



# Machine Learning Suite Overview and Tutorial

Richard E. Edwards

Distributed Intelligence Lab  
Department of Electrical Engineering and Computer Science  
University of Tennessee, Knoxville TN, USA

May 4, 2013

Funded by Whole Building & Community Integration Group,  
Oak Ridge National Laboratory, Oak Ridge TN, USA



# Outline

Introduction

Machine Learning Suite

XML Interface

MLSuite Results

Future Electrical Consumption

EnergyPlus Approximation

Inverse EnergyPlus

Closing Remarks



# Outline

## Introduction

## Machine Learning Suite

## XML Interface

## MLSuite Results

Future Electrical Consumption

EnergyPlus Approximation

Inverse EnergyPlus

## Closing Remarks



# Machine Learning

- ▶ Objective:
  - ▶ Learn some function:  $F$
- ▶  $F$ 's true **characteristics** are unknown
- ▶  $F$ 's known **characteristics**
  - ▶ Maps input set  $X$  to output set  $Y$



# Examples

- ▶ Cereal Brand Classification
  - ▶ Input: Consumer Information
    - ▶ Age
    - ▶ Region
    - ▶ Ethnicity
    - ▶ etc
  - ▶ Output:
    - ▶ Cereal Brand or Brands
  
- ▶ Predicting Residential Electrical Consumption
  - ▶ Input: Environmental Measurements
    - ▶ Dry Bulb Temperature
    - ▶ Humidity
    - ▶ Wind Speed
    - ▶ etc
  - ▶ Output:
    - ▶ Electrical Consumption or Expected Load



# Machine Learning Techniques

- ▶ Two-Types:
  - ▶ Classification
  - ▶ Regression
- ▶ Example Methods
  - ▶ Linear Regression
  - ▶ Logistic Regression
  - ▶ Decision Trees
  - ▶ Neural Networks
  - ▶ Support Vector Machines



# Existing Software

- ▶ Scattered



## Existing Software

- ▶ Scattered
- ▶ Example: Support Vector Machine Software
  - ▶ OCAS, LibLinear, SVMlin, libSVM, etc





## Existing Software

- ▶ Scattered
- ▶ Example: Support Vector Machine Software
  - ▶ OCAS, LibLinear, SVMlin, libSVM, etc
  - ▶ Different capabilities (i.e. regression vs classification support)
  - ▶ Different limitations (i.e. large scale data support)
- ▶ Shogun <http://www.shogun-toolbox.org/page/home/>
  - ▶ Bundles 6 different open-source SVM libraries



## Machine Learning Suite (MLSuite)

- ▶ Reduces software scattering even more!
- ▶ MLSuite:  
<http://web.eecs.utk.edu/~redwar15/MLSuite/MLSuite>
- ▶ Supported Learners:
  - ▶ Linear Regression
    - ▶ Support Subset Selection
  - ▶ Lasso Regression (Large and Small problems)
  - ▶ Feed Forward Neural Networks (FFNN) (Large and Small problems)
  - ▶ Support Vector Machines
  - ▶ Least Squares Support Vector Machines (LS-SVM)
  - ▶ Sparse Gaussian Graphical Model (GGM)
  - ▶ Ensemble Methods
    - ▶ FFNN
    - ▶ LS-SVM
    - ▶ Linear Regression



## MLSuite's Applications

- ▶ Predicting future hourly electrical consumption
  - ▶ Inputs: Campbell Creek's sensor data
  - ▶ Output: Next hour's electrical consumption
  - ▶ R. E. Edwards, J. New, and L. E. Parker, Predicting Future Hourly Residential Electrical Consumption: A Machine Learning Case Study, *Energy and Buildings*, vol. 49, pages 591-603, June 2012
- ▶ Approximating EnergyPlus
  - ▶ Inputs: 150 building simulation parameters
  - ▶ Outputs: 80 to 90 simulation variables
  - ▶ Paper submitted for review to *Energy and Buildings*
- ▶ Inverse EnergyPlus
  - ▶ Inputs: Simulation output, weather data, and operation schedule
  - ▶ Outputs: Building simulation parameter estimates
  - ▶ Preparing paper for Big Data Mining Workshop at KDD'13



# Outline

Introduction

Machine Learning Suite

XML Interface

MLSuite Results

Future Electrical Consumption

EnergyPlus Approximation

Inverse EnergyPlus

Closing Remarks



# Software Components

- ▶ Core
  - ▶ Developed in Matlab
  - ▶ Integrates multiple learning
  - ▶ Supports multiple data sources
- ▶ XML interface
  - ▶ Developed in Python
  - ▶ Stream lines learner execution



