“Prediction of building energy consumption”

- **Artificial neural network model for forecasting sub-hourly electricity usage in Commercial buildings**. Energy Build 2016;111:184–94.

- **A hybrid model approach for forecasting future residential electricity consumption.** Energy Build 2016;117:341–51.

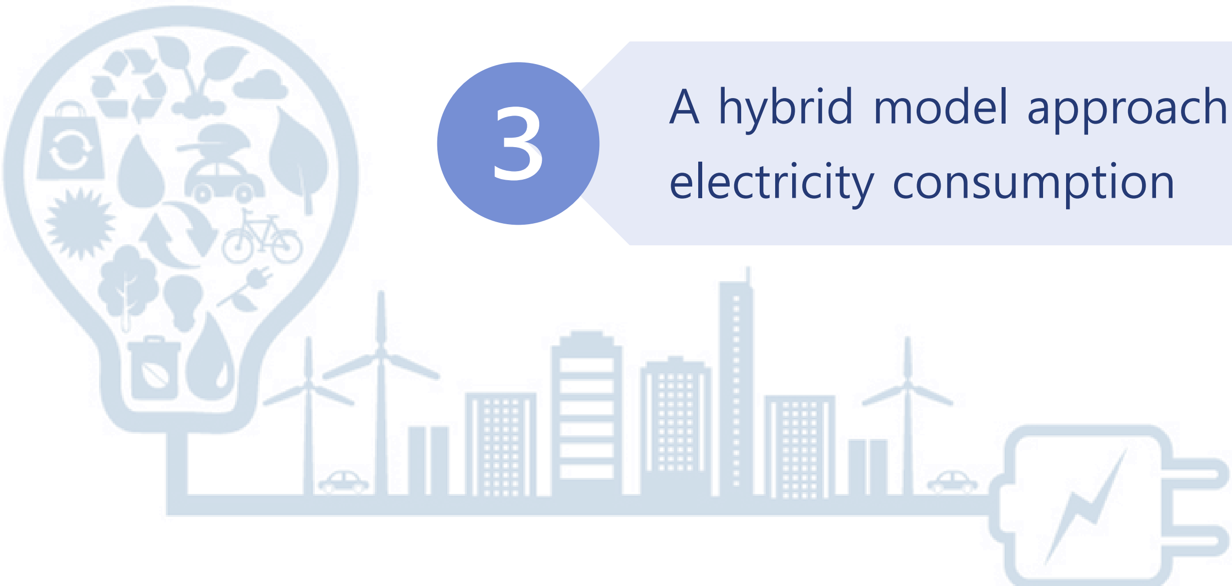
Energy AI Track **Kim Sohee**

AI application for energy system

22.06.14

Contents

1. Prediction models for forecasting building energy consumption
2. Artificial neural network model for forecasting sub-hourly electricity usage in Commercial buildings
3. A hybrid model approach for forecasting future residential electricity consumption

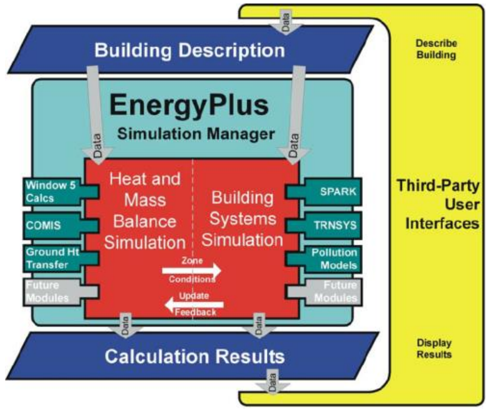


❖ Prediction models

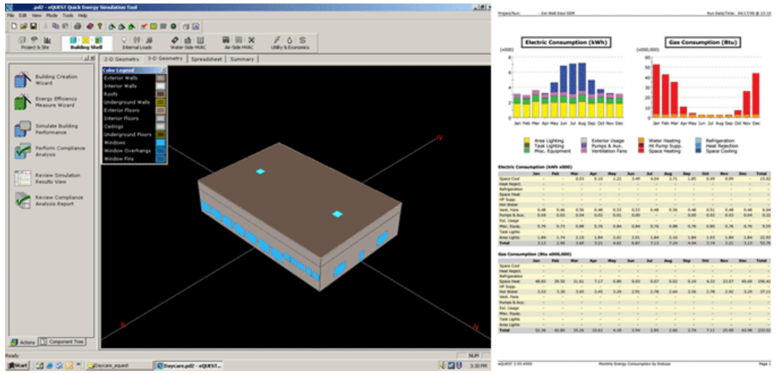
◆ Engineering models (Physical models)

- Rely on **thermodynamic rules** for detailed energy modelling and analysis
- Software : EnergyPlus, eQuest, and Ecotect
  - Calculate building energy consumption based on detailed building and environmental parameters
- Detailed data may not be available to the users at the time of simulation

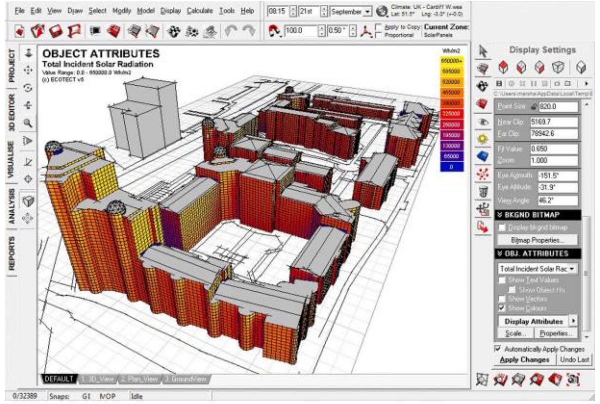
  ⇒ Failure to provide accurate input can result in poor prediction performance



> EnergyPlus
(whole building energy simulation program)



> eQuest
(Building energy analysis program)



> EcoTect

❖ Prediction models

◆ Data-driven models

- Learns from **historical/available data** for prediction
- Correlate the energy consumption or energy index with the influencing variables
- 4 steps : data collection, data preprocessing, model training, and model testing
- Machine learning algorithms : SVM, ANN, decision trees, and other statistical algorithms

Reference	Learning algorithm (type)	Building type	Temporal granularity	Type of energy consumption predicted	Purpose of prediction	Type of dataset (simulation tool)	Types of feature	Data size	Performance (metric)
[26]	SVM (RBF)	Non-residential	Hourly	Cooling	HVAC system operation improvement	Real (N/A)	Date, daily average temperature, daily lowest temperature, daily highest temperature	620 instances	0.17 (RMSE)
	PCA-SVM (RBF)								0.04 (RMSE)
	KPCA-SVM (RBF)								0.02 (RMSE)
[41]	SVM (RBF)	Non-residential	Hourly	Cooling	N/S	Real (N/A)	Date, daily average temperature, daily lowest temperature, daily highest temperature	620 instances	0.17 (RMSE)
	PCA-SVM (RBF)								0.04 (RMSE)
	PCA-WSVM (RBF)								0.03 (RMSE)
[20]	SVM (RBF)	Non-residential	Hourly	Cooling	HVAC system design	Simulated (DeST)	Dry-bulb temperature, relative humidity, solar radiation	5 months	1.15% - 1.18% (CV)
	ANN(BPNN)								2.22% - 2.36% (CV)
[37]	SVM (RBF)	Non-residential	Hourly	Cooling	HVAC system design	Simulated (DeST)	Dry-bulb temperature, relative humidity, solar radiation	5 months	1.15% - 1.18% (CV)
	ANN(BPNN)								2.22% - 2.36% (CV)
	ANN(RBFNN)								1.43% - 1.51% (CV)
	ANN(GRNN)								1.19% - 1.20% (CV)
[29]	LS-SVM (RBF)	Non-residential	Hourly	Cooling	HVAC system optimization	Simulated (DeST)	Dry-bulb temperature, relative humidity, solar radiation	4 months	5.56% (CV)
	[illegible]								[illegible]

❖ Prediction models

◆ Data-driven models Limitations

- Data-driven prediction models may not perform well outside of their training range.
- Black-box models – their internals are not known

⇒ **Hybrid or grey-box modelling approaches** (physical + data-driven)

Prediction models
- Data-driven models
- Physical models

➢ Black-box model

Artificial neural network model for forecasting sub-hourly electricity usage in Commercial buildings.

Chae YT, Horesh R, Hwang Y, Lee YM. Energy Build 2016;111:184–94.

➢ Grey-box model (hybrid)

A hybrid model approach for forecasting future residential electricity consumption.

Dong B, Li Z, Rahman SMM, Vega R. Energy Build 2016;117:341–51.

2. Artificial neural network model for forecasting sub-hourly electricity usage in Commercial buildings.

Chae YT, Horesh R, Hwang Y, Lee YM.
Energy Build 2016;111:184–94.

