Senior Project Proposal 2102490 Year 2016

Solar irradiance forecasting for Chulalongkorn University location using time series models

Vichaya Layanun ID 5630550721

Advisor: Assist. Prof. Jitkomut Songsiri

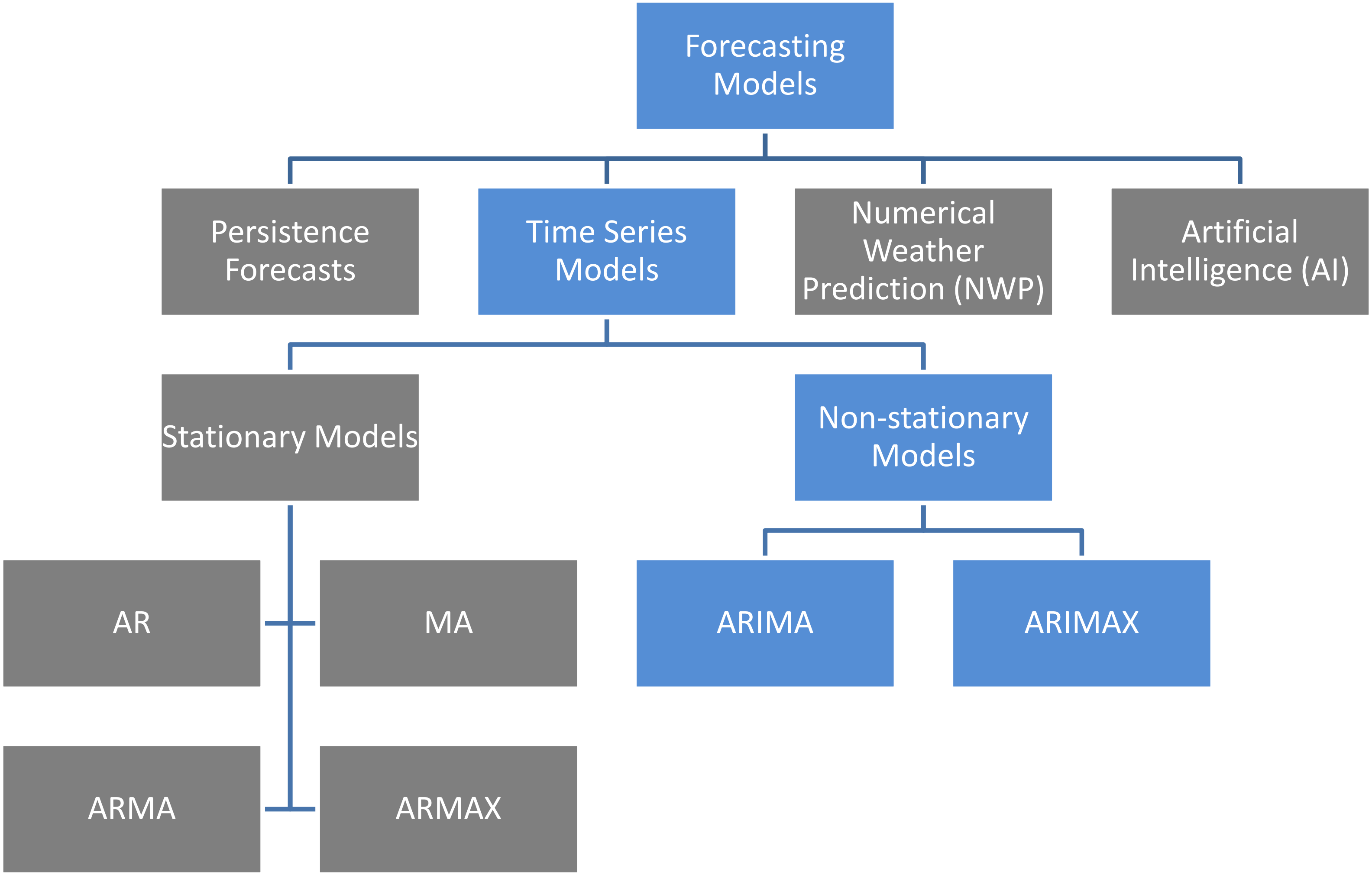
Department of Electrical Engineering

Faculty of Engineering, Chulalongkorn University

Why solar forecasting is important?

Because, **the unreliability of solar power generating** has made the entire electrical power generation difficult for power management.

Solar forecasting is a widely-common approach to deal with the problem. An improved accuracy in the forecast can provide a better management of electrical power production.



There are many forecasting models to predict the future solar irradiance. We focus on time series models.

Objective

- **To study the relevant variables** of solar irradiance forecasting.
- **To apply ARIMA** models, a **Seasonal ARIMA** models and **ARIMAX** models to forecast solar irradiance.
- **To validate results** of forecasting performance among the models using RMSE, MAE and a sample autocovarince function of the residual as model validation criterion.
- **To solve the practical issues** on data pre-processing.

Essential Elements

- Relevant variable
- Forecasting Horizon
- Forecasting Performance Evaluation Measures

Global Horizontal Irradiance (GHI) is the considered variable which is the geometric sum of Direct Normal Irradiance (DNI) and Diffuse Horizontal Irradiance (DHI).

