

บทที่ 3

ทฤษฎีที่เกี่ยวข้อง

3.1 โครงข่ายประสาทเทียม

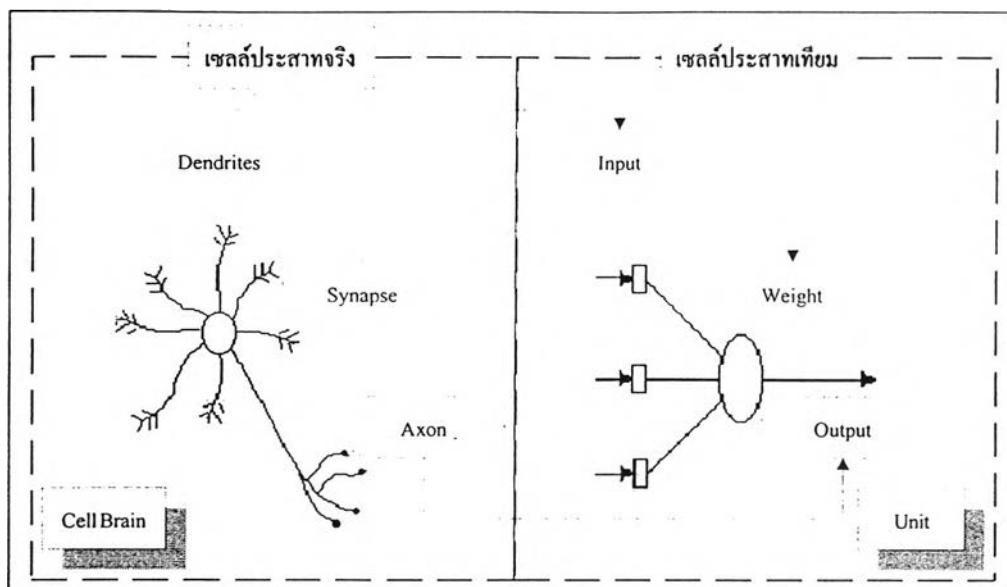
โครงข่ายประสาทเทียมคือระบบประมวลผลข้อมูลที่มีการกระจายการคำนวณแบบขนาน โดยเลียนแบบการทำงานของสมองมนุษย์ ซึ่งประกอบไปด้วยหน่วยประมวลผล (processing element) ซึ่งมีเซลล์หลายๆเซลล์ทำหน้าที่คล้ายกับเซลล์สมองของมนุษย์ โดยที่แต่ละเซลล์จะโยงใยติดต่อกันโดยส่งสัญญาณเป็นตัวแปรออก (output) ผ่านส่วนที่เรียกว่า ไซแนปส์ (Synapses) กลายมาเป็นตัวแปรเข้า (input) ของส่วนที่เรียกว่า เดนไดรต์ (Dendrites) และเมื่อผ่านกระบวนการประมวลผลจะได้ตัวแปรออกออกมาในส่วนที่เรียกว่า แอ็กซอน (Axon) ในแต่ละเซลล์จะรับรู้ข้อมูลจากหลายทาง แล้วส่งต่อไปยังเซลล์อื่นๆ โดยใช้หลักการของการเชื่อมโยงเซลล์สมอง ดังรูปที่ 3.1 สำหรับความสัมพันธ์ระหว่างเซลล์ประสาทกับประสาทเทียม แสดงดังตารางที่ 3.1

หลักการทำงานของโครงข่ายประสาทเทียมจะมีแนวคิดที่แตกต่างกับแบบจำลองทั่วไปโดยสิ้นเชิง ในการจำลองพฤติกรรมโดยการใช้แบบจำลองโครงข่ายประสาทเทียม เราไม่มีความจำเป็นที่จะต้องกำหนดหรือสร้างสมการแต่อย่างใด เพียงแต่เรารวบรวมข้อมูลตัวแปรเข้าและตัวแปรออก แต่ละเหตุการณ์ไว้เป็นคู่ๆ โดยทั่วไปการใช้แบบจำลองคณิตศาสตร์ หรือแบบจำลองทางสถิติจะต้องสร้างความสัมพันธ์ระหว่างตัวแปรเข้าและตัวแปรออก ซึ่งอยู่ในรูปของสมการแต่โครงข่ายประสาทเทียมจะทำการหาความสัมพันธ์ระหว่าง ตัวแปรเข้าและตัวแปรออก โดยกระบวนการเรียนรู้จากข้อมูลจำนวนมากที่มีอยู่ ความสัมพันธ์ดังกล่าวไม่มีการกำหนดในรูปของสมการ

ตารางที่ 3.1 ความสัมพันธ์ระหว่างเซลล์ประสาทกับประสาทเทียม

ลำดับ	เซลล์ประสาท	เซลล์ประสาทเทียม
1	ตัวเซลล์ (Cell body)	ยูนิต (unit)
2	เดนไดรต์ (Dendrites)	ตัวแปรเข้า (input)
3	แอ็กซอน (Axon)	ตัวแปรออก (output)
4	ไซแนปส์ (Synapse)	ค่าน้ำหนัก (weight)
5	ความเร็วในการทำงานช้า	ความเร็วในการทำงานสูง
6	มีเซลล์จำนวนมาก (10^9 ยูนิต)	มีเซลล์จำนวนน้อยกว่า

