

ข้อเสนอโครงการวิศวกรรมไฟฟ้า วิชา 2102490

การพยากรณ์กำลังไฟฟ้าแสงอาทิตย์ภายในวัน
ด้วยการใช้ภาพถ่ายเมฆจากดาวเทียมฮิมาวาริ

Intra-day solar power forecasting using cloud images
from Himawari satellite

นายณฐนนท โตงามรักษ์ เลขประจำตัวนิสิต 6232007121

นายณัฐพล เดชตระกูลวงศ์ เลขประจำตัวนิสิต 6232011621

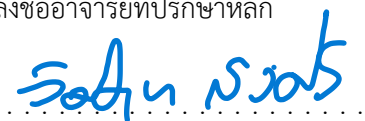
อาจารย์ที่ปรึกษา รศ.ดร. จิตโกมุท ส่งศิริ

ภาควิชาวิศวกรรมไฟฟ้า คณะวิศวกรรมศาสตร์

จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2565

ลงชื่ออาจารย์ที่ปรึกษาหลัก



(รศ.ดร. จิตโกมุท ส่งศิริ)

วันที่ 25 พ.ย. 2565

สารบัญ

1	บทนำ	3
1.1	ที่มาและความสำคัญของโครงการ	3
1.2	วัตถุประสงค์ของโครงการ	4
1.3	ขอบเขตของโครงการ	4
1.4	ผลลัพธ์ที่คาดหวังจากโครงการ	4
2	หลักการและทฤษฎีที่เกี่ยวข้อง	5
2.1	การพยากรณ์ความเข้มแสงอาทิตย์	5
2.2	แบบจำลองท้องฟ้าใส	6
2.3	เวกเตอร์ความเร็วเมฆ	6
2.3.1	การเทียบบล็อก	7
2.3.2	การไหลของแสง	8
2.4	การสกัดปริมาณเมฆที่ส่งผลกับตำแหน่งที่สนใจ	9
2.5	การคัดเลือกเซตย่อยที่ดีที่สุด	12
2.6	แบบจำลองการเรียนรู้ด้วยเครื่อง	12
2.6.1	แบบจำลองการถดถอยแบบเชิงเส้น	12
2.6.2	แบบจำลอง Support Vector Regression	13
2.6.3	แบบจำลองแบบป่าสุ่ม	13
2.7	สมรรถนะของแบบจำลอง	14
2.8	รายละเอียดชุดข้อมูล	14
3	ผลลัพธ์การดำเนินการเบื้องต้น	14
3.1	ผลการคัดเลือกเซตย่อยที่ดีที่สุด	15
3.2	ผลลัพธ์การพยากรณ์ความเข้มแสงอาทิตย์	16
3.3	ผลลัพธ์การวิเคราะห์การเคลื่อนที่ของเมฆ	17
4	บทสรุป	19
4.1	สรุปผลการดำเนินการ	19
4.2	แผนการดำเนินงาน	19
	เอกสารอ้างอิง	20

1 บทนำ

1.1 ที่มาและความสำคัญของโครงการ

ปัจจุบันพบว่ามีส่วนการใช้พลังงานทดแทนมากขึ้นในระบบโครงข่ายไฟฟ้าของประเทศต่าง ๆ ทั่วโลก พลังงานแสงอาทิตย์เป็นหนึ่งในพลังงานที่เริ่มเข้ามามีบทบาทอันเนื่องจากแนวโน้มต้นทุนการติดตั้งที่ต่ำลง อย่างไรก็ตาม เมื่อมีการรวมเข้ากันกับระบบโครงข่ายไฟฟ้าของพลังงานแสงอาทิตย์ ต้องคำนึงถึงสมดุลระหว่างกำลังไฟฟ้าที่ใช้และกำลังไฟฟ้าที่ผลิตได้ในระบบ พลังงานแสงอาทิตย์นั้นมีความไม่แน่นอนเนื่องมาจากความแปรปรวนของสภาพอากาศเมื่อนำมารวมเข้าไปในระบบโครงข่ายอาจส่งผลกระทบต่อเสถียรภาพของระบบ ดังนั้นจึงมีบริษัทต่าง ๆ ทั่วโลกที่มีบริการหรือผลิตภัณฑ์สำหรับการประมาณปริมาณแสงอาทิตย์ในช่วงเวลาต่าง ๆ เช่นบริษัท Reunivatt มีบริการตรวจวัดความเข้มแสงอาทิตย์ที่มีชื่อว่า SunSat™ โดยไม่ต้องติดตั้งตัวรับรู้ (sensor) บนภาคพื้น แต่จะใช้ภาพถ่ายจากดาวเทียมทางอุตุนิยมวิทยาที่มีวงโคจรแบบค้างฟ้า (geostationary satellite) ทั้ง 5 ดวง ได้แก่ Meteosat-10, Meteosat-8, GOES-E, GOES-W และ Himawari-8 ดาวเทียมทั้ง 5 สามารถให้บริการภาพถ่ายได้ครอบคลุมพื้นที่ทั่วโลก รวมถึงยังมีบริการพยากรณ์ความเข้มแสงอาทิตย์โดยใช้ข้อมูลทางอุตุนิยมวิทยาร่วมกับภาพถ่ายดาวเทียมในแต่ละระยะการพยากรณ์ล่วงหน้า หรือ บริษัท Solar Anywhere ที่ใช้ข้อมูลพยากรณ์จากแบบจำลองทำนายอากาศเชิงเลข (Numerical weather prediction หรือ NWP) ร่วมกับข้อมูลภาพถ่ายดาวเทียมสำหรับการสกัดเวกเตอร์ความเร็วเมฆ (Cloud motion vectors หรือ CMV) ช่วยในการพยากรณ์ความเข้มแสงอาทิตย์ในระดับชั่วโมง ซึ่งการพยากรณ์ในระยะดังกล่าวสามารถช่วยในการควบคุมการทำงานของระบบโครงข่ายไฟฟ้าและงานด้าน load-following ได้

ดังนั้นจึงได้มีงานวิจัยมากมายที่มีจุดมุ่งหมายในการพยากรณ์ความเข้มแสงอาทิตย์ เช่น มีการให้เหตุผลว่า นอกเหนือจากประสิทธิภาพการพยากรณ์ที่คุณภาพใกล้เคียงกับการใช้ข้อมูลที่วัดได้จากตัวรับรู้ที่ภาคพื้นแล้ว การใช้ข้อมูลจากภาพถ่ายดาวเทียมที่ผ่านการตรวจแก้ความคลาดเคลื่อน (bias-correction) จะช่วยกำจัดความคลาดเคลื่อนไม่เพียงแต่ที่ตำแหน่งจุดภาพที่มีตัวรับรู้ที่ภาคพื้นแต่จะรวมถึงตำแหน่งจุดภาพโดยรอบนั้นด้วย นั่นคือจะสามารถลดต้นทุนด้านการติดตั้งเครื่องมือวัดเพิ่มในบางพื้นที่ได้ [1] ในช่วงแรกที่เริ่มมีการนำข้อมูลภาพถ่ายดาวเทียมมาใช้ในการทดลอง โดยวิธีการเทียบบล็อก (block-matching) เป็นวิธีในการวิเคราะห์การเคลื่อนที่ของเมฆในแต่ละจุดภาพจากการใช้ภาพถ่ายดาวเทียมที่มีลำดับติดกันในการสร้างเวกเตอร์การเคลื่อนที่ของเมฆ ด้วยหลักการหาทิศทางที่ทำให้ผลรวมความต่างของความเข้มของจุดภาพโดยรอบของภาพในลำดับก่อนหน้ากับความเข้มโดยรอบในภาพลำดับถัดไปมีค่าต่ำสุดจากนั้นทำการประมาณค่านอกช่วง (extrapolate) เวกเตอร์ที่ได้ในการพยากรณ์ภาพถ่ายดาวเทียมที่ระยะพยากรณ์ล่วงหน้าต่าง ๆ จากนั้นใช้แบบจำลอง Heliosat ในการแปลงค่าดัชนีเมฆในแต่ละจุดภาพเป็นค่าความเข้มแสงอาทิตย์ที่ต้องการ [2] วิธีการเทียบบล็อกสำหรับการหา CMV นี้ยังปรากฏในอีกหลายงานวิจัยเช่น เปรียบเทียบให้เห็นว่าการใช้ความเข้มแสงอาทิตย์ที่ได้มาจากการใช้ CMV ให้ผลลัพธ์ที่ดีกว่าการใช้ค่าความเข้มแสงอาทิตย์จาก NWP ในการใช้เป็นตัวแปรต้นสำหรับแบบจำลอง Support vector regression (SVR) ในการพยากรณ์ค่ากำลังไฟฟ้าที่ผลิตได้จากเซลล์แสงอาทิตย์ แต่ถึงอย่างนั้นได้มีการทดลองที่แสดงให้เห็นว่าประสิทธิภาพการพยากรณ์โดยใช้เวกเตอร์ดังกล่าวจะขึ้นกับสภาพภูมิอากาศในแต่ละพื้นที่ด้วย [3]

อีกเทคนิคของการวิเคราะห์การเคลื่อนที่ของเมฆที่ได้รับความนิยมเป็นเทคนิคที่ชื่อว่า การไหลของแสง (Optical flow) ซึ่งเป็นเทคนิคจากการประมวลผลภาพที่ใช้สำหรับติดตามการเคลื่อนที่ของวัตถุในภาพหรือวิดีโอ และสามารถใช้ในการหา CMV ของภาพถ่ายดาวเทียมได้ดีอีกด้วย เมื่อเปรียบเทียบการใช้การไหลของแสงที่มีเงื่อนไขบังคับที่ต่างกัน 4 เงื่อนไขได้แก่ Horn และ Schunck [4], Lucas-Kanade [5], Farneback [6] และ TVL1 [7] รวมไปถึงวิธีการเทียบบล็อก [2] จากการวิเคราะห์พบว่าการใช้การไหลของแสงที่ใช้เงื่อนไข TVL1 ให้ผลลัพธ์การพยากรณ์ที่ดีที่สุดสำหรับการพยากรณ์ในระดับชั่วโมง [8] หรือมีงานวิจัยที่ได้มีการปรับปรุงสมการการไหลของแสง โดยคำนึงถึงความต่อเนื่องของความสว่างภายในภาพผ่านการใส่ฟังก์ชันแกร่ง (robust function) [9] อีกทั้งยังมีงานวิจัยที่ประยุกต์ใช้การประมาณ CMV ด้วยเทคนิคการไหลของแสง วิธี Horn-Schunck [4] ในงาน

พยากรณ์ความเข้มแสงอาทิตย์เชิงความน่าจะเป็น (Probabilistic solar forecasting) อีกด้วย [10]

