A comparison of intraday solar power forecasting methods

Chayanont Potawananont ID 5930084921
Sararut Pranonsatid ID 5930515021
Advisor : Assist. Prof. Jitkomut Songsiri

Department of Electrical Engineering
Faculty of Engineering
Chulalongkorn University

Outline

- Introduction
- Objectives
- Methodology
- Plan
- Preliminary Result
  - Features selection
  - Prediction model

Introduction

Introduction

Why solar forecasting is important?

- As in Alternative Energy Development Plan (AEDP) 2015, the government of Thailand plans to increase the proportion of solar energy.
- Variability in solar power has made the generation system difficult to manage.

Introduction

Very short-term forecasting (Intraday)

- Provide a better management of electrical power production
- Increases stability of electrical power systems

Introduction

There are many method used to predict the future solar power.

- AR, ARMA, ARIMA
- Artificial neural network (ANN)
- Support vector regression (SVR)
- Random Forest (RF)
- k-nearest neighbors (kNN)

Several input features:

- Previous solar power
- Previous irradiance
- Temperature

Objectives

- To study the relevant variables of intraday solar irradiance forecasting.
- To apply SVR models and RF models to forecast solar irradiance.
- To compare results of forecasting performance between SVR models and RF models.
- To compare the computational complexity between the models

Features selection : Partial Correlation

- **Partial correlation coefficient** is a measure of the strength and direction of the linear relationship between two variables after "adjusting" for linear relationships involving all the other variables

The partial covariance of Y_i, Y_j given X is defined as

$$\mathbf{cov}(Y_i, Y_j | X) = \mathbf{cov}(Y_i - \hat{Y}_i(X), Y_j - \hat{Y}_j(X))$$

Where $\hat{Y}_i(X)$ is the estimate of Y_i from X

Features selection : Partial Correlation

- **Partial correlation coefficient** is a measure of the strength and direction of the linear relationship between two variables after "adjusting" for linear relationships involving all the other variables

The partial correlation coefficient is the scaled partial covariance

$$\rho_{Y_i,Yj|X} = \frac{\mathbf{cov}(Y_i, Y_j|X)}{\sqrt{\mathbf{var}(Y_i - \hat{Y}_i(X))\,\mathbf{var}(Y_j - \hat{Y}_j(X))}}$$

Features selection : Stepwise linear regression

- Stepwise linear regression is a method of fitting linear regression models in which the choice of predictive variables is carried out by an automatic procedure.

- Stepwise linear regression is a strategy for selecting variables for prediction model

Prediction model

Support vector regression [Vapnik,1995]

The SVR is a supervised learning machine in the framework of statistical learning theory which can solve the non-linear regression

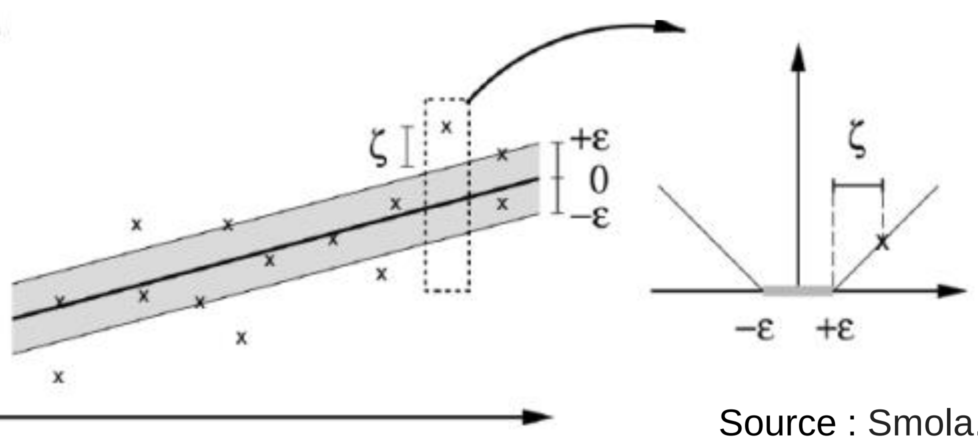
Principle of SVR is to map input data X into new space then apply linear regression with incentive loss function in the new space to find the estimate of Y from the following function :

$$f(x) = \langle w, \varphi(x) \rangle + b$$

Support vector regression [Vapnik,1995]

Incentive loss function

$$|y - f(x)|_\varepsilon = \begin{cases} |y - f(x)| - \varepsilon, & \text{if } |y - f(x)| > \varepsilon \\ 0, & \text{otherwise} \end{cases}$$



Source : Smola, Alex J., and Bernhard Schölkopf. "A tutorial on support vector regression."

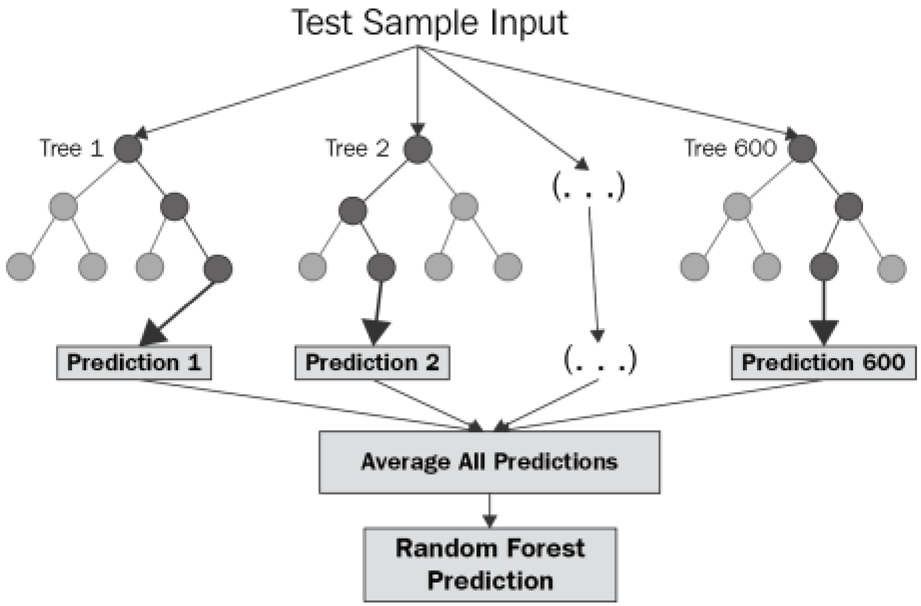
Support vector regression [Vapnik,1995]

