2016 Annual Conference

Joshua New, Ph.D.
Oak Ridge National Laboratory
newjr@ornl.gov

Seminar 22 – Large-Scale Computing

Design of Experiments: Statistical Confidence with Fewer Simulations

St. Louis, Missouri

Learning Objectives

1. Define design of experiment and statistical resolution
2. Describe capabilities for a resolution IV statistical design.
3. Describe sensitivity screening method options.
4. Explain additional requirements for design when variables are correlated.

Acknowledgments

- Innovative and Novel Computation Impact on Theory and Experiment (INCITE)
- Oak Ridge Leadership Computing Facility (OLCF)
- Extreme Science and Engineering Discovery Environment (XSEDE)
- Pragnesh Patel – UT
- Robert Mee – UT
- Jibonananda Sanyal – ORNL
- Mark Adams – ORNL
- George Ostrouchov – ORNL
- Kwai Wong – UT

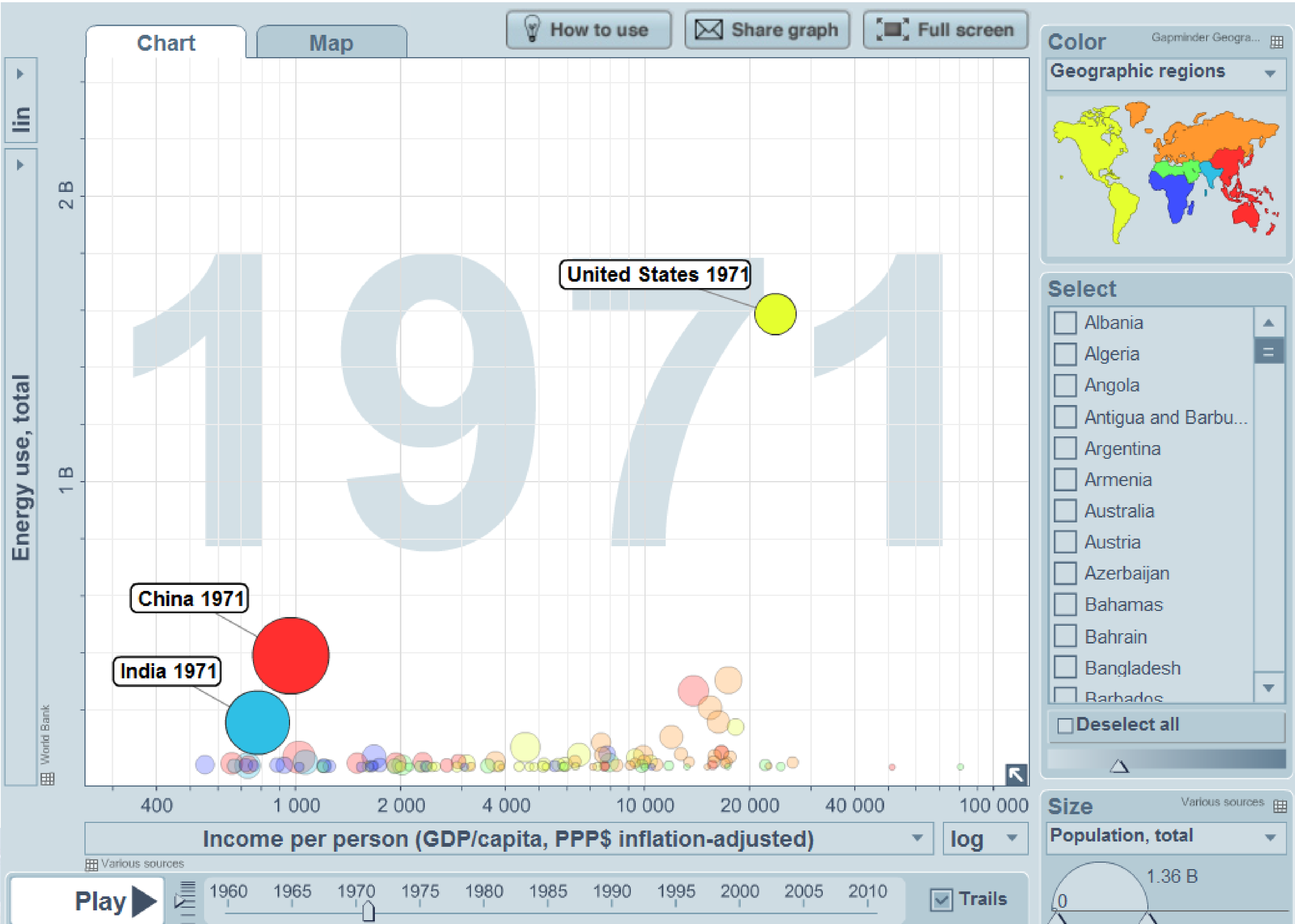
Outline/Agenda

- Motivation
- Parameterization
- Design of Experiments
  - Limitations
- Examples

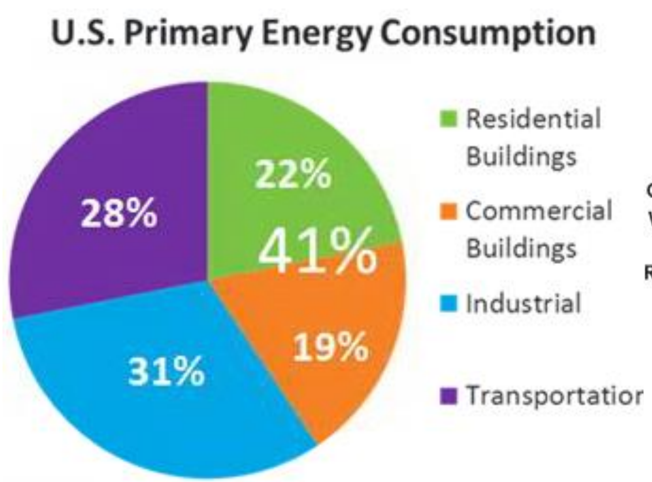
Outline/Agenda

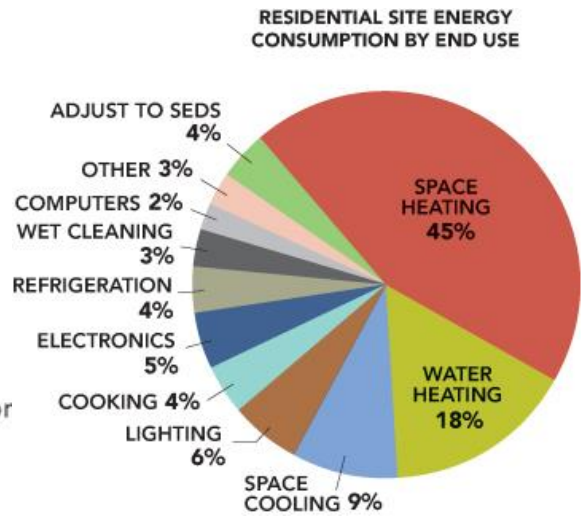
- Motivation
- Parameterization
- Design of Experiments
  - Limitations
- Examples

