EC-ASEAN Energy Facility (EAEF)

Projects 64 and 68
Commencement Meeting

Development of Baselining Methodologies In Singapore

Bing Dong
Energy Sustainability Unit
Department of Building
National University of Singapore

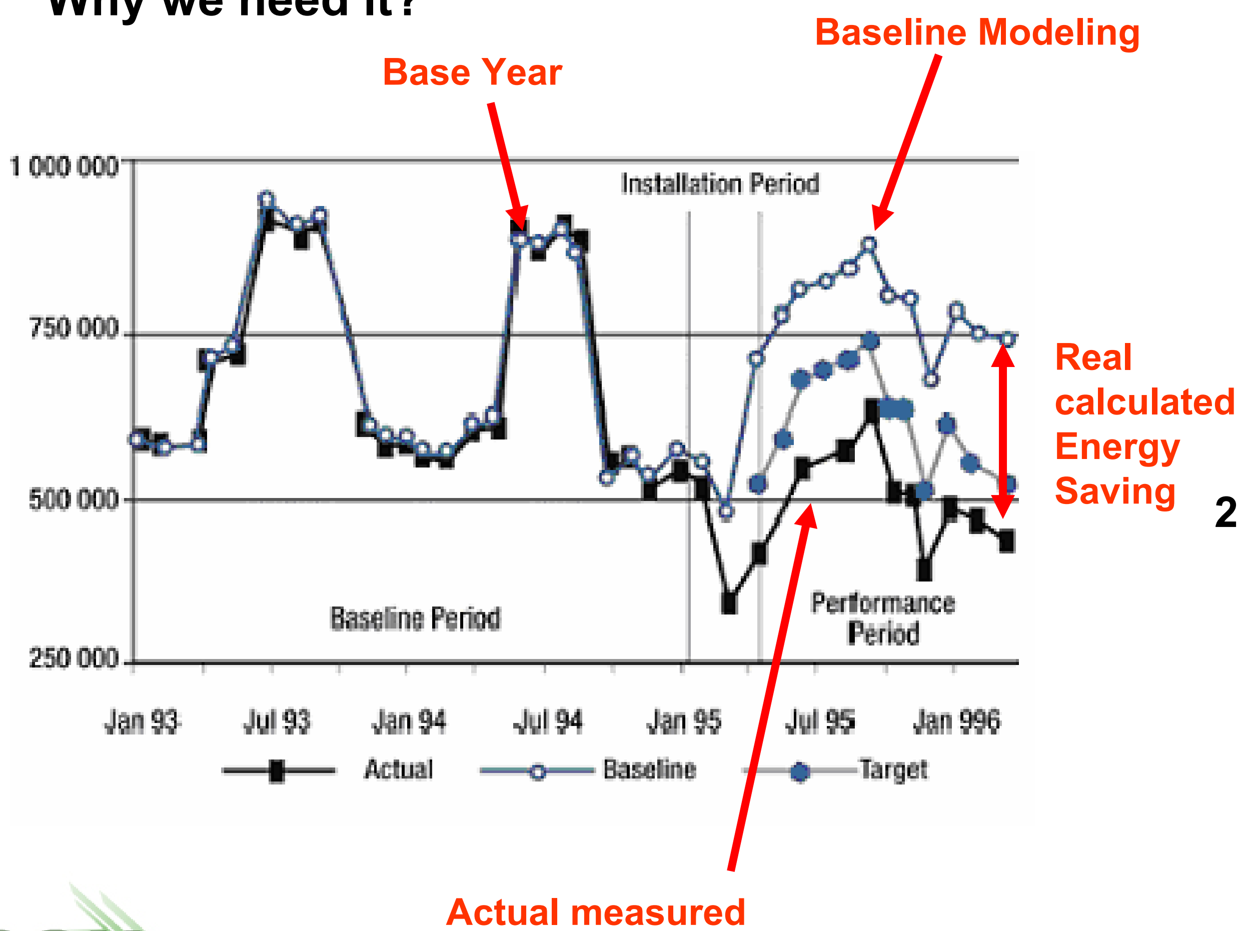
Outline

1. **Introduction**
2. **Baseline Models for Whole Building Energy Consumption**
3. **Baseline Models for Landlord Energy Consumption**
4. **Results and Discussion**
5. **Conclusion**

Introduction

What is a Baseline Model?

Why we need it?



1. A methodology to verify savings from energy conservation programs. It helps to compute the consumption of the building at any period, assuming that the building has not been retrofitted.