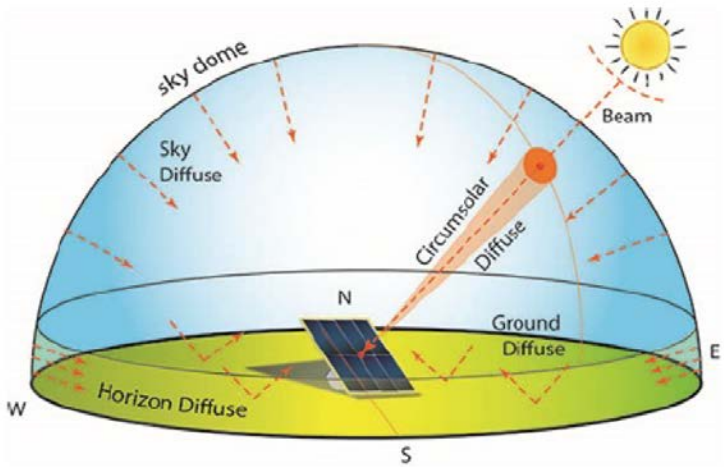


Figure 4: Solar irradiance component

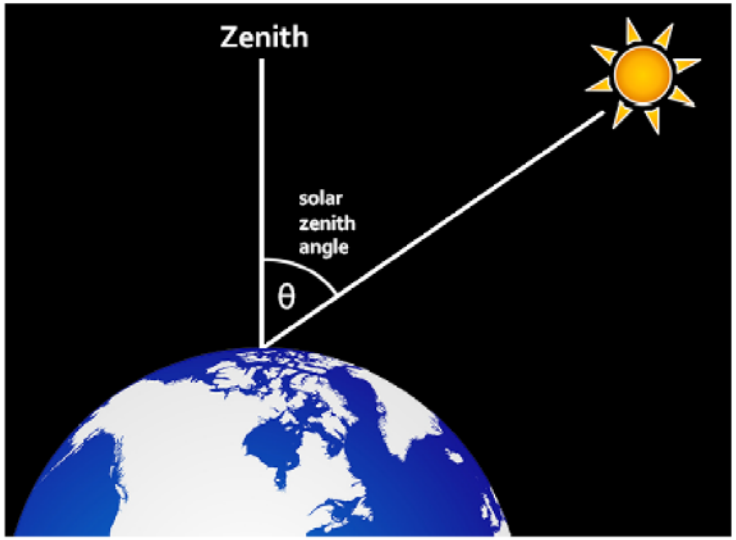


Figure 5: Solar zenith angle

This study predicted GHI and we will use the notation of $I(t)$ throughout this project.

Forecasting Horizon

- To decide how far to do prediction, forecasting horizon is the key. **Using time series model for solar forecasting is suitable and widely used for very-short-term forecasting.**
- Different forecasting horizons lead to different suitable forecasting method.

Terms of forecasting horizon.

Forecasting Horizon	Vert-short	Short	Medium	Long
Time	From a few second to 6 hours	Up to 48-72 hours ahead	Up to week ahead	Up to month or year

Forecasting Performance Evaluation Measures

Common evaluation measures are **Root Mean Squared Error (RMSE)** and **Mean Absolute Error (MAE)** which are used to validate a forecasting method. RMSE and MAE can be defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(I(t) - \hat{I}(t)\right)^2} \quad (1)$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{t=1}^{N}\left|I(t) - \hat{I}(t)\right| \quad (2)$$

where N is the length of the time horizon. **Desired value of RMSE and MAE is minimized**.

Forecasting Performance Evaluation Measures

The residual analysis can also be used for validation. The residuals is defined as

$$e(t) = I(t) - \hat{I}(t) \tag{3}$$

If $\hat{I}(t)$ can capture the dynamic in the true model, then $e(t)$ should behave like white noise. Because we do not know the true model, a sample autocovariance function (ACV) of the residual is determined. ACV of residual is defined as

$$R_e(\tau) = \frac{1}{N}\sum_{t=\tau}^{N} e(t)e(t-\tau). \tag{4}$$

Because ACV of white noise is zero after $\tau > 0$,
a desired model is the model of which $R_e(\tau)$ is small for $\tau > 0$.

Time Series Models

- ARIMAX description
- Estimation of integrated term
- Seasonal ARIMA model

ARIMAX description

An autoregressive integrated moving average model with an exogenous input (ARIMAX) is employed to predict the future solar irradiance. The model ARIMAX(*p,d,q*) is defined by

$$A(q^{-1})(1-q^{-1})^d y(t) = B(q^{-1})u(t) + C(q^{-1})v(t) \tag{5}$$

where

$$A(q^{-1}) = I - (a_1 q^{-1} + a_2 q^{-2} + \cdots + a_p q^{-p})$$
$$B(q^{-1}) = B_1 q^{-1} + B_2 q^{-2} + \cdots + B_m q^{-p}$$
$$C(q^{-1}) = I + c_1 q^{-1} + c_2 q^{-2} + \cdots + c_q q^{-p}$$

- q^{-1} is lag operator, $y(t)$ is the solar irradiance (GHI), $u(t)$ are exogenous inputs, $v(t)$ is white noise with zero mean and σ^2 variance.
- In this project, the exogenous inputs consist of the local temperature, relative humidity, wind speed, and air pressure.
- $(1-q^{-1})^d$ is integrated term where d is degree of differencing.

Estimation of integrated term

$$A(q^{-1})y(t) = C(q^{-1})v(t) \qquad (6)$$

$$A(q^{-1})(1 - q^{-1})^d y(t) = C(q^{-1})v(t) \qquad (12)$$

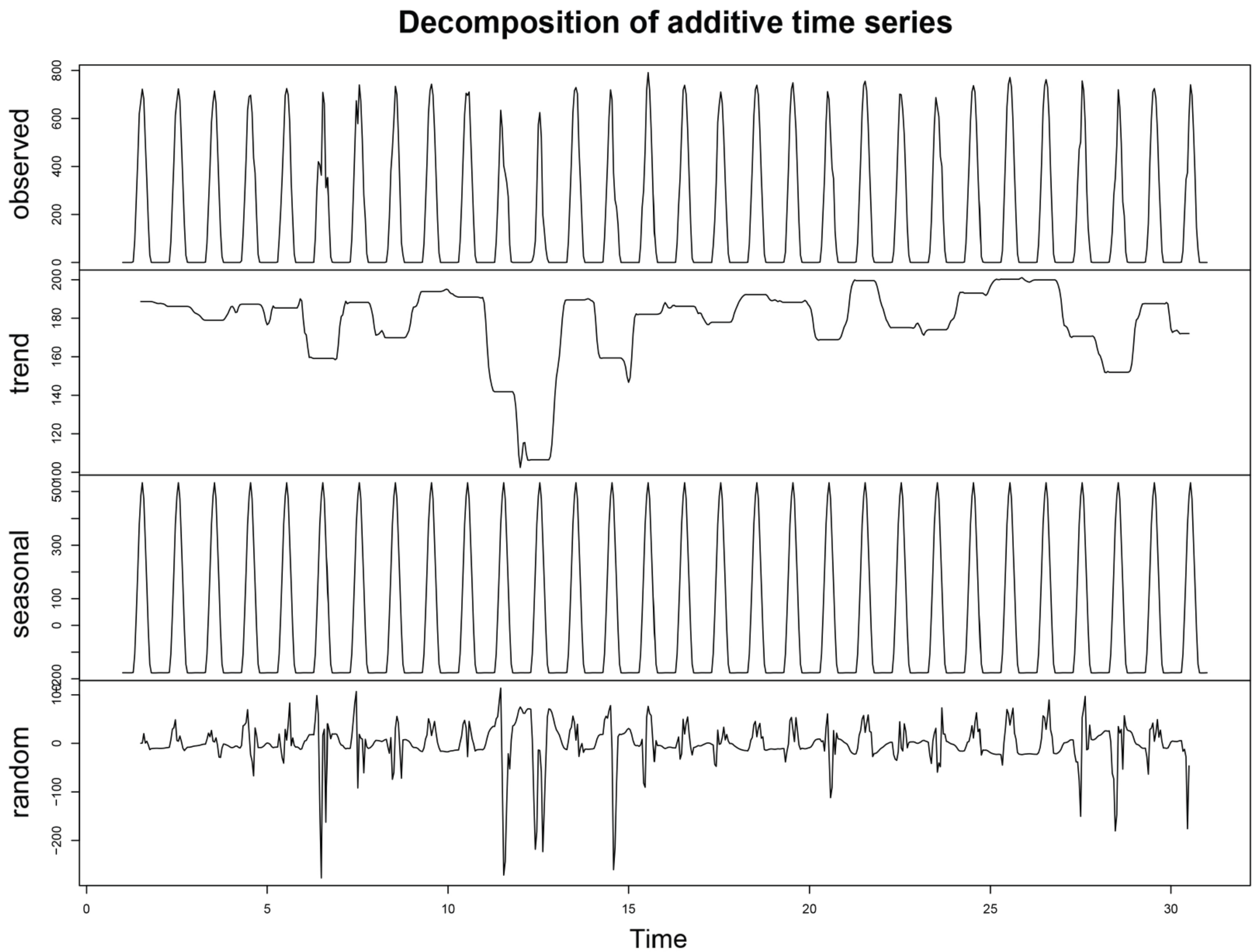
After differencing for certain order, if the resulting ACF behaves similarly to that of a stationary process in the table, then the actual process is ARIMA.

