

ข้อเสนอโครงการวิศวกรรมไฟฟ้า วิชา 2102490

การปรับปรุงแบบจำลองลดทอนจากเมฆเพื่อประมาณความเข้มแสง
อาทิตย์ทั่วประเทศไทยด้วยภาพถ่ายเมฆจากดาวเทียมฮิมาวาริ

Improving cloud attenuation model for ground
irradiance estimation across Thailand using cloud
images from Himawari satellite

นางสาวณัฐมน ตั้งแก เลขประจำตัวนิสิต 6230184121

นายบุญฤทธิ์ สร้างสุขสันติ เลขประจำตัวนิสิต 6230297421

อาจารย์ที่ปรึกษา รศ.ดร. จิตโกมุท ส่งศิริ

ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2565

ลงชื่ออาจารย์ที่ปรึกษาหลัก



(รศ.ดร. จิตโกมุท ส่งศิริ)

วันที่ 25 พ.ย. 2565

สารบัญ

1	บทนำ	4
1.1	ที่มาและความสำคัญของโครงการงาน	4
1.2	วัตถุประสงค์ของโครงการงาน	6
1.3	ขอบเขตของโครงการงาน	7
1.4	ผลลัพธ์ที่คาดหวังจากโครงการงาน	7
2	หลักการและทฤษฎีที่เกี่ยวข้อง	7
2.1	ความรู้พื้นฐานเกี่ยวกับความเข้มแสงอาทิตย์	7
2.2	แบบจำลองระบบ	10
2.3	ระเบียบวิธี	12
2.4	รายละเอียดชุดข้อมูล	14
3	ผลลัพธ์จากการดำเนินงานเบื้องต้น	16
3.1	ผลการตรวจสอบแบบจำลองรุ่นที่ 1 และรุ่นที่ 2	17
3.2	ผลการทดลองแบบจำลองที่นำเสนอ	19
4	บทสรุป	23
4.1	สรุปผลการดำเนินงาน	23
4.2	แผนการดำเนินงาน	24
5	เอกสารอ้างอิง	25

สารบัญรูป

1	แผนภาพขั้นตอนการประมาณความเข้มแสงอาทิตย์จากแบบจำลองลดทอนแสงด้วยเมฆ	4
2	กราฟความสัมพันธ์ระหว่าง CI และ k ของสมการ (1) และสมการ (2)	5
3	แผนภาพมุม zenith ความเข้มแสงอาทิตย์ตั้งฉากโดยตรง และความเข้มแสงกระจายแนวราบ	8
4	กราฟความสัมพันธ์ระหว่าง $AM(t)$ กับ $\cos(\theta(t))$ และ $CI(t)$ กับ $k(t)$	9
5	ภาพถ่ายเมฆแต่ละประเภทจากดาวเทียมฮิมาวาริ 8 จากสถานีรับ CUEE ณ วันที่เกิดเมฆมหึมา (Super cloud) และ ตัวอย่างดัชนีเมฆที่สถานีรับ 48 บริเวณซอยจุกฟ้าฯ 36 แขวงวังใหม่ เขตปทุมวัน กรุงเทพมหานคร	9
6	กราฟเปรียบเทียบ I และ I_{cr} ในวันที่ท้องฟ้าใสและท้องฟ้ามีเมฆ	10
7	แผนภาพการกระจายระหว่างดัชนีเมฆและดัชนีฟ้าใสของสถานีตัวอย่าง	11
8	ภาพพิกัดตำแหน่ง $CI(t-1)$ บนกริด 5×5 พิกเซล และมี $CI_{33}(t-1)$ เป็นพิกัดที่สนใจ	12
9	แผนภาพขั้นตอนฝึก ตรวจสอบ ทดสอบ และเลือกแบบจำลองการลดทอนความเข้มแสงด้วยเมฆ	13
10	แผนภาพขั้นตอนการนำแบบจำลองไปใช้งาน	14
11	กราฟความเข้มแสงที่วัดได้จากสถานีที่ 32 และ 54	15
12	กราฟ MAE และ NMAE รายชั่วโมงของแบบจำลองรุ่นที่ 1 และรุ่นที่ 2	17
13	ผลการประมาณ I จากแบบจำลองรุ่นที่ 1	18
14	ผลการประมาณ I จากแบบจำลองรุ่นที่ 2	18
15	กราฟค่า cross validation error ของแบบจำลองถดถอยเชิงพหุนามเมื่อปรับเปลี่ยนอันดับพหุนาม	20
16	กราฟค่า MAE รายชั่วโมงของ 56 สถานี ของแต่ละแบบจำลองที่นำเสนอ	21

17	กราฟค่า MAE รายชั่วโมงของ 56 สถานี ของแบบจำลองรุ่นที่ 2 และแบบจำลองที่นำเสนอ	22
18	ผลการประมาณ I จากแบบจำลองที่นำเสนอ	22

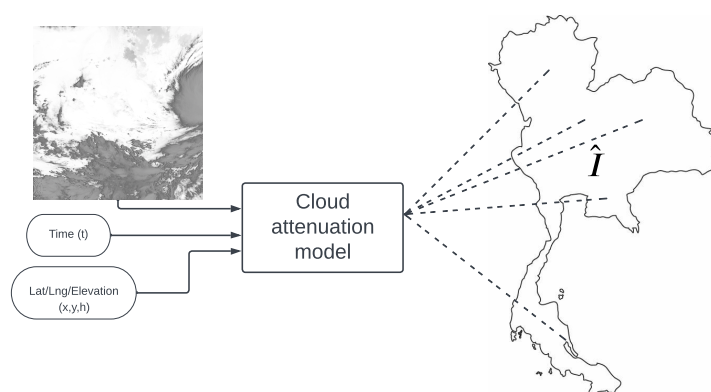
สารบัญตาราง

1	ตัวแปรที่เกี่ยวข้องกับการประมาณ	7
2	ตัวอย่างข้อมูลเชื่อมโยงตำแหน่งระหว่างประเทศไทยและภาพถ่ายเมฆ	12
3	รายละเอียดข้อมูลที่แบ่งตามรุ่นของแบบจำลอง	16
4	ความละเอียดเชิงเวลาและพื้นที่ของข้อมูล	16
5	เวลาของคู่ข้อมูลในการฝึกแบบจำลอง	16
6	ค่า cross validation error และ model selection scores ของแบบจำลองการถดถอยเชิงเส้น เมื่อมีการปรับเปลี่ยน ตัวแปรเข้า	19
7	ค่า training error และ cross validation error ของแบบจำลองป่าสุ่มเมื่อมีการปรับ q, r	20
8	ค่า training error และ cross validation error ของแบบจำลองป่าสุ่มเมื่อมีการปรับ d	21
9	ค่า MAE ของแต่ละแบบจำลองที่นำเสนอ	21
10	Gantt chart	24

1 บทนำ

1.1 ที่มาและความสำคัญของโครงการ

พลังงานทดแทนที่ใช้กันอย่างแพร่หลายในปัจจุบันคือพลังงานแสงอาทิตย์ และพื้นที่ส่วนใหญ่ของประเทศไทยเป็นพื้นที่ที่ได้รับแสงอาทิตย์ตลอดทั้งปี ประเทศไทยจึงเป็นตัวเลือกที่ดีต่อการติดตั้งโรงงานไฟฟ้าพลังงานแสงอาทิตย์ ความสามารถในการประมาณความเข้มแสงอาทิตย์สามารถนำข้อมูลไปวิเคราะห์ได้หลากหลาย เช่น เป็นส่วนช่วยสำคัญที่ใช้ในการประเมินความเสี่ยงในการจัดการระบบ ต้นทุนและสถานที่ที่จะติดตั้งระบบพลังงานแสงอาทิตย์เพื่อสามารถผลิตพลังงานได้อย่างมีประสิทธิภาพ หรือใช้ในประมาณกำลังผลิตไฟฟ้าของเซลล์แสงอาทิตย์แบบทันที เนื่องจากไม่สามารถติดตั้งตัวรับรู้เพื่อวัดค่าความเข้มแสงอาทิตย์ได้ทุกพื้นที่ ดังนั้นความสามารถในการประมาณค่าความเข้มแสงอาทิตย์จึงเป็นหัวข้อที่สนใจดังรูปที่ 1 และในปัจจุบันมีบริการทางการค้าเปิดให้บริการประมาณค่าความเข้มแสงอาทิตย์ เช่น Solar Anywhere, Solargis เป็นต้น



รูปที่ 1: แผนภาพขั้นตอนการประมาณค่าความเข้มแสงอาทิตย์จากแบบจำลองลดทอนแสงด้วยเมฆ

Solar Anywhere [13] มีบริการชื่อว่า Solar Anywhere Forecast Service ใช้ประโยชน์จากภาพถ่ายทางดาวเทียมช่วยในการประมาณและพยากรณ์ค่าความเข้มแสงอาทิตย์ มีให้บริการฟรีในทุกทวีปแต่จะมีเฉพาะพื้นที่ที่กำหนด กรุงเทพมหานครมีบริการฟรีทั้งหมด 9 จุด ละติจูดและลองจิจูดที่สามารถเลือกได้ เช่น (13.755 100.485) (13.755 100.495) (13.755 100.505) เป็นต้น ข้อมูลฟรีที่สามารถดาวน์โหลดได้เลือกค่าความละเอียดเชิงเวลาได้เฉพาะ 60 นาที มีผู้นำวิจัยคือ R. Perez [10] และทีมงาน ที่ the University at Albany (SUNY) นอกจากนี้ยังมีบริการที่ชื่อว่า Solar Anywhere Energy Modeling Services นำค่าความเข้มแสงอาทิตย์ที่ได้จาก Solar Anywhere Forecast ช่วยในการทำนายค่าต่าง ๆ ในเซลล์แสงอาทิตย์ เช่น กำลังไฟฟ้ากระแสตรงหรือกำลังไฟฟ้ากระแสสลับ เป็นต้น

Solargis [14] เปิดให้บริการประมาณและพยากรณ์ความเข้มแสงอาทิตย์ทั่วโลก รวมถึงประเทศไทย บริการทำความสะอาดและแสดงผลข้อมูลที่ได้จากการวัดความเข้มแสงอาทิตย์จากภาคพื้นดินและจากดาวเทียม บริการให้คำปรึกษาและช่วยวิเคราะห์ความไม่แน่นอนของแบบจำลองในแต่ละสถานที่โดยใช้ข้อมูลสภาพแวดล้อมของท้องถิ่นนั้น นอกจากนี้สามารถตรวจจับข้อมูลที่ผิดพลาดจากการวัด เพื่อเพิ่มประสิทธิภาพของการประมาณกำลังผลิตพลังงานไฟฟ้าจากระบบเซลล์แสงอาทิตย์ มีผู้นำวิจัยคือ M. Sári, T. Cebeauer, J.A. Ruiz-Arias [16] ที่ University of Málaga และทีมงาน

แบบจำลองที่ใช้ในการประมาณค่าความเข้มแสงอาทิตย์ที่ตกกระทบบนพื้นดินแบ่งออกเป็นสามประเภท ได้แก่ แบบจำลองเชิงกายภาพ (Physical model) แบบจำลองเชิงประจักษ์ (Empirical model) และแบบจำลองกึ่งเชิงประจักษ์ (Semi-empirical model) แบบจำลองเชิงกายภาพเป็นการประมาณความเข้มแสงอาทิตย์โดยตรงจากการสังเกตด้วยดาวเทียม และใช้ความรู้ทฤษฎีการส่งผ่านรังสี (Radiative transfer theory) ซึ่งคำนึงถึงลักษณะทางกายภาพของพื้นที่นั้น แบบจำลองเชิงประจักษ์เป็นการประมาณความเข้มแสงอาทิตย์จากการสังเกตความสัมพันธ์ของข้อมูลจากดาวเทียม และความเข้มแสงอาทิตย์ที่วัดได้จากภาคพื้นดิน [9] และ

แบบจำลองกึ่งเชิงประจักษ์เป็นการผสมผสานระหว่าง แบบจำลองเชิงกายภาพ และแบบจำลองเชิงประจักษ์ ซึ่งเป็นการประมาณความเข้มแสงอาทิตย์จากการพิจารณาลักษณะทางกายภาพของพื้นที่นั้น ร่วมกับการสังเกตความสัมพันธ์ระหว่างข้อมูลจากดาวเทียม และความเข้มแสงอาทิตย์ที่วัดได้จากภาคพื้นดิน (I)

พิจารณาตัวแปรสำคัญที่ใช้อธิบายในงานวิจัยกลุ่มนี้ ได้แก่ ความเข้มแสงในสภาวะฟ้าใส (I_{clr}) คือ ความเข้มแสงอาทิตย์ที่ตกกระทบบนพื้นดิน เป็นขอบบนของค่าความเข้มแสงอาทิตย์ที่พื้นดิน ดัชนีฟ้าใส (Clear-sky index, $k = I/I_{clr}$) คือค่าที่บ่งบอกความใสของท้องฟ้า และค่าดัชนีเมฆ (Cloud index, CI) บ่งบอกความหนาของเมฆ ดังนั้น ถ้ามีเมฆน้อย ค่าความเข้มแสงอาทิตย์ควรจะมาก ดัชนีฟ้าใสจึงควรจะสูง ตัวอย่างแบบจำลองการลดทอนความเข้มแสงอาทิตย์ด้วยเมฆที่มีการสร้างและพัฒนาขึ้น ดังนี้

K.F. Dagestad [2] พัฒนาแบบจำลองสำหรับโครงการ HELIOSAT ของฝรั่งเศสในปี 1982 กล่าวถึงการสกัดดัชนีเมฆ (CI) จากค่าความเข้มแสงที่ดาวเทียมได้รับ ในการหาดัชนีฟ้าใส (k) แบ่งตามช่วงของดัชนีเมฆดัง (1) จะเห็นว่า k มีความสัมพันธ์กับ CI ที่แตกต่างกันไปในแต่ละช่วงของ CI ที่จุดและเวลานั้น นำไปสู่การประมาณความเข้มแสงอาทิตย์ที่พื้นดินที่ได้รับ ($\hat{I} = \hat{k}I_{clr}$) และสังเกตว่าเมื่อ $-0.2 \leq CI \leq 1.1$ มีความสัมพันธ์ที่ผกผันกันกับ k เป็นไปตามฟังก์ชันของแต่ละช่วง CI และเมื่อ $CI < -0.2$, $CI > 1.1$ หมายความว่าสภาวะท้องฟ้าใส และท้องฟ้ามีเมฆปริมาณมากตามลำดับ ในสภาวะท้องฟ้าใส \hat{I} มีค่าเป็นไปตาม I_{clr} ที่มีขนาดใหญ่ขึ้น และในสภาวะท้องฟ้ามีเมฆปริมาณมาก \hat{I} ถูกลดทอนมากด้วยเมฆ มีค่าเป็นไป I_{clr} ที่มีขนาดเล็กลงดังรูปที่ 2(a)