# Supported Platforms

- ▶ Core Software
  - ▶ Windows, OSX, Linux, and Nautilus
  - ▶ Requires Matlab
- ▶ Core Software & Python XML Interface
  - ▶ Nautilus and networked Linux computers



# MLSuite's Matlab Core

- ▶ Data
  - ▶ Supports csv, excel 97, and mat files
  - ▶ Supports SQL and MySQL database access
  - ▶ Supports integrating multiple data sources
  - ▶ Supports data standardization
- ▶ Learners
  - ▶ All accept a standard object as input
  - ▶ All access data via the same interface
  - ▶ All have access to data
- ▶ Results
  - ▶ Supports CV, MAPE, and MBE metrics
  - ▶ Easily extensible to support additional metrics



# Data

- ▶ Data abstraction
  - ▶ All data is either a matrix or vector.
  - ▶ This includes files and database connections
- ▶ Data object types:
  - ▶ File Containers
    - ▶ EPlusFileContainer.m
  - ▶ Database Containers
    - ▶ EPlusContainer.m (eplusruns)
    - ▶ MOrder1Container.m (house1markovorder1)
    - ▶ MOrder2Container.m (house1)
  - ▶ Multi-Source Container
    - ▶ MultiData.m





# Data Preparation

- ▶ BuildDataStruct.m
  - ▶ Constructs Data, a Matlab struct
- ▶ Handled via
  - ▶ GeneralDataGeneration.m
  - ▶ PrepareData.m
  - ▶ PrepareDataLarge.m
- ▶ Supports:
  - ▶ K-Folds, Jack-Knife, and Bootstrap
  - ▶ Data normalization



# Learners

- ▶ BuildLConfig.m
  - ▶ Constructs LConfig, learner option container
- ▶ All learners use the LConfig interface
- ▶ All learners check the LConfig for their required parameters
  - ▶ If a parameter isn't present, learner should use a default value
- ▶ All learners define their own result saving rules
  - ▶ Current suite learners fixed naming conventions per learner



## Standard Learner Options

- learner
- data
- ofile
- fold
- order
- tshift
- target
- omit\_variable\_list
- scalefile



## Result Interface

- ▶ BuildRConfig.m
  - ▶ Constructs RConfig, result extraction option container
- ▶ Each learner has its own GenResultsFile.m file
  - ▶ REGResults.m
  - ▶ FFNNResults.m
  - ▶ SVRResults.m
- ▶ Compresses multiple results
- ▶ Final file stores
  - ▶ Best model
  - ▶ Best parameter settings



# Outline

Introduction

Machine Learning Suite

**XML Interface**

MLSuite Results

Future Electrical Consumption

EnergyPlus Approximation

Inverse EnergyPlus

Closing Remarks



# XML Parsing Software

- ▶ Two Launchers
  - ▶ LaunchBatch.py (Nautilus)
  - ▶ LaunchTask.py (OSX/Linux)
- ▶ Layout
  - ▶ BatchParser.py (Nautilus)
  - ▶ JobParser.py (OSX/Linux)
  - ▶ Task.py
  - ▶ Parameter.py
  - ▶ ProcManager.py



## Job Tags

- ▶ `<Batch> </Batch>`
  - ▶ `bpath` — base path
  - ▶ `lpath` — execution status storage path
- ▶ `<Job> </Job>`
  - ▶ Describe a groups of tasks that can be execute in parallel
- ▶ `<Task> </Task>`
  - ▶ Describes the parameters to be passed to a single executable

```
○○○○○○○○  
○○○○○○○○○○○○  
○○○○○
```

## XML Job Tag Example

```
<Batch bpath="~/ ">  
  <Job>  
    <Task>  
      Data Generation  
    <Task/>  
    :  
  <Job/>  
  <Job>  
    <Task>  
      Learner  
    <Task/>  
    :  
  <Job/>  
</Batch/>
```





# Parameter Tags

- ▶ `<gparameter ... / >`
  - ▶ Defines global Job parameters
- ▶ `<parameter ... / >`
  - ▶ Defines local task parameters
- ▶ Attributes
  - ▶ id
  - ▶ value
  - ▶ type
  - ▶ step
  - ▶ maxvalue



## XML Parameter Examples

```
<parameter id="fold" value="--fold" type="string" / >
```

```
<parameter id="foldv" value="1" maxvalue="10" step="1"
  type="numeric" / >
```

```
<parameter id="order" value="--order" type="string" / >
```

```
<parameter id="orderv" value="1" maxvalue="3" step="1"
  type="numeric" / >
```



