An Ensemble Model for Day-ahead Electricity Demand Time Series Forecasting

Wen Shen, Vahan Babushkin
Zeyar Aung, Wei Lee Woon
Masdar Institute of Science and Technology
Abu Dhabi, United Arab Emirates

Outline

- Time Series Forecasting
- Related work
- Pattern Forecasting Ensemble Model
- Experimental Results
- Conclusions and Future Work

Time Series

- Time Series: a sequence of successive data points over uniform time intervals.

(Source: Google Finance)

Time Series Forecasting

- Definition: the process of predicting future values based on previously observed values.
- Applications: numerous applications in real world, e.g.,
  - financial markets forecasting
  - weather forecasting
  - earthquake prediction
  - electricity demand and price forecasting
- In this work, we focus on electricity demand forecasting.

Challenges

- It is difficult to predict the electricity demand using very few features.

- It is challenging to produce accurate and reliable forecasts for the hourly electricity demand due to the variations.

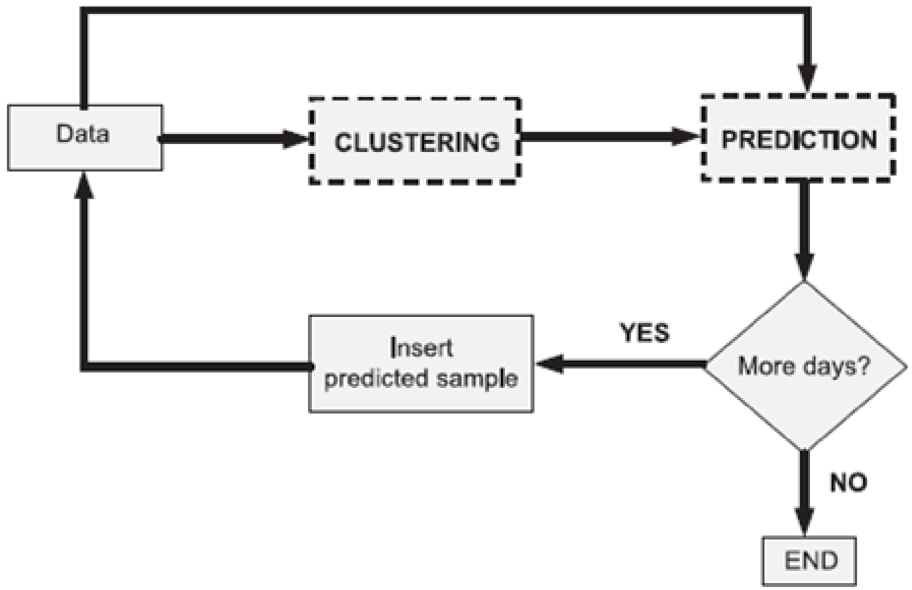
Outline

- Time Series Forecasting
- Related Work
- Pattern Forecasting Ensemble Model
- Experimental Results
- Conclusions and Future Work

