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

Natanon Tongamrak ID 6232007121
Natthapol Dejtrakulwongse ID 6232011621

Advisor : Assoc. Prof. Jitkomut Songsiri

Department of Electrical Engineering,
Faculty of Engineering Chulalongkorn University

OVERVIEW

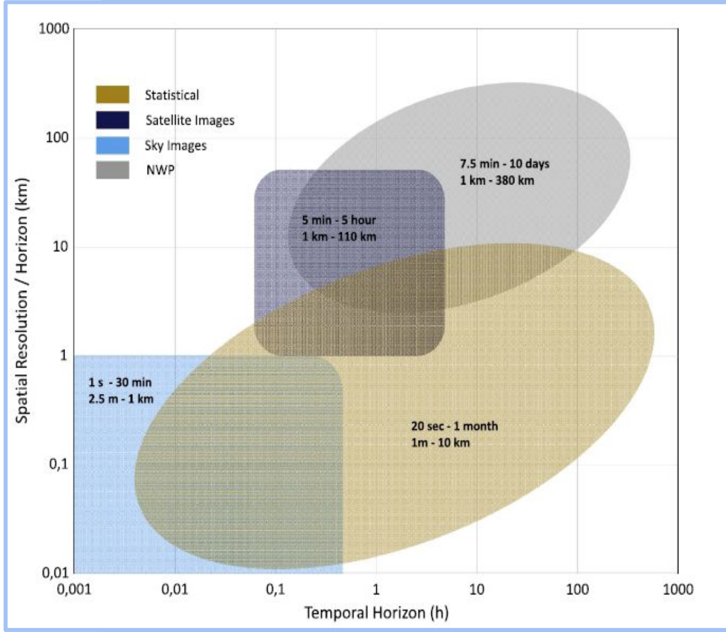
MOTIVATION

MOTIVATION

Why we have to do the Intra-day forecast ?

- The forecast horizon of around 1 - 6 hours (Intra-day) can help grid operation management e.g. load-following.

- Cloud information is necessary to estimate the future irradiance in the horizon of an Intra-day which can be extracted from satellite images.



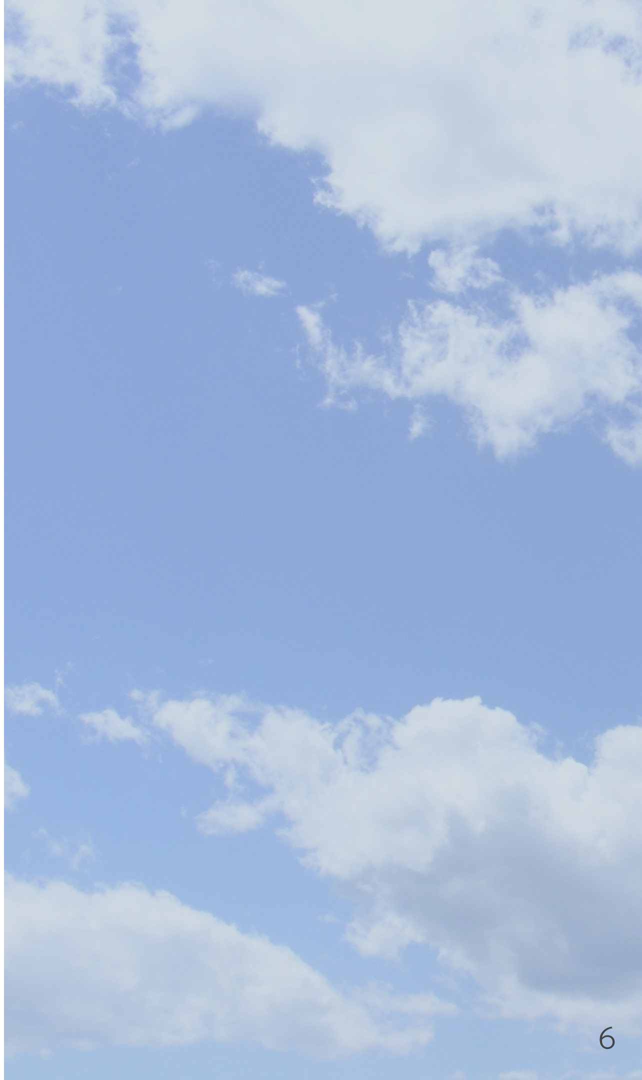
Source : J. Antonanzas et al. "Review of photovoltaic power forecasting," Solar energy, vol. 136, pp. 78–111, 2016.

OBJECTIVE

OBJECTIVE
Project goals

- Analyze the extracted information from cloud images then use it to develop forecasting model.

- Improve and compare the forecast model performance from various ML methods i.e. Linear regression, Support vector regression, Random forest, and also CNN.



METHODOLOGY

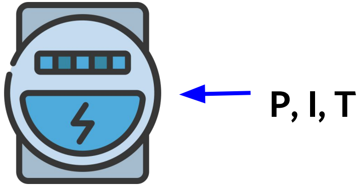
METHODOLOGY
Data acquisition

1. IMPACT SOLAR GROUP

IMPACT SOLAR GROUP

- They provided the measurement from sensors; Generated power (P [kW]), Irradiance (I [W/m^2]), and Temperature (T [°C]).