## Executable Tag

- ▶ `<binary> executable </binary>`
  - ▶ path
- ▶ Example: `<binary path="" >GenerateData.py</binary>`
- ▶ Provided Executables
  - ▶ GenerateData.py
  - ▶ Launch.py
  - ▶ GenerateResults.py



# Outline

Introduction

Machine Learning Suite

XML Interface

**MLSuite Results**

Future Electrical Consumption

EnergyPlus Approximation

Inverse EnergyPlus

Closing Remarks



# Data

- ▶ Two Subdivisions
  - ▶ Wolf Creek
  - ▶ Campbell Creek
- ▶ Wolf Creek
  - ▶ approximately 250 sensors
  - ▶ 15 minute resolution
- ▶ Campbell Creek
  - ▶ approximately 140 sensors
  - ▶ 15 minute resolution





# Predictor Experimental Setup

- ▶ Data Set:
  - ▶ Campbell Creek Houses
- ▶ Training/Testing
  - ▶ K-Folds: 10 Folds





# Campbell Creek House 1

Order 1

	CV(%)	MBE(%)	MAPE(%)
Regression	32.38±1.91	-0.06±1.08	30.52±1.41
FFNN	25.10±2.34	0.66±1.43	21.08±1.14
SVR	24.60±1.78	-2.46±0.95	17.05±0.94
LS-SVM	23.39±1.26	0.01±0.84	18.21±0.89
HME-REG	32.35±1.82	-0.05±1.02	30.57±1.42
HME-LSSVM	23.68±1.41	-0.03±0.99	18.69±0.85
HME-FFNN	22.77±1.56	0.15±0.98	17.74±0.65
FCM-REG	31.91±1.67	-0.09±0.91	29.74±0.86
FCM-FFNN	22.65±1.42	0.81±0.95	18.18±0.75
FCM-LSSVM	24.03±1.20	0.01±0.87	19.52±0.92

Order 2

	CV(%)	MBE(%)	MAPE(%)
Regression	27.63±1.95	-0.03±1.09	26.18±1.51
FFNN	24.32±2.61	0.53±1.74	22.28±2.67
SVR	21.58±1.40	-1.41±0.89	16.41±0.95
<b>LS-SVM</b>	<b>20.05±0.81</b>	<b>0.06±0.62</b>	16.11±0.85
HME-REG	27.60±2.13	-0.03±1.01	26.11±1.67
HME-LSSVM	20.23±0.85	0.07±0.56	16.40±0.80
HME-FFNN	20.15±1.65	0.46±0.93	17.07±1.19
FCM-REG	27.33±1.48	-0.14±0.72	25.62±0.80
FCM-FFNN	20.53±1.76	0.74±0.87	17.57±1.42
FCM-LSSVM	20.54±0.83	0.04±0.62	16.91±0.84

Order 3

	CV(%)	MBE(%)	MAPE(%)
Regression	26.27±1.19	-0.11±1.45	24.33±0.96
FFNN	25.24±1.59	1.00±1.05	22.29±1.81
SVR	21.32±1.32	-1.50±0.80	15.48±0.87
LS-SVM	20.36±1.46	0.11±0.63	15.73±1.11
HME-REG	26.14±1.10	-0.08±1.44	24.21±0.93
HME-LSSVM	20.58±1.19	0.03±0.94	16.03±0.98
HME-FFNN	20.39±1.67	0.70±0.92	17.09±0.81
FCM-REG	26.33±1.72	-0.20±1.10	23.91±1.22
FCM-FFNN	21.03±1.29	0.47±1.49	18.27±1.06
FCM-LSSVM	20.50±1.47	0.07±0.69	16.11±1.15

▶ LS-SVM is the best (ASHRAE Metrics)



# Campbell Creek House 1

Order 1

	CV(%)	MBE(%)	MAPE(%)
Regression	32.38±1.91	-0.06±1.08	30.52±1.41
FFNN	25.10±2.34	0.66±1.43	21.08±1.14
SVR	24.60±1.78	-2.46±0.95	17.05±0.94
LS-SVM	23.39±1.26	0.01±0.84	18.21±0.89
HME-REG	32.35±1.82	-0.05±1.02	30.57±1.42
HME-LSSVM	23.68±1.41	-0.03±0.99	18.69±0.85
HME-FFNN	22.77±1.56	0.15±0.98	17.74±0.65
FCM-REG	31.91±1.67	-0.09±0.91	29.74±0.86
FCM-FFNN	22.65±1.42	0.81±0.95	18.18±0.75
FCM-LSSVM	24.03±1.20	0.01±0.87	19.52±0.92