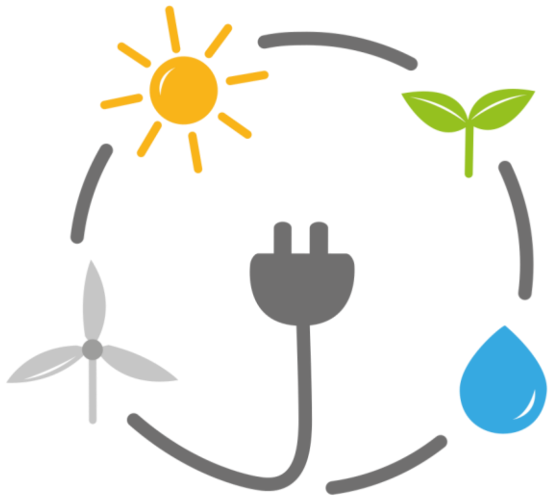
❖ Introduction

- **Forecasting electricity load**
  – commercial & industrial buildings
  - To be able to better manage energy usage

- **The short-term load forecast (STLF)**
  – a period shorter than a day → more interest
  ∵ The utility prices may change by seasonality, time-of-use in on/off peak period, and contract demand

❖ Introduction

- **The short-term load forecast (STLF) – sub-hourly electricity consumption of buildings**



✓ *Escriva-Escriva et al.* – STLF model using ANNs

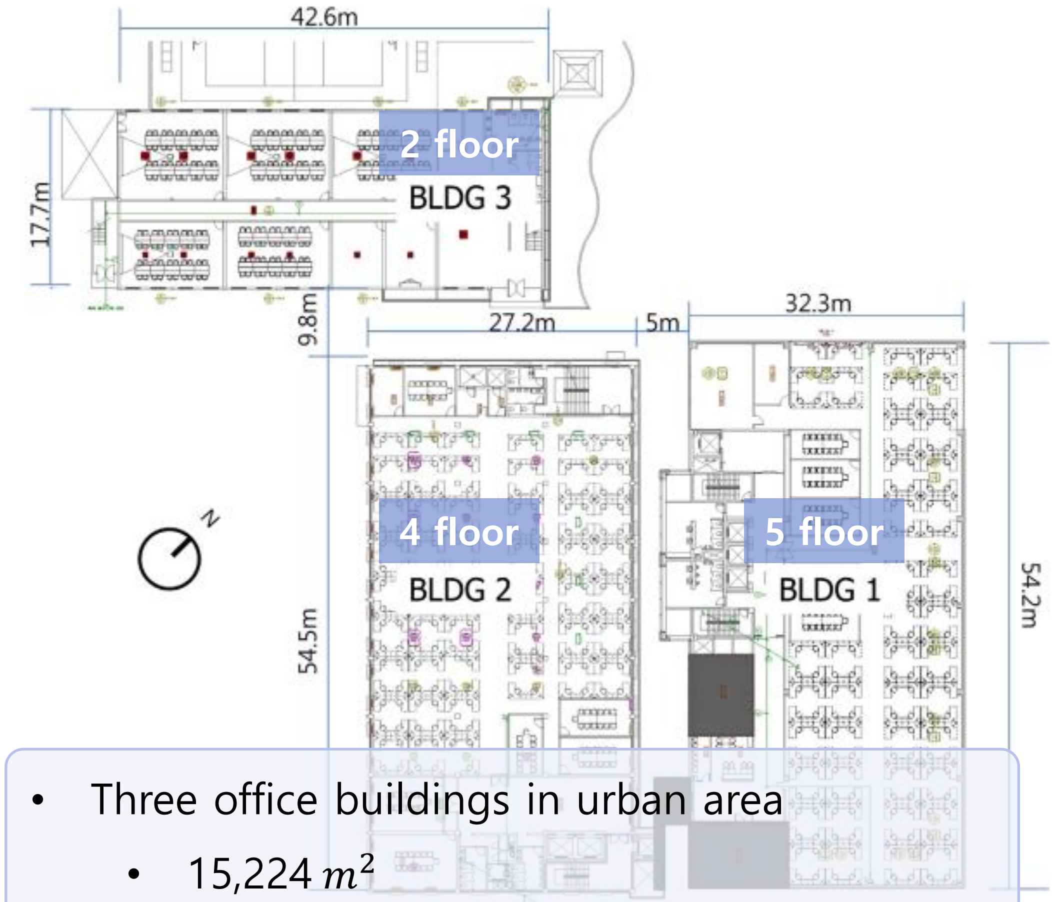
- Requires an entire whole year's data set
- Not be stable when the energy consumption pattern has large daily or annual variability

⇒ **A short-term load forecasting model** using data mining and machine learning technique while assuming limited availability of data.

⇒ ANN models – commercial building complex

❖ Methodological approach

1. Description of a case study: a building complex



- Three office buildings in urban area
  - 15,224 m^2
  - All managed by one utility billing system

- Two absorption chiller systems
  - Constant air volume (CAV) – BLDG 1
  - Fan coil units (FCU) – BLDG 2
- Electric heat pump (EHP) – BLDG 1, 2, 3
- Building management system (BMS)
  : monitor operational conditions

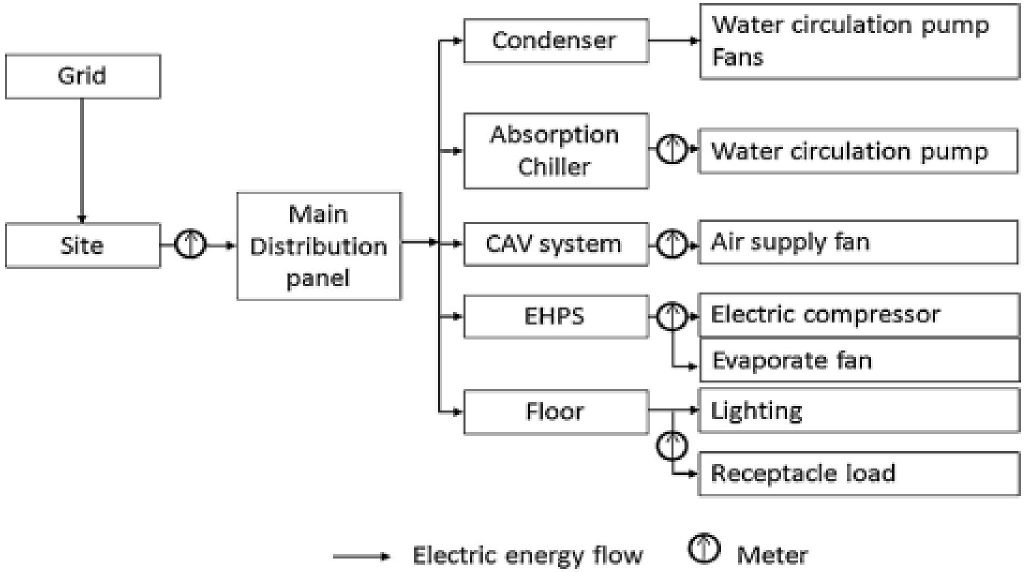


Fig. 2. Electricity delivery hierarchy and meter installation at the site.

❖ Methodological approach

2. Data collection and processing

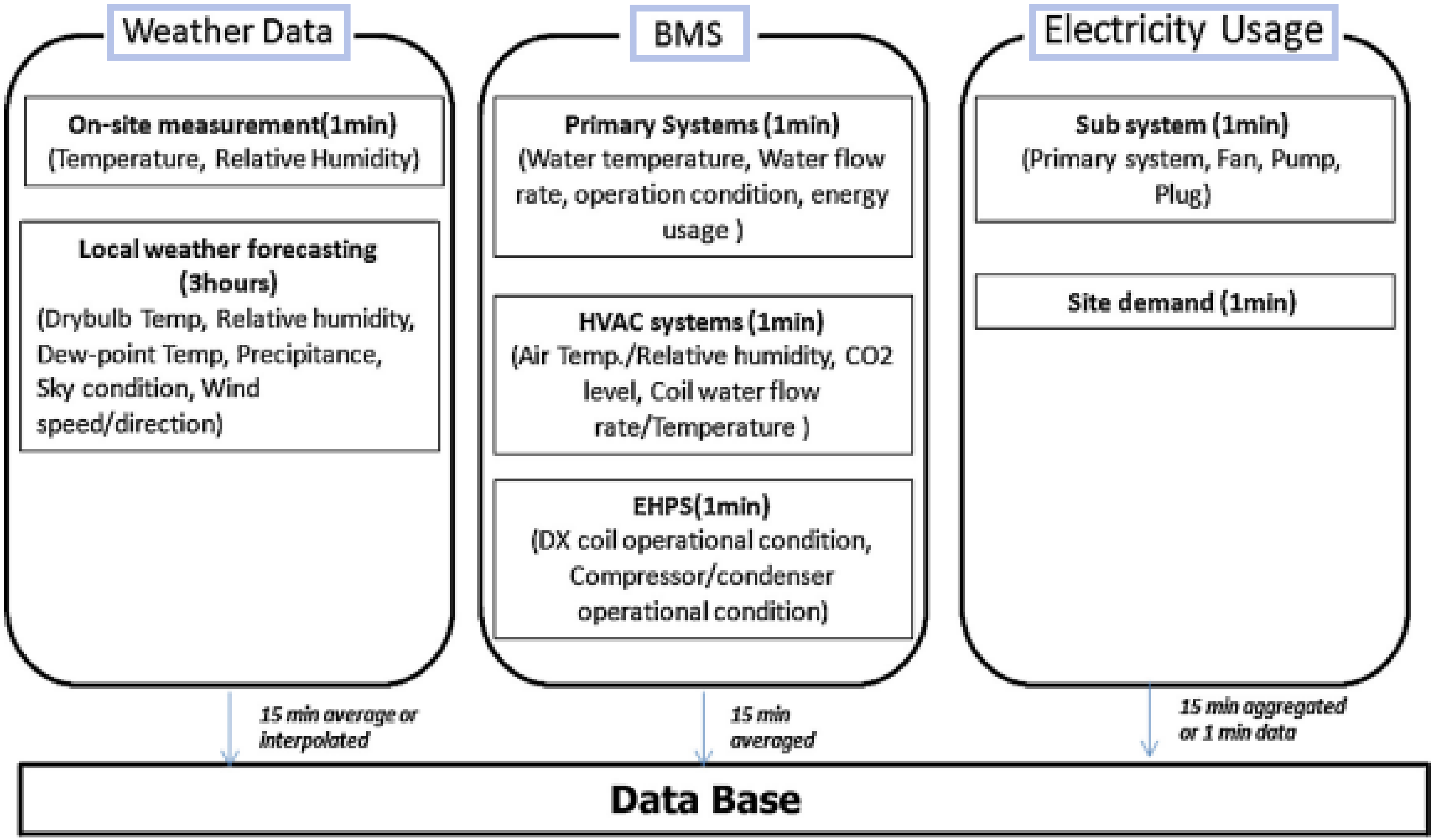


Fig. 3. Data collecting system.