- The data were obtained from site No.48 located at Pathumwan Bangkok with a period of 15 minutes.



METHODOLOGY

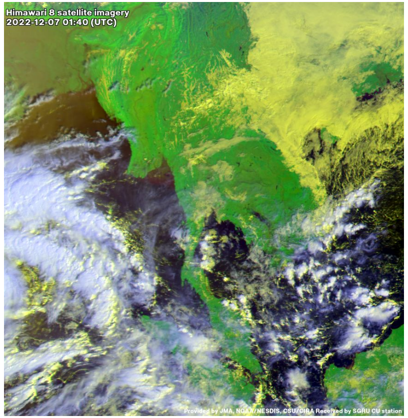
Data acquisition

2. SGRU

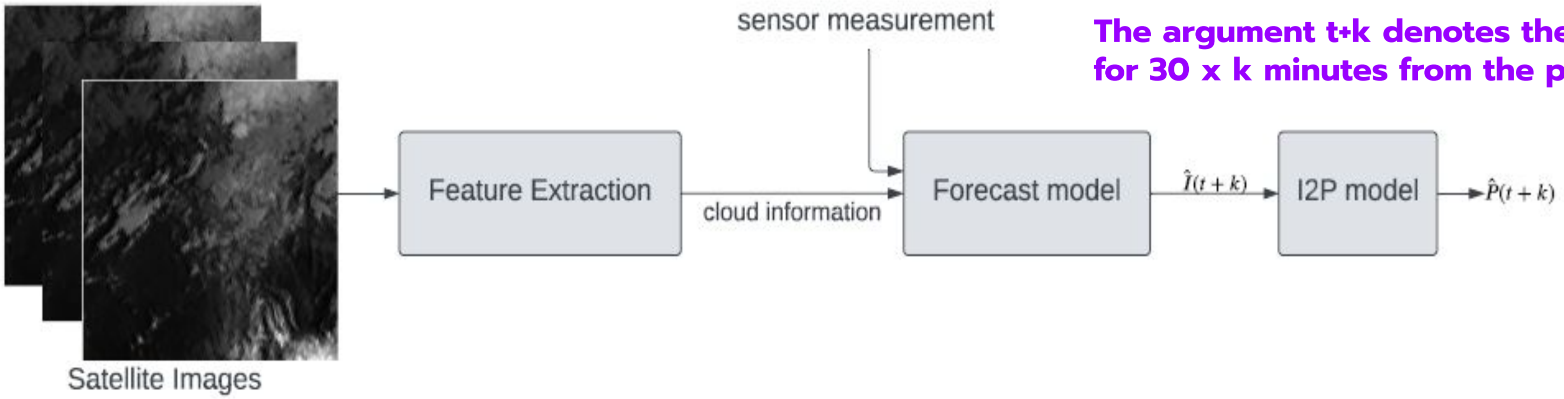
SGRU
Smart Grid Research Unit

- Cloud images from Himawari satellite were received at the ground station located at CUEE by SGRU

- The pictures came with a period of 10 minutes, resolution of 1725x1670 pixels and each pixel represents area with size 2x2 km^2



