Solar Irradiance Forecasting at Chulalongkorn University

Naebboon Hoonchareon

David Banjerdpongchai

Wanchalerm Pora

Suchin Arunsawatwong

Jitkomut Songsiri

Department of Electrical Engineering

CHULA ENGINEERING

Foundation toward Innovation

Outline

- Chula team
- resources and facilities
- data preprocessing
- forecasting methods
  - ANN with weather classification
  - time series models
  - model output statistics (MOS)
  - forecasting with cloud motion detection
- discussions on methods and implementation

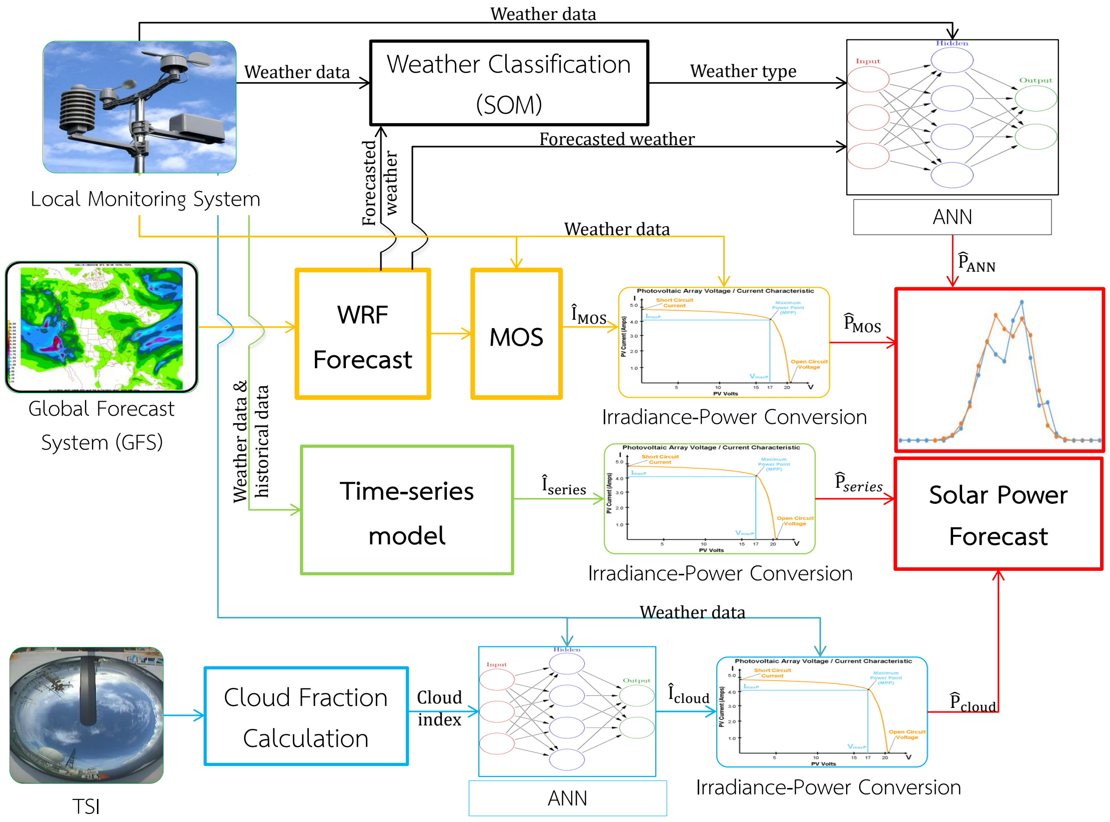
professors: Jitkomut, David, Wanchalerm, Suchin, and Naebboon

Chula team members

RAs and engineers: Supachai, Suchakrey, Rujipart, Sarawoot, and Sarawoot

Project overview

our current framework consists of four forecasting methods



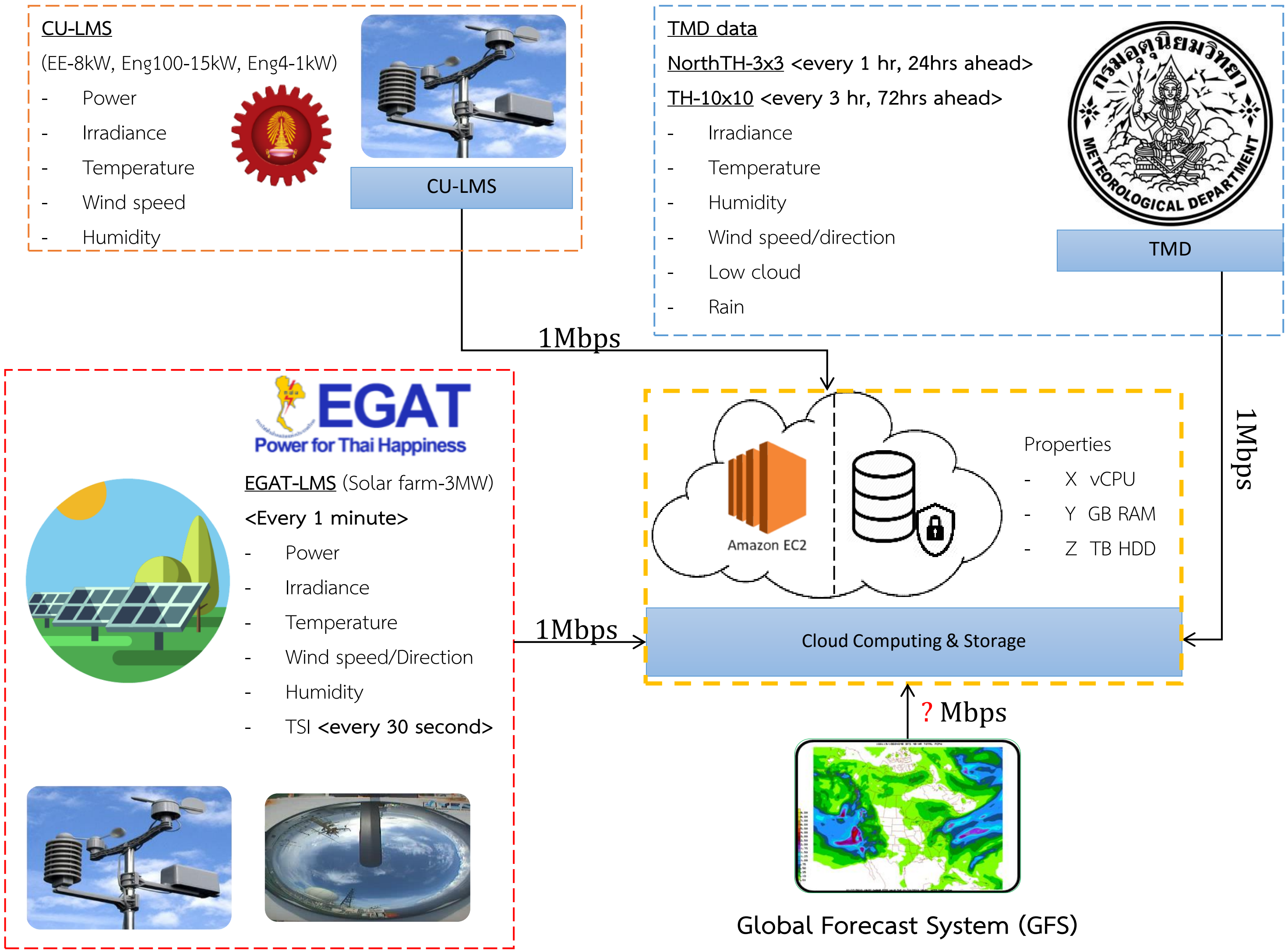
- ANN with weather classification
- time series forecasting
- model output statistics (MOS)
- forecasting with cloud motion prediction

Resources

available equipment: rooftop PV system (8+15 kw) + pyrometer

with sensors of: wind speed, wind direction, temperature, relative humidity, UV index, irradiance, power of solar cells

ICT Structure



Available data of CU location

measurements: sampling period (mostly) is 3 mins

1. solar data:
   - solar irradiance
   - solar power
2. weather data:
   - temperature
   - relative humidity
   - wind speed
   - wind direction
   - UV index

prediction: provided by TMD (Thailand Meteorological Department)

- forecasted WRF of weather and solar irradiance (10×10 km^2 on every 3-hour)