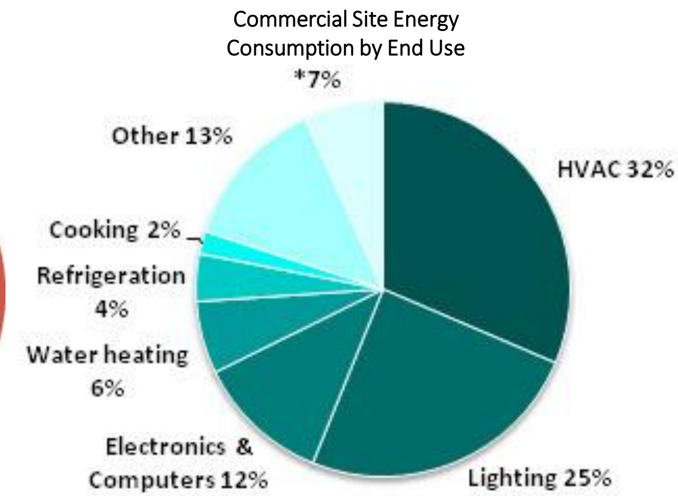
40 Years: Energy and Life

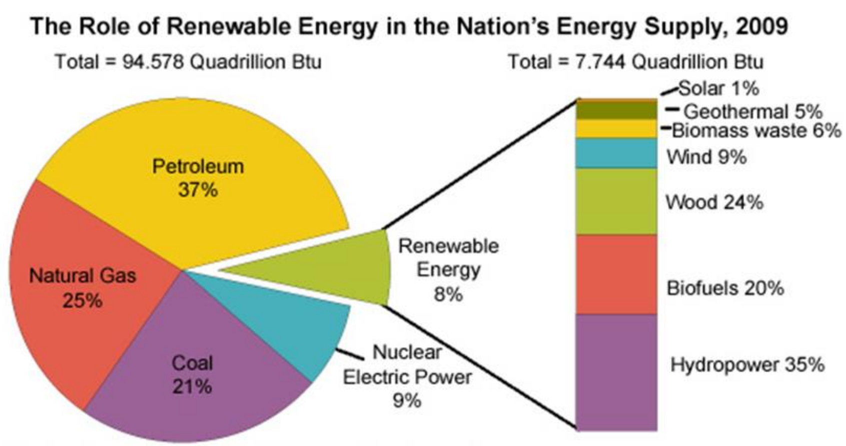


U.S. Energy Consumption

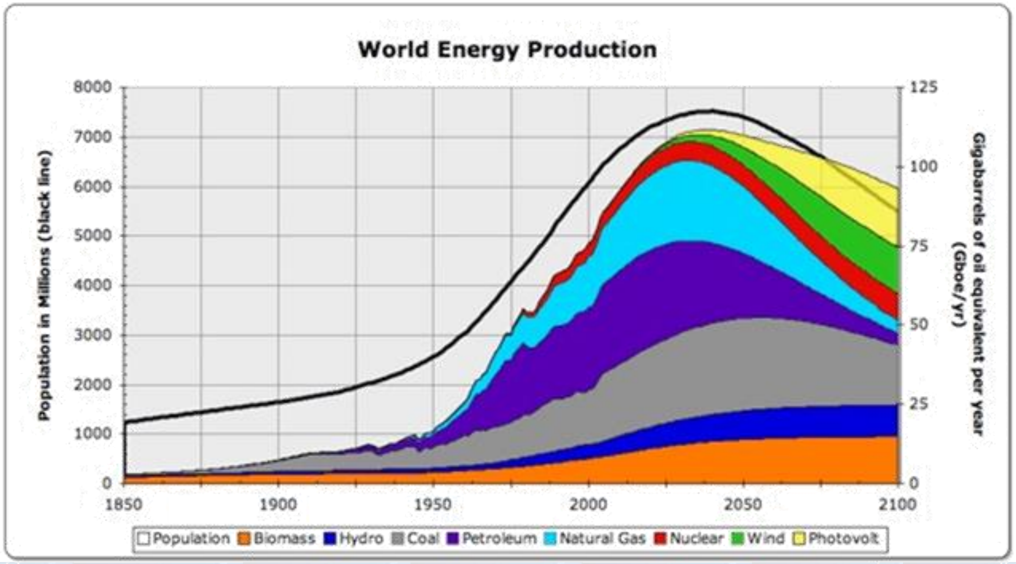




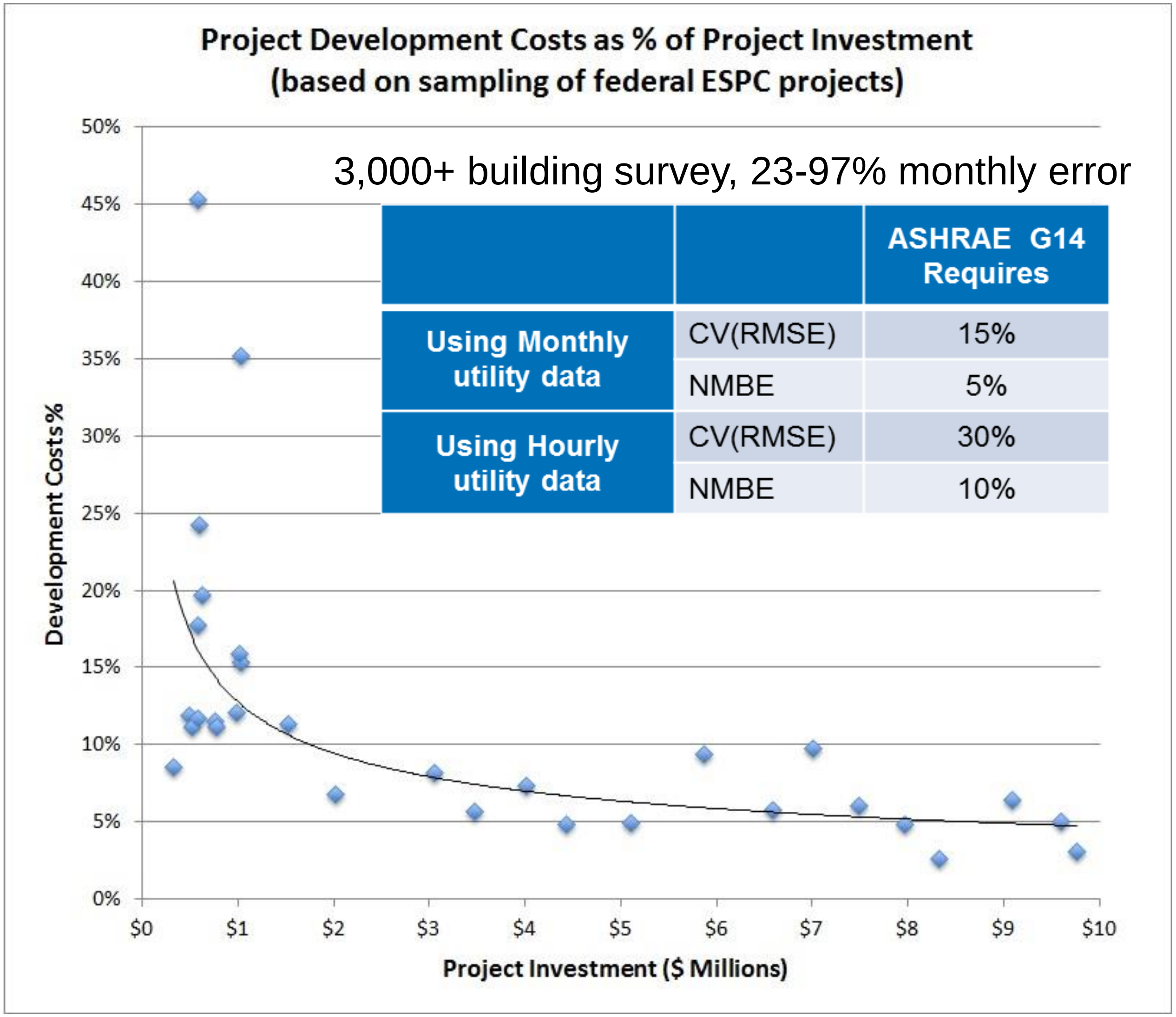




Note: Sum of components may not equal 100% due to independent rounding.
Source: U.S. Energy Information Administration, *Annual Energy Review 2009*, Table 1.3, Primary Energy Consumption by Energy Source, 1949-2009 (August 2010).



BEM Limited



3,000+ building survey, 23-97% monthly error

		ASHRAE G14 Requires
Using Monthly utility data	CV(RMSE)	15%
	NMBE	5%
Using Hourly utility data	CV(RMSE)	30%
	NMBE	10%

Outline/Agenda

- Motivation
- Parameterization
- Design of Experiments
  - Limitations
- Examples

Most "Important" Inputs

- 3,000+ inputs per building
- Use prototype buildings
  - many assumptions differences from an existing or planned/new building
- Calibrate important inputs
- Pick qualitatively important parameters, quantify energy use impact (sensitivity) or quantify uncertainty of estimates (UQ)