METHODOLOGY
Overall scheme of project

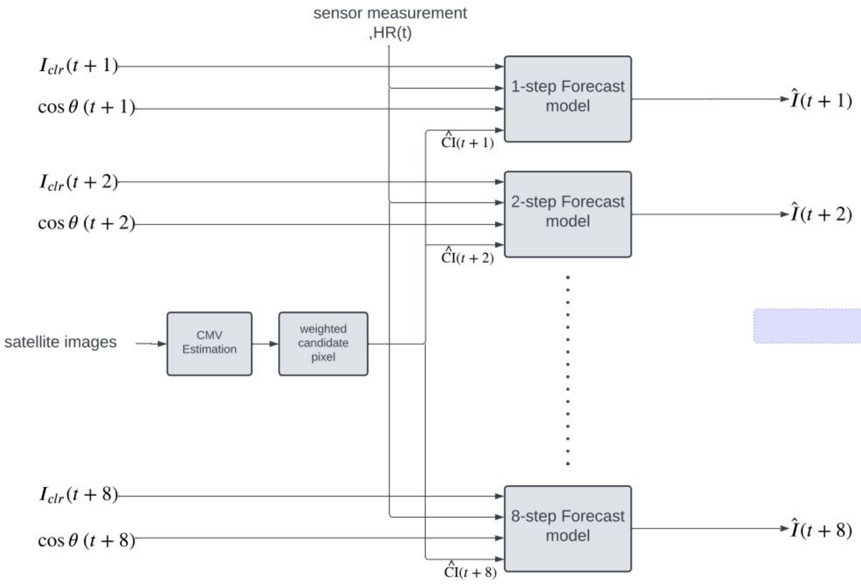


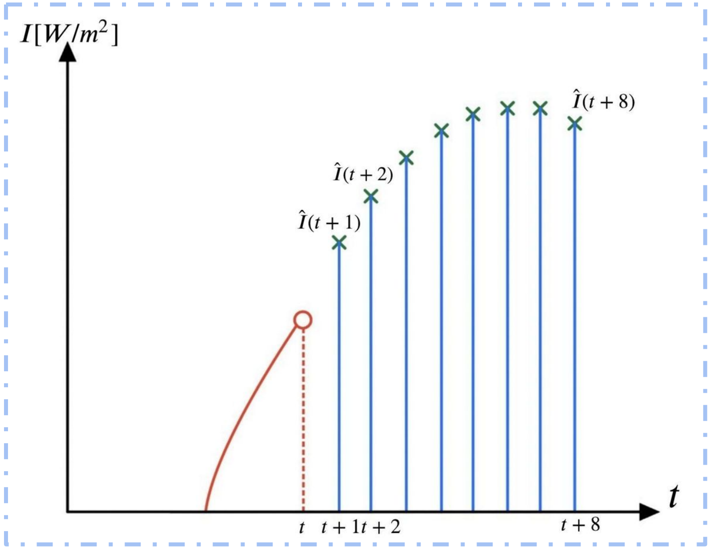
The argument t+k denotes the future time for 30 x k minutes from the present time t

- The forecast model used sensor measurement and cloud information extracted from satellite images as predictors to predict the next k steps of irradiance.
- To get the generated power, Irradiance to Power (I2P) model was applied to convert the estimated irradiance.

METHODOLOGY

Forecasting in each horizon



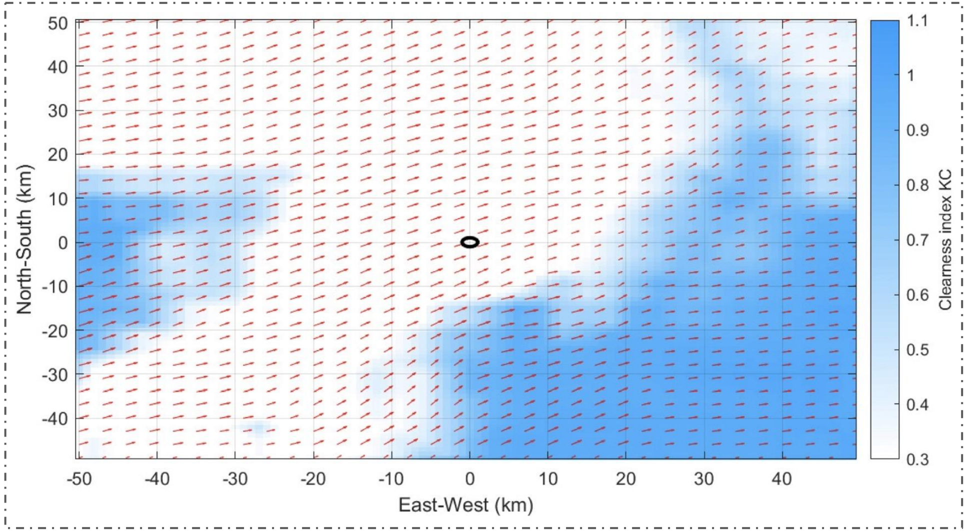


In this project, we will forecast the irradiance in a horizon of 30, 60, ..., 240 minutes (8 steps), the estimated irradiance comes from separated forecast models.

Cloud motion vector

CMV is the vector that represents velocity in each pixel which can be determined by 2 methods including

- Block matching
- Optical flow

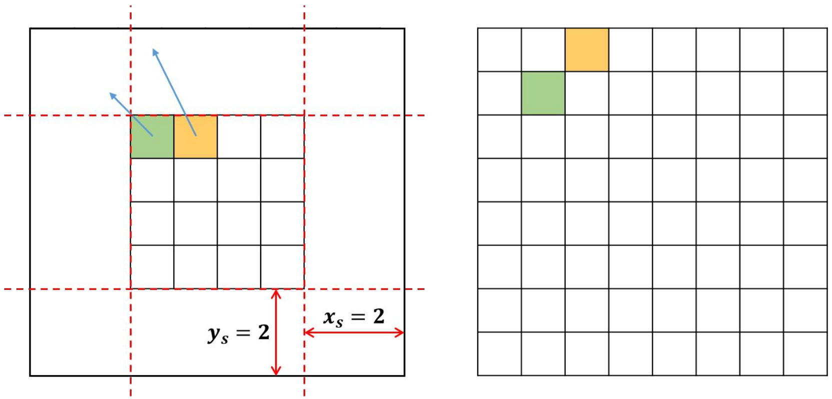


Source : T. Carrière, R. A. e Silva, F. Zhuang, Y.-M. Saint-Drenan, and P. Blanc, "A new approach for satellite-based probabilistic solar forecasting with cloud motion vectors", Energies, vol. 14, no. 16, p. 4951, 2021.

