Hourly Probabilistic Solar Power Forecasts

Mohamed Abuella
Prof. Badrul Chowdhury

Energy Production and Infrastructure Center
Department of Electrical and Computer Engineering
University of North Carolina at Charlotte

UNC CHARLOTTE
The WILLIAM STATES LEE COLLEGE *of* ENGINEERING

September 18, 2017

Presentation Outline

- Overview of Solar Power Forecasting
- Hourly Probabilistic Forecasting of Solar Power

Variable Generations (V.G.) Forecasting

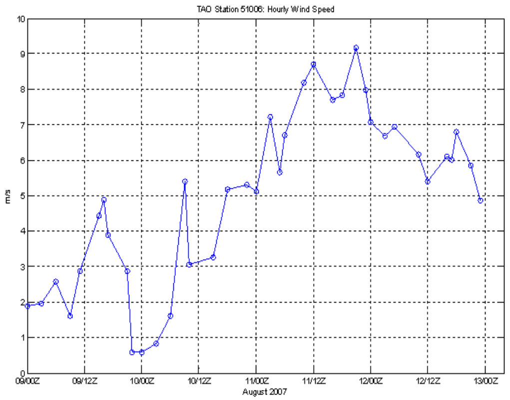
Why Forecast?

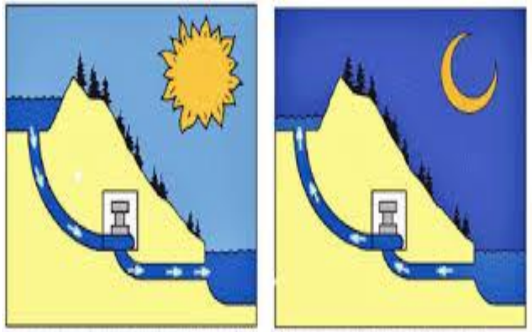
$$P_G = P_D + P_{loss}$$

Renewables Generations (Wind and Solar) are Too Variable

High Efficiency and Large Energy Storage Still not Exist

Reducing Cost and Pollution



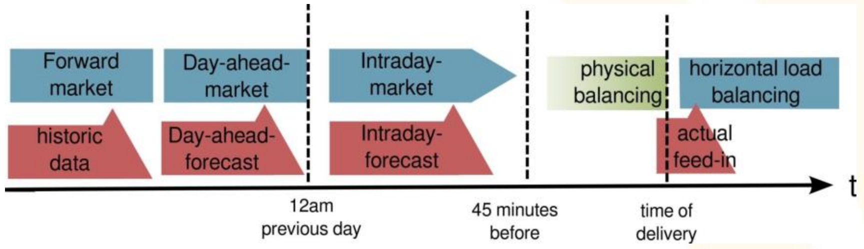


Variable Generations (V.G.) Forecasting

Where Forecast?

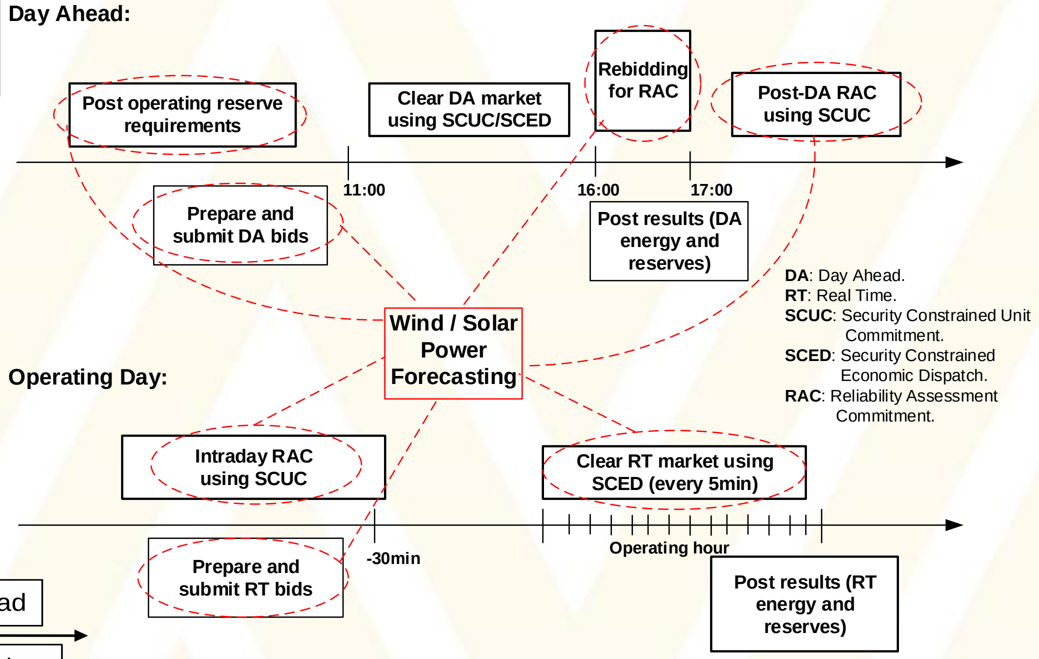
Forecast horizons, forecasting models and the related applications of VG forecasts

Forecast Horizons



	Intra-hour	Intra-day	Day ahead
Forecasting horizon	15 min to 2 h	1 h to 6 h	1 day to 3 day
Granularity-Time step	30 s to 5 min	hourly	hourly
Related to	Ramping events, variability related to operations	Load following forecasting	Unit commitment, transmission scheduling, day ahead markets