2. Since everything keeps changing (weather data, O&M etc.), direct comparison between measured energy consumption is not REAL calculated energy saving

Introduction

How to develop and use a baseline model?

1. A good record of base year conditions

- **Energy and demand profile (utility bills or spot measurement)**
- **User or occupancy density**
- **Equipment inventory and operation methods and periods**

2. Select the critical indicators of changing of energy consumption.

- **Weather data**
- **Occupancy rate or density**
- **Operation hours, etc.**

3. Select the proper baseline methodologies

4. Continuous recording and adjustments in the post-installation period

introduction

Existing Models

Models	Examples	Strength	Weakness
Statistical Regression models	Linear, Multiple-linear, Change-point, degree-day etc.	Easy to establish Most commonly used	Fair accuracy
Computer simulation	DOE-2, EnergyPlus, BEST,etc.	High accuracy, detailed description	Time consuming, Large building information
Other models	Neural Network, Fourier Series, Support Vector Machines (SVMs), etc.	High accuracy	Large pool of hourly data

Introduction

How to asset a good Baseline Model?

Model Performance/Prediction Assessment Criteria (1)

1. Mean Absolute Error (MAE) (Rough Estimation)

$$MAE(\%) = \frac{\overline{E}_{predicted} - \overline{E}_{measured}}{\overline{E}_{measured}} \times 100$$

2. Coefficient of Variance (CV) ***(Using both in model development and prediction)***

CV-RMSE= $\frac{RMSE}{\overline{E}}.100$

$$RMSE = [MSE]^{1/2} = \left[\frac{\sum_{i=1}^{n} \left(E_i - \hat{E}\right)^2}{n-p} \right]^{1/2}$$

Introduction

How to asset a good Baseline Model?

Model Performance/Prediction Assessment Criteria (2)

3. Variance of forecast error (VAR) ***(Using in Prediction)***

$$VAR(\hat{E}-E)=S^2\left[1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right] \qquad S=\sqrt{\frac{\sum e_i{}^2}{n-p}}$$

4. 90% uncertainty bands ***(Using in Prediction)***

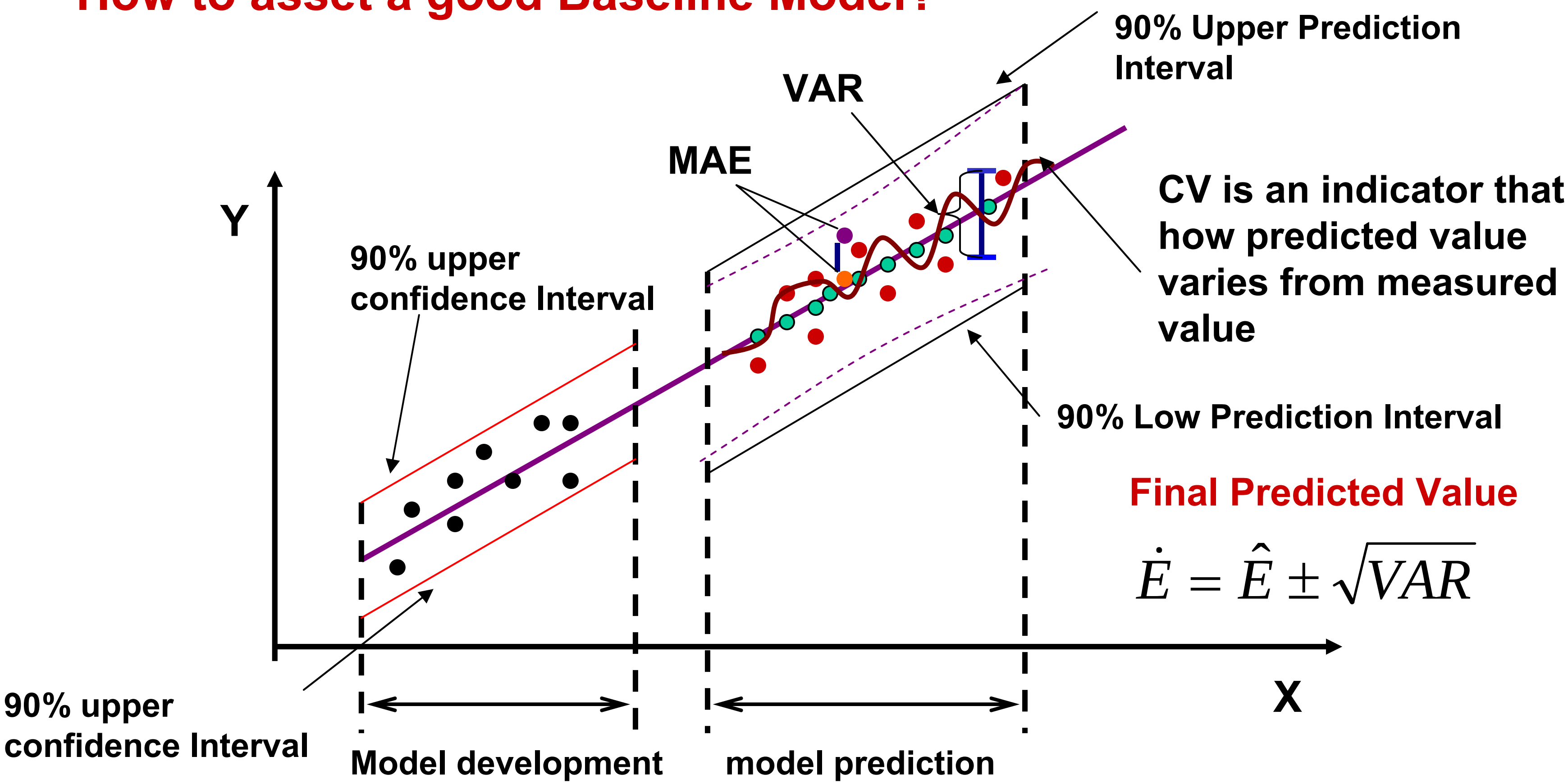
$$PI=t\left(1-\frac{\alpha}{2},n-p\right)\cdot\left\{\frac{MSE}{m}\cdot\left[m+\frac{m}{n}+\frac{\sum_{0=1}^{m}(x_0-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}\right]\right\}^{1/2}$$