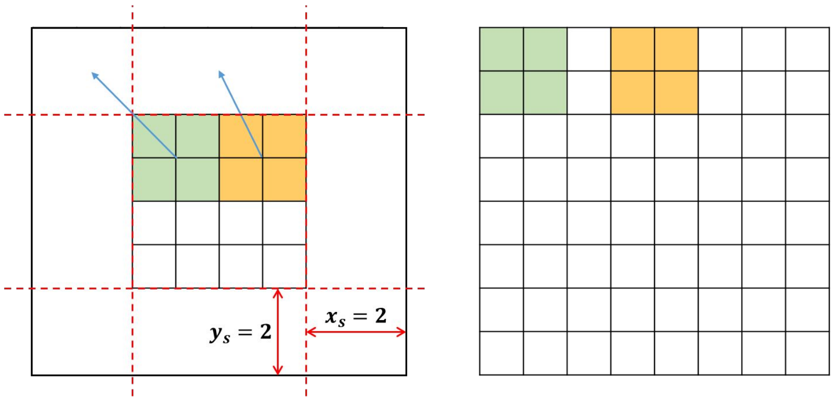
Block matching

Block matching is an algorithm which searches a paired pixel that yields the highest Correlation

- Exhaustive search - search every pixel in the search domain
- Box search - grouping pixels into a box and then search



Exhaustive search



Box search

Optical flow

Optical flow is the method that determine CMV in every pixel in picture with the following assumptions :

- The intensity remains the same between two consecutive images
- No formation/deformation and no spreading of cloud
- Motion of cloud is treated the same as motion of rigid body

All of these assumptions can be written in the mathematical expression as

$$\boxed{I_{\text{pix}}(x+dx, y+dy, t+dt) = I_{\text{pix}}(x, y, t)} \Rightarrow \boxed{\frac{\partial I_{\text{pix}}}{\partial x} v_x + \frac{\partial I_{\text{pix}}}{\partial y} v_y + \frac{\partial I_{\text{pix}}}{\partial t} \approx 0}$$

Horn-Schunck method

Horn-Schunck proposed the optimization problem

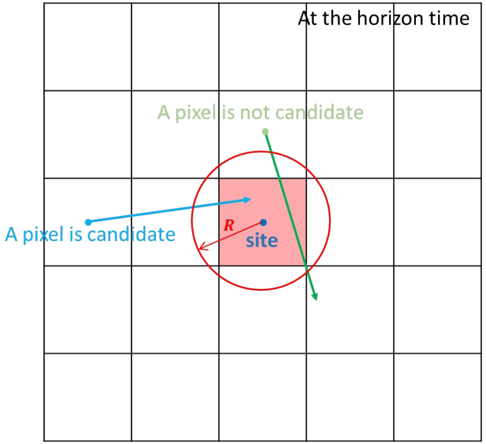
$$\arg\min_{v_x, v_y} \int_{\text{pic}} (\nabla_x v_x)^2 + (\nabla_y v_y)^2 + \lambda \cdot \left(\frac{\partial I_{\text{pix}}}{\partial x} \cdot v_x + \frac{\partial I_{\text{pix}}}{\partial y} \cdot v_y + \frac{\partial I_{\text{pix}}}{\partial t} \right)^2$$

For solving this optimization problem together with the OF assumptions, this will yield a unique solution of velocity.

METHODOLOGY
Candidate pixel

A pixel will get candidate if it satisfies both the spatial and temporal conditions

- Temporal condition - a pixel get closest to the interested location at the time around the forecast horizon
- Spatial conditions - a minimum distance between pixel and the location must not exceed a threshold R



METHODOLOGY
Candidate pixel

- The distance between pixel and the interested location can be expressed as

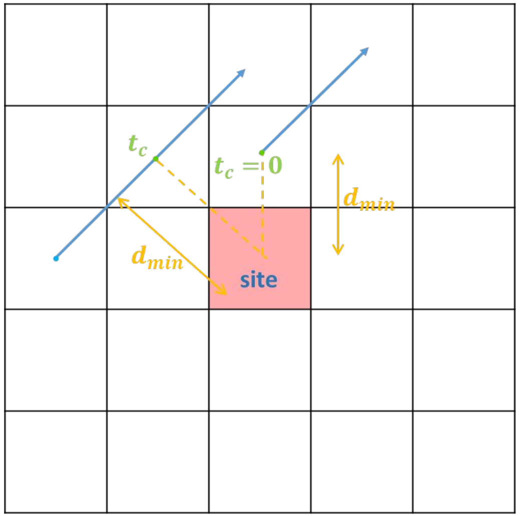
$$d(t) = \sqrt{[(x - x_{\text{site}}) + tv_x]^2 + [(y - y_{\text{site}}) + tv_y]^2}$$