ที่มา : เสรี, 2544



ที่มา: เสรี, 2544

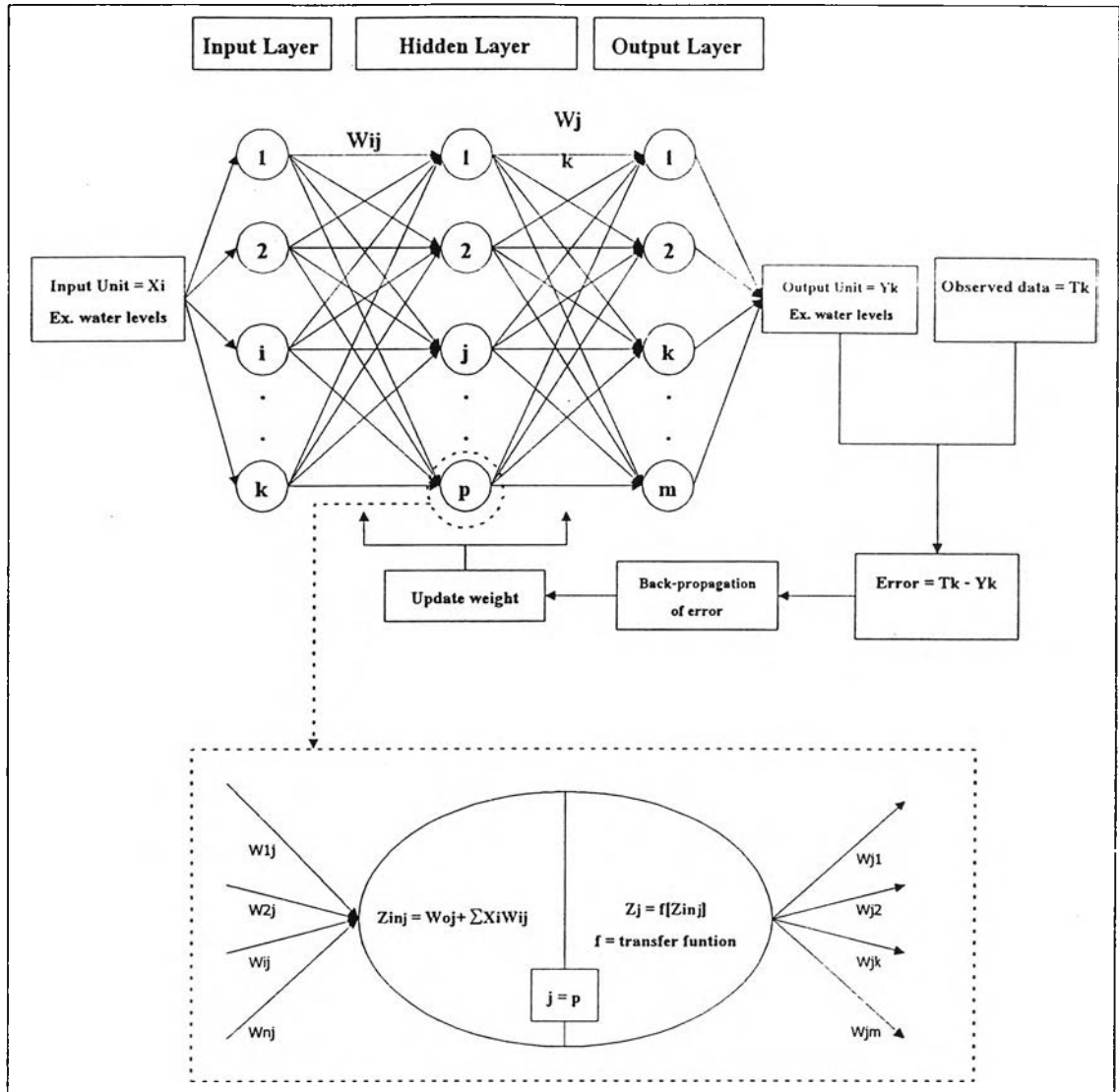
รูปที่ 3.1 เซลล์ประสาทและเซลล์ประสาทเทียม