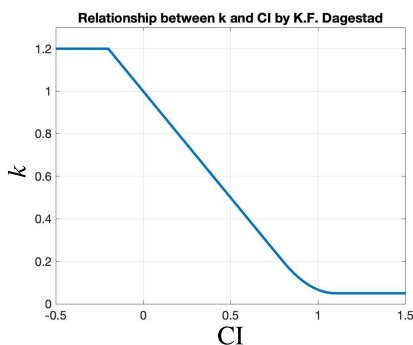
$$\hat{k} = \begin{cases} 1.2, & CI < -0.2 \\ 1 - CI, & CI \in [-0.2, 0.8] \\ 2.0667 - 3.6667CI + 1.6667CI^2, & CI \in [0.8, 1.1] \\ 0.05, & CI > 1.1 \end{cases} \quad (1)$$

R. Perez [9] กล่าวถึงการสกัดดัชนีเมฆ จากค่าความเข้มแสงอาทิตย์ที่ถูกสะท้อนโดยเมฆเข้าสู่ตัวรับรู้ และนำไปสู่การหาดัชนีฟ้าใสจากดัชนีเมฆที่จุดนั้น ๆ โดยแบบจำลองจะเป็นฟังก์ชันพหุนามอันดับห้า แบบจำลองนี้ใช้ใน Suny model มีสมการดังนี้

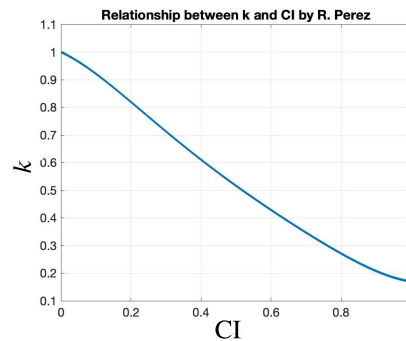
$$\hat{k} = a_5CI^5 + a_4CI^4 + a_3CI^3 + a_2CI^2 + a_1CI + a_0 \quad (2)$$

$$\hat{I} = \hat{k}I_{clr}(0.0001 \cdot \hat{k}I_{clr} + 0.9) \quad (3)$$

$(a_5, a_4, a_3, a_2, a_1, a_0) = (2.36, -6.3, 6.22, -2.63, -0.58, 1)$ ในการประมาณพารามิเตอร์ดังกล่าวใช้ข้อมูลของ 8 สถานีในสหรัฐอเมริกา และทวีปยุโรป สังเกตว่าเมื่อ $0 < CI < 1$ มีความสัมพันธ์ที่ผกผันกับ k เป็นไปตามฟังก์ชันพหุนามอันดับ 5 ดังรูปที่ 2(b)



(a) แบบจำลองของ K.F. Dagestad



(b) แบบจำลองของ R. Perez

รูปที่ 2: กราฟความสัมพันธ์ระหว่าง CI และ k ของสมการ (1) และสมการ (2)

M. Noia [12] และ C. Raphael [7] กล่าวถึงการสกัดดัชนีเมฆจากภาพถ่ายเมฆที่ถ่ายจากดาวเทียม ภาพถ่ายเมฆสามารถแบ่งภาพเป็นพิกเซลย่อย และสามารถคำนวณดัชนีเมฆ ณ พิกเซลนั้น จากอัตราส่วนระหว่างผลรวมของจำนวนพิกเซลที่มีเมฆ และจำนวนพิกเซลทั้งหมดในกรอบพื้นที่นั้น ซึ่งจำนวนพิกเซลที่มีเมฆจะเป็นผลรวมของจำนวนพิกเซลที่มีเมฆปานกลาง และจำนวนพิกเซลที่มีเมฆมากถูกถ่วงน้ำหนักโดยให้ความสำคัญกับจำนวนพิกเซลที่มีเมฆมากกว่า จำแนกแบบจำลองนี้เป็นสามประเภทตามดัชนีเมฆได้แก่ สภาวะฟ้าใส สภาวะเมฆบางส่วน และสภาวะเมฆมาก ได้แบบจำลองประมาณความเข้มแสงอาทิตย์ถดถอยแบบ Tarpley คือ

$$\hat{I} = \begin{cases} a_1 + b_1 \cos(\theta) + c_1\psi + d_1CI + e_1(B_{mo})^2, & CI < 0.4 \\ a_2 + b_2 \cos(\theta) + c_2CI(B_{cn})^2, & 0.4 \leq CI < 1.0 \\ a_3 + b_3 \cos(\theta) + c_3(B_{cn})^2, & CI = 1.0 \end{cases} \quad (4)$$

I คือค่าเฉลี่ยของความเข้มแสง ณ จุดที่สนใจ B_{mo} คือสัดส่วนระหว่างค่าเฉลี่ยความสว่างที่จุดสนใจกับค่าความสว่างสภาวะฟ้าใส B_{cn} คือสัดส่วนระหว่างค่าเฉลี่ยความสว่างเมฆกับค่าความสว่างสภาวะฟ้าใส ψ คือสัดส่วนปริมาณแสงที่ผ่านมาจากชั้นบรรยากาศ CI คือดัชนีเมฆ และ a_i, b_i, c_i, d_i และ e_i คือค่าสัมประสิทธิ์

แบบจำลองการประมาณความเข้มแสงอาทิตย์ที่พิจารณาในโครงการนี้เป็นประเภทแบบจำลองกึ่งเชิงประจักษ์ (Semi-empirical model) พิจารณาลักษณะทางกายภาพของพื้นที่นั้น เช่น ละติจูด ลองจิจูด ความสูงจากระดับน้ำทะเล สัมประสิทธิ์มวลอากาศ และสัมประสิทธิ์ความขุ่นมัวของชั้นบรรยากาศ ผ่านแบบจำลองฟ้าใส (Clear-sky model) เพื่อคำนวณค่าความเข้มแสงในสภาวะฟ้าใส (I_{clr}) ประกอบกับการสังเกตความสัมพันธ์ระหว่างดัชนีเมฆ (Cloud index, CI) ที่ถูกสกัดจากภาพถ่ายเมฆจากดาวเทียมฮิมาวาริ 8 และ ดัชนีฟ้าใสที่จุดและเวลาเดียวกัน โดยสังเกตว่าที่แต่ละค่าของ CI จะสามารถลดทอน I เท่าใด เป็นเพราะเหตุนี้แบบจำลองประมาณค่าความเข้มแสงอาทิตย์ที่พื้นดิน จึงมีชื่อว่า แบบจำลองการลดทอนความเข้มแสงด้วยเมฆ (Cloud attenuation model)

โครงการนี้ต่อยอดจากโครงการ Thailand Solar Map by CUEE [11] ซึ่งทดลองเกี่ยวกับการประมาณความเข้มแสงอาทิตย์ทั่วพื้นที่ประเทศไทย ได้ประเมินค่าความคลาดเคลื่อนของการประมาณความเข้มแสงอาทิตย์ของแบบจำลองแบ่งเป็น 2 รุ่นคือ

- แบบจำลองรุ่นที่ 1 แบบจำลองลอจิสติก พัฒนาในโครงการ Thailand Solar Map เฟส 1 ช่วงปี 2563 พัฒนาโดย นายสรารุต พรานนท์สถิตย์ ได้ความคลาดเคลื่อนการประมาณความเข้มแสงอาทิตย์โดยเฉลี่ยมีค่า MAE (Mean absolute error) เท่ากับ 159.37 W/m^2
- แบบจำลองรุ่นที่ 2 แบบจำลองป่าสุ่ม พัฒนาในโครงการ Thailand Solar Map เฟส 2 ช่วงปี 2564 - 2565 โดย นายปรินทร์ มโนมัยเสาวภาคย์ และเป็นแบบจำลองที่ใช้งานในปัจจุบัน ได้ความคลาดเคลื่อนการประมาณความเข้มแสงอาทิตย์โดยเฉลี่ยมีค่า MAE เท่ากับ 149.40 W/m^2 (อันเป็นการวัดสมรรถนะจากข้อมูลในช่วง 2022/09/07-2022/10/22)

การคำนวณแนวโน้มความเข้มแสงอาทิตย์ในสภาวะฟ้าใส จำเป็นต้องคำนึงถึงสองพารามิเตอร์หลัก คือพารามิเตอร์ที่บ่งบอกสภาพพื้นที่ และสภาพความขุ่นมัวของชั้นบรรยากาศ จากแบบจำลองรุ่นที่ 1 และรุ่นที่ 2 ได้พิจารณาความสูงของแต่ละพื้นที่เท่ากันทั้งหมด และใช้ค่าสภาพความขุ่นมัวของชั้นบรรยากาศในอดีต นอกจากนี้แบบจำลองทั้งสองยังไม่ได้คำนึงถึงเงื่อนไขการรับข้อมูลภาพถ่ายเมฆ มีเวลาประวิงเท่ากับ 30 นาที สามปัจจัยดังกล่าวควรนำไปใช้ในการพัฒนาและปรับปรุงแบบจำลองการลดทอนความเข้มแสงให้มีประสิทธิภาพเพิ่มขึ้น

1.2 วัตถุประสงค์ของโครงการ

1. วิเคราะห์ความคลาดเคลื่อนของการประมาณความเข้มแสงอาทิตย์ด้วยแบบจำลองในโครงการ Thailand Solar Map by CUEE รุ่นที่ 1 และรุ่นที่ 2
2. ปรับปรุงแบบจำลองการลดทอนความเข้มแสงอาทิตย์จากเมฆให้มีประสิทธิภาพมากกว่าแบบจำลองที่ใช้ในโครงการ Thailand Solar Map by CUEE

1.3 ขอบเขตของโครงการ

1. แบบจำลองการลดทอนความเข้มแสงอาทิตย์ด้วยเมฆสามารถประมาณความเข้มแสงอาทิตย์เฉพาะพื้นที่ประเทศไทย ประมวลผลตั้งแต่ 8:00 น. ถึง 17:00 น. และมีความละเอียดทุก 30 นาที
2. ข้อมูลที่ใช้ฝึกสอนแบบจำลอง ได้แก่ ข้อมูลค่าความเข้มแสงอาทิตย์ที่วัดได้จากภาคพื้นดินทั้งหมด 56 สถานี และ ภาพถ่ายเมฆจากดาวเทียมฮิมวาริ 8
3. เปรียบเทียบแบบจำลองการลดทอนความเข้มแสงอาทิตย์ด้วยเมฆในกลุ่มแบบจำลองทางสถิติ ได้แก่ แบบจำลองในกลุ่ม Regression และ Machine learning

1.4 ผลลัพธ์ที่คาดหวังจากโครงการ

1. แบบจำลองและชุดคำสั่งในรูปแบบ Python ของแบบจำลองการลดทอนความเข้มแสงอาทิตย์ด้วยเมฆ ซึ่งพร้อมใช้ในการแสดงผลค่าประมาณความเข้มแสงอาทิตย์ทั่วประเทศไทย
2. ผลการวิเคราะห์ประสิทธิภาพของแบบจำลองการลดทอนความเข้มแสงอาทิตย์จากเมฆที่น่าเสนอ

2 หลักการและทฤษฎีที่เกี่ยวข้อง

ในส่วนนี้อธิบายถึงความรู้พื้นฐานเกี่ยวกับความเข้มแสงอาทิตย์ แบบจำลองระบบ เปรียบวิธีการสร้างแบบจำลองการลดทอนความเข้มแสงด้วยเมฆ และรายละเอียดชุดข้อมูล ในการทดลองมีจำนวนตัวแปรที่เกี่ยวข้องกับการประมาณเป็นจำนวนมาก จึงกำหนดสัญลักษณ์และตัวแปรทางคณิตศาสตร์เพื่อความกะทัดรัดดังนี้

กำหนดให้ x คือตัวแปรหนึ่ง ตัวแปรที่ถูกเขียนอยู่ในรูป \hat{x} คือ ค่า x ที่ถูกประมาณโดยแบบจำลอง $x(t)$ คือ ค่า x ณ เวลา t ที่เป็นตัวชี้เวลาวิยุต มีเวลาในการชักตัวอย่าง 30 นาที $x(t - \tau)$ คือ ค่า x เวลาก่อนหน้า 30τ นาที และ x_i คือ ค่า x ลำดับที่ i และตัวแปรที่เกี่ยวข้องกับการประมาณดังตารางที่ 1

ตารางที่ 1: ตัวแปรที่เกี่ยวข้องกับการประมาณ

ตัวแปร	ความหมาย	หน่วย
I	ความเข้มแสงที่วัดได้จากภาคพื้นดิน (Global horizontal irradiance)	W/m ²
I_{clr}	ความเข้มแสงที่พื้นดินในสภาวะฟ้าใส (Clear sky irradiance)	W/m ²
k	ดัชนีฟ้าใส (Clear-sky index)	-
CI	ดัชนีเมฆ (Cloud index)	-
θ	มุมของดวงอาทิตย์เทียบกับแนวตั้งฉากของพื้นโลก (Zenith angle)	degree
AM	ค่าสัมประสิทธิ์มวลอากาศ (Airmass coefficient)	-
T_L	ค่าสัมประสิทธิ์ความขุ่นมัวของชั้นบรรยากาศ (Linke turbidity coefficient)	-
h	ความสูงจากระดับน้ำทะเล (Altitude)	m

2.1 ความรู้พื้นฐานเกี่ยวกับความเข้มแสงอาทิตย์

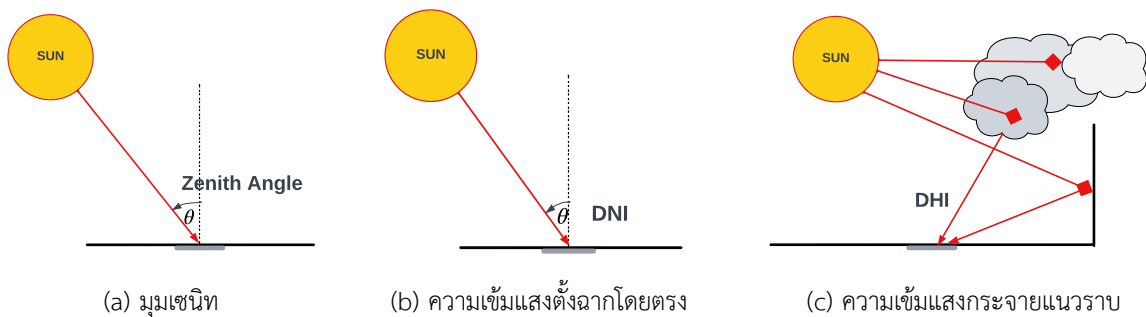
ในส่วนนี้อธิบายความรู้พื้นฐานเกี่ยวกับความเข้มแสงอาทิตย์ ตัวแปรที่ส่งผลหรือเกี่ยวข้องกับความเข้มแสงอาทิตย์ และแบบจำลองระบบที่จำเป็นในการสร้างแบบจำลองประมาณความเข้มแสงอาทิตย์ อ้างอิงจาก [1] และ [4] เป็นหลัก

มุมเซนิต (Zenith angle) คือ มุมของดวงอาทิตย์เทียบกับแนวตั้งฉากของพื้นโลกดังรูปที่ 3(a)