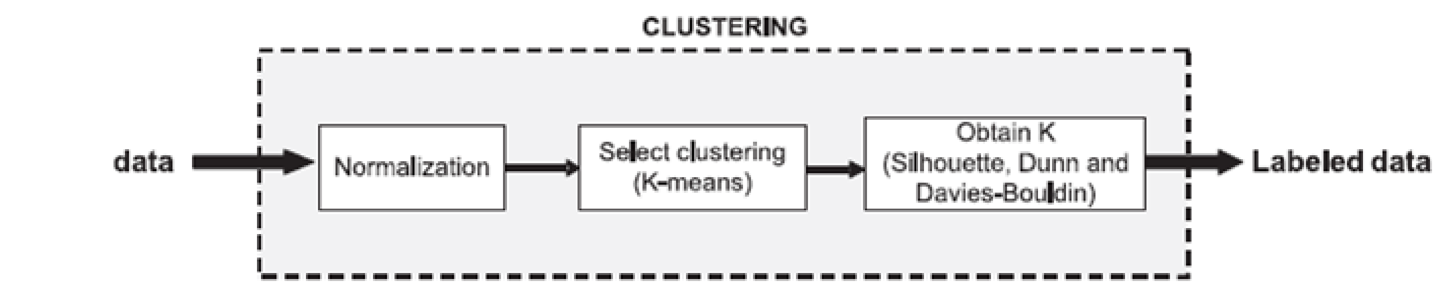
Related Work

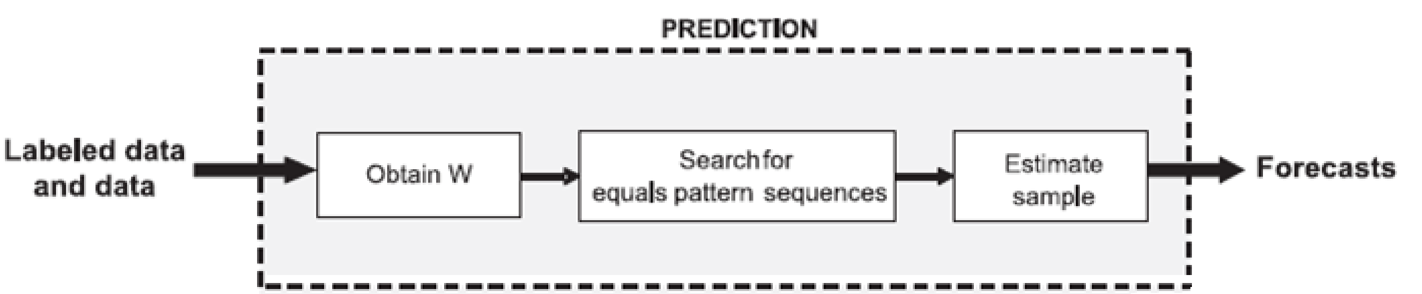
- Linear Models
  - ARIMA (AutoRegressive Integrated Moving Average)
  - GARCH (Generalized AutoRegressive Conditional Heteroskedasticity)
- Artificial Neural Network and its hybrid models
- Support Vector Machines
- Nearest Neighbors
- Clustering

Pattern Sequence-based Forecasting (PSF)



(Source: Mart ńez- Álvarez etc.)