ในการศึกษาครั้งนี้ ใช้โครงข่ายใยประสาทเทียมชนิดวิธีปรับแก้ค่าย้อนกลับ (Back Propagation Neural Network, BPNN) โดยโครงสร้างประกอบด้วย ชั้นตัวแปรเข้า (input layer) ชั้นตัวแปรซ่อน (hidden layer) และชั้นตัวแปรออก (output layer) ในแต่ละชั้นประกอบด้วยหน่วยย่อย (nodes) ซึ่งแต่ละหน่วยของแต่ละชั้นนั้นจะถูกเชื่อมโยงด้วยค่าน้ำหนัก (weight) ลักษณะการทำงานจะทำงานแบบเคลื่อนที่ไปข้างหน้าอย่างเดียวแต่มีการปรับแก้ค่าน้ำหนักย้อนกลับซึ่งแสดงรายละเอียดในหัวข้อถัดไปและแสดงโครงสร้างของโครงข่ายใยประสาทเทียมดังรูปที่ 3.2

3.1.1 ขั้นตอนการเรียนรู้ภายในแบบจำลอง BPNN

สำหรับขั้นตอนการเรียนรู้ภายในแบบจำลอง BPNN นั้นประกอบด้วยสองขั้นตอนหลักคือการคำนวณไปข้างหน้า (forward pass) ซึ่งจะเป็นการประมวลผลข้อมูลจากชั้นตัวแปรเข้าสู่ชั้นตัวแปรออก และขั้นตอนที่สองคือ การคำนวณย้อนกลับ (backward pass) ซึ่งจะใช้ค่าผลต่างระหว่างค่าสังเกต (ข้อมูลจริง) และค่าคำนวณจากแบบจำลองในชั้นตัวแปรออกเป็นการปรับแก้ค่าความสัมพันธ์ระหว่างชั้น ดังแสดงในรูปที่ 3.2 โดยมีขั้นตอนดังต่อไปนี้

1. แปลงข้อมูลนำเข้า และข้อมูลออก ให้อยู่ในรูปแบบมาตรฐาน (normalization) โดยในที่นี้จะมีค่าอยู่ระหว่าง 0.15-0.85
2. สมมุติค่าเริ่มต้นของค่าถ่วงน้ำหนักและค่าพารามิเตอร์ต่างๆ ในโครงสร้างแบบจำลอง
3. กำหนดเงื่อนไขการหยุดการทำงานของกระบวนการเรียนรู้ เช่น ระดับความผิดพลาดที่ยอมรับได้ หรือจำนวนรอบการคำนวณ



ที่มา: เสรี, 2544

รูปที่ 3.2 ลักษณะโครงสร้างของแบบจำลอง BPNN และหลักการทำงานในหน่วยย่อย

การคำนวณไปข้างหน้า (forward pass)

4. ในชั้นตัวแปรเข้า (input layer) แต่ละหน่วย ($X_i, i = 1, 2, \dots, n$) รับค่าตัวแปรเข้า x_i และจะส่งต่อไปยังชั้นตัวแปรซ่อนซึ่งอยู่ถัดไป
5. ในชั้นตัวแปรซ่อน (hidden layer) แต่ละหน่วย ($Z_j, j = 1, 2, \dots, n$) รับค่าผลรวมของผลคูณระหว่างตัวแปรเข้า x_i และค่าน้ำหนักระหว่างชั้นตัวแปรเข้าและชั้นตัวแปรซ่อน w_{ij}

$$Z_{in_j} = w_{oj} + \sum x_i w_{ij} \tag{3.1}$$

เมื่อ Z_{in_j} คือ ผลรวมของผลคูณระหว่างตัวแปรเข้า x_i และค่าน้ำหนักระหว่างชั้นตัวแปรเข้าและชั้นตัวแปรซ่อน w_{ij}

w_{oj} คือ ค่าเอียงในชั้นตัวแปรซ่อน j

x_i คือ ค่าตัวแปรเข้าในชั้นตัวแปรเข้า

w_{ij} คือ ค่าน้ำหนักระหว่างหน่วย i ในชั้นตัวแปรเข้าและหน่วย j ในชั้นตัวแปรซ่อน

หลังจากนั้นค่า Zin_j จะถูกแปลงผ่านฟังก์ชันกระตุ้นแบบซิกมอยด์ ค่าที่ได้จะเป็นค่าตัวแปรออกจากชั้นตัวแปรซ่อนเพื่อส่งผ่านไปให้ชั้นต่อไป

$$Z_j = f(Zin_j) \quad (3.2)$$

เมื่อ f คือ ฟังก์ชันกระตุ้นแบบซิกมอยด์ ดังแสดง

$$f(Zin_j) = \frac{1}{1 + \exp(-x)} \quad (3.3)$$

6. ในชั้นตัวแปรออก (output layer) ก็คำนวณเช่นเดียวกับชั้นตัวแปรซ่อนในข้อ 5. โดยแต่ละหน่วย (Y_k , $k = 1, 2, \dots, n$) รับค่าผลรวมของผลคูณระหว่างตัวแปรเข้า Z_j และค่าน้ำหนักระหว่างชั้นตัวแปรซ่อนและชั้นตัวแปรออก w_{jk}

$$Yin_k = w_{ok} + \sum Z_j w_{jk} \quad (3.4)$$

เมื่อ Y_{mk} คือ ผลรวมของผลคูณระหว่างตัวแปรเข้า Z_j และค่าน้ำหนักระหว่างชั้นตัวแปรซ่อนและชั้นตัวแปรออก w_{jk}

w_{ok} คือ ค่าเอียงเอียงในชั้นตัวแปรซ่อน k

w_{jk} คือ ค่าน้ำหนักระหว่างหน่วย j ในชั้นตัวแปรซ่อนและหน่วย k ในชั้นตัวแปรออก

หลังจากนั้นค่า Yin_k จะถูกแปลงผ่านฟังก์ชันกระตุ้นแบบซิกมอยด์ ค่าที่ได้จะเป็นผลลัพธ์จากแบบจำลอง

$$Y_k = f(Yin_k) \quad (3.5)$$

การคำนวณย้อนกลับ (backward pass)