Parametric List

	Restaurant	Hospital	Large Hotel	Large Office	Medium Office	Midrise Apartment	Primary School	Quick Service
#Inputs	49	227	110	85	81	155	166	54
#Groups	49	139	67	43	36	78	109	54
	Secondary School	Small Hotel	Small Office	Stand-alone Retail	Strip Mall	Super Market	Warehouse	TOTAL
#Inputs	231	282	72	59	113	78	47	1809
#Groups	122	131	58	55	85	72	44	1142

	A	B	C	D	E	F	G	H	I	J
1	Class	Object	Field	Default	Minimum	Maximum	Distribution	Type	Group	Constraint
2	Lights	Bakery_Lights	Watts per Zone	18.29	12.803	23.777	uniform	float		
3	Lights	Deli_Lights	Watts per Zone	18.29	12.803	23.777	uniform	float		
4	ElectricEquipment	Bakery_MiscPlug_Ec	Design Level	11244	7870.8	14617.2	uniform	float		
5	ElectricEquipment	Deli_MiscPlug_Equi	Design Level	12105	8473.5	15736.5	uniform	float		
6	GasEquipment	Bakery_MiscGas_Eq	Design Level	5622	3935.4	7308.6	uniform	float		
7	GasEquipment	Deli_MiscGas_Equip	Design Level	6053	4237.1	7868.9	uniform	float		
8	Exterior:Lights	Exterior Facade Ligh	Design Level	13577	9503.9	17650.1	uniform	float		
9	ZoneInfiltration:Desig	Bakery_Infiltration	Flow per Exteri	0.000302	0.000211	0.000393	uniform	float	G0001	
10	ZoneInfiltration:Desig	Deli_Infiltration	Flow per Exteri	0.000302	0.000211	0.000393	uniform	float	G0001	
11	Schedule:Compact	CLGSETP_SCH	Field 4	30	21	39	uniform	float	CA1	
12	Schedule:Compact	CLGSETP_SCH	Field 7	30	21	39	uniform	float	CA2	HA2 - CA2 < - 1
13	Schedule:Compact	CLGSETP_SCH	Field 9	24	16.8	31.2	uniform	float	CA3	HA3 - CA3 < - 1
14	Schedule:Compact	CLGSETP_SCH	Field 11	30	21	39	uniform	float	CA4	HA4 - CA4 < - 1
15	Schedule:Compact	HTGSETP_SCH	Field 4	15.6	10.92	20.28	uniform	float	HA1	
16	Schedule:Compact	HTGSETP_SCH	Field 7	15.6	10.92	20.28	uniform	float	HA2	
17	Schedule:Compact	HTGSETP_SCH	Field 9	21	14.7	27.3	uniform	float	HA3	
18	Schedule:Compact	HTGSETP_SCH	Field 11	15.6	10.92	20.28	uniform	float	HA4	

Outline/Agenda

- Motivation
- Parameterization
- Design of Experiments
  - Limitations
- Examples

DoE - Full

- Design of Experiments (a.k.a. experimental designs)
  - Full factorial, consists of 2+ variables with discrete values whose experimental units take on all possible combinations of those values across all variables. $\#vals^{\#vars}$
    - Example: 3 vars, 2 vals each = 8 sims (16 mins)
    - Example: 5 vars, 2 vals each = 32 sims (1.1 hrs)
    - Example: 10 vars, 2 vals each = 1,024 sims (1.4 days)
    - Residential Building Calibration:
      156 parameters, $5x10^{52}$ sims, $4x10^{28}$ (LOKU=13.8 billion years) on world's fastest supercomputer!

DoE – Full to Fractional

- Full factorial sampling matrix for 3 variables: 2 values each (+1, -1) represents (max, min)

	X_1	X_2	X_3	$X_4 = X_1X_2$	$X_5 = X_1X_3$
1	-1	-1	-1	+1	+1
2	+1	-1	-1	-1	-1
3	-1	+1	-1	-1	+1
4	+1	+1	-1	+1	-1
5	-1	-1	+1	+1	-1
6	+1	-1	+1	-1	+1
7	-1	+1	+1	-1	-1
8	+1	+1	+1	+1	+1

16 minutes   1.1 hours (or 8 minutes if X_4 and X_5 are confounded)

Fractional Factorial

L^{k-p} where:

L = #vals

k = #vars

p = #generators

This example is a 2^{5-2} design

DoE - Fractional

- Fractional Factorial Design
  - Resolution III (2^{5-2}) estimates main effects (A,B,C,D,E) but may be confounded with 2-factor interactions (D,E)
  - Resolution IV: main effects unconfounded, two-factor interactions confounded
  - Resolution V: main effects and 2-factor interactions unconfounded, three-factor interactions confounded

…

DoE - Fractional

- Fractional Factorial Design
  - Change the alias (why X_4 and X_5?) for new designs

$$M= \begin{matrix}
-1 & -1 & 1 & -1 & -1 & -1 & 1 & -1 & 1 \\
-1 & -1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \\
-1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 & -1 \\
-1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 & 1 \\
-1 & 1 & -1 & -1 & 1 & -1 & 1 & 1 & 1 \\
-1 & 1 & 1 & -1 & 1 & 1 & 1 & -1 & -1 \\
-1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\
-1 & -1 & 1 & 1 & 1 & 1 & -1 & -1 & 1 \\
1 & 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\
1 & 1 & -1 & -1 & -1 & 1 & -1 & -1 & -1 \\
1 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & -1 \\
1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 \\
1 & -1 & -1 & -1 & 1 & -1 & -1 & -1 & 1 \\
1 & -1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 \\
1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{matrix}$$

M is 16x9 matrix…
can be used to construct
one of arbitrary size
(#rows = #vars)

$$M' = \begin{bmatrix} M & M \\ M & -M \end{bmatrix}$$

  - #vars not factor of 16, throw out columns based on minimizing aberration
  - Helped extend “FrF2” package in R

DoE – Fractional Limitations

- Charts which show design of experiment scalability (confidence level based on #sims for #vars)
  - None go into hundreds of variables (extending)
- De-barred designs for 2-level (min,max) to 3-level (min,default,max)
- Accommodating distributions (more likely to be default)
- Correlated input variables

Outline/Agenda

- Motivation
- Parameterization
- Design of Experiments
  - Limitations
- Examples

Parametric List (retail building)