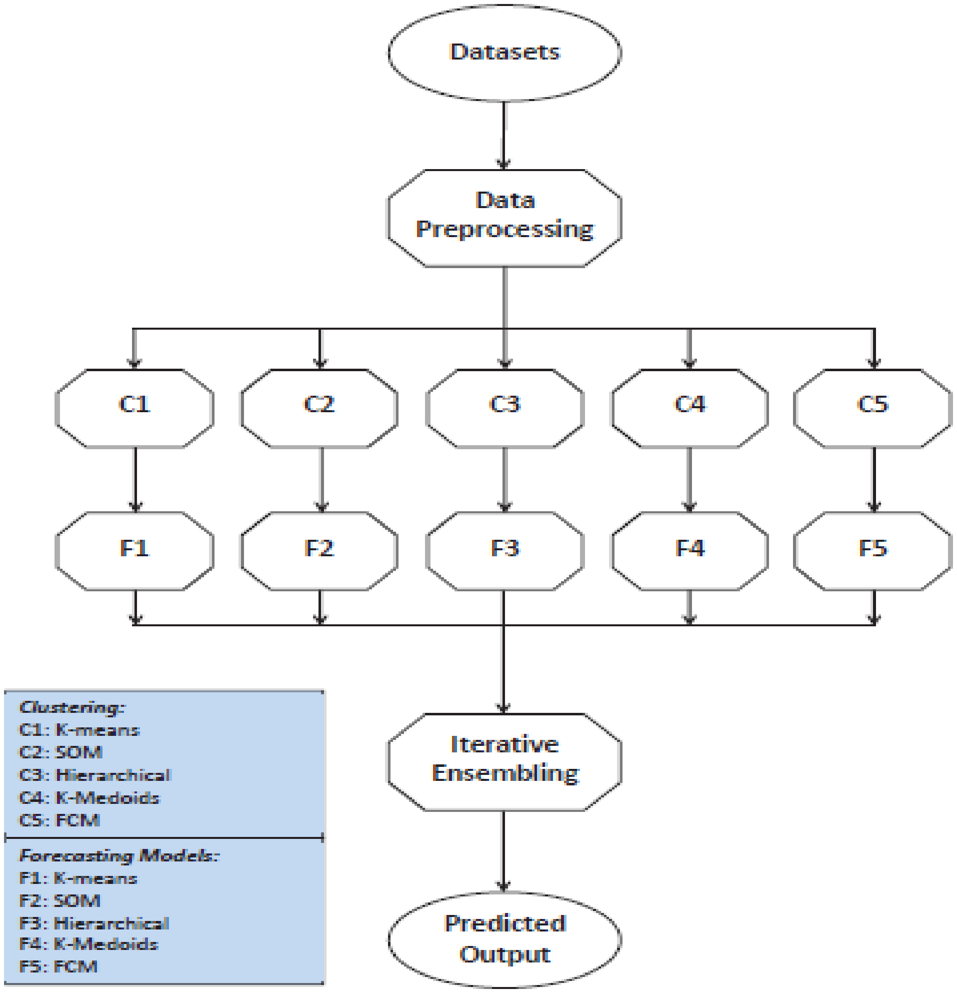
Outline

- Time Series Forecasting
- Related Work
- Pattern Forecasting Ensemble Model
- Experimental Results
- Conclusions and Future Work

Pattern Forecasting Ensemble Model (PFEM)

- Data preprocessing
- Applying Individual Clustering Methods
- Building Individual Forecasting Models
- Iterative Ensembling

Pattern Forecasting Ensemble Model



Iterative Ensembling

1. Initialize the vector of observation weights $\mathbf{w}^{(0)} = (w_1^{(0)}, w_2^{(0)}, \ldots, w_n^{(0)})$, where n is the number of participating forecasting models for ensemble learning ($n = 5$ in our case):

$$w_i^{(0)} = 1/n \quad \forall i = 1, 2, \ldots, n. \tag{1}$$

2. For a training dataset with M days, for $m = 1$ to M:

$$\mathbf{P}^{(m)} = \sum_{i=1}^{n} w_i^{(m-1)} \mathbf{P}_i^{(m)} \tag{2}$$

where $\mathbf{P}^{(m)} = (P_1^{(m)}, P_2^{(m)}, \ldots, P_{24}^{(m)})$ is the vector of combined predicted values for 24 hours in day m, and $\mathbf{P}_i^{(m)} = (P_{i1}^{(m)}, P_{i2}^{(m)}, \ldots, P_{i24}^{(m)})$ is the vector of predicted values for 24 hours in day m generated by the individual forecasting model i.

Iterative Ensembling

(a) Define the prediction error rate for the iteration m:

$$err^{(m)} = \frac{1}{24}\sum_{h=1}^{24} \frac{\mid P_h^{(m)} - A_h^{(m)} \mid}{\bar{A}^{(m)}} \tag{3}$$

where $\mathbf{A}^{(m)} = (A_1^{(m)}, A_2^{(m)}, \ldots, A_{24}^{(m)})$ is the vector of actual values for 24 hours in day m and $\bar{A}^{(m)}$ is the average of actual values for 24 hours in that day.

(b) Calculate new weights for the iteration m:

$$\mathbf{w}^{(m)} = \underset{w_i^{(m-1)}, i=1,\ldots,n}{\text{argmin}} \quad err^{(m)}, \tag{4}$$

such that $0 \leq w_i^{(m)} \leq 1, \sum_{i=1}^{n} w_i^{(m)} = 1$

Iterative Ensembling

3. Calculate the initial weights for testing:

$$\mathbf{w} = \frac{1}{m} \sum_{m=1}^{M} \mathbf{w}^{(m)} \tag{5}$$

   Re-normalize $\mathbf{w}$.

4. Produce the predicted values using Equation (2) with the weights obtained in Step 3.

5. Add testing sample to the training dataset, increase M by 1, recalculate the initial weights for testing using Step 2, 3 and 4 for the next testing sample.

Outline

- Time Series Forecasting
- Related Work
- Pattern Forecasting Ensemble Model
- Experimental Results
- Conclusions and Future Work

Metrics

- Mean Error Relative (MER):

$$MER = 100 \times \frac{1}{N} \sum_{h=1}^{N} \frac{|\hat{x}_h - x_h|}{\overline{x}}$$

- Mean Absolute Error(MAE):

$$MAE = \frac{1}{N} \sum_{h=1}^{N} |\hat{x}_h - x_h|$$

- Mean Absolute Percentage Error(MAPE):

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{x_i - \hat{x}_i}{x_i} \right|$$

Experimental Settings

- Datasets:
  - New York Independent System Operator (NYISO)
  - Australian Energy Market Operator (ANEM)
  - Ontario's Independent Electricity System Operator (IESO)
- Number of Clusters:
  - Silhouette index
  - Dunn index
  - Davies-Bouldin index