In Electricity Markets Operation: **You Plan What You Forecast**

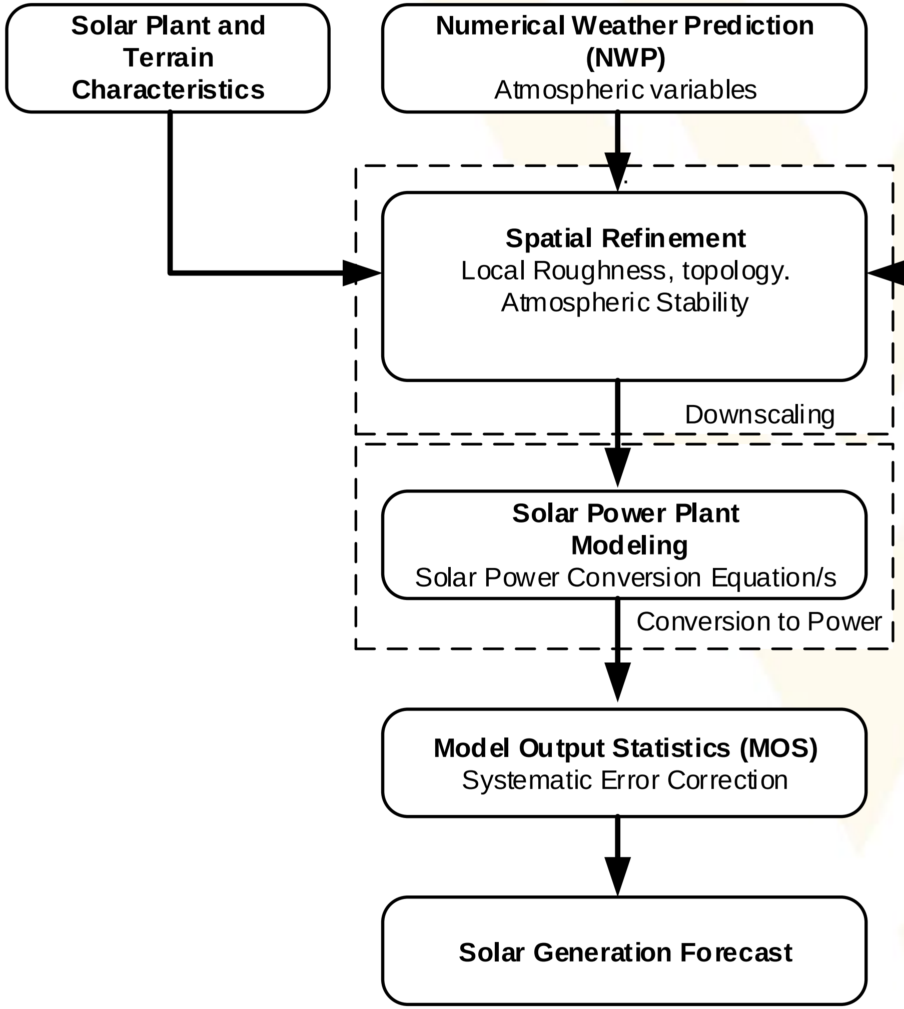


VG forecasting in US. electric utilities and ISO, such as CAISO, ERCOT, MISO, ISO-NE, NYISO,…etc.

Botterud, J. Wang, V. Miranda, and R. J. Bessa, “Wind power forecasting in US electricity markets,” The Electricity Journal, vol. 23, no. 3, pp. 71–82, 2010.
Elke Lorenz, “Solar Resource Forecasting” International Solar Energy Society (ISES) Webinar, 2016.
Voyant, C., Notton, G., Kalogirou, S., Nivet, M. L., Paoli, C., Motte, F., & Fouilloy, A. (2017). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, *105*, 569-582.

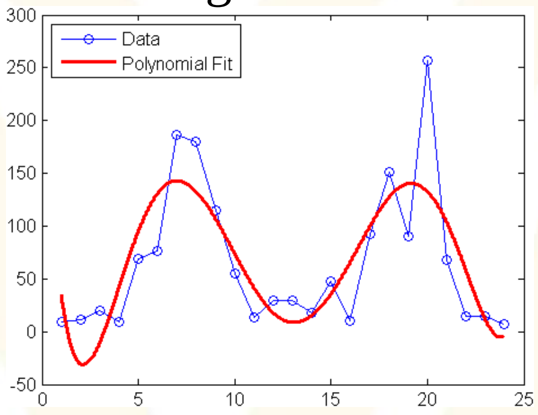
Solar Power Forecasting

How Forecast?

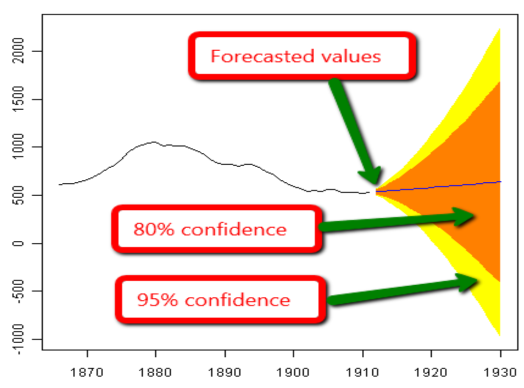


Flowchart of the Combined Approach (Physical and Statistical) of the Solar Power Forecasting

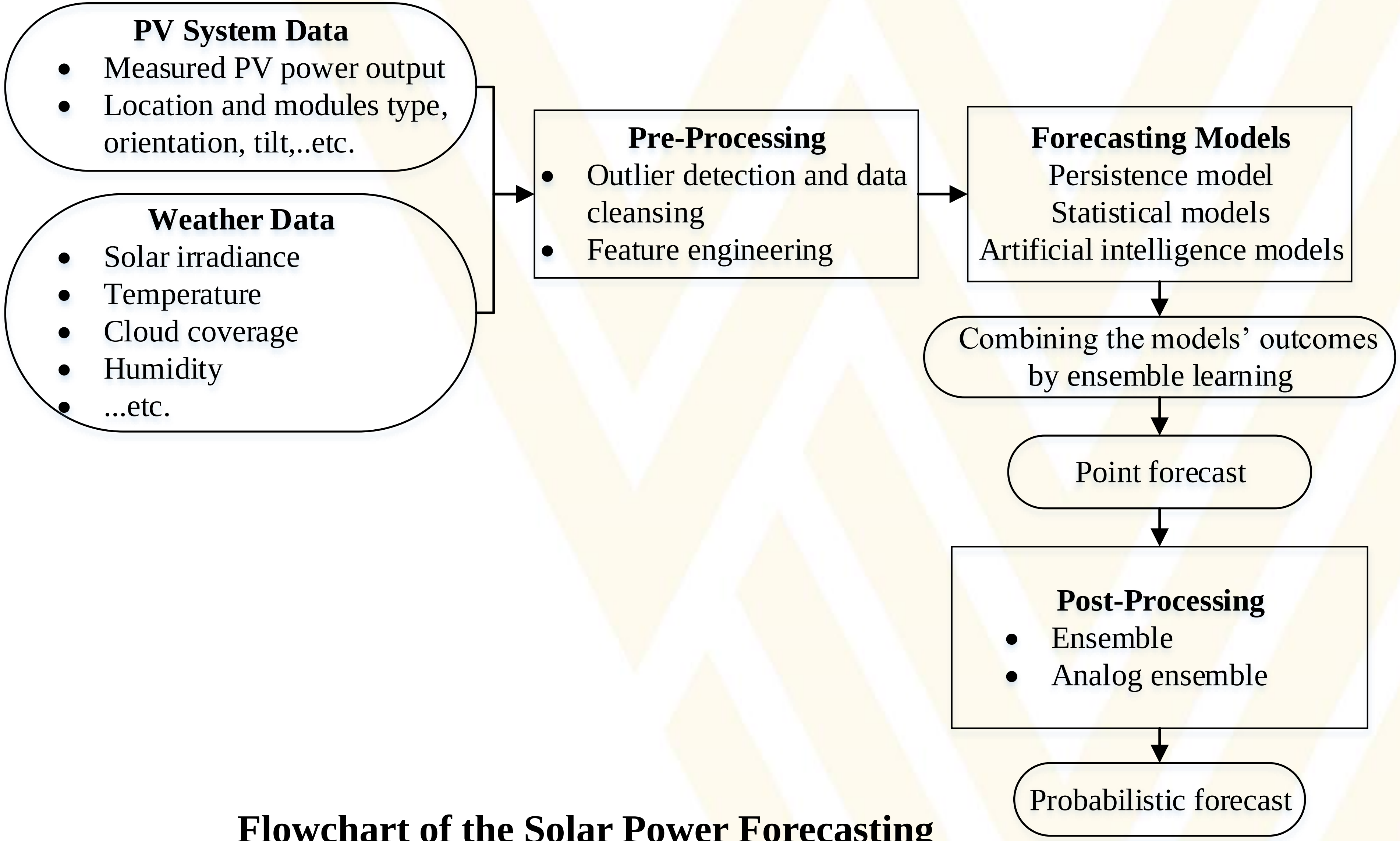
Regression



Extrapolation



Hourly Probabilistic Forecasting of Solar Power



Flowchart of the Solar Power Forecasting

Hourly Probabilistic Forecasting of Solar Power

Data Description:

PV solar system is near Canberra, Australia, consisting of 8 panels, its nominal power of (1560W), and panel orientation 38° clockwise from the north, with panel tilt (of 36°). The historical observed solar power data are normalized to the rated capacity (i.e., 1560W).