ส่วนประกอบของรังสีดวงอาทิตย์ ความเข้มแสงจากดวงอาทิตย์ทั้งหมดที่ตกกระทบชั้นบรรยากาศ (I_0) มีค่าประมาณ 1360 W/m^2 [1] จะถูกลดทอนเนื่องจากมีปฏิสัมพันธ์กับชั้นบรรยากาศเช่น การสะท้อน การดูดกลืน โดยชั้นบรรยากาศ เป็นต้น ความเข้มแสงอาทิตย์บางส่วนจะหายไป และความเข้มแสงอาทิตย์บางส่วนจะเข้ามาถึงพื้นโลก ความเข้มแสงอาทิตย์ที่ตกกระทบบนพื้นโลกนั้นจึงมีชื่อว่า ความเข้มแสงบนพื้นแนวนอนของโลก (Global Horizontal Irradiance, GHI) มีองค์ประกอบหลักสองส่วน ได้แก่ ความเข้มแสงตั้งฉากโดยตรง (Direct Normal Irradiance, DNI) และความเข้มแสงกระจายแนวราบ (Diffuse Horizontal Irradiance, DHI) ดัง (5)

$$\text{GHI} = \text{DNI} \cos(\theta) + \text{DHI} \quad (5)$$

เมื่อพิจารณาพื้นที่หนึ่ง ความเข้มแสงตั้งฉากโดยตรง คือ ความเข้มแสงจากดวงอาทิตย์มายังพื้นที่นั้น โดยตรงดังรูปที่ 3(b) และความเข้มแสงกระจายแนวราบ คือ ความเข้มแสงจากดวงอาทิตย์ที่ถูกสะท้อนด้วยวัตถุต่าง ๆ เช่น เมฆ และพื้นทีราบชื้น มายังพื้นที่นั้นดังรูปที่ 3(c)



รูปที่ 3: แผนภาพมุมเซนิต ความเข้มแสงอาทิตย์ตั้งฉากโดยตรง และความเข้มแสงกระจายแนวราบ

ค่าสัมประสิทธิ์มวลอากาศ (Airmass coefficient, AM) คืออัตราส่วนระหว่างความยาวของเส้นแสง (optical path length) จากการแผ่รังสีของดวงอาทิตย์ ณ ตำแหน่งมุมหนึ่ง ๆ กับความยาวของเส้นแสงที่แผ่มาตามแนวแกนเซนิต ค่าสัมประสิทธิ์มวลอากาศเป็นตัวแปรที่ขึ้นกับมุมเซนิตเป็นหลักดัง (6) ค่าสัมประสิทธิ์มวลอากาศจะเพิ่มขึ้นเมื่อมุมเซนิตเพิ่มขึ้นดังรูปที่ 4(a) มีสมการดังนี้ [8]

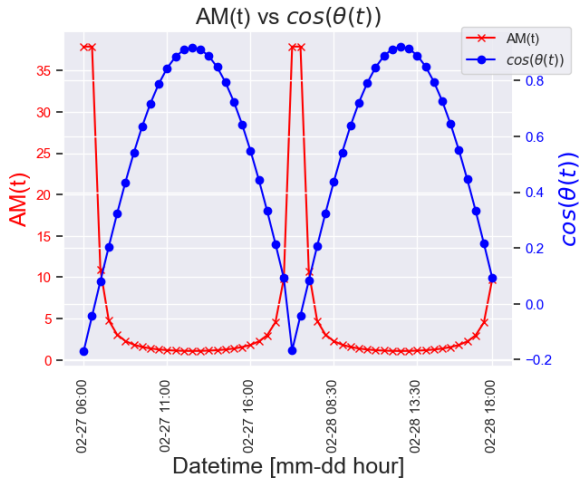
$$\text{AM}(t) = 1 / [\cos(\theta(t)) + 0.50572(96.07995 - \theta(t))^{-1.6364}] \quad (6)$$

สัมประสิทธิ์ความขุ่นมัวของชั้นบรรยากาศ (Linke turbidity coefficient, T_L) บ่งชี้ถึงค่าความสะอาดและความแห้งของชั้นบรรยากาศ กล่าวคือ ความขุ่นมัวของชั้นบรรยากาศจะลดค่าความเข้มแสงอาทิตย์ที่ผ่านชั้นบรรยากาศ จากการกระเจิงและดูดกลืนของอนุภาคที่อยู่ในอากาศ เช่น หยดน้ำ เกล็ดน้ำแข็ง ฝุ่น หรือหมอกควัน เมื่อมีค่าความขุ่นมัวเพิ่มขึ้นการกระเจิงและดูดกลืนแสงจากอนุภาคเหล่านั้นมากขึ้น I จึงมีค่าลดลง [17] ซึ่งเป็นค่าที่ต้องทราบเพื่อที่ใช้ในการสังเกตการณ์ I และ T_L เป็นพารามิเตอร์หลักที่ส่งผลกับความแม่นยำของแบบจำลองฟ้าใส ในปี 1992 Linke เสนอการหาค่าสัมประสิทธิ์ความขุ่นมัวของชั้นบรรยากาศ จากความหนาของชั้นบรรยากาศที่สะอาดและแห้ง (δ_{cda}) [4] ดัง (7)

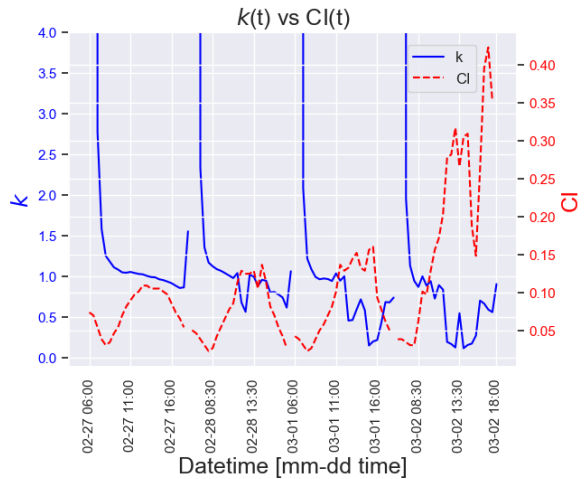
$$\text{DNI} = I_0 \cdot \exp(-\delta_{\text{cda}} \cdot T_L \cdot \text{AM}) \quad (7)$$

ดัชนีฟ้าใส (Clear-sky index, k) คือ อัตราส่วนระหว่าง I และ I_{clr} โดยที่ k ณ เวลาและตำแหน่งหนึ่ง แสดงถึงค่าความเข้มแสงที่พื้นดินได้รับขณะนั้นมีปริมาณมากน้อยอย่างไร เมื่อเทียบกับความเข้มแสงสภาวะท้องฟ้าใส แสดงตัวอย่างค่า k ของวันที่ 8 กุมภาพันธ์ 2565 ถึง 10 กุมภาพันธ์ 2565 ดังรูปที่ 4(b)

$$k = \frac{I}{I_{clr}} \quad (8)$$



(a) กราฟระหว่าง $AM(t)$ กับ $\cos(\theta(t))$



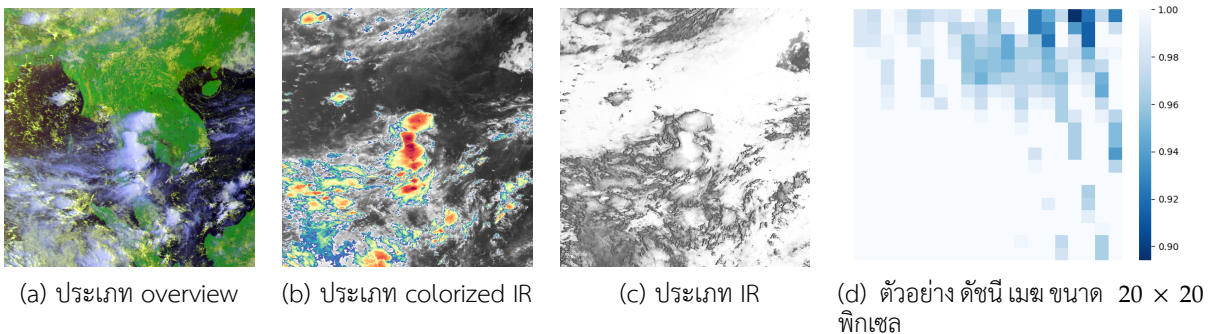
(b) กราฟความสัมพันธ์ระหว่าง $CI(t)$ กับ $k(t)$

รูปที่ 4: กราฟความสัมพันธ์ระหว่าง $AM(t)$ กับ $\cos(\theta(t))$ และ $CI(t)$ กับ $k(t)$

ดัชนีเมฆ (cloud index, CI) คือค่าที่บ่งบอกความหนาแน่นของเมฆ ดัชนีเมฆสามารถหาได้หลายวิธี จะกล่าวสองวิธีหลัก ได้แก่ หาดัชนีเมฆจากภาพถ่ายท้องฟ้า ใช้ภาพถ่ายท้องฟ้าจากพื้นดิน และหาดัชนีเมฆจากภาพถ่ายดาวเทียม การหาดัชนีเมฆจากภาพถ่ายท้องฟ้า วิธีนี้มีข้อจำกัดคือ สามารถบ่งบอกได้เพียงว่าภาพมีหรือไม่มีเมฆ นักวิจัยหลายท่านพยายามแก้ไขปัญหานี้ เช่น เปรียบเทียบสัดส่วนของพิกเซลเมฆจากสีน้ำเงินถึงสีแดง เพื่อแยกความแตกต่างของความหนาแน่นของเมฆ [4] และการหาดัชนีเมฆจากภาพถ่ายดาวเทียมเป็นวิธีที่ใช้ในการศึกษานี้ ใช้ภาพจากเลเยอร์ระดับสีเทา ดังรูปที่ 5(c) และทำให้เป็นมาตรฐานดัง (9)

$$CI = \frac{X - LB}{UB - LB} \quad (9)$$

X คือ ค่าของความหนาแน่นเมฆของพิกเซล LB และ UB คือขอบล่างและขอบบนของค่าความหนาแน่นของเมฆ เลือกใช้ $LB = 0$ และ $UB = 255$ หมายความว่าสภาวะท้องฟ้าใส และสภาวะที่ท้องฟ้ามีเมฆปริมาณมาก



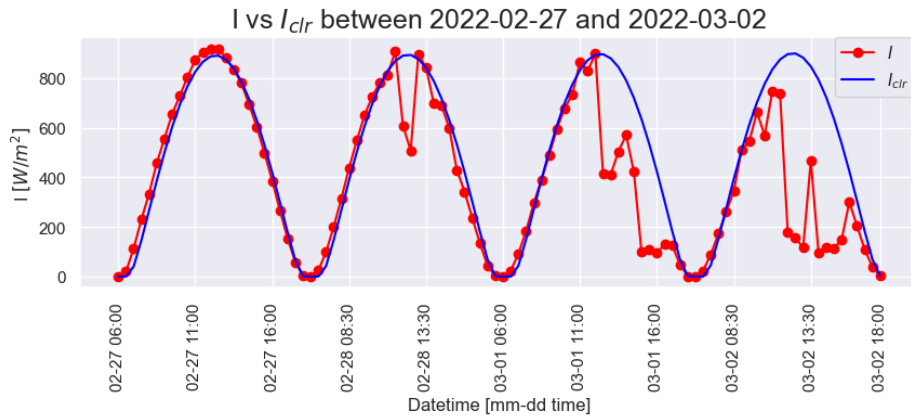
(a) ประเภท overview

(b) ประเภท colorized IR

(c) ประเภท IR

(d) ตัวอย่าง ดัชนี เมฆ ขนาด 20×20 พิกเซล

รูปที่ 5: ภาพถ่ายเมฆแต่ละประเภทจากดาวเทียมฮิมวาริ 8 จากสถานีรับ CUEE ณ วันที่เกิดเมฆมหึมา (Super cloud) และ ตัวอย่างดัชนีเมฆที่สถานีรับ 48 บริเวณซอยจุฬาฯ 36 แขวงวังใหม่ เขตปทุมวัน กรุงเทพมหานคร



รูปที่ 6: กราฟเปรียบเทียบ I และ I_{clr} ในวันที่ท้องฟ้าใสและท้องฟ้ามีเมฆ

2.2 แบบจำลองระบบ

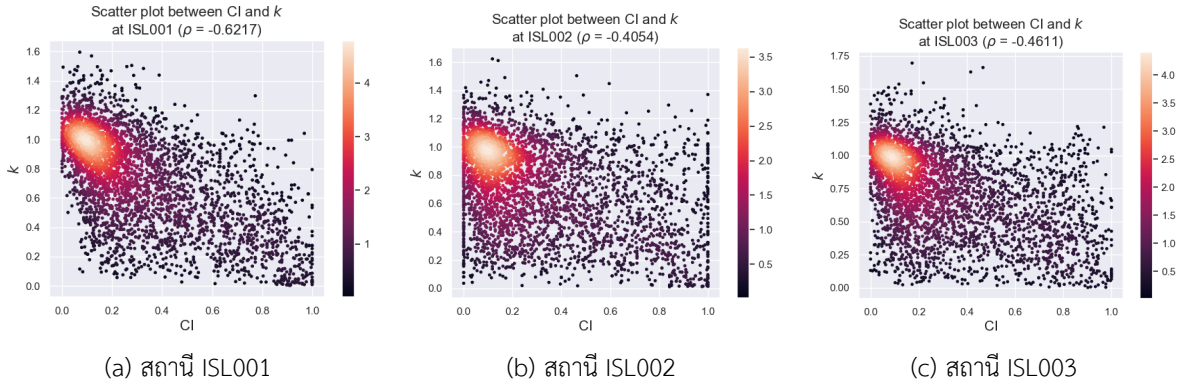
แบบจำลองฟ้าใส (Clear-sky model) เป็นแบบจำลองที่ให้ค่าความเข้มแสงอาทิตย์โดยมีเงื่อนไขท้องฟ้าอยู่ในสภาวะฟ้าใส หรือท้องฟ้าไม่มีเมฆ (I_{clr}) แบบจำลองนี้จะขึ้นกับตัวแปรหลายตัว เช่น ปริมาณโอโซน ปริมาณไอน้ำ และตำแหน่งเชิงเรขาคณิตของดวงอาทิตย์ เป็นต้น ในวันที่ท้องฟ้าใสจะเห็นว่าความเข้มแสงมีแนวโน้มเป็นไปตามแบบจำลองฟ้าใส และในวันที่มีเมฆ ค่าความเข้มแสงจะมีค่าที่ลดลง และผันผวนแสดงดังรูปที่ 6 I_{clr} เป็นตัวแปรที่สำคัญในการประมาณความเข้มแสงเพราะเป็นแนวโน้มเอียงของแต่ละเวลาและตำแหน่งนั้น จึงมีผู้พัฒนาแบบจำลองนี้ขึ้นมากมาย แต่ละแบบจำลองแตกต่างกันที่พารามิเตอร์ที่ใช้ และความแม่นยำของแบบจำลองฟ้าใสขึ้นกับค่า T_L เป็นหลัก เกณฑ์ในการเลือกใช้แบบจำลองจึงสามารถปรับเปลี่ยนตามการนำไปใช้งาน [4] ดังนั้นจึงเลือกใช้แบบจำลองฟ้าใสที่พัฒนาโดย P. Ineichen และ R. Perez มีตัวแปรเข้าที่สำคัญคือ ละติจูด ลองจิจูด ความสูงจากระดับน้ำทะเล (Altitude) และมุมเซนิท (θ) [3] ดังนี้

$$I_{clr}(t) = a_1 I_0 \cos(\theta(t)) e^{-a_2(f_{h_1} + f_{h_2}(T_L - 1))AM(t)} \quad (10)$$