  (using the majority rule)
- Window Size
  - 12 folds cross validation

Results (2009)

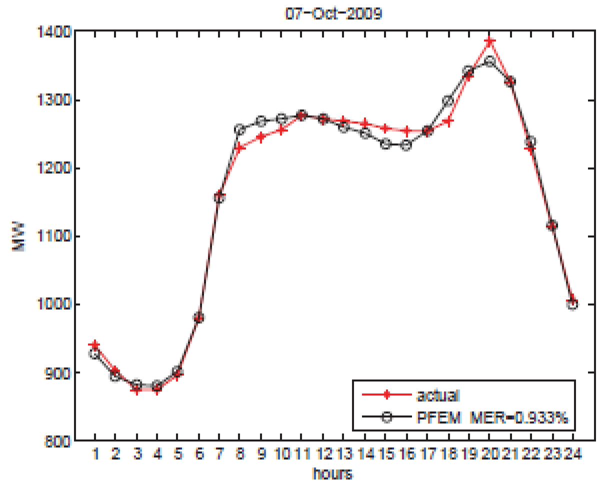


Figure 1: Best prediction of PFEM for NYISO dataset(2009).

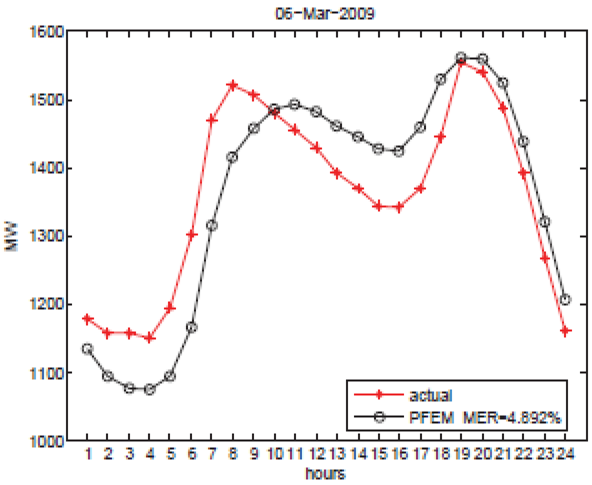


Figure 2: Worst prediction of PFEM for NYISO dataset(2009).

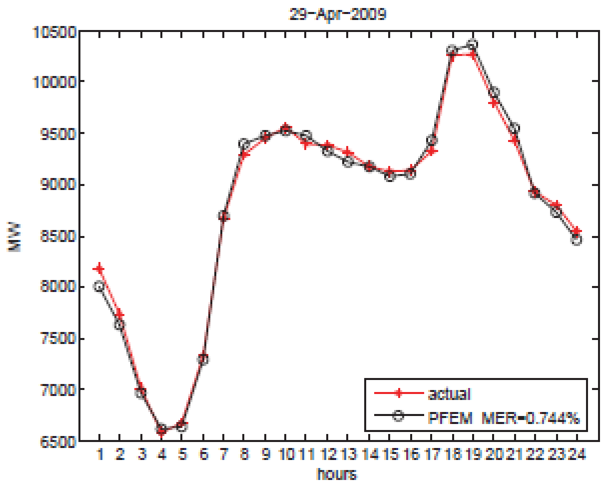


Figure 3: Best prediction of PFEM for ANEM dataset(2009).

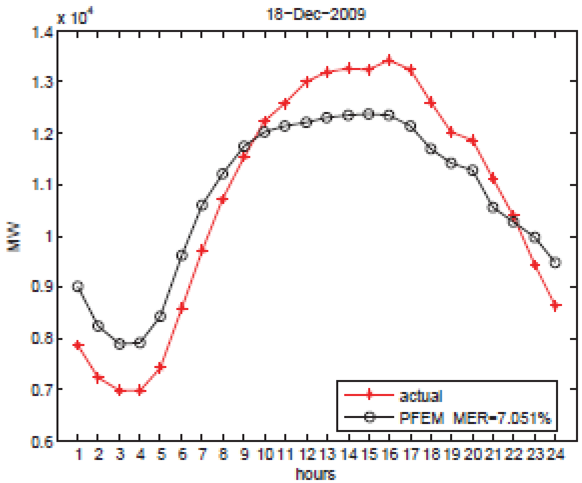


Figure 4: Worst prediction of PFEM for ANEM dataset(2009).