7. เมื่อคำนวณถึงชั้นตัวแปรออก ค่าผลลัพธ์จากแบบจำลองแต่ละหน่วยในชั้นตัวแปรออกจะถูกนำไปคำนวณพจน์ความผิดพลาด (δ_k) โดยเทียบกับข้อมูลจริง

$$\delta_k = (T_k - y_k) f'(Yin_k) \quad (3.6)$$

เมื่อ T_k คือ ค่าที่ได้จากข้อมูลจริง

$f'(\)$ คือ อนุพันธ์ของฟังก์ชันกระตุ้นแบบซิกมอยด์

$$f'(Yin_k) = f(Yin_k) \times (1 - f(Yin_k)) \quad (3.7)$$

เมื่อได้พจน์ความผิดพลาด (δ_k) แล้วจึงทำการคำนวณค่าปรับแก้ค่าน้ำหนักระหว่างชั้นตัวแปรออกและชั้นตัวแปรซ่อน $\Delta w_{jk}(n+1)$ เพื่อใช้ปรับแก้ค่า $w_{jk}(\)$

$$\Delta w_{jk}(n+1) = \eta \delta_k Y_k + \alpha \Delta w_{jk}(n) \quad (3.8)$$

เมื่อ η คือ ค่าอัตราการเรียนรู้ มีค่าระหว่าง 0 ถึง 1
 α คือ ค่าโมเมนตัม มีค่าระหว่าง 0 ถึง 1

$\Delta w_{jk}(n)$ คือ ค่าปรับแก้น้ำหนักระหว่างหน่วย j ในชั้นตัวแปรซ่อนและหน่วย k ในชั้นตัวแปรออก ในรอบที่ n (ในรอบที่ $n = 1$ มีค่าเท่ากับ 0)

8. สำหรับชั้นตัวแปรซ่อนก็คำนวณพจน์ความผิดพลาด (δ_j) เช่นเดียวกับชั้นตัวแปรออก

$$\delta_j = \sum \delta_k w_{jk} f'(Z_{in_j}) \quad (3.9)$$

นำค่าพจน์ความผิดพลาด (δ_j) ที่ได้คำนวณค่าปรับแก้น้ำหนักระหว่างชั้นตัวแปรเข้าและชั้นตัวแปรซ่อน $\Delta w_{ij}(n+1)$ เพื่อใช้ปรับแก้ค่า $w_{ij}(n)$

$$\Delta w_{ij}(n+1) = \eta \delta_j Z_j + \alpha \Delta w_{ij}(n) \quad (3.10)$$

$\Delta w_{ij}(n)$ คือ ค่าปรับแก้น้ำหนักระหว่างหน่วย i ในชั้นตัวแปรเข้าและหน่วย i ในชั้นตัวแปรซ่อน ในรอบที่ n (ในรอบที่ $n = 1$ มีค่าเท่ากับ 0)

9. สำหรับแต่ละหน่วย ($Y_k, k = 1, 2, \dots, n$) ในชั้นตัวแปรออกทำการปรับแก้ค่าน้ำหนักเพื่อใช้ในการคำนวณใหม่ในรอบที่ $n+1$