We can find threshold value (b) and regression coefficient vector (w) by solving the constrained optimization problem as follow:

$$\begin{aligned}
\underset{w,b,u_i,v_i}{\text{minimize}} \quad & (1/2)||w||^2 + C\sum_{i=1}^{n}(u_i + v_i) \\
\text{subject to} \quad & y_i - \langle w, \varphi(x_i)\rangle - b \le \varepsilon + u_i, \quad i = 1, 2, \ldots, n, \\
& \langle w, \varphi(x_i)\rangle + b - y_i \le \varepsilon + v_i, \quad i = 1, 2, \ldots, n, \\
& u_i, v_i \ge 0 \quad , \quad i = 1, 2, \ldots, n
\end{aligned}$$

Random forest [Tin Kam Ho, 1995]

Random forest is an ensemble classifier/regressor that consists of many decision/regression trees.



Source : Afriz Chakure
Random forest regression

Regression tree

Process of building a regression tree

1. We divide the predictor space
   (the set of possible values for $X_1, X_1, \ldots, X_p$)
   into distinct and non-overlapping region, $R_1, R_2, \ldots, R_J$
2. For every observation that falls into region R_J
   We make the same prediction, which is simply the mean of the response values for the training observations in R_J

$$\text{RSS} = \sum_{j=1}^{J} \sum_{i \in R_j} ||y_i - \hat{y}_{R_j}||_2^2$$

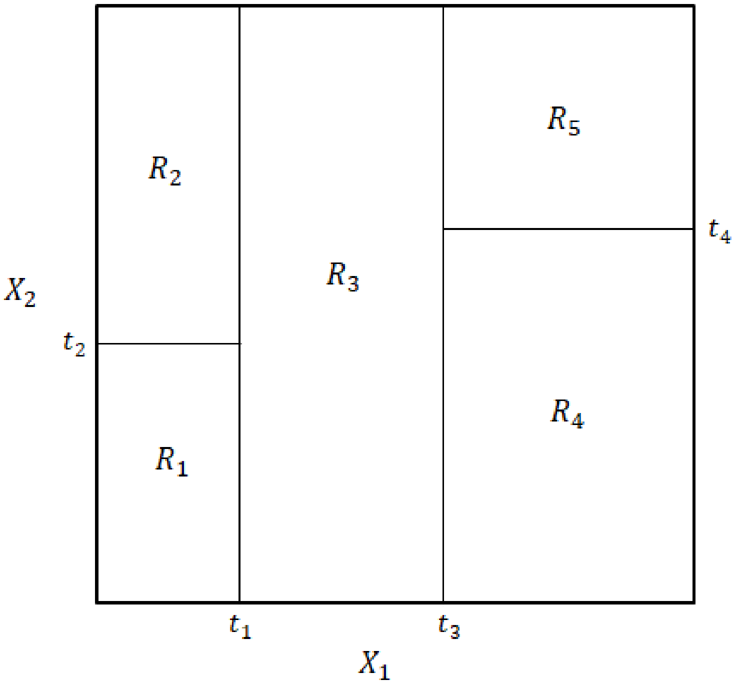
Regression tree

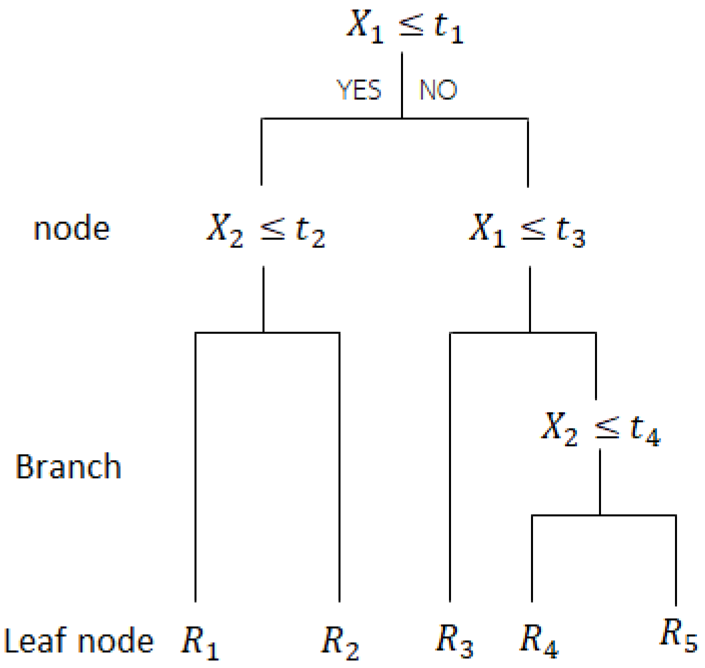
Unfortunately, it is computationally infeasible to consider
Every possible partition of the feature space into J boxes.
In practical, we use **recursive binary splitting** to split a region.