Behavior of the ACF for ARMA Models.

	AR(p)	MA(q)	ARMA(p,q)
ACF	Exponentially decay	Cuts off after lag q	Exponentially decay

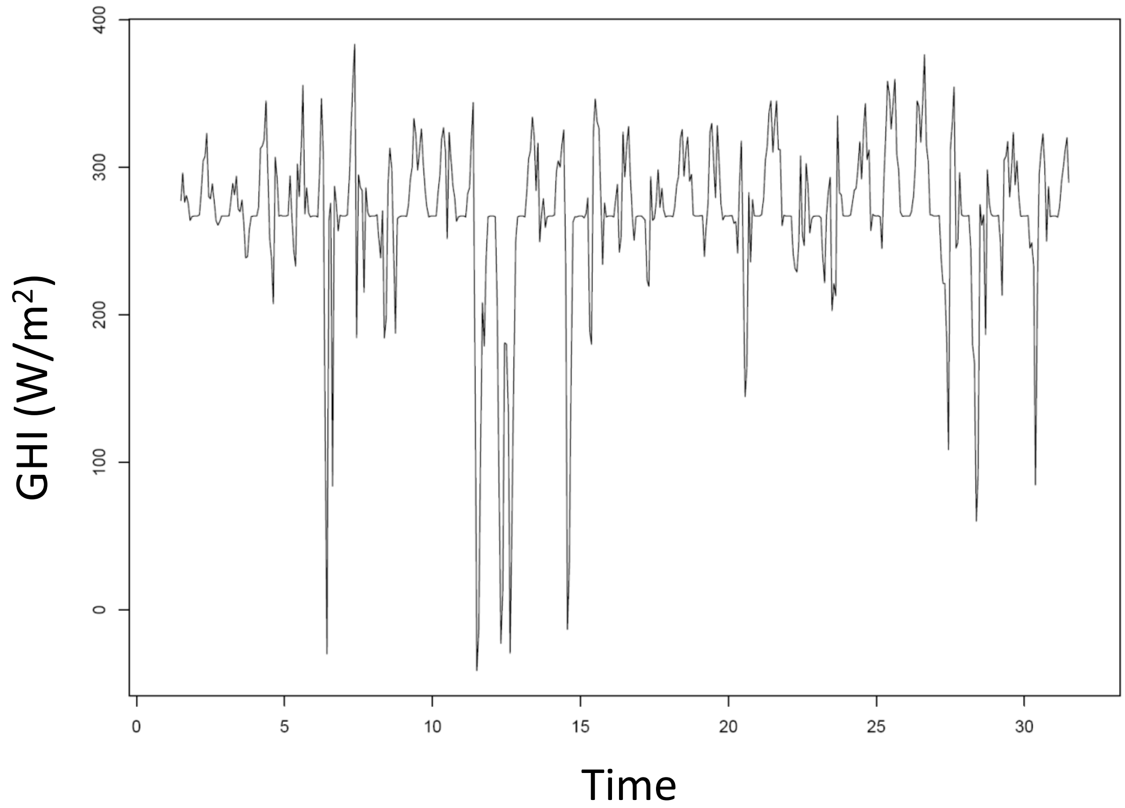
In other words, **if ACF decrease apparently slowly,** the transfer function has pole lying near or on unit circle. Hence, **the model needs more differencing.**

Seasonal ARIMA model



Decomposition of Solar Irradiance in January 2014 in Bangkok.

Seasonal ARIMA model



GHI after removing seasonal trend in January 2014 in Bangkok.

We imply that GHI can be described as ARMA models containing s season:

$$A(q^{-1})y(t) = \mathrm{s(t)} + \alpha + C(q^{-1})v(t) \tag{7}$$

Seasonal ARIMA model

In this study, **we used a seasonal ARIMA models to remove a seasonal trend.** The seasonal term and constant are removed by using this transformation. This method is called a Seasonal ARIMA $(p, d, q)(P, D, Q)_T$ models which can be defined as

$$\tilde{A}\left(q^{-T}\right)A(q^{-1})(1-q^{-T})^{D}(1-q^{-1})^{d}y(t) \\ = \tilde{C}\left(q^{-T}\right)C(q^{-1})v(t) \qquad (8)$$

where

$$\tilde{A}(q^{-1}) = I - (\tilde{a}_1 q^{-T} + \tilde{a}_2 q^{-2T} + \cdots + \tilde{a}_p q^{-pT})$$
$$\tilde{C}(q^{-1}) = I + \tilde{c}_1 q^{-T} + \tilde{c}_2 q^{-2T} + \cdots + \tilde{c}_q q^{-qT}$$

T is a seasonal period and D is integrated seasonal order.