$$w_{jk}(n+1) = w_{jk}(n) + \Delta w_{jk}(n+1) \quad (3.11)$$

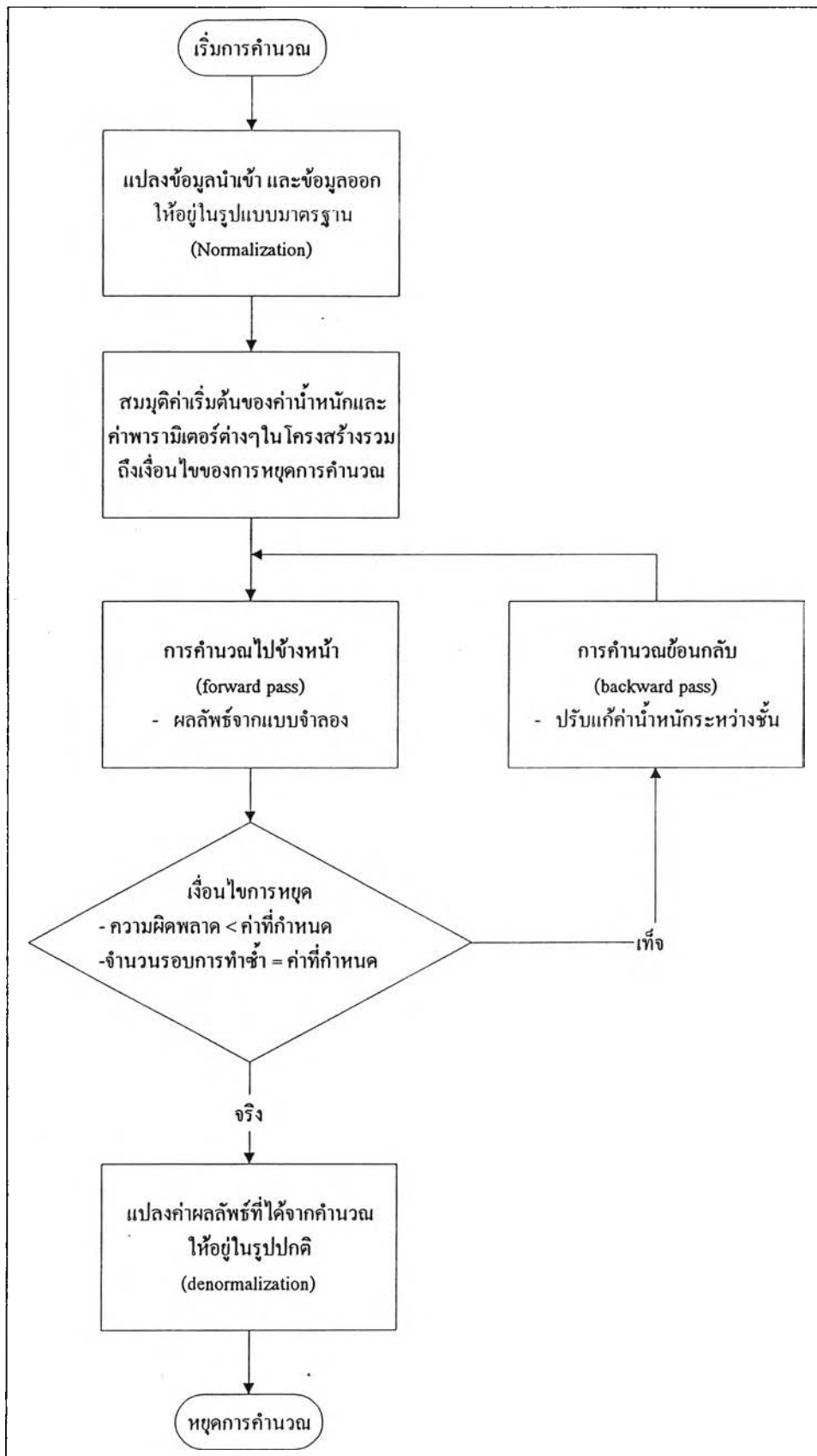
สำหรับแต่ละหน่วย ($Z_j, j = 1, 2, \dots, n$) ในชั้นตัวแปรซ่อนก็ทำการปรับแก้ค่าน้ำหนักเพื่อใช้ในการคำนวณใหม่ในรอบที่ $n+1$ เช่นเดียวกัน

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n+1) \quad (3.12)$$

10. ทำการคำนวณใหม่ตั้งแต่ข้อที่ 5-9 และจะหยุดการทำซ้ำตามเงื่อนไขการหยุดการทำงานของกระบวนการเรียนรู้ในข้อที่ 3

11. แปลงค่าผลลัพธ์ที่ได้จากคำนวณ Y_k กลับให้อยู่ในรูปปกติ (denormalization)

จากขั้นตอนการคำนวณทั้งหมดสามารถสรุปผังรูปที่ 3.3



รูปที่ 3.3 ฟังการไหลแสดงการทำงานของกระบวนการเรียนรู้ของแบบจำลอง BPNN

3.1.2 ตัวแปรและองค์ประกอบภายในแบบจำลอง BPNN

การพิจารณาตัวแปรและองค์ประกอบภายในแบบจำลอง BPNN ซึ่งอาจส่งผลต่ออัตราการลู่เข้าหรือความถูกต้อง ประกอบด้วย

1. จำนวนหน่วยในชั้นตัวแปรซ่อนและจำนวนชั้นตัวแปรซ่อน

ในช่วงแรกยังไม่มีความชัดเจนในการกำหนดโครงสร้างในชั้นตัวแปรซ่อนที่ชัดเจน ส่วนใหญ่ใช้การลองผิดลองถูกซึ่งกินเวลานาน เพื่อหาโครงสร้างที่เหมาะสม ต่อมาจึงมีการกำหนดโครงสร้างให้สัมพันธ์กับจำนวนหน่วยในชั้นตัวแปรเข้า เช่น จำนวนหน่วยในชั้นตัวแปรซ่อนเป็นจำนวนเท่าของจำนวนหน่วยในชั้นตัวแปรเข้า ต่อมามีการใช้ทฤษฎีการหาค่าที่เหมาะสมมาประยุกต์ใช้ เช่น Genetic Algorithm (เสรี, 2002)