Available data of MHS location

Mae Hong Son project has a solar farm of size 3MW (provided by EGAT)

measurements: sampling period is 10 mins (but data are mostly incomplete)

1. solar data: irradiance and power
2. weather data:
   - temperature
   - relative humidity
   - wind speed
   - wind direction
   - UV index
   - pressure

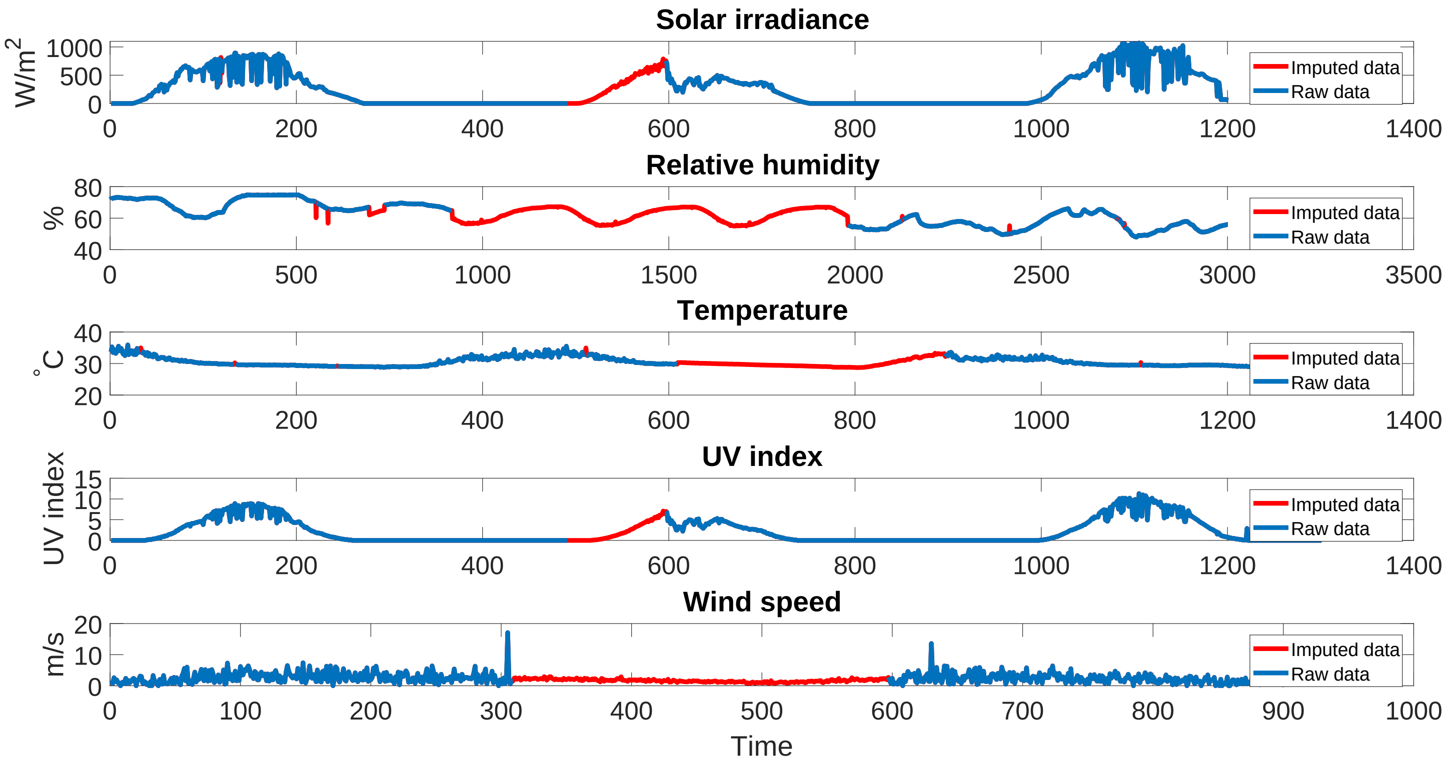
prediction: provided by TMD (Thailand Meteorological Department)

- forecasted WRF of weather and solar irradiance (3×3 km^2 and hourly)

Data preprocessing

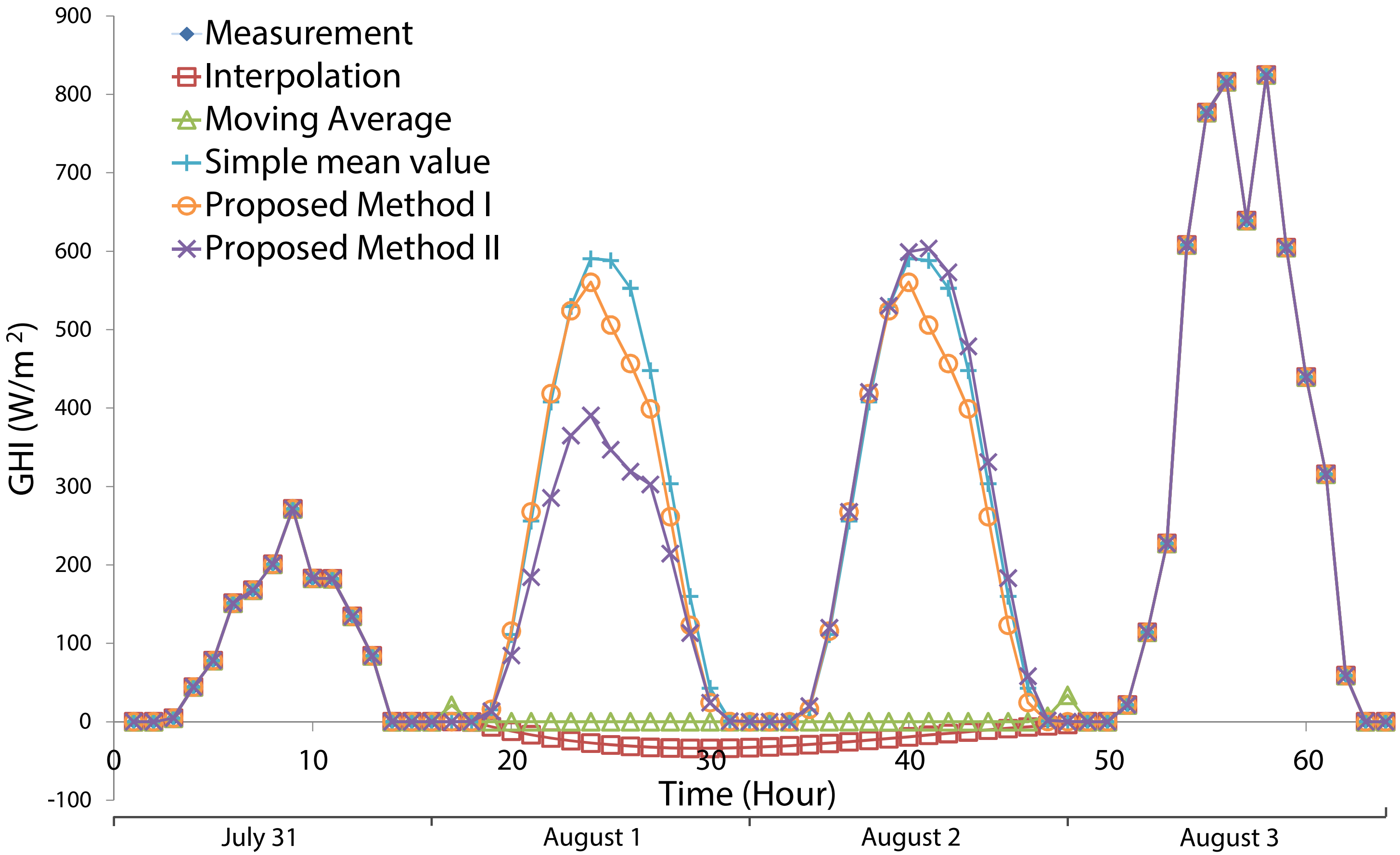
- missing-data imputation: irradiance, humidity, temperature, UV, wind speed
- spatial averaging on $\hat{I}_{\mathrm{wrf}}$ (over 9 grids)
- data smoothing
- others: depending on the forecasting techniques