Weather predictions are produced by a global numerical weather prediction system, European Centre for Medium-Range Weather Forecasts (ECMWF).

No.	Input Variable, (X)
1	Cloud Water Content
2	Cloud Ice Content
3	Surface Pressure
4	Relative Humidity
5	Cloud Cover
6	10m - U Wind
7	10m - V Wind
8	2-m Temperature
9	Surface solar radiation down

No.	Input Variable, (X)
10	Surface thermal radiation down
11	Top net solar radiation
12	Total precipitation
13	Heat Index
14	Wind Speed
15	Hours
16	Months
17	Days of Month
18	Days of Year

Data partition into training and testing sets

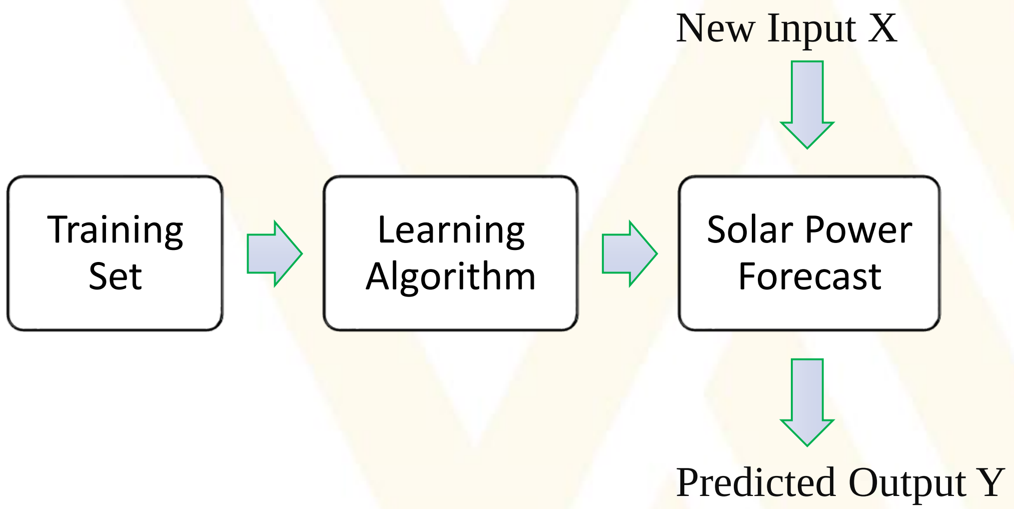
Timeline	Month	Year	Partition
From	April	2012	Training Set
To	May	2013	
From	June	2013	Testing Set
To	May	2014	

https://crowdanalytix.com/contests/global-energy-forecasting-competition-2014-probabilistic-solar-power-forecasting
http://www.ecmwf.int (European Centre for Medium-Range Weather Forecasts)

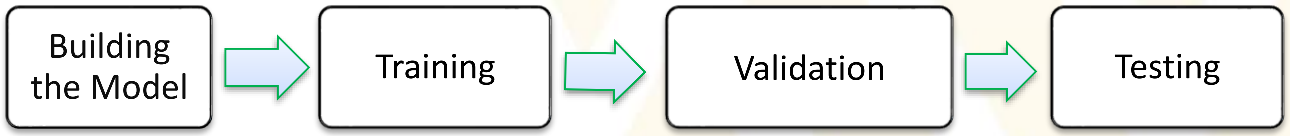
Forecasting Models

Parametric and Nonparametric Models

Multiple Linear Regression (**MLR**) Analysis, Artificial Neural Networks (**ANN**), and Support Vector Regression (**SVR**) are deployed for the short-term solar power forecasting.



Flowchart of the forecasting models

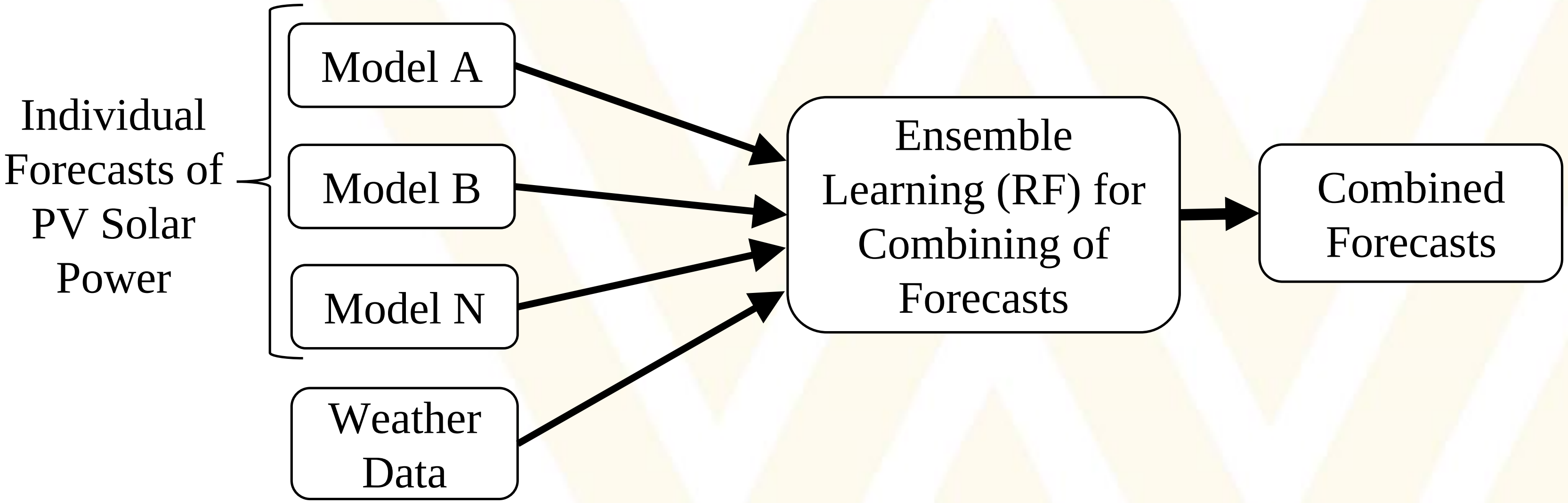


Flowchart of solar forecasting model building steps

T. Hastie, R. Tibshirani, J. Friedman, and others, *The elements of statistical learning*, 2nd Edition. Springer-Verlag New York, 2009.