ในการวิจัยครั้งนี้จะใช้วิธีการลองผิดลองถูกควบคู่กับการกำหนดโครงสร้างให้สัมพันธ์กับจำนวนหน่วยในชั้นตัวแปรเข้าในการกำหนดจำนวนหน่วยในชั้นตัวแปรซ่อนและจำนวนชั้นตัวแปรซ่อน

2. ฟังก์ชันกระตุ้น (activation function)

ฟังก์ชันกระตุ้นที่นิยมใช้โดยทั่วไปมีอยู่หลายแบบทั้งแบบเชิงเส้นและไม่เชิงเส้น โดยคุณสมบัติของฟังก์ชันกระตุ้นจะต้องอยู่ในขอบเขตและสามารถหาอนุพันธ์ได้ โดยฟังก์ชันซิกมอยด์เป็นที่นิยมมากที่สุดเนื่องจากมีความสะดวกในการหาอนุพันธ์ ดังแสดงในสมการที่ 3.7

ในการวิจัยครั้งนี้จะใช้ฟังก์ชันกระตุ้นซิกมอยด์ (Sigmoid function) ดังแสดงในตารางที่ 3.2

ตารางที่ 3.2 ฟังก์ชันกระตุ้นที่ใช้ในการวิจัย

ชนิดของฟังก์ชันกระตุ้น	ฟังก์ชันกระตุ้น
Sigmoid	$f(x) = 1/(1+\exp(-x))$

3. อัตราการเรียนรู้ (Learning rate , ETA)

อัตราการเรียนรู้จะเป็นตัวแปรหนึ่งในการกำหนดขนาดการปรับค่าถ่วงน้ำหนักดังสมการที่ 3.8 และ 3.15 โดยทั่วไปกำหนดให้เป็นค่าคงที่ระหว่าง 0.05 - 0.90 หากเลือกค่าที่สูงเกินไปจะทำให้การเดินทางไปหาความคลาดเคลื่อนที่ต่ำที่สุดมีการแกว่งและอาจไม่มีการลู่เข้าหาค่าตอบที่ต้องการ ในทางตรงข้ามการเลือกค่าอัตราการเรียนรู้ที่น้อยเกินไปก็เป็นการสิ้นเปลืองเวลาในการคำนวณ

4. ค่าโมเมนตัม (Momentum , ALPHA)

ในกระบวนการเรียนรู้ ค่าโมเมนตัมอาจจะใช้หรือไม่ใช้ก็ได้ อย่างไรก็ตามการเพิ่มค่าโมเมนตัมจะช่วยป้องกันการแกว่งของระบบ การเพิ่มค่าโมเมนตัมเป็นการเพิ่มสัดส่วนน้ำหนักของค่าถ่วงน้ำหนัก ดังสมการที่ 3.8 และ 3.10 การเพิ่มสัดส่วนดังกล่าวจะช่วยป้องกันการเปลี่ยนแปลง

ค่าถ่วงน้ำหนักที่รุนแรงหรือผิดปกติ ค่าโมเมนต์จะเป็นค่าบวกที่น้อยกว่า 1 โดยทั่วไปมีค่าระหว่าง 0.5 ถึง 0.9

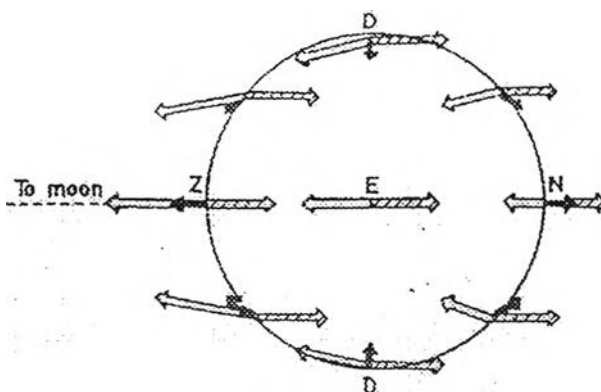
3.2 ทฤษฎีน้ำขึ้นน้ำลง

น้ำขึ้นน้ำลงเป็นคลื่นยาว (long wave) ที่ก่อให้เกิดการเปลี่ยนแปลงระดับของน้ำทะเล ในขณะที่ระดับน้ำทะเลกำลังสูงขึ้น เรียกว่า น้ำขึ้น (flood tide) และในขณะที่ระดับน้ำทะเลกำลังลดต่ำลง เรียกว่า น้ำลง (ebb tide) ระยะในแนวตั้งระหว่างตำแหน่งที่น้ำขึ้นสูงสุดและน้ำลงต่ำสุด เรียกว่า พิสัยน้ำขึ้นน้ำลง (tide range) ในแต่ละวัน ถ้ามีน้ำขึ้นสูงสุด หรือน้ำลงต่ำสุดสองครั้ง เรียกว่า น้ำคู่ (semi-diurnal) แต่ถ้ามีเพียงครั้งเดียว จะเรียกว่า น้ำเดี่ยว (diurnal)