- A time that pixel get closest to the interested location is given by

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

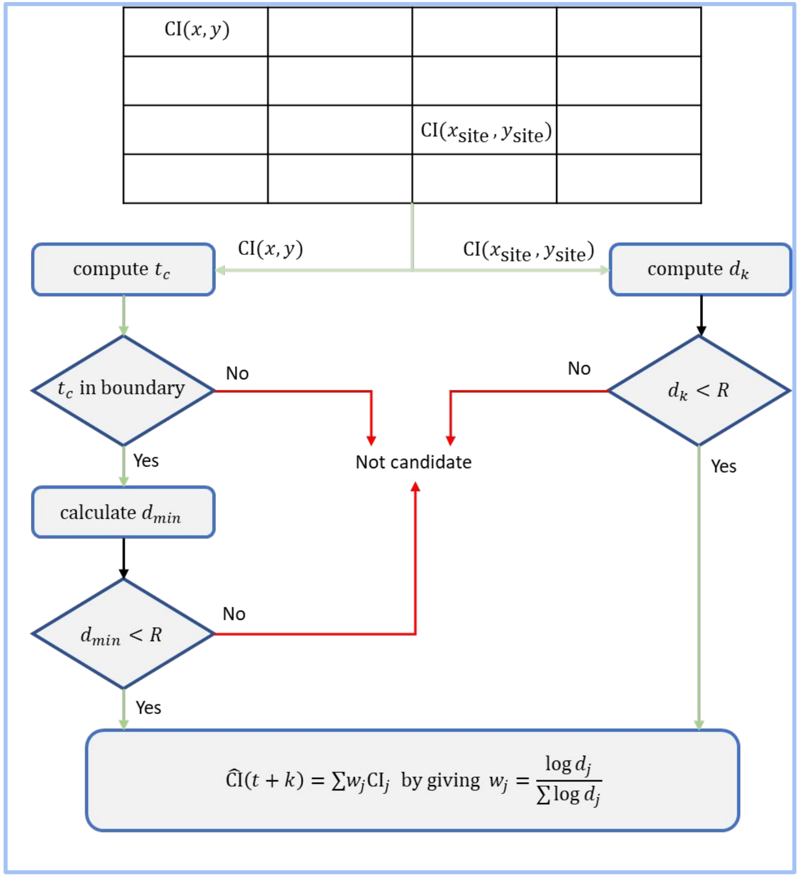
- Minimum distance between pixel and the interested location is given by

$$d_{\min} \triangleq d(t_c)$$



METHODOLOGY

The procedure of weighted average CI

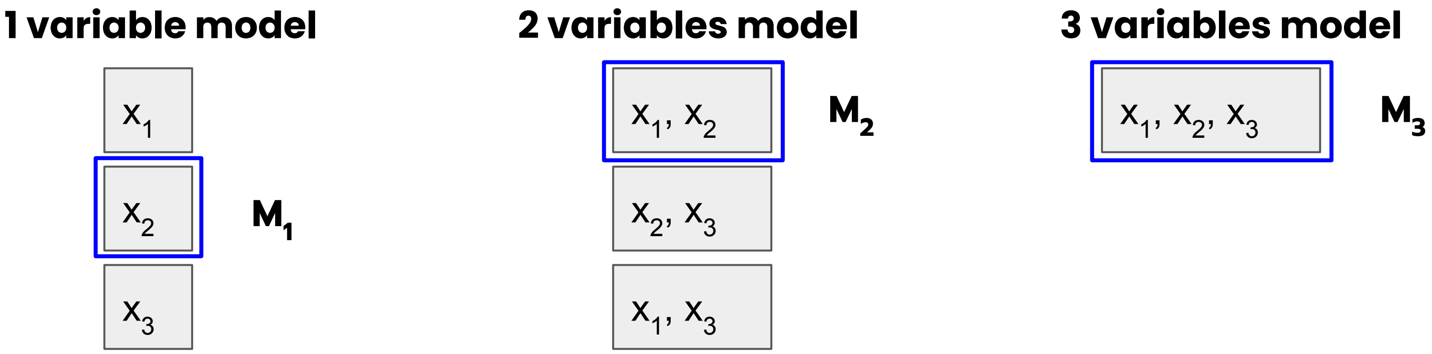


- For each pixel we compute time t_c and minimum distance d to determine whether such pixel is candidate or not.
- After completing the determination process, we will calculate the estimated cloud index by the proposed expression.

METHODOLOGY
Best subset selection

Let the predictor X has p variable(s) then there are 2^p possible subset of predictor, We want to select the subset that provide the best prediction result in the following steps

1. Denote M_0 be the blank model for not using any predictor.
2. For the model with k variable(s), choose only one group of predictor that give the best model performance from all possible group $\binom{p}{k}$ then define as M_k.
3. For model M_1, M_2, ..., M_p , apply the model selection criterion in order to choose the best model.



Model selection criterion

The selection criterion provides the score that weighted between model loss and model complexity, In this work we use 2 criterions that is Akaike information criterion (AIC) and Bayes information criterion (BIC)

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

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

- N is the number of validation data and d is the number of parameter used.
- Noted that this is true under the assumption that the linear regression model is disturbed by noise that is gaussian distributed.