Order 2

	CV(%)	MBE(%)	MAPE(%)
Regression	27.63±1.95	-0.03±1.09	26.18±1.51
FFNN	24.32±2.61	0.53±1.74	22.28±2.67
SVR	21.58±1.40	-1.41±0.89	<b>16.41±0.95</b>
LS-SVM	20.05±0.81	0.06±0.62	<b>16.11±0.85</b>
HME-REG	27.60±2.13	-0.03±1.01	26.11±1.67
HME-LSSVM	20.23±0.85	0.07±0.56	16.40±0.80
HME-FFNN	20.15±1.65	0.46±0.93	17.07±1.19
FCM-REG	27.33±1.48	-0.14±0.72	25.62±0.80
FCM-FFNN	20.53±1.76	0.74±0.87	17.57±1.42
FCM-LSSVM	20.54±0.83	0.04±0.62	16.91±0.84

Order 3

	CV(%)	MBE(%)	MAPE(%)
Regression	26.27±1.19	-0.11±1.45	24.33±0.96
FFNN	25.24±1.59	1.00±1.05	22.29±1.81
SVR	21.32±1.32	-1.50±0.80	15.48±0.87
LS-SVM	20.36±1.46	0.11±0.63	15.73±1.11
HME-REG	26.14±1.10	-0.08±1.44	24.21±0.93
HME-LSSVM	20.58±1.19	0.03±0.94	16.03±0.98
HME-FFNN	20.39±1.67	0.70±0.92	17.09±0.81
FCM-REG	26.33±1.72	-0.20±1.10	23.91±1.22
FCM-FFNN	21.03±1.29	0.47±1.49	18.27±1.06
FCM-LSSVM	20.50±1.47	0.07±0.69	16.11±1.15

▶ LS-SVM and SVR are best (MAPE)





## Campbell Creek House 2

Order 1

	CV(%)	MBE(%)	MAPE(%)
Regression	36.73±2.26	-0.13±1.00	31.01±3.48
FFNN	33.24±1.26	0.50±0.91	27.28±3.12
SVR	30.36±1.83	-2.95±1.03	20.44±2.81
<b>LS-SVM</b>	<b>27.88±1.24</b>	<b>-0.05±0.91</b>	20.47±2.37
HME-REG	35.82±1.04	0.15±0.88	30.48±3.20
HME-LSSVM	27.98±1.39	0.01±0.99	20.84±2.89
HME-FFNN	29.30±1.28	0.09±1.01	22.71±2.92
FCM-REG	35.20±0.87	0.05±1.99	29.77±2.41
<b>FCM-FFNN</b>	<b>28.14±1.21</b>	<b>0.40±0.97</b>	21.96±2.74
FCM-LSSVM	28.05±1.17	-0.03±1.00	21.01±2.33

Order 2

	CV(%)	MBE(%)	MAPE(%)
Regression	34.15±1.66	0.05±1.61	28.36±3.72
FFNN	33.83±1.98	0.21±1.45	27.07±4.14
SVR	29.22±1.06	-3.00±1.12	19.42±3.27
LS-SVM	27.43±1.90	0.20±1.03	20.17±2.26
HME-REG	34.15±1.74	0.14±1.38	28.29±3.86
HME-LSSVM	27.63±1.28	0.10±0.89	20.41±3.42
HME-FFNN	28.17±2.04	0.26±0.58	22.43±2.44
FCM-REG	33.49±1.52	0.01±1.59	27.46±2.77
FCM-FFNN	28.34±1.67	-0.20±1.27	22.30±3.28
FCM-LSSVM	27.19±1.90	0.16±1.14	20.17±2.34

Order 3

	CV(%)	MBE(%)	MAPE(%)
Regression	33.15±1.33	-0.02±0.96	27.87±2.40
FFNN	34.23±1.63	2.01±2.45	29.62±2.16
SVR	28.59±2.05	-2.33±1.09	19.58±2.07
LS-SVM	27.68±1.91	-0.02±1.71	20.23±2.56
HME-REG	33.20±1.32	-0.08±0.97	27.95±2.31
HME-LSSVM	27.19±1.87	0.37±0.84	20.67±2.30
HME-FFNN	29.64±2.21	-0.12±1.64	24.81±0.38
FCM-REG	32.70±1.66	-0.00±2.02	27.12±2.91
FCM-FFNN	28.94±1.46	0.45±1.27	22.76±2.03
FCM-LSSVM	27.24±1.93	-0.01±1.76	19.70±2.53

- ▶ LS-SVM and FCM-FFNN is best (ASHRAE Metrics)



