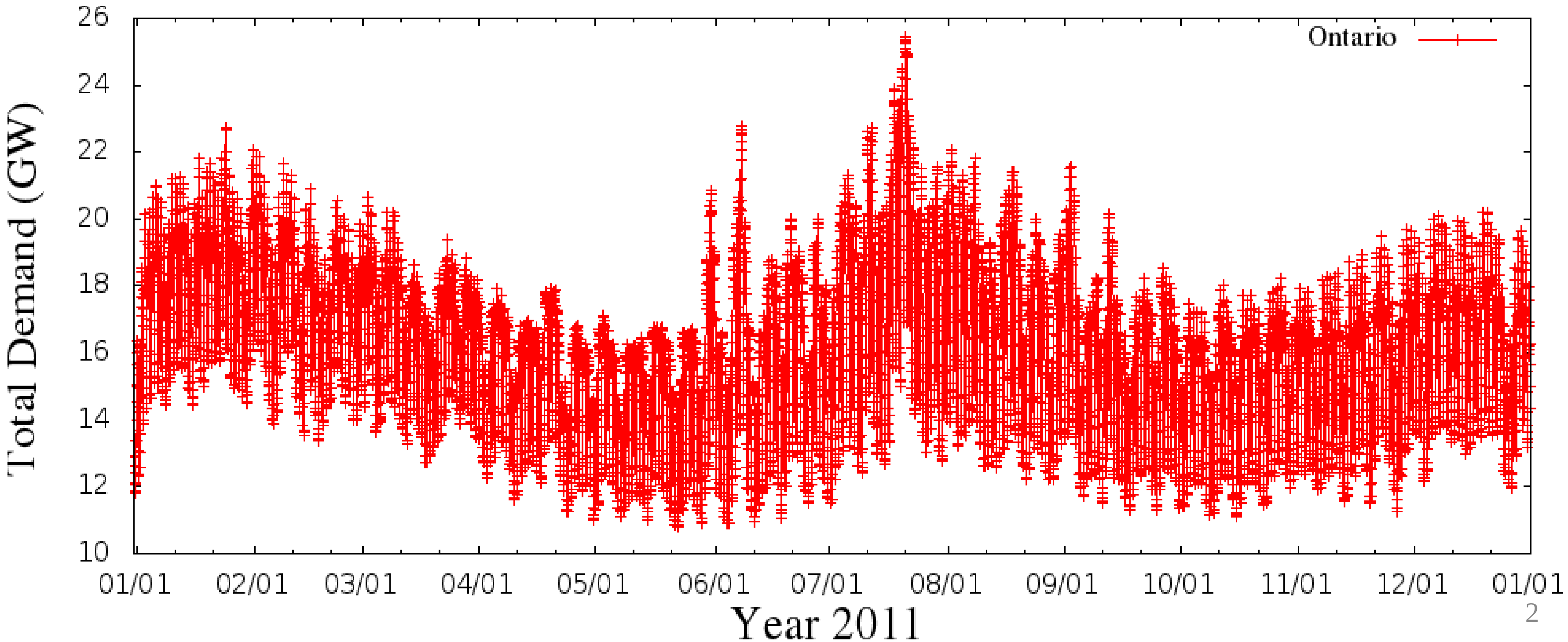
On Hourly Home Peak Load Prediction

Rayman Preet, Peter Gao, Dan Lizotte

Peak Load

- Highest *Aggregate Demand*
  - *Area*
  - *Timescale*
- Costs

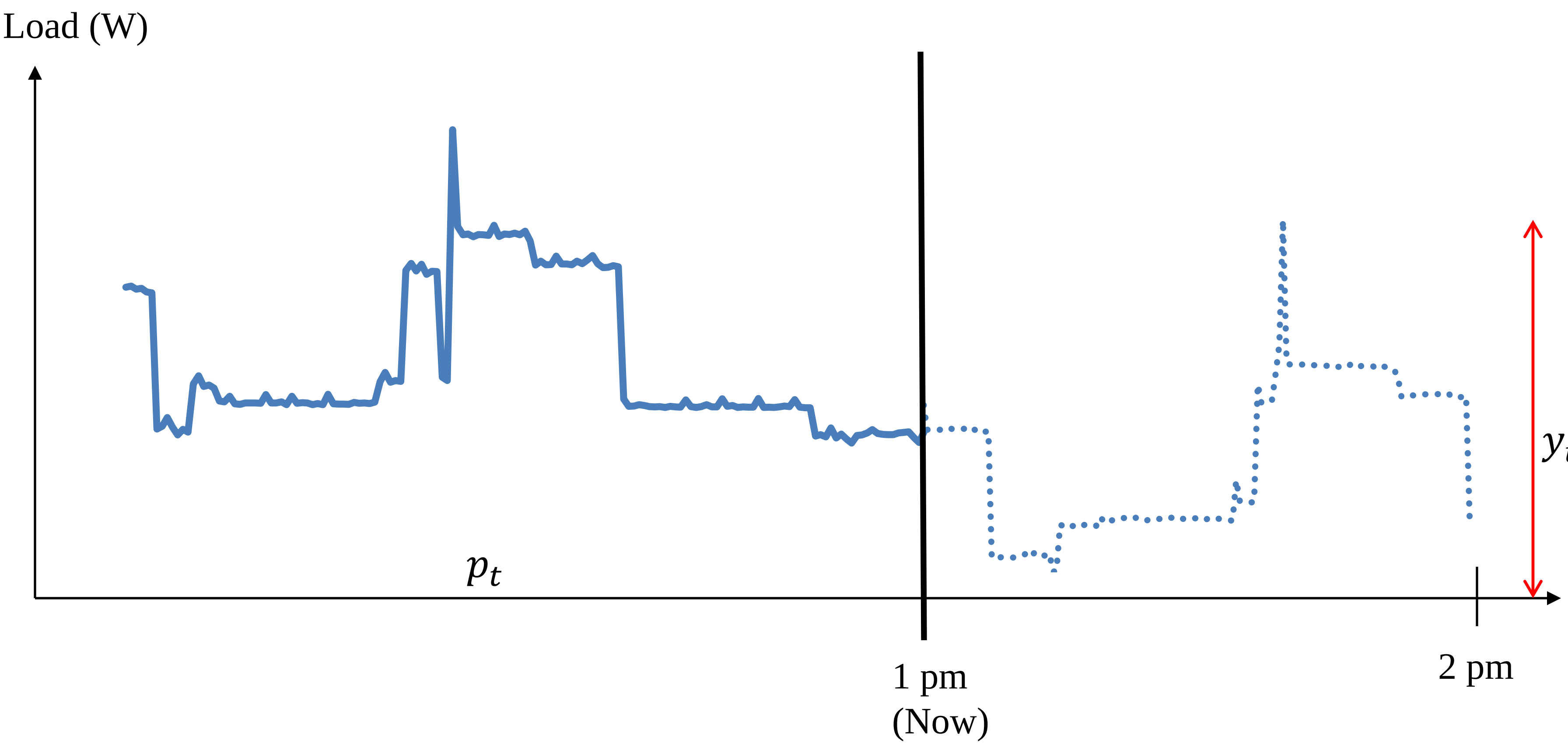


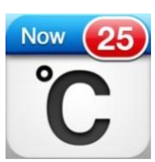
Solutions

- Storage