Methodological approach

2. Data collection and processing

> Potential predictor variables

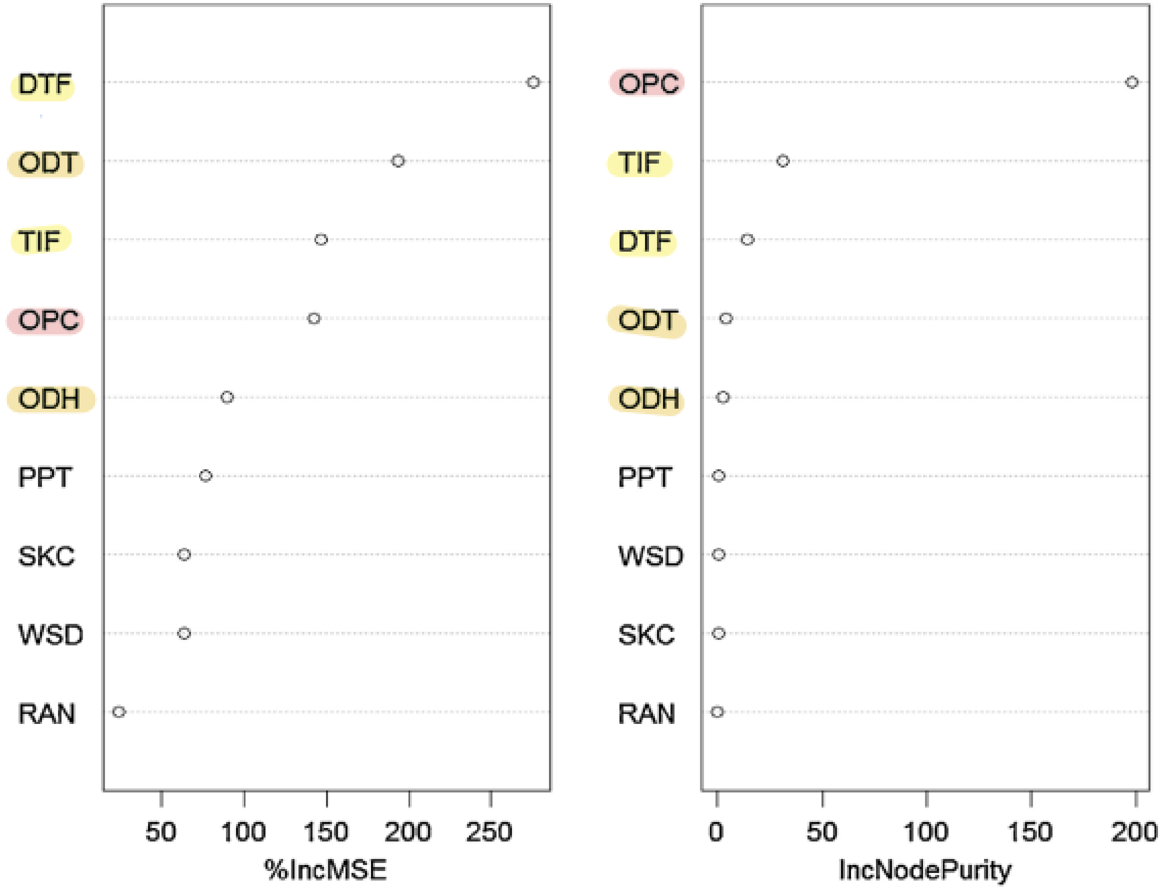
Category	Variables	Unit/index
Environment	Outdoor dry-bulb temp. (ODT)	°C
	Outdoor relative humidity (ODH)	%
	Precipitation probability (PPT)	%
	Rain indicator (RAN)	0: No rain 1: Rain 2: Snow 3: Rain and snow
	Wind speed (WSD)	km/h
	Sky condition (SKC)	0: Clear sky1: Partially cloud2: Cloud3: Overcast
Time indicator	Day indicator (DTF)	0: Weekdays 1: Saturday 2: Sunday
	Interval stamp (TIF)	0-95
Operational condition	HVAC operation schedule(OPC)	Discretized

- Environmental data
  - ODT, ODH, PPT, RAN, WSD, SKC
- Time indicator
  - DTF, TIF
- Operational condition
  - OPC

9 potential variables

❖ Methodological approach

3. Feature extraction for dimension reduction



(a) Permutation importance **(b) Gini importance**

Variables	Conditioned permutation-importance
TIF	482.27
DTF	354.03
OPC	296.50
ODT	222.01
ODH	212.01
PPT	92.56
WSD	90.42
SKC	81.95
RAN	32.36

Random forests algorithm

Variables
Outdoor dry-bulb temp. (ODT)
Outdoor relative humidity (ODH)
Precipitation probability (PPT)
Rain indicator (RAN)
Wind speed (WSD)
Sky condition (SKC)
Day indicator (DTF)
Interval stamp (TIF)
HVAC operation schedule(OPC)

5 variables

+ Previous electricity usages

❖ Methodological approach

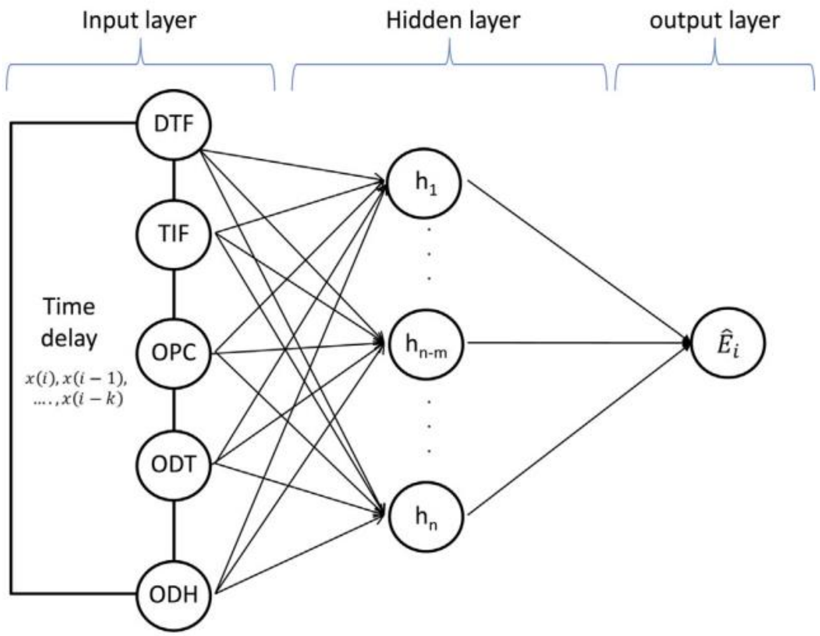
4. Predictive model selection

Table 4
Evaluation of different machine-learning algorithms.

