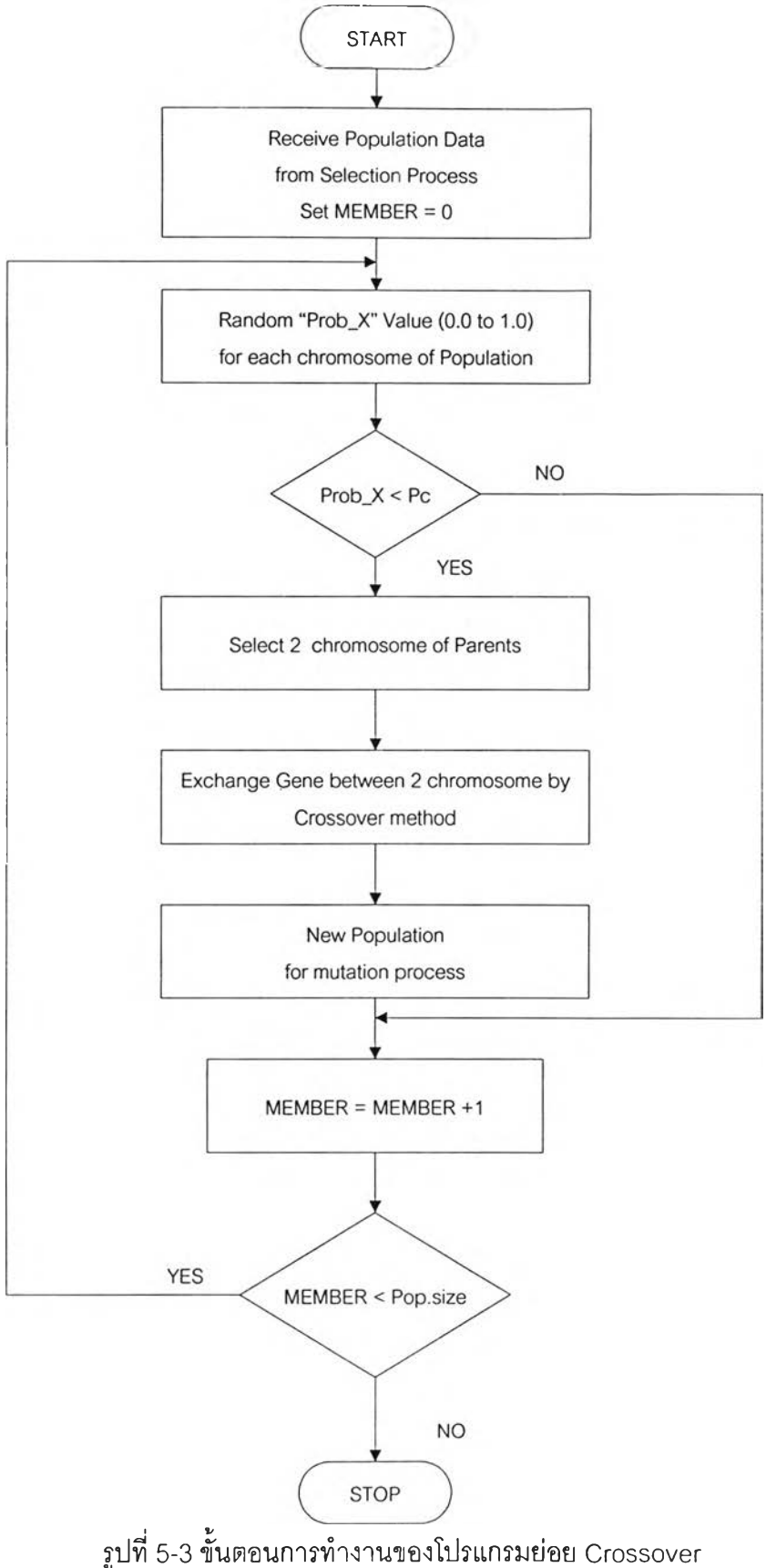
(2) วิธีปรับปรุงพันธุ์

ชุดคำตอบที่ผ่านการผสมข้ามพันธุ์ ซึ่งประกอบด้วย ชุดคำตอบชุดใหม่ แต่ยังมีค่าอยู่ในกลุ่มของชุดคำตอบเริ่มต้น วิธีการปรับปรุงพันธุ์จะทำการเปลี่ยนแปลงค่าถ่วงน้ำหนักบางค่าให้มีค่าแตกต่างจากชุดคำตอบเริ่มต้น ดังรูปที่ 5-4 โดยมีวิธีการดังนี้