- I_0 คือ ค่าความเข้มแสงก่อนตกระทบชั้นบรรยากาศ มีค่าเท่ากับ 1366.1 W/m^2
- h คือ ความสูงจากระดับน้ำทะเล นำข้อมูลแต่ละละติจูดและลองจิจูดจากเซตข้อมูล SRTM30m [6]
- $T_L = 4.7737$ (คำนวณโดยวิธีดรอยเชิงเส้นอย่างง่ายจากข้อมูลของแบบจำลองที่นำเสนอ)
- $f_{h_1} = e^{-h/8000}$
- $f_{h_2} = e^{-h/1250}$
- $a_1 = 5.09 \times 10^{-5}h + 0.868$
- $a_2 = 3.92 \times 10^{-5}h + 0.0387$

แบบจำลองการลดทอนความเข้มแสงด้วยเมฆ (Cloud attenuation model) แบบจำลองนี้เป็นแบบจำลองที่ประมาณความเข้มแสงอาทิตย์ที่พื้นดิน มีตัวแปรทำนายหลักคือ CI และ I_{clr} และตัวแปรออกคือ I หรือ k จากการสำรวจข้อมูลพบว่า CI และ k มีความสัมพันธ์ที่แปรผกผันกันดังรูปที่ 7 แสดงตัวอย่างของแผนภาพการกระจายระหว่าง CI กับ k ของ IMPACT SOLAR station (ISL) ที่ 1 2 และ 3 ตามลำดับ ซึ่งเป็นตัวอย่างของสถานีที่มีตัวรับรู้ที่วัดค่าความเข้มแสงอาทิตย์ กล่าวคือ ถ้าเมฆมีปริมาณมากส่งผลให้ความเข้มแสงที่พื้นดินมีค่าน้อย ดังนั้นจึงสามารถเชื่อมโยงความสัมพันธ์ระหว่างตัวแปรเข้าและตัวแปรออกผ่านแบบจำลองการลดทอนความเข้มแสงด้วยเมฆ และเนื่องจากภาพถ่ายเมฆจากดาวเทียมมีเวลาประวิง 30 นาที เช่น เวลาที่ปรากฏ

บนภาพถ่ายเมฆ ณ เวลา 8:00 น. คือภาพที่ถูกถ่ายจากดาวเทียม ณ เวลา 7:30 น. ดังนั้นแบบจำลองที่นำเสนอจะพิจารณาความสัมพันธ์ของ $CI(t-1)$ และ $I(t)$ แบบจำลองที่พิจารณามีดังนี้



รูปที่ 7: แผนภาพการกระจายระหว่างดัชนีเมฆและดัชนีฟ้าใสของสถานีตัวอย่าง

1. แบบจำลองการถดถอยเชิงเส้น (Linear regression) ใช้ฟังก์ชันเชิงเส้นเป็นเกณฑ์มาตรฐานของแบบจำลองมีโครงสร้างทั่วไปดัง (11)

$$\hat{I} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n \quad (11)$$

x_i คือ ตัวแปรเข้า สำหรับ $i = 1, \dots, n$ β_i คือ ค่าสัมประสิทธิ์การถดถอย สำหรับ $i = 0, \dots, n$ เลือกกลุ่ม x_i จากชุดย่อยของกลุ่มตัวแปร $I_{clr}(t)$, $CI(t-1)$, $I_{clr}(t)CI(t-1)$ ที่เป็นไปได้ดังนี้

$$\{I_{clr}(t)\}, \{CI(t-1)\}, \{I_{clr}(t)CI(t-1)\}, \{I_{clr}(t), CI(t-1)\}, \{I_{clr}(t), I_{clr}(t)CI(t-1)\}, \\ \{CI(t-1), I_{clr}(t)CI(t-1)\}, \{I_{clr}(t), CI(t-1), I_{clr}(t)CI(t-1)\} \quad (12)$$

2. แบบจำลองการถดถอยเชิงพหุนาม (Polynomial regression) ใช้ฟังก์ชันพหุนามอันดับ m เป็นค่าบ่งชี้ความซับซ้อนของแบบจำลอง เชื่อมโยงความสัมพันธ์ระหว่างตัวแปรเข้าคือ $CI(t-1)$ และตัวแปรออก $\hat{k}(t)$ ดัง (13)

$$\hat{k}(t) = \beta_0 + \beta_1 CI(t-1) + \beta_2 CI(t-1)^2 + \dots + \beta_m CI(t-1)^m \quad (13)$$

$$\hat{I}(t) = \hat{k}(t) \cdot I_{clr}(t) \quad (14)$$

β_i คือ ค่าสัมประสิทธิ์การถดถอย $i = 0, \dots, n$

3. แบบจำลองการสุ่มป่าไม้ (Random forest) ในการประมาณความเข้มแสงอาทิตย์ เลือกตัวแปรเข้าดังนี้

$$I_{clr}(t), I_{clr}(t-1), HR(t), CI_{33}(t-2), CI_{11}(t-1), CI_{12}(t-1), CI_{13}(t-1), \dots, CI_{55}(t-1) \quad (15)$$

ตัวแปรเข้าของแบบจำลองป่าสุ่มมีทั้งหมด 29 ตัวแปร แบ่งเป็น 3 กลุ่มหลัก คือ ความเข้มแสงในสภาวะฟ้าใส ลำดับชั่วโมง (HR) และค่า $CI(t-1)$ ที่แต่ละพิกัด แสดงพิกัดของ $CI(t-1)$ ดังรูปที่ 8

แบบจำลองการสุ่มป่าไม้ มีไฮเปอร์พารามิเตอร์คือ จำนวนต้นไม้ตัดสินใจ (Number of trees in the forest, n_{tree}) ระดับความลึกที่เป็นไปได้มากที่สุดของแต่ละต้นไม้ตัดสินใจ (Maximum number of levels in each decision tree, d) จำนวนตัวแปรเข้าที่มากที่สุดในแต่ละโนด (Maximum number of features considered for splitting node, p) จำนวนจุดข้อมูลที่น้อยที่สุดในแต่ละโนดก่อนที่จะมีการแตกแยก

$CI_{11}(t-1)$	$CI_{12}(t-1)$			$CI_{15}(t-1)$
$CI_{21}(t-1)$	$CI_{22}(t-1)$			
		$CI_{33}(t-1)$		
$CI_{51}(t-1)$				$CI_{55}(t-1)$

รูปที่ 8: ภาพพิกัดตำแหน่ง $CI(t-1)$ บนกริด 5×5 พิกเซล และมี $CI_{33}(t-1)$ เป็นพิกัดที่สนใจ

(Minimum number of data points placed in a node before the node is split, q) และ จำนวนจุดข้อมูลที่น้อยที่สุดในโนดใบไม้ (Minimum number of data points allowed in a leaf node, r)

ไฮเปอร์พารามิเตอร์หลักที่มีผลกับสมรรถนะของแบบจำลองได้แก่ จำนวนต้นไม้ตัดสินใจ จำนวนตัวแปรเข้าที่มากที่สุดในแต่ละโนด และจำนวนจุดข้อมูลที่น้อยที่สุดในโนดใบไม้ [15] เมื่อ n_{tree} มีค่ามากส่งผลให้แบบจำลองมีสมรรถนะที่ดี แต่ทำให้ใช้เวลาในการฝึกแบบจำลองนานขึ้น ไม่ควรเลือกค่าเกินไปจะส่งผลให้แบบจำลองมีความแปรปรวนมาก เมื่อ p มีค่ามากแต่ละโนดจะมีจำนวนตัวแปรเข้าที่พิจารณาเป็นจำนวนมาก ส่งผลให้แบบจำลองมีสมรรถนะที่ดี แต่จะทำให้ไม่มีความหลากหลายของแต่ละต้นไม้ และทำให้ใช้เวลาในการฝึกแบบจำลองนานขึ้นเช่นกัน r เมื่อจำนวนจุดข้อมูลในโนดใบไม้มีค่าน้อยจะทำให้แบบจำลองจับสัญญาณรบกวนในชุดข้อมูลได้ง่าย จำนวนจุดข้อมูลที่น้อยที่สุดในโนดใบไม้จึงไม่ควรมีค่าต่ำเกินไป

2.3 ระเบียบวิธี

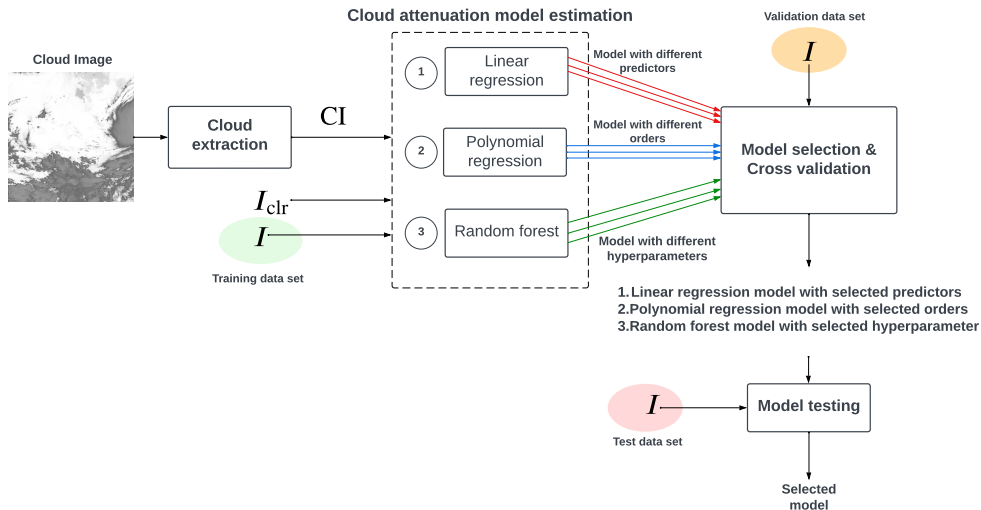
การประมาณความเข้มแสงอาทิตย์ที่ตกกระทบทั่วประเทศไทย สามารถประมาณได้จากแบบจำลองการลดทอนความเข้มแสงด้วยเมฆ มีตัวแปรเข้าได้แก่ ภาพถ่ายเมฆจากดาวเทียมฮิมาวาริ 8 ข้อมูลของเวลา (t) และข้อมูลของตำแหน่ง (x, y, h) ที่สนใจ แบบจำลองการลดทอนความเข้มแสงด้วยเมฆให้ I แต่ละตำแหน่งนั้นตั้งรูปที่ 1 ขึ้นตอนในการประมาณความเข้มแสงอาทิตย์ที่ตกกระทบทั่วประเทศไทย แบ่งเป็น 3 ส่วน ได้แก่ การเตรียมชุดข้อมูล การประมาณแบบจำลองการลดทอนความเข้มแสงด้วยเมฆ และการนำแบบจำลองไปใช้งาน

การเตรียมชุดข้อมูล ชุดข้อมูลในการประมาณแบบจำลองการลดทอนความเข้มแสงด้วยเมฆ ประกอบด้วยตัวแปรเข้าคือ ความเข้มแสงในสภาวะฟ้าใส (I_{clr}) ดัชนีเมฆ (CI) และตัวแปรออกคือ ความเข้มแสงที่วัดได้จากภาคพื้นดิน (I) การประมาณความเข้มแสงอาทิตย์ทั่วประเทศไทยจำเป็นต้องมีการเชื่อมโยงตำแหน่งบนประเทศไทยและภาพถ่ายเมฆจากดาวเทียม ประเทศไทยได้ถูกตีกริดโดยใช้ระบบซุมระดับ 14 ประเทศไทยมีกริดทั้งหมดประมาณ 270,000 กริด แต่ละกริดมีขนาด $2.4 \times 2.4 \text{ km}^2$ มีรหัสกริด และพิกัดละติจูดและลองจิจูดสำหรับแต่ละกริดเป็นแบบหนึ่งต่อหนึ่ง มีข้อมูลที่เชื่อมโยงระหว่างรหัสกริด ไปยังตำแหน่งแถวและคอลัมน์บนรูปภาพ นอกจากนี้เมื่อทราบตำแหน่งพิกัดละติจูดและลองจิจูด สามารถทราบค่าความสูงจากระดับน้ำทะเล (h) ที่ตำแหน่งนั้นได้จากการเรียก เอพีโอจาก SRTM30m [6] ดังตารางที่ 2

ตารางที่ 2: ตัวอย่างข้อมูลเชื่อมโยงตำแหน่งระหว่างประเทศไทยและภาพถ่ายเมฆ

รหัสกริด	ละติจูด	ลองจิจูด	h (m)	แถวบนรูปภาพ	คอลัมน์บนรูปภาพ
13220103001320	20.4990	97.3059	1134.0	455	685
13220103001321	20.4990	97.3278	1297.0	455	686
13220103001330	20.4990	97.3498	1458.0	455	687
13220103001331	20.4990	97.3718	1431.0	455	688
13220103010220	20.4990	97.3938	1448.0	455	689

หลังจากเชื่อมโยงตำแหน่งบนประเทศไทยและภาพถ่ายเมฆจากดาวเทียม สามารถคำนวณ I_{clr} ของแต่ละเวลา พิกัดละติจูด ลองจิจูด และ h จาก (10) และในส่วนของารสกัดดัชนีเมฆ เมื่ออ่านค่าจากภาพถ่ายเมฆพบว่าค่าที่อ่านเป็นเมทริกซ์ที่มี 2 เลเยอร์ ใช้เลเยอร์ที่ 1 เป็นค่าความหนาแน่นของเมฆ และเลือกค่าพิกเซลในตำแหน่งที่มีลำดับแถวและคอลัมน์ตรงกับรหัสกริดของแต่ละสถานีการวัดค่าความเข้มแสงทั้ง 56 สถานี หมายความว่าแต่ละภาพถ่ายมีค่าพิกเซลที่ถูกเลือกทั้งหมด 56 ค่า และนำค่าของพิกเซลที่เลือกคำนวณ CI ดัง (9)



รูปที่ 9: แผนภาพขั้นตอนฝึก ตรวจสอบ ทดสอบ และเลือกแบบจำลองการลดทอนความเข้มแสงด้วยเมฆ