  - *When?*
  - *How much?*

- Demand Response
  - *Scheduling*
- *Microgrids*

Goal





Given, ***past p_t*** values, and ***all physical observables***, predict ***y_t***

Existing Work

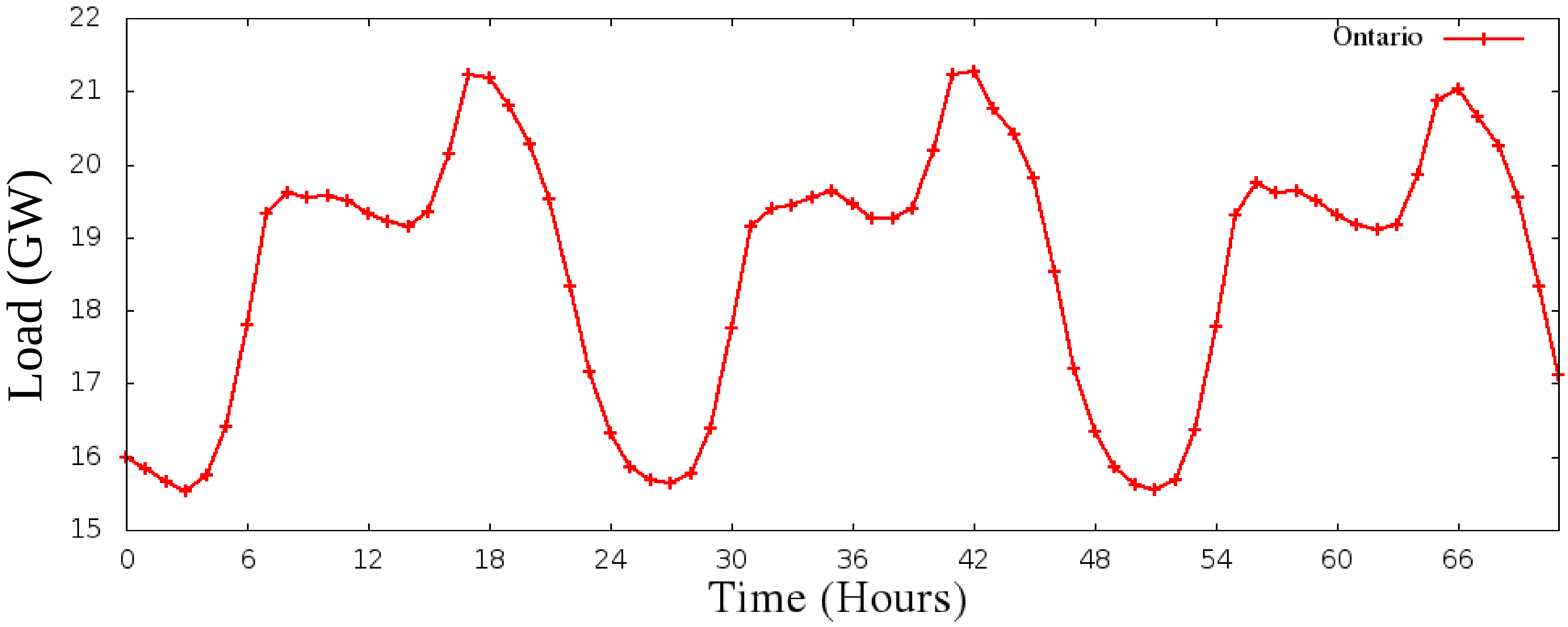
- *Area*
  - Building, Block, Region, Province v/s *Home*