นอกเหนือจากการวิเคราะห์การเคลื่อนที่ของเมฆที่จะส่งผลต่อความเข้มแสงอาทิตย์ ณ ตำแหน่งที่สนใจด้วยการคำนวณ CMV แล้ว ยังได้มีการใช้แบบจำลองการเรียนรู้เชิงลึก (Deep learning) ที่มีข้อมูลนำเข้าเป็นลำดับของภาพถ่ายดาวเทียมที่ติดกันเช่น การใช้ข้อมูลค่าประมาณความเข้มแสงอาทิตย์ในอดีตร่วมและความเข้มแสงอาทิตย์ในสภาวะท้องฟ้าใส (Clear-sky irradiance) ร่วมกับภาพถ่ายดาวเทียมเป็นข้อมูลนำเข้าของแบบจำลองโครงข่ายประสาทเทียมแบบสังวัตนาการ (Convolutional Neural network หรือ CNN) ในการพยากรณ์ความเข้มแสงอาทิตย์ระดับชั่วโมง [11] การใช้ข้อมูลภาพถ่ายท้องฟ้าที่เป็นลำดับ ในการพยากรณ์ความเข้มแสงอาทิตย์ภายในชั่วโมง (Intra-hour) โดยประยุกต์ใช้เทคนิค CNN โดยเทียบระหว่าง CNN แบบ 2 มิติ กับ CNN แบบ 3 มิติ พบว่า CNN แบบ 2 มิติ ให้ประสิทธิภาพการพยากรณ์และต้นทุนการคำนวณที่ดีกว่า [12] หรือ การสกัดคุณลักษณะโดยเริ่มต้นจากสมมติฐานของสมการการไหลของแสง ผ่านวิธีเรียนรู้เชิงกราฟ สำหรับเป็นข้อมูลนำเข้าในแบบจำลองการเรียนรู้เชิงลึกร่วมกับข้อมูลกำลังไฟฟ้าที่ผลิตได้ในอดีตและความเข้มแสงอาทิตย์ในสภาวะท้องฟ้าใส ในการพัฒนาแบบจำลองพยากรณ์กำลังไฟฟ้าภายในชั่วโมงเป็นต้น [13]

จากข้อมูลที่ได้ศึกษามาในข้างต้น จะเห็นได้ว่าการสกัดเวกเตอร์ความเร็วเมฆจากภาพถ่ายดาวเทียมด้วยเทคนิคต่าง ๆ หรือการวิเคราะห์การเคลื่อนที่ของเมฆโดยใช้กลุ่มแบบจำลองการเรียนรู้เชิงลึก เป็นวิธีที่ได้รับความนิยมอย่างแพร่หลายในงานพยากรณ์ความเข้มแสงอาทิตย์ในระดับชั่วโมง และด้วยเหตุผลความแตกต่างของปัจจัยด้านภูมิอากาศและพื้นที่ โครงการนี้จึงสนใจจะเปรียบเทียบเทคนิคการวิเคราะห์ความเร็วเมฆรวมถึงการสกัดคุณลักษณะจากภาพถ่ายดาวเทียม ในการพัฒนาการพยากรณ์ความเข้มแสงอาทิตย์ โดยใช้เทคนิคทาง ML ได้แก่ แบบจำลองการถดถอยแบบเชิงเส้น (Linear regression) แบบจำลอง Support vector regression และแบบจำลองแบบป่าสุ่ม (Random forests) รวมถึงแบบจำลองการเรียนรู้เชิงลึก (CNN) เพื่อเสนอตัวแปรทำนายและแบบจำลองที่ให้ผลลัพธ์การพยากรณ์ในระดับชั่วโมงที่ดีที่สุด

1.2 วัตถุประสงค์ของโครงการ

1. วิเคราะห์ข้อมูลเมฆจากภาพถ่ายดาวเทียมสำหรับสกัดข้อมูลเป็นตัวแปรขาเข้าในการพยากรณ์ความเข้มแสงอาทิตย์ในระดับชั่วโมง
2. พัฒนาและเปรียบเทียบประสิทธิภาพของแบบจำลองการพยากรณ์ความเข้มแสงอาทิตย์ที่ใช้คุณลักษณะที่สกัดจากข้อมูลภาพถ่ายดาวเทียมจากวิธีต่าง ๆ

1.3 ขอบเขตของโครงการ

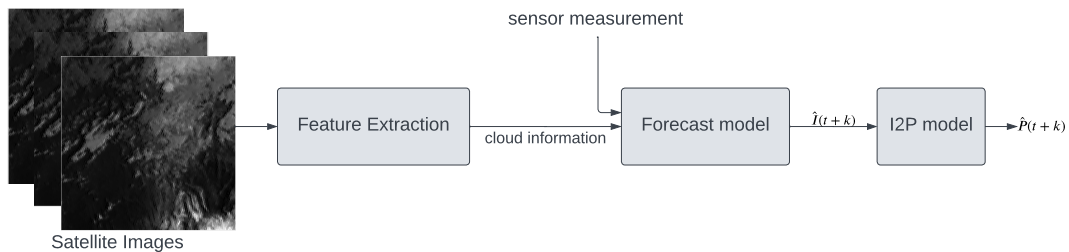
1. การพยากรณ์แบบแหล่งที่ตั้งเดียวที่ใช้ข้อมูลกำลังไฟฟ้าที่ผลิตได้ ความเข้มแสงอาทิตย์ และอุณหภูมิ จากตัวรับรู้บนแหล่งที่ตั้งลำดับที่ 48 และข้อมูลภาพถ่ายดาวเทียมจากดาวเทียมฮิมาวาริ-8
2. แบบจำลองการพยากรณ์ความเข้มแสงอาทิตย์ระยะตั้งแต่ 30, 60, ..., 240 นาทีล่วงหน้า จากข้อมูลภาพถ่ายดาวเทียม และข้อมูลจากตัวรับรู้ โดยจะพยากรณ์ตั้งแต่ช่วงเวลา 7.00 น. ถึง 17.00 น.

1.4 ผลลัพธ์ที่คาดหวังจากโครงการ

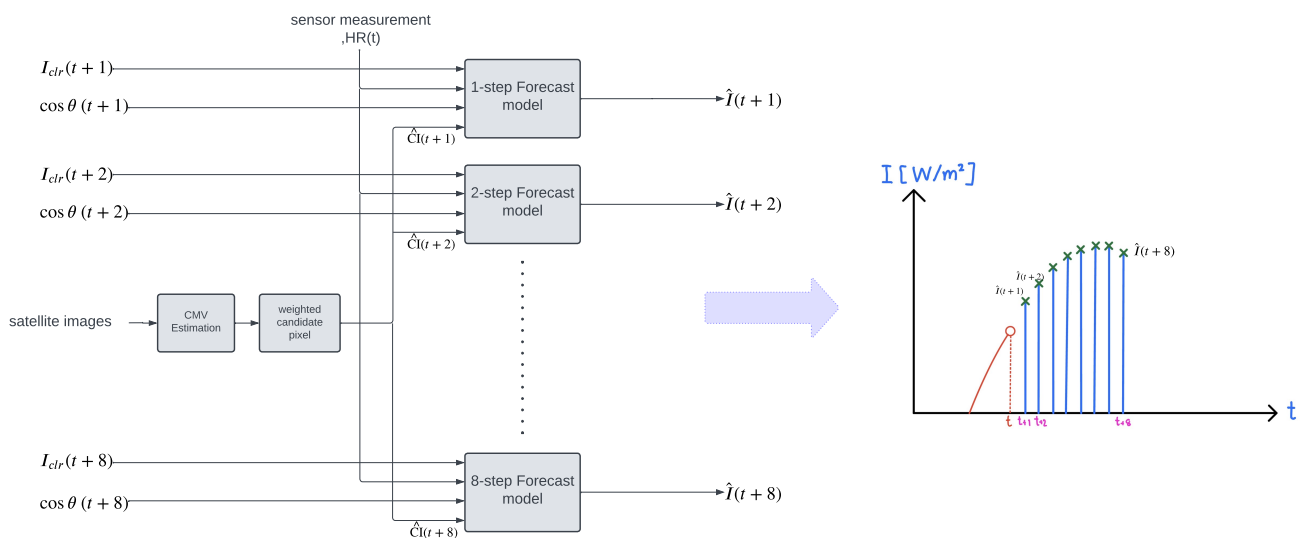
1. ผลการวิเคราะห์ความสำคัญของข้อมูลปริมาณเมฆที่ส่งผลต่อสมรรถนะการพยากรณ์ความเข้มแสงอาทิตย์ในระดับชั่วโมง
2. ชุดคำสั่งแบบจำลองพยากรณ์ค่าความเข้มแสงอาทิตย์ด้วยแบบจำลองที่เสนอไว้ข้างต้น

2 หลักการและทฤษฎีที่เกี่ยวข้อง

ในการพัฒนาแบบจำลองการพยากรณ์นั้นจะมีขั้นตอนหลักดังรูปที่ 1 นั่นคือการสกัดคุณลักษณะจากภาพถ่ายดาวเทียมและการฝึกแบบจำลองการเรียนรู้ด้วยเครื่อง จากนั้นจะพยากรณ์ออกมาเป็นค่าความเข้มแสงอาทิตย์ด้วยตัวแปรต้นจากคุณลักษณะที่สกัดได้ร่วมกับจากข้อมูลตัวรับรู้และข้อมูลจากแบบจำลองท้องฟ้าใสดังรูปที่ 2 จากนั้นจึงใช้แบบจำลองในการแปลงความเข้มแสงอาทิตย์เป็นค่ากำลังไฟฟ้าที่ผลิตได้



รูปที่ 1: ภาพรวมขั้นตอนในการทำโครงการ



รูปที่ 2: การพยากรณ์ความเข้มแสงอาทิตย์ในแต่ละระยะการพยากรณ์

การพัฒนาแบบจำลองการพยากรณ์ที่มีตัวแปรต้น x และตัวแปรตาม y ที่เป็นตัวแปรที่ขึ้นกับเวลา t จะมีข้อกำหนดทางสัญลักษณ์ตัวแปรเพิ่มเติมคือ

- เขียน $x(t+n)$ เมื่อ $n = 1, 2, \dots, 8$ สำหรับค่าของตัวแปร x ในเวลา 30, 60, ..., 240 นาทีข้างหน้าตามลำดับ
- เขียน $x(t-p)$ เมื่อ $p = 1, 2, \dots, 8$ สำหรับค่าของตัวแปร x ในเวลา 30, 60, ..., 240 นาทีย้อนหลังตามลำดับ

โดยข้อกำหนดข้างต้นเป็นข้อกำหนดที่ใช้กับทั้งตัวแปรต้นและตัวแปรตาม ตัวแปรที่เกี่ยวข้องกับโครงการนี้จะปรากฏดัง ตารางที่ 1

2.1 การพยากรณ์ความเข้มแสงอาทิตย์

การพยากรณ์ (Forecasting) คือ การคำนวณหรือการประมาณค่าใด ๆ ในอนาคต โดยสามารถใช้ตัวแปรต้น X เป็นข้อมูลในอดีต (past data) หรือข้อมูลปัจจุบันในการประมาณค่าตัวแปรตาม ณ ตำแหน่งเวลาในอนาคต