Ensemble Forecasts

Combining Various Models



General diagram of combining different models

$$F_{comb}=W_A*M_A+W_B*M_B+W_C*M_C \ldots .+W_N*M_N$$

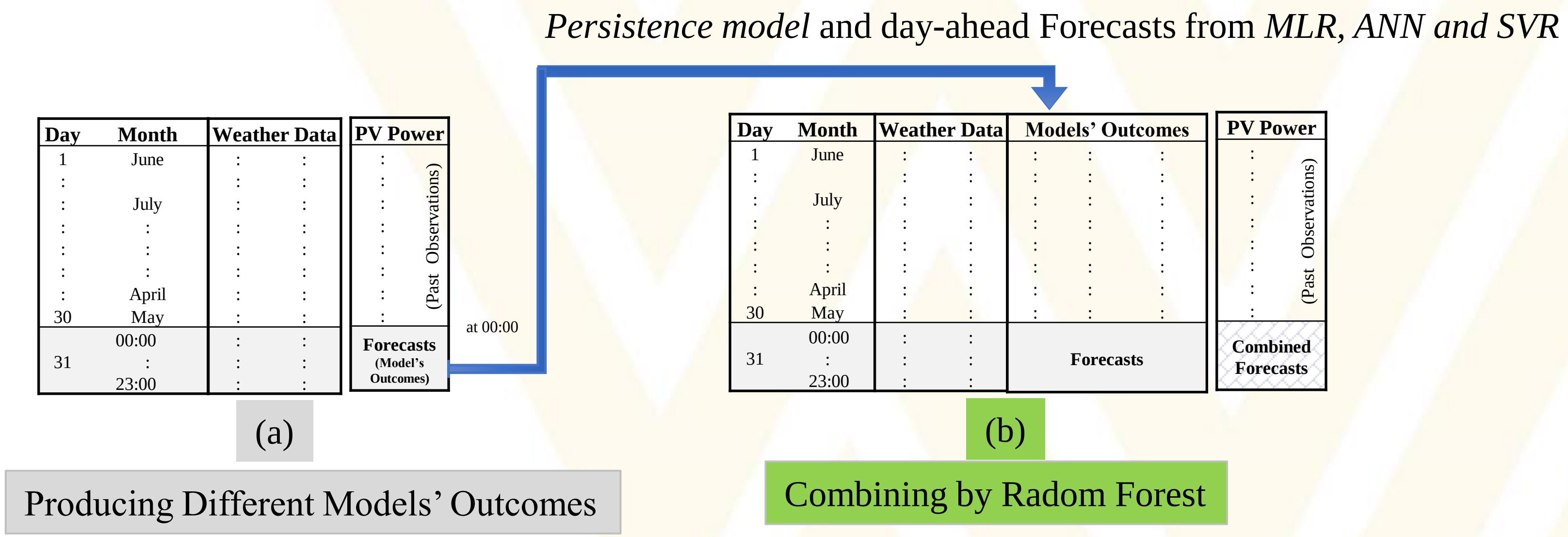
W_N is a weight is assigned to the outcome of a model M_N

Methods of Combining The Models

Random forest (RF) is chosen to be the ***ensemble learning*** method for combining the various models' outcomes.

T. Hastie, R. Tibshirani, J. Friedman, and others, *The elements of statistical learning*, 2nd Edition. Springer-Verlag New York, 2009.

Ensemble Forecasts



Schematic diagram of producing and ensemble different models' outcomes

$$Persistence\ model, \quad F(t) = P(t - horizon)$$

$$Here\ the\ horizon = 1hour$$

Probabilistic Forecasts

Ensemble-based probabilistic forecasts method:

a) Diagram of ensemble-based probabilistic forecasts,

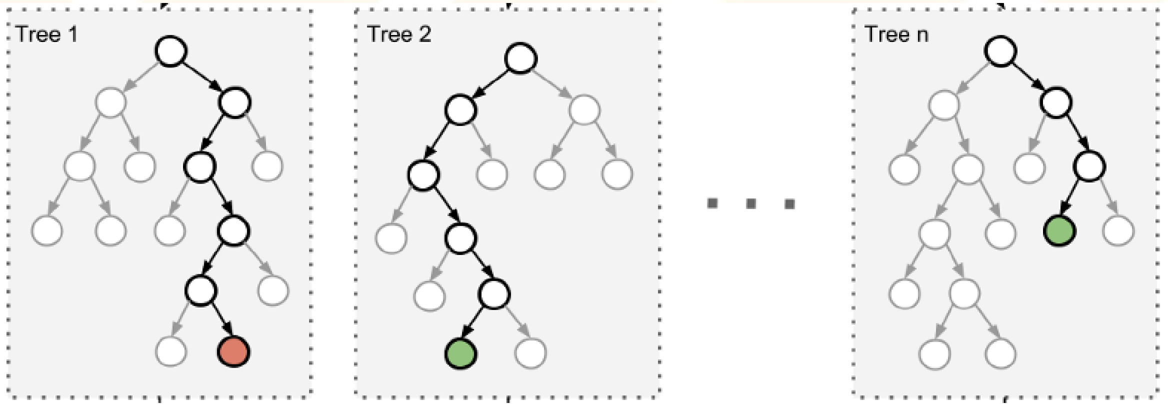
b) Splitting mechanism of trees in random forest,

c) Sample of ensemble-based probabilistic forecasts of solar power of 3 days

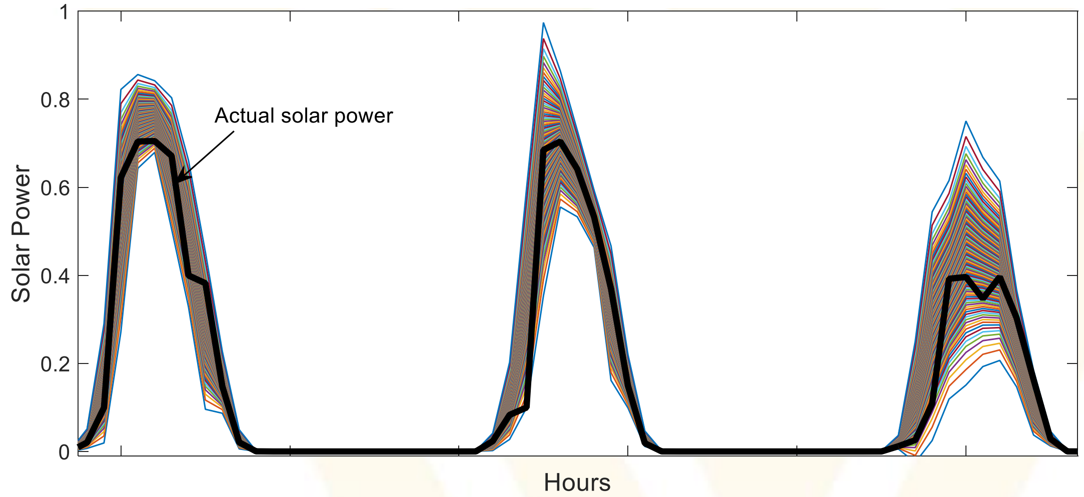
$$\hat{f}_{RF} = \frac{1}{B}\sum_{b=1}^{B} T_b(Hr)$$

(a) NWP point forecasts → Solar power point forecasts → Probabilistic model → Probabilistic forecasts of solar power

(b)