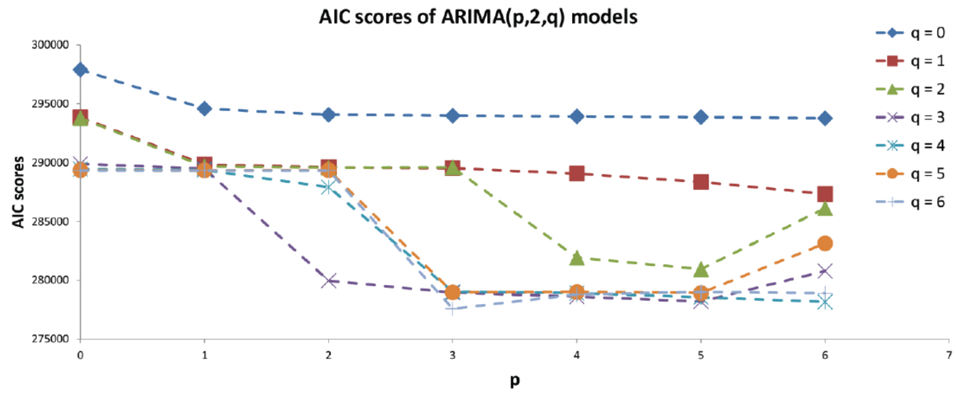
Preliminary Results

- Model Selection
- Model Validation

Model selection

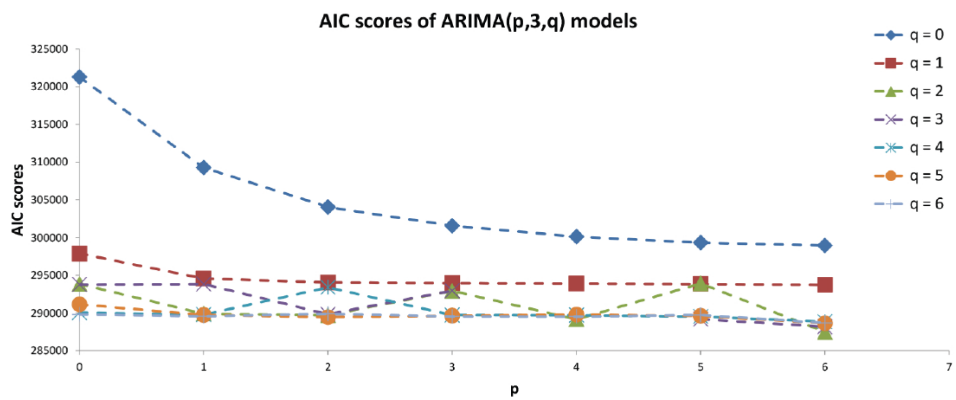
The goal of the experiment is to apply the Akaike information criterion (AIC) and Bayesian information criterion (BIC) to find the optimum order of ARIMA models and Seasonal ARIMA models. AIC and BIC are measures of the performance of the candidate models. **The best performance model gives the lowest AIC or BIC.**

AIC scores of the candidate models



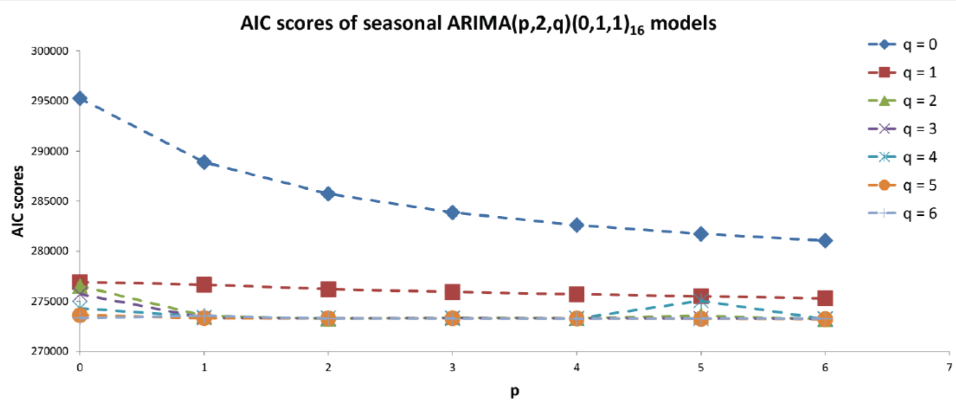
Order : ARIMA(p,2,q)

Selected order : ARIMA(3,2,6)



Order : ARIMA(p,3,q)

Selected order : ARIMA(6,3,2)