PRELIMINARY RESULTS

PRELIMINARY RESULTS

Outcomes of this semester

- Best subset selection result
- Baseline model
- Extracting cloud information

Data augmentation

1. Training data - We only use it in the feature selection process. In this early semester, data from February 5, 2022 - July 15, 2022 was used from both sensor and satellite image data.

2. Validating data - This dataset is used for measuring the performance of each sub-model in the selection process. The data was used from 16 July 2022 - 31 August 2022.

3. Testing data - This dataset is used to test the performance of a model that has been validated. The data was used from 1 September 2022 - 30 October 2022.

PRELIMINARY RESULTS

Best subset selection : candidate features

For building model to forecast irradiance next 30 minutes (1 step at time t+1) , the candidate features are

- Previous irradiance values : I(t), I(t-1), I(t-2), ..., I(t-6)
- Clear sky irradiance at that time : $I_{clr}(t+1)$
- Cosine of solar zenith angle at that time : $\cos(\theta(t+1))$
- Temperature : T(t)

I(t)	**I(t-1)**	**I(t-2)**	**I(t-3)**	**I(t-4)**	**I(t-5)**	**I(t-6)**	**$\cos(\theta(t+1))$**	**$I_{clr}(t+1)$**	**T(t)**

PRELIMINARY RESULTS

Best subset selection : result

For linear regression model, the set of parameters that gives the lowest value of AIC and BIC is obtained when k = 7 with score AIC_{scaled} = 10.2469 and BIC_{scaled} = 10.3122, The chosen predictors are

- Some previous irradiance : I(t-5), I(t-4), I(t-3), I(t-1), I(t)
- Clear sky irradiance at that time : $I_{clr}(t+1)$
- Cosine of solar zenith angle at that time : cos(θ(t+1))

I(t)	**I(t-1)**	**I(t-2)**	**I(t-3)**	**I(t-4)**	**I(t-5)**	**I(t-6)**	**cos(θ(t+1))**	**I_{clr}(t+1)**	**T(t)**

Baseline model : model mapping

To evaluate the capability of cloud information, the baseline model will **NOT** include cloud feature. The model for estimating irradiance I(t+k) has features which were inferred from the selection method and then use its for LR model

Output : I(t+k)

Input : I(t), I(t-1), I(t-3), I(t-4), I(t-5),

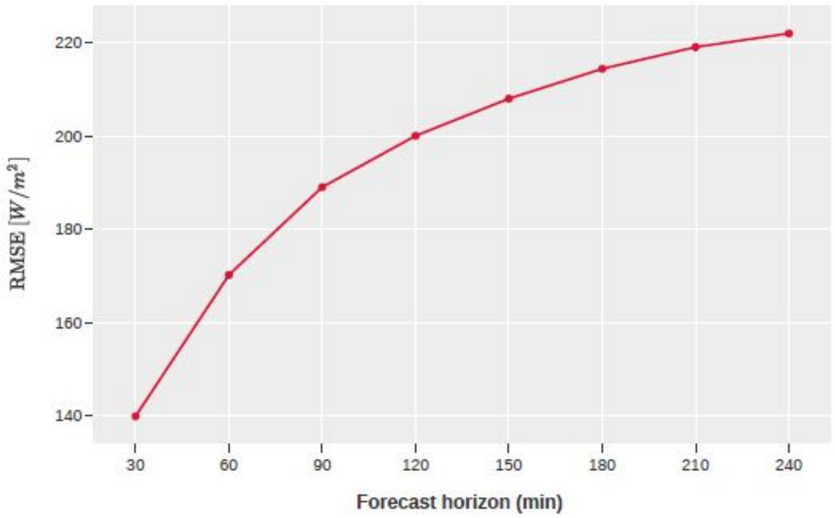
$\cos(\theta(t+k))$, $I_{clr}(t+k)$,

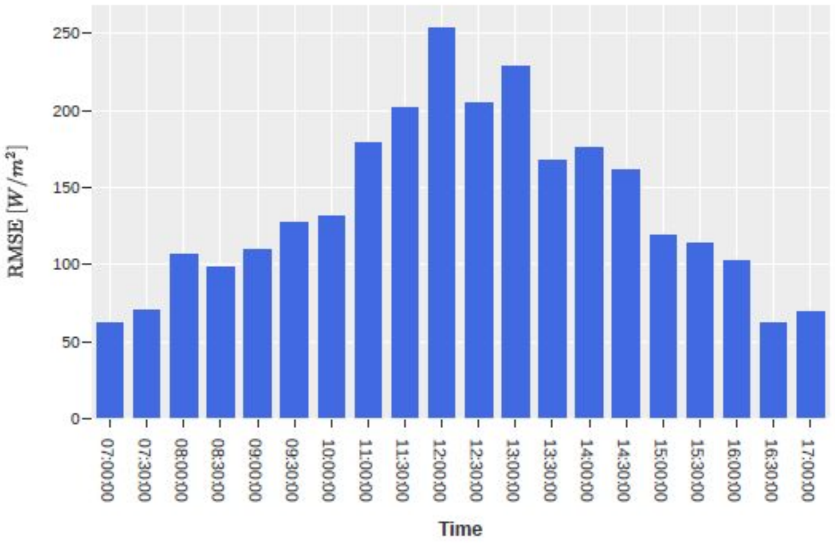
HR(t)