Algorithm	Correlation coefficient	CV(RMSE)
Gaussian process with radial basis function (RBF) kernel	0.94	0.11
Gaussian process with polynomial kernel	0.87	0.15
Linear regression	0.83	0.16
Artificial neural network	0.96	0.08
Support vector machine (SVM) with normalized polynomial kernel	0.92	0.13
SVM with RBF kernel	0.88	0.14
K-Star classifier	0.92	0.12
Nearest neighbour ball tree	0.94	0.11
Simple model	0.81	0.18

> 9 machine-learning algorithms

5. ANN model architecture



- 15-min interval data set E_i(energy consumption)
- 5 predictor variables (July 1st ~ July 31st, 2012)

  ⇒ Train : 2880 data points for 6 input variables

❖ Results and discussion

1. Network design parameters - Hidden layer architecture and time-delayed inputs

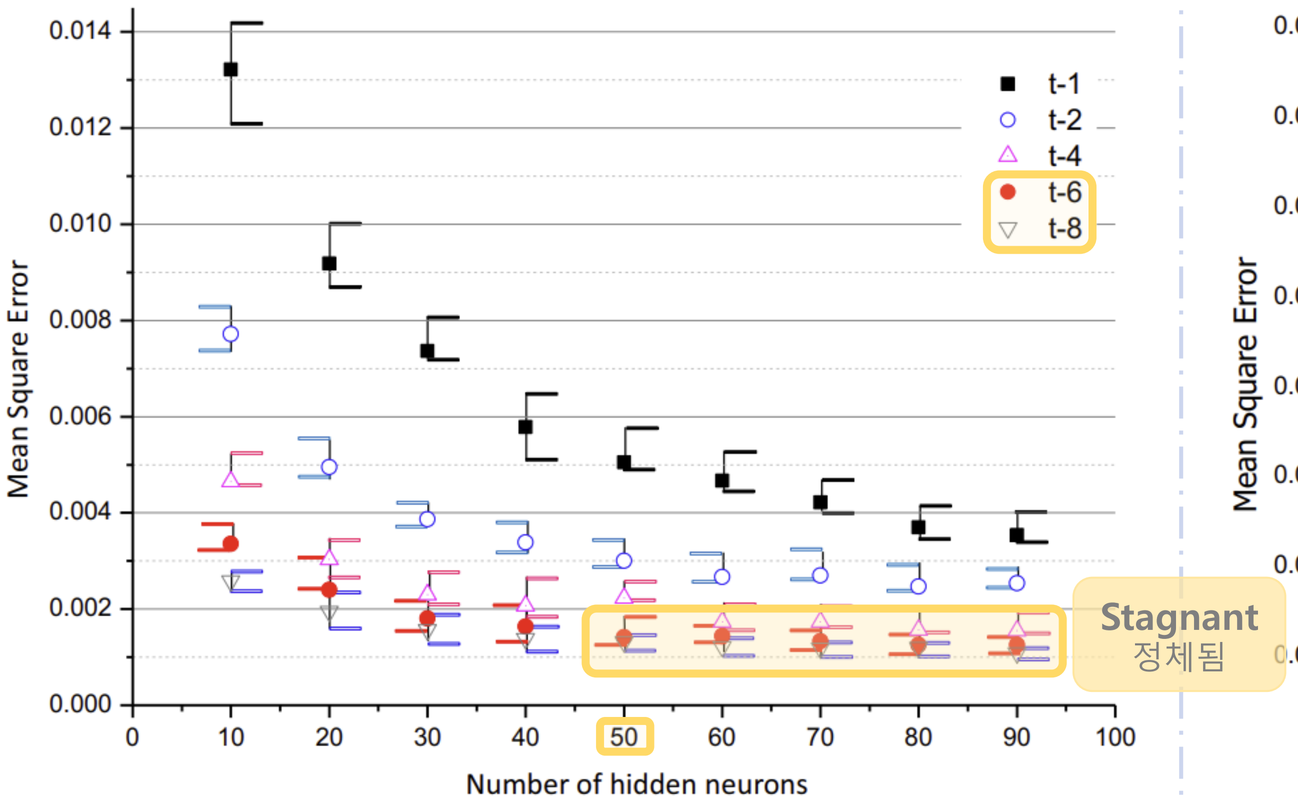


Fig. 6. Average MSE in the training with neuron numbers and time delays.

Training

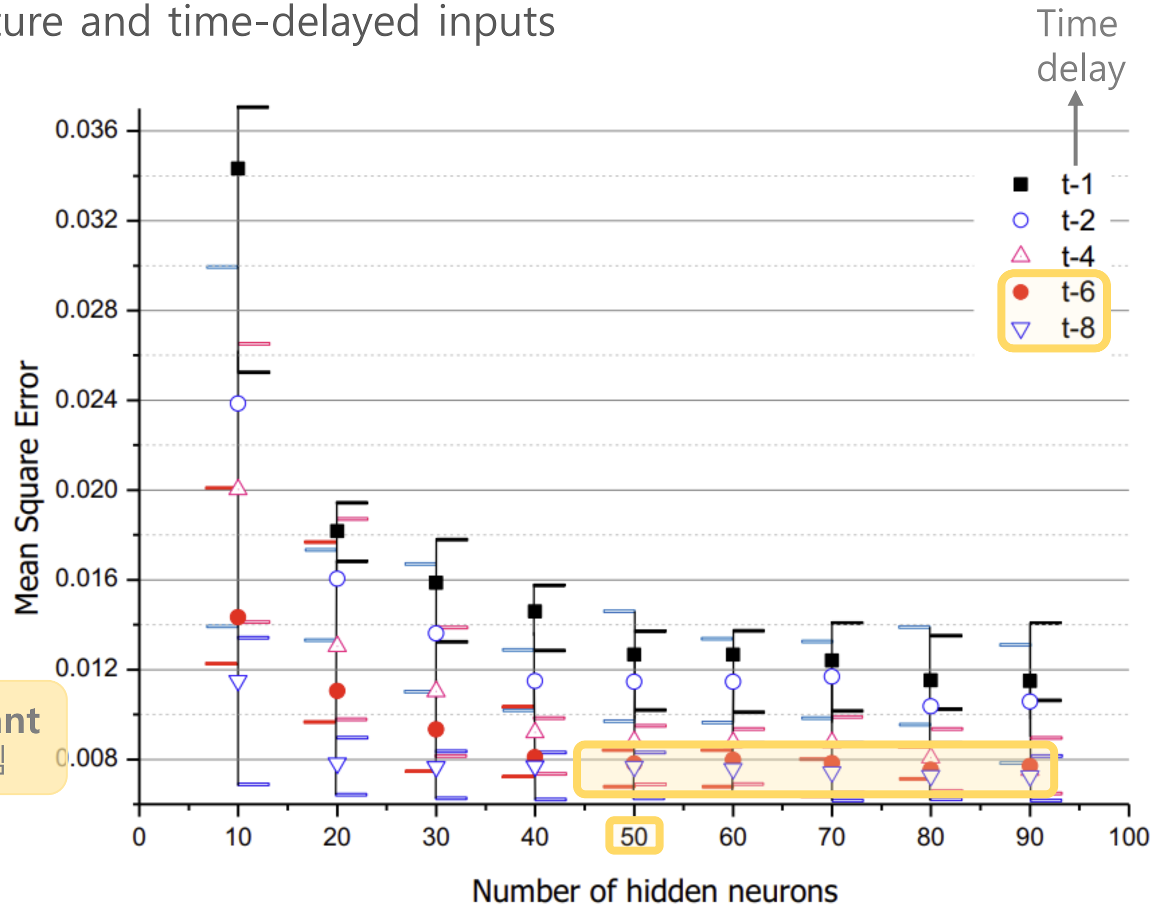


Fig. 7. Average MSE in the evaluation with neuron numbers and time delays.