การประมาณแบบจำลองการลดทอนความเข้มแสงด้วยเมฆ ภาพรวมในการประมาณแบบจำลองการลดทอนความเข้มแสงด้วยเมฆ ขั้นตอนแรกการตรวจสอบแบบจำลอง คือการปรับพารามิเตอร์ของแต่ละแบบจำลอง เพื่อคัดเลือกพารามิเตอร์ที่ดีที่สุดของแต่ละแบบจำลอง ขั้นตอนที่สอง การประมาณแบบจำลอง คือการฝึกแบบจำลองที่กำหนดค่าพารามิเตอร์ที่ดีที่สุดจากขั้นตอนที่หนึ่ง ขั้นตอนที่สามคือการทดสอบและคัดเลือกแบบจำลอง คือการทดสอบแบบจำลองที่ถูกประมาณ และเปรียบเทียบสมรรถนะเพื่อคัดเลือกแบบจำลองที่ดีที่สุดดังรูปที่ 9

ชุดข้อมูล (Dataset) ถูกแบ่งออกเป็น 2 ส่วนได้แก่ ชุดข้อมูลฝึก (training set) และชุดข้อมูลทดสอบ (testing set) ในอัตราส่วน 90:10 ตามลำดับ และคละชุดข้อมูลจากชุดข้อมูลทั้งหมดที่แบ่งตามวันและสถานี เพื่อให้ขั้นตอนการฝึกแบบจำลองและขั้นตอนการทดสอบแบบจำลอง ประมวลผลจากข้อมูลพื้นที่และสภาพอากาศที่หลากหลาย ชุดข้อมูลฝึกใช้ในการฝึกแบบจำลอง และใช้สำหรับตรวจสอบแบบจำลองที่มีการปรับพารามิเตอร์ ชุดข้อมูลฝึกบางส่วนจะถูกแบ่งออกเป็นชุดข้อมูลตรวจสอบ เพื่อคัดเลือกพารามิเตอร์ที่ดีที่สุด และชุดข้อมูลทดสอบใช้เพื่อทดสอบสมรรถนะของแบบจำลองที่ประมาณได้ และเปรียบเทียบสมรรถนะกับแบบจำลองอื่น

การตรวจสอบแบบจำลอง จะเห็นว่าแบบจำลองการถดถอยเชิงเส้นอย่างง่าย การถดถอยเชิงพหุนาม และป่าสุ่ม มีการคัดเลือกพารามิเตอร์ได้แก่ ตัวแปรเข้า อันดับของพหุนาม และไฮเปอร์พารามิเตอร์ตามลำดับ ดังนั้นจึงต้องตรวจสอบแบบจำลองเพื่อคัดเลือกพารามิเตอร์ที่ทำให้แบบจำลองมีสมรรถนะที่ดีที่สุด ใช้วิธีในการตรวจสอบแบบจำลองคือ 5-fold cross validation และมีตัวชี้วัดคือค่าความคลาดเคลื่อนสัมบูรณ์เฉลี่ย (mean absolute error, MAE) และในแบบจำลองถดถอยเชิงเส้น ใช้วิธีในการตรวจสอบแบบจำลอง model selection criteria ประกอบด้วย

แบบจำลองการถดถอยเชิงเส้น จาก (11) จะเห็นว่าแบบจำลองมีการเลือกตัวแปรเข้าทั้งหมด 7 รูปแบบ ดัง (12) ทำการตรวจสอบแบบจำลองโดยวิธี 5-fold cross validation และ model selection criteria เพื่อพิจารณาค่า MAE, AIC, BIC สำหรับแต่ละตัวแปรเข้าซึ่งมีทั้งหมด 7 ค่า และนำมาเปรียบเทียบ เพื่อคัดเลือกตัวแปรเข้าที่ทำให้ 5-fold cross validation error และ model selection scores เหมาะสมที่สุด

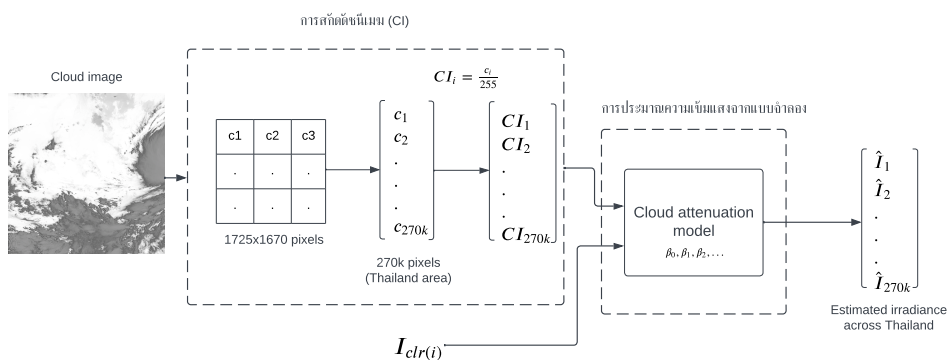
แบบจำลองการถดถอยเชิงพหุนาม จาก (13) จะเห็นว่าแบบจำลองมีการเลือกอันดับของพหุนาม (m) แบบจำลองถดถอยเชิงพหุนามจะถูกปรับเปลี่ยนค่า $m = 1, 2, 3, 4, 5, \dots, 15$ ที่แต่ละค่า m ทำการตรวจสอบแบบจำลองโดยวิธี 5-fold Cross validation เพื่อพิจารณาค่า MAE ที่แต่ละค่า m ซึ่งจะมีค่า MAE ทั้งหมด 15 ค่า นำมาเปรียบเทียบ และคัดเลือกค่า m ที่ทำให้ cross validation error น้อยที่สุด

แบบจำลองการสุ่มป่าไม้ ไฮเปอร์พารามิเตอร์ของแบบจำลองคือ $n_{\text{tree}}, d, p, q, r$ ในการปรับไฮเปอร์พารามิเตอร์ กำหนด $n_{\text{tree}} = 100, q = 2, r = 1$ และปรับเลือกค่า p ด้วยวิธี 3-fold cross validation เมื่อเลือก p ที่ทำให้ cross validation error น้อยที่สุดและปรับเปลี่ยน d, q, r ที่เป็นไปได้ทั้งหมดในชุดตั้ง (16) เพื่อคัดเลือกไฮเปอร์พารามิเตอร์ที่ดีที่สุด

$$d = \{10, 40, 80\}, q = \{2, 4, 10\}, r = \{1, 2, 4\} \quad (16)$$

หลังจากคัดเลือกพารามิเตอร์ที่ดีที่สุดของแต่ละแบบจำลอง ชุดข้อมูลฝึกจะนำมาฝึกให้กับแบบจำลองอีกครั้ง เพื่อที่จะประมาณแบบจำลอง และขั้นตอนต่อไป ชุดข้อมูลทดสอบจะถูกนำมาทดสอบแบบจำลองที่ถูกประมาณทั้งหมด และเปรียบเทียบค่า MAE ของแต่ละแบบจำลอง เพื่อคัดเลือกแบบจำลองที่ถูกประมาณที่มีค่า MAE ต่ำที่สุด และนำไปเปรียบเทียบกับแบบจำลองรุ่นที่ 2

การนำแบบจำลองไปใช้งาน (Implementation) เมื่อได้แบบจำลองที่นำเสนอแล้ว ขั้นตอนการนำแบบจำลองไปใช้งานของแบบจำลองการลดทอนความเข้มแสงด้วยเมฆ บอกถึงขั้นตอนของการได้มาซึ่ง I ผ่านแบบจำลองการลดทอนความเข้มแสงด้วยเมฆที่นำเสนอ ตัวแปรเข้าของแบบจำลองมีสองตัวแปร ได้แก่ ภาพถ่ายจากดาวเทียมฮิมาวาริ 8 ณ เวลาที่สนใจ และ I_{clr} ณ เวลาและสถานที่ที่สนใจ เริ่มจากรับภาพถ่ายเมฆเข้าโปรแกรมที่สร้างไว้ในรูปแบบ Python เพื่ออ่านค่าความหนาแน่นของเมฆในแต่ละพิกเซลของรูป มีทั้งหมด 1725×1670 พิกเซล จากนั้นแปลงค่าความหนาแน่นเมฆในพิกเซลนั้น เป็นค่าดัชนีเมฆตั้ง (9) เลือกเฉพาะพิกเซลที่ครอบคลุมพื้นที่ของประเทศไทยมาประมวลผล มีทั้งหมดประมาณ 270,000 กริด ในแต่ละกริดมี ละติจูด ลองจิจูด และ h แสดงตัวอย่างดังตารางที่ 2 จากนั้นสามารถนำข้อมูลละติจูด ลองจิจูด และ h คำนวณ I_{clr} ดัง (10) นำค่า $CI(t-1)$ และ $I_{\text{clr}}(t)$ เป็นตัวแปรเข้าของแบบจำลอง เพื่อให้แสดงค่าประมาณความเข้มแสงอาทิตย์ของจุดและเวลานั้น $I(t)$ และคำนวณทั้งหมด 270,000 กริด เพื่อแสดงผลการประมาณ I ทั่วประเทศไทยดังรูปที่ 10



รูปที่ 10: แผนภาพขั้นตอนการนำแบบจำลองไปใช้งาน

2.4 รายละเอียดชุดข้อมูล

ข้อมูลภาพถ่าย ภาพถ่ายเมฆจากดาวเทียมฮิมาวาริ 8 ไฟล์ภาพอยู่ในสกุล (.TIFF) เป็นภาพระดับสีเทา ขนาด 1725×1670 พิกเซล มี 2 เลเยอร์ โดยเลเยอร์ของความหนาแน่นของเมฆอยู่ในเลเยอร์ที่ 1 ดังนั้นจึงเลือกใช้เลเยอร์

นี้ ในแต่ละจุดพิกเซลของภาพจะมีค่าความหนาแน่นของเมฆอยู่ระหว่าง 0 ถึง 255 มีค่าความละเอียดเชิงพื้นที่เท่ากับ $2 \times 2 \text{ km}^2$ ซึ่งข้อมูลแบ่งออกเป็น 2 ประเภท ตามสถานีรับภาพ ได้แก่ สถานีรับภาพจากสำนักงานพัฒนาเทคโนโลยีอวกาศและภูมิสารสนเทศ (องค์การมหาชน) หรือ GISDA ซึ่งเป็นหน่วยงานของรัฐในรูปแบบขององค์การมหาชน ที่ให้บริการข้อมูลสารสนเทศและบริการวิชาการต่าง ๆ รวมถึงการวิจัยและพัฒนาเทคโนโลยีอวกาศให้เป็นประโยชน์ต่อประชาชน และสถานีรับภาพจากคณะวิศวกรรมไฟฟ้า จุฬาลงกรณ์มหาวิทยาลัย หรือ CUEE

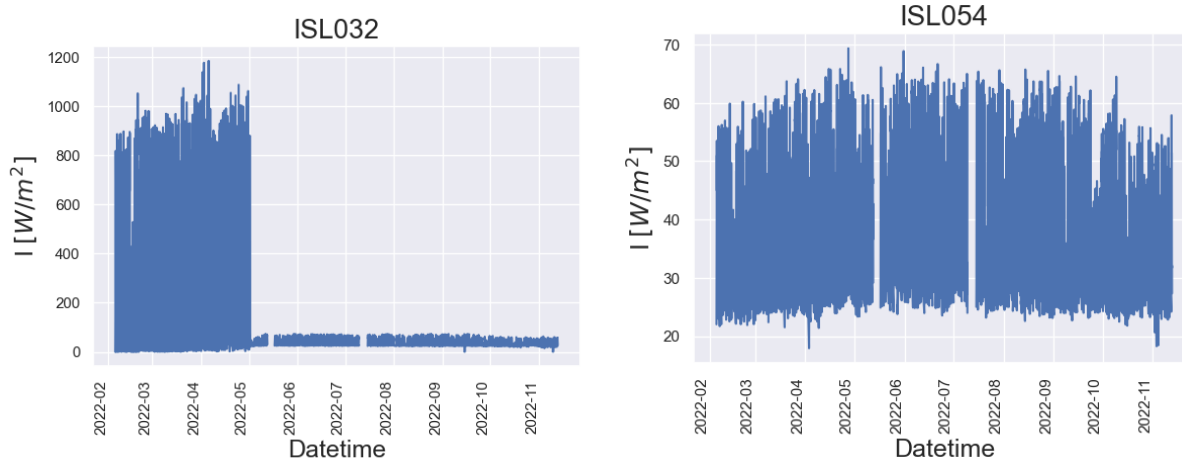
- สถานีรับ GISDA ให้ข้อมูลช่วง 2021/6/15 - 2022/1/26
- สถานีรับ CUEE ให้ข้อมูลช่วง 2022/2/5 - ปัจจุบัน

สำหรับโครงการที่นำเสนอ ใช้ภาพถ่ายเมฆที่รับจากสถานีรับ CUEE ช่วง 2022/02/05-2022/10/22 ค่าความละเอียดของภาพเท่ากับ 10 นาที เนื่องจากมีภาพบางส่วนสูญหายและภาพที่ใช้งานไม่ได้ หลังจากการคัดแยกภาพเหล่านั้นแล้วเหลือภาพในการใช้งานทั้งหมด 17,443 ภาพ

ค่าความเข้มแสงที่วัดได้จากภาคพื้นดิน (I) คือค่าความเข้มแสงอาทิตย์ที่วัดได้จากตัวรับรู้บนภาคพื้นดิน เป็นข้อมูลอนุกรมเวลา ให้ข้อมูลตั้งแต่ 2022 ถึง ปัจจุบัน มีค่าความละเอียดเชิงเวลาเท่ากับ 15 นาที และข้อมูลถูกวัดจาก IMPACT SOLAR Station ทั้งหมด 56 สถานี (ISL 56 stations) พิจารณา I แบ่งตามรุ่นของแบบจำลอง

- ใช้ข้อมูล I ช่วง 2020/11/16 - 2022/09/20 เพื่อวิเคราะห์ความคลาดเคลื่อนของแบบจำลองรุ่นที่ 1
- ใช้ข้อมูล I ช่วง 2022/09/20 - 2022/11/12 เพื่อวิเคราะห์ความคลาดเคลื่อนของแบบจำลองรุ่นที่ 2
- ใช้ข้อมูล I ช่วง 2022/02/05 - 2022/10/22 เพื่อฝึกและทดสอบแบบจำลองที่นำเสนอ

การฝึกแบบจำลองที่นำเสนอใช้ข้อมูล I ช่วงเวลา 8:00 น. ถึง 16:00 น. เนื่องจากค่าความเข้มแสงในสภาวะฟ้าใสในช่วงเช้า (6:00 น. ถึง 8:00 น.) และช่วงเย็น (16:00 น. ถึง 18:00 น.) มีค่าเข้าใกล้ศูนย์ ทำให้ค่า k มีค่าสูงมาก ข้อมูลช่วงเวลานี้มีประโยชน์น้อยต่อการฝึกแบบจำลอง จึงตัดช่วงเวลาเหล่านี้จากการศึกษา นอกจากนี้จากการสำรวจข้อมูลของแต่ละสถานีพบว่า สถานีที่ 45 ไม่มีข้อมูลการวัด และค่าความเข้มแสงที่วัดได้จากสถานีที่ 32 และ 54 มีข้อมูลบางส่วนผิดพลาดดังรูปที่ 11 จึงคัดข้อมูลที่วัดได้จากสถานีที่ 45 และ 54 ออก สำหรับสถานีที่ 32 นำไปใช้เฉพาะค่าที่ไม่เกิดข้อผิดพลาด