$$\mathrm{RSS} = \sum_{j=1}^{J} \sum_{i \in R_j} ||y_i - \hat{y}_{R_j}||_2^2$$

Regression tree : Recursive binary splitting

Example : 2-Dimensional predictor space



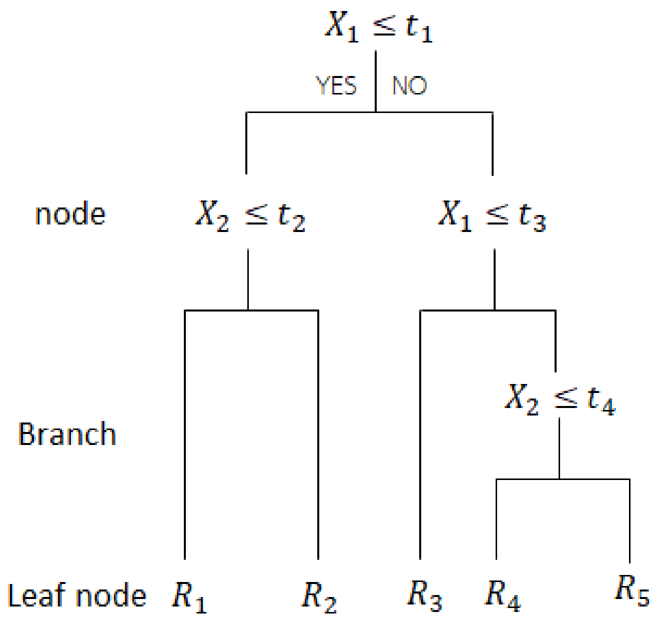


Random forest [Tin Kam Ho, 1995]

Random forest is an ensemble regressor that consists of many regression trees.

- Feature **Random** selecting

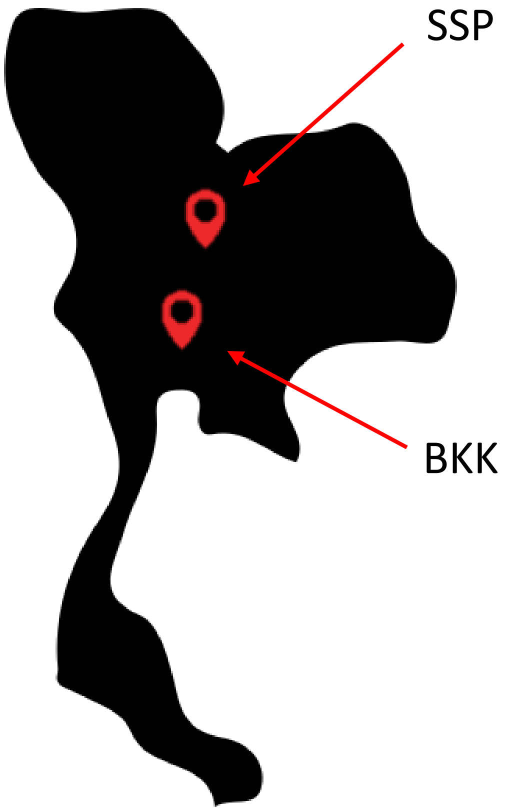
  For each node of the tree,
  randomly choose m variables ($m < p$).
  Calculate the best split based on
  these m variables in the training set.



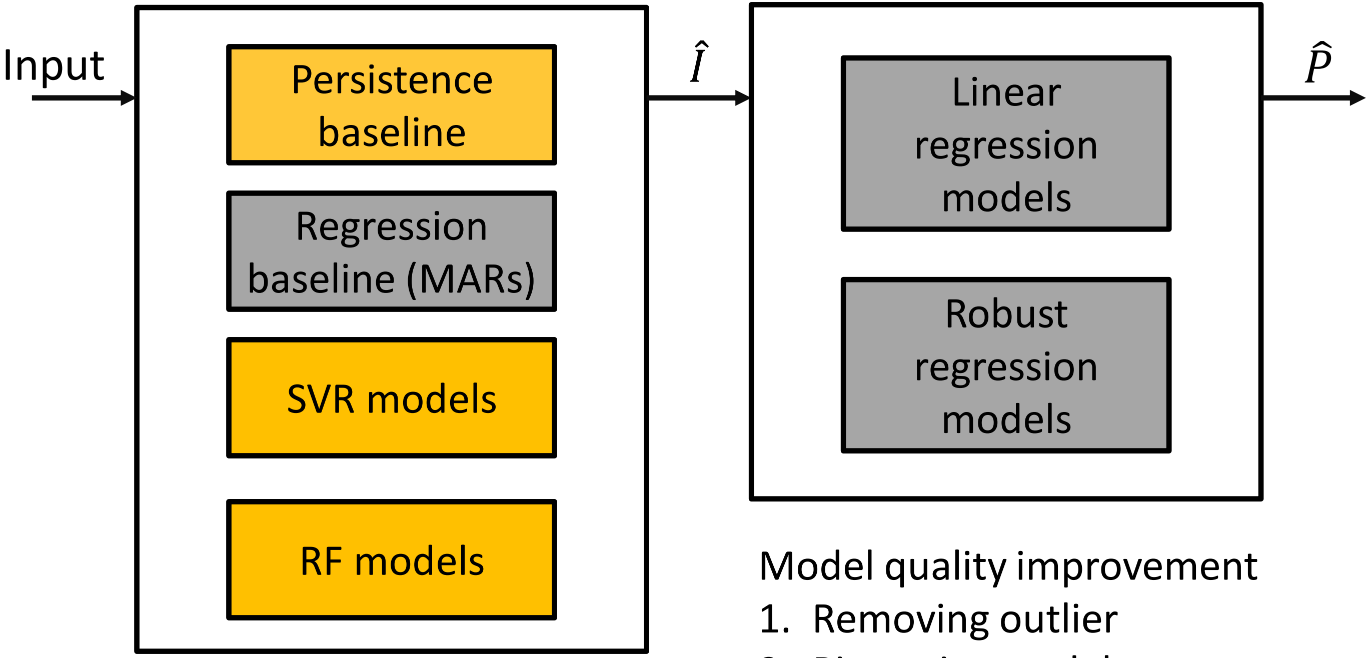
Plan

Plan

Dataset



Plan



Model quality improvement

1. Piecewise SVR,RF models
2. Feature extraction

Model quality improvement

1. Removing outlier
2. Piecewise models

Preparing model for other station (SERM solar plant)

Plan

Model comparison

1. Prediction model (Irradiance)
   - Baseline, SVR, RF, ANN
2. Power converting model (Irradiance to power)
3. Computational complexity