n = number of observations
m = number of month
p = number of parameters in the model

If 90% PI adopted, 9 out of 10 times the predicted value E will be between (E+PI) and (E-PI)

Introduction

How to asset a good Baseline Model?



Final Predicted Value

$$\dot{E} = \hat{E} \pm \sqrt{VAR}$$

Baseline Models for Whole Building Energy Consumption

Simple-Linear Regression

- **Use the outdoor dry-bulb temperature (T) to baseline whole building energy consumption (E) based on monthly utility bills.**

$$\hat{E} = \beta_0 + \beta_1 T_0$$

Strengthen: **simple; world-wide adopted and used**

Accuracy: Low absolute error within 5% with large error bars at 90% confidence level (explain)

Application: Simple comparison and estimation, preliminary analysis

Baseline Models for Whole Building Energy Consumption

Multiple Linear Regression

•Correlationships between the climate data, which are monthly mean outdoor dry-bulb temperature (T), relative humidity (RH) and global solar radiation (GSR), and whole building energy consumption are derived.

$$\hat{E} = \beta_0 + \beta_1 T_0 + \beta_2 RH + \beta_3 GSR$$

Strengthen: Holistic; removing weather effects

Accuracy: An improved prediction accuracy around 2% is achieved.

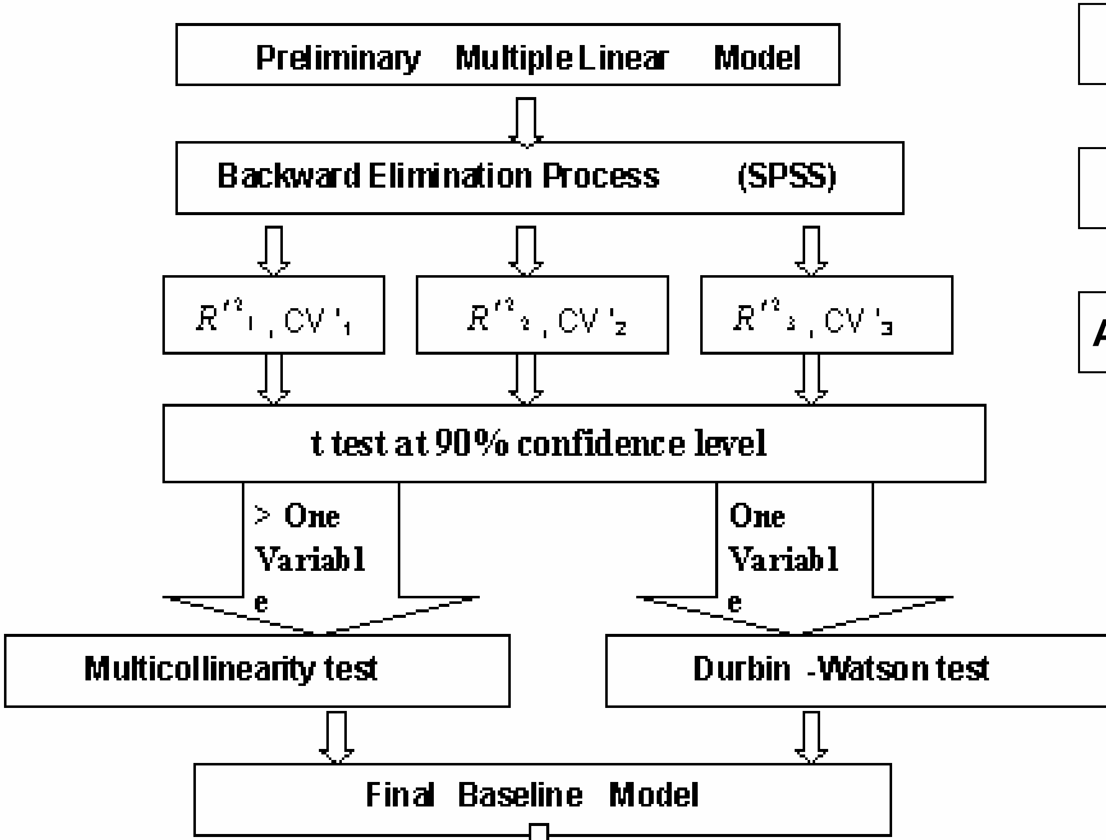