ตัวแปร	ความหมาย	หน่วย
P	กำลังผลิตไฟฟ้า	กิโลวัตต์
I	ความเข้มแสงอาทิตย์	วัตต์ต่อตารางเมตร
T_{ambient}	อุณหภูมิภายนอก	องศาเซลเซียส
T_{module}	อุณหภูมิของแผงเซลล์แสงอาทิตย์	องศาเซลเซียส
I_{pix}	ความเข้มของแต่ละจุดภาพในภาพ	มีค่าตั้งแต่ 0-255
CI	ค่าที่บ่งบอกถึงปริมาณเมฆในจุดภาพ	มีค่าตั้งแต่ 0-1

ตารางที่ 1: ความหมายและสัญลักษณ์ของตัวแปรที่ใช้

ระยะเวลาพยากรณ์ล่วงหน้า (Forecasting horizon) คือ ระยะเวลาที่แบบจำลองสามารถพยากรณ์ค่าตัวแปรตามออกมาได้ ในส่วนของการพยากรณ์ความเข้มแสงอาทิตย์นั้นจะแบ่งออกเป็น 4 ช่วงเวลา [14]

1. การพยากรณ์ในระยะไม่เกิน 1 ชั่วโมง (Intra-hour หรือ nowcasting) เป็นการพยากรณ์ในระยะเวลาที่สั้นมาก มีจุดประสงค์เพื่อความมีเสถียรภาพ (stability) ของโครงข่าย และการตอบสนองด้านโหลด (demand response)
2. การพยากรณ์ในระดับชั่วโมง (Intra-day) เป็นการพยากรณ์ในระดับ 1 ถึง 6 ชั่วโมง มีความสำคัญสำหรับการควบคุมการทำงานของระบบโครงข่ายไฟฟ้า (grid operation) หรืองานด้าน load-following
3. การพยากรณ์ในระดับ 6 ชั่วโมงถึง 1 วัน (Six hours to day ahead) เป็นการพยากรณ์ในระดับ 6 ถึง 48 ชั่วโมง มีความสำคัญสำหรับการวางแผนหรือเดินเครื่องการผลิตภายในโรงไฟฟ้า (unit commitment)
4. การพยากรณ์ในระยะ 2 วันขึ้นไป (Two days ahead) เป็นการพยากรณ์ในระดับ 48 ชั่วโมงขึ้นไป มีความสำคัญต่อการวางแผนเดินเครื่องการผลิตภายในโรงไฟฟ้า หรือการวางแผนบำรุงรักษาโครงข่าย โดยในการพยากรณ์ประเภทนี้ยังมีจำนวนงานวิจัยที่เกี่ยวข้องไม่มากนัก

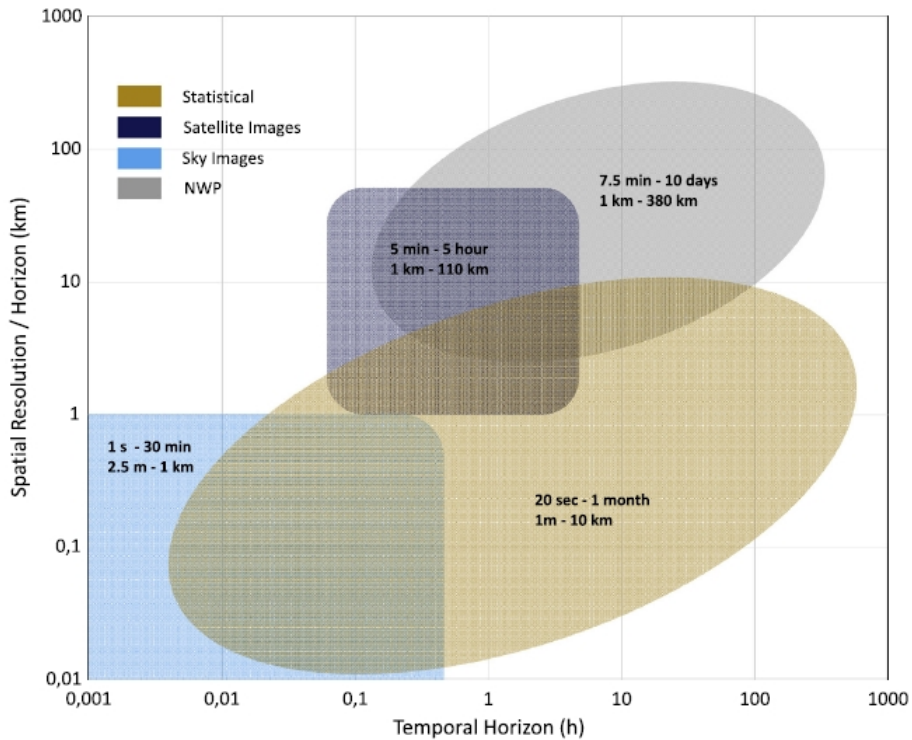
จากงานวิจัยในอดีตสามารถสรุปประเภทข้อมูลนำเข้าหรือตัวแปรต้นที่นิยมสำหรับแต่ละระยะเวลาพยากรณ์ล่วงหน้าได้ดัง รูปที่ 3 โดยจะเห็นได้ว่าประเภทของข้อมูลที่เป็นที่นิยมสำหรับระยะเวลาพยากรณ์ในระดับชั่วโมงจะเป็นข้อมูลภาพถ่ายดาวเทียม (Satellite Image) และข้อมูลพยากรณ์จากแบบจำลองทำนายอากาศเชิงเลข (NWP) แต่ในการทดลองนี้ไม่ได้ถูกนำมาใช้เนื่องด้วยประสิทธิภาพของข้อมูลและความละเอียดเชิงเวลาที่ต่ำกว่า ซึ่งถ้านำมาเพิ่มอัตราสุ่มจะยิ่งทำให้สมรรถนะของแบบจำลองต่ำลง โครงการนี้จึงใช้เพียงข้อมูลภาพถ่ายดาวเทียมสำหรับการพยากรณ์ในระดับชั่วโมง

2.2 แบบจำลองท้องฟ้าใส

แบบจำลองท้องฟ้าใส (Clear-sky model) คือ แบบจำลองที่ทำนายค่าความเข้มแสงในสภาวะฟ้าใสออกมาได้ โดยในทางอุดมคติความเข้มแสงในสภาวะฟ้าใสนั้นเป็นค่าสูงสุดที่เป็นไปได้ของความเข้มแสงอาทิตย์ ณ เวลาต่างๆ ซึ่งขึ้นอยู่กับลักษณะทางกายภาพของตำแหน่งที่สนใจ เช่น มุมเซนนิธ (Solar-zenith angle) ความสูงจากระดับน้ำทะเล องค์ประกอบของไอน้ำและโอโซนภายในอากาศ เป็นต้น เราจึงนำค่าความเข้มแสงในสภาวะฟ้าใสมาเป็นคุณลักษณะหนึ่งของแบบจำลอง ในโครงการนี้ใช้แบบจำลองท้องฟ้าใสที่ถูกพัฒนาโดย P. Ineichen และ R. Perez [15] ที่มีตัวแปรเข้าที่สำคัญคือ ละติจูด ลองจิจูด ความสูงจากระดับน้ำทะเล และ ลำดับวันในปี (day of year)

2.3 เวกเตอร์ความเร็วเมฆ

เวกเตอร์ความเร็วเมฆ (CMV) คือเวกเตอร์แสดงความเร็วของเมฆแต่ละส่วนในภาพ การที่เรารู้ค่า CMV จะบอกได้ว่าภายใต้อะไรเวลาที่เราสงเกตนั้น เมฆจะเคลื่อนที่ไปอยู่ในตำแหน่งใด ถ้าหากเรารู้ว่าในระยะเวลาที่เราสงเกตนั้น



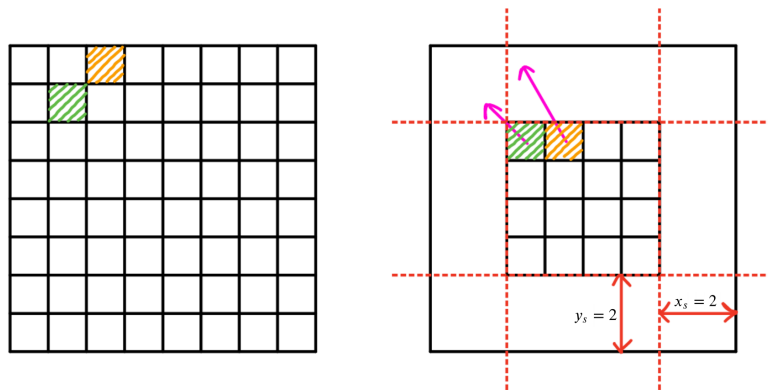
รูปที่ 3: ประเภทข้อมูลนำเข้าตามระยะการพยากรณ์ล่วงหน้าและขอบเขตเชิงพื้นที่ ที่มาภาพจาก [14]

จะมีเมฆบริเวณใดเคลื่อนที่มายังตำแหน่งที่เราสนใจได้ เราก็จะสามารถพยากรณ์ได้ว่าเวลาที่เราน่าสนใจนั้นมีแนวโน้มที่ความเข้มแสงอาทิตย์จะถูกลดทอนโดยเมฆมากน้อยเพียงใด CMV จึงมักถูกนำมาเป็นคุณลักษณะในแบบจำลองการพยากรณ์ในระดับชั่วโมง สำหรับการสกัดหา CMV นั้นมีอยู่ 2 วิธี ได้แก่ วิธีการเทียบบล็อกและวิธีการไหลของแสง

2.3.1 การเทียบบล็อก

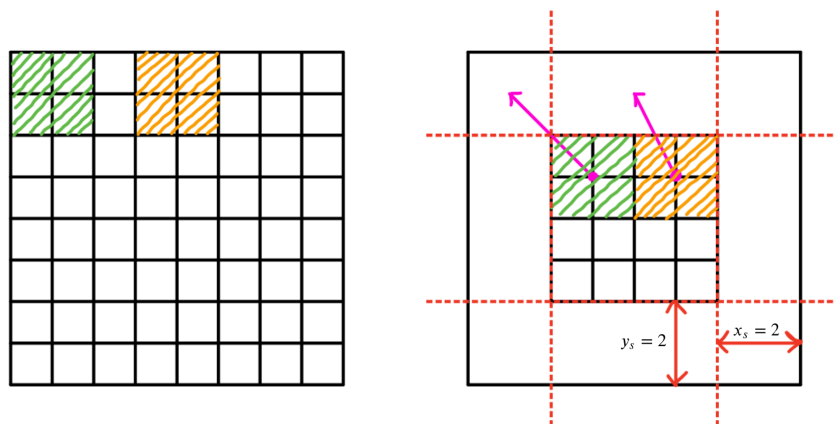
วิธีการเทียบบล็อกนั้นจะมีข้อมูลเข้าเป็นภาพจากดาวเทียม โดยจะคำนวณหา CMV ผ่าน cross-correlation method (CCM) ผ่านภาพดาวเทียม 2 ภาพที่นำเข้า โดย CCM จะบอกถึงตำแหน่งที่ จุดภาพ ณ เวลา $t - 1$ จะเคลื่อนที่ไปยังตำแหน่งใด ๆ ณ เวลา t ซึ่งตัวแปรที่จะใช้มี x, y คือระยะทางในแกน X และ Y ของโดเมนในภาพทั้งหมด และ x_s, y_s คือระยะค้นหา (search radius) ในระยะแกน X และ Y ตามลำดับ โดยกระบวนการที่กล่าวมาทั้งหมดสามารถทำได้ 2 วิธี [16] คือ