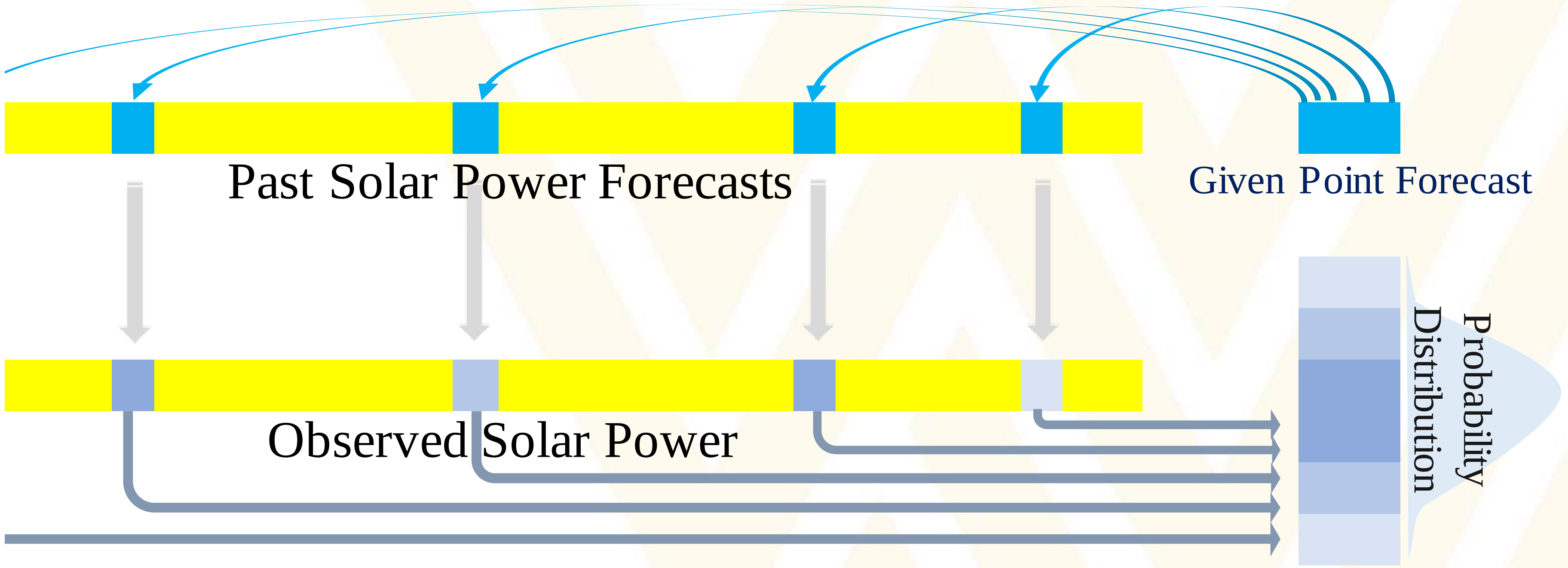


(c)



Probabilistic Forecasts

Analog Ensemble (AnEn) method:



Schematic diagram of analog ensemble method

$$\left|F_{\text{Given}}^{Hr} - F_{\text{Past}}^{Hr}\right| \leq \varepsilon$$

$\varepsilon = 0.1$

where F_{Given}^{Hr} denotes the given point forecast at an hour Hr, for which the prediction interval will be estimated, F_{Past}^{Hr} the point forecasts at the same hour of the day.

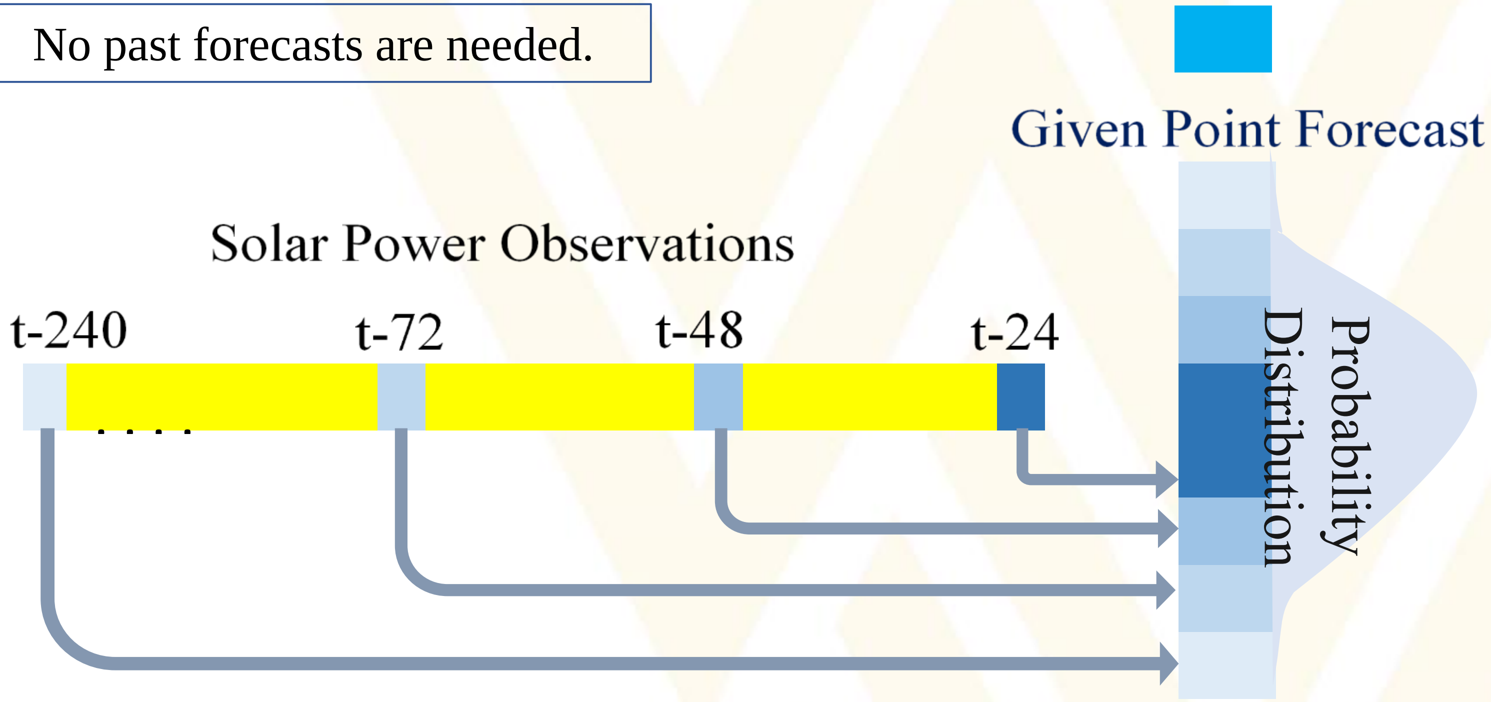
Notice that all values are normalized in the range [0, 1].

S. Alessandrini, L. Delle Monache, S. Sperati, and G. Cervone, "An analog ensemble for short-term probabilistic solar power forecast," Appl. Energy, vol. 157, pp. 95–110, 2015.

Probabilistic Forecasts

Persistence probabilistic forecasts method:

No past forecasts are needed.



Schematic diagram of analog ensemble method

The 10, 20 and 30 recent observed powers are carried out.