(a) กราฟ I ที่สถานีที่ 32

(b) กราฟ I ที่สถานีที่ 54

รูปที่ 11: กราฟความเข้มแสงที่วัดได้จากสถานีที่ 32 และ 54

ค่าประมาณความเข้มแสง (I) แบ่งตามรุ่นของแบบจำลองที่นำไปใช้งาน แบบจำลองรุ่นที่ 1 พัฒนาโดย นาย สรรสุด พรานนท์สถิตย์ และแบบจำลองรุ่นที่ 2 ที่พัฒนาโดย นายปรินทร์ มโนมัยเสาวภาคย์ดังตารางที่ 3

ตารางที่ 3: รายละเอียดข้อมูลที่แบ่งตามรุ่นของแบบจำลอง

แบบจำลอง	สถานีที่ใช้ข้อมูลฝึก	สถานีรับภาพ	ช่วงเวลาของข้อมูลฝึก	ช่วงเวลาที่ใช้งานแบบจำลอง
รุ่นที่ 1	1-20	GISDA	2020/06/10 - 2020/06/18	2022/11/16 - 2021/12/31
รุ่นที่ 2	1-56	GISDA	2020/10/01 - 2021/08/01	2022/09/07 - ปัจจุบัน
รุ่นที่ 3	1-56	CUEE	2022/02/05 - 2022/10/22	-

เวลาของข้อมูล พิจารณาเวลาของข้อมูล มีสองประเด็นที่สำคัญที่ต้องคำนึง ได้แก่ (1). ค่าความละเอียดเชิงเวลา (Temporal resolution) ที่แตกต่างกันระหว่างภาพถ่ายเมฆกับค่าความเข้มแสงอาทิตย์ที่วัดได้ และ (2). เวลาประวิงของภาพถ่ายเมฆในการใช้งานจริง

1. **ค่าความละเอียดเชิงเวลา (Temporal resolution)** ค่าความละเอียดเชิงเวลาของภาพถ่ายเมฆและค่าความเข้มแสงอาทิตย์ที่วัดได้มีความแตกต่างกัน มีค่าความละเอียดเชิงเวลาเท่ากับ 10 นาทีและ 15 นาทีตามลำดับ ดังนั้นจึงเลือกค่าความละเอียดเชิงเวลาของคู่ข้อมูลของทั้งสองเท่ากับ 30 นาทีดังตารางที่ 4

ตารางที่ 4: ความละเอียดเชิงเวลาและพื้นที่ของข้อมูล

ข้อมูล	ความละเอียดเชิงเวลา(นาที)	ความละเอียดเชิงพื้นที่(km ²)	แหล่งที่มา
ภาพถ่ายเมฆ	10	2x2	ดาวเทียมฮิมาวาริ 8
I	15	-	สถานี ISL 56 สถานี
\hat{I}	30	2.4x2.4	แบบจำลองรุ่นที่ 1 และรุ่นที่ 2

2. **เวลาประวิงของภาพถ่ายเมฆ (Delay time)** ในการใช้งานจริงพบว่าภาพถ่ายเมฆจากดาวเทียมฮิมาวาริ 8 จะผ่านการประมวลผลก่อนที่จะส่งมายังสถานีรับ CUEE จึงมีเวลาประวิงของภาพถ่ายเมฆ 30 นาที ดังนั้น $\hat{I}(t)$ จะถูกประมาณจาก $CI(t - 1)$ ดังนั้น ในขั้นตอนฝึกแบบจำลองจะจับคู่ข้อมูลทั้งสองดังตารางที่ 5

ตารางที่ 5: เวลาของคู่ข้อมูลในการฝึกแบบจำลอง

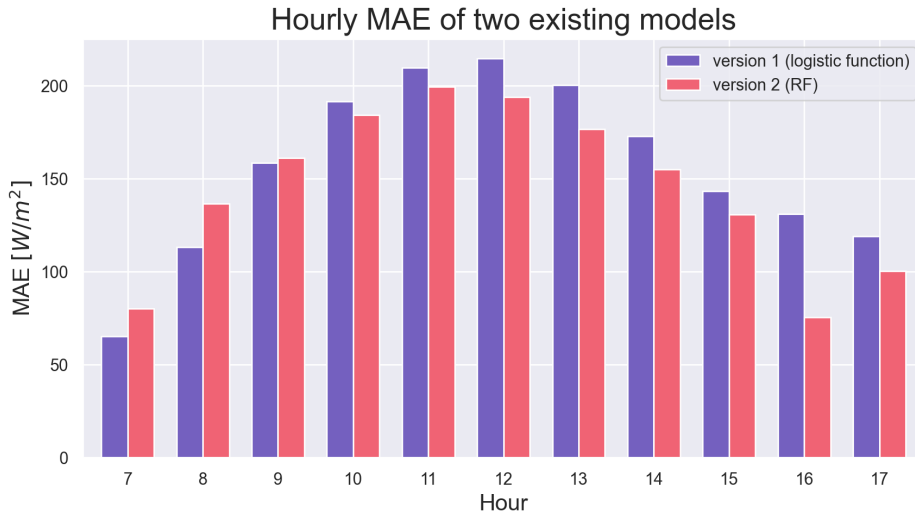
ณ เวลา t	$I(t)$	$CI(t - 1)$
8:00	$I(8:00)$	$CI(7:30)$
8:30	$I(8:30)$	$CI(8:00)$
9:00	$I(9:00)$	$CI(8:30)$
9:30	$I(9:30)$	$CI(9:00)$

3 ผลลัพธ์จากการดำเนินงานเบื้องต้น

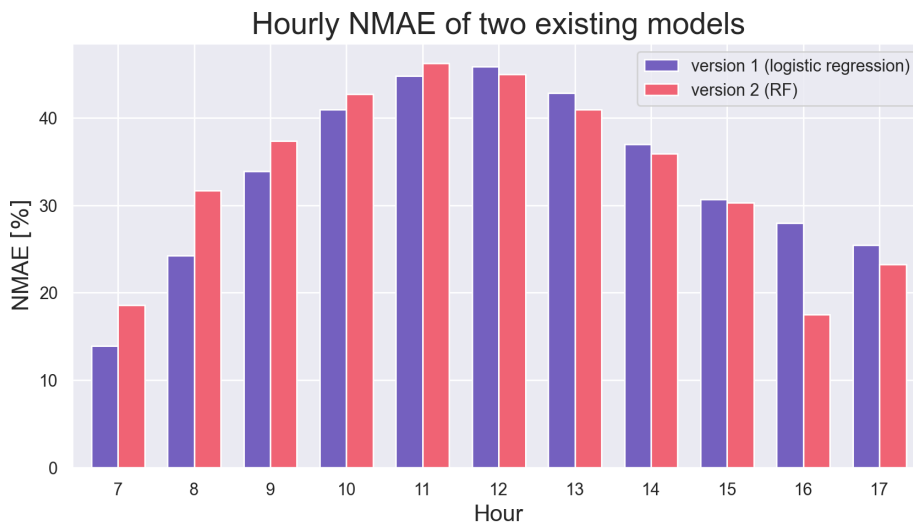
ผลลัพธ์จากการดำเนินงานประกอบด้วย 2 ส่วน ได้แก่ ผลการตรวจสอบสมรรถนะของแบบจำลองจากแบบจำลองรุ่นที่ 1 และรุ่นที่ 2 เพื่อบ่งชี้สมรรถนะของแบบจำลองดังกล่าว และผลการทดลองแบบจำลองที่นำเสนอ เพื่อแสดงถึงสมรรถนะของแบบจำลองที่นำเสนอ ผลลัพธ์จากการดำเนินงานเบื้องต้นทั้งหมดคิดเป็นความก้าวหน้า 60% ของแผนการดำเนินงานทั้งหมด

3.1 ผลการตรวจสอบแบบจำลองรุ่นที่ 1 และรุ่นที่ 2

การตรวจสอบสมรรถนะของแบบจำลองจากแบบจำลองรุ่นที่ 1 และรุ่นที่ 2 แบ่งตามช่วงดังตารางที่ 3 มีตัวชี้วัดคือค่า MAE และค่า NMAE (Normalize mean absolute error) ที่ได้จากการหาร MAE ด้วยค่าเฉลี่ย I แบบจำลองรุ่นที่ 1 มีค่าเฉลี่ย MAE เท่ากับ 159.37 W/m^2 และแบบจำลองรุ่นที่ 2 มีค่าเฉลี่ย MAE เท่ากับ 149.40 W/m^2 แบบจำลองรุ่นที่ 2 มีค่า MAE ต่ำกว่าแบบจำลองรุ่นที่ 1 คิดเป็นร้อยละ 6.26 เทียบกับ MAE ของแบบจำลองรุ่นที่ 1 และสมรรถนะรายชั่วโมงของแบบจำลองรุ่นที่ 1 และรุ่นที่ 2 แสดงดังรูปที่ 12



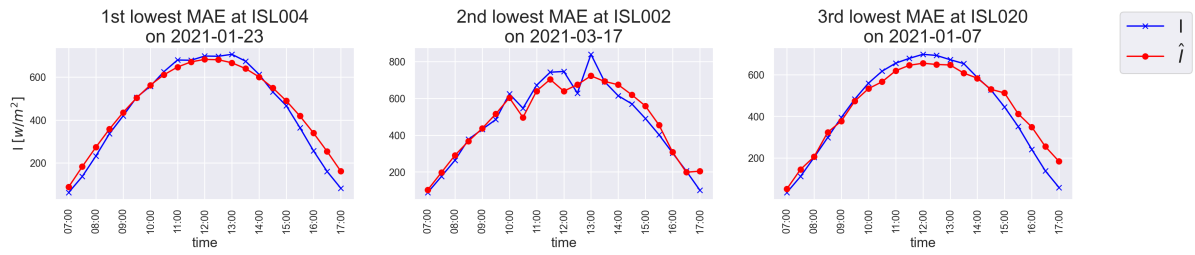
(a) กราฟ MAE รายชั่วโมง ของแบบจำลองรุ่นที่ 1 และรุ่นที่ 2



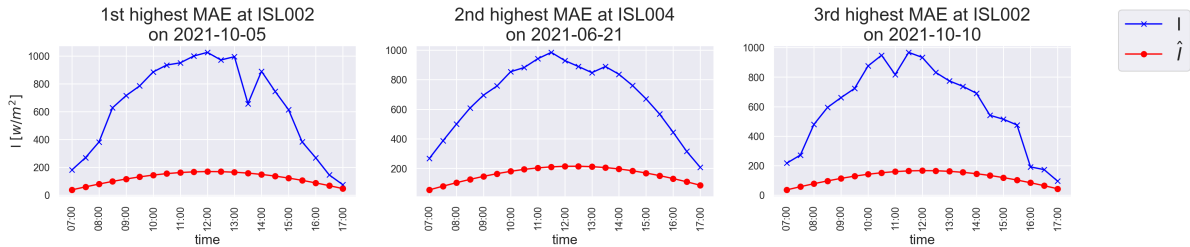
(b) กราฟ NMAE รายชั่วโมง ด้วยการนำ MAE ให้เป็นมาตรฐานเทียบกับค่าเฉลี่ย I เท่ากับ 467.59 W/m^2 (แบบจำลองรุ่นที่ 1) 431.25 W/m^2 (แบบจำลองรุ่นที่ 2)

รูปที่ 12: กราฟ MAE และ NMAE รายชั่วโมงของแบบจำลองรุ่นที่ 1 และรุ่นที่ 2

ผลการตรวจสอบแบบจำลองรุ่นที่ 1 ดังรูปที่ 13 แสดงผลการประมาณ I ของวันและสถานที่ที่มี MAE ต่ำที่สุดและมากที่สุด 3 อันดับแรก ผลการประมาณ I ของวันที่มีค่า MAE น้อยที่สุด 3 อันดับแรกดังรูปที่ 13(a) จะเห็นว่าช่วงบ่าย I มีแนวโน้มผิดปกติ มีสมมติฐานว่าแบบจำลองฟ้าใสมีความผิดพลาด ในส่วนผลการประมาณ I ของวันที่มีค่า MAE มากที่สุด 3 อันดับแรกดังรูปที่ 13(b) จะเห็นว่าแนวโน้มของค่า I ต่ำเกินไป มีสมมติฐานว่าเกิดจากเวลานั้นมีค่า CI มาก จึงทำให้แบบจำลองรุ่นที่ 1 (ลอจิสติก) ให้ค่า k ที่ต่ำ ส่งผลให้ I มีค่าต่ำไปด้วย



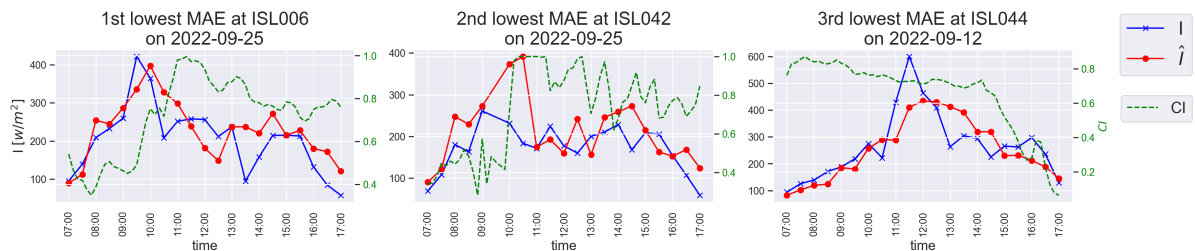
(a) 3 วันและสถานที่ที่มีค่า MAE ต่ำที่สุด



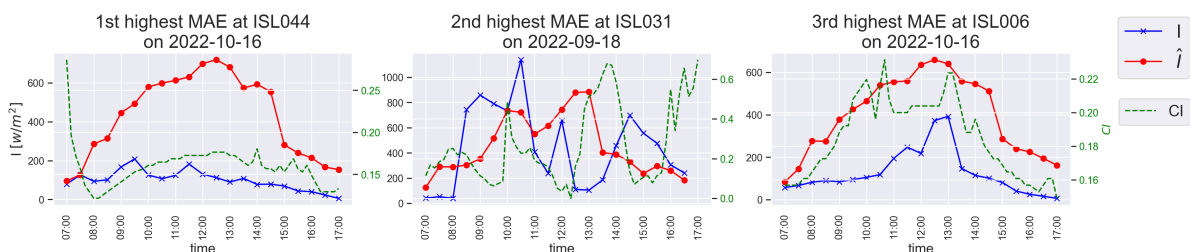
(b) 3 วันและสถานที่ที่มีค่า MAE สูงที่สุด