Results (2009)

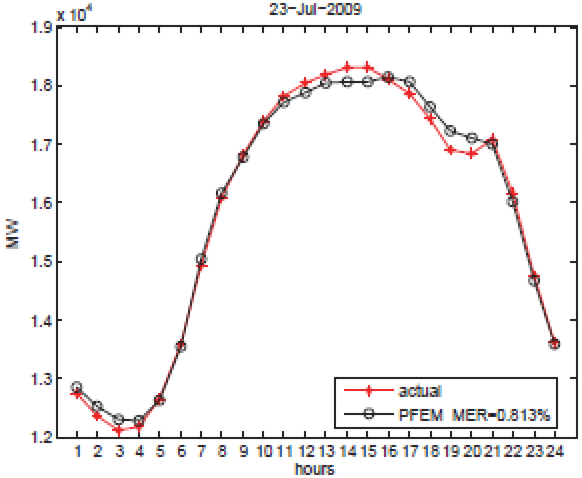


Figure 5: Best prediction of PFEM for IESO dataset(2009).

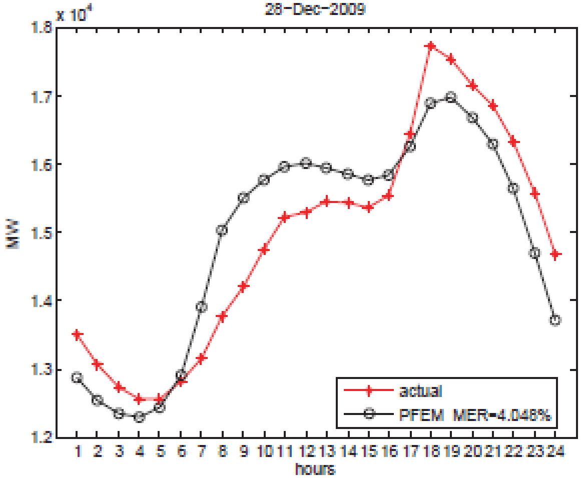


Figure 6: Worst prediction of PFEM for IESO dataset(2009).

Market	Error	K-means		SOM		Hierarchical		K-medoids		Fuzzy C-means		PFEM	
		Err	σ	Err	σ	Err	σ	Err	σ	Err	σ	Err	σ
NYISO	MER	3.11	0.41	3.06	0.44	2.92	0.31	2.97	0.37	3.18	0.43	2.76	0.35
	MAE	39.16	6.88	38.5	7.38	36.79	5.92	37.27	6.21	39.89	6.61	34.78	6.31
	MAPE	3.18	0.42	3.12	0.44	2.99	0.32	3.03	0.38	3.26	0.43	2.82	0.36
ANEM	MER	2.98	0.862	3.18	0.76	2.76	0.91	2.79	0.73	2.58	0.67	2.55	0.80
	MAE	259.66	85.92	283.23	76.98	244.83	85.89	249.25	71.88	229.28	65.22	228.35	79.00
	MAPE	2.96	0.902	3.25	0.81	2.78	0.95	2.86	0.77	2.63	0.71	2.61	0.84
IESO	MER	2.42	0.36	2.67	0.36	2.5	0.41	2.34	0.29	2.31	0.27	2.23	0.25
	MAE	384.02	59.27	422.61	49.22	394.26	58.62	371.21	45.23	364.71	39.31	354.99	46.92
	MAPE	2.49	0.38	2.74	0.36	2.58	0.43	2.41	0.31	2.37	0.30	2.30	0.27

Table 1: Summary performance results of models tested on demand data of NYISO, ANEM and IESO markets for 2009.