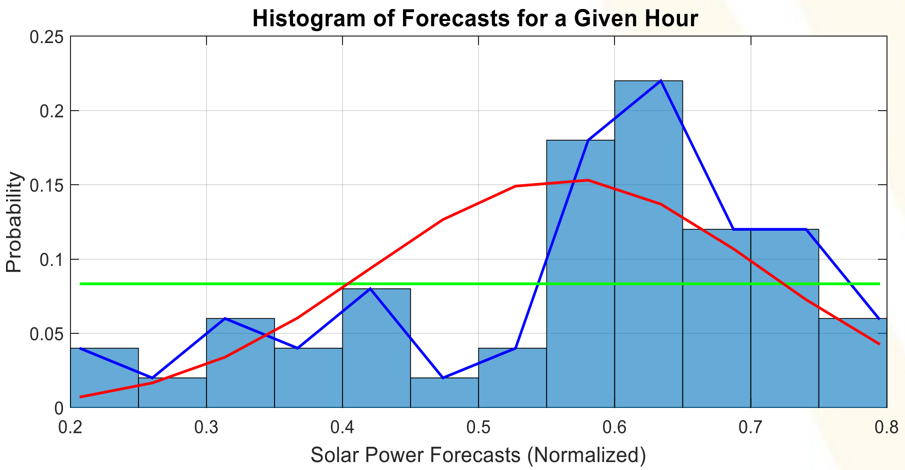
It is found that the recent **10** observed solar powers at the given hour with CDF distribution achieve more accurate persistence probabilistic forecasts.

Probabilistic Forecasts

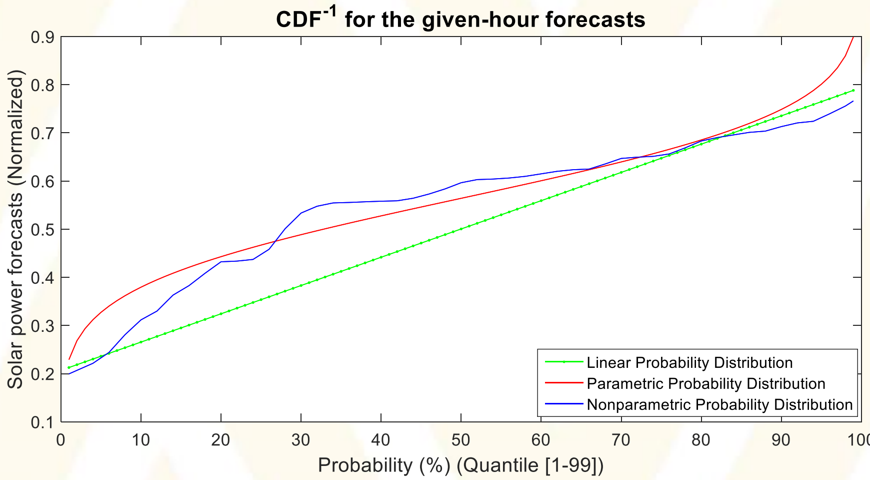
Probability distributions by the cumulative distribution function (CDF)

➔For example, for a given point forecast **at 14:00, June 2nd 2013**:

Histogram of the ensemble of RF's outcomes



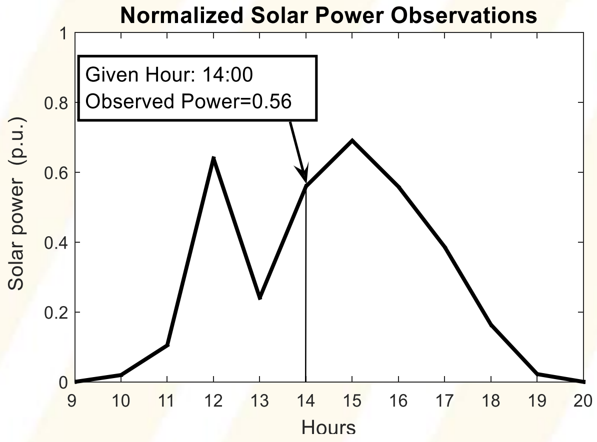
Different distributions of probability



Linear CDF
Max
Min
to derive CDF

Parametric normal-distributed CDF
Mean
Std. dev.
to derive CDF.

Nonparametric CDF
No mean
neither Std. Dev.
CDF is estimated by piecewise nonparametric method



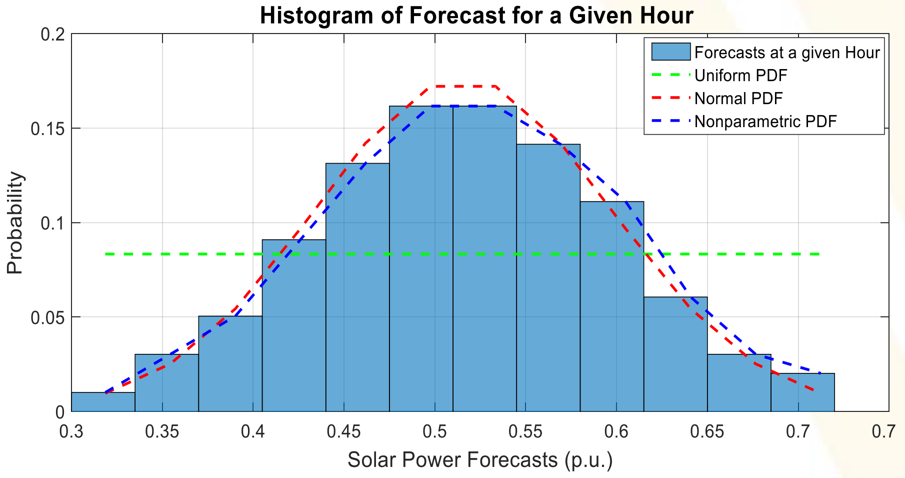
The probabilistic forecasts are estimated by using CDF^{-1}

Probabilistic Forecasts

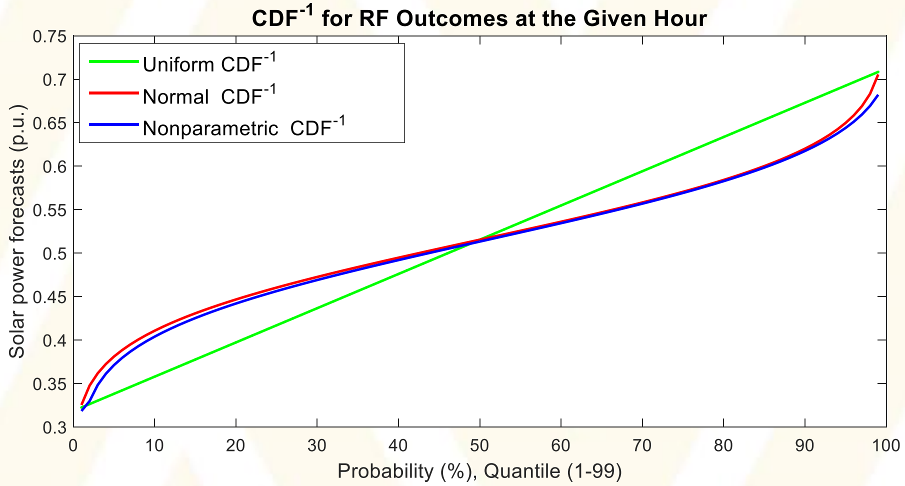
Probability distributions by the cumulative distribution function (CDF)

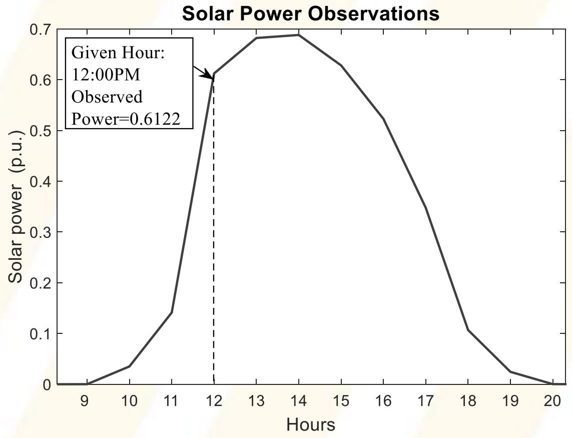
➔For example, for a given point forecast at **12:00, May 29th 2014**:

Histogram of the ensemble of RF's outcomes



Different distributions of probability





Linear CDF
Max
Min
to derive CDF

Parametric normal-distributed CDF
Mean
Std. dev.
to derive CDF.