## Campbell Creek House 2

Order 1

	CV(%)	MBE(%)	MAPE(%)
Regression	36.73±2.26	-0.13±1.00	31.01±3.48
FFNN	33.24±1.26	0.50±0.91	27.28±3.12
SVR	30.36±1.83	-2.95±1.03	<b>20.44±2.81</b>
LS-SVM	27.88±1.24	-0.05±0.91	<b>20.47±2.37</b>
HME-REG	35.82±1.04	0.15±0.88	30.48±3.20
HME-LSSVM	27.98±1.39	0.01±0.99	20.84±2.89
HME-FFNN	29.30±1.28	0.09±1.01	22.71±2.92
FCM-REG	35.20±0.87	0.05±1.99	29.77±2.41
FCM-FFNN	28.14±1.21	0.40±0.97	<b>21.96±2.74</b>
FCM-LSSVM	28.05±1.17	-0.03±1.00	21.01±2.33

Order 2

	CV(%)	MBE(%)	MAPE(%)
Regression	34.15±1.66	0.05±1.61	28.36±3.72
FFNN	33.83±1.98	0.21±1.45	27.07±4.14
SVR	29.22±1.06	-3.00±1.12	19.42±3.27
LS-SVM	27.43±1.90	0.20±1.03	20.17±2.26
HME-REG	34.15±1.74	0.14±1.38	28.29±3.86
HME-LSSVM	27.63±1.28	0.10±0.89	20.41±3.42
HME-FFNN	28.17±2.04	0.26±0.58	22.43±2.44
FCM-REG	33.49±1.52	0.01±1.59	27.46±2.77
FCM-FFNN	28.34±1.67	-0.20±1.27	22.30±3.28
FCM-LSSVM	27.19±1.90	0.16±1.14	20.17±2.34

Order 3

	CV(%)	MBE(%)	MAPE(%)
Regression	33.15±1.33	-0.02±0.96	27.87±2.40
FFNN	34.23±1.63	2.01±2.45	29.62±2.16
SVR	28.59±2.05	-2.33±1.09	19.58±2.07
LS-SVM	27.68±1.91	-0.02±1.71	20.23±2.56
HME-REG	33.20±1.32	-0.08±0.97	27.95±2.31
HME-LSSVM	27.19±1.87	0.37±0.84	20.67±2.30
HME-FFNN	29.64±2.21	-0.12±1.64	24.81±0.38
FCM-REG	32.70±1.66	-0.00±2.02	27.12±2.91
FCM-FFNN	28.94±1.46	0.45±1.27	22.76±2.03
FCM-LSSVM	27.24±1.93	-0.01±1.76	19.70±2.53

- ▶ LS-SVM, SVR, and FCM-FFNN is best (MAPE)



# Campbell Creek House 3

## House 3

### Order 1

	CV(%)	MBE(%)	MAPE(%)
Regression	40.07±2.21	0.07±1.15	32.49±1.88
FFNN	37.15±1.57	0.35±2.03	28.92±2.55
SVR	33.71±1.72	-3.36±0.99	21.49±1.80
LS-SVM	31.60±2.07	-0.15±1.10	22.25±1.33
HME-REG	39.17±2.17	0.33±1.38	31.72±2.07
HME-LSSVM	31.85±1.83	0.14±1.12	23.03±2.48
HME-FFNN	32.98±1.28	-0.12±0.84	23.99±1.63
FCM-REG	39.69±3.11	0.12±1.30	31.58±1.88
FCM-FFNN	33.03±1.67	0.93±1.52	25.28±2.14
FCM-LSSVM	31.75±2.01	-0.12±1.09	22.76±1.29

### Order 2

	CV(%)	MBE(%)	MAPE(%)
Regression	39.26±4.19	0.11±1.86	31.34±2.58
FFNN	38.02±2.49	2.05±2.67	29.83±2.02
SVR	32.38±2.96	-3.12±1.73	20.72±1.38
<b>LS-SVM</b>	<b>30.66±2.53</b>	<b>-0.05±0.93</b>	21.33±1.40
HME-REG	38.48±4.34	1.03±1.72	30.53±3.07
HME-LSSVM	30.61±2.23	-0.25±1.74	21.22±1.34
HME-FFNN	32.99±2.17	1.07±1.17	24.76±1.94
FCM-REG	38.74±2.67	0.08±1.90	30.56±1.76
FCM-FFNN	32.92±2.49	0.76±2.03	24.20±2.06
FCM-LSSVM	30.48±2.39	-0.04±0.99	21.24±1.36