- *Type*
  - Mean v/s *Peak*

- *Timescale*
  - Year, Month, Week, Day v/s *Hour*

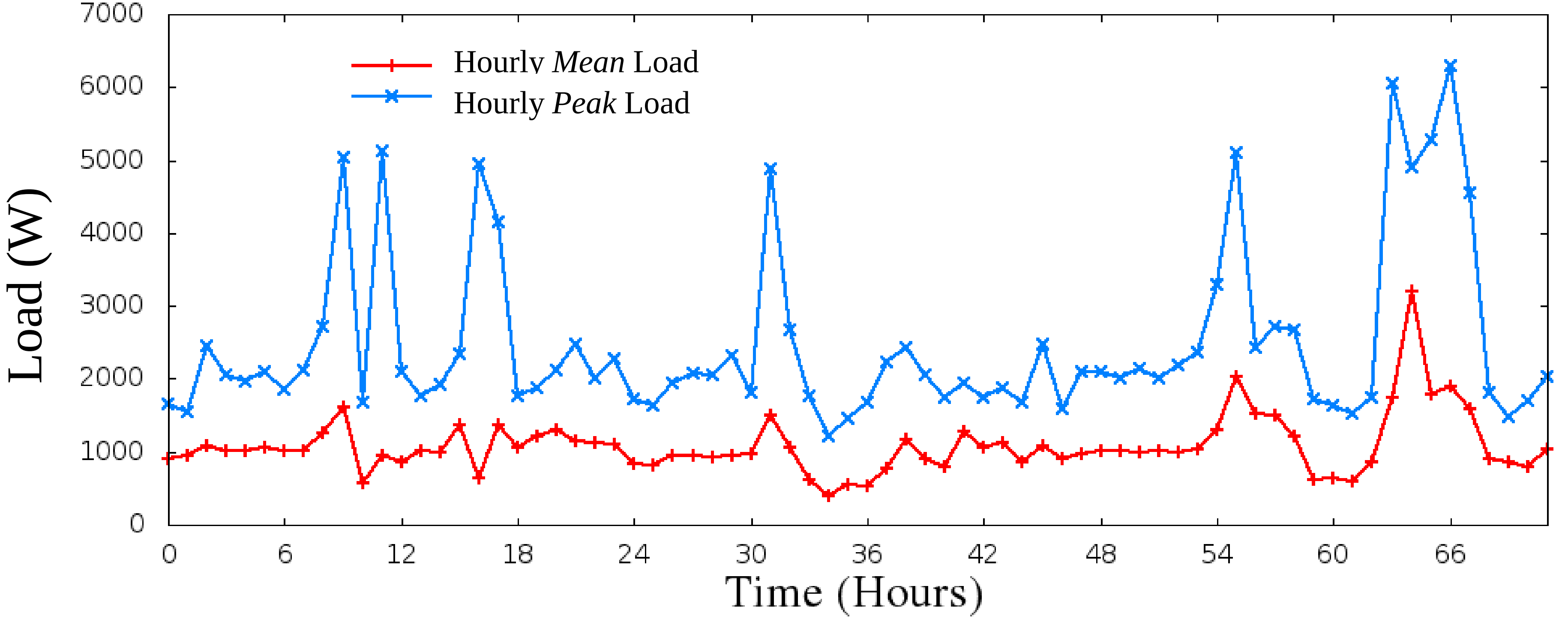
Hourly Load

- Ontario



Hourly Load

- Home



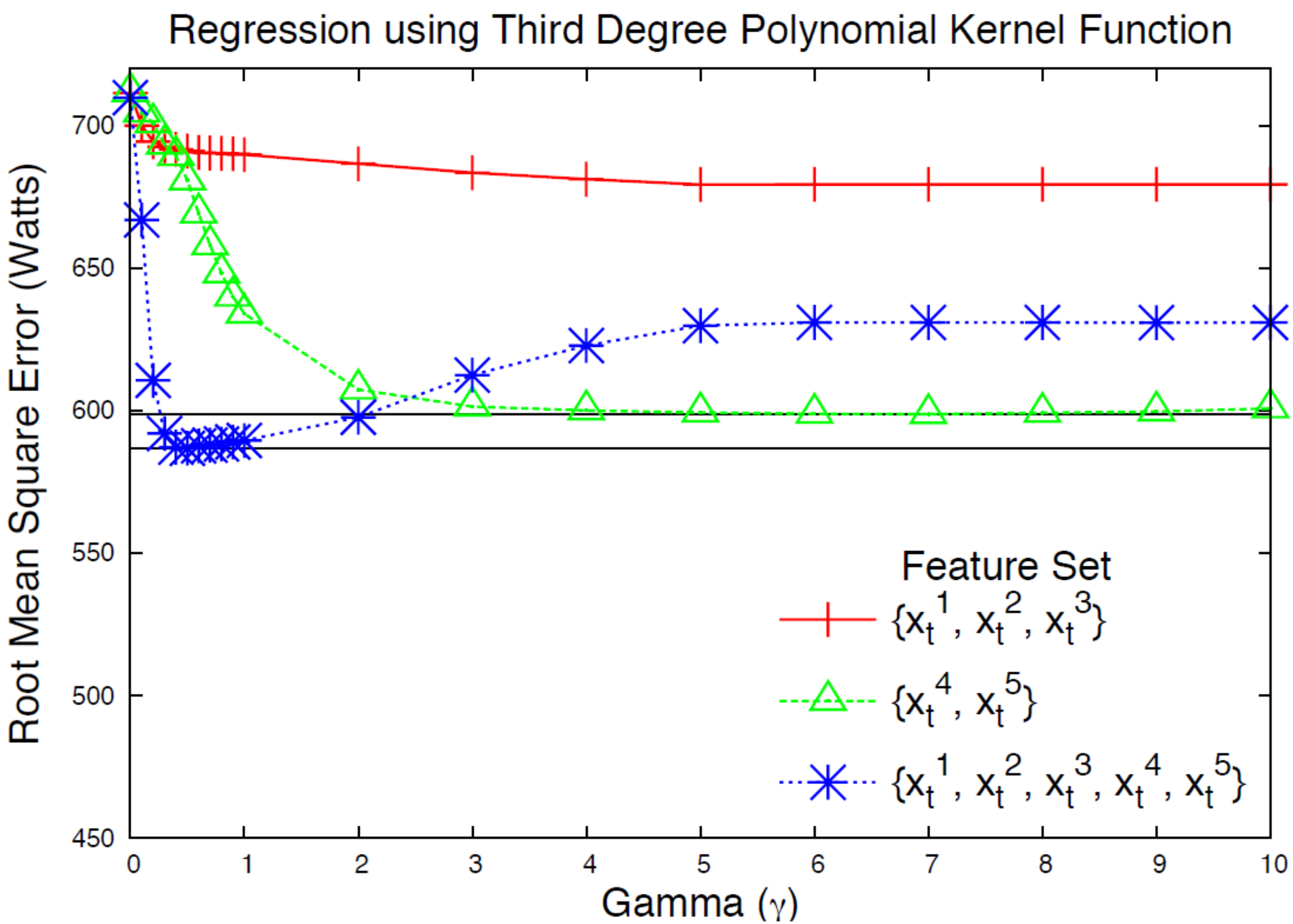
Straw man Proposal

- *Time of Day*
  - $x_t{}^1 \in \{1, 2 \ldots 24\}$
- *Day of Week*
  - $x_t{}^2 \in \{1, 2 \ldots 7\}$
- *Ambient Temperature*
  - $x_t{}^3 \sim T_{t-1}$
- *Consumer Activity*
  - $x_t{}^4 \sim Variance(p_j)$, $t\text{-}1 < j < t$
- *Previous Peak Load*
  - $x_t{}^5 = y_{t-1}$

Learning Methods

- Find *f* :