Application: Whole building diagnostic and more accurate prediction, detailed analysis including sub-systems.

Baseline Models for Whole Building Energy Consumption

Whole building holistic baseline method

-- A combination of simple regression and multiple regression method

Model Establishing



Model Verification

input: Weather data of projected year

⇩

Output: Projected year's energy use

⇩

Accuracy assessment: CV,PI, VAR, MAE

Baseline Models for Landlord Energy Consumption

Landlord Energy Consumption

A buildings' landlord energy consumption refers to the energy utilized by The common facilities, systems, services and space provided by the landlord

a) Air-conditioner central plant system which supply air-conditioning inside the building;

b) Vertical transportation service i.e. escalator and lift;

c) Ventilation system such as exhaust fan and ventilator;

d) Artificial lighting system in the common area i.e. corridor or public common service area i.e. toilet and lift;

Baseline Models for Landlord Energy Consumption

Primary Analysis

1. Utility bills are mostly in the form of landlord energy consumption

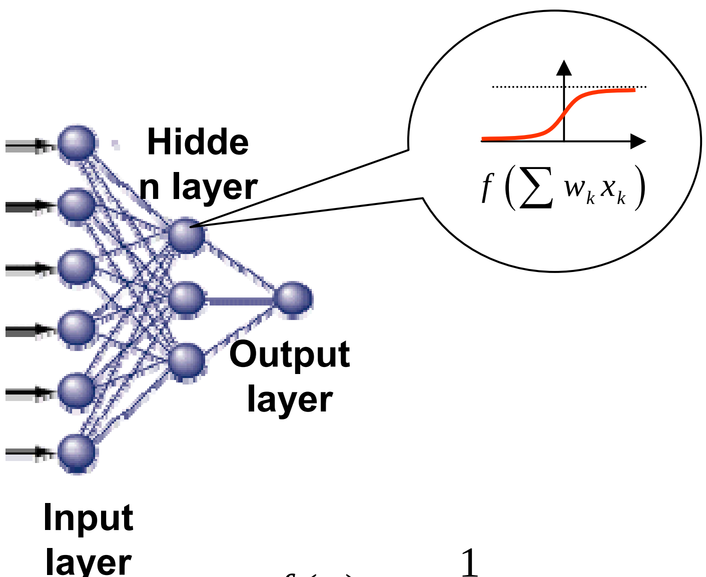
2. Linear regression models are not adopted. ($R2<0.5$)

Non-linear models were investigated including Neural Networks and Support Vector Machines based on four years utility bill data