simple method: filling the mean of 10-day before and after data



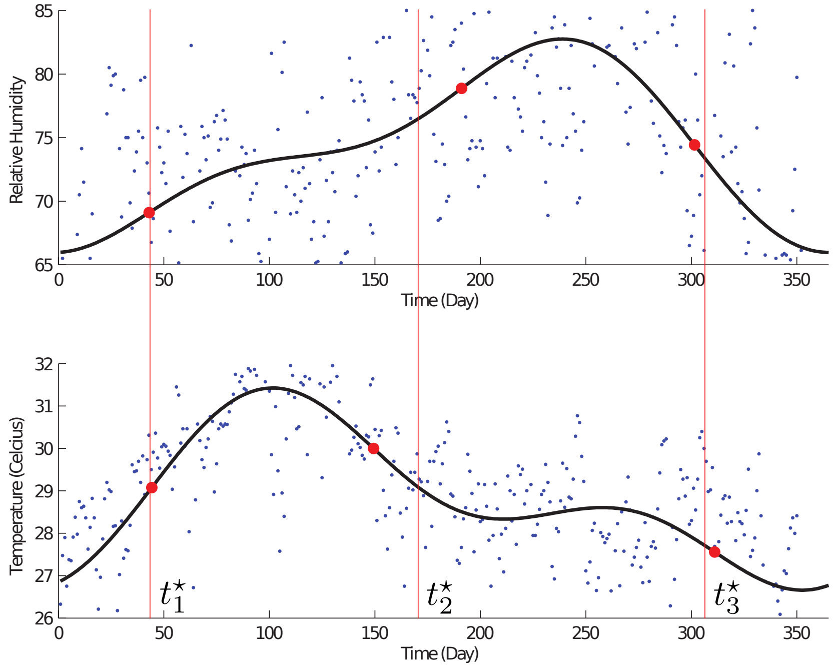
selected after comparisons among linear, spline and PCHIP interpolations

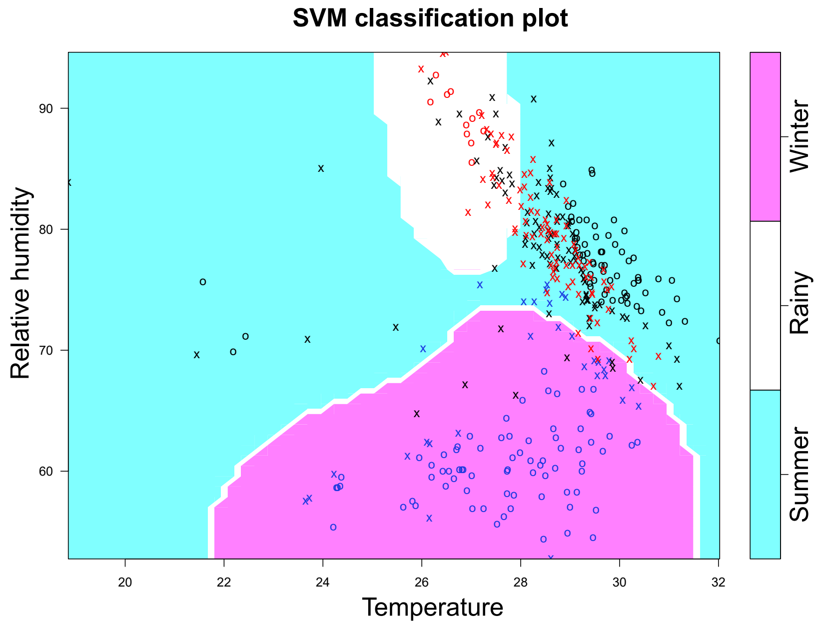
Missing value imputation during August 1-2 (TMD data)



proposed method: filling the mean of data from the same weather type

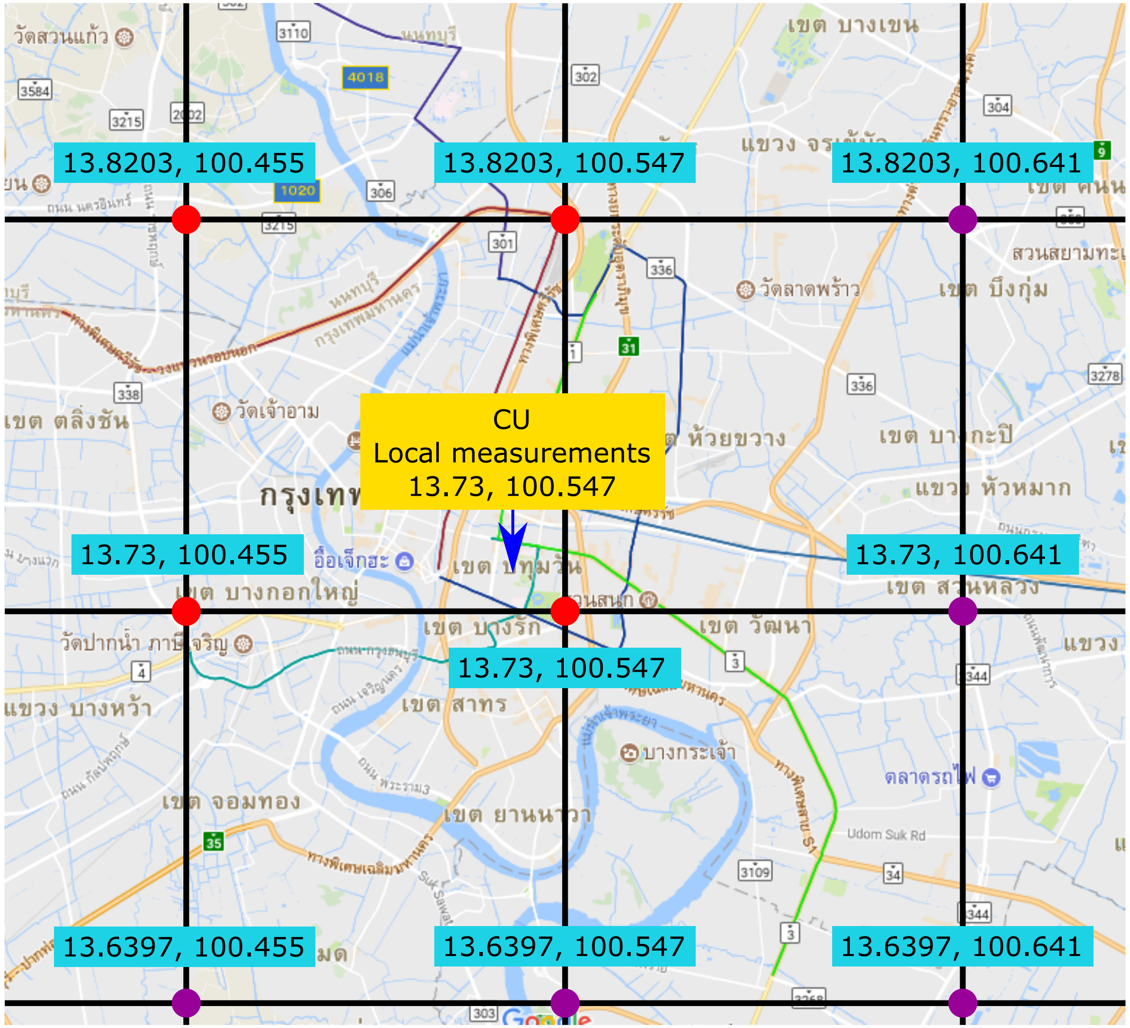
Fourier representations (solid lines) of the temperature and humidity data



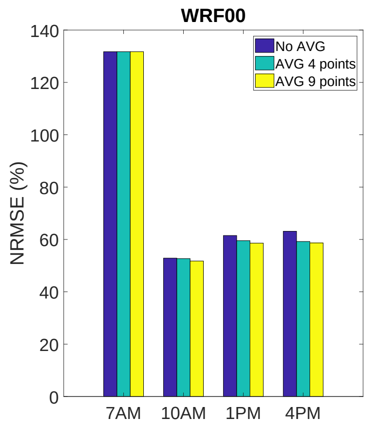


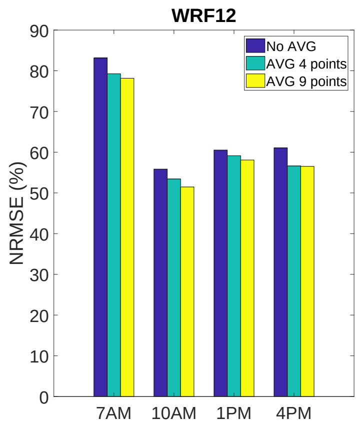
- fit Fourier representations of T and RH
- detect a prior seasonal change from their monotonicity change
- use the prior seasons as labels to train SVM as a weather classifier