1. ค้นหาทั้งโดเมน จะทำการค้นหาทั้งโดเมน แต่เพื่อที่จะทำให้การค้นหาไม่ตกขอบของโดเมน เราจึงสร้างกล่องที่มีขนาด $(x - 2x_s, y - 2y_s)$ ในการค้นหา โดยความเร็วที่ได้จะได้อาจมาจากคู่จุดภาพที่มีค่า cross-correlation coefficient สูงที่สุด



รูปที่ 4: การเทียบบล็อกโดยค้นหาทั้งโดเมน

- ค้นหาโดเมนย่อย โดยแบ่งโดเมนทั้งหมดเป็นกล่องขนาดเล็ก เช่น กล่องขนาด 5x5 จุดภาพ โดยยิ่งแบ่งกล่องขนาดเล็กเท่าใดจะส่งผลให้ CMV ที่ได้มีความละเอียดมากขึ้น



รูปที่ 5: การเทียบบล็อกโดยค้นหาโดเมนย่อยด้วยกล่องขนาด 2×2

2.3.2 การไหลของแสง

การไหลของแสง (Optical flow) คือ ความเร็วจากการเคลื่อนที่ของความเข้ม (intensity) ที่ปรากฏ (apparent) จากภาพสองภาพที่ต่อเนื่องกัน เทคนิคการไหลของแสงในโครงการนี้จึงมีไว้เพื่อหาความเร็วของแต่ละจุดภาพ หรือก็คือเมฆ โดยมีสมมุติฐานดังนี้

- ค่าความเข้มในแต่ละจุดภาพของสองภาพที่ต่อเนื่องกันมีค่าเท่ากัน (uniform)
- เมฆไม่มีการก่อตัว/สลายตัว และไม่มีการแผ่ขยาย

- การเคลื่อนที่ของเมฆจะถูกจำลองเหมือนการเคลื่อนที่ของวัตถุแข็งเกร็ง (motion of rigid body)

เราสามารถเขียนสมมติฐานนี้เป็นความสัมพันธ์ทางคณิตศาสตร์ได้ดัง (1)

$$I_{\text{pix}}(x + dx, y + dy, t + dt) = I_{\text{pix}}(x, y, t) \quad (1)$$

โดยที่ I_{pix} คือค่าความเข้ม (Intensity) ในแต่ละจุดภาพโดย x และ y คือค่าระยะทางเชิงพื้นที่ และ t คือเวลา เมื่อกระจายอนุกรมเทเลอร์อันดับหนึ่งและหาอนุพันธ์เทียบเวลาของ (1)

$$\frac{\partial I_{\text{pix}}}{\partial x} v_x + \frac{\partial I_{\text{pix}}}{\partial y} v_y + \frac{\partial I_{\text{pix}}}{\partial t} \approx 0 \quad (2)$$

เนื่องจาก (2) มีตัวแปรที่ไม่ทราบค่าอยู่ 2 ตัวแปร $[v_x, v_y]$ ทำให้มีคำตอบที่ไม่เป็นหนึ่งเดียว (unique) จึงมีการเพิ่มปัญหาค่าเหมาะสมสุดต่าง ๆ ที่คิดค้นขึ้นมาเพื่อแก้ข้อผิดพลาดนี้อีกหนึ่งปัญหา เช่น Horn-Schunck, Lucas-Kanade, Farneback และ TVL-1 เป็นต้น ในโครงงานนี้จะปัญหาค่าเหมาะสมสุดของ Horn-Schunck เพราะสามารถทำการทดลองได้ง่าย มี toolbox ที่ใช้งานได้เลยในโปรแกรม MATLAB

วิธี Horn-Schunck เป็นการหาความเร็วเมฆ v_x, v_y ที่สอดคล้องกับ (2) และยังทำให้ค่าเหมาะสมสุดของ (3) เกิดขึ้นด้วย [4]

$$\arg \min_{v_x, v_y} \int_{\text{pic}} (\nabla_x v_x)^2 + (\nabla_y v_y)^2 + \lambda \cdot \left(\frac{\partial I_{\text{pix}}}{\partial x} \cdot v_x + \frac{\partial I_{\text{pix}}}{\partial y} \cdot v_y + \frac{\partial I_{\text{pix}}}{\partial t} \right)^2 \quad (3)$$

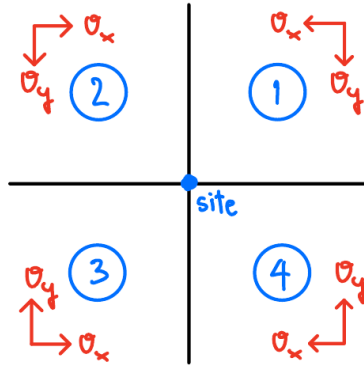
โดยที่ λ คือพารามิเตอร์ถ่วงน้ำหนัก

2.4 การสกัดปริมาณเมฆที่ส่งผลกับตำแหน่งที่สนใจ

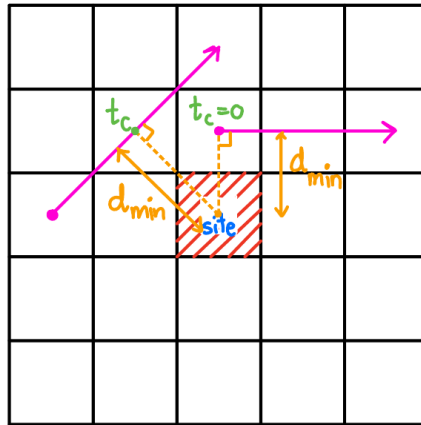
หลังจากที่คำนวณ CMV แล้ว ขั้นแรกจะกำหนดขอบเขตที่เราคัดเลือกจุดภาพที่ส่งผลต่อแบบจำลอง (candidate pixel) จากนั้นสำหรับทุกจุดภาพในขอบเขตที่เลือกที่มี CMV ในแกน X และ แกน Y เป็น v_x, v_y ตามลำดับ ทำการกำหนดฟังก์ชันในตัวแปรเวลา t ของระยะห่างระหว่างจุดภาพที่กำลังคัดเลือกกับจุดภาพที่ตำแหน่งที่ตั้งที่สนใจดัง (4) [10]

$$d(t) = \sqrt{[(x - x_{\text{site}}) + tv_x]^2 + [(y - y_{\text{site}}) + tv_y]^2} \quad (4)$$

เมื่อพิจารณาเงื่อนไขด้านตำแหน่งและความเร็วที่ส่งผลให้จุดภาพเข้าใกล้ตำแหน่งที่ตั้งที่สนใจมากขึ้นเรื่อย ๆ $[\frac{d}{dt}(t) < 0]$ จะได้ว่าถ้าจุดภาพมีตำแหน่งและความเร็วเริ่มต้นดัง รูปที่ 6 แล้วจุดภาพจะเข้าใกล้ตำแหน่งที่สนใจมากขึ้นเรื่อย ๆ เสมอ



รูปที่ 6: ตำแหน่งและความเร็วเริ่มต้นที่ทำให้จุดภาพเข้าใกล้ตำแหน่งที่ตั้งที่สนใจโดยแบ่งเป็นจตุภาค



รูปที่ 7: ตัวอย่างค่า t_c ทั้งสองกรณี

จากเงื่อนไขจำเป็นของอนุพันธ์อันดับหนึ่งในการเป็นจุดต่ำสุด จะได้ว่าเวลาที่แต่ละจุดภาพจะเข้าใกล้ตำแหน่งที่ตั้งที่สนใจมากที่สุดเป็น

$$t_c = -\frac{(x - x_{\text{site}})v_x + (y - y_{\text{site}})v_y}{v_x^2 + v_y^2} \quad (5)$$

ในกรณีที่จุดภาพไม่ได้มีความเร็วที่สอดคล้องกับตำแหน่งเริ่มต้นดัง รูปที่ 6 จะมีโอกาสที่จุดภาพเคลื่อนที่ห่างจากตำแหน่งที่สนใจมากขึ้นเรื่อย ๆ หรือก็คือ $t_c = 0$ แต่ถ้าคำนวณ t_c จาก(5) จะได้ $t_c \leq 0$ เราจึงปรับปรุง (5) ใหม่ให้อยู่ในรูปฟังก์ชัน ReLU ได้

$$t_c = \max\left(0, -\frac{(x - x_{\text{site}})v_x + (y - y_{\text{site}})v_y}{v_x^2 + v_y^2}\right) \quad (6)$$

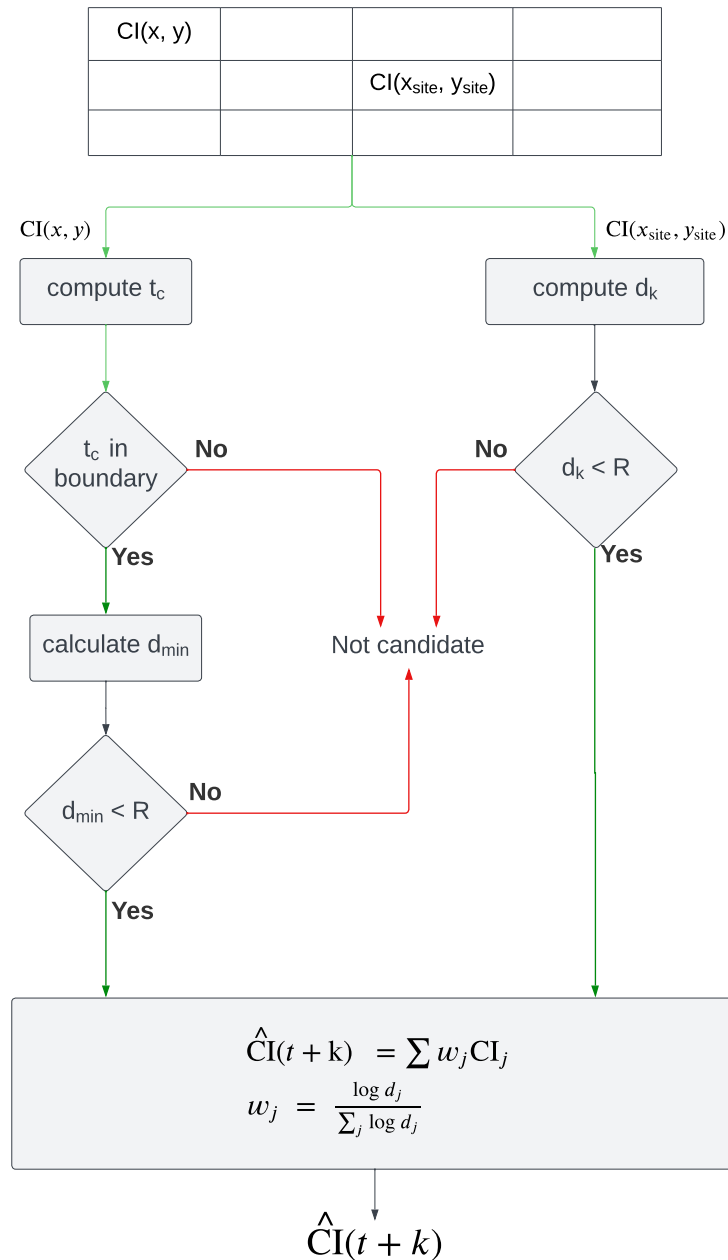
เกณฑ์การคัดเลือก candidate จุดภาพ นั่นคือจุดภาพนั้นจะต้องมีเวลา t_c ไม่เกินค่าขีดแบ่ง (threshold) T_{upper} ค่าหนึ่งแต่มีค่าเกินขอบเขตล่างทางเวลา T_{lower} ด้วย ที่เวลา t_c นั้นจะต้องมี $d_{\text{min}} \triangleq d(t_c)$ ไม่เกินระยะ R โดยค่าขีดแบ่งทั้งสองค่าจะถูกปรับให้เหมาะสมสำหรับระยะเวลาการพยากรณ์ล่วงหน้าที่แตกต่างกัน หลังจากคัดเลือกตัวแทนจุดภาพจากจุดภาพทั้งหมดในขอบเขตที่กำหนดไว้ นำค่าดัชนีเมฆ ของจุดภาพที่ถูกเลือกมาหาค่าเฉลี่ย และกำหนดเป็นตัวแปร CI ที่บ่งบอกถึงปริมาณเมฆ ณ ตำแหน่งที่สนใจในอนาคต