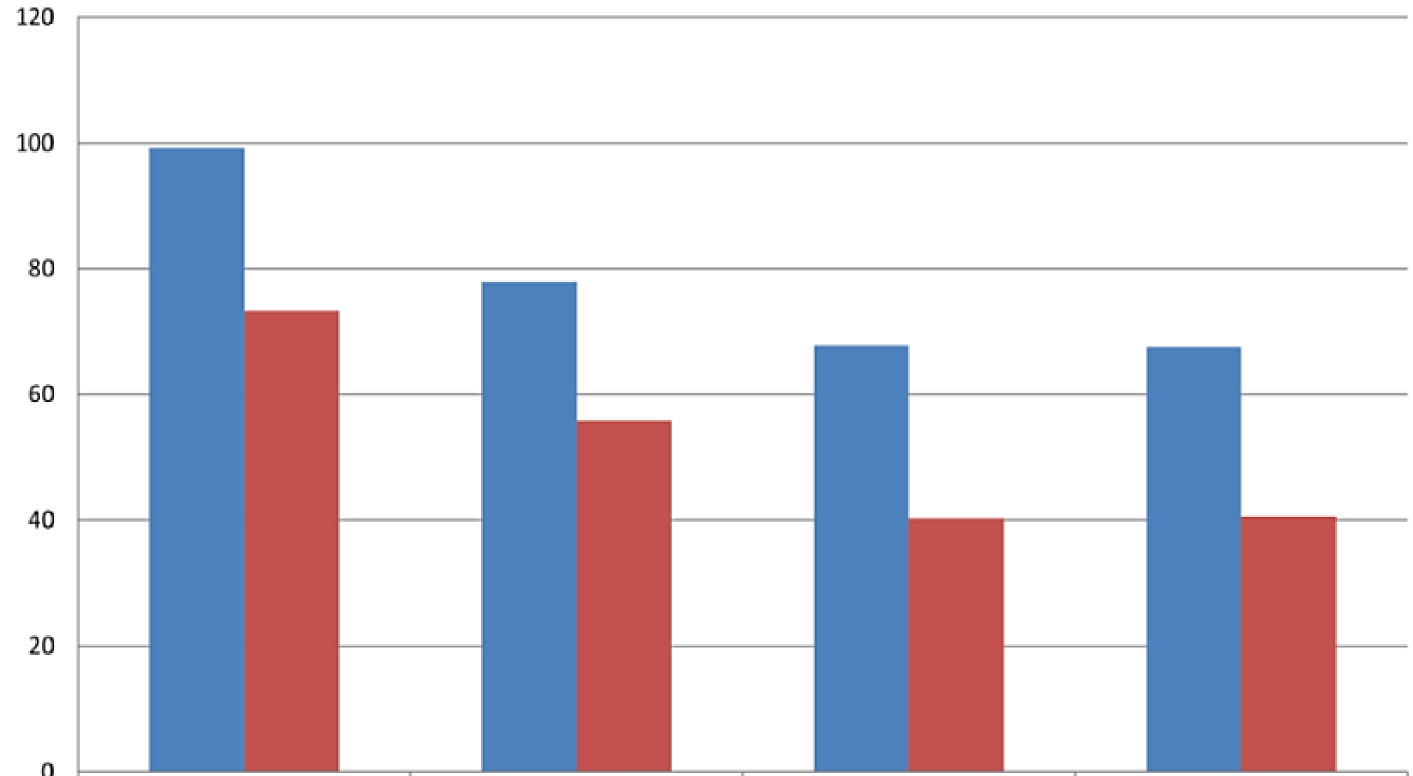
Order : Seasonal ARIMA(p,2,q)$(0,1,1)_{16}$

Selected order : Seasonal ARIMA(6,2,5)$(0,1,1)_{16}$

Model validation

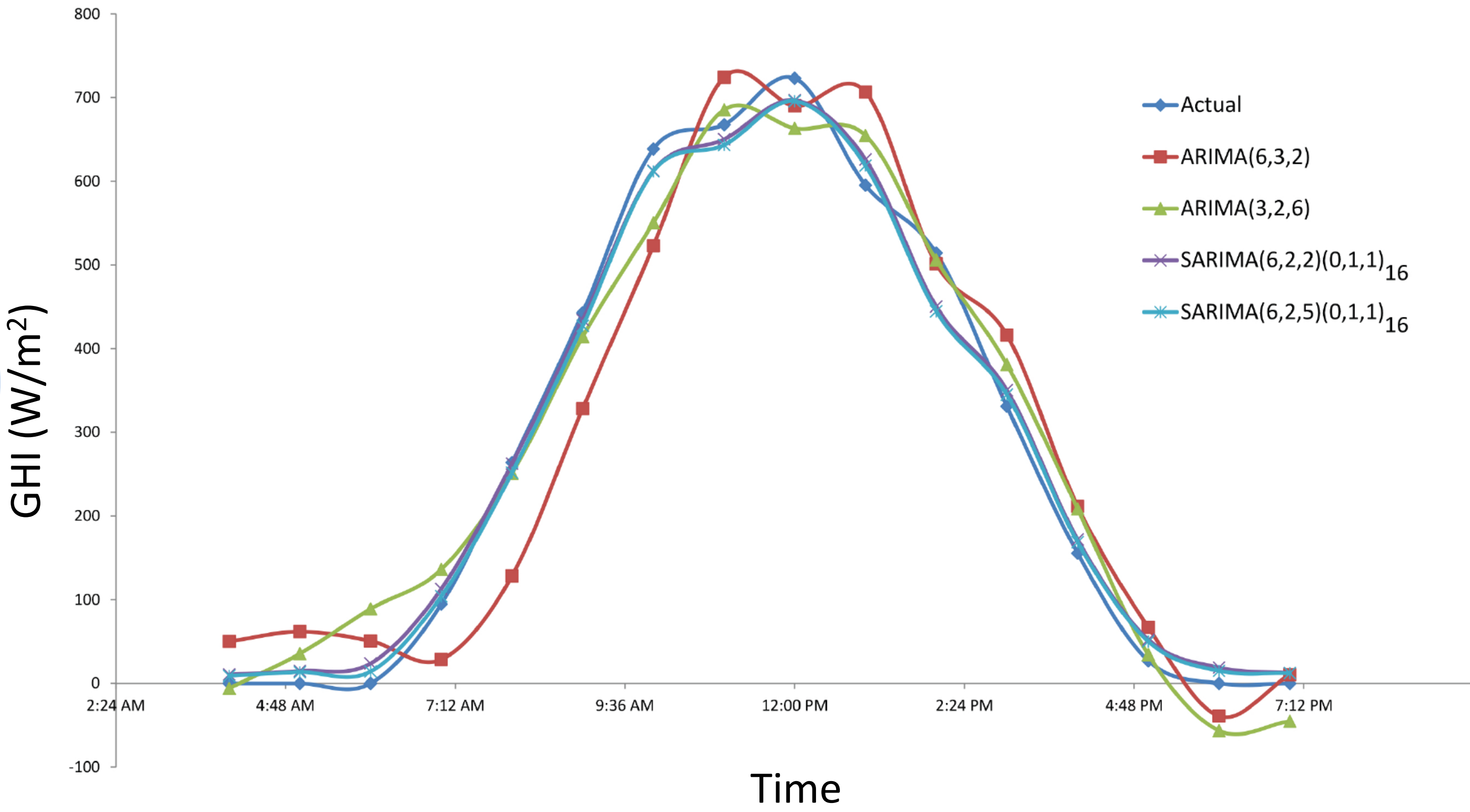
After we select the order of the candidate models, we would compare their performance on the unseen data sets. **We used RMSE and MAE to validate the performance of the candidate models.** The validation data is historical GHI data in year 2015.