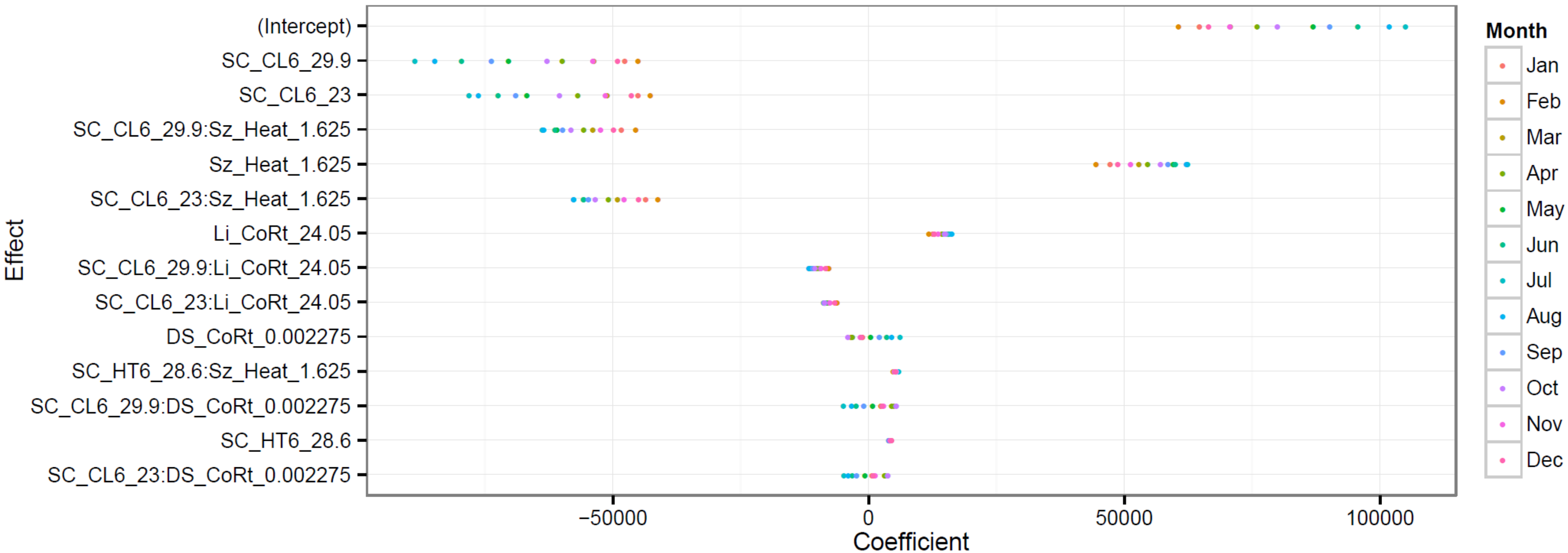
Class	Name	Short Name	Field	Default	Min	Max
Schedule: Compact	CLGSETP_SCH	SC_CL6	Field 6	23	16.1	29.9
	HTGSETP_SCH	SC_HT4	Field 4	16	11.2	20.8
	HTGSETP_SCH	SC_HT6	Field 6	22	15.4	28.6
Lights	Back_Space_Lights	Li_BaSp	Watts per Zone Floor Area	9	6.3	11.7
	Core_Retail_Lights	Li_CoRt	Watts per Zone Floor Area	18.5	12.95	24.05
	Front_Entry_Lights	Li_FrEn	Watts per Zone Floor Area	12	8.4	15.6
	Front_Retail_Lights	Li_FrRt	Watts per Zone Floor Area	18.5	12.95	24.05
	Point_Of_Sale_Lights	Li_POS	Watts per Zone Floor Area	18.5	12.95	24.05
Electric Equipment	BackSpace_MiscPlug	Eq_BaSp	Watts per Zone Floor Area	8.2	5.74	10.66
	CoreRetail_MiscPlug	Eq_CoRt	Watts per Zone Floor Area	3.3	2.31	4.29
	FrontRetail_MiscPlug	Eq_FrRt	Watts per Zone Floor Area	3.3	2.31	4.29
	PointOfSale_MiscPlug	Eq_POS	Watts per Zone Floor Area	22	15.4	28.6
ZoneInfil: FlowRate	Back_Space_Infil	ZF_BaSp	Flow per Ext Surface Area	0.00033	0.000231	0.000429
	Front_Entry_Infil	ZF_FrEn	Air Changes per Hour	1.1	0.77	1.43
	Front_Entry_Infil	ZF_FrRtA	Constant Term Coefficient	0	0	1
	Front_Retail_Infil	ZF_FrRtC	Flow per Ext Surface Area	0.00033	0.000231	0.000429
	Point_Of_Sale_Infil	ZF_POS	Flow per Ext Surface Area	0.00033	0.000231	0.000429
DsgnSpec: OutdrAir	SZ DSOA Back_Space	DS_BaSp	Outdr Airflow per Area	0.0008	0.00056	0.00104
	SZ DSOA Core_Retail	DS_CoRt	Outdr Airflow per Area	0.00175	0.001225	0.002275
Sizing: Parameters	Sizing:Parameters	Sz_Heat	Heating Sizing Factor	1.25	0.875	1.625

Building Sim Example

- Retail building, 20 variables
  - Resolution VI design
  - 1024 simulations (instead of 1 million)
  - Record monthly energy (electricity and gas) use as response
  - All (380) 2-way interactions for 20 variables
  - 778 degrees of freedom in the error term
  - 58 of 380 interactions significant at <0.01