Evaluation

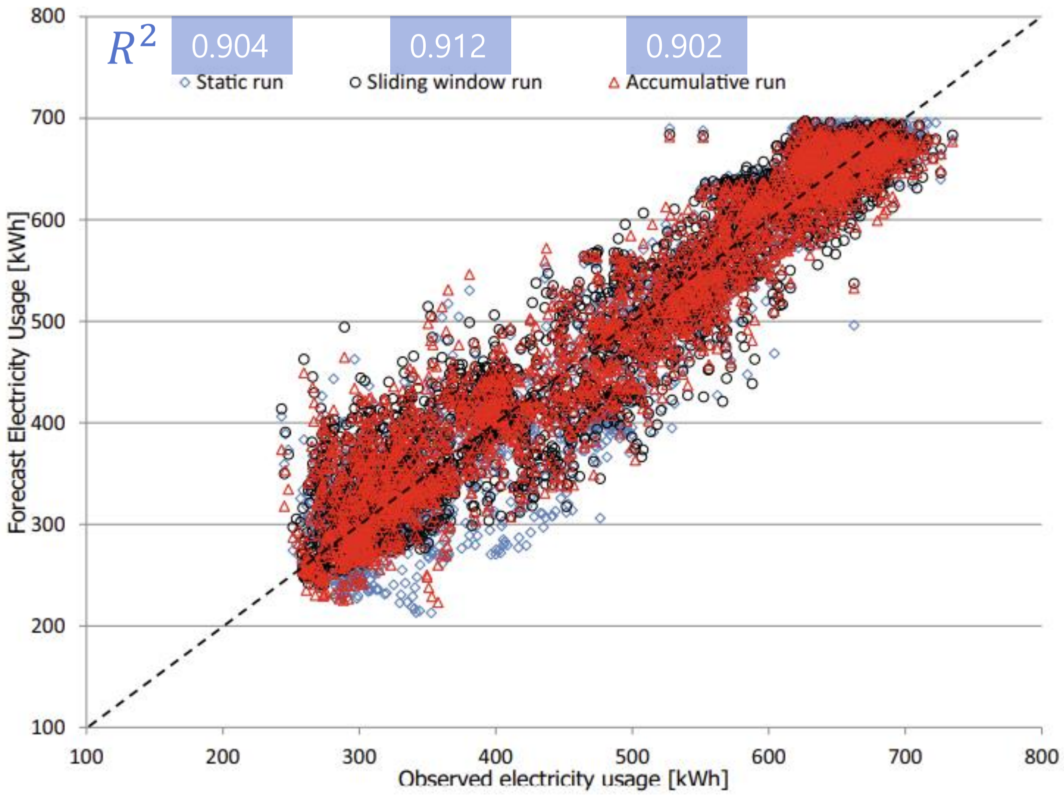
⇒ 50 neurons & 6h time delay

❖ Results and discussion

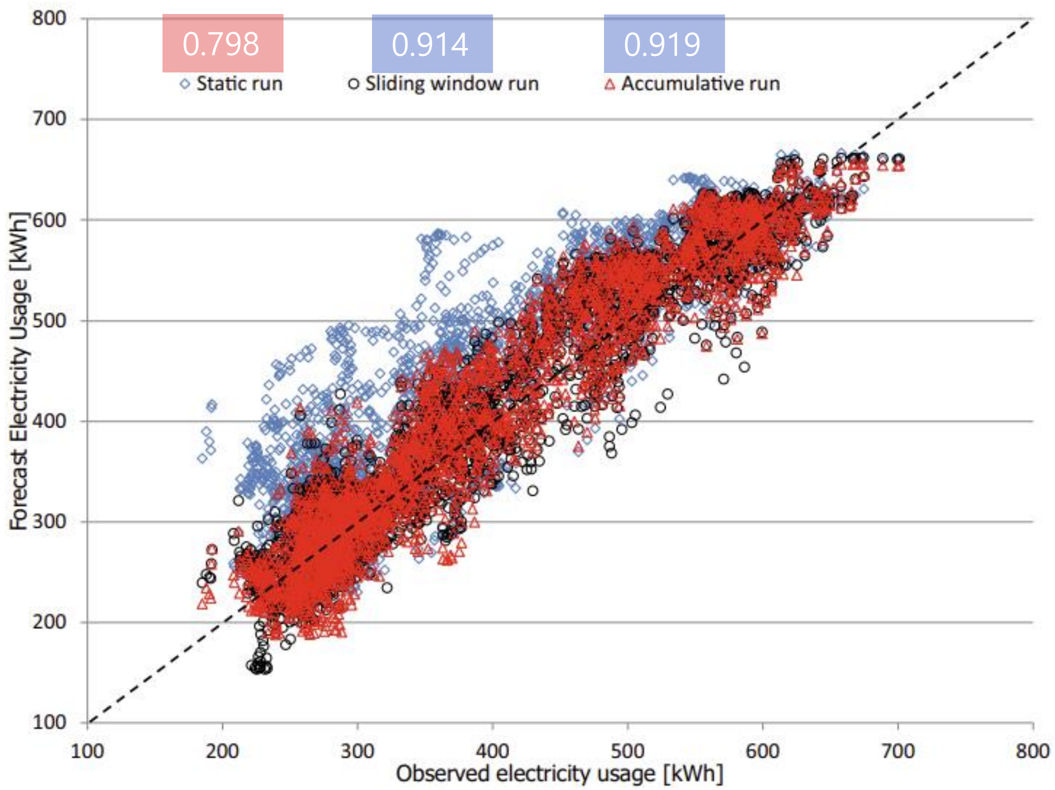
2. Model implementation

Training method
- Static
- Accumulative
- Sliding windows

> Comparisons of actual observed and forecasted electricity usage



August

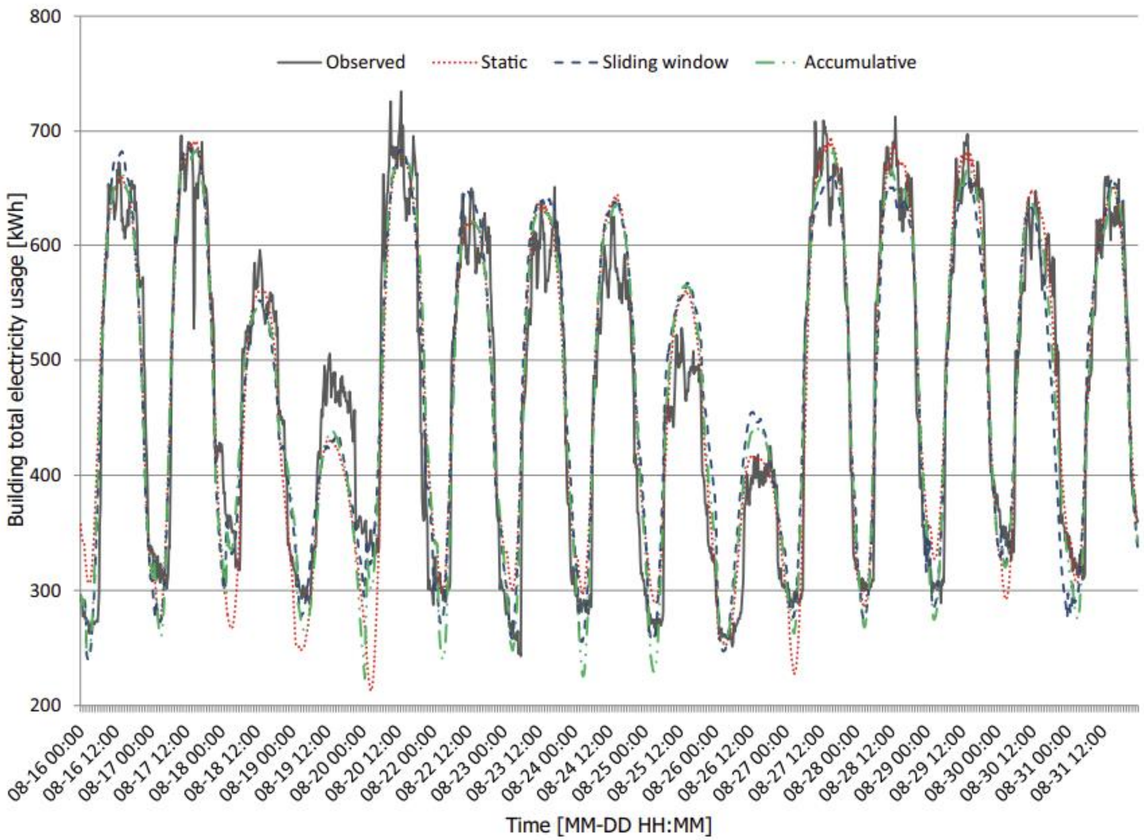


September

❖ Results and discussion

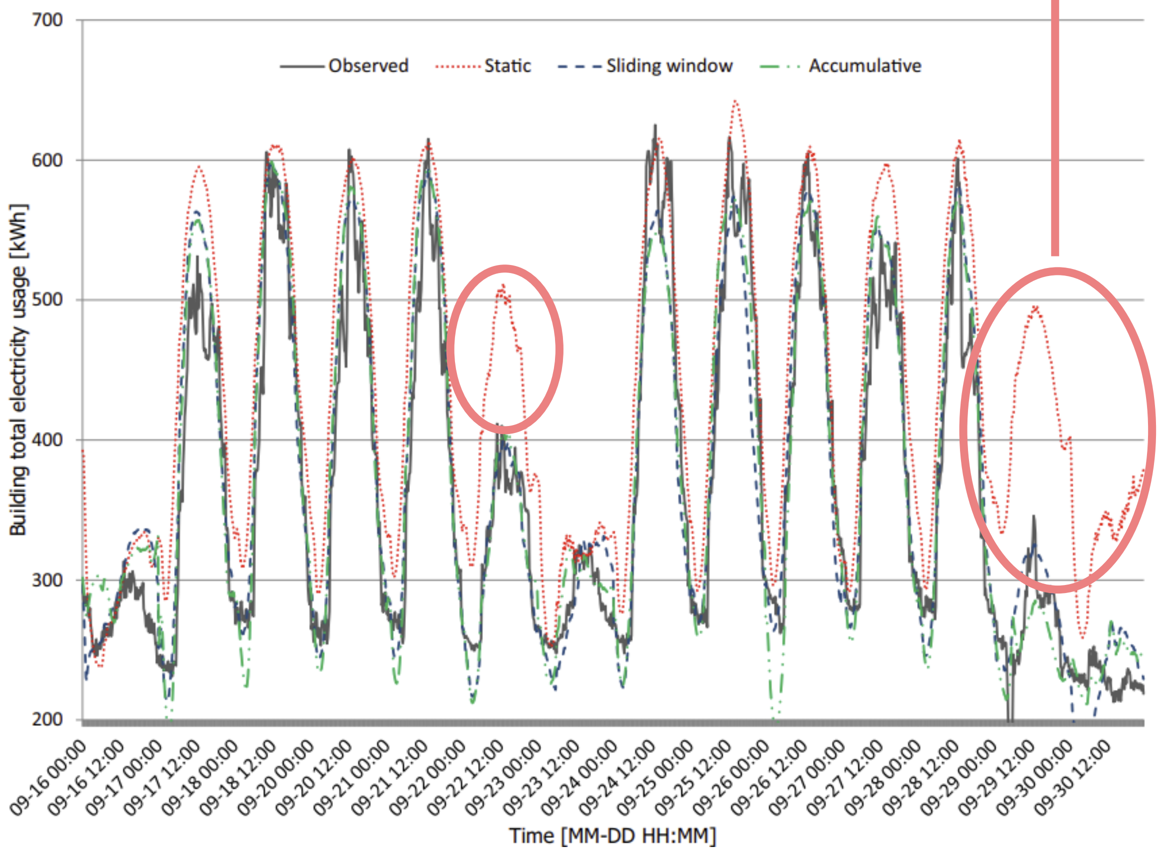
2. Model implementation

> Actual forecast result



August

Over estimated



September

❖ Results and discussion

2. Model implementation

Daily

Table 6
15 min prediction results with each training type.