Baseline Models for Landlord Energy Consumption

Neural Networks

1. **Structure of the proposed Neural Networks**



$$f(x) = \frac{1}{1 - e^{-x}}$$

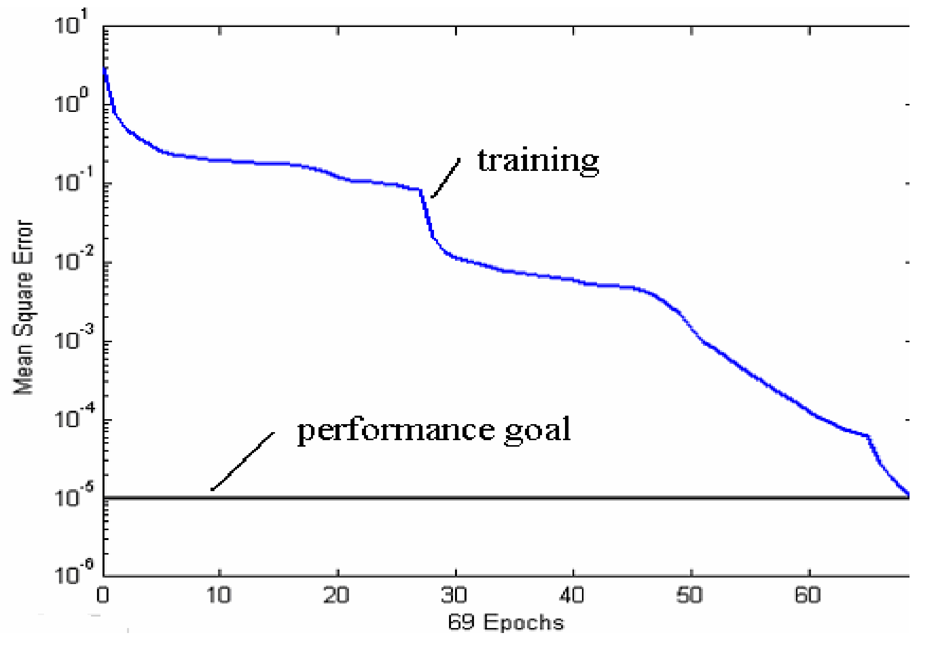
$$f'(x) = f(x)(1 - f(x))$$

- Three layers, fully-connected feed-foreword, Backpropagation (BP) NN
- Neurons in layers: 4, n, 1;n is defined by the best performance
- Transfer functions: tan-sigmoid, tan-sigmoid and linear

Baseline Models for Landlord Energy Consumption

1. **Structure of the proposed Neural Networks**

- **The stop criteria is 0.00001 (MSE_1-MSE_2<10^{-5}), Epochs are set to 500**
- **Up to 3 year's bills for training The best five results are selected from 30 runs.**



2. **Main findings**

- **NN approach gives fairly good annual predictions with MAE<2%, on the other hand rather high CV>9%**

Baseline Models for Landlord Energy Consumption

Support Vector Machines

- A new neural network algorithm developed by Vapnik
- Widely applied in the financial time series forecasting

Principle of SVM for regression:

•Given a set of data points (x1,y1),(x2,y2)…(xn,yn) are randomly and independently generated from an unknown function. SVM approximates the unknown function using the following form

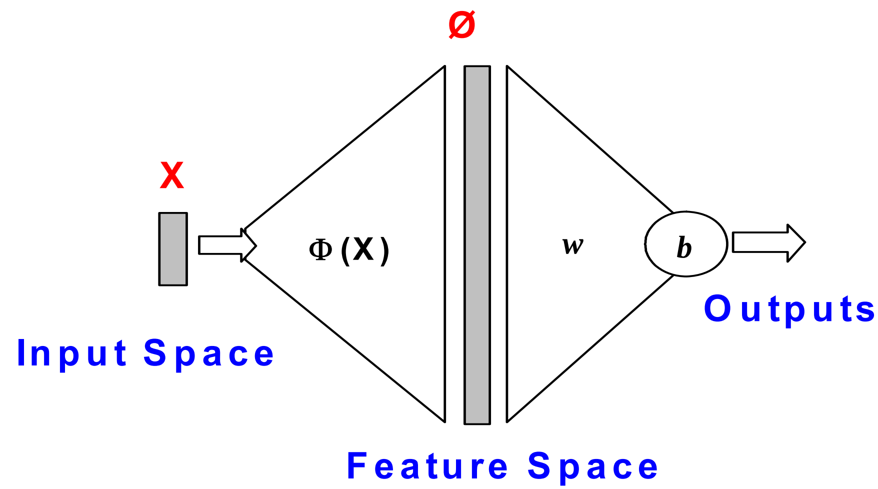
$$f(x) = \omega \cdot \phi(x) + b$$

Baseline Models for Landlord Energy Consumption

Support Vector Machines

Principle of SVM for regression:

$$f(x) = \omega \cdot \phi(x) + b$$



- Ø(x) represents the high-dimensional feature spaces which is nonlinearly mapped from the input space X using kernel function K(x).