Preliminary result
Feature selection

Feature candidates

Target : $I(t+1)$

- Previous solar irradiance values :

  - $I(t), I(t-1), I(t-2)$, ...
  - $I^{(d-1)}(t+1)$
- Temperature : $\mathrm{T}(t)$
- Wind speed : $\mathrm{WS}(t)$
- Relative humidity : $\mathrm{RH}(t)$
- Ultraviolet Index : $\mathrm{UV}(t)$
- Cosine of solar zenith angle : $\cos(\theta(t+1))$

Data set : EECU data 2017- 2018

Feature selection : Result

Features that have significant relationship with : $I(t+1)$

- Previous solar irradiance values :
  - $I(t), I(t-3), I(t-5), I(t-6), I(t-7)$
  - $I^{(d-1)}(t+1)$
- Relative humidity : $\mathrm{RH}(t)$
- Ultraviolet Index : $\mathrm{UV}(t)$
- Cosine of solar zenith angle : $\cos(\theta(t+1))$

Preliminary result
Prediction model

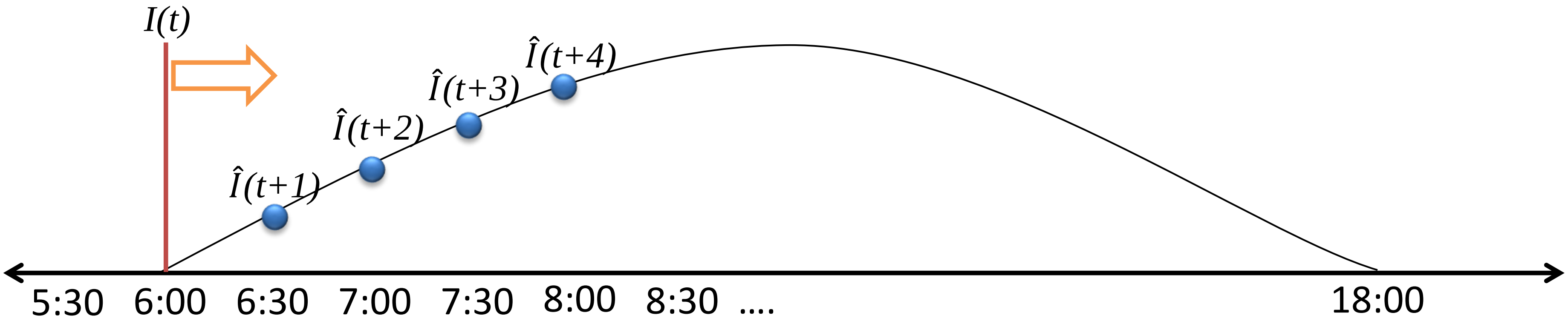
Experiment setting

Goal: predict solar power with the horizon of 4 hours every 30 minutes

$$\hat{I}(t+1),\ \hat{I}(t+2),\ \ldots.,\ \hat{I}(t+8)$$

Time of forecast values : 6:00 - 18:00: Forecast every 30 min

Execution time : 5:30 - 17:30



Data augmentation

Data set : EECU data 2017- 2018

Training set (80%) :

- feature selection
- fitting model

Validation set (10%) :

- hyper-parameter tuning

Testing set (10%) :

- testing prediction model

Forecasting Performance Evaluation Measures

1. Root Mean Square Error (RMSE)

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

2. Mean Bias Error (MBE)

$$\text{MBE} = \frac{1}{N}\sum_{t=1}^{N}\left(\hat{x}(t) - x(t)\right)$$

Prediction model : Support vector regression

Data-preprocessing :

Standardize the “input” features by removing the mean and scaling by standard deviation of the training samples

Experiment setting for SVR :

Train 8 models separately to forecast $\hat{I}(t+n)$; n = 1,2,...,8

Input features :

- $I(t), I(t\text{-}1), I(t\text{-}2)$, ..., $I(t\text{-}7)$,$I^{(d\text{-}1)}(t+n)$
- $\cos(\theta(t+n))$,