การนำเสนอผลรวมดัชนีเมฆ ในโครงการนี้จะนำเสนอวิธีการรวมดัชนีเมฆจากจุดภาพที่ผ่านการคัดเลือกโดยการหาค่าเฉลี่ยถ่วงน้ำหนักที่มีสมมติฐานว่ายังมีปริมาณเมฆเข้าใกล้ตำแหน่งที่สนใจมาก ก็ยิ่งส่งผลต่อการลดทอน

ความเข้มแสงอาทิตย์มากขึ้นตามไปด้วย ซึ่งจะพิจารณาทั้งตำแหน่งจุดภาพที่ทำการทดลองและตำแหน่งโดยรอบ โดยมีแผนผังขั้นตอนการประมาณค่าดัชนีเมฆ ที่เวลา k ชั้นล่องหนาดังรูปที่ 8

$$\hat{CI}(t+k) = \sum_{j \in \text{candidate}} w_j CI_j \text{ เมื่อ } k = 1, 2, \dots, 8$$

โดยที่ $w_j = \frac{\log d_j}{\sum_j \log d_j}$ (7)



รูปที่ 8: กระบวนการนำมาซึ่งค่าดัชนีเมฆที่ส่งผลต่อตำแหน่งที่สนใจในเวลา k ชั้นล่องหน้า

2.5 การคัดเลือกเซตย่อยที่ดีที่สุด

สำหรับการสร้างแบบจำลองการถดถอยเชิงเส้น กรณีตัวทำนาย X ที่มีจำนวน p ตัวแปร เราต้องการที่จะทราบว่าเซตของกลุ่มตัวแปรทำนายเซตใดที่จะทำให้แบบจำลองที่ได้มีผลลัพธ์ที่ดีที่สุดจากทั้งหมด 2^p เซตความเป็นไปได้ โดยแบบจำลองทำการเรียนรู้ด้วยข้อมูลชุดฝึกสอน (training data) และประเมินสมรรถนะบนชุดข้อมูลตรวจสอบ (validation data) ซึ่งมีขั้นตอนการเลือกเซตย่อยที่ดีที่สุดโดยสรุปดังนี้ [17]

1. กำหนดแบบจำลอง M_0 ที่เป็นแบบจำลองว่างที่ไม่ใช้ตัวแปรทำนายใดเลย นั่นคือแบบจำลองจะให้ผลการทำนายเป็นค่าเฉลี่ยของกลุ่มตัวอย่างบนชุดฝึกสอน
2. ดำเนินการเพิ่มตัวแปรทำนายตั้งแต่ $k = 1, 2, \dots, p$ สำหรับแบบจำลองที่มี k ตัวแปร เลือกหนึ่งเซตตัวแปรทำนายจากทั้งหมด $\binom{p}{k}$ เซต ที่ทำให้สมรรถนะของแบบจำลองที่ดีที่สุดซึ่งสรุปได้จากผลรวมความคลาดเคลื่อนกำลังสอง Residual Sum Square (RSS) มีค่าต่ำสุด หรือค่า R^2 มีค่าสูงสุดบนชุดข้อมูลตรวจสอบ กำหนดเป็นแบบจำลอง M_k
3. ในแต่ละแบบจำลอง M_0, M_1, \dots, M_p เลือกแบบจำลองที่จะใช้งานผ่านเกณฑ์การคัดเลือกแบบจำลอง (model selection criterion) นั่นคือ เกณฑ์สารสนเทศของอะกะอิเกะ Akaike Information (AIC) และ เกณฑ์สารสนเทศของเบส์ Bayesian Information Criterion (BIC) โดยเลือกจากแบบจำลองที่มีค่าเกณฑ์การคัดเลือกต่ำที่สุด

ภายใต้สมมติฐานว่าแบบจำลองถดถอยเชิงเส้นนั้นถูกรบกวนโดยสัญญาณรบกวนที่มีการแจกแจงแบบเกาส์เซียน เราจะสามารถคำนวณค่าเกณฑ์สารสนเทศทั้งสองได้จาก

$$AIC_{\text{scaled}} = \log(\text{MSE}) + 2 \frac{d}{N} \quad (8)$$

และ

$$BIC_{\text{scaled}} = \log(\text{MSE}) + 2 \frac{\log(N)}{N} \quad (9)$$

เมื่อ N เป็นจำนวนข้อมูลของชุดตรวจสอบและ d เป็นจำนวนตัวแปรทำนายที่ใช้ในแบบจำลองที่กำลังทดสอบ

2.6 แบบจำลองการเรียนรู้ด้วยเครื่อง

2.6.1 แบบจำลองการถดถอยแบบเชิงเส้น

เนื่องจากแบบจำลองการถดถอยแบบเชิงเส้นนั้นใช้ทรัพยากรในการคำนวณที่น้อยเมื่อเทียบกับแบบจำลองอื่น ๆ จึงมีงานวิจัยที่ได้ทำการทดลองในการพยากรณ์แบบทันที โดยเปรียบเทียบการใช้แบบจำลองการถดถอยในตัวที่มีสัญญาณรบกวนระบบ (ARX model) ที่มีตัวแปรทำนาย คือค่ากำลังไฟฟ้าที่ถูกทำให้เป็นกลาง (normalize) โดยแบบจำลองฟ้าใส (clear sky model) เพื่อที่ว่าตัวแปรทำนายจะมีลักษณะนิ่ง (stationary) มากขึ้น และสัญญาณรบกวนระบบคือ NWP กับแบบจำลองการถดถอยในตัว (AR model) ที่มีตัวแปรทำนายคือค่ากำลังไฟฟ้า จะถูกทำให้เป็นกลาง พบว่าแบบจำลอง ARX มีสมรรถนะดีขึ้นในระยะการพยากรณ์ไม่เกิน 2 ชั่วโมงอย่างมีนัยสำคัญ [18]

แบบจำลองการถดถอยแบบเชิงเส้น คือ แบบจำลองการเรียนรู้เชิงสถิติที่ใช้ในการวิเคราะห์เชิงถดถอยที่มีข้อมูลเข้าและข้อมูลออกมีความสัมพันธ์แบบเชิงเส้นเท่านั้น ซึ่งสามารถอธิบายได้ดัง (10)

$$y = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1n} \\ X_{21} & X_{22} & \cdots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m1} & X_{m2} & \cdots & X_{mn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} \triangleq y = X\beta \quad (10)$$

โดยที่

- $y \in \mathbb{R}^m$ คือข้อมูลออก
- $X \in \mathbb{R}^{m \times n}$ คือตัวแปรทำนาย
- $\beta \in \mathbb{R}^n$ คือสัมประสิทธิ์การถดถอยที่ต้องการประมาณ

โดยเราจะพัฒนาแบบจำลองการถดถอยแบบเชิงเส้น วัตถุประสงค์เพื่อใช้เป็นแบบจำลองฐานสำหรับทุกกระบวนการพยากรณ์ล่วงหน้า

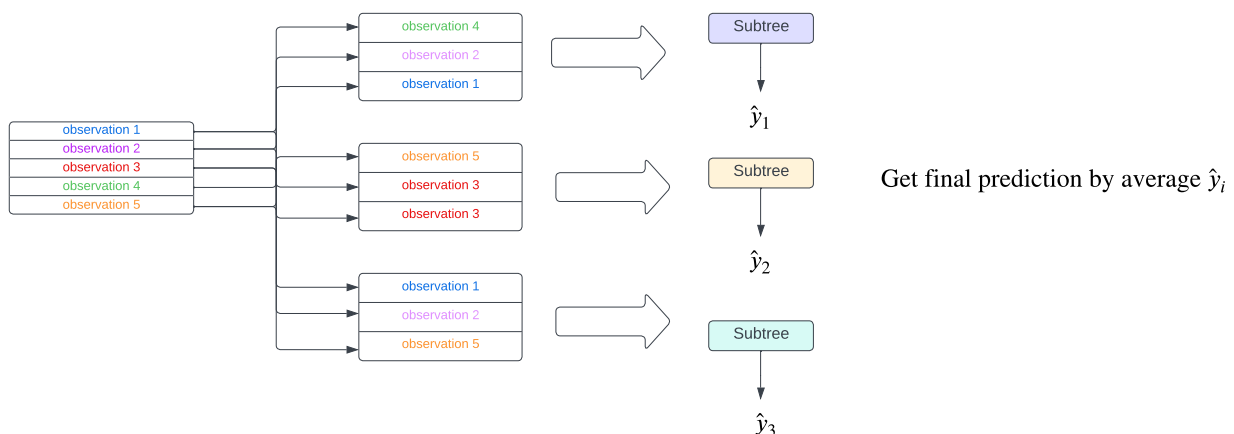
2.6.2 แบบจำลอง Support Vector Regression

มีการใช้ SVR เพื่อพยากรณ์กำลังไฟฟ้าในระดับ 1 ชั่วโมง โดยตัวแปรทำนายที่ถูกหวนวง ได้แก่ ค่ากำลังไฟฟ้าที่ผลิตได้ ความเข้มแสงอาทิตย์ อุณหภูมิ ความชื้น และความเร็วลม จะถูกคัดเลือกโดยอัลกอริทึมเลือกคุณลักษณะอิงสหสัมพันธ์ (Correlation-based Feature Selection Algorithm: CFS Algorithm) จากนั้นจึงถูกป้อนเข้าสู่แบบจำลอง โดยกระบวนการดังกล่าวพบว่าเมื่อตัวแปรทำนายมีเพียงแค่ค่ากำลังไฟฟ้าที่ถูกหวนวง แบบจำลองสามารถพยากรณ์ได้อย่างแม่นยำโดยไม่ต้องนำคุณลักษณะอื่น ๆ เช่น ค่าความเข้มแสงอาทิตย์ และ NWP มาใช้เพิ่ม [19]