- SVM will construct a linear function in the high dimensional space Ø to solve the non-linear problem in the low-dimensional space X

Baseline Models for Landlord Energy Consumption

Support Vector Machines

Kernel function:

$$K(\mathbf{x}_i, \mathbf{x}) = (\Phi(\mathbf{x}_i) \cdot \Phi(\mathbf{x})) = \Phi^T(\mathbf{x})\Phi(\mathbf{x}_i)$$

Types of Kernel function:

Linear: $K(\mathbf{x}_i,\mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j$

Polynomial of power *p*: $K(\mathbf{x}_i,\mathbf{x}_j) = (1+ \mathbf{x}_i^T \mathbf{x}_j)^p$

Gaussian (radial-basis function (selected network):

$K(xi,xj) = e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}}$

Two-layer perception: $K(\mathbf{x}_i,\mathbf{x}_j) = \tanh(\beta_0 \mathbf{x}_i^T \mathbf{x}_j + \beta_1)$

Baseline Models for Landlord Energy Consumption

Support Vector Machines

The coefficients w and b are estimated by minimizing the regularized risk function

$$\tfrac{1}{2}\left\|\omega\right\|^2 + C\,\tfrac{1}{l}\sum_{i=1}^{l} L_\varepsilon(y_i, f(x_i))$$

Where ε-insensitive loss function

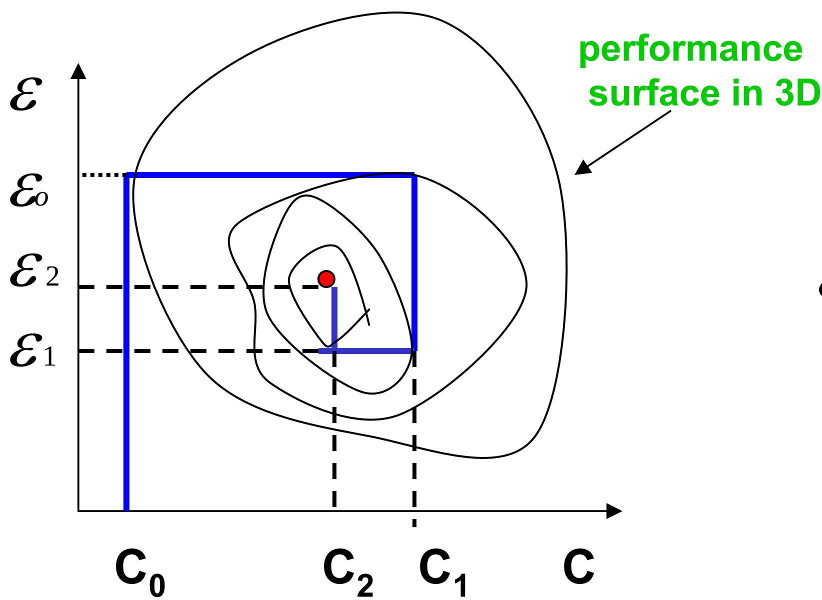
$$L_\varepsilon(y_i, f(x_i)) = \begin{cases} \left|y_i - f(x_i)\right| - \varepsilon, & \left|y_i - f(x_i)\right| \geq \varepsilon \\ 0 \end{cases}$$

Because $\left\|\omega\right\|^2$ is defined by kernel function and constant, the best pair of (C and ε) should be found

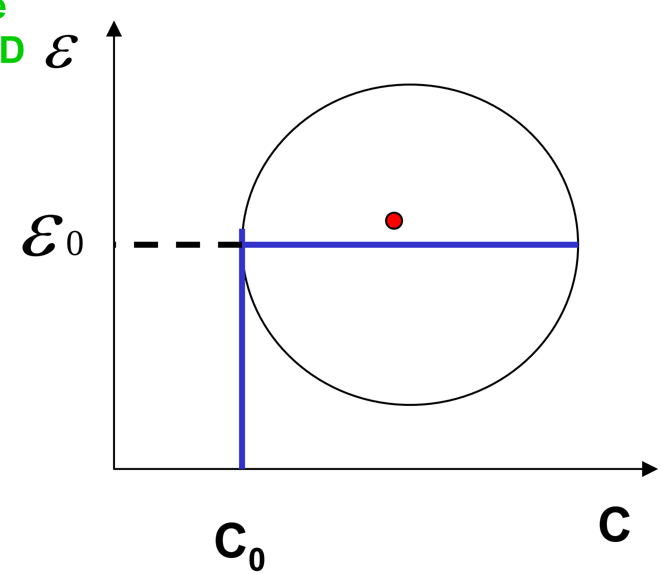
Baseline Models for Landlord Energy Consumption

Support Vector Machines

Step Wise Search for the (C and ε)



One time search (Normally Used)



Main Findings

Good prediction with MAE<4% and CV<3%

Baseline Models for Landlord Energy Consumption

Support Vector Machines

Characters of SVMs

1. SVMs estimates the regression using a set of linear functions that are defined in a high-dimensional feature space, while the inputs have nonlinear performance.