  - At *t* , $y_t \approx f\ (x_t{}^1, x_t{}^2, x_t{}^3, x_t{}^4, x_t{}^5)$
- Regression
  - *Support Vector*
  - *Least Square Support Vector*
- Artificial Neural Networks

Contribution of Features



Overfitting →

Auto-Regressive Moving Average

Home Load = *Latent Pattern* + *Stochastic Pattern*

ARMA = *Auto-Regressive* + *Moving Average*

Auto-Regressive Moving Average

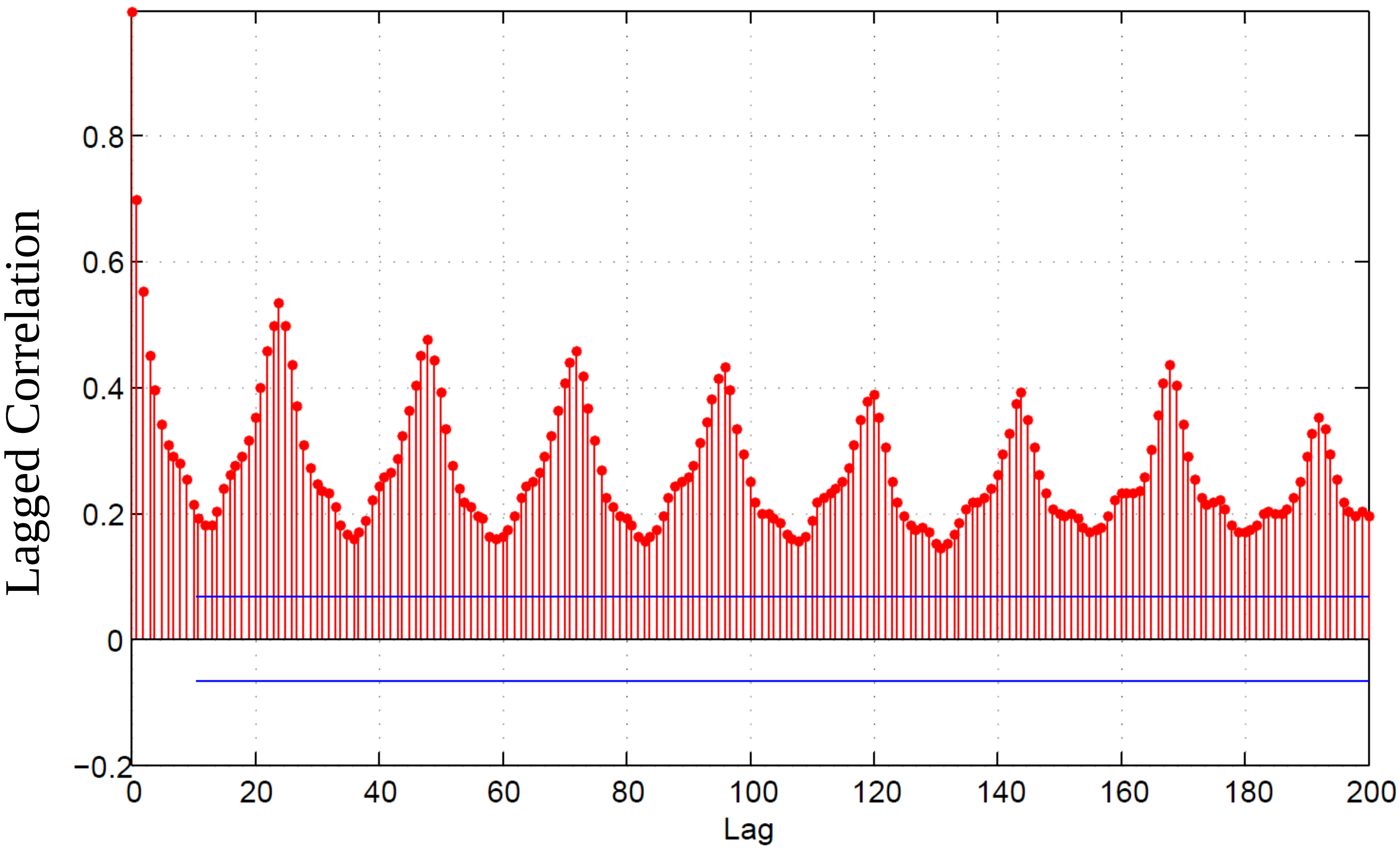
$$y_t = \underbrace{\delta + \sum_{i=1}^{p} \phi_i y_{t-i}}_{AR\ (routine\ activities)} + \underbrace{\epsilon_t - \sum_{i=1}^{q} \theta_i \epsilon_{t-i}}_{MA\ (stochastic\ activities)}$$