### Order 3

	CV(%)	MBE(%)	MAPE(%)
Regression	38.53±3.47	0.15±1.22	30.49±2.15
FFNN	38.58±2.07	-0.08±2.46	30.57±2.51
<b>SVR</b>	<b>31.88±2.01</b>	<b>-2.84±0.97</b>	20.47±1.69
LS-SVM	30.78±2.56	-0.21±1.04	21.36±1.50
HME-REG	38.22±3.58	1.20±1.49	29.52±2.47
HME-LSSVM	30.97±1.37	-0.21±0.97	21.37±1.61
HME-FFNN	33.34±1.83	1.09±1.24	25.15±2.13
FCM-REG	37.66±1.88	0.04±1.06	29.82±1.67
FCM-FFNN	33.66±2.09	1.17±1.30	25.51±1.72
FCM-LSSVM	30.57±2.55	-0.19±1.02	21.22±1.58

▶ LS-SVM is best (ASHRAE Metrics)



# Campbell Creek House 3

House 3

Order 1

	CV(%)	MBE(%)	MAPE(%)
Regression	40.07±2.21	0.07±1.15	32.49±1.88
FFNN	37.15±1.57	0.35±2.03	28.92±2.55
SVR	33.71±1.72	-3.36±0.99	21.49±1.80
LS-SVM	31.60±2.07	-0.15±1.10	22.25±1.33
HME-REG	39.17±2.17	0.33±1.38	31.72±2.07
HME-LSSVM	31.85±1.83	0.14±1.12	23.03±2.48
HME-FFNN	32.98±1.28	-0.12±0.84	23.99±1.63
FCM-REG	39.69±3.11	0.12±1.30	31.58±1.88
FCM-FFNN	33.03±1.67	0.93±1.52	25.28±2.14
FCM-LSSVM	31.75±2.01	-0.12±1.09	22.76±1.29

Order 2

	CV(%)	MBE(%)	MAPE(%)
Regression	39.26±4.19	0.11±1.86	31.34±2.58
FFNN	38.02±2.49	2.05±2.67	29.83±2.02
SVR	32.38±2.96	-3.12±1.73	20.72±1.38
<b>LS-SVM</b>	30.66±2.53	-0.05±0.93	<b>21.33±1.40</b>
HME-REG	38.48±4.34	1.03±1.72	30.53±3.07
HME-LSSVM	30.61±2.23	-0.25±1.74	21.22±1.34
HME-FFNN	32.99±2.17	1.07±1.17	24.76±1.94
FCM-REG	38.74±2.67	0.08±1.90	30.56±1.76
FCM-FFNN	32.92±2.49	0.76±2.03	24.20±2.06
FCM-LSSVM	30.48±2.39	-0.04±0.99	21.24±1.36

Order 3

	CV(%)	MBE(%)	MAPE(%)
Regression	38.53±3.47	0.15±1.22	30.49±2.15
FFNN	38.58±2.07	-0.08±2.46	30.57±2.51
<b>SVR</b>	31.88±2.01	-2.84±0.97	<b>20.47±1.69</b>
LS-SVM	30.78±2.56	-0.21±1.04	21.36±1.50
HME-REG	38.22±3.58	1.20±1.49	29.52±2.47
HME-LSSVM	30.97±1.37	-0.21±0.97	21.37±1.61
HME-FFNN	33.34±1.83	1.09±1.24	25.15±2.13
FCM-REG	37.66±1.88	0.04±1.06	29.82±1.67
FCM-FFNN	33.66±2.09	1.17±1.30	25.51±1.72
FCM-LSSVM	30.57±2.55	-0.19±1.02	21.22±1.58

- ▶ LS-SVM and SVR is best according to MAPE





# Data

- ▶ Markov Order 1
  - ▶ Adjust parameters independently
  - ▶ min,max adjustment
- ▶ Markov Order 2
  - ▶ Adjust two parameters together
  - ▶ min,max adjustments
- ▶ Fine Grain (Brute Force)
  - ▶ Adjust 14 parameters
  - ▶ Small incremental adjustments
- ▶ All datasets cover 150 building parameters
- ▶ All use the same weather and operation schedule



# Dataset Sizes

	Number Outputs	Number Simulations	Gigabytes
Markov Order 1	95	299	3.9
Markov Order 2	95	29,727	387.2
Fine Grain	82	11,989	136.0

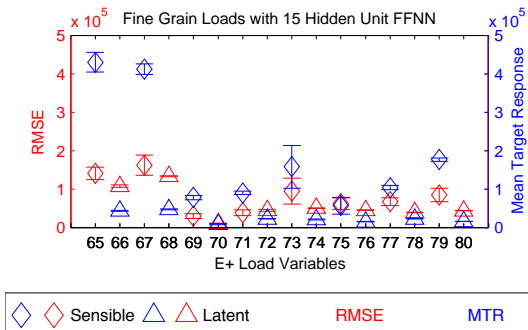


# Experimental Setup

- ▶ FG Experiments
  - ▶ Training set 250 simulations
  - ▶ Testing set 750 simulations
- ▶ MO1 & MO2 Experiments
  - ▶ Training set MO1 data set
  - ▶ Testing set 250 MO2 simulations