PRELIMINARY RESULTS

Baseline model : forecasting results

The graphs show the performance of baseline model

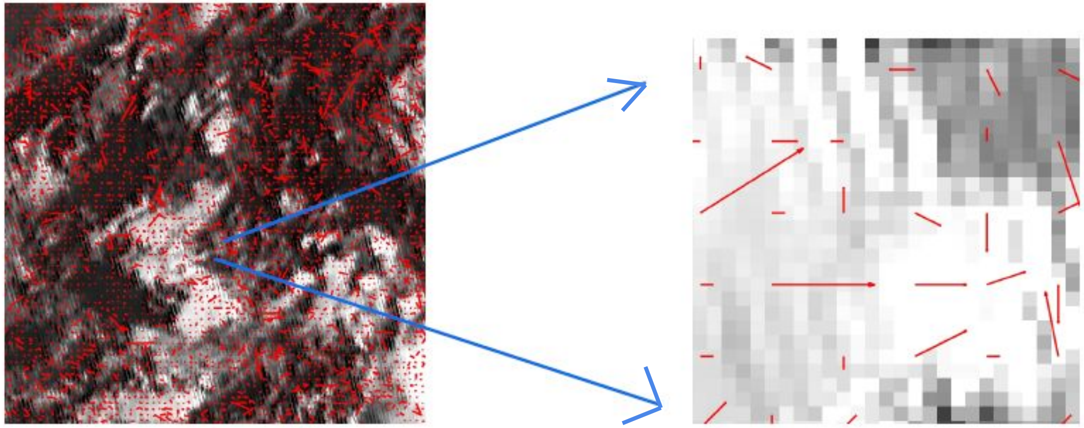




- Errors increase as the time horizon increases.
- Errors tend to be relatively high during the daytime.

Extracting cloud information

Consider the area around the experiment location



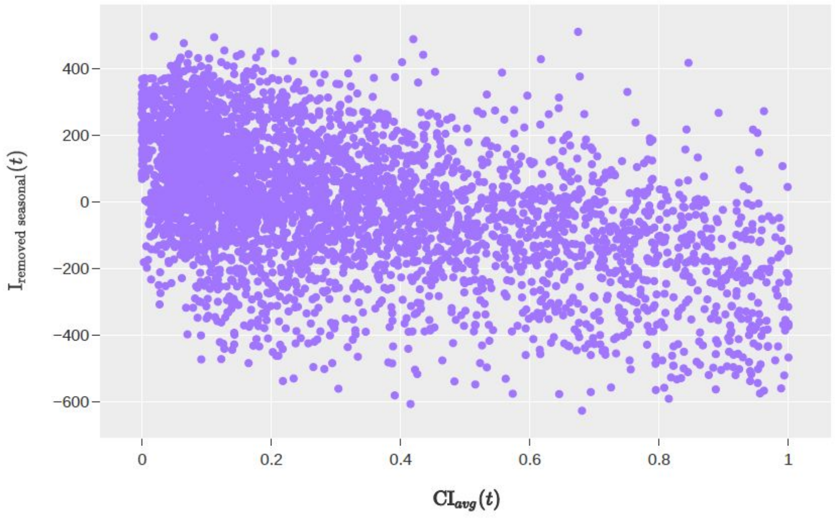
then extract CMV by HS method

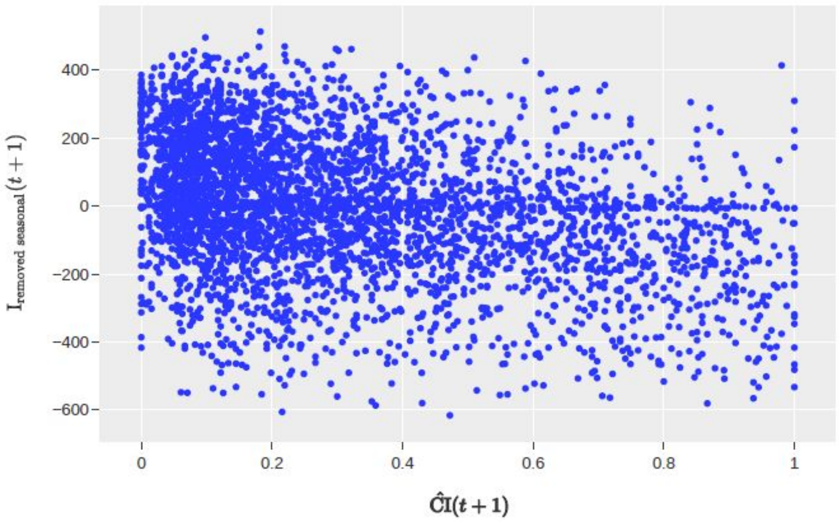
- Calculate CMV from two consecutive images by MATLAB computer vision toolbox.
- Every pixels has its own CMV.