2.6.3 แบบจำลองแบบป่าสุ่ม

แบบจำลองแบบป่าสุ่ม เป็นแบบจำลองที่พัฒนามาจากต้นไม้ตัดสินใจ (Decision tree) เพื่อเพิ่มประสิทธิภาพและลดปัญหาความแปรปรวนของการทำนายที่สูงจากการใช้ชุดข้อมูลฝึกทั้งหมดในการฝึกแบบจำลองหลักเพียงแบบจำลองเดียว (high-variance) โดยมีขั้นตอนคือเริ่มจากการสร้างชุดข้อมูล bootstrap ซึ่งชุดข้อมูลดังกล่าวจะเป็นการสุ่มเลือกมาจากชุดข้อมูลเดิมและสามารถเลือกตัวอย่างซ้ำได้ต่อชุดข้อมูล bootstrap 1 ชุดซึ่งเราจะทำการสุ่มสร้างมาทั้ง B ชุดข้อมูลสำหรับแบบจำลองต้นไม้ย่อย (subtree) B ต้น ในแต่ละต้นไม้ย่อยนั้นเราจะสุ่มใช้บางตัวแปรทำนาย (Predictor) เท่านั้น โดยเราจะได้ค่าทำนายสุดท้ายจากการเฉลี่ยผลลัพธ์จากต้นไม้ย่อยแต่ละต้นดัง (11) เราจะเรียกวิธีการได้ผลลัพธ์สุดท้ายจากหลาย ๆ ผลการทำนายของแบบจำลองย่อยนี้ว่าเป็นการใช้เทคนิคแบบ ensemble [17]

$$\hat{y}_{\text{avg}} = \frac{1}{B} \sum_{i=1}^B \hat{y}_i \quad (11)$$



รูปที่ 9: แผนผังขั้นตอนของแบบจำลองแบบป่าสุ่มโดยในที่นี้ยกตัวอย่างการใช้ชุดข้อมูล bootstrap B = 3 ชุด

2.7 สมรรถนะของแบบจำลอง

สำหรับโครงการนี้จะใช้ 3 ดัชนีในการประเมินความแม่นยำการพยากรณ์ของแบบจำลอง ได้แก่

- ผลบวกกำลังสองของส่วนเหลือ (Residual sum of squares: RSS)

$$RSS = \sum_{t=1}^N (\hat{X}_t - X_t)^2$$

- รากที่สองของค่าเฉลี่ยความผิดพลาดกำลังสอง (Root Mean Square Error: RMSE)

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (\hat{X}_t - X_t)^2}$$

- ค่าเฉลี่ยของความผิดพลาดสัมบูรณ์ (Mean Absolute Error: MAE)

$$MAE = \frac{1}{N} \sum_{t=1}^N |\hat{X}_t - X_t|$$

2.8 รายละเอียดชุดข้อมูล

ข้อมูลที่ถูกใช้ในการสร้างแบบจำลองการพยากรณ์มาจาก 2 แหล่งด้วยกัน ได้แก่

- ข้อมูลภาพถ่ายดาวเทียม ใช้เป็นข้อมูลที่เป็นภาพถ่ายจากดาวเทียมฮิมาวาริ-8 ซึ่งเป็นดาวเทียมค้างฟ้า โดยข้อมูลที่ได้รับจะเป็นไฟล์ภาพสกุล (.TIFF) ขนาด 1725 × 1670 จุดภาพ ในแต่ละจุดภาพมีค่าความหนาแน่นของเมฆอยู่ระหว่าง 0 ถึง 255 และขนาดเชิงพื้นที่เท่ากับ 2 × 2 ตารางกิโลเมตร โดยข้อมูลภาพถ่ายดาวเทียมที่ได้มานี้จะมีสถานีรับภาพอยู่ที่คณะวิศวกรรมศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย ให้บริการตั้งแต่วันที่ 5 กุมภาพันธ์ 2565 เป็นต้นมา ด้วยความละเอียดเชิงเวลาเท่ากับ 10 นาที แต่จะมีข้อจำกัดในด้านเวลา คือ แต่ละข้อมูลภาพถ่ายดาวเทียมที่ได้รับจะมีความหน่วงทางเวลา 30 นาทีนับตั้งแต่เวลาจริงที่ภาพนั้นถูกถ่าย

พารามิเตอร์	ค่า
ความละเอียดเชิงเวลา	10 นาที
ขนาดพื้นที่ต่อ 1 พิกเซล	2 × 2 ตารางกิโลเมตร
ขนาดของภาพถ่ายทางดาวเทียม	1725 × 1670 พิกเซล
ตำแหน่งภายในภาพของไซด์ที่ 48 ซึ่งเป็นตำแหน่งที่ตั้งหลักที่ทำการทดลอง	แถวที่ 864, หลักที่ 847

ตารางที่ 2: ข้อมูลภาพถ่ายโดยดาวเทียมฮิมาวาริ-8

- ข้อมูลจากตัวรับรู้ของบริษัท Impact solar โดยจะเป็นค่ากำลังไฟฟ้าที่ผลิตได้, ค่าความเข้มแสงอาทิตย์ และค่าอุณหภูมิ ข้อมูลที่ได้รับจะอยู่ในรูปแบบอนุกรมเวลาจากแผงเซลล์แสงอาทิตย์ของไซด์ที่ 48 โดยมีความละเอียดเชิงเวลาเท่ากับ 15 นาทีซึ่งมีตำแหน่งที่ตั้งอยู่ที่เขตปทุมวัน กรุงเทพมหานคร และมีความจุติดตั้งของกำลังไฟฟ้าที่ผลิตได้ (installed capacity) อยู่ที่ 252 กิโลวัตต์

การปรับปรุงชุดข้อมูลเบื้องต้น เนื่องจากข้อมูลทั้งสองมีความละเอียดเชิงเวลาที่ไม่เท่ากัน โครงการนี้จะทำการรวมข้อมูลทั้งสองแหล่งเฉพาะตำแหน่งเวลานั้นคือทุก ๆ 30 นาที

3 ผลลัพธ์การดำเนินการเบื้องต้น

สำหรับผลการดำเนินงานในภาคการศึกษาต้นปีนี้เราได้ทำการเลือกเซตย่อยที่ดีที่สุดในการพัฒนาแบบจำลองฐาน โดยได้ทำการแบ่งชุดข้อมูลสำหรับการพัฒนาแบบจำลองดังนี้

1. ชุดข้อมูลฝึก เป็นชุดข้อมูลที่ใช้เฉพาะในขั้นตอนการคัดเลือกตัวแปรต้นในแบบจำลอง โดยในภาคการศึกษาต้นนี้ใช้ข้อมูลของ วันที่ 5 กุมภาพันธ์ พ.ศ.2565 - 15 กรกฎาคม พ.ศ.2565 สำหรับข้อมูลจากทั้งตัวรับรู้และข้อมูลภาพถ่ายดาวเทียม
2. ชุดข้อมูลตรวจสอบ เป็นชุดข้อมูลที่ไม่ถูกใช้ในการฝึกแบบจำลองแต่ใช้สำหรับการวัดผลของแต่ละแบบจำลองย่อยในขั้นตอนการคัดเลือกเท่านั้น โดยจะใช้ข้อมูลตั้งแต่วันที่ 16 กรกฎาคม พ.ศ.2565 - 31 สิงหาคม พ.ศ.2565
3. ชุดข้อมูลทดสอบ เป็นชุดข้อมูลที่ใช้ในการวัดประสิทธิภาพการพยากรณ์ของแบบจำลองที่ทำการคัดเลือกตัวแปรเสร็จสิ้นแล้ว โดยทั่วไปแล้วจะเลือกใช้เป็นส่วนท้ายสุดจากชุดข้อมูลที่มี ในที่นี้ใช้ข้อมูลตั้งแต่วันที่ 1 กันยายน พ.ศ.2565 - 30 ตุลาคม พ.ศ.2565

สำหรับการแบ่งข้อมูลข้างต้นนี้จะใช้สำหรับภาคการศึกษาต้นเท่านั้น สำหรับภาคการศึกษาปลายเมื่อจำนวนข้อมูลจากตัวรับรู้และ ภาพถ่ายดาวเทียมมีจำนวนมากขึ้นจะทำการแบ่งชุดข้อมูลในการพัฒนาแบบจำลองอีกครั้งให้มีสัดส่วนที่เหมาะสม จากนั้นจะเป็นผลการดำเนินงานในส่วนการวิเคราะห์การเคลื่อนที่ของเมฆเบื้องต้นเพื่อใช้สำหรับการพัฒนาแบบจำลองในขั้นตอนต่อไป

3.1 ผลการคัดเลือกเซตย่อยที่ดีที่สุด

เมื่อดำเนินการใช้เทคนิคการหาเซตย่อยที่ดีที่สุดกับตัวแปรต้นที่พิจารณานั้นคือ $I(t - 6)$, $I(t - 5)$, ..., $I(t - 1)$, $I(t)$, $I_{clr}(t + 1)$, $\cos(\theta(t + 1))$, $T(t)$ สำหรับแบบจำลองถดถอยเชิงเส้นที่มีตัวแปรขาออกเป็น $I(t + 1)$ ได้ผลลัพธ์การดำเนินการดัง ตารางที่ 3 สรุปได้ว่าเซตของตัวแปรที่ให้ค่าเกณฑ์สารสนเทศทั้งสองต่ำสุดบนชุด

k	RSS($\times 10^5$)	R^2	คุณลักษณะ	AIC _{scaled}	BIC _{scaled}
1	226.1	0.5671	$I(t)$	10.4187	10.4841
2	197.2	0.6224	$I(t)$, $\cos(\theta(t + 1))$	10.2820	10.3474
3	193.5	0.6295	$I(t - 5)$, $I(t)$, $\cos(\theta(t + 1))$	10.2629	10.3283
4	191.8	0.6328	$I(t - 5)$, $I(t - 1)$, $I(t)$, $\cos(\theta(t + 1))$	10.2541	10.3195
5	190.8	0.6346	$I(t - 5)$, $I(t - 1)$, $I(t)$, $I_{clr}(t + 1)$, $\cos(\theta(t + 1))$	10.2491	10.3145
6	190.4	0.6354	$I(t - 5)$, $I(t - 3)$, $I(t - 1)$, $I(t)$, $I_{clr}(t + 1)$, $\cos(\theta(t + 1))$	10.2471	10.3124
7	190.3	0.6354	$I(t - 5)$, $I(t - 4)$, $I(t - 3)$, $I(t - 1)$, $I(t)$, $I_{clr}(t + 1)$, $\cos(\theta(t + 1))$	10.2469	10.3122
8	190.4	0.6354	$I(t - 6)$, $I(t - 5)$, $I(t - 4)$, $I(t - 3)$, $I(t - 1)$, $I(t)$, $I_{clr}(t + 1)$, $\cos(\theta(t + 1))$	10.2470	10.3123
9	190.4	0.6353	$I(t - 6)$, $I(t - 5)$, $I(t - 4)$, $I(t - 3)$, $I(t - 2)$, $I(t - 1)$, $I(t)$, $I_{clr}(t + 1)$, $\cos(\theta(t + 1))$	10.2472	10.3126
10	192.1	0.6321	$I(t - 6)$, $I(t - 5)$, $I(t - 4)$, $I(t - 3)$, $I(t - 2)$, $I(t - 1)$, $I(t)$, $I_{clr}(t + 1)$, $\cos(\theta(t + 1))$, $T(t)$	10.2559	10.3213