## FFNN FG Result

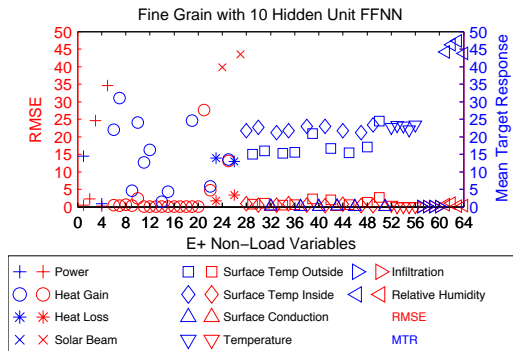


- ▶ FFNN with 15 and 5 hidden units fit the Fine Grain loads best





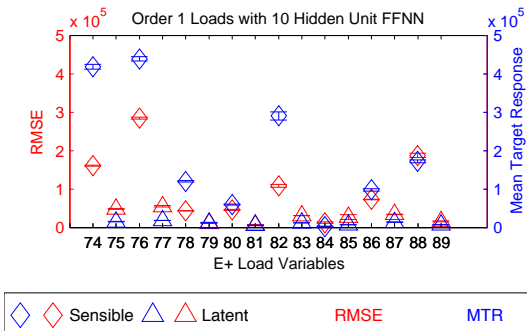
## FFNN FG Result



- ▶ Fits non-loads better than the 5 hidden unit model
- ▶ The 15 hidden unit model is very similar to the 10 hidden unit