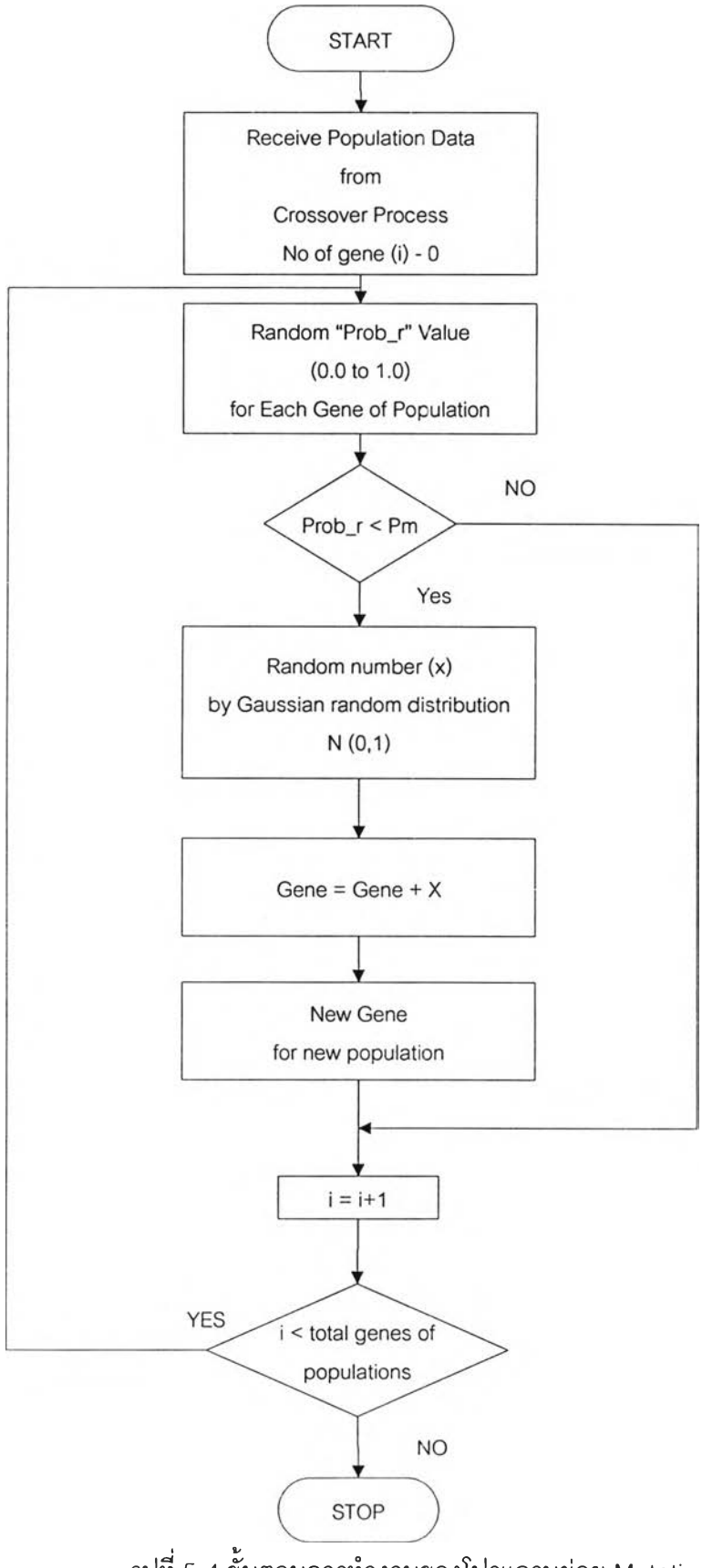
(ก) สุ่มค่า Prob_r ระหว่าง 0 ถึง 1 ให้กับทุกยีนในจำนวนประชากรทั้งหมด