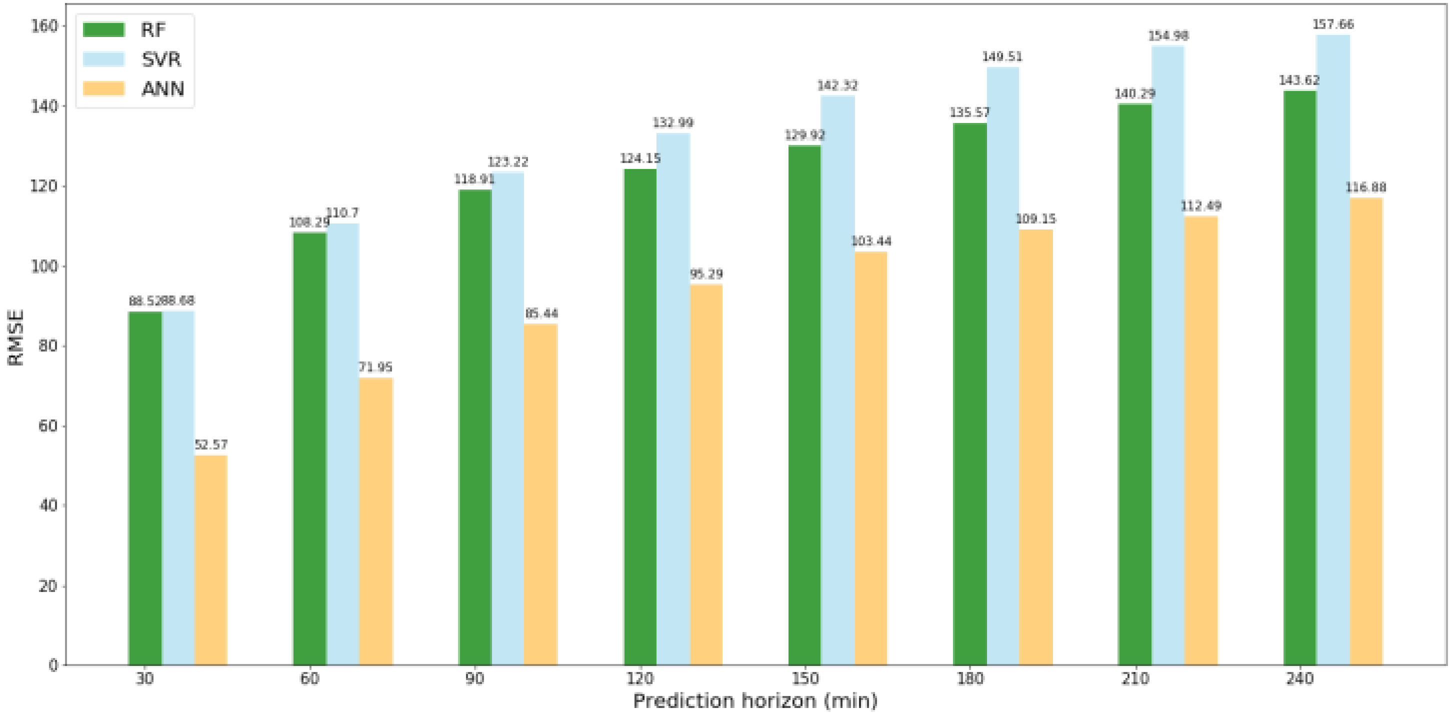
Prediction model : Random Forest

Target : $\hat{I}(t+1), \hat{I}(t+2), \ldots, \hat{I}(t+8)$

Input

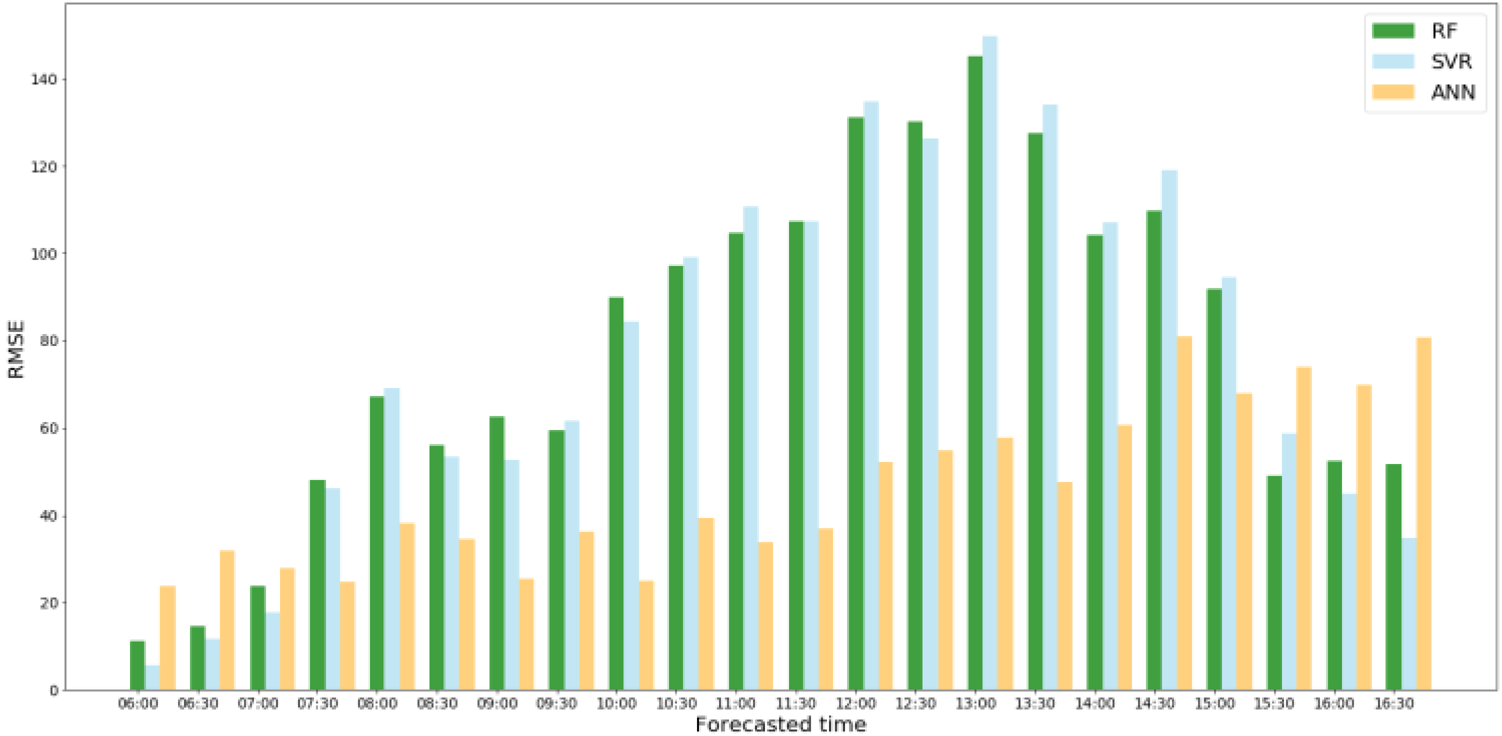
- $I(t), I(t-1), I(t-2), \ldots, I(t-7)$
- $I^{(d-1)}(t+1), I^{(d-1)}(t+2), \ldots, I^{(d-1)}(t+8)$
- $\cos(\theta(t+1)), \cos(\theta(t+2)), \ldots, \cos(\theta(t+8))$
- Hour stamp : $\mathrm{HR}(t)$