ตารางที่ 3: สรุปเซตของคุณลักษณะที่ดีที่สุดในแต่ละจำนวนตัวแปร

ข้อมูลตรวจสอบเกิดขึ้นเมื่อใช้จำนวนตัวแปร $k = 7$ โดยมีเซตย่อยที่ดีที่สุดคือ

$$I(t - 5), I(t - 4), I(t - 3), I(t - 1), I(t), I_{clr}(t + 1), \cos(\theta(t + 1)) \quad (12)$$

พบว่าตัวแปรที่ถูกเลือกจากการคัดเลือกเซตย่อยที่ดีที่สุดคือ ค่าความเข้มแสงอาทิตย์ที่เวลาปัจจุบัน และ เวลาในอดีตบางเวลา, ค่าความเข้มแสงอาทิตย์ในสภาวะท้องฟ้าใสกับค่าโคไซน์ของมุมเซนิตที่เวลา 1 ชั่วโมงหน้า โดย

ตัวแปรที่ถูกเลือกเข้าแบบจำลองท้ายสุดคือตัวแปร อุณหภูมิที่เวลาปัจจุบัน เราจะใช้เซตของคุณลักษณะดังกล่าวเป็นตัวแปรต้นสำหรับแบบจำลองฐานสำหรับทุกระยะการพยากรณ์ล่วงหน้าแต่จะทำการเปลี่ยนตัวแปร ความเข้มแสงอาทิตย์ในสภาวะท้องฟ้าใสกับค่าโคไซน์มุมของดวงอาทิตย์เทียบกับแนวตั้งฉากพื้นโลก เป็นค่าที่ตำแหน่งเวลา $t + n$ สำหรับการพยากรณ์ล่วงหน้า n ชั้น

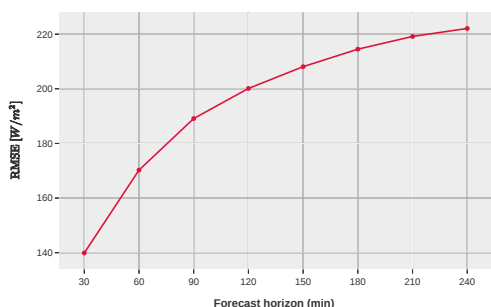
3.2 ผลลัพธ์การพยากรณ์ความเข้มแสงอาทิตย์

ผลการพยากรณ์ของแบบจำลองฐาน เราได้สร้างแบบจำลองฐานจากแบบจำลองการถดถอยแบบเชิงเส้น โดยใช้ตัวแปรต้นและตัวแปรตามในแต่ละระยะการพยากรณ์ล่วงหน้าที่ได้โดยดังนี้

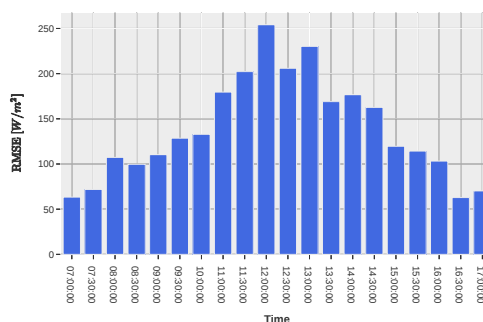
แบบจำลองการถดถอยเชิงเส้น (Linear regression) สำหรับการพยากรณ์ความเข้มแสงอาทิตย์ที่เวลา $t + n$ จะมีตัวแปรที่เกี่ยวข้องคือ

- ตัวแปรต้น x : $I(t - 5), I(t - 4), I(t - 3), I(t - 1), I(t), I_{clr}(t + n), \cos(\theta(t + n)), HR(t)$
- ตัวแปรตาม y : $\hat{I}(t + n)$

สำหรับตัวแปร $HR(t)$ เป็นตัวแปรที่บอกเวลาที่เป็นชั่วโมง ณ ตำแหน่งเวลาปัจจุบัน โดยมีผลลัพธ์การพยากรณ์ดังรูปที่ 10 ซึ่งประสิทธิภาพของแบบจำลองฐานที่ใช้เฉพาะข้อมูลตัวรับรู้ที่ภาคพื้นนี้จะเป็นตัวอภิปรายผลลัพธ์ของแบบจำลองที่กำลังจะถูกพัฒนาขึ้นซึ่งจะใช้ข้อมูลจากภาพถ่ายดาวเทียมเพิ่มเติม ว่ามีสมรรถนะการพยากรณ์ที่แตกต่างจากเดิมอย่างไร



(a) RMSE ในแต่ละระยะการพยากรณ์



(b) RMSE ของแต่ละตำแหน่งเวลาในการพยากรณ์ $I(t+1)$

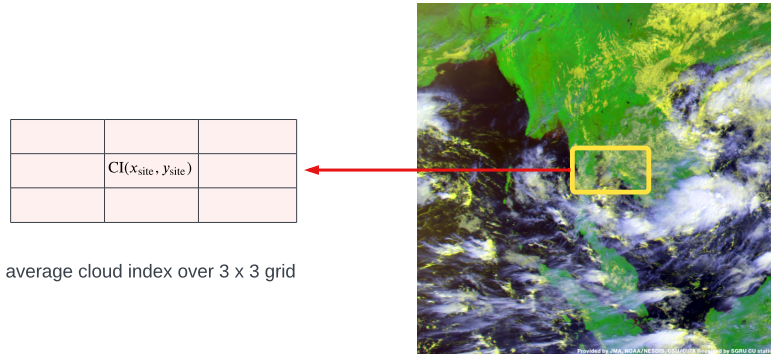
รูปที่ 10: ประสิทธิภาพการพยากรณ์ของแบบจำลองถดถอยเชิงเส้นซึ่งเป็นแบบจำลองฐานโดยใช้ตัวแปรต้นที่กำหนดไว้ข้างต้น

3.3 ผลลัพธ์การวิเคราะห์การเคลื่อนที่ของเมฆ

จากการศึกษาวิจัยที่เกี่ยวข้องกับการวิเคราะห์การเคลื่อนที่ของเมฆ ซึ่งเบื้องต้นในโครงการนี้เราได้ทำการทดลอง 2 วิธีคือ ใช้ข้อมูลภาพถ่ายดาวเทียมวิเคราะห์ความสัมพันธ์ระหว่างความเข้มแสงอาทิตย์ ณ ตำแหน่งที่ทำการทดลองกับ ปริมาณเมฆเฉลี่ยของจุดภาพโดยรอบที่เวลาเดียวกันผ่านค่าเฉลี่ยดัชนีเมฆดังรูปที่ 11 อธิบายโดย (13)

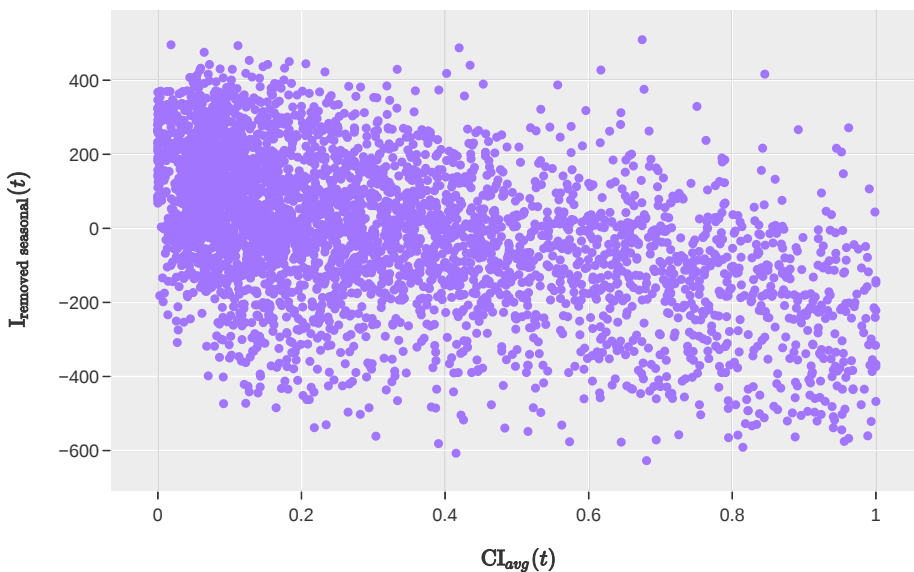
$$CI_{avg}(t) = \frac{1}{9} \sum_{(x_i, y_i) \in w} CI(x_i, y_i, t) \quad (13)$$

เมื่อ $w = \{(x, y) \mid x \in [x_{site} - 1, x_{site} + 1], y \in [y_{site} - 1, y_{site} + 1]\}$



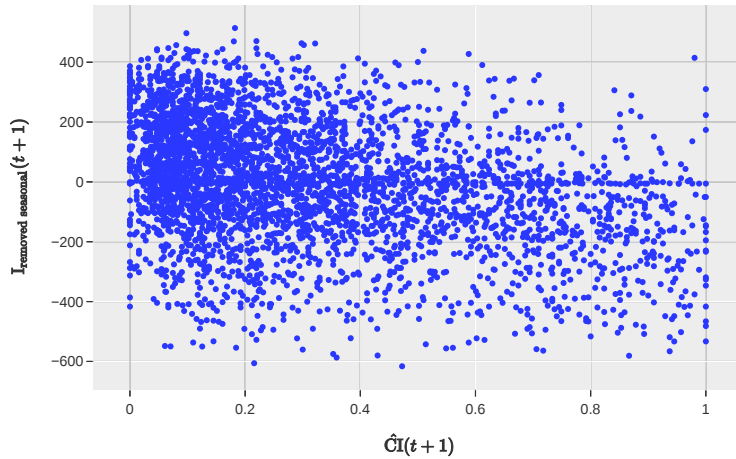
รูปที่ 11: การเฉลี่ยดัชนีเมฆรอบจุดภาพที่ทำการทดลอง

พบว่าความสัมพันธ์ระหว่างทั้งสองตัวแปรนั้นมีความใกล้เคียงลักษณะเชิงผกผันแต่มีความฟุ้งกระจายอยู่ กล่าวคือในค่าดัชนีเมฆที่เท่ากันหรือใกล้เคียงกัน จะมีช่วงค่าความเข้มแสงอาทิตย์ที่เป็นไปได้หลายค่าดังรูปที่ 12