Comparison among the models

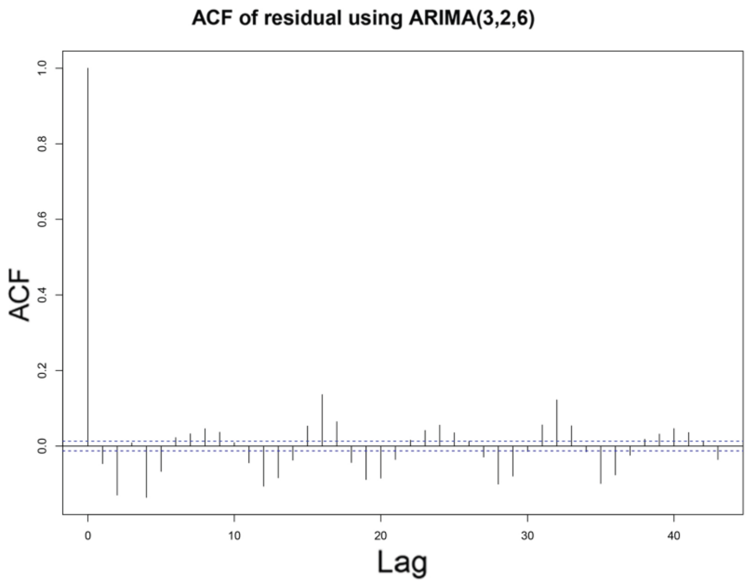


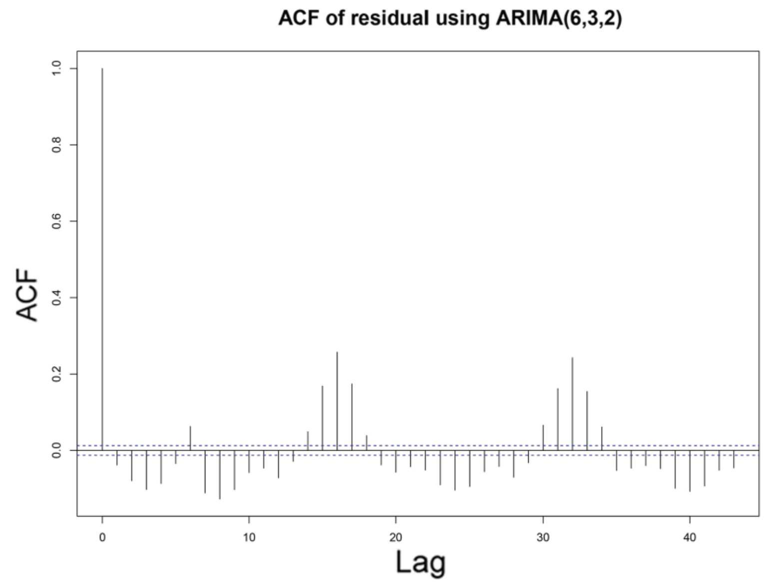
	ARIMA(6,3,2)	ARIMA(3,2,6)	SARIMA(6,2,5)(0,1,1)16	SARIMA(6,2,2)(0,1,1)16
RMSE	99.27683062	77.88499233	67.7761789	67.59103781
MAE	73.3338842	55.90355152	40.29232815	40.62253123

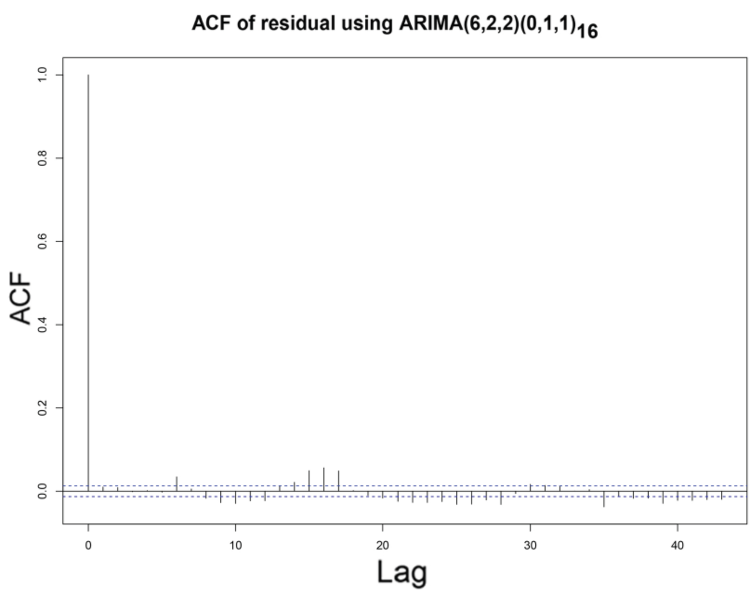
Comparison among the models

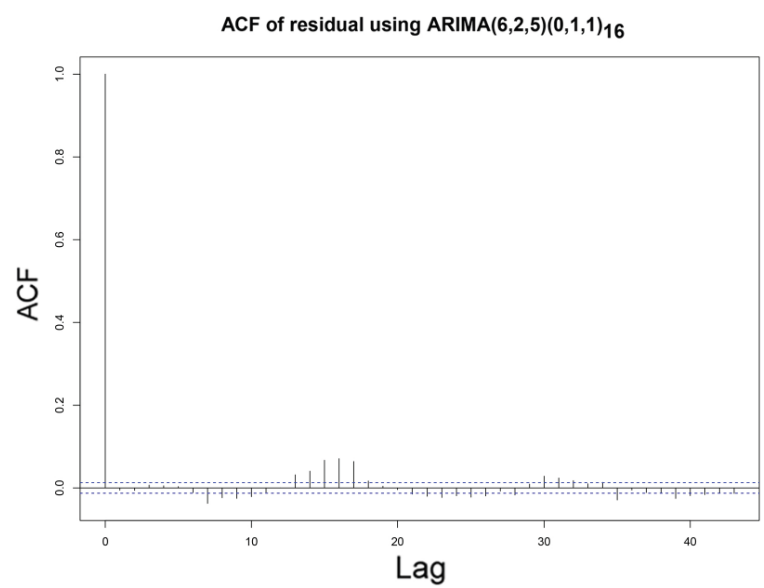


Comparison among the models









Project overview

- Scope of work
- Expected outcomes

Scope of work

- We focus on solar irradiance in the area of Chulalongkorn University location by using ARIMAX and seasonal ARIMAX models.
- The exogenous inputs consist of the local temperature, relative humidity, wind speed and air pressure will be included in the models.
- We will conduct experiments to verify our approach by using data obtained from Thai Meteorological Department (TMD).
- We will conduct an experiment by using data obtained from department of electrical engineering if they are available.
- We will study the consequence of seasons of the year in the location.