Prediction model



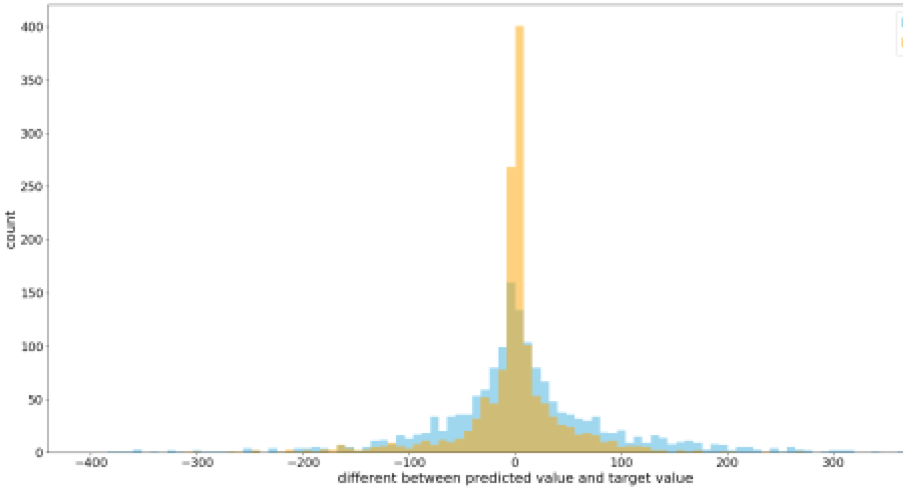
RMSE of solar irradiance forecast in each prediction horizon