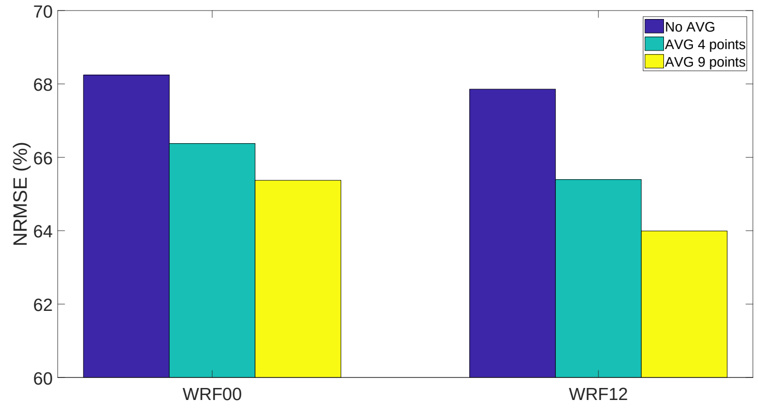
forecasted GHI by WRF are available at two times: WRF00 and WRF12



Data preprocessing — Results of spatial averaging





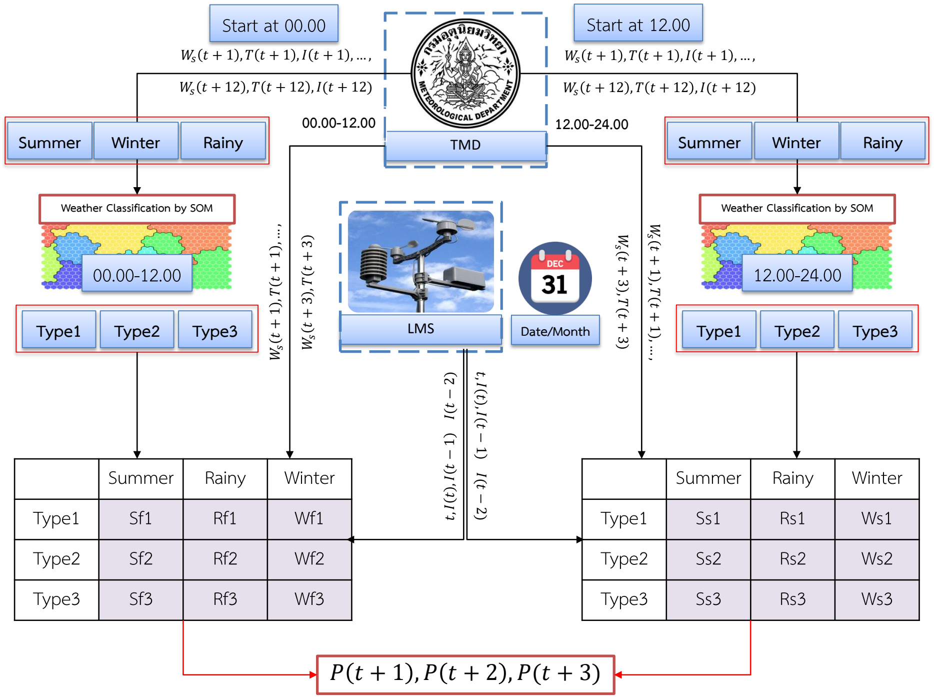


- comparison: no averaging, averaging over 4 grids, averaging over 9 grids
- performance: compare normalized error to the measured GHI
- WRF12 has less error
- improvement on performance when averaging over 9 grids
- significant improvement is seen at 1PM and 4PM

Forecasting methods

- ANN with weather classification
- time series forecasting
- model output statistics (MOS)
- forecasting with cloud motion detection
- irradiance-power conversion model

ANN with weather classification



assumption:
forecasting models should differ by the following factors

1. seasons (summer, rainy, winter)
2. weather type of the day (3 types classified by self-organized map or SOM)
3. time of the day (morning and afternoon)

these results in $3 \times 3 \times 2 = 18$ ANN models

ANN with weather classification inputs

inputs of weather classification: forecasted weather variables by WRF

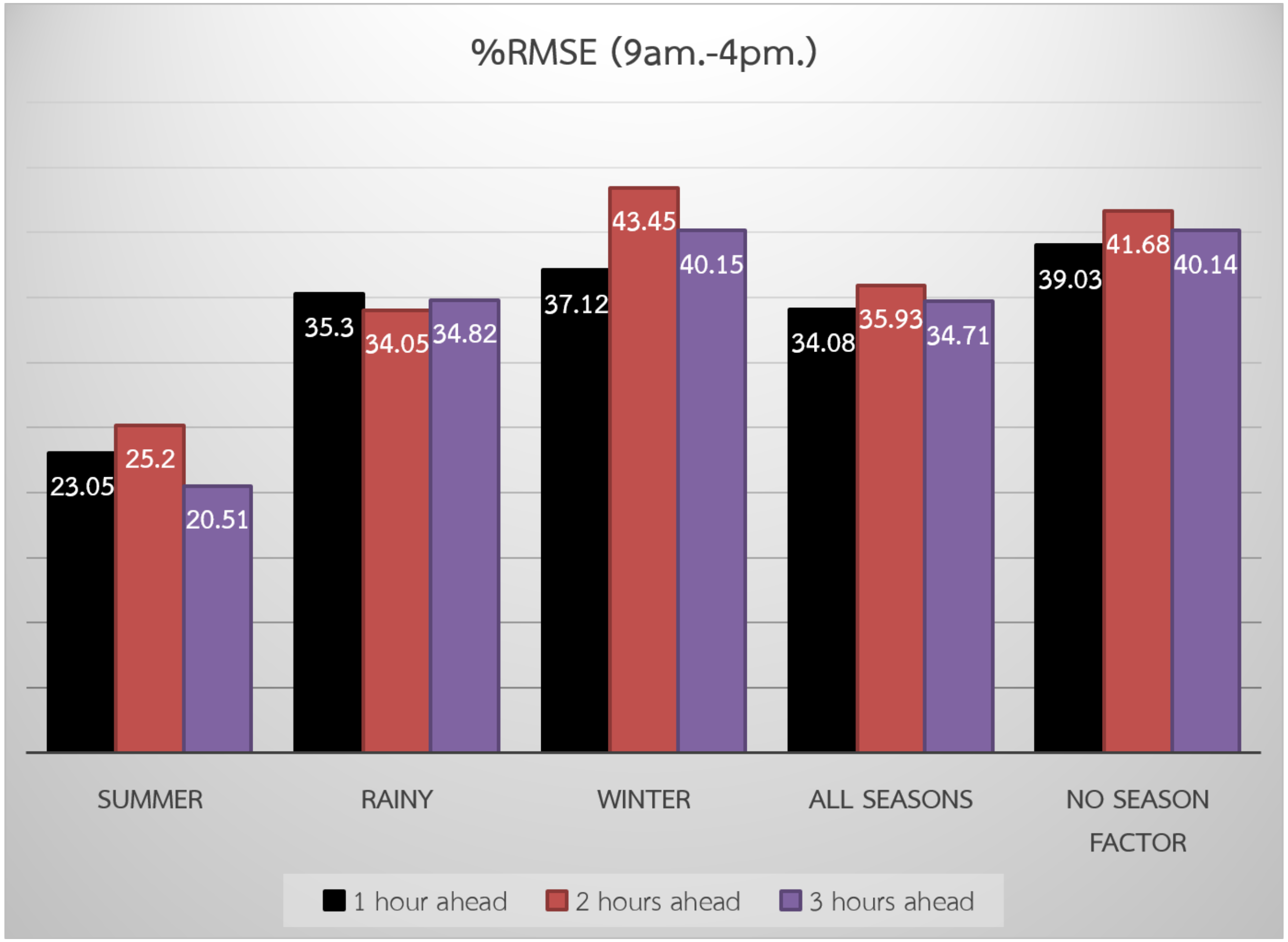
- $\mathrm{WS}(t+1), \mathrm{WS}(t+2), \mathrm{WS}(t+3)$ wind speed
- $\mathrm{T}(t+1), \mathrm{T}(t+2), \mathrm{T}(t+3)$ temperature

inputs of ANN:

- local measurements: $t, I(t), I(t-1), I(t-2)$
- classified weather type (to specify which ANN model will be used)
- forecasted weather data by WRF
  - $\mathrm{WS}(t+1), \mathrm{WS}(t+2), \mathrm{WS}(t+3)$
  - $\mathrm{T}(t+1), \mathrm{T}(t+2), \mathrm{T}(t+3)$

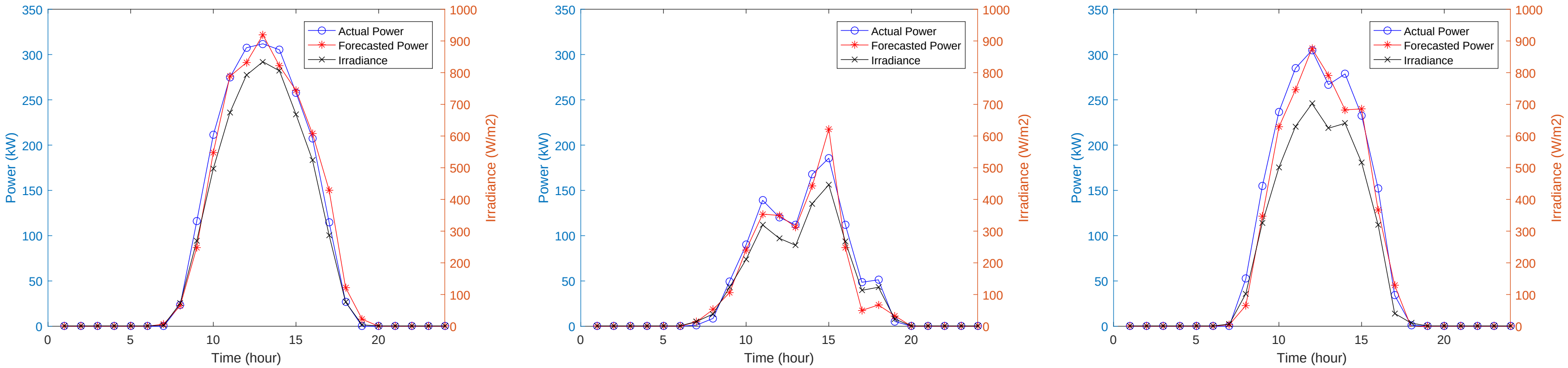
target of ANN: solar power $P(t+1), P(t+2), P(t+3)$

ANN with weather classification

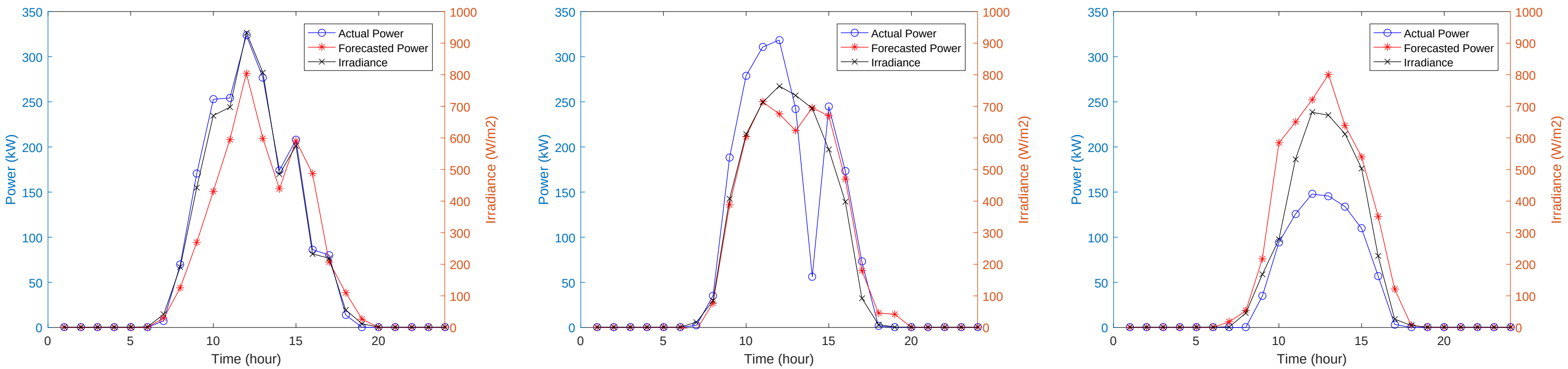


using individual ANN (specific to weather type) has reduced the error

best-day prediction in summer, rainy and winter seasons



worst-day prediction in summer, rainy and winter seasons



by assumption, solar irradiance clearly has

- seasonal trends (at least, daily and annual circles)
- been influenced from weather variables (temperature, relative humidity, wind speed, air pressure)

so we consider a seasonal ARIMAX model:

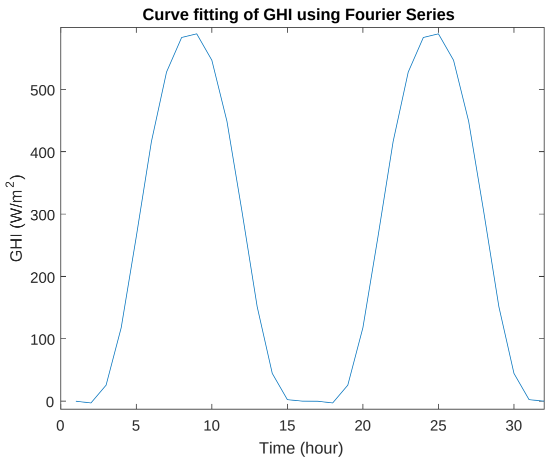
$$\tilde{A}(L)A(L)(1-L^T)^D(1-L)^d I(t) = B(L)u(t) + \tilde{C}(L)C(L)v(t) \tag{1}$$

where L is a lag operator and $A, \tilde{A}, B, \tilde{C}, C$ are polynomials in L

- d is integrated order, determined by differencing I and see the autocorrelation
- T specifies the seasonal period and D is integrated seasonal order
- u represents weather variables

to estimate the seasonal period,

- determine dominant frequencies of GHI from FFT, which are $\omega = 0.125\pi, 0.25\pi$ and 0.375π
- fit a seasonal trend to: $s(t) = \sum_{i=1}^{3} \sigma_i \sin \omega_i t + \beta_i \cos \omega_i t + \alpha$ by regression



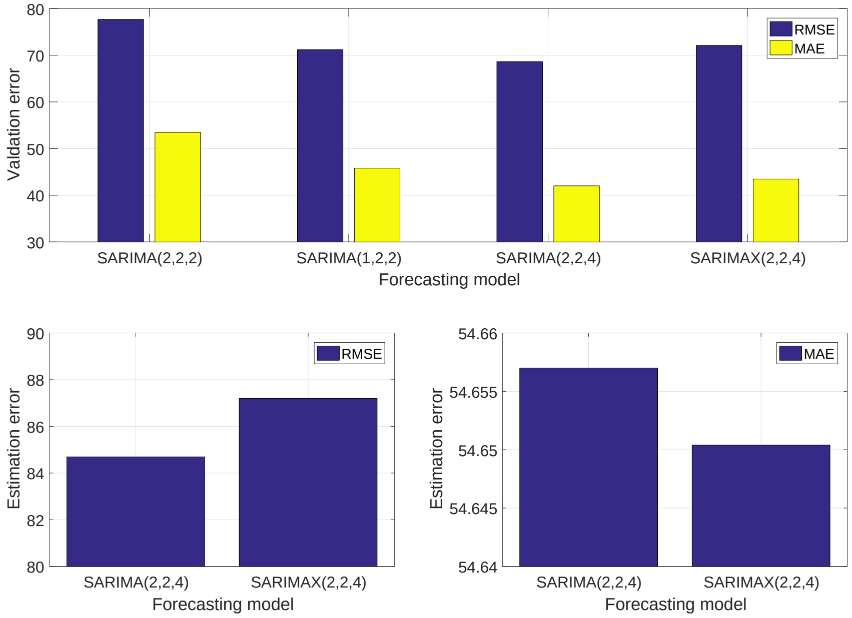
- the dominant frequencies correspond to the periods of $16, 8$ and 5.3 hours two possible ways to handle seasonal effects