Expected outcomes

- **Comparison results of forecasting performance** among Persistence forecast, ARIMA models, a Seasonal ARIMA models and ARIMAX models using Root Mean Squared Error (RMSE), Mean Absolute Error and a sample autocovariance function (ACV) of the residual as model validation criterion.

- **Schemes for solving practical issues** on data pre-processing which are 1) missing data and 2) timely asynchronous data.

Q&A

Backup

Relevant Variables

Global Horizontal Irradiance (GHI) is the considered variable which is the geometric sum of Direct Normal Irradiance (DNI) and Diffuse Horizontal Irradiance (DHI).

$$\mathrm{GHI} = \mathrm{DHI} + \mathrm{DNI}\cos(\theta) \tag{9}$$

where θ is the solar zenith angle. The unit of GHI, DNI and DHI are W/m^2. **This study predicted GHI and we will use the notation of $I(t)$** throughout this project.

Persistence Forecasts

Persistence forecast as a baseline prediction method is used for a comparison to more advanced methods. There are many equations of persistence forecast in accordance with each study.

$$\hat{I}(t+h) = I(t) \qquad (10)$$

1) Persistence model supposes that **solar irradiance at time $t+h$ can be predicted by it value at time t.**

2) **Clearness persistence forecast** supposes that **clearness index is constant, i.e.**

$$K(t+h) = K(t),$$

where clearness index is defined as :

$$K(t) = \frac{I(t)}{I_{\mathrm{EX}}(t)}. \qquad (11)$$

$I_{\mathrm{EX}}(t)$ = $I_0 \cos\theta(t)$ where I_0 is an average value of the extraterrestrial irradiance and is suggested to be 1360 and $\theta(t)$ is solar zenith angle at time t. Thus,

$$\hat{I}(t+h) = K(t) I_{\mathrm{EX}}(t+h) \qquad (12)$$

Persistence Forecasts

2) **Clear sky persistence forecast** supposes that **clear sky index is constant, i.e.**

$$k(t+h) = k(t),$$

where clearness index is defined as :

$$k(t) = \frac{I(t)}{I_{\mathrm{clr}}(t)}. \tag{13}$$

$I_{\mathrm{clr}}(t)$ is the value from a clear sky irradiance model at time t which can be widely estimated by Ineichen model. Thus,

$$\hat{I}(t+h) = k(t) I_{\mathrm{clr}}(t+h) \tag{14}$$

Seasonal ARIMA model

An estimate of the daily seasonal effect $\hat{s}(t)$ can be obtained by averaging $\tilde{s}(t)$ which is obtained by subtracting $\widehat{m}(t)$:

$$\hat{s}(t) = \frac{1}{N+1}\sum_{k=0}^{N} \tilde{s}(t+kT)$$

where T is the seasonal period,

$$\tilde{s}(t) = y(t) - \widehat{m}(t)$$

$$\widehat{m}(t) = \frac{\frac{1}{2}y\left(t-\frac{T}{2}\right) + y\left(t-\frac{T}{2}+1\right) + \cdots + y\left(t+\frac{T}{2}-1\right) + \frac{1}{2}y\left(t+\frac{T}{2}\right)}{T}$$

Seasonal ARIMA model

In this study, we used a seasonal ARIMA models to remove a seasonal trend. Giving the following equation as an example:

$$y(t) = s(t) + \alpha + v(t) \tag{15}$$

where $s(t) = s(t - kT)$, k is an integer and α is a constant.

We subtract $y(t)$ by $y(t - T)$.

$$y(t) - \mathrm{y}(\mathrm{t} - \mathrm{T}) = s(t) + \alpha + v(t) - s(t - T) - \alpha - v(t - T)$$

$$y(t) - \mathrm{y}(\mathrm{t} - \mathrm{T}) = v(t) - v(t - T)$$

Model Selection

The AIC is defined as

$$\mathrm{AIC} = 2L + 2d \tag{16}$$

where L is the loglikelyhood function and d is the number of effective parameters.

The BIC is defined as

$$\mathrm{BIC} = 2L + d\log N \tag{17}$$

where N is the number of sample.

Autocorrelation Function

An autocorrelation function is defined by

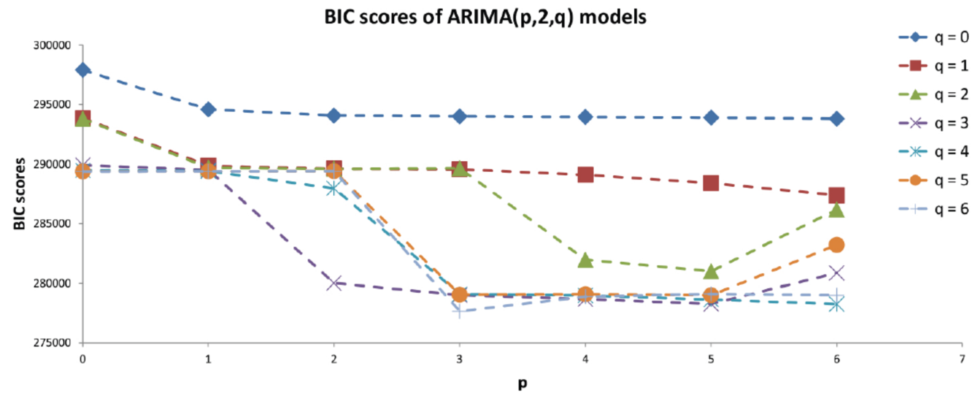
$$\mathrm{ACF} = \frac{R(\tau)}{R(0)} \tag{18}$$

where the sample autocovariance function (ACVF) is defined by

$$R(\tau) = \frac{1}{N}\sum_{t=\tau}^{N} y(t)y(t-\tau) \tag{19}$$

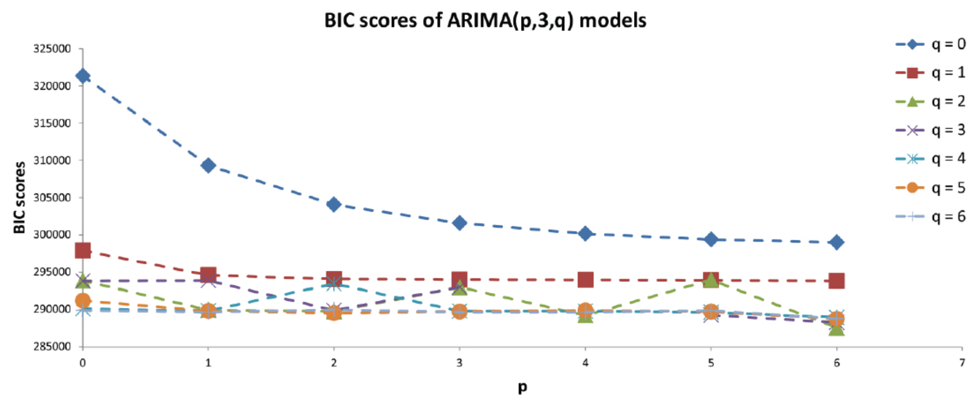
where the expectation of y is zero and the variance of y which are the same at all time t.