(ข) ถ้ายีนตัวใดมีค่า Prob_r < Pm ทำการปรับปรุงค่ายีนที่ตำแหน่งนั้นด้วยการเพิ่มหรือลดค่าดังกล่าว ด้วยค่าที่ได้จากการสุ่มซึ่งมีการกระจายแบบโค้งปกติ Normal Distribution เรียกว่า วิธี Gaussian Mutation

นำโครโมโซมทั้งหมดที่ผ่านกระบวนการของ GA ไปคำนวณซ้ำในหัวข้อ 5.2.2 ต่อไปจนครบจำนวนรอบการคำนวณสูงสุดหรือค่าความเหมาะสมของโครโมโซมที่ดีที่สุดมีค่าน้อยกว่าค่าที่ยอมรับได้



รูปที่ 5-3 ขั้นตอนการทำงานของโปรแกรมย่อย Crossover



รูปที่ 5-4 ขั้นตอนการทำงานของโปรแกรมย่อย Mutation

5.3 การพัฒนาโปรแกรม GA+BPNN

โปรแกรม GA+BPNN ที่พัฒนาการบูรณาการการประยุกต์ใช้ GA ร่วมกับ BPNN โดยใช้ GA จะสร้างชุดค่าถ่วงน้ำหนักขึ้นมาหลายชุด ค่าถ่วงน้ำหนักแต่ละชุดจากการสร้างด้วย GA จะเป็นค่าถ่วงน้ำหนักเริ่มต้นของ BPNN จากนั้นแบบจำลอง BPNN จะปรับค่าถ่วงน้ำหนักในแต่ละชุดด้วยวิธีย้อนกลับตามจำนวนรอบการคำนวณซ้ำของ BPNN เพื่อหาค่าถ่วงน้ำหนักสุดท้าย และใช้ค่าถ่วงน้ำหนักสุดท้ายจากการเรียนรู้ด้วย BPNN แต่ละชุด คำนวณค่าผลลัพธ์เทียบกับข้อมูลจริงแล้วเปรียบเทียบความผิดพลาดของทุกชุดเป็นค่า RMSE แล้วนำกลับไปเข้าใน GA และเข้าสู่ขั้นตอนการ คัดเลือกพันธุ์ การผสมข้ามพันธุ์และการปรับปรุงพันธุ์ตามลำดับ เพื่อสร้างค่าถ่วงน้ำหนักเริ่มต้นสำหรับ BPNN ในรอบถัดไป ทำการคำนวณซ้ำจนครบจำนวนรอบการคำนวณสูงสุด ดังรูปที่ 5-5