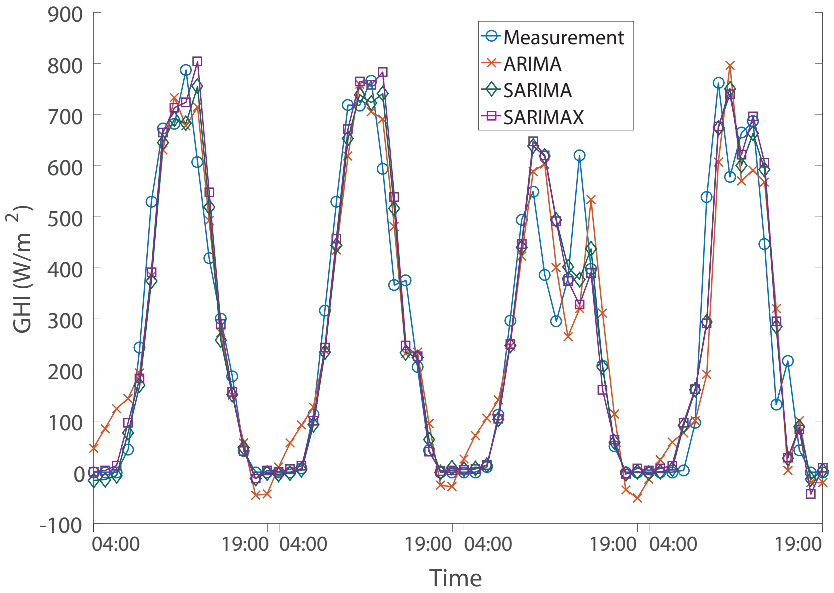
- use seasonal ARIMAX models with $T = 16$ hours
- remove the fitted seasonal trend, $s(t)$ from $I(t)$ and fit to ARIMAX

to include impacts of one-lag weather variables, we consider four models

1. seasonal ARIMA with $T = 16$
2. seasonal ARIMAX with $T = 16$
3. ARIMA (after the fitted seasonal trend is removed)
4. ARIMAX (after the fitted seasonal trend is removed)

exogenous inputs: temperature, relative humidity, wind speed and air pressure data: GHI and weather data from TMD during 2011-2014

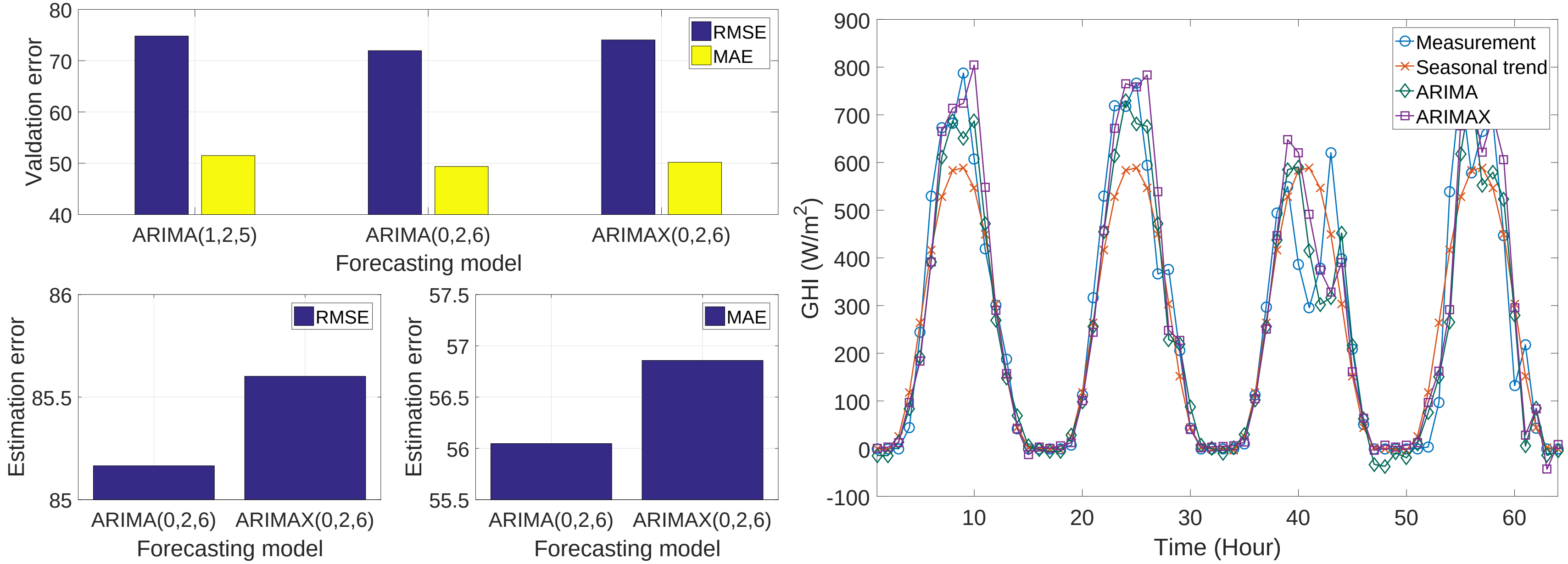




selected seasonal models: (from model selection criterions)

- SARIMA$(2, 2, 4)(0, 1, 1)_{16}$: 2-order AR, 4-order MA
- SARIMAX$(2, 2, 4)(0, 1, 1)_{16}$: 2-order AR, 4-order MA, 1-order exogenous term

data: GHI and weather data from TMD during 2011-2014



selected models: when being fitted to I after seasonal trends are removed

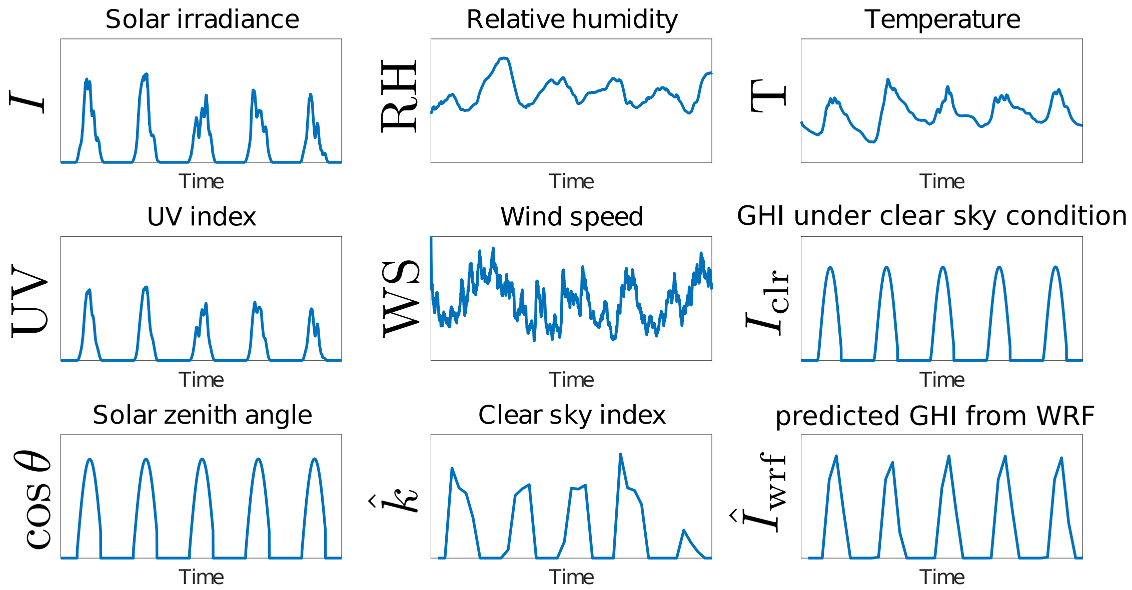
- ARIMA$(0, 2, 6)$: no AR, 6-order MA
- ARIMAX$(0, 2, 6)$: no AR, 6-order MA, 1-order exogeneous term

Model output statistics

MOS is a multiple linear regression (regress I on relevant variables)

$$I(t) = \beta_1 I(t-1) + \beta_2 \mathrm{RH}(t-1) + \beta_3 \mathrm{T}(t-1) + \beta_4 \mathrm{UV}(t-1) + \beta_5 \mathrm{WS}(t-1) + \beta_6 I_{\mathrm{clr}}(t) + \beta_7 \cos\theta(t) + \beta_8 \hat{k}(t) + \beta_9 \hat{I}_{\mathrm{wrf}}(t)$$

goal: use MOS to improve the predicted I from local weather data and $\hat{I}_{\mathrm{wrf}}$