Nonparametric CDF
No mean
neither Std. Dev.
CDF is estimated by piecewise nonparametric method

The probabilistic forecasts are estimated by using CDF^{-1}

Probabilistic Forecasts

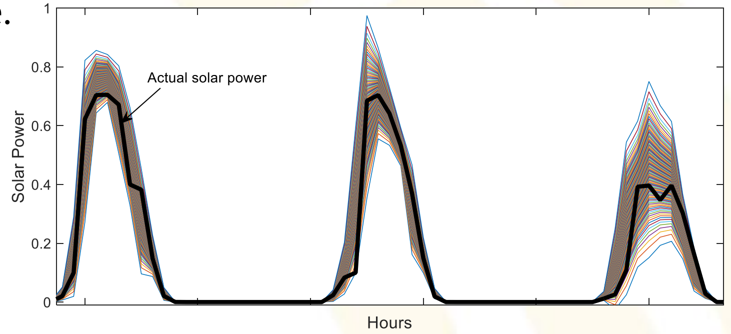
Evaluation of probabilistic forecasts:

The objective is to determine probabilistic solar forecasts in the form of probabilistic distribution (in quantiles) in incremental time steps through the forecast horizon.

A Pinball loss function is used to evaluate the accuracy of the probabilistic forecasts. It is a piecewise linear function which is often used to evaluate the accuracy of quantile forecasts.

$$\boxed{Pb_q(F,P) = \begin{cases} q(F-P), & if\ P \leq F \\ (1-q)(P-F), & if\ P > F \end{cases}}$$

where $Pb_q(F, P)$ is the pinball loss function to the probabilistic forecasts for each hour; F is the forecasted value at the certain q quantile of the probabilistic solar power forecasts, and P is the observed value of the solar power. The quantile q has discrete values $q \in [0.01, 0.99]$.For instance, $q = 0.9$ means that there is a 90% probability that the observed solar power will be less than the value of the 90th quantile.



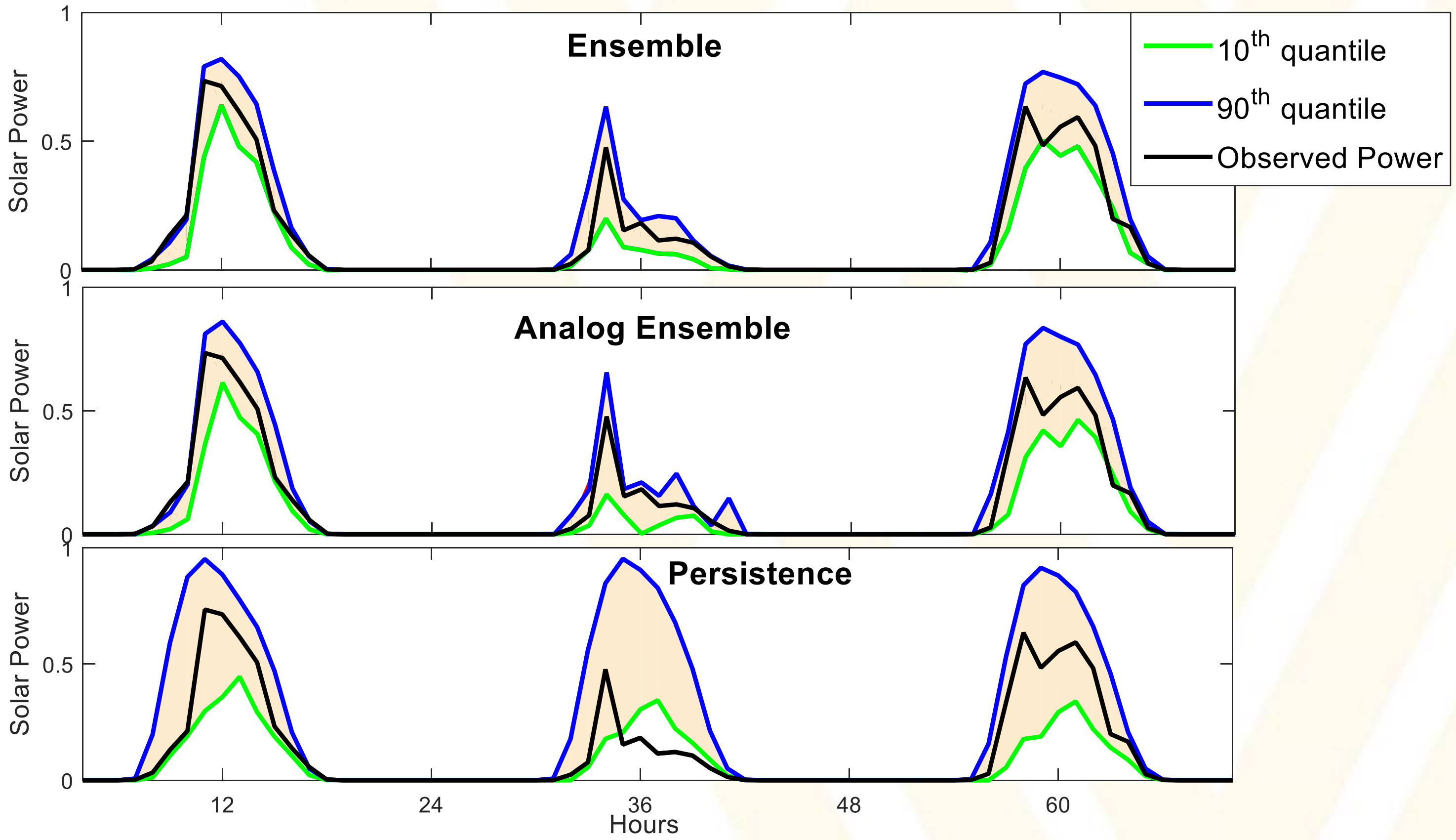
J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, Integrating renewables in electricity markets - Operational problems, vol. 205. Boston, MA: Springer US, 2014.

Results and Evaluation

Pinball loss function (Pb):

$$Pb_q(F,P) = \begin{cases} q(F-P), & if\ P \leq F \\ (1-q)(P-F), & if\ P > F \end{cases}$$

The lower Pinball (Pb) is, the more accurate probabilistic forecasts are.