Month	Day type (days)	Static		Accumulative		Sliding window	
		CV(RMSE) [%]	σ	CV(RMSE) [%]	σ	CV(RMSE) [%]	σ
August	Weekday (21)	8.44	2.41	7.97	2.53	8.74	2.53
	Weekend (8)	11.16	3.61	9.91	2.11	9.62	2.31
September	Weekday (19)	13.76	5.03	9.35	2.14	9.20	2.25
	Weekend (10)	26.74	11.97	11.06	2.44	11.20	3.60

Daily peak

Table 7
Daily peak prediction results with each training type.

Month	Day type (days)	Static			Accumulative			Sliding window		
		Max/min APE [%]	Averaged APE [%]	σ	Max/min APE [%]	Averaged APE [%]	σ	Max/min APE	Averaged APE [%]	σ
August	Weekday (21)	7.57/0.16	2.84	2.79	7.62/0.05	3.37	2.38	8.58/0.02	3.30	2.81
	Weekend (8)	20.33/2.0	8.77	6.67	13.28/3.03	7.28	3.78	13.67/3.75	8.10	2.83
September	Weekday (19)	13.17/0.2	4.40	3.92	11.6/0.2	4.52	2.71	9.7/0.6	4.26	2.64
	Weekend (10)	59.4/4.4	24.9	1.77	17.4/0.6	6.48	3.01	7.7/0.8	4.25	2.80

❖ Conclusion

- **The short-term load forecast (STLF)** – sub-hourly electricity consumption of buildings
  - Still a challenging problem – complexity of usage pattern & highly noisy input data
- A new approach – using feature extraction and an ANN

  → **A-day-ahead forecast of the electricity usage for a commercial building complex** (15 min)

✓ Developed model for 2 months

- Daily error : stable around averaged CV of 10%
- Daily peak demand : averaged APE of 5%

⇒ **The model can provide a day-ahead electricity usage profile with sub-hourly intervals and daily peak electricity consumption with a reasonable accuracy**

- A good predictive model of energy consumption in buildings

  → Useful in developing a good model predictive control (MPC) method that can reduce energy costs in buildings.

3. A hybrid model approach for forecasting future residential electricity consumption

Dong B, Li Z, Rahman SMM, Vega R.
Energy Build 2016;117:341–51..

❖ Introduction

- **Data-driven - Black-box** : machine learning driven, derived by measured data
  - Limited in one type of methods (neural networks)
- **Physical + Data-driven = Grey-box**
  - Built upon physical relationship – parameters of physical model are unknown or uncertain.
  - + Measured data → identify those parameters, model tuning for better accuracy

- **Commercial vs Residential building**

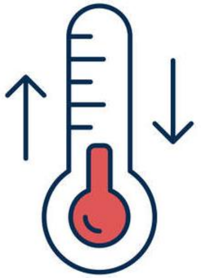
1. The lack of hourly or more granular data
2. The growing percentage of miscellaneous electrical load (MEL) in residential building (~52%)

기타 전기부하 비율

⇒ **A new hybrid modeling approach**

⇒ Forecast hour ahead and 24 ahead residential electrical load

❖ Current state of the art



- **Data-driven models** – machine learning algorithms
  - ANN, SVR, LS-SVM, GPR, GMM
  - Only conducted on **total** building energy consumption forecast
- **Physical models (Forward models) -** Use laws of thermodynamics (열역학)
  - 2 main parts : Building thermal zone envelopes / Air conditioning modeling
  - Physical phenomena (convection, conduction, radiation) → 1 layer
  - All layers (walls, floor, roof, etc.) → Thermal envelope zone (thermal resistance-capacity (RC) networks)
  - Suffers from a lack of building information
  - Limited to **thermal usage** prediction such as **air conditioning**

⇒ **A new hybrid modeling approach**

⇒ Forecast both total and air-conditioning (AC) energy consumption

❖ Methodology and approach

◆ Test bed and performance metrics

Fig. 1. Four test houses.

- **4 residential houses** in San Antonio (110m^2 each)
  - SIP (Structure Insulated Panel)
  - ACC (Autoclaved Aerated Concrete)
  - Container
  - Stick

- **Material information**
  - Calibrated thermal resistances and capacities
    – manufactures (119mm SIP U-value is 0.22W/m^2K)
  - Energy consumption – 5 min intervals for all rooms
  - Outdoor air temperature, global horizontal solar radiation
    – weather station near the position
  - Solar radiations on the surface – sensors installed on site

- Hour ahead, 24-h ahead models – evaluated

$$\mathrm{MAPE}(\%) = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|\overset{\text{actual}}{y_i} - \overset{\text{predicted}}{y_p}\right|}{y_i} \times 100$$

⇒ % of error per prediction
⇒ Smaller is better

$$\mathrm{CV}(\%) = \frac{\sqrt{1/N-1\sum_{i=1}^{N}(y_i - y_p)^2}}{\bar{y}} \times 100$$

⇒ Variation of overall prediction
⇒ Smaller, more similar dispersions

❖ Methodology and approach

◆ Hybrid model

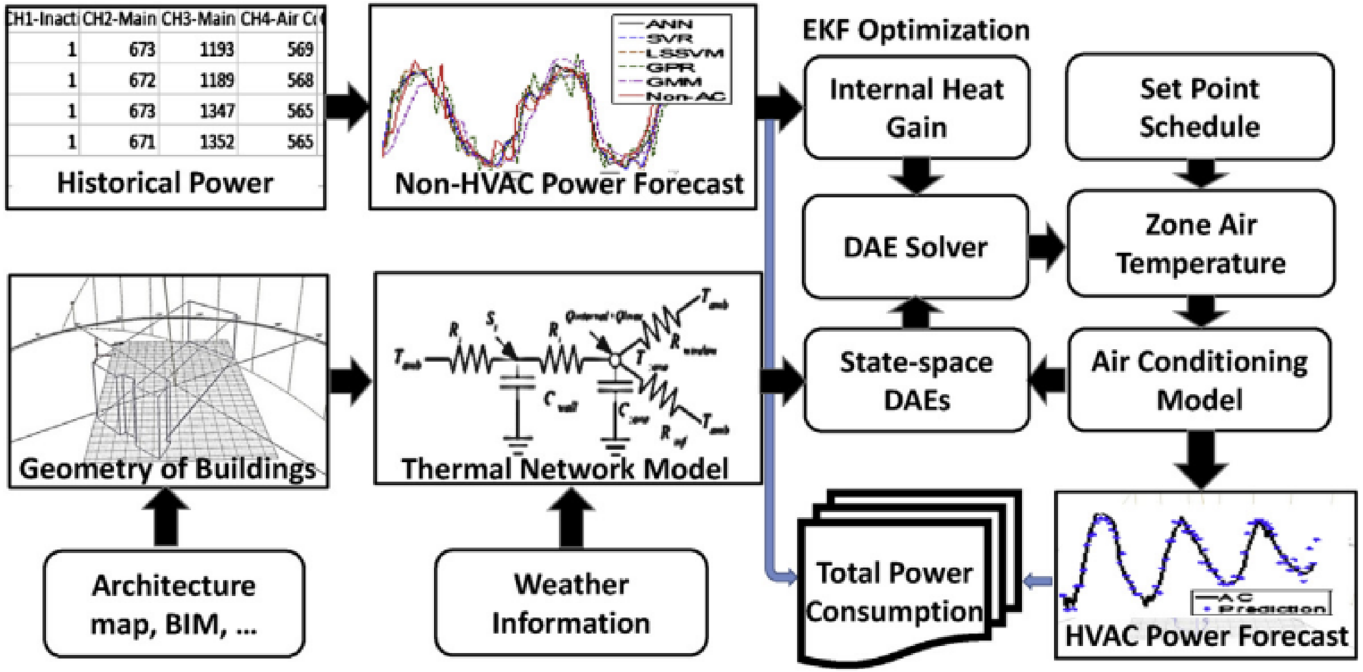


Fig. 3. Hybrid approach.