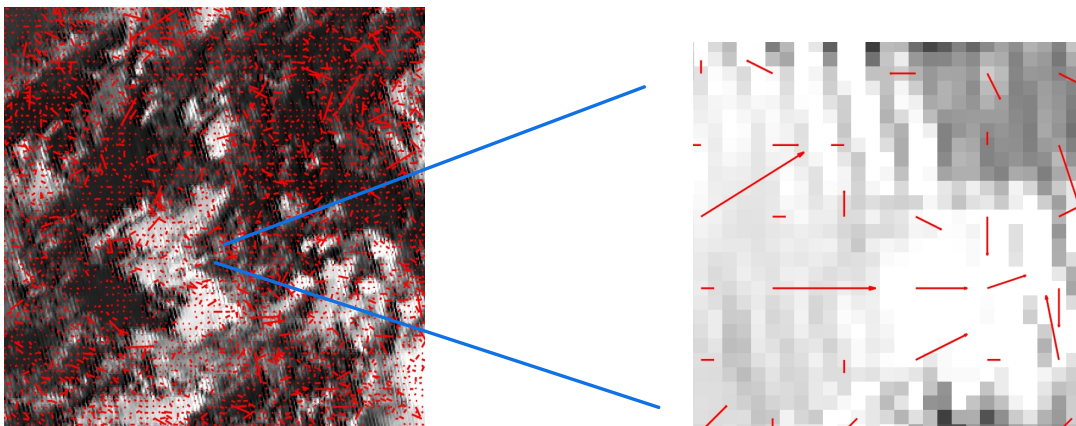


รูปที่ 12: ความสัมพันธ์ระหว่างความเข้มแสงอาทิตย์และค่าเฉลี่ยดัชนีเมฆรอบจุดที่สนใจ

จากนั้นพิจารณาผลลัพธ์เทียบกับค่าเฉลี่ยดัชนีเมฆ CI แต่ในเบื้องต้นนี้จะเป็นผลลัพธ์จากการถ่วงน้ำหนักที่เท่ากันของแต่ละจุดภาพที่ผ่านเกณฑ์การคัดเลือกจากหัวข้อที่ 2.4 มีตัวอย่างผลการสกัดเวกเตอร์ดังรูปที่ 14 โดยแสดงความสัมพันธ์เทียบกับค่าความเข้มแสงอาทิตย์ที่เวลา 30 นาทีถัดไปดังรูปที่ 13 พบว่ากราฟความสัมพันธ์มีลักษณะคล้ายรูปที่ 12 แต่มีความฟุ้งกระจายมากกว่า ในส่วนของทั้งสองกราฟที่เรามองเห็นความผกผันนั้นเป็นผลจากการที่ว่ากรณีที่ดัชนีเมฆมีค่าสูงสามารถนำไปสู่ข้อสรุปได้ว่ามีหรือกำลังจะมีปริมาณเมฆมาบดบัง ณ ตำแหน่งที่สนใจนั้นคือจะส่งผลให้ค่าความเข้มแสงอาทิตย์มีค่าต่ำนั่นเอง



รูปที่ 13: ความสัมพันธ์ระหว่างความเข้มแสงอาทิตย์และ CI ในอีก 30 นาทีข้างหน้าในที่นี้เป็นการคิดค่าเฉลี่ยด้วยการถ่วงน้ำหนักที่เท่ากันในแต่ละจุดภาพ



รูปที่ 14: ตัวอย่างผลการสกัด CMV จากภาพถ่ายดาวเทียม

4 บทสรุป

4.1 สรุปผลการดำเนินการ

จากผลลัพธ์การดำเนินการที่ได้นำเสนอไว้ข้างต้น ได้แก่แบบจำลองฐานสำหรับการพยากรณ์ความเข้มแสงอาทิตย์ ระยะการพยากรณ์ล่วงหน้าตั้งแต่ 30, 60, ..., 240 นาทีและผลการดำเนินการวิเคราะห์ความสัมพันธ์ระหว่างการเคลื่อนที่ของเมฆและความเข้มแสงอาทิตย์ โดยขั้นตอนถัดไปจะพัฒนาแบบจำลองการพยากรณ์ที่ใช้ข้อมูลที่สกัดได้จากภาพถ่ายดาวเทียมที่มาจากเทคนิคที่แตกต่างกัน สำหรับเป็นตัวแปรต้นในแต่ละแบบจำลองการเรียนรู้ด้วยเครื่องและแบบจำลองการเรียนรู้เชิงลึกต่อไป

4.2 แผนการดำเนินงาน

สำหรับแผนการดำเนินงานในภาคการศึกษาต้นจะแบ่งเป็น 3 ส่วนหลักดังตารางที่ 4 ได้แก่ การศึกษางานวิจัยที่เกี่ยวข้อง การทดลองเกี่ยวกับข้อมูลภาพถ่ายดาวเทียม และการพัฒนาแบบจำลองการเรียนรู้ โดยนายณัฐพล เดชตระกูลวงศ์เป็นผู้รับผิดชอบในส่วนของการสกัดข้อมูลของภาพถ่ายดาวเทียม และนายณัฐนนท โตงามร์รักษ์เป็นผู้รับผิดชอบในส่วนของการพัฒนาแบบจำลองการเรียนรู้ต่าง ๆ

ขั้นตอนการดำเนินการ	วิชา 2102490					วิชา 2102499				
	ส.ค.	ก.ย.	ต.ค.	พ.ย.	ธ.ค.	ม.ค.	ก.พ.	มี.ค.	เม.ย.	พ.ค.
1.ศึกษางานวิจัยที่เกี่ยวข้อง	■	■								
2.วิเคราะห์ข้อมูลเบื้องต้นที่จะเป็นข้อมูลนำเข้าของแบบจำลอง		■								
3.ทดลองวิเคราะห์การเคลื่อนที่ของเมฆด้วย CMV เทคนิคต่าง ๆ			■	■	■					
4.พัฒนาแบบจำลองฐาน				■						
5.ทดลองนำ CMV ที่ได้มาเป็นคุณลักษณะของแบบจำลอง				■	■					
6.จัดทำข้อเสนอโครงการฉบับสมบูรณ์และสื่อการนำเสนอโครงการ				■	■					
7.จัดทำแบบจำลองโครงข่ายประสาทเทียม-แบบสังวัตนาการ CNN						■	■			
8.เปรียบเทียบผลลัพธ์ภายในตำแหน่งที่ตั้งของแบบจำลองต่าง ๆ						■	■			
9.จัดทำรายงานและการนำเสนอโครงการฉบับสมบูรณ์								■	■	

ตารางที่ 4: แผนการดำเนินงานในการจัดทำโครงการสำหรับปีการศึกษา 2565

เอกสารอ้างอิง

- [1] G. M. Yagli, D. Yang, O. Gandhi, and D. Srinivasan, “Can we justify producing univariate machine-learning forecasts with satellite-derived solar irradiance?,” *Applied Energy*, vol. 259, p. 114122, 2020.
- [2] E. Lorenz, A. Hammer, D. Heinemann, *et al.*, “Short term forecasting of solar radiation based on satellite data,” in *EUROSUN2004 (ISES Europe Solar Congress)*, vol. 1, Freiburg Germany, 2004.
- [3] S. Cros, J. Badosa, A. Szantaï, and M. Haeffelin, “Reliability predictors for solar irradiance satellite-based forecast,” *Energies*, vol. 13, no. 21, p. 5566, 2020.
- [4] B. K. Horn and B. G. Schunck, “Determining optical flow,” *Artificial intelligence*, vol. 17, no. 1-3, pp. 185–203, 1981.
- [5] B. D. Lucas, T. Kanade, *et al.*, *An iterative image registration technique with an application to stereo vision*, vol. 81. Vancouver, 1981.
- [6] G. Farneböck, “Two-frame motion estimation based on polynomial expansion,” in *Scandinavian conference on Image analysis*, pp. 363–370, Springer, 2003.
- [7] C. Ballester, L. Garrido, V. Lazcano, and V. Caselles, “A TV-L1 optical flow method with occlusion detection,” in *Joint DAGM (German Association for Pattern Recognition) and OAGM Symposium*, pp. 31–40, Springer, 2012.
- [8] D. Aicardi, P. Musé, and R. Alonso-Suárez, “A comparison of satellite cloud motion vectors techniques to forecast intra-day hourly solar global horizontal irradiation,” *Solar Energy*, vol. 233, pp. 46–60, 2022.
- [9] C. W. Chow, S. Belongie, and J. Kleissl, “Cloud motion and stability estimation for intra-hour solar forecasting,” *Solar Energy*, vol. 115, pp. 645–655, 2015.
- [10] T. Carrière, R. A. e Silva, F. Zhuang, Y.-M. Saint-Drenan, and P. Blanc, “A new approach for satellite-based probabilistic solar forecasting with cloud motion vectors,” *Energies*, vol. 14, no. 16, p. 4951, 2021.
- [11] E. Pérez, J. Pérez, J. Segarra-Tamarit, and H. Beltran, “A deep learning model for intra-day forecasting of solar irradiance using satellite-based estimations in the vicinity of a pv power plant,” *Solar Energy*, vol. 218, pp. 652–660, 2021.
- [12] H.-M. Zuo, J. Qiu, Y.-H. Jia, Q. Wang, and F.-F. Li, “Ten-minute prediction of solar irradiance based on cloud detection and a long short-term memory (LSTM) model,” *Energy Reports*, vol. 8, pp. 5146–5157, 2022.
- [13] L. Cheng, H. Zang, Z. Wei, T. Ding, and G. Sun, “Solar power prediction based on satellite measurements – a graphical learning method for tracking cloud motion,” *IEEE Transactions on Power Systems*, vol. 37, no. 3, pp. 2335–2345, 2022.
- [14] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de Pison, and F. Antonanzas-Torres, “Review of photovoltaic power forecasting,” *Solar energy*, vol. 136, pp. 78–111, 2016.
- [15] P. Ineichen and R. Perez, “A new airmass independent formulation for the linke turbidity coefficient,” *Solar Energy*, vol. 73, no. 3, pp. 151–157, 2002.
- [16] M. Jamaly and J. Kleissl, “Robust cloud motion estimation by spatio-temporal correlation analysis of irradiance data,” *Solar Energy*, vol. 159, pp. 306–317, 2018.

- [17] T. H. R. T. Gareth James, Daniela Witten, *An Introduction to Statistical Learning - with Applications in R*. Springer Texts in Statistics, Springer Science+Business Media, 2 ed., 2021.
- [18] P. Bacher, H. Madsen, and H. A. Nielsen, “Online short-term solar power forecasting,” *Solar energy*, vol. 83, no. 10, pp. 1772–1783, 2009.
- [19] M. Rana, I. Koprinska, and V. G. Agelidis, “Univariate and multivariate methods for very short-term solar photovoltaic power forecasting,” *Energy Conversion and Management*, vol. 121, pp. 380–390, 2016.