Graphs of the probabilistic forecasts of the three methods for three days

Results and Evaluation

Pinball loss function (Pb):

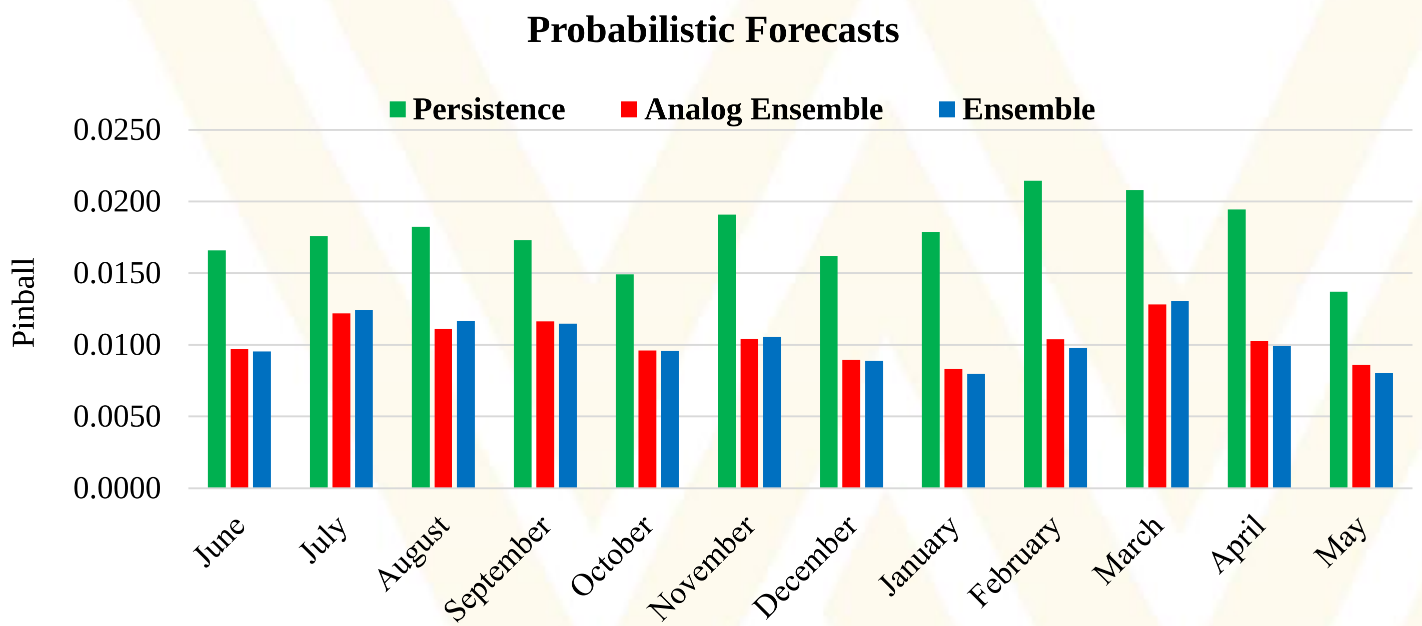
$$Pb_q(F,P) = \begin{cases} q(F-P), & if\ P \le F \\ (1-q)(P-F), & if\ P > F \end{cases}$$

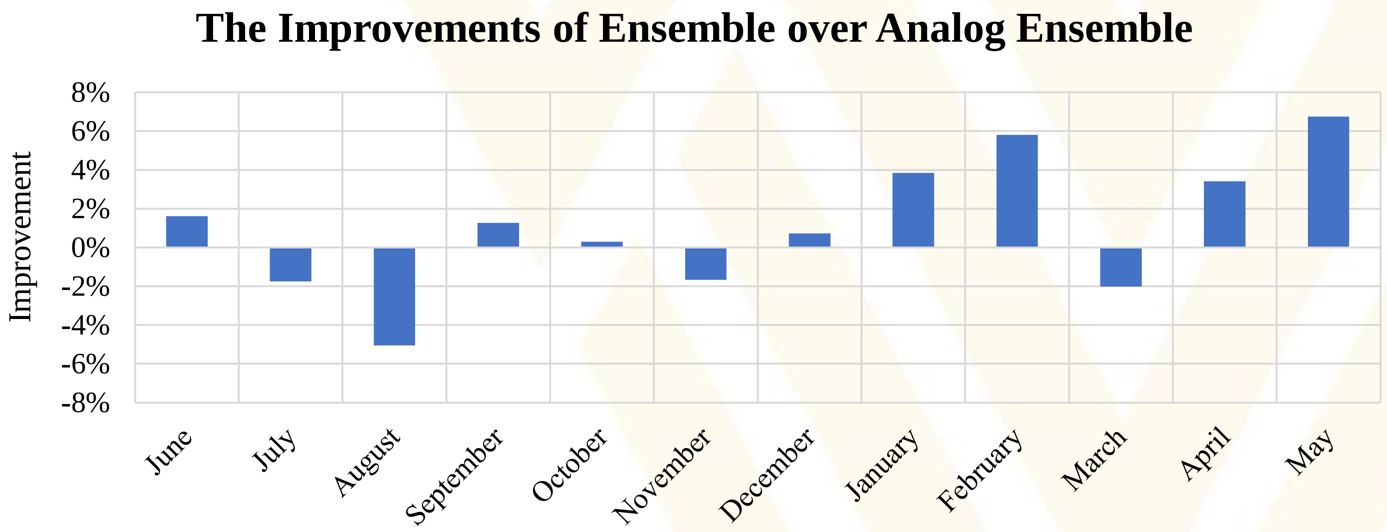
The lower Pinball (Pb) is, the more accurate probabilistic forecasts are.

Month	Pinball (Pb)			Improvement of Ensemble Over	
	Persistence	Analog Ensemble	Ensemble	Persistence	Analog Ensemble
June	0.0166	0.0097	0.0095	42%	2%
July	0.0176	0.0122	0.0124	29%	-2%
August	0.0182	0.0111	0.0117	36%	-5%
September	0.0173	0.0116	0.0115	34%	1%
October	0.0149	0.0096	0.0096	36%	0%
November	0.0191	0.0104	0.0106	45%	-2%
December	0.0162	0.0090	0.0089	45%	1%
January	0.0179	0.0083	0.0080	55%	4%
February	0.0215	0.0104	0.0098	54%	6%
March	0.0208	0.0128	0.0131	37%	-2%
April	0.0194	0.0103	0.0099	49%	3%
May	0.0137	0.0086	0.0080	42%	7%
Average Pb.	**0.0178**	**0.0103**	**0.0102**	**42%**	**1%**

Monthly Pinball of the probabilistic forecasts of the three methods

Results and Evaluation





Conclusions

- ✓ The probabilistic forecasting are quantifying the uncertainty associated with point forecasts.
- ✓ Combining the forecasts of various models leads to accurate point and probabilistic forecasts.
- ✓ Throughout the complete year, the ensemble based-probabilistic forecasts are more accurate. than the analog ensemble and persistence probabilistic forecasts.
- ✓ The random forest is a powerful ensemble learning method.
- ✓ The CDF with the assumption of a normal distribution is better than the linear distribution to produce the probabilistic forecasts.
- ✓ The nonparametric estimation of CDF without the normality assumption yields a small improvement (Pb=0.0100 vs. Pb=0.0102 with a normality assumption of CDF).
- ✓ With additional historical data, the forecasting performance could be improved.

Thanks for Your Listening

Any Question?

Mohamed Abuella

mabuella@uncc.edu

http://epic.uncc.edu/

Energy Production and Infrastructure Center

University of North Carolina at Charlotte

UNC CHARLOTTE