ϵ_t → *white noise*

p, q → *model degrees*

δ, ϕ_i, θ_i → *model parameters*

Seasonal ARMA



Peak Load *may* show *seasonality !*

Seasonal ARMA

- Decompose y_t

$$y_t = S_t + N_t$$

$$Let\ B^i y_t = y_{t-i}\ ,$$

$$S_t = S_{t+s} \quad \Rightarrow \quad (1 - B^s)S_t = 0$$

$$(1 - B^s)y_t = (1 - B^s)S_t + (1 - B^s)N_t = (1 - B^s)N_t$$

Not Seasonal

Dataset

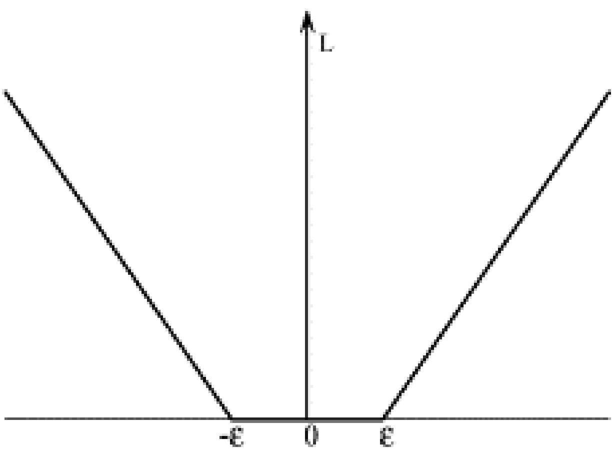
- *Twenty Four* homes in Kitchener-Waterloo, Ontario
- *Stratified* samples
- *Sampled* every six seconds

Evaluation: Regression

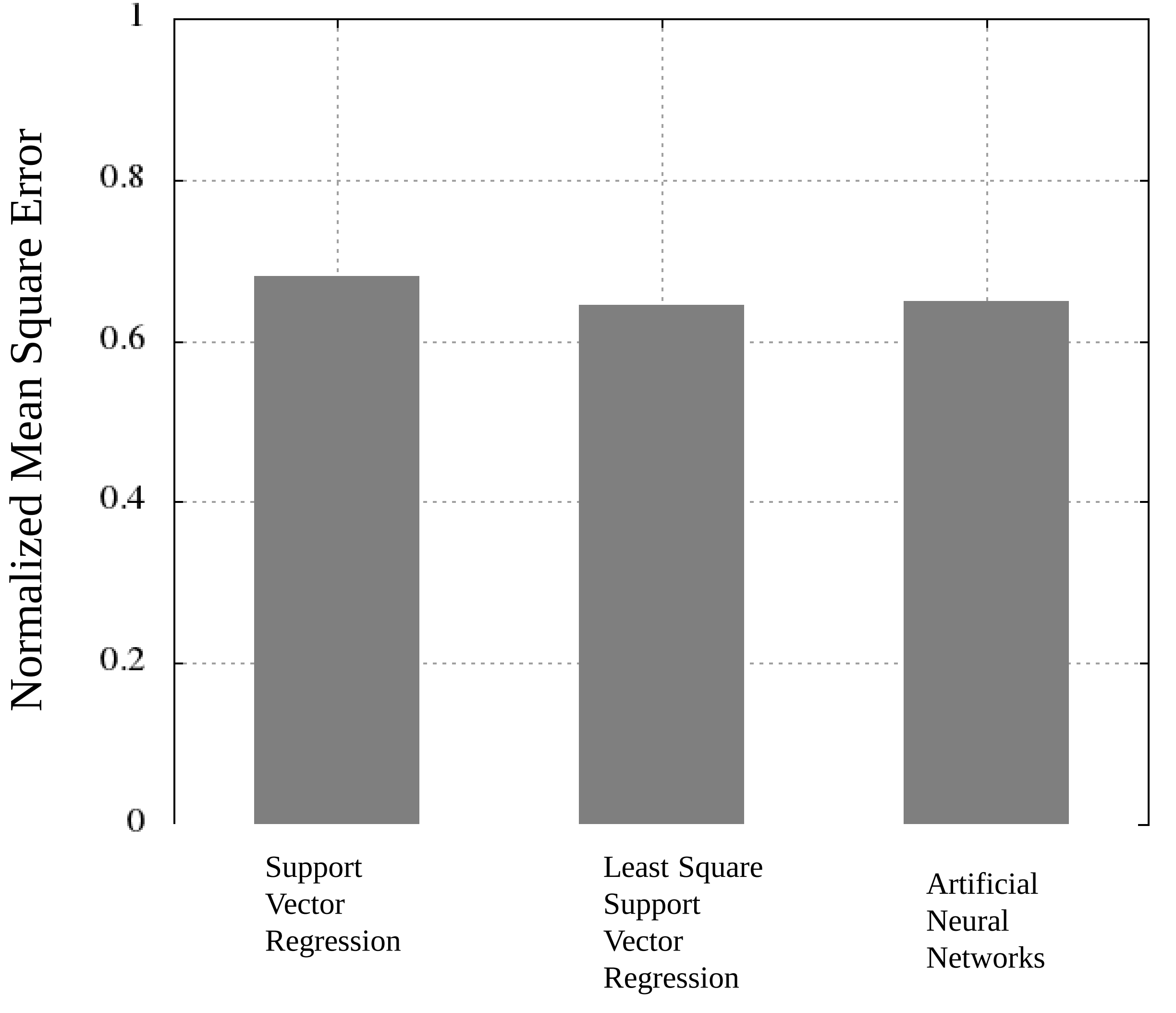
- *Randomized* Subsets and Cross-validation
- *Normalized Root Mean Square Error*