3.2.1 แรงที่ทำให้เกิดน้ำขึ้นน้ำลง

แรงที่ทำให้เกิดน้ำขึ้นน้ำลง คือ แรงดึงดูดระหว่างกัน (attractive force) และแรงหนีศูนย์กลาง (centrifugal force) ของโลกกับดวงจันทร์ และดวงอาทิตย์ ซึ่งการเคลื่อนที่ที่สัมพันธ์กันระหว่างโลก ดวงจันทร์ และดวงอาทิตย์ ทำให้เกิดการเปลี่ยนแปลงระดับของน้ำขึ้นน้ำลง เมื่อพิจารณาแรงระหว่างโลกและดวงจันทร์ จะได้สมมูลย์ของแรงดังรูปที่ 3.4 และเมื่อพิจารณาว่าสถานะของโลกเป็นทรงกลมและมีน้ำทะเลครอบคลุมผิวโลกทั้งหมด ก็จะได้ผิวน้ำสมมูลย์ (equilibrium surface) ดังรูปที่ 3.5 จะเห็นว่าในกรณีที่มุมเท่ากับ 0 องศา จะเกิดน้ำขึ้นมากที่สุด และเมื่อมุมเท่ากับ 90 องศา จะเกิดน้ำลงต่ำสุด เมื่อพิจารณาการเคลื่อนที่ของโลกและดวงจันทร์ พบว่า ดวงจันทร์โคจรรอบโลก 1 รอบ ใช้เวลา 29.53 วัน ถ้าเทียบกับการหมุนรอบตัวเองของโลก ที่ตำแหน่งคนยืน ดังรูปที่ 3.6 จะพบว่าตำแหน่งของดวงจันทร์จะตรงศีรษะคนยืนอีกครั้ง ดวงจันทร์จะเคลื่อนที่ไป $1/29.53$ ของวงโคจรของดวงจันทร์รอบโลก ลักษณะเช่นนี้ จะทำให้น้ำขึ้นสูงสุด หรือน้ำลงต่ำสุดที่ตำแหน่งใดๆ บนผิวโลก จะเกิดช้ากว่าวันก่อน 50 นาที (เมื่อคำนึงเฉพาะระบบโลกและดวงจันทร์) ในทำนองเดียวกัน เมื่อพิจารณาระบบโลกและดวงอาทิตย์ ลักษณะสมมูลย์ของแรงและผิวน้ำสมมูลย์ จะเหมือนกับกรณีระหว่างโลกและดวงจันทร์ แต่การเคลื่อนที่สัมพันธ์ระหว่างโลกและดวงอาทิตย์มีคาบเท่ากับ 24 ชั่วโมง ดังนั้นน้ำขึ้นสูงสุดหรือน้ำลงต่ำสุดที่ตำแหน่งใดๆ บนผิวโลก จะเกิดที่เวลาเดียวกันทุกวัน (เมื่อคำนึงเฉพาะโลกและดวงอาทิตย์) อย่างไรก็ตาม จากการศึกษาพบว่าอัตราส่วนของระดับน้ำ เนื่องจากอิทธิพลของดวงอาทิตย์ต่อระดับน้ำเนื่องจากอิทธิพลของดวงจันทร์เท่ากับ 0.457 (Ippen, 1966) แสดงว่าดวงจันทร์มีอิทธิพลต่อระดับน้ำบนโลก 54.3 % เมื่อพิจารณาการเคลื่อนที่สัมพันธ์กันในระหว่างโลก ดวงจันทร์และดวงอาทิตย์ ดังรูปที่ 3.7 ถ้าตำแหน่งของดวงจันทร์ โลก และดวงอาทิตย์ อยู่ในแนวตั้งฉากกัน คือดวงจันทร์อยู่ที่ตำแหน่ง