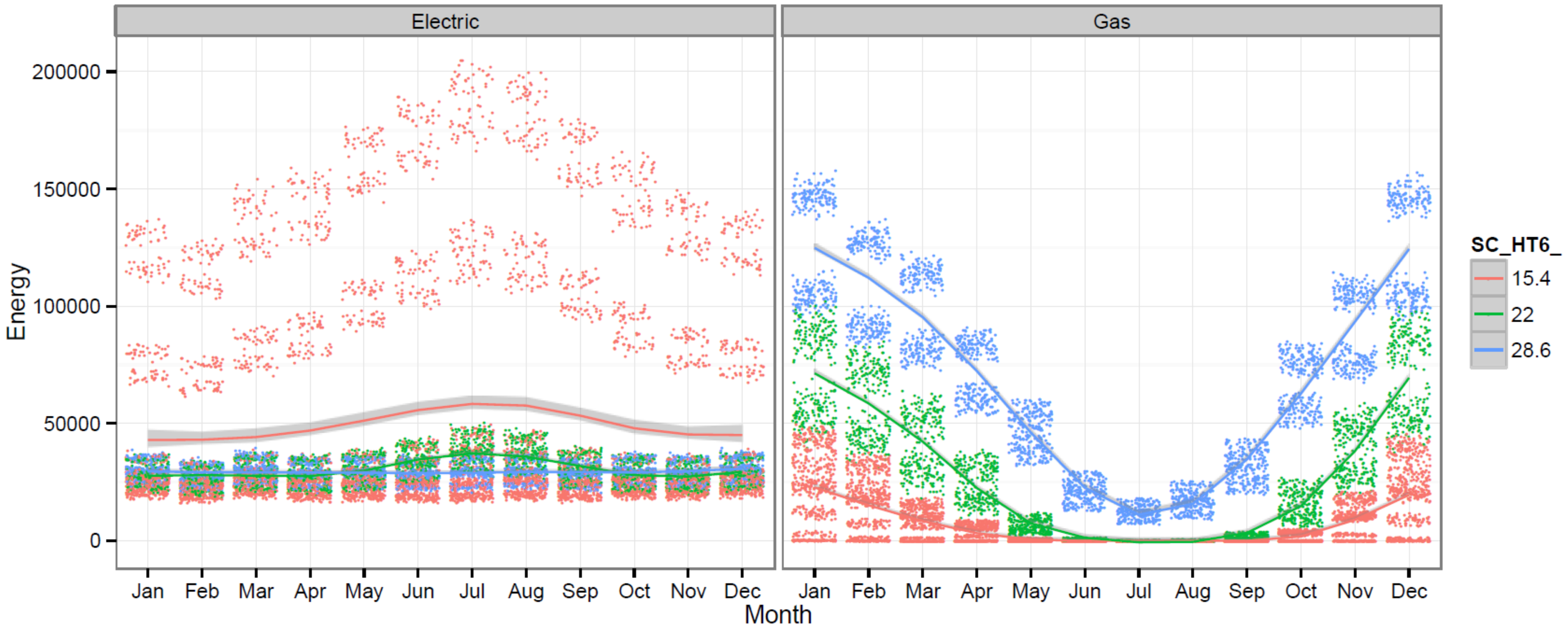
Most Important

- Y-axis descending from most important input variable (Cooling setpoint)
- X-axis in kWh



Sensitivity Analysis

- Sensitivity based on heating setpoint levels



Conclusions

- Sampling parametric spaces is important for sensitivity, uncertainty, and calibration of building energy models
- Design of experiments can usually allow strong results without supercomputers
- There are methods to overcome or mitigate most limitations

Bibliography

- Sanyal, Jibonananda and New, Joshua R. (2014). "Building Simulation Modelers - Are We Big Data Ready?" In *Proceedings of the ASHRAE/IBPSA-USA Building Simulation Conference*, pp. 449-456, Atlanta, GA, September 10-12, 2014. [ASHRAE] [PDF] [PPT]
- Ostrouchov, George, New, Joshua R., Sanyal, Jibonananda, and Patel, Pragnesh (2014). "Uncertainty Analysis of a Heavily Instrumented Building at Different Scales of Simulation." In *Proceedings of the 3rd International High Performance Buildings Conference*, Purdue, West Lafayette, IN, July 14-17, 2014. [PDF] [PPT]
- Castello, Charles C., Sanyal, Jibonananda, Rossiter, Jeffrey S., Hensley, Zachary P., and New, Joshua R. (2014). "Sensor Data Management, Validation, Correction, and Provenance for Building Technologies." Technical paper TRNS-00223-2013.R1. In*Proceedings of the ASHRAE Annual Conference and ASHRAE Transactions 2014*, Seattle, WA, June 28-July 2, 2014. [PDF] [PPT]

Questions?

Joshua New
newjr@ornl.gov