2. SVMs carries out the regression estimation by risk minimization, based on statistical learning theory, where the risk is measured using Vapnik's Epsilon-insensitive loss function

3. Training SVMs is equivalent to solving a linearly constrained quadratic programing problem so that the solution of SVMs is always unique and globally optimal, while the results of NN are not unique.

Results and Discussion

Results of Baselining Whole Building Energy Consumption

1. Results of simple linear regression (SR)

Building	Annual Energy Use (kWh/m^2/month)		MAE (%)	CV	VAR(Model) (kWh/m^2/month)	90%PI (%)
	Measured	Modeled				
A	20.26	20.59	1.64	3.80	1.35	10.06
B	32.87	32.09	2.42	2.78	2.60	2.9
C	28.59	24.22	15.33	3.49	29.79	48.15
D	25.55	25.76	0.8	3.76	2.05	2.59
E	18.81	18.34	2.5	3.32	2.83	3.04

For example: the final predicted result of building A are expressed as 20.59±√1.35, within 20.59±10.06 prediction inerval

Results and Discussion

Results of Baselining Whole Building Energy Consumption

2. Result of Multiple linear regression (MLR)

	Baseline Year	Prediction Year		MAE(%)		CV(%)		VAR		PI	
		SR	MLR	SR	MLR	SR	MLR	SR	MLR	SR	MLR
G	23.52	24.11	24.14	2.51	2.63	7.11	5.18	3.96	2.05	4.15	3.30
H	30.31	28.63	30.35	5.53	0.14	8.26	3.97	8.9	2.05	5.72	3.72
I	22.86	23.66	23.66	7.44	3.49	12.63	11.37	13.44	9.71	6.73	5.90

Improvements 97% 52% 77% 35%

Results and Discussion

Results of Baselining Landlord Energy Consumption

1. **Results of Neural Network (NN)**

Building Ref. No.	Actual Value ($kWh/month/m^2$)	Predicted value	CV (%)	MAE (%)
K	10.55	10.39	9.67	-1.5
L	11.05	11.12	15.5	0.64
M	9.54	10.23	15.12	7.33
N	12.59	11.82	14.17	-6.2

Results and Discussion

Results of Baselining Landlord Energy Consumption

2. Results of Support Vector Machines (SVMs)(1)

Building Ref. No.	Actual Value (kWh/month/ m^2)	Predicted value	CV (%)	MAE (%)
K	10.55	10.26	2.69	-2.72
L	11.05	11.43	2.39	3.44
M	9.54	9.61	1.28	0.68
N	12.59	12.35	0.99	-1.89