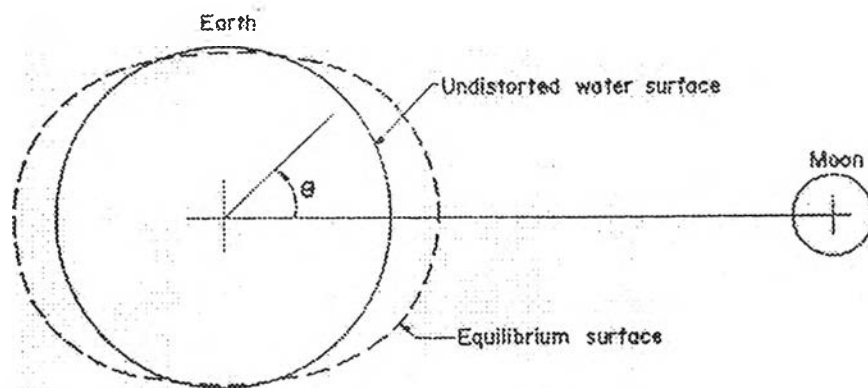
เสี้ยวที่ 1 (first quarter) และเสี้ยวสุดท้าย (last quarter) ทำให้เกิดพิสัยน้ำขึ้นน้ำลง (tidal range) น้อยที่สุด เรียกว่าน้ำตาย (neap tide) แต่ถ้าดวงจันทร์ โลกและดวงอาทิตย์ อยู่ในแนวเดียวกัน คือ ดวงจันทร์อยู่ที่ตำแหน่งเริ่มต้นดวงใหญ่อีกครั้ง (New moon) และตำแหน่งดวงจันทร์เต็มดวง (Full moon) จะทำให้เกิดพิสัยน้ำขึ้นน้ำลงมากที่สุด เรียกว่าน้ำเกิด (spring tide) การเกิดน้ำเกิดและน้ำตาย นี้ ก็เป็นองค์ประกอบหนึ่งที่ต้องพิจารณา เรียกว่า องค์ประกอบรายปักษ์ทางจันทร์คติ (lunar fortnightly constituents) ซึ่งมีคาบ (period) 13 วัน 15 ชั่วโมง 52 นาที



Determination of intensity and direction of the tide generating forces resulting from the difference between attractive and centrifugal forces.
 ⇨, open arrows : attractive forces : ⇨, hatched arrows : centrifugal forces : ⇨, black arrows : tide generating forces (after Defant, 1951).

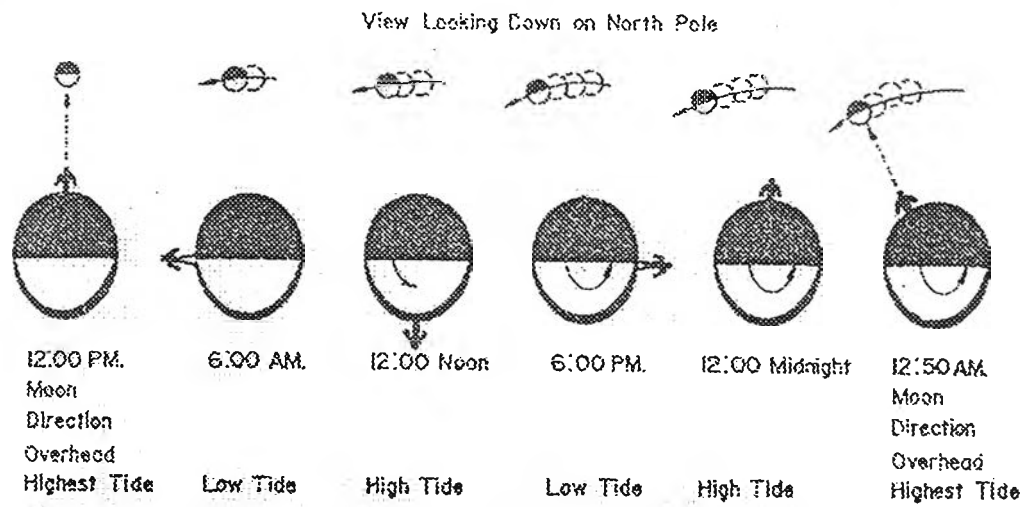
จาก : Kiyoshi (1978)

รูปที่ 3.4 แรงที่ทำให้เกิดน้ำขึ้นน้ำลง



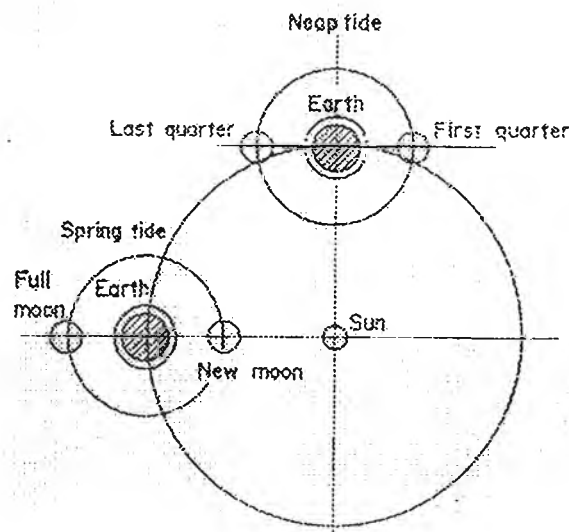
จาก : Tppen (1966)

รูปที่ 3.5 ผิวน้ำสมดุลย์ในระบบโลกและดวงจันทร์



จาก : ฮัลลรดคา (2524)

รูปที่ 3.6 การหมุนของโลกรอบตัวเองและการหมุนของดวงจันทร์รอบโลก



จาก : Kiyoshi (1978)

รูปที่ 3.7 การเกิดน้ำตาย (neap tide) และน้ำเกิด (spring tide)