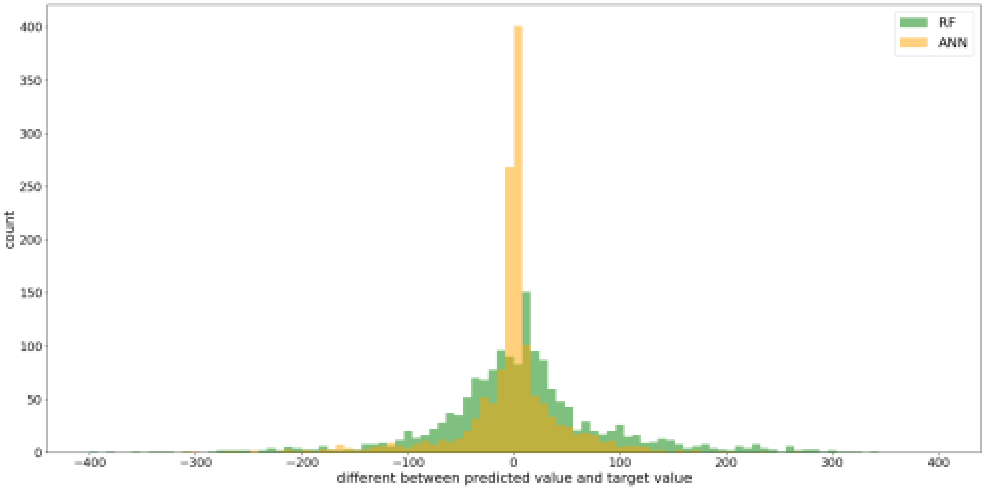
Prediction model



RMSE of 1-step solar irradiance forecast in each prediction time

Prediction model





Histogram of error distribution in each models

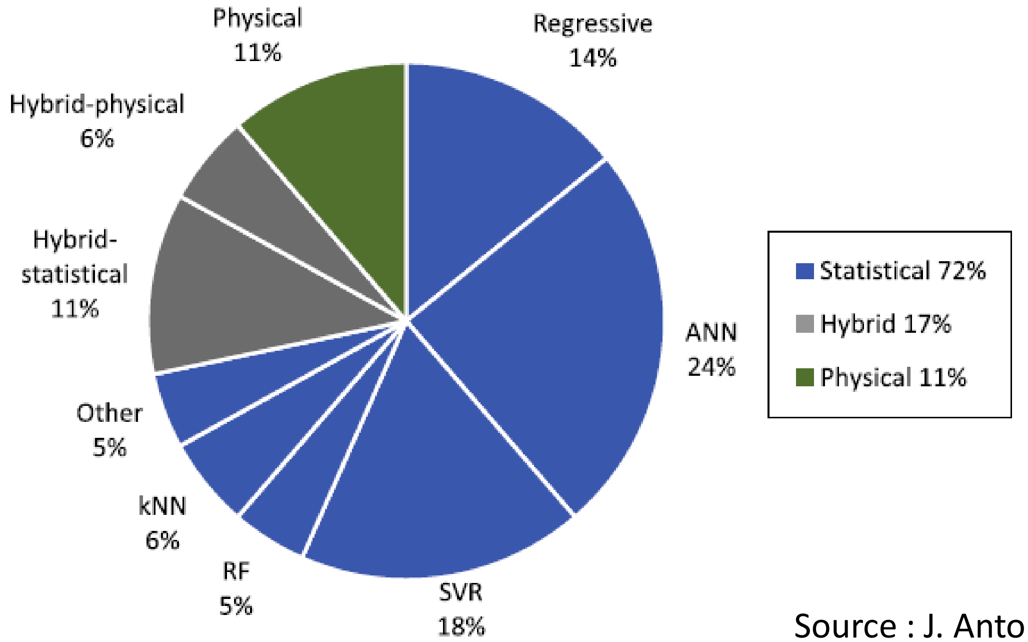
Q/A

Thank you

Backup slide

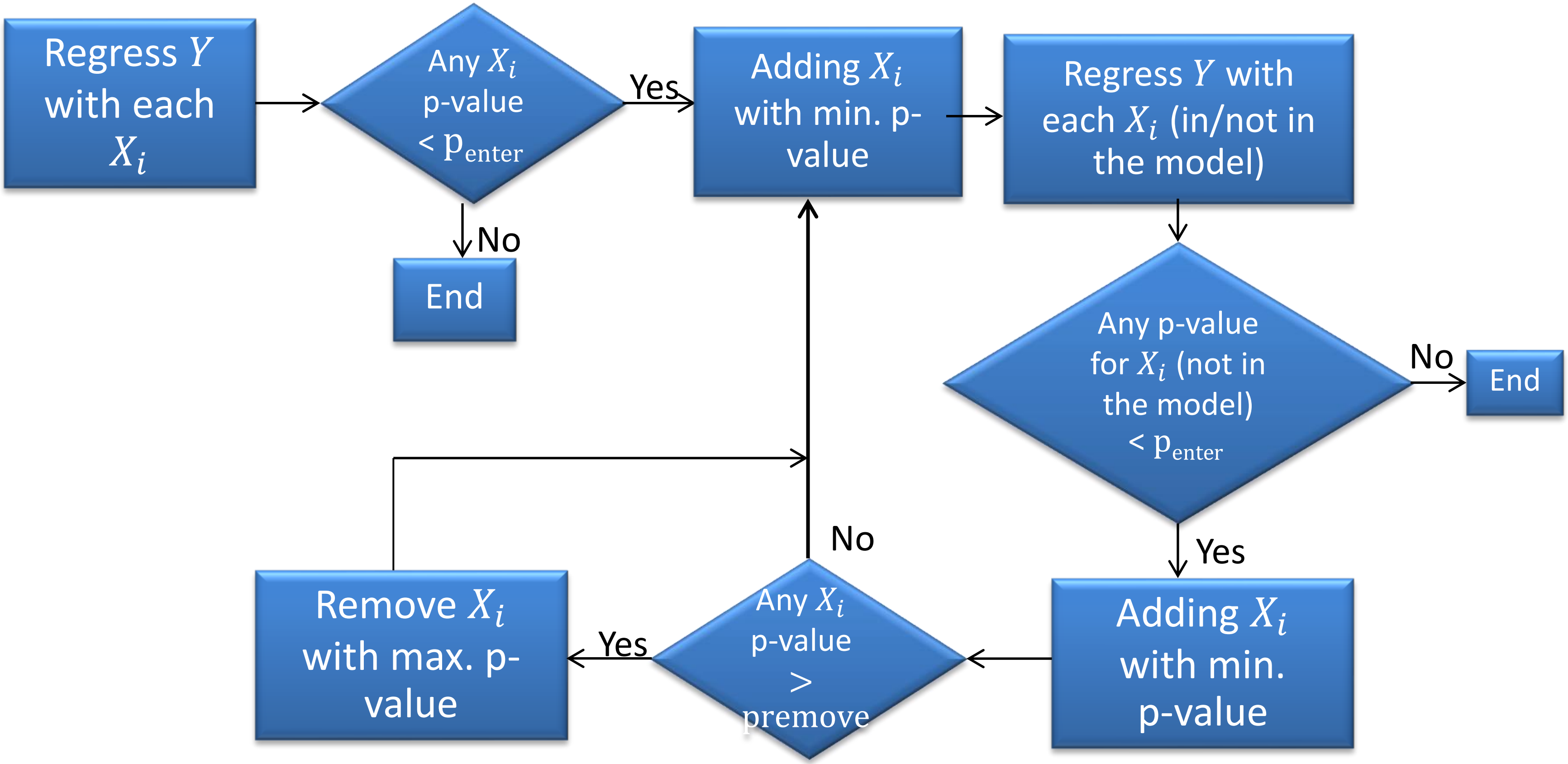
Introduction

Distribution of studies with respect to the technique used



Source : J. Antonanzas (2016) :
Review of photovoltaic power forecasting

Features selection : Stepwise linear regression



Matlab : Stepwiselm function (1)

stepwiselm

Fit linear regression model using stepwise regression

Syntax

```
mdl = stepwiselm(tbl)
mdl = stepwiselm(X,y)
mdl = stepwiselm(___,modelspec)
mdl = stepwiselm(___,Name,Value)
```

Matlab : Stepwiselm function (2)

'Criterion' — Criterion to add or remove terms

'sse' (default) | 'aic' | 'bic' | 'rsquared' | 'adjrsquared'

Criterion to add or remove terms, specified as the comma-separated pair consisting of 'Criterion' and one of these values:

- 'sse' — *p*-value for an *F*-test of the change in the sum of squared error that results from adding or removing the term
- 'aic' — Change in the value of Akaike information criterion (AIC)
- 'bic' — Change in the value of Bayesian information criterion (BIC)
- 'rsquared' — Increase in the value of R^2
- 'adjrsquared' — Increase in the value of adjusted R^2

Example: 'Criterion','bic'

Support vector regression [Vapnik,1995]

After we solve dual form of those constrained optimization problem, the estimate function is correspondingly converted into :

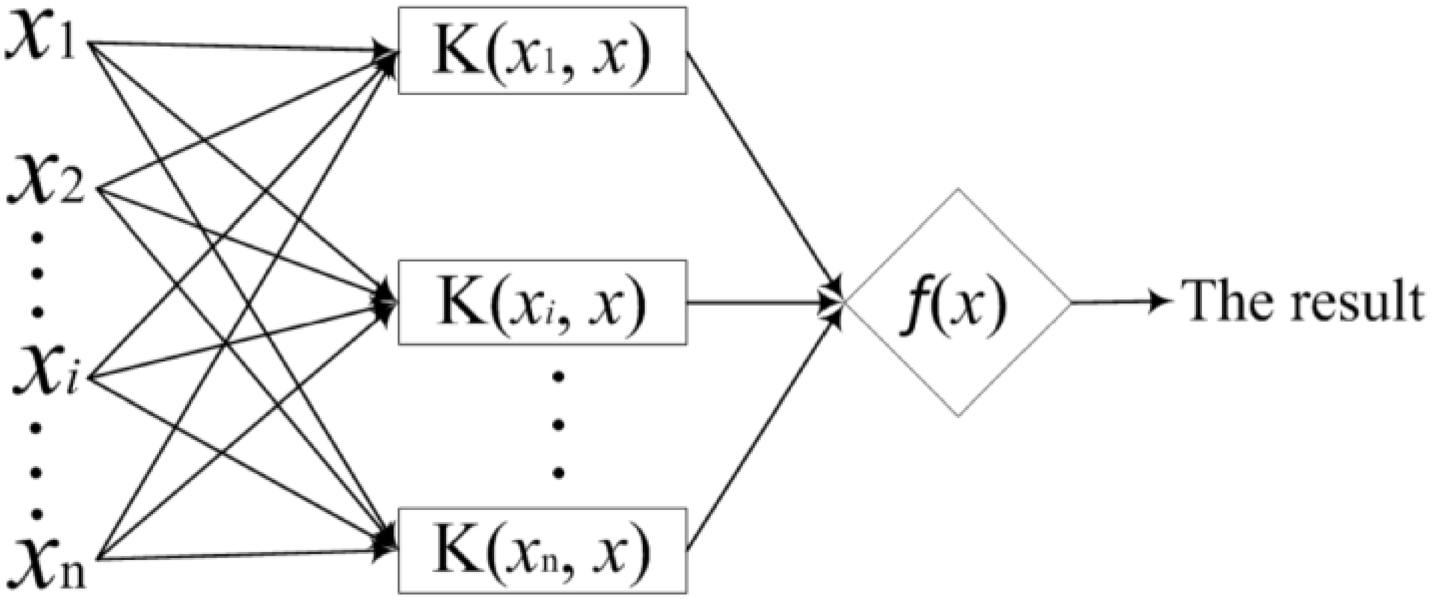
$$f(x) = \sum_{i=1}^{n} (\lambda_i - \nu_i) k(x_i, x) + b$$