BIC scores of the candidate models



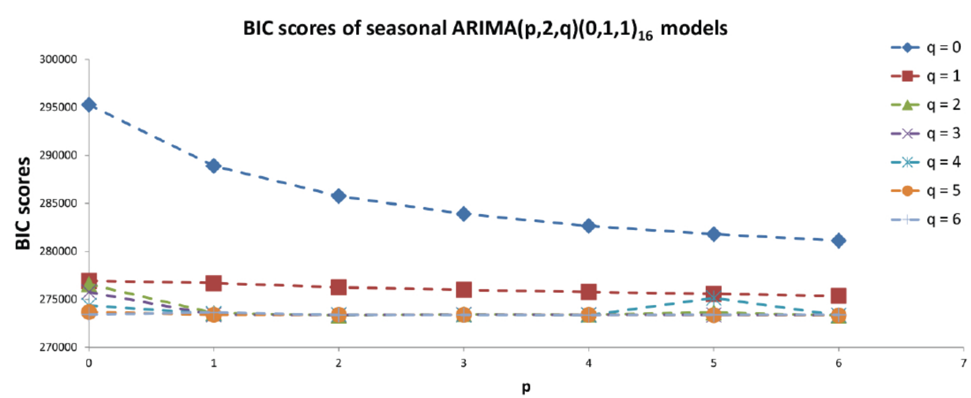
Order : ARIMA(p,2,q)

Selected order : ARIMA(3,2,6)



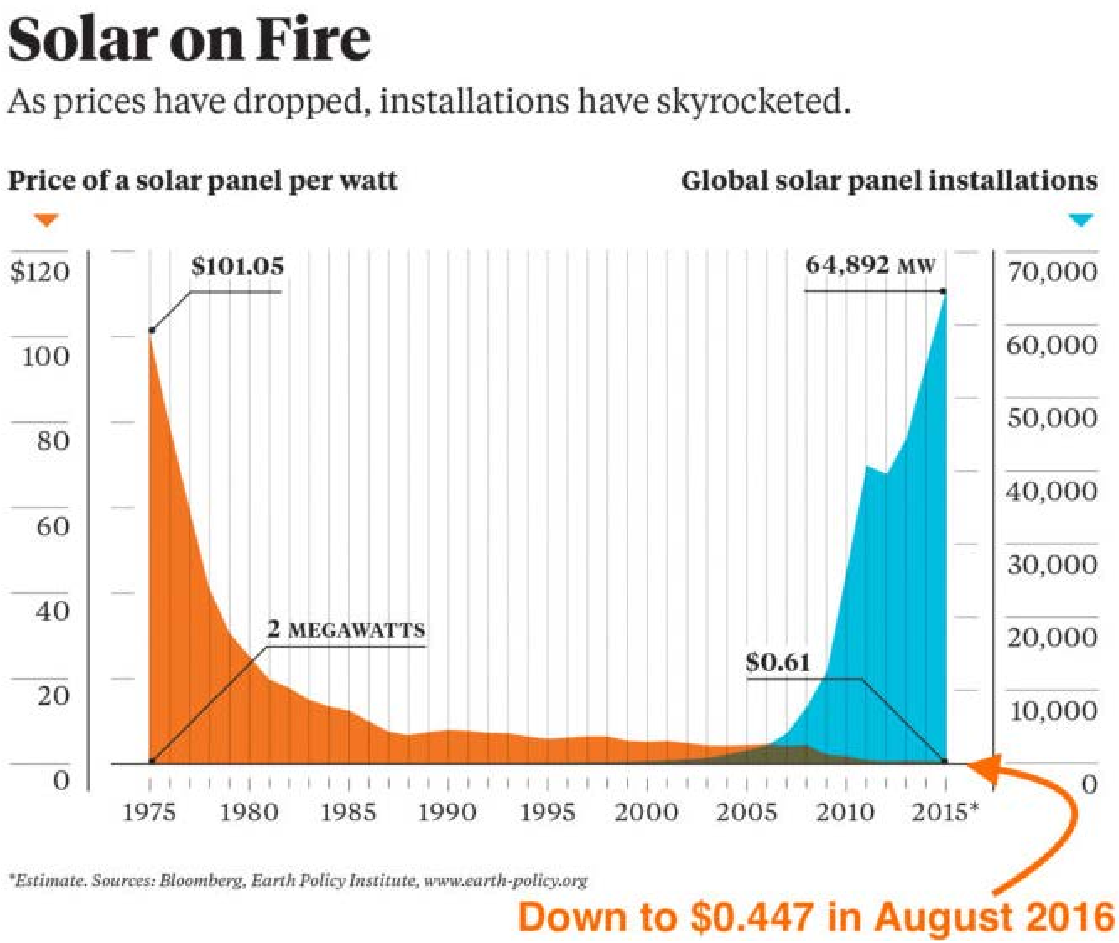
Order : ARIMA(p,3,q)

Selected order : ARIMA(6,3,2)



Order : Seasonal ARIMA(p,2,q)(0,1,1)$_{16}$

Selected order : Seasonal ARIMA(6,2,2)(0,1,1)$_{16}$



The price of a solar panel have obviously dropped from $101.05 in 1975 to $0.447 in 2016 and global solar panel installations have risen from 2 MW in 1975 to 64,892 MW in 2015.

Source: http://cleantechnica.com/2016/08/17/10-solar-energy-facts-charts-everyone-know/

Total installed renewable power plant capacity from AEDP 2015-2036 in Thailand.

Alternative Energy Development Plan during 2015-2036 (MW)								
Year	Solar	Wind	Hydro	Waste	Biomass	Biogas	Energy crops	Total
2014	1,298.5	224.5	3,048.4	65.7	2,541.8	311.5	-	7,490.4
2036	6,000	3,002	3,282.4	500	5,570	600	680	19,634.4

Source : Thailand Ministry of Energy