Results (2010)

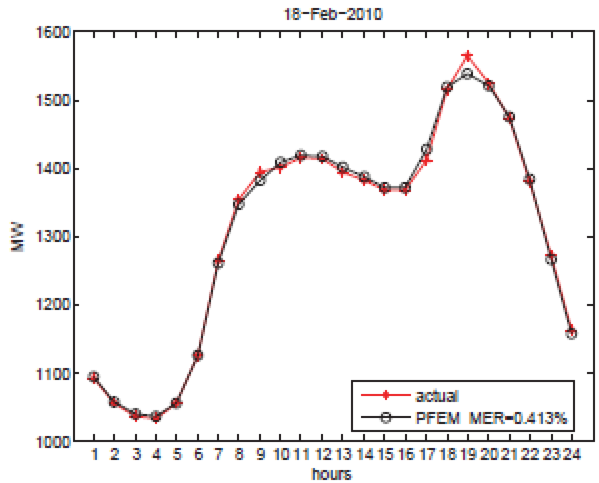


Figure 7: Best prediction of PFEM for NYISO dataset(2010).

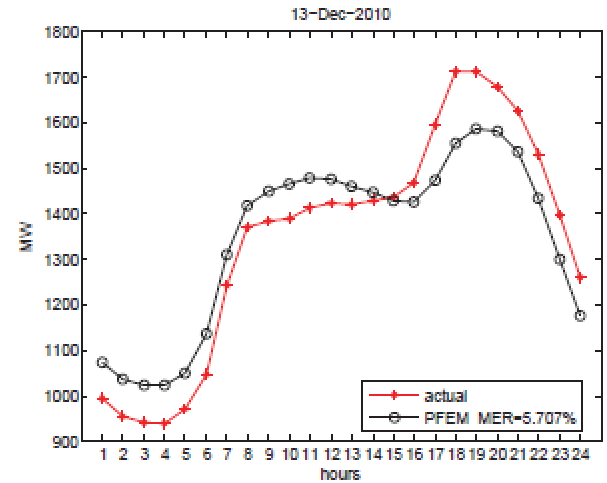


Figure 8: Worst prediction of PFEM for NYISO dataset(2010).

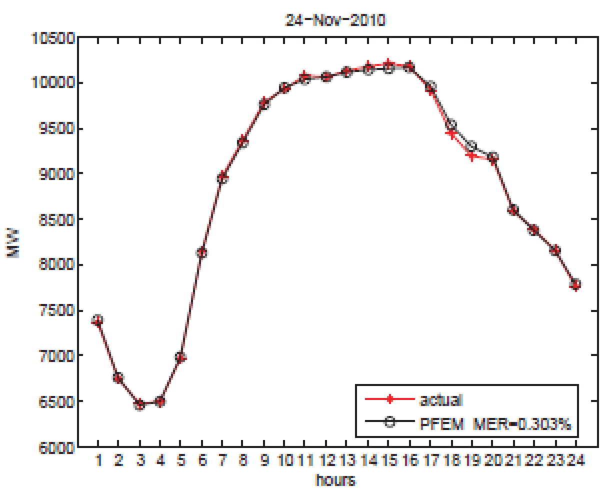


Figure 9: Best prediction of PFEM for ANEM dataset(2010).

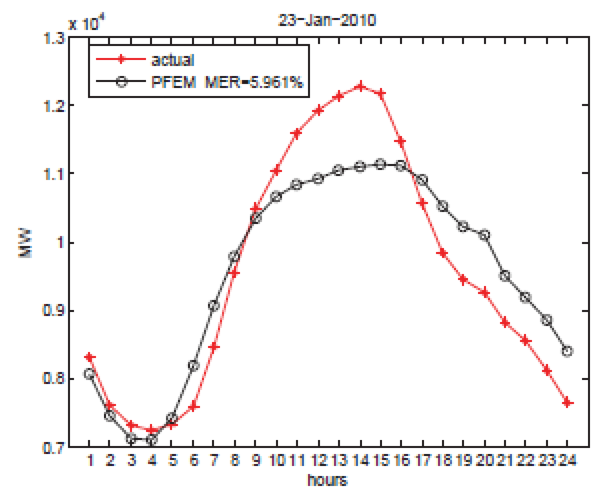


Figure 10: Worst prediction of PFEM for ANEM dataset(2010).

Results (2010)

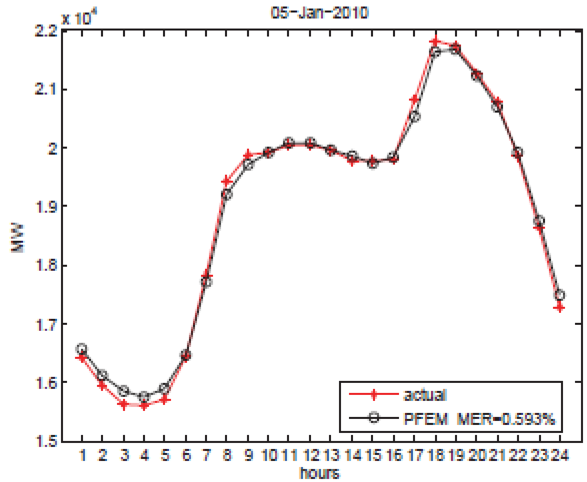


Figure 11: Best prediction of PFEM for IESO dataset(2010).