Where $k(x_i, x)$ is the kernel function which can be any of the following :

1. Linear kernel : $k(x, x') = \langle x, x' \rangle$
2. Polynomial kernel : $k(x, x') = (\gamma \langle x, x' \rangle + r)^d$
3. RBF kernel : $k(x, x') = \exp(-\gamma \|x - x'\|^2)$
4. Sigmoid kernel : $k(x, x') = \tanh(\gamma \langle x, x' \rangle + r)$

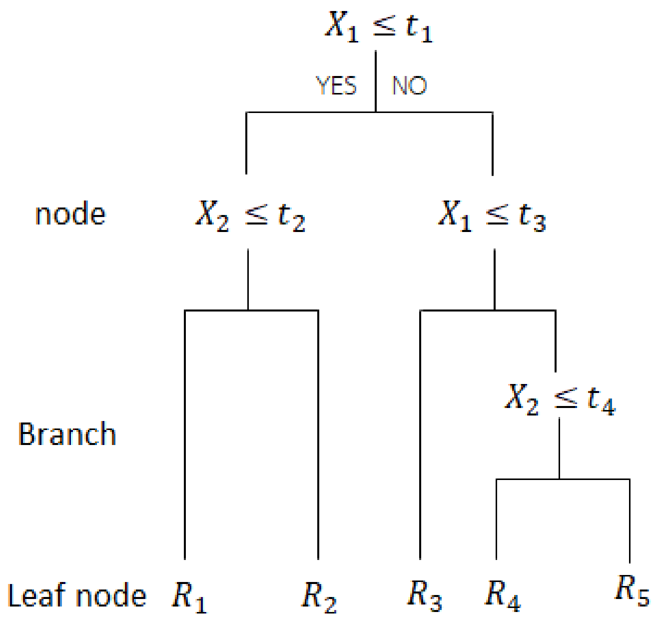
Support vector regression [Vapnik,1995]

The architecture of SVR



Random forest : Hyper parameter

- The number of tree in the forest
- The number of features (m)
- The maximum depth of the tree
- The minimum number of samples required to split



Feature selection : Correlation

Result from correlation coefficient

Correlation	$I(t+1)$	ระดับนัยสำคัญ (p-value)
$I(t)$	0.8953	0
$I(t-1)$	0.7789	0
$I(t-2)$	0.6478	0
$I(t-3)$	0.5018	0
$I(t-4)$	0.3610	0
$I(t-5)$	0.2260	10^{-155}
$I(t-6)$	0.1039	10^{-33}
$I(t-7)$	-0.0059	0.4955
$I^{(d-1)}(t+1)$	0.07369	0
$T(t)$	0.4290	0
$RH(t)$	-1.1291	10^{-51}
$UV(t)$	0.8540	0
$WS(t)$	0.1388	10^{-59}
$\cos(\theta(t+1))$	0.7810	0

Feature selection : Partial correlation

Result from partial correlation coefficient

Partial correlation	$I(t+1)$	ระดับนัยสำคัญ (p-value)
$I(t)$	0.4366	0
$I(t-1)$	-0.0014	0.8675
$I(t-2)$	0.0101	0.2397
$I(t-3)$	0.0172	0.0466
$I(t-4)$	-0.0099	0.2533
$I(t-5)$	-0.0344	0.0001
$I(t-6)$	-0.0202	0.0192
$I(t-7)$	-0.0720	0
$I^{(d-1)}(t+1)$	0.1021	0
$T(t)$	0.0035	0.6825
$RH(t)$	-0.0638	0
$UV(t)$	0.1090	0
$WS(t)$	-0.0088	0.3090
$\cos(\theta(t+1))$	0.0910	0

Feature selection : Stepwise linear regression

Result from Stepwise linear regression

variable	coefficient	p-value
constant term	118.72	1.1089×10^{-34}
$I(t)$	0.6574	0
$I(t-3)$	-0.019084	1.443×10^{-2}
$I(t-5)$	-0.050983	2.3593×10^{-7}
$I(t-6)$	-0.028777	1.1015×10^{-2}
$I(t-7)$	-0.083408	7.7209×10^{-21}
$I^{(d-1)}(t+1)$	0.087565	1.1994×10^{-50}
$RH(t)$	-1.2015	4.3465×10^{-17}
$UV(t)$	1.4957	4.9252×10^{-39}
$\cos(\theta(t+1))$	109.78	1.4576×10^{-75}

Non-selected variables are $I(t-1), I(t-2), I(t-4), WS(t), T(t)$