รูปที่ 13: ผลการประมาณ I จากแบบจำลองรุ่นที่ 1

ผลการทดสอบแบบจำลองรุ่นที่ 2 ดังรูปที่ 14 แสดงผลการประมาณ I ของวันและสถานที่ที่มี MAE ต่ำที่สุด 3 อันดับแรก และมากที่สุด 3 อันดับแรก ผลการประมาณ I ของวันที่มีค่า MAE น้อยที่สุด 3 อันดับแรกดังรูปที่ 14(a) ถึงแม้เป็นวันที่มีค่าความแปรปรวนของความหนาแน่นเมฆสูง แบบจำลองรุ่นที่ 2 (ป่าสุ่ม) ยังคงให้ค่าการประมาณที่ใกล้เคียงกับค่าวัดจริง ผลการประมาณ I ของวันที่มีค่า MAE มากที่สุด 3 อันดับแรกดังรูปที่ 14(b) พบว่า \hat{I} และ CI ให้ค่าที่สอดคล้องกัน กล่าวคือเมื่อ CI น้อย ค่า \hat{I} ถูกลดทอนน้อย แบบจำลองจึงให้ค่า \hat{I} ตาม CI ที่อ่านได้จากภาพถ่ายเมฆ



(a) 3 วันและสถานที่ที่มีค่า MAE ต่ำที่สุด



(b) 3 วันและสถานที่ที่มีค่า MAE สูงที่สุด

รูปที่ 14: ผลการประมาณ I จากแบบจำลองรุ่นที่ 2

3.2 ผลการทดลองแบบจำลองที่นำเสนอ

ผลการทดลองแบบจำลองที่นำเสนอประกอบด้วย การตรวจสอบแบบจำลอง คัดเลือกแบบจำลองที่มีสมรรถนะดีที่สุด และการเปรียบเทียบระหว่างแบบจำลองที่ใช้งานอยู่กับแบบจำลองที่นำเสนอ

แบบจำลองการถดถอยเชิงเส้น การเลือกตัวแปรเข้าตั้ง (12) พบว่าตัวแปรเข้า $I_{clr}(t)$, $CI(t-1)$, $I_{clr}(t)CI(t-1)$ ทำให้แบบจำลองมีสมรรถนะดีที่สุดเพราะว่ามีค่า MAE AIC และ BIC น้อยที่สุดดังตารางที่ 6 และสามารถประมาณค่าพารามิเตอร์ของแบบจำลองได้ดังนี้

$$\hat{I}(t) = \beta_0 + \beta_1 I_{clr}(t) + \beta_2 CI(t-1) + \beta_3 I_{clr}(t)CI(t-1) \quad (17)$$

$$\beta_0 = 93.218, \beta_1 = 0.82103, \beta_2 = -180.317, \beta_3 = -0.43357$$

β_0 มีหน่วยเป็น W/m^2 คือค่าคงที่การถดถอย (intercept term) β_1 ไม่มีหน่วยและมีค่าจะเข้าใกล้ 1 ส่งผลให้ความสัมพันธ์มีแนวโน้มตามค่าความเข้มแสงในสภาวะฟ้าใส β_2 มีหน่วยเป็น W/m^2 และ β_3 ไม่มีหน่วย β_2, β_3 มีค่าเป็นลบเนื่องจากพจน์ของ $CI(t-1)$ กล่าวคือ เมื่อค่า $CI(t-1)$ เพิ่มขึ้น ส่งผลให้ค่าประมาณค่าความเข้มแสงอาทิตย์ลดลง สังเกตว่า $CI(t-1)$ มีค่ามากที่สุดเท่ากับ 1 ดังนั้น β_2 สามารถลดทอน $\hat{I}(t)$ ได้มากที่สุดเพียง $180.32 W/m^2$ การมีพจน์ $I_{clr}(t)CI(t-1)$ จะช่วยลดทอน $\hat{I}(t)$ แตกต่างกันตามเวลาและตำแหน่งเป็นผลมาจากที่มี $CI(t-1)$ เป็นตัวประกอบ ส่งผลให้แบบจำลองการถดถอยเชิงเส้นมีสมรรถนะดีขึ้น

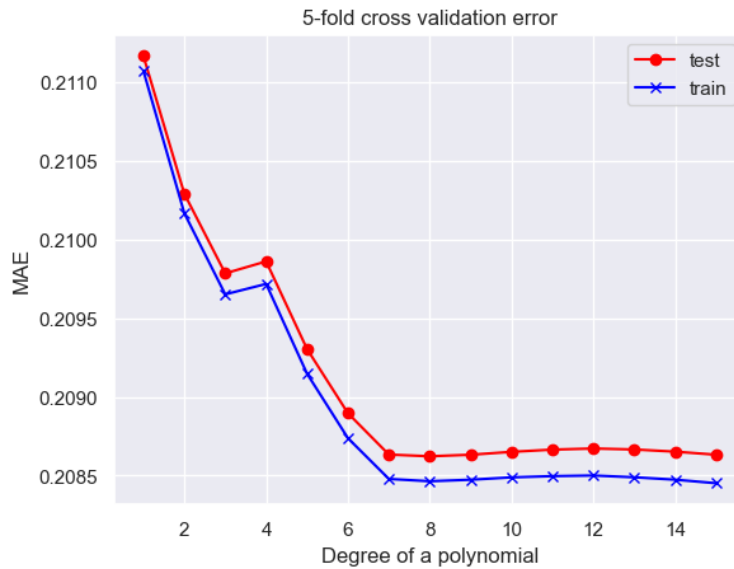
ตารางที่ 6: ค่า cross validation error และ model selection scores ของแบบจำลองการถดถอยเชิงเส้น เมื่อมีการปรับเปลี่ยน ตัวแปรเข้า

ตัวแปรเข้า	training error	cross validation error	AIC	BIC
$I_{clr}(t)$	195.00	195.09	13.741	13.741
$CI(t-1)$	195.94	196.01	13.759	13.759
$I_{clr}(t)CI(t-1)$	213.20	213.29	13.902	13.902
$I_{clr}(t), CI(t-1)$	153.03	153.10	13.356	13.357
$I_{clr}(t), I_{clr}(t)CI(t-1)$	151.81	151.89	13.347	13.348
$CI(t-1), I_{clr}(t)CI(t-1)$	181.81	181.89	13.619	13.619
$I_{clr}(t), CI(t-1), I_{clr}(t)CI(t-1)$	151.27	151.36	13.344	13.344

แบบจำลองการถดถอยเชิงพหุนาม จากการปรับค่าอันดับของพหุนาม $m = 1, 2, \dots, 15$ ดัง (13) จะเห็นว่าจุดที่อันดับของพหุนามมีค่าเท่ากับ 8 ทำให้แบบจำลองมีสมรรถนะที่ดี และไม่ซับซ้อนเกินไปดังรูปที่ 15 จึงเลือกอันดับของพหุนามเท่ากับ 8 และสามารถประมาณพารามิเตอร์ของแบบจำลองได้เป็น

$$\hat{k}(t) = \beta_0 + \beta_1 CI(t-1) + \beta_2 CI(t-1)^2 + \dots + \beta_8 CI(t-1)^8 \quad (18)$$

$$\beta_0 = 0.94333, \beta_1 = 3.5608, \beta_2 = -73.378, \beta_3 = 475.15, \beta_4 = -1575.4, \\ \beta_5 = 2951.2, \beta_6 = -3155.7, \beta_7 = 1795.0, \beta_8 = -421.055$$



รูปที่ 15: กราฟค่า cross validation error ของแบบจำลองถดถอยเชิงพหุนามเมื่อปรับเปลี่ยนอันดับพหุนาม

แบบจำลองป่าสุ่ม จากการปรับค่าไฮเปอร์พารามิเตอร์ กำหนด $n_{\text{tree}} = 100$, $q = 2$, $r = 1$ และปรับเปลี่ยนค่า p พบว่า $p = 9$ ทำให้ cross validation error น้อยที่สุด หลังจากนั้นจากการปรับไฮเปอร์พารามิเตอร์ที่เป็นไปได้ทั้งหมดตั้ง (16) พบว่า $d = 40$, $q = 2$, $r = 2$ เป็นไฮเปอร์พารามิเตอร์ที่ดีที่สุดของชุดที่เป็นไปได้ ทำการปรับ q , r , d แบบละเอียดอีกครั้ง กำหนด $n_{\text{tree}} = 100$, $d = 40$, $p = 9$ จะเห็นว่า $q = 4$, $r = 2$, $d = 37$ ทำให้แบบจำลองป่าสุ่มมี cross validation error น้อยที่สุดดังตารางที่ 7 และดังตารางที่ 8 และเลือกไฮเปอร์พารามิเตอร์ที่ดีที่สุดดัง (19)

ตารางที่ 7: ค่า training error และ cross validation error ของแบบจำลองป่าสุ่มเมื่อมีการปรับ q , r

q	r	training error	cross validation error
2	2	54.985	122.24
3	2	54.985	122.24
4	2	54.985	122.24
5	2	57.324	122.32
4	1	46.043	122.40
4	2	54.985	122.24
4	3	64.500	122.67

ตารางที่ 8: ค่า training error และ cross validation error ของแบบจำลองป่าสุ่มเมื่อมีการปรับ d

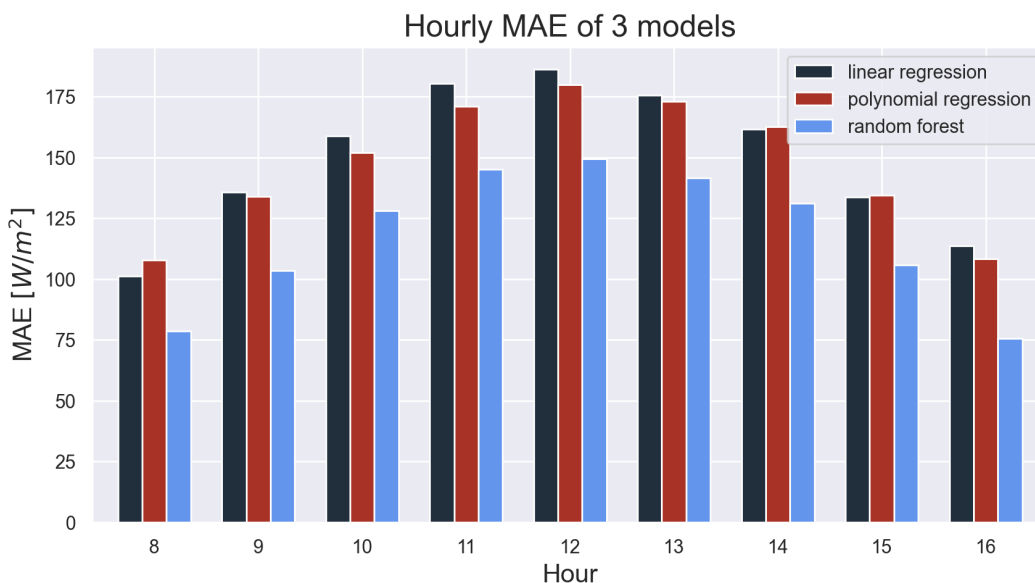
d	training error	cross validation error
34	55.115	122.28
35	55.177	122.29
36	55.050	122.25
37	55.025	122.17
38	55.001	122.24
39	54.979	122.27
40	54.985	122.24
41	54.983	122.26
42	54.974	122.26

$$n_{\text{tree}} = 100, d = 37, p = 9, q = 4, r = 2 \quad (19)$$

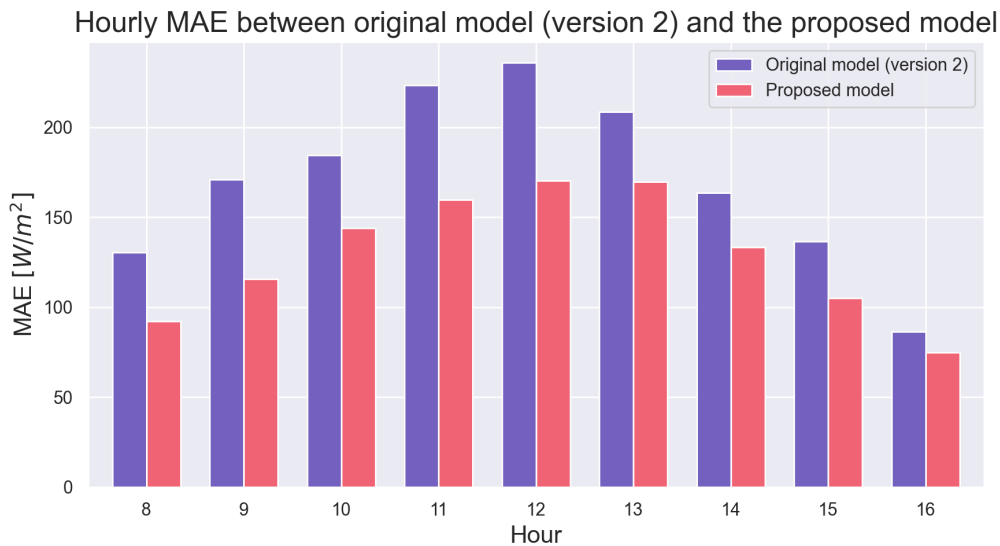
การคัดเลือกแบบจำลองที่มีสมรรถนะดีที่สุด หลังจากการคัดเลือกไฮเปอร์พารามิเตอร์ที่ดีที่สุดสำหรับแต่ละแบบจำลอง ค่า MAE ของแต่ละแบบจำลองดังตารางที่ 9 และแสดงค่า MAE รายชั่วโมงดังรูปที่ 16 จะเห็นว่าแบบจำลองป่าสุ่มมีสมรรถนะที่ดีกว่าแบบจำลองอื่นมาก จึงเลือกแบบจำลองป่าสุ่มในการนำไปใช้เป็นแบบจำลองการลดทอนความเข้มแสงด้วยเมฆ และนำไปเปรียบเทียบกับแบบจำลองที่ใช้งานอยู่ (รุ่นที่ 2)

ตารางที่ 9: ค่า MAE ของแต่ละแบบจำลองที่นำเสนอ

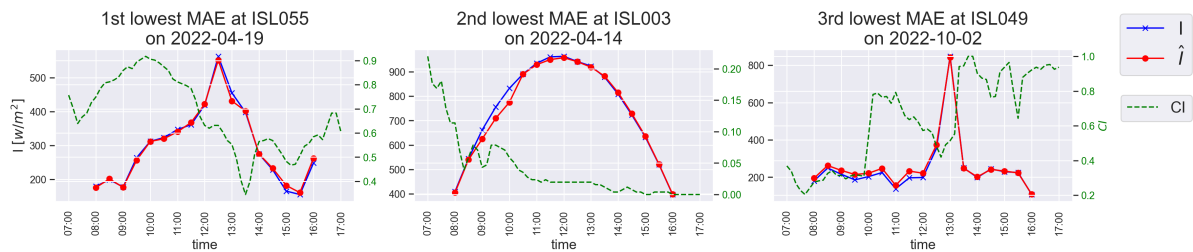
แบบจำลอง	MAE (W/m^2)		
	ช่วงเวลาทั้งวัน	ช่วงเวลากลางวัน	ช่วงเวลาเช้าและบ่าย
การถดถอยเชิงเส้นอย่างง่าย	151.68	176.00	130.17
การถดถอยเชิงพหุนาม	149.14	171.65	129.24
ป่าสุ่ม	119.97	141.88	100.61