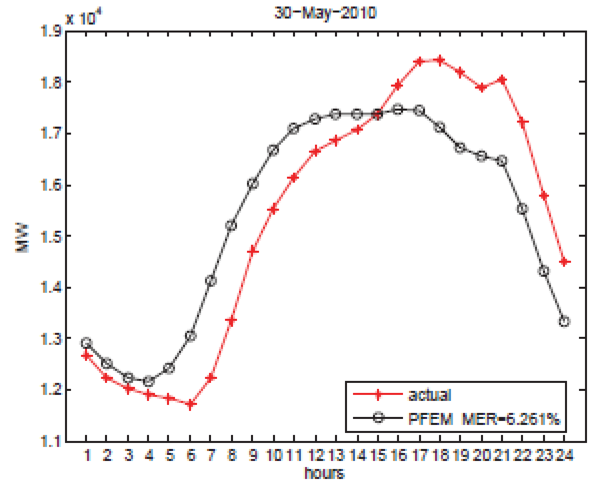


Figure 12: Worst prediction of PFEM for IESO dataset(2010).

Market	Error	K-means		SOM		Hierarchal		K-medoids		Fuzzy C-means		PFEM	
		Err	σ	Err	σ	Err	σ	Err	σ	Err	σ	Err	σ
NYISO	MER	3.09	0.61	3.11	0.55	2.97	0.54	2.77	0.54	3.07	0.56	2.74	0.57
	MAE	40.72	11.45	40.87	10.86	40.87	10.86	36.57	10.38	40.35	10.84	36.27	11.11
	MAPE	3.16	0.49	3.12	0.55	2.99	0.58	2.79	0.57	3.26	0.63	2.78	0.58
ANEM	MER	2.78	0.56	2.89	0.39	2.45	0.71	2.64	0.52	2.73	0.51	2.39	0.46
	MAE	243.88	52.78	254.17	38.63	214.61	60.99	232.11	47.29	239.93	46.91	210.92	44.11
	MAPE	2.83	0.58	2.96	0.41	2.48	0.74	2.69	0.55	2.78	0.52	2.44	0.48
IESO	MER	2.29	0.48	2.52	0.47	2.19	0.45	2.28	0.46	2.59	0.53	2.10	0.44
	MAE	372.48	91.21	410.51	89.67	354.15	83.38	368.67	86.72	422.13	102.99	345.30	87.04
	MAPE	2.35	0.52	2.59	0.51	2.24	0.48	2.34	0.49	2.66	0.56	2.18	0.48

Table 2: Summary performance results of models tested on demand data of NYISO, ANEM and IESO markets for 2010.

Outline

- Time Series Forecasting
- Related Work
- Pattern Forecasting Ensemble Model
- Experimental Results
- Conclusions and Future Work

Conclusions and Future Work

- The performance of PFEM outperforms the performance of the other five individual models in terms of MER, MAE and MAPE.

- It's more accurate and reliable to use PFEM for day-ahead electricity demand forecasting than other five individual PSF-style methods.

- Future work is left to investigate non-linear combinations of the individual models.

Introducing our institute: Masdar Institute of Science and Technology

- One of the leading research-based universities in the Middle East.

- Located in Masdar City – one of the most sustainable communities on the planet.
- Research focuses: renewable energy, sustainability, environment, water resources and microelectronics, etc.

- Currently carrying out many collaborative research projects with Massachusetts Institute of Technology (MIT), Cambridge, USA.

Welcome to visit our website:
http://www.masadar.ac.ae

We are offering scholarships!

- MSc and PhD scholarships.
  - Tuition fees fully covered
  - Free accommodation
  - Monthly stipend
  - Other benefits

- Students are to participate in the collaborative projects with MIT.
- For information on admission:
  - http://www.masdar.ac.ae/Admission.html

Thank you!