Model output statistics — Predictor selection

select highly relevant variables in the model using

- partial correlation
- stepwise regression (backward and forward)
- subset regression (and use AIC/BIC)

summary of influential variables on solar irradiance

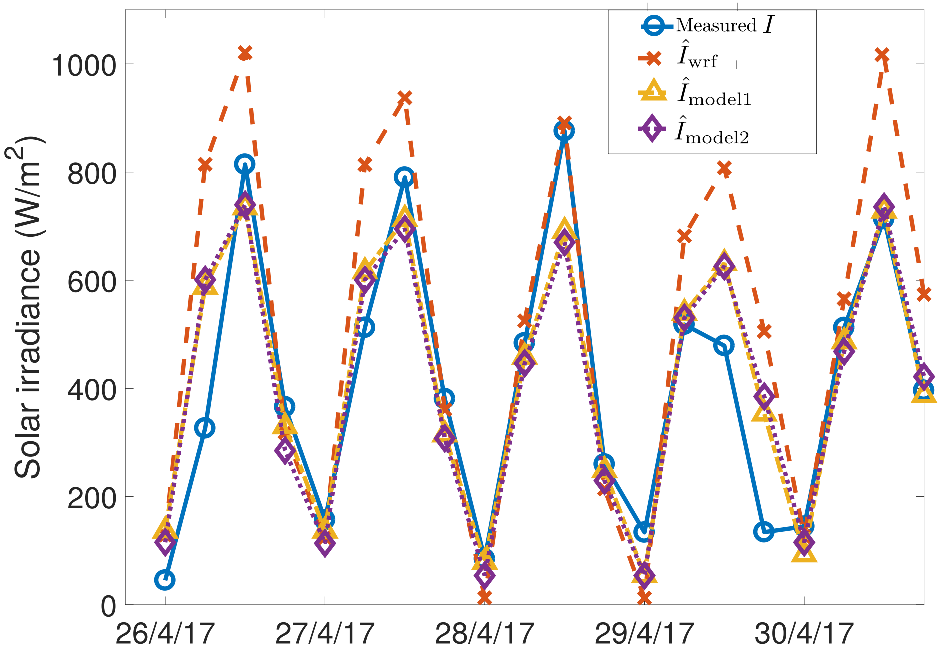
Methods	**Predictors**								
	I	RH	T	UV	WS	I_{clr}	$\cos\theta$	$\hat{k}$	$\hat{I}_{\mathrm{wrf}}$
Partial correlation	✓	✓	✓	✓					✓
Forward stepwise							✓		✓
Backward wtepwise	✓	✓		✓			✓		✓
Subset selection									
SSE validation	✓	✓	✓	✓	✓	✓	✓		✓
AIC training	✓	✓		✓			✓		✓
AIC validation							✓		✓
BIC training	✓	✓		✓			✓		✓
BIC validation							✓		✓

based on the listed method, our selected models are

model 1:

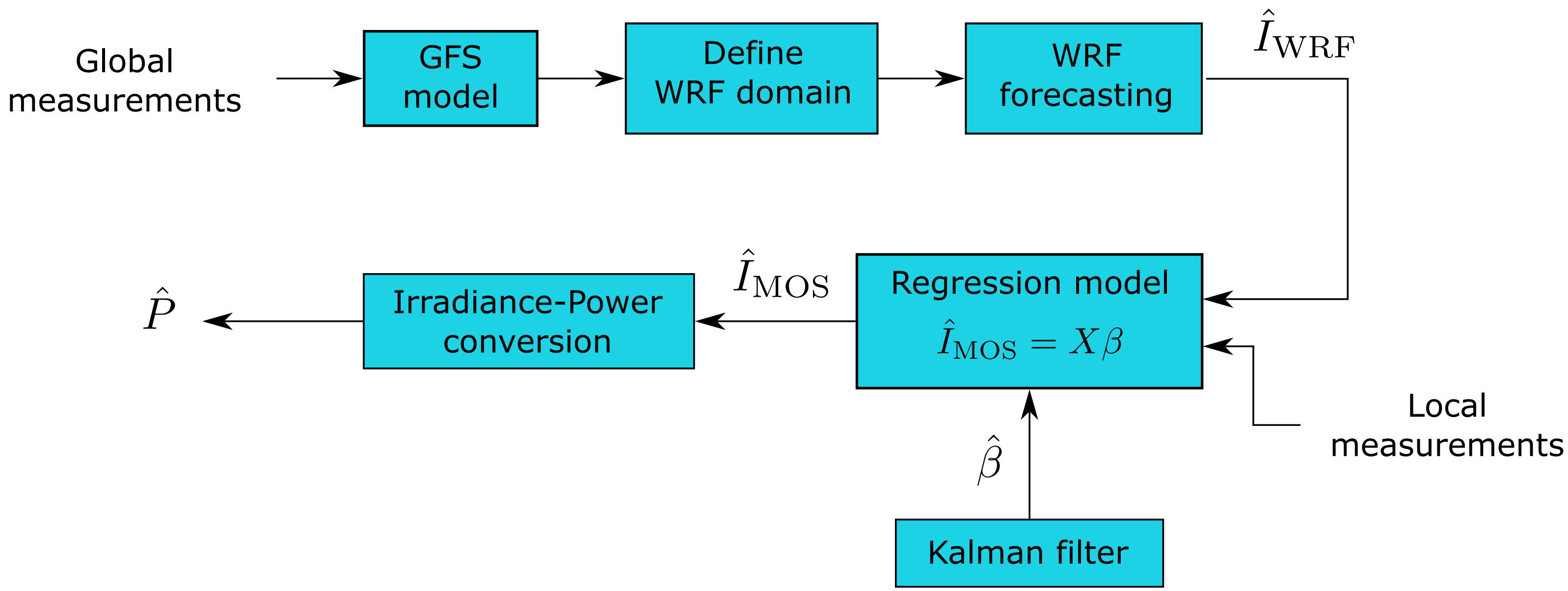
$$\hat{I}_{\mathrm{mos}}(t) = \beta_1 I(t-1) + \beta_2 \mathrm{RH}(t-1) + \beta_3 \mathrm{UV}(t-1) + \beta_4 \cos\theta(t) + \beta_5 \hat{I}_{\mathrm{wrf}}(t)$$

model 2: $\hat{I}_{\mathrm{mos}}(t) = \alpha_1 \cos\theta(t) + \alpha_2 \hat{I}_{\mathrm{wrf}}(t)$



evaluated on validation set:

- WRF mostly over-estimate I
- $\hat{I}_{\mathrm{mos}}$ improves the prediction from $\hat{I}_{\mathrm{wrf}}$

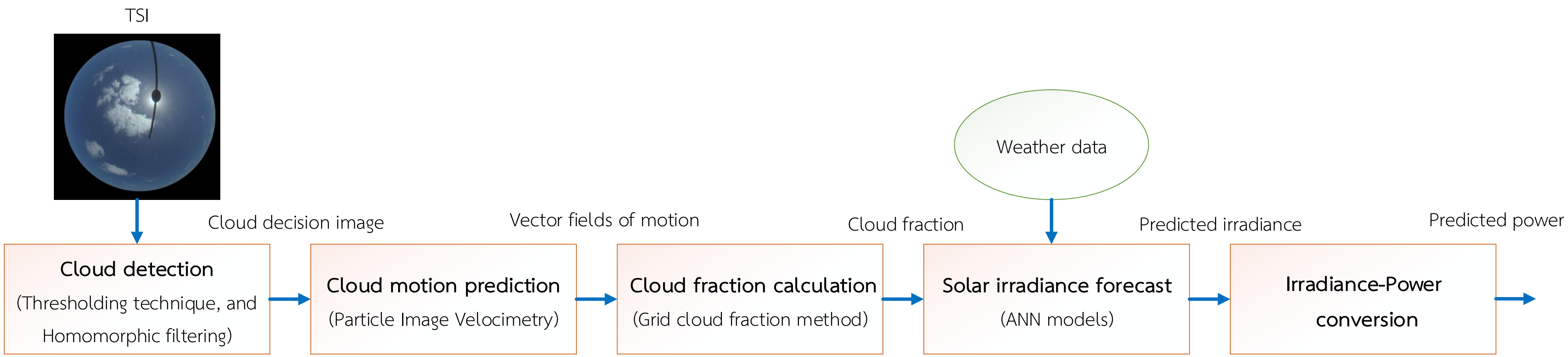


after selecting relevant variables,

- use the regression model to predict I
- the regression coefficients are allowed to be adaptive as new data arrive
- apply Kalman filter to recursively estimate those coefficients