5.4 ตรวจสอบความถูกต้องของโปรแกรม BPNN

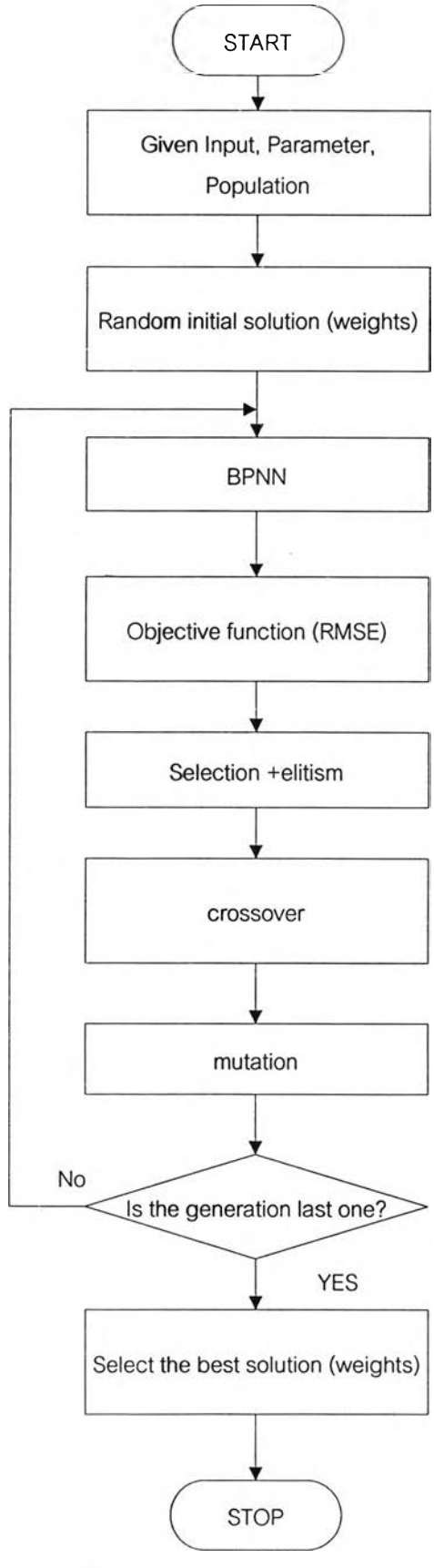
ขั้นตอนการคำนวณของโปรแกรม BPNN มีความสลับซับซ้อนและมีค่าพารามิเตอร์ที่ใช้ในการสร้างแบบจำลองเป็นจำนวนมาก ทำให้ผลการคำนวณอาจมีความผิดพลาดได้ เพื่อตรวจสอบความถูกต้องของโปรแกรม BPNN ที่พัฒนาขึ้น ในการขั้นตอนนี้จะเลือกแบบจำลองสำเร็จรูป BPNN ชื่อ Qnet2000 ซึ่งได้รับความนิยมและเชื่อถือในงานศึกษาและวิจัยทั่วไป มาปรับเทียบแบบจำลอง โดยนำผลลัพธ์จากแบบจำลองสำเร็จรูป Qnet2000 และผลลัพธ์ของโปรแกรม BPNN ที่พัฒนามาเปรียบเทียบความถูกต้องทางสถิติและชลภาพน้ำท่ากับข้อมูลที่วัดจริง

5.5 การคัดเลือกขั้นตอนและพารามิเตอร์ของ GA เหมาะสม

เป็นการศึกษาคัดเลือกขั้นตอนและพารามิเตอร์ของ GA ที่เหมาะสมสำหรับแบบจำลอง ANN และ BPNN ในการพยากรณ์น้ำท่าในฤดูฝนและฤดูแล้ง และผลการเลือกใช้ขั้นตอนและพารามิเตอร์ GA แบบต่างๆ ต่อความถูกต้องของการพยากรณ์ของทุกแบบจำลอง

5.6 การประยุกต์ใช้ในการพยากรณ์

นำแบบจำลองที่พัฒนาทั้ง 3 โปรแกรมคือ BPNN, GA+ANN และ GA+BPNN มาใช้ในการพยากรณ์น้ำท่าล่วงหน้า 1-7 วันของสถานีวัดน้ำท่า 6 สถานีในลุ่มน้ำปราจีนบุรี โดยใช้โครงสร้างและพารามิเตอร์ที่ดีที่สุดจากขั้นตอนการคัดเลือกของโปรแกรมทั้งสาม



รูปที่ 5-5 ขั้นตอนโปรแกรม GA+BPNN

5.7 การเปรียบเทียบผลการพยากรณ์ระหว่างโปรแกรมทั้งสาม

เมื่อได้ทำการพัฒนาแบบจำลองและปรับเทียบโปรแกรมทั้งสามเสร็จเรียบร้อยแล้ว จะนำผลการพยากรณ์ที่ได้จากทุกแบบจำลองมาเปรียบเทียบผลการพยากรณ์ด้วยค่าความถูกต้องทางสถิติและซลภาพน้ำท่าเทียบกับข้อมูลที่วัดจริงในช่วงข้อมูลเรียนรู้ (ปรับเทียบ) และช่วงข้อมูลทดสอบ เพื่อพิจารณาถึงความแตกต่างของแบบจำลองทั้งสาม ในการนำมาประยุกต์ใช้สำหรับการพยากรณ์อัตราการไหลน้ำท่ารายวัน