PRELIMINARY RESULTS

Extracting cloud information

After calculate the CMV, then use its to estimate the future cloud index





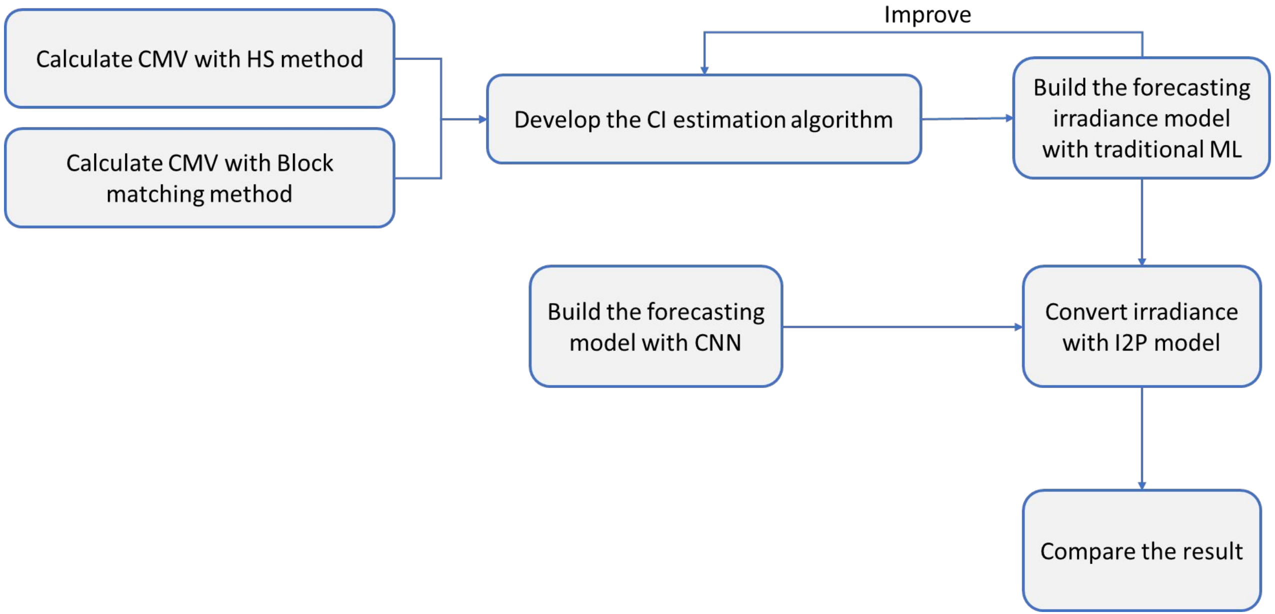
The correlation between the estimated cloud index and irradiance at time t + k should appear in the same aspect as the correlation between the averaging cloud index and irradiance at the present time.

Future work

Future work

The plan for next semester are

- Try to extract CMV with Block matching method
- Experiment and improve CI estimation algorithm
- Build forecasting model with CNN
- Convert irradiance with I2P model



REFERENCE

[1] B. K. Horn and B. G. Schunck, “Determining optical flow”, Artificial intelligence, vol. 17, no. 1-3, pp. 185–203, 1981.

[2] 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.

[3] 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.

[4] 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.

Q&A

THANK YOU

Backup

BACKUP

Model selection score

k	RSS($\times 10^5$)	R^2	Feature	$\text{AIC}_{\text{scaled}}$	$\text{BIC}_{\text{scaled}}$
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	$I(t-5), I(t-1), I(t), I_{\text{clr}}(t+1), \cos(\theta(t+1))$	10.2491	10.3145
6	190.4	0.6354	$I(t-5), I(t-3), I(t-1), I(t), I_{\text{clr}}(t+1), \cos(\theta(t+1))$	10.2471	10.3124
7	190.3	0.6354	$I(t-5), I(t-4), I(t-3), I(t-1), I(t), I_{\text{clr}}(t+1), \cos(\theta(t+1))$	**10.2469**	**10.3122**
8	190.4	0.6354	$I(t-6), I(t-5), I(t-4), I(t-3), I(t-1), I(t), I_{\text{clr}}(t+1), \cos(\theta(t+1))$	10.2470	10.3123
9	190.4	0.6353	$I(t-6), I(t-5), I(t-4), I(t-3), I(t-2), I(t-1), I(t), I_{\text{clr}}(t+1), \cos(\theta(t+1))$	10.2472	10.3126
10	192.1	0.6321	$I(t-6), I(t-5), I(t-4), I(t-3), I(t-2), I(t-1), I(t), I_{\text{clr}}(t+1), \cos(\theta(t+1)), T(t)$	10.2559	10.3213

MATLAB CV toolbox for calculating CMV