Forecasting with cloud motion prediction

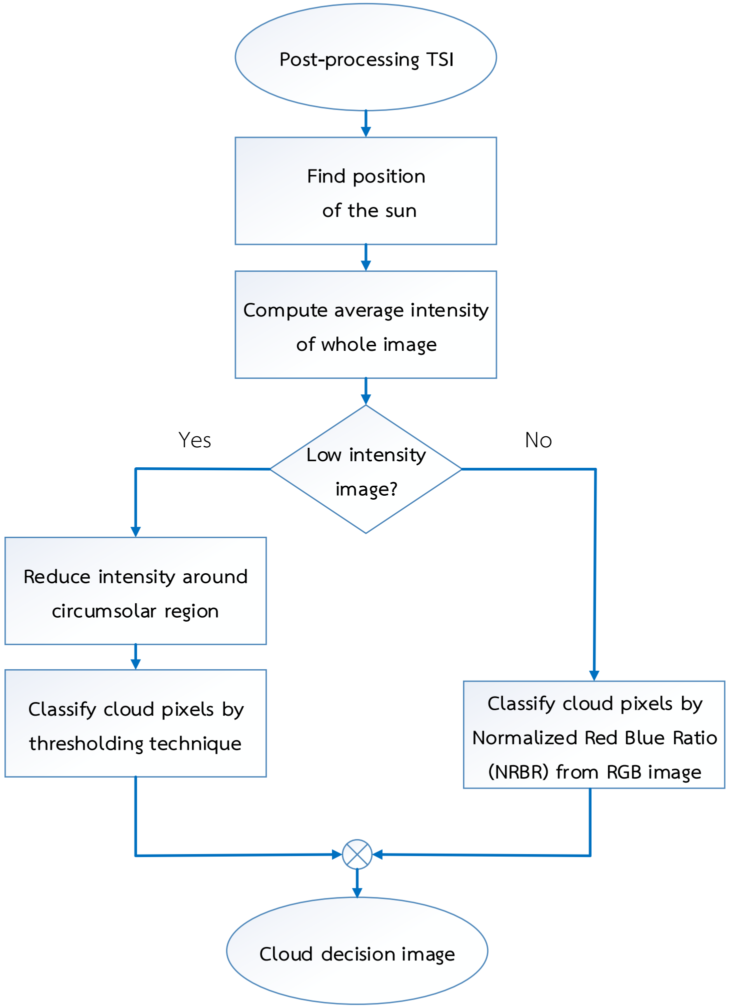
our current scheme: forecast every 10 mins



inputs of ANN:

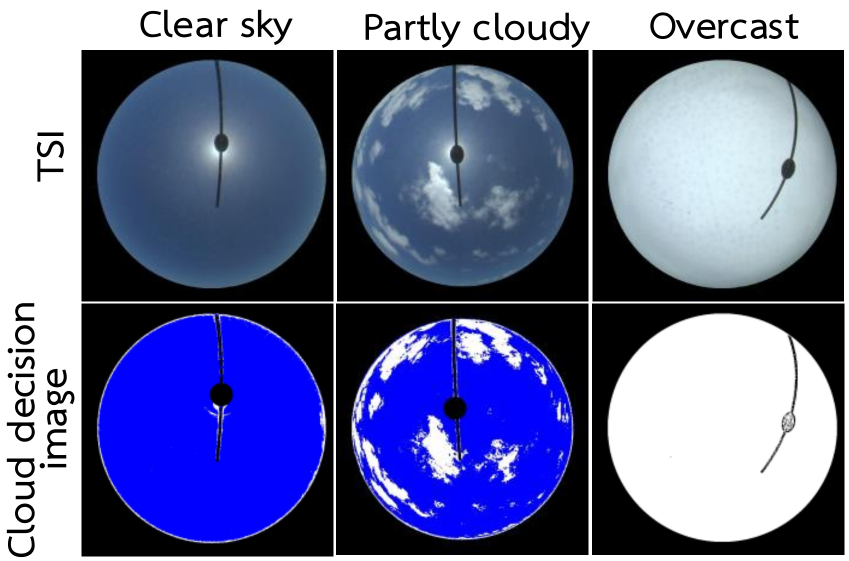
- grid cloud fraction at time t
- $I(t), I(t-1), I(t-2), I(t-3), I(t-4)$
- weather data at time t

Cloud detection

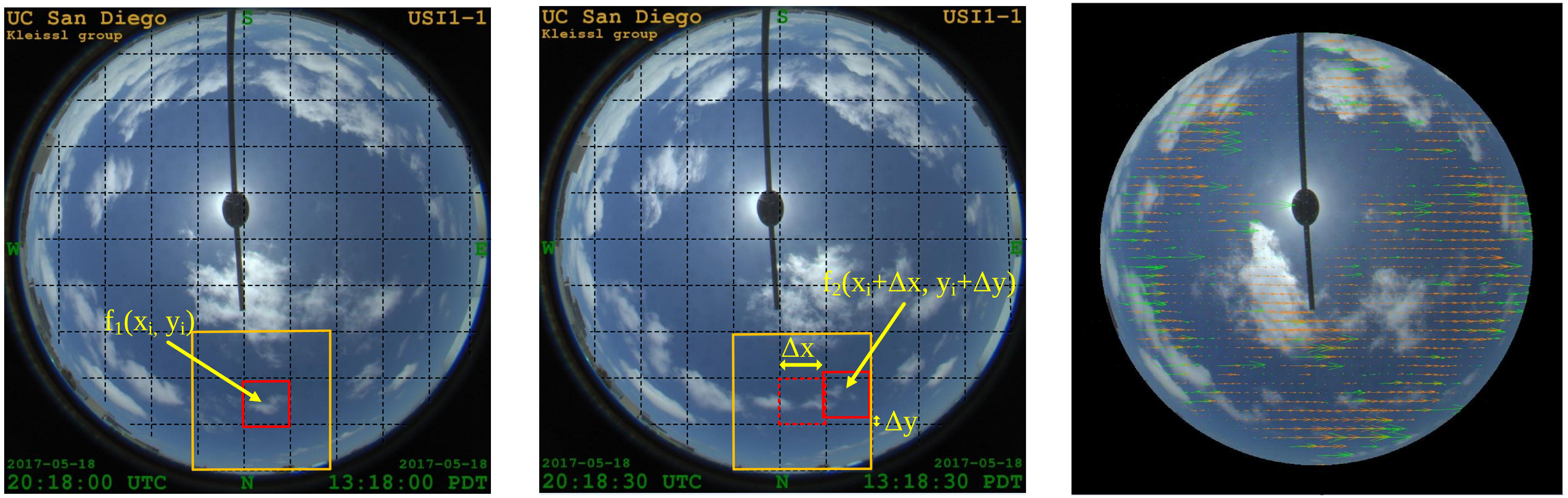


Results:

we can identify cloud pixels under various sky conditions



Cloud motion prediction



- RGBs of two sky images from consecutive times are compared using MQD

$$D(\Delta x, \Delta y) = \sum_{i=1}^{N} \sum_{j=1}^{N} |f_1(x_i, y_i) - f_2(x_i + \Delta x), y_i + \Delta y)|$$

- if MQD is low in a certain direction then the cloud should move toward such direction

Irradiance-power conversion model

under investigation

Summary

as of October 19, 2017

- data: limited and contain uncertainty in some variables (need cleaning)
- WRF prediction: limited and available in coarse temporal resolution (every 3 hours)
- relevant variables: surprisingly, temperature is not selected but solar zenith angle and predicted WRF are the most influential variables
- ANN forecasting: should develop specific ANN model for each weather type
- time series forecasting: may require recursive forecasting in the online implementation
- MOS and forecasting with cloud: under being experimented

Acknowledgement

special Thank to

- Director Somkuan Tonjan, Numerical Division, Weather Forecast Bureau, Thai Meteorological Department (TMD) for providing the GHI data, WRF forecasting and practical information
- Vichaya Laohanun for missing-value imputation and time series forecasting experiments