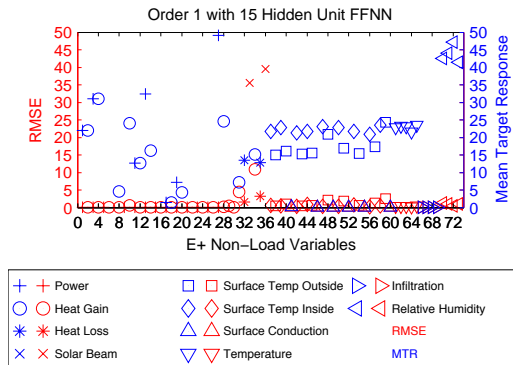
# FFNN MO2 Result



- ▶ MO1 results are similar to FG results



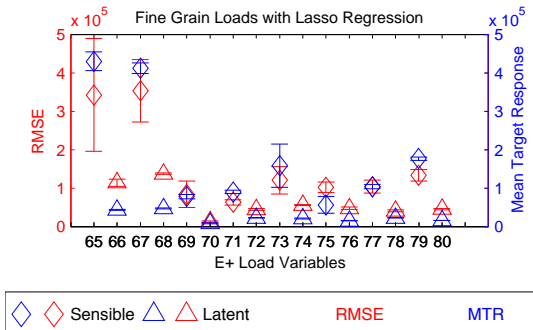
# FFNN MO2 Result



## ► Best non-load model



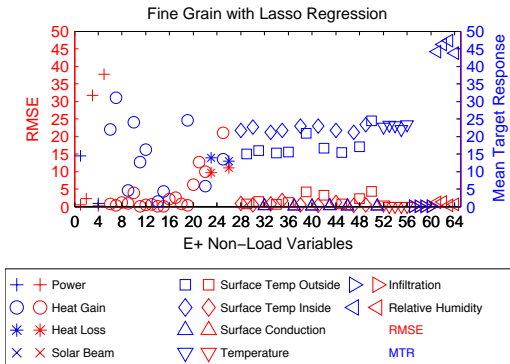
## Lasso FG Results



- ▶ Does not estimate FG loads as well as FFNN
- ▶ Based on variable 65 and 67



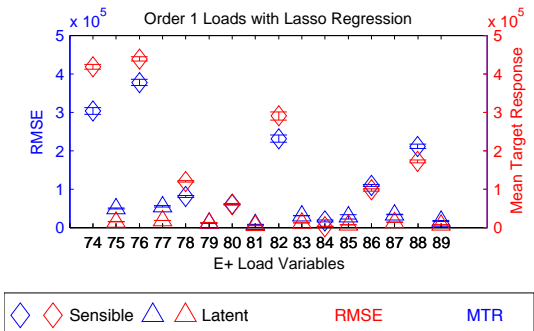
# Lasso FG Results



- ▶ Estimates non-load variables worse than FFNN



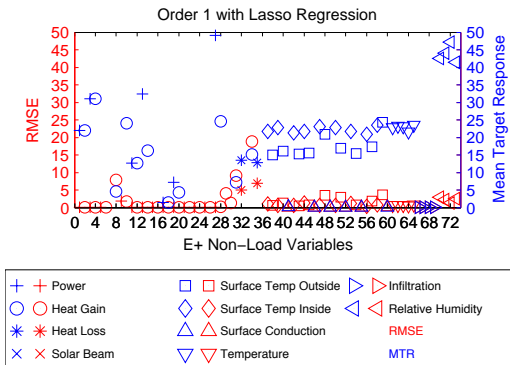
## Lasso MO2 Results



- ▶ Estimates MO1 loads better than FG loads
- ▶ Worse than FFNN



# Lasso MO2 Results



- ▶ Estimates non-load variables as well as FFNN



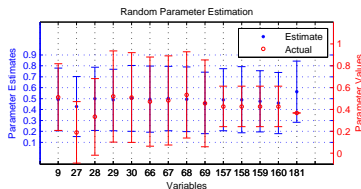
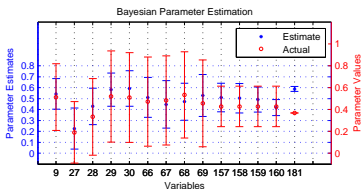
# Experiment Setup

- ▶ FG Experiments
  - ▶ Training set 250 simulations
  - ▶ Testing set 750 simulations
- ▶ MO1 & MO2 Experiments
  - ▶ Training set MO1 data set
  - ▶ Testing set 250 MO2 simulations





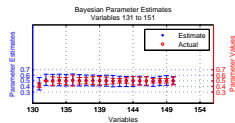
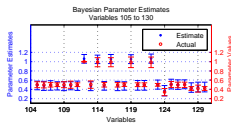
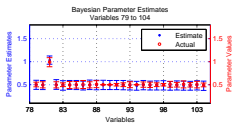
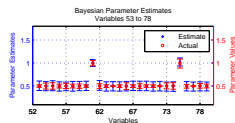
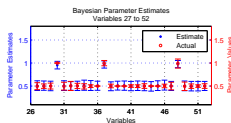
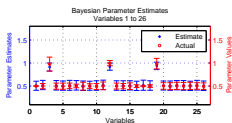
## FG Results



- ▶ Random works best on 0.5 mean variables
- ▶ Bayesian tracks means better
- ▶ Appears to infer building parameters well



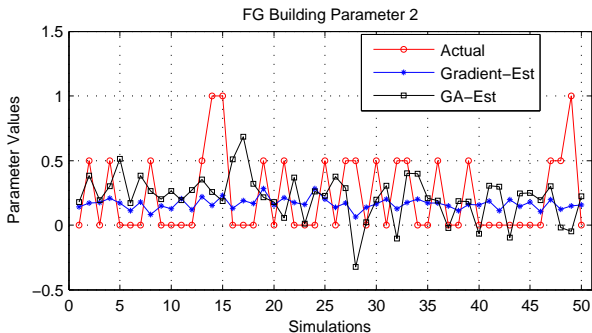
# MO2 Results



- ▶ Appears to infer building parameters well
- ▶ Tracks means well

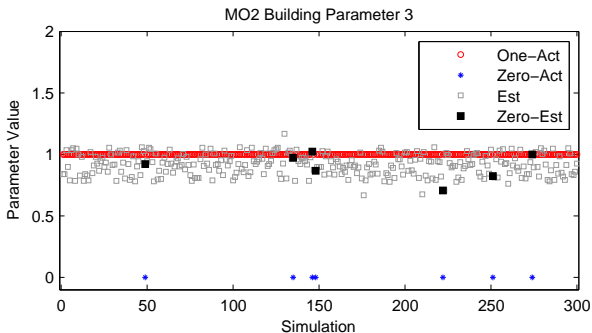


# Genetic Algorithm vs Gradient



- ▶ Gradient estimates near the mean often
- ▶ GA introduces more variance
- ▶ Gradient better for large parameter inference
  - ▶ Variance scales with number of parameters
  - ▶ MO1 and FG used GA, MO2 used Gradient

# Estimating Distant Values



- ▶ Values concentrate on the mean closely
- ▶ Distant values hard to estimate



# Outline

Introduction

Machine Learning Suite

XML Interface

MLSuite Results

Future Electrical Consumption

EnergyPlus Approximation

Inverse EnergyPlus

Closing Remarks



- ▶ MLSuite characteristics
  - ▶ Supports a wide range of learning options
  - ▶ Supports running across multiple networked computers
  - ▶ Supports running on Nautilus
  - ▶ Supports a wide range of data options