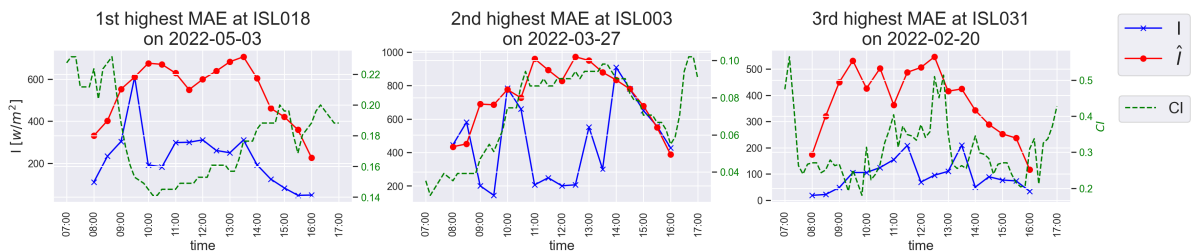
รูปที่ 16: กราฟค่า MAE รายชั่วโมงของ 56 สถานี ของแต่ละแบบจำลองที่นำเสนอ



รูปที่ 17: กราฟค่า MAE รายชั่วโมงของ 56 สถานี ของแบบจำลองรุ่นที่ 2 และแบบจำลองที่นำเสนอ



(a) 3 วันและสถานีที่มีค่า MAE ต่ำที่สุด



(b) 3 วันและสถานีที่มีค่า MAE สูงที่สุด

รูปที่ 18: ผลการประมาณ I จากแบบจำลองที่นำเสนอ

การเปรียบเทียบแบบจำลองที่ใช้งานอยู่ และแบบจำลองที่นำเสนอ ชุดข้อมูลที่ใช้ในการเปรียบเทียบคือ ชุดข้อมูลทดสอบของแบบจำลองที่นำเสนอเชื่อมกับชุดข้อมูล I ของแบบจำลองรุ่นที่ 2 ซึ่งมีจำนวนจุดข้อมูล ประมาณ 3000 จุด ผลการเปรียบเทียบแบบจำลองที่ใช้งานอยู่ และแบบจำลองที่เสนอพบว่า แบบจำลอง RF ที่ นำเสนอมีสรรณะดีกว่าแบบจำลองที่ใช้งานปัจจุบัน (รุ่นที่ 2) ดังรูปที่ 17 กล่าวคือแบบจำลองที่นำเสนอมีค่า MAE น้อยกว่าแบบจำลองรุ่นที่ 2 ดังรูปที่ 17 คิดเป็นร้อยละ 24.91 เทียบกับแบบจำลองรุ่นที่ 2 ผลการปรับปรุง นี้สามารถอธิบายได้เป็น 3 ประเด็นหลักดังนี้

- ประเด็นที่ 1 แบบจำลองที่นำเสนอใช้ค่าความสูงในแบบจำลองฟ้าใสของแต่ละพื้นที่จาก [6] ในขณะที่แบบจำลองรุ่นที่ 2 ใช้ค่าความสูงของแต่ละพื้นที่เท่ากันทั้งหมด เท่ากับ 35 เมตร

- ประเด็นที่ 2 สัมประสิทธิ์ความขุ่นมัวของชั้นบรรยากาศ (T_L) ในแบบจำลองฟ้าใส สามารถคำนวณ T_L จาก การแก้สมการการถดถอยเชิงเส้นดัง (10) แบบจำลองที่นำเสนอประมาณ T_L จากข้อมูลปัจจุบันได้เท่ากับ 4.7738 ขณะที่แบบจำลองรุ่นที่ 2 ประมาณ T_L จากข้อมูลในอดีตได้เท่ากับ 5.124 ค่า T_L จากแบบจำลอง ที่นำเสนอจึงมีค่าเป็นปัจจุบัน การปรับค่า T_L เป็นปัจจุบันมีความสำคัญ เนื่องจาก I_{clr} ในแบบจำลองฟ้าใส มีสภาพไวต่อ T_L พิจารณาการเปลี่ยนแปลงของ I_{clr} สัมพัทธ์ เมื่อ T_L เปลี่ยน ($\Delta I_{relative}$) ดัง (20)

$$\Delta I_{relative} = \frac{dI_{clr}(T_L)}{I_{clr}(T_L)} = -\exp(-h/1250) \cdot AM(t) \cdot dT_L \quad (20)$$

จะเห็นว่า $\Delta I_{relative}$ ขึ้นกับ $AM(t)$ และ h ดัง (20) เมื่อพิจารณาเชิงเวลาจะพบว่า $AM(t)$ มีค่าสูงในช่วง เช้าและเย็นดังรูปที่ 4 ทำให้ I_{clr} มีสภาพไวต่อ T_L ที่สูงกว่าช่วงกลางวัน ถ้าหาก T_L มีค่าไม่แม่นยำส่งผล ให้ช่วงเช้าและเย็น I มีความคลาดเคลื่อนมากกว่าช่วงกลางวัน เมื่อพิจารณาในเชิงพื้นที่พบว่าตำแหน่งที่มี h น้อย ส่งผลให้ I_{clr} มีสภาพไวต่อ T_L สูงกว่าตำแหน่งที่มี h มาก ถ้าหาก T_L มีค่าไม่แม่นยำส่งผลให้ ตำแหน่งที่มี h น้อย I จะมีค่าความคลาดเคลื่อนมากกว่าตำแหน่งที่มี h มาก และแบบจำลองที่นำเสนอใช้ ค่า h ที่ใกล้เคียงกับสภาพภูมิประเทศจริง และ T_L ที่ปัจจุบัน ดังนั้นแบบจำลองที่นำเสนอจึงมีสมรรถนะที่ ดีขึ้นจากแบบจำลองรุ่นที่ 2

- ประเด็นที่ 3 ดัชนีเมฆรอบจุดที่สนใจ แบบจำลองที่นำเสนอมีตัวแปรเข้าเพิ่มเติมจากแบบจำลองรุ่นที่ 2 คือ CI รอบจุดที่สนใจ มีสมมติฐานว่า CI รอบจุดที่สนใจส่งผลต่อ I ในเชิงพื้นที่ กล่าวคือเมื่อจุดที่สนใจ มี CI น้อย ขณะที่พื้นที่รอบ ๆ มี CI มาก จะส่งผลให้ I มีค่าน้อยลงเมื่อเทียบกับ I ขณะที่พื้นที่รอบ ๆ มี CI น้อย ด้วยเหตุผลดังกล่าวแบบจำลองที่นำเสนอจึงมีสมรรถนะที่ดีขึ้นจากแบบจำลองรุ่นที่ 2

ตัวอย่างความเข้มแสงอาทิตย์ที่ประมาณได้เทียบกับปริมาณเมฆแสดงดังรูปที่ 18 แบบจำลองทำงานสมเหตุ สมผล ค่าประมาณความเข้มแสงอาทิตย์มีความสัมพันธ์แบบแปรผกผันกับดัชนีเมฆ

4 บทสรุป

บทสรุปการดำเนินงานประกอบด้วย สรุปผลการดำเนินงาน เป็นการสรุปภาพรวมการดำเนินงานและผลลัพธ์ ทั้งหมดในภาคเรียนต้น และ แผนการดำเนินงาน เป็นการวางแผนการดำเนินงานในภาคเรียนปลาย เพื่อที่สามารถ จัดทำโครงการได้อย่างต่อเนื่องและเป็นระบบ

4.1 สรุปผลการดำเนินงาน

จากการวิเคราะห์ความคลาดเคลื่อนของการประมาณความเข้มแสงอาทิตย์ด้วยแบบจำลองในโครงการ Thailand Solar Map by CUEE รุ่นที่ 1 และรุ่นที่ 2 พบว่าในการประมาณความเข้มแสงมีค่าเฉลี่ย MAE เท่ากับ 159.37 W/m^2 และ 149.40 W/m^2 ตามลำดับ เนื่องจากแบบจำลองรุ่นที่ 1 เป็นแบบจำลองลอจิสติก พิจารณา ตัวแปรเข้า 1 ตัวแปร คือดัชนีเมฆ ($CI(t)$) ในขณะที่แบบจำลองรุ่นที่ 2 เป็นแบบจำลองป่าสุ่ม พิจารณาตัวแปร เข้าคือ ดัชนีเมฆ ($CI(t), CI(t-1)$) ความเข้มแสงสถานะฟ้าใส ($I_{clr}(t), I_{clr}(t-1)$) และลำดับชั่วโมง (HR) แบบ จำลองป่าสุ่มจึงมีความซับซ้อนมากกว่า ดังนั้นแบบจำลองรุ่นที่ 2 จึงมีสมรรถนะที่ดีกว่าแบบจำลองรุ่นที่ 1

จากการปรับปรุงแบบจำลองลดทอนความเข้มแสงจากเมฆ ในการศึกษาเปรียบเทียบ 3 แบบจำลองคือ แบบจำลองถดถอยเชิงเส้น แบบจำลองถดถอยเชิงพหุนาม และแบบจำลองป่าสุ่ม ในการประมาณความเข้มแสง มีค่าเฉลี่ย MAE ของทั้งวันเท่ากับ 151.68 W/m^2 149.14 W/m^2 และ 119.97 W/m^2 ตามลำดับ เนื่องจากแบบ จำลองป่าสุ่มมีสมรรถนะที่ดีที่สุดจึงเลือกแบบจำลองป่าสุ่ม และเปรียบเทียบแบบจำลองที่นำเสนอ (RF) และแบบ จำลองรุ่นที่ 2 พบว่า ในการประมาณความเข้มแสงค่าเฉลี่ย MAE ของแบบจำลองที่นำเสนอมีค่าน้อยกว่าแบบ จำลองรุ่นที่ 2 คิดเป็นร้อยละ 24.91 เนื่องจากแบบจำลองที่นำเสนอมีการปรับปรุง 3 ประเด็นหลักคือ ความสูง จากระดับน้ำทะเลของพื้นที่ที่สนใจ สัมประสิทธิ์ความขุ่นมัวของชั้นบรรยากาศ และดัชนีเมฆรอบจุดที่สนใจ

4.2 แผนการดำเนินงาน

การดำเนินงานในขั้นถัดไป คือ ทดลองและปรับปรุงแบบจำลองในกลุ่ม Machine learning เช่น Support vector machine (SVR) เตรียมชุดคำสั่งเพื่อใช้งานแบบจำลอง สรุปผลและอภิปรายและเตรียมนำเสนอโครงการ

ตารางที่ 10: Gantt chart

ขั้นตอนการดำเนินงาน	วิชา 2102490					วิชา 2102499			
	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.	เม.ย.
1.ศึกษาภาพรวมของวิธีการประมาณความเข้มแสงอาทิตย์ และแบบจำลองการลดทอนความเข้มแสงด้วยเมฆ									
2.ทำ pre-process data เช่น การเก็บข้อมูลภาพถ่ายการคัดเลือกภาพการจัดการ outlier และคัดเลือกข้อมูลที่ผิดปกติออก									
3.วิเคราะห์ ค่า ความคลาดเคลื่อน ของ แบบจำลองที่ใช้งานปัจจุบัน									
4.สร้าง แบบ จำลอง Linear regression, Polynomial regression และ Random forest ปรับ พารามิเตอร์ เพื่อ เปรียบ เทียบ ประสิทธิภาพและเลือกแบบจำลองที่ดีที่สุด									
5.วิเคราะห์เปรียบเทียบแบบจำลองที่ใช้งานอยู่ และแบบจำลองที่นำเสนอ									
6.จัดทำรูปเล่มและเตรียมนำเสนอโครงร่างรายงาน									
7.ทดลอง และ ปรับปรุง แบบ จำลอง ใน กลุ่ม Machine learning และเลือกแบบจำลอง									
8.เตรียมชุดคำสั่งเพื่อใช้งานแบบจำลอง									
9.สรุปผลและอภิปรายและเตรียมนำเสนอโครงการ									

หมายเหตุ

- คือ ความก้าวหน้าที่วางแผนไว้
- คือ ความก้าวหน้าปัจจุบัน

5 เอกสารอ้างอิง

- [1] J. Antonanzas et al. “Review of photovoltaic power forecasting”. In: *Solar energy* 136 (2016), pp. 78–111.
- [2] K.F. Dagestad and J.A. Olseth. “An alternative algorithm for calculating the cloud index”. In: *Heliosat-3 Projekt-Bericht* (2005).
- [3] P. Ineichen and R. Perez. “A new airmass independent formulation for the Linke turbidity coefficient”. In: *Solar Energy* 73.3 (2002), pp. 151–157.
- [4] R.H. Inman, H.T.C Pedro, and C.F.M Coimbra. “Solar forecasting methods for renewable energy integration”. In: *Progress in energy and combustion science* 39.6 (2013), pp. 535–576.
- [5] S.D. Miller, A.K. Heidinger, and M. Sengupta. “Physically based satellite methods”. In: *Solar Energy Forecasting and Resource Assessment*. Ed. by J. Kleissl. 2013, pp. 49–80.
- [6] A. Nisbet. *Open topo data api*. 2021. URL: <https://www.opentopodata.org/>.
- [7] M. Noia, C.F. Ratto, and R. Festa. “Solar irradiance estimation from geostationary satellite data: I. Statistical models”. In: *Solar Energy* 51.6 (1993), pp. 449–456.
- [8] M. Paulescu et al. *Weather modeling and forecasting of PV systems operation*. Vol. 358. Springer, 2013.
- [9] R. Perez, T. Cebeauer, and M. Šúri. “Chapter 2 - Semi-Empirical Satellite Models”. In: *Solar Energy Forecasting and Resource Assessment*. Ed. by J. Kleissl. Boston: Academic Press, 2013, pp. 21–48.
- [10] R. Perez, T. Hoff, and J. Kleissl. “SolarAnywhere forecasting”. In: *Solar Energy Forecasting and Resource Assessment*. Ed. by J. Kleissl. Academic Press, 2013, pp. 233–265.
- [11] S. Pranonsatid, P. Manomaisaowapak, and J. Songsiri. *Thailand Solar Map by CUEE*. Tech. rep. Department of Electrical Engineering, Chulalongkorn University, 2022.
- [12] C. Raphael and J.E. Hay. “An assessment of models which use satellite data to estimate solar irradiance at the earth’s surface”. In: *Journal of Applied Meteorology and Climatology* 23.5 (1984), pp. 832–844.
- [13] *SolarAnywhere*. <https://www.solaranywhere.com/>. Accessed: 22-Novemeber-2022.
- [14] *Solargis*. <https://solargis.com/>. Accessed: 22-Novemeber-2022.
- [15] T. Srivastava. “Tuning the parameters of your Random Forest model”. In: *Analytics Vidhya* 9 (2015).
- [16] M. Šúri and T. Cebeauer. “SolarGIS: New web-based service offering solar radiation data and PV simulation tools for Europe, North Africa and Middle East”. In: *Proceedings of the International Conference on Solar Heating, Cooling and Buildings EUROSUN*. Vol. 28. 2010.
- [17] Inc. The Gale Group. *Atmospheric Turbidity The Great Soviet Encyclopedia, 3rd Edition*. URL: <https://encyclopedia2.thefreedictionary.com/Atmospheric+Turbidity> (visited on 11/17/2022).