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}{N}} \qquad \mathrm{NMSE} = \frac{\frac{1}{N}\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}{\mathrm{Var}(y_t)} = \frac{\mathrm{MSE}}{\mathrm{Var}(y_t)}$$

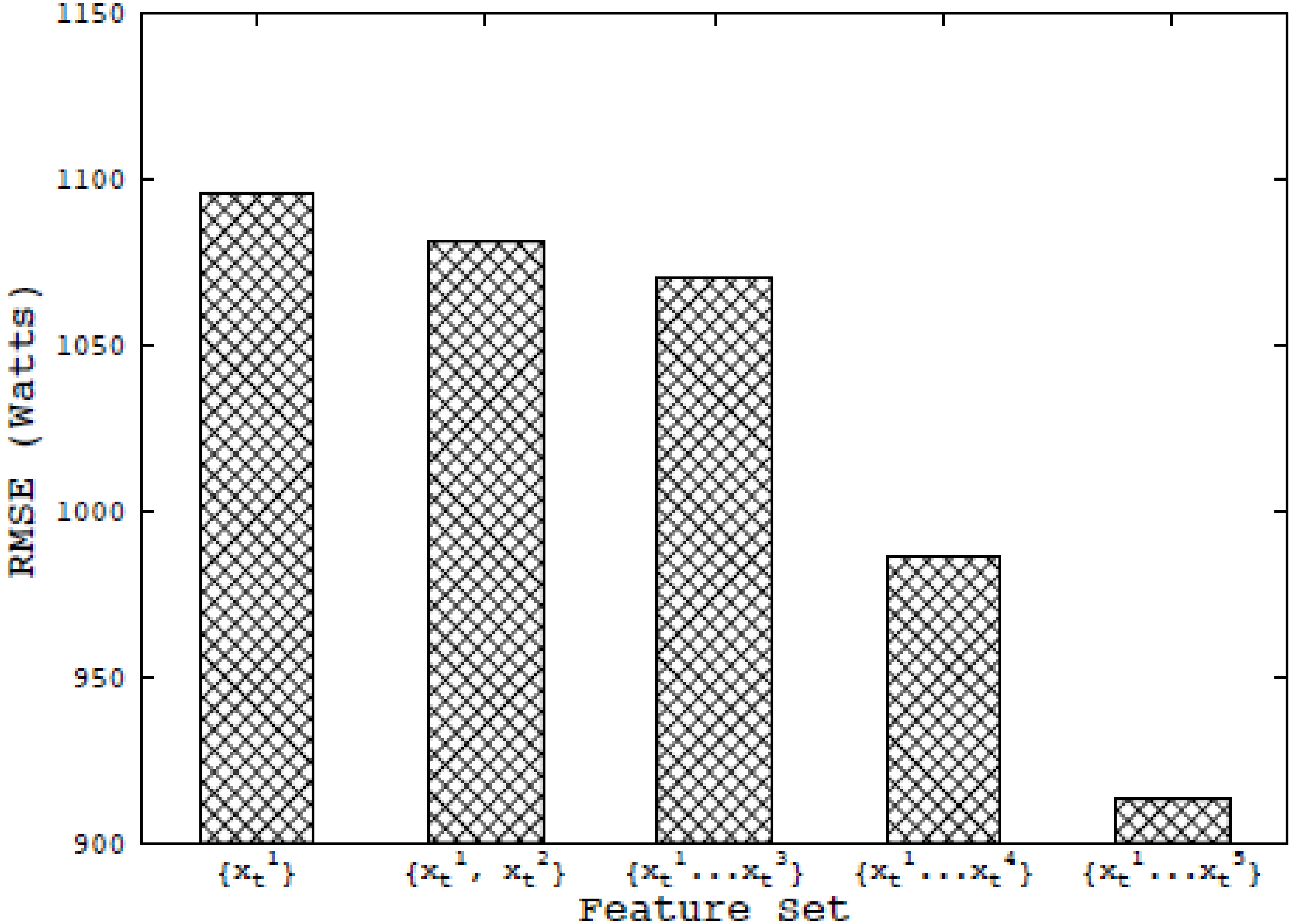
- *Kernels*
  - Linear, Polynomial, Radial Basis, and Sigmoid
- *Vapnik Loss*