1. Historical information
   → Forecast non-AC electricity consumption
   → Internal heat gain
2. Predicted weather + internal gain forecast
   → The thermal network differential algebraic equations (DAEs)
   → Zone temperature + Set point schedule
   → AC regression model

➢ **AC + non-AC**
= a total electricity consumption

❖ Methodology and approach

◆ Hybrid model

- Non-AC forecast

➢ Hour ahead

$$\mathrm{H1}: f(t, L_{t-1}, L_{t-2}, \ldots, L_{t-5})$$

- Previous 1-5 h historical load information

➢ 24-hour ahead

$$\mathrm{H2}: f(t, L_{t-24}, \ldots, L_{t-31}, \mathrm{avg}(L_{t-24}, L_{t-25}), \mathrm{avg}(L_{t-24}, L_{t-25}, L_{t-26}), \ldots, \mathrm{avg}(L_{t-24}, \ldots, L_{t-31}))$$

- Previous 24-31 h historical load + moving averages

- AC forecast

$$P = \frac{m \sum_{i=1}^{n} q_i}{\mathrm{COP}}$$

- P : electricity consumption forecast
- q_i : minutes resolution AC cooling or heating load in the forecast time window
- COP : coefficient of performance
- m : time scale

(if q = 5 min temporal load, m=1/12)

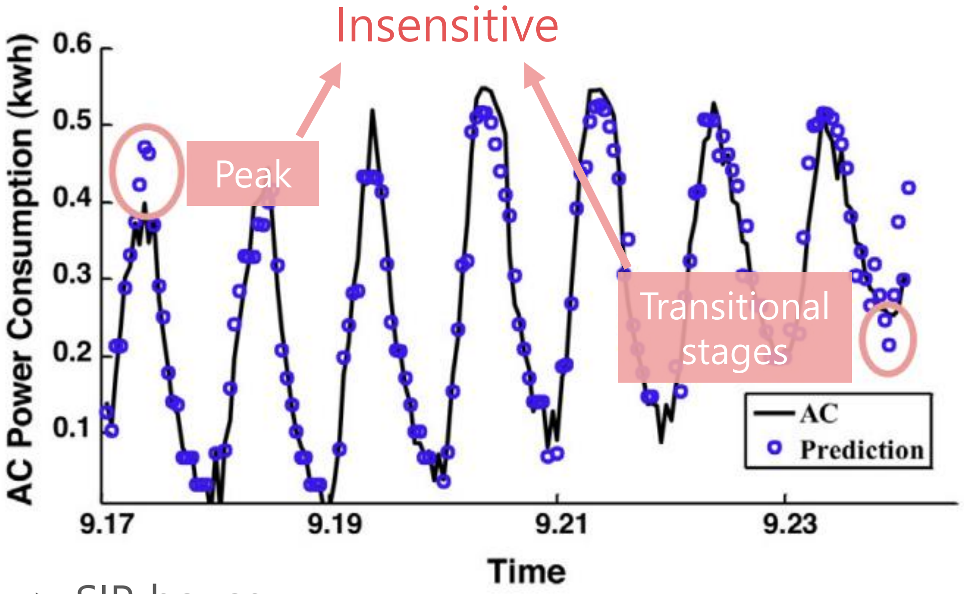
❖ Results and discussions

- ➢ Train - 1 week ~ 1 month before tested period (before 2013.9.17)
- ➢ Test - 1 week (2013.9.17 ~ 2013.9.24)
- ➢ AC / non-AC / total building energy consumption