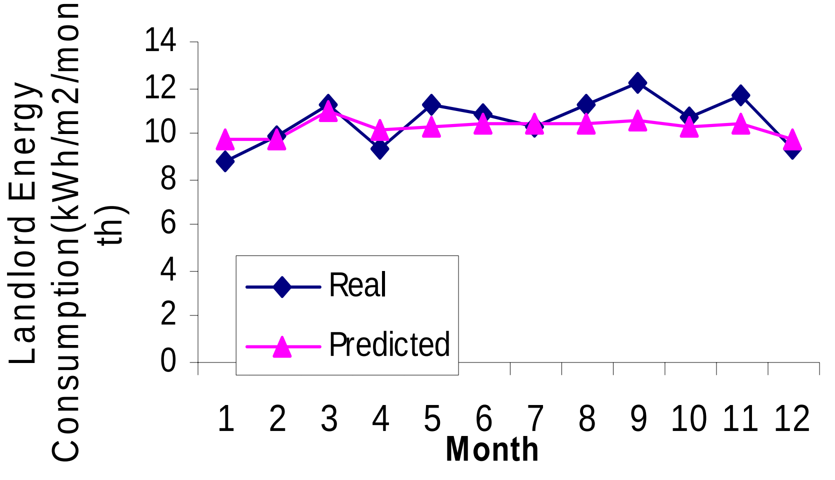
<3%

<4%

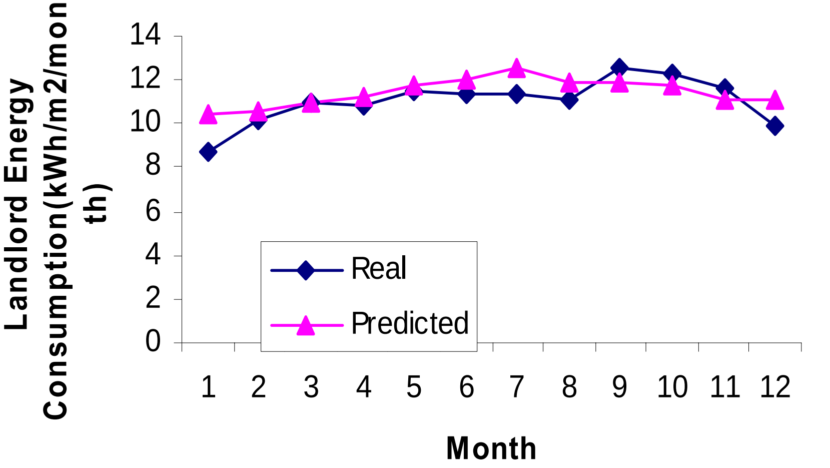
Results and Discussion

Results of Baselining Landlord Energy Consumption

2. Results of Support Vector Machines (SVMs)(2)



Building K

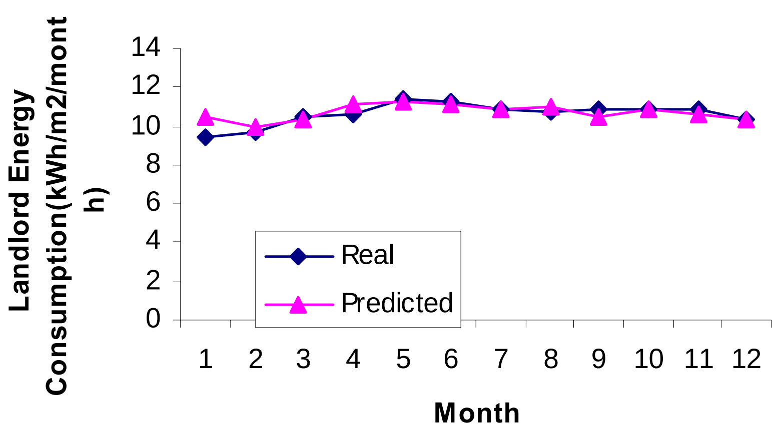


Building L

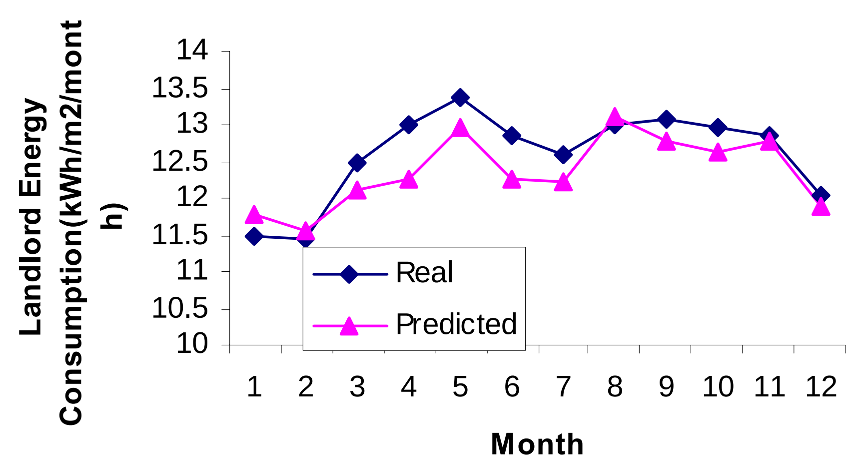
Results and Discussion

Results of Baselining Landlord Energy Consumption

2. Results of Support Vector Machines (SVMs)(2)



Building M



Building N

Conclusion

1. **A holistic baseline methodology for office buildings in Singapore has been established**

2. **The weather data as the only normalization parameters can generalize a fairly well prediction results.**

3. **It is important to keep recording the future energy consumption and weather data after the installation of ECM. (Continuous Commission)**

4. **For enhanced accuracy, future research will focus on hourly data analysis and model development**

Thank You

Energy Sustainability Unit

NUS
National University of Singapore