Evaluation: Regression

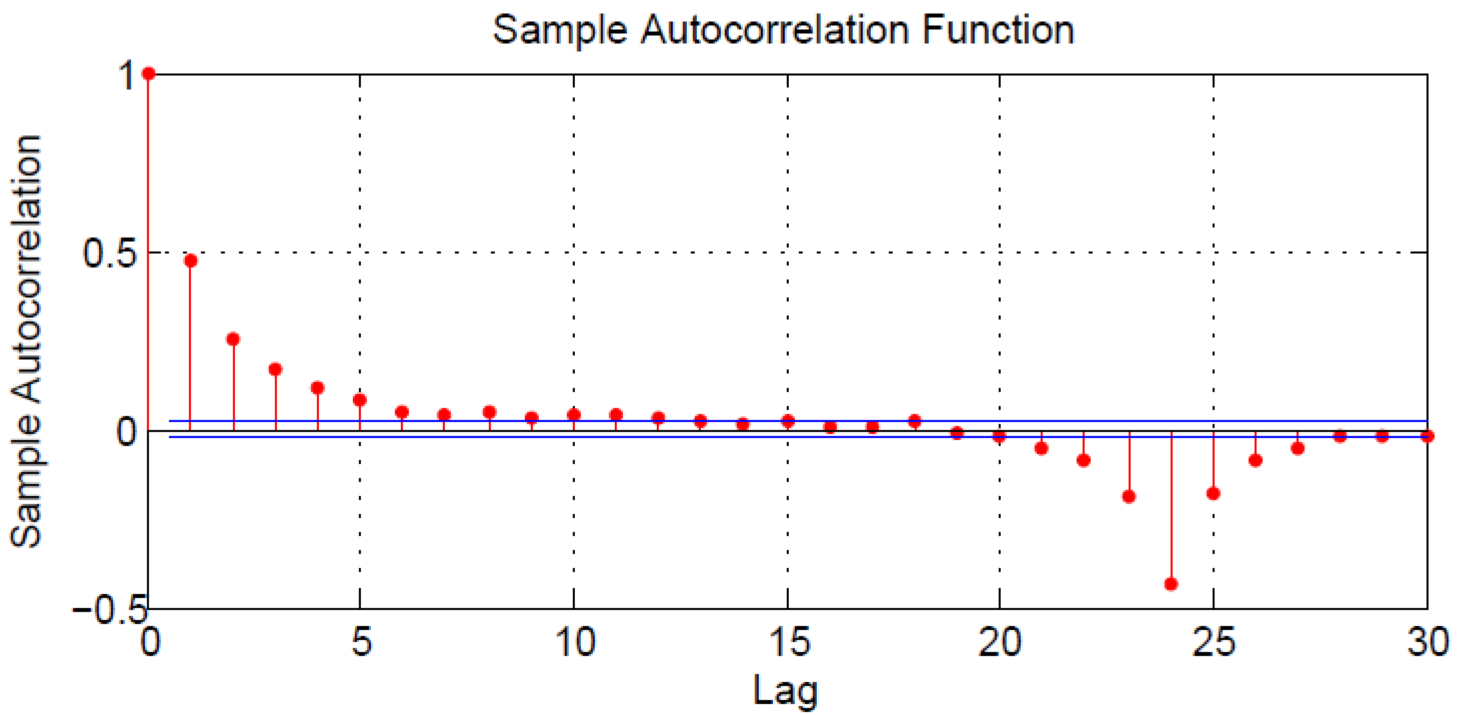


Evaluation: Regression



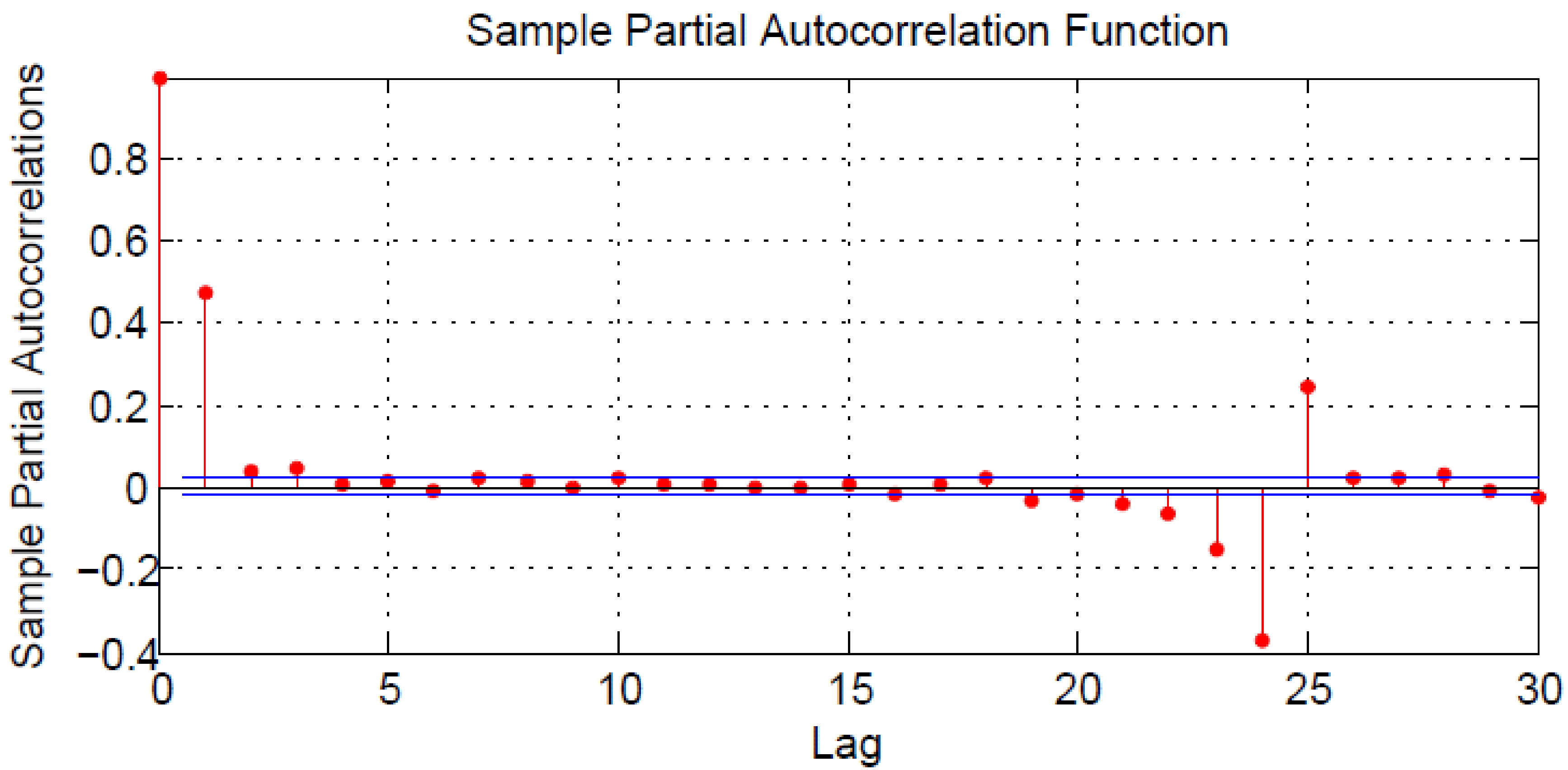
Evaluation: ARMA

- Model degrees *p*,*q*
- Error *minimization* to get δ, ϕ_i, θ_i



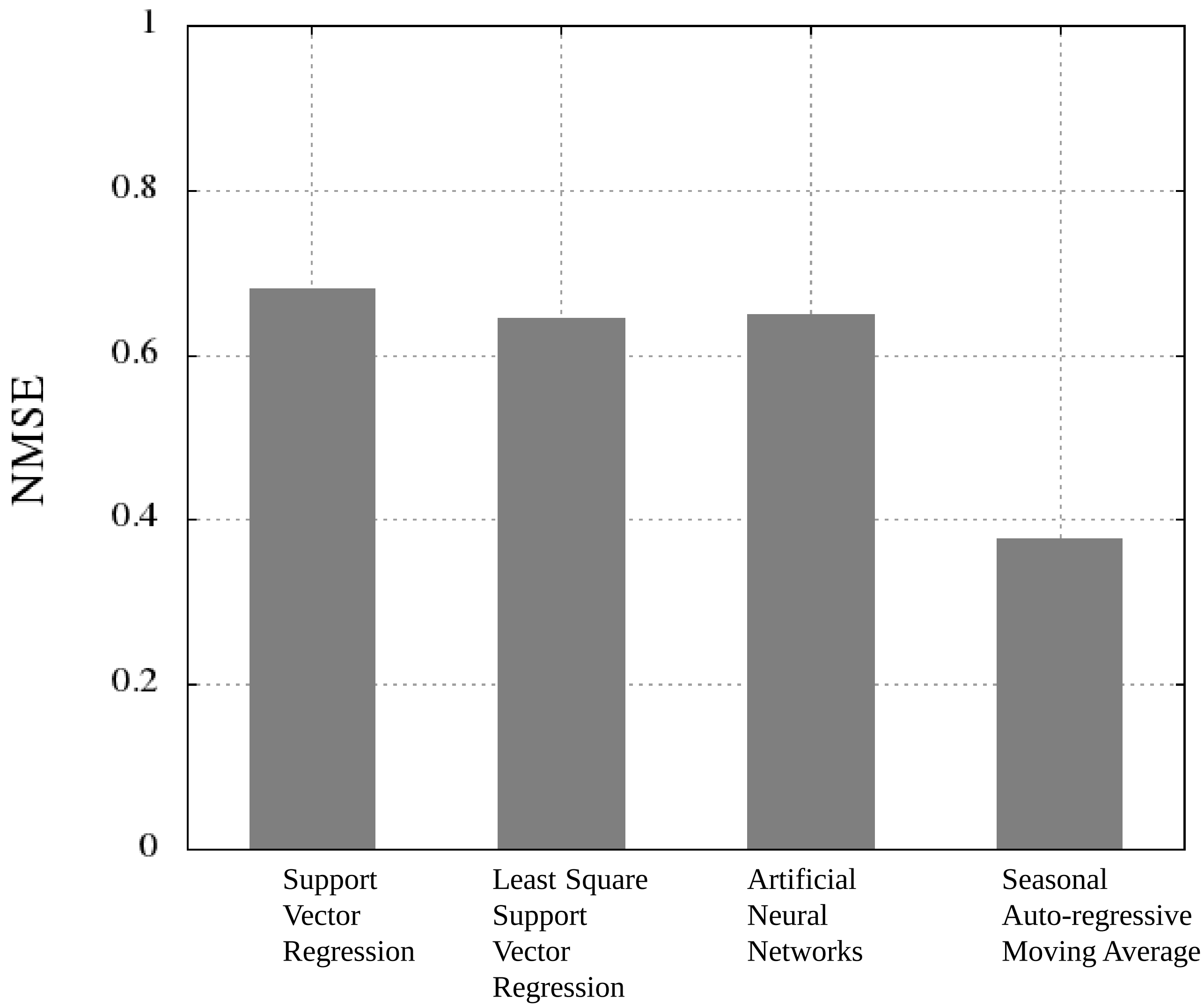
Auto-Correlation with Seasonality Removed

Evaluation: ARMA



Partial Auto-Correlation with Seasonality Removed

Comparison


NMSE
1
0.8
0.6
0.4
0.2
0
Support Vector Regression
Least Square Support Vector Regression
Artificial Neural Networks
Seasonal Auto-regressive Moving Average

Conclusion

- Home load is *highly stochastic!*
- *Moving-average* based solutions are better than *learning-based*
- For hourly prediction, both the *short-term* and *long-term* trends are important
- SARMA works ~30% better than *others*