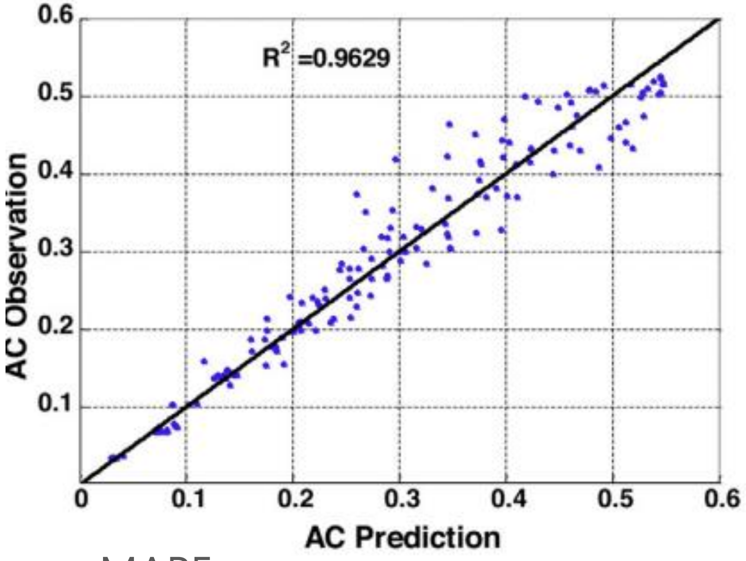
◆ Air-conditioning energy consumption

 # MAPE

- SIP : 7.03%
- ACC : 8.29%
- Container : 9.15%
- Stick : 8.03%



> SIP house

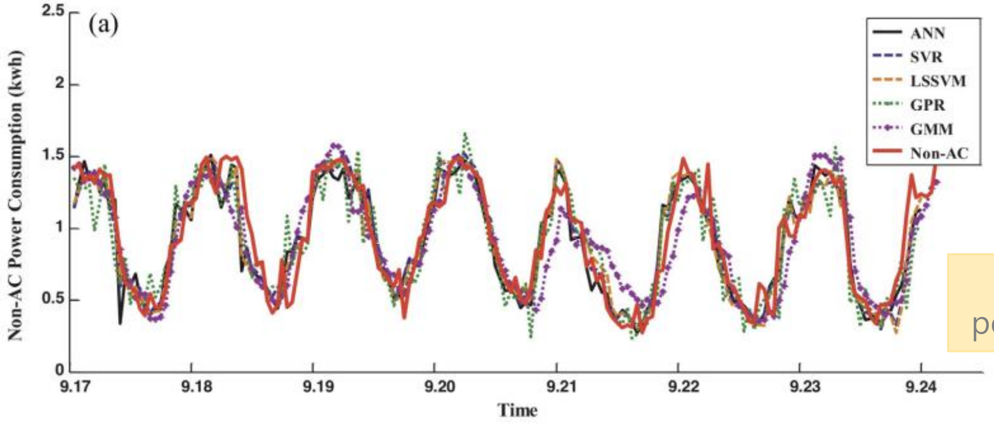


> MAPE

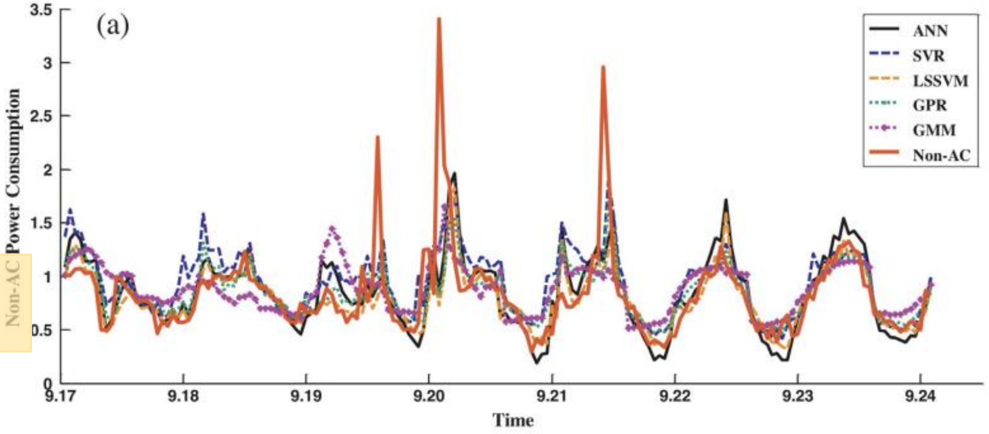
❖ Results and discussions

◆ **Non-air-conditioning** energy consumption

Hour

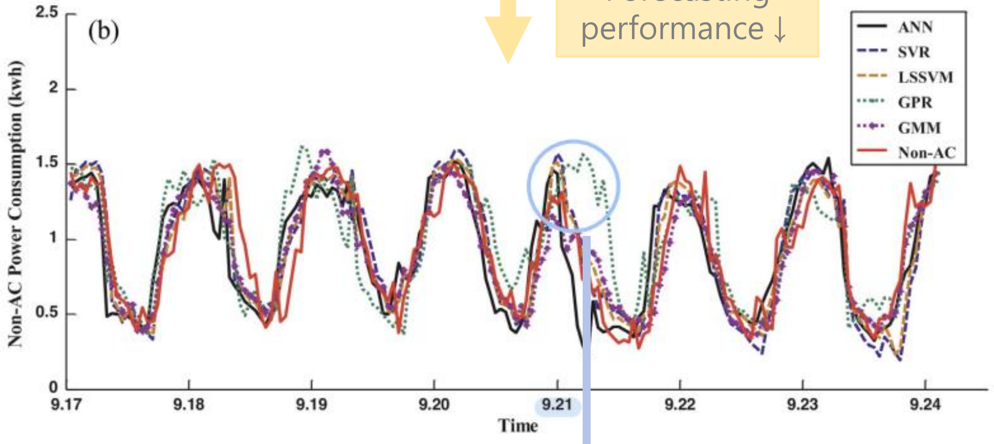


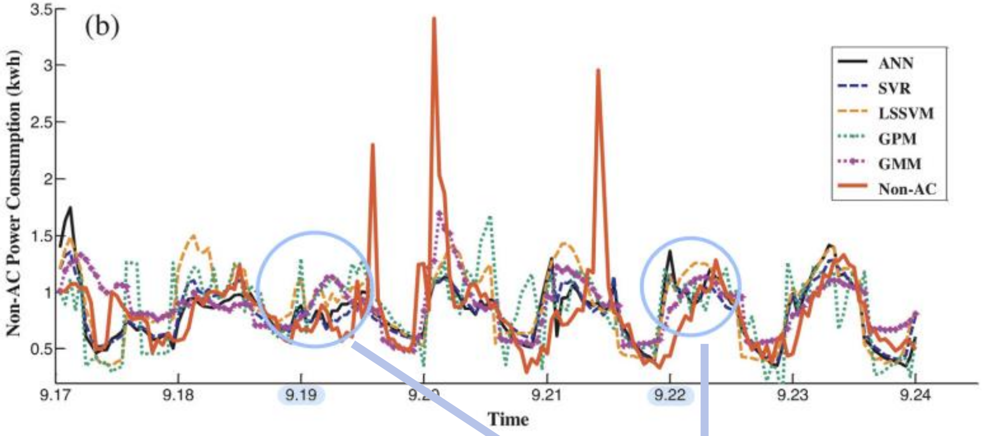
Forecasting performance ↓



Forecasting performance ↓

24 h





> **Sinusoid shape**

Over fitting

> **Irregular shape**

Insensitive

❖ Results and discussions

◆ **Non-air-conditioning** energy consumption

Reduced performance due to

✓ Constraints on **input feature**

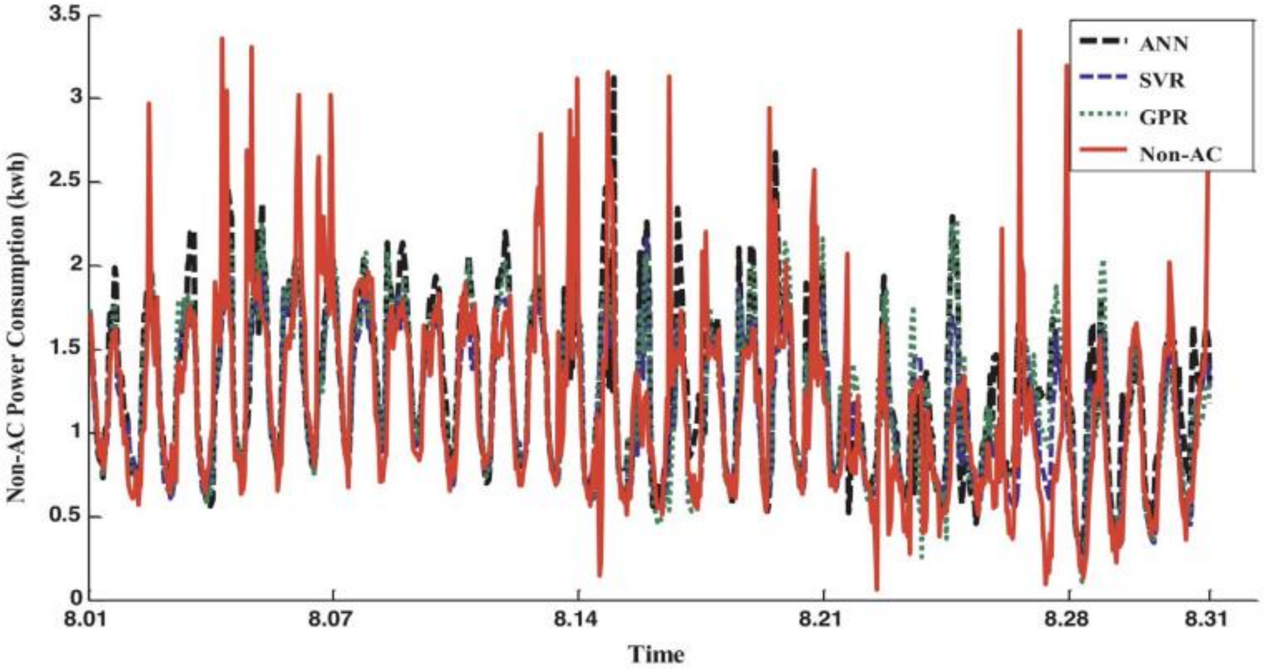
- Hour : previous 1~5 h
- 24-h : yesterday

→ Key assumption :

Daily usages are similar without large deviation

✓ Random behavior of residents creates the randomness

> **1 month** of SIP

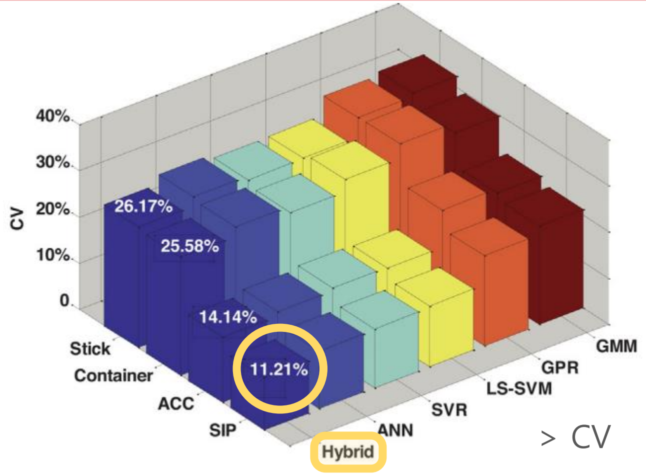


Inability to capture the random spikes

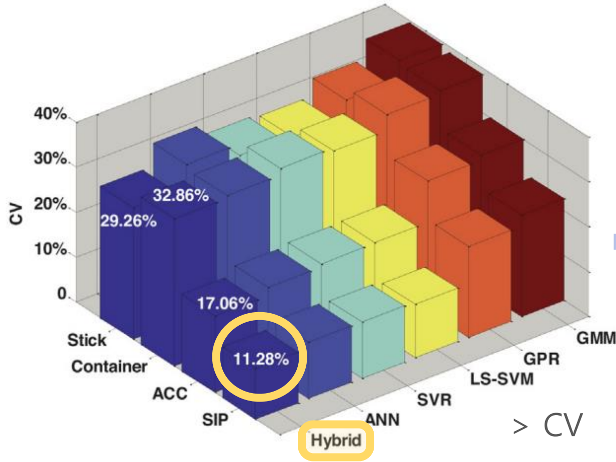
→ Short period of time (5 min)

❖ Results and discussions

◆ Total building energy consumption



> CV



> CV

Temporal variance has a significant impact on the predictive capability

> Hour ahead

Model	MAPE	CV	Model	MAPE	CV
House: SIP			House: container		
Hybrid	8.16%	11.21%	Hybrid	21.44%	25.58%
ANN	9.26%	12.49%	ANN	22.80%	27.54%
SVR	9.58%	12.76%	SVR	22.11%	26.07%
LSSVM	9.09%	13.18%	LSSVM	23.26%	28.70%
GMM	14.29%	19.50%	GMM	28.44%	31.91%
GPR	15.86%	21.39%	GPR	26.24%	29.75%
House: ACC			House: stick		
Hybrid	12.60%	14.14%	Hybrid	21.51%	26.17%
ANN	12.50%	15.08%	ANN	23.86%	28.19%
SVR	12.70%	15.72%	SVR	23.22%	27.28%
LSSVM	12.94%	15.47%	LSSVM	23.44%	27.42%
GMM	19.21%	23.50%	GMM	27.22%	31.60%
GPR	18.57%	22.62%	GPR	30.11%	32.27%

> 24-hour ahead

Model	MAPE	CV	Model	MAPE	CV
House: SIP			House: container		
Hybrid	10.02%	11.28%	Hybrid	26.43%	32.82%
ANN	11.92%	12.59%	ANN	27.39%	33.35%
SVR	11.97%	12.42%	SVR	28.25%	34.91%
LSSVM	10.39%	11.83%	LSSVM	28.95%	34.19%
GMM	17.80%	20.24%	GMM	34.01%	37.81%
GPR	18.85%	22.69%	GPR	35.94%	38.57%
House: ACC			House: stick		
Hybrid	14.36%	17.06%	Hybrid	27.52%	29.26%
ANN	14.41%	18.85%	ANN	30.22%	34.40%
SVR	15.69%	19.55%	SVR	28.56%	31.87%
LSSVM	15.10%	19.72%	LSSVM	28.32%	31.36%
GMM	25.47%	28.85%	GMM	31.13%	35.15%
GPR	27.12%	30.03%	GPR	35.76%	39.34%

❖ Results and discussions

◆ Discussions

Table 3
Hour ahead forecast performance of recent studies.

Reference	Year	Type	Location	Method	Forecast	Performance
Ghofrani et al. [52]	2011	Dwelling	Navada	Kalman filter	Hour ahead	MAPE: 12–30%
Edwards et al. [1]	2012	Two-floor house	Tennessee	Linear regression various types of ANN, SVR, LSSVM	Hour ahead	MAPE: 9–30%; CV: 11–38%
Jain et al. [38]	2014	Multi-family	New York	SVR	Hour ahead	CV: 10–11%

- **3 previous research** for hour ahead total energy consumption forecasting of residential buildings
  - MAPE : 9% ~ 30%
  - CV : 10% ~ 38%
  - Our hybrid approach is slightly better than other existing models

Hybrid approach

- Improvements – decomposition of total building energy consumption
  - Isolate certain appliances (AC)
- Simultaneous prediction 동시 예측
  – both **total** & **sub-meter appliances** (AC)

❖ Conclusion

- An innovative **hybrid** modeling approach for **residential** building energy consumption forecasting
  - Data-driven + physic-based models
  - Single-family residential houses
  - Validated through 1 month measured data – 4 residential buildings

- Hybrid modeling approach
  - Hour ahead : slightly better
  - 24-h ahead : similar
- AC prediction - 5 min interval data

  → Not typically available to homeowners and utility companies

✓ Future
- Using hourly data
- Improve 24h ahead forecasting

Thank You! 😊 🎉